Tradeoff Between Age of Information and Operation Time for UAV Sensing Over Multi-Cell Cellular Networks

Cheng Zhan, *Member, IEEE*, Han Hu, *Member, IEEE*, Jing Wang, *Member, IEEE*, Zhi Liu, *Senior Member, IEEE*, and Shiwen Mao, *Fellow, IEEE*

***Abstract*—Unmanned aerial vehicles (UAVs) have a significant potential for sensing applications in further cellular networks due to their extensive coverage and flexible deployment. In this paper, we consider a multi-cell cellular network with a cellular-connected UAV, which senses data with onboard sensors and uploads sensory data to the ground base stations (BSs). To evaluate the freshness of sensory data, we employ the concept of age of information (AoI), which is defined as the time elapsed since the latest successful transmission of sensory data. A lower AoI implies fresher sensory data, which may lead to the increase of UAV operation time. To balance such tradeoff, we aim to minimize the weighted sum of operation time and total AoI for the UAV by jointly optimizing transmission scheduling, BS association, as well as UAV trajectory. The problem is formulated as a mixed-integer nonlinear programming (MINLP) problem, which is difficult to solve due to the time-varying propagation channels. To this end, we first characterize the average communication performance with statistic channel information, and then develop a search algorithm to obtain the optimal solution via employing the optimal structure as well as convex optimization techniques, while a low-complexity Double Graph based Algorithm (DGA) is developed to obtain a suboptimal solution. Then, by taking into account the site-specific performance and making fast decisions online, we propose a Deep reinforcement Learning Algorithm (DLA). Compared to DGA, DLA can adapt to the specific local environment and obtain a solution more rapidly once the training process is completed. Simulation results show that the proposed algorithms outperform the benchmarks about 30%, and achieve flexible tradeoff between operation time and AoI of UAV sensing, which is not available by considering just one objective.**

***Index Terms*—Multi-cell cellular network, UAV sensing, operation time, age of information (AoI).**

Manuscript received 5 October 2022; revised 27 March 2023; accepted 12 April 2023. Date of publication 17 April 2023; date of current version 6 March 2024. This work was supported in part by the National Natural Science Foundation of China under Grants 62172339 and 61971457, and in part by the Natural Science Foundation of Chongqing under Grant CSTB2022NSCQ-MSX0338. Recommended for acceptance by M. Rossi. *(Corresponding author: Han Hu.)*

Cheng Zhan is with the School of Computer and Information Science, Southwest University, Chongqing 400715, China (e-mail: zhanc@swu.edu.cn).

Han Hu is with the School of Information and Electronics, Beijing Institute of Technology, Beijing 100811, China (e-mail: hhu@bit.edu.cn).

Jing Wang is with the Renming University of China, Beijing 100872, China (e-mail: jwang_ruc@163.com).

Zhi Liu is with the Graduate School of Informatics and Engineering, The University of Electro-Communications, Tokyo 182-8585, Japan (e-mail: liu@ieee.org).

Shiwen Mao is with the Department of Electrical and Computer Engineering, Auburn University, Auburn, AL 36849 USA (e-mail: smao@ieee.org).

Digital Object Identifier 10.1109/TMC.2023.3267656

I. Introduction

THANKS to the advantages of high mobility and large service coverage, unmanned aerial vehicles (UAVs) are being applied in more and more civilian and commercial applications [1], [2], including precision agriculture, forest fire surveillance, aerial imaging, disaster response, monitoring social distancing for combating COVID-19 [3], etc. Due to their high altitude above the ground, UAVs are more likely to have line-of-sight (LoS) links with their ground users, which can be efficiently exploited to enhance the communication performance [4]. As the number of UAVs increases dramatically, it is more promising and cost-effective to exploit the 5G-and-beyond cellular networks to support UAV communications. In this case, UAVs can be integrated into cellular networks as cellular-connected aerial users [5], [6], [7], where high-rate air-to-ground communications and ultra-reliable beyond-visual-line-of-sight (BVLOS) operation can be achieved by cellular networks, and remote sensing with cellular-connected UAVs over wide range areas becomes more feasible.

Equipped with cameras or a multitude of sensors of various types, UAVs have been widely employed to execute multiple sensing tasks, where the sensory data is transmitted to ground base stations (BSs) over a cellular network for further processing [8], [9]. Compared with conventional terrestrial sensing, UAV sensing typically has a wider field of view due to its elevated altitude and reduced signal blockage. In addition, sensing performance can be further enhanced by introducing new degree of freedom brought about by trajectory optimization due to the high mobility of UAVs. For example, compared to the conventional ground image/video surveillance system whose cameras are deployed at fixed and pre-configured locations, aerial surveillance system (e.g., UAVs equipped with cameras) is able to fly flexibly to avoid blockage along the view path due to its high mobility, and thus the the surveillance performance can be drastically improved and even the surveillance targets can be actively tracked. Since the collected sensory data at the UAV changes rapidly with time, it is important to evaluate the performance of data freshness, i.e., age of information (AoI), in UAV sensing. To be specific, the AoI of sensing applications is defined as the time that has elapsed since the most recent successful transmission of a sensing result [10]. Several existing works have been focused on analyzing the AoI

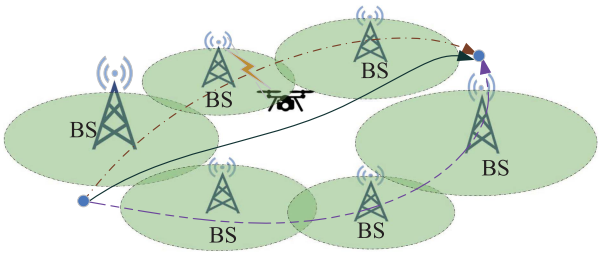


Fig. 1. Illustration of different trajectories for cellular-connected UAV sensing.

metric for UAV sensing [11], [12], [13], [14], [15]. However, the aforementioned works assumed a single cell network, while how to effectively leverage cellular-connected UAVs for remote sensing over multi-cell cellular networks has been neither well understood nor thoroughly studied.

The UAV operation time for completing specific tasks is a critical performance metric, and there are various surveillance and sensing applications which are concerned with the timeliness of the collected information and require minimizing the UAV operation time [16], [17], [18], [19]. For example, in some emergency situations, such as geological surveys, disaster investigations, industrial emergencies, and even military conflicts, gathering data requires strict deadline of UAV operation time. Specifically, to ensure that the critical information at Point-of-Interests (PoIs) in disaster areas are collected in time for aiding rescue activities, it is important to monitor PoIs as quickly as possible. As shown in Fig. 1, there exists a fundamental tradeoff between the AoI and the operation time of the UAV. Intuitively, to minimize AoI, a UAV tends to move closer to the associated BSs to achieve a smaller duration of sensory data transmission, which leads to a longer UAV operation time since it may moves farther from the target locations or destination. It is a non-trivial problem to achieve the optimal performance while balancing the AoI and operation time of the UAV. In this context, how to achieve the above tradeoff through appropriate joint design of transmission scheduling and association as well as UAV trajectory becomes one of the major challenges in UAV remote sensing over multi-cell cellular networks.

On the other hand, the BS antenna is generally designed to point downward towards ground users for good performance. As a result, aerial users will be served by the sidelobes of ground BSs where significant performance degradation may be suffered [20]. In urban areas, the building blockage generally deteriorates the local radio propagation environment. Although probabilistic LoS model is usually adopted [21], it is only meaningful for a large number of realizations of similar communication environments. In practice, for a particular environment under consideration, the existence/absence of a LoS link is in fact determined by checking whether any obstacles exist between ground BSs and the UAV, which can be obtained by environment interaction [22], [23]. In order to properly design and optimize the UAV sensing system, environment-aware realistic air-to-ground (A2G) propagation channels are essential.

To overcome the above challenges, in this paper, we propose a sensing framework with a cellular-connected UAV over multi-cell cellular networks in urban areas, which provides flexible sensing services for wide areas through adjusting UAV-BS association as well as the UAV's position dynamically. Our aim is to minimize the weighted sum of AoI and operation time of the UAV via jointly optimization of UAV trajectory and operation time as well as transmission scheduling and association, where the sensory data are transmitted to the BSs in the urban area. We summarize the main contributions of this paper as follows.

- First, we propose a novel design framework for cellular-connected UAV remote sensing over multi-cell cellular networks. To balance the tradeoff between operation time and AoI of the UAV, we minimize the weighted sum of the above two metrics and formulate the problem as a mixed-integer nonlinear programming (MINLP) problem, which is challenging to solve due to the existence of environment-aware channels.
- Second, we characterize the average communication performance and proposed two optimization based algorithms to obtain useful insights. Through utilizing the optimal structure of the problem, we first develop a search algorithm to obtain the optimal solution using convex optimization techniques. To further reduce the computational complexity, we develop a low-complexity Double Graph based Algorithm (DGA) to obtain a suboptimal solution by employing shortest path and Traveling Salesman Problem (TSP) path techniques in graph theory.
- Third, by taking into account the site-specific performance and making fast decisions online, we propose a Deep reinforcement Learning Algorithm (DLA). We decompose the problem into subproblems and reformulate each subproblem as a Markov Decision Process (MDP), based on which the UAV is trained by Double Deep Q-Network (DDQN) to be aware of its surrounding environment, and intelligently transmits sensory data to the cellular network by taking advantage of environment awareness.
- Finally, the tradeoff between operation time and total AoI for UAV sensing has been demonstrated through simulations and the impact of different weighting factors is analyzed.

The remainder of this paper is organized as follows. Section II discusses the related work. Section III establishes UAV sensing models with the practical A2G channel model. Section IV formulates and solves the weighted sum minimization problem for average communication performance. Section V introduces the Deep Reinforcement Learning (DRL) based solution framework for site-specific performance. Section VI presents extensive simulation results and analyzes the system performance. Finally, conclusions are given in Section VII.

II. Related Work

In the literature, there are flurry of works on UAV enabled sensing, which can be mainly divided into two categories: UAVs serving as data collection platforms for wireless sensor networks

(WSNs) [24], and UAVs directly acting as aerial sensing platforms to provide wireless sensing service. In the first scenario, the UAVs serve as aerial communication platforms, such as data collectors or data fusion center, to collect sensory data generated from ground sensor nodes. The work in [25] studied the age-optimal data collection problem for UAV-enabled WSNs, where the AoI was defined as a weighted sum of sensor nodes' data uploading time and the corresponding flight time. The fine-grained trajectory plan for multi-UAVs was investigated in [26] to minimize the maximum flight time of UAVs for data collection. An AoI-optimal trajectory planning in UAV-assisted wireless sensor networks was studied in [27], which consists of a clustering module and a neural trajectory solver. The work in [28] proposed a time-efficient data collection scheme, where sensory data from multiple ground devices were uploaded to the UAV via uplink non-orthogonal multiple access (NOMA). The work in [29] studied the minimum UAV deployment problem to find the data collection tours for multiple UAVs to collect data from ground devices within a delay threshold. However, UAVs were only considered as data collection platforms in the above works.

The other emerging scenario is that the UAVs equipped with various sensors are regarded as aerial nodes to directly provide wireless sensing support from the sky, such as image/video surveillance, etc. The work in [30] studied UAV enabled surveillance of uneven surface to achieve a maximal compact coverage, where a centralized algorithm and two distributed algorithms were proposed. The work in [19] minimized the maximum time spent by the UAVs for disaster area surveillance through developing approximation algorithm. The work in [31] developed a UAV-assisted surveillance framework by utilizing random walks with consideration of battery constraints of the UAVs. The deployment of UAVs for anisotropic monitoring tasks was investigated in [32], where the anisotropy of monitoring angle was taken into account. The work in [33] studied the problem of UAV-assisted multi-task allocation for mobile crowd sensing to maximize sensing coverage with deep reinforcement learning. However, the air-to-ground sensory data transmission to BSs have not been captured in the above works.

Cellular-connected UAVs have been increasingly considered for UAV sensing due to their operability and applicability to UAV operations over wide areas [34], where UAVs act as aerial users in cellular networks and the sensory data are transmitted to ground BSs by leveraging the 5G high-speed wireless infrastructures. AoI was minimized in [11] by joint design of sensing time, transmission time, UAV trajectory, as well as task scheduling. The work in [12] considered underlay UAV-to-device communications and studied the AoI minimization problem over a cellular internet of UAVs via trajectory design. AoI-driven quality of service (QoS) provisioning schemes over UAV sixth generation (6G) multimedia mobile networks was proposed in [13] to efficiently support massive ultra-reliable and low latency communications. The work in [14] studied delay-sensitive energy-efficient UAV crowdsensing to maximize the data collection ratio from PoIs, while keeping data freshness and minimizing energy consumption of all UAVs. The work in [15] studied the framework of image surveillance UAVs for relay communication between ground users and a remote BS. However, the aforementioned works only considered a single cell network. In this work, different from the aforementioned works, we study the UAV sensing over multi-cell cellular networks in urban areas, where flexible tradeoff between AoI and operation time for UAV sensing has been investigated.

