IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

*Digital Object Identifier*

# Deep Reinforcement Learning for Rechargeable UAV-Assisted Data Collection from Dense Mobile Sensor Nodes

**Shanshan Bai[1], Xueyuan Wang[1] (Member, IEEE), M. Cenk Gursoy[2] (Senior Member, IEEE), Guangqi Jiang[1], and Shoukun Xu [1]**

[1] School of Computer Science and Artificial Intelligence, Changzhou University, Changzhou, China

[2] Department of Electrical Engineering and Computer Science, Syracuse University, Syracuse NY

Corresponding author: Shoukun Xu (e-mail: xsk@cczu.edu.cn).

"This work was supported in part by the Natural Science Foundation of Jiangsu Province under Grant BK20240962, in part by Changzhou Sci & Tech Program under Grant CJ20220241, in part by National Natural Science Foundation of China under Grant 62306048, and in part by Changzhou Applied Basic Research Fund Project under Grant (CQ20230092; CJ20235036)."

**ABSTRACT** In the realm of the Internet of Things (IoT), unmanned aerial vehicles (UAVs) have garnered significant attention due to their high mobility and cost-effectiveness. However, the limited onboard energy, kinematic constraints, and highly dynamic environments present significant challenges for UAVs in the context of continuous real-time data collection scenarios. To address this issue, we investigate the utilization of a rechargeable UAV for data collection tasks in scenarios with densely mobile sensor nodes. This study formulates the problem as a Markov decision process and designs a reinforcement learning approach called guided search twin-dueling-double deep Q-Network (GS-TD3QN). Within this framework, the goal is to optimize the flight path, charging strategy, and data upload intervals to collectively maximize the total number of uploaded data packets, improve energy efficiency, and minimize the average age of information. Additionally, we propose an action filter to mitigate collision risks and explore various scheduling strategies. Ultimately, by evaluating the performance with simulation results, we confirm the effectiveness of the proposed algorithm and validate its applicability across varying quantities of nodes.

**INDEX TERMS** Deep reinforcement learning, rechargeable UAV, data collection, mobile sensor nodes.

## I. INTRODUCTION

UNMANNED Aerial Vehicles (UAVs) are considered a promising technology for enhancing network coverage and providing on-demand connectivity, having been extensively deployed as mobile communication base stations within the Internet of Things (IoT) [1] [2]. In response, extensive research has focused on UAV-assisted data collection in IoT networks. Initially, traditional algorithms were widely adopted for single-objective optimization tasks. For instance, Le et al. proposed a wireless power transfer (WPT)-enabled mobile crowdsensing (MCS)-assisted sustainable federated learning architecture to minimize the total task completion time [3]. In [4], the average age of information (AoI) was minimized using dynamic programming and genetic algorithms. Li et al. transformed the optimization problem into a traveling salesman problem (TSP) to minimize the AoI [5]. However, these approaches often require retraining when environmental conditions change, leading to substantial computational costs [6]. To address these limitations, deep reinforcement learning (DRL) has gained traction due to its adaptability in handling dynamic environments [7]. Research shifted toward multi-objective optimization. For example, Hu et al. employed guided search DRL to jointly optimize the AoI of high-priority nodes and the total uploaded data with different priorities [8]. Zhang et al. applied the Deep Deterministic Policy Gradient (DDPG) algorithm to minimize completion time while maximizing data collection [9].

The above studies generally assumed that a UAV could accomplish its task from the starting point to the end point in a single flight. However, in practical applications, as network scales expand [10], sensor nodes (SNs) continuously generate data packets, necessitating that UAVs are capable of prolonged data collection [11]. The limited onboard energy of UAVs poses a significant challenge for continuous data collection, making UAV recharging essential [12]. Due to this, some studies had considered allowing UAVs to recharge.The strategies discussed in [13] and [14] both involved comparing the remaining battery level of the UAV with the minimum

required battery level needed to return to the charging station (CS), and choosing a straight-line path back for recharging. However, this approach may not be optimal. On one hand, data can not be collected during the return trip, increasing the time cost of round trips. On the other hand, these strategies do not consider obstacles, making the straight-line path infeasible in environments with obstacles. Existing studies mainly focus on optimizing data acquisition, energy consumption, AoI, and task completion time [15], with little attention to energy efficiency (EE). Fu et al. proposed Q-learning to optimize EE and charging costs but overlooked AoI and environmental obstacles [16]. Yi et al. employed the D3QN to optimize EE and AoI, but they did not consider obstacles or the sustainable generation of data packets [17]. Due to this, this paper incorporates these factors into the optimization.

The aforementioned studies primarily focused on scenarios with a limited number of nodes. In dense network environments, researchers have proposed various optimization approaches. As an illustration, Wang et al. proposed a method where UAVs alternate between charging and data collection, clustering SNs during charging and selecting cluster heads as target nodes for immediate data collection afterward [18]. Zhu et al. clustered the nodes and used ground vehicles to carry spare batteries to accompany the UAV [19]. Nevertheless, these approaches are not applicable to the mobile SNs scenarios addressed in this study. Pre-clustering mobile nodes introduces substantial challenges, such as inefficiencies arising from dynamic changes in node locations. Therefore, we propose guided search to to assist the UAV in data collection. With the increasing prevalence of mobile sensors and intelligent devices, the scope of IoT has expanded to include mobile networks, where IoT terminals are no longer fixed to static infrastructures but instead move dynamically across different environments. This introduces additional challenges for UAV operations, as UAVs must continuously adjust their flight paths to maintain communication with SNs [20]. Existing studies on mobile SNs have primarily focused on objectives such as movement planning or security. For example, [21] employed a twin-delayed deep stochastic policy gradient (TDDS) to guide SNs to specific positions to collect data and upload it to edge stations. Similarly, [22] addressed secrecy rate maximization in scenarios involving relay UAVs and interference UAVs. Currently, no research has thoroughly investigated UAV trajectory optimization and data collection in scenarios involving dense mobile SNs. Inspired by the aforementioned studies, this paper explores UAV trajectory optimization in such environments, offering a novel approach to data collection.

Based on the above considerations, this paper investigates the problem of data collection by rechargeable UAV trained with DRL in scenarios with dense mobile nodes. The main contributions of this paper are summarized as follows:

- Unlike previous studies, we introduce mobile SNs into the environment and assumed that the information of SNs is only partially known. Additionally, to better reflect real-world scenarios, obstacles are added, and charging stations are implemented to support the UAV's sustainable data collection tasks.
- We formulate the rechargeable UAV data collection problem as a Markov decision process (MDP) and propose the guided search twin-dueling-double deep Q-Network (GS-TD3QN) algorithm. This algorithm incorporates a twin architecture into the D3QN framework to enhance stability, allowing for the joint optimization of the total number of uploaded data packets, energy efficiency, and average AoI.
- By comparing various scheduling strategies, the optimal strategy is selected as the guided search (GS) to assist the UAV in efficient data collection. Additionally, to ensure flight safety and minimize risks in real-world environments, an action filter system is designed to predict and prevent UAV collisions with obstacles, providing real-time feedback during decision-making.

