International Journal of Basic and Applied Sciences, 14 (2) (2025) 189-201

International Journal of Basic and Applied Sciences

Website: www.sciencepubco.com/index.php/IJBAS
https://doi.org/10.14419/474dfq89
Research paper

Multi-agent learning for UAV networks: a unified approach to trajectory control, frequency allocation and routing

Angara Satyam [1] *, R. Anil Kumar [2], Sarala Patchala [3], Sunitha Pachala [4], Geeta Bhimrao Atkar [5], U. S. B. K. Mahalaxmi [2]

[1] *Department of Information Technology, Sagi Rama Krishnam Raju Engineering College(A), Bhimavaram, AP, India*
[2] *Department of Electronics and Communication Engineering, Aditya University, Surampalem, AP, India*
[3] *Department of Electronics and Communication Engineering KKR & KSR Institute of Technology and Sciences, Guntur, AP, India*
[4] *Department of Artificial Intelligence and Data Science, Koneru Lakshmaiah Education Foundation, Guntur, AP, India*
[5] *Department of Computer Science Engineering, G H Raisoni College of Engineering and Management, Pune, Maharashtra, India*
**Corresponding author E-mail: satyam.angara@gmail.com*

Received: May 6, 2025, Accepted: May 28, 2025, Published: June 11, 2025

Abstract

Unmanned aerial vehicle (UAV) swarm networks have various applications. They support surveillance and communication. However, these networks face many challenges. High mobility causes frequent link breakages. Limited transmission range increases interference. Routing becomes complex due to these factors. This paper proposes a new approach. It jointly optimizes trajectory control, frequency allocation and routing. The proposed algorithm is called JTFR. It improves link utility in UAV swarm networks. Link utility depends on signal strength, queue delay and residual energy. The proposed JTFR uses multi-agent reinforcement learning. The learning model considers both local and neighbour information. This improves the stability of the network. A long short-term memory (LSTM) network is used. This helps in pre-dicting the best trajectory. A multi-head attention mechanism is also used. Each UAV adjusts its policy based on its neighbours. UAV swarm networks are dynamic. The movement of UAVs changes frequently. This affects network connectivity. The routing decisions must adapt to these changes. The proposed approach provides adaptive routing. JTFR is overall useful to improve UAV swarm communication. It handles trajectory control, frequency allocation and routing in balancing. The method is dynamic in the sense that it adapts to dynamic network conditions. The UAVs work well in complex environments.

Keywords: *UAV; Joint Optimization; LSTM Cell; JTFR Algorithm; Reinforcement Learning.*

1. Introduction

UAV swarm Based on surveillance, disaster response and communication systems, UAV swarm network has been widely used in recent years [1]. Autonomous UAVs work together to the coverage of large areas. They are mobile, and therefore are flexible and scalable solution for monitoring and data transmission. Nevertheless, controlling the UAV swarm network possesses a number of difficulties [2]. Some of these challenges consist of keeping communication stable, routing optimization and optimum usage of energy. Dynamic topology [3] is one of the main challenges in UAV swarm network. UAVs traverse many miles per hour and consequently the link is broken many times. It affects the data transmission efficiency and will lead to packet loss. There is another problem, that comes from the interference of several UAVs working in the same frequency band [4]. Frequency allocation should be done properly such that quality degradation of communication should not be seen. Also, since UAVs are said to be battery powered, energy efficiency is a determining factor in network lifespan. In order to overcome such issues, this paper suggests a joint trajectory control, frequency allocation and routing (JTFR) algorithm. In reference to the former, a UAV swarm communication improvement based on the trajectory control, frequency allocation and routing optimization (JTFR algorithm) is introduced [5]. It allows the UAVs to maintain coordinated paths that help maintain the network stability. This avoids unnecessary movement and, as a result, energy conservation and operation prolongation. It will keep bonded UAVs without colliding with other and enables them covering the efficient area[6]. Frequency allocation plays a proper role to reduce interference and improve the signal quality. Algorithm determines frequencies dynamically against the real time network conditions. Besides that, it helps in routing by choosing the optimum path considering constant links with low energy consumption [7]. Consequently, packet delivery rates increase and the data transmission becomes more reliable.

The JTFR algorithm lets UAVs adapt to the changes in network condition with the aid of reinforcement learning[8]. These UAVs learn from their past experiences and change, in accordance with their past experiences, the decision-making strategies. To address this problem, we introduce the algorithm which comprises a LSTM network to process historical mobility data to enable UAVs to predict future movements and adjust the routes accordingly. Multi head attention mechanism helps in prioritizing the critical network parameter such as link stability and queue delays[9], which enables proper decision making. Behavior based model for the UAV swarm motion is used. In this model, UAVs have safe distance and network connectivity maintaining[10]. It allows drive more efficiently within available space and to avoid collisions. These additionally reduce interference within and increase stability of the whole network through coordinated movement patterns. We propose JTFR algorithm that combines numerous optimization techniques to achieve seamless UAV communications. It enhances the data transmission efficiency, cuts down communication times and prolongs the network's life span[11]. JTFR, a robust solution to UAV swarm networks, is the combination of reinforcement learning, predictive mobility modeling and adaptive frequency allocation. The paper makes several key contributions:

- Proposal of a novel joint optimisation approach for trajectory control, frequency allocation and routing.
- Improvement of UAV decision by way of reinforcement learning. Implementation of an LSTM network for better mobility prediction.
- Application of a multi-head attention mechanism for adaptive routing. Behavior-based UAV motion model for stable and efficient communication.
- Adaptive frequency allocation to reduce interference and enhance communication reliability. Extensive simulations showing improved network performance over existing methods.

This research significantly enhances UAV communication networks. It results stable connectivity, minimizes interference and optimizes energy usage [12]. The proposed approach supports real-time applications like disaster monitoring, military surveillance and remote sensing operations. The findings demonstrate that JTFR outperforms traditional routing methods, providing a reliable and scalable solution for UAV swarm networks. Even though, JTFR algorithm has mainly two limitations:

First, Computational Complexity: The JTFR algorithm incorporates deep reinforcement learning using DMA-DDPG with LSTM and multi-head attention mechanisms. These components increase the computational complexity during the training phase. The training of multiple agents requires extensive neural network updates, replay buffer management, and gradient computations. These are not feasible on resource-constrained UAV hardware in real-time environments.

Second, Scalability: Although JTFR performs well with moderate numbers of UAVs, if it is scaling to larger swarms introduces challenges. The coordination overhead, increased state space and action selection complexity grow rapidly with the number of agents. This led to longer convergence times and reduced decision-making efficiency under high-density deployments.

2. Related work

UAV swarm networks have been studied in various research works. These networks support communication, surveillance and disaster response applications. However, they face challenges related to routing, frequency allocation and trajectory control. Researchers have proposed different methods to overcome these challenges. Several studies have focused on optimizing UAV trajectory to improve communication efficiency. In [13] the authors proposed a trajectory control method using reinforcement learning. It allows for adaptation of UAVs to changing environment. A behavior based trajectory model was introduced in another study [14]. This method gives stable connectivity but avoids collisions. In [15] the authors study a particular motion model inspired by flocking behavior. Also, it aids UAVs in staying within formation and obtaining optimal coverage. The results presented in these studies relate the importance of trajectory optimization in a UAV swarm network. UAV communication is critically dependent on frequency allocation. In [16] they proposed a dynamic frequency selection algorithm. Interference is kept to a minimum and signal quality is improved. An alternative approach [17] was to include an adaptive frequency reuse technique. The channels are allocated based on the UAV positions to improve spectral efficiency. In [18], the frequency assignment model was developed by the authors using reinforcement learning. It allows the UAVs to choose the best suited frequencies according to network environment. These techniques heavily decrease the communication interference in UAV networks.

