Wireless Personal Communications (2022) 122:3513–3527
https://doi.org/10.1007/s11277-021-09098-4

Online Power Control and Optimization for Energy Harvesting Communication System Based on State of Charge

Jing Guo[1] · **Xu Zhang**[2]

Accepted: 5 September 2021 / Published online: 15 September 2021

Abstract
In this paper, the online power control problem for energy harvesting wireless communication system with a finite storage capacity battery is addressed, where the channel state and energy harvesting rate are both unknown. A low complexity algorithm based online convex optimization is proposed to guarantee energy availability of energy harvesting node and maximize average long-term throughput. The proposed algorithm restricts maximum transmission power with the information of state of charge, and allocates transmission power based on historical information. In addition, energy availability constraint is given by rigorous theoretical analysis to guarantee the optimization of average long-term throughput. Simulations have been conducted to demonstrate the effectiveness of the algorithm without considering probability distribution of energy arrival or channel coefficients. The proposed algorithm outperforms counterparts in different energy harvested rates.

Keywords Power control · Energy harvesting · Finite capacity · Throughput optimization Wireless communication

Abbreviations

D	Upper bound on the diameter of $\mathbb{P}$
EH	Energy harvesting
E^{max}	Battery capacity
$\mathbb{P}$	Feasible set
G	Upper bound of the gradient of $U_t(p;s)$
MDP	Markov decision process
OCO	Online convex optimization
ODG	Online descend gradient
P^{max}	Maximal transmission power of each time slot

✉ Xu Zhang
eckel_zhang@163.com

Jing Guo
guojing_cc@163.com

[1] Department of Automation, Foshan University, Guangyun 33 Foshan, Guangdong, China

[2] Present Address: Department of Mechanical and Electronic Engineering, Central South University, Hunan, China

SOC State of charge
$U_t(p;s)$ Throughput with power p at channel s

1 Introduction

Recently, energy harvesting (EH) technologies extend the lifetime of node by collecting ambient energy (solar, vibration, RF, etc.) from environment [1]. The system with EH nodes has been considered to provide extremely cost-effective and fully passive solutions in Internet of Things (IoT) [2], especially in places where conventional power source are not accessible. Though the operation, efficiency and maturity of energy harvesting systems vary with each other, it is indispensability to support long-term node availability in most of real scenarios. Within the operating time, the power supply of EH node should be no less than the power consumption [3]. However, the harvested energy is usually not continuous and sometimes limited due to the intermittent nature of energy sources. Nodes powered by harvested energy need energy storage and power management algorithm to provide uninterrupted power supply [4].

The efficiency of EH technologies is important to enhance node availability, while optimal power control plays a critical role in improving overall performance, especially in EH based wireless communication system [5]. Computing and communication are the main sources of power consumption during node operation. In most cases, the amount of power consumption on communication is bigger than other tasks, such like sensing and computing. Furthermore, it varies large along with the dynamic and unpredictable wireless channel. As a result, the power control problem of a EH based transmitter, which harvests energy from ambient environment and transmits information to remote wireless receivers with harvested energy is a big challenge [6]. The unknown wireless channel and intermittent nature of ambient environmental energy lead to the challenges of ensuring node continuity and maximizing communication throughput [7–9].

In practical scenarios, the characteristics of energy arrival rate and the coefficients of wireless channel are hard to be predicted. The learning theoretic approaches improve the power control strategies by identifying the model of energy arrival rate and channel distribution [10, 11]. However, the computing complexity increases sharply with the states number of energy arrival rate and wireless channel, that is unacceptable for low cost transmitter. Online convex optimization (OCO) [12] is a low complexity way to allocate power effectively by projecting power vector into feasible set, which is determined by constraints. However, the energy of battery may be exhausted if the average of harvested energy is below the maximum of feasible set. Then the OCO with stochastic constraints [13] is deeply analyzed, and a new algorithm of guaranteeing node continuity is proposed which required a costly big-capacity battery [14].

In order to remove the constraint of battery capacity, this paper proposes an improved gradient descent algorithm based on OCO for power control problem. The algorithm adjusts the size of feasible set with the State of charge (SOC) (range is [0, 1]). When the SOC is close to 1, the feasible set is almost full size, which indicates maximum power is allocated. Otherwise, the feasible set is a fractional subset of full size set, extra energy would be stored for future when the harvested energy is bigger than maximum of current fractional subset. By rigorous theoretical analysis, the conditions of energy availability guaranty and the lower bounder of average long term throughput are explicitly

given. Furthermore, the throughput and battery state are analyzed, the simulations demonstrate the excellent results.

The benefits of proposed algorithm are as follows:

- Firstly, comparing with learning theoretic approaches [11], the proposed algorithm has low complexity and only relies last historical information.
- Secondly, the node continuity or energy available guaranty is fulfilled without a complex setting of battery capacity, which is indispensable in algorithm [14].
- Thirdly, the proposed algorithm is scalable, that is easily to be deployed into the scenarios with one transmitter to many receivers.

