Relay Scheduling for Cooperative Communications in Sensor Networks with Energy Harvesting

Huijiang Li, Neeraj Jaggi and Biplab Sikdar

***Abstract*— This paper considers wireless sensor networks (WSNs) with energy harvesting and cooperative communications and develops energy efficient scheduling strategies for such networks. In order to maximize the long-term utility of the network, the scheduling problem considered in this paper addresses the following question: given an estimate of the current network state, should a source transmit its data directly to the destination or use a relay to help with the transmission? We first develop an upper bound on the performance of any arbitrary scheduler. Next, the optimal scheduling problem is formulated and solved as a Markov Decision Process (MDP), assuming that complete state information about the relays is available at the source nodes. We then relax the assumption of the availability of full state information, and formulate the scheduling problem as a Partially Observable Markov Decision Process (POMDP) and show that it can be decomposed into an equivalent MDP problem. Simulation results are used to show the performance of the schedulers.**

I. Introduction

Wireless sensor networks have a wide range of applications in various fields but are typically limited by the size of the battery and the power it can store. Existing research has addressed this problem in two ways: (a) development of energy efficient communication and networking protocols for WSNs and (b) improved battery technologies. In terms of communication technologies, existing research has shown that cooperative diversity gains can be achieved in distributed networks where nodes help each other by relaying transmissions [1], resulting in either higher network capacity or lower energy consumption with the same capacity. On the other hand, in the field of battery technologies, energy harvesting or energy scavenging has become a promising and feasible approach to address the energy supply problem [2], [3]. Exploiting both energy harvesting and cooperative communications is thus a promising future direction for developing energy efficient sensor nodes. However, to achieve the full energy saving potential of such sensors, fundamental scheduling issues related to the usage of cooperative communications as a function of the system state need to be addressed, and this is the focus of this paper.

The cooperative communication considered in this paper is the discrete memoryless three-terminal relay network developed in the landmark papers by van der Meulen [4], and Cover and El Gamal [5]. For such networks, the capacity, the strategies on the relay, energy efficiency and distribution among the network, have been the focus of intensive research [6], [7], [8]. The most common cooperation protocols are amplify-and-forward (AF) and decode-and-forward (DF). With AF, the relay node simply amplifies the source's transmission and retransmits it to the destination; with the DF strategy, the relay node receives and decodes the source's transmission, then re-encodes and transmits it to the destination. The destination then combines the reception from the sender and the reception from the relay to decode the data.

This paper considers sensor networks with energy harvesting capability and addresses the problem of scheduling cooperative, relay based communications. We consider a time-slotted source-relay-destination system, where a sensor (the source) has the option to have another sensor (the relay) help to transmit its data to the destination. All sensor nodes under consideration are equipped with energy harvesting capability. From an energy efficiency perspective, the source may achieve the same bit error rate (BER) for a lower transmission power if it uses a relay, as compared to a direct transmission. However, this increases the power consumed by the relay and as a result, the relay sensor may not have energy to report its own data in the future. At any given instant of time, the problem of interest is for the source to determine whether to transmit data on its own or cooperatively with the relay in order to maximize the long term ratio of the data that is successfully delivered, to the total data that is generated.

In order to optimally determine if the relay should be used or not at a given time, in addition to its own state information (e.g. current battery level), the source also needs to know the state information at the relay. We consider two scenarios with different system observabilities. In the ideal, *fully observable* system, we assume that the source can always obtain complete knowledge about the current status of the battery level, battery recharge and event occurrences at both the source and the relay. The transmission scheduling problem is then formulated as a Markov Decision Process. However, a fully observable system is unrealistic and in a more practical system, it is reasonable to assume that when a relay transmits or relays data, the headers of the packets may include the relay's state information. However, in periods without data, conveying the state information of the relay in real time represents a significant overhead. Thus the source may have to base its decision on stale state information. In such a *partially observable* system, the scheduling problem to determine the optimal decision based on partial information about the system is formulated as a Partially Observable Markov Decision

H. Li and B. Sikdar are with Department of Electrical Computer and Systems Engineering Rensselaer Polytechnic Institute, 110 8th Street, Troy, NY, 12180 USA: e-mail: {lih4, sikdarb}@rpi.edu.

N. Jaggi is with the Department of Department of Electrical Engineering and Computer Science, Wichita State University, Wichita KS 67260, USA E-mail: neeraj.jaggi@wichita.edu.

Manuscript received May 09, 2010; revised October 22, 2010; accepted December 22, 2010.

Process. We show that the POMDP can be decomposed into a equivalent MDP, the solution of which also gives the optimal solution for the POMDP. In addition, we provide a theoretical upper bound on the performance of any arbitrary scheduler.

The rest of the paper is organized as follows. Section II presents the related work, Section III describes the system model, and an upper bound on the performance is presented in Section IV. A MDP formulation of the scheduling problem for the fully observable system is presented in Section V. A POMDP formulation for the partially observable system is presented in Section VI and Section VII considers a multiple node scenario. Finally, simulation results are presented in Section VIII and Section IX concludes the paper.

II. Related Work

The use of energy harvesting has been proposed in the general framework of wireless sensor networks (WSNs) [3], [9]. However, existing results concerning the optimal use of cooperative communication techniques for energy harvesting networks are limited. The problem of duty-cycling in sensor networks with energy harvesting is considered in [10], [11], [12]. In [13] it is shown that using cooperative automatic repeat request (ARQ) protocols, sensor nodes can match their energy consumption to their energy harvesting rate, thereby improving the throughput. The authors of [14] address the problem of sensor activation with battery recharging assuming temporally correlated events. Throughput-optimal and delay minimizing energy management policies are proposed for sensors with energy harvesting in [15] while [16] develops throughput maximizing policies assuming knowledge of the energy harvesting process and the time varying channel conditions.

Physical layer issues for cooperative communications in the context of sensor and ad hoc networks has been addressed extensively (see [17] for a survey). Cooperative communication using collaborative MIMO is considered in [18], [19] while [20] considers broadcasting using cooperative communications. A survey of various power allocation strategies for cooperative communications can be found in [21]. Medium access control (MAC) layer issues for cooperative communications are considered in [23], [24]. Note that the notion of cooperative communications in [22] and similar papers is different: the focus of these papers is to use a multi-hop path providing a better bit rate than a single hop path. To the best of our knowledge, the problem of scheduling cooperative transmissions in networks with energy harvesting has not been addressed in existing literature.

III. System Model

We consider a general WSN whose primary function is to monitor and report events of interest, such as detecting and reporting abnormal conditions and emergencies, industrial monitoring etc. Each node in the network may serve as a source, a destination, or a relay or any combination of the roles. We consider one hop transmissions and any node that is within the transmission range (i.e. a neighbor) of both the source and the destination may act as a cooperative relay. In this paper, we assume that decode-and-forward relaying is adopted [7] at the physical layer. To develop our models, we first consider a three-node network with one source, one relay and one destination node. In Section VII we consider networks with more than three nodes and where nodes can take on multiple roles.

A source sensor has two transmission modes: the *direct mode* in which the sensor transmits the packet directly to the destination and consumes δ_1^s units of energy and the *relay mode* (which consumes δ_2^s units of energy) in which the packet is transmitted by the source and relayed by the relay sensor. A relay sensor also has two transmission modes: *own-traffic mode* and *relay mode*. In the own-traffic mode, the relay sensor transmits its own packet to the destination consuming δ_1^r units of energy while in the relay mode the relay sensor's own traffic is discarded and δ_2^r units of energy is consumed to relay another sensor node's packet. We have $\delta_1 > \delta_2$ where we have dropped the superscript (s or r to indicate source and relay sensors, respectively) to indicate that the relation holds for both source and relay sensors. We assume that the sensors are working in real-time monitoring scenarios. Thus no retransmission is attempted for packets with errors. Also, if a packet is not transmitted in the slot in which it arrives, it is dropped to allow more recent data to be transmitted. Finally, a sensor is considered available for operation if it has enough energy to transmit or relay a packet.

A discrete time model is assumed where time is slotted in intervals of unit length. Each slot is long enough so that a source node and a relay node can either cooperatively transmit one data packet for the source, or both can transmit one of their own packets. At most one data packet is generated at a node in a slot. Each sensor has a rechargeable battery and an energy harvesting device. The energy generation process at each sensor is modeled by a correlated, two-state process with parameters (q_{on}, q_{off}). In the *on* state (i.e. when ambient conditions are conducive to energy harvesting), the sensor generates energy at a constant rate of c units per time slot. In the *off* state, no energy is generated. If the sensor harvested energy in the current slot, it harvests energy in the next slot with probability q_{on} and no energy is harvested with probability $1 - q_{on}$. On the other hand, if no energy was harvested in the current slot, no energy is harvested in the next slot with probability q_{off}, and energy is harvested with probability $1 - q_{off}$, and we assume $0.5 < q_{on}, q_{off} < 1$. We assume that the energy generated during a recharge event is available at the end of the slot. The correlation probabilities and the charge amount c for the different sensor nodes (source or relay) could be different, and are denoted q_{on}^s, q_{off}^s, q_{on}^r, q_{off}^r, c^s, c^r respectively. The two-state model can account for harvesting methods such as vibration-based event-driven harvesting, as well as harvesting devices that convert light if it is sufficiently bright and little or none otherwise. More complicated harvesting models can be accommodated by increasing the number of states in the Markov model.

