A Jackson Network Model and Threshold Policy for Joint Optimization of Energy and Delay in Multi-Hop Wireless Networks *

Li Xia [†] and Basem Shihada[‡]

Abstract

This paper studies the joint optimization problem of energy and delay in a multi-hop wireless network. The optimization variables are the transmission rates, which are adjustable according to the packet queueing length in the buffer. The optimization goal is to minimize the energy consumption of energy-critical nodes and the packet transmission delay throughout the network. In this paper, we aim at understanding the well-known decentralized algorithms which are threshold based from a different research angle. By using a simplified network model, we show that we can adopt the semi-open Jackson network model and study this optimization problem in closed form. This simplified network model further allows us to establish some significant optimality properties. We prove that the system performance is monotonic with respect to (w.r.t.) the transmission rate. We also prove that the threshold-type policy is optimal, i.e., when the number of packets in the buffer is larger than a threshold, transmit with the maximal rate (power); otherwise, no transmission. With these optimality properties,

we develop a heuristic algorithm to iteratively find the optimal threshold. Finally, we conduct some simulation experiments to demonstrate the main idea of this paper.

Keywords   Queueing, Markov processes, energy and delay optimization, threshold-type policy

1   Introduction

Energy efficiency is one of the most important requirements of wireless networks since most of the mobile end users are powered by battery (Montemanni et al., 2008; Shakkottai et al., 2003). On the other hand, low latency of data transmission is another important requirement for many real-time applications. Here, the data delay includes only the transmission delay and the queueing delay, the propagation delay and the precessing delay are omitted since they are very small in our scenario. High transmission rate can reduce the end-to-end delay with the cost of consuming more energy. Therefore, it is a fundamental problem that how to schedule the packet transmission to optimize the transmission delay and the energy consumption in a wireless network (Berry and Gallager, 2002; Ren and Meng, 2009).

Ad-hoc is a preferred communication mode for wireless networks when the end users are mobile. The decentralized control scheme is suitable for ad-hoc networks since it has a good scalability and flexibility. In this paper, we study a hybrid network. That is, the network consists of two types of communication nodes, one with adequate energy supply and the other with limited energy. We call it *partial energy-critical* wireless network. This partial energy-critical wireless network exists in the real world. For example, the hybrid ad-hoc network deployed in desert may have some nodes with infrastructure and power supply, and other nodes have neither infrastructure nor power supply (Ben Salem et al., 2006). Another example is that a heterogeneous wireless sensor network may have some sensor nodes whose remaining energy is relatively low (Mhatre et al., 2005). The goal is to optimize the energy consumption of these nodes in order to extend their life times without largely increasing the data delay of the entire network. Fig. 1 is an example of the network topology of a partial energy-critical wireless network.

The energy and delay optimization problem in wireless networks has been studied in the
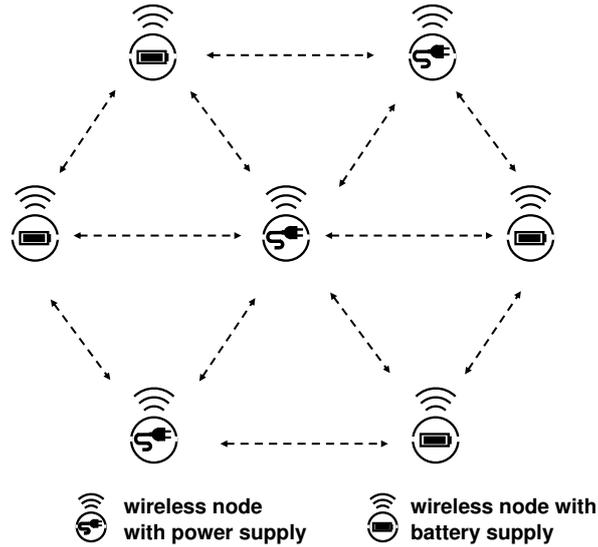
Figure 1: Example topology of a partial energy-critical wireless network.

literature. Some of the studies discuss the optimization of energy consumption with a delay constraint (Chen et al., 2007; Karmokar et al., 2004; Zhong and Xu, 2008). As a comparison, some other studies propose to minimize the delay under a constraint of energy consumption (Goyal et al., 2008; Yang and Ulukus, 2010). However, all of the studies above consider only a single node as the optimization target. In a multi-hop wireless network, communication nodes are interconnected and the dynamics of nodes have mutual influences. Therefore, it is necessary to consider the mutual influence of nodes during the energy and delay optimization (Neely, 2007; Rajan, 2007). Cross-layer optimization is a commonly adopted way to study this problem in a multi-hop wireless network (Cruz and Santhanam, 2003; Cui et al., 2007; Kwon et al., 2006). Xia and Shihada (2014) study the joint optimization of energy and delay in a multi-hop wireless network with the full observation of system buffer status and channel status. However, all these studies are based on a centralized control scheme. The network needs a central controller to acquire the full system information and to determine the node's scheduling policy. Such a centralized control scheme does not have a good scalability and flexibility.

As a comparison, decentralized control is a preferred regime for wireless networks, and it has the advantages on scalability and flexibility (Abouei et al., 1969; Kho et al., 2009). How to develop a decentralized optimization approach under a multi-hop network scenario is an important topic in the research field of wireless networks (Shakkottai et al., 2003).

Moreover, there are some other works studying the rate control and scheduling problem in networking systems from other viewpoints. The work by Xia and Shihada (2013a) studies the optimality properties of service rate control problem in a closed Jackson network. But this paper does not discuss the optimization algorithm to find the optimal policy. The works by Eryilmaz and Srikant (2007); Lin and Shroff (2004) jointly consider the congestion control and the resource scheduling problem. But these studies do not take the energy efficiency as their first optimization target in a multi-hop network scenario.

In this paper, we aim at using a closed-form analysis to study the decentralized optimization problem of energy and delay in multi-hop wireless networks, with the cost of a simplified network model. The transmission rates of nodes are adjustable and the optimization goal is to minimize the average energy consumption and the data delay of the entire network. Every node observes only their own local information including the buffer status and the system reward to determine their own transmission strategy. The network model is simplified under some assumptions. For example, we assume that the channel is stable and the interference is not considered since this issue is not severe after adopting proper physical or data-link layer techniques (such as the synchronization and TDMA).

With proper assumptions, we use a simplified network model, semi-open Jackson network, to formulate this multi-hop wireless network. The associated optimization problem of energy and delay is formulated as a Markov decision process (MDP) (Guo and Hernandez-Lerma, 2009; Puterman, 1994). We utilize the special structure of product-form solution of Jackson networks to study this problem in closed form. Some interesting optimality properties are derived as follows. The system performance is monotonic w.r.t. the transmission rate. The optimal scheduling strategy has a threshold form. That is, when the queue length in the buffer is larger than a threshold, the optimal transmission rate is maximal; otherwise, it is minimal. This indicates that a "bang-bang" control (Artstein, 1980; Bellman et al., 1956) is an optimal control for this problem. The optimality of threshold-type policy greatly reduces the optimization complexity. We further develop a heuristic iterative algorithm to find the optimal threshold-type policy. This algorithm does not require the full system information and it can be implemented in an online manner. This paper is substaintially extended based on our previous conference paper (Xia and Shihada, 2013b) and the new contents include the online estimation algorithm, the detailed proofs of optimality properties, the analysis of

numerical experiments, etc. The limitation of this paper mainly exists in the simplification of the network model. Since our emphasis is the closed-form analysis of the threshold-type policy, some assumptions are required to simplify the network model. If we consider more practical issues, such as the interference and the fading channel, we need more sophisticated investigations in the future.

The remainder of the paper is organized as follows. In Section 2, we give a detailed description of our problem and formulate it with a simplified network model. In Section 3, we utilize the special structure of Jackson network to obtain the performance difference formula of this optimization problem. We also derive some optimality properties in this section. In Section 4, we propose a heuristic and iterative algorithm to distributely optimize the threshold-type policy. The detailed algorithm and its online implementation are also discussed. In Section 5, we discuss some limitations of our simplified network model and the optimization approach. In Section 6, we conduct some numerical experiments to demonstrate the effectiveness of our approach. Finally, we conclude the paper in Section 7.

2 Problem Description and Model Formulation

Consider a partial energy-critical wireless network. The network consists of M communication nodes. The number of energy-critical nodes is denoted as M_1 and the number of energy-noncritical nodes is $M - M_1$. Without loss of generality, we assume that the first M_1 indexed nodes are energy-critical. That is, server i is energy-critical and server j is energy-noncritical, $i = 1, 2, \cdots, M_1$, $j = M_1 + 1, \cdots, M$. We assume that the data packet to be transmitted is independently generated at each node. The data generation process is assumed as a Poisson process and its rate at node i is denoted as λ_i, $i = 1, 2, \cdots, M$. The network usually has different types of data packets (packets for HTTP, FTP, video, audio, control protocols, etc.) and different types of packets have different packet sizes. We assume that the packet size is exponentially distributed with a unit mean. Thus, the transmission time of each packet obeys an exponential distribution. Since the network has multi-hop transmission routes, the generated packet will be relayed among nodes according to a routing protocol. We assume that the routing table is relatively stable and the statistics of the routing traffic has the following distribution. When a packet arrives at node i, this

packet will be relayed to adjacent node j with routing probability q_{ij}, $i, j = 1, 2, \cdots, M$. On the other hand, a packet may arrive at its final destination, say node i, with probability q_{i0} and disappear from the network, $i = 1, 2, \cdots, M$. Obviously, we have $\sum_{j=0}^{M} q_{ij} = 1$ for all $i = 1, 2, \cdots, M$. This assumption of routing probability is reasonable in multi-hop wireless networks, especially for a network adopting the probabilistic forwarding scheme (Chen et al., 2006; Liu and Wu, 2009).

