Lyapunov Optimization for Energy Harvesting Wireless Sensor Communications

Chengrun Qiu, Yang Hu, Yan Chen, *Senior Member, IEEE*, and Bing Zeng, *Fellow, IEEE*

***Abstract*—With the development and popularity of the renewable energy harvesting devices, the energy harvesting wireless sensor communications that can make use of the energy harvested from the nearby environments have gained more and more attentions. One key problem in the energy harvesting wireless sensor communications is the transmission strategy management, i.e., how to manage the transmission strategy at each time slot to optimize the transmission performance. In this paper, we propose to use Lyapunov optimization theory to maximize the expected good bits per packet transmission for the source node in an energy harvesting wireless communication system. Considering the channel and battery states, we adapt the transmission power and modulation type to achieve such a goal. The problem is formulated as an optimization where the objective function is the long-term average good bits per packet transmission and the constraints are the bounded long-term average battery level and bit error rate. To solve the optimization, we introduce virtual queues and employ the Lyapunov optimization theory to transform the optimization with long-term average format into optimizing the drift-plus-penalty problem. The drift-plus-penalty is further upper bounded with variables only related to current time slot, which greatly simplifies the optimization problem. Theoretic analysis is also conducted to show that the optimal solution is limited by an upper bound that is independent of the operation time index. Finally, simulation results with real solar irradiance data show that the proposed algorithm can achieve much better performance than existing approaches based on Markov decision process and water-filling.**

***Index Terms*—Continuous energy, drift-plus-penalty, energy harvesting, energy management, Lyapunov optimization, online algorithm.**

I. Introduction

ENERGY harvesting sensor communication techniques have recently gained much attention, and the available renewable energy used to power up the wireless sensor communications can be collected from a large scale of energy sources, including solar/light, thermoelectric, mechanical motion/vibration, electromagnetic radiation, etc. [1]. With the renewable battery replacing the nonrechargeable batteries, the harvested energy can be utilized to charge batteries periodically, saving the time of battery replacement. There have been three typical energy harvesting architectures [1], [2]. One is harvest-use (HU) architecture, in which case the system consumes the harvested energy directly. The other two are harvest-store-use architecture and HU-store architecture, where the system can use part of the harvested energy and store the remaining for future use. The management of the harvested energy is challenging due to the uncertainty of the coming harvested energy. How to efficiently and effectively utilize the harvested renewable energy is the key problem in energy harvesting sensor communications.

Various objective functions have been used in the energy scheduling algorithms literature. The most common one is to maximize the throughput of the energy harvesting sensor communication system [3]. In [4], the throughput of the communication system with a one-bit channel feedback is maximized by searching for a suitable MQAM modulation. Rezaee *et al.* [5] considered the single-user throughput maximization of an energy harvesting system with continuous energy and data arrivals. A three-step optimal energy scheduling algorithm was derived with the knowledge of the harvested energy and data arrival. Blasco *et al.* [6] utilized the Markov decision process (MDP) to maximize the long-term expected throughput to derive the optimal power level. Specifically, the state space is discretized and the transition probabilities among different states are derived, with which the long-term expected throughput can be described with a Bellman equation and value iteration can be used to obtain the optimal solution [7], [8]. However, the computational complexity of the MDP-based approaches is generally high due to the large volume of the state and action space.

Another commonly adopted method to maximize the throughput is the water-filling technique where the KKT optimality conditions are used to obtain the optimal power or energy. Ozel *et al.* [9] proposed a directional water-filling algorithm to maximize the number of bits sent in a finite duration and minimize the time taken to sent a given amount of data in an energy harvesting system with fading channels. Ho and Zhang [10] considered the problem of energy allocation over a finite horizon by taking into account the time-varying channel conditions and energy sources to maximize the throughput. It was shown that when the battery is with unlimited capacity and the full information of the channel conditions and harvested energy is known, the water-filling energy allocation solution is optimal. Wang *et al.* [11] proposed

Manuscript received December 24, 2017; accepted March 16, 2018. Date of publication March 20, 2018; date of current version June 8, 2018. This work was supported in part by the National Natural Science Foundation of China under Grant 61602090 and Grant 61672137 and in part by the 111 Project under Grant B17008, and the Thousand Youth Talents Program of China (to Yan Chen). *(Corresponding author: Yang Hu.)*

The authors are with the School of Information and Communication Engineering and the Center of Future Media, University of Electronic Science and Technology of China, Chengdu 611731, China (e-mail: cr_qiu@std.uestc.edu.cn; yanghu@uestc.edu.cn; eecyan@uestc.edu.cn; eezeng@uestc.edu.cn).

Digital Object Identifier 10.1109/JIOT.2018.2817590

an iterative dynamic water-filling algorithm for the optimal energy scheduling for fading multiple access channels with energy harvesting. Such an iterative algorithm was shown to be converged in a few iterations with much better sum-rate compared with various suboptimal energy scheduling algorithms. While achieving promising results, the water filling algorithm is limited only to the optimization with finite time slots and cannot be used for the problem with long-term optimization. Other objective functions have also been investigated [12]–[16]. In [12], the average number of good bits per packet is optimized by managing the power and modulation for the energy harvesting sensor communications, while the total energy cost of a wireless communication system powered by a combination of harvested energy and conventional energy is minimized in [17]. The battery lifetime of an energy harvesting wireless communication system is maximized in [18] with the constraint of the time average reward larger than a certain threshold.

To resolve the optimization with the long-term objective function and/or constraints with continuous variables, Lyapunov optimization has been utilized [19]. Since Lyapunov optimization does not rely on the state transition, its computational complexity is generally much smaller than the MDP-based approaches. He *et al.* [20] utilized the Lyapunov optimization to optimize the time average expected transmission cost for the mobile video streaming system. It was shown that such a method can well balance the data wastage cost and the freezing time of the mobile user. Wu *et al.* [21] proposed to maximize the received video quality of an energy harvesting video communication system with the Lyapunov optimization theory, where the received video quality takes into account the playback interruption and fluctuation. Cui *et al.* [22] studied the delay-aware resource control problem in wireless systems with Lyapunov optimization, where the system throughput, the sum delay, and the power consumption are jointly optimized in one problem with Lyapunov optimization.

In this paper, we propose to use Lyapunov optimization theory to maximize the good bits per packet transmission for the source node in an energy harvesting wireless communication system. We achieve such an objective function through adapting the transmission power and modulation type while considering the channel and battery state. Specifically, we optimize the long-term average good bits per packet transmission while satisfying the long-term average battery and bit error rate (BER) constraints. Note that different from the MDP-based algorithms such as [12], the energy management is continuous other than discrete, and different from the water-filling-based algorithms such as [9], we consider a long-term objective function and thus cannot be solved with offline algorithms. To solve the optimization, we introduce virtual queues and employ the Lyapunov optimization theory to transform the optimization with long-term average format into optimizing the drift-plus-penalty problem. Furthermore, we propose an online algorithm by finding an upper bound of the drift-plus-penalty with all the terms that is only related to current time slot, which greatly simplifies the optimization problem. We also conduct the theoretic analysis to show that the optimal solution is limited by a upper bound that is independent of the operation time index. Finally, we conduct multiple simulations to evaluate the proposed algorithm, and the simulation results with real solar irradiance data show that the proposed algorithm can achieve much better performance than the MDP-based approach in [12] and the water-filling approach in [9]. The energy in our model is continuous thus brings higher requirement of the components in the system.

