Joint Power Control, Scheduling and Routing for Multicast in Multihop Energy Harvesting Sensor Networks

Vinay Joseph[1], Vinod Sharma[2], Utpal Mukherji[3] and Manjunath Kashyap[4]
Department of Electrical Communication Engineering, Indian Institute of Science, Bangalore, India.[1,2,3]
Department of Electrical Engineering, Indian Institute of Technology, Madras.[4]
Email:{vinay,vinod,utpal}@ece.iisc.ernet.in[1,2,3], manjunathkashyap@smail.iitm.ac.in[4]

***Abstract*—We consider the problem of joint power control, scheduling and routing in energy harvesting sensor networks allowing for multicast of data generated at the sensor nodes to a set of sink nodes. In this setup we exploit broadcast nature of the channels and network coding to show performance improvement. We also develop computationally efficient suboptimal algorithms and study their performance.**

Keywords: Energy harvesting, Sensor networks, Power Control, Multicast, Energy efficient routing, Joint routing-scheduling, Network Coding.

I. Introduction

We study wireless multihop energy harvesting sensor networks employed for random field estimation. The sensors sense a random field and generate data that is to be sent to sink node for estimation. Each sensor has an energy harvesting source and can operate in two modes: Wake and Sleep. The sleep mode is a power saving mode in which the sensor only harvests energy and performs no other functions so that the energy consumption is negligible. In the wake mode, a sensor node can sense and process data and communicate with other nodes.

The problem of scheduling and routing in adhoc multihop wireless networks has been extensively studied in recent years (see [2], [3] for general surveys and tutorials). In [4], the authors obtain a static multipath routing algorithm for networks with energy replenishment which is optimal with respect to a set of pre-computed paths computed heuristically. But, they do not consider power control. The general problem of optimal joint routing and scheduling with power control in wireless multihop networks is NP-complete ([5]). In [6], the authors considered this problem with the objective of minimizing the total average transmission power with a simplifying linear approximation for the rate on a wireless link. The problem was considered in a sensor network setting in [7] with an objective of minimizing the energy consumption. But, they consider only Time Division Multiple Access schemes. In [8], non-linear column generation is used to obtain an approach that converges to the global optimum solution. However, the column generation subproblem is also typically NP-complete. In [5], the authors formulate the problem as an optimization problem involving time sharing across different transmission modes and provide a computationally tractable approach to obtain good solutions to the problem.

9781-4244-3941-6/09/$25.00 Â©2009 IEEE

Most of the early wireless sensor networks involved only a single sink that collected data from a number of nodes. Recently, however, scenarios with multiple sinks are increasingly being proposed. Multiple sinks are required to implement advanced applications and programming abstractions ([9]). An example is that of a sensor network with actuator nodes where different actuators are likely to need data coming from the same set of source nodes. Geocasting ([10]), data centric storage ([11]) are some other applications of wireless sensor networks where multicasting is used.

The emerging area of Network Coding has often been found useful in networks with multiple sinks and multicast connections. The idea of network coding was introduced in the seminal paper [12]. Network coding can improve throughput, robustness, complexity, and security [1]. It has been shown in [13] that linear network coding is enough to achieve the Min-Cut, Max-Flow bounds in wired networks. Network coding for wireless networks is developed in [15] and [16]. But these papers, like all the papers mentioned above, have restricted their attention to intra-session network coding. When there are multiple sessions, intra-session coding is suboptimal: in general inter-session coding is necessary to achieve optimal throughput ([17]). However, performing inter-session coding is difficult and to perform inter-session coding optimally, linear coding operations are not sufficient ([18]). Thus in this paper we will limit ourselves to intra-session network coding.

In this paper, we consider the problem of joint power control, scheduling and routing for multicast in energy harvesting sensor networks. This generalizes the work in [24]. We consider a general model for the network and traffic flowing through the network allowing for multicast of data generated at the sensor nodes to a set of sink nodes. Our problem can be considered as a generalization of multihop wireless networks where sleep and wake option may or may not be there and modes may not be harvesting energy.

We consider three approaches to solve the problem. Though these we will demonstrate the advantages of braodcasting and network coding in the present setup. The approaches become computationally intractable for large networks. For two of the approaches, we also develop a computationally efficient suboptimal approach to solve the problem.

The main contributions of the paper can be summarized as follows. We study and compare three important approaches to solve the problem of joint power control, scheduling and routing in energy harvesting sensor networks with special attention given to scenarios where multicast traffic originates at the nodes in the network. In our formulations, we consider several important aspects of energy harvesting sensor networks including power saving modes, and power consumed for various activities like transmission, reception and sensing. We compare the approaches through insightful examples. Further, we develop computationally tractable suboptimal approaches to solve the problem.

This paper is organized as follows. In Section II, we discuss the problem setting. In the next three sections, we present the three approaches to solve the problem. In Section VI, we evaluate the three approaches using examples. Section VII concludes the paper.

II. Model

We consider a wireless sensor network consisting of N stationary energy harvesting sensor nodes and t sink nodes, with the energy harvesting nodes sensing a random field. Each sensor node generates packets to be transmitted to a certain subset of sink nodes. The sink nodes only collect data and do not generate or send data. Let the set of sink nodes corresponding to sensor node n be $\mathbf{D}_n$. We consider multicast connections where the data generated at node n is sent to all the sink nodes in $\mathbf{D}_n$. The system is slotted with each slot being of length T seconds. In any slot, a sensor node is in one of two modes: Wake or Sleep. In the sleep mode, a node can only generate energy via an energy harvesting source. In the wake mode, it can sense, process, receive and/or transmit data. We assume that the energy consumed by a sensor node in the sleep mode is negligible. In the wake mode, a sensor cannot transmit and receive in the same slot (i.e., it has a half duplex channel). For transmission, a sensor node chooses its transmit power from a finite set of transmit powers (CC2420 ([19]) and ADF7020 ([20]) are commercially available low power transceiver chips commonly used in sensor nodes that have a programmable output power). We assume that all the nodes choose their transmit powers from the same set consisting of K elements. In most of these low power transceivers, during transmission, only a small fraction of the total power consumed appears as transmit power. We will denote the extra power consumed in the transceiver as e_t (which is independent of the transmit power chosen). The power consumption during reception is approximately constant and is denoted by e_r. Let the constant power consumed by the sensor for sensing data be e_s and due to other activities (e.g., processing, leakage etc) be e_w.