The interference effect of radio channel is a common issue in wireless communication and it has been intensively studied in the literature (Berry and Gallager, 2002; Kwon et al., 2006; Neely, 2007; Rajan, 2007). However, in some literature, interference-free links are studied for simplicity (Cui et al., 2007). It is reasonable for radio channel which is highly synchronized by certain schemes, such as TDMA (Time Division Multiple Access), where each node is assigned a separate time slot to communicate and there is no interference among nodes. In this paper, we also assume that the links are interference-free and the radio channels are statistically stable. All the issues of radio media are handled well by the techniques on the physical or data-link layers. This assumption allows us more feasibility to analytically study this optimization problem in a closed form.

When a node is busy in transmitting packets, newly arriving packets will be queued in the buffer. For simplicity, we assume that the total buffer size of all the nodes in the network has an upper limit denoted as N. That is, when the total number of packets in the network reaches N, any newly generated packets will be rejected by the network instantly. This assumption is reasonable since the buffer is always finite.

As the packet size is exponentially distributed, the transmission time of data packet is also exponentially distributed. The transmission rates of energy-noncritical nodes are fixed and denoted as μ_i, $i = M_1 + 1, \cdots, M$. The transmission rates of energy-critical nodes are varied according to the number of queueing packets in the buffer. We call it *load-dependent* transmission rates and denote it as μ_{i,n_i}, where n_i is the number of packets at node i (including the packet being transmitted), $i = 1, 2, \cdots, M$, $n_i = 0, 1, \cdots, N$. Obviously, $\mu_{i,0} = 0$.

Energy-critical nodes can adjust their transmission rates in an interval $[0, \mu_i^{max}]$, $i = 1, 2, \cdots, M_1$. It is known that the energy consumption in wireless communication heavily

depends on the transmission rate. In this paper, we assume that the transmission power is linear to the transmission rate, which is true for certain situations. For example, in the M-ary orthogonal modulation, it is known that when $\sqrt{E/N_0} \geq 3$, we have

$$P_w = \frac{2N_0}{\log_2 M} \ln \frac{M}{4P_b} \cdot \mu, \tag{1}$$

where P_w is the transmission power, N_0 is the spectrum density of the channel noise after including the channel gain, M denotes the M-ary modulation with a little abuse of notation, P_b is a given bit-error-probability. Therefore, to maintain a certain transmission rate μ_{i,n_i}, a node has an energy consumption rate $b \cdot \mu_{i,n_i}$ per unit time, where b is a coefficient given in (1). In order to save energy, nodes may adopt a lower transmission rate. However, low transmission rates will decrease the throughput of the network and increase the end-to-end delay. We assume that each node has to pay a cost rate c per unit time for each packet in the buffer. Our target is to find the optimal transmission rates of energy-critical nodes, which minimizes the average system cost including the energy consumption of energy-critical nodes and the cost of the packet delay in the entire network. The diagram about the decision logic of an energy-critical node is illustrated by Fig. 2.
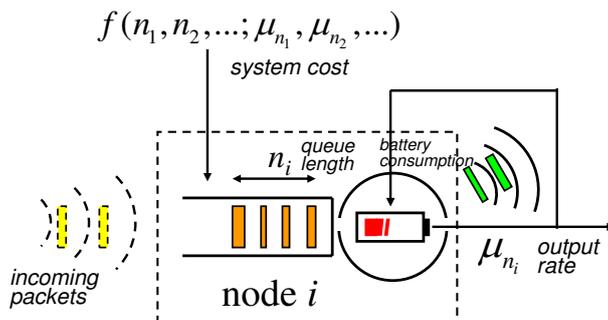

Figure 2: Diagram of decision logic of an energy-critical node.

As we know, queueing model is widely used to formulate and analyze the performance of many computer systems (Gross et al., 2008). In this paper, with the aforementioned problem descriptions and assumptions, we use a semi-open Jackson network model to formulate our problem. We define a virtual node 0 to model the external data generation (see page 21 in Chen and Yao (2001)). The service rate of node 0 is $\mu_0 := \sum_{i=1}^{M} \lambda_i$. The routing probabilities from node 0 to other nodes are $q_{0i} := \lambda_i/\mu_0$, $i = 1, 2, \cdots, M$. The queue length of node 0 is defined as $n_0 := N - \sum_{i=1}^{M} n_i$. Therefore, this multi-hop wireless network can be modeled as

a closed Jackson network with $M+1$ servers and N customers. The system state is denoted as $\boldsymbol{n} := (n_0, n_1, \cdots, n_M)$. The state space is denoted as $\mathcal{S} = \{\text{all } \boldsymbol{n} : \sum_{i=0}^{M} n_i = N\}$. The cost function $\boldsymbol{f}$ is a column vector and its element $f(\boldsymbol{n})$ is defined as follows.

$$f^{\boldsymbol{\mu_n}}(\boldsymbol{n}) = c \sum_{i=1}^{M} n_i + b \sum_{i=1}^{M_1} 1_{n_i>0} \mu_{i,n_i}, \quad \boldsymbol{n} \in \mathcal{S}, \tag{2}$$

where $\boldsymbol{\mu_n}$ is a vector of service rates of all nodes at state $\boldsymbol{n}$, i.e., $\boldsymbol{\mu_n} := (\mu_{1,n_1}, \mu_{2,n_2}, \cdots, \mu_{M,n_M})$, $1_{n_i>0}$ is an indicator function which is defined as $1_{n_i>0} = 1$ when $n_i > 0$, otherwise $1_{n_i>0} = 0$. The first term in the right-hand-side of (2) is the cost rate for the delay of all the packets in the network and the second term is the cost rate for the energy consumption of all the energy-critical nodes based on (1).

As we know, a Jackson network is a continuous time Markov process. The state transition rate matrix of this Markov process is denoted as matrix $\boldsymbol{B}$ and the values of its elements are as follows. For any state $\boldsymbol{n} \in \mathcal{S}$, the value of the row vector $B(\boldsymbol{n},:)$ is determined as: $B(\boldsymbol{n}, \boldsymbol{n}) = -\sum_{i=0}^{M} 1_{n_i>0} \mu_{i,n_i}$; $B(\boldsymbol{n}, \boldsymbol{n}_{-i+j}) = 1_{n_i>0} \mu_{i,n_i} q_{ij}$, where $\boldsymbol{n}_{-i+j}$ is the neighboring state of $\boldsymbol{n}$ and defined as $\boldsymbol{n}_{-i+j} := (n_0, \cdots, n_i - 1, \cdots, n_j + 1, \cdots, n_M)$ for $n_i > 0$, $i, j = 0, 1, \cdots, M$; $B(\boldsymbol{n}, \boldsymbol{n}') = 0$ for all the other states $\boldsymbol{n}'$. Note that for the energy-noncritical nodes and node 0, their service rates μ_{i,n_i} are fixed as μ_i, $i = 0, M_1 + 1, \cdots, M$ for all n_i.

The steady state distribution of the system is denoted as a row vector $\boldsymbol{\pi}$ and its element $\pi(\boldsymbol{n})$ is the probability that the steady system stays at state $\boldsymbol{n}$. The long-run average cost of the system is $\eta = \boldsymbol{\pi} \boldsymbol{f}$. From the structure of $\boldsymbol{B}$, we see that different μ_{i,n_i} will induce different $\boldsymbol{B}$. We define the system policy as a set of transmission rates $\mathcal{L} := \{\mu_{i,n_i}, \text{ for all } i = 1, \cdots, M_1, n_i = 1, \cdots, N\}$. The policy space is defined as $\Psi := \{\text{all } \mathcal{L}\}$. Different policy $\mathcal{L}$ has different $\boldsymbol{\pi}$, different $\boldsymbol{f}$, and different η. The optimization goal is to find the optimal $\mathcal{L}^*$ which minimizes the long-run average cost η.

$$\mathcal{L}^* = \underset{\mathcal{L} \in \Psi}{\arg\min} \{\eta^{\mathcal{L}}\} = \underset{\mathcal{L} \in \Psi}{\arg\min} \{\boldsymbol{\pi}^{\mathcal{L}} \boldsymbol{f}^{\mathcal{L}}\}, \tag{3}$$

where $\eta^{\mathcal{L}}$, $\boldsymbol{\pi}^{\mathcal{L}}$, and $\boldsymbol{f}^{\mathcal{L}}$ are the system average cost, the steady state distribution, and the cost function, when the system adopts policy $\mathcal{L}$, respectively.

