This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2016.2603152, IEEE Transactions on Mobile Computing

Near Optimal Data Gathering in Rechargeable Sensor Networks with a Mobile Sink

Yongmin Zhang, *Member, IEEE,* Shibo He, *Member, IEEE,* Jiming Chen, *Senior Member, IEEE*

Abstract—We study data gathering problem in Rechargeable Sensor Networks (RSNs) with a mobile sink, where rechargeable sensors are deployed into a region of interest to monitor the environment and a mobile sink travels along a pre-defined path to collect data from sensors periodically. In such RSNs, the optimal data gathering is challenging because the required energy consumption for data transmission changes with the movement of the mobile sink and the available energy is time-varying. In this paper, we formulate data gathering problem as a network utility maximization problem, which aims at maximizing the total amount of data collected by the mobile sink while maintaining the fairness of network. Since the instantaneous optimal data gathering scheme changes with time, in order to obtain the globally optimal solution, we first transform the primal problem into an approximate network utility maximization problem by shifting the energy consumption conservation and analyzing necessary conditions for the optimal solution. As a result, each sensor needs not estimate the amount of harvested energy and the problem dimension is reduced. Then, we propose a Distributed Data Gathering Approach (DDGA), which can be operated distributively by sensors, to obtain the optimal data gathering scheme. Extensive simulations are performed to demonstrate the efficiency of the proposed algorithm.

Index Terms—Rechargeable sensor networks, mobile sink, data gathering, network utility maximization, Distributed algorithm.

1 Introduction

DATA gathering is one of the most challenging issues in Wireless Sensor Networks (WSNs) since it relies on multiple factors, such as energy constraints of sensors, network topology, links conditions, routing scheme, transmission scheduling and so on. In battery-powered WSNs with fixed sink node(s), sensors near sink node(s) consume a large portion of energy for relaying data for other sensors, resulting in a critical issue of short network lifetime. To address this issue, the approach of mobile sink(s) has been introduced to deal with inadequate network resources and unbalanced traffic distribution in [1]–[4], which gives rise to the data gathering optimization problem in battery-powered WSNs with mobile sink(s) [5]–[10]. However, most of these works focus on the network lifetime maximization by minimizing the energy consumption for data gathering due to limited energy of sensors.

As battery-powered sensors are not desirable for long-term applications, energy harvesting technologies have been recently introduced to power sensors, thus enabling the perpetual operation [11]–[13]. In such networks, each sensor can harvest energy from the surrounding environment and store it in the battery for further use, thus preventing sensor from running out of energy. Due to the time-varying characteristics of energy harvesting sources, such as solar and wind, the optimal data gathering in RSNs is much more challenging than that in battery-powered WSNs and has been studied in [14]–[17]. However, most of these works just considered the data gathering in RSNs with fixed sink node(s) and are not suitable for RSNs with mobile sink(s). Since the total amount of harvested energy during a given period is limited, the network performance will be declined due to inadequate network resources and unbalanced traffic distribution. Thus, to further optimize the network performance for perpetual RSNs, the approach of mobile sink(s) can be introduced to RSNs.

Y. Zhang, S. He, and J. Chen are with State Key Lab. of Industrial Control Technology, Department of Control Science and Engineering, Zhejiang University, Hangzhou, 310027, China. (Email: ymzhang@zju.edu.cn, shibohe@ieee.org, jmchen@iipc.zju.edu.cn).

In such RSNs, rechargeable sensors are deployed into a region of interest for environmental monitoring and a mobile sink travels along a pre-defined path to collect data from sensors periodically[1]. In [18], [19], Ren *et al.* considered the data collection problem with one-hop data transmission for RSNs with a mobile sink, and proposed an offline algorithm as well as an online distributed solution to maximize data collection. However, they only consider one-hop data transmission scheme, which may lead to limited network performance. Typically, sensors can transmit data to the mobile sink either directly or in multi-hop manners. Thus, we explored the optimal data transmission problem with multi-hop data transmission to improve the network performance in this paper.

As we know, the required transmission energy consumption increases with the increase of transmission distance [20]. Due to the movement of the mobile sink, the required transmission energy consumption for each sensor to transmit one unit data to the mobile sink will change and this affects the routing selection and the total amount of transmitted data. In addition, due to the stochastic nature of harvested energy, the total amount of available energy will change with time, making the problem even more complicated. In order to maximize network performance, the sensors need to jointly decide the total amount of data that should be sent out, routing path and transmission energy consumption, as well as the total amount of energy use.

There are two challenges to address in the data gathering optimization problem in such RSNs. First, *in order to obtain the globally optimal solution, each sensor needs to estimate the total amount of energy that it can harvested during each period with high accuracy*. Secondly, *it is difficult for the mobile sink or one sensor to know all the information about other sensors (including*

1. In this paper, we do not consider the path optimization, since the path may be a special path, such as road, wirerope, pipeline, which could not be changed.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2016.2603152, IEEE Transactions on Mobile Computing

position, total amount of harvested energy and so on) and the mobile sink in large-scale RSNs. Therefore, the problem should be solved in a distributed way. To address these two challenges, we firstly shift the energy consumption conservation, in which each sensor only consumes the energy that it has harvested before. In such a way, each sensor needs not estimate the amount of energy in the harvesting process. Secondly, we analyze necessary conditions of optimality and transform the network utility maximization problem into an approximate network utility maximization problem, which is shown to be a convex optimization problem, then we propose a Distributed Data Gathering Approach (DDGA) to solve the data gathering optimization problem in a distributed way. In such a way, each sensor only needs to cooperate with its neighbours to distributively calculate the globally optimal solution. Specifically, our contributions are summarized as follows:

- We transform the network utility maximization problem into an approximate convex optimization problem by shifting the energy consumption conservation and analyzing the necessary conditions of optimality to avoid estimation of harvested energy, and thus simplify the problem.
- We design a Distributed Data Gathering Approach (D-DGA) to obtain the near optimal solution for data gathering in a distributed manner. More importantly, it is unnecessary for any sensor/mobile sink to know the global information related to the entire network.
- Extensive simulations based on real experimental data are performed to demonstrate the efficiency of the proposed algorithms. Simulation results show that sensors obtained by DDGA can maximize the total amount of data while maintaining the fairness of network.

The rest of this paper is organized as follows. We introduce the related work in Sec. 2 and describe the network model and problem formulation in Sec. 3. We transform the network utility maximization problem into an approximate network utility maximization problem by shifting the energy consumption conservation and analyzing the necessary conditions for the optimal solution in Sec.4. Then, we prove the convexity of transformed problem in Sec.4.3. A distributed data gathering approach is proposed to obtain the globally optimal solution in Sec. 5. Simulation results are given to demonstrate the performance of the proposed algorithms in Sec. 6. We conclude this work in Sec. 7.

2 Related works

Recently, since mobile sink(s) can deal with inadequate network resources and unbalanced traffic distribution in WSNs, great research interest has been drawn on optimal data gathering in WSNs with mobile sink(s). Most of these works focus on the network lifetime maximization by minimizing the energy consumption for data gathering. In [6], Chakrabarti *et al.* modeled the process of data collection in WSNs with a mobile observer as a queue with deadlines, and then discovered a constraint on minimum sensor separation that guarantees on data loss. In [7], Gao *et al.* proposed a novel data collection scheme, called the Maximum Amount Shortest Path (MASP), that increases network throughput and conserves energy as well by optimizing the assignment of sensors. In [8], Guo *et al.* proposed a data gathering cost minimization (DaGCM) framework with concurrent data uploading and then presented a distributed algorithm composed of cross-layer data control, routing, power control and compatibility determination sub-algorithms with explicit message passing. However, since the total amount of energy in each sensor's battery is limited and given, most of the works focus on the network lifetime maximization problem in battery-powered sensor networks rather than maximizing the network performance. In addition, the initial energy levels of battery-powered sensors are given while the energy level of rechargeable sensors are uncertain. The data gathering problem in battery-powered WSNs can be solved in offline or centralized manners. However, the data gathering problem in rechargeable WSNs should be designed in online and distributed manners. Hence, most of data gathering schemes in battery-powered WSNs with a mobile sink do not work in RSNs with a mobile sink.

For the RSNs with fixed sink node(s), there are many works on the data gathering optimization. In [14], Lin *et al.* studied the optimal data gathering problem and designed routing algorithms to optimally utilize the available energy. In [16], Liu *et al.* formulated the optimal data gathering problem as a network utility maximization problem, and developed a QuickFix algorithm to compute the optimal sampling rate and route, and a SnapIt algorithm to adjust the sampling rate. In [21], Chen *et al.* considered the network utility maximization problem in RSNs and developed a joint energy allocation and routing algorithm, which is a low-complexity online solution. In [22], Zhang *et al.* focus on the network utility maximization problem and proposed a distributed algorithm to jointly optimize energy management, data sensing and routing for sensors. In [23], Joseph *et al.* designed a joint optimal power control, routing and scheduling algorithm to ensure that the network resources can be fairly utilized. In [24], Zhao *et al.* considered the network utility maximization in RSNs and proposed a distributed algorithm to adjust data rates, link scheduling and routing according to the up-to-date energy replenishing states of the sensors. However, these works on data gathering in RSNs with fixed sink node(s) are not suitable for RSNs with mobile sink(s) since the movement of mobile sink brings dynamically changes the network topology, which affects the optimal routing selection.

