2022 IEEE 19th International Conference on Mobile Ad Hoc and Smart Systems (MASS) | 978-1-6654-7180-0/22/$31.00 ©2022 IEEE | DOI: 10.1109/MASS56207.2022.00053

Data Collection in Wireless Sensor Networks: A Truck-Assisted Multi-UAV Method

Yuchao Zhu and Shaowei Wang
School of Electronic Science and Engineering, Nanjing University, Nanjing 210023, China
E-mail: dz20230030@smail.nju.edu.cn, wangsw@nju.edu.cn

***Abstract*—Unmanned aerial vehicles (UAVs) are endorsed as means to support efficient data collection from wireless sensor networks due to their flexibility and agility. However, the limited battery capacity of the UAVs becomes a bottleneck to the applications on many occasions, especially the networks deployed on the vast land. To this end, we consider a multi-UAV assisted data collection framework and propose a synchronized network to handle the energy issue, in which a truck carrying backup batteries moves together with the UAVs, acting as a "mobile charging station". Our goal is to minimize the total mission time for data collection from all the sensor nodes. We first develop a clustering algorithm to divide the region of interest into multiple clusters in a load-balanced way so as to minimize the number of movements of the UAVs, then formulate the trajectory optimization task as a coordinated multiple traveling salesman problem which is heuristically solved by a utilization-balancing trajectory planning algorithm. Numerical results indicate that our proposed scheme can effectively find the trajectories for the truck and UAVs. Moreover, our proposal sheds light on how to design a high-efficient and practical way for data gathering from large-scale wireless sensor networks.**

***Index Terms*—Data gathering, trajectory optimization, unmanned aerial vehicles, wireless sensor network.**

I. Introduction

Internet of Things (IoT), in which a large number of physical objects are connected to the Internet at an unprecedented rate, is expected to be an indispensable technology in the coming years [1]. As a major part of IoT, wireless sensor networks (WSNs) are widely deployed to collect data with only occasional human interventions, in which numerous small-size, low cost, low-power sensor nodes (SNs) are spatially distributed to detect physical phenomena in the ambient environment [2]. Efficient data collection consequently becomes a critical and challenging task in large-scale WSNs due to the explosive growth of the number of SNs. Traditionally, the WSN is organized as a routing network in a static manner, where the sensed data is forwarded and relayed via multihopping to the destination [2]. However, gathering data from massive sparsely or widely deployed SNs would inevitably cause overuse of the rendezvous nodes and unreliable wireless links, which consequently impairs the performance of WSN. Hence, ground mobile sinks are employed to alleviate these problems. By introducing one or more ground vehicles equipped with data sinks to gather data, the static routing network is partially or fully taken over to eliminate the non-uniformity of energy consumption [3]. Nevertheless, one of the common features of the WSN application scenarios is that the SNs are usually distributed on a vast land. The deployment of terrestrial Internet infrastructures is expensive as well as difficult in such remote areas, which poses an obstacle to static data collection. Moreover, some WSNs are deployed in harsh terrains, which increases the difficulty of data collection by ground vehicles.

This work was partially supported by the National Natural Science Foundation of China under Grants 61931023 and U1936202.

In recent years, unmanned aerial vehicle (UAV) has received much attention in diverse domains thanks to the fast-paced progress in design and production [4], [5]. Due to the great mobility and on-demand service ability, UAV as a flying base station can provide an effective approach for gathering data from WSNs. In [6], the time-varying UAV speed and the transmission interval are jointly optimized to minimize the flight time of the UAV in a one-dimensional WSN. In [7], a unified scheme for age-optimal data gathering is proposed, where a trajectory planning algorithm based on dynamic programming is developed to minimize the UAV's flight time. These works are limited to data collection with a single UAV, but in actual application scenarios, single-UAV based data collection system is hard to cope with the excessive number of SNs. In [8], the multi-UAV enabled data collection mission is formulated as a capacitated vehicle routing problem, where the sensed data have time deadlines. Assuming that the UAV can gather data in both flying and hovering mode, [9] jointly optimizes the allocation of SNs, the data collection mode and each UAV's trajectory to minimize the completion time while considering SNs' limited energy.

Nevertheless, the finite onboard battery capacity tends to be an intractable limitation. Recent works envisage that the UAV's battery can be recharged or replaced whenever it drains up during its flight, enabled by an automatic battery replacement system. Given one fixed charging station, an energy-efficiency oriented scenario is studied in [10], where multiple rechargeable UAVs collaborate to provide seamless and long-term coverage for ground SNs. A block descent-based iterative algorithm is presented to optimize the node assignment, UAVs' trajectory and transmission power, while only three SNs are considered. In [11], the authors investigate the optimal deployment of UAV charging stations, where the flying distance of UAV and the number of charging stations are jointly minimized. However, the method lacks flexibility since the placement of charging stations is fixed.

2155-6814/22/$31.00 ©2022 IEEE
DOI 10.1109/MASS56207.2022.00053

Authorized licensed use limited to: Nanjing University. Downloaded on February 13,2023 at 08:11:33 UTC from IEEE Xplore. Restrictions apply.

Motivated by the above observations, we attempt to endow the UAV charging station with mobility so as to provide energy support for the UAV. Different from prior works which focus on single UAV [12], [13], in this paper, we study a novel truck-assisted multi-UAV approach to realize high-efficient periodic data collection in large-scale WSNs, where UAVs acting as aerial mobile sinks are dispatched for data gathering and a truck carrying backup batteries is employed to serve as a mobile battery swap station. We elaborate that the optimization task to minimize the mission time for collecting data from all the SNs under the constraint of UAVs' flight endurance can be fulfilled by two stages, i.e., optimal selection of hovering positions and optimal trajectory planning. The former is solved by a clustering method, and the latter is formulated as an extension of the multiple traveling salesman problem (mTSP) and is solved by a simple and efficient utilization-balancing coordinated route planning algorithm. We thoroughly analyze the upper and lower bound of the problem to show the performance guarantee. Numerical results verify that our proposal is practical and high-efficient, which provides a novel and promising way for data gathering in real application scenarios.

