# Distributed Algorithms for Maximum Lifetime Routing in Wireless Sensor Networks

Ritesh Madan and Sanjay Lall

***Abstract*— A sensor network of nodes with wireless transceiver capabilities and limited energy is considered. We propose distributed algorithms to compute an optimal routing scheme that maximizes the time at which the first node in the network drains out of energy. The problem is formulated as a linear programming problem and subgradient algorithms are used to solve it in a distributed manner. The resulting algorithms have low computational complexity and are guaranteed to converge to an optimal routing scheme that maximizes the network lifetime. The algorithms are illustrated by an example in which an optimal flow is computed for a network of randomly distributed nodes. We also show how our approach can be used to obtain distributed algorithms for many different extensions to the problem. Finally, we extend our problem formulation to more general definitions of network lifetime to model realistic scenarios in sensor networks.**

***Index Terms*— routing, linear programming, dual decomposition, subgradient algorithm, sensor networks, network lifetime.**

## I. Introduction

CONSIDER a network of wireless sensor nodes distributed in a region. Each node has a limited battery energy supply and can generate information that needs to be communicated to a sink node. It is assumed that each wireless node has the capability to relay packets. Also each node is assumed to be able to dynamically adjust its transmission power depending on the distance over which it transmits a packet. We focus on the problem of computing a flow that maximizes the *lifetime* of the network - the lifetime is taken to be the time at which the first node runs out of energy. Since sensor networks need to self configure in many situations, the goal of this paper is to find algorithms that do this computation in a *distributed manner*. We analyze a *partially distributed* algorithm and a *completely distributed* algorithm to compute such a flow. The algorithms described in this paper can be used in static networks, or in networks in which the topology changes slowly enough such that there is enough time between topology changes to optimally balance the traffic.

Manuscript received September 28, 2004; revised May 1, 2005; accepted June 29, 2005. The associate editor coordinating the review of this paper and approving it for publication was J. Zhang. A part of this work was presented at IEEE GLOBECOM, Dallas, TX, November, 2004. Both authors were partially supported by the Stanford URI *Architectures for Secure and Robust Distributed Infrastructures*, AFOSR DoD award number 49620-01-1-0365. Ritesh Madan is also partially supported by a Sequoia Capital Stanford Graduate Fellowship.

R. Madan is with the Department of Electrical Engineering, Stanford University, Stanford, CA 94305 (e-mail: rkmadan@stanford.edu).

S. Lall is with the Department of Aeronautics and Astronautics, Stanford University, Stanford, CA 94305 (e-mail: lall@stanford.edu).

Digital Object Identifier 10.1109/TWC.2006.04667.

### A. Prior Work

Energy efficient algorithms for routing in wireless networks have received considerable attention over the past few years. Distributed algorithms to form sparse topologies containing minimum-energy routes were proposed in [1], [2]. An approximate approach based on discretization of the coverage region of a node into cones was described in [3], [4].

All the above mentioned works focused on minimizing the total energy consumption of the network. However, as pointed out in [5], this can lead to some nodes in the network being drained out of energy very quickly. Hence instead of trying to minimize the total energy consumption, routing to maximize the network lifetime was considered in [6], [5]. The problem was formulated as a linear program, and heuristics were proposed to select routes in a distributed manner to maximize the network lifetime. However, as illustrated in these papers, these heuristics do not always lead to selection of routes that are globally optimum. A similar problem formulation for selection of relay nodes was given in [7].

We note that distributed iterative algorithms for the computation of the maximum lifetime routing flow were described in [8]. Each iteration involved a bisection search on the network lifetime, and the solution of a max-flow problem to check the feasibility of the network lifetime. The complexity of the algorithm was shown to be polynomial in the number of nodes in the special case of one source node. In this paper, we use a different approach based on the subgradient algorithm for the solution of the dual problem. We exploit the separable nature of the problem using dual decomposition to obtain partially and fully distributed algorithms. This is similar to the dual decomposition approaches applied to other problems in communication networks (see, for example, [9], [10], [11]).

### B. Outline

In Section II, we describe the system model and formulate the problem of maximizing the network lifetime as an optimization problem. The optimization problem we consider is the same as the one in [6], [5]. Section III gives a brief overview of the subgradient algorithm to solve a convex optimization problem via the dual problem. In Sections IV and V, we derive the partially and fully distributed algorithms. Section VI describes various extensions to the problem for which we can obtain distributed algorithms using the approach described in this paper. In Section VII, we extend the simplistic definition of network lifetime to more general definitions which model more realistic scenarios in sensor networks. Finally, in Section VIII, we conclude and give possible directions for future work.

## II. System Model

Consider a static wireless network modeled as an undirected graph $\mathcal{G}(V, L)$ where $V$ is the set of nodes, and $L$ is the set of links. Two nodes $i$ and $j$ are connected by a link if they can transmit a packet to each other with a transmission power less than the maximum transmission power at each node. Thus all links are assumed to be bi-directional. This assumption is not necessary for the convergence of the distributed algorithms; however it makes the presentation clearer. The set of nodes connected to node $i$ by a link is denoted as $N_i$. We assume that the network graph is connected, i.e. there always exists a path between any pair of nodes $i$ and $j$ in $V$.

We only consider the routing layer. We assume that the transmission rate $X_{ij}$ on each link from node $i$ to a neighboring node $j$ is fixed. Also, the maximum fraction of time $\tau_{ij}$ for which a link can transmit is fixed. In addition it is assumed that links which interfere with each other are never scheduled simultaneously. Thus the maximum flow that a link from node $i$ to node $j$ can support is $R_{ij} = X_{ij}\tau_{ij}$. For a fixed data rate, modulation scheme, and bit error rate (BER), the minimum receiver power needed for decoding can be determined (see for example, [12]). Along with the path loss model, the minimum receiver power can be used to determine the required transmission power $P_{ij}$ for data transmission at rate $X_{ij}$ between nodes $i$ and $j$. Thus if the data flow from node $i$ to node $j$ is $r_{ij}$, the average power consumption at node $i$ is

$$\frac{r_{ij}}{X_{ij}} P_{ij} = E_{ij} r_{ij}$$

where $E_{ij}$ is a constant depending only on the pair of nodes $i$ and $j$. We assume that the links are symmetric in the sense that $E_{ij} = E_{ji}$; our methods apply equally well to the asymmetric case as well.

