

# Multiple Timescale Energy Scheduling for Wireless Communication with Energy Harvesting Devices

Hua XIAO<sup>1</sup>, Huaizong SHAO<sup>1</sup>, Kai YANG<sup>2</sup>, Fan YANG<sup>1</sup>, Wenqin WANG<sup>1</sup>

<sup>1</sup> School of Communication and Information Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China

<sup>2</sup> Bell Laboratories, Murray Hill, NJ, USA

founder@uestc.edu.cn

**Abstract.** *The primary challenge in wireless communication with energy harvesting devices is to efficiently utilize the harvesting energy such that the data packet transmission could be supported. This challenge stems from not only QoS requirement imposed by the wireless communication application, but also the energy harvesting dynamics and the limited battery capacity. Traditional solar predictable energy harvesting models are perturbed by prediction errors, which could deteriorate the energy management algorithms based on this model. To cope with these issues, we first propose in this paper a non-homogenous Markov chain model based on experimental data, which can accurately describe the solar energy harvesting process in contrast to traditional predictable energy models. Due to different timescale between the energy harvesting process and the wireless data transmission process, we propose a general framework of multiple timescale Markov decision process (MMDP) model to formulate the joint energy scheduling and transmission control problem under different timescales. We then derive the optimal control policies via a joint dynamic programming and value iteration approach. Extensive simulations are carried out to study the performances of the proposed schemes.*

## Keywords

Energy scheduling, energy harvesting, multiple timescale Markov decision process, transmission control, wireless communication.

## 1. Introductions

Due to the time-varying nature of wireless channel and the limited energy capacity of wireless transmission nodes, one of the challenges in wireless data networks is to efficiently manage transmission energy consumption and support the specific missions [1-5], e.g. transmission delay constraints, QoS guarantee. Energy harvesting technologies [6], [13] obtain energy from environment, which provide effective means to extend the life-span of wireless commu-

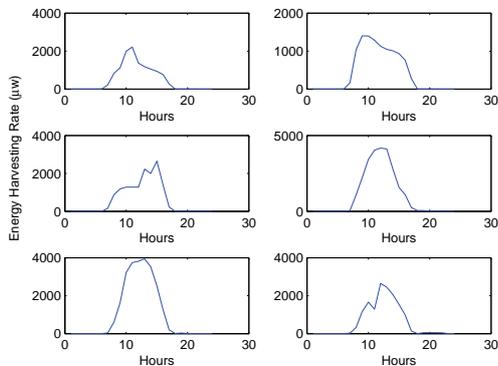
nication and consequently make the networks more resilient and sustainable.

In this paper, we focus on the energy scheduling and transmission control problems in wireless data networks with solar energy harvesting devices. Our objective is to design efficiently energy scheduling and transmission policies, which make transmission decisions according to the energy harvesting dynamic and meet transmission requirements.

In what follows, we survey the related works on the wireless energy scheduling problem with energy harvesting devices. The paper [7] studies the energy management policies for wireless transmission systems, which maximize the achievable rate under the stability of the data queue line. These policies are derived on the assumption that the energy harvesting process is a stationary and ergodic process. Different from the assumption that the energy buffer is infinite in [7], the battery capacity constraint is introduced into the energy management problems in [8], [9]. In addition, the transmission policies that maximize short-term throughput in a two-user interference channel are proposed in [8]. The paper [10] proposes the off-line scheduling policies for a single-user communication channel, which minimize the time by which all transmission packets are delivered. It is assumed that how much and when the energy will be harvesting is exactly known. The paper [11] extends this energy management problem to multiple AWGN broadcast channels. The paper [12] considers the similar problem in a fading channel and proposes policies that maximize throughput with a limited battery capacity constraint. Existing approaches assume that the energy traces are exactly known and they consider the transmission process and the energy harvesting process in the same time-scale. However, the true energy harvesting profiles are quite different from these assumptions.

For example, the paper [7] assumes that the solar energy harvesting process is stationary. When we observe the solar energy harvesting profile, the statistic of the harvesting process is time-varying as shown in Fig. 1. Moreover, the predictable energy harvesting model is proposed in recent works [13-15]. It assumes that the energy harvest

ing rate is deterministic and predictable. This assumption is supported by the periodicity of position change and activities of the sun. However, due to weather influences, season variation, and measurement errors, the actual energy harvesting traces in Fig. 1 are quite different.



**Fig. 1.** Solar energy harvesting data in six days (the area of energy harvesting panel is  $2 \text{ cm}^2$ ).

Although the exponentially weighted moving-average (EWMA) filter is introduced to obtain the average energy harvesting rate from past data, the prediction errors are still high, for example, 16% in [14]. The model errors could considerably deteriorate the performances of the above energy management algorithms. Therefore, we propose a non-homogenous Markov chain to characterize the solar energy harvesting process, which can accurately track the main trend and random event in this process compared to existing models.

Another issue in the existing approaches is that they consider the transmission process and the energy harvesting process in the same time-scale. In practice, the change speed of the energy harvesting process is much slower than the channel variation. That is, the energy harvesting rate keeps stable in a relatively long time period when the wireless channel experiences very fast changing. However, previous works ignore the different timescale issue between the transmission process and the energy harvesting process when the transmission control is considered into energy management problems. These motivate us to investigate the energy scheduling problem under different time-scale. As far as we know, our proposed approaches in this paper first address the different time-scale issue in the energy harvesting management problems.

