

Received 8 June 2025; revised 8 July 2025; accepted 26 July 2025. Date of publication 1 August 2025; date of current version 25 August 2025.

Digital Object Identifier 10.1109/OJCOMS.2025.3594373

Deep Reinforcement Learning-Based Joint Trajectory Design and Resource Allocation for Secure and Energy-Efficient UAV Networks

ABDULMALIK ALWARAFY¹ (Member, IEEE), SUHIB BANI MELHEM² (Member, IEEE),
 RAND ABOU CHAHINE¹, BATOOL SAID¹, MAITHA ALHARETHI¹, LATIFA ALMAZROUEI¹,
 AND SARA ALBLOOSHI¹

¹Computer and Network Engineering Department, College of Information Technology, United Arab Emirates University, Al Ain, UAE

²Department of Cybersecurity, College of Engineering, Al Ain University, Abu Dhabi, UAE

CORRESPONDING AUTHOR: A. ALWARAFY (e-mail: aalwarafy@uaeu.ac.ae).

This work was supported by the Sure Plus Grant from the United Arab Emirates University (UAEU).

ABSTRACT Unmanned Aerial Vehicles (UAVs) have been extensively used recently for wireless networks. However, such networks encounter several challenges that remain unsolved. In this paper, we address the issue of joint optimization of trajectory design and resource allocation in UAV-based wireless networks in the presence of eavesdroppers. We first formulate an optimization problem with the objective to maximize a utility function defined in terms of secrecy rate, energy utilization efficiency, and interference. Due to the high dimensionality and non-convex nature of the formulated problem, we propose a Proximal Policy Optimization (PPO)-based Deep Reinforcement Learning (DRL) algorithm to solve the problem and learn the environment. Our proposed PPO algorithm solves the problem by jointly controlling the 3D position of UAVs, power, and energy harvesting. Simulation results demonstrate the efficiency of the proposed algorithm in solving the problem, learning the environment dynamics, and its superiority over some existing conventional and DRL-based methods.

INDEX TERMS Deep reinforcement learning, secrecy rate, resource allocation, UAV, trajectory design, wireless communication.

I. INTRODUCTION

IN RECENT years, UAV-based wireless networks have gained significant research interest and are being progressively adopted across various sectors of society. Their broad range of applications includes aerial photography, precision agriculture, environmental monitoring [1], and search and rescue operations [2]. UAV networks are also used in military applications [3], where secure and efficient communication is critical. As UAV technology advances, the ability to coordinate multiple UAVs in a networked environment opens up even more possibilities, such as disaster response, delivery services, and infrastructure inspection. However, because of the high demand for multi-UAV networks, there must be effective countermeasures to ensure the safety of the network from potential security issues that could threaten data and communication integrity. Protecting UAV networks

from unauthorized access and interference is crucial for their reliable and secure operation in civilian and defense sectors.

Despite their impressive capabilities, UAV-enabled networks face various challenges, with security being one of the most critical concerns. Because of the open nature of wireless communication and the high mobility of UAVs, these networks are particularly vulnerable to different attacks. Threats such as jamming, eavesdropping, hijacking, and spoofing [4], [5] pose serious risks to the integrity and confidentiality of the data transmitted in the UAV network. Such vulnerabilities can compromise the stability and reliability of the network, which makes implementing effective security measures very crucial and necessary. One of the primary goals for secure UAV communication is achieving a high secrecy rate, which involves ensuring that only authorized users can access and interpret sensitive information.

By enhancing the secrecy rate, UAV networks can enable secure communication channels, safeguard confidential data transmission, and reduce the risk of unauthorized interception or manipulation by eavesdroppers, thus promoting a more secure network environment.

Existing approaches often encounter substantial challenges within UAV environments due to the high-dimensional nature of the problem space and the rapid, unpredictable mobility of UAVs. These factors make many traditional techniques both inefficient and impractical, as they often struggle to adapt to the real-time demands of complex UAV networks. Deep Reinforcement Learning (DRL), however, offers a promising solution to these limitations by allocating the learning process across multiple agents or computing nodes, allowing it to address the large-scale and high-dimensional demands typical of UAV scenarios [6], [7], [8]. Unlike traditional reinforcement learning, DRL enables UAVs to learn adaptive, optimal behaviors without the need for extensive pre-defined rules, which can restrict and limit flexibility in evolving environments. Through DRL, UAVs can autonomously learn and execute tasks such as efficient path planning, collision avoidance, and adaptive resource allocation. This adaptability is especially valuable in dynamic and densely populated scenarios, where real-time decision-making and reactivity are essential. Consequently, DRL significantly enhances the operational efficiency and versatility of UAV networks, making them far better suited to handle the complexities of modern, large-scale applications.

A. RELATED WORK

DRL techniques have been widely used lately in addressing various issues and challenges in UAV networks. In this section, we focus only on research works that address the secrecy rate issue in UAV networks using DRL and conventional methods. However, related works on DRL-based secrecy rate in non-UAV networks can be found in [9], [10], [11], [12], [13], [14]. The authors in [15] introduced an objective of maximizing secrecy rates in UAVs by optimizing the UAVs' beamforming and trajectory. However, optimizing the power of the UAVs was not studied. In [16], a resource allocation scheme to enhance the secrecy rate for UAV-assisted cognitive radios was presented by optimizing the spectrum sensing time, the power splitting ratio, and the UAVs' hovering locations. A security algorithm employing DRL is proposed in [17] that enables UAVs to find the optimal flight strategy in order to maximize the average secrecy rate in communications. In [18], the authors introduced a DRL-based approach to correct UAV's localization errors. In [19], the authors formed an optimization problem aiming to minimize the UAV's flight energy consumption and proposed a DRL-based method to address that problem. The work in [20] maximizes the average secrecy rate by optimizing trajectory and power under constraints by employing a Proximal Policy Optimization (PPO) algorithm to handle the continuous action space. In [21], the authors maximized secrecy rates

in a UAV-assisted relay network by jointly designing the transmit beamformer and artificial noise with a transmit power-constrained UAV. The authors in [22] proposed a deep unfolding DRL algorithm to maximize the secrecy rate in a network of multiple UAVs considering a rate-splitting multiple access (RSMA) network. The authors in [23] studied a ground-to-UAV (G2U) communication system in the presence of a malicious eavesdropper. They proposed a DRL-based algorithm to optimize UAV trajectory planning for enhanced network security. In [24], the authors proposed a multi-agent DRL algorithm to secure communications in UAV networks in the presence of smart UAV eavesdroppers. To optimize the UAV-base station trajectory for effective communication in catastrophe scenarios, the authors in [25] proposed a DRL-based method to address this issue. In [26], the authors proposed a novel model based on a post-decision state deep Q-network combined with a Fourier feature mapping algorithm (referred to as PDS-DQN-FFM) to enable reliable and secure communication in UAV-RIS networks. Their goal was to maximize the secrecy rate while ensuring the QoS requirements of the legitimate user in the presence of both eavesdroppers and a jammer. Interestingly, the authors in [27] proposed a multi-agent deep deterministic policy gradient (MADDPG)-based approach to maximize secure capacity by jointly optimizing the trajectories of UAVs, the transmit power of UAV transmitters, and the jamming power of UAV jammers in the presence of ground eavesdroppers.

However, none of these works have addressed the joint optimization of secrecy rate, energy efficiency, and interference in the presence of both a ground eavesdropper and a wireless charging base station using an advanced DRL algorithm, such as the PPO algorithm. This paper fills that research gap, as detailed in the following sections.

B. CONTRIBUTIONS

This paper addresses the challenge of jointly maximizing a utility function defined by secrecy rate, energy utilization efficiency, and interference management through the combined optimization of 3D position for transmitting and jamming UAVs, resource allocation, and energy harvesting. The study considers a UAV network operating in the presence of an eavesdropper and multiple transmitting and jamming UAVs. Given the problem's high non-convexity, complexity, and difficulty in obtaining a globally optimal solution, we propose a DRL-based approach utilizing the PPO algorithm to solve the problem and learn network dynamics. The key contributions of this paper are outlined as follows:

- We formulate a multi-objective non-convex optimization problem aimed at simultaneously maximizing the network's secrecy rate and energy utilization efficiency while minimizing interference by jointly optimizing UAVs' 3D positioning, power allocation, and energy harvesting.
- Given the highly non-convex nature and computational complexity of the formulated problem, along with the challenges in using conventional approaches, we

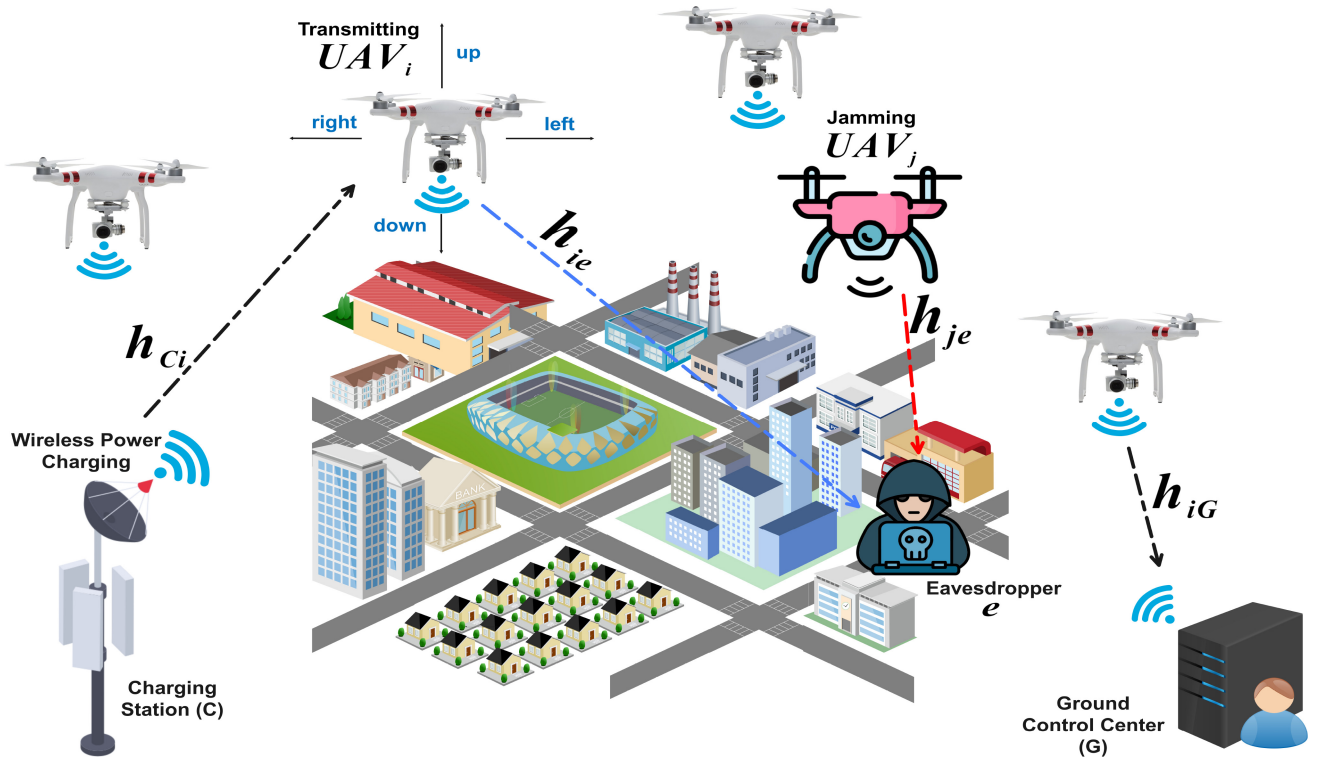


FIGURE 1. Illustration of our proposed UAV system model.

propose a centralized PPO-based DRL architecture to address it and capture network dynamics. Unlike most existing studies that assume perfect channel state information (CSI) availability, our method operates with limited network and CSI knowledge.

