



The Scientific Temper

JULY 2026 Vol. 17 (7): 6618-6631

ISSN: 0976-8653, E-ISSN: 2231-6396

DOI: 10.58414/SCIENTIFICTEMPER.2026.17.7.2499

<https://scientifictemper.com/>

RESEARCH ARTICLE

Optimized Multi-Path Relay Node Selection Using Deep Reinforcement Learning for Reliable and Energy-Efficient MANET Routing

A. ABDUL SAMATHU¹, G. RAVI², A. R. MOHAMED SHANAVAS³

^{1,2,3}Jamal Mohamed College (Autonomous), Affiliated to Bharathidasan University, Tiruchirappalli – 620020, India

Corresponding Author*: aas@jmc.edu

ABSTRACT

Mobile Ad-hoc Network (MANET) are highly dynamic and infrastructure-less wireless network in which frequent topology changes, node mobility, packet collision, and energy constraints significantly affect routing performance and network reliability. This research suggests a Deep Reinforcement Learning (DRL)-based Optimized Multi-Path Relay Node Selection method for dependable and energy-efficient MANET routing. The suggested method takes into account important network metrics such as residual energy, node mobility, link stability, congestion level, and packet collision probability in order to intelligently choose the best relay nodes and different routing options using a Deep Q-Network (DQN)-based learning model. In comparison to traditional MANET routing protocols, simulation results show that the suggested DRL-based relay node selection technique greatly improves network lifetime, Packet Delivery Ratio (PDR), throughput and routing stability while lowering packet collision, end-to-end delay and energy consumption.

Keywords: Mobile Ad Hoc Network, Multi-Path Routing, Deep Reinforcement Learning, Relay Node Selection, Energy-Efficient Routing, Packet Collision Avoidance, Deep Q-Network

INTRODUCTION

The potential of machine learning and artificial intelligence approaches to learn network behaviour and make intelligent routing decisions has recently attracted a lot of attention in MANET routing research. Multipath routing in MANET is depicted in Figure 1. Supervised

learning and unsupervised learning approaches have been applied for link prediction, congestion detection, intrusion detection, and traffic classification in MANET (Sun et al., 2019). However, these approaches often require large labelled dataset and may not effectively support real-time routing decisions in continuously changing network conditions.

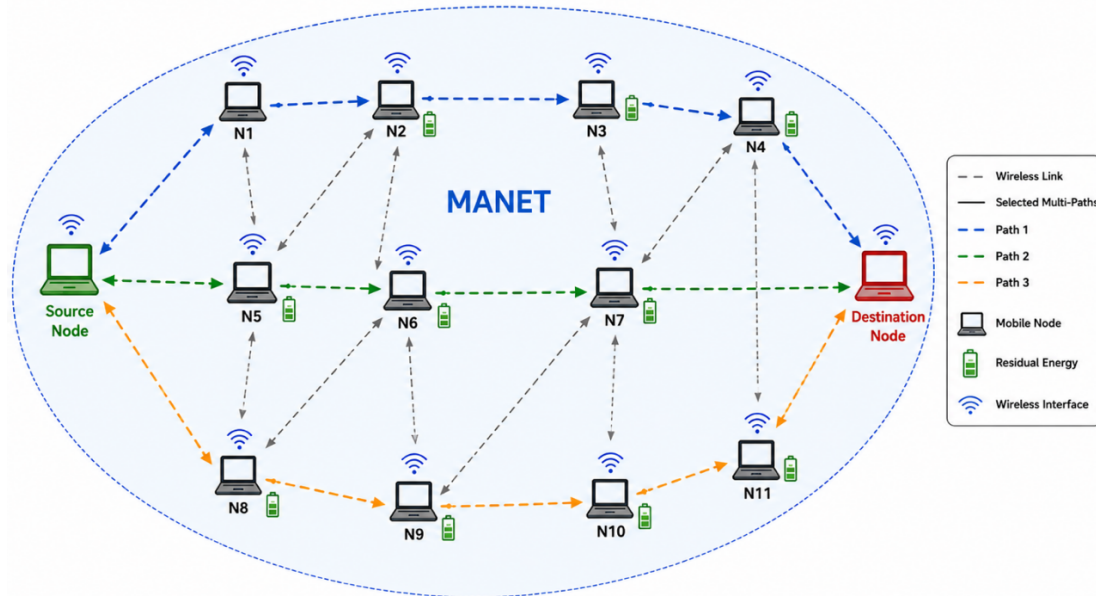


Figure 1: Multipath Routing in MANET

Reinforcement Learning (RL) has emerged as a potential solution to adaptive routing in MANET. RL allows network agents to discover the most effective routing policies by maximizing cumulative rewards while communicating with the environment (Sutton & Barto, 2018). Q-learning-based routing protocols have been proposed to improve packet forwarding decisions using metrics such as delay, queue length, and energy availability. These approaches enable nodes to dynamically select routes based on changing network conditions without requiring prior knowledge of the environment. Nevertheless, conventional RL methods suffer from scalability issues and slow convergence when applied to large-scale MANET with complex state spaces.

DRL combines RL with deep neural network, has recently shown remarkable success in solving complex decision-making problems. DRL techniques such as DQN, Deep Deterministic Policy Gradient (DDPG), and Actor-Critic methods have been increasingly explored for intelligent routing and resource management in wireless network. DRL-based routing algorithms in MANET can efficiently learn optimal relay node selection rules by taking into account many network metrics at the same time, such as residual energy, connection reliability, node mobility, congestion, and signal strength. In comparison to traditional routing strategies, these methods offer greater versatility, faster convergence, and improved decision-making capability.

LITERATURE REVIEW

N. C. Luong et al. demonstrated a thorough investigation on the use of DRL in

communication and networking systems. The study revealed a DRL's capability to enhance network resource management, routing, and decision-making processes in dynamic contexts. However, the work was primarily concerned with conventional communication networks and did not address relay node selection or energy-efficient multi-path routing in highly dynamic MANET environments (Luong et al., 2019).

DRL approaches for wireless networks and the Internet of Things were thoroughly surveyed by A. Alwarafy et al. The study emphasized how DRL may be used to solve problems including resource allocation, network optimization, and adaptive decision-making. However, the study mostly offered a theoretical overview and lacked real-world applications for dependable routing and adaptive relay node selection in MANET systems (Alwarafy et al., 2021).