The effectiveness and stability of the proposed GS-D3QN are validated through comparisons with baseline algorithms. Furthermore, the universality of the proposed algorithm was tested across different numbers of nodes.

The remainder of this paper is structured as follows: Section II introduces the system model and the problem formulation. Section III presents the proposed GS-TD3ON algorithm along with the action filter mechanism. Section IV discusses the simulation results, and finally, Section V concludes the paper with a summary.

## II. SYSTEM MODEL AND PROBLEM FORMULATION

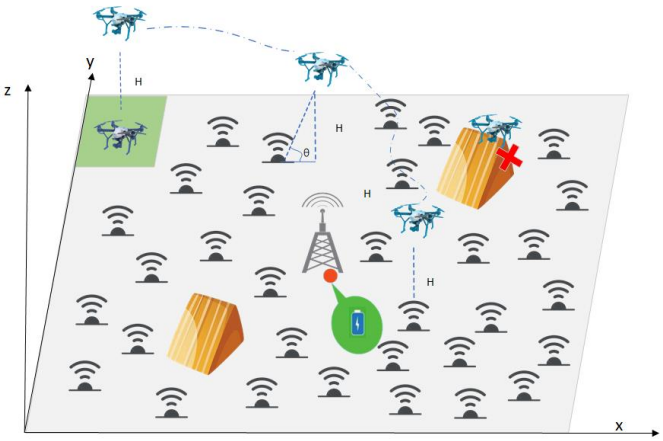



FIGURE 1: An illustration of a rechargeable UAV-assisted data collection network.

### A. SYSTEM MODEL

As illustrated in Fig. 1, a rechargeable UAV is deployed from an initial position (indicated by the green diamond area at the top-left corner) to collect data generated by SNs within a designated region. When the UAV intends to upload the collected data packets to the base station (BS), located at $P_B = [x_B, y_B]$, it must be within the communication range

of the BS, which means that the distance between the UAV and the BS must not exceed $d_B$. For the sake of simplicity, we assume that the data upload occurs fast. Under the assumption that the UAV flies at a fixed altitude $H$ and maintains constant speed during data collection, its horizontal position in time slot $t$[1] is $q(t) = [x_u(t), y_u(t)]^T$. The trajectory of the UAV during the entire collection process can be represented as $\mathcal{T} = \{q(0), \cdots, q(T-1)\}$, where $T$ is the total number of time slots. Considering the limited battery capacity $u_c$ of the UAV, a charging station (CS) is set up below the BS (marked by a red dot). The UAV can proceed to the CS as needed and is supposed to be fully charged within the time slot $t$ upon entering the proximity of the CS. To mitigate the wear and tear on the battery caused by frequent charging, it is stipulated that data upload to the base station and charging cannot occur simultaneously. Within the designated area of size $L_1 \times L_2$, $N$ SNs are randomly distributed. These SNs generate data packets intermittently and can move a certain distance $l$ in any direction, where $l \leq L_0$ and $L_0$ denotes the maximum distance each SN can move within each time slot $t$. The information of mobile SNs is represented as $\mathbb{N} = \{I_1, I_2, \cdots, I_n\}$, where $I_n$ represents the information of the $n$-th node, specifically, $I_n = (n_{th}, x_n(t), y_n(t), D_n(t), T_n(t))$, $n \in [1, N]$. To simplify the notation, the following descriptions omit the use of $(t)$. In the information $I_n$, $n_{th}$ represents the index of $n$-th SN, $x_n$ and $y_n$ represent the position of the SN, $D_n$ is the list of data packet generated by the SN that have not yet been collected, and $T_n$ denotes the timestamp of packets generation. Additionally, several obstacles exist within the bounded area, which the UAV cannot fly over and must navigate around. To prevent interference during transmission, it is assumed that the UAV adopts a time-division multiple access (TDMA) protocol [23]. More specifically, it is stipulated that within a given time slot $t$, at most one SN is allowed to communicate with the UAV.

### B. COMMUNICATION MODEL

Similar to [24], due to the absence of buildings, non-line-of-sight (NLoS) links can be neglected, and only line-of-sight (LoS) links are considered in this study. Consequently, with the LoS links the path loss can be expressed as follows [25]:

$$L(d) = (d^2 + H^2)^{\alpha/2}, \tag{1}$$

where $H$ represents the altitude of the UAV and $d$ is the horizontal distance between the UAV and the SN. To be more specific, we project the position of the UAV onto the ground plane, and denote the projected position as $q = (x_u, y_u)$. The instantaneous distance between the UAV and the $n$-th SN is given by $d = \sqrt{(x_u - x_n)^2 + (y_u - y_n)^2 + H^2}$. $\alpha$ represents the path loss exponent.

We assume that each SN is equipped with an omnidirectional antenna, while the UAV is equipped with a directional antenna. According to [26], the antenna gain from the SN to the UAV, denoted as $G_V(d)$, can be approximated as follows:

$$G_V(d) = \sin(\theta) = \frac{H}{\sqrt{d^2 + H^2}}, \tag{2}$$

where $\theta$ denotes the elevation angle between the UAV and the SN, as depicted in Fig. 1.

Given $P_n$ as the transmit power of each SN and $\mathcal{N}_s$ as the noise power, the signal-to-noise ratio (SNR) between the UAV and the $n^{th}$ SN is expressed using the following formula [27]:

$$\mathbb{S}_n \triangleq \frac{P_n}{\mathcal{N}_s} G_V(d) L^{-1}(d) = \frac{P_n}{\mathcal{N}_s} H (d^2 + H^2)^{-\frac{1+\alpha}{2}}. \tag{3}$$

We define a fixed threshold $\bar{S}$ as the SNR criterion for determining whether the UAV can collect data from SNs. Specifically, the UAV can successfully gather data packets generated by SNs only if the SNR of the $n$-th SN is no smaller than the threshold $\bar{S}$, and this can be expressed as

$$C_n(t) = \begin{cases} L_{D_n}, & \text{if packet generated and } \mathbb{S}_n \geq \bar{S} \\ 0, & \text{otherwise.} \end{cases} \tag{4}$$

here $L_{D_n}$ is the number of data packets generated by the SN.

The UAV's data storage buffer is assumed to have sufficient capacity, and data can only be successfully uploaded to the BS when the UAV is at a distance no greater than $d_B$. Then the total amount of data uploaded, denoted as $C_d$, is expressed as follows:

$$C_d = \begin{cases} \sum_{t=0}^{T-1} \sum_{n=1}^{N} C_n(t), & \text{if } d_g \leq d_B \\ 0, & \text{otherwise.} \end{cases} \tag{5}$$

### C. ENERGY EFFICIENCY (EE)

During flight and hovering, the UAV expends energy to overcome gravity and air resistance in order to stay aloft or move forward [28]. Since the energy consumption related to communication is negligible compared to that of propulsion, it is not taken into account in this paper. This paper focuses on the flexibility and adaptability of rotary-wing UAVs. Based on [29], we employ the following basic energy consumption model. The energy consumption $E_v$ during the time interval $\tau$ can be represented as

$$E_v(t) = P(v)\tau, \tag{6}$$

where $P(v)$ represents the propulsion power of the UAV. When $v = 0$, the UAV is considered to be hovering, which is regarded as a specific case of flying.