We have formulated this problem as an optimization problem above. Although our system is a Markov system, this optimization problem is not a standard MDP. This is because this problem violates a fundamental requirement of MDP that *actions should be*

chosen independently at different states (Puterman, 1994). In our problem, the action is the load-dependent transmission rate μ_{i,n_i}, $i = 1, 2, \cdots, M_1$, $n_i = 1, 2, \cdots, N$. For example, if one node has the same queue length under two system states $\boldsymbol{n}_1$ and $\boldsymbol{n}_2$, such as $\boldsymbol{n}_1 = (1, 3, 2, 4)$ and $\boldsymbol{n}_2 = (1, 3, 4, 2)$ for a network with 3 nodes and finite buffer size $N = 10$, the queue lengths of node 1 at these two states are both 3. Thus, the transmission rates of node 1 at these two states $\boldsymbol{n}_1$ and $\boldsymbol{n}_2$ should be the same as $\mu_{1,3}$. The actions (values of $\mu_{1,3}$) at these two states should be identical and they cannot be chosen independently. Therefore, we cannot use the standard approach of MDP, such as policy iteration or value iteration, to solve this problem. In the next section, we will discuss how to utilize the special structure of Jackson networks to handle this problem.

3   Optimality Structure and Optimization Method

In this section, we utilize the property of product-form solution of Jackson networks to derive a performance difference formula for our problem, which describes the change of system performance when the transmission rates are changed. Based on the difference formula, we study our optimization problem and derive some special properties of the optimal policy.

3.1   Property Based on Product-Form Solution

Based on the product-form solution, we discuss a special property of conditional probability in Jackson networks. This special property can help us derive a difference formula to analyze our optimization problem.

Consider a closed Jackson network with M servers and N customers (Gordon and Newell, 1967; Gross et al., 2008). The service rates are load-dependent. The normalized *visit ratio* v_i is determined by the following traffic equations

$$
\begin{aligned}
&v_i = \sum_{j=1}^{M} v_j q_{ji}, \quad i = 1, 2, \cdots, M; \\
&\sum_{i=1}^{M} v_i = 1.
\end{aligned}
\tag{4}
$$

Define $A_i(0) = 1$ and

$$A_i(k) = \prod_{j=1}^{k} \mu_{i,j}, \quad i = 1, 2, \cdots, M, \quad k = 1, 2, \cdots, N. \tag{5}$$

Further define a normalization constant as below.

$$G_M(N) = \sum_{n_1 + \cdots + n_M = N} \prod_{i=1}^{M} \frac{v_i^{n_i}}{A_i(n_i)}. \tag{6}$$

It is well known that the steady-state distribution of closed Jackson networks has the following product-form solution (Gross et al., 2008).

$$\pi(\boldsymbol{n}) = \frac{1}{G_M(N)} \prod_{i=1}^{M} \frac{v_i^{n_i}}{A_i(n_i)}. \tag{7}$$

Now, we consider the conditional probability of the system state when the queue length of server i is fixed as n_i. This conditional probability is denoted as $\pi(\boldsymbol{n}|n_i)$, where $\boldsymbol{n} \in \mathcal{S}_{n_i}$ and $\mathcal{S}_{n_i}$ is the set of states, where the queue length of server i equals n_i. We have

$$\pi(\boldsymbol{n}|n_i) = \frac{\pi(\boldsymbol{n})}{\pi(n_i)} = \frac{\pi(\boldsymbol{n})}{\sum\limits_{\boldsymbol{n} \in \mathcal{S}_{n_i}} \pi(\boldsymbol{n})}, \tag{8}$$

where $\pi(n_i)$ is the marginal probability that the queue length of server i equals n_i. Substituting (7) into (8), we obtain

$$\pi(\boldsymbol{n}|n_i) = \frac{\prod\limits_{j=1}^{M} \frac{v_j^{n_j}}{A_j(n_j)}}{\sum\limits_{\boldsymbol{n} \in \mathcal{S}_{n_i}} \prod\limits_{j=1}^{M} \frac{v_j^{n_j}}{A_j(n_j)}} = \frac{\frac{v_i^{n_i}}{A_i(n_i)} \prod\limits_{j \neq i}^{M} \frac{v_j^{n_j}}{A_j(n_j)}}{\frac{v_i^{n_i}}{A_i(n_i)} \sum\limits_{\boldsymbol{n} \in \mathcal{S}_{n_i}} \prod\limits_{j \neq i}^{M} \frac{v_j^{n_j}}{A_j(n_j)}} = \frac{\prod\limits_{j \neq i}^{M} \frac{v_j^{n_j}}{A_j(n_j)}}{\sum\limits_{\boldsymbol{n} \in \mathcal{S}_{n_i}} \prod\limits_{j \neq i}^{M} \frac{v_j^{n_j}}{A_j(n_j)}}. \tag{9}$$

From (9) we observe that $\pi(\boldsymbol{n}|n_i)$ has no relation to the value of μ_{i,n_i}. Therefore, for a particular server i, if we change μ_{i,n_i} to μ'_{i,n_i} and fix other servers' service rates, the conditional probability $\pi(\boldsymbol{n}|n_i)$ has the following property

$$\pi(\boldsymbol{n}|n_i) = \pi'(\boldsymbol{n}|n_i), \text{ when } \mu_{i,n_i} \text{ is changed at a particular server } i, n_i = 1, 2, \cdots, N. \tag{10}$$

The above property means that the conditional probability $\pi(\boldsymbol{n}|n_i)$ remains unchanged when only server i's service rate is changed. In the next subsection, we further utilize this property to analyze the system performance of our optimization problem.

3.2 Difference Formula and Optimality Structure

First we give a brief introduction of the direct-comparison theory, which can be viewed as an extension of the theory of perturbation analysis (Cao, 2007; Cao and Chen, 1997; Xia et al., 2009; Xia and Cao, 2012). The performance difference formula is the key part in the direct-comparison theory and it describes the performance difference of a Markov system under any two different sets of system parameters or policies. We apply this theory to analyze the performance change of our optimization problem when the transmission rates are changed.

Consider a continuous-time Markov process $\{X_t, t \geq 0\}$, where X_t is the system state at time t. When the policy of Markov system is changed from $\mathcal{L}$ to $\mathcal{L}'$, the corresponding state transition rate matrix will be changed from $\boldsymbol{B}$ to $\boldsymbol{B}'$ and the cost function will be changed from $\boldsymbol{f}$ to $\boldsymbol{f}'$. We have the following difference formula to quantify the change of the system long-run average performance (Cao and Chen, 1997).

$$\eta' - \eta = \boldsymbol{\pi}'[(\boldsymbol{B}' - \boldsymbol{B})\boldsymbol{g} + (\boldsymbol{f}' - \boldsymbol{f})], \tag{11}$$

where the column vector $\boldsymbol{g}$ is called *performance potential* and its associated element $g(\boldsymbol{n})$, $\boldsymbol{n} \in \mathcal{S}$, is defined as below.

$$g(\boldsymbol{n}) = \lim_{T \to \infty} E \left\{ \int_0^T [f(X_t) - \eta]dt \bigg|_{X_0 = \boldsymbol{n}} \right\}. \tag{12}$$

From (12), we find that $g(\boldsymbol{n})$ can be obtained from the sample path of the current system and it has no relation to the perturbed system under new policy $\mathcal{L}'$.

As we described in Section 2, our system to be optimized is a Markov process. We apply the difference formula (11) to analyze this problem. When the transmission rates of an energy-critical node i are changed from μ_{i,n_i} to μ'_{i,n_i}, $n_i = 1, 2, \cdots, N$, the transition rate matrix $\boldsymbol{B}$ and the cost function $\boldsymbol{f}$ will be changed to $\boldsymbol{B}'$ and $\boldsymbol{f}'$. Substituting the value of $\boldsymbol{B}$ and the cost function (2) into (11), we have

$$
\begin{aligned}
\eta' - \eta &= \sum_{n_i=1}^{N} \sum_{\boldsymbol{n} \in \mathcal{S}_{n_i}} \pi'(\boldsymbol{n}) \left\{ (\mu'_{i,n_i} - \mu_{i,n_i}) \sum_{j=0}^{M} q_{ij}[g(\boldsymbol{n}_{-i+j}) - g(\boldsymbol{n})] + b(\mu'_{i,n_i} - \mu_{i,n_i}) \right\} \\
&= \sum_{n_i=1}^{N} \pi'(n_i) \sum_{\boldsymbol{n} \in \mathcal{S}_{n_i}} \pi'(\boldsymbol{n}|n_i)(\mu'_{i,n_i} - \mu_{i,n_i}) \left\{ \sum_{j=0}^{M} q_{ij}[g(\boldsymbol{n}_{-i+j}) - g(\boldsymbol{n})] + b \right\}. \tag{13}
\end{aligned}
$$

Since we only change the transmission rates of node i and keep other nodes' rates unchanged, we can apply the property (10) to (13) and obtain

$$\eta' - \eta = \sum_{n_i=1}^{N} \pi'(n_i)(\mu'_{i,n_i} - \mu_{i,n_i}) \sum_{\boldsymbol{n} \in \mathcal{S}_{n_i}} \pi(\boldsymbol{n}|n_i) \left\{ \sum_{j=0}^{M} q_{ij}[g(\boldsymbol{n}_{-i+j}) - g(\boldsymbol{n})] + b \right\}. \quad (14)$$

We further define an aggregated parameter $\alpha(i, n_i)$ as follows.

