A Distributed Optimal Framework for Mobile Data Gathering with Concurrent Data Uploading in Wireless Sensor Networks

Songtao Guo and Yuanyuan Yang
High Performance Networking Research Center and State Key Lab. of Power Transmission Equipment and System Security, College of Computer Science, Chongqing University, Chongqing, 400044, P.R. China
Department of Electrical and Computer Engineering, Stony Brook University, Stony Brook, NY 11794, USA

***Abstract*—In this paper, we consider mobile data gathering in wireless sensor networks (WSNs) by using a mobile collector with multiple antennas. By taking into account the elastic nature of wireless link capacity and the power control for each sensor, we first propose a data gathering cost minimization (DaGCM) framework with concurrent data uploading, which is constrained by flow conservation, energy consumption, link capacity, compatibility among sensors and the bound on total sojourn time of the mobile collector at all anchor points. One of the main features of this framework is that it allows concurrent data uploading from sensors to the mobile collector to sharply shorten data gathering latency and significantly reduce energy consumption due to the use of multiple antennas and space-division multiple access technique. We then relax the DaGCM problem with Lagrangian dualization and solve it with the subgradient iteration algorithm. Furthermore, we present a distributed algorithm composed of cross-layer data control, routing, power control and compatibility determination subalgorithms with explicit message passing. We also give the subalgorithm for finding the optimal sojourn time of the mobile collector at different anchor points. Finally, we provide numerical results to show the convergence of the proposed DaGCM algorithm and its advantages over the algorithm without concurrent data uploading and power control in terms of data gathering latency and energy consumption.**

***Index Terms*—Mobile data gathering, data gathering cost, convex optimization, distributed algorithm, wireless sensor networks.**

I. Introduction

Recent advances in sensor technology and wireless communications have enabled wireless sensor networks (WSNs) to play an increasingly important role in a wide range of applications, such as remote habitat monitoring and battlefield monitoring [1]. In such applications, hundreds or even thousands of low-cost sensors normally powered by energy-constrained batteries are dispersed over the monitoring area, and these nodes self-organize into a wireless network, in which each sensor node periodically reports its sensed data to the sink(s). Thus, how to efficiently aggregate the sensed data from scattered sensors with lowest energy consumption is one of the most critical challenges in the applications of large-scale resource-limited sensor networks, which motivates our work in this paper.

In recent years, much research effort has been devoted to efficient data gathering in WSNs and various types of data gathering mechanisms have been proposed for large-scale sensor networks. Most of them focused on static data gathering based on efficient relay routing [2] [3] and hierarchical infrastructures [4] [5]. The core idea of the relay routing is that data packets are forwarded to the data sink by single or multiple hop relays among sensors. In hierarchical infrastructures, a hierarchical or cluster-based routing approach is usually employed in which sensors are organized into clusters and cluster heads take the responsibility of forwarding data to the outside data sink. Although the approaches in these two categories can perform effective data forwarding in some applications, the major disadvantages of these approaches are that they induce and increase the non-uniformity of energy consumption among sensors, as a result, the neighboring nodes of the data sink will suffer from high congestion and packet loss, which severely degrades the network performance.

To overcome the drawbacks of these two types of mechanisms, mobile data gathering schemes have been proposed in [6]- [15]. The idea of the mobile data gathering approach is to deploy a special type of mobile nodes (usually called mobile collectors) to collect data from sensors at low cost via short range communications in the network. The mobile collector could be a mobile robot or vehicle with powerful transceivers (antennas) and batteries, which will be called SenCar in this paper. The advantage of this approach is that it can reduce energy consumption and traffic load, remove the burden of the sensor nodes closer to the data sink, increase network lifetime, and collect data in both connected networks and disconnected networks.

These existing mobile data gathering schemes can perform valid data gathering in WSNs, however, there still exist some inefficiencies. Specifically, some of these mobile data gathering schemes may lead to long data gathering latency, which is partly due to that the mobile collector uses a single antenna to collect data from only one sensor at a time. In practice, a mobile collector can be equipped with multiple transceivers or antennas, which enables the mobile collector to receive the data from multiple sensors simultaneously. Clearly, this can significantly shorten the data gathering time.

In addition, although link capacity constraint was considered in the optimization for mobile data gathering [11] - [15], it was regarded as a constant. In fact, link capacity is "elastic" in WSNs because it depends on transmission powers and wireless channel conditions such as link gains and thermal noises. As a result, these schemes are not suitable for the WSNs with fading channels. Moreover, saving energy in these schemes is achieved by only improving the routing paths of the packets or reducing the gathered data amount, rather than considering power control. In practice, power control can minimize the overall transmission power given a constant signal-to-interference-and-noise ratio (SINR) requirement for each sensor. These observations further motivate us to design a low latency mobile data gathering scheme with concurrent data uploading by considering multiple antennas and integrating power control.

In this paper, in order to reduce and balance the energy consumption among sensors and shorten the data uploading time, we deploy a SenCar with two antennas to collect data from sensors at some specific locations, i.e., anchor points, for a period of sojourn time in a sensing field, and utilize space-division multiple access (SDMA) technique to schedule data transmissions. By jointly using the SenCar with two antennas and the SDMA technique, we first propose a data gathering cost minimization (DaGCM) framework by considering the elastic nature of link capacity and the power control for each sensor, which is one of the major differences between the frameworks in [14] and [15] and this work. The data gathering cost (energy consumption cost) can be regarded as a function of the data amount that a sensor uploads to the mobile collector during its sojourn time at an anchor point. We then convert the original non-convex DaGCM problem into a convex one by introducing some auxiliary variables. Finally, we provide optimal algorithms and solutions to data control, routing, power control and sojourn time allocation subproblems by decomposing the convex DaGCM problem into separate optimization subproblems, which is another major difference between [15] and this work. In [15], although the SDMA technique was employed, only the heuristic algorithms and approximate solutions were given since the mobile data gathering problem was formulated into an NP-hard integer linear problem. To the best of our knowledge, our work is the first one that provides the optimal solutions to the mobile data gathering problem with concurrent data uploading.

The contributions and findings of the paper can be briefly summarized below:

- We adopt a SenCar with two antennas and the SDMA technique, which enables concurrent data uploading from sensors to the SenCar, to shorten data gathering latency and eliminate nonuniformity of energy consumption among sensors.
- We propose a DaGCM optimization framework by integrating the constraints of flow conservation, energy consumption, elastic link capacity and compatibility required by SDMA technique along with the bound on total sojourn time at all anchor points.
- We provide distributed cross-layer optimal subalgorithms of data control, routing and power control for each sensor, and sojourn time allocation for the SenCar by solving the non-convex DaGCM problem.
- Our numerical results demonstrate that the proposed DaGCM algorithm can converge to the optimum and remarkably reduce data gathering time and total energy consumption compared to the algorithm without concurrent data uploading and power control.

The reminder of this paper is organized as follows. Section II introduces the related work. Section III outlines the DaGCM framework and transforms the non-convex DaGCM into the convex one. Section IV proposes the distributed DaGCM algorithm, and Section V provides numerical results. Finally, Section VI concludes this paper.

II. Related Work

In this section, we review some recent work on mobile sinks for data collection in WSNs.

