Data-Driven Stochastic Scheduling for Solar-Powered Sensor Communications

Meng-Lin Ku⋆, Yan Chen†, and K. J. Ray Liu*

Department of Communication Engineering, National Central University, Taiwan⋆
Department of Electrical and Computer Engineering, University of Maryland, College Park, USA† *
E-mail: mlku@ce.ncu.edu.tw⋆, yan@umd.edu†, kjrliu@umd.edu*

***Abstract*—This paper presents a data-driven approach of finding optimal scheduling policies for a solar-powered sensor node that attempts to maximize net bit rates by adapting its transmission to the changes of channel fading and battery recharge. The problem is formulated as a discounted Markov decision process (MDP) framework, whereby the energy harvesting process is stochastically quantized into several representative solar states with distinct energy arrivals and is totally driven by historical data records at a sensor node. We evaluate the average net bit rate of the optimal transmission scheduling policy, and computer simulations show that the proposed policy significantly outperforms other schemes with or without the knowledge of short-term energy harvesting and channel fading patterns.**

I. Introduction

Recently, energy harvesting has become an attractive alternative to overcome energy exhaustion problems in wireless sensor communications by scavenging ambient energy sources to replenish the power supply [1]. Though energy supplies from the environments enable sensor nodes to function for an infinite lifetime, management of the harvested energy remains a crucial issue due to the uncertainty of battery replenishment. In fact, ambient sources occur randomly and sporadically in nature, and this leads to various design considerations for energy management, since overly aggressive or conservative use of the harvested energy may either run out of the energy in the battery or fail to utilize the excess energy.

Energy generation models play an important role in the implementation of an energy harvesting system. In general, they can be categorized into two classes: deterministic models and stochastic models. For deterministic models, energy arrival instants and amounts are assumed to be known in advance by the transmitter [2]. Recently attention has shifted to stochastic models by accommodating energy management to the randomness of energy renewal processes. Commonly used models include Bernoulli process [3], Poisson process [4], exponential process [5], and two-state Markov models [6], etc. However, there has been little research to validate the assumptions, along with exact physical interpretation, of the aforementioned stochastic models. The mismatch problem between the theoretical models and the real-world data can severely degrade the performance of energy harvesting nodes.

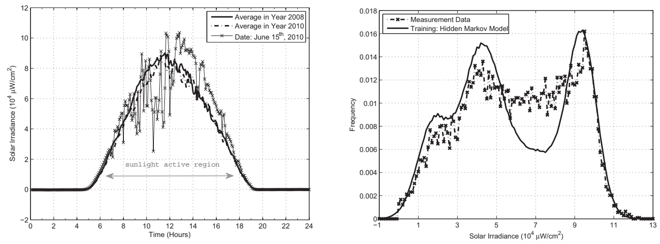
Resource management has been investigated in the literature to harmonize the energy consumption with the battery recharge rate. With deterministic energy profiles, a utility maximization framework was investigated in [7] to achieve smooth energy spending. With stochastic models, the authors of [5] designed a threshold to decide the transmission timing based on the importance of a data message. Some works resorted to a Markov decision process (MDP) for maximizing long-term utilities in energy harvesting systems [8]. However, none of these works linked the solar irradiance data, gathered by an energy harvesting node, to the constructions of the design frameworks and the optimal transmission policies.

In this paper, we present a data-driven stochastic scheduling policy for point-to-point energy harvesting sensor communications. We maximize the long-term net bit rates of the communication link by adapting transmission to the source's knowledge of its current battery and channel status. A data-driven stochastic model is proposed by quantifying energy harvesting conditions into several representative solar states, whereby the underlying parameters enable us to capture the dynamics of the harvested energy through real data records. The transmission scheduling problem is then formulated as a discounted MDP and solved by a value iteration algorithm. Since the exact solar state is unknown to the sensor, the belief state information is computed to decide whether to transmit or keep silent based on the present irradiance measurement. Finally, we compare the proposed policy with other existing schemes by computer simulations.