$$\alpha(i, n_i) := \sum_{\boldsymbol{n} \in \mathcal{S}_{n_i}} \pi(\boldsymbol{n}|n_i) \left\{ \sum_{j=0}^{M} q_{ij}[g(\boldsymbol{n}_{-i+j}) - g(\boldsymbol{n})] + b \right\}, \quad n_i = 1, 2, \cdots, N. \quad (15)$$

Difference formula (14) can be rewritten as below.

$$\eta' - \eta = \sum_{n_i=1}^{N} \pi'(n_i)(\mu'_{i,n_i} - \mu_{i,n_i})\alpha(i, n_i). \quad (16)$$

From the definition (15), we see that $\alpha(i, n_i)$ is uniquely determined by the current system and it has no relation to the perturbed system with μ'_{i,n_i}'s. Therefore, we can calculate or estimate the value of $\alpha(i, n_i)$'s based on the sample path of the system with the current parameters μ_{i,n_i}'s. With (16), we observe that $\pi'(n_i)$ is always positive, although we do not know its exact value (it is computationally exhaustive to enumeratively calculate the value of $\pi'(n_i)$ under very possible policies). Therefore, if $\alpha(i, n_i) < 0$, we choose $\mu'_{i,n_i} > \mu_{i,n_i}$ which induces $\eta' < \eta$; otherwise if $\alpha(i, n_i) > 0$, we choose $\mu'_{i,n_i} < \mu_{i,n_i}$ which induces $\eta' < \eta$. With the above scheme, we can find a better setting of parameters which make the system average cost smaller. Repeating this process, we can iteratively reduce the system average cost.

Difference formula (16) gives a very clear picture about the relation of the system average cost and the transmission rates. Based on (16), we can thoroughly study the structure of this optimization problem and obtain some optimality properties as follows.

Theorem 1. *The system average cost η is monotonic w.r.t. the transmission rate μ_{i,n_i}, $i = 1, 2, \cdots, M_1$, $n_i = 1, 2, \cdots, N$.*

Proof. We consider a special scenario, where only one service rate, say $\mu_{i,k}$, is changed to $\mu'_{i,k}$, and all the other service rates μ_{i,n_i} are fixed, $n_i = 1, 2, \cdots, N$ and $n_i \neq k$. Therefore, the difference formula (16) can be rewritten as below.

$$\eta' - \eta = \pi'(k)(\mu'_{i,k} - \mu_{i,k})\alpha(i, k). \quad (17)$$

12

On the other hand, we consider a reverse scenario, where the original value of this transmission rate is $\mu'_{i,k}$ and other transmission rates are μ_{i,n_i}'s, $n_i \neq k$. Suppose $\mu'_{i,k}$ is changed to $\mu_{i,k}$ and all the other rates remain as μ_{i,n_i}'s. Similarly, with (16), we have the following difference formula

$$\eta - \eta' = \pi(k)(\mu_{i,k} - \mu'_{i,k})\alpha'(i,k). \tag{18}$$

Comparing (17) and (18), we have

$$\frac{\alpha'(i,k)}{\alpha(i,k)} = \frac{\pi'(k)}{\pi(k)} > 0, \tag{19}$$

since $\pi'(k)$ and $\pi(k)$ are always positive. Therefore, the sign of $\alpha(i,k)$ remains unvaried if we only change a particular transmission rate $\mu_{i,k}$. From (17), we can obtain the following derivative formula by letting $\mu'_{i,k} \to \mu_{i,k}$

$$\frac{d\eta}{d\mu_{i,k}} = \pi(k)\alpha(i,k). \tag{20}$$

Since the sign of $\alpha(i,k)$ remains unvaried and $\pi(k)$ is always positive, the sign of the derivative $\frac{d\eta}{d\mu_{i,k}}$ also remains unvaried when $\mu_{i,k}$ is changed. Since k is a general symbol and it can be any value from $1, 2, \cdots, N$, we obtain that η is monotonic w.r.t. μ_{i,n_i} for any $i = 1, 2, \cdots, M_1$ and $n_i = 1, 2, \cdots, N$. The theorem is proved. $\qquad\square$

With Theorem 1, we derive the following corollary

Corollary 1. *The optimal transmission rate* μ^*_{i,n_i} *is either* 0 *or* μ^{max}_i, $i = 1, 2, \cdots, M_1$, $n_i = 1, 2, \cdots, N$.

This is a direct conclusion from Theorem 1. The proof is neglected for simplicity.

With Corollary 1, we see that a "bang-bang" control is optimal for this optimization problem (Artstein, 1980; Bellman et al., 1956). We only have to search the optimal value from 0 and μ^{max}_i for every parameter μ_{i,n_i}, without considering the intermediate values between 0 and μ^{max}_i. This greatly reduces the optimization complexity.

With (16) and Theorem 1, we further prove that the optimal transmission rates of this problem has a threshold form, which is described by the following theorem.

Theorem 2. *The optimal transmission rates of our problem have a threshold form. That is, there exists a threshold* θ_i *for each energy-critical node* i, $i = 1, 2, \cdots, M_1$, *which has* $\mu^*_{i,n_i} = 0$, *if* $n_i < \theta_i$; $\mu^*_{i,n_i} = \mu^{max}_i$, *if* $n_i \geq \theta_i$, *where* $\theta_i \in \{1, 2, \cdots, N+1\}$.

Proof. With Theorem 1, we know that the optimal transmission rate can be either 0 or μ_i^{max}. We assume that $\mu_{i,k}^* = 0$ for a particular node i and $n_i = k$. We observe that when the queue length of node i reaches k, the transmission rate of node i becomes 0 and the customers at node i will always accumulate. Thus, n_i will increase to $k + 1$ and it will never drop below k. Since this problem is a Markov process, this phenomena means that the states in $\bigcup_{n_i < k} \mathcal{S}_{n_i}$ are *transient*. Because the transient states have no effect on the long-run average performance of the system, the optimal transmission rates at these states can be chosen as 0. That is, $\mu_{i,n_i}^* = 0$ when $n_i \leq k$. Therefore, we have proved that if there exists $\mu_{i,k}^* = 0$, then $\mu_{i,n_i}^* = 0$ for all $n_i \leq k$. It is straightforward that the optimal transmission rates have a threshold form. $\qquad\square$

Threshold-type policy is a commonly used policy in practice because its form is simple and it is easy to be adopted by practitioners. It is significant to prove the optimality of a threshold-type policy, such as the famous (s, S) policy in the inventory problem and other cases in the literature (Iglehart, 1963; Puterman, 1994; Chen and Yao, 2001; Xia, 2014). In the transmission scheduling problem in wireless networks, the threshold-type policy is also commonly used. However, the optimality of such type policy is rarely proved in the literature. For our formulated problem (3) with the simplified network model, Theorem 2 rigorously guarantees the optimality of the threshold-type policy and significantly reduces the optimization complexity. The original scheduling policy is $\mathcal{L} := \{\mu_{i,n_i}, i = 1, 2, \cdots, M_1, n_i = 1, 2, \cdots, N\}$, where $\mu_{i,n_i} \in [0, \mu_i^{max}]$. It has an NM_1-dimensional continuous search space. With Theorem 2, the policy is simplified as a threshold-type policy which is defined as a vector as below.

$$\boldsymbol{\theta} := (\theta_1, \theta_2, \cdots, \theta_{M_1}). \tag{21}$$

The search space of the threshold-type policy is further denoted as $\boldsymbol{\Theta}$

$$\boldsymbol{\Theta} := \{\text{all } \boldsymbol{\theta} \mid \theta_i = 1, 2, \cdots, N + 1 \text{ and } i = 1, 2, \cdots, M_1\}. \tag{22}$$

Compared with the original NM_1-dimensional continuous search space, the threshold-type policy space $\boldsymbol{\Theta}$ is a finite space with size $(N+1)^{M_1}$. This is a significant reduction. Therefore, the original optimization problem (3) can be further rewritten as the following simplified problem

$$\boldsymbol{\theta}^* = \underset{\boldsymbol{\theta} \in \boldsymbol{\Theta}}{\arg\min}\{\eta^{\boldsymbol{\theta}}\}. \tag{23}$$

Therefore, we can focus our computation resource on the optimization of threshold-type policy, as described by (23). Please note, this threshold-type policy is not a standard one in MDP theory, because it violates the assumption of state-independent action selection, as we discussed in the last part of Section 2. We cannot directly use the classical approach of MDP, such as policy iteration, to solve this problem. However, the direct-comparison theory is still valid to analyze this problem. Below, we discuss the change in the system average cost when the threshold-type policy $\boldsymbol{\theta}$ is changed.

Suppose the threshold of node i is changed from θ_i to θ_i'. Without loss of generality, we assume $\theta_i < \theta_i'$. This change of threshold means that the transmission rates are changed from $\mu_{i,n_i} = \mu_i^{max}$ to $\mu_{i,n_i}' = 0$, for $n_i = \theta_i, \cdots, \theta_i' - 1$. Therefore, the difference formula (16) can be rewritten as follows.

$$\eta' - \eta = \sum_{n_i=\theta_i}^{\theta_i'-1} \pi'(n_i)(0 - \mu_i^{max})\alpha(i, n_i). \tag{24}$$

The system with threshold θ_i' has two types of states, the transient states $\mathcal{S}_{n_i < \theta_i'-1}$ and the positive recurrent states $\mathcal{S}_{n_i \geq \theta_i'-1}$. Therefore, the marginal probability $\pi'(n_i)$ has the following property, $\pi'(n_i) = 0$, for $n_i = 0, \cdots, \theta_i' - 2$; $\pi'(n_i) > 0$, for $n_i = \theta_i' - 1, \cdots, N$. The difference formula (24) is rewritten as

$$\eta' - \eta = -\pi'(\theta_i' - 1)\mu_i^{max}\alpha(i, \theta_i' - 1), \tag{25}$$

where the threshold of a particular server i is changed from θ_i to θ_i', $\theta_i < \theta_i'$, and $\theta_i, \theta_i' \in \{1, 2, \cdots, N+1\}$.