In [7], Luo and Hubaux studied how routing should be adjusted to improve the trajectory of the mobile collector. To obtain more flexible data gathering tours for mobile collectors, Ma and Yang [8] proposed a heuristic algorithm for planning the moving path of mobile collectors and balancing traffic load in a multi-hop network. Furthermore, they also [9] formalized the single-hop data gathering problem (SHDGP) into a mixed integer programming and presented a heuristic tour-planning algorithm for the mobile collector.

To reduce data collection latency, Xing, et al. proposed a rendezvous-based mobile data collection mechanism in [10] to minimize the distance in multi-hop routing paths for local data aggregation under the given tour length bound of the mobile collector. For WSNs with higher node densities, Sharifkhani and Beaulieu [11] proposed a transmission scheduling algorithm to optimize a metric that trades off between transmission power and reliability. In [12], Sugihara and Gupta formulated the problem of minimizing the latency of the data mule scheme into a scheduling problem under both location and time constraints.

In [13] and [15], Zhao, et al. provided optimization based distributed algorithms for data gathering where the mobile collector stays at each anchor point for a period of sojourn time and collects data from nearby sensors via multi-hop communications. In [14], the mobile data gathering problem was formulated into a cost minimization problem constrained by channel capacity, minimum amount of data gathered from each sensor and bound on total sojourn time at all anchor points. However, the work did not consider utilizing concurrent data uploading from sensors to the SenCar as this paper.

III. System Model and Problem Formulation

A. Network Model and Assumptions

Consider a wireless sensor network which consists of a set of static sensors, denoted as N, and a set of anchor points, denoted as A. We assume that the SenCar, denoted as s, is equipped with two antennas while each sensor has a single antenna and is statically scattered over the entire sensing field. When the SenCar moves to an anchor point a, it will stay at the anchor point for a period of sojourn time t^a to gather the data uploaded by nearby sensors in a multi-hop fashion. All sensors in the coverage area of an anchor point form the neighbor set of the anchor point, as shown in Fig 1. There are several ways to determine the sequence of visiting anchor points for

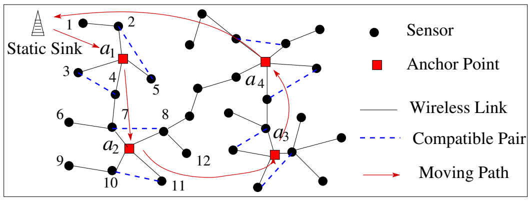

Fig. 1. Illustration of mobile data gathering with two antennas.

the SenCar, such as simply following the sequence of anchor point identifiers in an increasing order [16] or following the sequence in which the total moving tour length is minimized [17]. Our proposed algorithms can work for any given visiting sequence.

We model the sensor network with the SenCar located at an anchor point a ($a \in A$) as a directed graph $G^a = (V^a, E^a)$, where $V^a = N \cup \{s^a\}$ is the set of nodes, including all the sensors and the SenCar at anchor point a, denoted by s^a. E^a represents the set of directed links among the sensors and the SenCar. A directed link $(i,j) \in E^a$ exists if $d_{ij} \le r_{tx}$, where d_{ij} denotes the distance between node i and node j and r_{tx} stands for the transmission range of the sensor nodes. Each link is associated with a weight $e_{ij} = d_{ij}^2$, which is the energy consumed per unit flow on link (i,j). All the links are assumed to be symmetrical, i.e., $e_{ij} = e_{ji}$. Moreover, f_{ij}^a represents the flow rate over link (i,j) at anchor point a. Sensor $i \in N$ generates data for the SenCar at a data rate of R_i^a when the SenCar moves to anchor point a. Here, the rate vector $[R_i^a]_{\forall i \in V^a}$ and the flow vector $[f_{ij}]_{\forall (i,j) \in E^a}$ are regarded as the variables that can be adjusted to minimize the optimization objective.

We assume that each sensor i attains a data gathering cost function $NC_i(\cdot)$, which is twice-differentiable, increasing and strictly-convex with respect to the total amount of data gathered from sensor i in a data gathering tour (i.e.,$\sum_{a \in A} R_i^a t^a$). Cost function $NC_i(\cdot)$ can be regarded as the energy cost or the suitability for sensor i to upload data towards the SenCar at a data gathering tour [14]. Our work in this paper aims to minimize the data gathering cost by means of dynamically adjusting the data rate R_i^a of sensor i, the flow rate f_{ij}^a over link (i,j) and the sojourn time t^a at different anchor points, which is constrained by flow conservation, energy consumption, link capacity and compatibility among sensors. In the following, we will elaborate these constraints.

B. Flow Conservation Constraint

The forwarded data flow for any sensor comes from two sources: the incoming traffic from other sensors and the data sensed by the sensor. For each sensor $i \in V^a$ at anchor point $a \in A$, the total outgoing traffic flow $f_{i,out}^a = \sum_{j:(i,j) \in E^a} f_{ij}^a$ must equal the sum of the data flow generated by sensor i, R_i^a, and the aggregated incoming traffic flow, $f_{i,in}^a = \sum_{j:(j,i) \in E^a} f_{ji}^a$, i.e.,

$$f_{i,out}^a = R_i^a + f_{i,in}^a, \forall i \in N, \forall a \in A \tag{1}$$

Flow conservation constraint ensures that the outgoing and incoming flows are balanced.

C. Energy Constraint

In order to better enjoy the benefit of SDMA, it is necessary to ensure the number of concurrent data traffic flows is no more than the number of antennas. Since the SenCar is equipped with two antennas, at most two sensors are allowed to concurrently send data to the SenCar at a time. Therefore, if two sensors associated with the same anchor point want to upload their data simultaneously to the SenCar, they are required to be compatible and the two sensors are called a *compatible pair*, which will be formally defined in Section III-E.

It is assumed that sensor i has non-renewable battery energy. To guarantee a specific network lifetime, we impose an energy consumption budget W_i for sensor i, i.e., the maximum energy consumed by a sensor in a data gathering tour. Let $\rho(i,m)$ denote the ratio of the time spent by the compatible pair (i,m) or the isolated sensors i and m on uploading data to the sojourn time t^a of the SenCar at anchor point a. Note that nodes i and m here refer to the sensor nodes that are close to the SenCar and can upload data in a single hop manner to the SenCar instead of all the sensor nodes in the neighbor set V^a of the SenCar at anchor point a. The energy consumed by sensor i on transmitting data is $\sum_{j:(i,j) \in E^a} f_{ij}^a e_{ij} \rho(i,m) t^a$. Clearly, the energy consumed by the compatible pair (i,m) at anchor point a should not exceed their energy budgets, i.e.,

$$\sum_{j:(i,j) \in E^a} f_{ij}^a e_{ij} \rho(i,m) t^a \le W_i \tag{2}$$

$$\sum_{n:(m,n) \in E^a} f_{mn}^a e_{mn} \rho(i,m) t^a \le W_m \tag{3}$$

D. Link Capacity Constraint

The capacity of a fading channel is defined as the maximum rate that can be transmitted over the channel. Let p_{ij} denote the transmission power allocated to link $(i,j) \in E^a$ subject to $p_{ij}^{\min} \le p_{ij} \le p_{ij}^{\max}$, where $p_{ij}^{\min}$ and $p_{ij}^{\max}$ denote the minimum and maximum transmission powers, respectively, and γ_{ij} is the SINR of link $(i,j) \in E^a$. For the successful transmission on link (i,j), it is necessary that the received signal at node j is not garbled by another concurrent transmission not involving i and j. To characterize the condition of non-interference, the average SINR of link (i,j) is given [18] by

$$\gamma_{ij}(\boldsymbol{P}) = \frac{h_{ij} p_{ij}}{\theta \sum_{i_1 \in S_{ij}} \sum_{(i_1,j_1) \in E^a} h_{i_1 j} p_{i_1 j_1} + \sigma_{ij}^2} \tag{4}$$

where σ_{ij}^2 is the thermal noise power associated with link (i,j), $\boldsymbol{P} = (p_{11}, p_{12}, \ldots, p_{ij}, \ldots)^T$ denotes the vector of transmission powers, h_{ij} indicates the link gain between transmitter i and receiver j, S_{ij} denotes the set of sensor nodes whose transmissions may interfere with the receiver of link (i,j), and θ is the orthogonality factor.