For the RSNs with mobile sink(s), there are a few works on the data gathering optimization with controllable harvested energy. Guo *et al.* considered the data gather optimization problem by joint optimizing mobile data gathering and energy provisioning [24], [25]. Zhao *et al.* proposed a framework for jointly optimizing mobile energy replenishment and data gathering [26]. However, the harvested energy for each sensor in these works can be estimated by the mobile sink. Thus the proposed data gathering schemes are not applicable for the RSNs in this paper. There are few works on the data gathering optimization with uncontrollable harvested energy. Ren *et al.* considered the data collection problem with one-hop data transmission for RSNs with a mobile sink, and proposed an offline algorithm as well as an online distributed solution to maximize data collection [18], [19]. However, they only considered one-hop data transmission, which may lead to limited network performance. Thus we consider the data collection problem with multi-hop data transmission scheme to improve the data gathering performance.

3 System model and Problem statement

We consider an RSN with a mobile sink, in which there are N rechargeable sensors deployed into an area to sense their surrounding environment and one mobile sink traveling along a pre-defined path at a constant speed v without stops to collect

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2016.2603152, IEEE Transactions on Mobile Computing

TABLE 1
Notation definitions

Symbol	Definition
$\rho_i(t)$	The energy collected by sensor i at slot t
$\overline{\rho}_i(t)$	The upper bound of $\rho_i(t)$
B_i^{ini}	The initial battery level for sensor i
$B_i(t)$	The battery level for sensor i at slot t
B_i^{cap}	The battery capacity for sensor i
$d_{ij}(t)$	The Euclidean distance between sensor i and its next-hop j at slot t
d_{ij}^{max}	The maximal Euclidean distance between sensor i and its next-hop j
$E_{ij}^{th}(t)$	The threshold for transmission energy consumption for sensor i to transmit one unit data to its next-hop j directly at slot t
E_{ij}^{tr}	The transmission energy consumption for sensor i to transmit one unit data to its next-hop j directly
E_i^{re}	The receiving energy consumption for sensor i to receive one unit data
E_i^{sn}	The sensing energy consumption for sensor i in one period
$P_i(t)$	The total energy consumption for senor node i at slot t

data periodically. Due to limited data preprocessing and analysis ability of sensor, all the sensors need to transmit their sensory data to the mobile sink directly or in multi-hop manners for further computation and analysis. Each sensor senses its surrounding environment periodically with a constant energy consumption E_i^{sn}, and the total amount of sensory data is large enough data to be sent out. Hence, the total amount of data sent out by one sensor indicates the monitoring performance, and more data reported by the sensor means better monitoring performance. Furthermore, each sensor can change its transmission range by adjusting its transmitting power, as well as the data transmission path. For simplicity, we assume that the rechargeable sensors are powered by solar cell and rechargeable batteries and the mobile sink travels along a straight path in this paper.

3.1 Energy consumption model

Generally, the mobile sink travels along the pre-defined straight path for one time to collect data from all the sensors periodically. Since the cycle of solar energy harvesting process is one day, one day is one period in this paper. Let K denote the K-th period. which can be divided into T slots. Let t denote the t-th slot, and $t = 1, 2, \cdots, T$. Let j denote a sensor or the mobile sink, i or k denote a sensor, respectively. A summary of notation related to energy consumption model is given in Table 1.

Assume that, the battery capacity B_i^{cap} is large enough to reserve all the harvested energy $\rho_i(t)$, such as $B_i^{cap} \geq B_i^{ini} + \sum_{t=1}^{T} \overline{\rho}_i(t)$ where $\overline{\rho}_i(t)$ is the upper bound of $\rho_i(t)$. Thus, the battery level of sensor i at slot $t+1$ can be given by

$$B_i(t+1) = [B_i(t) + \rho_i(t) - P_i(t)]^+, \quad (1)$$

where $B_i(1) = B_i^{ini}$ and $[\cdot]^+ = \max(\cdot, 0)$.

Since the mobile sink travels along the road at a constant speed v periodically, sensor i can estimate the location of the mobile sink, as well as the Euclidean distance $d_{is}(t)$ between them at any slot t. If sensor i needs to transmit one unit data to its next-hop j at slot t, the threshold for the transmission energy consumption $E_{ij}^{th}(t)$ can be given by

$$E_{ij}^{th}(t) = \beta + \mu\big(d_{ij}(t)\big)^{\alpha}, \quad (2)$$

where β is a distance-independent term, μ is a distance-dependent term, α is the path-loss exponent ($2 \leq \alpha \leq 4$ for the free-space and short-to-medium-range radio communication) and $d_{ij}(t)$ is the Euclidean distance between sensor i and its next-hop j at slot t [20], [27].

From Eq. (2), it can be found that $E_{ij}^{th}(t)$ is an increasing and convex function of $d_{ij}(t)$. For transmission energy consumption E_{ij}^{tr}, we assume that it satisfies the following properties:

- The transmission energy consumption E_{ij}^{tr} in one period is a constant;
- If sensor i need to transmit one unit data to its next-hop j directly at slot t, the transmission energy consumption E_{ij}^{tr} should satisfy $E_{ij}^{tr} \geq E_{ij}^{th}(t)$;

Since the Euclidean distance between sensor i and sensor k is fixed, i.e., $d_{ik}(t) = d_{ik}(t')$, $t' = 1, 2, \cdots, T$, the threshold for the transmission energy consumption E_{ik}^{th} is a constant, which can be given by

$$E_{ik}^{th} = \beta + \mu\big(d_{ik}(t)\big)^{\alpha}. \quad (3)$$

Due to the movement of the mobile sink, the Euclidean distance $d_{is}(t)$ between sensor i and mobile sink s will change with slot t, thus the threshold for transmission energy consumption $E_{is}^{th}(t)$ at different slots may be different, i.e.,

$$E_{is}^{th}(t) = \beta + \mu\big(d_{is}(t)\big)^{\alpha}. \quad (4)$$

Hence, for a given transmission energy consumption E_{is}^{tr}, we have the following result:
Let $\underline{t}_i$ denote the earliest slot, in and after which sensor i can transmit data to mobile sink s, and $\overline{t}_i$ denote the latest slot, after which sensor i can not transmit data to mobile sink s.

Lemma 1. For any given transmission energy consumption E_{is}^{tr} in one period, sensor i can only transmit data to mobile sink s at slot t', $t' \in [\underline{t}_i, \overline{t}_i]$.

Proof: For a given transmission energy consumption E_{is}^{tr}, the maximal Euclidean distance d_{is}^{max} can be calculated by

$$d_{is}^{max} = \Big(\frac{E_{is}^{tr} - \beta}{\mu}\Big)^{\frac{1}{\alpha}}. \quad (5)$$

Taking sensor i as the center and maximal Euclidean distance d_{is}^{max} as the radius to draw a circle, it can be found that there are at most two crossover points between the circle and the straight road [See Fig.1]. Denoted corresponding slots of crossover points by $\underline{t}_i$ and $\overline{t}_i$, respectively. For the values of $\underline{t}_i$ and $\overline{t}_i$, there are three possible cases:

1) $\underline{t}_i = \overline{t}_i = \emptyset$ when $d_{is}^{max} < h_i$;
2) $\underline{t}_i = \overline{t}_i$ when $d_{is}^{max} = h_i$;
3) $\underline{t}_i < \overline{t}_i$ when $d_{is}^{max} > h_i$

where $\emptyset$ is a null set and h_i denotes the vertical distance between sensor i and the straight road. Since sensor i can only send data to mobile sink s directly when $d_{is}^{max} \geq h_i$, we only consider the transmission energy consumption d_{is}^{max}, satisfying $d_{is}^{max} \geq h_i$, in the following paper.

Let $(y(t), 0)$ denote the position of mobile sink s at slot t and (y_i^0, h_i) denote the position of sensor node i, where $y(t)$ is an increasing function of slot t, y_i^0 is a constant and $y(1) < y_i^0 < y(T)$. The Euclidean distance $d_{is}(t)$ between mobile sink s and sensor i can be given by

$$d_{is}(t) = \sqrt{(y(t) - y_i^0)^2 + (0 - h_i)^2} \quad (6)$$

Based on the above equation, it can be found that $d_{is}(t)$ is a decreasing function of t when $y(1) \leq y_i^0$ and an increasing function of t when $y_i^0 \leq y(T)$. Hence, we can derive that the Euclidean distance $d_{is}(t') > d_{is}^{max}$ when $t' < \underline{t}_i$ or $t' > \overline{t}_i$ and $d_{is}(t') \leq d_{is}^{max}$ when $t' \in [\underline{t}_i, \overline{t}_i]$. If sensor i needs to transmit data to mobile sink s at slot t', the Euclidean distance $d_{is}(t')$ should satisfy

$$d_{is}(t') \leq d_{is}^{max}, \tag{7}$$

Thus, sensor i can only transmit data to mobile sink s at slot t', $t' \in [\underline{t}_i, \overline{t}_i]$. □

Let $\theta_i(E_{is}^{tr})$ denote the maximal number of slots between $\underline{t}_i$ and $\overline{t}_i$. Thus we have

$$\begin{aligned}\theta_i(E_{is}^{tr}) &= \overline{t}_i - \underline{t}_i + 1 \\ &= \frac{2\sqrt{(d_{is}^{max})^2 - h_i^2}}{v},\end{aligned} \tag{8}$$

where h_i denotes the vertical distance between sensor i and the straight road and v denotes the speed of mobile sink. The relationship between $\theta_i(E_{is}^{tr})$ and d_{is}^{max} can be illustrated in Fig. 1.