X. Li et al. suggested a Q-Learning-Based Routing Protocol for MANET that would improve route maintenance and discovery by utilizing RL. The protocol allowed nodes to discover efficient routing paths depending on network conditions, reducing reliance on traditional routing metrics. However, when applied to large-scale and highly mobile MANET, classical Q-learning produced delayed convergence and restricted scalability (Li et al., 2020).

A Deep Reinforcement Learning-based resource allocation system for communications in Vehicle-to-Vehicle (V2V) was presented by H. Ye et al. The suggested approach successfully increased communication reliability and spectrum usage in dynamic wireless situations. The study showed the efficacy of DRL, but it ignored the particular

difficulties related to MANET routing and energy efficiency and concentrated on resource allocation rather than relay node selection (Ye et al., 2019).

An Energy-Efficient Multipath Routing Protocol was introduced for MANET by K. S. Vaisakh and R. Venkatesan to increase network lifetime and reduce energy consumption. In order to improve communication reliability and balance traffic loads, the protocol used numerous routing paths. Despite this, the relay node selection procedure lacked intelligent adaptive mechanisms that could react to quickly shifting network conditions and instead depended on predetermined routing metrics (Vaisakh & Venkatesan, 2021).

Routing protocols, threats to security, mobility management, and new research approaches were all covered in S. Roy et al.'s comprehensive review of MANET. The study offered insightful information about the strengths and weaknesses of current MANET routing strategies. However, the survey found that while sophisticated AI-based relay node optimization strategies are still not extensively investigated, the majority of conventional routing protocols suffer from frequent route failures and low adaptability (Roy et al., 2020).

Y. Sun et al. examined the usage of machine learning techniques in wireless networks and discussed about the manner in which they could be used for resource allocation, traffic control, and network optimization. The study highlighted the advantages of intelligent decision-making processes and listed a number of unresolved research issues. However, neither energy-aware multi-path routing nor adaptive relay node selection techniques for MANET systems were particularly examined in this work (Sun et al., 2019).

The AODV routing protocol, developed by C. Perkins et al., reduces routing cost in MANETs by establishing routes only when necessary. The protocol became one of the most popular reactive routing strategies due to its effectiveness and simplicity. However, AODV uses single-path routing, it is susceptible to frequent route breakages, packet losses, and higher latency in situations with high mobility (Perkins et al., 2003).

D. B. Johnson et al. developed the DSR protocol which employs route caching and source routing mechanisms to discover and maintain communication paths in MANET. In comparatively stable network conditions, the protocol successfully lowers routing overhead. However, in highly dynamic MANET applications, its performance is severely degraded by inactive route cache information

and increased packet header overhead (Johnson et al., 2007).

The Optimized Link State Routing protocol which uses Multipoint Relays (MPR) to increase routing efficiency and lower flooding overhead was suggested by T. Clausen and P. Jacquet. OLSR offers proactive route maintenance and operates well in networks with modest mobility. However, in dense and highly mobile MANET environments, the static MPR selection technique may result in inadequate flexibility and excessive control overhead which would increase energy consumption and lower routing performance (Clausen & Jacquet, 2003).

Routing protocols for Mobile Ad Hoc Networks (MANETs) were thoroughly reviewed by M. Abolhasan, T. Wysocki, and E. Dutkiewicz, who categorized them into proactive, reactive, and hybrid routing categories. Under various network situations, the study emphasized the strengths and challenges of current routing techniques. Nevertheless, intelligent methods for adaptive relay node selection and energy-aware routing decisions have been excluded from the majority of protocols studied, which relied on predetermined routing metrics (Abolhasan et al., 2004).

A review of existing routing protocols in ad hoc mobile wireless networks was proposed by E. M. Royer and C. K. Toh, with an emphasis on scalability, maintenance, and route discovery. Frequent topology changes have a substantial impact on routing performance, as their investigation showed. However, in highly mobile MANETs, this resulted in higher routing overhead and worse dependability because the protocols under investigation could not dynamically learn from network conditions (Royer & Toh, 1999).

The Temporally Ordered Routing Algorithm (TORA) was developed by S. R. Das, C. E. Perkins, and E. M. Royer to offer highly dispersed and adaptive routing in dynamic MANET environments. Through localized route maintenance, the protocol effectively responded to topological changes. However, in high mobility scenarios, TORA produced significant control message overhead, which had an adverse effect on network scalability and energy efficiency (Park & Corson, 2001).

The Zone Routing Protocol (ZRP), which combines reactive and proactive routing techniques to increase routing efficiency in MANETs, was suggested by V. D. Park and M. S. Corson. The hybrid strategy preserved scalability while lowering route finding latency. However, ZRP's performance was highly dependent on the ideal zone radius configuration, making it

challenging to adjust to quickly evolving network topologies (Haas et al., 2002).

Multipoint relay (MPR)-based routing techniques were presented by P. Jacquet et al. to lower flooding overhead in wireless ad hoc networks. The method greatly reduced redundant transmissions and increased the effectiveness of routing. However, especially in highly dynamic and crowded network contexts, static MPR selection frequently led to relay node selections that were not ideal (Jacquet et al., 2001).

The Ad Hoc On-Demand Multipath Distance Vector (AOMDV) routing protocol was developed by M. K. Marina and S. R. Das to offer numerous loop-free pathways between sender and receiver nodes. The methodology decreased the frequency of route discovery and increased fault tolerance. However, especially in large-scale MANET deployments, maintaining numerous routes increased routing overhead and energy usage (Marina & Das, 2001).

A. Nasipuri, R. Castaneda, and S. R. Das proposed multipath routing strategies for wireless ad hoc networks to improve load balancing and route reliability. The approach effectively distributed traffic across several routes to reduce traffic. Despite this, the protocol lacked adaptive techniques to choose the best relay nodes in response to changing network conditions (Nasipuri et al., 2001).

Techniques for multi-hop wireless ad hoc network routing were compared in terms of performance by J. Broch, D. A. Maltz, D. B. Johnson, Y. Hu, and J. Jetcheva. According to their research, mobility and traffic conditions had a major impact on protocol performance. Nevertheless, energy-aware intelligent routing issues were not addressed by the assessed protocols, which depended on traditional routing metrics (Broch et al., 1998).