TABLE I
BRIEF SUMMARY OF MAJOR NOTATIONS

T_L : Policy management duration	B : Battery virtual queues
E_H: Energy harvested during T_L	P : BER virtual queues
ω : Transmission energy during T_L	Ir : Solar irradiance
T_P : Packet duration	R : Reward function
L_S : Number of symbols during T_P	P_e: Expected bit error rate
χ_m: Information bits per data symbol	Θ : Penalty drift
L : Lyapunov function	V : Parameter of penalty weight
θ : Battery constraint	p : Transmission power during T_L
δ : BER constraint	m : Modulation type
γ_0 : Average channel power	n_0: Noise power
γ_C : Average channel-to-noise ratio	$\mathcal{M}$: Modulation set
γ_I : Average signal-to-noise ratio	φ : Chernoff bound of Q-function

The rest of this paper is organized as follows. In Section II, we describe in detail the model of the energy harvesting wireless communication system, including the components of the solar irradiance, the battery, the channel power and the reward function. Then, we discuss the problem formulation in Section III, where two constraints for battery energy and BER as well as the objective function are introduced. In Section IV, we show how to utilize the Lyapunov optimization theory to solve the proposed optimization problem. Simulations about the evolution of the reward function, the performance comparison, and the parametric sensitivity analysis are illustrated in Section V. Finally, the conclusion is drawn in Section VI.

A summary of major notations used in this paper is listed in Table I.

II. SYSTEM MODEL

Harvesting and transmitting are two main components of a stochastic solar energy harvesting communication system as shown in Fig. 1. In the harvesting part, the energy is harvested from the environment and kept in a battery for the future potential transmission. In the transmitting part, the transmitter decides the transmission power according to certain criteria and consumes the energy in the battery for wireless transmission. The information is transmitted through wireless channel, which is assumed to be Rayleigh fading, and received by the receiver. The time is divided into slots, and the channel is assumed to keep unchanged within one time slot. At the beginning of each time slot, the transmitter will determine the optimal transmission power and modulation to optimize the objective function which will be discussed in detail later.

A. Harvestable Solar Irradiance

The solar energy harvested in each time slot is influenced by many factors such as weather, sunshine duration, harvesting efficiency, etc. In this paper, we utilize the real data of solar energy harvested in a time slot observed in Elizabeth City State University in September from 2010 to 2012 [23]. Fig. 2 illustrates the evaluation of the solar irradiance in a particular

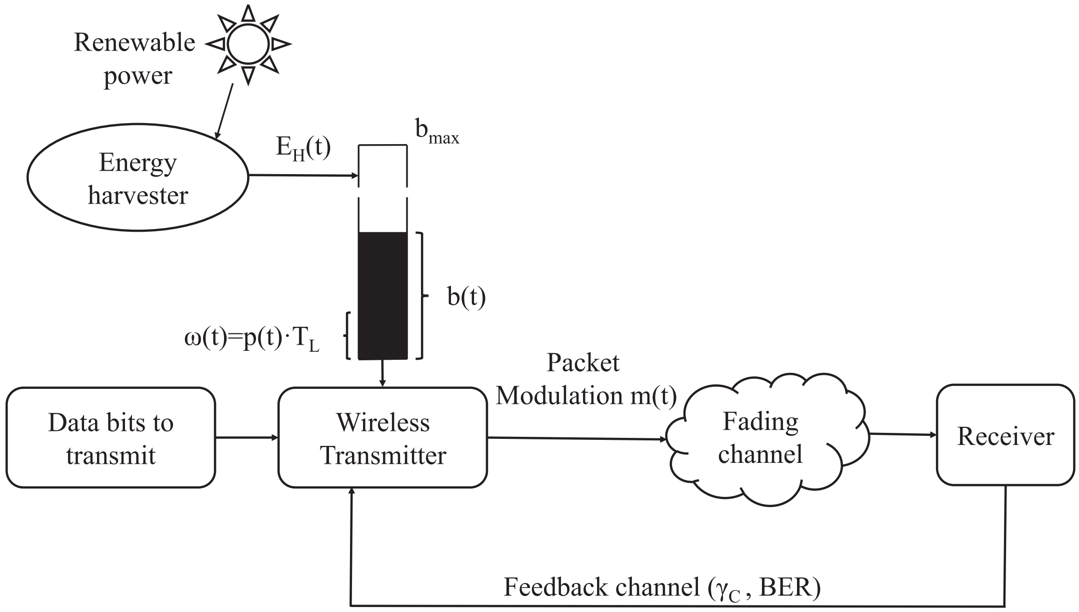


Fig. 1. Illustration of the system model.

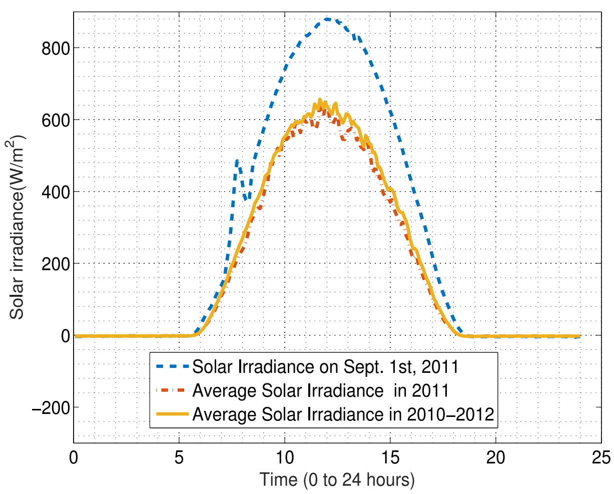


Fig. 2. Average solar irradiance observed in Elizabeth City State University in September from 2010 to 2012 [23].

day, the average solar irradiance in 2011 and the average solar irradiance in 2010–2012. We can see from the figure that the average solar irradiance is un-trivial from 6:00 to 18:00, which can be seen as the active time period. Based on the observed data, we only study the solar energy harvesting system during the sunlight active period.

B. Wireless Channel and Rechargeable Battery Model

We assume that the wireless channel keeps unchanged within each time slot and varies independently from time slot to time slot according to the Rayleigh fading channel model as

$$f(\gamma) = \frac{1}{\gamma_0} \exp\left(\frac{-\gamma}{\gamma_0}\right) \tag{1}$$

where γ is the instantaneous channel power, $f(\gamma)$ represents the probability that the channel power equals to γ, and $\gamma_0 = \mathbb{E}(\gamma)$ is the expected channel power. The channel power is assumed to be i.i.d at all time slots.

In this paper, we assume that the transmitter finishes an energy harvesting period at the end of each time slot, and consumes some energy for transmission at the beginning of the coming time slot. The energy harvesting and consuming are both sustained process where $E_H(t) = Ir(t) \cdot T_L$ and $\omega(t) = p(t) \cdot T_L$ with $Ir(t)$ and $\omega(t)$ being the average harvested energy density and the transmission power density in a time slot, respectively. In such a case, the $E_H(t)$ cannot be used immediately at the time slot t. Since the battery has a capacity limit where the overflow energy would be discarded, the energy amount of the battery in the time slot $(t+1)$ can be written as follows:

$$b(t+1) = \min\{b(t) - \omega(t) + E_H(t), b_{\max}\} \tag{2}$$

where $b_{\max}$ is the capacity of the battery.

C. Reward Function

The reward function is defined to be the expected good bits per packet transmission. The channel condition is assumed to be fed back periodically by feedback channel with no loss for the preparation of the next transmission. The expected BER [12] is defined as

$$\begin{aligned} P_e(\gamma, b, \omega, m) &= \sum_r \alpha(m,r) \cdot Q\left(\sqrt{\frac{\beta(m,r) \cdot p \cdot \gamma}{n_0}}\right) \\ &= \sum_r \alpha(m,r) \cdot Q\left(\sqrt{\frac{\beta(m,r) \cdot \omega \cdot \gamma}{n_0 T_L}}\right) \end{aligned} \tag{3}$$

where n_0 is the noise power, $[(\omega \cdot \gamma)/n_0]$ stands for the input signal-to-noise ratio, (γ/n_0) stands for the channel-to-noise ratio, and $\alpha(m,r)$ and $\beta(m,r)$ are two parameters related to the modulation type shown in Table II [12].