II. System Model And Problem Formulation

A. Network description

Let us consider a WSN deployed on the ground over a region $\mathcal{R} \in \mathbb{R}^2$, consisting of N SNs denoted by $\mathcal{S} = \{s_1, s_2, ..., s_N\}$, one data center located at $\boldsymbol{c}_0$, one truck and K rotary-wing UAVs denoted by $\mathcal{K} = \{1, 2, ..., K\}$. The location of SN $s_n \in \mathcal{S}$ is assumed to be fixed, which is denoted by $\boldsymbol{z}_n = (s_n^x, s_n^y, 0)$. UAVs can operate in hovering to gather data from ground SNs, thus the entire region is supposed to be partitioned into M clusters as depicted in Fig. 1, so that the UAVs can collect data from SNs in each cluster through one taking-off and landing. The set of clusters and UAVs' hovering positions above corresponding clusters are denoted by $\mathcal{R}_c = \{R_1, R_2, ..., R_M\}$ and $\mathcal{H} = \{\boldsymbol{h}_1, \boldsymbol{h}_2, ..., \boldsymbol{h}_M\}$, respectively. When one UAV hovers at $\boldsymbol{h}_i = (h_i^x, h_i^y, h_i^z)$, it establishes communication links with the SNs lied in cluster R_i, where the coverage radius of the UAV is related to its altitude. A

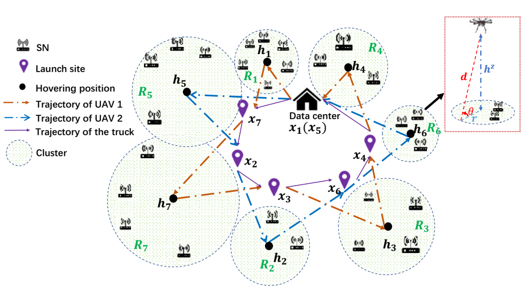


Fig. 1. Illustration of system model.

truck is dispatched to assist the UAVs since the UAVs have to fly back to the moving truck for battery replacement before visiting the next target cluster. The locations at which the UAV and truck meet each other are represented by "launch sites" $\mathcal{X} = \{\boldsymbol{x}_1, \boldsymbol{x}_2, ..., \boldsymbol{x}_M\}$, where $\boldsymbol{x}_i = (x_i^x, x_i^y, 0)$. The time required for collecting data from one SN is set as T_s. The number of clusters that allocated to UAV k is denoted by n_k and the corresponding indexes of clusters are collected in a set $\mathcal{M}_k$. Let $\mathcal{U}_k \ni \sigma_{k(\cdot)}$ denote the permutations of $\mathcal{M}_k$ representing all the possible sequences for the UAV to visit its assigned clusters, where we set $\sigma_k(n_k+1) = \sigma_k(1)$ to simplify the notation. Thus, the trajectory of UAV $k \in \mathcal{K}$ is represented by $[\boldsymbol{x}_{\sigma_k(1)}, \boldsymbol{h}_{\sigma_k(1)}, \boldsymbol{x}_{\sigma_k(2)}, ..., \boldsymbol{h}_{\sigma_k(n_k)}, \boldsymbol{x}_{\sigma_k(n_k+1)}]$, where the data center $\boldsymbol{c}_0$ is the start/finish point.

B. Air-to-ground channel model

The UAV's air-to-ground channel takes both line-of-sight (LoS) links and non-line-of-sight (NLoS) links into consideration. Following the model in [14], the path loss for the air-to-ground communication can be expressed as follows:

$$L_u(d_{m,n}) = (4\pi f_c d_{m,n}/c)^2 \beta_u, \ u \in \{LoS, NLoS\}, \quad (1)$$

where $d_{m,n} = \|\boldsymbol{z}_n - \boldsymbol{h}_m\|$ is the Euclidean distance between the SN and the UAV, c is the speed of light, f_c represents the carrier frequency, β_{LoS} and β_{NLoS} refer to the excessive path loss of LoS and NLoS links, respectively. The probability of LoS link is given by:

$$P_{LoS}(\theta_{m,n}) = \frac{1}{1 + a\exp\big(-b(\theta_{m,n} - a)\big)}, \quad (2)$$

where a, b are parameters related to the environment, $\theta_{m,n} = \frac{180^\circ}{\pi}\arcsin(h_m^z/d_{m,n})$ represents the elevation angle between the SN s_n and the UAV hovering at $\boldsymbol{h}_m$. Also, the probability of NLoS links is $P_{NLoS}(\theta) = 1 - P_{LoS}(\theta)$, so the average path loss model (measured in dB) can be expressed as follows [14]:

$$L_{pl}(d_{m,n}) = \sum_{u \in \{LoS, NLoS\}} L_u(d_{m,n}) P_u(\theta_{m,n}). \quad (3)$$

To guarantee the communication link between the UAV and SNs, a threshold Γ is introduced to represent the maximum allowable path loss corresponding to the minimum transmission rate requirement of data collection. We assume that the SNs could establish connection with the UAV if $L_{pl}(d_{m,n}) \leq \Gamma$. It is obvious that $L_{pl}(d_{m,n})$ from (3) can be rewritten as an implicit function of h_m^z and $r_{m,n}$ (the ground distance between the SN s_n and UAV at $\boldsymbol{h}_m$), so the UAV's maximal coverage radius can be mathematically expressed as $r_{max} = r_{m,n}|_{L_{pl}(r_{m,n}, h_m^z) = \Gamma}$. Note that the coverage radius rises first and then descends as the UAV altitude increases, h^z could be adjusted to achieve different coverage areas while satisfying the same path loss requirement. The maximal coverage radius and corresponding UAV altitude can be obtained by numerically solving $\partial r_{max}/\partial h^z = 0$, which gives a reference to the number of clusters that need to be divided.

C. Problem formulation

We focus on the time to complete the data collection from all the SNs since the mission is delay-tolerant. And we try to balance the number of clusters that each UAV serves so as to efficiently utilize the battery of all the UAVs.

Authorized licensed use limited to: Nanjing University. Downloaded on February 13,2023 at 08:11:33 UTC from IEEE Xplore. Restrictions apply.

Define $l = M \bmod K$, then the number of clusters that UAV k needs to visit can be written as:

$$n_k = \begin{cases} \lfloor \frac{M}{K} \rfloor, & l < k, \\ \lfloor \frac{M}{K} \rfloor + 1, & l \geq k, \end{cases} \forall k \in \mathcal{K}, \tag{4}$$

and the index of the UAV that visits the last cluster is represented by $\kappa = \begin{cases} K, & l = 0 \\ l, & \text{otherwise} \end{cases}$.