III. System Model

In this paper, we consider a multi-cell cellular network enabled UAV sensing scenario, which consists of a target sensing region $\hat{\mathcal{S}} \subseteq \mathbb{R}^{2\times 1}$, $M > 1$ ground BSs denoted by set $\mathcal{M} = \{s_1, \ldots, s_M\}$, and one UAV embedded with sensors and communication devices. The UAV collects various required data (e.g., aerial surveillance) from $\hat{\mathcal{S}}$ with their sensors in each time instant [8], where the sensing data are transmitted to the ground BSs for further processing or delivery. The UAV is assumed to fly at a constant altitude of H since frequent descending and ascending are energy-inefficient [35]. The horizontal location of the UAV in time instant t is denoted by $\{\mathbf{u}(t)|\mathbf{u}(t) \in \mathbb{R}^{2\times 1}, 0 \le t \le T\}$, where T denotes the total time horizon of the UAV flight, also named as UAV operation time (or mission completion time). All BSs are assumed to have the same height of H_G, $H_G \ll H$, where $\mathbf{g}_m \in \mathbb{R}^{2\times 1}$ denotes the horizontal coordinate of BS s_m, $1 \le m \le M$. There are K important target locations (e.g., well-known attractions for aerial filming) in the sensing region $\hat{\mathcal{S}}$, which are required to be reached during the flight [36], denoted by $\mathcal{K} = \{\rho_1, \ldots, \rho_K\}$. Denote $\mathbf{w}_k \in \mathbb{R}^{2\times 1}$ as the horizontal coordinate of ρ_k, $1 \le k \le K$. The mission of the UAV is to sense and transmit data over region $\hat{\mathcal{S}}$ from an initial location $\mathbf{u}_I$ to a final location $\mathbf{u}_F$, while accessing all important target locations in $\mathcal{K}$. For example, the aerial filming not only film the target locations in $\mathcal{K}$, but also the region along its flight [15]. At the end of the mission, the whole aerial film can be generated. Another example is the aerial virtual reality (VR) applications with cellular-connected UAV [37], where the VR users can experience a aerial view flight along target locations in $\mathcal{K}$, and the UAV's vision is transmitted to BSs all the time. Thus, we have the target constraint

$$\{\mathbf{w}_k|1 \le k \le K\} \subseteq \{\mathbf{u}(t)|t \in [0, T]\}, \tag{1}$$

In practice, $\mathbf{u}_I$ and $\mathbf{u}_F$ may correspond to different charging stations or the target locations for its pre- and post-mission. As such, the UAV operation time T corresponds to the total time horizon for the UAV flying from $\mathbf{u}_I$ to $\mathbf{u}_F$, while accessing all important target locations in $\mathcal{K}$, which is a design variable in this paper. We assume that the UAV battery capacity is enough for the UAV to visit all target points. Note that our design framework can also be extended to the multi-UAV scenario with different missions by separating different UAVs in operation time or flying altitude to avoid collision. Let $\mathbf{v}(t) \triangleq \dot{\mathbf{u}}(t)$ be the UAV velocity at time t, with $\|\mathbf{v}(t)\| \le V_{\max}, \forall t \in [0, T]$ where $V_{\max}$ is the maximum speed due to the mechanical limitation. The distance

between BS s_m and the UAV at time t is calculated as $d_m(t) = \sqrt{(H - H_G)^2 + \|\mathbf{u}(t) - \mathbf{g}_m\|^2}$.

A. Channel Model

Through A2G channels, the data sensed by the UAV are transmitted to the ground BSs. The UAV is assumed to be assigned to a dedicated subchannel without inter-cell interference [37], [38]. As we consider UAV sensing for urban environments, buildings will affect the radio propagation by blocking LoS paths. We assume that the buildings are distributed in the considered area, where the locations and heights of buildings follow the model suggested by the International Telecommunication Union (ITU) [39]. With a given building distribution, the UAV will have an unobstructed LoS link with BS s_m at time t if there exist no buildings intersecting the straight line between horizon location $\mathbf{g}_m$ at height H_G and $\mathbf{u}(t)$ at height H. Otherwise, if there is at least one building that intersects with this line, then the A2G channel is non-Line-of-Sight (NLoS).

Assuming that the complete channel state information (CSI) for the A2G channels in the considered area is not known a priori, while the UAV can estimate the instantaneous CSI by receiving signals from the BSs within its communication coverage or leveraging the existing handover mechanisms with continuous reference signal received power (RSRP) measurements [23]. We characterize the A2G channels by both large-scale and small-scale fading. Denote $\beta_m(t)$ as the large-scale channel gain at time t between BS s_m and the UAV, which can be represented by the LoS and NLoS components. Specifically, $\beta_m(t) = \beta_0 d_m(t)^{-\alpha}$ for LoS links while $\beta_m(t) = \mu\beta_0 d_m(t)^{-\alpha}$ for NLoS links as in [22], [40], where $\alpha \geq 2$ denotes the path loss exponent, β_0 is the channel gain at the reference distance 1 m, and $\mu < 1$ denotes the additional attenuation factor which is brought from NLoS propagation. The small-scale fading $\tilde{h}_m(t)$ for the A2G link between the UAV and s_m at time t can be modelled as Rayleigh fading for the NLoS case and Rician fading for the LoS case with Rician factor K_c [41], where $\mathbb{E}[|\tilde{h}_m(t)|^2] = 1$.

Typically, ground BSs are equipped with downtilted antennas to ensure good performance of ground users. We adopt a practical BS antenna radiation pattern with downtilt angle $\phi_D \in [0°, 90°]$, and each BS is assumed to be equipped with a vertical uniform linear array (ULA) with n_0 elements [16], [42]. Then the power gain of each antenna element at time t along the direction between the UAV and BS s_m is calculated by $G_e(\phi_m(t)) \triangleq -\min\left(12\left(\frac{\phi_m(t)}{HPBW_v}\right)^2, G_0\right)$, where $G_e(\phi_m(t))$ is measured in dB and $\phi_m(t) \triangleq \arcsin\left(\frac{H-H_U}{d_m(t)}\right)$ represents the elevation angle between the BS s_m and UAV at time t, $HPBW_v$ and G_0 denote the half power beamwidth and antenna nulls threshold, respectively. As derived in [16], the array factor at time t for the ULA on BS s_m is calculated by $A_f(\phi_m(t)) = \frac{1}{\sqrt{N_0}} \frac{\sin(\frac{n_0\pi}{2}(\sin\phi_m(t) - \sin\phi_D))}{\sin(\frac{\pi}{2}(\sin\phi_m(t) - \sin\phi_D))}$. As a result, the overall antenna gain can be given by $G_m(t) = 10^{\frac{G_e(\phi_m(t))}{10}} A_f(\phi_m(t))^2$, and the baseband equivalent channel between the UAV and s_m at time t, denoted by $h_m(t)$, can be expressed as $h_m(t) = \sqrt{G_m(t)\beta_m(t)}\tilde{h}_m(t)$.

B. Data Transmission

Define $x_m(t) \in \{0, 1\}$ as the data transmission scheduling indicator, where $x_m(t) = 1$ if the UAV transmits the sensed data to BS s_m at time t and $x_m(t) = 0$ otherwise. We assume that at most one ground BS is scheduled for reception at each time instant t, then we have $\sum_{m=1}^{M} x_m(t) \leq 1, t \in [0, T]$, where the BS association is also implied in transmission scheduling indicator. Denote $R_m(t)$ as the achievable rate from the UAV to BS s_m if scheduled at time t, then $R_m(t)$ is calculated as

$$R_m(t) = B\log_2\left(1 + \frac{P|h_m(t)|^2}{\sigma^2}\right), \tag{2}$$

where σ^2 and B denote the noise power and channel bandwidth, respectively, and P denotes the UAV's maximum transmit power.

We consider the real-time transmission with just-in-time transmission policy where data is required to be transmitted immediately after it is generated from the sensors on the UAV. In particular, we assume that the UAV generates sensory data with a sensing rate of R_s in each time instant, while the UAV is required to finish the data transmission process concurrently with the data sensing process to keep the sensory data up-to-date for further processing at BSs [43]. In practice, R_s is determined by the size of UAV sensing data, which depends on intrinsic proprieties such as the resolution of the sensors on the UAV. As a result, the transmission rate of the UAV should be no less than its sensing rate, i.e.,

$$\sum_{m=1}^{M} x_m(t)R_m(t) \geq R_s, \forall t, \tag{3}$$

We assume that the above immediate sensory data transmission happens perfectly once (3) is satisfied; otherwise, the immediate data transmission fails. The corresponding problem for non-real-time transmission will be left as our future work. However, the A2G channel quality are highly dependent on the distance between the UAV and BSs, as well as the channel randomness, which is brought from random small-scale fading and building blockages, and thus the constraints in (3) may not be always satisfied. To tackle such an issue, we adopt the AoI to measure how timely the transmission of sensory data is, i.e., the freshness of the sensory data [44]. It is worthwhile to note that the sensory data is only generated from the sensors or cameras on the UAV (e.g., video surveillance), and data is transmitted immediately after it is generated from the UAV that is successfully received once (3) is satisfied. Denote $\lambda(t)$ as the latest time t that the immediate data transmission of the UAV is successfully received at BSs, i.e., $\lambda(t) \triangleq \max_\tau\{\tau | \sum_{m=1}^{M} x_m(\tau)R_m(\tau) \geq R_s, \tau \in [0, t]\}$. Thus, for each time instant t, the time duration $t - \lambda(t)$ can be regarded as the time elapsed since the latest time when *freshest* information is successfully received, which specifies the age of such freshest information received and thus we refer it as AoI. As such, denote the AoI of the UAV at time t as $A(t)$, which can be defined as

$$A(t) \triangleq t - \lambda(t). \tag{4}$$

Thus, we have $\lambda(t) \leq t$, $A(t) \geq 0, \forall t$. If $\lambda(t) = t$, then the UAV transmits the sensory data on time, and AoI $A(t) = 0$; otherwise,

the AoI $A(t)$ will increase with time. Therefore, the total AoI throughout the UAV mission can be calculated as $\int_0^T A(t)dt$. In this paper, we aim to minimize the total AoI to reduce the overall effect of the disconnected events/durations for freshness of sensory data.

C. Problem Formulation

Note that the UAV makes its sensory data fresher with a smaller AoI. To ensure the freshness of sensory data, the UAV should transmit data to ground BSs as quickly as possible using a clean channel such that the AoI can be kept small. On the other hand, the UAV operation time T should be minimized to increase the operation efficiency. Intuitively, with a larger T, the UAV has more flexibility to move closer to ground BSs for better channel quality, leading to a timely transmission of sensory data and a smaller AoI. Thus, a fundamental tradeoff exists between minimizing UAV operation time T and minimizing the total AoI $\int_0^T A(t)dt$. To balance such a tradeoff, in this paper, we associate each of them with a weighting factor and minimize the weighted sum, by optimizing UAV trajectory $\{\mathbf{u}(t)\}$ and transmission scheduling $\{x_m(t)\}$ as well as UAV operation time T, subject to the target location constraint and UAV's mechanical constraints. The optimization problem is formulated as

$$(\text{P1}): \min_{T,\{\mathbf{u}(t)\},\{x_m(t)\}} \theta T + (1-\theta)\int_0^T A(t)dt$$

$$\text{s.t. } x_m(t) \in \{0,1\}, \forall m, t, \tag{5}$$

$$\sum_{m=1}^{M} x_m(t) \le 1, \forall t, \tag{6}$$

$$\{\mathbf{w}_k | 1 \le k \le K\} \subseteq \{\mathbf{u}(t) | t \in [0,T]\}, \tag{7}$$

$$\|\mathbf{v}(t)\| \le V_{\max}, \forall t, \tag{8}$$

$$\mathbf{u}(0) = \mathbf{u}_I, \mathbf{u}(T) = \mathbf{u}_F, \tag{9}$$

where θ and $1-\theta$, $0 \le \theta \le 1$, denote the weights of the operation time and the total AoI of the UAV, respectively. The tradeoff between T and $\int_0^T A(t)dt$ can be obtained by solving problem (P1) for a given value of θ. With a large θ, the UAV is expected to navigate through the target points where the AoI issue is almost ignored. Otherwise, the UAV tends to keep connectivity with the BS for better channel quality, leading to a smaller AoI.

Note that the optimal solution to problem (P1) is difficult to obtain, since (P1) consists of continuous optimization variables $\{\mathbf{u}(t)\}$ and binary variables $\{x_m(t)\}$. Furthermore, (P1) is an MINLP which contains the complicated target location constraint and an integral upper limited by a design variable T. In addition, an accurate channel coefficient $h_m(t)$ is usually unavailable due to the channel randomness and the frequent switching between LoS/NLoS connections during a flight.

To address such difficulties, we first consider the average communication performance with statistic channel information and obtain a statistically favorable formulation in Section IV, based on the expected communication rate over the probabilistic LoS channel model. By analyzing the optimal structure, we develop a search algorithm to obtain the optimal solution and a low-complexity DGA to obtain a suboptimal solution by employing shortest path and TSP path techniques. Next, we address the site-specific performance for a specific local environment in Section V, and the UAV acts as an agent to interact with the environment by CSI measurement. We propose a DLA to learn from the specific local environment and make fast decisions. In particular, the sampled rate measurements by the UAV is used as input, based on which a function that maps the input local environment to the output flying decisions can be learnt with a DDQN. Note that the DDQN-based offline model training can be conducted at the UAV control center or BSs, which usually have a more powerful computation capability. Then the well-trained learning model can be transferred to and executed on the UAV to perform model inference locally and make fast decisions.

IV. Proposed Graph Based Algorithm

In this section, we first characterize the average communication performance with statistic channel information, and then reformulate the problem into a more tractable form by analyzing the structure of the optimal solution. We then obtain the optimal and low-complexity suboptimal solutions by employing the graph based algorithms.

A. Problem Reformulation

We adopt the probabilistic LoS channel model [21] to characterize the average communication performance under building blockages over a large number of similar communication environments, which is assumed to be known in advance. In particular, the LoS probability at time t between ground BS s_m and the UAV is denoted by $\mathbb{P}_m^L(t) = \frac{1}{1+a\exp(-b[|\phi_m(t)|-a])}$, where a and b are two environment dependent parameters, and $\phi_m(t)$ is the elevation angle at time t. As a result, $\beta_m(t) = \beta_0 d_m(t)^{-\alpha}$ for LoS link with probability $\mathbb{P}_m^L(t)$ while $\beta_m(t) = \mu\beta_0 d_m(t)^{-\alpha}$ for NLoS link with probability $\mathbb{P}_m^N(t) \triangleq 1 - \mathbb{P}_m^L(t)$. As such, we obtain $\mathbb{E}[|h_m(t)|^2] = G_m(t)\hat{P}_m^L(t)\beta_0 d_m(t)^{-\alpha}$, where $\hat{P}_m^L(t) \triangleq \mathbb{P}_m^L(t) + (1-\mathbb{P}_m^L(t))\mu$ is a regularized LoS probability with additional factor μ for NLoS occurrence.

