Mesopotamian journal of Big Data
Vol. **(2025)**, 2025, **pp**. 241–260
DOI: https://doi.org/10.58496/MJBD/2025/016 , ISSN: 2958-6453
https://mesopotamian.press/journals/index.php/BigData

Research Article

# PSOA-CRL: A Hybrid Multi-Objective Routing Mechanism Using Particle Swarm Optimization and Actor-Critic Reinforcement Learning For VANETs

Mustafa Maad Hamdi[1], Baraa Saad Abdulhakeem[2], Ahmed Adil Nafea[3, *]

[1] *Department of Computer Science, College of Computer Science and IT, University of Anbar, Ramadi, Iraq*
[1] *Department of Information Technology, College of Computer Science and IT, University of Anbar, Ramadi, Iraq*
[1] *Department of Artificial Intelligence, College of Computer Science and IT, University of Anbar, Ramadi, Iraq*

**ARTICLE INFO**

**Article History**

Received 25 Jun 2025
Revised 17 Jul 2025
Accepted 25 Aug 2025
Published 23 Sep 2025

**Keywords**

VANETs

Routing Protocols

Intelligent Transportation Systems

Actor-Critic Reinforcement Learning

Packet Delivery Ratio

**ABSTRACT**

Vehicular ad hoc networks (VANETs) serve vehicles and infrastructure systems to communicate in real time for critical safety functions and traffic control. The highly mobile nature of VANETs with rapid topology changes, high mobility, and frequent disconnections is a very challenging situation for routing protocols. Most of the current approaches are static and tend to focus on a single metric rather than being flexible in practical environments. This paper introduces a hybrid routing method that is capable of maintaining a high packet delivery rate, low delay, and stable connectivity in VANETs with dynamic traffic situations. To address these problems, in this paper, we propose the PSOA-CRL, which is a hybrid multi-objective routing algorithm that integrates particle swarm optimization (PSO) with actor-critic reinforcement learning (A-CRL). The offline PSO component generates a variety of optimal routes. where the adaptive CRL just-in-time chooses the best available path. The two-way protocol maximizes the trade-off between the packet delivery ratio (PDR), end-to-end delay (E2E), link reliability, energy consumption, and routing overhead. A performance evaluation of PSOA-CRL with benchmarks under multi-objective optimization (MOO) through network metrics reveal the dominance of PSOA-CRL in most of the performance evaluation metrics. The obtained result reveals that the PSOA-CRL has a 97.8% packet delivery ratio, 41.3 ms end-to-end delay, and 96.1% link reliability. These results indicate that the PSOA-CRL is efficient in realizing reliable, real-time VANET routing and can be practically utilized in intelligent transportation systems (ITS).

## 1. INTRODUCTION

Vehicular ad hoc networks (VANETs) are of great value to the present intelligent transportation systems (ITS) for vehicular communication among vehicles (V2V), and vehicles to infrastructure nodes (V2I) [1]. This connectivity has been considered crucial for a number of safety and efficiency applications, including collision avoidance systems, dynamic traffic management, and emergency response coordination [2]. However, the nature of vehicular environments is that high node mobility, fast topology changes, and intermittent links occur frequently [3], which presents significant challenges to reliable and efficient data routing delivery [4]. The most crucial problem is how to model a routing algorithm to handle the multi-objective features of VANET scenarios [5], [6]. Quality of service (QoS) metrics, such as the packet delivery ratio (PDR), end-to-end delay (E2E), link reliability, energy consumption, and overhead, are significantly influenced by routing in VANETs [7].

Good performance in all of them is a constant challenge and an unexplored area for network fluctuations. Many current methods integrate multi-objectives into one weighted function, making single weighted function optimization easier; however, the final decision-making process is biased or sub-optimal all the time [8]. Deep reinforcement learning (DRL) has also been utilized in the optimization of VANET routing to better accommodate a dynamic network environment and improve the QoS features, including PDR, E2E delay, and energy efficiency [9], [10], [11]. In addition, most of the existing works are not scalable and do not adapt in real time, which significantly limits their application in VANETs in cases of large data and fast mobility. In addition, the stability of the links, which is essential for successful communication, is generally underreported by few and can be (wrongfully) ignored even though it leads to more packet loss and overhead in the routing

*Corresponding author. Email: ahmed.a.n@uoanbar.edu.iq

[12]. To address these problems, this paper proposed PSOA-CRL, which is a hybrid multi-objective routing algorithm based on particle swarm optimization (PSO) and actor-critic reinforcement learning (A-CRL).

While PSO is effective in complex solution space exploitation in routing optimization, ACRL is applicable for dynamic environment RTAM decision making [13]. The PSO by itself may not respond fast enough to real-time changes, whereas the RL may not converge rapidly, as it is randomly initialized. The hybrid PSO and actor-critic RL take advantage of the global optimization ability of PSO, as well as the adaptive learning capability of RL, so that the performance of the proposed model can be balanced and aimed at an efficient approach. The approach is capable of simultaneously optimizing multiple routing objectives (PDR, E2E delay, link reliability, energy, and routing overhead) while maintaining the Pareto-based approximation of the solution space. It takes trade-offs between opposing performance into consideration automatically and reacts in a timely manner to the variations in network topology induced by vehicle mobility and topology changes [14].

The design of the PSOA-CRL is organized in two complementary phases. The first phase is an offline phase that leverages PSO to produce a set of diversified initial strategies, while the second phase is an online phase that applies actor-critic RL to make fine-grained routing decisions in real time. The proposed approach architecture achieves stable and adaptive routing even when the VANET is highly volatile. The PSOA-CRL was evaluated through extensive simulations via utilizing the OMNeT++ network simulator combined with the SUMO traffic simulator, which supports realistic vehicular mobility and network conditions. The dual simulation environment also ensures the realistic VANET performance, enabling to demonstration that the proposed approach is applicable under different traffic congestion and mobility patterns. The objective of this work is to develop a routing scheme that can guarantee a high packet delivery ratio, low delay, and good routing stability and that can scale and adapt in real-time. The contributions of this paper are as follows:

- This work proposed a hybrid multi-objective routing algorithm (PSOA-CRL) for VANETs based on the integration of PSO with A-CRL.
- Design a decision-making process via utilizing a two-phase approach that decouples offline global exploration from real-time policy refinement to maximize routing performance in dynamic environments.
- Enhance flexibility and deliverability via exploiting RL feedback based on real-time link quality, mobility pattern, and network state.
- The efficiency of the PSOA-CRL is validated via extensive simulation experiments, and it can achieve better performance than conventional and metaheuristic-based routing protocols.
- The PSOA-CRL has achieved improvements in performance evaluation over the RBMOORPV technique concerning PDR, E2E delay, and responsiveness to network dynamics via integrating dual-phase optimization with real-time adaptability.

The significance of this paper lies in that it can provide the possibility of practicing and scalable implementation of VANET routing in ITS, requiring low-latency and high-reliability decisions. The novelty of the PSOA-CRL approach is due to its two-phase nature that applies an offline global optimization method along with an online adaptive decision mechanism. This is in contrast to previous modes of study on VANET routing, which have applied metaheuristic methods or reinforcement learning alone. By combining the two disciplines and integrating trust and QoS awareness into the decision-making procedure, PSOACRL gives a more comprehensive and realistic routing solution that considers reliability, scalability, and adaptability to highly dynamic vehicular environments.

The rest of the paper is organized as follows: In Section 2, the related work is presented in detail. Next, the PSOA-CRL model is provided in Section 3. The experimental results and evaluation are explained in Section 4. The limitations are introduced in Section 5. Finally, the conclusion is provided in Section 6.

## 2. RELATED WORK

This section presents related work on the problem of routing as multi-objective optimization (MOO) in VANETs. This section is divided into two sub-sections. VANET routing involves trade-offs among PDR, E2E, link reliability, energy consumption, routing overhead, and connectivity in highly dynamic environments. MOO methods are often applied to address these problems. In [15], the NSGA-II is an algorithm that proposed fast elitist non-dominated sorting, a standard approach for multi-objective problems, due to its strong convergence and diversity-preserving properties. However, NSGA-II is designed for static vehicular environments and is not suitable for dynamic vehicular environments to be applied to VANET routing with high mobility. It is effective for static multi-objective optimization problems. In [16], Gaussian mutation according to the harmony search algorithm was proposed to enhance the solution diversity and convergence rate in the MOO, but it did not consider the environmental adaptation towards the vehicular domain. Although the scheme accelerates the statistical convergence and increases the solution variety, it has the problem that it is not geared toward a real VANET environment; thus, its practical performance in terms of metrics such as PDR and E2E delay would be degraded under dynamic traffic. Another work in [17] proposed several designs for communication scenarios, but these methods could not be easily scaled to real-time routing in VANETs. As they may be effective in small-scale or simulated cases, such techniques do not perform well under high node mobility or dense traffic and suffer from both delay and reliability. In [18], the PSO-NSGA-II global hybrid algorithm for intelligent parking applications was proposed, which makes decisions about delay and energy efficiency in an inherently efficient way; however, this approach is a domain-specific solution, so its generality over different contexts of research is limited. The hybrid method can achieve the optimal trade-off between delay and energy efficiently in intelligent parking, but it has not been verified for distribution VANET routing, and its scalability and flexibility in dynamic traffic are still limited. A subsequent framework introduced adaptive mechanisms towards optimizing both the delivery ratio and overhead in a multi-objective manner, but it was found to introduce scalability issues under dense traffic [19]. Although it achieves better performance for PDR and overhead under moderate conditions, the proposed approach has difficulties in terms of large-scale networks and densification of vehicular traffic, limiting its practical deployment.

