Available online at www.sciencedirect.com
ScienceDirect
Computer Networks 00 (2023) 1–22
Computer Networks

Deep Learning-based Energy Harvesting with Intelligent Deployment of RIS-assisted UAV-CFmMIMOs

Alvi Ataur Khalil[a], Mohamed Y. Selim[b], Mohammad Ashiqur Rahman[a]

[a]*Florida International University, USA.*
[b]*Iowa State University, USA.*

Abstract

The ever-evolving internet of things (IoT) has spawned hundreds of wireless sensors that communicate via the internet infrastructure. The lifetime and self-sustainability of these sensors are pivotal factors dictating the performance of respective application infrastructure. Radio frequency energy harvesting (RFEH) technology has exhibited the capability of effectively augmenting the battery lifetime of these sensors. In this work, we introduce a novel framework called CURE, which combines the advantages of cell-free massive multiple-input multiple-output (CFmMIMO) and reconfigurable intelligent surfaces (RISs) to provide uninterrupted energy harvesting for IoT devices through RFEH. CFmMIMO integrates the advantages of distributed systems and massive MIMO, while RIS improves the signal strength of the information transfer and RFEH via its passive reflection capabilities. Moreover, we consider unmanned aerial vehicles (UAVs) equipped with CFmMIMO as mobile access points (APs) to better serve the moving sensory devices. To further enhance RFEH, we propose DENCE, a channel estimation technique based on deep learning (DL) that eliminates the need for traditional closed-form equation-based channel estimation methods. Through evaluation, we first validate the performance of CURE by comparing it with the modified bisection search for max-min fairness (MBS-MMF) algorithm and later corroborate that DENCE significantly improves the performances of both models. Finally, to optimize the UAV deployment and ensure continuous RFEH coverage, we propose dARL, a deep reinforcement learning (DRL)-based scheduling framework that enables UAV-CFmMIMO swarms to perform continuous energy harvesting in the coverage area collaboratively.

Keywords: Energy harvesting, deep learning, reinforcement Learning, unmanned aerial vehicles, reconfigurable intelligent surfaces, cell-free massive MIMO.

1. Introduction

Energy harvesting through means of channel estimation plays a crucial role in the system performance of wireless networks. Currently, an estimation of more than 50 billion internet of things (IoT) devices are connected to the internet [1, 2], including diverse applications such as healthcare, agricultural, and logistics, etc [3, 4, 5, 6]. Most IoT devices connect to other IoT devices or the network wirelessly [7]. Moreover, most of these IoT devices are transportable or are installed and left in remote areas. Hence, recharging the battery of those IoT nodes is crucial for guaranteeing the continuous operations of these devices [8]. Although replacing the IoT nodes' batteries is an option, this solution becomes infeasible for remotely placed nodes. Moreover, the replacement mechanism will cause even higher expenses owing to transportation and associated labor cost. To resolve these issues, a lot of efforts were exerted to find different ways of delivering power to IoT devices by wireless means. The visionary Nikola-Tesla proposed to transmit energy to free space and convert this energy into direct current [9]. That provision led to the techniques of acquiring energy from new sources, such as wireless power transfer (WPT) and energy harvesting (EH).

Harvesting energy from radio frequency (RF) is a reliable option for charging IoT nodes in indoors, outdoor, stationary, or mobile situations [10]. However, radio frequency energy harvesting (RFEH) conversion efficiency is

low when the RF source is far from the node. Cell-free network systems with a central processing unit (CPU) and front-haul connection can help cope with the distant problem by deploying access points (APs). Some energy harvesting mechanisms aim to maximize total RFEH while selecting APs under transmission power constraints (e.g.,[11]); however, they don't consider uplink communications or imperfect channel state information (CSI). Others use a linear energy harvesting model to minimize transmitted energy for wirelessly-powered IoT devices (e.g., [12]).

Table 1: Abbreviations of Models.

Abbreviations	*Full form*
CURE	(**C**)FmMIMO-mount (**U**)AV assisted by (**R**)IS framework for RF (**E**)nergy harvesting
DENCE	(**De**)ep (**N**)eural network-based (**C**)hannel (**E**)stimation architecture
dARL	(**D**)eployment planning for UAV (**A**)Ps using (**RL**)

Since the massive multiple-input multiple-output (mMIMO) technique can direct signal power with narrow beams [13, 14], it has been considered in both wireless information and energy transfer [15, 16]. However, it suffers from inter-cell interference, which reduces energy harvesting efficiency for IoT nodes. To address this, a cell-free mMIMO (CFmMIMO) architecture has been proposed [17, 18], where many distributed single-antenna access points (APs) serve a small number of users. Leveraging this infrastructure, we have proposed CURE, a framework for RF energy harvesting (in our published work [19]) that utilizes CFmMIMO-mount Unmanned aerial vehicle (UAV) assisted by Reconfigurable intelligent surfaces (RIS). Mounting an AP on UAV achieves much better energy harvesting efficiency due to its mobility and reachability [20]. The UAVs used in our framework, equipped with CFmMIMO, provide better signal strength at the edge of the cell. Furthermore, the distributed nature of CFmMIMO allows for a light-weighted UAV (carrying only a few antennas) that does not affect flying time. Additionally, RISs are utilized to reflect AP signals (using controllable intelligent elements) to desired IoT devices and provide an alternative line-of-sight link with coherent interference. There are a lot of recent advancements toward UAV and RIS fusion. For instance, Diamanti et al. propose an energy-efficient multi-user communication system using the aforementioned fusion [21]. The work aims to minimize the system's energy consumption by optimizing the UAV's location and the RIS's phase shift. However, the paper does not address the issue of inter-cell interference, which limits the energy harvesting efficiency of the IoT nodes. You et al. improve the communication performance of an integrated air-ground wireless network using RIS and UAVs [22]. The paper proposes a novel approach that leverages UAVs' mobility to enhance the performance of the IRS-assisted communication system. Mei et al. propose optimizing the 3D-trajectory and phase-shift design for RIS-assisted UAV systems using deep reinforcement learning [23]. However, none of these works provide any solutions to improve the energy harvesting process, which could be a significant challenge in the practical implementations of the proposed communication models. Our framework addresses the inter-cell interference issue by leveraging CFmMIMO, while proposing a communication model for energy harvesting at the IoT devices. We evaluate the performance of CURE by comparing it with the modified bisection search for max-min fairness (MBS-MMF) algorithm proposed by Demir et al. [24] since we leverage the CFmMIMO-based system model proposed by them.

Channel estimation is vital to the energy harvesting system design [25]. Among the well-known channel estimation methods, least squares (LS) estimation is known for its low computational complexity method [26, 27]. However, in many actual applications, especially for multi-path channels, LS estimation produces a rather high channel estimation errors. By minimizing channel estimate errors on average, minimum mean square error (MMSE) estimation, as an alternative to LS estimation, yields substantially better channel estimation quality [28]. The closed-form equations of the channel estimates derived by MMSE are based on assumptions such as the propagation channels being described by a linear system and each channel response following a circularly symmetric complex Gaussian distribution [26]. MMSE estimation, however, has a high computational complexity since CSI is required, which is either exceedingly difficult to get or fluctuates rapidly in a short coherence time, making it difficult to implement [29, 30]. Deep learning (DL) has been shown to enhance communication reliability and reduce the computational complexity of 5G and beyond networks significantly [31]. Consequently, in this work, we propose DENCE, a Deep Neural network (DNN)-based Channel Estimation architecture for estimating the channel between the UEs and AP antennas as well as RIS elements. This channel estimation scheme outperforms classic CSI processing approaches like linear minimum mean square error (LMMSE) estimation, resulting in a *Beyond 5G* networking paradigm in which machine learning drives networking optimization. We validate DENCE's efficiency by contrasting the performance of the MBS-MMF algorithm with "LMMSE and LS closed-form based channel estimation" versus "DENCE-based channel estimation". Moreover, we quantify the performance augmentation (with respect to spectral efficiency) of proposed CURE with

DeNCE-based channel estimation.

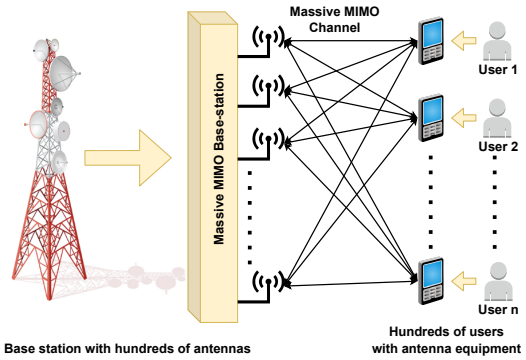
Another key aspect of the UAV-mounted CFmMIMOs (UAV-CFmMIMOs)-based energy harvesting is the deployment scheduling of the APs. This is crucial as the optimal deployment of UAV access points can ensure continuous coverage without interruptions for UE energy harvesting. Since there is a limitation in the battery capacity of the UAVs, efficient scheduling becomes more significant for UAV-CFmMIMOs, where the mounted CFmMIMO brings in some additional burden. To this end, we propose dARL, a Deployment planning for UAV Access points using Reinforcement Learning (RL), which dynamically routes the UAVs according to the UE requirements in different regions in the coverage area. The RL agents are trained using a proximal policy optimization (PPO)-based deep RL (DRL) framework to perform a collaborative energy harvesting job. Finally, we show the intelligent behavior exhibited by the agent swarm through the iterative training phases.

The contributions of this work can be summarized as follows:

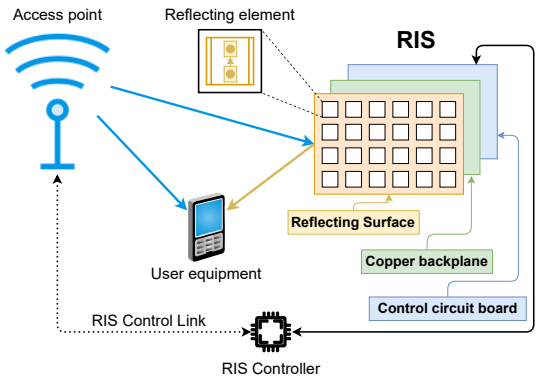
- We design and implement CURe, a novel RFEH framework that leverages the combined benefits of UAVs, CFmMIMO, and RIS. We validate the framework's performance by comparing it with the MBS-MMF algorithm with respect to spectral efficiency.
- To further augment the RFEH, we design and implement DeNCE, a DL-based channel estimation framework, eliminating the need for deterministic equations-based channel estimation.
- To optimize the UAV-cfmMIMO deployment cost and ensure continuous coverage, we design and implement dARL, a DRL-based scheduling framework that dynamically deploys UAVs in different regions.

The first contribution has been published earlier [19], while the latter two contributions are made in this extension. The full form of each of the proposed models are listed in Table 1.

We organize the rest of the manuscript as follows: We present adequate preliminary and background information in Section 2. The proposed framework is introduced in Section 3. In Section 4, we discuss in detail the proposed framework and the complete analysis of the proposed algorithms. In Section 5, we demonstrate the system setup along with the experimental analysis and findings related to our proposed framework. Recent related literature are presented in Section 6. Finally, the paper is concluded in Section 7.



(a)



(b)

Figure 1: Graphical Architecture of (a) massive Multiple-input Multiple-output (mMIMO) and (b) reconfigurable intelligent surface (RIS).

2. Background

In this section, we first discuss the necessary information that helps to explain the CURe framework. Later, we present preliminary information regarding DL-based channel estimation and RL-based UAV scheduling, which will motivate the implementation of the DeNCE and dARL frameworks.

2.1. Cell-Free massive Multiple-input Multiple-output

The mMIMO technology has drawn paramount attention over the course of the past ten years due to its high spectral efficiency (SE), achieved by the spatial multiplexing of a considerably large number of connected devices [32]. The architecture of the mMIMO system is shown in Figure 1(a). Although mMIMO technology has significant advantages, it suffers from inter-cell interference, especially at the edge of the cell [33]. In addition, the complexity of the symbol detector increases dramatically in the uplink direction of the MIMO receiver, owing to a large number of antennas and RF chain [34]. Alternatively, CFmMIMO network architecture has been introduced by Ashikhmin et al. [17], where a small number of IoT

devices will be served with a larger number of geographically distributed antennas throughout the whole coverage region. All these antennas are connected to a central processing unit (CPU) via fronthaul links. The antennas (i.e., the APs) send the IoT devices' uplink data to the CPU via the fronthaul links while the CPU sends downlink data and power coefficients to the APs. One of the significant advantages of the CFmMIMO is that it has a high SE in addition to other advantages, such as low deployment cost, high flexibility, and appealingly uniform quality of service (QoS) [35].

