

# Joint Energy Allocation for Sensing and Transmission in Rechargeable Wireless Sensor Networks

Shaobo Mao, Man Hon Cheung, *Member, IEEE*, and Vincent W.S. Wong, *Senior Member, IEEE*

**Abstract**—Different from a traditional wireless sensor network (WSN) powered by non-rechargeable batteries, the energy management policy of a rechargeable WSN needs to take into account the process of energy harvesting. In this paper, we study the energy allocation for sensing and transmission in an energy harvesting sensor node with a rechargeable battery and a finite data buffer. The sensor aims to maximize the expected total amount of data transmitted until the sensor stops functioning subject to time-varying energy harvesting rate, energy availability in the battery, data availability in the data buffer, and channel fading. Since the lifetime of the sensor is a random variable, we formulate the energy allocation problem as an infinite-horizon Markov decision process (MDP), and propose an optimal energy allocation (OEA) algorithm using the value iteration. We then consider a special case with infinite data backlog and prove that the optimal transmission energy allocation (OTE) policy is monotone with respect to the amount of battery energy available. Finally, we conduct extensive simulations to compare the performance of our OEA algorithm, OTE algorithm, the finite-horizon transmission energy allocation (FHTEA) algorithm extended from [2], and the finite-horizon optimal energy allocation (FHOEA) algorithm from [1]. Simulation results show that the OEA algorithm transmits the largest amount of data, and the OTE algorithm can achieve a near-optimal performance with a low computational complexity.

**Index Terms**—Energy harvesting, wireless sensor networks, resource allocation, Markov decision process (MDP).

## I. INTRODUCTION

TRADITIONALLY, a wireless sensor network (WSN) is composed of a large number of sensor nodes powered by non-rechargeable batteries with limited energy storage capacities. As a result, a WSN can only function for a limited amount of time. The idea of *energy harvesting* was proposed to address the problem of finite lifetime in a WSN by enabling the sensor nodes to replenish energy from ambient sources, such as solar, wind, and vibrations [3], [4]. The design considerations of an

energy harvesting WSN are different from a non-rechargeable battery powered WSN in many ways. First, with a potentially infinite amount of energy available to the sensor nodes, an energy harvesting WSN can remain functional for a long period of time. Hence, energy conservation is not the prime design issue. Second, the energy management strategy for an energy harvesting WSN needs to take into account the energy replenishment process. For example, an overly conservative energy expenditure may limit the transmitted data by failing to take the full advantage of the energy harvesting process. On the other hand, an overly aggressive use of energy may result in an energy outage, which prevents some sensor nodes from functioning properly. Third, the energy availability constraint, which requires the energy consumption to be less than the energy stored in the battery, must be met at all time. This constraint complicates the design of an energy management policy, since the current energy consumption decision would affect the outcome in the future.

A lot of research efforts have been devoted to study the energy management and data transmission in energy harvesting WSNs. Kansal *et al.* in [5] proposed analytically tractable models to characterize the complex time varying nature of energy sources. Distributed algorithms were developed to utilize the harvested energy efficiently. Sharma *et al.* in [6] proposed energy management schemes for a single energy harvesting sensor node that achieves the maximum throughput and minimum mean delay. Gatzianas *et al.* in [7] presented an online adaptive transmission scheme for wireless networks with rechargeable batteries that maximizes total system utility and stabilizes the data queue using Lyapunov techniques. Huang *et al.* in [8] proposed an online algorithm that achieves a close-to-optimal utility performance in finite capacity energy storage devices. The Lyapunov optimization techniques with weight perturbation were used. In [9], utility-optimal energy allocation algorithms were proposed for systems with predictable or stochastic energy availability. Srivastava *et al.* in [10] analyzed the limits of the performance of energy harvesting sensor nodes with finite data and energy storage. An energy management scheme was proposed that achieves the optimal utility asymptotically. Mao *et al.* in [11] studied the joint data buffer and rechargeable battery control problem that aims to maximize the long-term average sensing rate of a wireless sensor network. Joint rate control, power allocation, and routing algorithms were proposed for both single hop and multihop networks. Chen *et al.* in [12] addressed the joint energy allocation and routing problem for network utility

Manuscript received on December 21, 2012; revised on April 18, 2013, July 28, 2013, and November 13, 2013; and accepted on November 16, 2013. This work was supported by the Natural Sciences and Engineering Research Council (NSERC) of Canada. Part of this paper was presented at the *IEEE International Conference on Communications (ICC'12)*, Ottawa, Canada, June 2012 [1]. The review of this paper was coordinated by Prof. Riku Jäntti.

S. Mao is with Teradici Corporation, Burnaby, BC, Canada, e-mail: shaobomao@gmail.com. M. H. Cheung is with the Department of Information Engineering, the Chinese University of Hong Kong, Hong Kong, China, e-mail: mhcheung@ie.cuhk.edu.hk. This work was completed when S. Mao and M. H. Cheung were with the Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, BC, Canada, V6T 1Z4. V. W. S. Wong is with the Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, BC, Canada, V6T 1Z4, e-mail: vincentw@ece.ubc.ca.

maximization. An online solution was proposed that achieves asymptotic optimality. Khouzani *et al.* in [13] proposed routing and scheduling policies that do not require explicit knowledge of the statistics of the energy replenishment or the traffic generation processes. They were able to learn and adapt to time variations in the physical and network environments dynamically, so as to achieve the long-term optimal data rates. In [14], energy management policies were identified that guarantees a minimum average distortion while ensuring the stability of the data buffer. Wang *et al.* in [15] considered the near-optimal power control policies with a saturated data queue in both the finite-horizon and infinite-horizon cases.

Some of the related works on energy harvesting WSNs have formulated the energy management problem as a Markov decision process (MDP) [16], [17]. Ho *et al.* in [2] proposed a throughput-optimal energy allocation algorithm for a time-slotted system under time-varying fading channel and energy source by using MDP. In [18], a throughput-optimal energy allocation policy was derived in a continuous time model and suboptimal online waterfilling schemes were proposed to address the dimensionality problem inherent in the MDP solution. Chen *et al.* in [19] studied the energy allocation problem of a single node using the shortest path approach. A simple distributed heuristic scheme was proposed that solves the joint energy allocation and routing problem in a rechargeable WSN. Li *et al.* in [20] proposed energy efficient scheduling strategies for cooperative communications in energy harvesting WSNs to maximize the long-term utility. Jaggi *et al.* in [21] considered the node activation problem for a rechargeable wireless sensor in the presence of temporal correlations in the sensed phenomena, and proposed optimal policies by using MDP. Kashef *et al.* in [22] studied a communication link that operates over a Gilbert-Elliot channel. The problem of maximizing the number of successfully delivered packets per time slot was formulated as a MDP problem. The proposed optimal policy was proved to exhibit the threshold structure that depends on the channel state and the energy queue length. Balsco *et al.* in [23] considered the transmission decisions of an energy harvesting sensor node with random data arrival, energy arrival, and channel conditions. A learning-theoretic approach, as well as online and offline problems were studied.

Most of these results from [2], [6]–[9], [13]–[15], [18], [19], [22] only considered the special case that there is either an infinitely long data backlog or data buffer. Yet, it is more practical to consider a *finite data buffer*. Besides, the energy consumed in *data sensing* has always been overlooked in the literature, such as in [23]. This motivates us to design an optimal energy allocation (OEA) algorithm for energy harvesting WSNs which takes into account both the data sensing energy consumption and the finite capacity of the data buffer. However, these considerations introduce new challenges. For instance, if the sensor node consumes an insufficient amount of energy for sensing but an excessive amount of energy for transmission, then the data buffer may be empty, which leads to a reduction in the total amount of data transmitted. Thus, the sensor node needs to maintain a good *balance* between the energy consumed for sensing and the energy for transmission. Moreover, different from [21], we assume that the sensor

nodes are always turned on, and we consider the case without temporally correlated events.

In this paper, we consider a point-to-point wireless link between an energy harvesting sensor node and a sink. The channel and energy harvesting rate may vary over time. The sensor node has a rechargeable battery and a data buffer with finite capacity. Our objective is to maximize the expected total amount of data transmitted until the sensor node stops functioning. The sensor node needs to decide the amount of energy it should allocate for sensing and transmission in each allocation interval by taking into account the battery energy level, data buffer level, energy harvesting rate, and channel condition. The main contributions of our work are as follows:

- We study the energy allocation problem for sensing and transmission in an energy harvesting sensor node. We propose optimal energy allocation algorithms that maximize the total amount of data transmitted over a random length of time, for a general case with a finite data buffer, and a special case with an infinite data backlog.
- For the general case, we formulate the problem as an infinite-horizon discounted MDP and propose the OEA algorithm by using the value iteration in MDP.
- For the special case, we prove that the optimal policy has a monotone structure and propose an optimal transmission energy allocation (OTEA) algorithm, which has a lower complexity than the value iteration algorithm. When applied to practical systems where the sensor needs to allocate energy for sensing and the data buffer is finite, the OTEA algorithm can also achieve a near-optimal performance.
- We provide extensive simulation results to compare the performance of the OEA algorithm, the OTEA algorithm, the finite-horizon transmission energy allocation (FHTEA) algorithm extended from [2], and the finite-horizon optimal energy allocation (FHOEA) algorithm from [1]. We study the impact of the average energy harvesting rate, the battery capacity, the data buffer size, the lifetime of the sensor node, and the data-sensing efficiency (i.e., the amount of data that the sensor can sense per unit energy) on the performance of total transmitted data. The results show that the OEA algorithm transmits the largest amount of data, and the OTEA algorithm achieves a near-optimal performance.

Compared with the existing works in the literature [2], [6]–[11], [13], [14], [18]–[20], [22], [24], [25], our system model is more realistic because we take into account a *finite* battery storage capacity, a *finite* data buffer, the *unpredictable* nature of the energy harvesting rate and channel condition, and the energy consumed for *sensing*, while the literature mentioned above does not consider all of these practical issues. In our earlier work in [1], we considered the energy allocation problem in a finite horizon, where the lifetime of the sensor is fixed. In this paper, however, we consider a more practical setting, where the lifetime of the sensor is a *random variable*, since the sensors are typically deployed in hostile outdoor environments, and thus are highly susceptible to the random physical destructions. Moreover, the optimal policies in an

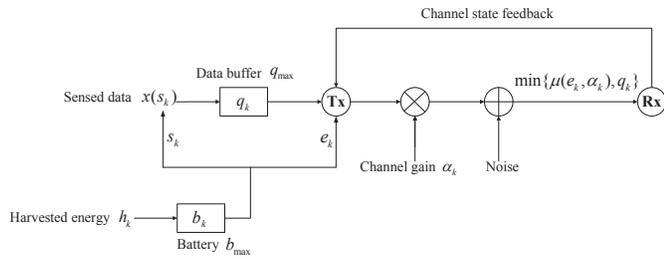


Fig. 1. The system model of an energy harvesting wireless sensor node transmitting data to the receiver  $\mathbf{R}\mathbf{x}$  of the sink. At allocation interval  $k$ , the random variables are the energy to be harvested  $h_k$  and the channel gain  $\alpha_k$ . Due to channel feedback delay and the time required to track the energy harvesting rate, we assume that the values of  $\alpha_{k-1}$  and  $h_{k-1}$  are only known at the beginning of allocation interval  $k$ . The optimization variables (or actions) are the energy consumed for transmission  $e_k$  and sensing  $s_k$ . The stored battery level is  $b_k$  and the amount of data available in the buffer is  $q_k$ .  $x(s_k)$  is the amount of data obtained by using  $s_k$  amount of energy. The amount of data transmitted in allocation interval  $k$  is  $\min\{\mu(e_k, \alpha_k), q_k\}$ . Our problem is to optimally allocate  $e_k$  and  $s_k$  such that the expected total amount of data transmitted till the sensor node stops functioning is maximized.

infinite-horizon MDP problem are typically stationary, and thus are simpler to implement than those in a finite-horizon MDP problem that vary in each allocation interval.