Another work in [20] aimed to optimize and disseminate roadside unit (RSU) placement by means of an evolutionary algorithm to ensure the maximization of network coverage and the minimization of infrastructure deployment cost, but such an approach does not incorporate dynamic conditioning based on routing behavior, which is still missing. It is extended to reduce the deployment cost, and the network coverage is increased; however, the absence of route-aware dynamic adaptation makes it less effective in highly mobile VANET scenarios where real-time consideration of QoS metrics is affected. To improve real-time performance, a digital twin-based routing model for real-time synchronization and predictive optimization was proposed, with scalability issues due to the computational cost [21]. The digital twin method improves predictive routing, but it is highly computational and cannot be used in real time for the entire VANET. This work in [22] suggested that metaheuristic methods, such as Harris hawk optimization, be employed in cognitive radio VANETs for route selection and that load balancing be investigated, which requires many parameters to be tuned to reach optimal results. As an excellent route selection algorithm and load balancing algorithm, its performance is strongly affected by its parameters, which makes it lack practical value in the dynamic VANET environment. In [23], reliability-focused routing strategies were also explored via an enhanced Gaussian mutation harmony search, which improved the performance in high-mobility settings; however, these studies were not verified in large-scale environments. This work proposed a scheme that enhances the robustness in high mobility, without being evaluated in large-scale VANET networks, with a loss of confidence in the scalability and adaptability. In the work of [24], clustering-based multi-objective techniques were also presented to ensure security and reliability in VANETs to detect and prevent malicious activities. Although such models enhanced the attack detection, the extra communication overhead degraded the performance in a delay-sensitive environment. The method improves security and availability; however, it adds additional communication overhead, which weakens delay-sensitive applications and overall network performance. The work of [25] suggested a multi-objective routing approach using the Pelican optimization algorithm to achieve maximum route reliability for dynamic traffic conditions, but the associated algorithmic complexity does not guarantee responsiveness in real-life scenarios. However, most of the existing methods cannot solve these problems very well when considering scalability, real-time adaptability, security integration, and computational overhead. While the RBMOORP can enhance the reliability of the route, its high computational complexity and low adaptability make the RBMOORP less effective for large-scale, real-time VANETs.

These research gaps warrant further study of lightweight, adaptive, and trustworthy multi-objective routing algorithms designed for large-scale VANETs. MOO methods have long been essential for handling conflicting objectives in VANET routing. In [26], a new approach called MOTD-DE, which blends Kubernetes-based clusters with differential evolution to handle tasks smarter in VANETs, was proposed. It uses deep learning to judge how hard each task is, cutting both run-time and resource drain. Tests indicate that it outperforms classical techniques such as PSO, the GA, ACO, and ABC in terms of speed and efficiency. The approach reveals good computational efficiency and task planning optimization; however, its applicability in the context of real-time vehicular routing with high mobility has not been thoroughly verified. The DRLbased work in [27] first presented MOEA-DRL, but it has been attempted on VRPTW, and the multi-objective counterpart, MOVRPTW, is still challenging. The MTMO/DRL-AT method combines DRL with a multi-task evolutionary algorithm to reduce travel distance, decrease vehicle number, and shorten delays simultaneously. When validated against the real world, it is demonstrated to be better than previous methods on standard test sets. Despite the promising results, the application of MOEA-DRL to VRPTW is limited, and the MOVRPTW, when deployed in VANETs, is still a challenge. MTMO/DRL-AT demonstrates multi-objective enhancements and calls for real-time verification in dynamic vehicular networks. During the past decade, various recent and future generational studies have examined hierarchical and Q-learning-based VANET routing to address dynamic vehicle motion mobility and urban environments [28]. By combining roadside units (RSUs) with V2X communications and exchanging Q-vectors, distributed multi-agent reinforcement learning at intersections is applied to the cross-layer optimization of PDR, reduction in broadcasting overhead, and acceleration of learning toward efficient data routing. This hierarchical RL strategy enhances cross-layer optimization as well as learning efficiency; however, it may suffer from extremely dense traffic or dynamic topologies.

DRL-based routing strategies have recently been suggested for VANETs to increase efficiency in dynamic network environments [29]. By using the deep Q-network approach for the configuration of next-hop intersections, it integrates intersection forwarding and traffic awareness to improve PDR, decrease E2E, and reduce overhead. This work demonstrates superiority to existing methods with simulations based on real-world taxi trajectory data. The rate-adaptive scheme is proven to achieve PDR and delay gains significantly under both scenarios, whereas its performance for larger and diverse VANETs **has** not **been** fully verified. Finally, in [30], it was proposed that OptiE2ERL is a reinforcement learning model that maximizes energy-efficient routing by considering energy levels, bandwidth, mobility, and traffic. The simulation results reveal that OptiE2ERL performs better than models such as LEACH, PEGASIS, and EER-RL in terms of network lifetime, energy consumption, and scalability. Dynamic real-time road-level energy consumption adaptation, as achieved by OptiE2ERL, is another key contribution **to** energy-efficient routing for VANETs.

A comparative summary of the popular multi-objective routing approaches in VANETs is shown in Table 1. The main benefits, drawbacks, and performance characteristics of each approach are presented to illustrate its efficiency and limitations in dynamic vehicle scenarios. This study offers a clear reflection on current methodologies and the gaps they have compared with (PSOA-CRL).

TABLE I. SUMMARY OF EXISTING MULTI-OBJECTIVE ROUTING STUDIES

| References | Technique | Advantages | Limitations | Key Metrics |
|---|---|---|---|---|
| **[15]** | NSGA-II | Strong convergence, preserves diversity | Not suited for dynamic VANETs | General MOO |
| **[16]** | Gaussian Mutation with Harmony Search | Enhances diversity and convergence | Lacks real-time VANET adaptation | Convergence, diversity |
| **[17]** | Multi-objective Harmony Search | Effective in small-scale scenarios | Poor scalability; affected by high mobility | PDR, delay |
| **[18]** | PSO-NSGA-II Hybrid | Efficient delay & energy optimization | Domain specific; limited generality | Delay, energy |
| **[19]** | Adaptive MOO Framework | Optimizes PDR & overhead adaptively | Scalability issues in dense traffic | PDR, overhead |
| **[20]** | RSU Deployment Evolutionary | Maximizes coverage; reduces cost | No routing-aware adaptation | Coverage, deployment cost |
| **[21]** | Digital Twin Routing | Predictive routing & sync | High computational cost | Delay, predictive performance |
| **[22]** | Harris Hawks Optimization | Route selection & load balancing | Needs parameter tuning; limited practicality | PDR, load balancing |
| **[23]** | Reliability-Focused Harmony Search | Improved reliability in mobility | Not validated large-scale | Reliability, mobility |
| **[24]** | Clustering-Based Secure Routing | Enhances security & attack detection | Extra overhead; delays sensitive traffic | Security, delay, overhead |
| **[25]** | Pelican Optimization (RBMOORP) | Improves route reliability | High complexity; limited real-time adaptability | Route reliability |
| **[26]** | MOTD-DE (Diff. Evolution + DL) | Efficient computation & task allocation | Not fully validated in real-time VANETs | Task allocation, efficiency |
| **[27]** | MOEA-DRL/MTMO-DRL-AT | Multi-objective optimization; reduces travel & delays | Needs real-time validation in VANETs | Distance, vehicle number, delay |
| **[28]** | Hierarchical RL with RSU | Better cross-layer optimization & learning | Challenges in dense traffic | PDR, broadcast reduction |
| **[29]** | Deep Q-Network Intersection Routing | Improves PDR, delay & overhead | Scalability in large/heterogeneous VANETs | PDR, delay, overhead |
| **[30]** | OptiE2ERL | Enhances energy efficiency & lifetime | Needs further real-time evaluation | Energy, lifetime, mobility |

## 3. METHODOLOGY

This section explains the proposed hybrid algorithms (PSOA-CRL) based on hybrid multi-objective routing in VANETs.

## 3.1 PSO Hybrid Framework

This section describes a three-layer framework that combines PSO with A-CRL to enhance VANET routing, as shown in Figure 1. Otherwise, the typical solutions that integrate performance measures into one objective, this method remains the PSO as a Pareto-solution set by pursuing five objectives (PDR, E2E delay, link reliability, energy, and routing overhead) independently. The combinations of PSO and ACRL complement each other. PSO is very quick for global exploration in the high solution space dimension but does not respond quickly to real-time network changes. ACRL, on the other hand, is good at making adaptive decisions by using continuously updated environmental information, but it has a relatively slower convergence rate under random initial conditions. By combining them, the proposed PSOA-CRL finds a middle way: global optimization employing PSO and responsiveness in real time using ACRL, which makes it especially interesting for highly dynamic and large-scale VANET environments. The first layer provides a simulation of the realistic VANET environment with vehicle mobility and wireless link dynamics. The second layer performs off-line multi-objective PSO to seek an effective routing strategy, considering further factors such as energy, trust, and QoS priority, thus enabling context-aware behavior. The third layer uses ACRL to change routing in real time. The actor chooses next-hop nodes according to the live network state, and the critic estimates the results to improve the policy. A monitoring mechanism ensures the robustness of re-optimizing if performance deteriorates and if the system remains responsive in dynamic environments.