The paper is organized as follows. In the second section, related work is discussed. In the third section, system model is given and the problem is formulated. The new algorithm is proposed in the fourth section. In the fifth section, the performances are analyzed including average long-term throughput and energy-availability guaranty. In the sixth section, simulations are shown and discussed to verify the advantageous properties. Finally, conclusion is given.

2 Related Work

The power control or the energy allocation problem in EH based wireless system has been widely researched. Existing research can be divided into two groups based on whether the information (about the energy arrival and channel state) assumed to be available at the transmitter or not.

In offline optimization group [15], EH transmitter has almost ideal knowledge of future energy arrival amount or perfect prediction of channel state. Solar-based EH systems work well under such assumption, where the amount of harvested energy is almost predictable. In [16], the authors model the energy arrival rate and channel distribution as Markov Process with known transition probability, optimize the power control by dynamic programming. The transmitter allocates more energy for transmission when the wireless channel is predicted to be good, less energy in bad channel. When the wireless channel or energy arrival rate is known partially, the power control strategies are implemented based on prediction of unknown part, such like the work on unknown channel distribution, on unknown energy arrival rate. In [17], the authors focus on the situation of hybrid energy storage, and battery imperfections situation is discussed in [18]. In [19], the authors consider a multihop EH communication system and cover all possible harvesting profiles including continuous and discrete cases. Overall the solution of offline strategies show the upper bound of optimized throughput.

In online optimization group [8, 10], EH transmitter is assumed to know the statistics of underling EH arrival process or to have causal information about their realizations. In this case, the EH arrival process is model as approximated model, and online methods make decision on energy allocation based on predefined models. The model could be Markov decision process (MDP) or regression model based on statistic data [20]. In practical situation, the future channel state is unknown, and the learning theoretic approach is suitable for unpredictable case [21]. In this case, the transmitter learns the optimal energy allocation policies by performing actions and observing their rewards. Learning-based algorithms manage to gain rewards by minimizing the gap between online rewards and offline optimal

throughput. In [11], the authors assumed energy arrival and channel state as individual MDPs, and did not know the transition ratios. After a period of learning, the expected average throughput showed convergence in lower bounder. Recently deep reinforcement learning algorithm is implemented in a point-to-point EH wireless communication system where prior information about distribution on energy arrival process and channel coefficient both are not available [22].

For most EH-based wireless systems, the computational capability of node is limited. As a result, the computational complexity of power control algorithm must be considered. In most existing algorithms, the computational loads of value iteration or policy iteration algorithms increases sharply with the number of quantized states and/or actions [15]. The online convex optimization opens up a brand new way of optimizing energy allocation and long-term throughput. The Online descend gradient (ODG) algorithm, a traditional OCO algorithm, achieved acceptable regret on average long-term throughput in energy unlimited case [12]. However, the energy continuity is ignored. In limited energy capacity case, the authors in [13, 14] proposed an updated ODG version. By subtracting a vector, proposed algorithm help restrict total allocated energy. Related analysis showed the performance lower bounder based on assumption of huge battery capacity.

3 System Model

In this paper, a point-to-point EH communication system is considered. The general configuration is similar as [14]. There are n sub-channels between the transmitter and receiver. In the beginning of time slot t, the EH transmitter allocates energy with vector $p[t] = [p_1[t], p_2[t], \dots, p_i[t], \dots, p_n[t]]$, where $p_i[t]$ is the energy allocated for sub-channel i in time slot t. Maximal transmission power of each time slot (P^{max}) is defined. The Feasible set ($\mathbb{P}$) of energy allocation is defined as (1).

$$\mathbb{P} = \left\{ p \in \mathbb{R}^n : \sum_{i=1}^{n} p_i \le P^{max} \right\}. \tag{1}$$

where $p_i \ge 0, \forall i \in \{1, 2, \dots, n\}$.

Battery capacity (E^{max}) is defined, and the harvested energy in time slot t is $e[t]$, which is known only at the end of time slot t. Assume available energy in the beginning of time slot t is $E[t]$. The dynamics of $E[t]$ is described in the following Eq. (2).

$$E[t+1] = E[t] + e[t] - P[t]. \tag{2}$$

where $P[t] = \sum_{i=1}^{n} p_i[t]$ and $E^{max} \ge E[t+1] \ge 0$.

The states of all sub-channels in time slot t are displayed as a vector $s[t] = [s_1[t], s_2[t], \dots, s_n[t]]$, which contains n sub-channels. The corresponding channel capacity of sub-channel i is $log(1 + p_i[t]s_i[t])$ when assigning $p_i[t]$ energy into channel state $s_i[t]$.