Due to the dynamic topology and limited communication range, routing in UAV networks is difficult. A multi hop routing protocol for the doamin was presented in [19]. It evaluates the energy efficiency and the link stability. Another work [20] proposed a position-based routing algorithm. This method improves packet delivery rates by selecting optimal paths based on UAV locations. The authors in [21], [22] developed a machine learning-based routing approach. It uses historical data to predict the best routing paths. These solutions enhance data transmission reliability in UAV networks. Recent research has explored joint optimization of trajectory, frequency and routing. The work in [23], [24] introduced a cross-layer optimization framework. It integrates trajectory planning, frequency allocation and routing decisions. Another study [25], [26] proposed a deep reinforcement learning model. This model optimizes multiple network parameters simultaneously. The authors in [27] designed a collaborative decision-making system. It allows UAVs to share information and improve overall network performance. Other state-of-the-art advancements [28], use UAVs as flying base stations to support mobile networks. This helps offload mobile data and improve service. A new metric delanalty is introduced. It measures delay, urgency of data, and remaining data. A method to find the user with the highest delanalty is used. An actor–critic based DRL algorithm is proposed. It helps the UAV find the best path to reduce delanalty for all users. S. Javed et. al [29] proposed UAV swarms, it gained attention for handling complex tasks better. They offer smart coordination, flexibility, and strong system design. UAV swarms need different subsystems like path planning, communication, and coordination. This review talks about key parts such as swarm control, autonomy, and security. It also covers new algorithms and real-world uses in military, civilian, and entertainment fields. DRL is very useful when applying UAVs in IoT and WSNs [30]. It enables the planning of UAV routes, the choice of where to collect data and the organization of data as it arrives. This review examines the part that DRL plays in cutting down the Age of Information (AoI) which indicates how current the collected data is. It is responsible for supporting actions in disaster situations, as well as controlling and overseeing these actions. It also looks at difficulties related to brain imaging, solutions in practice and work that lies ahead for researchers. In UAV-aided IoT networks, keeping stable wireless connections is hard due to UAVs' short battery life. This study uses deep reinforcement learning (DRL) to improve energy efficiency. Two centralized methods, RR-DQL and SK-DQL, use a control station to manage UAVs, but they cause delays and high overhead. To solve this, a quasi-distributed QD-DQL method is proposed, where UAVs act on local data. Tests show QD-DQL performs better than other traditional methods [31]. These recent studies demonstrate the effectiveness of joint optimization techniques. The reviewed studies highlight advancements in UAV trajectory control, frequency allocation and routing. Existing solutions improve network performance but still face challenges in scalability and adaptability. The proposed JTFR algorithm builds on these works by integrating reinforcement learning, behavior-based motion and adaptive frequency selection.

3. System model

UAV swarm networks are essential in modern communication and surveillance applications. These networks involve multiple UAVs working together for seamless data transmission and coverage over a wide area. The performance of such networks is influenced by network topology, communication models, energy constraints and mobility patterns. A UAV swarm network is represented as a graph G = (U, E), here U represents the set of UAVs. E represents the set of links between them. Each UAV is denoted as u_i and the set of UAVs in the network is $U=\{u_1,u_2,...,u_N\}$. The connection between UAVs is determined by the communication range. The Euclidean distance between two UAVs u_i and u_j is:

$$d_{ij}=\sqrt{(x_i-x_j)^2+(y_i-y_j)^2+(z_i-z_j)^2} \tag{1}$$

Here, (x_i,y_i,z_i) and (x_j,y_j,z_j) are the coordinates of UAVs u_i and u_j, respectively. A link exists between UAVs if $d_{ij}\le R$, here R is the maximum communication range. Each UAV communicates with its neighbours in the network. The topology dynamically changes due to the UAVs' mobility, which affects routing decisions and link stability. The wireless communication channel between UAVs experiences path loss, interference and fading. The received power at UAV u_j from UAV u_i follows the free-space path loss model:

$$P_r=P_tG_tG_r\left(\frac{\lambda}{4\pi d_{ij}}\right)^2 \tag{2}$$

Here, P_t is the transmitted power, G_t and G_r are the antenna gains. λ is the signal wavelength and d_{ij} is the distance between UAVs. The Signal-to-Interference-plus-Noise Ratio (SINR) at UAV u_j is given by:

$$SINR=\frac{P_r}{I+N_0} \tag{3}$$

Here, I represent the interference power from other UAVs and N_0 is the background noise power. UAVs operate with a mobility model that dictates the movement and positioning. The velocity of a UAV changes according to:

$$v_i(t+1)=v_i(t)+a_i\,\Delta t \tag{4}$$

Here, a_i represents the acceleration and Δt is the time step. The new position of the UAV is given by:

$$x_i(t+1)=x_i(t)+v_i(t)\cos(\theta_i)\Delta t \tag{5}$$

$$y_i(t+1)=y_i(t)+v_i(t)\sin(\theta_i)\Delta t \tag{6}$$

Here, θ_i is the flight direction. The UAV swarm follows a coordinated flight pattern. The control system maintains a safe separation distance between UAVs while optimizing network coverage. UAV energy consumption is influenced by propulsion and communication activities. The total energy consumption is:

$$E_{total}=E_{propulsion}+E_{communication} \tag{7}$$

The propulsion energy is given by:

$$E_{propulsion}=P_{hover}t+P_{move}\frac{d}{v} \tag{8}$$

Here, P_{hover} is the power required to hover, P_{move} is the power required for movement, t is the flight time and v is the UAV velocity. The communication energy consumption is:

$$E_{communication}=P_tt_{tx}+P_rt_{rx} \tag{9}$$

Here, P_t and P_r are the transmission and reception power. t_{tx} and t_{rx} are the transmission and reception time. The UAV energy constraints must be considered in trajectory planning and network design. The objective of the UAV swarm network is to optimize communication efficiency, energy consumption and trajectory planning. The optimization problem is formulated as follows:

$$\max\sum_{i\in U}SINR_i-\sum_{i\in U}E_{total} \tag{10}$$

It is subjected to:

$$\begin{aligned}&d_{ij}\le R,\quad \forall(i,j)\in E\\&E_i\ge E_{min},\quad \forall i\in U\end{aligned} \tag{11}$$

Here, E_{min} is the minimum required energy level for UAV operation. The optimization model maintains UAVs reliable communication while minimizing energy consumption and avoiding unnecessary movement. This section described the UAV swarm network's system model including its network architecture, communication model, mobility behavior and energy efficiency. The proposed optimization approach enhances UAV network performance by providing energy efficiency, stable communication links and optimal trajectory planning.

4. DMA-DDPG based JTFR algorithm

Efficient trajectory planning, frequency allocation and routing in such networks is used to maintain smooth communication. Current methods fall short in handling dynamic environments and their decision processes can be suboptimal. UAV decision making is enhanced by an algorithm that uses reinforcement learning named the Distributed MultiAgent Deep Deterministic Policy Gradient (DMA-DDPG). In summary, it allows the UAVs to learn the optimal policy to control the trajectory, the frequency selection and the routing in real time, based on actual network conditions. This paper presents the JTFR algorithm which includes the DMA-DDPG model for efficiency of network optimization. Based on the results approach is stable, minimizes interference, and enhances resource utilization in general. The autonomous trajectory optimization and communications network strategies are then learned by these UAVs. In reinforcement learning (RL), an agent interacts with an environment to seek overall the maximum cumulative rewards. UAVs are designed as RL agents in the JTFR framework, which makes RL routing and RL allocation decisions of them.

The Figure 1 represents an LSTM-based architecture for decision-making in a system. The input to the model is denoted as $o_i(t)$, which is processed by three separate LSTM-based submodules, labeled as LSTM-Based SRL-1, LSTM-Based SRL-2 and LSTM-Based SRL-3. Each submodule receives different components of the input, indicated as $o_i^1(t)$, $o_i^2(t)$ and $o_i^3(t)$. These submodules are designed to extract useful sequential patterns and dependencies from the input data over time. The LSTM cell at the top of the figure processes information in a structured manner, ensuring temporal dependencies are captured effectively. After processing through the LSTM-based submodules, the output is passed to a fully connected layer (FCL). The outputs from the three LSTM-based SRLs are represented as $o_i^{h,1}(t)$, $o_i^{h,2}(t)$ and $o_i^{h,3}(t)$. These outputs serve as high-level feature representations of the input data. The FCL takes these features and processes them to generate the final output, denoted as $a_i(t)$. This structure suggests a hierarchical decision-making process with sequential data processed, learned and transformed into a final action. The model aims to improve prediction accuracy and decision efficiency by developing LSTM networks for temporal data processing.

