International Journal of Information Systems and Computer Technologies (IJISCT) Vol. 5, No.1
ISSN: 2791-3635(Print), 2791-3643(Online). DOI: https://doi.org/10.58325/ijisct.004.02.00140

Optimizing Throughput and Quality of Service in RIS-Assisted UAV Systems Using Deep Q-Network (DQN)

Ameer Ullah Khan[1], Muhammad Bilal Qureshi[2], Muhammad Shuaib Qureshi[3], Mohammad Faizan[4], Nauman Zareen[5], Azam Khan[6], Aamir Malik[7]

[1]Department of Computer Science, University of Science & Technology Bannu KPK, Pakistan
[2,4]Department of Computer Science & IT, University of Lakki Marwat, KPK, Pakistan
[3,7]School of Computing Sciences, Pak-Austria Fachhochschule Institute of Applied Sciences and Technology, Haripur, Pakistan.
[5]Department of computer science & IT, Khushal Khan Khattak University Karak, Pakistan
[6]Department of Computer Science & Information Technology, University of Balochistan Quetta, Pakistan.

Corresponding Author: Muhammad Bilal Qureshi* (muhdbilal.qureshi@gmail.com)

ABSTRACT
A DQN-based algorithm is presented in this paper for maximizing throughput and providing quality of service (QoS) in an RIS-assisted UAV system. The system includes a UAV-BS coordinating multiple mobile users in a dynamic 3D environment. By utilizing DQN to dynamically optimize the trajectory, transmission power, and user scheduling online, the system is able to accommodate the changes of the environment and to achieve the maximum performance while keeping QoS. The proposed system model integrates an RIS element to optimize signal quality and coverage, especially in NLoS cases, by additionally enabling diversity. The performance of the proposed method is coordinated with the throughput maximization, the power efficiency , and the QoS guarantee in extensive simulations against traditional optimization techniques. It is shown that the proposed scheme has better adaptability and efficiency compared to traditional schemes in a complex and dynamic communication environment.

Keywords: UAVs, RIS, Deep Q-Network, Throughput Optimization, QoS, Reinforcement Learning, Trajectory Optimization, Power Allocation, User Scheduling, Wireless Systems

1. INTRODUCTION

The rapid development of Unmanned Aerial Vehicles (UAVs) and Reconfigurable Intelligent Surfaces (RIS) has opened up new opportunities in mobile communications, especially in scenarios with limited or unavailable infrastructure. UAVs with the role of MBSs (UAV-BS) enable continuous coverage for ground users with geographical isolation. Meanwhile, as an innovative technology, RIS has the potential ability to improve the quality of communication effectively by passively adjusting the reflecting signals between UAVs and UEs. This method is proven to be particularly useful in areas with a lack of direct LoS communication. However, how to bring the throughput to the optimal and keep high quality of service (QoS) in such dynamic environments becomes very challenging since user mobility, UAV movement, and channel variations cause very uncertain situations [1].

Traditionally, to supply and expand the network coverage of the remote areas and to increase the network capacity, UAVs were used. Such systems have shown encouraging prospects in regions with the absence of terrestrial infrastructure [2, 3]. However, because the area of operation condition of UAV is uncertain, and the variation rule of the area of operation is uncertain, the communication efficiency depends greatly on intelligent task scheduling and real-time scheduling. In addition, RIS-assisted UAV systems have been considered as a novel solution for boosting the received signal power, suppressing interference, and optimizing coverage by controlling the reflected electromagnetic waves smartly. Such a reconfiguration of the electromagnetic field is particularly beneficial when a line-of-sight (LoS) path between the UAV and the user does not exist, but a strong signal strength can still be supported via non-LoS (NLoS) links [4, 5].

