Performance Modeling of Transmission Schedulers for Sensor Networks Capable of Energy Harvesting

Alireza Seyedi† and Biplab Sikdar†
†Department of ECE, University of Rochester, Rochester NY 14627 USA
‡ Department of ECSE, Rensselaer Polytechnic Institute, Troy, NY 12180 USA

***Abstract*— Energy harvesting is one of the most promising solutions for the enhancing the lifetime of sensor networks by overcoming the limitations of current batter technology. This paper investigates the performance of scheduling strategies for sensor networks with energy harvesting. The problem of selecting the power level at which a sensor should transmit is formulated as a Markov Decision Process (MDP) and the performance of the transmission policy thus derived is compared with that of an energy balancing policy as well as an aggressive policy. Our results show that the quality of coverage associated with the MDP formulation outperforms the other policies.**

I. Introduction

A major hurdle for the wide adoption of Wireless Sensor Networks (WSNs) is their energy supply. At present, the battery technology does not provide a high enough energy density to develop WSN nodes (for many applications such as health monitoring) with sufficiently long life and acceptable cost and form factor. The most promising approach to deal with the energy supply problem, particularly for sensors for medical applications, is energy harvesting or energy scavenging [8]. In this approach the nodes have an energy harvesting device that collects energy from ambient sources such as vibration and motion, light, and heat. However, to improve the performance of energy harvesting WSNs to acceptable levels, progress needs to be made both in energy harvesting techniques and communication policies and protocols.

This paper considers the problem of evaluating the performance of transmission schedulers in WSNs with energy harvesting devices. The sensors are assumed to have the ability to choose from a set of available transmission modes for transmitting their data, with each scheme consuming a different amount of energy. However, each scheme has a packet error probability that is a decreasing function of the energy used on transmission. For a given data packet, the schedulers select the appropriate transmission mode so that the probability that the sensor does not have any energy to report future events when they occur is minimized while maximizing the likelihood of data reports being correctly transmitted. This paper evaluates the performance of three scheduling policies using analytic models as well as simulations.

The rest of the paper is organized as follows. Section II presents the related work and Section III describes the system model. Sections IV and V evaluate the energy balancing and aggressive strategies. A MDP formulation of the problem is presented in Section VI, simulation results are presented in Section VII and Section VIII concludes the paper.

II. Related Work

The use of energy harvesting has been explored in the general framework of WSNs. The problem of duty-cycling in general sensor networks with energy harvesting is considered in [3], [6]. In [5], the authors show that using cooperative ARQ protocols, sensor nodes can match their energy consumption to their energy harvesting rate. The authors of [2] address the problem of sensor activation with battery recharging assuming temporally correlated events. However, none of the existing literature considers the problem of energy aware communication mode selection for wireless devices with energy harvesting.

III. System Model

We consider a discrete time model where time is slotted in intervals of unit length. Each slot is long enough to transmit one data packet and at most one data packet is generated in a slot. Each sensor has a rechargeable battery and an associated energy harvesting device. The energy generation process of the sensor is modeled by a correlated, two-state process. In its on state (i.e. when ambient conditions are conducive to energy harvesting), the sensor generates energy at a constant rate of c units in a 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}, with $0.5 < q_{on} < 1$, 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}, $0.5 < q_{off} < 1$, and energy is harvested with probability $1 - q_{off}$.

The data packets that sensors report to the sink are also generated according to a correlated, two-state process. If an event if generated in the current slot, another event is generated in the next slot with probability p_{on}, $0.5 < p_{on} < 1$, and no event is generated with probability $1 - p_{on}$. Similarly, if no event is generated in the current slot, no event is generated in the next slot with probability p_{off}, $0.5 < p_{off} < 1$, while an event is generated with probability $1 - p_{off}$. The average duration of a period of continuous events, $E[N]$, is given by

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

and the steady-state probability of event occurrence is

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

Similarly, the average length of a period without events is $\frac{1}{1-p_{off}}$ and $\pi_{off} = 1-\pi_{on}$. Along the same lines, 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.

