IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 30 December 2023, accepted 17 January 2024, date of publication 19 January 2024, date of current version 25 January 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3356175

RESEARCH ARTICLE

Age of Information Oriented Data Collection via Energy-Constrained UAVs in Wireless Sensor Networks

ZHENG ZHOU, JUAN LIU, (Member, IEEE), AND CHIXIONG MAO
School of Electrical Engineering and Computer Science, Ningbo University, Ningbo, Zhejiang 315211, China

Corresponding author: Juan Liu (liujuan1@nbu.edu.cn)

This work was supported in part by the Zhejiang Provincial Natural Science Foundation of China under Grant LQ21F010006, and in part by the National Natural Science Foundation of China under Grant 62101289.

ABSTRACT In this paper, an age of information (AoI) oriented data collection scheme is proposed for wireless sensor networks with energy-constrained unmanned aerial vehicles (UAVs). Different from traditional approaches that focus on one single UAV, we advocate the activation of multiple candidate UAVs to improve the AoI performances. However, coordinating multiple UAVs complicates the data collection strategies of each UAV, rendering AoI minimization a formidable challenge. To address the intractable formulated optimization problem, we propose a heuristic two-step algorithm based on greedy search for multiple UAVs. Specifically, we address the coordination of multiple UAVs through a novel algorithm grounded in graph theory and kernel K-means methods. Concurrently, we devise an energy-constrained trajectory planning algorithm to optimize data collection for each activated UAV. The numerical results validate the accuracy of the proposed algorithm, and demonstrate that both average AoI and energy consumption per UAV can be reduced when more energy-constrained UAVs are involved in the data collection process for wireless sensor networks.

INDEX TERMS Age of information, graph theory, wireless sensor network, multi-UAV.

I. INTRODUCTION

In recent years, unmanned aerial vehicles (UAVs) aided wireless sensor networks (WSNs) have attracted much attention [1]. Equipped with wireless communication platforms, the UAVs can be employed as mobile base stations or data collectors to communicate with the ground sensor nodes (SNs) in a flexible and reliable manner. In UAV-aided WSNs, most existing works focused on optimizing traditional performance metrics, such as system throughput, coverage, and delay [2], [3], [4], [5]. However, these performance metrics are not suitable for UAV-aided WSNs with mission-critical applications [6], such as smart factories, city traffic surveillance, disasters monitoring and so on. In the mission-critical applications, the systems usually run in real time and the freshness of data is of great importance. To measure the data freshness, a new metric age of information (AoI) is becoming a popular concept for UAV-aided WSNs, which characterizes the time interval from the generation of the latest received information to the current time [7].

In UAV-aided WSNs, the UAV is assigned to collect data from SNs within their communication ranges, and return to the data center. AoI of the UAV-aided WSNs is determined by the time interval between data collection from SNs to UAVs and data offloading from UAVs to data center. As a result, the UAVs' flight trajectories during the execution of data collection tasks need to be designed carefully.

Considering the limited energy storage at UAVs, it's very challenge for one single UAV to complete data collection. Therefore, the multi-UAV-aided data collection has emerged to improve the AoI performances in the WSNs. Most existing works employ optimization methods [8], [9] and/or deep reinforcement learning (DRL) methods to find the UAVs' flight trajectories [10], [11]. However, utilizing the overlapped area of communication ranges of different SNs is ignored in AoI minimization. Generally, the UAVs visit all N

The associate editor coordinating the review of this manuscript and approving it for publication was Renato Ferrero.

SNs sequentially to fulfill the data collection task. However, in a dense network, the communication ranges of the SNs might overlap with each other, and the UAV need not to visit each SN individually. Instead, $L \leq N$ hovering points (HPs) can be selected at which each UAV collects data from multiple SNs concurrently. However, this scenario introduces new challenges. Firstly, determining the optimal number of HPs is crucial; generally, fewer HPs results in reduced AoI. Secondly, designing the positions of these HPs significantly impacts AoI performance.

To solve the above challenges, an average AoI minimization problem is studied in a multi-UAV-aided WSN subject to the energy capacity of each UAV. To utilize the overlapped area effectively, graph theory and optimization approaches are introduced in this paper for AoI minimization. The main contributions can be summarized as follows:

- We propose an AoI-oriented data collection scheme for WSNs with multiple energy-constrained UAVs. In our system, UAVs do not directly fly to SNs but rather navigate to designated hovering points where they establish communication with all SNs within range.
- We propose a heuristic two-step algorithm to minimize the AoI of the proposed scheme. In Step 1, the hovering points are determined based on graph theory and optimization methods, and the UAV-SN association is established via the hovering points based on the application of kernel K-means. In Step 2, after establishing the UAV-SN association, the AoI-oriented trajectories for energy-constrained UAVs are found via optimization methods.
- We conduct extensive simulations to evaluate the performances of the proposed algorithm. Simulation results confirm the correctness of our proposed algorithm and showed that with a finite energy supply, our proposed data collection scheme can achieve significant AoI performance gains.

The rest of this paper is organized as follows. Section II introduces the system model. The multi-UAV aided data collection problem is formulated in Section III. Then, we design algorithms to solve the multi-UAV aided data collection problem in Section IV. Simulation results and conclusions are presented in Section V and Section VI, respectively. Throughout this paper, we use $|\mathcal{A}|$ to denote the number of elements in the set $\mathcal{A}$ and $\|\cdot\|$ to represent the Euclidean distance, respectively.

A. RELATED WORKS

In UAV-aided WSN networks, it is very challenging to find the UAVs' flight trajectories during the execution of data collection tasks. One is to establish the association between the UAVs and SNs. The other is to find optimal flight trajectories for the UAVs. To deal with these two challenges, one avenue of research incorporates optimization methods to find optimal trajectories for the UAVs. For example, in our previous work [9], the hovering points of the UAV ware determined by applying the Affinity Propagation algorithm and the age-optimal trajectory was found using dynamic programming without considering energy constraint on the UAV. In [12], the peak and average AoIs were derived for the UAV-aided data gathering and dissemination in a graph, and minimized by optimizing the UAV's randomized flight trajectories, respectively. Jia et al. studied the UAV path planning and data acquisition problem using dynamic programming methods [13]. In [14], the authors studied the joint optimization of flight trajectory and time allocation for AoI-oriented UAV-assisted wireless powered Internet of Things (IoT) systems, where one UAV was applied to transfer energy and collect data from the ground SNs. In [15], the authors found the path of a UAV that will minimize the max AoI from any data collected from the field, given the data generation times and locations of IoT devices. The problem was modeled as a mixed integer convex optimization problem using graph theory and solved with CVX tools. Also, a heuristic-based, fast-running, practical solution was provided. In [8], Zhan et al. delved into the optimization methods for addressing the multi-UAV-aided data collection problem. Their work focused on the joint optimization of SN-UAV associations, UAV trajectory, and wake-up scheduling of SNs, aimed at minimizing the maximum completion time.

Different from the above works, data collection via energy-constrained UAVs have been considered in [16], [17], and [18]. In [16], Li et al. focused on developing efficient heuristic and approximation algorithms for the data collection maximization problems in which the hovering locations with sojourn durations were optimized for an energy-constrained UAV. Wherein, the hovering region of the UAV was partitioned into finite numbers of equal squares and the square centers were identified as the potential hovering points of this UAV. And a version of Traveling Salesman Problem based tour was found for the energy-constrained UAV. In [17], the authors studied the problem of minimizing the average peak AoI for a source-destination pair in a UAV-assisted IoT network, where the UAV's flight trajectory, energy allocations and service time durations at the source and the UAV were jointly optimized. To minimize the total AoI within a period of time, the authors of [18] proposed to jointly optimize the UAV's sensing time, transmission time, flight trajectory, and task scheduling in a cellular Internet with UAVs.

In the other lines of works, popular DRL methods, like deep Q networks (DQN), have been recently introduced to make real-time decisions on UAVs' movements and resource allocation strategies [10], [11]. The authors of [19] designed the UAV's online flight trajectory in the DRL framework to achieve the SNs' minimum weighted sum of AoIs. Moreover, in [20], Shokry et al. designed a DRL algorithm to find the trajectories of the deployed UAVs and scheduling of status updates to minimize the expected weighted sum AoI for UAV-assisted vehicular networks. Although these widely used DRL methods can do real-time decision-making regarding UAV movement and resource

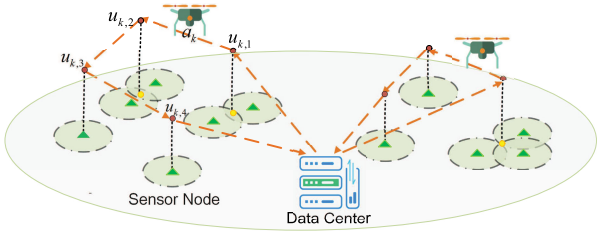


FIGURE 1. The multi-UAV-aided data collection model.