Due to the broadcast nature of wireless links, each node can potentially transmit to every other node although the link quality will depend upon the distances between the nodes and possible shadowing and other physical obstacles in the way. Let $\mathcal{N}$ denote the set of all sensor nodes in the network. Let $\mathcal{O}(n)$ denote the set of outgoing links and $\mathcal{I}(n)$ the set of incoming links of sensor node n. We denote the set of all links in the network by $\mathcal{L}$ and let $L = |\mathcal{L}|$. We assume that a node can receive from at most one other node at any time.

If sensor node n is sensing in slot k (which denotes time interval $[(k-1)T, kT]$ seconds), X_k^n bits are generated in the slot. We assume that $\{X_k^n; k \geq 1\}$ is a stationary, ergodic random process. Then, sensor n when it is sensing is effectively generating data at a rate $d_n = E[X_k^n]/T$. Energy harvesting can replenish energy by Y_k^n at sensor node n in slot k. We assume that $\{Y_k^n; k \geq 1\}$ is a stationary, ergodic random process. Then energy is harvested at a rate $P_n^{av} = E[Y_k^n]/T$ by sensor node n. We assume that each node has infinite data and energy buffers. This is a good approximation if the buffers are large enough (see [22] for more discussion and generalizations).

We define a mode of the network as a possible combination of active links along with their transmit powers such that a node cannot transmit and receive at the same time and can receive from at most one other node at any time. Since each sensor node chooses its transmit power from a finite set, the number of modes is finite. Let $\mathcal{M}$ denote the set of all modes. Corresponding to any transmission mode m, we have a link transmit power vector $\mathbf{P}^m = (P_1^m, ..., P_L^m)$ where the lth element, P_l^m, is the transmit power of link l in mode m.

Let the channel gain from the transmitting sensor node of link k to the receiver node of link l be h_{kl}. This includes the effect of fading, shadowing and attenuation due to distance. We assume slow fading which is realistic for a sensor network with stationary nodes. Hence, in the following, h_{kl} will be assumed constant $\forall\ k, l \in \mathcal{L}$ (thus the algorithm developed will run periodically with updated channel gains). For any link transmit power vector $\mathbf{P}$, the SINR (Signal to Interference + Noise Ratio) at the receiver of link l is

$$\gamma_l(\mathbf{P}) = \frac{h_{ll} P_l}{\sum_{k \neq l} h_{kl} P_k + \sigma_l^2}$$

where σ_l^2 is the noise power at the receiver of link l. The data rate, $R_l(\mathbf{P})$, over link l can be taken as a non-decreasing concave function of $\gamma_l(\mathbf{P})$. (Many low power transceivers, e.g., ADF7020 ([20]), ATA542X ([21]) can transmit at multiple rates using Adaptive Modulation and/or Coding). In this paper, we use the Shannon formula

$$R_l(\mathbf{P}) = W log_2(1 + \gamma_l(\mathbf{P})),$$

where W is the bandwidth of the wireless channel.

The focus of this paper is the problem of joint power control, scheduling and routing for multicast connections in an energy harvesting sensor network to maximize the throughput of the network in a fair manner. We have seen in [22] and [23] that for energy neutral operation in steady state, it is necessary and sufficient for the nth node to use average power less than P_n^{av}. Given the traffic requirements d_n of each sensor node n, we obtain energy neutral policies which satisfy these requirements. If it is not possible, then these policies satisfy their requirements in a "fair" (to be explained below) way. Also, we exploit the energy saving sleep mode to obtain policies that can satisfy the traffic requirements better. We will find that complexity of these problems is high and hence we develop a computationally efficient suboptimal approach to obtain good solutions to this problem.

We discuss three approaches to solve this problem. The first approach involves a simple extension of the problem solution discussed in [24], which provides a routing scheme using Multicommodity flow model. We refer to this approach as APP-R. Next, we discuss an approach involving the routing of data belonging to a multicast session through Steiner trees ([27]). We refer to this approach as APP-T. In the last approach, we consider the use of intra-session network coding and we refer to this approach as APP-NC. These approaches are discussed in the following sections and compared in Section VI.

III. Approach 1: Using Multicommodity flow model

In this section, we discuss a solution of the problem using an extension of the approach used in [24]. We consider all the data flowing to a sink in the network as belonging to one flow. Denote the sink of a flow f by $t(f)$ and the set of source nodes of flow f by $\mathbf{S}(f)$. Denote the set of all flows by $\mathcal{F}$. Let r_n^f denote the rate at which data is transmitted to $t(f)$ from a node $n \in \mathbf{S}(f)$ so that the total rate of flow arriving at $t(f)$ is $\sum_{n\in\mathbf{S}(f)} r_n^f$. To obtain the routing, we use a multi-commodity flow model ([25]) where each flow f corresponds to a commodity in the network. Let x_l^f denote the traffic assigned to link l by the routing scheme for flow f. The flow assignment given by the routing scheme must satisfy the flow conservation constraints:

$$\sum_{l\in\mathcal{O}(n)} x_l^f - \sum_{l\in\mathcal{I}(n)} x_l^f = 0, \forall\ n\in\mathcal{N}\backslash\mathbf{S}(f),\ \forall\ f\in\mathcal{F},$$
$$\sum_{l\in\mathcal{O}(n)} x_l^f - \sum_{l\in\mathcal{I}(n)} x_l^f = r_n^f, \forall\ n\in\mathbf{S}(f),\ \forall\ f\in\mathcal{F},$$
$$\sum_{l\in\mathcal{I}(t(f))} x_l^f = \sum_{n\in\mathbf{S}(f)} r_n^f, \forall\ f\in\mathcal{F};$$
$$\sum_{l\in\mathcal{O}(t(f))} x_l^f = 0,\ \forall\ f\in\mathcal{F}.$$

The above equations can be compactly written as $\boldsymbol{A}\mathbf{x}^f = \mathbf{y}_R^f,\ \forall\ f\in\mathcal{F}$ where $\boldsymbol{A}$ is $(N+t)$xL node-link incidence matrix with $a_{nl}=1$ if $l\in\mathcal{O}(n)$, $a_{nl}=-1$ if $l\in\mathcal{I}(n)$ and $a_{nl}=0$ otherwise. $\mathbf{x}^f$ is length L column vector with its lth element being x_l^f and $\mathbf{y}_R^f$ is an $N+t$ length column vector with $\left[\mathbf{y}^f\right]_n = r_n^f,\ \forall\ n\in\mathbf{S}(f),\ \left[\mathbf{y}^f\right]_{t(f)} = -\sum_{n\in\mathbf{S}(f)} r_n^f$ and all other elements zero.