Reinforcement learning methods for adaptive decision-making in dynamic contexts were first presented by R. S. Sutton and A. G. Barto. Their framework made it possible for agents to interact with their environment and learn the optimal action. However, when applied to complicate MANET routing issues, traditional reinforcement learning techniques suffer from enormous state-space restrictions and slow convergence (Sutton & Barto, 2018).

The Deep Q-Network (DQN) architecture, developed by V. Mnih et al., combines deep neural networks and reinforcement learning to address challenging decision-making problems. Compared to traditional RL techniques, the strategy showed notable performance gains. However, the original DQN framework needed

additional modification for MANET setups because it was not intended for dynamic routing scenarios (Mnih et al., 2015).

The Deep Deterministic Policy Gradient (DDPG) technique was developed by T. Lillicrap et al. to manage ongoing action spaces in reinforcement learning applications. The approach enhanced policy optimization capabilities and learning efficiency. However, it was not directly applicable to highly dynamic network routing problems due to training instability and sensitivity to hyperparameter settings (Lillicrap et al., 2015).

A thorough study on Deep Reinforcement Learning was provided by V. François-Lavet, P. Henderson, R. Islam, M. G. Bellemare, and J. Pineau. They highlighted recent developments in policy learning and optimization. The study showed the potential of DRL for challenging decision-making tasks. Nevertheless, energy-efficient relay node selection in MANETs has not yet received much attention in actual applications (François-Lavet et al., 2018).

Mobility prediction approaches were proposed by Y. Gwon, J. Kempf, and A. D. Sarma for MANET routing in order to increase route stability and decrease route failures. The protocol made an effort to anticipate node mobility patterns in order to proactively maintain dependable communication channels. However, poor routing performance was frequently the consequence of prediction errors in highly dynamic mobility conditions (Gwon et al., 2004).

Hierarchical routing techniques for mobile wireless networks were studied by R. Ramanathan and M. Steenstrup in order to increase communication efficiency and scalability. By grouping nodes into structured clusters, the method simplified routing. However, its efficacy in rapidly changing MANET systems was restricted by cluster maintenance overhead and frequent reconfiguration (Ramanathan & Steenstrup, 1998).

The investigation of mobile ad-hoc networking designs and routing challenges by S. Giordano focused on problems including energy limitations, mobility control, and service quality. The necessity of intelligent and adaptable routing algorithms was emphasized in the study. Nevertheless, learning-based optimization methods that could simultaneously take energy efficiency, dependability, and route stability in MANETs were not present in the systems that were in use at the time (Giordano, 2002)

This paper presents an Optimized Multi-Path Relay Node Selection method using DRL. By taking into account factors including residual energy, connection stability, node mobility, congestion level, and packet collision probability,

the suggested method uses a DQN-based model to intelligently choose relay nodes and routing paths. In order to enhance PDR, throughput, routing stability, and network longevity while lowering delay, collision, and energy consumption, the DRL agent continuously learns from network conditions and dynamically modifies routing decisions. The proposed model provides an intelligent and scalable routing solution for next-generation MANET applications.

OBJECTIVES

- To develop an optimized multi-path routing framework that dynamically selects reliable relay nodes based on parameters such as residual energy, link stability, node mobility, and congestion level in MANET.
- To design and implement a DRL based relay node selection mechanism for adaptive and intelligent routing decisions in highly dynamic MANET environments.
- To minimize energy consumption, routing overhead, packet collisions, and end-to-end delay while improving throughput and PDR through efficient multi-path communication.
- To compare and evaluate the performance of proposed DRL-based routing approach with existing MANET routing protocols such as AODV, DSR, and AOMDV using standard network performance metrics.

METHODOLOGY

The proposed Deep Reinforcement Learning-based Optimized Multi-Path Relay Node Selection (DRL-OMRNS) algorithm is an intelligent routing framework designed to enhance reliable and energy-efficient communication in MANET. The proposed approach integrates DRL with multi-path routing to dynamically select optimal relay nodes based on multiple network parameters including residual energy, link quality, node mobility, congestion level, link stability, packet forwarding ratio, and hop count. The suggested algorithm, in contrast to traditional routing methods, continuously learns the best routing policies from network interactions and modifies relay node selection choices in response to real-time topology changes. The framework is appropriate for highly mobile environments like vehicular networks, military communication, disaster recovery systems and IoT-enabled MANET applications because it enhances load balancing, fault tolerance, communication reliability and collision reduction through the use of multiple disjoint routing paths.

Algorithm 1: Deep Reinforcement Learning-Based Optimized Multi-Path Relay Node Selection (DRL-OMRNS)

Algorithm: DRL-OMRNS

Input:

Network topology $G(N, E)$

Source node S

Destination node D

Thresholds: Energy_T, Stability_T, LinkQuality_T

DRL parameters: α, γ, ϵ

Weights: $w_1, w_2, w_3, w_4, w_5, w_6, w_7$

Output:

Optimal relay node set R

Begin

1. Initialize CandidateList $\leftarrow \emptyset$

2. Initialize RelaySet $R \leftarrow \emptyset$

Step 1: Compute Composite Relay Score

For each node $n \in N$ do

 Compute normalized residual energy:

$$E_{\text{norm}}(n) = \frac{E_{\text{res}}(n)}{E_{\text{Max}}}$$

 Compute link stability:

$$LS(n) = \frac{T_{\text{stable}}}{T_{\text{total}}}$$

 Compute normalized mobility:

$$M(n) = \left(\frac{1}{t}\right) \times \sum |Pn(i) - Pn(i-1)|$$

$$M_{\text{norm}}(n) = \frac{M(n)}{M_{\text{max}}}$$

 Compute normalized link quality:

$$LQ_{\text{norm}}(n) = \frac{SNR(n)}{SNR_{\text{max}}}$$

 Compute relay score:

$$\begin{aligned} RS(n) = & w_1 \times E_{\text{norm}}(n) + w_2 \times LQ_{\text{norm}}(n) \\ & + w_3 \times LS_{\text{norm}}(n) \\ & + w_4 \times PFR_{\text{norm}}(n) \\ & - w_5 \times M_{\text{norm}}(n) \\ & - w_6 \times QL_{\text{norm}}(n) \\ & - w_7 \times HC_{\text{norm}}(n) \end{aligned}$$

End For

Step 2: Candidate Relay Node Discovery

For each neighbouring node n_i of source node S do

 If $E(n_i) > \text{Energy}_T$

 AND $LS(n_i) > \text{Stability}_T$

 AND $LQ(n_i) > \text{LinkQuality}_T$

 Then

 Compute $RS(n_i)$

Add ni to CandidateList
 End If
 End for
 End for

Step 3: DRL Environment Formation

Define state space using:
 residual energy,
 link quality,
 node mobility,
 queue length,
 path stability,
 neighbor density,
 and hop count

Define action space:
 relay node selection,
 routing path switching,
 forwarding strategy update

Define reward function:

$$Reward = \alpha \times PDR + \beta \times Throughput - \gamma \times Delay - \delta \times Energy_{Consumption} - \mu \times Collision_{Rate}$$

Initialize DRL agent

Step 4: Multi-Path Relay Set Construction

Sort CandidateList in descending order according to RS(n)
 Choose the top K nodes from the CandidateList.
 Construct RelaySet R

Ensure:
 Paths have low overlap or are node-disjoint.

