# A Sub-Gradient Algorithm For Maximal Data Extraction In Energy-limited Wireless Sensor Networks

Wei Ye
University of Southern California
GER 240, 3715 McClintock Ave.
Los Angeles, CA90089, USA
Email: yewei@usc.edu

Fernando Ordóñez
University of Southern California
GER 247, 3715 McClintock Ave.
Los Angeles, CA90089, USA
Email: fordon@usc.edu

***Abstract*— We present an efficient and implementable algorithm for maximizing data extraction from energy limited wireless sensor networks. A distinguishing feature of this algorithm is that it arrives at efficient routing solutions after few iterations, which is vital for efficient performance in energy limited networks. The algorithm uses sub-gradient optimization to solve the dual of a data extraction problem constructed by relaxing the energy constraints. We show through computational experiments that, for the problem considered, both centralized and distributed versions of the algorithm arrive at routing solutions that are on average better than 10% from optimal after only 10 iterations.**

## I. Introduction

Wireless Sensor Networks (WSN) constitute a paradigm that is already revolutionizing the availability and quality of information for many applications including battlefield and homeland security surveillance, wildfire monitoring, air quality/environmental control, manufacturing monitoring and control, and structural integrity monitoring. It is possible that in the future WSN will become integral to our everyday lives in ways that are difficult to imagine today [1].

We consider a WSN composed of nodes with sensing, processing, and communicating functions integrated into a small unit with a finite energy supply. These nodes are to be deployed in large numbers without close human supervision, possibly into unfriendly territory, for various information gathering tasks. Within its energy limits, the WSN must sense, process, and transmit information to a base station or sink node, where a remote end-user can access it. Since communication is often the most expensive operation for a sensor node, an efficient algorithm to route the data gathered is crucial to efficiently use the limited energy [2]. For instance, if every node transmits its data directly to the sink node, nodes with little data to send will be left with unused energy while data can be stranded in a node that depleted its energy. In addition, since the limited energy in effect limits the life of the network, the WSN can not spend much time coordinating a routing policy and must reach an efficient data gathering mechanism rapidly. If not it would spend a significant percentage of its lifetime operating inefficiently.

In this work we consider the problem of maximizing data extraction in energy-limited WSN. In particular, we are concerned with algorithms that can achieve close to optimal performance after only a few iterations (rounds of communications). In fact, since the actual operating conditions are subject to uncertainty, one might argue that an algorithm that reaches near-optimal routing solutions quickly without optimality guarantee is preferable to an algorithm that is guaranteed to converge to the optimal solution but takes a long time to reach efficient solutions. The paper is organized as follows: We review recent results in routing algorithms and contrast them to our work in Section II. In Section III we describe the problem, model, and a general sub-gradient algorithm. In Section IV we present three centralized variants of the sub-gradient algorithm for the maximal data extraction problem. We discuss a distributed version of the algorithm in Section V and present computational results comparing these implementations in Section VI. We finish with concluding remarks in Section VII.

## II. Literature Review

We classify recent routing protocols for sensor networks in two broad classes, which we discuss below in more detail: network dependent protocols [3,4,9-11] and optimization based protocols [2,12-14].

Network dependent routing protocols exploit various network properties for efficiency. For example LEACH [3], PEGASIS [4], and *Directed Diffusion* [9] exploit the possibility of data aggregation at the nodes to achieve important energy savings. The algorithms proposed in [10,11] rely on some high-energy agents in the network which create directed paths between the source of information and the sink. Thus the algorithms can save the energy typically used for flooding queries through the network. In general terms, since network dependent protocols exploit particular features of the network, they could conceivably perform poorly in a network without these properties, in addition it is difficult to obtain performance bounds for these routing heuristics as they do not compute or approximate an optimal solution.

Optimization based protocols are implementing an iterative optimization algorithm on some problem over the sensor network. Examples of such protocols for sensor networks include: sub-gradient algorithms for the maximum lifetime problem, i.e. maximizing the time until the first node runs out of energy, see [12-14], and approximate solutions to maximum data extraction problem [2]. The prior work on sub-gradient based methods considers additional assumptions, such as the use of potential functions, to ensure efficient performance. The sub-gradient optimization method is classic in non-linear optimization [7], and has been used to develop distributed algorithms for network flow [15] and flow control in networks without energy constraints but with fixed capacity [6]. In [2], the authors propose an approximate algorithm that uses network topology and current energy information to derive a metric with which to route the information. Optimization based protocols are both general and can provide performance bounds based on the optimization problem over the network.