The objective of this paper is to maximize the throughput for DL users subject to the QoS requirement in a RIS-assisted UAV system. To this end, we use the Deep Q-Network (DQN) algorithm, which is a reinforcement learning (RL) method presented in [6], to jointly optimize the UAV trajectory, the transmission power, and the user scheduling over time. This attempt wants to satisfy not only for maximum throughput [7], but also to satisfy each user's QoS

requirements. DQN is of great application value in such dynamic systems since its ability of learning flexible policies and decision-making with no delay, which is vital for complex systems such as UAV communication networks [8]. It is also important to note that the established policy must constantly interact with the environment in order to optimize and produce the maximum system parameter. DQN is a variant of Q-learning based on deep neural networks to learn Q-values, which is effective in solving problems with high-dimensional state and action spaces, such as UAV trajectory optimization, power control, and user scheduling in RIS-assisted communication systems [9, 10]. This approach enables the UAV to learn from its experiences and take actions that optimize the performance of the system and provide QoS.

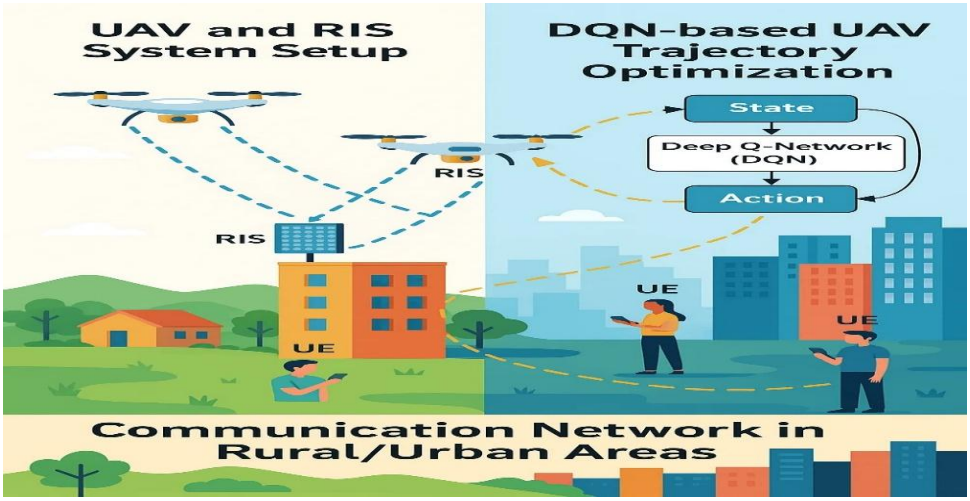


Figure 1. UAV and RIS system optimized using DQN for signal transmission

1.1 Motivation

The RIS introduced, integrated with the unmanned aerial vehicle (UAV) systems, and has drawn tremendous attention due to the capability to improve the communication efficiency, especially in environments such as urban and rural areas where infrastructure is scarce. RIS can manipulate electromagnetic waves to enhance signal intensity for mobile users, for example, with cellphones in an area in which the conventional network infrastructure is insufficient or missing completely. Despite so, the highly dynamic nature of the system and the random mobility of the users along with the repositioning scheme of the UAV pose challenges to sustain the maximum throughput and to meet the QoS requirements of all the users. Furthermore, UAVs are subject to limited energy backup and transmission power; thus, resource allocation and intelligent scheduling play a key role in reducing energy consumption while improving system performance. Such challenges indicate that traditional optimization techniques can be inappropriate in dynamic environments. Deep Q-Network (DQN) offers a potential approach to address these issues by learning a policy that explores and exploits optimally while doing optimization. RL, with DQN in specific, has been shown to be successful for the optimization of UAV trajectory, transmission power, and resource allocation, especially when conventional methods are infeasible. Therefore, this work aims to maximize throughput and guarantee QoS in an RIS-aided UAV system by actively adjusting the UAV trajectory, transmit power, and user scheduling online with a Deep Q-Network (DQN). The objective is to maximize throughput subject to minimum QoS for users and minimum power consumption. The mobility of the UAV system and users, as well as the consideration of the RIS for the improvement in the signal quality, make the optimization problem complex. At each time step, the UAV must decide on how to orient itself, assign power, and schedule users so that it can achieve the highest system performance, while being subject to QoS constraints.

