Article

Average Throughput Performance of Myopic Policy in Energy Harvesting Wireless Sensor Networks

Omer Melih Gul * and Mubeccel Demirekler

Department of Electrical and Electronics Engineering, Middle East Technical University (METU),
06531 Cankaya, Ankara, Turkey; demirek@metu.edu.tr
* Correspondence: omgul@metu.edu.tr; Tel.: +90-312-210-4590

Abstract: This paper considers a single-hop wireless sensor network where a fusion center collects data from M energy harvesting wireless sensors. The harvested energy is stored losslessly in an infinite-capacity battery at each sensor. In each time slot, the fusion center schedules K sensors for data transmission over K orthogonal channels. The fusion center does not have direct knowledge on the battery states of sensors, or the statistics of their energy harvesting processes. The fusion center only has information of the outcomes of previous transmission attempts. It is assumed that the sensors are data backlogged, there is no battery leakage and the communication is error-free. An energy harvesting sensor can transmit data to the fusion center whenever being scheduled only if it has enough energy for data transmission. We investigate average throughput of Round-Robin type myopic policy both analytically and numerically under an average reward (throughput) criterion. We show that Round-Robin type myopic policy achieves optimality for some class of energy harvesting processes although it is suboptimal for a broad class of energy harvesting processes.

1. Introduction

1.1. Motivation

The Internet of Things (IoT) is an intelligent large-scale communication infrastructure of uniquely identifiable devices capable of communicating with each other wirelessly through the Internet [1]. The devices in an IoT structure are typically equipped with wireless sensors [2]. Wireless Sensor Networks (WSNs) provide the opportunity of efficient data collection and transmission anywhere [3]. Thus, WSNs have various applications, such as agriculture [4], ambient air monitoring [5,6], frost monitoring [7], structural health monitoring [8–10], remote assistance for elderly people [11], home monitoring [3,11–13] and smart cities [14,15]. Being frugal with energy consumption is important to several WSN deployments. Energy harvesting (EH) [16] can particularly facilitate WSN applications where replacing battery is not practical. Therefore, energy harvesting is a promising approach for the emerging IoT technology [17]. Energy may be harvested from the environment in several different ways (solar, piezoelectric, wind, etc.) [17]. As energy harvesters generally depend on uncontrollable energy resources and the amount of harvested energy is generally low [17,18], WSNs need robust, self-adaptive, energy efficient policies to optimize their reliable operation lifetime [19,20].

In this paper, we consider a fusion center (FC) collecting data from M EH wireless sensors. At each time slot (TS), K sensors are scheduled for data transmission by the FC, which does not have the direct knowledge of the battery states of the sensors or the statistics of their EH processes. It is assumed that the communication is error-free and the sensors are data backlogged but limited in available

energy. Each sensor has an infinite-capacity battery to store the harvested energy and battery leakage is ignored. When a sensor is scheduled in TS t, it sends data to the FC in the TS t as long as it has enough energy in that TS. Sending one packet takes up one TS. The objective of the FC is to maximize the average throughput over a time horizon.

In fact, battery states can be made available to the FC through some additional cost (i.e., feedback) and complexity in some WSNs. However, sending information about the battery state will cause extra time and energy consumption, which we avoid. Assume that the header containing only battery state is H bytes and the remaining part (payload + other headers) of the data packet is P bytes, then sending information about battery state will cause $\frac{H}{P}$ times more time and energy consumption than those consumption not sending no information about battery state. We can avoid extra time consumption by consuming significantly more energy. As it is well known in communication field, data transmission rate is a concave function of transmission power. In fact, the well known Shannon's capacity formula [21] (Shannon's capacity formula is $C = B\log_2(1 + \frac{S}{N})$, where C is the maximum capacity of the channel in bits/second otherwise called Shannon's capacity limit for the given channel, B is the bandwidth of the channel in Hertz, S is the signal power in Watts and N is the noise power, also in Watts. The ratio $\frac{S}{N}$ is called Signal to Noise Ratio (SNR)) indicates that this concave function is a logarithmic function. Therefore, when sending info about battery state, we can avoid the extra time consumption only by consuming much more extra energy than $\frac{H}{P}$. For example, assume that the overhead containing knowledge of battery state is one fourth of the exact data, i.e., $\frac{H}{P} = \frac{1}{4}$. Then, sending both overhead and exact data instead of sending only exact data in the same time duration may cause two times more energy consumption. Thus, it can be said that although energy consumption and network lifetime are not performance metrics in the problem at hand, the problem definition (with sending no information about battery states) helps sensors decrease energy consumption per data packet transmission and thus network lifetime can be increased. Therefore, it is more relevant from a practical perspective that the FC makes scheduling decisions without any knowledge about battery states or statistics of their EH processes [22].

To set up the problem, a model about generation and usage of energy is needed. Each sensor accesses the energy state of its own battery only at the beginning of the time slots in which it is scheduled by the FC. Moreover, independent from the functional form (linear or other) and type of energy harvesting resource (solar, wind, piezoelectric, RF, etc.), the net amount of harvested energy minus used energy is stored losslessly. This assumption is consistent with typical batteries in use today for which leakage is negligibly small over several minutes because battery leakage causes the battery to self-discharge less than 10% (10% for Nickel-based batteries and 5% for Lithium-ion batteries) in 24-h from the results in [23]. Based on these mild assumptions about EH processes, an appropriate performance criterion is the average throughput (reward) criterion over a time horizon rather than expected discounted throughput (reward) for the problem at hand [24].

1.2. Related Work

Although EH processes are not limited to be Markovian in this work, under Markovian assumption, the problem could be formulated as a partially observable Markov decision process (POMDP) [25]. In this case, dynamic programming (DP) [26] may be employed for optimal solution. However, DP has exponential complexity, which limits its scalability [27].

A second approach is reinforcement learning by considering the problem as a POMDP. Q-learning [28], one of the most effective model-free reinforcement learning algorithms, would guarantee convergence to an optimal solution in this problem. However, its very slow convergence [29] deems it non-ideal for a problem with a sizeable state space, especially as the discount factor approaches 1. R-learning [30], which maximizes the average reward, may be considered; however, there is no guarantee on the convergence of R-learning. Therefore, reinforcement learning do not seem to be suitable for obtaining an efficient solution to this problem. There are other approaches that can, in the long run, guarantee convergence to optimal behavior. However, in many practical

applications, a policy that achieves near optimality very quickly is preferable to the one that converges too slowly to exact optimality [29].

Another approach to this problem is to set it up as a restless multi-armed bandit (RMAB) problem. An optimal solution was proposed for RMAB problem under certain assumptions by Whittle [31]. It is shown that finding the optimal solution to a general RMAB problem is PSPACE-hard [32] (In complexity theory, PSPACE is the set of all decision problems that can be solved by a Turing machine using a polynomial amount of space). As a policy with a reasonable complexity, myopic policy (MP) has been suggested for various RMAB problems. While MP is not optimal in general since it focuses only on the present state [33], it can be proven to be optimal in certain special cases.

A very similar problem to the problem at hand is investigated in [34,35]. In fact, we pose the same problem in [34,35] with the exception that we assume infinite capacity battery without leakage at the sensors, in contrast to [34,35] where either no battery or unit capacity batteries with leakage are assumed. Both [34,35] formulate the problem as a POMDP, and, due to the myopic approach in these work, the focus is on the immediate reward instead of future rewards. In [35], a single-hop WSN consisting of EH sensors with unit capacity batteries (i.e., able to store only one transmissions's worth of energy) and a fusion center is posed as a RMAB problem. The optimality of a round-robin (RR) based MP is proved under certain specific assumptions. Then, it is shown that this RR based MP coincides with the Whittle index policy, which is generally suboptimal for RMAB problems [36], for a specific case. In [34], the problem is formulated as a POMDP and the optimality of a MP is proven for two cases: (1) the sensors are unable to harvest and transmit simultaneously, and transition probabilities of the EH processes are affected by the scheduling decisions, and (2) the sensors have no batteries.

In [37–39], we investigate quite a similar problem with the problem at hand, although the problem in [38] has some differences due to its system model. In this paper, we consider more general class of energy harvesting processes than [37,39] do (as it is explained in the rest of this paper, we consider energy harvesting processes with intensities both $\rho \leq 1$ and $\rho > 1$ in this paper, whereas we consider only energy harvesting processes with intensities $\rho \leq 1$ in [37–39]. Besides this, in this paper, we also consider the cases for which finding exact throughput performance of the myopic policy is not possible with using only intensities. For these cases, we find an upper bound for the throughput performance of the myopic policy).

1.3. Our Contributions

Main contributions of the paper are summarized as follows:

- The EH WSN problem is studied under average throughput (reward) criterion and no battery leakage assumption for the most general class of EH processes whereas the problem is studied under discounted throughput (reward) criterion and battery leakage assumption for certain specific cases in [34,35].
- This paper considers a battery capacity (infinite-capacity) larger than unit capacity (which is the maximum battery capacity considered in [34,35]) for the EH WSN problem.
- We show that under average throughput criterion and infinite-capacity battery assumption, RR policies including the MP in [34,35] achieve optimality for some class of EH processes although they are suboptimal for a broad class of EH processes.
- We obtain an upper bound for throughput performance of the RR policies under average throughput criterion for quite general (Markov, i.i.d., nonuniform, uniform, etc.) EH processes. Furthermore, we show that all RR policies including the myopic policy achieve almost the same throughput performance under an average throughput criterion.
- Compared with [34,35], we consider more reasonable finite capacity battery case in the numerical results and show that there is a slight difference in throughput performance between the finite capacity battery case and infinite capacity battery case.

