Transmission Scheduling for Optimizing Sensor Network Lifetime: A Stochastic Shortest Path Approach

Yunxia Chen, *Member, IEEE*, Qing Zhao, *Member, IEEE*, Vikram Krishnamurthy, *Fellow, IEEE*, and Dejan Djonin

***Abstract*—We present transmission scheduling algorithms for maximizing the lifetime of a sensor network. In each data collection, only one group of sensors are scheduled to transmit their measurements directly to an access point (AP) through a fading channel, causing a reduction in battery energy levels of these sensors. We formulate the problem of dynamically choosing which group of sensors should communicate with the AP to maximize network lifetime as a stochastic shortest path Markov decision process. We consider three types of channel information structure: global channel state information (CSI), channel statistics, and local CSI. For optimal scheduling using global CSI (i.e., the scheduler has the information of all sensors' instantaneous channel realizations), we propose an algorithm that obtains the optimal stationary scheduling policy with computational complexity reduced from exponential to polynomial in network size when the network is dense. Due to the large implementation overhead in acquiring global CSI, we consider scheduling algorithms that use channel statistics (i.e., each sensor only knows its own channel distribution) or local CSI (i.e., each sensor knows its own instantaneous channel realization). The optimal scheduling using channel statistics is formulated as a shortest path multiarmed bandit problem, which has an indexable optimal policy. We derive a closed-form expression for the Gittins index. We further show that the Gittins index is simply the residual energy when the channel fading is identically (but not necessarily independently) distributed across sensors, i.e., the optimal scheduling policy in this case is to choose the sensor with the most residual energy. Finally, we study an asymptotically optimal scheduling protocol using local CSI. This protocol has a distributed implementation yet achieves a comparable performance as the optimal scheduling using global CSI.**

***Index Terms*—Network lifetime, sensor scheduling, shortest path multiarmed bandit process, stochastic shortest path, wireless sensor network.**

Manuscript received January 5, 2006; revised August 7, 2006. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Ananthram Swami. This work was supported by Army Research Laboratory CTA on Communication and Networks under Grant DAAD19-01-2-0011 and the National Science Foundation under Grants ECS-0622200. Part of this result was presented at the Thirty-First International Conference on Acoustics, Speech, and Signal Processing, Toulouse, France, May 2006.

Y. Chen and Q. Zhao are with the Department of Electrical and Computer Engineering, University of California, Davis, CA 95616 USA (e-mail: yxchen@ece.ucdavis.edu; qzhao@ece.ucdavis.edu).

V. Krishnamurthy is with the Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, BC V6T 1Z4, Canada (e-mail: vikramk@ece.ubc.ca).

D. Djonin is with Dypative, Inc., Vancouver, BC, Canada.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2007.893213

I. Introduction

A WIRELESS sensor network (WSN) consists of a large number of low-cost, low-power, energy-constrained sensors. Sensors are responsible for monitoring a certain phenomenon and transmitting their measurements to the access point (AP). Due to node redundancy, it is often sufficient to retrieve data from a fraction of sensors. Hence, one of the key problems arising in WSNs is sensor scheduling: which set of sensors should be scheduled to transmit their measurements for an optimal performance specified by the underlying network applications.

A. Related Work

There is a growing body of literature on the design of medium access control and sensor scheduling algorithms [1]–[6]. In principle, the optimal sensor scheduling can be formulated as a stochastic control problem under a certain performance metric and solved via dynamic programming [7]–[10]. For large sensor networks, the dynamic programming recursion can be computationally prohibitive. The underlying structures of the problems have thus been exploited to simplify the computation of the optimal policies. For example, the multiarmed bandit process, which has numerous applications in job scheduling and resource allocation, is shown to have a simple indexable optimal policy [11]–[13]. More recently, nearly optimal sensor management algorithms for energy minimization are developed in [14] and [15] under the formulation of partially observable Markov decision process (POMDP) with a multiarmed bandit structure. The Markov decision process (MDP) has also been used in the design of node activation schemes for sensor networks [16], [17].

Different performance metrics have been used in the design of scheduling policies. Examples include mutual information between the *a priori* and the *a posteriori* target states [18], mean-squared error in state estimation [10], [19], [20], the difference between the desired and the estimated covariance matrix [21], sensor usage cost [22], sum capacity [23], and network lifetime [5], [6], [24], [25]. In this paper, we use network lifetime, a design objective of paramount importance for sensor networks, as the performance measure.

Research efforts have been made to design network protocols for lifetime maximization. Most of the existing results focus on the design of optimal routing protocols in multihop networks where sensors collaboratively relay all their measurements to the AP (see [24] and [25] and references therein). In this paper, we consider sensor scheduling for applications where only a fraction of sensor measurements need to be collected and sensors communicate directly with the AP. While suboptimal and asymptotically optimal scheduling protocols have been proposed and studied in [26] and [27], the fundamental performance limit on network lifetime and the optimal sensor scheduling protocol remain unknown. This paper redresses this gap by developing lifetime-optimal transmission scheduling algorithms for single-hop networks under the framework of the MDP. An advantage of our formulation in terms of lifetime maximization is the computational tractability of the scheduling algorithms—they have polynomial or linear complexity in network size. In comparison, apart from special restrictive cases and certain toy examples, POMDPs are PSPACE-hard problems [28] and hence computationally intractable.

B. Contributions

In this paper, we address optimal transmission scheduling for maximizing sensor network lifetime. We consider applications where the AP initiates the data collection process. In each data collection slot, only a subset of sensors need to transmit their data to the AP through a fading channel, causing a reduction in their battery energy levels. We assume that the channel fading is independently and identically distributed (i.i.d.) across data collection slots, but not necessarily independent *or* identical across sensors. This block fading model holds when data collections are carried out infrequently. Extensions to Markovian fading channels are also discussed.

We formulate the problem of dynamically choosing which group of sensors should communicate with the AP for maximum network lifetime as a *stochastic shortest path* (SSP) *MDP*. We consider three types of channel information structure: global channel state information (CSI), channel statistics, and local CSI.

1) Transmission Scheduling Using Global CSI: In this case, the scheduler has the information of all sensors' instantaneous channel realizations in each data collection slot. The transmission scheduling problem is then formulated as an SSP, which is a special type of MDP that terminates with probability one in a random but finite stopping time. A general SSP problem can be solved iteratively via the stochastic dynamic programming value iteration algorithm, which generally does not converge in a finite number of iterations. Noticing that the total energy in the network decreases after each data collection slot, we show that the transition graph of the underlying Markov chain is acyclic (i.e., loop-free). Due to this acyclic structure, the value iteration algorithm converges to the optimal transmission scheduling policy in one iteration. Furthermore, by exploiting the sparsity of the transition probability matrix, the spatial aggregation, the invariance of channel statistics with respect to sensor permutation, and the fact that the channel evolution is independent of the scheduling policy, we show that the computational complexity per iteration can be reduced from exponential to polynomial in network size when the network is dense. To give more structural insights, we also give upper bounds for the maximum expected network lifetime.

The optimal scheduling algorithm using global CSI defines the limiting performance in network lifetime and provides a benchmark for all sensor scheduling algorithms. We emphasize that the optimal scheduling algorithm can be obtained offline without imposing any computation burden on sensors.

2) Transmission Scheduling Using Channel Statistics: Acquiring global CSI may result in a large implementation overhead. This motivates us to study optimal sensor scheduling using channel statistics where channel distributions rather than realizations are exploited in the transmission scheduling policy. We show that, in this case, the transmission scheduling problem is a shortest path *multiarmed bandit* process.

Multiarmed bandits have been widely studied in operations research in the context of an infinite-horizon discounted-cost stochastic control problems [11], [14], [15], [29]. Multiarmed bandits have a remarkable "indexable" property that dramatically simplifies the computation and implementation of the optimal policy for scheduling N processes (e.g., sensors). This means that associated with the state of each of the N processes is an index called the *Gittins index* such that the optimal scheduling policy is to choose the largest Gittins index at each time instant.

However, unlike discounted infinite-horizon multiarmed bandits, the optimal scheduling problem using channel statistics has a nondiscounted reward and a finite but random stopping time. Recently, nondiscounted stochastic shortest path multiarmed bandits have been considered in the context of playing golf with multiple balls [12]. It turns out, from an abstract point of view, that the optimal transmission scheduling using channel statistics is identical to the problem of playing golf with multiple balls. We can thus use the results in [12] to determine computationally efficient scheduling policies. Although the closed-form expression for the Gittins index does not exist for a general multiarmed bandit problem, we show that the rich structure of sensor scheduling problem leads to a closed-form expression for the Gittins index. We further demonstrate that choosing the sensor with the most residual energy is an optimal strategy when the channel fading is identically (but not necessarily independently) distributed across sensors. In this case, the optimal transmission scheduling policy does not even require the knowledge of channel statistics.

3) Transmission Scheduling Using Local CSI: A distributed scheduling algorithm that exploits local CSI can retain the benefit of using CSI without suffering from large implementation overhead. The basic idea is to allow each sensor to determine, based on its own channel state and residual energy, whether to transmit in a data collection slot. We formulate this problem by introducing the concept of the energy-efficiency index, which is a function of a sensor's channel state and residual energy. In each data collection slot, the sensor with the largest energy-efficiency index is scheduled for transmission using the distributed opportunistic carrier sensing scheme [30]. The scheduling problem is thus reduced to the design of the energy-efficiency index. We present an asymptotically optimal design of the energy-efficiency index and compare its performance with the limiting performance achieved by the optimal scheduling policy using global CSI.

In summary, we study sensor scheduling that exploits channel information with different levels of resolution. We aim at defining the optimal performance of sensor scheduling under each of these three formulations. By examining and comparing these three approaches to sensor scheduling, we study the tradeoff across computational complexity, implementation overhead, and performance in network lifetime.

C. Notation and Organization

All random variables (RVs) and their realizations are denoted by capital and small letters, respectively. The distribution of a discrete RV is denoted by $p_X(x) = \Pr\{X = x\}$. All vectors are denoted by boldfaced letters. For two equal-length vectors $\mathbf{x} = (x_1, x_2, \ldots, x_N)$ and $\mathbf{y} = (y_1, y_2, \ldots, y_N)$, we say $\mathbf{x} < \mathbf{y}$ if $x_k < y_k$ for all indexes $1 \leq k \leq N$. Let $1_{[x]}$ denote the indicator function: $1_{[x]} = 1$ if x is true and zero otherwise. Let

$$\mathbf{I}_n = (0, \ldots, 0, \underbrace{1}_{n-\text{th}}, 0, \ldots, 0)$$

be a $1 \times N$ unit vector and $\mathbf{I} = (1, \ldots, 1)$ be a $1 \times N$ vector of all ones.

The rest of this paper is organized as follows. In Section II, we present the network model and define the network lifetime. In Section III, we study the optimal transmission scheduling using global CSI and examine its computational complexity and implementation overhead. Section IV focuses on the optimal transmission scheduling using channel statistics, while Section V addresses the distributed scheduling using local CSI. The paper is concluded in Section VI.

II. Problem Statement

A. Network Model

Consider a WSN with N sensors, each powered by a nonrechargeable battery with initial energy $\mathcal{E}_0$. In each data collection slot, N_0 out of N sensors are scheduled to transmit their measurements to the AP through a fading channel. We assume that sensor measurement is encoded in a packet with fixed size. The channels between the AP and the sensors follow the block fading model with block length equal to the transmission time of one packet [2]. That is, channel gains are i.i.d. across data collection slots (but not necessarily i.i.d. across sensors). In Section III-E, we address the extension to Markovian fading channels.

The network model considered here applies to target detection and field estimation [31]–[33]. For example, consider a network deployed for estimating a certain parameter in a fixed area. In each data collection, every sensor in the network observes the same parameter with independent observation noise. The AP can thus collect measurements from any N_0 out of N sensors, where the number N_0 of samples is determined by the desired estimation performance (e.g., mean-square error).

For ease of presentation, we assume that $N_0 = 1$, i.e., only one sensor is chosen in each data collection slot. Our results can be extended to the general case where $N_0 > 1$ (see details in subsequent sections). When multiple sensors are scheduled in a data collection slot, we assume that the transmissions are orthogonalized (using FDMA or CDMA) so that they do not interfere with each other.

B. Energy Model

We assume that sensors can adjust their transmission power according to the fading condition and noise/interference level to ensure successful transmission. Let $\mathbf{W} = (W_1, \ldots, W_N)$ be the transmission energy requirement in a data collection slot, where W_n is the energy required for sensor n to successfully transmit a packet to the AP if it is scheduled. Note that W_n is an RV determined by the current channel gain associated with sensor n. Since channel gains are i.i.d. across data collection slots, the transmission energy requirements W_n's are i.i.d. across data collection slots. In general, the better the channel state, the lower the transmission energy requirement. In practice, sensors can only transmit at a finite number L of power levels due to hardware and power limitations. Hence, the energy requirement W_n has realizations restricted to a finite set $\mathcal{W} \triangleq \{\varepsilon_k\}_{k=1}^{L}$, where $0 < \varepsilon_1 < \cdots < \varepsilon_L < \infty$ and ε_k is the energy consumed by a sensor in transmitting a packet at the kth power level.