The paper is structured as follows. The system model is described in Section 2, where a comprehensive overview of the UAV, RIS, channel model, and user selection is presented, detailing their interplay with the RIS-empowered UAV system. The DQN framework is presented in Section 3, and its use for the real-time optimization problem of UAVs in 3D space is discussed. Section 4 states the optimization problem while mentioning the goals of maximizing throughput and minimizing power and of providing each user its Quality of Service (QoS). The system constraints

are also comprehensively explained in this section. Section 5 describes the simulation setup to evaluate the proposed approach, and the results obtained, together with a comparison of the performance with baseline methods to show that it is effective. Finally, Section 6 presents the conclusion of this paper by reviewing the main results and proposing some future research directions to improve the system's performance and scalability.

2. LITERATURE REVIEW

The collaboration between Unmanned Aerial Vehicle (UAV) technology and Reconfigurable Intelligent Surfaces (RIS) has become one of the hot topics recently, as they could resolve communication problems, specifically in areas with no or limited infrastructure. Drones have been the poster-child vehicle for network coverage extension, especially in rural areas or where infrastructure is in short supply. Preliminary works have also considered optimization of the UAV trajectory and power allocation to improve system performance. Among them, [11] studied throughput maximization and power minimization for UAV-enabled systems. Analog-optical processing was also used to realize optimal system performance via conventional optimization methods, such as concave programming and gradient descent. Although such methods have performed well in practice, they are usually developed for set scenarios and are not adaptable to the real-time and time-variant environment caused by environmental changes that are prevalent in UAV networks [12].

As environments become increasingly dynamic, RL is considered an appropriate method to produce real-time decisions for UAV systems. In particular, the Deep Q-Networks (DQN) have been extensively adopted to solve UAV trajectory design. The DQN, a modified Q learning, uses DNNs to approximate Q values, so that UAVs can achieve the throughput maximization and carefully mitigate power consumption and QoS constraints at the same time [13][14]. DQN-based algorithms have proven themselves to be suitable candidates for optimizing 3D UAV trajectory planning to adapt to dynamic environments and make resource usage more efficient. The application of RIS in UAV networks increases the signal quality and coverage, so it is very popular. The RIS can actively control the reflected electromagnetic wave signals and is capable of enhancing the reliability of communication in NLoS conditions. Many works [15],[16] have proved the advantages of RIS in UAV communication and shown the ability of signal enhancement and interference reduction. Yet, most of these works are confined on the theoretical part of RIS [17] and do not provide practical implementation, including dynamic environments that need quick decision-making.

In the literature, a recent upsurging branch of research directions in UAV systems is towards the use of multi-agent reinforcement learning (MARL) models, in which multiple UAVs and RISs are jointly optimized for improving the performance of communication. The authors of [18] also introduced a multi-agent RL model to learn resource allocation in UAV-aided networks. Although this paper has shown the advantages of the multi-agent system, it only concentrated on the user scheduling and did not consider the RIS effect in designing the communication performance. The position of RIS in a multi-agent system is a relatively less-studied area in the literature. Although UAV-RISs are very effective in achieving the desired performance, there are several directions that are still not fully covered by the existing works [19]. A majority of existing work is theory-driven or simulation-based approaches that do not consider the dynamic nature of realistic settings. Furthermore, all existing related work on RIS does not apply reinforcement learning (RL) based DQN, which can support real-time decision and optimization. The sparse research on practical application underscores the necessity of more research on adaptive optimization for RIS-aided UAV systems.

2.1 Contributions

The contributions of this paper are:

- To introduce the use of Deep Q-Networks (DQN) for the first time to perform real-time UAV trajectory, power, and user control in RIS-aided UAV systems.
- To present a complete system model, including UAV mobility, RIS-assisted communication, and user scheduling, which are crucial for optimizing throughput and ensuring QoS compliance.
- To propose a reinforcement learning (RL) framework based on DQN, designed to learn optimal decisions adaptively in dynamic environments.
- To conduct performance evaluation through extensive simulations to assess the performance of the proposed approach, demonstrating its superiority in throughput maximization, power efficiency, and QoS compliance compared to conventional optimization techniques.