The system model is shown in Fig. 1, where EH transmitter chooses $p[t+1]$ at the beginning of time slot $t+1$. We define the reward $U_t(p[t];s[t])$ as the sum throughput of all n sub-channels in (3).

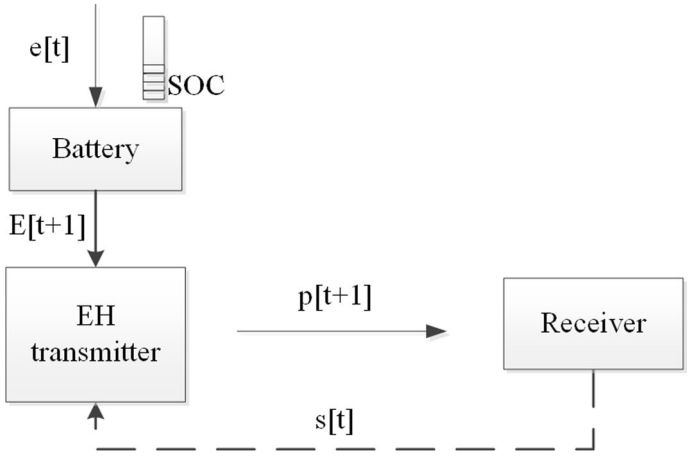


Fig. 1 Basic system model

$$U_t(p[t];s[t]) = \sum_{i=1}^{n} \log(1 + p_i[t]s_i[t]). \tag{3}$$

Throughput with power p at channel s ($U_t(p;s)$) is defined, ant it is a non-negative, non-decreasing, and concave utility function. The $U_t(p;s)$ is obtained at the end of time slot t by calculating (3), while power decision $p[t]$ is decided at the beginning of time slot, where channel state and energy arrival of previous time slot is known, as shown in Fig. 1. So the objective of power control is to find a policy π, a sequence of $p[t], t = 1, 2, \ldots T$, so that the transmitter can send out as more information as possible by all sub-channels in T time slots. The mathematical description is as follow:

$$\max J_\pi = \frac{1}{T} \sum_{t=1}^{T} U_t(p;s), \tag{4}$$

s.t (1) and (2).

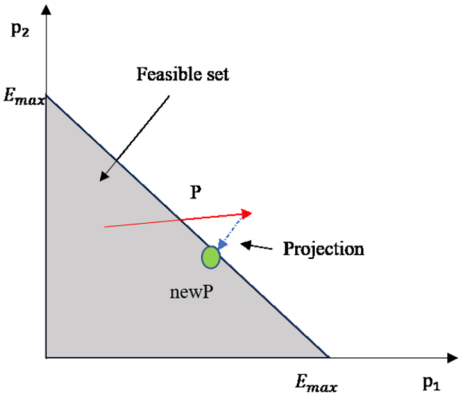
4 Proposed Algorithm

In traditional OCO, the allocated power of next time slot $P[t+1]$ is the projection in fixed feasible set of power vector, which is obtained by ODG method. The feasible set is restricted by E^{max}. As a convex function, the allocated power $P[t+1]$ would be equal to E^{max} without any restriction, for which a sketch in two-dimensional space is shown in Fig. 2a. In energy limited case, the energy availability conditions shown in (5) must be satisfied, that is energy allocation should be no more than available energy.

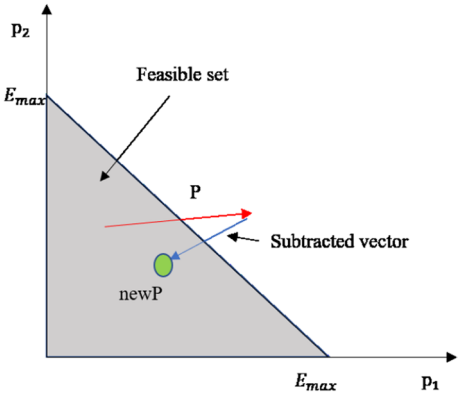
$$E[t+1] \geq P[t+1], \quad \forall t \in 0, 1, 2, \ldots T. \tag{5}$$

In [14], the authors subtract an additional vector from the power vector, so that the new power vector is moved into the inner of feasible set, and a sketch is shown in Fig. 2b. The

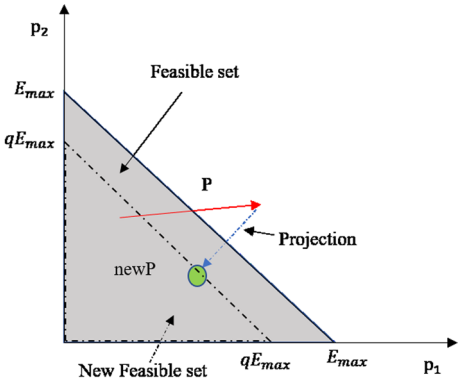
Fig. 2 Power control with OCO algorithms



(a) Original OCO



(b) OCO in [14]



(c) Proposed OCO

modification ensures (5) with the buff provided by a big capacity battery. Inspired by fractional allocation policy in [15], this paper adjusts the feasible set size to ensure (5). The new algorithm obtains SOC (range is [0, 1]) at the beginning of time slot $t+1$, and restricts feasible set of power control via SOC. The power control strategy is handled under the new feasible set. A sketch for the idea of proposed algorithm is shown in Fig. 2c.