We can formulate the problem as follows:

OPT-R: maximize λ subject to

$$\frac{r_n^f}{d_n} \geq \lambda,\ \forall\ n\in\mathbf{S}(f),\ \forall f\in\mathcal{F}; \tag{1}$$

$$\boldsymbol{A}\mathbf{x}^f = \mathbf{y}_R^f,\ \forall f\in\mathcal{F}; \tag{2}$$

$$\sum_{f\in\mathcal{F}} x_l^f \leq \sum_{m\in\mathcal{M}} \alpha_m R_l(\mathbf{P}^m) - \delta,\ \forall l\in\mathcal{L}; \tag{3}$$

$$\sum_{m\in\mathcal{M}} \alpha_m \left[\mathbf{q}^m\right]_n + u_{sn}e_s + u_{wn}e_w \leq P_n^{av} - \epsilon,\ \forall\ n\in\mathcal{N}; \tag{4}$$

$$\frac{r_f^n}{d_n} \leq u_{sn},\ \ \forall\ n\in\mathbf{S}(\{\}),\forall\ f\in\mathcal{F};\ u_{sn}\leq u_{wn},\ \ \forall\ n\in\mathcal{N}; \tag{5}$$

$$\sum_{m\in\mathcal{M}} \alpha_m \left(I(\sum_{l\in\mathcal{O}(n)} P_l^m > 0) + \sum_{l\in\mathcal{I}(n)} I(P_l^m > 0)\right) \leq u_{wn},\ \ \forall\ n\in\mathcal{N}; \tag{6}$$

$$\sum_{m\in\mathcal{M}} \alpha_m \leq 1 \tag{7}$$

$$x_l^f \geq 0,\ \forall l\in\mathcal{L},\ \forall\ f\in\mathcal{F};\ r_n^f\geq 0,\ \forall\ n\in\mathbf{S}(f),\ \forall f\in\mathcal{F}; \quad \alpha_m\geq 0,\ \forall m\in\mathcal{M}; u_{wn}\geq 0, u_{sn}\geq 0\ \forall\ n\in\mathcal{N} \tag{8}$$

where $\{\mathbf{q}^m : m\in\mathcal{M}\}$ are given by

$$\left[\mathbf{q}^m\right]_n = \sum_{l\in\mathcal{O}(n)} (P_l^m + e_t I(P_l^m > 0)) + e_r \sum_{l\in\mathcal{I}(n)} I(P_l^m > 0).$$

This is a linear programming problem where we maximize the minimum, λ, of fractions r_n^f/d_n of data carried by sensor nodes to the associated destinations (thus the solution is *fair* in this sense). The variable α_m is the fraction of time the network operates in mode m. The inequalities in (3) state that the traffic through any link should be less than its capacity. This is a necessary condition for stability of the data queue corresponding to link l. Under some weak conditions, it is also sufficient. The inequalities (4) state that the average power consumed at each node n should be slightly less than P_n^{av}. This ensures energy neutral operation at each sensor node ([26], [22]). We do not consider any power constraints on the sink nodes whose main activity is reception (we assume that it has enough power to receive data in every slot). u_{wn} and u_{sn} denote the fractions of time sensor node n is in wake mode and the fraction of time it is performing sensing, respectively. Thus, the sensor node spends energy at the rate of e_w in u_{wn} fraction of time and at the rate of e_s in u_{sn} fraction of time. This is taken into account while writing inequalities

(4). Since a sensor n is sensing and thus generating data only for u_{sn} fraction of time, r_n^f/d_n for each associated flow f is upper bounded by u_{sn} and a sensor can perform sensing only when it is in the wake mode. Thus, we have the inequalities (5). Similarly, the fraction of time a sensor is transmitting or receiving is limited by the fraction of time it is in the wake mode which is stated in inequalities (6). In addition, the sleep slots of the sensors can be scheduled so that the sensors are awake for the fraction u_{wn} and they sense for u_{sn}/u_{wn} fraction of time they are awake.

By solving **OPT-1**, we obtain:
1) Optimal λ which we denote by λ^*: If $\lambda^* \geq 1$, then the network can carry all the traffic originating from the nodes to the associated sink nodes. But, if $\lambda^* < 1$, there are sensor nodes that are generating more data than what can be sent by the network to the sink nodes. These nodes will drop the excess data.
2) Optimal (mode scheduling) fractions of time $\{\alpha_m^*, m \in \mathcal{M}\}$ that the network operates in the various valid modes.
3) Optimal x_l^{*f} and the fair share of traffic r_f^* carried by the network for flow f. We briefly describe a simple scheduling and routing scheme that can be obtained using the optimal solution. At each node n, corresponding to each outgoing link l, we have a queue for each flow f for which $x_l^{*f} > 0$ and the fraction $x_l^{*f} / \sum_{i \in \mathcal{O}(n)} x_i^{*f}$ of flow f arriving at node n gets buffered in this queue. Whenever the outgoing l gets served (i.e., the network is in a mode with link l active), the queues corresponding to link l can be served using Weighted Round Robin with the weight for flow f being $x_l^{*f} / \sum_{j \in \mathcal{F}} x_l^{*j}$.