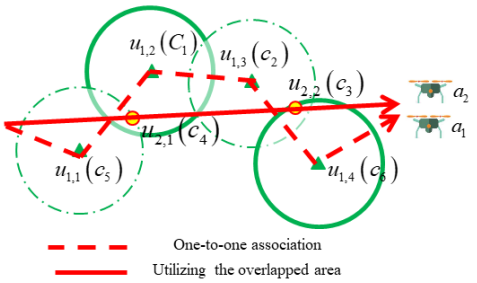


FIGURE 2. Illustration of the HP selection.

allocation strategies, but often necessitate extensive training periods before practical implementation.

In contrast to these works, we focus on applying graph theory and optimization approaches to deal with the above two challenges in multi-UAV-aided data collection scenarios. Different from our previous work [21] that focuses on the maximum AoI scenario, this work discusses the average AoI performances. Besides, the major difference between this work and [22] lies in the design of UAV-SN association.

II. SYSTEM MODEL

A. NETWORK DESCRIPTION

We consider a multi-UAV-enabled wireless sensor network consisting of a data center v_0, K rotary-wing UAVs and N ground sensor nodes (SNs). The set of SNs is denoted by $\mathcal{V} = \{v_1, v_2, \cdots, v_N\}$, where v_n is the n-th SN. The locations of the data center and the SNs are denoted by $p_n = (x_n, y_n) \in \mathbb{R}^2 (n = 0, 1, \cdots, N)$ unifiedly.The set of the K candidate UAVs is denoted by $\mathcal{A} = \{a_1, a_2, \cdots, a_K\}$, where a_k is the k-th UAV. Suppose that each UAV a_k flies at a fixed velocity V_k and altitude H_k. Each UAV initially carries a certain amount of energy, denoted by E_{max} (Joules). Similar to the assumptions in [24], each SN has a communication range denoted by R, and the neighborhood area with radius r is used to characterize the maximum horizontal distance between each SN and its receiver in accordance with the communication range. Considering that different UAVs may fly at different altitudes, the neighborhood area of SN v_n is denoted by $\mathcal{N}_n = \{p = (x, y) | \parallel p - p_n \parallel = \sqrt{(x - x_n)^2 + (y - y_n)^2} \leq r\} (n = 1, \cdots, N)$, where $r = \sqrt{R^2 - (\max_k H_k)^2}$.

During data collection, each UAV flies to a position within the communication range of a SN, and hovers to gather sensing data from the SN. Once the data is acquired, the UAV proceeds to other SNs before returning to the data center for data offloading. These strategic positions for UAVs to receive messages are termed "hovering points" (HPs). An example of this data collection process is depicted in Fig. 1, UAV a_k starts from the data center, navigating and hovering successively at HPs $u_{k,1}, u_{k,2}, u_{k,3}, u_{k,4}$ to collect data from the SNs, and finally returning to the data center. Consequently, the trajectory of UAV a_k can be described by $u_k = [v_0, u_{k,1}, u_{k,2}, u_{k,3}, u_{k,4}, v_0]$, where $u_{k,j}$ denotes the j-th HP in the trajectory of UAV a_k. For simplicity, we use $u'_k = [u_{k,1}, u_{k,2}, u_{k,3}, u_{k,4}]$ to describe the trajectory of UAV a_k. It is noteworthy that the selection of HPs significantly influences UAV flight time, thereby affecting the AoI performance. Generally, the fewer HPs are selected, the less the AoI is induced. Moreover, the UAV flight trajectory should be designed to determine the optimal sequence for each UAV's traversal across HPs. This is because, for each SN, the sooner it is visited by any UAV, the larger AoI it has. Consequently, the UAV flight trajectory defines the sequence of HPs to be visited, which should be optimized to reduce the average AoI of SNs.

B. SN-HP ASSOCIATION AND HP-UAV ASSOCIATION

As can be seen in Fig. 2, UAV a_1 can access four SNs individually via four HPs (namely, $u_{1,1}, u_{1,2}, u_{1,3}, u_{1,4}$), denoting a one-to-one association. However, when the communication ranges of different SNs overlap, a UAV can efficiently cover this overlapped area, collecting data from multiple SNs simultaneously. For instance, these four SNs can be visited by UAV a_2 using only two HPs (i.e., $u_{2,1}, u_{2,2}$ in Fig. 2). Consequently, the selection of HPs, both in terms of number and locations, should be carefully designed.

Let $\mathcal{C} = \{c_1, c_2, \cdots, c_L\}$ denote the set of potential HPs, where L denotes the number of potential HPs. The coordinates of HP c_l are denoted by $s_l = (x_l^c, y_l^c) \in \mathbb{R}^2 (l = 1, \cdots, L)$. The subset of HPs visited by UAV a_k is denoted as $\mathcal{C}_k \subseteq \mathcal{C}$. Hence, $u_{k,j} \in \mathcal{C}_k$ denotes the j-th HP in the trajectory of UAV a_k. Moreover, $u'_k = [u_{k,1}, u_{k,2}, \cdots, u_{k,|\mathcal{C}_k|}]$ denotes all the HPs in trajectory u_k, and therefore u'_k is a permutation of the set $\mathcal{C}_k$. Let's take Fig. 2 as an example. The set of potential HPs is $\mathcal{C} = \{c_1, c_2, \cdots, c_6\}$. In the data collection scheme based on one-to-one association, the set of selected HPs is denoted by $\mathcal{C}_1 = \{c_1, c_2, c_5, c_6\}$. These HPs are visited in the order c_5, c_1, c_2, c_6. Thus, the trajectory of UAV a_1 can be described as $u'_1 = [u_{1,1} = c_5, u_{1,2} = c_1, u_{1,3} = c_2, u_{1,4} = c_6]$, and u'_1 is a permutation of $\mathcal{C}_1$. Similarly, for the scheme based on utilizing the overlapped area, $u'_2 = [u_{2,1} = c_4, u_{2,2} = c_3]$.

Then, we introduce the concept of SN-HP association. In particular, we use $\zeta_{n,l}$ to represent the association between SN v_n and HP c_l. If $\zeta_{n,l} = 1$, it indicates that SN v_n can be visited via HP c_l by a single UAV; conversely, if $\zeta_{n,l} = 0$, the association doesn't exist. For example, consider the SNs represented as v_1, v_2, v_3, v_4 from left to right in Fig. 2. In the scheme utilizing the overlapped area, the UAV hovers at HP c_4 to communicate with SN 1 and 2, and hovers at HP c_3

to communicate with SN 3 and 4. Consequently, the SN-HP association can be described as $\zeta_{1,4}=1, \zeta_{2,4}=1, \zeta_{3,3}=1, \zeta_{4,3}=1$. Similarly, in the one-to-one association scheme, we have $\zeta_{1,5}=1, \zeta_{2,1}=1, \zeta_{3,2}=1, \zeta_{4,6}=1$.

To describe the activation of the potential UAVs, we introduce the HP-UAV association. In particular, we use $\eta_{l,k}$ to represent the association between HP c_l and UAV a_k. If $\eta_{l,k}=1$, it means that UAV a_k is assigned to visit HP c_l, conversely, if $\eta_{l,k}=0$, the association doesn't exist. Consider Fig. 2 as an example. If the scheme utilizing the overlapped area is adopted, UAV a_2 is assigned to visit v_1, v_2, v_3, v_4 through c_4, c_3, while UAV a_1 might be assigned to visit the rest SNs. Consequently, we have $\eta_{4,1}=1, \eta_{3,1}=1$. Similarly, if the scheme based on one-to-one association is adopted, we have $\eta_{5,2}=1, \eta_{1,2}=1, \eta_{2,2}=1, \eta_{6,1}=1$.

In conclusion, to complete data collection, we need to find the optimal SN-HP association and HP-UAV association, and then find the optimal trajectory of UAVs. Keeping in mind that each SN is associated to one HP and each HP is visited by one single UAV, we have

$$\sum_{l=1}^{L}\zeta_{n,l}=1, \quad \forall n, \quad \sum_{k=1}^{K}\eta_{l,k}=1, \quad \forall l. \tag{1}$$

The set of SNs covered by HP c_l is denoted as $\mathcal{V}_l=\{v_n|\zeta_{n,l}=1, v_n\in\mathcal{V}\}$. Similarly, the set of SNs collected by UAV a_k is denoted by $\widetilde{\mathcal{V}}_k=\{v_n|\sum_l\zeta_{n,l}\eta_{l,k}=1, v_n\in\mathcal{V}\}$. Therefore, the number of SNs visited by the UAVs satisfies

$$\sum_{k=1}^{K}\sum_{l=1}^{L}\sum_{n=1}^{N}\zeta_{n,l}\cdot\eta_{l,k}=N. \tag{2}$$

Let ζ and η denote the vectors of variables $\{\zeta_{n,l}\}$ and $\{\eta_{l,k}\}$, respectively.