In each slot, a sensor consumes δ_0 units of energy to run its circuits and additional energy is expended if the sensor decides to transmit data. Each sensor is assumed to have to capability to communicate at two transmission modes: "mode 1" consumes δ_1 units of energy on the modulation, coding and transmission and achieves an expected packet error rate of $1-\rho_1$ while "mode 2" consumes δ_2 units of energy with an expected packet error rate of $1-\rho_2$. We have $\delta_1 > \delta_2$ and $\rho_1 > \rho_2$ allowing a tradeoff between the energy consumed and reliability. For many applications it is more important to deliver the most recent data without delay rather than queue them behind retransmission attempts. Since data is generated in continuous bursts in our model, we thus assume that no retransmissions are attempted for packets with error. Also, a sensor is considered available for operation if its energy is greater that $\delta_0+\delta_2$. If a sensor's energy level falls below this threshold, it moves to the *dead* state where it is incapable of detecting and reporting events and stays there until it harvests enough energy. No energy in spent in the dead state.

The communication strategy of a sensor is governed by a policy Π that decides on the transmission mode to be used for reporting an event. The action taken by the sensor in time slot t is denoted by a_t with $a_t \in \{0,1,2\}$ denoting no transmission, and transmissions at mode 1 and 2, respectively. The decision may be based on the current battery level of the sensor and the states of the recharge as well as the event generation process, with the basic objective of maximizing the *quality of coverage*, defined as follows. Let $\mathcal{E}_o(T)$ denote the number of events that occurred in the sensing region of the sensor over a period of T slots in the interval $[0,T]$. Let $\mathcal{E}_d(T)$ denote the total number of events that are detected and correctly reported by the sensor over the same period under policy Π. The time average of the fraction of events detected and correctly reported represents the quality of coverage and is given by

$$U(\Pi) = \lim_{T\to\infty} \frac{\mathcal{E}_d(T)}{\mathcal{E}_o(T)} \tag{3}$$

IV. Energy Balancing Policies

In order to utilize the available energy efficiently, one strategy is to use an energy balancing (or energy neutral) policy, Π_{EB}, that assigns γ_1 and γ_2, the fraction of slots with data events in which the policy uses transmission mode 1 and 2 respectively, such that the total energy spent is equal to the energy generated, while maximizing the likelihood of detecting and reporting events without errors.

To develop an energy-balancing policy, first consider the data event generation process. This process strictly alternates between periods with events (*on state*) and periods without events (*off state*). The instances when the event process enters the off state can be considered renewal instants of the event process state. The expected length of a renewal period is

$$E[T_R] = \frac{1}{1-p_{on}} + \frac{1}{1-p_{off}} = \frac{2-p_{on}-p_{off}}{(1-p_{on})(1-p_{off})} \tag{4}$$

The expected energy generated in a renewal period is then

$$E[C] = \mu_{on} c E[T_R] = \mu_{on} c \frac{2-p_{on}-p_{off}}{(1-p_{on})(1-p_{off})} \tag{5}$$

The maximum possible energy that may be spent on running the on-board electronics of the sensor during a renewal period of $E[T_R]$ slots (i.e. ignoring slots in which the sensor is in the dead state) is $\delta_0 E[T_R]$. The expected amount of energy available for communications is thus at least

$$E[A] \geq \mu_{on} c E[T_R] - \delta_0 E[T_R] = \frac{(\mu_{on}c-\delta_0)(2-p_{on}-p_{off})}{(1-p_{on})(1-p_{off})} \tag{6}$$

The expected number of events to be reported and those correctly reported in a renewal period are $E[N]$ and $E[N](\gamma_1\rho_1+\gamma_2\rho_2)$, respectively. The number of events detected and correctly reported in the period $[0,T]$ is then

$$\mathcal{E}_d(T) = \frac{\gamma_1\rho_1+\gamma_2\rho_2}{1-p_{on}} \frac{T}{E[T_R]} \tag{7}$$

Also, the number of events generated in the period $[0,T]$ is $\mathcal{E}_o(T) = \pi_{on} T$. The performance of the policy is then

$$U(\Pi_{EB}) = \frac{\gamma_1\rho_1+\gamma_2\rho_2}{\pi_{on}(1-p_{on})} \frac{1}{E[T_R]} = \gamma_1\rho_1+\gamma_2\rho_2 \tag{8}$$