We see that the problem formulation presented above is an extension of the approach presented in [24] to multiple destinations. But, it would often lead to a suboptimal solution. This is due to the fact that the above formulation does not consider the use of broadcasts. Intuitively, it is clear that in a problem setting with multiple destinations, broadcasting will be useful. We illustrate this using a simple example. Consider the network shown in Fig. 1 with one sensor node S and two sinks $t(1)$ and $t(2)$. Suppose that the energy harvesting rate at S is large enough so that it can sense at 100 bps and transmit data at a maximum rate 100 bps in all the slots. Then, using the above approach, we see that the optimal λ is only 0.5, whereas it is clear that using a broadcast all the time, we can achieve $\lambda = 1$. (Note that in this example, we have made an assumption on the rate of broadcast, viz rate of broadcast is equal to the minimum of the rates at which the transmitter can unicast to the receivers individually. We will consider this point further in the next section.)

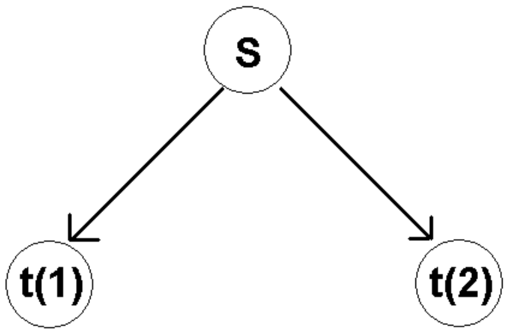

Fig. 1. Example: Need for broadcasts in multicast connections

IV. Approach 2: Using Steiner Trees for routing

In this section, we use the idea discussed in [27] for formulating the problem involving multicast connections in a network where the nodes also use broadcasts. Broadcast is useful only when there are multiple sinks.

[27] considers directed trees rooted at a source and reaching all associated sinks as the basic unit of flow for multicast network flow. They define a multicast flow as a collection of flows along directed trees (rooted at a source and reaching all its associated sinks) obeying the capacity constraints for each link. Thus, any multicast flow f from a source node $n = s(f)$ is identified with such a collection of directed trees $\mathbf{K}_n$ and with the associated flows $\{u_n^k : k \in \mathbf{K}_n\}$ obeying the following capacity constraints:

$$\sum_{\{k \in \mathbf{K}_n, n \in \mathcal{N} : l \in k\}} u_n^k \leq z_l, \quad \text{for every link } l \tag{9}$$

$$\text{and} \quad u_n^k \geq 0, \ \forall k \in \mathbf{K}_n \tag{10}$$

where z_l denotes the rate of link l (which can be obtained from the values of $\{\alpha_m : m \in \mathcal{M}\}$). Let the set of all such multicast flows be denoted by $\mathcal{F}_T$.

We have assumed that a node can broadcast to any number of its neighbours and if it can transmit at rates r_1, r_2, ..., r_{k-1} and r_k to k of its neighbours individually, then it can broadcast at the rate $min(r_1, r_2, ..., r_k)$ to those neighbours. Along with the introduction of broadcasts at a node, we also introduce virtual nodes (and associated virtual links) as done in [16]. The virtual node plays the role of an artificial bottleneck which constrains the rate of distinct information going out of the transmitter. We will illustrate the need for such a bottleneck using a small example. In Fig. 2, node S is multicasting data to sinks $t(1)$ and $t(2)$. Assume that the $(s, 1)$ and $(s, 2)$ links have very high capacity (so that they are not bottlenecks). Also, whenever nodes 1 and 2 transmit, they broadcast to $t(1)$ and $t(2)$ at the rate of 2 units. Let us consider two of the trees satisfying the conditions discussed above, shown in Fig. 2(b) and Fig. 2(c). Without a virtual node, if we just use the value $z_l = 2$ while forming the rate constraint equations (9) on links $l \in \{(1, t(1)), (2, t(1)), (1, t(2)), (2, t(2))\}$, we can see that node S can send 2 distinct units of data to each of the destinations, i.e., 4 distinct units of information are carried to the two sinks combined. This clearly is not possible as the maximum rate of the nodes that transmit to the sinks is only 2 units so that the maximum number of distinct units of information reaching sinks has to be less than or equal to 2. Once we place the virtual nodes (as shown in Fig. 2(d)), the virtual links $(1, 1')$ and $(2, 2')$ act as bottleneck links and limit the total rate of data flowing to the two destinations to 2 distinct units. See [16] for more details.

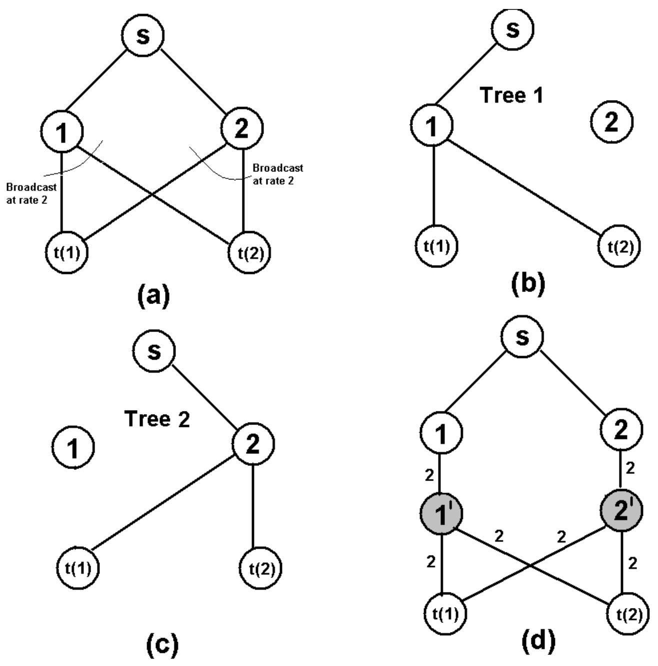

Fig. 2. Example: Need for virtual nodes

Once we allow for broadcasts, the set of all modes is much larger as we have to include the modes that involve broadcasts too. Let us denote the set of all modes in the present setup by $\mathcal{M}_B$. Further, due to the addition of virtual nodes, we also have a larger set of nodes $\mathcal{N}_B = \mathcal{N} \cup \mathcal{N}_V$ where $\mathcal{N}_V$ is the set of virtual nodes. Similarly, denote the new larger set of links by $\mathcal{L}_B$.