Therefore, the total energy consumed over the entire flight duration $T$ can be expressed as

$$E(T) = \sum_{t=0}^{T-1} E_v(t). \tag{7}$$

It should be noted that only one SN can communicate with the UAV at any given time slot. Then, the maximum amount

[1]To simplify the problem, the continuous collection period is discretized into T discrete time slots.

of data the UAV can receive during the entire mission duration $T$ is expressed as

$$R(T) = \int_0^T B \log_2 (1 + \mathbb{S}_n(t))\, dt, \tag{8}$$

where $B$ is the bandwidth of the communication channel.

In communication systems, optimizing energy efficiency (EE) involves balancing energy consumption with data transmission/collection [30]. The aim is to achieve higher data collection while minimizing energy expenditure. Consequently, EE, which measures the amount of data collected per unit of energy, can be defined as:

$$EE = \frac{R(T)}{E(T)}. \tag{9}$$

### D. AGE OF INFORMATION (AOI)

To characterize the freshness of the collected data, AoI is defined as the time that has passed since SN has generated the latest information [31]:

$$AoI_n(t) = t - T_u, \tag{10}$$

where $t$ represents the current time, and $T_u$ denotes the timestamp for the most recent update of the information source.

As illustrated in Fig. 2, the SN generates a data packet in time slot $t_g$, and AoI of the packet increases over time. UAV collects the data packet at time $t_i$ and uploads to the BS at time $t_u$. The AoI in this scenario is $t_u - t_g$.

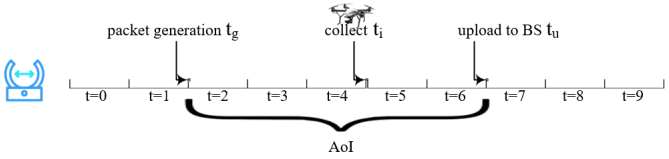



FIGURE 2: AoI of the sensor nodes.

### E. PROBLEM FORMULATION

In this paper, we investigate the problem of efficient data collection with a rechargeable UAV. Our objective is to ensure that the UAV can recharge effectively within a given collection period, while balancing EE and AoI, and maximizing the collection of data packets generated by mobile SNs. Specifically, we consider the following optimization problem:

$$\mathrm{P}_1 : \max_{\mathcal{T}} \; EE - \frac{1}{N}\sum_{n=1}^{N} AoI(n) + C_d \tag{11}$$

$$\text{s. t.} \quad q(t) \in L_1 \times L_2 \; \forall t, \tag{11a}$$

$$u_c \geq 0 \; \forall t, \tag{11b}$$

$$\mathbb{S}_n \geq \bar{S} \; \forall n, \tag{11c}$$

$$||q(t) - p_o||_2 > 0 \; \forall t \text{ and } \forall o, \tag{11d}$$

Above, $\mathcal{T} = \{q(0), \cdots, q(T)\}$ represents the trajectory of the UAV. $C_d$ indicates the total number of data packets successfully transmitted by the UAV to the BS. In (11b), $u_c$ represents the remaining battery level of the UAV, and the constraint is imposed to ensure that the UAV has sufficient energy to sustain its flight and complete its mission without running out of power. The constraint in (11c) represents the condition that the UAV collects data only when the SNR in the communication link with a node exceeds the required threshold. The constraint in (11d) ensures that the UAV avoids collisions with obstacles, where $p_o = (x_o, y_o)^T$ represents the coordinates of an obstacle.

## III. REINFORCEMENT LEARNING-BASED AGENT

Considering the complexity of the nonlinear programming problem in (11), we propose an DRL approach to determine the optimal trajectory based on the specified performance metrics.

### A. GS-TD3QN

In this section, we provide a detailed description of a UAV trajectory design algorithm based on GS-TD3QN. The architecture of this algorithm features a twin structure that includes two independent Q-networks. These networks share the same architecture but have separate parameters and are updated independently [32]. The proposed RL framework is depicted in Fig. 3.

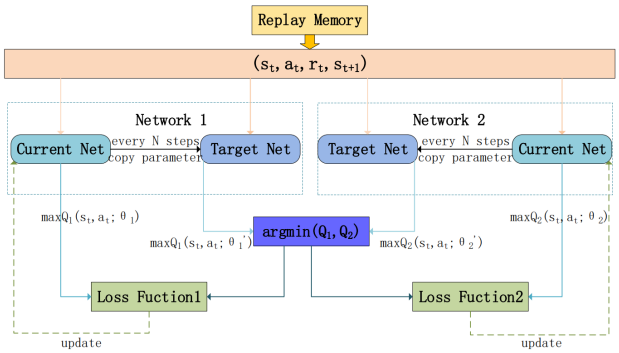



FIGURE 3: The framework structure of GS-TD3QN.

During the training process, the UAV obtains state from the simulated environment and selects action by GS-TD3QN. The action is then processed through action filter to eliminate those that may result in collisions or other hazardous scenarios. Subsequently, the reward is computed based on the outcomes of the filtered action. Through repeated interactions with the environment, the GS-TD3QN's network parameters are continuously optimized, ultimately yielding a well-performing pre-trained policy. The detailed training process is further described in Algorithm 1. In practical deployment, the UAV utilizes this pre-trained policy with necessary fine-tuning, significantly reducing the latency associated with algorithm training and effectively meeting the requirements of real-world tasks.

Now, we construct this DRL framework with state, action and reward, and the relevant specifics are described below:

1) State Space($\mathcal{S}$)

$$\mathcal{S} \triangleq [S_u, [S_n^n \ \forall n]]. \tag{12}$$

Here, $S_u$ represents the state of the UAV, which can be expressed as:

$$S_u = [u_x, u_y, g_x, g_y, d_g, \theta_g, \theta_u, t_l, EE, u_{CS}, e_{emp}, a_u, u_{BS}], \tag{13}$$

where $u_x$ and $u_y$ respectively represent the flight distance along the $x$ and $y$ axis at time step $t$. Explicitly, $u_x = v\cos\theta_u$ and $u_y = v\sin\theta_u$, where $\theta_u$ denotes the flying direction of the UAV and $v$ is the UAV speed. $g_x$ and $g_y$ respectively represent the $x$ and $y$ coordinates of the BS or CS (BS/CS), and $d_g$ represents the distance from the UAV to the BS/CS. $\theta_g$ denotes the angle between the UAV and the BS/CS, $t_l$ denotes the remaining flight time of the UAV, and $EE$ is the energy efficiency of the UAV over the entire period from start to present. $u_{CS}$ indicates whether recharging is needed; when the UAV's battery level drops below 2000, $u_{CS}$ is set to 1; otherwise, it is set to 0. Similarly, $e_{emp}$ is set to 0 if the battery has power, and 1 otherwise. $a_u$ represents the average AoI of the data packets stored in the UAV's data buffer. $u_{BS}$ indicates whether the stored data packets in storage should be transmitted to the base station. If the AoI exceeds a predetermined threshold, $u_{BS}$ is set to 1; otherwise, it is set to 0.

