Modeling and Optimization of Transmission Schemes in Energy-Constrained Wireless Sensor Networks

Ritesh Madan, *Member, IEEE*, Shuguang Cui, *Member, IEEE*, Sanjay Lall, *Senior Member, IEEE*, and Andrea J. Goldsmith, *Fellow, IEEE*

***Abstract*—We consider a wireless sensor network with energy constraints. We model the energy consumption in the transmitter circuit along with that for data transmission. We model the bottom three layers of the traditional networking stack—the link layer, the medium access control (MAC) layer, and the routing layer. Using these models, we consider the optimization of transmission schemes to maximize the network lifetime. We first consider the optimization of a single layer at a time, while keeping the other layers fixed. We make certain simplifying assumptions to decouple the layers and formulate optimization problems to compute a strategy that maximizes the network lifetime. We then extend this approach to cross-layer optimization of time division multiple access (TDMA) wireless sensor networks. In this case, we construct optimization problems to compute the optimal transmission schemes to an arbitrary degree of accuracy and efficiently. We then consider networks with interference, and propose methods to compute approximate solutions to the resulting optimization problems. We give numerical examples that illustrate the computational approaches as well as the benefits of cross-layer design in wireless sensor networks.**

***Index Terms*—Cross-layer design, energy efficiency, network lifetime, optimization, wireless sensor networks.**

I. Introduction

THERE are a number of fundamental optimization problems that arise when designing or controlling a wireless sensor network. In this paper, the objective is to construct these problems and the associated constraints from the essential characteristics of the wireless network model. Specifically, we focus on those problems that arise due to energy constraints. We formulate a general system model and consider the optimization of different network objectives. Efficient computational methods are obtained to solve these problems exactly or in an approximate manner. Some of the optimization problems formulated in this paper can be solved using partially or fully distributed algorithms. Hence, we can implement these algorithms as protocols to implicitly solve the relevant optimization problems in wireless networks in a distributed manner. In very special cases, analytical solutions can also be obtained.

Over the past few years, optimization techniques have been used to solve many problems arising in wireless as well as wireline networks. The problem of flow control in the Internet was formulated as a convex optimization problem in [1] and [2]. In these papers, dual decomposition methods were used to derive distributed congestion control algorithms. In wireless networks, achievable rate combinations were computed in [3]. Cross-layer optimizations to maximize throughput have been considered in [4]–[9].

In this paper, we consider the optimization of various layers for energy-constrained wireless networks. A good overview of the design challenges and the importance of cross-layer design in energy-constrained wireless networks was given in [10]. Recent results show that the processing/circuit energy consumption can have a great impact on the capacity as well as the optimum transmission schemes for medium access control (MAC) channels [11], [12]. For wireless sensor networks, we may not always need to operate on the boundary of the achievable rates region. Hence, we have a choice among various transmission strategies (routing, power control, scheduling) that we can exploit to increase the lifetimes of such networks. An overview of the synergy between the various layers in a wireless network was given in [13].

We consider a wireless sensor network of nodes distributed in a certain region (see, for example, [14], [15]). We assume that each node has a limited energy supply and generates information that needs to be communicated to a sink node. Also, each node can vary its transmission power, duty cycle, and modulation scheme. We will use transmission scheme/strategy to refer to the data rates, transmission powers, and link schedule for a network. For energy-constrained wireless networks, we can increase the network lifetime by using transmission schemes that have the following characteristics.

1) *Multihop routing*: In wireless environments, the received power typically falls off as the mth power of distance, with $2 \leq m \leq 6$. Hence, we can conserve transmission power by using multihop routing in long-range applications [16], [17].
2) *Load Balancing*: If a node is on the routes of many source-destination pairs, it will run out of energy very quickly.

Manuscript received January 7, 2005; revised January 4, 2006; approved by IEEE/ACM Transactions on Networking Editor D. Towsley. This work was supported by National Semiconductor, Toyota Corporation, Robert Bosch Corporation, the Sequoia Capital Stanford Graduate Fellowship, the Stanford URI Architecture for Secure and Robust Distributed Infrastructures (AFOSR DoD award number 49620-01-1-0365), and the University of Arizona Foundation.

R. Madan was with the Department of Electrical Engineering, Stanford University, Stanford, CA 94305 USA. He is now with Qualcomm–Flarion Technologies, Bedminster, NJ 07921 USA (e-mail: rkmadan@stanfordalumni.org).

S. Cui was with the Department of Electrical and Computer Engineering, University of Arizona, Tucson, AZ 85721 USA. He is now with the Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX 77843 USA (e-mail: cui@ece.tamu.edu).

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

A. J. Goldsmith is with the Wireless System Laboratory, Department of Electrical Engineering, Stanford University, Stanford, CA 94305 USA (e-mail: andrea@wsl.stanford.edu).

Digital Object Identifier 10.1109/TNET.2007.897945

Hence, load balancing is necessary to avoid the creation of hot spots where some nodes die out quickly and cause the network to fail [18].

3) *Interference mitigation*: Links that strongly interfere with each other should be scheduled at different times to decrease the average power consumption on these links [19].
4) *Scheduling*: Given a fixed number of bits to transmit over a link, we can reduce the average transmission energy over the link by scheduling it for as long as possible. Hence, weakly interfering links should be scheduled together so that each link has a longer duration of time to transmit the same amount of data.

We use the standard approach of viewing the wireless system as consisting of layers. The lowest layer is the physical or the link layer, where the objective is to choose the transmission power and rate to minimize the energy consumption to transmit at a given average rate. The next layer is the MAC layer, where the objective is to select a link schedule. The uppermost layer is the routing layer. We first consider the optimization of one layer at a time, keeping the other layers fixed. We then extend our models to compute cross-layer schemes that are energy/lifetime optimal. The problems considered in this paper are either convex optimization problems, or can be approximated by convex optimization problems.

We note that parts of the work in this paper appear in [11] and [20]–[23]. These papers considered different resource allocation problems in energy-constrained networks. In particular, optimization of the link layer and the routing layer were considered in [11] and [21], respectively. In this paper, the goal is to connect these different special instances of network optimization problems and put them in a common framework. Thus, in this paper, we provide a network-wide view of seemingly disparate results. Also, we neglect the energy consumption in the receiver circuit to simplify notation and make the presentation clearer. The extensions to modeling receiver circuit energy are quite straightforward; we refer the reader to [22] and [24].

A. Outline

The rest of the paper is organized as follows. The system model for the wireless sensor network is described in Section II. In Section III, we model one layer at a time and formulate optimization problems to optimize each layer to maximize the network lifetime. In Section IV, we extend the models to cross-layer optimizations in TDMA-based wireless sensor networks. Section V describes the methodology for cross-layer optimization in interference-limited sensor networks.

II. System Model

A. Network Graph

We consider a static wireless sensor network, modeled by a directed graph $\mathcal{G} = (V, L)$, where V is the set of wireless nodes and L is the set of directed links. Each link l is from a transmitter (initial) node i to a receiver (destination) node j. A link exists from node i to node j if the received power at node j, when node i transmits at maximum transmission power and no interference is present, is greater than a predefined threshold. Note that, in general, we can consider a very low threshold; this would correspond to a fully connected network. However, this introduces many links in the network graph and hence, many optimization variables, making the cross-layer optimization problem more complex. By introducing a reasonable threshold, we do not consider links that are very weak—these links will not only consume a lot of power, but also cause a lot of interference to the other links. For an information theoretic justification of this approach, we refer the reader to [25].

We define the incidence matrix of the graph $\mathcal{G} \in \mathbb{R}^{|V|\times|L|}$ as follows:

$$A_{v,l} = \begin{cases} 1, & \text{if } v \text{ is the transmitter of link } l \\ -1, & \text{if } v \text{ is the receiver of link } l \\ 0, & \text{otherwise.} \end{cases}$$

Let us write $A = A^+ - A^-$, such that $A^+_{v,l}, A^-_{v,l} = 0$ if $A_{v,l} = 0$, and A^+, A^- have only 0 and 1 entries.

We assume that each node $i \in V$ generates information at rate S_i and has total battery energy E_i. The goal of the network is to communicate the information generated by each node to a sink node. For the sink node, we take $S_{\text{sink}} = -\sum_{i\in V, i\neq \text{sink}} S_i$. Thus, we consider a single commodity flow to simplify notation. However, the methods in this paper apply to the multicommodity flow scenario as well.

B. Slotted Time

We consider time to be slotted into discrete time slots numbered 0, 1, 2, . . ., where each slot is of equal length h. Let $p_l(m)$ denote the transmission power used by the transmitter of link l to send data over link l and time slot m. Similarly, let $r_l(m)$ denote the corresponding rate. We will restrict $p_l(m)$ (and $r_l(m)$) to be periodic functions of the slot index m, with period M. Thus, the network time shares between M transmission modes in a periodic fashion. We will use the term *schedule frame* to refer to a period of M time slots. Let x_l denote the average flow rate over link l. Thus, we have $x_l = (1/M)\sum_{m=1}^{M} r_l(m)$.