### A. Problem Definition

Each node $i$ is assumed to have an initial battery energy $B_i$. Let $S_i$ be the rate at which information is generated at node $i$; this information needs to be communicated to the sink node. We will write $S_{\text{sink}} = -\sum_{i \in V, i \neq \text{sink}} S_i$. For a network flow $r$, let $r_{ij}$ denote the rate of information flow from node $i$ to node $j$. The energy spent by node $i$ to transmit a unit of information directly to node $j$ is $E_{ij}$. Then the lifetime of node $i$ under flow $r = \{r_{ij}\}$ is given by

$$T_i(r) = \frac{B_i}{\sum_{j \in N_i} E_{ij} r_{ij}}$$

We define the *network lifetime* $T_{\text{net}}$ under flow $r$ to be the time until the first node runs out of energy, i.e.

$$T_{\text{net}}(r) = \min_{i \in N} T_i(r)$$

Our goal is to find a distributed algorithm to compute a flow $r = \{r_{ij}\}$ that maximizes the network lifetime. Hence we aim to solve the following problem in a distributed manner [5].

$$\begin{array}{ll} \text{maximize} & T_{\text{net}}(r) \\ \text{subject to} & \displaystyle\sum_{j \in N_i} (r_{ij} - r_{ji}) = S_i, \quad \forall i \in V \\ & 0 \leq r_{ij} \leq R_{ij}, \quad \forall i \in V, \forall j \in N_i \end{array} \tag{1}$$

The constraints are the flow conservation equations at each node. The variables are $r_{ij}$ for $i \in V, j \in N_i$. Thus we have two flow variables for each undirected edge $(i, j)$ in the set $L$.

## III. Subgradient Algorithm for Convex Optimization Problems

Consider the following convex optimization problem, with variable $x \in D \subseteq \mathbb{R}^n$, where $D$ is a convex set.

$$\begin{array}{ll} \text{minimize} & w_0(x) \\ \text{subject to} & w_i(x) \leq 0, \quad i = 1, \ldots, m \\ & v_i(x) = 0, \quad i = 1, \ldots, p \\ & x \in D \end{array}$$

where $w_0, w_i$'s are convex functions and $v_i$'s are affine functions of $x$. We assume that $w_0$ is strictly convex, and $D$ is polyhedral and bounded. The Lagrangian, for $\lambda \succeq 0$, is given by

$$L(x, \lambda, \nu) = w_o(x) + \sum_{i=1}^{m} \lambda_i w_i(x) + \sum_{i=1}^{p} \nu_i v_i(x)$$

The dual function is $g(\lambda, \nu) = \inf_{x \in D} L(x, \lambda, \nu)$, and the dual problem is

$$\begin{array}{ll} \text{maximize} & g(\lambda, \nu) \\ \text{subject to} & \lambda \succeq 0 \end{array} \tag{2}$$

Let us define, for $\lambda \succeq 0$,

$$x^*(\lambda, \nu) = \arg \inf_{x \in D} L(x, \lambda, \nu)$$

This is unique under the assumptions on $w_0$ and $D$. We describe here the projected subgradient algorithm [13] for the dual problem. We start with an initial point $(\lambda^{(0)}, \nu^{(0)})$. At each iteration $k$ we compute

$$\begin{array}{ll} \lambda_i^{(k+1)} = \left(\lambda_i^{(k)} - \alpha_k h_i^{(k)}\right)_+ & \forall i = 1, \ldots, m \\ \nu_i^{(k+1)} = \nu_i^{(k)} - \alpha_k f_i^{(k)} & \forall i = 1, \ldots, p \end{array} \tag{3}$$

where $(x)_+ = x$ if $x \geq 0$, and $(x)_+ = 0$, otherwise. Also, $\alpha_k$ is a positive scalar step-size. The vector $[h^{(k)^T}, f^{(k)^T}]^T$ is a subgradient of $-g$ at $(\lambda^{(k)}, \nu^{(k)})$. A subgradient is given by

$$\begin{aligned} h_i^{(k)} &= -w_i(x^*(\lambda^{(k)}, \nu^{(k)})), \forall i = 1, \ldots, m \\ f_i^{(k)} &= -v_i(x^*(\lambda^{(k)}, \nu^{(k)})), \forall i = 1, \ldots, p \end{aligned}$$

One simple condition for convergence is (e.g., [13])

$$\alpha_k \to 0, \quad \sum_{k=1}^{\infty} \alpha_k = \infty$$

Under the assumptions that $w_0$ is strictly convex and $D$ is a bounded polyhedron, the dual function is differentiable (see, for example, Appendix C in [14]). Then if strong duality holds, the sequence of primal iterates $x^{(k)} = x^*(\lambda^{(k)}, \nu^{(k)})$ converges to the optimal solution of the primal problem (see, for example, Chapter 3 in [14]).

## IV. Partially Distributed Algorithm

### A. Linear Programming Formulation

The problem in (1) can be rewritten as

$$\begin{aligned} \text{maximize} \quad & T \\ \text{subject to} \quad & \sum_{j\in N_i}(r_{ij}-r_{ji}) = S_i, \quad \forall i \in V \\ & 0 \le r_{ij} \le R_{ij}, \quad \forall i \in V, \forall j \in N_i \\ & T\sum_{j\in N_i} E_{ij} r_{ij} \le B_i, \quad \forall i \in V \end{aligned}$$

where the last set of constraints model energy conservation at each node. Changing variable to $q = 1/T$, we obtain an equivalent linear programming formulation.

$$\begin{aligned} \text{minimize} \quad & q \\ \text{subject to} \quad & \sum_{j\in N_i}(r_{ij}-r_{ji}) \;=\; S_i, \quad \forall i \in V \\ & 0 \le r_{ij} \le R_{ij}, \quad \forall i \in V, \forall j \in N_i \\ & \sum_{j\in N_i} E_{ij} r_{ij} \;\le\; qB_i, \quad \forall i \in V \end{aligned} \tag{4}$$

This is the optimization problem that we will solve in a distributed manner. We can interpret the above problem as minimizing the maximum ratio of power consumption to energy supply at a node.

### B. Dual Problem

We construct the dual problem by introducing Lagrange multipliers $\lambda_i$ for the energy constraint and $\nu_i$ for the flow conservation constraint at each node $i$. This results in the *Lagrangian*

$$\begin{aligned} L(q,r,\lambda,\nu) &= q + \sum_{i\in V}\lambda_i\Big\{\sum_{j\in N_i} E_{ij}r_{ij} - qB_i\Big\} \\ &\quad + \sum_{i\in V}\nu_i\Big\{\sum_{j\in N_i}\Big(r_{ij}-r_{ji}\Big) - S_i\Big\} \\ &= -\sum_{i\in V}\nu_i S_i + q\Big\{1 - \sum_{i\in V}\lambda_i B_i\Big\} \\ &\quad + \sum_{i\in V}\sum_{j\in N_i} r_{ij}\Big(\lambda_i E_{ij} + \nu_i - \nu_j\Big) \end{aligned} \tag{5}$$

for $\lambda \succeq 0, 0 \le r_{ij} \le R_{ij}, q \ge 0$. The *dual function* is given by

$$g(\lambda,\nu) = \inf_{0\le r_{ij}\le R_{ij}, q\ge 0} L(q,r,\lambda,\nu)$$

for $\lambda \succeq 0$. It is clear from the expression of the Lagrangian that the dual function can be evaluated separately in each of the variables $r_{ij}$ and $q$. We assume that there always exists a feasible flow satisfying the flow conservation equations and the rate constraints with strict inequality in Problem (4). Also, we can always satisfy the energy conservation inequality in a strict manner by choosing a large enough value of $q$. Thus the Slater's condition for constraint qualification is satisfied (see, for example, [15]) and hence strong duality holds for this problem. Thus we can solve the primal problem (4) via the dual given in (2).