II. Stochastic Energy Harvesting Models

We begin with an example to justify the rationality of the proposed energy harvesting models. Consider a real data record of irradiance (in units μW/cm^2) for the month of June from 2008 to 2010, measured by a solar site in Elizabeth City State University, with the measurements taken at five-minute intervals [9]. In Fig. 1(a), the time series of the irradiance is sketched over twenty-four hours for June 15^{th}, 2010, along with the average results for the month of June in 2008 and 2010. The evolution of the diurnal irradiance follows a very similar time-symmetric mask, whereas the short-term profiles of different days can be very different and unpredictable. By considering the irradiance from seven o'clock to seventeen o'clock for June in 2008, 2009 and 2010, Fig. 1(b) shows the histogram plotted against the irradiance on the x-axis, which represents the percentage of the occurrences of data samples in each bin of width 10^3 μW/cm^2. It is seen that the irradiance

978-1-4799-7088-9/14/$31.00 ©2014 IEEE



(a) (b)

Fig. 1. Toy examples of solar irradiance measured by a solar site in Elizabeth City State University. (a) Time series of the daily irradiance in June. (b) Histogram of the irradiance during a time period of seven o'clock to seventeen o'clock for the month of June from 2008 to 2010.

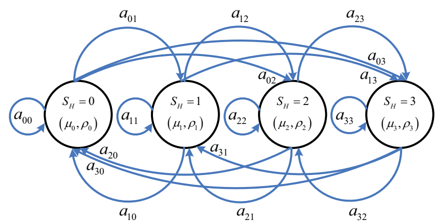


Fig. 2. Gaussian mixture hidden Markov chain of the solar power harvesting model with the underlying parameters (μ_j, ρ_j) $(N_H = 4)$.

behaves like a mixture random variable generated by a number of normal distributions.

The aforementioned observations motivate us to describe the evolution of the irradiance via an N_H-state solar power harvesting hidden Markov model in Fig. 2, where the underlying normal distribution for the j^{th} state is specified by the parameters of the mean μ_j and the variance ρ_j. The solar irradiance can be classified into several states S_H to represent energy harvesting conditions such as 'Excellent', 'Good', 'Fair', 'Poor', etc. Let $S_H^{(t)}$ be the solar state at time instant t. We assume that the hidden Markov model is time homogeneous and governed by the state transition probability $P\left(S_H^{(t)} = j \middle| S_H^{(t-1)} = i\right) = a_{ij}$, for $i, j = 0, \ldots, N_H - 1$. Let $\mathbf{x} = \left\{X^{(1)} = x_1, \ldots, X^{(T)} = x_T\right\}$ be observed data over a measurement period of T, corresponding to hidden states $\mathbf{s} = \left\{S_H^{(1)} = s_1, \ldots, S_H^{(T)} = s_T\right\}$. The parameters of the model are thus defined as $\boldsymbol{\Theta} = \{\boldsymbol{\mu}, \boldsymbol{\rho}, \mathbf{a}\}$, where $\boldsymbol{\mu} = \left[\mu_0, \ldots, \mu_{(N_H-1)}\right]^T$, $\boldsymbol{\rho} = \left[\rho_0, \ldots, \rho_{(N_H-1)}\right]^T$, and $\mathbf{a} = \left[a_{00}, a_{01}, \ldots, a_{(N_H-1)(N_H-1)}\right]^T$. The probabilistic model can be trained by an expectation-maximization (EM) algorithm, which is a general method of finding the maximum-likelihood (ML) estimate for the state parameters of underlying distributions from incomplete observed data, as follows [10]:

$$\begin{aligned}\boldsymbol{\Theta}^{(n)} &= \arg\max_{\boldsymbol{\Theta}} \mathbb{E}_{\mathbf{s}} \left[\log P\left(\mathbf{x}, \mathbf{s} \middle| \boldsymbol{\Theta}\right) \middle| \mathbf{x}, \boldsymbol{\Theta}^{(n-1)}\right] \\ &= \arg\max_{\boldsymbol{\Theta}} \sum\nolimits_{\mathbf{s}} \log P\left(\mathbf{x}, \mathbf{s} \middle| \boldsymbol{\Theta}\right) \cdot P\left(\mathbf{x}, \mathbf{s} \middle| \boldsymbol{\Theta}^{(n-1)}\right),\end{aligned} \tag{1}$$

