2024 IEEE 21st International Conference on Mobile Ad-Hoc and Smart Systems (MASS) | 979-8-3503-6399-9/24/$31.00 ©2024 IEEE | DOI: 10.1109/MASS62177.2024.00045

Optimizing UAV-Assisted Data Collection in IoT Sensor Networks using Dual Cluster Head Strategy

Keiwan Soltani*, Federico Corò†, and Sajal K. Das*
*Missouri University of Science and Technology, Rolla, USA; Email: {ksoltani@mst.edu, sdas@mst.edu}
†University of Padua, Padua, Italy; Email: federico.coro@unipd.it

***Abstract*—The proliferation of the Internet of Things (IoT) has significantly impacted the integration of digital and physical realms, with Wireless Sensor Networks (WSNs) playing a crucial role. However, these sensor nodes often face challenges related to battery constraints and deployment in inaccessible terrains. The advent of Unmanned Aerial Vehicles (UAVs) presents a transformative solution, particularly for data collection from remote IoT devices. This work explores the application of UAVs to improve data collection in dense IoT sensor networks. We propose a novel approach called optimizing UAV-assisted data collection in IoT sensor networks using Dual Cluster Head (UAVDCH) that utilizes dual cluster heads within each cluster to optimize the UAV's energy consumption. The primary cluster head is responsible for collecting data within the cluster, while the secondary cluster head is tasked with transmitting the data to the UAV. Our objective is to maximize the available data the UAV collects with respect to its energy constraints. We develop a strategy for selecting appropriate secondary cluster heads, determining UAV's hovering points, and designing flight trajectories that maximize data collection. By adopting a multi-channel technique, we facilitate simultaneous data collection from multiple clusters, reducing hovering and transmission times. Experimental results demonstrate that our algorithm outperforms existing methods, offering a promising solution for energy-efficient data collection in IoT sensor networks.**

***Index Terms*—IoT, UAV, Clustering, Dual-cluster heads, Flight trajectory, Data collection**

I. Introduction

The IoT represents a paradigm shift in integrating digital and physical worlds, with WSNs playing a crucial role. These networks consist of small sensor nodes that can sense, process, and transmit data [1]. The sensed data from sensors needs to be forwarded to the base station for further analysis. According to the transmission range limitation and ground obstacles like ponds, sensors cannot send their data directly or by using multi-hop relays to the base station, necessitating the implementation of intermediary devices or protocols for data retrieval. This retrieval process becomes pivotal in ensuring that valuable data from these remote sensors can be accessed and utilized effectively for various applications and analyses.

Unmanned Aerial Vehicles (UAVs) have emerged as a transformative solution due to their rapid deployment, high flexibility, and adaptability in various scenarios [2]. Their line-of-sight link establishment, reliable connectivity, and ability to access remote regions have made them a focal point of research and applications. Especially in scenarios where sensor nodes are dispersed in inaccessible areas, UAVs serving as mobile relay nodes offer a solution to the traditional bottleneck problems faced by ground-based multi-hop communication systems. However, the convergence of UAV and IoT systems presents challenges, particularly the need for energy-efficient and reliable routing protocols to ensure optimal data collection and delivery from ground sensor nodes. This intersection opens up a new approach to exploration and innovation, promising to redefine the landscape of networks and applications.

This paper explores how to maximize data collection within IoT sensor networks using an energy-constrained UAV. In our proposed scenario, a UAV is used to collect data from sensors while navigating predefined routes. However, due to the limited battery capacity of the drone during flight and hovering, it is essential to plan an efficient route to optimize data collection efficiency while conserving energy resources. Our approach begins by utilizing a clustering technique to effectively organize IoT sensors into clusters, employing the K-means clustering algorithm. Then, within each cluster, we designate a Primary Cluster Head (PCH) and a Border Cluster Head (BCH). Following this, we pinpoint hovering points for the UAV to facilitate data collection. Hovering points are then identified for the UAV from the BCH points set. Finally, we adopt a multi-channel approach that enables the UAV to simultaneously collect data from multiple $BCHs$. One possible scenario for deploying the proposed method is agricultural monitoring. Consider a large agricultural field equipped with numerous IoT sensors for monitoring soil moisture, temperature, humidity, and crop health. By utilizing the proposed method, UAVs can efficiently collect data from the sensors even in expansive and difficult-to-access areas of the field. This allows for timely and accurate data collection, enabling better crop management and irrigation planning.

The novelty of this work lies in defining a new method for selecting the hovering points for the UAV utilizing dual clustering. To the best of our knowledge, this is the first work that uses dual clustering in IoT sensor networks for data collection by a UAV. This approach leads to a decrease in the number of hovering points and the UAV's travel trajectory. We demonstrate the effectiveness of our algorithm experimentally, showing significant improvements in data collection efficiency compared to the existing methods.

The remainder of this paper is organized as follows: Section II reviews related work. Section III introduces the system model, energy model, and problem definition. Section IV discusses the proposed method in detail, including cluster

2155-6814/24/$31.00 ©2024 IEEE
DOI 10.1109/MASS62177.2024.00045

formation, UAV trajectory design, and data collection. Experimental results are given in Section V. Finally, Section VI discusses the conclusion and future work.

II. Related Work

Recently, the focus has shifted toward using UAVs in WSNs. These UAVs can easily navigate over barriers, making them effective for data collection [1], [3], [4]. The data collection problem has been studied under different variants, such as Age of Information (AoI) and delay-sensitive data. In [5], a UAV-assisted IoT data collection mechanism based on an aerial collaborative relay and AoI-sensitive data collection (ADC) scheme is investigated. The authors in [6] studied the deployment of UAVs to collect data from IoT devices by finding a data collection tour for each UAV. To ensure the freshness of the collected data, the total time spent on the tour must not exceed a given delay.