Comparing difference formulas (16), (24), and (25), we observe that (25) is the simplest one. When the threshold of a server, say server i, is increased from θ_i to θ_i', the performance difference only involves three terms, $\pi'(\theta_i' - 1)$ which is always positive, μ_i^{max} which is a known parameter, and $\alpha(i, \theta_i' - 1)$ which can be calculated or estimated based on the sample path of the system under the current parameters μ_{i,n_i}'s. There is no need to consider the terms $\pi'(n_i)$ and $\alpha(i, n_i)$, $n_i = \theta_i, \cdots, \theta_i' - 2$, which appear in (24). Since $\pi'(\theta_i' - 1)$ is always positive, we do not calculate or estimate its exact value. We only calculate or estimate the value of $\alpha(i, \theta_i' - 1)$. Then, we can compare the performance of the new system as follows. If $\alpha(i, \theta_i' - 1)$ is positive, $\eta' < \eta$ and the system average cost is reduced by increasing the

threshold to θ'_i. Otherwise, $\eta' > \eta$ and the new system is worse than the current system and we have to try other thresholds θ'_i. This is the key idea to develop an iterative algorithm for our optimization problem (23). More details will be discussed in the next section.

4   Decentralized Algorithm and Online Implementation

In a network scenario, the decentralized control is preferred to the centralized control since it has a better scalability and flexibility. After we derive the concise difference formula (25), we further discuss how to develop an optimization algorithm which can be distributed on every node to find their optimal thresholds. The difference formula (25) clearly quantifies how the system performance will be changed when the threshold-type policy is changed. Based on (25), we directly develop the following algorithm to find the optimal threshold for a particular energy-critical node.

Algorithm 1. An iterative algorithm to find the optimal threshold of node i.

Step 1. Initialization

- Choose the initial threshold as $\theta^{(0)} = 1$, set $\Theta = \{2, \cdots, N + 1\}$ and $k = 0$.

Step 2. Evaluation

- For the current policy with threshold $\theta^{(k)}$, calculate or estimate the aggregated performance potential $\alpha(i, n_i)$ based on its definition (15), $n_i = \theta^{(k)}, \cdots, N$.

Step 3. Reduction

- If $\alpha(i, j) \leq 0$, remove the element $j + 1$ from the set Θ, where $j = \theta^{(k)}, \cdots, N$.

Step 4. Stopping Rule

- If $\Theta = \varnothing$, set $\theta^* = \theta^{(k)}$ and stop;

- Otherwise, choose the new threshold as $\theta^{(k+1)} = \min\{\Theta\}$ and remove $\theta^{(k+1)}$ from the set Θ, set $k := k + 1$ and go to Step 2.

In Step 3, the thresholds with a worse system performance are identified based on (25) and removed from the feasible search space Θ. Every element of the refined set Θ has a dominant performance over that of the current threshold $\theta^{(k)}$. In Step 4, we choose the minimal element of Θ as the new threshold. Thus, the system average cost will be reduced through the update in Step 4. The system performance will be strictly improved after each iteration. Since the size of the search space Θ is finite (the maximal size of the space is $N+1$), Algorithm 1 can converge to the optimal threshold after a finite number of iterations.

From the above discussion, we can observe that Algorithm 1 is very similar to the policy iteration in the classical MDP theory. That is, it can decompose the combination optimization problem (23) into a series of subproblems in Step 2 and 3, which can great improve the algorithm efficiency.

Remark 1. We have two tiers of complexity reductions for our problem. The first one is the optimality of threshold-type policy which simplifies the original problem (3) to (23). The second one is the policy iteration type algorithm which decomposes the combination problem into a series of subproblems.

With Algorithm 1, we can find the optimal threshold for a particular energy-critical node i. However, the optimization problem (23) is to find the optimal thresholds for all the energy-critical nodes i, $i = 1, \cdots, M_1$. We have to consider how to distribute Algorithm 1 to all the energy-critical nodes. Thus, we propose a heuristic scheme as follows.

Algorithm 2. A heuristic scheme to distribute Algorithm 1 on every energy-critical node.

Step 1. Initialization

- Choose the initial policy as $\theta_i^{*(0)} = 1$, for all $i = 1, \cdots, M_1$. Set $k = 0$.

Step 2. Rotation Optimization

- Set $i = 1$.

- For a particular node i, set the threshold of other nodes j as: $\theta_j^{*(k+1)}$ for $j = 1, \cdots, i-1$ and $\theta_j^{*(k)}$ for $j = i+1, \cdots, M_1$.

- Run Algorithm 1 for node i and obtain its optimal threshold as $\theta_i^{*(k+1)}$.

- If $i < M_1$, set $i := i + 1$ and go to Step 2.2; Otherwise, go to Step 3.

Step 3. Stopping Rule

- If $\theta_i^{*(k+1)} = \theta_i^{*(k)}$, for all $i = 1, \cdots, M_1$, stop and output the current thresholds $\theta_i^{*(k)}$, $i = 1, \cdots, M_1$, as the final results;

- Otherwise, set $k := k + 1$ and go to Step 2.

Algorithm 2 describes a rotation procedure to distribute Algorithm 1 among all the energy-critical nodes. The performance of the total network will be repeatedly improved at every iteration. When the rotation procedure stops, the system performance cannot be further improved under the current optimization framework.

From the description of Algorithm 2, we can see that this rotation optimization process looks like a scenario in game theory. Each node adjusts their own strategy (threshold) by observing and analyzing the system behavior and the payoff obtained. The decisions of nodes have mutual effects. Such mutual effects make the nodes dynamically adjust their strategies, according to other nodes' response. When Algorithm 2 stops, we can verify that the algorithm finds a stable solution, which is called *Nash equilibrium* in game theory. Thus, we directly have the following theorem.

Theorem 3. *The heuristic and decentralized Algorithm 2 can converge to the Nash equilibrium of our optimization problem.*

This theorem is very straightforward as we observe the fact that Algorithm 1 can strictly improve the system performance. When Algorithm 2 stops, according to the stopping rule in

Step 3, we know that every node has no incentive to change their strategy since their optimal threshold remains unvaried by using Algorithm 1. This exactly satisfies the definition of Nash equilibrium in game theory. Therefore, our algorithm provides an effective way to find the equilibrium of this decentralized optimization problem. Compared with a standard game theoretic problem, every node in our problem has the same payoff function (2), which is also the social welfare function of the entire network. Therefore, it is easy to verify that the equilibrium is also a *local optimum* of our optimization problem. For some special cases, such as the convex problem, this local optimum obtained by Algorithm 2 is also the global optimum.

Remark 2. The Nash equilibrium and the local optimum are unified in our optimization problem since the payoff function of each node is identical to the social welfare function. This is a special scenario in game theoretic model.

To perform Algorithm 1, we have to know the values of $\alpha(i, n_i)$ in Step 2, where $n_i = \theta^{(k)}, \cdots, N$. Obviously, $\alpha(i, n_i)$ can be numerically calculated according to its definition (15). Below, we discuss how to online estimate $\alpha(i, n_i)$ from the system sample path. From (15), we observe that we have to know the value of $\pi(\boldsymbol{n}|n_i)$ and $g(\boldsymbol{n})$ in order to calculate $\alpha(i, n_i)$. $\pi(\boldsymbol{n}|n_i)$ and $g(\boldsymbol{n})$ only depend on the current system and can be estimated based on the sample path of the current system. First, we introduce how to estimate the difference of potentials of any two states, $g(\boldsymbol{n}) - g(\boldsymbol{n}')$, $\boldsymbol{n}, \boldsymbol{n}' \in \mathcal{S}$. With the definition of performance potential (12), $g(\boldsymbol{n}) - g(\boldsymbol{n}')$ can be written as follows (Cao, 2007).

$$g(\boldsymbol{n}) - g(\boldsymbol{n}') = E\left\{\int_0^{T(\boldsymbol{n}',0)} [f(X_t) - \eta]dt \bigg|_{X_t=\boldsymbol{n}}\right\}, \tag{26}$$

where X_t is the system state at time t and $T(\boldsymbol{n}', 0)$ is the time when the system state X_t first reaches $\boldsymbol{n}$. $T(\boldsymbol{n}', t_0)$ is further defined as $T(\boldsymbol{n}', t_0) := \min\{t : X_t = \boldsymbol{n}', t > t_0\}$, which indicates the time that X_t first reaches $\boldsymbol{n}'$ after t_0. Define $r_0^{\boldsymbol{n}} := 0$ and $r_k^{\boldsymbol{n}} := \min\{t : X_t = \boldsymbol{n}, t > r_{k-1}^{\boldsymbol{n}}\}$, $k = 1, 2, \cdots$. We know that $r_k^{\boldsymbol{n}}$ is called the regenerative point of state $\boldsymbol{n}$ in a Markov process. For an ergodic Markov process, (26) can be further rewritten as follows.

