SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •

February 2020, Vol. 63 122302:1–122302:14
https://doi.org/10.1007/s11432-019-9936-y

Joint utility optimization for wireless sensor networks with energy harvesting and cooperation

Pengcheng ZHU, Bingqian XU, Jiamin LI* & Dongming WANG

National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096*, China*

Received 26 March 2019/Revised 3 June 2019/Accepted 17 July 2019/Published online 16 January 2020

Abstract In wireless sensor networks (WSNs), the limited battery capacity restricts the lifetime of the sensor nodes and thus degrades the system performance. The energy harvesting and cooperation techniques are promising solutions to prolonging the battery life, by collecting energy from ambient environment and exchanging energy among sensor nodes. This paper studies the joint utility maximization problem for WSNs in consideration of energy harvesting and cooperation. We first derive an upper bound on the Lyapunov drift for the network stability, and then formulate the optimization as a stochastic optimization problem. Furthermore, we propose an energy harvesting and energy transfer, data transmission, power control, routing and scheduling (EDPR) online algorithm by combining Lyapunov optimization technique with drift-plus-penalty method and perturbation technique. It contributes to optimal utility in a distributed manner along with a balanced trade-off between network utility and queue backlog, with no need for any statistical information about dynamic systems and no concern of curse of dimensionality under large queue backlog. Simulation results also show the practicality of the proposed algorithm in real implementation since data transmission has a linear relationship with battery life.

Keywords drift-plus-penalty, energy management, Lyapunov analysis, utility optimization, WSNs

Citation Zhu P C, Xu B Q, Li J M, et al. Joint utility optimization for wireless sensor networks with energy harvesting and cooperation. Sci China Inf Sci, 2020, 63(2): 122302, https://doi.org/10.1007/s11432-019-9936-y

1 Introduction

With the accessibility of low-cost hardware, development demands of data transmission are ever increasing. This enables wireless sensor networks (WSNs) to become a promising system framework, where many interconnected sensor nodes work together to obtain and fuse data originated from heterogeneous sources [1]. Thus, WSNs can be widely applied to diverse areas including multimedia monitoring networks, real-time tracking of objects, traffic management systems, and disaster management [2–4]. However, WSNs also face some challenges, such as limited bandwidth, data storage, battery power consumption and limited processing capabilities. Since the sensor nodes in WSNs are almost all powered by batteries, the limited lifetime of the WSNs has been the key problem to be addressed urgently, which involves energy efficient transmission and low-energy consumption [5]. As one of the new approaches in energy management, the energy harvesting technique is exploited in various wireless networks, including WSNs [6], cellular networks [7, 8] and ad-hoc networks [9]. Efforts are made to solve this problem via energy harvesting, that is sensor nodes are equipped with energy-harvesting devices and then replenish energy from external environment, thus prolonging the battery life and improving the performance of WSNs.

* Corresponding author (email: lijiamin@seu.edu.cn)

However, the dynamic characteristics in energy harvesting framework increase the difficulty in developing some transmission policy from many aspects of nodes, such as the collected energy, energy demands, and surrounding environment. For instance, an excessive use of energy may cause an energy outage, with nodes not working properly. In addition, sensor nodes have limited available energy due to the battery capacity, and thus the strategy is required to adapt to future state of the nodes. All these enhance the difficulty in dealing with the utility optimization in WSNs.

There has been some previous work on dealing with utility optimization strategies for energy harvesting nodes. An adaptive energy harvesting system was proposed in [6], which extended the lifetime of WSNs by solar panels or batteries. Michelusi et al. [10] put forward transmission policies when sensor nodes are supplied by time-correlated energy. The reinforcement learning approach was adopted for developing heuristic algorithms to handle the energy management in [11, 12]. Typically, such optimization problems can be tackled by dynamic programs (DPs), whereas they are often premised on great knowledge of statistical information about dynamic systems [13] and may lead to curse of dimensionality under large queue backlog. Among the proposed schemes designed for energy management and utility optimization, the algorithms based on quadratic Lyapunov function attracted much attention [14–16]. Additionally, the drift-plus-penalty technique can be utilized to balance the utility performance and the network stability, which can be interpreted as the first-order Lagrangian dual method for constrained optimization [17–19]. Le et al. [20] established a joint flow control, routing, and power control algorithm by introducing the Lyapunov drift technique and achieved the optimal utility in a finite queue length, whose network model could be only applied to single-hop WSNs. In WSNs, Ref. [21] proposed a new heterogeneous energy supply model by the coexistence of renewable energy and electricity grid. According to [22], the energy transfer efficiency in WSNs is around 40%, which means that energy cooperation can be of great benefit to utility optimization. However, most of the researches assume that the energy consumption is independent and the energy exchange does not exist among nodes. Ref. [23] provided an overview of state-of-the-art RF-enabled wireless energy transfer (WET) in wireless powered communication. However, energy harvesting and cooperation are not taken into consideration together, which may provide WSNs with potentially infinite lifetime. Therein, the available energy in the buffer of the sensor nodes in such dynamic systems greatly complicates the design of an efficient scheduling algorithm due to the fact that the current energy expenditure decision may cause energy outage in the future and thus affect the future decisions.

In this paper, we take the multi-hop WSNs system into consideration. The main contributions of this paper are listed as follows.

- We consider both energy harvesting and cooperation in WSNs, where nodes not only exchange energy among each other but also harvest energy from the surrounding environment.
- We provide an upper bound on the Lyapunov drift to maintain the network stability, that is nodes can gather energy only when the size of energy queue is lower than the threshold. It gives specific network queue backlog bounds and provides explicit calculation of the required energy storage capacities for the sensor nodes in multi-hop WSNs.
- We formulate the joint utility optimization as a stochastic optimization problem so as to maximize the system utility concerning receiving rate of data. And the problem is handled by using Lyapunov technique, combined with the ideas of drift plus penalty [24] and perturbation technique [25]. We design a joint energy harvesting and energy transfer, data transmission, power control, routing and scheduling (EDPR) online algorithm. The algorithm is proposed based on a quadratic Lyapunov function and then carefully perturbs the penalty weights used for decision making, so as to push the target queue length towards certain nonzero value to avoid underflow. It is able to achieve the optimal utility in a distributed manner by jointly optimizing the energy storage and power allocation policy with the battery capacity limitation, the data rate bound, the network-stability requirement, and the energy-availability constraint. In addition, there is no need for any statistical information about the dynamic systems and it can avoid curse of dimensionality even with large queue backlog.
- Simulation results show that EDPR algorithm achieves a good time-averaged $[\mathcal{O}(V), \mathcal{O}(1/V)]$ trade-off between network utility and queue size. It also reveals a linear relationship between data transmission

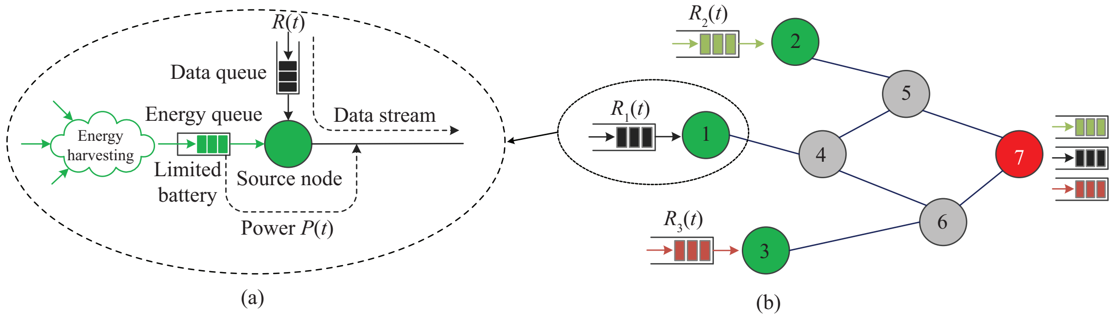

Figure 1 (Color online) (a) The energy harvesting and data receiving process of a source node; (b) the topology of a WSN.

and battery size $\mathcal{O}(V)$ for any tunable parameter $V > 0$.

The remainder of this paper is organized as follows. Section 2 first provides the network model including the link capacity model as well as the data and energy queue model in WSNs with energy harvesting and cooperation. And we then formulate the utility maximization with energy management problem. In Section 3, we derive the upper bound of Lyapunov drift and propose an online algorithm to solve the stochastic problem based on Lyapunov technique. Section 4 investigates the performance of the proposed algorithm by numerical analysis. Simulation results are presented to validate our scheme in Section 5. Finally, Section 6 concludes our work.