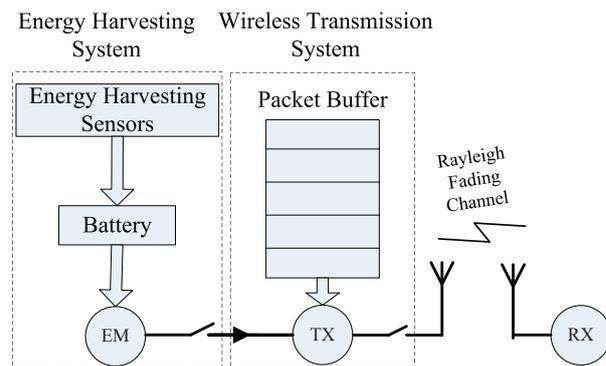
Different from existing energy harvesting management approaches, the main contributions of this paper are as follows. We first propose a novel non-homogenous Markov chain model for the solar energy harvesting process. In contrast to the predictable energy harvesting model [13], [14], it can accurately characterize the main trend and random events in the harvesting process. Secondly, we consider the different timescale issue between the energy harvesting process and the wireless transmission process. Different from the energy management problem in single level [7-11], [13], [16], we propose a general framework

based on the multiple timescale Markov decision process (MMDP) [17] to model the coupled energy scheduling and transmission control problems. Next, we derive the optimal solution for the MMDP problem. Finally, we use the true energy harvesting data to evaluate the proposed methods and the simulation results verify the effectiveness of these approaches.

The remainder of the paper is organized as follows. In Section 2, we describe the system model. The MMDP problem is formulated in Section 3. In Section 4, we derive the optimal solution for the MMDP problem. Simulation results are given in Section 5. Section 6 concludes the paper.

## 2. System Descriptions

In this section, we present the wireless transmission system model and introduce the multiple timescale MDP model. As shown in Fig. 2, the system is composed of the energy harvesting subsystem and the wireless transmission subsystem.



**Fig. 2.** The system model including the energy harvesting subsystem and the wireless transmission subsystem.

We assume that all system parameters are discretized and the time is slotted. The system parameters keep constant in each time slot. Let  $n$  denote the  $n$ -th energy harvesting time slot, where  $n=1, \dots, N$ , and  $T$  denote the length of this time slot. Let  $H_n$  denote the energy harvesting rate in time slot  $n$  and  $a_n$  denote the energy consumption rate in time slot  $n$ . The energy harvesting devices obtain environmental energy with rate  $H_n$  in time slot  $n$  and store it into the battery with a limited capacity. Let  $C_H$  denote the maximum battery capacity. The energy management (EM) part schedules energy consumption rate  $a_n$  over a whole day. The transmission subsystem makes transmission decisions according to the waiting packet number and channel conditions under the total energy constraint. Due to the difference of time-varying characteristics between the energy harvesting process and the wireless channel, each energy harvesting time slot is further divided into sub-time slots as shown in Fig. 3. Let  $t^n$  denote the  $t$ -th sub-time slot during the energy harvesting time slot  $n$  and  $T_0$  denote the length of the sub-time slot. Our objective is to design the joint energy scheduling and transmission control policies that

minimize the expected total cost over  $N$  time slots, where the cost takes account of both the transmission cost and the delay cost. In what follows, we cast this problem as an MMDP, which consists of two level MDPs as shown in Fig. 3, where  $X_n$  and  $Y_t$  denote the system state of the upper level MDP and the lower level MDP, respectively.  $p(X_{n+1}|X_n, a_n)$  and  $p(Q_{t+1}|Q_t, I)$  denote the related transition probability, respectively. The upper level MDP is a finite horizon MDP, which is corresponding to the energy harvesting dynamic with the timescale  $T$ . The lower level MDP is an infinite horizon MDP problem, which is triggered at the beginning of each energy harvesting time slot and corresponding to the channel dynamic with the fast timescale  $T_0$ .

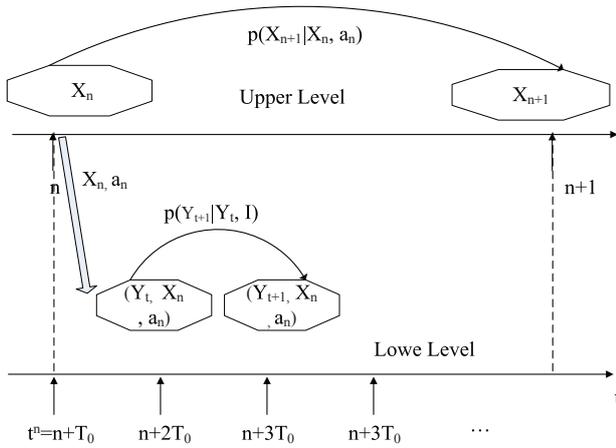


Fig. 3. The multiple timescale MDP model. The upper level MDP based on time scale  $T$  and the lower level MDP based on time scale  $T_0$ .

### 3. The MMDP Formulation

In general, an MDP is defined by a 5-tuple set  $\{X, A, \mathbf{P}, f, \Pi\}$ , where  $X$  denotes the system state space,  $A$  denotes the action set,  $\mathbf{P}$  denotes the transition matrix set of system states,  $f$  is the cost function of system states and actions,  $\Pi$  is the policy space that is a set of decision sequences, where each decision in the sequence is a mapping from system states to actions. In what follows, we will describe the two level MDPs according to the definition of the 5-tuple set.