C. Physical Layer

Associated with the graph $\mathcal{G}$ is the link gain matrix $G \in \mathbb{R}^{|L|\times|L|}$. G_{lk}, $l \neq k$, denotes the power gain from the transmitter of link k to the receiver of link l. Note that G_{lk} depends upon the path loss from the transmitter of link k to the receiver of link l. It is also a function of the links l and k—it depends on the antenna gains, and the correlation between the code sequences if code division multiple access (CDMA) is used. The channel over each link is assumed to be an additive white Gaussian noise (AWGN) channel, with noise power spectral density N_0 over the bandwidth of operation B. The total interference and noise power at the receiver of link l during slot m is given by $I_l(m) = \sum_{k\neq l} G_{lk}p_k(m) + N_0 B$. We assume that the interference is Gaussian. From an information theoretic viewpoint, this is a conservative assumption since Gaussian noise is the "worst noise" [26]. The signal to interference and noise ratio (SINR) at the receiver of link l during slot m is defined to be

$$\gamma_l(m) = \frac{G_{ll}p_l(m)}{\sum_{k\neq l} G_{lk}p_k(m) + N_0 B}. \quad (1)$$

We assume that the bit error rate (BER) is kept constant and same for all links. Then the rate that can be supported over link l with SINR $\gamma_l(m)$ satisfies

$$r_l(m) \leq B \log\left(1 + K\gamma_l(m)\right) \quad (2)$$

where $K = -1.5(\ln(5\text{BER}))$. This is a standard model for modulation schemes such as M-quadrature amplitude modulation (MQAM) with constellation size greater than or equal to 4 [27]. Note that we have relaxed the constraint that the constellation size is an integer; thus $r_l(m)$ can take all non-negative real values. This reduces the computational complexity from that of solving integer programs to that of solving convex optimization problems in many cases. Real values of constellation sizes are technologically feasible but need high complexity hardware [28]. If only integer values of constellation size can be used, then our methods provide approximate solutions.

Another approach to model interference is to use a combinatorial interference model, for example, protocol model [29]. The scheduling constraints can then be modeled by a conflict graph [6], [30]. We note that many combinatorial models like the protocol model do not model all the interference in the network. For throughput optimization algorithms for such models we refer the reader to [6].

D. Power Consumption Model

For each link, we model the transmission and circuit energy consumption at the transmitter. We do not consider the energy consumption for reception; the methods described in this paper can be easily extended to model the energy consumption at the receiver as well. The nature of the problems or the computational approaches to solve them do not change when we include the receiver circuit energy as well; only the optimal operating point of the network changes. Consider a slot m and link l with the transmission rate over link l given by $r_l(m)$. Assume that the transmission powers over all other links are fixed. From (1) and (2), we can see that the minimum transmission power over the link to transmit at rate $r_l(m)$ is given by $p_l(m) = I_l(m)(2^{r_l(m)/B} - 1)/(KG_{ll})$. Hence, the transmission energy needed to transmit at rate $r_l(m)$ for one slot is $hp_l(m)$. Then the total energy consumption in slot m at the transmitter of link l can be approximated as [11, sec. 5]

$$h\left((1+\alpha)p_l(m) + \beta\right) = \frac{(1+\alpha)I_l(m)h}{KG_{ll}}\left(2^{\frac{r_l(m)}{B}} - 1\right) + h\beta \quad (3)$$

where α and β are system constants. The physical interpretation for these constants is as follows:

$\alpha =$ ratio of overhead power consumption in the power amplifier to the transmission power;

$\beta =$ energy consumption in the transmitter circuit, excluding that in the power amplifier.

We denote the vector of transmission powers over the links in slot m by $p(m)$. Let us define the function $\mathbf{1} : \mathbb{R}^{|V|} \mapsto \{0,1\}^{|V|}$ as follows: $(\mathbf{1}(x))_i = 1$ if $x_i > 0$, and 0 otherwise. Then $\mathbf{1}(A^+p(m))$ gives the set of nodes active in slot m. The average power consumption at node i during a schedule frame of M slots is given by

$$\frac{1}{M}\sum_{m=1}^{M}\left(\sum_{l\in\mathcal{O}(i)}(1+\alpha)p_l(m) + \left(\mathbf{1}\left(A^+p(m)\right)\right)_i\beta\right) \quad (4)$$

where $\mathcal{O}(i)$ denotes the set of outgoing links at node i. Here, we have neglected the overhead energy consumed each time a node turns on to transmit in a schedule frame [11]. Usually, this value is small and hence can be neglected if a node transmits for a time much longer than the time it takes to turn on.

E. Flow and Energy Conservation

The total outgoing flow at a node should be equal to the sum of the total incoming flow and the flow generated at the node itself. Thus, we have the flow conservation constraint at node i given by

$$\sum_{l\in\mathcal{O}(i)} x_l - \sum_{l\in\mathcal{I}(i)} x_l = S_i \quad (5)$$

where $\mathcal{O}(i)$ and $\mathcal{I}(i)$ denote the set of outgoing and incoming links, respectively, at node i. Note that the flow conservation equations are satisfied at the frame timescale. This necessitates the use of buffers at each node.

If a node is alive for time T_i, the energy dissipated by it in time T_i should be less than or equal to the initial energy E_i. Hence, we have the following energy conservation constraint at each node i.

$$\frac{T_i}{M}\sum_{m=1}^{M}\left(\sum_{l\in\mathcal{O}(i)}(1+\alpha)p_l(m) + \left(\mathbf{1}\left(A^+p(m)\right)\right)_i\beta\right) \leq E_i. \quad (6)$$

F. Objectives

The optimization objectives for networks with finite energy may differ depending on the circumstances and the data model. We consider three different objective functions in this paper.

1) Total Power Consumption: For a scenario, where the source rates change randomly with time, minimizing the total power consumption at all nodes will save power, on an average, at each node. However, we assume that the rate at which the flow changes is much slower than the rate of convergence of the network optimization algorithms. A weighted sum of power consumptions at nodes can be used to give more importance to nodes close to the sink, as they will, on an average, support higher data rates.

2) Minimum Node Lifetime: Let T_i denote the time at which node i runs out of energy. Then the minimum node lifetime is given by $T_{\min} = \min_{i\in V} T_i$. This is a good metric when all the nodes in the network are critical to network operation. This definition makes the analysis tractable for many different scenarios. Another interpretation of the optimization problem that maximizes the time at which the first node dies is that it minimizes the maximum ratio of average power consumption to initial energy among all nodes—it thus balances the data flow in the network such that no node incurs a very high power consumption.

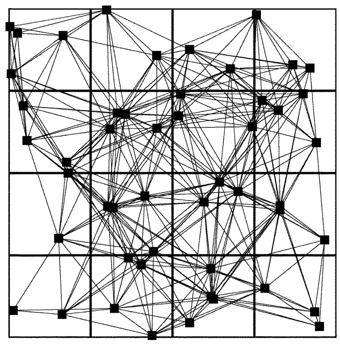


Fig. 1. Network lifetime for a gridded network.

Hence, for a large network with many redundant nodes, a suboptimal but reasonable approach would be to use the transmission scheme that maximizes the time at which the first node dies, and recompute the transmission scheme once the topology has changed significantly after many nodes run out of energy. Until then, the same transmission scheme can be used with greedy readjustment of flows.

3) Concave Functions of Node Lifetimes: In many scenarios, we can consider an operational definition of network lifetime as a concave function of node lifetimes. In particular, we define $T_{\rm net} = f(T_1, \ldots, T_{|V|})$, where $f : \mathbb{R}^{|V|} \rightarrow \mathbb{R}$ is a concave function in the vector of node lifetimes, and T_i is now defined to be the *total time for which flow of node i is supported*. Note that minimum node lifetime is a special case of this general definition. For example, consider a sensor network of randomly distributed nodes in a certain region, as shown in Fig. 1. The deployment region of the network has been divided into a grid of 16 regions. Two nodes are connected by an edge if they can directly communicate with each other using a transmission power less than the maximum allowable transmission power at each node. Let $\mathcal{G}_i$ denote the set of nodes in region i. We assume that the network is functional as long as at least one node in each region can sense data. This would correspond to a network lifetime given by the concave function of node lifetimes, $T_{\rm net} = \min_{i=1,\ldots,16}[\sum_{j\in\mathcal{G}_i} T_j]$. Note that the time for which at least one node is active in region i is $\sum_{j\in\mathcal{G}_i} T_j$. This is because all nodes in a particular region can take turns to transmit data. Now lets further consider a network to be partially functional as long as some regions have at least one alive node. Moreover, sensing in some regions can be more important than that in others. In such a scenario, we can consider the network lifetime to be a weighted linear sum of the lifetimes of the individual regions, i.e., $T_{\rm net} = \sum_{i=1}^{16} k_i \sum_{j\in\mathcal{G}_i} T_j$.

G. System Parameters

The system parameters that we use for the computational results in this paper are listed in Table I. For the rest of the paper, to simplify notation, we take the bandwidth $B = 1$. The values of transmission rates can then be interpreted as being normalized with respect to the bandwidth. Also, we take $N_0 B = N_0 = 0.5$ mW. In addition, we take $G_{lk} = (\kappa)/(d_{lk}^4)$, where κ is a constant given in Table I and d_{lk} is the distance between the receiver of link l and transmitter of link k. This corresponds to a deterministic path loss model where the receiver power falls off as the fourth power of the distance from the transmitter.

TABLE I
SYSTEM PARAMETERS

Parameter	Value
α	3
$N_0 B$	0.5 mW
K	0.15
β	100 mW
κ	$10^{-3} m^4$

III. Layer-Wise Optimization

We first consider optimizing each layer separately to maximize the network lifetime. The layers that we consider are the link and routing layers. For the routing layer, we obtain *distributed algorithms* to maximize network lifetime. Also, in this section we consider transmission schemes without interference—we assume TDMA for the MAC layer. Thus, we allow only one link in the network to transmit at any given time. This is a good paradigm in small networks in which links interfere strongly, i.e. the terms G_{lk}, $l \neq k$, are comparable to G_{ll}. For TDMA transmission schemes, we have $I_l(m) = N_0$ if link l is active in slot m.

From (3), we see that the power consumption to transmit at rate r over a link is of the form $f(r) = k_1(e^{r/k_2} - 1) + k_3$, where $k_1,\ k_2 > 0$. The function $f(r)$ is convex for $r \geq 0$. Hence, time sharing between two transmission modes with different rates consumes more power than transmission at a uniform rate equal to the average rate. Thus, if a link is active over m_l slots, the energy efficient scheme is to transmit at a uniform rate during the m_l slots. In this section, we will denote by r_l and p_l the transmission rate and power, respectively, over link l when it is active. Thus, we have suppressed the dependence of the transmission rate and power on the slot index m, since the optimal scheme for a link is to transmit at the same rate and hence same power over all active time slots.