Since the channel gains $h_m(t)$ are random variables, the channel achievable rates $R_m(t)$ are also random variables. Thus we are interested in the expected achievable rate, i.e., $\mathbb{E}[R_m(t)]$, whose closed form expressions are also difficult to obtain due to the difficulty of deriving its probability distribution. To address such issue, we obtain its approximation based on the following result from [45].

Proposition 1. ([45, Theorem 1]): The approximation $\mathbb{E}[\log_2(1+\frac{X}{Y})] \approx \log_2(1+\frac{\mathbb{E}[X]}{\mathbb{E}[Y]})$ holds, where X and Y are independent random variables, $X \ge 0, Y > 0$.

Let $X = P|h_m(t)|^2$ and $Y = \sigma^2$. By applying Proposition 1 over $\mathbb{E}[R_m(t)]$, we obtain $\mathbb{E}[R_m(t)] \approx B\log_2(1+\frac{P\mathbb{E}[|h_m(t)|^2]}{\sigma^2}) = B\log_2(1+\frac{PG_m(t)\hat{P}_m^L(t)\beta_0}{\sigma^2 d_m(t)^\alpha})$. Note that $\hat{P}_m^L(t)$ is still a complicated functions with UAV trajectory $\mathbf{u}(t)$, and it is difficult to tackle. By applying homogeneous approximation approach as in [40], [41], we let $\hat{P}_m^L(t) \approx \bar{P}_m^L$ and $G_m(t) \approx \bar{G}_m$, where $\bar{P}_m^L$ and $\bar{G}_m$ are determined as the average value for the

TSP path visiting the locations in $\mathcal{K}$ exactly once from $\mathbf{u}_I$ to $\mathbf{u}_F$ due to the target location constraint (7), and this way, a satisfactory approximation accuracy can be guaranteed [41]. Thus, we have $\mathbb{E}[R_m(t)] \approx B\log_2\left(1+\frac{P\bar{G}_m\bar{P}_m^L\beta_0}{\sigma^2 d_m(t)^\alpha}\right) \triangleq \bar{R}_m(t)$, and the UAV's total AoI can be approximated as $\int_0^T (t-\bar{\lambda}(t))dt$, where $\bar{\lambda}(t) \triangleq \max_\tau\{\tau|\tau\in[0,t], \sum_{m=1}^M x_m(\tau)\bar{R}_m(\tau) \geq R_s\}$. As such, the transmission scheduling can be approximately obtained as

$$x_m(t) = \begin{cases} 1, & m = \arg\max_{m'\in\{1,2,\ldots,M\}} \bar{R}_{m'}(t), \\ 0, & \text{otherwise}, \end{cases} \tag{10}$$

and $\bar{\lambda}(t)$ is written as $\bar{\lambda}(t) = \max\{\tau|\tau\in[0,t], \max_m \bar{R}_m(\tau) \geq R_s\}$.

It can be shown that the QoS requirement $\max_m \bar{R}_m(\tau) \geq R_s$ is equivalent to the constraint on the distance between the UAV and its closest BS, i.e., $\min_m \|\mathbf{u}(t)-\mathbf{g}_m\| \leq \bar{d}_m$, where $\bar{d}_m \triangleq \sqrt{(\frac{\gamma_m}{2^{R_s/B}-1})^{\frac{2}{\alpha}} - (H-H_G)^2}$ and $\gamma_m = \frac{P\bar{G}_m\bar{P}_m^L}{\sigma^2}$ denotes the reference signal-to-noise ratio (SNR) at 1 m. As a result, the UAV selects the closest BS for sensory data transmission at each time t. Note that the above approximation solution for transmission scheduling is only applied for the average communication performance, which is not suitable for the site-specific performance with given a specific local environment. Define *service area* $\mathcal{A}_m \triangleq \{\mathbf{u}|\mathbf{u}\in\mathbb{R}^{2\times 1}, \|\mathbf{u}-\mathbf{g}_m\| \leq \bar{d}_m\}$ for ground BS s_m, which is exactly a disk region on the horizontal plane centered at $\mathbf{g}_m$ with radius $\bar{d}_m$. The QoS requirement $\max_m \bar{R}_m(\tau) \geq R_s$ can always be satisfied as long as the UAV's horizontal position lies in the region $\hat{\mathcal{A}} \triangleq \bigcup_{m=1}^M \mathcal{A}_m$. If the UAV is out of region $\hat{\mathcal{A}}$, then no immediate sensory data transmission happen and thus the AoI of the UAV will increase.

We denote J, $J \leq M$, as the number of service areas that the UAV flies over during the mission, and the corresponding BSs set is denoted as $\mathcal{J} = \{s_{\omega_1}, s_{\omega_2}, \ldots, s_{\omega_J}\} \subseteq \mathcal{M}$, where $|\mathcal{J}| = J$ and $\omega_i \in \{1,\ldots,M\}$ is the BS index in $\mathcal{M}$, $1 \leq i \leq J$. Denote $t^s_{\omega_i}$ and $t^l_{\omega_i}$ as the time instants that the UAV starts to enter and leave service area $\mathcal{A}_{\omega_i}$, respectively. Thus, we can obtain the visiting order of $\mathcal{J}$, $\boldsymbol{\pi} = (\pi_1, \pi_2, \ldots, \pi_J)$ by re-arranging $\{\omega_i\}$ with increasing $t^s_{\omega_i}$, which is exactly a permutation of $(\omega_1, \omega_2, \ldots, \omega_J)$. Thus, we have $t^s_{\pi_i} \leq t^l_{\pi_i} \leq t^s_{\pi_{i+1}} \leq t^l_{\pi_{i+1}}, 1 \leq i \leq J-1$, and $\|\mathbf{u}(t)-\mathbf{g}_{\pi_i}\| \leq \bar{d}_{\pi_i}, t\in[t^s_{\pi_i}, t^l_{\pi_i}]$. Due to the definition of AoI, we have $\bar{\lambda}(t) = t$ when $t\in[t^s_{\pi_i}, t^l_{\pi_i}]$, $1\leq i\leq J$, and $\bar{\lambda}(t) = t^l_{\pi_{i-1}}$ when $t\in[t^l_{\pi_{i-1}}, t^s_{\pi_i}]$, $2\leq i\leq J$. As a result, the total AoI of the UAV can be expressed as

$$\int_0^T (t-\bar{\lambda}(t))\,dt = \sum_{i=1}^{J+1} \frac{(t^s_{\pi_i} - t^l_{\pi_{i-1}})^2}{2}, \tag{11}$$

where we define $t^l_{\pi_0} \triangleq 0$ and $t^s_{\pi_{J+1}} \triangleq T$. Based on the above discussions, problem (P1) can be reformulated in the following equivalent form.

$$\text{(P2)}: \min_{T,\mathcal{J},\boldsymbol{\pi},\{\mathbf{u}(t)\},\{t^s_{\pi_i},t^l_{\pi_i}\}} \theta T + (1-\theta)\sum_{i=1}^{J+1}\frac{(t^s_{\pi_j} - t^l_{\pi_{j-1}})^2}{2}$$

$$\text{s.t. } \mathcal{J}\subseteq\mathcal{M}, \tag{12}$$

$$\left\|\mathbf{u}(t)-\mathbf{g}_{\pi_j}\right\| \leq \bar{d}_{\pi_j}, t\in[t^s_{\pi_j}, t^l_{\pi_j}], i=1,\ldots,J, \tag{13}$$

$$t^s_{\pi_{j-1}} \leq t^l_{\pi_{j-1}} \leq t^s_{\pi_j} \leq t^l_{\pi_j}, 2\leq j\leq J,$$

$$(7)\text{–}(9). \tag{14}$$

Theorem 1: Problem (P2) is NP-hard.

Proof: We show the proof by a reduction from TSP, which is a well known NP-hard problem [46]. We consider a special case of problem (P2) by assuming that the service area of each BS is sufficiently large, such that the UAV lies in service areas of all BSs during its flight. In this case, the UAV can transmit the freshest sensory data to any BS and the total AoI of the UAV is always zero. Problem (P2) in this special case reduces to a UAV operation time minimization problem under the target location constraint (7), where the UAV should reach all the target locations in $\mathcal{K}$. To minimize the UAV operation time, the UAV is required to fly at the maximum speed and all the target locations in $\mathcal{K}$ are visited exactly once. Thus, this special problem is equivalent to the TSP with given initial and end locations [47], which is to find the shortest possible route that visits each city exactly once with a given list of cities and the distance between each pair of cities. By adding a dummy city with its distance to the initial and end locations set to 0, while that to all other cities set to a sufficiently large value, TSP with given initial and end locations is equivalent to the standard TSP. Hence, problem (P2) is NP-hard. ■

In the following, we first analyze the structure of the optimal solution to (P2), based on which both optimal and low-complexity suboptimal solutions are derived by employing graph theory.

B. Proposed Solution

To gain more insights into problem (P2), we consider a *special case* when $\mathcal{K} = \emptyset$, i.e., $K = 0$, where there exist no important target locations that the UAV needs to visit. The above special case of (P2) is named as *problem (P2-s)*. Define the UAV waypoints $\mathbf{u}^s_{\pi_j} \triangleq \mathbf{u}(t^s_{\pi_j})$, $\mathbf{u}^l_{\pi_j} \triangleq \mathbf{u}(t^l_{\pi_j})$, $1\leq j\leq J$. It is easy to prove by contradiction that the optimal UAV trajectory to (P2-s) can be assumed to contain line segments connecting $\mathbf{u}_I, \mathbf{u}^s_{\pi_1}, \mathbf{u}^l_{\pi_1}, \ldots, \mathbf{u}^s_{\pi_J}, \mathbf{u}^l_{\pi_J}, \mathbf{u}_F$ and the UAV flies at the maximum speed. As such, with given $\mathcal{J}$ and $\boldsymbol{\pi}$, problem (P2-s) can be written as the following problem (P3), where $\mathbf{u}^l_{\pi_0} \triangleq \mathbf{u}_I$ and $\mathbf{u}^s_{\pi_{J+1}} \triangleq \mathbf{u}_F$.

$$\text{(P3)}: \min_{\{\mathbf{u}^s_{\pi_j},\mathbf{u}^l_{\pi_j}\}} \frac{\theta}{V_{\max}}\left(\sum_{j=1}^{J}\left\|\mathbf{u}^l_{\pi_j}-\mathbf{u}^s_{\pi_j}\right\| + \sum_{j=1}^{J+1}\left\|\mathbf{u}^s_{\pi_j}-\mathbf{u}^l_{\pi_{j-1}}\right\|\right)$$

$$+\frac{(1-\theta)}{2V^2_{\max}}\sum_{j=1}^{J+1}\left\|\mathbf{u}^s_{\pi_j}-\mathbf{u}^l_{\pi_{j-1}}\right\|^2$$

$$\text{s.t. } \left\|\mathbf{u}^s_{\pi_j}-\mathbf{g}_{\pi_j}\right\| \leq \bar{d}_{\pi_j}, \left\|\mathbf{u}^l_{\pi_j}-\mathbf{g}_{\pi_j}\right\| \leq \bar{d}_{\pi_j},$$

$$j=1,\ldots,J. \tag{15}$$

Note that (P3) is a standard convex optimization problem, which can be solved with CVX [48]. As a result, the optimal solution to (P2-s) can be obtained by an exhaustively search over all possible subsets $\mathcal{J} \subseteq \mathcal{M}$ and the visiting order $\boldsymbol{\pi}$ of each $\mathcal{J}$, and then solving (P3) to determine the minimum objective value. However, searching all possible subsets of $\mathcal{J}$ has an exponential complexity of $O(2^M)$, which is infeasible for large values of M. Therefore, we propose an efficient suboptimal solution to (P2-s) by using a graph based algorithm. With given BS location $\{\mathbf{g}_m\}$ and corresponding service area radius $\{\bar{d}_m\}$, we define a weighted graph $\hat{G}(\hat{V}, \hat{E}, \hat{w})$ as follows, and obtain the following proposition.

Definition 1: $\hat{G}(\hat{V}, \hat{E}, \hat{w})$ is constructed by

- $\hat{V} \triangleq \{\hat{v}_0, \hat{v}_1, \ldots, \hat{v}_M, \hat{v}_{M+1}\}$. We introduce a vertex $\hat{v}_m$ for the location of each BS s_m, i.e., $\mathbf{g}_m$, $1 \le m \le M$, where $\hat{v}_o$ and $\hat{v}_{M+1}$ represent the initial and final locations, respectively.
- $\hat{E} \triangleq \{(\hat{v}_i, \hat{v}_j) | 0 \le i \ne j \le M+1\}$. For any two different vertices $\hat{v}_i$ and $\hat{v}_j$, $\hat{v}_i \in \hat{V}$, $\hat{v}_j \in \hat{V}$, $i \ne j$, there exists an edge $(\hat{v}_i, \hat{v}_j)$.
- $\hat{w} : \hat{E} \to \mathbf{R}^+$, where $\hat{w}$ is a weight function. In particular, $\hat{w}(\hat{v}_i, \hat{v}_j) = \frac{\theta}{V_{\max}} \|\mathbf{g}_i - \mathbf{g}_j\| + \frac{(1-\theta)}{2V_{\max}^2} \|\mathbf{g}_i - \mathbf{g}_j\|^2, 0 \le i \ne j \le M+1$, where $\mathbf{g}_0$ and $\mathbf{g}_{M+1}$ denote the initial and final locations, respectively.

