Transmit Power Control with ARQ in Energy Harvesting Sensors: A Decision-Theoretic Approach

Anup Aprem, Chandra R. Murthy, *Senior Member, IEEE*, and Neelesh B. Mehta, *Senior Member, IEEE*

***Abstract*—This paper addresses the problem of finding optimal power control policies for wireless energy harvesting sensor (EHS) nodes with automatic repeat request (ARQ)-based packet transmissions. The EHS harvests energy from the environment according to a Bernoulli process; and it is required to operate within the constraint of energy neutrality. The EHS obtains partial channel state information (CSI) at the transmitter through the link-layer ARQ protocol, via the ACK/NACK feedback messages, and uses it to adapt the transmission power for the packet (re)transmission attempts. The underlying wireless fading channel is modeled as a finite state Markov chain with known transition probabilities. Thus, the goal of the power management policy is to determine the best power setting for the current packet transmission attempt, so as to maximize a long-run expected reward such as the expected outage probability. The problem is addressed in a decision-theoretic framework by casting it as a partially observable Markov decision process (POMDP). Due to the large size of the state-space, the exact solution to the POMDP is computationally expensive. Hence, two popular approximate solutions are considered, which yield good power management policies for the transmission attempts. Monte Carlo simulation results illustrate the efficacy of the approach and show that the approximate solutions significantly outperform conventional approaches.**

***Index Terms*— Energy harvesting sensors, Power control, ARQ, Retransmission, POMDP.**

I. Introduction

Wireless energy harvesting sensors (EHS) operate using energy harvested from environmental sources such as solar, wind, piezoelectric, etc. Due to their promise of a potentially infinite life of operation, they are fast emerging as viable options for a variety of sensing-related applications. However, due to the sporadic/random nature of the harvesting process, energy management becomes critical to ensure continuous and reliable operation of these nodes.

The energy replenishment process of the natural phenomena and the energy storage constraints of the node need to be taken into consideration when designing efficient transmission strategies. This requires a cross-layer, energy-aware protocol design that optimizes energy consumption for reliable packet transmission as a function of the statistics of the harvested energy. Transmission policies for EHS have to satisfy the constraint of *energy neutrality:* at every point in time, the total energy consumed by a node cannot exceed the total energy harvested by it. In addition, the transmission policy of the EHS needs to be able to handle the randomness in the channel fading process. This motivates the use of an automatic repeat request (ARQ)-protocol, as it resilient to channel variations and is known to be energy efficient [1].

The problem of transmit power management in EHS nodes was studied in [2], where deterministic models for energy harvesting were proposed. A simple Bernoulli injection model, in which the node harvests a fixed amount of energy with some probability or does not harvest at all, was proposed in [3]. Other models and metrics for EHS in the literature include minimizing the transmission time [4]; maximizing the short term throughput [5] or the quality of coverage [6]; and throughput and delay-optimal policies [7].

In this work, we study the problem of finding optimal power management policies with ARQ-based packet transmission, where the EHS node makes at most K attempts to transmit each packet. After each attempt, the node receives an ACK/NACK depending on whether the packet is successfully received or not. In the latter case, it retransmits the packet. If the packet is not successfully delivered after the K attempts, it is declared to be in outage. Our goal, in this context, is to find the optimal transmit power for each packet transmission attempt, given the history of packet transmission energy levels, the ACK/NACK messages received, and the energy available in the battery. Intuitively, receiving an ACK or NACK on a packet transmitted at a certain power level provides the EHS implicit information about the channel state, which can be used to reduce the odds that the retransmitted packet also suffers an outage. The knowledge of the time correlation of the channel can also be exploited to judiciously increase or decrease the power for the next transmission attempt. The key tradeoff involved here is that increasing the transmission power improves the odds of successful packet reception, but drains energy from the battery and decreases the probability that there will be sufficient energy to deliver future packets. On the other hand, a conservative approach of transmitting at a low or minimal power could lead to packet outages and wastage of energy if the battery gets full and energy arrivals occur after an ACK has been received.

An ad hoc approach to power management in EHS nodes with fixed-size packets and ARQ-based retransmission was studied in [8], where all the packets were transmitted at the same power level. The packet transmission power was restricted to be an integer multiple or integer fraction of the harvested power. The wireless channel to the destination either remained constant for all the transmission attempts of a given packet, or it changed independently from one attempt to the next. The outage probability analysis derived by the

The authors are with the Dept. of ECE, Indian Institute of Science, Bangalore, India. Emails: {aaprem, cmurthy, nbmehta}@ece.iisc.ernet.in

This work was financially supported by a research grant from the Aerospace Network Research Consortium.