Now, we can state the problem formulation as we have done in the previous section:

OPT-T: maximize λ subject to

$$\frac{r_{s(f)}}{d_{s(f)}} \geq \lambda, \ \forall f \in \mathcal{F}_T; \tag{11}$$

$$r_n \leq \sum_{k\in\mathbf{K}_n} u_n^k \ \forall \ n \in \mathcal{N} \tag{12}$$

$$\sum_{\{k\in\mathbf{K}_n, n\in\mathcal{N}:\updownarrow\in||\}} u_n^k \leq \sum_{m\in\mathcal{M}_B} \alpha_m R_l(\mathbf{P}^m) - \delta, \ \forall \ l \in \mathcal{L}_B \tag{13}$$

$$\sum_{m\in\mathcal{M}_B} \alpha_m \left[\mathbf{q}^m\right]_n + u_{sn}e_s + u_{wn}e_w \leq P_n^{av} - \epsilon, \ \forall \ n \in \mathcal{N}; \tag{14}$$

$$\frac{r_n}{d_n} \leq u_{sn}, \ \ \forall \ n \in \mathcal{N}; \ u_{sn} \leq u_{wn}, \ \ \forall \ n \in \mathcal{N}; \tag{15}$$

$$\sum_{m\in\mathcal{M}_B} \alpha_m \left(I\left(\sum_{l\in\mathcal{O}(n)} P_l^m > 0\right) + \sum_{l\in\mathcal{I}(n)} I(P_l^m > 0)\right) \leq u_{wn}, \ \ \forall \ n \in \mathcal{N}; \tag{16}$$

$$\sum_{m\in\mathcal{M}_B} \alpha_m \leq 1 \tag{17}$$

$$r_n \geq 0, \ \forall n \in \mathcal{N}; \ \alpha_m \geq 0, \ \forall m \in \mathcal{M}_B;$$
$$u_{wn} \geq 0, u_{sn} \geq 0 \ \forall \ n \in \mathcal{N}, \ \ u_n^k \geq 0 \ \forall k \in \mathbf{K}_n, \forall \ n \in \mathcal{N}.$$

Here, (9) has been rewritten as (13), and (12) states that the rate of multicast from each node n should be less than the sum of the rates associated with the collection of directed trees $\mathbf{K}_n$ from the node.

Note that for large networks, finding the $\mathbf{K}_n$ for each node n has high computational complexity and the number of such trees is very large. For instance, solving the above problem for $N > 4$ with two sinks takes too long to compute. Thus, for large networks, we need some computationally efficient suboptimal approaches like the ones given in [27].

We will see that this approach provides improvements on the approach provided in the previous section. We present some examples in Section VI to illustrate this. But even with this approach, we cannot always attain the best performance. For instance, use of network coding could significantly improve the performance and this will be discussed in the next section.

V. Approach 3: Using Network Coding

In this section, we formulate and solve the problem of supporting multicast connections over the network by allowing the nodes to perform network coding. The traffic from the sensor nodes multicasting to the same set of sinks forms a session. We denote the set of all sessions by $\mathcal{S}$.

When there are multiple sessions sharing the network, we allocate disjoint subsets of the network capacity to each session. If each session's allocations satisfy the Max-Flow Min-Cut condition for each sink (associated with the session) separately we can obtain a solution (transmitting flows of a session to all its destinations simultaneously) with intra-session network coding among information symbols of each session separately. Although suboptimal (because we are not doing inter-session network coding even when there are multiple sessions), this approach provides considerable gains compared to the two approaches presented earlier.

Given a session $s \in \mathcal{S}$, for each sink which is part of the multicast connection, we have an associated flow $f \in \mathcal{F}(s)$ where $\mathcal{F}(s)$ denotes the set of all flows of session s. Thus, the number of flows in a session is equal to the number of sinks associated with the multicast connection in session s. Denote the sink of a flow f by $t(f)$ and the set of source nodes of flow f by $\mathbf{S}(f)$. Let r_n^f denote the rate at which data is transmitted to $t(f)$ from a node $n \in \mathbf{S}(f)$; the total rate of flow arriving at $t(f)$ is then $\sum_{n\in\mathbf{S}(f)} r_n^f$. The rest of the notation is as in the last section.

The flow conservation equations for each flow $f \in \mathcal{F}(s)$ associated with each session $s \in \mathcal{S}$ are given as in Section III, by $\boldsymbol{A}\mathbf{x}^f = \mathbf{y}_R^f$.

Now, we formulate the problem as follows:

OPT-NC: maximize λ subject to

$$\frac{r_n^f}{d_n} \geq \lambda, \ \forall \ n \in \mathbf{S}(f), \ \forall f \in \mathcal{F}(s), \ \forall \ s \in \mathcal{S}; \tag{18}$$

$$\boldsymbol{A}\mathbf{x}^f = \mathbf{y}_R^f, \ \forall f \in \mathcal{F}(s), \ \forall \ s \in \mathcal{S}; \tag{19}$$

$$x_l^f \leq g_l^s \ \forall f \in \mathcal{F}(s), \ \forall \ s \in \mathcal{S}, \ \forall l \in \mathcal{L}_B / \bigcup_{n \in \mathcal{N}_v} \mathcal{O}(n); \tag{20}$$

$$\sum_{f \in \mathcal{F}(s)} x_l^f \leq g_l^s, \ \forall \ s \in \mathcal{S}, \ \forall l \in \bigcup_{n \in \mathcal{N}_v} \mathcal{O}(n); \tag{21}$$

$$\sum_{s \in \mathcal{S}} g_l^s \leq \sum_{m \in \mathcal{M}_B} \alpha_m R_l(\mathbf{P}^m) - \delta, \ \forall l \in \mathcal{L}_B; \tag{22}$$

$$\sum_{m \in \mathcal{M}_B} \alpha_m \left[\mathbf{q}^m\right]_n + u_{sn} e_s + u_{wn} e_w \leq P_n^{av} - \epsilon, \ \forall \ n \in \mathcal{N}; \tag{23}$$