Proposition 2: The objective value of (P2-s) is upper bounded by the weight of the shortest path between vertex $\hat{v}_o$ and $\hat{v}_{M+1}$ over graph $\hat{G}$.

Proof: Considering a special case of (P2-s) with $\bar{d}_{\pi_j} = 0, \forall j$, which we call *problem (P2-ss)*. Then $\mathbf{u}^s_{\pi_j} = \mathbf{u}^l_{\pi_j} = \mathbf{g}_{\pi_j}, \forall j$. Due to Definition 1, it is easy to see that problem (P2-ss) is equivalent to finding the shortest path over graph $\hat{G}$, and thus the optimal value to (P2-ss) equals to the weight of the shortest path $\vec{p}^*$ between vertex $\hat{v}_0$ and $\hat{v}_{M+1}$, i.e., $\mathcal{J}^* = \{\hat{v}_i | \hat{v}_i \in \vec{p}^*\}$, while $\boldsymbol{\pi}^*$ can be obtained by rearranging the vertices in $\mathcal{J}^*$ with the order appearing in $\vec{p}^*$. In addition, the optimal solution to (P2-ss) is also a feasible solution to (P2-s), since (P2-ss) is a special case of (P2-s). Thus the optimal objective value of (P2-ss) serves as an upper bound of that of (P2-s), which concludes the proof. □

Note that Dijkstra algorithm can be used to find the shortest path from $\hat{v}_0$ to $\hat{v}_{M+1}$ in $\hat{G}$ with complexity $O(M^2)$ [49]. Thus, a suboptimal solution to (P2-s) can be obtained via finding the shortest path in $\hat{G}$ using Dijkstra algorithm to determine the serving BS set $\mathcal{J}$ as well as the visiting order $\boldsymbol{\pi}$, and solving the convex optimization problem (P3) to determine the waypoints $\{\mathbf{u}^s_{\pi_j}, \mathbf{u}^l_{\pi_j}\}$. It is worthwhile to note that such suboptimal solution is general for any initial location $\tilde{\mathbf{u}}_I$ and final location $\tilde{\mathbf{u}}_F$.

Let us consider the *general case* when $\mathcal{K} \ne \emptyset$, i.e., $K > 0$, where there exist K important target locations that the UAV needs to visit. Then the UAV trajectory can be partitioned into $K+1$ segments, with $\boldsymbol{\Pi} = (\Pi_1, \ldots, \Pi_K)$ denoting the visiting order of $\mathcal{K}$. Specifically, the lth segment starts at $\mathbf{w}_{\Pi_{l-1}}$ and ends at $\mathbf{w}_{\Pi_l}$, $1 \le l \le K+1$, where $\mathbf{w}_{\Pi_0} \triangleq \mathbf{u}_I$ and $\mathbf{w}_{\Pi_{K+1}} \triangleq \mathbf{u}_F$. Denote the optimal UAV trajectory for the lth segment as $\{\mathbf{u}^*_l(t)\}$ and the minimum weighted sum of operation time and total AoI along the lth segment is denoted by Λ^*_l. Thus,

Algorithm 1: Search Algorithm for Problem (P2).

1: $\Lambda^* \leftarrow \infty$;
2: **for** each permutation $\boldsymbol{\Pi}$ of the target locations in $\mathcal{K}$ **do**
3: **for** $l = 1$ to $l = K+1$ **do**
4: Set initial location $\tilde{\mathbf{u}}_I = \mathbf{w}_{\Pi_{l-1}}$ and final location $\tilde{\mathbf{u}}_F = \mathbf{w}_{\Pi_l}$. Let $\Lambda^*_{\Pi,l} \leftarrow \infty$;
5: **for** each subset $\mathcal{J} \subseteq \mathcal{M}$ **do**
6: **for** each visiting order $\boldsymbol{\pi}$ for $\mathcal{J}$ **do**
7: With given $\tilde{\mathbf{u}}_I$, $\tilde{\mathbf{u}}_F$, $\mathcal{J}$, and $\boldsymbol{\pi}$, solve the standard convex optimization problem (P3) with CVX to obtain the objective value $\Lambda_{\Pi,l}$ and waypoints $\{\mathbf{u}^s_{\pi_j}, \mathbf{u}^l_{\pi_j}\}$;
8: **if** $\Lambda_{\Pi,l} < \Lambda^*_{\Pi,l}$ **then**
9: $\Lambda^*_{\Pi,l} \leftarrow \Lambda_{\Pi,l}$, $\mathcal{J}^*_{\Pi,l} \leftarrow \mathcal{J}$, $\boldsymbol{\pi}^*_{\Pi,l} \leftarrow \boldsymbol{\pi}$, $\mathcal{U}^*_{\Pi,l} \leftarrow \{\mathbf{u}^s_{\pi_j}, \mathbf{u}^l_{\pi_j}\}$;
10: **end if**
11: **end for**
12: **end for**
13: **end for**
14: **if** $\sum_{l=1}^{K+1} \Lambda^*_{\Pi,l} < \Lambda^*$ **then**
15: $\boldsymbol{\Pi}^* \leftarrow \boldsymbol{\Pi}$, $\Lambda^* \leftarrow \sum_{l=1}^{K+1} \Lambda^*_{\Pi,i}$, $\mathcal{J}^*_\Pi \leftarrow \{\mathcal{J}^*_{\Pi,l}\}$, $\boldsymbol{\pi}^*_\Pi \leftarrow \{\boldsymbol{\pi}^*_{\Pi,l}\}$, $\mathcal{U}^*_\Pi \leftarrow \{\mathcal{U}^*_{\Pi,l}\}$;
16: **end if**
17: **end for**
18: Construct the optimal UAV trajectory based on $\boldsymbol{\Pi}^*$, $\mathcal{J}^*_\Pi$, $\boldsymbol{\pi}^*_\Pi$, $\mathcal{U}^*_\Pi$ with line segments and the maximum speed;

we can conclude that the optimal objective value of (P2) for the entire flight is equal to $\sum_{l=1}^{K+1} \Lambda^*_l$, since otherwise we can always replace the lth segment with $\{\mathbf{u}^*_l(t)\}$, which results in a smaller weighted sum. As a result, we can search all the possible permutations of the K target locations and minimize the weighted sum along each segment of UAV trajectory. With given permutation $\boldsymbol{\Pi}$, the weighted sum minimization problem for the lth segment can be regarded as (P2-s) with initial location $\mathbf{w}_{\Pi_{l-1}}$ and finial location $\mathbf{w}_{\Pi_l}$, where the solution has already been obtained. Based on the above discussions, the optimal solution to (P2) can be obtained by the following Algorithm 1, where the optimal solution to (P2-s) is obtained from Step 5 to Step 12. The complexity of Algorithm 1 is given by $O(2^M M! K! K \log(1/\zeta))$, where ζ is the solution accuracy.

To further reduce the complexity, especially when M and K are large, we propose a DGA for finding a suboptimal solution to (P2). Specifically, we define another graph $\check{G}(\check{V}, \check{E}, \check{w})$ as follows.

Definition 2: $\check{G}(\check{V}, \check{E}, \check{w})$ is constructed by

- $\check{V} \triangleq \{\check{v}_0, \check{v}_1, \ldots, \check{v}_K, \check{v}_{K+1}\}$. We introduce a vertex $\check{v}_k$ for the target location ρ_k, i.e., $\mathbf{w}_k$, $1 \le k \le K$, where $\check{v}_0$ and $\check{v}_{K+1}$ represent $\mathbf{u}_I$ and $\mathbf{u}_F$, respectively.
- $\check{E} \triangleq \{(\check{v}_i, \check{v}_j) | 0 \le i \ne j \le K+1\}$. For any two different vertices $\check{v}_i$ and $\check{v}_j$, $\check{v}_i \in \check{V}$, $\check{v}_j \in \check{V}$, $i \ne j$, there exists an edge $(\check{v}_i, \check{v}_j)$.
- $\check{w} : \check{E} \to \mathbf{R}^+$, where $\check{w}$ is a weight function. In particular, $\check{w}(\check{v}_i, \check{v}_j)$ represents the weighted sum of the operation time and total AoI starting from $\check{v}_i$ and ending at $\check{v}_j$.

Algorithm 2: Double Graph Algorithm for Problem (P2).

1: Construct graph $\check{G}(\check{V},\check{E},\check{w})$ with Definition 2;
2: **for** each edge $(\check{v}_i,\check{v}_j)\in\check{E}$ **do**
3: Let $\tilde{\mathbf{u}}_I$ and $\tilde{\mathbf{u}}_F$ be the initial and final locations corresponding to vertices $\check{v}_i$ and $\check{v}_j$;
4: Construct graph $\hat{G}(\hat{V},\hat{E},\hat{w})$ with Definition 1, where $\hat{v}_0$ corresponds to $\tilde{\mathbf{u}}_I$ and $\hat{v}_{M+1}$ corresponds to $\tilde{\mathbf{u}}_F$;
5: Obtain the shortest path in $\hat{G}(\hat{V},\hat{E},\hat{w})$ by utilizing Dijkstra algorithm to determine the visited BS set $\mathcal{J}_{i,j}$ and visiting order $\boldsymbol{\pi}_{i,j}$;
6: With given $\tilde{\mathbf{u}}_I$, $\tilde{\mathbf{u}}_F$, $\mathcal{J}_{i,j}$, and $\boldsymbol{\pi}_{i,j}$, solve convex optimization problem (P3) with CVX to obtain objective value $\Lambda_{i,j}$ and waypoints $\mathcal{U}_{i,j}$.
7: $\check{w}(\check{v}_i,\check{v}_j)\leftarrow\Lambda_{i,j}$;
8: **end for**
9: Find a TSP path $\vec{p}$ over $\check{G}(\check{V},\check{E},\check{w})$ starting from $\check{v}_0$ and ending at $\check{v}_{K+1}$ with TSP algorithm [50] to determine total weighted sum Λ and visiting order $\boldsymbol{\Pi}$;
10: $\mathcal{J}_\Pi\leftarrow\{\mathcal{J}_{i,j}|(\check{v}_i,\check{v}_j)\in\vec{p}\}$, $\boldsymbol{\pi}_\Pi\leftarrow\{\boldsymbol{\pi}_{i,j}|(\check{v}_i,\check{v}_j)\in\vec{p}\}$, $\mathcal{U}_\Pi\leftarrow\{\mathcal{U}_{i,j}|(\check{v}_i,\check{v}_j)\in\vec{p}\}$;
11: Construct the UAV trajectory based on $\boldsymbol{\Pi}$, $\mathcal{J}_\Pi$, $\boldsymbol{\pi}_\Pi$, $\mathcal{U}_\Pi$ with line segments and maximum speed;

Note that weight $\check{w}(\check{v}_i,\check{v}_j)$ can be obtained by solving problem (P2-s) with given $\check{v}_i$ and $\check{v}_j$. Due to Definition 2, solving (P2) is equivalent to finding a TSP path starting from $\check{v}_0$ and ending at $\check{v}_{K+1}$, while each vertex in $\check{V}$ is visited only once. Note that although finding a TSP path is NP-hard, it has been well studied and there exist many efficient algorithms to find an approximate solution, e.g., with complexity $O(K^2)$ [50]. Based on the above discussions, the suboptimal solution to (P2) can be obtained by the following DGA given by Algorithm 2, where the suboptimal solution to (P2-s) is obtained from Step 3 to Step 7. Note that the main steps in Algorithm 2 are Step 2 to Step 8 for weight calculation, whose complexity is given by $O(K^2M^2)$. Furthermore, the complexity of the TSP algorithm is $O(K^2)$. In summary, the complexity of Algorithm 2 is given by $O(K^2M^2)$.

Note that the above algorithms rely on the statistic channel information, thus only average communication performance over a large number of similar scenarios can be obtained, which may not be suitable for the site-specific performance. Furthermore, these algorithms are time consuming, thus not suitable for onboard implementation on the UAV, where fast decision making is needed. The aforementioned limitations motivate the DLA developed based on DRL, which will be presented in the following section.

V. Proposed DDQN-Based DLA

To analyze the site-specific performance for a specific local environment and make fast decisions, the DLA is proposed in this section. We first reformulate the problem as a Markov decision process (MDP), and then propose a DDQN-based DLA to learn from the local radio propagation environment by sampled rate measurements, where no prior knowledge about the radio propagation environment is required. Specifically, the sampled rates of the A2G channels measured by the UAV is used as input, based on which a function that maps the local environment input to the output flying decisions can be learnt with a DDQN-based model. After suitable training, the DDQN can be utilized to make quick decisions for the UAV.

A. Problem Reformulation

In practice, the rate $R_m(t)$ can be measured in a specific local environment by leveraging the existing handover mechanisms with continuous RSRP measurements [23]. Similar as in Section IV, the transmission scheduling for (P1) can be obtained by

$$x_m(t)=\begin{cases}1, & m=\arg\max_{m'\in\{1,2,\ldots,M\}}R_{m'}(t),\\ 0, & \text{otherwise},\end{cases}\tag{16}$$

where $R_{m'}(t)$ is the actual measured rate. As a result, the UAV selects the BS with the maximum measured rate for sensory data transmission at each time t. Then, $\lambda(t)=\max_\tau\{\tau|\max_m R_m(\tau)\geq R_s,\tau\in[0,t]\}$. Due to (7), the optimal UAV trajectory to (P1) can be partitioned by the target locations in $\mathcal{K}$ into $K+1$ segments, and we can solve $K+1$ independent subproblems for each segment of the UAV trajectory as in Section IV. As such, we first obtain the visiting order of the target locations $\boldsymbol{\Pi}=(\Pi_1,\ldots,\Pi_K)$ by the offline Algorithm 2 in Section IV. The subproblem for each segment l can be written as the following (P4), $1\leq l\leq K$, where we omit the subscripts of segment l for ease of exposition, i.e.,

$$\text{(P4)}:\min_{\tilde{T},\{\mathbf{u}(t)\}}\theta\tilde{T}+(1-\theta)\int_0^{\tilde{T}}(t-\lambda(t))\,dt$$

$$\text{s.t. }\|\mathbf{v}(t)\|=V_{\max},\forall t\in[0,\tilde{T}],\tag{17}$$

$$\mathbf{u}(0)=\tilde{\mathbf{u}}_I,\mathbf{u}(\tilde{T})=\tilde{\mathbf{u}}_F.\tag{18}$$

where $\tilde{T}$ is the UAV operation time along the lth segment, $\tilde{\mathbf{u}}_I$ and $\tilde{\mathbf{u}}_F$ are the start and end points of the lth segment of the UAV trajectory, respectively. Note that $\tilde{\mathbf{u}}_I$ and $\tilde{\mathbf{u}}_F$ are exactly determined by the given visiting order $\boldsymbol{\Pi}$. In the following, we will only focus on solving problem (P4).

