ARTICLE

Artificial Intelligence (AI)-Enabled Unmanned Aerial Vehicle (UAV) Systems for Optimizing User Connectivity in Sixth-Generation (6G) Ubiquitous Networks

Zeeshan Ali Haider[1]**, Inam Ullah**[2,*]**, Ahmad Abu Shareha**[3]**, Rashid Nasimov**[4] **and Sufyan Ali Memon**[5,*]

[1]Department of Computer Science, Qurtuba University of Science & IT, Peshawar, 25000, Pakistan

[2]Department of Computer Engineering, Gachon University, Seongnam, 13120, Republic of Korea

[3]Department of Data Science and Artificial Intelligence, Hourani Center for Applied Scientific Research, Al-Ahliyya Amman University, Amman, 19328, Jordan

[4]Department of Artificial Intelligence, Tashkent State University of Economics, Tashkent, 100066, Uzbekistan

[5]Department of Defense Systems Engineering, Sejong University, Gwangjin-gu, Seoul, 05006, Republic of Korea

*Corresponding Authors: Inam Ullah. Email: inam@gachon.ac.kr; Sufyan Ali Memon. Email: sufyanahmedali@sejong.ac.kr

ABSTRACT: The advent of sixth-generation (6G) networks introduces unprecedented challenges in achieving seamless connectivity, ultra-low latency, and efficient resource management in highly dynamic environments. Although fifth-generation (5G) networks transformed mobile broadband and machine-type communications at massive scales, their properties of scaling, interference management, and latency remain a limitation in dense high mobility settings. To overcome these limitations, artificial intelligence (AI) and unmanned aerial vehicles (UAVs) have emerged as potential solutions to develop versatile, dynamic, and energy-efficient communication systems. The study proposes an AI-based UAV architecture that utilizes cooperative reinforcement learning (CoRL) to manage an autonomous network. The UAVs collaborate by sharing local observations and real-time state exchanges to optimize user connectivity, movement directions, allocate power, and resource distribution. Unlike conventional centralized or autonomous methods, CoRL involves joint state sharing and conflict-sensitive reward shaping, which ensures fair coverage, less interference, and enhanced adaptability in a dynamic urban environment. Simulations conducted in smart city scenarios with 10 UAVs and 50 ground users demonstrate that the proposed CoRL-based UAV system increases user coverage by up to 10%, achieves convergence 40% faster, and reduces latency and energy consumption by 30% compared with centralized and decentralized baselines. Furthermore, the distributed nature of the algorithm ensures scalability and flexibility, making it well-suited for future large-scale 6G deployments. The results highlighted that AI-enabled UAV systems enhance connectivity, support ultra-reliable low-latency communications (URLLC), and improve 6G network efficiency. Future work will extend the framework with adaptive modulation, beamforming-aware positioning, and real-world testbed deployment.

KEYWORDS: 6G networks; UAV-based communication; cooperative reinforcement learning; network optimization; user connectivity; energy efficiency

1 Introduction

With the increasing demand from consumers for seamless connectivity, high data rates, and stable services, sixth-generation (6G) networks have emerged as a response to these needs [1]. Unlike the

connectivity of previous generations, 6G will achieve seamless connectivity between ground-to-ground, air-to-ground, and space-based users, resulting in a fully interconnected environment [2,3]. It will accommodate massive machine-type communications (mMTC), ultra-reliable low-latency communications (URLLC), and high-speed broadband, whose deployment requires high-efficiency spectrum and energy utilization [4,5]. However, dynamic user mobility, interference, and dense resource allocation challenges still restrict large-scale implementation [6]. While 5G has revolutionized mobile broadband and mMTC, its spectrum efficiency, latency, interference, and scalability limitations in dense and high-mobility environments pose significant challenges to the technology's complete potential [7,8]. Advanced solutions are required to meet the growing demand for URLLC and massive connectivity in 6G applications, including autonomous driving, smart cities, immersive services, and others [9]. These difficulties fuel the adoption of artificial intelligence, UAVs, and decentralized decision-making systems to maximize connectivity and resource utilization [10].

Unmanned aerial vehicles (UAVs) offer a promising solution as mobile base stations to extend network coverage in underserved or densely populated areas, thereby increasing network capacity [11,12]. Dynamic service provisioning based on user needs and network conditions can be realized through their flexible location in terms of distributed intelligence, but without centralized control to handle accessibility and resources [13] and [14]. However, centralized UAV control in dense 5G networks has led to user drop rates of over 18%–22% [15] as well as latency bursts of over 50–70 ms in high-mobility environments, which does not meet the URLLC requirement of below 1 ms for 6G [16]. Additionally, UAV-assisted networks are energy-inefficient, with power consumption increasing by 25%–35% compared to reinforcement learning-based methods [17]. These constraints, in turn, underscore the critical need for distributed and cooperative intelligence in UAV-assisted 6G, which the proposed CoRL-based framework aims to address.

The paper explores optimizing user connectivity in 6G networks using AI-based UAVs with cooperative reinforcement learning (CoRL). UAVs autonomously adjust position, power, and resources based on local data and shared information. Simulations show that CoRL outperforms traditional methods in user coverage, convergence, latency, and energy efficiency. The contributions of this paper are as follows:

- This study proposes a novel framework for AI-enabled UAV systems in 6G networks, leveraging CoRL to optimize user connectivity.
- The work introduces collaborative information exchange techniques that allow UAVs to learn and adapt based on both local and shared information, improving scalability and efficiency.
- The study demonstrates the effectiveness of this framework through simulation results, showing superior performance compared to traditional approaches in dynamic environments.

This paper is structured as follows: Section 2 presents the related work in the field, highlighting the current research and gaps in UAV-based systems, particularly within the context of 6G networks. Section 3 explains the system model, the network architecture of the UAV-based network, and the significant assumptions underlying it. Section 4 provides a discussion of formulating the problem, specifying the objective function, and constraints. The CoRL framework and the algorithm used to make UAV decisions are discussed in section 5. Section 6 contains the simulation performance and comparison results. Lastly, section 7 concludes the study and proposes future work.

2 Related Work

Recent studies have demonstrated the UAV's potential to enhance wireless connectivity by supporting 5G and 6G networks. For instance, reference [18] showed the ability of UAVs to strengthen throughput and user connectivity in 5G through deep reinforcement learning (DRL), while reference [19] presented heuristics for bandwidth-efficient resource allocation in large UAV networks. Energy efficiency has also been

highlighted, including AI-based power control optimization in [20] and machine learning-based network performance optimization and energy consumption in [21]. For issues of interference and bandwidth allocation, the UAV altitude was optimized in [22] to minimize interference with ground users. A power and bandwidth management scheme with bounded optimal performance under challenging conditions was presented in [23]. Security has also been a primary concern; reference [24] proposed secure communication protocols to maintain data integrity in UAV networks. In contrast, reference [25] suggested encryption-based schemes to prevent data leakage in multi-UAV systems.