Let $\mathbf{E} = (E_1, \ldots, E_N)$ denote the network energy profile at the beginning of a data collection slot, where E_n is the residual energy of sensor n. Note that E_n is an RV depending on the channel states in all previous data collection slots and the transmission scheduling protocol employed in the network. Since the transmission energy consumption of a sensor is restricted to the set $\mathcal{W}$, the residual energy E_n takes values from finite set $\mathcal{E}$

$$\mathcal{E} \triangleq \left\{ e : e = \mathcal{E}_0 - \sum_{k=1}^{L} \alpha_k \varepsilon_k \geq 0 \ \text{ for some } \alpha_k \in \mathbb{Z} \text{ and } \alpha_k \geq 0 \right\}. \tag{1}$$

C. Lifetime Definition

According to the transmission energy requirement W_n and the residual energy E_n at the beginning of a data collection slot, sensor n can be in one of the following states: active, inactive, and dead. Sensor n is considered active if it has enough energy for transmission in the current data collection, i.e., $E_n \geq W_n$. Sensor n is considered dead if its residual energy E_n drops below the minimum transmission energy requirement ε_1. In other words, it does not have enough energy for transmission under any channel condition. If sensor n has residual energy higher than ε_1 but insufficient for current transmission ($\varepsilon_1 \leq E_n < W_n$), then it is considered inactive in the current data collection slot.

The definition of lifetime $\mathcal{L}$ depends on the underlying network application. One of the commonly used lifetime definitions is the number of data collections until the number of dead sensors reaches a threshold N_T [34]. For the example application of parameter estimation given in Section II-A, a natural network lifetime definition would be given by $N_T = N - N_0 + 1$, i.e., the remaining $N_0 - 1$ sensors can no longer achieve the targeted estimation performance. We also assume that the network lifetime terminates when a failure in data collection occurs, i.e., there is no active sensor in a data collection[1] ($\mathbf{E} < \mathbf{W}$). All results in this paper, however, can be extended straightforwardly without posing this condition on network lifetime.

[1]This condition on lifetime allows us to ignore the tail portion of the network lifetime when sensors only have enough energy for exceptionally good channel realizations. In this case, data collection may suffer from large delay.

D. Objectives

The goal of this paper is to design transmission scheduling algorithms that dynamically choose a sensor for transmission in each data collection slot to maximize network lifetime. We consider the following three types of channel information structure.

1) Global CSI: In this case, the scheduler has network-level knowledge of transmission energy requirement $\mathbf{W} = \mathbf{w}$ and energy profile $\mathbf{E} = \mathbf{e}$. Given $\mathbf{w}$ and $\mathbf{e}$ in a particular data collection slot, a scheduling policy chooses sensor a given by

$$a = \mu(\mathbf{e}, \mathbf{w}), \quad \mu : \mathcal{E}^N \times \mathcal{W}^N \rightarrow \{1, \ldots, N\} \tag{2}$$

where μ may be time varying (for nonstationary policies). Section III deals with computing optimal transmission scheduling policy μ for this case.

2) Channel Statistics: In this case, the scheduler only knows the distribution of the transmission energy requirement $\mathbf{W}$, along with the energy profile $\mathbf{E} = \mathbf{e}$. Admissible policies are of the form of

$$a = \mu(\mathbf{e}), \quad \mu : \mathcal{E}^N \rightarrow \{1, \ldots, N\} \tag{3}$$

where again a is the index of the scheduled sensor. We show in Section IV that the optimal transmission scheduling policy under this formulation is indexable, i.e., the optimal policy is given by

$$a^* = \arg\max\{\gamma_1(e_1), \gamma_2(e_2), \ldots, \gamma_N(e_N)\}, \ \gamma_n(e_n) : \mathcal{E} \rightarrow \mathbb{R} \tag{4}$$

where $\gamma_n(e_n)$ is the Gittins index associated with sensor n when its residual energy is e_n. In other words, the sensor with the largest Gittins index is scheduled for transmission in each data collection slot. This indexable property of the optimal policy leads to a distributed implementation as detailed in Section IV. A closed-form expression of the Gittins index is also obtained.

3) Local CSI: Admissible policies under this formulation are in the form of

$$a = \arg\max\{\gamma_1, \gamma_2, \ldots, \gamma_N\}, \ \gamma_n \triangleq g(e_n, w_n) : \mathcal{E} \times \mathcal{W} \rightarrow \mathbb{R} \tag{5}$$

where $\gamma_n \triangleq g(e_n, w_n)$ is the energy-efficiency index of sensor n, a real-valued function $g(\cdot)$ of its residual energy e_n and channel state given by w_n. Note that differing from the Gittins index, the energy-efficiency index exploits the local CSI. An asymptotically optimal design of the energy-efficiency index $g(\cdot)$ is presented in Section V along with its distributed implementation.

III. Optimal Transmission Scheduling Using Global CSI

In this section, we develop optimal transmission scheduling algorithm for maximizing network lifetime when the AP has the information of all sensors' instantaneous channel realizations. The optimal policy using global CSI defines the limiting performance in network lifetime for the model specified in Section II. It thus provides a benchmark for all sensor scheduling algorithms under the current network setup.

We formulate this problem as an SSP, which is a special class of MDP with a positive nondiscounted reward and a terminating state that is reached with probability one in a finite time [35]. The goal of an SSP problem is to maximize the total reward until the system reaches the terminating state. We then propose a complexity-reduced algorithm to obtain the optimal policy and derive upper bounds for the limiting performance achieved by the optimal policy.

A. SSP Formulation

We model the state evolution of the network by an MDP, i.e., a controlled Markovian dynamic system.

1) Sensor Network State Space: In each data collection slot, the network state is characterized by energy profile $\mathbf{e}$ and the transmission energy requirement $\mathbf{w}$. The state space $\mathcal{S}$ is defined as

$$\mathcal{S} \triangleq \{\text{state } i = (\mathbf{e}, \mathbf{w}) : \mathbf{e} \in \mathcal{E}^N, \mathbf{w} \in \mathcal{W}^N\}. \tag{6}$$

The size of the state space grows exponentially with network size N: $|\mathcal{S}| = M^N L^N$, where $M = |\mathcal{E}|$ and $L = |\mathcal{W}|$ denote, respectively, the numbers of possible residual energies and transmission energy levels.

The network enters a terminating state when its lifetime expires. According to the network lifetime definition, we define the set of terminating states $\mathcal{S}_t \subset \mathcal{S}$ as

$$\mathcal{S}_t \triangleq \{(\mathbf{e}, \mathbf{w}) : |\{e_n : e_n < \varepsilon_1\}| \geq N_T \text{ or } \mathbf{e} < \mathbf{w}\} \tag{7}$$

where $\mathbf{e} < \mathbf{w}$ indicates a failure in data collection, and $|\{e_n : e_n < \varepsilon_1\}| \geq N_T$ indicates that the number of dead sensors in the network reaches threshold N_T.

2) Action Space: In each data collection slot, according to the current network state $\mathbf{e}$ and $\mathbf{w}$, the transmission scheduling protocol chooses a sensor from the set of active sensors to transmit its packet to the AP. We thus define the action space $\mathcal{A}(i) \subseteq \{1, 2, \ldots, N\}$ in state $i = (\mathbf{e}, \mathbf{w})$ as

$$\mathcal{A}(i) = \mathcal{A}[(\mathbf{e}, \mathbf{w})] \triangleq \{n : e_n \geq w_n\} \tag{8}$$

which consists of the indexes of all active sensors.[2] According to the definition of the terminating set $\mathcal{S}_t$, we can see that there exists at least one active sensor when the network is in a nonterminating state. In other words, the action space of any nonterminating state is nonempty.

3) Controlled Markovian Dynamics: After the transmission of the scheduled sensor, the network transits to a new state according to the channel fading statistics. Let $P_{ij}^{(n)}$ denote the probability that the network transits from state $i = (\mathbf{e}, \mathbf{w})$ to state $j = (\mathbf{e}', \mathbf{w}')$ after sensor n is chosen for transmission. Since the terminating states are absorbing, we have $P_{ii}^{(n)} = 1$ for any terminating state $i \in \mathcal{S}_t$ and any action $n \in \mathcal{A}(i)$. In each data collection slot, only the residual energy e_n of the scheduled sensor n changes according to its transmission energy requirement w_n while other sensors' residual energies remain the same. Hence, for any nonterminating state $i = (\mathbf{e}, \mathbf{w}) \in \mathcal{S} \backslash \mathcal{S}_t$, state $j = (\mathbf{e}', \mathbf{w}') \in \mathcal{S}$, and action $n \in \mathcal{A}(i)$, the transition probability $P_{ij}^{(n)}$ is given by

$$P_{ij}^{(n)} = p_{\mathbf{W}}(\mathbf{w}') 1_{[\mathbf{e}' = \mathbf{e} - \mathbf{I}_n w_n]} \tag{9}$$

[2]We can extend our analysis to the general case where a group of N_0 sensors are chosen for transmission in each data collection slot by modifying the definition of the action space as $\mathcal{A}[(\mathbf{e}, \mathbf{w})] \triangleq \{(n_1, \ldots, n_{N_0}) : e_{n_k} \geq w_{n_k} \text{ for all } 1 \leq k \leq N_0\}$.

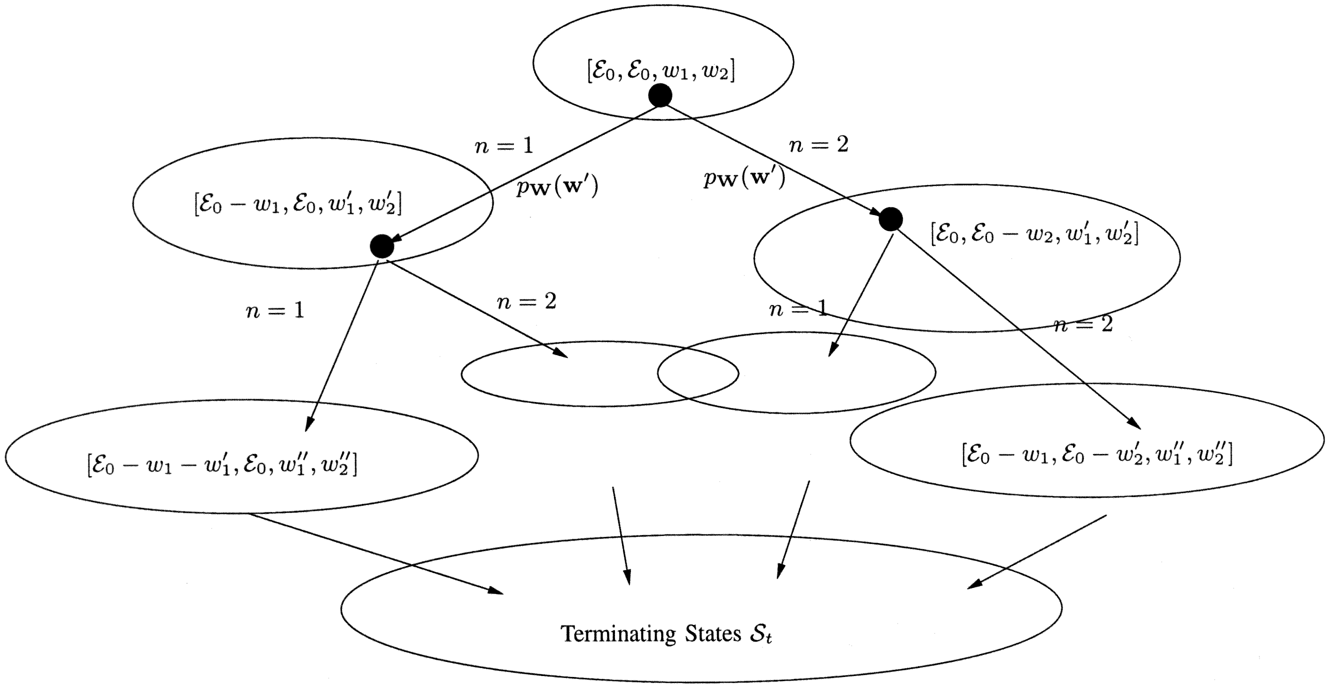


Fig. 1. Stochastic shortest path formulation of sensor scheduling using global CSI: $N = 2$.

where $p_{\mathbf{W}}(\mathbf{w}') = \Pr\{\mathbf{W} = \mathbf{w}'\}$ is the probability mass function (PMF) of $\mathbf{W}$ determined by the channel-fading statistics for a given set $\mathcal{W}$ of transmission energy levels.

4) Example: Fig. 1 illustrates a network comprising of two sensors ($N = 2$). The network state transits from $(\mathcal{E}_0, \mathcal{E}_0, w_1, w_2)$ to $(\mathcal{E}_0 - w_1, \mathcal{E}_0, w_1', w_2')$ with probability $p_{\mathbf{W}}(\mathbf{w}')$ if sensor $n = 1$ is chosen; it transits to state $(\mathcal{E}_0, \mathcal{E}_0 - w_2, w_1', w_2')$ with probability $p_{\mathbf{W}}(\mathbf{w}')$ if sensor $n = 2$ is chosen. The bottom ellipse in Fig. 1 indicates the set $\mathcal{S}_t$ of terminating states. The fact that sensor batteries have finite initial energy and that each transmission consumes nonzero energy implies that the network always reaches a terminating state in a finite but random time. The inevitable termination makes the sensor scheduling problem an instance of an SSP problem. The design objective of transmission scheduling policy is to maximize the time before the network reaches a terminating state in $\mathcal{S}_t$, i.e., to maximize the network lifetime.