$$\frac{r_n^f}{d_n} \leq u_{sn}, \ \forall \ n \in \mathcal{S}(f), \ \forall f \in \mathcal{F}(s), \ \forall s \in \mathcal{S}; \quad u_{sn} \leq u_{wn}, \ \forall \ n \in \mathcal{N}; \tag{24}$$

$$\sum_{m \in \mathcal{M}_B} \alpha_m \left(I \left(\sum_{l \in \mathcal{O}(n)} P_l^m > 0 \right) + \sum_{l \in \mathcal{I}(n)} I(P_l^m > 0) \right) \leq u_{wn}, \ \forall \ n \in \mathcal{N}; \tag{25}$$

$$\sum_{m \in \mathcal{M}_B} \alpha_m \leq 1 \tag{26}$$

$$x_l^f \geq 0, \ \forall \ l \in \mathcal{L}_B, \ \forall \ f \in \mathcal{F}(s), \ \forall \ s \in \mathcal{S};$$
$$\alpha_m \geq 0, \ \forall m \in \mathcal{M}_B;$$
$$r_n^f \geq 0, \ \forall \ n \in \mathbf{S}(f), \ \forall f \in \mathcal{F}(s), \ \forall \ s \in \mathcal{S};$$
$$u_{wn} \geq 0, u_{sn} \geq 0 \ \forall \ n \in \mathcal{N}; g_l^s \geq 0 \ \forall \ s \in \mathcal{S}, \forall \ l \in \mathcal{L}_B.$$

where $\{\mathbf{q}^m : m \in \mathcal{M}_B\}$ are given by

$$[\mathbf{q}^m]_n = (P_l^m + e_t) I(\sum_{l \in \mathcal{O}(n)} P_l^m > 0) + e_r \sum_{l \in \mathcal{I}(n)} I(P_l^m > 0).$$

Note that all equations in the above formulation are similar to the ones in Section III except for (20) and (21). We have introduced new variables $\{g_l^s; \ \ s \in \mathcal{S}, \ l \in \mathcal{L}_B\}$ in this formulation. For a link l that is not an outgoing link from a virtual node (i.e., $l \notin \bigcup_{n \in \mathcal{N}_v} \mathcal{O}(n)$), (20) ensures that the variable g_l^s is greater than or equal to all the $\left\{x_l^f; f \in \mathcal{F}(s), s \in \mathcal{S}\right\}$. Thus, for these links, only the maximum of the variables $\left\{x_l^f; f \in \mathcal{F}(s), s \in \mathcal{S}\right\}$ are considered in the rate constraints (22). This ensures that allocations of each session satisfy the Max-Flow Min-Cut condition for each sink (or each flow) associated with the session and then, we obtain a solution with intra-session network coding among information symbols of each session separately. Also note that for a link l that is an outgoing link from a virtual node (i.e., $l \in \bigcup_{n \in \mathcal{N}_v}$), we have the usual flow conservation and rate constraint equations as the virtual nodes do not exist and thus are not capable of performing network coding. It has been shown in [16] that there is no need to perform network coding on the virtual nodes as far as throughput is concerned.

A. Remarks

The problem formulations discussed in the previous sections have high computational complexity and hence become intractable for large networks. For OPT-R and OPT-NC, we can use a suboptimal approach similar to the one discussed in [24], [5] . The method uses an iterative approach that considers a smaller set of modes when solving the optimization problem and then updates the set of modes using column generation. Since the method can be easily extended to the current setting also, we will skip a detailed discussion. We will refer to the method as ALGO-M in the next section.

VI. Examples

In this section, we study the performance of the three approaches presented in the paper using some examples. We focus on
(a) The improvement in performance with each approach and the dependence of the improvements on the network topology
(b) Performance of the suboptimal approach that uses ALGO-M to find solution for OPT-R and OPT-NC.

For all the examples in this section, the following settings apply: $e_t = 60mW$, $e_r = 60mW$, $e_s = 2mW$ and $e_w = 20mW$. Each node in the network chooses its transmit power levels from the set $\{0, 0.1, 1, 10\}$ mW.

In Figs. 3(a)-3(c), we consider networks consisting of 4 sensor nodes that multicast to two sinks and in Fig. 3(d), we consider a network consisting of 3 sensor nodes that multicast to two sinks. All the nodes generate data at the rate of 1000 *bps*, i.e., $d_n = 1000$ *bps* for all $n \in \mathcal{N}$. Also, $P_n^{av} = 20$ mW for all $n \in \mathcal{N}$.

For the network shown in Fig. 3(a), the solution obtained using OPT-T provides significant gains over the one obtained using OPT-R. The optimal λ (or fraction of throughput obtained) is found to be 0.4681, 0.6043 and 0.6043 with OPT-R, OPT-T and OPT-NC respectively. In this example, the two sinks are almost at the same distance from the sensor nodes. Hence, if a sensor node broadcasts to the sinks, the rate of broadcast will be comparable to the rate of unicasts to each sink. Hence, broadcast is an attractive option in this network topology as we get same rate of multicast (rate of multicast of a sensor node is the minimum of the rates of flows from it to the sinks associated with the multicast) as in the case when we use individual unicasts with less resource consumption. This explains the gains we see using the OPT-T formulation.

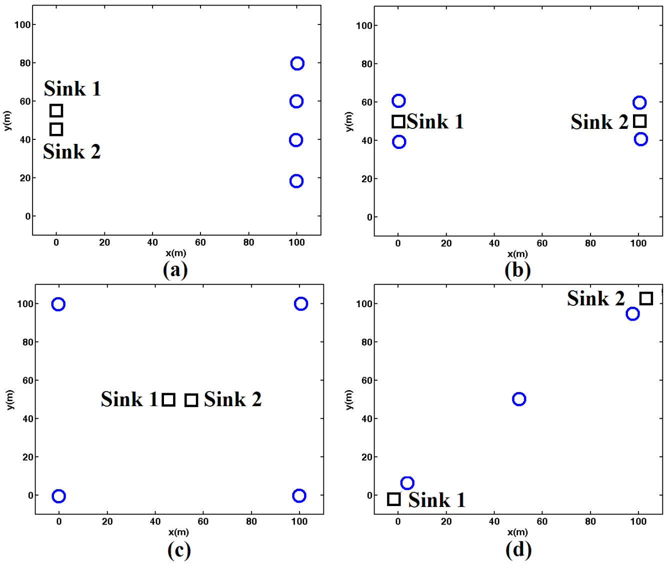

Fig. 3. (a) Example 1: For this network, using OPT-T provides significant gains over using OPT-R (b) Example 2: For this network, solution obtained using OPT-T and OPT-R are comparable (c) Example 3: For this network, solution obtained using OPT-T and OPT-NC are comparable (d) Example 4: For this network, using OPT-NC provides gains over using OPT-T