Discussions related to the use of UAVs within 6G networks have gained significant attention. In addition, due to the increasing significance of security and privacy in UAV-assisted communication environments, a privacy-preserving access control protocol for intelligent UAV networks supported by 6G was recently proposed by [26]. A recent study by [27] on AI-driven seamless and massive access in space-air-ground integrated networks discussed AI-driven seamless and massive access in space-air-ground integrated networks, highlighting the way in which artificial intelligence can facilitate large-scale heterogeneous connectivity in a future 6G world. Additionally, reference [28] highlighted UAV integration in 6G networks, enabling seamless communication and data delivery in smart urban scenarios. To illustrate, reference [29] has designed an AI-enabled traffic routing system for UAVs, where repositioning is performed optimally according to the traffic conditions. This study demonstrated the capability of AI to optimize urban transport systems and alleviate traffic congestion. Similar to [30], which covered innovative city technologies in real-time urban traffic control, UAV applications are likely to attract a greater level of interest.

Although generative artificial intelligence (GAI) has numerous applications in autonomous systems [31], its role in UAV-assisted 6G networks remains underexplored. Utilized generative adversarial networks (GANs) to model sophisticated UAV communication environments and train synthetic data, while reference [32] employed variational autoencoders (VAEs) to predict UAV networks based on data extrapolation. Generative diffusion models (GDMs) have been studied for real-time UAV placement and routing optimization [33]. Additionally, transformer models have been applied to predictive situational awareness, enabling adaptive redeployment and predicting user demand in dynamic environments [34]. Although some of the research above has discussed UAV energy efficiency, interference mitigation, and security, the application of AI-driven optimization to UAV networks remains largely unexplored. Thus, the target of this study is to incorporate AI models into UAV systems, which are necessary for 6G, to enhance decision-making, positioning, and user connection.

3 System Model

This section explains the system architecture of AI-supported UAV-based communication networks with 6G ubiquitous connectivity, as illustrated in Fig. 1. The system is the set of UAVs and intelligent base stations (IBS) introduced in the dynamic 6G environment. These UAVs are highly flexible and offer services on demand due to their sophisticated wireless communication solutions. They improve connectivity by optimizing resource utilization, enhancing user mobility, and adjusting network conditions.

3.1 Network Architecture

The UAV system architecture includes UAVs, $\mathcal{U} = \{U_1, U_2, \ldots, U_I\}$ with I the total number of UAVs, and ground users, $\mathcal{U}_g = \{u_1, u_2, \ldots, u_U\}$ with U the total number of ground users. The UAVs are located in a 3D scenario that encompasses a large area (urban or rural), providing coverage and resource distribution to both static and dynamic users. Each UAV can modify its altitude and location accordingly to maximize connectivity for users and network requirements. UAVs equipped with directional antennas effectively serve users by providing power and minimizing interference. The connection with users is established through

orthogonal frequency division multiple access (OFDMA) technology, where resource blocks (RBs) are assigned according to the users' data rate requirements. The communication connection can be damaged and distorted, and has a limited transmission capacity that needs to be addressed.
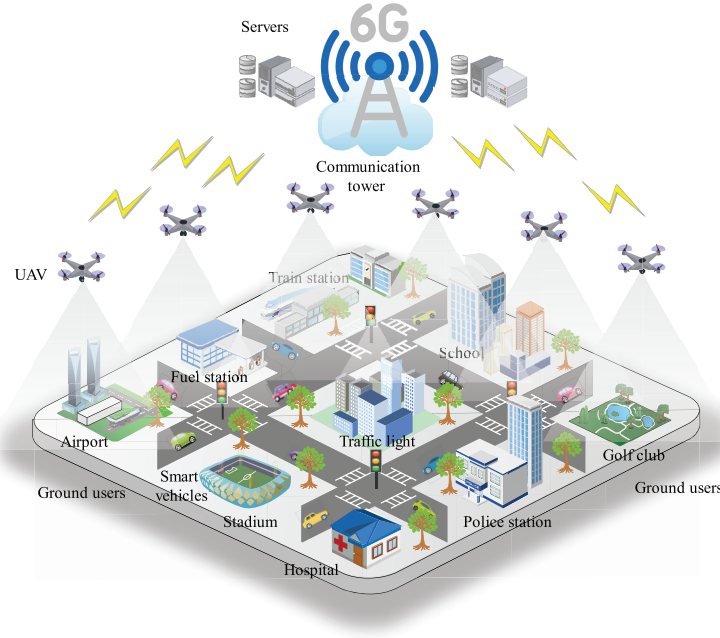

Figure 1: System model of AI-enabled UAVs optimizing connectivity in 6G networks

3.2 UAV Movement and Positioning

Deployed UAVs navigate in three-dimensional (3D) space, where the location of UAV i at time t is described by $(x_{i,t}, y_{i,t}, z_{i,t})$, with $x_{i,t}, y_{i,t}$ where are horizontal coordinates and $z_{i,t}$ is altitude. They are autonomous, using active positioning, and take movement decisions based on local observations and cooperative information with neighbours. The altitude for the UAV is the result of the environmental and user density: in urban areas, a UAV can operate at 100–150 m to avoid blockage and reflections; in rural areas, 200–300 m to ensure maximal coverage; in dense networks, UAVs spread horizontally to ensure no overlap, and in sparse networks UAVs cluster together to ensure connectivity. The CoRL framework informs the strategies and is demand-driven, congestion-aware, and resource-aware. The action space of the UAVs $A_i(t)$ consists of forward, backward, left, right, and hovering actions, and each action $a_i(t)$ is chosen to optimize network performance.