The objective is to minimize the total mission time, which is determined by the UAV that serves the last cluster, i.e., UAV κ. Denote $v_0, v_1 > 0$ as the speed of the truck and the UAV, respectively, with $v_0 < v_1$. Let τ^i_{truck} and τ^i_{UAV} denote the time required for the truck and UAV κ to finish the i-th piece of route, respectively. Denote $t_h^{\sigma_k(i)}$ as the hovering time of UAV k above the i-th cluster among the clusters assigned to it. Then the optimization problem can be written as follows:

$$\begin{aligned} &\underset{\sigma_k \in \mathcal{U}_k,\ \boldsymbol{x}}{\text{minimize}} \quad \sum_{i=1}^{n_\kappa} \max\{\tau^i_{truck}, \tau^i_{UAV}\} \\ &s.t. \quad E_f t_f^{\sigma_k(i)} + E_h t_h^{\sigma_k(i)} \leq E_{max},\ \forall 1 \leq k \leq K, 1 \leq i \leq n_k, \end{aligned} \tag{5}$$

where $t_f^{\sigma_k(i)} = \frac{1}{v_1}(||\boldsymbol{x}_{\sigma_k(i)} - \boldsymbol{h}_{\sigma_k(i)}|| + ||\boldsymbol{h}_{\sigma_k(i)} - \boldsymbol{x}_{\sigma_k(i+1)}||)$, E_f, E_h are the energy consumption of flying and hovering, respectively, and the total capacity of the battery is denoted as E_{max}. The first term in $\max\{\cdot, \cdot\}$ represents the time needed for the truck to move from one launch site to another to rendezvous with UAV κ, the second term $\tau^i_{UAV} = \frac{1}{v_1}(\|\boldsymbol{x}_{\sigma_\kappa(i)} - \boldsymbol{h}_{\sigma_\kappa(i)}\| + \|\boldsymbol{h}_{\sigma_\kappa(i)} - \boldsymbol{x}_{\sigma_\kappa(i+1)}\|) + t_h^{\sigma_\kappa(i)}$ corresponds to the amount of time for UAV κ to fly away from one launch site, reach its hovering position and collect data from SNs, then return to meet the truck at another launch site for battery replacement. The constraint guarantees that all the UAVs could fly back to the truck before their energy drains.

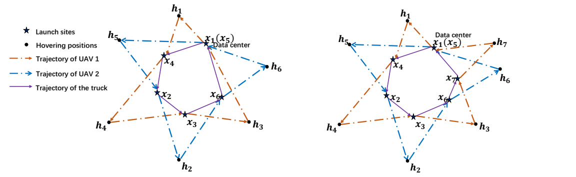
Then, we elaborate τ^i_{truck} in detail as given in **Theorem 1.**

Theorem 1. *Denote $\boldsymbol{x}_{(\cdot)}$ as the location of launch sites. As the accumulative mission time is determined by UAV κ, the entire path can be divided into n_κ pieces. The amount of time required for the truck to complete the i-th piece of path can be expressed as:*

$$\tau^i_{truck} = \begin{cases} \frac{\|\boldsymbol{x}_{\sigma_\kappa(1)} - \boldsymbol{x}_{\sigma_1(2)}\| + \sum_{a=2}^{\kappa}\|\boldsymbol{x}_{\sigma_a(2)} - \boldsymbol{x}_{\sigma_{a-1}(2)}\|}{v_0}, & i = 1, \\ \frac{1}{v_0}(\sum_{a=\kappa}^{K-1}\|\boldsymbol{x}_{\sigma_a(i)} - \boldsymbol{x}_{\sigma_{a+1}(i)}\| + \|\boldsymbol{x}_{\sigma_1(i+1)} - \boldsymbol{x}_{\sigma_K(i)}\| \\ + \sum_{b=1}^{\kappa-1}\|\boldsymbol{x}_{\sigma_b(i+1)} - \boldsymbol{x}_{\sigma_{b+1}(i+1)}\|), & 2 \leq i \leq n_\kappa - 1, \\ \frac{\|\boldsymbol{x}_{\sigma_\kappa(n_\kappa+1)} - \boldsymbol{x}_{\sigma_\kappa(n_\kappa)}\|}{v_0}, & i = n_\kappa. \end{cases} \tag{6}$$

Proof. According to the number of clusters assigned to each UAV, we discuss the time required for the truck to finish its i-th pieces of route in two conditions, as shown in Fig. 2.

Condition 1: $l = 0$. The completion time is determined by UAV K, and it is obvious that $n_1 = n_2 = ... = n_K = \lfloor \frac{M}{K} \rfloor$. As illustrated in Fig. 2(a), when $1 \leq i \leq n_K - 1$, the truck route is composed of K segments (e.g., $\overline{\boldsymbol{x}_5\boldsymbol{x}_4}$ and $\overline{\boldsymbol{x}_4\boldsymbol{x}_2}$ when



(a) Condition 1 (b) Condition 2

Fig. 2. Illustration of two conditions, an instance of two UAVs.

$i = 1$), and when $i = n_K$, the truck route consists of one segment (e.g., $\overline{\boldsymbol{x}_6\boldsymbol{x}_5}$). So we have

$$\begin{aligned} \tau^i_{truck} = \frac{1}{v_0}\Big(&\sum_{a=1}^{K-1}\|\boldsymbol{x}_{\sigma_a(i+1)} - \boldsymbol{x}_{\sigma_{a+1}(i+1)}\| \\ &+ \|\boldsymbol{x}_{\sigma_1(i+1)} - \boldsymbol{x}_{\sigma_K(i)}\|\Big),\ 1 \leq i \leq n_K - 1, \end{aligned} \tag{7}$$

and

$$\tau^{n_K}_{truck} = \frac{\|\boldsymbol{x}_{\sigma_K(n_K+1)} - \boldsymbol{x}_{\sigma_K(n_K)}\|}{v_0}. \tag{8}$$

Note that, if we set $i = n_K$ in (7), the first term in the brackets is equals to 0, since $x_{\sigma_a(n_K+1)} = x_{\sigma_a(1)}$ $(\forall a \in \{1, ..., K-1\})$ is the location of the data center. So it can be derived that:

$$\begin{aligned} \tau^i_{truck} = \frac{1}{v_0}\Big(&\sum_{a=1}^{K-1}\|\boldsymbol{x}_{\sigma_a(i+1)} - \boldsymbol{x}_{\sigma_{a+1}(i+1)}\| \\ &+ \|\boldsymbol{x}_{\sigma_1(i+1)} - \boldsymbol{x}_{\sigma_K(i)}\|\Big),\ 1 \leq i \leq n_K. \end{aligned} \tag{9}$$