2.2. Reconfigurable Intelligent Surfaces

RIS is a completely new concept in the field of wireless technologies, which is attracting a lot of interest from the wireless research community. It is a controllable metasurface that allows controlling the amplitude and phase of the reflected signal from its surface [36]. Figure 1(b) shows the basic architecture of the RIS. It has several layers of planes that can be manufactured with advanced printing and lithography techniques [37]. The elements can make some changes to the incident signal without affecting the transmit power. The change can be in phase, frequency, amplitude, or even polarization [36]. A Salient feature of the RIS is its capability to control the environment via the network operator that shapes the EM response of the objects distributed throughout the network [38]. For example, when obstacles block the line-of-sight communication, the RIS can be installed on walls or ceilings indoors or on facades of the buildings outdoors [39] to assist the transmission between the transmitter and the receiver.

2.3. Unmanned Aerial Vehicle as Access Point

The UAV technology has recently received a lot of attention for its use in numerous military and commercial applications [40]. Because of the high probability of line-of-sight and flexible deployment, both fixed-wing and rotary-wing UAVs became strong candidates to assist 5G and beyond networks. UAV-backed cellular communications are specifically adapted to provide service to geographic areas of poor coverage or high-traffic hotspots. The application of UAV-mounted APs is shown in Figure 2. These APs communicate with the CPU through fronthaul links to provide cellular services in such geographic areas. However, this method has several weaknesses, including battery capacity and service quality. The limited capacity of the UAVs' batteries restricts their operational time, while the service quality will be limited by the capacity of the fronthaul link between the UAV and the terrestrial base station (TBS) [40]. An applicable solution to these two problems could be tethered UAVs located on a dedicated mobile station or on the roof of a building, where those tethered UAVs will receive power and data supply via cable from the TBS [41]. Our proposed model can utilize either untethered or tethered UAVs. Untethered UAVs can be used for time-limited missions, while tethered UAVs can be used for long-lasting missions.

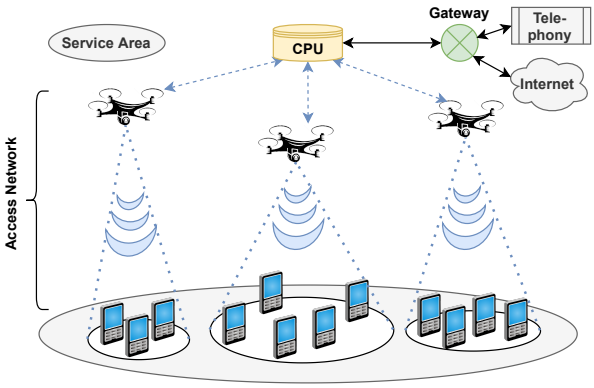


Figure 2: UAV-mounted Access Point.

2.4. Channel Estimation with DL

To deal with the distortion in received signals, due to the channel estimation, DL is becoming a popular choice, specifically under imperfect environments. Despite DL's significant success, no analytical interpretation is available to confirm the benefits or drawbacks of DL methods when applied to communications [42]. Nevertheless, a lot of recent research works are leveraging DL to supplant closed-form-based channel estimations. More research is indicating that DL approaches are particularly well suited to channel estimation, and DNNs are increasingly being used in communication systems [43, 44, 45, 46, 47, 48]. The DNN can learn the channel structure with sufficient training data and provide reasonable estimates, which outperforms the traditional channel estimation techniques.

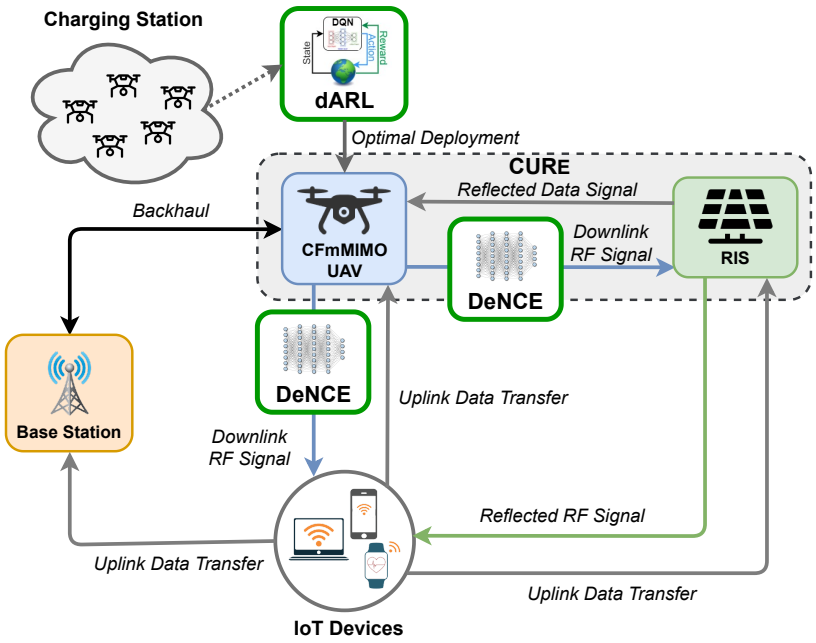


Figure 3: Proposed framework for improved energy harvesting.

2.5. Reinforcement Learning in Scheduling

With the introduction of UAVs into wireless networks, a plethora of planning tasks such as optimal deployment, flight trajectory design, and energy efficiency are posed alongside. Traditional performance measures like network coverage, completion rate, and delay have primarily been used in the literature to evaluate each of these assignments. However, such performance metrics can not always assess the UAV swarm's performance concerning their primary objective, i.e., surveying assigned areas, providing coverage of energy harvesting through designated regions, etc. RL, on the other hand, facilitates a rewards-based performance measure, to make UAV swarms intelligent enough to stride towards a particular objective collaboratively [49, 50, 51]. The UAVs must complete tasks based on real-time observation since instantaneous action is mandatory while functioning in a dynamic environment (without centralized control). RL can perform real-time learning and decision-making depending on the environment, which makes it a suitable and practical solution for scheduling the energy harvesting job for a swarm of UAV-CFmMIMOs.

3. Framework

In this part, we go over our proposed model, comprising three frameworks, namely CURE, DENCE, and dARL. Our system, as shown in Figure 3, provides downlink RF signals to battery-limited IoT devices for energy harvesting. The APs are UAVs (each equipped with a CFmMIMO, i.e., UAV-CFmMIMO), which are not linked to any direct energy source. As a result, they always need to come back to recharge themselves after they hover at "a certain location" or "multiple locations" to serve as APs. They also have a wireless connection to the core backhaul network, i.e., the base station. The UAV-CFmMIMOs are responsible for transmitting RF signals directly to IoT devices. The RISs, on the other hand, improve energy harvesting in the downlink and provide signals to non-line-of-sight locations in the coverage zones. They intelligently reflect the RF signal to the IoT devices. The combined RF signal from the access points and RISs is harvested at the user end. This power is then used to transfer information from user devices to access points via uplink. Depending on the location, the uplink signal can be transmitted either directly to the base station or through the UAV access points or the RIS panel. In summary, the proposed wireless communication utilizes UAVs that carry only one or a few antenna elements instead of a regular base station. This limited payload keeps the UAV's power consumption low, extending its flying time. The CPU is responsible for the cooperation of the individual antennas (distributed over multiple UAVs) and determines which user each set of antennas will serve. The proposed

Table 2: List of notations and definitions.

Notation	***Definition***	***Unit***
δ_c	Total number of samples per coherence interval.	N/A
δ_p	Pilot sequence length	N/A
δ_d	Downlink sequence length	N/A
δ_u	Uplink sequence length	N/A
$\mathbf{h}_{ju}$	Channel between the u_{th} AP and the j_{th} UE	GHz
$e^{k\phi_{ju}}\bar{\mathbf{h}}_{ju}$	Line-of-sight components	GHz
$\tilde{\mathbf{h}}_{ju}$	Non-line-of-sight components	GHz
γ_{ju}	Large-scale fading co-efficient	dB
ϕ_{ju}	Phase shift by the Rician fading in the line-of-sight component	radian
$\mathbf{x}_u^E$	Signal sent by the u^{th} AP	dBm
$\mathbf{w}_{ju}^*$	Downlink precoding vector for the phase of the sent signal by AP	N/A
s_j	Zero-mean unit-variance energy signal for the j_{th} IoT device	dBm
p_{ju}	Power control coefficient of the u_{th} AP corresponding to the j_{th} device	N/A
P_u^E	Average transmitted power for the u_{th} AP	dBm
ρ_d	Maximum power limit	dBm
r_j^E	Received signal for the j_{th} IoT device	dBm
n_j^E	Additive noise at the j_{th} IoT device	dBm
I_j	Average input power to the j_{th} device's energy harvesting rectifier circuit	dBm
E_j	Total harvested energy for the j_{th} IoT device in the δ_d channel	Joules
N	Number of N discrete elements in RIS	N/A
$\mathbf{h}_{ur}$	Deterministic channel from source to RIS	GHz
$\mathbf{h}_{jr}$	Channel between the destination and the RIS	GHz
Θ	RIS properties	N/A
θ_k	Fixed amplitude reflection coefficient	N/A
S_r	Received signal at the destination	dBm
x	Transmit power	dBm
y	Unit-power information signal	dBm
n	Reception noise	dBm

model differs from regular wireless models as it distributes processing instead of concentrating on the base station. However, the communication model remains the same. The model combines the benefits of cell-free techniques with the dynamic deployment of UAVs to allow for real-time dynamic scheduling of access points based on real-time UE/IoT requirements, adjusting scheduling based on real-time region density and population. This concludes the workflow of CURE [19], which is improved in this work to augment the RFEH further.

Our previous work leveraged closed-form-based deterministic methods for the channel estimation part of the energy harvesting for both the downlink RF signals of UAV-CFmMIMO and RISs. However, the method has a considerable amount of estimation error. Consequently, we propose a new channel estimation architecture aided by DL to improve the channel estimates. We introduce DENCE, a convolutional neural network (CNN)-based channel estimation method that is trained with the H matrices produced by the implementation of Demir et al. [24]. This framework effectively produces the estimated channels when given the appropriate input set. We propose utilizing this framework for the channel estimation of both the direct UEs and the UEs via RIS. The deployment planning of the UAV-CFmMIMOs is of paramount importance since they are the only source of RF signals in the proposed model, and they need to be in the vicinity of UEs to be able to serve them. Since the user movement is hardly deterministic, it would not be practical to deploy the UAV-CFmMIMOs at specific coverage locations and change their position in a fixed, routine way. Due to the stochastic mobility of UEs, fixed scheduling will never be able to facilitate them with continuous energy harvesting support. Consequently, we introduce real-time dynamic scheduling of the UAV-CFmMIMOs, where there is no fixed schedule. Rather the UAVs are deployed with respect to the real-time UE requirements. To this end, we introduce dARL, a DRL-based scheduling and deployment framework, which schedules and routes the UAV-CFmMIMOs in different regions according to the UE density and UE population of the regions. One of the prominent aspects of this deployment planning scheme is the dynamic capability that allows it to adapt the plan according to real-time region density and population. It is worth mentioning that the operation of dARL is independent of the rest of the model, as the deployment is solely dependent on the density and population of UEs in

different regions of the coverage area. We assume that efficient scheduling will (by means of continuous coverage) incur increased harvested energy. We validate our assumption through experimentation, presented in Section 5.