978-1-4673-0921-9/12/$31.00 ©2012 IEEE

authors was useful in demonstrating that it is important to correctly tune the communication parameters to obtain the best possible link performance. However, the complexity of the formulation made direct optimization of the parameters analytically intractable. Also, the use of a fixed transmission power precluded the EHS from exploiting the time correlation of the channel or the information from previous ACK/NACK messages in designing its transmission strategy. In this paper, we overcome all of the aforementioned drawbacks.

A natural framework within which to handle the above power management problem at the EHS is to formulate it as a sequential decision process. Since channel state information (CSI) is only partially available at the EHS through the ACK/NACK information, we model the problem as a partially observable Markov decision process (POMDP) [9]. This makes our approach fundamentally different from past work employing decision theory for EHS nodes to obtain decision policies [6], [7], [10]–[12]. To the best of our knowledge, this is the first time in the literature that the power management problem in EHS with ARQ-based packet retransmissions has been cast in a POMDP framework. Our study leads to useful insights about optimal link-layer protocols for EHS nodes. Moreover, packet retransmissions and power control are already enabled in present-day low power communication standards such as the IEEE 802.15.4 [13].

The POMDP formulation allows us to choose from a gamut of available techniques for finding optimal and near-optimal power management policies. Due to the large size of the state-space, the exact solution to POMDP turns out to be computationally infeasible. Hence, we consider two heuristic solutions to the POMDP: the voting policy and the maximum likelihood policy. Through simulations, we show that these policies significantly outperform existing schemes.

The organization of the paper is as follows. Section II introduces the system model and the problem definition. Section III presents the POMDP formulation of the power management problem. Section IV discusses the approximate solutions of the POMDP. Section V presents simulation results, and concluding remarks are offered in Section VI.

II. System Model

Consider an EHS node that transmits a packet of ℓ data bits periodically once in a *frame* of duration T_m (s) to a destination node over a wireless fading channel. Each packet transmission attempt happens during a *slot* of duration T_p (s). The slot duration includes sending the packet and receiving an ACK from the destination. Time is discretized in multiples of T_p, and let the integer $K \triangleq \lfloor T_m/T_p \rfloor$ represent the maximum number of transmission attempts in a frame. An Automatic Repeat Request (ARQ) protocol is assumed at the link-layer. If the EHS receives no acknowledgment (NACK) from the receiver, it retransmits the packet until it receives an ACK, or it is time to transmit the next packet. If it runs out of energy, it suspends transmission till it harvests sufficient energy to attempt transmission again. If it receives an ACK, the node stops transmitting and just accumulates the energy harvested during the rest of the frame. An outage event is said to occur if the receiver fails to successfully decode the packet within the interval T_m. Apart from the ACK/NACK, no CSI is assumed to be available at the EHS. Fig. 1 illustrates the packet transmission model.

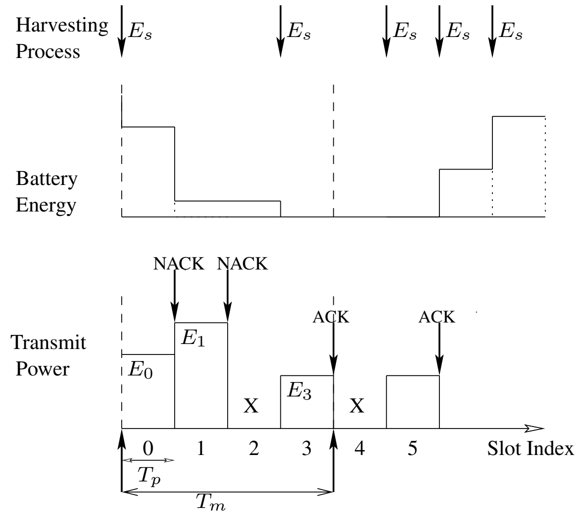

Fig. 1. Transmission timeline of the EHS for $K = 4$, showing the random energy harvesting process $(\downarrow)$ and periodic data arrivals $(\uparrow)$. The marker "X" denotes slots where the EHS does not transmit.

For the energy harvesting process, an independent and identically distributed Bernoulli model is considered, in which an energy E_s is injected into the EHS node at the beginning of every slot with probability ρ, and with probability $1 - \rho$, no energy is harvested [3]. Alternative models for the energy harvesting process include the leaky-bucket model [2], switch-based or vibration-based models [14], etc. While the above Bernoulli model is simple, it does capture the sporadic and random nature of energy availability at the EHS.