Condition 2: $l \neq 0$. The completion time is determined by UAV l, and we have $n_l = \lfloor \frac{M}{K} \rfloor + 1$. As illustrated in Fig. 2(b), when $i = 1$, the truck route is consisted of l segments, when $1 \leq i \leq n_l - 1$, the truck route is composed of K segments, and when $i = n_l$, the truck route consists of one segment:

$$\tau^1_{truck} = \frac{\|\boldsymbol{x}_{\sigma_l(1)} - \boldsymbol{x}_{\sigma_1(2)}\| + \sum_{a=2}^{l}\|\boldsymbol{x}_{\sigma_a(2)} - \boldsymbol{x}_{\sigma_{a-1}(2)}\|}{v_0}, \tag{10}$$

$$\begin{aligned} \tau^i_{truck} = \frac{1}{v_0}\Big(&\sum_{a=l}^{K-1}\|\boldsymbol{x}_{\sigma_a(i)} - \boldsymbol{x}_{\sigma_{a+1}(i)}\| + \|\boldsymbol{x}_{\sigma_1(i+1)} - \boldsymbol{x}_{\sigma_K(i)}\| \\ &+ \sum_{b=1}^{l-1}\|\boldsymbol{x}_{\sigma_b(i+1)} - \boldsymbol{x}_{\sigma_{b+1}(i+1)}\|\Big), 2 \leq i \leq \left\lfloor \frac{M}{K} \right\rfloor, \end{aligned} \tag{11}$$

$$\tau^{n_l}_{truck} = \frac{\|\boldsymbol{x}_{\sigma_l(n_l+1)} - \boldsymbol{x}_{\sigma_l(n_l)}\|}{v_0}. \tag{12}$$

Then, it can be seen that **Condition 1** is a special case of. We can derive (9) when setting $l = K$ in (10)-(12).

Thus, define $\kappa = \begin{cases} K, & l = 0 \\ l, & \text{otherwise} \end{cases}$, the amount of time required for the truck to finish its i-th pieces of route can be written as (6). □

The optimization task (5) is NP-hard since it is a generalization of the mTSP that requires taking the locations of launch sites as well as the battery life of the UAV into consideration.

Authorized licensed use limited to: Nanjing University. Downloaded on February 13,2023 at 08:11:33 UTC from IEEE Xplore. Restrictions apply.

III. Our proposed scheme

To address the optimization task (5), we need to determine the hovering positions $\mathcal{H}$ at first. After selecting the target hovering positions, we develop a utilization-balancing trajectory planning strategy to heuristically solve the NP-hard optimization problem. We also give the analysis of the upper bound and lower bound to indicate the performance guarantee.

A. Selection of hovering positions

First, we estimate the number of clusters to be partitioned. Let s_{max} denote the maximum number of SNs that a UAV can serve considering battery capacity and A_R denote the area of region $\mathcal{R}$. The approximate number of divided clusters can be expressed as $N_a = \max\left\{\left\lceil \frac{N}{s_{max}} \right\rceil, \left\lceil \frac{A_R}{\pi r_{max}^2} \right\rceil\right\}$. The first term and second term in $\max\{\cdot,\cdot\}$ represent the capacity requirement and coverage constraint, respectively. We try to partition the region of interest into at least N_a clusters in a balanced way since the number of clusters would be minimal when the SNs are equally assigned between the clusters. We determine $\mathcal{H}$ based on a clustering method called ISODATA [15]. The flowchart of the clustering algorithm is shown in Fig. 3. The parameters that need to be set are denoted by

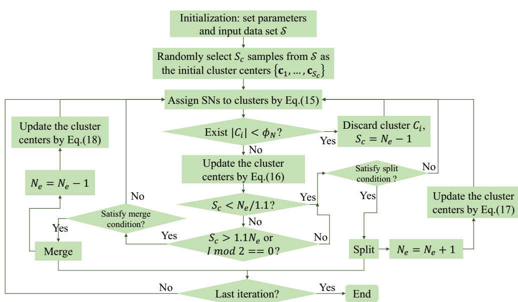


Fig. 3. The flowchart of ISODATA.

I, N_e, ϕ_N, ϕ_d, ϕ_s, which represent the maximum number of iterations, the desired number of clusters, the minimum size of clusters, the minimum distance between two cluster centroids, the allowed standard deviation, respectively.

First, randomly select S_c samples from the input data set to initialize the centers, the coordinates of which are denoted as $\boldsymbol{c}_i$, and the corresponding clusters are denoted by C_i. Then each SN would be assigned to the closest cluster center:

$$C_i = \{s \in \mathcal{S} |\ \|\boldsymbol{z} - \boldsymbol{c}_i\| \le \|\boldsymbol{z} - \boldsymbol{c}_j\|,\ \forall i \neq j\}. \tag{13}$$

And if the cluster C_i possesses SNs fewer than the given threshold ϕ_N, a discarding procedure would be triggered along with a reduction of cluster number. We explain the split and merge operations in detail. In a split procedure, we pick up the maximum cluster standard deviation denoted as δ_{max}. If $\delta_{max} > \phi_s$ and the cluster size is not smaller than $1.1\phi_N$, the split operation is triggered. In a merge procedure, we calculate the distance between all the cluster centers and denote them as matrix D_{dis}, where $D_{dis}(i,i) = 0$. If $D_{dis}(i,j) < \phi_d$, the merge operation is triggered. The cluster centers' updating rules are as follows:

$$\boldsymbol{c}_i = \frac{1}{|C_i|} \sum_{s \in C_i} \boldsymbol{z}, \tag{14}$$

$$\boldsymbol{c}_i^+ = \boldsymbol{c}_i + \delta_{max},\ \ \boldsymbol{c}_i^- = \boldsymbol{c}_i - \delta_{max}, \tag{15}$$

$$\boldsymbol{c}_{new} = \frac{1}{|C_i| + |C_j|}(|C_i|\boldsymbol{c}_i + |C_j|\boldsymbol{c}_j), \tag{16}$$

which correspond to the normal operation, split operation and merge operation, respectively.

We use SN set $\mathcal{S}$ as the input and the desired number of clusters is set as N_a, the output centroids are considered as the horizontal part of the hovering positions $\mathcal{H}$. Thus, the required coverage radius of the UAV above cluster C_m is $r_m = \max_{s_n \in C_m} \{r_{m,n} | r_{m,n} = \|\boldsymbol{c}_m - \boldsymbol{z}_n\|\}$, which is determined by the SNs on the edge. The corresponding altitude h_m^z of the UAV can be calculated by $L_{pl}(h_m^z, r_m) = \Gamma$.