- By carefully defining the DRL network parameters, we use simulations to showcase how the proposed algorithm effectively learns the optimal policy to address the formulated optimization problem.
- Through simulations, we showcase that our proposed PPO-based DRL method outperforms conventional and DRL-based approaches in both convergence speed and utility values. Additionally, we demonstrate the capability of our DRL model to rapidly adapt to the unpredictable mobility of eavesdroppers within the UAV network.

The rest of the paper is structured as follows. Section II introduces our proposed system model and problem formulation. Section III details our proposed PPO DRL-based approach, explains the key components of the DRL model, and provides a brief overview of the PPO algorithm. Section IV explains the simulation setup and provides a comprehensive analysis of the numerical results. Finally, Section V concludes this paper.

NOTATIONS: $\mathbf{A}$, $\mathbf{a}$, a , and $\mathcal{A}$ denote a matrix, vector, scalar, and set, respectively. $\text{Tr}(\mathbf{a})$ denotes the trace of matrix $\mathbf{a}$ and $\text{diag}(\mathbf{a})$ denotes a diagonal matrix whose diagonal is the vector $\mathbf{a}$. The superscript $(\cdot)^H$ and $(\cdot)^T$ denote the Hermitian and transpose operations, respectively. a and $\|\mathbf{a}\|$

represent the magnitude of a complex number and Euclidean norm of a vector, respectively. $\mathbb{C}^{x \times y}$ is the space of $x \times y$ complex matrices, and $\otimes$ represents the outer product of two vectors.

II. SYSTEM MODEL AND PROBLEM FORMULATION

This section explains in details our proposed system model along with the associated optimization problem formulation.

A. NETWORK MODEL

Our proposed network model is illustrated in Fig. 1. The network consists of several transmitting UAVs hovering within the coverage area to collect data and transmit it to the ground control center (G) for further processing. We assume the presence of an unauthorized ground passive eavesdropper attempting to intercept and compromise the data transmitted from the UAVs over the communication channel to the ground control center (G). To counter this threat, we deploy jamming UAVs within the network. These jamming UAVs continuously transmit known artificial noise over the null UAV-to-ground control center (G) communication channel to ensure the confidentiality and secrecy of the transmitted information. This can be achieved using precoding or beamforming techniques to ensure the artificial noise transmitted by the jamming UAVs lies in the null space of the legitimate UAV-to-ground control center (G) channel. The ground control center (G) is presumed to know the channel state information (CSI) and artificial noise in advance. This enables it to nullify the noise and retrieve

TABLE 1. List of mathematical notations used in this paper.

Symbol	Definition	Symbol	Definition
N_t	Total number of transmitting UAVs.	N_j	Total number of jamming UAVs.
N_e	Total number of ground unauthorized eavesdroppers.	n_{LoS}	LoS path loss exponent.
h_{iG}	Channel gain between transmitting UAV i & ground center G .	n_{NLoS}	NLoS path loss exponent.
h_{ie}	Channel gain between transmitting UAV i & eavesdropper e .	p_i	Transmit power of i th transmitting UAV.
h_{je}	Channel gain between jamming UAV j & eavesdropper e .	p_j	Transmit power of j th jamming UAV.
h_{Ci}	Channel gain between charging station C & transmitting UAV i .	P^{max}	Maximum power budget of UAVs.
$PL_{LoS}(d_{TXRX})$	LoS path loss in dB between TX & RX.	v_{max}	The UAV's velocity.
$PL_{NLoS}(d_{TXRX})$	NLoS path loss in dB between TX & RX.	B^{max}	Maximum battery capacity of UAVs.
$\overline{PL}(d_{TXRX})$	Average path loss for LoS & NLoS.	I_{tot}	Total interference of the network.
p_{LoS}	Probability of LoS link.	p_{NLoS}	Probability of NLoS link.
Γ_{iG}	SNR between i th UAV & ground center.	Γ_{ie}	SNR between i th UAV & e th eavesdropper.
d_{min}	Min distance between two transmitting/jamming UAVs.	R_{ie}	Rate between i th UAV & e th eavesdropper.
d_{iG}	Total distance between i th UAV & ground center.	E_{tot}	Total network energy utilization efficiency.
l_i/l_j	Size of data packet transmitted from the i th UAV/ j th UAV.	R_{tot}^{sec}	Total network secrecy rate.
e_i^t	Energy consumed by i th transmitting UAV to transmit data.	e_i^f	Energy consumed by transmitting UAVs for flight.
e_j^t	Energy consumed by j th jamming UAV to send artificial noise.	e_j^f	Energy consumed by jamming UAVs for flight.
e_i^h	Energy transferred from charging station to i th transmitting UAV.	ζ_i/ζ_j	Energy harvesting efficiency of i th/ j th UAV.
E_i^h/E_j^h	Energy consumed by i th/ j th UAV harvesting technique.	b_i	Battery level (status) of i th UAV.
e_j^h	Energy transferred from charging station to j th transmitting UAV.	R_i^{sec}	Secrecy rate for i th transmitting UAV.
L	Max horizontal distance UAVs cross in one time slot.	R_{iG}	Rate between i th UAV & ground center.

our confidential transmitted data. Meanwhile, the passive eavesdropper intercepting the communication channel will experience degraded signal quality due to the artificial noise, which will eventually hinder its data extraction capability. Note that the artificial noise from jamming UAVs can interfere with legitimate transmitting UAVs, causing a degradation of their signal quality. To address this issue, we explicitly model and constrain this interference in the optimization problem **P1** by ensuring it remains below a predefined threshold (as shown in constraint C_3). In our framework, we assume that the system has knowledge of the jamming UAVs' transmit power, which we use as a proxy to estimate the interference. This approximation enables tractable modeling while preserving relevance to practical scenarios. Such information can be obtained in practical scenarios through the centralized ground control center (G) or coordinated mission planning. In many UAV network architectures, especially those involving ground control centers, the transmit power levels of cooperative jamming UAVs are typically pre-configured or dynamically managed by this controller. This makes it reasonable to assume that the system is aware of the jamming UAVs' transmit power, particularly when these UAVs are part of a coordinated security mechanism within the same network infrastructure. On the other hand, as depicted in Fig. 1, it is assumed that the UAV network is equipped with a charging station (C) to recharge both transmitting and jamming UAVs and extend the lifespan of their batteries. Additionally, it is assumed that each transmitting and jamming UAV is equipped with energy harvesting functionality to transform the received energy from the charging station (C) into usable power. In our network framework, activities are executed in discrete time slots t . At each time slot, the current network

state remains static but evolves in the subsequent slot. During each time slot, the UAV network determines the following for the transmitting and jamming UAVs: the 3D locations, transmission and jamming power, and the energy transferred from the charging station (C) to the transmitting and jamming UAVs.

Let's denote N_t as the total number of transmitting UAVs, N_j as the total number of jamming UAVs, N_e as the total number of ground unauthorized eavesdroppers, H_{ij} as the channel gain between the i th transmitting UAV and j th jamming UAV, h_{ie} as the channel gain between the i th transmitting UAV and e th eavesdropper, h_{je} as the channel gain between the j th jamming UAV and e th eavesdropper, p_i as the transmit power of the i th transmitting UAV, and p_j as the transmit power of the j th jamming UAV. In this paper, we will reserve the letters i to denote transmitting UAVs $\forall i \in N_t$, j to denote jamming UAVs $\forall j \in N_j$, and e to denote ground eavesdropper $\forall e \in N_e$. For clarity and completeness, Table 1 lists the notations used in this paper.

B. SECRECY COMMUNICATION MODEL

In this study, we consider downlink communication between the transmitting UAVs and the ground control center (G). The air-to-ground channel model characterizes the channel conditions between the UAVs and the ground control center (G). We adopt the widely used probabilistic model, as in [28], [29], [30], [31], [32], which assumes that received signals can be either line-of-sight (LoS) or non-line-of-sight (NLoS) with certain probabilities. These probabilities are functions of the environment, the density and height of surrounding buildings, and the elevation angle. The probabilities of LoS (p_{LoS}) and NLoS (p_{NLoS}) are determined using the probabilistic model

as follows [28], [29], [30], [31], [32]:

$$p_{LoS} = \frac{1}{(1 + \beta \exp(-a[\frac{180}{\pi} \arctan(\frac{h}{d}) - \beta])} \quad (1)$$

where the constants a and β are determined by the urban environment, h is the hovering altitude of UAVs, and d represents the horizontal distance between the UAVs and the ground control center (G). The p_{NLoS} is calculated using $p_{NLoS} = 1 - p_{LoS}$.

The path loss equations for both LoS and NLoS cases are defined as follows [28], [29], [30], [31], [32], [33], [34]:

$$PL_{LoS}(d_{TXRX}) = \lambda_{LoS} + 10n_{LoS} \log_{10} \left(\frac{4\pi f_c d_{TXRX}}{c} \right) \quad (2)$$

$$PL_{NLoS}(d_{TXRX}) = \lambda_{NLoS} + 10n_{NLoS} \log_{10} \left(\frac{4\pi f_c d_{TXRX}}{c} \right) \quad (3)$$

where d_{TXRX} is the transmitter-receiver separation distance, λ_{LoS} and λ_{NLoS} are an environmental constant, f_c is the carrier frequency, and c is the light speed. Note that the probability of receiving a LoS signal increases with a higher elevation angle in 1. Consequently, the average path loss is calculated using [28], [29], [30], [31], [32]:

$$\overline{PL}(d_{TXRX}) = p_{LoS} \times PL_{LoS} + p_{NLoS} \times PL_{NLoS} \quad (4)$$

It is noteworthy that modeling the channel as either LoS or NLoS with certain probabilities (i.e., using a Bernoulli selection per time slot) is also a valid option to reflect the instantaneous CSI in each slot. However, we intentionally adopted the averaged path loss model shown in (4), as also used in [28], [35], for these two reasons: 1) to simplify the system dynamics for tractable optimization and stable DRL training, and 2) to accommodate the cases of partial LoS/NLoS conditions.

The channel gain between any transmitter-receiver communication link as shown in Fig. 1 is given by [36], [37], [38]:

$$h_{TXRX}(t) = 10^{\frac{-\overline{PL}(d_{TXRX})}{10}} |\Gamma_{TXRX}(t)|^2, \quad (5)$$

where $\Gamma_{TXRX}(t)$ is the small scale fading of the transmitter-receiver communication channel.