Furthermore, the capacity of link (i,j) can be determined by

$$C_{ij}(\boldsymbol{P}) = B \log(\varsigma \cdot \gamma_{ij}(\boldsymbol{P})) \tag{5}$$

where ς is "SINR-gap" that reflects a particular modulation and coding scheme. Without loss of generality, we assume that

$\varsigma = 1$. B is the base-band bandwidth, which is normalized by a fixed packet size.

For a data transmission, the flow rate f_{ij}^a over link (i, j) at anchor point a is restricted by link capacity $C_{ij}(\boldsymbol{P})$, i.e.,

$$0 \leq f_{ij}^a \leq C_{ij}(\boldsymbol{P}), \forall (i,j) \in E^a, \forall a \in A \tag{6}$$

Constraint (6) implies that data is correctly received if $C(\boldsymbol{P})$ is greater than or equal to the ingress flow rate f. It can be seen from (5) that the capacity of link (i, j) is varying and determined by transmission power and channel condition.

E. Compatibility Constraint

We apply the SDMA technique to improve the throughput of the uploaded data of sensors to the SenCar by incorporating spatial channels. Let $H_i = [h_{i1}, h_{i2}]^T$ represent the link gain vector between sensor i and the two antennas of the SenCar, and any two compatible sensors, i and m, simultaneously upload data to the SenCar at the rate x_i and x_m, respectively. The received data at the SenCar can be given by

$$\mathbf{y} = H_i x_i + H_m x_m + \mathbf{n} \tag{7}$$

where $\mathbf{n}$ is independent and identically distributed (i.i.d.) $CN(0, \sigma^2 \mathbf{I}_2)$ channel noise. It is clear from (7) that a data stream suffers the interference from other data streams. To remove the interference, we let U_i and U_m be the filter vectors for sensors i and m, respectively, which satisfies $U_i^T H_m = 0$ and $U_m^T H_i = 0$, i.e., U_i is any vector in the space orthogonal to H_m and U_m is any vector in the space orthogonal to H_i. After the filtering process by the two filters, data rates x_i and x_m can be separated from each other. Similar to [15], without loss of generality, U_i and U_m can be chosen as a unit vector

$$\begin{aligned} U_i &= \left(\sqrt{|h_{m2}|^2 + |h_{m1}|^2}\right)^{-1} [h_{m2}, -h_{m1}]^T \\ U_m &= \left(\sqrt{|h_{i2}|^2 + |h_{i1}|^2}\right)^{-1} [h_{i2}, -h_{i1}]^T \end{aligned}$$

As a result, for a given transmission power of each sensor, not any two sensors can successfully transmit to the SenCar with two antennas simultaneously. To guarantee that the SenCar can successfully separate the received signals, the following criteria should be satisfied [15]

$$SNR_i = p_i \left|U_i^T H_i\right|^2 / \sigma^2 \geq \delta_1 \tag{8}$$

$$SNR_m = p_m \left|U_m^T H_m\right|^2 / \sigma^2 \geq \delta_1 \tag{9}$$

where p_i, p_m, SNR_i and SNR_m are the transmission powers and SNR of the data from sensors i and m respectively. δ_1 indicates the SINR threshold for the SenCar to correctly decode the received data. Any two sensors that satisfy these criteria can successfully make concurrent data uploading to the SenCar. Such a sensor pair (i, m) is referred to as *compatible pair*.

F. Cross-Layer Optimization Model

By combining the optimization objective with constraints, the cost minimization (DaGCM) of the mobile data gathering with two antennas can be formulated as the following joint cross-layer optimization problem

$$\textbf{P1:} \min \sum_{i \in N} NC_i \left(\sum_{a \in A} R_i^a t^a \right) \tag{10}$$

subject to constraints (1), (2), (3), (6), (8), (9) and $\sum_{a \in A} t^a \leq T$. The last constraint ensures that the total sojourn time at all anchor points is no more than the bound of data gathering latency T.

The objective of the data gathering problem is to minimize the total data gathering cost or energy consumption, which can be achieved by dynamically allocating the flow rates over links and the data rates of sensors at the transport layer, controlling the transmission powers of sensors at the physical layer and adjusting the sojourn time of the SenCar in a distributed manner.

In general, the DaGCM problem (10) is non-convex since there exist the couplings among variables R_i^a, t^a, f_{ij}^a and p_i in both objective function (10) and constraint (2). In other words, the objective function is not convex with respect to R_i^a and t^a since its Hessian is not positively semidefinite [19]. Therefore, the problem needs to be converted into an equivalent convex problem under some conditions by employing auxiliary variables and appropriate transformation. Define $x_{ij}^a = f_{ij}^a t^a$, $y_i \phi_i^a = R_i^a t^a$, $\phi_i^a \geq 0$, $\sum_{a \in A} \phi_i^a = 1$, where x_{ij}^a denotes the data amount over link (i, j) destined to the SenCar at anchor point a, which can be considered as a routing variable, y_i represents the total amount of data generated by sensor i in a data gathering tour, and ϕ_i^a is a data split variable with $\phi_i^a \geq 0$, which determines the proportion of the data uploaded by sensor i to the SenCar at anchor point a over the total amount of data generated by sensor i in a gathering tour.

By multiplying the flow conservation and link capacity constraints by t^a, the original optimization problem (**P1**) can be transformed into a convex optimization problem with respect to x, y, ϕ, t and p as follows

$$\textbf{P2:} \ \min \sum_{i \in N} NC_i (y_i) \tag{11}$$

subject to

$$\sum_{j:(i,j) \in E^a} x_{ij}^a - \sum_{j:(j,i) \in E^a} x_{ji}^a = y_i \phi_i^a, \forall i \in N, \forall a \in A \tag{12}$$

$$\sum_{j:(i,j) \in E^a} x_{ij}^a e_{ij} \rho(i,m) \leq W_i, \forall i, m \in N \tag{13}$$

$$\sum_{n:(m,n) \in E^a} x_{mn}^a e_{mn} \rho(i,m) \leq W_m, \forall i, m \in N \tag{14}$$

$$0 \leq x_{ij}^a \leq C_{ij}(\boldsymbol{P}) t^a, \forall (i,j) \in E^a, \forall a \in A \tag{15}$$

$$min\{SNR_i, SNR_m\} \geq \delta_1, \forall i, m \in N \tag{16}$$

$$\sum_{a \in A} t^a \leq T, \ \sum_{a \in A} \phi_i^a = 1 \tag{17}$$

with

$$\sum_{i,m \in N} \rho(i,m) = 1, \phi_i^a \geq 0, \forall i \in N, \forall a \in A$$

where constraint (16) indicates that filters U_i and U_m are for not only the channel between the sensor itself and the

SenCar but also the channel between its compatible peer and the SenCar.