A finite energy buffer (e.g., a battery) is used to store the harvested energy, and it is assumed that there are no storage inefficiencies in the buffer. Let B_i denote the battery energy level at the beginning of the i^{th} slot, and let $E_i \leq B_i$ denote the energy used for packet transmission. The battery energy itself gets replenished whenever the node harvests energy, and, consequently, obeys the following Markovian evolution:

$$B_{i+1} = \begin{cases} \min\left(B_i + E_s - E_i, B_{\max}\right) & \text{w. prob. } \rho \\ B_i - E_i & \text{w. prob. } 1-\rho \end{cases} \quad (1)$$

where $B_{\max}$ denotes the battery capacity.

For the foregoing POMDP formulation, we need the state-space to be finite. To facilitate this, we discretize the channel into N levels, $\gamma_1, \ldots, \gamma_N$. In the *correlated channel model*, we model the channel as the finite state Markov chain (FSMC) shown in Fig. 2, with known state transition probabilities[1] $P_{k,k+1}, P_{k,k-1}$ and $P_{k,k}$. Such a first-order model is known

[1] In this paper, for convenience, we use the notations $P_{i,j}$ and P_{γ_i,γ_j} interchangeably, to represent the probability of going from state γ_i to γ_j.

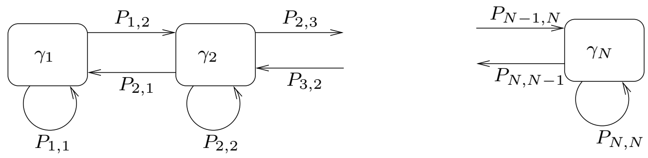

Fig. 2. Finite state Markov model for Rayleigh fading channel

to be accurate for packet-level simulation studies [15] and in cross-layer optimization [16], [17] with slowly-varying channels. The channel levels $\gamma_1, \ldots, \gamma_N$ and the transition probabilities can be computed based on the underlying fading distribution and Doppler frequency, following the procedure in [18], [19]. In the *block fading channel model*, we assume that the channel remains fixed for the duration of a frame, and changes independently and statistically identically from one frame to the next. This assumption is valid when the channel coherence time equals the frame duration, due to which, the initial packet transmission and all of the retransmission attempts see the same channel state [8]. To facilitate comparisons, we assume that the stationary distribution of the quantized channel is the same for the two channel models.

In a slot, the receiver may fail to decode the packet if the EHS node does not have sufficient energy to transmit, or if the transmitted packet is corrupted by the channel or noise at the receiver. The packet error probability depends on the modulation and coding scheme (MCS) used for packet transmission. For example, with uncoded BPSK modulation at a transmit energy E_i per bit, and when the channel state is γ_k, the packet error probability can be written as

$$P_e(E_i; \gamma_k) = 1 - \left(1 - Q\left(\sqrt{\frac{2\gamma_k E_i}{N_0}}\right)\right)^{\ell} \quad (2)$$

where Q is the standard Gaussian tail function and N_0 is the noise power spectral density. The extension to other MCS is straightforward provided one has an expression that analytically relates the packet error probability to the transmit energy and the channel state.

The next section presents the POMDP formulation of the problem considered in this paper.

III. POMDP Formulation

Our goal is to sequentially select the packet transmit power levels based on the retransmission index, battery energy level, and the history of transmission energies and ACK/NACK messages received, to minimize the expected outage probability. The channel state γ_i in the i^{th} slot is only partially observed at the EHS transmitter through the ACK/NACK messages. Hence, the sequential decision problem is cast as a partially observable Markov decision process (POMDP). For convenience, we normalize all energy values with respect to a minimum possible transmit energy E, which is typically imposed by the lower end of the linearity range of RF amplifier on the EHS node. Let $L \triangleq E_s/E$ be an integer with $L \geq 1$. The POMDP consists of the following components:

a) State-Space: The finite set of system states denoted by $\mathcal{S} \triangleq \mathcal{B} \times \mathcal{G} \times \mathcal{K} \times \mathcal{U}$, where

- $\mathcal{B} \triangleq \{0, 1, \ldots, B_{\text{max}}\}$ is the set of battery states, normalized with respect to the minimum transmit energy E. Recall that B_{max} is the battery capacity.
- $\mathcal{G}$ is the set of channel states. Under the channel model explained in the previous section, $\mathcal{G} \triangleq \{\gamma_1, \gamma_2, \ldots, \gamma_N\}$.
- $\mathcal{K}$ is the set of packet transmission attempt indices within a frame, and, hence, $\mathcal{K} \triangleq \{0, 1, \ldots, K-1\}$.
- $\mathcal{U} \triangleq \{0, 1\}$ is set of *packet reception states*. The packet reception state takes the value 1 when an ACK is received by the EHS, and is 0 otherwise. At the beginning of the frame, i.e., when the packet transmission attempt index is $k = 0$, the EHS node is always in the packet reception state 0, since an ACK has not yet been received. If the receiver successfully decodes the current packet and an ACK is received by the EHS, the packet reception state changes to 1 for the rest of the frame. Irrespective of the system state at $k = K-1$, the packet reception state is reset to 0 at the beginning of the next slot, as it corresponds to the beginning of a new frame.

