International Journal of Computer Networks and Applications (IJCNA)
DOI: 10.22247/ijcna/2025/06 Volume 12, Issue 1, January – February (2025)

RESEARCH ARTICLE

Cooperative Device-to-Device Communication Using Joint Relay Assignment and Channel Allocation Using Deep Learning

Md. Tabrej Khan
Department Faculty of Computer Science, Pacific Academy of Higher Education and Research University, Udaipur (Rajasthan), India.
tabrejmlkhan@gmail.com

Ashish Adholiya
Department Faculty of Computer Science, Pacific Academy of Higher Education and Research University, Udaipur (Rajasthan), India.
asia_1983@rediffmail.com

Received: 20 November 2024 / Revised: 18 January 2025 / Accepted: 06 February 2025 / Published: 28 February 2025

Abstract – Fusion centers support sensing and signal processing in decentralized mobile user communications. The backend network is made up of device-to-device (D2D) connections, which use appropriate channel selection to guarantee smooth connectivity. Delays and decreased dependability result from an imbalance in the distribution of channels. In cognitive radio networks, the behavior of primary users affects stability, whereas relay communication maximizes resources. In this regard, a joint channel selection and routing protocol, is proposed in this research based on deep reinforcement learning. The goal of this research is to minimize interference and optimize network performance by developing a deep reinforcement learning-based joint channel allocation and relay selection framework for D2D communication. Initially, the channel allocation technique is proposed using the enhanced hunter prey optimization (EnHpo) algorithm. The adaptive weighting method is integrated with the traditional hunter-prey optimization in the design of the proposed EnHpo to improve the convergence rate and produce the global optimum solution with balanced local search and randomization phases. Here, the multi-objective fitness function based on factors like priority, bandwidth and transmission rate are considered for the optimal channel allocation. Followed by, the relay selection is devised using the deep reinforcement learning criteria based on the channel gain based on the bit error rate. Here, the relay sub-set selection in the using the deep reinforcement learning improves D2D communication efficiency. The performance evaluation of the proposed joint channel allocation and relay selection mechanism in terms of Average Residual Energy, Latency, Network Life Time, Packet Delivery Ratio, and Throughput acquired the values of 0.998, 2.709, 99.592, 0.999, and 23015 respectively. The maximum throughput estimated by the proposed method is 23015, which is 54.73%, 41.63%, 29.98%, and 8.00% superior outcome compared to the conventional DDPG Approach, Zigbee/WiFi Routing, Decode and Forward method, and Game-based Framework methods with 100 nodes.

Index Terms – Deep Reinforcement Learning, Channel Allocation, Joint Optimization, Device to Device Communication, Cooperative Networks, Enhanced Hunter Prey Optimization, Residual Energy.

1. INTRODUCTION

The current generation of mobile communications is finding it challenging due to the quick expansion of smart devices and multimedia services. to keep up with the increasing demand for traffic. In cellular networks, new methods must be created to improve energy and spectrum efficiency and lower transmission latency [1-3]. One of the most promising approaches among the many suggested technologies to satisfy the growing demands is D2D communication. With D2D communications, two mobile users (MUs) can send data directly to each other utilizing the same frequency resources in a cellular network, whether or not the base station (BS) is under control [4-5]. In general, there are two kinds of D2D communications that offer D2D links while maintaining cellular link performance: underlay D2D communications, which share spectrum bands with cellular links, and overlay D2D communications, which employ specialized channels or time slots [6-8].

Cooperative D2D communication has drawn a lot of interest lately as a way to improve network performance, with cooperative relay across D2D lines being its key component [4]. When there is data to send to a base station (BS) or another MU (also called a target node) but the direct link has poor channel conditions, a MU (also called a source node) can rely on other MUs (also called relay nodes) with D2D communication capabilities to help forward the data to the target node. If the data is sent across multiple MUs or the

RESEARCH ARTICLE

target node is a MU, this communication type is sometimes referred to as multi-hop D2D [9-10]. A novel approach to D2D communications is offered by cooperative D2D communications. Handovers must be carried out more precisely and effectively, nevertheless, because of the MUs' erratic mobility and the D2D links' limited coverage. Thus, the two main problems listed below emerge [11].

An essential phase in handovers is the evaluation of network service capabilities. Relay nodes and BSs typically have varying service capabilities because to differences in data processing capacity, traffic loads, etc [12-15]. MUs can only choose the best BS or cooperative peer MU(s) after a thorough assessment of the service capabilities. Even though it has been extensively researched for cellular networks, further investigation is still needed to determine how to use D2D communication capabilities and MU collaboration to increase evaluation accuracy [16-20]. Mode and peer selection: There are two fundamental ways that cooperative D2D communications operate: cooperative relay via peer MUs and direct transmission to the destination node. Based on the network's current status, including the source node's service demand, MU locations, and channel conditions, the source node must select one of the two modes. A peer MU must be chosen by the source node as its relay node before data can be sent cooperatively to the destination node [21-27].

In order to address these issues, this research proposes a cooperative channel allocation and routing as a realistic way to sufficiently reduce radio interference from the networks. In this research a unique method that combines deep reinforcement learning with an EnHpo algorithm to integrate relay selection and joint channel allocation. With the help of this innovative technique, network performance maximized and total system efficiency raised by efficiently controlling channel access and relay communication. By taking into account important variables including priority, bandwidth, and transmission rate using a multi-objective fitness function, the proposed approach aims to produce better results in terms of energy efficiency, packet delivery ratio, and network reliability.

The major contributions of the research are:

- Design of Optimal Channel Allocation: The optimal channel allocation is designed using the enhanced hunter prey optimization (EnHpo) by taking into account the priority-based, bandwidth-based multi-objective fitness function and transmission rate.
- Design of joint channel allocation and relay selection for D2D communication: DRL is used to create the joint channel allocation and relay selection, with the chosen channel gain serving as the basis for the relay selection.

The organization of the research is: The related works along with the problem statement is detailed in Section 2 and proposed methodology is detailed in Section 3 and its experimental outcome is presented in Section 4. Finally, Section 5 concludes the research with future scope.

2. LITERATURE REVIEW

Some of the related works based on the relay based routing techniques are reviewed in this section. In [28], a multi-hop routing system based on Stackelberg game theory was created for cooperative networking, employing machine learning to allocate resources as efficiently as possible. The optimization problems were addressed with a feed-forward neural network. The model performed well in real-world applications by taking relay utilities and throughput into account, enhancing results through the most efficient use of power and cost. However, its effectiveness is constrained by coalitions that overlap in the distribution of resources.

Computational resources, task division, and bandwidth are all taken into account in [29]'s adaptive bandwidth allocation and relay selection technique. By employing an evolutionary strategy to solve transmission latency, the model performs better in terms of task completion time. Its disregard for energy efficiency, however, shortens the network's lifespan.

In [30], an ideal relay selection method takes into account variables including enhanced SINR, channel gain, and minimum distance. It uses bitwise XOR encoding to increase throughput and mode switching to reduce traffic. By using a decode-and-forward approach, the model decreases co-channel interference and improves coverage probability. Its performance is, however, constrained by interference from user interaction dynamics.

Using a Markov decision process (MDP) to choose a communication mode and allocate resources, the D2D communication model in [31] concentrates on energy-efficient communication. The MDP is solved using continuous state and action spaces and a DDPG algorithm. For deterministic actions, the model uses an actor network; for performance evaluation, it uses a critic network. However, its performance is limited by bandwidth limitations.

Instead of calculating the allocation for every channel realization, a DL framework presented in [32] approximates the best way to allocate resources in a variable channel. using deep neural networks (DNNs). In order to attain near-optimal performance while controlling integer optimization variables and guaranteeing the QoS requirements of cellular users, It combines a local CSI sharing mechanism with supervised and unsupervised learning.

In [33], the hybrid flow direction with the chameleon swarm algorithm (HFDCSA) is used to minimize link rates in D2D and cellular networks. The chameleon swarm method and the flow direction algorithm (FDA) are combined in this optimization scheme to pick relays optimally and allocate

RESEARCH ARTICLE

resources together. The adaptive multi-layer perceptron (AMLP) simplifies relay selection and resource allocation, and, to guarantee exceptional performance, the model takes energy efficiency and sum rate constraints into account. A hybrid centralized-distributed system that uses the Kuhn-Munkres (KM) algorithm with deep reinforcement learning (DRL) have been suggested by [34]. Using only local data, the CUs and DUs use the former to optimize power control and spectrum allocation independently. The BS then determines the connection matching using the latter.