### C. Subgradient Algorithm

We use the subgradient algorithm to solve the dual problem. However, since the objective function is not strictly convex in the primal variables, the dual function is non-differentiable. Hence, the primal solution is not immediately available; it can be recovered using approaches in [13], [16]. We use a simple approach similar to that used in [11]. We change the primal objective function to $q^2$, since minimizing $q$ is the same as minimizing $q^2$, and also add a small convex quadratic *regularization term* to the primal objective function. We minimize $q^2 + \epsilon\sum_{i\in V}\sum_{j\in N_i} r_{ij}^2$. Let us denote by $\hat{q}, \hat{r}$ the solution of the optimization problem with the regularized objective, and by $\tilde{q}, \tilde{r}$ the solution of the original linear program (4). Then we have

$$\begin{aligned} & \tilde{q}^2 + \epsilon\sum_{i\in V}\sum_{j\in N_i} R_{ij}^2 \ge \hat{q}^2 + \epsilon\sum_{i\in V}\sum_{j\in N_i}\hat{r}_{ij}^2 \\ \Rightarrow \quad & \hat{T}^2 \ge \frac{\tilde{T}^2}{1+\epsilon\tilde{T}^2\sum_{i\in V}\sum_{j\in N_i} R_{ij}^2} \end{aligned}$$

Thus by choosing $\epsilon$ small enough, we can make the solution of the regularized problem $\tilde{T}$ arbitrarily close to the optimal lifetime $\hat{T}$ given by problem (4). For a detailed discussion on augmented Lagrangian (with regularization) methods, refer [14].

Also, we consider the domain of $q$ to be $[0\ Q]$, where $Q$ is a loose upper bound on the value of $q$. Then it follows that the dual function is differentiable. The Lagrangian for the regularized objective function is given by

$$\begin{aligned} L(q,r,\lambda,\nu) = -\sum_{i\in V}\nu_i S_i \quad & + \Big(q^2 - q\sum_{i\in V}\lambda_i B_i\Big) \\ & + \sum_{i\in V}\sum_{j\in N_i}\Big(\epsilon r_{ij}^2 + r_{ij}\Big(\lambda_i E_{ij} + \nu_i - \nu_j\Big)\Big) \end{aligned}$$

Note that the Lagrangian is still separable in the variables $r_{ij}$ and $q$; the Lagrange dual function can still be evaluated separately in each of these variables.

For given $\lambda, \nu$, the dual function for the regularized problem is given by

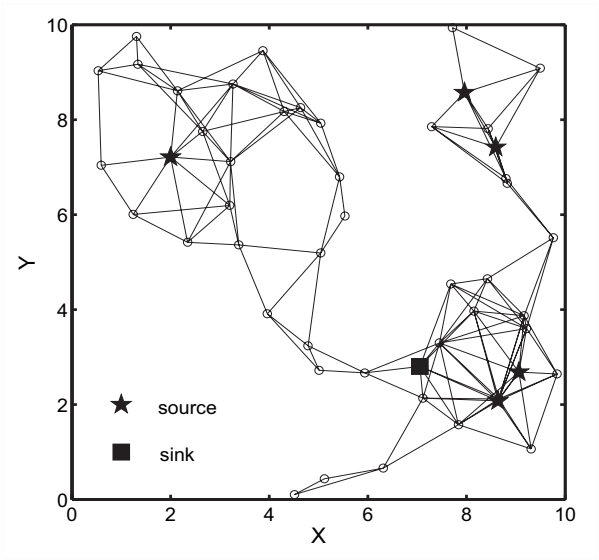
$$\begin{aligned} g(\lambda,\nu) = -\sum_{i\in V}\nu_i S_i \quad & + \min_{0\le q\le Q}\Big(q^2 - q\sum_{i\in V}\lambda_i B_i\Big) \\ & + \sum_{i\in V}\sum_{j\in N_i}\min_{0\le r_{ij}\le R_{ij}}\Big(\epsilon r_{ij}^2 + r_{ij}\Big(\lambda_i E_{ij} + \nu_i - \nu_j\Big)\Big) \end{aligned}$$

The sequence of primal iterates is given by

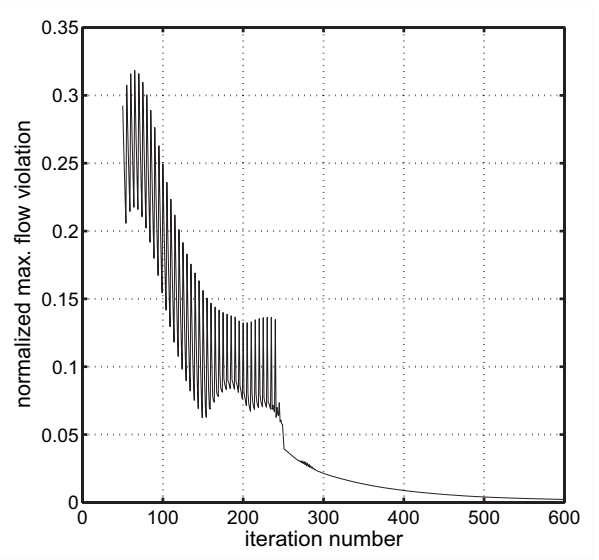
$$\begin{aligned} q^{(k)} &= \arg\min_{0\le q\le Q}\Big(q^2 - q\sum_{i\in V}\lambda_i^{(k)} B_i\Big) \\ r_{ij}^{(k)} &= \arg\min_{0\le r_{ij}\le R_{ij}}\Big(\epsilon r_{ij}^2 + r_{ij}\big(\lambda_i^{(k)} E_{ij} + \nu_i^{(k)} - \nu_j^{(k)}\big)\Big) \end{aligned} \tag{6}$$

The subgradient of $-g$ at $(\lambda^{(k)}, \nu^{(k)})$ is given by

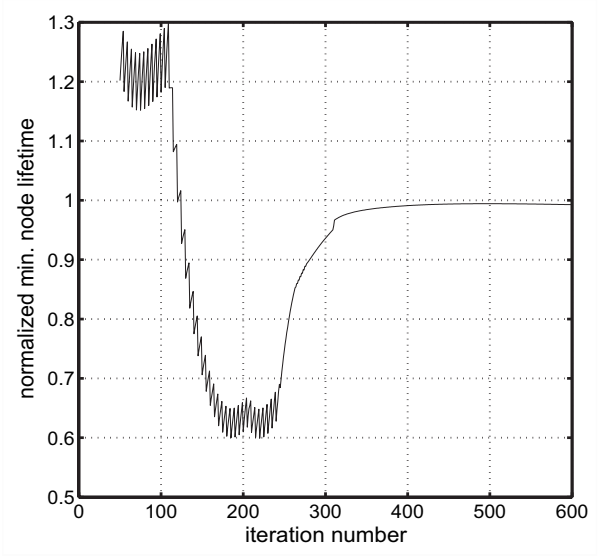
$$\begin{aligned} h_i^{(k)} &= q^{(k)} B_i - \sum_{j\in N_i} E_{ij} r_{ij}^{(k)} \\ f_i^{(k)} &= S_i - \sum_{j\in N_i}\Big(r_{ij}^{(k)} - r_{ji}^{(k)}\Big) \end{aligned} \tag{7}$$