Assuming an OFDM-based system with flat fading, the upper bound on the downlink data rate for all communication links depicted in Fig. 1 at time slot t is formulated as in (6). Accordingly, the transmission data rate from the i th transmitting UAV to the ground control center (G) at time slot t is given by [28], [37], [38]:

$$R_{iG}(t) = \omega \log_2 \left(1 + \frac{p_i(t)h_{iG}(t)}{\omega N_0} \right), \forall i \in N_t \quad (6)$$

where ω and $p_i(t)$ represent the bandwidth allocation (which is assumed to be fixed and pre-allocated in our system model) and transmit power for the i th UAV, respectively, and N_0 denotes the noise power spectral density. The transmission

data rate from the i th UAV to the e th eavesdropper at time slot t is expressed as:

$$R_{ie}(t) = \omega \log_2 \left(1 + \frac{p_i(t)h_{ie}(t)}{\omega N'_0 + p_j(t)h_{je}(t)} \right), \forall i, j \in N_t, N_j \quad (7)$$

where $h_{ie}(t)$ and $h_{je}(t)$ are defined by (5) and they represent the channel gain between the i th transmission UAV and the e th eavesdropper, and between the j th jamming UAV and the e th eavesdropper, respectively, as depicted in Fig. 1. $p_j(t)$ is the artificial transmit noise power of the j th jamming UAV, which is intentionally introduced to reduce the ability of the eavesdropper to intercept the legitimate signals.

To guarantee secure communication between the UAV network and the ground control center (G), maximizing the secrecy rate of the legitimate channel is crucial. The average secrecy rate is calculated as the difference between the wireless transmission rate from the i th UAV to the ground control center (G) and the wireless transmission rate from the i th UAV to the e th eavesdropper, i.e., [28], [39]:

$$R_i^{sec}(t) = R_{iG}(t) - R_{ie}(t), \quad \forall i, e \in N_t, N_e \quad (8)$$

Consequently, the overall secrecy rate attained by all UAVs within the network illustrated in Fig. 1 is given by:

$$R_{tot}^{sec}(t) = \sum_{i=1}^{N_t} R_i^{sec}(t) = \sum_{i=1}^{N_t} R_{iG}(t) - R_{ie}(t), \quad (9)$$

C. ENERGY MODEL

This section presents the energy model adopted in our proposed system framework depicted in Fig. 1. We consider three energy models: the energy consumption for data transmission, the energy utilized by jamming UAVs, and the energy harvested by UAVs.

1) ENERGY OF DATA TRANSMISSION

The energy consumed by the i th transmitting UAV to send data to the ground control center (G) is determined by:

$$e_i^t(t) = \frac{p_i(t)l_i(t)}{R_{iG}(t)} = \frac{p_i(t)l_i(t)}{\omega \log_2 \left(1 + \frac{p_i(t)h_{iG}(t)}{\omega N_0} \right)} \quad (10)$$

where l_i represents the size of data packet transmitted from the i th transmitting UAV to the ground control center (G).

2) ENERGY OF JAMMING UAVS

The primary function of the jamming UAVs illustrated in Fig. 1 is to prevent the ground eavesdropper from accurately receiving the data sent by the transmitting UAVs. The energy consumed by the j th jamming UAV to send artificial noise and disrupt the signal intercepted by the e th eavesdropper is expressed as [28], [39]:

$$e_j^t(t) = \frac{p_j(t)l_j(t)}{R_{ie}(t)} = \frac{p_j(t)l_j(t)}{\omega \log_2 \left(1 + \frac{p_i(t)h_{ie}(t)}{\omega N'_0 + p_j(t)h_{je}(t)} \right)} \quad (11)$$

where l_j represents the size of data packet transmitted from the j th jamming UAV to the ground control center

(G). In our model, the jamming UAVs are assumed to transmit artificial noise during the same time slots as the transmitting UAVs. In our model, the jamming UAVs are assumed to transmit artificial noise during the same time slots as the transmitting UAVs. However, the energy consumption in (11) is calculated based on the effective jamming duration required to ensure sufficient interference at the eavesdropper side. This modeling choice allows us to account for the trade-off between jamming efficiency and energy usage under realistic transmission constraints. In particular, this definition captures a practical energy-aware design philosophy, where jamming is applied only as much as necessary to achieve secure communication in order to preserve energy resources. The effective jamming duration is determined by the minimum time required to reduce the eavesdropper's received signal quality below a secrecy rate threshold. This is implicitly governed by factors such as jamming power, CSI conditions, and the eavesdropper's decoding capability.

3) HARVESTING ENERGY

In the system model illustrated in Fig. 1, we employ wireless radio frequency-based energy harvesting technology to recharge all UAVs in the network, thereby extending their operational lifespan [35], [40], [41], [42]. The charging station (C) transmits radio frequency signals, serving as a controllable energy source to power all UAVs and enhance the overall quality of service (QoS). It is assumed that UAVs are equipped with energy harvesting capabilities to convert the received radio frequency signals from the charging station (C) into usable energy [41], [43].

In this work, energy harvesting management is also integrated as a key component, considering that the harvested energy primarily depends on two factors: the path loss between the charging station (C) and the UAVs, and the amount of energy transferred by the charging station (C) [29]. Under the assumption of a linear energy harvesting model [35], the energy harvested by the i th UAV at time slot t is given by $E_i^h(t) = h_{Ci}(t)\zeta_i e_i^h(t)$, where ζ_i represents the UAV's energy harvesting efficiency, $e_i^h(t)$ denotes the energy transferred from the charging station (C) to the i th UAV, and $h_{Ci}(t)$ is the channel gain between the charging station C and the i th transmitting UAV. The channel gain for wireless energy transfer is expressed as $h_{Ci}(t) = 10^{\frac{-\overline{PL}(d_{Ci})}{10}} |\Gamma_{Ci}(t)|^2$ where $\overline{PL}(d_{Ci})$, $\Gamma_{Ci}(t)$, and d_{Ci} represent the average path loss, the small scale fading, and the distance between the charging station (C) and the i th UAV, respectively.

In Fig. 1, we assume that the batteries for all UAVs can wirelessly recharge from the charging station (C). The battery status for the i th transmitting and j th jamming UAV is given by $b_{i/j}(t) = [b_1(t), b_2(t), \dots, b_{N_i}(t)/b_{N_j}(t)]$, where $b(t) \in [0, 1]$ represents the UAV's battery percentage at the beginning of each time slot t .

The flight energy consumption of the i th transmitting and j th jamming UAVs is defined to encompass propulsion energy, hovering energy, blade profile energy, induced

energy, and mean rotor-induced energy. Accordingly, the total flight energy consumption is expressed as [17], [44], [45]:

$$e_{i/j}^f(t) = \frac{P_f \|\mathbf{q}_{i/j}(t+1) - \mathbf{q}_{i/j}(t)\|}{v}, \quad (12)$$

where P_f is the flight power of the UAVs, v is the velocity or speed of the UAV, and $\|\mathbf{q}_{i/j}(t)\|$ defines the 3D coordinates of the i th transmitting and j th jamming UAVs at time slot t , such that $\mathbf{q}_{i/j}(t) = [x_{i/j}(t), y_{i/j}(t), z_{i/j}(t)]$. The 3D coordinates $\|\mathbf{q}_{i/j}(t)\|$ has the following constraints:

$$\|\mathbf{q}_{i/j}(t+1) - \mathbf{q}_{i/j}(t)\|^2 \leq L^2, \quad (13)$$

$$\|\mathbf{q}_{i/j}(t) - \mathbf{q}_{i/j}(t)\|^2 \geq d_{\min}^2, \quad (14)$$

where $L = \tau v_{\max}$ represents the maximum 3D Euclidean distance a UAV can cross in one time slot, with $v_{\max}$ denoting the UAV's velocity, τ represents the length of the time slots, and $d_{\min}$ represents the minimum required distance between two transmitting and jamming UAVs at all times.

Consequently, the battery energy level at each time slot t is determined by the energy harvested, the current battery status, the energy consumption for data transmission or jamming, and the energy consumed for flight by the transmitting or jamming UAVs. To determine the UAV's energy level in the next time slot t , we consider two cases;

- First, if the i th transmitting UAV is fully charged to $B_{\max}$, where $B_{\max}$ denotes the maximum battery capacity of the UAV, its battery energy level at time slot t is expressed as $B_{\max} - e_i^t(t) - e_i^f(t)$, whereas for the j th jamming UAV, the battery energy in this case is $B_{\max} - e_j^t(t) - e_j^f(t)$.
- Second, if the UAV's battery is not fully charged, then the energy level for the i th transmitting and j th jamming UAVs is given by $E_{i/j}^h(t) + b_{i/j}(t) - e_{i/j}^t(t) - e_{i/j}^f(t)$.

Accordingly, the battery level for the i th transmitting and j th jamming UAVs at the next time slot $t+1$ is given by:

$$\begin{aligned} b_i(t+1) &= \min\{E_i^h(t) + b_i(t) - e_i^t(t) - e_i^f(t), B_{\max}\}, \\ b_j(t+1) &= \min\{E_j^h(t) + b_j(t) - e_j^t(t) - e_j^f(t), B_{\max}\} \end{aligned} \quad (15)$$

The total energy consumed by the network encompasses all the energies discussed above. Now, we define the energy utilization efficiency as the ratio of total energy consumption to the associated transmit power. It serves as a power-normalized metric that reflects how effectively the available energy is utilized in the network under a given power budget. A higher energy utilization efficiency, indicates more favorable energy usage behavior. This makes it also suitable as a reward component in the DRL-based optimization framework, as discussed in the next section. Consequently,

the total energy utilization efficiency of the network is expressed as:

$$E_{tot}(t) = \sum_{i=1}^{N_t} \left(\frac{e_i^t(t) + e_i^f(t) + e_i^h(t)}{p_i(t)} \right) + \sum_{j=1}^{N_j} \left(\frac{e_j^t(t) + e_j^f(t) + e_j^h(t)}{p_j(t)} \right), \quad (16)$$

It is noteworthy that we adopted in this paper the term energy utilization efficiency, which differs from the classical definition of the energy efficiency. Traditionally, energy efficiency is defined as the ratio of secrecy rate to total energy consumption. However, our total utility function will be formulated as a weighted balance of key metrics, namely the secrecy rate, energy utilization efficiency, and interference, to allow flexible policy learning within the DRL framework. In particular, dividing energy consumption by power is used to normalize energy usage relative to transmit power. This allows us to capture the trade-off between energy consumed and the power budget within the DRL framework. This definition of energy utilization efficiency is intentionally designed as a task-specific utility term to guide the learning process of the DRL agent and promote convergence toward energy-aware policies.

The total interference of the network is the total interference from the jamming UAVs to all the legitimate transmitting UAVs (I_{ji}) as shown in Fig. 1, which is expressed as:

$$I_{tot}(t) = \sum_{i=1}^{N_t} \sum_{j=1}^{N_j} I_{ji}(t), \quad (17)$$

D. MOBILITY MODEL

In order to emulate dynamic network conditions, we adopt the Random Waypoint mobility model for the eavesdropper [46]. In this model, we assume that the eavesdropper

moves from one random waypoint to another within the network coverage. At each waypoint, the eavesdropper pauses for a certain time before moving to the next waypoint.