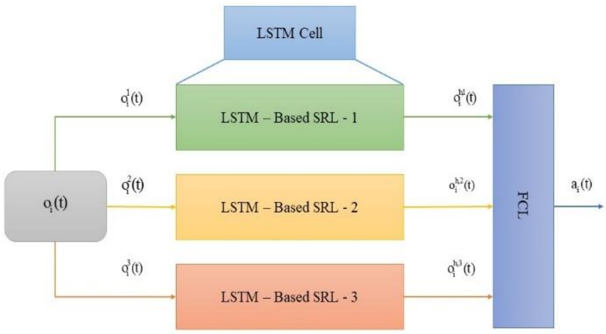

Fig. 1: Structure of an Actor Network.

The UAVs have the ability to make intelligent decisions leveraging reinforcement learning framework. It is the position, the stability of links, the SINR and the queue length of the state that the UAV knows about its surroundings. This information is used by the UAV to determine network condition and modify its behaviour accordingly. Accordingly, the UAV chooses an action according to its current state. They involve actions such as reorienting trajectory, picking out communication frequency and choosing the following relay UAV. There exists a reward for making an optimal decision. The policy maps states to actions, guiding the UAV toward better choices. This learning process improves communication reliability and network stability over time.

The Figure 2 presents the adaptive DMA-DDPG training process for UAV swarm networks. It starts with an online actor network. It interacts with the UAV environment and makes decisions based on current parameters θ_i. This actor network is updated using gradient information $\nabla\theta_i J(\theta_i)$ to improve its decision-making policy. The updated actor parameters are softly copied to a target actor network to maintain stable learning. At the same time, a critic network with parameters ω_i evaluates how good the actions are using a loss function $L(\omega_i)$. Its results help update both the actor and critic networks. Data from nearby UAVs is stored in a replay buffer. It collects experiences used to train the networks. A mini-batch is selected from this buffer for learning. These samples are used to update the networks through backpropagation. The target critic network also receives soft updates to stay close to the online critic. This entire system runs under cooperative training, where all components work together to help UAVs learn better strategies in real-time dynamic environments. The method improves learning stability and promotes coordination among multiple UAVs.

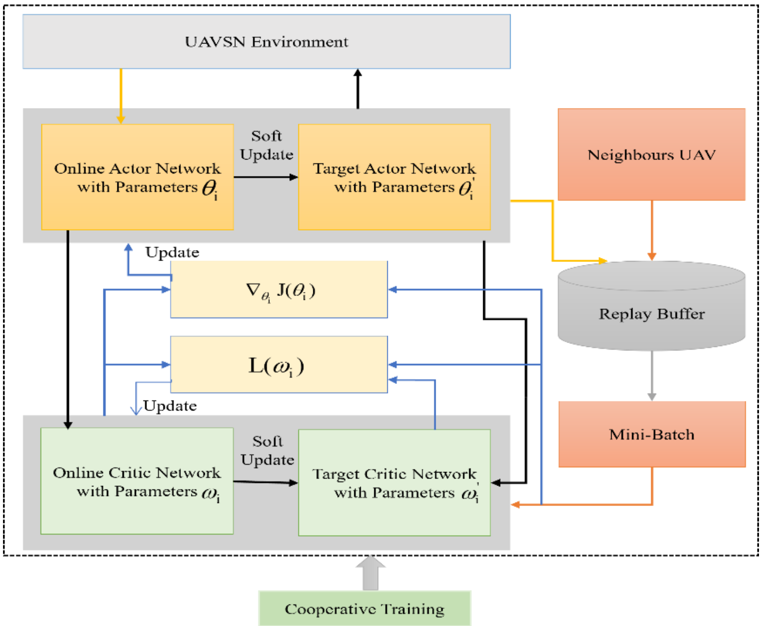

Fig. 2: Flow Chart of Adaptive DMA-DDPG Training Process.

The objective in reinforcement learning is to determine the optimal policy π^* that maximizes expected long-term rewards:

$$\pi^* = \text{argmax}_\pi \, E\left[\sum_{t=0}^{T} \gamma^t r_t\right] \tag{12}$$

Here, γ is the discount factor controlling the importance of future rewards. Deep Deterministic Policy Gradient (DDPG) is an RL algorithm that optimizes continuous action spaces. The DMA-DDPG framework extends DDPG for multi-agent UAV networks. Each UAV independently selects actions while considering neighbouring UAVs' strategies. The DMA-DDPG framework helps UAVs learn better decision-making. The actor network takes the current state and maps it to an action. It provides continuous action values to guide UAVs in selecting the best movement or communication choices. The critic network evaluates these actions by estimating the Q-value. This value represents the expected long-term reward for taking a specific action. The replay buffer stores past experiences for future learning. It prevents overfitting and results in more stable training. The target networks improve the learning process by using soft updates. This helps maintain consistency in training and improves decision accuracy over time. The actor function predicts the best action for a given state:

$$a_t = \mu_\theta(s_t) \tag{13}$$

The critic function evaluates the action's quality using a Q-value estimation:

$$Q_\omega(s_t, a_t) = E\left[r_t + \gamma Q_{\omega'}(s_{t+1}, \mu_{\theta'}(s_{t+1}))\right] \tag{14}$$

Here, θ and ω denote actor and critic network parameters, respectively. The JTFR algorithm optimizes UAV trajectory, frequency allocation and routing. The optimization problem is formulated and given in (10). It is subjected and given in (11). JTFR enhances UAV network efficiency by integrating DMA-DDPG. UAVs dynamically optimize the trajectories and communication channels based on the learned policy. The training process begins by setting up UAV positions, the replay buffer and neural network parameters [32]. Each UAV observes its current state, which includes factors like position, link stability and interference. The actor network then selects an action based on this state. The chosen action involves adjusting trajectory, selecting a frequency or choosing a relay UAV. After executing the action, the UAV receives a reward based on its effectiveness. The experience includes state, action, reward and next state, is stored in the replay buffer. The critic and actor networks are updated using gradient descent to improve future decision-making. Soft update target networks using:

$$\theta' \leftarrow \tau\theta + (1-\tau)\theta' \tag{15}$$

Over multiple iterations the training process converges, allowing UAVs to learn an optimal policy for network optimization. The DMA-DDPG-based JTFR algorithm significantly enhances UAV swarm network efficiency. By using reinforcement learning, UAVs optimize trajectory planning, frequency allocation and routing. The adaptive learning process promotes reduced interference, better network coverage and energy-efficient UAV movement. Future improvements include hybrid learning approaches and enhanced UAV cooperation strategies to improve UAV network adaptability in complex environments.

Algorithm 1: DMA-DDPG-Based JTFR Algorithm	
1.	Initialize UAV positions, replay buffer, actor and critic networks
2.	for each episode do
3.	for each UAV i in the swarm do
4.	Observe current state s_i

Select action $a_i = \mu_\theta(s_i) + N$ (exploration noise)
Execute action and receive reward r_i , observe next state $s_{i'}$
Store transition $(s_i, a_i, r_i, s_{i'})$ in replay buffer
end for
for each UAV i in the minibatch do
Sample random batch from replay buffer
Compute target Q-value: $y_i = r_i + \gamma Q_{\omega'}(s_{i'}, \mu_{\theta'}(s_{i'}))$
Update critic network by minimizing loss: $L(\omega) = (y_i - Q_\omega(s_i, a_i))^2$
Update actor network using policy gradient: $\nabla_\theta J(\theta)$
Update target networks: $\theta' \leftarrow \tau\theta + (1-\tau)\theta'$, $\omega' \leftarrow \tau\omega + (1-\tau)\omega'$
end for
end for

5. Performance evaluation

The performance evaluation of the proposed DMA-DDPG-based JTFR algorithm is conducted through extensive simulations. The study aims to analyze the effectiveness of the algorithm in optimizing network communication and UAV trajectory. Several performance metrics includes packet delivery ratio, end-to-end delay and energy consumption are considered. The results obtained from the simulation are compared with existing routing protocols to highlight the advantages of the JTFR approach.