5) Transmission Reward: Maximizing the expected network lifetime is equivalent to assigning a unit reward to each data collection slot until the network reaches a terminating state after which no reward is earned. Accordingly, if the network is in state $i = (\mathbf{e}, \mathbf{w})$ in a particular data collection slot, we define the instantaneous reward $R(i)$ in this data collection slot as

$$R(i) = R[(\mathbf{e}, \mathbf{w})] \triangleq 1_{[i \in \mathcal{S} \backslash \mathcal{S}_t]}. \quad (10)$$

Hence, the total reward accumulated until the network reaches a terminating state in $\mathcal{S}_t$ represents network lifetime.

6) SSP Formulation: We have formulated the sensor transmission scheduling problem as an SSP. A transmission scheduling protocol is thus a policy $\boldsymbol{\pi}$ of this SSP. A policy $\boldsymbol{\pi}$ is given by a sequence of functions $\boldsymbol{\pi} = \{\mu_1, \mu_2, \ldots, \}$, where $\mu_l : \mathcal{S} \rightarrow \{1, \ldots, N\}$ specifies the sensor chosen in the lth data collection slot. If μ_l is identical for all l, π is a stationary policy.

Let $\mathcal{L}_{\boldsymbol{\pi}}(i)$ denote the expected network lifetime (i.e., the total expected reward in the SSP problem) starting from state i with policy π. The maximum expected lifetime $\mathcal{L}^*(i)$ starting from state i is given by

$$\mathcal{L}^*(i) = \max_{\boldsymbol{\pi}} \mathcal{L}_{\boldsymbol{\pi}}(i). \quad (11)$$

A policy $\boldsymbol{\pi^*}$ is called optimal if it achieves the maximum expected lifetime at all nonterminating states in $\mathcal{S} \backslash \mathcal{S}_t$, i.e.,

$$\mathcal{L}_{\boldsymbol{\pi}^*}(i) = \mathcal{L}^*(i), \quad \forall i \in \mathcal{S} \backslash \mathcal{S}_t. \quad (12)$$

Thus, the optimal sensor scheduling protocol is given by the optimal policy $\boldsymbol{\pi^*}$ of the above SSP problem.

B. Stationary Optimal Policy

The maximum expected lifetime $\mathcal{L}^*(i)$ starting from state $i = (\mathbf{e}, \mathbf{w})$ can be obtained as the unique solution to the Bellman's optimality equation [35]

$$\begin{aligned}\mathcal{L}^*(i) &= \mathcal{L}^*[(\mathbf{e}, \mathbf{w})] \\ &= R(i) + \max_{n \in \mathcal{A}(i)} \left\{ \sum_{j \in \mathcal{S}} P_{ij}^{(n)} \mathcal{L}^*(j) \right\}, \quad \forall i \in \mathcal{S}. \quad (13)\end{aligned}$$

Since the terminating states are absorbing states with a zero reward, the maximum expected lifetime starting from a terminating state is zero, i.e., $\mathcal{L}^*(i) = 0$ for $i \in \mathcal{S}_t$.

Stationary policies are generally preferred due to reduced memory requirements and low complexity in implementation. It has been shown in [35] that any SSP has a stationary optimal policy due to its inevitable termination. A stationary optimal policy μ for scheduling using global CSI is given by [35]

$$\mu(i) = \arg \max_{n \in \mathcal{A}(i)} \left\{ \sum_{j \in \mathcal{S}} P_{ij}^{(n)} \mathcal{L}^*(j) \right\}, \quad \forall i \in \mathcal{S} \backslash \mathcal{S}_t. \quad (14)$$

Clearly, if we obtain the maximum expected lifetime $\mathcal{L}^*(j)$ for all states, the optimal policy can be readily computed. Obtaining the optimal scheduling policy thus hinges on an efficient computation of the maximum expected lifetime given in (13).

The value iteration algorithm [35, p. 303] is a widely used iterative procedure to compute the maximum expected total reward for an SSP problem. Specifically, we initialize the value iteration algorithm at $\mathcal{L}_0(i) = 0$ for all $i \in \mathcal{S}$. In a typical iteration k of the value iteration algorithm, we calculate [35]

$$\mathcal{L}_k(i) = R(i) + \max_{n\in\mathcal{A}(i)} \left\{ \sum_{j\in\mathcal{S}} P_{ij}^{(n)} \mathcal{L}_{k-1}(j) \right\}, \quad \forall i \in \mathcal{S}\backslash\mathcal{S}_t. \tag{15}$$

It has been shown [35] that the value iteration algorithm always converges, i.e.,

$$\mathcal{L}^*(i) = \lim_{k\to\infty} \mathcal{L}_k(i), \quad \forall i \in \mathcal{S}\backslash\mathcal{S}_t. \tag{16}$$

Unfortunately, it generally requires an infinite number of iterations to converge [35]. Furthermore, in each iteration, the computational complexity (measured as the number of multiplications) is quadratic in the number $|\mathcal{S}\backslash\mathcal{S}_t|$ of nonterminating states and linear in the number $|\mathcal{A}(i)|$ of actions (as can be seen from (15)). Hence, the complexity of computing maximum expected lifetime $\mathcal{L}^*(i)$ for every state $i \in \mathcal{S}$ is on the order of $N(LM)^{2N}$ per iteration, which increases exponentially with network size N.

C. Computing Maximum Network Lifetime With Polynomial Complexity in Network Size

By exploiting the underlying structure of the sensor scheduling problem, we can significantly reduce the computational complexity of the value iteration algorithm. We first show that for the scheduling problem, the value iteration algorithm converges in one iteration. We then reduce the computational complexity of this iteration from exponential to polynomial in network size.

1) Acyclic Transition Graph: Due to the fact that the total residual energy in the network decreases after each data collection slot, we can show that the value iteration algorithm converges in one iteration.

Proposition 1: For any transmission scheduling policy, the transition graph of the underlying Markov chain is acyclic. As a consequence, the expected network lifetime achieved by the optimal sensor scheduling using global CSI can be obtained in one iteration.

Proof: Since the total energy in the network decreases after each data collection slot, it is easy to see that in the transition graph corresponding to any transmission scheduling policy, there is no arc from state i to state j if state j has higher total energy, i.e., the corresponding transition graph is acyclic.

It has been shown in [36, p. 89] that value iteration converges in one step if the transition graph corresponding to the optimal policy is acyclic. Hence, for sensor scheduling problem, the maximum expected network lifetime can be obtained in one iteration by calculating (13) in an increasing order of total energy in the network. ■

2) Sparse Transition Matrix: We have reduced the number of iterations required to compute the maximum expected lifetime to one. Next, we focus on reducing the computational complexity of the Bellman's optimality (13) by reducing the size of the state space over which the maximum expected lifetime is executed.

We notice that the transition matrix given in (9) is sparse. Substituting (9) into (13), we obtain the maximum expected lifetime as

$$\begin{aligned}&\mathcal{L}^*[(\mathbf{e},\mathbf{w})]\\ &= R[(\mathbf{e},\mathbf{w})] + \max_{n\in\mathcal{A}[(\mathbf{e},\mathbf{w})]} \left\{ \sum_{\mathbf{w}'} p_{\mathbf{W}}(\mathbf{w}') \times \mathcal{L}^*[(\mathbf{e}-\mathbf{I}_n w_n, \mathbf{w}')] \right\},\\ &\quad \forall (\mathbf{e},\mathbf{w}) \in \mathcal{S}\backslash\mathcal{S}_t.\end{aligned} \tag{17}$$

Note that the summation in the curly parenthesis of (17) is taken over $|\mathcal{W}|^N = L^N$ channel states while that of (13) is taken over $|\mathcal{S}| = (LM)^N$ network states. Hence, the computational complexity of (17) is on the order of $NL^{2N}M^N$ instead of $N(LM)^{2N}$ as in (13).

3) Uncontrollable Noncorrelated Channel States: Recall that a network state $(\mathbf{e},\mathbf{w})$ consists of two components: the network energy profile $\mathbf{e}$ and the transmission energy requirement $\mathbf{w}$. The transition of $\mathbf{e}$ is affected by the chosen action, but the transition of $\mathbf{w}$ is not since $\mathbf{w}$ is determined solely by fading statistics. Define a new value function $\hat{\mathcal{L}}[\mathbf{e}]$, the maximum expected lifetime starting from network energy profile $\mathbf{e}$, as

$$\hat{\mathcal{L}}[\mathbf{e}] = \sum_{\mathbf{w}} p_{\mathbf{W}}(\mathbf{w}) \mathcal{L}^*[(\mathbf{e},\mathbf{w})]. \tag{18}$$

Averaging (17) over all channel realizations $\mathbf{w}$, we obtain an equivalent modified Bellman's optimality equation [35] for $\hat{\mathcal{L}}[\mathbf{e}]$ as

$$\hat{\mathcal{L}}[\mathbf{e}] = \sum_{\mathbf{w}} p_{\mathbf{W}}(\mathbf{w}) \left\{ R[(\mathbf{e},\mathbf{w})] + \max_{n\in\mathcal{A}[(\mathbf{e},\mathbf{w})]} \hat{\mathcal{L}}[\mathbf{e} - \mathbf{I}_n w_n] \right\}. \tag{19}$$

Equation (19) is executed over the space of $\mathbf{e}$ rather than $(\mathbf{e},\mathbf{w})$ as in (17). Hence, the computational complexity is reduced to $\mathcal{O}(NL^N M^N)$. The optimal policy can also be computed by

$$\mu[(\mathbf{e},\mathbf{w})] = \arg\max_{n\in\mathcal{A}[(\mathbf{e},\mathbf{w})]} \hat{\mathcal{L}}[\mathbf{e} - \mathbf{I}_n w_n]. \tag{20}$$

4) Invariance to Sensor Permutation: Equation (19) holds for any channel distribution. We can further simplify the calculation of the maximum expected lifetime when the channel distribution satisfies certain conditions.

Suppose that the joint channel distribution $p_{\mathbf{W}}(\mathbf{w})$ is invariant to sensor permutations, i.e., $p_{\mathbf{W}}(\mathbf{w}) = p_{\mathbf{W}}(\tilde{\mathbf{w}})$ if $\tilde{\mathbf{w}}$ is a permutation of $\mathbf{w}$. Note that this condition is satisfied when the channel fading is i.i.d. across sensors. From (18), we can thus show that $\hat{\mathcal{L}}[\mathbf{e}] = \hat{\mathcal{L}}[\tilde{\mathbf{e}}]$ if $\tilde{\mathbf{e}}$ is a permutation of $\mathbf{e}$, or equivalently $\tilde{\mathbf{e}}$ and $\mathbf{e}$ have the same pattern. Hence, we only need to compute the maximum expected lifetime $\hat{\mathcal{L}}[\mathbf{e}]$ for different patterns of the network residual energy profile $\mathbf{e}$ rather than all possible $\mathbf{e}$. Since the number of patterns of the network energy profile $\mathbf{e}$ is polynomial $\mathcal{O}(N^{M-1})$ in network size, we reduce the complexity of the maximum expected lifetime from $\mathcal{O}(NL^N M^N)$ as in (19) to $\mathcal{O}(N^M L^N)$ in network size N.

TABLE I
COMPUTATIONAL COMPLEXITY OF MAXIMUM EXPECTED LIFETIME

Special structure	Iterations	Complexity per Iteration in N
General SSP problem	∞	$\mathcal{O}(N(LM)^{2N})$
Acyclic transition graph	1	$\mathcal{O}(N(LM)^{2N})$
Sparse transition matrix	1	$\mathcal{O}(NL^{2N}M^{N})$
Uncontrollable non-correlated channel state	1	$\mathcal{O}(NL^{N}M^{N})$
Invariance to sensor permutation	1	$\mathcal{O}(N^{M}L^{N})$
Spatial aggregation	1	$\mathcal{O}(N^{M})$

N: number of sensors

L: number of power levels

M: number of possible residual energies

5) Spatial Aggregation: In dense WSNs, closely spaced sensors may experience approximately the same channel fading. According to sensor locations, we can classify sensors into $\tilde{N} \ll N$ spatial clusters such that the transmission energy requirement is identical for sensors within a given spatial cluster. As a consequence, the space size of the transmission energy requirement $\mathbf{w}$ can be reduced from L^N to $L^{\tilde{N}}$. Suppose that the distribution of the channel fading is invariant to cluster permutations, i.e., $p_{\mathbf{W}_c}(\mathbf{w}_c) = p_{\mathbf{W}_c}(\tilde{\mathbf{w}}_c)$ if $\tilde{\mathbf{w}}_c$ is a permutation of $\mathbf{w}_c = (w_{c_1}, \ldots, w_{c_{\tilde{N}}})$, where w_{c_j} is the transmission energy requirement for cluster c_j. The computational complexity of the value iteration can then be reduced to $\mathcal{O}(N^M L^{\tilde{N}}) = \mathcal{O}(N^M)$ in network size N if the number $\tilde{N}$ of clusters is independent of the network size N, which holds in dense networks deployed over fixed geographic areas.

