scientific reports

Check for updates

OPEN

Enhanced clustering approach for efficient relay vehicle selection in vehicular ad hoc networks

Virender Kumar [1], Rachit Manchanda[2], Hemant Patel[3], Gowrishankar J.[4], S. Srinadhraju[5,6], Rajesh Singh[7], Nitin Kumar[8], Rupesh Gupta[9], Abhijit Bhowmik[10,11] & Balewgize A. Zeru[12]✉

Vehicular Ad-hoc Networks (VANETs) face significant challenges in ensuring reliable data delivery due to the highly dynamic mobility of vehicles and frequent link disruptions. Traditional clustering approaches often suffer from unstable Cluster Head (CH) selection, which degrades packet delivery, increases communication overhead, and other relative metrics also get effected. To address these limitations, this paper proposes an Enhanced Clustering-based Efficient Relay Vehicle (ECERV) selection framework that integrates a dual-strategy CH selection mechanism with predictive relay support from Road Side Units (RSUs). The proposed method combines stability and proximity factors to improve cluster lifetime, while RSUs dynamically predict optimal relay vehicles for data forwarding in uncovered regions. Extensive NS-2 simulations demonstrate that the proposed ECERV selection scheme outperforms its competitive approaches and achieves 23% higher throughput, 12% higher packet delivery ratio, 25% higher requested data completeness, and 14% lower delay compared to baseline protocol, while also reducing control overhead by 24%, energy consumption by 8.5%, and extending the cluster stability period by 50% under a 100-vehicles scenario. These results confirm that the proposed ECERV selection scheme provides a scalable and robust solution for enhancing data dissemination in highly mobile vehicular communication scenarios.

Keywords VANETs, Road Side Units, V2V, V2I, Valid Connection Time, Message Deadline, Static factor, Link Life Time, Cluster Vehicle, Cluster Head, Relay Vehicle

The development of Intelligent Transportation Systems (ITSs) has been largely driven by factors such as increasing traffic congestion, rising road accidents, and the growing demand for internet connectivity to support various applications. As a result, Vehicular Ad-hoc Networks (VANETs) have emerged as a key solution for enabling these applications[1]. VANETs are a specialized subset of Mobile Ad-hoc Networks (MANETs) and exhibit unique characteristics that distinguish them from conventional MANETs. These differences primarily arise from the high mobility of vehicles, fluctuating vehicle density, and the fact that power consumption is not a major constraint in vehicles. VANETs have garnered significant research interest, as they facilitate a wide range of internet-based services for vehicles through fixed Road Side Units (RSUs). These services include both safety-critical applications, such as collision prevention and real-time traffic monitoring, as well as non-safety applications like weather updates and internet access. To ensure seamless communication between vehicles and infrastructure, several protocols and standards have been established. Key examples include Dedicated Short Range Communication (DSRC), IEEE

[1]Department of Electronics and Communication Engineering, Chandigarh University, Mohali 140413, Punjab, India. [2]Department of Computer Science Engineering, Chandigarh University, Mohali 140413, Punjab, India. [3]School of Engineering and Technology, Dr. Subhash University, Junagadh 362001, India. [4]Department of Computer Science Engineering, School of Engineering and Technology, JAIN (Deemed to be University), Bangalore 560002, Karnataka, India. [5]Department of CSE, Raghu Engineering College, Visakhapatnam 531162, Andhra Pradesh, India. [6]Sharda School of Engineering and Sciences, Sharda University, Knowledge Park III, Greater Noida, Uttar Pradesh 201310, India. [7]Uttaranchal Institute of Technology, Uttar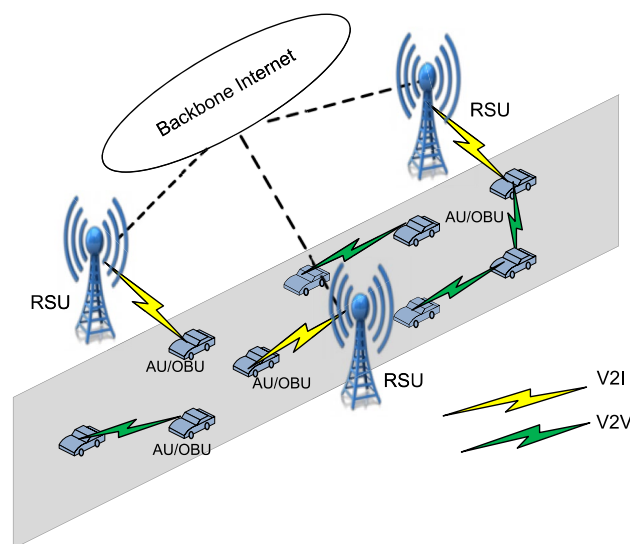anchal University, Dehradun 248007, Uttarakhand, India. [8]Graphic Era (Deemed to be University), Dehradun 248002, Uttarakhand, India. [9]Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India. [10]Department of additive manufacturing, Mechanical Engineering, SIMATS, Saveetha Institute of Medical and Technical Sciences, Thandalam, Chennai, India. [11]Division of Research and Development, Lovely Professional University, Phagwara, Punjab, India. [12]Faculty of Mechanical Engineering, Jimma Institute of Technology, Jimma University, Jimma, Ethiopia. ✉email: balewgize.zeru@ju.edu.et

802.11p, and the IEEE 1609 standard, all of which play a crucial role in ensuring reliable and efficient communication within VANETs[2]. To further enhance communication capabilities in vehicular networks, the Institute of Electrical and Electronics Engineers (IEEE) introduced the Wireless Access in Vehicular Environments (WAVE) standard, which is built upon DSRC technology. WAVE extends the IEEE 802.11 standard, incorporating the 802.11p protocol along with the IEEE 1609 standard, to specifically address the unique requirements and challenges of vehicular communication[3]. Beyond improving road safety through collision avoidance and real-time traffic control, VANETs also facilitate non-safety functions such as providing internet access and environmental data to users. Communication within VANETs can take three forms: Vehicle-to-Infrastructure (V2I), Vehicle-to-Vehicle (V2V), and Hybrid Vehicle (HV) communication. For effective data exchange, an On-Board Units (OBUs), sensors or radio interfaces are installed on vehicles that enable interaction with RSUs or other vehicles. HV communication integrates both V2I and V2V communications, allowing vehicles to leverage multiple communication methods simultaneously[4]. The architecture of the vehicular network is shown in Fig. 1.

The rapid movement of vehicles in VANETs results in the formation of a dynamically changing network topology and short communication times for both V2V and V2I interactions. To tackle the challenge of frequent network partitions caused by this high mobility, Vehicular Delay-Tolerant Networks (VDTNs) leverage the characteristics of Delay-Tolerant Networks (DTNs). VDTNs incorporate the key features and strategies used in DTNs to address issues related to intermittent connectivity and delays in message delivery, ensuring robust communication in VANETs. In VDTNs[5], several RSUs are installed along the highways, connecting vehicles to the internet. The high cost of deploying RSUs makes it difficult to achieve full coverage across the entire road network. Consequently, there are gaps between adjacent RSUs where no coverage is available, referred to as uncovered areas. The duration for which a vehicle stays in this uncovered region is referred to as the outage time. If the connection between an SV and a DV is interrupted while in an uncovered region, Relay Vehicles (RVs) serve a vital function by forwarding messages from the SV to the DV. Numerous algorithms and methods have been put forth to select appropriate RVs for efficient data dissemination within the uncovered area, with the aim of reducing outage time in VANETs. Furthermore, in scenarios where there are multiple SVs and DVs, clustering becomes essential for maintaining connectivity among vehicles in VDTNs. Clustering helps to organize vehicles into groups, facilitating effective communication and coordination within the network. In VANETs, clustering[6] is employed to partition the extensive network into smaller groups of mobile vehicles. This approach aims to enhance routing efficiency, improve information dissemination, and facilitate data gathering within the network.

By organizing vehicles into clusters, communication and coordination among vehicles can be more effectively managed. Clustering allows for localized decision-making, optimized resource allocation, and efficient data exchange within the smaller groups, resulting in improved overall network effectiveness with respect to routing, information dissemination, and data gathering. In VANETs, the selection of CHs is accomplished through various clustering techniques or algorithms. In cluster-based VANETs, vehicles are first organized into clusters, and then the CHs are chosen using a beaconing process. Clustering is important in VANETs due to factors such as reduced mobility impact, lower vehicle congestion, higher reliability, greater stability, and the ability to facilitate cooperative communication. In this paper, an Enhanced Clustering approach for Efficient Relay Vehicle (ECERV) selection scheme is proposed for VANETs. The scheme involves forming clusters and selecting CHs based on their position and stability. RSUs form clusters and select a CH based on their coverage range, the remaining connection duration and Cluster Head Coefficient ($CH_{coefficient}$). After identifying the requested or missing data, the RSU determines the DV, and the CHs transmit the required data to the corresponding DVs. The proposed ECERV selection scheme aims to optimize RV selection in VANETs, leveraging the benefits of clustering to enhance data dissemination and improve overall communication efficiency within the network. This paper makes notable contributions, which can be summed up as follows:

Fig. 1. The VANET Architecture.

- This paper proposes a novel strategy for CHs selection in clusters by combining two different strategies. This new approach aims to enhance the efficiency and effectiveness of CH selection in VANETs
- This paper presents an enhanced clustering strategy for efficient RV selection, tailored for vehicular communication environments. In this framework, the CHs linked to their corresponding RSUs are designated to serve as RVs.
- The proposed ECERV protocol integrates two complementary CH selection strategies: (i) Centroid-Proximity, which minimizes intra-cluster distances through the Closeness Factor (*CF*), and (ii) Stability-based, which maximizes cluster lifetime through the Weighted Stabilization Factor (β_{WSF}). These are combined via a tunable parameter, ϵ, in the formula $CH_{coefficient} = \epsilon \cdot CF_j^m + (1 - \epsilon) \cdot \beta_{WSF_j}$. Unlike the comparative approaches, which rely on a single criterion, The proposed ECERV selection approach uniquely fuses spatial and temporal metrics and further resolves ties using Vehicle Degree and Available Bandwidth.
- The proposed *ECERV* selection scheme demonstrates superior performance over existing approaches by attaining higher throughput and packet delivery ratio, along with reduced delay and significant enhancement in requested data completeness. Furthermore, it minimizes Control Overhead and Energy Consumption, while extending the cluster stability period compared to baseline protocols.

The rest of the paper is structured as follows: Section II discusses the related work. Section III explains the network architecture and system model utilized in the proposed approach. Section IV outlines the proposed methodology. Section V assesses the performance of the suggested strategy and compares it to other established methods. Section VI brings this effort to a close.