2 Network model and problem formulation

We consider a general multi-hop energy harvesting and cooperation WSN system which operates in slotted time. The source nodes can harvest energy from the environment, such as electromagnetic energy, wind and solar energy. The intermediate nodes can receive and deliver energy, whereas they cannot collect energy from the outside as shown in Figure 1. In the network, the harvested energy is assumed to be utilized for the transmission of data streams. The network is modeled by an aperiodic directed graph $G = (\mathcal{N}, \mathcal{L})$, where $\mathcal{N} = \{1, 2, \ldots, N\}$ is the set of the network nodes. We set $\mathcal{L} = \{(n, m), n, m \in \mathcal{N}\}$ as the links between two nodes, where (n, m) represents that the data streams or energy transfers from node n to node m. $\mathcal{N}_n^{(\text{in})}$ denotes nodes where streams flow into node n, and the set $\mathcal{N}_n^{(\text{out})}$ stands for nodes where streams flow out node n. Then, we define

$$D_{\max} = \max_n \big(|\mathcal{N}_n^{(\text{in})}|, |\mathcal{N}_n^{(\text{out})}|\big) \tag{1}$$

as the maximum of the in-degree and out-degree of all nodes in the network.

2.1 Link capacity model

In the WSN system, each node needs to allocate power for data transmission. We set the transmission power allocation matrix at time slot t as $\boldsymbol{P}_n(t) = \{P_{nm}(t), n, m \in \mathcal{N}\}$, where $P_{nm}(t)$ is the power allocated on link $(n, m) \in \mathcal{L}$. It is assumed that each node has a finite capacity buffer and then $P_{nm}(t)$ satisfies the following constraint:

$$\text{C1}: \ 0 \leqslant \sum_{m \in \mathcal{N}_n^{(\text{out})}} P_{nm}(t) \leqslant P_{\max}, \quad \forall n \in \mathcal{N}. \tag{2}$$

We then define the link state set as $\{S_{nm}(t), n, m \in \mathcal{N}\}$, where $S_{nm}(t)$ is the channel gain on link (n, m) at time t, i.i.d for each time slot. The capacity of data stream on link $(n, m) \in \mathcal{L}$ is a function depending on the link state and power allocation. In this paper, we assume that there exists no mutual interference among all data transmission process. Therefore, at time slot t, the signal interference noise rate (SINR) can be written as

$$\gamma_{nm}(t) \triangleq \gamma_{nm}(P_{nm}(t), S_{nm}(t)) = \frac{P_{nm}(t) S_{nm}(t)}{N_0 + \sum_{(n,k) \in \mathcal{L}} P_{nk}(t) S_{nk}(t)}, \tag{3}$$

where N_0 is the spectral density of noise.

Thus, we can obtain the link capacity

$$C_{nm}(t) = \log_2(1 + \gamma_{nm}(t)), \tag{4}$$

where $0 \leqslant C_{nm}(t) \leqslant C_{\max}$. Since the capacity of data stream on link (n, m) has an upper limit, we assume that it is bounded by a certain linear function of the assigned power according to [26], that is

$$C_{mn}(t) \leqslant \delta P_{nm}(t), \tag{5}$$

where the positive constant δ restricts the data transmitted per unit of energy. Let $\lambda_{nm}^{(d)}(t)$ be the rate of data stream d transmitted over the link (n, m) at time slot t, and then it is obvious that

$$\mathrm{C2}: \sum_d \lambda_{nm}^{(d)}(t) \leqslant C_{nm}(t), \quad \forall (n, m) \in \mathcal{L}. \tag{6}$$

2.2 Data queue model

Data streams are transmitted to destination nodes in packets through the links in WSNs. Here we define the backlog of the data streams as $\boldsymbol{q}(t) = \{q_n^{(d)}(t), n, d \in \mathcal{N}\}$, $t = 0, 1, 2, \ldots$, where $q_n^{(d)}(t)$ refers to the quantity of the data stream d accumulated on node n by time sot t. In the initial state, the data storage devices are assumed to have no data stream, that is $q_n^{(d)}(0) = 0$. To ensure the stability of the network, the backlog of the data streams needs to fulfill the following constraint [17]:

$$\overline{\boldsymbol{q}} = \lim_{t\to\infty} \sup \frac{1}{t} \sum_{\tau=0}^{t-1} \sum_{n,d} \mathrm{E}\{q_n^{(d)}(\tau)\} < \infty. \tag{7}$$

Hence, the stability requirement for each node n can be written as

$$\mathrm{C3}: \lim_{t\to\infty} \sup \frac{1}{t} \sum_{\tau=0}^{t-1} \mathrm{E}\{q_n^{(d)}(\tau)\} < \infty. \tag{8}$$

With data d being transmitted to node n, the data queue is updated constantly over time as

$$q_n^{(d)}(t+1) = \left[q_n^{(d)}(t) - \sum_{m \in \mathcal{N}_n^{(\mathrm{out})}} \lambda_{nm}^{(d)}(t) + \sum_{k \in \mathcal{N}_n^{(\mathrm{in})}} \lambda_{kn}^{(d)}(t) + R_n^{(d)}(t) \right]^+, \quad \forall n, d \in \mathcal{N}, \tag{9}$$

where $R_n^{(d)}(t)$ denotes the input data d in packets on node n at time slot t and satisfies $0 \leqslant R_n^{(d)}(t) \leqslant R_{\max}$. Here the notation represents $[f]^+ = \max\,[0, f]$, which is due to the fact that the outflow data should not exceed the current backlog at slot t if there exists little or no data stream in the system.

2.3 Energy queue model

Each sensor node is equipped with a battery with limited capacity for energy harvest, storage and transfer. We define $\boldsymbol{e}(t) = \{e_n(t), n \in \mathcal{N}\}$, $t = 0, 1, 2, \ldots$ as the energy queue backlog and express $\boldsymbol{\varepsilon} = \{\varepsilon_{nm}(t), n, m \in \mathcal{N}\}$ as the energy transfer energy transfer process, where $\varepsilon_{nm}(t)$ stands for the energy transferred from node n to node m. Therefore, the sum of the consumed power and transferred energy cannot exceed what has already been stored in the buffer. In other words, the power allocation and energy transfer for every node must satisfy the following energy-availability constraint at any time slot t:

$$\mathrm{C4}: \sum_{m \in \mathcal{N}_n^{(\mathrm{out})}} (P_{nm}(t) + \varepsilon_{nm}(t)) \leqslant e_n(t), \quad \forall n. \tag{10}$$

We make the assumption that the battery of sensor nodes has no energy in the initial state, namely $e_n(0) = 0$, $\forall n$. Based on a more reasonable situation later discussed in Section 3, where sensor nodes actually do not need to harvest energy all the time, we have

$$h_n(t) = \begin{cases} \eta_n^{(\mathrm{H})} \widehat{h}_n(t), & \text{ON}, \\ 0, & \text{OFF}, \end{cases} \tag{11}$$

where $h_n(t)$ and $\widehat{h}_n(t)$ denote the theoretically available energy and actually harvested energy by node n respectively, and the efficiency of energy harvest $\eta_n^{(\mathrm{H})}$ satisfies $0 \leqslant \eta_n^{(\mathrm{H})} \leqslant 1$. Besides, there exists $0 \leqslant \widehat{h}_n(t) \leqslant h_{\max}$, for some $h_{\max} < \infty$.

We assume that both the harvested and transferred energy, possessed by sensor nodes, can be used in the latter time slots. Then, the energy queue backlog of node n evolves as

$$e_n(t+1) = e_n(t) + h_n(t) - \sum_{m \in \mathcal{N}_n^{(\mathrm{out})}} P_{nm}(t) - \sum_{m \in \mathcal{N}_n^{(\mathrm{out})}} \varepsilon_{nm}(t) + \sum_{k \in \mathcal{N}_n^{(\mathrm{in})}} \hat{\varepsilon}_{kn}(t), \quad \forall n \in \mathcal{N}, \tag{12}$$

where $\widehat{\varepsilon}_{kn}(t) = \eta_{kn} \varepsilon_{kn}(t)$ denotes the energy transmission via the link (k, n), and η_{kn} is the energy transfer efficiency with $0 \leqslant \eta_{kn} \leqslant \eta_{\max}$. Thereinto, the infinite capacity of batteries brings about $\sum_{m \in \mathcal{N}_n^{(\mathrm{out})}} \varepsilon_{nm}(t) \leqslant \varepsilon_{\max}$, $\forall t, n$, for some $\varepsilon_{\max} < \infty$.