B. Trajectory planning

To jointly plan the trajectories of multiple UAVs and the truck when given target hovering positions, the key points of which lie in allocating target clusters to UAVs and finding the optimal locations of the launch sites. We propose an efficient trajectory planning algorithm to address the optimization task (5) based on the fact that (5) over variables $\boldsymbol{x}_i$ will be convex under fixed permutations σ_k. The algorithm is given in Algorithm 1, which can be summarized as follows:

Algorithm 1 Utilization-balancing coordinated routing

Input: Region of interest $\mathcal{R}$, hovering positions $\mathcal{H}$, number of UAVs K, parameters of truck and UAV: $v_0, v_1, E_f, E_h, E_{max}, T_s$.
Step 1: Using LKH algorithm to solve the TSP tour of $\mathcal{H}$ to initialize the visiting sequence;
Return the solution σ;
Step 2: Assign $\mathcal{H}$ to the UAVs:
for $k = 1 : K$ **do**
  Calculate n_k by (4);
  for $i = 1 : n_k$ **do**
    $\sigma_k(i) \leftarrow \sigma(k + (i-1)K)$;
  end for
end for
Return the permutations $\sigma_1, ..., \sigma_K$;
Step 3: Using CVX to solve (17) with σ_k $(1 \le k \le K)$ to find the optimal trajectories;
Output: $\sigma_1, ..., \sigma_K$, $\boldsymbol{x}$, t_M.

Step 1: Initialize the order of visiting all the targets $\mathcal{H}$.

We set the initial sequence of $\{\boldsymbol{h}_1, ..., \boldsymbol{h}_M\}$ the same as the optimal TSP tour. The methods to solve the TSP are diverse, Lin-Kernighan heuristic (LKH), first proposed in [16], is one of the best methods for solving TSP. Here, we use an LKH solver constructed in [17] to find a promising TSP solution and initialize the visiting sequence σ.

Step 2: Assign the target hovering positions to each UAV.

Authorized licensed use limited to: Nanjing University. Downloaded on February 13,2023 at 08:11:33 UTC from IEEE Xplore. Restrictions apply.

Given the visiting sequence, the hovering positions are ordered as $\boldsymbol{h}_{\sigma(1)}, \boldsymbol{h}_{\sigma(2)}, ..., \boldsymbol{h}_{\sigma(M)}$. Since we try to balance the utilization of all the UAVs, the list of all the target hovering positions assigned to the UAV k are denoted as $\{\boldsymbol{h}_{\sigma(k)}, \boldsymbol{h}_{\sigma(k+K)}, ...\boldsymbol{h}_{\sigma(k+(n_k-1)K)}\}$. That is, $\sigma_k(i) = \sigma(k + (i-1)K)$. Each UAV flies to its target and returns to the truck, at which point its new target becomes the next point that is not currently a target of any of the UAVs.

Step 3: Solve problem (5) with fixed σ_k $(\forall k \in \mathcal{K})$.

Under fixed permutations $\sigma_{k(i)}$ $(k \in \mathcal{K})$, denote $m_{\sigma_k(\cdot)}$ as the number of SNs located in cluster $R_{(\cdot)}$, and the optimization problem can be rewritten as:

$$\begin{aligned}
&\underset{\boldsymbol{x}_1,...,\boldsymbol{x}_M,t_1,...,t_M}{\text{minimize}} \quad t_M \\
s.t.\ &C_1: \ t_{k+1} = t_1 + t_f^{\sigma_k(1)} + T_s m_{\sigma_{k(1)}}, \ \forall k \in \mathcal{K}, \\
&C_2: \ t_q \geq t_p + t_f^{\sigma_k(j+1)} + T_s m_{\sigma_{k(j+1)}}, \\
&\quad \{(p,q) | p = (j-1)K + k + 1, \ q = jK + k + 1, \\
&\quad \ j = 1, ..., n_k - 1, \ \forall k \in \mathcal{K}\}, \\
&C_3: \ t_{i+1} \geq t_i + \frac{\|\boldsymbol{x}_i - \boldsymbol{x}_{i+1}\|}{v_0}, \forall i = 1, ..., M-1, \\
&C_5: \ t_f^{\sigma_k(j)} \leq \frac{E_{max} - E_h T_s m_{\sigma_{k(j)}}}{E_f}, j = 1, ..., n_k, \forall k \in \mathcal{K}, \\
&C_6: \ t_1 = 0, \\
&C_7: \ \boldsymbol{x}_{\sigma_k(1)} = \boldsymbol{x}_{\sigma_k(n_k+1)}, \ \forall k \in \mathcal{K},
\end{aligned} \tag{17}$$

where $t_f^{\sigma_k(j)} = \frac{1}{v_1}\big(\|\boldsymbol{x}_{\sigma_k(j)} - \boldsymbol{h}_{\sigma_k(j)}\| + \|\boldsymbol{h}_{\sigma_k(j)} - \boldsymbol{x}_{\sigma_k(j+1)}\|\big)$, $t_{(\cdot)}$ represents the accumulative time at each ordered launch site. $\boldsymbol{x}_{\sigma_k(1)} = \boldsymbol{x}_{\sigma_k(n_k+1)}$ is the data center with fixed location. We can solve this convex optimization problem by using standard techniques such as CVX, then the optimal coordinated routes under fixed ordered visiting assignments can be found.

C. Analysis of bounds

Note that weakening the constraint would expand the solution space, which will not influence the bounds of the problem. So we set the constraint of the problem aside to obtain crude lower bound and upper bound for proof convenience. We prove the bounds based on the following lemma [18], which describes the relationship of the average distance between a point sampled from $g(x)$ and a loop $\mathcal{L}$. For notation convenience, let $d(x,\mathcal{L}) = \min_{x' \in \mathcal{L}} \|x - x'\|$ denotes the distance between a point x and the loop $\mathcal{L}$.

Lemma 1. *Denote $\mathcal{D}$ as a compact planar region and let $g(\cdot)$ be an absolutely continuous probability density function defined on $\mathcal{D}$. Let OPT(L) denoted the optimal objective value to the problem*

$$\begin{aligned}
&\underset{\mathcal{L}}{\text{minimize}} \quad \iint_{\mathcal{D}} g(x) d(x,\mathcal{L}) dx \\
&s.t. \quad len(\mathcal{L}) = L,
\end{aligned} \tag{18}$$

where the optimization variable $\mathcal{L}$ is taken over the set of all loops in $\mathcal{D}$ whose length $len(\mathcal{L})$ is well defined. Then

$$OPT(L) \sim \frac{1}{4L}\Big(\iint_{\mathcal{D}} \sqrt{g(x)} dx\Big)^2 \tag{19}$$

as $L \to \infty$.