3.3 Resource Allocation and User Connectivity

In the proposed system, resource allocation is concerned with the assignment of discrete resource blocks (RBs) and transmit power that will fulfill users' data rate requirements. In this paper, the available bandwidth is divided into RBs, and the UAVs allocate resources of some RBs to each user according to the signal-to-interference-plus-noise ratio (SINR) and throughput requirement, while minimizing interference and energy consumption. The UAV assigns part of its RBs to users within its coverage area, which is dynamically estimated based on the UAV's position and the density of users in that area. The throughput $r_{j,i}(t)$ of user u_j served by the UAV i at a time t is obtained as follows, as per the Shannon capacity equation of the UAV communication channel:

$$r_{j,i}(t) = N_{i,j}(t) \cdot BW \cdot \log_2\left(1 + \mathrm{SINR}_{j,i}(t)\right) \tag{1}$$

where $N_{i,j}(t)$ is the number of resource blocks (RBs) allocated to user u_j by UAV i, BW is the bandwidth of each RB, $\mathrm{SINR}_{j,i}(t)$ is the SINR between UAV i and user u_j, which is given by:

$$\mathrm{SINR}_{j,i}(t) = \frac{P_i(t)G_{j,i}(t)}{\sigma + \sum_{k \neq i} P_k(t)G_{j,k}(t)} \tag{2}$$

Here $P_i(t)$ is the transmit power of the UAV i, $G_{j,i}(t)$ is the channel gain between UAV i and user u_j, σ is the noise power, The summation term represents the interference from other UAVs that may affect user u_j. A user u_j is considered connected to UAV i if its data rate $r_{j,i}(t)$ exceeds its minimum throughput requirement $r_{\min}$. In this case, the user connectivity is represented as a binary variable $x_{i,j}(t)$, where:

$$x_{i,j}(t) = \begin{cases} 1 & \text{if user } u_j \text{ is connected to UAV } i, \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

The goal is to maximize the number of connected users in all UAVs in the network while minimizing latency, interference, and energy consumption.

3.4 Collaborative Information Exchange and Cooperative Learning

UAVs are modeled as intelligent agents that optimize the network's performance through distributed decision-making, based on local observations (e.g., position, connected users, available bandwidth), and by exchanging state information among neighbors. In the proposed CoRL framework, UAVs exchange data within a 1 km proximity of a collaboration radius, such as (i) 3D position, (ii) user load, (iii) available bandwidth, and (iv) interference levels. This enables the coordination of user handoffs, eliminates redundant position checking, and facilitates dynamic resource balancing without the need for a global controller. In such cooperation, UAVs can notify of congestion, report their status, and trigger neighbors to change stance to reduce the load. A reward function is defined to minimize interference, power consumption, and latency, while maximizing benefits for users. CoRL therefore trains UAVs to learn optimal strategies of resource allocation, user association, and navigation; the reward function $R_i(t)$ for UAV i at time t is defined as:

$$R_i(t) = \sum_{u_j \in \mathcal{U}_i} x_{i,j}(t) - \lambda \cdot \text{interference}(t) - \mu \cdot \text{energy}(t) \tag{4}$$

where $\mathcal{U}_i$ is the set of users connected to UAV i, λ and μ are weighting factors that penalize interference and energy consumption, respectively. The UAVs use these rewards to update their policies through Q-learning or Deep Q-learning algorithms, enabling them to improve connectivity over time and adapt to changing user and environmental conditions.

Weighting rationale and anti-starvation guard. The coefficients λ and μ control the interference and energy emphasis, respectively, while maintaining connectivity as the dominant signal. To prevent pathological policies (e.g., conserving energy by not serving users), we add two guard terms to the reward: a penalty for unmet demand and a small fairness bonus. The resulting molded reward is:

$$\tilde{R}_i(t) = \sum_{u_j \in \mathcal{U}_i} x_{i,j}(t) - \lambda \cdot interference(t) - \mu \cdot energy(t) - \kappa \cdot unmet(t) + \beta \cdot fair(t), \tag{4a}$$

where $unmet(t) = \sum_j \mathbf{1}\{r_{j,i}(t) < r_{\min}\}$ penalizes users falling below $r_{\min}$, and $fair(t) = \frac{(\sum_j x_{i,j})^2}{N \sum_j x_{i,j}^2}$ (Jain's index across locally served users) encourages balanced service. We use a small β to avoid overshadowing the primary objective. This shaping empirically eliminated starvation behaviors while maintaining energy awareness.

3.5 UAV Position Constraints

The positions of UAVs are constrained by the area in which they operate. Let the operational area be a 3D space, where the UAVs are limited to a square region on the ground with a side length of L and an altitude of H. The position of each UAV must satisfy the following constraints:

$$0 \le x_{i,t}, y_{i,t} \le L, \quad \forall i \in \mathcal{U}, \quad 0 \le z_{i,t} \le H, \quad \forall i \in \mathcal{U} \tag{5}$$

4 Problem Formulation

In this section, the optimization problem of AI-enhanced UAVs in 6G networks is formalized and seeks to maximize the connectivity of users, minimize interference, energy cost, and latency. UAVs must make decisions about where and whom to serve, as well as which resources to allocate. These choices should be based on location, user identification, and coverage in both urban and rural areas. SINR and reward shaping are implemented in the CoRL framework, which is flexible to changing environments and long-term network performance.

4.1 Objective Function

The goal is to maximize the total number of connected users in the network while minimizing interference and energy consumption. The objective function $J(t)$ for the entire network at time t can be expressed as:

$$J(t) = \sum_{i=1}^{I} \sum_{j=1}^{U} x_{i,j}(t) - \lambda \cdot \sum_{i=1}^{I} \sum_{j \ne i} \text{interference}_{i,j}(t) - \mu \cdot \sum_{i=1}^{I} \text{energy}_i(t) \tag{6}$$

where $x_{i,j}(t)$ is a binary variable indicating whether the user u_j is connected to the UAV i at time t (1 if connected, 0 otherwise), λ is a weight factor for interference between UAVs, $\text{interference}_{i,j}(t)$ represents the interference caused by UAV i on the user u_j due to proximity or shared bandwidth, μ is a weight factor for energy consumption, and $\text{energy}_i(t)$ represents the energy consumption of UAV i at time t.

4.2 Decision Variables

This problem has decision variables that include UAV positioning, UAV user association, resource allocation, and UAV power control. The location of the UAVs at a particular time t, $(x_{i,t}, y_{i,t}, z_{i,t})$, where $i \in \mathcal{U}$ also determines the coverage of the respective UAVs as well as the connectivity with the users. Whether each user is assigned to the UAV is described by $x_{i,j}(t)$ whereby; $x_{i,j}(t) = 1$ each time user u_j is evenly associated with UAV i at a given time t, otherwise it will be 0. Such an association between this user influences the quality of service (QoS) delivered to the user as well as defines the RB assignment. The data rate to each user u_j is influenced by the number of RBs $N_{i,j}(t)$ assigned by UAV i at time t, together with transmit power, as part of the time slot t by the UAV i, that the user gets. Finally, the transmit power $P_i(t)$ of each UAV at time t will affect the quality of the signal and interference, directly affecting users and those near the UAVs.