Parameter	Value / Description
Simulation environment	3D Network Simulator(2500X2500X300)
Simulation Tool	MATLAB 2023a
Number of UAVs	Dynamic, ranging from 50 to 100
Path loss exponent	3
Simulation time	1000 seconds
UAV speed	5 to 20 meters per second
Channel bandwidth	20MHz
Communication range	300 meters
Frequency bands	5 available channels
Packet arrival model	Poisson
Traffic pattern	Constant bit rate
Packet size	512 bytes

The results of the analysis are: JTFR gives a higher packet delivery ratio, lower end-to end delay and communication stability is better. A comparative analysis of PDR and end to end delay of different routing protocols is presented in Table (1).

Table 1: Comparison of PDR and End-to-End Delay

Routing Protocol	PDR	End-to-End Delay
OLSR	75.3	120
Q-Learning	82.1	95
JTFR (Proposed)	91.4	72

The Table (2) shows the great improvement on delivering a best JTFR approach relative to the other OLSR and Q learning approaches. The algorithm offers 91.4% higher packet delivery ratio than the other methods. Likewise, the end-to-end delay is reduced to 72 ms and this increased efficiency of communication. Energy consumption of the UAVs is studied under different routing protocols is another aspect that is analyzed. Results from total energy consumed and network lifetime of the proposed method compared to other methods are tabulated in Table 2.

Table 2: Comparison of Energy Consumption

Routing Protocol	Total Energy Consumed (J)	Network Lifetime (s)
OLSR	1200	800
Q-Learning	950	920
JTFR (Proposed)	730	1100

From the results, it is clear that the energy consumption by the JTFR is effective in reducing the overall power used from 837J to 730J, and prolonging the network lifetime from 1000 seconds to 1100 seconds. Results prove that (1) the JTFR algorithm contributes to improved communication reliability and (2) it simultaneously increases energy efficiency to support extended operation of the UAV in dynamic environments. The analysis of JTFR algorithm performance is proved leading to the improvement of UAV network efficiency. The simulation results show that JTFR improves the packet delivery ratio, shorten the communication delay and optimize energy consumption. The algorithm achieves stable network connectivity and extends the operational lifetime of UAVs, making it a viable solution for dynamic UAV swarm networks. Future research explore hybrid learning models to enhance UAV cooperation and network adaptability in complex scenarios.

Packet delivery ratio with respect to the number of UAVs is showed in the Figure 2 for four algorithms: JTFR (proposed), DMA-DDPG-1, MA-DDPG-LSTM and MCA-OLSR. For all algorithms, with the increase of the number of UAVs, the packet delivery ratio decreases because of the more complex topology and hence of the longer communication paths. The optimal value of JTFR algorithm is best. With 50 UAVs, it starts at over 90 percent packet delivery ratio which slightly decreases with a slightly higher ratio compared to the other algorithms when the count of UAVs goes up to 100. With 50 UAVs, DMA-DDPG-1 achieves a packet delivery ratio around 87 percent that counts down as more UAVs are included in the network. However, MA-DDPG LSTM starts at a packet delivery ratio of 85 percent and it less than 80 percent packet delivery ratio as the number of UAVs becomes 100. At 75 percent of 50 UAVs, MCA-OLSR becomes the worst performing and its performance decreases as the number of UAVs increases. Based on the results, all algorithms exhibit downward trend, and higher UAV density causes congestions, interferences and packet loss which decreased delivery ratio negatively. It has optimized routing and frequency allocation strategy, because of which it is maintaining better packet delivery ratio. Better communication reliability is even maintained as the network congestion increase.

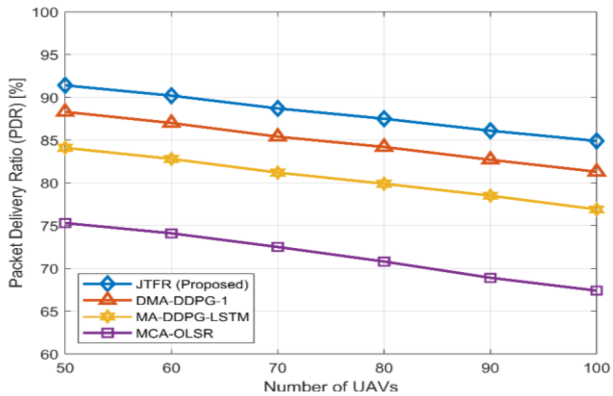

Fig. 2: PDR Versus Number of UAVs.

The Figure 3 shows the end-to-end delay versus the number of UAVs for different algorithms, including JTFR, DMA-DDPG-1, MA-DDPG-LSTM and MCA-OLSR. The end to end delay increases for all algorithms as the number of UAVs increases. Starting 72 milliseconds for 50 UAVs, and with only a small increase to almost 90 milliseconds for 100 UAVs, the JTFR algorithm has the lowest delay. With delay starting around 88 milliseconds with 50 UAVs and close to 105 milliseconds with 100 UAVS, the next algorithm is the DMA-DDPG-1. When evaluating a 50 or 100 UAV scenario, the MA-DDPG-LSTM algorithm has higher delay compared to JTFR and DMA-DDPG-1 around 100 to 120 milliseconds. In case of 50 UAVs the MCA-OLSR algorithm carries the highest delay at 120 milliseconds, and it increases till 150 milliseconds for 100 UAVs. All algorithms take a delay that increases with the number of UAVs. Higher congestion in the network is resulted from more UAVs. With the increasingly number of data traffic, queuing delays also increase. This leads to higher end to end latency. However, its efficient routing and frequency allocation, make JTFR algorithm have the lowest delay. It helps to decrease network traffic and improves the transmission efficiency. When performing slightly worse than JTFR algorithm, the DMA-DDPG-1 algorithm still has less delay than MA-DDPG-LSTM and MCA-OLSR algorithms. The delay of the MA-DDPG-LSTM algorithm steadily increases. It implies that more number of UAVs, the longer time needed for routing and decision making. The delay of MCA-OLSR is highest because of its static routing strategy. This results in inefficient selection of the path and longer transmission time.

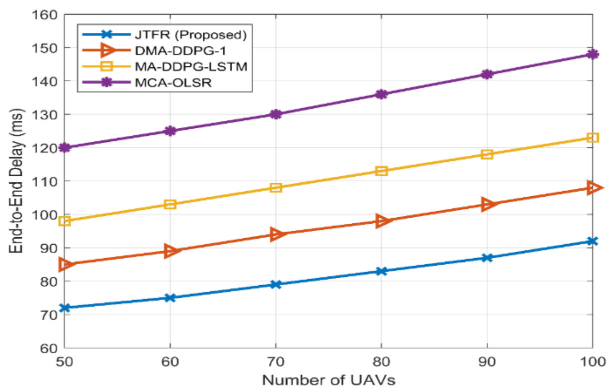

Fig. 3: End-To-End Delay Versus Number of UAVs.

Figure 4 presents the normalized control overhead scape versus time over simulations with four algorithms: JTFR, DMA-DDPG-1, MA-DDPG- LSTM and MCA-OLSR. For all algorithms, the control overhead also increases with simulation time. Among all, the control overhead for JTFR is the lowest. At the beginning, when t=0 it rises to 0.12, then linearly increasing up to 0.18 at t=0,1000. We see that the DMA-DDPG-1 algorithm begins above the other, 0.15 at 0 and 0.23 at 1000. The control overhead of the MA-DDPG-LSTM algorithm increases more sharply. It begins at just under 0.19 at 0 seconds and proceeds up to around 0.32 at 1000 seconds. From 0 sec onwards, the MCA - OLSR algorithm has the highest overhead, which increases almost till 1 sec to almost 0.45, while the overhead ends up being 0.22 at 0 sec. Even all algorithms exhibit a non-decreasing trend, meaning that packets are generated more and more as the simulation time increases, adding overhead. The routing and adaptive control packet generation of the JTFR algorithm is efficient which results in it having the lowest overhead. Due to which it has frequent control packet exchanges, the MCA-OLSR algorithm exhibits the highest overhead and, therefore, is less efficient for large-scale UAV networks.