Related work

This section reviews related research efforts that aim to enhance data transmission and accuracy in cluster-based VANETs, as well as optimizing RV selection in VANETs. Previously, the capacity of path links was a key factor in selecting an RV. A new relaying scheme has been suggested, which takes into account both the path link capacity and the vehicle's location. According to this scheme[7], vehicles with higher link capacity are preferred as RVs. However, it has been observed that this selection method performs poorly when it comes to transmitting missing content to the DV. To address this limitation, a data transmission method utilizing clustering is proposed. In this approach, a CH is chosen as the RV. The CH then takes responsibility for transmitting data to the DV if the DV is located outside the coverage area or is unreachable by other means. This clustering-based technique[8] aims to improve the efficiency of content delivery to the DV even in challenging situations. The clustering-based data transmission scheme has certain drawbacks. Firstly, it lacks a mechanism to transmit missing data effectively. Secondly, the primary focus is on selecting the CH, which might overshadow other critical aspects of the transmission process. To address these issues, a different approach is proposed in reference[9]. The Adaptive Carry Store Forward (ACSF) scheme is specifically designed for two-hop Vehicular Delay-Tolerant Networks (VDTNs). In this approach, a passing vehicle is selected as the RV to temporarily store and forward missing data when the DV comes within range of a RSU. The selection of the RV is based on minimizing the outage or run-out time for the DV. However, despite these enhancements, further reductions in outage time are necessary to optimize data transmission efficiency. In addition to the previous schemes, reference[10] proposes a cooperative relay vehicle selection scheme tailored for LTE-A (Long-Term Evolution Advanced) networks. This approach considers both network performance and outage reduction. Efficient utilization of the relay terminals helps minimizing outage time significantly. This cooperative approach aims to optimize the network's performance and ensure more reliable data transmission between the base station and the eNodeB (Evolved Node B), leading to improved overall communication efficiency in LTE-A networks. In reference[11], the outage time is effectively minimized through the implementation of a relaying scheme considering many-hop-store and carry-forward approach. An opportunistic relay node selection approach is introduced in this context. To identify potential relay candidates, the scheme leverages the variability of the rapidly changing fading channel, which exists due to variations in signal strength among multiple users. RSUs play a critical role in this process by optimally rate-adapting their transmission of packets to the vehicles within their coverage range. When RSUs broadcast packets, nearby relay candidates are assessed based on their ability to successfully decode these packets. The RSU then selects one of the relay candidates in close proximity that demonstrates successful packet decoding capability. This selection process ensures that the chosen RV is well-suited to enhance data transmission, leading to a reduction in outage time and improved overall network performance.

In reference[12], the Bivious RV selection scheme introduced a novel approach where both trailing and leading vehicles are utilized as RVs. Interestingly, in this scheme, it is the DVs that undertake the RV selection process instead of the RSUs. The concept of speed optimization is employed in this scheme to make RV selections. However, a disadvantage of this RV selection method is that DVs must wait until they arrive within the coverage area of the very next RSU, when a RV just moving behind finishes sending the data. This waiting time for DVs can result in delays and reduced efficiency in data transmission. As a result, further optimizations or improvements may be required to minimize waiting periods and enhance the overall performance of the Bivious RV scheme. In reference[13], the authors of the paper propose an algorithm for selection of hybrid RV that combines two strategies to select two RVs. The first strategy involves selecting the vehicle nearest to the SV as the RV for the next hop in the data transmission process. The second strategy involves selecting the vehicle closest to the DV as the RV for disseminating packets. However, a disadvantage pointed out in the paper is the selection of two RVs. This could potentially introduce complexities and overhead in the system, as managing and coordinating multiple RVs can be more challenging than a single RV selection approach. In references[14] and[15], the limitation of selecting only two RVs in the bivious scheme for VDTNs has been overcome. These papers propose two different approaches to address this issue. (1) Multiple RVs Selection Scheme: This scheme involves RSUs selecting multiple vehicles as RVs to ensure authentic content retrieval in an uncoverage region between two neighbouring RSUs. By having multiple RVs involved in the data transmission process, the likelihood of successful data delivery is increased, especially in areas with limited coverage.

(2) Improved Bivious Scheme: This enhanced version of the bivious scheme optimizes the RV selection process to reduce outage time. RV selection is initiated when a vehicle cannot obtain all the requested content or data within the range of RSUs. The improved scheme employs various metrics to choose multiple RVs that can efficiently relay data, resulting in considerably reduced outage time. By leveraging these novel approaches, both[14] and[15] aim to enhance the reliability and efficiency of data transmission in VDTNs, particularly in challenging situations where traditional methods might fall short. The authors introduce a RV selection algorithm[16] that focuses on optimizing throughput in vehicular networks. They also present a model for evaluating the transmission performance of RVs in these networks. To analyze the message dissemination performance of RVs, they apply the theory of network calculus, which provides a mathematical framework for understanding network behavior and performance. Additionally, the paper proposes several clustered-based schemes or algorithms. These schemes aim to enhance the stability and data dissemination efficiency in VANETs. By organizing vehicles into clusters, data can be more effectively distributed and relayed among the vehicles, leading to improved network stability and overall performance. By combining these approaches, the authors seek to contribute to the advancement of efficient and reliable communication in vehicular networks, thereby facilitating various applications and services that rely on seamless data transmission in VANETs.

In reference[17], the authors propose an algorithm for CH selection in a vehicular network. The selection process involves grouping neighboring CVs and selecting two CHs for each group: the primary CH and the Backup CH. To achieve the stability of the cluster, the algorithm takes into consideration the knowledge of CVs' behavior during the CH selection process. This means that certain factors related to the behavior of CVs are considered when deciding which vehicles will act as CH and Backup CH for the group. By carefully selecting CHs based on the knowledge of CV behavior, the algorithm aims to create stable clusters within the vehicular network. Stable clusters are essential for efficient communication, data dissemination, and coordination among vehicles in the network, which ultimately contributes to the overall performance and reliability of the vehicular communication system. The authors propose an efficient and reliable data transmission algorithm[18] for vehicular networks, utilizing the concept of clustering. The clustering process is based on the evaluation of link reliability between vehicles. The algorithm's primary objective is to improve data transmission efficiency and reliability by addressing the issue of unstable neighbouring CVs. To achieve this, the algorithm identifies and handles redundant or unstable CVs during the clustering process. By using link reliability as a basis for clustering and incorporating redundancy management, CH selection, and cluster maintenance, the proposed algorithm aims to achieve efficient and reliable data transmission in vehicular networks, ultimately enhancing the overall performance of the communication system. The following section of the paper focuses on the network architecture and the general model employed in the proposed work.

Network architecture and general model

There are various difficulties associated with VANETs, and one of these challenges involves the choice of the best RV among multiple vehicles for transmitting data either from SV to DV or delivering requested data to DV. This issue arises due to the extensive network coverage area and the high mobility of vehicles, which often results in frequent disconnections between vehicles and RSUs. The recurring disconnection hinders the transfer of requested data from RSUs to the DV, particularly when the DV is at a considerable distance or within the coverage range of another RSU. Additionally, the criteria used by RSUs to select the efficient RV among numerous vehicles pose another problem in this context.

In this system architecture it is assumed that RSUs are deployed in such a way so that every vehicle is in RSU's range. A formation of cluster occurs when vehicles enter in RSU's range and, thereafter, a CH is selected for individual clusters. The vehicles in a cluster are referred as Cluster Vehicles (CVs) which are allowed to communicate with their corresponding CHs. Further, only CHs are allowed to be in communication with RSUs. Figure 2, illustrates the network structure along with the overall system model.

The communication channel between the CH and the RSU is modeled as a Nakagami-m Fading Channel, where "h_1" represents the gain of channel. The probability distribution function (PDF) for this channel[20], is given as:

$$f(h_1) = \frac{2m^m}{\Omega(d)^m \Gamma(m)} (h_1)^{2m-1} \exp\left(-\frac{m}{\Omega(d)}(h_1)^2\right) \tag{1}$$

where, the parameter m in the Nakagami fading channel represents the Nakagami fading parameter, and its value should be greater than 1/2 and $\Gamma(m)$ denotes the Gamma function, as indicated in book[21]. $\Omega(d)$ denotes the power loss, which can be determined using the following equation[22].
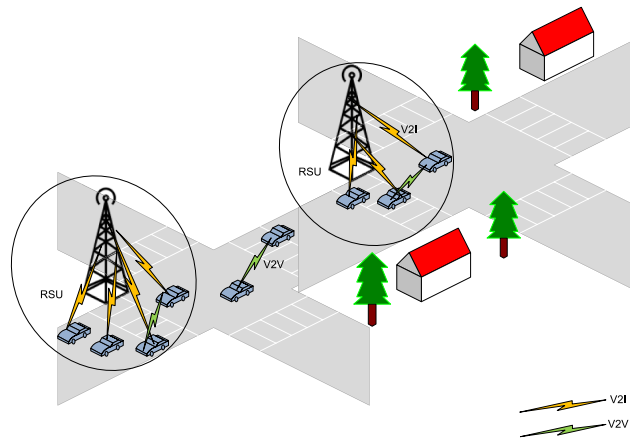
$$\Omega(d) = \frac{P_t G_t G_r h_t^2 h_r^2}{d^\theta L} \tag{2}$$

where, P_t stands for transmitted power, G_t represents the transmitter's antenna gain, G_r represents the receiver's antenna gain, h_t is the height of the transmitter's antenna, h_r is the height of the receiver's antenna. The θ and L are used to denote the path loss exponent and system loss, respectively.

Within a cluster, the communication between CVs or between a CV and its associated CH can be represented as a cascaded Nakagami-m-fading channel with a cascading factor of 2. The PDF for the channel gain h_2 which includes all relevant parameters [reference[22] and reference[23], can be expressed as follows:

$$f(h_2) = \frac{2}{h_2 \Gamma(m_1) \Gamma(m_2)} G_{0,2}^{2,0} \left[\frac{m_1 m_2 (h_2)^2}{\Omega_1 \Omega_2} \middle| \begin{array}{c} - \\ m_1, m_2 \end{array} \right] \tag{3}$$

where, $G_{0,2}^{2,0}$ denotes Meijer G-function.
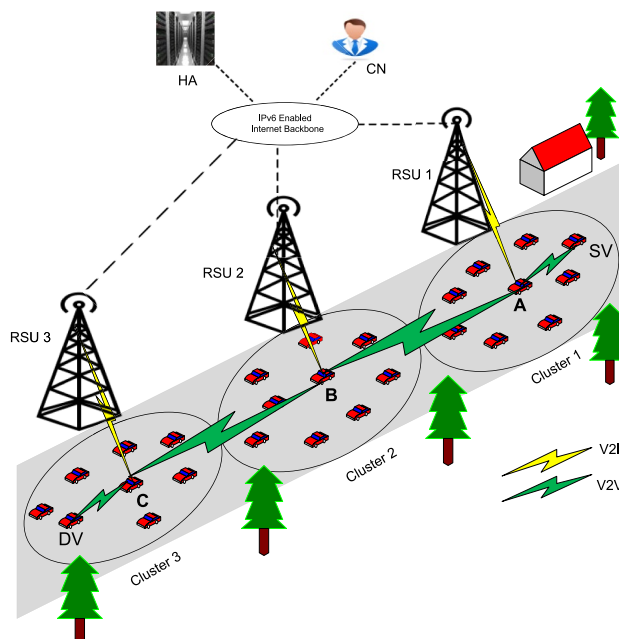
Fig. 2. Architecture with System Model[19].

Uncovered area

Uncovered Area refers to a road segment that lies outside the transmission range of any RSU. Vehicles traveling through uncovered areas are unable to communicate directly with infrastructure, which can lead to disrupted data delivery. To overcome this limitation, the proposed ECERV selection scheme employs RVs, often chosen from CHs, to forward data until connectivity with the next RSU is restored.