3. SYSTEM MODEL

3.1 UAV and User Model

The system consists of a UAV-mounted base station (UAV-BS) and K mobile users. The UAV-BS serves users within a 3D environment, where its position at time step n is denoted by:

$$[x_U(n), y_U(n), h_U(n)] \quad (1)$$

where:

- $x_U(n), y_U(n)$ represent the horizontal coordinates of the UAV at time step n,
- $h_U(n)$ Is the altitude of the UAV at time step n.

Each user k is represented by its position at time step n as:

$$[x_k(n), y_k(n), 0] \quad (2)$$

where x_k and $y_k(n)$ Represent the horizontal coordinates of user k at time step n, and the altitude is assumed to be 0 since the user is on the ground.

The Reconfigurable Intelligent Surface (RIS) is placed on a building with a fixed position. $[x_R, y_R, z_R]$ where x_R, y_R, z_R are the coordinates of the RIS in the horizontal plane, and its height above the ground?

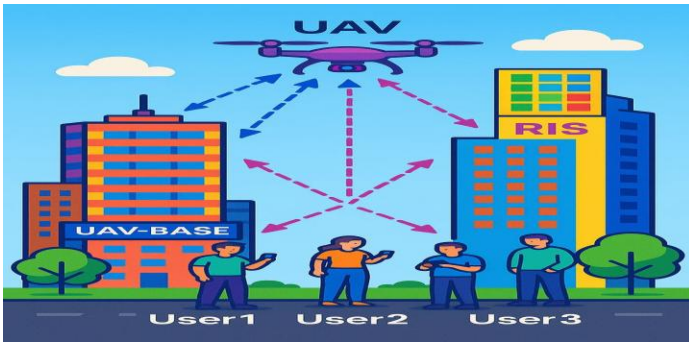


Figure 2. RIS-assisted UAV-BS system architecture

3.2 Channel Model

The communication between the UAV-BS and users is enhanced by the RIS. The channel gains between the UAV-BS, RIS, and users are modeled based on the path loss, distance, and angle of arrival at each time step.

- ***Channel Gain Between UAV-BS and RIS:***

The channel gain $g_{UR}(n)$ Between the UAV-BS and RIS at time step n is modeled as:

$$g_{UR}(n) = \frac{1}{d_{UR}(n)^2} e^{j\phi_{UR}(n)} \quad (3)$$

Where

- $d_{UR}(n)$ is the distance between the UAV and RIS at time step n,
- $\phi_{UR}(n)$ is the angle of arrival of the signal from the UAV to the RIS at time step n.

Channel Gain Between RIS and User k:

Similarly, the channel gain $g_{RK_k}(n)$ between the RIS and user k at time step n is given by:

$$g_{RK_k}(n) = \frac{1}{d_{RK_k}(n)^2} e^{j\phi_{RK_k}(n)} \quad (4)$$

3.3 Scheduling Model

The UAV-BS can serve only one user at each time step n. The scheduling decision for each user k at time step n is represented by a binary variable $c_{k,n}$, where:

$$c_{k,n} = \begin{cases} 1, & \text{if user } k \text{ is served at time step } n, \\ 0, & \text{otherwise.} \end{cases}$$

This implies that at each time step, the UAV-BS serves only one user:

$$\sum_{k=1}^{K} c_{k,n} = 1, \quad \forall n \in N \tag{5}$$

Where K is the number of all users and N is the set of times. This constraint guarantees that no more than one user is assigned for service in each time step, and the total number of served users in each time step is equal to 1.

