

Article

Toward 6G: Latency-Optimized MEC Systems with UAV and RIS Integration

Abdullah Alshahrani

Department of Computer Science and Artificial Intelligence, College of Computer Science and Engineering, University of Jeddah, Jeddah 23218, Saudi Arabia; asalshahrani2@uj.edu.sa

Abstract: Multi-access edge computing (MEC) has emerged as a cornerstone technology for deploying 6G network services, offering efficient computation and ultra-low-latency communication. The integration of unmanned aerial vehicles (UAVs) and reconfigurable intelligent surfaces (RISs) further enhances wireless propagation, capacity, and coverage, presenting a transformative paradigm for next-generation networks. This paper addresses the critical challenge of task offloading and resource allocation in an MEC-based system, where a massive MIMO base station, serving multiple macro-cells, hosts the MEC server with support from a UAV-equipped RIS. We propose an optimization framework to minimize task execution latency for user equipment (UE) by jointly optimizing task offloading and communication resource allocation within this UAV-assisted, RIS-aided network. By modeling this problem as a Markov decision process (MDP) with a discrete-continuous hybrid action space, we develop a deep reinforcement learning (DRL) algorithm leveraging a hybrid space representation to solve it effectively. Extensive simulations validate the superiority of the proposed method, demonstrating significant latency reductions compared to state-of-the-art approaches, thereby advancing the feasibility of MEC in 6G networks.

Keywords: network protocols; artificial neural networks and deep learning; deep reinforcement learning; task offloading; mobile edge computing; reconfigurable intelligent surfaces; unmanned aerial vehicle; problem-solving in the context of artificial intelligence

MSC: 68M12; 68T07; 68T20



Received: 12 February 2025

Revised: 1 March 2025

Accepted: 4 March 2025

Published: 5 March 2025

Citation: Alshahrani, A. Toward 6G: Latency-Optimized MEC Systems with UAV and RIS Integration. *Mathematics* **2025**, *13*, 871. <https://doi.org/10.3390/math13050871>

Copyright: © 2025 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The rapid advancement of wireless technologies has driven the proliferation of portable intelligent devices, marking the dawn of the Internet of Things (IoT) era. According to the International Telecommunication Union (ITU), global mobile subscribers and data traffic are projected to grow substantially in the coming years, with mobile data traffic expected to reach approximately 5 zettabytes per month by 2030 [1]. This explosive growth, combined with the rising demand for data-intensive applications such as extended reality (XR), autonomous vehicles, and industrial automation, poses significant challenges to the capacity and sustainability of current 5G networks. To address the requirements of these emerging applications, future networks must achieve ultra-low latency and unparalleled reliability. Consequently, the development of 6G systems has become imperative to ensure robust, seamless, and efficient connectivity for the next generation of communication technologies (Ahmed et al. [2]).

The evolution toward 6G networks introduces significant advancements in communication efficiency and intelligent data processing, enabling seamless support for emerging

technologies and applications. A crucial component of this vision is multi-access edge computing (MEC), which offloads computationally intensive tasks from resource-constrained Internet of Things (IoT) devices to edge servers. By processing data closer to the source, MEC effectively minimizes latency and reduces energy consumption, addressing two critical challenges faced by modern IoT ecosystems. As next-generation networks aim to deliver ultra-reliable, low-latency, and energy-efficient communication, MEC stands out as a foundational technology (Tariq et al. [3]). Its integration with advanced 6G capabilities ensures robust performance for diverse applications, ranging from real-time analytics to autonomous operations.

The deployment of 6G networks demands innovations to improve wireless channel propagation, particularly in overcoming high penetration losses associated with THz frequencies in complex environments. To address these challenges, solutions such as massive multiple-input multiple-output (mMIMO) and hybrid analog-digital beamforming have been proposed, ensuring efficient and reliable communication (Ahmed et al. [4]). However, designing efficient multi-antenna transceivers for THz bandwidth introduces significant challenges due to the complexity of beamforming in such high-frequency domains. To tackle these issues, UAV-based communications, where unmanned aerial vehicles act as flying base stations to establish line-of-sight (LoS) links and reconfigurable intelligent surface (RIS)-assisted environments, which dynamically control signal propagation to enhance physical layer communication, have been proposed. These innovations aim to overcome the limitations of THz communication while enhancing network performance. Moreover, UAVs improve signal strength and network efficiency by providing flexible and dynamic communication links, while RIS, with its programmable metasurfaces, mitigates path loss and channel sparsity by redirecting signals to strengthen connections between base stations and users (Mehrabian and Wong [5]). By combining these technologies, RIS-assisted UAV communications effectively leverage the strengths of both approaches, offering a powerful solution for addressing key challenges in next-generation networks and garnering significant research attention (Xu et al. [6]).

However, despite these advantages, the practical deployment of RIS and UAVs introduces several challenges. UAV-based communication is constrained by high power consumption, limited battery life, and stringent airspace regulations, which impact operational feasibility (Banafaa et al. [7]). Additionally, RIS requires real-time reconfigurability and advanced control mechanisms to dynamically adjust signal propagation, posing implementation difficulties. Efficient energy management strategies, regulatory compliance frameworks, and intelligent control algorithms are essential to overcoming these obstacles and fully realizing the potential of RIS-UAV-assisted networks.

The integration of MEC solutions with RIS-assisted UAV communications holds immense potential for next-generation wireless networks. This approach combines the computational efficiency of MEC with the enhanced signal propagation and flexibility of RIS and UAV technologies, paving the way for innovative solutions to meet the stringent demands of future wireless systems.

1.1. Contributions

The implementation of MEC systems assisted by both UAV and RIS technologies has garnered significant attention from the research community due to their potential to enhance network performance. Recent studies (Xu et al. [8]), Hu et al. [9] have demonstrated that the complete integration of these technologies offers substantial advantages compared to the exclusive use of either UAV-based MEC systems or RIS-assisted MEC systems (Wang et al. [10]). However, most existing works that integrate UAVs and RIS have primarily considered RIS mounted on the facades of buildings. This limitation underscores the need

for exploring alternative RIS deployment strategies, such as UAV-assisted RIS, to fully realize the potential of these combined technologies.

Under these perspectives, this paper addresses the optimization problem of task offloading and resource allocation in MEC systems supported by a RIS-equipped UAV. The proposed optimization problem considers practical constraints, assuming perfect CSI at the MIMO base station (MBS), UAV energy limitations affecting its movement and service time, and configurable RIS phase shifts without additional delay overhead. These constraints ensure feasibility in real-world deployments.

The key contributions of this work to the current state of the art are as follows:

- We propose a novel optimization framework for a MEC system hosted within a massive MBS and assisted by a RIS-equipped UAV capable of flying within the coverage area. For this system, we formulate an optimization problem aimed at minimizing system latency by jointly optimizing the power allocation for each user, user association, phase shift configuration of RIS reflecting elements, and computing resource allocation at the MBS, all subject to the MBS's computing resource constraints and QoS requirements. Moreover, this optimization problem is formulated into a Markov decision process (MDP).
- Addressing the challenge of discrete-continuous hybrid action spaces, the paper proposes a unique solution that integrates a mechanism to represent discrete actions using an embedding table and leverages a conditional variational autoencoder to handle continuous actions. By integrating these into a unified hybrid action space, the framework leverages the twin delayed deep deterministic policy gradient (TD3) algorithm to solve the joint optimization problem effectively.
- Extensive simulation results reveal that the proposed algorithm, when compared with benchmark approaches, effectively minimizes the total latency for executing tasks across all users.

1.2. Paper Organization

The remainder of this paper is structured as follows: Section 2 provides a review of related work. Section 3 describes the system model under consideration, followed by the formulation of the optimization problem for latency minimization in Section 4. The proposed methodology is detailed in Section 5, while Section 6 evaluates the effectiveness of the proposed algorithm in minimizing total system latency. Finally, Section 7 presents the conclusion and discusses potential future research directions.

2. Related Work

The integration of line-of-sight (LoS) transmissions enabled by UAVs with the smart and controllable propagation environment created by reconfigurable intelligent surfaces (RIS) is a growing research focus for next-generation wireless networks. UAV-assisted MEC systems have demonstrated significant potential to enhance network computing capacity dynamically and support emergency scenarios. Task offloading in such systems has become a critical research area, addressing challenges in resource allocation optimization and ensuring network efficiency.

Recent advancements have leveraged UAVs in 5G networks to enhance task offloading, addressing critical challenges like resource constraints and latency. Ning et al. [11] proposed a 5G-enabled task offloading network that clusters users geographically and models the offloading problem as a mixed-integer non-linear programming (MINLP) problem to maximize throughput. Similarly, Song et al. [12] introduced an evolutionary multi-objective RL algorithm to optimize UAV trajectory and task offloading, with objectives to minimize task delay, reduce UAV energy consumption, and maximize task collection efficiency. These

works highlight the growing focus on integrating UAVs for efficient task management in next-generation networks.

Task dependencies have gained significant attention in UAV-assisted MEC systems due to their impact on resource allocation and system efficiency. Xu et al. [13] addressed this by formulating a joint optimization problem for resource allocation and UAV trajectory design, aiming to minimize system energy consumption while meeting delay and dependency constraints. Their approach combined dynamic programming and optimization algorithms to provide efficient solutions. However, their work was limited to a single UAV with constrained power and computational resources, underscoring the need for scalable and resource-efficient alternatives to support more complex and demanding scenarios.