Amount of traffic is balanced.

backup routes are available.

Step 5: DRL-Based Adaptive Relay Selection

During data transmission do
 For each routing path $p_i \in R$ do
 Monitor:
 link stability, packet delivery ratio, congestion level, delay, energy consumption
 If performance is below the threshold, then
 Relay scores are recalculated by the DRL agent
 Choose different relay nodes.

Update the route path dynamically

End If

End For

End During

Step 6: Collision Avoidance and Load Balancing

Spread traffic over several routes.
 Avoid relay nodes that are congested
 Dynamically activate alternate routes
 Cut down on retransmissions

Step 7: Loop Detection and Redundancy Elimination

For each routing path do

Detect loops using:

sequence numbers,
 visited-node tables

Remove redundant relay nodes

Avoid unnecessary retransmissions

End For

Step 8: Energy Consumption Optimization

Compute total energy consumption:

$$E_{total} = E_{tx} + E_{rx} + E_{idle}$$

Avoid low-energy relay nodes

Balance traffic load among paths

Step 9: Periodic Route Re-evaluation

Periodically or when topology changes do
 Calculate relay scores again.
 Modify the DRL policy
 Change any weak relay nodes.
 Create new, optimal multi-path routes

End

Return RelaySet R

End Algorithm

Step 1: Composite Relay Score Calculation

In this stage, a composite relay score is used to assess each mobile node in the MANET. Important routing factors such as residual energy, link quality, link stability, packet forwarding ratio, node mobility, congestion level, and hop count are taken into account while calculating the score. Highly mobile or congested nodes have lower scores than nodes with more energy, solid connectivity, and superior communication quality. This helps in

determining which relay nodes are the most dependable for routing.

Step 1(a): Residual Energy Calculation

The residual energy of each node is normalized with respect to the maximum available energy. This increases energy efficiency and network lifetime overall. Higher residual energy nodes are favoured since they can engage in routing for longer periods of time without failing. This increases energy efficiency and network lifetime overall.

Step 1(b): Link Stability Calculation

The duration of time the communication link stays connected during observation time is used to gauge how stable it is between nearby nodes. Routing reliability is increased by stable links because they lower retransmissions, packet loss, and route breakdown.

Step 1(c): Node Mobility Calculation

The mobility of a node is computed using its position changes over time. Nodes with a high degree of mobility are penalized because they can disrupt routing routes. In order to keep communication lines steady, lower mobility nodes are preferred.

Step 1(d): Link Quality Calculation

The Signal-to-Noise Ratio (SNR) is used to measure the quality of the link. Higher SNR nodes offer reduced packet loss and improved communication quality. As a result, relay selection favours nodes with high link quality.

Step 2: Candidate Relay Node Discovery

Neighbouring nodes of the source node are checked against predefined thresholds for energy, link stability, and link quality. Candidate relay nodes are only chosen from those nodes that meet all threshold requirements. Weak and unreliable nodes are eliminated from routing through this filtering process.

Step 3: DRL Environment Formation

The MANET routing problem is modelled as a DRL environment. The DRL environment is used to model the MANET routing challenge. Network conditions including energy, mobility, congestion, and path stability are all contained in the state space. Relay node selection and routing path modification are included in the action space. The goal of a reward function is to reduce latency, energy usage, and collisions while increasing packet delivery and throughput. The DRL agent learns the optimal routing strategy dynamically.

Step 4: Multi-Path Relay Set Construction

Candidate relay nodes are sorted according to their relay scores. Multiple routing pathways between the sender and the receiver are formed by selecting the top-ranked nodes. The algorithm assures that paths are node-disjoint or minimally overlapping to improve fault tolerance, load balancing and reliability,

Step 5: DRL-Based Adaptive Relay Selection

During data transmission, network performance metrics such as energy consumption, congestion, delay and PDR are continuously monitored. The DRL agent dynamically chooses other relay nodes and modifies the routing path if the routing performance deteriorates. This enables adaptive routing in highly dynamic MANET environments.

Step 6: Collision Avoidance and Load Balancing

Traffic is distributed across multiple routing paths instead of using a single path. Alternative routes are triggered when needed, and congested nodes are avoided. This enhances network speed and communication stability while lowering packet collisions and retransmissions.

Step 7: Loop Detection and Redundancy Elimination

The approach uses visited-node tables and sequence numbers to check routing paths for loops. Reducing redundant relay nodes and needless retransmissions lowers routing overhead and boosts communication effectiveness.

Step 8: Energy Consumption Optimization

The energy consumed totally by a given node is calculated by considering reception, transmission, and idle-state energy usage. Traffic is distributed over several pathways and low-energy nodes are avoided during routing. This contributes to the MANET's longer lifetime.

Step 9: Periodic Route Re-evaluation

Relay scores and DRL policies are updated on a regular basis or whenever the network topology changes. Weak relay nodes are replaced with better nodes, and new optimal routing paths are reconstructed. This ensures continuous adaptability and robust routing efficiency in dynamic MANET scenarios. The proposed DRL-Based Optimized Multi-Path Relay Node Selection (DRL-OMRNS) algorithm intelligently selects and maintains optimal relay nodes for energy-efficient and reliable multi-path MANET routing. The system constantly learns the best routing choices based on shifting network conditions by combining DRL with adaptive multi-path routing. The relay node selection

process utilizes multiple performance parameters including congestion level, node mobility, residual energy, path stability and link quality to ensure robust communication and prolonged network lifetime.