The proposed enhanced clustering approach for efficient relay vehicle selection in VANETs

This section outlines the proposed Enhanced Clustering approach for Efficient Relay Vehicle (ECERV) selection scheme in VANETs, where CHs are chosen as RVs. The primary objective of this approach is to enhance network performance, focusing on parameters like Throughput, Data Communication Delay, Requested Data Completeness, Packet Delivery Ratio, Cluster Stability Period, Control Overhead and Energy Consumption. It achieves this by selecting CHs as RVs instead of opting for any arbitrary CV as an RV. By incorporating the concept of clustering into the proposed scheme, it reduces the frequent disconnections between CVs and their respective RSUs. Instead, it establishes a direct connection between DV and SV through CHs acting as RVs, regardless of the number of clusters and their corresponding RSUs. Figure 3, depicts the VANET scenario with vehicles organized into clusters, illustrating the proposed scheme.

In this vehicular network setup, there are three clusters consisting of moving vehicles, all traveling at the same speed. These clusters are serviced by three RSUs. Each CV communicates with other CVs or CHs using OBUs installed in the vehicles. However, only CHs are permitted to communicate with their corresponding RSUs.

Fig. 3. Clustering Scenario in Vehicular Networks for the Proposed ECERV Selection Scheme[24].

There are two types of communications: V2I and V2V, and the characteristics of the communication channels in both V2I and V2V scenarios are modeled according to the specifications outlined in previous Section. The RSUs store information about nearby RSUs and the CHs of their respective clusters, including their location and distance from the RSU's coverage range. Importantly, only the CH in one cluster is allowed to communicate with another CH in a neighboring cluster served by a different RSU. If any CV needs to request data, it does so through its corresponding CH. If the requested data doesn't reach within the RSU's coverage range, the CV becomes a DV. Clustering provides increased network stability, enabling SVs to communicate with DVs through CHs instead of relying solely on RSUs. To determine the CH within a cluster, a novel CH selection procedure is introduced in the proposed approach. This section provides a comprehensive description of the proposed enhanced clustering approach for efficient relay vehicle selection in VANETs.

Methodology

The proposed ECERV selection framework operates in multiple stages, integrating clustering, relay selection, and RSU-assisted prediction to achieve reliable communication in dynamic VANET environments. The methodology is explained step by step as follows:

Cluster formation and CH selection
In the first stage, vehicles periodically broadcast beacon messages containing position, velocity, and residual energy information. A dual-strategy CH selection process is applied:

- **Stability-based metric:** Vehicles with lower relative velocity variance and stronger link lifetime are preferred, ensuring longer cluster stability.
- **Proximity-based metric:** Vehicles closer to the geometric centroid of their neighbors are prioritized, reducing intra-cluster communication cost.
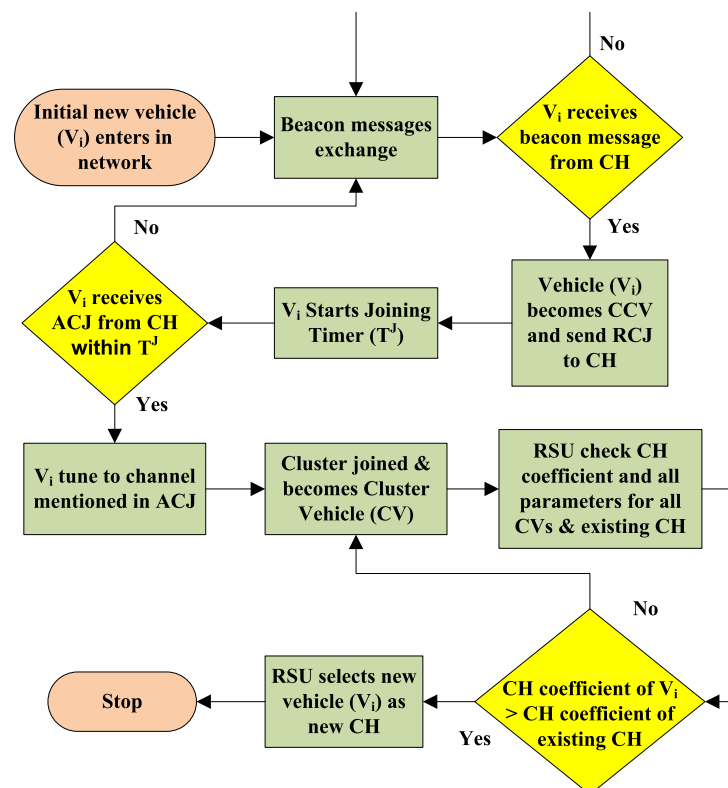
By combining both metrics, the proposed ECERV selection scheme ensures that selected CHs are both stable and centrally positioned.

RSU prediction and relay vehicle determination
RSUs periodically collect cluster information and predict potential link breakages using vehicle mobility patterns. For Destination or Designated Vehicles (DVs) that fall outside RSU coverage, the RSU dynamically determines a RV, usually the CH or a nearby stable vehicle, to forward data towards the DV. This prediction mechanism minimizes packet loss due to frequent topology changes.

Data forwarding process
Once relay vehicles are identified, data packets are forwarded as follows:

Fig. 4. Cluster Formation and CH Selection Flowchart.

- CH transmits collected data to the RSU.
- RSU either delivers data directly (if DV is in range) or selects an RV/CH for forwarding.
- The selected RV ensures successful packet delivery to the DV.

This multi-stage relay mechanism improves delivery ratio and reduces delay compared to existing methods.

Role of clustering in relay selection
Clustering is not only used to organize vehicles into manageable groups but also plays a direct role in efficient RV selection. By limiting the candidate set to CHs, the search space for potential relays is greatly reduced, thereby minimizing control overhead. Moreover, because vehicles within a cluster exhibit similar mobility characteristics, CHs are inherently more stable than arbitrary vehicles, ensuring longer link lifetimes and fewer relay switches. Finally, since CHs already act as coordinators for their clusters, they can seamlessly forward data to DVs, avoiding additional discovery procedures. This makes clustering an integral component of the relay selection strategy in ECERV.

Cluster formation

The proposed RV selection scheme harnesses the clustering based topology for selecting an efficient RV in vehicular networks. Initially, when a new vehicle V_i enters in RSU's range, it receives beaconing messages either from RSU or CH_j of its neighboring cluster. On receiving beaconing messages, V_i becomes Candidate Cluster Vehicle (CCV). Thereafter, V_i sends a message Request for Cluster Joining (RCJ) to CH_j and initialize Joining Timer, T^J. Joining Timer is a time assumed in V_i during which vehicle receives Acceptance for Cluster Joining (ACJ) from CH_j. After receiving ACJ, new vehicle tunes its frequency on channel mentioned in the message ACJ and joins cluster of CH_j. If it does not receive ACJ message from CH_j during T^J, then, it is not allowed to become even CCV. Besides, RSU contains a list of CH coefficient values of all vehicles that are used for the selection of CH in a cluster. Cluster Initialization, Overlap Resolution, and Stable CH Reassignment is mentioned in algorithm 1 and flowchart with CH selection is given in figure 4.

Input: RSU coverage; vehicle set $\mathscr{V}$; transmission range R; CH fusion weight ε; reassignment threshold $\tau > 0$; CH minimum hold time $T_{\text{hold}} > 0$
Output: Clusters $\{C_1, \ldots, C_M\}$ with one CH per cluster

Definitions: For any cluster C, let its centroid (X, Y) be computed by Eq. (6). For any vehicle $v_j \in C$, compute distance l_j by Eq. (7), Closeness Factor CF_j^m by Eq. (9), and weighted stabilization $\beta_{WSF,j}$. Compute the fused CH score:

$$CH_{coefficient}(j) = \varepsilon \cdot CF_j^m + (1-\varepsilon) \cdot \beta_{WSF,j} \quad \text{(Eq. (10))}$$

Periodic beaconing (RSU/CHs) every Δt: advertise cluster IDs, (X, Y), current CH, and summary stats.

On vehicle arrival v into RSU range:
1. *Discover candidate clusters:* $\mathscr{C}_{\text{cand}} \leftarrow \{ C \mid \|(x_v, y_v) - (X_C, Y_C)\| \leq R \}$.

2. *If* $\mathscr{C}_{\text{cand}} = \emptyset$ *then* create new cluster $C_{\text{new}} = \{v\}$; set v as provisional CH; **go to** Step 4.

3. *Overlap resolution (single membership):* For each $C \in \mathscr{C}_{\text{cand}}$, temporarily place v in C and compute $CH_{\text{coefficient}}$ for the current CH of C and for v. Assign v permanently to the cluster $C^\star$ that yields the *largest resulting $CH_{\text{coefficient}}$* for its winning CH (ties broken by Vehicle Degree, then Available Bandwidth).

4. *Update cluster state:* recompute (X, Y) for the chosen cluster; update members and cached metrics.

Boundary maintenance (continuous): For each cluster C and each member $u \in C$:
    if $\|(x_u, y_u) - (X_C, Y_C)\| > R$ **then**
        Remove u from C (candidate for other clusters via the arrival rule).

Stable CH Reassignment (event-driven or every $k\Delta t$): For each cluster C with current CH c:

1. Compute $CH_{\text{coefficient}}(j)$ for all $j \in C$ using Eq. (10).

2. Let $j^\star = \arg\max_{j \in C} CH_{\text{coefficient}}(j)$.

3. **Hysteresis rule: if** $CH_{coefficient}(j^\star) - CH_{coefficient}(c) \geq \tau$ **and** c has served for T_{hold} **then**
    Promote $j^\star$ to new CH; demote c to member; reset CH hold timer.

Return updated clusters and CHs.

Algorithm 1. Cluster Initialization, Overlap Resolution, and Stable CH Reassignment

Notes. Cluster boundaries follow the transmission range R around the centroid (Eq. (6)–(9)); the CH score fuses spatial and stability terms via Eq. (10). To avoid CH flapping, a vehicle only replaces the current CH when its score advantage exceeds the hysteresis threshold τ, and after the current CH has served at least T_{hold}. Single-cluster membership is enforced by selecting the cluster that yields the highest resulting CH score, with ties resolved by Vehicle Degree and then Available Bandwidth.

Lemma 1 (Geometric packing upper bound) The maximum number of (pairwise disjoint) clusters M that can be placed in an area Ω of size A with cluster radius R satisfies

$$M \leq \left\lfloor \frac{\delta A}{\pi R^2} \right\rfloor, \quad \text{where } \delta = \frac{\pi}{\sqrt{12}} \approx 0.9069.$$

Sketch of Proof The densest packing of equal discs in the plane has area density $\delta = \pi/\sqrt{12}$. Hence the total area covered by M disjoint radius-R discs cannot exceed δA. Dividing both sides by πR^2 (area per disc) gives $M\pi R^2 \leq \delta A$, i.e., $M \leq \delta A/(\pi R^2)$. $\square$

Lemma 2 (Counting bound from minimum cluster size) If each cluster must contain at least $s_{\min} \geq 1$ vehicles and there are N vehicles in Ω, then

$$M \leq \left\lfloor \frac{N}{s_{\min}} \right\rfloor.$$

Proof This follows directly from counting: if each cluster contains at least $s_{\min}$ vehicles, then $Ms_{\min} \leq N \Rightarrow M \leq N/s_{\min}$. $\square$

Theorem 1 (Combined bound on the number of clusters) Under the assumptions above, the number of clusters M is bounded by

$$M \leq \min\left\{ \left\lfloor \frac{\delta A}{\pi R^2} \right\rfloor, \left\lfloor \frac{N}{s_{\min}} \right\rfloor \right\}.$$

If the vehicle field is approximately homogeneous with intensity λ (vehicles/m²), then $N \approx \lambda A$ and

$$\mathbb{E}[M] \lesssim \min\left\{ \frac{\delta A}{\pi R^2}, \frac{\lambda A}{s_{\min}} \right\}.$$

Proof This result follows by combining the geometric bound from Lemma 1 with the counting bound from Lemma 2. $\square$

Remark 1   (i) If clusters may geometrically overlap but membership remains exclusive, the effective territories are still disjoint, so the packing bound remains a safe upper bound.