The data packets that the sensors report to a sink are also generated according to a correlated, two-state process with parameters (p_{on}, p_{off}) with $0.5 < p_{on}, p_{off} < 1$, where in the *on* state an event (i.e. data packet) is generated in each slot, and no events are generated in the *off* state. We assume that an event is generated and detected at the beginning of a time slot. The average duration of a period of continuous events, $E[N]$, is

$$E[N] \ = \ \sum_{i=1}^{\infty} i(p_{on})^{i-1}(1-p_{on}) \ = \ \frac{1}{1-p_{on}}, \qquad (1)$$

and the steady-state probability of event occurrence, π_{on}, is

$$\pi_{on} = \frac{1-p_{off}}{2-p_{on}-p_{off}}. \qquad (2)$$

Similarly, the average length of a period without events and its steady-state probability are $\frac{1}{1-p_{off}}$ and $\pi_{off} = 1-\pi_{on}$, respectively. Also, the average length of a period with energy harvesting and the steady-state probability of such events are $\frac{1}{1-q_{on}}$ and $\mu_{on} = \frac{1-q_{off}}{2-q_{on}-q_{off}}$, respectively. Finally, the expected length of periods without recharging and its steady-state probability are $\frac{1}{1-q_{off}}$ and $\mu_{off} = 1-\mu_{on}$, respectively. The parameters corresponding to the source and relay nodes are denoted with a superscript of s and r, respectively (e.g. p^s_{on}).

The communication strategy of a sensor pair {source, designated relay} is governed by a policy Π that decides on the transmission mode to be used for reporting events. We define the set of actions as $\mathcal{A} = \{0,1,2,3,4\}$, denoting {no transmission, no transmission}, {direct, no transmission}, {relay, relay}, {direct, own-traffic}, and {no transmission, own-traffic}, respectively. The action taken by the sensor pair in time slot t is denoted by a_t, $a_t \in \mathcal{A}$. A transmission action can be taken only if the corresponding sensor has enough power (δ_1 for direct/own-traffic mode and δ_2 for relay mode) and an event occurs at the beginning of the slot. The basic objective of the decision policy Π is to maximize the *quality of coverage*, defined as follows. Let $\mathcal{E}_o(T)$ denote the number of events that occurred in the sensing region of a sensor over a period of T slots in the interval $[0,T]$. Also, let $\mathcal{E}_d(T)$ denote the total number of events that are detected and correctly reported by a sensor over the same period under policy Π. The time average of the fraction of events detected and correctly reported by both the source and relay nodes represents the quality of coverage and is given by

$$U(\Pi) = \lim_{T\to\infty} \frac{\mathcal{E}^s_d(T)+\mathcal{E}^r_d(T)}{\mathcal{E}^s_o(T)+\mathcal{E}^r_o(T)}. \qquad (3)$$

A node is said to be *active* in a time slot if the action is taken such that it has a packet transmission (either its own traffic or relaying), and *inactive* if there is no transmission. For example, when the action equals 2, 3 or 4, the relay is active, and inactive otherwise. We assume that the communication strategy is decided at the source sensor and consider following two systems with different observability at the source.

(1) In a *fully observable* system, the source is able to observe the state at the relay even if the relay is inactive in a slot. Thus, the transmission scheduling is based on perfect system state information, specifically, the current battery levels, and the states of the recharge and the event generation processes of both sensors.

(2) To consider a more practical system, we assume that when a sensor transmits a packet, its current state information is included in the packet's header. Then, a source sensor can obtain state information about the relay including: (a) the relay's current energy level, (b) whether there is an event generated at the relay in the current slot, and (c) whether the relay's battery is currently recharging or not. However, if the relay is inactive in a slot, the source will not have the updated state information of the relay. Thus the decision at the source may be based on: (a) the current battery level, the states of the energy and the event generation processes at the source; (b) the partial information of the battery, the energy and event generation processes at the relay based upon the information which was obtained when the relay was last active. We refer to such a system as a *partially observable* system.

For any system, we next find an upper bound on the performance of an arbitrary scheduling policy.

IV. An Upper Bound on the Performance

Let T_i with $i \in \{0,1,2,3,4\}$ denote the number of time slots in which action a_i is successfully taken over the period $[0,T]$, under the optimal policy Π_{OPT}. Then events reported by source and relay nodes will be $\mathcal{E}^s_d(T) = T_1+T_2+T_3$ and $\mathcal{E}^r_d(T) = T_3+T_4$ respectively. As $T\to\infty$, the number of events occurring in the interval $[0,T]$ satisfies

$$\lim_{T\to\infty} \frac{\mathcal{E}^s_o(T)}{T} = \pi^s_{on} = \frac{1-p^s_{off}}{2-p^s_{on}-p^s_{off}} \qquad (4)$$

$$\lim_{T\to\infty} \frac{\mathcal{E}^r_o(T)}{T} = \pi^r_{on} = \frac{1-p^r_{off}}{2-p^r_{on}-p^r_{off}}. \qquad (5)$$

We denote the available energy at the sensor at the beginning of slot t by L_t and assume that the initial energy level is L_0. The expected energy level of the source sensor at time T is given by

$$E[L^s_T] = L^s_0 - (T_1+T_3)\delta^s_1 - T_2\delta^s_2 + T\mu^s_{on}c^s, \qquad (6)$$

and the expected energy level of the relay sensor at time T is given by

$$E[L^r_T] = L^r_0 - (T_3+T_4)\delta^r_1 - T_2\delta^r_2 + T\mu^r_{on}c^r. \qquad (7)$$

Using the fact that $E[L^s_T] \geq 0$ and $\delta^s_1 > \delta^s_2$ we have

$$\lim_{T\to\infty} \frac{(T_1+T_3)\delta^s_1 + T_2\delta^s_2}{T} \leq \mu^s_{on}c^s \qquad (8)$$

and

$$\lim_{T\to\infty} \frac{T_1+T_3+T_2}{T} \leq \frac{\mu^s_{on}c^s}{\delta^s_2}. \qquad (9)$$

Additionally, since $\frac{T_1+T_2+T_3}{T} \leq \frac{\mathcal{E}^s_o(T)}{T}$, we have

$$\lim_{T\to\infty} \frac{T_1+T_3+T_2}{T} \leq \min\{\frac{\mu^s_{on}c^s}{\delta^s_2}, \pi^s_{on}\}. \qquad (10)$$

Using the fact that $E[L^r_T] \geq 0$ and $T_2 \geq 0$, we have

$$\lim_{T\to\infty} \frac{(T_3+T_4)\delta^r_1 + T_2\delta^r_2}{T} \leq \mu^r_{on}c^r \qquad (11)$$

$$\lim_{T\to\infty} \frac{T_3+T_4}{T} \leq \frac{\mu^r_{on}c^r}{\delta^r_1}. \qquad (12)$$

Since $\frac{T_3+T_4}{T} \leq \frac{\mathcal{E}^r_o(T)}{T}$, we also have

$$\lim_{T\to\infty} \frac{T_3+T_4}{T} \leq \min\{\frac{\mu^r_{on}c^r}{\delta^r_1}, \pi^r_{on}\}. \qquad (13)$$

Finally, combining Eqns. (10) and (13), the performance of the optimal policy and thus any arbitrary policy is bounded by

$$U(\Pi_{OPT}) = \lim_{T\to\infty} \frac{\mathcal{E}_d^s(T)+\mathcal{E}_d^r(T)}{\mathcal{E}_o^s(T)+\mathcal{E}_o^r(T)} \leq \frac{\min\{\frac{\mu_{on}^s c^s}{\delta_2^s}, \pi_{on}^s\}+\min\{\frac{\mu_{on}^r c^r}{\delta_1^r}, \pi_{on}^r\}}{\pi_{on}^s+\pi_{on}^r} \quad (14)$$

V. Markov Decision Process Formulation with Perfect State Information

The problem of developing the optimal scheduling problem is fairly challenging due to the number of variables involved and their complex interactions. In this section we show that the problem may be modeled as a MDP, thereby allowing us to obtain the optimal policy[1].