The load-based dynamic channel allocation (LB-DCA) model, which was presented in [35], combines channel load balancing, interference estimation, and control systems to enhance D2D communication in WPAN. By estimating active nodes and channel utilization through distributed coordination and a cell-splitting approach, it strives for low energy consumption, high throughput, and little interference. Table 1 shows the Summary of recent research work.

Table 1 Summary of Recent Research Work

Author	Techniques	Findings	Limitations
KUMAR R, SINGH H [28]	Artificial Neural Network (ANN)	Optimal power distribution in cooperative multi-relay situations.	Not scalable
Imtiaz HH, Tang S [29]	Partial Offloading with Relay and Adaptive Bandwidth Allocation (PORAB)	Improved task offloading in IoT contexts with MEC support.	Demands a balance between the efficiency of local computation and offloading.
Sarma SS, et al. [30]	Dynamic relay selection (DRS)	Enhanced D2D communication efficiency in 5G millimeter-wave networks.	Significant power usage as a result of mmWave technology.
Zhang T, et al. [31]	Deep deterministic policy gradient (DDPG)	Mode selection and resource allocation for D2D were accomplished in an energy-efficient manner.	Computatiol complexity.
Lee, W. and Schober, R., [32]	Deep neural network (DNN)	Suggested a DL-based strategy for D2D communication resource optimization.	Restricted to particular D2D situations and resource limitations.
Chennaboin, R.B. and Nandakumar, S., [33]	HFDCSA	For D2D, resource allocation and relay selection are optimized together.	Possible inefficiency in situations those are quite dynamic.
Yu, Y. and Tang, X., [34]	Deep reinforcement learning (DRL)	Resource allocation that is optimized for collaborative D2D communication.	Reliance on the quality of training data and on performance in real time.
Logeshwaran J, et al. [35]	LB-DCA	Suggested a concept for dynamic channel allocation to improve WPAN performance.	If there are more users or larger networks, it does not scale well.

2.1. Problem Statement

Applications of the CRNs include femtocells, cooperative networks, smart grid communications, dynamic spectrum access, public safety systems, and intelligent transportation systems. Researchers have developed a number of strategies for effective D2D communication among CRN, but the model's functionality is still constrained by a number of difficult issues. High packet loss, high latency, and high

International Journal of Computer Networks and Applications (IJCNA)
DOI: 10.22247/ijcna/2025/06 Volume 12, Issue 1, January – February (2025)

resource consumption were discovered to be the constraints of the conventional methods. The dynamic nature of the nodes and lack of a route recovery method are caused the packet loss. Due to the additional routing pathways detection, the network also experiences a significant delay. Therefore, this research presents a method for effective D2D communication that combines relay selection and channel allocation.

3. PROPOSED METHODOLOGY FOR JOINT CHANNEL ALLOCATION AND RELAY SELECTION TECHNIQUE

The joint channel allocation and relay selection of the cooperative network utilizes the orthogonal frequency division multiplexing channel (OFDMA) cognitive model with source destination pairs along with multiple relays. Here, the power utilized for the transmission is fixed, wherein the device selects the channel in the idle state for making the communication between the devices more effectively. Besides, each node holds a antenna and the channel gain is considered constant for each time slot. The resource allocation and the controlling operations are controlled by the base station. Initially, during the first hop of communication, the factors like bandwidth, priority and transmission rate are considered for the reduction of latency in the network. Followed by, in the second hop, the bit error-based relay node selection is devised for meeting the QoS requirements. The illustration of the two-hop joint channel allocation and the relay selection technique is depicted in Figure 1.

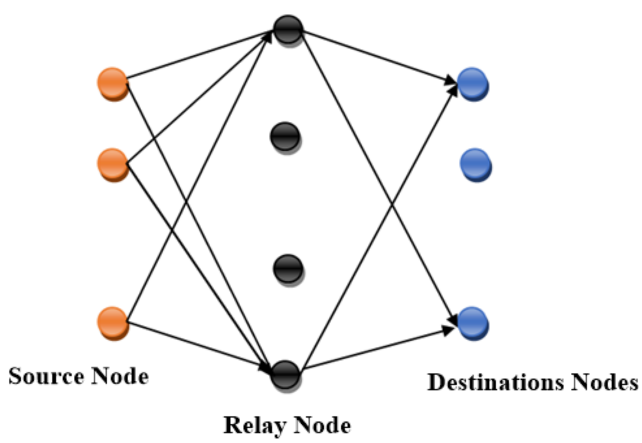


Figure 1 System Model for the Proposed Joint Channel Allocation and Relay Selection Technique

3.1. Multi-Objective Fitness Function

Priority, bandwidth, and transmission rate are some of the criteria that are taken into consideration when designing the multi-objective fitness function for D2D communication channel allocation. As an example, let's look at the channels that are used for the proposed joint channel allocation and relay selection method.

Priority: For the incoming request, the priority is estimated based on the packet loss. The higher priority is assigned to the node that has the faster packet loss and is estimated in equation (1).

$$A_a = \frac{1}{\min\left\{\frac{D_a - B_a}{D_b}, \frac{C_a - E_a}{S_a}\right\}} \tag{1}$$

where, the device is indicated as a, the tolerable delay is notated as D_a, the packets arrives at the node is represented asS_a, and the present size of the data is represented as E_a. The buffer size is indicated as C_a, the frames duration is indicated as D_b, and the delay associated with the node is indicated as B_a.

Bandwidth: Let us consider the bandwidth required for the node for communication be indicated as B_{G_a}. For the better channel allocation, the device chooses the channel with bandwidth higher than $B_{G_a}$and hence the condition utilized for the bandwidth requirement is expressed in equation (2).

$$A = \{A_a | B_a > B_{G_a}\} \tag{2}$$

Transmission rate: The data size is used to estimate the transmission duration, and the greatest transmission rate is taken into account for optimal channel allocation. The estimation of the transmission rate is shown in equation (3).

$$a = \arg\max\left\{g_{a,e} \left| \frac{f_a}{g_{a,e}} \leq T_a \right.\right\} \tag{3}$$

Where, the available time is indicated as T_a, and transmission rate of device ais indicated as $g_{a,e}$.

Then, the multi-objective fitness function for the channel allocation is formulated in equation (4).

$$Fit = \{\max(\text{priority}, \text{bandwidth}, \text{transmissionrate})\} \tag{4}$$

3.2. Optimal Channel Allocation Using EnHpo

The proposed Enhanced hunter prey optimization (EnHpo) algorithm is created by combining traditional hunter-prey optimization techniques using an adaptive weighting technique to accelerate the pace of convergence. The hunting behaviour animals for capturing the prey is considered in the hunter prey optimization in order to resolve the optimization problems. Here, the prey considered in the optimization are gazelle, stag, and deer, while the hunters like wolves, leopards and lions follow the hunting strategy utilized in the proposed algorithm. In order to get the best solution globally without becoming stuck at the local optimal solution, the optimal algorithm is one that combines balanced randomization with local search capabilities. The conventional hunter prey optimization has the capability of pre-mature convergence, which is limited by incorporating the adaptive weighting strategy for improving the randomization standards in order to get the best solution globally.

3.3. Mathematical Modelling of EnHpo

The search agents are arbitrarily placed in the search space and then, the multi-objective based fitness is estimated for all the hunters to identify the feasibility of the solution. Here, the localization of the hunters is represented as $\left(\vec{P}\right) = \left\{\vec{P}_1, \vec{P}_2, \ldots\ldots \vec{P}_R\right\}$. Also, the maximal iterations considered for the algorithm is initialized as τ^{max}.The solution accomplished by the hunter in the randomization phase during the arbitrary phase is represented in equation (5).

$$P_k = m(1, L) * (Max_r - Min_r) + Min_r \quad (5)$$

Where, Min_r and Max_rrefers to the minimum and maximum dimension of the solution and the position of the hunter is indicated asP_k and$L$refers to the variables. The mathematical modelling of the minimum and maximum dimension of the solution is written as in equation (6).

$$Min_r = [Min_1, Min_2, \ldots. Min_L] \quad (6)$$

$$Max_r = [Max_1, Max_2, \ldots. Max_L] \quad (7)$$

Fitness Evaluation: The multi-objective function, which is represented in equation (4), is used to estimate fitness.

