# Reducing Information Gathering Latency through Mobile Aerial Sensor Network

Zhaoquan Gu*, Qiang-Sheng Hua*, Yuexuan Wang* and Francis C.M. Lau†
*Institute for Theoretical Computer Science, Institute for Interdisciplinary Information Sciences,
Tsinghua University, Beijing, 100084, P.R. China
†Department of Computer Science, The University of Hong Kong, Hong Kong, P.R. China

***Abstract*—Gathering information in a sensing field of interest is a fundamental task in wireless sensor networks. Current methods either use multihop forwarding to the sink via stationary nodes or use mobile sinks to traverse the sensing field. The multihop forwarding method intrinsically has the energy hole problem and the mobile sinks method has a large gathering latency due to its low mobility velocity. In addition, all the mobile sinks methods assume unlimited power supply and memory which is unrealistic in practice. In this paper, we propose a new approach for information gathering through a Mobile Aerial Sensor Network (MASN). We adopt the *Hive-Drone* model [5] where a centralized station (Hive) responsible for serving and recharging Micro-Aerial Vehicle (MAV) sensor nodes (Drones) is strategically placed in the sensing field. We then face the challenges of how to control the mobility of each MAV and devising interference-free scheduling for wireless transmissions that can substantially reduce the latency. We present a family of algorithms with constant memory to reduce both *gathering latency*, which is the duration from dispatching the MAVs to the moment when all the sensed information are gathered at the central station, and *information latency*, which is the duration from when some information is sensed to when it is received by the station. We also consider how to extend the single Hive to multiple Hives for monitoring an arbitrarily large area. Extensive simulation results corroborate our theoretical analysis.**

***Index Terms*—Micro-Aerial Vehicle, Information Gathering, Sensor Networks, Gathering Latency, Information Latency.**

## I. Introduction

Wireless sensor networks have been widely used in environmental monitoring, urban surveillance, disaster recovery, industrial sensing and control, etc. [7], [15], [18]. Their primary task is to gather crucial information of the sensing field for subsequent decisions to be made such as where pests or diseases exist in crops or in order to provide a bird's eye view of the environment for disaster recovery.

Traditional methods for information gathering begin with deploying a large number of stationary sensors randomly or uniformly in the target field. Two approaches of data gathering are widely applied. One is through the wireless sensor network. In this approach, all information will be forwarded hop by hop to the sink node. However, this kind of multi-hop forwarding method may suffer from the energy hole problem [19] where the nodes near the sink will deplete their power more quickly than the nodes far away. Thus the sensor network may become disconnected after being deployed for some length of time. The other approach utilizes mobile elements [8], [20], in which the land-based mobile sinks will traverse each sensing field to collect the needed information. This approach, however, suffers from the following problems. First, the gathering latency, i.e., the time for gathering all the information, may be large as the mobility velocity is much smaller than the network transmission speed. The prolonged gathering time may affect significantly subsequent decisions to be made using the gathered information. Second, since the mobile sink will traverse each sensing field to collect the field's information, it has to have a large enough energy supply and memory space; for a relatively large field, this assumption will limit the method's applicability in practice.

In this paper, we propose information gathering through a Mobile Aerial Sensor Network (MASN) based on the *Hive-Drone* model [5]. MASN comprises a swarm of Micro-Aerial Vehicles (MAVs), each of which carries a sensor and can fly to specific regions for sensing. In the *Hive-Drone* model, the Hive is a centralized station which is suitably placed in the sensing field for recharging the MAVs' batteries and deciding how to dispatch the MAVs (*Drones*) to perform specific tasks. This model is attractive for two reasons. First, it is more convenient to deploy since the Hive can decide each MAV's flying track and it can fly to a region without the need for deploying the stationary sensor nodes in advance. Second, the MAV sensors (Drones) whenever necessary can fly back to the Hive to be recharged, which avoids the energy hole problem.

Although attractive, the MASN method faces some tough challenges. Actuation dominates the energy budget of each MAV, leaving the bare minimum for sensing, control and communication. Such a limit calls for a more refined mobility control for reducing energy consumption. In addition, in order to reduce the information gathering latency, this kind of mobility control needs to be jointly considered with the design of interference-free wireless transmission scheduling. For example, should the MAV fly to another sensing field to gather new information or stay in the current sensing field acting as the relay node for transmitting information from the other MAV nodes needs precise coordination.

The contribution of this paper is as follows: First of all, to overcome the drawbacks of the two frequently used methods for information gathering, i.e., the multihop-forwarding method via the stationary wireless sensor network and the mobile sinks method, we introduce the *Hive-Drone* model which not only can avoid the energy hole problem, but also reduce the information gathering latency with only constant

978-1-4673-5946-7/13/$31.00 ©2013 IEEE

memory. In particular, for the case where there is only one MAV (Drone), we present a greedy algorithm which yields a constant approximation ratio. For the case where there are enough MAVs to cover the entire monitored area, we present constant approximation algorithms under both graph based and physical (SINR) interference models. Finally, for the case where there is only a small number of MAVs, we propose another approach with a slightly increased gathering latency at the cost of uncovered boundaries of the monitored area. Moreover, we extend this approach to the multiple Hives scenario in order to monitor a much larger area.

The remainder of the paper is organized as follows. The next section gives the related work. Preliminaries are provided in Section III. A greedy constant approximate algorithm is described in Section IV for the case of only one MAV. When there are many MAVs to cover the monitored area, we present how to jointly control the MAV mobilities and design interference-free wireless transmission schedules for reducing both gathering and information latencies in Section V. We analyze the case where there is only a small number of MAVs and consider the multiple Hives scenario for monitoring an arbitrarily large area in Section VI. We conclude the paper with some future work suggestions in Section VII.

## II. Related Work

Information gathering is a fundamental task in wireless sensor networks. Many algorithms have been proposed in the past decades. Generally speaking, there are two categories of work in this area. The first uses stationary wireless sensor networks where the generated data will be forwarded to the sink node hop by hop. We call this method the multi-hop forwarding method. The other category uses mobile elements (sinks) to traverse each sensing field in the monitored area for information collection. We briefly survey each of the two methods. A well known shortcoming of the multi-hop forwarding method is that it has the notorious energy hole problem in which the sensor nodes near the sink will deplete their power much quickly than the nodes that are far away. This phenomenon significantly reduces the wireless sensor network's lifetime. A number of methods have been proposed to mitigate the energy hole effect, such as using non-uniform node distributions [19]. Aiming at minimizing the gathering latency, Dominating Tree is proposed in [16] and [17] under the physical interference model and for a multi-channel multihop network, respectively. [4] derives the upper and constructive lower bounds for data collection capacity theoretically in arbitrary networks and analyzes the lower bounds for both the protocol interference model and the general graph model. Moreover, degree-constrained spanning trees and capacitated minimal spanning trees are proposed to show the improvements, and the impacts of different interference models are also considered; however, all these methods did not consider how to eliminate the energy hole.