Denote the system state at time t by $X_t = (L_t^s, E_t^s, C_t^s, L_t^r, E_t^r, C_t^r)$ where $L_t^s, L_t^r \in \{0,1,2,\cdots,K\}$ represents the energy available in the sensors at time t, and $E_t^s, E_t^r \in \{0,1\}$ equals one if an event to be reported during time interval $[t,t+1)$ occurred at time t and zero otherwise. Also, $C_t^s, C_t^r \in \{0,1\}$ equals one if the sensor recharged during time interval $[t-1,t)$ and zero otherwise. (i.e. we assume that the sensor does not know at time t if it will recharge during interval $[t,t+1)$). The battery capacity of the sensor is assumed to be K. Then the state space $\mathcal{X}$ is given by $\mathcal{X} = \{(0,0,0,0,0,0),(1,0,0,0,0,0),\cdots,(K,1,1,K,1,1)\}$ with $|\mathcal{X}| = 16(K+1)^2$. The action taken at time t is denoted by $a_t \in \mathcal{A}$ as described in Section III. The next state of the system depends only on the current state and the action taken. Thus the system constitutes a Markov Decision Process [25].

Let θ^s and θ^r denote the reward gained by the system for each source sensor and relay sensor event that is successfully reported. The values of θ^s and θ^r may be chosen to reflect the importance of the observations of each sensor. Alternatively, θ^s and θ^r may also be made equal to the probability that a transmitted packet is received without errors, in order to account for channel errors. The reward function $R(X_t,a_t)$ is then given by,

$$R(X_t,a_t) = \begin{cases} \theta^s & \text{if } a_t = 1 \text{ or } 2, E_t^s = 1 \\ \theta^s+\theta^r & \text{if } a_t = 3, E_t^s = E_t^r = 1 \\ \theta^r & \text{if } a_t = 4, E_t^r = 1 \\ 0 & \text{otherwise.} \end{cases} \quad (15)$$

Let g_t^s and g_t^r be the amount of energy gained by the source and relay sensors in the interval $[t,t+1)$ respectively. Then,

$$g_t^s = \begin{cases} c^s & \text{w.p. } C_t^s q_{on}^s + (1-C_t^s)(1-q_{off}^s) \\ 0 & \text{otherwise} \end{cases} \quad (16)$$

$$g_t^r = \begin{cases} c^r & \text{w.p. } C_t^r q_{on}^r + (1-C_t^r)(1-q_{off}^r) \\ 0 & \text{otherwise} \end{cases} \quad (17)$$

where $w.p.$ stands for "with probability". Let l_t^s and l_t^r be the amount of energy spent by the source and relay sensors in the interval $[t,t+1)$ respectively. Then,

$$l_t^s = \begin{cases} \delta_1^s & \text{if } a_t = 1 \text{ or } a_t = 3 \\ \delta_2^s & \text{if } a_t = 2 \\ 0 & \text{otherwise} \end{cases} \quad (18)$$

$$l_t^r = \begin{cases} \delta_1^r & \text{if } a_t = 3 \text{ or } a_t = 4 \\ \delta_2^r & \text{if } a_t = 2 \\ 0 & \text{otherwise.} \end{cases} \quad (19)$$

To complete the MDP formulation, the system state at time $t+1$ is given by

$$X_{t+1} = (L_{t+1}^s, E_{t+1}^s, C_{t+1}^s, L_{t+1}^r, E_{t+1}^r, C_{t+1}^r), \quad (20)$$

where

$$L_{t+1}^s = \max\{\min\{L_t^s + g_t^s - l_t^s, 0\}, K\} \quad (21)$$

$$L_{t+1}^r = \max\{\min\{L_t^r + g_t^r - l_t^r, 0\}, K\} \quad (22)$$

$$E_{t+1}^s = \begin{cases} 1 & \text{w.p. } E_t^s p_{on}^s + (1-E_t^s)(1-p_{off}^s) \\ 0 & \text{otherwise} \end{cases} \quad (23)$$

$$E_{t+1}^r = \begin{cases} 1 & \text{w.p. } E_t^r p_{on}^r + (1-E_t^r)(1-p_{off}^r) \\ 0 & \text{otherwise} \end{cases} \quad (24)$$

$$C_{t+1}^s = \begin{cases} 1 & \text{w.p. } C_t^s q_{on}^s + (1-C_t^s)(1-q_{off}^s) \\ 0 & \text{otherwise} \end{cases} \quad (25)$$

$$C_{t+1}^r = \begin{cases} 1 & \text{w.p. } C_t^r q_{on}^r + (1-C_t^r)(1-q_{off}^r) \\ 0 & \text{otherwise.} \end{cases} \quad (26)$$

The objective is to maximize the average reward criteria over an infinite horizon. The optimal solution can be computed by using value iteration [25]. Since the induced Markov chain is unichain, from Theorem 8.5.2 of [25], there exists a deterministic, Markov, stationary optimal policy Π_{MD} which also leads to a steady-state transition probability matrix. Considering the average expected reward criteria, the optimality equations are given by [26]

$$\lambda^* + h^*(X) = \max_{a\in\{0,1,2,3,4\}} \left[R(X,a) + \sum_{X'=(0,0,0,0,0,0)}^{(K,1,1,K,1,1)} p_{X,X'}(a) h^*(X') \right], \quad \forall X \in \mathcal{X} \quad (27)$$

where $p_{X,X'}(a)$ represents the transition probability from state X to X' when action a is taken, λ^* is the optimal average reward and $h^*(i)$ are the optimal rewards when starting at state $i = (0,0,0,0,0,0)$, $\cdots$, $(K,1,1,K,1,1)$. For the purpose of evaluation, the relative value iteration technique [26] is used to solve Eqn. (27).

VI. Partially Observable Markov Decision Process Formulation

For the partially observable system, we first formulate the decision problem as a POMDP, and then present the equivalent MDP formulation.

A. System States and Observations

The system state at time t is denoted by $X_t = (L_t^s, E_t^s, C_t^s, L_t^r, E_t^r, C_t^r)$, as in Section V, except that the variable $E_t^r \in \{0,1\}$ is defined similarly for the relay but

[1]Readers are referred to [25] for an introduction to MDPs.

equals one if the event process is *on* during time interval $[t-1,t)$. Note that while E_t^r, C_t^s and C_t^r are based on the interval $[t-1,t)$, E_t^s is based on the interval $[t,t+1)$. The state of the relay at time t is defined in terms of the previous slot since that is the latest information the source may have about the relay. We assume that the battery at a sensor has a finite capacity K. Then the state space is $\mathcal{X}$ as defined in Section V. In subsequent discussions, we also refer to $X_t^s = (L_t^s, E_t^s, C_t^s)$ as the source sensor's state and $X_t^r = (L_t^r, E_t^r, C_t^r)$ as the relay's state at time t. The action taken at time t is denoted by $a_t \in \mathcal{A}$ as described in Section III.

The *system observation* at time t at the source sensor is denoted by Y_t. The source is assumed to always have full information about itself. If the action taken at time $t-1$ is $2,3$ or 4, then the relay was active, and the observation matches the state and equals X_t. However, if the action taken was 0 or 1, the relay was inactive. Thus the state of the event and energy generation processes at the relay are not known, along with the energy level at the relay (due to the possibility of recharging). Thus the observation Y_t is characterized by,

$$Y_t = \begin{cases} X_t & \text{if } a_{t-1} \in \{2,3,4\} \\ (L_t^s, E_t^s, C_t^s, \phi_L, \phi_E, \phi_C) & \text{if } a_{t-1} \in \{0,1\} \end{cases}$$

where ϕ_ω denotes that a variable ω is unknown. The observation space $\mathcal{Y}$ is given by,

$$\mathcal{Y} = \mathcal{X} \cup \{(0,0,0,\phi_L,\phi_E,\phi_C), (1,0,0,\phi_L,\phi_E,\phi_C), \cdots, (K,1,1,\phi_L,\phi_E,\phi_C)\}, \quad (28)$$

with $|\mathcal{Y}| = 4(K+1)(4K+5)$. Let $q_{x,y}(a)$ be the probability distribution of the observation ($Y_t = y$) at time t, conditioned on the current state ($X_t = x$) and the action taken at time $t-1$ ($a_{t-1} = a$). Thus, $q_{x,y}(a) = Pr[Y_t = y | X_t = x, a_{t-1} = a]$. If the relay was active in time interval $[t-1,t)$, the source has perfect information at time t. Then we have

$$q_{x,y}(0) = q_{x,y}(1) = \begin{cases} 1 & y = (x.L^s, x.E^s, x.C^s, \phi_L, \phi_E, \phi_C) \\ 0 & \text{otherwise} \end{cases}$$

$$q_{x,y}(2) = q_{x,y}(3) = q_{x,y}(4) = \begin{cases} 1 & y = x \\ 0 & \text{otherwise} \end{cases} \quad (29)$$

where $x.L^s = L_t^s$, $x.E^s = E_t^s$ and $x.C^s = C_t^s$ when $x = (L_t^s, E_t^s, C_t^s, L_t^r, E_t^r, C_t^r)$.