For the upper level, given the finite system state space  $X$ , let  $X_n = \{C_n, H_n\} \in X$  denote the system state in time slot  $n$ , where  $C_n$  is the available energy level in the battery. Due to the battery capacity limitation, we have  $0 \leq C_n \leq C_H$ . Given the finite action space  $A$ , let  $a_n \in A$  denote the energy consumption rate in time slot  $n$ , which determines the total available energy for the lower level MDP in time slot  $n$ . Thus, the dynamic of the energy level  $C_n$  is given by  $C_{n+1} = C_n - (a_n - H_n)T$  and the total energy for the transmission process in time slot  $n$  is constrained by  $a_n T$ . The transition probability of system states is given by

$$p(X_{n+1} | X_n, a_n) = p(C_{n+1} | C_n, a_n) p(H_{n+1} | H_n) \quad (1)$$

where the energy harvesting dynamic is independent with  $a_n$ .

At the beginning of the time slot  $n$ , given the action  $a_n$  and  $X_n$  of the upper level, the lower level MDP with the total energy constraint  $a_n T$  is triggered. Given the finite system space  $Y$ , let  $Y_t \equiv \{G_t, B_t, Q_t\} \in Y$  denote the system state in sub-time slot  $t^n$ , where  $G_t$  is the channel state and  $G_t \in G$ ,  $B_t$  is the remaining budget of the total transmission energy constraint during the time slot  $n$ , and  $Q_t$  is the number of waiting packets. Note that  $X \cap Y = \emptyset$  and the lower level MDP actions cannot influence the system state and system dynamics of the upper level. Let  $I(Y_t) \in \{0, 1\}$  denote the action of system state  $Y_t$ , where  $I = 1$  is corresponding to the action of sending packet and  $I = 0$  is corresponding to the opposite operation. We assume that the required SNR of receivers is set as a constant. Therefore, we have  $E_c = G_t P_t / \sigma^2$ , where  $E_c$  is the required SNR of receivers,  $P_t$  is the transmission power, and  $\sigma^2$  is the noise power. Thus, the dynamic of  $B_t$  is given by  $B_{(t+1)^n} = B_t - I(Y_t) P_t T_0$ , where  $B_{(t+1)^n} \geq 0$  and  $B_{1^n} = a_n T$ . Furthermore, the transition probability of the system states is given by

$$p(Y_{(t+1)^n} | Y_t) = p(G_{(t+1)^n} | G_t) \times p(Q_{(t+1)^n} | Q_t, I(Y_t)) p(B_{(t+1)^n} | B_t, I(Y_t), a_n) \quad (2)$$

where  $Q_{(t+1)^n} = Q_t - I(Y_t) + \zeta_t$  and  $\zeta_t$  is the number of arrival packets in sub-time slot  $t^n$ . The immediate cost  $f_L(Y_t, I(Y_t))$  is defined by

$$f_L(Y_t, I(Y_t)) = I(Y_t) P_t + \mu f_Q(Q_t) \quad (3)$$

where  $f_Q$  is a non-decreasing function of the waiting packet length, and  $\mu$  is an adjustable factor, which effects the proportion of the energy cost and the delay cost. (3) is used to achieve the tradeoff between the energy consumption and the delayed packets under the energy constraint. In order to control the number of delayed packets, we use different penalty when the number of waiting packets exceeds the prescribed delay threshold. Thus,  $f_Q(Q_t)$  is defined by a piecewise function as follows

$$f_Q(Q_t) = \begin{cases} Q_t P_w, & Q_t \leq Q_c \\ P_c, & Q_t > Q_c \end{cases} \quad (4)$$

where  $P_w$  is the transmission power based on the worst case channel conditions,  $Q_c$  is the prescribed delay threshold, and  $P_c$  is the penalty, where  $P_c \gg Q_c P_w$ .

For the solar energy harvesting process,  $T \gg T_0$  as shown in Fig. 1,  $T$  is set as 1 hour and  $T_0$  is set as 5 ms. It is intractable to find the non-stationary transmission policy for such huge amount of sub-time slots. Thus, we model the lower level MDP as a discount infinite horizon MDP problem to obtain the stationary transmission control policy. Let  $\pi_L^n \in \Pi_L$  denote the stationary policy of the lower level in time slot  $n$ , where  $\Pi_L$  is the policy space of the lower level MDP. In time slot  $n$ , given system state  $X_n$  and action  $a_n$ , the total value function of the lower level MDP is

defined by

$$V(Y_{1^n}) = \mathbb{E} \left\{ \sum_{t^n=1}^{+\infty} \eta^{t^n} f_L(Y_{t^n}, I(Y_{t^n})) | a_n \right\} \quad (5)$$

where  $0 < \eta < 1$  is the discount factor,  $Y_{1^n}$  is the initial system state, and the expectation is over the system state  $Y_{t^n}$ .