where $\boldsymbol{\Theta}^{(n)}$ is the estimation update at the n^{th} iteration. The estimation can be efficiently carried out using well-known forward and backward procedures. The reader can refer to [10] for details. The training results with respect to the example above are shown in Table I and Fig. 1(b), where the irradiance measurements are performed from seven o'clock to seventeen o'clock. We can observe that the histograms of the training results and the measurement data in Fig. 1(b) behave quite similarly. Also in Table I, the transition probabilities from the current solar state to the other adjacent states are very small.

TABLE I
TRAINING RESULTS OF THE HIDDEN MARKOV SOLAR POWER HARVESTING MODEL.

(a) Mean, variance and steady state probability.

State $(S_H = i)$	0	1	2	3
μ_i (10^4 μW/cm^2)	1.75	4.21	7.02	9.38
ρ_i	0.65	1.04	2.34	0.54
$P(S_H = i)$	0.16	0.36	0.21	0.27

(b) State transition probability.

a_{ij}	$j = 0$	$j = 1$	$j = 2$	$j = 3$
$i = 0$	0.979	0.015	0.006	0
$i = 1$	0.005	0.988	0.007	0
$i = 2$	0.006	0.009	0.975	0.010
$i = 3$	0	0	0.007	0.993

The transmission strategy is usually designed on the basis of the required numbers of energy quanta during a management period T_L. Here, we map the solar power harvesting model into a discrete energy harvesting model. Let P_U be the basic transmission power level of sensor nodes, corresponding to one unit of the energy quantum $E_U = P_U T_L$ during the management period. In addition, for the harvested solar power P_H, the obtained energy over T_L is given by $E_H = P_H T_L$. The numbers of harvested energy quanta, Q, at $t = nT_L$ are given as

$$E_C^{(n)} = E_R^{(n-1)} + E_H\,; \tag{2}$$

$$E_R^{(n)} = E_C^{(n)} - QE_U,\ Q = \left\lfloor E_C^{(n)}/E_U \right\rfloor, \tag{3}$$

where $E_C^{(n)}$ and $E_R^{(n)}$ are the accumulated and the residual energy in the capacitor at $t = nT_L$, and $\lfloor\cdot\rfloor$ is the floor function. By assuming that the harvested power level is quasi-static over many power management runs, it can be analyzed that if $qE_U \le E_H \le (q+1)\,E_U$ for some q, then the probability of the number of energy quanta, Q, can be computed as

$$P\left(Q = i\right) = \begin{cases} \frac{E_H - qE_U}{E_U}, & i = q+1\,; \\ 1 - \frac{E_H - qE_U}{E_U}, & i = q\,; \\ 0, & \text{otherwise}\,. \end{cases} \tag{4}$$

When a sensor node is operated at the j^{th} solar state with the normal distribution $\mathcal{N}\left(x; \mu_j, \rho_j\right)$, the obtained energy E_H is again a normally distributed random variable, which is equal to the solar power per unit area x multiplied by the solar panel area Ω_S, the time duration T_L and the energy conversion efficiency ϑ, i.e., $E_H = x\Omega_S T_L \vartheta$. The conversion efficiency of an energy harvesting device typically ranges between 15% and 20% [1]. Thus, the mean and variance of E_H are respectively given as $\bar{\mu}_j = \mu_j \Omega_S T_L \vartheta$ and $\bar{\rho}_j = \rho_j \Omega_S^2 T_L^2 \vartheta^2$, and the probability of the number of energy quanta is calculated by using (4), as follows (The detailed derivation is skipped due to the limited space.):