b) Observation Space: The observations are the ACK/NACK messages received by the EHS node, after each packet transmission attempt. The observation space is the finite set $\mathcal{O} \triangleq \{\text{ACK}, \text{NACK}\}$. Since errors in the ACK/NACK messages are not modeled in this work, the observation always matches the packet reception state of the system.

c) Action Space: An *action* a by the EHS node corresponds to sending a packet at power level aE. The action space is the set of possible actions, and is denoted by $\mathcal{A} \triangleq \{0, 1, \ldots, B\}$, with $B \in \mathcal{B}$ representing the battery level in the current slot.

d) State Transition Function: Let two arbitrary states in $\mathcal{S}$ be $s \triangleq (b, \gamma, k, u)$ and $s' \triangleq (b', \gamma', k', u')$. The state transition function is the probability that the system starts in state s, takes an action a, and lands in state s'. Under the correlated fading channel model, it is given by

$$\mathcal{T}(s, a, s') = \delta(k', k_+)\, P_{\gamma,\gamma'}\, \psi\left((b, u), a, (b', u'), k, \gamma\right) \quad (3)$$

where $k_+ \triangleq (k+1) \bmod K$, $\delta(k', k)$ is the Kronecker delta function, and $P_{\gamma,\gamma'}$ is the channel transition probability as defined in Section II. In the above, the term $\delta(k', k_+)$ captures the fact that the packet transmission index always increases by one at a time until the end of the frame, when it resets to 0. Also, $\psi\left((b, u), a, (b', u'), k, \gamma\right)$ represents the probability that the EHS node starts from battery state b and packet reception state u, takes an action a, and lands in the state (b', u') when the current channel state and packet transmission index are γ and k, respectively. An expression for this term is provided in the next page. In the block fading model, the state transition function is given by

$$\mathcal{T}(s, a, s') = \delta(k', k_+)\, \zeta(\gamma, \gamma'; k)\, \psi\left((b, u), a, (b', u'), k, \gamma\right)$$

where $\zeta(\gamma, \gamma'; k) = \delta(\gamma, \gamma')$ when $k \neq K-1$, i.e., when not at the end of the frame, and $= \pi_{\gamma'}$, otherwise. Here, $\pi_{\gamma'}$ represents the stationary probability of the channel state γ'.

In the previous page, $\psi\left((b,u),a,(b',u'),k,\gamma\right)$ is as follows. Define

$$\eta(b,a,b') \triangleq \rho\delta(b',b+L-a)+(1-\rho)\delta(b',b-a). \quad (4)$$

Then, if $k = K-1$, it can be shown that $\psi\left((b,u),a,(b',u'),K-1,\gamma\right) = \eta(b,a,b')$ when $u'=0$, and is 0 otherwise. If $k \neq K-1$, it can be shown that $\psi\left((b,u),a,(b',u'),k,\gamma\right)$

$$= \begin{cases} \eta(b,a,b') & u'=1,u=1 \\ \eta(b,a,b')(1-P_e(aE;\gamma)) & u'=1,u=0 \\ \eta(b,a,b')P_e(aE;\gamma) & u'=0,u=0 \\ 0 & \text{else.} \end{cases} \quad (5)$$

The above expression is obtained by tracking the probabilities of the following events: i) Whether energy has been harvested in the current slot or not; ii) The packet reception state of the system; iii) The probability of successful packet reception at the receiver given the channel state and action. Also, the expression in the $k=K-1$ case arises because the packet reception state always resets to zero at the end of the frame. For example, to get the first term in (5), note that, when the EHS has already received an ACK, it transits from battery state b to $b+L-a$ upon taking an action a if it harvests energy (which occurs with probability ρ). If it does not harvest energy, it transits to the state $b-a$ (which occurs with probability $1-\rho$).

e) Observation Function: The observation function is the probability of observing an ACK or a NACK given the current state and action. Since this probability depends only on the current channel state and action, it is given by

$$\begin{aligned} P\left(\text{NACK}|a,\gamma\right) &= P_e(aE;\gamma) \\ P\left(\text{ACK}|a,\gamma\right) &= 1-P_e(aE;\gamma) \end{aligned} \quad (6)$$

where $P_e(aE;\gamma)$ is given by (2).