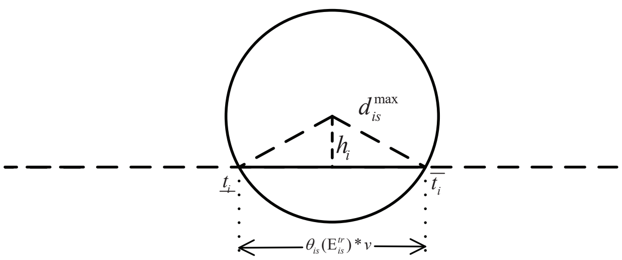


Fig. 1. Relationship between $\theta_{is}(E_{is}^{tr})$ and d_{is}^{max}.

Let E_i^{re} denote receiving energy consumption for sensor i to receive one unit data from other sensors, which can be arranged at a constant. For simplicity, we assume that the signal interference between sensors can be eliminated by underlying MAC layer (e.g., by TDMA or FDMA mechanism by [28]). Note that, we only consider the energy consumption for sensing, transmitting and receiving data in this paper since they dominate the energy consumption [7], [14].

3.2 Data transmission model

Since some of sensors are far away from the road, it will be energy-costly for these sensors to transmit data to the mobile sink directly. Hence, some sensors close to the road can serve as relay nodes to forward sensory data for other sensors to improve the network performance of the entire network.

In this paper, the network topology can be established according to their vertical distance h_i and h_k and the Euclidean distance d_{ik} between sensors i and k, where h_i (h_k) denotes the distance between sensor i (k) and the road and d_{ik} denotes the distance between sensor i and sensor k. If $h_i > h_k$ and $h_i \geq d_{ik}$, $k \in O(i)$; otherwise, $k \notin O(i)$; where $k \in O(i)$ denotes that logical link (i,k) can be established to transmit data and sensor k can serves as the next hop to forward sensory data for sensor i to mobile sink s. For mobile sink s, h_s always equals to 0, i.e., $h_s = 0$, thus, $s \in O(i)$ always holds.

Let $f_{ij}(t)$ denote the total amount of sensory data transmitted from sensor i to its next-hop j through the logical link (i,j) at slot t. Typically, each logical link has an upper bound of data transmission, such as link capacity. For simplicity, we assume that all the logical links have the same link capacity, denoted by C. Note that, our algorithm is also applicable to the network with different link capacities. For a given transmission energy consumption E_{ij}^{tr}, $f_{ij}(t)$ can be given by

$$\begin{cases} f_{ij}(t) \in [0, C], & \text{if } E_{ij}^{tr} \geq E_{ij}^{th}(t), \\ f_{ij}(t) = 0, & \text{otherwise.} \end{cases} \tag{9}$$

Since all the data received by sensor i should be transmitted to next hop immediately, all the sensors should satisfy the flow conservation, i.e.,

$$\sum_{j:j\in O(i)} f_{ij}(t) = \sum_{k:i\in O(k)} f_{ki}(t) + x_i(t), \tag{10}$$

where $x_i(t)$ denotes the total amount of sensory data that sensed and sent out by sensor i at slot t. The total energy consumption for sensor i at slot t can be given by

$$P_i(t) = \sum_{j:j\in O(i)} f_{ij}(t)E_{ij}^{tr} + E_i^{re} \sum_{k:i\in O(k)} f_{ki}(t) + \frac{E_i^{sn}}{T}. \tag{11}$$

The total energy consumption for sensor i at slot t could not be larger than the sum of the energy reserved in the battery and energy collected by sensor i at slot t, i.e.,

$$P_i(t) \leq B_i(t) + \rho_i(t). \tag{12}$$

Based on the Eqs. (1) and (12), we have

$$P_i(t') \leq B_i^{ini} + \sum_{t=1}^{t'} \rho_i(t) - \sum_{t=1}^{t'-1} P_i(t), \tag{13}$$

which can be rewritten as

$$\sum_{t=1}^{t'} P_i(t) \leq B_i^{ini} + \sum_{t=1}^{t'} \rho_i(t), \tag{14}$$

where $t' = 1, 2, \cdots, T$. From Eq. (14), it can be found that, in any slot t' during one period, the total amount of energy consumption for sensor i should be smaller than the sum of initial energy and the energy that it can collect.

In order to ensure that each sensor can operate perpetually, the total amount of energy consumed by sensor i during one period could not exceed the energy that it can collect, i.e.,

$$\sum_{t=1}^{T} P_i(t) \leq \sum_{t=1}^{T} \rho_i(t). \tag{15}$$

3.3 Utility maximization problem

Let X_i denote the total amount of sensory data sensed and sent out by sensor i and received by the mobile sink during K-th period, i.e., $X_i = \sum_{t=(K-1)T+1}^{KT} x_i(t)$. Let $U(X_i)$ denote the network utility function for sensor i, which is a continuously differentiable, increasing and strictly concave function of X_i. In this paper, we set $U(X_i) = \log(X_i)$, since it can be used to guarantee the fairness of network [29]. Based on the energy consumption model and data transmission model, we formulate the network utility maximization problem for one period, which can be given by

$$\textbf{P0:} \max_{E_{ij}^{tr}, x_i(t), f_{ij}(t)} \sum_i U(X_i) \tag{16}$$

$$s.t. \sum_{j:j\in O(i)} f_{ij}(t) = \sum_{k:i\in O(k)} f_{ki}(t) + x_i(t), \quad \forall i, t \tag{17}$$

$$\sum_{t=1}^{t'} P_i(t) \leq B_i^{ini} + \sum_{t=1}^{t'} \rho_i(t'), \quad \forall i, t' \tag{18}$$

$$\sum_{t=1}^{T} P_i(t) \leq \sum_{t=1}^{T} \rho_i(t), \quad \forall i, t \tag{19}$$

$$0 \leq f_{ij}(t) \leq C, \quad \forall i, j, t \tag{20}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2016.2603152, IEEE Transactions on Mobile Computing

where $t \in [(K-1)T+1, KT]$. Constraint (17) ensures that all the sensors should transmit the data to their next hop successfully and immediately. Constraints (18) and (19) enforce that the total energy consumption is reachable and permanent. Constraint (20) shows the available range for $f_{ij}(t)$. In this problem, the variables are coupled by the slot t since $\rho_i(t)$ will change with slot t.

From the above problem formulation, it is easy to find that utility maximization problem is a dynamic programming problem. Furthermore, due to mobility of mobile sink, required transmission energy consumption for sensor to transmit data to mobile sink will change with time, thus the instantaneous optimal data gathering scheme will change with time. In order to obtain the globally optimal solution, we first transform the primal problem into an approximate network utility maximization problem and then design a Distributed Data Gathering Approach (DDGA), which can be operated by sensor, to obtain the globally optimal solution for data gathering.

4 Problem Transformation

Since the objective function only depends on the total amount of data sensed and sent out by each sensor during the entire period rather than that in each slot, we can firstly solve the network utility maximization problem for the entire period, and then propose an available scheduling scheme to manage the data gathering in each slot.

Let $F_{ij} = \sum_{t=(K-1)T+1}^{KT} f_{ij}(t)$ denote the total amount of data transmitted through link (i, j) and $\overline{P}_i = \sum_{t=(K-1)T+1}^{KT} P_i(t)$ denote the total amount of energy consumption for sensor i during entire period, respectively. The value of $\overline{P}_i$ can be given by

$$\overline{P}_i = \sum_{j:j\in O(i)} E_{ij}^{tr} F_{ij} + E_i^{re} \sum_{k:i\in O(k)} F_{ki} + E_i^{sn} T \tag{21}$$

According to Eq.(18) in problem **P0**, all the sensors need to satisfy the energy consumption in every slot, thus each sensor needs to estimate the total amount of harvested energy during the period with high accuracy to solve the dynamic programming problem. Furthermore, since it is difficult for one sensor or mobile sink to know the information about all the sensors in large-scale RSNs, centralized algorithm could not be implemented. Thus, we firstly shift the energy consumption conservation, in which each sensor only consumes the energy that it has harvested before. In such a way, each sensor needs not estimate harvesting process. Secondly, we analyze necessary conditions for the globally optimal solutions to transform the primal problem into an approximate network utility maximization problem. Then, we prove that the transformed problem is a convex optimization problem, which can be solved in a distributed way.

4.1 Shift of energy consumption conservation

There exists an implicit assumption in Eq. (19) that each sensor needs to estimate the harvested energy during the entire period with high accuracy. However, this may be difficult to realized without enough additional information. To address this issue, we shift the energy consumption conservation as follows:

$$\overline{P}_i \le \mathbb{E}_i(K), \tag{22}$$

where K, $K \ge 1$, denotes the K-th period, and

$$\mathbb{E}_i(K) = \begin{cases} B_i^{ini} + \sum_{t=1}^{t'} \rho_i(t), & \text{if } K = 1 \\ \sum_{t=(K-2)T+t'+1}^{(K-1)T+t'} \rho_i(t), & \text{if } K \ge 2 \end{cases} \tag{23}$$

where t' denotes the slot t', before which the sensor does not transmit data to other sensors or mobile sink. In such a way, each sensor needs not estimate the total amount of harvested energy. Instead, each sensor only needs to decide the values of t', which is easier to realize. In this paper, we set $t' = 1$, which means that all sensors use the energy harvested during the last period.