E. PROBLEM FORMULATION

The main goal of this paper is to jointly determine the optimal 3D coordinates (x, y, and z), power allocations, and harvesting energy for the transmitting and jamming UAVs to maximize the overall utility of the UAV network shown in Fig. 1, which is defined in terms of total secrecy rate, energy utilization efficiency, and interference. In particular, we aim to maximize a multi-objective utility function that balances the following aspects:

- Secrecy rate: Maximize the difference between the legitimate transmission rate and the rate intercepted by the eavesdropper, as expressed in (9).
- Energy utilization efficiency: Maximize the energy usage of UAVs supported by energy harvesting, as expressed in (16).
- Interference mitigation: Minimize the interference caused by jamming to legitimate communication, as expressed in (17).

Consequently, our optimization problem is expressed as shown in (18), where constraint C_1 ensures that the downlink data rate from all transmitting UAVs exceeds the minimum required secrecy rate of the network R_{min}^{sec} . C_2 is the jamming constraint that ensures the jamming UAVs disrupt the signals intercepted by eavesdroppers by minimizing their received SNR below their minimum required SNR (i.e., sensitivity) Γ_e^{min} . C_3 ensures that the interference from the jamming to transmitting UAVs is below a certain threshold I_i^{max} . C_4 ensures that the battery levels of transmitting and jamming UAVs at the next time slot are chosen to be the minimum of the harvested energies and the maximum energy budget of the UAVs. C_5 shows that the efficiency of the batteries of the transmitting and jamming UAVs must always be positive. C_6 ensures that the maximum

$$\begin{aligned} P1 : \quad & \underset{\mathbf{Q}, \mathbf{p}, \mathbf{e}^h}{\text{maximize}} \sum_{t=1}^T U(t) = \sum_{t=1}^T \left(\underbrace{w_1 R_{tot}^{sec}(t)}_{\text{Total secrecy rate}} + \underbrace{w_2 E_{tot}(t)}_{\text{Total energy utilization efficiency}} - \underbrace{w_3 I_{tot}(t)}_{\text{Total interference}} \right) \\ \text{s.t.} \quad & C_1 : R_i^{sec}(t) \geq R_{min}^{sec}, \quad \forall i \in N_t \\ & C_2 : \Gamma_{ie}(t) < \Gamma_e^{min}, \quad \forall i \in N_t, \forall e \in N_e \\ & C_3 : I_{ji}(t) < I_i^{max}, \quad \forall i \in N_t, \forall e \in N_e \\ & C_4 : b_{i/j}(t+1) = \min(b_{i/j}(t) + E_{i/j}^h(t) - e_{i/j}^t(t) - e_{i/j}^f(t), B^{max}), \quad \forall i \in N_t, \quad \forall j \in N_j \\ & C_5 : b_{i/j}(t) + E_{i/j}^h(t) - e_{i/j}^t(t) - e_{i/j}^f(t) \geq 0, \quad \forall i \in N_t, \forall j \in N_j \\ & C_6 : \|\mathbf{q}_{i/j}(t) - \mathbf{q}_{i/j}(t-1)\|^2 \leq L^2, \quad \forall i \in N_t, \forall j \in N_j, \forall t \in N \\ & C_7 : \|\mathbf{q}_u(t) - \mathbf{q}_v(t)\|^2 \geq d_{min}^2, \quad u, v \in (i, j), u \neq v, \forall i \in N_t, \forall j \in N_j, \forall t \in N \\ & C_8 : x_{min} \leq x_{i/j}(t) \leq x_{max}, y_{min} \leq y_{i/j}(t) \leq y_{max}, z_{min} \leq z_{i/j}(t) \leq z_{max}, \quad \forall i \in N_t, \forall j \in N_j \\ & C_9 : 0 \leq p_{i/j}(t) \leq P^{max}, \quad \forall i \in N_t, \forall j \in N_j. \end{aligned} \quad (18)$$

distance the transmitting and jamming UAVs can cross in one time slot is within a pre-defined limit, where $L = \tau v_{max}$. C_7 ensures that the minimum distances between any two different transmitting and jamming UAVs at all times are kept within a predefined limit d_{min} . C_8 guarantees that 3D coordinates of the transmitting and jamming UAVs are within the allowable spatial coverage. C_9 guarantees that the transmit power of transmitting and jamming UAVs do not violate their allowable power budgets. Also, w_1 , w_2 , and w_3 in the utility function $U(t)$ are weighting factors for the secrecy rate, energy utilization efficiency, and interference mitigation, respectively, such that $w_1 + w_2 + w_3 = 1$.

F. WHY DRL TO SOLVE PROBLEM P1?

Problem **P1** is inherently non-convex, primarily due to the non-concavity of the objective function and constraints $C1$ and $C2$. While an optimal solution to **P1** could theoretically be achieved through exhaustive search methods using classical mathematical techniques [47], this becomes infeasible as the search space expands exponentially. This growth is driven by the high dimensionality, large-scale nature, and continuous variables within the underlying network depicted in Fig. 1. Consequently, standard methods are inadequate for solving such complex non-convex optimization problems [8]. Note that compared to the problem formulated in previous works, such as [28], we enhanced the objective function in **P1** and incorporated optimization variables for both transmitting and jamming UAVs. Additionally, we introduced a mobility model for the eavesdropper. These enhancements made the problem more realistic but significantly increased its complexity. Therefore, we propose a PPO model to address **P1** and effectively learn the network dynamics, as detailed in Section III. In particular, DRL (specifically PPO) offers key advantages: it handles continuous state and action spaces, learns optimal policies without requiring full model knowledge, and adapts to dynamic environments through experience. These capabilities make DRL particularly suitable for capturing the complex and stochastic nature of UAV networks, as it enables joint optimization over time while learning environment dynamics [6], [8], [48].

III. PPO FOR JOINT TRAJECTORY DESIGN AND RESOURCE ALLOCATION FOR UAV NETWORKS

In this section, we demonstrate how to implement the DRL technique to solve **P1** and learn the dynamics of the UAV network shown in Fig. 1. We first formulate the problem as a Markov decision process (MDP) represented by the tuple $\{S, A, p, r\}$, where S , A , p , and r represent the state space, action space, probability transition, and reward function, respectively [8], [49]. We consider the UAV network as the environment and the ground control center (G) as our acting agent that observes the environment and takes appropriate actions based on the current state of the environment. Subsequently, a reward signal will be produced according to the actions executed by the agent within the UAV environment to produce the next network

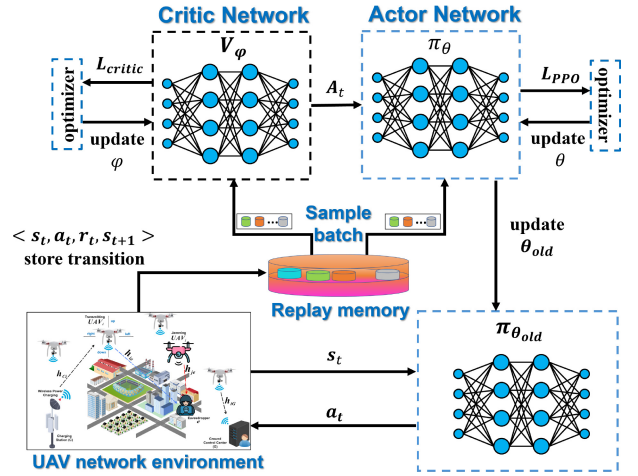


FIGURE 2. Architecture of our proposed PPO model.

state. We define these elements as follows. Fig. 2 shows the architecture of our proposed PPO algorithm used to solve **P1**. Note that we have adopted a single-agent approach in this work, where the ground control center (G) acts as the central agent. However, a MADRL approach remains a valid alternative, depending on how the system model in Fig. 1 is interpreted. We believe that extending this work to a MADRL framework represents a promising direction for future research in this context.

1) STATE SPACE

The state space is designed to contain effective information about the UAV networks to enable our agent to learn the optimal policy. The ground agent will be trained considering that only limited information about the dynamics of the UAV is available. Specifically, assume the agent only knows historical data on achieved rates and battery levels. Note that the equations of data rate discussed in Section (II) still include the channel gain terms, which are crucial for the agent's learning process. At time step t , the agent's state space comprises four components; the vector of downlink secrecy data rate from all transmitting UAVs to the ground center at the previous time slot $R^{sec}(t-1) \in \mathbb{C}^{1 \times 1}$, vector of battery levels of transmitting and jamming UAVs at the previous time slot $\mathbf{b}(t-1) \in \mathbb{C}^{(N_t+N_j) \times 1}$, vector of energy consumed by transmitting and jamming UAVs at the previous time slot $E^H(t-1) \in \mathbb{C}^{(N_t+N_j) \times 1}$, and vector of minimum secrecy rate requirements of transmitting UAVs $R_{min}^{sec} \in \mathbb{C}^{1 \times 1}$. The size of the state space is a vector of length $3N_t + 2N_j + 1$. Consequently, the state space is expressed as:

$$S(t) = \left[R^{sec}(t-1) \in \mathbb{C}^{N_t \times 1}, \mathbf{b}(t-1) \in \mathbb{C}^{(N_t+N_j) \times 1}, \right. \\ \left. E^H(t-1) \in \mathbb{C}^{(N_t+N_j) \times 1}, R_{min}^{sec} \in \mathbb{C}^{1 \times 1} \right], \quad (19)$$

2) ACTION SPACE

In our proposed system model shown in Fig. 1, the actions taken by the ground control center are designed to maximize

the objective function of **P1** by jointly controlling the optimization variables. Hence, the actions of the ground center agent at each time step are the matrix of the 3D coordinates of all transmitting and jamming UAVs $\mathbf{Q}(t) \in \mathbb{C}^{(N_t+N_j) \times 3}$, vector of power allocations of transmitting and jamming UAVs $\mathbf{p}(t) \in \mathbb{C}^{(N_t+N_j) \times 1}$, and vector of the amount of energy transferred from the charging station to all transmitting and jamming UAVs $\mathbf{e}^h(t) \in \mathbb{C}^{(N_t+N_j) \times 1}$. The size of the action space is a vector of length $5(N_t + N_j)$. Consequently, the action space is expressed as:

$$\mathcal{A}(t) = \left[\mathbf{Q}(t) \in \mathbb{C}^{(N_t+N_j) \times 3}, \mathbf{p}(t) \in \mathbb{C}^{(N_t+N_j) \times 1}, \mathbf{e}^h(t) \in \mathbb{C}^{(N_t+N_j) \times 1} \right], \quad (20)$$

3) REWARD FUNCTION

The objective of our optimization problem in **P1** is to maximize a multi-objective utility function in terms of secrecy rate, energy utilization efficiency, and interference mitigation. The agent's reward at each time step is calculated based on observing the current state of the UAV network as given by (19) and executing the corresponding action as given by (20). Hence, the instantaneous reward function is expressed by:

$$r(t) = \sum_{i=1}^T \left(w_1 R_{tot}^{sec}(t) + w_2 E_{tot}(t) - w_3 I_{tot}(t) \right), \quad (21)$$

The utility maximization problem can be formulated based on the agent's instantaneous reward. The expected cumulative discounted reward throughout the time horizon T is expressed as $\mathcal{R} = \mathbb{E}[\sum_{t=1}^T \gamma^{t-1} r(t)]$, where $0 \leq \gamma \leq 1$ is the discount factor [6], [8]. Our objective is to derive the optimal decision policy, π^* , which determines the UAVs' 3D locations, power allocation, and harvested energy to maximize the ground control center agent's long-term reward. This is expressed as $\pi^* = \underset{\pi}{\operatorname{argmax}} \mathcal{R} = \underset{\pi}{\operatorname{argmax}} \sum_{t=1}^T \gamma^{t-1} r(s(t), a(t))$.