3.4 Deep Q-Network (DQN) Algorithm

3.4.1 Overview

We adopt the deep Q-network (DQN) algorithm for trajectory, power allocation and user scheduling optimization in a RIS-aided UAV system. DQN is a method of estimating the Q-function by a deep neural network and chooses actions that maximize future expected rewards. Updating of the Q-value function is performed with the help of the Bellman equation:

$$Q(s_t, a_t) = r_t + \gamma \cdot \max_{a'} Q(s_{t+1}, a') \tag{6}$$

Where s_t is the state, a_t is the action, r_t is the immediate reward, γ is the discount factor, and a' represents future actions.

3.4.2 Experience Replay

To guarantee the stability of learning, DQN utilizes experience replay in which previous experiences are recorded in a replay buffer. Mini-batches are drawn randomly from the buffer for updating the Q-network, which decorrelates temporally and facilitates training stability.

3.4.3 Target Network

DQN uses a target network to alleviate overestimation bias. The target network is updated with the Q-network's weights every C training steps, which gives stable targets of Q-values and more constant updates in the course of training.

3.5 Reward Function

The reward function encourages the action that can maximize the throughput while minimizing the power, and meanwhile satisfy the QoS:

$$r_t = \text{Throughput}_t - \lambda_1 \cdot \text{Power Usage}_t - \lambda_2 \cdot \text{QoS Violation Penalty}_t \tag{7}$$

where λ1 and λ2 are the penalty factors of the power consumption and PS violation.

3.5.1 Training the DQN Agent

The training procedure of the DQN consists of several steps. The Q-network is first initialized with random weights, and the target network is then generated with these same weights. A replay buffer for the agent's forwarded experience is also built. The agent randomly samples actions and transitions: it explores the environment during training. As the agent gets trained, it begins to act in a greedy way by taking actions that lead to maximum rewards given the current policy. Experiences are collected in the replay buffer, and mini-batches are sampled randomly from the buffer to train the Q-network with stochastic gradient descent (SGD). Occasionally, the target network is updated with the weights from the Q-network to make more stable Q-value targets. Then, the Q-network is updated with the target Q-values that are obtained by the Bellman equation, and the loss function is minimized by computing the error of the predicted Q-value and target Q-value.

$$\text{Loss} = \frac{1}{2} \sum (Q_{\text{target}} - Q_{\text{predicted}})^2 \tag{8}$$

This iterative process enables the DQN agent to learn and optimize the UAV's decision-making over time, improving its performance in dynamic environments.

3.6 Optimization

The Deep Q-Network (DQN) technique is applied to jointly optimize the trajectory, power control, and user scheduling in the RIS-aided UAV system. To achieve action selection that maximizes throughput and minimizes power consumption with certain QoS for users in UAV-aided networks, DQN leverages a deep neural network to approximate the Q-values. The goal is to maximize the sum-rate under the constraints on UAV trajectory, power

allocation, user index scheduling, and QoS satisfaction. The UAV trajectory is limited by its feasible position in 3D space in each slot, and the power allocation of each user has to meet some constraints. The constraint here is that the UAV should fully serve exactly one user at each time step, and the throughput of each user should not fall below a minimum QoS level. Such constraints ensure that the UAV attains the best service based on the resources available and the nature of the service. The optimization problem is successively addressed by the DQN algorithm in a way to well learn a policy that regards the UAV trajectory, power allocation, and user scheduling over time in such a manner that maximizes the throughput, minimizes the power, and satisfies QoS requirements of all users.

4. RESULTS AND DISCUSSION

4.1 Simulation Settings

Simulation Parameters The simulation parameters adopted to evaluate the performance of the proposed DQN-based solution in optimizing throughput and QoS for the RIS-assisted UAV system is presented in this sub-section. The system is composed of K=10K = 10K=10 users that are uniformly distributed over a target area of size 200×200×120 meters. The initial location of the UAV-BS and the RIS is defined, and the maximum flying height of the UAV is limited to 60 meters. The considered scenarios are simulated with N=200N=200N=200 time slots, and the maximum transmit power of the UAV-BS is 100 watts.