TABLE II
MODULATION SPECIFIC CONSTANTS

Modulation schemes	Parameters$(\alpha(m,r),\beta(m,r))$
QPSK	$(\alpha(m,0),\beta(m,0))=(1,1)$
8PSK	$(\alpha(m,0),\beta(m,0))=(\frac{2}{3},2sin^2(\frac{\pi}{8}))$ $(\alpha(m,1),\beta(m,1))=(\frac{2}{3},2sin^2(\frac{3\pi}{8}))$
16QAM	$(\alpha(m,0),\beta(m,0))=(\frac{3}{4},\frac{1}{5})$ $(\alpha(m,1),\beta(m,1))=(\frac{1}{2},\frac{9}{5})$

Based on the average BER in (4), the reward function can be written as follows [12]:

$$R(\gamma,b,\omega,m)=\begin{cases}\frac{\chi_m L_S}{T_p}(1-P_e(\gamma,b,\omega,m))^{\chi_m L_S}, & \omega\neq 0\\ 0, & \omega=0\end{cases} \tag{4}$$

where χ_m is the information bits per data symbol, L_S is the number of symbols during a packet duration, and T_P is the packet duration.

From (4), we can see that the reward function $R(\omega,m)$ calculates the net bit rate, i.e., the information bits in the effective data packets per unit time in a management period. Reward function $R(\omega,m)$ equals to 0 when $\omega=0$. This is reasonable since the information bits at the receiver would be 0 if the transmitter decides not to transmit any packet. Given the transmission energy ω, the immediate reward $R(\omega,m)$ is independent of the battery $b(t)$ and an improved channel power γ can bring a higher immediate reward.

III. POWER AND MODULATION OPTIMIZATION PROBLEM FORMULATION

The solar energy harvesting sensor communications problem is to determine the optimal power and modulation at each time slot to optimize the transmission performance while satisfying certain constraints, which will be discussed in detail in this section.

A. Battery Constraint for Energy Harvesting Sensor Communications

In practice, the channel can be in deep fading due to the stochastic characteristic. In such a situation, the transmission performance can be very bad, and thus the transmitter would store the energy in the battery waiting for a better channel. Moreover, to maintain proper operations of the transmitter, there should be always some power stored in the battery. Therefore, a reasonable constraint for the battery is that the average energy amount in the battery should be not less than a certain threshold as follows:

$$\lim_{T\to+\infty}\frac{1}{T}\sum_{t=0}^{T-1}b(t)\geq\theta \tag{5}$$

where θ is a predefined energy level threshold. A higher θ means more long-term average available energy in the battery. Because the transmission energy is regulated as $0\leq\omega(t)\leq b(t)$, higher θ means stronger instantaneous transmission ability. Generally, the transmitter tends to spend more energy for transmission since larger power means better transmission performance. With constraint (5), in order to consume more energy under good channel condition, the transmitter will store energy when the channel is bad. In other words, such a constraint can better allocate the energy at different channel conditions and thus lead to better transmission performance.

B. BER Constraint for Energy Harvesting Sensor Communications

In addition to the average energy constraint in the battery, another constraint we need to consider is the BER constraint, which is directly related to the transmission performance. In a real communication system, a few times of high BER can be tolerated. However, if the high BER situation turns to be very frequent, the entire transmission performance would be poor and thus unacceptable. Therefore, a reasonable constraint for the BER is that the average BER should be not greater than a certain threshold as follows:

$$\lim_{T\to+\infty}\frac{1}{T}\sum_{t=0}^{T-1}P_e(t)\leq\lim_{T\to+\infty}\frac{1}{T}\sum_{t=0}^{T-1}\varphi(\omega(t),m(t))\ \leq\ \delta \tag{6}$$

where δ is a predefined threshold, $P_e(t)$ is the BER at time slot t defined in (3), and $P_e(t)\leq\varphi(\omega(t),m(t))$, where $\varphi(\omega(t),m(t))=\sum_r\alpha(m,r)\exp(-[(\omega(t)\cdot\gamma\cdot\beta(m,r))/2n_0])$ is the Chernoff bound of the Q-function.

C. Problem Formulation

The objective of the transmitter is to determine the optimal transmission power and modulation to maximize the time average net bit rate while satisfying the battery constraint in (5) and the BER constraint in (6), i.e., the optimization problem can be formulated as follows:

$$\begin{aligned}\max_{\omega(t),m(t)}\quad & \lim_{T\to+\infty}\frac{1}{T}\sum_{t=0}^{T-1}R(t)\\ \text{s.t.}\quad & \lim_{T\to+\infty}\frac{1}{T}\sum_{t=0}^{T-1}b(t)\geq\theta\\ & \lim_{T\to+\infty}\frac{1}{T}\sum_{t=0}^{T-1}\varphi(t)\leq\delta\\ & 0\leq b(t)\leq b_{\max}\\ & 0\leq\omega(t)\leq b(t)\\ & m(t)\in\mathcal{M}\end{aligned} \tag{7}$$

where $R(t)$ is the net bit rate at each time slot t defined in (4) and $\mathcal{M}=\{\text{QPSK, 8PSK, 16QAM}\}$ represents the stochastic available modulation set at every time slot.

IV. TRANSMISSION QUALITY OPTIMIZATION STRATEGIES

The optimization problem in (7) involves the long-term averaging in both objective function and constraints, which cannot be directly solved by the traditional optimization techniques. To solve such an optimization problem, we propose to use the Lyapunov optimization theory, which will be discussed in detail in the rest of this section.

A. Lyapunov Optimization Formulation

To use Lyapunov optimization theory, we first transform the long-term averaging constraints in (7) into virtual queues. Let us define two virtual queues $B(t)$ and $P(t)$ for battery constraint and BER constraint, respectively, as follows:

$$B(t+1) = \max\{B(t) + \theta - b(t+1), 0\} \tag{8}$$

and

$$P(t+1) = \max\{P(t) + \varphi(t) - \delta, 0\}. \tag{9}$$

The virtual queues can be seen as the signs to judge how well the constraints are satisfied in the previous slots. With a larger $B(t)$, the battery constraint is sacrificed for the continuous previous transmission.

Lemma 1: If the virtual queues $B(t)$ and $P(t)$ are rate stable, i.e., $\lim_{T\to+\infty}[B(T)/T] \le 0$ and $\lim_{T\to+\infty}[P(T)/T] \le 0$, then we have $\lim_{T\to+\infty}(1/T)\sum_{t=0}^{T-1} b(t) \ge \theta$ and $\lim_{T\to+\infty}(1/T)\sum_{t=0}^{T-1} \varphi(t) \le \delta$.