We consider a WSN model similar to the one in [2], however we introduce a routing protocol based on sub-gradient optimization as opposed to the heuristic developed in [2]. Prior work on sub-gradient based protocols for WSN consider maximizing lifetime of the network, which simplifies the problem to be solved. In contrast, the network problem considered here is to maximize data extraction. Although maximum data is related to maximizing lifetime of the network, it has one significant difference: in maximum data extraction the network operates until energy in all nodes is depleted, not until the first node exhausts its energy. Maximizing data extraction is a reasonable situation in surveillance applications, for example, where nodes provide intrusion information, although with lower quality, until the last node ceases to operate. Another difference with prior optimization based algorithms on WSN is that the sub-gradient algorithm implemented in this paper does not use potential functions or additional assumptions, it solves a dual of the maximal data extraction problem.

Finally we mention that, to our knowledge, our work is the first to focus on the transient behavior of protocols in particular whether they exhibit fast convergence to efficient solutions. Prior work is usually concerned with the asymptotic convergence of the protocol. We use the optimal solution of the maximal data extraction problem as a benchmark for the protocol, as suggested in [5].

## III. PROBLEM DEFINITION

We consider a WSN with $n$ fixed sensor nodes that gather data to be sent to a sink node, denoted as node $n$+1. Let $D^i_{\max}$ be the total amount of the data (bytes) collected by the node $i$, and $E^i_{\max}$ be the total energy of the node $i$. Let $d_{ij}$ be the Euclidean distance between nodes $i$ and $j$. We denote by $N$, the set of sensor nodes, and $A$ the set of directed arcs ($i$, $j$), in the complete graph $i \in N, j \in N \cup \{n+1\}$. The energy consumed in transmitting data from one sensor to another depends on the distance between them according to the following radio model, as presented in [2,8]. We consider that a radio dissipates $\varepsilon_{\text{elec}}$=400nJ/byte to run the transmitter or receiver circuitry and $\varepsilon_{\text{amp}}$=800pJ/byte/m$^2$ for the transmitter amplifier. The transmission energy costs and receiving energy costs for a $k$-byte message and distance $d$ are given by

Transmitting: $E_{Tx}(k,d) = \varepsilon_{\text{elec}}k + \varepsilon_{\text{amp}}kd^2$

Receiving: $E_{Rx}(k) = \varepsilon_{\text{elec}}k$ .

We also assume that $k = 1$ byte and the radio channel is symmetric, so that the energy required to transmit a byte of information is the same from node $i$ to node $j$ and from node $j$ to node $i$ for a given signal to noise ratio.

### A. Mathematical Programming Model

The problem of maximizing data output given limited energy at the nodes and the above energy expenditure in transmissions and receptions can be written as a linear programming problem, see [2]. The problem constraints are:

(1) Energy Constraint: the amount of data transmitted and received by a sensor is limited by the energy available at the sensor node.

(2) Flow conservation: the amount of data transmitted by a sensor minus the amount of the data received by the sensor must be less than or equal to all data collected by the sensor and also greater than or equal to 0.

Let $f_{ij}$ be the amount of the data transmitted from node $i$ to node $j$, then the maximal data extraction problem is:

$$\begin{aligned}
\max \quad & \sum_{(i,n+1)\in A} f_{in+1} \\
\text{s.t.} \quad & \sum_{\{j|(i,j)\in A\}} f_{ij}(1+\beta d_{ij}^2) + \sum_{\{j|(j,i)\in A\}} f_{ji} \le E^i \qquad i \in N \quad (1) \\
& 0 \le \sum_{\{j|(i,j)\in A\}} f_{ij} - \sum_{\{j|(j,i)\in A\}} f_{ji} \le D^i_{\max} \qquad i \in N \quad (2) \\
& f_{ij} \ge 0 \qquad (i,j) \in A
\end{aligned}$$

where $\beta = \frac{\varepsilon_{amp}}{\varepsilon_{\text{elec}}}, E^i = \frac{E^i_{\max}}{\varepsilon_{\text{elec}}}$.

To simplify notation, we have normalized the energy in terms of receptions, that is to say, each reception consumes a unit of energy, while each transmission from $(i,j)$ consumes $1+\beta d_{ij}^2$. This model considers that each node has a maximal amount of data $D^i_{\max}$ to be transmitted. Hence if there are no energy limit constraints (1), then we can extract all data available in the network: $\max \sum_{(i,n+1)\in A} f_{in+1} = \sum_{i\in N} D^i_{\max}$.