Multi-UAV MEC systems address these limitations by enabling resource sharing and improving task offloading efficiency. For example, Goudarzi et al. [14] proposed a two-layer framework incorporating a queue-based algorithm to minimize delay, while Bai et al. [15] tackled non-convex load balancing challenges using Lyapunov optimization for online decision-making. Furthermore, Zhao et al. [16] introduced a multi-agent DRL approach based on the twin delayed deep deterministic policy gradient (TD3) algorithm to optimize UAV trajectories, power allocation, and task offloading, achieving significant cost reductions and scalability.

RIS technology has been explored to enhance MEC systems. Wang et al. [10] proposed a RIS-aided MEC system for heterogeneous networks (HetNet), focusing on the joint optimization of caching, task offloading, computing resources, and RIS/base station resource allocation. To address the NP-hard nature of this problem, they developed a two-stage optimization algorithm. Their findings indicate that integrating RIS into MEC HetNet systems can significantly reduce task computing delays, highlighting the potential of RIS to improve system performance.

The integration of RIS and UAV technologies has proven to significantly enhance MEC system performance through joint optimization. Xu et al. [8] proposed a RIS-enhanced, UAV-assisted MEC framework utilizing NOMA communication, focusing on optimizing RIS phase shifts, resource allocation, decoding order, and UAV deployment. Building on this, Hu et al. [9] incorporated dynamic UAV trajectories into the framework, further optimizing computation/offloading bits, RIS phase shifts, bandwidth allocation, and UAV trajectories. These advancements have demonstrated improved computation capacity, underscoring the effectiveness of combining RIS and UAV technologies in MEC systems for superior performance and resource efficiency.

3. System Model

We consider a mMIMO communication system that extends coverage between a MIMO base station (MBS) and distributed users, such as mobile users, vehicles, and IoT sensors, with the assistance of a RIS-equipped UAV. This integration significantly enhances network performance and quality of service (QoS), especially in areas affected by propagation blockages, improving reliability for UEs beyond direct MBS coverage (Zhai et al. [17]). In this system, the MBS is equipped with a large antenna array of L elements to serve K single-antenna users, which are grouped into M clusters represented by $\mathbf{K}_U = \{\mathcal{K}_1, \mathcal{K}_2, \dots, \mathcal{K}_M\}$. Each cluster has a varying number of users. The RIS-equipped UAV functions as a small-cell base station, with an RIS panel of N discrete elements reflecting signals from UEs to the MBS. The (m, k) -th UE denotes the k -th UE in the m -th cluster, where $u_{m,k}$ is the user association indicator, determining whether the UE offloads its computing task to the MEC server ($u_{m,k} = 1$ for offloading, $u_{m,k} = 0$ otherwise). User

associations are represented by the vector $u_m = [u_{m,k}]_{k=1}^{K_m}$ for cluster m and the matrix $\mathbf{u} = [u_m]_{m=1}^M$ for all clusters, it is expressed as follows:

$$u_{m,k} = \begin{cases} 1 & \text{if computing task offloading is needed,} \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

A time division multiple access (TDMA) scheme is employed when the UAV serves the m -th cluster, allowing users to offload tasks to the MBS. The RIS plays a critical role by steering beams to reflect only the desired signal toward the MBS, minimizing interference from other users and ensuring efficient communication (Di Renzo et al. [18]).

3.1. Channel Model

We consider a 3D Cartesian coordinate system to represent the spatial positions of the MBS, UAV-RIS, and users. The MBS is located at (x_0, y_0, H_0) , where H_0 denotes the antenna height, while the UAV-RIS is positioned at (x_U, y_U, H_U) , with H_U being the UAV's altitude. The users (UEs) are distributed at ground level, with their positions represented as $(x_k, y_k, 0)$, $k = 1, 2, \dots, K$. The positions of all entities, including the MBS, UAV-RIS, and UEs, are assumed to be accurately tracked using the GPS and stored locally at the MBS. Communication between the MBS and the UAV-RIS is modeled under a LoS assumption, ensuring high-quality signal propagation. For the channel between the MBS and UAV-RIS, we adopt a free-space path loss model to describe signal attenuation. This model accounts for the LoS condition, and the specific mathematical expression for the free-space path loss model is expressed as follows:

$$\beta_{m,0} = \frac{\beta_0}{d_{m,0}^2 + (H_0 - H_m)^2}, \quad m = 1, \dots, M, \tag{2}$$

where $d_{m,0} = \sqrt{(x_0 - x_m)^2 + (y_0 - y_m)^2}$, d_0 is defined as the distance, and β_0 signifies the power gain associated with the wireless channel. Additionally, H_m refers to the altitude of the UAV as it hovers above the m -th cluster.

The communication channel between the UAV-RIS and the (m, k) -th UE, classified as an air-to-ground channel, is more complex due to the effects of propagation blockages. To address this, the path-loss formulation incorporates both the air-to-air and air-to-ground links, represented as follows:

$$\beta_{m,k} = PL_{m,k} + \eta_{LoS} P_{m,k}^{LoS} + \eta_{NLoS} P_{m,k}^{NLoS}, \tag{3}$$

We define the average additional losses for the LoS and NLoS paths as η_{LoS} and η_{NLoS} , respectively. Based on these, the distance-based path loss can be expressed as follows:

$$PL_{m,k} = 10 \log \left(\frac{4\pi f_c D_{m,k}}{c} \right)^\alpha, \tag{4}$$

Here, α is the path loss exponent with $\alpha \geq 2$, while c represents the speed of light (in meters per second) and f_c denotes the carrier frequency (in Hertz). The Euclidean distance from the UAV-RIS to the (m, k) -th UE is represented as $d_{m,k}$, and the total Euclidean distance, considering the UAV height H_m , is given by $D_{m,k} = \sqrt{d_{m,k}^2 + H_m^2}$. As a result, the probability of LoS given as follows (Al-Hourani et al. [19]):

$$P_{m,k}^{LoS} = \frac{1}{1 + a \exp \left[-b \left(\arctan \left(\frac{H_m}{d_{m,k}} \right) - a \right) \right]} \tag{5}$$

where the values of the constants a and b are environment-dependent. Consequently, the probability of an NLoS connection can be expressed as $P_{m,k}^{\text{NLoS}} = 1 - P_{m,k}^{\text{LoS}}$. The free-space path loss model is chosen for the MBS-to-UAV-RIS link due to the highly probable LoS between the two, given the UAV-RIS's elevated positioning and the absence of significant obstacles. This assumption aligns with established aerial communication models and ensures accurate signal attenuation estimation. In contrast, the UAV-to-UE link is more susceptible to environmental factors such as urban obstructions and terrain variations. To account for these uncertainties, a probabilistic LoS/NLoS model is employed, effectively capturing the dynamic nature of air-to-ground signal propagation in UAV-assisted networks.

For UEs requiring assistance from the RIS to communicate with the MBS, the small-scale fading coefficients for the channels are represented as $h_{m,k} \in \mathbb{C}^{N \times 1}$ for the link from the (m, k) -th UE to the UAV-RIS, and $h_{m,0}^H \in \mathbb{C}^{N \times L}$ for the link from the UAV-RIS to the MBS, where $m \in M$ and $k \in K_m$. These coefficients are assumed to be independent and identically distributed (i.i.d.) random variables with zero mean and unit variance. The superscript H denotes the conjugate transpose operation. Furthermore, we define $H_{m,k} \in \mathbb{C}^{L \times N}$ and $H_{m,0}^H \in \mathbb{C}^{L \times N}$ as the channel matrices for the (m, k) -th UE to UAV-RIS link and the UAV-RIS to MBS link, respectively, where these matrices describe the small-scale fading effects over the respective multi-antenna channels are as follows:

$$H_{m,k} = \sqrt{\beta_{m,k}} h_{m,k}, \quad \text{and} \quad H_{m,0}^H = \sqrt{\beta_{m,0}} h_{m,0}^H. \tag{6}$$

Hence, the cascaded channel matrix of the link from the (m, k) -th UE to the MBS via the UAV-RIS, $G_{m,k} \in \mathbb{C}^{L \times 1}$, was described by Wu and Zhang [20]:

$$\mathbf{G}_{m,k} = \mathbf{H}_{m,0}^H \mathbf{\Phi}_m \mathbf{H}_{m,k}, \quad m \in \mathcal{M}, \tag{7}$$

where $\mathbf{\Phi}_m = \text{diag}[\phi_{1m}, \phi_{2m}, \dots, \phi_{Nm}]$ represents the phase shift matrix at the UAV-RIS. Each element $\phi_{nm} = \alpha_{nm} e^{j\theta_{nm}}$ characterizes the reflection properties of the n -th reflecting element, where $\alpha_{nm} \in [0, 1]$ denotes the reflection amplitude, and $\theta_{nm} \in [0, 2\pi]$ represents the phase shift. Here, $n = 1, 2, \dots, N$ and $m \in M$. Assuming that the RIS reflecting elements modify only the phase of the reflected signals, we set $\alpha_{nm} = 1$, ensuring maximum signal reflection without amplitude attenuation.