Proof: We first rewrite (8) as follows:

$$B(t+1) = \begin{cases} B(t) + \theta - b(t+1), & \text{if } B(t) \ge b(t+1) - \theta \\ 0, & \text{if } B(t) < b(t+1) - \theta. \end{cases} \tag{10}$$

Then, we have

$$\begin{aligned} B(t+1) - B(t) &= \begin{cases} \theta - b(t+1), & \text{if } B(t) \ge b(t+1) - \theta \\ -B(t), & \text{if } B(t) < b(t+1) - \theta \end{cases} \\ &= \max\{\theta - b(t+1), -B(t)\} \\ &\ge \theta - b(t+1). \end{aligned} \tag{11}$$

By taking the average from time index 0 to time index $T-1$ on both sides of (11) and letting T goes to infinite, we have

$$\begin{aligned} \lim_{T\to+\infty} \frac{B(T)}{T} &\ge \theta - \lim_{T\to+\infty} \frac{1}{T}\sum_{t=1}^{T} b(t) \\ &= \theta - \lim_{T\to+\infty} \frac{1}{T}\sum_{t=0}^{T-1} b(t). \end{aligned} \tag{12}$$

With (12), if the virtual queue $B(t)$ is rate stable, i.e., $\lim_{T\to+\infty}(B(T)/T) \le 0$, we have

$$\lim_{T\to+\infty} \frac{1}{T}\sum_{t=0}^{T-1} b(t) \ge \theta. \tag{13}$$

Similarly, if the virtual queue $P(t)$ is rate stable, i.e., $\lim_{T\to+\infty}(P(T)/T) \le 0$, we can also show that

$$\lim_{T\to+\infty} \frac{1}{T}\sum_{t=0}^{T-1} \varphi(t) \le \delta. \tag{14}$$

This completes the proof. ■

According to Lemma 1, the optimization problem in (7) can be transformed into the following optimization problem by enhancing the constraints:

$$\begin{aligned} \max_{\omega(t),m(t)} \quad & \lim_{T\to+\infty} \frac{1}{T}\sum_{t=0}^{T-1} R(t) \\ \text{s.t.} \quad & \lim_{t\to+\infty} \frac{B(t)}{t} \le 0 \\ & \lim_{t\to+\infty} \frac{P(t)}{t} \le 0 \\ & 0 \le b(t) \le b_{\max} \\ & 0 \le \omega(t) \le b(t) \\ & m(t) \in \mathcal{M}. \end{aligned} \tag{15}$$

Then, we can utilize the penalty drift in Lyapunov optimization theory to solve (15). Let us first define a concatenated vector of the virtual queues as $\Theta(t) = [B(t), P(t)]$, and the corresponding Lyapunov function can be written as

$$L(\Theta(t)) = \frac{1}{2}\Big(B^2(t) + P^2(t)\Big). \tag{16}$$

With (16), the Lyapunov penalty drift can be obtained as [19]

$$\Delta\Theta(t) = \mathbb{E}\{L(\Theta(t+1)) - L(\Theta(t))|\Theta(t)\}. \tag{17}$$

According to the Lyapunov optimization theory, enforcing the constraints is equivalent to minimize the drift penalty $\Delta\Theta(t)$, and optimizing the objective function with the constraints is equivalent to optimize the "drift-plus-penalty" defined as follows:

$$\Delta\Theta(t) - V\mathbb{E}\{R(\omega(t), m(t))|\Theta(t)\} \tag{18}$$

where the parameter $V \ge 0$ stands for the penalty weight, which represents the importance of the objective function compared to the constraints at every time slot.

Therefore, the optimization problem in (15) can be rewritten as follows:

$$\begin{aligned} \min_{\omega(t),m(t)} \quad & \Delta\Theta(t) - V\mathbb{E}\{R(t)|\Theta(t)\} \\ \text{s.t.} \quad & 0 \le b(t) \le b_{\max} \\ & 0 \le \omega(t) \le b(t) \\ & m(t) \in \mathcal{M}. \end{aligned} \tag{19}$$

According to (16) and (17), we can see that the objective function in (19) involves the variables at time slot $t+1$, which needs to be eliminated for finding the solution to the optimization problem in (19). To do so, we first find the upper bound of $B^2(t+1) - B^2(t)$ as follows:

$$\begin{aligned} &B^2(t+1) - B^2(t) \\ &\quad = \{\max\{B(t) + \theta - b(t+1), 0\}\}^2 - B^2(t) \\ &\quad \le [B(t) + \theta - b(t+1)]^2 - B^2(t) \\ &\quad \le \theta^2 + b^2(t+1) + 2B(t)[\theta - b(t+1)] \\ &\quad = \theta^2 + \{\min\{\max\{b(t) - \omega(t), 0\} + E_H(t), b_{\max}\}\}^2 \\ &\qquad + 2B(t)[\theta - b(t+1)] \\ &\quad \le \theta^2 + \{\min\{\max\{b(t) - \omega(t), 0\} + E_H(t), b_{\max}\}\}^2 \\ &\qquad + 2B(t)[\theta - b(t) + \omega(t)] \end{aligned}$$

$$
\begin{aligned}
&\leq \theta^2 + \{\max\{b(t) - \omega(t), 0\} + E_H(t)\}^2 \\
&\quad + 2B(t)[\theta - b(t) + \omega(t)] \\
&= \theta^2 + \{\max\{b(t) - \omega(t), 0\}\}^2 + E_H(t)^2 + 2E_H(t) \\
&\quad \times \max\{b(t) - \omega(t), 0\} + 2B(t)[\theta - b(t) + \omega(t)] \\
&\leq \theta^2 + [b(t) - \omega(t)]^2 + E_H(t)^2 + 2E_H(t) \\
&\quad \times b(t) + 2B(t)[\theta - b(t) + \omega(t)] \\
&\leq \theta^2 + E_{Hmax}^2 + b_{\max}^2 + \omega(t)^2 - 2b(t)\omega(t) \\
&\quad + 2b_{\max} \cdot E_{Hmax} + 2B(t)[\theta - b(t) + \omega(t)] \qquad (20)
\end{aligned}
$$

where the first inequality is due to $\{\max\{a, 0\}\}^2 \leq a^2$, the third inequality is due to $b(t+1) = \min\{\max\{b(t) - \omega(t), 0\} + E_H(t), b_{\max}\}$ and $\max\{b(t) - \omega(t), 0\} + E_H(t) \geq b(t) - \omega(t)$, and the fourth inequality is due to $\{\min\{a, b\}\}^2 \leq a^2$.

Similarly, the upper bound of $P^2(t+1) - P^2(t)$ can be derived as follows:

$$
\begin{aligned}
P^2(t+1) - P^2(t) &= \{\max\{P(t) + \varphi(t) - \delta, 0\}\}^2 - P^2(t) \\
&\leq [P(t) + \varphi(t) - \delta]^2 - P^2(t) \\
&\leq \delta^2 + \varphi^2(t) + 2P(t)[\varphi(t) - \delta] \qquad (21)
\end{aligned}
$$

where the first inequality is due to $\{\max\{a, 0\}\}^2 \leq a^2$, and the last inequality is due to $\varphi^2(t) \leq 1$.

With (20) and (21), we can derive the upper bound of the drift-plus-penalty as follows:

$$
\begin{aligned}
&\Delta\Theta(t) - VR(\omega(t), m(t)) \\
&\leq \frac{1}{2} \cdot \Big\{\theta^2 + E_{H\max}^2 + b_{\max}^2 + \omega^2(t) - 2b(t)\omega(t) \\
&\quad + 2b_{\max}E_{H\max} + 2B(t) \cdot [\theta - b(t) + \omega(t)] + \delta^2 + 1 \\
&\quad + 2P(t) \cdot [\varphi(\omega(t), m(t)) - \delta] \\
&\quad - VR(\omega(t), m(t))\Big\} \\
&= \frac{1}{2} \cdot \Big\{B_1 + B_2 + \omega(t)^2 - 2b(t)\omega(t) + 2B(t)\omega(t) \\
&\quad + 2P(t) \cdot [\varphi(\omega(t), m(t)) - \delta]\Big\} \\
&\quad - VR(\omega(t), m(t)) \qquad (22)
\end{aligned}
$$

where $B_1 = \theta^2 + E_{H\max}^2 + b_{\max}^2 + \delta^2 + 1 + 2 \cdot b_{\max} \cdot E_{H\max}$ and $B_2 = 2B(t) \cdot (\theta - b(t)) - 2P(t) \cdot \delta$ are two fixed parameters at time slot t.