2.4 Utility optimization with energy management

This paper aims to design a transmission strategy to deal with the utility optimization with energy management (UOEM) problem, where we need to determine the appropriate amount of $R_n^{(d)}(t)$, plan the energy transfer $\varepsilon_{nm}(t)$, and allocate proper power $P_{nm}(t)$ at every time slot t. Thus, we introduce a generic utility function $f(r)$, which needs to be increasing, continuously differentiable and strictly concave in r, with $f(0) = 0$. And then the utility function of the WSNs system is given by

$$f_{\mathrm{tot}}(\overline{\boldsymbol{r}}) = \sum_{n,d} f(\overline{r}_n^{(d)}), \tag{13}$$

where $\overline{\boldsymbol{r}} = \{\overline{r}_n^{(d)}, \forall n, d \in \mathcal{N}\}$ is the mean rate vector, and $\overline{r}_n^{(d)} = \lim_{t \to \infty} \inf \sum_{\tau=0}^{t-1} \mathrm{E}\{R_n^{(d)}(\tau)\}$ is the average receiving rate of data stream d on node n. Therefore, our goal is to maximize the utility function (13) under some constrained conditions, such as the finite capacity constraint (2), the data rate bound (6), the network-stability requirement (8), and the energy-availability constraint (10). Thus it can be formulated as the following stochastic problem:

$$\mathrm{P}_1: \underset{\varepsilon_{nm}, P_{nm}, R_n^{(d)}}{\text{maximize}} \sum_{n,d} f(\overline{r}_n^{(d)}) \quad \text{s.t. C1, C2, C3, C4.} \tag{14}$$

However, the solution of the UOEM problem cannot be easily obtained due to the complexity of the objective function as well as the constraints. To begin with, C1 generally can only be achieved in WSNs with centralized control by supervising all the sensor nodes, which is inferior to distributed computation with preferable flexibility and extendibility. In addition, C3 demands the guarantee that the whole network is stable over average time. This can be satisfied only when the network has immediate knowledge of data queue at all time slots, which is infeasible though. Furthermore, the power allocation and energy transfer couple both current and future actions in C4, which causes that the current state will affect the future. Then the energy queue may be empty, thus resulting in no energy for power allocation in the next time slot. Under usual circumstances, such problems can be transformed and classified into DPs. Nevertheless, DPs may have high complicacy when faced with a large queue backlog. All of these enhance the difficulty in solving the UOEM problem.

So we take Lyapunov drift-plus-penalty method into consideration in that it helps to convert the objective function (13) and constraint C3 into instantaneous expressions, which can be solved in each time slot regardless of the time-average effect. Additionally, we combine the modified quadratic Lyapunov function with weighted perturbation analysis to satisfy the constraint C1 and C4. The energy backlog can be pushed towards certain non-negative values by perturbing the weights used for decision actions, thus successfully avoiding underflow. More details will be explained in Section 3.

3 EDPR algorithm based on Lyapunov optimization

In this section, we propose an online algorithm to solve the UOEM problem based on Lyapunov optimization. The Lyapunov technique proposed in [20] is a commendable tool for stochastic queue network problems. It has several advantages as: (i) the ability to achieve an approximate optimal performance; (ii) the robustness in time-varying networks; (iii) the mathematical tractability. However, if applied to our system directly, the Lyapunov optimization tool is not suitable for networks with "no underflow" constraint, which means it is difficult to deal with the energy-availability constraint C4. Hence, we establish the online algorithm additionally combining the idea of drift-plus-penalty method and perturbation technique. For one thing, the drift-plus-penalty method makes for the trade-off between the network backlog and the utility optimization. For another, the energy queue level can be pushed towards certain non-negative value to get rid of underflow by perturbing the weight of power transmission strategy at each time slot. Finally, a joint EDPR online algorithm is proposed.

3.1 Upper bound of Lyapunov drift

We first introduce a perturbation parameter $\{\theta_n, n \in \mathcal{N}\}$ related to energy queue, and then define a perturbed quadratic Lyapunov function as

$$L(\boldsymbol{Q}(t)) = \frac{1}{2}\sum_{n,d\in\mathcal{N}} [q_n^{(d)}(t)]^2 + \frac{1}{2}\sum_{n\in\mathcal{N}} [e_n(t) - \theta_n]^2, \tag{15}$$

where $\boldsymbol{Q}(t) = \{\boldsymbol{q}(t), \boldsymbol{e}(t)\}$ denotes the queue backlog in the network. We can speculate it from (15) that a smaller value of $L(\boldsymbol{Q}(t))$ may bring about less backlog of energy and data queue, thus increasing stability of the network. For one thing, the reduction in $L(\boldsymbol{Q}(t))$ helps to keep the data queue fluctuating in a narrow range in that the data queue is in a low congestion state. For another, it causes the energy queue to tend towards some corresponding perturbation parameter, which is the target threshold we set, so as to satisfy the energy-availability constraint.

For the sake of further analysis, we define the Lyapunov drift as the variation of $\boldsymbol{Q}(t)$ between time slots as

$$\Delta(\boldsymbol{Q}(t)) = \mathrm{E}\{L(\boldsymbol{Q}(t+1)) - L(\boldsymbol{Q}(t)) | \boldsymbol{Q}(t)\}. \tag{16}$$

By continuously pushing the Lyapunov drift (16) to a lower value at every time slot t, we can guarantee the stability of the network [17], which means that we need to minimize the Lyapunov drift $\Delta(\boldsymbol{Q}(t))$ for maintaining network stability as C3.

Therefore, we provide an upper bound on the Lyapunov drift in Theorem 1.

Theorem 1. At any time slot t, the Lyapunov drift $\Delta(\boldsymbol{Q}(t))$ is upper-bounded by (17) under any feasible energy management, data transmission, power control, routing and scheduling circumstances.

$$\begin{aligned} \Delta(\boldsymbol{Q}(t)) \leqslant B &- \mathrm{E}\left\{\sum_{n\in\mathcal{N}}\left[\widehat{E}_n(t)\sum_{m\in\mathcal{N}_n^{(\mathrm{out})}} P_{nm}(t) - h_n(t)\right]\middle|\boldsymbol{Q}(t)\right\} \\ &- \mathrm{E}\left\{\sum_{n\in\mathcal{N}}\sum_{m\in\mathcal{N}_n^{(\mathrm{out})}} \varepsilon_{nm}(t)\eta_{nm}\left[\widehat{E}_n(t) - \widehat{E}_m(t)\right]\middle|\boldsymbol{Q}(t)\right\} \\ &- \mathrm{E}\left\{\sum_n\sum_d q_n^{(d)}(t)\left[\sum_{m\in\mathcal{N}_n^{(\mathrm{out})}} \lambda_{mn}^{(d)}(t) - \sum_{k\in\mathcal{N}_n^{(\mathrm{in})}} \lambda_{kn}^{(d)}(t) - R_n^{(d)}(t)\right]\middle|\boldsymbol{Q}(t)\right\}. \end{aligned} \tag{17}$$

Therein, B is a finite positive constant, which satisfies

$$B \geqslant N^2\left(\frac{3}{2}D_{\max^2}C_{\max^2}+R_{\max^2}\right)+\sum_{n=1}^{N}\sum_{m=1}^{N}(1-\eta_{nm})\varepsilon_{\max} + N\left\{\left[D_{\max}\left(P_{\max}+\varepsilon_{\max}\right)\right]^2+\left[D_{\max}\varepsilon_{\max}+h_{\max}\right]^2\right\}. \quad (18)$$

Proof. Please refer to Appendix A.

Since the network stability is the precondition of the utility optimization, we need to minimize the right hand side (R.H.S.) of the drift bound in our algorithm.

3.2 Drift-plus-penalty optimization

After handling the dynamic queue and network stability, we then further analyze the utility maximization problem in this subsection. On the basis of the drift-plus-penalty method, the UOEM problem can be converted into the minimization of an upper bound on penalty function in each time slot t.