The description of the algorithm is shown as follows.

Algorithm:

Setup of the initial time $t = 1$:

$p[1] = [0, 0, \dots 0]^T$ are the initial power of the channels, $E[1] = E_{ini}$ is the initial energy.

Procedures:

At the end of each time slot $t \in \{1, 2, \dots T\}$, do the follows:

Step 1: obtain current SOC $q[t+1]$

$$q[t+1] = E[t+1]/E^{max} \tag{6}$$

Step 2: Restrict the maximum of transmission power.

$$P_q^{max} = q[t+1]P^{max}$$
$$\mathbb{P}_q = \left\{ p \in \mathbb{R}^n : \sum_{i=1}^{n} p_i \le P_q^{max}, \right\} \tag{7}$$

where $p_i \ge 0, \forall i \in 1, 2, \dots, n$

Step 3: power control

$$p[t+1] = Proj_{\mathbb{P}_q}\{p[t] + q[t+1]\nabla_p U_t(p;s)\} \tag{8}$$

where $Proj_{\mathbb{P}_q}\{\cdot\}$ represents the projection onto feasible set $\mathbb{P}_q$ and $\nabla_p U_t(p;s)$ represents a gradient of function $\nabla_p U_t(p;s)$ at point $p = p[t]$.

As shown in Fig. 1 and the algorithm description, the SOC value ($0 \le q[t+1] \le 1$) is available when deciding energy allocation $p[t+1]$, because SOC relies on $E[t+1]$ only. In the step 2, the feasible set of power control is restricted by SOC, which is dynamic along with harvested energy and allocated energy. In step 3, projection is done in the restricted feasible set.

5 Algorithm Analysis

When deploying OCO into energy allocation problem in EH based wireless system, the energy availability guarantee should be fulfilled as (5). Then the regret of actions should be analysed in order to provide the lower bounder of performance. In the following part, the power control policy of proposed algorithm is analysed in above two aspects.

5.1 Energy Availability Guaranty

In order to implement the power control decision in proposed algorithm, the SOC in time slot $t+1$ must satisfy following energy availability constraint (5).

Theorem 1 (Energy guaranty) *If* $P^{max} \leq E^{max}$ *, then the energy availability is guaranteed.*

Proof Based on the definition of $Proj\{\cdot\}$ and algorithm design, we obtain

$$P[t+1] \leq \frac{E[t+1]}{E^{max}} P^{max} = \frac{P^{max}}{E^{max}} E[t+1]$$

If $P^{max} \leq E^{max}$, then

$$P[t+1] \leq \frac{P^{max}}{E^{max}} E[t+1] \leq E[t+1].$$

□

5.2 Lower Bound for Long-Term Average Expected Throughput

Upper bound on the diameter of $\mathbb{P}(D)$ is defined, that is, $\forall x, y \in \mathbb{P}, \|x - y\| \leq D$. Upper bound of the gradient of $U_t(p;s)(G)$ is defined, that is, $\|\nabla_p U_t(p;s)\| \leq G, \forall p \in \mathbb{P}, t \in \{1, 2, \ldots T\}$. Then we give the main result of this paper.

Theorem 2 (Main Result) *If the long term average throughput is defined as* (4)*, then under proposed algorithm, when* $T \to \infty$*, the lower bounder of average long term throughput can be given by*

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} U_t(p;s) \geq \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} U_t(p^*;s) \\ (GD + \frac{G^2}{2}), \tag{9}$$

where $p^* = \arg\max_{p\in\mathbb{P}} \sum_{t=1}^{T} U_t(p;s)$ *in energy unlimited situation.*

Before proving the Theorem 2, we need to introduce the following three lemmas first.

Lemma 1 Let $p_z = Proj_{\mathbb{P}}\{p\}$, and $p_y = Proj_{\mathbb{P}_q}\{p\}$, where $q \in [0, 1]$, Then for any $p \in R^n$, $\|p_z - p_y\| \leq (1 - q)D$.

Proof

1. If $p \notin \mathbb{P}$, then $\|p_z - p_y\| \leq (1 - q)D$.
2. If $p \in \mathbb{P}$, but $p \notin \mathbb{P}_q$, then $\|p_z - p_y\| \leq (1 - q)D$.
3. If $p \in \mathbb{P}_q$, then $\|p_z - p_y\| = 0$.