We denote the set of flows that satisfy the routing conditions by $X$, that is:

$$X = \left\{ f \in \Re^{|A|} \mid \begin{array}{ll} 0 \le \sum_{\{j|(i,j)\in A\}} f_{ij} - \sum_{\{j|(j,i)\in A\}} f_{ji} \le D^i_{\max} & i \in N \\ f_{ij} \ge 0 & (i,j) \in A \end{array} \right\}.$$

### B. Partial Lagrangian Relaxation

We now consider the problem obtained by the Lagrangian relaxation of the energy constraints, that is incorporating these constraints in the objective with the a multiplier, or price, $p$. Let $\xi(f)$ denote the vector of energy consumption at each node given flow $f$, that is $\xi^i(f) = \sum_{\{j|(i,j)\in A\}} f_{ij}(1+\beta d_{ij}^2) +$

$\sum_{\{j|(j,i)\in A\}} f_{ji}$ for any $i \in N$. Then the Lagrangian dual function is given by

$$\begin{aligned} D(p) &= \max_{f\in X} L(f,p) \\ &= \max_{f\in X} \left\{ \sum_{(i,n+1)\in A} f_{in+1} - \sum_{i\in N} p^i \left[\xi^i(f) - E^i\right] \right\} \\ &= \max_{f\in X} \left\{ \sum_{(i,n+1)\in A} f_{in+1}(-p^i - \beta p^i d_{ij}^2 + 1) + \right. \\ &\qquad \left. + \sum_{\{(i,j)\in A|j\neq n+1\}} f_{ij}(-p^i - \beta p^i d_{ij}^2 - p^j) + \sum_{i\in N} p^i E^i \right\} . \end{aligned}$$

Set $p^{n+1} = -1$ and define

$$B(p) = \max_{f\in X} \sum_{(i,j)\in A} f_{ij}(-p^i - \beta p^i d_{ij}^2 - p^j)$$

as the part of $L(f,p)$ which involves $f$, then the Lagrangian dual becomes

$$D(p) = \max_{f\in X} L(f,p) = B(p) + \sum_{i\in N} p^i E^i.$$

Since the original problem is a linear program (LP), then its Lagrangian dual, $D : \min_{p\geq 0} D(p)$, is also an LP. Also, since both are feasible, the primal and dual attain the same finite optimal objective function value, see [7]. Therefore, we solve the dual problem to obtain the optimal objective function value, and in the process provide a routing solution that achieves this value.

### C. Sub-gradient projection method

We use the sub-gradient projection optimization method to solve $D : \min_{p\geq 0} D(p)$, see [6,7]. This method is an iterative algorithm where at each iteration $t+1$ the prices per node, $p(t+1) \in \Re^n_+$, are set by the recursion

$$p(t+1) = [p(t) - \alpha_t g(t)]^+ \ . \qquad (3)$$

Here, $[z]^+ = \max\{z, 0\}$ denotes the positive part of $z$, $g(t) \in \partial D(p(t))$ is a sub-gradient of $D(p)$ at $p(t)$, and $\alpha_t > 0$ is the step-size at the $t$-th iteration. Thus at each iteration of the sub-gradient method, we take a step in the direction of a negative sub-gradient. A sub-gradient of $D(p)$ at $p(t)$ is defined as any vector $g$ that satisfies the inequality $g^T(\overline{p} - p(t)) \leq D(\overline{p}) - D(p(t))$ for any $\overline{p}$. Given a price $p(t)$, let $f^*(t)$ denote the solution that maximizes the Lagrangian $L(f, p(t))$ over $f \in X$, and $p^*$ the optimal solution that minimizes dual problem $D(p)$.

We now present results that outline the correctness of the sub-gradient algorithm and detail how to compute the sub-gradient for this problem. We omit the proof of the well known result on convergence of the sub-gradient projection algorithm, and refer the reader to [7].

**Proposition 1**: The sub-gradient projection method with iterates defined by (3) where $0 < \alpha_t = \frac{D(p(t)) - D(p^*)}{\sum_i (g^i(t)^2)}$, converges and $p(t) \longrightarrow p^*$.