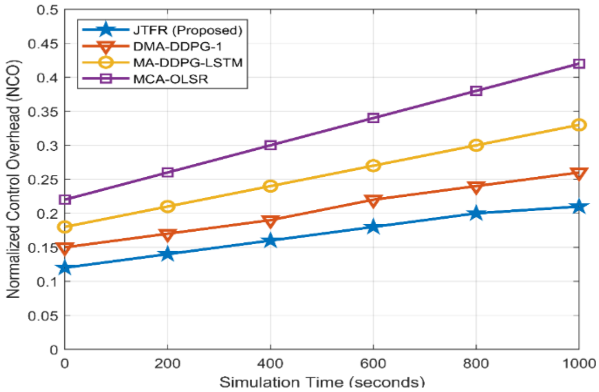

Fig. 4: Normalized Control Overhead Versus Simulation Time.

The result is shown in the Figure 5 where normalized residual energy is plotted against the number of UAVs by using four different algorithms namely, JTFR, DMA-DDPG-1, MA-DDPG-LSTM and MCA-OLSR. For all algorithms, the normalized residual energy also decreases when the number of UAVs increases. Else, it maintains the highest residual energy. However, it begins below 0.88 for 50 UAVs and decreases further to 0.80 for 100 UAVs. The DMA-DDPG-1 algorithm starts from around 0.78 for 50 UAVs and then reduces almost to 0.70 for 100 UAVs. For 50 UAVs, the residual energy largely declines from about 0.72, whereas it declines less drastically but buckles down from 0.63 for 100 UAVs with the MA-DDPG-LSTM algorithms. In this sense, MCA-OLSR has the lowest residual energy throughout. For 50 UAVs, it starts around 0.65 and decreases significantly to almost 0.40 for 100 UAVs. Further, the fact that the trend is downward across all algorithms indicates that as the number of UAVs operating in the network increases, energy consumption increases as more and more communication overhead as well as data transmission and network path adjustments are required. Because its respective trajectory planning and dynamic resource allocation are optimized, the JTFR algorithm will preserve the most energy. MCAOLSR has the highest energy depletion, its routing is static and less efficient and hence it is not suitable for UAV networks at large scale.

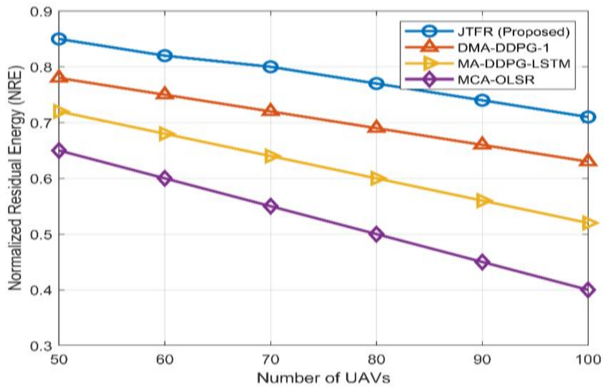

Fig. 5: Normalized Residual Energy Versus Number of UAVs.

The Figure 6 presents the average reward versus the number of episodes for four different algorithms: JTFR, DMA-DDPG-1, MA-DDPG-LSTM and MCA-OLSR. As the number of episodes increases the average reward also increases for all algorithms. The JTFR algorithm achieves the highest reward starting at around 5 for 0 episodes and reaching 98 at 5000 training episodes. The DMA-DDPG-1 algorithm follows closely, beginning at nearly 4 for 0 episodes and rising to about 94 at 5000 training episodes. The MA-DDPG-LSTM algorithm shows a slightly lower reward trajectory. It starts at 3 for 0 episodes and increasing steadily to around 90 at 5000 training episodes. The MCA-OLSR algorithm has the lowest reward among all. It begins at nearly 2 for 0 episodes and gradually increasing to about 82 at 5000 training episodes. The overall trend suggests that all algorithms improve with training, but the JTFR algorithm learns faster and achieves better results. The MCA-OLSR algorithm lags in performance, indicating slower learning and less efficient decision-making. The figure highlights the advantage of reinforcement learning-based approaches in achieving higher rewards through continuous training.

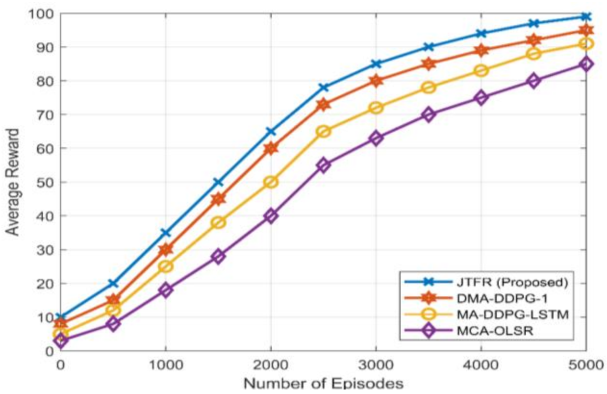

Fig. 6: Average Reward Versus Number of Episodes.

The Figure 7 presents the traveling distance fairness versus the number of UAVs for four different algorithms: JTFR, DMA-DDPG-1, MA-DDPG-LSTM and MCA-OLSR. As the number of UAVs increases the fairness index decreases for all algorithms. The JTFR algorithm achieves the highest fairness starting at 0.92 for 50 UAVs and gradually decreasing to about 0.85 for 100 UAVs. The DMA-DDPG-1 algorithm follows closely with fairness values starting at 0.90 for 50 UAVs and reducing to nearly 0.83 for 100 UAVs. The MA-DDPG-LSTM algorithm shows a lower fairness trend beginning at around 0.85 for 50 UAVs and decreasing steadily to about 0.78 for 100 UAVs. The MCA-OLSR algorithm exhibits the lowest fairness among all starting at 0.78 for 50 UAVs and decreasing to nearly 0.65 for 100 UAVs. The downward trend in all algorithms shows that fairness decreases as more UAVs operate. Increased congestion makes maintaining fairness more difficult. Mobility adjustments also affect movement distribution. The JTFR algorithm keeps the highest fairness with balanced trajectory planning. It provides equal movement among UAVs. The MCA-OLSR algorithm has the lowest fairness. Its static path selection leads to uneven UAV movements. This causes higher disparities in travel distances.

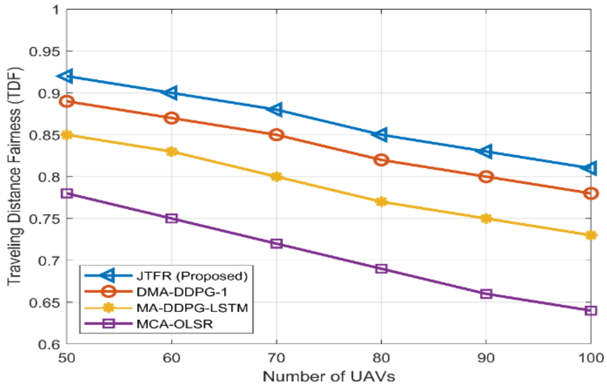

Fig. 7: Travelling Distance Fairness Versus Number of UAVs.

The Figure 8 presents the signal-to-interference-plus-noise ratio versus simulation time for four different algorithms: JTFR, DMA-DDPG-1, MA-DDPG-LSTM and MCA-OLSR. As simulation time increases the SINR improves for all algorithms. At 0 seconds and beyond, the SINR achieved with the JTFR algorithm is the highest, beginning around 25 dB and increasing up to around 35 dB at 1000 seconds. Following the start around 22 dB and move almost to 32 dB at 1000 seconds, the DMA-DDPG-1 algorithm results are obtained. The trend in the SINR for the MA-DDPG-LSTM algorithm is lower. Its onset is at around 18 dB and rise to roughly 28 dB at 1000 s. Starting at around 15 dB and ending at close to 25 dB at 1000 seconds, the MCA-OLSR algorithm shows the lowest SINR throughout. An upward trend is also seen in all algorithms, indicating that as the simulation proceeds, learning based methods make the interference mitigation better, which leads to better signal quality. As the JTFR algorithm has an efficient frequency allocation and interference management, it maintains the highest SINR. Among all the five algorithms, the MCA-OLSR algorithm has the lowest SINR value, suggesting that its static routing method cannot alleviate the interference and keep the data communication links stable.