For the network shown in Fig. 3(b), solutions obtained using OPT-T and OPT-R are comparable. The solutions were found as 0.6268, 0.6269 and 0.6269 with OPT-R, OPT-T and OPT-NC respectively. In this network, we see that the sinks are well separated and further, with respect to any sensor node, there is one sink which is much farther away than the other. In this case the performance of the three approaches is almost identical.

For the network shown in Fig. 3(c), solutions obtained using OPT-T and OPT-NC are both approximately equal to 0.8011. Solution obtained using OPT-R is approximately equal to 0.7419. Due to the symmetry in the topology, no routing of data through an intermediate node is needed and hence only trivial network coding operations occur at the nodes in OPT-NC.

For the network shown in Fig. 3(d) which consists of 3 sensor nodes that multicast to two sinks, solutions obtained using OPT-R, OPT-T and OPT-NC are 0.6134, 0.6475 and 0.6811 respectively. Unlike in the previous examples, here for each sensor node, routing of its data through the other nodes can be beneficial (provided there is adequate power available at the other nodes). This could explain the gain that we observe when using the OPT-NC formulation. We would like to point out that for many examples with small N, we found that the solutions obtained using OPT-T and OPT-NC were comparable. We feel that the gains using network coding should be higher in larger networks.

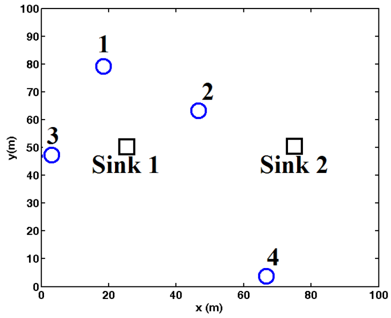

Fig. 4. Layout of the network: Four sensors multicasting to two sinks

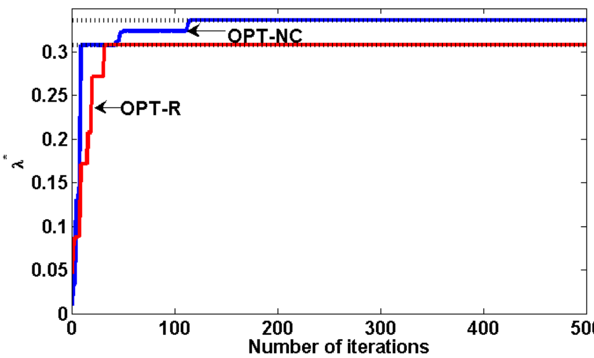

Fig. 5. Performance of ALGO-M: Used to solve OPT-R and OPT-NC

Next, we present results showing the performance of the computationally efficient suboptimal approach that uses ALGO-M. We use ALGO-M to obtain good suboptimal solutions for OPT-R and OPT-NC. Fig. 4 shows the layout of the network consisting of 4 sensor nodes and 2 sinks. All the nodes generate data at the rate of 5000 bps. Also, $P_n^{av} = 10\ mW$ $\forall n \in \mathcal{N}$. The optimal solution for OPT-NC was obtained as 0.3364 whereas the optimal solution for OPT-R was obtained as 0.3076. In Fig. 5, we show the solution obtained by using ALGO-M. The optimal solutions for the two formulations are also shown using dashed lines. We see that using ALGO-M, we get close to the optimal solution quickly (within the first 100 iterations). We observed similar performance for ALGO-M when tested for several different settings involving small networks (small N and small number of destinations) for which we could compute the optimal solution. In the next example, we study the performance of ALGO-M when it is used for OPT-R and OPT-NC for a large network. Fig. 6 shows the layout of the network which consists of 20 nodes and 3 sink nodes. Here, 10 of the sensor nodes have a multicast connection to sinks 1 and 2, and the remaining 10 sensor nodes have a multicast connection to sinks 2 and 3. In Fig. 7, we depict the solutions obtained for OPT-R and OPT-NC using ALGO-M. Note that when we consider the OPT-NC formulation, there are two sessions and we perform intra-session network coding within each session. For this network also, $d_n = 1000\ bps$, $P_n^{av} = 10\ mW$ $\forall n \in \mathcal{N}$. Although, there is no optimal solution to compare against, we do see that most of the increase in the solution obtained using

ALGO-M is during the first 200 iterations (similar to what we observed in the previous example). Also, observe that there is a significant performance improvement when using OPT-NC for this network.

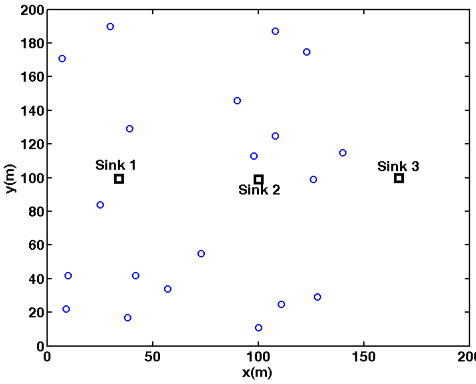

Fig. 6. Layout of the network: 20 sensors, 3 sinks

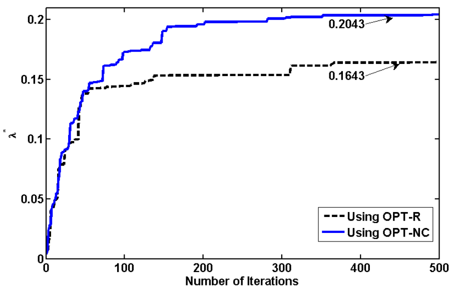

Fig. 7. Performance of ALGO-M: Used to solve OPT-R and OPT-NC

VII. Conclusions and Future Work

We considered the problem of joint power control, scheduling and routing for multicast in energy harvesting sensor networks allowing for multicast of data generated at the sensor nodes to a set of sink nodes. We presented three approaches to solve the problem. We compared the performance of the three approaches using examples. For two of the approaches discussed in this paper, we also considered a computationally efficient suboptimal approach and studied its performance through examples.