Since the objective function $NC_i(\cdot)$ is strictly convex with respect to y_i and the constraint set composed by (12)-(17) is also convex, the transformed DaGCM problem (**P2**) is a strictly convex optimization problem, however, with nonlinear constraints, which is induced by the variable coupling of constraints (12) and (15). It is easy to see that the Slater condition [20] is satisfied, the primal and dual problems have optimal solutions. Thus, it can be efficiently tackled by using the convex programming technique.

IV. Distributed Algorithms for DaGCM Problem

In this section, we propose distributed algorithms for data control, routing, sojourn time and power control by using the Layering as Optimization Decomposition (LOD) [21] approach with Lagrange-duality [20] to solve problem (**P2**) in a distributed manner. The necessary conditions for applying LOD are convexity and separability. However, problem (**P2**) is not separable on optimization variables in its dual because of the coupling between variables y_i and ϕ_i^a in the flow conservation constraint and between variables p_{ij} and t^a in the link capacity constraint. To decompose the coupling, we let $\widetilde{x}_{ij}^a = \log x_{ij}^a$, take logarithm transformation for constraints (12) and (15) and obtain the transformed constraints as follows

$$\log\Big(\sum_{j:(i,j)\in E^a} \exp(\widetilde{x}_{ij}^a) - \sum_{j:(j,i)\in E^a} \exp(\widetilde{x}_{ji}^a)\Big) \le \log y_i + \log \phi_i^a \quad (18)$$

for $\forall i \in N$, and

$$\widetilde{x}_{ij}^a \le \log C_{ij}(\boldsymbol{P}) + \log t^a \quad (19)$$

for $\forall (i,j) \in E^a$, $\forall a \in A$.

We then have the following theorem.

Theorem 1: After a log transformation of x in equality (12) and inequality (15), the transformed DaGCM problem (**P2**) with constraints (13)-(14) and (16)-(19) is a convex optimization problem.

Proof: It is clear that objective function $NC_i(y_i)$ is a convex function with regard to y_i. Constraints (13)-(14) and (16)-(19) are also convex since they all are linear functions. We further find that the left and right sides of (18) are log-sum-exp and (log, x) terms, respectively, which implies that constraint (18) is convex [20]. Similarly, since term $\log C_{ij}(\boldsymbol{P})$ is (log, log)-concave and term $\log t^a$ is (log, t)-concave, constraint (19) is convex. Hence the DaGCM problem (**P2**) is convex. ■

A. Lagrangian Dual Decomposition

Note that data routing variables x, total data amount variables y_i for sensor i at a data gathering tour and data split variable ϕ_i^a are coupled by inequality (18), while data routing variables x, power allocation variables p and sojourn time variables t^a are correlated each other in (19). We use the Lagrangian dual method to decouple them in the dual problem.

We first introduce Lagrangian Multipliers λ_i^a, $\mu_{(i,m)}$, υ_{ij}, ξ_i, η_m, $i,m \in N$, $(i,j) \in E^a$, $a \in A$ for sensors and links with respect to constraints (18), (13)-(14), (19) and (16) , respectively. Here, λ_i^a, $\mu_{(i,m)}$ and υ_{ij} can be interpreted as flow conservation cost for each sensor node $i \in V^a$ at anchor point $a \in A$, the cost of the energy consumed by compatible sensor pair (i,m) at anchor point a no more than their total energy budget, and the cost for keeping the outgoing rate no more than the link's average capacity over link (i,j) at anchor point a, respectively. Moreover, ξ_i and η_m can be referred to as the costs of power control required by sensors i and m to be compatible. The partial Lagrangian of the DaGCM problem (**P2**) can then be expressed as

$$L(x,y,\phi,t,p,\lambda,\mu,\upsilon,\xi,\eta) = L_y(y,\lambda) + L_x(x,\lambda,\mu,\upsilon) + L_\phi(\phi,\lambda) + L_t(t,\upsilon) + L_p(p,\eta,\upsilon,\xi)$$

where

$$L_y(y,\lambda) = \sum_{i\in N} NC_i(y_i) - \sum_{a\in A}\sum_{i\in N} \lambda_i^a \log y_i$$

$$L_x(\widetilde{x},\lambda,\mu,\upsilon) = \sum_{a\in A}\sum_{i,j\in N} \upsilon_{ij}\widetilde{x}_{ij}^a + \sum_{a\in A}\sum_{i,m\in N} \mu_{(i,m)} F\left(\widetilde{x}_{ij}^a, \widetilde{x}_{mn}^a\right)$$
$$+ \sum_{a\in A}\sum_{i\in N} \lambda_i^a \log\left(\sum_{j:(i,j)\in E^a} \exp(\widetilde{x}_{ij}^a) - \sum_{j:(j,i)\in E^a} \exp(\widetilde{x}_{ji}^a)\right)$$

$$L_p(p,\eta,\upsilon,\xi) = \sum_{i\in N} \xi_i \left(\delta_1 - p_i \left|U_i^T H_i\right|^2/\sigma^2\right) -$$
$$\sum_{a\in A}\sum_{(i,j)\in E^a} \upsilon_{ij} \log C_{ij}(\boldsymbol{P}) + \sum_{m\in N} \eta_m \left(\delta_1 - p_m \left|U_m^T H_m\right|^2/\sigma^2\right)$$

$$L_\phi(\phi,\lambda) = -\sum_{a\in A}\sum_{i\in N} \lambda_i^a \log \phi_i^a$$

$$L_t(t,\upsilon) = -\sum_{a\in A}\sum_{(i,j)\in E^a} \upsilon_{ij} \log t^a$$

where

$$F\left(\widetilde{x}_{ij}^a, \widetilde{x}_{mn}^a\right) = -W_i - W_m$$
$$+ \left(\sum_{j:(i,j)\in E^a} \exp(\widetilde{x}_{ij}^a) e_{ij} + \sum_{n:(m,n)\in E^a} \exp(\widetilde{x}_{mn}^a) e_{mn}\right) \rho(i,m)$$

We define the objective function of the Lagrangian dual problem as

$$D(\lambda,\mu,\upsilon,\xi,\eta) = \min_{\lambda,\mu,\upsilon,\xi,\eta\ge 0} L(x,y,\phi,t,p,\lambda,\mu,\upsilon,\xi,\eta)$$

with constraint (17). By the linearity of the differentiation operator, the objective can be decomposed into five separate minimization subproblems due to the separable nature of the function.

$$DP1: D_y(\lambda) = \min_{\lambda\ge 0} L_y(y,\lambda)$$
$$DP2: D_x(\lambda,\mu,\upsilon) = \min_{\lambda,\mu,\upsilon\ge 0} L_x(\widetilde{x},\lambda,\mu,\upsilon)$$
$$DP3: D_p(\eta,\upsilon,\xi) = \min_{\upsilon,\xi,\eta\ge 0} L_p(p,\eta,\upsilon,\xi)$$
$$\text{s.t. } p_{ij}^{\min} \le p_{ij} \le p_{ij}^{\max}, \forall (i,j) \in E^a, \forall a \in A$$
$$DP4: D_\phi(\lambda) = \min_{\lambda\ge 0} L_\phi(\phi,\lambda), \text{ s.t. } \sum_{a\in A} \phi_i^a = 1$$
$$DP5: D_t(\upsilon) = \min_{\lambda\ge 0} L_t(t,\upsilon) \text{ s.t. } \sum_{a\in A} t^a \le T$$

It can be seen that the minimization problem in the dual function has been decomposed into several simple subproblems, which implies that the optimal DaGCM problem can be separated and solved in a distributed manner.