**Proof**: See [7], pages 610-612 and 629. ■

**Proposition 2:** The sub-gradients of $D(p(t))$ at $p(t)$ are given by

$$\left(\frac{\partial D(p(t))}{\partial p}\right)_i = \frac{\partial D(p(t))}{\partial p^i} = -(\xi^i(f^*(t)) - E^i)$$

**Proof:** Let $g^i(t) = \frac{\partial D(p(t))}{\partial p^i} = -(\xi^i(f^*(t)) - E^i)$ where $f^*(t) = \operatorname{argmax}_{f\in X} L(f, p(t))$. We know that

$$\begin{aligned} g^T(t)(\overline{p} - p(t)) &= - \sum_{i\in N} (\overline{p}^{\ i} - p^i(t))(\xi^i(f^*(t)) - E^i) \\ &= \sum_{j=n+1} f^*_{ij}(t) - \sum_{i\in N} (\overline{p}^{\ i} \left(\xi^i(f^*(t)) - E^i\right) \\ &\quad - [\sum_{j=n+1} f^*_{ij}(t) - \sum_{i\in N} p^i(t) \left(\xi^i(f^*(t)) - E^i\right)] \ . \end{aligned}$$

Given $\overline{p}$ let $\overline{f} = \operatorname{argmax}_{f\in X} L(f, \overline{p})$. From the definition of $D(\overline{p})$ we have

$$\begin{aligned} D(\overline{p}) &= \sum_{j=n+1} \overline{f_{ij}} - \sum_{i\in N} \overline{p}^{\ i}(\xi^i(\overline{f}) - E^i) \\ &\geq \sum_{j=n+1} f^*_{ij}(t) - \sum_{i\in N} \overline{p}^{\ i}(\xi^i(f^*(t)) - E^i) \ . \end{aligned}$$

Replacing this inequality in the previous equation and substituting the definition of $D(p(t))$ yields

$$g^T(t)(\overline{p} - p(t)) \leq D(\overline{p}) - D(p(t)) \ ,$$

which proves that $g^i(t) = -(\xi^i(f^*(t)) - E^i)$ is the sub-gradient of $D(p(t))$ at $p(t)$. ■

Although sub-gradient type algorithms converge to the optimal solution, this convergence can be very slow [7]. We are interested in studying whether, for the problem in question, this convergence is sufficiently efficient in the first iterations as to provide solutions that are already reasonably close to the optimal to be considered efficient routing heuristics.

## IV. CENTRALIZED ALGORITHM

We now describe the implementation details for the sub-gradient projection method for the maximal data extraction problem. The following implementations use a routing solution for the problem without energy constraints and decide a unique step size at each iterate. Hence, these are centralized algorithms. Although these are not realistic algorithms from an implementation point of view, we seek to determine whether these algorithms can achieve fast enough convergence to an efficient solution as to justify pursuing this approach in a distributed setting.

### A. Method 1: optimal value

Our first centralized algorithm, which is referred to as Method 1, uses the optimal value of the dual problem $D(p^*)$ to calculate the step-size. It is therefore a purely theoretical algorithm which will be used for comparison purposes. The algorithm considers a fixed integer value $m$ to control the rate of decrease of the diminishing step size, an iteration limit ITLIM, and an optimal tolerance value of TOL.

**Algorithm 1** Centralized Algorithm-Method 1:

At $t$=0, set $p^i(0) = 0$, $\forall i \in N$, or other initial values.
**while** $t \le$ ITLIM and $|D(p(t)) - D(p(t-1))| >$ TOL **do**
  Solve problem $B(p(t))$, let $f^*(t)$ be the optimal solution.
  Set
$$D(p(t)) = B(p(t)) + \sum_{i \in N} p^i(t) E^i$$
  Set $\alpha_t = \frac{m}{m+t} \frac{D(p(t)) - D(p^*)}{\sum_i (g^i(t)^2)}$,
  Set $g^i(t) = -(\xi^i(f^*(t) - E^i)$
  Compute a new price: $p^i(t+1) = [p^i(t) - \alpha_t g^i(t)]^+$
  Set $t$=$t$+1

In Algorithm 1, the formula $\frac{m}{m+t}$, satisfies the usual conditions for a diminishing step size: $\frac{m}{m+t} \xrightarrow[t \to \infty]{} 0$, and $\sum_{t=0}^{\infty} \frac{m}{m+t} = \infty$, see [7]. In the experimental section below we use $m = 1$. Since we cannot use the value $D(p^*)$ to set step sizes, our next two centralized methods consider variations of Model 1 which use a lower bound of $D(p^*)$ instead.

### B. Method 2: $\delta\_LB$

We modify Method 1 above simply by replacing a lower bound $\delta\_LB$ instead of the optimal solution value $D(p^*)$ when computing the step size.