C. TRANSMISSION MODEL

Similar to the assumptions in [26], the channel gain between the UAVs and SNs is expressed as

$$h_{n,l}^{k}=\begin{cases}\beta_0\left[D_{n,l}^{k}\right]^{-\varrho}, & \text{line-of-sight link},\\ \zeta\beta_0\left[D_{n,l}^{k}\right]^{-\varrho}, & \text{non-line-of-sight link},\end{cases} \tag{3}$$

where $D_{n,l}^{k}=\sqrt{\|p_n-s_l\|^2+H_k^2}$ is the distance between SN v_n and UAV a_k that is hovering over HP c_l, β_0 denotes the channel gain at the reference distance of one meter, ϱ is the path loss exponent, and $\zeta\in(0,1)$ is the additional attenuation factor due to the non-line-of-sight effects.

Based on the generate-at-will policy [25], we assume that when reaching an HP, the UAV sends a message to all the associated SNs. Upon receiving the message, the SNs begin to sample the environment and upload their sensing data to the UAV by some multiple access scheme. In particular, assuming that the signaling time is negligible, the hovering time T_l^k can be evaluated as the uploading time of the associated SNs. Moreover, when the simple time division multiple access scheme is adopted, the hovering time T_l^k is expressed as

$$T_l^k=\sum_{v_n\in\mathcal{V}_l}\frac{W_n}{R_{n,l}^{k}}, \tag{4}$$

where W_n is the size of one packet at SN v_n, and $R_{n,l}^{k}$ is the uploading data rate. The UAV always gets successful reception if each SN transmits at a data rate adapted to the worse non-line-of-sight channel condition. In this case, the data rate is calculated as

$$R_{n,l}^{k}=B\log\left(1+\frac{\zeta\beta_0\left[D_{n,l}^{k}\right]^{-\varrho}}{\sigma^2+I^k}P_s\right), \tag{5}$$

where B is the system bandwidth, σ^2 and I^k denote the noise power and interference at the receiver of UAV a_k, respectively, and P_s is the transmit power at each SN. Similarly, we can calculate the offloading time when the UAV flies back and offloads to the data center. Considering the line-of-sight link between UAV a_k and the data center, the offloading time can be expressed as

$$T_o^k=\frac{\sum_{v_n\in\widetilde{\mathcal{V}}_k}W_n}{B\log\left(1+\frac{\beta_0H_k^{-\varrho}}{\sigma^2+I^o}P_u\right)}, \tag{6}$$

where I^o denotes the inteference at the receiver of the data center and P_u represents the transmit power at the UAV. Suppose that the K UAVs are dispatched separately in time and space so as to avoid crash and severe interference between each other. Hence, the interference in (5) and (6) can be modeled as additive white gaussian noise.

III. PROBLEM FORMULATION

In this section, the performance of the multi-UAV aided data collection scheme is measured by the SNs' AoI subject to the energy capacity of each UAV.

A. AVERAGE AOI OF THE SNS

We denote by $X_n^k(t)$ the AoI of SN v_n whose data is collected by UAV a_k. It can be defined as $X_n^k(t)=(t-t_{n,l}^{k})^+$, where $t_{n,l}^{k}$ is the instant at which SN v_n samples the environment and packs the sensing data into a data packet, and $(x)^+=\max\{x,0\}$. From [23], if SN v_n is covered by the j-th HP in the trajectory u_k, the AoI of this SN can be expressed as

$$X_n^k(u_k)=(t^k-t_{n,l}^{k})^+=\sum_{l=j}^{|\mathcal{C}_k|}\left(T_l^k+\tau_{l,(l+1)}^{k}\right)+T_o^k, \tag{7}$$

where t^k is the instant when UAV a_k finishes offloading all the collected data to the data center, T_l^k is the hovering time of UAV a_k at HP $u_{k,l}$, and $\tau_{l,(l+1)}^{k}$ is the flight time of UAV a_k from HP $u_{k,l}$ to HP $u_{k,(l+1)}$ in the trajectory u_k.

Furthermore, based on (7), the average AoI of the SNs collected by UAV a_k can be computed as

$$X_{k,ave}(u_k) = \frac{1}{|\widetilde{\mathcal{V}}_k|}\sum_{j=1}^{|\mathcal{C}_k|}\phi_j^k\left[\sum_{l=j}^{|\mathcal{C}_k|}\left(T_l^k + \tau_{l,(l+1)}^k\right) + T_o^k\right]$$
$$= \frac{1}{|\widetilde{\mathcal{V}}_k|}\sum_{j=1}^{|\mathcal{C}_k|}\sum_{l=1}^{j}\phi_l^k\left(T_j^k + \frac{d_{j,(j+1)}^k}{V_k}\right) + T_o^k, \quad (8)$$

where ϕ_j^k denotes the number of SNs covered by the j-th HP $u_{k,j}$ in the trajectory u_k, and the second equality holds due to $|\widetilde{\mathcal{V}}_k| = \sum_{l=1}^{|\mathcal{C}_k|}\phi_l^k$. Notice that ϕ_j^k depends on the SN-HP association ζ and can be counted as $\phi_j^k = |\mathcal{V}_l| = \sum_{n=1}^{N}\zeta_{n,l}$ when HP c_l is selected as $u_{k,j}$.

B. ENERGY CONSUMPTION MODEL

For a rotary-wing UAV, the propulsion power consumption mainly depends on the UAV's flight velocity and the acceleration. Similar to the assumptions in [26], we ignore the power consumed for acceleration or deceleration. This is because the acceleration or deceleration time is very small in comparison to flight and data transmission times. As a result, the propulsion power can be expressed as

$$P(V_k) = P_0\left(1 + \frac{3V_k^2}{U_{tip}^2}\right) + P_1\sqrt{\sqrt{1 + \frac{V_k^4}{4V_0^4}} - \frac{V_k^2}{2V_0^2}}$$
$$+ \frac{1}{2}d\rho s A_r V_k^3, \quad (9)$$

where P_0 and P_1 are the blade profile and induced powers of the UAV in the hovering status, respectively, V_0 is the mean rotor induced velocity of the UAV in hovering, U_{tip} is the tip speed of the rotor blade, ρ is the air density, d, s and A_r are the fuselage drag ratio, rotor solidity, and the rotor disc area, respectively.

When UAV a_k flies along the trajectory u_k to collect data, the amount of aggregate energy can be calculated as

$$e_k(u_k) = P(V_k)\sum_{l=0}^{|\mathcal{C}_k|}\frac{d_{l,(l+1)}^k}{V_k} + P(0)\cdot\sum_{l=1}^{|\mathcal{C}_k|}T_l^k + P_u\cdot T_o^k, \quad (10)$$

where $d_{0,1}^k$ is the flight distance when UAV a_k flies from the data center to the first HP $u_{k,1}$, P_u denotes the transmit power at each SN. $P(0)$ is the power consumption when the UAV hovers at velocity $V_k = 0$, i.e., $P(0) = P_0 + P_1$. The three terms in (10) mean the propulsion energy, hovering energy, and transmission energy, respectively.

C. ENERGY-CONSTRAINED AOI MINIMIZATION PROBLEM

In this work, we attempt to jointly design the UAV-SN association and the UAVs' flight trajectories via the HPs to minimize the SNs' average AoI subject to the energy capacity at the UAVs. In particular, the HPs' positions, the SN-HP and HP-UAV associations, and the permutation of the HPs visited by each UAV should be found. Let $s = [s_1, s_2, \cdots, s_L]$ and $u = [u_1, u_2, \cdots, u_K]$ denote the vectors of the HPs' locations and the UAVs' trajectories, respectively. To find the optimal multi-UAV-aided data collection solution, we formulate an optimization problem as follows:

$$\min_{s,\zeta,\eta,u} X_{ave} = \frac{1}{N}\sum_{k=1}^{K}\left[\sum_{j=1}^{|\mathcal{C}_k|}\sum_{l=1}^{j}\phi_l^k\left(T_j^k + \frac{d_{j,(j+1)}^k}{V_k}\right) + T_o^k\right]$$
$$s.t. \begin{cases} e_k(u_k) \leqslant E_{max}, \forall k, & (a) \\ (1), (2), & (b) \\ s_l \in \mathbb{R}^2, \zeta_{n,l}, \eta_{l,k} \in \{0,1\}, u_k' \in \sigma(\mathcal{C}_k). & (c) \end{cases} \quad (11)$$

where $u_k = [v_0, u_k', v_0]$ is determined by the permutation of HPs u_k', and $\sigma(\mathcal{C}_k)$ is the set including all the permutations of the set $\mathcal{C}_k$. Constraint (11.a) means that the amount of energy consumed by each UAV cannot exceed its energy storage. Constraints (11.b) points out the limitation on the SN-HP and HP-UAV associations, and (11.c) specifies the ranges of the variables.

Note that, the variables in Problem (11) are coupled, and the variables $s_l \in \mathbb{R}^2$ are continuous while the other variables are discrete. As a result, Problem (11) is a mixed-integer programming problem and is generally intractable.