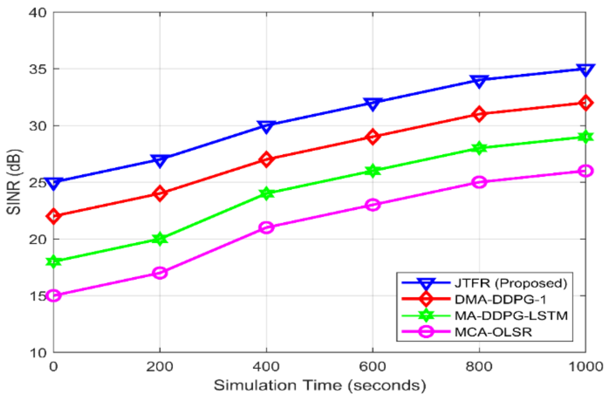

Fig. 8: SINR Versus Simulation Time.

The Figure 9 shows the convergence rate versus the number of episodes for four different algorithms: JTFR, DMA-DDPG-1, MA-DDPG-LSTM and MCA-OLSR. As the number of episodes increases the convergence rate also increases for all algorithms. The JTFR algorithm achieves the fastest convergence starting at nearly 0 at 0 episodes and reaching almost 1 at 5000 training episodes. The DMA-DDPG-1 algorithm follows closely showing a similar trend but with a slightly lower convergence rate throughout the training period. The MA-DDPG-LSTM algorithm exhibits a slower convergence pattern compared to JTFR and DMA-DDPG-1 taking longer to approach a stable value near 1. The MCA-OLSR algorithm has the slowest convergence, with the lowest values at all episodes. It indicates a less efficient learning process. The general upward trend in all algorithms suggests that as more episodes are completed, learning improves and the models become more effective. The JTFR algorithm reaches higher convergence levels earlier due to its optimized reinforcement learning framework. The MCA-OLSR algorithm converges the slowest, highlighting the limitations of its learning process and adaptation speed.

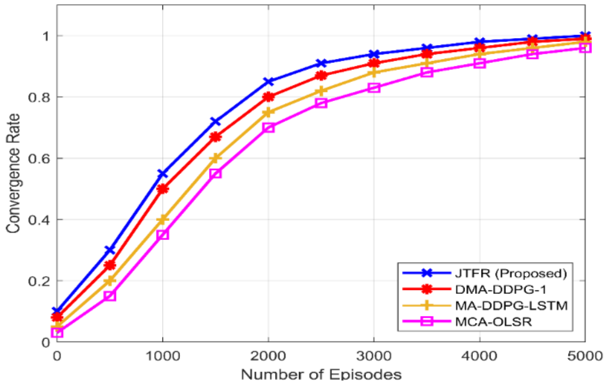

Fig. 9: Convergence Rate Versus Number of Episodes.

The Figure 10 presents throughput versus simulation time for four different algorithms: JTFR, DMA-DDPG-1, MA-DDPG-LSTM and MCA-OLSR. For all the algorithms, the throughput increases with simulation time. At 0 seconds, the JTFR algorithm also reaches a highest throughput of 2.5 Mbps starting at 0 seconds and rising to close to 5.5 Mbps at 1000 seconds. Starting from barely 2.2 Mbps for the DMA-DDPG-1 algorithm in 2.2 0 seconds, it comes within 4.8 Mbps at 1000 seconds. The trend on MA-DDPG-LSTM Algorithm is of lower throughput. At roughly 1.9 Mbps, it continues up to around 4.2 Mbps after 1000 seconds. The one with the lowest throughput throughout has been I MCA-OLSR algorithm. At around 1.5 Mbps and increasing nearly to 3.8 Mbps after 1000 seconds. The trend in most of the algorithms displays the tendency for learning based methods to better the data transmission efficiency of the network compared to other methods as the simulation goes on. Because its routing and frequency allocations are also optimized, the JTFR algorithm retains the highest throughput. As a result, the MCA-OLSR algorithm has the lowest throughput which confirms that its static routing method fails to transmit high data transmission rate.

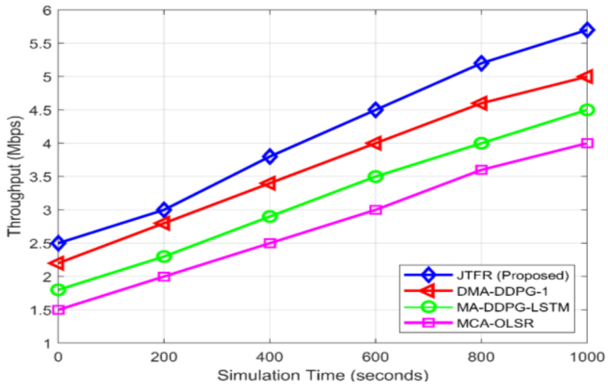

Fig. 10: Throughput Versus Simulation Time.

The Figure 11 presents the packet loss ratio versus the number of UAVs for four different algorithms: JTFR, DMA-DDPG-1, MA-DDPG-LSTM and MCA-OLSR. As the number of UAVs increases the packet loss ratio also increases for all algorithms. The JTFR algorithm has the lowest packet loss ratio. It starts at around 9 percent for 50 UAVs. It gradually increases to about 14 percent for 100 UAVs. The DMA-DDPG-1 algorithm follows with a slightly higher packet loss. It begins at around 11 percent for 50 UAVs. It increases to nearly 17 percent for 100 UAVs. The MA-DDPG-LSTM algorithm has a higher packet loss ratio. It starts at about 14 percent for 50 UAVs. It increases to around 22 percent for 100 UAVs. The MCA-OLSR algorithm has the highest packet loss. It begins at nearly 25 percent for 50 UAVs. It steadily rises to about 32 percent for 100 UAVs. The increasing trend in all algorithms shows higher packet loss with more UAVs. Network congestion and interference contribute to this rise. Transmission failures also lead to a higher packet loss ratio. The JTFR algorithm maintains the lowest packet loss ratio due to its efficient routing and resource allocation strategies. The MCA-OLSR algorithm shows the highest packet loss ratio, indicating that its static routing approach struggles to handle larger UAV networks efficiently.

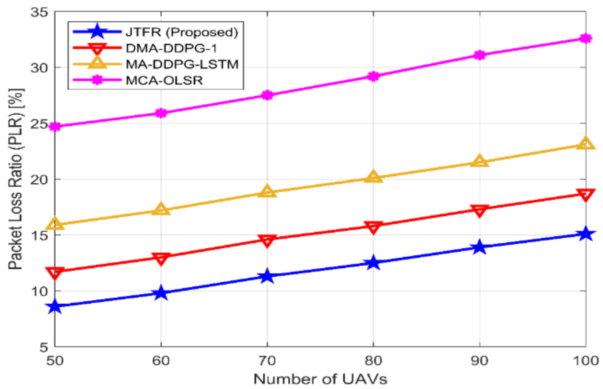

Fig. 11: PLR Versus Number of UAVs.

The Figure 12 presents energy consumption versus the number of UAVs for four different algorithms: JTFR, DMA-DDPG-1, MA-DDPG-LSTM and MCA-OLSR. As the number of UAVs increases, energy consumption also increases for all algorithms. The JTFR algorithm achieves the lowest energy consumption. It starts at around 720 joules for 50 UAVs and increasing steadily to about 900 joules for 100 UAVs. The DMA-DDPG-1 algorithm follows, with energy consumption beginning at 850 joules for 50 UAVs and increasing to nearly 1050 joules for 100 UAVs. The MA-DDPG-LSTM algorithm uses more energy than JTFR and DMA-DDPG-1. It starts at around 980 joules for 50 UAVs. It increases to about 1200 joules for 100 UAVs. The MCA-OLSR algorithm has the highest energy consumption. It begins at 1200 joules for 50 UAVs. It rises steadily to about 1450 joules for 100 UAVs. The increasing trend in all algorithms shows that more UAVs lead to higher energy use. Communication overhead and mobility adjustments contribute to this rise. Transmission demands also increase energy consumption. The JTFR algorithm has the lowest energy use. It benefits from optimized trajectory planning and resource allocation. The MCA-OLSR algorithm shows the highest energy consumption, indicating that its static routing approach leads to inefficient energy usage in larger UAV networks.