Since no prior knowledge about the specific radio propagation environment is given, then the UAV can make sequential decisions regarding its trajectory in each time step, and the trajectory design influences its states as well as the AoI in the future. As such, we first transform problem (P4) into an MDP. For ease of exposition, the time horizon $\tilde{T}$ is discretized into $\tilde{N}$ equal time slots, i.e., $\tilde{T}=\tilde{N}\delta$, where the slot duration δ is appropriately chosen such that the UAV's location can be assumed to be approximately unchanged within each time slot. As such, the UAV trajectory can be approximated by a sequence $\{\mathbf{u}[n]|1\leq n\leq\tilde{N}\}$. The discretized form of $R_m(t)$, $\lambda(t)$, $\beta_m(t)$, $h_m(t)$, $\tilde{h}_m(t)$ are represented by $R_m[n]$, $\lambda[n]$, $\beta_m[n]$, $h_m[n]$, $\tilde{h}_m[n]$, respectively. Thus, (P4) can be approximated as

$$\text{(P5)}:\max_{\{\mathbf{u}[n]\},\tilde{N}}-\theta\tilde{N}-(1-\theta)\sum_{n=1}^{\tilde{N}}(n-\lambda[n])$$

$$\text{s.t. }\mathbf{u}[n+1]=\mathbf{u}[n]+\delta V_{\max}\vec{\mathbf{v}}[n],\forall n,\tag{19}$$

$$\|\vec{\mathbf{v}}[n]\| = 1, \forall n, \tag{20}$$

$$\mathbf{u}[1] = \tilde{\mathbf{u}}_I, \mathbf{u}[\tilde{N}] = \tilde{\mathbf{u}}_F, \tag{21}$$

where $\vec{\mathbf{v}}[n]$ denotes the UAV flying direction at time slot n.

Note that the UAV is at waypoint $\mathbf{u}[n]$ at time slot n, and the LoS/NLoS states can be exactly determined by checking whether obstacles exist between BSs and the UAV in the specific local environment, where the large-scale channel power gain $\beta_m[n]$ is thus determined. On the other hand, we can also measure $\tilde{h}_m[n]$ with the existing handover mechanisms with continuous RSRP measurements. Note that each time slot may contain multiple fading blocks, e.g., $\Omega \geq 1$ fading blocks. Then the UAV can perform Ω measurements and adopt the average value for $R_m[n]$. In other words, $R_m[n] = \frac{1}{\Omega}\sum_{j=1}^{\Omega} R_m[n,j]$, where $R_m[n,j]$ is the jth measurement at time slot n, $1 \leq j \leq \Omega$. Although such an estimation causes a performance gap, the gap is practically negligible with a sufficiently small δ [22]. As a result, rate $R_m[n]$ can be measured at location $\mathbf{u}[n]$, which is utilized in the following for reward calculation of DDQN.

Problem (P5) can be modeled as an MDP. In particular, the UAV is regarded as an agent and all of the network settings (including the local radio propagation environment, target locations and BSs) are regarded as the environment. As such, we can characterize the UAV by a tuple $<\mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R}>$, where $\mathcal{S}$ is the state space, $\mathcal{A}$ is the action space, $\mathcal{P}$ is the state transition function, and $\mathcal{R}$ is the reward function. By formulating the problem as an MDP, the UAV acts like an agent which finds the peak of the reward by interacting with the environment. In the following, we define the above elements.

- *State Space:* We define the state of the UAV at the beginning of the nth time slot as the UAV's horizontal location, i.e., $\mathbf{u}[n]$. Then $\mathbf{u}[1] = \tilde{\mathbf{u}}_I$ is the initial state while $\tilde{\mathbf{u}}_F$ is the final state. As a result, $\mathcal{S} \subseteq \mathbb{R}^{2\times 1}$ denotes the continuous state space that contains all possible UAV horizontal locations.
- *Action Space:* We define the actions of the UAV as its flying directions. In particular, the UAV's action within the nth time slot is expressed as $\vec{\mathbf{v}}[n]$, where $\|\vec{\mathbf{v}}[n]\| = 1$. Note that the UAV updates its decisions at the beginning of each time slot, and will keep its decision unchanged within each time slot. As a result, $\mathcal{A} \triangleq \{\vec{\mathbf{v}} | \|\vec{\mathbf{v}}\| = 1\}$ denotes the action space that contains all the possible UAV flying directions. For simplicity, we uniformly discretize the set of actions into κ values to ensure finite action space, i.e., $\mathcal{A} = \{\vec{\mathbf{v}}^{(1)}, \ldots, \vec{\mathbf{v}}^{(\kappa)}\}$.
- *State Transition Function:* The UAV's current state before the $(n+1)$th time slot is determined by the UAV's state before the nth time slot. With given the current action $\vec{\mathbf{v}}[n]$, we can obtain the movement of the UAV within the time slot n with (19), i.e., $\mathbf{u}[n+1] = \mathbf{u}[n] + \delta V_{\max}\vec{\mathbf{v}}[n]$.
- *Reward Function:* For (P5), the reward function is defined to award the UAV for reaching its destination, and penalize the UAV for moving and the increase of AoI. At time slot n, we define an indicator $I_n = 1$ when $\max_m R_m[n] \geq R_s$, and $I_n = 0$ otherwise. Let $\lambda[1] = 1$. For $n \geq 2$, we have

$$\lambda[n] = \begin{cases} n, & I_n = 1, \\ \lambda[n-1], & I_n = 0. \end{cases} \tag{22}$$

Due to the objective function of (P5), the reward at time slot n is defined as

$$\Phi[n] = -\theta - (1-\theta)(n - \lambda[n]). \tag{23}$$

Therefore, the UAV is motivated to minimize its weighted sum of operation time and total AoI by making decisions on its flying directions. Note that if there exist two flying paths with the same total AoI, we prefer the one with minimum operation time. In this case, we take the total AoI as performance metric which tilts the problem towards solutions that reduce the mission duration.

After obtaining the discretized actions, the state space can be further discretized into finite space since position increments and directions are both discrete. In the following, a DLA is proposed based on DDQN with a dueling network architecture. Our proposed DLA consists of an offline training process and an online execution process. During the training process, the training data is collected and the DDQN model is trained offline. After the training process, we execute the well-trained DDQN model to learn the optimized flying strategy for the UAV according to its current state.

B. Proposed Solution

In the following, we adopt a DQN with the dueling network architecture [51], termed as dueling DQN, to approximate the state-action value $\hat{Q}(s,a)$ with weights $\boldsymbol{\eta}$. Compared to the standard DQN, the dueling DQN consists of two streams that represent the value and advantage functions, which are combined in a smart way via a special aggregating layer to estimate the state-action value function Q. In particular, $\hat{Q}(s,a;\alpha,\beta) = \mathcal{V}(s;\beta) + (\mathcal{G}(s,a;\alpha) - \frac{1}{|\mathcal{A}|}\sum_a \mathcal{G}(s,a;\alpha))$, where α and β are the parameters of fully-connected layers. The value function $\mathcal{V}$ corresponds to how good it is in a particular state s, while the advantage function $\mathcal{G}$ decouples the state value from the Q-function such that the importance of each action can be measured. In this case, more robust estimates of state value can be achieved, and thus stability and convergence rate can be significantly improved [51].

We adopt the multi-step bootstrapping which can effectively accelerate the training [52]. In particular, the truncated N_0-step return is given by

$$\Phi[n : n+N_0] = \sum_{k=0}^{N_0-1} \gamma \Phi[n+k+1], \tag{24}$$

with γ denoting the discount factor. In (P5), the objective function corresponds to $\gamma = 1$, which means that all rewards are equally important. Furthermore, it is shown that Q-values are always overestimated by DQN training. We adopt DDQN to tackle such issue [23], [52]. DDQN includes two neural networks named the primary neural network and the target neural network, both of which adopt the dueling structure. The key idea of DDQN is to select an action by using the primary network with weights $\boldsymbol{\eta}$, while uses the target network with $\boldsymbol{\eta}'$ to compute the target Q-value for the action. By decoupling Q value evaluation and action selection, DDQN can efficiently mitigate overestimation and enhance stability during model training.

For the training process of DDQN with the dueling network architecture, the UAV obtains a state from the state space $\mathcal{S}$ and selects an action from the action space $\mathcal{A}$ with the ϵ-greedy policy to balance exploitation and exploration. To be specific, the action $\vec{\mathbf{v}}[n]$ that maximizes the Q-value is chosen with a probability of $1-\epsilon$, while a random action is chosen with probability ϵ, i.e.,

$$\mathbb{P}(\vec{\mathbf{v}}[n]) = \begin{cases} 1-\epsilon, & \vec{\mathbf{v}}[n] = \max\limits_{\vec{\mathbf{v}}' \in \mathcal{A}} \hat{Q}(\mathbf{u}[n], \vec{\mathbf{v}}'; \boldsymbol{\eta}), \\ \epsilon/(|\mathcal{A}|-1), & \text{otherwise.} \end{cases} \tag{25}$$

A replay buffer $\mathcal{D}$, i.e., memory pool, is used to store the N_0-step transitions (or, experience) $(\mathbf{u}[n], \vec{\mathbf{v}}[n], \Phi[n:n+N_0], \mathbf{u}[n])$, and a mini-batch of experiences are randomly sampled from it to update the weights $\boldsymbol{\eta}$ by minimizing the loss function

$$\Big(\Phi[n:n+N_0] + \gamma^{N_0} \hat{Q}(\mathbf{u}[n+N_0], \vec{\mathbf{v}}^*; \boldsymbol{\eta}') - \hat{Q}(\mathbf{u}[n], \vec{\mathbf{v}}[n]; \boldsymbol{\eta})\Big)^2, \tag{26}$$

where $\vec{\mathbf{v}}^* = \arg\max_{\vec{\mathbf{v}}' \in \mathcal{A}} \hat{Q}(\mathbf{u}[n+N_0], \vec{\mathbf{v}}'; \boldsymbol{\eta})$. Here, the first part $\Phi[n:n+N_0] + \gamma^{N_0} \hat{Q}(\mathbf{u}[n+N_0], \vec{\mathbf{v}}^*; \boldsymbol{\eta}')$ represents a target that the Q-value needs to move and the second part $\hat{Q}(\mathbf{u}[n], \vec{\mathbf{v}}[n]; \boldsymbol{\eta})$ denotes the estimation of Q-value. As a result, the loss function indicates the estimation error of a DDQN model, where better estimation performance of DDQN can be achieved with smaller loss function. The pseudocode of DLA for problem (P5) is summarized in Algorithm 3.

In Algorithm 3, the training process consists of $\Psi^{\max}$ episodes, where ϵ-greedy approach is adopted to choose an action with the given current state in each episode. The algorithm starts with a fairly randomized policy at Step 7 and slowly move to a deterministic policy later due to the update of ϵ at Step 13. According to [51], it has been proved that the convergence of the learning Algorithm 3 with DDQN can be guaranteed. As for the DLA, the training procedure can be deployed in a simulator and runs offline at BSs, where its time complexity is proportional to the number of training time, i.e., $O(\Psi^{\max} N^{\max})$. Once the neural networks are trained, the converged neural networks are saved for testing and can be easily deployed at the UAV, where the optimized policy can be generated very fast with only some simple algebraic calculations. In particular, in each time slot, the UAV's action will be generated by the trained neural networks, i.e., through the operation at Step 7, which guides the UAV by making real-time decisions and will inevitably reduce the payoff of practical implementation.

Algorithm 3: Training Process for Problem (P5).