Algorithm 1 The Graph Theory Based SN-HP Association Algorithm

Require: The network topology $G = (\mathcal{V}, \epsilon)$, including the location p_n and communication radius r of each SN v_n;

Part I: Selection of Candidate HPs
Set the neighborhood area of each SN $\mathcal{N}_n$ ($v_n \in \mathcal{V}$);
Find a set of SNs $\mathcal{I}_{mk} \subseteq \mathcal{V}$ with their neighborhood areas constructing a maximal independent set;
Find a minimum spanning tree $\mathcal{I}_{tree}$ in the graph G;
Select the intersections of the edges in $\mathcal{I}_{tree}$, and the boundaries of the neighborhood areas of the SNs in $\mathcal{I}_{mk}$ as candidate HPs, and add them to the set $\mathcal{C}$.
Part II: HP pruning
Set $\mathcal{V}' = \mathcal{V}$ and $\mathcal{C}^* = \emptyset$;
repeat
  Select HP c_l that covers a maximum number of SNs, i.e., $c_l = \arg\max_{c_l \in \mathcal{C}} \sum_{v_n \in \mathcal{V}'} b_{n,l}$;
  Add c_l to $\mathcal{C}^*$, i.e., $\mathcal{C}^* = \mathcal{C}^* \cup \{c_l\}$;
  Find the SNs in the set $\mathcal{V}'$ satisfying $b_{n,l} = 1$, and set $\zeta_{n,l}^* = 1$ accordingly;
  Remove the SNs covered by HP c_l from $\mathcal{V}'$;
until $\mathcal{V}' = \emptyset$

Ensure: The set of HPs with their coordinates $\{s_l\}(c_l \in \mathcal{C}^*)$ and the SN-HP association ζ^*.

IV. ALGORITHM DESIGN

The formidable nature of Problem (11) necessitates the adoption of sub-optimal algorithms. It's important to note that

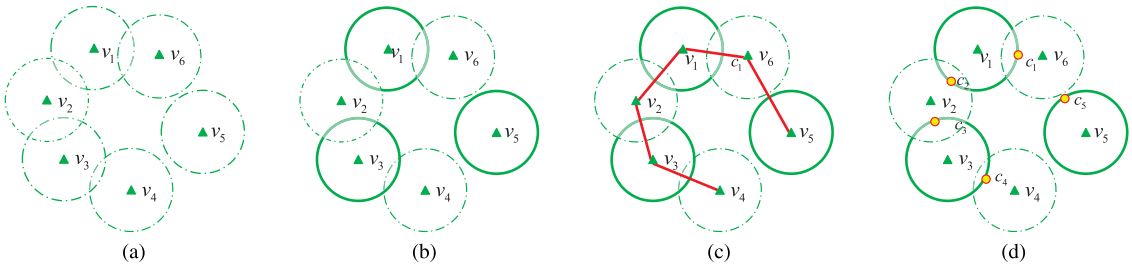


FIGURE 3. Illustrate how to find candidate HPs: (a) Set a neighborhood area for each SN; (b) Construct a maximal independent set of neighborhood areas of the SNs in $\mathcal{I}_{mk} = \{v_1, v_3, v_5\}$; (c) Find a minimum spanning tree $\mathcal{I}_{tree}$ consisting of the six SNs $\{v_1, \cdots, v_6\}$ and five edges drawn by red lines; (d) Select the intersections of the edges in $\mathcal{I}_{tree}$ and boundaries of the SNs in $\mathcal{I}_{mk}$ as candidate HPs, i.e., $\{c_1, c_2, c_3, c_4, c_5\}$.

the pursuit of an optimal UAV-SN association (comprising both SN-HP and HP-UAV associations) carries a heavy computational burden. To achieve a better balance between the computation complexity and the performances improvement, we present a novel algorithm based on a combination of graph theory and kernel K-means clustering. This optimization results in a refined UAV-SN association solution, which simplifies Problem (11) into an age-optimized trajectory design problem.

A. UAV-SN ASSOCIATION VIA GRAPH THEORY AND KERNEL K-MEANS CLUSTERING

Initially, our focus is on optimizing the SN-HP association. It is essential to recognize that by selecting a smaller number of potential HPs for data collection, we can reduce the UAVs' flight time, thus ensuring the freshness of the SNs' information while avoiding unnecessary energy expenditure. Consequently, we choose the optimized solution for SN-HP association that minimizes the activation of HPs as the solution to Problem (11).

Specifically, our approach involves identifying a set of potential HPs through graph theory. Subsequently, we refine this set by selecting a subset of these candidates to ensure comprehensive coverage of all N SNs, establishing the corresponding SN-HP association.

(1) We first construct a maximal independent set (MIS) of the SNs' neighborhood areas (Line 3). Specifically, the sets of unmarked and marked SNs are initialized as $\mathcal{I}_{un} = \mathcal{V}$ and $\mathcal{I}_{mk} = \emptyset$, respectively. One SN v_n is randomly selected from the set $\mathcal{I}_{un}$, and is marked and added to the set $\mathcal{I}_{mk}$. Then, the SNs with their neighborhood areas $\mathcal{N}_j$ ($v_j \in \mathcal{I}_{un}$) overlapping with the neighborhood area $\mathcal{N}_n$ are deleted from the set $\mathcal{I}_{un}$. This procedure is repeated till the set $\mathcal{I}_{un}$ becomes empty. The neighborhood areas of the SNs in $\mathcal{I}_{mk}$ are disjoint and constitute a maximal independent set [27], as plotted in bold in Fig. 3(b).

(2) Next, we find a minimum spanning tree (MST) $\mathcal{I}_{tree}$ in the graph $G = (\mathcal{V}, \epsilon)$ (Line 4). Using classic algorithms, such as Prim's algorithm, we can found an MST $\mathcal{I}_{tree}$ that connects the N SNs with $N-1$ edges [27]. An exemplar MST $\mathcal{I}_{tree}$ is plotted in Fig. 3(c), where the vertices represent the SNs and the edges in $\mathcal{I}_{tree}$ are drawn in red lines.

(3) Last, we select the intersections of the edges in $\mathcal{I}_{tree}$ and the neighborhood boundaries of the marked SNs in $\mathcal{I}_{mk}$ as candidate HPs (Line 5), as shown in Fig. 3(d).

By applying the above three steps, we find a set of candidate HPs, and the relationship between the candidate HPs and the SNs is determined. As shown in Fig. 3, each SN could be covered by several candidate HPs. For example, SN v_2 can be visited from either c_2 or c_3, and all the six SNs can be covered by candidate HPs $\{c_1, c_3, c_4, c_5\}$ or $\{c_1, c_2, c_4, c_5\}$. However, it is enough for one UAV to collect data from each SN at one HP. Hence, it is possible to prune out some of the candidate HPs to save energy and flight time for the UAVs.

The number of candidate HPs is denoted by L_c. Let $b_{n,l}$ be a binary indicator specifying whether SN v_n is covered by the candidate HP c_l. We have $b_{n,l} = 1$, if candidate HP c_l is in the neighborhood area of SN v_n. Otherwise, $b_{n,l} = 0$. To minimize the number of HPs while providing coverage to all the SNs, we formulate a linear integer programming problem as follows:

$$\min_{\zeta_{n,l}, C_l} \sum_{l=1}^{L_c} C_l$$
$$s.t. \begin{cases} \sum_{l=1}^{L_c} \zeta_{n,l} b_{n,l} = 1, \forall n, & (a) \\ \zeta_{n,l} \le C_l, \forall n, l, & (b) \\ \zeta_{n,l}, C_l \in \{0,1\}, \forall n, l, & (c) \end{cases} \tag{12}$$

where C_l is a binary variable indicating whether HP c_l is activated. The first constraint (12.a) means each SN is associated with one single HP. The second constraint (12.b) makes sure that each HP c_l is able to provide coverage for some SN v_n only when it is activated, i.e., $C_l = 1$. The third constraint (12.c) points out the binary variables $\{\zeta_{n,l}\}$ and $\{C_l\}$. By solving the problem (12), we obtain the optimal variables $\{C_l^*\}$ and $\{\zeta_{n,l}^*\}$. Then, the candidate HPs with $C_l^* = 0$ are removed. Thus, $L = \sum_{l=1}^{L_c} C_l^*$ HPs are determined. The optimal SN-HP association $\zeta^* = [\zeta_{n,l}^*]$ is achieved accordingly.