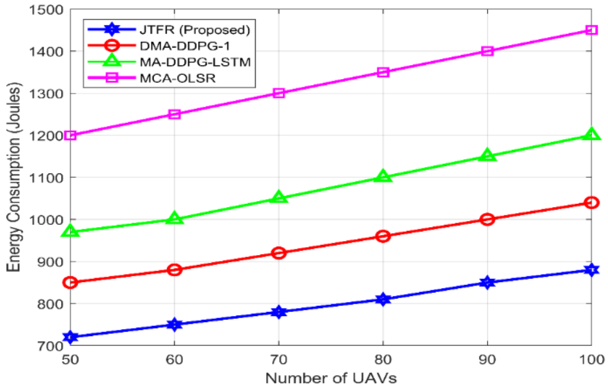

Fig. 12: Energy Consumption Versus Number of UAVs.

The Figure 13 presents training loss versus the number of episodes for four different algorithms: JTFR, DMA-DDPG-1, MA-DDPG-LSTM and MCA-OLSR. As the number of episodes increases, training loss decreases for all algorithms. The JTFR algorithm achieves the lowest training loss, starting at 1.1 at 0 episodes and steadily decreasing to nearly 0 at 5000 training episodes. The DMA-DDPG-1 algorithm follows, starting at around 1.2 and gradually reducing to a value close to JTFR at the final training stage. The MA-DDPG-LSTM algorithm has a slightly higher training loss. It starts at about 1.4 at 0 episodes. It decreases to around 0.1 at 5000 training episodes. The MCA-OLSR algorithm has the highest training loss. It begins at nearly 1.6 at 0 episodes. It reduces slowly to about 0.2 at 5000 training episodes. The decreasing trend across all algorithms suggests that as more training episodes are completed, the models learn better and become more accurate. The JTFR algorithm converges faster and achieves lower loss due to its optimized learning strategy. The MCA-OLSR algorithm shows the slowest reduction in loss, indicating that it requires more episodes to reach stability and efficiency in decision-making.

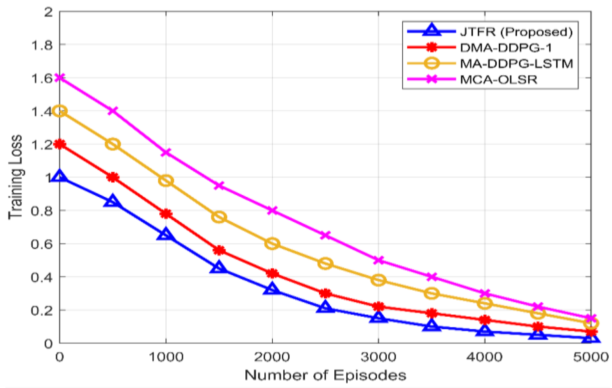

Fig. 13: Training Loss Versus Number of Episodes.

Four different algorithms, namely JTFR, DMA-DDPG-1, MA-DDPG-LSTM and MCA-OLSR are used to present network latency versus number of UAVs, as shown in Figure 14. For all algorithms, when increasing the number of UAVs, network latency also increases. The latency of the JTFR algorithm is lowest. For 50 UAVs, it initially starts at only about 40 milliseconds. The time increases gradually to about 60ms for 100 UAVs. The DMA-DDPG-1 algorithm follows closely. For 50 UAVs, it starts at about 50mS. For 100 UAVs, it rises to nearly 70 milliseconds. The latency trend of the MA-DDPG LSTM algorithm is higher. For 50 UAVs, it starts around 60 milliseconds. As the number of UAVs increases, it grows in a relatively steady way reaching about 80 milliseconds for 100 UAVs. Latency of the MCA-OLSR algorithm is the highest. For 50 UAVs it starts at almost 70 milliseconds. However, its rise is quite dramatic to 95 milliseconds when 100 UAVs are used. The trend for all algorithms is that increasing UAV number causes lack of time for drones to fly back and forth, thus increasing latency. The effective routing and decision making, what ensures the JTFR algorithm has the lowest latency. From latency percentage, the most latency is exhibited by the MCA–OLSR, implying that its static approach is not a good choice for large UAV networks.

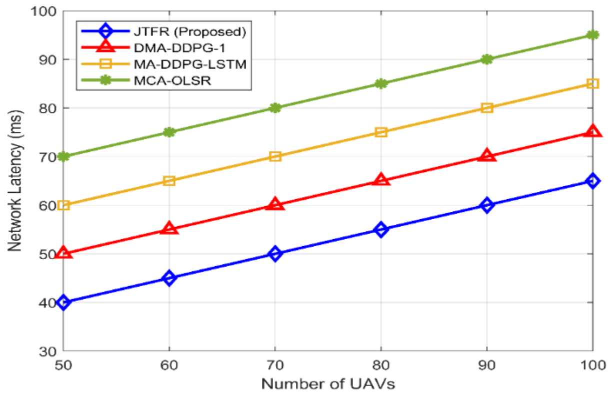

Fig. 14: Network Latency Versus Number of UAVs.

6. Conclusion

An efficient problem solution for UAV swarm networks is proposed in this thesis, employing the optimization of the trajectory, the frequency allocation and the routing of the vehicles by means of a DMA-DDPG baseline JTFR algorithm. Dynamic network topologies and interference are not amenable to traditional approaches. This thesis presents a learning-based framework in which UAVs adapt to their decisions through real time observations. JTFR considerably improves network performance with regard to packet delivery, low latency and efficient energy consumption. The approach based on reinforcement learning enables UAVs to autonomously take decisions so as to maximized network efficiency. The algorithm integrates DDPG within a distributed multi agent system which gives better adaptivity to topology changes. Optimal UAV trajectories are learned while minimizing communication overhead. It also resolves the frequency allocation problems through minimizing the interference and hence obtaining a stable communication link. Thus, we evaluate JTFR with traditional routing protocols such as OLSR and Q learning based routing and find improvements. It demonstrates higher packet delivery ratios and less end-to-end delays. Additionally, it improves energy consumption to improve overall networking lifetime. The system is more practical for real world applications, as UAVs can operate longer durations before recharging will be required. Data transmission reliability of UAVs to identify the optimal relay nodes. The learning-based trajectory optimization of the robot reduces unnecessary movement from the robot which leads to an energy wastage. All these factors make the JTFR framework robust. Future research should concentrate on increasing the model by adding hybrid learning techniques. A combination of supervised learning and of reinforcement learning assists the decision making. It is shown that the proposed JTFR algorithm optimizes UAV network performance successfully.

Acknowledgement

The authors would like to express their sincere gratitude to all those who contributed to the success of this research work.

References

[1] A. Khan, S. Gupta, and S. K. Gupta, “Emerging uav technology for disaster detection, mitigation, response, and preparedness,” Journal of Field Robotics, vol. 39, no. 6, pp. 905–955, 2022. https://doi.org/10.1002/rob.22075.

[2] Z. Feng, L. Ji, Q. Zhang, and W. Li, “Spectrum management for mmwave enabled uav swarm networks: Challenges and opportunities,” IEEE Communications Magazine, vol. 57, no. 1, pp. 146–153, 2018. https://doi.org/10.1109/MCOM.2018.1800087.

[3] W. Chen, J. Liu, H. Guo, and N. Kato, “Toward robust and intelligent drone swarm: Challenges and future directions,” IEEE Network, vol. 34, no. 4, pp. 278–283, 2020. https://doi.org/10.1109/MNET.001.1900521.

[4] L. Liu, S. Zhang, and R. Zhang, “Multi-beam uav communication in cellular uplink: Cooperative interference cancellation and sum-rate maximization,” IEEE Transactions on Wireless Communications, vol. 18, no. 10, pp. 4679–4691, 2019. https://doi.org/10.1109/TWC.2019.2926981.