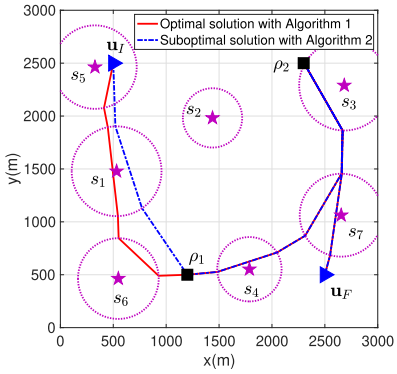
1: Initialize replay memory $\mathcal{D}$ with capacity $|\mathcal{D}|$; Initialize exploration probability ϵ and decaying rate ν;
2: Initialize the primary Q-network with weights $\boldsymbol{\eta}$;
3: Initialize the target Q-network with weights $\boldsymbol{\eta}' = \boldsymbol{\eta}$;
4: **for** episode $= 1$ to $\Psi^{\max}$ **do**
5: Initialize the initial state $\mathbf{u}[1] = \tilde{\mathbf{u}}_I$. Initialize $n = 1$;
6: **while** $n \le N^{\max}$ and $\mathbf{u}[n] \neq \tilde{\mathbf{u}}_F$ **do**
7: Choose a random action $\vec{\mathbf{v}}[n]$ in $\mathcal{A}$ with probability ϵ; otherwise select $\vec{\mathbf{v}}[n] = \max_{\vec{\mathbf{v}}' \in \mathcal{A}} \hat{Q}(\mathbf{u}[n], \vec{\mathbf{v}}'; \boldsymbol{\eta})$;
8: Perform action $\vec{\mathbf{v}}[n]$ and observe the next state $\mathbf{u}[n+1]$;
9: Obtain $\lambda[n]$ due to (22) based on measurement of $R_m[n]$, and calculate reward $\Phi[n] = -\theta - (1-\theta)(n - \lambda[n])$;
10: Calculate the N_0-Step reward $\Phi[n-N_0:n]$ according to (24), and store $(\mathbf{u}[n-N_0], \vec{\mathbf{v}}[n-N_0], \Phi[n-N_0:n], \mathbf{u}[n])$ into the replay buffer $\mathcal{D}$;
11: Sample a random minibatch of transitions $(\mathbf{u}[j], \vec{\mathbf{v}}[j], \Phi[j:j+N_0], \mathbf{u}[j+N_0])$ from replay buffer $\mathcal{D}$;
12: Update the weights $\boldsymbol{\eta}$ of the primary Q-network by gradient descent with the loss function defined by (26);
13: Update $n \leftarrow n+1$; $\epsilon \leftarrow \epsilon\nu$;
14: Every $\hat{B}$ steps, update the target network weights $\boldsymbol{\eta}' = \boldsymbol{\eta}$;
15: **end while**
16: **end for**

VI. Simulation Study

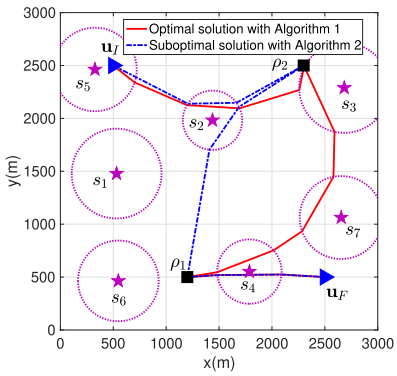
In this section, we provide the system settings and evaluate the performance of both the DGA for average performance and DLA for site-specific performance through simulation. We also study the impact of different parameters on the performance of the proposed algorithms.

A. Simulation Setting

As shown in Fig. 2, we consider a cellular-connected UAV sensing system with $M = 7$ ground BSs, which are uniformly distributed in a square urban area (denoted by $\hat{\mathcal{S}} \in \mathbb{R}^2$) with width 3.0 km. The UAV collects sensory data from $\hat{\mathcal{S}}$ with its sensors in each time instant, while the sensing data are transmitted to the ground BSs for further processing. The target locations in $\mathcal{K}$ are denoted by black squares in Fig. 2, which are required to be visited during the flight. The mission of the UAV is to sense and transmit data in region $\hat{\mathcal{S}}$ from $\mathbf{u}_I$ to $\mathbf{u}_F$, while accessing all the target locations in $\mathcal{K}$. The heights and locations of the buildings in the considered urban area are generated based on one realization of the statistical model specified by ITU [39]. As such, the LoS/NLoS states at any location can be exactly determined by checking whether obstacles exist between ground BSs and the UAV, while the statistical model can be used to obtain the LoS probability, which reflects the average communication performance over a large number of realizations. In the following, the results of average communication performance with the statistical model using DGA are given in Section VI-B, while the results of site-specific performance using DLA are given in Section VI-C. Similar to [53], the antenna patterns related parameters are set as: $\phi_D = 10^\circ$, $G_0 = 30$ dB, $HPBW_v = 65^\circ$, $n_0 = 8$. The A2G channels parameters are set

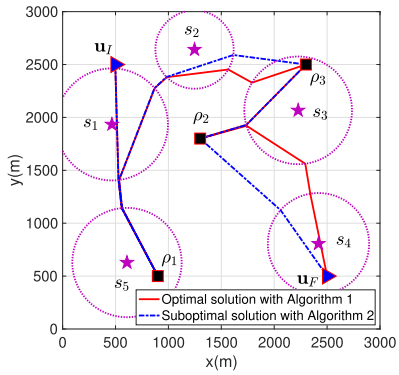


(a) $\theta = 0.2$

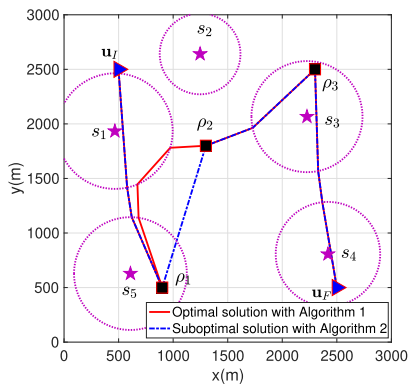


(b) $\theta = 0.8$

Fig. 2. Optimized trajectories with different values of weighting factor θ. The blue triangles denote the initial and final UAV locations.



(a) $\theta = 0.2$



(b) $\theta = 0.8$

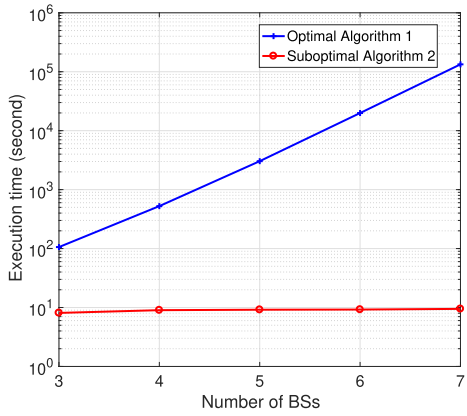
Fig. 3. Optimized trajectories with different values of weighting factor θ ($M = 5$, $K = 3$, $R_s = 1$ Mbps).

as: $\beta_0 = -60$ dB, $\sigma^2 = -110$ dBm, $a = 10$, $b = 0.6$, $\alpha = 2.2$, $\mu = 0.01$, and $K_c = 15$ dB as in [21]. Unless otherwise stated, other parameters are set as: $H_G = 25$ m, $H = 100$ m, $B = 1$ MHz, $P = 0.1$ W, $R_s = 1.5$ Mbps, $\delta = 1$ s, and $V_{\max} = 50$ m/s.

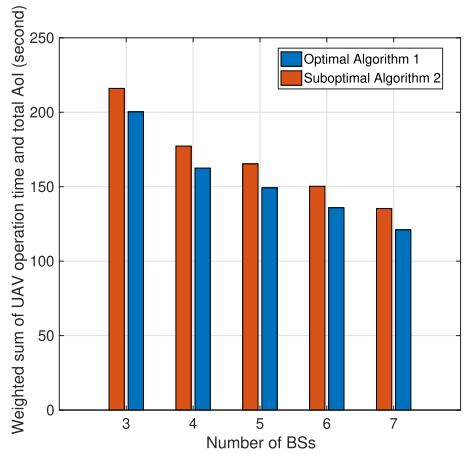
B. Average Performance of the Proposed DGA

We first consider the offline design with the average communication performance. To balance the tradeoff between the operation time and total AoI of the UAV, the optimized trajectories with different values of weighting factor θ are shown in Fig. 2. The colored stars denote the ground BSs and the circle around BS s_m represents the service area $\mathcal{A}_m$ (obtained in Section IV-A), where timely transmission requirement can be always satisfied if the UAV is within $\mathcal{A}_m, \forall m$. If the UAV is out of the service areas, the sensory data cannot be transmitted timely and thus the AoI of the UAV increases. It is observed that the ground BSs have service areas with heterogeneous size even when the UAV generates sensory data with a fixed sensing rate R_s and transmits data with the same transmit power P, since the local radio environment around different BSs are heterogeneous (e.g., with different building distribution). From Fig. 2, we can see that the UAV enters and leaves the service areas of BSs sequentially, where all the target locations are reached. With different values of weighting factor θ, the visiting order of the target locations and BSs as well as waypoints may be different, which can be obtained by the optimal Algorithm 1 and the suboptimal Algorithm 2. It is observed that the suboptimal solution obtained by Algorithm 2 achieves similar trajectory as the optimal solution obtained by Algorithm 1, thus validating the efficiency of the DGA. The performance with different scenario configuration are provided in Fig. 3 (e.g., with $M = 5, K = 3, R_s = 1$ Mbps). It is observed from Fig. 3 that similar trends can be obtained as in Fig. 2, where the details are omitted for brevity. Thus, unless otherwise stated, we only consider the scenario configuration as in Fig. 2 to avoid redundancy.

Fig. 4 depicts the average performance comparison between Algorithms 1 and 2 under different numbers of BSs when $\theta = 0.8$, where the BSs are randomly distributed in the area and the results are obtained by averaging over 100 random realizations of the BS locations. Similar results can be obtained for $\theta = 0.2$, which are omitted for brevity. The required execution times for the two algorithms are given in Fig. 4(a), which are obtained over a computer with dual core CPU 3.4 GHz. It is observed that the execution time of the search Algorithm 1 increases exponentially with M due to the rapidly increased search space, which is also verified by the complexity analysis in Section IV. In contrast, our proposed suboptimal Algorithm 2 requires much less execution time than Algorithm 1 and achieves a close performance as Algorithm 1 (see Fig. 4(b)). Thus, the proposed

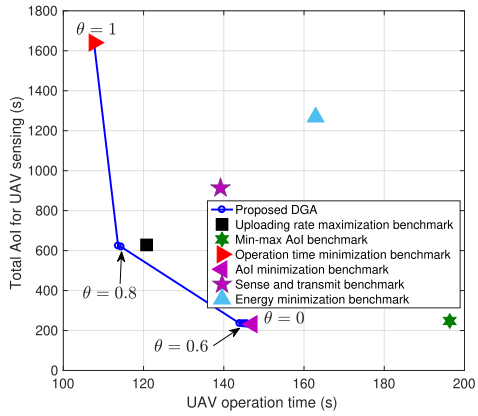


(a) Execution time comparison

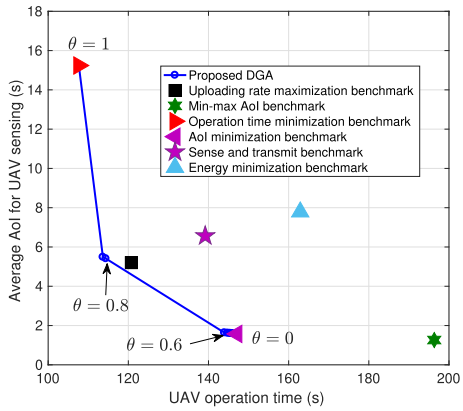


(b) Weighted sum comparison

Fig. 4. Performance comparisons between optimal Algorithm 1 and suboptimal Algorithm 2.



(a) Total AoI versus UAV operation time



(b) Average AoI versus UAV operation time

Fig. 5. Tradeoff between UAV operation time and AoI for UAV sensing.

DGA is a practically appealing solution for the offline problem (P2) from both complexity and performance considerations. In the following, we only use DGA for performance comparison with other benchmark schemes.

It can be seen in Fig. 2 that when the weighting factor θ is large, the UAV is expected to fly directly to the target locations to reduce the flying distance since minimizing UAV operation time is more important, while the total AoI of the UAV increases due to the long time out of service areas. This is also verified in the tradeoff curves plotted in Fig. 5, which show the total AoI and average AoI (i.e., $\frac{1}{T}\int_0^T A(t)dt$) versus UAV operation time obtained for the proposed DGA by using different values of θ. For any point in Fig. 5(a), its values for x-coordinate and y-coordinate represent for the contributions for operation time and total AoI, respectively. As expected, the UAV operation time decreases with θ while the total AoI increases with θ, which shows that the decrease of total AoI is at the cost of an increase of the UAV operation time. Similar trend is observed for average AoI in Fig. 5(b), where the details are omitted for brevity. Thus, θ can be flexibly set to achieve a good balance between the AoI and operation time of the UAV according to practical system requirements.

To further illustrate the performance gain achieved by the proposed DGA, we consider the following six benchmark schemes, which are referred to as the *Operation time minimization benchmark*, the *AoI minimization benchmark*, the *Min-max AoI benchmark*, the *Uploading rate maximization benchmark*, the *Sense and transmit benchmark*, as well as the *Energy minimization benchmark*. In the operation time minimization benchmark, the UAV operation time is minimized by visiting all target locations with the TSP method as in [47], which corresponds to the case when $\theta = 1$. In the AoI minimization benchmark, the total AoI of UAV sensing is minimized as in [54], which corresponds to the case when $\theta = 0$. In the min-max AoI benchmark, the maximum AoI along the UAV's flight is minimized with graph algorithm as in [55]. In the uploading rate maximization benchmark, the UAV flies to the top of the visited BSs for uploading rate maximization of sensory data as in [22]. In the sense and transmit benchmark, the UAV flies directly to the target points and then transmits sensory data to the nearby BSs, similar as in [11]. For the energy minimization benchmark, the UAV flies to the target points with the minimum-energy speed, which is found numerically by one-dimensional search based on the propulsion energy model derived in [41].

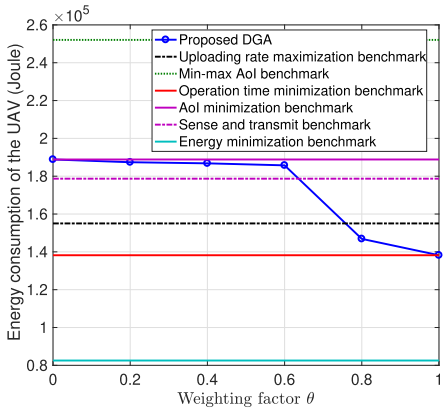


Fig. 6. Energy consumption of the UAV for different schemes.

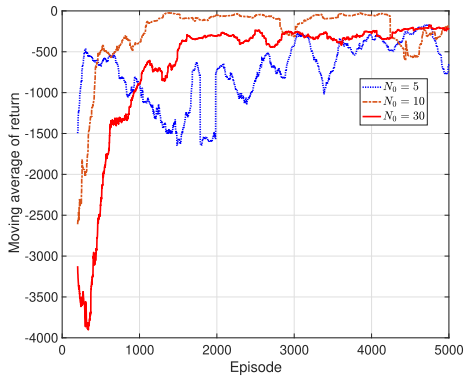


Fig. 7. Average reward trends versus training episodes for DLA.