$$P\left(Q=i|\,S_H=j\right) = \begin{cases} \int_{iE_U}^{(i+1)E_U} \frac{(i+1)E_U-E_H}{E_U}\mathcal{N}\left(E_H;\bar{\mu}_j,\bar{\rho}_j\right)dE_H, & i=0\,; \\ \int_{iE_U}^{(i+1)E_U} \frac{(i+1)E_U-E_H}{E_U}\mathcal{N}\left(E_H;\bar{\mu}_j,\bar{\rho}_j\right)dE_H \\ +\int_{(i-1)E_U}^{iE_U} \frac{E_H-(i-1)E_U}{E_U}\mathcal{N}\left(E_H;\bar{\mu}_j,\bar{\rho}_j\right)dE_H, & i\neq 0\,. \end{cases} \tag{5}$$

III. Data-Driven Stochastic Scheduling Policy Using Markov Decision Process

Consider a point-to-point communication link with two sensor nodes, where a source node intends to convey data packets to its sink node. Each data packet consists of L_S data symbols at a rate of R_S (symbols/sec), and the packet duration is given by $T_P = L_S/R_S$. The design framework is solved through the MDP, which is mainly composed of the state space, the action set, and the state transition probabilities. It is operated on the time scale T_L, covering D data packets, i.e., $T_L = DT_P$. Let $\mathcal{S}$ be the composite state space of the solar state $\mathcal{H} = \{0,\ldots,N_H-1\}$, the channel state $\mathcal{C} = \{0,\ldots,N_C-1\}$ and the battery state $\mathcal{B} = \{0,\ldots,N_B-1\}$, i.e., $\mathcal{S} = \mathcal{H}\times\mathcal{C}\times\mathcal{B}$, where $\times$ denotes the Cartesian product. We describe each component of the MDP in the following.

Scheduling Action: When the action $w \in \{0,1\}$ is chosen by the sensor node, the transmission power is set as wP_U during the management period T_L. If $w = 0$, the node remains silent without transmitting data packets; otherwise, it consumes one energy quantum for data transmission.

Channel State: The wireless channel is quantized using a finite number of thresholds $\mathbf{\Gamma} = \{0 = \Gamma_0, \Gamma_1, \ldots, \Gamma_{N_C} = \infty\}$, where $\Gamma_i < \Gamma_j$ for all $i < j$. The Rayleigh fading channel is said to be in the i^{th} channel state, if the instantaneous channel power, γ, belongs to the interval $[\Gamma_i, \Gamma_{i+1})$. The stationary probability and the transition probability for the finite-state Markov channel can be determined according to [11].

Battery State: When the sensor node is run at the n^{th} battery state, the available energy in the battery is stored up to n energy quanta. The state transition probability for the n^{th} battery state and the w^{th} scheduling action under the j^{th} solar state can be constructed by exploiting (5), as follows:

$$P_w\left(S_B=k|\left(S_H,S_B\right)=(j,n)\right) = \begin{cases} P\left(Q=k-n+w|\,S_H=j\right), \\ \qquad\qquad k=n-w,\ldots,N_B-2\,; \\ 1-\sum_{i=0}^{N_B-2-n+w} P\left(Q=i|\,S_H=j\right), & k=N_B-1\,, \end{cases} \tag{6}$$

for $n = 0,\ldots,N_B-1$ and $w \in \mathcal{W} \triangleq \{0, \min\{n,1\}\}$.

MDP State Transition: Since the transition of the channel and battery states are independent of each other, the transition probability from the state $(S_H,S_C,S_B) = (j,i,n)$ to the state $(S_H,S_C,S_B) = (j',i',n')$ with respect to the action $W = w$ can be formulated as

$$\begin{aligned} &P_w\left(\left(S_H,S_C,S_B\right)=(j',i',n')|\left(S_H,S_C,S_B\right)=(j,i,n)\right) \\ &= P\left(S_H=j'|\,S_H=j\right)P\left(S_C=i'|\,S_C=i\right) \\ &\qquad \cdot P_w\left(S_B=n'|\left(S_H,S_B\right)=(j,n)\right)\,, \end{aligned} \tag{7}$$

where the transition probability of the solar states can be directly obtained from the training results in (1).