□

Lemma 2 *Based on proposed algorithm description,* $q[t+1]$ *follows the below inequality constraint.*

$$q[t+1] \geq \left(1 - \frac{P^{max}}{E^{max}}\right) q[t].$$

Proof Based on projection definition, we obtain

$$P[t] = \sum_{i=1}^{n} p_i[t] \leq P^{max} q[t] = P^{max} \frac{E[t]}{E^{max}}.$$

Recall

$$E[t+1] = E[t] + e[t] - P[t].$$

Then

$$\begin{aligned} E[t+1] &\geq E[t] + e[t] - P^{max} \frac{E[t]}{E^{max}} \\ &\geq \left(1 - \frac{P^{max}}{E^{max}} E[t]\right) \end{aligned}$$

Therefore,

$$q[t+1] \geq \left(1 - \frac{P^{max}}{E^{max}}\right) q[t].$$

□

Lemma 3 [12] *Let $c_1, c_2, \ldots c_T : \mathbb{P} \to \mathbb{R}$ be an arbitrary sequence of convex differentiable functions. Let $p_1, \ldots, p_T \in \mathbb{P}$ be defined by $p_1 = 0$ and $p_{t+1} = Proj_{\mathbb{P}}\{p_t + \eta \nabla c_t(p_t)\}$, then for $G = \max \|\nabla c_t(p_t)\|$,*

$$\max_{p \in \mathbb{P}} \sum_{t=1}^{T} c_t(p) - \sum_{t=1}^{T} c_t(p_t) \leq \frac{D^2}{\eta} + T\eta \frac{G^2}{2}.$$

Below we give the proof of our main result.

Proof of Theorem 2 Let energy vector $p[t]$ is the output of the proposed algorithm, based on Lemma 1, we can get

$$\|p_z - p\| \leq (1-q)D.$$

Then for all $p, p_z \in \mathbb{P}_q$ and at time slot t, we have

$$\begin{aligned} |U_t(p_z;s) - U_t(p;s)| &\leq G\|p_z - p\| \\ &\leq GD(1 - q[t]). \end{aligned} \tag{10}$$

Sum (10) from $t = 1$ to $t = T$, and consider Lemma 2, we obtain

$$\begin{aligned}\sum_{t=1}^{T} U_t(p;s) &\geq \sum_{t=1}^{T} U_t(p_z;s) - GDT + GD\sum_{t=1}^{T} q[t] \\ &= \sum_{t=1}^{T} U_t(p_z;s) - GDT \\ &\quad + GDq[1]\sum_{1}^{T}\left(1-\frac{P^{max}}{E^{max}}\right)^{T-t} \\ &\geq \sum_{t=1}^{T} U_t(p_z;s) - GDT + GD\frac{E[1]}{P^{max}}\end{aligned}$$

The p_z and p are the output of Zinkevich's policy [12] and proposed policy, respectively. Considering Lemma 3, we get

$$\sum_{t=1}^{T} U_t(p^*;s) - \sum_{t=1}^{T} U_t(p_z;s) \leq \frac{D^2}{2q[t+1]} + \frac{G^2}{2}\sum_{t=1}^{T} q[t]. \tag{11}$$

Combine (11) and (12), then we get

$$\begin{aligned}\sum_{t=1}^{T} U_t(p;s) &\geq \sum_{t=1}^{T} U_t(p^*;s) + GD\frac{E[1]}{P^{max}} - \frac{D^2}{2q[t+1]} \\ &\quad - GDT - \frac{G^2}{2}\sum_{t=0}^{T} q[t] \\ &\geq \sum_{t=1}^{T} U_t(p^*;s) + GD\frac{E[1]}{P^{max}} - \frac{D^2}{2q[t+1]} \\ &\quad - (GD + \frac{G^2}{2})T\end{aligned}$$

From Theorem 1, one gets $E[t] \geq P[t]$. Then $\frac{1}{q[t+1]}$ is bounded. Finally,

$$\lim_{T\to\infty}\frac{1}{T}U_t(p;s) \geq \lim_{T\to\infty}\frac{1}{T}U_t(p^*;s) - (GD + \frac{G^2}{2})$$

□

Table 1 Simulation parameters

Name	Value	Description
n	2	Number of sub-channels
σ_1, σ_2	1	Rayleigh info of sub-channel 1, 2
P^{max}	10	Maximum transmitting power
E^{max}	$\geq P^{max}$	Battery capacity
$e[t]$	Uniform	Energy arrival rate