From (22), we can see that the upper bound is only partly related to the variable at time slot t, i.e., $\omega(t)$ and $m(t)$. Therefore, by relaxing the drift-plus-penalty with the corresponding upper bound, the optimization problem becomes

$$
\begin{aligned}
\min_{\omega(t), m(t)} \quad & J(\omega(t), m(t)) \\
& = \omega(t)^2 - 2b(t)\omega(t) + 2B(t)\omega(t) \\
& + 2P(t) \cdot \varphi(\omega(t), m(t)) \\
& - 2VR(\omega(t), m(t)) \\
\text{s.t.} \quad & 0 \leq b(t) \leq b_{\max} \\
& 0 \leq \omega(t) \leq b(t) \\
& m(t) \in \mathcal{M}. \qquad (23)
\end{aligned}
$$

From (23), we can see that the optimization problem is nonconvex, and thus the convex optimization techniques cannot be directly applied. While some convex relaxation techniques may be used, in this paper we simply use the full search method by discretizing the remaining energy with a fixed step size and finding the pair of transmission energy and modulation that can achieve the optimal performance. The maximal computational complexity of such a full search algorithm is $\mathcal{O}([(b_{\max}/(\nu))] \times M)$, where ν is the step size and M is the number of possible modulation.

B. Performance Analysis

There exists a lower bound for the optimal average reward derived in (23) as shown in the following theorem.

Theorem 1: The optimal average reward function obtained by (23) is limited by a lower bound that is independent of the operation time

$$
\lim_{T \to +\infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}\{R(\omega(t), m(t))\} \geq R^{\text{opt}} - \frac{B_1 + 2b_{\max}B_{\max}}{2V} \qquad (24)
$$

where $B_{\max}$ is the upper threshold of the virtual queue B.

Proof: By substituting the expression of B_2, (22) can be rewritten as

$$
\begin{aligned}
&\Delta\Theta(t) - V\mathbb{E}\{R(\omega(t), m(t))|\Theta(t)\} \\
&\quad \leq \frac{1}{2}\Big\{B_1 + 2B_{\max}\mathbb{E}\{(\theta - b(t))|\Theta(t)\} + \omega(t)^2 \\
&\qquad - 2b(t)\omega(t) + 2B_{\max}\omega(t) \\
&\qquad + 2P(t) \cdot \mathbb{E}\{[\varphi(\omega(t), m(t)) - \delta]|\Theta(t)\}\Big\} \\
&\quad - V\mathbb{E}\{R(\omega(t), m(t))|\Theta(t)\}. \qquad (25)
\end{aligned}
$$

Since "$\theta - b(t)$" and "$\varphi(\omega(t), m(t)) - \delta$" are two factors of inequality constraints in (8), once the constraints are satisfied and the distribution of γ is i.i.d. (25) can be simplified by applying ω-only policy as follows [19]:

$$
\begin{aligned}
&\Delta\Theta(t) - V\mathbb{E}\{R(\omega(t), m(t))|\Theta(t)\} \\
&\quad \leq \frac{1}{2}\Big\{B_1 + \omega(t)^2 - 2b(t)\omega(t) + 2B(t)\omega(t)\Big\} - VR^{\text{opt}}. \qquad (26)
\end{aligned}
$$

With some manipulations, (26) can be rearranged as follows:

$$
\begin{aligned}
&\mathbb{E}\{R(\omega(t), m(t))\} \\
&\quad \geq R^{\text{opt}} + \frac{2\Delta\Theta(t) - B_1 + \omega(t)[-\omega(t) + 2b(t) - 2B(t)]}{2V} \\
&\quad \geq \begin{cases} R^{\text{opt}} + \frac{2\Delta\Theta(t) - B_1 + b^2(t) - 2b(t)B(t)}{2V}, & B(t) \geq \frac{1}{2}b(t) \\ R^{\text{opt}} + \frac{2\Delta\Theta(t) - B_1}{2V}, & \text{others} \end{cases} \\
&\quad \geq R^{\text{opt}} + \frac{2\Delta\Theta(t) - B_1 - 2b_{\max}B_{\max}}{2V}. \qquad (27)
\end{aligned}
$$

By taking the average from time index 0 to time index $T - 1$ on both sides of (27) and letting T goes to infinite, we have

$$
\begin{aligned}
&\lim_{T \to +\infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}\{R(\omega(t), m(t))\} \\
&\quad \geq R^{\text{opt}} - \frac{B_1 + 2b_{\max}B_{\max}}{2V} + \lim_{T \to +\infty} \frac{1}{T} \sum_{t=0}^{T-1} \frac{\Delta\Theta(t)}{V} \\
&\quad = R^{\text{opt}} - \frac{B_1 + 2b_{\max}B_{\max}}{2V} + \lim_{T \to +\infty} \frac{L(\Theta(T)) - L(\Theta(0))}{VT} \\
&\quad = R^{\text{opt}} - \frac{B_1 + 2b_{\max}B_{\max}}{2V}. \qquad (28)
\end{aligned}
$$

This completes the proof. ∎

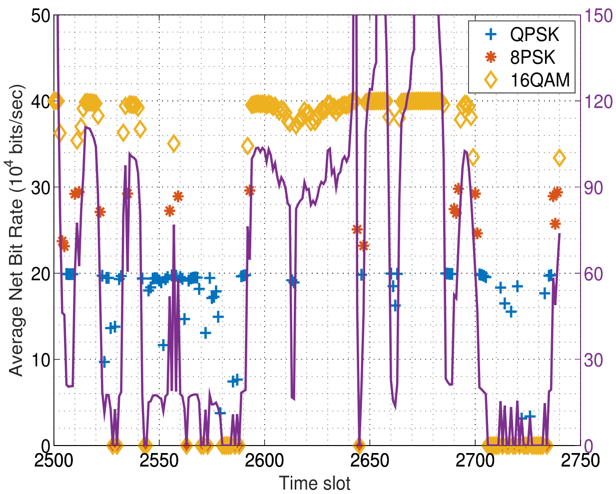


Fig. 3. Average net bit rate performance and modulation versus channel condition.

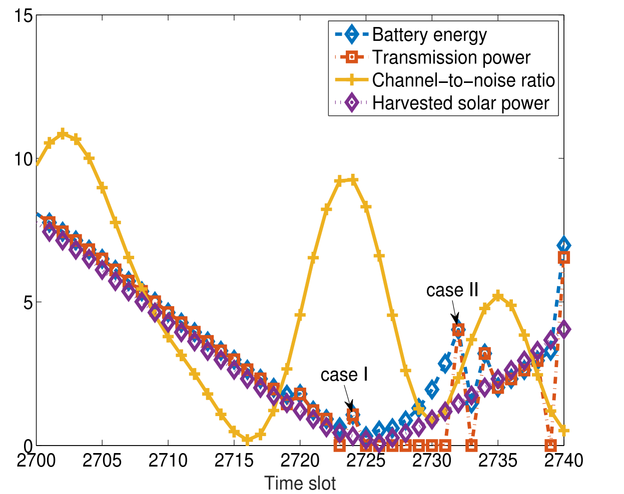


Fig. 4. Transmission power versus channel condition, battery state, and harvested energy.

V. Simulation Results

In this section, we conduct simulations to evaluate the performance of the proposed algorithm. The solar power collected in every 5 min in Elizabeth City in September from 2010 to 2012 are utilized in the simulations. The fading channel is produced automatically with the method of Jakes model [24]. The harvesting panel size is set to be 1 cm^2 and the energy conversion efficiency is assumed to be $\eta = 20\%$. The symbol rate R_S is set to be 100 kHz, and each packet is assumed to contain 1000 data symbols ($L_S = 1000$), i.e., the packet duration $T_P = 0.01$ s. The modulation types could be QPSK, 8PSK, and 16QAM and the policy management duration T_L is set as 5×60 s. The average reward function is used to characterize the performance of the algorithm.

A. Performance Evaluation

We first examine the system performance under different conditions such as channel conditions, battery conditions, etc. The parameters are set as $V = 10\,000$, $\delta = 0.01$, and $\theta = 40 * T_L(mJ)$. Intuitively, if the channel condition is bad and/or the energy in battery is low, the system tends to save energy to wait for better channel conditions or satisfying the battery threshold. On the other hand, if the BER turns too bad for a long time, the system tends to spare some energy to improve the QoS.