In Table I, we summarize the computational complexity of the maximum expected lifetime. By exploiting the special structures of the sensor transmission scheduling problem, we show that the value iteration algorithm converges in one iteration. For dense networks deployed over fixed geographic areas, we reduce the computational complexity from exponential to polynomial in network size.[3] The memory required to store the optimal policy grows linearly with the number of different network states, which is also polynomial in network size N for dense networks. We point out that the optimal policy can be obtained offline and stored at the AP without imposing any computation burden or memory requirement on sensors.

D. Upper Bound on Maximum Expected Lifetime

The maximum expected lifetime not only serves as the stepping stone for obtaining the optimal scheduling policy but also provides a benchmark to which suboptimal sensor scheduling protocols can be compared. A simple upper bound on the maximum expected lifetime is thus desirable.

Noticing the increasing property of network lifetime $\hat{\mathcal{L}}[\mathbf{e}]$ in energy profile $\mathbf{e}$, i.e., $\hat{\mathcal{L}}[\mathbf{e}] \le \hat{\mathcal{L}}[\mathbf{e}']$ if $\mathbf{e} \le \mathbf{e}'$, we obtain an upper bound on $\hat{\mathcal{L}}[\mathbf{e}]$ from (19) as

$$\begin{aligned}\hat{\mathcal{L}}[\mathbf{e}] &\le R[\mathbf{e}] + \sum_{\mathbf{w}} p_{\mathbf{W}}(\mathbf{w}) \max_{n\in\mathcal{A}[(\mathbf{e},\mathbf{w})]}\left\{\hat{\mathcal{L}}[\mathbf{e}-\mathbf{I}_n w_{\min}]\right\} \\ &= R[\mathbf{e}] + \sum_{k=1}^{L} p_{W_{\min}}(\varepsilon_k) \max_{n\in\mathcal{A}[(\mathbf{e},\mathbf{w})]}\left\{\hat{\mathcal{L}}[\mathbf{e}-\mathbf{I}_n\varepsilon_k]\right\} \end{aligned} \quad (21)$$

where $R[\mathbf{e}] = \sum_{\mathbf{w}} p_{\mathbf{W}}(\mathbf{w}) R[(\mathbf{e},\mathbf{w})]$, $W_{\min} = \min\{\mathbf{W}\}$ is the minimum transmission energy requirement among sensors, and $p_{W_{\min}}(x)$ is the PMF of $W_{\min}$. To further simplify the calculation of (21), we define $d(\mathbf{e}) \triangleq \sum_{n=1}^{N} e_n$ as the total energy in state $\mathbf{e}$, and

$$R(d) \triangleq \max_{\mathbf{e}:d(\mathbf{e})=d}\{R[\mathbf{e}]\}, \quad \hat{\mathcal{L}}_m(d) \triangleq \max_{\mathbf{e}:d(\mathbf{e})=d}\{\hat{\mathcal{L}}[\mathbf{e}]\}. \quad (22)$$

Recall that the network is considered dead if the number of dead sensors reaches threshold N_T. Hence, for a fixed total energy d in the network, the most sensor residual energy in a nonterminating network state is bounded above by $d-(N-N_T)\varepsilon_1$. Noting that $R[(\mathbf{e},\mathbf{w})] \le 1_{[\max \mathbf{e} \ge \min \mathbf{w}]}$, we bound $R(d)$ as

$$\begin{aligned} R(d) &\le \max_{\mathbf{e}:d(\mathbf{e})=d}\left\{\sum_{\mathbf{w}} p_{\mathbf{W}}(\mathbf{w}) 1_{[\max \mathbf{e}\ge\min\mathbf{w}]}\right\} \\ &\le \sum_{\mathbf{w}} p_{\mathbf{W}}(\mathbf{w}) 1_{[d-(N-N_T)\varepsilon_1 \ge \min \mathbf{w}]} \\ &= \sum_{k=1}^{L} p_{W_{\min}}(\varepsilon_k) 1_{[d-(N-N_T)\varepsilon_1\ge\varepsilon_k]} \triangleq R_m(d). \end{aligned} \quad (23)$$

With the definitions in (22) and the upper bound on $R(d)$, we obtain an upper bound on the maximum expected lifetime $\hat{\mathcal{L}}[\mathbf{e}]$ when the network starts with an energy profile $\mathbf{e}$, as follows:

$$\hat{\mathcal{L}}[\mathbf{e}] \le \hat{\mathcal{L}}_m(d), \quad \text{where} \quad d = \sum_{n=1}^{N} e_n. \quad (24)$$

This upper bound can be computed by a simple linear recursion over d

$$\begin{aligned}\hat{\mathcal{L}}_m(d) &= R_m(d) + \sum_{k=1}^{L} p_{W_{\min}}(\varepsilon_k)\hat{\mathcal{L}}_m(d-\varepsilon_k) \\ &= \sum_{k=1}^{L} p_{W_{\min}}(\varepsilon_k)\left[1_{[d-(N-N_T)\varepsilon_1\ge\varepsilon_k]} + \hat{\mathcal{L}}_m(d-\varepsilon_k)\right] \end{aligned} \quad (25)$$

with $\hat{\mathcal{L}}_m(d) = 0$ if $d \le 0$.

Below we derive a closed-form expression for the upper bound $\hat{\mathcal{L}}_m(d)$ given in (25). Without loss of generality, we assume that the initial energy $\mathcal{E}_0$ and the transmission energy consumption $\varepsilon_k = k$ are integers. Hence, the total energy $d = d(\mathbf{e})$ for a nonterminating state $\mathbf{e}$ is a positive integer. Taking the Z-transform on both sides of (25), we obtain that

$$\mathcal{Z}[\hat{\mathcal{L}}_m(d)](z) = \frac{\mathcal{Z}[R_m(d)](z)}{1 - \sum_{k=1}^{L} p_{W_{\min}}(k) z^{-k}} \quad (26)$$

[3]Note that in the general case where $N_0 > 1$ measurements need to be collected, the computational complexity of the maximum expected lifetime can still be reduced to polynomial in network size. Specifically, for a network deployed to monitor a phenomenon within a fixed geographic area, the amount of required information characterized by the number N_0 of samples is a fixed value determined by the targeted detection and estimation performance. As a consequence, the size of the action space given by $\binom{N}{N_0}$ grows polynomially with network size. The complexity of scheduling using global CSI thus remains to be polynomial in network size due to the fact that the complexity of MDP is linear in the size of action space.

where $\mathcal{Z}[f(d)](z) = \sum_{d=0}^{\infty} f(d)z^{-d}$ is the Z-transform of the sequence $\{f(d)\}_{d=0}^{\infty}$. Hence, the upper bound (25) can be computed as

$$\hat{\mathcal{L}}_m(d) = R_m(d) * \mathcal{Z}^{-1}\left[H(z)\right](d) \tag{27}$$

where $f(d) * g(d) = \sum_{n=0}^{d} f(n)g(d-n)$ is the convolution of two sequences $\{f(d)\}_{d=0}^{\infty}$ and $\{g(d)\}_{d=0}^{\infty}$, $\mathcal{Z}^{-1}\left[H(z)\right](d)$ is the inverse Z-transform of $H(z)$ defined as

$$H(z) \triangleq \left[1 - \sum_{k=1}^{L} p_{W_{\min}}(k) z^{-k}\right]^{-1} \tag{28}$$

which can be readily calculated via common mathematical software such as Mathematica.

Next, we study the asymptotic behavior of the upper bound given in (27). By the Perron–Frobenius theorem, $H(z)$ has a pole at 1 and all other poles inside the unit circle. Hence, as d goes to infinity, the contribution of all other poles diminishes, and the inverse Z-transform of $H(z)$ approaches a constant c. We also notice that when the total energy d is sufficiently large, the expected immediate reward $R_m(d)$ reaches 1. From (27), we can see that when d is sufficiently large, the upper bound $\hat{\mathcal{L}}_m(d)$ behaves as a linear function in d with an increasing rate c. Suppose that $\hat{\mathcal{L}}_m(d) = c_0 + cd$ for sufficiently large d. Substituting it into (25), we obtain the asymptotic increase rate c of $\hat{\mathcal{L}}_m(d)$ with total energy d as

$$c = \frac{1}{\sum_{k=1}^{L} \varepsilon_k p_{W_{\min}}(\varepsilon_k)} = \frac{1}{\mathbb{E}[W_{\min}]}. \tag{29}$$

Since $\hat{\mathcal{L}}_m(d)$ is an upper bound on the maximum expected lifetime, (29) provides an upper bound on the asymptotic increase rate of the maximum expected lifetime with respect to the total energy in the network.

E. Extension to Markovian Fading Model

In Markovian fading channels, the transmission energy requirement $\mathbf{W}$ follows a Markov chain with $|\mathcal{W}|^N = L^N$ states and transition probabilities $p(\mathbf{w}' \mid \mathbf{w})$ indicating the probability that the transmission energy requirement transits from $\mathbf{w}$ to $\mathbf{w}'$. Following Section III-A, we can formulate sensor scheduling using global CSI in Markovian fading channels as an SSP problem by modifying the definition of the transition probability from state $i = (\mathbf{e}, \mathbf{w})$ to state $j = (\mathbf{e}', \mathbf{w}')$ under action n as

$$P_{ij}^{(n)} = p(\mathbf{w}' \mid \mathbf{w}) 1_{[\mathbf{e}' = \mathbf{e} - \mathbf{I}_n w_n]}. \tag{30}$$

Next, we show that the complexity of scheduling using global CSI in Markovian channels can still be reduced to polynomial in network size for dense networks. Exploiting the acyclic structure of the transition graph and the sparseness of the transition matrix, we obtain the maximum expected lifetime starting from state $(\mathbf{e}, \mathbf{w}) \in \mathcal{S} \backslash \mathcal{S}_t$ as

$$\mathcal{L}^*[(\mathbf{e}, \mathbf{w})] = R[(\mathbf{e}, \mathbf{w})] + \max_{n \in \mathcal{A}[(\mathbf{e}, \mathbf{w})]} \left\{ \sum_{\mathbf{w}'} p(\mathbf{w}' \mid \mathbf{w}) \times \mathcal{L}^*[(\mathbf{e} - \mathbf{I}_n w_n, \mathbf{w}')] \right\}. \tag{31}$$

The complexity of (31) is $\mathcal{O}(NL^{2N}M^N)$ in network size N. Assuming that the Markovian fading channels are invariant to sensor permutations, *i.e.,* $p(\mathbf{w}' \mid \mathbf{w}) = p(\tilde{\mathbf{w}}' \mid \tilde{\mathbf{w}})$ where $\tilde{\mathbf{w}}$ is a permutation of $\mathbf{w}$, we have $\mathcal{L}^*[(\mathbf{e}, \mathbf{w})] = \mathcal{L}^*[(\tilde{\mathbf{e}}, \tilde{\mathbf{w}})]$. Hence, we only need to calculate (31) for different patterns of $(\mathbf{e}, \mathbf{w})$ and its complexity reduces to $\mathcal{O}(N^{LM}L^N)$ in network size. Furthermore, applying spatial aggregation in dense networks and assuming invariant Markovian fading channels across clusters, we can reduce the complexity of (31) to a polynomial in network size: $\mathcal{O}(N^{\tilde{N}M})$, where $\tilde{N}$ is the number of clusters.

F. Implementation and Overhead

To implement the optimal sensor scheduling using global CSI, we need the information on network energy profile $\mathbf{e}$ and transmission energy requirement $\mathbf{w}$. The latter requires the knowledge of all sensors' instantaneous channel realizations. One possible implementation is described below. At the beginning of a data collection slot, the AP broadcasts a beacon signal to activate sensors in the network. Every sensor responds to the beacon signal by sending pilot signals to the AP for global CSI acquisition. Using the responses from sensors, the AP estimates the channel realizations of all sensors and obtain the transmission energy requirement $\mathbf{w}$. It then determines which sensor to schedule according to the *precomputed* optimal policy μ and the current network state $(\mathbf{e}, \mathbf{w})$. Finally, the AP broadcasts the ID of the chosen sensor and the required transmission power level. The chosen sensor then transmits its measurements to the AP at the required power level. Since the AP knows the scheduled sensor's ID and all sensors' channel realizations, it can keep track of the network energy profile $\mathbf{e}$ using energy consumption characteristics. Specifically, let v denote the energy consumed by a sensor in transmitting pilot signals to the AP for global CSI acquisition. If the network energy profile at the beginning of a data collection slot is $\mathbf{e}$, then at the end of this slot, the residual energy profile is given by $\mathbf{e}' = \mathbf{e} - v\mathbf{I} - \mathbf{I}_n w_n$, where sensor n is the scheduled one.