The optimization problem for K-th period can be given by

$$\textbf{P1:} \max_{E_{ij}^{tr}, X_i, F_{ij}} \sum_i U(X_i) \tag{24}$$

$$s.t. \sum_{j:j\in O(i)} F_{ij} = \sum_{k:i\in O(k)} F_{ki} + X_i, \quad \forall i, \tag{25}$$

$$\overline{P}_i \le \mathbb{E}_i(K), \quad \forall i, \tag{26}$$

$$0 \le F_{ij} \le TC, \quad \forall i, j. \tag{27}$$

The difference between the problem **P0** and **P1** is the energy consumption constraint. It can be found that, when $t' = 1$, the optimal solution for problem **P1** for $(K+1)$-th period is the optimal solution for problem **P0** for K-th period. Thus, the gap between network utility for problem **P0** and that for **P1** depends on the gap between initial battery level B_i^{ini} and the total amount of harvested energy $E(K+1)$.

4.2 Analysis of optimal solution

Let E_{ij}^{tr*}, X_i^* and F_{ij}^* denote the optimal solution of problem **P1** for K-th period. We have the following results:

Lemma 2. For optimal solution, the total amount of energy consumption $\overline{P}_i$ should be equal to total amount of available energy $\mathbb{E}_i(K)$, i.e., $\overline{P}_i = \mathbb{E}_i(K)$

Proof: The proof can be found in Appendix A. □

Lemma 3. The optimal transmission energy consumption E_{ij}^{tr*} should be equal to the threshold for transmission energy consumption $E_{ij}^{th}(t)$, i.e., $E_{ij}^{tr*} = E_{ij}^{th}(t)$.

Proof: The proof can be found in Appendix B. □

Lemma 4. For any given E_{is}^{tr} in one period, the total amount of the data transmitted to the mobile sink directly by sensor i has an upper bound $\theta_i(E_{is}^{tr})C$.

Proof: According to Lemma 1, given a transmission energy consumption E_{is}^{tr}, the maximal transmit distance d_{is}^{max} can be calculated by Eq. (5) and the maximal number of slots $\theta_i(E_{is}^{tr})$ can be given by Eq. (8). Thus, the amount of data F_{is} that can be transmitted to mobile sink s directly by sensor i can be given by

$$F_{is} = \sum_{t=1}^{T} f_{is}(t) = \sum_{t=\underline{t}_i}^{\overline{t}_i} f_{is}(t) \le \theta_i(E_{is}^{tr})C, \tag{28}$$

where $\theta_i(E_{is}^{tr}) = 0$, when $d_{is}^{max} \le h_i$, and

$$\theta_i(E_{is}^{tr}) = \frac{2\sqrt{(d_{is}^{max})^2 - h_i^2}}{v} = \frac{2\sqrt{(\frac{E_{is}^{tr}-\beta}{\mu})^{\frac{2}{\alpha}} - h_i^2}}{v}, \tag{29}$$

when $d_{is}^{max} \ge h_i$. From Eq. (29), it can be found that $\theta_i(E_{is}^{tr})$ is an increasing function of transmission energy consumption E_{is}^{tr} when $d_{is}^{max} \ge h_i$. Thus, for any given E_{is}^{tr}, F_{is} has an upper bound $\theta_i(E_{is}^{tr})C$. □

Lemma 5. For the optimal solution, F_{is}^* should be equal to the upper bound $\theta_i(E_{is}^{tr}) * C$.

Proof: Since $\theta_i(E_{is}^{tr})C$ is an upper bound of F_{is} and $\theta_i(E_{is}^{tr})$ is an increasing function of transmission energy consumption E_{is}^{tr} when $d_{is}^{max} \geq h_i$, F_{is} should be located in $[0, \theta_i(E_{is}^{tr})C]$. Thus, we only need to prove that F'_{is}, which is smaller than $\theta_i(E_{is}^{tr*})C$, is not the optimal solution. Assuming that, there exists an optimal solution $F'_{is} = \theta'_i(E_{is}^{tr})C$ and $\theta'_i(E_{is}^{tr}) < \theta_i(E_{is}^{tr})$. The corresponding energy consumption for F'_{is} is $F'_{is}E_{is}^{tr*}$. Since $\theta_i(E_{is}^{tr})$ is an increasing function of E_{is}^{tr}, there exists another transmission energy consumption $\hat{E}_{is}^{tr}$, $\hat{E}_{is}^{tr} < E_{is}^{tr*}$, for sensor i to transmit F'_{is} unit data to mobile sink, which is a contradiction of Lemma 3. Thus, the optimal F_{is}^* should be equal to the upper bound $\theta_i(E_{is}^{tr})C$. □

4.3 Analysis of convexity

In this section, we analyze the convexity of the transformed problem. Firstly, we prove that the energy consumption for data transmission is a convex function of F_{ij}, then we show that the problem is a convex optimization problem.

Lemma 6. $\theta_i(E_{is}^{tr})$ is a concave function of transmission energy consumption E_{is}^{tr} when $d_{is}^{max} \geq h_i$.

Proof: The proof can be found in Appendix C. □

Lemma 7. The optimal transmission energy consumption E_{is}^{tr*} is an increasing and convex function of F_{is}.

Proof: The proof can be found in Appendix D. □

From Lemma 7, it can be found that, for any given F_{is}, we can find the optimal transmission energy consumption E_{is}^{tr*} accordingly, which can be given by

$$E_{is}^{tr*} = \mu\Big((\frac{F_{is}v}{2C})^2 + h_i^2 \Big)^{\frac{\alpha}{2}} + \beta. \quad (30)$$

Thus, the total amount of energy consumption for sensor i to transmit F_{is} unit data to mobile sink s can be given by

$$F_{is}E_{is}^{tr*} = F_{is}\bigg(\mu\Big((\frac{F_{is}v}{2C})^2 + h_i^2 \Big)^{\frac{\alpha}{2}} + \beta \bigg) \quad (31)$$

For the total amount of energy consumption, we have the following results:

Lemma 8. The total amount of energy consumption $F_{is}E_{is}^{tr*}$ is an increasing and convex function of F_{is}.

Proof: The proof can be found in Appendix E. □

Theorem 1. The transformed problem **P1** is a convex optimization problem.

Proof: Since $E_i^{re} \sum_{k:i \in O(k)} F_{ki} + E_i^{sn}T$ is a linear function of F_{ki} and $\sum_{j:j\in O(i)} E_{ij}^{tr} F_{ij}$ is a convex function of F_{ij}, the constraint (26) is a convex function of F_{ij}. It can be found that the constraint (25) is a linear constraint for X_i and F_{ij} and constraint (27) is a linear constraint for F_{ij}. Since the objective function (24) is concave function of X_i, the transformed problem **P1** is a convex optimization problem. □

5 Distributed data collection approach

In this section, we proposed a Distributed Data Gathering Approach (DDGA) to solve the formulated network utility maximization problem and an available scheduling scheme to manage the data transmission in any slot.

According to Lemma 3, the optimal transmission energy consumption E_{ij}^{tr*} should be equal to the threshold for transmission energy consumption $E_{ij}^{th}(t)$. Thus, E_{ik}^{tr} is a constant since E_{ik}^{th} is a constant given by Eq. (3) and E_{is}^{tr*} is a function of F_{is} given by Eq. (30). In addition, since $TC >> \theta_i(E_{is}^{tr})C$, we can omit the constraint (27). Hence, the data gathering optimization problem for K-th period can be rephrased as

$$\textbf{P2:} \max_{X_i, F_{ij}} \sum_i U(X_i) \quad (32)$$

$$s.t. \sum_{j\in O(i)} F_{ij} = \sum_{i \in O(k)} F_{ki} + X_i, \ \forall i \quad (33)$$

$$\overline{P}_i = \mathbb{E}_i(K) \qquad \forall i \quad (34)$$

In this data gathering optimization problem, the variables are: 1) The total amount of sensory data sensed and sent out by each sensor, denoted by X_i; 2) The total amount of sensory data goes through each link (i, j), denoted by F_{ij}. The objective function is an increasing, concave function of X_i. The first constraint (33) is the flow balance conservation, which is a linear constraint. The second constraint (34) is the energy consumption constraint, which ensures that the total amount of energy consumption for each sensor consumed should equals to the energy that it collected.