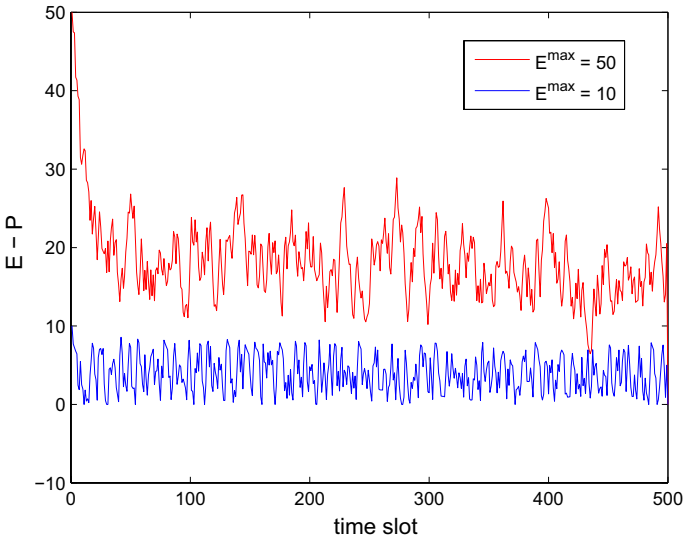


Fig. 3 Energy availability

6 Simulations

In this section, the properties of proposed power control algorithm on EH based wireless system are analyzed. In EH transmitter, there are two sub-channels, which are two independent Rayleigh fading channels. The maximum transmitting power is set to 10. The energy harvesting rate follows uniform distribution within [0, E^{max}]. Firstly, the energy availability guarantee of proposed algorithm is demonstrated. Then the comparison on

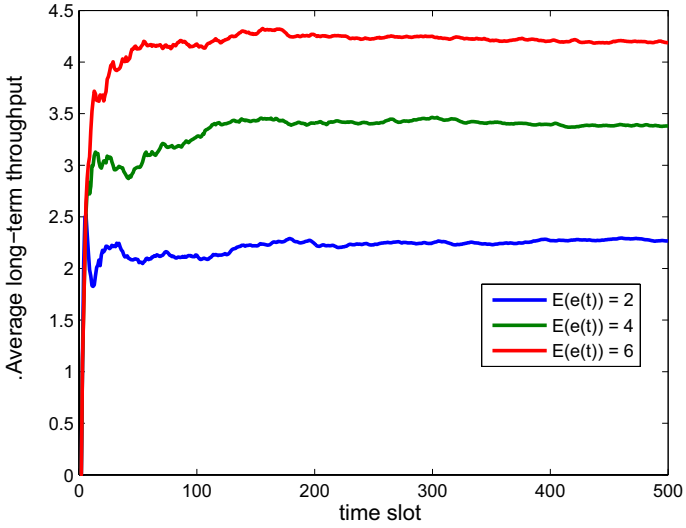


Fig. 4 Average long-term throughput with different energy harvest rate $e[t]$

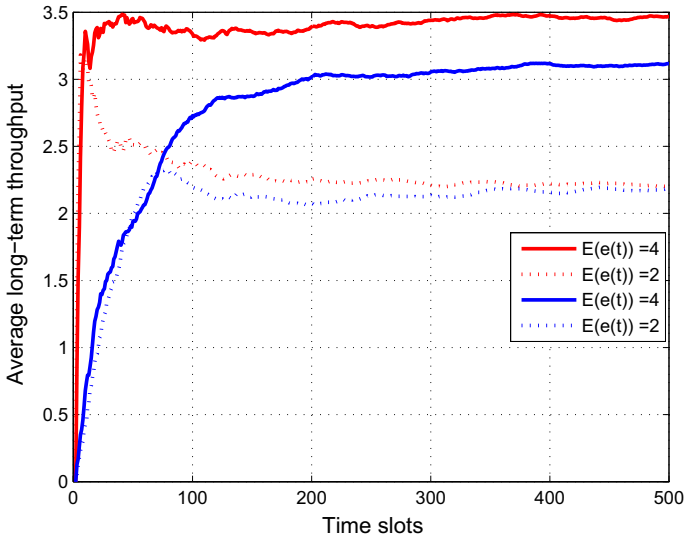


Fig. 5 Average long-term throughput comparison

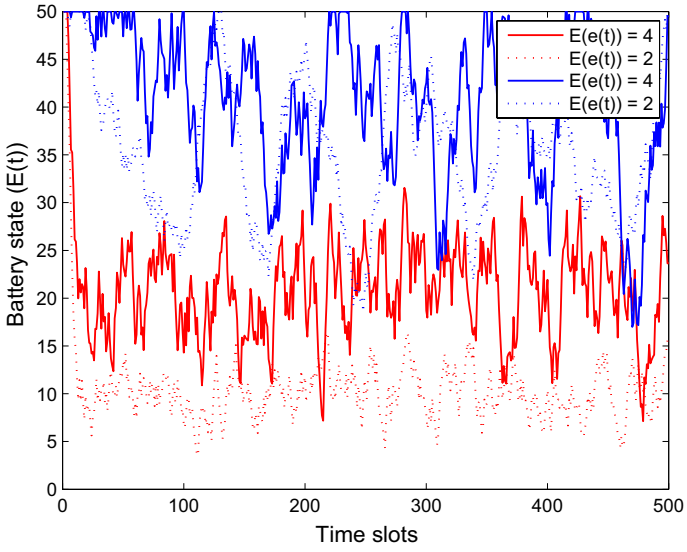


Fig. 6 Battery state comparison

average long-term throughput between proposed algorithm and algorithm in [14] are shown under different setups. The simulation settings are listed in Table 1.