For the upper level, given the system state  $X_n$  and the action  $a_n$ , the immediate cost in time slot  $n$  is defined by  $f_U(X_n, a_n, \pi_L^n) = \mathbb{E}\{V(Y_{1^n})\}$ , where the expectation is over the initial system state. Note that different  $\pi_L^n$  is corresponding to different  $V(Y_{1^n})$ . Thus, the two level MDPs are connected by  $a_n$  and  $V(Y_{1^n})$ . Let  $\pi_U \equiv (a_1, \dots, a_N)$  denote the policy of the upper level, where  $\Pi_U$  is the policy space. The joint energy scheduling and transmission control problem is cast as follows

$$\min_{\pi_U \in \Pi_U} \min_{\pi_1^1, \dots, \pi_L^N \in \Pi_L} \mathbb{E} \left\{ \sum_{n=1}^N f_U(X_n, a_n, \pi_L^n) \right\} \quad (6)$$

where the expectation is over the system states  $X_n$ . In what follows, we will describe the energy harvesting dynamic and channel dynamic.

### 3.1 The Non-homogenous Markov Energy Harvesting Model

The solar energy harvesting process can be decomposed into two parts, i.e., the deterministic process and the random process. Let  $D_1, \dots, D_N$  denote the deterministic process, which takes account for the periodical behaviors such as the sun position change. Let  $e_1, \dots, e_N$  denote the random process, which is determined by environmental change such as the weather transition from sunshine to rainfall. We have  $H_n = D_n + e_n$ . Note that both  $D_n$  and  $e_n$  are discretized. Non-homogenous Markov chain can be used to model the weather state variation [18-20]. Since  $e_n$  is mainly determined by the weather change, it is reasonable to assume that  $e_n$  has non-homogenous Markov property. Thus, the transition probability of the energy harvesting process is such that

$$p(H_n | H_{n-1} = D_{n-1} + e_{n-1}, \dots, H_1 = D_1 + e_1) = p(H_n | H_{n-1}). \quad (7)$$

In practice, we can use historical measurement data to estimate  $p(H_n | H_{n-1})$ . Let  $\mathbb{H} = \{h_0, \dots, h_{K-1}\}$  denote the energy harvesting state space, where  $H_n \in \mathbb{H}$  and  $|\mathbb{H}|=K$ . The energy harvesting dynamic can be characterized by  $\{\mathbf{P}_1, \dots, \mathbf{P}_N\}$ , where  $\mathbf{P}_n$  is the transition matrix in time slot  $n$ , which is given by

$$\mathbf{P}_n = [p_{i,j,n}]_{h_i, h_j \in \mathbb{H}} \quad (8)$$

where  $p_{i,j,n} \equiv p(H_n = h_j | H_{n-1} = h_i)$ . (7) and (8) characterize the non-homogenous Markov chain model for the solar energy harvesting process. In contrast to the predictable energy model [13], this model can accurately capture the deterministic trend and the random events during each time slot. In the next subsection, we will describe the channel model based on the time-scale  $T_0$ .

### 3.2 The Discrete-time Markov Chain Channel

We assume that the channel experience Rayleigh fading and characterize it by a discrete-time Markov chain model. Let  $\mathbb{G} = \{g_0, \dots, g_{L-1}\}$  denote the channel state space, where  $|\mathbb{G}| = L$ . We use the received SNR to characterize the channel states. Let  $\tau_0, \dots, \tau_L$  denote the thresholds of the received SNR in ascending order, where  $\tau_0 = 0$  and  $\tau_L = +\infty$ . The thresholds divide the received SNR into  $L$  intervals, which are corresponding to the channel states, for example,  $g_i$  is corresponding to the SNR interval  $[\tau_i, \tau_{i+1})$ . We assume that the sub-time slot is sufficiently long to guarantee that current channel state can only transit to its adjacent states or itself. Thus, the transition probability  $p(g_{i+1}|g_i)$  and  $p(g_{i-1}|g_i)$  are given by [21]

$$p(g_{i+1} | g_i) = \frac{N(\lambda_{i+1})T_0}{p_i} \quad (9)$$

and  $p(g_{i-1} | g_i) = \frac{N(\lambda_i)T_0}{p_i}$

where  $N(\lambda_i)$  is the threshold crossing rate of the received SNR [21] and  $p_i$  is the stationary probability in the state  $g_i$ , which is given by

$$p_i = \int_{\tau_i}^{\tau_{i+1}} p(\lambda) d\lambda. \quad (10)$$

Due to Rayleigh fading, the received SNR has exponential distribution. We have

$$p(\lambda) = \frac{1}{\lambda} e^{-\frac{\lambda}{\lambda_0}} \quad (11)$$

where  $\lambda$  is the received SNR and  $\lambda_0$  is the average SNR. The channel dynamic in each energy harvesting time slot is completely characterized by the transition matrix, which is given by

$$[p(g_j | g_i)]_{g_i, g_j \in \mathbb{G}} \quad (12)$$

where  $g_j$  is the current channel state and  $g_i$  is the channel state in the previous sub-time slot. In the next section, we will derive the optimal solution of the problem (6).

## 4. Optimal Solution for the MMDP

The optimal solution for the problem (6) is a sequence of  $(a_n^*, \pi_L^{n*})$  that minimizes the expected sum of  $f_U(X_n, a_n, \pi_L^n)$  over  $N$  time slots, which can be obtained via the following method.