1.4. Organization of the Paper

The rest of this paper is organized as follows. The system model and problem formulation are given in Section 2. In Section 3, we show that RR based MP in [34,35] cannot achieve 100% throughput for a broad class of EH processes under average throughput (reward) criterion. Moreover, we obtain an upper bound for throughput performance of RR policies including the myopic policy under average throughput criterion. Furthermore, we show that RR policies including the myopic policy achieve almost the same throughput as each other. In Section 4, numerical results show that the myopic policy is suboptimal for a broad class of EH processes, which supports the results found in Section 3. Section 5 concludes the paper and provides some future directions.

2. System Model and Problem Formulation

We consider a single-hop WSN where a fusion center (FC) collects data from M EH-capable sensors (please see Figure 1). The index set of all sensors is denoted by $S = \{1, 2, \ldots, M\}$. The WSN operates in a time-slotted fashion indexed as $t = 1, 2, \ldots, T$. At the beginning of each TS, the FC schedules K sensors for data transmission by assigning each sensor to one of its K mutually orthogonal channels. As the research community working on multi-channel protocols generally either assume that channels are perfectly orthogonal (interference-free) or consider the use of only orthogonal channels [40], we assume that the channels are mutually orthogonal, i.e., there is no interference. If the sensors send data at a low data transmission rate and interference management is applied, very low-error transmission can be achieved. Therefore, we assume error-free transmission in the WSN. We assume that the sensors always have data to send as it is assumed in [34,35]. When you consider a single hop wireless sensor network in a wide lowland (flat cropland), there will be no obstacles like buildings, hills which may cause shadowing, reflection, refraction or absorption/diffractions, etc. In a single hop WSN with a central scheduler, the sensors are expected to send the same type of data such as humidity, temperature, pressure, etc. Considering the applications of WSNs in agriculture and frost monitoring [41–45], the sensors have nearly the same propagation conditions to send the same type of data in large croplands. Therefore, we assume that data packets have equal size and sending one packet takes up one TS. A *unit energy* is defined as the energy required for a sensor to send one packet in one TS.
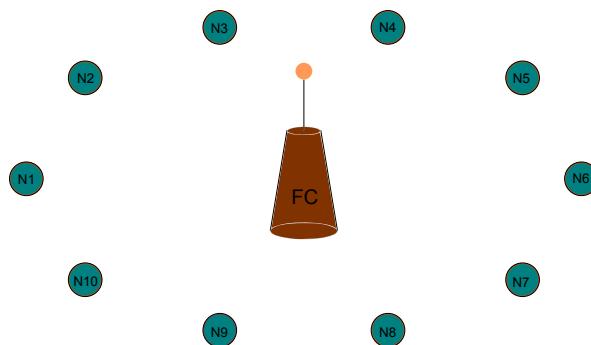
Figure 1. An example single hop WSN where an FC collects data from 10 EH sensors.

The energy harvested by sensor i in TSs 1 through t is denoted by $E_i(t)$, and the energy harvested in TS t is denoted by $E_i^h(t)$; i.e., $E_i^h(t) = E_i(t+1) - E_i(t)$. For $t = 1, \ldots, T$, we define the *activation set*, denoted by $\pi(t)$, as set of the sensors scheduled in TS t under a policy π.

As it is assumed in [34,35], if a sensor has sufficient energy and scheduled in TS t, it sends one data packet to the FC in TS t. The number of data packets sent by sensor i in TS t under a policy π can be written as $\mathbb{I}_{\{i \in \pi(t)\}} \mathbb{I}_{\{B_i^\pi(t) \geq 1\}} \in \{0, 1\}$, where $\mathbb{I}_{\{.\}}$ is the indicator function and $B_i^\pi(t)$ is the stored

energy in infinite-capacity battery of sensor i in TS t under a policy π. Under the policy π, $B_i^\pi(t)$ is evolved as

$$B_i^\pi(t+1) = B_i^\pi(t) + E_i^h(t) - \mathbb{I}_{\{i \in \pi(t)\}} \mathbb{I}_{\{B_i^\pi(t) \geq 1\}} \quad \forall i, t. \tag{1}$$

The number of data packets sent by all sensors to the FC within the first t TSs under a policy π is denoted by $V^\pi(t)$, which is $V^\pi(t) = \sum_{i=1}^{M} V_i^\pi(t)$, where the number of packets sent by sensor i in TSs 1 through t under a policy π is denoted by $V_i^\pi(t) = \sum_{\tau=1}^{t} \mathbb{I}_{\{i \in \pi(\tau)\}} \mathbb{I}_{\{B_i^\pi(\tau) \geq 1\}}$.

In [34,35], the objective is to find a policy that maximizes the total throughput over the time horizon under expected discounted reward criteria, where the discount factor corresponds to battery leakage. (Ref. [34] considers the problem under discounted throughput (reward) criteria since [34] assume battery leakage with discount factor 0.9 such that stored energy decreases to 90% in a time slot which is generally less than 1 ms. However, this is not realistic with recent battery technology.) On the other hand, battery leakage in typical batteries causes less than 10% decrease in the stored energy in 24 h [23]. This decrease implies that battery leakage in a 1 ms-long time slot is less than 0.000000005% and so the discount factor is greater than 0.9999999988, i.e., $0.9999999988 \leq \beta < 1$ (Twenty-four hours equals to 86400000 ms. If length of a time slot is chosen as 1 ms, then $0.90 \leq \beta^{86400000} < 1$ which implies that $0.9999999988 \leq \beta < 1$). Therefore, we neglect battery leakage in our problem formulation, which is practical in terms of engineering aspects. As the problem at hand assumes infinite data backlog and no battery leakage from [19,23], it is delay insensitive by nature. Hence, from [19,23,24], we formulate the scheduling problem as follows.

Problem 1. *Average throughput (reward) maximization over a time horizon, T*

$$\max_{\{\pi(t)\}_{t=1}^{T}} \quad \frac{1}{T} V^\pi(T),$$

$$s.t. \quad Label \text{ (1)}.$$

The following notions are used in the rest of the paper.

Definition 1. *For a given sequence of energy harvests, an **optimal policy**, π^*, is a policy that maximizes the total throughput of all sensors upto KT over a time horizon, T, i.e., $\pi^* \triangleq \arg\max_{\pi \in G} \dfrac{V^\pi(T)}{T}$, where G is set of feasible policies.*

Definition 2. *A **fully efficient policy**, π^{FE}, is a policy under which the sensors use up all of their harvested energy at the end of the time horizon which yields $V^{FE}(T) = \sum_{i=1}^{M} V_i^{FE}(T)$ (Although we use $V^\pi(t)$ to denote the total throughput achieved in first t TSs under a policy π, the total throughput achieved in first t TSs under a policy π^{FE} is denoted by $V^{FE}(T)$ instead of $V^{\pi^{FE}}(T)$ for simplicity. Moreover, $V^*(T)$ and $V^{RR}(T)$ are used instead of $V^{\pi^*}(T)$ and $V^{\pi^{RR}}(T)$, respectively. Similarly, $V_i^{FE}(T)$, $V_i^*(T)$ and $V_i^{RR}(T)$ are used instead of $V_i^{\pi^{FE}}(T)$, $V_i^{\pi^*}(T)$ and $V_i^{\pi^{RR}}(T)$, respectively.), where $V_i^{FE}(T) = \lfloor E_i(T) \rfloor$.*

For certain EH processes, an optimal policy may not be a fully efficient policy, as it is explained in Remark 1.

Definition 3. *Efficiency of a policy π, denoted by $\eta(\pi)$, is defined as the ratio of the throughput of a policy π over the throughput of a fully efficient policy, π^{FE}, over the time horizon, T. It can be expressed as*

$$\eta(\pi) \triangleq \frac{V^\pi(T)}{V^{FE}(T)}, \tag{2}$$

where $V^\pi(T)$ and $V^{FE}(T)$ are the number of collected data packets (throughput) over a time horizon T under a policy π, and fully efficient policy π^{FE}, respectively (When K and T are in order of tens and thousands,

respectively, throughput of an optimal policy is expected to be in the order of ten thousands. The term, efficiency, provides us the opportunity of dealing with small numbers less than or equal to 1 instead of large throughput numbers. Efficiency of a policy also gives us the relative throughput of that policy to the throughput of a fully efficient policy, which provide convenience in numerical results).

The efficiency term itself can also be considered as relative energy consumption of the system to total energy harvested by the system.

The number of data packets which can be sent by all sensors from TS $t+1$ to TS T is denoted by $Y(t)$, i.e.,

$$Y(t) \triangleq V^{FE}(T) - V^*(t). \tag{3}$$

The number of data packets which can be sent by sensor i from TS $t+1$ to TS T is denoted by $Y_i(t)$, i.e.,

$$Y_i(t) \triangleq V_i^{FE}(T) - \max_{\pi^* \in G^*} V_i^*(t), \tag{4}$$

where G^* is the set of all throughput-optimal policies (under different throughput-optimal policies, the throughput of a sensor i in first t TSs may be differed since sensor i may be scheduled by the FC different times under different throughput-optimal policies).

Notice that $Y(0) = \sum_{i=1}^{M} \lfloor E_i(T) \rfloor$ and $Y_i(0) = \lfloor E_i(T) \rfloor$.