Several studies have explored clustering IoT devices to enhance data collection efficiency, with UAVs concentrating on cluster centers for optimal data collection. The authors in [1], [3], [4] proposed the clustered IoT (CIoT) routing protocol to enhance message propagation efficiency in IoT networks and reduce control overhead messages. It is shown in [7] that the construction tree in the WSN using a clustering scheme substantially reduces the energy consumption of the network. A layered multi-hop clustering method for structuring large-scale networks is proposed in [8], which established a multi-hop uplink communication based on layer clustering in long-range communication (LoRa) networks for emerging IoT applications. An adaptive opportunistic clustering approach is proposed in [9] that uses computational intelligence for industrial IoT networks to increase mobility support and network lifetime. A multi-hop constant-time clustering algorithm for IoT networks [10] used smart load balancing to reduce computational complexity and enhance scalability. The authors in [11] proposed an optimum rotation scheduling (ORS) that utilizes Integer Linear Programming to find the optimum rotation strategy for selecting the cluster head to prolong the network lifetime.

While most studies on the use of UAVs for data collection have focused on the one-to-one data collection scheme, the authors in [12] proposed a data collection method in WSNs by adopting a one-to-many data collection scheme to maximize the volume of data collected, subject to the energy capacity of the UAV. The authors in [13] studied a method that considers a one-to-many data collection scheme by adopting the orthogonal frequency division multiple access (OFDMA) technique to maximize data collection in IoT sensor networks through an energy-constrained UAV. In the field of UAV path planning using the clustering method, many research studies have been conducted. The authors in [3] proposed an optimization strategy for UAV path planning, which searches for the optimal solution of the information age by combining a clustering algorithm and a genetic algorithm. A self-adaptive algorithm is presented in [14] based on clustering and symbiotic organism search optimization strategy to reduce the search time of UAVs in performing target-searching missions. By considering dual clustering, our proposed method reduces the number of hovering points and the total hovering time, leading to a decreased flight trajectory for the UAV and allowing it to collect more data in the network.

Over the years, the selection of a cluster head per cluster has been widely used in many studies. However, recent research suggests employing two cluster heads within a cluster. The work in [15] applied a fuzzy c-mean clustering approach combined with multi-objective particle swarm optimization to determine the roles of these cluster heads. In this model, the primary cluster head collects and aggregates data, which is then sent to the second cluster head. The second cluster head then transmits data through a mobile sink to the base station. Although the method presented in [15] decreases the trajectory of the mobile sink, the number of visiting points by the mobile sink to collect data remains the same. The authors in [4] proposed an energy-efficient dual cluster head algorithm, based on K-means and Canopy optimization. The primary cluster head is responsible for communicating with nodes in the same cluster. Since the vice cluster head is in charge of sending data to the base station, it is located in the nearest region in the cluster to the base station to consume less energy. This algorithm is inefficient when the base station is far from the sensor deployment area. Due to the limited transmission range of sensors, they are unable to send data directly to the base station. To address these problems, this paper proposes a dual clustering technique and deploys a UAV to collect data from the sensors. This limits UAV's hovering points for data collection from specific locations and leads to shorter flight distances. Ultimately, using OFDMA, the UAV can collect data from several cluster heads simultaneously.

III. Preliminaries

A. System Model

This study considers an IoT sensor network consisting of N sensors $S = \{s_1, s_2, \ldots, s_N\}$ randomly distributed in a monitoring area M. Let (x_i, y_i) denote the coordinates of an IoT sensor s_i, for $1 \leq i \leq N$. Sensors have to periodically transfer the recorded data to external devices, e.g., a server. Let $0 < \rho_i \leq \rho$ be the size of the data that each s_i wants to transfer. The size of the data depends on the type of sensor and the recorded data. For example, IoT sensors are used to monitor the temperature and heat in forests to detect bushfires [16].

We assume that all IoT sensors are homogeneous, with initial energy, data processing, and communication capabilities. A UAV is dispatched from the base station depot point $O = (0, 0)$ to collect information from the IoT sensors S. The flight mission of the UAV starts and finishes at O. The UAV operates by either flying from one point to another or hovering over a sensor to collect data. For simplicity in deployment, the UAV maintains a constant flight velocity v and operates at an optimal altitude h with a transmission range R. According to the altitude h of the UAV, the data reception range Rr of the UAV is $Rr = \sqrt{R^2 - h^2}$.

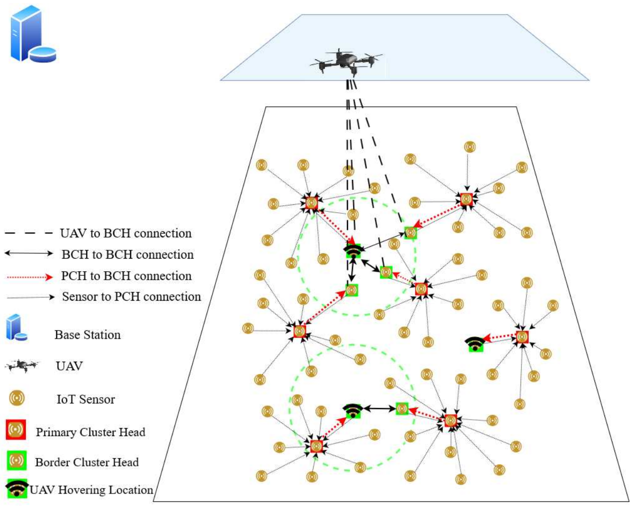

Fig. 1. Simultaneous Data Collection by the UAV from Multiple $BCHs$ in an IoT Clustered Sensor Network.

Throughout the paper, we use $d_{x,y}$ to compute the Euclidean distance between any two points x and y. The UAV is able to collect data from sensor s_i if $d_{UAV,s_i} \leq Rr$, i.e., if the Euclidean distance between UAV and s_i is within the reception range. Moreover, communication problems including multi-path propagation, fading, and shadowing are not addressed in this study. Given that the UAV is powered by a battery with limited energy, it can only fly for a limited amount of time within its energy constraint. Therefore, the maximum flight time of the UAV does not exceed a specified constraint B.