4. Technical Details

The technical details of our proposed framework are discussed in this section. First, we go through the system model that governs it all. Then, we explain our downlink energy harvesting approach. The notations used in the system model and the downlink energy harvesting approach are defined in Table 2. Later, we describe the DeNCE-based channel estimation for improved energy harvesting. Afterward, we discuss the dARL-based optimal deployment of RIS-assisted UAV-CFmMIMO. Finally, we provide a detailed description of the 'improved' CURe algorithm through means of DeNCE and dARL models. We conclude the section by discussing the complexity of the algorithm and probable deployment in practice.

4.1. System Model

As access points, we propose cell-free mMIMOs installed on UAVs. We assume that U number of APs are spread across the coverage area to serve J number of IoT devices with harvesting capability. Each UAV-CFmMIMO is considered to have N antennas and an error-free fronthaul link to the CPU. We utilize the implementation of Demir et al. scheme, proposed in [24]. As a result, we assume a time division duplex (TDD) operation to force channel reciprocity. δ_c refers to the total number of samples per coherence interval. Each coherent interval is divided into three phases: (i) uplink training, (ii) downlink wireless power transfer, and (iii) uplink wireless information transfer. All IoT devices broadcast pilot sequences of length δ_p to the UAV access points in the first phase to estimate the channel and create precoding vectors for efficient energy transfer and data reception. δ_d and δ_u samples are used for downlink and uplink transfers, respectively. As a result, the total samples for each coherent interval are:

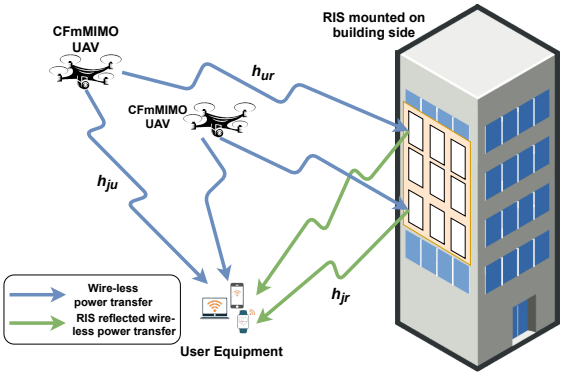


Figure 4: Proposed RIS-assisted wireless power transfer with UAV-CFmMIMOs.

$$\delta_c = \delta_p + \delta_d + \delta_u \tag{1}$$

We represent the channel between the u_{th} AP and the j_{th} user by $\mathbf{h}_{ju}$, where the channels do not change across a single time-frequency coherence interval. We follow the spatially uncorrelated rician fading channels described by Demir et al. [24] with unknown phase shifts in the context of CFmMIMO UAVs with multiple antennae. As a result, each channel's realization can be stated as follows:

$$\mathbf{h}_{ju} = e^{k\phi_{ju}}\bar{\mathbf{h}}_{ju} + \tilde{\mathbf{h}}_{ju} \tag{2}$$

Here, the line-of-sight (LOS) and non-line-of-sight (NLOS) components are denoted by $e^{k\phi_{ju}}\bar{\mathbf{h}}_{ju}$ and $\tilde{\mathbf{h}}_{ju}$ respectively. The small-scale fading for the NLOS component is modeled as $\mathcal{N}_C(\mathbf{0}_N, \gamma_{ju}\mathbf{I}_N)$, where γ_{ju} denotes the large-scale fading co-efficient. We assume that the UAV access points have perfect knowledge of the LOS component and large-scale fading coefficient corresponding to the channel between the IoT devices and themselves, explaining the long-term channel effects in accordance with previous literature. We investigate the realistic scenario where, due to user movement, phase shift ϕ_{ju} by the Rician fading in the LOS component is unknown, as opposed to previous research that considers the neglect of ϕ_{ju}. When the receiver and transmitter move over distances on the order of the wavelength, a small amount of random ϕ_{ju} is induced on both the LOS and NLOS components, which are constructed by separate routes.

4.2. Downlink EH

During this phase, all APs begin transmitting energy to IoT devices using CSI for downlink precoding. To increase the gathered energy, the coherent energy transmission will be examined first, in which the same energy symbol is transmitted by all APs for each IoT device in a synchronous way. The signal sent by the u^{th} AP can be written as follows:

$$\mathbf{x}_u^E = \sum_{j=1}^{J} \sqrt{p_{ju}} \mathbf{w}_{ju}^* s_j \tag{3}$$

Here, the downlink precoding vector for this phase is denoted by $\mathbf{w}_{ju}^*$. The zero-mean unit-variance energy signal for the j_{th} IoT device and the power control coefficient of the u_{th} AP corresponding to the j_{th} device are represented by s_j and p_{ju} respectively. In the long run, each CFmMIMO AP's transmission power should meet the following maximum power limit:

$$P_u^E \triangleq \mathrm{E}\left\{\left\|\mathbf{x}_u^E\right\|^2\right\} = \mathrm{E}\left\{\left\|\sum_{j=1}^{J} \sqrt{p_{ju}} \mathbf{w}_{ju}^* s_j\right\|^2\right\} = \sum_{j=1}^{J} p_{ju} \mathrm{E}\left\{\left\|\mathbf{w}_{ju}\right\|^2\right\} \leq \rho_d \tag{4}$$

Here, the average transmitted power for the u_{th} AP is denoted by P_u^E, and the maximum power limit is denoted by ρ_d. The received signal for the j_{th} IoT device is:

$$r_j^E = \sum_{u=1}^{U} \mathbf{h}_{ju}^T \mathbf{x}_u^E + n_j^E = \sum_{u=1}^{U} \sum_{m=1}^{J} \sqrt{p_{mu}} \mathbf{w}_{mu}^H \mathbf{h}_{ju} s_m + n_j^E \tag{5}$$

Here, the additive noise at the j_{th} IoT device is indicated by n_j^E. The average input power to the j_{th} device's energy harvesting rectifier circuit can be written as:

$$I_j = \mathrm{E}\left\{\left|\sum_{u=1}^{U} \sum_{m=1}^{J} \sqrt{p_{mu}} \mathbf{w}_{mu}^H \mathbf{h}_{ju} s_m\right|^2\right\} \tag{6}$$

We will use the following non-linear energy harvesting model similar to [52] because it correlates with real measured data. The total harvested energy for the j_{th} IoT device in the δ_d channel can be stated as:

$$E_j = \frac{\delta_d A_j I_j}{B_j I_j + C_j} \tag{7}$$

Here, $A_j > 0$, $B_j \geq 0$, and C_j are constants determined by curve fitting of the rectifier circuit of the j_{th} device [52].

We use the expressions from [53] for the RIS-supported transmission. Figure 4 presents a detailed view of the proposed RIS-assisted energy harvesting with UAV-CFmMIMOs (includes different channel symbols). The RIS consists of N discrete elements, and the deterministic channel from source to RIS is denoted by $\mathbf{h}_{ur}$ (n_{th} component is presented by $[\mathbf{h}_{ur}]_n$). As described in the previous equations, the channel between the destination and the RIS is denoted by $\mathbf{h}_{jr}$, and the deterministic flat-fading channel is designated by $\mathbf{h}_{ju}$. The following are the RIS properties:

$$\Theta = \alpha diag(e^{j\theta_1}, ..., e^{j\theta_N}) \tag{8}$$

Here, the diagonal matrix is represented by Θ, the phase-shift variables and the fixed amplitude reflection coefficient are represented by $\theta_1, ..., \theta_N$ and $\alpha \in (0,1]$ respectively. It is worth mentioning that we do not optimize the Θ, given that the locations of IoT devices, APs, and RIS are already known. The received signal at the destination can be expressed as follows:

$$S_r = (\mathbf{h}_{ju} + \mathbf{h}_{ur}^T \Theta \mathbf{h}_{jr}) \sqrt{xy} + n \tag{9}$$

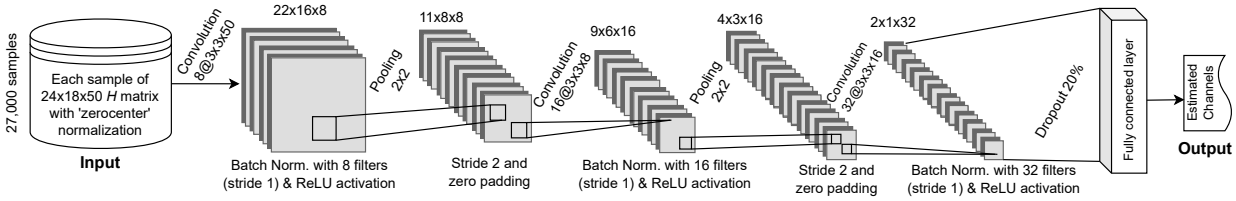


Figure 5: Proposed DeNCE network.

The transmit power, unit-power information signal, and reception noise are represented by x, y, and n, respectively. The RIS-supported network's channel capacity is represented:

$$R_{RIS}(N) = \max_{\theta_1,\ldots,\theta_N} \log_2\left(1 + \frac{x|\mathbf{h}_{ju} + \mathbf{h}_{ur}^T \Theta \mathbf{h}_{jr}|^2}{\sigma^2}\right) \tag{10}$$

$$= \log_2\left(1 + \frac{x(|\mathbf{h}_{ju}| + \alpha \sum_{n=1}^{N} |[\mathbf{h}_{ur}]_n [\mathbf{h}_{jr}]_n|)^2}{\sigma^2}\right) \tag{11}$$

The maximum rate is achieved when the phase-shifts are set as $\theta_n = arg(\mathbf{h}_{ju}) - arg([\mathbf{h}_{ur}]_n [\mathbf{h}_{jr}]_n)$. For brevity, the above equation can be re-written as follows:

$$|h_{ju}| = \sqrt{\beta_{ju}}, |h_{ur}| = \sqrt{\beta_{ur}}, |h_{jr}| = \sqrt{\beta_{jr}} \tag{12}$$

$$\frac{1}{N} \sum_{n=1}^{N} |[\mathbf{h}_{ur}]_n [\mathbf{h}_{jr}]_n| = \sqrt{\beta_{RIS}} \tag{13}$$

The re-written equation would be:

$$R_{RIS}(N) = \log_2\left(1 + \frac{x(\sqrt{\beta_{ju}} + N\alpha\sqrt{\beta_{RIS}})^2}{\sigma^2}\right) \tag{14}$$

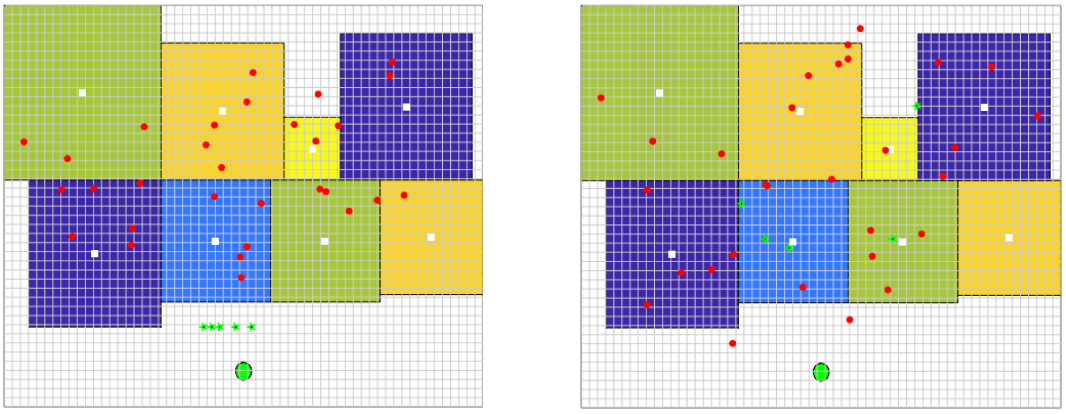
We utilize this equation for calculating the rates for our RIS-assisted energy harvesting.