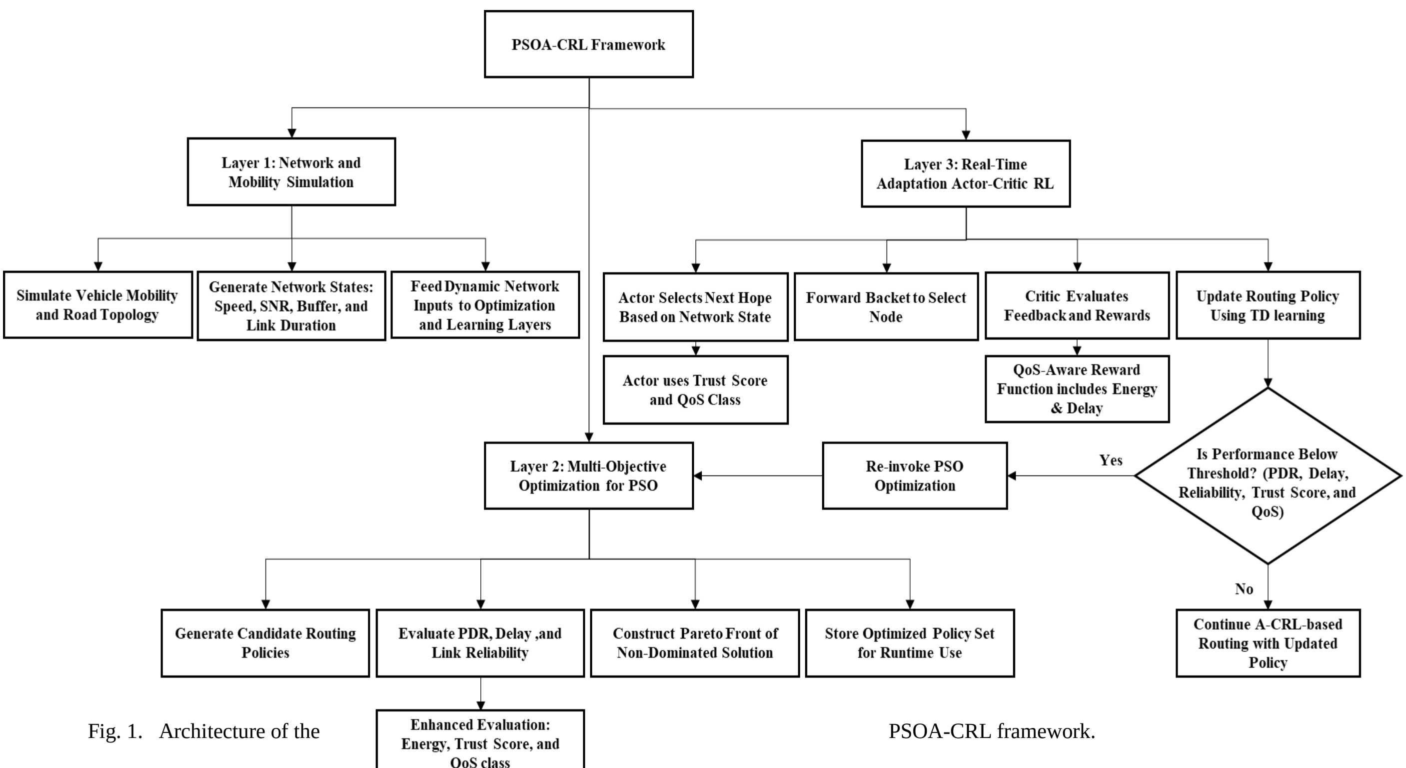



Fig. 1. Architecture of the PSOA-CRL framework.

## 3.2 PSOA-CRL Algorithms Flowchart

Figure 2 illustrates the proposed PSOA-CRL framework as a flowchart and indicates the interaction between PSO and A-CRL to achieve adaptive and intelligent routing in highly dynamic VANET environments. There are two main phases of the process. Initially, the system leverages PSO to construct a wide range of candidate routing strategies, promoting diversity and adaptability to VANET constraints. There are different parameters, such as transmission range, SNR threshold, buffer occupancy limit, and link reliability weight, are associated with the policy. This set of candidate policies is overlaid with some real-world constraints to ensure that the policies are feasible and robust. Specifically, the problem settings include the minimum SNR requirements, maximum transmission range, buffer size limitations, and minimum link duration constraints. Each candidate policy is evaluated against five key objectives, PDR, E2E delay, link stability, energy usage, and routing overhead, to identify optimal trade-offs. Rather than combining these objectives into a single metric, we resort to a Pareto-based multi-objective selection approach to obtain non-dominated policies with Pareto optimality concerning these objectives. The learned policies serve as the initial knowledge for the following reinforcement learning. In the subsequent phase, A-CRL dynamically selects the optimal next hop on the basis of real-time network states. When there is more than one next-hop node, the system selects the node with the highest priority. Before forwarding, a trust-based filtering mechanism assesses each node on the basis of a trust score and checks whether the node is listed in the Certificate Revocation List (CRL). A packet is dropped or rerouted via another trusted path if a node is non-trusted. If the trust check is successful for the node, the actor selects it for forwarding. The critic assesses the state action pair by looking at the network state that the action has caused and computes a QoS-aware reward via performance metrics that are weighted by the class priority of that packet, namely, high, medium or low. This reward is used to train the actor and critic modules via temporal difference (TD) learning. The performance is tracked in real time to maintain adaptiveness. When some metrics of PDR, E2E delay, reliability, energy, and overhead are lower than the predefined threshold, the PSO optimization is re-executed to refresh the routing policies concerning the current network state. This feedback mechanism guarantees that the PSOA-CRL framework remains responsive, efficient, and robust under the challenging and dynamic conditions frequently encountered in the VANET scenario.

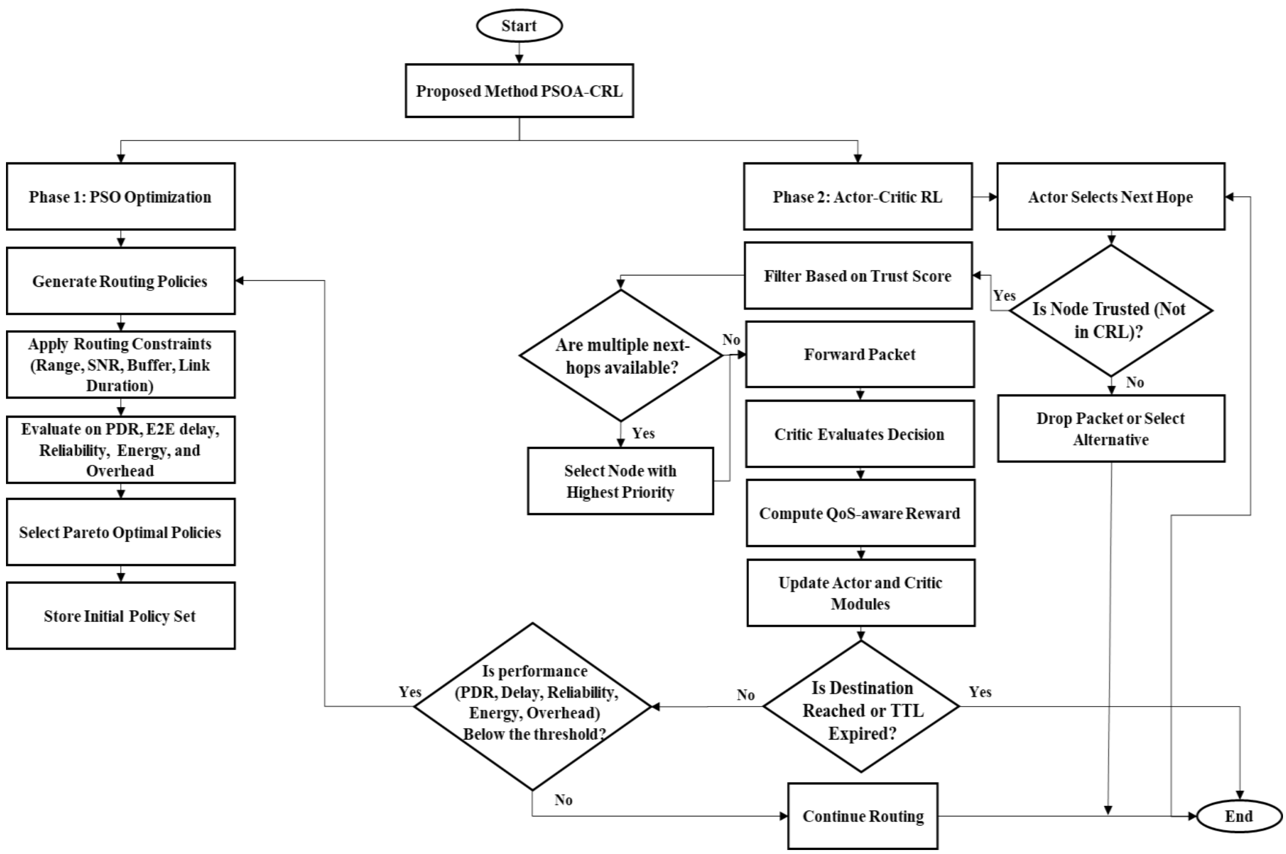


Fig. 2. Flowchart of the proposed PSOA-CRL.

### 3.3 Offline Optimization Using PSO

Algorithm 1 executes offline optimization for routing policies within VANETs via a selection process that is based on Pareto's method. The algorithm begins by creating a specific number of random routing policies. Each policy is initialized with parameters within the prescribed bounds and sampled uniformly. Candidate routing strategies are tested within a simulated VANET scenario and scored across five essential performance dimensions: packet delivery, delay, reliability, energy efficiency, and protocol overhead. The results gathered from previous computations are stored and analysed in pairs to extract non-dominated policies, those that represent the optimal trade-offs among the selected objectives. The policies selected for the Pareto front represent non-dominated solutions that achieve balanced compromises across all the objectives and serve as the final optimized set. The identification of robust and high-quality routing strategies is achieved through this offline process before real-time implementation. To evaluate each routing policy candidate during the offline PSO phase, a multi-objective fitness function is used. This function quantifies the performance of a given policy in five critical routing metrics. It is formally (1) defined as:

$$f(x) = [PDR(X), -Delay(x), Reliability(x), -Energy(x), -Overhead(x)] \quad (1)$$

where:

$x$: candidate routing policy.