Definition 4. *Intensity of sensor* i, ρ_i, *is defined as the integer part of the total energy harvested by sensor i over the time horizon, T, normalized by $\frac{KT}{M}$, i.e.,*

$$\rho_i \triangleq \frac{M \lfloor E_i(T) \rfloor}{KT}.$$

Definition 5. *Intensity*, ρ, *is defined as the sum of integer parts of the total energy harvested by all sensors over the time horizon, T, normalized by KT, i.e.,*

$$\rho \triangleq \frac{\sum_{i=1}^{M} \lfloor E_i(T) \rfloor}{KT} = \frac{\sum_{i=1}^{M} \rho_i}{M}.$$

Remark 1. *If both of the following conditions, $Y_i(t) \leq (T - t) \; \forall i \in S, \; \forall t$ and $Y(t) \leq K(T - t) \; \forall t$, are satisfied, then an optimal policy becomes a fully efficient policy, i.e., $V^*(T) = V^{FE}(T)$. Otherwise, an optimal policy cannot achieve throughput of $V^{FE}(T) = \sum_{i=1}^{M} \lfloor E_i(T) \rfloor$, i.e., $V^*(T) < V^{FE}(T)$. In the cases violating at least one of these conditions, comparing a policy with a fully efficient policy is much simpler than comparing it with an optimal policy. Therefore, we also introduce the notion of fully efficient policy.*

For ease of reference, our commonly used notation is summarized in Table 1.

Table 1. Summary of commonly used symbols and notation.

Symbol	Definition
M	The number of energy harvesting nodes
K	The number of mutually orthogonal channels of FC
S	The index set of all nodes
T	The time horizon
$V^\pi(t)$	Throughput of all nodes in TSs 1 through t under a policy π
$V_i^\pi(t)$	Throughput of node i in TSs 1 through t under a policy π
$\eta(\pi)$	Efficiency of a policy π
$Y_i(t)$	The number of packets which can be sent by node i in $(t, T]$
ρ_i	Intensity of node i
ρ	Intensity

3. Efficiency of Myopic and Round Robin Policies

A similar problem to the problem at hand is studied in [34,35] for certain specific cases under discounted reward criterion. RR based MP is proposed in both papers in which they prove the optimality of this policy for certain specific cases. We applied this RR based myopic policy to the problem at hand. As the MP in [34,35] is an RR policy with quantum = 1 TS, we investigate only RR policies with quantum = 1 TS, denoted by π^{RR}, in this paper.

Definition 6. *For the network that consists of M sensors and an FC with K channels, a **Round Robin (RR) policy with quantum = 1 TS** is an RR policy under which the FC schedules the sensors by allocating one TS to each sensor for data transmission in a period of $\frac{M}{K}$ TSs (Quantum is defined as the number of TSs allocated to each sensor in a period (round) by an RR policy. An RR policy with quantum=n TSs is an RR policy that allocates n TSs to each sensor in a period (round) of $\frac{Mn}{K}$, and so on. For applicability of RR policies with quantum = $n \geq 1$ TS, $\frac{M}{K}$ must be an integer).*

In this section, we show that RR policies with quantum = 1 TS are generally suboptimal by Theorem 1. Next, we study their efficiencies more precisely and obtain an upper bound for their efficiencies by Theorem 2. Then, we show that an RR policy with quantum = 1 TS achieves almost the same efficiency as another RR policy with quantum = 1 TS by Theorem 3, which implies that the MP in [34,35] is generally suboptimal and the upper bound obtained for RR policies with quantum = 1 TS is also valid for the MP in [34,35].

Theorem 1. *If there exist some sensors $i \in S$ such that $Y_i(t) > \left\lceil \frac{K(T-t)}{M} \right\rceil$ for some $t < T$, all RR policies with quantum = 1 TS have efficiency lower than 100% even if fully efficient policy exists for Problem 1 (they are suboptimal).*

Proof. Under an RR policy with quantum = 1 TS, each sensor is visited by the FC either $\left\lfloor \frac{K(T-t)}{M} \right\rfloor$ or $\left\lceil \frac{K(T-t)}{M} \right\rceil$ times if $\frac{K(T-t)}{M}$ is not an integer. If $\frac{K(T-t)}{M}$ is an integer, then FC allocates $\frac{K(T-t)}{M}$ TSs to each sensor. This means that total number of transmissions from any sensor cannot exceed $\left\lceil \frac{K(T-t)}{M} \right\rceil$. $Y_i(t) > \left\lceil \frac{K(T-t)}{M} \right\rceil$ implies that sensor i need to send more than $\left\lceil \frac{K(T-t)}{M} \right\rceil$ data packets in TSs $t+1$ through T so as to send $V_i^{FE}(T) = \lfloor E_i(T) \rfloor$ packets, which must be sent by each sensor for full efficiency. Hence, even if a fully efficient policy exists, any RR policy with quantum = 1 TS is not fully efficient for Problem 1 (they are suboptimal). □

3.1. Efficiency Bounds of RR Policies with Quantum = 1 TS

In this subsection, efficiency bounds of RR policies with quantum = 1 TS are studied precisely for general EH processes.

Lemma 1. *There exists a class of EH processes with intensity ρ_i, such that, for these EH processes, some sensor i transmits lower than $\min \{Y_i(0), q_i\}$ data packets over a time horizon, T, by an RR policy with quantum = 1 TS, where q_i is the number of TSs allocated to sensor i over the time horizon, i.e., $q_i \in \left\{ \left\lfloor \frac{KT}{M} \right\rfloor, \left\lfloor \frac{KT}{M} \right\rfloor + 1 \right\}$.*

Proof. Please see Appendix A. □

Now, we introduce two sets, H_1 and H_2, which are used in Theorem 2, Corollaries 1 and 2. H_1 denotes the index set of the sensors i for which $\left\lfloor \frac{KT}{M} \right\rfloor + 1$ TSs are allocated and $Y_i(0) > \left\lfloor \frac{KT}{M} \right\rfloor + 1$. Moreover, H_2 denotes the index set of the sensors i for which $\left\lfloor \frac{KT}{M} \right\rfloor$ TSs are allocated and $Y_i(0) > \left\lfloor \frac{KT}{M} \right\rfloor$. By definition, $\rho_i > 1$ for sensors $i \in H \triangleq H_1 \cup H_2$.

Theorem 2. *For general EH processes,*

(i) If $\frac{KT}{M} \notin \mathbb{Z}$, efficiency of an RR policy with quantum=1 TS satisfies

$$\eta(\pi^{RR}) \leq 1 - \frac{\sum\limits_{i \in H} (\rho_i - 1) - \frac{M}{KT} |H_1|}{M\rho}.$$

(ii) If $\frac{KT}{M} \in \mathbb{Z}$, efficiency of an RR policy with quantum=1 TS satisfies

$$\eta(\pi^{RR}) \leq 1 - \frac{\sum\limits_{i \in H} (\rho_i - 1)}{M\rho}.$$

Proof. Please see Appendix B. □

From Theorem 2, we derive the following corollaries.

Corollary 1. *For sufficiently large T, we have $M \ll KT$. By the definition, $|H_1| < \sum_{i \in H_1} \rho_i \leq M\rho$ so $\frac{|H_1|}{KT\rho} \ll 1$. Hence, upper bound of efficiency of an RR policy with quantum=1 TS can approximately be written as $1 - \frac{\sum_{i \in H}(\rho_i - 1)}{M\rho}$.*

Corollary 2. *For sufficiently large T, $|H_1| < M\rho \ll KT\rho$. If $\frac{|H_1|}{KT\rho}$ in part (i) of Theorem 2 is neglected and*

(i) If $H = \emptyset$, then $\eta(\pi^{RR}) \leq 1$.
(ii) If $\rho_i = c$, $\forall i \in H$ and $\rho_i = 0$, $\forall i \in S - H$, then $\eta(\pi^{RR}) \leq \min \left\{ 1, \frac{1}{c} \right\}$ where $c \in \mathbb{R}^+$.

3.2. Throughput Difference of RR Policies with Quantum = 1 TS

We will prove that the throughput difference between any two RR policies with quantum = 1 TS cannot be greater than $M - K$ in a time horizon. Recall that, for all RR policies with quantum = 1 TS, the scheduling is periodic with a period of $\frac{M}{K}$ TSs. The only difference between any two RR policies with quantum = 1 TS is their initial scheduling time, t_0; therefore, an RR policy with quantum = 1 TS that starts to send first packet in TS t_0 can be labeled as RR_{t_0}, where $1 \leq t_0 \leq \frac{M}{K}$.

Lemma 2. *The number of transmissions of sensor i can be varied at most one under two different RR policies with quantum = 1 TS over the time horizon T, i.e.,*

$$\max_{1 \leq t_0 \leq \frac{M}{K}} V_i^{RR_{t_0}}(T) - \min_{1 \leq t_0 \leq \frac{M}{K}} V_i^{RR_{t_0}}(T) \leq 1.$$

Proof. Please see Appendix C. □

The following example is given to illustrate Lemma 2.