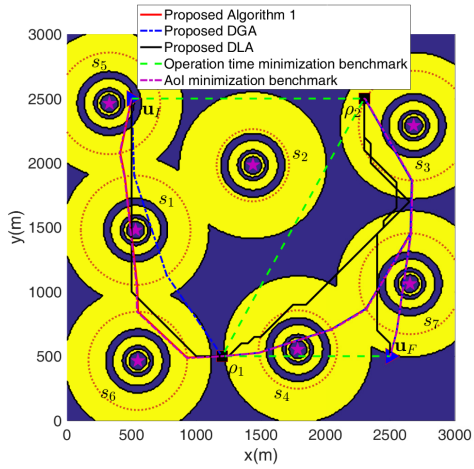
From Fig. 5 we can see that each of the six benchmark schemes only result in one singleton trade-off point. Unlike DGA, no flexible trade-off between the operation time and AoI of the UAV can be achieved with these benchmark schemes, since they only optimize either the UAV operation time or the AoI. Uploading rate maximization benchmark achieves a comparable AoI as DGA, but it leads to a higher UAV operation time and such gain is brought by the waypoint optimization of DGA. The performance gap between the min-max AoI benchmark and DGA illustrates the additional gain of joint optimization of the AoI and the visiting order of target locations. The min-max AoI benchmark corresponds to egalitarian bargaining for fairness efficiency, while the AoI minimization benchmark corresponds to utilitarian bargaining. It is also observed that the sense and transmit benchmark results in higher AoI and UAV operation time than DGA, which is mainly attributed to the performance gain brought by the optimized BS selection in DGA. Although energy minimization benchmark results in the minimum energy consumption of the UAV, higher AoI and UAV operation time are incurred. This can be verified in Fig. 6, where a comparison of energy consumption among different schemes is presented, utilizing the energy model in [41].

C. Site-Specific Performance of the Proposed DLA

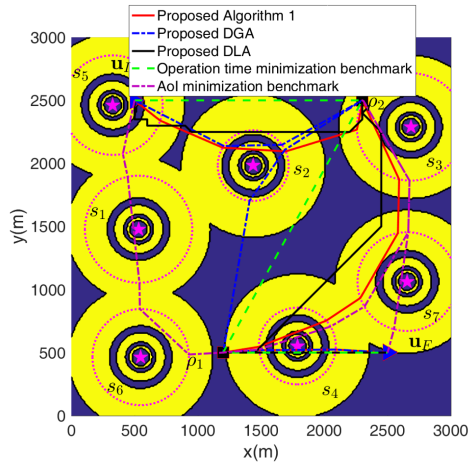
Next, the site-specific performance is studied for a specific urban local environment mentioned in Section VI-A, and we evaluate the performance of the proposed DLA. For the DLA, we set $\Psi^{\max} = 5000$, $N^{\max} = 800$, $\Omega = 100$, $\kappa = 8$, $\epsilon = 0.5$, $\nu = 0.998$, and $\hat{B} = 5$. There are 5 hidden layers in the DDQN, where the numbers of neurons of the first 4 hidden layers are 512, 256, 128, and 128, respectively. There are $\kappa + 1$ neurons in the last dueling layer, among which κ neurons corresponds to the action advantages of the κ actions and the extra neuron is for the estimation of the state-value.

Fig. 7 depicts the convergence trends of the training process of the proposed DLA. In particular, the moving average return per episode for DLA with different values of N_0 is given in Fig. 7, where the length of moving window is 200 episodes. It can be seen that the average return gradually increases with the training episode and finally tends to be stable with a sufficiently large amount of episodes, which demonstrates the convergence of the proposed DLA. This is because the UAV can learn from its experience to improve the long-term reward. In addition, the influence of the step size N_0 of multi-step bootstrapping on the convergence speed is also shown in Fig. 7. Intuitively, small N_0 leads to poor predictability. When N_0 is very large, the algorithm converges slower since the amount of calculation is particularly large. Thus, N_0 should be appropriately chosen to improve performance throughout the learning process. It is observed in Fig. 7 that the curve value fluctuates greatly with smaller N_0 due to the poor predictability. According to the results, we choose $N_0 = 30$ in the following to ensure faster learning. Even for the curve with $N_0 = 30$, it can be seen that the moving average return decreases first and finally tends to be stable as the value of episode increases. The reason is that when the UAV starts to interact with the environment, due to the lack of experience, it will fly off the area border prematurely such that more punishment can be avoided, and thus the episode terminates prematurely. As more experience has been accumulated, the UAV tends to select a better strategy avoiding the blue areas to reach the destination.

Fig. 8(a) shows the actual global service map for the considered area, which denotes whether the immediate data transmission constraint in (3) is satisfied at any horizontal location $\mathbf{u}[n] \in \hat{\mathcal{S}}$, i.e., $\sum_{m=1}^{M} x_m(t) R_m(t) \geq R_s, \forall t$, at $\mathbf{u}[n]$. In other words, if the UAV lies in the yellow areas in Fig. 8, then immediate data transmission constraint can be achieved; otherwise the AoI will increase with time. It is worthwhile to note that such a global service map is numerically obtained based on the building locations and heights as well as channel realizations mentioned in Section VI-A via computer simulations, and it is not available prior to our proposed algorithms are executed. The circles around BSs represent the service areas obtained by the statistical model in Section IV. It can be seen that the actual global service map are similar as that obtained from the statistical model, which verifies the approximation accuracy of average communication performance. On the other hand, the difference between actual global service map and that obtained from statistical model (e.g., not the same regular disc areas as red circle parts and many coverage holes exist) demonstrates



(a) $\theta = 0.2$



(b) $\theta = 0.8$

Fig. 8. Optimized trajectories for site-specific performance.

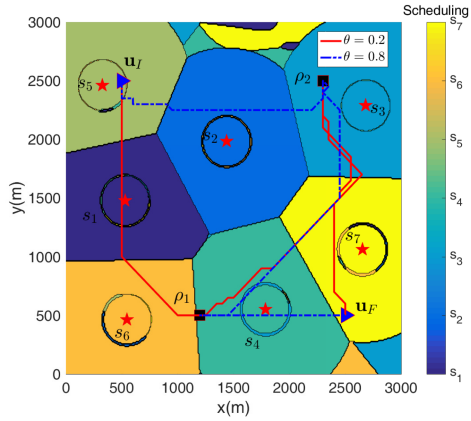


Fig. 9. Transmission scheduling for site-specific performance with DLA.

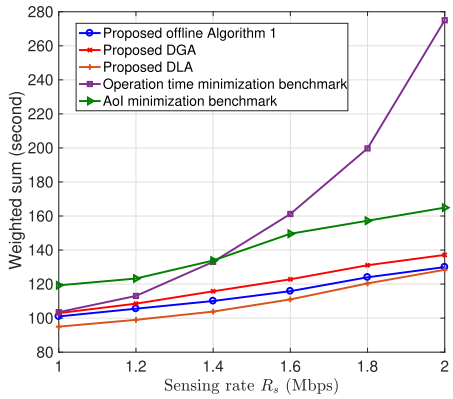


Fig. 10. Performance comparisons with different schemes for site-specific performance.

the characteristic of site-specific performance, which depends on the actual radio propagation environment with buildings and downtilted BS antenna radiation pattern.

It is observed in Fig. 8 that the optimized trajectory obtained from DLA follows the similar trend as the offline optimal Algorithm 1 and DGA. In particular, the UAV tries to visit all the target locations by avoiding service holes due to the values of θ, such that the weighted sum of operation time and total AoI is minimized. The solution obtained by Algorithm 1 and DGA may lead to a longer distance or a larger AoI, which results in a larger weighted sum, since it is only based on the statistical model without no specific local environment information. Fig. 9 depicts the transmission scheduling for DLA, and we can see that the UAV may not always transmit to its closest BS due to the specific positioning of buildings as well as practical BS antenna radiation pattern.

Fig. 10 depicts site-specific performance with different schemes. In particular, Fig. 10 shows the weighted sum of operation time and total AoI of the UAV versus the sensing rate R_s when $\theta = 0.8$. Similar results can be obtained for other values of θ, which is omitted for brevity. When R_s increases, the service areas shrink and thus it is more difficult to satisfy the immediate transmission constraint, leading to the increase of AoI. It is observed that DLA performs better than not only the benchmark schemes but also the offline Algorithm 1 and DGA. The reason is that the benchmark schemes and offline algorithms are all based on the graph model constructed by the statistical model related disk coverage area with no specific local environment information. With DLA, the UAV can learn to effectively avoid service holes from accumulated experience by interacting with the specific local environment, and minimize the weighted sum of operation time and total AoI, which demonstrates the effectiveness of DLA. In addition, once the DDQN are properly trained, it only needs a small amount of algebra calculations to obtain the solution.

VII. Conclusion

In this paper, we proposed an AoI-driven UAV sensing framework with a cellular-connected UAV to provide remote sensing services over multi-cell cellular networks. In particular, we developed a remote UAV sensing model over multi-cell cellular networks in urban environments. Taking into account the tradeoff between the AoI and operation time of the UAV,

we formulated a weighted sum minimization problem through jointly optimizing the UAV trajectory and operation time as well as transmission scheduling and BS association. A search algorithm and the low-complexity DGA are developed to obtain optimal and suboptimal solutions for average performance, respectively. A DLA is then proposed to analyze the site-specific performance by employing a DDQN with dueling network model. Simulation results validate the proposed schemes in supporting remote UAV sensing and demonstrate the flexible tradeoff between the AoI and operation time of the UAV. The design framework can be extended by taking into the energy issue such that both accurate propulsion energy model and communication related energy model are included, which will be left as future work. For scenario with multi-UAV sharing the same mission, a comprehensive cooperative sensing design deserves further study, where game theory may be adopted to tackle the coordination among the UAVs. Furthermore, it would be possible to construct a radio map of considered area with given the locations of ground BS and the buildings [56], which may provide environment awareness for the corresponding problem. In the future work, the accurate radio map construction and radio map based optimization needs to be further considered in practice.

REFERENCES

[1] Y. Zeng, Q. Wu, and R. Zhang, "Accessing from the sky: A tutorial on UAV communications for 5G and beyond," *Proc. IEEE*, vol. 107, no. 12, pp. 2327–2375, Dec. 2019.

[2] Q. Wu et al., "A comprehensive overview on 5G-and-beyond networks with UAVs: From communications to sensing and intelligence," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 10, pp. 2912–2945, Oct. 2021.

[3] K. Rezaee, S. J. Mousavirad, M. R. Khosravi, M. K. Moghimi, and M. Heidari, "An autonomous UAV-assisted distance-aware crowd sensing platform using deep shuffleNet transfer learning," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 7, pp. 9404–9413, Jul. 2022.

[4] M. Mozaffari, W. Saad, M. Bennis, Y.-H. Nam, and M. Debbah, "A tutorial on UAVs for wireless networks: Applications, challenges, and open problems," *IEEE Commun. Surv. Tut.*, vol. 21, no. 3, pp. 2334–2360, Third Quarter 2019.

[5] Y. Zeng, J. Lyu, and R. Zhang, "Cellular-connected UAV: Potential, challenges, and promising technologies," *IEEE Wireless Commun.*, vol. 26, no. 1, pp. 120–127, Feb. 2019.

[6] W. Mei and R. Zhang, "Aerial-ground interference mitigation for cellular-connected UAV," *IEEE Wireless Commun.*, vol. 28, no. 1, pp. 167–173, Feb. 2021.

[7] W. K. New, C. Y. Leow, K. Navaie, Y. Sun, and Z. Ding, "Interference-aware NOMA for cellular-connected UAVs: Stochastic geometry analysis," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 10, pp. 3067–3080, Oct. 2021.

[8] S. Zhang, H. Zhang, B. Di, and L. Song, "Cellular UAV-to-X communications: Design and optimization for multi-UAV networks," *IEEE Trans. Wireless Commun.*, vol. 18, no. 2, pp. 1346–1359, Feb. 2019.

[9] F. Wu, H. Zhang, J. Wu, and L. Song, "Cellular UAV-to-device communications: Trajectory design and mode selection by multi-agent deep reinforcement learning," *IEEE Trans. Commun.*, vol. 68, no. 7, pp. 4175–4189, Jul. 2020.

[10] J. Hu, H. Zhang, L. Song, R. Schober, and H. V. Poor, "Cooperative internet of UAVs: Distributed trajectory design by multi-agent deep reinforcement learning," *IEEE Trans. Commun.*, vol. 68, no. 11, pp. 6807–6821, Nov. 2020.

[11] S. Zhang, H. Zhang, Z. Han, H. V. Poor, and L. Song, "Age of information in a cellular internet of UAVs: Sensing and communication trade-off design," *IEEE Trans. Commun.*, vol. 19, no. 10, pp. 6578–6592, Oct. 2020.

[12] F. Wu, H. Zhang, J. Wu, Z. Han, H. V. Poor, and L. Song, "UAV-to-device underlay communications: Age of information minimization by multi-agent deep reinforcement learning," *IEEE Trans. Commun.*, vol. 69, no. 7, pp. 4461–4475, Jul. 2021.

[13] X. Zhang, J. Wang, and H. V. Poor, "AoI-driven statistical delay and error-rate bounded QoS provisioning for mURLLC over UAV-multimedia 6G mobile networks using FBC," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 11, pp. 3425–3443, Nov. 2021.

[14] Z. Dai, C. H. Liu, R. Han, G. Wang, K. Leung, and J. Tang, "Delay-sensitive energy-efficient UAV crowdsensing by deep reinforcement learning," *IEEE Trans. Mobile Comput.*, vol. 22, no. 4, pp. 2038–2052, Apr. 2023.