A. THE PPO ALGORITHM

As highlighted in Section II-F, the optimization problem in **P1** is computationally challenging and unsuitable for traditional solution methods [6], [8]. To address this, we propose employing the PPO algorithm to derive and learn the optimal policy, π^* . The PPO method is a robust and efficient DRL approach specifically designed to handle problems characterized by high dimensionality and continuous state-action spaces [50].

The PPO algorithm is a policy gradient approach based on the actor-critic framework [50], [51], and it represents an enhanced version of the trust region policy optimization (TRPO) algorithm. It relies on the use of the importance sampling concept to quantify the difference between policies before and after updates, thereby regulating policy adjustments and ensuring the updated policy remains within the

trust region of the original. The desired policy (π^*) is derived by training the DNNs to reduce the following Bellman loss equation:

$$L(\theta) = \mathbb{E}_t[\log \pi_\theta(a_t | s_t) A_t], \quad (22)$$

where θ is the training weights of the actor and critic networks, and A_t is an estimator of the advantage function at step t . TRPO represents the ratio of probabilities between the current policy and the previous one as $\psi(\theta) = \frac{\pi_\theta(a_t | s_t)}{\pi_{\theta_{old}}(a_t | s_t)}$. An alternative objective function is defined as [50], [51]:

$$L_{TRPO}(\theta) = \mathbb{E}_t[\psi(\theta) A_t]. \quad (23)$$

However, maximizing $L_{TRPO}(\theta)$ can result in an overly large policy update. To address this, $L_{PPO}(\theta)$ was introduced to adjust the objective function and penalize policy changes that cause $\psi(\theta)$ to deviate from 1. It is defined as [50]:

$$L_{PPO}(\theta) = \mathbb{E}_t[\min(\psi(\theta) A_t, \text{clip}(\psi(\theta), 1 - \varepsilon, 1 + \varepsilon) A_t)], \quad (24)$$

where ε is the policy clip hyperparameter, the $\text{clip}()$ function is used to guarantee that the values of $\psi(\theta)$ remain within $(1 - \varepsilon, 1 + \varepsilon)$ in order to mitigate excessive policy updates. The actor network optimizes its policy parameter θ by minimizing the loss function $L_{PPO}(\theta)$. A truncated version of generalized advantage estimation to compute A_t is used as, which is expressed as [50]:

$$A_t = \delta_t + \gamma \delta_{t+1} + \dots + \gamma^{T-t+1} \delta_{T-1} \quad (25)$$

where $\delta_t = r_t + \gamma V_\varphi(s_{t+1}) - V_\varphi(s_t)$,

where $V_\varphi(s_t)$ is the state-value function at time slot t . The critic network is updated using the empirical average A_t for each sample batch. The loss function for the critic network is expressed as:

$$L_{critic}(\varphi) = \frac{\sum_{t=1}^T \sqrt{A_t}}{T}. \quad (26)$$

This loss function is based on the standard formulation used in PPO implementations to train the critic network [50], where the goal is to minimize the error between predicted and actual value estimates. In particular, the use of the advantage estimator A_t ensures that the critic learns accurate state-value predictions to support stable policy updates in the actor network. This formulation is designed to improve learning efficiency and convergence stability. We refer the interested reader to the original paper [50] for more details on the PPO algorithm.

The pseudo-code flow of the proposed PPO model designed to solve our optimization problem **P1** is outlined in Algorithm 1 along with Fig. 2. Lines 1 to 3 involve initializing the agent's hyperparameters, including the state space, DNN weights, and replay memory $\mathcal{D}$. In Lines 4 to 10, the agent interacts with the UAV environment by observing states, performing actions, and receiving next states and numerical rewards, which are then stored in $\mathcal{D}$. Lines 11

Algorithm 1 PPO for Joint Trajectory Design and Resource Allocation in UAV Networks

Input: $N_t, N_j, N_e, R_{min}^{sec}$.

Output: Optimal $\mathbf{Q}, \mathbf{p}, \mathbf{e}^h$.

```

1: Initialization: Set  $t = 0$  and initialize states of the agent  $s_0$  with
   initial observations.
2: Randomly initialize weights of the actor and critic networks  $\theta$ .
3: Initialize replay memory  $\mathcal{D}$ .
4: for episode = 1 to  $T$  do
5:   Reset the state of the agent  $s(t)$ .
6:   for  $t = 1$  to  $M$  do
7:     Observe state  $s(t)$  from the environment using (19).
8:     Execute action  $a(t)$  using (20) according to  $\pi_{\theta_{old}}$ .
9:     Receive reward using (21), then observe next states  $s(t+1)$ .
10:    Store transitions  $\langle s(t), a(t), r(t), s(t+1) \rangle$  in  $\mathcal{D}$ .
11:    if  $\text{len}(\text{batch}) \geq \text{batch size}$  then
12:      Calculate the advantage  $A_t$  using (25).
13:      Sample random mini-batch of transitions from  $\mathcal{D}$ .
14:      Update the weights  $\theta$  of the actor and critic networks by
        minimizing the losses  $L_{ppo}(\theta)$  and  $L_{critic}(\varphi)$  using (24) and
        (26).
15:    end if
16:    Update weights of target networks  $\theta_{old} \leftarrow \theta$ .
17:  end for
18: end for

```

to 16 detail the learning phase, where the ground agent samples mini-batch transitions from the stored data to refine its policy. This iterative process continues until the agent converges to the optimal global policy for solving **P1**.

B. COMPUTATIONAL COMPLEXITY OF THE PPO MODEL

The computational complexity of our proposed PPO algorithm, detailed in Algorithm 1 and illustrated in Fig. 2, is analyzed as follows. PPO is typically implemented using the Actor-Critic framework [50], with complexity determined by the number of multiplications performed in each episode. Since the employed DNNs consist of fully connected layers, the computational complexity is given by $\mathcal{O}(\sum_{k=0}^{K-1} C_k^{\text{in}} C_{k+1}^{\text{out}})$ where C_k^{in} and C_k^{out} represent the number of input and output neurons in the k th hidden layer, and K denotes the total hidden layers. In our work, both the actor and critic networks utilize identical fully connected DNNs with K layers, leading to a complexity of $\mathcal{O}(2 \sum_{k=0}^{K-1} C_k^{\text{in}} C_{k+1}^{\text{out}})$. As outlined in Algorithm 1, T training episodes and M time slots are considered. Hence, the overall complexity order of our proposed PPO model is expressed as $\mathcal{O}(2TM \sum_{k=0}^{K-1} C_k^{\text{in}} C_{k+1}^{\text{out}})$.

IV. PERFORMANCE EVALUATION

This section outlines the simulation setup and the corresponding numerical results used to evaluate the performance of the proposed PPO-based DRL algorithm, detailed in 1, for solving the network optimization problem **P1**. We first describe the simulated UAV network's specifications and the architecture of the PPO model. Next, the numerical results are presented and analyzed. Furthermore, we consider a scenario where the PPO algorithm operates online, showcasing its adaptivity to dynamic and abrupt network changes, such as eavesdropper mobility. The proposed PPO algorithm is compared against three well-known benchmarks:

TABLE 2. Network's simulation parameters.

Notation	Value	Notation	Value
N_t	10	N_j	2
N_e	1	v_{max}	4 m/s
N_0, N'_0	-174 dBm/Hz	L	5 m
l_i/l_j	[100, 200]Mbps	n_{LoS}	2
h_{min}	10 m	n_{NLoS}	3
h_{max}	70 m	P_{max}	30dBm
a	0.28	ζ_i/ζ_j	0.5
β	9.6	B^{max}	500J
f_c	2.4 GHz	λ_{LoS}	1 dB
Bandwidth	20 MHz	λ_{NLoS}	20 dB

- A DRL-based method, specifically the DDPG algorithm similar to [28],
- Another DRL-based method, which is the policy-based Asynchronous Advantage Actor-Critic (A3C), similar to [52], [53], and
- A conventional method, namely the genetic algorithm (GA), similar to [54], [55], [56].

A. SIMULATION SETUP

We simulate the UAV network scenario depicted in Fig. 1, which consists of ten transmitting UAVs, two jamming UAVs, and one ground eavesdropper, all operating within a 200-meter-diameter area. This setup provides a proof of concept and can be scaled to accommodate more UAVs. The speeds of the transmitting and jamming UAVs are controlled by commands issued from the ground center (C). The remaining simulation parameters, largely inspired by [28], [45], [57], [58], [59], are detailed in Table 2. Additionally, a single agent is simulated using the PPO DRL algorithm and implemented in Python with the PyTorch library. Tanh activation functions, coupled with Batch-normalization layers, are employed in all DNNs, and the Adam optimizer is utilized to update the weights of the actor and critic networks [60]. The remaining model hyperparameters are outlined in Table 3.

B. NUMERICAL RESULTS

This section presents and analyzes simulation results to validate the functionality and evaluate the accuracy of the proposed PPO algorithm compared to the other three benchmarks. It is important to note that the GA algorithm often fails to converge to a feasible solution within a reasonable time, so it will only be used as a benchmark when suitable.

Fig. 3 illustrates the reward training convergence of our PPO agent. The figure highlights how the agent effectively learns a global policy to address the optimization problem **P1**. The agent achieves rapid convergence to the optimal policy after just 84 episodes.

Fig. 4 illustrates the total utility achieved by our proposed PPO algorithm compared to the DDPG, A3C, and GA

TABLE 3. Hyperparameters of the PPO model.