The average power consumption at node i, given by (4), can be written as

$$\frac{1}{M}\sum_{m=1}^{M}\left(\sum_{l\in\mathcal{O}(i)}(1+\alpha)p_l(m) + \left(\mathbf{1}\left(A^{+}p(m)\right)\right)_i\beta\right) = \frac{1}{M}\sum_{l\in\mathcal{O}(i)} m_l\left((1+\alpha)p_l+\beta\right) \quad (7)$$

where m_l is the number of slots per frame during which link l transmits at rate r_l using transmission power

$$p_l = \frac{N_0}{KG_{ll}}(2^{r_l} - 1). \quad (8)$$

A. Link Layer

We fix the MAC and the routing layer, i.e., the average flow rate x_l and the number of time slots m_l allocated to each link l in every frame are fixed. We assume that the MAC layer mitigates interference between active links. Hence, we neglect the

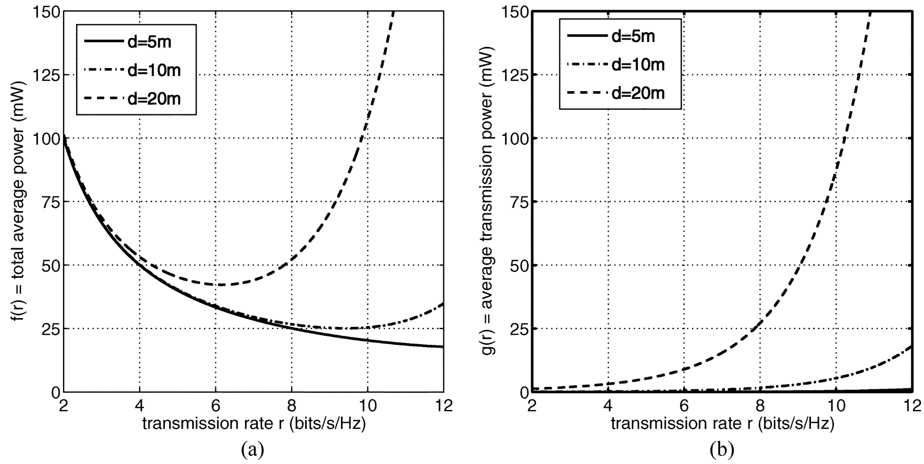


Fig. 2. Power consumption on a single link. (a) Total average power. (b) Average transmission power.

cross-link interference terms and approximate $I_l(m) = N_0$ if link l is active in slot m. The goal is to select the transmission rate for each link to minimize the power consumption on the links and hence maximize the network lifetime. The only constraint is that each link l has to transmit at an average rate x_l. This can be written as $r_l \geq (Mx_l)/(m_l)$. Also, in Section II-B, we assumed that the transmission power $p_l(m)$ and rate $r_l(m)$ remain constant over a single slot of length h. Hence, the amount of time that each link is active in a given frame should be a multiple of h, i.e.,

$$\frac{x_l M}{r_l} \in \{0, \dots, m_l\}. \tag{9}$$

The objective is to minimize the total power consumption in the network. This can be written as the following optimization problem with variables $r_l, \forall l \in L$.

$$\begin{aligned} \text{minimize} \quad & \frac{1}{M}\left(\sum_{l\in L} \frac{(1+\alpha)N_0 x_l}{KG_{ll} r_l}(2^{r_l}-1) + \frac{x_l}{r_l}\beta\right), \\ \text{subject to} \quad & r_l \geq \frac{Mx_l}{m_l}, \\ & \frac{x_l M}{r_l} \in \{0,\dots,m_l\}, \qquad \forall l \in L. \end{aligned} \tag{10}$$

Note that the objective function and the constraints are separable in the variables r_l's. Hence, we can solve the optimization problem for each link separately to minimize the transmission power over each link. Moreover, for each link we have an optimization problem in one variable r_l, with r_l being allowed to take only finitely many values. This problem can be solved easily using bisection search since the objective function is convex.

We note that similar observations were made in previous works, for example [11] and [31]; we have included link-layer optimization here for completeness.

1) Example: Here, we illustrate the trade-off between reducing transmission power by transmitting at a low rate for a longer time, and reducing circuit power by transmitting in bursts. Consider two nodes separated by a distance d, with one node transmitting data to the other node at an average rate of $x = 2$ bits/s/Hz. If r is the rate of transmission, the transmitter turns on for a fraction x/r of time. The total average power consumption is given by [see (3)]

$$f(r) = \frac{x}{r}\left(\frac{(1+\alpha)N_0}{KG}(2^r - 1) + \beta\right)$$

where $G = \kappa/d^4$ is the link gain. Also, the average transmission power used by the link is given by $g(r) = (xN_0(1+\alpha)(2^r-1))/(rKG)$. For $r \geq 2$, $g(r)$ is minimized at $r = 2$. Thus, to minimize the average transmission power, we should choose the lowest possible data transmission rate. Fig. 2 shows the total average power consumption and the average transmission power as a function of the transmission rate, for different distances d. We can see that as the distance between the two nodes increases, the optimal transmission rate decreases. This is because as the distance increases the transmission power increases but the circuit power remains constant. Hence, lowering of transmission power by reducing the transmission rate becomes increasingly important to decrease the average power consumption. Thus, the optimal transmission scheme for a small distance is more bursty than that for a large distance.

B. Routing Layer

We now fix the link and the MAC layers. Thus, the transmission rate r_l and the maximum number of slots m_l for which each link can transmit during each frame, is fixed. We have the following constraint on the maximum flow over a link l : $x_l \leq x_l^{\max} = r_l m_l/M$. We also have the constraint given in (9). However, in this section, we will assume that each frame is divided into a large enough number of slots such that we can drop this constraint while computing the optimal routing flow. We can thus approximate the number of slots for which link l transmits data by Mx_l/r_l. This is a good approximation if $\lceil Mx_l/r_l \rceil - Mx_l/r_l$ is small compared to M, which would be

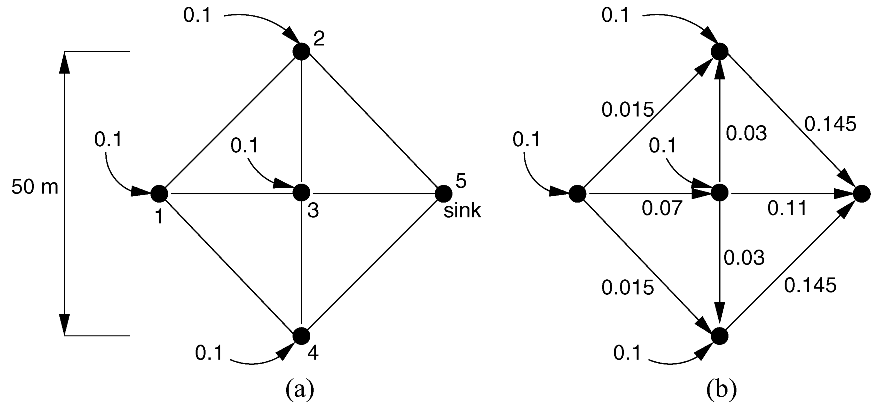


Fig. 3. Load balancing to maximize the network lifetime. (a) Topology. (b) Optimal flow.

the case if each frame is divided into a large number of slots. The average power consumption at node i follows from (7):

$$\frac{1}{M}\sum_{l\in\mathcal{O}(i)} m_l\left((1+\alpha)p_l+\beta\right)=\sum_{l\in\mathcal{O}(i)} c_l x_l$$

where

$$c_l=\frac{(1+\alpha)p_l+\beta}{r_l}. \quad (11)$$

Since r_l and hence p_l are fixed, c_l is a constant for a given link l. Also, since the total initial energy at node i is E_i, the lifetime of node i under flow $\{x_l\}$ is given by $T_i(x)=E_i/(\sum_{l\in\mathcal{O}(i)} c_l x_l)$.

1) Maximizing Minimum Node Lifetime: Here, $T_{\rm net}(x)=\min_{i\in V} T_i(x)$. The goal is to select a flow $\{x_l\}$ such that it satisfies the maximum rate constraint and the flow conservation constraints in (5), and maximizes the network lifetime $T_{\rm net}(x)$. This can be written as the following optimization problem in variables x_l, $\forall l\in L$:

$$\begin{aligned}&\text{maximize} && T_{\rm net}(x),\\ &\text{subject to} && Ax=S, \quad 0\preceq x\preceq x^{\max}.\end{aligned} \quad (12)$$

We note that this problem formulation was first considered in [18]. Using the definition of $T_{\rm net}$, the above problem can be rewritten as follows:

$$\begin{aligned}&\text{maximize} && T,\\ &\text{subject to} && Ax=S, \quad 0\preceq x\preceq x^{\max}, \quad TA^{+}\text{diag}(c)x\preceq E\end{aligned}$$

where $\text{diag}(c)$ denotes a diagonal matrix with the ith diagonal value equal to c_i. The variables are T, x_l, $\forall l\in L$. The last constraint implies that no node runs out of energy as long as the network is alive. The value of T obtained by solving this problem gives the optimal value of the network lifetime. Using a change of variable $q=1/T$, we obtain the following linear programming problem:

$$\begin{aligned}&\text{minimize} && q,\\ &\text{subject to} && Ax=S, \quad 0\preceq x\preceq x^{\max}, \quad A^{+}\text{diag}(c)x\preceq qE.\end{aligned} \quad (13)$$

This problem can be solved efficiently in a centralized manner to compute a maximum lifetime routing flow.