Reward Function: We adopt the average net bit rate as our reward function. Let $P_{e,b}\left((S_C,S_B)=(i,n)\right)$ be the average bit error rate (BER) at the i^{th} channel state and the n^{th} battery state when the action $w=1$ is taken. By using the upper bound of the Q-function $Q(x) \leq \frac{1}{2}\exp\left(-\frac{x^2}{2}\right)$, it is computed as

$$\begin{aligned} P_{e,b}\left((S_C,S_B)=(i,n)\right) &= \frac{\int_{\Gamma_i}^{\Gamma_{i+1}} \alpha Q\left(\sqrt{\frac{\beta P_U\gamma}{N_0}}\right)\frac{1}{\gamma_0}e^{\left(-\frac{\gamma}{\gamma_0}\right)}d\gamma}{\int_{\Gamma_i}^{\Gamma_{i+1}}\frac{1}{\gamma_0}e^{\left(-\frac{\gamma}{\gamma_0}\right)}d\gamma} \\ &\leq \frac{\frac{\alpha}{\beta\gamma_U+2}}{e^{\left(-\frac{\Gamma_i}{\gamma_0}\right)}-e^{\left(-\frac{\Gamma_{i+1}}{\gamma_0}\right)}} \\ &\quad \cdot\left(e^{\left(-\frac{(\beta\gamma_U+2)\Gamma_i}{2\gamma_0}\right)}-e^{\left(-\frac{(\beta\gamma_U+2)\Gamma_{i+1}}{2\gamma_0}\right)}\right) \triangleq \eta(i,n)\,, \end{aligned} \tag{8}$$

where α and β are modulation specific constants for a 2^{χ}-ary modulation scheme, N_0 is the noise power, and $\gamma_U = P_U\gamma_0/N_0$ is the average signal-to-noise power ratio (SNR) with respect to the basic transmission power level. Hence, the probability of successful packet transmission (i.e., all χL_S bits in a packet are successfully detected) is expressed as

$$P_{f,k}\left((S_C,S_B)=(i,n)\right) = \left(1-P_{e,b}(i,n)\right)^{\chi L_S}\,. \tag{9}$$

The retransmission mechanism is employed if the sensor fails to decode the received data packet. Let Z be the total number of retransmissions required to successfully convey a data packet. The number of effective data packets during each management period is in average given as [12]

$$D_E = \frac{D}{\mathbb{E}[Z]} = \frac{T_L}{\mathbb{E}[Z]\,T_P} = \frac{T_L P_{f,k}(i,n)}{T_P}\,. \tag{10}$$

From (8)-(10), the net bit rate is lower bounded by

$$\begin{aligned} G_w\left((S_C,S_B)=(i,n)\right) &= \frac{1}{T_L}D_E\chi L_S \\ &\geq \frac{1}{T_P}\chi L_S\left(1-\eta(i,n)\right)^{\chi L_S}\,. \end{aligned} \tag{11}$$

The reward function can therefore be defined as

$$R_w\left((S_C,S_B)=(i,n)\right) = \begin{cases} 0, & w=0\,; \\ \frac{1}{T_P}\chi L_S\left(1-\eta(i,n)\right)^{\chi L_S}, & \text{otherwise}\,. \end{cases} \tag{12}$$

IV. Optimization of Scheduling Policies

The main goal of the MDP is to find a decision policy $\pi(s) : \mathcal{S} \rightarrow \mathcal{W}$ that specifies the optimal action in the state s and maximizes the objective function. The expected discounted infinite-horizon reward is formulated by using (12):

$$V_\pi(s_0) = \mathbb{E}_\pi\left[\sum_{k=0}^{\infty}\lambda^k R_{\pi(s_k)}(s_k)\right], \quad s_k\in\mathcal{S}, \quad \pi(s_k)\in\mathcal{W}\,, \tag{13}$$

where $V_\pi(s_0)$ is the expected reward starting from the initial state s_0 and continuing with the policy π from then on, and