f) Reward: Let $s \triangleq (b,\gamma,k,u)$ be the state of the system. The expected immediate reward is defined as:

$$\mathcal{R}(s,a) = \begin{cases} 1-P_e(aE;\gamma) & a\le b, u=0 \\ -1 & (a>b,u=0) \text{ or } (a\neq 0, u=1) \\ 0 & \text{else.} \end{cases} \quad (7)$$

The immediate reward of -1 is used to preempt the EHS from using a non-zero energy when it has already received an ACK, or from attempting to use more than the energy available in the battery, to transmit a packet.

g) Objective: The objective is to maximize the expected reward collected by the EHS node over an infinite time horizon, and is given by

$$J = \lim_{m\to\infty}\frac{1}{m}\mathbb{E}\left\{\sum_{n=1}^{m}\mathcal{R}(s_n,a_n)\right\} \quad (8)$$

where $n \in \{1,2,\dots\}$ denotes the slot index, s_n is the state sequence, and a_n is the action sequence. The expectation in the above is over the distributions of the channel and energy harvesting processes. The next section discusses the techniques for solving the POMDP considered in this work.

IV. Solution Techniques

In a POMDP, the system state s is not fully observable. However, given the history of actions and observations, a so-called *belief state* $\beta(s)$ can be computed, that represents the probability that the system is in state s. It is known that the belief state is a sufficient statistic for finding the optimal policy [20]. The belief state can be updated at the end of each slot based on the previous belief state, the current observation, and the transition probability matrix. Then, the solution of the POMDP is the solution of a fully observable Markov decision process (MDP) on the belief states [9].

Although several exact algorithms [9] for solving the belief MDP exist, these algorithms are computationally feasible only when the cardinality of the state-space is of the order of ten or so [21], [22]. Even approximate solution methods can only handle a state-space with a cardinality of about a hundred [23]. In our case, the state-space is much larger, as it is indexed by the number of battery energy levels, the number of channel states, the packet retransmission index, and the ACK/NACK state. As a result, finding an exact or even approximate solution to the POMDP is computationally infeasible. Hence, we explore two popular heuristic solutions for the POMDP. For this, we first describe the solution to the MDP that is obtained when the system state is fully observable; both the heuristic methods rely on solving this underlying MDP.

The solution to the MDP when the system state is fully observable yields an optimal *policy*, μ^*_{MDP}, which prescribes the mapping from the state-space $\mathcal{S}$ to the action space $\mathcal{A}$ that maximizes the expected long-term reward J defined in (8). The optimal policy is the solution to the following Bellman equation [24]

$$\lambda^* + h^*(s) = \max_{a\in\mathcal{A}, a\le B(s)}\left[\mathcal{R}(s,a)+\sum_{s'\in\mathcal{S}}\mathcal{T}(s,a,s')h^*(s')\right] \quad (9)$$

for all $s\in\mathcal{S}$, where λ^* is the optimal average reward and h^* is an optimal reward vector satisfying the Bellman equation. Here, with a slight abuse of notation, the battery energy level is written as $B(s)$ to indicate that it is one of the components of the system state s.

The value iteration method [9] can be used to solve the Bellman equation (9). This involves iteratively solving

$$J_{k+1}(s) = \max_{a\in\mathcal{A}, a\le B(s)}\left[\mathcal{R}(s,a)+\sum_{s'\in\mathcal{S}}\mathcal{T}(s,a,s')J_k(s')\right] \quad (10)$$

for all $s\in\mathcal{S}$, where J_k is the value function at the k^{th} iteration, $k=0,1,\dots$. It can be shown that [24]

$$\lim_{k\to\infty}\frac{J_k(s)}{k} = \lambda^*, \quad \forall s\in\mathcal{S}. \quad (11)$$

In practice, it is standard to use relative value iteration to solve the MDP, which is a numerically stable version of the above procedure [24]. We denote the solution to the MDP obtained using the relative value iteration as $\mu^*_{\text{MDP}}(s)$.

The next step is to use the solution to the MDP obtained in the fully observable case to solve the original POMDP. Recall that, in our system model, the battery state, packet reception state and the transmission attempt index are fully observable, while the channel state component of the system state is only partially observable, through the ACK/NACK messages. Therefore, we maintain a belief $\beta(\gamma)$ over the channel state of the system, and use it to approximately solve the POMDP. The update equations for the belief $\beta(\gamma)$ are obtained as follows.