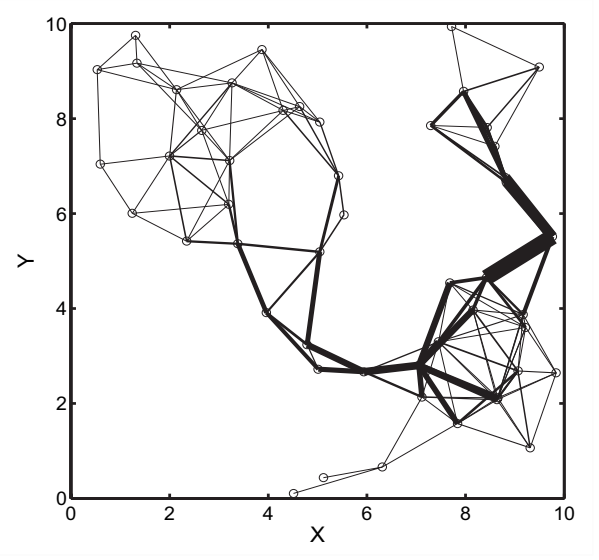

(a) connectivity graph - 50 nodes




(b) violation of flow conservation equations




(c) minimum node lifetime




(d) computed routing flow

Fig. 1. Partially distributed computation of optimal flows over a 50 node network

We can then use the subgradient algorithm in equation (3) to solve the dual problem. Since the dual function is differentiable, the sequence of primal iterates converges to a solution of the regularized primal problem.

All the computations in the equations (6) and (7), involve either computing a linear function, or the optimal value of a quadratic function in one variable. Both these operations need very low computational power. Note that each node $i$ can compute $r_{ij}^{(k)}, j \in N_i$, using the values of the Lagrange multipliers $\lambda_i^{(k)}, \nu_i^{(k)}, \nu_j^{(k)}$, which are either local or computed by its neighboring nodes (see equation 6). However the computation of $q^{(k)}$ needs the values of $\lambda_i^{(k)}$ for all $i \in V$. Hence this is a centralized computation that needs to be carried out by one node in the network. Similarly to compute the subgradients $h_i, f_i$, each node $i$ needs the values of the rates computed by itself and its neighbors at the previous iteration. Also, it needs the value of the variable $q$ at the previous iteration. Hence after each iteration $q^{(k)}$ needs to be broadcast to all nodes. We will address this issue in the next section.

### D. Simulation Results

The subgradient algorithm described in the previous section was simulated over a network of 50 nodes randomly distributed over a square of side 10 units. Each node was assumed to have a maximum communication radius of 2 units. The energy consumption for sending a packet from node $i$ to node $j$ was taken to be $E_{ij} = c_1 + c_2 d^4$, with $c_2 = 0.1c_1$. This corresponds to a deterministic path loss model over an AWGN channel, where the received power decays as $1/d^4$, over a distance $d$. The constant $c_1$ corresponds to the fixed energy cost of transmitting a unit of information between any two nodes. Five nodes were selected as source nodes; these nodes generated information packets which were routed to

the sink node. All the nodes were assumed to have equal initial energy. For the computation, the node energies and data rates were normalized with respect to their maximum values; this gives better convergence. The step-size at iteration $k$ was $\alpha_k = \max\left(0.01, \frac{0.5}{\sqrt{k}}\right)$. Also, an exponentially decreasing sequence of $\epsilon$ was used for a certain number of iterations; after that a constant value of $\epsilon$ was used for the rest of the iterations. It is worth pointing out that convergence is faster for larger values of $\epsilon$. Thus there is a tradeoff between the rate of convergence and the suboptimality of the solution.

Fig. 1(a) shows the connectivity graph of the network. Fig. 1(b) shows the maximum violation of the flow conservation equations at each iteration, normalized w.r.t. the total flow in the network. After about 600 iterations, the flow conservation equations are satisfied with a high degree of accuracy. The minimum node lifetime corresponding to the routing scheme computed at each iteration, normalized w.r.t. the optimal value, is shown in Fig. 1(c). After about 600 iterations, the minimum node lifetime is close to the optimal value of the network lifetime given by a centralized solution to the linear program. In Fig. 1(d), the computed flows are represented as a graph, where the thickness of an edge is proportional to the amount of flow on the corresponding wireless link. We see that the optimal load-balanced solution makes use of multipath routing.

This example shows that the number of iterations required for convergence is substantial. Hence energy should be invested in computing an optimal flow only if the total amount of data communicated over the network during its lifetime is much larger than the overhead incurred in computing the routing scheme. This would be the case in static sensor networks, for example, a network to monitor the environment.

## V. Fully Distributed Algorithm

The algorithm proposed in the previous section requires the values of the Lagrange multipliers $\lambda_i$ to be communicated to one central node at each iteration; similarly the variable $q$ needs to be broadcast to every node. In this section, we describe a way to completely decentralize the problem by introducing additional variables corresponding to an upper bound on the inverse lifetime of each node. The problem of maximizing the network lifetime can be reformulated as the following convex quadratic optimization problem.

$$\begin{aligned} \text{minimize} \quad & \sum_{i\in V} q_i^2 \\ \text{subject to} \quad & \sum_{j\in N_i}(r_{ij}-r_{ji}) = S_i, \quad \forall i \in V \\ & 0 \le r_{ij} \le R_{ij}, \quad \forall i \in V, \forall j \in N_i \\ & \sum_{j\in N_i} E_{ij} r_{ij} \le q_i B_i, \quad \forall i \in V \\ & q_i = q_j, \quad \forall i \in V, \forall j \in N_i \end{aligned} \tag{8}$$

Here again, the constraints are flow conservation and energy conservation constraints. In addition, we have local variables $q_i$'s such that for a flow $r$, $\frac{1}{q_i} \le T_i(r)$. There are additional constraints that enforces all $q_i$'s to be equal. Again we consider a quadratic objective function that is strictly convex in the $q_i$'s. Also, to ensure that the dual function is differentiable, we restrict the domain to $0 \le q_i \le Q$, for some loose upper bound $Q$ (as in Section IV-C).