The main drawback of the centralized scheduling is its implementation overhead. At the beginning of each data collection slot, all sensors need to transmit pilot signals over fading channels to the AP for global CSI acquisition. Besides the complex issue of multiple access, the transmission of pilot signals by every sensor can be energy consuming. Hence, for large sensor networks, the energy consumed in acquiring global CSI may override the benefit of exploiting CSI as demonstrated in Section IV-F. Nevertheless, assuming cost-free channel acquisition, the optimal scheduling algorithm using global CSI defines the limiting performance of any sensor scheduling protocols in terms of network lifetime.

G. Numerical Examples

In Figs. 2 and 3, we study several scheduling protocols by comparing their lifetime with the limiting performance $\hat{\mathcal{L}}[\mathcal{E}_0\mathbf{I}]$ given in (19). We ignore the overhead in channel acquisition in these examples and focus on the gap between suboptimal performances and the limiting performance. Specifically, we consider the following three scheduling protocols: 1) the layered approach which assumes that sensors are indistinguishable at the physical layer and randomly schedules a sensor for transmission; 2) the pure opportunistic protocol which exploits the

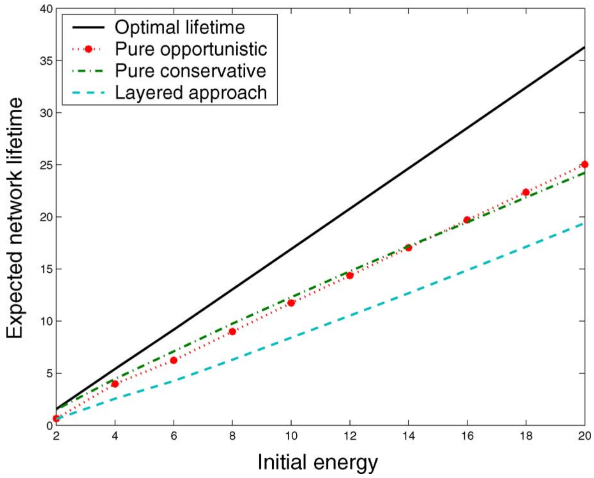


Fig. 2. Network lifetime comparison. $N = 3$ (number of sensors), $N_T = 1$ (threshold on the number of dead sensors in the lifetime definition), $L = 3$ (number of power levels), and $\mathcal{W} = \{1, 2, 3\}$ (transmission energy levels). Channel distribution: $p_{W_n}(1) = p_{W_n}(2) = 1/4$, $p_{W_n}(3) = 1/2$, for all n.

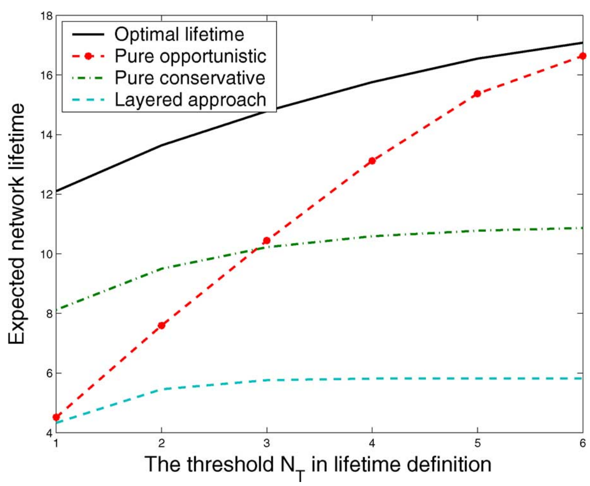


Fig. 3. Network lifetime comparison under different lifetime definitions. $N = 6$ (number of sensors) , $\mathcal{E}_0 = 3$ (sensor initial energy), $L = 2$ (number of power levels), and $\mathcal{W} = \{1, 2\}$ (transmission energy levels). Channel distribution: $p_{W_n}(1) = 0.6$, $p_{W_n}(2) = 0.4$, for all n.

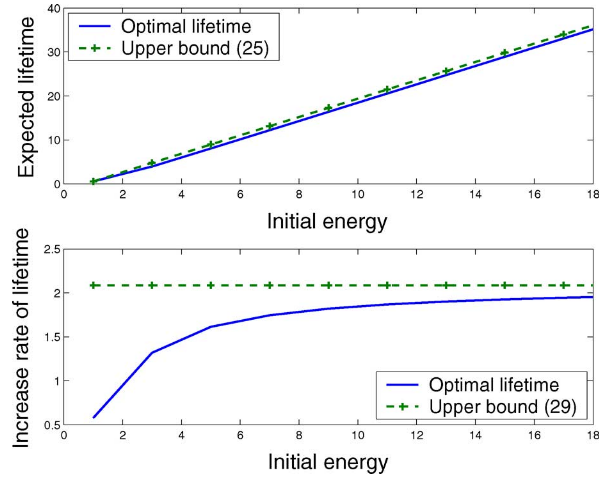


Fig. 4. Upper bound on the maximum expected lifetime. $N = 3$ (number of sensors), $N_T = 1$ (threshold on the number of dead sensors in the lifetime definition), $L = 3$ (number of power levels), and $\mathcal{W} = \{1, 2, 3\}$ (transmission energy levels). Channel distribution: $p_{W_n}(1) = p_{W_n}(3) = 1/4$, $p_{W_n}(2) = 1/2$, for all n.

channel diversity among sensors by scheduling the active sensor with the best channel realization; and 3) the pure conservative protocol which captures the diversity among sensor residual energies by scheduling the sensor with the most residual energy. As shown in [26], maximizing network lifetime requires an optimal tradeoff between two conflicting goals: minimizing the energy consumed in each data collection slot and minimizing the total unused energy left in the network when it dies. The former requires the use of CSI by favoring sensors with better channels for transmission while the latter requires the use of residual energy information (REI) by prioritizing sensors with more residual energy. The pure opportunistic and the pure conservative schemes are at the two ends of the spectrum, each focusing solely on one of the two conflicting goals. As a consequence, neither is optimal as demonstrated below.

Fig. 2 shows the expected network lifetime as a function of the initial energy. Ignoring the diversities at the physical layers, the layered approach performs the worst. The pure conservative protocol outperforms the pure opportunistic protocol when initial energy is small. This is because the benefit of exploiting channel diversity in a relatively short network lifetime is subdued by the small number of transmission power levels. Significant performance loss of the pure opportunistic and the pure conservative schemes can be seen, and the loss increases with initial energy.

In Fig. 3, we study the expected network lifetime of these scheduling protocols under different lifetime definitions by varying N_T, where N_T is the threshold on the number of dead sensors in the lifetime definition. As expected, the expected lifetime of all scheduling protocols increases with N_T. It is interesting to observe how the performance gap between the optimal and suboptimal algorithms varies with N_T. Specifically, when N_T is small, exploiting REI is crucial to network lifetime; it can balance energy consumption across sensors so that the amount of unused energy left in the network when the lifetime expires is minimized. When N_T increases, however, exploiting REI becomes less crucial. As a consequence, we see in Fig. 3 that the gap between the optimal and the pure opportunistic algorithms decreases with N_T while the gap between the optimal and the pure conservative algorithms increases with N_T. Furthermore, the pure conservative algorithm (solely focusing on REI) outperforms the pure opportunistic algorithm (solely focusing on CSI) when N_T is small. Their performances cross when N_T increases.

Fig. 4 shows the upper bound $\hat{\mathcal{L}}_m(N\mathcal{E}_0)$ given in (25) on the limiting performance and the upper bound given in (29) on the asymptotic increase rate of the limiting performance with initial energy. As shown in the upper figure, the upper bound (25) on the limiting performance is tight. Extensive simulation examples with arbitrarily chosen parameters have indicated the tightness of the upper bound. We also notice that the difference between the upper bound and the maximum expected lifetime remains almost the same as the initial energy increases. The lower figure shows that the increase rate of the limiting performance approaches the upper bound given in (29).

IV. Optimal Transmission Scheduling Using Channel Statistics

To implement the optimal scheduling algorithms using global CSI developed in Section III, all sensors have to respond to the AP at the beginning of each data collection slot for global CSI acquisition. This can result in a large implementation overhead. We are thus motivated to study sensor scheduling using channel statistics where the channel distribution $p_{\mathbf{W}}(\mathbf{w})$ of sensors is exploited and the instantaneous channel realization of only the scheduled sensor is used to adjust its transmission power level.

For Markovian fading channels, scheduling using channel statistics can be modeled as a POMDP problem,[4] and results in [37] can be applied to obtain the optimal scheduling policy. When channel fading is i.i.d. across data collection slots, however, scheduling using channel statistics can be formulated as an SSP problem in a similar fashion to scheduling using global CSI, and the results developed in Section III can be applied after modifying the state space to include only the residual energy profile $\mathbf{e}$. In this section, we show that under a further assumption of $N_T = 1$ in the lifetime definition, we can formulate the resulting problem as a shortest path multiarmed bandit process which has an indexable optimal policy [12]. As mentioned in Section I, the transmission scheduling problem turns out to be identical to the problem of playing golf with multiple balls. We also show in Section IV-E that the optimal sensor scheduling using channel statistics can be implemented in a distributed fashion via opportunistic carrier sensing.

A. SSP Multiarmed Bandit Formulation

The SSP multiarmed bandit is different from a standard discounted-cost multiarmed bandit where an infinite-horizon discounted cost is optimized. An SSP multiarmed bandit process consists of N Markov systems, each has an inevitable absorbing terminating state. In each decision interval, one system is chosen and a positive nondiscounted reward is obtained. Only the state of the chosen system evolves according to its transition probabilities; the states of other systems remain fixed. The goal is to maximize the total expected reward until any system reaches its terminating state. For sensor scheduling using channel statistics, we model the network by N Markov systems, each represents the state evolution of a sensor in the network. Below we characterize the SSP multiarmed bandit by specifying its state space, action space, transition probabilities, and reward.

1) Sensor State Space: In each data collection slot, the state of a sensor is characterized by its residual energy. Since the channel realizations are unknown, the scheduled sensor may be inactive. In this case, a failure in data collection occurs and the network lifetime terminates.[5] We introduce state t to represent this situation. The state space $\mathcal{S}$ of each sensor is thus given by

$$\mathcal{S} \triangleq \{\text{state } i = e : e \in \mathcal{E}\} \cup \{\text{state } t\}. \tag{32}$$

The set $\mathcal{S}_t$ of terminating states is defined as

$$\mathcal{S}_t \triangleq \{e : e < \varepsilon_1\} \cup \{\text{state } t\} \tag{33}$$

where the states in $\{e : e < \varepsilon_1\}$ indicate that the death of this sensor (i.e., it does not have enough energy for transmitting at the lowest power level) and state t indicates a failure in data collection. By the lifetime definition, if any sensor reaches a terminating state, the network dies.

2) Action Space: In each data collection slot, one sensor is chosen for transmission.[6] For an MDP representing a particular sensor, the action space consists of two elements: $\mathcal{A} \triangleq \{0, 1\}$. Action $a = 0$ indicates that this sensor is not scheduled for transmission. Consequently, its state will not change. Action $a = 1$ indicates that this sensor is chosen for transmission; it then transits to a new state according to the channel statistics.

3) Controlled Markovian Dynamics: If sensor n is not chosen, i.e., $a = 0$, its state remains unchanged. Otherwise, the state of sensor n transits from $i = e_n$ to j with probability $p_{ij}^{(n)}$ given by

$$p_{ij}^{(n)} = \begin{cases} \Pr\{W_n = \varepsilon_k\} 1_{[e'_n = e_n - \varepsilon_k]}, & j = e'_n \\ \Pr\{W_n > e_n\}, & j = t \end{cases} \tag{34}$$

where ε_k is the energy consumed by a sensor in transmitting a packet at the kth power level. We point out that the Markovian dynamics are determined by the marginal distributions of $\{W_n\}$; channel correlation across sensors does not affect the SSP multiarmed bandit formulation.

4) Transmission Reward: We assign a unit reward to a data collection slot in which the scheduled sensor is active. That is, if sensor n is chosen ($a = 1$), the reward $r_{ij}^{(n)}$ obtained when it transits from state i to state j is given by

$$r_{ij}^{(n)} = 1_{[j \neq t]}. \tag{35}$$

If sensor n is not chosen ($a = 0$), then the reward is zero. Since only one sensor is scheduled in each data collection slot, the total reward obtained until the network dies (i.e., any sensor reaches a terminating state) represents the network lifetime.

With the above formulation, we see that sensor transmission scheduling using channel statistics is an SSP multiarmed bandit process. It is shown in [12] that the optimal policy is given by an "indexable" strategy that chooses the sensor whose current state has the largest Gittins index (as shown in (4)). Below we give a brief description of Gittins index for completeness.

B. Gittins Index

Consider a particular sensor n and modify the MDP associated with sensor n in the following way: under action $a = 0$ (*i.e.,* the sensor is not chosen), sensor n transits to a terminating state with probability 1 and a terminating reward g is received. Our goal here is to maximize the total expected reward before sensor n reaches any terminating state given in (33).

[4]Due to the temporal correlation of fading, the channel state of the scheduled sensor in the previous data collection slot provides information on the current channel states. The network state is thus characterized by both residual energy profile $\mathbf{e}$ and transmission energy requirement $\mathbf{w}$. Since $\mathbf{w}$ is not observable, the problem becomes a POMDP.