The noise density σ is assumed to be -174 dBm/Hz, and both Line-of-Sight (LoS) and Non-Line-of-Sight (NLoS) conditions are considered. The system operates with a bandwidth B=1B = 1B=1 MHz. The UAV's maximum and minimum speeds are set to 21 m/step and 8 m/step, respectively, while the users' speed is capped at 0.5 m/step. For the s-DDQN network, both hidden layers consist of 64 neurons. The deep reinforcement learning (DRL) settings include a learning rate of 0.0001, a batch size of 64, and a discount factor γ=0.95\gamma = 0.95γ=0.95. The exploration factor decays with a constant of 0.0001, and the replay buffer size is set to 20,000. The number of training episodes is 2,000.The RIS is configured with M=25M = 25M=25 elements, and the Quality of Service (QoS) requirement for the RIS-assisted UAV system is set to 10,000.

4.2 Simulation Results

The performance of the reward curves of the optimal RIS phase shift in the system is depicted in Figure 3. At the beginning, the UAV-BS explores the environment through random actions, and a low value is assigned to the exploration reward as the agent always reaches the border of the target area with few satisfied users. As the exploration phase ends, the UAV-BS starts to take actions according to the learned policy to maximize the value of R. The reward function gradually converges to a certain value, which demonstrates that the proposed method outperforms other baseline methods such as s-DQN, rs-DDQN, fs-DDQN, and ss-DDQN.

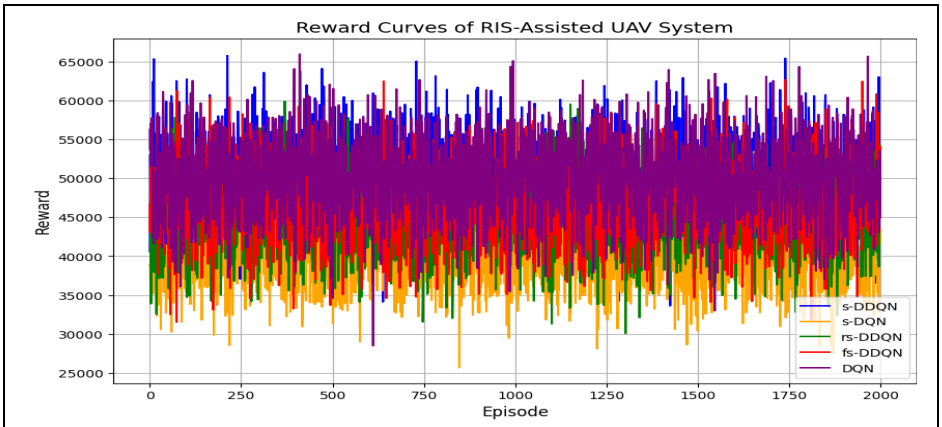


Figure 3. Reward curves for different algorithms in the RIS-assisted UAV system.

Figure 4 shows the 2D paths of the UAV with respect to the different algorithms. The new algorithm achieves convergence toward regions of interest faster, and accordingly, UAV can find the best configuration of user QoS while

minimizing the transmission power. From the path planning, the proposed UAV can achieve more efficient flight and provide higher QoS for users than other baselines.