We then have the following bound on the performance of any energy balancing policy:

Claim 1: The performance of an energy balancing policy is bounded by

$$U(\Pi_{EB}) \geq \begin{cases} \Big[\frac{(\mu_{on}c-\delta_0)(\rho_1-\rho_2)}{\pi_{on}(\delta_1-\delta_2)} & \frac{\rho_1}{\delta_1} \leq \frac{\rho_2}{\delta_2} \text{ and} \\ -\frac{\pi_{on}(\delta_2\rho_1-\delta_1\rho_2)}{\pi_{on}(\delta_1-\delta_2)}\Big] & \frac{\delta_2}{1-p_{on}} \leq E[A] < \frac{\delta_1}{1-p_{on}} \\ \frac{\mu_{on}c-\delta_0}{\delta_2\pi_{on}}\rho_2 & \frac{\rho_1}{\delta_1} \leq \frac{\rho_2}{\delta_2} \text{ and } \frac{\delta_2}{1-p_{on}} > E[A] \\ \frac{\mu_{on}c-\delta_0}{\delta_1\pi_{on}}\rho_1 & \frac{\rho_1}{\delta_1} > \frac{\rho_2}{\delta_2} \text{ and } \frac{\delta_1}{1-p_{on}} > E[A] \\ \rho_1 & \frac{\delta_1}{1-p_{on}} \leq E[A] \end{cases} \tag{9}$$

where the relation for the last case holds with a strict equality.

Proof: We consider the four cases and prove the result by obtaining the optimal choice of γ_1 and γ_2 in each case.

Case I: $\left(\frac{\rho_1}{\delta_1} \leq \frac{\rho_2}{\delta_2} \text{ and } \frac{\delta_2}{1-p_{on}} \leq E[A] < \frac{\delta_1}{1-p_{on}}\right)$ In this case the energy available for communications is such that all attempts may be made using transmission mode 2 but not enough to make all transmissions using transmission mode 1. A linear programming formulation (LP1) for obtaining the γ_1 and γ_2 that maximizes the quality of coverage is given by

LP1: maximize $\quad \gamma_1\rho_1+\gamma_2\rho_2$

subject to $\quad \gamma_1+\gamma_2 \leq 1$ and $\frac{\gamma_1\delta_1+\gamma_2\delta_2}{1-p_{on}} \leq E[A]$

The objective function is maximized when the available energy is divided such that the largest number of transmissions with energy consumption δ_1 are accommodated while leaving enough energy to transmit in the remaining slots with events with energy consumption δ_2. Using $\gamma_2 = 1 - \gamma_1$, this implies

$$\gamma_1 E[N]\delta_1 + (1-\gamma_1)E[N]\delta_2 = E[A] \tag{10}$$

Solving for γ_1 gives

$$\gamma_1 = \frac{E[A]}{E[N](\delta_1-\delta_2)} - \frac{\delta_2}{\delta_1-\delta_2} = \frac{\mu_{on}c-\delta_0}{\pi_{on}(\delta_1-\delta_2)} - \frac{\delta_2}{\delta_1-\delta_2} \tag{11}$$

$$\gamma_2 = 1-\gamma_1 = \frac{\delta_1}{\delta_1-\delta_2} - \frac{\mu_{on}c-\delta_0}{\pi_{on}(\delta_1-\delta_2)} \tag{12}$$

The above values for γ_1 and γ_2 are achievable because in energy balancing policies the sensor always has energy to transmit, with probability one. To justify this claim, consider each sensor as a queue where the arrivals correspond to the energy harvested and the departures correspond to the energy spent. Thus the sensor represents a G/G/1 queue where the arrival rate equals the departure rate (due to the energy balancing property). The results of [7], page 422, then imply that the queue remains non-empty with probability one and the expected queue length becomes unbounded. This in turn implies that we always have enough energy to transmit data with probability one. The quality of coverage, $U(\Pi_{EB}) = \gamma_1\rho_1 + \gamma_2\rho_2$, is then given by

$$U(\Pi_{EB}) \geq \left[\frac{\mu_{on}c-\delta_0-\pi_{on}\delta_2}{\pi_{on}(\delta_1-\delta_2)}\right]\rho_1 + \left[\frac{\pi_{on}\delta_1-\mu_{on}c+\delta_0}{\pi_{on}(\delta_1-\delta_2)}\right]\rho_2$$
$$= \frac{(\mu_{on}c-\delta_0)(\rho_1-\rho_2)-\pi_{on}(\delta_2\rho_1-\delta_1\rho_2)}{\pi_{on}(\delta_1-\delta_2)} \tag{13}$$

The inequality arises because $E[A] \geq (\mu_{on}c-\delta_0)E[T_R]$.