$S_n^n$ represents the state of the SNs, which can be defined as:

$$S_n = [x_n, y_n, d_n, \theta_n^n, D_n, \mathbb{S}_n, C_n, a_n], \tag{14}$$

where, $x_n$ and $y_n$ represent the coordinates of the $n$-th SN along the $x$-axis and $y$-axis, respectively. The distance between SN and UAV is denoted by $d_n$, while $\theta_n^n$ represents the elevation angle between them. The average AoI of the data packets generated by the SN is indicated by $a_n$. Considering that the information on SNs is only partially known to the UAV, we stipulate that the UAV can observe the nearest $N_c$ SNs that have generated data packets. If there are fewer than $N_c$ SNs, we use zero padding.

2) Action Space ($\mathcal{A}$)

The action space for the UAV is defined as follows:

$$\mathcal{A} = [a_f, a_c, a_b]. \tag{15}$$

The action set is discretized into 10 specific options. Among them, the UAV's flight direction $\theta_u$ is constrained within $(-\theta_c, \theta_c)$ and discretized into 7 directions. The flight action $a_f$ is defined as $a_f = [v, \theta_u]$, where $v$ represents the UAV's constant flight speed. Having $v = 0$ indicates a hovering operation. The 9th action, $a_c$, represents the decision to charge, while the 10th action, $a_b$, indicates uploading data to the base station.

3) Reward ($\mathcal{R}$)

We define the reward as:

$$R = p_a + r_a + r_o + p_b + r_u + r_c + r_n + r_e. \tag{16}$$

Above, when the average AoI of the data stored in the UAV's storage buffer exceeds $\mathbb{A}$, a penalty of $p_a$ is applied. $r_a$ indicates the negative average AoI of the data collected by the UAV and is specifically designed to optimize the AoI within the objective function. $r_o$ denotes the penalty for encountering an obstacle, while $p_b$ represents the penalty for leaving the designated area, both designed to satisfy the constraints specified in Equations (11d) and (11a). $r_u = C_d r_b$ is the reward given for uploading data, aiming to optimize $C_d$ as defined in Equation (11). $r_c$ is granted for Equation (11b) and is defined as:

$$r_c = \begin{cases} a_1, & \text{if } dis_c > d_{th} \\ a_2, & \text{if } dis_c \le d_{th} \text{ and battery} \le B_{th}, \\ a_3, & \text{if } dis_c \le d_{th} \text{ and battery} > B_{th}, \\ a_4, & \text{battery } \le 0. \end{cases} \tag{17}$$

Here, $a_1$, $a_3$, and $a_4$ are negative constants, while $a_2$ is a positive constant. $dis_c$ and $d_{th}$ represent the distance between the UAV and the BS, and the communication distance threshold, respectively. battery and $B_{th}$ represent the remaining battery level and the specified battery threshold, respectively. To simplify the modeling process, we assume that the UAV can recharge only when it is within a certain range of the CS and that its battery is instantaneously fully charged, restored to its maximum capacity $u_c$. While recharging helps ensure the UAV's operational continuity, charging too frequently may negatively impact task efficiency [33]. Therefore, reward is granted only when the UAV recharges with its remaining battery level below a specified threshold $B_{\text{th}}$, whereas frequent charging with sufficient battery results in a penalty.

$r_n = C_n(t)\, r_d$ is the reward for collecting new data, where $r_d$ is a positive constant designed to encourage the UAV to maximize data collection. To optimize the EE, the reward $r_e$ is introduced to maximize the EE in Equation (11).

### B. ACTION FILTER

Notably, this paper introduces an action filter to evaluate the chosen actions. This action filter can detect potential risks in real-time and provide feedback, allowing the UAV to reselect its actions, thereby effectively mitigating potential losses or avoiding collisions.

The action filter $A_t^{\text{filtered}}$ can be defined as follows:

$$A_t^{\text{filtered}} = \begin{cases} a_t, & \text{if} P(a_t) = 0 \\ \text{HoldPosition}, & \text{if} P(a_t) = 1. \end{cases} \tag{18}$$

When the UAV selects its current action, the action is fed to the action filter for preprocessing. The system pre-executes the action to assess potential risks such as collisions or boundary violations. If such risks are detected, that is, when $P(a_t) = 1$, the system will force the UAV to maintain its current position and cancel the execution of the selected action. Additionally, the system will send a penalty signal to the UAV, indicating the imprudence of the chosen action and guiding the UAV toward more optimal flight decisions.

**Algorithm 1** GS-TD3QN algorithm

**Input:** Initialize the replay buffer $D$, capacity $N$, update steps $C$, max. flight time $T_F$, max. training episode $E$, exploration probability $\varepsilon$ and learning rate $\lambda$.

**for** episode in 0,1,...,T **do**
  Initialize $Q_1, Q_2$ network parameters $\xi_1, \xi_2$
  Initialize $Q_1$ target network and $Q_2$ target network parameters $\xi_1{}'$ and $\xi_2{}'$
  Initialize the UAV position, node states, goal position, flight time $T_f$
  **while** True **do**
    With probability $\varepsilon$, randomly select action $a_t$
    otherwise select $a_t = \arg\max Q_1(s_t, a; \xi_1)$
    Action filter system
    The UAV takes action $a_t$
    The UAV receives reward $R$, moves to next state $s_{t+1}$
    Store $(s_t, a_t, R, s_{t+1})$ in replay buffer
    Randomly sample $K$ samples from D
    Calculate the target values $Q_1, Q_2$ of each sample according to
    $Q = \arg\min(\arg\max Q_1, \arg\max Q_2)$
    $y_t = \begin{cases} r_t, \text{ if episode terminates at time+1} \\ r_t + \gamma \max_{a'} Q(s_{t+1}, a'; \xi^-), \text{otherwise} \end{cases}$
    Update the network parameters using $L(\xi) = (y_t - Q)^2$
    $s_{t+1} \leftarrow s, f \leftarrow f + 1$
    Every C steps, reset $\xi_1{}' \leftarrow \theta\xi_1, \xi_2{}' \leftarrow \xi_2$
    **if** flight time $T_f > T_F$ or battery $<0$ **then**
      break
    **end if**
  **end while**
**end for**

### C. SCHEDULING POLICIES

Given the large number of mobile SNs, the UAV must not only focus on data collection but also learn to recharge efficiently and upload data in a timely manner. To effectively address these challenges and enhance the UAV's performance and efficiency, we introduce five distinct scheduling policies (SP), with the most effective strategy designated as Guided Search (GS). The GS strategy serves as the key approach to optimize the UAV's operations.