The rest of the paper is organized as follows: We describe the system model in Section II and formulate our problem in Section III. In Section IV, we first propose the OEA algorithm that maximizes the expected total amount of data transmitted using infinite-horizon MDP and then consider a special case that assumes an infinite data backlog and no energy allocated for sensing. In Section V, we discuss the possible extensions of our model. In Section VI, we evaluate the performance of the OEA algorithm and compare it with the OTEA algorithm, the FHOEA algorithm, and the FHTEA algorithm. Conclusions are given in Section VII.

## II. SYSTEM MODEL

As shown in Fig. 1, we consider an energy harvesting sensor node [2], [6], [18], [25] which contains a rechargeable battery with capacity  $b_{\max}$  Joule and a data buffer with size  $q_{\max}$  Mbits. We assume that the system is time-slotted, where the duration of a time slot, i.e., an allocation interval, is  $\tau$  s. Since sensor nodes are typically deployed in hostile outdoor environments (e.g., for forestry fire and volcano monitoring and detection, and battlefield surveillance [26]) in an unattended and distributed manner, they are highly susceptible to the physical destructions [27]–[29]. We let  $\nu$  be the probability that the sensor node can survive from the physical destruction or hardware failure and continue to function in an allocation interval, where  $\nu \in [0, 1)$ . For simplicity, we assume that  $\nu$  is fixed in all allocation intervals [23], [29]. Thus, the lifetime  $K$  of the sensor node is a geometrically distributed random variable with mean  $1/(1-\nu)$ . Let  $k \in \mathcal{K} \triangleq \{0, 1, \dots, K-1\}$  be the allocation interval index. The sensor node performs sensing in the field, stores the sensed data in the buffer, and transmits the data to the receiver  $\mathbf{R}\mathbf{x}$  of the sink over a wireless channel. Since the sensor node is exposed to hazardous environment and easy to get destroyed, it aims to

maximize the total amount of data transmitted until the sensor gets destroyed and stops functioning.

We consider an additive white Gaussian noise (AWGN) channel with block flat fading. That is, the channel remains constant for the duration of each allocation interval, but may change at the slot boundaries. Let  $\alpha_k$  be the channel gain in allocation interval  $k$ . We assume that the sink sends *delayed* channel state information (CSI) of the previous allocation interval back to the sensor node. At the beginning of allocation interval  $k$ , the sensor node only knows the value of  $\alpha_{k-1}$ , but not  $\alpha_k$ . The stored battery level is  $b_k$  and the amount of stored data in the data buffer is  $q_k$ . During the whole allocation interval  $k$ , the sensor node is able to replenish energy by  $h_k$ , which can be used for sensing or transmission in allocation interval  $k+1$  onward. As a result, the sensor node does not know the value of  $h_k$  until the beginning of the next allocation interval  $k+1$ . In other words, at the beginning of allocation interval  $k$ , the sensor node knows the value of  $h_{k-1}$ , but not  $h_k$ .

If the channel gain is  $\alpha_k$  and the allocated transmission energy is  $e_k$  in allocation interval  $k$ , then the average instantaneous transmission power is  $\frac{e_k}{\tau}$ . We consider that the sensor node is able to transmit  $\mu(e_k, \alpha_k)$  bits of data in allocation interval  $k$ , where  $\mu(e_k, \alpha_k)$  is a monotonically non-decreasing and concave function in  $e_k$  given  $\alpha_k$  in general. One such function is given by [30, pp. 172], [31], [32]:

$$\mu(e_k, \alpha_k) = \tau W \log_2 \left( 1 + \frac{\alpha_k e_k}{N_0 W \tau \Gamma} \right) \text{ bits}, \quad (1)$$

where  $N_0$  is the power spectral density of the Gaussian noise,  $W$  is the bandwidth of the channel,  $\Gamma$  is the signal-to-noise ratio (SNR) gap used to measure the reduction of SNR with respect to capacity, and it only depends on the error probability requirements.

For sensing in allocation interval  $k$ , we let  $x(s_k)$  be the amount of data generated when  $s_k$  units of energy are used for sensing. In general,  $x(s_k)$  is a monotonically non-decreasing and concave function in  $s_k$ . The data obtained by sensing in allocation interval  $k$  will be stored in the data buffer until they are transmitted in the subsequent allocation intervals. The exact expressions of  $h_k$ ,  $\alpha_k$ , and  $x(s_k)$  used for simulation will be discussed in more details in Section VI. Since sensing and transmission are the major tasks of a typical wireless sensor node, we focus on the energy consumption of these two aspects, and consider that other circuits in the sensor node consume negligible energy.

At allocation interval  $k$ , the sensor node needs to choose  $e_k$  and  $s_k$ , for all  $k \in \mathcal{K}$  such that the expected total amount of data transmitted is maximized. To achieve this goal, the sensor has to maintain a good *tradeoff* between the energy allocation for  $e_k$  and  $s_k$ . Given a fixed energy budget in an allocation interval, if  $e_k$  is too small, then the transmitted data in allocation interval  $k$  will be small. However, if  $e_k$  is too large, then  $s_k$  will be small such that insufficient amount of sensing data is stored in the buffer for transmission in the next allocation interval, which leads to a reduction in the amount of data transmitted in future allocation intervals. In addition, the total energy budget  $e_k + s_k$  in allocation interval  $k$  should

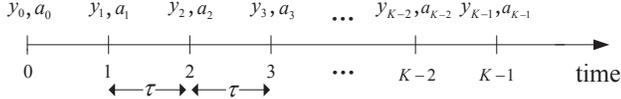


Fig. 2. Timing diagram of a Markov decision process (MDP).

also be carefully controlled. If the energy management policy is overly aggressive such that the rate of energy consumption is greater than the rate of energy harvesting, the sensor node may stop functioning because of the energy outage. On the other hand, an overly conservative energy management policy would limit the amount of data transmitted in each allocation interval. Thus, it is a challenging problem to decide the values of  $e_k$  and  $s_k$  optimally in each allocation interval  $k \in \mathcal{K}$ .

### III. PROBLEM FORMULATION

In this section, we formulate the problem of finding the optimal energy allocation for sensing and transmission as an MDP [16] [17], which consists of five elements: decision epochs, states, actions, state transition probabilities, and rewards. Referring to Fig. 2, the *decision epochs* are  $k \in \mathcal{K} = \{0, 1, \dots, K-1\}$ . The *state* of the system is denoted as  $\mathbf{y} = (b, q, h, \alpha)$ , which includes the battery energy state  $b$  and data buffer state  $q$  for the *current* allocation interval, as well as the energy harvesting state  $h$  and channel state  $\alpha$  in the *previous* allocation interval.  $b$ ,  $q$ ,  $h$ , and  $\alpha$  take discrete values and are all bounded. We denote the state space as  $Y = \mathcal{B} \times \mathcal{Q} \times \mathcal{H} \times \mathcal{A}$ , where  $\mathcal{B}$  is the set of battery energy states,  $\mathcal{Q}$  is the set of data buffer states,  $\mathcal{H}$  is the set of energy harvesting states, and  $\mathcal{A}$  is the set of channel states.  $Y$  is discrete and countable. Let  $\mathbf{y}_k$  denote the state of the system at allocation interval  $k$ , i.e.,  $\mathbf{y}_k = (b_k, q_k, h_{k-1}, \alpha_{k-1})$ . First, for the battery energy state in allocation interval  $k$ , the sensor node harvests  $h_k$  units of energy from the environment. On the other hand, it consumes  $e_k$  units of energy for data transmission and  $s_k$  units of energy for sensing. Because the battery energy state is discrete,  $e_k$  and  $s_k$  also take discrete values. Since the battery has a finite capacity  $b_{\max}$ , the energy stored in the battery is updated as

$$b_{k+1} = \min\{b_k - (e_k + s_k) + h_k, b_{\max}\}, \quad \forall k \in \mathcal{K}. \quad (2)$$

Eq. (2) ensures that the maximum stored energy  $b_{\max}$  is not exceeded. We assume that the initial energy  $b_0$  is known and satisfies the constraint  $0 \leq b_0 \leq b_{\max}$ . Moreover, the amount of energy consumed for sensing and transmission must be no more than the battery level:

$$e_k + s_k \leq b_k, \quad \forall k \in \mathcal{K}. \quad (3)$$

Second, for the data buffer state in allocation interval  $k$ ,  $x(s_k)$  amount of sensing data is generated and queued up in the data buffer if  $s_k$  units of energy are allocated for sensing. On the other hand, if the amount of data available in the data buffer for transmission at allocation interval  $k$  is  $q_k$ , and  $e_k$  units of energy are used for transmission, then the amount of data transmitted and removed from the data buffer at allocation interval  $k$  is given by  $\min\{\mu(e_k, \alpha_k), q_k\}$ . Since

the data buffer is finite with capacity  $q_{\max}$ , the amount of data in the buffer is then updated as

$$q_{k+1} = \min\{[q_k - \mu(e_k, \alpha_k)]^+ + x(s_k), q_{\max}\}, \quad \forall k \in \mathcal{K}, \quad (4)$$

where  $[z]^+ = \max\{z, 0\}$ . We assume that the initial amount of data in the data buffer  $q_0$  is known and satisfies  $0 \leq q_0 \leq q_{\max}$ . Eq. (4) implies that if the sensor allocates too much energy for transmission such that  $\mu(e_k, \alpha_k) > q_k$ , then energy will be wasted. Thus, the sensor should make a proper energy allocation decision at each allocation interval. Third, since the energy harvesting rate and the current channel state information at allocation interval  $k$  is not known to the sensor, we use two independent first-order stationary Markovian models to model  $h_k$ <sup>1</sup> and  $\alpha_k$  [2], [20], [34]. The transition probability of these two independent random variables are denoted as  $P(h_k | h_{k-1})$  and  $P(\alpha_k | \alpha_{k-1})$ .

Based on the current state  $\mathbf{y}_k$  at allocation interval  $k$ , an *action*  $\mathbf{a}_k = (e_k, s_k)$  is taken for transmission and sensing energy allocation from its feasible set  $U(\mathbf{y}_k)$ . We have

$$\mathbf{a}_k \in U(\mathbf{y}_k) = \{(e, s) | e + s \leq b_k, e \geq 0, s \geq 0\}, \quad (5)$$

where  $U(\mathbf{y}_k)$  represents the feasible set of the action  $\mathbf{a}_k$  given the current state  $\mathbf{y}_k$  at allocation interval  $k$ , and is discrete and finite. In addition, it is possible to impose additional constraints on  $\mathbf{a}_k$ . For example, a constraint on the minimum amount of energy for sensing or transmission to ensure a minimum amount of sensed data or transmitted data for each allocation interval, respectively. Also, a maximum transmission power constraint can be imposed.

The *state transition probability*  $P(\mathbf{y}_{k+1} | \mathbf{y}_k, \mathbf{a}_k)$  is the probability that the system will go into state  $\mathbf{y}_{k+1}$  if action  $\mathbf{a}_k$  is taken at state  $\mathbf{y}_k$  at allocation interval  $k$ . Due to the independence between  $(b_{k+1}, h_k)$  and  $(q_{k+1}, \alpha_k)$  for all  $k \in \mathcal{K}$ , we can simplify the state transition probability as

$$\begin{aligned} & P(\mathbf{y}_{k+1} | \mathbf{y}_k, \mathbf{a}_k) \\ &= P(b_{k+1}, q_{k+1}, h_k, \alpha_k | b_k, q_k, h_{k-1}, \alpha_{k-1}, e_k, s_k) \\ &= P(b_{k+1}, h_k | b_k, h_{k-1}, e_k, s_k) P(q_{k+1}, \alpha_k | q_k, \alpha_{k-1}, e_k, s_k) \\ &= P(b_{k+1} | b_k, h_k, e_k, s_k) P(h_k | h_{k-1}) \\ & \quad \times P(q_{k+1} | q_k, \alpha_k, e_k, s_k) P(\alpha_k | \alpha_{k-1}), \end{aligned} \quad (6)$$

where

$$P(b_{k+1} | b_k, h_k, e_k, s_k) = \begin{cases} 1, & \text{if Eq. (2) is satisfied,} \\ 0, & \text{otherwise,} \end{cases} \quad (7)$$

and

$$P(q_{k+1} | q_k, \alpha_k, e_k, s_k) = \begin{cases} 1, & \text{if Eq. (4) is satisfied,} \\ 0, & \text{otherwise.} \end{cases} \quad (8)$$

Eqs. (7) and (8) are due to the deterministic state element transitions described in Eqs. (2) and (4), respectively. Otherwise, there will not be any state transitions as stated in the second line of Eqs. (7) and (8).