Above all, we introduce a control parameter V, which can be any non-negative constant. By adding the penalty expression to the upper bound on drift in (17), we obtain the drift-plus-penalty expression, namely

$$\Delta_V(\boldsymbol{Q}(t))=\Delta(\boldsymbol{Q}(t))-V\mathrm{E}\left\{\sum_{n,d\in\mathcal{N}}f(R_n^{(d)}(t))|\boldsymbol{Q}(t)\right\}. \quad (19)$$

When we set V as zero, Eq. (19) is equivalent to the previous problem that minimizes the Lyapunov drift. Additionally, when V is a positive constant, Eq. (19) shows a certain trade-off between the network backlog and the utility maximization, which reflects how much we emphasize the system stability compared to the utility performance. Combining (19) with (17), we derive an upper bound on Lyapunov drift-plus-penalty

$$\begin{aligned}\Delta_V(\boldsymbol{Q}(t)) \leqslant B &- \sum_{n,d}\left[Vf(R_n^{(d)}(t))-q_n^{(d)}(t)R_n^{(d)}(t)\right]-\sum_{n}\sum_{m\in\mathcal{N}_n^{(\mathrm{out})}}\varepsilon_{nm}(t)\eta_{nm}\left[\widehat{E}_n(t)-\widehat{E}_m(t)\right]\\ &-\sum_{n}\left\{\sum_{d}\sum_{k\in\mathcal{N}_n^{(\mathrm{out})}}\lambda_{nk}^{(d)}(t)\left[q_n^{(d)}(t)-q_k^{(d)}(t)\right]-\left[e_n(t)-\theta_n\right]\sum_{m\in\mathcal{N}_n^{(\mathrm{out})}}P_{nm}(t)\right\}\\ &+\sum_{n=1}^{N}\left[e_n(t)-\theta_n\right]h_n(t).\end{aligned} \quad (20)$$

On account of the minus sign in the R.H.S. of the above expression, the utility maximization problem P_1 can be transferred into a minimization problem. Thus, the UOEM problem can be formulated as follows, with constraint C3 incorporated into the Lyapunov optimization framework and the original utility function contained in the R.H.S. of (20).

$$\mathrm{P}_2:\ \underset{\varepsilon_{nm},P_{nm},R_n^{(d)}}{\text{minimize}}\ \text{R.H.S of } \Delta_V(\boldsymbol{Q}(t)) \quad \text{s.t. C1, C2, C4.} \quad (21)$$

Considering the difficulty in minimizing the drift-plus-penalty at every time slot directly, we thus design an online algorithm in the next subsection to solve the problem P_2.

3.3 EDPR algorithm

In this subsection, we propose an online algorithm as the transmission strategy to achieve joint utility optimization. This joint energy management algorithm is named after EDPR algorithm, present as follows:

• **Initialization.** Select V and define the perturbation parameter θ_n as

$$\theta_n = \delta(\beta V + R_{\max}) + P_{\max} + \varepsilon_{\max}, \tag{22}$$

where β is the maximum first derivative of the utility function for all nodes, that is $\beta = \max_{n,d} \beta_n^{(d)} = \max_{n,d} \frac{\partial f}{\partial r}$. Note that we have known the utility function and the threshold of other parameters, the value of θ_n can be easily determined.

• **Energy management.** If $\mathcal{E}_n(t) - \theta_n < 0$, node n will perform the energy harvesting action and store energy, that is $h_n(t) = \eta_n^{(\mathrm{H})}\widehat{h}_n(t)$; otherwise, there is no need to harvest energy, and set $h_n(t) = 0$.

• **Data transmission.** At time slot t, let node n receive appropriate rate $R_n^{(d)}(t)$ of data stream d according to the solution to the following constraint satisfaction problem:

$$\text{maximize } Vf(R_n^{(d)}(t)) - q_n^{(d)}(t)R_n^{(d)}(t) \quad \text{s.t. } 0 \leqslant R_n^{(d)}(t) \leqslant R_{\max}. \tag{23}$$

• **Power control.** Define the weight of energy transfer and of the data stream d on link (n, m) respectively as

$$W_{(n,m)}(t) = \left[\eta_{nm}\left(\widehat{E}_n(t) - \widehat{E}_m(t)\right) - \zeta\right]^+, \tag{24}$$

and

$$\widetilde{W}_{(n,m)}^{(d)}(t) = \left[q_n^{(d)}(t) - q_m^{(d)}(t) - \tau\right]^+, \tag{25}$$

where $\zeta = D_{\max}\eta_{\max}\varepsilon_{\max} + \theta_{\max}$, and $\tau = D_{\max}C_{\max} + R_{\max}$. Thus, the corresponding weight of node n and of the data link (n, m) are accordingly

$$W_n(t) = \max_m W_{(n,m)}(t), \tag{26}$$

and

$$\widetilde{W}_{(n,m)}(t) = \max_d \widetilde{W}_{(n,m)}^{(d)}(t). \tag{27}$$

Then, seek proper energy transfer and power allocation actions $P_{nm}^*(t)$, $\varepsilon_{nm}^*(t)$ under the finite-capacity constraint C1 and energy-availability constraint C4 in order to maximize

$$G(\varepsilon_{nm}(t), P_{nm}(t)) = \sum_n \left[\sum_{m \in \mathcal{N}_n^{(\mathrm{out})}} \varepsilon_{nm}(t)W_{(n,m)}(t) + \sum_{m \in \mathcal{N}_n^{(\mathrm{out})}} \lambda_{nm}(t)\widetilde{W}_{(n,m)}(t) + \widehat{E}_n(t) \sum_{m \in \mathcal{N}_n^{(\mathrm{out})}} P_{nm}(t)\right]. \tag{28}$$

Although each node can perform energy harvesting independently, the maximization of (28) requires centralized control which might be challenging in practical implementation. Therefore, we further simplify the power control part in the EDPR algorithm in a distributed manner as below.

• In case of $e_n(t) < \theta_n$, the node n can transfer neither its energy nor data stream to other nodes. Thus, set $\varepsilon_{nm}(t) = P_{nm}(t) = 0$. Moreover, it is obvious that $W_n(t) > 0$ on this occasion.

• In case of $e_n(t) \geqslant \theta_n$ and $W_n(t) < 0$, the node n cannot transfer energy to other nodes with $\varepsilon_{nm}(t) = 0$, whereas it is able to transmit data streams at $P_{nm}(t) = P_{nm}^*(t)$.

• In case of $W_n(t) > 0$, node n not only transfers energy to node m with the energy of $\varepsilon_{nm}(t) = \varepsilon_{nm}^*(t)$, but also transmits data streams at the power of $P_{nm}(t) = P_{nm}^*(t)$.

Hence, the problem (28) can be solved in a distributed manner by synthesizing the routing and scheduling strategy. By simplifying the power control process, the redundant constraint C1 and C4 can be eliminated, which is further explained in the next section.

• **Routing and scheduling.** For node n, find the optimal $d^* \in \arg\{\max_d \widetilde{W}_{(n,m)}^{(d)}(t)\}$. The specific transmission scheduling strategy at tme slot t is implemented as: If $\widetilde{W}_{(n,m)}^{(d^*)}(t) > 0$, set $\lambda_{nm}^{(d^*)}(t) = C_{nm}(t)$, which means allocating the full rate to the data stream so as to attain the maximum weight; else if $\widetilde{W}_{(n,m)}^{(d^*)}(t) = 0$, set $\lambda_{nm}^{(d^*)}(t) = 0$, that is, no data transmission action is performed.

• **Queue update.** Update the data queue $q_n^{(d)}(t)$ and energy queue $e_n(t)$ according to (9) and (12).

The above proposed EDPR algorithm for solving the original optimization problem P_1 can be summarized in Algorithm 1.

Algorithm 1 EDPR algorithm with optimal utility at time slot t

---

Require: Initialize V and obtain the corresponding θ_n $(n \in \mathcal{N})$ for nodes according to (22);
Ensure: Calculate the optimal amount of data transmission $R_n^{(d)}(t)$, the energy transfer $\varepsilon_{nm}(t)$, and the power allocation $P_{nm}(t)$.
1: Determine $h_n(t)$ for node n through energy management;
2: Determine $R_n^{(d)}(t)$ according to (23) through data transmission;
3: **if** the link from node n to node m exists **then**
4: Calculate the weight $W_n(t)$ and $\widetilde{W}_{(n,m)}(t)$ according to (26) and (27) through power control and routing scheduling;
5: **end if**
6: Obtain $\varepsilon_{nm}(t)$, $P_{nm}(t)$, $\lambda_{nm}^d(t)$ based on the value of $e_n(t)$, θ_n, $W_n(t)$;
7: Update $q_n^{(d)}(t)$ and $e_n(t)$ according to (9) and (12).