The t-th iteration lower bound $\delta\_LB(t)$ is obtained by scaling down the flow $f(t)$ obtained at each iteration to obtain a feasible flow. This guarantees that it provides a lower bound, which is $\delta\_LB(t) = \sum_{\{i,j=n+1|(i,j)\in A\}} \delta f^*_{ij}(t)$ where $\delta = \min_{g^i(t)>0, \forall i} \frac{E^i}{\xi^i(f^*(t))}$.

Then the lower bound $\delta\_LB = \max_{t=0,1,2,\ldots} \delta\_LB(t)$

### C. Method 3: $Hop\_LB$

Method 3 also modifies Method 1 simply by constructing a lower bound to $D(p^*)$ for computing the step size. At the beginning of the centralized algorithm, we use a flow obtained by transmitting directly from each node to the sink (see Fig.1), which gives the following lower bound on $D(p^*)$: $DT\_LB = \sum_i \min\left\{D^i_{\max}, \frac{E^i}{1+\beta d^2_{ij}}\right\}$. However, we improve this lower bound by considering the feasible flow $Hop\_LB$ obtained from $DT\_LB$ by allowing flow to take a single hop to the sink if its beneficial to the system.

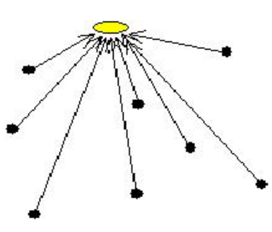

Fig.1: Direct Transmit

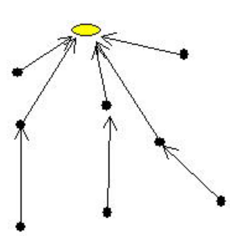

Fig.2: One-hop transmit

In Fig.2, we consider some nodes that cannot send all their collected data directly to sink node because they have limited energy and large distance to the sink. Let these nodes be the set $K$ and let nodes that have sent all their information and have residual energy be the set $L$. The $Hop\_LB$ initial solution allows a node $k \in K$ to send part of its information to the sink through some node $l \in L$, if it is beneficial to the network. We obtain this improved lower bound $Hop\_LB$ using the procedure below:

**Algorithm 2** $Hop\_LB$ procedure:

Initially, set $LB = 0, E^i = E^i_{\max}$
**while** $\frac{E^i}{1+\beta d^2_{ij}} > D^i_{\max}$ and $i \in N$ **do**
  $E^i = E^i - (1+\beta d^2_{ij}) D^i_{\max}$, $D(i) = 0$, $LB = LB + D^i_{\max}$.
**while** $D(i) = 0$ and $E^i > 0$ and $i \in N$ **do**
  find the close node $j$ with $D(j) > 0$ to sink node,
  let $f_{ij} = \min(\frac{E^i}{2+\beta d^2_{i(n+1)}}, \frac{E^i}{1+\beta d^2_{ij}})$
  $E^i = E^i - (2+\beta d^2_{i(n+1)}) f_{ij}$, $E^j = E^j - (1+\beta d^2_{ij}) f_{ij}$, $D(j) = D(j) - f_{ij}$, $LB = LB + f_{ij}$.
**while** $i \in N$ **do**
  **if** $\frac{E^i}{1+\beta d^2_{i(n+1)}} < D(i)$ **then**
    $LB = LB + \frac{E^i}{1+\beta d^2_{i(n+1)}}$
  **else**
    $LB = LB + D(i)$
$LB$ is the $Hop\_LB$.

Note that the lower bound for Method 3, $Hop\_LB$ is fixed, independent of the current iterate of the sub-gradient algorithm, and can be determined a priori. This is in contrast to Method 2 which is updated at each iteration from the current optimal solution $f^*(t)$.

## V. DISTRIBUTED ALGORITHM

The centralized algorithm assumes we can compute the step-size globally and get the optimal flow $f^*(t)$ at each time period $t$. In this section, we extend the centralized algorithm to a distributed model. In the experimental section below we show that we maintain in part the quick convergence of the sub-gradient algorithm in this distributed implementation.

The centralized algorithms discussed above are coordinated in two steps: in determining $f^*(t)$, the optimal solution to $B(p(t))$, and in setting the step size. The steps of computing the sub-gradient and the new price can be done separately at each node.