The suggested technique uses load balancing, dynamic route switching, collision avoidance, and adaptive path monitoring to further increase communication dependability. In contrast to conventional routing protocols, DRL-OMRNS uses reward-driven learning to continually update routing algorithms, allowing for rapid response to network congestion and topology changes. Stable and effective communication is ensured even in highly mobile MANET circumstances through intelligent relay replacement techniques and periodic route re-evaluation. As a result, the suggested framework is ideal for mission-critical and real-time MANET

applications that demand reliable, scalable, and energy-efficient routing performance.

EXPERIMENTAL SIMULATION SYSTEM

The DRL-OMRNS method is implemented and evaluated in the proposed study under dynamic MANET settings using NS-3. The simulator offers flexible integration of DRL with routing algorithms, multi-hop communication capabilities, and robust wireless channel and mobility models. Its open-source architecture allows customization of routing algorithms, energy-aware communication models, and DRL-based optimization techniques. As a result, NS-3 provides an effective platform for verifying the suggested framework's dependability, scalability, energy efficiency, congestion management, and communication stability in dynamic MANET environments. Table 1 given below provides a tabulation of the simulation parameters.

Table 1: Simulation Parameters

Parameter	Value
Simulator	NS-3 Version 3.38
Number of Nodes	50 – 100
Area Size	1000 m × 1000 m
Mobility Model	Random Waypoint Mobility Model
MAC Protocol	IEEE 802.11
Energy Model	Basic Energy Source + Wifi Radio Energy Model
Routing Protocol	DRL-Based Optimized Multi-Path Routing (DRL-OMRNS)
Traffic Type	CBR over UDP
Simulation Time	200 seconds
Packet Size	512 bytes
Packet Interval	0.2 s
Transmission Range	250

RESULTS AND DISCUSSION

The performance of the suggested DRL-OMRNS approach is examined and compared with existing MANET routing protocols like AODV, DSR, and AOMDV using key network performance measures including throughput, PDR, and end-to-end delay.

Packet Delivery Ratio (PDR)

PDR is the ratio of all packets sent by the source nodes to all packets successfully received at the destination nodes.

$$\text{PDR (\%)} = \frac{\sum_{i=1}^N P_{\text{recv}}^{(i)}}{\sum_{i=1}^N P_{\text{send}}^{(i)}} \times 100 \dots \dots \dots (1)$$

Where

- $P_{\text{recv}}^{(i)}$ = Number of packets successfully received by destination node (i)
- $P_{\text{send}}^{(i)}$ = Number of packets transmitted by source node (i)
- N = Total number of source-destination communication flows

Table 2: Packet Delivery Ratio (%)

Number of Packets Sent	OLSR	AODV	DSR	CAMRNS	DRL-OMRNS
100	89.5	91.0	90.2	96.1	98.2
200	87.8	89.4	88.5	95.3	97.6

300	86.2	88.0	86.7	94.5	96.9
400	84.9	86.3	85.2	93.7	96.3
500	83.5	85.1	84.0	93.0	95.7
600	82.1	83.7	82.8	92.3	95.1
700	80.8	82.3	81.6	91.6	94.5
800	79.3	80.9	80.4	90.8	93.8
900	77.9	79.5	79.1	90.0	93.2
1000	76.5	78.2	77.8	89.3	92.6

In table 2, the PDR performance of the suggested DRL-OMRNS protocol is assessed for packet transmission rates between 100 and 1000 packets and contrasted with OLSR, AODV, DSR, and CAMRNS. In dynamic MANET contexts, PDR is a crucial indicator of packet forwarding capabilities, communication stability, and routing reliability. Reduced packet loss, better route stability, and enhanced network performance are all indicated by a higher PDR. According to the simulation results, DRL-OMRNS regularly outperforms all other procedures in terms of PDR. Thanks to its clever DRL-based relay node selection process, DRL-OMRNS achieves a PDR of 98.2% with a lower traffic load of 100 packets.

The DRL agent dynamically selects stable and energy-efficient relay nodes by considering residual energy, link quality, congestion level, node mobility, and path stability. DRL-OMRNS maintains a high PDR of 92.6% when the traffic load reaches 1000 packets, while CAMRNS, AODV, DSR, and OLSR achieve 89.3%, 78.2%, 77.8%, and 76.5%, respectively. Through adaptive relay selection, effective congestion management, and optimized multi-path routing, DRL-OMRNS exhibits superior routing reliability and packet delivery performance, making it highly suitable for dynamic MANET environments, even though the PDR of all protocols decreases with increasing network load.

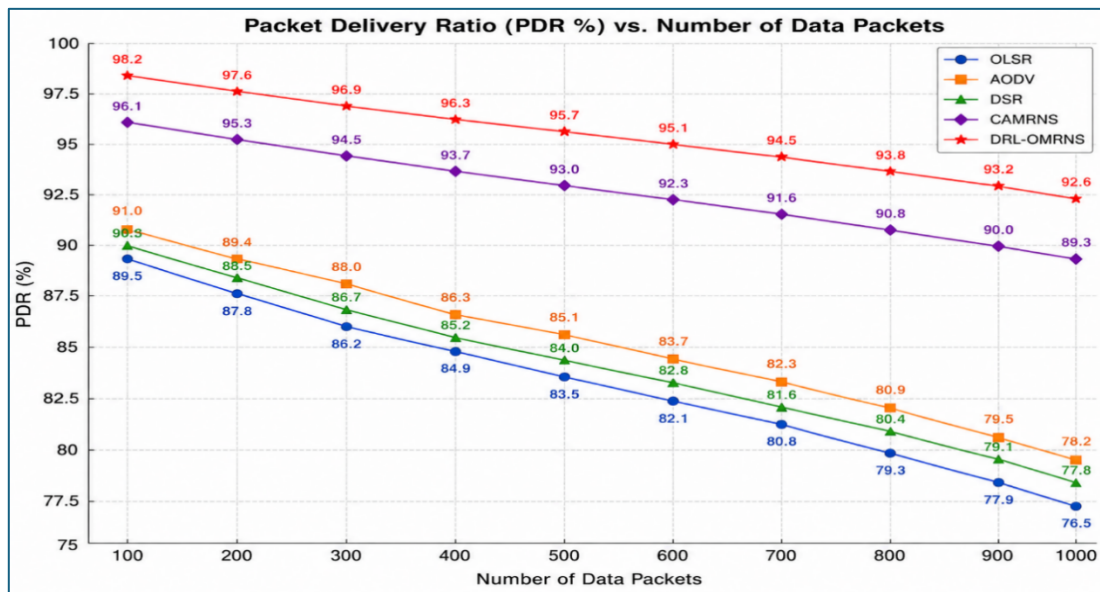


Figure 2: Packet Delivery Ratio Vs Number of data packets

The graph in Figure 2 compares the PDR results of OLSR, AODV, DSR, CAMRNS, and the proposed DRL-OMRNS protocol for packet loads ranging from 100 to 1000. Due to increased congestion, routing overhead, and link failures, the PDR of all protocols eventually drops as the number of transmitted packets rises. Nonetheless, DRL-OMRNS exhibits improved routing reliability by continuously achieving the highest PDR under all traffic scenarios. The protocol outperforms all current techniques, achieving a PDR of 98.2% at 100 packets and maintaining a high PDR of 92.6% even at 1000