<sup>1</sup>For example, it was shown in [33] that the stationary Markovian model is suitable for modeling the solar energy harvesting rate. Moreover, the work in [9] showed how to treat a stochastic model with an independent and identically distributed (i.i.d.) energy source as an MDP.

Given the current state  $\mathbf{y}_k$  and the action  $\mathbf{a}_k$ ,  $\mathbb{E}_{\alpha_k}[\mu(e_k, \alpha_k) | \alpha_{k-1}]$  is the expected amount of data that can be transmitted when  $e_k$  units of energy are used for transmission. However, since the data available in the data buffer for transmission at allocation interval  $k$  are  $q_k$ , the expected amount of data transmitted at allocation interval  $k$  is given by  $\mathbb{E}_{\alpha_k}[\min\{\mu(e_k, \alpha_k), q_k\} | \alpha_{k-1}]$ . We define the *reward* at allocation interval  $k$ ,  $r(\mathbf{y}_k, \mathbf{a}_k)$  to be the expected amount of data transmitted at allocation interval  $k$ . That is,

$$r(\mathbf{y}_k, \mathbf{a}_k) = \mathbb{E}_{\alpha_k}[\min\{\mu(e_k, \alpha_k), q_k\} | \alpha_{k-1}]. \quad (9)$$

A *decision rule* prescribes a procedure for the selection of an action in each state at a specified allocation interval. We denote the most general decision rule, i.e., the randomized history dependent decision rule [17, pp.21] at allocation interval  $k$  as  $\delta_k$ . A general *policy*  $\pi = (\delta_0, \delta_1, \dots, \delta_{K-1})$  is a sequence of decision rules to be used at all the allocation intervals. A *feasible* policy should satisfy Eq. (5) at all the allocation intervals. Let  $\Pi$  be the feasible set of  $\pi$ . Then, for any given state  $\mathbf{y}_0 = (b_0, q_0, h_{-1}, \alpha_{-1})$  at the first allocation interval, the *expected total reward* between the first allocation interval till the sensor stops functioning with policy  $\pi \in \Pi$  is given by [17, pp.125]:

$$J^\pi(\mathbf{y}_0) = \mathbb{E} \left\{ \mathbb{E}_K \left\{ \sum_{k=0}^{K-1} r(\mathbf{y}_k, \mathbf{a}_k) \right\} \middle| \mathbf{y}_0, \pi \right\}, \quad (10)$$

where  $\mathbb{E}\{\cdot\}$  denotes the statistical expectation taken over all relevant random variables given initial state  $\mathbf{y}_0$  and policy  $\pi$ .  $\mathbb{E}_K\{\cdot\}$  denotes the expectation with respect to the random variable  $K$ , which is the lifetime of the sensor node. It should be noted that with a different policy  $\pi$  and initial state  $\mathbf{y}_0$ , a different action will be chosen in each allocation interval in general, which results in a different state transition probability when the expectation  $\mathbb{E}\{\cdot\}$  is computed.

Based on the geometric distribution of the lifetime  $K$  of the sensor node with mean  $1/(1-\nu)$ , Eq. (10) is equivalent to the objective function of infinite-horizon MDP with discounted reward given by [17, Proposition 5.3.1]:

$$J^\pi(\mathbf{y}_0) = \mathbb{E} \left\{ \sum_{k=0}^{\infty} \nu^k r(\mathbf{y}_k, \mathbf{a}_k) \middle| \mathbf{y}_0, \pi \right\}. \quad (11)$$

Here, we can interpret  $\nu$  as the *discount factor* of the model. Since the sensor node will stop functioning at some time in the future, the reward at allocation interval  $k$  is discounted by a factor  $\nu^k$ .

*Lemma 1:*  $J^\pi(\mathbf{y}_0)$  defined in Eq. (11) is finite. That is  $|J^\pi(\mathbf{y}_0)| < \infty$ .

*Proof:* Since

$$\begin{aligned} & \sup_{\mathbf{a} \in U(\mathbf{y}), \mathbf{y} \in Y} |r(\mathbf{y}, \mathbf{a})| \\ &= \max_{\alpha \in \mathcal{A}} \left\{ \mathbb{E}_{\alpha'} [\min\{\mu(b_{\max}, \alpha'), q_{\max}\} | \alpha] \right\} < \infty, \end{aligned} \quad (12)$$

the objective function  $J^\pi(\mathbf{y}_0)$  of the infinite-horizon MDP converges to a finite value [17, pp.121]. ■

The sensor node aims to find an optimal sensing and transmit energy allocation policy that maximizes the expected

total discounted reward given in Eq. (11). That is, given the initial state  $\mathbf{y}_0$ , the sensor aims to obtain the *optimal expected total discounted reward*  $J(\mathbf{y}_0)$  and the *optimal policy*  $\pi^*$  defined as

$$J(\mathbf{y}_0) = \max_{\pi \in \Pi} J^\pi(\mathbf{y}_0) \quad \text{and} \quad \pi^* = \arg \max_{\pi \in \Pi} J^\pi(\mathbf{y}_0). \quad (13)$$

A policy is said to be *stationary deterministic* if  $\delta_k$  is deterministic Markovian [17, pp.21] and  $\delta_k = \delta$  for all  $k \in \mathcal{K}$  such that  $\pi = (\delta, \delta, \dots)$ . For the rest of the paper, a general policy is denoted by  $\pi$ , while a stationary deterministic policy is denoted by  $\delta^{SD}$ . For an infinite-horizon MDP, the only case of interest is when an optimal stationary deterministic policy exists. Thus, our objective is to find an *optimal stationary deterministic policy*  $\delta^{SD*}$ , which maximizes the expected total discounted reward in Eq. (11).

#### IV. ENERGY ALLOCATION ALGORITHMS

In this section, we obtain the *optimal stationary deterministic policies* for energy allocation. First, we consider a general case that takes into account a finite data buffer and the energy allocated for sensing. Next, we study a special case where we assume that there is an infinite data backlog.

##### A. General Case

In this subsection, we obtain the *optimal stationary deterministic policy* for the general case. An OEA algorithm that achieves the maximum expected total discounted reward in Eq. (13) is proposed based on the *value iteration* algorithm [17, pp.161].

The optimal expected total discounted reward  $J(\mathbf{y})$  given current state  $\mathbf{y}$  satisfies the *Bellman's equation of optimality* [17]:

$$J(\mathbf{y}) = \max_{\mathbf{a} \in U(\mathbf{y})} \left\{ r(\mathbf{y}, \mathbf{a}) + \nu \sum_{\mathbf{y}' \in Y} P(\mathbf{y}' | \mathbf{y}, \mathbf{a}) J(\mathbf{y}') \right\}. \quad (14)$$

In Eq. (14), the first and second terms on the right hand side represent, respectively, the *immediate reward* at the current allocation interval and the *expected total discounted future reward* if action  $\mathbf{a}$  is chosen. Hence, Eq. (14) describes the *tradeoff* between the current reward and the expected future reward. As mentioned in Section III, for an infinite-horizon MDP, the only case of interest is when an optimal stationary deterministic policy exists.

*Theorem 1:* There exists an *optimal stationary deterministic policy*  $\delta^{SD*}$  that maximizes the right hand side of Eq. (14), given by

$$\delta^{SD*}(\mathbf{y}) = \arg \max_{\mathbf{a} \in U(\mathbf{y})} \left\{ r(\mathbf{y}, \mathbf{a}) + \nu \sum_{\mathbf{y}' \in Y} P(\mathbf{y}' | \mathbf{y}, \mathbf{a}) J(\mathbf{y}') \right\}. \quad (15)$$

*Proof:* Notice that the system state space  $Y$  is countable and discrete, and  $U(\mathbf{y})$  is finite for each  $\mathbf{y} \in Y$ . From [17, Theorem 6.2.10], an optimal stationary deterministic policy exists. ■

---

**Algorithm 1** Optimal Energy Allocation (OEA) Algorithm for Energy Harvesting Sensor Node.

---

- 1: Planning Phase:
- 2: Arbitrarily select  $J_0(\mathbf{y})$  for each  $\mathbf{y} \in Y$ , specify  $\varepsilon > 0$ , and set  $n := 0$ .
- 3: For each  $\mathbf{y} \in Y$ , compute  $J_{n+1}(\mathbf{y})$  by

$$J_{n+1}(\mathbf{y}) := \max_{\mathbf{a} \in U(\mathbf{y})} \left\{ r(\mathbf{y}, \mathbf{a}) + \nu \sum_{\mathbf{y}' \in Y} P(\mathbf{y}' | \mathbf{y}, \mathbf{a}) J_n(\mathbf{y}') \right\}. \quad (16)$$

- 4: If  $\|\mathbf{J}_{n+1} - \mathbf{J}_n\| < \frac{\varepsilon(1-\nu)}{2\nu}$ , go to line 5. Otherwise increment  $n$  by 1 and go to line 3.
- 5: For each  $\mathbf{y} \in Y$ , choose stationary  $\varepsilon$ -optimal policy

$$\delta^{SD^*}(\mathbf{y}) := \arg \max_{\mathbf{a} \in U(\mathbf{y})} \left\{ r(\mathbf{y}, \mathbf{a}) + \nu \sum_{\mathbf{y}' \in Y} P(\mathbf{y}' | \mathbf{y}, \mathbf{a}) J_{n+1}(\mathbf{y}') \right\}, \quad (17)$$

and stop.

- 6: Sensing and Transmission Phase:
  - 7: Set  $k := 0$ .
  - 8: **while**  $k \leq K - 1$  **do**
  - 9: Track the energy harvesting rate of the previous allocation interval  $h_{k-1}$ .
  - 10: Track the energy available for use in the battery  $b_k$ .
  - 11: Track the amount of data in the buffer  $q_k$ .
  - 12: Obtain the channel feedback  $\alpha_{k-1}$  from the sink.
  - 13: Set  $\mathbf{y} := (b_k, q_k, h_{k-1}, \alpha_{k-1})$ .
  - 14: Obtain action  $\delta^{SD^*}(\mathbf{y}) := (e^*(\mathbf{y}), s^*(\mathbf{y}))$  based on the optimal policy.
  - 15: Consume  $e^*(\mathbf{y})$  amount of energy for transmission and  $s^*(\mathbf{y})$  amount of energy for sensing.
  - 16: Update battery energy  $b_{k+1}$  using Eq. (2) and the amount of data in the buffer  $q_{k+1}$  using Eq. (4).
  - 17: Set  $k := k + 1$ .
  - 18: **end while**
- 

We then propose the OEA algorithm in Algorithm 1. In the planning phase, the sensor solves for the optimal stationary deterministic policy  $\delta^{SD^*}$  based on value iteration algorithm, and records it as a look-up table. Specifically, in line 2, we initialize  $J_0(\mathbf{y})$  for all  $\mathbf{y} \in Y$  arbitrarily, specify the error bound  $\varepsilon$ , and set the iteration sequence  $n$  to be 0. In line 3, we compute  $J_{n+1}(\mathbf{y})$  for each  $\mathbf{y} \in Y$  based on the knowledge of  $J_n(\mathbf{y})$ . In line 4, we first check whether  $\|\mathbf{J}_{n+1} - \mathbf{J}_n\| < \frac{\varepsilon(1-\nu)}{2\nu}$  holds where  $\mathbf{J}_{n+1} = (J_{n+1}(\mathbf{y}), \forall \mathbf{y} \in Y)$  and  $\mathbf{J}_n = (J_n(\mathbf{y}), \forall \mathbf{y} \in Y)$ , and the norm function is defined to be  $\|\mathbf{J}\| = \max_{\mathbf{y} \in Y} |J(\mathbf{y})|$  for  $\mathbf{y} \in Y$ . If the inequality holds, which means that the value iteration algorithm has converged, then we proceed to obtain the optimal stationary deterministic policy  $\delta^{SD^*}$  in line 5 and stop. Otherwise, we go back to line 3 and continue to iterate. In the sensing and transmission phase, the sensor node chooses the action  $\delta^{SD^*}(\mathbf{y}) = (e^*(\mathbf{y}), s^*(\mathbf{y}))$  based on current system state  $\mathbf{y}$  and the optimal stationary deterministic policy  $\delta^{SD^*}$  in line 14. That is, it consumes  $e^*(\mathbf{y})$  and  $s^*(\mathbf{y})$  amount of energy for transmission and sensing, respectively.