$$g(\boldsymbol{n}) - g(\boldsymbol{n}') = \lim_{K \to \infty} \frac{1}{K} \sum_{k=1}^{K} \int_{r_k^{\boldsymbol{n}}}^{T(\boldsymbol{n}', r_k^{\boldsymbol{n}})} [f(X_t) - \eta]dt, \tag{27}$$

where the system average performance η can be estimated as

$$\eta = \lim_{T \to \infty} \frac{1}{T} \int_0^T f(X_t)dt. \tag{28}$$

19

Below, we apply (27) to estimate $\alpha(i, n_i)$. (15) can be rewritten as

$$\alpha(i, n_i) = b - \sum_{\boldsymbol{n} \in \mathcal{S}_{n_i}} \pi(\boldsymbol{n}|n_i) \sum_{j=0}^{M} q_{ij}[g(\boldsymbol{n}) - g(\boldsymbol{n}_{-i+j})]. \tag{29}$$

We only have to estimate the second term of the right-hand side of the above equation. Define $T(n_i, i, t_0) := \min\{t : X_{t^-} \in \mathcal{S}_{n_i}, X_t \in \mathcal{S}_{n_i-1}, t > t_0\}$, which indicates the first time after t_0 when node i finishes a packet transmission and makes the queue length of node i decreased to $n_i - 1$. Similarly, define $r_0^{\mathcal{S}_{n_i}} := 0$ and $r_k^{\mathcal{S}_{n_i}} := \min\{t : X_t \in \mathcal{S}_{n_i}, t > r_{k-1}^{\mathcal{S}_{n_i}}\}$, $k = 1, 2, \cdots$, which indicates the kth time when the system state reaches the subset $\mathcal{S}_{n_i}$. To simplify the notations, we denote the time series of $r_k^{\mathcal{S}_{n_i}}$ as t_k, $k = 0, 1, \cdots$. By applying (27), we can estimate $\alpha(i, n_i)$ as follows.

$$\alpha(i, n_i) = b - \lim_{K \to \infty} \frac{1}{K} \sum_{k=1}^{K} \int_{t_k}^{T(n_i, i, t_k)} [f(X_t) - \eta]dt. \tag{30}$$

Therefore, with (28) and (30), we can obtain an estimate of $\alpha(i, n_i)$ based on the sample path of the current system. Integrating the above estimation with Algorithm 2, every energy-critical node can optimize their own scheduling policy (threshold) with a rotation scheme. Each node only needs their own local information (queue length) and the feedback of the system cost to make decision.

From the estimation formula (30), we find that the conditional probability $\pi(\boldsymbol{n}|n_i)$ and the routing probability q_{ij} in (29) disappear. This is because such information can be obtained from the sample path when we estimate $g(\boldsymbol{n}) - g(\boldsymbol{n}_{-i+j})$ with (30). Our Algorithm 2 is online implementable. We can complete the estimation and optimization procedure almost simultaneously, without disturbing the normal operation of the system. The online implementation makes our optimization approach more applicable in practice. For more details about the online estimation of performance potentials and sensitivities, please refer to Cao (2007); Li and Cao (2013).

5   Limitation and Discussion

In this section, we discuss the limitations of our approach, especially the side effects of the assumptions which simplify our network model in Section 2.

The first issue is the assumption of the linear relation between the transmission power and rate. Although this assumption is reasonable as we explained it in (1) when the M-array modulation is adopted, there truly exist other forms of relations which are adopted in the literature. One of the typical forms is that the transmission power is exponential w.r.t. the transmission rate from the theoretical viewpoint of Shannon channel. With the *Shannon-Hartley Theorem*, the relationship between the transmission power and the transmission rate is (Berry and Gallager, 2002)

$$P_w = \frac{\sigma^2}{|\hbar|^2} \left(2^{\mu/B} - 1 \right), \tag{31}$$

where σ^2 is the variance of the channel noise, $|\hbar|^2$ is the channel gain, B is the theoretical channel bandwidth. Therefore, the cost function (2) should be modified as below.

$$f^{\boldsymbol{\mu_n}}(\boldsymbol{n}) = c \sum_{i=1}^{M} n_i + b' \sum_{i=1}^{M_1} 1_{n_i > 0} 2^{\mu_{i,n_i}/B}, \tag{32}$$

where the constant term $-\frac{\sigma^2}{|\hbar|^2}$ in (31) is ignored in the above cost function. With the above cost function, we can also obtain a difference formula which is similar to (14) but with a more complicated form. Other following analysis is similar. However, the monotonicity and the optimality of threshold-type policy might not hold for this cost function. Therefore, our approach might be applicable for different cost functions, but the algorithm complexity may increase significantly in some situations.

The second issue is the physical affecting factors of wireless channel on our simplified network model. One of the primary factors is the interference and the fading of wireless channels. As we explained previously, the issue of channel interference can be handled by the physical and data-link layers of the network. However, if we want to model this optimization problem from a cross-layer viewpoint, it is necessary to consider the interference of channels. This issue will significantly increase the complexity of our optimization problem, because the interference makes the states of neighboring nodes heavily joint. This makes the complexity reduction discussed in Remark 1 infeasible. Therefore, we need further investigations to study this problem as a future research topic. Other factors, such as the buffer limit of individual nodes and the physical constraint of the total limited power, are also important for us to consider during the problem formulation procedure. These will certainly increase the model complexity of the problem. The optimality properties and the complexity reduction techniques proposed in this paper might not be valid for these new formulations.

In summary, our simplified network model has limitations since we made some assumptions and simplifications to formulate this model. However, this simplified model gives us opportunity to thoroughly study this joint optimization problem. The meaningful results obtained in this paper, such as the monotonicity, the optimality of threshold-type policy, and the iterative optimization algorithm, can provide some theoretical insights for the practical and complicated model, even these are obtained based on a simplified model.

6 Numerical Experiment

For the purpose of illustration, we conduct numerical experiments on a small-scale network. Consider a network with 3 nodes. Node 1 and 2 are energy-critical, while node 3 is energy-noncritical. That is, $M = 3$ and $M_1 = 2$. The total buffer size of the network is $N = 6$ (e.g., $N = 6$ may mean that the total buffer size of the network is 6K(kilo) packets and we use 1K packets as a basic unit). The data generation rates at the node are $\lambda_1 = 15$, $\lambda_2 = 20$, $\lambda_3 = 15$, respectively. The routing probabilities of the node are $q_{10} = 0.3$, $q_{12} = 0.5$, $q_{13} = 0.2$, $q_{20} = 0.3$, $q_{21} = 0.3$, $q_{23} = 0.4$, $q_{30} = 0.2$, $q_{31} = 0.4$, $q_{32} = 0.4$. Since the external data generation process can be treated as a virtual node 0, the transmission rate of node 0 is $\mu_0 = \lambda_1 + \lambda_2 + \lambda_3 = 50$ and $q_{01} = 0.3$, $q_{02} = 0.4$, $q_{03} = 0.3$. The transmission rate of node 3 is fixed as $\mu_3 = 100$. The transmission rates of node i are adjustable within $[0, \mu_i^{max}]$ and $\mu_i^{max} = 100$, $i = 1, 2$. The coefficients in cost function (2) are $b = 0.2$ and $c = 10$.

We apply Algorithm 2 to optimize the thresholds of energy-critical nodes. The optimization procedure is illustrated in Fig. 3. During the first round of optimization, Algorithm 2 selects node 1 as the target node and uses Algorithm 1 to find the optimal threshold of node 1. The aggregated performance potentials $\alpha(1, n_1)$, $n_1 = 1, 2, \cdots, 6$, can be obtained under the current system setting. We find that the signs of $\alpha(1, n_1)$ are all negative, which are illustrated by "–" on the first row of the middle table in Fig. 3. Therefore, all the candidate thresholds for $\theta_1 = 2, 3, \cdots, 7$ are removed according to Step 3 of Algorithm 1. These removed candidate thresholds are marked with "×" on the middle table of Fig. 3. Therefore, with the system sample path under the current threshold, we can simultaneously compare 6 candidate solutions with the current solution at the same time. However, in our common sense, we usually have to run the system under every candidate solution to obtain their

comparisons. This is a great improvement which is discussed in Remark 1. This benefit comes from the difference formula (25) which clearly describes the change of the system performance w.r.t. thresholds.
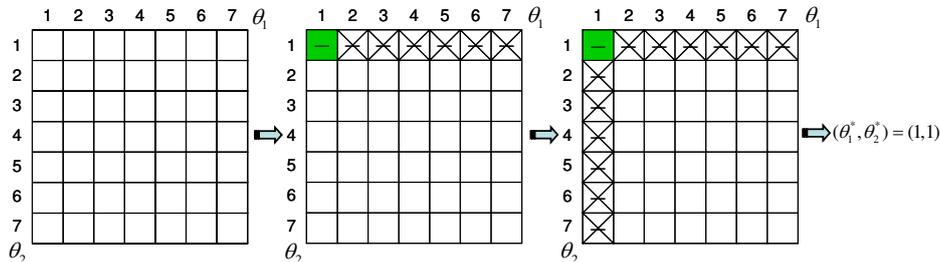

Figure 3: Optimization process of thresholds based on Algorithm 2

Since $\theta_1 = 1$ cannot be improved anymore at the current situation, we fix $\theta_1 = 1$ and optimize the threshold of node 2. This is the second round of the optimization in Algorithm 2. We run the system under the current threshold setting and obtain $\alpha(2, n_2)$, $n_2 = 1, 2, \cdots, 6$. We find that all the signs of $\alpha(2, n_2)$ are negative, which are illustrated by the first column of the right table in Fig. 3. According to Step 3 of Algorithm 1, all these candidate thresholds θ_2 with negative $\alpha(2, \theta_2 - 1)$ can be removed as they have worse performance than that of the current threshold, $\theta_2 = 2, 3, \cdots, 7$. Therefore, $\theta_2 = 1$ is also optimal for node 2 at the current situation. Algorithm 2 stops at the current thresholds since it cannot be improved anymore according to the heuristic scheme. That is, $\boldsymbol{\theta}^* = (1, 1)$ is the output of Algorithm 2.