---

4 Performance analysis

In this section, we evaluate the performance of the algorithm proposed in Section 3. We can observe it from the energy management strategy that sensor node n harvest energy only when its energy capacity is less than the perturbation parameter θ_n, which indicates that the algorithm can achieve near optimal utility with finite energy buffer. And the data transmission strategy ensures the utility maximization, namely our optimization goal. In addition, it is apparent that the full rate action in routing and scheduling satisfies constraint C2. The combination of power control along with routing and scheduling strategies can minimize the terms including $\varepsilon_{nm}(t)$ and $P_{nm}(t)$ on condition that $\zeta = \tau = 0$. Additionally, we introduce positive constants ζ and τ in corresponding queue defined in (24) and (25), intending to obtain the deterministic upper bounds of the queues, since that of energy storage devices is very practical in the communication systems deployment.

Theorem 2. We have the following findings with respect to the EDPR algorithm.

(a) All the data streams and the energy queues in the update obey the following boundary constraint:

$$0 \leqslant q_n^{(d)}(t) \leqslant \beta V + R_{\max}, \quad \forall (n,d), \tag{29}$$

$$0 \leqslant e_n(t) \leqslant \theta_n + h_{\max} + D_{\max}\varepsilon_{\max}\eta_{\max}, \quad \forall n. \tag{30}$$

As for routing and scheduling, $e_n(t)$ subjects to

$$e_n(t) \geqslant P_{\max} + \varepsilon_{\max}, \tag{31}$$

when the data transmission or energy transfer action is performed.

(b) Let $\overline{\boldsymbol{r}}^*$ denote the optimal time-averaged receiving rate of data streams. We define $\overline{\boldsymbol{r}}(T) = \{\overline{r}_n^{(d)}(T), \forall (n,d)\}$ as the time-averaged receiving rate under EDPR algorithm up to time T, and then we can obtain that

$$\lim_{T\to\infty} \inf f_{\rm tot}(\overline{\boldsymbol{r}}(T)) = \lim_{T\to\infty} \inf \sum_{n,d} f(\overline{r}_n^{(d)}(T)) \geqslant f_{\rm tot}(\boldsymbol{r}^*) - \frac{\widehat{B}}{V}, \tag{32}$$

where $\overline{r}_n^{(d)}(T) = \frac{1}{T}\sum_{t=0}^{T-1} {\rm E}\{R_n^{(d)}\}$, and $\widehat{B} = B + ND_{\max}\varepsilon_{\max}\zeta + N^2\tau D_{\max}C_{\max}$, independent of V.

Proof. Please refer to Appendix B.

Combining part (a) in Theorem 2 with (B5), we can see that $P_{nm}^*(t) > 0$ cannot be the optimal choice if $e_n(t) < P_{\max} + \varepsilon_{\max}$. It turns out that the constraint C1 and C4 are indeed redundant as mentioned earlier in the power control part, which indicates that the energy-harvesting network invariably has enough energy for data transmission in our algorithm. We overcome the energy-availability requirement by carefully perturbing the weight of decision making at each time slot, to push the energy queue level towards certain non-negative value to get rid of underflow.

From Theorem 2, we can see that the time-averaged backlogs of energy and data stream queues are all bounded by $\mathcal{O}(V)$. And part (b) of Theorem 2 implies that the proposed EDPR algorithm achieves a $[\mathcal{O}(V), \mathcal{O}(1/V)]$ trade-off between backlog and utility.

5 Simulation results

We investigate the performance of the proposed EDPR algorithm through simulation. In consideration of the WSN system shown in Figure 1, we suppose that there are totally 7 nodes in the network, of which the first three are source nodes, the other two are relay nodes, and the remaining one is the sink node. Nodes 1, 2, 3 receive data streams (video, audio, images, etc.) and deliver them to the sink node 7 via relay nodes 4, 5, 6. The link state $S_{nm}(t)$ is assigned to be $\{1, 2\}$ with equal probability. It indicates that unit power is able to two packets under good link state whereas it can only send one single packet otherwise. For simplicity, we also assume that there is no interference among all links. Besides, we can get that $D_{\max} = 2$ from the topology in Figure 1, and then figure out the value of θ_n, τ and ζ. The harvested energy $h_n(t)$ takes values in a finite set $\boldsymbol{H} = \{h_1, h_2, \ldots, h_{\text{end}}\}$ with equal probability, i.i.d every time slot. As for the utility function, those of nodes 1, 2, 3 are given as $f_n(\overline{r}_n^{(d)}) = \log(1+\overline{r}_n^{(d)})$ related to the average receiving rate of data streams, which satisfies the conditions of increasing, continuously differentiable and strictly concave mentioned in Subsection 2.4. And those of nodes 4, 5, 6 are given by $f_n(\overline{r}_n^{(d)}) = 0$, since they are relay nodes. As a result, $\beta = 1$ is the maximum first derivative of utility functions. The efficiency of energy transfer is set to be $\eta_{1,4} = \eta_{2,5} = 0.2$, $\eta_{4,5} = \eta_{4,6} = 0.3$, $\eta_{3,6} = \eta_{5,7} = 0.4$. We also specify $\varepsilon_{\max} = 10$ J/slot, $R_{\max} = 3$ packet/slot, $P_{\max} = 10$ J/slot and $\delta = 2$ packet/J. The control parameter V takes value in {10, 20, 40, 60, 80, 100, 200}. Each simulation point represents the average value of 10^5 independent simulation runs.

Figure 2 demonstrates the utility performance of the proposed EDPR algorithm under various values of the control parameter V. At the very beginning, the utility function is zero when V is relatively small and no data transmission is underway. As V increases, the time-averaged utility of the total network quickly converges to optimal value, about 2.27 packets/slot in this scenario.

In Figure 3, we plot the average data queue bounds and energy queue bounds with regard to different V. It is intuitive that a larger V brings about more queue congestion. Both of them grow near linear with V, which corresponds to Theorem 2. The consequence is quite useful for practical implement because we can easily design the size of energy storage devices for sensors.

In order to further study the dynamic update process feature of queues, we illustrate the renewal process of data stream queue q_1, q_2, q_5, q_6 in Figure 4, and energy queue e_1, e_2, e_5, e_6 in Figure 5, when in the same scenario setting with $V = 100$. It can be seen from the figures that the backlog of queues is all upper bounded by a specific limit, and thus the network can maintain stable. And we also observe that the backlog of source nodes $1, 2$ is larger than those of relay nodes $5, 6$. Additionally, whether data or energy queue, the posterior nodes (such as nodes $5, 6$) gradually tend to stabilization after anterior nodes (such as nodes $1, 2$) become stable.

Figures 6 and 7 respectively show the average backlog of data queue and energy queue against t under several different control parameter V. We can observe from the two figures that the average queue backlog first increases with time slot t, and then fluctuates around a deterministic value eventually. As for influence of V, a larger V contributes to larger capacity of the stable network, whereas a small V leads to earlier arrival of the stable state. This observation verifies that V indeed controls the trade-off between utility performance and network stability. Compared with data queues, energy queues have a larger fluctuation in amplitude, which is reasonable because the nodes in the WSN harvest and transfer energy at the same time slot.

6 Conclusion

In this paper, we focus on joint utility optimization in multi-hop WSNs with energy harvesting and cooperation. An upper bound on the Lyapunov drift is derived for the stability of the network, which can calculate the precise battery capacities in demand for sensor nodes. Then, we propose an online algorithm by combining drift-plus-penalty method and perturbation technique under Lyapunov framework, so as to deal with the stochastic optimization problem by jointly optimizing the energy storage and power

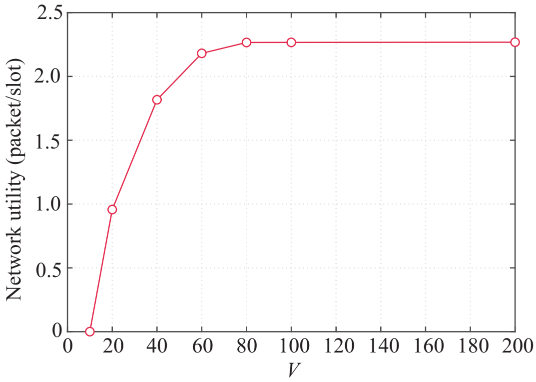

Figure 2 (Color online) Utility of EDPR under various V.