Let $o_n \in \mathcal{O}$ be the observation and let γ_n denote the channel state component of the system state s_n at time-slot n, i.e., $s_n = (b_n, \gamma_n, k_n, u_n)$. Let $\beta_n(\gamma)$ denote the belief of the channel state $\gamma \in \mathcal{G}$ at time-slot n. In the correlated fading model, the belief state update is given by

$$\beta_n(\gamma_j) = \frac{\sum_i P_{\gamma_i,\gamma_j} P(o_{n-1}|a_{n-1},\gamma_i)\beta_{n-1}(\gamma_i)}{\sum_j \sum_i P_{\gamma_i,\gamma_j} P(o_{n-1}|a_{n-1},\gamma_i)\beta_{n-1}(\gamma_i)}, \quad (12)$$

for $j = 1, 2, \ldots, N$, where $P(o_n|a_n, \gamma_j)$ is given by (6). In the block fading model, P_{γ_i,γ_j} in the above is replaced by $\zeta(\gamma_i, \gamma_j; k)$ defined in the previous section, and one also needs to keep track of the retransmission index for the current packet, k.

The final task is to use the channel belief state obtained above to convert the POMDP to an MDP, and use the solution to the MDP as an approximate solution the POMDP. To this end, we present two popular heuristic approaches.

1) Maximum Likelihood (ML) Heuristic [25]: Here, at each instant of time, we find the most probable channel state (i.e., the largest entry of the belief vector over the channel states), $\gamma_{\text{ML}} \triangleq \arg\max_{\gamma \in \mathcal{G}} \beta(\gamma)$ of the system. Then, the ML state of the system is defined as $s_{\text{ML}} \triangleq (b, \gamma_{\text{ML}}, k, u)$, where b, u and k are the current battery, packet reception and slot state, respectively. The ML heuristic method adopts the action corresponding to the solution of the MDP with the ML state as the solution of the POMDP. Thus,

$$\mu_{\text{ML}} = \mu^*_{\text{MDP}}(s_{\text{ML}}). \quad (13)$$

2) Voting Policy Heuristic [26]: Here, for a given b, u and k, we consider the set of states $s = (b, \gamma, k, u)$, $\gamma \in \mathcal{G}$. Each of these states votes for an action a, as determined by the optimal policy $\mu^*_{\text{MDP}}(s)$ of the underlying MDP corresponding to that particular state. In any given time-slot, these votes are weighed by the component of the belief state corresponding to each $\gamma \in \mathcal{G}$, and the sum of the weighted votes for each action is determined. The action with the largest sum, denoted by μ_{voting}, is selected as the optimal action:

$$\mu_{\text{voting}} = \arg\max_{a \in \mathcal{A}} \sum_{\substack{s=(b,\gamma,k,u)\\ \gamma \in \mathcal{G}}} \beta(s)\delta(\mu^*_{\text{MDP}}(s), a). \quad (14)$$

The solution to the POMDP can be computed offline. The online computations involve updating the belief using (12) and adopting the heuristic policy in (13) or (14) corresponding to the updated belief state. Hence, the computational overhead in implementing the above power control policy is not a burden on the EHS.

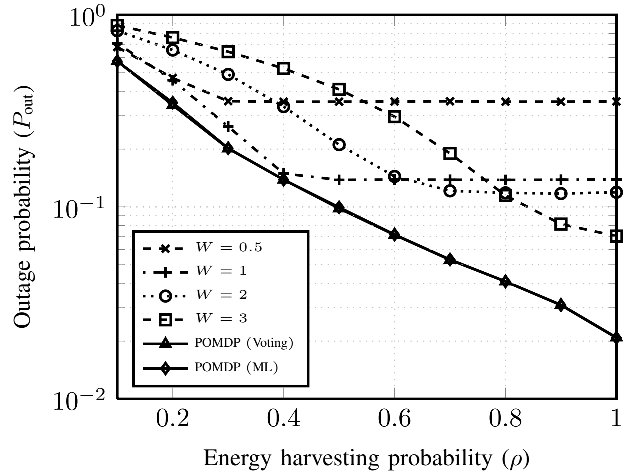

Fig. 3. Block fading channel: Comparison of ML and voting policy heuristic solutions against fixed-power transmission schemes.

V. Simulation Results

In this section, we compare the performance of the heuristic algorithms described in Sec. IV against the conventional fixed-power retransmission scheme, through Monte Carlo simulations. Let the fixed power be denoted by E_w and define $W \triangleq E_w/E_s$. For comparison with past work [8], we assume that W is an integer multiple or integer fraction of E_s.

Figure 3 shows the simulation results for the block fading channel case. The following system parameters are used: $K = 4$, $L = 4$, $N = 7$, $\ell = 50$, $E_s = 12$ dB, $N_0 = 1$, and $B_{\max} = 20E_s$. In order to enable a comparison with the results presented in [8], in this experiment, we use the uncoded on-off keying modulation. In this case, the packet outage probability in (2) is replaced with $P_e(aE;\gamma) = 1-(1-Q(\sqrt{aE\gamma/N_0}))^{\ell}$. The outage probability of the system is plotted against the probability of harvesting energy in a slot, ρ. The heuristic solution of POMDP is compared against fixed-power scheme with $W = \frac{1}{2}, 1, 2$ and 3. The voting policy and ML policy perform almost equally well, and significantly outperform the fixed-power transmission scheme.