Energy Availability Guarantee Figure 3 shows the trajectories of $E[t] - P[t]$ with different E^{max} under the condition of $E^{max} \geq P^{max}$. When $E^{max} = 50$, the remaining power is fluctuating between 10 and 30 as shown in red line. When $E^{max} = 10$, the remaining power is fluctuating between 0 and 10 as shown in blue line. It can be seen that $E[t] - P[t] \geq 0$

always holds as time goes on, then it is clearly that the energy availability is guaranteed as long as the inequality $E^{max} \geq P^{max}$ is satisfied.

The minimal required battery capacity is P^{max}, which is 10 in simulation. However, the minimal required battery capacity [14] exceeds 100, which is much bigger than 10. A small required battery capacity is easy to reduce the overall node cost.

Average Long-Term Throughput Analysis Figure 4 shows the average long term throughput trajectories of the proposed algorithm with different $E(e[t])$, where the $E[e(t)]$ is the mean value of e(t). We can get that the performance relies on the energy harvest rate, that is, when average value of $e(t)$ increases, the average throughput also increases. The sum throughput of all sub-channels is a convex function of allocated power.

In the following part, the comparisons of average long-term throughput and battery state are demonstrated in Figs. 5 and 6, respectively. The average harvest rate $e[t]$ equals to 4 is an example of low energy harvesting rate, while 2 is an example of extremely low energy harvesting rate. In both figures, red color solid and dotted lines are the results of our proposed algorithm, the blues are the results of the compared algorithm in [14]. Regardless of the average harvest rate $e[t]$, we see that both algorithms achieve throughput convergence along with the time slot. In Fig. 5, the solid lines indicate that proposed algorithm outperforms counterpart, the dotted lines share similar performance. Overall, proposed algorithm utilizes harvested energy better.

In order to deeply analyze the throughput performance, the battery state is shown. In Fig. 6, the SOC follows a flat oscillation in proposed algorithm, while in other algorithm, the battery level fluctuates in a wider range. The proposed algorithm uses the SOC as a negative input so as to reduce the dynamic of SOC. In compared algorithm, the allocated powers rely on the subtracted vector, which is mostly affected by the channel distribution. As a result, the blue solid and dotted lines show similar fluctuation level when the channel distribution is same. The SOC dynamic of our proposed algorithm is more conducive to extend battery lifetime comparing with the other algorithm, as deep charge-discharge cycle may reduce the battery lifetime. Furthermore, the average SOC of proposed algorithm is below another algorithm's, that is the main reason of achieving higher throughput.

7 Conclusions

In this paper, a power control problem in EH based wireless communication system aiming to maximize throughput is discussed in this paper. In our setup, the wireless transmitter does not know any future information about channel state and energy arrival rate. The setup is reasonable in actual situation. This paper proposes a simple online algorithm which fulfils energy availability guarantee, and achieves outstanding performance in average long-term throughput. The required battery capacity is small, which is suitable for low-cost wireless sensor node. The battery states are smooth in simulations, which is good for battery lifetime. Furthermore, the analysis model is scalable and is also suitable for one transmitter sends out information to many receivers respectively.

Acknowledgements The work was supported by National Natural Science Foundation of China (NO 61803087) and Guangdong Basic and Applied Basic Reasearch Fund 2019A1515110180 and 2020B1515310003.