If we assume that there is a flow x such that $0\prec x\prec x^{\max}$, and it satisfies the flow conservation equations, we can choose a value of q large enough to satisfy the energy conservation constraints with strict inequality. Then the Slater's condition for constraint qualification is satisfied (see, for example, [32]) and hence, strong duality holds. Thus, we can solve the primal problem via the dual. It can be seen that the dual problem is separable in the variables x_l. Hence, we can exploit dual decomposition methods to obtain partially distributed subgradient-based algorithms to solve problem (13). Adding a regularization term $\epsilon\sum_{i\in V}\sum_{j\in N_i} x_{ij}^2$ for small ϵ (for which the suboptimality of the solution can be bounded as in [21]), the sequence of primal iterates is given by

$$\begin{aligned} q^{(k)} &= \arg\min_{0\le q\le Q}\left(q^2-q\sum_{i\in V}\lambda_i^{(k)}E_i\right)\\ x_{ij}^{(k)} &= \arg\min_{0\le x_{ij}\le x_{ij}^{\max}}\left(\epsilon x_{ij}^2+x_{ij}\left(\lambda_i^{(k)}c_{ij}+\nu_i^{(k)}-\nu_j^{(k)}\right)\right)\end{aligned} \quad (14)$$

where λ_i's and ν_j's are dual variables. The subgradient of the dual function $-g$ at $(\lambda^{(k)},\nu^{(k)})$ is given by

$$\begin{aligned} h_i^{(k)} &= q^{(k)}E_i-\sum_{j\in N_i} c_{ij}x_{ij}^{(k)}\\ f_i^{(k)} &= S_i-\sum_{j\in N_i}\left(x_{ij}^{(k)}-x_{ji}^{(k)}\right).\end{aligned} \quad (15)$$

Since the dual function is differentiable, the sequence of primal iterates converges to a solution of the regularized primal problem. For details, see [21]. Also, note that similar methods can be used to obtain fully distributed algorithms as well.

We note that distributed algorithms to compute a routing scheme to maximize the minimum node lifetime were also proposed in [18], [33], and [34]. However, these algorithms do not readily extend to more general definitions of network lifetime. As we will show later in this section, the subgradient algorithm is applicable to much more general scenarios.

2) Example: We now illustrate the importance of load balancing in maximizing the network lifetime. We consider a simple topology shown in Fig. 3(a). Nodes 1, 2, 3, and 4 are source nodes generating data at an average rate of 0.1 bits/s/Hz. There are 16 links in the network, each pair of links in opposite directions is shown by a single edge in the figure. We assume that the transmission rate over each link is 8 bits/s/Hz when active. Thus, we model a scenario where the network supports low data rates compared to the achievable rates over the network. From (8) and (11), we have

$$c_l = \frac{(1+\alpha)d_{ll}^4 N_0}{8K\kappa}(2^8-1) + \frac{\beta}{8}$$

where d_{ll} is the distance between the transmitter and the receiver of link l. We assume that each source node has the same amount of initial battery energy. The optimal flow that maximizes the network lifetime is shown in Fig. 3(b). The flow direction and value is indicated for each link. We see that node 1 distributes its flow over three different paths instead of sending all its flow along the minimum cost path $1 \to 3, 3 \to 5$. This helps to conserve energy at node 3 and hence increase the network lifetime, which is the time at which the first node in the network runs out of energy.

3) Maximizing Concave Function of Node Lifetimes: Until now we considered the network lifetime to be the 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 may be more important than others. We now 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. In addition, we will drop the constraint on the maximum rate over a link. We will assume that the network operates in a low-data-rate regime, and hence we can easily re-allocate m_l such that $x_l \le (m_l r_l)/M$.

Each node $k \in V$ generates data at rate S_k. Let x_l^k be the rate at which link l transmits data, originally generated by node k. The flow conservation equations at node i for the flow originating at node k can be written as follows:

$$\sum_{l \in \mathcal{O}(i)} x_l^k - \sum_{l \in \mathcal{I}(i)} x_l^k = (\delta_k)_i S_k \qquad \forall i, k \in V \tag{16}$$

where $\delta_k \in \mathbb{R}^{|V|}$ is such that $(\delta_k)_i = 1$ if $i = k$, and 0 otherwise. 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 over link l is $\sum_{k \in V} x_l^k T_k$. The energy constraint at each node is given by

$$\sum_{l \in \mathcal{O}(i)} c_l \sum_{k \in V} x_l^k T_k \le E_i \qquad \forall i \in V$$

where c_l is defined in (11). We consider a generic definition of network lifetime given by a concave function of the node lifetimes. In particular, we define $T_{\rm net} = f(T_1, \ldots, T_{|V|})$, 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 $T_{\rm net} = \min(T_1, \ldots, T_{|V|})$.

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

$$\begin{array}{ll} \text{maximize} & f\left(T_1, \ldots, T_{|V|}\right), \\ \text{subject to} & Ax^k = S_k \delta_k, \qquad x^k \succeq 0, \quad \forall k \in V, \\ & \sum_{k \in V} T_k A^+ \mathrm{diag}(c) x^k \preceq E. \end{array}$$

We apply a change of variables $y_l^k = x_l^k T_k$. We can interpret y_l^k as the total number of bits generated by node k, transmitted over link l. We can rewrite the above problem as the following convex optimization problem:

$$\begin{array}{ll} \text{maximize} & f\left(T_1, \ldots, T_{|V|}\right), \\ \text{subject to} & Ay^k = T_k S_k \delta_k, \qquad y^k \succeq 0, \quad \forall k \in V, \\ & \sum_{k \in V} A^+ \mathrm{diag}(c) y^k \preceq E. \end{array} \tag{17}$$

Then, we can interpret the first 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

$$\begin{array}{ll} \text{maximize} & f\left(T_1, \ldots, T_{|V|}\right), \\ \text{subject to} & Az = \mathrm{diag}(S)T, \qquad z \succeq 0, \quad A^+ \mathrm{diag}(c) z \preceq E \end{array} \tag{18}$$

where $z_l = \sum_{k \in V} y_l^k$ is the total number of bits transmitted over link l. From the solution z_l of this problem, we can easily obtain an optimal flow y_l^k, $k \in V$, which is a solution to problem (17).

IV. TDMA-Based Wireless Sensor Networks

The general cross-layer problem of optimizing the physical, medium access control, and routing layers to minimize network energy consumption is complex and hard to solve. Heuristics to compute approximate solutions and high-complexity algorithms to compute exact solutions have been proposed in the literature. A representative set of such results include [13], [19], and [35]. In this section, we restrict to networks with interference-free link scheduling and practical MQAM link transmission schemes. We show that the cross-layer optimization problem can be closely approximated by convex optimization problems that can be efficiently solved using standard interior point methods. We note that interference-free scheduling is suboptimal but still performs well in small scale wireless networks where the links strongly interfere with each other. This is the most general cross-layer optimization problem that has been solved efficiently; no efficient computational methods are known to solve the problem for more general scheduling schemes. While we do not discuss in detail, the optimization problems in this section can be solved in a decentralized manner using the techniques in [36]. Also, we note that similar optimization problems for throughput maximization with more restrictive system models were considered in [8] and [37].

A. Power Consumption

Consider the case where the average flow rate over link l is x_l. We consider the simple case where the MAC is restricted to TDMA. Then if a link l is allocated m_l slots, the rate at which it transmits data is

$$r_l = \frac{Mx_l}{m_l}. \tag{19}$$

Substituting the above relation in (7) and (8), the average power consumption at node i is

$$\begin{aligned} \frac{1}{M} \sum_{l \in \mathcal{O}(i)} & m_l \left((1+\alpha)p_l, +\beta\right) \\ &= \frac{1}{M} \sum_{l \in \mathcal{O}(i)} m_l \left((1+\alpha)\frac{N_0}{KG_{ll}}(2^{r_l}-1)+\beta\right), \\ &= \frac{1}{M} \sum_{l \in \mathcal{O}(i)} m_l \left((1+\alpha)\frac{N_0}{KG_{ll}}\left(2^{\frac{Mx_l}{m_l}}-1\right)+\beta\right), \\ &= \sum_{l \in \mathcal{O}(i)} g_l(x_l, m_l) \end{aligned} \tag{20}$$

where $g_l(x_l, m_l) = a_l m_l (2^M x_l / m_l + b_l)$ gives the average power consumption over link l as a function of x_l and m_l. Also, a_l and b_l are suitably defined constants for link l. The function $f(x,y) = ye^{x/y}$, defined for $y > 0$, is the perspective of the exponential function, and so is convex in x and y (see, e.g., [32, sec. 3.2.6]). The function g_l is obtained from f by an affine composition, and the addition of a linear term, and so is convex.

B. Cross-Layer Lifetime Maximization

As in Section III-B, we can write the lifetime of node i as $T_i = E_i / (\sum_{l \in \mathcal{O}(i)} g_l(x_l, m_l))$. We now consider the joint optimization of the link, MAC, and routing layers. Thus, the variables are r_l, x_l, and m_l for each link $l \in L$. Note that the values of two of these variables automatically determine the third one. For example, if the average flow rate and the number of slots is fixed, we have r_l given by (19). Also, the flow variables x_l have to satisfy the flow conservation equations in (5). Hence, the problem of maximizing the network lifetime can be written as

$$\begin{aligned} \text{minimize} \quad & q, \\ \text{subject to} \quad & Ax = S, \qquad \sum_{l \in L} m_l \le M, \quad x \succeq 0, \\ & \sum_{l \in \mathcal{O}(i)} g_l(x_l, m_l) \le qE_i, \qquad \forall i \in V, \\ & m_l \in \{0, \ldots, M\}, \qquad \forall l \in L \end{aligned} \tag{21}$$

where the variables are q, x_l's, and m_l's. Hence, it can be solved using branch and bound methods. Also, relaxing $m_l \in \{0, \ldots, M\}$ to $m_l > 0$ again gives a convex optimization problem which can be solved efficiently. The resulting solution gives the optimal transmission scheme where the MAC is variable-length TDMA.