3.2. Transmission Scheme

Given that the (m, k) -th UE in the m -th group lacks a direct link to the MBS due to propagation blockages like large buildings, it offloads its computing task to the MBS by transmitting its signal via the RIS-equipped UAV. Consequently, the signal received at the MBS from the (m, k) -th UE is represented as follows:

$$y_{m,k} = \sqrt{P_{m,k}} \mathbf{G}_{m,k}^H \mathbf{f}_{m,k} s_{m,k} + \sum_{l=1, l \neq k}^{K_m} \sqrt{P_{l,m}} \mathbf{G}_{l,m}^H \mathbf{f}_{l,m} s_{l,m} + n_0, \tag{8}$$

here, $P_{m,k}$ represents the transmit power of the (m, k) -th UE, $\mathbf{f}_{m,k} \in \mathbb{C}^{L \times 1}$ is the MBS beamforming vector for the (m, k) -th UE, and $s_{m,k}$ is the offloading information transmitted by the UE with $|s_{m,k}|^2 \leq 1$. The AWGN at the MBS is denoted by $n_0 \sim \mathcal{CN}(0, \sigma_0^2)$, and the maximum transmit power of the UE is $P_{m,k}^{\text{max}}$. The first term in (8) represents the signal transmitted by the (m, k) -th UE via the RIS panel, while the second term captures intra-cell interference from other UEs in the m -th group transmitting simultaneously. To mitigate this interference, the zero-forcing (ZF) technique is defined as follows:

$$\mathbf{G}_m = [\mathbf{G}_{m,1}, \dots, \mathbf{G}_{m,K_m}] \in \mathbb{C}^{L \times K_m} \quad (m = 0, 1, \dots, M). \tag{9}$$

As the eigenvalue distribution of the square matrix $\mathbf{G}_m^H \mathbf{G}_m \in \mathbb{C}^{K_m \times K_m}$ becomes increasingly deterministic with larger L , we leverage this property of the mMIMO system to design the beamforming vector $\mathbf{f}_{m,k}$ using the zero-forcing (ZF) technique, expressed as follows:

$$\tilde{\mathbf{f}}_m = \mathbf{G}_m (\mathbf{G}_m^H \mathbf{G}_m)^{-1}, \tag{10}$$

where $\tilde{\mathbf{f}}_m = [\tilde{\mathbf{f}}_{m,1}, \dots, \tilde{\mathbf{f}}_{m,K_m}] \in \mathbb{C}^{L \times K_m}$, $\tilde{\mathbf{f}}_{m,k} \in \mathbb{C}^{L \times 1}$, $m = 0, 1, \dots, M, k \in K_m$. The matrix $\tilde{\mathbf{f}}_m$ is obtained from the ZF beamforming solution given in (10), where $\tilde{\mathbf{f}}_m$ represents the precoding matrix computed based on the pseudo-inverse of the channel matrix \mathbf{G}_m . Each column $\tilde{\mathbf{f}}_{m,k}$ corresponds to the beamforming vector for user k within cluster m . To ensure unit-norm beamforming vectors, we normalize $\tilde{\mathbf{f}}_{m,k} = \tilde{\mathbf{f}}_{m,k} / \|\tilde{\mathbf{f}}_{m,k}\|$ and then compute the final beamforming vector $\mathbf{f}_{m,k}$ as follows:

$$\mathbf{f}_{m,k} = \sqrt{p_{m,k}} \tilde{\mathbf{f}}_{m,k}, \quad m = 0, 1, \dots, M, k \in K_m, \tag{11}$$

where $p_{m,k}$ represents the power control coefficient of the (m, k) -th UE. Consequently, Equation (8) can be rewritten as follows and the intra-cell interference has been successfully canceled.

$$y_{m,k} = \sqrt{P_{m,k}} \sqrt{p_{m,k}} \mathbf{G}_{m,k}^H \mathbf{f}_{m,k} s_{m,k} + n_0, \tag{12}$$

Let $\mathbf{p}_m = [p_{m,k}]_{k=1}^{K_m}$ represent the power control coefficients, and $\mathbf{p} = [\mathbf{p}_m]_{m=0}^M$ aggregate these coefficients across all panels. Similarly, let $\Phi = [\Phi_m]_{m=1}^M$ denote the phase shift configuration of the RIS panels. The achievable throughput for the transmission of the (m, k) -th UE to the MBS can be expressed as follows:

$$R_{m,k}(p_{m,k}, \Phi_m) = W \log_2 \left(1 + \frac{P_{m,k} p_{m,k} |\mathbf{G}_{m,k}^H \tilde{\mathbf{f}}_{m,k}|^2}{\sigma_0^2} \right), \tag{13}$$

where W is the allocated bandwidth to the (m, k) -th UE.

3.3. Offloading Model

In the context of computation modeling, it is assumed that a specific task of size $\mathcal{I}_{m,k}$ can be processed either locally by the (m, k) -th UE or offloaded for execution to the MEC server housed at the MBS. Accordingly, two models are introduced to characterize the computation latency, as outlined below.

3.3.1. Local Computing

Let $F_{m,k}$ denote the number of CPU cycles required to process each bit of the task for the (m, k) -th UE. Consequently, the time taken to execute the task locally can be calculated as follows, according to Raza et al. [21]:

$$T_{m,k}^l = \frac{\mathcal{I}_{m,k} \mathcal{F}_{m,k}}{c_{m,k}}, \quad m = 0, 1, \dots, M, k \in K_m, \tag{14}$$

where $c_{m,k}$ denotes the upper limit of computational resources available to the (m, k) -th UE.

3.3.2. Offloading to MBS

Alternatively, when the task is offloaded from the (m, k) -th user equipment (UE) to the MBS, the transmission time for offloading must be considered, which is given as follows, according to Raza et al. [21]:

$$T_{m,k}^{tx}(p_{m,k}, \Phi_m) = \frac{\mathcal{I}_{m,k}}{R_{m,k}(p_{m,k}, \Phi_m)}, \tag{15}$$

where $R_{m,k}(p_{m,k}, \Phi_m)$ represents the communication rate defined in (11). After the task is successfully transmitted to the MBS, the computation time required to process the offloaded task at the MBS can be expressed as follows:

$$T_{m,k}^{com}(\zeta_{m,k}^{bs}) = \frac{\mathcal{I}_{m,k} \mathcal{F}_{m,k}}{\zeta_{m,k}^{bs}}, \quad m = 0, 1, \dots, M, \quad k \in \mathcal{K}_m, \tag{16}$$

where $\zeta_{m,k}^{bs}$ represents the computing capacity of the MBS allocated to process the task for the (m, k) -th UE. For simplicity, let $\zeta_m = [\zeta_{m,k}^{bs}]_{k=1}^{K_m}$ and $\zeta = [\zeta_m]_{m=0}^M$ represent the overall MBS computing capacity allocation. The total latency for executing the task of the (m, k) -th UE can therefore be expressed as follows:

$$T_{Total} = T_{m,k}^{tot}(p_{m,k}, u_{m,k}, \Phi_m, \zeta_{m,k}^{bs}) = (1 - u_{m,k})T_{m,k}^l + u_{m,k} \left(T_{m,k}^{tx}(p_{m,k}, \Phi_m) + T_{m,k}^{com}(\zeta_{m,k}^{bs}) \right). \tag{17}$$

The time required to transmit the computation results from the MBS back to the UEs can be neglected, as this latency is significantly smaller than the total latency incurred during task execution (Raza et al. [22,23]).

4. Problem Formulation

The UAV-RIS-assisted MEC model optimization aims to minimize total latency for task execution by jointly optimizing power allocation (\mathbf{p}), user association (\mathbf{u}), RIS phase shift matrix (Φ), and computing resource allocation (\mathbf{i}). Constraints ensure the MBS computing capacity is not exceeded and QoS requirements for all users are met. The problem is formulated to balance resources and reduce latency effectively. The detailed formulation is as follows:

$$\min_{\mathbf{p}, \mathbf{u}, \Phi, \zeta} \sum_{m=0}^M \sum_{k=1}^{K_m} T_{Total} \tag{18}$$

$$\text{s.t.} \quad 0 \leq p_{m,k} \leq 1, \tag{18a}$$

$$R_{m,k}(p_{m,k}, \Phi_m) \geq \bar{r}_0, \quad m = 0, 1, \dots, M, \quad k \in \mathcal{K}_m, \tag{18b}$$

$$0 \leq \theta_{nm} \leq 2\pi, \quad \forall n = 1, 2, \dots, N, \quad m \in \mathcal{M}, \tag{18c}$$

$$\sum_{k=1}^{K_m} u_{m,k} \zeta_{m,k}^{bs} \leq \zeta_{max}, \tag{18d}$$

The constraints are defined as follows: $u_{m,k}$ represents the user association coefficient as per (1). Constraint (18a) specifies the allowable power range for each user, while (18b) ensures that users meet the QoS requirement in terms of a minimum achievable uplink rate (\bar{r}_0). Constraint (18c) defines the valid range for each RIS phase-shift coefficient. Finally, (18d) reflects the MEC server’s computing resource limits that can be allocated to each UE.