Although the mobile elements methods [6] could avoid the energy hole, as mentioned earlier, they implicitly assume unlimited power and memory in the mobile elements, which makes them impractical when the monitored area is relatively large. In this line of work, three kinds of mobilities have been studied: random mobility [13], predictable mobility [3] and controlled mobility [8]. [13] initiated this topic by proposing a three-tier architecture for sparse sensor networks, where the mobile entities, referred to as Data Mobile Ubiquitous LAN Extensions (MULEs), move around in the network to collect information from sensor nodes and drop off the data to wired access points. [3] models the information gathering process as a queuing system with predictable mobility where random arrivals model randomness in the spatial distribution of sensor nodes. Power consumption of the network is quantified with the success analysis of information collection. [8] adopts controlled mobility and formulates the collection process as the Traveling Salesman Problem with Neighborhoods (TSPN). A combine-skip-substitute (css) scheme is proposed to reduce the tour length and data gathering time with mobile elements, which is the best known heuristic algorithm.

Instead of mobile sinks, mobile sensors are considered in [10] for improving the the network coverage. Several systems are designed around the notion of dispatching mobile sensors to generate and gather information. TriopsNet is an example for pipeline monitoring [9] and it can cover the sensing area with fewer sensors while maintaining a high data collection rate with a stable network topology. Along with the development of mobile sensing systems, Micro-aerial Vehicles (MAVs) has become an emerging topic at the forefront of systems research. In [12], MAVs are used in disaster management applications by providing a bird's view of the environment. Flight data exchanges can happen between MAVs and one major concern is how to place the sensing points in order to monitor the entire area as efficiently as possible. [5] proposed the Karma system based on the *Hive-Drone* model in crop monitoring which collects flower information for searching for pests and diseases. MAVs are dispatched from a station in the field and move around the sensing field to sense and bring back information. A pollen task will be activated when flowers in bloom are being detected. However, the Karma system suffers from large gathering latency and the information of blooming flowers may not reach the station quickly enough due to the slow mobility velocity. Thus, we propose information gathering through a Mobile Aerial Sensor Network (MASN) consisting of all the MAVs in the sensing field (monitored area) for reducing information gathering latency while providing the station with the crucial sensed information as quickly as possible.

## III. Preliminaries

### A. Hive-Drone Model

The *Hive-Drone* Model is introduced in [5] where a centralized station (Hive) is placed in the monitored area to serve the MAVs (Drones), in particular to recharge their batteries. Table I gives some notations used in this model. Generally speaking, $D_f \simeq v_f \cdot T_f$, $v_t > v_f$ and $P$ is much smaller than the flying power consumption.

TABLE I
NOTATIONS IN HIVE-DRONE MODEL

| Notations | Meaning |
|---|---|
| $N$ | Number of MAVs |
| $T_f$ | The maximum time for flying |
| $D_f$ | The maximum distance a MAV can fly |
| $r$ | The sensing radius of each MAV |
| $v_f$ | The flying speed of each MAV |
| $v_t$ | The network transmission speed |
| $P$ | Maximum power for each transmission |

### B. Interference Models

There are two commonly adopted interference models: the graph based model [1] and the physical model [11]. The graph based model assumes the network consists of two graphs: connectivity graph $G_C$ and interference graph $G_I$. A node $v$ can receive from $u$ successfully if it is connected with $u$ in $G_C$ but not with any other simultaneous transmitting nodes in $G_I$. *Hop distance model* and *protocol model* are two specific graph based interference models.

The $h$-hop distance model assumes that if the hop distance of two links is less than $h$ hops, then they cannot be scheduled in the same timeslot [14]. Note that 802.11 protocols mainly assume $h = 2$. The protocol model assumes the transmission range and the interference range vary with signal power. Two nodes are connected in $G_C$ ($G_I$) if their distance is no more than the transmission (interference) range. Typically, we assume the maximum transmission range is $R_T$ and the interference range $R_I$ varies with $R_T$ ($R_I \geq R_T \geq r$).

The physical model is more practical compared with graph based models. The classic one is the signal-to-interference-plus-noise ratio (SINR) model which is based on the study of fading channel models. The energy of a signal fades with the distance to the power of the path-loss exponent $\alpha$. If the signal strength received over the interfering strength of all other transmitters plus the background noise $I_b$ is above some threshold $\beta$, the recipient can decode the signal successfully. More precisely, it can be formulized as:

$$\frac{P_u/d(u,v)^\alpha}{I_b + \sum_{w \in S \setminus \{u\}} P_w/d(w,v)^\alpha} \geq \beta \tag{1}$$

where $d(\cdot)$ means the Euclidean distance and $S$ is the set of simultaneous transmitters.

### C. Problem Definition

In this paper, we first consider the case of only one Hive, as proposed in [5], and a bounded area for information gathering $\mathcal{F} = Disk(c, D)$ where $c$ is the center and $D$ is the radius satisfying $D = \frac{D_f}{2}$ in the worst case; we extend this to arbitrarily large monitoring area in Section VI. The Hive has three operations:

- $Dispatch(MAV_i, P_i, t_d(i))$: Dispatch $MAV_i$ to the sensing point $P_i$ at time $t_d(i)$;
- $Receive(m_i)$: Receive information $m_i$ and record the receive time as $t_r(m_i)$;
- $Recharge(MAV_i)$: Recharge $MAV_i$ when it is back to the Hive;

The operations of each MAV are as follows:

- $Fly(P_s, P_d)$: Fly from point $P_s$ to $P_d$ with speed $v_f$;
- $Land(P_i)$: Land at point $P_d$ and three operations are available:
  - $GI(m_i)$: Generate Information of the sensed area $F_i = Disk(P_i, r)$ (such as taking a photo of its sensed area) and assume $m_i$ costs one unit of memory;
  - $SI(m_i, MAV_j)$: Send Information $m_i$ to node $MAV_j$;
  - $RI(m_j)$: Receive Information from some other MAV node;
- $Back$: Fly back to the station for recharging and may carry some information to the station at the same time.

When the Hive performs $Dispatch$ $n$ times, a sequence of $n$ sensing points $\{P_1, P_2, \cdots, P_n\}$ is generated. For each sensing point $P_i$, suppose $MAV_i$ arrives there at time $t_a(i)$ and generates information at $t_g(m_i)$. For the sake of brevity, we omit the sensing delay and $t_g(m_i) = t_a(i)$. We assume these points can fully cover the sensing field $\mathcal{F}$ if $\mathcal{F} \subseteq \bigcup_{i=1,\cdots,n} F_i$.