SP1: This policy takes into account the distance between UAV and SN, the average AoI of the data packets stored in the SN, and the number of data packets generated by SNs. It selects the target SN based on the minimum sum of these three factors. This policy also constitutes the GS proposed in our study.

SP2: This policy determines the target SN by selecting the SN with the minimum sum of the distance from the UAV to the SN and the average AoI of the data packets stored in the SN.

SP3: This policy identifies the target SN by selecting the SN with the minimum sum of the distance between the UAV and the SN and the number of data packets generated by the SNs.

SP4: This policy addresses only the average AoI of the data packets stored in each SN and selects the SN with the smallest average AoI as the target SN.

TABLE 1: **The Environmental Parameters**

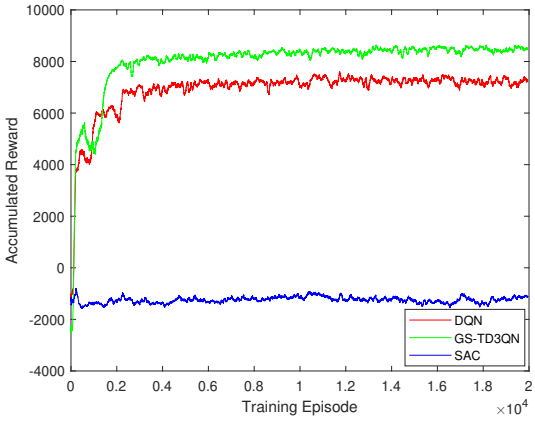
| Notation | Definition | Value |
|---|---|---|
| $\alpha$ [25] | Exponent of path loss | 2 |
| $B$ [25] | Channel bandwidth | 1MHz |
| $d_B$ | The maximum distance the UAV can upload data to the BS. | 5m |
| $u_c$ | The battery of the UAV | 120000mAh |
| $N$ | The number of SNs | 50 |
| $v$ | The flying speed of UAV | 5m/s |
| $H$ | The flying height of UAV | 50m |
| $T$ | The flight time of UAV | 1000s |
| $P_n$ [27] | Transmitting power of sensor equipment | 1e-3 |
| $\mathcal{N}_s$ [27] | Noise power of sensor equipment | 1e-6 |
| $P(0)$ [29] | Propulsion power in hovering state | 222W |
| $P(5)$ [29] | Propulsion power during flight at a speed of 5 | 215W |
| $\bar{S}$ | SNR Threshold | 0.37 |
| $\theta_c$ | The maximum turning angle the UAV can achieve during flight | $\frac{\pi}{3}$ |
| $\mathbb{A}$ | The AoI threshold for data packets stored in the UAV's storage pool | 100s |
| $d_{th}$ | The given communication distance threshold | 5m |
| $B_{th}$ | The specified battery threshold | 20000mAh |
| $N_c$ | The maximum number of SNs observable by the UAV | 10 |
| $p_a$ | The penalty for the average AoI of data packets in the UAV's storage pool exceeding threshold | -2 |
| $r_o$ | The reward of collision with obstacles | -0.2 |
| $p_b$ | Out of Bounds Reward | -0.2 |
| $r_b$ | The reward of upload data to BS | 5 |
| $r_d$ | The reward of collect data | 2 |
| $L_0$ | The maximum moving distance of nodes | 1m |
| $P_o$ | The coordinate of obstacle 1 | [-23,-23,6] |
| | The coordinates of obstacle 2 | [20,20,6] |
| $P_B$ | The coordinate of BS | [0,0,6] |
| $P_C$ | The coordinate of CS | [0,0,2] |
| $a_1, a_3, a_4$ | The negative reward of $r_c$ | -5,-40,-500 |
| $a_2$ | The positive reward of $r_c$ | 5 |

SP5 : This policy depends only on the Euclidean distance and chooses the SN closet to the UAV as the target SN.

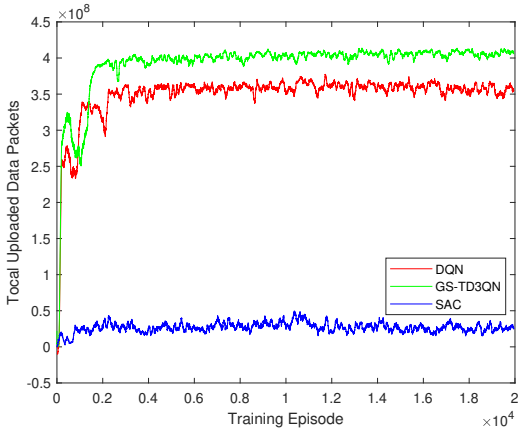
## IV. SIMULATION RESULTS AND ANALYSIS

In this section, we present comprehensive simulation results and evaluate the performance of the proposed GS-TD3QN algorithm, implemented using the PyTorch framework. The simulation environment consists of a 100m $\times$ 100m area, divided into 20m$\times$20m grids and treated as a 2D plane due to the UAV's constant flight altitude. The UAV takes off from a random position in a rectangular region, which is defined by $x$-axis coordinates ranging from -50 to -30 and $y$-axis coordinates ranging from 30 to 50, as illustrated by the green area in Figure 1. Mobile SNs are randomly distributed within the scenario, and there are two obstacles in the environment. Additional details are provided in Tables 1 and 2.

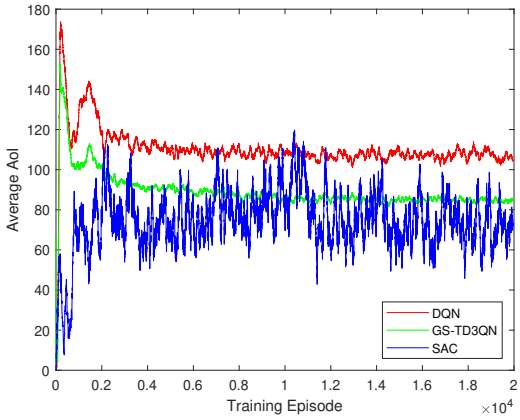
We define the total number of data packets successfully uploaded to the BS as CD, the number of charging instances as CH, the ability of a UAV to maintain continuous power supply throughout the mission duration as Power Sustainment Success Rate (PSSR), the average number of data uploads as UP.



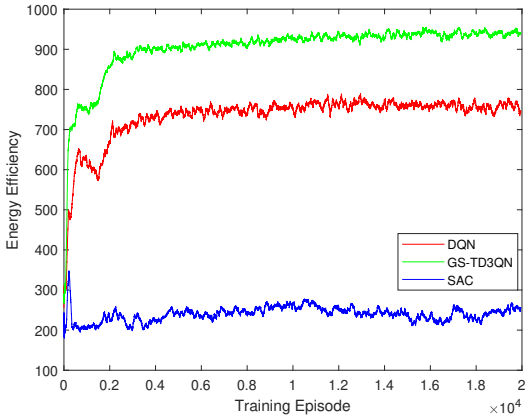

(a) Accumulated reward during training.