However, it is not easy to solve the integer programming problem (12) when the number of SNs and that of candidate HPs are very large. Taking into account that one SN is associated to one single HP, we would like to select the

activated HPs from the candidates greedily (Line 5-Line 11). Initially, we set $\mathcal{V}' = \mathcal{V}$ and $\mathcal{C}^* = \emptyset$. The HP covering a maximum number of SNs is selected from the set of candidates, i.e., $c_l = \arg\max_{c_l \in \mathcal{C}} \sum_{v_n \in \mathcal{V}'} b_{n,l}$. By checking the SNs in the set $\mathcal{V}'$, i.e., $b_{n,l} = 1$ ($v_n \in \mathcal{V}'$), we find the SNs covered by this HP, and set the SN-HP association variable as $\zeta^*_{n,l} = 1$ for these SNs. Then, the HP is removed from the set of candidates $\mathcal{C}$ and its associated SNs are removed from the set $\mathcal{V}'$, respectively. This procedure repeats till all the SNs are checked, i.e., $\mathcal{V}' = \emptyset$. The above procedures are summarized in Algorithm 1, where the number of the selected HPs, and the SN-HP association can be optimized. By applying Algorithm 1, we get L activated HPs with their coordinates, and the SN-HP association. The computational complexity of Algorithm 1 is about $\mathcal{O}(N^2)$, since each step in HP selection and pruning requires at most N^2 scalar operations. Hence, this algorithm is appropriate when the number of SNs N is quite large.

Secondly, a kernel K-means [28] based algorithm is designed to find the optimized HP-UAV association. The main idea of kernel K-means is to map the input space to a higher-dimensional feature space using a nonlinear kernel function, and partition them by linear separators in the kernel space [28]. Let m_k denote the center of the k-th cluster in the feature space. The cluster center m_k is the best representative of each cluster in the feature space, and can be calculated as

$$m_k = \frac{\sum_{c_l \in \mathcal{C}_k} \Phi(s_l) + \Phi(p_0)}{|\mathcal{C}_k| + 1}, \tag{13}$$

where $\Phi(\cdot)$ denote the nonlinear function. From [28], the objective function of kernel K-means is defined as

$$J(\eta_{l,k}) = \sum_{k=1}^{K} \left(\sum_{l=1}^{L} \eta_{l,k} D^2_{l,k} + D^2_{0,k} \right), \tag{14}$$

where $D_{l,k} = \| \Phi(s_l) - m_k \|$ is the distance between HP c_l and the cluster center m_k in the feature space, and $D_{0,k} = \| \Phi(p_0) - m_k \|$ is the distance between m_k and the data center v_0 in the feature space. By substituting the kernel function $\kappa(s_i, s_j)$, we can further express the square of the distance as

$$D^2_{l,k} = \| \Phi(s_l) - m_k \|^2 = \kappa(s_l, s_l) - 2\kappa(s_l, s^h_k) + \kappa(s^h_k, s^h_k), \tag{15}$$

where s^h_k is the coordinate of the center in the k-th cluster. Among the popular examples, polynomial and Gaussian kernel functions are widely applied.

The objective is to minimize the function $J(\eta_{l,k})$ by optimizing the HP-UAV association $\eta_{l,k}$ gradually. In the first iteration with $i = 0$, the set of HPs $\mathcal{C}^*$ is randomly divided into K clusters and each cluster center $m_k(i)$ is calculated by (13). Then, each HP is assigned to one of the clusters according to the nearest neighbor principle, i.e.,

$$\eta_{l,k}(i) = \begin{cases} 1, & \text{if } D_{l,k} \le D_{l,j},\ \forall k \ne j \\ 0, & \text{otherwise.} \end{cases} \tag{16}$$

Algorithm 2 The Kernel K-Means Based HP-UAV Association Algorithm for K UAVs

Require: The number of UAVs K, the set of HPs $\mathcal{C}^*$, the coordinates of the HPs $\{s_l\}$, the coordinate of the data center p_0, and kernel function $\kappa(\cdot, \cdot)$.

1: Pick up K HPs from the set $\mathcal{C}^*$ randomly and set them as the initial cluster centers $\{s^h_k(0)\}$;
2: Set the iteration index $i = 0$, and calculate the cluster centers in the feature space $m_k(i) = \Phi(s^h_k(i))$;
3: **repeat**
4: Assign the association variable $\eta_{l,k}(i)$ by (16);
5: Update the cluster centers in the kernel space $m_k(i)$ by (13);
6: Increase the iteration index by one: $i = i + 1$;
7: **until** The K cluster centers do not change.

Ensure: The HP-UAV association variable $\eta^*_{l,k}$ and the clustering result $\mathcal{C}^*_k$.

In this way, the clustering result $\mathcal{C}_k(i) = \{c_l \in \mathcal{C}^* | \eta_{l,k}(i) = 1\}$ is found, and the new cluster centers $\{m_k(i+1)\}$ are calculated. This iterative procedure continues till the cluster centers do not change any more. The kernel K-means clustering based HP-UAV association algorithm is presented in detail in Algorithm 2. Notice that Algorithm 2 needs $2(N+1)K$ extra operations to calculate the distances between the $(N+1)$ nodes and the K cluster centers in the 2-dimensional space. The overall complexity of Algorithm 2 is about $\mathcal{O}((N+1)K(T_{iter}+2))$ scalar operations, where T_{iter} is the maximum number of iterations.

B. TRAJECTORY PLANNING FOR ENERGY-CONSTRAINED UAVS

Note that, by applying Algorithms 1 and 2, we establish the SN-HP and HP-UAV associations after optimizing HPs. However, the AoI performances are also affected by the trajectory planning for the UAVs. Moreover, energy capacity of UAVs affects the trajectory planning. As a result, we would like to find the age-optimal trajectory for each UAV subject to its energy capacity E_{max}.

In particular, the optimization problem (11) is reduced to the following energy-constrained age-optimal trajectory planning problem for each activated UAV $a_k (k = 1, \ldots, K)$:

$$\min_{u_k} X_{k,ave}(u_k) = \frac{1}{|\widetilde{\mathcal{V}}_k|} \sum_{j=1}^{|\mathcal{C}^*_k|} \sum_{l=1}^{j} \phi^k_l \left(T^k_j + \frac{d^k_{j,(j+1)}}{V_k} \right) + T^k_o,$$
$$s.t. \begin{cases} e_k(u_k) \leqslant E_{max}, \\ u'_k \in \sigma(\mathcal{C}^*_k). \end{cases} \tag{17}$$

Let $u^*_{k,ave}$ denote the optimal solution to the problem (17), which corresponds to the age-optimized trajectory of UAV a_k.

1) CASE I

When the UAV's energy capacity is sufficiently large, we do not need to consider the energy constraint. As discussed in [21] and [23], the UAV's ave-AoI-optimal trajectory is a stage-weighted shortest Hamiltonian path, when no energy constraint is imposed. And the two age-optimal trajectories can be found for each UAV using dynamic programming (DP) or other appropriate methods.

The ave-AoI-optimal trajectories of UAV a_k without energy constraint are denoted by $u_{k,ave}$. Then, we briefly describe how to find $u_{k,ave}$ using the DP method. Let $g(c_l, \mathcal{S}_k)$ $(c_l \notin \mathcal{S}_k \subset \mathcal{C}_k^*)$ denote the minimum path cost when UAV a_k arrives at HP c_l, flies across all the HPs in the set $\mathcal{S}_k$ exactly once and goes back to the data center v_0. To find the ave-AoI-optimal trajectory, we calculate the minimum weighted path cost $g(c_l, \mathcal{S}_k)$ iteratively as:

$$g(c_l, \mathcal{S}_k) = \begin{cases} T_l^k + \tau_{l,0}^k + T_o^k, & \mathcal{S}_k = \emptyset, \\ \min\limits_{c_j \in \mathcal{S}_k} \{g(c_j, \mathcal{S}_k - \{c_j\}) + \\ (1 - \dfrac{1}{|\widetilde{\mathcal{V}}_k|} \sum\limits_{c_i \in \mathcal{S}_k} \phi_i^k)(T_l^k + \tau_{l,j}^k)\}, & \mathcal{S}_k \neq \emptyset, \end{cases} \tag{18}$$

where $\tau_{l,0}^k$ is the flight time from HP c_l to the data center, and $T_l^k + \tau_{l,j}^k$ is the time interval between the arrival at HP c_l and the arrival at HP $c_j \in \mathcal{S}_k$ for UAV a_k. Moreover, $\frac{1}{|\widetilde{\mathcal{V}}_k|} \sum_{c_i \in \mathcal{S}_k} \phi_i^k$ is the ratio of the number of SNs covered by the set of HPs $\mathcal{S}_k$ to the total number of SNs collected by UAV a_k. Accordingly, the average AoI of the SNs collected by UAV a_k is expressed as

$$X_{k,ave}(u_{k,ave}) = \min_{c_l \in \mathcal{C}_k} g(c_l, \mathcal{C}_k - \{c_l\}), \tag{19}$$

where the path cost function is given by (18). Accordingly, the ave-AoI-optimal trajectory $u_{k,ave}$ is obtained.