4.3 Constraints

The problem is subject to certain essential constraints. First, user connectivity requirements are the constraint that each user can be connected to at most one UAV at any given time (i.e., throughout their lifetime) which is written as $\sum_{i=1}^{I} x_{i,j}(t) \leq 1$, $\forall j \in \mathcal{U}_g$. Second, the RB assignment constraint is that each UAV should not allocate more than its available pool of RBs to the UAVs, i.e., $\sum_{j=1}^{U} N_{i,j}(t) \leq N_{\max}$, $\forall i \in \mathcal{U}$. Third, UAV motion is constrained to the area of deployment, which in the case of a square area of side L and altitude H, the condition $0 \leq x_{i,t}, y_{i,t} \leq L$, $0 \leq z_{i,t} \leq H$, $\forall i \in \mathcal{U}$ must be satisfied. Fourth, power and interference constraints are such that UAV transmission power is bounded above by $P_{\max}$, and cumulative interference is bounded as follows: $\sum_{k \neq i} \text{interference}_{i,k}(t) \leq \varepsilon$. Finally, for energy consumption, it is considered that, at any time t, $\text{energy}_i(t) \leq E_{\max}$, $\forall i \in \mathcal{U}$ meaning that UAVs operation is constrained to stay within tolerable energy limits.

4.4 Reinforcement Learning-Based Decision Making

The UAVs in this network are intelligent agents that make decisions under CoRL. The cooperation between agents involves sharing information about the state, which can take the form of positions, user associations, and rewards, depending on factors such as connectivity, interference, and energy consumption. To find a policy that maximizes user connectivity while complying with the aforementioned constraints, each UAV learns a policy. The learning process can be defined as a Markov Decision Process (MDP), where the state space comprises the location of each UAV, the group of connected users, and the available bandwidth. The action space involves the movement decision space, the user relationship decision space, and the resource allotment decision space.

5 Cooperative Reinforcement Learning (CoRL) Algorithm

In this section, the CoRL algorithm is described, in which the UAVs on the 6G network optimize user connections, resources, and mobility in a distributed way. UAVs are autonomous intelligent agents that work collectively to optimize the overall network performance. It is by feeding on the environment that they learn the best policies of decision-making, which alter their positions, user associations, and resource allocation. Fig. 2 illustrates the stepwise methodology followed in the proposed study. The process begins with designing the UAV-based 6G network and initializing the UAV and user nodes. It then models resource allocation, formulates the optimization problem, and sets up the CoRL framework. UAVs collaboratively learn optimal actions through Q-learning, information sharing, and cycles of exploration and exploitation. Finally, the model is evaluated through simulations, producing optimized UAV policies for enhanced user connectivity in dynamic 6G environments.

5.1 Overview of Cooperative Reinforcement Learning (CoRL)

In traditional RL, an agent learns a policy to maximize cumulative rewards by interacting with the world. In multi-agent systems, such as UAV networks, agents must not only optimize their own actions but also coordinate with their neighbors to maximize their performance. CoRL achieves this by enabling UAVs to exchange information and coordinate their policies to optimize network connectivity and efficiency through mutual actions. Unlike generic multi-agent reinforcement learning (MARL), in which UAVs selfishly optimize their own policies, which often leads to coverage overlaps or resource wastage, CoRL adds two mechanisms to prevent these issues: (i) Collaborative State Sharing (local state exchange to prevent coverage overlaps), and (ii) Conflict-Aware Reward Shaping (rewards to prevent redundant associations or excessive interference). Together, these mechanisms ensure that the UAVs maximize the utility of their local operations while supporting the global objective of achieving well-balanced user coverage, making

CoRL a methodically designed cooperation paradigm for 6G UAV networks, rather than a trivial adaptation of MARL.

The CoRL framework is modeled as a multi-agent Markov Decision Process (MDP), where:

- State Space: The state $s_i(t)$ of UAV i at time t includes position $(x_{i,t}, y_{i,t}, z_{i,t})$, connected users, available bandwidth, and local environmental data.
- Action Space: The action $a_i(t)$ for each UAV includes movement (e.g., forward, backward, left, right), resource allocation, and power control.
- Reward Function: The reward $r_i(t)$ is based on the number of connected users, resource efficiency, interference, and energy consumption.
- Collaboration: UAVs exchange detailed state information (position, user load, RB availability, and interference levels) and reward values with neighboring UAVs within a 1 km range to improve convergence and avoid overlapping actions.
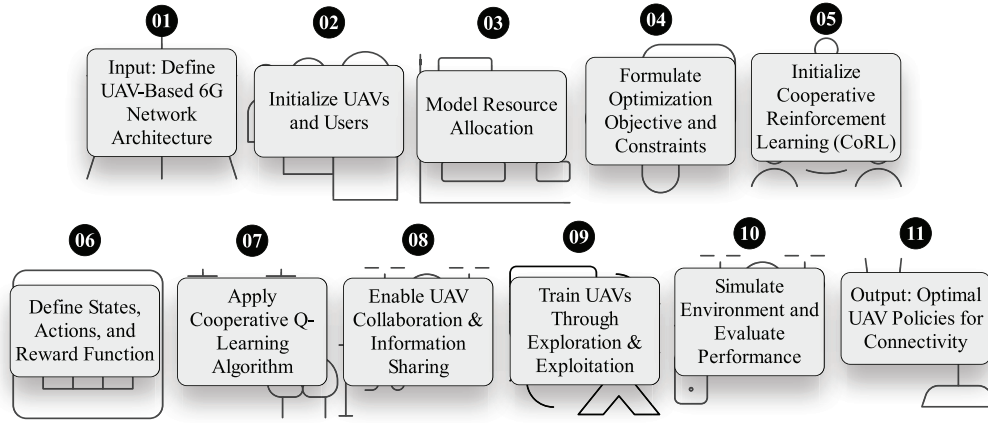

Figure 2: Methodology of the proposed CoRL-based AI-enabled UAV system for 6G user connectivity optimization

5.2 Multi-Agent Learning Process

Each UAV is a reinforcement learning agent that runs the Q-learning algorithm, in which state-action pairs are indexed with Q-values in a Q-table (or Q-network in deep Q-learning). The learning process balances exploration, in which UAVs randomly change positions and allocation of resources to acquire experience, and exploitation, in which the UAVs use learned policies to maximize rewards. Cooperation is attained through the exchange of state, action, and reward information with adjacent UAVs, which allows policy synchronism to enhance connectivity and to minimize network inefficiency.