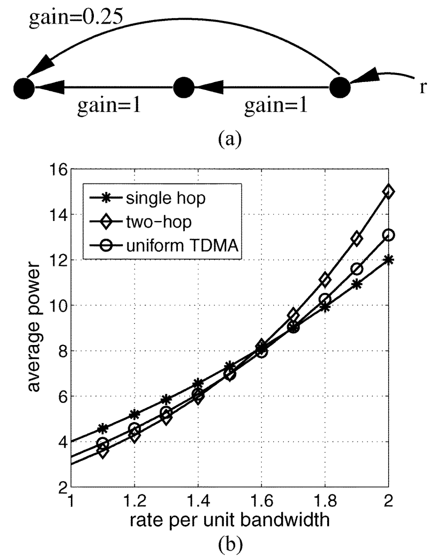


Fig. 4. Cross-layer optimization for TDMA networks. (a) Topology. (b) Power consumption.

C. Cross-Layer Power Minimization

We now consider a related problem, that of minimizing the total power consumption. We will see that for this objective function, it is sometimes possible to obtain analytical solutions. The problem of optimizing the different layers to minimize the total power consumption can be stated as the following optimization problem:

$$\begin{aligned} \text{minimize} \quad & \sum_{l \in L} g_l(x_l, m_l), \\ \text{subject to} \quad & Ax = S, \qquad \sum_{l \in L} m_l \le M, \quad x \succeq 0, \\ & m_l \in \{0, \ldots, M\}, \qquad \forall l \in L \end{aligned} \tag{22}$$

where the variables are x_l's, and m_l's. Again, for computational tractability, we make an approximation and relax m_l to take real values.

1) Example: We now demonstrate the advantages of cross-layer optimization for TDMA networks through a very simple example topology shown in Fig. 4(a). There is one source node generating data at rate r bits/Hz/s, which it needs to transmit to the destination. There is an additional relay node which can be used to relay part of the data. The channel power gains are shown adjacent to each link. We consider the following three transmission schemes.

1) *One-Hop*: Here, the source transmits directly to the destination. Hence, all the spectrum is allocated to the link from the source to the destination.
2) *Two-Hop*: Here, the source transmits all the data to the relay node which then forwards it to the destination. Since, two identical links in the network carry the same amount of data, the optimal MAC scheme is to allocate half the spectral resources each to both these transmissions.
3) *Uniform TDMA*: Unlike the above schemes, where routing is fixed, we now fix the MAC layer. Each of the three links are allocated equal time fractions to transmit. As a function of the source rate, we obtain the optimal fractions of the

data sent directly to the destination and that sent through the relay.

Thus, the optimal routing scheme is a function of the link schedule and vice versa. Also, as we see from Fig. 4(b), the optimal schemes are a function of the source rate. Hence, cross-layer optimization is necessary to determine the optimal transmission scheme to send the data from the source to the destination.

2) MAC Optimization With Link Adaptation: Here, we consider the routing scheme to be fixed. This would include the case of single hop transmissions where each node directly transmits data to the sink node. In this case, the power optimization problem can be simplified to

$$\begin{aligned} \text{minimize} \quad & \sum_{l\in L} g_l(x_l, m_l), \\ \text{subject to} \quad & \sum_{l\in L} m_l \le M \end{aligned} \tag{23}$$

where the variables are the m_l's. A specialized computational algorithm to solve the above convex program was derived in [31].

We now consider special cases where analytical solutions to the problem can be obtained.

Special Case: Identical Equidistant Nodes

Here, we assume that the functions g_l are the same for all links $l \in L$, i.e., $g_l = g, a_l = a, b_l = b$. This would be true if all nodes have identical circuits and are located at the same distance from the sink node. We consider the following two scenarios.

(a) *Minimization of the total transmission power*: Neglecting the circuit energy consumption for reception corresponds to setting $b_l = -1$. The corresponding optimization problem is as follows:

$$\begin{aligned} \text{minimize} \quad & \sum_{l\in L} am_l \left(2^{\frac{Mx_l}{m_l}} - 1\right), \\ \text{subject to} \quad & \sum_{l\in L} m_l \le M. \end{aligned} \tag{24}$$

We have the following result on the optimal solution.[1]

Lemma 4.1: The optimal solution to problem (24) is given by $m_l^* = (x_l M)/\sum_{j\in L} x_j$.

Proof: The Karush–Kuhn–Tucker (KKT) conditions (see, for example, [32]) for the problem can be written as

$$\sum_{l\in L} m_l = M,$$
$$a2^{\frac{x_l M}{m_l}}\left(1 - \ln 2\frac{x_l M}{m_l}\right) = a - \nu, \qquad i = 1, \ldots, N-1$$

for some $\nu \in \mathbb{R}$. It is easy to see that the KKT conditions are satisfied by $m_l^* = \frac{x_l}{\Sigma_{j\in L} x_j} M$, with $\nu = a - a2^{Mx_\Sigma}(1 - \ln 2x_\Sigma)$ and $x_\Sigma = \sum_{l\in L} x_l$. Since the optimization problem in (24) is convex, these m_l^*'s are optimal. ■

[1]Note that constraints such as minimum constellation size and maximum allowable average power can be incorporated easily.

(b) *Minimization of the total transmission and circuit power*: Note that $g(x_l, m_l)$ is not a monotonically decreasing function of m_l. In this case, we have the following lemma.

Lemma 4.2: Let m^* be the value of m that minimizes $am(2^{(M\Sigma_{l\in L} x_l)/m} + b)$. Then the optimal solution to the problem in (23) is given by $m_l^* = (m^* x_l)/\sum_{j\in L} x_j$.

Proof: Consider the following arbitrary feasible solution that satisfies the constraints of the optimization problem in (23): link l transmits at instantaneous rate $r_l = Mx_l/m_l$ for m_l slots. Let $\hat{M} = \sum_{l\in L} m_l$. Note that $a(2^{r_l} + b)$ is a convex function of r_l. Hence, the total power consumption across all links to transmit at average rates $(x_1, \ldots, x_{|L|})$ satisfies

$$\frac{1}{\hat{M}} \sum_{l\in L} m_l a(2^{r_l} + b) \ge a\left(2^{\frac{\sum_{l=\in L} m_l r_l}{\hat{M}}} + b\right).$$

where the inequality follows from Jensen's inequality. The quantity on the right side is the average power consumption when each nodes transmits at an instantaneous rate of $(\sum_{i\in L} m_l r_l)/\hat{M}$. Combining this with the definition of m^* in the Lemma, we can see that the optimal solution is given by $m_l^* = (m^* x_l)/\sum_{j\in L} x_j$. ■

V. Interference-Limited Wireless Sensor Networks

We now develop models for wireless networks where interfering links are allowed to transmit simultaneously. For such networks, we state the optimization problem exactly, and then suggest an iterative method to obtain approximate solutions to the problem. The transmission scheme computed during each iteration is feasible. Hence, the suboptimality of the solution can be traded off with the required computational power. The main step of our algorithm involves the solution of a convex optimization problem for which the number of variables grows as $2LN$, where N is the number of slots and L is the number of links in the network. Also, the number of constraints grows as $NL + 2|V|$, where $|V|$ is the number of nodes in the network.

Here, we consider a MAC layer that is more general than TDMA. TDMA divides the spectral resources in an orthogonal manner by scheduling interfering links at different times. In a large wireless network, consider two links such that the transmitter of one link is separated by a large distance from the receiver of another link, and vice versa. Then if the two links are scheduled simultaneously, they do not interfere much. Thus, we can take advantage of frequency reuse to schedule each link for a longer time to reduce the average transmission power on the links. If each of these links supports a high enough rate, the power consumption in the power amplifier dominates the power consumption in the rest of the transmitter circuit [see (3)]. Then the simultaneous scheduling of weakly interfering links leads to a reduction in the net power consumption.

We consider a slotted time model, and guarantee the satisfaction of average rate requirements over a predefined frame duration; this is similar to the model in [13]. We note that a slotted model enables us to use a link schedule that mitigates interference by scheduling strongly interfering links in different slots.

Thus, it is more general than the model in [38]. Also, the algorithm in [38] assumes that a feasible solution can be computed for which each link has a SINR of at least 1. This is highly unlikely in a network with many nodes and high data rates where many links will strongly interfere with each other. Moreover, the algorithm in [38] is for a single slot and does not extend to a slotted time model.

The interference model in this paper is more general than the interference model considered in [37], where the schedule for links in the same neighborhood was assumed to be orthogonal, and interference from distant links was neglected. Also, unlike in [13], we consider *rate adaptation* on links and a fixed bit error rate (BER) requirement. As shown in [11], rate adaptation can lead to significant decrease in energy consumption. However, allowing for rate adaptation on links makes the problem considerably more complex. We no longer have a linear constraint on transmission powers [13] that guarantees an SINR greater than a threshold. Instead we have a non-linear and non-convex constraint on the rate and power of each link (see Section II-D for details). In addition, we consider joint routing along with link scheduling and power control (with rate adaptation). Also, instead of minimizing the total average power consumption over the network, we maximize the minimum node lifetime. We note that only scheduling and power control to reduce energy consumption and increase just single hop throughput was considered in [8], while [9] considered a low SINR assumption and an infinite horizon for which they proposed an extremely high complexity solution to maximize throughput.