$0 \le \lambda \le 1$ is a discount factor. It is known that the optimal value of the expected reward is unrelated to the initial state if the states of the Markov chain are assumed to be recurrent. The well-known value iteration approach can be applied to iteratively find the optimal scheduling policy:

$$V_{i+1}^{w}(s) = R_w(s) + \lambda \sum\nolimits_{s' \in \mathcal{S}} P_w(s'|\, s)\, V_i(s'), \qquad s \in \mathcal{S},\ \ w \in \mathcal{W}; \quad (14)$$

$$V_{i+1}(s) = \max_{w \in \mathcal{W}} \left\{ V_{i+1}^{w}(s) \right\},\ \ s \in \mathcal{S}, \quad (15)$$

where i is the iteration index, and the initial value of $V_0(s)$ is set as zero for all $s \in \mathcal{S}$. The update rule is repeated for several iterations until a stop criterion is satisfied, i.e., $|V_{i+1}(s) - V_i(s)| \le \varepsilon$.

In practice, the channel state can be reliably known at the transmitter via feedback channels, while the belief of the solar state can be estimated from the observation. Let $x^{(t)}$ be the average value of the measured solar data during the t^{th} management period, and $\zeta_j^{(t-1)} = P\left(S_H^{(t-1)} = j \middle|\, x^{(1)}, \ldots, x^{(t-1)} \right)$ be the belief of the j^{th} solar state according to the historical observation up to the $(t-1)^{th}$ management period. With the solar power harvesting model, the belief information at the t^{th} period can be updated via Bayes' rule as follows:

$$\zeta_j^{(t)} = \frac{\sum_{i=0}^{N_H-1} \zeta_i^{(t-1)} a_{ij} f_j\left(x^{(t)}\right)}{\sum_{j'=0}^{N_H-1} \sum_{i'=0}^{N_H-1} \zeta_{i'}^{(t-1)} a_{i'j'} f_{j'}\left(x^{(t)}\right)}, \quad (16)$$

where $f_j(x) \triangleq \mathcal{N}(x; \mu_j, \rho_j)$ is the normal distribution function. The final task is to apply the belief information $\zeta_j^{(t)}$ for deciding the scheduling action at each management period. Assuming that the current channel and battery states are known at each period, the action corresponding to the j^{th} solar state is chosen with probability proportional to $\zeta_j^{(t)}$.

V. Simulation Results

We evaluate the performance of the proposed data-driven scheduling policy by computer simulations, and set $N_H = 4$, $N_B = 12$, and $N_C = 6$. The data record of the irradiance collected by the solar site in Elizabeth City State University in June from 2008 to 2012 is adopted throughout the simulation [9]. A four-state solar power harvesting model is trained using the data in 2008, 2009 and 2010, where the underlying parameters are given in Table I. The irradiance data of the subsequent two years, 2011 and 2012, are then applied for performance evaluation. The stopping criterion ε is selected as 10^{-6}. Other simulation parameters are listed in Table II. A normalized SNR γ_C is defined with respect to the transmission power of 10^3 μW. A myopic policy is included for performance comparisons, and it attempts to transmit data packets at the basic transmission power level as long as the energy storage is non-empty. We also compare the proposed scheme with a deterministic scheme in [7], called t-time fair rate assignment (t-TFR), which requires perfect knowledge of the channel fading and energy harvesting patterns over a short-term period t for transmission scheduling.

TABLE II
SIMULATION PARAMETERS

Symbol rate (R_S)	100 kHz
Packet size (L_S)	10^3 symbols
Modulation type (α, β)	QPSK: $(1, 2)$ 16QAM: $(\frac{3}{4}, \frac{3}{15})$
Policy management duration (T_L)	300 sec
Basic action power (P_U)	40×10^3 μW
Solar panel area (Ω_S)	1 cm^2
Energy conversion efficiency (ϑ)	20%
Channel quantization levels ($\mathbf{\Gamma}$)	$\{0, 0.3, 0.6, 1.0, 2.0, 3.0, \infty\}$
Channel Model	Jakes' model
Normalized Doppler frequency (f_D)	0.05
Discount factor (λ)	0.99