5.3 Cooperative Q-Learning Algorithm

The cooperative Q-learning algorithm is the central approach used in the CoRL framework. Each UAV learns a Q-function $Q_i(s_i, a_i)$ that estimates the expected cumulative reward of taking action a_i in state s_i and following the optimal policy thereafter. The Q-values are updated iteratively using the Bellman equation:

$$Q_i(s_i, a_i) \leftarrow Q_i(s_i, a_i) + \alpha \left(\tilde{R}_i(t) + \gamma \cdot \max_{a_i'} Q_i(s_i', a_i') - Q_i(s_i, a_i) \right) \qquad (7)$$

where $Q_i(s_i, a_i)$ is the Q-value of UAV i for state s_i and action a_i, α is the learning rate that controls the rate at which Q-values are updated, $\tilde{R}_i(t)$ is the reward received by UAV i at time t, and γ is the discount factor,

representing the importance of future rewards. Also, $\max_{a'_i} Q_i(s'_i, a'_i)$ is the greatest future expected reward to be attained in the next state, s'_i.

Algorithm 1 outlines the key steps of the Cooperative Reinforcement Learning (CoRL) framework for UAV decision-making in 6G-enabled networks. Each UAV learns its optimal policy for user connectivity, resource allocation, and movement through cooperative interactions with neighboring UAVs.

Algorithm 1: Cooperative Reinforcement Learning (CoRL) for UAV decision-making

---

Input: Environment parameters, learning rate α, discount factor γ, exploration rate ε, and reward function.

Output: Optimal policies for UAV decision-making and network performance improvement.

Data: Initialize Q-values $Q_i(s_i, a_i)$ for each UAV i for all state-action pairs (s_i, a_i).

Data: Set exploration rate ε for balancing exploration and exploitation.

Data: Initialize replay buffer to store experiences $(s_i, a_i, r_i(t), s'_i)$.

1: **for** *each episode* **do**
2:     Reset environment and initialize UAV states;
3:     **for** *each time step t* **do**
4:         **for** *each UAV i* **do**
5:             Select action $a_i(t)$ using ε-greedy policy:;
6:             **if** *with probability ε* **then**
7:                 Take random action;
8:             **else**
9:                 Select $a_i(t) = \arg\max_{a_i} Q_i(s_i(t), a_i)$;
10:            Execute $a_i(t)$ and observe new state $s'_i(t)$ and reward $r_i(t)$;
11:            Share local state information (x_i, y_i, z_i), user load, available RBs, and interference levels with UAVs located within a 1 km radius, along with reward values;
12:            Compute shaped reward $\tilde{R}_i(t)$ using $\lambda, \mu, \kappa, \beta$ (Eq. 4a), including unmet-demand and fairness terms;
13:            Update Q-values using Bellman rule:;
14:            $Q_i(s_i, a_i) \leftarrow Q_i(s_i, a_i) + \alpha \left(\tilde{R}_i(t) + \gamma \cdot \max_{a'_i} Q_i(s'_i, a'_i) - Q_i(s_i, a_i) \right)$
15:            Store $(s_i(t), a_i(t), r_i(t), s'_i(t))$ in replay buffer;
16:        Periodically update target Q-networks for stability;

---

6 Results

In this section, the efficiency of the AI-enabled UAV system is evaluated to optimize user connectivity in 6G networks through extensive simulations. The relative success of the CoRL algorithm is compared to that of popular centralized and decentralized algorithms, and the algorithms are parameterized in dynamic environments. The findings indicate the benefits of UAV collaboration and distributed learning in massive dynamic networks.

6.1 Simulation Setup

The simulation models a smart city with UAVs providing 6G connectivity to mobile users. A total of $I = 10$ UAVs are used by $U = 50$ ground users randomly spread out in a space of size $L \times L$ where $L = 1000$ m. Each UAV is equipped with directional antennas and mobility to maximize coverage under a maximum transmit power of $P_{\max} = 20$ dBm and a maximum bandwidth of $N_{\max} = 100$ RBs. The user needs at least a throughput of $r_{\min} = 250$ kbps. UAV can go left, right, up, down, or hover. Learning parameters

are set to $\alpha = 0.01$ (learning rate), $\gamma = 0.95$ (discount factor), and $\varepsilon = 0.1$ (exploration rate decaying over episodes). The reward function uses connectivity, interference, and energy sensitivity. The simulation model was developed using Python and standard libraries, along with its own modules. Computation and data were implemented in NumPy and Pandas, visualization was implemented in Matplotlib, and reinforcement learning with multi-agent adaptations to CoRL was implemented in TensorFlow. The 6G network supporting the UAV mobility model, RB allocation model, and interference model was modeled based on the OFDMA with custom modules. It was a reproducible and flexible design. Reward weights $(\lambda, \mu, \kappa, 2\beta)$ were calibrated by a two-stage grid search over 100 validation episodes, with coarsely and then finely tuned grid$\{\lambda, \mu \in \{0.1, 0.3, 0.5, 0.7\}, \kappa \in \{0.1, 0.25, 0.5\}, \beta \in \{0, 0.05, 0.1\}\}$. The optimum environment $(0.40, 0.30, 0.25, 0.05)$ maximized connectivity while maintaining the other KPIs within 2% of their single-metric optima. Performance is compared to six baselines: centralized Q-learning (C-QA), decentralized Q-learning (D-QA), Greedy (choose reward based on immediate reward), decentralized scheduling with QoS (DS-QoS) using MARL and Graph Neural Network (GNNs), multi-agent coordinated exploration (MACE) with novelty sharing, and decentralized Monte Carlo tree search (MCTS) for partial observability. Table 1 contains the summary of the main simulation parameters.

Table 1: Simulation parameters

Parameter	Value
Number of UAVs (I)	10
Number of users (U)	50
Area size ($L \times L$)	1000 m × 1000 m
Maximum UAV power ($P_{\max}$)	20 dBm
Maximum bandwidth per UAV ($N_{\max}$)	100 RBs
UAV altitude (H)	150 m
Minimum user throughput ($r_{\min}$)	250 kbps
Learning rate (α)	0.01
Discount factor (γ)	0.95
Exploration rate (ε)	0.1
Interference weight (λ)	0.40
Energy weight (μ)	0.30
Unmet-demand penalty (κ)	0.25
Fairness bonus (β)	0.05
Maximum episodes	500
Number of time steps per episode	100

6.2 Performance Metrics

In the evaluation, the following performance metrics were taken into consideration:

- Total connected users: number of successful connections to users in the course of the simulation time.
- Network convergence speed: number of episodes before the UAVs achieve a converged policy where the throughput performance reaches a steady state.
- Energy efficiency: the overall amount of energy used by UAVs, both during transmission and the energy required to keep the flight operations going.
- Latency: mean end-to-end delay to link users, the sum of the user association stage and the resource block (RB) allocation stage delays.