4.3. DL-based Channel Estimation for Improved EH

In this subsection, we discuss the DNN considered in the proposed DeNCE framework, which estimates the channel matrices between the UAV-CFmMIMOs and UEs (both the direct case and the via RIS case). First, we generate the training data using the closed-form equations discussed in previous sections. We randomly initialize the setup with different system configurations and store the estimated channel matrices for training purposes. We generated a data set of twenty-seven thousand samples, i.e., twenty-seven thousand channel matrices, each representing the channel's information among every UE and every antenna of the APs, as well as the channel's information among every RIS element and every UE. Finally, the data set is split into the training, validation, and testing data sets. As mentioned before, we design a CNN as the DL model in this work, and its specific architecture is presented in Figure 5. We utilized CNN owing to its incredible capability of automatically detecting the important features with even unstructured untabular data, which is the inherent nature of the channel matrices. In a real-world scenario, many environmental randomnesses make it impossible to calculate the absolute channels among the UEs and antennas. Although the closed-form equations provide a good approximation of the channel matrix, it hardly takes environmental randomness into account. Hence, to introduce that aspect, we

Table 3: Modeling Parameters of the DeNCE model.

Parameter Name	Value Assigned
Initial Learn Rate	1e-3
Learn Rate Schedule	'piecewise'
Learn Rate Drop Factor	0.1
Learn Rate Drop Period	20
Shuffle	'every-epoch'
MiniBatch Size	100
Validation Frequency	30
Training Iterations	3000



(a) (b)

Figure 6: The simulated environment, designed for DRL-based UAV-CFmMIMO deployment, includes the base station (green circle with black border), DRL agents (green stars), UEs (red circles), different square regions (colored squares), and centroids of each region (small white cells). The grids show the possible different states an agent can be present in, and the small white square cells present the center of each region. The sub-figures present (a) the 10th step and (b) the 60th step in a particular training iteration.

include an environmental noise to the input data during the training phase. This inclusion of random noise resulted in a better approximation of the channel matrices from the perspective of harvested energy. Choosing a dropout layer just before the final fully connected layer was also crucial. We chose to have a 20% dropout, which ensured the training did not begin to memorize the pattern for channel estimation. The values of important hyper-parameters of the DeNCE framework are provided in Table 3.

4.3.1. Network Architecture of DeNCE

As mentioned above, we considered twenty-seven thousand different setups for data generation. The first convolution layer is equipped with 3×3-sized kernels (filters) to detect low-level features of the H matrix. Setting the stride to 1 ensures the kernel does not miss any shifting feature that can trigger the kernel, i.e., detecting the presence of a specific feature of the H matrix. Unlike image data kernels, where well-known kernel patterns detect pre-defined low-level features, in the case of the H matrix, the CNN model initializes the kernel weights randomly at the beginning of the training phase. Through the gradient descent of the RMSprop optimizer, the kernel weights are updated to minimize the loss. Since we utilize eight kernels (of size 3×3) in the first layer, the sample depth is increased. In order to downsample the spatial dimensions of the processed sample while retaining the most important features, we add an average pooling of size 2×2 and stride 2. We utilize average pooling instead of max pooling since the detection of a feature's presence is more important than the exact trigger of the feature. Average pooling can help to downsample the feature map and extract a more general representation of the input sample. Another convolution layer follows the pooling with a 3×3 kernel where the number of filters is increased to 16 to allow the network to learn more complex and abstract features of the H matrix. After having another average pooling of size 2×2 (stride 2) for a similar intention as the previous, the final convolution layer is stacked to detect the higher-level features that can help the fully connected layer to produce the activation of the estimated channel. The final convolution has 32 filters to let all the features be mapped into the input of the fully connected layer, which has 64 ($2 \times 1 \times 32$) processing elements. All three convolution layers utilize ReLU activation to make sure the "vanishing gradient" problem does not happen. We provided the detailed structure of the CNN for regeneration purposes for the other researchers since even missing only one convolution layer will drastically reduce the performance of the network. Finally, the output layer produced the estimated channel matrix according to the dimension of the input configuration.

4.4. DRL-based Optimal Deployment of UAV-CFmMIMO

In this part, we introduce the proposed DRL-based deployment of the UAV-CFmMIMOs in the coverage area. The conventional UAV scheduling schemes generally assume that the locations of the goal points do not vary, meaning the optimal trajectory predicted at the beginning of the trip remains constant. However, in real-world communication applications, the network users (to be served by UAV APs) move along the coverage area, and the movement is not

deterministic. Hence, the final trajectory of the serving UAVs can not be predicted at the beginning of the trip, rather it has to be improvised depending on the current circumstances. Consequently, we develop dARL for deploying the UAV-CFmMIMOs in different regions of the coverage area to maximize the total energy harvesting. When developing a DRL-based model, the initial task is to design a set of key elements: the **environment** (where the agents will interact and learn), the **states** (that represent the observation of an agent), the **action** (that the agents can perform), and the **reward function** (that drives an agent to act intelligently). The formulation of each key element is as follows:

Environment: The environment, in our case, is the coverage area where the UAV-CFmMIMOs provide the energy harvesting facility. In the simulated environment, we abstract the intended regions in the coverage area into different-sized square regions (presented by different colored squares in Figure 6). We assume the UAV-CFmMIMO agents (green stars) are placed at the charging station (presented by a green circle with a black border placed around the bottom-middle) at the beginning of the DRL episode. The charging station is outside the coverage area, so the UAV-CFmMIMO agents have to travel inside the area. The entry point is not fixed, which facilitates the UAVs to choose any path that leads to the region it plans to serve at the beginning. The UEs (small red circles) are placed at random locations at the beginning of the episode and move around the regions throughout the episode's time steps. They can even move out of the coverage region, making the UAV-CFmMIMO agents ignore their presence.

Table 4: Modeling Parameters of the dARL model.

Parameter Name	Value Assigned
Reinforcement learning algorithm	PPO
Experience Horizon	128
Clip Factor	0.2
Entropy Loss Weight	0.01
MiniBatchSize	64
Advantage Estimate Method	gae
GAE Factor	0.95
Initial Discount Factor	0.995
Steps per episode	1000
Learning Rate	1e-4
Gradient Threshold	1

States: The state, in our case, includes the observation of three components: the current locations of different UEs that are present within the intended coverage regions, the current locations of each of the UAV-CFmMIMO agents, and the areas of each intended region.

Actions: The UAV-CFmMIMO agents can take nine different actions depending on the current observation. The regions are divided into small grid cells, and the UAVs can move in all eight directions (forward, backward, left, right, and four diagonal directions) to proceed from one grid cell to another. As the ninth action, the agent can decide not to move (i.e., hovering at the same location).

Rewards: The PPO-based DRL is designed in a way to make the UAVs work collaboratively. That means the agents will not compete with one another to increase their personal rewards, rather they will work together to maximize the collaborative reward. The reward function is designed in a way to consider two aspects of the UEs; the first one is the density of the UEs in each of the regions, and the second one is the population of the UEs in different regions. Considering the density will not suffice as there might be some regions with very few UEs; however, the density might become very high due to the even smaller area of the region. Hence, we also consider the area's population and weigh these two factors in the reward function.

A swarm of UAV-CFmMIMO agents starts at the base station at the beginning of each training episode and, depending on the initial observation of the UEs, the agents start to move towards different regions as per the collaborative scheduling plan provided by the dARL framework. Figure 6(a) presents the 10th step of a random episode, where the agents can be seen to move from the base station towards the coverage regions. The goal of the swarm is to provide continuous coverage of energy harvesting facilities to the largest portion of the UEs. The agents can move to the centroid of each region and harvest energy for the UEs within. Depending on the real-time movement of the UEs, the agents dynamically choose to hover in the same location or move to the centroid of a different region. Figure 6(b) presents the 60th step of the same episode, where the agents can be seen to spread around and cover a larger area to serve more UEs. The values of important hyper-parameters of the dARL framework are provided in Table 4.

4.5. Implementation of the EH Model

Finally, in this section, we provide a detailed description of Algorithm 1, the energy harvesting algorithm that we considered in this work. To begin, we set up the essential system parameters ($Param_{sys}$), such as realization, transmit power ($Pow_{transmit}$), block coherence ($Block_{coher}$), and so on, based on the needs of various experiments. We must keep the values of some of the parameters constant across all of the experiments, as shown in Table 5. The client IoT devices in the coverage area are placed at random, which are represented by each of the setups. We take into

Algorithm 1: Improved energy harvesting of CURE with DENCE & dARL

$Param_{sys}$ = [$Realizations, Pow_{transmit}, Pow_{noise}$, $Block_{coher}, Param_{rectifier}, Carrier$];
$Param_{sys} \leftarrow$ Initialize;
$Matrix_{rate} \leftarrow \phi$;
$AllSetups \leftarrow$ Number of setups;
for *each setup* $\in AllSetups$ **do**
  $rates \leftarrow \phi$;
  $Param_{UEs} \leftarrow random(Count_{UEs})$;
  for *each timeStep* $\in EpiosdeSteps$ **do**
    $DeploymentPlan_{AP} \leftarrow$ **dARL**($Count_{AP}, Param_{sys}, Param_{UEs}$);
    $I_{max} \leftarrow EpiosdeSteps$;
    while I_{max} **do**
      $Gain_{Channel}, Realization_{Channel} \leftarrow$ **SetupFunc**($Param_{AP}, Param_{RIS}$);
      $H \leftarrow$ **DENCE**($DeploymentPlan_{AP}$ $Param_{AP}, Param_{UEs}, Param_{RIS}$,);
      $StatTerms, HarvestedEnergy \leftarrow$ **CalculateStats**($Gain_{Channel}, H$, $Realization_{Channel}$, $Param_{sys}$);
      $rates \leftarrow$ **SpectralEfficiency**($StatTerms$, $HarvestedEnergy, \beta$);
      $Solution \leftarrow$ **Feasibility**($rates$);
      if *Solution is Feasible* **then**
        $Matrix_{rate}$.append($rates$);
        $Param_{UEs} \leftarrow current(Count_{UEs})$;
        break;
      else
        $I_{max} \leftarrow I_{max}$ -1;
        continue;

account a variety of setups to ensure that the performance of a particular placement is not biased. We keep a matrix of rates across the setups for saving and later averaging the energy harvesting model's attained rates. We feed the dARL framework with the UEs' information (random at the first time-step), along with the UAV-CFmMIMOs information and system parameters. It produces an immediate deployment plan for the agents for that time-step; however, it is updated as UEs move into different positions in the next time-step. For each *timeStep* of each *setup* a feasible solution, given the calculated rates, is sought until I_{max} iterations. The $SetupFunc$ function is used to calculate the channel gain and realization for each setup, with parameters for the APs ($Param_{AP}$) and RISs ($Param_{RIS}$) being given as arguments.

Then we use the DENCE framework to estimate the channel matrices. The framework is fed with the current UAV-CFmMIMO deployment plan along with the parameters of UAV-CFmMIMOs, RISs, and UEs. The $CalculateStats$ function is then used to calculate the statistical terms and the amount of harvested energy. The system parameters, channel gain, and channel realization are passed as arguments here. The *SpectralEfficiency* function calculates the final rates based on the value of the large-scale fading co-efficient (β), statistical factors, and the amount of captured energy. The rate solutions are then checked for feasibility, and the iteration for that particular configuration is ended when feasible solutions are found. For the following time-step, we take note of the UEs' current locations. As stated earlier, we store the rates of various setups and calculate the model's average harvested energy.

Table 5: Modeling Parameters of the CURE Framework.