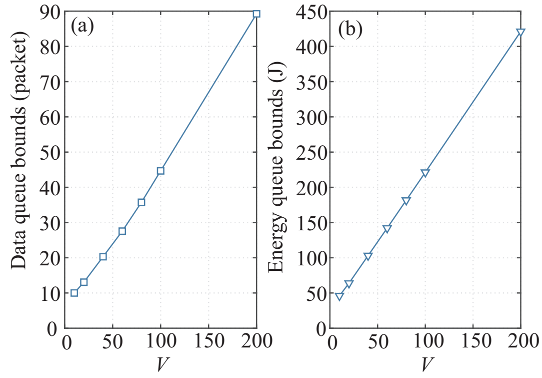

Figure 3 (Color online) (a) Average data queue bounds and (b) energy queue bounds of EDPR under various V.

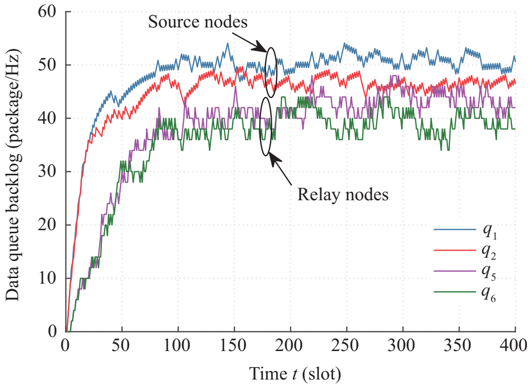

Figure 4 (Color online) Renewal process of data stream queue backlog q_1, q_2, q_5, q_6, when V=100.

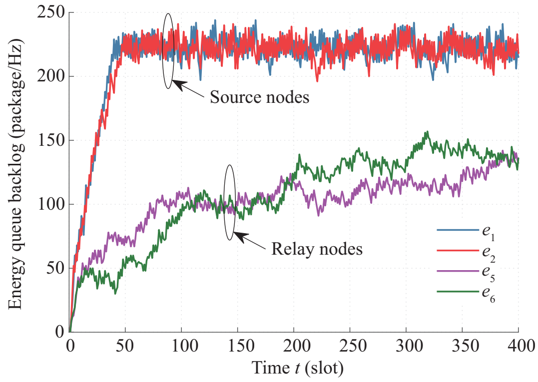

Figure 5 (Color online) Renewal process of energy queue backlog e_1, e_2, e_5, e_6, when V=100.

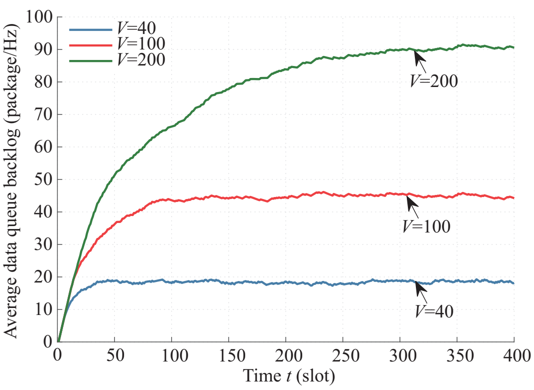

Figure 6 (Color online) Average data queue backlog under different control parameter V.

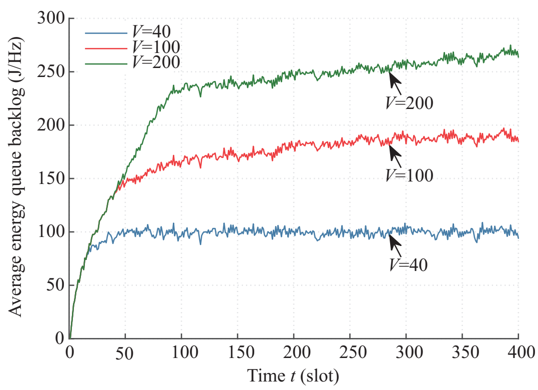

Figure 7 (Color online) Average energy queue backlog under different control parameter V.

allocation policy, which takes time-average and instantaneous stability constraints and power constraints into account. The EDPR algorithm does not require any statistical information about dynamic systems and can achieve the optimal utility in a distributed manner. Furthermore, it contributes to a time-averaged $[\mathcal{O}(V), \mathcal{O}(1/V)]$ trade-off between network utility and queue backlog. In addition, simulation results also reflect the linear relationship between data transmission and queue sizes, which can be very useful in practical implementations.

Acknowledgements This work was supported in part by National Science and Technology Major Project of China (Grant No. 2018ZX03001008-002), in part by Natural Science Foundation of Jiangsu Province (Grant No. BK20180011), and in part by National Natural Science Foundation of China (Grant Nos. 61571120, 61871122).

References

1 Almalkawi I T, Guerrero Zapata M, Al-Karaki J N, et al. Wireless multimedia sensor networks: current trends and future directions. Sensors, 2010, 10: 6662–6717
2 Zhang Y S, He Q, Xiang Y, et al. Low-cost and confidentiality-preserving data acquisition for internet of multimedia things. IEEE Int Thing J, 2018, 5: 3442–3451
3 Zhu B, Xie L H, Han D M, et al. A survey on recent progress in control of swarm systems. Sci China Inf Sci, 2017, 60: 070201
4 Zhang Y, Xiang Y, Zhang L Y, et al. Secure wireless communications based on compressive sensing: a survey. IEEE Commun Surv Tut, 2019, 21: 1093–1111
5 Yousaf R, Ahmad R, Ahmed W, et al. A unified approach of energy and data cooperation in energy harvesting WSNs. Sci China Inf Sci, 2018, 61: 082303
6 Alippi C, Galperti C. An adaptive system for optimal solar energy harvesting in wireless sensor network nodes. IEEE Trans Circ Syst I, 2008, 55: 1742–1750
7 Yang H H, Lee J, Quek T Q S. Heterogeneous cellular network with energy harvesting-based D2D communication. IEEE Trans Wirel Commun, 2016, 15: 1406–1419
8 Dhillon H, Li Y, Nuggehalli P, et al. Fundamentals of heterogeneous cellular networks with energy harvesting. IEEE Trans Wirel Commun, 2014, 13: 2782–2797
9 Vaze R. Transmission capacity of wireless ad hoc networks with energy harvesting nodes. In: Proceedings of Global Conference on Signal and Information Processing (GlobalSIP), Austin, 2013. 353–358
10 Michelusi N, Stamatiou K, Zorzi M. Transmission policies for energy harvesting sensors with time-correlated energy supply. IEEE Trans Commun, 2013, 61: 2988–3001
11 Prabuchandran K J, Meena S K, Bhatnagar S. Q-learning based energy management policies for a single sensor node with finite buffer. IEEE Wirel Commun Lett, 2013, 2: 82–85
12 Hsu R C, Liu C T, Wang H L. A reinforcement learning-based ToD provisioning dynamic power management for sustainable operation of energy harvesting wireless sensor node. IEEE Trans Emerg Top Comput, 2014, 2: 181–191
13 Yu H, Neely M J. Learning aided optimization for energy harvesting devices with outdated state information. In: Proceedings of IEEE Conference on Computer Communications, 2018. 1853–1861
14 Neely M J. Super-fast delay tradeoffs for utility optimal fair scheduling in wireless networks. IEEE J Sel Areas Commun, 2006, 24: 1489–1501
15 Huang L, Moeller S, Neely M J, et al. LIFO-backpressure achieves near-optimal utility-delay tradeoff. IEEE/ACM Trans Netw, 2013, 21: 831–844
16 Huang L. Optimal sleep-wake scheduling for energy harvesting smart mobile devices. IEEE Trans Mobile Comput, 2017, 16: 1394–1407
17 Neely M J. Stochastic network optimization with application to communication and queueing systems. Synth Lect Commun Netw, 2010, 3: 1–211
18 Eryilmaz A, Srikant R. Joint congestion control, routing, and MAC for stability and fairness in wireless networks. IEEE J Sel Areas Commun, 2006, 24: 1514–1524
19 Liu J, Shroff N B, Xia C H, et al. Joint congestion control and routing optimization: an efficient second-order distributed approach. IEEE/ACM Trans Netw, 2016, 24: 1404–1420
20 Le L B, Modiano E, Shroff N B. Optimal control of wireless networks with finite buffers. IEEE/ACM Trans Netw, 2012, 20: 1316–1329
21 Xu W Q, Zhang Y S, Shi Q J, et al. Dynamic optimization for heterogeneous powered wireless multimedia sensor networks with correlated sources and network coding. 2014. ArXiv:1410.5697
22 You C S, Huang K B, Chae H. Energy efficient mobile cloud computing powered by wireless energy transfer. IEEE J Sel Areas Commun, 2016, 34: 1757–1771
23 Bi S Z, Ho C K, Zhang R. Wireless powered communication: opportunities and challenges. IEEE Commun Mag, 2015, 53: 117–125
24 Huang L B, Neely M J. Utility optimal scheduling in energy-harvesting networks. IEEE/ACM Trans Netw, 2013, 21: 1117–1130
25 Gutiérrez M A, Steen K. Stochastic finite element methods. In: Encyclopedia of Computational Mechanics. Hoboken: Wiley, 2018
26 Tapparello C, Simeone O, Rossi M. Dynamic compression-transmission for energy-harvesting multihop networks with correlated sources. IEEE/ACM Trans Netw, 2014, 22: 1729–1741