Randomization: The promising locations are explored by the hunters to find the global best solution for optimal channel allocation. The solution accomplished by the hunters in the randomization phase is updated as:

$$P_{k,h}(\tau + 1) = P_{k,h}(\tau) + 0.5 \begin{bmatrix} \left(2HW \cdot N_{v(h)} - P_{k,h}(\tau)\right) + \\ \left(2(1 - H)W \cdot M_h - P_{k,h}(\tau)\right) \end{bmatrix} \quad (8)$$

Where, the position updated by the hunters at $(\tau + 1)^{th}$ iteration is notated as $P_{k,h}(\tau + 1)$ and τ^{th} iteration is notated as$P_{k,h}(\tau)$. The mean of the solutions acquired by the hunters in the present iteration is indicated as M_h and the adaptive parameter is notated as$W$. The definitions for the $W$and$M_h$are expressed in equation (9) and (10).

$$N = \vec{Y}_1 < H; u = (N == 0);$$
$$W = Y_2 \otimes u + \vec{Y}_3 \otimes (\sim u) \quad (9)$$

$$M = \frac{1}{n}\sum_{k=1}^{n} \vec{P}_k \quad (10)$$

Where, the factor utilized for balancing the randomization and local search criteria to the acquisition of the global best solution is represented as $H$and the solution considered as target is represented as$N$. The value of index is defined as$u$for $\vec{Y}_1$ that maintains the $(N == 0)$ assumption. The random numbers are mentioned as $Y_1, Y_2$and$Y_3$with the limit of $[0,1]$. Then, the balancing parameter that degrades the value 1to0.02 throughout is formulated automatically in equation (11).

$$H = 1 - \tau\left(\frac{0.98}{\tau_{max}}()\right) \quad (11)$$

Where, the processing iteration is defined as$\tau$and its highest value is indicated asτ_{max}.

By considering the mean of the solution identified by all the individual hunters and the distance between the prey is utilized for evaluating the position of the prey and is outlined in equation (12).

$$\vec{N}_v = \vec{P}_k | k \text{ is index of} Max(End) sort(X) \quad (12)$$

The distance between the mean solution and the prey is measured by evaluating the Euclidean distance and is defined in equation (13).

$$X(k) = \left(\sum_{h=1}^{L}\left(P_{k,h} - M_h\right)^2\right)^{1/2} \quad (13)$$

The position of the prey is updated if the distance measure yields a lower result. When, the outcome is larger, the convergence of the algorithm gets delayed that is eliminated by evaluating the assumption in equation (14).

$$Z = round(H \times q) \quad (14)$$

Where, the distance limiting factor is indicated as $Z$and the hunters population is indicated as$q$. To improve the algorithm's pace of convergence, the distance constraint is gradually reduced from its starting value over the course of the iteration. After incorporating the distance limiting factor, the solution accomplished by the prey is outlined in equation (15).

$$\vec{N}_v = \vec{P}_k | k \text{ is sorted } X(Z) \quad (15)$$

Thus, the solution acquired by the hunters in the randomization phase is defined and it is given in equation (16).

$$P_{k,h}(\tau + 1) = I_{v(h)} + HW\cos(2\pi Y_4) \times \left(I_{v(h)} - P_{k,h}(\tau)\right) \quad (16)$$

The pre-mature convergence of the solution at the randomization phase is eliminated by incorporating the adaptive weighting strategy and is outlined in equation (17).

$$Y = \left(1 - \tau/\tau_{max}()^{1 - \tan(\pi \times (l - 0.5) \times b/\tau_{max}())}\right) \quad (17)$$

Where, $b$is the factor utilized for making the hunters move towards the prey, which is added to the solution accomplished by the hunters while solution updation is not employed. In contrast, the factor bis divided by 2when solution updation is devised by the hunters. Here, the maximal limit of the adaptive weight factor Yis 1 and its minimal value is 0. Thus, after incorporating the adaptive weighting strategy, the position of the hunters is updated using the proposed EnHpo is defined and it is given in equation (18).

$$P_k(\tau+1)_{EnHpo} = Y * P_k(\tau+1) \tag{18}$$

Where, the solution updation by the proposed EnHpo algorithm is indicated as$P_k(\tau+1)_{EnHpo}$.

Local Search: Based on the solution accomplished at the randomization, the local search is devised for obtaining the indepth solution in allocating the channel. Then, the definition for the position updation is outlined in equation (19).

$$P_k(\tau+1) = \begin{cases} P_k(\tau) + 0.5\left[\left(2 \cdot H \cdot W \cdot N_v - P_k(\tau)\right) + \left(2(1-w)W \cdot M - P_k(\tau)\right)\right] if Y_5 < \eta \\ I_v + H \cdot W \cdot cos(2\pi Y_4) \times \left(I_v - P_k(\tau)\right) otherwise \end{cases} \tag{19}$$

Feasibility Evaluation: The feasibility of the solution acquired in the local search phase is evaluated using the fitness estimation devised in equation (4).

Termination: Following the attainment of the finest possible solution or the purchase of τ_{max}. The proposed EnHpo algorithm's pseudo-code is shown in Algorithm 1.

Pseudo-code for EnHpo algorithm

1. The parameters like $H, \tau^{max}$are initialized
2. The dimensions limit$Max_v$and $Min_v$are initialized
3. Evaluate the fitness using equation (4)
4. While
5. {
6. Update the solution in randomization phase using equation (18)
7. Update the solution in local serach phase using equation (19)
8. Re-estimate the feasibility using equation (4)
9. }
10. $I = I++$
11. Return the best solution
12. End

Algorithm 1 Pseudo-Code for EnHpo Algorithm

3.4. DQL Based Relay Selection Using Channel Gain

The proposed multi-hop routing algorithm utilizes the DQL for identifying the best relay for communication between the devices by considering the selected channel along with its probability. The issue concerning the Markov decision control performance for the efficient routing is resolved and enhanced through reinforcement learning.

A learning agent is the fundamental module in the design of reinforcement learning and is capable of perceiving the status of the surrounding environment and acting in a way that has the potential to impact the controlled environment in the reverse direction. The reward signal is defined and directs the agent to acquire greater cumulative values using a trial-and-error procedure in order to enhance relay selection performance. It is designed to solve Markov choice problems using the well-known reinforcement learning algorithm Q-learning (Watkins, 1989). Q-learning is anticipated to increase the overall reward because it is one of the most widely utilized off-policy RL. As a result, the distribution across the given current state and control action may be described as the optimal function value that directs the policy decision-making process. The fundamental principle underlying Deep Reinforcement Learning is displayed in Figure 2.

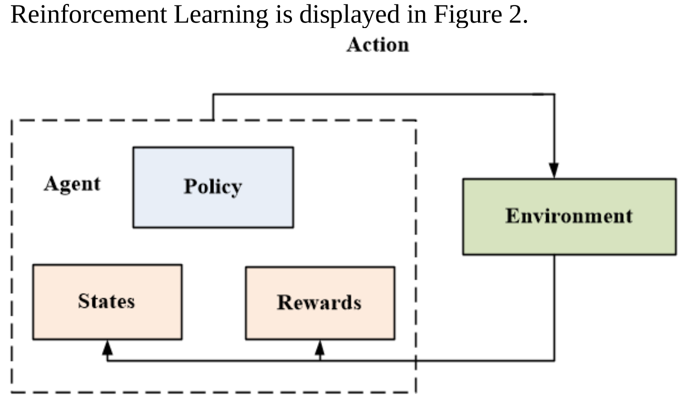


Figure 2 Basics Architecture of Deep Reinforcement Learning

Agent: The solution is referred as the agent, wherein the decision making takes place for solving the issues based on uncertainty. Thus, the environment (problem) gets affected by the agent. The maximization of the reward is the goal of the agent, which is crucial for obtaining the efficient relay selection based on channel allocation.

Action: The possible operations for performing the relay selection are termed as action. From all the possible actions, the agent chooses the best action.

Environment: The issue is referred as the environment, in the proposed method, the relay selection based on the channel gain is considered as an issue. The agent's choice has an impact on the environment in terms of rules, incentives, or conditions.

Policy: Choosing the right course of action to increase the reward is the policy's responsibility.

RESEARCH ARTICLE

States: The set of variables that defines the environment is referred as states.

Rewards: The feedback provided by the environment for the agent's action in each state is defined as the reward.

3.5. Deep Reinforcement Learning

The combined behaviour of deep learning with the Reinforcement learning constitutes the deep reinforcement learning. Deep reinforcement learning yields a variety of actions for the given state$Q$, and the best action is chosen for the relay selection. The network parameter is referred to φ.In the proposed D2D communication protocol, the Deep reinforcement learning is utilized for choosing the relay based on the channel gain of the allocated channel. Figure 3 shows the Deep Reinforcement Learning architecture.