In Fig. 3, we illustrate the reward function value under different channel conditions at each time slot. To clearly see the details, we choose the data between the 2500th time slot and the 2800th time slot when the system runs into a stable state. The reward function value, i.e., the average net bit rate, varies from 0 to 4×10^5. The maximal value is $4L_S/T_P$, where 4 is the information bits per symbol when 16QAM is used. From Fig. 3, we can see that the value of the reward function reaches the maximal value only when the channel state is good enough. We also illustrate the modulation of the transmission at each time slot in Fig. 3. We can see that when the system chooses 16QAM, 8PSK, and QPSK, the reward function value varies between 3×10^5 and 4×10^5, 2×10^5, and 3×10^5, as well as 0 and 2×10^5, respectively. We also observe that when the channel condition is bad, medium, and good, the system chooses QPSK, 8PSK, and 16QAM, respectively. In other words, the channel condition determines the modulation to a large extent, and thus determines the average net bit rate. In Fig. 4, we can see that the transmitter tends to consume energy when either the battery energy is high or the channel condition is good. For example, at the 2724th time slot, the transmitter chooses to transmit for an excellent channel condition, while at the 2732nd time slot the transmitter chooses to transmit mainly because of high amount of energy in the battery instead of an excellent channel condition.

B. Performance Comparison

To validate the advantages of the proposed algorithm, we compare with four benchmarks: the first one is the greedy strategy where all harvested energy will be used to transmit at every time slot, the second one is the on–off policy with modulation selection based on MDP in [12], the third one is the water-filling algorithm in [9], and the fourth one is the online water-filling algorithm in [9].

The parameters of our strategy are set as $V = 10\,000$ and $\delta = 0.01$. The on–off strategy in [12] considers the state $S = \{S_H \times S_C \times S_B\}$, where S_H is the state of solar irradiance, S_C is the state of channel, and S_B is the state of battery. Then the transition probabilities $P(s' = \{i' \times j' \times k'\}|s = \{i \times j \times k\})$ is derived, and the expected reward function $V_{\pi^*}(s) = \max_{a \in A}(R_a(s) + \lambda \sum_{s' \in S} P_a(s'|s)V_{\pi^*}(s'))$ is calculated based on the present state and their transition probabilities. The optimal strategy is the strategy that maximizes the expected reward function. Similar to [12], the number of the states of solar, channel and battery is $\{N_H \times N_C \times N_B\} = \{4 \times 6 \times 12\}$, and the channel interval is set as $\{0, 0.3, 0.6, 1.0, 2.0, 3.0, \infty\}$. The water-filling algorithm in [9] is an offline algorithm, which cannot be directly used in our problem. Here we run the water-filling algorithm with $T = 10$ time slots under the ideal assumption that all the

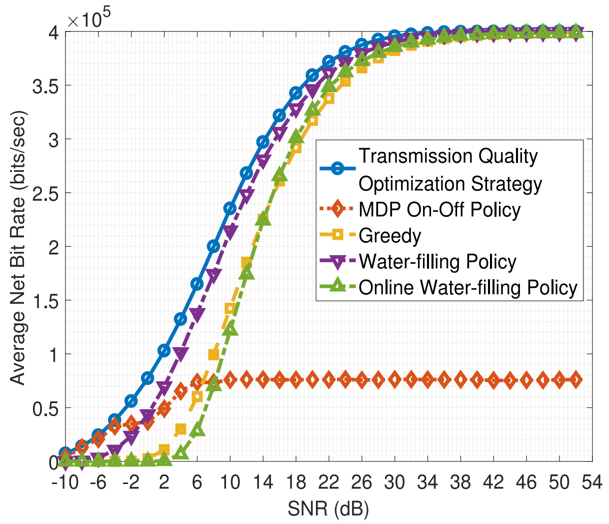


Fig. 5. Performance comparison at different SNRs (QPSK, 8PSK, and 16QAM).

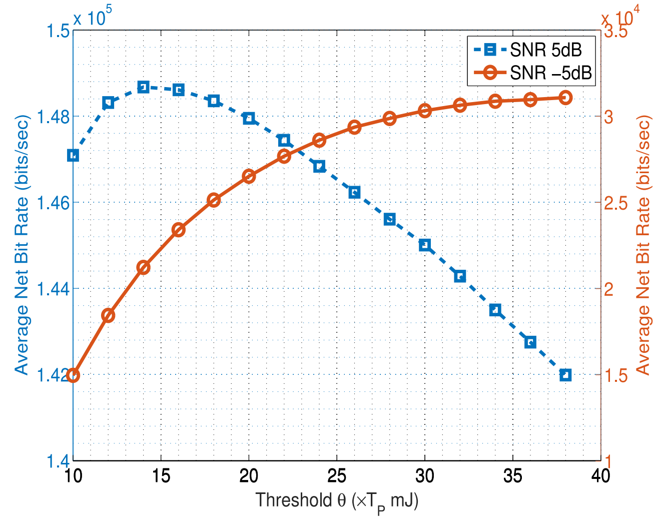


Fig. 6. Performance sensitivity to θ.

information in these ten time slots are known. Note that such an assumption is impractical in reality.

Fig. 5 shows the performance comparison in term of the average net bit rate between the proposed algorithm and the four benchmarks at different SNRs. From Fig. 5, we can see that the proposed strategy achieves better performance than all benchmarks for all SNRs varying from -10 to 54 dB. We also observe that with the proposed strategy, the average net bit rate increases when the SNR increases from -10 to 28 dB and finally saturates at 4×10^5 bits/s, which is $[(4L_S)/(T_P)]$. With the greedy strategy, the average net bit rate increases only when the SNR is larger than -2 dB and the performance finally arrives at the same value as that of the proposed strategy. With the on–off policy [12], the average net bit rate increases when the SNR varies from -10dB to 10dB while the performance finally converges at about 0.75×10^5 bits/s. The on–off policy has a lower converged performance since the upper bound of the on–off policy is $R_{\text{net},m} \leq \min\{\bar{q}, 1\}\cdot([\chi_m L_S/T_P]\cdot(1-P_e(\gamma_{\max}, 1, m))^{\chi_m L_S})$ [12], where $\bar{q} = [\overline{E_H}/(T_P \cdot P_U)]$ is the energy quantum harvesting rate. The greedy algorithm can be achieved by setting the $\theta = 0$ and $\delta = 1$. With the offline water-filling algorithm, we find that its performance lies between our proposed algorithm and the greedy algorithm. The average net bit rate starts to increase when the SNR is larger than -6 dB, and finally saturates at 4×10^5 bits/s. With the online water-filling algorithm, we find that the average net bit rate starts to increase only if the SNR exceeds 4 dB, and finally saturates at 4×10^5. Compared with the greedy algorithm, its performance is worse when the SNR is smaller than 14 dB, and better when the SNR is larger than 14 dB. The main reason that the proposed algorithm performs better than the water-filling algorithm is that the water-filling algorithm cannot take infinite time slots into consideration. Such an offline algorithm can be somehow treated as one kind of greedy algorithms which only focuses on the reward in the period within the length T. Moreover, we would like to emphasize again that the offline algorithm needs more information than online algorithm, which is generally impractical.

When the SNR varies from -10 to -4 dB, the performance of proposed strategy is slightly better than that of the MDP on–off policy while the greedy strategy performs poorly. This is because when channel condition is very poor, accumulating energy and consuming it effectively can lead to a better performance than consuming all energy at every time slot. On the other hand, when the SNR is larger than 8 dB, the performance of the greedy strategy goes beyond that of the MDP on–off policy. Such a phenomenon is because that when the SNR is high, a small amount of energy for transmission can lead to a better performance improvement. In such a case, the saving energy strategy in MDP on–off policy may waste the good chances of transmission and thus leads to a lower performance upper bound.