Denote $\boldsymbol{h}^p_{\sigma(i)}$ as the projection of $\boldsymbol{h}_{\sigma(i)}$ on the 2D region plane $\mathcal{R}$. Let M_κ be the total number of SNs that UAV κ serves, then we have $\sum_{i=1}^{n_\kappa} t_h^{\sigma_\kappa(i)} = M_\kappa T_s$.

1) Lower bound: Refer to Fig. 2, it can be observed that for each $\boldsymbol{h}_{\sigma_k(i)}$, we can always find a $\boldsymbol{x}'_{\sigma_k(i)}$ that closest to $\boldsymbol{h}^p_{\sigma_k(i)}$ on the line segment $\overline{\boldsymbol{x}_{\sigma_k(i)}\boldsymbol{x}_{\sigma_k(i+1)}}$ of the truck route. Thus, according to the geometric relation, it is straightforward that the formula below always holds:

$$\|\boldsymbol{x}_{\sigma_k(i)} - \boldsymbol{h}^p_{\sigma_k(i)}\| + \|\boldsymbol{h}^p_{\sigma_k(i)} - \boldsymbol{x}_{\sigma_k(i+1)}\| \geq 2\|\boldsymbol{x}'_{\sigma_k(i)} - \boldsymbol{h}^p_{\sigma_k(i)}\|. \tag{20}$$

Then the objective function of (5) satisfies:

$$\begin{aligned}
Obj &\geq \max\Big\{\sum_{i=1}^{n_\kappa} \tau^i_{truck}, \sum_{i=1}^{n_\kappa} \tau^i_{UAV}\Big\} \\
&\geq \max\Big\{\sum_{i=1}^{n_\kappa} \tau^i_{truck}, \sum_{i=1}^{n_\kappa} \frac{1}{v_1}(\|\boldsymbol{x}_{\sigma_\kappa(i)} - h^p_{\sigma_\kappa(i)}\| \\
&\quad + \|\boldsymbol{h}^p_{\sigma_\kappa(i)} - \boldsymbol{x}_{\sigma_\kappa(i+1)}\|) + M_\kappa T_s\Big\} \\
&\geq \max\Big\{\sum_{i=1}^{n_\kappa} \tau^i_{truck}, \sum_{i=1}^{n_\kappa} \frac{2}{v_1}\|\boldsymbol{x}'_{\sigma_\kappa(i)} - \boldsymbol{h}^p_{\sigma_\kappa(i)}\| + M_\kappa T_s\Big\} \\
&= \max\Big\{\frac{1}{v_0} len(\mathcal{L}), \sum_{i=1}^{n_\kappa} \frac{2}{v_1} d(\boldsymbol{h}^p_{\sigma_\kappa(i)}, \mathcal{L}) + M_\kappa T_s\Big\},
\end{aligned} \tag{21}$$

where $\mathcal{L}$ here is regarded as the moving route of truck.

Then the lower bound of problem (5) can be expressed by the optimal objective value of problem (22):

$$\underset{\mathcal{L}}{\text{minimize}} \ \max\Big\{\frac{len(\mathcal{L})}{v_0}, \sum_{i=1}^{n_\kappa} \frac{2}{v_1} d(\boldsymbol{h}^p_{\sigma_\kappa(i)}, \mathcal{L}) + M_\kappa T_s\Big\}. \tag{22}$$

Eq. (22) is still a combinatorial optimization problem, the optimal value of which is hard to compute. As is standard in the continuous approximation paradigm, it is assumed that the projection of hovering positions $\boldsymbol{h}^p_{(\cdot)}$ are independent samples from an absolutely continuous probability density function $g(\cdot)$ defined on $\mathcal{R}$. The summation in (22) can be written as an integral over $\mathcal{R}$, which is given as follows:

$$\underset{\mathcal{L}}{\text{minimize}} \ \max\Big\{\frac{len(\mathcal{L})}{v_0}, \frac{2n_\kappa}{v_1} \iint_{\mathcal{R}} g(x) d(x,\mathcal{L}) dx + M_\kappa T_s\Big\}. \tag{23}$$

According to (19) of **Lemma 1.**, problem (23) can be transformed to:

$$\underset{L \geq 0}{\text{minimize}} \ \max\Big\{\frac{L}{v_0}, \frac{n_\kappa}{2v_1 L}\Big(\iint_{\mathcal{R}} \sqrt{g(x)} dx\Big)^2 + M_\kappa T_s\Big\}. \tag{24}$$

Obviously, the solution is $L^* = \frac{M_\kappa T_s v_0 + \sqrt{(M_\kappa T_s v_0)^2 + 2n_\kappa v_0/v_1 (\iint_{\mathcal{R}} \sqrt{g(x)} dx)^2}}{2}$, and the optimal objective value can be calculated by:

$$opt(n_\kappa) = \frac{M_\kappa T_s + \sqrt{(M_\kappa T_s)^2 + \frac{2n_\kappa}{v_0 v_1}(\iint_{\mathcal{R}} \sqrt{g(x)} dx)^2}}{2}. \tag{25}$$

Authorized licensed use limited to: Nanjing University. Downloaded on February 13,2023 at 08:11:33 UTC from IEEE Xplore. Restrictions apply.

2) Upper bound: It would take the most time to complete the mission when the truck keeps stationary whenever the UAVs fly away for data collection. So we can obtain an upper bound by replacing the $\max\{\cdot,\cdot\}$ in (5) with a summation:

$$\begin{aligned}
Obj &\leq \tau_{truck}^{1} + \frac{2}{v_1}\|\boldsymbol{h}_{\sigma_\kappa(1)} - \boldsymbol{x}'_{\sigma_\kappa(1)}\| + t_h^{\sigma_\kappa(1)} \\
&+ \sum_{i=2}^{n_\kappa-1}\left(\tau_{truck}^{i} + \frac{2}{v_1}\|\boldsymbol{h}_{\sigma_\kappa(i)} - \boldsymbol{x}'_{\sigma_\kappa(i)}\| + t_h^{\sigma_\kappa(i)}\right) \\
&+ \tau_{truck}^{n_\kappa} + \frac{2}{v_1}\|\boldsymbol{h}_{\sigma_\kappa(n_\kappa)} - \boldsymbol{x}'_{\sigma_\kappa(n_\kappa)}\| + t_h^{\sigma_\kappa(n_\kappa)} \\
&\leq \sum_{i=1}^{n_\kappa}\tau_{truck}^{i} + \frac{2}{v_1}\sum_{i=1}^{n_\kappa}(\|\boldsymbol{h}^p_{\sigma_\kappa(i)} - \boldsymbol{x}'_{\sigma_\kappa(i)}\| + H_{\sigma_\kappa(i)}) + M_\kappa T_s \\
&= \frac{1}{v_0}len(\mathcal{L}) + \frac{2}{v_1}\sum_{i=1}^{n_\kappa} d(\boldsymbol{h}^p_{\sigma_\kappa(i)}, \mathcal{L}) + \frac{2}{v_1}H_\kappa + M_\kappa T_s, \quad (26)
\end{aligned}$$

where $H_\kappa = \sum_{i=1}^{n_\kappa} H_{\sigma_\kappa(i)}$ denotes the sum of flight altitudes of the UAV κ.