Parameter	Value
Batch size	64
Horizon (N)	20
Learning rate	actor & critic $= 1 \times 10^{-3}$
γ	0.99
GAE parameter (λ)	0.95
No. epochs	2
Number of hidden layers	actor & critic = 3
Number of neurons	(64, 64, 64)
Policy clip	0.2
Optimizer	Adam

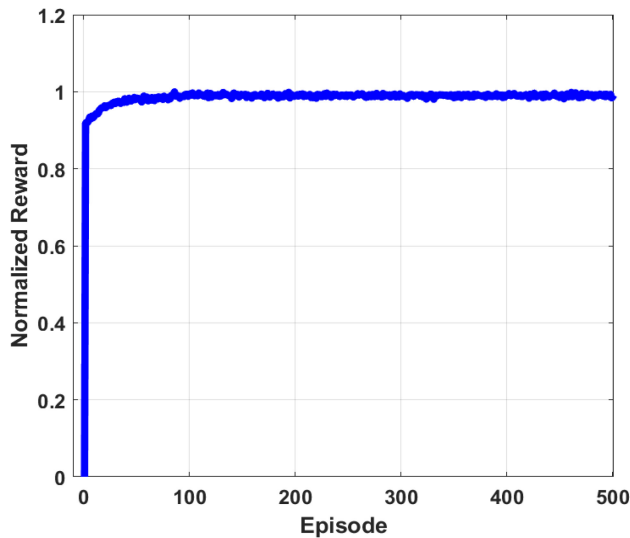


FIGURE 3. Reward convergence of our proposed PPO algorithm. The agent rapidly converges to the optimal policy after just 84 episodes.

algorithms. PPO, A3C, and DDPG algorithms converge to the optimal policy that solves P1 within a good timeframe, but the PPO algorithm achieves higher utility values. Specifically, after convergence, the PPO algorithm outperforms the DDPG and A3C algorithms by a utility values gap of 1.4 and 3.3, respectively. Furthermore, the GA algorithm converges much slower than PPO, A3C, and DDPG algorithms, with lower utility values and reduced stability. Notably, the GA algorithm begins converging after 126 episodes, exhibiting instability due to rapid fluctuations. The utility values achieved by the PPO algorithm are approximately 20.7 higher than those obtained by the GA algorithm. These results clearly demonstrate the superiority of the PPO algorithm over the DDPG, A3C, and GA algorithms in terms of both performance and stability.

In the optimization problem **P1**, the objective is to jointly maximize the total secrecy rate and energy utilization

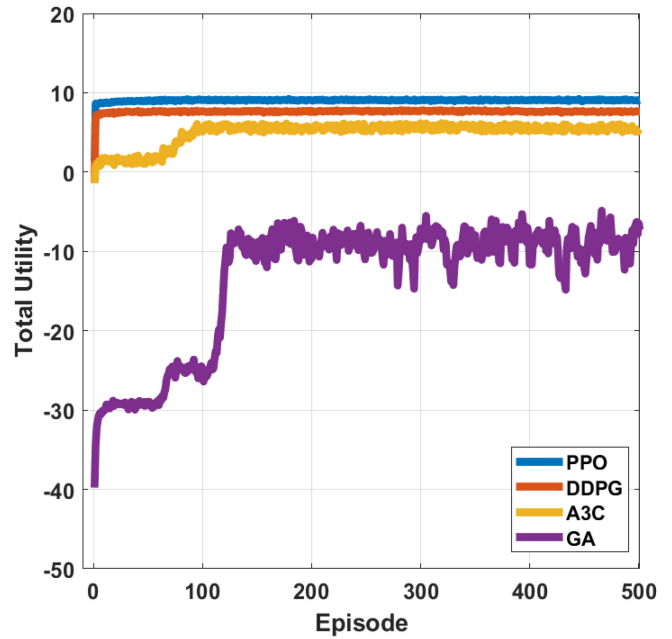


FIGURE 4. Utility of our proposed PPO algorithm compared to the DDPG, A3C, and GA solutions. The PPO solution converges faster than the two benchmarks and achieves higher utility values.

efficiency while minimizing overall interference. Figs 5, 6, and 7 depict the total normalized secrecy rate, energy utilization efficiency, and interference, respectively, for our proposed PPO algorithm in comparison to the DDPG and A3C algorithms. It is important to note that the GA, as demonstrated earlier, fails to converge to feasible solutions within a reasonable timeframe and has thus been excluded from these comparisons.

As shown in Fig. 5, the proposed PPO algorithm achieves faster convergence and higher total secrecy rate values compared to the DDPG and A3C algorithms. Specifically, the PPO algorithm provides an average secrecy rate enhancement of 5.78% and 13.96% over the DDPG and A3C algorithms, respectively, across the last 100 episodes. Similarly, Fig. 6 highlights that the PPO algorithm not only converges more rapidly but also achieves superior energy utilization efficiency, with an average improvements of 15.05% and 20.78% over the DDPG and A3C algorithms, respectively, in the last 100 episodes. Lastly, Fig. 7 demonstrates the effectiveness of the PPO algorithm in significantly reducing total network interference. The PPO algorithm exhibits faster convergence, lower interference values, and greater stability than the DDPG and A3C algorithms. Over the last 100 episodes, the PPO algorithm achieves a remarkable 94.55% and 96.73% reduction in interference compared to the DDPG and A3C algorithms, respectively.

C. MODEL ADAPTIVITY IN ONLINE DEPLOYMENT

To demonstrate the robustness of the proposed PPO model in an online setting, we evaluate its adaptability to sudden changes in the location of the ground eavesdropper. To

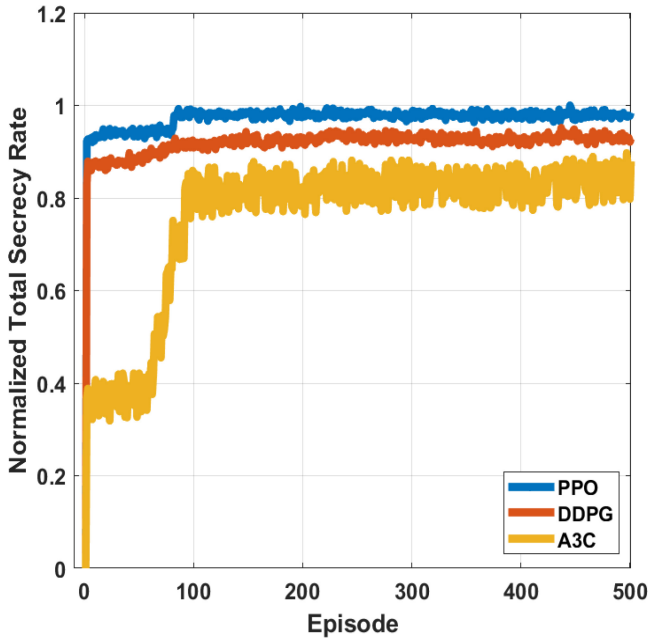


FIGURE 5. Total secrecy rate of our proposed PPO algorithm compared to the DDPG and A3C. The PPO algorithm converges faster with higher values.

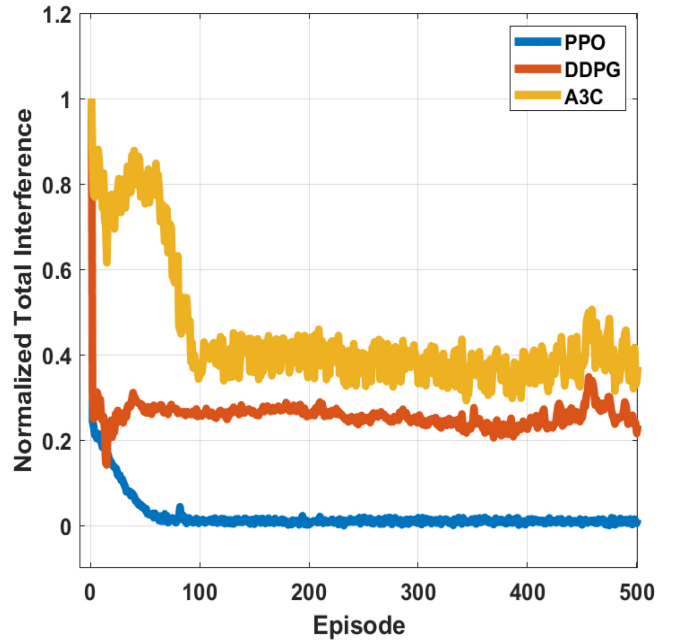


FIGURE 7. Total interference of our proposed PPO algorithm compared to the DDPG and A3C. The PPO algorithm converges faster with lower values and more stability.

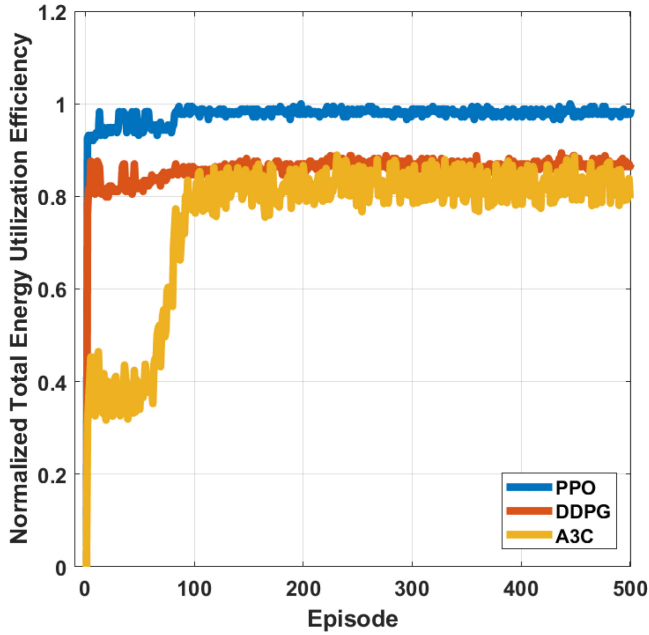


FIGURE 6. Total energy utilization efficiency of our proposed PPO algorithm compared to the DDPG and A3C. The PPO algorithm converges faster with higher values.

simulate realistic dynamic network conditions, we use the Random Waypoint Mobility Model [46] discussed in Section II-D, where the ground eavesdropper changes position every 500 episodes. The convergence speed is defined as the number of episodes required to reach a steady state, where the total utility remains constant over the last 200 episodes, similar to [56]. Fig. 8 shows the mobility pattern of the ground eavesdropper within the designated area, while

Fig. 9 presents the normalized total utility, i.e., the objective function of problem **P1**, convergence results for the PPO model compared to the two established benchmarks. Two key observations emerge from the results. First, the PPO algorithm adapts rapidly to changes in the mobility of the eavesdropper and dynamically finds feasible solutions to problem **P1**. This behavior is evident both during the initial episodes and after the eavesdropper relocates at episode 500. During the initial training phase, the PPO model takes approximately 39 episodes to converge to a steady state. However, after the eavesdropper updates its location, the PPO algorithm re-converges and reaches a steady state by episode 514. This indicates that the convergence speed of the PPO algorithm, after training and learning the optimal policy, is approximately 2.8 times faster than during initial training. This trend persists in later episodes (e.g., 1000, 1500, 2000, etc.), where the PPO model leverages its accumulated learning experience to adapt quickly and re-solve our optimization problem. Second, Fig. 9 demonstrates that the performance of our proposed PPO algorithm surpasses the DDPG, A3C, and GA. Specifically, the PPO algorithm achieves higher utility values compared to the DDPG, A3C, and GA solutions. Additionally, the PPO algorithm exhibits faster convergence speeds than the DDPG, A3C, and GA algorithms. This is particularly evident in episodes 1000 to 1059 and 2000 to 2066, where the DDPG and A3C algorithms require more time to converge. Furthermore, as previously concluded, the conventional GA algorithm consistently fails to converge to a feasible solution within the 500-episode time frame. These results highlight the dynamic adaptability and robustness of our proposed PPO algorithm in handling abrupt changes within the UAV network.