Our goals are to reduce:

- **Gathering Latency**: $T_G = \max_i t_r(m_i)$ denotes the last information arrival time;
- **Information Latency**: $T_I = \max_i\{t_r(m_i) - t_g(m_i)\}$ is the longest delay for some information generated to be collected at the station.

$T_G$ is an important measure representing the time to gather all information. Low gathering latency implies high efficiency. Information latency is also important in practical applications because low information latency can facilitate subsequent tasks. For example, as shown in [5], if a MAV detects a blossomed flower, the Hive will arrange a Drone to pollen the flower as quickly as possible.

## IV. SINGLE MAV GATHERING

Consider the simplest case where only one MAV is available to fly back and forth for information sensing and gathering. This is quite similar to the mobile sink method, but as mentioned earlier, an MAV has only limited power and memory which means that it cannot visit each sensing field within one flight without recharging. Since the MAV has to reach at least $\frac{D^2}{r^2}$ points from simple area calculation, if only one unit of memory is allowed and it has to reach every point, then this would turn out to be a coverage problem. Thus, we assume the MAV can fly to different points to gather information with more memory on board. The velocity will decrease when the MAV performs the $Land(P_i)$ operation and increase to $v_t$ again when it $Fly$ to another point. Thus, it cannot fly the maximum distance $D_f$. Supposing the time from $Land$ to $Fly$ is $t_d$, denote the distance delay as $r_d \simeq \frac{v_f \cdot t_d}{2}$. If $k$ different points are reached for sensing, the maximum fly distance is $D'_f = D_f - kr_d$.

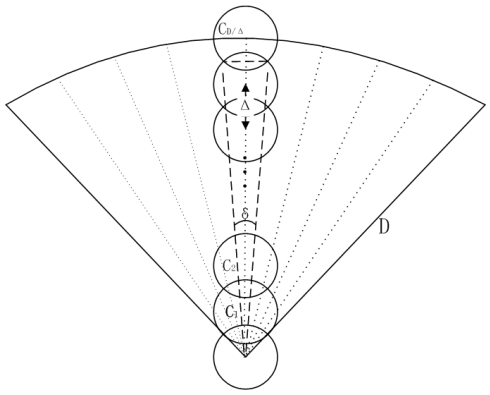



Fig. 1. Schedule to cover all points within the angle $\delta$

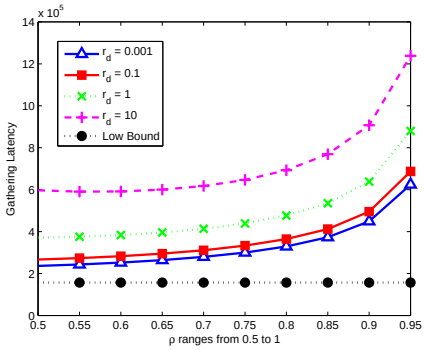



Fig. 2. Gathering Latency comparisons: $D = 3000m, \rho \in (1/2, 1)$

*Theorem 1:* The lower bounds for Single MAV Gathering are: $T_G = \Omega(\frac{D^2}{rv_f})$ and $T_I = \Omega(\frac{D}{v_f})$.

*Proof:* In order to cover all marginal points, the largest angle is $\gamma = 4\arcsin\frac{r/2}{D}$ when the MAV is actually at the margin and sensing a disk of radius $r$. Thus at least $\frac{2\pi}{\gamma} = \frac{\pi}{2\arcsin\frac{r/2}{D}} \simeq \frac{\pi D}{r}$ flights are needed. Each flight costs at least $\frac{D}{v_f}$ time reaching the margin and at least $\frac{D}{v_f}$ time when it carries back the information. Thus $T_G \geq \frac{2\pi D^2}{rv_f}$ and $T_I \geq \frac{D}{v_f}$, concluding the theorem. ■

We propose a greedy algorithm to cover the area within a certain angle. In Fig. 1, the distance of every two consecutive disks' centers is $\Delta = 2\rho r, \rho \in (1/2, 1)$. Obviously, the area within the angle $\delta = 2\arcsin\frac{\sqrt{1-\rho^2}r}{D} \simeq \frac{2\sqrt{1-\rho^2}r}{D}$ can be covered. We label all centers along the radius as $c_1, c_2, \cdots, c_{\frac{D}{\Delta}}$ and these points will be reached greedily for information gathering.

**Algorithm 1** Greedy Information Gathering for Single MAV

$\Delta = 2\rho r, \rho \in (1/2, 1)$; $a = 2\Delta + r_d$;
**for** $round$ = 1 to $\frac{\pi D}{\sqrt{1-\rho^2}r}$ **do**
  $i = 1$, $k_i = \frac{2D}{a}$;
  **while** $k_i \geq 1$ **do**
    $D_i = D - k_i \cdot r_d$, $c_{r,i} = \frac{D_i}{\Delta}$;
    $Fly(H, c_{r,i})$, $Land(c_{r,i})$ and $GI(m(c_{r,i}))$;
    **for** $j = 1$ to $k_i - 1$ **do**
      $Fly(c_{r,i-j+1}, c_{r,i-j})$, $Land(c_{r,i-j})$ and generate information $GI(m(c_{r,i-j}))$;
    **end for**
    $Back$ for recharge and exchange information;
    $i = i + 1$, $k_i = k_{i-1} \cdot \frac{r_d}{a}$;
  **end while**
**end for**

In Alg. 1, $k_i$ means the number of points to reach for sensing during the $i$-th flight in each round. The main idea is the MAV reaches the farthest center $c_{r,i}$ and then flies to the nearby $k_i - 1$ centers for sensing along its way back to the Hive. Since each round can cover the whole area within $\delta = \frac{2\sqrt{1-\rho^2}r}{D}$, $\frac{\pi D}{\sqrt{1-\rho^2}r}$ rounds are enough to gather the sensing field's information. The relations between $k_i$ and $r_d$ are formulated as:

$$\begin{aligned} k_1 \cdot r_d + 2k_1 \cdot \Delta &= 2D \\ k_i \cdot r_d + 2k_i \cdot \Delta &= 2(D - \sum_{j=1}^{i-1} k_j \Delta) \end{aligned}$$

Thus all centers along the radius can be covered greedily. The number of flights in each round is the smallest value $i$ such that $k_i = \frac{2D}{a} \cdot (\frac{r_d}{a})^i < 1$:

- $r_d$ is comparable with $r$ ($r_d = \Theta(r)$); let $i = c' \log\frac{D}{r}$ where $c'$ is an appropriate constant, $k_i < 1$ can be satisfied; thus all information can be gathered with $\log\frac{D}{r}$-approximation when compared with the lower bound in Theorem 1.
- $r_d = o(r)$, $(\frac{r_d}{2\Delta + r_d})^i = (1 - \frac{2\Delta}{2\Delta + r_d})^i \leq 1 - \frac{2\Delta}{2\Delta + r_d} \cdot i$; we can find a constant $c'$ such that $1 - \frac{2\Delta}{2\Delta + r_d} \cdot c' < \frac{a}{2D}$ implies constant approximation to the lower bound.