Since the problem **P2** is a convex optimization problem, we design a distributed data gathering approach, by employing dual decomposition method and subgradient method, to obtain the globally optimal solution. Let $\boldsymbol{\lambda} = \{\lambda_i | i = 1, 2, \cdots, N\}$ and $\boldsymbol{\xi} = \{\xi_i | i = 1, 2, \cdots, N\}$ be the Lagrange multipliers, and the dual problem is given as follows:

$$L(\boldsymbol{\lambda}, \boldsymbol{\xi}) = \max_{X_i, F_{ij}} \sum_i \Big(U(X_i) + \xi_i(\mathbb{E}_i(K) - \overline{P}_i + \lambda_i(\sum_{j\in O(i)} F_{ij} - \sum_{i\in O(k)} F_{ki} - X_i)) \Big). \quad (35)$$

The dual problem of **P2** is

$$\min_{\boldsymbol{\lambda}, \boldsymbol{\xi}} L(\boldsymbol{\lambda}, \boldsymbol{\xi}). \quad (36)$$

For given $\boldsymbol{\lambda}$ and $\boldsymbol{\xi}$, the dual problem can be divided into two independent terms, such as

$$L_1(\boldsymbol{\lambda}) = \max_{X_i} \sum_i \Big(U(X_i) - \lambda_i X_i \Big) \quad (37)$$

$$L_2(\boldsymbol{\lambda}, \boldsymbol{\xi}) = \max_{F_{ij}} \sum_i \Big(\lambda_i(\sum_{j\in O(i)} F_{ij} - \sum_{i \in O(k)} F_{ki}) + \xi_i(\mathbb{E}_i(K) - \overline{P}_i) \Big) \quad (38)$$

For the second sub-dual problem, we can change the description as follows:

$$L_2(\boldsymbol{\lambda}, \boldsymbol{\xi}) = \max_{F_{ij}} \sum_i \sum_{j\in O(i)} \Big((\lambda_i - \lambda_j)F_{ij} - (\xi_i E_{ij}^{tr*} + \xi_j e_j^{re})F_{ij} + \xi_i \mathbb{E}_i(K) - \xi_i e_i^{sn} \Big) \quad (39)$$

It can be found that all the sub-dual problems can be decomposed into some subproblems, which can be given by

$$\hat{L}_1^i(\boldsymbol{\lambda}) = \max_{X_i} U(X_i) - \lambda_i X_i \quad (40)$$

$$\hat{L}_2^i(\boldsymbol{\lambda}, \boldsymbol{\xi}) = \max_{F_{ij}} \sum_{j\in O(i)} \Big((\lambda_i - \lambda_j)F_{ij} - (\xi_i E_{ij}^{tr*} + \xi_j e_j^{re})F_{ij} + \xi_i \mathbb{E}_i(K) - \xi_i e_i^{sn} \Big) \quad (41)$$

By employing dual decomposition approach, the data collection optimization problem has been decomposed into some sub-dual problems, which can be solved by sensors. Thus, we proposed a distributed algorithm by employing the subgradient method. For each iteration,

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2016.2603152, IEEE Transactions on Mobile Computing

1) The total amount of data X_i, sensed and sent out by sensor i, is determined by

$$\max_{X_i} U(X_i) - \lambda_i X_i \tag{42}$$

2) The total amount of data F_{ij} transmitted through the link (i, j) is determined by

$$F_{ik}(k+1) = F_{ik}(k) + \sigma(\lambda_i(k) - \lambda_k(k) - \xi_i(k)E_{ik}^{tr} - \xi_k(k)E_k^{re}) \tag{43}$$

$$F_{is}(k+1) = F_{is}(k) + \sigma(\lambda_i(k) - \lambda_j(k) - \lambda_i(k)F_{is}\frac{\mu\alpha}{2}\Big((\frac{F_{is}v}{2C})^2 + h_i^2\Big)^{\frac{\alpha-2}{2}}\frac{F_{is}v}{C}) - \xi_j(k)E_i^{re} - \lambda_i(k)E_{is}^{tr}(k)) \tag{44}$$

3) The Lagrange multipliers can be updated iteratively according to the following equations:

$$\lambda_i(k+1) = \lambda_i(k) - \epsilon(\sum_{j\in O(i)} F_{ij} - \sum_{i\in O(k)} F_{ki} - X_i)$$

$$\xi_i(k+1) = \xi_i(k) - (\mathbb{E}_i(K) - \overline{P}_i) \tag{45}$$

where k is the iteration number.

Until all the variables converge to their optimal solution.

Theorem 2. For a sufficiently small positive constant ϵ, DDGA algorithm will converge to optimal solution for the entire period.

Proof: Since the primal problem is a convex optimization problem, the global optimal solution can be obtained by employing dual decomposition and sub-gradient method [30] [31]. we omit the details of the proof. □

By now, the optimal solution $X_i^*, \forall i$ and $F_{ij}^*, \forall i, j$, are known. Hence, the optimal transmission energy consumption E_{is}^{tr*} can be calculated by Eq. (30) and the maximal number of slots $\theta_i(E_{is}^{tr*})$ can be calculated by Eq. (29).

Since the mobile sink is traveling along the straight path periodically, the required transmission energy consumption will change with time. If optimal transmission energy consumption is given, the available slots for sensor to transmit sensory data to mobile sink will be given. Thus, we need to design an available scheduling scheme to manage the data transmission for each sensor to make sure the data transmission satisfy the flow conservation in any slot.

As is known, for any sensor i, there exists only one nearest position in the path, which also is the nearest point between the mobile sink and the sensor. Let D_i denote the nearest point and $\hat{t}_i$ denote the corresponding slot, at which the mobile sink reaches that point. For given $\theta_i(E_{is}^{th})$, $\underline{t}_i$ and $\overline{t}_i$ can be defined as follows:

$$\underline{t}_i = \hat{t}_i - \frac{\theta_i(E_{is}^{th})}{2}, \tag{46}$$

$$\overline{t}_i = \hat{t}_i + \frac{\theta_i(E_{is}^{th})}{2}. \tag{47}$$

Since sensor i only knows its own available slots $[\underline{t}_i, \overline{t}_i]$ and the values of F_{ki} and F_{ij}, it needs to make the decision based on these information. Thus, we proposed a simple scheduling scheme to schedule the data transmission according to the total amount of data that have been transmitted and the total amount of data that need to be transmitted. Let $\hat{F}_{ki}(t)$ denote the total amount of data have been transmitted through link (k, i), which can be given by

$$\hat{F}_{ki}(t) = \sum_{t'=1}^{t-1} f_{ki}(t'), \tag{48}$$

Note that, X_i will be treated as F_{ii} in this section.

The ratio between $\hat{F}_{ki}(t)$ and F_{ki} can be given by

$$W_{ki}(t) = \begin{cases} \frac{\hat{F}_{ki}(t)}{F_{ki}}, & \text{if } F_{ki} > 0 \\ +\infty, & \text{otherwise.} \end{cases} \tag{49}$$

where $k = 1, 2, \cdots, N$. At each slot t, $t = 1, 2, \cdots, T$, if sensor i can transmit to other sensors/the mobile sink, sensor i will select sensor k, whose value in $\{W_{ki}(t), k = 1, 2, \cdots, N\}$ is smallest, and notice sensor k to transmit data to him through link (k, i). The process can be given as in Algorithm 1.

Algorithm 1 Scheduling Scheme

repeat
  for $i = 1, 2, \cdots, N$
  1) If $t \in [\underline{t}_i, \overline{t}_i]$, goes to step 2);
  2) Sets $f_{is}(t) = C$, and goes to step 3);
  3) Finds the smallest value(s) in set $\{W_{ki}(t), k = 1, 2, \cdots, N\}$ and marks it as $W_{k'i}(t)$, then updates $W_{k'i}(t) = W_{k'i}(t) + \frac{C}{F_{k'i}}$ correspondingly and goes to step 4);
  4) If $k' = i$, sets $x_i(t) = C$ and goes to **end**; Otherwise, sets $f_{k'i}(t) = C$ and goes to step 5).
  5) Sets $i = k'$ and goes to step 3).

  end
  - Set $W_{ki}(t+1) = W_{ki}(t), \forall k, i$, and $t = t + 1$.

until $t = T + 1$
return $f_{ij}(t), \forall i, j, t$

Note that, if there are two sensors or more than two sensors with the same smallest value, sensor i will select one sensor with smallest value randomly. In this scheduling scheme, it is easy for sensor i to select its $x_i(t)$ and $f_{ki}(t)$ in slot t.

Theorem 3. The globally optimal solution is obtained by DDGA when all the flows F_{ij} can be divided by link capacity C with no remainder.

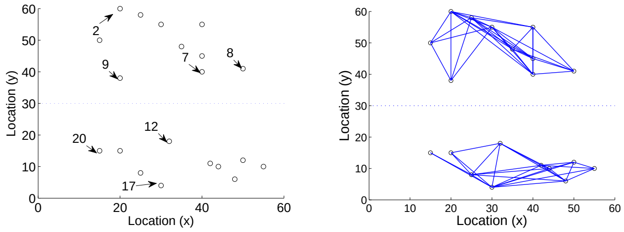
Proof: If all the flows F_{is} can be divided by link capacity C with no remainder, according to Lemma 5, $\theta_i(E_{is}^{tr})$ will be an integer. Similarly, if all the flows F_{ki} can be divided by link capacity C with no remainder, F_{ki} can be redescribed as $c_{ki}C$, where c_{ki} is an integer. Since the flow constraints $\sum_{j:j\in O(i)} f_{ij}(t) = \sum_{k:i\in O(k)} f_{ki}(t) + x_i(t)$ always hold, all X_i can be redescribed as $c_{ii}C$, where c_{ii} is an integer. Thus, the flow constraint $\sum_{j:j\in O(i)} f_{ij}(t) = \sum_{k:i\in O(k)} f_{ki}(t) + x_i(t)$ is equivalent to $\sum_{j:j\in O(i)} c_{ij} = \sum_{i\in O(k)} c_{ki} + c_{ii}$. According to the scheduling scheme, in each slot t, $t \in [\underline{t}_i, \overline{t}_i]$, sensor i will select only one sensor as source node and relay data for him. Thus, $\theta_i(E_{is}^{tr})C$ unit data will be transmitted to the mobile sink by sensor i during $[\underline{t}_i, \overline{t}_i]$. Since all the sensors should satisfy the flow conservation, X_i unit data will be sent out by sensor i and received by sink node, then globally optimal solution will be obtained. □

6 Performance Evaluation

In this section, numerical results are shown to demonstrate the performance of proposed algorithms by comparing with the direct transmission scheme [18], [19], in which an optimal one-hop data transmission scheme for RSNs with mobile sink is given, as well as the optimal data gathering for an RSN with a static sink node at $(30, 30)$. All the results are obtained by MATLAB.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2016.2603152, IEEE Transactions on Mobile Computing

6.1 Simulation setting



(a) Network topology.