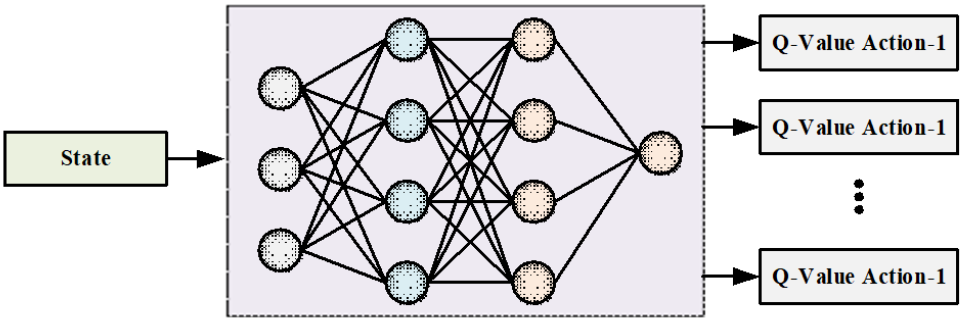


Figure 3 Architecture of Deep Reinforcement Learning

Thus, the decision making regarding the relay selection is devised based on the reward $R^{h}_{Q,Q''}$for the state$Q$. Here, the probability of choosing the relay based on the action state pair is notated as $B^{h}_{Q,Q'}$, whereas the parameter H indicates the action. In the proposed joint channel allocation and relay selection technique, the optimum relay is chosen based on its channel gain for effective D2D communication.

Reward and Q value Evaluation: For every action, the reward is estimated for making the decision, wherein the action with highest reward chooses the best relay based on the channel gain. For the source device m_c, the receiver device m_b is considered for the efficient relay selection. Here, the reward for the action is outlined in equation (20)

$$R^{h}_{Q,Q''} = -p - \alpha_1\left[\left(S_{d,a}\right)_c + \left(S_{d,a}\right)_b\right] + \alpha_2[n(m_c) + n(m_b)] \quad (20)$$

Where, factor that defines the reward for the relay selection is indicated as $R^{h}_{Q,Q''}$, and the corresponding weight factors are defined as$\alpha_1$andα_2. The parameter utilized for the estimation of the punishment factor is defined as $p$and the corresponding action-state pair is notated as(m, f_s).

When the process fails to choose the optimal relay according to the channel gain, then its reward function is outlined and it is given in equation (21).

$$R^{h}_{Q,Q''} = -p \times \eta - \gamma_1\left(S_{d,a}\right)_c + \gamma_2 n(m_c) \quad (21)$$

Here, the channel gain considered for the communication is indicated as $\left(S_{d,a}\right)_c$ and the drop case of the relay selection is indicated asη. Then, the formulation for defining the channel gain based on the bit error rate requirement is outlined in equation (22).

$$K_a = \underset{a\in\{M_1,M_2,\ldots M_A\}}{\operatorname{argmin}} \{S_a\} \quad (22)$$

Where, the set of relay chosen for communicating with the destination is represented as $K_a \in D$, wherein the destination device is represented as a. Here, for the reduction of computational complexity of the model, the relay selection is evaluated based on the channel coefficient and is expressed in equation (23).

$$S_{d,a} = m_e \exp\left(-c_e \frac{q_{d,a}|p_{d,a}|}{\sigma^2}\right) \quad (23)$$

Where, the channel coefficient concerning the d^{th}relay to the destination device a, the bit error rate is indicated as $S_{d,a}$, the power allocation is indicated as $q_{d,a}$, the parameters utilized for modulation and coding is indicated as $c_e$and m_e. Here, for the relay selection based on the channel gain is essential to effective D2D communication. The estimation of the reward is outlined in equation (24).

$$Re\,ward = B_Q \times R^{h}_{m} + B\left(1 - B_Q\right) \times R^{h}_{m} \quad (24)$$

Where, the initial channel gain varies from [0,1].

Q-Value: Q is estimated in order to achieve the maximum reward and is outlined in equation (25).

International Journal of Computer Networks and Applications (IJCNA)
DOI: 10.22247/ijcna/2025/06 Volume 12, Issue 1, January – February (2025)

RESEARCH ARTICLE

$$Q - V(Q, h) = \mathrm{Re\,w\,ard} + \beta\left[Q - V(Q, h) + \mathrm{Max}_{h'}\left(Q - V(Q', :)\right)\right] \quad (25)$$

To choose the best relay for D2D communication, $Q - V$ is very useful to estimate the Q-value.

4. RESULTS

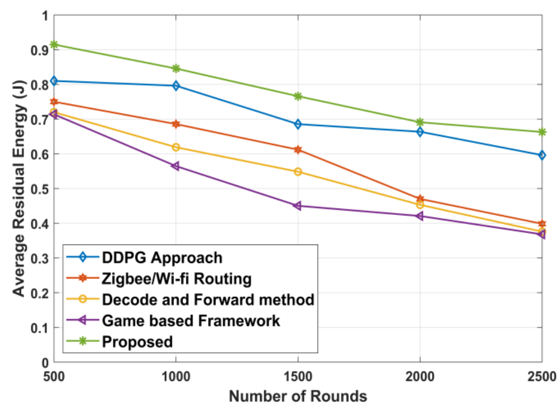
The evaluation of the proposed joint channel allocation and relay selection is devised by implementing in MATLAB programming tool based on various assessment measures. Here, the conventional resource allocation methods like DDPG Approach [29], Zigbee/WiFi Routing [30], Decode and Forward method [28], and Game based Framework [26] are utilized for evaluating the proposed protocol's routing performance. The simulation parameters of the proposed approach are portrayed in Table 2.

Table 2 Simulation Parameters

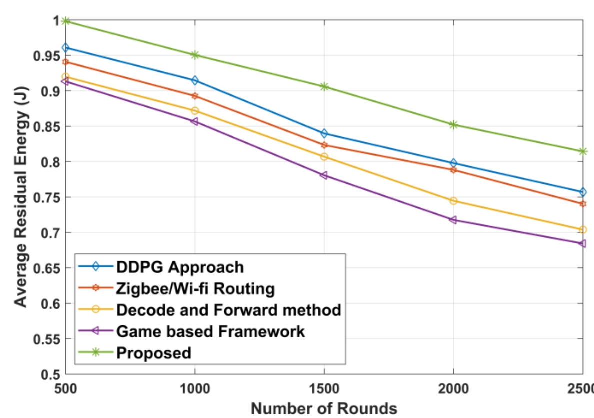
Parameters	Value
Number of Nodes	50, 100
Number of Relay	5, 10
Number of Rounds	2500
Initial Energy	1J
Data Rate	1kbps
Simulation area	$100x100m^2$

4.1. Assessment Based on Average Residual Energy

Figure 4 displays the average amount of remaining energy, which is the amount of energy left over after the devices have communicated, wherein Figure 4(a) illustrates the analysis with 50 nodes and Figure 4(b) illustrates the analysis with 100 nodes. Let us consider for example, the average residual energy estimated by proposed approach is 0.663 with 2500 rounds, which is 10.06%, 39.89%, 43.29%, and 44.47% superior outcome compared to the conventional DDPG Approach, Zigbee/WiFi Routing, Decode and Forward method, and Game-based Framework methods with 50 nodes. The average residual energy estimated by proposed approach is 0.950 with 1000 rounds, which is 3.77%, 6.08%, 8.27%, and 9.88% superior outcome compared to the conventional DDPG Approach, Zigbee/WiFi Routing, Decode and Forward method, and Game-based Framework methods with 100 nodes. Table 3 presents the detailed analysis.