*Theorem 2:* Algorithm 1 can gather all information within $T_G = O(\log\frac{D}{r} \cdot \frac{D^2}{rv_f})$ when $r_d = \Theta(r)$ and $T_G = O(\frac{D^2}{rv_f})$ when $r_d = o(r)$; $T_I = \Theta(\frac{D}{v_f})$.

We simulate this algorithm with different $r_d$ values and evaluate the gathering latency varying with $\rho \in (1/2, 1)$. In Fig. 2 ($D = 3000m$, $v_f = 2m/s$, $r = 20m$), when $r_d$ becomes larger, the gathering latency increases under the same $\rho$ value. For each $r_d$, the latency increases as $\rho$ grows. In Fig. 3, when $r_d$ is small (such as $r_d = 0.001$ or $r_d = 0.1$), the gathering latency will not increase too much as compared with the rough lower bound as shown. In practice, $r_d$ will not be comparable to the sensing range $r$. For simplicity, we omit the impact of $r_d$ in the following when the MAV can fly to more than one point for sensing.

*Corollary 1:* If we only utilize the mobility to sense and gather information with no wireless transmission, $T_I = \Omega(\frac{D}{v_f})$.

*Remark 4.1:* The memory needed in Alg. 1 is $O(\frac{D}{r})$. However, the lower bound in Theorem 1 will not change even if we have unlimited memory, which is constrained by the limited energy.

*Remark 4.2:* In view of the rapid development of battery recharging techniques, we do not take the recharging time into consideration in our algorithm design. Nevertheless, the greedy

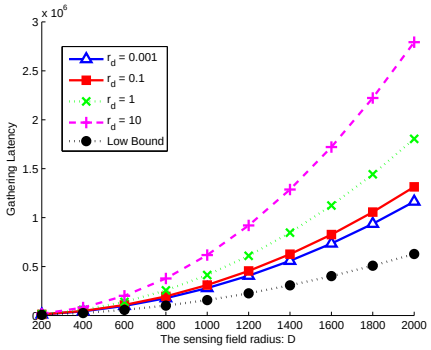



Fig. 3. Gathering Latency comparisons: $\rho = 0.6, D \in [200, 2000]$

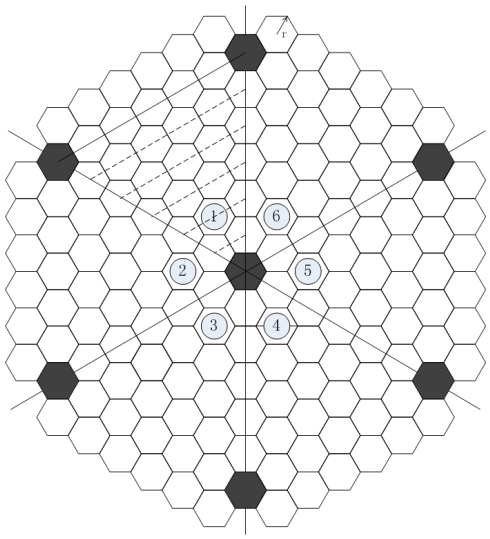



Fig. 4. Spatial Decomposition of the space

algorithm also yields a constant approximation ratio when the recharging time for each MAV is the same when $r_d = o(r)$.

## V. MASN Reduces the Latency

We have bounded the information latency as $\Omega(\frac{D}{v_f})$ when no communication exists between the MAVs. The flight velocity is about 1 to 5 meters per second, which is extremely slow as compared with the wireless transmission speed. We propose the MASN method for information gathering to reduce both gathering latency and information latency.

### A. Spatial Decomposition

The field $\mathcal{F}$ can be covered by regular hexagons with radius $r$ as in Fig. 4 (the Hive is in the center). Divide these hexagons into six regions with left side open and right side closed, as shown, and assume they are placed layer by layer. Take the $j$-th layer in Region 1 as an example, there are $j$ hexagons and the distance between two consecutive hexagon center is $\sqrt{3}r$.

*Lemma 5.1:* For each hexagon center $c$ in the $j$-th layer, the distance to the Hive ($dis_j(c, H)$) satisfies:

$$\frac{3r}{2} \cdot j \leq dis_j(c, H) \leq \sqrt{3}r \cdot j \tag{2}$$

*Proof:* Consider the triangle in Region 1 (c.f. Fig. 4), the length of the edges is equal to $l = \sqrt{3}r \cdot j$ due to the distance of two consecutive hexagon centers being the same. The largest distance from the Hive to the opposite edge is $d_{max} = l$ and the shortest one is the line from the Hive to the foot point with $d_{min} = l \cdot \sin\frac{\pi}{3} = \frac{3r}{2} \cdot j$. Thus, $\frac{3r}{2} \cdot j \leq dis_j(c, H) \leq \sqrt{3}r \cdot j$. ■

We bound the number of hexagons ($N_h$) needed to cover $\mathcal{F}$ by:

*Lemma 5.2:* $\frac{D^2}{r^2} \leq N_h \leq \frac{4D^2+6Dr}{3r^2}$

*Proof:* Each hexagon can be covered by a disk of radius $r$ and the area of the sensing field is $\pi D^2$, so $N_h \geq \frac{\pi D^2}{\pi r^2} = \frac{D^2}{r^2}$ hexagons are necessary.

As in Fig. 4, there are at most $j_{max} = \frac{D}{3r/2} = \frac{2D}{3r}$ layers from the distance relation in Lemma 5.1. $j$-th layer has $6j$ hexagons, thus $N_h \leq \sum_{i=1}^{j_{max}} 6j = \frac{4D^2+6Dr}{3r^2}$. ■

### B. Layer by Layer Gathering

We assume the number of MAVs is enough to cover all the hexagons, which means $N \geq N_h$. We present the Layer by Layer Gathering Algorithm based on the following notations:

- $h(i, j, k)$: The $i$-th region, $j$-th layer, $k$-th hexagon (clockwise); $(1 \leq i \leq 6, 1 \leq j \leq \frac{2D}{3r}, 1 \leq k \leq j)$
- $Ch(i, j, k)$: The center point of $h(i, j, k)$;
- $Dh(i, j, k)$: The distance from $Ch(i, j, k)$ to the Hive;
- $Ph(i, j, k)$: The parent sensor node for the MAV dispatched to $Ch(i, j, k)$;
- $DTh(i, j, k)$: The time delay to invoke transmission schedule after the arrival;

The Dispatching Schedule at the Hive is presented in Alg. 2 ($c$ is a constant in Line 3, to be defined later for different interference models).