B. Energy Model

In this paper, two energy models are described: the first is for sending and receiving data among IoT sensors, and the second is for a UAV to collect data from the IoT sensors in the network. We apply the model proposed in [17] to calculate the energy required to send and receive a packet of length ρ among IoT sensors. Given the transmission range TR of IoT sensors, each sensor is capable of transmitting data to adjacent sensors within the coverage area. In addition, energy consumption varies based on the distance. The computation of energy consumption for sending and receiving data, relative to the distance between two adjacent IoT sensors, is as follows:

$$E_{TX}(\rho, d_{s_i,s_j}) = \begin{cases} \rho e_c + \rho \epsilon_{fs} d^2_{s_i,s_j} & d_{s_i,s_j} < d_0 \\ \\ \rho e_c + \rho \epsilon_{mp} d^4_{s_i,s_j} & d_{s_i,s_j} \geq d_0 \end{cases}$$

$$E_{RX}(\rho) = \rho e_c$$

where E_{TX} is the energy consumption for sending data between two sensors s_i, s_j with data length ρ at distance d_{s_i,s_j}. If d_{s_i,s_j} is less than the threshold (d_0), the free space mode is used. Otherwise, the multi-path fading channel model is applied to calculate E_{TX}. Parameters ϵ_{fs} and ϵ_{mp} denote the exhausted energy by the amplifier for free space and multi-path fading channel models, respectively. Moreover, e_c denotes the consumed energy by the electronic circuit and E_{RX} the energy consumed by the IoT sensor s_j to receive ρ bits of data. Table I shows the parameter values related to the energy model.

TABLE I
THE VALUES OF THE PARAMETERS OF THE IOT SENSOR ENERGY MODEL

Parameters	***Value***
e_c	50 nJ/bit
ϵ_{fs}	10 $pJ/bit/m^2$
ϵ_{mp}	0.0013 $pJ/bit/m^4$
d_0	$\sqrt{\frac{\epsilon_{fs}}{\epsilon_{mp}}}$

The energy consumption of a UAV is affected by the (i) energy required for traveling between hover points; (ii) energy expended while hovering at these points to gather data; and (iii) energy consumption for data reception.

The UAV starts its mission from point O to collect data from IoT sensors, then returns to the departure zone to recharge and deliver the collected data for further processing at the base station.

Let's denote the time taken by the UAV to complete its route as T, with T_h representing the time spent hovering, T_t the time spent traveling, and T_c the time spent collecting data. Hence, we can express T as $T = T_h + T_t + T_c$. Moreover, the total energy utilized by the UAV during the mission must be less than its energy budget B. Thus, $T_h \cdot \eta_h + T_t \cdot \eta_t + T_C \cdot E_{RX}(\rho) < B$, where η_h and η_t denote the energy used per unit of time while hovering and traveling, respectively [13].

C. Problem Definition

In this paper, we investigate the challenge of maximizing data collection in an IoT sensor network using a UAV. Given the set S of IoT sensor nodes dispersed across a geographical area, a UAV flies to the location of each IoT device to collect data, recall that the data size of sensor s_i is represented by ρ_i.

We aim to identify a set of optimal hovering points within the network, $\mathcal{H} = \{\mathcal{H}_0, \mathcal{H}_1, \dots \mathcal{H}_q\}$, where the UAV can hover to collect data from the sensor nodes, thereby maximizing the overall data collection efficiency. The key challenge in this problem lies in determining the optimal hovering points for the UAV to effectively communicate with the sensor nodes and collect their data. This task involves considering factors such as the sensing range (SR) and transmission range (TR) of the sensor nodes, the data reception range (Rr) of the UAV, and the UAV energy constraints (B).

The UAV must follow an energy-efficient trajectory to visit these optimal points and collect data from IoT sensor nodes while ensuring that its total energy consumption does not exceed its capacity. Formally, the problem can be defined as:

$$\max \sum_{i=1}^{N} \rho_i \quad \text{subject to} \quad \psi \leq B,$$

where ψ is the total energy consumption of the UAV, for completing a tour to collect data from IoT sensors within the network. Thus, the solution to this problem involves identifying the points for data collection, $\mathcal{H}$, and designing a flight path for the UAV that maximizes data collection efficiency while adhering to the system constraints.

IV. Proposed Method

This section elaborates on the details of optimizing UAV-assisted data collection in IoT sensor networks using the Dual Cluster Head (UAVDCH) method, which consists of methods for cluster formation, UAV trajectory design, and data collection in the IoT sensor network. In the cluster formation section, we devise clustering techniques to select two types of cluster heads, namely Primary Cluster Heads ($PCHs$) and Border Cluster Heads ($BCHs$). We select $PCHs$ by employing a K-means algorithm. After clustering the network and defining $PCHs$, which are responsible for receiving the data within the cluster, several IoT sensors are selected as $BCHs$ within the network. The PCH transmits the data to the BCH. Subsequently, a UAV collects data from the clusters by hovering over the BCH points.

Following the network clustering and designation of cluster heads, the UAV leverages the Orthogonal Frequency Division Multiple Access (OFDMA) technique [13] to concurrently gather data from multiple clusters. When the UAV hovers over a location with more than one BCH, it can simultaneously collect data from each BCH using distinct frequency channels as shown in Figure 1. These strategies not only enhance the UAV's operational efficiency but also maximize data collection in expansive and sparsely equipped monitoring environments. Furthermore, to perform data collection efficiently, it is essential to determine an optimal trajectory for the UAV that maximizes data collection while adhering to its energy constraints. To address this issue, we propose an approach where the UAV departs from the base station and selects the next hovering point as the one with the best ratio between travel distance and the number of nodes that could be covered. The UAV continues this process to gather more data, considering its energy reserves for the return trip to the base station. In the following, we elaborate on the aforementioned steps.

A. Cluster Formation

The first phase of the proposed algorithm organizes the IoT sensor nodes into clusters. IoT sensors generate data, and if data is transmitted directly, it can overload the network and energy consumption of the sensors. Clustering offers a hierarchical structure that divides the network into manageable groups, ensuring efficient data transmission and reduced energy consumption. Clustering involves partitioning the network into smaller groups, called clusters. Within each cluster, a node is selected as the Cluster Head (CH). The $CH's$ role is to collect data from member nodes and send it to the base station or other higher-level nodes. The objectives of clustering include energy efficiency, scalability, and latency reduction.