A. Cross-Layer Optimization

The transmission rate over link l, during slot m, is bounded by [see (1) and (2)]

$$\begin{aligned} r_l(m) &\leq \log\left(1+K\gamma_l^m\right) \\ &= \log\left(1+K\frac{G_{ll}p_l(m)}{\sum_{k\neq l}G_{lk}p_k(m)+N_0}\right). \end{aligned} \tag{25}$$

The average rate over link l is $x_l=\frac{1}{M}\sum_{m=1}^{M}r_l(m)$. Thus, the flow conservation equations in (5) can be written as

$$\begin{aligned} \frac{1}{M}\left(\sum_{l\in\mathcal{O}(i)}\sum_{m=1}^{M}r_l(m)-\sum_{l\in\mathcal{I}(i)}\sum_{m=1}^{M}r_l(m)\right)=S_i, \qquad \forall i\in V, \\ \Rightarrow \qquad A\sum_{m=1}^{M}r(m)=MS. \end{aligned} \tag{26}$$

Also, we have the energy conservation constraint in (6).

$$\begin{aligned} \frac{T_i}{M}\sum_{m=1}^{M}\left(\sum_{l\in\mathcal{O}(i)}(1+\alpha)p_l(m)+\left(\mathbf{1}\left(A^+p(m)\right)\right)_i\beta\right)\leq E_i, \\ \Rightarrow \quad \frac{T_{\rm net}}{M}\sum_{m=1}^{M}\left(A^+(1+\alpha)p(m)+\beta\mathbf{1}\left(A^+p(m)\right)\right)\leq E, \end{aligned} \tag{27}$$

where $T_{\rm net}=\min(T_1,\ldots,T_{|V|})$.

The goal is to compute a transmission scheme, i.e. $p_l(m)$, $r_l(m)$ (and hence x_l), to maximize the minimum node lifetime. The constraints are flow conservation and energy conservation constraints given by (26) and (27), respectively. Moreover, any physically possible transmission scheme satisfies the relation (25). This can be written as the following optimization problem in the variables $T_{\rm net}$, $p_l(m)$, $r_l(m)$, for all $l\in L$, $m=1,\ldots,M$.

$$\begin{aligned} \text{maximize} \quad & T_{\rm net}, \\ \text{subject to} \quad & A\sum_{m=1}^{M}r(m)=MS, \qquad r(m)\succeq 0, \\ & \log\left(1+K\frac{G_{ll}p_l(m)}{\sum_{k\neq l}G_{lk}p_k(m)+N_0}\right)\geq r_l(m), \\ & \frac{T_{\rm net}}{M}\sum_{m=1}^{M}\left((1+\alpha)A^+P^m+\beta\mathbf{1}\left(A^+p(m)\right)\right)\preceq E \end{aligned} \tag{28}$$

for all $m=1,\ldots,M$ and $l\in L$. As in Section III-B, we can use the change of variables $q=1/T$ to rewrite the above problem as the following optimization problem:

$$\begin{aligned} \text{minimize} \quad & q \\ \text{subject to} \quad & A\sum_{m=1}^{M}r(m)=MS, \qquad r(m)\succeq 0, \\ & \log\left(1+K\frac{G_{ll}p_l(m)}{\sum_{k\neq l}G_{lk}p_k(m)+N_0}\right)\geq r_l(m), \\ & \sum_{m=1}^{M}\left((1+\alpha)A^+P^m+\beta\mathbf{1}\left(A^+p(m)\right)\right)\preceq qME \end{aligned} \tag{29}$$

for all $m=1,\ldots,M$ and $l\in L$. This problem (as stated above) is not a convex optimization problem, and so it is hard to solve. We propose an iterative approach below where, in each iteration, we solve an approximate problem that is convex.

B. Routing and Power Control

The rate constraint in the problem formulation (29) is not convex. For a fixed link schedule, let $L(m)$, $m=1,\ldots,M$ denote the sets of links allowed to be active in each slot m. We approximate the rate constraint by

$$r_l(m)\leq\log\left(\frac{G_{ll}p_l(m)}{\sum_{k\in L(m),k\neq l}G_{lk}p_k(m)+N_0}\right). \tag{30}$$

This is a good approximation if the SINR over link l and time slot m is high. For any SINR γ, $\log(\gamma)$ is a lower bound on the achievable rate. Hence, the feasible set corresponding to the optimization problem with the above approximation is a subset of the feasible set of the original optimization problem (29). Thus, the network lifetime computed under this approximation is a lower bound on the optimum network lifetime. Using a change of variables $Q_l(m) = \log(p_l(m))$, we can rewrite the approximate rate constraint for link $l \in L(m)$ as follows (see, for example, [39]):

$$\log\left(\frac{N_0}{G_{ll}}e^{r_l(m)-Q_l(m)} + \sum_{k\in L(m),k\neq l}\frac{G_{lk}}{G_{ll}}e^{r_l(m)+Q_k(m)-Q_l(m)}\right) \leq 0.$$

The function $\log(\sum_i a_i e^{x_i})$ is convex if $a_i \geq 0, x_i \in \mathbb{R}$ (see, for example, [32]). Composition with an affine function preserves convexity. Hence, the function

$$\log\left(\frac{N_0}{G_{ll}}e^{r_l(m)-Q_l(m)} + \sum_{k\in L(m),k\neq l}\frac{G_{lk}}{G_{ll}}e^{r_l(m)+Q_k(m)-Q_l(m)}\right)$$

is convex over $r(m)$, $Q(m)$. Using the above change of variables and the approximate rate constraint (30) in problem (29), we obtain the following optimization problem.

$$\begin{aligned}
&\text{minimize } q,\\
&\text{subject to } A\sum_{m=1}^{M} r(m) = MS\\
&\qquad r(m) \succeq 0, \quad m = 1,\ldots,M,\\
&\qquad r_l(m) = 0, \qquad \forall l \notin L(m), \quad m = 1,\ldots,M,\\
&\qquad \log\Bigg(\frac{N_0}{G_{ll}}e^{r_l(m)-Q_l(m)}\\
&\qquad\qquad + \sum_{k\in L(m),k\neq l}\frac{G_{lk}}{G_{ll}}e^{r_l(m)+Q_k(m)-Q_l(m)}\Bigg),\\
&\qquad \leq 0, \qquad \forall l \in L(m), \quad m = 1,\ldots,M,\\
&\qquad \sum_{m=1}^{M}\left(\sum_{l\in\mathcal{O}(v)\cap L(m)}\left((1+\alpha)e^{Q_l(m)}+\beta\right)\right)\\
&\qquad \leq qME_i, \qquad \forall i \in V.
\end{aligned} \tag{31}$$

The variables are q, $r_l(m)$, $Q_l(m)$, for $l \in L(m)$, $m = 1,\ldots,M$. Using interior point methods (e.g., [32]), we can efficiently solve the approximate problem for optimal transmission powers and rates over each link, for a given link schedule. Geometrically, for a fixed link schedule and the convex approximation to the rate constraint, the feasible set in problem (31) is a convex subset of the feasible set in problem (29).

C. Link Scheduling

The convex optimization problem (31) is feasible only if the constraints $r_l(m) \geq 0, l \in L(M), m = 1,\ldots,M$ are feasible. For the approximate rate constraint, these constraints imply that each link has an SINR ≥ 1 during the scheduled slots. If we schedule all links during all slots, the problem may be infeasible. There is no simple characterization of the set of link schedules for which the constraints $r_l(m) \geq 0, l \in L(m), m = 1,\ldots,M$, are feasible. Hence, in order to use problem formulation (31) to compute a transmission scheme corresponding to the best link

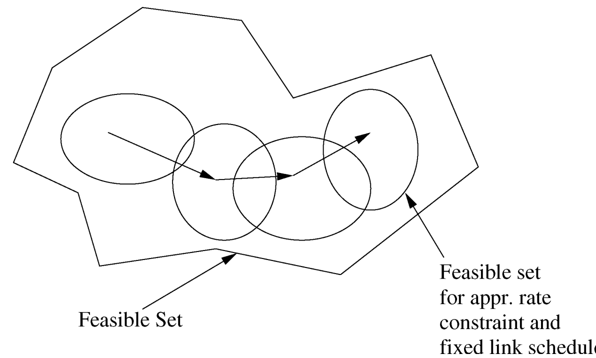


Fig. 5. Algorithm—Solving a series of approximate convex optimization problems.

schedule, we need to solve this problem for all possible link schedules. However, for a network of L links, and for a schedule frame that has M time slots, there are 2^{ML} different link schedules. Hence, the complexity of this approach is *doubly exponential* in the number of slots and the number of links. This results in a tradeoff between the computational complexity and the quality of the solution. When we compare the solution for say M slots and $2M$ slots, the solution corresponding to $2M$ slots gives a network lifetime greater than or equal to the lifetime corresponding to the solution for M slots. This is because we have more freedom in choosing the link schedules, and hence can find a transmission scheme that gives a larger network lifetime. Here, we take M to be a system constant.

We use a suboptimal approach to iterate between scheduling and computation of rates and powers. The links that carry a larger amount of traffic should be scheduled over a greater number of time slots—this decreases the average transmission power consumption over the links. Hence, the link schedule is adapted to the solution of problem (31) at each iteration, and in turn the convex optimization problem is solved for the new link schedule. Note that we motivate our link adaptation heuristic for a scenario in which transmission power dominates the power consumption in the circuit. This is a realistic scenario for interference-limited networks with high data rates. For a discussion of the tradeoffs between decreasing transmission power and decreasing circuit power, see [11].

The algorithm is illustrated in Fig. 5. Geometrically, the feasible set in problem (31) is a convex subset of the feasible set in problem (29). The heuristic approach proposed below solves a series of convex optimization problems with feasible regions given by different convex subsets of the original optimization problem in (29). Each convex subset corresponds to a link schedule and approximation of the rate constraints by convex constraints.

D. Algorithm

The iterative approach used to compute an approximate optimal strategy is summarized in the flowchart in Fig. 6. The steps of the algorithm are as follows.