(b) Total uploaded data packets during training.




(c) Average AoI during training.




(d) Energy efficiency during training.

FIGURE 4: Training performance metrics.

TABLE 2: **The Hyperparameters of The GS-TD3QN**

| Parameter | Value |
|---|---|
| Number of trainning round episodes | 20000 |
| Target network update frequency | 1000 |
| Size of repaly buffer | 1e6 |
| Mini-batch size | 256 |
| Discount factor | 0.99 |
| Learning rate | 0.0003 |
| The number of neurons in the first hidden layer | 256 |
| Initial epsilon | 0.6 |

### *A. PERFORMANCE ANALYSIS OF DIFFERENT ALGORITHMS*

We use traditional DQN and soft actor-critic (SAC) algorithms as baseline comparisons to validate the effectiveness and convergence of the GS-TD3QN algorithm. The red, green, and blue curves represent DQN, GS-TD3QN, and SAC, respectively. Fig. 4a, Fig. 4b, Fig. 4c, and Fig. 4d illustrate the performance of the three algorithms across various dimensions during the training process. These figures comprehensively reflect the effects of each algorithm across multiple aspects, providing an in-depth revelation of their performance differences and characteristics throughout the training. Although SAC achieves the lowest AoI, this is attributed to its extremely limited data collection. The results indicate that in our system model, both DQN and GS-TD3QN algorithms perform better than SAC; however, DQN underperforms compared to GS-TD3QN in most performance metrics. This suggests that the proposed twin architecture places greater emphasis on long-term benefits and favors a more diversified approach. GS-TD3QN incorporates two Q-networks with identical structures but completely independent parameters, reducing mutual interference through independent updates. Additionally, the algorithm employs a conservative estimation approach, effectively mitigating the risk of overestimated Q-values during the update process. This design enhances the algorithm's exploration capabilities in complex state spaces, enabling it to avoid local optima and more effectively identify global optimal solutions. Therefore, although GS-TD3QN converges slightly slower than DQN, it ultimately achieves higher overall gains.

TABLE 3: **Test Results of Different Algorithms**

| Algorithm | DQN | GS-TD3QN | SAC |
|---|---|---|---|
| **Reward** | 7027.26 | 8894.92 | -1131.45 |
| **CD** | 3.537e8 | 4.150e8 | 2.854e7 |
| **CH** | 1.84 | 1.66 | 0.27 |
| **EE** | 714.17 | 975.2 | 213.16 |
| **PSSR%** | 97.7 | 98.8 | 92.3 |
| **Avg AoI(s)** | 111.48 | 80.12 | 65.73 |
| **UP** | 71.8 | 72.9 | 54.3 |

TABLE 4: **Test Results of NLOS**

| Algorithm | LOS | LOS+NLOS |
|---|---|---|
| **Reward** | 8895 | 7933 |
| **CD** | 4.150e8 | 3.214e8 |
| **CH** | 1.66 | 1.92 |
| **EE** | 975.2 | 746 |
| **PSSR%** | 98.8 | 99.2 |
| **Avg AoI(s)** | 80 | 88 |

TABLE 6: **Test Results for Varying Sensor Nodes**

| Num | 10 | 30 | 70 | 100 | 120 |
|---|---|---|---|---|---|
| **Reward** | **8055** | 8569 | 8359 | 8290 | 8337 |
| **CD** | **4.292e8** | 4.163e8 | 3.924e8 | 3.913e8 | 3.908e8 |
| **CH** | **1.89** | 1.72 | 1.53 | 1.53 | 1.59 |
| **EE** | **515** | 835 | 1130 | 1230 | 1275 |
| **PSSR%** | **98.7** | 97.2 | 92.8 | 92.0 | 93.8 |
| **Avg AoI(s)** | **51** | 74 | 75 | 80 | 84 |

To provide a more intuitive evaluation of algorithm performance, we have summarized the test results in Table 3. Compared to DQN and SAC, the GS-TD3QN's reward values are higher by approximately 26.5% and 887.6%, CD is higher by approximately 17.33% and 1354.10%, energy efficiency is higher by approximately 36.6% and 357.6%, and PSSR is higher by approximately 1.13% and 7.04%, respectively. These results demonstrate that GS-TD3QN exhibits superior advantages in optimizing drone performance.

Our model demonstrates good scalability and can adapt to more realistic NLOS scenarios. To evaluate the performance of GS-DT3QN under such conditions, we compared the experimental results for LOS and LOS+NLOS scenarios, as shown in table 4. Overall, the introduction of NLOS conditions had some impact on the model's performance, with all metrics showing varying degrees of decline. This can be attributed to the obstruction of signal propagation by obstacles, which degrades the quality of communication links and negatively affects overall performance. Nevertheless, the changes are relatively minor, indicating that the model maintains reasonable adaptability and stability even in complex environments.

### B. VALIDATION OF THE EFFECTIVENESS OF THE PROPOSED GS.

We further compare the different scheduling policies (SPs) outlined in Section III.C and summarize the results in Table 5. The SP1 not only collects the largest amount of packets but also achieves the smallest AoI for the uploaded data packets and demonstrates excellent EE. It fully meets the objectives of our study. When considering a single metric, such as in SP4, where only the AoI of data packets generated by SNs is considered, the AoI may be optimal, but the performance in other aspects is poor. In SP2, where both distance and AoI are taken into account, multiple data packets may have similar conditions, leading to the selection of suboptimal packets and resulting in an AoI performance that is inferior to that of SP3, which considers both distance and the total number of data packets. Observations show that the total number of data packets in SP3 is also much lower than in SP1, likely due to the similar conditions of multiple data packets, requiring a comprehensive consideration of the AoI. Therefore, selecting SP1 as the GS strategy in this study is experimentally justified.

TABLE 5: **Test Results of Different Scheduling Policies**

| Policy | SP1 | SP2 | SP3 | SP4 | SP5 |
|---|---|---|---|---|---|
| **Reward** | **8895** | 4365 | 6602 | 1442 | 4130 |
| **CD** | **4.150e8** | 3.072e8 | 3.696e8 | 8.871e7 | 3.236e8 |
| **CH** | **1.66** | 1.83 | 1.39 | 0.095 | 1.76 |
| **EE** | **975.2** | 289 | 780 | 438 | 251 |
| **PSSR%** | **98.8** | 99 | 99.2 | 8.7 | 97.8 |
| **Avg AoI(s)** | **80.12** | 198 | 114 | 68.2 | 220 |

### C. VALIDATION OF THE ALGORITHM'S GENERALIZATION ACROSS DIFFERENT NUMBERS OF SNS.

Finally, we train and test the proposed algorithm in environments with varying numbers of SNs to validate its universality, as shown in Table 6. Regardless of the number of SNs being as few as 10 or as many as 120, it can be observed that the UAV typically only needs to recharge approximately once to serve all SNs within the designated period. Although the total amount of data packets successfully uploaded to the BS decreases as the number of SNs increases, the UAV is still able to effectively complete the data upload for the majority of data packets. As the number of SNs in the environment increases, the AoI of the uploaded data steadily rises but remains within an acceptable range. However, when the number of SNs exceeds 70, the PSSR shows a significant decline, indicating that the capacity of a single UAV is limited and cannot optimally balance recharging and data collection, resulting in a lower PSSR. Thus, in future work, we aim to determine the optimal UAV-to-sensor node ratio to enhance resource utilization and system performance across various deployment scenarios. This includes simulations to evaluate different configurations for scalability and efficiency. Additionally, we plan to develop advanced algorithms for dynamic UAV allocation and task assignment based on real-time sensor data, improving operational effectiveness in complex environments.