(b) Available paths.

Fig. 2. Topology for simulation and the corresponding available path without the mobile sink

TABLE 2
Network parameter values

solar panel	$37 \times 33mm^2$	β	0.003mW/kb
B_{max}	2000mWh	μ	0.0002mW/kb
α [32]	3.14	e^{re}/e^{sn}	0.276/0.022mW/kb

In such an RSN with a mobile sink, there are 20 sensors deployed into a $60m \times 60m$ area randomly, and the mobile sink runs along the road ($y = 30, x \in (0, 100)$) with a give speed $v = 1m/s$ to collect data sent out by the sensors. The maximal data gathering rate for each link is $20kbps$. All the sensors have the same wireless module, such as TelosB from Crossbow [33]. The values of network parameters are given in Table 2. Fig. 2(a) shows a simple network topology for the simulation. The corresponding available path (without the mobile sink) can be shown in Fig. 2(b). Since all the sensors can transmit data to mobile sink directly, the link between each sensor and the mobile sink is always available. From the Fig. 2(b), it is easy to find that some of the sensors, close to the road, need to serve as relay node for other sensors.

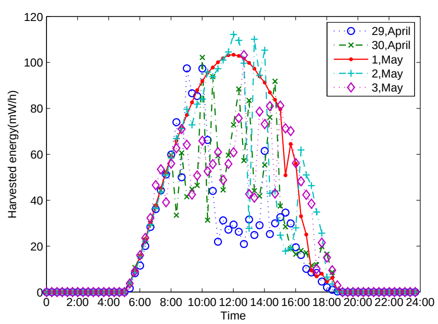


Fig. 3. Experimental solar profile from 29th, April to 3rd, May 2014.

In this paper, we use the solar profile obtained from Baseline Measurement System (BMS) of Solar Radiation Research Laboratory (SRRL) for a period from 29th, April to 3rd, May 2014 [34], and the solar profile can be found in Fig. 3. The total amount of harvested energy for the five days are $449.1131mWh$, $571.7443mWh$, $817.5944mWh$, $761.7379mWh$ and $660.7781mWh$, respectively. Let the initial energy of the rechargeable battery for all senors be $1000mWh$, and the utility function be $U(X_i) = \log(X_i)$.

6.2 Performance evaluation of distributed data gathering approach

Fig. 4 shows the results of the routing path for the optimal solution in 30th, April. Comparing with Fig. 2(b), it can be found that some of the available paths will not be used during the entire period.

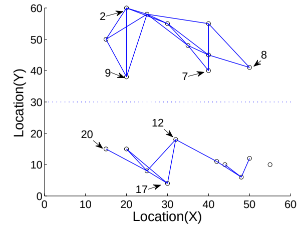
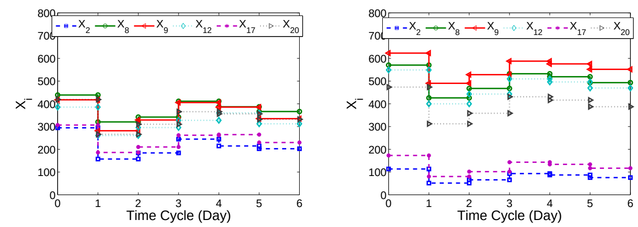


Fig. 4. the routing path for the sensor network in 30th, April.

That is because each sensor needs to select its optimal routing path according to its available energy and its transmission power, as well as the traffic load and energy consumption of its next-hop. Taking sensor 2 as an example, there are more than 8 available paths, but in fact, it only selects three links as its routing paths.

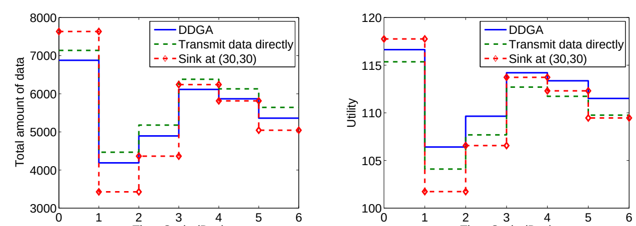


(a) Data for DDGA.

(b) Data for direct transmission.

Fig. 5. Total amount of data for each day obtained by DDGA and direct transmission, respectively.

Fig. 5 shows the total amount of data sent out by sensors employing DDGA and direct transmission scheme, respectively. In the same day, the gap between the largest value and the smallest value for the sensors employing direct transmission scheme is much larger than that employing DDGA. Taking sensor 9, which is closest to the road in the network, and sensor 2, which is farthest from the road, as an example: by employing direct transmission scheme, sensor 9 obtains the largest value, while sensor 2 obtains the smallest value; by employing DDGA, sensor 9 does not obtain the largest value since it needs to relay data for other sensors and its value decreases, while sensor 2's values increases greatly, thus maintains the fairness of the network.



(a) Total amount of data (Kbps).

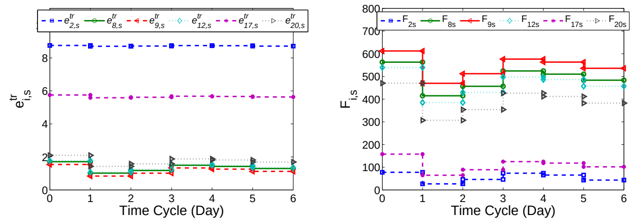
(b) Utility.

Fig. 6. Total amount of data for the entire network in each day and corresponding utility obtained by DDGA, direct transmission scheme and optimal data gathering for an RSN with a static sink node at $(30, 30)$, respectively.

Fig. 6 shows the total amount of data collected by the mobile sink and the utility values obtained by DDGA, direct transmission scheme and optimal data gathering for an RSN with a static sink node at $(30, 30)$, respectively. From these two figures, it can be found that the total amount of data collected by the mobile sink employing direct transmission approach is larger than that employing DDGA. However, the utility for sensor employing the direct transmission approach is much smaller than that employing

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2016.2603152, IEEE Transactions on Mobile Computing

DDGA. That is because the sensors employing direct transmission approach only aims at maximizing the total amount of data, and the sensors employing DDGA aims at maximizing the total amount of data while maintaining the fairness of the network. Most of time, the total amount of utility obtained by DDGA is larger than that obtained by optimal data gathering for an RSN with a static sink node at $(30, 30)$. The statistical characteristic of data given in the section 6.4 demonstrates the efficient of proposed algorithm for an RSN with a mobile sink.



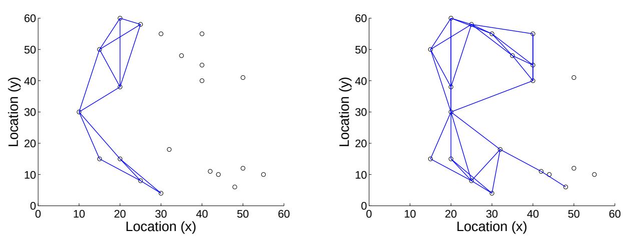
(a) E_{is}^{tr}.

(b) F_{is}.

Fig. 7. The optimal transmission energy consumption E_{is}^{tr} and the optimal total amount of flow F_{is}.

Fig. 7(a) shows the transmission energy consumption E_{is}^{tr} for each sensor i to transmit data to mobile sink directly and Fig. 7(b) shows the total amount of data F_{is} transmitted by sensor i to mobile sink. From Fig. 7(a), it can be found that the transmission power energy consumption E_{is}^{tr} for the sensor, which is closer to the road, is much smaller than that for the sensor, which is far away from the road. It can be seen from Fig. 7(b) that, the sensor, which is closest to the road, will transmit largest amount of data to the mobile sink. Joint with Fig. 5, it can be found that the largest values in $\boldsymbol{X}$ is X_8 instead of X_9, since sensor 9 needs to relay more data for other sensors than sensor 8.

6.3 Performance evaluation of scheduling



(a) Mobile sink at $(10, 30)$.

(b) Mobile sink at $(20, 30)$.

Fig. 8. Real-time routing scheme when mobile sink at the different location.

Fig. 8 shows the real-time routing scheme for the mobile sink at $[10, 30]$ and $[20, 30]$, respectively. From these two figures, it can be found that the real-time routing scheme will change with the location of the mobile sink. When the mobile sink reaches $(10, 30)$, there are only a few sensors will transmit their data to the mobile sink since most of them are far away from the mobile sink. But when the mobile sink reaches $(20, 30)$, most of the sensors will transmit their data to the mobile sink. That is because the distance between the sensor and the mobile sink decreases, thus decreases transmission energy consumption for the sensor to transmit data to the mobile sink. In order to transmit data to mobile sink as much as possible, the sensor needs to select its routing scheme based on the location of the mobile sink.