2) CASE II

From (10), the UAV's energy consumption $e_k(u_k)$ depends on the trajectory length when the flight velocity V_k is fixed. A longer trajectory usually leads to more energy consumption. As is known to us, the shortest trajectory of one UAV is the solution to the Traveling Salesman Problem (TSP), referred to the TSP trajectory denoted by $u_{k,tsp}$. The energy consumption spent on the longest TSP trajectory can be used as an energy threshold calculated as $e_{th} = \max_k e_k(u_{k,tsp})$. When the UAV's energy capacity E_{max} is greater than or equal to the threshold e_{th}, there exists an age-optimal flight trajectory for each UAV. Specifically, if the UAV's energy capacity satisfies $e_{th} \leq E_{max} \leq \max_k e_k(u_{k,max})$ or $e_{th} \leq E_{max} \leq \max_k e_k(u_{k,ave})$, we always find an energy-constrained age-optimal trajectory for each UAV using some intelligent search method like genetic algorithm [23].

Algorithm 3 The Greedy Search Based Multi-UAV-Assisted Data Collection Algorithm

Require: The network topology $G = (\mathcal{V}, \epsilon)$, the number of UAVs $K = 1$.

1: Find the SN-HP association via Algorithm 1.
2: Find the HP-UAV association via Algorithm 2.
3: Find the optimal Hamiltonian trajectory $u_{k,hp}$ using DP or other appropriate methods, and calculate the energy consumption $e_k(u_{k,hp})$;
4: Calculate the energy consumption of the UAV when flying along the TSP trajectory e_{th};
5: **if** $e_k(u_{k,hp}) \leq E_{max}$ **then**
6: $\quad u_{k,ave}^* = u_{k,ave}$;
7: **else if** $E_{max} \geq e_{th}$ **then**
8: $\quad$ Find a trajectory for each UAV by genetic algorithm;
9: **else if** $E_{max} < e_{th}$ **then**
10: $\quad$ No feasible solution exists and jump to line 2 with $K = K + 1$;
11: **end if**

Ensure: The SN-HP and HP-UAV associations, the trajectory of each UAV, and SNs' average AoI.

3) CASE III

When the UAV's energy capacity falls below the energy threshold, such as $E_{max} < e_{th}$, it becomes impossible to identify any feasible trajectory that complies with the energy constraint. In simpler terms, the optimal solution to Problem (17) ceases to exist.

As evident in Case III above, it is often unfeasible to find solutions when each UAV is energy-constrained, unless more UAVs are engaged in data collection. Consequently, we introduce a greedy search method to optimize the activation of UAVs. Specifically, a new UAV is activated in response to Case III, and this process iterates until either all the candidate UAVs are activated or a feasible solution to Problem (11) exists. In summary, the greedy search based multi-UAV-assisted data collection algorithm is described in Algorithm 3.

V. SIMULATION RESULTS

In this section, we present simulation results to demonstrate the performances of our proposed multi-UAV aided data collection scheme. Utilizing the widely employed software, Matlab 2012, we conducted comprehensive simulations. In a WSN, $N = 80$ SNs are randomly located in a bounded area of size 1000m×1000m, and the data center is located at the origin $(0, 0)$. The coverage radius r of each SN is defined as 40m unless otherwise specified. The noise-plus-interference power is set to be −100dbm. Similar to [21] and [26], the major simulation parameters are presented in Table 1.

TABLE 1. Major simulation parameters.

Notation	Parameter	Value
H	The constant flying altitude (m)	100
V_k	The flight velocity of each UAV (m/s)	40
ϱ	Path loss exponent	2.3
P_0	Blade profile power when hovering (Watt)	79.9
P_1	Induced power when hovering (Watt)	88.6
U_{tip}	The tip speed of the rotor blade (m/s)	120
ρ	The air density (kg/m^3)	1.225
V_0	Mean rotor induced velocity when hovering (m/s)	4.3
d	The fuselage drag ratio	0.6
s	Rotor solidity	0.05
A_r	The rotor disc area (m^2)	0.503
W_n	The size of data packet (bits)	20M
B	System bandwidth (Hz)	1M
P_u	Transmit power at each SN (Watt)	1

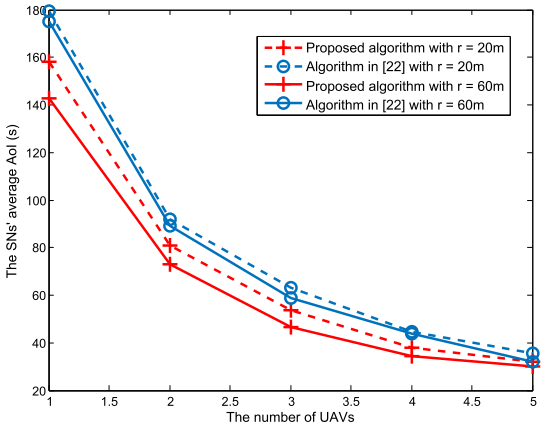
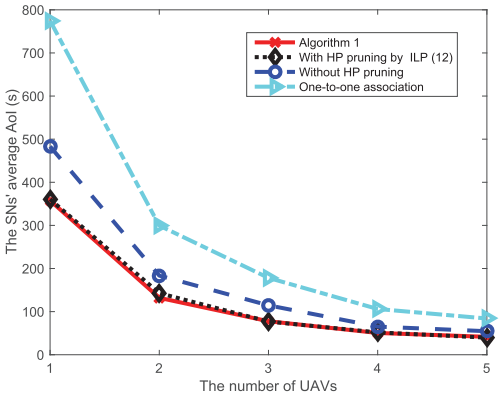
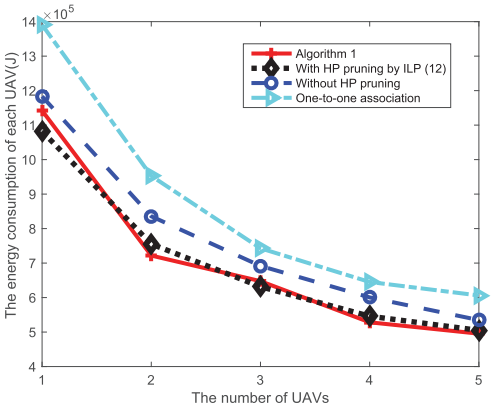


FIGURE 4. The AoI performance vs. the number of UAVs with different data collection strategies.



(a) The SNs' average AoI vs. the number of UAVs



(b) The energy consumption per UAV vs. the number of UAVs

FIGURE 5. The AoI and energy consumption performances vs. the number of UAVs with different SN-HP association methods.

A. PERFORMANCE COMPARISON WITH CONVENTIONAL STRATEGIES

To demonstrate the benefits of our proposed strategy, we compare our work with the state-of-art strategy proposed in [22]. In [22], the locations of the HPs and the SN-HP association are determined based on the affinity propagation based clustering algorithm, and the HP-UAV association is conducted through K-means clustering, then the trajectory planning is designed via an ant colony algorithm. From Fig. 4, the SNs' average AoI decreases as the number of UAVs grows, and our proposed algorithm outperforms the algorithm in [22]. This is because our proposed scheme can utilize the overlapped areas effectively with the help of graph theory and optimization approaches. Besides, when the SN's communication ranges are enlarged from 20 meters to 60 meters, the improvement of our proposed scheme is more obvious. This is because more overlapped areas can be utilized when the communication ranges are enlarged to reduce the number of HPs, thereby striking well balance between the SNs' data transmission time and the UAV's flight time.

B. THE IMPACT OF THE SN-HP ASSOCIATION

In Fig. 5, we plot the AoI and energy consumption curves for different SN-HP association schemes. In this experiment, we implement Algorithm 1 to find appropriate SN-HP associations, divided into two key parts: 1) Selection of candidate HPs; and 2) HP pruning. For comparison, we introduce the one-to-one association scheme which selects one single HP for each SN. To highlight the significance of HP pruning in the SN-HP association process, we further compare two HP pruning methods: the greedy HP pruning approach employed in Algorithm 1 and the exact HP pruning achieved by solving the ILP problem (11). This comparison is made against the approach that involves no HP pruning. Subsequently, the HP-UAV associations and ave-AoI-UAV trajectories are determined following the procedures outlined in Algorithm 3. As shown in this figure, the SNs' average AoI and the energy consumption per UAV decreases significantly when more UAVs are dispatched to collect data, since each UAV flies along a much shorter trajectory. Moreover, from Fig. 5(a) and Fig. 5(b), our proposed SN-HP association scheme performs the best, and the one-to-one association scheme achieves the worst performance in terms of the SNs' average AoI and energy consumption per UAV. Hence, it is important to select as few HPs as possible and establish the SN-HP association appropriately to save the flight time and energy for each UAV. Meanwhile, we can see that our proposed scheme achieves nearly the same AoI and energy consumption performance as

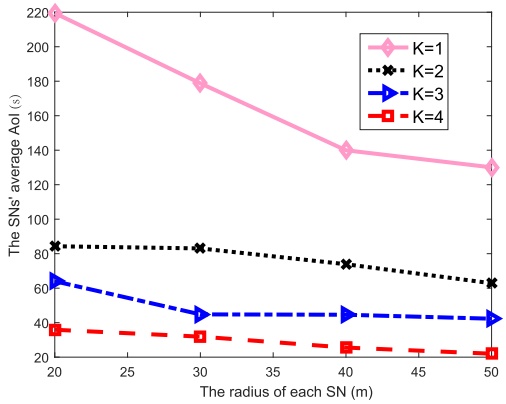


FIGURE 6. The AoI performance vs. the radius of each SN with different number of UAVs.