We conduct some enumerative calculations to verify the main results of this paper. First, we verify the monotonicity described in Theorem 1. We choose arbitrary rates for node 1 and 2, e.g., all the transmission rates of node 1 and 2 are set as 50. Then we enumeratively calculate the system average cost when the transmission rates of node 2 are changed within the value domain $[0, 100]$. The curves of the system average cost w.r.t. different transmission rates of node 2 are illustrated in Fig. 4. From Fig. 4, we see that the average cost η is monotonic w.r.t. the transmission rate $\mu_{2,n}$, $n = 1, 2, \cdots, 6$. Some of the curves are monotonically decreasing and others are monotonically increasing. Therefore, Theorem 1 is verified in this experiment.

Second, we conduct an experiment to verify the optimality of threshold-type policy described in Theorem 2. The transmission rates of node 1 are set as 100. Since the monotonicity
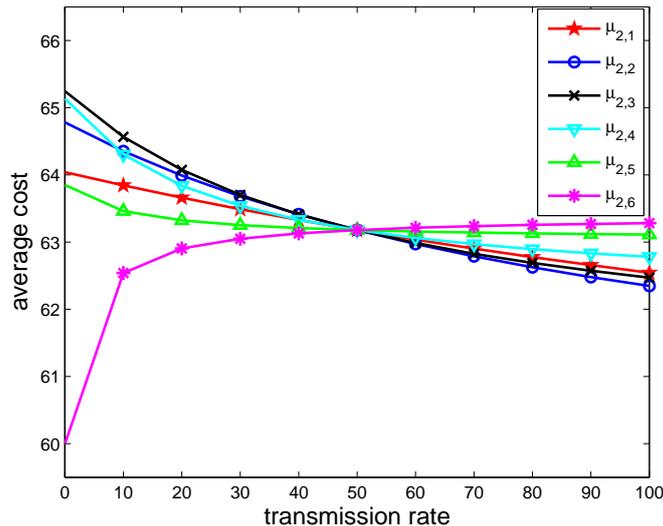
Figure 4: The monotonicity of the average cost w.r.t. the transmission rate

is already verified, we enumerate all the possible rates $\mu_{2,n}$ from the 2-element set $\{0, 100\}$ to do verification, $n = 1, 2, \cdots, 6$. Thus, the total number of enumerations is $2^6 = 64$. We sort the index of enumerations in a binary order of $\mu_{2,n}$'s, i.e., in a order of $(0, 0, 0, 0, 0, 0)$, $(0, 0, 0, 0, 0, 100)$, $\cdots$, $(100, 100, 100, 100, 100, 100)$. Fig. 5 is the enumeration result. We observe that the last index achieves the minimum, i.e., $\mu_{2,n} = 100$ is the optimal rate for all $n = 1, 2, \cdots, 6$. Therefore, the optimal rates of node 2 have a threshold-type policy with $\theta_2 = 1$ and Theorem 2 is verified in this experiment. Please note, many points in Fig. 5 with value 60 indicate a special situation, where the system state is absorbed to $\boldsymbol{n} = (0, 0, 6, 0)$ and all the other states are transient.

To demonstrate the efficiency of our approach, we conduct a comparison experiment as a baseline. The baseline experiment uses brute-force enumeration to search from the NM_1-dimensional continuous space (i.e., 12-dimensional space and each dimension in domain $[0, 100]$). The continuous space needs to be discretized. For example, we can discretize at each integer in $[0, 100]$ and the total number of candidate solutions is $101^{12} \approx 10^{24}$, which is infeasible to enumerate. For simplicity, we utilize Corollary 1 to reduce the complexity and the number of candidate solutions becomes $2^{12} = 4096$. The system performance under every candidate solution is numerically calculated. All the numerical experiments are programmed with *MATLAB (R2009b)* and are executed on a *DELL* computer with 2.92GHz *Intel Core2* CPU, 4GB RAM, 32-bit *Win7* OS. The running-time of our approach with Algorithm 2 is
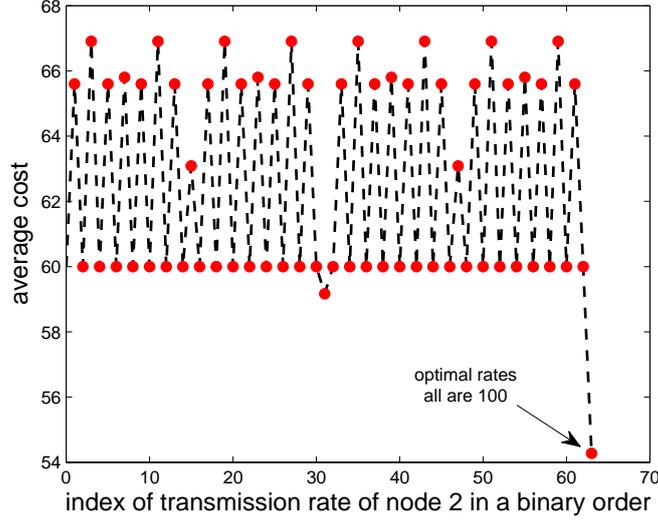
24

Figure 5: The system average cost under all possible min or max rates of node 2

1.217 second, while the running-time of the baseline experiment is 1438.729 second. Therefore, we observe that even with the great simplification of Corollary 1, the computation time of the baseline experiment is still 1000 times more than that of our approach with Algorithm 2. In general, the computation complexity of our approach is $(N + 1)^{M_1}$ in the worst case, which is significantly smaller than that of an NM_1-dimensional continuous search space in the baseline situation.

To verify the effectiveness of our approach, we enumerate the system performance under all of the possible thresholds, $\theta_1, \theta_2 = 1, 2, \cdots, 7$. The curve plane of the system performance under different thresholds is illustrated in Fig. 6. We find that $\boldsymbol{\theta}^* = (1, 1)$ is truly the optimal threshold. Consider the edge of the curve plane. For example, we study the left side of the edge curve depicting η w.r.t. θ_2 when θ_1 is fixed as 1, which is highlighted by the curve of red dashed line in Fig. 6. We find that this curve is an *unimodal* function and $\theta_2 = 1$ is truly optimal. This also demonstrates the effectiveness of Algorithm 1.

It is worth pointing out that Algorithm 2 can only find the local optimum, not the global optimum. The above experiment is a special example, where the output of Algorithm 2 is exactly the global optimum.

From the above experiments, we find that Algorithm 2 has a good searching capability to find the optimal threshold-type policy. Based on the difference formula (25), the algorithm
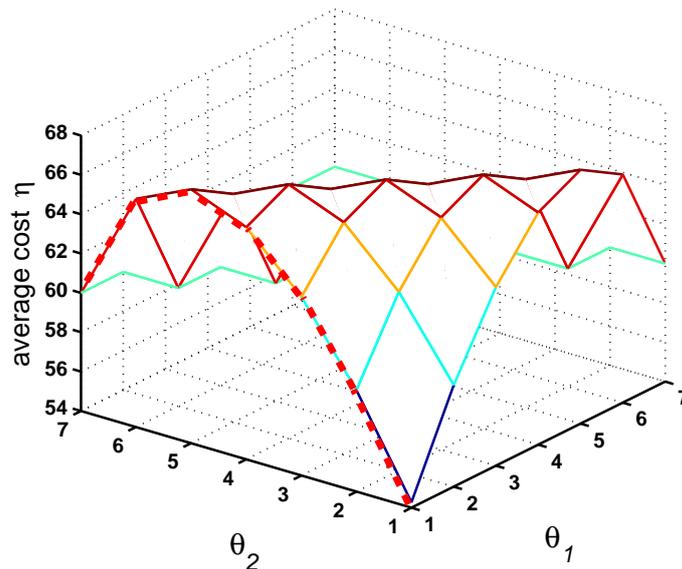
25

Figure 6: The curve plane of the average cost w.r.t. thresholds

can efficiently utilize the performance sensitivity information from the system sample path. We only need to run the system under the current threshold and observe the system information from the sample path. Based on the sample path, we can estimate or calculate $\alpha(i, n_i)$'s and directly find the better solutions which are indicated by the signs of the corresponding $\alpha(i, n_i)$'s. It is commonly observed that the *sign of an estimate* is much more robust than the *exact value of an estimate*. Thus, our estimation or calculation of $\alpha(i, n_i)$'s does not require a very high accuracy. This makes our optimization approach more stable in practice.