Figure 4 shows the simulation results for the correlated channel case. The system parameters are the same as in Fig. 3, except for $K = 3$ and $B_{\max} = 10E_s$. To capture the time correlation, $T_p = 10$ ms and $f_d T_p = 0.03$ are used as the parameters of the FSMC, where f_d denotes the Doppler frequency. Again the outage probability of the system is plotted against ρ. As before, the ML heuristic and the voting policy heuristic perform nearly equally well and outperform the fixed-power transmission scheme. The ML heuristic is simple to implement and would, therefore, be a good choice for an EHS, given its limited hardware capabilities. Typically, to achieve the same outage probability, the POMDP solution requires only about 80% of the average energy harvesting rate required by the fixed-power scheme. Finally, Fig. 5 shows how the outage probability varies as a function of ρ with various battery capacities. The plot highlights the role of the battery capacity in improving the outage probability.

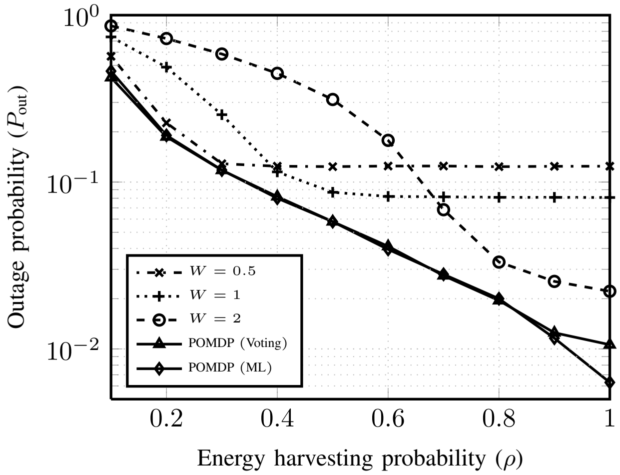

Fig. 4. Correlated channel: Comparison of ML and Voting policy heuristic solutions against fixed-power transmission schemes.

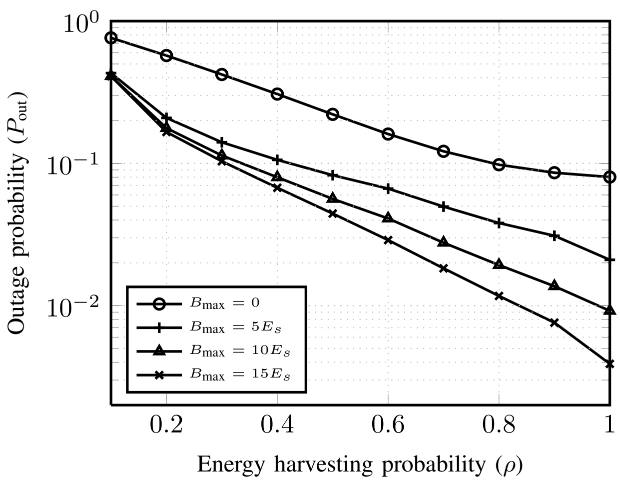

Fig. 5. Correlated channel: Effect of battery capacity.

VI. Conclusions

In this paper, we considered the problem of power management policies for EHS nodes with packet retransmissions. Since the channel state is only partially observable through the ACK/NACK messages, we formulated the problem as a POMDP. Exact solutions were found to be computationally infeasible, motivating us to explore two heuristic solutions: the voting policy and the ML policy. The heuristic solutions were obtained by solving MDPs based on the belief state of the channel. Simulation results showed that the proposed POMDP solution significantly outperforms existing fixed-power retransmission schemes. Thus, the decision-theoretic approach adopted in this paper is a promising technique for the design of power management policies for EHS nodes that need to operate under stringent energy constraints and yet achieve high reliability or throughput.

References

[1] I. Stanojev, O. Simeone, Y. Bar-Ness, and D. H. Kim, "On the energy efficiency of hybrid-ARQ protocols in fading channels," in *Proc. ICC*, 2007, pp. 3173–3177.

[2] A. Kansal, J. Hsu, S. Zahedi, and M. B. Srivastava, "Power management in energy harvesting sensor networks," *ACM Trans. Embed. Comput. Syst.*, vol. 7, no. 4, pp. 1–38, 2007.