Appendix A Proof of Theorem 1

The Lyapunov function (15) consists of two parts, the data stream queue and the energy queue. We separately handle these two parts in order to obtain the upper bound of the Lyapunov drift (16).

For one thing, we discuss the energy queue part. By introducing θ_n, squaring both sides of (12) and rearranging it, we then have

$$\begin{aligned}&\frac{1}{2}[e_n(t+1)-\theta_n]^2-\frac{1}{2}[e_n(t)-\theta_n]^2\\&\leqslant\left[\sum_{m\in\mathcal{N}_n^{(\mathrm{out})}}(P_{nm}(t)+\varepsilon_{nm}(t))\right]^2+\left[\sum_{k\in\mathcal{N}_n^{(\mathrm{in})}}\varepsilon_{kn}(t)+h_n(t)\right]^2\\&\quad-[e_n(t)-\theta_n]\left[\sum_{m\in\mathcal{N}_n^{(\mathrm{out})}}P_{nm}(t)-h_n(t)\right]-[e_n(t)-\theta_n]\left[\sum_{m\in\mathcal{N}_n^{(\mathrm{out})}}\varepsilon_{nm}(t)-\sum_{k\in\mathcal{N}_n^{(\mathrm{in})}}\widehat{\varepsilon}_{kn}(t)\right].\end{aligned}\tag{A1}$$

Since $P_{nm}(t)$, $\varepsilon_{nm}(t)$, $h_n(t)$ have upper bounds $P_{\max}$, $\varepsilon_{\max}$, $h_{\max}$ respectively, we can obtain that

$$\left[\sum_{m\in\mathcal{N}_n^{(\mathrm{out})}}(P_{nm}(t)+\varepsilon_{nm}(t))\right]^2+\left[\sum_{k\in\mathcal{N}_n^{(\mathrm{in})}}\widehat{\varepsilon}_{kn}(t)+h_n(t)\right]^2\leqslant[D_{\max}(P_{\max}+\varepsilon_{\max})]^2+[D_{\max}\varepsilon_{\max}+h_{\max}]^2.\tag{A2}$$

Let B_1 denote $[D_{\max}(P_{\max}+\varepsilon_{\max})]^2+[D_{\max}\varepsilon_{\max}+h_{\max}]^2$. Hence, we sum both sides of (A2) over n and it can be recast as

$$\begin{aligned}\frac{1}{2}\sum_{n=1}^N[e_n(t+1)-\theta_n]^2-\frac{1}{2}\sum_{n=1}^N[e_n(t)-\theta_n]^2\leqslant NB_1-\sum_{n=1}^N[e_n(t)-\theta_n]\left[\sum_{m\in\mathcal{N}_n^{(\mathrm{out})}}P_{nm}(t)-h_n(t)\right]\\-\sum_{n=1}^N\sum_{m\in\mathcal{N}_n^{(\mathrm{out})}}\varepsilon_{nm}(t)\,[\widehat{\varepsilon}_n(t)-\eta_{nm}\widehat{\varepsilon}_m(t)].\end{aligned}\tag{A3}$$

We define $\widehat{E}_n(t)=e_n(t)-\theta_n$. And due to the fact that $0\leqslant\eta_{kn}\leqslant1$, we have $(1-\eta_{nm})\widehat{E}_n(t)\geqslant-(1-\eta_{nm})\theta_n$. Let B_2 denote $B_2=NB_1+\sum_{n=1}^N\sum_{m=1}^N(1-\eta_{nm})\varepsilon_{\max}$, and thus (A3) can be rewritten as

$$\begin{aligned}\frac{1}{2}\sum_{n=1}^N[e_n(t+1)-\theta_n]^2-\frac{1}{2}\sum_{n=1}^N[e_n(t)-\theta_n]^2\leqslant B_2-\sum_{n=1}^N\widehat{E}_n(t)\left[\sum_{m\in\mathcal{N}_n^{(\mathrm{out})}}P_{nm}(t)-h_n(t)\right]\\-\sum_{n=1}^N\sum_{m\in\mathcal{N}_n^{(\mathrm{out})}}\eta_{nm}\varepsilon_{nm}(t)\big[\widehat{E}_n(t)-\widehat{E}_m(t)\big].\end{aligned}\tag{A4}$$

For another, the data queue part is studied and proved. Since the inequality $([f(x)]^+)\leqslant f^2(x)$ holds for any $x\in\mathbb{R}$, Eq. (9) can be converted into

$$\begin{aligned}\big[q_n^{(d)}(t+1)\big]^2-\big[q_n^{(d)}(t)\big]^2\leqslant-2q_n^{(d)}(t)\left[\sum_{m\in\mathcal{N}_n^{(\mathrm{out})}}\lambda_{nm}^{(d)}(t)-\sum_{k\in\mathcal{N}_n^{(\mathrm{in})}}\lambda_{kn}^{(d)}(t)-R_n^{(d)}\right]\\+\left[\sum_{m\in\mathcal{N}_n^{(\mathrm{out})}}\lambda_{nm}^{(d)}(t)\right]^2+\left[\sum_{k\in\mathcal{N}_n^{(\mathrm{in})}}\lambda_{kn}^{(d)}(t)+R_n^{(d)}\right]^2.\end{aligned}\tag{A5}$$

Rearrange (A5) similarly to how we deal with that in the energy queue part, and it can be written as

$$\frac{1}{2}\big[q_n^{(d)}(t+1)\big]^2-\frac{1}{2}\big[q_n^{(d)}(t)\big]^2\leqslant B_3-q_n^{(d)}(t)\left[\sum_{m\in\mathcal{N}_n^{(\mathrm{out})}}\lambda_{nm}^{(d)}(t)-\sum_{k\in\mathcal{N}_n^{(\mathrm{in})}}\lambda_{kn}^{(d)}(t)-R_n^{(d)}\right],\tag{A6}$$

where $B_3=\frac{3}{2}D_{\max}^2C_{\max}^2+R_{\max}^2$.

Hence, we sum (A6) over all (n,d), and then obtain the following bound combined with (A3).

$$\begin{aligned}L(\boldsymbol{Q}(t+1))-L(\boldsymbol{Q}(t))\leqslant B-\sum_{n,d}q_n^{(d)}(t)\left[\sum_{m\in\mathcal{N}_n^{(\mathrm{out})}}\lambda_{nm}^{(d)}(t)-\sum_{k\in\mathcal{N}_n^{(\mathrm{in})}}\lambda_{kn}^{(d)}(t)-R_n^{(d)}\right]\\-\sum_{n=1}^N\hat{E}_n(t)\left[\sum_{m\in\mathcal{N}_n^{(\mathrm{out})}}P_{nm}(t)-h_n(t)\right]-\sum_{n=1}^N\sum_{m\in\mathcal{N}_n^{(\mathrm{out})}}\eta_{nm}\varepsilon_{nm}(t)[\hat{E}_n(t)-\hat{E}_m(t)],\end{aligned}\tag{A7}$$

where

$$\begin{aligned}B&=N^2B_3+B_2\\&=N^2\left(\frac{3}{2}D_{\max}^2C_{\max}^2+R_{\max}^2\right)+N\big\{[D_{\max}(P_{\max}+\varepsilon_{\max})]^2+[D_{\max}\varepsilon_{\max}+h_{\max}]^2\big\}+\sum_{n=1}^N\sum_{m=1}^N(1-\eta_{nm})\varepsilon_{\max}.\end{aligned}\tag{A8}$$

According to (16), we take expectations over the network state $\boldsymbol{Q}(t)$ on both sides of (A7), and then obtain (18).