B. POMDP Formulation

In the presence of only partial observations, the optimal action depends on the current and past observations, and on past actions. Existing work has shown that a POMDP may be formulated as a completely observable MDP with the same finite action set [26], [27], [28]. The state space for the equivalent MDP comprises of the space of probability distributions on the original state space. Thus in the general case, the state space of the equivalent MDP may become uncountable or infinite. In our case, the structure of the POMDP leads to a countable state space for the equivalent MDP, guaranteeing the existence of an optimal solution to the average cost (reward) optimality equation [28]. As a result, the solution to the equivalent MDP with complete state information provides the optimal actions to take in the POMDP, and with the optimal reward.

Denote the state space of the equivalent MDP as Δ, and its state at time t as Z_t. Then $Z_t \in \Delta$ is a information vector of length $|\mathcal{X}|$, whose i-th component is given by,

$$Z_t^{(i)} = Pr[X_t = i | y_t, \cdots, y_1; a_{t-1}, \cdots, a_0], \quad i \in \mathcal{X}. \quad (30)$$

We have $\mathcal{I}' Z_t = 1$, where $\mathcal{I}'$ denotes a row vector of length $|\mathcal{X}|$ with all elements equal to 1, since the elements of Z_t are mutually exclusive whose union is the universal set. The state Z_{t+1} is recursively computable given the state transition probability matrices $P(a)$, action taken a_t, and the observation y_{t+1} and is given by [28],

$$Z_{t+1} = \sum_{y \in \mathcal{Y}} \frac{\bar{Q}_y(a_t) P'(a_t) Z_t}{\mathcal{I}' \bar{Q}_y(a_t) P'(a_t) Z_t} I[Y_{t+1} = y], \quad (31)$$

where $I[B]$ denotes the indicator function of the event B and the matrices $\bar{Q}_y(a) = \text{diag}\{q_{x,y}(a)\}$, with $q_{x,y}(a)$ as defined in Eqn. (29). $P(a)$ is the state transition probability matrix with action a, and for $i,j \in \mathcal{X}$, is defined as $[P(a)]_{ij} = Pr[X_{t+1} = j | X_t = i, X_{t-1}, \cdots, X_0, a_t = a, a_{t-1}, \cdots, a_0] = Pr[X_{t+1} = j | X_t = i, a_t = a] \triangleq P_{ij}(a)$, i.e. it is the transition probability of the system state, given the history of previous states and actions. We use $\bar{T}(y, Z_t, a_t) = \bar{Q}_y(a_t) P'(a_t) Z_t$ to denote the numerator and $V(y, Z_t, a_t) = \mathcal{I}' \bar{Q}_y(a_t) P'(a_t) Z_t$ to denote the denominator in Eqn. (31), with $P'(a_t)$ denoting the transpose of $P(a_t)$. Then $\bar{T}(y, Z_t, a_t)$ represents the probability of the event $X_{t+1} = i, Y_{t+1} = y$ given past actions and observations, i.e. $\bar{T}(y, Z_t, a_t) = Pr[X_{t+1} = i, Y_{t+1} = y | Y_t = y_t, Y_{t-1}, \cdots, Y_0, a_t = a, a_{t-1}, \cdots, a_0] = Pr[X_{t+1} = i, Y_{t+1} = y | Y_t = y_t, a_t = a]$. $V(y, Z_t, a_t)$ is interpreted as the probability of $Y_{t+1} = y$ given the past actions and observations, $V(y, Z_t, a_t) = Pr[Y_{t+1} = y | Y_t = y_t, Y_{t-1}, \cdots, Y_0, a_t = a, a_{t-1}, \cdots, a_0]$, i.e.

$$V(y, Z_t, a_t) = Pr[Y_{t+1} = y | Y_t = y_t, a_t = a]. \quad (32)$$

Denote,

$$W(y, Z_t, a_t) = \frac{\bar{T}(y, Z_t, a_t)}{V(y, Z_t, a_t)}. \quad (33)$$

Then Eqn. (31) can be written as,

$$Z_{t+1} = \sum_{y \in \mathcal{Y}} W(y, Z_t, a_t) I[Y_{t+1} = y]. \quad (34)$$

Thus $\{Z_t\}$ forms a completely observable controlled Markov process with state space Δ.

In Section VI-C, we will show that the state space of the equivalent MDP is countable and that the state at time t, Z_t, can be represented in the form $Z_t = (L_t^s, E_t^s, C_t^s, L^r, E^r, C^r, i)$, representing the following: (a) the relay had no transmissions in the past i slots; (b) the state of the relay when it last transmitted was (L^r, E^r, C^r); (c) the current state at the source is (L_t^s, E_t^s, C_t^s).

The POMDP is then transformed to an equivalent MDP with state space Δ and the optimality equations for this MDP are given by [28]:

$$\Gamma^* + h^*(Z) = \max_{a\in\mathcal{A}}\left[\bar{R}(Z,a) + \sum_{y\in\mathcal{Y}} V(y,Z,a)h^*(W(y,Z,a))\right], \quad \forall Z \in \Delta. \tag{35}$$

where $h^*(Z)$ is the optimal reward when starting at state Z, $V(y,Z,a)$ and $W(y,Z,u)$ are defined in Eqns. (32) and (33), and $\bar{R}(Z,a) = Z'[R(i,a)]_{i\in\mathcal{X}}$ is the reward function which will be discussed in Section VI-D. These equations can be solved using the relative value iteration algorithm [26], however, exact closed-form expressions for h^* and Γ^* may not exist, particularly for finite K.

C. Formulation of the State Space

Given $X_t^s = (L_t^s, E_t^s, C_t^s)$ as the state of the source at time t and observation $y_t = (L_t^s, E_t^s, C_t^s, \phi_L, \phi_E, \phi_C)$, the observation vector has $4(K+1)$ possibilities with different combinations of $0 \le \phi_L \le K$, $\phi_E \in \{0,1\}$ and $\phi_C \in \{0,1\}$. We number these states as: state $1 = (L_t^s, E_t^s, C_t^s, 0, 0, 0)$, state $2 = (L_t^s, E_t^s, C_t^s, 0, 0, 1)$, $\cdots$, and state $4(K+1) = (L_t^s, E_t^s, C_t^s, K, 1, 1)$.

Let e^j denote the unit column vector with all zeros except the j^{th} element being one. Then $Z_t = e^{y_t} = e^{x_t}$ for $a_{t-1} \in \{2,3,4\}$. However, if $a_{t-1} \in \{0,1\}$, the state Z_t of the equivalent MDP has a maximum of $4(K+1)$ non-zero components, and can be represented by,

$$Z_t = \alpha_1 e^1 + \alpha_2 e^2 + \cdots + \alpha_{4(K+1)} e^{4(K+1)} \tag{36}$$

where $\sum_{i=1}^{4(K+1)} \alpha_i = 1$. To obtain the values of $\alpha_i, 1 \le i \le 4(K+1)$, we first evaluate the transition probabilities of the event process and then those for the energy generation process and the battery level.

Let $F_{1,0}^{(i)}$ be the probability that the event process at a sensor (source or relay) at time $t+i$ is *off*, given that it was *on* at time t. Similarly, $F_{0,1}^{(i)}$ denotes the i-step transition probability of the event generation process at the node from *off* to *on* state in i time slots. Then the $(i+1)$-step transition probabilities are recursively given by,

$$\begin{aligned} F_{1,0}^{(i+1)} &= p_{off} F_{1,0}^{(i)} + (1-p_{on})(1-F_{1,0}^{(i)}) \\ F_{0,1}^{(i+1)} &= p_{on} F_{0,1}^{(i)} + (1-p_{off})(1-F_{0,1}^{(i)}). \end{aligned}$$

Equivalently,

$$F_{E,1-E}^{(i+1)} = p_{1-E} F_{E,1-E}^{(i)} + (1-p_E)(1-F_{E,1-E}^{(i)}) \tag{37}$$

where $E \in \{0,1\}$, $p_0 = p_{off}$ and $p_1 = p_{on}$. We also have $F_{E,E} = 1 - F_{E,1-E}^{(i)}$, $F_{1,0}^{(1)} = 1 - p_{on}$ and $F_{0,1}^{(1)} = 1 - p_{off}$. Since $0 < p_{off} + p_{on} - 1 < 1$, it can be shown that,

$$F_{E,1-E}^{(i)} = \frac{(1-p_E)[1-(p_E+p_{1-E}-1)^i]}{2-p_E-p_{1-E}}, \tag{38}$$

and $\lim_{i\to\infty} F_{0,1}^{(i)} = \pi_{on}$ and $\lim_{i\to\infty} F_{1,0}^{(i)} = \pi_{off}$.