References

1. Li, L., Zhang, X., Song, C., & Huang, Y. (2020). Progress, challenges, and perspective on metasurfaces for ambient radio frequency energy harvesting. *Applied Physics Letters, 116*(6), 060501.
2. Zhang, X., Li, H.-X., & Chung, H.S.-H. (2021). Setup-independent UHF RFID sensing technique using multidimensional differential measurement. *IEEE Internet of Things Journal, 8*(13), 10509–10517.
3. Wang, H., Zhou, G., Bhatia, L., Zhu, Z., Li, W., & McCann, J. A. (2020). Energy-neutral and QOS-aware protocol in wireless sensor networks for health monitoring of hoisting systems. *IEEE Transactions on Industrial Informatics, 16*(8), 5543–5553.
4. Calautit, K., Nasir, D. S. N. M., & Hughes, B. R. (2021). Low power energy harvesting systems: State of the art and future challenges. *Renewable and Sustainable Energy Reviews, 147,* 111230.
5. Lee, K., Lee, J.-R., & Choi, H.-H. (2020). Learning-based joint optimization of transmit power and harvesting time in wireless-powered networks with co-channel interference. *IEEE Transactions on Vehicular Technology, 69*(3), 3500–3504.
6. Ku, M.-L., & Lin, T.-J. (2021). Neural network-based power control prediction for solar-powered energy harvesting communications. *IEEE Internet of Things Journal,* p. 1.
7. Gunduz, D., Stamatiou, K., Michelusi, N., & Zorzi, M. (2014). Designing intelligent energy harvesting communication systems. *IEEE Communications Magazine, 52*(1), 210–216.
8. Luo, Y., Lina, P., Zhao, Y., Wang, W., & Yang, Q. (2020). A nonlinear recursive model based optimal transmission scheduling in RF energy harvesting wireless communications. *IEEE Transactions on Wireless Communications, 19*(5), 3449–3462.
9. Ashraf, M., Jung, J., Shin, H. M., & Lee, I. (2018). Energy efficient online power allocation for two users with energy harvesting. *IEEE Signal Processing Letters, 26*(1), 24–28.
10. Dong, M., Li, W., & Amirnavaei, F. (2017). Online joint power control for two-hop wireless relay networks with energy harvesting. *IEEE Transactions on Signal Processing, 66*(2), 463–478.
11. Blasco, P., Gunduz, D., & Dohler, M. (2013). A learning theoretic approach to energy harvesting communication system optimization. *IEEE Transactions on Wireless Communications, 12*(4), 1872–1882.
12. Zinkevich, M. (2003). Online convex programming and generalized infinitesimal gradient ascent. In *Proceedings of the 20th international conference on machine learning (icml-03)* (pp. 928–936).
13. Yu, H., Neely, M. J., & Wei, X. (2017). Online convex optimization with stochastic constraints. arXiv preprint arXiv:1708.03741.
14. Yu, H., & Neely, M. J. (2019). Learning-aided optimization for energy-harvesting devices with outdated state information. *IEEE/ACM Transactions on Networking, 27*(4), 1501–1514.
15. Shaviv, D., & Özgür, A. (2018). Online power control for block IID energy harvesting channels. *IEEE Transactions on Information Theory, 64*(8), 5920–5937.
16. Shaviv, D., & Özgür, A. (2017). Approximately optimal policies for a class of markov decision problems with applications to energy harvesting. In *2017 15th international symposium on modeling and optimization in mobile, ad hoc, and wireless networks (WiOpt)* (pp. 1–8). IEEE.
17. Annapandi, P., Banumathi, R., Pratheeba, N. S., & Manuela, A. A. (2021). An efficient optimal power flow management based microgrid in hybrid renewable energy system using hybrid technique. *Transactions of the Institute of Measurement and Control, 43*(1), 248–264.
18. Ozel, O., Shahzad, K., & Ulukus, S. (2014). Optimal energy allocation for energy harvesting transmitters with hybrid energy storage and processing cost. *IEEE Transactions on Signal Processing, 62*(12), 3232–3245.
19. Rezaee, M., Mirmohseni, M., Aggarwal, V., & Aref, M. R. (2018). Optimal transmission policies for multi-hop energy harvesting systems. *IEEE Transactions on Green Communications and Networking, 2*(3), 751–763.
20. Baknina, A., & Ulukus, S. (2013). Online scheduling for energy harvesting two-way channels with processing costs. In *2016 IEEE Globecom workshops (GC Wkshps)* (pp. 1–6). IEEE.
21. Wu, K., Tellambura, C., & Jiang, H. (2017). Optimal transmission policy in energy harvesting wireless communications: A learning approach. In *2017 IEEE international conference on communications (ICC)* (pp. 1–6). IEEE.

22. Li, M., Zhao, X., Liang, H., & Fengye, H. (2019). Deep reinforcement learning optimal transmission policy for communication systems with energy harvesting and adaptive mqam. *IEEE Transactions on Vehicular Technology, 68*(6), 5782–5793.

Jing Guo is currently an Assistant Professor with the Department of Automation, Foshan University, China. She received the Ph.D degree in Control Theory and Control Engineering from Zhejiang University, Hangzhou, China, in 2011. From 2011 to 2012, she was an Assistant Researcher in Shenzhen Institutes of Advanced Technology, Chinese Academy of Science, Shenzhen, China. From 2012 , she join Foshan University, China. From 2017 to 2018, she was a Guest Researcher with the Engineering and Technology institute of Groningen, University of Groningen. Her research interests include wireless sensor network, control of multi-agent and network systems and distributed decision-making and coordination.

Xu Zhang is currently a Ph.D student of Central South University and City University of Hong Kong. His research interests include wireless sensor network, energy harvesting and RFID.