Similar to the lower bound, the optimal objective value of problem (27) is the upper bound of problem (5):

$$\underset{\mathcal{L}}{\text{minimize}}\ \frac{len(\mathcal{L})}{v_0} + \frac{2}{v_1}\sum_{i=1}^{n_\kappa} d(\boldsymbol{h}^p_{\sigma_\kappa(i)}, \mathcal{L}) + \frac{2H_\kappa}{v_1} + M_\kappa T_s, \quad (27)$$

whose continuous approximation can be written as follows:

$$\underset{\mathcal{L}}{\text{minimize}}\ \frac{len(\mathcal{L})}{v_0} + \frac{2n_\kappa \iint_{\mathcal{R}} g(x) d(x,\mathcal{L}) dx}{v_1} + \frac{2H_\kappa}{v_1} + M_\kappa T_s. \quad (28)$$

By applying **Lemma 1.** again, (28) is equivalent to:

$$\underset{L\geq 0}{\text{minimize}}\ \max\left\{\frac{L}{v_0} + \frac{n_\kappa\left(\iint_{\mathcal{R}}\sqrt{g(x)}dx\right)^2}{2v_1 L} + \frac{2H_\kappa}{v_1} + M_\kappa T_s\right\}, \quad (29)$$

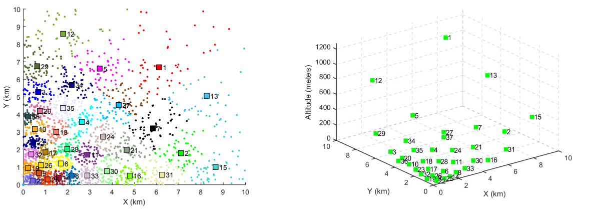
whose optimal solution is $L^* = \sqrt{\frac{n_\kappa v_0}{2v_1}}\iint_{\mathcal{R}}\sqrt{g(x)}dx$, and the corresponding objective value is:

$$opt(n_\kappa) = \sqrt{\frac{2n_\kappa}{v_0 v_1}}\iint_{\mathcal{R}}\sqrt{g(x)}dx + \frac{2H_\kappa}{v_1} + M_\kappa T_s. \quad (30)$$

IV. Numerical Results

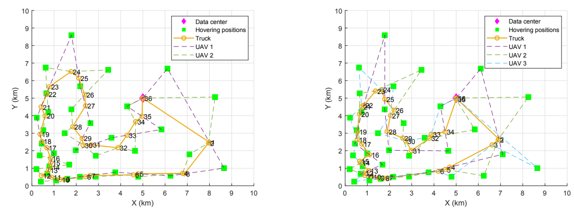
We consider a geographical region in suburban of size $10 \times 10\text{km}^2$, where 2000 SNs are deployed and the data center is located at the center of the area. The parameters in suburban environment with carrier frequency $f_c = 2$ GHz are $a = 4.88$, $b = 0.43$, $\eta_{LoS} = 0.1$ and $\eta_{NLoS} = 20$, which are calculated based on [14]. The path loss threshold is set as $\Gamma = 111$ dB. The battery parameters of the UAV are $E_f = 80$ W, $E_h = 160$ W, $E_{max} = 40$ Wh, respectively. The time required to finish data transmission with one SN is $T_s = 10$ seconds. Unless otherwise specified, the velocity of the UAVs and truck are $v_1 = 80$ km/h and $v_0 = 20$ km/h, respectively.

For preliminary preparation, we first partition the region of interest into multiple clusters with approximately equal load. Applying (3), we can obtain that the maximum achievable coverage radius of UAV is 3864 meters, then the estimated number of clusters is initialized as $n = 34$, where s_{max} is set as 60 to reserve a margin. Thus, for the clustering method, parameters are set as $I = 100$, $N_e = 34$, $\phi_N = 30$, $\phi_d = 1$, $\phi_s = 1$, respectively. Fig. 4 shows the hovering position selection result of our proposed scheme, where SNs are non-uniformly deployed. The corresponding number of clusters is 37. Then, the altitude of the UAV is adjusted to cover all the SNs in each cluster. The 3D hovering positions of Fig. 4(a) are given in Fig. 4(b), which indicate that the UAV will decrease its hovering altitude when the density of SNs increases. Finally, the trajectory design results under different numbers of UAVs are illustrated in Fig. 5. It can be seen that the truck tends to take a shorter route as the number of UAVs increases since the speed of UAVs is faster than the truck's and they could catch up with the moving truck.



(a) (b)

Fig. 4. (a) shows the clustering result, where the dots and squares in different colors represent SNs and the cluster centers, respectively. (b) shows the corresponding altitudes of the UAV.



(a) (b)

Fig. 5. The coordinated routing results under different number of UAVs.

Then, we compare our proposal with an intuitive greedy-based trajectory planning method in which the UAVs choose the nearest hovering positions as their next target cluster. Here we take uniformly distributed SNs as an example for the convenience of calculating upper and lower bounds so as to validate the performance guarantee of our proposal. Fig. 6 shows the total mission time as a function of the number of UAVs, where the lower bound and upper bound are calculated by (25) and (30), respectively. As expected, the mission time decreases as the number of UAVs increases. About 45% and 61% reduction of mission time can be achieved when increasing K from one to two and from one to three, respectively. The performance of our proposal is generally better than the greedy-based method and is quite close to the lower bound. Fig. 6 also indicated that there is no need to blindly increase the number of UAVs since the gain obtained by increasing the number of UAVs becomes slight when the number of UAVs reaches a particular value (e.g., $K = 5$).

Authorized licensed use limited to: Nanjing University. Downloaded on February 13,2023 at 08:11:33 UTC from IEEE Xplore. Restrictions apply.

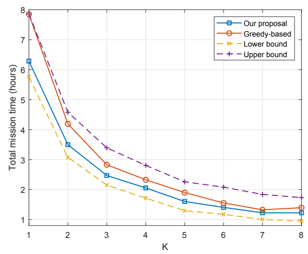


Fig. 6. Total mission time as a function of number of UAVs.