To obtain $f^*(t)$, we have to solve the following problem which has a linear objective function, which can be separated grouping all outgoing arcs of each node:

$$B(p) = \max_{f \in X} \sum_{(i,j)\in A} [f_{ij}(-p^i - \beta p^i d^2_{ij} - p^j)] \ .$$

Hence the only coordination has to do with the flow constraints $f \in X$. Our distributed algorithm approximately solves this problem by increasing or decreasing the flow at each arc $f_{ij}$ independently according to the sign of the objective cost coefficient $v_{ij} := -p^i - \beta p^i d^2_{ij} - p^j$, while maintaining a close to feasible flow with the information available. If we denote $\mu_i(t-1) = \sum_{\{j|(i,j)\in A\}} f_{ij}(t-1) - \sum_{\{j|(j,i)\in A\}} f_{ji}(t-1)$ the amount of flow that originates at node $i$ given flow $f(t-1)$, then the flow update heuristic is given by

$$f_{ij}(t)=\begin{cases}\min\left\{f_{ij}(t-1)+\max\left\{D^i_{\max}-\mu_i(t-1),1\right\},\frac{E^i}{1+\beta d^2_{ij}}\right\} & \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{if } v_{ij}>0 & \\ \max\{f_{ij}(t-1)-\max\{\mu_i(t-1),1\},0\} \quad \text{if } v_{ij}<0 & \\ f_{ij}(t-1) \qquad\qquad\qquad\qquad\qquad\qquad \text{if } v_{ij}=0 & \end{cases} \tag{4}$$

Note that the solution obtained by this heuristic can vary depending on the order in which the arcs are updated, and that in fact $f(t)$ could violate slightly the flow constraints.

To determine the step size in a distributed algorithm, we select a predetermined value at every iteration $\alpha_t=\alpha_0\frac{m}{m+t}$, where $\alpha_0$ and $m$ are fixed parameters. This rule for selecting step sizes has been shown to lead to convergent sub-gradient algorithms, see [6,7]. Algorithm 3 presents the details of the distributed algorithm.

**Algorithm 3** Distributed Algorithm

At $t$=0, set $p^i(0)=0$, $\forall i\in N$, or other initial values.
**while** $t\leq$ ITLIM and $|D(p(t))-D(p(t-1))|>$ TOL **do**
  Compute $v_{ij}=-p^i(t)-\beta p^i(t)d^2_{ij}-p^j(t)$ objective coefficients of $B(p(t))$.
  Update $f(t)$ according to (4)
  Adjust so $f(t)\in X$ by modifying $f_{ij}$'s with $v_{ij}=0$.
  Set $B(p(t))=\sum_{(i,j)\in A}f_{ij}(t)v_{ij}$.
  Set $D(p(t))=B(p(t))+\sum_{i\in N}p^i(t)E^i$.
  Set $\alpha_t=\alpha_0\frac{m}{m+t}$.
  Set $g^i(t)=-(\xi^i(f(t))-E^i)$.
  Compute a new price, $p^i(t+1)=[p^i(t)-\alpha_t g^i(t)]^+$
  Set $t$=$t+1$.

The distributed algorithm above suggests the following protocol which can be implemented on each sensor node $i$. Algorithm 4 running on each node leads to a synchronized protocol since it requires that all nodes update flows before broadcasting new prices.

**Algorithm 4** Synchronous Distributed Protocol at node $i$:

At $t$=0, set $p^i(0)=0$ and $f_{ij}(0)=0$ for $j\in N\setminus\{i\}$.
**while** $E^i>0$ **do**
  Broadcasts price $p^i(t)$ and receives $p^j(t)$ from node $j$ via arc $(i,j)$.
  Compute $v_{ij}=-p^i(t)-\beta p^i(t)d^2_{ij}-p^j(t)$.
  Computes the new flow rate $f_{ij}(t)$ for $j\in N\setminus\{i\}$ from (4).
  Transmits $f_{ij}(t)$ through arc $(i,j)$ and receives $f_{ji}(t)$.
  Set $\alpha_t=\alpha_0\frac{m}{m+t}$.
  Set new price $p^i(t+1)=[p^i(t)+\alpha_t(\xi^i(f(t))-E^i)]^+$.
  Set $t=t+1$.