5. Proposed Scheme

The proposed scheme aims to enhance task offloading efficiency in a UAV-RIS-assisted MEC model by systematically addressing decision-making complexities. By reformulating the problem as an MDP (Ahmed et al. [24]), we enabled a structured framework to optimize offloading decisions, power allocation, user associations, and resource distribution. The following sections outline this MDP formulation in detail:

5.1. MDP Formulation

In the UAV-RIS-assisted MEC model, we optimize the offloading decision, the power allocated by each user (\mathbf{p}), user association (\mathbf{u}), the phase shift matrix of the RIS (Φ), and computing resource allocation at the MBS ($\mathbf{1}$) subject to the MBS computing resource constraints and QoS requirements. The task offloading problem is modeled as an MDP, where the system state at the next time slot is determined solely by the current state and the action taken at the present time. In each time slot, the system observes the current state, selects an action based on this observation, and generates a reward that reflects the effectiveness of the action. The primary objective is to minimize system latency by utilizing an optimization scheme that effectively maps system states to latency-reducing actions.

5.1.1. State Space

The state captures essential information required for decision-making while avoiding redundancy. It includes three main factors: the task size ($\mathcal{I}_{m,k}$), CPU cycles per bit ($\mathcal{F}_{m,k}$), and the minimum required uplink rate (\bar{r}_0).

The task size ($\mathcal{I}_{m,k}$) represents the data amount that needs to be processed, measured in bits. This directly impacts the computation and transmission latency. CPU cycles per bit ($\mathcal{F}_{m,k}$) denote the computational effort required to process one bit of data, influencing the processing time both locally and at the MEC server. The minimum required uplink rate (\bar{r}_0) serves as a QoS constraint to ensure adequate communication performance. These three factors are mathematically expressed as follows:

$$s(t) = \{\mathcal{I}_{m,k}, \mathcal{F}_{m,k}, \bar{r}_0\}. \quad (19)$$

These components are sufficient to derive or calculate other interrelated variables, such as path loss, LoS probability, and computational capacity.

5.1.2. Action State

The action space includes the control variables necessary for optimizing system performance. The first component is the transmission power allocation ($p_{m,k}$), which specifies the transmission power for each user and is bounded by the user's maximum allowable power:

$$0 \leq P_{m,k} \leq P_{m,k}^{\max}. \quad (20)$$

The second component, user association ($u_{m,k}$), determines whether the user's task is offloaded to the MEC server or processed locally. The third component involves optimizing the phase shifts (Φ_m) of the RIS panel to improve signal quality. This is expressed in (7). Lastly, the MEC resource allocation ($\zeta_{m,k}^{\text{bs}}$) determines the computational capacity assigned to each user, constrained by the MEC server's total capacity as expressed in Equation (18d).

Thus, the action space is represented as follows:

$$a(t) = \{p_{m,k}, u_{m,k}, \Phi_m, \zeta_{m,k}^{\text{bs}}\}. \quad (21)$$

5.1.3. Reward Function

The reward function is structured to align with the objective of minimizing the total latency for all users. The total latency for a user (m, k) is defined as the sum of local computation, transmission, and MEC processing times as expressed in (17). To ensure the agent minimizes this latency, the reward function includes a latency term:

$$r(t) = - \sum_{m=0}^M \sum_{k=1}^{K_m} T_{m,k}^{\text{tot}}. \quad (22)$$

5.2. DRL-Based Algorithm

The proposed DRL-based algorithm addresses the challenges of the hybrid discrete-continuous action space in the MDP, which conventional DRL methods cannot handle effectively. Converting to a discrete space causes scalability issues, while using a continuous space increases approximation complexity, both of which degrade performance. To resolve this, a novel algorithm is designed specifically for hybrid action spaces, ensuring efficient and robust decision-making.

5.2.1. Hybrid Space Modeling in UAV-RIS-Assisted MEC

In the UAV-RIS-assisted MEC model, converting the MDP into a hybrid space representation facilitates efficient optimization of the hybrid discrete-continuous action space. This approach simplifies the complexity inherent in optimizing offloading decisions, user associations, power allocation, RIS phase shifts, and MEC resource allocations while ensuring scalability and effective policy learning.

The hybrid space is constructed by first encoding the discrete components of the action space, such as user association $(u_{m,k})$ and RIS phase shifts (Φ_m) . An embedding table $G_\omega \in \mathbb{R}^{N_u \times l_1}$ is utilized to map each discrete action to a continuous vector representation, ensuring that the discrete variables are represented compactly and consistently. Each embedding $g_{\omega,u} = G_\omega(u)$ captures the essential features of the discrete actions, reducing the dimensionality and enabling effective integration with continuous variables.

For the continuous components of the action space, including transmission power $(p_{m,k})$ and MEC resource allocation $(\zeta_{m,k}^{\text{bs}})$, a conditional variational autoencoder (VAE) is employed. The VAE encoder $q_\phi(\mathbf{z}|\mathbf{p}, \mathbf{1}, s, g_{\omega,u})$ maps these variables into a hybrid representation $\mathbf{z} \in \mathbb{R}^{l_2}$, conditioned on the current state $s(t)$ and the embeddings of discrete actions $g_{\omega,u}$. This hybrid representation encapsulates the interdependencies between discrete and continuous variables, maintaining a unified and compact action representation.

The resulting hybrid representation combines the discrete embeddings and the continuous hybrid variables into a single hybrid vector:

$$h = [g_{\omega,u}, \mathbf{z}] \in \mathbb{R}^{Ml_1+l_2}. \quad (23)$$

This vector serves as the input to the reinforcement learning algorithm. During execution, the hybrid vector is decoded back into the original action space. Discrete actions $(u_{m,k}, \Phi_m)$ are retrieved through a nearest neighbor search in the embedding table:

$$u_{m,k} = \arg \min_u \|g_{\omega,u} - g\|_2, \quad (24)$$

while continuous actions $(p_{m,k}, \zeta_{m,k}^{\text{bs}})$ are reconstructed using the VAE decoder q_ψ :

$$\mathbf{p}, \mathbf{1} = q_\psi(\mathbf{z}, s, g_{\omega,u}). \quad (25)$$

The hybrid space is trained to accurately reconstruct the original action space through a combined loss function. The reconstruction loss ensures the decoded actions match the originals:

$$\mathcal{L}_{\text{recon}} = \|\mathbf{p} - \hat{\mathbf{p}}\|_2^2 + \|\mathbf{1} - \hat{\mathbf{1}}\|_2^2, \tag{26}$$

and the Kullback–Leibler (KL) divergence term regularizes the hybrid variables for smooth and meaningful representations:

$$\mathcal{L}_{\text{KL}} = D_{\text{KL}}(q_{\phi}(\mathbf{z}|\mathbf{p}, \mathbf{1}, s, g_{\omega, u}) || \mathcal{N}(0, I)). \tag{27}$$

The final training loss for the hybrid space is a combination of these terms:

$$\mathcal{L} = \mathcal{L}_{\text{recon}} + \beta \mathcal{L}_{\text{KL}}, \tag{28}$$

where β is a weighting parameter to balance reconstruction and regularization.

Integrating the hybrid space with reinforcement learning algorithms like TD3 ensures that the compact representation is directly used for policy and value function approximations. The actor network outputs the hybrid representation $h = [g_{\omega, u}, \mathbf{z}]$, while the critic evaluates the Q-value for the hybrid actions. This hybrid representation enables efficient optimization of the hybrid action space while preserving the relationships between discrete and continuous components.

The hybrid space significantly reduces the dimensionality of the action space, enhancing computational efficiency and convergence stability. Furthermore, it captures the relationships between discrete and continuous actions, improving the overall performance in dynamic and large-scale UAV-RIS-assisted MEC environments. This methodology not only simplifies hybrid action optimization but also ensures scalability and adaptability in complex, resource-constrained scenarios, making it ideal for minimizing latency and optimizing system performance in UAV-RIS MEC systems.

5.2.2. Long-Term Latency Minimization Algorithm for UAV-RIS-Assisted MEC Systems

The long-term latency minimization algorithm for UAV-RIS-assisted MEC systems integrates the hybrid action representation space with the TD3 algorithm to address the offloading problem. The TD3, designed for deterministic strategy, excels in handling high-dimensional continuous action spaces. Its key components include an actor network, which maps system states to actions and drives decision-making, and a critic network, which estimates rewards for specific state-action pairs to refine decisions. This integration ensures effective task offloading and optimization, aiming to minimize long-term system latency.

The proposed approach employs a hybrid action space representation to enhance decision-making in UAV-RIS-assisted MEC systems. The actor network utilizes the system state s as input and outputs a hybrid action vector (g, z) , where $g \in \mathbb{R}^{M_1}$ and $z \in \mathbb{R}^l$. A decoder processes (g, z) , mapping it to a hybrid action $a = \{i_{m,k}, p_{m,k}\}, \forall m, k$.

The hybrid action value is approximated using twin critic networks, Q_{θ_1} and Q_{θ_2} , which utilize the hybrid action a as input and evaluate its value. The training process leverages experience tuples (s, a, r, s') stored in a replay buffer D . The critic networks are optimized through the Double Q-Learning algorithm, ensuring stability and precise value approximation.

The loss function for critic networks is defined as follows:

$$L_{CDQ}(\theta_j) = \mathbb{E} \left[(\zeta - Q_{\theta_j}(s, g, z))^2 \right], \quad \forall j = 1, 2, \tag{29}$$

where $\zeta = r + \gamma \min Q_{\hat{\theta}_j}(s', \pi_{\zeta}(s'))$, and $\hat{\theta}_j$ represent the parameters of the target network.