Case II: $\left(\frac{\rho_1}{\delta_1} \leq \frac{\rho_2}{\delta_2} \text{ and } \frac{\delta_2}{1-p_{on}} > E[A]\right)$ In this case the available energy is not enough to report all events even when transmission mode 2 is used. Since $\frac{\rho_1}{\delta_1} \leq \frac{\rho_2}{\delta_2}$, i.e., the utility per unit energy is higher with transmission mode 2, no transmissions are made using transmission mode 1 in order to maximize the objective function. We prove the optimality of this strategy using contradiction. If all transmissions are made using transmission mode 2, $\frac{E[A]}{\delta_2}$ transmissions can be made resulting in an objective function of $\frac{E[A]}{\delta_2(1-p_{on})}\rho_2$. Assume now that there exists a policy that assigns k_1 of the $\frac{1}{1-p_{on}}$ slots with data events to transmission mode 1 with $k_1 > 0$ so that the resulting objective function is greater than $\frac{E[A]}{\delta_2(1-p_{on})}\rho_2$. Comparing the objective functions for the two cases

$$\frac{k_1}{1-p_{on}}\rho_1 + \frac{k_2}{1-p_{on}}\rho_2 \;>\; \frac{E[A]}{\delta_2(1-p_{on})}\rho_2 \tag{14}$$

$$\Rightarrow \quad k_1\rho_1 - \frac{k_1\delta_1\rho_1}{\delta_2} \;>\; 0 \quad \Rightarrow \quad \frac{\rho_1}{\delta_1} > \frac{\rho_2}{\delta_2} \tag{15}$$

This is a contradiction of the initial assumption of $\frac{\rho_1}{\delta_1} \leq \frac{\rho_2}{\delta_2}$ that defines Case II and the proof is thus complete. Thus in this case we have $\gamma_1 = 0$ and

$$\gamma_2 = \frac{E[A]}{\delta_2 E[N]} \geq \frac{(\mu_{on}c-\delta_0)(2-p_{on}-p_{off})}{\delta_2(1-p_{on})(1-p_{off})}(1-p_{on}) = \frac{\mu_{on}c-\delta_0}{\delta_2\pi_{on}} \tag{16}$$

The quality of coverage for this case is then given by

$$U(\Pi_{EB}) = \gamma_2\rho_2 \geq \frac{\mu_{on}c-\delta_0}{\delta_2\pi_{on}}\rho_2 \tag{17}$$

Case III: $\left(\frac{\rho_1}{\delta_1} > \frac{\rho_2}{\delta_2} \text{ and } \frac{\delta_1}{1-p_{on}} > E[A]\right)$ In this case the available energy is not enough to report all events using transmission mode 1. However, since $\frac{\rho_1}{\delta_1} > \frac{\rho_2}{\delta_2}$, the utility per unit energy is higher with transmission mode 1. Thus the solution for LP1 in this case assigns transmission mode 1 to all slots with data to be transmitted, as long as the available energy, L, satisfies $L \geq \delta_0 + \delta_1$. The optimality of this strategy can be proved using contradiction as was done for Case II. Assume that there exists a policy that transmits using transmission mode 2 during k_2 slots with $k_2 > 0$ so that its objective function for LP1 is greater than that of the policy that only uses transmissions using transmission mode 1. Thus

$$\frac{k_1}{1-p_{on}}\rho_1 + \frac{k_2}{1-p_{on}}\rho_2 \;>\; \frac{E[A]}{\delta_1(1-p_{on})}\rho_1 \tag{18}$$

$$\Rightarrow \quad k_2\rho_2 - \frac{k_2\delta_2\rho_1}{\delta_1} \;>\; 0 \quad \Rightarrow \quad \frac{\rho_1}{\delta_1} < \frac{\rho_2}{\delta_2} \tag{19}$$