## VI. COMPUTATIONAL EXPERIMENTS

Our computational experiments consider 50 sensor nodes randomly deployed in a 0.5km×0.5km area with the sink node at (0.25km, 0.5km). Each sensor node has limited power and the ability to transmit data to any other node, including the sink node. To send data, the sensor has to run its transmitter and amplifier circuitry, with parameters $\varepsilon_{\rm elec}=400$nJ/byte and $\varepsilon_{\rm amp}=800$pJ/byte/m$^2$ respectively. We assume that a reception of a single byte consumes one unit of energy. The value $\varepsilon_{\rm elec}$ is so that each 0.01J of energy allows about 25,000 receptions [2].

We present numerical results for two different scenarios: one considers homogeneous nodes, that is, all nodes have the same energy (25,000) and the same available data 10,000 (approximately 10k). The second scenario considers nodes with different ratios of $\frac{E^i}{D^i_{\max}}$. This heterogeneous scenario consider three types of nodes: high energy nodes with a ratio of 2,500, with $E^i=250,000$ and $D^i_{\max}=100$; medium energy nodes with ratio 2.5 for $E^i=25,000$ and $D^i_{\max}=10,000$; and low energy nodes with ratio 0.5 for $E^i=2,500$ and $D^i_{\max}=5,000$. The type of node in a heterogeneous scenario is selected randomly maintaining even proportions: 17 high energy nodes, 17 medium energy nodes, and 16 low energy nodes. It is clear that high energy nodes have residual energy but no data, while low energy nodes are exhausted even with substantial data left to transmit to sink node. We use AMPL and LOQO software in our computational experiments. For each problem setting, we generate 30 random instances and execute the algorithms on each instance. We report the average of the relative error to true optimal solutions: $\mathrm{Rate}=\frac{\mathrm{curr\_valueD(p(t))-optimal\_value}}{\mathrm{optimal\_value}}\times 100\%$.

### A. Computational Results

We conduct two computational experiments. The first investigates whether the centralized versions of the algorithm achieve fast convergence to an efficient solution. Our second set of experiments compare the centralized and distributed versions of the algorithm against simple routing heuristics.

1. The performance of different centralized algorithm methods. As shown in Fig.3 for the homogeneous experiment, and Fig.4 for the heterogeneous experiment, we set initial value $p^i(0)=0$. The three centralized methods converge to efficient solutions at different speeds with Method 2 being the slowest to converge. Method 1 and Method 3 both converge to 10% of the optimal value within 10 iterations.

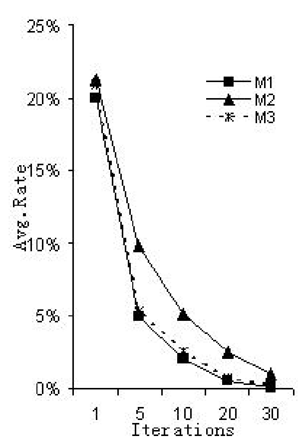


Fig.3: Homogeneous

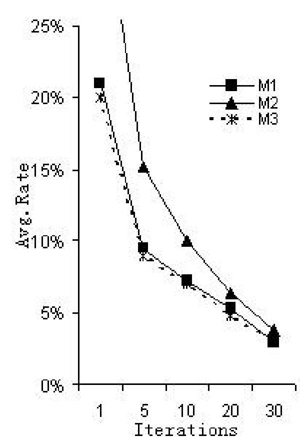


Fig.4: Heterogeneous

2. Comparison of different algorithms: Distributed Algorithms (DA), Centralized Algorithms (CA), Direct trans-

mit(DT), and LEACH.

Now, we contrast the performance of these sub-gradient based heuristics with two common protocols: algorithm DT, where each sensor node sends its information directly to the sink, and LEACH. If the sink node is located far from certain nodes, DT will only be able to transmit small amounts of data from these distant nodes, which would impair the performance of the network. Algorithm LEACH uses randomization to designate some nodes as cluster heads, which receive data from sensors in the cluster and transmit it to the sink node. The algorithm greatly improves on simple routing solutions partly because it aggregates data at the cluster heads, thus reducing the total data sent long distances to the sink. Here we consider that LEACH does not perform data aggregation at cluster heads and instead uses the centralized Method 3 to route from cluster heads.

From the first set of experiments we know that the centralized algorithm converges to within 10% of the optimal value in 10 iterations. Based on the step sizes used in these centralized methods, we set the step size for the distributed method as $\alpha_0 = 0.5 \times 10^{-7}$, with $m = 1$. We run all four algorithms on each random problem created and average the results.