## V. CONCLUSIONS

This paper addresses the path planning problem for rechargeable UAV-assisted data collection from dense mobile SNs. To enable UAV to effectively avoid obstacles while performing long-term data collection tasks, we employ reinforcement learning techniques to jointly optimize the uploaded data, EE,

and average AoI. Specifically, we propose the GS-TD3QN algorithm, which utilizes GS to assist UAVs in path planning by designing a reward function and action filter. Through comparisons of various performance metrics under different scheduling policies, the effectiveness of GS is experimentally validated. Simulation results indicate that the GS-TD3QN algorithm outperforms the DQN and SAC algorithms in the studied environment, demonstrating superior performance. Furthermore, the algorithm exhibits robust performance in environments with varying numbers of nodes, proving its efficiency and strong convergence capabilities. However, as the number of nodes increases, the performance of a single UAV declines. Therefore, future research will focus on determining the optimal number of UAVs based on the scale of sensor nodes to effectively address this issue.

## REFERENCES


[1] Y. Zeng, R. Zhang, and T. J. Lim, "Wireless communications with unmanned aerial vehicles: Opportunities and challenges," *IEEE Communications magazine*, vol. 54, no. 5, pp. 36–42, 2016.

[2] L. Gupta, R. Jain, and G. Vaszkun, "Survey of important issues in uav communication networks," *IEEE communications surveys & tutorials*, vol. 18, no. 2, pp. 1123–1152, 2015.

[3] M. Le, D. T. Hoang, D. N. Nguyen, W.-J. Hwang, and Q.-V. Pham, "Wirelessly powered federated learning networks: Joint power transfer, data sensing, model training, and resource allocation," *IEEE Internet of Things Journal*, 2023.

[4] J. Liu, X. Wang, B. Bai, and H. Dai, "Age-optimal trajectory planning for uav-assisted data collection," in *IEEE INFOCOM 2018-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*. IEEE, 2018, pp. 553–558.

[5] M. Li, X. Liu, and H. Wang, "Completion time minimization considering gns' energy for uav-assisted data collection," *IEEE Wireless Communications Letters*, 2023.

[6] K. K. Nguyen, T. Q. Duong, T. Do-Duy, H. Claussen, and L. Hanzo, "3d uav trajectory and data collection optimisation via deep reinforcement learning," *IEEE Transactions on Communications*, vol. 70, no. 4, pp. 2358–2371, 2022.

[7] S. Xu, X. Zhang, C. Li, D. Wang, and L. Yang, "Deep reinforcement learning approach for joint trajectory design in multi-uav iot networks," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 3, pp. 3389–3394, 2022.

[8] Y. Hu, Y. Liu, A. Kaushik, C. Masouros, and J. S. Thompson, "Timely data collection for uav-based iot networks: A deep reinforcement learning approach," *IEEE Sensors Journal*, vol. 23, no. 11, pp. 12 295–12 308, 2023.

[9] S. Zhang, W. Liu, and N. Ansari, "Completion time minimization for data collection in a uav-enabled iot network: A deep reinforcement learning approach," *IEEE Transactions on Vehicular Technology*, vol. 72, no. 11, pp. 14 734–14 742, 2023.

[10] B. Zhu, E. Bedeer, H. H. Nguyen, R. Barton, and Z. Gao, "Uav trajectory planning for aoi-minimal data collection in uav-aided iot networks by transformer," *IEEE Transactions on Wireless Communications*, vol. 22, no. 2, pp. 1343–1358, 2022.

[11] Y. Zhu, B. Yang, M. Liu, and Z. Li, "Uav trajectory optimization for large-scale and low-power data collection: An attention-reinforced learning scheme," *IEEE Transactions on Wireless Communications*, 2023.

[12] Z. Wei, M. Zhu, N. Zhang, L. Wang, Y. Zou, Z. Meng, H. Wu, and Z. Feng, "Uav-assisted data collection for internet of things: A survey," *IEEE Internet of Things Journal*, vol. 9, no. 17, pp. 15 460–15 483, 2022.

[13] Z. Qin, X. Zhang, X. Zhang, B. Lu, Z. Liu, and L. Guo, "The uav trajectory optimization for data collection from time-constrained iot devices: A hierarchical deep q-network approach," *Applied Sciences*, vol. 12, no. 5, p. 2546, 2022.

[14] B. Khamidehi and E. S. Sousa, "Reinforcement-learning-aided safe planning for aerial robots to collect data in dynamic environments," *IEEE Internet of Things Journal*, vol. 9, no. 15, pp. 13 901–13 912, 2022.

[15] K. Messaoudi, O. S. Oubbati, A. Rachedi, A. Lakas, T. Bendouma, and N. Chaib, "A survey of uav-based data collection: Challenges, solutions and future perspectives," *Journal of network and computer applications*, vol. 216, p. 103670, 2023.

[16] S. Fu, Y. Tang, Y. Wu, N. Zhang, H. Gu, C. Chen, and M. Liu, "Energy-efficient uav-enabled data collection via wireless charging: A reinforcement learning approach," *IEEE Internet of Things Journal*, vol. 8, no. 12, pp. 10 209–10 219, 2021.

[17] M. Yi, X. Wang, J. Liu, Y. Zhang, and R. Hou, "Deep reinforcement learning for energy-efficient fresh data collection in rechargeable uav-assisted iot networks," in *2023 IEEE Wireless Communications and Networking Conference (WCNC)*. IEEE, 2023, pp. 1–6.

[18] R. Wang, D. Li, and K. Meng, "Rechargeable uav trajectory optimization for real-time persistent data collection of large-scale sensor networks," *arXiv preprint arXiv:2404.15761*, 2024.

[19] Y. Zhu and S. Wang, "Flying path optimization of rechargeable uav for data collection in wireless sensor networks," *IEEE Sensors Letters*, vol. 7, no. 2, pp. 1–4, 2023.

[20] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, "Mobile unmanned aerial vehicles (uavs) for energy-efficient internet of things communications," *IEEE Transactions on Wireless Communications*, vol. 16, no. 11, pp. 7574–7589, 2017.

[21] T. Wang, Y. Zhang, H. Shen, and G. Bai, "Task partitioning and scheduling based on stochastic policy gradient in mobile crowdsensing," *IEEE Transactions on Computational Social Systems*, 2024.