Fig. 9 shows the relationship between $f_{is}(t)$ and the location of the mobile sink in 30th, April. From this figure, it can be found

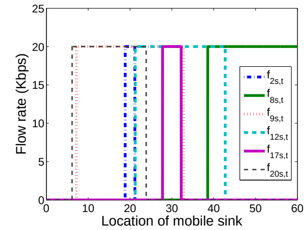
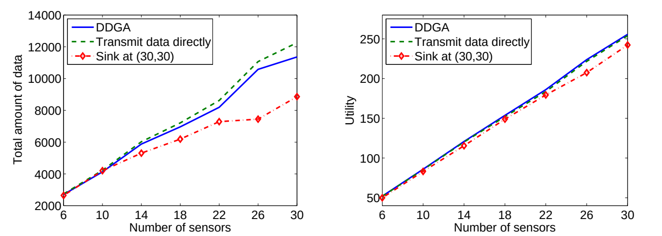


Fig. 9. $f_{is}(t)$ and location of the mobile sink.

that the $f_{is}(t)$ equals to the upper bound when the sensor begins to transmit data. Furthermore, the sensor only transmits data to mobile sink directly when the mobile sink is closer to him and the sensor, which is far away from the road, only transmits a few data to the mobile sink directly, since it is energy-consuming to do this.

6.4 Numerical simulation results

In order to explore the performance of proposed algorithm for an RSN with a mobile sink, we perform simulations in a $60m \times 60m$ area in which $14, 18, 22, 26, 30$ sensors are randomly deployed. The simulation results can be found in Fig. 10. It can be found that, with the increase of the number of sensors, the total amount of data, as well as the total amount of utility increases. It can be seen from the simulation results, that sensors employing DDGA obtains the largest utility, which is much larger than that employing optimal data gathering scheme for an RSN with a static sink node.

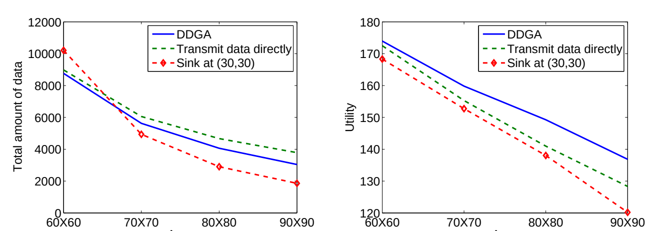


(a) Total amount of data (Kbps).

(b) Utility.

Fig. 10. Total amount of data and corresponding utility obtained by DDGA, direct transmission scheme and optimal data gathering for an RSN with a static sink node at $(30, 30)$, respectively.

In order to explore the performances of optimal multi-hops and one-hop data transmission schemes for an RSN with a mobile sink, we randomly deploy 20 sensors into several different areas (such as $60m \times 60m$, $70m \times 70m$, $80m \times 80m$ and $90m \times 90m$, respectively). The simulation results can be found in Fig. 11. With increase of area, the gap between total amount of utility obtained by DDGA and that obtained by direct transmission increases. Again, sensors employing DDGA obtains the largest utility.



(a) Total amount of data (Kbps).

(b) Utility.

Fig. 11. Total amount of data and corresponding utility for different areas obtained by DDGA, direct transmission scheme and optimal data gathering for an RSN with a static sink node at $(30, 30)$, respectively.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2016.2603152, IEEE Transactions on Mobile Computing

7 Conclusion

In this paper, we have studied the data gathering problem in RSNs with a mobile sink. We first formulated the problem as a network utility maximization problem. Then, we transformed the network utility maximization problem into an approximate one by shifting the energy consumption conservation and analyzing necessary conditions for the optimal solution. As a result, each sensor needs not estimate the amount of harvested energy and the problem dimension is reduced. And then we proved that the transformed problem is a convex optimization problem and designed a Distributed data gathering Approach to obtain the optimal data gathering. Extensive simulations based on real experimental data are performed to demonstrate the efficiency of the proposed algorithm.

Appendix A

Proof: Since the total amount of available energy $\mathbb{E}_i(K)$ is the upper bound of total amount of energy consumption $\overline{P}_i$, $\overline{P}_i \leq \mathbb{E}_i(K)$ holds. If there exist an optimal $\overline{P}_i^{'}$, $\overline{P}_i^{'} < \mathbb{E}_i(K)$, and corresponding X_i, the total amount of energy reserved by sensor i is $\mathbb{E}_i(K) - \overline{P}_i^{'}$. As all the sensor can transmit data to the mobile sink directly, the sensor can make use of the reserved energy $\mathbb{E}_i(K) - \overline{P}_i^{'}$ to transmit more data to mobile sink directly. Thus, there exists X_i^*, $X_i^* \geq X_i + \frac{\mathbb{E}_i(K) - \overline{P}_i^{'}}{E_{is}^{max}}$, where E_{is}^{max} can be any transmission energy consumption and $E_{is}^{max} > E_{is}^{tr*}$. Hence, $X_i^* > X_i$ holds. Because the objective function is an increasing function of X_i, the optimal solution for $\overline{P}_i = \mathbb{E}_i(K)$ is better than that for $\overline{P}_i^{'} < \mathbb{E}_i(K)$. Thus, $\overline{P}_i^{'}$ is not optimal energy consumption. □

Appendix B

Proof: Since $E_{ij}^{th}(t)$ is the lower bound of transmission energy consumption for sensor i to forward data to its next-hop j at slot t, thus optimal transmission energy consumption should not be smaller than E_{ij}^{th}, i.e., $E_{ij}^{tr*} \geq E_{ij}^{th}$ holds. If there exists an optimal transmission energy consumption $E_{ij}^{tr'} > E_{ij}^{th}$, the total amount of energy consumption is $\overline{P}_i^{'} = \sum_{j \in O(i)} E_{ij}^{tr'} F_{ij}^* + E_i^{re} \sum_{i \in O(k)} F_{ki}^* + E_i^{sn}$. According to Lemma 2, $\overline{P}_i^{'} = \mathbb{E}_i(K)$ holds. Let $\overline{P}_i^* = \sum_{j \in O(i)} E_{ij}^{th} F_{ij}^* + E_i^{re} \sum_{i \in O(k)} F_{ki} + E_i^{sn}$, $\mathbb{E}_i(K) > \overline{P}_i^*$ holds since $E_{ij}^{tr'} > E_{ij}^{th}$. Thus, similar to the proof in Lemma 2, it is possible for sensor i to transmit additional δ_i unit data to mobile sink directly, where $\delta_i \geq \frac{\mathbb{E}_i(K) - \overline{P}_i^*}{E_{is}^{max}}$. Since the utility function is an increasing function of X_i, the total utility for $E_{ij}^{tr'}$ is smaller than that for E_{ij}^{tr*}. Thus, the transmission energy consumption E_{ij}^{tr*} should be equal to E_{ij}^{th}. □

Appendix C

Proof: For any given transmission energy consumption E_{is}^{tr}, we can calculate the maximal Euclidean distance d_{is}^{max}, using Eq. (5), correspondingly. Thus, the partial derivative of $\theta_i(E_{is}^{tr})$ with respect to E_{is}^{tr} is

$$\frac{\partial \theta_i(E_{is}^{tr})}{\partial E_{is}^{tr}} = \frac{2}{v\alpha\mu} \frac{1}{\sqrt{(\frac{E_{is}^{tr}-\beta}{\mu})^{\frac{2}{\alpha}} - h_i^2}} (\frac{E_{is}^{tr} - \beta}{\mu})^{\frac{2-\alpha}{\alpha}} \quad (50)$$

and the second derivative is

$$\begin{aligned} \frac{\partial^2 \theta_i(E_{is}^{tr})}{\partial {E_{is}^{tr}}^2} &= \frac{2}{v\alpha\mu} \Big[\frac{\partial \frac{1}{\sqrt{(\frac{E_{is}^{tr}-\beta}{\mu})^{\frac{2}{\alpha}} - h_i^2}}}{\partial E_{is}^{tr}} (\frac{E_{is}^{tr} - \beta}{\mu})^{\frac{2-\alpha}{\alpha}} \\ &+ \frac{1}{\sqrt{(\frac{E_{is}^{tr}-\beta}{\mu})^{\frac{2}{\alpha}} - h_i^2}} \frac{\partial (\frac{E_{is}^{tr}-\beta}{\mu})^{\frac{2-\alpha}{\alpha}}}{\partial E_{is}^{tr}} \Big]. \quad (51) \end{aligned}$$

Since $\alpha \geq 2$, $\frac{1}{\sqrt{(\frac{E_{is}^{tr}-\beta}{\mu})^{\frac{2}{\alpha}} - h_i^2}} > 0$ and $(\frac{E_{is}^{tr}-\beta}{\mu})^{\frac{2-\alpha}{\alpha}} > 0$, thus Eq. (50) is positive. Furthermore, since $\frac{\partial \frac{1}{\sqrt{(\frac{E_{is}^{tr}-\beta}{\mu})^{\frac{2}{\alpha}} - h_i^2}}}{\partial E_{is}^{tr}} < 0$ and $\frac{\partial (\frac{E_{is}^{tr}-\beta}{\mu})^{\frac{2-\alpha}{\alpha}}}{\partial E_{is}^{tr}} \leq 0$, Eq. (51) is negative. Therefore, $\theta_i(E_{is}^{tr})$ is a concave function of transmission energy consumption E_{is}^{tr}. □

Appendix D

Proof: According to Lemma 3, the optimal transmission energy consumption E_{is}^{tr*} should be equal to the threshold for transmission energy consumption E_{is}^{th}, and according to Lemma 5, F_{is} should equal to $\theta_i(E_{is}^{tr})C$. Thus, the optimal transmission energy consumption for sensor i to transmit F_{is} unit data to mobile sink directly is E_{is}^{tr*}. With the increase of F_{is}, E_{is}^{tr*} will increase. According to Lemma 6, $\theta_i(E_{is}^{tr})$ is an increasing and concave function of transmission energy consumption E_{is}^{tr} when $d_{is}^{max} \geq h_i$. Thus, E_{is}^{tr} is a convex function of $\theta_i(E_{is}^{tr})$. Since C is a constant, the optimal transmission energy consumption E_{is}^{tr*} is an increasing and convex function of F_{is}. □