The state of the battery level at the relay is related to the recharge process, given the last known battery level. In the slots where the relay does not transmit, the battery level increases whenever the recharge process is *on*. The i-step transition probabilities of the recharge process at the relay, $G_{E,1-E}^{(i)}$, are then also given by Eqn. (38) with p_E replaced by q_E. To obtain the battery level at time $t+i$, given the battery level and recharging state at time t, we need to evaluate the number of slots with recharge events during the i-step time interval. This problem is solved recursively as follows.

Consider a time interval with u slots. Let $B(u,v,0)$ and $B(u,v,1)$ denote the probabilities that in v out of u slots, the recharge process at a sensor was in the *on* state and the state in the u-th (the final) slot is *off* ($C^r = 0$) and *on* ($C^r = 1$), respectively. These probabilities can be recursively written as,

$$\begin{aligned} B(u,v,0) &= (1-q_{on})B(u-1,v,1) + q_{off}B(u-1,v,0) \\ B(u,v,1) &= q_{on}B(u-1,v-1,1) + (1-q_{off})B(u-1,v-1,0), \end{aligned}$$

while satisfying the following initial conditions:

$$B(u,0,1) = 0 \tag{39}$$

$$B(u,u,0) = 0 \tag{40}$$

$$B(u,u,1) = (q_{on})^u B_1(0) + (1-q_{off})(q_{on})^{u-1} B_0(0) \tag{41}$$

$$B(u,0,0) = (q_{off})^u B_0(0) + (1-q_{on})(q_{off})^{u-1} B_1(0) \tag{42}$$

with $u, v \in \{1,2,\cdots\}$, $u \ge v$ and $B_1(0) = 1$ if the energy harvesting process was in the *on* state in the interval just preceding the u-slot interval, 0 otherwise. Also, $B_0(0) = 1 - B_1(0)$.

Given that a relay sensor did not transmit for i slots, let $\beta_{j,k}^{(i)}$ denote the probability of transition of the relay from state j to state k in i slots. Let the states be $j = X_t^r = (L,E,C)$ and $k = X_{t+i}^r = (L',E',C')$. Then, $\beta_{j,k}^{(i)}$ is given by,

$$\beta_{j,k}^{(i)} = \begin{cases} F_{E,E'}^{(i)} B(i,v,C') & \text{if } L' = L + vc^r < K \\ \sum_{v=\lceil \frac{K-L}{c^r} \rceil}^{i} F_{E,E'}^{(i)} B(i,v,C') & \text{if } L' = K \\ 0 & \text{otherwise} \end{cases} \tag{43}$$

for all $v \in [0,1,\cdots,i]$. Since the state of the source is independent of that of the relay, given that the relay was in state $j = (L^r, E^r, C^r)$ the time last it transmitted and that the relay has been inactive for the last i slots, we can rewrite Eqn. (36) as,

$$Z_t = \beta_{j,1}^{(i)} e^1 + \beta_{j,2}^{(i)} e^2 + \cdots + \beta_{j,4(K+1)}^{(i)} e^{4(K+1)}. \tag{44}$$

Now we can represent Z_t as,

$$Z_t = (L_t^s, E_t^s, C_t^s, L^r, E^r, C^r, i) \tag{45}$$

where $X_t^s = (L_t^s, E_t^s, C_t^s)$ is the source's state at time t and $j = X^r = (L^r, E^r, C^r)$ is the relay's state when it last transmitted, which was i slots ago. We then have the following result:

Lemma 1: The state-space Δ is countable.

Proof: Let $Z_t = (L_t^s, E_t^s, C_t^s, L^r, E^r, C^r, i)$ for some $(L_t^s, E_t^s, C_t^s, L^r, E^r, C^r) \in \mathcal{X}$ and integer $i \ge 0$. Let $X_{t+1} = (L_{t+1}^s, E_{t+1}^s, C_{t+1}^s, L_{t+1}^r, E_{t+1}^r, C_{t+1}^r)$.

Case(i) : $a_t = \{0,1\}$. Since the relay is not observable, $y_{t+1} = (L_{t+1}^s, E_{t+1}^s, C_{t+1}^s, \phi_L, \phi_E, \phi_C)$. Then, Z_{t+1} in the form of

Eqn. (44) can be expanded, using $j=(L^r,E^r,C^r)$, as,

$$\begin{aligned} Z_{t+1} &= \sum_{k=1}^{4(K+1)} \beta^{(i)}_{j,k}\beta^{(1)}_{k,1}e^1 + \cdots + \sum_{k=1}^{4(K+1)} \beta^{(i)}_{j,k}\beta^{(1)}_{k,4(K+1)}e^{4(K+1)} \\ &= \beta^{(i+1)}_{j,1}e^1 + \beta^{(i+1)}_{j,2}e^2 + ... + \beta^{(i+1)}_{j,4(K+1)}e^{4(K+1)} \\ &= (L^s_{t+1}, E^s_{t+1}, C^s_{t+1}, L^r, E^r, C^r, i+1). \end{aligned} \tag{46}$$

Case(ii) : $a_t=2$. In this case the transmission of a source's packet relies on the relay and the action can be taken if the relay has enough power (by assumption, the relay discards its own traffic, if any). Thus,

$$Z_{t+1} = \begin{cases} (L^s_{t+1}, E^s_{t+1}, C^s_{t+1}, L^r_{t+1}, E^r_{t+1}, C^r_{t+1}, 0) \\ \qquad \text{w.p. } Pr[L^r_t \ge \delta^r_2 | L^r_{t-i} = L^r] \\ (L^s_{t+1}, E^s_{t+1}, C^s_{t+1}, L^r, E^r, C^r, i+1) \\ \qquad \text{w.p. } 1 - Pr[L^r_t \ge \delta^r_2 | L^r_{t-i} = L^r]. \end{cases} \tag{47}$$

Case(iii) : $a_t=\{3,4\}$. In these two cases, the relay will transmit its own traffic if it has enough power and has an event to report. Thus,

$$Z_{t+1} = \begin{cases} (L^s_{t+1}, E^s_{t+1}, C^s_{t+1}, L^r_{t+1}, E^r_{t+1}, C^r_{t+1}, 0) \\ \text{w.p. } Pr[L^r_t \ge \delta^r_1 | L^r_{t-i} = L^r] Pr[E^r_t = 1 | E^r_{t-i} = E^r] \\ (L^s_{t+1}, E^s_{t+1}, C^s_{t+1}, L^r, E^r, C^r, i+1) \\ \text{w.p. } 1 - Pr[L^r_t \ge \delta^r_1 | L^r_{t-i} = L^r] Pr[E^r_t = 1 | E^r_{t-i} = E^r]. \end{cases}$$

To sum up, Z_{t+1} is completely described by Z_t, a_t and y_{t+1}. Since L^s_t, L^r_t, E^s_t, E^r_t, C^s_t, C^r_t and i are individually finite or countable, and all $Z \in \Delta$ has the form of Eqn. (45), we have the result. ■

D. Equivalent MDP Reward Function

Recall that θ^s and θ^r denote the rewards gained by the system for each source sensor and relay sensor event that is successfully reported, respectively. For the partially observable system, the reward associated with the states $Z \in \Delta$ of the equivalent MDP, denoted as $\bar{R}(Z,a)$, is the same as that of the optimal reward for the original POMDP [28]. Then, the reward function of the equivalent MDP at time t is given by,

$$\bar{R}(Z,a) = \begin{cases} \theta^s & \text{if } a_t = 1, E^s_t = 1, L^s_t \ge \delta^s_1 \\ \theta^s Pr[L^r_t \ge \delta^r_2 | L^r_{t-i} = L^r] & \text{if } a_t = 2, E^s_t = 1, L^s_t \ge \delta^s_2 \\ \theta^s + \theta^r Pr[L^r_t \ge \delta^r_1 | L^r_{t-i} = L^r] \cdot & \text{if } a_t = 3, E^s_t = 1, L^s_t \ge \delta^s_1 \\ \quad Pr[E^r_t = 1 | L^r, E^r_{t-i} = E^r] & \\ \theta^r Pr[L^r_t \ge \delta^r_1 | L^r_{t-i} = L^r] \cdot & \text{if } a_t = 4 \\ \quad Pr[E^r_t = 1 | E^r_{t-i} = E^r] & \\ 0 & \text{otherwise.} \end{cases} \tag{48}$$

E. Computational Complexity of MDP and POMDP

It has been shown in [29] that the infinite horizon average cost MDP is solvable in polynomial time by successive approximation techniques such as value iteration. However, no known strongly polynomial time algorithm exists for solving MDP [30]. Practical algorithms (such as approximation algorithms) based on alternative methods of analysis that rely on the structure of the MDP are possible [30]. In addition, solving the POMDP (corresponding to our relay scheduling problem) is PSPACE-hard, and efficient online implementation is not possible even if arbitrary amount of precomputation is allowed [29]. [31] shows that finding the optimal strategy for POMDP is NP-hard, but polytime approximations are possible.