The actor network, employing a hybrid policy, is updated using the deterministic policy gradient method, expressed as follows:

$$\nabla_{\zeta} J(\zeta) = \mathbb{E} \left[\nabla_{\pi_{\zeta}(s)} Q_{\theta_1}(s, \pi_{\zeta}(s)) \nabla_{\zeta} \pi_{\zeta}(s) \right]. \quad (30)$$

By integrating the hybrid representation space with the TD3, we propose a novel approach to address the optimization problem in UAV-RIS-assisted MEC systems. This approach effectively balances exploration and exploitation in hybrid action spaces, ensuring robust performance in dynamic network environments.

In Algorithm 1, the actor and critic networks, along with the discrete action table and VAE parameters, are initialized to manage the hybrid action space (lines 1–2). The initial system state and replay buffer are then prepared (lines 3–4). A warm-up phase follows, where the VAE parameters are trained using reconstruction and regularization losses (lines 5–7). During each environment step, the current state is observed, and hybrid actions are generated by the actor network with added exploration noise (lines 8–10). These hybrid actions are decoded into discrete actions (e.g., user association and RIS phase shifts) and continuous actions (e.g., power and resource allocation) for execution, producing rewards and the next state (lines 11–13). Transition tuples are stored in the replay buffer for training (line 14). Critic networks are updated using loss functions, while the actor network is refined via policy gradients (lines 15–17). Representation parameters are periodically updated for efficient action decoding (lines 18–20). The algorithm iterates until convergence or the maximum steps are reached (lines 8–22). The framework of the proposed algorithm is illustrated in Figure 1.

Algorithm 1 Proposed algorithm for the UAV-RIS-assisted MEC optimization.

- 1: Initialize the actor network π_{ζ} and critic networks $Q_{\theta_1}, Q_{\theta_2}$ with $\zeta, \theta_1, \theta_2$ parameters;
 - 2: Initialize embedding table G_{ω} and ϕ, ψ, ω ;
 - 3: Initialize system state s_1 as in Equation (19);
 - 4: Initialize a replay buffer, denoted as \mathcal{D} ;
 - 5: **while** warm-up iterations not complete **do**
 - 6: Update ω, ϕ, ψ with samples in \mathcal{D} as per Equation (28);
 - 7: **end while**
 - 8: **while** environment steps not complete **do**
 - 9: Monitor the current system state s ;
 - 10: Choose hybrid actions by actor network
 - 11: $g, z = \pi_{\zeta}(s) + \epsilon_g, \epsilon_g \sim \mathcal{N}(0, \sigma)$;
 - 12: Decode $u_{m,k}, \Phi_m = f_D(g), p_{m,k}, \zeta_{m,k}^{\text{bs}} = q_{\psi}(z, s, g)$ as per (20), (21), (24), and (25);
 - 13: Execute $\{u_{m,k}, \Phi_m, p_{m,k}, \zeta_{m,k}^{\text{bs}}\}$, obtain reward r as per Equation (22) and next state s' ;
 - 14: Store (s, g, z, r, s') in \mathcal{D} ;
 - 15: Analyze hybrid actions using the critic network.
 - 16: Sample mini-batch from \mathcal{D} ;
 - 17: Update $Q_{\theta_1}, Q_{\theta_2}$ as per Equation (29);
 - 18: Update π_{ζ} with according to Equation (30);
 - 19: **while** training iterations not complete **do**
 - 20: Update ω, ϕ, ψ using \mathcal{D} as per Equation (28);
 - 21: **end while**
 - 22: **end while**
-

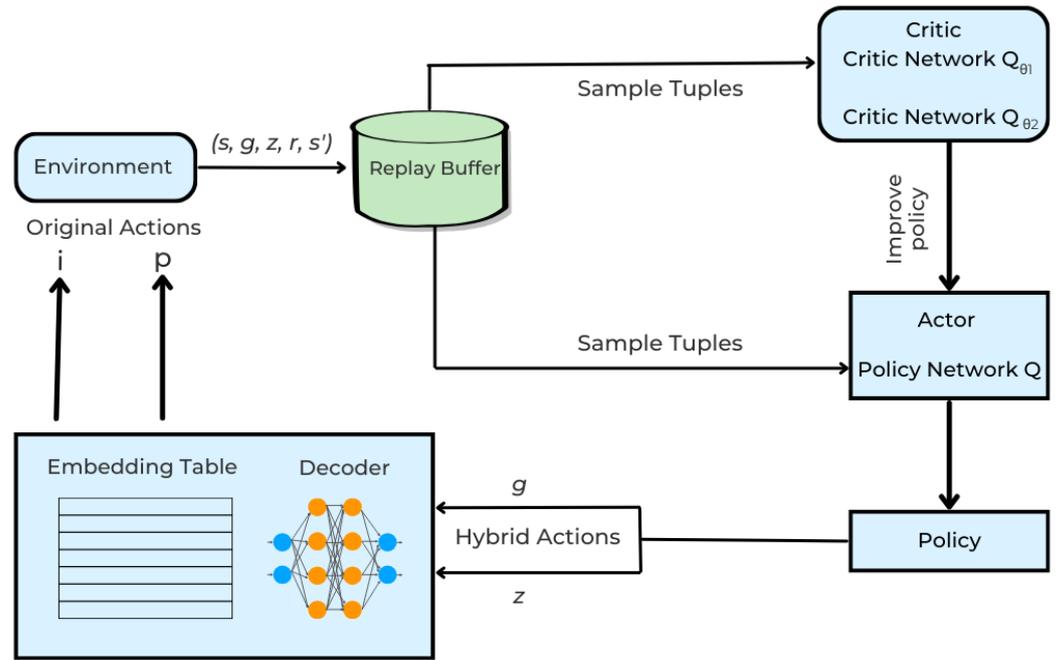


Figure 1. Framework of the proposed algorithm.

5.2.3. Computation Complexity

The computational complexity of the proposed algorithm is analyzed across two key phases: the warm-up phase and the environment interaction phase. The warm-up phase involves updating the parameters ω, φ, ψ using stored samples from the replay buffer \mathcal{D} via stochastic gradient descent (SGD). Let N_{warm} denote the number of warm-up iterations, and assume each neural network layer has $O(F)$ parameters with L layers. The complexity of this phase is approximately $O(N_{\text{warm}} \cdot L \cdot F^2)$.

In the environment interaction phase, the actor network π_{ζ} selects hybrid actions with a complexity of $O(F)$. Decoding and retrieving discrete actions (e.g., user association $u_{m,k}$, RIS phase shifts Φ_m) and reconstructing continuous actions (e.g., power allocation $p_{m,k}$, computing resource allocation $\zeta_{m,k}^{bs}$) require matrix operations and nonlinear activations, leading to $O(F^2)$. The reward computation in (22) incurs a negligible $O(1)$ cost.

Once the transition (s, g, z, r, s') is stored in \mathcal{D} , the critic networks $Q_{\theta_1}, Q_{\theta_2}$ update using mini-batches of size B , following (29), with a complexity of $O(B \cdot L \cdot F^2)$. The actor network update, as per (30), incurs the same cost, while additional training iterations contribute $O(N_{\text{train}} \cdot L \cdot F^2)$.

Summing all contributions, the overall complexity for T environment steps can be expressed as follows:

$$O(N_{\text{warm}} \cdot L \cdot F^2) + O(T \cdot (F + F^2 + B \cdot L \cdot F^2 + N_{\text{train}} \cdot L \cdot F^2)). \tag{31}$$

Since the dominant term is $O(T \cdot B \cdot L \cdot F^2)$, scalability depends on batch size B , network depth L , and feature count F . Additionally, the number of user equipment (UE) units K and RIS elements M directly influence the action space complexity, increasing computational overhead in action decoding and network updates.

Moreover, in large-scale mMIMO systems, ZF beamforming introduces an additional computational burden due to the required matrix inversion, resulting in a complexity of $O(K^3)$, where K is the number of UEs. This cubic complexity can significantly impact real-time processing, especially as the number of UEs grows. Despite this, the proposed framework remains computationally feasible due to its structured hybrid action space and

efficient optimization strategy, making it suitable for large-scale deployments in MEC-assisted networks.

6. Numerical Results

This section provides numerical results and a discussion of our proposed algorithm in terms of the latency minimization when compared with other state-of-the-art schemes.

The simulation scenario is designed to evaluate the performance of the proposed MEC system. A massive MBS is positioned at coordinates (0,0,30) meters, providing direct service to users within a 500 m radius. To extend coverage for users located up to 2000 m away, a RIS-equipped UAV operates within the area at an altitude ranging from 50 to 150 m, with user equipment (UE) randomly distributed across the extended region (Nguyen et al. [25]). The transmission parameters include an uplink transmission power of 30 dBm, a central frequency of 2.4 GHz, a bandwidth of 1 MHz, and a noise spectral density of -130 dBm/Hz. To ensure quality of service (QoS), the minimum achievable uplink rate per UE is set at 1 Mbps. The computational tasks involve a task size of 100 kB with a computational complexity of 600 cycles per bit. The MEC server, hosted at the MBS, has a maximum processing capacity of 30 Giga cycles per second, while each UE is limited to a maximum of 0.5 Giga cycles per second. This setup provides a realistic and challenging environment for evaluating task offloading and resource allocation in the RIS-assisted UAV-enabled MEC framework (Liu et al. [26]). The detailed configuration of the simulation parameters is presented in Table 1.