The values of parameter b and c have effects on the optimal threshold $\boldsymbol{\theta}^*$. When b is larger, it means that the price of energy is higher and the importance of the power consumption is increased. Therefore, nodes are more willing to save energy, with the cost of having more data delay even losses. Thus, the optimal threshold will become larger. Otherwise, when b is smaller, the optimal threshold will become smaller. That is, the change direction of optimal threshold is the same as that of the parameter b. Please note, the optimal thresholds of different nodes are generally different and they are affected by the value of routing probabilities q_{ij}'s. If the routing probabilities make some nodes' traffic intensity larger (through the traffic equation (4)), the optimal thresholds of these nodes will be smaller than those of other nodes.

26

7 Conclusion

In this paper, we present a decentralized optimization framework for the energy and delay optimization problem in multi-hop wireless networks. We formulate this problem as an MDP and prove that the optimal policy has a threshold form. This optimality lets the algorithm focus on only searching the optimal thresholds, which significantly reduces the optimization complexity. Based on the performance difference formula under any different threshold-type policies, we develop an iterative and decentralized algorithm to find the optimal thresholds. This algorithm is rotatively executed on every node and it will converge to the Nash equilibrium or the local optimum. We further implement this algorithm in an online manner based on the system sample path.

Although our approach provides some insights to handle the decentralized optimization of energy and delay in wireless networks, we take some assumptions to simplify the physical issues, such as the interference of radio channels. These assumptions let us concentrate our analysis on a simplified network model and make it possible to derive some significant results, such as the optimality of threshold-type policy and the iterative optimization algorithm to find the optimal thresholds. It is a future research topic that how to extend this simplified network model to a more complicated case by considering the practical factors, such as the channel interference and the buffer limit of individual node. Moreover, the relation between our decentralized optimization problem and the game theory deserves further investigations, such as the existence of equilibrium and the convergence analysis of the optimization algorithm.

References

Abouei, J., Ebrahimi, M., and Khandani, A. K. (1969). A new decentralized power allocation strategy in single-hop wireless networks. *Proceedings of the 41st Annual Conference on Information Sciences and Systems*, 288-293.

Artstein, Z. (1980). Discrete and continuous bang-bang and facial spaces, or: Look for the extreme points. *SIAM Review* 22, 172-185.

Bellman, R., Glicksberg, I., and Gross, O. (1956). On the bang-bang control problem. *Quarterly of Applied Mathematics* 14, 11-18.

Ben Salem, N., Buttyan, L., Hubaux, J.-P., and Jakobsson, M. (2006). Node cooperation in hybrid ad hoc networks. *IEEE Transactions on Mobile Computing* 5, 365-376.

Berry, R. A. and Gallager, R. G. (2002). Communication over fading channels with delay constraints. *IEEE Transactions on Information Theory* 48, 1135-1149.

Cao, X. R. (2007). *Stochastic Learning and Optimization – A Sensitivity-Based Approach.* New York: Springer.

Cao, X. R. and Chen, H. F. (1997). Perturbation realization, potentials, and sensitivity analysis of Markov processes. *IEEE Transactions on Automatic Control* 42, 1382-1393.

Chen, H. and Yao, D. D. (2001). *Fundamentals of Queueing Networks: Performance, Asymptotics, and Optimization.* New York: Springer Verlag.

Chen, X., Ho, Y. C., Zhang, J., and Bai, H. (2006). Probabilistic Forwarding (ProFor) for large scale sensor networks. *Proceedings of 2006 IEEE International Conference on Networking, Sensing and Control*, April 23-25, 2006, Ft. Lauderdale, Florida, USA, 963-968.

Chen, W., Neely, M. J., and Mitra, U. (2007). Delay-constrained energy-efficient scheduling over a multihop link. *Proceedings of IEEE International Symposium on Information Theory*, 1406-1410.

Cruz, R. and Santhanam, A. (2003). Optimal routing, Link scheduling, and power control in multi-hop wireless networks. *Proceedings of IEEE INFOCOM 2003*, San Francisco, CA, 702-711.

Cui, S., Madan, R., Goldsmith, A. J., and Lall, S. (2007). Cross-layer energy and delay optimization in small-scale sensor networks. *IEEE Transactions on Wireless Communications* 6, 3688-3699.

Eryilmaz, A. and Srikant, R. (2007). Fair resource allocation in wireless networks using queue-length-based scheduling and congestion control. *IEEE/ACM Transactions on Networking* 15, 1333-1344.

Gordon, W. J. and Newell, G. F. (1967). Closed queueing systems with exponential servers. *Operations Research* 15, 252-265.

Goyal, M., Kumar, A., and Sharma, V. (2008). Optimal cross-layer scheduling of transmissions over a fading multiaccess channel. *IEEE Transactions on Information Theory* 54, 3518-3537.

Gross, D., Shortle, J. F., Thompson, J. M., and Harris, C. M. (2008). *Fundamentals of Queueing Theory, 4th Edition.* Hoboken: Wiley.

Guo, X. and Hernandez-Lerma, O. (2009). *Continuous-Time Markov Decision Processes: Theory and Applications.* Springer.

Iglehart, D. L. (1963). Optimality of (s, S) policies in the infinite horizon dynamic inventory problem. *Management Science* 9, 259-267.

Karmokar, A. K., Djonin, D. V., and Bhargava, V. K. (2004). Delay constrained rate and power adaptation over correlated fading channels. *Proceedings of IEEE Globecom 2004*, 3448-3453.

Kho, J., Rogers, A., and Jennings, N. R. (2009). Decentralized control of adaptive sampling in wireless sensor networks. *ACM Transactions on Sensor Networks* 5, 19-35.

Kwon, H., Kim, T. H., Choi, S., and Lee, B. G. (2006). A cross-layer strategy for energy-efficient reliable delivery in wireless sensor networks. *IEEE Transactions on Wireless Communications* 5, 3689-3699.

Li, Y. and Cao, F. (2013). A basic formula for performance gradient estimation of semi-Markov decision processes. *European Journal of Operational Research* 224, 333-339.

Lin, X. and Shroff, N. B. (2004). Joint rate control and scheduling in multihop wireless networks. *Proceedings of the 43rd IEEE Conference on Decision and Control*, 1484-1489.

Liu, C. and Wu, J. (2009). An optimal probabilistic forwarding protocolin delay tolerant networks. *Proceedings of the 10th ACM international symposium on mobile ad hoc networking and computing (MobiHoc'09)*, 105-114.

Montemanni, R., Leggieri, V., and Triki, C. (2008). Mixed integer formulations for the probabilistic minimum energy broadcast problem in wireless networks. *European Journal of Operational Research* 190, 578-585.

Mhatre, V. P., Rosenberg, C., Kofman, D., Mazumdar, R., and Shroff, N. (2005). A minimum cost heterogeneous sensor network with a lifetime constraint. *IEEE Transactions on Mobile Computing* 4, 4-15.

Neely, M. J. (2007). Optimal energy and delay tradeoffs for multi-user wireless downlinks. *IEEE Transactions on Information Theory* 53, 3095-3113.

Puterman, M. L. (1994). *Markov Decision Processes: Discrete Stochastic Dynamic Programming*, New York: John Wiley & Sons.

Rajan, D. (2007). Power efficient delay allocation in multihop wireless networks. *IEEE Transactions on Vehicular Technology* 56, 1813-1825.

Ren, H. and Meng, M. Q.-H. (2009). Game-theoretic modeling of joint topology control and power scheduling for wireless heterogeneous sensor networks. *IEEE Transactions on Autoamtion Science and Engineering* 6, 610-625.

Shakkottai, S., Rappaport, T. S., and Karlsson, P. C. (2003). Cross-layer Design for wireless networks. *IEEE Communications Magazine* 41, 74-80.

Xia, L. (2014). Event-based optimization of admission control in open queueing networks. *Discrete Event Dynamic Systems: Theory and Applications*, DOI 10.1007/s10626-013-0167-1, Vol. 24, No. 2, pp. 133-151, 2014.

Xia, L. and Shihada, B. (2014). Power and delay optimisation in multi-hop wireless networks. *International Journal of Control*, Vol. 87, No. 6, pp. 1252-1265, 2014.

Xia, L. and Cao, X. R. (2012). Performance optimization of queueing systems with perturbation realization. *European Journal of Operational Research* 218, 293-304.

Xia, L., Chen, X., and Cao, X. R. (2009). Policy iteration for customer-average performance optimization of closed queueing systems. *Automatica* 45, 1639-1648.

Xia, L. and Shihada, B. (2013a). Max-Min optimality of service rate control in closed queueing networks. *IEEE Transactions on Automatic Control* 58, 1051-1056.

Xia, L. and Shihada, B. (2013b). Decentralized control of transmission rates in energy-critical wireless networks. *Proceedings of the 2013 American Control Conference*, June 17-19, 2013, Washington, DC., USA, 113-118.

Yang, J. and Ulukus, S. (2010). Delay-minimal transmission for average power constrained multi-access communications. *IEEE Transactions on Wireless Communications* 9, 2754-2767.

Zhong, X. and Xu, C. Z. (2008). On-line energy efficient packet scheduling with delay constraints in wireless networks. *Proceedings of IEEE INFOCOM 2008*, 421-429.