### A. Dual Problem

As in Section IV-C, we add a regularization term $\epsilon \sum_{i\in V}\sum_{j\in N_i} r_{ij}^2$ to the primal objective function. We form the partial Lagrangian by introducing Lagrange multipliers $\nu_i, \gamma_{ij}$ for the flow conservation and lifetime equality constraints, respectively. We do not need to relax the energy conservation constraint because this constraint can be satisfied locally at each node $i$. For $\lambda \succeq 0, 0 \le q_i \le Q$,we have

$$\begin{aligned} & L(q,r,\nu,\gamma) = \\ & \sum_{i\in V} q_i^2 + \epsilon \sum_{i\in V}\sum_{j\in N_i} r_{ij}^2 + \sum_{i\in V}\nu_i\left\{\sum_{j\in N_i}\Big(r_{ij}-r_{ji}\Big) - S_i\right\} \\ & \qquad + \sum_{i\in V}\sum_{j\in N_i}\gamma_{ij}(q_i - q_j) \\ & = -\sum_{i\in V}\nu_i S_i + \sum_{i\in V}\Bigg\{q_i^2 + \epsilon\sum_{j\in N_i} r_{ij}^2 + q_i\sum_{j\in N_i}(\gamma_{ij}-\gamma_{ji}) \\ & \qquad + \sum_{j\in N_i} r_{ij}(\nu_i - \nu_j)\Bigg\} \end{aligned}$$

For $\lambda \succeq 0$, the dual function is given by

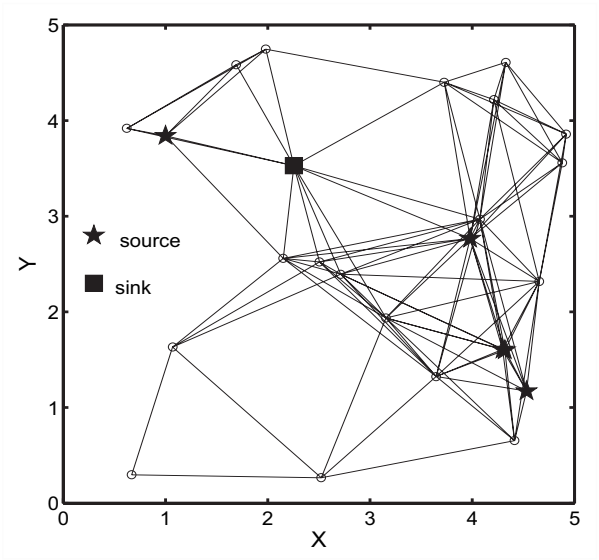
$$\begin{aligned} & g(\nu,\gamma) \\ & = \inf_{r, 0\le q_i\le Q}\left\{ L(q,r,\nu,\gamma) \;\middle|\; \begin{array}{l} 0 \le r_{ij} \le R_{ij}, \forall i \in V, \forall j \in N_i \\ \sum_{j\in N_i} E_{ij} r_{ij} \le q_i B_i, \ \forall i \in V \end{array}\right\} \\ & = -\sum_{i\in V}\nu_i S_i \\ & \quad + \sum_{i\in V}\inf_{0\le q_i\le Q, r_{ij}, j\in N_i}\Bigg\{q_i^2 + \epsilon\sum_{j\in N_i} r_{ij}^2 + q_i\sum_{j\in N_i}(\gamma_{ij}-\gamma_{ji}) \\ & \qquad + \sum_{j\in N_i} r_{ij}(\nu_i-\nu_j)\;\Bigg|\; 0 \le r_{ij} \le R_{ij}, \sum_{j\in N_i} E_{ij} r_{ij} \le q_i B_i\Bigg\} \end{aligned}$$

Thus the dual function can be evaluated separately in the variables corresponding to each node $i \in V$. The variables local to node $i$ are $q_i$ and $r_{ij}, j \in N_i$. The dual problem is then given by (2). Again, by arguments similar to those in Section IV, the Slater's condition for constraint qualification is satisfied, and hence strong duality holds.

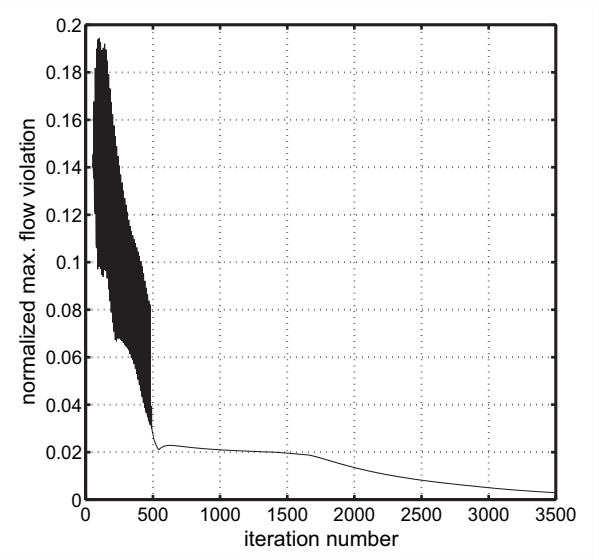
### B. Subgradient Algorithm

Again, we will use the subgradient algorithm (3) to solve the dual problem. During the $k$'th iteration of the subgradient algorithm, each node solves the following convex quadratic program with variables $q_i, r_{ij}$, for $i \in V, j \in N_i$.

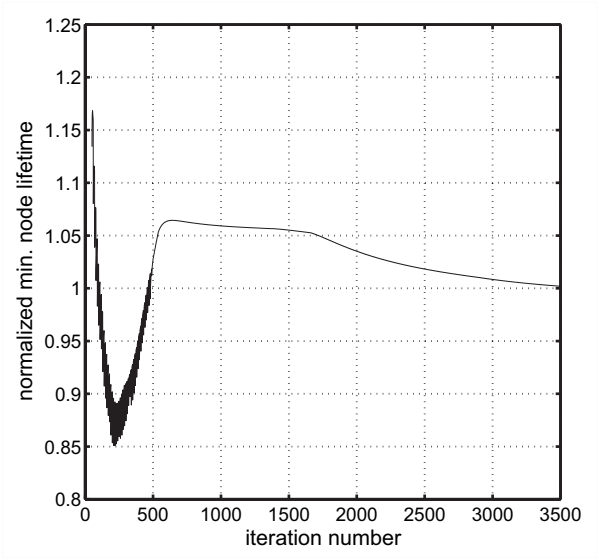
$$\begin{aligned} \text{minimize} \quad & q_i^2 + \epsilon\sum_{j\in N_i} r_{ij}^2 + q_i\sum_{j\in N_i}(\gamma_{ij}^{(k)} - \gamma_{ji}^{(k)}) \\ & + \sum_{j\in N_i} r_{ij}(\nu_i^{(k)} - \nu_j^{(k)}) \\ \text{subject to} \quad & \sum_{j\in N_i} E_{ij} r_{ij} \le q_i B_i \\ & 0 \le r_{ij} \le R_{ij}, \quad \forall i \in V, \forall j \in N_i \end{aligned}$$