(a)



(b)

Figure 4 Comparison Based on Average Residual Energy (a) 50 Nodes and (b) 100 Nodes

Table 3 Comparison Based on Average Residual Energy

Methods/ Rounds	500	1000	1500	2000	2500
50 Nodes					
DDPG Approach	0.810	0.796	0.686	0.664	0.596
Zigbee/Wi-fi Routing	0.750	0.686	0.612	0.470	0.399
Decode and Forward method	0.720	0.619	0.549	0.453	0.376
Game based Framework	0.714	0.565	0.450	0.421	0.368
Proposed	0.915	0.846	0.766	0.691	0.663

RESEARCH ARTICLE

100 Nodes					
DDPG Approach	0.961	0.915	0.840	0.798	0.757
Zigbee/Wi-fi Routing	0.941	0.893	0.823	0.788	0.740
Decode and Forward method	0.920	0.872	0.807	0.744	0.704
Game based Framework	0.913	0.857	0.780	0.717	0.684
Proposed	0.998	0.950	0.906	0.852	0.814

4.2. Assessment Based on Latency

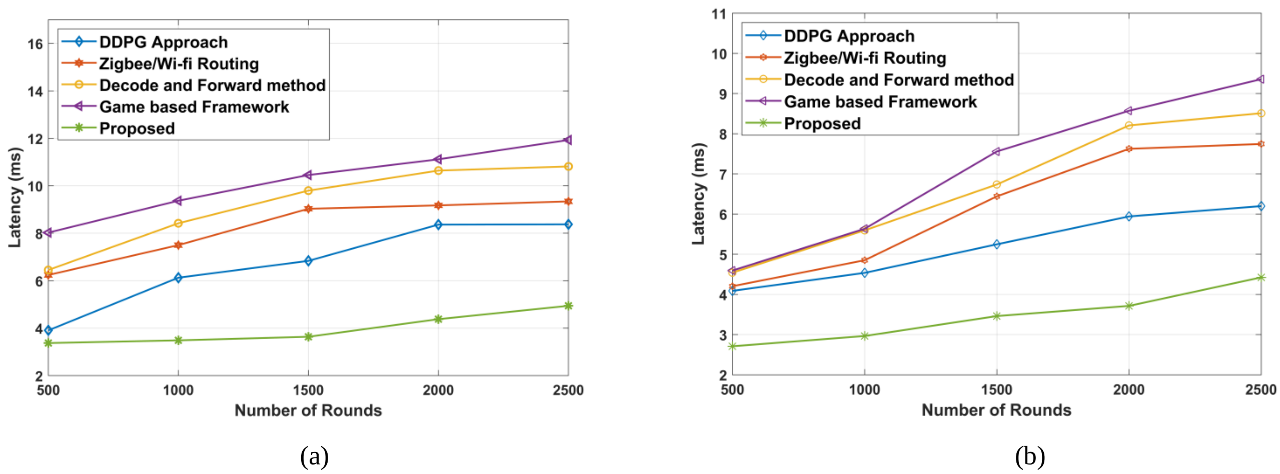


Figure 5 Comparison Based on Latency (a) 50 Nodes and (b) 100 Nodes

Table 4 Comparison Based on Latency

Methods/ Rounds	500	1000	1500	2000	2500
50 Nodes					
DDPG Approach	3.906	6.127	6.836	8.364	8.376
Zigbee/Wi-fi Routing	6.240	7.499	9.031	9.174	9.343
Decode and Forward method	6.451	8.420	9.797	10.642	10.816
Game based Framework	8.027	9.371	10.454	11.117	11.931
Proposed	3.372	3.486	3.637	4.379	4.942
100 Nodes					
DDPG Approach	4.089	4.536	5.247	5.941	6.198
Zigbee/Wi-fi Routing	4.542	4.851	6.441	7.624	7.745
Decode and Forward method	4.204	5.634	6.737	8.207	8.510
Game based Framework	4.594	5.594	7.558	8.572	9.358
Proposed	2.709	2.965	3.461	3.715	4.424

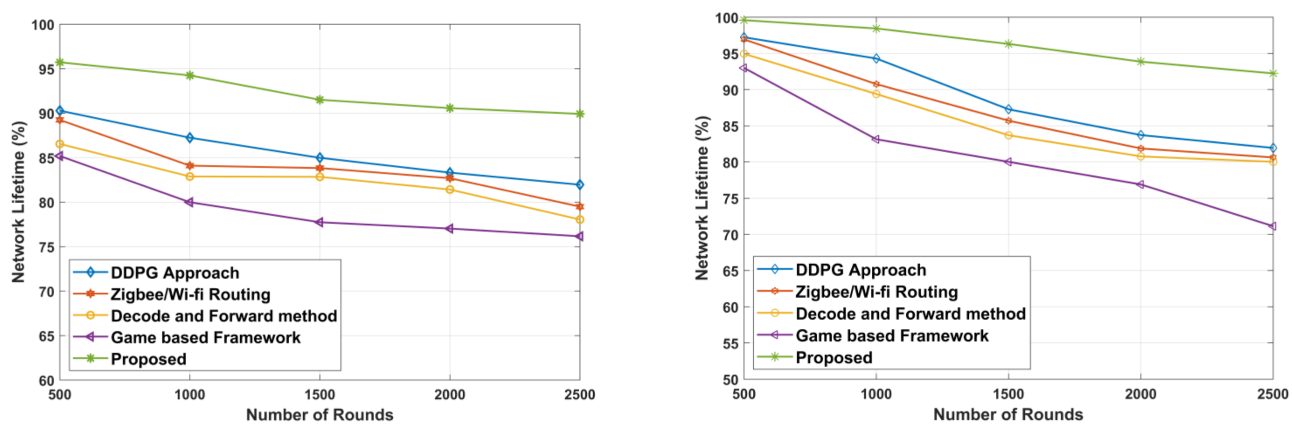
Figure 5 illustrates latency, which is the amount of time it takes for the sender to arrive at the destination, where the evaluation with 50 nodes is shown in Figure 5(a), and the evaluation with 100 nodes is shown in Figure 5(b). Here, the Latency estimated by proposed method is 3.372 with 500 rounds, which is 13.68%, 45.96%, 47.73%, and 57.99%

RESEARCH ARTICLE

superior outcome compared to the conventional DDPG Approach, Zigbee/WiFi Routing, Decode and Forward method, and Game-based Framework methods with 50 nodes. The Latency estimated by proposed method is 3.715 with 2000 rounds, which are 37.47%, 51.28%, 54.74%, and 56.66% superior outcome compared to the conventional DDPG Approach, Zigbee/WiFi Routing, Decode and Forward method, and Game-based Framework methods with 100 nodes. The analysis is presented in Table 4.

4.3. Assessment Based on Network Lifetime

Network lifetime is the amount of time that a network lasts until its first node burns out from energy depletion. Figure 6 shows this, with Figure 6(a) showing the analysis with 50 nodes and Figure 6(b) showing the analysis with 100 nodes. Here, the Network Lifetime estimated by proposed method is 91.507 with 1500 rounds, which is 7.13%, 8.39%, 9.47%, and 15.05% superior outcome compared to the conventional DDPG Approach, Zigbee/WiFi Routing, Decode and Forward method, and Game-based Framework methods with 50 nodes. The Network Lifetime estimated by proposed method is 93.860 with 2000 rounds, which are 10.80%, 12.78%, 13.94%, and 18.07% superior outcome compared to the conventional DDPG Approach, Zigbee/WiFi Routing, Decode and Forward method, and Game-based Framework methods with 100 nodes. The entire analysis is presented in Table 5.



(a) (b)

Figure 6 Comparison Based on Network Lifetime (a) 50 Nodes and (b) 100 Nodes

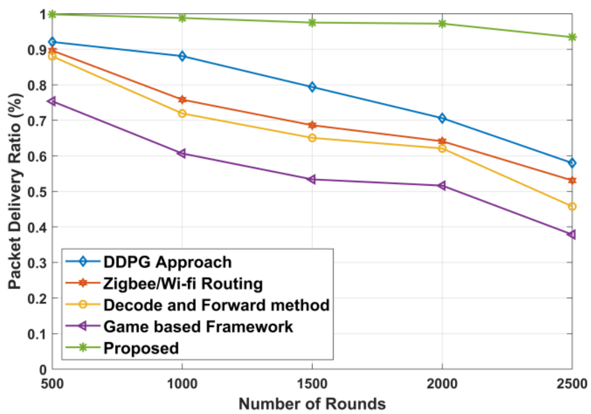
Table 5 Comparison Based on Network Lifetime

Methods/ Rounds	500	1000	1500	2000	2500
50 Nodes					
DDPG Approach	90.285	87.241	84.982	83.316	81.958
Zigbee/Wi-fi Routing	89.244	84.098	83.830	82.690	79.491
Decode and Forward method	86.547	82.883	82.842	81.416	78.044
Game based Framework	85.187	79.989	77.738	77.026	76.161
Proposed	95.720	94.245	91.507	90.567	89.920
100 Nodes					
DDPG Approach	97.232	94.297	87.283	83.721	81.940
Zigbee/Wi-fi Routing	96.938	90.763	85.708	81.867	80.629
Decode and Forward method	94.947	89.399	83.695	80.772	80.027
Game based Framework	92.999	83.140	80.024	76.896	71.131
Proposed	99.592	98.448	96.310	93.860	92.238