Accordingly, by duality, the dual problem of the primal problem is given by

$$\max_{\lambda,\mu,\upsilon,\xi,\eta\geq 0} D(\lambda,\mu,\upsilon,\xi,\eta). \tag{20}$$

Because the primal problem is convex and also satisfies the Slater's condition [20], we can design primal-dual algorithms to find the optimal data amount, routing, power, data split amount, sojourn time and costs, which we will discuss next.

B. Data Control Subalgorithm (DCSA)

The data control subalgorithm aims to determine the optimal amount of data generated by each sensor in a data gathering tour by solving the *DP1* problem. Since the *DP1* is convex, a subgradient projection with a sufficiently small step-size can be employed to solve the problem. Taking the partial derivative of $L_y(y,\lambda)$ with respect to y, we have the subgradient on y as follows

$$\nabla D_1(y_i) = NC_i^{'}(y_i) - \lambda_i^a/y_i, \forall i \in N, \forall a \in A$$

where $NC_i^{'}(y_i)$ denotes the first-order derivative of $NC_i(y_i)$ with regard to y_i. Similarly, the subgradient on the cost λ_i^a for sensor i at anchor point a is given by

$$\nabla D_1(\lambda_i^a) = -\log y_i, \forall i \in N, \forall a \in A$$

Accordingly, the data amount variables y_i and the cost variables λ_i^a can be updated according to

$$y_i(k+1) = [y_i(k) + \varepsilon(k)\nabla D_1(y)]^+ \tag{21}$$

$$\lambda_i^a(k+1) = [\lambda_i^a(k) - \varepsilon(k)\nabla D_1(\lambda_i^a)]^+ \tag{22}$$

where ε is the step size and $[a]^+ = \max(0,a)$.

Theorem 2: The update iterations of y_i and λ_i^a by (21) and (22) will converge to the optimal solution (y_i^*, λ_i^{a*}) provided that a sequence of the step sizes satisfies

$$\varepsilon(k) \to 0, \sum_{k=1}^{\infty}\varepsilon(k) \to \infty \text{ and } \sum_{k=1}^{\infty}\varepsilon(k)^2 < \infty$$

Proof: Due to the space limit, we omit the proof here. A similar proof can be found in [20]. ■

C. Routing Subalgorithm (RSA)

We now proceed to solve *DP2* problem, i.e., determine the routing of data from sensors to the SenCar so as to adjust the flow amount over sensors' outgoing links destined to each anchor point. Taking the derivatives of $L_x(x,\lambda,\mu,\upsilon)$ with respect to $\widetilde{x}$, μ and υ, we obtain the subgradients with respect to $\widetilde{x}$, μ and υ as follows

$$\nabla D_2(\widetilde{x}_{ij}) = \mu_{(i,m)}\rho(i,m)G(\widetilde{x}_{ij}^a, \widetilde{x}_{mn}^a) + \sum_{j\in N}\upsilon_{ij} + \lambda_i^a H(\widetilde{x}_{ij}^a)$$

$$\nabla D_2(\mu_{(i,m)}) = F\left(\widetilde{x}_{ij}^a, \widetilde{x}_{mn}^a\right)$$

$$\nabla D_2(\upsilon_{ij}) = \widetilde{x}_{ij}^a$$

where

$$G(\widetilde{x}_{ij}^a, \widetilde{x}_{mn}^a) = \sum_{j:(i,j)\in E^a}\exp(\widetilde{x}_{ij}^a)e_{ij} + \sum_{n:(m,n)\in E^a}\exp(\widetilde{x}_{mn}^a)e_{mn}$$

$$H(\widetilde{x}_{ij}^a) = \frac{\sum\limits_{j^{'}:(j^{'},i^{'})\in E^a\setminus(i,j)}\exp(\widetilde{x}_{j^{'}i^{'}}^a) - \sum\limits_{j^{'}:(i^{'},j^{'})\in E^a\setminus(i,j)}\exp(\widetilde{x}_{i^{'}j^{'}}^a)}{\sum\limits_{j:(j,i)\in E^a}\exp(\widetilde{x}_{ji}^a) - \sum\limits_{j:(i,j)\in E^a}\exp(\widetilde{x}_{ij}^a)}$$

By using the subgradient projection method to solve the dual problem (20), the routing variables $\widetilde{x}_i$ and the costs $\mu_{(i,m)}$ and υ_{ij} can be updated by

$$\widetilde{x}_{ij}(k+1) = [\widetilde{x}_{ij}(k) + \varepsilon(k)\nabla D_2(\widetilde{x}_{ij})]^+ \tag{23}$$

$$\mu_{(i,m)}(k+1) = \left[\mu_{(i,m)}(k) - \varepsilon(k)\nabla D_2(\mu_{(i,m)})\right]^+ \tag{24}$$

$$\upsilon_{ij}(k+1) = [\upsilon_{ij}(k) - \varepsilon(k)\nabla D_2(\upsilon_{ij})]^+ \tag{25}$$

It is clear that the optimal routing of the data from sensors to SenCar depends on the incoming and outgoing data amount of sensors, which is related to the costs of the energy consumed by the compatible sensor pairs and the congestion. Thus the cost λ_i^a plays a critical role in the balance between the amount of the data generated by each sensor and the routing of the data from sensors to SenCar.

D. Power Control and Compatibility Subalgorithm (PCSA)

The goal of this subalgorithm is to give a distributed protocol for optimal power allocation in the physical layer. The objective for power control is to ensure the two sensors to be compatible while determining suitable link capacities. This objective can be achieved by solving the *DP3* problem. Note that *DP3* is also convex. Similar to the above subalgorithm, the subgradients with respect to p, ξ and η can be given by

$$\nabla D_3(p_i) = -\xi_i\left|U_i^T H_i\right|^2/\sigma^2 - \upsilon_{ij}C_{ij}^{'}(\boldsymbol{P})/C_{ij}(\boldsymbol{P})$$

$$\nabla D_3(p_m) = -\eta_m\left|U_m^T H_m\right|^2/\sigma^2 - \upsilon_{mn}C_{mn}^{'}(\boldsymbol{P})/C_{mn}(\boldsymbol{P})$$

$$\nabla D_3(\xi_i) = \delta_1 - p_i\left|U_i^T H_i\right|^2/\sigma^2$$

$$\nabla D_3(\eta_m) = \delta_1 - p_m\left|U_m^T H_m\right|^2/\sigma^2$$

where $C_{ij}^{'}(\boldsymbol{P})$ denotes the derivative of link capacity $C_{ij}(\boldsymbol{P})$ with respect to power p_i.