The complexity of the proposed solution approach can be varied depending upon the accuracy of the desired relay scheduling scheme. In particular, the complexity of the solution can be reduced by limiting the number of iterations incurred by the value iteration . This can be achieved by choosing a suitable value for the convergence parameter ϵ, which could address this complexity vs. performance tradeoff. We propose that a precomputed table of optimal actions is loaded onto the source nodes in order for them to make online decisions. The complexity of the precomputation stage would depend upon the exact approach chosen to solve for the optimal solution, and one of the above mentioned approaches could be chosen to tradeoff performance in order to improve time complexity.

VII. Multi-node Networks

This section describes how the schedulers developed in the previous sections may be applied in a more realistic, multi-node network. We consider a network with an arbitrary number of nodes that may be arbitrarily distributed in a geographical region. When a node wishes to transmit a packet to a neighboring node, it has the option of using any of their common neighbors as a relay. In case no such node exists, the two nodes use a direct (non-cooperative) transmission. We do not assume the use of any particular medium access control (MAC) protocol and any of the ones proposed in literature such as [23], [24] may be used. We now present an overview of the data transmission for the case where a relay node exists.

When a node wishes to transmit data to another node, it first determines whether to use a direct transmission or use a relay. For each of its neighbors that are available to act as a relay, the source uses the schedulers proposed earlier to determine if a direct transmission is preferable. If the direct transmission is preferable for all the relays, the source directly sends its packet to the destination. Otherwise, the source selects the available relay with the highest battery power to help with the transmission. The actual sequence of transmissions depends on the MAC protocol in use.

There are variations possible to the methodology described above. The source needs the MAC layer's help to determine which nodes are available to help with a transmission. In order to simplify the MAC layer implementation, a source may choose to always use the same relay. Alternatively, the source may pick one of the relay nodes at random for each packet transmission and check if it is available to help. Irrespective of the scenario, the use of the proposed scheduler is to determine whether to use a direct transmission or a given relay node.

Finally we note that our models do not consider the scenario where a source node may in turn act as the relay for another node. This is a complex scenario and depends on a number of factors including the routing mechanism. However, when a source node (S_1) for one transmission is acting as the relay for the source node (S_2) of another transmission, node S_2 executes the proposed schedulers while using S_1's state variables for the relay. Thus effectively, each node tries to maximize its own quality of coverage in a distributed way. Even with the distributed scheme we can achieve better performance as compared to direct transmissions, as shown in the next section. For scenarios where the roles of sensors as source and relay nodes is strictly defined, even in the distributed setting our model leads to optimal results.

VIII. Simulation Results

This section explores the impact of various parameters on the performance of the proposed schedulers using simulations. We first present the results when only a single three-node-group (i.e. source, relay and destination) is present in the network and then consider networks with multiple groups. The simulations were done using an event-driven, packet-level simulator developed by us, primarily because energy harvesting is not well supported in existing simulators[2]. All simulations were run for a duration of 5000000 time units and physical layer aspects such as bit errors were not considered. All figures show the quality of coverage (unless noted otherwise) defined as the ratio of the number of events successfully reported to the total number of events generated. Since there is no existing literature addressing the same problem, there is no performance comparison with other schemes.

Figure 1 demonstrates the effect of the event generation process on the performance of a fully observable system with MDP formulated policy, along with the theoretical upper bound from Section IV, and all parameters are specified in the caption. In all the four cases, the recharge process parameters and transmission energies are the same for both the source and the relay sensors. Since p_{on} and p_{off} are constrained in the range (0.5,1.0), the four choices of (0.6,0.6), (0.6,0.9), (0.9,0.6) and (0.9,0.9) in Figure 1 give an indication of the performance in diverse settings of low-low, low-high, high-low and high-high correlation probabilities at the relay. For each of the four cases, the parameters p^s_{on} and p^s_{off} of the event generation process of the source node are varied from 0.55 to 0.95. The quality of coverage decreases as p^s_{on} increases, since an increase in p^s_{on} increases both π^s_{on} and the average length of periods with continuous packets, while decreasing π^s_{off} and the interval between bursts of traffic. With a relatively higher (close to 1) p^r_{on} and lower (close to 0.5) p^r_{off}, as in Figure 1(c), the event generation rate at the relay node increases, thereby reducing the energy available at the relay for helping the source sensor. Thus the quality of coverage degrades in this case, and in contrast, improves with a lower p^r_{on} and higher p^r_{off} (Figure 1(b)).

In Figure 1, the theoretical bound is tighter when p^s_{on} is low and p^s_{off} is high. For an intuition behind this observation, we note that the bound in Eqn. (14) uses two approximations: (a) the first term in its numerator uses Eqn. (9) that omits the term $\frac{T_1+T_3}{T}\frac{\delta^s_1-\delta^s_2}{\delta^s_2}$ in Eqn. (8), and, (b) the second term in the numerator of Eqn. (14) uses Eqn. (12) that neglects the term $\frac{T_2}{T}\frac{\delta^r_2}{\delta^r_1}$ in Eqn. (11). When p^s_{on} is low and p^s_{off} is high, we have $\pi^s_{on} < \frac{\mu^s_{on}c^s}{\delta^s_2}$. The source sensor has enough energy and tends to transmit the packet directly and the relay spends most of its energy and time slots on its own traffic. Consequently, T_2 is small and the approximation error from the second term is very small, and the bound is thus tight. On the other hand, as π^s_{on} increases, the first term of the numerator of Eqn. (14) is determined by $\frac{\mu^s_{on}c^s}{\delta^s_2}$ which introduces an error corresponding to the term $\frac{T_1+T_3}{T}\frac{\delta^s_1-\delta^s_2}{\delta^s_2}$ in Eqn. (9). Also, when π^s_{on} is large, the packet generation rate at the source is high, resulting in a low battery level at the source and a higher use of the relay to transmit the source's traffic. Thus the fraction $\frac{T_2}{T}$ increases and the error introduced by the second term in the numerator of Eqn. (14) also increases, and the bound becomes looser.

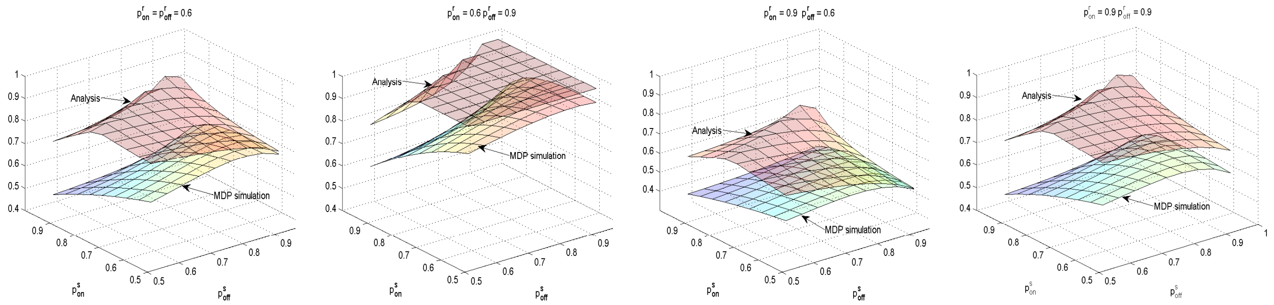
[2]The simulators may be downloaded from http://networks.ecse.rpi.edu/∼bsikdar/pomdp.

TABLE I
Relay usage summary. (Parameters used: $q^s_{on} = q^r_{on} = 0.85$, $q^s_{off} = q^r_{off} = 0.7$, $p^s_{on} = 0.85$, $p^s_{off} = 0.7$, $c^s = c^r = 1$, $\delta^s_1 = \delta^r_1 = 2$, $\delta^s_2 = \delta^r_2 = 1$, $\theta^s = \theta^r = 1$)

(p^r_{on}, p^r_{off})	(0.6,0.6)	(0.6,0.9)	(0.9,0.6)	(0.9,0.9)
QoC	0.5662	0.7309	0.4510	0.5179
Source QoC	0.5945	0.7765	0.4934	0.4932
Relay usage	0.3372	0.7477	0.0000	0.0000
Relay QoC	0.5284	0.5793	0.4157	0.5507