Parameter Name	Value Assigned
Pilot transmit power (W)	10^(-7)
Pilot Length	5
Total power limit per AP (W)	10/U
Length of coherence block	200
Compute noise power (dBm)	-96
Carrier frequency (GHz)	3.4
Number of downlink samples	25
Standard deviation of shadow fading for LOS (dB)	3
Standard deviation of shadow fading for NLOS (dB)	4

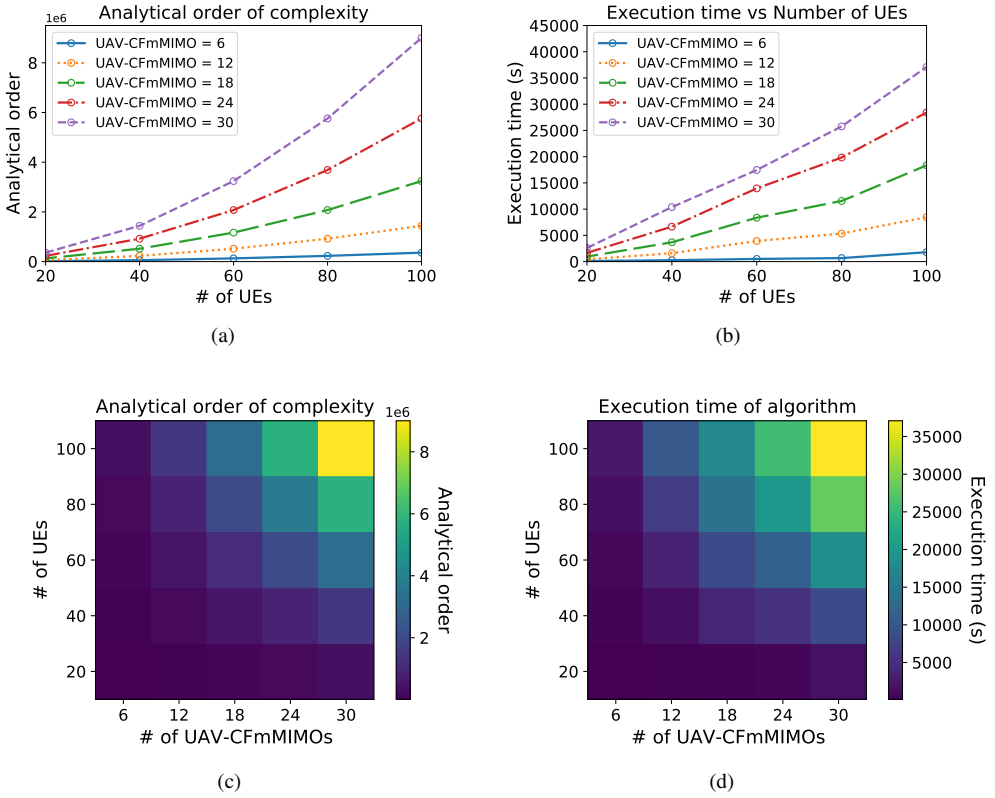


Figure 7: Complexity analysis of the algorithm: (a) Analytical order of complexity and (b) Real execution time, w.r.t a different number of UEs and UAV-CFmMIMOs. Comparison of the theoretical analysis and numerical evaluation of complexity through (c) Heatmap of analytical order of complexity and (d) Heatmap of real execution time, w.r.t a different number of UEs and UAV-CFmMIMOs.

4.6. Complexity of the Proposed Algorithm

The total time complexity of Algorithm 1 will take into account the time required for DeNCE model execution, the time required for dARL-based schedule optimization, and the time required for feasible rates obtainment. Let the time required for CNN model execution be T_C, the time for schedule optimization be T_S, and the time to obtain feasible rates is T_R. Then the overall time complexity of the framework T will be $O(T_C \times T_S \times T_R)$. Since the DeNCE model needs to calculate the rates for all the channels between each UE and antenna pair, its execution time depends on the total number of antennas (distributed over the set of UAV APs) and the total number of UEs. Accordingly, T_C can be written as $J \times N$, where J is the number of UEs and N is the total number of antennas carried by U number of UAV-CFmMIMOs (i.e., if ϵ is number of antennas per UAV-CFmMIMO, then $N = \epsilon \times U$). Next, the dARL model schedules according to the distribution of UEs over different regions and available UAV-CFmMIMOs (assuming region areas in the real world do not change, hence do not impact the computation time). Consequently, T_S can be written as $J \times U$. Finally, the algorithm will require constant time to obtain feasible rates, regardless of the input size (as the parameters $AllSetups$, $EpisodeSteps$, and I_{max} are fixed finite values, these can be considered constants and duly ignored because they do not affect the overall behavior of the algorithm). Thus, T_R can be written as 1. Accordingly, we get the total time complexity, $T = O((J \times N) \times (J \times U) \times 1) = O(J^2 \times U^2)$, since ϵ is constant (in the real world, the number of antennas equipped to each UAV-CFmMIMO is fixed). The time complexity of $O(J^2 \times U^2)$ can be simplified as $O(n^2)$, where n is the maximum of J and U. It is important to note that the upper bound of $O(J^2 \times U^2)$ may not always reflect the actual time complexity of the algorithm, as it only provides an upper limit.

To validate the theoretical analysis above, we experimented with different numbers of UEs and UAV-CFmMIMOs, as presented in Figure 7. We used 11th Gen Intel(R) Core (TM) i7-1195G7 @2.90GHz with 16 GB memory to perform the execution of the algorithm. The real execution time observed from the framework for a variable number of UEs and UAV-CFmMIMOs is presented in Figure 7(b). The resultant curves are well within the worst-case scenario, i.e.,

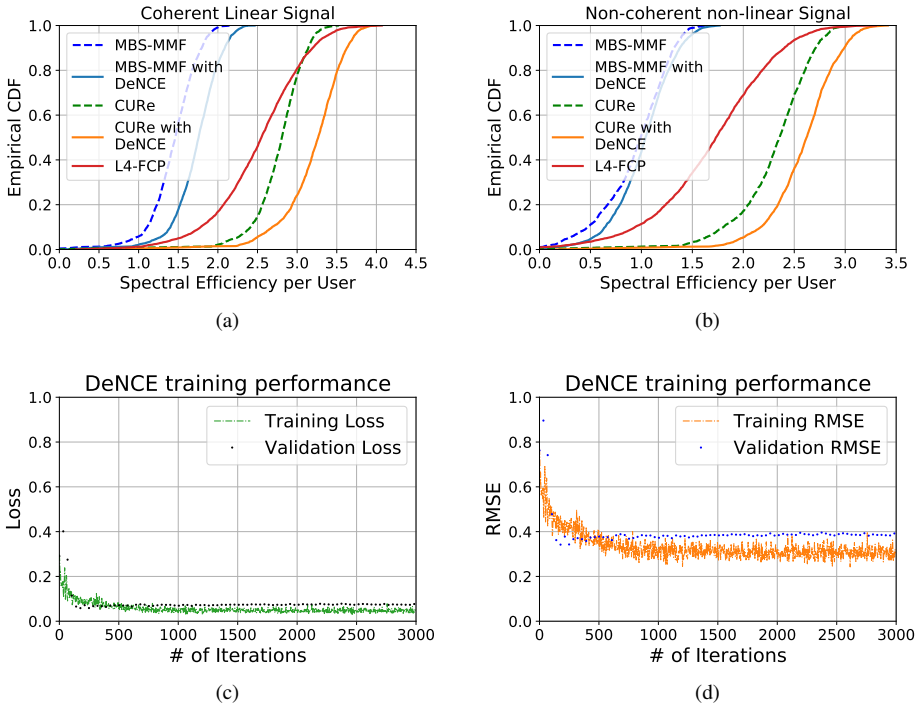


Figure 8: Empirical cumulative distributive function of spectral efficiency per user for (a) coherent linear signal, (b) non-coherent non-linear signal, for 100 randomly deployed user equipment setups. (c) The loss calculation for the training and validation (after every 30 iterations) data through the 3000 training iterations of DeNCE, (d)The RMSE calculation for the training and validation (after every 30 iterations) data through the 3000 training iterations of DeNCE.

$O(J^2 \times U^2)$ as depicted in Figure 7(a). Since the metrics are different, we provide the heatmaps of the analytical order (Figure 7(c)) and real execution time (Figure 7(d)) for further encapsulating the comparison. The heatmaps validate the soundness of the theoretical analysis with respect to the real execution time of Algorithm 1.

4.7. Probable Deployment Details in Practice

This section will discuss the probable real-world deployment details of the proposed models.

i **DeNCE:** The model will be trained offline, with the data generated through closed-form-based equations. We assume the training will occur at the base station, in any computer system with standard configuration. The trained model will be loaded into the microprocessor unit equipped in the UAV with minimal testing and validation error. Later, the UAVs will utilize the model to perform channel estimation for transmitting downlink RF signals to UEs.

ii **dARL:** Similar to the previous model, dARL will also be trained offline. Since the standard practice for training the DRL agents is in the simulated environment (to avoid unnecessary expenses during the training and testing phase), we perform the training in MATLAB utilizing the "Reinforcement Learning Toolbox" (Figure 6 presents the simulated environment). We also assume that the base station has global observation of the coverage area, and it continuously communicates the regions' states with the UAVs. From the state observation, the UAVs (i.e., DRL agents) can intelligently decide if it needs to move to another region or not, and if so, which region.

5. Evaluation, Results, and Discussion

In this section, we present the experimental setup and assess the energy harvesting performance of the proposed combined framework. We validate the performance enhancement through DeNCE by comparing it with the spectral efficiency per user achieved by MBS-MMF algorithm [24], CURe [19], and Level four Fully Centralized Processing

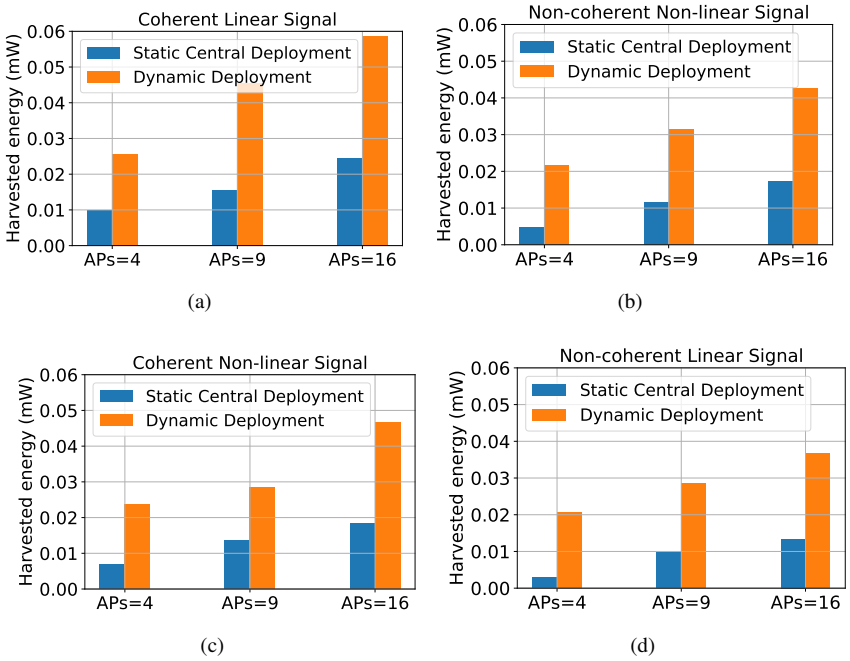
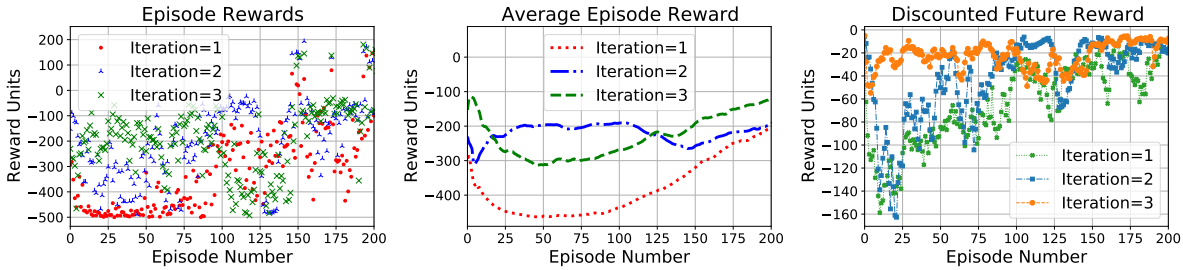


Figure 9: Amount of harvested energy with static (central) deployment strategy and dARL-based dynamic deployment strategy with (a) coherent linear signal, (b) non-coherent non-linear signal, (c) coherent non-linear signal, and (d) non-coherent linear signal, for 30 randomly deployed UE setups in 8 regions having served by 5 UAVs.