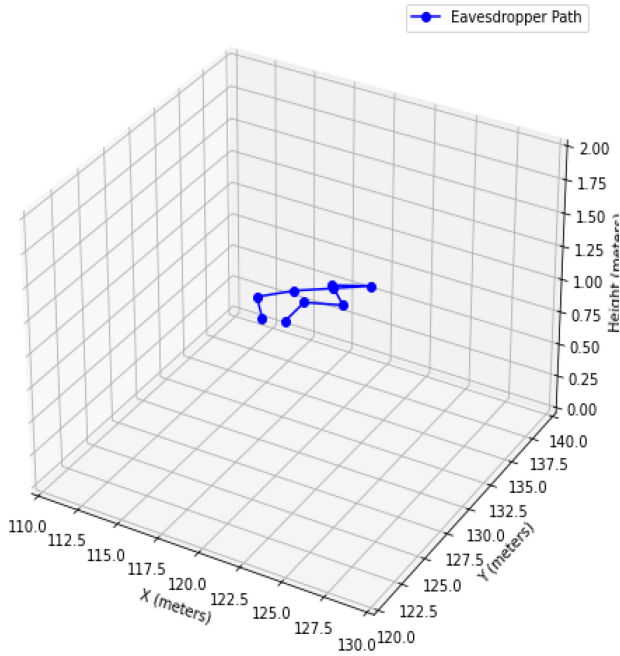


FIGURE 8. Mobility pattern of the eavesdropper based on the Random Waypoint Mobility Model.

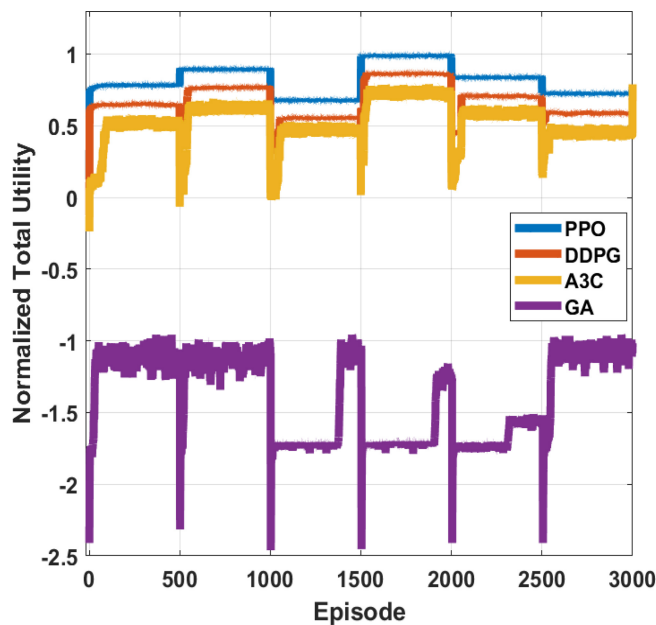


FIGURE 9. Adaptivity of our proposed PPO algorithm to the dynamic mobility and environment changes. The eavesdropper randomly moves every 500 episodes, and the PPO algorithm quickly adapts to these abrupt changes. Our proposed PPO algorithm converges faster with higher utility values than the DDPG, A3C, and GA benchmarks.

V. CONCLUSION

In this paper, we proposed a DRL model based on the PPO algorithm to address the problem of multi-objective utility maximization for a UAV network in the presence of ground eavesdroppers and energy-harvesting capabilities. The optimization problem was first formulated as a non-convex problem with the objective of jointly maximizing

the network's secrecy rate and energy utilization efficiency while mitigating total interference by controlling the 3D positions of UAVs, power allocation, and harvested energy. Given the high dimensionality and non-convex nature of the problem, coupled with the limitations of traditional methods in solving it efficiently, we proposed a PPO-based DRL algorithm to tackle the problem and learn network dynamics. Through simulation results, we demonstrated that our proposed algorithm rapidly converges to the optimal policy, achieving higher utility values when compared to the DDPG- and A3C-based DRL algorithms and the GA algorithm. As part of future work, we aim to extend the optimization problem to include the optimal selection of UAVs that can function as jamming UAVs. Also, to consider more complex scenarios, such as multiple ground eavesdroppers and distributed cooperation among multiple UAVs.

REFERENCES

- [1] Y. Bai, H. Zhao, X. Zhang, Z. Chang, R. Jäntti, and K. Yang, "Toward autonomous multi-UAV wireless network: A survey of reinforcement learning-based approaches," *IEEE Commun. Surveys Tuts.*, vol. 25, no. 4, pp. 3038–3067, 4th Quart., 2023.
- [2] S. Hayat, E. Yanmaz, and R. Muzaffar, "Survey on unmanned aerial vehicle networks for civil applications: A communications viewpoint," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 4, pp. 2624–2661, 4th Quart., 2016.
- [3] L. Gupta, R. Jain, and G. Vaszkun, "Survey of important issues in UAV communication networks," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 2, pp. 1123–1152, 2nd Quart., 2016.
- [4] A. Fotouhi et al., "Survey on UAV cellular communications: Practical aspects, standardization advancements, regulation, and security challenges," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 4, pp. 3417–3442, 4th Quart., 2019.
- [5] Z. Yu, Z. Wang, J. Yu, D. Liu, H. H. Song, and Z. Li, "Cybersecurity of unmanned aerial vehicles: A survey," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 39, no. 9, pp. 182–215, Sep. 2024.
- [6] N. C. Luong et al., "Applications of deep reinforcement learning in communications and networking: A survey," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 4, pp. 3133–3174, 4th Quart., 2019.
- [7] P. Yadav, A. Mishra, J. Lee, and S. Kim, "A survey on deep reinforcement learning-based approaches for adaptation and generalization," 2022, *arXiv:2202.08444*.
- [8] A. Alwarafy, M. Abdallah, B. S. Çiftler, A. Al-Fuqaha, and M. Hamdi, "The frontiers of deep reinforcement learning for resource management in future wireless HetNets: Techniques, challenges, and research directions," *IEEE Open J. Commun. Soc.*, vol. 3, pp. 322–365, 2022.
- [9] H. Yang, Z. Xiong, J. Zhao, D. Niyato, L. Xiao, and Q. Wu, "Deep reinforcement learning-based intelligent reflecting surface for secure wireless communications," *IEEE Trans. Wireless Commun.*, vol. 20, no. 1, pp. 375–388, Jan. 2021.
- [10] Z. Zhang and L. Li, "Distributed RISs-aided secure wireless network under practical phase shift model: A deep reinforcement learning approach," in *Proc. 14th Int. Conf. Wireless Commun. Signal Process. (WCSP)*, 2022, pp. 1–5.
- [11] Q. Liu, Y. Zhu, M. Li, R. Liu, Y. Liu, and Z. Lu, "DRL-based secrecy rate optimization for RIS-assisted secure ISAC systems," *IEEE Trans. Veh. Technol.*, vol. 72, no. 12, pp. 16871–16875, Dec. 2023.
- [12] H. Sharma, N. Kumar, I. Budhiraja, and A. Barnawi, "Secrecy rate maximization in THz-aided heterogeneous networks: A deep reinforcement learning approach," *IEEE Trans. Veh. Technol.*, vol. 72, no. 10, pp. 13490–13505, Oct. 2023.
- [13] R. Saleem, W. Ni, M. Ikram, and A. Jamalipour, "Deep-reinforcement-learning-driven secrecy design for intelligent-reflecting-surface-based 6G-IoT networks," *IEEE Internet Things J.*, vol. 10, no. 10, pp. 8812–8824, May 2023.