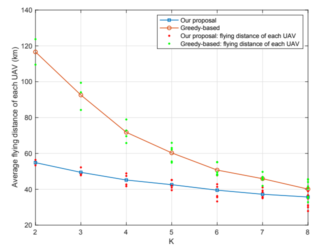


Fig. 7. Average flying distance of one UAV as a function of number of UAVs.

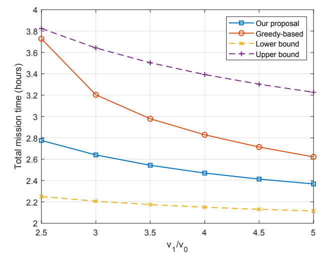


Fig. 8. Total mission time as a function of $\frac{v_1}{v_0}$, with $v_0 = 20$ km/h, $K = 3$.

In Fig. 7, we show the average flying distance of each UAV as a function of the number of UAVs. As can be seen, the required flying distance for each UAV is approximately equal by our proposal, which validates the utilization balancing of our scheme. And the flying distance of UAVs in our proposal is much shorter than that of greedy-based method, especially when K is relatively small. This is because the trajectory in our proposal generally follows a TSP route, while the truck and UAVs in greedy-based method always move back and forth. In Fig. 8, we show the total mission time as a function of $\frac{v_1}{v_0}$ when the number of UAVs is 3. On the whole, the total mission time is inversely proportional to $\frac{v_1}{v_0}$. Our proposal outperforms the greedy-based method especially when the speed of the UAV and truck is close, the gap is larger than 25% when $\frac{v_1}{v_0} = 2.5$. Fig. 8 also indicates that the proper configuration of v_0 and v_1 is important in practice, and it is unnecessary to increase the speed of the truck and the UAV rashly.

V. Conclusion

In this paper, we investigated the multi-UAV based data collection mission in large-scale WSNs deployed in the wild, which aimed to gather data in the shortest possible time considering the data transmission and battery capacity of the UAV. A truck carrying backup batteries was dispatched to aid the UAVs so that each UAV could return to the truck for battery replacement whenever its battery drains. Specifically, the proposed novel collaborative trajectory planning scheme can be summarized as two stages. First, UAV's hovering positions are determined based on a load-balanced clustering method. Then, a heuristic utilization-balancing trajectory planning algorithm was developed to optimize the assignment of clusters among UAVs as well as the location of launch sites. We also gave the analysis of the lower and upper bound of the formulated trajectory planning problem to show the performance guarantee. Numerical results validated that the performance of our proposal is close to the optimal, which can provide a guideline for data collection design in practice.

References

[1] S. H. Alsamhi *et al.*, "Green internet of things using UAVs in B5G networks: A review of applications and strategies," *Ad Hoc Netw.*, vol. 117, no. October 2020, p. 102505, 2021.

[2] J. Yick, B. Mukherjee, and D. Ghosal, "Wireless sensor network survey," *Comput. Netw.*, vol. 52, no. 12, pp. 2292–2330, 2008.

[3] M. Zhao, Y. Yang, and C. Wang, "Mobile data gathering with load balanced clustering and dual data uploading in wireless sensor networks," *IEEE Trans. Mob. Comput.*, vol. 14, no. 4, pp. 770–785, 2015.

[4] Y. Sun, T. Wang, and S. Wang, "Location optimization and user association for unmanned aerial vehicles assisted mobile networks," *IEEE Trans. Veh. Technol.*, vol. 68, no. 10, pp. 10 056–10 065, Oct. 2019.

[5] Y. Lin, T. Wang, and S. Wang, "UAV-assisted emergency communications: An extended multi-armed bandit perspective," *IEEE Commun. Lett.*, vol. 23, no. 5, pp. 938–941, 2019.

[6] J. Gong *et al.*, "Flight time minimization of UAV for data collection over wireless sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 9, pp. 1942–1954, Sept. 2018.

[7] J. Liu *et al.*, "UAV-aided data collection for information freshness in wireless sensor networks," *IEEE Trans. Wirel. Commun.*, vol. 20, no. 4, pp. 2368–2382, Apr. 2021.

[8] O. Ghdiri *et al.*, "Energy-efficient multi-UAV data collection for IoT networks with time deadlines," in *Proc. IEEE GLOBECOM'20*, Taibei, China, Dec. 2020.

[9] Y. Wang *et al.*, "Multi-UAV collaborative data collection for IoT devices powered by battery," in *Proc. IEEE WCNC'20*, Seoul, Korea (South), May 2020.

[10] X. Li *et al.*, "Rechargeable multi-UAV aided seamless coverage for QoS-guaranteed IoT networks," *IEEE Internet Things J.*, vol. 6, no. 6, pp. 10 902–10 914, Dec. 2019.

[11] M. Won, "UBAT: On jointly optimizing UAV trajectories and placement of battery swap stations," in *Proc. IEEE ICRA'20*, Paris, France, May-Aug. 2020.

[12] Y. Zhu and S. Wang, "Aerial data collection with coordinated UAV and truck route planning in wireless sensor network," in *Proc. IEEE GLOBECOM'21*, Madrid,Spain, Dec. 2021.

[13] ——, "Efficient aerial data collection with cooperative trajectory planning for large-scale wireless sensor networks," *IEEE Trans. Commun.*, vol. 70, no. 1, pp. 433–444, Jan. 2022.

[14] A. Al-Hourani, S. Kandeepan, and S. Lardner, "Optimal LAP altitude for maximum coverage," *IEEE Wirel. Commun. Lett.*, vol. 3, no. 6, pp. 569–572, Dec. 2014.

[15] G. H. Ball and J. Hall, "Isodata: a novel method of data analysis and pattern classification," Stanford Res. Inst., Stanford, CA, USA, Tech. Rep. AD-699616, 1965.

[16] S. Lin and B. W. Kernighan, "An effective heuristic algorithm for the traveling-salesman problem," *Oper. Res.*, vol. 21, no. 2, pp. 498–516, Mar.-Apr. 1973.

[17] K. Helsgaun, "General k-opt submoves for the Lin-Kernighan TSP heuristic," *Math. Program. Comput.*, vol. 1, no. 2-3, pp. 119–163, Jul. 2009.

[18] J. G. Carlsson and S. Song, "Coordinated logistics with a truck and a drone," *Manage. Sci.*, vol. 64, no. 9, pp. 4052–4069, Sept. 2018.

Authorized licensed use limited to: Nanjing University. Downloaded on February 13,2023 at 08:11:33 UTC from IEEE Xplore. Restrictions apply.