Appendix E

Proof: According to the total amount of energy consumption $F_{is}E_{is}^{tr*}$, given in Eq. (31), the partial derivative of $F_{is}E_{is}^{tr*}$ with respect to F_{is} is

$$\begin{aligned} \frac{\partial F_{is}E_{is}^{tr*}}{\partial F_{is}} &= \mu\Big((\frac{F_{is}v}{2C})^2 + h_i^2\Big)^{\frac{\alpha}{2}} + \beta \\ &+ \frac{v\mu\alpha}{2C} F_{is}^2 \Big((\frac{F_{is}v}{2C})^2 + h_i^2\Big)^{\frac{\alpha-2}{2}}. \end{aligned}$$

and the second derivative is

$$\begin{aligned} \frac{\partial^2 F_{is}E_{is}^{tr*}}{\partial {F_{is}}^2} &= \frac{3v\mu\alpha}{4C} F_{is}\Big((\frac{F_{is}v}{2C})^2 + h_i^2\Big)^{\frac{\alpha-2}{2}} \\ &+ F_{is}^3 \frac{v^2\mu\alpha(\alpha-2)}{4C^2}\Big((\frac{F_{is}v}{2C})^2 + h_i^2\Big)^{\frac{\alpha-4}{2}}. \end{aligned}$$

Since α is located in $[2, 4]$, all of the partial derivatives and the second derivatives are larger than 0. Thus, the total amount of energy consumption $F_{is}E_{is}^{tr*}$ is an increasing and convex function of F_{is}. □

References

[1] A. Chakrabarti, A. Sabharwal, and B. Aazhang. Using predictable observer mobility for power efficient design of sensor networks. In *Information Processing in Sensor Networks*, pages 129–145. Springer, 2003.

[2] A. Kansal, A. A Somasundara, D. D Jea, M. B Srivastava, and D. Estrin. Intelligent fluid infrastructure for embedded networks. In *Proceedings of the 2nd international conference on Mobile systems, applications, and services*, pages 111–124. ACM, 2004.

[3] E. Ekici, Y. Gu, and D. Bozdag. Mobility-based communication in wireless sensor networks. *IEEE Communications Magazine*, 44(7):56, 2006.

[4] Y. Yun, Y. Xia, B. Behdani, and J. C. Smith. Distributed algorithm for lifetime maximization in a delay-tolerant wireless sensor network with a mobile sink. *IEEE Transactions on Mobile Computing*, 12(10):1920–1930, 2013.

[5] M. Di Francesco, S. K Das, and G. Anastasi. Data collection in wireless sensor networks with mobile elements: A survey. *ACM Transactions on Sensor Networks (TOSN)*, 8(1):7, 2011.

[6] A. Chakrabarti, A. Sabharwal, and B. Aazhang. Communication power optimization in a sensor network with a path-constrained mobile observer. *ACM Transactions on Sensor Networks (TOSN)*, 2(3):297–324, 2006.

[7] S. Gao, H. Zhang, and S. K Das. Efficient data collection in wireless sensor networks with path-constrained mobile sinks. *IEEE Transactions on Mobile Computing*, 10(4):592–608, 2011.

[8] S. Guo and Y. Yang. A distributed optimal framework for mobile data gathering with concurrent data uploading in wireless sensor networks. In *Proceedings of IEEE INFOCOM*, pages 1305–1313. IEEE, 2012.

[9] X. Li, J. Yang, A. Nayak, and I. Stojmenovic. Localized geographic routing to a mobile sink with guaranteed delivery in sensor networks. *IEEE Journal on Selected Areas in Communications*, 30(9):1719–1729, 2012.

[10] M. Zhao, Y. Yang, and C. Wang. Mobile data gathering with load balanced clustering and dual data uploading in wireless sensor networks. *IEEE Transactions on Mobile Computing*, 14(4):770–785, 2015.

[11] K. Lin, J. Yu, J. Hsu, S. Zahedi, D. Lee, J. Friedman, A. Kansal, V. Raghunathan, and M. Srivastava. Long-lived sensor networks through solar energy harvesting. In *Proceedings of ACM SenSys*, pages 309–309, 2005.

[12] C. Park and P. Chou. Ambimax: Autonomous energy harvesting platform for multi-supply wireless sensor nodes. In *Proceedings of IEEE SECON*, pages 168–177, Reston, VA, 2006.

[13] S. He, J. Chen, F. Jiang, D. KY Yau, G. Xing, and Y. Sun. Energy provisioning in wireless rechargeable sensor networks. *IEEE Transactions on Mobile Computing*, 12(10):1931–1942, 2013.

[14] L. Lin, N. B. Shroff, and R. Srikant. Asymptotically optimal energy-aware routing for multihop wireless networks with renewable energy sources. *IEEE/ACM Transactions on Networking*, 15(5):1021–1034, 2007.

[15] R. S. Liu, K. W. Fan, Z. Z. Zheng, and P. Sinha. Perpetual and fair data collection for environmental energy harvesting sensor networks. *IEEE/ACM Transactions on Networking*, 19(4):947–960, 2011.

[16] R. S. Liu, P. Sinha, and C. E. Koksal. Joint energy management and resource allocation in rechargeable sensor networks. In *Proceedings of IEEE INFOCOM*, pages 1–9, San Diego, CA, 2010.

[17] Z. Mao, C. E. Koksal, and N. B. Shroff. Near optimal power and rate control of multi-hop sensor networks with energy replenishment: Basic limitations with finite energy and data storage. *IEEE Transactions on Automatic Control*, 57(4):815–829, 2012.

[18] X. Ren, W. Liang, and W. Xu. Use of a mobile sink for maximizing data collection in energy harvesting sensor networks. In *Proceedings of ICPP*, pages 439–448. IEEE, 2013.

[19] X. Ren and W.and Xu W. Liang. Data collection maximization in renewable sensor networks via time-slot scheduling. *IEEE Transactions on Computing, to appear*.

[20] Y. Shi and Y. T. Hou. Some fundamental results on base station movement problem for wireless sensor networks. *IEEE/ACM Transactions on Networking (TON)*, 20(4):1054–1067, 2012.

[21] S. Chen, P. Sinha, N. B. Shroff, and C. Joo. A simple asymptotically optimal energy allocation and routing scheme in rechargeable sensor networks. *IEEE/ACM Transactions on Networking*, 22(4):1325–1336, 2014.

[22] C. Zhang, J. Kurose, Y. Liu, D. Towsley, and M. Zink. A distributed algorithm for joint sensing and routing in wireless networks with non-steerable directional antennas. In *Proceedings of IEEE ICNP*, pages 218–227, 2006.

[23] V. Joseph, V. Sharma, U. Mukherji, and M. Kashyap. Joint power control, scheduling and routing for multicast in multihop energy harvesting sensor networks. In *Proceedings of ICUMT*, pages 1–8, 2009.

[24] S. Guo, C. Wang, and Y. Yang. Joint mobile data gathering and energy provisioning in wireless rechargeable sensor networks. *IEEE Transactions on Mobile Computing*, 99(1):1, 2014.

[25] Songtao Guo, Cong Wang, and Yuanyuan Yang. Mobile data gathering with wireless energy replenishment in rechargeable sensor networks. In *INFOCOM, 2013 Proceedings IEEE*, pages 1932–1940. IEEE, 2013.

[26] M. Zhao, J. Li, and Y. Yang. A framework of joint mobile energy replenishment and data gathering in wireless rechargeable sensor networks. *IEEE Transactions on Mobile Computing*, 13(12):2689–2705, 2014.

[27] J. Pan, Y. T. Hou, L. Cai, Y. Shi, and S. X. Shen. Topology control for wireless sensor networks. In *Proceedings of ACM MobiCom*, pages 286–299, 2003.

[28] L. Orihuela, F. Gómez-Estern, and Francisco R. R. Scheduled communication in sensor networks. *IEEE Transactions on Control Systems Technology*, 22(2):801–808, 2014.

[29] S. H. Low and D. E. Lapsley. Optimization flow control-I: Basic algorithm and convergence. *IEEE/ACM Transactions on Networking*, 7(6):861–874, 1999.

[30] D. Bertsekas. *Nonlinear Programming*. Athena Scientific, 1999.

[31] M. Chiang. Balancing transport and physical layers in wireless multihop networks: Jointly optimal congestion control and power control. *IEEE Journal on Selected Areas in Communications*, 23(1):104–116, 2005.

[32] B. Gaudette, V. Hanumaiah, M. Krunz, and S. Vrudhula. Maximizing quality of coverage under connectivity constraints in solar-powered active wireless sensor networks. *ACM Transactions on Sensor Networks (TOSN)*, 10(4):59, 2014.

[33] R. Liu, K. Fan, Z. Zheng, and P. Sinha. Perpetual and fair data collection for environmental energy harvesting sensor networks. *IEEE/ACM Transactions on Networking*, 19(4):947–960, 2011.

[34] NREL Solar Radiation Research Laboratory. Baseline measurement system (BMS). $www.nrel.gov/midc/srrl_bms/$.