**Theorem 1:** For each initial system state  $X_1$ , the optimal policy pair  $(a_n^*, \pi_L^{n*})$  can be obtained by implementing the following backward recursions from the time slot  $N$  to 1

$$\begin{cases} \pi_L^{N*} = \arg \min_{\pi_L^N \in \Pi_L} \mathbb{E} \left\{ \sum_{t^N=1}^{+\infty} \eta^{t^N} f_L(Y_{t^N}, I(Y_{t^N})) | a_N \right\}, \\ a_N^* = \arg \min_{a_N \in \mathcal{A}} \{f_U(X_N, a_N, \pi_L^{N*})\}, \\ J_N(X_N) = f_U(X_N, a_N^*, \pi_L^{N*}) \end{cases} \quad (13)$$

and

$$\left\{ \begin{array}{l} \pi_L^{n*} = \arg \min_{\pi_L^n \in \Pi_L} E \left\{ \sum_{t^n=1}^{+\infty} \eta^{t^n} f_L(Y_{t^n}, I(Y_{t^n})) \mid a_n \right\}, \\ a_n^* = \arg \min_{a_n \in A} \left\{ f_U(X_n, a_n, \pi_L^{n*}) + \sum_{X_{n+1} \in X} p(X_{n+1} \mid X_n, a_n) J_{n+1}(X_{n+1}) \right\}, \\ J_n(X_n) = f_U(X_n, a_n^*, \pi_L^{n*}) + \sum_{X_{n+1} \in X} p(X_{n+1} \mid X_n, a_n^*) J_{n+1}(X_{n+1}), \\ n=1, \dots, N-1. \end{array} \right. \quad (14)$$

**Proof:** At the beginning of the time slot  $n$  in the upper level, given system state  $X_n$  and action  $a_n$ , let  $\pi_L^n(a_n) \in \Pi_L$  denote the lower level policy corresponding to  $a_n$ . We define a new finite horizon MDP over  $N$  time slots, which has the same system state space as the upper level MDP. The action of the new MDP is a composite of  $a_n$  and  $\pi_L^n(a_n)$ , defined by

$$\{(a_n, \pi_L^n(a_n)) \mid a_n \in A, \pi_L^n(a_n) \in \Pi_L\}. \quad (15)$$

Since the lower level policy  $\pi_L^n(a_n)$  is triggered by the action  $a_n$  and independent with the dynamic of  $X_n$ , we have  $p(X_{n+1} \mid X_n, a_n) = p(X_{n+1} \mid X_n, a_n, \pi_L^n(a_n))$ . Furthermore, given  $X_n$  and  $(a_n, \pi_L^n(a_n))$ , the cost function of the new MDP is defined by  $f_U(X_n, a_n, \pi_L^n)$ . Thus, the multiple level MDP problem (6) can be reduced to the one level MDP problem and the optimal policy of the new MDP is the solution of problem (6). For each time slot  $n$ , the cost function  $f_U(X_n, a_n, \pi_L^n)$  is obtained via value iteration algorithm [22] and the optimal action  $(a_n^*, \pi_L^n(a_n^*))$ , can be efficiently solved by dynamic programming [23] as shown in (13) and (14). ■

In (13) and (14), given  $X_n$  and  $a_n$ , there is an infinite MDP problem defined by (5), of which the optimal value function is given by

$$V^*(y) = \min_{\pi_L^n \in \Pi_L} \left\{ f_L(y, I(y) \mid a_n) + \sum_{z \in Y} p(z \mid y, I(y)) V^*(z) \right\} \quad (16)$$

where  $y, z \in Y$ . The optimal policy  $\pi_L^{n*}$  can be obtained via the value iteration algorithm [22]. Furthermore, for each system state  $X_n$ , the optimal solution of the problem (6) can be obtained by the following algorithm.

**Algorithm 1:** Input: the total number of upper level time slots  $N$ , the system states space  $X$  and  $Y$ , the action set  $A$ ; Output: the optimal policy.

1. Initialize  $n \leftarrow N$ ;
2. While  $(n \neq 0)$  do
3. For  $(X_n \in X)$  do
4. For  $(a_n \in A)$  do
5. Call value iteration algorithm [22] to obtain  $\pi_L^{n*}$  and  $V^*(y)$ , where  $y \in Y$ ;
6. End for

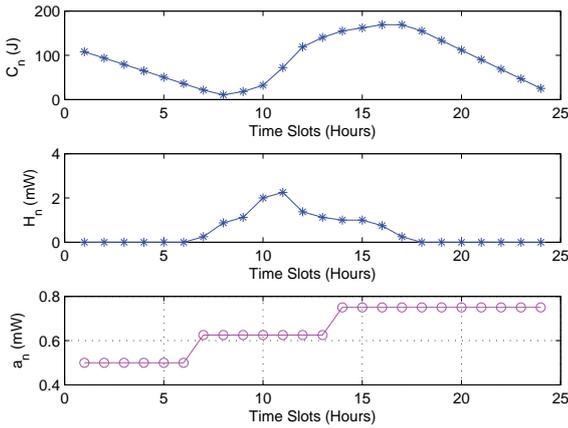
7. if  $(n == N)$  do
8. Call (13)
9. Else
10. Call (14)
11. End if
12. Save  $(a_n^*, \pi_L^{n*})$  into the optimal policy table
13. End for
14.  $n \leftarrow n - 1$
15. End while

Note that Algorithm 1 is implemented in an off-line way and the optimal policy is stored in a look-up table. When the system is implementing the policies, it simply searches the table for optimal actions according to the current system state. In the next section, we will show the simulation results to verify our proposed approach.