(L4-FCP) [54] with respect to the empirical cumulative distribution function value of the harvested energy. For calculating the downlink RF energy harvesting from the CFmMIMO UAVs, we utilize the optimization scheme from [24]. Additionally, we present the training performance of the DeNCE framework with respect to the training and validation loss and root mean square error (RMSE). Afterward, we validate the performance of dARL by comparing it with the standard static central deployment of APs, with respect to the amount of harvested energy. Then we analyze the performance of the dARL agents with respect to the achieved episode reward, average reward through the training, and discounted future reward. Next, we experiment with a variable number of UEs and agents while dividing the coverage area into different regions to observe the performance variation of the dARL framework. Finally, we experiment with the different deployment strategies of RISs (installments of RISs) and evaluate the effect of the number of RISs on the energy harvesting performance.

5.1. Spectral Efficiency per User with DeNCE

In this section, we assess the performance of the DeNCE framework from two perspectives; first, the evaluation of achieved spectral efficiency, and then the validation of its training efficiency, as represented in Figure 8. We compare the empirical cumulative distribution function of individual spectral efficiency per user, achieved by the MBS-MMF mechanism, the CURe framework, and the DeNCE augmented versions of both of the techniques. We further compare our results with L4-FCP, which involves sending the received pilot and data signals to the CPU for channel estimation and data signal detection. We choose to compare with L4-FCP since Bjornson et al. advocate it as the optimal choice among the multiple cell-free implementations analyzed [54]. We compare the spectral efficiency for the combinations of coherent linear and non-coherent non-linear energy transmissions. In Figure 8(a), for the coherent linear energy transmission, we observe that both techniques are achieving a higher amount of spectral efficiency compared to their standard versions when combined with DeNCE. Similarly, in Figure 8(b), it is observed that DeNCE aided MBS-MMF and CURe achieved higher spectral efficiency per user for the non-coherent non-linear signal as well. When the vertical axis is at 0.1 or 0.05, representing the 90% or 95% likely SE points, respectively, CURe in combination with DeNCE provides by far the highest efficiency among all other methods. Conversely, regular MBS-MMF has the lowest efficiency in these scenarios. Although L4-FCP achieves higher spectral efficiency per user than even the DeNCE augmented MBS-MMF, it performs worse than the regular CURe at the 90% or 95% likely SE points. All the models have comparable distributions, with the coherent linear signal having a more skewed distribution than the



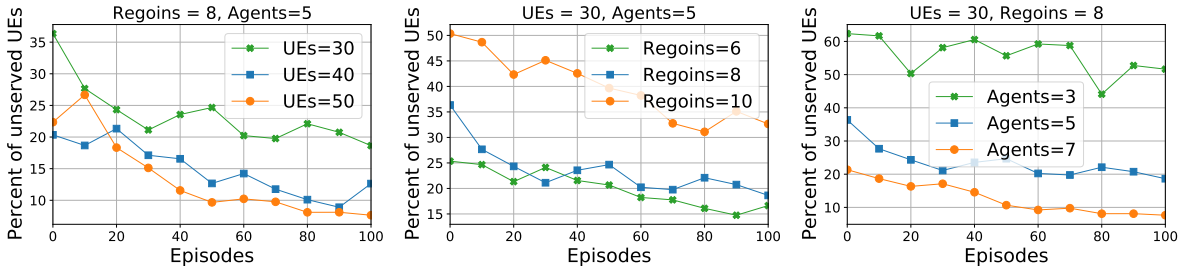
(a) (b) (c)

Figure 10: Reward accumulation analysis for dARL: (a) Comparison of the amount of total reward achieved in different iterations in each of the 200 training episodes, (b) Comparison of average reward achieved in different iterations through the 200 training episodes, and (c) Comparison of discounted long-term reward in different iterations in each of the 200 training episodes; with a collaborative agent set of 5 UAVs, while providing downlink RF signals to 30 UEs in a coverage area divided into eight regions.

non-coherent non-linear signal. The upper tail of the curves shows the users with favorable channel conditions, who achieve better spectral efficiency than other users due to their operation in the saturation area. Finally, we present the training effectiveness of the DeNCE framework in Figures 8(c) and 8(d). For the training dataset, we generated the closed-form-based channel estimation data using both the MBS-MMF and CURe mechanism and trained DeNCE on both datasets, each comprising twenty-seven thousand data samples. We split the data set into training, validation, and testing subsets, where the test set comprises 75% of the data set and the validation set comprises 10% of the data set. From Figure 8(c), it can be observed that the training loss starts to converge at around the 1000-th training iteration. We calculate the validation loss after every 30 iterations, and it can be observed that this loss converges around the 1500-th iteration. This signifies that the proposed model fits both the training data and the new unknown data really well. Later, in Figure 8(d), we observe that training RMSE has moderate but tolerable magnitude, referring to the fact that the model is introducing some sort of randomness in the form of residuals, which might mimic the natural environmental disturbance. This helps in achieving higher spectral efficiency than standard techniques. The validation RMSE approaches the convergence even before 1000 iterations.

5.2. *Performance Evaluation of d*ARL

In this part, we evaluate the performance of dARL framework in deploying the UAV-CFmMIMOs efficiently. Similar to the previous section, we perform the evaluation from two perspectives; first, by comparing the amount of harvested energy by dARL with the standard practice of AP deployment and then by validating the DRL's training efficiency. We compare the amount of harvested energy for all the combinations of coherent, non-coherent, linear, and non-linear energy transmission. It is evident from Figure 9 that the DRL-based dynamic deployment of UAV-CFmMIMOs facilitates a substantial amount of increased energy harvesting than static (central) deployment for all types of energy transmissions. We observe from Figure 9(a) that the highest amount of harvested energy is achieved for the coherent linear energy transmission, whereas the lowest amount of harvested energy is observed for the non-coherent linear energy transmission (Figure 9(d)). Moreover, we have observed that the amount of energy harvested is higher when a coherent signal is used compared to a non-coherent signal (i.e., energy harvested is greater in Figure 9(a) as compared to Figure 9(d), and in Figure 9(c) as compared to Figure 9(b)). The reason is that coherent signals have a well-defined amplitude and phase, which allows them to be efficiently utilized for energy harvesting. In contrast, non-coherent signals have randomly fluctuating amplitudes and phases, resulting in a lower amount of harvested energy. We also increase the number of UAV-CFmMIMO APs and observe that with an increased number of elements to deploy, dARL is facilitating exponential growth in the amount of harvested energy. Finally, for the second part of the evaluation, we present different training observations in Figure 10. We compare the evaluation results from three different iterations to show how quickly dARL agents learn to act towards the surging of rewards. In Figure 10(a), we show the exact accumulated reward by the collaborative agents in each of the 200 episodes. There are 1000 steps in each episode, and the agents try to serve all the UEs in different regions collaboratively. It can be observed from Figure 10(a) that, in the first iteration, the reward gain was pretty low (around -500 Reward Units) till almost the 75th episode. Then it slowly increases till the 100th episode. Afterward, an unstable reward accumulation trend is seen till the 200th episode. In the case of the second iteration, we see higher reward accumulation, meaning agents



(a) (b) (c)

Figure 11: Performance evaluation of dARL w.r.t. percentage of unserved UEs throughout 100 training episodes: (a) variable number of UEs having 8 regions and 5 DRL agents, (b) variable number of regions having 30 UEs and 5 DRL agents, and (c) a variable number of DRL agents having 30 UEs and 8 regions.

started exploiting the learned patterns (through weights and biases of the PPO network from the previous iteration). However, a similar unstable trend is seen throughout all 200 episodes. In the third iteration, we observe even increased rewards with a more stable trend. In Figure 10(b), we show the average reward of the collaborative agents through the episodes. Unlike the previous figure, here we can not actually perceive the exact increased rewards achieved towards the 200th episode (which is almost 200 Reward Units); however, we can better see the learned behavior of the collaborative agents. We observe that by the third iteration, agents show a pretty high average reward than the previous iterations. In Figure 10(c), we show the learning behavior toward the discounted long-term reward for the first three iterations. A discounted factor (DF) of value zero indicates that the current reward is prioritized, while a DF of value one indicates that future rewards are prioritized. In practice, zero-valued DF will never learn since it only considers current rewards, whereas a DF of one will continue to examine future rewards, possibly leading to infinity. Since in Figure 10(c), we show the discounted long-term reward, a stable discounter reward is expected throughout the training. We observe that the discounted long-term rewards are unstable in iterations one and two, whereas, in iteration three, the discounted long-term rewards are moving towards a stable trend. This manifests the effectiveness of the training stage.

*5.3. Optimal Parameter Setting for d*ARL

In this section, we experiment with variable environmental setups to evaluate the performance of dARL with respect to the portion of unserved UEs, as shown in Figure 11. By "unserved UEs", we refer to the UEs which are within the intended coverage regions, however, are not provided the energy harvesting facility (due to the absence of APs in those regions). We show the results throughout the different episodes, with random initialization in each episode, keeping some of the setup parameters constant. In Figure 11(a), we keep the number of regions and the number of agents constant while we alter the number of UEs to be served. An interesting trend can be seen from this experimentation, as the setup with the largest number of UEs has a minimal percentage of unserved UEs. In other words, the dARL framework performs better, having a larger set of UEs to serve. The reasoning could be with the considered setup, the "region density" and "UE count factor" of the reward function are able to schedule the DRL agents better. The percentage decreases to almost zero by the 100th training episode. In the case of 30 UEs, a slight converging behavior can be seen from the 20-25% range. In Figure 11(b), we keep the number of agents and the number of UEs constant while we vary the number of regions. We observe that by dividing the coverage area into a higher number of regions (i.e., 10), the percentage of unserved UEs increases. That is reasonable since the number of UAVs is half the number of regions to be served. However, after around 80 episodes of training, the unserved UEs percentage is reduced to around 30%, meaning the designed reward function ensures that the densest and most populated regions are covered. On the other hand, we observe almost similar performance from dARL with the setup we considered in Figure 11(b), having divided the coverage area into 6 and 8 regions. The number of regions and the number of agents considered are correlated, which was evident in the last experiment. In Figure 11(c), we keep the number of regions and the number of UEs constant while we alter the number of agents. It can be observed from Figure 11(c) that a higher number of agents are facilitating a higher percentage of UEs being served. In the case of 7 agents, we observe that close to 80% of the UEs are served, even in the first episode. That is apparent as the dARL agents might learn to just hover over different regions, each in a separate area, and still get the reward as all (or most

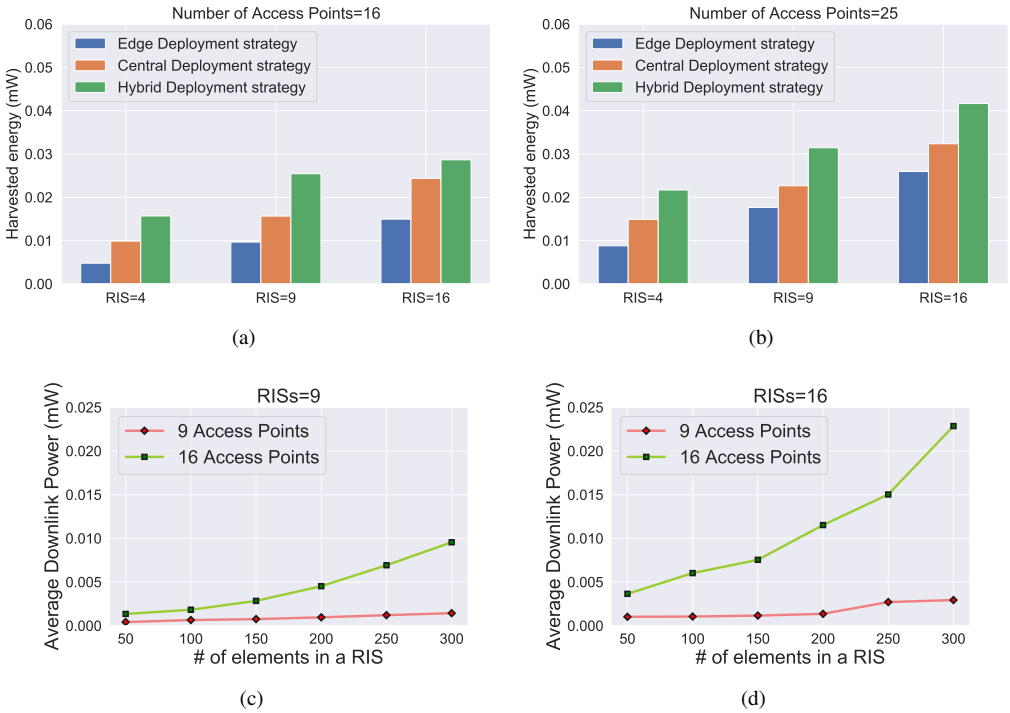


Figure 12: Comparison of different RIS deployment strategies with (a) 16 APs and (b) 25 APs, having a variable number of RISs. Evaluation of the effect on energy harvesting performance having (c) 9 RISs and (d) 16 RISs, with a variable number of meta elements.

of) the regions are served. Conversely, in the case of 3 agents, we observe that the unserved percentage of UEs is more than 50%, even after 100 training episodes. This emphasizes the importance of choosing the right parameter setting for the training environment.