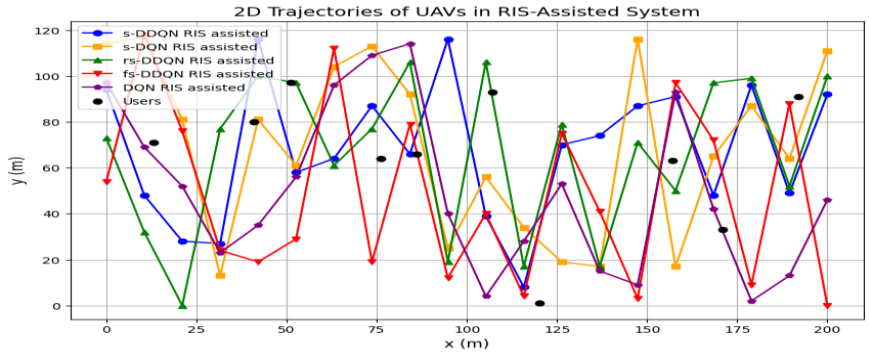


Figure 4. 2D trajectories of UAV-BS in the RIS-assisted system

Whereas 3D trajectories of the UAV in the RIS-aided UAV system are illustrated in Figure 5. Extra degrees of freedom added by the mobility of the user and the adjustable altitude of the UAV enable the UAV to maximize its flight trajectory by escalating its altitudinal level to alleviate the effect of user mobility on the power allocation, although the path loss may become larger. Result: This approach can not only maximize effective power consumption, but also provide superior QoS for the users.

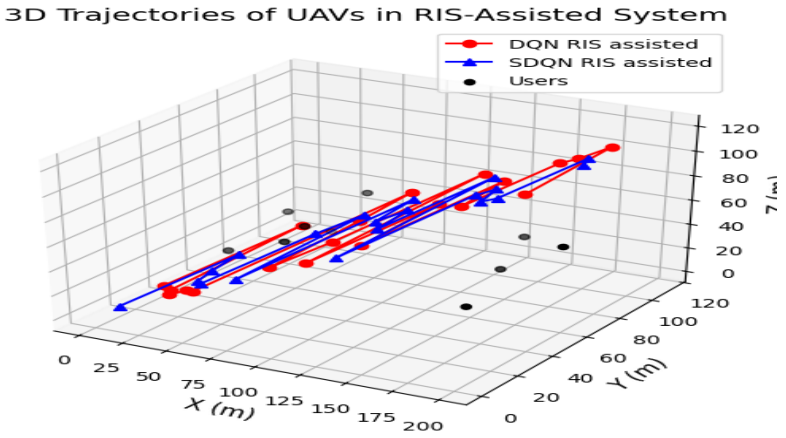


Figure 5. Proposed 3D trajectories of UAV-BS in the RIS-assisted system

These results confirm that the DQN-based optimization is indeed an effective approach for the RIS-empowered UAV systems, and has been observed to achieve a higher performance in comparison to throughput maximization, power minimization, and QoS satisfaction. The presented algorithm outperforms our baseline in both 2D and 3D trajectories and reward maximization.

5. CONCLUSION

This study reveals that the performance of RIS-aided UAV systems can be significantly enhanced by utilizing the Deep Q-Network (DQN) algorithm. The proposed DQN method is effective in achieving maximum throughput, minimum power consumption, and guaranteed minimum QoS performance to all the users. We demonstrate through extensive simulation results that the DQN method performs better than traditional S-DQN, RS-DDQN, FS-DDQN, and ss-DDQN methods in both two and three-dimensional trajectory optimization. Facilitated by the RL, the UAV is capable of making dynamic decisions, which are adaptable to complex and fast time-varying environments and contribute to more efficient resource utilization and better user service. This work demonstrates that DQN-driven

optimization will be an effective tool to improve the RIS-assisted UAV systems in practice and inspire further works in other communication systems.

REFERENCES

[1] Adil, M., Song, H., Jan, M. A., Khan, M. K., He, X., Farouk, A., & Jin, Z. (2024). UAV-assisted IoT applications, QoS requirements and challenges with future research directions. *ACM Computing Surveys*, *56*(10), 1-35.

[2] Alavizadeh, H., Alavizadeh, H., & Jang-Jaccard, J. (2022). Deep Q-learning based reinforcement learning approach for network intrusion detection. *Computers*, *11*(3), 41.

[3] Yuan, B., He, R., Ai, B., Chen, R., Wang, G., Ding, J., & Zhong, Z. (2022). A UAV-assisted search and localization strategy in non-line-of-sight scenarios. *IEEE Internet of Things Journal*, *9*(23), 23841-23851.