[3] J. Lei, R. Yates, and L. Greenstein, "A generic model for optimizing single-hop transmission policy of replenishable sensors," *IEEE Trans. Wireless Commun.*, vol. 8, no. 2, pp. 547–551, 2009.

[4] J. Yang and S. Ulukus, "Transmission completion time minimization in an energy harvesting system," in *Proc. Conf. on Inform. Sci. and Syst. (CISS)*, 2010, pp. 1–6.

[5] K. Tutuncuoglu and A. Yener, "Short-term throughput maximization for battery limited energy harvesting nodes," in *Proc. ICC*, 2011, pp. 1–5.

[6] A. Seyedi and B. Sikdar, "Energy efficient transmission strategies for body sensor networks with energy harvesting," *IEEE Trans. Commun.*, vol. 58, no. 7, pp. 2116–2126, 2010.

[7] 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, 2010.

[8] B. Medepally, N. B. Mehta, and C. R. Murthy, "Implications of energy profile and storage on energy harvesting sensor link performance," in *Proc. Globecom*, 2009, pp. 1–6.

[9] L. P. Kaelbling, M. L. Littman, and A. R. Cassandra, "Planning and acting in partially observable stochastic domains," *Artificial Intelligence*, vol. 101, no. 1-2, pp. 99–134, 1998.

[10] H. Li, N. Jaggi, and B. Sikdar, "Cooperative relay scheduling under partial state information in energy harvesting sensor networks," in *Proc. Globecom*, 2010, pp. 1–5.

[11] N. Jaggi, K. Kar, and A. Krishnamurthy, "Rechargeable sensor activation under temporally correlated events," *Wireless Networks*, vol. 15, no. 5, pp. 619–635, 2009.

[12] A. Sinha and P. Chaporkar, "Optimal power allocation for a renewable energy source," in *Proc. NCC*, 2012, pp. 1–5.

[13] "*IEEE standard 802, part 15.4: wireless medium access control (MAC) and physical layer (PHY) specifications for low rate wireless personal area networks (WPANs)*," Tech. Rep., 2003.

[14] J. A. Paradiso and T. Starner, "Energy scavenging for mobile and wireless electronics," *IEEE Pervasive Comput.*, vol. 4, no. 1, pp. 18–27, 2005.

[15] C. C. Tan and N. C. Beaulieu, "On first-order Markov modeling for the Rayleigh fading channel," *IEEE Trans. Commun.*, vol. 48, no. 12, pp. 2032–2040, 2000.

[16] A. Farrokh, V. Krishnamurthy, and R. Schober, "Optimal adaptive modulation and coding with switching costs," *IEEE Trans. Commun.*, vol. 57, no. 3, pp. 697 –706, 2009.

[17] A. K. Karmokar, D. V. Djonin, and V. K. Bhargava, "POMDP-based coding rate adaptation for type-I hybrid ARQ systems over fading channels with memory," *IEEE Trans. Wireless Commun.*, vol. 5, no. 12, pp. 3512–3523, 2006.

[18] H. S. Wang and N. Moayeri, "Finite-state Markov channel- A useful model for radio communication channels," *IEEE Trans. Veh. Technol.*, vol. 44, no. 1, pp. 163–171, 1995.

[19] Q. Zhang and S. A. Kassam, "Finite-state Markov model for Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 47, no. 11, pp. 1688–1692, 1999.

[20] R. D. Smallwood and E. J. Sondik, "The optimal control of partially observable Markov processes over a finite horizon," *Operations Research*, vol. 21, no. 5, pp. 1071–1088, 1973.

[21] C. H. Papadimitriou and J. N. Tsitsiklis, "The complexity of optimal queuing network control," *Mathematics of Operations Research*, vol. 24, no. 2, pp. pp. 293–305, 1999.

[22] M. L. Littman, A. R. Cassandra, and L. Kaelbling, "Learning policies for partially observable environments: Scaling up," in *Proc. Int. Conf. Machine Learning*, 1995, pp. 362–370.

[23] J. Pineau, G. Gordon, and S. Thrun, "Point-based value iteration: An anytime algorithm for POMDPs," in *Proc. Int. Joint Conf. on Artificial Intelligence*, vol. 18, 2003, pp. 1025–1032.

[24] D. P. Bertsekas, *Dynamic Programming and Optimal Control 3rd Edition, Volume I*, 2005.

[25] A. R. Cassandra, "Exact and approximate algorithms for partially observable Markov decision processes," Ph.D. dissertation, Brown University, 1998.

[26] R. Simmons and S. Koenig, "Probabilistic robot navigation in partially observable environments," in *Proc. Int. Joint Conf. on Artificial Intelligence*, 1995, pp. 1080–1087.