**Algorithm 2** Dispatching Schedule at the Hive

1: $\delta_f = \frac{\sqrt{3}r}{v_f}$,$\delta_t = \frac{\sqrt{3}r}{v_t}$, $DT = 0$;
2: **for** $j = 1$ to $\frac{2D}{3r}$ **do**
3: $Dt_i = \max\{0, cj\delta_t - \delta_f/6\}$;
4: **for** $i = 1$ to $6$ **do**
5: $t(i, j) = \frac{\sqrt{3}r}{v_f} \cdot j + (i-1) \cdot \frac{\delta_f}{6}$;
6: **for** $k = 1$ to $j$ **do**
7: **if** $Dh(i, j, k) \leq D$ **then**
8: $t_{dp} = t(i, j) - \frac{Dh(i,j,k)}{v_f}$;
9: $Ph(i, j, k) = MAV(i, j-1, 2\lceil\frac{k}{2}\rceil - 1)$;
10: $DTh(i, j, k) = DT + (i-1)Dt_i$;
11: $Dispatch(MAV(i, j, k), Ch(i, j, k), t_{dp})$;
12: **end if**
13: **end for**
14: **end for**
15: $DT = DT + 6Dt_i$;
16: **end for**

**Algorithm 3** MAV Operations

*Dispatch(MAV,Ch(i,j,k),$t_{dp}$)*:

1: **When** $time = t$
2: $Fly(H, Ch(i,j,k))$;
3: **When** $time = t + \frac{Dh(i,j,k)}{v_f}$
4: $Land(Ch(i,j,k))$ and $GI(m(i,j,k))$;
5: **if** $DTh(i,j,k) + cj\delta_t < D/v_f$ **then**
6: Invoke Hop Delay Schedule;
7: **else**
8: $Back$ for carrying information to the Hive;
9: **end if**

**On** *RI($m$)*:

1: $SI(m, Ph(i,j,k))$ according to the Hop Delay Schedule;

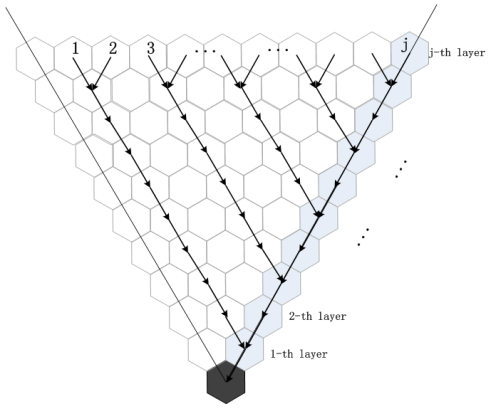



Fig. 5. Topology for $j$-th layer Information Gathering

The operations at $MAV(i,j,k)$ is presented in Alg. 3:

The main idea of the schedule at the Hive is dispatching $j$ MAVs to $j$-th layer of each Region and control them to arrive at the same time. (The arriving times for two consecutive regions differ by $\delta_f/6$.) We assume the sensed information by $j$ MAVs can be gathered in $cj\delta_t$ time and $Dt_i$ denotes the delay when the next region MAVs should begin transmitting. ($Dt_i = 0$ means the sensed information is gathered by the Hive before the next region's MAVs arrive.) The topology is constructed as in Fig. 5 and we present the Hop Delay Schedule (c.f. Alg. 4) to gather $j$ pieces of information from the same layer in $cj\delta_t$ time under different interference models.

*1) **Hop Distance Model**:* Suppose any two nodes within $h$-hops will interfere with each other.

**Algorithm 4** Hop Delay Schedule

1: $r = 1$, $\delta_t = \frac{\sqrt{3}r}{v_t}$;
2: **for** $k = 1$ to $j$ **do**
3: $MAV(i,j,k)$ executes $SI(m, Ph(i,j,k))$ at time $r \cdot \delta_t$;
4: $r = r + h + 1$;
5: **end for**
6: **When** The $MAV_i$ receives $m$
7: Execute $SI(m, Ph(i))$ after $\delta_t$;

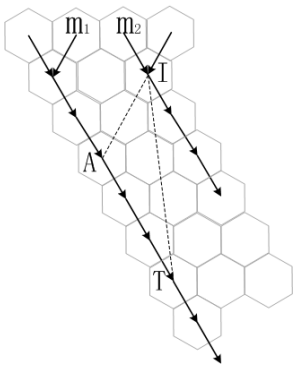



Fig. 6. Hop Delay in Protocol Model

*Lemma 5.3:* The Hop Delay Schedule can gather the information from $j$-th layer in $(h+2)j\delta_t$ time.

*Proof:* We can regard the scheduling operates in rounds where each round takes $\delta_t$ time. From the delay defined Line 4 in Alg. 4 , the $j$ pieces of information will begin transmitting in round $1, h+2, 2(h+1)+1, \cdots, (j-1)(h+1)+1$. The time used for transmitting for 1 hop is exactly $\delta_t$, and thus when each piece of information is delayed for $h+1$ rounds, the hop distance of any two consecutive transmissions will be at least $h$ and they can transmit successfully under the hop distance model. So we can gather all the information in $(j-1)\cdot(h+1)+1+j \le (h+2)j$ rounds and the time is bounded by $(h+2)j\delta_t$. ■

*2) **Protocol Model**:* The Protocol Model is more complicated. Without loss of generality, suppose $R_T \ge \sqrt{3}r$ and we can adjust the power such that the equality is satisfied (if $R_T < \sqrt{3}r$, reconstruct the decomposition such that the hexagon radius equals $\frac{R_T}{\sqrt{3}}$). The corresponding interference range $R_I \ge R_T = \sqrt{3}r$. Denote $R_I = \epsilon R_T$, where $\epsilon$ is a constant and $\epsilon \ge 1$.

*Lemma 5.4:* The information from the $j$ MAVs of the same region can be gathered at the Hive in $(\epsilon+3)j\delta_t$ time.