packets. Optimized multi-path routing and adaptive relay node selection based on deep reinforcement learning are responsible for its exceptional performance. As a result, in dynamic MANET situations, DRL-OMRNS improves packet delivery efficiency, overall network performance and communication stability.

Throughput

Throughput, which is commonly expressed in kbps, is the average successful data delivery rate in the communication network.

Throughput

$$(\text{kbps}) = \frac{\sum_{i=1}^N D_{\text{recv}}^{(i)} \times 8}{T_{\text{end}} - T_{\text{start}}} \quad \dots \dots \dots (2)$$

where

- $D_{\text{recv}}^{(i)}$ = Total amount of data successfully received by destination node (i) in bytes
- $T_{\text{end}} - T_{\text{start}}$ = Total simulation time duration in seconds

Throughput is a critical performance metric in MANETs as it indicates the efficiency of packet delivery, routing reliability, and overall network performance. Table 3 shows the throughput performance of the suggested DRL-OMRNS protocol for packet transmission loads between 100 and 1000 packets, and it is contrasted with OLSR, AODV, DSR, and CAMRNS. The simulation results show that throughput increases with the number of transmitted packets for all routing protocols; however, the suggested

DRL-OMRNS consistently obtains the maximum throughput under all traffic scenarios. At a lower network load of 100 packets, DRL-OMRNS achieves approximately 156 kbps throughput due to its intelligent relay node selection and adaptive routing optimization. DRL-OMRNS achieves 286 kbps when the packet load hits 1000 packets, while CAMRNS, AODV, DSR, and OLSR obtain 263 kbps, 225 kbps, 218 kbps, and 202 kbps, respectively.

The deep reinforcement learning-based adaptive relay node selection, continuous monitoring of network conditions, and improved multi-path routing strategy of DRL-OMRNS are responsible for its exceptional performance. In highly dynamic MANET circumstances, these capabilities lead to increased scalability, throughput, and communication efficiency through effective traffic distribution, decreased congestion, greater route stability, and reliable packet forwarding.

Table 3: Throughput (kbps)

Number of Packets Sent	OLSR	AODV	DSR	CAMRNS	DRL-OMRNS
100	122	129	125	144	156
200	135	142	138	158	172
300	145	155	150	173	188
400	155	168	162	186	202
500	165	180	172	200	217
600	172	190	183	213	231
700	180	200	192	226	245
800	188	208	201	239	259
900	195	217	210	251	272
1000	202	225	218	263	286

The throughput performance of OLSR, AODV, DSR, CAMRNS, and the suggested DRL-OMRNS protocol for packet loads between 100 and 1000 is depicted in the graph in Figure 3. All protocols experience a steady increase in throughput as the quantity of packets delivered rises. DRL-OMRNS reliably delivers the maximum throughput under

all traffic conditions. DRL-OMRNS outperforms OLSR, AODV, DSR, and CAMRNS at 100 packets, recording 156 kbps. DRL-OMRNS reaches 286 kbps, CAMRNS reaches 263 kbps, AODV, DSR, and OLSR reach 225 kbps, 218 kbps, and 202 kbps, respectively, when the packet load hits 1000.

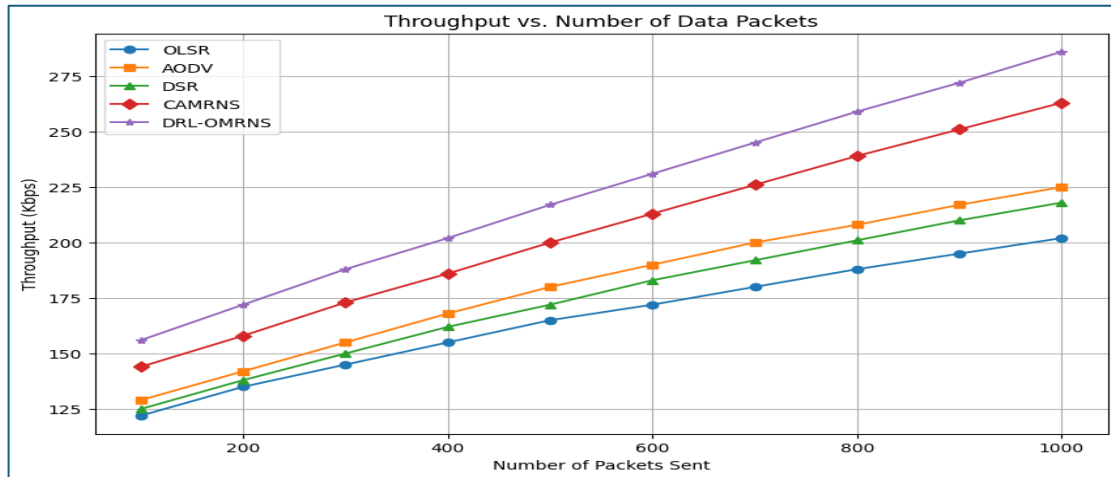


Figure 3: Throughput Vs Number of Packets

The deep reinforcement learning-based relay node selection technique is responsible for the enhanced performance. Network traffic is effectively distributed and congestion is decreased by the optimized multi-path routing technique. These characteristics improve resource usage, packet forwarding effectiveness, and route stability. As a result, in dynamic MANET situations, DRL-OMRNS offers excellent throughput, scalability, and dependable communication.