$f(x)$: - objective vector for policy x.

$PDR(X)$:- To maximize the PDR.

$Delay(x)$: - To minimize E2E.

$Reliability(x)$: -To maximize link stability.

$Energy(x)$: -To minimize energy consumption.

$Overhead(x)$: - To minimize routing overhead.

This vector-based formulation of the objective here makes it possible to utilize Pareto-based MOO, in which PSO seeks out and selects multiple non-dominated policies across the entire performance trade-off frontier.

**Algorithm 1: PSO offline optimization for VANETs**

**Start**
**Input:**
  $N\_particles$   # Number of policies to generate
  $param\_bounds$   # Parameter ranges for each policy variable
  $VANET\_Model$   # Simulation environment
**Output:** $Pareto\_Set$   # Set of Pareto-optimal routing policies
**Process:**
1 .Initialize an empty Swarm list
2 .For $i = 1$ $to$ $N\_particles$:
- Randomly generate a policy with parameters within$param_bounds$
- Add policy to Swarm
3 .Initialize empty$Fitness_List$
4 .For each policy in Swarm:
- Simulate $the\ VANET\ model$ via policy$x$
- Measure performance: $f(x)$   //calculate by using Eq. (1)
- Store results in$Fitness_List$
5 .Initialize empty$Pareto\ Set$
6 .For each $policy\ i$ in Swarm:
-$dominated = False$
- For each $policy_j$ in Swarm where $j\ != i$:
-   If$FitnessList[j]\ dominates\ FitnessList[i]$:
-     $dominated = True$
-     Break inner loop;
- If not dominated:
-   Add$policy\ i\ to\ Pareto_Set$
7. Return$Pareto\ Set$
End

## 3.4 Online Decision-Making via the Actor-Critic Model

The algorithm 2 presents the online decision-making phase of the A-CRL model in The PSOA-CRL for VANET routing is designed. It has a policy starting from offline PSO that is utilized as a baseline for routing. The algorithm continues monitoring the state of its current network, such as the vehicle speed, signal-to-noise ratio (SNR), buffer occupancy, and link duration. The actor module chooses the next-hop node to forward packets on the basis of this state. This is then rewarded by the system after it forwards, where it rates the success or quality of that action. The temporal difference (TD) error is then calculated to update the value function and the policy via (2).

$$\delta = V(s) - V(s') \cdot \gamma + r \tag{2}$$

where:

$\delta$: TD error is used for learning.

$r$: An immediate reward after the action.

$\gamma$: Discount factor for future value ($0 < \gamma \leq 1$).

$V(s)$: Value of the current state.

$V(s')$: Value of the next state after the action.

The critic updates its value while the actor tunes the policy with this TD error, so routing steadily sharpens. The loop continues to run during a VANET session, enabling ongoing learning and adaptation.

**Algorithm 2: ACRL Online Decision-Making**

**Start**
**Input:**
  $policybase$ = initial routing policies from PSO
  $currentstate$ = real-time network state (speed, SNR, buffer, link duration)
**Output:** $updated\ policy$ = continuously improved routing policy
**Process:**
Initialize the actor and critic modules with$policy_base$
While the VANET session is active, do
  Observe $currentstate$ = [$vehiclespeed$, SNR, $buffer\_ccupancy$, $link\_duration$]
  action = $Actor.select\_action$ ($current\_state$)
  result = $forward\ packet$ (action)
  reward = $calculatereward$ (result)
  $nextstate = get\ new\ networkstate()$ //Calculate the temporal difference (TD) error δ via Eq. (2)
  $Critic.update$ ($\delta$) //Update value function using$\delta$
  $Actor.update$ ($currentstate$, action, $\delta$) / /Update policy using$\delta$
    $currentstate$ =$nextstate$
End While
Return$updated\ policy$
**End**

## 3.5 Integration of the PSOA-CRL Algorithms

The initial stage is offline optimization for exploring a variety and creating a particle swarm that represents potential routing policies and is evaluated according to multiple metrics (PDR, E2E, link reliability, energy and routing overhead). On the basis of such evaluations, the algorithm manages to approximate a Pareto-optimal collection of policies that compromise between multiple and conflicting objectives. These trained policies are further encoded into a format that is compatible with reinforcement learning and helps initialize the actor module of A-CRL. This initialization allows the actor to have sensible policy preferences rather than random weights. At the same time, the critic is bootstrapped with value estimates obtained from the fitness of the Pareto-optimal policies. The second phase is the online process that enters the ongoing training loop, where it watches the current network state, takes action through the actor, and obtains rewards. The Critic's value function is updated based on the TD error given in Eq. (2). This TD error helps the Actor to update its policy. The PSO process can be reinvoked periodically to regenerate and encode new policies, which are utilized to update the actor and critic from existing policies, ensuring adaptation. Finally, this fusion allows the system to exploit the advantages of global off-line optimization as well as real-time learning for a more stable and reactive VANET routing strategy.

**Algorithm 3: Integrate PSO with A-CRL Algorithms.**

**Start**
**Input:**
  $Network\ Environment$ //Network simulation environment (e.g., VANET scenario)
  $PSOParams$ //Parameters for PSO
  $RL\_Params$ / /Parameters for Actor-Critic RL
  $Evaluation\ Metrics$ //Metrics for assessing routing quality (e.g., delay, PDR)

**Output:** $Trained\ ACRL\ Model$ //The final Actor-critic model with optimized routing policy
**Process:**
1. **Phase 1: Offline Policy Optimization Using PSO**
Initialize Swarm
For each particle p in Swarm **do**
Simulate p in$NetworkEnvironment$
Evaluate the fitness of p using$EvaluationMetrics$
End For
$ParetoFront$ = Extract Pareto-optimal policies from Swarm
$EncodedPolicies = Encode\ Policies\ For\ RL\ (ParetoFront)$
2. **Phase 2: RL Model Initialization Using PSO Results**
$ActorWeights = Initialize\ Actor\ With\ Encoded\ Policies\ (EncodedPolicies)$
$CriticValues = Compute\ Critic\ Values(ParetoFront)$
$CriticWeights = Initialize\ Critic$(CriticValues)
$ACRL\_Model = Create\ Actor\ Critic\ Model\ (ActorWeights, CriticWeights)$
//**Phase 3: Online Policy Learning and Adaptation**
For episode = 1 to $MaxEpisodes$ **do**
state = $Observe\ Initial\ State$ ($NetworkEnvironment$)
While the state is not terminal do
action = $Actor.select\ action$(state)
$nextstate, reward = Execute\ Action\ (state, action, NetworkEnvironment)$
TD error calculation by using eq (2)

$Critic.update$ (State, reward, $next\_state$)

advantage =$Critic.advantage(state, action)$
$Actor.update(state, action, advantage)$
state =$nextstate$
End While
If episode mod $UpdateCycle$ = 0 then
$NewPolicies\ = Run\ PSO\ (NetworkEnvironment, PSO\_Params)$
Encoded =$Encode\ Policies\ For\ RL(NewPolicies)$
$Update\ Actor$ (Actor, Encoded)
$NewValues$ =$Compute\ Critic\ Values(NewPolicies)$
$Update\ Critic(Critic, NewValues)$
End If
End For
Return$ACRL_Model$
End

### 3.6 Enhanced PSOA-CRL Methodology

This section provides an improved PSOA-CRL methodology for VANETs to augment their routing in an integrative VANET term deployment and improve it in terms of energy efficiency, trust, QoS awareness, scalability, and world-based deployment. The enhanced model has a similar two-phase architecture, with PSO as the offline optimizer, whereas A-CRL is an online learning platform, but it is more capable of resisting network failure and adaptation through the complete routing function. The main design criteria include maximizing PDR, minimizing E2E, improving link reliability, minimizing energy, minimizing routing overhead, and new performance metrics for more realistic conditions.

#### 3.6.1 QoS-Awareness integration

To handle different QoS requirements ranging from emergency messages to information technology, in VANETs, the routing mechanism adopts a service classification strategy. A QoS class is associated with each data packet and affects its routing restrictions. High-priority traffic, such as emergency messaging, is delivered over paths with low delay and high reliability, while for low-priority services, energy efficiency becomes more important. The actor module uses the QoS class as an input for adaptive and service-aware routing decision-making. Algorithm 4 develops such logic by interpreting the packet's QoS level, fulfilling the routing demand on the basis of the QoS level, and then scoring all potential paths to select the best path; see (3).

$$\text{Fitness}(p) = \{f_1(p), f_2(p), f_3(p), f_4(p), f_5(p)\} \tag{3}$$

where:

$f_1(p)$: $PDR(maximize)$

$f_2(p)$: Delay $(minimize)$

$f_3(p)$: Reliability $(maximize)$

$f_4(p)$: Energy $(minimize)$

$f_5(p)$: Overhead $(minimize)$

$p$ : A routing-in policy.

$f_1 to\ f_5$ : Individual objective functions.