[15] N. Van Cuong, Y.-W. Peter Hong, and J.-P. Sheu, "UAV trajectory optimization for joint relay communication and image surveillance," *IEEE Trans. Wireless Commun.*, vol. 21, no. 12, pp. 10177–10192, Dec. 2022.

[16] X. Xu, Y. Zeng, Y. L. Guan, and R. Zhang, "Overcoming endurance issue: UAV-enabled communications with proactive caching," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 6, pp. 1231–1244, Jun. 2018.

[17] H. Wang, J. Wang, G. Ding, J. Chen, F. Gao, and Z. Han, "Completion time minimization with path planning for fixed-wing UAV communications," *IEEE Trans. Wireless Commun.*, vol. 18, no. 7, pp. 3485–3499, Jul. 2019.

[18] J. Gong, T.-H. Chang, C. Shen, and X. Chen, "Flight time minimization of UAV for data collection over wireless sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 9, pp. 1942–1954, Sep. 2018.

[19] Q. Guo et al., "Minimizing the longest tour time among a fleet of UAVs for disaster area surveillance," *IEEE Trans. Mobile Comput.*, vol. 21, no. 7, pp. 2451–2465, Jul. 2022.

[20] R. Amer, W. Saad, and N. Marchetti, "Mobility in the sky: Performance and mobility analysis for cellular-connected UAVs," *IEEE Trans. Commun.*, vol. 68, no. 5, pp. 3229–3246, May 2020.

[21] C. You and R. Zhang, "Hybrid offline-online design for UAV-enabled data harvesting in probabilistic LoS channels," *IEEE Trans. Wireless Commun.*, vol. 19, no. 6, pp. 3753–3768, Jun. 2020.

[22] C. Zhan and Y. Zeng, "Energy-efficient data uploading for cellular-connected UAV systems," *IEEE Trans. Wireless Commun.*, vol. 19, no. 11, pp. 7279–7292, Nov. 2020.

[23] Y. Zeng, X. Xu, S. Jin, and R. Zhang, "Simultaneous navigation and radio mapping for cellular-connected UAV with deep reinforcement learning," *IEEE Trans. Wireless Commun.*, vol. 20, no. 7, pp. 4205–4220, Jul. 2021.

[24] M. Samir, C. Assi, S. Sharafeddine, and A. Ghrayeb, "Online altitude control and scheduling policy for minimizing AoI in UAV-assisted IoT wireless networks," *IEEE Trans. Mobile Comput.*, vol. 21, no. 7, pp. 2493–2505, Jul. 2022.

[25] J. Liu, P. Tong, X. Wang, B. Bai, and H. Dai, "UAV-aided data collection for information freshness in wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 20, no. 4, pp. 2368–2382, Apr. 2021.

[26] C. Luo, M. N. Satpute, D. Li, Y. Wang, W. Chen, and W. Wu, "Fine-grained trajectory optimization of multiple UAVs for efficient data gathering from WSNs," *IEEE/ACM Trans. Netw.*, vol. 29, no. 1, pp. 162–175, Feb. 2021.

[27] T. Wu et al., "A novel AI-based framework for AoI-optimal trajectory planning in UAV-assisted wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 21, no. 4, pp. 2462–2475, Apr. 2022.

[28] W. Wang, N. Zhao, L. Chen, X. Liu, Y. Chen, and D. Niyato, "UAV-assisted time-efficient data collection via uplink NOMA," *IEEE Trans. Commun.*, vol. 69, no. 11, pp. 7851–7863, Nov. 2021.

[29] W. Xu et al., "Minimizing the deployment cost of UAVs for delay-sensitive data collection in IoT networks," *IEEE/ACM Trans. Netw.*, vol. 30, no. 2, pp. 812–825, Apr. 2022.

[30] D. Saha, D. Pattanayak, and P. S. Mandal, "Surveillance of uneven surface with self-organizing unmanned aerial vehicles," *IEEE Trans. Mobile Comput.*, vol. 21, no. 4, pp. 1449–1462, Apr. 2022.

[31] S. Hosseinalipour, A. Rahmati, D. Y. Eun, and H. Dai, "Energy-aware stochastic UAV-assisted surveillance," *IEEE Trans. Wireless Commun.*, vol. 20, no. 5, pp. 2820–2837, May 2021.

[32] W. Wang et al., "Deployment of unmanned aerial vehicles for anisotropic monitoring tasks," *IEEE Trans. Mobile Comput.*, vol. 21, no. 2, pp. 495–513, Feb. 2022.

[33] H. Gao, J. Feng, Y. Xiao, B. Zhang, and W. Wang, "A UAV-assisted multi-task allocation method for mobile crowd sensing," *IEEE Trans. Mobile Comput.*, early access, Feb. 01, 2022, doi: 10.1109/TMC.2022.3147871.

[34] B. Chang, W. Tang, X. Yan, X. Tong, and Z. Chen, "Integrated scheduling of sensing, communication, and control for mmWave/THz communications in cellular connected UAV networks," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 7, pp. 2103–2113, Jul. 2022.

[35] J. Liu, M. Sheng, R. Lyu, Y. Shi, and J. Li, "Access points in the air: Modeling and optimization of fixed-wing UAV network," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 12, pp. 2824–2835, Dec. 2020.

[36] Z. Zhou et al., "When mobile crowd sensing meets UAV: Energy-efficient task assignment and route planning," *IEEE Trans. Commun.*, vol. 66, no. 11, pp. 5526–5538, Nov. 2018.

[37] M. Chen, W. Saad, and C. Yin, "Echo-liquid state deep learning for 360° content transmission and caching in wireless VR networks with cellular-connected UAVs," *IEEE Trans. Commun.*, vol. 67, no. 9, pp. 6386–6400, Sep. 2019.
[38] S. Zhang, Y. Zeng, and R. Zhang, "Cellular-enabled UAV communication: A connectivity-constrained trajectory optimization perspective," *IEEE Trans. Commun.*, vol. 67, no. 3, pp. 2580–2604, Mar. 2019.
[39] Propagation Data and Prediction Methods Required for the Design of Terrestrial Broadband Radio Access Systems Operating in a Frequency Range From 3 to 60GHz, document ITU-R Rec. 5-1410, Radiowave Propagation, Feb. 2012.
[40] J. -H. Lee, K. -H. Park, Y. -C. Ko, and M. -S. Alouini, "Throughput maximization of mixed FSO/RF UAV-aided mobile relaying with a buffer," *IEEE Trans. Wireless Commun.*, vol. 20, no. 1, pp. 683–694, Jan. 2021.
[41] Y. Zeng, J. Xu, and R. Zhang, "Energy minimization for wireless communication with rotary-wing UAV," *IEEE Trans. Wireless Commun.*, vol. 18, no. 4, pp. 2329–2345, Apr. 2019.
[42] M. M. Azari, F. Rosas, and S. Pollin, "Cellular connectivity for UAVs: Network modeling, performance analysis, and design guidelines," *IEEE Trans. Wireless Commun.*, vol. 18, no. 7, pp. 3366–3381, Jul. 2019.
[43] S. Zhang, H. Zhang, B. Di, and L. Song, "Joint trajectory and power optimization for UAV sensing over cellular networks," *IEEE Commun. Lett.*, vol. 22, no. 11, pp. 2382–2385, Nov. 2018.
[44] S. Kaul, R. Yates, and M. Gruteser, "Real-time status: How often should one update?," in *Proc. IEEE Conf. Comput. Commun.*, Orlando, FL, 2012, pp. 2731–2735.
[45] M. Hua, L. Yang, Q. Wu, and A. L. Swindlehurst, "3D UAV trajectory and communication design for simultaneous uplink and downlink transmission," *IEEE Trans. Commun.*, vol. 68, no. 9, pp. 5908–5923, Sep. 2020.
[46] P. Vansteenwegen, W. Souffriau, and D. V. Oudheusden, "The orienteering problem: A survey," *Eur. J. Oper. Res.*, vol. 209, no. 1, pp. 1–10, 2011.
[47] Y. Zeng, X. Xu, and R. Zhang, "Trajectory design for completion time minimization in UAV-enabled multicasting," *IEEE Trans. Wireless Commun.*, vol. 17, no. 4, pp. 2233–2246, Apr. 2018.
[48] M. Grant and S. Boyd, "CVX: MATLAB software for disciplined convex programming," 2016. [Online] Available: http://cvxr.com/cvx
[49] D. B. West, *Introduction to Graph Theory*. Hoboken, NJ, USA: Prentice Hall, 2001.
[50] C. Rego, D. Gamboa, F. Glover, and C. Osterman, "Traveling salesman problem heuristics: Leading methods, implementations and latest advances," *Eur. J. Oper. Res.*, vol. 211, no. 3, pp. 427–441, 2011.
[51] Z. Wang et al., "Dueling network architectures for deep reinforcement learning," in *Proc. 33rd Int. Conf. Mach. Learn.*, 2016, pp. 1995–2003.
[52] H. Hasselt, A. Guez, and D. Silver, "Deep reinforcement learning with double Q-learning," in *Proc. 30th AAAI Conf. Artif. Intell.*, 2016, pp. 2094–2100.
[53] Technical Specification Group Radio Access Network: Study on Enhanced LTE Support for Aerial Vehicles, document TR 36.777 V15.0.0, 3GPP, Dec. 2017.
[54] 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 Trans. Wireless Commun.*, vol. 22, no. 2, pp. 1343–1358, Feb. 2023.
[55] S. Zhang and R. Zhang, "Trajectory design for cellular-connected UAV under outage duration constraint," in *Proc. IEEE Int. Conf. Commun.*, Shanghai, China, 2019, pp. 1–6.
[56] N. Dal Fabbro, M. Rossi, G. Pillonetto, L. Schenato, and G. Piro, "Model-free radio map estimation in massive MIMO systems via semi-parametric Gaussian regression," *IEEE Wireless Commun. Lett.*, vol. 11, no. 3, pp. 473–477, Mar. 2022.

Cheng Zhan (Member, IEEE) received the BEng and PhD degrees in computer science from the School of Computer Science, University of Science and Technology of China, Anhui, China, in 2006 and 2011 respectively. From 2009 to 2010, he was a research assistant with the Department of Computer Science, City University of Hong Kong. From 2016 to 2017, he was a visiting scholar with the Department of Electrical and Computer Engineering, National University of Singapore. He is currently an associate professor with the School of Computer and Information Science, Southwest University, China. His research interests include unmanned aerial vehicle communications, multimedia communications, wireless sensor networks, and network coding. He served as a TPC member for the IEEE ICC, GLOBECOM, WCNC, and UIC.

Han Hu (Member, IEEE) received the BE and PhD degrees from the University of Science and Technology of China, China, in 2007 and 2012 respectively. He is currently a professor with the School of Information and Electronics, Beijing Institute of Technology, China. His research interests include multimedia networking, edge intelligence and space-air-ground integrated network. He received several academic awards, including Best Paper Award of IEEE TCSVT 2019, Best Paper Award of IEEE Multimedia Magazine 2015, Best Paper Award of IEEE Globecom 2013, etc. He served as an associate editor of *IEEE Transactions on Multimedia* and *Ad Hoc Networks*, and a TPC member of Infocom, ACM MM, AAAI, IJCAI, etc.

Jing Wang (Member, IEEE) received the PhD degree from Peking University, in 2011. She is currently an associate professor with the School of Information, Renmin University of China. Her research interests include edge intelligence and data analytics, computer system for Artificial Intelligence, and energy-efficient computing. She received Beijing Nova Award, she has published papers on top computer conferences such as MICRO, ISCA, HPCA, and IEEE/ACM Transactions journals. She received the Best Paper Award of ICCD, Featured paper of *IEEE Transactions on Computer*, etc. She served as a TPC member of ISCA, NAS, and ASDB.

Zhi Liu (Senior Member, IEEE) received the BE degree from the University of Science and Technology of China, China, and the PhD degree in informatics from the National Institute of Informatics. He is currently an associate professor with The University of Electro-Communications, Japan. His research interest includes video network transmission, vehicular networks and mobile edge computing. He was a recipient of the IEEE StreamComm 2011 Best Student Paper Award, the 2015 IEICE Young Researcher Award, and the ICOIN 2018 Best Paper Award. He is now an editorial board member of *Wireless Networks* (Springer) and *IEEE Open Journal of the Computer Society*.

Shiwen Mao (Fellow, IEEE) received the PhD degree in electrical and computer engineering from Polytechnic University, Brooklyn, New York, in 2004. Currently, he is a professor and Earle C. Williams Eminent Scholar, and director of the Wireless Engineering Research and Education Center, Auburn University, Auburn, Alabama. His research interest includes wireless networks, multimedia communications, and smart grid. He is on the editorial Board of *IEEE Transactions on Wireless Communications*, *IEEE Transactions on Network Science and Engineering*, *IEEE Transactions on Mobile Computing*, *IEEE Internet of Things Journal*, *IEEE Open Journal of the Communications Society*, *IEEE Multimedia*, *IEEE Networking Letters*, ACM GetMobile, and KeAi Digital Communications and Networks Journal, etc. He is the TPC co-chair of IEEE INFOCOM 2018 and TPC vice chair of IEEE GLOBECOM 2022. He received the IEEE ComSoc TC-CSR Distinguished Technical Achievement Award, in 2019 and NSF CAREER Award, in 2010. He is a co-recipient of the IEEE Vehicular Technology Society 2020 Jack Neubauer Memorial Award, the IEEE ComSoc MMTC 2018 Best Journal Award and 2017 Best Conference Paper Award, the Best Demo Award from IEEE SECON 2017, the Best Paper Awards from IEEE GLOBECOM 2019, 2016 & 2015, IEEE WCNC 2015, and IEEE ICC 2013, and the 2004 IEEE Communications Society Leonard G. Abraham Prize in the Field of Communications Systems.