This is a contradiction of the initial assumption of $\frac{\rho_1}{\delta_1} > \frac{\rho_2}{\delta_2}$ that defines Case III. Using the expected number of slots where a transmission can be attempted with energy consumption δ_1, we have

$$\gamma_1 = \frac{E[A]}{\delta_1 E[N]} \geq \frac{(\mu_{on}c-\delta_0)(2-p_{on}-p_{off})}{\delta_1(1-p_{on})(1-p_{off})}(1-p_{on}) = \frac{\mu_{on}c-\delta_0}{\delta_1\pi_{on}} \tag{20}$$

The quality of coverage for this case is then given by

$$U(\Pi_{EB}) = \gamma_1\rho_1 \geq \frac{\mu_{on}c-\delta_0}{\delta_1\pi_{on}}\rho_1 \tag{21}$$

Case IV: $\left(\frac{\delta_1}{1-p_{on}} \leq E[A]\right)$ Transmissions using transmission mode 1 are more likely to be successful and in this case the sensor has enough available energy to make all transmissions using this transmission mode. The solution to LP1 is thus trivial and we have $\gamma_1 = 1$ and $\gamma_2 = 0$. The quality of coverage for this case is thus

$$U(\Pi_{EB}) = \gamma_1\rho_1 = \rho_1 \tag{22}$$

This completes the proof. ∎

V. Aggressive Transmission Policies

A transmission policy that uses transmission mode 1 for all transmissions as long as the available energy $L \geq \delta_0 + \delta_1$ is termed an aggressive policy Π_A. Since $\gamma_2 = 0$ in an aggressive policy, the expected number of events correctly reported in a renewal period is $E[N]\gamma_1\rho_1$. The exact value of γ_1 depends on the system parameters and $E[A]$. The number of events detected and correctly reported in the period $[0, T]$ is then

$$\mathcal{E}_d(T) = \frac{\gamma_1\rho_1}{1-p_{on}}\frac{T}{E[T_R]} \tag{23}$$

Since the expected number of events generated in the period $[0, T]$ is $\mathcal{E}_o(T) = \pi_{on}T$, the performance of the aggressive policy is given by

$$U(\Pi_A) = \lim_{T\to\infty}\frac{\mathcal{E}_d(T)}{\mathcal{E}_o(T)} = \frac{\gamma_1\rho_1}{\pi_{on}(1-p_{on})}\frac{1}{E[T_R]} = \gamma_1\rho_1 \tag{24}$$

Next, we obtain a bound on the performance of the aggressive policy.

Claim 2: The performance of an aggressive transmission policy is bounded by

$$U_L(\Pi_A) \geq \begin{cases} \frac{\mu_{on}c-\delta_0}{\delta_1\pi_{on}}\rho_1 & \frac{\delta_1}{1-p_{on}} > E[A] \\ \rho_1 & \text{otherwise} \end{cases} \tag{25}$$

where the relation for the last case holds with a strict equality.

Proof: **Case I:** $\left(E[A] < \frac{\delta_1}{1-p_{on}}\right)$. The energy available in this case is not sufficient to transmit in all slots with events using transmission mode 1. Since the policy always schedules transmissions with transmission mode 1, we have

$$\gamma_1 = \frac{1}{E[N]}\frac{E[A]}{\delta_1} \geq \frac{(\mu_{on}c-\delta_0)(2-p_{on}-p_{off})}{\delta_1(1-p_{off})} = \frac{\mu_{on}c-\delta_0}{\delta_1\pi_{on}}$$

Then

$$U(\Pi_A) = \gamma_1\rho_1 \geq \frac{\mu_{on}c-\delta_0}{\delta_1\pi_{on}}\rho 1 \tag{26}$$

Case II: $\left(E[A] \geq \frac{\delta_1}{1-p_{on}}\right)$. Since there is enough energy to transmit all packets using transmission mode 1, the aggressive policy in this case results in $\gamma_1 = 1$ and $\gamma_2 = 0$. Thus

$$U(\Pi_A) = \gamma_1\rho_1 = \rho_1 = U(\Pi_{EB}) \tag{27}$$

This completes the proof. ■

Finally, we note that the sub-optimality of transmitting in all slots using transmission mode 1 in cases I and II of Claim 1 implies that the performance of the aggressive policy cannot exceed that of the energy balancing policy.