[22] M. Shao, J. Yan, and X. Zhao, "Secrecy rate maximization by cooperative jamming for uav-enabled relay system with mobile nodes," *IEEE Internet of Things Journal*, vol. 10, no. 15, pp. 13 168–13 180, 2023.

[23] D. D. Falconer, F. Adachi, and B. Gudmundson, "Time division multiple access methods for wireless personal communications," *IEEE Communications Magazine*, vol. 33, no. 1, pp. 50–57, 1995.

[24] G. Chen, X. B. Zhai, and C. Li, "Joint optimization of trajectory and user association via reinforcement learning for uav-aided data collection in wireless networks," *IEEE Transactions on Wireless Communications*, vol. 22, no. 5, pp. 3128–3143, 2022.

[25] Y. Li, A. H. Aghvami, and D. Dong, "Path planning for cellular-connected uav: A drl solution with quantum-inspired experience replay," *IEEE Transactions on Wireless Communications*, vol. 21, no. 10, pp. 7897–7912, 2022.

[26] J. Chen, D. Raye, W. Khawaja, P. Sinha, and I. Guvenc, "Impact of 3D UWB antenna radiation pattern on air-to-ground drone connectivity," in *2018 IEEE 88th Vehicular Technology Conference (VTC-Fall)*. IEEE, 2018, pp. 1–5.

[27] X. Wang, M. C. Gursoy, T. Erpek, and Y. E. Sagduyu, "Learning-based UAV path planning for data collection with integrated collision avoidance," *IEEE Internet of Things Journal*, vol. 9, no. 17, pp. 16 663–16 676, 2022.

[28] M. Sun, X. Xu, X. Qin, and P. Zhang, "Aoi-energy-aware uav-assisted data collection for iot networks: A deep reinforcement learning method," *IEEE Internet of Things Journal*, vol. 8, no. 24, pp. 17 275–17 289, 2021.

[29] C. Di Franco and G. Buttazzo, "Energy-aware coverage path planning of UAVs," in *2015 IEEE international conference on autonomous robot systems and competitions*. IEEE, 2015, pp. 111–117.

[30] C. Zhan, H. Hu, S. Mao, and J. Wang, "Energy-efficient trajectory optimization for aerial video surveillance under qos constraints," in *IEEE INFOCOM 2022-IEEE Conference on Computer Communications*. IEEE, 2022, pp. 1559–1568.

[31] Y. Sun, E. Uysal-Biyikoglu, R. D. Yates, C. E. Koksal, and N. B. Shroff, "Update or wait: How to keep your data fresh," *IEEE Transactions on Information Theory*, vol. 63, no. 11, pp. 7492–7508, 2017.

[32] X. Liang, Y. Yang, Q. Wang, W. Wang, and W. Han, "Research on anti-pursuit evasion strategy of unmanned surface vehicle based on T-D3QN," in *2022 5th International Conference on Intelligent Autonomous Systems (ICoIAS)*. IEEE, 2022, pp. 172–178.

[33] D. Han, T. Shi, T. Han, and Z. Zhou, "Joint optimization of trajectory and node access in uav-aided data collection system," *IEEE Systems Journal*, vol. 17, no. 2, pp. 2574–2585, 2023.

**SHANSHAN BAI** received the B.S. degree in Software Engineering from Changzhou University, China, in 2022, and is currently pursuing an M.S. degree in Computer Science and Artificial Intelligence at Changzhou University. Her research interests include reinforcement learning and Internet of Things."

**XUEYUAN WANG** received the B.S. degree in electrical and electronics engineering from Beijing University of Posts and Telecommunications, Beijing, China, in 2013, the M.S. degree in electrical engineering from Syracuse University, Syracuse, NY, in 2016, and the Ph.D. degree in electrical and computer engineering from Syracuse University in 2021. She is currently an instructor in the School of Computer Science and Artificial Intelligence at Changzhou University. Her primary research interests include unmanned aerial vehicles-enabled networks, 5G and beyond communications, Internet of Things networks, multi-agent joint control, and reinforcement learning.

**M. CENK GURSOY** received the B.S. degree with high distinction in electrical and electronics engineering from Bogazici University, Istanbul, Turkey, in 1999 and the Ph.D. degree in electrical engineering from Princeton University, Princeton, NJ, in 2004. He was a recipient of the Gordon Wu Graduate Fellowship from Princeton University between 1999 and 2003. He is currently a Professor in the Department of Electrical Engineering and Computer Science at Syracuse University. His research interests are in the general areas of wireless communications, information theory, communication networks, signal processing, optimization and machine learning. He is an Editor for IEEE Transactions on Communications, and an Area Editor for IEEE Transactions on Vehicular Technology. He is on the Executive Editorial Committee of IEEE Transactions on Wireless Communications. He also served as an Editor for IEEE Transactions on Green Communications and Networking between 2016–2021, IEEE Transactions on Wireless Communications between 2010–2015 and 2017–2022, IEEE Communications Letters between 2012–2014, IEEE Journal on Selected Areas in Communications - Series on Green Communications and Networking (JSAC-SGCN) between 2015-2016, Physical Communication (Elsevier) between 2010–2017, and IEEE Transactions on Communications between 2013–2018. He has been the co-chair of the 2017 International Conference on Computing, Networking and Communications (ICNC) - Communication QoS and System Modeling Symposium, the co-chair of 2019 IEEE Global Communications Conference (Globecom) - Wireless Communications Symposium, the co-chair of 2019 IEEE Vehicular Technology Conference Fall - Green Communications and Networks Track, and the co-chair of 2021 IEEE Global Communications Conference (Globecom), Signal Processing for Communications Symposium. He received an NSF CAREER Award in 2006. More recently, he received the EURASIP Journal of Wireless Communications and Networking Best Paper Award, 2020 IEEE Region 1 Technological Innovation (Academic) Award, 2019 The 38th AIAA/IEEE Digital Avionics Systems Conference Best of Session (UTM-4) Award, 2017 IEEE PIMRC Best Paper Award, 2017 IEEE Green Communications & Computing Technical Committee Best Journal Paper Award, UNL College Distinguished Teaching Award, and the Maude Hammond Fling Faculty Research Fellowship. He is a Senior Member of IEEE, and is the Aerospace/Communications/Signal Processing Chapter Co-Chair of IEEE Syracuse Section.

**GUANGQI JIANG** received the Ph.D. degree from Dalian Maritime University, Dalian, Liaoning, China. Now, she is a lecturer in Changzhou University, Changzhou, Jiangsu, China. Her research interests include computer vision and machine learning.

**SHOUKUN XU** received the B.S. degree in Mineral Processing Engineering from China University of Mining and Technology, China, in 1995, the M.S. degree in Mineral Processing Engineering from China University of Mining and Technology in 1998, and the Ph.D. degree in Mineral Processing Engineering from China University of Mining and Technology in 2001. He is currently a Professor in the School of Computer Science and Artificial Intelligence at Changzhou University, China. His research interests include computer vision and digital twins.