For the *convergence*,  $J_n(\mathbf{y})$  generated in line 3 converges in norm to  $J(\mathbf{y})$  for all  $\mathbf{y} \in Y$ . The stationary policy  $\delta^{SD^*}$  defined in line 5 is  $\varepsilon$ -optimal, and whenever the convergence criterion  $\|\mathbf{J}_{n+1} - \mathbf{J}_n\| < \frac{\varepsilon(1-\nu)}{2\nu}$  is satisfied,  $\|\mathbf{J}_{n+1} - \mathbf{J}\| <$

$\varepsilon/2$  holds [17, Theorem 6.3.1(d)], where  $\mathbf{J} = (J(\mathbf{y}), \forall \mathbf{y} \in Y)$  is the vector of *optimal* expected total discounted reward defined in Eq. (14). Besides, the convergence is linear at rate  $\nu$  [17, Theorem 6.3.3]. In practice, choosing  $\varepsilon$  small enough ensures to obtain a policy that is very close to optimal.

The following property described in Lemma 2 is intuitive and is used for establishing the structural result of the optimal policy in Theorem 2. If more energy is available in the battery (i.e., a larger  $b$ ), we can allocate more energy for sensing and transmission so that the total reward  $J$  increases. Similarly, if more data are available in the data buffer for transmission (i.e., a larger  $q$ ), we can then allocate less energy for sensing and more energy for transmission, which would result in a larger total reward  $J$ .

*Lemma 2:* (a)  $J(b, q, h, \alpha)$  is increasing in battery state  $b$  for any given data buffer state  $q$ , energy harvesting state  $h$ , and channel state  $\alpha$ . (b)  $J(b, q, h, \alpha)$  is increasing in  $q$  for any given  $b, h$  and  $\alpha$ .

*Proof:* We prove Lemma 2 by mathematical induction. In order to show that  $J(b, q, h, \alpha)$  is increasing in  $b$  and  $q$ , we aim to prove that  $J_n(b, q, h, \alpha)$  generated by Eq. (16) in Algorithm 1 is increasing in  $b$  and  $q$  for all  $n$ . Since for any initialization  $J_0(b, q, h, \alpha)$ ,  $J_n(b, q, h, \alpha)$  converges to the same optimal expected total discounted reward  $J(b, q, h, \alpha)$  [17], we can select  $J_0(b, q, h, \alpha)$  which is increasing in  $b$  and  $q$ . Assume  $J_n(b, q, h, \alpha)$  is increasing in  $b$  and  $q$ . We expand Eq. (16) as

$$J_{n+1}(b, q, h, \alpha) = \max_{\mathbf{a} \in U(\mathbf{y})} \left\{ \mathbb{E}_{\alpha'}[\min\{\mu(e, \alpha'), q\} | \alpha] + \nu \mathbb{E}_{h', \alpha'} \left[ J_n \left( \min\{b - (e + s) + h', b_{\max}\}, \min\{[q - \mu(e, \alpha')]^+ + x(s), q_{\max}\}, h', \alpha' \right) | h, \alpha \right] \right\}. \quad (18)$$

Note that the first term on the right hand side of Eq. (18) is independent of  $b$  and increasing in  $q$ , and the second term is increasing in  $b$  and  $q$  based on the assumption that  $J_n(b, q, h, \alpha)$  is increasing in  $b$  and  $q$ . Therefore,  $J_{n+1}(b, q, h, \alpha)$  is increasing in  $b$  and  $q$ . By induction,  $J_n(b, q, h, \alpha)$  is increasing in  $b$  and  $q$  for all  $n$ . Thus  $J(b, q, h, \alpha) = J_\infty(b, q, h, \alpha)$  is increasing in  $b$  for any given  $q, h$  and  $\alpha$ . Moreover, it is increasing in  $q$  for any given  $b, h$  and  $\alpha$ . ■

### B. Special Case: Infinite Data Backlog

In this subsection, we consider a special case where the sensor has an infinite data backlog. As a result, we do not need to consider the sensing energy  $s$  and the data buffer state  $q$ . So the system state is left with three elements: the battery energy  $b$  for current allocation interval, the energy harvesting rate  $h$ , and the channel state  $\alpha$  for the previous allocation interval. Based on the current system state, the sensor will choose  $e$  units of energy for transmission. We denote the expected optimal total discounted reward as  $\hat{J}(b, h, \alpha)$ , which satisfies the following Bellman's equation of optimality:

$$\hat{J}(b, h, \alpha) = \max_{0 \leq e \leq b} \left\{ \mathbb{E}_{\alpha'}[\mu(e, \alpha') | \alpha] + \nu \bar{J}(b - e, h, \alpha) \right\}, \quad (19)$$

where

$$\bar{J}(\hat{b}, h, \alpha) = \mathbb{E}_{h', \alpha'}[\hat{J}(\min\{\hat{b} + h', b_{\max}\}, h', \alpha') | h, \alpha]. \quad (20)$$

The first term on the right hand side of Eq. (19) represents the immediate reward for allocating  $e$  units of energy for transmission, and the second term represents the total future discounted reward. Eq. (19) can be solved via the value iteration algorithm as in Section IV-A. However, we can prove some properties related to  $\hat{J}(b, h, \alpha)$  and  $\bar{J}(\hat{b}, h, \alpha)$  in Lemmas 3 and 4, which leads to the *monotone policy* [17] in Theorem 2.

*Lemma 3:* (a)  $\hat{J}(b, h, \alpha)$  is concave in  $b$  for any given  $h$  and  $\alpha$ . (b)  $\bar{J}(\hat{b}, h, \alpha)$  is concave in  $\hat{b}$  for any given  $h$  and  $\alpha$ .

The proof of Lemma 3 is given in Appendix A. Since  $\mu(e, \alpha')$  is concave in  $e$ ,  $\mathbb{E}_{\alpha'}[\mu(e, \alpha') | \alpha]$  is also concave in  $e$ . By applying Lemma 3(b),  $\nu \bar{J}(b - e, h, \alpha)$  is concave in  $(b - e)$ . Thus, the concavities of the two terms in Eq. (19) translate into a diminishing marginal reward for consuming energy at the current allocation interval, and saving energy for the future allocation intervals, respectively. Balancing these two terms properly results in an optimal policy.

*Lemma 4:*  $\hat{J}(b, h, \alpha)$  is increasing in  $b$  for any given  $h$  and  $\alpha$ .

The proof is quite similar with the proof of Lemma 2, and is omitted due to page limitation.

*Theorem 2:* For the optimal stationary deterministic policy

$$\hat{e}^*(b, h, \alpha) = \min \left\{ e' \in \arg \max_{0 \leq e \leq b} \{ \mathbb{E}_{\alpha'}[\mu(e, \alpha') | \alpha] + \nu \bar{J}(b - e, h, \alpha) \} \right\}, \quad (21)$$

it is monotone increasing in  $b$  for any given  $h$  and  $\alpha$ . That is, for any  $b' \geq b$ , we have

$$\hat{e}^*(b', h, \alpha) \geq \hat{e}^*(b, h, \alpha), \quad \forall h \in \mathcal{H}, \forall \alpha \in \mathcal{A}. \quad (22)$$

The proof of Theorem 2 is given in Appendix B. With this monotone structure, we can significantly reduce the computational complexity of the value iteration algorithm, and propose our OTEA algorithm in Algorithm 2. The planning phase of the OTEA algorithm and OEA algorithm (i.e., Algorithm 1) are similar. The main difference is the procedure in computing  $\hat{J}_{n+1}(b, h, \alpha)$  in Eq. (23) in Algorithm 2, which has a lower complexity than that of the OEA algorithm. Specifically, in line 6 of Algorithm 2, for any given  $h \in \mathcal{H}$  and  $\alpha \in \mathcal{A}$ , we have  $\hat{e}_{n+1}(b + \Delta_b, h, \alpha) \geq \hat{e}_{n+1}(b, h, \alpha)$  from the proof of Theorem 2, where  $\Delta_b$  is the quantization resolution of battery energy. When we compute  $\hat{J}_{n+1}(b + \Delta_b, h, \alpha)$  and search for  $\hat{e}_{n+1}(b + \Delta_b, h, \alpha)$ , we can find the optimal solution in the interval of  $[\hat{e}_{n+1}(b, h, \alpha), b + \Delta_b]$  instead of the longer interval  $[0, b + \Delta_b]$ . In the sensing and transmission phase, when we apply our OTEA algorithm to a practical system, we still need to take into account the energy for sensing. In this way, we fix the percentage of energy allocated for sensing to be  $p$  in line 19. Notice that the proper value of  $p$  can be obtained by simulation-based approaches [35]. For example, we can run the algorithm with different  $p$ , and find the optimal value of  $p$  that achieves performance which is close to the optimal performance of the OEA algorithm. The other operations in

---

## Algorithm 2 Optimal Transmission Energy Allocation (OTEA) Algorithm for Energy Harvesting Sensor Node.

---

- 1: **Planning Phase:**
  - 2: Arbitrarily select  $\hat{J}_0(b, h, \alpha)$  for each  $b \in \mathcal{B}$ ,  $h \in \mathcal{H}$ ,  $\alpha \in \mathcal{A}$ , specify  $\varepsilon > 0$ ,  $\Delta_b > 0$ , and set  $n := 0$ .
  - 3: **for each**  $h \in \mathcal{H}$ ,  $\alpha \in \mathcal{A}$  **do**
  - 4:   Set  $b := 0$  and  $l := 0$ .
  - 5:   **while**  $b \leq b_{\max}$  **do**
  - 6:     Compute
 
$$\hat{J}_{n+1}(b, h, \alpha) := \max_{l \leq e \leq b} \left\{ \mathbb{E}_{\alpha'}[\mu(e, \alpha') | \alpha] + \nu \mathbb{E}_{h', \alpha'}[\hat{J}_n(\min\{b - e + h', b_{\max}\}, h', \alpha') | h, \alpha] \right\}, \quad (23)$$

$$\hat{e}_{n+1}(b, h, \alpha) := \min \left\{ e' \in \arg \max_{l \leq e \leq b} \left\{ \mathbb{E}_{\alpha'}[\mu(e, \alpha') | \alpha] + \nu \mathbb{E}_{h', \alpha'}[\hat{J}_n(\min\{b - e + h', b_{\max}\}, h', \alpha') | h, \alpha] \right\} \right\}.$$
  - 7:     Set  $b := b + \Delta_b$ ,  $l := \hat{e}_{n+1}(b, h, \alpha)$ .
  - 8:     **end while**
  - 9: **end for**
  - 10: If  $\|\hat{J}_{n+1} - \hat{J}_n\| < \frac{\varepsilon(1-\nu)}{2\nu}$ , go to line 11. Otherwise increment  $n$  by 1 and go to line 3.
  - 11: For each  $b \in \mathcal{B}$ ,  $h \in \mathcal{H}$ ,  $\alpha \in \mathcal{A}$ , choose stationary  $\varepsilon$ -optimal policy
 
$$\hat{\delta}^{SD*}(b, h, \alpha) := \arg \max_{0 \leq e \leq b} \left\{ \mathbb{E}_{\alpha'}[\mu(e, \alpha') | \alpha] + \nu \mathbb{E}_{h', \alpha'}[\hat{J}_{n+1}(\min\{b - e + h', b_{\max}\}, h', \alpha') | h, \alpha] \right\}, \quad (24)$$
 and stop.
  - 12: **Sensing and Transmission Phase:**
  - 13: Set  $k := 0$ .
  - 14: **while**  $k \leq K - 1$  **do**
  - 15:   Track the energy harvesting rate of the previous allocation interval  $h_{k-1}$ .
  - 16:   Track the energy available for use in the battery  $b_k$ .
  - 17:   Track the amount of data in the buffer  $q_k$ .
  - 18:   Obtain the channel feedback  $\alpha_{k-1}$  from the sink.
  - 19:   Choose the amount of energy for sensing to be  $\hat{s}^* := pb_k$ , where  $p$  is the fixed percentage of energy for sensing.
  - 20:   Choose the amount of energy for transmission to be  $\hat{e}^* := \hat{\delta}^{SD*}((1-p)b_k, h_{k-1}, \alpha_{k-1})$  based on the policy  $\hat{\delta}^{SD*}$  obtained in Step 11.
  - 21:   Consume  $\hat{e}^*$  amount of energy for transmission and  $\hat{s}^*$  amount of energy for sensing.
  - 22:   Update battery energy  $b_{k+1}$  using Eq. (2) and the amount of data in the buffer  $q_{k+1}$  using Eq. (4).
  - 23:   Set  $k := k + 1$ .
  - 24: **end while**
- 

the sensing and transmission phase are the same as that in the OEA algorithm.