C. Evaluation of the Parameter Sensitivity

In this section, we conduct simulations to study how the performance of the proposed strategy varies with the parameters (θ, δ, V). The parameters are set as $V = 10\,000, \theta = 40 \times T_P$, and $\delta = 0.01$ unless one of them is the variable.

1) Performance Sensitivity to θ: We conduct simulations to analyze how the reward function varies with θ changing from $10 \times T_P$ to $38 \times T_P$, and the results are shown in Fig. 6 where the SNR is set to -5 and 5 dB, respectively. We can see that when the SNR is -5 dB, as θ increases from $10 \times T_P$ to $32 \times T_P$, the average reward function increases rapidly. This is because the parts close to zero of the reward function is convex. When the channel condition is not good, concentrating energy for usage and sacrifice other time slots has a better performance than dividing the energy equally. In other words, sacrificing the continuous energy for stronger instantaneous transmission ability can get more fed back. However, if θ is too large, considering the penalty weight V, the system will turn to consume energy to improve transmission. There, the average reward function starts to saturate when θ is large. On the other hand, with a higher SNR (5 dB), we can see that starting from $\theta = 14 \times T_P$, the average reward function keeps decreasing. Such a phenomenon is mainly because that when

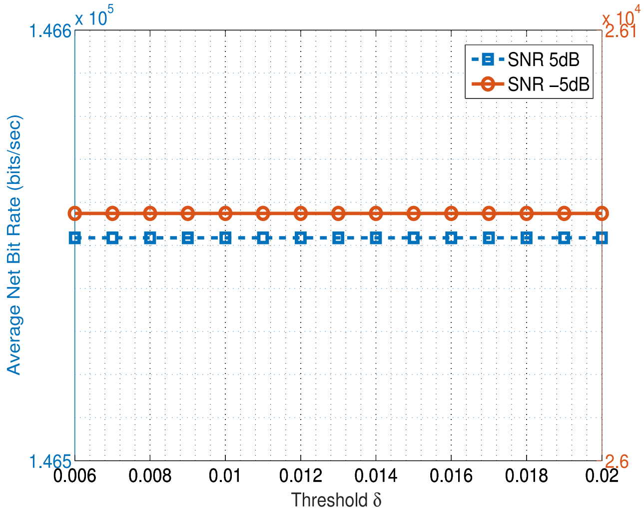


Fig. 7. Performance sensitivity to δ.

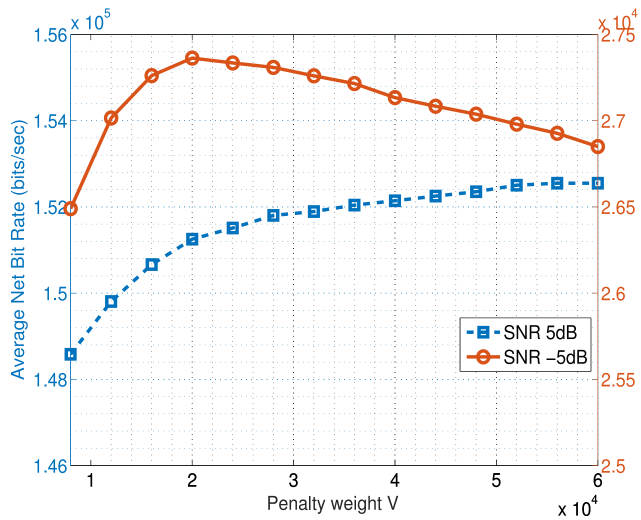


Fig. 8. Performance sensitivity to V.

the channel condition is good, saving energy may miss good transmission chances and thus lead to a lower performance.

2) Performance Sensitivity to δ: Fig. 7 illustrates the average reward function versus the δ which varies from 0.006 and 0.02. We can see that the average reward function almost keeps constant regardless of the change of the δ. This is mainly because that comparing with the battery virtual queue, the value of BER virtual queue is much smaller and thus has little effect on the entire policy.

3) Performance Sensitivity to V: We evaluate how the penalty weight V affects the average reward function under different channel conditions, and the results are shown in Fig. 8. We can see that when the SNR is -5 dB, the average reward function first increases as the V increases. However, when the V is larger than 2×10^4, the average reward function begins to decrease as the V increases. The reason is that when the SNR is -5 dB, either not considering the constraints or considering the constraints too much cannot bring the best performance. Instead, we need to find a balance by choosing a suitable V. When the SNR is 5 dB, the average reward function increases monotonically as the V varies from 8000 to 60 000. This is because when the channel condition is good, grabbing the present transmission chance rather than considering the constraints can achieve better performance.

VI. Conclusion

In this paper, we investigated the transmission strategy management problem in energy harvesting wireless communications. We formulated the strategy management problem as a constrained stochastic optimization problem with the objective to maximize the long-term expected good bits per packet transmission under long-term battery and BER constraints. Utilizing the Lyapunov optimization theory, we derived an optimal strategy to manage the transmission power and modulation based on the instantaneous channel information and battery state. The simulation results using real solar irradiance data demonstrate that the proposed algorithm achieves much better performance compared with the existing MDP-based approach and water-filling approach.

References

[1] M.-L. Ku, W. Li, Y. Chen, and K. J. R. Liu, “Advances in energy harvesting communications: Past, present, and future challenges,” *IEEE Commun. Surveys Tuts.*, vol. 18, no. 2, pp. 1384–1412, 2nd Quart., 2016.

[2] S. Sudevalayam and P. Kulkarniand, “Energy harvesting sensor nodes: Survey and implications,” *IEEE Commun. Surveys Tuts.*, vol. 13, no. 3, pp. 443–461, 3rd Quart., 2011.

[3] S. Ulukus *et al.*, “Energy harvesting wireless communications: A review of recent advances,” *IEEE J. Sel. Areas Commun.*, vol. 33, no. 3, pp. 360–381, Mar. 2015.

[4] R. Ma and W. Zhang, “Adaptive MQAM for energy harvesting wireless communications with 1-bit channel feedback,” *IEEE Trans. Wireless Commun.*, vol. 14, no. 11, pp. 6459–6470, Nov. 2015.

[5] M. Rezaee, M. Mirmohseni, and M. R. Aref, “Energy harvesting systems with continuous energy and data arrivals: The optimal offline and heuristic online algorithms,” *IEEE J. Sel. Areas Commun.*, vol. 34, no. 12, pp. 3739–3753, Dec. 2016.

[6] P. Blasco, D. Gunduz, and M. Dohler, “A learning theoretic approach to energy harvesting communication system optimization,” *IEEE Trans. Wireless Commun.*, vol. 12, no. 4, pp. 1872–1882, Apr. 2013.

[7] A. Munir and A. Gordon-Ross, “An MDP-based dynamic optimization methodology for wireless sensor networks,” *IEEE Trans. Parallel Distrib. Syst.*, vol. 23, no. 4, pp. 616–625, Apr. 2012.

[8] J. Lee and S. Bahk, “On the MDP-based cost minimization for video-on-demand services in a heterogeneous wireless network with multihomed terminals,” *IEEE Trans. Mobile Comput.*, vol. 12, no. 9, pp. 1737–1749, Sep. 2013.

[9] O. Ozel, K. Tutuncuoglu, J. Yang, S. Ulukus, and A. Yener, “Transmission with energy harvesting nodes in fading wireless channels: Optimal policies,” *IEEE J. Sel. Areas Commun.*, vol. 29, no. 8, pp. 1732–1743, Sep. 2011.

[10] C. K. Ho and R. Zhang, “Optimal energy allocation for wireless communications with energy harvesting constraints,” *IEEE Trans. Signal Process.*, vol. 60, no. 9, pp. 4808–4818, Sep. 2012.