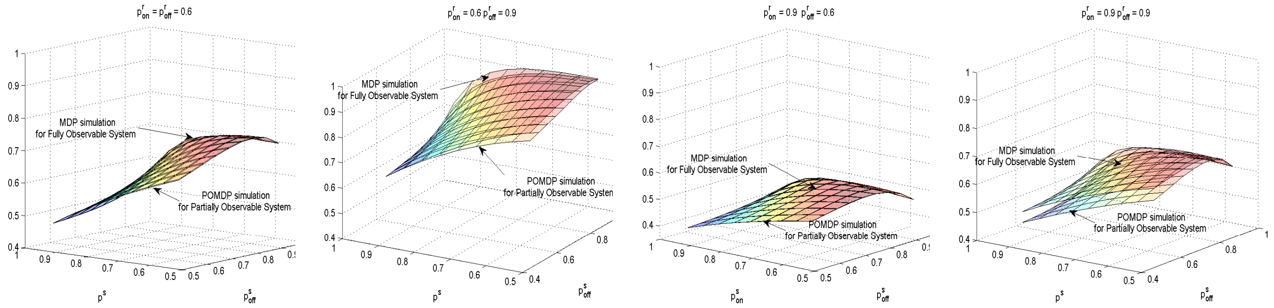
Another set of results (omitted due to constraints on the number of figures) explored the impact of the recharge process on the performance. For these results, the event generation process parameters and the transmission energies were the same for both the source and relay sensors. It was observed that the quality of coverage increases as q^s_{on} increases and q^s_{off} decreases, i.e. when the steady state probability of recharging and the average length of continuous recharging slots of the source node increase. The rate of increase in the quality of coverage with q^s_{on} is approximately linear when q^s_{off} is close to 0.5 and approximately exponential when it is close to 1. Also, as q^r_{on} increases and q^r_{off} decreases, the quality of coverage increases in general. The theoretical bound is tighter when both q^s_{on} is small (close to 0.5) and q^s_{off} is large (close to 1). The intuition behind the results is similar to that given for Figure 1.

Figure 2 demonstrates the effect of the event generation process on the performance for a partially observable system formulated as POMDP, compared with results of a fully observable system. All parameters are specified in the caption. In all the four cases, the recharge process parameters and the transmission energies are the same for both the source and the relay sensors. The four choices of $(0.6, 0.6)$, $(0.6, 0.9)$, $(0.9, 0.6)$ and $(0.9, 0.9)$ give an indication of the performance in diverse settings of low-low, low-high, high-low and high-high correlation probabilities at the relay. For each of the four cases, the parameters p^s_{on} and p^s_{off} of the event generation process at the source node are varied from 0.55 to 0.95. Overall, the quality of coverage is higher when the relay has a lower p^r_{on} and higher p^r_{off}. The percentage of relay usage in the four cases are shown in Table I. The table shows the overall quality of coverage (QoC) defined earlier along with the individual QoCs defined as the following: *Source QoC* is the ratio of the number of packets transmitted to the number



(a) $p_{on}^r = 0.6$, $p_{off}^r = 0.6$ (b) $p_{on}^r = 0.6$, $p_{off}^r = 0.9$ (c) $p_{on}^r = 0.9$, $p_{off}^r = 0.6$ (d) $p_{on}^r = 0.9$, $p_{off}^r = 0.9$

Fig. 1. Effect of p_{on} and p_{off} on the quality of coverage (z-axis) of a fully observable system and the theoretical upper bound. Parameters used: $q_{on}^s = q_{on}^r = 0.85$, $q_{off}^s = q_{off}^r = 0.7$, $c^s = c^r = 1$, $\delta_1^s = \delta_1^r = 2$, $\delta_2^s = \delta_2^r = 1$, $\theta^s = \theta^r = 1$.

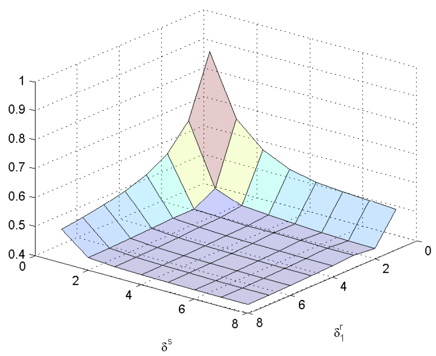


(a) $p_{on}^r = 0.6$, $p_{off}^r = 0.6$ (b) $p_{on}^r = 0.6$, $p_{off}^r = 0.9$ (c) $p_{on}^r = 0.9$, $p_{off}^r = 0.6$ (d) $p_{on}^r = 0.9$, $p_{off}^r = 0.9$

Fig. 2. Effect of p_{on} and p_{off} on the quality of coverage (z-axis) of a partially observable system and a fully observable system. Parameters used: $q_{on}^s = q_{on}^r = 0.85$, $q_{off}^s = q_{off}^r = 0.7$, $c^s = c^r = 1$, $\delta_1^s = \delta_1^r = 2$, $\delta_2^s = \delta_2^r = 1$, $\theta^s = \theta^r = 1$.

of packets generated by the source and *Relay QoC* is the ratio of the number of packets transmitted (its own) to the number of packets generated by the relay. *Relay usage* is defined as the ratio of the number of source packets transmitted using the relay to the total number of packets transmitted by the source. For cases $(0.6, 0.6)$ and $(0.9, 0.9)$, the steady-state probabilities of the event occurrence at the relay (π_{on}^r) are the same ($1/2$), but as the length of continuous events at the relay ($E(N)$, defined in Eqn. (1)) increases, the source tends to transmit the traffic directly as long as it has enough energy. When both $E(N)$ and π_{on}^r are low ($(0.6, 0.9)$ case), the relay is used intensively by the source. The reason behind this is that when the reward for transmitting one source packet equals that of transmitting one relay packet, from Eqn. (48) we see that the system reward is maximized when the relay has enough energy such that the relay and the source could transmit their own traffic, respectively, in the same slot. Thus, intuitively, as the traffic rate at the relay increases, the relay tends to use its energy for its own traffic. The QoC of a fully observable system is slightly higher than that of the partially observable system. The difference is larger when p_{off}^r is higher. This is so because when p_{off}^r is higher, the relay has more energy available most of the times, and thus can be used more often for its own transmissions or for relaying transmissions. In the MDP case, the sensor has complete information about the relay's state and can utilize the relay fully, whereas in the POMDP case, the source is not able to utilize the relay fully due to partial state information availability, resulting in larger performance difference.

Figure 3 shows how the different ratios between δ_1, c and



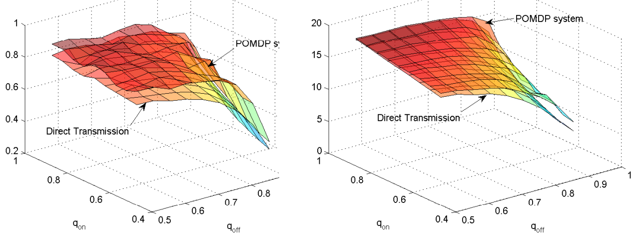
$c^s = 1$, $c^r = 1$

Fig. 3. Effect of δ_1 and c on the quality of coverage (z-axis) of a partially observable system. Parameters used: $\delta_2^s = \delta_2^r = 1$, $p_{on}^s = p_{on}^r = 0.85$, $p_{off}^s = p_{off}^r = 0.7$, $q_{on}^s = q_{on}^r = 0.85$, $q_{off}^s = q_{off}^r = 0.7$, $\theta^s = \theta^r = 1$.

δ_2 affect the quality of coverage of the partially observable system. The source node and the relay node have the same event generation and recharge process. We fix δ_2 as 1, and vary δ_1^s and δ_1^r from 1 to 8 with the recharge units c^s and c^r kept either low (1) or high (8) [3]. In the case tested, $\mu_{on} = 0.67$ and $\pi_{on} = 0.67$ for both the source and relay nodes. We see that when c^s/δ_2^s and c^r are fixed, the performance is mainly decided by δ_1^r, monotonically decreasing as δ_1^r increases, and slightly degrades as δ_1^s increases.

Our results show that when the traffic rate is fixed in both the source and the relay, as the steady state recharge probability of the source decreases, it uses the relay more frequently, thereby

[3]The result for high c^s is not shown due to limit of number of figures.



(a) Average throughput (b) Average battery level

Fig. 4. Effect of q_{on} and q_{off} on the average per node throughput and average per node battery level (z-axis) on a multi-node network with 50 nodes, for a partially observable relay based system and a network without cooperative communication. Parameters used: $\delta_1^s = \delta_1^r = 4$, $\delta_2^s = \delta_2^r = 1$, $p_{on}^s = p_{on}^r = 0.55$, $p_{off}^s = p_{off}^r = 0.95$, $c^s = c^r = 1$, $\theta^s = \theta^r = 1$.

increasing the quality of coverage and the theoretical bound gets tighter as well. When the recharge process is fixed in both nodes, as the steady state probability of events in the source node decreases, the source node transmits the packet directly. Finally, for a particular event generation and recharge process, if the source node has a much higher recharge unit than the relay node, it sends out the packets mostly by itself. In the opposite case, as the energy required for direct transmission goes up, the quality of coverage decreases and the ratio of events dealt with by relaying increases approximately linearly.