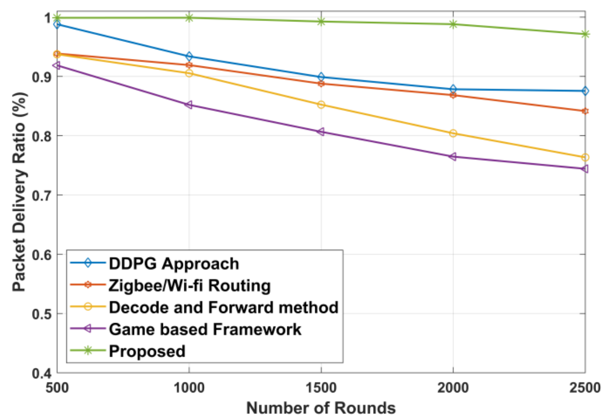
RESEARCH ARTICLE

4.4. Assessment Based on Packet Delivery Ratio

The packet delivery ratio is calculated by dividing the total number of packets sent by the sender by the total number of packets received at the destination, as shown in Figure 7, wherein Figure 7(a) illustrates the analysis with 50 nodes and Figure 7(b) illustrates the analysis with 100 nodes. Here, the proposed method's expected packet delivery ratio is 0.934 with 2500 rounds, which is 37.92%, 43.17%, 51.00%, and 59.46% superior outcome compared to the conventional DDPG Approach, Zigbee/WiFi Routing, Decode and Forward method, and Game-based Framework methods with 50 nodes. The proposed approach's expected packet delivery ratio is 0.999 with 1000 rounds, which are 6.53%, 8.00%, 9.37%, and 14.71% superior outcome compared to the conventional DDPG Approach, Zigbee/WiFi Routing, Decode and Forward method, and Game-based Framework methods with 100 nodes. The entire analysis is presented in Table 6.



(a)



(b)

Figure 7 Comparison Based on Packet Delivery Ratio (a) 50 Nodes and (b) 100 Nodes

Table 6 Comparison Based on Packet Delivery Ratio

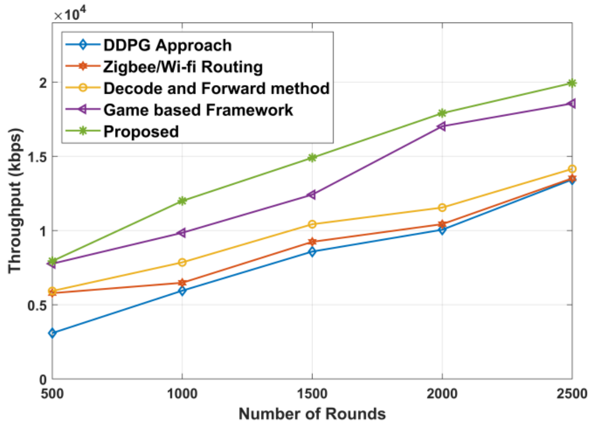
Methods/ Rounds	500	1000	1500	2000	2500
50 Nodes					
DDPG Approach	0.921	0.881	0.794	0.706	0.580
Zigbee/Wi-fi Routing	0.896	0.758	0.686	0.641	0.531
Decode and Forward method	0.880	0.719	0.651	0.621	0.458
Game based Framework	0.754	0.607	0.534	0.516	0.379
Proposed	0.998	0.988	0.975	0.972	0.934
100 Nodes					
DDPG Approach	0.988	0.934	0.899	0.878	0.876
Zigbee/Wi-fi Routing	0.939	0.919	0.888	0.868	0.842
Decode and Forward method	0.937	0.905	0.852	0.804	0.763
Game based Framework	0.919	0.852	0.807	0.765	0.744
Proposed	0.999	0.999	0.993	0.988	0.971

4.5. Assessment Based on Throughput

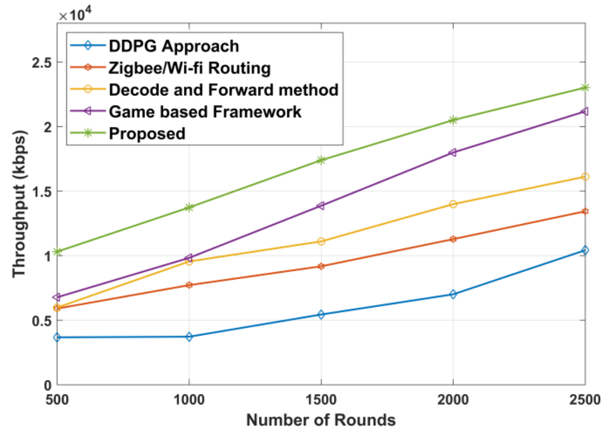
Throughput refers to the amount of data received at the destination with the pre-defined time and is depicted in Figure 8, wherein Figure 8(a) illustrates the analysis with 50 nodes and Figure 8(b) illustrates the analysis with 100 nodes. Here, the throughput calculated using the proposed approach is 11996 with 1000 rounds, which is 50.40%, 45.97%, 34.51%, and 17.89% superior outcome compared to the conventional DDPG Approach, Zigbee/WiFi Routing, Decode and Forward

RESEARCH ARTICLE

method, and Game-based Framework methods with 50 nodes. The throughput calculated using the proposed approach is 13728 with 1000 rounds, which are 72.91%, 43.82%, 30.47%, and 28.44% superior outcome compared to the conventional DDPG Approach, Zigbee/WiFi Routing, Decode and Forward method, and Game-based Framework methods with 100 nodes. The entire analysis is presented in Table 7.



(a)



(b)

Figure 8 Comparison Based on Throughput (a) 50 Nodes and (b) 100 Nodes

Table 7 Comparison Based on Throughput

Methods/ Rounds	500	1000	1500	2000	2500
50 Nodes					
DDPG Approach	3093	5950	8585	10056	13439
Zigbee/Wi-fi Routing	5785	6481	9241	10431	13521
Decode and Forward method	5927	7856	10426	11547	14154
Game based Framework	7766	9850	12418	17026	18562
Proposed	7940	11996	14906	17908	19941
100 Nodes					
DDPG Approach	3671	3719	5437	7002	10418
Zigbee/Wi-fi Routing	5896	7713	9171	11275	13433
Decode and Forward method	5975	9545	11098	13986	16114
Game based Framework	6770	9824	13863	17986	21173
Proposed	10303	13728	17389	20497	23015

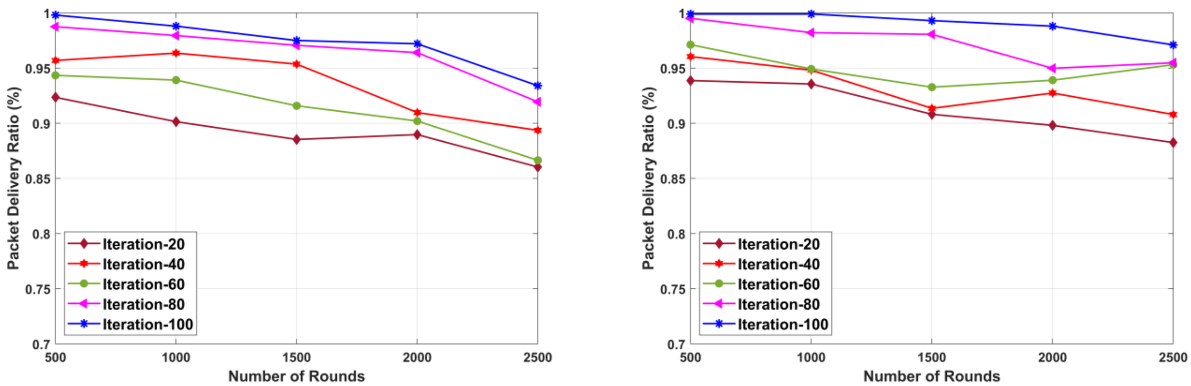
4.6. Iteration Based Analysis

In Figure 9, the proposed routing protocol is analyzed using different iterations. Figure 9 (a) shows a general decrease in PDR for all iterations in a 50-node network as the number of rounds increases.

This suggests that network performance is gradually declining. At the outset, the PDR is exceptionally high, surpassing 95% for every iteration, indicating robust performance. In a 100-node network, figure 9(b) shows a steady decrease in PDR over all iterations as the number of rounds rises. This pattern indicates that network performance has been steadily declining over time.

In this, the analysis depicts that the superior performance is acquired with higher number of iterations. The entire analysis is presented in Table 8.