For clustering, several criteria and metrics are considered such as energy levels, proximity to neighbors, and other metrics which depend on the problem state. For example, IoT nodes with higher energy might be preferred as CHs. On the other hand, proximity to neighbors may improve communication efficiency. Although several algorithms like LEACH [18] or HEED [19] exist for clustering in IoT sensor networks, in our proposed method, we consider a K-means algorithm where a new approach is devised for determining K. In this regard, sensing range SR of IoT sensor is the critical metric for obtaining K by having $K = \sqrt{\frac{M}{AC}}$, with $AC = \pi \cdot (SR)^2$, where, M is the monitoring area. By obtaining K through this method, all sensors within the cluster can communicate with their cluster head with respect to the transmission range TR.

Therefore, after finding the proper K, the network is divided into K clusters utilizing the K-means algorithm. Then, we first define PCH, and later select BCH in each cluster. In the end, the data in the network will be collected by a UAV from the BCH in each cluster. The PCH is chosen in each cluster based on the distance to cluster members and the remaining energy of the candidate IoT sensor. The criteria for selecting a PCH in a cluster are computed by Eq. (1). The IoT sensor with the highest Fitness F is selected as PCH for the cluster. This process is repeated for each cluster to select the PCH.

$$F_{s_i} = (1 - \hat{d}_{s_i}) + Re_{s_i}, \qquad 1 \leq s_i \leq N \tag{1}$$

Where, $\hat{d}_{s_i}$ represents the mean Euclidean distance from candidate sensor s_i to all members (m) of the candidate's cluster, and is defined as:

$$\hat{d}_{s_i} = \frac{\triangle d_{s_i} - \min_{1 \leq s_j \leq m}(d_{s_i,s_j})}{\max_{1 \leq s_j \leq m}(d_{s_i,s_j}) - \min_{1 \leq s_j \leq m}(d_{s_i,s_j})}$$

where $\triangle d_{s_i} = \frac{\sum_{j=1}^{m} d_{s_i,s_j}}{m}$ for all pairs of sensors such that $1 \leq s_i, s_j \leq m$ and $s_i \neq s_j$. The remaining energy of s_i is defined and normalized as $Re_{s_i} = \frac{Re_{s_i}}{Ie}$, where Ie is its initial energy.

After selecting the PCH for each cluster, our next step is to select one IoT sensor as the BCH in every cluster. We aim to find the degree δ of cluster members in each cluster, defined as the number of other clusters that a node can cover based on the UAV's coverage range. Within each cluster, the degree δ is computed for all cluster members δs_i.

Then, the node with the highest degree δs and remaining energy Re above the threshold is selected as the BCH for the cluster. This process is repeated for each cluster in the network until all clusters have a designated BCH. The degree of each node in the cluster is computed by:

$$\delta s_c = \sum_{1 \leq i \leq k \,:\, i \neq z} |\{\forall s_j \in C_i : d_{s_j,s_c} \leq Rr\}|$$

Where C_i represents cluster i, C_z is the cluster that contains the s_c, and Rr is the reception range of the UAV. Our objective is to identify a BCH within each cluster that can cover the maximum number of available neighboring $BCHs$ concerning the UAV coverage range. Consequently, the network structure will resemble a forest composed of BCH trees. Each tree consists of direct edges (connections) to neighboring $BCHs$. Hence, each tree has a BCH as its root, which possesses the highest degree, and several $BCHs$ as leaves $\mathcal{T} = \{BCH_1, BCH_2, \ldots, BCH_k\}$, as depicted in Figure 1. According to Figure 1, the roots of the trees are labeled as UAV hovering points, and there is a BCH-to-BCH connection between the root and the leaves of each tree.

Algorithm 1 Find degree δ for each IoT sensor node within each cluster in the network.

Input: IoT sensors $S = \{s_1, \ldots, s_N\}$, consist of $PCH = \{PCH_1, \ldots, PCH_K\}$ and cluster members, as well as reception range of the UAV Rr.
Output: Degree δs of all IoT sensors
Create a Flag list for each sensor node with a length K, containing default zero values except for sensor's cluster.
for s_i=1 to N **do**
  Let $Flag_{S_i} = \{0_1, 0_2, \ldots, 0_K\}$.
  for s_j=1 to N **do**
    if $d_{s_i,s_j} \leq Rr$, $cluster_i \neq cluster_j$,$i \neq j$ **then**
      Turn the related Flag of $Cluster_j$ to 1, $Flag_{s_i}[cluster_j] = 1$
    end if
  end for
  δs_i = Count number of 1 in the $Flag_{s_i}$
end for
Return δs

After obtaining the degree of each cluster member from Algorithm 1, a node with the highest degree is selected as BCH for each cluster, and the number of created trees in the network is computed. The process of selecting BCH and computing the number of formed trees is explained in the Algorithm 2. This procedure is repeated until the minimum number of trees are found that cover all clusters. Since the root of the tree is the hovering point for the UAV, our ultimate goal is to minimize the number of hovering locations. For instance, considering a network consisting of multiple sensors, as depicted in Figure 1, the number of UAV hovering points using conventional clustering algorithms like Primary Cluster Head (PCH) is 7. However, by adopting the proposed clustering method, the number of hovering points needed to collect data from all clusters decreases from 7 to 3. In the subsequent phase, we define the UAV's flight path to travel and collect data from the IoT sensors in the network.

B. UAV Trajectory Design

In this phase, a trajectory within the network is designed to enable the UAV to optimally collect data from IoT sensor nodes. The objective is to devise a strategy for the UAV to gather the maximum amount of data by covering the maximum number of sensors in the network, considering its energy constraints. To this end, we determine an optimal path for the UAV to select specific hovering points that can cover the maximum number of nodes in the network. In our proposed method, the UAV does not visit each cluster head to collect data; instead, it visits only the root of the trees in the network.

Algorithm 2 Determining BCH points.