**Algorithm 4: QoS-Aware Routing Decision**

**Start**
**Input:** $Packet, Available\ Paths, NetworkState$
**Output:** $Selected\ Path$
**Process:**
$QoSClass$ =$Packet.QoS\_Class$
If $QoS\ Class$ = "$Emergency$", then
Requirements = $\{Low\_Delay, High\_Reliability\}$
If $QoSClass$ = "$RealTime$", then
Requirements = $\{Moderate\_Delay, Stable\_Link\}$
If $QoSClass$ = "$Infotainment$", then
Requirements = $\{Energy\_Efficiency\}$
Else
Requirements = $\{Best\_Effort\}$
End If

3.$Best\ Score\ = \infty$
$Selected\ Path$ = NULL

4 .For each path in $Available\_Paths$ do
Score = 0
If "$Low\ Delay$" in Requirements then
$Score\ = Score\ + path.Delay$
End If
If "$High\ Reliability$" in Requirements then
$Score\ = Score\ - path.Reliability$
End If
If "$Energy\ Efficiency$" in Requirements then
$Score\ = Score\ + path.Energy\ Consumption$
End If
If "$Stable\ Link$" in Requirements then
$Score\ = Score\ - path.Link\ Stability$
End If

If $Score\ < Best\_Score$ then
$Best\ Score$ = Score
$Selected\ Path$ = path
End If

End For
5 .Return$Selected\ Path$
**End**

### 3.6.2 Energy-Aware Fitness Evaluation:

The energy consumption metric is considered an important criterion that needs to be optimized to make the PSOA-CRL framework more effective. The overall energy expenditure to form a routing path is obtained through the sum of the transmission energy for every hop, comprising the base transmission energy and the amplifier energy multiplied by the squared distance raised to the path loss exponent. The procedure is described in Algorithm 5. On the basis of this metric, the framework favours energy-aware paths to maximize network lifetime and evenly distribute energy consumption among nodes. The total energy is calculated between a routing path via (4):

$$E_{total} \sum_{i=1}^{n} \left( E_{elec\ i^{n}} .\ k + \varepsilon_{\text{amp}} .\ k . d \right) \quad (4)$$

where:

$E_{\text{total}}$ : Total energy consumed for a full route.

$E_{\text{elec}}$ : Electronic circuitry energy per bit.

$\varepsilon_{\text{amp}}$ : Transmit amplifier energy per bit per m2.

$k$: Packet size in bits.

$d$: Distance of hop $i$.

$n$: Path loss exponent (2.5).

**Algorithm 5: Energy-aware fitness**

**Start**
**Input:**
$Path = list\ of\ nodes\ in\ routing\ policy$
$E_{tx} = base\ transmission\ energy$
$E_{amp} = amplifier\ energy\ constant$
$n = path\ loss\ exponent$
**Output:** $E_{total} = total\ energy\ consumption$
**Process:**
1.$E_{total} = 0$
2 .For all consecutive hop $(i, j)$ in Path do
$d_{i,j} = Distance(i, j)$ //calculate via eq (4)
End For
3 .Return$E_total$
**End**

### 3.6.3 QoS-Aware Reward Computation in (A-CRL)

This part presents a revised reward model for the actor-critic framework, which integrates quality-of-service (QoS)-based traffic priorities in addition to some fundamental routing performance metrics so that the agent can learn rewards that are reflective of actual traffic requirements. It begins by giving each packet a weight, determined via its class of service (low, middle and high), with high-priority data receiving the most weight. The final reward is defined as a weighted sum that balances the packet delivery rate and link reliability with maximum delay and energy consumption, with weighing factors determining the impact of the different factors. This mechanism helps the learning model select valuable traffic with a trade-off between efficiency and reliability. The QoS-aware reward is calculated via (5):

$$R = \text{Energy} \cdot \delta - \text{Reliability} \cdot \gamma + \text{Delay} \cdot \beta - \text{PDR} \cdot \alpha \quad (5)$$

where:

$R$: Total reward value assigned by the agent.

$\delta, \gamma, \beta, \alpha$: Scaling factors depending on the QoS class (high, medium, or low).

The weighting factors in Eq. (5) choose the QoS metrics that represent high and medium importance in vehicular networks. Since packet delivery and link reliability directly represent the success of communication, more emphasis is placed on them, whereas delay and energy are considered secondary penalties to prevent excessive latency or resource waste. This prioritization was chosen a priori, on the basis of qualitative results obtained in preliminary trials and is consistent with the general practice in actor–critic models of multiple objective optimization within the reward function to ensure diverse outputs but without a single component dominating [1].

**Algorithm 6: QoS-Aware Reward Computation**

**Start**
**Input:**
$QoSClass$ //High, Medium, or Low
$PDR$ //Packet delivery ratio
$Delay$ //End-to-end delay
$Reliability$ //Link reliability
$Energy$ //Energy consumption
$alpha, beta, gamma, delta$ //Scaling coefficients
**Output:** $Reward$ //Computed reward value
**Process**:
If $QoSClass$ is high:
$weight$ = 1.0
Otherwise, if $QoSClass$ is medium:
$weight$ = 0.7
Else:
$weight$ = 0.4
Calculate $Reward$ via eq (5):
Return Reward
**End**

### 3.6.4 Trust-Aware Decision Making

Algorithm 7 adds trust checking to the same reinforcement-learning routing engine for vehicular networks. It begins by booting both the actor and the critic with policies pruned by the PSO to boot start learning. Trust scores for each node surface from its past forwards, current battery state, and the QoS targets of messages. Nodes rated below a set threshold drop out of the next-hop pool, shrinking choices but increasing reliability. The actor then picks the best path among the trusted contenders. Continuous updates to the actor and critic ensure adaptability to network changes; promote routing through reliable, energy-efficient, and QoS-compliant paths; and improve overall network stability and performance. The trust score is calculated by combining its forwarding behavior and remaining energy to evaluate the trustworthiness of a neighboring node, as in (6).

$$T_i = \left(\frac{E_i}{E_{max}}\right).W_2 + \frac{F_i}{S_i}.W_1 \qquad (6)$$

where:

$T_i$: Trust score for node $i$.

$F_i$: Packets successfully forwarded by node$i$.

$S_{i:}$ Packets received by node$i$.

$E_i$: Residual energy of node$i$.

$E_{max}$: Initial (maximum) energy of nodes.

$W_1, W_{2:}$ Weights for behavior and energy impact.

**Algorithm 7: Trust score computation and filtering**

**Start**
**Input:**
$Neighbors$ //list of neighbor nodes
$PacketsSent$ //packets sent to each neighbor
$PacketsForwarded$ //packets forwarded by each neighbor
$TrustThreshold$ //minimum acceptable trust score
**Output:**
$TrustedNodes$ //list of neighbors with $trust\ >=\ threshold$
**Process:**
Initialize $TrustedNodes$ as empty list
: For each node in Neighbors
by using eq (6) $Calculate\ TrustScore$
If $TrustScore\ >=\ TrustThreshold$
Add a node to $TrustedNodes$
Return $TrustedNodes$
**End**

### 3.6.5 Scalability considerations

PSOA-CRL was explicitly designed to work for large-scale and dense VANETs in a scalable manner. Unlike the pseudorandom key generation stage, there is no continuous execution of PSO re-optimization, as its events are driven only by predefined quality measures when vital performance indicators such as routing overhead or PDR fall below certain thresholds. This guarantees that expensive re-optimizations are called rarely and that the cost is spread out in time. The actor–critic module, on the other hand, only executes its local neighborhoods where each forwarding decision is based on the neighbors that have exchanged their trust information and are within the communication range. Thus, the computational complexity is ranked with the local-density rather than the total number of vehicles. Our design allows an upper bound on the latency and cures complexity into constant factors even in high VANET dense situations to retain real-time responsiveness. Most importantly, this scaling behavior is consistent with current benchmarks for VANET routing protocols in general, hence can indeed validate that PSOA-CRL remains effective and stable when networks are scaled up.

## 3.7 Benchmarking

This paper shows benchmark results of the proposed hybrid PSOA-CRL routing with four well-known MOO algorithms for a dynamic VANET scenario. MOHS generalizes Harmony search using a Pareto-based selection; however, it suffers from the difficulty of preserving population diversity. GMHS improves MOHS by adding a Gaussian mutation for finer search adjustments; however, it may still converge prematurely without diversity control. NSGA-II is a well-known evolutionary algorithm that effectively balances convergence and diversity but requires large populations and longer computation times, limiting its real-time applicability. RBMOORPV employs Pelican optimization, which focuses on reliable links and achieves stable routing offline but lacks online learning, reducing responsiveness to fast network changes. In general, PSOA-CRL achieves superior performance by generating well-distributed, high-quality Pareto-optimal routes with dynamic adaptability suited for VANETs.

## 3.8 Simulation parameters

The PSOA-CRL protocol was evaluated in an urban VANET via a simulation built with OMNeT++ and SUMO over a 5×5 grid for 600 seconds. A total of 200 vehicles and a maximum speed of 50 km/h used IEEE 802.11p (300 m range, 20 dBm), with CBR traffic generating 10 packets/sec (512 bytes). The routing was updated every 2 seconds. The key metrics included PDR, delay, reliability, energy, and overhead, with a 0.25 overhead threshold triggering re-optimization. Trust filtering of ≥0.6 and 1000 units of initial energy ensures secure, efficient routing. The reinforcement-learning agent learned across 300 episodes, whereas the particle swarm optimizer handled 40 particles for 50 iterations, enabling adaptive and high-quality routing under dynamic conditions, as presented in Table 2.

TABLE II. SIMULATION PARAMETERS AND CONFIGURATION FOR THE VANET.

| Parameter | Value |
|---|---|
| **Simulation Tool** | OMNeT++ **version 5.6.2** with SUMO (Veins) **version 1.14.1** |
| **Simulation Time** | 600 seconds |
| **Map Type** | Urban grid (5x5) |
| **Number of Vehicles** | 200 |
| **Mobility Model** | SUMO (Car Following Model) |
| **Max Speed** | 50 km/h |
| **Communication Range** | 300 meters |
| **MAC Protocol** | IEEE 802.11p |
| **Transmission Power** | 20 dBm |
| **Packet Size** | 512 Bytes |
| **Traffic Type** | CBR (Constant Bit Rate) |
| **Packet Rate** | 10 packets/sec |
| **Routing Update Interval** | 2 seconds |
| **Evaluation Metrics** | PDR, E2E Delay, Link Reliability, Energy, and Overhead |
| **Trust Threshold** | 0.6 |
| **Energy Initial per Node** | 1000 units |
| **Path Loss Exponent (n)** | 2.5 |
| **RL Episodes** | 300 |
| **PSO Particles** | 40 |
| **PSO Iterations** | 50 |
| **Overhead Calculation** | Control pkts/Data pkts ratio |
| **Overhead Metric Threshold** | 0.25 |

For all the experiments and protocols, the parameter values in Table 1 were kept fixed to ensure fairness and reproducibility. By doing so, we can ensure a balanced comparison of the algorithms, where any differences in performance are due to the algorithms themselves and not some selective parameter tuning. While an exhaustive sensitivity analysis would be outside of our scope here, we recognize the need for it and parameter tuning strategies to ensure that performance gains are consistent across different datasets.