- Interference: mean number of cells in the network that cause interference with each other.

6.3 Detail Performance Evaluation

The simulation executed for 500 episodes, each consisting of 100 time steps. The results for different algorithms are listed in Table 2, and details are summarized below.

Table 2: Comparison of UAV-based algorithm performance in 6G networks

Algorithm	Total connected users	Network convergence speed (Episodes)	Energy efficiency (Joules)	Latency (ms)	Interference (dBm)
CoRL	47	200	15	20	−92
C-QA	44	350	25	30	−100
D-QA	42	300	20	40	−108
Greedy	30	500	35	50	−110
DS-QoS	45	280	19	25	−99
MACE	43	290	21	30	−101
MCTS	44	275	20	28	−98

6.3.1 Total Connected Users

The CoRL-supported UAV system surpasses the other systems in supporting a larger number of connected users. C-QA is effective, but it involves a high communication overhead due to its centralized approach. D-QA slightly underperforms because it does not collaborate, and the greedy algorithm is the worst due to its myopic decision-making. In contrast, CoRL achieves maximum connectivity, dynamic adaptation, and resource-efficient allocation even surpassing the performance of recent decentralized schedulers such as DS-QoS, MACE, and MCTS.

6.3.2 Network Convergence Speed

The CoRL-based UAV system converges significantly less than C-QA and D-QA, primarily due to its distributed learning and cooperation. The greedy algorithm shows the least convergence and, in many cases, gets trapped in a local optimum. CoRL performs optimally in 200 episodes, compared to C-QA and D-QA, which exhibit more steady performance with slower convergence. Among the new decentralized schedulers, MACE shows higher convergence speed. Still, it is inferior to CoRL in terms of network performance because of the novelty-sharing mechanism, which needs more episodes to balance agents.

6.3.3 Energy Efficiency

CoRL-based UAV system utilizes the UAV's position and resource allocation optimization to reduce energy consumption. More power is required to run D-QA and C-QA due to insufficient power management, with the C-QA consuming the most power because it is centralized. The greedy algorithm also consumes the most energy due to its random decision-making process. It is shown that MCTS is a bit more energy efficient than D-QA and C-QA, but it still falls short compared to CoRL because of its pathfinding operations with partial observability.

6.3.4 Latency and Interference

The CoRL UAV system has the lowest latency and interference among all, primarily due to its cooperative behavior, in which UAVs share their states. For completeness, the latency metric was further broken down into signal transmission latency and algorithmic overhead. Signal transmission latency is defined as the end-to-end delay between UAVs and ground users. In contrast, algorithmic overhead is defined as the time spent on Q-value updates and transmission of the cooperative information between UAVs. The results demonstrate that transmission latency makes up the central part of total delay (around 85%–90%). In comparison, the algorithmic overhead is negligible (around 2–3 ms per decision step) even on resource-restricted onboard processors. Fig. 3 shows that the higher latency performance of the CoRL framework does not mask the hidden computational cost, confirming that it is suitable for real-time deployment in practical UAV systems. However, this low latency is associated with some algorithmic communication and coordination overheads. Specifically, the latency of CoRL is the sum of algorithm overhead (caused by the exchange of shared state and coordination among UAVs) and transmission delay of signals. This communication overhead can be a significant issue for resource-constrained UAVs, particularly when bandwidth or processing capacity is limited. In contrast, C-QA and D-QA have longer latencies, mainly due to their centralized or independent computation mechanisms, respectively. Due to the non-cooperative nature of UAVs, the Greedy Algorithm suffers from higher interference. MACE and DS-QoS exhibit higher latency than CoRL; however, DS-QoS offers better interference control, which still falls short of CoRL's overall efficiency. Future work would involve studying methods to minimize algorithmic overhead and thereby enhance deployability in constrained environments.
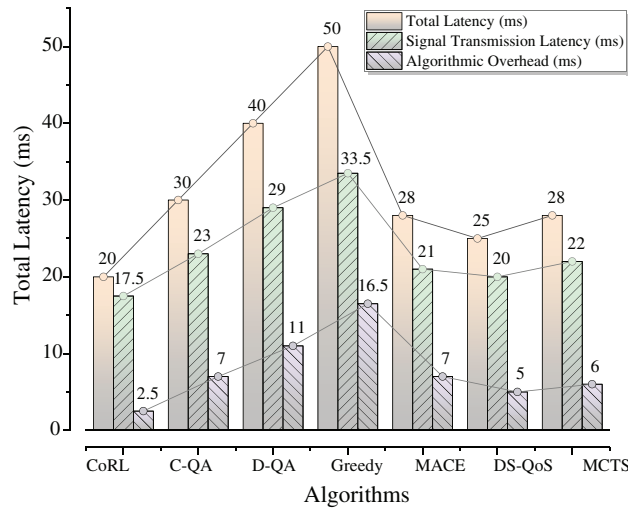

Figure 3: Latency breakdown for UAV-based algorithms in 6G networks

The results in Fig. 4 show that the CoRL-based UAV system offers better results by all metrics for the baseline algorithms. CoRL improves connectivity and convergence in UAVs, improving speed, energy efficiency, and latency by enabling them to collaborate and share information, making it an excellent choice for a 6G network. $\lambda \in [0.2, 0.6]$ and $\mu \in [0.1, 0.5]$ were varied while fixing $\kappa = 0.25$ and $\beta = 0.05$. Across 5 × 5 combinations (25 runs), CoRL maintained high connectivity (45–47 users), with a monotonic trade-off: higher λ reduced interference by 9%–14% but increased latency by 6%–9%, while higher μ saved 7%–12% energy at the cost of 1–2 fewer connected users when $\mu > 0.45$. The chosen setting ($\lambda = 0.40$, $\mu = 0.30$) sits