The primal-dual method [20] can be used to solve the *DP3* problem, i.e., the power allocation variables and the costs can be updated by

$$p_i(k+1) = [p_i(k) + \varepsilon(k)\nabla D_3(p_i)]_{p_{ij}^{\min}}^{p_{ij}^{\max}} \tag{26}$$

$$p_m(k+1) = [p_m(k) + \varepsilon(k)\nabla D_3(p_m)]_{p_{mn}^{\min}}^{p_{mn}^{\max}} \tag{27}$$

$$\xi_i(k+1) = [\xi_i(k) - \varepsilon(k)\nabla D_3(\xi_i)]^+ \tag{28}$$

$$\eta_m(k+1) = [\eta_m(k) - \varepsilon(k)\nabla D_3(\eta_m)]^+ \tag{29}$$

where

$$[a]_{p^{\min}}^{p^{\max}} = \min\left(p^{\max}, \max\left(a, p^{\min}\right)\right)$$

In the physical layer and network layer, respectively, routing and power control operate independently to update sensors' routing strategies and power allocation. The interface variables υ_{ij} are used to control the performance of the two subalgorithms and regulate the power allocation and routing strategies towards the optimal solution.

E. Sojourn Time Allocation and Data Split Subalgorithms

The SenCar is responsible for allocating the sojourn time for each anchor point to satisfy the optimization objective

$$\max\sum_{a\in A}\sum_{(i,j)\in E^a}\upsilon_{ij}\log t^a, \text{ s.t. } \sum_{a\in A}t^a \leq T$$

Since the optimization problem is convex, by employing the dual decomposition and introducing a new Lagrangian Multiplier ϖ^a, the sojourn time t^a and Lagrangian Multiplier ϖ^a can be updated by

$$t^a(k+1) = [t^a(k) - \varepsilon(k)\nabla D_4(t^a)]^+ \quad (30)$$

$$\varpi^a(k+1) = [\varpi^a(k) + \varepsilon(k)\nabla D_4(\varpi^a)]^+ \quad (31)$$

where $\nabla D_4(t^a) = \upsilon_{ij}/t^a + \varpi^a$ and $\nabla D_4(\varpi^a) = t^a - T$ represent the subgradients with respect to t^a and ϖ^a.

For the SenCar to determine the sojourn time, the values of the cost υ_{ij} need to be routed to the SenCar in each subgradient iteration. To avoid the communication overhead, an alternative method is to let each sensor determine the sojourn time at each anchor point. However, it results in slower convergence than the former. Thus, this is a tradeoff between communication overhead and convergence speed.

Similarly, the data split amount ϕ_i^a and Lagrangian Multiplier ω_i^a can be updated according to

$$\phi_i^a(s+1) = [\phi_i^a(s) - \varepsilon(s)\left(\lambda_i^a/\phi_i^a(s) + \omega_i^a(s)\right)]^+ \quad (32)$$

$$\omega_i^a(s+1) = [\omega_i^a(s) + \varepsilon(k)\left(\phi_i^a(s) - 1\right)]^+ \quad (33)$$

In addition, for sensor i to determine the data amount uploaded to the SenCar at anchor point a, the cost message λ_i^a should be sent from the data control subalgorithm to the data split subalgorithm of sensor i. Note that the information required for these updates includes: measurements of average SINR and channel link gain of each link for PCSA, and ingress rate of each link for RSA. These updates can be incorporated into a distributed protocol by making use of explicit message passing.

To summarize, the proposed algorithm is implemented by a three-layer Lagrange-duality method. At the first layer, the DaGCM algorithm searches iteratively for $\{\lambda_i^{a*}\}$ and $\{y_i^*\}$ with which the optimal amount of data generated by each sensor in a data gathering tour is determined. In each iteration, an update for $\{\lambda_i^{a*}\}$ is generated and then passed to the second layer where the routing subalgorithm starts a parallel search for $\{\widetilde{x}_{ij}^*\}$, $\{\mu_{(i,m)}^*\}$ and $\{\upsilon_{ij}^*\}$ to determine the optimal routing of the data from sensors to the SenCar while an update of $\{\lambda_i^{a*}\}$ and $\{\upsilon_{ij}^*\}$ is passed to the third layer where a parallel search for $\{\phi_i^{a*}\}$, $\{p_{ij}^*\}$, $\{t^{a*}\}$, $\{\xi_i^*\}$ and $\{\eta_m^*\}$ is performed to determine optimal transmission power, ensure the compatibility of the two sensors and give the optimal sojourn time of the SenCar at an anchor point. The distributed algorithm for the DaGCM problem is summarized in Table I.

The complexity of the DaGCM algorithm can be derived as follows. It is known from [19] that the number of iterations required by the primal-dual method for convergence is a linear function of the problem size. Thus, the total complexity of the algorithm is $O\left(|A|\,|N|^2\right)$. Note that the solution exploration of each distributed subalgorithm only needs to be executed when the energy budget is updated, thus does not need to be frequently repeatedly executed by sensors.

TABLE I
DISTRIBUTED ALGORITHM FOR DAGCM PROBLEM

For each sensor $i \in N$ do
  Initialize T, $p_{ij}^{\min}$ and $p_{ij}^{\max}$ for link $(i,j) \in E^a$, and
    $\phi_i^a(0)$ such that $\sum_{a\in A}\phi_i^a(0) = 1$, for all $a \in A$;
  Repeat
    Initialize Lagrangian multipliers $\lambda_i^a(0), \mu_{(i,m)}(0), \upsilon_{ij}(0), \xi_i(0)$
      and $\eta_m(0)$ to non-negative values, for all j: $(i,j) \in E^a$,
      $m \in N$ and $a \in A$;
    Repeat: for all $j : (i,j) \in E^a$, $n : (m,n) \in E^a$ and $a \in A$
      /* **Data control subalgorithm***/
      Compute $y_i(k)$ by Eq. (21);
      Update Lagrangian multipliers $\lambda_i^a(k+1)$ by Eq. (22);
      Send the updated λ_i^a to RSA in network layer and data split
        subalgorithm in transport layer;
      /***Routing subalgorithm***/
      Obtain the ingress rate of each link in E^a;
      Compute $\widetilde{x}_{ij}(k)$ by Eq. (23);
      Update Lagrangian multipliers $\mu_{(i,m)}(k+1)$ by Eq. (24);
      Update Lagrangian multipliers $\upsilon_{ij}(k+1)$ by Eq. (25);
      Send the updated υ_{ij} to power allocation subalgorithm
        in physical layer and the SenCar;
      /***Power control and compatibility subalgorithm***/
      Obtain the average SINR and link gain of each link in E^a;
      Compute $p_i(k)$ by (26);
      Compute $p_m(k)$ by (27);
      If compatibility constraint is satisfied, sensors i and m are
        a compatible pair and can transmit simultaneously;
      Update Lagrangian multipliers $\xi_i(k+1)$ by (28);
      Update Lagrangian multipliers $\eta_m(k+1)$ by (29);
      /***Sojourn time allocation subalgorithms***/
      SenCar simultaneously computes $t^a(k)$ by (30);
      SenCar updates Lagrangian multiplier $\varpi^a(k+1)$ by (31);
    Until $\{\lambda(k)\}$ converges to λ^* and $\{y_i(k)\}$ converges to y^*;
    Adjust data split variables $\phi_i^a(s+1)$, $a \in A$ by (32);
    Update Lagrangian multiplier $\omega_i^a(s+1)$, $a \in A$ by (33);
  Until reach the equilibrium, i.e., $\{\phi(s)\}$ converges to ϕ^*.
End For

TABLE II
PARAMETER SETTINGS

Notation	Value	Notation	Value
W_i ,$\forall i \in N$	1.25*10^4	T	80
$p_{ij}^{\min}, (i,j) \in E^a$	1mw	σ_{ij}	10^{-7}
$p_{ij}^{\max}, (i,j) \in E^a$	100mw	B	1M
e_{ij}	0.007d_{ij}^2	$\rho(i,m)$	0.05*$max(x_i, x_m)$

V. PERFORMANCE EVALUATION

In this section, we provide some numerical results to demonstrate the convergence and efficiency of the proposed distributed algorithm and compare its performance with that of the algorithm without concurrent uploading and power control [14] in terms of the total sojourn time and the total energy consumption. We consider a WSN, where a total of 30 sensors are randomly scattered over a $100m \times 100m$ square area and 4 anchor points are selected, as shown in Fig. 1. We set the radius of the coverage area of each anchor point to 30m and assume that each link experiences Rayleigh fading. The moving velocity of the SenCar is 0.8m/sec. In this simulation, for clarity, we take anchor points a_1 and a_2 as the observation objects and other parameter settings are listed in Table II. All performance measures are obtained by averaging over 100 simulation runs.