*Proof:* Modify the hop delay in Alg. 4 based on following calculation:

Suppose we delay each information piece by $d$ hops. Consider two consecutive information $m_1$ and $m_2$. If they share a common parent in the topology, $d = \epsilon + 2$ can make the transmissions of $m_1$ and $m_2$ successful during forwarding to the Hive. If they have two different parents as in Fig. 6 (the only case from the topology), hop delay should be large enough such that $TI > R_I$. From the Cosine theorem:

$$AT^2 + AI^2 - 2AT \cdot AI \cdot \cos\frac{2\pi}{3} = TI^2$$

Plug in $AI = 2\sqrt{3}r$, then $AT = (\epsilon-1)\sqrt{3}r$ is enough to satisfy $TI > R_I$. Thus, adding the first three hop delays, if each information piece is delayed by $d = \epsilon + 2$ hops, the correctness of Alg. 4 is guaranteed. Thus, the last piece of information can be gathered in $(j-1)\cdot(\epsilon+2)+1+j \le (\epsilon+3)j$ rounds which implies all information can be gathered in $(\epsilon+3)j\delta_t$ time. ■

*3) **The SINR Model**:* Based on the topology constructed, all link lengths are $r_l = \sqrt{3}r$. The maximum power used in one transmission slot is $P$ and the background interference is $I_b$. Define $R_m = (\frac{P}{\beta I_b})^{\frac{1}{\alpha}}$ be the maximum transmission range when no other MAVs are transmitting simultaneously. Generally $R_m > r$ and define $\lambda = 1 + (\frac{\beta(16\zeta(\alpha-1)+8\zeta(\alpha)-6)}{1-(r_l/R_m)^\alpha})^{\frac{1}{\alpha}}$.

*Lemma 5.5:* If the receiving ends of a set of links have mutual distances greater than $\lambda r_l$, they can be scheduled simultaneously.

*Proof:* Let $S$ be the set of links. Pick $(u, v) \in S$; divide all other links into sets $S_1, S_2, \cdots, S_\infty$ such that $S_i = \{(u', v') \in S | i\lambda r_l \le d(v, v') \le (i+1)\lambda r_l\}$. $|S_1| \le 18$ and $|S_i| \le 8(2i+1)$ from [2]. Thus, $d(u', v) \ge d(v, v') - d(u', v') \ge i\lambda r_l - r_l \ge i(\lambda - 1)r_l$. The interference caused by other transmitters is

$$\begin{aligned} I_t &= \sum_{(u',v') \in S \setminus \{(u,v)\}} \frac{P}{d(u', v)^\alpha} \\ &= \sum_{i=1}^{\infty} \sum_{(u',v') \in S_i} \frac{P}{d(u', v)^\alpha} \\ &\le \sum_{i=1}^{\infty} |S_i| \frac{P}{(i(\lambda - 1)r_l)^\alpha} \\ &\le \frac{P}{\beta} \frac{R_m^\alpha - r_l^\alpha}{R_m^\alpha r_l^\alpha} \end{aligned} \tag{3}$$

Plug in $I_b = \frac{P}{\beta R_m^\alpha}$, and we have $\frac{P/r_l^\alpha}{I_b + I_t} \ge \beta$, and so the lemma follows. ∎

Based on Lemma 5.5, the SINR Model can be transformed to the Protocol Model where $R_I = \lambda r_l$. Thus these information can be gathered to the Hive in $(\lambda + 3)j\delta_t$ time.

*Remark 5.1:* The constant $c$ in Line 3 of Alg. 4 is determined by the interference models: $c = h + 2$ in the $h$-Hop Distance Model, $c = \epsilon + 3$ in the Protocol Model and $c = \lambda + 3$ in the SINR Model.

*Theorem 3:* When $\frac{v_t}{v_f} = \Omega(\frac{D}{r})$, all information can be gathered through the MASN with $T_G = \Theta(\frac{D}{v_f})$ and information latency $T_I = \Theta(\frac{D}{v_t})$ under both graph based models and the SINR model.

*Proof:* From the Dispatching Schedule in Alg. 2, all information can be gathered in $j_{max} \cdot \frac{\sqrt{3}r}{v_f} + c \cdot j_{max}\delta_f \le \frac{2\sqrt{3}}{3} \cdot \frac{D}{v_f} + \frac{\sqrt{3}r}{v_f} = \Theta(\frac{D}{v_f})$ time while the information latency is $c \cdot j_{max} \frac{\sqrt{3}r}{v_t} = \Theta(\frac{D}{v_t})$ where $c$ is a constant based on the interference model chosen. ∎

*Remark 5.2:* The lower bound to gather $j$ pieces of information is $j\delta_t$, and thus Alg. 4 achieves constant approximation when compared with the lower bound. In addition, one unit of memory is enough to buffer the information during the scheduling process.

*Remark 5.3:* After the information from the $j$-th layer is gathered at the Hive, $\lceil \frac{j}{2} \rceil - 1$ MAVs (in each region) can fly back from the topology constructed, which improves the efficiency when performing some other tasks at the same time.

Table II compares the proposed algorithm, the mobile sinks method (fly back and forth for gathering) and the rough lower bound through the MASN when $v_t \gg v_f$ ($\frac{v_t}{v_f} = \Omega(\frac{D}{r})$). We simulate the three methods in Fig. 7 where $v_t = 3 \times 10^8 m/s$ for general wireless transmissions. The results show that our algorithm achieves lower gathering latency when compared with the mobile sinks method, and the information latency is reduced substantially which is a constant approximation with respect to the lower bound through MASN.

TABLE II
COMPARISON OF $G_L$ AND $G_I$

| **Methods** | $T_G$ | $T_I$ |
|---|---|---|
| Layer by Layer Gathering | $\frac{2\sqrt{3}}{3}\frac{D}{v_f} + \frac{\sqrt{3}r}{v_f}$ | $\frac{2\sqrt{3}}{3}c \cdot \frac{D}{v_t}$ |
| Mobile Sinks Method | $\frac{2D}{v_f}$ | $\frac{D}{v_f}$ |
| MASN Lower Bound | $\frac{D}{v_f} + \frac{D}{v_t}$ | $\frac{D}{v_t}$ |

If we extend this algorithm to some other applications such as underwater sensor networks where acoustic waves may be used, we get some interesting result. We simulate the case when $v_t = 340m/s$ in Fig. 8. The result shows that both gathering latency and information latency will increase as the sensing radius grows. Information latency will converge to the mobile sinks method and gathering latency will be slightly larger when the sensing radius is large enough due to the time delay for dispatching. In general, we claim our algorithm can achieve both low gathering latency and low information latency through the MASN.

## VI. FEWER MAVS, LARGER FIELD?

The algorithm described in the previous section can achieve both low gathering latency and low information latency. However, the number of MAVs used is as large as $N_h = \Theta(\frac{D^2}{r^2})$ and the sensing field is limited by the maximum distance the MAV can reach. In this section, we extend the *Hive-Drone* model to sense an arbitrarily large field with a small number of MAVs.

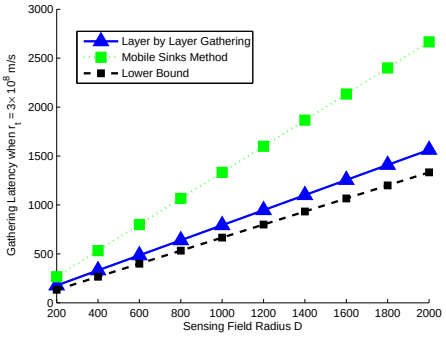
Given a sensing field $\mathcal{F}'$ where the diameter $D'$ is much larger than a single Hive's sensing radius $D$, multiple Hives (central stations) need to be placed in $\mathcal{F}'$. We describe the Controlled Path Gathering algorithm for each Hive where only $\frac{2D}{\sqrt{3}r} = \Theta(\frac{D}{r}) \ll N_h$ MAVs are used.