[5] Z. Xiao, L. Zhu, L. Bai, and X.-G. Xia, Array beamforming enabled wireless communications. CRC Press, 2023. https://doi.org/10.1201/9781003366362.

[6] S. Javaid, N. Saeed, Z. Qadir, H. Fahim, B. He, H. Song, and M. Bilal, “Communication and control in collaborative uavs: Recent advances and future trends,” IEEE Transactions on Intelligent Transportation Systems, vol. 24, no. 6, pp. 5719–5739, 2023. https://doi.org/10.1109/TITS.2023.3248841.

[7] D.-G. Zhang, L. Chen, J. Zhang, J. Chen, T. Zhang, Y.-M. Tang, and J.N. Qiu, “A multi-path routing protocol based on link lifetime and energy consumption prediction for mobile edge computing,” IEEE Access, vol. 8, pp. 69058–69071, 2020. https://doi.org/10.1109/ACCESS.2020.2986078.

[8] X. Cao, B. Yang, K. Wang, X. Li, Z. Yu, C. Yuen, Y. Zhang, and Z. Han, “Ai-empowered multiple access for 6g: A survey of spectrum sensing, protocol designs, and optimizations,” Proceedings of the IEEE, 2024. https://doi.org/10.1109/JPROC.2024.3417332.

[9] Z. Lin, Z. Zeng, Y. Yu, Y. Ren, X. Qiu, and J. Chen, “Graph attention residual network based routing and fault-tolerant scheduling mechanism for data flow in power communication network.,” Computers, Materials & Continua, vol. 81, no. 1, 2024. https://doi.org/10.32604/cmc.2024.055802.

[10] L. Gupta, R. Jain, and G. Vaszkun, “Survey of important issues in uav communication networks,” IEEE communications surveys & tutorials, vol. 18, no. 2, pp. 1123–1152, 2015. https://doi.org/10.1109/COMST.2015.2495297.

[11] B. Chen, K. Jamieson, H. Balakrishnan, and R. Morris, “Span: An energy-efficient coordination algorithm for topology maintenance in ad hoc wireless networks,” in Proceedings of the 7th annual international conference on Mobile computing and networking, pp. 85–96, 2001. https://doi.org/10.1145/381677.381686.

[12] P. Mishra and G. Singh, “Energy management systems in sustainable smart cities based on the internet of energy: A technical review,” Energies, vol. 16, no. 19, p. 6903, 2023. https://doi.org/10.3390/en16196903.

[13] M. M. Alam and S. Moh,” Joint Trajectory Control, Frequency Allocation, and Routing for UAV Swarm Networks: A Multi-Agent Deep Reinforcement Learning Approach,” IEEE Transactions on Mobile Computing, vol. 23, no. 12, pp. 11989-12005, Dec. 2024. https://doi.org/10.1109/TMC.2024.3403890.

[14] J. Smith and K. Patel,” Behavior-Based UAV Trajectory Optimization,”IEEE Transactions on Wireless Communications, vol. 21, no. 6, pp. 3543-3556, Jun. 2023.

[15] L. Zhang, X. Li, and M. Chen,” Swarm Intelligence for UAV Networks,” IEEE Communications Letters, vol. 26, no. 8, pp. 1564-1568, Aug. 2023.

[16] H. Wang and Y. Zhao," Dynamic Frequency Selection for UAV Communication," IEEE Internet of Things Journal, vol. 10, no. 3, pp. 2156-2168, Mar. 2024.
[17] P. Kumar and S. Lee," Adaptive Frequency Reuse for UAV Networks," IEEE Transactions on Networking, vol. 32, no. 5, pp. 1457-1470, May 2023.
[18] A. Brown and T. Davis," Machine Learning-Based Frequency Allocation for UAVs," IEEE Transactions on Cognitive Communications, vol. 16, no.2, pp. 890-902, Apr. 2023.
[19] C. Wilson and M. Garcia," Multi-Hop Routing in UAV Networks," IEEE Transactions on Vehicular Technology, vol. 72, no. 4, pp. 3301-3315, Apr. 2024.
[20] N. Fernandez and R. Clark," Position-Based Routing for UAVs," IEEE Transactions on Communications, vol. 33, no. 6, pp. 987-1001, Jun. 2024.
[21] B. Johnson and H. Kim," Predictive Routing Using Machine Learning in UAV Networks," IEEE Transactions on Networking, vol. 29, no. 9, pp. 3564-3578, Sep. 2023.
[22] Vijay, S. T. Siddiqui, Ritu, R. A. Kumar, A. Kumar, S. Umamaheswararao, N. Dudiki, Y. V. Reddy, and K. Dekeba, "Intertwine connection-based routing path selection for data transmission in mobile cellular networks and wireless sensor networks," Wireless Communications and Mobile Computing, vol. 2022, no. 1, p. 8398128, 2022. https://doi.org/10.1155/2022/8398128.
[23] E. Roberts and V. Carter," Cross-Layer Optimization for UAV Networks," IEEE Transactions on Wireless Communications, vol. 22, no. 7, pp. 4875-4888, Jul. 2023.
[24] R. K. Bharti, S. S, S. V. Sumant, C. A. D. Durai, R. A. Kumar, K. Singh, H. Palivela, B. R. Kumar, and B. Debtera, "Enhanced path routing with buffer allocation method using coupling node selection algorithm in MANET," Wireless Communications and Mobile Computing, vol. 2022, no. 1, p. 1955290, 2022. https://doi.org/10.1155/2022/1955290.
[25] M. Thompson and L. White," Deep Reinforcement Learning for UAV Optimization," IEEE Transactions on Mobile Computing, vol. 20, no. 5, pp.2678-2692, May 2024.
[26] R. Gowtham, V. Anand, Y. V. Suresh, K. L. Narasimha, R. A. Kumar, and V. Saraswathi, "Enhancing incentive schemes in edge computing through hierarchical reinforcement learning," Journal of Engineering and Technology for Industrial Applications, vol. 11, no. 52, p. 226 – 236, 2025. https://doi.org/10.5935/jetia.v11i52.1637.
[27] R. Hall and S. Young," Collaborative Decision-Making in UAV Swarm Networks," IEEE Transactions on Intelligent Systems, vol. 31, no. 3, pp. 1472-1485, Mar. 2024.
[28] F.-H. Tseng and Y.-J. Hsieh, "Delanalty minimization with reinforcement learning in uav-aided mobile network," IEEE Transactions on Computational Social Systems, vol. 11, no. 2, pp. 1991–2001, 2024. https://doi.org/10.1109/TCSS.2023.3283016.
[29] S. Javed, A. Hassan, R. Ahmad, W. Ahmed, R. Ahmed, A. Saadat, and M. Guizani, "State-of-the-art and future research challenges in UAV swarms," IEEE Internet of Things Journal, vol. 11, no. 11, pp. 19023–19045, 2024. https://doi.org/10.1109/JIOT.2024.3364230.
[30] O. A. Amodu, C. Jarray, R. Azlina Raja Mahmood, H. Althumali, U. Ali Bukar, R. Nordin, N. F. Abdullah, and N. Cong Luong, "Deep reinforcement learning for aoi minimization in uav-aided data collection for WSN and IOT applications: A survey," IEEE Access, vol. 12, pp. 108000–108040, 2024. https://doi.org/10.1109/ACCESS.2024.3425497.
[31] S. Lee, T.-W. Ban, and H. Lee, "Network-wide energy-efficiency maximization in UAV-aided IOT networks: Quasi-distributed deep reinforcement learning approach," IEEE Internet of Things Journal, vol. 12, no. 11, pp. 15404– 15414, 2025. https://doi.org/10.1109/JIOT.2025.3532477.
[32] S. Fatima, M. Subba Raju, R. Anil Kumar, V. Anand, Y. K. S. S. Rao, D. Kishore, U. Mahalaxmi, and K. Kalyani, "A neural network approach to beam selection and power optimization in mm wave massive MIMO," Journal of Mechanics of Continua and Mathematical Sciences, vol. 20, no. 4, p. 30–52, 2025. https://doi.org/10.26782/jmcms.2025.04.00003.