Example 1. *Assume that $T = 14$, $M = 3$ and $K = 1$. There are three RR policies with quantum=1 TS which schedule the same sensor node i in different TSs as follows.*

RR_1: Sensor i is scheduled in TSs $1, 4, 7, 10, 13,$
RR_2: Sensor i is scheduled in TSs $2, 5, 8, 11, 14,$
RR_3: Sensor i is scheduled in TSs $3, 6, 9, 12.$

Notice that sensor i have the chance of transmission five times under RR_1 and RR_2 while it has a chance of transmission only four times under RR_3. Denote the transmission times of sensor i under different RR policies by t_j^n where $n = 1, 2, 3$ are the indices (initial scheduling times) of RR policies and $j = 1, \ldots, 4$

or 5 are transmission times of sensor i under the applied RR policy in this example. $E_i(t_j^2) \geq E_i(t_j^1)$ for $j = 1, \ldots, 5$ and $E_i(t_{j+1}^2) \geq E_i(t_j^3)$ for $j = 1, \ldots, 4$ and $E_i(t_{j+1}^1) \geq E_i(t_j^3)$ for $j = 1, \ldots, 4$; therefore, $V_i^{RR_2}(14) \geq V_i^{RR_1}(14) = V_i^{RR_1}(13) \geq V_i^{RR_3}(14) = V_i^{RR_3}(12)$. On the other hand, $E_i(t_j^2) \leq E_i(t_j^3)$ for $j = 1, \ldots, 4$; therefore, $V_i^{RR_2}(11) \leq V_i^{RR_3}(12)$. The difference between the most and the least efficient RR policies with quantum = 1 TS are

$$
\begin{aligned}
V_i^{RR_2}(14) - V_i^{RR_3}(14) &= V_i^{RR_2}(14) - V_i^{RR_3}(12) \\
&\leq V_i^{RR_2}(14) - V_i^{RR_2}(11) \\
&\leq 1.
\end{aligned}
$$

Hence, it is observed that throughput of a sensor i can be varied at most 1 under any RR policies with quantum = 1 TS.

The following theorem is based on the extension of Lemma 2 for the whole network.

Theorem 3. *An RR policy with quantum = 1 TS achieves at most $M - K$ more throughput than another RR policy with quantum = 1 TS over the time horizon T, i.e.,*

$$
\max_{\pi^{RR} \in G^{RR}} V^{RR}(T) - \min_{\pi^{RR} \in G^{RR}} V^{RR}(T) \leq M - K,
$$

where G^{RR} is the set of all RR policies with quantum = 1 TS.

Proof. In this proof, we first consider the case of transmitting messages of $\frac{M}{K}$ sensors over a single channel under an RR policy with quantum = 1 TS. Then, we extend the result of Lemma 2 to the case of multiple (K) channels. Notice that

$$
t_0 = mod_{\left(\frac{M}{K}\right)} T \mathbb{I}_{\left\{ mod_{\left(\frac{M}{K}\right)} T \neq 0 \right\}} + \frac{M}{K} \mathbb{I}_{\left\{ mod_{\left(\frac{M}{K}\right)} T = 0 \right\}}
$$

is the best choice regardless of the sensor i and this most efficient RR policy must be applied to one of the $\frac{M}{K}$ sensors over a single channel. When K channels of the FC are considered, this most efficient RR policy must be applied to K of M sensors. By considering this fact and Lemma 2, an RR policy with quantum = 1 TS transmits at most $M - K$ more data packets than another RR policy with quantum = 1 TS. □

Remark 2. *From Theorem 3 and Definition 2,*

$$
\max_{\pi^{RR} \in G^{RR}} \eta(\pi^{RR}) - \min_{\pi^{RR} \in G^{RR}} \eta(\pi^{RR}) \leq \frac{M - K}{V^{FE}(T)} = \frac{M - K}{\rho K T},
$$

where G^{RR} is the set of all RR policies with quantum = 1 TS.

As the MP in [34,35] is also an RR policy with quantum = 1 TS, from Remark 2, it has almost same efficiency as another RR policy with quantum=1 TS for sufficiently large time horizon T, i.e., $\eta(\pi^{MP}) \approx \eta(\pi^{RR})$ if $K, M \ll T$. For sufficiently large time horizons, these results can be extended for RR policies with quantum = $n > 1$ TSs.

4. Numerical Results

In this section, efficiency of the myopic policy (MP) is evaluated for the cases of infinite battery and finite battery with $B = 50$ ($B = 50$ implies that the battery of a sensor can store energy enough to send 50 data packets since we assume that each data packet transmission requires one unit of energy)

at the time horizons varying from 0 to 2000 TSs via simulations (as efficiency of a policy are defined only for the time horizon, T, we obtain efficiency vs. time horizon figures in this section. Notice that efficiency of the MP at $T = 0$ is taken as 0 for these simulations).

For each node i, the Markovian EH process is modelled by a state space $\{0, 1, 2\}$ and a 3×3

transition matrix $\begin{bmatrix} 0.90 & 0.05 & 0.05 \\ 0.05 & 0.90 & 0.05 \\ 0.05 & 0.05 & 0.90 \end{bmatrix}$. The harvested energy for sensor i in TS t is $E_i^h(t) = \frac{K}{M} \times \rho_i \times$

$M_i(t)$ where EH state of sensor i in TS t is denoted by $M_i(t) \in \{0, 1, 2\}$.

Notice that efficiency of the MP in [34,35] is almost the same as efficiency of an RR policy with quantum=1 TS since $\eta(\pi^{RR}) \approx \eta(\pi^{MP})$ for sufficiently large T from Theorem 3 and Remark 2. We observe efficiency of the MP under nonuniform EH processes (it is obvious that the MP achieve efficiencies close to efficiency of an optimal policy for uniform EH processes. As this case is trivial, we did not show the results for this case).

The simulations are made by taking $M = 100$ and $K = 10$ under Markovian and i.i.d EH processes with various intensities which are adjusted by choosing intensities of some sensors as 3.0 and choosing others as 0.3 as explained in Table 2 (In WSNs, it is highly possible that some EH sensors can harvest energy much efficiently than others due to their energy harvesting resource (solar, piezoelectric, RF, wind, etc.) and environmental conditions. For example, solar energy harvesting is generally more efficient than the others. Therefore, we choose intensities of some sensors much larger than the others (3.0 for some sensors and 0.3 for the remaining ones) in order to represent the difference between the amount of energy harvested by sensors).

Table 2. W and L denote the numbers of sensors with $\rho_i = 0.3$ and $\rho_i = 3.0$, respectively. ρ denotes the resultant intensity.

| W	95	85	75	65	55
L	5	15	25	35	45
ρ	0.435	0.705	0.975	1.245	1.515

4.1. Infinite Capacity Battery

In Figure 2, for i.i.d. EH process with $\rho \leq 1$, efficiencies of the MP at $T = 2000$ are 0.758, 0.564 and 0.469 for $\rho = 0.435$, $\rho = 0.705$ and $\rho = 0.975$, respectively. In Figure 3, for Markov EH process with $\rho \leq 1$, at $T = 2000$, the MP achieves efficiency of 0.758, 0.552 and 0.467 for $\rho = 0.435$, $\rho = 0.705$ and $\rho = 0.975$, respectively. The dramatic difference between efficiencies of the MP in these three intensities is expected since Theorem 2 and Corollary 1 state that, as the number of nodes with intensity $\rho_i > 1$ increases, the efficiency of RR policies with quantum = 1 TS decrease. Notice that, from Theorem 2 and Corollary 1, efficiencies of an RR policy with quantum = 1 TS at $T = 2000$ are expected to be $\eta(\pi^{RR}) \leq 0.770$, $\eta(\pi^{RR}) \leq 0.574$ and $\eta(\pi^{RR}) \leq 0.487$ for $\rho = 0.435$, $\rho = 0.705$ and $\rho = 0.975$, respectively. When EH processes have memory (Markov processes), we observe similar results to the results in memoriless (i.i.d.) EH processes with same intensity.