Table 1. Simulation parameters.

| Parameter | Value |
|---|-------------------|
| MBS coverage radius for direct connection | 500 m |
| Maximum distance of users from MBS | 2000 m |
| 3D Cartesian coordinates of the MBS | (0, 0, 30) |
| Altitude range of RIS-equipped UAV (H^{\min}, H^{\max}) | (50, 150) m |
| Maximum transmission power ($P_{m,k}^{\max}$) | 30 dBm |
| Central frequency (f_c) | 2.4 GHz |
| Bandwidth (W) | 1 MHz |
| White noise spectral density (σ_0^2) | -130 dBm/Hz |
| Minimum achievable uplink rate (\tilde{r}_0) | 1 Mbps |
| Task size ($\mathcal{I}_{m,k}$) | 100 kB |
| Task computation complexity ($\mathcal{F}_{m,k}$) | 600 cycles/bit |
| Maximum MEC server computing resource (ζ_{\max}) | 30 Giga cycles/s |
| Maximum UE computing resource ($c_{m,k}$) | 0.5 Giga cycles/s |

The proposed scheme is compared with the following baseline schemes.

1. Deep Q-network (DQN) (Khan et al. [27]): DQN algorithm is designed for discrete action spaces, where a neural network approximates the Q-value function to derive optimal policies.
2. Soft actor-critic (SAC) (Heidarpour et al. [28]): SAC is a model-free, off-policy algorithm optimized for continuous action spaces. It employs a stochastic policy with an entropy-based objective to balance exploration and exploitation. While SAC performs well in purely continuous environments.
3. Multi-Agent deep deterministic policy gradient (MADDPG) (Tariq et al. [3]): MADDPG extends deterministic policy gradient methods to multi-agent systems, facilitating coordinated decision-making among RIS-equipped UAVs and MEC servers.
4. Proposed scheme: The proposed approach employs a hybrid discrete-continuous optimization framework with hybrid space representation to optimize task offload-

ing, power allocation, and RIS phase configurations. By integrating a conditional variational autoencoder and the TD3 algorithm, it addresses hybrid action spaces.

Figure 2 illustrates the average reward progression for various DRL algorithms, i.e., the proposed scheme, DQN, SAC, and MADDPG over 2000 training episodes. The average reward reflects the capability of each method in addressing the task offloading and resource allocation challenges in a UAV-RIS-assisted MEC environment. The results demonstrate the superiority of the proposed scheme, which achieves the highest average reward by the end of the training. This performance is attributed to its hybrid discrete-continuous optimization framework, effectively utilizing hybrid representations and advanced policy learning to navigate complex decision spaces. Among the baseline methods, the MADDPG algorithm shows competitive performance due to its multi-agent coordination capabilities. However, its effectiveness is hindered by inefficiencies in handling hybrid action spaces, an area where the proposed scheme excels. SAC achieves moderate results, optimized for continuous actions but constrained by its lack of inter-agent coordination. Meanwhile, DQN, with its discrete action space, shows the weakest performance, highlighting its limitations in handling the complexity and scale of the optimization problem. For clarity, Figure 2 incorporates shaded confidence intervals that represent the variability in rewards across multiple training trials. These confidence intervals highlight the stability of each algorithm, with the proposed scheme exhibiting a narrower range of fluctuations, indicating more consistent performance. MADDPG and SAC show moderate variance, while DQN not only achieves the lowest average reward but also displays the widest confidence interval, reinforcing its instability in dynamic UAV-RIS-MEC environments. Despite varying performance levels, all algorithms demonstrate steady learning progression, with rewards increasing as training episodes advance. However, the faster convergence, higher average reward, and reduced variability of the proposed scheme underscore its robustness for latency-critical, resource-intensive applications. These findings emphasize the significance of hybrid optimization in dynamic MEC environments, where decision efficiency and stability are crucial.

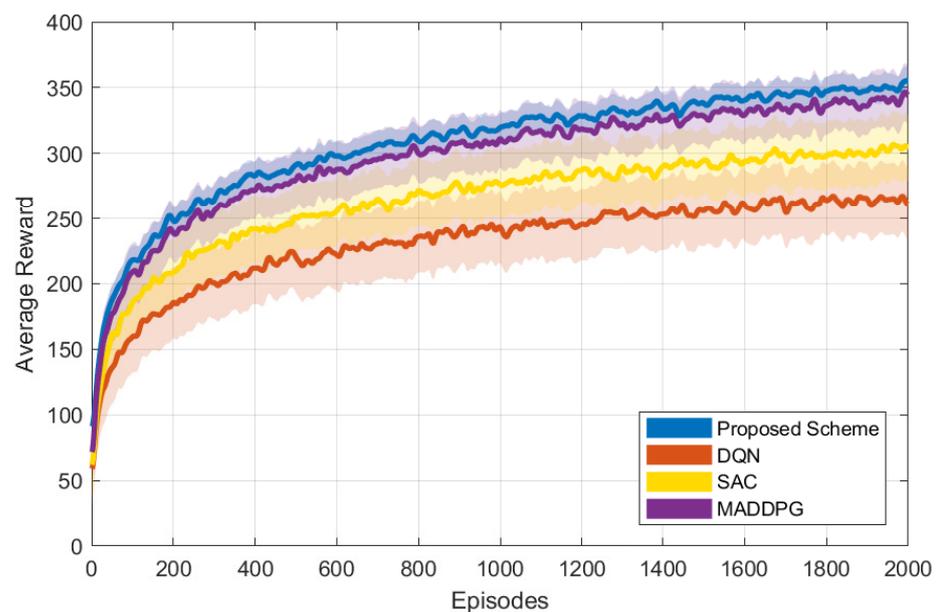


Figure 2. Average rewards vs. no. of episodes.

In Figure 3, the outperformance of the proposed method compared with the state-of-the-art schemes is demonstrated in terms of total network latency. For instance, at $F_{m,k} = 600$, the total latency with the proposed scheme and the benchmark schemes are

approximately 28.8 ms, 32.2 ms, 40.8 ms, and 49.3 ms for the proposed scheme, MADDPG, SAC, and DQN, respectively. Moreover, it is evident from the figure that as $F_{m,k}$ increases, the total network latency for all schemes grows, yet the proposed scheme consistently achieves the lowest latency. This performance gap highlights the efficiency of the proposed approach in minimizing latency. To this end, it is worth mentioning that these results are valid within the range of considered CPU cycles per bit ($F_{m,k}$) values, demonstrating the robustness of the proposed method across varying computational demands.

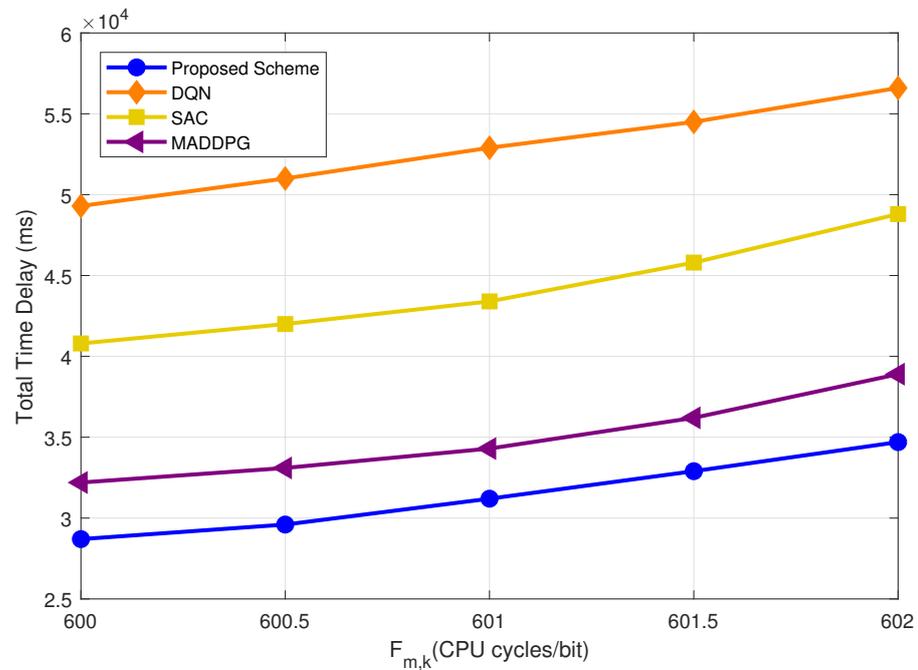


Figure 3. The total time delay according to different schemes vs. $F_{m,k}$, with $K = 100$ and $\zeta_{max} = 30$ Giga cycles/s.

In Figure 4, the total network latency is evaluated for the proposed scheme in comparison with other benchmark schemes across varying computing capacities of the MBS (ζ_{max}) measured in megacycles per second. The evaluation considers a fixed number of UEs ($K = 100$) and CPU cycles per bit ($F_{m,k} = 600$ cycles/bit). Figure 4 illustrates that increasing the computing capacity of the MBS (ζ_{max}) leads to a noticeable reduction in total network latency for all schemes. This trend underscores the critical role of computational resources at the MBS in minimizing latency. Among the evaluated schemes, the proposed method consistently outperforms the benchmark schemes (DQN, SAC, and MADDPG), achieving the lowest latency across all considered values of ζ_{max} . Moreover, the latency reduction with the proposed scheme is more pronounced compared to the other schemes, highlighting its efficiency in leveraging enhanced computational resources.