      (ii) If a design (or bandwidth) cap imposes a maximum cluster size $s_{\max}$, then $M \geq \lceil N/s_{\max} \rceil$ gives a lower bound.

Numeric example (based on Table 2 settings). Let the road area be $A = 5000 \times 5000 = 25,000,000 \text{ m}^2$.

- If $R = 300$ m (typical DSRC transmission range):

$$M_{\max} \leq \frac{\delta A}{\pi R^2} \approx \frac{0.9069 \times 25 \times 10^6}{\pi \times 300^2} \approx 80 \text{ clusters.}$$

- If $R = 1000$ m:

$$M_{\max} \approx \frac{0.9069 \times 25 \times 10^6}{\pi \times 1000^2} \approx 7 \text{ clusters.}$$

If a minimum cluster size constraint $s_{\min}$ is applied, the bound can be further tightened as $M \leq \lfloor N/s_{\min} \rfloor$.

Enhanced cluster's stability
In the proposed work, cluster stability is improved by adopting the strategy given in reference[25]. The network comprises N total vehicles, each capable of obtaining its mobility information. Let V_i be a moving vehicle with coordinate location (x_i, y_i) and following a normal distribution $\mathcal{N}(\mu_i, \sigma_i^2)$. Initially, vehicles are divided into M clusters, where $M < N$. Each cluster is represented as C_i for $i = 1, 2, \ldots, M$, and the vehicles within cluster C_i are denoted as V_i^k, where $k = 1, 2, \ldots, S_i$. All symbols and their description are given in table 1. The procedure to enhance cluster's stability is mentioned in algorithm 2.

The average velocity, V_i^{avg}, of the ith vehicle in cluster C_i is determined as follows:

$$V_i^{avg} = \frac{1}{S_i} \sum_{k=1}^{S_i} \mu_{i,k} \qquad (4)$$

The velocity deviation factor ϕ is given by:

$$\phi = \left| \frac{\mu_{i,k} - V_i^{avg}}{V_i^{avg}} \right| \qquad (5)$$

Input: V, information about vehicles
Output: M, C_i, S_i where $i = 1, 2, \ldots, M$
Step 1: Divide vehicles into M clusters based on geographical location;
Step 2: Calculate the average velocity of vehicles in each cluster C_i;
for $i = 1$ **to** M **do**
    for $K = 1$ **to** S_i **do**
        if $Q_{i,K} \geq 0.25$ **then**
            Delete vehicle $V_{i,K}$ from cluster C_i;
        end if
    end for Update cluster C_i and the number of vehicles S_i;
end for

Algorithm 2. Vehicle Clustering and Filtering

Cluster head selection

In the proposed ECERV selection scheme, CH selection is based on the integration of two complementary strategies. The first strategy is the *centroid-proximity strategy*, represented by the Closeness Factor, CF_j, which ensures that the CH is located near the cluster centroid to minimize intra-cluster distance. The second strategy is the *stability-based strategy*, represented by the Weighted Stabilization Factor (β_{WSF}), which measures the velocity stability of vehicles to enhance cluster lifetime. We fuse these strategies through a tunable coefficient ϵ, where the cluster head coefficient is defined as: $CH_{coefficient} = \epsilon \cdot CF_j^m + (1 - \epsilon) \cdot \beta_{WSF_j}$, with $0 \leq \epsilon \leq 1$. The vehicle with the highest $CH_{coefficient}$ is selected as CH. In case of a tie, we employ Vehicle Degree and Available Bandwidth as secondary tie-breakers. This multi-criteria fusion is the novelty of proposed approach ECERV, as existing protocols generally rely on a single selection metric.

Cluster head coefficient
A clustering algorithm based on the relative velocity of a vehicle is presented[26] for determining a factor referred to as Weighted Stabilization Factor (β_{WSF}). The vehicle bearing the highest β_{WSF} is chosen as the CH. When a new CV joins a cluster, it can assume the role of a CH if it meets all the necessary criteria and satisfies the $CH_{coefficient}$ requirements. The selection process for CHs is structured to minimize frequent changes, ensuring cluster stability.

In this proposed approach, a $CH_{coefficient}$ is utilized to determine the CH by considering both the transmission range of vehicles and their velocity. Vehicles with higher $CH_{coefficient}$ values are prioritized for CH selection[27]. Let N represent the total number of vehicles in vehicular network, and their positions be $(X_1, Y_1), (X_2, Y_2), \ldots, (X_N, Y_N)$. The centroid of the cluster can be determined using the following equations:

$$X = \frac{\sum_{i=1}^{N} X_i}{N}, \quad Y = \frac{\sum_{i=1}^{N} Y_i}{N} \qquad (6)$$

Symbol	Description
N	Total number of vehicles
V	Vehicle
μ_i	Mean
σ_i	Standard deviation
M	Total number of clusters
C_i	Cluster i^{th}
S_i	Total count of vehicles within cluster C_i
x_i	Coordinate along the direction of the moving vehicle
y_i	Coordinate that is orthogonal to the road's direction

Table 1. Notations and Their Descriptions.

where, The point (*X, Y*) represents the center of the cluster and the distance of every j^{th} vehicle from centroid i.e. l_j, is determined as:

$$l_j = \sqrt{(x_j - X)^2 + (y_j - Y)^2} \tag{7}$$

Closeness Factor (CF_j) is the factor which indicates that how closer is the vehicle to centroid and is determined according to reference[27] as:

$$CF_j^m = \left(1 - \frac{l_j}{R}\right) \tag{8}$$

In this proposed scheme this CF_j is multiplied with Link Reliability, *R*(*l*), so that, selection of *CH* ensures stronger and stable connection in cluster. The proposed Modified Closeness Factor, CF_j^m is calculated by the equation mentioned below:

$$CF_j^m = \left(1 - \frac{l_j}{R}\right) \times R(l) \tag{9}$$

where, *R*(*l*) is Link Reliability, *R* is transmission range of vehicle and CF_j^m is Modified Closeness Factor of jth vehicle. Higher value of CF_j^m of vehicle indicates that it is closer to centroid. Figure 5 illustrates the integrating of two strategies leading to improved $CH_{coefficient}$ with the modified CF_j^m. An improved $CH_{coefficient}$ in terms of CF_j^m and β_{WSFj} is determined in the following equation. Vehicles having higher values of $CH_{coefficient}$ are prioritized as CH[28].

$$CH_{\text{coefficient}} = \epsilon.CF_j^m + (1 - \epsilon).\beta_{\text{WSF}_j} \tag{10}$$

where, ϵ determines the relative importance of velocity compared to range and $\beta.WSF_j$ defines a weighted stabilization factor for the jth vehicle, which should be greater for a chosen CH[29]. In a network with dense vehicle movement, multiple clusters exist, each with its own CH. The RSU identifies clusters within its coverage area based on vehicle velocities and the beaconing process[30]. After detecting a cluster, the RSU selects a CH from within its coverage range.

Practical interpretations of the metrics
Eq. (4)–(5) summarize average and relative velocity within a cluster, used to quantify local mobility dispersion. Eq. (6)–(7) define the centroid and per-vehicle distance to the centroid, and Eq. (8)-(9) normalizes this proximity as the *Closeness Factor*, $CF_j^m \in [0, 1]$ (larger CF_j^m means closer to the centroid). Eq. (10) fuses spatial compactness and temporal stability via $CH_{\text{coefficient}}(j) = \epsilon\,CF_j^m + (1 - \epsilon)\,\beta_{\text{WSF}_j}$; vehicles with larger $CH_{\text{coefficient}}$ are preferred as CH. Here, β_{WSF_j} increases as the vehicle's relative-velocity dispersion decreases, indicating better cluster stability.

Vehicle mobility
Within a cluster, all vehicles travel at the same velocity, resulting in a minimal velocity difference between CVs and the CH. This enhances cluster stability. When selecting the efficient CH, the vehicle with the least velocity difference is preferred. If v_j^k and v_i^k represent the velocities of the jth and i^{th} vehicles in the kth cluster, respectively, average velocity difference (∂V) can be determined as follows:

$$\partial V = \frac{\sum_i |v_j^k - v_i^k|}{2N_{\text{max}}V_{\text{max}}} \tag{11}$$

where, *i* is an element of the set containing all parameters within the kth cluster, N_{max} denotes the total number of CVs in the cluster and V_{max} signifies the highest velocity among the CVs.
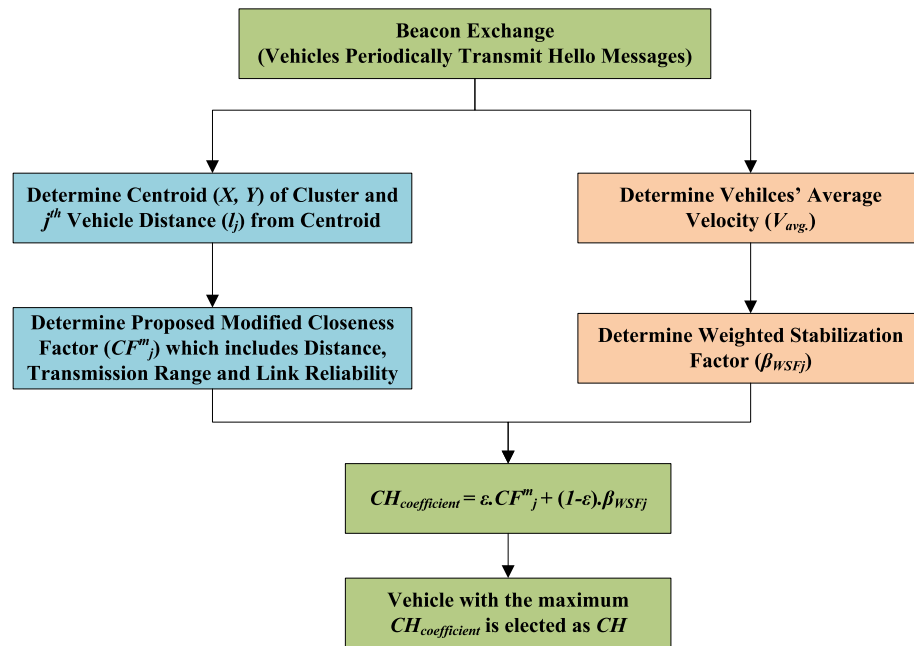
Vehicle degree
Vehicle degree refers to the highest number of CVs in a cluster that can establish a direct link with the CH. Therefore, the CH should be selected based on the highest vehicle degree. The vehicle degree, represented as VD_j^k, indicates the total count of vehicles within the range of the jth vehicle in the kth cluster. To maintain a stable connection among vehicles in a cluster, the maximum number of vehicles considered is N_{max}.

Available bandwidth
When transmitting messages or requested data to DVs, the CH utilizes a certain amount of bandwidth. In this approach, the selection of an efficient CH as a RV also takes into account the average bandwidth consumed by the CH. The calculation for average consumed bandwidth is as follows:

$$B_j^k = \frac{B_{\text{avail}}}{B_{\text{total}}} \tag{12}$$

where, B_j^k denotes the average bandwidth utilized by the jth vehicle in the cluster k^{th}, B_{avail} represents the available bandwidth and B_{total} refers to the total or maximum bandwidth capacity.