- [14] H. Sharma, N. Kumar, and R. K. Tekchandani, "SecBoost: Secrecy-aware deep reinforcement learning based energy-efficient scheme for 5G HetNets," *IEEE Trans. Mobile Comput.*, vol. 23, no. 2, pp. 1401–1415, Feb. 2024.
- [15] A. Iqbal, A. Al-Habashna, G. Wainer, F. Bouali, G. Boudreau, and K. Wali, "Deep reinforcement learning-based resource allocation for secure RIS-aided UAV communication," in *Proc. IEEE 98th Veh. Technol. Conf. (VTC-Fall)*, 2023, pp. 1–6.
- [16] X. He, X. Li, H. Ji, and H. Zhang, "Resource allocation for secrecy rate optimization in UAV-assisted cognitive radio network," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, 2021, pp. 1–6.
- [17] L. Jing, X. Jia, Y. Lv, and N. Wan, "Maximizing the average secrecy rate for UAV-assisted MEC: A DRL method," in *Proc. IEEE 5th Adv. Inf. Technol., Electron. Autom. Control Conf. (IAEAC)*, 2021, pp. 2514–2518.
- [18] D. Tsiakmakis, N. Passalis, and A. Tefas, "Improving inertial-based UAV Localization using data-efficient deep reinforcement learning," in *Proc. 31st Eur. Signal Process. Conf. (EUSIPCO)*, 2023, pp. 1355–1359.
- [19] C. Zhang, Y. Liu, and Z. Zhang, "A deep reinforcement learning approach for federated learning optimization with UAV trajectory planning," in *Proc. IEEE 34th Annu. Int. Symp. Pers., Indoor Mobile Radio Commun. (PIMRC)*, 2023, pp. 1–7.
- [20] M. Shao, J. Yan, and X. Zhao, "Secrecy rate maximization by cooperative jamming for UAV-enabled relay system with mobile nodes," *IEEE Internet Things J.*, vol. 10, no. 15, pp. 13168–13180, Aug. 2023.
- [21] B. Li, Z. Fei, Y. Dai, and Y. Zhang, "Secrecy-optimized resource allocation for UAV-assisted relaying networks," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, 2018, pp. 1–6.
- [22] A. B. Adam and M. A. Elhassan, "Enhancing secrecy in UAV RSMA networks: Deep unfolding meets deep reinforcement learning," in *Proc. Int. Conf. Commun., Comput., Cybersecurity, Inform. (CCCI)*, 2023, pp. 1–5.
- [23] Y. You, R. Zhao, and H. Sun, "Deep reinforcement learning-based trajectory planning for secure UAV communication," in *Proc. 4th Int. Conf. Inf. Commun. Signal Process. (ICICSP)*, 2021, pp. 528–532.
- [24] C. Wen, Y. Fang, and L. Qiu, "Securing UAV communication based on multi-agent deep reinforcement learning in the presence of smart UAV eavesdropper," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, 2022, pp. 1164–1169.
- [25] T.-Y. Kim, J. Lee, and J.-H. Kim, "Deep reinforcement learning-based full-duplex communication UAV base station trajectory optimization in disaster environments," in *Proc. VTS Asia Pacific Wireless Commun. Symp. (APWCS)*, 2023, pp. 1–5.
- [26] H. Yang, S. Liu, L. Xiao, Y. Zhang, Z. Xiong, and W. Zhuang, "Learning-based reliable and secure transmission for UAV-RIS-assisted communication systems," *IEEE Trans. Wireless Commun.*, vol. 23, no. 7, pp. 6954–6967, Jul. 2024.
- [27] Y. Zhang, Z. Mou, F. Gao, J. Jiang, R. Ding, and Z. Han, "UAV-enabled secure communications by multi-agent deep reinforcement learning," *IEEE Trans. Veh. Technol.*, vol. 69, no. 10, pp. 11599–11611, Oct. 2020.
- [28] N. Abouelenen, A. Alwarafy, and M. Abdallah, "Secure and energy-efficient communication for Internet of Drones networks: A deep reinforcement learning approach," in *Proc. Int. Wireless Commun. Mobile Comput. (IWCMC)*, 2023, pp. 818–823.
- [29] T. D. P. Perera, D. N. K. Jayakody, S. K. Sharma, S. Chatzinotas, and J. Li, "Simultaneous wireless information and power transfer (SWIPT): Recent advances and future challenges," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 1, pp. 264–302, 1st Quart., 2018.
- [30] J. Yao and N. Ansari, "QoS-aware power control in Internet of Drones for data collection service," *IEEE Trans. Veh. Technol.*, vol. 68, no. 7, pp. 6649–6656, Jul. 2019.
- [31] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, "Drone small cells in the clouds: Design, deployment and performance analysis," in *Proc. IEEE global Commun. Conf. (GLOBECOM)*, 2015, pp. 1–6.
- [32] A. Al-Hourani, S. Kandeepan, and S. Lardner, "Optimal LAP altitude for maximum coverage," *IEEE Wireless Commun. Lett.*, vol. 3, no. 6, pp. 569–572, Dec. 2014.
- [33] A. Alwarafy, M. Alreshedi, A. F. Abas, and A. Alsanie, "Performance evaluation of space time coding techniques for indoor visible light communication systems," in *Proc. Int. Conf. Opt. Netw. Design Model. (ONDM)*, 2018, pp. 88–93.
- [34] A. Alwarafy, A. I. Sulyman, A. Alsanie, S. A. Alshebeili, and H. M. Behairy, "Path-loss channel models for receiver spatial diversity systems at 2.4 GHz," *Int. J. Antennas Propag.*, vol. 2017, no. 1, 2017, Art. no. 6790504.
- [35] J. Yao and N. Ansari, "Wireless power and energy harvesting control in IoD by deep reinforcement learning," *IEEE Trans. Green Commun. Netw.*, vol. 5, no. 2, pp. 980–989, Jun. 2021.
- [36] J. Kong, Z.-Y. Wu, M. Ismail, E. Serpedin, and K. A. Qaraqe, "Q-learning based two-timescale power allocation for multi-homing hybrid RF/VLC networks," *IEEE Wireless Commun. Lett.*, vol. 9, no. 4, pp. 443–447, Apr. 2020.
- [37] A. Alwarafy, B. S. Ciftler, M. Abdallah, M. Hamdi, and N. Al-Dhahir, "Hierarchical multi-agent DRL-based framework for joint multi-RAT assignment and dynamic resource allocation in next-generation HetNets," *IEEE Trans. Netw. Sci. Eng.*, vol. 9, no. 4, pp. 2481–2494, Jul./Aug. 2022.
- [38] A. Alwarafy, B. S. Ciftler, M. Abdallah, and M. Hamdi, "DeepRAT: A DRL-based framework for multi-RAT assignment and power allocation in HetNets," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, 2021, pp. 1–6.
- [39] M. Bloch, J. Barros, M. R. Rodrigues, and S. W. McLaughlin, "Wireless information-theoretic security," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2515–2534, Jun. 2008.
- [40] M. Lu, M. Bagheri, A. P. James, and T. Phung, "Wireless charging techniques for UAVs: A review, reconceptualization, and extension," *IEEE Access*, vol. 6, pp. 29865–29884, 2018.
- [41] U. Gordhan and S. Jayalath, "Wireless power transfer system for an unmanned aerial vehicle," in *Proc. 12th Power Electron., Drive Syst., Technol. Conf. (PEDSTC)*, 2021, pp. 1–5.
- [42] S. Pang, J. Xu, Z. Xie, J. Lu, H. Li, and X. Li, "Lightweight UAV's wireless power transfer system for constant current charging without secondary feedback control," *IEEE Trans. Veh. Technol.*, vol. 72, no. 12, pp. 15611–15621, Dec. 2023.
- [43] Z. Yang, W. Xu, and M. Shikh-Bahaei, "Energy efficient UAV communication with energy harvesting," *IEEE Trans. Veh. Technol.*, vol. 69, no. 2, pp. 1913–1927, Feb. 2020.
- [44] B. Duo, Q. Wu, X. Yuan, and R. Zhang, "Energy efficiency maximization for full-duplex UAV secrecy communication," *IEEE Trans. Veh. Technol.*, vol. 69, no. 4, pp. 4590–4595, Apr. 2020.
- [45] M. Yan, L. Zhang, W. Jiang, C. A. Chan, A. F. Gygax, and A. Nirmalathas, "Energy consumption modeling and optimization of UAV-assisted MEC networks using deep reinforcement learning," *IEEE Sensors J.*, vol. 24, no. 8, pp. 13629–13639, Apr. 2024.
- [46] X. Zhong, F. Chen, S. Xiong, T. Yang, and F. Ji, "On performance analysis for random 3D mobile AUV networks with limited data buffers," *IEEE Trans. Veh. Technol.*, vol. 72, no. 9, pp. 12261–12275, Sep. 2023.
- [47] C. Huang, R. Mo, and C. Yuen, "Reconfigurable intelligent surface assisted multiuser MISO systems exploiting deep reinforcement learning," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1839–1850, Aug. 2020.
- [48] A. Alwarafy, A. Albaser, B. S. Ciftler, M. Abdallah, and A. Al-Fuqaha, "AI-based radio resource allocation in support of the massive heterogeneity of 6G networks," in *Proc. IEEE 4th 5G World Forum (5GWF)*, Oct. 2021, pp. 464–469.
- [49] A. Alwarafy, B. S. Ciftler, M. Abdallah, M. Hamdi, and N. Al-Dhahir, "DRL-based joint RAT association, power and bandwidth optimization for future HetNets," *IEEE Wireless Commun. Lett.*, vol. 11, no. 7, pp. 1503–1507, Jul. 2022.
- [50] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," 2017, *arXiv:1707.06347*.
- [51] X. Lin, F. Yan, Z. Li, X. Xu, W. Xia, and L. Shen, "PPO based computation offloading and resource allocation algorithm in smart grids," in *Proc. 9th Int. Conf. Comput. Commun. (ICCC)*, 2023, pp. 243–248.
- [52] B. Mahmoudi, A. Khonsari, F. Zeinali, M. R. Mili, M. B. Mashhadi, and P. Xiao, "Beamforming and reflection design for short packet ISAC with non-ideal RIS: An A3C-based approach," *IEEE Trans. Veh. Technol.*, vol. 74, no. 5, pp. 8379–8384, May 2025.
- [53] H. Yang, J. Zhao, J. Nie, N. Kumar, K.-Y. Lam, and Z. Xiong, "UAV-assisted 5G/6G networks: Joint scheduling and resource allocation based on asynchronous reinforcement learning," in *Proc. IEEE Conf. Comput. Commun. Workshops (INFOCOM WKSHPS)*, 2021, pp. 1–6.

- [54] B. S. Ciftler, A. Alwarafy, and M. Abdallah, "Distributed DRL-based downlink power allocation for hybrid RF/VLC networks," *IEEE Photon. J.*, vol. 14, no. 3, pp. 1–10, Jun. 2022.
- [55] N. Zhao, Y.-C. Liang, D. Niyato, Y. Pei, M. Wu, and Y. Jiang, "Deep reinforcement learning for user association and resource allocation in heterogeneous cellular networks," *IEEE Trans. Wireless Commun.*, vol. 18, no. 11, pp. 5141–5152, Nov. 2019.
- [56] A. Alwarafy, M. Abdallah, N. Al-Dhahir, T. Khattab, and M. Hamdi, "Rate control for RIS-empowered multi-cell dual-connectivity HetNets: A distributed multi-task DRL approach," *IEEE Trans. Wireless Commun.*, vol. 23, no. 10, pp. 14109–14124, Oct. 2024.
- [57] Q. An, Y. Pan, H. Han, and H. Hu, "Secrecy capacity maximization of UAV-enabled relaying systems with 3D trajectory design and resource allocation," *Sensors*, vol. 22, no. 12, p. 4519, 2022.
- [58] M. Li, X. Tao, N. Li, H. Wu, and J. Xu, "Secrecy energy efficiency maximization in UAV-enabled wireless sensor networks without eavesdropper's CSI," *IEEE Internet Things J.*, vol. 9, no. 5, pp. 3346–3358, Mar. 2022.
- [59] P. Liu, Z. Li, J. Si, N. Al-Dhahir, and Y. Gao, "Joint information-theoretic secrecy and covertness for UAV-assisted wireless transmission with finite blocklength," *IEEE Trans. Veh. Technol.*, vol. 72, no. 8, pp. 10187–10199, Aug. 2023.
- [60] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.



ABDULMALIK ALWARAFY (Member, IEEE) received the Ph.D. degree in computer science and engineering from Hamad Bin Khalifa University, Doha, Qatar. He is currently an Assistant Professor with the Department of Computer and Network Engineering, College of Information Technology, United Arab Emirates University, UAE. With over 25 publications in high-impact peer-reviewed journals and premier international conferences, his research specializes in AI-driven optimization for self-organizing heterogeneous wireless networks,

intelligent radio resource management, and security and privacy in next-generation wireless systems. His work advances the theoretical and practical foundations of autonomous network orchestration, machine learning-based resource allocation, and resilient communication protocols for 6G and beyond.



SUHIB BANI MELHEM (Member, IEEE) received the M.Eng. and Ph.D. degrees in electrical and computer engineering from Concordia University, Canada. He is an Assistant Professor with Al Ain University. He completed a Postdoctoral Fellowship with York University in collaboration with the National Research Council Canada. In 2023, he earned a Certificate in Cybersecurity from Polytechnique Montréal. From May to August 2023, he served as a Scientific Researcher in cybersecurity with Zayed University, UAE,

where he focused on integrating cutting-edge advancements in cybersecurity. His research interests include cloud computing, cybersecurity, information security, machine learning, IoT, and 5G networks.

RAND ABOU CHAHINE is currently pursuing the B.Sc. degree with the College of Information Technology, United Arab Emirates University, Al Ain, UAE. Her current research interests include wireless networks and networks security.

BATOOL SAID is currently pursuing the B.Sc. degree with the College of Information Technology, United Arab Emirates University, Al Ain, UAE. Her current research interests include wireless and IoT networks.

MAITHA ALHARETHI is currently pursuing the B.Sc. degree with the College of Information Technology, United Arab Emirates University, Al Ain, UAE. Her current research interests include wireless and IoT networks.

LATIFA ALMAZROUEI is currently pursuing the B.Sc. degree with the College of Information Technology, United Arab Emirates University, Al Ain, UAE. Her current research interests include wireless and IoT networks.

SARA ALBLOOSHI is currently pursuing the B.Sc. degree with the College of Information Technology, United Arab Emirates University, Al Ain, UAE. Her current research interests include wireless and IoT networks.