[5]By redefining the terminating state, we can readily remove this condition in the network lifetime definition.

[6]When $N_0 > 1$ sensors need to be scheduled in each data collection slot, we have a restless bandit process. Existing results on restless bandit processes [38] may be applicable here.

The optimal expected reward $\mathcal{L}_s^*(i)$ of the modified system starting from a nonterminating state $i \in \mathcal{S}\backslash\mathcal{S}_t$ is given by the Bellman's optimality equation

$$\mathcal{L}_s^*(i) = \max\left\{g, R_n(i) + \sum_{j\in\mathcal{S}\backslash\mathcal{S}_t} p_{ij}^{(n)}\mathcal{L}_s^*(j)\right\} \tag{36}$$

where $R_n(i)$ is the expected reward associated with state $i = e_n$ of sensor n under action $a = 1$

$$R_n(i) = R_n[e_n] = \sum_{j\in\mathcal{S}} r_{ij}^{(n)} p_{ij}^{(n)} = \Pr\{W_n \leq e_n\}. \tag{37}$$

Note that the optimal reward $\mathcal{L}_s^*(i) = 0$ for all terminating states $i \in \mathcal{S}_t$ and $\mathcal{L}_s^*(i) \geq g$ for all nonterminating states $i \in \mathcal{S}\backslash\mathcal{S}_t$. The Gittins index $\gamma_n(i)$ of state i is defined as the smallest value of terminating reward g at which action $a = 0$ is optimal [12], as follows:

$$\gamma_n(i) = \min_g \{g : \mathcal{L}_s^*(i) = g\}. \tag{38}$$

It has been shown in [12] that the optimal policy for the modified Markov system is to choose $a = 1$ if the Gittins index $\gamma_n(i)$ of the current state i is greater than the terminating reward g and to choose $a = 0$ otherwise. Hence, the Gittins index of state i is also the value of g at which we are indifferent between action $a = 0$ and action $a = 1$.

C. Indexable Optimal Policy and a Closed-Form Expression for Gittins Index

Since the optimal policy for an SSP multiarmed bandit process is to choose the system with the largest Gittins index, the calculation of the Gittins index is the key to optimal transmission scheduling using channel statistics. For a general multiarmed bandit problem, closed-form expression for the Gittins index does not exist, and the computational complexity of the most efficient algorithm (up to date) for calculating the Gittins index is cubic in the number of nonterminating states [13]. For the scheduling problem, this means that the complexity is on the order of $O(NM^3)$, where $M = |\mathcal{E}|$ is the number of possible residual energies. Fortunately, the rich structure of the sensor scheduling problem enables us to derive a closed-form expression for the Gittins index and hence reduce the computational complexity to linear in both network size N and the number M of possible residual energies.

Theorem 1: For sensor scheduling using channel statistics, the Gittins index $\gamma_n(i)$ of state $i = e_n$ associated with sensor n is given by

$$\gamma_n(i) = \gamma_n[e_n] = \frac{\Pr\{W_n \leq e_n\}}{\Pr\{W_n > e_n - \varepsilon_1\}} \tag{39}$$

where ε_1 is energy consumed by a sensor in transmitting a packet at the lowest power level.[7]

Proof: We provide two proofs of Theorem 1 in Appendix A: a constructive proof which directly calculates the Gittins index based on the algorithm developed in [13], and a nonconstructive proof, suggested by an anonymous reviewer, which uses the interchange argument [39] to show the optimality of the index defined as (39). ■

We see from (39) that the computation of the Gittins index only requires the marginal instead of joint distribution of the channel states. We show in Corollary 1 that when channel fading is i.i.d. across sensors, the optimal scheduling algorithm can be implemented even when the channel distribution is unknown.

Corollary 1: When channel fading is identically (but not necessarily independently) distributed across sensors, the optimal sensor scheduling using channel statistics reduces to the pure conservative transmission protocol that chooses the sensor with the most residual energy in each data collection slot.

Proof: Since channel fading is identically distributed, the Gittins index $\gamma_n(i) = \gamma_m(i)$ for all $1 \leq n, m \leq N$. From (39), we find that Gittins index $\gamma_n(i)$ of state $i = e$ increases with the residual energy e of sensor n and Corollary 1 follows. ■

D. Maximum Lifetime Using Channel Statistics

We now obtain the maximum expected network lifetime that can be achieved by a scheduling algorithm using channel statistics. Applying the Bellman's equation, we obtain the maximum expected lifetime $\mathcal{L}^*[\mathbf{e}]$ starting with a network energy profile $\mathbf{e}$ as (40), shown at the bottom of the page. Due to the acyclic graph structure of the underlying Markov chain as shown in Proposition 1, the maximum expected lifetime can be obtained in one iteration by calculating (40) in an increasing order of total energy in the network. Applying the indexable policy with Theorem 1 yields

$$\mathcal{L}^*[\mathbf{e}] = \sum_{k=1}^{L} \Pr\{W_a = \varepsilon_k\} 1_{[e_a - \varepsilon_k \geq 0]} \{1 + \mathcal{L}^*[\mathbf{e} - \mathbf{I}_a\varepsilon_k]\} \tag{41}$$

[7]For sufficiently large residual energy $e_n \geq \varepsilon_1 + \varepsilon_L$, the probability $\Pr\{W_n > e_n - \varepsilon_1\} = 0$, resulting in an infinite Gittins index. At the early stage of the network lifetime, the Gittins indexes of several sensors can be infinite. A random selection among them suffices.

$$\begin{aligned}\mathcal{L}^*[\mathbf{e}] &= \max_n \left\{R_n[e_n] + \sum_{e'_n\in\mathcal{S}\backslash\mathcal{S}_t} p_{e_n e'_n}^{(n)} \mathcal{L}^*[\mathbf{e} - \mathbf{I}_n(e_n - e'_n)]\right\} \\ &= \max_n \left\{R_n[e_n] + \sum_{k=1}^{L} \Pr\{W_n = \varepsilon_k\} \times 1_{[e_n - \varepsilon_k \geq 0]} \mathcal{L}^*[\mathbf{e} - \mathbf{I}_n\varepsilon_k]\right\}. \end{aligned} \tag{40}$$

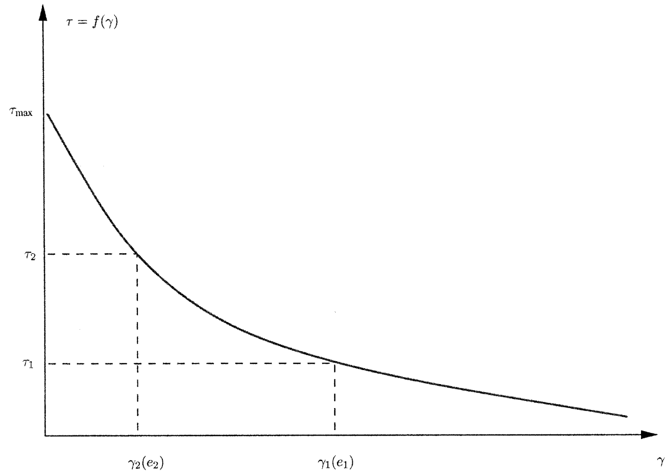


Fig. 5. Opportunistic carrier sensing.

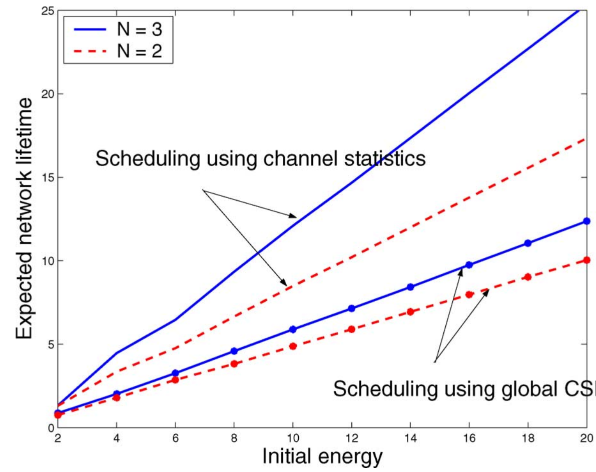


Fig. 6. Network lifetime comparison with channel acquisition cost. $N = 2, 3$ (number of sensors), $N_T = 1$ (threshold on the number of dead sensors in the lifetime definition), $L = 3$ (number of power levels), $\mathcal{W} = \{1, 2, 3\}$ (transmission energy levels), and $v = 1$ (channel acquisition cost per sensor). Channel distribution: $p_{W_n}(1) = p_{W_n}(2) = 1/4$, $p_{W_n}(3) = 1/2$, for all n.

where $a = \arg\max_n\{\gamma_n[e_n]\}$ is the index of the sensor with the largest Gittins index. Hence, the computational complexity of the maximum expected lifetime is linear in the network size N.

E. Implementation and Overhead

Compared with the optimal scheduling using global CSI developed in Section III, the optimal scheduling using channel statistics has not only significantly reduced computational complexity but also lowered implementation overhead. Specifically, at the beginning of each data collection slot, the AP chooses a sensor according to the network energy profile $\mathbf{e}$ and broadcasts a beacon signal and the ID of the chosen sensor. The scheduled sensor then estimates its channel state using the beacon signal and transmits its measurements at a power level determined by its current channel state. To help the AP keep track of the network energy profile $\mathbf{e}$, the scheduled sensor can piggyback its residual energy level to the data packet. We see that in each data collection slot, scheduling using channel statistics only requires the scheduled sensor to estimate its channel; other sensors do not need to estimate their channels or respond to the AP.

Sensor scheduling using channel statistics can also be implemented in a distributed way via opportunistic carrier sensing [30]. Specifically, in each data collection slot, every sensor maps its Gittins index to a backoff time τ based on a common decreasing function $f(\gamma)$ as shown in Fig. 5. Hence, a sensor with larger Gittins index has a shorter backoff time. A sensor will transmit with its chosen backoff delay if and only if no one transmits before its backoff time expires. If the channel propagation delay among sensors is negligible, this opportunistic carrier sensing scheme ensures that the sensor with the largest Gittins index seizes the channel and transmits. We point out that when the propagation delay is significant, $f(\gamma)$ needs to be designed judiciously (see [30]). Furthermore, a random backoff strategy can be overlayed with the opportunistic carrier sensing to avoid collision among sensors with the same Gittins index.

F. Numerical Examples

Recall that the optimal scheduling using channel statistics reduces to the pure conservative protocol when channel fading is i.i.d. across sensors. Fig. 2 thus demonstrates the significant performance gain of the optimal scheduling using global CSI over the optimal scheduling using channel statistics (the pure conservative scheme) when the channel acquisition cost is negligible. We now compare the performance of these two approaches when the channel acquisition cost v is high. As shown in Fig. 6, when the energy consumed in channel acquisition is comparable to that in packet transmission, scheduling using channel statistics outperforms scheduling using global CSI. The performance loss of scheduling using global CSI increases with network size N and initial energy $\mathcal{E}_0$ since more energy is wasted in global CSI acquisition as the network lifetime increases. Hence, the optimal scheduling using channel statistics is preferred when the channel acquisition cost is high. We point out the tradeoff between using global CSI and using channel statistics depends on the channel acquisition cost.

In Fig. 7, we compare the performance of the pure conservative protocol and the optimal scheduling using channel statistics when channel fading is not i.i.d across sensors. We consider the case of two sensors, each with two transmission energy levels: $\varepsilon_1 = 1$ and $\varepsilon_2 = 3$. The channel distributions are given by $p_{W_1}(1) = 0.4$ and $p_{W_1}(3) = 0.6$ for sensor 1, and $p_{W_2}(1) = p$ and $p_{W_2}(3) = 1 - p$ for sensor 2. As p increases, the channel of sensor 2 improves and so does the expected lifetime. When $p = 0.4$, i.e., the channel is i.i.d. across sensors, the pure conservative protocol has the optimal lifetime performance as dictated by Corollary 1. Clearly, when the channel is not identically distributed, the pure conservative protocol performs worse than the optimal scheduling using channel statistics, but the performance degradation is small.

In Fig. 8, we investigate the performance of the indexable policy under a Markovian fading model as compared with the optimal performance obtained via POMDP formulation. We consider two power levels $\mathcal{W} = \{1, 2\}$. The transition matrix of the transmission energy requirements for each sensor is given by

$$\begin{pmatrix} p & 1-p \\ 1-p & p \end{pmatrix} \qquad (42)$$

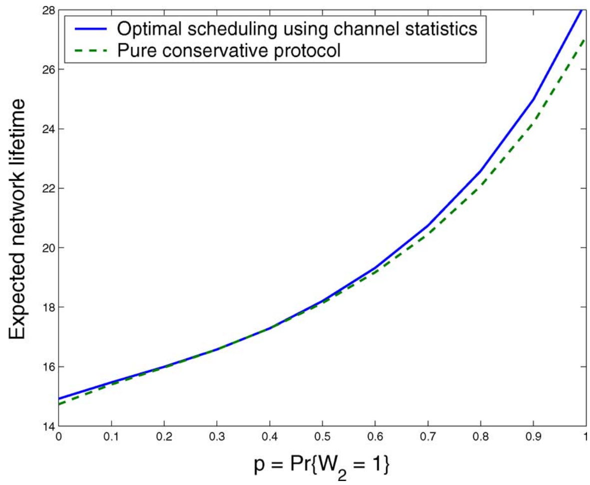


Fig. 7. Expected lifetime of the pure conservative protocol. $N = 2$ (number of sensors), $N_T = 1$ (threshold on the number of dead sensors in the lifetime definition), $L = 2$ (number of power levels), and $\mathcal{W} = \{1, 3\}$ (transmission energy levels). Channel distribution of sensor 1: $p_{W_1}(1) = 0.4$, $p_{W_1}(3) = 0.6$. Channel distribution of sensor 2: $p_{W_2}(1) = p$, $p_{W_2}(3) = 1 - p$.