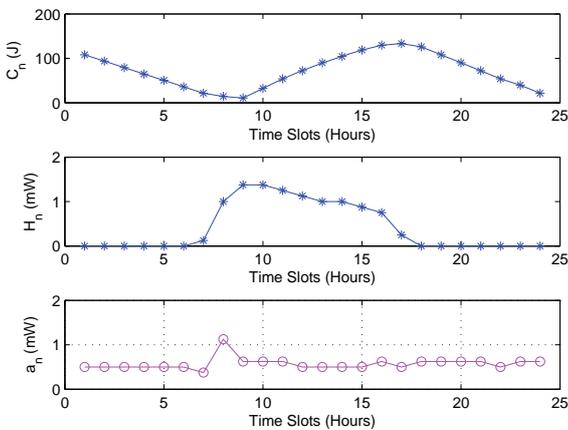
## 5. Simulation Results

In the simulation studies, we use the actual energy harvesting data to evaluate the performance of the proposed approaches, which was measured at the conference room of the department of electrical engineering at Columbia University [13]. The TAOS TSL230rd photometric sensors are equipped on LabJack U3 DAQ devices to harvest the light and solar energy, and the unit of the measurement data is  $\mu\text{W}/\text{cm}^2$ . The area of the energy harvesting panel in our simulations is set as  $2 \text{ cm}^2$ . For the upper level, the time slot length is set as 1 hour and the total time period is 24 hours. The energy harvesting data traces for simulation studies were measured from November 6, 2009 to September 13, 2010 [13]. For the lower level, the channel experiences Rayleigh fading and the sub-time slot length is set as 5 ms. The received channel SNR thresholds are given by  $\tau_0 = 0$ ,  $\tau_i = \tau_{i-1} + 4 \text{ dB}$ ,  $i = 1, 2, 3$ ,  $\tau_4 = +\infty$ , which divide the channel into 4 states. The transition probabilities of the channel dynamic are obtained by (9). We assume that only one packet can be sent in one sub-time slot. The packet arrival model is set as a Poisson process with the arrival rate 0.1, the discount factor  $\eta$  is set as 0.97, and  $\mu$  in (3) is set as 1. The delay threshold depends on the specific transmission application. Without loss of generality, the delay threshold is set as 50 ms and 75 ms, where the maximum amount of waiting packets is 10 and 15, respectively. We assume that the new arrival packets will be dropped when the amount of waiting packets exceeds the delay threshold. We use the number of delayed packets to evaluate the performance of delay control.

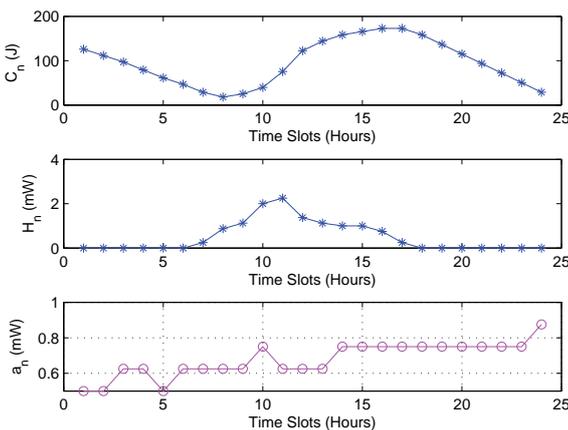
Fig. 4 and Fig. 5 show the optimal energy consumption rate and the corresponding energy level of the battery in different days. The energy harvesting traces are different in the two days. In both figures, the energy harvesting rate reaches peak in the noon, and the optimal policy is inclined to reserve energy for future usage. It is due to the fact that there is less harvesting energy in the mid-night and the



**Fig. 4.** The optimal actions  $\{a_1, \dots, a_{24}\}$  based on the energy harvesting trace in a whole day (the area of energy harvesting panel is  $2 \text{ cm}^2$ ,  $\mu = 1$ ,  $\eta = 0.97$ ,  $C_1 = 108 \text{ J}$ ).



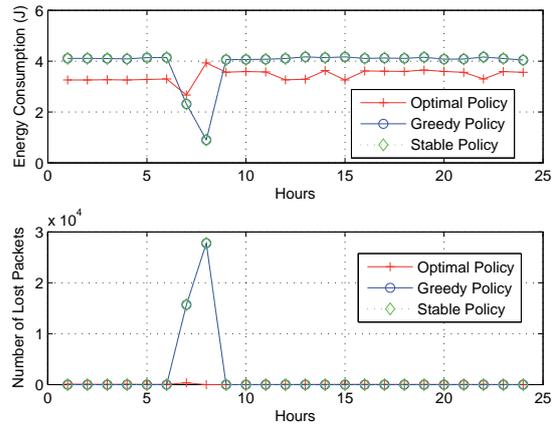
**Fig. 5.** The optimal actions  $\{a_1, \dots, a_{24}\}$  based on the energy harvesting trace in a whole day (the area of energy harvesting panel is  $2 \text{ cm}^2$ ,  $\mu = 1$ ,  $\eta = 0.97$ ,  $C_1 = 108 \text{ J}$ ).