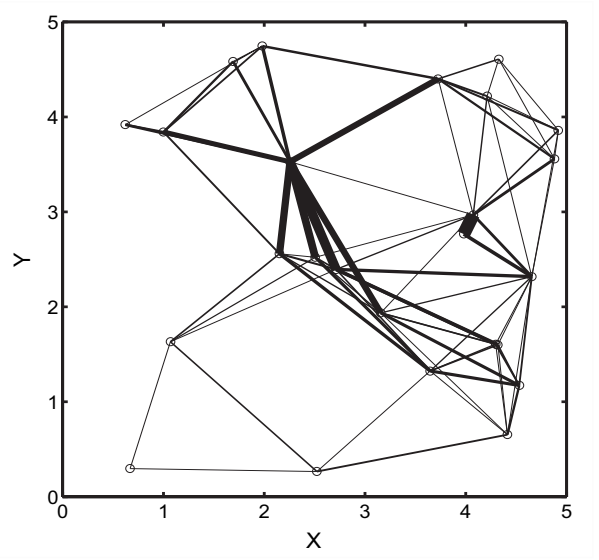

(a) connectivity graph - 25 nodes




(b) violation of flow conservation equations




(c) minimum node lifetime




(d) computed routing flow

Fig. 2. Fully distributed computation of optimal flows over a 25 node network

Let us denote the optimal values of the variables for the above problem as $q_i^{(k)}, r_{ij}^{(k)}$. $f_i^{(k)}$ and $h_{ij}^{(k)}$ are components of the subgradient of $-g$ evaluated at $(\nu^{(k)}, \gamma^{(k)})$. They are given by

$$f_i^{(k)} = S_i - \sum_{j \in N_i} (r_{ij}^{(k)} - r_{ji}^{(k)})$$

$$h_{ij}^{(k)} = q_j^{(k)} - q_i^{(k)}$$

### C. Simulation Results

The fully distributed algorithm was simulated over a network of 25 nodes. The network parameters were chosen to be the same as those in Section IV-D. The results are shown in Fig. 2. Again, the flow conservation violation is normalized with respect to the total flow in the network and the minimum node lifetime is normalized with respect to the optimal value of the network lifetime given by a centralized solution to the problem. As we can see from the graphs in Fig. 2, the convergence of this algorithm is much slower than that for the partially distributed algorithm. Thus we have a trade-off between the amount of communication during each iteration, and the number of iterations.

## VI. Extensions

In this section, we show how we can exploit the problem structure to obtain decentralized subgradient algorithms for many different extensions to the problem considered in the previous sections. We only show how to extend the methodology for the partially distributed algorithm to these problem

formulations. In each case, the analysis for the fully distributed algorithm follows in a similar way. Note that some of these extensions were also considered in [5] and [8].

### A. Multicommodity Flows

Here, we consider the case where the sources in the network generate data of different *commodities*. We assume that in general, this data needs to be communicated to a different sink node for each commodity. Let $S_i^c$ be the rate at which data of commodity $c \in \{1,\ldots,C\}$ is generated at node $i$. Also, if node $k$ is the sink for commodity $c$, we take $S_k^c = -\sum_{i\in V, i\neq k} S_i^c$. Then the problem of maximizing the network lifetime can be written as follows.

$$\begin{aligned}
\text{minimize} \quad & q \\
\text{subject to} \quad & \sum_{j\in N_i} (r_{ij}^c - r_{ji}^c) = S_i^c, \quad \forall i \in V, c = 1,\ldots,C \\
& r_{ij}^c \geq 0, \forall i \in V, \quad \forall j \in N_i, c = 1,\ldots,C \\
& \sum_{c=1}^{C} r_{ij}^c \leq R_{ij}, \quad \forall i \in V, \forall j \in N_i \\
& \sum_{j\in N_i} E_{ij} \sum_{c=1}^{C} r_{ij}^c \leq qB_i, \quad \forall i \in V
\end{aligned}$$

where the variables are $q, r_{ij}^c$. The interpretation of the variables and constraints is similar to that in Problem (4). We relax the flow conservation and energy conservation constraints by introducing Lagrange multipliers multipliers $\nu_i^c$ (for each commodity $c$) and $\lambda_i$, respectively, for each node $i$. The flow positivity and maximum rate constraint for each link can be satisfied locally. This results in the partial Lagrangian

$$\begin{aligned}
L(q,r,\lambda,\nu) &= q + \sum_{i\in V} \lambda_i \left\{ \sum_{j\in N_i} E_{ij} \sum_{c=1}^{C} r_{ij}^c - qB_i \right\} \\
&\quad + \sum_{i\in V} \sum_{c=1}^{C} \nu_i^c \left\{ \sum_{j\in N_i} \left( r_{ij}^c - r_{ji}^c \right) - S_i^c \right\} \\
&= -\sum_{i\in V} \sum_{c=1}^{C} \nu_i^c S_i^c + q\left\{ 1 - \sum_{i\in V} \lambda_i B_i \right\} \\
&\quad + \sum_{i\in V} \sum_{j\in N_i} \sum_{c=1}^{C} r_{ij}^c \left( \lambda_i E_{ij} + \nu_i^c - \nu_j^c \right)
\end{aligned}$$

Note that the Lagrangian is still separable as in equation (5). The only difference is that we now have a separate dual variable for the flow conservation constraint for each commodity at a given node. Hence, methods similar to those in Section IV can be used to obtain a partially distributed algorithm for solving this problem.

### B. Multiple Sinks - Anycast Flows

Consider a sensor network with more than one sink. The objective is for the data generated at each node to arrive at least one sink out of a given set of sinks. For such networks we can modify Problem (4) by dropping the flow conservation equations at the sink nodes from the problem formulation. It is easy to see that the solution of the resulting problem formulation gives the optimal routing flow for the multiple sink case. Alternatively, we can consider an equivalent network, with an additional *super-sink* node (similar to the idea of a super-origin node in [8]) that is connected to the sink nodes by links with zero energy consumption. We can then solve the problem for this new network, where the super-sink node is considered to be the only sink node in the network. This problem is then identical to the problem with one sink node considered throughout this paper. Also, note that we can use the same idea to extend the analysis for the multicommodity flow problem formulation described above, to the case where for each commodity $c$, we have multiple sinks given by the set $M^c$. In this case, the data of commodity $c$ can be routed to any node in the set $M^c$.

### C. Capacitated Nodes and Links

Each node can be limited in the maximum average power that its battery can supply. This can be modeled by the following local constraint.

$$\sum_{j\in N_i} E_{ij} r_{ij} \leq P_i^{\max}$$

We can introduce the above constraint in problem (4) without changing the separable structure of the dual problem. Alternatively, this can be incorporated into the domain of the problem, again maintaining the separable structure. This would mean solving equation (6) with the additional constraint above.