on the "elbow," balancing interference and energy without harming connectivity. It is decentralized, making it scalable and efficient, especially in large, dynamic networks.
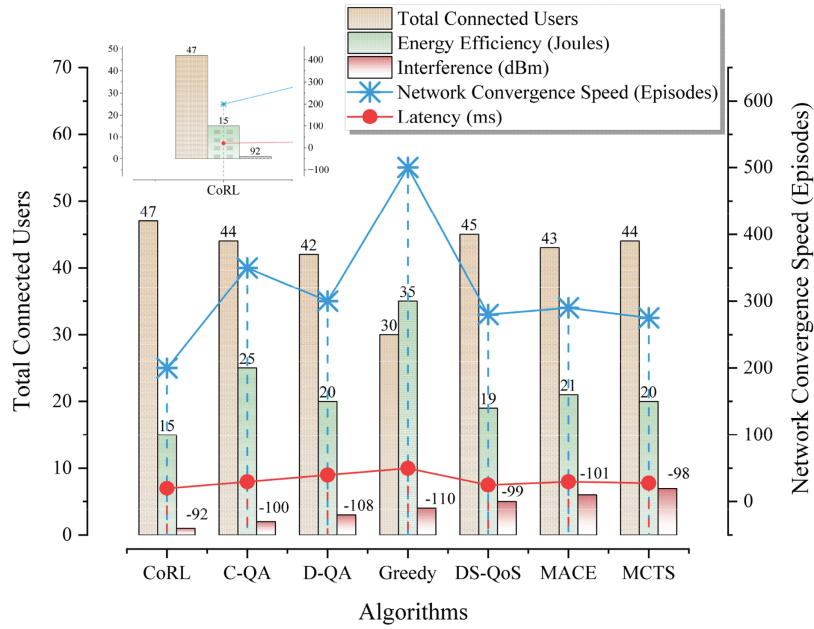

Figure 4: Comparison of UAV-based algorithm performance in 6G networks

6.4 Discussion

The findings demonstrate that CoRL has consistently outperformed centralized, decentralized, and greedy baselines. In addition to the numerical gains, some real-world advantages surface. Increased user connectivity has been used to demonstrate that CoRL can help ease congestion in dense networks. Additionally, lower latency and reduced interference have contributed to the development of CoRL applications in URLLC areas, such as autonomous driving and telemedicine. Enhanced energy efficiency increases UAV service time, reduces costs, and contributes to long, sustainable deployments. Increasingly faster convergence also demonstrates how CoRL can adapt to dynamically changing situations, and its capacity to operate in real-time 6G orchestration. Collectively, these findings suggest that CoRL can go beyond theoretical analysis to handle real 6G networks based on UAVs.

7 Conclusion

In this study, a cooperative reinforcement learning (CoRL)-based AI-controlled UAV framework has been proposed to leverage the ubiquitous 6G networks user connectivity. The method proposed is based on distributed and cooperative intelligence, in which UAVs learn policies of resource block allocation, user association, and mobility independently, and exchange state and reward information with their neighbors to enhance the overall network performance. The simulation findings proved that the CoRL-based system is always superior to the centralized and decentralized baselines on the main performance metrics, such as connectivity, latency, energy efficiency, and interference management. The framework can quickly adapt to evolving conditions of a network, making it applicable to large and dynamic environments like smart cities. Future directions will build upon this by adding adaptive modulation and coding scheme (MCS) selection at the per-RB level to give a more detailed physical-layer model, and beamforming-conscious positioning

to maximize UAV coverage in three-dimensional space. The extensions will improve additional spectral efficiency, equity, and real-time functionality in third-generation 6G networks.

Author Contributions: The authors confirm contribution to the paper as follows: Conceptualization, Zeeshan Ali Haider and Inam Ullah; methodology, Zeeshan Ali Haider and Inam Ullah; software, Zeeshan Ali Haider; validation, Zeeshan Ali Haider, Inam Ullah and Ahmad Abu Shareha; formal analysis, Rashid Nasimov and Sufyan Ali Memon; investigation, Ahmad Abu Shareha and Rashid Nasimov; resources, Zeeshan Ali Haider; data curation, Sufyan Ali Memon; writing—original draft preparation, Zeeshan Ali Haider; writing—review and editing, Inam Ullah; visualization, Rashid Nasimov and Sufyan Ali Memon; supervision, Inam Ullah; project administration, Inam Ullah; funding acquisition, Inam Ullah. All authors reviewed the results and approved the final version of the manuscript.

Availability of Data and Materials: Data will be available on reasonable request from the authors.

Ethics Approval: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest to report regarding the present study.

References

1. Cao X, Yang B, Wang K, Li X, Yu Z, Yuen C, et al. AI-empowered multiple access for 6G: a survey of spectrum sensing, protocol designs, and optimizations. Proc IEEE. 2024;112(9):1264–302.
2. Chen T, Wang S, Fan X, Zhang X, Luo C, Hong Y. UAV-assisted multi-object computing offloading for blockchain-enabled vehicle-to-everything systems. Comput Mater Contin. 2024;81(3):3927–50. doi:10.32604/cmc.2024.056961.
3. Lohan P, Kantarci B, Ferrag MA, Tihanyi N, Shi Y. From 5G to 6G networks: a survey on AI-based jamming and interference detection and mitigation. IEEE Open J Commun Society. 2024;5(19):3920–74. doi:10.1109/ojcoms.2024.3416808.
4. Haider ZA, Zeb A, Rahman T, Singh SK, Akram R, Arishi A, et al. A survey on anomaly detection in IoT: techniques, challenges, and opportunities with the integration of 6G. Comput Netw. 2025;270(7):111484. doi:10.1016/j.comnet.2025.111484.
5. Dai M, Huang N, Wu Y, Gao J, Su Z. Unmanned-aerial-vehicle-assisted wireless networks: advancements, challenges, and solutions. IEEE Internet Things J. 2022;10(5):4117–47. doi:10.1109/jiot.2022.3230786.
6. Cano MD, Guillen-Perez A, Tasic I, Villafranca A. A conceptual framework for the development of autonomous driving in 6G: the role of AI and edge computing. In: 2023 8th International Conference on Control, Robotics and Cybernetics (CRC); 2023 Dec 22–24; Changsha, China. p. 97–105.
7. Zhang H, Qi Z, Li J, Aronsson A, Bosch J, Olsson HH. 5G network on Wings: a deep reinforcement learning approach to the UAV-based integrated access and backhaul. IEEE Trans Machine Learning Commun Netw. 2024;2:1109–26. doi:10.1109/tmlcn.2024.3442771.
8. Baghnoi FM, Jamali J, Taghizadeh M, Fatehi MH. Multi-agent based optimal UAV deployment for throughput maximization in 5G communications. Wirel Netw. 2024;30(4):2285–96. doi:10.1007/s11276-023-03641-w.
9. Masaracchia A, Van Huynh D, Duong TQ, Dobre OA, Nallanathan A, Canberk B. The role of digital twin in 6G-based URLLCs: current contributions, research challenges, and next directions. IEEE Open J Commun Soc. 2025;6:1202–15. doi:10.1109/ojcoms.2025.3540287.