Input: Degree of IoT sensors δs.
Output: $\mathcal{T}$ containing BCH point for each cluster.
Sort δs in descending order
$Uncovered\_Cls \leftarrow$ all clusters
while $Uncovered\_Cls \neq$ empty **do**
  $BCH_i \leftarrow \max \delta s$, $s \in Uncovered\_Cls$
  $cluster_i \leftarrow$ Cluster which BCH_i is in it.
  create $\mathcal{T}_i$ and select BCH_i as the root, $1 \leq i \leq K$
  $Uncovered\_Cls \leftarrow Uncovered\_Cls \backslash cluster_i$
  for $c = 1 \quad to \quad (K - cluster_i)$ **do**
    for j=1 to N **do**
      if $d_{s_j,BCH_i} \leq Rr$ and $s_j \notin cluster_i$ and $cluster_c \notin Covered\_Cls_i$ **then**
        $Covered\_Cls_i \leftarrow cluster_c$
      end if
    end for
  end for
  for j= 1 to $|\, Covered\_Cls_i \,|$ **do**
    $candidates_j \leftarrow$ IoT sensors in $Covered\_Cls_j$ that has connection with BCH_i
    $BCH_j \leftarrow \arg\max_{s_a \in candidates_j}(Re(s_a))$
    Select BCH_j as the leaf of the $\mathcal{T}_i$
  end for
  $Uncovered\_Cls \leftarrow Uncovered\_Cls \backslash Covered\_Cls_i$
end while
Return $\mathcal{T}$

More precisely, with the network resembling a forest comprising q trees $\mathcal{T}$, the UAV visits only the root of these trees. The set of hovering locations for the UAV is a set of root BCH, represented as $\mathcal{H} = \{\mathcal{H}_0, \mathcal{H}_1, \ldots \mathcal{H}_q\}$. Consequently, the number of visiting and hovering locations for the UAV is significantly reduced. To pinpoint the precise hovering locations for the UAV, the positions of the root BCH in each tree serve as ideal spots. For each hovering point $\mathcal{H}$, a ratio is calculated using Eq. (2) and Eq. (3), which define the value of each hovering point. In other words, the UAV chooses hovering points which have the minimum ratio when considering the travel distance and the number of nodes that could be covered at the hovering points, as described in Eq. (4).

$$\mathcal{H}\_r_i = \frac{d_{UAV,\mathcal{H}_i}}{\alpha s_i} \tag{2}$$

$$\alpha s_i = \sum_{j=1}^{N} s_j \tag{3}$$

$$\min(\mathcal{H}\_r_i), \qquad \forall i \in \mathcal{H}_q \tag{4}$$

where αs_i is the number of sensors that can be covered when the UAV hovers at $\mathcal{H}_i$, and $d_{UAV,\mathcal{H}_i}$ is the Euclidean distance from the current location of the UAV to the hovering point $\mathcal{H}_i$. After computing the ratio for all designated hovering points, the UAV selects the next hovering point with minimum ratio using Eq. (4) while considering its energy constraint. More precisely, the UAV can collect data from more IoT sensor nodes while minimizing the travel distance. Each hovering point is the root of a $\mathcal{T}$, whose vertices are within each other's transmission range. When the UAV hovers over the root of a $\mathcal{T}$, it can simultaneously collect data from all vertices of the tree using the multi-channel method. This method enables the UAV to gather data from multiple clusters simultaneously, resulting in savings of both energy and time. After the UAV collects data from $\mathcal{T}_\infty$, it proceeds to the next hovering point. This process is repeated until the UAV has only enough battery to return to the base station to offload the collected data.

Recall that every IoT sensor node has a label corresponding to its coordinates (x_i, y_i). Let (x', y', h) represent the hovering coordinates of the UAV, where h is the UAV's flight altitude. This altitude, h, shouldn't exceed the transmission range, TR, of each IoT sensor node. Additionally, Tr represents the data transmission rate of the IoT sensors. Each BCH can forward the accumulated data to the UAV at the data transmission rate Tr when the drone is inside its communication range TR. Therefore, the hovering altitude h for collecting data is constrained by $h \leq TR$ and it is also hypothesized that this altitude, h, remains consistent [19]. Following the data gathering paradigm, if all IoT devices are within the UAV transmission range, their data is accessible to the UAV, under the assumption that distinct communication channels are utilized by each device. With the help of the OFDMA technique [20], the UAV can collect the data of all $BCHs$ in its coverage area simultaneously as shown in Figure 1.

Since the $BCHs$ are tasked with sending the data from each cluster, the UAV requires a certain amount of time to collect the total data Td from each cluster. Therefore, the UAV hovers over the $BCH's$ location to collect the Td. The duration of the UAV's hover to collect data from BCH_i is $Ht_i = \frac{Td_i}{Tr}$. As the UAV's hovering location is the root of the tree, and since a tree may have more than one BCH, the UAV can collect data from several $BCHs$ simultaneously. Thus, the total hovering duration time at the root of a tree $\mathcal{T}_i$ is equal to $H\mathcal{T}_i = \frac{\triangle\mathcal{T}_i}{tc_i}$, where $\triangle\mathcal{T}_i = \sum_{j=1}^{C} Td_j$ and $tc_i = \max_{1\leq j\leq C}\left(Td_j/Tr\right)$. Where C represents the number of clusters that are vertices of the tree $\mathcal{T}_i$ and tc_i denotes the maximum amount of time required to collect data from the $BCHs$ of a $\mathcal{T}_\rangle$. The total energy consumption for hovering at the location $\mathcal{T}_i$ is computed as $\xi H\mathcal{T}_i = H\mathcal{T}_i \cdot \eta_h$.

The UAV saves a considerable amount of hovering time by collecting data from several $BCHs$ in $\mathcal{T}_\rangle$ using the OFDMA approach. Additionally, it eliminates the need to fly to each cluster's location to collect data from that specific cluster. Consequently, travel time, the number of hovering points visited, and overall travel distance are reduced. Furthermore, the UAV consumes energy to receive data from the IoT sensors while hovering at each location. The amount of energy the UAV consumes for collecting data $\triangle\mathcal{T}_i$ from each tree is computed as $\xi c_{\mathcal{T}_i} = \triangle\mathcal{T}_i \cdot e_c$, where e_c indicates the consumed energy by the electronic circuit of the UAV.

The flight path is predetermined before the UAV departs from the base station. At each hovering zone, the UAV assesses its remaining energy before proceeding to the next location. If the UAV has enough remaining energy to collect data at the next point and return to the base station, it continues; if not, it returns to the base station for recharging.