### D. Energy for Reception

For a packet transmission from node $i$ to node $j$, let the energy consumption per packet at the receiver $j$ be $E_{ij}^r$. This can be used to model the effect of the overhead at a receiver to receive and decode a packet. For many radio models, the value of $E_{ij}^r$ does not depend on the transmitting node $i$. Also, again let $E_{ij}$ be the transmission energy consumption per packet at the transmitter $i$. Then the problem of maximizing the network lifetime can be written as follows.

$$\begin{aligned}
\text{minimize} \quad & q \qquad (9) \\
\text{subject to} \quad & \sum_{j\in N_i} (r_{ij} - r_{ji}) = S_i, \quad \forall i \in V \\
& 0 \leq r_{ij} \leq R_{ij}, \forall i \in V, \forall j \in N_i \\
& \sum_{j\in N_i} \left( E_{ij} r_{ij} + E_{ji}^r r_{ji} \right) \leq qB_i, \quad \forall i \in V \qquad (10)
\end{aligned}$$

The interpretation of the constraints and the variables is similar to Problem (4) . However, the constraints corresponding to energy conservation now model the energy consumption in

the reception of a packet as well. The Lagrangian is given by

$$\begin{aligned} L(q,r,\lambda,\nu) &= q + \sum_{i\in V}\lambda_i\left\{\sum_{j\in N_i}(E_{ij}r_{ij} + E^r_{ji}r_{ji} - qB_i\right\} \\ &\quad + \sum_{i\in V}\nu_i\left\{\sum_{j\in N_i}\Big(r_{ij} - r_{ji}\Big) - S_i\right\} \\ &= -\sum_{i\in V}\nu_i S_i + q\left\{1 - \sum_{i\in V}\lambda_i B_i\right\} \\ &\quad + \sum_{i\in V}\sum_{j\in N_i} r_{ij}\Big(\lambda_i E_{ij} + \lambda_j E^r_{ji} + \nu_i - \nu_j\Big) \end{aligned} \tag{11}$$

Thus the Lagrangian is separable, and the dual function can be evaluated in a partially distributed manner as in Section IV.

### E. Packet Loss

We first consider the case where for every packet from node $i$ sent to neighboring node $j$, an acknowledgment is sent back by node $j$ to node $i$. It is easy to modify the analysis below to the case, where an acknowledgment is sent for more than one packet together. We assume that the transmission powers (and hence the values of $E_{ij}$) are such that all links have the same SNR at the receiver. This assumption is for simplicity; the techniques in this paper can also be generalized to the case where the SNR can be different for each link. For the given fixed SNR, let $p_e$ be the probability that a packet is received in error. Then the average number of transmissions $\bar{n}$ required over each hop to transmit a packet reliably is given by

$$\bar{n} = (1-p_e) + 2p_e(1-p_e) + 3p_e^2(1-p_e) + \ldots = \frac{p_e}{1-p_e}$$

If each packet is $b_1$ bits long, and the acknowledgment is $b_2$ bits long, then the average energy consumption per packet on a link from node $i$ to node $j$ at the transmitter and the overhead at the receiver, respectively, are given by

$$\bar{E}_{ij} = b_1 E_{ij}\frac{p_e}{1-p_e} \quad \text{and} \quad E^r_{ij} = b_2 E_{ji}$$

where $E_{ij}$ and $E_{ji}$ are the energy consumption values in transmitting a bit on link $(i,j)$ in the two directions. To simplify the analysis, we assumed that an acknowledgment is always received correctly. If the network is alive for a long time, we can assume that the lifetime of each node depends on the average rate of energy consumption. Hence we can replace the energy consumption values in Problem (10) by their corresponding means calculated above; the solution of the new problem will give the routing flow that maximizes the network lifetime. Note that the above equations do not consider the overhead in receiving packets; we can easily modify them to include the overhead too (as discussed in VI-D). For the case, where we have end-to-end acknowledgment, the analysis is more complex. This is because the probability of packet loss, and hence the average number of retransmissions, depends on the total number of hops on a path.

## VII. Generalized Network Lifetime

Until now we considered the network lifetime to be time at which the first node runs out of energy. Thus we assumed that all nodes are of equal importance and critical to the operation of the sensor network. However, for a heterogeneous wireless sensor network, some nodes maybe more important than others. Also, if there are two nodes collecting highly correlated data, the network can remain functional even if one node runs out of energy. Moreover, for the case of nodes with highly correlated data, we may want only one node to forward the data at a given time. Thus we can activate the two nodes in succession, and still be able to send the necessary data to the sink. In this section, we will model the lifetime of a network to be a function of the times for which the nodes in the network can forward their data to the sink node. In order to state this precisely, we redefine the node lifetime and the network lifetime for the analysis in this section. We will relax the constraint on the maximum flow over a link at a given time. We assume that the data rates are low enough such that we can always find a TDMA scheme to support the given data rates. This is similar to the model assumed in [5].

Each node $k \in V$ generates data at rate $S_k$. Let $r^k_{ij}$ be the rate at which node $i$ sends data, originally generated by node $k$, from $i$ to $j$. Then we have the flow conservation constraints

$$\sum_{j\in N_i}\left(r^k_{ij} - r^k_{ji}\right) = \delta_{ik}S_k, \qquad \forall i,k \in V \tag{12}$$

where $\delta_{ik} = 1$ if $i = k$, and 0 otherwise. Also, as before $E_{ij}$ is the amount of energy used by node $i$ to transmit a bit to node $j$, and $B_i$ is the total energy available at node $i$. We define $T_k$, the lifetime of node $k$, to be the total time for which node $k$ generates data that is transmitted over the network to the sink node. Then the total number of bits sent from $i$ to $j$ is $\sum_{k\in V} r^k_{ij}T_k$. Thus we have the energy constraint

$$\sum_{j\in N_i} E_{ij}\sum_{k\in V} r^k_{ij}T_k \le B_i$$

We consider a generic definition of network lifetime given by a concave function of the node lifetimes. In particular, we define

$$T_{\text{net}} = f(T_1,\ldots,T_{|V|})$$

where $f : \mathbb{R}^{|V|} \to \mathbb{R}$ is a concave function in the vector of node lifetimes. In the previous sections, we considered the special case of $T_{\text{net}} = \min(T_1,\ldots,T_{|V|})$.