5.4. Effect of RIS deployment and reflecting elements

Since RIS is an immobile smart material sheet that can be installed on walls, buildings, or ceilings [55], we conducted experiments to determine the best deployment strategy for it using three different approaches: edge, central, and hybrid deployment. In edge deployment, all RISs are placed at the edges of the coverage area, while in central deployment, they are placed in the center of the coverage regions. A hybrid deployment is a mixed approach where half of the RISs are placed at the edges and the other half in a centralized fashion. Figure 12 shows that the hybrid deployment strategy is optimal for all RIS options with 16 APs (Figure 12(a)) and 25 APs (Figure 12(b)). It can also be observed that the hybrid deployment strategy becomes more prominent with an increasing number of APs in the coverage areas. We also experimented with different numbers of meta elements in the RISs, assisting varying numbers of APs. Figure 12(c) shows that with 9 APs, the harvested energy is low, even with a higher number of meta elements in each of the 9 RISs, while with 16 APs, the amount of harvested energy seems to have an upward trend with the increasing number of elements in the RISs. In comparison, Figure 12(d) demonstrates a steeper curve as the number of RISs increases to 16, assisting 16 APs.

6. Literature Review

In recent years, the downlink RFEH with mMIMO has become an extensively researched subject in the scientific community. To name a few, Chen et al. examined the enhancement of energy efficiency for energy harvesting with mMIMO while ensuring satisfactory QoS via energy beamforming in [56]. Amarasuriya et al. investigated the performance of WPT for multi-cell multi-way mMIMO relaying by deriving the achievable sum rate vs. the harvested energy trade-offs [57]. The authors in [16] study mMIMO's energy harvesting capabilities by optimizing the minimum collected energy among energy-requiring devices while maintaining a minimum achievable rate for information-requiring devices. There has also been a lot of research into energy harvesting utilizing CFmMIMO. Like

in [17], Ngo et al. investigate the performance of CFmMIMO with conjugate beamforming in downlink by taking into account the combined effects of channel estimation, power control, and pilot sequence nonorthogonality. In [58], Nayebi et al. explore the downlink performance of CFmMIMO with zero-forcing (ZF) precoding and conjugate beamforming and present a low-complexity power allocation method. To enhance the performance of CFmMIMO at a low and affordable cost, RISs have been introduced with CFmMIMO systems. While most works concentrated on instantaneous performance measurements and depended on alternating optimization techniques ([59, 60]), Noh et al. proposed a unique two-step approach that offers long-term passive beamformers at RISs and short-term active precoders and long-term power allocation at APs to optimize the lowest possible rate.in [61] to deal with the huge computational complexity and signaling overhead imposed. To maximize the minimum achievable rate among users, Jin et al. proposed a joint optimization algorithm in [62] for RIS-aided CFmMIMO systems, maintaining transmit power constraints at different APs. For investigating the influence of the CSI uncertainty on IRS-assisted cell-free networks, Xie et al. adopt a stochastic programming method in [63] by maximizing the expectation of the sum rate, which ensured robust performance over the average. Elwekeil et al. introduced a novel scheme for power control in CFmMIMO that supported ultra-reliable low-latency communication applications for both traditional ground users and unmanned aerial vehicles in [64]. To efficiently allocate resources for RIS-assisted WPT of IoT, Zhu et al. introduce a new scheme to maximize the energy efficiency of the system by optimizing the time allocation, power splitting ratio, and reflection coefficients of the RIS [65]. Chu et al., on the other hand, propose a RIS-assisted wireless powered sensor network (WPSN) for IoT [66], which reduces energy consumption and cost while ensuring reliable communication between the sensors and the base station. Recently, Zhu et al. proposed a robust beamforming design for SWIPT in terahertz (THz) systems assisted by RIS. The proposed design aims to maximize the secure communication rate of the system while considering the non-linear energy harvesting model at the receiver [67]. However, none of the existing works incorporate the benefits of CFmMIMO, UAV-AP, and RIS technologies together.

Contrary to previously mentioned studies, Shrestha et al. evaluate the effectiveness of simultaneous wireless information and power transfer (SWIPT) for training-based CFmMIMO in [68]. They did, however, assume that information and energy users are separated. Similarly, Alageli et al. [69] investigated SWIPT with CFmMIMO, in which only the energy users harvest energy while the information users do not. In contrast, we address energy harvesting for all types of users, regardless of their primary requirement, by considering power control for maximizing the minimum uplink spectral efficiency for wireless power transfer using CFmMIMO, following the methodology of Demir et al. [24]. Furthermore, we present the concept of using RIS to support the CFmMIMO placed on UAV access points for increased energy harvesting performance. To the best of our knowledge, this is the first effort that combines all three methods for RFEH.

As mentioned earlier, channel estimation plays a crucial role in energy harvesting. Jiang et al. proposed using a trained deep neural network (DNN) model along with a pilot signal for efficient underwater channel estimation [43]. The authors of [44] recommended using a DNN model to utilize channel correlation in both frequency and time domains to conduct channel estimates for the standard of IEEE 802.11p. Furthermore, the authors in [45] evaluated the implications of the channel estimating phase for a wireless energy transfer system and proved that a downlink channel estimate is required to collect energy feedback information. A DNN structure produces better channel estimates than standard estimations like LS and LMMSE. However, none of the works considered a system with UAV-CFmMIMO and RIS-aided energy harvesting.

One of the major challenges for UAVs is the limited battery capacity, which revolves around the technical specification of the same. Thus for a swarm of UAVs, performing a global task together, an effective scheduling technique is pivotal for continuous operation without interruption. RL has been regarded as a prominent tool to surpass centralized control and allow local intelligence to accomplish collaborative global tasks in the UAV swarm domain. Like in [49], Bouhamed et al. proposed an RL-based framework that enables UAVs to autonomously arrange their schedules in order to cover the maximum number of pre-scheduled events in a particular geographic area but during a pre-determined time horizon. Ferdowsi et al. proposed a deep Q-learning (DQN)-based optimal scheduling policy for UAVs in the UAV-assisted wireless network by formulating the scheduling policy with a convex optimization-based solution in [70]. None of these works consider the mobility of the goals. Although Nguyen et al. proposed a DRL-based model for scheduling UAVs in device-to-device communications with a random walk of users in [71], they did not consider multiple UAVs collaboratively completing the task of energy harvesting. In this work, we propose a model for a swarm of UAVs that collectively try to achieve the energy harvesting goal.

7. Conclusion

In this work, we provide a comprehensive RF energy harvesting framework that takes advantage of the combined benefits of UAV-mounted CFmMIMO and RIS. The UAV-CFmMIMOs give a strong LoS signal in the broad coverage area, taking advantage of mobility in a cell-free fashion. Accordingly, we first devise a method for sending directive signals to target devices using RISs, which help with energy harvesting and information transfer. The empirical evaluation findings have validated that our framework can deliver a higher level of energy harvesting than the MBS-MMF [24]. For further improvement of energy harvesting performance, we propose to utilize the benefits of DL in estimating the channels, which is crucial in energy harvesting. Our analytical results confirm that the data-driven approach yields a considerably higher amount of energy harvesting compared to the traditional LS and LMMSE channel estimation techniques. Finally, to optimize the cost and augment the effective energy harvesting of the UEs, we propose DRL-based dynamic scheduling of UAV-CFmMIMOs, where the agents collaboratively maintain continuous coverage of downlink energy harvesting by observing the movement of the UEs in the coverage area.

References

[1] E. Kavyashree, H. Vidyashree, B. A. Kumar, A survey of internet of things (iot)-applications merits demerits & challenges, International Journal of Innovative Research in Computer and Communication Engineering 6 (2) (2018) 903–907.
[2] S. Balaji, K. Nathani, R. Santhakumar, Iot technology, applications and challenges: a contemporary survey, Wireless personal communications 108 (1) (2019) 363–388.
[3] J. Alzubi, R. Manikandan, O. Alzubi, N. Gayathri, R. Patan, A survey of specific iot applications, International Journal on Emerging Technologies 10 (1) (2019) 47–53.
[4] H. D. Kotha, V. M. Gupta, Iot application: a survey, Int. J. Eng. Technol 7 (2.7) (2018) 891–896.
[5] A. A. Khalil, A. J. Byrne, M. A. Rahman, M. H. Manshaei, Replanner: Efficient uav trajectory-planning using economic reinforcement learning, in: 2021 IEEE International Conference on Smart Computing (SMARTCOMP), IEEE, 2021, pp. 153–160.
[6] A. A. Khalil, M. A. Rahman, Fed-up: Federated deep reinforcement learning-based uav path planning against hostile defense system, in: 2022 18th International Conference on Network and Service Management (CNSM), IEEE, 2022, pp. 268–274.
[7] M. Sikimić, M. Amović, V. Vujović, B. Suknović, D. Manjak, An overview of wireless technologies for iot network, in: 2020 19th International Symposium INFOTEH-JAHORINA (INFOTEH), IEEE, 2020, pp. 1–6.
[8] D. Zorbas, C. Douligeris, Computing optimal drone positions to wirelessly recharge iot devices, in: IEEE INFOCOM 2018-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), IEEE, 2018, pp. 628–633.
[9] J. Huang, Y. Zhou, Z. Ning, H. Gharavi, Wireless power transfer and energy harvesting: Current status and future prospects, IEEE wireless communications 26 (4) (2019) 163–169.
[10] J. Ren, J. Hu, D. Zhang, H. Guo, Y. Zhang, X. Shen, Rf energy harvesting and transfer in cognitive radio sensor networks: Opportunities and challenges, IEEE Communications Magazine 56 (1) (2018) 104–110.
[11] H.-V. Tran, G. Kaddoum, Green cell-less design for rf-wireless power transfer networks, in: 2018 IEEE Wireless Communications and Networking Conference (WCNC), IEEE, 2018, pp. 1–6.
[12] X. Wang, A. Ashikhmin, X. Wang, Wirelessly powered cell-free iot: Analysis and optimization, IEEE Internet of Things Journal 7 (9) (2020) 8384–8396.
[13] T. L. Marzetta, Noncooperative cellular wireless with unlimited numbers of base station antennas, IEEE transactions on wireless communications 9 (11) (2010) 3590–3600.
[14] K. Zheng, L. Zhao, J. Mei, B. Shao, W. Xiang, L. Hanzo, Survey of large-scale mimo systems, IEEE Communications Surveys & Tutorials 17 (3) (2015) 1738–1760.
[15] G. Yang, C. K. Ho, R. Zhang, Y. L. Guan, Throughput optimization for massive mimo systems powered by wireless energy transfer, IEEE Journal on Selected Areas in Communications 33 (8) (2015) 1640–1650.
[16] L. Zhao, X. Wang, K. Zheng, Downlink hybrid information and energy transfer with massive mimo, IEEE Transactions on Wireless Communications 15 (2) (2015) 1309–1322.
[17] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, T. L. Marzetta, Cell-free massive mimo versus small cells, IEEE Transactions on Wireless Communications 16 (3) (2017) 1834–1850.
[18] E. Nayebi, A. Ashikhmin, T. L. Marzetta, B. D. Rao, Performance of cell-free massive mimo systems with mmse and lsfd receivers, in: 2016 50th Asilomar Conference on Signals, Systems and Computers, IEEE, 2016, pp. 203–207.
[19] A. A. Khalil, M. Y. Selim, M. A. Rahman, Cure: Enabling rf energy harvesting using cell-free massive mimo uavs assisted by ris, in: 2021 IEEE 46th Conference on Local Computer Networks (LCN), IEEE, 2021, pp. 533–540.
[20] E. Krijestorac, S. Hanna, D. Cabric, Uav access point placement for connectivity to a user with unknown location using deep rl, in: 2019 IEEE Globecom Workshops (GC Wkshps), IEEE, 2019, pp. 1–6.
[21] M. Diamanti, M. Tsampazi, E. E. Tsiropoulou, S. Papavassiliou, Energy efficient multi-user communications aided by reconfigurable intelligent surfaces and uavs, in: 2021 IEEE International Conference on Smart Computing (SMARTCOMP), IEEE, 2021, pp. 371–376.
[22] C. You, Z. Kang, Y. Zeng, R. Zhang, Enabling smart reflection in integrated air-ground wireless network: Irs meets uav, IEEE Wireless Communications 28 (6) (2021) 138–144.
[23] H. Mei, K. Yang, Q. Liu, K. Wang, 3d-trajectory and phase-shift design for ris-assisted uav systems using deep reinforcement learning, IEEE Transactions on Vehicular Technology 71 (3) (2022) 3020–3029.