Fig. 5. Flowchart Illustrating the Integrating of Two Strategies leading to Improved $CH_{coefficient}$ with the Modified CF_j^m.

Estimation of remaining duration of connectivity for each CH

Based on velocity, the RSU estimates the Remaining Duration of Connectivity (T) for a CH when a cluster enters its coverage area and transmits a beacon message. T, is defined as the duration for which the CH remains connected to the RSU before moving out of its coverage. In this proposed approach, clusters and their CHs are assumed to be traveling in a single direction, specifically from left to right. the T of V_d can be computed as:

$$T_{V_d} = \frac{D_{R_r} - D_{V_d}}{V_{\text{avg}}} \tag{13}$$

where, D_{R_r} represents the position of the RSU's right boundary, V_d is the DV, D_{V_d} represents the location of V_d and V_{avg} is the average velocity of V_d.

Identification of destination vehicle by RSU and delivery of requested data

DVs refer to those that have moved beyond the RSU's coverage area after requesting data. The RSU identifies these vehicles based on their remaining duration of connectivity and the data they have requested.

$$T_{V_d} \times R_{dt} < \text{Req}_{V_d} \tag{14}$$

where, R_{dt} represents the data transmission rate and Req_{V_d} denotes the data requested by the DV i.e. V_d.

If the RSU determines that the remaining duration of connectivity T_{V_d} is insufficient for V_d to receive all requested data, the vehicle is classified as a DV. Upon receiving a data request from V_d, the RSU selects the nearest CH based on its distance and velocity relative to the DV. The RSU then transmits the data to the chosen relay CH. If the distance between the relay CH and the DV is significant, the data may be relayed through an additional CH before reaching the DV.

Selection of cluster head as relay vehicle

To transmit data from a *SV* or a *RSU* to a *DV*, an efficient *RV* is selected among multiple vehicles in a cluster. In the proposed scheme *ECERV*, the *CHs* are selected as *RVs* for each cluster based on specific parameters, including:

- Link Life Time (LLT_{ij}),
- Link Reliability ($LREL_i$) metric of vehicle V_i, and
- Cluster Head Coefficient ($CH_{coefficient}$).

The Link Life Time (LLT_{ij}) is also known as the expiration time of the link between two adjacent vehicles, the jth and i^{th} vehicles, in a cluster. It is basically a predicted duration time beyond which two adjacent vehicles are not connected in cluster and determined[31] as:

$$LLT_{ij} = \frac{|\partial V| \times R - \partial V \times \partial d_{ij}}{(\partial V)^2} \tag{15}$$

where, ∂V is the difference in average velocity between the jth and i^{th} vehicle and ∂d_{ij} is the difference in distance between the jth and i^{th} vehicle. The parameter Link Life Time (LLT) maintains a list of stable neighboring (SN_i) vehicles for each V_i vehicle. The larger the value of LLT_{ij}, the more sustainable the link is.

Stable links and the stable-neighbor set
Following Eq. (15), we deem the link (i, j) *stable* if $LLT_{ij} \geq T_{stable}$. For each vehicle V_i we form the stable-neighbor set $SN_i = \{ V_j \mid LLT_{ij} \geq T_{stable} \}$, which is then used in the link-reliability aggregation (Eqs. (16)–(18)). Intuitively, larger LLT_{ij} implies a longer residual time that two vehicles remain within range, which promotes consistent forwarding opportunities.

Link reliability
The model of Link Reliability between the jth and i^{th} vehicles in urban vehicular networks is presented[32]. The model of Link Reliability may be defined as the conditional probability applied to equation (16), which gives the probability of continuous connectivity of the link between two vehicles.

$$R(l) = P\{l \text{ continues to } t + LLT\} \tag{16}$$

where, $R(l)$ indicates Link Reliability and 'l' represents the link between j^{th} and i^{th} vehicle on condition that 'l' is available at 't'. The above equation indicates that if link is available at time 't' then it will also be available at time (t + LLT). To determine link reliability, speed of vehicles is main parameters. In the proposed ECERV selection scheme, a specific metric i.e. Link Reliability Metric for vehicle V_i is considered for selecting CH as RV. Link Reliability (LREL) metric is determined[18] as:

$$LREL_i(t) = \sum_{v_j \in SN_i} R_t(l_{ij}) \tag{17}$$

where, $R_t(l_{i,j})$ can be calculated as:

$$R_t(l_{ij}) = \begin{cases} \int_t^{t+LLT} f(t)\,d(t), & \text{if } LLT > 0 \\ 0, & \text{otherwise} \end{cases} \tag{18}$$

The CHs of each clusters having higher value of *LREL* and $CH_{coefficient}$, will be selected as RV in the proposed ECERV selection scheme for the purpose of data transmission. RSU contains all the information about metrics of each vehicle in clusters. In case of selection out of *LREL* and $CH_{coefficient}$ by RSU, the preference will be given to *LREL*.

Overall operation of proposed approach
The overall operation of the proposed ECERV selection scheme is shown in Fig. 6 and various steps for data forwarding are depicted in algorithm 3. As illustrated in the flowchart, cluster formation is initiated through a beaconing message process. Once clusters are established, the selection of a CH takes place. In the proposed approach, the efficiently chosen CH also serves as a reliable RV, responsible for forwarding missing data to the DV or transmitting messages from the SV to the DV. Additionally, each RSU predicts the nearby clusters and their CHs while updating CHs with information regarding the velocity, location, and distance of CVs. Vehicles that previously requested data from an RSU within its coverage area and later moved out of range become DVs. To facilitate seamless data delivery, the RSU identifies DVs based on their remaining duration of connectivity and transmits the requested data to the CH. These CHs then act as RVs, relaying the data to CHs of adjacent RSU clusters while searching for the respective DVs. Consequently, data is efficiently delivered to the DV without requiring frequent RSU involvement. The communication process occurs between CHs of different clusters, independent of the number of RSUs or clusters.

The proposed selection scheme integrates two different CH selection strategies and introduces an innovative approach for cluster formation and CH selection. Subsequently, CHs of each RSU-associated cluster are designated as RVs. Due to the high deployment cost of RSUs along roadways, maintaining continuous vehicle connectivity with RSUs at all times is impractical. This results in uncovered regions between adjacent RSUs where no direct connection can be established between RSUs and vehicles. When vehicles enter such uncovered regions, RSUs cannot efficiently provide the requested data. To ensure a stable vehicular network, cluster formation is necessary when vehicles move into a new RSU's range. Following the clustering process, CHs are selected for each cluster, and only these CHs function as RVs to relay the requested data to the DV. The Performance evaluation of the proposed ECERV selection scheme demonstrates significant improvements in throughput, data communication delay, requested data completeness, and packet delivery ratio, making it a more efficient solution compared to conventional approaches.

Data: Vehicle clusters and RSU information
Result: Data forwarding to the DV
Step 1: Cluster of vehicles is formed;
Step 2: RSU selects CH;
Step 3: The RSU detects clusters and their respective CHs within its communication range;
Step 4: RSU determines T for every CHs;
Step 5: DV is determined by RSU which is out of its coverage range;
Step 6: Total number of RVs are determined;
if *there is demand for missing or requested data by DV* **then**
  | RSU identifies closest CH to DV and assigned it as RV;
  | RSU provides requested data to CH designated as RV;
  | Designated RV forward this missing or requested data to DV;
else
  | There is no requirement for assigning any CH as RV;
DV receives missing or Requested data;

Algorithm 3. Cluster Formation and Data Forwarding

Results analysis and discussion

In this section, the performance of the proposed ECERV selection scheme for VANETs is compared with previous RV selection schemes. The motive behind such comparison is to analyze the impact of stability of cluster on selecting optimal RV and variation in different performance metrics due to variation in speed and number of vehicles. It is shown that, the variation in speed of vehicles and number of vehicles affects the Delay in Data Communication, Completeness of Missing or Requested Data, Throughput, Packet Delivery Ratio, Cluster Stability Period, Control Overhead and Energy Consumption in the vehicular network. The results demonstrated in this section, strongly argue that the proposed ECERV selection scheme for VANETs shows significant improvements in terms of specific performance metrics, when it is compared with previous research work.

Parameterization & thresholds

We select four operational parameters to balance data delivery, latency, and control stability: (i) the fusion weight ϵ in CH_{coef}, (ii) the CH-switch hysteresis τ (minimum advantage to replace the current CH), (iii) the minimum CH hold time T_{hold}, and (iv) the stable-link threshold T_{stable} used in SN_i.

Selection rationale

We conducted a small grid search on representative traces (vehicle counts 60–100; speeds 25–35 m/s) and chose values that maximized PDR while constraining CH churn and end-to-end delay. Specifically, we explored $\epsilon \in \{0.3, 0.5, 0.7\}, \tau \in \{0.03, 0.05, 0.08\}$ (with CH_{coef} normalized to $[0,1]$), $T_{\text{hold}} \in \{2, 3, 5\}$ s, and T_{stable} as the p-th percentile of the empirical LLT distribution with $p \in \{10, 25, 40\}$. Across densities, $\epsilon = 0.5$ provided a robust spatial/temporal balance; $\tau = 0.05$ and $T_{\text{hold}} = 3$ s limited unnecessary CH switches; $T_{\text{stable}} = \text{percentile}_{25}(LLT)$

Fig. 6. Overall Operation of the Proposed ECERV Selection Scheme.

filtered transient links yet preserved enough neighbors for forwarding. Operational parameters and final values used in experiments have been mentioned in table 2.

In summary, $CH_{\text{coefficient}}$ (Eq. (10)) weights centroid proximity and velocity stability (Eqs. (7)–(9)) while τ and T_{hold} ensure CH stability; LLT_{ij} (Eq. (15)) and T_{stable} define stable neighbors used in the link-reliability aggregation (Eqs. (16)–(18)).

Criterion for performance metric selection

In this study, seven key metrics such as: Throughput, Packet Delivery Ratio, Data Communication Delay, Requested Data Completeness, Cluster Stability Period, Control Overhead and Energy Consumption were selected for performance evaluation. These metrics were prioritized because the primary objective of ECERV selection scheme is to enhance data delivery reliability and connectivity in uncovered regions, where packet-level performance directly indicates protocol effectiveness. We note that other metrics, such as energy consumption (important for RSUs/OBUs) and control overhead (relevant to clustering), are also valuable. However, unlike wireless sensor networks, VANET devices are generally less energy-constrained, and clustering-based control is already implicitly reflected through cluster stability and reduced CH reassignments in the proposed ECERV selection approach.

Statistical validation

To ensure that the reported improvements are statistically reliable, each experiment was repeated five times, and we computed the mean along with 95% confidence intervals (CIs). Additionally, we performed two-tailed t-tests between ECERV and each baseline schemes Ahmed et al.[15], Chai et al.[16], and the CORV[24]. The results confirm that the observed gains in throughput, PDR, and delay are statistically significant ($p < 0.05$). This statistical validation strengthens the claim that ECERV consistently outperforms existing approaches. Table 3 shows the statistical validation of the propose ECERV selection approach using two-tailed t-tests (p-values).