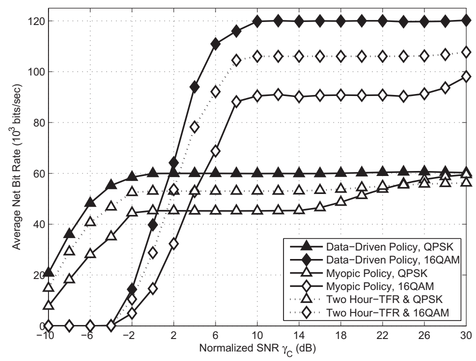


Fig. 3. Average net bit rates of the proposed and benchmark policies.

The average net bit rates of the proposed and benchmark scheduling policies are shown in Fig. 3. It is seen that the maximum net bit rates provided by our proposed policy are approximately given by 60×10^3 bits/sec and 120×10^3 bits/sec for QPSK and 16QAM, respectively. With a fixed modulation scheme, the proposed policy offers significant performance gains over the myopic policy by taking advantage of channel diversity gains. A closer look at this figure reveals that the performance gap between these two policies becomes wider as the modulation level increases. When compared with the Two Hour-TFR scheme, the proposed policy can still achieve better performance, no matter which modulation type is used.

VI. Conclusions

In this paper, we have studied the problem of maximizing long-term net bit rates in sensor communications that solely rely on solar energy for data transmission. Based on a developed node-specific energy harvesting model, a data-driven MDP framework was formulated to obtain the optimal transmission scheduling policy in response to the dynamics of channel fading and battery storage. The proposed data-driven approach was rigorously justified by the real data of solar irradiance. The proposed optimal scheduling policy was shown to achieve significant gains over other radical approaches, while it did not require non-causal knowledge of energy harvesting and channel fading patterns.

References

[1] S. Sudevalayam and P. Kulkarni, "Energy harvesting sensor nodes: survey and implications, " *IEEE Commun. Surveys Tutorials*, vol. 13, no. 3, pp. 443-461, Third Quad. 2011.

[2] M. Tacca, P. Monti, and A. Fumagalli, "Cooperative and reliable ARQ protocols for energy harvesting wireless sensor nodes, " *IEEE Trans. Wireless Commun.*, vol. 6, no. 7, pp. 2519-2529, July 2007.

[3] B. Medepally, N. B. Mehta, and C. R. Murthy, "Implications of energy profile and storage on energy harvesting sensor link performance, " *Proc. IEEE Glob. Commun. Conf.*, pp. 1-6, 2009.

[4] K. J. Prabuchandran, S. K. Meena, and S. Bhatnagar, "Q-learning based energy management policies for a single sensor node with finite buffer, " *IEEE Wireless Commun. Lett.*, vol. 2, no. 1, pp. 82-85, Feb. 2013.

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

[6] S. Mao, M. H. Cheung, and V. W. S. Wong, "An optimal energy allocation algorithm for energy harvesting wireless sensor networks, " *Proc. IEEE Int. Conf. Commun.*, pp. 265-270, 2012.

[7] M. Gorlatova, A. Wallwater, and G. Zussman, "Networking low-power energy harvesting devices: measurements and algorithms, " *IEEE Trans. Mobile Comput.*, vol. 12, no. 9, pp. 1853-1865, Sept. 2013.

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

[9] N. R. E. Laboratory. (2012, Feb.) Solar radiation resource information. [Online]. Available: http://www.nrel.gov/rredc/.

[10] J. A. Bilmes, "A gentle tutorial of the EM algorithm and its application to parameter estimation for Gaussian mixture and hidden Markov models, " International Computer Science Institute, Tech. Rep. ICSI-TR-97-021, Apr. 1998.

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

[12] C. Pandana and K. J. R. Liu, "Near optimal reinforcement learning framework for energy-aware wireless sensor communications, " *IEEE J. Sel. Areas Commun.*, vol. 23, no. 4, pp. 788-797, Apr. 2005.