In Figure 2, for i.i.d. EH process with $\rho > 1$, efficiencies of the MP at $T = 2000$ are 0.417 and 0.380 for $\rho = 1.245$ and $\rho = 1.515$, respectively. In Figure 3, for Markov EH process with $\rho > 1$, at $T = 2000$, the MP achieves efficiency of 0.415 and 0.381 for $\rho = 1.245$ and $\rho = 1.515$, respectively. From Corollary 1, efficiencies of an RR policy with quantum = 1 TS at $T = 2000$ are expected to be $\eta(\pi^{RR}) \leq 0.438$ and $\eta(\pi^{RR}) \leq 0.406$ for $\rho = 1.245$ and $\rho = 1.515$, respectively. For EH processes with $\rho > 1$, from Definitions 1–3, efficiency of an optimal policy is $\eta(\pi^*) = \frac{V^*(T)}{V^{FE}(T)} \leq \frac{1}{\rho}$. Hence, at $T = 2000$, $\eta(\pi^*) \leq 0.803$ and $\eta(\pi^*) \leq 0.660$ for $\rho = 1.245$ and $\rho = 1.515$, respectively. Markov EH processes have similar results to the results for i.i.d. EH processes with same intensity.
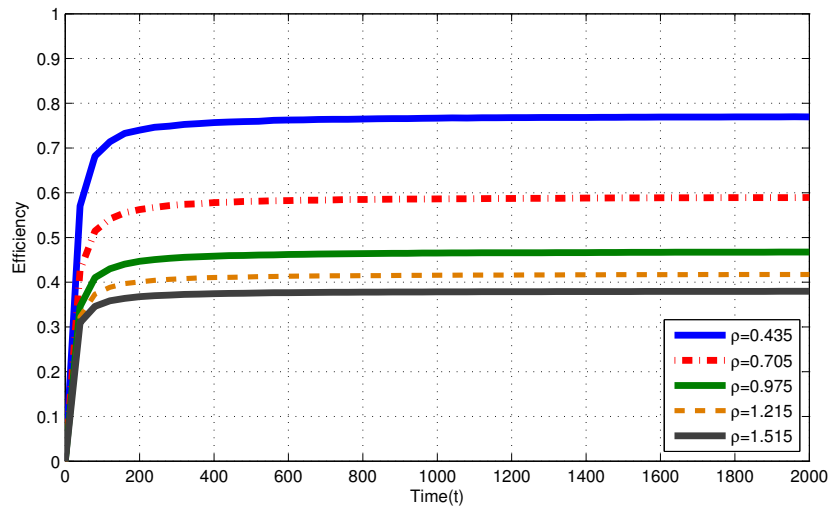
Figure 2. Efficiency of the myopic policy (MP) for i.i.d. EH processes under infinite capacity battery assumption.
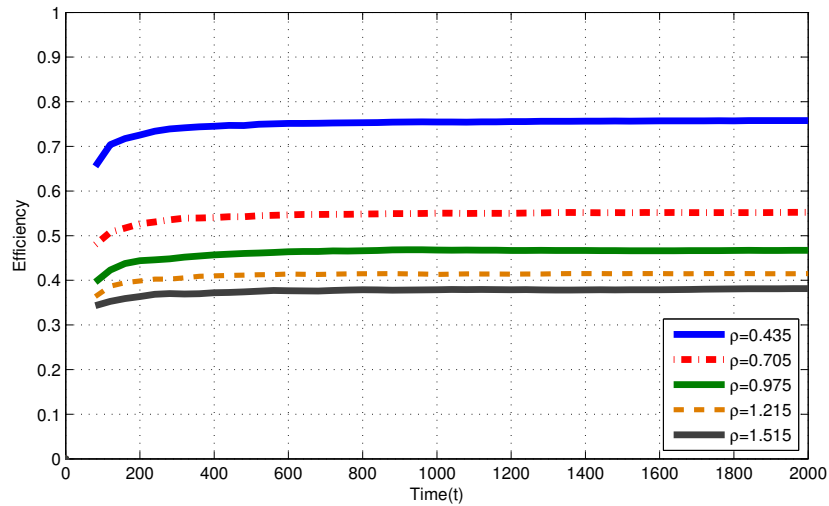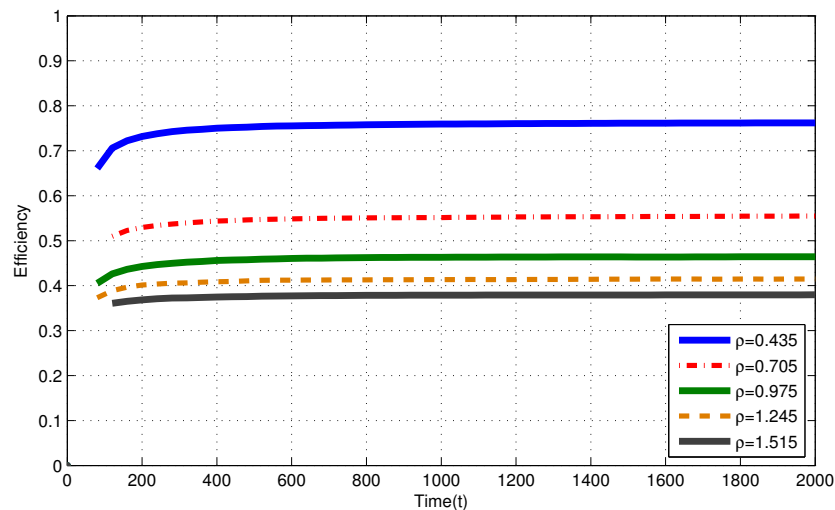

Figure 3. Efficiency of the myopic policy (MP) for Markov EH processes under infinite capacity battery assumption.
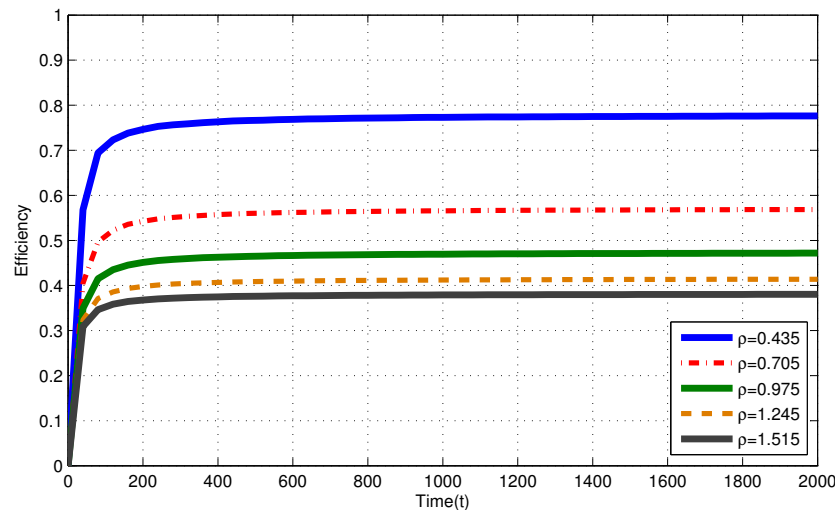
4.2. Finite Capacity Battery

In this subsection, the simulations are made by considering a finite battery capacity with $B = 50$ under Markovian and i.i.d EH processes with the intensities in Table 2.

In Figure 4 (i.i.d. EH process), efficiencies of the MP at $T = 2000$ are 0.757, 0.555, 0.464, 0.415 and 0.380 for $\rho = 0.435$, $\rho = 0.705$, $\rho = 0.975$, $\rho = 1.245$ and $\rho = 1.515$, respectively. In Figure 5 (Markov EH process), at $T = 2000$, the MP achieves efficiency of 0.756, 0.564, 0.462, 0.414 and 0.380 for $\rho = 0.435$, $\rho = 0.705$, $\rho = 0.975$, $\rho = 1.245$ and $\rho = 1.515$, respectively. From Corollary 1, efficiencies of an RR policy with quantum = 1 TS at $T = 2000$ are expected to be $\eta(\pi^{RR}) \leq 0.758$, $\eta(\pi^{RR}) \leq 0.564$, $\eta(\pi^{RR}) \leq 0.469$ $\eta(\pi^{RR}) \leq 0.438$ and $\eta(\pi^{RR}) \leq 0.406$ for $\rho = 0.435$, $\rho = 0.705$, $\rho = 0.975$, $\rho = 1.245$ and $\rho = 1.515$, respectively.

Figure 4. Efficiency of the myopic policy (MP) for Markov EH processes under finite capacity ($B_i = 50$) battery assumption.

Figure 5. Efficiency of the myopic policy (MP) for i.i.d. EH processes finite capacity ($B_i = 50$) battery assumption.

Markov EH processes have similar results to the results for i.i.d. EH processes with the same intensity.

4.3. Discussion

In this subsection, the efficiencies of the myopic policy in both infinite and finite capacity battery cases are compared with each other based on the numerical results in Table 3. Besides this, these numerical results are compared with the expected upper bounds for efficiency of myopic policy. Finally, the complexity of the Round-Robin based myopic policy is investigated.

From Table 3, it can observed that the maximum efficiency difference (0.009) occurs between $B = \infty$ (0.564) and $B = 50$ (0.555) for i.i.d. EH processes with intensity $\rho = 0.705$. For this intensity, the efficiency in finite battery case is 1.596% less than the efficiency in infinite battery case. Besides this, the minimum difference (0.000) occurs between $B = \infty$ (0.380) and $B = 50$ (0.380) for i.i.d. EH processes with $\rho = 1.515$. Therefore, the efficiency of MP for $B = 50$ is only 1.596% less than that for $B = \infty$ at most. Hence, we can conclude that the MP can achieve almost the same throughput performance with a reasonable finite capacity ($B = 50$) battery as that with infinite capacity battery.

Table 3. Efficiency of MP for IID and Markov EH processes under both infinite and finite capacity battery assumptions $B = \infty$ and $B = 50$ stands for infinite and finite capacity batteries, respectively. ρ denotes the intensity. Max. efficiency difference between $B = \infty$ and $B = 50$ represents the efficiency difference between $B = \infty$ and $B = 50$ cases for the same intensity. Max. efficiency difference (%) btw. $B = \infty$ and $B = 50$ represents the percentage of efficiency difference between $B = \infty$ and $B = 50$ cases over the efficiency in $B = \infty$ case for the same intensity. Max. deviation between the bound and efficiency of MP represents the difference between the upper bound of efficiency of MP and minimum efficiency result of MP for the same intensity.

ρ	0.435	0.705	0.975	1.245	1.515
Efficiency of MP for Markov EH process, $B = \infty$	0.758	0.562	0.467	0.415	0.381
Efficiency of MP for Markov EH process, $B = 50$	0.756	0.561	0.462	0.414	0.380
Efficiency of MP for IID EH process, $B = \infty$	0.758	0.564	0.469	0.417	0.380
Efficiency of MP for IID EH process, $B = 50$	0.757	0.555	0.464	0.415	0.380
Max. efficiency difference between $B = \infty$ and $B = 50$	0.002	0.009	0.005	0.002	0.001
Max. efficiency difference (%) btw. $B = \infty$ and $B = 50$	0.264	1.596	1.071	0.480	0.262
Upper bound for efficiency of MP	0.770	0.574	0.487	0.438	0.406
Max. deviation between the bound and efficiency of MP	0.014	0.019	0.025	0.024	0.026

Moreover, from Table 3, it can observed that the maximum deviation (difference) occurs between the upper bound (0.406) and efficiency of MP for i.i.d. EH processes (0.380) with intensity $\rho = 1.515$. For this intensity, the efficiency is 6.40% less than the upper bound. Besides this, the minimum deviation occurs between the upper bound (0.770) and efficiency of MP for i.i.d. EH processes (0.758) with $\rho = 0.435$. For this intensity, the efficiency is only 1.56% less than the upper bound. Based on these results, we can say that the upper bounds for efficiency of the MP are generally tight.