Experimental setup

To analyze the impact of variations in vehicle speed and density on key performance metrics such as: throughput, data communication delay, requested data completeness packet delivery ratio, cluster stability period, control overhead and energy consumption, simulations were conducted, as these metrics are the primary focus of this study. The simulations were carried out using Network Simulator[33] (NS-2), version 2.34. A portion of the simulated vehicular environment is illustrated in Fig. 7.These figures are entirely simulation-based outputs and not derived from external satellite imagery. This ensures that no copyright permissions are required, as the maps and backgrounds originate solely from the simulation tools employed. Vehicular mobility traces were generated using a realistic mobility model developed with the SUMO traffic simulation tool, which is integrated with NS-2 [reference[34] and reference[35]. This model produces vehicles moving at speeds ranging from 10 to 40 m/s.The simulations were performed on bi-directional roads with a single lane in each direction, covering a simulation area of 5000 m × 5000 m over a duration of 150 seconds. Each message had a size of 100 bytes and was transmitted at a data rate of 3 Mbps. The ad hoc communication range and data transfer rate were set according to the IEEE 802.11p standard, with vehicle transmission ranges varying between 300 m and 1000 m. A summary of the simulation parameters is provided in Table 4.

Simulated results

This section gives the comparison of the simulated results of the proposed ECERV selection scheme with other previous relevant schemes Ahmed et al.[15], Chai et al.[16] and CORV[24]. These four performance metrics are: delay in data communication, completeness of missing or requested data, throughput, and packet delivery ratio.

In Fig. 8, the performance of Packet Delivery Ratio (PDR) is observed with the increasing number of vehicles. Initially, PDR is low but as the number of vehicles increase, it increases and it remains almost constant. The reason is that, in the proposed scheme cluster's stability is enhanced and CH is selected as RV and if the number CHs increase, relay vehicles also increases.

Figure 9 shows the performance of PDR with speed of vehicles. PDR decreases when speed of vehicles increase. PDR decreases with the increase of speed because of quickly disconnection from other vehicles or RSUs. The proposed ECERV shows considerable improvement with respect to other schemes, because in a cluster, speed of all the vehicles is constant. So, there is lesser effect on PDR.

Figure 10 illustrates the effect of increasing vehicle density on network throughput. In VANETs, throughput generally improves as the number of vehicles rises since more data packets are transmitted, making higher throughput desirable. The approach proposed by Chai et al.[16] primarily considers the selection of a single CH as an RV within a cluster, without enabling communication between CHs of different clusters. Similarly, the

Parameter	Range explored	Selected value
ϵ (fusion weight in CH_{coef})	$\{0.3, 0.5, 0.7\}$	0.5
τ (CH switch hysteresis)	$\{0.03, 0.05, 0.08\}$	0.05
T_{hold} (min CH tenure)	$\{2, 3, 5\}$ s	3 s
T_{stable} (stable link threshold)	$\{P10, P25, P40\}$ of LLT	$P25(LLT)$

Table 2. Operational parameters and final values used in experiments.

Metric	Ahmed et al.[15]	Chai et al.[16]	CORV[24]
Packet Delivery Ratio (PDR)	$p < 0.01$	$p < 0.01$	$p < 0.01$
Throughput	$p < 0.05$	$p < 0.01$	$p < 0.01$
Data Communication Delay	$p < 0.05$	$p < 0.05$	$p < 0.01$
Requested Data Completeness	$p < 0.01$	$p < 0.01$	$p < 0.05$

Table 3. Statistical Validation of ECERV Improvements using Two-Tailed t-tests (P-Values).

Fig. 7. Screenshot of the vehicular network simulation environment. The scenario was generated using SUMO http://sumo.dlr.de) integrated with NS-2. The figure represents simulation outputs only; no third-party or copyrighted satellite imagery has been used.

method introduced by Ahmed et al.[15] emphasizes RV selection, relying on RSU services whenever more than two RSUs are available. In contrast, the proposed ECERV selection scheme incorporates both CHs as RVs and facilitates intra-cluster communication through CHs, enhancing network stability. As a result, a SV in one cluster can communicate with a DV in another cluster without consistently depending on RSUs. This approach significantly improves throughput compared to existing methods.

Figure 11 reveals that throughput decreases as vehicle speed increases, with vehicle speeds considered up to 40 m/s. This reduction in throughput occurs due to frequent disconnections of CVs from RSUs. In the proposed ECERV scheme, the CH collects all requested data within the RSU's coverage and efficiently forwards it to the DV, minimizing collision probability due to the low relative velocity of CVs within clusters. By selecting CHs as RVs, the adverse effects of RSU disconnections are mitigated, leading to a notable enhancement in throughput when compared to competing schemes, including those by Ahmed et al.[15], Chai et al.[16], and the CORV[24] approach.

Figure 12 illustrates the influence of the number of neighboring vehicles around a DV on Requested Data Completeness (RDC), which is defined as the successful reception of all requested data by a DV after moving out of RSU coverage. This data was initially requested when the vehicle was still a CV within the RSU range and later became a DV. Compared to existing schemes, the proposed ECERV approach observes that lower vehicle density results in higher RDC. However, as the number of vehicles increases, RDC decreases due to a rise in packet collisions caused by higher transmission rates in the network.
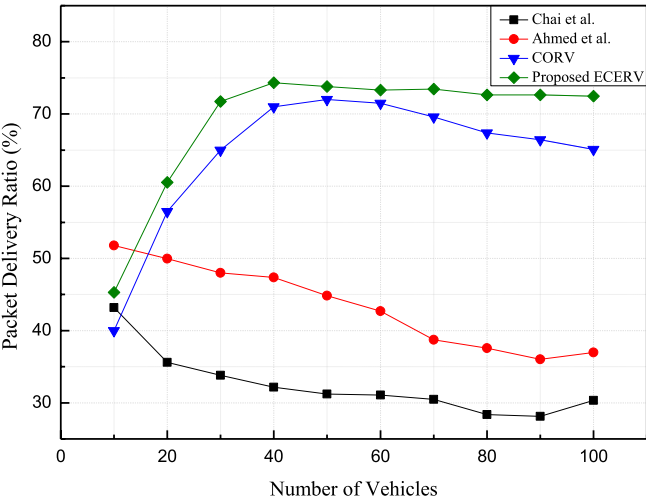
Figure 13 investigates the impact of vehicle speed on RDC. Vehicles with lower speeds exhibit improved RDC performance since they maintain longer connection times with RSUs, allowing them to store and retrieve requested data effectively. Conversely, as vehicle speed increases, connection time with RSUs decreases, leading to reduced RDC. The proposed ECERV scheme achieves a higher RDC percentage by maintaining low relative velocity within clusters and reducing the number of relay vehicles, thereby minimizing packet collisions.

Figure 14 examines the effect of vehicle density on data communication delay, where lower delay is preferred. In Ahmed et al.'s[15] scheme, as vehicle density increases within an RSU's coverage, data communication delay rises due to multiple service requests being processed simultaneously and multiple RVs being selected. Additionally, their method lacks CH selection as RVs for transmitting requested data in uncovered areas. The approach by Chai et al.[16] involves CH selection as RVs, but it does not facilitate inter-cluster communication between CHs. In contrast, the proposed ECERV scheme uses clustering and efficient selection of CH as RVs, reducing the number of required RVs and consequently lowering data communication delay. Furthermore, service requests are consolidated through a single CH, rather than multiple CVs, leading to significantly reduced and more consistent communication delay.

Figure 15 presents the relationship between vehicle speed and data communication delay, demonstrating that the proposed ECERV scheme achieves lower delay compared to the methods of Ahmed et al.[15], Chai et al.[16], and CORV[24]. In these competing schemes, as vehicle speed increases, delay rises due to vehicle mobility and the absence of a parameter accounting for remaining duration of connectivity when determining DV locations. The proposed ECERV approach addresses this by minimizing mobility effects through clustering and forwarding

Parameter	Value
Simulation area	5000 m × 5000 m
Speed of Vehicle	15–40 m/s
Simulation time	150 s
Number of Vehicles	2 (1 in each direction)
Scenario	Urban
MAC interface	Mac/802_11Ext
Network interface	Phy/WirelessPhyExt
Interface queue	Queue DSRC
Message size (non-safety)	1023 bytes
Message size (safety)	100 bytes
Propagation model	Propagation/Nakagami
Transmission range	300 and 1000 m
Channel type	Channel/wireless channel
Antenna type	Antenna/Omni antenna
Data rate	3 Mbps
Modulation type	BPSK

Table 4. Simulation parameters for proposed ECERV selection scheme.

Fig. 8. Number of Vehicles versus Packet Delivery Ratio.

data via CHs to DVs. Additionally, the proposed ECERV algorithm assesses whether a request can be served based on the remaining duration of connectivity and the volume of data that can be transmitted to DVs. This optimization results in a significantly lower and more stable data communication delay in the proposed scheme.

Cluster Stability Period (CSP) refers to the duration (in simulation rounds or time units) for which a cluster remains intact before a re-clustering or CH re-selection event occurs. Figure 16 demonstrates the variations of CSP against number of vehicles. The proposed ECERV shows higher CSP in comparison to other schemes. A longer CSP indicates that the clustering algorithm ensures more stable groupings of vehicles, thereby reducing frequent cluster maintenance overhead and improving routing reliability in VANETs.

Figure 17 presents the graph which gives the variations of Control Overhead (CO) with increasing number of vehicles. CO represents the fraction of control packets (e.g., beacon messages, CH announcements, cluster join requests) relative to the total packets transmitted in the network. A lower CO implies higher efficiency, as fewer resources are wasted on signaling, making the protocol more scalable for dense vehicular networks. The proposed ECERV approach results in lower CO compared to other schemes.
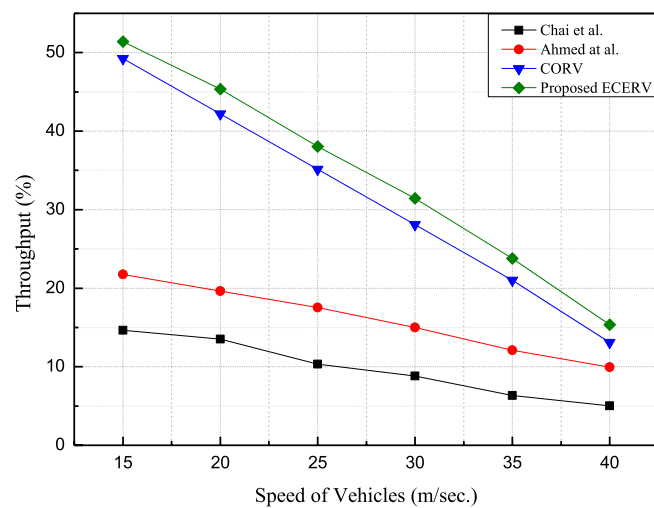
Figure 18 demonstrates the Energy Consumption (EC) with respect to Simulation Time. Energy Consumption refers to the total energy expended by all nodes (vehicles and RSUs) during communication, including transmission, reception, and idle listening activities. Lower energy consumption indicates better utilization of resources and is crucial in extending the operational lifetime of vehicular devices and RSUs. The proposed ECERV approach results in lower energy consumption compared to other schemes.
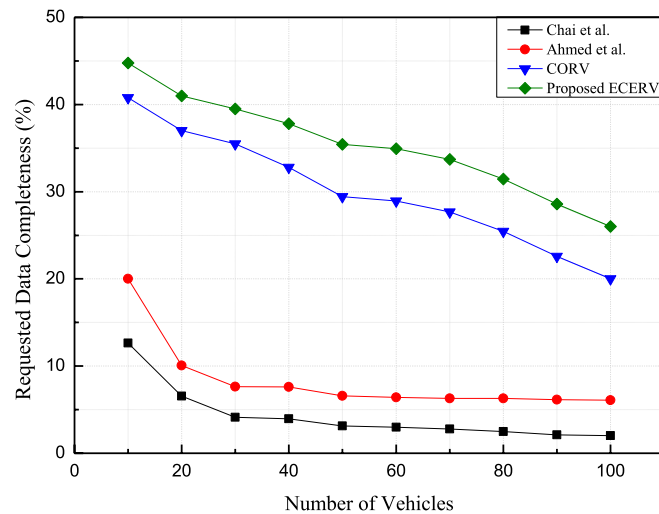
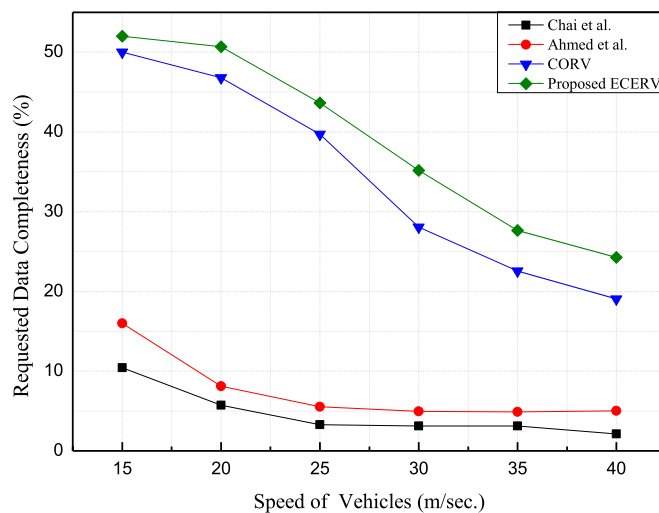**Fig. 9**. Speed of Vehicles versus Packet Delivery Ratio.