[24] Ö. T. Demir, E. Björnson, Joint power control and lsfd for wireless-powered cell-free massive mimo, IEEE Transactions on Wireless Communications 20 (3) (2020) 1756–1769.
[25] F. A. Khan, Y. Chen, M.-S. Alouini, Novel receivers for af relaying with distributed stbc using cascaded and disintegrated channel estimation, IEEE transactions on wireless communications 11 (4) (2012) 1370–1379.
[26] E. Björnson, J. Hoydis, L. Sanguinetti, Massive mimo has unlimited capacity, IEEE Transactions on Wireless Communications 17 (1) (2017) 574–590.
[27] S. M. Kay, Fundamentals of statistical signal processing: estimation theory, Prentice-Hall, Inc., 1993.
[28] T. Van Chien, E. Björnson, E. G. Larsson, Joint pilot design and uplink power allocation in multi-cell massive mimo systems, IEEE Transactions on Wireless Communications 17 (3) (2018) 2000–2015.
[29] S. Wu, C.-X. Wang, H. Haas, M. M. Alwakeel, B. Ai, et al., A non-stationary wideband channel model for massive mimo communication systems, IEEE transactions on wireless communications 14 (3) (2014) 1434–1446.
[30] M. J. Peacock, I. B. Collings, M. L. Honig, Unified large-system analysis of mmse and adaptive least squares receivers for a class of random matrix channels, IEEE transactions on information theory 52 (8) (2006) 3567–3600.
[31] U. Mutlu, Y. Kabalci, Deep learning aided channel estimation approach for 5g communication systems, in: 2022 4th Global Power, Energy and Communication Conference (GPECOM), IEEE, 2022, pp. 655–660.
[32] V. W. Wong, R. Schober, D. W. K. Ng, L.-C. Wang, Key technologies for 5G wireless systems, Cambridge university press, 2017.
[33] J. Zhang, E. Björnson, M. Matthaiou, D. W. K. Ng, H. Yang, D. J. Love, Prospective multiple antenna technologies for beyond 5g, IEEE Journal on Selected Areas in Communications 38 (8) (2020) 1637–1660.
[34] M. A. Albreem, M. Juntti, S. Shahabuddin, Massive mimo detection techniques: A survey, IEEE Communications Surveys & Tutorials 21 (4) (2019) 3109–3132.
[35] J. Zhang, S. Chen, Y. Lin, J. Zheng, B. Ai, L. Hanzo, Cell-free massive mimo: A new next-generation paradigm, IEEE Access 7 (2019) 99878–99888.
[36] E. Basar, M. Di Renzo, J. De Rosny, M. Debbah, M.-S. Alouini, R. Zhang, Wireless communications through reconfigurable intelligent surfaces, IEEE access 7 (2019) 116753–116773.
[37] C. Huang, A. Zappone, G. C. Alexandropoulos, M. Debbah, C. Yuen, Reconfigurable intelligent surfaces for energy efficiency in wireless communication, IEEE Transactions on Wireless Communications 18 (8) (2019) 4157–4170.
[38] C. Liaskos, S. Nie, A. Tsioliaridou, A. Pitsillides, S. Ioannidis, I. Akyildiz, A new wireless communication paradigm through software-controlled metasurfaces, IEEE Communications Magazine 56 (9) (2018) 162–169.
[39] J. Zhao, A survey of intelligent reflecting surfaces (irss): Towards 6g wireless communication networks, arXiv preprint arXiv:1907.04789.
[40] O. M. Bushnaq, M. A. Kishk, A. Celik, M.-S. Alouini, T. Y. Al-Naffouri, Optimal deployment of tethered drones for maximum cellular coverage in user clusters, IEEE Transactions on Wireless Communications 20 (3) (2020) 2092–2108.
[41] M. Kishk, A. Bader, M.-S. Alouini, Aerial base station deployment in 6g cellular networks using tethered drones: The mobility and endurance tradeoff, IEEE Vehicular Technology Magazine 15 (4) (2020) 103–111.
[42] Q. Hu, F. Gao, H. Zhang, S. Jin, G. Y. Li, Deep learning for channel estimation: Interpretation, performance, and comparison, IEEE Transactions on Wireless Communications 20 (4) (2020) 2398–2412.
[43] R. Jiang, X. Wang, S. Cao, J. Zhao, X. Li, Deep neural networks for channel estimation in underwater acoustic ofdm systems, IEEE access 7 (2019) 23579–23594.
[44] A. K. Gizzini, M. Chafii, A. Nimr, G. Fettweis, Deep learning based channel estimation schemes for ieee 802.11 p standard, IEEE Access 8 (2020) 113751–113765.
[45] J.-M. Kang, C.-J. Chun, I.-M. Kim, Deep-learning-based channel estimation for wireless energy transfer, IEEE Communications Letters 22 (11) (2018) 2310–2313.
[46] M. Soltani, V. Pourahmadi, A. Mirzaei, H. Sheikhzadeh, Deep learning-based channel estimation, IEEE Communications Letters 23 (4) (2019) 652–655.
[47] H. He, C.-K. Wen, S. Jin, G. Y. Li, Deep learning-based channel estimation for beamspace mmwave massive mimo systems, IEEE Wireless Communications Letters 7 (5) (2018) 852–855.
[48] H. Ye, G. Y. Li, B.-H. Juang, Power of deep learning for channel estimation and signal detection in ofdm systems, IEEE Wireless Communications Letters 7 (1) (2017) 114–117.
[49] O. Bouhamed, H. Ghazzai, H. Besbes, Y. Massoud, A generic spatiotemporal scheduling for autonomous uavs: A reinforcement learning-based approach, IEEE Open Journal of Vehicular Technology 1 (2020) 93–106.
[50] A. Ferdowsi, M. A. Abd-Elmagid, W. Saad, H. S. Dhillon, Neural combinatorial deep reinforcement learning for age-optimal joint trajectory and scheduling design in uav-assisted networks, IEEE Journal on Selected Areas in Communications 39 (5) (2021) 1250–1265.
[51] K. K. Nguyen, N. A. Vien, L. D. Nguyen, M.-T. Le, L. Hanzo, T. Q. Duong, Real-time energy harvesting aided scheduling in uav-assisted d2d networks relying on deep reinforcement learning, IEEE Access 9 (2020) 3638–3648.
[52] Y. Chen, N. Zhao, M.-S. Alouini, Wireless energy harvesting using signals from multiple fading channels, IEEE Transactions on Communications 65 (11) (2017) 5027–5039.
[53] E. Björnson, Ö. Özdogan, E. G. Larsson, Intelligent reflecting surface versus decode-and-forward: How large surfaces are needed to beat relaying?, IEEE Wireless Communications Letters 9 (2) (2019) 244–248.
[54] E. Björnson, L. Sanguinetti, Making cell-free massive mimo competitive with mmse processing and centralized implementation, IEEE Transactions on Wireless Communications 19 (1) (2019) 77–90.
[55] M. Di Renzo, A. Zappone, M. Debbah, M.-S. Alouini, C. Yuen, J. De Rosny, S. Tretyakov, Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and the road ahead, IEEE journal on selected areas in communications 38 (11) (2020) 2450–2525.
[56] X. Chen, X. Wang, X. Chen, Energy-efficient optimization for wireless information and power transfer in large-scale mimo systems employing energy beamforming, IEEE Wireless Communications Letters 2 (6) (2013) 667–670.
[57] G. Amarasuriya, E. G. Larsson, H. V. Poor, Wireless information and power transfer in multiway massive mimo relay networks, IEEE

Transactions on Wireless Communications 15 (6) (2016) 3837–3855.
[58] E. Nayebi, A. Ashikhmin, T. L. Marzetta, H. Yang, B. D. Rao, Precoding and power optimization in cell-free massive mimo systems, IEEE Transactions on Wireless Communications 16 (7) (2017) 4445–4459.
[59] Z. Zhang, L. Dai, A joint precoding framework for wideband reconfigurable intelligent surface-aided cell-free network, IEEE Transactions on Signal Processing 69 (2021) 4085–4101.
[60] Q. N. Le, V.-D. Nguyen, O. A. Dobre, R. Zhao, Energy efficiency maximization in ris-aided cell-free network with limited backhaul, IEEE Communications Letters 25 (6) (2021) 1974–1978.
[61] T. Noh, J. Choi, Cell-free mimo systems powered by intelligent reflecting surfaces, IEEE Communications Letters 26 (5) (2022) 1076–1080.
[62] S.-N. Jin, D.-W. Yue, H. H. Nguyen, Ris-aided cell-free massive mimo systems: Joint design of transmit beamforming and phase shifts, arXiv preprint arXiv:2112.06593.
[63] X. Xie, C. He, X. Li, Z. J. Wang, Robust joint design for intelligent reflecting surfaces assisted cell-free networks, arXiv preprint arXiv:2201.09685.
[64] M. Elwekeil, A. Zappone, S. Buzzi, Power control in cell-free massive mimo networks for uavs urllc under the finite blocklength regime, arXiv preprint arXiv:2111.10613.
[65] Z. Zhu, Z. Li, Z. Chu, G. Sun, W. Hao, P. Liu, I. Lee, Resource allocation for intelligent reflecting surface assisted wireless powered iot systems with power splitting, IEEE Transactions on Wireless Communications 21 (5) (2021) 2987–2998.
[66] Z. Chu, Z. Zhu, F. Zhou, M. Zhang, N. Al-Dhahir, Intelligent reflecting surface assisted wireless powered sensor networks for internet of things, IEEE Transactions on Communications 69 (7) (2021) 4877–4889.
[67] Z. Zhu, J. Xu, G. Sun, W. Hao, Z. Chu, C. Pan, I. Lee, Robust beamforming design for irs-aided secure swipt terahertz systems with non-linear eh model, IEEE Wireless Communications Letters 11 (4) (2022) 746–750.
[68] R. Shrestha, G. Amarasuriya, Swipt in cell-free massive mimo, in: 2018 IEEE Global Communications Conference (GLOBECOM), IEEE, 2018, pp. 1–7.
[69] M. Alageli, A. Ikhlef, F. Alsifiany, M. A. Abdullah, G. Chen, J. Chambers, Optimal downlink transmission for cell-free swipt massive mimo systems with active eavesdropping, IEEE Transactions on Information Forensics and Security 15 (2019) 1983–1998.
[70] A. Ferdowsi, M. A. Abd-Elmagid, W. Saad, H. S. Dhillon, Neural combinatorial deep reinforcement learning for age-optimal joint trajectory and scheduling design in uav-assisted networks, IEEE Journal on Selected Areas in Communications`doi:10.1109/JSAC.2021.3065049`.
[71] K. K. Nguyen, N. A. Vien, L. D. Nguyen, M.-T. Le, L. Hanzo, T. Q. Duong, Real-time energy harvesting aided scheduling in uav-assisted d2d networks relying on deep reinforcement learning, IEEE Access 9 (2021) 3638–3648. `doi:10.1109/ACCESS.2020.3046499`.