The running time of each iteration in the value iteration algorithm is  $O(N_{state}^2 N_{action})$  [36], where  $N_{state}$  is the total number of system states, and  $N_{action}$  is the total number of actions in the action space. In the general case discussed in Section IV-A, there are  $|\mathcal{B}| \times |\mathcal{Q}| \times |\mathcal{H}| \times |\mathcal{A}|$  system states, and the maximum number of actions is equal to  $\frac{(|\mathcal{B}|+1)|\mathcal{B}|}{2}$ . Since many of the state transition probabilities in Eq. (6) are equal to 0, the running time of each iteration in Algorithm 1 can be reduced to  $O(|\mathcal{B}||\mathcal{Q}||\mathcal{H}|^2|\mathcal{A}|^2 \times \frac{(|\mathcal{B}|+1)|\mathcal{B}|}{2})$ . On the

other hand, the running time of each iteration in Algorithm 2 is  $O(|\mathcal{B}||\mathcal{H}|^2|\mathcal{A}|^2 \times |\mathcal{B}|)$ . Thus, we can see that Algorithm 2 has a lower computational complexity than Algorithm 1, and meanwhile achieves a near-optimal performance when a proper value of  $p$  is chosen, which will be shown later in the simulation results in Section VI.

Although some energy is consumed and some delay is incurred in order to compute the optimal policy, it should be noted that since the optimal policy can be computed *offline*, we can handle the computation-intensive operations before the deployment of the sensor nodes. Specifically, we first obtain the statistics related to the variations of the energy harvesting state  $h$  and the channel state  $\alpha$ . Next, with this information, we compute the optimal policy  $\pi$  in the planning phase in Algorithms 1 or 2 *offline* with computers. Then, we load the optimal policy into the sensor nodes, and deploy them for the sensing application. In the field, each node determines its sensing and transmission decision in the sensing and transmission phase in Algorithms 1 or 2 by checking the optimal policy  $\pi$  as a table lookup, which has a small energy consumption and delay.

## V. MODEL EXTENSIONS

In this section, we discuss different possible extensions of our model.

1) Network Model: In this paper, we consider the communication between a sender and a receiver. For the extension to the multi-hop network model, issues such as scheduling and routing should be considered [11], [13].

2) Time Scales: We assumed in this work that the variations in wireless channel condition and energy harvesting rate are in the same time scale. To improve the accuracy of the modeling, we may consider a hierarchically structured problem with two different time scales, and apply the *multitime scale Markov decision processes* (MMDPs) [37]. Specifically, we may consider a *fast time-scale* involving the channel state  $\alpha$  and the transmission decision  $e$ . Then, we may consider a *slow time-scale* involving the energy harvesting state  $h$  and the sensing decision  $s$ . The two time scales are coupled by the battery energy state  $b$  and the data buffer state  $q$ .

3) Transmission Errors: Similar to the previous works (such as [2], [14]) in energy harvesting sensor networks, we do not take into account the details of the erroneous transmission in the physical layer, which is an interesting direction for further research. For the issues related to packet errors and retransmissions, we notice that they depend on the error control of the link layer protocol [38], which is chosen according to the requirements of the wireless sensor application.

## VI. PERFORMANCE EVALUATION

In this section, we evaluate the performance of our OEA, OTEA, FHTEA, and FHOEA algorithms in terms of the total amount of data transmitted. The FHTEA algorithm is extended based on the algorithm proposed in [2]. The FHOEA algorithm is from [1]. These two algorithms are used for comparison with our proposed algorithms. We consider a band-limited AWGN channel, where the channel bandwidth is  $W = 100$  kHz and

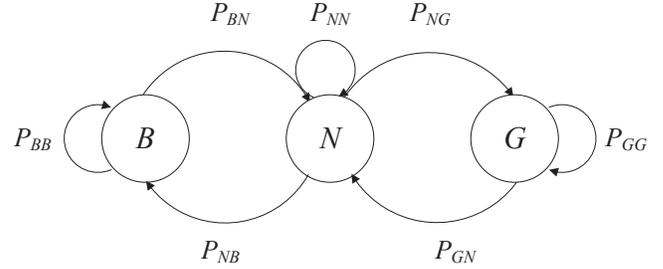


Fig. 3. A three-state Markov chain for the channel gain, where “B”, “N”, and “G” represent the channel in the bad, normal, and good states, respectively.

the noise power spectral density is  $N_0 = 10^{-18}$  W/Hz. The channel state can be “G = Good”, “N = Normal”, or “B = Bad”. It evolves according to the three-state Markov chain [39] as shown in Fig. 3 with the transition matrix of the Markov chain given by

$$P_\alpha = \begin{bmatrix} P_{BB} & P_{BN} & P_{BG} \\ P_{NB} & P_{NN} & P_{NG} \\ P_{GB} & P_{GN} & P_{GG} \end{bmatrix} = \begin{bmatrix} 0.3 & 0.7 & 0 \\ 0.25 & 0.5 & 0.25 \\ 0 & 0.7 & 0.3 \end{bmatrix}, \quad (25)$$

where  $P_{XZ}$  represents the probability of the channel state evolving from state  $X$  to state  $Z$ ,  $X$  and  $Z \in \{B, N, G\}$ . The channel gain  $\alpha$  at the “Bad”, “Normal”, and “Good” states are equal to  $2 \times 10^{-13}$ ,  $4 \times 10^{-13}$ , and  $6 \times 10^{-13}$ , respectively. We set the symbol error rate (SER) requirement to be  $10^{-3}$  [32]. Since the SNR gap  $\Gamma$  only depends on SER, i.e.,  $\Gamma = \frac{1}{3}(Q^{-1}(\frac{SER}{4}))^2$  [32],  $\Gamma$  is equal to 4. Unless specified otherwise, we assume that the battery buffer size  $b_{\max} = 30$  J [10], [11], and the data buffer size  $q_{\max} = 0.5$  Mbits. The initial amount of energy in the battery is 10 J and the initial amount of data in the buffer is 0.1 Mbits. For tractability, we assume that the energy harvesting rate  $h_k$  takes values from the finite set  $\mathcal{H} = \{H_1, H_2, H_3\} = \{4, 8, 12\}$  J/allocation interval [4], and evolves according to the three-state *Markov chain* with the state transition probability given by

$$P_h = \begin{bmatrix} P_{H_1H_1} & P_{H_1H_2} & P_{H_1H_3} \\ P_{H_2H_1} & P_{H_2H_2} & P_{H_2H_3} \\ P_{H_3H_1} & P_{H_3H_2} & P_{H_3H_3} \end{bmatrix} = \begin{bmatrix} 0.5 & 0.5 & 0 \\ 0.25 & 0.5 & 0.25 \\ 0 & 0.5 & 0.5 \end{bmatrix}, \quad (26)$$

where  $P_{H_iH_j}$  represents the probability of the energy harvesting state going from state  $H_i$  to state  $H_j$ ,  $\forall i, j \in \{1, 2, 3\}$ . The steady state probability is then given by  $[P_{H_1} P_{H_2} P_{H_3}] = [0.25 \ 0.5 \ 0.25]$ .  $x(s_k)$  is assumed to be a linear function of  $s_k$  given by  $x(s_k) = \gamma s_k$  [19], where  $\gamma$  is the data-sensing efficiency parameter (i.e., the amount of data that the sensor can sense per unit energy). We adopt  $\gamma = 0.08$  Mbits/J. For the value iteration algorithm, we choose  $\varepsilon$  to be  $10^{-3}$  and the discount factor  $\nu$  to be 0.95.

The FHTEA algorithm in [2] assumed infinite backlogged data and neglected the sensing energy. For fair comparison, we modify the FHTEA algorithm by allowing the data buffer to be finite with size  $q_{\max}$ . We assume that the sensor allocates a fixed percentage of battery energy for sensing in each allocation interval. The optimization problem is to maximize the total amount of transmitted data with the energy allocated for transmission in each allocation interval as the optimization

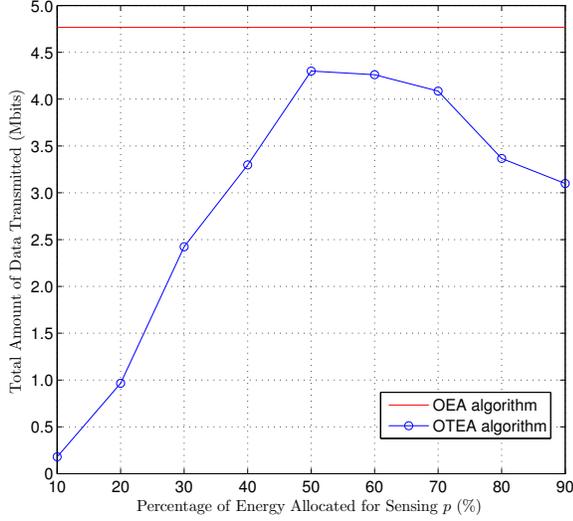


Fig. 4. The total amount of data transmitted of OTEA algorithm under different percentage of energy allocated for sensing  $p$ . Since the OEA algorithm does not allocate a fixed amount of energy for sensing, its total amount of data transmitted is independent of  $p$ .

variable. Besides, the FHTEA algorithm in [2] considered the energy allocation over a finite horizon, where the lifetime of the sensor node is known. So we fix the sensor lifetime in the FHTEA algorithm to be equal to the mean of the sensor lifetime in the OEA and OTEA algorithms. In the FHOEA algorithm, the sensor takes into account the energy allocated for sensing, but neglects the randomness of the sensor lifetime by fixing the sensor lifetime to be a constant. By comparing with the FHTEA algorithm and FHOEA algorithm, which have a similar system model as the OEA and OTEA algorithms, we can see clearly the performance gain by taking into the account the limited size of the data buffer, the sensing energy, and the randomness of the sensor lifetime.

Since the performance of the OTEA algorithm is related to the fixed amount of energy allocated for sensing  $p$ , we examine the total amount of data (computed in the expected sense) transmitted by the OTEA algorithm under different percentage of energy allocated for sensing, and compare with the OEA algorithm. As shown in Fig. 4, with around 50% of the available battery energy allocated for sensing, the OTEA algorithm transmits the largest amount data, and it is close to that of the OEA algorithm. This implies that we can apply the OTEA algorithm, which has a lower complexity than the OEA algorithm, and choose the optimal fixed percentage of energy for sensing to achieve a near-optimal performance.