VI. Markov Decision Process Formulation

The solution to the problem of assigning the transmission mode for each communication event so that the quality of coverage is maximized can be also obtained by formulating it as a Markov Decision Process. Denote the system state at time t by $X_t = (L_t, E_t, Y_t)$ where $L_t \in \{0, 1, 2, \cdots\}$ represents the energy available in the sensor at time t, $E_t \in \{0, 1\}$ equals one if an event to be reported occurred at time t and zero otherwise. Also, $Y_t \in \{0, 1\}$ equals one if the sensor is being charged at time t and zero otherwise. The action taken at time t is denoted by $a_t \in \{0, 1, 2\}$ where $a_t = 0$ corresponds to no transmission, $a_t = 1$ corresponds to a transmission using transmission mode 1 and $a_t = 2$ corresponds to a transmission using transmission mode 2.

The next state of the system depends only on the current state and the action taken. Thus the system constitutes a Markov Decision Process. The sensor gains a reward of one with probability ρ_1 if $E_t = 1$ and $a_t = 1$ and a reward of one with probability ρ_2 if $E_t = 1$ and $a_t = 2$. The reward function is then given by

$$r(X_t, a_t) = \begin{cases} p_{on}\rho_1 & \text{if } a_t = 1,\ L_t \geq \delta_0 + \delta_1 \\ & \text{and } E_{t-1} = 1 \\ p_{on}\rho_2 & \text{if } a_t = 2,\ L_t \geq \delta_0 + \delta_2 \\ & \text{and } E_{t-1} = 1 \\ (1-p_{off})\rho_1 & \text{if } a_t = 1,\ L_t \geq \delta_0 + \delta_1 \\ & \text{and } E_{t-1} = 0 \\ (1-p_{off})\rho_2 & \text{if } a_t = 2,\ L_t \geq \delta_0 + \delta_2 \\ & \text{and } E_{t-1} = 0 \\ 0 & \text{otherwise} \end{cases} \tag{28}$$

Let g_t and l_t be the amount of energy gained and lost by the sensor in the interval $[t, t+1)$ respectively. Then

$$g_t = \begin{cases} c & \text{w.p. } Y_t q_{on} + (1-Y_t)(1-q_{off}) \\ 0 & \text{otherwise} \end{cases} \tag{29}$$

$$l_t = \begin{cases} \delta_0 + \delta_1 & \text{w.p. } [E_t p_{on} + (1-E_t)(1-p_{off})]I_1(a_t) \\ & \quad \text{if } L_t \geq \delta_0 + \delta_1 \\ \delta_0 + \delta_2 & \text{w.p. } [E_t p_{on} + (1-E_t)(1-p_{off})]I_2(a_t) \\ & \quad \text{if } L_t \geq \delta_0 + \delta_2 \\ \delta_0 & \text{w.p. } I_0(a_t) \text{ if } L_t \geq \delta_0 + \delta_2 \\ 0 & \text{otherwise} \end{cases} \tag{30}$$

where w.p. stands for "with probability", $I_A(a_t)$ represents the indicator function that equals one only when $a_t = A$ and zero otherwise. To complete the MDP formulation, the next state of the system $X_{t+1} = (L_{t+1}, E_{t+1}, Y_{t+1})$ is given by

$$L_{t+1} = L_t + g_t - l_t \tag{31}$$

$$E_{t+1} = \begin{cases} 1 & \text{w.p. } E_t p_{on} + (1-E_t)(1-p_{off}) \\ 0 & \text{otherwise} \end{cases} \tag{32}$$

$$Y_{t+1} = \begin{cases} 1 & \text{w.p. } Y_t q_{on} + (1-Y_t)(1-q_{off}) \\ 0 & \text{otherwise} \end{cases} \tag{33}$$