The energy required for the UAV to travel between hovering locations is determined by $\xi\mathcal{T}_i = T_{t_{i,j}}.\eta_t$ where $T_{t_{i,j}} = \frac{d_{\mathcal{T}_i,\mathcal{T}_j}}{v}, \mathcal{T}_i, \mathcal{T}_j \in \mathcal{T}$ and v denotes the velocity of the UAV and d represents the Euclidean distance between trees $\mathcal{T}_i$ and $\mathcal{T}_j$, while η_t signifies the UAV's energy consumption rate during travel. The overall energy consumption of the UAV consists of the energy expended for traveling, hovering, and data collection. The total energy consumption of the UAV, for completing a tour to collect data from IoT sensors within the network, is given by $\psi = \sum_{i=1}^{q}\left(\xi\mathcal{T}_i + \xi H\mathcal{T}_i + \xi c_{\mathcal{T}_i}\right)$, where $\psi \leq B$, and B denotes the energy capacity of the UAV's.

V. Performance Evaluation

In this section, we evaluate the performance of the proposed algorithm in terms of collected data and sensor energy consumption through experimental simulations.

A. Experimental Settings

This paper considers a sparse sensor network that consists of $N = \{500, 600, 700, 800, 900, 1000\}$ sensor nodes randomly deployed in a 1000 $\times$ 1000 meters square area. The data volume of each sensor node is randomly drawn from two ranges, the first range is 10 KB to 200 KB, and the second range is 100 KB to 500 KB. We assume that the transmission range TR of each sensor node is 150 meters, the sensing range SR of each sensor node is 25 meters, and the data transmission rate Tr of the sensor is 150 Kb. The UAV is initially deployed at a depot $O = (0, 0)$, and we consider three different energy capacities for it, i.e., $B = 1.5 \times 10^5, 3 \times 10^5$, and 4.5×10^5 joules at constant flying speed $v = 10m/s$ and the flying altitude $h = 100$ meters. The energy capacity of the real drone is $B = 3 \times 10^5$ [13], and we examine the performance of the UAV by decreasing and increasing this value by 50 percent. The energy consumption rates of the UAV on traveling and hovering are $\eta_t = 100J/s$ and $\eta_h = 150J/s$, as used in [13], respectively. The value in each figure is the mean of the results from 20 network instances of the same size. For each instance, sensor nodes are randomly deployed with random data sizes within the defined ranges. Table II lists the default settings of the parameters in this paper. All the simulated evaluation experiments are conducted using Python and Jupyter Notebook on a MacBook Pro equipped with an Apple M3 Pro chip, featuring a 12-core CPU, and 18 GB of RAM.

To evaluate the performance of the proposed UAVDCH algorithm, we compare it with PCH algorithm, MRE-s algorithm [21], and its improved version called IMRE-s algorithm.

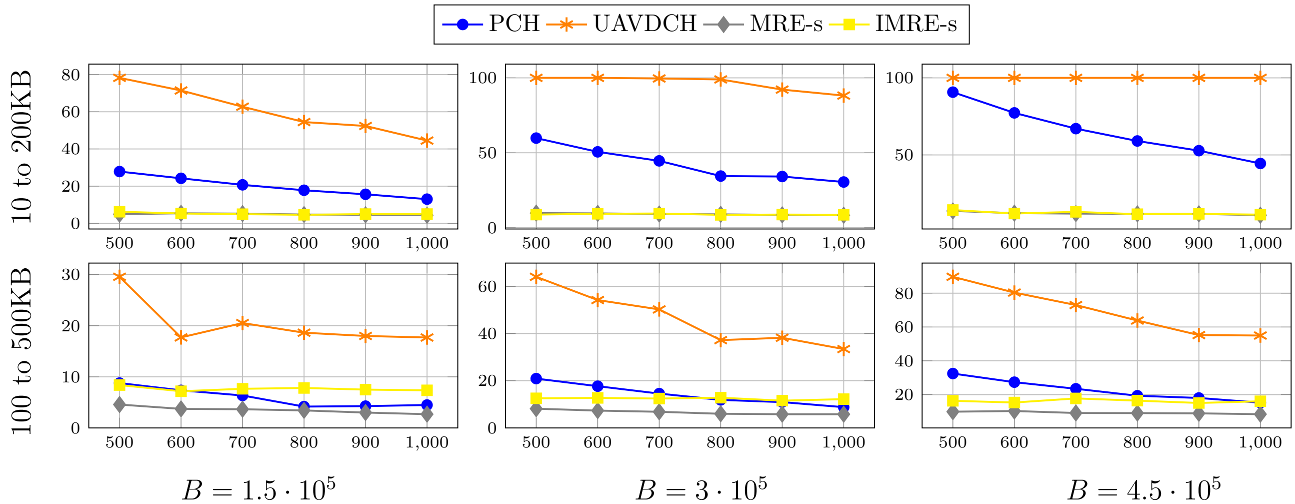

Fig. 2. UAV Data collection: comparison of all the algorithms when varying the drone's energy capacity and data size. In each plot, the y-axis represents the percentage of data collected in the network, and the x-axis represents the number of sensors.

TABLE II
TABLE OF PARAMETER SETTINGS

Parameters	***Value***
Sensing field	$1000m \times 1000m$
Network size	500 - 1000 Sensor nodes
Sensor data volume ρ	$10KB$ - $200KB$, $100KB$ - $500KB$
Transmission range TR	$150m$
Sensor sensing range SR	$25m$
Flying altitude h	$100m$
Data transmission rate Tr	$150Kbps$
Energy capacity B	1.5×10^5, 3×10^5, 4.5×10^5 joules
Flying speed v	$10m/s$
Traveling consumption rate η_t	$100J/s$
Hovering energy rate η_h	$150J/s$

- The Primary Cluster Head (PCH) algorithm is similar to the proposed method, but it considers only one layer of clustering. In the PCH algorithm, sensors in the network are categorized into several clusters using the K-means algorithm, following the same method as our proposed method. Each cluster has a PCH sensor node, which serves as a hovering point for the UAV. Therefore, the UAV collects data from the sensors by traveling and hovering over the PCH locations until its battery has limited energy for returning to the base station.
- The Max Ratio Reward-Energy with a single drone (MRE-s) algorithm is a heuristic approach designed to address the Maximum Data Collection Problem using a single UAV. It operates by iteratively selecting sensors based on the highest ratio of data reward to additional energy cost, ensuring efficient data collection while managing energy constraints. In this algorithm, the UAV hovers and collects data from sensors within its coverage area. Although the UAV may cover more than one sensor at a hovering point in MRE-s, it collects data from only one sensor at a time.
- In the IMRE-s algorithm, we integrate the OFDMA technique, enabling the UAV to simultaneously gather data from multiple sensors at the hovering point. This variant has been adopted to address a limitation of the original MRE-s algorithm, which could only collect data from one sensor at a time.