Moreover, the optimal percentage of energy allocated for sensing for the OTEA algorithm depends on the data-sensing efficiency  $\gamma$ , the average energy harvesting rate  $\bar{H}$ , and the average channel gain  $\bar{\alpha}$ . With a larger  $\gamma$ , the sensor can sense more data using the same amount of energy. Fig. 5 shows that as  $\gamma$  increases, the optimal percentage of energy for sensing decreases. Fig. 6 shows the optimal percentage of energy allocated for sensing under different average energy harvesting rate  $\bar{H}$ . As  $\bar{H}$  increases, the optimal percentage of energy allocated for sensing actually decreases. Fig. 7 shows that the optimal percentage of energy allocated for sensing increases

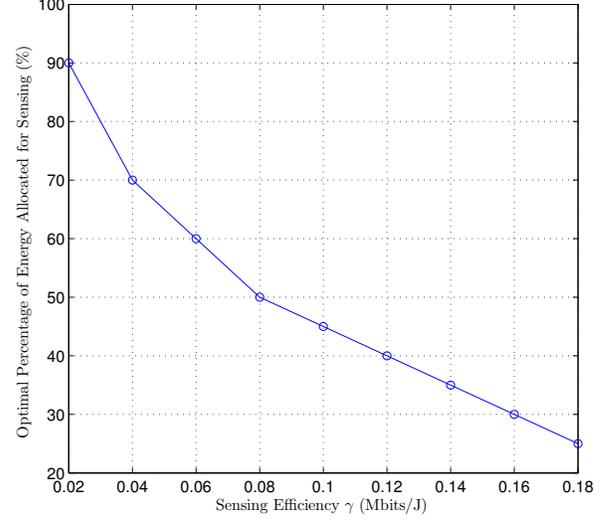


Fig. 5. The optimal percentage of energy allocated for sensing under different data-sensing efficiency  $\gamma$  for the OTEA algorithm.

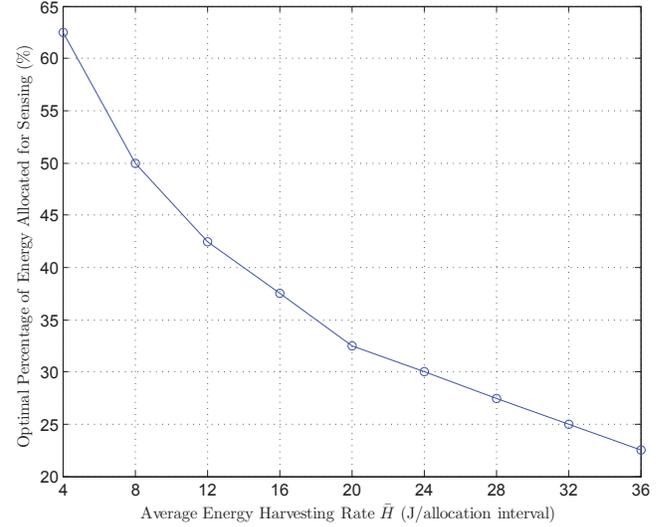


Fig. 6. The optimal percentage of energy allocated for sensing under different average energy harvesting rate  $\bar{H}$  for the OTEA algorithm.

as average channel gain  $\bar{\alpha}$  increases. As  $\bar{\alpha}$  increases, the data transmission is more efficient, and the sensor can allocate less energy for transmission and more energy for sensing. When  $\bar{\alpha}$  is very large, which means that the data transmission is extremely efficient, the sensor would allocate almost all of the incoming energy for sensing.

We then examine the total amount of data transmitted by the OEA algorithm, the OTEA algorithm, the FHOEA algorithm, and the FHTEA algorithm under different average energy harvesting rates  $\bar{H}$ , where  $\bar{H} = \sum_{i=1}^3 H_i P_{H_i}$ . For the OTEA and FHTEA algorithms, the percentage of energy allocated for sensing  $p$  is fixed to be 50%. In Fig. 8, we plot the total amount of data transmitted against different average energy harvesting rate for these three algorithms. We observe that the OEA algorithm performs better than the OTEA algorithm, the FHOEA algorithm, and the FHTEA algorithm, since the OEA algorithm achieves the optimal performance by solving

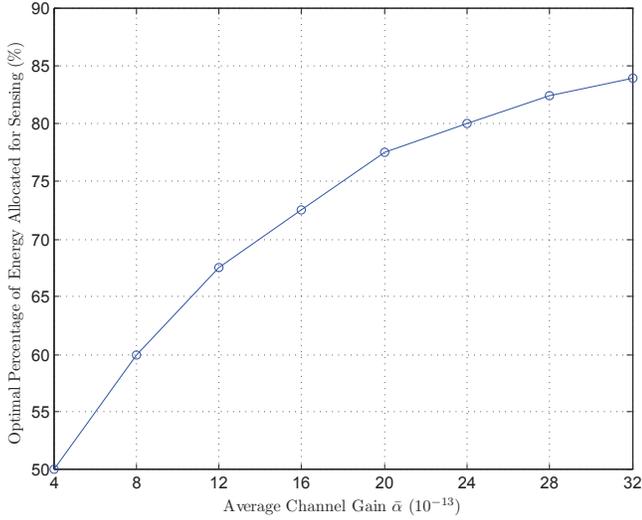


Fig. 7. The optimal percentage of energy allocated for sensing under different average channel gain  $\bar{\alpha}$  for the OTEA algorithm.

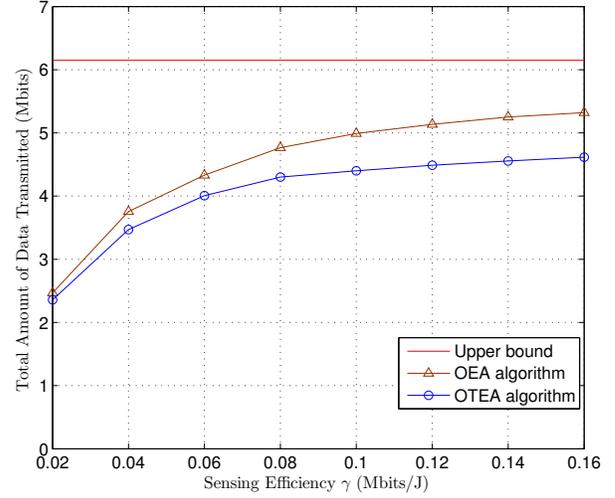


Fig. 9. The total amount of data transmitted of the OEA algorithm and the OTEA algorithm for different values of data-sensing efficiency parameter  $\gamma$ .

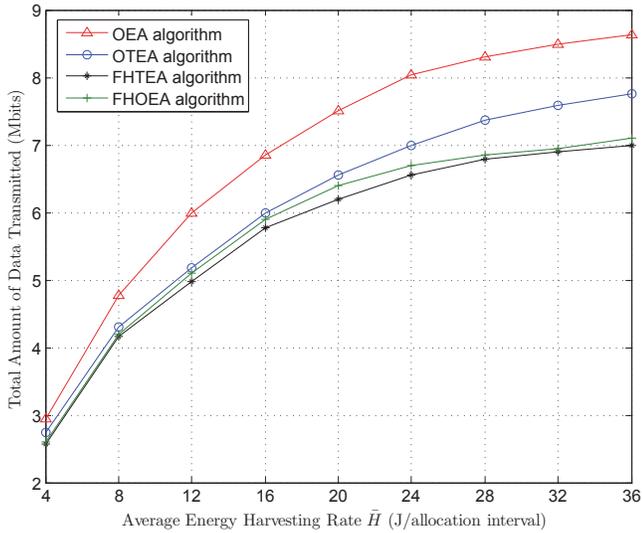


Fig. 8. The total amount of data transmitted of the four algorithms for different average energy harvesting rates  $\bar{H}$ .

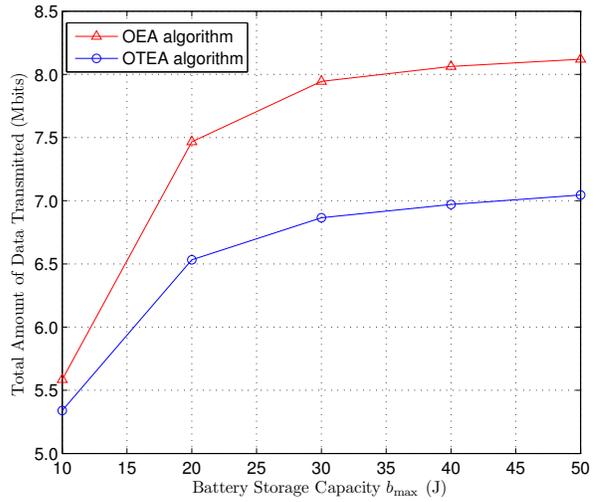


Fig. 10. The total amount of data transmitted of the OEA algorithm and the OTEA algorithm for different battery storage capacity  $b_{max}$ .

the problem (13). Moreover, the OTEA algorithm has a better performance than the FHOEA and FHTEA algorithms. It is because the OTEA algorithm takes into account the randomness of the lifetime of the sensor node, while the FHOEA and FHTEA algorithms just consider the lifetime of the sensor node as a constant. Moreover, the total amount of data transmitted by these four algorithms saturates as the average harvesting rate is increased beyond a certain level. It is because when the energy harvesting rate is larger than the battery capacity, part of the harvested energy cannot be accommodated, and is lost due to the overflow of the battery energy.

Next, we examine the total amount of data transmitted of the OEA and OTEA algorithms under different data-sensing efficiency  $\gamma$ . As shown in Fig. 9, when  $\gamma$  is increased, the total amount of data transmitted increases as well, because more energy is left for data transmission. However, the performance saturates as  $\gamma$  is increased beyond a certain value. To the

extreme when  $\gamma$  approaches infinity, it corresponds to the case where the sensing is extremely efficient. The total amount of data transmitted in this case provides an upper bound for the performance of the OEA algorithm for the sensor node with different sensing efficiency.

Fig. 10 shows the impact of the battery storage capacity  $b_{max}$  on the total amount of data transmitted. We consider that the value of  $h$  is taken from the set  $\mathcal{H} = \{20, 24, 28\}$  J/allocation interval. As shown in Fig. 10, the total amount of data transmitted increases as the battery storage capacity  $b_{max}$  increases. It is because with a larger battery storage capacity  $b_{max}$ , the sensor node can manage the harvested energy better since the sensor can save more energy for future use if necessary. In other words, the sensor has more freedom to manage the incoming energy when  $b_{max}$  is larger. The total amount of data transmitted saturates when  $b_{max}$  goes to large values because under current energy harvesting rates, the battery energy level never exceeds some value, thus for all the battery capacities  $b_{max}$  that are larger than that certain

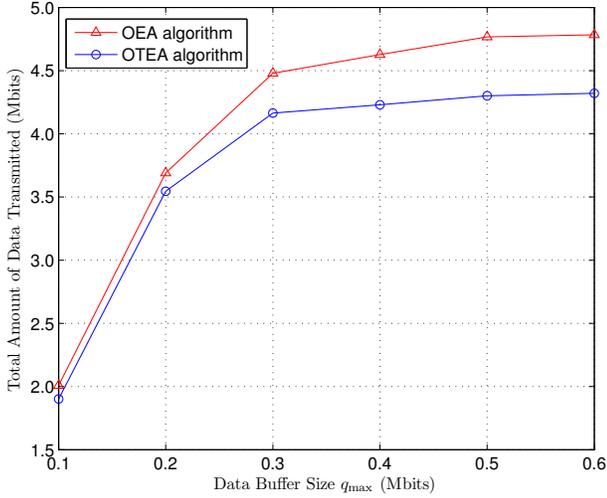


Fig. 11. The total amount of data transmitted of the OEA algorithm and the OTEA algorithm for different data buffer size  $q_{\max}$ .