## 4. EXPERIMENTAL RESULTS AND EVALUATION

This section presents the experimental work and the analysis of the results. The section is composed of two subsections: the first is the MOO optimization, described in Section 4.1, and the second is the simulation-based evaluation, outlined in Section 4.2.

### 4.1 Multi-objective optimization evaluation (MOO)

This section reviews key evaluation metrics for MOO algorithms and compares the proposed PSOA-CRL with the benchmarks NSGA-II [13], GMHS [14], MOHS [15], and RBMOORPV [23]. The assessment uses four core metrics, set coverage, hypervolume, the delta metric, and generational distance, to evaluate solution optimality, convergence, and diversity across the objective space.

Figure 3 shows that in the set coverage metric, the PSOA-CRL achieved a higher value of 0.91, while RBMOORPV is 0.78, NSGA-II is 0.76, GMHS is 0.74, and MOHS is 0.73. This indicates its ability to consistently present dominant Pareto-optimal solutions. The result reflects the strength of combining the PSO's global search with actor–critic learning's adaptive refinement, allowing the PSOA–CRL to deliver more robust routing strategies in dynamic VANET environments.

Figure 4 shows the hypervolume, which measures the volume in the objective space covered by the Pareto front. A higher value means better convergence and diversity. The PSOA-CRL achieved the highest value of 0.86, whereas NSGA-II 0.75, RBMOORPV 0.72, GMHS 0.70, and MOHS 0.66. The proposed method covers a larger and more efficient area of the solution space, which means that its results are both high quality and well spread. This is critical in real-time applications where diverse routing trade-offs (PDR, E2E delay, energy, link reliability, and routing overhead) must be available. The high hypervolume is a direct outcome of the exploration power of the PSO combined with the adaptive learning of the actor-critic model, enabling the PSOA-CRL to maintain a diverse but converged set of routing solutions suitable for dynamic VANET environments.

Figure 5 presents the generational distance metric, which reflects the proximity of solutions to the true Pareto-optimal front. The PSOA-CRL method achieved 0.08, whereas the NSGA-II reached 0.12, the GMHS reached 0.14, the RBMOORPV reached 0.16, and the MOHS reached 0.18. A lower value of the generation distance metric indicates better convergence. This result proves that the solutions produced by the PSOA-CRL are closer to the ideal, ensuring optimal routing decisions in practice. This superior convergence is driven by A-CRL, which continuously fine-tunes routing policies on the basis of real-time feedback, surpassing the static or nonadaptive nature of competing approaches.

Figure 6 presents the delta metric, which reflects the distribution uniformity among the Pareto solution sets. The PSOA-CRL achieved 0.18, whereas the NSGA-II reached 0.27, the GMHS reached 0.30, the RBMOORPV reached 0.31, and the MOHS reached 0.32. A lower value indicates better diversity and balanced spread. Therefore, PSOA-CRL not only optimizes the routes but also ensures that they are balanced, providing an ample weighted range for the decision maker's options. The increased diversity is due to the global search capability of PSO and local refinement provided by reinforcement learning working together to maintain a well-spread set of solutions along the Pareto front.

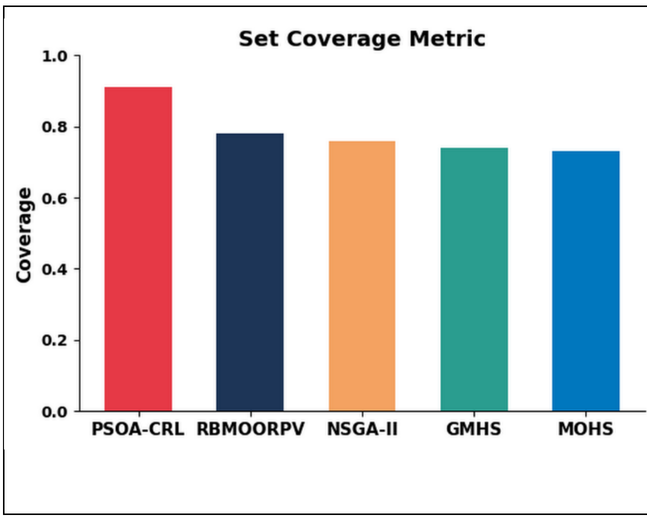



Fig. 3. Set the coverage metric.

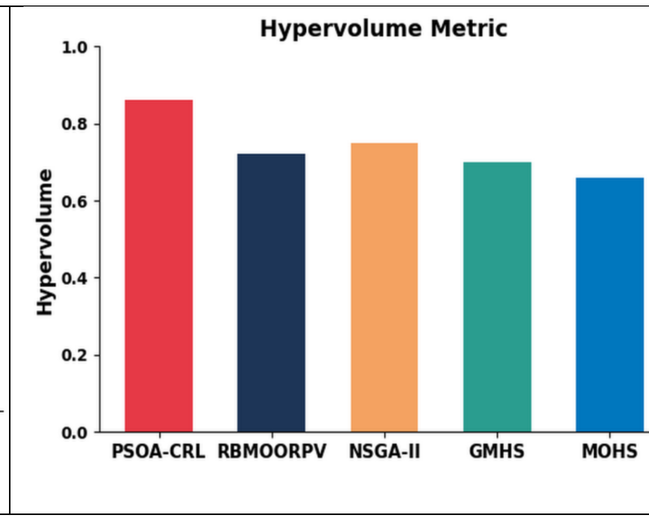



Fig. 4. Hypervolume metric.

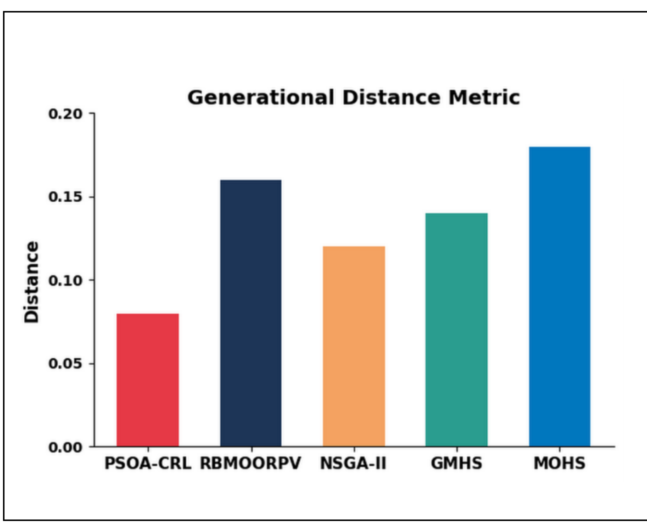



Fig. 5. Generation distance metric.

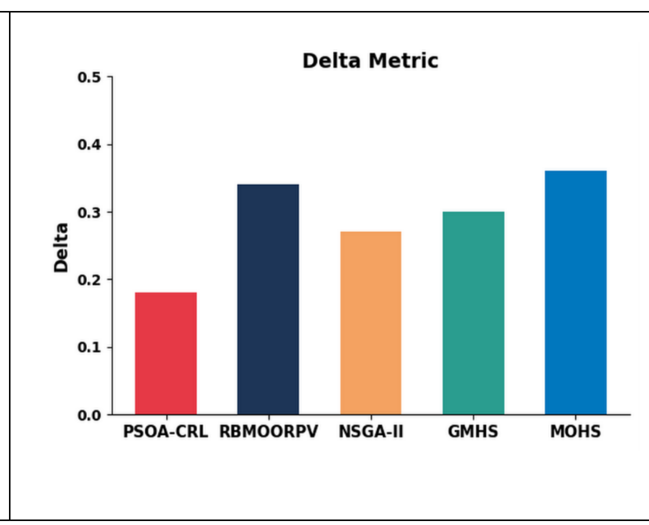



Fig. 6. Delta metric.

## 4.2 Simulation-based evaluation

This section evaluates the PSOA-CRL algorithm via five metrics and analyzes the results in an urban VANET environment. Its adaptability, effectiveness, and dominance over benchmark protocols in all metrics are clearly achieved.

As shown in Figure 7, PSOA-CRL achieves a PDR of 97.8%, which is superior to those of RBMOORPV (94.2%), NSGA-II (92.1%), GMHS (91.4%), and MOHS (89.5%). The highest value of PDR performance is derived from trust-aware routing and dynamic path optimization reinforced with QoS-aware reward reinforcement learning. This guarantees dependable results under dynamic VANET conditions.

In Figure 8, PSOA-CRL achieves a lower value of the E2E delay metric of 41.3 ms, which is superior to that of RBMOORPV at 54.8 ms, NSGA-II at 60.1 ms, GMHS at 66.7 ms, and MOHS at 70.2 ms. The reason for this reduction is real-time responsive adaptation to congestion and link quality, together with trust and QoS-aware reward mechanisms that strengthen low-latency stable routes crucial for delay-sensitive applications in VANETs.

Figure 9 shows that the link reliability metric of the PSOA-CRL, which is 96.1% higher than that of the RBMOORPV, is 91.3%, that of the NSGA-II is 89.6%, that of the GMHS is 88.4%, and that of the MOHS is 85.7%. This reflects the ability of the PSOA-CRL to maintain stable connections by avoiding low-trust or weak-signal nodes. Its actor-critical learning model adapts routes on the basis of real-time link quality and trust feedback, ensuring reliable communication, which is essential for safety and data integrity in VANETs.