In addition, the Round Robin based myopic policy is a simple policy for this problem. There is an initial ordering and the order is kept during the time interval when Round Robin based scheduling is performed. Sorting algorithms that are required for initial ordering has a worst case complexity of $O(M^2)$ and Round Robin algorithm has a complexity of $O(1)$. Therefore, the myopic policy is a low-complexity solution for the problem at hand.

5. Conclusions

This paper investigates a problem occurring in a single-hop WSN where an FC schedules a set of EH sensors to collect data from them. The FC does not know the instantaneous battery states or the statistics of EH processes at sensors that are data backlogged and the communication is error-free. There is no leakage from the infinite-capacity batteries. The problem at hand is set up as an average throughput (reward) maximization problem. The myopic policy in [34,35] that has an RR structure is applied to this problem as a solution. It is shown that RR policies with quantum = 1 TS are suboptimal for a broad class of EH processes. Next, an upper bound is obtained for efficiencies of RR policies with quantum = 1 TS. Then, it is shown that the myopic policy have almost equal efficiency as another RR policy with quantum = 1 TS. Furthermore, numerical results show that the myopic policy is suboptimal for a broad class of EH processes although it achieves optimality for certain specific cases.

As a future work, we search for a simple, optimal solution to this problem for quite general EH processes. As another future work, we look for extending the single hop problem to multi hop case. Moreover, we plan to investigate the same problem under finite capacity battery case and make the throughput performance analysis of the myopic policy under finite capacity battery case. In addition, we will work to extend the problem in our future works such that we can consider the network lifetime as a performance metric. We believe that novel approaches and concepts in this paper will give insight to the researchers who study similar scheduling problems.

Abbreviations

The following abbreviations are used in this manuscript:

IoT	Internet of Things
WSN	Wireless sensor network
EH	Energy harvesting
FC	Fusion center
TS	time slot
SNR	Signal to Noise Ratio
RF	Radio frequency
POMDP	Partially Observable Markov Decision Processes
DP	Dynamic Programming
RMAB	Restless Multi Armed Bandit
PSPACE	Polynomial Space
MP	Myopic policy
RR	Round-Robin
IID	independent and identically distributed

Appendix A. Proof of Lemma 1

Define a new function, $\rho_i(t)$, for a sensor i as

$$\rho_i(t) \triangleq \frac{V_i^{FE}(T) - \max\limits_{\pi^{RR} \in G^{RR}} V_i^{RR}(t)}{\left\lceil \frac{K(T-t)}{M} \right\rceil}, \tag{A1}$$

where G^{RR} is the set of all RR policies with quantum = 1 TS.

Assume $\rho_i(t) > 1 > \rho_i$ for some sensor i and t. From Label (A1),

$$\frac{V_i^{FE}(T) - \max\limits_{\pi^{RR} \in G^{RR}} V_i^{RR}(t)}{\left\lceil \frac{K(T-t)}{M} \right\rceil} \quad > \quad 1, \quad \tau \leq t < T,$$

$$V_i^{FE}(T) - \max\limits_{\pi^{RR} \in G^{RR}} V_i^{RR}(t) \quad > \quad \left\lceil \frac{K(T-t)}{M} \right\rceil, \tau \leq t < T, \tag{A2}$$

where $\tau = \min\limits_{\rho_i(t)>1} t$.

In the interval, $\tau \leq t < T$, sensor i can send at most $\left\lceil \frac{K(T-t)}{M} \right\rceil$ data packets under an RR policy with quantum = 1 TS. Therefore, Label (A2) implies that $\min\{Y_i(0), q_i\} = Y_i(0)$ packets cannot be sent by sensor i over a time horizon, T, where $q_i \in \left\{ \left\lfloor \frac{KT}{M} \right\rfloor, \left\lfloor \frac{KT}{M} \right\rfloor + 1 \right\}$. If the EH process at sensor i were a constant EH process with intensity $\rho_i < 1$ (for these EH processes, $\rho_i(t) = \rho_i$ for all $1 \leq t < T$), then sensor i would send $\min\{Y_i(0), q_i\} = Y_i(0)$ packets over a time horizon, T, from Definition 5 and Remark 1. Hence, efficiencies of RR policies with quantum=1 TS under such an EH process that $\rho_i(t) > 1 > \rho_i$ for some sensor i and t become lower than they do under a constant EH process with intensity ρ_i. This implies that there exists a class of EH processes for which some sensor i cannot achieve the throughput of $\min\{Y_i(0), q_i\}$ packets over time horizon, T.

Appendix B. Proof of Theorem 2

Define

$$q \triangleq \frac{KT}{M} - \frac{m}{M}, \tag{A3}$$

where q and m are integers and $0 \leq m \leq M - K$. Recall $\frac{M}{K} \in \mathbb{Z}$ for applicability of RR policies with quantum = 1 TS.

Case i: Assume $\frac{KT}{M} \notin \mathbb{Z}$ and the RR policy with quantum = 1 TS starts to scheduling with sensor 1 without loss of generality. Notice that $K \leq m \leq M - K$ if $\frac{KT}{M} \notin \mathbb{Z}$. If the order $1, 2, \ldots, M$ is followed by the RR policy with quantum=1 TS for the scheduling, sensors $1, \ldots, m$ are scheduled $q + 1$ times and sensors $m + 1, \ldots, M$ are scheduled q times over a time horizon, T. From Lemma 1, for some EH processes,

$$V^{RR}(T) < \sum_{i=1}^{m} \min\{Y_i(0), q+1\} + \sum_{i=m+1}^{M} \min\{Y_i(0), q\}; \tag{A4}$$

otherwise, Label (A4) becomes equality for other EH processes. Hence,

$$V^{RR}(T) \leq \sum_{i=1}^{r} \min\{Y_i(0), q+1\} + \sum_{i=r+1}^{M} \min\{Y_i(0), q\}. \tag{A5}$$

Let $S = H_1 \cup H_2 \cup (S - H_1 - H_2)$, where $H_1 \subset \{1, \ldots, m\}$ and $H_2 \subset \{m+1, \ldots, M\}$ are the index sets of sensors that have enough energy to send more than $q + 1$ and q data packets, respectively. With this specification, the total throughput is

$$V^{RR}(T) \leq \sum_{i \in H_1}(q+1) + \sum_{i \in H_2} q + \sum_{i \in S-(H_1 \cup H_2)} Y_i(0). \tag{A6}$$

From Label (A6) and Definition 3,

$$\begin{aligned}
\eta(\pi^{RR}) &= \frac{V^{RR}(T)}{Y(0)} \\
&\leq \frac{\sum_{i \in H_1}(q+1) + \sum_{i \in H_2} q + \sum_{i \in S} Y_i(0) - \sum_{i \in H_1 \cup H_2} Y_i(0)}{\sum_{i \in S} Y_i(0)} \\
&= 1 - \frac{\sum_{i \in H_1 \cup H_2} Y_i(0) - \sum_{i \in H_1}(q+1) - \sum_{i \in H_2} q}{\sum_{i \in S} Y_i(0)}.
\end{aligned} \tag{A7}$$

If the numerator and denominator of the second term of the right-hand side in Label (A7) are normalized by $\frac{KT}{M}$,

$$\eta(\pi^{RR}) \leq 1 - \frac{\sum_{i \in H_1 \cup H_2} \frac{MY_i(0)}{KT} - \sum_{i \in H_1} \frac{M(q+1)}{KT} - \sum_{i \in H_2} \frac{Mq}{KT}}{\sum_{i \in S} \frac{MY_i(0)}{KT}}. \tag{A8}$$

From Definition 4, $\rho_i = \frac{MY_i(0)}{KT}$. Labels (A3) and (A8) yields

$$\eta(\pi^{RR}) \leq 1 - \frac{A}{\sum_{i \in S} \rho_i}, \tag{A9}$$

where

$$
\begin{aligned}
A &= \sum_{i \in H_1 \cup H_2} \rho_i - \sum_{i \in H_1} \left(1 + \frac{M - m}{KT} \right) - \sum_{i \in H_2} \left(1 - \frac{m}{KT} \right) \\
&= \sum_{i \in H_1 \cup H_2} (\rho_i - 1) - \sum_{i \in H_1} \frac{M}{KT} + \sum_{i \in H_1 \cup H_2} \frac{m}{KT} \\
&\geq \sum_{i \in H_1 \cup H_2} (\rho_i - 1) - \frac{M}{KT} |H_1| .
\end{aligned}
$$

Hence, efficiencies of RR policies with quantum = 1 TS satisfy

$$
\eta(\pi^{RR}) \leq 1 - \frac{\displaystyle\sum_{i \in H} (\rho_i - 1) - \frac{M}{KT} |H_1|}{M\rho} .
$$