Fig. 10. Number of vehicles versus Throughput (Kbps).

Fig. 11. Speed of Vehicles versus Throughput.

Fig. 12. Number of Vehicles versus Requested Data Completeness.

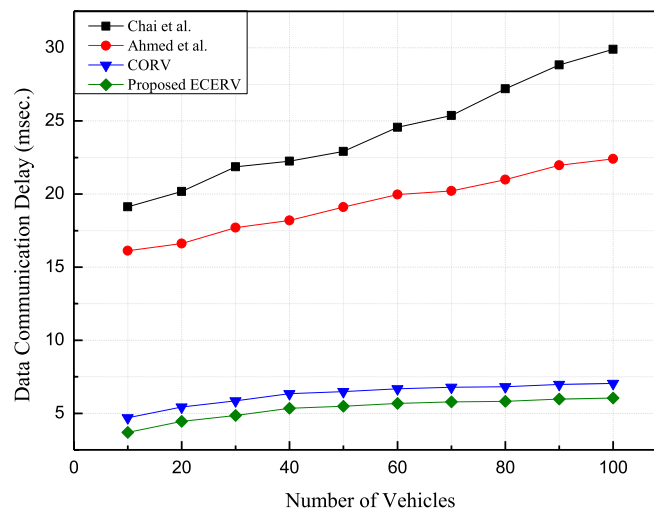**Fig. 13**. Speed of Vehicles versus Requested Data Completeness.

Performance comparison and improvements of metrics of ECERV across protocols

The methods of reference[15], reference[16], and reference[24] were selected as baselines because they are widely cited and represent established clustering-based or relay selection approaches in VANETs. These schemes provide a fair benchmark to highlight the contribution of our clustering and CH-fusion strategy. We note, however, that more recent state-of-the-art methods, particularly those employing Machine Learning for RV prediction and selection, are gaining prominence. Incorporating such ML-based baselines is an important direction for future work to further validate ECERV selection scheme under highly dynamic and large-scale VANETs scenarios. Table 5 gives the comparison of performance metrics across protocols and the percentage improvement of the proposed ECERV selection scheme are given in table6.
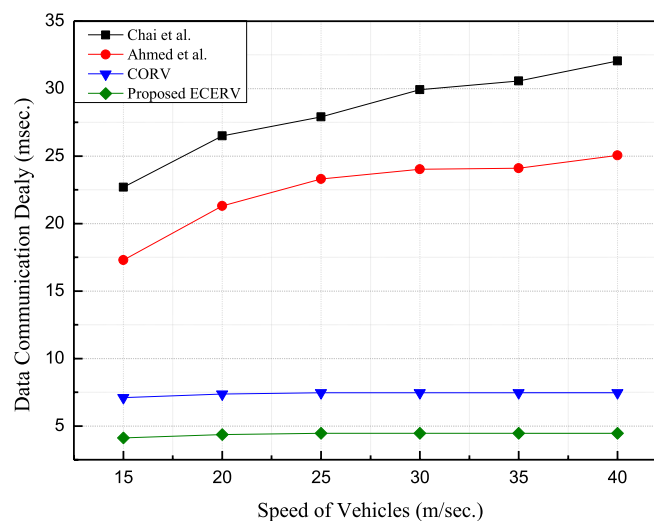
Conclusion

The Enhanced Clustering approach for Efficient Relaying Vehicle selection is introduced to improve data retrieval reliability for DVs in Vehicular Delay-Tolerant Networks. This method identifies the relay vehicle as the cluster head after forming clusters and selecting CHs by considering both vehicle transmission range and velocity. Since clustered vehicles exhibit lower mobility, greater stability, and a reduced risk of message collisions, the proposed scheme ensures more reliable data retrieval for DVs in uncovered regions. As a result, it enhances network performance by achieving higher throughput, reduced communication delay, improved data completeness, and a higher packet delivery ratio compared to conventional RV selection strategies.

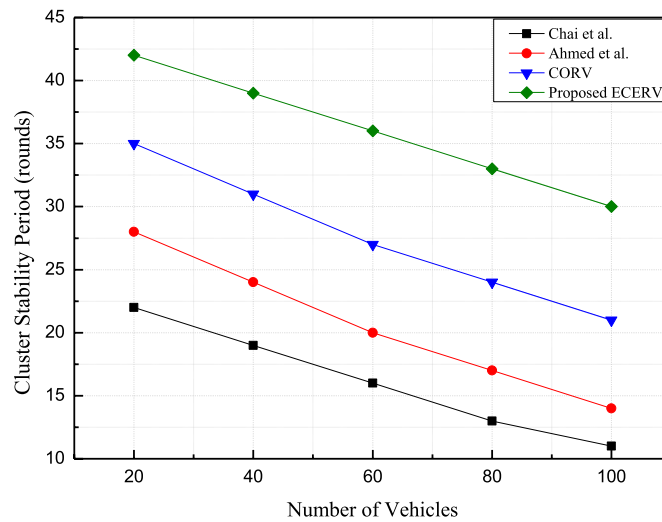**Fig. 14**. Number of Vehicles versus Data Communication Delay.

Fig. 15. Speed of Vehicles versus Data Communication Delay.

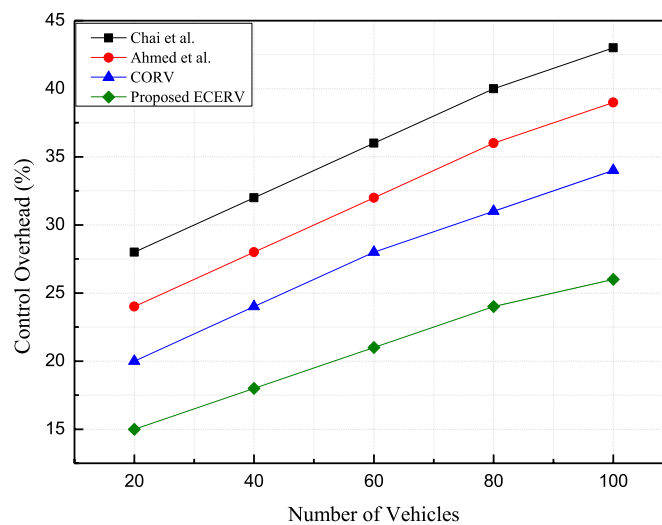### Extension to UAV-based communication

Although ECERV focuses on clustering and relay vehicle (RV) selection among ground vehicles, there may be extreme cases where neither a suitable RV nor an RSU is available in an uncovered area. In such scenarios, Unmanned Aerial Vehicles (UAVs) or drones can act as temporary relay nodes, providing on-demand connectivity and extending coverage. This hybrid approach has the potential to complement ECERV by ensuring communication continuity even in sparse or infrastructure-less regions. In future work, we plan to integrate UAV-assisted communication into the ECERV framework to enhance robustness in challenging environments.
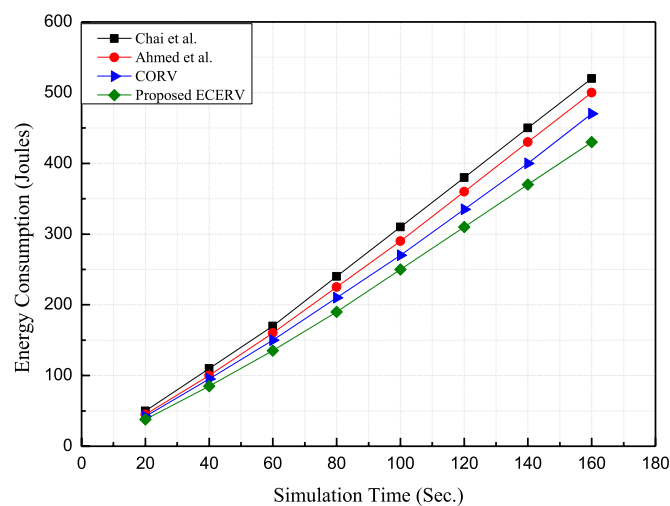
Limitations and scalability

While the proposed ECERV scheme has demonstrated significant improvements under a bi-directional, single-lane road scenario, this setting represents a simplified abstraction of real-world VANETs conditions. In practice, vehicular networks operate over multi-lane highways, dense urban intersections, and heterogeneous RSU deployments, which introduce additional complexities such as frequent lane changes, traffic signal effects, and varying densities of vehicles. Although the clustering and CH-fusion strategy in ECERV is inherently scalable to larger and more complex topologies, additional adaptations may be required. For example, lane-change dynamics can alter cluster boundaries more frequently, and urban intersections may necessitate intersection-aware CH reassignment policies. As part of future work, we intend to extend ECERV to such multi-lane and urban scenarios and to validate its performance in large-scale, mixed-mobility environments. This discussion clarifies the current scope of our simulations and outlines the scalability of the proposed approach.

Fig. 16. Number of Vehicles versus Cluster Stability Period.

Fig. 17. Number of Vehicles versus Control Overhead.

Fig. 18. Energy Consumption versus Simulation Time.

Metric	Chai et al.	Ahmed et al.	CORV	Proposed ECERV
PDR vs Vehicles (%)	30	37	65	73
PDR vs Speed (%)	40	55	70	80
Throughput vs Vehicles (Kbps)	28	35	65	80
Throughput vs Speed (Kbps)	13	20	45	50
RDC vs Vehicles (%)	5	8	22	28
RDC vs Speed (%)	5	7	20	25
Delay vs Vehicles (ms)	30	23	7	6
Delay vs Speed (ms)	32	25	7	5
Control Overhead vs Vehicles (%)	46	42	37	28
Cluster Stability Period (rounds)	9	12	18	27
Energy Consumption vs Time (Joules)	520	500	470	430

Table 5. Comparison of Performance Metrics across Protocols.