Acknowledgment

This work was partially supported by a research grant from Boeing Corporation.

References

[1] T. Ho and D. S. Lun, *Network Coding: An Introduction*, Cambridge University Press, Cambridge, U.K., 2008.

[2] L. Georgiadis, M. Neely, and L. Tassiulas, "Resource allocation and cross-layer control in wireless networks," *Foundations and Trends in Networking*, vol. 1, no. 1, pp. 1-144, 2006.

[3] X. Lin, N. Shroff and R. Srikant, "A tutorial on cross-layer optimization in wireless networks,“ *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 8, pp.1452-1463, Aug. 2006.

[4] L. Lin, N. B. Shroff, and R. Srikant, "Energy-aware routing in sensor networks: A large system approach," *Ad Hoc Networks*, vol. 5, no. 6, pp. 818 -831, 2007.

[5] M. Cao, V. Raghunathan, S. Hanly, V. Sharma, and P. Kumar, "Power control and transmission scheduling for network utility maximization in wireless networks," *46th IEEE Conference on Decision and Control*, pp. 5215-5221, Dec. 2007.

[6] R. Cruz and A. Santhanam, "Optimal routing, link scheduling and power control in multihop wireless networks," *Twenty-Second Annual Joint Conference of the IEEE Computer and Communications Societies INFOCOM*, vol. 1, March-3 April 2003, pp. 702-711.

[7] S. Cui, R. Madan, A. Goldsmith, and S. Lall, "Joint routing, mac, and link layer optimization in sensor networks with energy constraints," *IEEE International Conference on Communications, ICC* , vol. 2, May 2005, pp. 725-729.

[8] M. Johansson and L. Xiao, "Cross-layer optimization of wireless networks using nonlinear column generation," *IEEE Transactions on Wireless Communications*, vol. 5, no. 2, pp. 435-445, Feb. 2006.

[9] A. Das and D. Dutta, "Data acquisition in multiple-sink sensor networks," *SIGMOBILE Mob. Comput. Commun. Rev.*, vol. 9, no. 3, pp. 82-85, 2005.

[10] I. Stojmenovic, "Geocasting with guaranteed delivery in sensor networks," *IEEE Wireless Communications*, vol. 11, no. 6, pp. 29-37, Dec. 2004.

[11] Q. Cao, T. He, and T. Abdelzaher, "Ucast: Unified connectionless multicast for energy efficient content distribution in sensor networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 18, no. 2, pp.240-250, Feb. 2007.

[12] R. Ahlswede, N. Cai, S. yen Robert Li, and R. W. Yeung, "Network information flow," *IEEE Transactions on Information Theory*, vol. 46, pp. 1204-1216, 2000.

[13] R. Koetter and M. Mdard, "An algebraic approach to network coding," *IEEE/ACM Transactions on Networking*, vol. 11, pp. 782-795, 2003.

[14] T. Ho, R. Koetter, M. Medard, D. Karger, and M. Effros, "The benefits of coding over routing in a randomized setting," *IEEE International Symposium on Information Theory*, June-4 July 2003, pp. 442.

[15] J. Yuan, Z. Li, W. Yu, and B. Li, "A cross-layer optimization framework for multihop multicast in wireless mesh networks," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 11, pp. 2092-2103, Nov. 2006.

[16] Y. Wu, P. Chou, Q. Zhang, K. Jain, W. Zhu, and S.-Y. Kung, "Network planning in wireless ad hoc networks: a cross-layer approach," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 1, pp. 136-150, Jan. 2005.

[17] A. Eryilmaz, "Control for inter-session network coding," *in Proc Workshop on Network Coding, Theory and Applications*, 2007.

[18] R. Dougherty, C. Freiling, and K. Zeger, "Insufficiency of linear coding in network information flow," *IEEE Transactions on Information Theory*, vol. 51, no. 8, pp. 2745-2759, Aug. 2005.

[19] *CC2420 Datasheet*, Texas Instruments. [Online]. Available: http://focus.ti.com/lit/ds/symlink/cc2420.pdf

[20] *ADF7020 Datasheet*, Analog Devices. [Online]. Available: http://www.analog.com/static/imported-files/data_sheets/ADF7020.pdf

[21] *ATA542x Datasheet*, ATMEL. [Online]. Available: http://www.atmel.com/dyn/resources/prod_documents/doc4841.pdf

[22] V. Sharma, U. Mukherji, V. Joseph, and S. Gupta, "Optimal energy management policies for energy harvesting sensor nodes," 2008, Submitted. [Online] Available: http://arxiv.org/pdf/0809.3908v1

[23] V. Joseph, V. Sharma, and U. Mukherji, "Optimal Sleep-Wake policies for an energy harvesting sensor node," in *ICC 2009 Ad Hoc and Sensor Networking Symposium*, June 2009.

[24] V. Joseph, V. Sharma and U. Mukherji, "Joint power control, scheduling and routing for multihop energy harvesting sensor networks," 2009, Submitted.

[25] R. K. Ahuja, T. L. Magnanti, and J. B. Orlin, *Network Flows: Theory, Algorithms, and Applications*, Prentice Hall, 1993.

[26] V. Sharma, U. Mukherji, and V. Joseph, "Efficient energy management policies for networks with energy harvesting sensor nodes," in *Allerton Conference on Communication, Control, and Computing*, 2008, Invited Paper.

[27] M. Kodialam, T. V. Lakshman, and S. Sengupta, "Online multicast routing with bandwidth guarantees: a new approach using multicast network flow," *IEEE/ACM Trans. Netw.*, vol. 11, no. 4, pp. 676-686, 2003.