10. Alhammadi A, Shayea I, El-Saleh AA, Azmi MH, Ismail ZH, Kouhalvandi L, et al. Artificial intelligence in 6G wireless networks: opportunities, applications, and challenges. Int J Intell Syst. 2024;2024(1):8845070. doi:10.1155/2024/8845070.

11. Javaid S, Khalil RA, Saeed N, He B, Alouini MS. Leveraging large language models for integrated satellite-aerial-terrestrial networks: recent advances and future directions. IEEE Open J Commun Soc. 2025;6(2):399–432. doi:10.1109/ojcoms.2024.3522103.

12. Mao B, Tang F, Kawamoto Y, Kato N. AI models for green communications towards 6G. IEEE Commun Surv Tut. 2021;24(1):210–47.

13. Chen N, Cheng Z, Zhao Y, Huang L, Du X, Guizani M. Joint dynamic spectrum allocation for URLLC and eMBB in 6G networks. IEEE Trans Netw Sci Eng. 2023;11(6):5681–94. doi:10.1109/tnse.2023.3272013.

14. Tsekenis V, Barmpounakis S, Demestichas P. Flexible topologies for efficient network coverage expansion, sustainability and trust. In: 2025 IEEE Wireless Communications and Networking Conference (WCNC); 2025 Mar 24–27; Milan, Italy. p. 1–6.

15. Song F, Wang Z, Li J, Shi L, Chen W, Jin S. Dynamic trajectory and power control in ultra-dense UAV networks: a mean-field reinforcement learning approach. IEEE Trans Wirel Commun. 2025;24(7):5620–34. doi:10.1109/twc.2025.3548127.

16. Luat LB, Luong NC, Kim DI. Integrated radar and communication in ultra-reliable and low-latency communications-enabled UAV networks. IEEE Trans Veh Technol. 2025;74(8):13133–8. doi:10.1109/tvt.2025.3550585.

17. Li L, Xu G, Liu Z, Xu X, Meng X, Meng X. Multi-objective optimization of energy efficiency and fairness in UAV-assisted wireless powered MEC systems: a DRL-based approach. IEEE Internet Things J. 2025;12(14):28758–75. doi:10.1109/jiot.2025.3566865.

18. Fang X, Lei C, Feng W, Chen Y, Xiao M, Ge N, et al. Sensing-communication-computing-control closed-loop optimization for 6G digital twin-empowered unmanned robotic systems. IEEE J Sel Areas Commun. 2025. doi:10.1109/jsac.2025.3574601.

19. Lahmeri MA, Kishk MA, Alouini MS. Artificial intelligence for UAV-enabled wireless networks: a survey. IEEE Open J Commun Soc. 2021;2:1015–40. doi:10.1109/ojcoms.2021.3075201.

20. Lee S, Ban TW, Lee H. Network-wide energy efficiency maximization in UAV-aided IoT networks: quasi-distributed deep reinforcement learning approach. IEEE Internet Things J. 2025;12(11):15404–14. doi:10.1109/jiot.2025.3532477.

21. Zhou D, Sheng M, Bao C, Hao Q, Ji S, Li J. 6G non-terrestrial networks-enhanced IoT service coverage: injecting new vitality into ecological surveillance. IEEE Netw. 2024;38(4):63–71. doi:10.1109/mnet.2024.3382246.

22. Khan N, Coleri S, Abdallah A, Celik A, Eltawil AM. Explainable and robust artificial intelligence for trustworthy resource management in 6G networks. IEEE Commun Magaz. 2023;62(4):50–6. doi:10.1109/mcom.001.2300172.

23. Li J, Liu J, Wang J. Optimizing spectrum and energy efficiency in IRS-enabled UAV-ground communications. Comput Netw. 2025;256:110911. doi:10.1016/j.comnet.2024.110911.

24. Sehad N, Bariah L, Hamidouche W, Hellaoui H, Jantti R, Debbah M. Generative AI for immersive communication: the next frontier in internet-of-senses through 6G. IEEE Commun Magaz. 2025;63(2):31–43. doi:10.1109/mcom.001.2400199.

25. Haider ZA, Fayaz M, Zhang Y, Ali A. Advanced hyperelliptic curve-based authentication protocols for secure internet of drones communication. ICCK Trans Adv Computi Syst. 2024;1(4):1–16.

26. Mahmood K, Shamshad S, Anisi MH, Brighente A, Saleem MA, Das AK. A privacy-preserving access control protocol for 6G supported intelligent UAV networks. Veh Commun. 2025;54(12):100937. doi:10.1016/j.vehcom.2025.100937.

27. Lin Z, Feng Z, Guo K, Nauman A, Niyato D, Wang J. AI-driven seamless and massive access in space-air-ground integrated networks. IEEE Wirel Commun. 2025;32(3):72–9. doi:10.1109/mwc.001.2400371.

28. Nawaz MW, Zhang W, Flynn D, Zhang L, Swash R, Abbasi QH, et al. 6G edge-networks and multi-UAV knowledge fusion for urban autonomous vehicles. Phys Commun. 2024;67(6):102479. doi:10.1016/j.phycom.2024.102479.

29. Zhang G. 6G enabled UAV traffic management models using deep learning algorithms. Wirel Netw. 2024;30(8):6709–19. doi:10.1007/s11276-023-03485-4.

30. Waleed S, Ullah I, Khan WU, Rehman AU, Rahman T, Li S. Resource allocation of 5G network by exploiting particle swarm optimization. Iran J Comput Sci. 2021;4(3):211–9. doi:10.1007/s42044-021-00091-5.

31. Ullah I, Singh SK, Adhikari D, Khan H, Jiang W, Bai X. Multi-agent reinforcement learning for task allocation in the internet of vehicles: exploring benefits and paving the future. Swarm Evol Comput. 2025;94(1):101878. doi:10.1016/j.swevo.2025.101878.

32. Yilin D, Zhou Z, Rui W. Dynamic behavior recognition in aerial deployment of multi-segmented foldable-wing drones using variational autoencoders. Chin J Aeronaut. 2025;38(6):103397. doi:10.1016/j.cja.2025.103397.

33. Zuo Y, Guo J, Gao N, Zhu Y, Jin S, Li X. A survey of blockchain and artificial intelligence for 6G wireless communications. IEEE Commun Surv Tut. 2023;25(4):2494–528.

34. Danach K, Harb H, Rashid ASK, Al-Tarawneh MA, Aly WHF. Location planning techniques for Internet provider service unmanned aerial vehicles during crisis. Results Eng. 2025;25:103833. doi:10.1016/j.rineng.2024.103833.