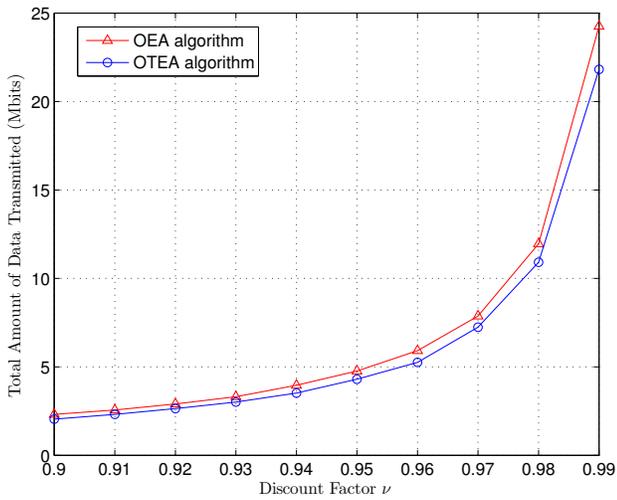


Fig. 12. The total amount of data transmitted of the OEA algorithm and the OTEA algorithm for different values of discount factor  $\nu$ .

value, the sensor has the same performance.

In Fig. 11, we study the impact of the data buffer size  $q_{\max}$  on the total amount of data transmitted by the OEA and OTEA algorithms. We can observe that the total amount of transmitted data increases when  $q_{\max}$  increases. The performance saturates when  $q_{\max}$  is increased to a certain large value. This means that the amount of data in the buffer never exceeds a certain level under the optimal energy allocation policy. Otherwise, we should have observed that the total amount of transmitted data would continue to increase with the data buffer size  $q_{\max}$ .

Finally, we study the total amount of transmitted data of the OEA algorithm and the OTEA algorithm under different discount factors  $\nu$ . Since  $1/(1-\nu)$  represents the average lifetime of the sensor node, a larger  $\nu$  corresponds to a longer lifetime, which leads to a larger total amount of data transmitted. In Fig. 12, the total amount of data transmitted increases as the discount factor  $\nu$  increases. When  $\nu$  approaches 1, where the lifetime of the sensor node approaches infinity, the total amount of transmitted data goes to infinity. Besides, the

number of iterations required for the value iteration algorithm to converge depends on the value of  $\nu$ . With a larger  $\nu$ , a larger number of iterations is required.

## VII. CONCLUSIONS

In this paper, we studied the problem of maximizing the expected total amount of data transmitted for an energy harvesting sensor node under energy harvesting rate variations and channel fluctuations in a time-slotted system. A finite data buffer and the energy consumed for sensing data were considered for the first time. In this case, the sensor should achieve a good tradeoff between the energy consumed for sensing and transmission so as to achieve a large amount of total transmitted data. Since the lifetime of the sensor node is a random variable with geometric distribution, we formulated the problem as an infinite-horizon MDP. We obtained the optimal energy allocation policy and proposed an OEA algorithm based on value iteration in MDP. We also studied the transmission energy allocation problem under the assumption that there was infinite data backlog. We obtained structural results for the OTEA policy and proved that the OTEA policy was a monotonically increasing function of the available battery energy. Finally, we provided extensive simulation results to compare the performances of the OEA, OTEA, FHOEA, and FHTEA algorithms. We studied the impact of the average energy harvesting rate, the data-sensing efficiency, the battery capacity, the data buffer size and the lifetime of the sensor node on the total amount of data transmitted. The results showed that the OEA algorithm transmitted the largest amount of data among the three algorithms. Moreover, we showed that the OTEA algorithm can also be applied to a practical system and achieves a near-optimal performance with a lower computational complexity than the OEA algorithm, when the fixed percentage of energy for sensing was chosen properly using simulation-based approaches. An interesting topic for future work is the extension of our model to a multi-hop setting for data transmission.

## APPENDIX

### A. Proof of Lemma 3

We prove Lemma 3 by mathematical induction. The optimal discounted reward  $\hat{J}(b, h, \alpha)$  is obtained by the value iteration algorithm, given by

$$\hat{J}_{n+1}(b, h, \alpha) = \max_{0 \leq e \leq b} \{ \mathbb{E}_{\alpha'} [\mu(e, \alpha') | \alpha] + \nu \bar{J}_n(b - e, h, \alpha) \}, \quad (27)$$

where

$$\bar{J}_n(\hat{b}, h, \alpha) = \mathbb{E}_{h', \alpha'} [\hat{J}_n(\min\{\hat{b} + h', b_{\max}\}, h', \alpha') | h, \alpha]. \quad (28)$$

Since for any initialization  $\hat{J}_0(b, h, \alpha)$ , the sequence  $\hat{J}_n(b, h, \alpha)$  generated by Eq. (27) converges to the optimal discounted reward  $\hat{J}(b, h, \alpha)$ , we can choose such  $\hat{J}_0(b, h, \alpha)$  that is concave in  $b$  for any given  $h$  and  $\alpha$ . Assume  $\hat{J}_n(b, h, \alpha)$  is concave in  $b$  for any given  $h$  and  $\alpha$ . We denote the optimal action that achieves  $\hat{J}_{n+1}(b_1, h, \alpha)$  by  $e_1$ , and the optimal

action that achieves  $\hat{J}_{n+1}(b_2, h, \alpha)$  by  $e_2$ . Then, we have

$$\hat{J}_{n+1}(b_1, h, \alpha) = \mathbb{E}_{\alpha'}[\mu(e_1, \alpha') | \alpha] + \nu \bar{J}_n(b_1 - e_1, h, \alpha), \quad (29)$$

$$\hat{J}_{n+1}(b_2, h, \alpha) = \mathbb{E}_{\alpha'}[\mu(e_2, \alpha') | \alpha] + \nu \bar{J}_n(b_2 - e_2, h, \alpha). \quad (30)$$

Since  $\mu(e, \alpha')$  is concave in  $e$  for any given  $\alpha'$ ,  $\mathbb{E}_{\alpha'}[\mu(e, \alpha') | \alpha]$  is also concave in  $e$  because it is a weighted sum of concave functions. We then prove that  $\bar{J}_n(\hat{b}, h, \alpha)$  is concave in  $\hat{b}$ . We can follow the procedure of the proof of Lemma 2 and easily prove that  $\hat{J}_n(b', h', \alpha')$  is increasing in  $b'$  for given  $h'$  and  $\alpha'$  for all  $n$ . We already assume at the beginning of the proof that  $\hat{J}_n(b', h', \alpha')$  is concave in  $b'$  for given  $h'$  and  $\alpha'$ . And  $b' = \min\{\hat{b} + h', b_{\max}\}$  is a concave function in  $\hat{b}$  [40]. Thus, by applying the results of composition [40, (3.10)], we can conclude that  $\hat{J}_n(\min\{\hat{b} + h', b_{\max}\}, h', \alpha')$  is concave in  $\hat{b}$  for given  $h'$  and  $\alpha'$ , which indicates that  $\bar{J}_n(\hat{b}, h, \alpha)$  is concave in  $\hat{b}$ , since it is a weighted sum of concave functions.

Now combining Eq. (29) and (30), and using the concavity of  $\mathbb{E}_{\alpha'}[\mu(e, \alpha') | \alpha]$  and  $\bar{J}_n(\hat{b}, h, \alpha)$ , we have

$$\begin{aligned} & \lambda \hat{J}_{n+1}(b_1, h, \alpha) + (1 - \lambda) \hat{J}_{n+1}(b_2, h, \alpha) \\ & \leq \mathbb{E}_{\alpha'}[\mu(e_\lambda, \alpha') | \alpha] + \nu \bar{J}_n(b_\lambda - e_\lambda, h, \alpha), \end{aligned} \quad (31)$$

where  $e_\lambda = \lambda e_1 + (1 - \lambda)e_2$ ,  $b_\lambda = \lambda b_1 + (1 - \lambda)b_2$ . Since  $0 \leq e_1 \leq b_1$  and  $0 \leq e_2 \leq b_2$ , we have  $0 \leq e_\lambda \leq b_\lambda$ . By applying the definition of maximum and  $\hat{J}_{n+1}(b, h, \alpha)$  in Eq. (27), we have

$$\begin{aligned} & \mathbb{E}_{\alpha'}[\mu(e_\lambda, \alpha') | \alpha] + \nu \bar{J}_n(b_\lambda - e_\lambda, h, \alpha) \\ & \leq \max_{0 \leq e \leq b_\lambda} \{ \mathbb{E}_{\alpha'}[\mu(e, \alpha') | \alpha] + \nu \bar{J}_n(b_\lambda - e, h, \alpha) \} \\ & = \hat{J}_{n+1}(b_\lambda, h, \alpha). \end{aligned} \quad (32)$$

Combining inequalities (31) and (32), we have

$$\begin{aligned} & \lambda \hat{J}_{n+1}(b_1, h, \alpha) + (1 - \lambda) \hat{J}_{n+1}(b_2, h, \alpha) \\ & \leq \hat{J}_{n+1}(\lambda b_1 + (1 - \lambda)b_2, h, \alpha). \end{aligned} \quad (33)$$

Inequality (33) shows that  $\hat{J}_{n+1}(b, h, \alpha)$  is concave in  $b$  for given  $h$  and  $\alpha$ . By induction, we can conclude that  $\hat{J}_n(b, h, \alpha)$  is concave in  $b$  for given  $h$  and  $\alpha$  for all  $n$ . Also,  $\bar{J}_n(\hat{b}, h, \alpha)$  is concave in  $\hat{b}$  for all  $n$ . Hence,  $\hat{J}(b, h, \alpha) = \hat{J}_\infty(b, h, \alpha)$  is concave in  $b$  for given  $h$  and  $\alpha$ , and  $\bar{J}(\hat{b}, h, \alpha) = \bar{J}_\infty(\hat{b}, h, \alpha)$  is concave in  $\hat{b}$  for given  $h$  and  $\alpha$ .

## B. Proof of Theorem 2

We prove Theorem 2 by applying [41, Theorem 2]. We aim to prove that  $\hat{e}_{n+1}(b, h, \alpha)$ , which is defined as

$$\begin{aligned} & \hat{e}_{n+1}(b, h, \alpha) = \\ & \min \left\{ e' \in \arg \max_{0 \leq e \leq b} \{ \mathbb{E}_{\alpha'}[\mu(e, \alpha') | \alpha] + \nu \bar{J}_n(b - e, h, \alpha) \} \right\}, \end{aligned} \quad (34)$$

is increasing in  $b$  for given  $h$  and  $\alpha$  for all  $n$ . We drop the arguments of  $h$  and  $\alpha$  from all functions. We denote  $f(e) = \mathbb{E}_{\alpha'}[\mu(e, \alpha') | \alpha]$ , and  $g_n(\hat{b}) = \nu \bar{J}_n(\hat{b}, h, \alpha)$ . Then, Eq. (27) can be written as

$$\hat{J}_{n+1}(b) = \max_{0 \leq e \leq b} \{ f(e) + g_n(b - e) \}. \quad (35)$$

Define  $e_l(b)$  and  $e_u(b)$  to be the lower bound and upper bound of the set of feasible actions  $e$ , respectively, when the available

energy in the battery is  $b$ . In Eq. (35), we have  $e_l(b) = 0$  and  $e_u(b) = b$ , which are both increasing in  $b$ . To apply [41, Theorem 2], it is sufficient to show that  $f(e) + g_n(b - e)$  has increasing difference in  $(b, e)$ , that is, for any  $b' \geq b, e' \geq e$ ,

$$\begin{aligned} & (f(e') + g_n(b' - e')) - (f(e') + g_n(b - e')) \\ & \geq (f(e) + g_n(b' - e)) - (f(e) + g_n(b - e)). \end{aligned} \quad (36)$$

Inequality (36) can be simplified as

$$\begin{aligned} & g_n(b' - e') - g_n(b - e') \geq g_n(b' - e) - g_n(b - e), \\ & \forall b' \geq b, e' \geq e. \end{aligned} \quad (37)$$

From the proof of Lemma 3,  $g_n(\hat{b}) = \nu \bar{J}_n(\hat{b}, h, \alpha)$  is concave in  $\hat{b}$  for all  $n$ . By applying the property of concave functions, we have

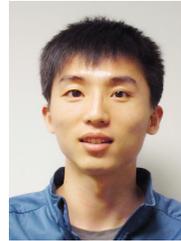
$$g_n(w + \Delta) - g_n(w) \geq g_n(v + \Delta) - g_n(v), \quad \forall w \leq v, \Delta \geq 0. \quad (38)$$

Substituting  $w = b - e', v = b - e, \Delta = b' - b$ , we obtain Eq. (37). Now, by applying the conclusion of [41, Theorem 2], we prove that  $\hat{e}_{n+1}(b, h, \alpha)$  is increasing in  $b$  for any given  $h$  and  $\alpha$  for all  $n$ . Thus,  $\hat{e}^*(b, h, \alpha) = \hat{e}_\infty(b, h, \alpha)$  is increasing in  $b$  for given  $h$  and  $\alpha$ .