End-to-End Delay

It represents the average time a data packet takes to move from the source node to the destination node in a MANET is known as end-to-end

delay. Route discovery, transmission, propagation, processing, queuing, and retransmission delays are only a few of the delays that can occur during communication.

$$\text{End-to-End Delay} = \frac{\sum_{i=1}^N (t_{r,i} - t_{s,i})}{N} \dots \dots \dots (3)$$

where

- N = Total number of successfully received packets
- $t_{s,i}$ = Timestamp when packet (i) was transmitted from the source, and
- $t_{r,i}$ = Timestamp when packet (i) was received at destination

Table 4: End-to-End Delay vs. Number of Data Packets

Packets Sent	OLSR	AODV	DSR	CAMRNS	DRL-OMRNS
100	190	175	182	142	128
200	197	183	191	148	133
300	205	191	200	151	137
400	210	198	207	155	141
500	217	205	213	158	145
600	222	211	219	162	149
700	228	218	225	166	153
800	233	224	230	169	156
900	239	231	236	172	159
1000	245	238	243	174	162

The End-to-End Delay performance of five routing protocols OLSR, AODV, DSR, CAMRNS, and DRL-OMRNS under various packet transmission loads ranging from 100 to 1000 packets is shown in the above table 4. Due to growing network traffic, routing overhead, and channel contention, the end-to-end delay steadily increases for all protocols as the number of packets sent grows. DRL-OMRNS consistently obtains the lowest latency among the compared approaches, rising only from 128 ms at 100

packets to 162 ms at 1000 packets, demonstrating its capacity to effectively choose the best relay nodes and maintain stable routing pathways.

OLSR has the most delay, rising from 190 ms to 245 ms, while CAMRNS performs the second best, with delays ranging from 142 ms to 174 ms. AODV and DSR, which reach 238 and 243 milliseconds, respectively, at 1000 packets, likewise exhibit mediocre performance. The findings show that by using deep reinforcement

learning for intelligent route optimization, congestion avoidance, and adaptive multi-path relay selection in dynamic MANET environments, the suggested DRL-OMRNS protocol considerably lowers packet delivery latency when compared to conventional and current routing techniques.

The relationship between the number of packets sent and the end-to-end delay for five routing protocols OLSR, AODV, DSR, CAMRNS, and DRL-OMRNS is illustrated in the graph in Figure 4. DRL-OMRNS consistently obtains the lowest

latency among the compared systems, increasing slightly from roughly 128 ms to 162 ms. This shows that it can effectively choose relay nodes and maintain stable routing paths under a range of traffic situations. While OLSR has the most delay during the experiment, closely followed by DSR and AODV, CAMRNS also outperforms the conventional methods. In comparison to current routing methods, the results demonstrate how well the suggested DRL-OMRNS solution reduces packet delivery latency.

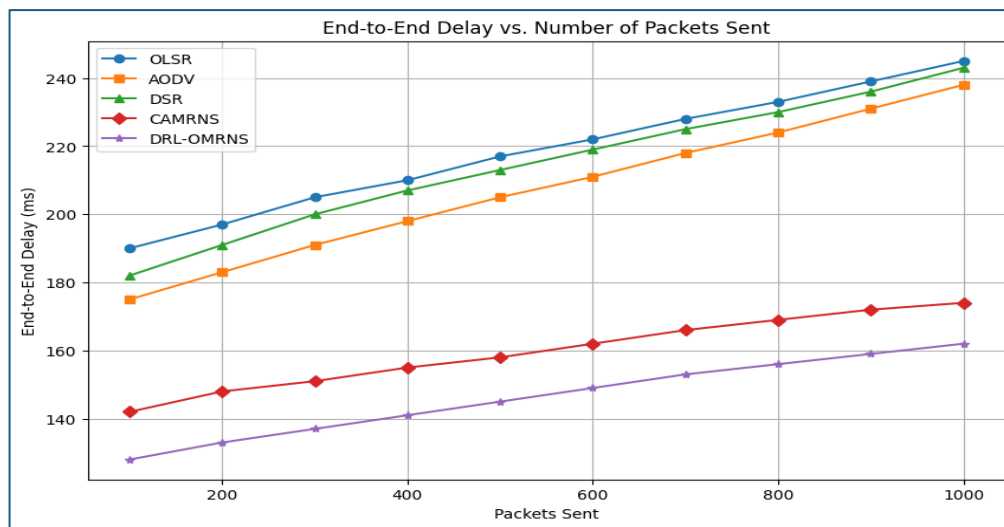


Figure 4: End-to-Delay Vs Number of Data Packets

DRL-OMRNS reduces congestion, increases route stability, and improves communication efficiency in dynamic MANET environments by employing deep reinforcement learning for adaptive multi-path relay selection. DRL-OMRNS scales better under higher network loads and is more appropriate for delay-sensitive applications where timely data transmission is crucial, as seen by the comparatively lesser rise in delay as traffic increases.

CONCLUSION

The DRL-OMRNS method is suggested in this study to enhance reliable and energy-efficient MANET routing. In highly dynamic network situations, the framework improves routing stability, communication reliability, congestion control, and energy efficiency by combining DRL with adaptive multi-path relay node selection. The proposed approach intelligently selects the best relay nodes and dynamically changes the routing choices based on current network conditions by taking into account crucial network elements such as residual energy, link quality, node mobility, congestion level, queue status, and path stability.

According to simulation results, the suggested DRL-OMRNS method performs better in terms of End-to-End Delay, collision reduction, throughput, routing overhead, PDR and network lifetime than traditional routing protocols like AODV, DSR, and AOMDV. While the multi-path routing technique enhances load balancing, congestion avoidance, and fault tolerance, the DRL mechanism allows for effective relay node selection, adaptive route maintenance, and quick recovery from link failures. The framework is appropriate for applications including military communication, disaster recovery, vehicular networks, healthcare monitoring, and IoT-enabled environments because it also shows improved scalability, adaptability, and energy utilization under high mobility and dense traffic scenarios. Security-aware routing, advanced DRL models, mobility prediction and federated learning approaches for next-generation wireless networks are potential future improvements.