Case ii: If $\frac{KT}{M} \in \mathbb{Z}$, then $q = \frac{KT}{M}$ since $m = 0$ in Label (A3) in this case. Let $S = H_2 \cup (S - H_2)$ where H_2 is index set of sensors that have enough energy to send more than q data packets. (Notice that $H_1 = \varnothing$ and $H = H_1 \cup H_2 = H_2$ if $\frac{KT}{M} \in \mathbb{Z}$.) By following similar steps in part (i), we obtain

$$
\begin{aligned}
V^{RR}(T) &\leq \sum_{i=1}^{M} \min \{ Y_i(0), q \} \\
&\leq \sum_{i \in H_2} q + \sum_{i \in S - H_2} Y_i(0). \tag{A10}
\end{aligned}
$$

From Label (A10) and Definition 3, efficiency of RR policies with quantum = 1 TS can be expressed as

$$
\begin{aligned}
\eta(\pi^{RR}) &= \frac{V^{RR}(T)}{Y(0)} \\
&= \frac{\displaystyle\sum_{i \in H_1} (\sigma + 1) + \sum_{i \in H_2} \sigma - \sum_{i \in H_1 \cup H_2} W_i(0) + \sum_{i \in H_1 \cup H_2} W_i(0) + \sum_{i \in S - H_1 - H_2} W_i(0)}{\displaystyle\sum_{i \in S} W_i(0)} \\
&\leq \frac{\displaystyle\sum_{i \in H_1 \cup H_2} q + \sum_{i \in S} Y_i(0) - \sum_{i \in H_1 \cup H_2} Y_i(0)}{\displaystyle\sum_{i \in S} Y_i(0)} \\
&\leq 1 - \frac{\displaystyle\sum_{i \in H_2} \frac{M Y_i(0)}{KT} - \sum_{i \in H_2} \frac{Mq}{KT}}{\displaystyle\sum_{i \in S} \frac{M Y_i(0)}{KT}} . \tag{A11}
\end{aligned}
$$

From Label (A3), Definitions 4 and 5, we obtain

$$
\eta(\pi^{RR}) \leq 1 - \frac{\displaystyle\sum_{i \in H} (\rho_i - 1)}{M\rho} .
$$

Appendix C. Proof of Lemma 2

Recall that $\frac{M}{K} \in \mathbb{Z}$ for applicability of RR policies with quantum = 1 TS. There are $\frac{M}{K}$ RR policies with quantum=1 TS which start to schedule a sensor i in different TSs. Recall that $1 \leq t_0 \leq \frac{M}{K}$ where t_0 is initial time for RR_{t_0} policy to schedule the sensor i. The proof is divided into two cases:

Case i: Assume that $\frac{KT}{M} \notin \mathbb{Z}$. Then, $\frac{KT}{M} \triangleq q + \frac{Kr}{M}$ where q and r are integers such that $1 \leq r \leq \frac{M}{K} - 1$ and $\frac{M}{K} \in \mathbb{Z}$. Some of the RR policies with quantum=1 TS schedules a sensor i $q+1$ times, whereas the other RR policies schedules the sensor i q times. The set of initial times of the RR policies which scheduled sensor i $q+1$ times is denoted by $P_{q+1}^{RR} = \{1, \ldots, r\}$, whereas the set of others is denoted by $P_q^{RR} = \left\{ r+1, \ldots, \frac{M}{K} \right\}$. t_j^n denotes the jth transmission time of sensor i under an RR policy, RR_n, which starts to schedule sensor i in TS n, where $j = 1, \ldots, q$ or $q+1$ are transmission times of sensor i under the applied RR policy, RR_n. By definition,

$$E_i\left(t_{j+1}^a\right) \geq E_i\left(t_j^b\right), \forall a \in P_{q+1}^{RR}, \forall b \in P_q^{RR}, \forall j \in \{1, \ldots, q\},$$

which yields

$$V_i^{RR_a}(T) \geq V_i^{RR_b}(T), \quad \forall a \in P_{q+1}^{RR}, \forall b \in P_q^{RR}. \tag{A12}$$

Notice that $r = mod_{\left(\frac{M}{K}\right)} T \in P_{q+1}^{RR}$ (in fact, $r = \max\limits_{n \in P_{q+1}^{RR}} n$). Therefore,

$$E_i\left(t_j^r\right) \geq E_i\left(t_j^a\right), \quad \forall a \in P_{q+1}^{RR}, \forall j \in \{1, \ldots, q+1\},$$

which yields

$$V_i^{RR_r}(T) \geq V_i^{RR_a}(T), \quad \forall a \in P_{q+1}^{RR}. \tag{A13}$$

Similarly, $r+1 = \min\limits_{n \in P_q^{RR}} n$. Therefore,

$$E_i\left(t_j^{r+1}\right) \leq E_i\left(t_j^b\right), \quad \forall b \in P_q^{RR}, \forall j \in \{1, \ldots, q\},$$

which yields

$$V_i^{RR_{r+1}}(T) \leq V_i^{RR_b}(T), \quad \forall b \in P_q^{RR}. \tag{A14}$$

From Labels (A12)–(A14),

$$\begin{aligned}
V_i^{RR_r}(T) &\geq V_i^{RR_a}(T) \geq V_i^{RR_b}(T) \geq V_i^{RR_{r+1}}(T), \\
\forall a &\in P_{q+1}^{RR}, \forall b \in P_q^{RR}.
\end{aligned} \tag{A15}$$

Furthermore,

$$E_i\left(t_j^{r+1}\right) \geq E_i\left(t_j^r\right), \forall j \in \{1, \ldots, q\},$$

which yields

$$V_i^{RR_{r+1}}(t_j^{r+1}) \geq V_i^{RR_r}(t_j^r), \forall j \in \{1, \ldots, q\}. \tag{A16}$$

From Label (A16),

$$\begin{aligned}
V_i^{RR_{r+1}}(T) &= V_i^{RR_{r+1}}\left(T - \frac{M}{K} + 1\right) \\
&\geq V_i^{RR_r}\left(T - \frac{M}{K}\right).
\end{aligned} \tag{A17}$$

From Labels (A15) and (A17),

$$\begin{aligned}
\max_n V_i^{RR_n}(T) - \min_n V_i^{RR_n}(T) &= V_i^{RR_r}(T) - V_i^{RR_{r+1}}(T) \\
&\leq V_i^{RR_r}(T) - V_i^{RR_r}\left(T - \frac{M}{K}\right) \\
&\leq 1. \tag{A18}
\end{aligned}$$

Hence, the lemma is proved for the first case, $\frac{KT}{M} \notin \mathbb{Z}$.

Case ii: If $\frac{KT}{M} \in \mathbb{Z}$, each of $\frac{M}{K}$ RR policies with quantum=1 TS schedules a sensor i $q = \frac{KT}{M}$ times. The set of initial times of RR policies with quantum=1 TS is denoted by $P_q^{RR} = \left\{1, 2, \ldots, \frac{M}{K}\right\}$. t_j^n denotes transmission time of sensor i under an RR policy, RR_n, where $j = 1, \ldots, q$ are transmission times of sensor i under the applied RR policy, RR_n, in this case

$$E_i\left(t_j^{\frac{M}{K}}\right) \geq E_i\left(t_j^n\right) \geq E_i\left(t_j^1\right), \forall n \in P_q^{RR}, \forall j \in \{1, \ldots, q\},$$

which yields

$$V_i^{RR_{\frac{M}{K}}}(T) \geq V_i^{RR_n}(T) \geq V_i^{RR_1}(T), \quad \forall n \in P_q^{RR}. \tag{A19}$$

Moreover,

$$E_i\left(t_{j-1}^{\frac{M}{K}}\right) \leq E_i\left(t_j^n\right), \quad \forall n \in P_q^{RR}, \forall j \in \{2, \ldots, q\},$$

which yields

$$V_i^{RR_{\frac{M}{K}}}\left(T - \frac{M}{K}\right) \leq V_i^{RR_n}(T), \quad \forall n \in P_q^{RR}. \tag{A20}$$

From Labels (A19) and (A20),

$$\begin{aligned} \max_n V_i^{RR_n}(T) - \min_n V_i^{RR_n}(T) &= V_i^{RR_{\frac{M}{K}}}(T) - V_i^{RR_1}(T) \\ &\leq V_i^{RR_{\frac{M}{K}}}(T) - V_i^{RR_{\frac{M}{K}}}\left(T - \frac{M}{K}\right) \\ &\leq 1. \end{aligned} \tag{A21}$$

Hence, the lemma is proved for the second case, $\frac{KT}{M} \in \mathbb{Z}$.