Metric	vs Chai et al. (%)	vs Ahmed et al. (%)	vs CORV (%)
PDR vs Vehicles (%)	143.3	97.3	12.3
PDR vs Speed (%)	100.0	45.5	14.3
Throughput vs Vehicles (Kbps)	185.7	128.6	23.1
Throughput vs Speed (Kbps)	284.6	150.0	11.1
RDC vs Vehicles (%)	460.0	250.0	27.3
RDC vs Speed (%)	400.0	257.1	25.0
Delay vs Vehicles (ms)	80.0	73.9	14.3
Delay vs Speed (ms)	84.4	80.0	28.6
Control Overhead vs Vehicles (%)	39.1	33.3	24.3
Cluster Stability Period (rounds)	200.0	125.0	50.0
Energy Consumption vs Time (Joules)	17.3	14.0	8.5

Table 6. Percentage Improvement of Proposed ECERV over baseline Protocols.

Data availability
The data that support the findings of this study are available from the corresponding author, [BAZ], upon reasonable request. The source code and simulation scripts used to implement and evaluate the proposed ECERV framework are publicly available at: https://github.com/rachitece-rgb/ECERV-implementation.git

References
1. Seth, I., Guleria, K. & Panda, S. N. A comprehensive review on vehicular ad-hoc networks routing protocols for urban and highway scenarios, research gaps and future enhancements. *Peer-to-Peer Netw. Appl. (Springer Nature)* **17**(4), 2090–2122. https://doi.org/10.1007/s12083-024-01683-1 (2024).
2. Khan, A. et al. Dsrc technology in vehicle-to-vehicle (v2v) and vehicle-to-infrastructure (v2i) iot system for intelligent transportation system (its): A review. *Recent Trends Mechatronics Towards Ind. 4.0. Lect. Notes Electr. Eng. (Springer)* 97–106. https://doi.org/10.1007/978-981-33-4597-3_10 (2022).
3. Kumar, V., Gupta, N. & Gottapu, S. An efficient rescheduling scheme for prioritizing safety messages in VANETs. In Koganti, K., E., S. & Gupta, N. (eds.) *Advanced Technologies in Electronics, Communications and Signal Processing. ICATECS 2024*, vol. 619 of *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, 14–26, (Springer, Cham) https://doi.org/10.1007/978-3-031-94280-8_2 (2026).
4. Kumar, V. & Manchanda, R. A relay selection based mobility management in vanets: A review. *AIP Conf. Proc. (AIP Publ.)* **3100**(1).https://doi.org/10.1063/5.0211804 (2024).
5. Pereira, P. et al. From delay-tolerant networks to vehicular delay-tolerant networks. *IEEE Commun. Surv. & Tutorials (IEEE)* **14**(4), 1166–1182. https://doi.org/10.1109/SURV.2011.081611.00102 (2011).
6. Cooper, C., Franklin, D., Ros, M., Safaei, F. & Abolhasan, M. A comparative survey of vanet clustering techniques. *IEEE Commun. Surv. & Tutorials (IEEE)* **19**(1), 657–681. https://doi.org/10.1109/COMST.2016.2611524 (2016).
7. Ge, Y., Wen, S., Ang, Y. & Liang, Y. ptimal relay selection in ieee 802.16j multihop relay vehicular networks. *IEEE Transactions on Veh. Technol. (IEEE)* **59**(5), 2198–2206. https://doi.org/10.1109/TVT.2010.2044822 (2010).
8. Chai, R., Yang, B., Li, L., Sun, X. & Chen, Q. Clustering-based data transmission algorithms for vanet. *Int. Conf. on Wirel. Commun. Signal Process. (IEEE)* 1–6. https://doi.org/10.1109/WCSP.2013.6677258 (2013).
9. Wu, D., Zhu, G. & Zhao, D. Adaptive carry-store forward scheme in two-hop vehicular delay tolerant networks. *IEEE Commun. Lett. (IEEE)* **17**(4), 721–724. https://doi.org/10.1109/LCOMM.2013.022713.130123 (2013).
10. Feteiha, M., Hassanein, H. & Kubbar, O. Opportunistic cooperation for infrastructure-to-relaying-vehicles over lte-a networks. *IEEE Int. Conf. on Commun. (ICC) (IEEE)* 6376–6380. https://doi.org/10.1109/ICC.2013.6655630 (2013).

11. Ahmed, S., Kang, H. & Kim, D. Vehicular delay tolerant network (vdtn): Routing perspectives. *12th Annu. IEEE Consumer Commun. Netw. Conf. (CCNC) (IEEE)* 898–903. https://doi.org/10.1109/CCNC.2015.7158095 (2015).

12. Bouk, S., Ahmed, S., Omoniwa, B. & Kim, D. Outage minimization using bivious relaying scheme in vehicular delay tolerant networks. *Wirel. Pers. Commun. (Springer Nature)* **84**, 2679–2692. https://doi.org/10.1007/s11277-015-2760-0 (2015).

13. Liu, K. & Niu, K. A hybrid relay node selection strategy for vanet. *Int. Conf. on Commun. China (ICCC) (IEEE)* 1–6. https://doi.org/10.1109/ICCChina.2017.8330400 (2017).

14. Mu, D., Ahmed, S., Lee, S., Guizani, N. & Kim, D. An adaptive multiple-relay selection in vehicular delay tolerant networks. *GLOBECOM 2017 - IEEE Glob. Commun. Conf. (IEEE)* 1–6. https://doi.org/10.1109/GLOCOM.2017.8254733 (2017).

15. Ahmed, S., Mu, D. & Kim, D. Improving bivious relay selection in vehicular delay tolerant networks. *IEEE Transactions on Intell. Transp. Syst. (IEEE)* **19**(3), 987–995. https://doi.org/10.1109/TITS.2018.2791925 (2018).

16. Chai, R., Qin, Y., Peng, S. & Chen, Q. Transmission performance evaluation and optimal selection of relay vehicles in vanets. *IEEE Wirel. Commun. Netw. Conf. (WCNC) (IEEE)* 1–6. https://doi.org/10.1109/WCNC.2017.7925784 (2017).

17. Rossi, G., Fan, Z., Chin, W. & Leung, K. Stable clustering for ad-hoc vehicle networking. *IEEE Wirel. Commun. Netw. Conf. (WCNC) (IEEE)*. https://doi.org/10.1109/WCNC.2017.7925786 (2017).

18. Ji, X., Yu, H., Fan, G., Sun, H. & Chen, L. Efficient and reliable cluster-based data transmission for vehicular ad hoc networks. *Mob. Inf. Syst. (Wiley)* **2018**, 1–15. https://doi.org/10.1155/2018/9826782 (2018).

19. Kumar, V., Srivastava, V., Gupta, N. & Mäkelä, J. Adaptive priority and optimum parameters based relay vehicle selection scheme in vehicular ad hoc networks. *Comput. Electr. Eng.* **126**, https://doi.org/10.1016/j.compeleceng.2025.110521 (2025).

20. Souza, E., Nikolaidis, I. & Gburzynski, P. A new aggregate local mobility (alm) clustering algorithm for vanets. *IEEE Int. Conf. on Commun. (IEEE)* 1–5. https://doi.org/10.1109/ICC.2010.5501789 (2010).

21. Kolbig, K., Gradshteyn, I., Ryzhik, I., Jeffrey, A. & Scripta Technica, I. Review of table of integrals, series, and products. *Math. Comput. (American Math. Soc.)* **64**(209), 439–441. https://doi.org/10.2307/2153347 (1995).

22. Yoo, J., Choi, B. & Gerla, M. An opportunistic relay protocol for vehicular road-side access with fading channels. *IEEE Int. Conf. on Netw. Protoc. (ICNP) (IEEE)* 233–242. https://doi.org/10.1109/ICNP.2010.5762772 (2010).

23. Karagiannidis, G., Sagias, N. & Mathiopoulos, P. N*nakagami: A novel stochastic model for cascaded fading channels. *IEEE Transactions on Commun. (IEEE)* **55**(8), 1453–1458. https://doi.org/10.1109/TCOMM.2007.902497 (2007).

24. Kumar, V. & Dahiya, P. Clustering-based optimal relay vehicle selection scheme for vehicular adhoc networks (vanets). *Int. J. Mob. Comput. Multimed. Commun. (IJMCMC) (IGI Glob.)* **11**(4), 67–83. https://doi.org/10.4018/IJMCMC.2020100104 (2020).

25. Zhang, D., Zhang, M., Ding, F., Li, S. & Li, K. A stability-based clustering scheme for vehicular networks. *3rd IEEE Int. Conf. on Electron. Technol. (ICET) (IEEE)* 809–813. https://doi.org/10.1109/ICET49382.2020.9119698 (2020).

26. Hafeez, K., Zhao, L., Mark, J. & Shen, X. Distributed multichannel and mobility-aware cluster-based mac protocol for vehicular ad hoc networks. *IEEE Transactions on Veh. Technol. (IEEE)* **62**(8), 3886–3902. https://doi.org/10.1109/TVT.2013.2258361 (2013).

27. Pal, R., Gupta, N., Prakash, A. & Tripathi, R. Adaptive mobility and range based clustering dependent mac protocol for vehicular ad hoc networks. *Wirel. Pers. Commun. (Springer Nature)* **98**, 1155–1170. https://doi.org/10.1007/s11277-017-4913-9 (2018).

28. Bhowmik, A. et al. Advancing sustainability in edm: A brief review of eco-friendly dielectric fluids. *AIP Adv.* **15** (2025).

29. Koundal, S. et al. Molecular interaction study of l-ornithine with bovine serum albumin using spectroscopic and molecular docking methods. *Sci. Reports* **15**, 11997 (2025).

30. Katoch, G. et al. Impact of lanthanum doping on crystal structure and magnetic anisotropy of mn-zn soft nanoferrites. *Sci. Reports* **15**, 11663 (2025).

31. Wang, S. & Lin, Y. Passcar: A passive clustering aided routing protocol for vehicular ad hoc networks. *Comput. Commun. (Elsevier)* **36**(2), 170–179. https://doi.org/10.1016/j.comcom.2012.08.013 (2013).

32. Eiza, M. & Ni, Q. An evolving graph-based reliable routing scheme for vanets. *IEEE Transactions on Veh. Technol. (IEEE)* **62**(4), 1493–1504. https://doi.org/10.1109/TVT.2013.2244625 (2013).

33. Network simulator 2 (ns2): Discrete event network simulator. Available at http://nsnam.isi.edu/nsnam/index.php/MainPage (2009).

34. Rapelli, M., Casetti, C. & Gagliardi, G. Vehicular traffic simulation in the city of turin from raw data. *IEEE Transactions on Mob. Comput.* **21**, 4656–4666. https://doi.org/10.1109/TMC.2021.3075985 (2021).

35. Baza-Solares, N., Velasquez-Martínez, R., Torres-Bohórquez, C., Martínez-Estupiñán, Y. & Poliziani, C. Traffic simulation with open-source and commercial traffic microsimulators: A case study. *Commun. - Sci. Lett. Univ. Zilina* **24**, 49–62. https://www.ceeol.com/search/article-detail?id=1120002 (2022).

Author contributions

VK, RM, HP: conceptualization, data curation, writing – original draft. GJ, SS, RS, NK: Validation, Visualization, Writing – review and editing. RG, AB, BAZ: Project Administration, supervision, Software, Methodology, formal analysis. All authors have read and approved the final version of the manuscript.

Declarations

Conflict of interest

The authors declare that there are no conflicts of interest.

Disclosure

This study was performed as part of the employment of the authors.

Additional information

Correspondence and requests for materials should be addressed to B.A.Z.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.