[4] Waqas, M., Abbas, Q., Qureshi, A., Amin, F., de la Torre Díez, I., Uc Rios, C., & Gongora, H. F. (2024). Advanced Line-of-Sight (LOS) model for communicating devices in modern indoor environment. *Plos one*, *19*(7), e0305039.

[5] Al-Hourani, A. (2024). Line-of-sight probability and holding distance in non-terrestrial networks. *IEEE Communications Letters*, *28*(3), 622-626.

[6] Elechi, P., & Onu, K. E. (2022). Unmanned Aerial Vehicle Cellular Communication Operating in Non-terrestrial Networks. In *Unmanned Aerial Vehicle Cellular Communications* (pp. 225-251). Cham: Springer International Publishing.

[7] Gupta, A., Dhawan, D., & Gupta, N. (2024). Review on UAV-based FSO links: recent advances, challenges, and performance metrics. *Optical Engineering*, *63*(4), 041204-041204.

[8] Wu, B., Qi, Y., Qiu, C., & Tang, Y. (2021). Wideband anti-jamming based on free space optical communication and photonic signal processing. *Sensors*, *21*(4), 1136.

[9] Zhang, X., Liu, W., Ren, J., Xing, H., Gui, G., Shen, Y., & Cui, S. (2025). Latency minimization for UAV-enabled federated learning: Trajectory design and resource allocation. *IEEE Internet of Things Journal*.

[10] Zhou, Y., Fan, P., & Ma, Z. (2025). A UAV-enabled integrated sensing and covert communication system. *IEEE Transactions on Vehicular Technology*.

[11] Hou, X., Wang, J., Zhang, Z., Wang, J., Liu, L., & Ren, Y. (2025). Split federated learning for UAV-enabled integrated sensing, computation, and communication. *arXiv preprint arXiv:2504.01443*.

[12] Ahmed, M., Soofi, A. A., Raza, S., Li, Y., Khan, F., Khan, W. U.,& Han, Z. (2025). A Comprehensive Survey on RIS-Enhanced Physical Layer Security in UAV-Assisted Networks. *IEEE Internet of Things Journal*.

[13] Ahmed, M., Soofi, A. A., Raza, S., Khan, F., Ahmad, S., Ullah, W.,& Han, Z. (2025). Advancements in RIS-Assisted UAV for Empowering Multi-Access Edge Computing: A Survey. *IEEE Internet of Things Journal*.

[14] Haider, Z. A., Khan, F. M., Zafar, A., & Khan, I. U. (2024). Optimizing Machine Learning Classifiers for Credit Card Fraud Detection on Highly Imbalanced Datasets Using PCA and SMOTE Techniques. *VAWKUM Transactions on Computer Sciences*, *12*(2), 28-49.

[15] Khan, I. U., Khan, Z. A., Ahmad, M., Khan, A. H., Muahmmad, F., Imran, A., ...& Hamid, M. K. (2023, May). Machine Learning Techniques for Permission-based Malware Detection in Android Applications. In *2023 9th International Conference on Information Technology Trends (ITT)* (pp. 7-13). IEEE.

[16] Khan, I. U., Zeb, A., Rahman, T., Khan, F. M., Haider, Z. A., & Bilal, H. (2024). ViTDroid and Hybrid Models for Effective Android and IoT Malware Detection. *ICCK Transactions on Advanced Computing and Systems*, *1*(1), 32-47.

[17] Khan, I. U., Khan, F. M., Haider, Z. A., Khattak, S., Naheed, G., & Kiani, S. S. Dynamic Malware Detection Using Effective Machine Learning Models with Feature Selection Techniques.

[18] Fang, Z., & Savkin, A. V. (2024). Strategies for optimized uav surveillance in various tasks and scenarios: A review. *Drones*, *8*(5), 193.

[19] Khan, I. U., Khan, F. M., Haider, Z. A., & Alturise, F. (2025). Integrating AI, Blockchain, and Edge Computing for Zero-Trust IoT Security: A Comprehensive Review of Advanced Cybersecurity Framework. *Computers, Materials and Continua*, *85*(3), 4307-4344.