The sequential forwarding is based on the $h$ Hop Distance model and we give an example for illustration in Fig. 10.

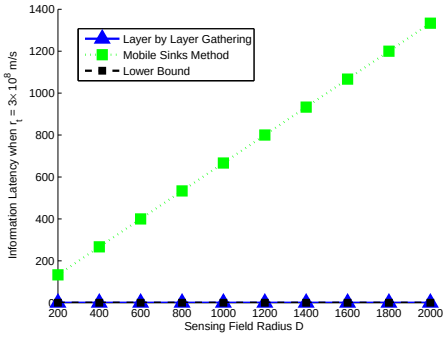
*Lemma 6.1:* The sequential forwarding method can gather the $N_l$ layers of information in $(h+1)N_l$ rounds where each round takes time $\delta_t = \frac{\sqrt{3}r}{v_t}$.

*Proof:* Each MAV in the $j$-th layer will transmit $N_l - j$ pieces of information from higher layers, and any messages transmitted in the same round have hop distance no less than $h$. Thus all $N_l$ information can be gathered in $(h+1)(N_l - 1) + 1 \le (h+1)N_l$ rounds and the time for each round is $\delta_t = \frac{\sqrt{3}r}{v_t}$. ∎

Moreover, one unit of memory is enough for each MAV to buffer the information and we can transform the protocol model and the SINR model to the hop distance model with

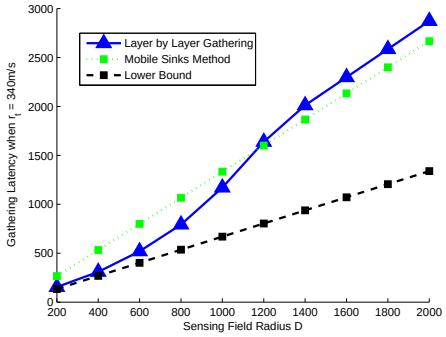



(a) Gathering Latency when $v_t = 3 \times 10^8 m/s$

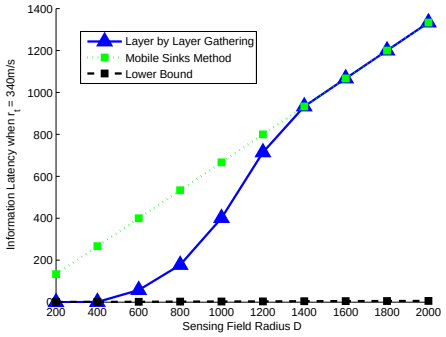



(b) Information Latency when $v_t = 3 \times 10^8 m/s$

Fig. 7. Latency Comparisons when $v_t = 3 \times 10^8 m/s, v_f = 1.5m/s$




(a) Gathering Latency when $v_t = 340m/s$




(b) Information Latency when $v_t = 340m/s$

Fig. 8. Latency Comparisons when $v_t = 340m/s, v_f = 1.5m/s$

**Algorithm 5** Controlled Path Gathering

S1. For one region, dispatch $N_l = \frac{D}{\sqrt{3}r}$ MAVs to the centers on the radius $L_1$ as Fig. 9;

S2. Gather information from $N_l$ layers by invoking *Sequential Forwarding*;

S3. MAV in the highest layer ($\frac{D}{\sqrt{3}r}$) goes back as the dotted line, others Fly to the higher layer center by one hop as the arrow in Fig. 9; a new MAV is dispatched along radius $L_2$ with a black point;

S4. Repeat S2 and S3 until all MAVs are on the radius $L_2$;

*Sequential Forwarding:*

1: **for** MAV in the $j$-th layer **do**
2: $r$ = j, $\delta_t = \frac{\sqrt{3}r}{v_t}$;
3: **for** $time$ = $j$ to $N_l$ **do**
4: $SI(m, Ph(MAV))$ at $r$-th round (after $r\delta_t$ time) ;
5: $r = r + (h + 1)$;
6: **end for**
7: **end for**

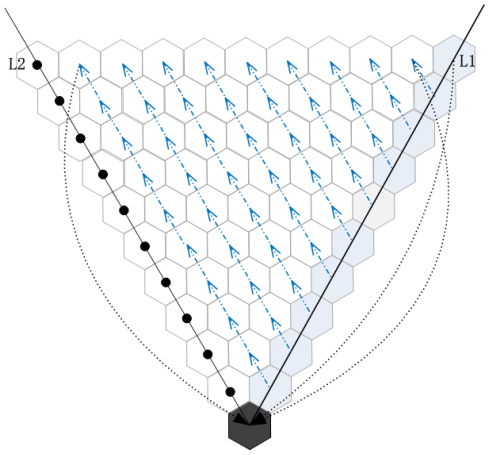



Fig. 9. Controlled Path Gathering

a constant factor difference. So Lemma 6.1 follows for both graph based models and the physical model (the SINR model).

Denote $Hex(H, D)$ as the hexagon where $H$ is the center and the radius is $D$.

*Theorem 4:* Alg. 5 can gather all information from $Hex(H, D)$ within $T_G = \Theta(\frac{D}{v_f} + \frac{D^2}{rv_t})$ and has low information latency $T_I = \Theta(\frac{D}{v_t})$.

*Proof:* Alg. 5 takes $\frac{D}{\sqrt{3}r}$ loops to gather the information of each region in $Hex(H, D)$ and the first loop will cost $\frac{D}{v_f}$

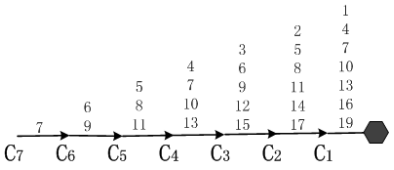



Fig. 10. Sequential Forwarding for 7 layers when $h = 2$

time for flying, while only $\delta_f = \frac{\sqrt{3}r}{v_f}$ time will be used during the other two consecutive loops. From Lemma 6.1, we can gather the information during each loop in $\frac{D}{\sqrt{3}r} \cdot (h+1)N_l \cdot \delta_t + \frac{D}{v_f} = \frac{D}{v_f} + \frac{(h+1)D^2}{\sqrt{3}rv_t} = \Theta(\frac{D}{v_f} + \frac{D^2}{rv_t})$ time. Refer to the 6 regions, all information can be gathered within $T_G = \Theta(\frac{D}{v_f} + \frac{D^2}{rv_t})$. Moreover, any information can arrive the Hive in $(h+1)N_l$ rounds where each round takes time $\frac{\sqrt{3}r}{v_t}$, and thus the information latency is $T_I = (h+1)N_l \cdot \delta_t = (h+1)\frac{D}{v_t} = \Theta(\frac{D}{v_t})$. ■

Any space can be fully covered by hexagons with the same radius. Similar to Fig. 4, an arbitrarily large field $\mathcal{F}'$ can be covered by $Hex(H_1, D), Hex(H_2, D), \cdots, Hex(H_n, D)$ such that $\mathcal{F}' \subseteq \bigcup_{i=1,2,\cdots,n} H_i$, where each Hive uses only $\Theta(\frac{D}{r})$ MAVs to gather information with a slight increase in gathering latency when $v_t$ is large. Thus, we can monitor an arbitrarily large area by placing multiple Hives and each Hive has $\Theta(\frac{D}{r})$ MAVs.