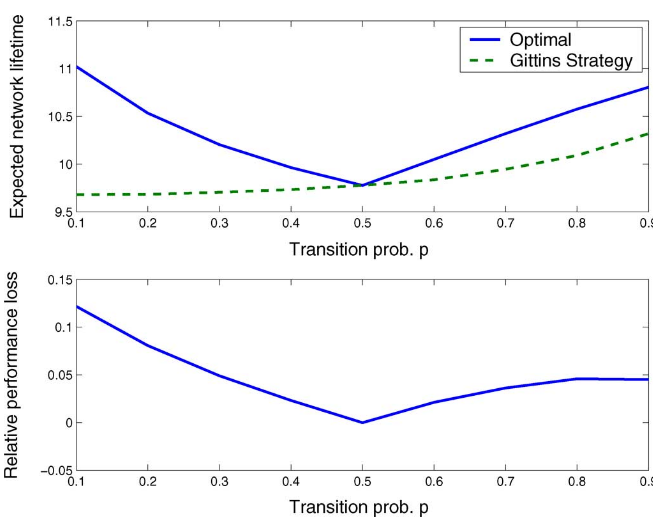


Fig. 8. Performance of the indexable policy in Markovian fading channels. $N = 2$ (number of sensors) , $\mathcal{E}_0 = 8$ (sensor initial energy), $L = 2$ (number of power levels), and $\mathcal{W} = \{1, 2\}$ (transmission energy levels).

which is determined by the underlying Markovian channel model. As expected, the indexable policy is optimal when channels are independent across data collections, i.e., $p = 0.5$. Without exploiting the temporal correlation of channel states, the indexable policy is suboptimal under a general Markovian fading model. However, its performance degrades gracefully with channel correlation, which is given by $1 - 2p$ as shown in the lower plot of Fig. 8.

V. Asymptotically Optimal Transmission Scheduling Using Local CSI

In this section, we consider sensor scheduling using local CSI. Specifically, each sensor, based on its own channel state and residual energy, needs to determine whether to transmit in a data collection slot. We formulate this problem by defining an energy-efficiency index γ_i associated with sensor i as

$$\gamma_i = g(e_i, w_i), \quad \text{where } g : \mathcal{E} \times \mathcal{W} \to \mathbb{R}. \tag{43}$$

In a particular data collection slot, a scheduling algorithm chooses the sensor with the largest energy-efficiency index,[8] i.e.,

$$a = \arg\max\{\gamma_1, \ldots, \gamma_N\} \tag{44}$$

where a is the index of the chosen sensor. The sensor scheduling problem is thus reduced to the design of the energy efficiency index $g(\cdot)$.

In [6], we have proposed a scheduling algorithm using local CSI. Referred to as the dynamic protocol for lifetime maximization (DPLM), this algorithm employs an energy-efficiency index defined as

$$\gamma_n = \frac{e_n}{w_n}. \tag{45}$$

In other words, it schedules the sensor whose current channel realization demands the least portion of its residual energy for transmission. It is shown in [27] that DPLM is asymptotically optimal: its relative performance loss as compared to the limiting performance achieved by the optimal policy using global CSI diminishes with the initial energy. Our goal here is to study its computational complexity, implementation overhead, and lifetime performance when initial energy is finite.

Since the energy-efficiency index defined in (45) is simply the ratio between the residual energy and the transmission energy requirement, DPLM has much lower computational complexity as compared to the optimal scheduling using global CSI.

Next, we consider the implementation overhead. Similar to scheduling using channel statistics, DPLM can be implemented in a distributed way via opportunistic carrier sensing. Specifically, at the beginning of each data collection slot, every sensor receives the beacon signal broadcasted by the AP and estimates its own channel state using the beacon signal. Sensors can then compute their energy-efficiency indexes and carry out opportunistic carrier sensing using backoff delays determined by their energy-efficiency indexes (see Fig. 5). Since the energy consumed in listening to the beacon signal is much lower than in transmitting through a fading channel, DPLM has lower implementation overhead than the optimal scheduling using global CSI. It, however, may have larger implementation overhead than scheduling algorithms using channel statistics, where only the scheduled sensor needs to estimate its channel state. Note that DPLM does not require the knowledge of channel distributions while both scheduling using global CSI and scheduling using channel statistics do. Hence, considering the potential overhead in estimating the channel distribution, one may further favor DPLM.

Finally, we compare in Fig. 9 the performance of DPLM with the limiting performance achieved by the optimal scheduling using global CSI. Since the upper bound given in (25) is

[8]When $N_0 > 1$ sensors need to be scheduled in each data collection slot, the AP can choose sensors with the largest N_0 energy efficiency indices. Scheduling multiple sensors via opportunistic carrier sensing can be found in [30]. An alternative implementation is to schedule these N_0 sensors sequentially.

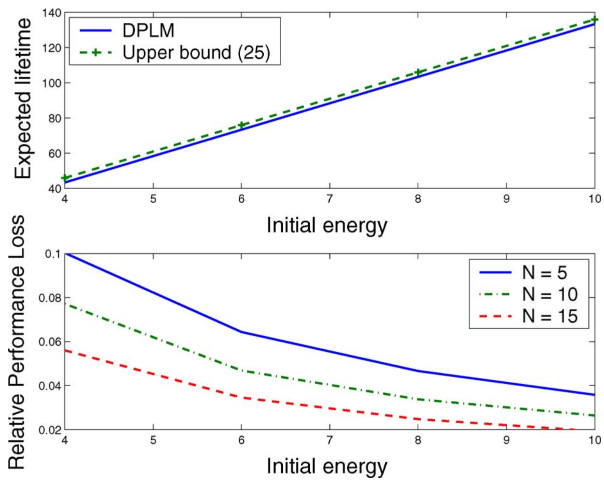


Fig. 9. Expected lifetime of DPLM. $N = 5, 10, 15$ (number of sensors), $N_T = 1$ (threshold on the number of dead sensors in the lifetime definition), $L = 2$ (number of power levels), and $\mathcal{W} = \{1, 2\}$ (transmission energy levels), $p_{W_n}(1) = 0.4, p_{W_n}(2) = 0.6$, for all n.

tight and has lower computational complexity, we use the upper bound as a benchmark. In the upper figure, we plot the expected lifetime of DPLM for $N = 15$. We can see that DPLM offers nearly optimal performance even when the initial energy is small. In the lower figure, we plot the relative performance loss of DPLM as compared to the upper bound for different network sizes. We find that the relative performance loss diminishes with the initial energy. These observations confirm the asymptotic optimality of DPLM that has been shown analytically in [27]. A perhaps more interesting observation is that the relative performance loss of DPLM decreases with network size N for a fixed initial energy $\mathcal{E}_0$. Combined with the distributed implementation of DPLM, this result suggests the optimality and feasibility of applying DPLM to large sensor networks. Furthermore, the small gap between the performance of DPLM and the lifetime upper bound shown in Fig. 9 provides another example for the tightness of the bound.

VI. Conclusion

Under the metric of network lifetime, we considered sensor transmission scheduling that exploits channel information with different levels of resolution, namely, global CSI, channel statistics, and local CSI. The optimal transmission scheduling problem was formulated as an MDP. By exploiting the structure of the problem, we proposed computationally tractable algorithms for computing the optimal scheduling policy. By studying the tradeoff across computational complexity, implementation overhead, and performance, we demonstrate the pros and cons and the different domains of applications of these three approaches to transmission scheduling for lifetime maximization. It is important to note that simply choosing the sensor with the best channel (the opportunistic scheduling) or the most residual energy (the conservative scheduling) in each data collection slot does not optimize the network lifetime. Maximizing network lifetime requires an optimal tradeoff between CSI and REI.

Appendix
Proof of Theorem 1

Proof: The first proof of Theorem 1 is based on the algorithm developed in [13] for calculating the Gittins indexes of a general SSP multiarmed bandit process. Consider the modified Markov system associated with sensor n, which has been introduced in Section IV-B.

Step 0: *Initiate* $\mathcal{U} = \phi$.

Step 1: *Calculate the largest Gittins index in the set* $\mathcal{S}\backslash\mathcal{U}$.

Suppose that state e^* has the largest Gittins index $\gamma_n[e^*]$ and the modified Markov system has a terminating reward $g = \gamma_n[e^*]$. Since the optimal policy of the modified Markov system is to choose action $a = 0$ whenever the Gittins index of the current state is smaller than the terminating reward g, the optimal expected reward starting from any nonterminating state $e \in \mathcal{S}\backslash\mathcal{S}_t\backslash\mathcal{U}$ is $\mathcal{L}[e] = g = \gamma_n[e^*]$. Since the Gittins index is also defined as the value of g at which we are indifferent between $a = 0$ and $a = 1$, we obtain the optimal total expected reward starting from state e^* as

$$\mathcal{L}[e^*] = \underbrace{\gamma_n[e^*]}_{a=0 \text{ is chosen}} = \underbrace{R_n[e^*] + \sum_{e\in\mathcal{S}\backslash\mathcal{U}} p_{e^*e}^{(n)}\mathcal{L}[e]}_{a=1 \text{ is chosen}}$$
$$= R_n[e^*] + F_n[e^*]\gamma_n[e^*] \tag{46}$$

where $F_n[e^*] = \Pr\{e^* - w_n \geq \varepsilon_1\} = \Pr\{w_n \leq e^* - \varepsilon_1\}$ is the probability that state e^* does not transit to a terminating state in one step. Hence, the largest Gittins index in $\mathcal{S}\backslash\mathcal{U}$ is given by

$$\gamma_n[e^*] = \frac{R_n[e^*]}{1 - F_n[e^*]} = \frac{\Pr\{W_n \leq e^*\}}{\Pr\{W_n > e^* - \varepsilon_1\}}. \tag{47}$$

Step 2: *Determine the state* e^* *that has the largest Gittins index and remove it from the modified Markov system.*

From (47), we see that the state e^* that has the largest Gittins index maximizes $R_n[e]/(1 - F_n[e])$ over all $e \in \mathcal{S}\backslash\mathcal{U}$, i.e.,

$$e^* \in \arg\max_{e\in\mathcal{S}\backslash\mathcal{U}} \frac{R_n[e]}{1 - F_n[e]}. \tag{48}$$

If there is more than one state that achieves the largest Gittins index (47), we will choose the one with the most residual energy. Since $R_n[e]/(1 - F_n[e]) = \Pr\{W_n \leq e\}/\Pr\{W_n > e - \varepsilon_1\}$ increases with e, the state e^* that has the largest Gittins index also has the most residual energy in $\mathcal{S}\backslash\mathcal{U}$. Hence, the state e^* is not reachable from any other state $e \in \mathcal{S}\backslash\mathcal{U}$. We can thus remove state e^* from the system without changing the immediate rewards and the transition probabilities of the remaining states.

Step 3: *Let* $\mathcal{U} = \mathcal{U} \cup \{e^*\}$. *Go to Step 1 until* $\mathcal{U} = \mathcal{S}\backslash\mathcal{S}_t$.

Following the above procedure, we find that the Gittins index of every state $e \in \mathcal{S}\backslash\mathcal{S}_t$ can be computed using (47), which is the same as (39). ■

Proof: The second proof is based on the interchange argument [39]. Let π^* be the policy that schedules the sensor with the largest index defined as (39) in each data collection slot. To prove the optimality of π^*, it suffices to show that starting from any given network energy profile $\mathbf{E} = \mathbf{e}$, policy $\pi^{(0)}$ that first chooses a sensor that does not have the largest index and then proceeds according to π^* performs no better than policy π^*, i.e., $\mathcal{L}_{\pi^{(0)}}[\mathbf{e}] \leq \mathcal{L}_{\pi^*}[\mathbf{e}]$ for any $\mathbf{e}$. To prove this, we construct a sequence of policies $\{\pi^{(k)}\}$, where $\pi^{(k)}$ starts by scheduling the sensor with the largest index and proceeds according to $\pi^{(k-1)}$, and then show that $\mathcal{L}_{\pi^{(k)}}[\mathbf{e}]$ monotonically increases to the limit $\mathcal{L}_{\pi^*}[\mathbf{e}]$. Specifically, since $\pi^{(k)}$ agrees with π^* for at least the first k data collection slots and termination is inevitable in SSP multiarmed bandit processes, we have $\pi^{(k)} \to \pi^*$ and hence $\mathcal{L}_{\pi^{(k)}}[\mathbf{e}] \to \mathcal{L}_{\pi^*}[\mathbf{e}]$ for any $\mathbf{e}$ as $k \to \infty$. Since the index given by (39) increases with the sensor residual energy, we find that given $\mathbf{e}$, policy $\pi^{(0)}$ will first choose a sensor $n \neq n^* \triangleq \arg\max_n(\gamma_n(e_n))$ and then sensor n^* while policy $\pi^{(1)}$ will first schedule sensor n^* and then sensor n. After these two data collection slots, both $\pi^{(0)}$ and $\pi^{(1)}$ follow policy π^*. Noting that channels are i.i.d. across data collection slots, we can show that $\mathcal{L}_{\pi^{(0)}}[\mathbf{e}] \leq \mathcal{L}_{\pi^{(1)}}[\mathbf{e}]$ for any $\mathbf{e}$. By induction, we can further show that $\mathcal{L}_{\pi^{(k)}}[\mathbf{e}]$ monotonically increases with k for any fixed $\mathbf{e}$, which completes the proof. ■

ACKNOWLEDGMENT

The authors would like to thank anonymous reviewers for their invaluable comments and suggestions. In particular, they thank a reviewer for pointing out an alternative to the proof of Theorem 1.