### A. Maximum Lifetime Routing

We can write the problem of maximizing the network lifetime as follows.

$$\begin{aligned} \text{maximize} \quad & f(T_1,\ldots,T_{|V|}) \\ \text{subject to} \quad & \sum_{j\in N_i}\left(r^k_{ij} - r^k_{ji}\right) = \delta_{ik}S_k, && \forall i,k \in V \\ & r^k_{ij} \ge 0, && \forall i,k \in V, \forall j \in N_i \\ & \sum_{j\in N_i} E_{ij}\sum_{k\in V} r^k_{ij}T_k \le B_i, && \forall i \in V \end{aligned}$$

We apply a change of variables $y^k_{ij} = r^k_{ij}T_k$. We can interpret $y^k_{ij}$ as the total number of bits generated by node $k$, transmitted over link $i$ to $j$. We can rewrite the above problem as a convex

optimization problem.

$$\begin{aligned} \text{maximize} \quad & f(T_1,\ldots,T_{|V|}) \\ \text{subject to} \quad & \sum_{j\in N_i}\left(y_{ij}^k - y_{ji}^k\right) = T_k\delta_{ik}S_k, \quad \forall i,k\in V \\ & y_{ij}^k \geq 0, \qquad \forall i,k\in V, \forall j\in N_i \\ & \sum_{k\in V}\sum_{j\in N_i} y_{ij}^k E_{ij} \leq B_i, \qquad \forall i\in V \end{aligned} \tag{13}$$

Then, we can interpret the last set of constraints as *bit conservation* equations. Since the data from all the nodes is routed to a single sink, solving the above problem is equivalent to solving

$$\begin{aligned} \text{maximize} \quad & f(T_1,\ldots,T_{|V|}) \\ \text{subject to} \quad & \sum_{j\in N_i}\left(z_{ij} - z_{ji}\right) = T_i S_i, \quad \forall i\in V \\ & z_{ij} \geq 0, \qquad \forall i\in V, \forall j\in N_i \\ & \sum_{j\in N_i} z_{ij}E_{ij} \leq B_i, \qquad \forall i\in V \end{aligned} \tag{14}$$

where $z_{ij} = \sum_{k\in V} y_{ij}^k$. From the solution $z_{ij}$ of this problem, we can easily obtain an optimal flow $y_{ij}^k, k\in V$, which is a solution to problem (13). If the function $f$ is partially or fully separable in the variables $T_i$, we can use the methodology in Sections IV and V to obtain distributed algorithms.

## VIII. Conclusions

In this paper, we proposed two subgradient algorithms to compute an optimal routing flow to maximize the network lifetime. The algorithms were derived to solve the dual problems of programs (4) and (8) in a partially and a fully decentralized manner, respectively. The computation results show that the rate of convergence of the fully distributed algorithm was slower than that for the partially distributed algorithm. However, each iteration of the partially distributed algorithm involves communication between all the nodes and a central node (e.g. sink node). Hence, it is not obvious which algorithm will have a lower total energy consumption cost. If the radius of the network graph is small, then it would be more energy efficient to use the partially distributed algorithm even though each iteration involves the update of a central variable. Conversely, for large network radius, the fully distributed algorithm would be a better choice. Also, we note that the computation at each node for the fully distributed algorithm involves the solution of a convex quadratic optimization problem. This is in contrast to the partially distributed algorithm, where each iteration consists of minimization of a quadratic function of a single variable, which can be done analytically.

We considered many different extensions to the original problem and showed how the subgradient approach can be used to obtain distributed algorithms. In addition, we considered a generalization of the definition of network lifetime to model realistic sensor network scenarios, and reformulated the problem as a convex optimization problem with separable structure.

### A. Discussion

We considered only synchronous subgradient algorithms for computing the maximum lifetime routing flow in a distributed manner. However, in real networks the nodes may not be synchronized and there might be packet losses. Asynchronous versions of these algorithms can be analyzed using techniques similar to that in [10]. Analysis of conditions under which the asynchronous algorithms will converge is beyond the scope of this paper. Also, asynchronous algorithms can be easily implemented as protocols on real networks. The actual design of the protocols is also beyond the scope of this paper.

## References


[1] V. Rodoplu and T. H. Meng, "Minimum energy mobile wireless networks," *IEEE J. Select. Areas Communi.*, vol. 17, no. 8, pp. 1333–1344, 1999.

[2] L. Li and J. Y. Halpern, "Minimum energy mobile wireless networks revisited," *IEEE International Conference on Communications (ICC)*, 2001.

[3] R. Wattenhofer et al., "Distributed topology control for power efficient operation in multihop wireless ad hoc networks," *IEEE INFOCOM*, 2001.

[4] L. Li et al., "Analysis of a cone-based distributed topology control algorithm for wireless multi-hop networks," *ACM Symposium on Principle of Distributed Computing (PODC)*, 2001.

[5] J.-H. Chang and L. Tassiulas, "Energy conserving routing in wireless ad-hoc networks," in *Proc. IEEE INFOCOM*, pp. 22–31, 2000.

[6] ——, "Routing for maximum system lifetime in wireless ad-hoc networks," in *Proc. of 37-th Annual Allerton Conference on Communication, Control and Computing*, 1999.

[7] J. Pan et al., "Topology control for wireless sensor networks," *MobiCom*, 2003.

[8] G. Zussman and A. Segall, "Energy efficient routing in ad hoc disaster recovery networks," *INFOCOM*, 2003.

[9] F. P. Kelly, A. Maulloo, and D. Tan, "Rate control for communication networks: shadow prices, proportional fairness and stability," *Journal of the Operational Research Society*, vol. 49, pp. 237–252, 1998.

[10] S. H. Low and D. E. Lapsley, "Optimization flow control - I: Basic algorithm and convergence," *IEEE/ACM Trans. Networking*, vol. 7, pp. 861–874, 1999.

[11] L. Xiao, M. Johansson, and S. Boyd, "Simultaneous routing and resource allocation via dual decomposition," *IEEE Trans. Commun.*, vol. 52, no. 7, pp. 1136–1144, 2004.

[12] T. S. Rappaport, *Wireless communications: principles and practice*. Prentice Hall, 1996.

[13] N. Z. Shor, *Minimization Methods for Non-differentiable Functions*. Springer-Verlag, 1985.

[14] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*. Athena Scientific, 1997.

[15] S. Boyd and L. Vandenberge, *Convex Optimization*. Cambridge University Press, 2004.

[16] H. D. Sherali and G. Choi, "Recovery of primal solutions when using subgradient optimization methods to solve Lagrangian duals of linear programs," *Operations Research Letters*, vol. 19, pp. 105–113, 1996.


**Ritesh Madan** received the B.Tech. degree from the Indian Institute of Technology (IIT), Mumbai, India, in 2001, and the M.S. degree from Stanford University, CA,in 2003, both in Electrical Engineering. He is currently working toward the Ph.D. degree in Electrical Engineering at Stanford University, CA. He is a Sequoia Capital Stanford Graduate Fellow. His research interests include wireless networks, optimization, distributed control, and dynamic programming.

**Sanjay Lall** is Assistant Professor of Aeronautics and Astronautics at Stanford University. Until 2000 he was a Research Fellow at the California Institute of Technology in the Department of Control and Dynamical Systems, and prior to that he was NATO Research Fellow at Massachusetts Institute of Technology, in the Laboratory for Information and Decision Systems. He received the Ph.D. in Engineering from the University of Cambridge, England. His research interests include optimization and distributed control.