Figure 10 shows the evaluation of the energy consumption metric. The PSOA-CRL achieves 184.2 J, outperforming RBMOORPV 207.5 J, NSGA-II 213.7 J, GMHS 219.3 J, and MOHS 224.1 J. This efficiency results from routing that avoids redundant hops and overloaded nodes, extends the network lifespan, and balances energy use. The QoS protocols and trust-aware decisions help select energy-efficient routes and evenly distribute network loads, enhancing VANET scalability.

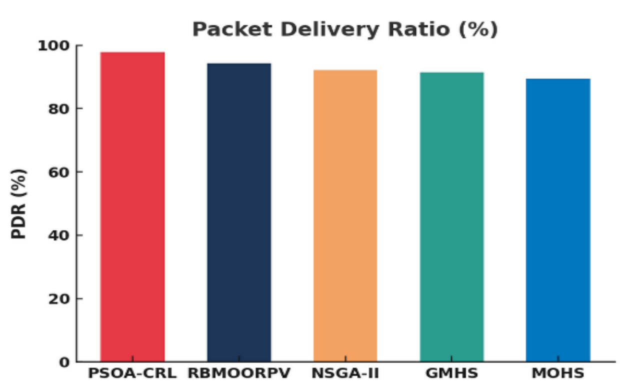



Fig. 7. Packet Delivery Ratio metric.

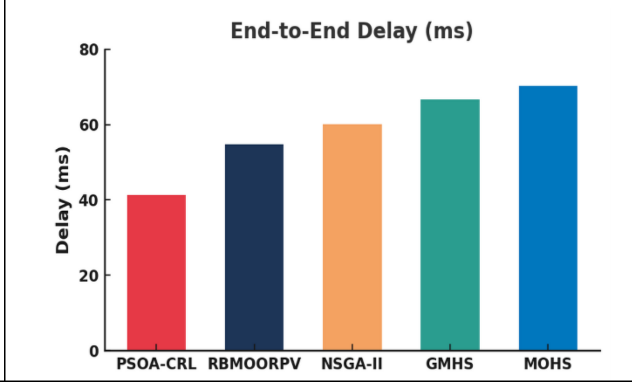
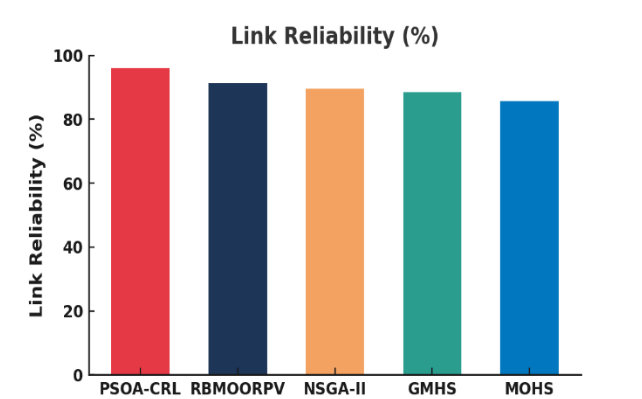



Fig. 8. End-to-End Delay metric.



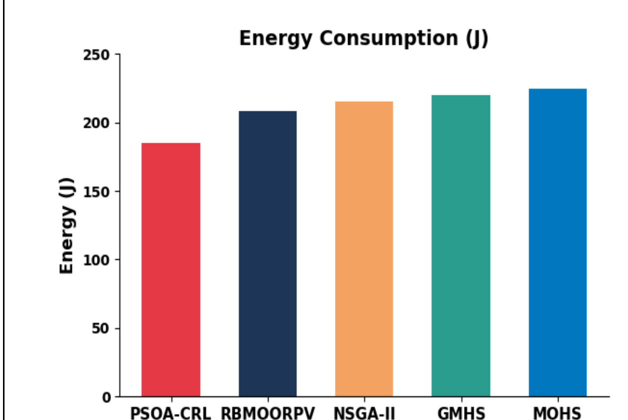



Fig. 9. Link reliability metric.

Fig. 10. Energy consumption metric.

Figure 11 presents the routing overhead metric. The overhead metric, which is the count of control packets divided by data successfully delivered in roadside networks, is crucial in VANETs to prevent network congestion and maintain scalability. The PSOA-CRL presents the lowest value of overhead of 0.21, whereas the RBMOORPV is 0.31, the NSGA-II is 0.35, the GMHS is 0.37, and the MOHS is 0.41. This efficiency stems from the hybrid design of the PSOA-CRL, which combines offline PSO policy generation with online actor-critical learning for adaptive route refinement. Its threshold-triggered re-optimization limits unnecessary updates, ensuring stable routing and minimal control traffic. Importantly, PSOA-CRL remains below the predefined acceptable overhead threshold of 0.25, demonstrating its effectiveness in managing communication resources in dynamic VANETs. Throughout the experiments, the routing overhead always remained below the set threshold, so re-optimization was very infrequent. In this way, for any number of vehicles considered, the runtime was highly dominated by lightweight local updates, which maintained constant growth as the number of vehicles increased and supported the PSOA-CRL scalability.

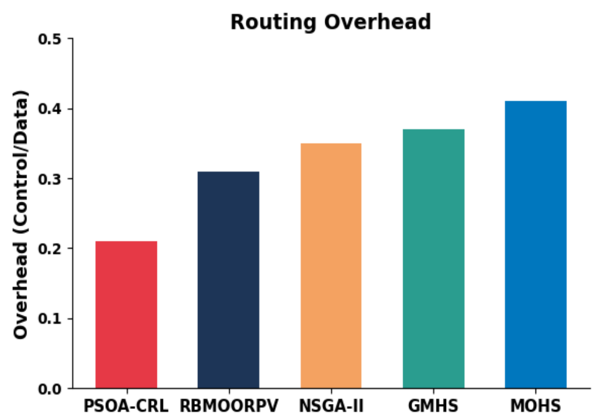



Fig. 11. Routing Overhead Metrics.

### 4.3 Performance Analysis of the PSOA-CRL Protocol

Table 3 shows that the PSOA-CRL significantly outperforms the other four multi-objective routing protocols. Among the MOO metrics, the proposed PSOA-CRL achieves the highest set coverage at 0.91, the lowest delta value of 0.18, and the smallest generational distance at 0.08, which are indicators of both diverse solutions and strong convergence toward the optimal front. Additionally, it has the highest hypervolume at 0.86, reflecting comprehensive and well-balanced coverage of the objective space, while from a network performance standpoint, the PSOA-CRL has the best network systems in terms of a packet delivery rate of 97.8%, the lowest end-to-end delay of 41.3 ms, and the highest link reliability at 96.1%, while consuming the least energy of 184.2 J. Additionally, it has the lowest routing cost at 0.21, demonstrating its high communication efficiency. This performance is largely attributed to its hybrid design, which combines an offline PSO system with a real-time A-CRL system and trust-based filtering to reduce unnecessary control traffic. The findings confirm that the PSOA-CRL is a balanced, adaptable, and efficient routing protocol, making it especially suitable for the fast, dynamic, and demanding environments of VANET systems. The RBMOORPV result values are shown in this paper, such as PDR, delay, and routing overhead, which are different from the results of this paper because RBMOORPV is implemented again on the basis of our simulation scenario with 200 vehicles and a unified configuration for all protocols. These variances are expected and acceptable; they represent the real-time performance of the protocols under the same conditions in our work. This suggested verification indicates that the PSOA-CRL is a well-balanced, provoking, and promising routing protocol, and its performance is relatively sufficient, especially for the VANET networks under high-speed and dynamic road traffic scenarios. Moreover, the algorithm also exhibits high and steady performance in high-vehicular-density scenarios, which demonstrates its adaptivity in large-scale VANET systems. This is because of the localized A-CRL decision and the event-driven PSO re-optimization, collectively resulting in fast convergence and reduction in computational cost, leading to real-time adaptability.

TABLE III. PERFORMANCE COMPARISON OF ROUTING PROTOCOLS IN VANETS.

| Metric | PSOA-CRL | RBMOORPV | NSGA-II | GMHS | MOHS | Summary |
|---|---|---|---|---|---|---|
| **Set Coverage** | 0.91 | 0.78 | 0.76 | 0.74 | 0.73 | Most dominant solutions |
| **Delta Metric** | 0.18 | 0.34 | 0.27 | 0.30 | 0.36 | Best diversity |
| **Hypervolume** | 0.86 | 0.72 | 0.75 | 0.70 | 0.66 | Largest covered space |
| **Generational Distance** | 0.08 | 0.16 | 0.12 | 0.14 | 0.18 | Closest to the optimal front |
| **PDR (%)** | 97.8 | 94.2 | 92.1 | 91.4 | 89.5 | Highest packet success |
| **E2E Delay (ms)** | 41.3 | 54.8 | 60.1 | 66.7 | 70.2 | Fastest routing decisions |
| **Link Reliability (%)** | 96.1 | 91.3 | 89.6 | 88.4 | 85.7 | Most stable routes |
| **Energy Consumption (J)** | 184.2 | 207.5 | 213.7 | 219.3 | 224.1 | Best energy efficiency |
| **Overhead** | 0.21 | 0.31 | 0.35 | 0.37 | 0.41 | Most communication-efficient |

## 5. LIMITATIONS

The PSOA-CRL framework was tested via simulation (OMNeT++ and SUMO). Although the improvements on the baselines are substantial, there remain some drawbacks. First, there remain some other uncertainties in real vehicle networks that were not well addressed through the simulations, e.g., hardware limitations of nodes, heterogeneous communication technologies, and even unpredictable driver behaviors. Second, even though the 200 vehicles scaled best among the high-density phase-based deployments, the computational overhead that such a model would place on the OBU in real deployments needs to be re-examined. Finally, only a few configuration parameters (the trust threshold and PSO iteration settings) were experimented; in order to prove the robustness in more general conditions in the network, a complete sensitivity analysis on the t variable is needed. The alleviation of these limitations in the future will also further strengthen the practicability and robustness of the PSOA-CRL for real VANET applications.