Figure 5 demonstrates the impact of varying the number of UEs on total task delay, with user counts ranging up to a maximum of 300 and RIS $N = 100$. The results reveal a clear trend: as the number of UEs increases, the total time delay rises significantly. Among the evaluated algorithms, the proposed scheme consistently achieves the lowest task delay across all scenarios, highlighting its superior ability to handle high user densities efficiently. In contrast, DQN exhibits the highest delay, reflecting its limited scalability for large-scale environments. SAC and MADDPG perform moderately, but their delays grow more sharply with increasing UEs compared to the proposed scheme. This analysis underscores the proposed scheme's robustness and scalability in managing resource allocation and task offloading in densely populated MEC systems.

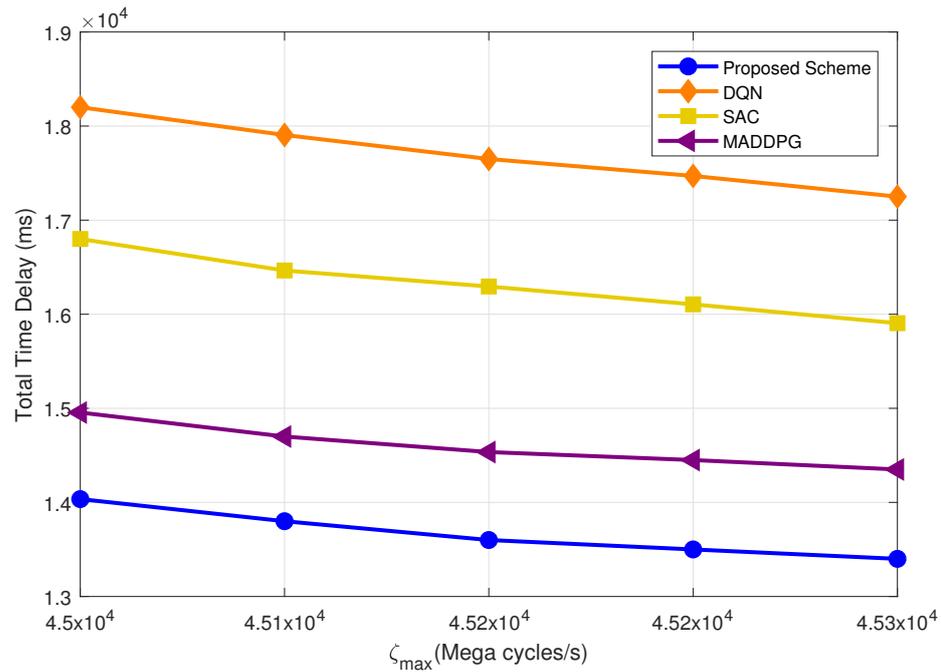


Figure 4. The total time delay according to different schemes vs. ζ_{max} , with $K = 100$ and $F_{m,k} = 600$ cycles/bit.

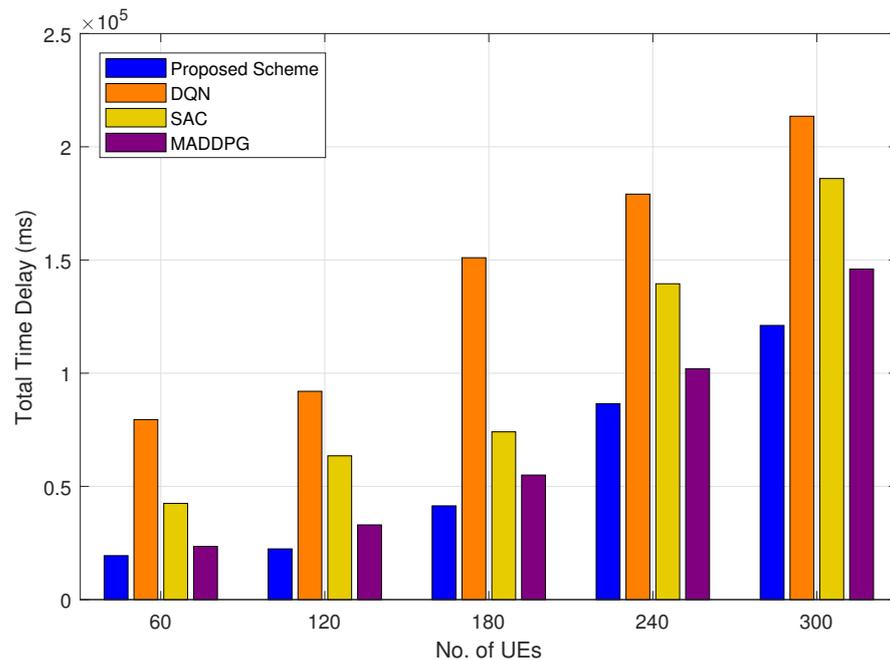


Figure 5. The total time delay vs. no. of UEs.

Figure 6 illustrates the task completion ratio as a function of the number of UEs in the system. The task completion ratio is defined as the percentage of tasks completed within the predefined delay threshold relative to the total number of tasks. As shown in Figure 6, the completion ratio decreases with an increasing number of UEs due to resource limitations within the UAV-RIS-assisted MEC framework. Specifically, the proposed scheme demonstrates the highest task completion ratio across all user densities, showcasing its robust capability to prioritize high-priority tasks and manage resource allocation effectively. In contrast, the benchmark algorithms—DQN, SAC, and MADDPG—exhibit significantly

lower completion ratios as the number of UEs grows, highlighting their limitations in handling resource scarcity and task prioritization. The proposed scheme's superior performance stems from its hybrid action space optimization and reward mechanism, prioritizing high-priority tasks. This scheme improves overall completion ratios by allocating more resources to critical tasks, ensuring effectiveness across varying UE densities.

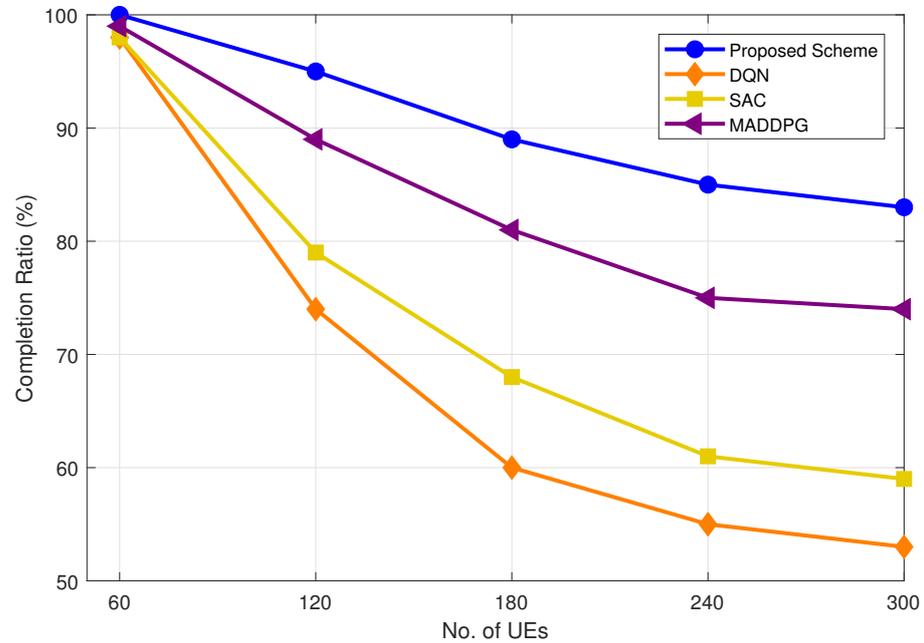


Figure 6. Task completion ratio vs. no. of UEs.

7. Conclusions

In this paper, we proposed an MEC system hosted within a massive MIMO base station, serving M groups of users with the assistance of a RIS-equipped UAV to enhance system coverage and performance. To optimize this communication scenario, we formulated the problem of minimizing total task execution latency across all UEs as an MDP and developed a novel DRL algorithm, leveraging a hybrid space representation. This approach effectively reduced latency while ensuring efficient resource utilization. Extensive simulations demonstrated the effectiveness of our approach in improving system performance. Specifically, our framework achieved a latency reduction of 10.56%, 29.41%, and 41.62% compared to MADDPG, SAC, and DQN, respectively. Additionally, it improved the task completion ratio by 12.16%, 40.68%, and 56.60% over MADDPG, SAC, and DQN, respectively. These results validate the superiority of the proposed method over state-of-the-art schemes, demonstrating its potential for next-generation communication networks. Moreover, the findings highlight the scalability of the proposed solution. As the number of UEs and RIS elements increases, our algorithm maintains a practical computational complexity, as analyzed in Section 5.2.3. Despite these promising outcomes, certain limitations remain. The model assumes idealized channel conditions, which may not fully capture real-world deployments where environmental uncertainties affect performance. In future work, we aim to develop adaptive solutions that incorporate real-time feedback to handle dynamic environments and fluctuating system states. Furthermore, the framework will be extended to consider energy efficiency and security as key performance metrics, enhancing its applicability to diverse 6G use cases.