A. Convergence and Performance Analysis

In this subsection, we observe the convergence property and analyze the performance of the proposed DaGCM algorithm. We define the cost function as $NC_i = w_i\left(\sum_a R_i^a t^a\right)^2$ or equivalently $NC_i = w_i\left(y_i\right)^2$, where w_i indicates the cost

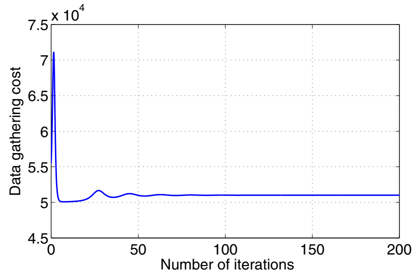

Fig. 2. Evolution of data gathering cost vs. number of iterations.

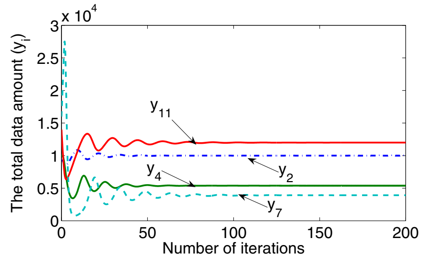

Fig. 3. Evolution of total amount of data generated vs. number of iterations.

weight for sensor i to upload data to the SenCar at a data gathering tour.

We examine the total data gathering costs of sensors 2 and 4 at anchor point a_1 and sensors 7 and 11 at anchor point a_2, whose cost weights w_2, w_4, w_7 and w_{11} are set to 4000, 7800, 8000 and 2000, respectively. Fig. 2 shows the evolution of the total data gathering cost at the two anchor points versus the number of iterations. It can be seen from Fig. 2 that the data gathering cost oscillates greatly at the beginning of iterations, then fluctuates slightly, and finally converges. In other words, the data gathering cost achieves the optimum after about 60 iterations. The reason for the fluctuation is that we use the incremental subgradient method to solve the DaGCM problem, which can lead to fluctuation when the solution is close to the optimum and the degree of the fluctuation is proportional to the iteration stepsize [19]. As the stepsize becomes smaller and smaller, the algorithm trends to be stable and achieves convergence.

Fig. 3 shows the evolution of total amount of data y_i generated from sensor i (i=2,4,7,11) in a data gathering tour with the number of iterations. It can be observed that since sensor 11 has a smaller cost weight compared to other sensors, in a data gathering tour, the SenCar preferentially gathers more data from it. In particular, y_7 has a larger fluctuation and a longer convergence time. This is because that in addition to the fluctuation induced by the incremental subgradient method, sensor 7 has to relay the data from sensors 4 and 6 while the available capacity of link $(7, a_2)$ is restricted by the link capacity constraint. We further observe that although sensors 4 and 7 have almost the same cost weight, the amount of data generated by sensor 4 is clearly more than the amount of data generated by sensor 7. It is justifiable since sensor 7 has to restrain from generating more data and reserve the suitable buffer size to avoid congestion, which in turn leads to that the SenCar would stay longer time at anchor point a_2 than at anchor point a_1. Thus, the optimal sojourn time for the two anchor points is t^{a_1*} =23.4 and t^{a_2*} =56.2, respectively.

TABLE III
OPTIMAL LINK FLOW AND NODE POWER

Anchor Point a_1				Anchor Point a_2			
$\widetilde{x}^{a_1*}_{(2,a_1)}$	5628	p_2^*	24	$\widetilde{x}^{a_2*}_{(7,a_2)}$	4512	p_7^*	52
$\widetilde{x}^{a_1*}_{(4,a_1)}$	4152	p_4^*	45	$\widetilde{x}^{a_2*}_{(11,a_2)}$	6854	p_{11}^*	28
$\widetilde{x}^{a_1*}_{(7,4)}$	2213	p_7^*	21	$\widetilde{x}^{a_2*}_{(4,7)}$	2721	p_4^*	23

Table III lists the optimal flow amount $\widetilde{x}^{a}_{ij}$ for each outgoing link and the optimal powers p_is allocated for sensors 2, 4, 7 and 11. It can be found that since the sojourn time can be adjusted dynamically among the different anchor points, more data from the sensors with smaller cost weights would be gathered by the SenCar. For example, sensor 4 has almost twice of the cost weight of sensor 2, as a result, the SenCar would collect 35.5% more data from sensor 2 than from sensor 4 at anchor point a_1. Furthermore, we observe that more data from sensor 7 are gathered by the SenCar at anchor point a_2. This is because that sensor 7 has to relay the data from sensor 4 to the SenCar over link $(7, a_2)$ at anchor point a_2. In particular, the gathered data amount from sensor 7 is nearly the same as that from sensor 11 in a data gathering tour, however, the energy consumed by the former is more than twice of the latter.

The compatible pairs determined by the proposed DaGCM algorithm are shown in Fig. 1. From the above simulation results, we find that whether the two sensors are compatible or not depends largely on the physical distance between the two sensors and also depends on the powers allocated to the two sensors. The reason is that for a deterministic fading model, wireless link gain is determined by the distance between the transmitter and the receiver of this link, the path loss exponent and the radio propagation properties of the environment. In practice, the last two factors also determine the cost weight of a sensor.

B. Performance Comparison

We now compare the performance of the proposed DaGCM algorithm with the pricing-based algorithm in [14]. In order to facilitate the comparison, we use the same network and parameter settings in the DaGCM algorithm as those in [14]. Fig. 4 plots the comparison of the total sojourn time between the two algorithms when the data gathering cost is varied from 1.0×10^6 to 5.0×10^6 for $T = 450$. We can observe that for a given data gathering cost, the total sojourn time in the DaGCM algorithm is remarkably less than that in the pricing-based algorithm in [14]. It is justifiable since the two compatible sensors can upload their data to the SenCar concurrently to shorten the sojourn time of the SenCar at an anchor point while the power control subalgorithms are executed to save the energy of sensors in DaGCM, which reveals the benefits of employing the SenCar with two antennas and the SDMA technique.