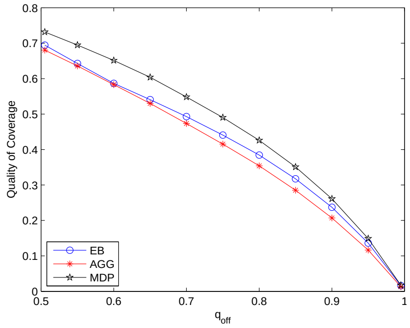
The optimal solution can be computed by using the well known value iteration technique [4]. The battery capacity of the sensor is assumed to be K. Since the induced Markov chain is unichain, from Theorem 8.5.2 of [4], there exists a deterministic, Markov, stationary optimal policy Π_{MD} which also leads to a steady-state transition probability matrix. Considering the average expected reward criteria, the optimality equations are given by [1]

$$h^*(X) = \max_{a\in\{0,1,2\}}\left[r(X,a) + \lambda^* + \sum_{X'=(0,0,0)}^{(K,1,1)} p_{X,X'}(a)h^*(X')\right]$$
$$\forall X \in \{(0,0,0), \cdots, (K,1,1)\} \tag{34}$$

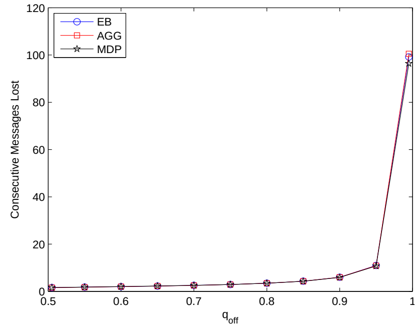
where $p_{X,X'}(a)$ represents the transition probability from state X to X' when action a is taken, λ^* is the optimal average reward and $h^*(i)$ are the optimal rewards when starting at state $i = (0,0,0), \cdots, (K,1,1)$. For the purpose of evaluation, the relative value iteration technique [1] is used to solve Eqn. (34).

VII. Simulation Results

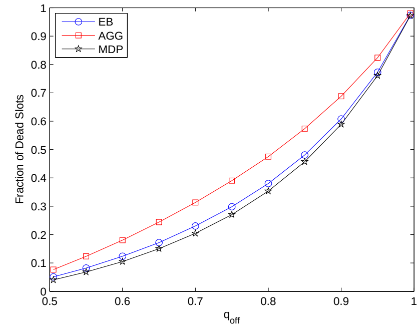
This section presents simulation results to compare the performance of of the three strategies. The simulations were



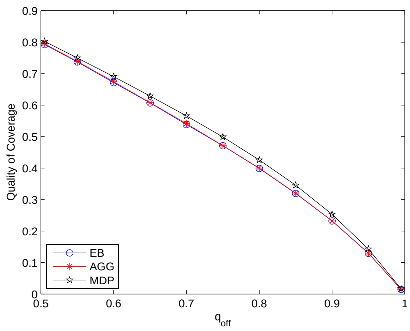
(a) $p_{on} = 0.6$, $\rho_2 = 0.4$, $\frac{\rho_1}{\delta_1} > \frac{\rho_2}{\delta_2}$



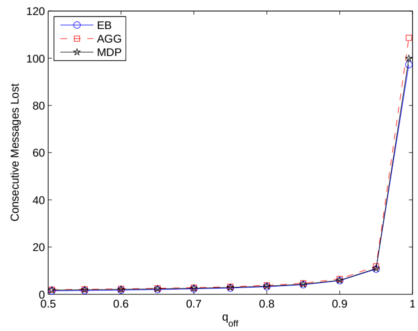
(b) $p_{on} = 0.6$, $\rho_2 = 0.4$, $\frac{\rho_1}{\delta_1} > \frac{\rho_2}{\delta_2}$



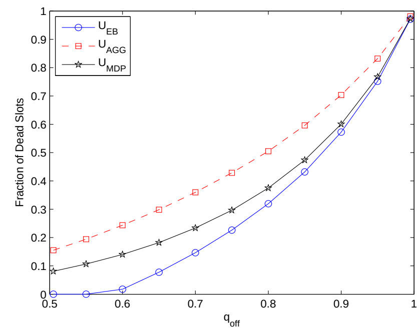
(c) $p_{on} = 0.6$, $\rho_2 = 0.4$, $\frac{\rho_1}{\delta_1} > \frac{\rho_2}{\delta_2}$