1) Find an initial suboptimal, feasible schedule. A good candidate would be a schedule in which all links are activated

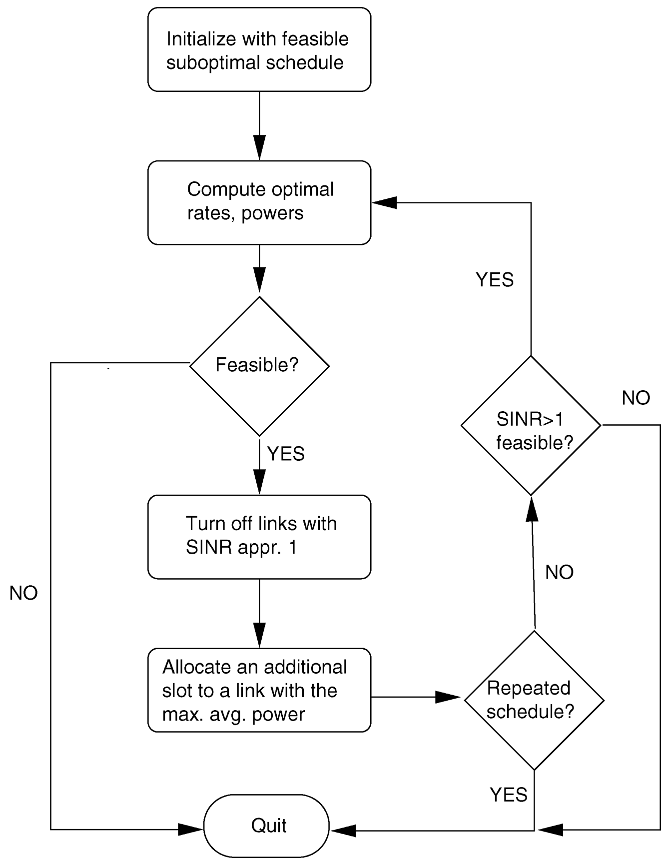


Fig. 6. Iterative approach to compute powers, rates and link schedule.

at least once in each frame of M slots, and also links that are activated in the same slot interfere only weakly.

2) Solve problem (31) to find an optimal routing flow and transmission powers during each slot for the approximate rate constraint. If the problem is infeasible, quit.
3) Remove l from $L(m)$ if $\frac{G_{ll}p_l(m)}{\Sigma_{k\in L(m),k\neq l}G_{lk}p_k(m)+N_0} \leq \gamma_0$, where $\gamma_0 > 1$ is a constant close to 1. Since we approximated the rate as $r = \log(\mathrm{SINR})$, links carry little traffic over the slots in which they have an SINR of about 1.
4) Find $\hat{l} = \arg\max_l \sum_{m=1}^{M} p_l(m)$. For this $\hat{l}$, find

$$\hat{m} = \arg\min_m \left(\sum_{k\in L(m),k\neq \hat{l}} G_{\hat{l}k}p_k(m) + N_0 \right).$$

Add $\hat{l}$ to $L(\hat{m})$. Thus, we find a link that consumes the maximum average power over the entire frame and schedule this link to be on during an additional time slot. The selected slot should be the one in which there is minimum interference to this link. If the resulting schedule is one that was used in a previous iteration, quit.
5) Check if $\frac{G_{ll}p_l(m)}{\Sigma_{k\in L(m),k\neq l}G_{lk}p_k(m)+N_0} \geq 1$ for all $l \in L(m)$ and $m = 1,\ldots,M$ is feasible. If yes, go to 2), else quit to prevent an infinite loop.

For small networks, we can use a TDMA schedule for step 1). If we assume that the maximum transmission power constraint is loose, the initial schedule is always feasible. However, for larger networks, a TDMA schedule may not be feasible due to the maximum power constraint. In this case, we can use edge coloring on the dual conflict graph [6], where only links which interfere very weakly are allowed be scheduled in the same slot. Note that we can use the gain value G_{lk} as a measure of interference. For example, we can say that a link k is said to interfere with a link l if $G_{lk} \geq \alpha$; then the total interference to any link will be bounded by $\alpha|V|^2 P_{\max}$. For a detailed discussion of this approach, we refer the readers to [6] and [7].

The algorithm uses a greedy heuristic to adaptively schedule links at each iteration, and then re-solves the convex optimization problem (31) to determine an optimal routing flow and link transmission powers and rates in each slot. Thus, our computational algorithm decouples the MAC layer from physical and routing layers. As we will see in the following subsection, even such a simple greedy heuristic can give strategies with a higher network lifetime than that given by static approaches to scheduling (e.g. TDMA and time sharing between modes in which links separated by a minimum distance are scheduled together). The gains in network lifetime are due to energy-efficient multihop routing, frequency reuse, and load balancing. Note that step 3) of the algorithm can be replaced by any general algorithm for updating the link schedule. Also, note that the algorithm will terminate in at most 2^{ML} steps. However, since the solution computed at each step is feasible, we can terminate the algorithm as soon as we have a competitive solution.

E. Example

Consider the topology shown in Fig. 7(a). We take the source rates to be $S_1,\ldots,S_9 = 0.15$ bits/s/Hz. Also, we take the distance between neighboring nodes as $d = 56$ m. Thus, the distance is large enough such that the transmission power dominates over the power consumption in the circuit. Note that the links closer to the sink support a higher average data rate than the links further away from the sink. We take the number of slots in each frame $M = 18$. We apply the algorithm proposed in the previous section to compute an efficient transmission scheme for this network. The results after 25 iterations are shown in Fig. 7. We plot the lifetime at each iteration in Fig. 7(b); the figure also shows the network lifetimes corresponding to uniform TDMA (each link is given the same amount of spectral resource), optimal TDMA, and a spatially periodic scheme, where frequency is reused on every third link. Our algorithm gives a network lifetime of 12% higher than the best among these other schemes. Also, the links carrying a higher data rate are scheduled for a higher number of slots to reduce the average transmission power. The total power consumption is shown in Fig. 7(d)—the last four nodes run out of energy at about the same time. Also, note that during each slot, more than one link is active. Thus, the algorithm does both *frequency reuse* and takes into account the *different data rate on each link*.

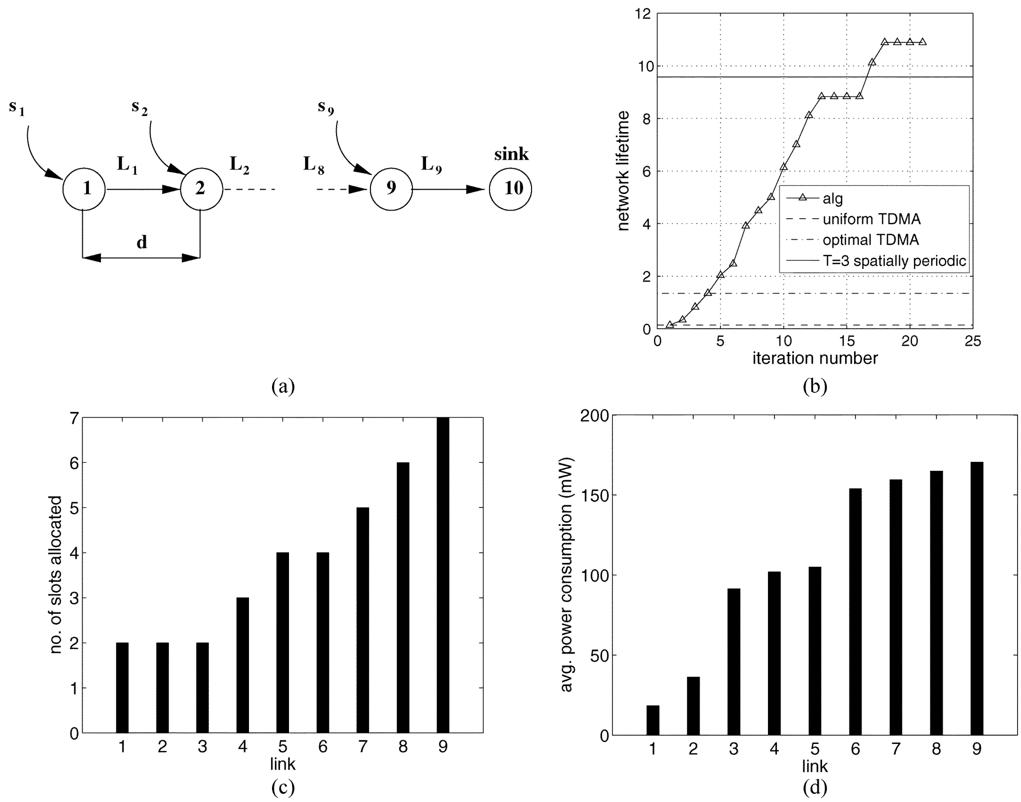


Fig. 7. Linear topology. (a) Topology. (b) Iterations. (c) Number of slots. (d) Average power.

VI. Conclusion

We described a framework for modeling energy-constrained wireless networks. The constraints imposed by the underlying system were studied and optimization problems were constructed to design such networks. In particular, we modeled the circuit energy consumption and the traditional physical, MAC, and the routing layers. We considered the optimization of individual layers as well as cross-layer optimization. The optimization problems can be solved exactly for TDMA networks. For networks with interference, we proposed approximation approaches to solve the optimization problems. The computational approaches were illustrated by several numerical examples.

References

[1] F. P. Kelly, A. Maulloo, and D. Tan, "Rate control for communication networks: Shadow prices, proportional fairness and stability," *J. Oper. Res. Soc.*, vol. 49, pp. 237–252, 1998.

[2] S. H. Low and D. E. Lapsley, "Optimization flow control. I. Basic algorithm and convergence," *IEEE/ACM Trans. Networking*, vol. 7, no. 6, pp. 861–874, Dec. 1999.

[3] S. Toumpis and A. Goldsmith, "Capacity regions for wireless ad hoc networks," *IEEE Trans. Wireless Commun.*, vol. 4, no. 4, pp. 736–748, Jul. 2003.

[4] 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, Jul. 2004.

[5] M. Johansson and L. Xiao, "Scheduling, routing and power allocation for fairness in wireless networks," in *Proc. IEEE Vehicular Technology Conf. (VTC 2004–Spring)*, 2004, pp. 1355–1360.

[6] K. Jain, J. Padhye, V. N. Padmanabhan, and L. Qiu, "Impact of interference on multi-hop wireless network performance," in *Proc. ACM MobiCom'03*, Sep. 2003, pp. 66–80.