## REFERENCES

- [1] S. Mao, M. H. Cheung, and V. W. S. Wong, "An optimal energy allocation algorithm for energy harvesting wireless sensor networks," in *Proc. of IEEE ICC*, Ottawa, Canada, Jun. 2012.
- [2] C. K. Ho and R. Zhang, "Optimal energy allocation for wireless communications powered by energy harvesters," in *Proc. of IEEE Int'l. Symp. on Inform. Theory*, Austin, TX, Jun. 2010.
- [3] D. Niyato, E. Hossain, M. M. Rashid, and V. K. Bhargava, "Wireless sensor networks with energy harvesting technologies: A game-theoretic approach to optimal energy management," *IEEE Wireless Communications*, vol. 14, no. 4, pp. 90–96, Aug. 2007.
- [4] S. Sudevalayam and P. Kulkarni, "Energy harvesting sensor nodes: Survey and implications," *IEEE Communications Surveys and Tutorials*, vol. 13, no. 3, pp. 443–461, 2011.
- [5] A. Kansal, J. Hsu, S. Zahedi, and M. B. Srivastava, "Power management in energy harvesting sensor networks," *ACM Trans. on Embedded Computing Systems*, vol. 6, no. 4, Sept. 2007.
- [6] V. Sharma, U. Mukherji, V. Joseph, and S. Gupta, "Optimal energy management policies for energy harvesting sensor nodes," *IEEE Trans. Wireless Communications*, vol. 9, no. 4, pp. 1326–1336, Apr. 2010.
- [7] M. Gatzianas, L. Georgiadis, and L. Tassiulas, "Control of wireless networks with rechargeable batteries," *IEEE Trans. on Wireless Communications*, vol. 9, no. 2, pp. 581–593, Feb. 2010.
- [8] L. Huang and M. J. Neely, "Utility optimal scheduling in energy harvesting networks," in *Proc. of ACM MobiHoc*, Paris, France, May 2011.
- [9] M. Gorlatova, A. Wallwater, and G. Zussman, "Networking low-power energy harvesting devices: Measurements and algorithms," in *Proc. of IEEE INFOCOM*, Shanghai, China, Apr. 2011.
- [10] R. Srivastava and C. E. Koksal, "Basic performance limits and tradeoffs in energy harvesting sensor nodes with finite data and energy storage," *IEEE/ACM Trans. on Networking*, vol. 21, no. 4, pp. 1049–1062, Aug. 2013.
- [11] Z. Mao, C. E. Koksal, and N. B. Shroff, "Near optimal power and rate control of multi-hop sensor networks with energy replenishment: Basic limitations with finite energy and data storage," *IEEE Trans. on Automatic Control*, vol. 57, no. 4, pp. 815–829, Apr. 2012.
- [12] S. Chen, P. Sinha, N. Shroff, and C. Joo, "A simple asymptotically optimal energy allocation and routing scheme in rechargeable sensor networks," in *Proc. of IEEE INFOCOM*, Orlando, FL, Mar. 2012.
- [13] M. Khouzani, S. Sarkar, and K. Kar, "Optimal routing and scheduling in multihop wireless renewable energy networks," in *Proc. of the Sixth Information Theory and Applications Workshop*, La Jolla, CA, Feb. 2011.

- [14] P. Castiglione, O. Simeone, E. Erkip, and T. Zemen, "Energy management policies for energy-neutral source-channel coding," *IEEE Trans. on Communications*, vol. 60, no. 9, pp. 2668–2678, Sep. 2012.
- [15] Q. Wang and M. Liu, "When simplicity meets optimality: Efficient transmission power control with stochastic energy harvesting," in *Proc. of IEEE INFOCOM*, Turin, Italy, Apr. 2013.
- [16] D. P. Bertsekas, *Dynamic Programming and Optimal Control: Volume 1, 2nd ed.* Athena Scientific, 2000.
- [17] M. L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. New York, NY: John Wiley and Sons, 2005.
- [18] O. Ozel, K. Tutuncuoglu, J. Yang, S. Ulukus, and A. Yener, "Adaptive transmission policies for energy harvesting wireless nodes in fading channels," in *Proc. of IEEE Information Sciences and Systems (CISS)*, Baltimore, MD, Mar. 2011.
- [19] S. Chen, P. Sinha, N. B. Shroff, and C. Joo, "Finite-horizon energy allocation and routing scheme in rechargeable sensor networks," in *Proc. of IEEE INFOCOM*, Shanghai, China, Apr. 2011.
- [20] H. Li, N. Jaggi, and B. Sikdar, "Relay scheduling for cooperative communications in sensor networks with energy harvesting," *IEEE Trans. on Wireless Communications*, vol. 10, no. 9, pp. 2918–2928, Sept. 2011.
- [21] N. Jaggi, K. Kar, and A. Krishnamurthy, "Rechargeable sensor activation under temporally correlated events," *Wireless Networks (WINET)*, vol. 15, no. 5, pp. 619–635, Jul. 2009.
- [22] M. Kashef and A. Ephremides, "Optimal packet scheduling for energy harvesting sources on time varying wireless channels," *Journal of Communication and Networks*, vol. 14, no. 2, pp. 121–129, Apr. 2012.
- [23] P. Blasco, D. Gunduz, and M. Dohler, "A learning theoretic approach to energy harvesting communication system optimization," *IEEE Trans. on Wireless Communications*, vol. 12, no. 4, pp. 1872–1882, Apr. 2013.
- [24] R.-S. Liu, P. Sinha, and C. E. Koksal, "Joint energy management and resource allocation in rechargeable sensor networks," in *Proc. of IEEE INFOCOM*, San Diego, CA, Mar. 2010.
- [25] Z. Wang, A. Tajer, and X. Wang, "Communication of energy harvesting tags," *IEEE Trans. on Communications*, vol. 60, no. 4, pp. 1159–1166, Apr. 2012.
- [26] I. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks: A survey," *Computer Networks*, vol. 38, no. 4, pp. 393–422, Mar. 2002.
- [27] X. Wang, W. Gu, S. Chellappan, K. Schosek, and D. Xuan, "Lifetime optimization of sensor networks under physical attacks," in *Proc. of IEEE ICC*, Seoul, Korea, May 2005.
- [28] X. Wang, S. Chellappan, W. Gu, W. Yu, and D. Xuan, "Search-based physical attacks in sensor networks," in *The 14th International Conference on Computer Communications and Networks (ICCCN)*, San Diego, CA, Oct. 2005.
- [29] R. D. Pietro and N. V. Verde, "Introducing epidemic models for data survivability in unattended wireless sensor networks," in *The 12th IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, Lucca, Italy, Jun. 2011.
- [30] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.
- [31] A. G. Armada, "SNR gap approximation for M-PSK-based bit loading," *IEEE Trans. on Wireless Communications*, vol. 5, no. 1, pp. 57–60, Jan. 2006.
- [32] J. J. E. Garzas, C. B. Calzon, and A. G. Armada, "An energy-efficient adaptive modulation suitable for wireless sensor networks with SER and throughput constraints," *EURASIP Journal on Wireless Communications and Networking*, 2007.
- [33] C. K. Ho, P. D. Khoa, and P. C. Ming, "Markovian models for harvested energy in wireless communications," in *Proc. of IEEE ICCS*, Singapore, Nov. 2010.
- [34] J. Lei, R. Yates, and L. Greenstein, "A generic model for optimizing single-hop transmission policy of replenishable sensors," *IEEE Trans. on Wireless Communications*, vol. 8, no. 2, pp. 547–551, Feb. 2009.
- [35] P. Marbach and J. N. Tsitsiklis, "Simulation-based optimization of Markov reward processes," *IEEE Trans. on Automatic Control*, vol. 46, no. 2, pp. 191–209, Feb. 2001.
- [36] M. L. Littman, T. L. Dean, and L. P. Kaelbling, "On the complexity of solving Markov decision problems," in *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence*, Montreal, Canada, Aug. 1995.
- [37] H. S. Chang, P. J. Fard, S. I. Marcus, and M. Shayman, "Multitime scale Markov decision processes," *IEEE Trans. on Automatic Control*, vol. 48, no. 6, pp. 976–987, Jun. 2003.
- [38] I. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *IEEE Communications Mag.*, vol. 40, no. 8, pp. 102–114, Aug. 2002.
- [39] Q. Zhang and S. A. Kassam, "Finite-state Markov model for Rayleigh fading channels," *IEEE Trans. on Communications*, vol. 47, no. 11, pp. 1688–1692, Nov. 1999.
- [40] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
- [41] R. Amir, "Supermodularity and complementarity in economics: An elementary survey," *Southern Economic Journal*, vol. 71, no. 3, pp. 636–660, Jan. 2005.



**Shaobo Mao** received the B.Eng. degree in Electrical Engineering from Zhejiang University in China in 2010, and the M.A.Sc. degree in Electrical and Computer Engineering from the University of British Columbia (UBC) in 2012. Currently, he is a design engineer in Teradici Corporation in Vancouver, Canada. His research interests include the design and analysis of energy harvesting wireless sensor networks using optimization techniques and dynamic programming.



**Man Hon Cheung** (S'06, M'13) received the B.Eng. and M.Phil. degrees in Information Engineering from the Chinese University of Hong Kong (CUHK) in 2005 and 2007, respectively, and the Ph.D. degree in Electrical and Computer Engineering from the University of British Columbia (UBC) in 2012. Currently, he is a postdoctoral fellow in the Department of Information Engineering in CUHK. He received the IEEE Student Travel Grant for attending *IEEE ICC 2009*. He was awarded the Graduate Student International Research Mobility Award by UBC, and the Global Scholarship Programme for Research Excellence by CUHK. He serves as a Technical Program Committee member in *IEEE ICC*, *Globecom*, and *WCNC*. His research interests include the design and analysis of wireless network protocols using optimization theory, game theory, and dynamic programming, with current focus on mobile data offloading and network economics.



**Vincent W.S. Wong** (SM'07) received the B.Sc. degree from the University of Manitoba, Winnipeg, MB, Canada, in 1994, the M.A.Sc. degree from the University of Waterloo, Waterloo, ON, Canada, in 1996, and the Ph.D. degree from the University of British Columbia (UBC), Vancouver, BC, Canada, in 2000. From 2000 to 2001, he worked as a systems engineer at PMC-Sierra Inc. He joined the Department of Electrical and Computer Engineering at UBC in 2002 and is currently a Professor. His research areas include protocol design, optimization, and resource management of communication networks, with applications to the Internet, wireless networks, smart grid, RFID systems, and intelligent transportation systems. Dr. Wong is an Associate Editor of the *IEEE Transactions on Communications*. He has served on the editorial boards of *IEEE Transactions on Vehicular Technology* and *Journal of Communications and Networks*. Dr. Wong is the Technical Program Co-chair of *IEEE SmartGridComm'14*. He has served as the Symposium Co-chair of *IEEE SmartGridComm'13 – Communications Networks for Smart Grid and Smart Metering Symposium* and *IEEE Globecom'13 – Communication Software, Services, and Multimedia Application Symposium*.