(d) $p_{on} = 0.7$, $\rho_2 = 0.6$, $\frac{\rho_1}{\delta_1} \leq \frac{\rho_2}{\delta_2}$



(e) $p_{on} = 0.7$, $\rho_2 = 0.6$, $\frac{\rho_1}{\delta_1} \leq \frac{\rho_2}{\delta_2}$



(f) $p_{on} = 0.7$, $\rho_2 = 0.6$, $\frac{\rho_1}{\delta_1} \leq \frac{\rho_2}{\delta_2}$

Fig. 1. Comparison of the performance of the three policies. Parameters used: $q_{on} = 0.75$, $p_{off} = 0.9$, $c = 2$, $\rho_1 = 0.9$, $\delta_0 = 1$, $\delta_1 = 2$ and $\delta_2 = 1$.

done using a simulator developed by us, primarily because energy harvesting is not well supported in existing simulators. The topology consisted of a single sensor and a sink. All simulations were run for a duration of 5000000 time units.

Figures 1(a) and 1(d) compare the performance of the three strategies (labeled EB: energy balancing, AGG: aggressive and MDP: Markov Decision Process) in terms of the quality of coverage as the recharge rate is varied by changing q_{off}. Two scenarios corresponding to $\frac{\rho_1}{\delta_1} \leq \frac{\rho_2}{\delta_2}$ and $\frac{\rho_1}{\delta_1} > \frac{\rho_2}{\delta_2}$ are considered. In both cases, the strategy obtained by the MDP outperforms the EB and AGG policies. The AGG policy performs worst because it always uses the mode with the higher energy consumption and is thus more likely to run out of energy.

Figures 1(b) and 1(e) compare the performance of the three strategies in terms of the average number of consecutive messages that are not successfully delivered by the sensor while Figures 1(c) and 1(f) compare the fraction of slots in which the sensor is in the dead state. The number of consecutive messages that are not delivered by the sensor is particularly important in certain medical applications. We observe that the performance of the three strategies is quite close though AGG has the worst performance. Also, while MDP performs better than EB when $\frac{\rho_1}{\delta_1} > \frac{\rho_2}{\delta_2}$, EB outperforms MDP in the other case. The smaller number of dead slots with EB when $\frac{\rho_1}{\delta_1} \leq \frac{\rho_2}{\delta_2}$ is because now EB transmits in more slots using transmission mode 2. While this decreases the fraction of dead slots, it does not necessarily result in better quality of coverage, as can be seen from Figure 1.

VIII. Conclusions

This paper developed models to study the performance of three transmission scheduling strategies for WSNs when energy harvesting devices are used by the sensors to generate energy. Simulation results show that while a strategy based on a MDP has better quality of coverage that both energy balancing and aggressive policies. In certain scenarios, the energy balancing policy may outperform the other two in terms of the number of dead slots and the average number of consecutive messages that are not reported correctly.

References

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

[2] N. Jaggi, K. Kar and A. Krishnamurthy, "Rechargeable Sensor Activation under Temporally Correlated Events," in Proc. *WiOpt*, Cyprus, Greece, April 2007.

[3] A. Kansal, J. Hsu, S. Zahedi, and M. B. Srivastava, "Power Management in Energy Harvesting Sensor Networks," *ACM Transactions on Embedded Computing Systems*, vol. 6, issue 4, December 2007.

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

[5] M. Tacca, P. Monti, and A. Fumagalli, "Cooperative and Reliable ARQ Protocols for Energy Harvesting Wireless Sensor Nodes," *IEEE Trans. on Wireless Comm.*, vol. 6, issue 7, pp. 2519-2529, July 2007.

[6] C. Vigorito, D. Ganesan and A. Barto "Adaptive Control of Duty Cycling in Energy-Harvesting Wireless Sensor Networks," in Proc. *IEEE SECON*, San Diego, CA, June 2007.

[7] R. Wolf, *Stochastic Modeling and the Theory of Queues*, Prentice Hall, 1989.

[8] G. Z. Yang, *Body Sensor Networks*, Springer, London, 2006.