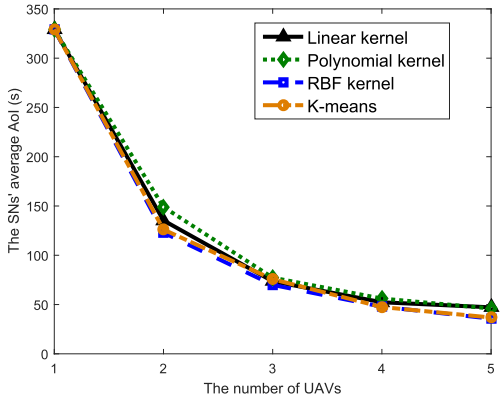


FIGURE 7. The AoIs performances vs. the number of UAVs with different clustering methods.

TABLE 2. Kernel functions.

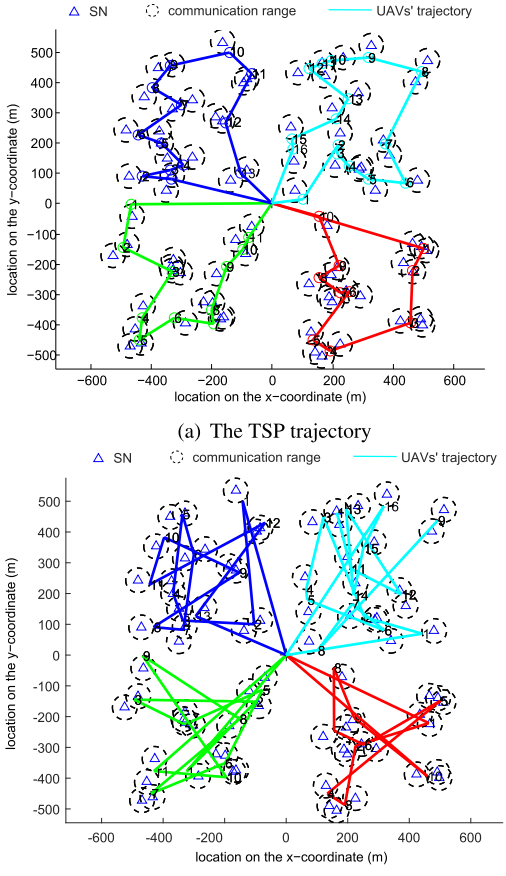
Kernel	Expression	Parameters
Linear	$\kappa(s_i, s_j) = s_i^T s_j + c$	$c = 0$
Polynomial	$\kappa(s_i, s_j) = \left(\lambda s_i^T s_j + c\right)^n$	$\lambda = \frac{1}{K}, c = 0, n = 3$
RBF	$\kappa(s_i, s_j) = \exp\left(-\frac{\|s_i - s_j\|^2}{2\Lambda^2}\right)$	$\Lambda = 11$

"with HP pruning by ILP (12)". The two SN-HP association schemes with HP pruning perform much better than the scheme without HP pruning. It means that HP pruning helps to shorten the flight trajectory of each UAV and is necessary to improve the system performance.

The SN-HP association is influenced by the coverage radius of each SN. A larger radius reduces the number of HPs required to cover all SNs. In Fig. 6, we illustrate the average AoI of SNs when our proposed scheme is applied with varying SN coverage radii. As each SN's radius expands, more SNs are associated with a single HP, resulting in reduced flight time for each UAV but increased transmission time for each SN. Consequently, as a weighted sum of SNs' transmission time and UAV flight time, the average AoIs of SNs gradually decrease with increased radii for any K UAVs, as depicted in Fig. 6. Additionally, we note that the communication range of each SN notably impacts AoI performance, particularly when only one UAV is utilized ($K = 1$). This discrepancy arises from the significant reduction in the UAV's flight trajectory length and flight time when fewer HPs are involved. This is due to the substantial reduction in the length of the flight trajectory (or flight time) of the UAV when fewer HPs are involved. In contrast, the reduction in trajectory length (or flight time) for each UAV is not as significant when multiple UAVs are utilized (i.e., $K > 1$).

C. THE IMPACT OF THE HP-UAV ASSOCIATION

In Fig. 7, we compare the kernel K-means method with the traditional K-means method, when different kernel functions are applied. In this experiment, Algorithm 2 is employed to determine HP-UAV associations utilizing kernel K-means. To facilitate comparison, we introduce three kernel functions:

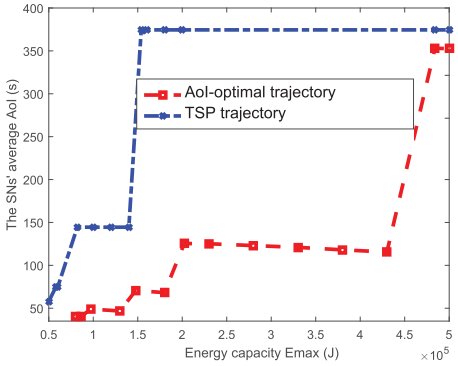


(a) The TSP trajectory

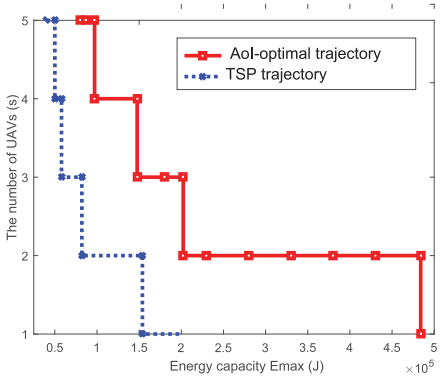
(b) The AoI-optimal trajectory

FIGURE 8. The UAV trajectories for N = 80 and K = 4 with different trajectory planning methods.

Linear, Polynomial, and Radial Basis Function (RBF), with their respective expressions presented in Table 2. Similar to Fig. 5, the SNs' average AoI decreases significantly when more UAVs are employed regardless of whether K-means or kernel K-means is applied. From Fig. 7, among the three kernel functions, the kernel K-means clustering method with RBF kernel achieves the best AoI performance for any K ($K > 1$) UAVs except the single-UAV scenario. When $K = 1$,



(a) The SNs' average AoI vs. the energy capacity



(b) The required number of UAVs vs. the energy capacity

FIGURE 9. The AoI performance and the required number of UAVs vs. energy capacity with different trajectory planning methods.

there is no need of conducting the HP-UAV association using the clustering method. One can see that our proposed scheme achieves a relatively higher average AoI when the polynomial and linear kernels are applied. And the performance gaps between the RBF kernel and polynomial and linear kernels become larger as the number of UAVs K increases. It is also interesting to see that the simple K-means method performs quite well, comparable to the kernel K-means method with RBF kernel.

D. ENERGY-CONSTRAINED TRAJECTORY PLANNING

In Fig. 8, we present the UAV trajectories with different trajectory planning methods. The term "AoI-optimal trajectory" refers to our proposed scheme that conducts trajectory planning using Algorithm 3, while "TSP trajectory" utilizes trajectory planning derived from the Traveling Salesman Problem solution. In this figure, the randomly distributed SNs are represented by the triangles and their associated HPs are marked by the small circles. Compared to the number of SNs, the number of HPs is significantly reduced by HP selection and pruning especially in a denser WSN. All the HPs are evenly classified into $K = 4$ clusters and each cluster of HPs is visited by one UAV. By trajectory planning, both the TSP trajectory and the ave-AoI-optimal trajectory are found for each UAV as plotted in Fig. 8 (a) and Fig.8 (b).

In Fig. 9, we present the system performance in relation to the energy capacity of each UAV. It is observed that the SNs' average AoIs increase monotonically with the increase of the energy capacity E_{max}, as shown in Fig. 9(a), respectively. Meanwhile, less UAVs are required in data collection as the energy capacity E_{max} is enlarged, as plotted in Fig. 9(b). This is due to the fact that each UAV carrying more energy is able to collect from more SNs distributed in the WSN. Hence, it spends more time on transmission and flight along a longer trajectory, which induces a higher average AoI. However, to keep information freshness, more UAVs should take part in the data collection task to achieve a smaller AoI. From Fig. 8 (a) and 8 (b), "TSP trajectory" has a shorter flight trajectory compared with "Ave-optimal trajectory". While from Fig. 9(a), "TSP trajectory" has a larger average AoI than the "Ave-optimal trajectory". This means that a shorter flight trajectory does not always result in an improved AoI performance. From Fig. 9(b), the shortest TSP trajectory consumes the least amount of energy.