[7] M. Kodialam and T. Nandagopal, "Characterizing achievable rates in multi-hop wireless networks: The joint routing and scheduling problem," in *Proc. ACM MobiCom'03*, Sep. 2003, pp. 42–54.

[8] T. ElBatt and A. Ephremides, "Joint scheduling and power control for wireless ad hoc networks," *IEEE Trans. Wireless Commun.*, vol. 3, no. 1, pp. 74–85, Jan. 2004.

[9] B. Radunovic and J. Y. LeBoudec, "Joint scheduling, power control and routing in symmetric, one-dimensional, multi-hop wireless networks," presented at the Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt'03), Nice, France, Mar. 2003.

[10] A. J. Goldsmith and S. B. Wicker, "Design challenges for energy-constrained ad hoc wireless networks," *IEEE Wireless Commun.Mag.*, vol. 9, no. 4, pp. 8–27, Aug. 2002.

[11] S. Cui, A. J. Goldsmith, and A. Bahai, "Energy-constrained modulation optimization," *IEEE Trans. Wireless Commun.*, vol. 4, no. 5, pp. 2349–2360, Sep. 2005.

[12] P. G. Youssef-Massaad, M. Médard, and L. Zheng, "Impact of processing energy on the capacity of wireless channels," presented at the Int. Symp. Information Theory and Its Applications (ISITA 2004), Parma, Italy, Oct. 2004, 6 pp.

[13] U. C. Kozat, I. Koutsopoulos, and L. Tassiulas, "A framework for cross-layer design of energy-efficient communication with QoS provisioning in multi-hop wireless networks," in *Proc. IEEE INFOCOM 2004*, Mar. 2004, pp. 1446–1456.

[14] D. Estrin, L. Girod, G. Pottie, and M. Srivastava, "Instrumenting the world with wireless sensor networks," in *Proc. ICASSP 2001*, May 2001, pp. 2033–2036.

[15] G. Pottie and W. Kaiser, "Wireless sensor networks," *Commun. ACM*, vol. 43, no. 5, pp. 51–58, 2000.

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

[17] L. Li and J. Y. Halpern, "Minimum energy mobile wireless networks revisited," in *Proc. IEEE Int. Conf. Communications (ICC 2001)*, Jun. 2001, pp. 278–283.

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

[19] R. L. Cruz and A. V. Santhanam, "Optimal routing, link scheduling and power control in multihop wireless networks," in *Proc. IEEE INFOCOM 2003*, Mar. 2003, pp. 702–711.

[20] S. Cui, R. Madan, A. J. Goldsmith, and S. Lall, "Cross-layer energy minimization in TDMA-based sensor networks," in *Proc. 42nd Annu. Allerton Conf. Communication, Control, and Computing*, 2004, pp. 1891–1900.

[21] R. Madan and S. Lall, "Distributed algorithms for maximum lifetime routing in wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 5, no. 8, pp. 2185–2193, Aug. 2006.

[22] R. Madan, S. Cui, S. Lall, and A. J. Goldsmith, "Cross-layer design for lifetime maximization in interference-limited wireless sensor networks," in *Proc. IEEE INFOCOM 2005*, Mar. 2005, pp. 1964–1975.

[23] S. Cui, A. Goldsmith, and A. Bahai, "Joint modulation and multiple access optimization under energy constraints," in *Proc. IEEE Globecom 2004*, Nov. 2004, pp. 151–155.

[24] R. Madan and S. Lall, "Distributed algorithms for maximum lifetime routing in wireless sensor networks," in *Proc. IEEE Globecom 2004*, Nov. 2004, pp. 748–753.

[25] S. H. A. Ahmad, A. Jovicic, and P. Viswanath, "On outer bounds to the capacity region of wireless networks," *IEEE Trans. Inf. Theory*, vol. 52, no. 6, pp. 2770–2776, Jun. 2006.

[26] B. Hughes and P. Narayan, "Gaussian arbitrarily varying channels," *IEEE Trans. Inf. Theory*, vol. IT-33, no. 2, pp. 267–284, Mar. 1987.

[27] G. J. Foschini and J. Salz, "Digital communications over fading radio channels," *Bell Syst. Tech. J.*, pp. 429–456, 1983.

[28] A. J. Goldsmith, *Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[29] P. Gupta and P. R. Kumar, "The capacity of wireless networks," *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 388–404, Mar. 2000.

[30] R. Madan and D. Shah, "Capacity-delay scaling in arbitrary wireless networks," presented at the 2005 Allerton Conf. Communication, Control, and Computing, Monticello, IL, Oct. 2005.

[31] Y. Yu, B. Krishnamachari, and V. Prasanna, "Energy-latency tradeoffs for data gathering in wireless sensor networks," in *Proc. IEEE INFOCOM 2004*, Mar. 2004, pp. 244–255.

[32] S. Boyd and L. Vandenberge, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[33] G. Zussman and A. Segall, "Energy efficient routing in ad hoc disaster recovery networks," in *Proc. IEEE INFOCOM 2003*, Mar. 2003, pp. 682–691.

[34] A. Sankar and Z. Liu, "Maximum lifetime routing in wireless ad-hoc networks," in *Proc. IEEE INFOCOM 2004*, Mar. 2004, pp. 1089–1097.

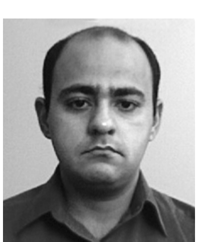
[35] M. Johansson and L. Xiao, "Scheduling, routing and power allocation for fairness in wireless networks," in *IEEE Vehicular Technology Conf. (VTC 2004–Spring)*, May 2004, pp. 1355–1360.

[36] B. Johansson, P. Soldati, and M. Johansson, "Mathematical decomposition techniques for distributed cross-layer optimization of data networks," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 8, pp. 1535–1547, Aug. 2006.

[37] R. Bhatia and M. Kodialam, "On power efficient communication over multi-hop wireless networks: joint routing, scheduling and power control," in *Proc. IEEE INFOCOM 2004*, Mar. 2004, pp. 1457–1466.

[38] M. Johansson, L. Xiao, and S. Boyd, "Simultaneous routing and power allocation in CDMA wireless data networks," in *Proc. IEEE Int. Conf. Communications (ICC'03)*, May 2003, pp. 51–55.

[39] D. O'Neill, D. Julian, and S. Boyd, "Adaptive management of network resources," in *IEEE Vehicular Technology Conf. (VTC 2003–Fall)*, Oct. 2003, pp. 1929–1933.



Ritesh Madan (S'04–M'06) received the B.Tech. degree from the Indian Institute of Technology (IIT), Mumbai, India, in 2001, and the M.S. and Ph.D. degrees from Stanford University, Stanford, CA, in 2003 and 2006, respectively, all in electrical engineering. He was a Sequoia Capital Stanford Graduate Fellow while at Stanford.

He is a Senior Systems Engineer at Qualcomm–Flarion Technologies, Bedminster, NJ. He has held visiting research positions at the Tata Institute of Fundamental Research, Mumbai, India, and Mitsubishi Electric Research Labs, Cambridge, MA. His research interests include wireless networks, optimization, distributed control, and dynamic programming.

Shuguang Cui (S'99–M'05) received the B.Eng. degree in radio engineering with the highest distinction from Beijing University of Posts and Telecommunications, Beijing, China, in 1997, the M.Eng. degree in electrical engineering from McMaster University, Hamilton, Canada, in 2000, and the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, in 2005.

From 2005 to 2007, he was working at the Department of Electrical and Computer Engineering, University of Arizona, Tucson, and is now working at the Department of Electrical and Computer Engineering, Texas A&M University, College Station. From 1997 to 1998, he worked at Hewlett-Packard, Beijing, China, as a system engineer. In the summer of 2003, he worked at National Semiconductor, Santa Clara, CA, as a wireless system researcher. His current research interests include cooperative communications, cognitive radios, cross-layer optimization for energy-constrained wireless networks, hardware and system synergies for high-performance wireless radios, and general communication theories.

Dr. Cui is the winner of the NSERC graduate fellowship from the Natural Sciences and Engineering Research Council of Canada and the Canadian Wireless Telecommunications Association (CWTA) graduate scholarship.

Sanjay Lall (S'93–M'95–SM'06) received the Ph.D. degree in engineering from the University of Cambridge, Cambridge, U.K.

He is Assistant Professor of aeronautics and astronautics at Stanford University, Stanford, CA. 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. His research interests include optimization and distributed control.

Andrea J. Goldsmith (S'90–M'93–SM'99–F'05) received the B.S., M.S., and Ph.D. degrees in electrical engineering from the University of California at Berkeley.

She is an Associate Professor of electrical engineering at Stanford University, Stanford, CA, and was previously an Assistant Professor of electrical engineering at Caltech. She has also held industry positions at Maxim Technologies and at AT&T Bell Laboratories. Her research includes work on capacity of wireless channels and networks, energy-constrained wireless communications, wireless communications for distributed control, and cross-layer design of wireless networks.

Dr. Goldsmith is a Fellow of the IEEE and of Stanford, and currently holds Stanford's Bredt Faculty Development Scholar Chair. She has received several awards for her research, including the National Academy of Engineering Gilbreth Lectureship, the Alfred P. Sloan Fellowship, the Stanford Terman Fellowship, the National Science Foundation CAREER Development Award, and the Office of Naval Research Young Investigator Award. She was also a co-recipient of the 2005 IEEEE Communications Society and Information Theory Society joint paper award. She currently serves as editor for the *Journal on Foundations and Trends in Communications and Information Theory and in Networks*, and was previously an editor for the IEEE TRANSACTIONS ON COMMUNICATIONS and for the *IEEE Wireless Communications Magazine*. She is active in committees, conference organization, and editorial boards for the IEEE Information Theory and Communications Societies and is an elected member of the Board of Governers for both societies.