Funding: The author extend his appreciation to the Deputyship for Research & Innovation, Ministry of Education in Saudi Arabia for funding this research work through the project number MoE-IF-UJ-R2-22-04100399-1.

Data Availability Statement: No new data were created or analyzed in this study.

Conflicts of Interest: The author declares no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

| | |
|------|---------------------------------------|
| DRL | deep reinforcement learning |
| IoT | Internet of Things |
| ITU | International Telecommunication Union |
| LoS | line-of-sight |
| MEC | multi-access edge computing |
| MDP | Markov decision process |
| MBS | MIMO base station |
| QoS | quality of service |
| RISs | reconfigurable intelligent surfaces |
| TDMA | time division multiple access |
| UAVs | unmanned aerial vehicles |
| UE | user equipment |

References

- Measuring Digital Development—Facts and Figures 2023. Available online: https://www.itu.int/hub/publication/d-ind-ict_mdd-2023-1/ (accessed on 20 January 2025).
- Ahmed, M.; Raza, S.; Soofi, A.A.; Khan, F.; Khan, W.U.; Xu, F.; Chatzinotas, S.; Dobre, O.A.; Han, Z. A survey on reconfigurable intelligent surfaces assisted multi-access edge computing networks: State of the art and future challenges. *Comput. Sci. Rev.* **2024**, *54*, 100668. [[CrossRef](#)]
- Tariq, M.N.; Wang, J.; Raza, S.; Siraj, M.; Altamimi, M.; Memon, S. Towards Optimal Resource Allocation: A Multi-agent DRL based Task Offloading Approach in Multi-UAV-Assisted MEC Networks. *IEEE Access* **2024**, *12*, 81428–81440. [[CrossRef](#)]
- Ahmed, M.; Raza, S.; Soofi, A.A.; Khan, F.; Khan, W.U.; Abideen, S.Z.U.; Xu, F.; Han, Z. Active reconfigurable intelligent surfaces: Expanding the frontiers of wireless communication—A survey. *IEEE Commun. Surv. Tutor.* **2024**. [[CrossRef](#)]
- Mehrabian, A.; Wong, V.W. Joint Spectrum, Precoding, and Phase Shifts Design for RIS-Aided Multiuser MIMO THz Systems. *IEEE Trans. Commun.* **2024**, *72*, 5087–5101. [[CrossRef](#)]
- Xu, F.; Ahmad, S.; Ahmed, M.; Raza, S.; Khan, F.; Ma, Y.; Khan, W.U. Beyond encryption: Exploring the potential of physical layer security in UAV networks. *J. King Saud Univ.-Comput. Inf. Sci.* **2023**, *35*, 101717. [[CrossRef](#)]
- Banafaa, M.K.; Pepeoğlu, Ö.; Shaye, I.; Alhammadi, A.; Shamsan, Z.A.; Razaz, M.A.; Alsagabi, M.; Al-Sowayan, S. A comprehensive survey on 5G-and-beyond networks with UAVs: Applications, emerging technologies, regulatory aspects, research trends and challenges. *IEEE Access* **2024**, *12*, 7786–7826. [[CrossRef](#)]
- Xu, Y.; Zhang, T.; Zou, Y.; Liu, Y. Reconfigurable intelligence surface aided UAV-MEC systems with NOMA. *IEEE Commun. Lett.* **2022**, *26*, 2121–2125. [[CrossRef](#)]
- Hu, H.; Sheng, Z.; Nasir, A.A.; Yu, H.; Fang, Y. Computation Capacity Maximization for UAV and RIS Cooperative MEC System with NOMA. *IEEE Commun. Lett.* **2024**, *28*, 592–596. [[CrossRef](#)]
- Wang, Y.; Niu, J.; Chen, G.; Zhou, X.; Li, Y.; Liu, S. RIS-Aided Latency-Efficient MEC HetNet with Wireless Backhaul. *IEEE Trans. Veh. Technol.* **2024**, *73*, 8705–8719. [[CrossRef](#)]
- Ning, Z.; Dong, P.; Wen, M.; Wang, X.; Guo, L.; Kwok, R.Y.; Poor, H.V. 5G-enabled UAV-to-community offloading: Joint trajectory design and task scheduling. *IEEE J. Sel. Areas Commun.* **2021**, *39*, 3306–3320. [[CrossRef](#)]
- Song, F.; Xing, H.; Wang, X.; Luo, S.; Dai, P.; Xiao, Z.; Zhao, B. Evolutionary multi-objective reinforcement learning based trajectory control and task offloading in UAV-assisted mobile edge computing. *IEEE Trans. Mob. Comput.* **2022**, *22*, 7387–7405. [[CrossRef](#)]
- Xu, B.; Kuang, Z.; Gao, J.; Zhao, L.; Wu, C. Joint offloading decision and trajectory design for UAV-enabled edge computing with task dependency. *IEEE Trans. Wirel. Commun.* **2022**, *22*, 5043–5055. [[CrossRef](#)]

14. Goudarzi, S.; Soleymani, S.A.; Wang, W.; Xiao, P. Uav-enabled mobile edge computing for resource allocation using cooperative evolutionary computation. *IEEE Trans. Aerosp. Electron. Syst.* **2023**, *59*, 5134–5147. [[CrossRef](#)]
15. Bai, Z.; Lin, Y.; Cao, Y.; Wang, W. Delay-Aware Cooperative Task Offloading for Multi-UAV Enabled Edge-Cloud Computing. *IEEE Trans. Mob. Comput.* **2024**, *23*, 1034–1049. [[CrossRef](#)]
16. Zhao, N.; Ye, Z.; Pei, Y.; Liang, Y.C.; Niyato, D. Multi-Agent Deep Reinforcement Learning for Task Offloading in UAV-Assisted Mobile Edge Computing. *IEEE Trans. Wirel. Commun.* **2022**, *21*, 6949–6960. [[CrossRef](#)]
17. Zhai, Z.; Dai, X.; Duo, B.; Wang, X.; Yuan, X. Energy-efficient UAV-mounted RIS assisted mobile edge computing. *IEEE Wirel. Commun. Lett.* **2022**, *11*, 2507–2511. [[CrossRef](#)]
18. Di Renzo, M.; Zappone, A.; Debbah, M.; Alouini, M.S.; Yuen, C.; De Rosny, J.; Tretyakov, S. Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and the road ahead. *IEEE J. Sel. Areas Commun.* **2020**, *38*, 2450–2525. [[CrossRef](#)]
19. Al-Hourani, A.; Kandeepan, S.; Lardner, S. Optimal LAP altitude for maximum coverage. *IEEE Wirel. Commun. Lett.* **2014**, *3*, 569–572. [[CrossRef](#)]
20. Wu, Q.; Zhang, R. Beamforming optimization for wireless network aided by intelligent reflecting surface with discrete phase shifts. *IEEE Trans. Commun.* **2019**, *68*, 1838–1851. [[CrossRef](#)]
21. Raza, S.; Liu, W.; Ahmed, M.; Anwar, M.R.; Mirza, M.A.; Sun, Q.; Wang, S. An efficient task offloading scheme in vehicular edge computing. *J. Cloud Comput.* **2020**, *9*, 1–14. [[CrossRef](#)]
22. Raza, S.; Wang, S.; Ahmed, M.; Anwar, M.R.; Mirza, M.A.; Khan, W.U. Task offloading and resource allocation for IoV using 5G NR-V2X communication. *IEEE Internet Things J.* **2021**, *9*, 10397–10410. [[CrossRef](#)]
23. Raza, S.; Ahmed, M.; Ahmad, H.; Mirza, M.A.; Habib, M.A.; Wang, S. Task offloading in mmwave based 5g vehicular cloud computing. *J. Ambient Intell. Humaniz. Comput.* **2023**, *14*, 12595–12607. [[CrossRef](#)]
24. Ahmed, M.; Raza, S.; Ahmad, H.; Khan, W.U.; Xu, F.; Rabie, K. Deep reinforcement learning approach for multi-hop task offloading in vehicular edge computing. *Eng. Sci. Technol. Int. J.* **2024**, *59*, 101854. [[CrossRef](#)]
25. Nguyen, L.D.; Tuan, H.D.; Duong, T.Q.; Dobre, O.A.; Poor, H.V. Downlink beamforming for energy-efficient heterogeneous networks with massive MIMO and small cells. *IEEE Trans. Wirel. Commun.* **2018**, *17*, 3386–3400. [[CrossRef](#)]
26. Liu, T.; Tang, L.; Wang, W.; Chen, Q.; Zeng, X. Digital-twin-assisted task offloading based on edge collaboration in the digital twin edge network. *IEEE Internet Things J.* **2021**, *9*, 1427–1444. [[CrossRef](#)]
27. Khan, I.; Raza, S.; Rehman, W.u.; Khan, R.; Nahida, K.; Tao, X. A Deep Learning-Based Algorithm for Energy and Performance Optimization of Computational Offloading in Mobile Edge Computing. *Wirel. Commun. Mob. Comput.* **2023**, *2023*, 1357343. [[CrossRef](#)]
28. Heidarpour, A.R.; Heidarpour, M.R.; Ardakani, M.; Tellambura, C.; Uysal, M. Soft actor–critic-based computation offloading in multiuser MEC-enabled IoT—A lifetime maximization perspective. *IEEE Internet Things J.* **2023**, *10*, 17571–17584. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.