VI. CONCLUSION

In this paper, we investigated the AoI-optimal data collection problem for multi-UAV-enabled WSNs, taking into account finite energy capacity of each UAV. To improve the AoI performance, we made efforts to optimize the HPs' locations, the SN-HP and HP-UAV associations, and AoI-optimal trajectory for each UAV so as to balance the data collection tasks between multiple UAVs. The efficient data collection problem is decomposed into two key components. Firstly, we employ graph theory to determine the locations of HPs and establish the SN-HP associations, and employ kernel K-means clustering to categorize all HPs into K clusters. Secondly, we optimize the trajectory for each UAV within a cluster of HPs to minimize the average AoI of SNs, all while adhering to energy constraints of UAVs. Simulation results confirmed the correctness of our proposed algorithm, and showed that with finite energy supply, our proposed multi-UAV-enabled data collection scheme can achieve significant AoI performance gains. While our adoption of the rotary-wing UAV model is simple, it provides practical means to evaluate SNs' average AoI and UAVs' energy consumption, facilitating the development of an efficient multi-UAV-aided fresh data collection scheme. Future investigations will delve into more sophisticated multi-UAV-assisted data collection schemes, accounting for more intricate UAV dynamics and characteristics.

REFERENCES

[1] Z. Wei, M. Zhu, N. Zhang, L. Wang, Y. Zou, Z. Meng, H. Wu, and Z. Feng, "UAV-assisted data collection for Internet of Things: A survey," *IEEE Internet Things J.*, vol. 9, no. 17, pp. 15460–15483, Sep. 2022.

[2] 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.

[3] C. H. Liu, Z. Chen, J. Tang, J. Xu, and C. Piao, "Energy-efficient UAV control for effective and fair communication coverage: A deep reinforcement learning approach," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 9, pp. 2059–2070, Sep. 2018.

[4] R. Ding, F. Gao, and X. S. Shen, "3D UAV trajectory design and frequency band allocation for energy-efficient and fair communication: A deep reinforcement learning approach," *IEEE Trans. Wireless Commun.*, vol. 19, no. 12, pp. 7796–7809, Dec. 2020.
[5] M. Pourghasemian, M. R. Abedi, S. S. Hosseini, N. Mokari, M. R. Javan, and E. A. Jorswieck, "AI-based mobility-aware energy efficient resource allocation and trajectory design for NFV enabled aerial networks," *IEEE Trans. Green Commun. Netw.*, vol. 7, no. 1, pp. 281–297, Mar. 2023.
[6] O. M. Bushnaq, A. Chaaban, and T. Y. Al-Naffouri, "The role of UAV-IoT networks in future wildfire detection," *IEEE Internet Things J.*, vol. 8, no. 23, pp. 16984–16999, Dec. 2021.
[7] M. Costa, M. Codreanu, and A. Ephremides, "On the age of information in status update systems with packet management," *IEEE Trans. Inf. Theory*, vol. 62, no. 4, pp. 1897–1910, Apr. 2016.
[8] C. Zhan and Y. Zeng, "Completion time minimization for multi-UAV-enabled data collection," *IEEE Trans. Wireless Commun.*, vol. 18, no. 10, pp. 4859–4872, Oct. 2019.
[9] 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.
[10] S. F. Abedin, M. S. Munir, N. H. Tran, Z. Han, and C. S. Hong, "Data freshness and energy-efficient UAV navigation optimization: A deep reinforcement learning approach," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 9, pp. 5994–6006, Sep. 2021.
[11] A. Ferdowsi, M. A. Abd-Elmagid, W. Saad, and H. S. Dhillon, "Neural combinatorial deep reinforcement learning for age-optimal joint trajectory and scheduling design in UAV-assisted networks," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 5, pp. 1250–1265, May 2021.
[12] V. Tripathi, R. Talak, and E. Modiano, "Age optimal information gathering and dissemination on graphs," in *Proc. IEEE INFOCOM Conf. Comput. Commun.*, Paris, France, Apr. 2019, pp. 2422–2430.
[13] Z. Jia, X. Qin, Z. Wang, and B. Liu, "Age-based path planning and data acquisition in UAV-assisted IoT networks," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, Shanghai, China, May 2019, pp. 1–6.
[14] H. Hu, K. Xiong, G. Qu, Q. Ni, P. Fan, and K. B. Letaief, "AoI-minimal trajectory planning and data collection in UAV-assisted wireless powered IoT networks," *IEEE Internet Things J.*, vol. 8, no. 2, pp. 1211–1223, Jan. 2021.
[15] A. Chapnevis and E. Bulut, "AoI-optimal cellular-connected UAV trajectory planning for IoT data collection," in *Proc. IEEE 48th Conf. Local Comput. Netw. (LCN)*, Daytona Beach, FL, USA, Oct. 2023, pp. 1–6.
[16] Y. Li, W. Liang, W. Xu, Z. Xu, X. Jia, Y. Xu, and H. Kan, "Data collection maximization in IoT-sensor networks via an energy-constrained UAV," *IEEE Trans. Mobile Comput.*, vol. 22, no. 1, pp. 159–174, Jan. 2023.
[17] M. A. Abd-Elmagid and H. S. Dhillon, "Average peak age-of-information minimization in UAV-assisted IoT networks," *IEEE Trans. Veh. Technol.*, vol. 68, no. 2, pp. 2003–2008, Feb. 2019.
[18] 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. Wireless Commun.*, vol. 19, no. 10, pp. 6578–6592, Oct. 2020.
[19] M. A. Abd-Elmagid, A. Ferdowsi, H. S. Dhillon, and W. Saad, "Deep reinforcement learning for minimizing Age-of-Information in UAV-assisted networks," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Waikoloa, HI, USA, Dec. 2019, pp. 1–6.
[20] M. Samir, C. Assi, S. Sharafeddine, D. Ebrahimi, and A. Ghrayeb, "Age of information aware trajectory planning of UAVs in intelligent transportation systems: A deep learning approach," *IEEE Trans. Veh. Technol.*, vol. 69, no. 11, pp. 12382–12395, Nov. 2020.
[21] C. Mao, J. Liu, and L. Xie, "Multi-uav aided data collection for age minimization in wireless sensor networks," in *Proc. Int. Conf. Wireless Commun. Signal Process. (WCSP)*, Nanjing, China, 2020, pp. 80–85.
[22] X. Gao, X. Zhu, and L. Zhai, "AoI-sensitive data collection in multi-UAV-assisted wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 22, no. 8, pp. 5185–5197, Aug. 2023.
[23] P. Tong, J. Liu, X. Wang, B. Bai, and H. Dai, "UAV-enabled age-optimal data collection in wireless sensor networks," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, May 2019, pp. 1–6.
[24] D. Kim, R. N. Uma, B. H. Abay, W. Wu, W. Wang, and A. O. Tokuta, "Minimum latency multiple data MULE trajectory planning in wireless sensor networks," *IEEE Trans. Mobile Comput.*, vol. 13, no. 4, pp. 838–851, Apr. 2014.
[25] Y. Sun, E. Uysal-Biyikoglu, R. D. Yates, C. E. Koksal, and N. B. Shroff, "Update or wait: How to keep your data fresh," *IEEE Trans. Inf. Theory*, vol. 63, no. 11, pp. 7492–7508, Nov. 2017.
[26] 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.
[27] D. B. West, *Introduction to Graph Theory*, 2nd ed. Upper Saddle River, NJ, USA: Prentice-Hall, 2001.
[28] I. S. Dhillon, Y. Guan, and B. Kulis, "Kernel k-means: Spectral clustering and normalized cuts," in *Proc. 10th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2004, pp. 551–556.

ZHENG ZHOU received the B.S. degree in information engineering and the Ph.D. degree in information and communication engineering from the Beijing University of Posts and Telecommunications (BUPT), Beijing, China, in 2012 and 2019, respectively.

Since January 2020, he has been with the College of Information Science and Engineering, Ningbo University, Ningbo, China, where he is currently a Lecturer. His research interests include simultaneous information and power transfer and cloud radio access networks and UAV communications.

JUAN LIU (Member, IEEE) received the B.S. degree in information and electronic engineering from Zhejiang University, Hangzhou, China, in 2000, the M.S. degree in information engineering from the Beijing University of Posts and Telecommunications, Beijing, China, in 2005, and the Ph.D. degree in electronic engineering from Tsinghua University, Beijing, in 2011.

From March 2012 to June 2014, she was with the ECE Department, NC State University (NCSU), Raleigh, NC, USA. From February 2015 to February 2016, she was with the ECE Department, The Hong Kong University of Science and Technology (HKUST), Hong Kong. Since March 2016, she has been with the College of Information Science and Engineering, Ningbo University, Ningbo, China, where she is currently a Professor. Her research interests include wireless communications and networking, UAV communications, and deep learning for wireless communications.

CHIXIONG MAO received the M.S. degree in electronics and communication engineering from Ningbo University, Ningbo, China, in 2021. His research interests include UAV communications and wireless communications.