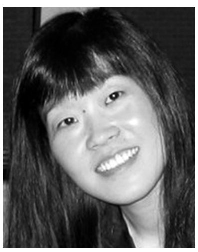
REFERENCES

[1] X. Qin and R. Berry, "Exploiting multiuser diversity for medium access control in wireless networks," in *Proc. IEEE INFOCOM*, San Francisco, CA, Mar. – Apr. 2003, vol. 2, pp. 1084–1094.
[2] P. Venkatasubramaniam, S. Adireddy, and L. Tong, "Sensor networks with mobile access: Optimal random access and coding," *IEEE J. Sel. Areas Commun. (Special Issue on Sensor Networks)*, vol. 22, pp. 1058–1068, Aug. 2004.
[3] Z. Yang, M. Dong, L. Tong, and B. Sadler, "On the MAC for optimal information retrieval pattern in sensor networks with mobile access," in *Proc. IEEE MILCOM*, Monterey, CA, Oct.-Nov. 2004, vol. 1, pp. 232–237.
[4] Y. Sung, L. Tong, and H. Poor, "A large deviations approach to sensor scheduling for detection of correlated random fields," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, Philadelphia, PA, Mar. 2005, vol. 3, pp. 649–652.
[5] Y. Chen and Q. Zhao, "Maximizing the lifetime of sensor network using local information on channel state and residual energy," in *Proc. 39th Conf. Information Science Systems*, Mar. 2005.
[6] ——, "Distributed transmission protocol for lifetime maximization in sensor networks," in *Proc. IEEE 6th Workshop Signal Processing Advances in Wireless Communications (SPAWC)*, Jun. 2005, pp. 895–899.
[7] L. Meier, III, J. Peschon, and R. Dressler, "Optimal control of measurement subsystems," *IEEE Trans. Autom. Control*, vol. 12, pp. 528–536, Oct. 1967.
[8] J. Baras and A. Bensoussan, "Optimal sensor scheduling in nonlinear filtering of diffusion processes," *SIAM J. Control Optim.*, vol. 27, pp. 786–813, Jul. 1989.
[9] D. A. Castanon, "Approximate dynamic programming for sensor management," in *Proc. 36th IEEE Conf. Decision Control*, Dec. 1997, pp. 1202–1207.
[10] V. Krishnamurthy, "Algorithms for optimal scheduling and management of hidden Markov model sensors," *IEEE Trans. Signal Process.*, vol. 50, no. 6, pp. 1382–1397, Jun. 2002.
[11] J. C. Gittins, *Multi-Armed Bandit Allocation Indices*. New York: Wiley, 1989.
[12] I. Dumitriu, P. Tetali, and P. Winkler, "On playing golf with two balls," *SIAM J. Discrete Math.*, vol. 16, pp. 604–615, Jul. 2003.
[13] A. -K. Katta and J. Sethuraman, "A note on bandits with a twist," *SIAM J. Discrete Math.*, vol. 18, pp. 110–113, Jul. 2004.
[14] V. Krishnamurthy, "POMDP multi-armed bandit formulation for energy minimization in sensor networks," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, Mar. 2005, pp. 793–796.
[15] ——, "Decentralized emission management for low probability of intercept sensor platforms in network centric warfare," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 41, no. 1, pp. 133–152, Jan. 2005.
[16] V. V. Veeravalli, J. A. Fuemmeler, and A. Visvanathan, "Tracking with sleepy sensors," in *Proc. Asilomar Conf. Signals, Systems, Computers*, Monterey, CA, Oct.–Nov. 2005, pp. 1546–1550.
[17] K. Kar, A. Krishnamurthy, and N. Jaggi, "Dynamic node activation in networks of rechargeable sensors," in *Proc. INFOCOM*, Mar. 2005, vol. 3, pp. 1997–2007.
[18] A. Logothetis and A. Isaksson, "On sensor scheduling via information theoretic criteria," in *Proc. Amer. Control Conf.*, San Diego, CA, Jun. 1999, pp. 2402–2406.
[19] J. S. Evans and V. Krishnamurthy, "Optimal sensor scheduling for hidden Markov model state estimation," *Int. J. Control*, vol. 74, no. 18, pp. 1737–1742, Dec. 2001.
[20] A. S. Chhetri, D. Morrell, and A. Papandreou-Suppappola, "Scheduling multiple sensors using particle filters in target tracking," in *Proc. 2003 IEEE Workshop Statistical Signal Processing*, Sep.–Oct. 2003, pp. 549–552.
[21] M. Kalandros and L. Y. Pao, "Covariance control for multisensor systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 38, pp. 1138–1157, Oct. 2002.
[22] Y. He and E. K. P. Chong, "Sensor scheduling for target tracking in sensor networks," in *Proc. 43rd IEEE Conf. Decision Control*, Dec. 2004, pp. 743–748.
[23] R. Knopp and P. Humblet, "Information capacity and power control in single cell multi-user communications," in *Proc. Int. Conf. Communications*, Seattle, WA, Jun. 1995, pp. 331–335.
[24] R. Madan, S. Cui, S. Lall, and A. Goldsmith, "Cross-layer design for lifetime maximization in interference-limited wireless sensor networks," in *Proc. IEEE INFOCOM 2005*, Mar. 2005, vol. 3, pp. 1964–1975.
[25] J. -H. Chang and L. Tassiulas, "Maximum lifetime routing in wireless sensor networks," *IEEE/ACM Trans. Netw.*, vol. 12, pp. 609–619, Aug. 2004.
[26] Y. Chen and Q. Zhao, "On the lifetime of wireless sensor networks," *IEEE Commun. Lett.*, vol. 9, pp. 976–978, Nov. 2005.
[27] ——, "An integrated approach to energy-aware medium access for wireless sensor networks," *IEEE Trans. Signal Process.*, submitted for publication.
[28] C. H. Papadimitrou and J. N. Tsitsiklis, "The complexity of Markov decision processes," *Math. Oper. Res.*, vol. 12, no. 3, pp. 441–450, 1987.
[29] D. Bertsimas and J. Nino-Mora, "Conservation laws, extended polymatroids and multiarmed bandit problems; a polyhedral approach to indexable systems," *Math. Oper. Res.*, vol. 21, no. 2, pp. 257–305, 1996.
[30] Q. Zhao and L. Tong, "Opportunistic carrier sensing for energy efficient information retrieval in sensor networks," *EURASIP J. Wireless Commun. Netw.*, no. 2, pp. 231–241, 2005.
[31] B. Chen, R. Jiang, T. Kasetkasem, and P. K. Varshney, "Fusion of decisions transmitted over fading channels in wireless sensor networks," in *Proc. 36th Asilomar Conf. Signals, Systems, Computers*, Nov. 2002, vol. 2, pp. 1184–1188.
[32] K. Liu and A. M. Sayed, "Optimal distributed detection strategies for wireless sensor networks," presented at the 42nd Annu. Allerton Conf. Commununications, Control, Computing, Monticello, IL, Oct. 2004.
[33] G. Mergen and L. Tong, "Type based estimation over multiaccess channels," *IEEE Trans. Signal Process.*, vol. 54, no. 2, pp. 613–626, Feb. 2006.
[34] Q. Dong, "Maximizing system lifetime in wireless sensor networks," in *Proc. 4h Int. Symp. Information Processing in Sensor Networks*, Apr. 2005, pp. 13–19.
[35] D. P. Bertsekas, *Dynamic Programming and Optimal Control*. Belmont, MA: Athena Scientific, 1995, vol. 1.
[36] ——, *Dynamic Programming and Optimal Control*. Belmont, MA: Athena Scientific, 1995, vol. 2.
[37] R. Smallwood and E. Sondik, "The optimal control of partially ovservable Markov processes over a finite horizon," *Oper. Res.*, pp. 1071–1088, 1971.
[38] P. Whittle, "Restless bandits: Activity allocation in a changing world," *J. Appl. Probab.*, vol. A25, pp. 287–298, 1988.
[39] J. C. Gittins and D. M. Jones, "A dynamic allocation index for the sequential design of experiments," in *Progress in Statistics, European Meeting of Statisticians*, North-Holland, 1974, vol. 1, pp. 241–266.

Yunxia Chen (S'03–M'04) received the B.Sc. degree from Shanghai Jiaotong University, Shanghai, China, in 1998 and the M.Sc. degree from the University of Alberta, Edmonton, AB, Canada, in 2004. She is currently working towards the Ph.D. degree in the Department of Electrical and Computer Engineering, University of California, Davis.

Her current research interests are in resource-constrained signal processing, communications, and networking. She is also interested in diversity techniques and fading channels, communication theory, and MIMO systems.

Ms. Chen received the Best Student Paper Award at the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) in May 2006 and First Place in the Student Paper Contest at the IEEE Asilomar Conference on Signals, Systems, and Computers, October 2006.



Qing Zhao (S'96–M'05) received the Ph.D. degree in electrical engineering from Cornell University, Ithaca, NY in 2001.

From 2001 to 2003, she was a Communication System Engineer with Aware, Inc., Bedford, MA. She returned to academe in 2003 as a Postdoctoral Research Associate with the School of Electrical and Computer Engineering at Cornell University. In 2004, she joined the Department of Electrical and Computer Engineering at the University of California, Davis, where she is currently an Assistant Professor. She is a co-editor of the book *Wireless Sensor Networks: Signal Processing and Communications Perspectives* (New York: Wiley, 2007). Her research interests are in the general area of signal processing, communication systems, wireless networking, and information theory.

Dr. Zhao received the 2000 IEEE Signal Processing Society Young Author Best Paper Award. She is an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING and an elected member of the Signal Processing for Communications technical committee of the IEEE Signal Processing Society. She is the lead Guest Editor for the *IEEE Signal Processing Magazine* Special Issue on Resource-Constrained Signal Processing, Communications, and Networking.

Vikram Krishnamurthy (S'90–M'91–SM'99–F'05) was born in 1966. He received the Bachelor's degree from the University of Auckland, New Zealand, in 1988 and the Ph.D. degree from the Australian National University, Canberra, in 1992.

Previously, he was a chaired Professor at the Department of Electrical and Electronic Engineering, University of Melbourne, Australia. Since 2002, he has been a Professor and Canada Research Chair at the Department of Electrical Engineering, University of British Columbia, Vancouver, BC, Canada. His research interests span several areas, including ion channels and nanobiology, stochastic optimization, scheduling and control, statistical signal processing, and wireless telecommunications.

Dr. Krishnamurthy has served as Associate Editor for IEEE TRANSACTIONS ON SIGNAL PROCESSING, IEEE TRANSACTIONS ON AEROSPACE AND ELECTRONIC SYSTEMS, IEEE TRANSACTIONS ON NANOBIOSCIENCE, IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS II, *Systems and Control Letters*, and the *European Journal of Applied Signal Processing*. He was Guest Editor of a special issue of IEEE TRANSACTIONS ON NANOBIOSCIENCE on bio-nanotubes, March 2005.

Dejan Djonin received the B.Sc. and M.Sc. degrees from the University of Belgrade, Belgrade, Serbia, in 1996 and 1999 and the Ph.D. degree from the University of Victoria, Victoria, BC, Canada, in 2003.

From 1998 to 2000, he was with the Department of Telecommunications, Institute Mihajlo Pupin, Belgrade, where he worked on the development of an antenna array processing system. From 2005 to 2006 he has held the NSERC Postdoctoral Fellowship at the Department of Electrical and Computer Engineering, the University of British Columbia, Vancouver, BC, Canada. He is currently with Dypative Inc., Vancouver, BC, Canada. His research interests span applications of control theory in multimedia communication systems, wireless communications, information theory, and machine learning.

Dr. Djonin has served as an Assistant Program Chair of the Wireless Communications and Networking Conference (WCNC) 2004 and as a Technical Program Committee Member of the International Conference of Communications (ICC) 2005 and the IEEE GLOBECOM 2003 Conference. In 2002, he was awarded the Outstanding Paper Award for Young Researchers at the International Symposium on Information Theory and its Application (ISITA) Conference.