TABLE III
SENSOR ENERGY CONSUMPTION RATIO BY NUMBER OF COVERED NODES

# Sensors	*PCH*	*UAVDCH*	*MRE-s*	*IMRE-s*
500	0.177	0.240	0.070	0.053
600	0.175	0.241	0.077	0.053
700	0.177	0.245	0.085	0.059
800	0.179	0.263	0.082	0.050
900	0.180	0.259	0.087	0.058
1000	0.185	0.273	0.089	0.056

B. Performance Evaluation of Different Algorithms for Data Collection and Sensor Energy Consumption

We first evaluate the data collection performance of the various algorithms. As depicted in Figure 2, our proposed algorithm, UAVDCH, demonstrates enhanced performance compared to other algorithms. With a network comprising 500 sensors, an energy capacity of $B = 1.5 \times 10^5$ joules, and data sizes ranging from $10KB$ to $200KB$, our method achieves a data collection rate of 80 percent, which is markedly higher than the other three algorithms. Similarly, for the same data size, the PCH algorithm ranks second in terms of data collection, outperforming both MRE-s and IMRE-s owing to its clustering-based model. Notably, both the proposed method and the PCH algorithm utilize a clustering technique leading to higher energy consumption for sensors in both algorithms compared to other algorithms, as detailed in Table III. Table III elaborates on the ratio of sensor energy consumption by the number of covered nodes when $B = 3 \times 10^5$ in a $1000m \times 1000m$ sensing field. It should be noted that nearly one-third of the sensor energy consumption in the network is related to transmitting data from PCH to BCH in the UAVDCH algorithm, while the rest is used to transmit the data from BCH points to the UAV. Moreover, by increasing the UAV's energy capacity to $B = 3 \times 10^5$ joules or more, the proposed method enables the UAV to completely collect data (exceeding 100 MB) from all sensors in the network, as demonstrated in Figure 2.

TABLE IV
NUMBER OF SENSORS COVERED BY THE UAV IN A $1000m \times 1000m$ SENSING FIELD

# Sensors	PCH	UAVDCH	MRE-s	IMRE-s
500	102	318	37	65
600	104	323	44	80
700	100	346	48	90
800	101	355	50	100
900	102	360	59	113
1000	93	375	60	126

Furthermore, as the data size increases to the range of $100KB$ - $500KB$, the UAVDCH algorithm consistently outperforms the others, capturing the highest percentage of total sensor data collection (more than 300 MB) in the network. For instance, UAVDCH can collect data from 318 sensors in $1000m \times 1000m$ sensing field with an energy capacity of $B = 3 \times 10^5$ as shown in Table IV. In contrast, the performance of the PCH algorithm substantially decreases, while the IMRE-s algorithm exhibits better performance than the PCH algorithm with larger data sizes, as presented in Table IV. The decline in PCH algorithm performance when handling large data sizes can be attributed to its operational mechanism, where the UAV is required to travel and hover solely over cluster heads. At each hovering point, the UAV receives data from only one cluster head, leading to increased time for hovering and data collection compared to the IMRE-s algorithm. In the IMRE-s algorithm, multiple sensors can communicate with the UAV at the hovering point, allowing the UAV to simultaneously collect data from several sensors. This approach consistently outperforms the MRE-s algorithm. Moreover, as can be seen from Figure 2 and Table III, the IMRE-s algorithm not only achieves higher data collection than the PCH algorithm but also demonstrates significantly lower energy consumption for sensors.

It can be observed that in both data ranges, the UAVDCH algorithm outperforms the others, and when the data size is small, the UAVDCH method can collect 100 percent of the data in most cases. The enhanced performance of the UAVDCH algorithm compared to the MRE-s and IMRE-s algorithms is attributed to the implementation of double clustering, which identifies potential hovering points for the UAV in advance and establishes the travel trajectory of the UAV by computing the ratio of each point. In addition, the number of hovering points and the travel distance are reduced in comparison with other algorithms. In contrast, the MRE-s and IMRE-s algorithms are greedy algorithms, and the UAV needs to decide the points step by step. Since these algorithms do not incorporate clustering, they consume more time for traveling and hovering over more points while collecting less data compared to our proposed method.

TABLE V
NUMBER OF HOVERING POINTS IN THE SENSING FIELD

Sensing field	PCH	UAVDCH
500 × 500	11	4
1000 × 1000	23	8
1500 × 1500	34	13
2000 × 2000	45	17

Although the PCH algorithm also employs clustering, it results in a greater number of hovering points than UAVDCH, leading to a longer UAV trajectory. Table V illustrates the effectiveness of the proposed method in reducing the number of hovering points for the UAV across sensing fields. For instance, within a $1000m \times 1000m$ sensing field, the number of hovering points was remarkably reduced by the UAVDCH algorithm compared to the PCH algorithm, from 23 to 8. Therefore, by traveling to fewer hovering points, the UAV can cover all sensors in the sensing field. As in the example depicted in Figure 1, the UAV can collect the data of all sensors in the network by just hovering over three points (UAVDCH) instead of seven (PCH) points, which decreases the total travel distance of the UAV. Consequently, our proposed method substantially enhances the energy efficiency of UAVs operating under energy constraints by decreasing and defining appropriate hovering points.

VI. CONCLUSION

In this study, we have introduced a novel approach for optimizing UAV-assisted data collection in IoT sensor networks using a dual cluster head strategy (UAVDCH). Our approach utilizes a dual-cluster head strategy to decrease the number of hovering points in the sensing field and to shorten the flight trajectory, thereby significantly enhancing data collection efficiency while optimizing the UAV's energy consumption. By employing a multi-channel technique, our method allows simultaneous data collection from multiple clusters, reducing hovering and transmission time.