Theorem 1 is thus proved.

Appendix B Proof of Theorem 2

First, we use mathematical induction (MI) to prove the boundary constraint of $q_n^{(d)}(t)$ in (29) and $e_n(t)$ in (30).

As for $q_n^{(d)}(t)$, it is clear to see that Eq. (29) holds for $t=0$, because $q_n^{(d)}(0)=0$ for all $n,d\in\mathcal{N}$. Then, we assume that at time slot t, $q_n^{(d)}(t)$ obeys $q_n^{(d)}(t)\leqslant \beta V+R_{\max}$, $\forall n,d\in\mathcal{N}$. Hence, two cases are considered in the next time slot $t+1$ classified by whether to receive data stream.

- If node n does not receive any data stream from other nodes through links at t, then it comes to $q_n^{(d)}(t+1)\leqslant \beta V+R_{\max}$ referring to (9).
- If node n receives data stream from node k, with $q_k^{(d)}(t)\leqslant \beta V+R_{\max}$, then it comes to

$$q_n^{(d)}(t)\leqslant q_k^{(d)}(t)-\tau\leqslant(\beta V+R_{\max})-(D_{\max}C_{\max}+R_{\max})\leqslant \beta V. \tag{B1}$$

Since the transmitted data cannot exceed $R_{\max}$, we have $q_n^{(d)}(t+1)\leqslant \beta V+R_{\max}$.

- If node n receives data packets from the outside of the network, we have $q_n^{(d)}(t)\leqslant Vf'(0)=\beta V$ and get the same result as the above case.

Hence, $q_n^{(d)}(t+1)\leqslant \beta V+R_{\max}$ holds for any $n,d\in\mathcal{N}$, and thus (29) is proved.

As for $e_n(t)$, since we have made the assumption that there is no energy in the initial state, we can get $e_n(t)=0$ for any $n\in\mathcal{N}$. For node n, we assume $e_n(t)\leqslant\theta_n+h_{\max}+D_{\max}\varepsilon_{\max}\eta_{\max}$ at time slot t.

- If $e_n(t)<\theta_n$, $e_n(t+1)$ holds due to the limit of the input energy, no more than $h_{\max}+D_{\max}\varepsilon_{\max}\eta_{\max}$.
- If $e_n(t)\geqslant\theta_n$, node n does not harvest energy but may receive energy from other nodes. As an example, $e_m(t)\leqslant\theta_m+h_{\max}+D_{\max}\varepsilon_{\max}\eta_{\max}$ and energy is transferred from node m to node n, which implies that

$$\begin{aligned}e_n(t)&<e_m(t)-\theta_m-\frac{\zeta}{\eta_{nm}}+\theta_n\\&\leqslant e_m(t)+\theta_n-\zeta\\&\leqslant(\theta_m+h_{\max}+D_{\max}\varepsilon_{\max}\eta_{\max})+\theta_n-(\theta_{\max}+D_{\max}\varepsilon_{\max}\eta_{\max})\\&\leqslant\theta_n+h_{\max}.\end{aligned}\tag{B2}$$

Since the transferred energy cannot surpass $D_{\max}\varepsilon_{\max}\eta_{\max}$, we then have $e_n(t+1)<\theta_n+h_{\max}+D_{\max}\varepsilon_{\max}\eta_{\max}$.

Hence, $e_n(t+1)<\theta_n+h_{\max}+D_{\max}\varepsilon_{\max}\eta_{\max}$ holds for any $n\in\mathcal{N}$, and thus (30) is proved.

Therefore, the boundary constraint of $q_n^{(d)}(t)$ and $e_n(t)$ are provided in (29) and (30).

Next we prove (31) by contradiction when the data transmission or energy transfer action is in progress, that is $P_{nm}>0$ or $e_{\max}>0$.

In the power control step of the EDPR algorithm, we strive to choose the optimal $\varepsilon^*_{nm}(t)$ and $P^*_{nm}(t)$ to solve the energy-availability problem (28) as

$$\begin{aligned}G(\varepsilon_{nm}(t),P_{nm}(t))=\sum_n\Bigg\{&\sum_{m\in\mathcal{N}_n^{(\text{out})}}\varepsilon_{nm}(t)\big[\eta_{nm}\big(\hat{E}_n(t)-\hat{E}_m(t)\big)-\zeta\big]\\&+\sum_{m\in\mathcal{N}_n^{(\text{out})}}C_{nm}(t)\left[q_n^{(d)}(t)-q_m^{(d)}(t)-\tau\right]+\hat{E}_n(t)\sum_{m\in\mathcal{N}_n^{(\text{out})}}P_{nm}(t)\Bigg\}.\end{aligned}\tag{B3}$$

If there exist the optimal power allocation $\boldsymbol{\varepsilon}^*$ and the transferred energy $\boldsymbol{P}^*$ that maximize (B3), it is sure that they satisfy $P^*_{nm}(t)>0$ and $\varepsilon^*_{nm}(t)>0$ on link (n,m). We now create a new solution $(\boldsymbol{\varepsilon}^*,\boldsymbol{P}^*)$ to the problem, by setting ${P_{nm}}^*(t)=0$, $\forall k\neq n$, $\forall m$ and $\varepsilon_{nm}(t)$, $\forall n,m$ remains the same. Then, we have

$$G(\boldsymbol{P}^*)-G(\boldsymbol{P})=\sum_n\left\{\hat{E}_n(t)\boldsymbol{P}^*+\sum_{m\in\mathcal{N}_n^{(\text{out})}}[C_{nm}(\boldsymbol{P}^*)-C_{nm}(\boldsymbol{P})]\widetilde{W}_{(n,m)}(t)\right\}.\tag{B4}$$

Contrary to Theorem 2, we assume $e_n(t)<P_{\max}+\varepsilon_{\max}$ in such case that node n cannot transfer energy to other nodes with energy transferred weight $W_{(n,m)}<0$. Combined with (22), then $\widehat{E}_n(t)<\delta(\beta V+R_{\max})$. In addition, we observe it from (29) that the backlog of all data queue is no more than $\beta V+R_{\max}$, so the backlog weight of data stream over link (n,m) satisfies $\widetilde{W}_{(n,m)}(t)\leqslant\beta V-D_{\max}C_{\max}$ at any time slot. Hence, the above problem can be simplified into

$$\begin{aligned}G(P^*_{nm}(t))-G(P_{nm}(t))&=C_{nm}P^*_{nm}(t)\ \widetilde{W}_{(n,m)}(t)+\hat{E}_n(t)P^*_{nm}(t)\\&\leqslant(\beta V-D_{\max}C_{\max})\,\delta P^*_{nm}(t)-\delta\,(\beta V+R_{\max})\,P^*_{nm}(t)\\&=\delta P^*_{nm}(t)\,(D_{\max}C_{\max}+R_{\max})\\&<0.\end{aligned}\tag{B5}$$

This implies that $\boldsymbol{P}^*$ is not the optimal solution of the problem (B4) since it conflicts with our assumption. The proof of the optimal transferred energy $\varepsilon^*_{nm}(t)$ can be performed similarly. Hence, $e_n(t)\geqslant P_{\max}+\varepsilon_{\max}$, namely (31), holds whether node n transfers any non-negative energy or allocates any non-zero power over its subsequent nodes. What's more, it turns out that all the power allocation and energy transfer actions are feasible, and the "energy-availability" constraint is indeed redundant.

Therefore, part (a) of Theorem 2 is proved. And the proof of part (b) is similar to the proof of Theorem 5.1 in [26], so we omit it here for brevity.