As is shown in Fig.5 for the homogeneous experiments, algorithm DT is constantly within 7-8% of the optimal value and DA and LEACH perform steadily within 4-5% of the optimal. As it was observed in the previous experiment CA converges to within 10% of the optimal in 10 iterations. We set initial value $p^i(0) = 0$ for high energy nodes $i$, $p^j(0)$=0.01 for other nodes $j$ to do heterogeneous experiments. Fig.6 shows that for heterogeneous experiments the behavior is different. Algorithms DT and LEACH perform badly, at about 60% and 30% from optimal respectively. However, algorithms CA and DA both converge achieving 10% of the optimal on average in less than 10 iterations.

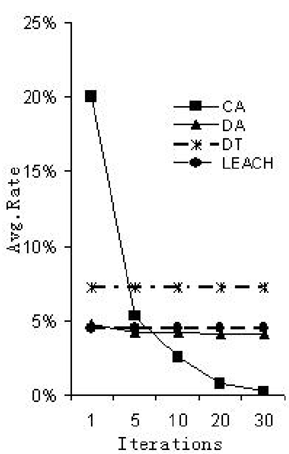


Fig.5: Homogeneous

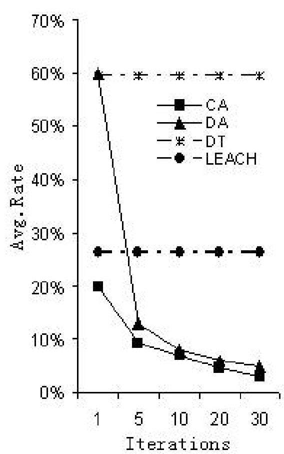


Fig.6: Heterogeneous

## VII. CONCLUSION

We formulate the maximal data extraction problem in energy limited WSNs as a linear programming problem. By using Lagrangian relaxation on the energy constraints we formulate a related dual problem amenable to a solution via a sub-gradient projection method. We present both a centralized and a distributed version of this algorithm and show through computational experiments that these algorithms achieve close to optimal performance quickly (achieving 10% of optimal in less than 10 iterations on average). Although convergence of sub-gradient methods can be slow in theory, for our problem they quickly arrive at an efficient routing heuristic.

## VIII. ACKNOWLEDGMENTS

The work was supported by a Research Initiation Grant from USC's James H. Zumberge Research and Innovation Fund.

## IX. REFERENCES

[1] J. Elson and D. Estrin, "WSN: A bridge to the physical world," *Wireless Sensor Networks*, Kluwer, 2004.

[2] N. Sadagopan, and B. Krishnamachari, "Maximizing data extraction in energy-limited sensor networks," *INFOCOM*, 2004.

[3] W. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for WSN," *Proc. of the Intl' Conference on System Sciences*, Hawaii, 2000.

[4] S. Lindsey and C. S. Raghavendra, "PEGASIS: Power Efficient GAthering in Sensor Information Systems," *ICC*, 2001.

[5] F. Ordóñez and B. Krishnamachari, "Optimal information extraction in energy-limited wireless sensor networks," *IEEE JSAC*, vol.22(6), pp.1121-1129, 2004.

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

[7] D. Bertsekas, *Nonlinear Programming*, Athena Scientific, second edition, 1999.

[8] S. Lindsey, C. Raghavendra and K. Sivalingam, "Data Gathering in sensor networks using energy delay metric," *Proc. of Parallel and Distributed Computing Issues in WNMC*, San Francisco, 2001.

[9] C. Intagagonwiwat, R. Govindan, and D. Estrin, "Directed Diffusion: A scalable and robust communication paradigm for sensor Networks," *MobiCom*,2000.

[10] D. Braginsky and D. Estrin, "Rumor Routing Algorithm for sensor networks, *Proc. of the 1st ACM Intl' Workshop on WSN & Applications*, Atlanta, 2002.

[11] D. Tian and N. D. Georganas, "Energy efficient routing with guaranteed delivery in WSN," *IEEE WCNC*, 2003.

[12] A. Sankar and Z. Liu, "Maximum lifetime routing in wireless ad-hoc networks," *INFOCOM*, 2004

[13] R. Madan and S. Lall, "Distributed algorithms for maximum lifetime routing in WSN," *IEEE GLOBECOM*, 2004.

[14] C. Hua and T. Yum, "Optimal routing for maximizing lifetime of WSN," *INFOCOM*, 2005.

[15] P. Tseng, D. P. Bertsekas, and J. N. Tsitsiklis, "Partially asynchronous, parallel algorithms for network flow and other problems," *SIAM J. Control and Optimization*, vol. 28, pp. 678-710, 1990.