Key Contributions:

- **Dual Cluster Head Strategy**: Implemented a dual cluster head strategy to minimize the number of UAV hovering points and shorten the flight trajectory.
- **Effective Hovering Points**: Utilized Border Cluster Heads (BCHs) as UAV hovering points to minimize the number of required hovering locations, thereby reducing flight duration and distance.
- **Multi-Channel Technique**: Developed a multi-channel approach for simultaneous data collection from multiple clusters, reducing both hovering and transmission times.
- **Efficiency Improvement**: Demonstrated significant improvements in data collection efficiency and UAV energy consumption.

Experimental results demonstrate that our proposed algorithm UAVDCH outperforms existing methods in terms of data collection efficiency, particularly in sparse sensor networks with varying data sizes. The use of Border Cluster Heads ($BCHs$) as UAV hovering points has proven to be effective in minimizing the number of required hovering locations, thereby reducing flight duration and distance.

Future work will focus on addressing the challenges of non-line-of-sight conditions, and dynamic environments, and exploring adaptive strategies to further optimize the UAV's flight path and energy consumption. Additionally, we aim to investigate the integration of machine learning techniques to enhance the decision-making process for UAV trajectory design. In addition, we plan to utilize multiple UAVs for collecting data in larger sensing fields.

Acknowledgements: This work was supported by NSF grants under award numbers SCC-1952045, OAC-2104078, and ECCS-2319995.

REFERENCES

[1] A. Abuashour, "An efficient clustered iot (ciot) routing protocol and control overhead minimization in iot network," *Internet of Things*, vol. 23, p. 100839, 2023.

[2] S. A. H. Mohsan, N. Q. H. Othman, Y. Li, M. H. Alsharif, and M. A. Khan, "Unmanned aerial vehicles (uavs): Practical aspects, applications, open challenges, security issues, and future trends," *Intelligent Service Robotics*, vol. 16, no. 1, pp. 109–137, 2023.

[3] J. Liu, P. Tong, X. Wang, B. Bai, and H. Dai, "UAV-aided data collection for information freshness in wireless sensor networks," *IEEE Trans. on Wireless Communications*, vol. 20, no. 4, pp. 2368–2382, 2020.

[4] M. Wu, Z. Li, J. Chen, Q. Min, and T. Lu, "A dual cluster-head energy-efficient routing algorithm based on canopy optimization and k-means for wsn," *Sensors*, vol. 22, no. 24, p. 9731, 2022.

[5] X. Huang and X. Fu, "Fresh data collection for uav-assisted iot based on aerial collaborative relay," *IEEE Sensors Journal*, vol. 23, no. 8, pp. 8810–8825, 2023.

[6] W. Xu, T. Xiao, J. Zhang, W. Liang, Z. Xu, X. Liu, X. Jia, and S. K. Das, "Minimizing the deployment cost of uavs for delay-sensitive data collection in iot networks," *IEEE/ACM Transactions on Networking*, vol. 30, no. 2, pp. 812–825, 2022.

[7] K. Soltani, L. Farzinvash, and M. A. Balafar, "Trust-aware and energy-efficient data gathering in wireless sensor networks using PSO," *Soft Computing*, vol. 27, no. 16, pp. 11 731–11 754, 2023.

[8] A. B. Bomgni, H. M. Ali, M. Shuaib, Y. Mtopi Chebu *et al.*, "Multihop uplink communication approach based on layer clustering in LoRa networks for emerging IoT applications," *Mobile Information Systems*, vol. 2023, 2023.

[9] P. Chithaluru, A.-T. Fadi, M. Kumar, and T. Stephan, "Computational intelligence inspired adaptive opportunistic clustering approach for industrial iot networks," *IEEE Internet of Things Journal*, 2023.

[10] Y. B. C. Mtopi, A. B. Bomgni, H. M. Ali, D. R. G. Zanfack, W. Ejaz, C. T. Djamegni, and E. Z. Gnimpieba, "Multihop optimal time complexity clustering for emerging iot applications," *Cluster Computing*, vol. 26, no. 2, pp. 993–1009, 2023.

[11] L. Mostarda, A. Navarra, and R. De Leone, "Optimal vs rotation heuristics in the role of cluster-head for routing in iot constrained devices," *Internet of Things*, vol. 22, p. 100757, 2023.

[12] M. Chen, W. Liang, and Y. Li, "Data collection maximization for uav-enabled wireless sensor networks," in *29th International Conference on Computer Communications and Networks (ICCCN)*, 2020, pp. 1–9.

[13] 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 Transactions on Mobile Computing*, vol. 22, no. 1, pp. 159–174, 2021.

[14] J. Chen, Y. Zhang, L. Wu, T. You, and X. Ning, "An adaptive clustering-based algorithm for automatic path planning of heterogeneous uavs," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 9, pp. 16 842–16 853, 2021.

[15] A. Zheng, Z. Zhang, W. Liu, J. Liu, Y. Xiao, and C. Li, "Dual cluster head optimization of wireless sensor networks based on multi-objective particle swarm optimization," *Sensors*, vol. 23, no. 1, p. 231, 2022.

[16] M. T. Lazarescu, "Design of a wsn platform for long-term environmental monitoring for iot applications," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 3, no. 1, pp. 45–54, 2013.

[17] W. B. Heinzelman, A. P. Chandrakasan, and H. Balakrishnan, "An application-specific protocol architecture for wireless microsensor networks," *IEEE Transactions on wireless communications*, vol. 1, no. 4, pp. 660–670, 2002.

[18] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in *Proceedings of the 33rd annual Hawaii international conference on system sciences*. IEEE, 2000, pp. 10–pp.

[19] A. Al-Hourani, S. Kandeepan, and S. Lardner, "Optimal lap altitude for maximum coverage," *IEEE Wireless Communications Letters*, vol. 3, no. 6, pp. 569–572, 2014.

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

[21] F. Betti Sorbelli, A. Navarra, L. Palazzetti, C. M. Pinotti, and G. Prencipe, "Wireless iot sensors data collection reward maximization by leveraging multiple energy- and storage-constrained uavs," *Journal of Computer and System Sciences*, vol. 139, p. 103475, 2024.