ACKNOWLEDGEMENT

We acknowledge the support of the & Research Department of Computer Science, Jamal Mohamed College (Autonomous), Trichy for providing computational resources. This research

was not supported by any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Source of support: None

Conflict of interest: None

REFERENCES

- Abolhasan, M., Wysocki, T., & Dutkiewicz, E. (2004). A review of routing protocols for mobile ad hoc networks. *Ad Hoc Networks*, 2(1), 1–22. [https://doi.org/10.1016/S1570-8705\(03\)00043-X](https://doi.org/10.1016/S1570-8705(03)00043-X)
- Alwarafy, A., Abdallah, M., Ciftler, B. S., Al-Fuqaha, A., & Guizani, M. (2021). Deep reinforcement learning for IoT and wireless network: A comprehensive survey. *IEEE Communications Surveys & Tutorials*, 23(3), 1659–1695. <https://doi.org/10.1109/COMST.2021.3070338>
- Broch, J., Maltz, D. A., Johnson, D. B., Hu, Y. C., & Jetcheva, J. (1998, October). A performance comparison of multi-hop wireless ad hoc network routing protocols. In *Proceedings of the 4th Annual ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom '98)* (pp. 85–97). Dallas, TX, USA. <https://doi.org/10.1145/288235.288256>
- Clausen, T., & Jacquet, P. (2003). *Optimized link state routing protocol (OLSR)* (RFC 3626). Internet Engineering Task Force. <https://doi.org/10.17487/RFC3626>
- François-Lavet, V., Henderson, P., Islam, R., Bellemare, M. G., & Pineau, J. (2018). An introduction to deep reinforcement learning. *Foundations and Trends® in Machine Learning*, 11(3–4), 219–354. <https://doi.org/10.1561/22000000071>
- Giordano, S. (2002). Mobile ad hoc networks. In I. Stojmenovic (Ed.), *Handbook of wireless networks and mobile computing* (pp. 325–346). Wiley. <https://doi.org/10.1002/0471224561.ch13>
- Gwon, Y., Kempf, J., & Sarma, A. D. (2004, March). Scalable mobility prediction for routing in mobile ad hoc networks. In *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC 2004)* (pp. 1914–1919). Atlanta, GA, USA. <https://doi.org/10.1109/WCNC.2004.1311754>
- Haas, Z. J., Pearlman, M. R., & Samar, P. (2002). *The zone routing protocol (ZRP) for ad hoc networks* (Internet-Draft). Internet Engineering Task Force (IETF) MANET Working Group.
- Jacquet, P., Mühlenthaler, P., Qayyum, A., Laouti, A., Clausen, T., & Viennot, L. (2001, December). Optimized link state routing protocol for ad hoc networks. In *Proceedings of the IEEE International Multi Topic Conference (INMIC 2001)* (pp. 62–68). Lahore, Pakistan. <https://doi.org/10.1109/INMIC.2001.995315>
- Johnson, D. B., Maltz, D. A., & Hu, Y. C. (2007). *The dynamic source routing protocol (DSR) for mobile ad hoc networks for IPv4* (RFC 4728). Internet Engineering Task Force. <https://doi.org/10.17487/RFC4728>
- Li, X., Xu, Y., & Zhang, H. (2020). Q-learning-based routing protocol for mobile ad hoc network. *IEEE Access*, 8, 116727–116738. <https://doi.org/10.1109/ACCESS.2020.3004647>
- Lillicrap, T. P., Hunt, J. J., Pritzel, A., Heess, N., Erez, T., Tassa, Y., Silver, D., & Wierstra, D. (2015). *Continuous control with deep reinforcement learning*. arXiv. <https://doi.org/10.48550/arXiv.1509.02971>
- Luong, N. C., Hoang, D. T., Gong, S., Niyato, D., Wang, P., Liang, Y. C., & Kim, D. I. (2019). Applications of deep reinforcement learning in communications and networking. *IEEE Communications Surveys & Tutorials*, 21(4), 3133–3174. <https://doi.org/10.1109/COMST.2019.2916583>
- Marina, M. K., & Das, S. R. (2001, November). On-demand multipath distance vector routing in ad hoc networks. In *Proceedings of the 9th IEEE International Conference on Network Protocols (ICNP 2001)* (pp. 14–23). Riverside, CA, USA. <https://doi.org/10.1109/ICNP.2001.992756>
- Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, I., King, H., Kumaran, D., Wierstra, D., Legg, S., & Hassabis, D. (2015). Human-level control through deep reinforcement learning. *Nature*, 518(7540), 529–533. <https://doi.org/10.1038/nature14236>
- Nasipuri, A., Castaneda, R., & Das, S. R. (2001). Performance of multipath routing for on-demand protocols in mobile ad hoc networks. *Mobile Networks and Applications*, 6(4), 339–349. <https://doi.org/10.1023/A:1011479229801>
- Park, V. D., & Corson, M. S. (2001). *Temporally-ordered routing algorithm (TORA) version 1 functional specification* (Internet-Draft). Internet Engineering Task Force (IETF) MANET Working Group.
- Perkins, C., Belding-Royer, E., & Das, S. (2003). *Ad hoc on-demand distance vector (AODV) routing* (RFC 3561). Internet Engineering Task Force. <https://doi.org/10.17487/RFC3561>

-
- Ramanathan, R., & Steenstrup, M. (1998). Hierarchically-organized, multihop mobile wireless networks for quality-of-service support. *Mobile Networks and Applications*, 3(1), 101–119. <https://doi.org/10.1023/A:1019108810885>.
- Roy, S., Dey, A., & Chaki, R. (2020). A comprehensive survey on MANET. *Wireless Personal Communications*, 114(4), 2673–2702. <https://doi.org/10.1007/s11277-020-07427-1>.
- Royer, E. M., & Toh, C. K. (1999). A review of current routing protocols for ad hoc mobile wireless networks. *IEEE Personal Communications*, 6(2), 46–55. <https://doi.org/10.1109/98.760423>.
- Sun, Y., Peng, M., Zhou, Y., Huang, Y., & Mao, S. (2019). Application of machine learning in wireless network: Key techniques and open issues. *IEEE Communications Surveys & Tutorials*, 21(4), 3072–3108. <https://doi.org/10.1109/COMST.2019.2924241>
- Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction* (2nd ed.). MIT Press.
- Vaisakh, K. S., & Venkatesan, R. (2021). Energy-efficient multipath routing protocol for mobile ad hoc network. *Wireless Networks*, 27(3), 1867–1881.
- Ye, H., Li, G. Y., & Juang, B. H. (2019). Deep reinforcement learning based resource allocation for V2V communications. *IEEE Transactions on Vehicular Technology*, 68(4), 3163–3173. <https://doi.org/10.1109/TVT.2019.2896837>
-