RESEARCH ARTICLE



(a) (b)

Figure 9 Analysis of PDR in Terms of Iteration (a) with 50 Nodes and (b) with 100 Nodes

Table 8 Analysis Based on Iteration

Iteration/ Rounds	500	1000	1500	2000	2500
50 Nodes					
Iteration-20	0.9235	0.9014	0.8853	0.8897	0.8603
Iteration-40	0.9569	0.9636	0.9536	0.9097	0.8936
Iteration-60	0.9434	0.9391	0.9158	0.902	0.8664
Iteration-80	0.9875	0.9795	0.9706	0.964	0.9194
Iteration-100	0.998	0.988	0.975	0.972	0.934
100 Nodes					
Iteration-20	0.9387	0.9356	0.9081	0.8981	0.8825
Iteration-40	0.9604	0.9481	0.9135	0.9274	0.9079
Iteration-60	0.9712	0.9492	0.9327	0.939	0.9531
Iteration-80	0.9951	0.9821	0.9806	0.9498	0.9548
Iteration-100	0.999	0.999	0.993	0.988	0.971

4.7. Discussion

The discussion based on the best outcome is illustrated in Table 9. The maximum throughput estimated by the proposed method is 23015, which is 54.73%, 41.63%, 29.98%, and 8.00% superior outcome compared to the conventional DDPG Approach, Zigbee/WiFi Routing, Decode and Forward method, and Game-based Framework methods with 100 nodes. As per the proposed approach, the highest packet delivery ratio is 0.999, which is 1.09%, 6.04%, 6.16%, and 8.04% superior outcome compared to the conventional DDPG Approach, Zigbee/WiFi Routing, Decode and Forward method, and Game-based Framework methods with 100 nodes. The maximum network lifetime estimated by the proposed method is 99.592, which is 2.37%, 2.67%, 4.66%, and 6.62% superior outcome compared to the conventional DDPG Approach, Zigbee/WiFi Routing, Decode and Forward method, and Game-based Framework methods with 100

RESEARCH ARTICLE

nodes. The minimum latency estimated by the proposed method is 2.709, which is 33.73%, 40.35%, 35.55%, and 41.02% superior outcome compared to the conventional DDPG Approach, Zigbee/WiFi Routing, Decode and Forward method, and Game-based Framework methods with 100 nodes. The maximum average residual energy estimated by the proposed approach is 0.998, which is 3.73%, 5.74%, 7.86%, and 8.51% superior outcome compared to the conventional DDPG Approach, Zigbee/WiFi Routing, Decode and Forward method, and Game-based Framework methods with 100 nodes.

The proposed method performs better in CRN because of a number of important aspects. First of all, it combines relay selection with joint channel allocation, enabling the model to learn optimal strategies adaptively depending on network conditions. When compared to existing approaches, this results in better decision-making and resource use. Furthermore, the EnHpo algorithm is used to optimize channel allocation by taking into account a number of goals, including transmission rate, bandwidth, and priority. A more equitable and effective distribution of resources is made possible by this multi-objective fitness function, which is essential for preserving high throughput and low latency.

Furthermore, relay communication minimizes interruptions from main users by efficiently managing data transit amongst cognitive users. By selecting the best relays based on channel gain and bit error rate, the relay selection process improves the overall dependability and effectiveness of D2D communication. In dynamic contexts where network circumstances might change quickly, the model's scalability the ability to sustain performance even as the number of nodes increases is particularly crucial. With an average residual energy of 0.998, which indicates efficient energy utilization, and a latency of 2.709, which suggests minimal delays in data transfer, key performance measures demonstrate notable gains. The model's robustness and capacity to sustain high performance under a variety of circumstances are attributed to its design, which takes into account a number of variables that impact network performance, including interference, node mobility, and channel fading.

Table 9 Comparative Discussion

Methods/ Metrics	Average Residual Energy	Latency	Network Life Time	Packet Delivery Ratio	Throughput
50 Nodes					
DDPG Approach	0.810	3.906	90.285	0.921	13439
Zigbee/Wi-fi Routing	0.750	6.240	89.244	0.896	13521
Decode and Forward method	0.720	6.451	86.547	0.880	14154
Game based Framework	0.714	8.027	85.187	0.754	18562
Proposed	0.915	3.372	95.720	0.998	19941
100 Nodes					
DDPG Approach	0.961	4.089	97.232	0.988	10418
Zigbee/Wi-fi Routing	0.941	4.542	96.938	0.939	13433
Decode and Forward method	0.920	4.204	94.947	0.937	16114
Game based Framework	0.913	4.594	92.999	0.919	21173
Proposed	0.998	2.709	99.592	0.999	23015

4.8. Limitations

The method described presents several advantages over conventional approaches, particularly in terms of throughput, packet delivery ratio, network lifetime, latency, and residual energy. However, there are some limitations to consider:

Scalability: Comparing the proposed approach to traditional methods, it performs better with 100 nodes; however, more research is necessary to determine its scalability. Performance impacted when the network expands due to the difficulty of handling multi-hop routes and relay selection, which result in decreased throughput and increased latency.

RESEARCH ARTICLE

Computational Overhead: As the network size increases, the computational expense of using DRL for relay selection and channel allocation result in higher latency and lower throughput, particularly in situations that call for prompt decision-making.

Resource Constraints: Although the proposed approach maximizes resource use in CRN, performance impacted by resource limitations such as restricted bandwidth and transmission rate as the number of nodes rises, particularly in situations with high traffic or dynamic conditions.

Generalization and Adaptability: The proposed method performs well in the scenarios that were investigated, although its efficacy changes depending on the network topology, traffic patterns, and environmental conditions. In order to maintain performance in a variety of environments, more research is required.

Reliability and Robustness: The reliability and robustness of the proposed approach under varying network conditions, including interference, node mobility, and channel fading, need thorough investigation. While the method aims to optimize performance, its resilience to unpredictable network events and adversarial conditions requires validation through extensive experimentation and testing.

Overall, while the proposed method shows promising results in improving network performance metrics, addressing these limitations through further research and experimentation is essential to ensure its effectiveness and practicality in real-world deployments.

5. CONCLUSION AND FUTURE WORK

This research proposed a DRL method for relay selection and channel allocation. Initially, the proposed EnHpo algorithm is used to develop the channel allocation, taking into account the multi-objective fitness function, which includes transmission rate, bandwidth, and priority. In order to speed up convergence and identify the best solution worldwide, the proposed EnHpo is created by combining the adaptive weight method with the traditional hunter-prey optimization. Followed by, the relay selection is devised using the deep reinforcement learning technique by considering the channel gain based on the bit error rate. As a result, the proposed strategy outperformed the traditional cooperative routing strategies. Average Residual Energy measures the average amount of energy left in the network nodes after communication activities. The value of 0.998 indicates that, on average, almost all nodes retain a significant amount of energy, suggesting efficient energy utilization. Data packet latency is the amount of time it takes for them to go from their source to their destination. A latency value of 2.709 indicates a relatively low delay, signifying efficient data transmission in the network. Network lifetime represents the duration for which the network can sustain its operations before nodes start to fail due to energy depletion or other factors. A value of 99.592 suggests that the network can operate efficiently for a prolonged period, indicating good longevity. The percentage of successfully delivered packets among all sent packets is known as the packet delivery ratio. A value of 0.999 indicates a very high packet delivery ratio, implying that the majority of packets reach their intended destinations successfully, ensuring reliable communication. Throughput measures the rate at which data is successfully transmitted through the network. The throughput value of 23015 signifies a high data transfer rate, indicating efficient utilization of network resources. In the future, the new deep learning framework will be used to develop D2D communication, however this will not take energy efficiency into account.

REFERENCES

[1] Yang D, Li C, Ren B, Li H, Guo K. Analysis of relay selection schemes in underlay cognitive radio non-orthogonal multiple access networks. Int J Distrib Sens Networks. 2022 May 13;18(5):155014772110663.

[2] Li X, Gao X, Shaikh SA, Zeng M, Huang G, Faseeh Qureshi NM, et al. NOMA-based cognitive radio network with hybrid FD/HD relay in industry 5.0. J King Saud Univ - Comput Inf Sci. 2023 Jun;35(6):101363.

[3] Chen G, Chen G. A Cooperative Automatic Retransmission Request Technique Based on Listening Node Participation Competition. Appl Sci. 2023 Feb 8;13(4):2161.

[4] Li S, Li H, Yang Q, Gaber J. Hybrid Cooperative Spectrum Sharing Protocols Based on OMA and NOMA in Cognitive Radio Networks. Appl Sci. 2022 Dec 10;12(24):12683.

[5] Tarek D, Hafez MI. Development of Spectrum Sharing Protocol for Cognitive Radio Internet of Things. 2021.