[11] Z. Wang, V. Aggarwal, and X. D. Wang, “Iterative dynamic water-filling for fading multiple-access channels with energy harvesting,” *IEEE J. Sel. Areas Commun.*, vol. 33, no. 3, pp. 382–395, Mar. 2015.

[12] M.-L. Ku, Y. Chen, and K. J. R. Liu, “Data-driven stochastic models and policies for energy harvesting sensor communications,” *IEEE J. Sel. Areas Commun.*, vol. 33, no. 8, pp. 1505–1520, Aug. 2015.

[13] M.-L. Ku, W. Li, Y. Chen, and K. J. R. Liu, “On energy harvesting gain and diversity analysis in cooperative communications,” *IEEE J. Sel. Areas Commun.*, vol. 33, no. 12, pp. 2641–2657, Dec. 2015.

[14] W. Li, M.-L. Ku, Y. Chen, and K. J. R. Liu, “On outage probability for stochastic energy harvesting communications in fading channels,” *IEEE Signal Process. Lett.*, vol. 22, no. 11, pp. 1893–1897, Nov. 2015.

[15] W. Li, M. L. Ku, Y. Chen, and K. J. R. Liu, “On outage probability for two-way relay networks with stochastic energy harvesting,” *IEEE Trans. Commun.*, vol. 64, no. 5, pp. 1901–1915, May 2016.

[16] W. Li, M.-L. Ku, Y. Chen, K. J. R. Liu, and S. Zhu, "Performance analysis for two-way network-coded dual-relay networks with stochastic energy harvesting," *IEEE Trans. Wireless Commun.*, vol. 16, no. 9, pp. 5747–5761, Sep. 2017.

[17] X. Kang, Y.-K. Chia, C. K. Ho, and S. Sun, "Cost minimization for fading channels with energy harvesting and conventional energy," *IEEE Trans. Wireless Commun.*, vol. 13, no. 8, pp. 4586–4598, Aug. 2014.

[18] N. Michelusi, L. Badia, R. Carli, L. Corradini, and M. Zorzi, "Energy management policies for harvesting-based wireless sensor devices with battery degradation," *IEEE Trans. Commun.*, vol. 61, no. 12, pp. 4934–4947, Dec. 2013.

[19] M. J. Neely, "Stochastic network optimization with application to communication and queueing systems," *Synth. Lectures Commun. Netw.*, vol. 3, no. 1, pp. 1–211, 2010.

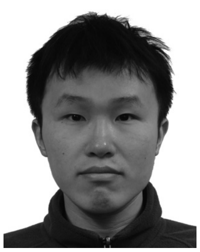
[20] J. He, Z. Xue, D. Wu, D. P. O. Wu, and Y. G. Wen, "CBM: Online strategies on cost-aware buffer management for mobile video streaming," *IEEE Trans. Multimedia*, vol. 16, no. 1, pp. 242–252, Jan. 2014.

[21] X. Wu, J. Yang, Y. Ran, and H. Xi, "Adaptive scalable video transmission strategy in energy harvesting communication system," *IEEE Trans. Multimedia*, vol. 17, no. 12, pp. 2345–2353, Dec. 2015.

[22] Y. Cui, V. K. N. Lau, R. Wang, H. Huang, and S. Zhang, "A survey on delay-aware resource control for wireless systems—Large deviation theory, stochastic Lyapunov drift, and distributed stochastic learning," *IEEE Trans. Inf. Theory*, vol. 58, no. 3, pp. 1677–1701, Mar. 2012.

[23] *Cooperative Networks for Renewable Resource Measurements (CONFRRM) Solar Energy Resource Data*, NREL, Elizabeth City, NC, USA, Sep. 2012. [Online]. Available: http://www.nrel.gov/rredc

[24] P. Dent, G. E. Bottomley, and T. Croft, "Jakes fading model revisited," *Electron. Lett.*, vol. 29, no. 13, pp. 1162–1163, Jun. 1993.

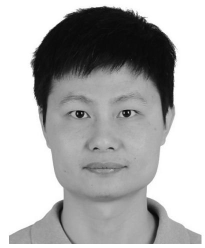


Chengrun Qiu is currently pursuing the B.S. degree in electronic information engineering from the University of Electronic Science and Technology of China, Chengdu, China.

His current research interests include reinforcement learning, stochastic optimization, and green communications.

Yang Hu received the B.S. and Ph.D. degrees in electrical engineering from the University of Science and Technology of China, Hefei, China, in 2004 and 2009, respectively.

She was a Research Associate with the Institute for Advanced Computer Studies, University of Maryland at College Park, College Park, MD, USA, from 2010 to 2015. She is currently an Associate Researcher with the School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu, China. Her current research interests include computer vision, machine learning, and multimedia signal processing.



Yan Chen (SM'14) received the bachelor's degree from the University of Science and Technology of China, Hefei, China, in 2004, the M.Phil. degree from the Hong Kong University of Science and Technology, Hong Kong, in 2007, and the Ph.D. degree from the University of Maryland at College Park, College Park, MD, USA, in 2011.

He was a Founding Principal Technologist with Origin Wireless Inc., Greenbelt, MD, USA. Since 2015, he has been a Full Professor with the School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu, China. His current research interests include multimedia, signal processing, game theory, and wireless communications.

Dr. Chen was a recipient of multiple honors and awards, including the Best Student Paper Award of the PCM in 2017, the Best Student Paper Award of IEEE ICASSP in 2016, the Best Paper Award of IEEE GLOBECOM in 2013, the Future Faculty Fellowship and Distinguished Dissertation Fellowship Honorable Mention from the Department of Electrical and Computer Engineering in 2010 and 2011, the Finalist of the Dean's Doctoral Research Award from the A. James Clark School of Engineering, University of Maryland at College Park, in 2011, and the Chinese Government Award for outstanding students abroad in 2010.

Bing Zeng (M'91–SM'13–F'16) received the B.Eng. and M.Eng. degrees in electronic engineering from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 1983 and 1986, respectively, and the Ph.D. degree in electrical engineering from the Tampere University of Technology, Tampere, Finland, in 1991.

He was a Post-Doctoral Fellow with the University of Toronto, Toronto, ON, Canada, from 1991 to 1992, and a Researcher with Concordia University, Montreal, QC, USA, from 1992 to 1993. He then joined the Hong Kong University of Science and Technology (HKUST), Hong Kong. After 20 years of service with HKUST, he returned to UESTC in 2013, through China's 1000-Talent-Scheme. At UESTC, he leads the Institute of Image Processing to research on image and video processing, 3-D and multiview video technology, and visual big data. During his tenure with HKUST and UESTC, he has graduated over 30 master's and Ph.D. students, received about 20 research grants, holds 8 international patents, and has authored or co-authored over 260 papers. Three representing works are as follows: one paper on fast block motion estimation, published in the IEEE Transactions on Circuits and Systems for Video Technology in 1994, has over SCI 1000 citations (over 2200 Google citations) and currently stands at the eighth position among all papers published in this transactions; one paper on smart padding for arbitrarily shaped image blocks, published in the IEEE Transactions on Circuits and Systems for Video Technology in 2001, led to a patent that has been successfully licensed to companies; and one paper on directional discrete cosine transform, published in the IEEE Transactions on Circuits and Systems for Video Technology in 2008.

Dr. Zeng was a recipient of the 2011 IEEE Transactions on Circuits and Systems for Video Technology Best Paper Award, the Best Paper Award of ChinaCom in 2009, 2010, and 2012, the Best Associate Editor Award in 2011, and a Second Class Natural Science Award (the first recipient) from the Chinese Ministry of Education in 2014. He served as an Associate Editor for the IEEE Transactions on Circuits and Systems for Video Technology for eight years. He was the General Co-Chair of VCIP 2016 and PCM 2017.