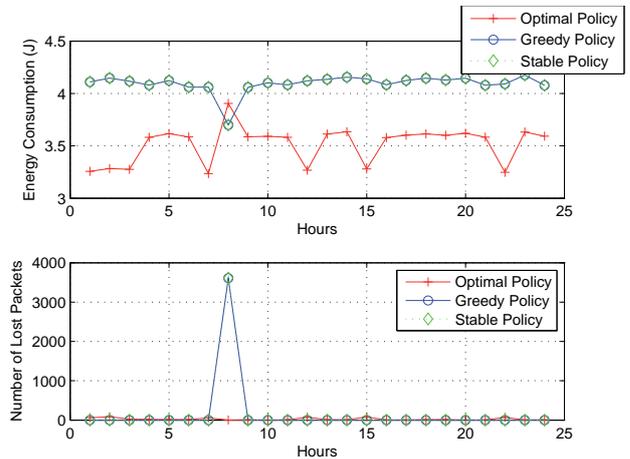
**Fig. 6.** The optimal actions  $\{a_1, \dots, a_{24}\}$  based on the energy harvesting trace in a whole day (the area of energy harvesting panel is  $2 \text{ cm}^2$ ,  $\mu = 1$ ,  $\eta = 0.97$ ,  $C_1 = 126 \text{ J}$ ).

optimal policy schedules the energy consumption to adapt to the future energy demands from the lower level transmission process. The available energy level in the battery can tightly meet the constraint. Fig. 6 shows the optimal actions based on the energy trace in Fig. 4 with different initial energy level  $C_1$ . Compared to Fig. 4, more available

energy is allocated to the lower level transmission process when increasing the initial available energy level  $C_1$ .



**Fig. 7.** Performance comparison among the optimal policy, the greedy policy, and the stable policy (the area of energy harvesting panel is  $2 \text{ cm}^2$ ,  $\mu = 1$ ,  $\eta = 0.97$ ,  $C_1 = 108 \text{ J}$ ,  $Q_c = 10$ ).



**Fig. 8.** Performance comparison among the optimal policy, the greedy policy, and the stable policy (the area of energy harvesting panel is  $2 \text{ cm}^2$ ,  $\mu = 1$ ,  $\eta = 0.97$ ,  $C_1 = 126 \text{ J}$ ,  $Q_c = 10$ ).

In order to compare the performance of the proposed approach, we consider two heuristic algorithms. First, we introduce a greedy algorithm, which implements a spend-to-go policy, that is, when the packet buffer is not empty, the waiting packets are sent immediately until the buffer is empty. When there is no available energy and the amount of waiting packets exceeds the prescribed threshold, the new arrival packet will be dropped. This policy is similar to the policies that minimize the transmission packet delay with battery capacity constraint [10-12]. Second, we consider the policy that maximize throughput with stable waiting packet length [7-9]. We call this policy as stable policy. Note that both the stable policy and the greedy policy do not consider the different timescale issue. We compare the average energy consumption rate and the average packet loss of these policies as shown in Fig. 7. The optimal policy has less energy consumption than the greedy policy and the stable policy most of the time. Since the greedy policy

and the stable policy aggressively use the available energy, there is no sufficient energy in the battery for transmission near 7 am. The energy consumption of the greedy policy and the stable policy drops down and the number of dropped packets simultaneously increases sharply. The performance of the greedy policy is very close to the one of the stable policy. It is because that the maximum throughput is constant, when the received SNR is fixed in the proposed model. And the stable policy sends the packet immediately to keep the waiting packet length stable, which is similar to the greedy policy. On the other hand, the optimal policy reduces the total energy consumption by 10% compared to the other policies even under the situation that the other policies are out of service for two hours. Fig. 8 shows the performances of these policies under different initial available energy  $C_1$  compared to Fig. 7. Since the initial available energy level increases, the packet loss of the greedy policy and the stable policy decreases. The simulation results that the delay threshold extends to 15 are shown in Fig. 9. Since the delay threshold is quite smaller than the time slot length of the upper level, the performance change is trivial compared to Fig. 7.

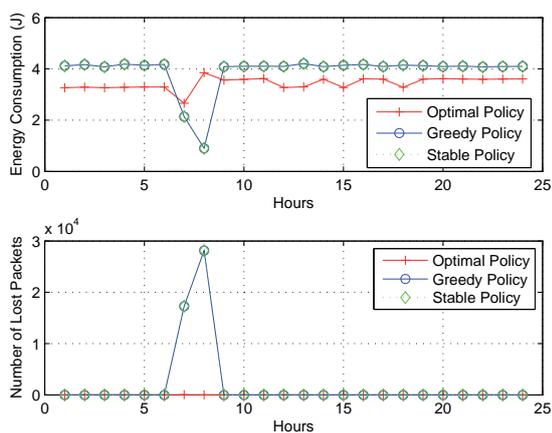


Fig. 9. Performance comparison among the optimal policy, the greedy policy, and the stable policy (the area of energy harvesting panel is  $2 \text{ cm}^2$ ,  $\mu = 1$ ,  $\eta = 0.97$ ,  $C_1 = 108 \text{ J}$ ,  $Q_c = 15$ ).

## 6. Conclusions

In this paper, we have studied the joint energy scheduling and transmission control problem for wireless communication with energy harvesting devices. We first model the solar energy harvesting process as a non-homogenous Markov chain, which can accurately characterize the harvesting dynamic in contrast to existing energy harvesting profiles. We consider the different timescale issue between the energy harvesting process and the transmission process, and propose a multiple timescale MDP framework to cast the joint problem. The optimal solution is derived via a modified dynamic programming and value iteration algorithms. Compared to existing energy management approaches, our proposed scheme has less packet loss and total energy consumption.

## Acknowledgements

This work is supported by the Fundamental Research Funds for the Central Universities under Grant ZYGX2009J006 and the National Natural Science Fund (No. 41101317).