[6] Jasim SI, Akawee MM, Hasan RA. A spectrum sensing approaches in cognitive radio network by using cloud computing environment. Bull Electr Eng Informatics. 2022 Apr 1;11(2):750–7.

[7] Cicioğlu M, Çalhan A, Miah MS. An effective routing algorithm for spectrum allocations in cognitive radio based internet of things. Concurr Comput Pract Exp. 2022 Dec 25;34(28).

[8] Yu X, Cai Y, Li W, Zhou X, Tang L. Research on Dynamic Spectrum Allocation Algorithm Based on Cyclic Neural Network. Jafari Navimipour N, editor. Wirel Commun Mob Comput. 2022 Mar 23;2022:1–14.

[9] Arshid K, Jianbiao Z, Hussain I, Pathan MS, Yaqub M, Jawad A, et al. Energy efficiency in cognitive radio network using cooperative spectrum sensing based on hybrid spectrum handoff. Egypt Informatics J. 2022 Dec;23(4):77–88.

[10] P. Venkatapathi, H. Khan, S. Srinivasa Rao, and G. Immadi, “Cooperative Spectrum Sensing Performance Assessment using Machine Learning in Cognitive Radio Sensor Networks,” Engineering, Technology & Applied Science Research, vol. 14, no. 1, pp. 12875–12879, Feb. 2024.

[11] Ferreira, M.V., Vieira, F.H. and Cardoso, Á.A., 2022. Power and delay optimization based uplink resource allocation for wireless networks with device-to-device communications. Computer Communications, 194, pp.226-239.

[12] Isnawati AF. A Survey of Game Theoretical Approach in Cognitive Radio Network and 5G-6G Communications. J Commun. 2022;17(10):830–43.

[13] Zafar MH, Khan I, Alassafi MO. An efficient resource optimization scheme for D2D communication. Digit Commun Networks. 2022 Dec;8(6):1122–9.

[14] Alotaibi S. Interference Mitigation Strategy for D2D Communication in 5G Networks. Eng Technol Appl Sci Res. 2023 Aug 9;13(4):11318–25.

RESEARCH ARTICLE

[15] Omidkar A, Khalili A, Nguyen HH, Shafiei H. Reinforcement-Learning-Based Resource Allocation for Energy-Harvesting-Aided D2D Communications in IoT Networks. IEEE Internet Things J. 2022 Sep 1;9(17):16521–31.

[16] Chatterjee S, De S. QoE-Aware Cross-Layer Adaptation for D2D Video Communication in Cooperative Cognitive Radio Networks. IEEE Syst J. 2022 Jun;16(2):2078–89.

[17] Arun Mozhi Selvi S, Alroobaea R, Rubaiee S, S. Hanbazazah A. Optimized Resource Allocation for IoT-D2D Communication Using WSN. Comput Syst Sci Eng. 2023;45(2):1789–804.

[18] Dejen A, Wondie Y, Förster A. Survey on D2D Resource Scheduling and Power Control Techniques: State-of-art and Challenges. ICST Trans Mob Commun Appl. 2018 Jul 13;173977.

[19] Khoshafa MH, Ngatched TMN, Ahmed MH. Relay Selection for Improving Physical Layer Security in D2D Underlay Communications. IEEE Access .2022;10:95539–52.

[20] Ahmad N, Sidhu GAS, Khan WU. A Learning Based Framework for Enhancing Physical Layer Security in Cooperative D2D Network. Electronics. 2022 Nov 30;11(23):3981.

[21] Eappen G, T S, Nilavalan R. Cooperative relay spectrum sensing for cognitive radio network: Mutated MWOA-SNN approach. Appl Soft Comput. 2022 Jan;114:108072.

[22] Farooq-i-Azam M, Ni Q, Dong M, Pervaiz H, Zarakovitis C, Alsolami F. Interference Avoidance in Cognitive Radio Networks Using Cooperative Localization. IEEE Syst J. 2022 Dec;16(4):5120–30.

[23] Ramzan MR, Naeem M, Chughtai O, Ejaz W, Altaf M. Radio resource management in energy harvesting cooperative cognitive UAV assisted IoT networks: A multi-objective approach. Digit Commun Networks. 2023 Jan.

[24] Salameh HB, Abdel-Razeq S, Al-Obiedollah H. Integration of Cognitive Radio Technology in NOMA-Based B5G Networks: State of the Art, Challenges, and Enabling Technologies. IEEE Access. 2023;11:12949–62.

[25] Amini T, Abouei J. A matching-based context-aware relay assignment scheme with power control in ad-hoc wireless networks. IET Commun. 2023 Feb 28;17(3):348–61.

[26] Alsunbuli BN, Fakhruldeen HF, Ismail W, Mahyuddin NM. Hybrid beamforming with relay and dual-base stations blockage mitigation in millimetre-wave 5G communication applied in (VIOT). Comput Electr Eng. 2022 May;100:107953.

[27] Sun G, Li J, Wang A, Wu Q, Sun Z, Liu Y. Secure and Energy-Efficient UAV Relay Communications Exploiting Collaborative Beamforming. IEEE Trans Commun. 2022 Aug;70(8):5401–16.

[28] Kumar R, Singh H. Machine learning solution of a coalitional game for optimal power allocation in multi-relay cooperative scenario. Sādhanā. 2022 Nov 22;47(4):250.

[29] Imtiaz HH, Tang S. Multi-Task Partial Offloading with Relay and Adaptive Bandwidth Allocation for the MEC-Assisted IoT. Sensors. 2022 Dec 24;23(1):190.

[30] Sarma SS, Hazra R, Chong PHJ. Performance Analysis of DF Relay-Assisted D2D Communication in a 5G mmWave Network. Futur Internet. 2022 Mar 24;14(4):101.

[31] Zhang T, Zhu K, Wang J. Energy-Efficient Mode Selection and Resource Allocation for D2D-Enabled Heterogeneous Networks: A Deep Reinforcement Learning Approach. IEEE Trans Wirel Commun. 2021 Feb;20(2):1175–87.

[32] Lee, W. and Schober, R., 2022. Deep learning-based resource allocation for device-to-device communication. IEEE Transactions on Wireless Communications, 21(7), pp.5235-5250.

[33] Chennaboin, R.B. and Nandakumar, S., 2024. An adaptive MLP-based joint optimization of resource allocation and relay selection in device-to-device communication using hybrid meta-heuristic algorithm. EURASIP Journal on Wireless Communications and Networking, 2024(1), p.50.

[34] Y. Yu and X. Tang, "Hybrid Centralized-Distributed Resource Allocation Based on Deep Reinforcement Learning for Cooperative D2D Communications," in IEEE Access, vol. 12, pp. 196609-196623, 2024.

[35] Logeshwaran, J., R. N. Shanmugasundaram, and Jaime Lloret. "Load based dynamic channel allocation model to enhance the performance of device-to-device communication in WPAN." Wireless Networks 30.4 (2024): 2477-2509.

Authors

Dr. Md. Tabrej Khan is a researcher and academic specializing in machine learning, artificial intelligence, and data science. He holds a PhD in Computer Science from Pacific Academy of Higher Education and Research University, India, and an MSc in Computer Science from Jamia Hamdard University. Dr. Khan has extensive teaching and research experience, having served at King Abdulaziz University, Saudi Arabia, and currently contributing to AISECT University, Jharkhand, and Liwa College, UAE. His research focuses on deep learning, NLP, and 5G technologies, with numerous publications in high-impact journals. He is also actively involved in consultancy projects and professional organizations in computing and engineering.

Dr. Ashish Adholiya is right now working as Assistant Professor at Pacific Academy of Higher Education and Research University Udaipur (Rajasthan), India. He pursued his Ph.D. in the area of Database Flexibility from JRN Rajasthan Vidyapeeth (Deemed to be University, NAAC –A Grade). He has a total experience of 12 years in academics and 3 years of IT companies. He has authored 16 research articles for international journals and 23 research articles for national journals with impact factor. He has been conducting Management Development Program to various organizations like IOC. He is managing editor of two national journals published by Pacific University, Udaipur since last 3 years. He has been the editor of two books published by the Pacific University, Udaipur.

How to cite this article:

Md. Tabrej Khan, Ashish Adholiya, "Cooperative Device-to-Device Communication Using Joint Relay Assignment and Channel Allocation Using Deep Learning", International Journal of Computer Networks and Applications (IJCNA), 12(1), PP: 77-92, 2025, DOI: 10.22247/ijcna/2025/06.