A. Multiple-node networks

To evaluate the proposed schemes when multiple nodes exists in the networks, we now consider a network with 30 and 50 nodes, spread randomly over a 1000×1000 meter region. The transmission range of each sensor is 100 meters. We consider one-hop traffic and each node picks one of its neighbors as the destination. From the set of neighbors that may serve as relays, the source picks one at random and uses it for transmissions. A node may serve as the source node for its packets (and use another node as a relay) and also as the relay or destination for other transmissions. The methodology described in Section VII is used by the nodes to determine if a direct or a relay-based transmission is to be used. The following backoff based MAC protocol was used by each node. Time is slotted with each slot having a fixed number of backoff slots at the beginning, followed by two mini-slots where the source and the relay may transmit data. At the beginning of each slot, nodes with packets select a random backoff value and transmit a request to send (RTS) packet once the counter decrements to zero, as long as none of its neighbors has transmitted so far. Collisions are handled by repeating the backoff procedure in the next time slot. The RTS packet also specifies the relay node chosen by the source. If the medium around the destination is free, it sends a clear to send (CTS) packet to the source. Additionally, if the medium around the relay is idle, it sends an acknowledgement (ACK) confirming its participation. In the following two mini-slots, the packet is transmitted by the source and the relay.

Figure 4(a) shows the per node throughput, averaged over all nodes in the network, for a network with 50 nodes. Here throughput is defined as the ratio of the number of packets successfully transmitted by a node to the total number of packets generated by the node. The figures also show the corresponding throughputs for the same networks when only direct transmissions are used. Note that the throughput metric is equivalent to the quality of coverage metric when multiple nodes are being considered. We observe that the proposed scheme performs better than just using direct transmissions. For another perspective at the energy saving achieved by using the proposed schedulers, Figure 4(b) shows the average battery levels at the sensor nodes for the relay based and direct transmission based networks. We see that the proposed scheme leads to higher battery levels at the nodes. Corresponding results for a network with 30 nodes are presented in the Supplementary Document.

IX. Conclusions

While WSNs are expected to facilitate new applications and transform many aspects of daily life, they are constrained by the limited onboard battery. This paper addressed the problem of developing transmission strategies for WSNs when energy harvesting devices are used by sensors to generate energy. We consider the case where a node may use either a direct transmission or a cooperative relay for its transmissions. A theoretical upper bound is obtained on the performance of any arbitrary strategy. Scheduling policies are then developed to choose the appropriate transmission mode depending on the available energy at the sensors as well as the states of their energy harvesting and event generation processes. We consider the cases where the state of the relay in terms of its battery level and the states of the energy harvesting and event processes is either fully or partially observable by formulating the problem as a MDP or POMDP, respectively. Results from our simulation study may be used towards developing a practical relay usage strategy.

References

[1] J. Laneman, D. Tse and G. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062-3080, December 2004.

[2] I. Akyildiz, W. Su, Y. Sankarasubramaniam and E. Cayirci, "Wireless Sensor Networks and Applications: a Survey," *Int. J. Comp. Sci. and Network Security*, vol. 7, no. 3, pp. 264-273, March 2007.

[3] J. Paradiso and T. Starner, "Energy scavenging for mobile and wireless electronics," *IEEE Perv. Comp.*, vol. 4(1), pp. 18-27, 2005.

[4] E. van der Meulen, "Three-terminal communication channels," *Adv. Appl. Prob.*, vol. 3, pp. 120-154, 1971.

[5] T. Cover and A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. 25, no. 5, pp. 572-584, Sep. 1979.

[6] A. Host-Madsen, "On the Capacity of Wireless Relaying," *Proc. IEEE VTC*,. 2002.

[7] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative Strategies and Capacity Theorems for Relay Networks," *Information Theory, IEEE Transactions on*, Vol. 51, Issue 9, pp. 3037- 3063, August 2005.

[8] L. Guo, X. Ding, H. Wang, Q. Li, S. Chen and X. Zhang , "Cooperative Relay Service in a Wireless LAN," *Selected Areas in Communications, IEEE Journal on*, Vol.25, Issue 2, pp.: 355-368, February 2007.

[9] X. Jiang, J. Polastre and D. Culler, "Perpetual environmentally powered sensor networks," *Proc. IPSN*, pp. 463-468, Apr. 2005.

[10] A. Kansal et. al., "Power management in energy harvesting sensor networks," *ACM Tran. Embedded Comp. Sys.*, Dec. 2007.

[11] C. Vigorito et. al., "Adaptive control of duty cycling in energy-harvesting wireless sensor networks," *Proc. IEEE SECON*, 2007.

[12] D. Niyato, E. Hossain and A. Fallahi, "Sleep and wakeup strategies in solar-powered wireless sensor/mesh networks: performance analysis and optimization," *IEEE Transactions on Mobile Computing,* vol. 6, no. 2, pp. 221-236, Feb. 2007.

[13] M. Tacca, P. Monti and A. Fumagalli, "Cooperative and reliable ARQ protocols for energy harvesting wireless sensor nodes," *IEEE Tran. Wireless Comm.,* vol. 6, issue 7, pp. 2519-2529, July 2007.

[14] N. Jaggi, K. Kar and A. Krishnamurthy, "Rechargeable Sensor Activation under Temporally Correlated Events," *Proc. WiOpt,* 2007.

[15] V. Sharma, U. Mukherji, V. Joseph, and S. Gupta, "Optimal energy management policies for energy harvesting sensor nodes," *IEEE Trans. Wireless Commun.,* vol. 9, no. 4, pp. 1326-1336, Apr. 2010.

[16] C. Ho and R. Zhang, "Optimal energy allocation for wireless communications powered by energy harvesters," *Proc. IEEE ISIT,* June 2010.

[17] A. Nosratinia, T. Hunter, A. Hedayat, "Cooperative communication in wireless networks," *IEEE Comm. Mag.,* vol. 42, no. 10, pp. 74-80, Oct. 2004.

[18] S. Jayaweera, "Virtual MIMO-based cooperative communication for energy-constrained wireless sensor networks," *IEEE Trans. Wireless Comm.,* vol.5, no.5, pp. 984-989, May 2006.

[19] H. Yang, H.-Y. Shen, B. Sikdar and S. Kalyanaraman, "A threshold based MAC protocol for cooperative MIMO transmissions," *Proceedings of IEEE INFOCOM minisymposium,* 2009.

[20] Y.-W. Hong and A. Scaglione, "Energy-efficient broadcasting with cooperative transmissions in wireless sensor networks," *IEEE Trans. Wireless Comm.,* vol.5, no.10, pp.2844-2855, Oct. 2006.

[21] Y.-W. Hong et. al., "Cooperative communications in resource-constrained wireless networks," *IEEE Signal Processing Magazine,* vol.24, no.3, pp.47-57, May 2007.

[22] P. Liu et. al., "CoopMAC: A cooperative MAC for wireless LANs," *IEEE JSAC,* vol.25, no.2, pp.340-354, Feb. 2007.

[23] N. Shankar, C.-T. Chou, and M. Ghosh,"Cooperative communication MAC (CMAC) - a new MAC protocol for next generation wireless LANs," *International Conference on Wireless Networks, Communications and Mobile Computing,* pp. 133-144, 2005.

[24] H. Li and B. Sikdar, "A Performance Guarantee for Maximal Schedulers in Sensor Networks with Cooperative Relays," to appear in *Proceedings of IEEE GLOBECOM,* 2010.

[25] M. Puterman, *Markov Decision Processes - Discrete Stochastic Dynammic Programming,* John Wiley and Sons, 1994.

[26] D. Bertsekas, *Dynamic Programming and Optimal Control,* Athena Scientific, Belmon MA, 2000.

[27] A. Cassandra, L. Kaelbling and M. Littman, "Acting optimally in partially observable stochastic domains," *Proceedings of the 12th National Conference on Artificial Intelligence*, vol. 2. pp. 1023-1028, Seattle, WA, 1994.

[28] E. Fernandez-Gaucherand, A. Arapostathis and S. Marcus, "On the average cost optimality equation and the structure of optimal policies for partially observable Markov decision processes," *Annals of Operations Research,* vol. 29, no. 1-4, pp. 439-470, April 1991.

[29] C. H. Papadimitriou and J. N. Tsitsiklis, "The Complexity of Markov Decision Processes", *Mathematics of Operations Research,* Vol. 12, No. 3, pp. 441 - 450, 1987.

[30] M. L. Littman, T. L. Dean and L. P. Kaelbling, "On the complexity of solving Markov decision problems", *in Proc. of the Eleventh International Conference on Uncertainty in Artificial Intelligence*, pp. 394 - 402, 1995.

[31] D. Burago, M. D. Rougemont and A. Slissenko, "On The Complexity of Partially Observed Markov Decision Processes", *Theoretical Computer Science*, Vol. 157, pp. 161 - 183, 1996.