Fig. 5 shows the total energy consumption comparison in a data gathering tour between the DaGCM algorithm and the pricing-based algorithm in [14] when the bound of data gathering latency T is varied from 50s to 300s. We assume that the total uploaded data amount of all the sensors is identical in the two algorithms. From the figure, we can draw some

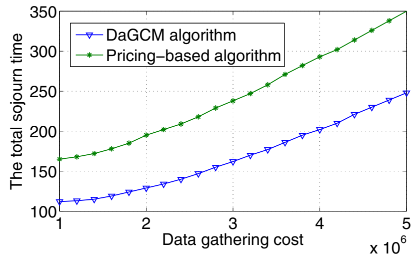

Fig. 4. Comparison of the total sojourn time under given data gathering cost.

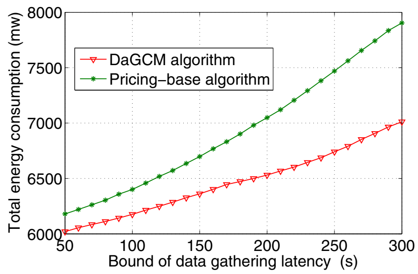

Fig. 5. Total energy consumption vs. data gathering latency bound.

observations. First, the increase of the total energy consumption in the DaGCM algorithm is clearly slower than that in the pricing-based algorithm. Second, the DaGCM algorithm always achieves lower energy consumption, which plays a critical role in prolonging network lifetime. For example, when T =200, the DaGCM algorithm results in 7.3% less total energy consumption with respect to the pricing-based algorithm. The underlying reason for such superiority of the DaGCM algorithm is that each sensor can dynamically adjust its transmission power, adaptively split its data and then send the data to the SenCar at different neighboring anchor points with the optimal transmission power.

VI. Conclusions

In this paper, we have addressed a cross-layer design problem for the WSNs involving data control, routing, power control and compatibility decisions. By utilizing the SenCar with two antennas and the SDMA technique, we first formulate the mobile data gathering problem with concurrent data uploading as the DaGCM problem under the constrains of flow conservation, energy consumption, link capacity and compatibility among sensors, whose objective is to shorten the data gathering time and reduce the energy consumption. Then, by introducing some auxiliary variables, we transform the non-convex DaGCM problem into a convex one and further decompose it into several subproblems of data control and data split at the transport layer, routing at the network layer, and power control and compatibility decisions at the physical layer. We also provide the optimal solution to the sojourn time allocation subproblem for determining the optimal sojourn time of SenCar at different anchor points. Moreover, by utilizing the subgradient iterative approach, we present several corresponding distributed subalgorithms with explicit message passing. Numerical results illustrate that the proposed DaGCM algorithm converges and outperforms the algorithm without concurrent data loading and power control in terms of data gathering latency and energy consumption.

VII. Acknowledgments

The work described in this paper was supported by the grants from the National Natural Science Foundation of China (60903213, 61170248), New Century Excellent Talents in University (No. NCET-10-0877), the Postdoctoral Science Foundation of China (201003314, 20090460706), US NSF under grant number ECCS-0801438 and US Army Research Office under grant number W911NF-09-1-0154.

References

[1] I. Vasilescu, K. Kotay, D. Rus, M. Dunbabin and P. Corke, "Data collection, storage, and retrieval with an underwater sensor network," *Proc. ACM SenSys*, 2005, pp. 154-165.

[2] D. England, B. Veeravalli and J. B. Weissman, "A robust spanning tree topology for data collection and dissemination in distributed environments," *IEEE Trans. Parallel Distrib. Syst.*, vol. 18, pp. 608-620, 2007.

[3] E. Lee, S. Park, F. Yu and S.-H. Kim, "Data gathering mechanism with local sink in geographic routing for wireless sensor networks," *IEEE Trans. Consum. Electron.*, vol. 56, no. 3, pp. 1433-1441, Aug. 2010.

[4] Z. Zhang, M. Ma and Y. Yang, "Energy-efficient multihop polling in clusters of two-layered heterogeneous sensor networks," *IEEE Trans. Comput.*, vol. 57, no. 2, pp. 231-245, Feb. 2008.

[5] Y. Wu, Z. Mao, S. Fahmy and N. Shroff, "Constructing maximum-lifetime data-gathering forests in sensor networks," *IEEE/ACM Trans. Netw.*, vol. 18, no. 5, pp. 1571-1584, Oct. 2010.

[6] W. Zhao, M. Ammar and E. Zegura, "A message ferrying approach for data delivery in sparse mobile ad hoc networks," *Proc. ACM MobiHoc*, pp. 187-198, 2004.

[7] J. Luo and J.-P. Hubaux, "Joint mobility and routing for lifetime elongation in wireless sensor networks," *Proc. IEEE INFOCOM*, vol. 3, pp. 1735-1746, March 2005.

[8] M. Ma and Y. Yang, "Sencar: An energy-efficient data gathering mechanism for large-scale multihop sensor networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 18, no. 10, pp. 1476-1488, Oct. 2007.

[9] M. Ma and Y. Yang, "Data gathering in wireless sensor networks with mobile collectors," *Proc. IEEE IPDPS*, pp. 1-9, April 2008.

[10] G. Xing, T. Wang, W. Jia and M. Li, "Rendezvous design algorithms for wireless sensor networks with a mobile base station," *Proc. ACM MobiHoc*, 2008, pp. 231-240.

[11] A. Sharifkhani and N. Beaulieu, "A mobile-sink-based packet transmission scheduling algorithm for dense wireless sensor networks," *IEEE Trans. Veh. Technol.*, vol. 58, no. 5, pp. 2509-2518, Jun. 2009.

[12] R. Sugihara and R. Gupta, "Optimal speed control of mobile node for data collection in sensor networks," *IEEE. Trans. Mob. Comput.*, vol. 9, no. 1, pp. 127-139, Jan. 2010.

[13] M. Zhao and Y. Yang, "Optimization based distributed algorithms for mobile data gathering in wireless sensor networks," to appear in *IEEE Trans. Mob. Comput.*.

[14] M. Zhao, D. Gong and Y. Yang, "A cost minimization algorithm for mobile data gathering in wireless sensor networks," *IEEE MASS*, Nov. 2010, pp. 322-331.

[15] M. Zhao, M. Ma and Y. Yang, "Efficient data gathering with mobile collectors and space-division multiple access technique in wireless sensor networks," *IEEE Trans. Comput.*, vol. 60, no. 3, pp. 400-417, Mar. 2011.

[16] M. Gatzianas and L. Georgiadis, "A distributed algorithm for maximum lifetime routing in sensor networks with mobile sink," *IEEE Trans. Wirel. Commun.*, vol. 7, no. 3, pp. 984-994, March 2008.

[17] S. Nesamony, M. Vairamuthu and M. Orlowska, "On optimal route of a calibrating mobile sink in a wireless sensor network," in *Fourth International Conference on Networked Sensing Systems (INSS '07)*, June 2007, pp. 61-64.

[18] C. Long, B. Li, Q. Zhang, B. Zhao, B. Yang and X. Guan, "The end-to-end rate control in multiple-hop wireless networks: Cross-layer formulation and optimal allocation," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 4, pp. 719-731, May 2008.

[19] D. P. Bertsekas, *Nonlinear Programming, 2nd Edition*. Athena Scientific, 1999.

[20] S. Boyd and L. Vandenbergh, *Convex Optimization*. Cambridge University Press, 2004.

[21] M. Chiang, S. Low, A. Calderbank and J. Doyle, "Layering as optimization decomposition: A mathematical theory of network architectures," *Proc. of the IEEE*, vol. 95, no. 1, pp. 255-312, Jan. 2007.