References

1. Sheng, Z.; Yang, S.; Yu, Y.; Vasilakos, A.; Mccann, J.; Leung, K. Survey on the IETF Protocol Suite for the Internet of Things: Standards, Challenges, and Opportunities. *IEEE Wirel. Commun.* **2013**, *20*, 91–98.
2. Kamalinejad, P.; Mahapatra, C.; Sheng, Z.; Mirabbasi, S.; Leung, V.C.M.; Guan, Y.L. Wireless Energy Harvesting for the Internet of Things. *IEEE Commun. Mag.* **2015**, *53*, 102–108.
3. Tsai, C.W.; Hong, T.P.; Shiu, G.N. Metaheuristics for the lifetime of WSN: A review. *IEEE Sens. J.* **2016**, *16*, 2812–2831.
4. Wark, T.; Corke, P.; Sikka, P.; Klingbeil, L.; Guo, Y.; Crossman, C.; Valencia, P.; Swain, D.; Bishop-Hurley, G. Transforming Agriculture through Pervasive Wireless Sensor Networks. *IEEE Pervasive Comput.* **2007**, *6*, doi:10.1109/MPRV.2007.47.
5. Chaiwatpongsakorn, C.; Lu, M.; Keener, T.C.; Khang, S.-J. The deployment of carbon monoxide wireless sensor network (CO-WSN) for ambient air monitoring. *Int. J. Environ. Res. Public Health* **2014**, *11*, 6246–6264.
6. Liu, X.-Y.; Zhu, Y.; Kong, L.; Liu, C.; Gu, Y.; Vasilakos, A.V.; Wu, M.-Y. CDC: Compressive data collection for wireless sensor networks. *IEEE Trans. Parallel Distrib. Syst.* **2015**, *26*, 2188–2197.
7. Valente, J.; Sanz, D.; Barrientos, A.; del Cerro, J.; Ribeiro, A.; Rossi, C. An Air-Ground Wireless Sensor Network for Crop Monitoring. *Sensors* **2011**, *11*, 6088–6108.
8. Fu, T.; Ghosh, A.; Johnson, E.A.; Krishnamachari, B. Energy-Efficient Deployment Strategies in Structural Health Monitoring using Wireless Sensor Networks. *J. Struct. Control Health Monit.* **2012**, *20*, 971–986.
9. Aderohunmu, F.; Balsamo, D.; Paci, G.; Brunelli, D. Long Term WSN Monitoring for Energy Efficiency in EU Cultural Heritage Buildings. *Lect. Notes Electr. Eng.* **2013**, *281*, 253–261.

10. Balsamo, D.; Paci, G.; Benini, L.; Davide, B. Long Term, Low Cost, Passive Environmental Monitoring of Heritage Buildings for Energy Efficiency Retrofitting. In Proceedings of the 2013 IEEE Workshop on Environmental Energy and Structural Monitoring Systems (EESMS), Trento, Italy, 11–12 September 2013; pp. 1–6.

11. Suryadevara, N.K.; Mukhopadhyay, S.C. Wireless sensor network based home monitoring system for wellness determination of elderly. *IEEE Sens. J.* **2012**, *12*, 1965–1972.

12. Yetgin, H.; Cheung, K.T.K.; El-Hajjar, M.; Hanzo, L. Network lifetime maximization of wireless sensor networks. *IEEE Access* **2015**, *3*, 2191–2226.

13. Zhou, F.; Chen, Z.; Guo, S.; Li, J. Maximizing Lifetime of Data-Gathering Trees With Different Aggregation Modes in WSNs. *IEEE Sens. J.* **2016**, *16*, 8167–8177.

14. Hancke, G.P.; de Carvalho e Silva, B.; Hancke, G.P., Jr. The role of advanced sensing in smart cities. *Sensors* **2013**, *13*, 393–425.

15. Gomez, C.; Paradells, J. Urban Automation Networks: Current and Emerging Solutions for Sensed Data Collection and Actuation in Smart Cities. *Sensors* **2015**, *15*, 22874–22898.

16. Paradiso, J.A.; Starner, T. Energy scavenging for mobile and wireless electronics. *IEEE Pervasive Comput.* **2005**, *4*, 18–27.

17. Sudevalayam, S.; Kulkarni, P. Energy Harvesting Sensor Nodes: Survey and Implications. *IEEE Commun. Surv. Tutor.* **2011**, *13*, 443–461.

18. Akyildiz, I.F.; Su, W.; Sankarasubramaniam, Y.; Cayirci, E. A survey on sensor networks. *IEEE Commun. Mag.* **2002**, *40*, 102–114.

19. Kansal, A.; Hsu, J.; Zahedi, S.; Srivastava, M.B. Power management in energy harvesting sensor networks. *ACM Trans. Embed. Comput. Syst.* **2007**, *6*, 32, doi:10.1145/1274858.1274870.

20. Garcia-Hernandez, C.F.; Ibargengoytia-Gonzalez, P.H.; Garcia-Hernandez, J.; Perez-Diaz, J.A. Wireless Sensor Networks and Applications: A Survey. *IJCSNS Int. J. Comput. Sci. Netw. Secur.* **2007**, *7*, 264–273.

21. Shannon, C.E. A Mathematical Theory of Communication. *Bell Syst. Technol. J.* **1948**, *27*, 379–426.

22. Alsheikh, M.A.; Hoang, D.T.; Niyato, D.; Tan, H.; Lin, S. Markov Decision Processes With Applications in Wireless Sensor Networks: A Survey. *IEEE Commun. Surv. Tutor.* **2015**, *17*, 1239–1267.

23. Battery University (Cadex Electronics). BU-802b: What Does Elevated Self-discharge Do? Available online: http://batteryuniversity.com/learn/article/elevating_self_discharge (accessed on 4 April 2017).

24. Arapostathis, A.; Borkar, V.S.; Fernández-gaucherand, E.; Ghosh, M.K.; Marcus, S.I. Discrete-time controlled Markov processes with average cost criterion: A survey. *SIAM J. Control Optim.* **1993**, *31* 282–334.

25. Monahan, G.E. State of the art-A survey of partially observable Markov decision processes: Theory, models, and algorithms. *Manag. Sci.* **1982**, *28*, 1–16.

26. Bellman, R.E. *Dynamic Programming*; Princeton University Press: Princeton, NJ, USA, 1957.

27. Littman, M.L.; Dean, T.L.; Kaelbling, L.P. In the complexity of solving Markov decision problems. In Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence, Montreal, QC, Canada, 18–20 August 1995; pp. 394–402.

28. Watkins, C.J. Learning from Delayed Rewards. Ph.D. Thesis, University of Cambridge, Cambridge, UK, 1989.

29. Kaelbling, L.P.; Littman, M.L.; Moore, A.W. Reinforcement learning: A survey. *J. Artif. Intell. Res.* **1996**, *4*, 237–285.

30. Mahadevan, S. Average reward reinforcement learning: Foundations, algorithms, and empirical results. *Mach. Learn. Spec. Issue Reinf. Learn.* **1996**, *22*, 159–196.

31. Whittle, P. Restless bandits: Activity allocation in a changing world. *J. Appl. Probab.* **1988**, *25*, 287–298.

32. Papadimitriou, C.H.; Tsitsiklis, J.N. The complexity of optimal queueing network control. *Math. Oper. Res.* **1999**, *24*, 293–305.

33. Hero, A.; Castanon, D.; Cochran, D.; Kastella, K. *Foundations and Applications of Sensor Management*; Springer: New York, NY, USA, 2007.

34. Blasco, P.; Gunduz, D.; Dohler, M. Low-Complexity Scheduling Policies for Energy Harvesting Communication Networks. In Proceedings of the IEEE International Symposium on Information Theory (ISIT), Istanbul, Turkey, 7–12 July 2013; pp. 1601–1605.

35. Iannello, F.; Simeone, O.; Spagnolini, U. Optimality of myopic scheduling and whittle indexability for energy harvesting sensors. In Proceedings of the 46th Annual Conference on Information Sciences and Systems(CISS), Princeton, NJ, USA, 21–23 March 2012; pp. 1–6.

36. Gittins, J.; Glazerbrook, K.; Weber, R. *Multi-Armed Bandit Allocation Indices*; Wiley: West Sussex, UK, 2011.

37. Gul, O.M.; Uysal-Biyikoglu, E. A randomized scheduling algorithm for energy harvesting wireless sensor networks achieving nearly 100% throughput. In Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC), Istanbul, Turkey, 6–9 April 2014; pp. 2456–2461.

38. Gul, O.M.; Uysal-Biyikoglu, E. Achieving nearly 100% throughput without feedback in energy harvesting wireless networks. In Proceedings of the IEEE International Symposium on Information Theory (ISIT), Honolulu, HI, USA, 29 June–4 July 2014; pp. 1171–1175.

39. Gul, O.M.; Uysal-Biyikoglu, E. UROP: A Simple, Near-Optimal Scheduling Policy for Energy Harvesting Sensors. *arXiv* **2014**, arXiv:1401.0437.

40. Durmaz Incel, O. A survey on multi-channel communication in wireless sensor networks. *Comput. Netw.* **2011**, *55*, 3081–3099.

41. Beckwith, R.; Teibel, D.; Bowen, P. Unwired Wine: Sensor Networks in Vineyards. In Proceedings of the 2014 IEEE Sensors, Vienna, Austria, 24–27 October 2004; pp. 1–4.

42. Chaudhary, D.D; Nayse, S.P.; Waghmare, L.M. Application of Wireless Sensor Networks for Greenhouse Parameter Control in Precision Agriculture. *Int. J. Wirel. Mob. Netw. IJWMN* **2011**, *3*, 140–149.

43. Srbinovska, M.; Gavrovski, C.; Dimcev, V.; Krkoleva, A.; Borozan, V. Environmental parameters monitoring in precision agriculture using wireless sensor networks. *J. Clean. Prod.* **2015**, *88*, 297–307.

44. Project RHEA. Robot Fleets For Highly Effective Agriculture And Forestry Management. From 2010-08-01 to 2014-07-31, Closed Project. Available online: http://cordis.europa.eu/project/rcn/95055_en.html (accessed on 29 May 2017).

45. Project CROPS. Intelligent Sensing and Manipulation for Sustainable Production and Harvesting of High Value Crops, Clever Robots for Crops. From 2010-10-01 to 2014-09-30, Closed Project. Available online: http://cordis.europa.eu/project/rcn/96216_en.html (accessed on 25 May 2017).