## VII. Conclusion

In this paper, we discuss how to gather information through a Mobile Aerial Sensor Network (MASN) based on the Hive-Drone model. This method can overcome the energy hole problem which can easily occur in stationary wireless sensor networks, and can reduce both the gathering latency and the information latency substantially as compared with the mobile sinks method. In addition, We present a greedy constant approximation algorithm for the single MAV case, which is similar to the mobile sinks method but with limited energy and bounded memory. For the case where there are enough MAVs ($\Theta(\frac{D^2}{r^2})$) to cover the entire monitored area, under three frequently used interference models, i.e., the hop distance model, the protocol model and the SINR model, we propose the Layer by Layer Gathering Algorithm through the MASN. This algorithm can achieve both low gathering latency and low information latency while using only one unit of memory to buffer information during transmission scheduling. For the case where there are only a small number of MAVs ($\Theta(\frac{D}{r})$), we describe an approach with only a slight increase in gathering latency and extend the single Hive to multiple Hives in order to monitor an arbitrarily large area. Our extensive simulation results clearly indicate that our methods can indeed achieve low information and gathering latencies while using only one unit of memory.

Note that all our proposed algorithms are centralized algorithms since we use the Hive-Drone model where the Hive is a central station. In the future, when we consider multiple Hives, it will be interesting to see whether we can design efficient distributed algorithms, such as for dispatching the MAVs (Drones) and scheduling the wireless transmissions between MAVs, with guaranteed low gathering and information latencies.

## VIII. Acknowledgments

The authors thank Prof. Thomas Moscibroda for pointing us to reference [5]. This work was supported in part by the National Basic Research Program of China Grant 2011CBA00300, 2011CBA00302, the National Natural Science Foundation of China Grant 61073174, 61103186, 61202360, 61033001, 61061130540, and Hong Kong RGC-GRF grants 714009 and 714311.

## References

[1] H. Balakrishnan, C. L. Barrett, V. S. Anil Kummar, M. V. Marathe, and S. Thite. The distance-2 matching problem and its relationship to the MAC-layer capacity of ad hoc wireless networks. *IEEE Journal on Selected Areas in Communications*, 22(6):1069-1079, 2004

[2] P. Bateman, and P. Erdös. Geometrical extrema suggested by a lemma of besicovitch. *The American Mathematical Monthly*, 306-314, May, 1951.

[3] A. Chakrabarti, A. Sabharwal, and B. Aazhang. Using predictable observer mobility for power efficient design of sensor networks. In *IPSN*, 2003.

[4] S. Chen, S. Tang, M. Huang, and Y. Wang. Capacity of Data Collection in Arbitrary Wireless Sensor Networks. In *INFOCOM*, 2010.

[5] K. Dantu, B. Kate, J. Waterman, P. Bailis, and M. Welsh. Programming Micro-Aerial Vehicle Swarms With Karma. In *SenSys*, 2011.

[6] M. D. Francesco, S. K. Das, and G. Anastasi. Data Collection in Wireless Sensor Networks with Mobile Elements: A Survey. *TOSN* 8(1): 7, 2011.

[7] J. K. Hart, and K. Martinez. Environmental Sensor Networks:A revolution in the earth system science? *Earth-Science Reviews*, 78. pp. 177-191, 2006.

[8] L. He, J. Pan, and J. Xu. Reducing Data Collection Latency in Wireless Sensor Networks with Mobile Elements. In *WiSARN*, 2011.

[9] T. T. Lai, W. J. Chen, K. H. Li, P. Huang, and H. H. Chu. TriopusNet: Automating Wireless Sensor Network Deployment and Replacement in Pipeline Monitoring. In *IPSN*, 2012.

[10] B. Liu, D. Towsley, and O. Dousse. Mobility improves coverage of sensor networks. In *MobiHoc*, 2005.

[11] T. Moscibroda and R. Wattenhofer. The complexity of connectivity in wireless networks. In *INFOCOM*, 2006.

[12] M. Quaritsch, K. Kruggl, D. Wischounig-Strucl, S. Bhattacharya, M. Shah, and B. Rinner. Networked UAVs as aerial sensor network for disaster management applications. *Elektrotechnik und Informationstechnik*, 2010.

[13] R. Shah, S. Roy, S. Jain, and W. Brunette. Data MULEs: modeling a three-tire architecture for sparse sensor network. In *SNPA*, 2003.

[14] X. Ta, G. Mao, and B. Anderson, Evaluation of the probability of k-hop connection in homogeneous wireless sensor networks. In *GLOBECOM*, 2007.

[15] I. Vasilescu, K. Kotay, D. Rus, M. Dunbabin, and P. Corke. Data collection, storage, and retrieval with an underwater sensor network. In *SenSys*, 2005.

[16] P.-J. Wan, L. Wang, and O. Frieder, Fast Group Communications in Multihop Wireless Networks Subject to Physical Interference, *MASS*, 2009.

[17] P.-J. Wan, Z. Wang, Z. Wan, S. C.-H. Huang, and H. Liu. Minimum-Latency Schedulings for Group Communications in Multi-channel Multihop Wireless Networks. In *WASA*, 2009.

[18] G. Werner-Allen, K. Lorincz, M. Welsh, O. Marcillo, J. Johnson, M. Ruiz, and J. Lees. Deploying a Wireless Sensor Network on an Active Volcano, *IEEE Internet Computing*, 10(2): 18-25, 2006.

[19] X. Wu, G. Chen, and S. Das. Avoiding Energy Holes in Wireless Sensor Networks with Nonuniform Node Distribution. *IEEE Transactions on Parallel and Distributed Systems*, 19(5):710-720, 2008.

[20] X. Xu, J. Luo, and Q. Zhang. Delay Tolerant Event Collections in Sensor Networks with Mobile Sink. In *INFOCOM*,2010.