## References

- [1] SURIYACHAI, U., SCOTT, A. A survey of mac protocols for mission-critical applications in wireless sensor networks. *IEEE Commun. Surveys and Tutorials*, 2011, vol. PP, no. 99, p. 1-25.
- [2] SZYMANSKI, T., GILBERT, D. Provisioning mission-critical telerobotic control systems over internet backbone networks with essentially-perfect QoS. *IEEE J. Select. Areas Commun.*, 2010, vol. 28, no. 5, p. 630-643.
- [3] SHIANG, H., VAN-DER-SCHAAR, M. Online learning in autonomic multi-hop wireless networks for transmitting mission-critical applications. *IEEE J. Select. Areas Commun.*, 2010, vol. 28, no. 5, p. 728-741.
- [4] DHAINI, A., HO, P. Mc-fiwiban: an emergency-aware mission-critical fiber-wireless broadband access network. *IEEE Commun. Magazine*, 2011, vol. 49, no. 1, p. 134-142.
- [5] OKORAFOR, U., KUNDUR, D. Security-aware routing and localization for a directional mission critical network. *IEEE J. Select. Areas Commun.*, 2010, vol. 28, no. 5, p. 664-676.
- [6] NIYATO, M., HOSSAIN, E., BHARGAVA, V. Wireless sensor networks with energy harvesting technologies: a game-theoretic approach to optimal energy management. *IEEE Tran. Wireless Commun.*, 2007, vol. 14, no. 4, p. 90-96.
- [7] SHARMA, V., MUKHERJI, U., JOSEPH, V., GUPTA, S. Optimal energy management policies for energy harvesting sensor nodes. *IEEE Trans. Wireless Comm.*, 2010, vol. 6, no. 4, p. 1326-1336.
- [8] TUTUNCUOGLU, K., YENER, A. Optimal power control for energy harvesting transmitters in an interference channel. In *Proc. of ASILOMAR*. CA(USA), 2011, p.378-382.
- [9] TUTUNCUOGLU, K., YENER, A. Optimum transmission policies for battery limited energy harvesting nodes. *IEEE Trans. Wireless Comm.*, 2012, vol. 11, no. 3, p. 1180-1189.
- [10] YANG, J., ULUKUS, S. Optimal packet scheduling in an energy harvesting communication system. *IEEE Trans. Comm.*, 2012, vol. 60, no. 1, p. 220-230.
- [11] OZEL, O., YANG, J., ULUKUS, S. Optimal broadcast scheduling for an energy harvesting rechargeable transmitter with a finite capacity battery. *IEEE Trans. Wireless Comm.*, to be appeared.
- [12] OZEL, O., TUTUNCUOGLU, K., YANG, J., ULUKUS, S., YENER, A. Resource management for fading wireless channels with energy harvesting nodes In *Proc. IEEE INFOCOM' 11*. Shanghai (China), 2011.
- [13] GORLATOVA, A., ZUSSMAN, G. Networking ultra low power energy harvesting devices: Measurements and algorithms. In *Proc. IEEE INFOCOM' 11*. Shanghai (China), 2011.
- [14] KANSAL, A., HSU, J., ZAHEDI, S., SRIVASTAVA, M. Power management in energy harvesting sensor networks. *ACM Trans. Embed. Comput. Syst.*, 2007, vol. 6.
- [15] REDDY, S., MURTHY, C. Profile-based load scheduling in wireless energy harvesting sensors for data rate maximization. In *Proc. of IEEE ICC 2010*. Cape Town (South Africa), 2010, p. 1-5.

- [16] SHENOY, V., MURTHY, C. Throughput maximization of delay-constrained traffic in wireless energy harvesting sensors. In *Proc. of IEEE ICC 2010*. Cape Town (South Africa), 2010, p. 1-5.
- [17] CHANG, S., FARD, P. Multitime scale Markov decision processes. *IEEE Trans. Automatic Control*, 2003, vol. 48, no. 6, p. 976-987.
- [18] HUGHES, J., GUTTORP, P., CHARLES, S. A non-homogeneous hidden Markov model for precipitation occurrence. *Journal of the Royal Statistical Society: Series C*, 1999, vol. 48, no. 1, p. 15-30.
- [19] LAMBERT, M., WHITING, J., METCALFE, A. A non-parametric hidden Markov model for climate state identification. *Hydrology and Earth System Sciences*, 2003, vol. 7, no. 5, p. 652-667.
- [20] BELLONE, E., HUGHES, J., GUTTORP, P. A hidden Markov model for downscaling synoptic atmospheric patterns to precipitation amounts. *Climate Research*, 2000, vol. 15, no. 1, p. 1-12.
- [21] ZHANG, Q., KASSAM, S. Finite-state Markov model for Rayleigh fading channels. *IEEE Trans. Commun.*, 1999, vol. 47, no. 11, p. 1688-1692.
- [22] PUTERMAN, M. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. Wiley, 1994.
- [23] BERTSEKAS, D. *Dynamic Programming and Optimal Control*. Athena Scientific, 2007.

### About Authors ...

**Hua XIAO** was born in Sichuan, China. He received his M.Sc. from University of Electronic Science and Technology of China (UESTC) in 2008. He is now pursuing PhD degree at the same place. His research interests include wireless communication and optimization.

**Huaizong SHAO** was born in Sichuan, China. He received his PhD degree from UESTC. He is now an associate professor at UESTC. His research interests include signal processing and wireless communication.