## 6. CONCLUSION

This study proposed the PSOA-CRL, a hybrid multi-objective routing algorithm designed for highly dynamic VANET environments. The combination of PSO global optimization with A-CRL adaptive learning enables the method to efficiently handle rapid topology changes, high mobility, and dynamic traffic conditions, outperforming previous approaches in key performance metrics, including packet delivery, end-to-end delay, link reliability, energy efficiency, and routing overhead. On the other hand, the proposed framework achieves real-time performance even on large-scale networks, incurs reasonable computational complexity, and maintains a clear modular and localized structure, which may ensure its gradual and scalable deployment in practical ITS. Deployment is still a challenge due to limitations in hardware and communication, but the flexibility of the framework enables integration with new vehicle communication standards and technologies. In the future, we will experimentally validate the approach in practical VNs, integrate the proposed approach with 5G/6G communication technologies, improve the optimization procedures, extend the evaluation to more details of the QoS, and investigate the systematic sensitivity analysis of key parameters to analyze the influence on performance metrics and evaluate the robustness and practical significance of the proposed method.

### Conflicts of interest

The authors declare that they have no conflicts of interest.

### Funding

None.

### Acknowledgement

None

### References

[1] S. Al-Sultan, M. M. Al-Doori, A. H. Al-Bayatti, and H. Zedan, "A comprehensive survey on vehicular Ad Hoc network," J. Netw. Comput. Appl., vol. 37, pp. 380–392, 2014. DOI: 10.1016/j.jnca.2013.02.008.

[2] R. C. L. Oliveira, G. C. M. Moura, and G. Pujolle, "VANETs: Survey, protocol taxonomy, and research challenges," Telecommun. Syst., vol. 62, pp. 15–32, 2016. DOI: 10.1007/s11235-015-9992-2.

[3] R. S. Raw, S. Balakrishnan, and S. P. Singh, "Security challenges, issues and their solutions for VANET," J. Netw. Comput. Appl., vol. 36, no. 6, pp. 1677–1694, 2013. DOI: 10.1016/j.jnca.2013.02.002.

[4] Aliesawi, Salah Ahmed, et.al, “Iterative Multipacket Detection With FDE Based MAC Protocol In Vehicular Ad Hoc Networks,” Periodicals of Engineering and Natural Sciences, vol. 7, no. 3, pp. 1041 – 1053, 2019, doi: 10.21533/pen.v7i3.623.
[5] A. M. Awaad, K. M. A. Alheeti, and A. K. A. Najem, “Anomaly-based IDS (intrusion detection system) for cyber-physical systems,” Mesopotamian J. Big Data, vol. 2024, pp. 230–240, Dec. 2024. DOI: 10.58496/MJBD/2024/017.
[6] M. Rizwan, K. R. Naqvi, and H. A. Khawaja, “Multi-objective Particle Swarm Optimization for routing in vehicular ad hoc networks,” Appl. Soft Comput., vol. 97, 2020. DOI: 10.1016/j.asoc.2020.106836.
[7] M. Q. AL-Shammari and R. C. Muniyandi, “Optimised Tail-based Routing for VANETs using Multi-Objective PSO,” IJACSA, vol. 11, no. 6, 2020. DOI: 10.14569/IJACSA.2020.0110629.
[8] S. A. Rashid, M. Alhartomi, L. Audah, and M. M. Hamdi, “Reliability-aware multi-objective optimization-based routing protocol for VANETs using enhanced Gaussian mutation harmony searching,” IEEE Access, vol. 10, pp. 26613–26627, 2022. DOI: 10.1109/ACCESS.2022.3167895.
[9] Ye, S., Xu, L., Xu, Z., & Wang, F. (2024). A deep reinforcement learning-based intelligent QoS optimization algorithm for efficient routing in vehicular networks. Alexandria Engineering Journal, 107, 317-331.
[10] M. Useng and S. Abdulrahman, “A survey on distributed reinforcement learning,” Mesopotamian J. Big Data, vol. 2022, pp. 44–50, Nov. 2022. DOI: 10.58496/MJBD/2022/006.
[11] Zhang, Z., Wu, Q., Fan, P., Cheng, N., Chen, W., & Letaief, K. B. (2025). DRL-Based Optimization for AoI and Energy Consumption in C-V2X Enabled IoV. IEEE Transactions on Green Communications and Networking.
[12] H. Yu, Z. Zheng, and K. Ren, “A hybrid reinforcement learning approach for adaptive routing in wireless networks,” IEEE/ACM Trans. Netw., vol. 29, no. 6, pp. 2761–2774, 2021. DOI: 10.1109/TNET.2021.3110158.
[13] Hamdi, M.M.; Rashid, S.A.; Nafea, A.A. Resource Allocation and Edge Computing for Dual Hop Communication in Satellite Assisted UAVs enabled VANETs. Iraqi Journal For Computer Science and Mathematics, vol. 5, no. 3, Art. no. 7, 2024.
[14] Ali Alheeti, K. M., Gruebler, A., & McDonald-Maier, K. (2016). Intelligent intrusion detection of grey hole and rushing attacks in self-driving vehicular networks. Computers, 5(3), 16.
[15] K. Deb, S. Agrawal, A. Pratap, and T. Meyarivan, “A Fast Elitist Non-dominated Sorting Genetic Algorithm for Multi-objective Optimization: NSGA-II,” in Parallel Problem Solving from Nature PPSN VI, Lecture Notes in Computer Science, vol. 1917, M. Schoenauer et al., Eds., Springer, Berlin, Heidelberg, 2000, pp. 849–858. DOI: 10.1007/3-540-45356-3_83.
[16] X. Dai, X. Yuan, and L. Wu, “A novel harmony search algorithm with Gaussian mutation for multi-objective optimization,” Soft Computing, vol. 21, no. 6, pp. 1549–1567, 2017.
[17] J. Ricart, G. Hüttemann, J. Lima, and B. Barán, “Multiobjective harmony search algorithm proposals,” Electron. Notes Theor. Comput. Sci., vol. 281, pp. 51–67, 2011.
[18] A. Kharrufa, A. Al-Fuqaha, and M. Guizani, “Multi-objective Optimization of VANET Networks for Intelligent Parking with Hybrid PSO-NSGA II Algorithm,” Vehicular Commun., vol. 34, 2022.
[19] A. Sharma, S. K. Singh, and A. Kumar, “An improved multi-objective optimization framework with a hybrid model for vehicular ad hoc network routing services,” Wireless Netw., vol. 28, 2022.
[20] S. Huang, Z. Wang, and Y. Chen, “A multi-objective optimization approach for roadside unit deployment strategy in IoV,” Ad Hoc Netw., vol. 131, 2022.
[21] A. Iftikhar, M. Z. Shakir, and A. Imran, “An optimized framework for VANET routing: A multi-objective hybrid model for data synchronization with digital twin,” Comput. Commun., vol. 192, 2022.
[22] Y. Zhang, L. Wang, and X. Liu, “Multi-objective Harris hawks optimization algorithm based 2-hop routing algorithm for CR-VANET,” Future Gener. Comput. Syst., vol. 136, 2023.
[23] R. Verma, A. Kumar, and P. Singh, “Reliability-Aware Multi-Objective Optimization-Based Routing Protocol for VANETs Using Enhanced Gaussian Mutation Harmony Searching,” Wireless Pers. Commun., vol. 128, 2023.
[24] M. Kaur, A. Singh, and R. Kaur, “A Multi-objective Approach for Secure Cluster Based Routing & Attack Classification in VANETs,” J. Ambient Intell. Humaniz. Comput., vol. 14, 2023. DOI: 10.1007/s12652-022-03850-5.
[25] F. Wang, X. Zhang, and Y. Liu, “Reliability-Based Multi-Objective Optimized Routing Protocol for VANETs Using Pelican Optimization Algorithm (RBMOORP),” IEEE Access, vol. 11, 2023.
[26] M. B. Taha, C. Talhi, H. Ould-Slimane, S. Alrabaee, and K. K. R. Choo, “A multi-objective approach based on differential evolution and deep learning algorithms for VANETs,” IEEE Transactions on Vehicular Technology, vol. 72, no. 3, pp. 3035–3050, 2022.
[27] J. Deng, J. Wang, X. Wang, Y. Cai, and P. Liu, “Multi-Task Multi-Objective Evolutionary Search Based on Deep Reinforcement Learning for Multi-Objective Vehicle Routing Problems with Time Windows,” Symmetry, vol. 16, no. 8, p. 1030, 2024.

[28] Q. Yang and S. J. Yoo, “Hierarchical reinforcement learning-based routing algorithm with grouped RSU in urban VANETs,” IEEE Trans. Intell. Transp. Syst., vol. 25, no. 8, pp. 10131–10146, 2024.
[29] Y. J. Song, C. E. Yen, Y. H. Hsieh, C. Kuo, and I. C. Chang, “An intersection-based traffic awareness routing protocol in VANETs using deep reinforcement learning,” Wireless Pers. Commun., vol. 138, no. 1, pp. 659–683, 2024.
[30] Q. Hussain, A. S. M. Noor, M. M. Qureshi, J. Li, A. U. Rahman, A. Bakry, M. A. Khan, and A. Rehman, “Reinforcement learning based route optimization model to enhance energy efficiency in internet of vehicles,” Scientific Reports, vol. 15, no. 1, p. 3113, 2025.