

Resource Allocation in Sensor Networks with Renewable Energy

(Invited Paper)

Zhoujia Mao
Department of ECE
The Ohio State University
maoz@ece.osu.edu

Can Emre Koksal
Department of ECE
The Ohio State University
koksal@ece.osu.edu

Ness B. Shroff
Departments of ECE and CSE
The Ohio State University
shroff@ece.osu.edu

Abstract—Renewable energy sources can be attached to sensor nodes to provide energy replenishment for prolonging the lifetime of sensor networks. However, for networks with replenishment, conservative energy expenditure may lead to missed recharging opportunities due to battery capacity limitations, while aggressive usage of energy may result in reduced coverage or connectivity for certain time periods. Thus, new power allocation schemes need to be designed to balance these seemingly contradictory goals, in order to maximize sensor network performance. In this paper, we study the problem of how to jointly control the data queue and battery buffer to maximize the long-term average sensing rate of a single communication link in rechargeable sensor networks. The coupling between the battery and data buffers does not lend itself amenable to traditional resource optimization techniques. Thus, we develop a new power and rate allocation scheme that explicitly takes this coupling into account. The new scheme is a simple myopic scheme whose performance is shown to be arbitrarily close to optimal analytically and via simulations.

I. INTRODUCTION

Wireless sensor networks are widely used in monitoring [1], maintenance [2], and environmental sensing [3]. The lack of easy access to a continuous power source and the limited lifetime of batteries have hindered the wide-scale deployment of such networks. However, new and exciting developments in the area of renewable energy [4] [5] provide an alternative to a limited power source, and may help alleviate some of the deployment challenges. These renewable sources of energy could be attached to the nodes and would typically provide energy replenishment at a slow rate (compared to the rate at which energy is consumed by a continuous stream of packet transmissions) that could be variable and dependent on the surroundings.

Energy management in networks equipped with renewable sources is substantially and qualitatively different from energy management in traditional networks. For example, conservative energy expenditure could lead to (i) instability of the data queue because the energy is not being fully used to transmit at high enough data rates, or (ii) missed recharging opportunities because the battery buffer is full. On the other hand, an over aggressive use of energy may lead to lack of coverage or connectivity for certain time periods, which could hurt the

application's requirements. Further, if the battery of a node discharges completely, it could be temporarily incapable of transferring time-sensitive data to the sink. This may have undesirable consequences for many applications. Thus, new techniques and protocols must be developed for networks with replenishment to balance these seemingly contradictory goals.

In this paper, we consider single link communication, in which the transmitting node has an infinite data buffer that holds the incoming variable-rate sensing data and a finite battery buffer, which is being replenished at a variable rate. One may experience instability of the data queue or frequent occurrences of battery discharge (batteries becoming empty) if power consumption and sensing rates are not managed properly. We investigate the problem of maximizing the long-term average sensing rate, subject to data queue stability and constraints on the desired rate of visits to zero battery state. We provide a simple joint rate control and power management scheme, and show that the performance of our scheme is close to optimum.

While the problem of energy management in sensor networks has seen considerable attention, there are relatively few works [6], [7], [8], [9], [10], [11] that also include energy replenishment. In [6], the authors consider the problem of dynamic node activation in rechargeable sensor networks. They provide a distributed threshold policy that achieves a performance within a certain factor of the optimal solution for a set of sensors whose coverage area overlap completely. In [7], the authors study the problem of computing lexicographically maximum data collection rate for each node such that no node will be out of energy. In [8], the authors consider the problem of energy-aware routing with distributed energy replenishment. They provide an algorithm that achieves a logarithmic competitive ratio and is asymptotically optimal with respect to the number of nodes in the network. However, the problems considered in these works is very different from the one considered here in that the coupling between the data and battery queue is not taken into account. Prior work on power allocation for wireless networks without replenishment has been widely studied, e.g., [12], [13], [14]. In [12], the authors develop approximated algorithms to minimize average

allocated power, or maximize throughput given average power constraint, and at the same time keep the data queue stable. The approximation improves at the cost of increasing the data queue length and queuing delay. In [13], [14] the authors assume that the data buffer is large enough so that packet loss does not occur and provide a dynamic programming based solution. Most of these works assume a constant energy supply, i.e., there is no battery issue. In this paper, we explicitly model energy replenishment and consider jointly managing the data and battery buffers. This coupling between the data and battery buffers is what makes the problem notoriously difficult to solve using standard optimization based approaches. For example, unlike prior works that utilize the fact that a static allocation policy is optimal, and then develop dynamic policies based on their performance against a dynamic policy, it isn't even clear whether a static policy would be optimal in our setting. Specifically, with a battery buffer, there is an additional energy constraint that the allocated energy should be within the battery state, and this constraint is even more difficult than the average power constraint.

The main contributions of the paper are as follows:

- We formulate a sensing rate maximization problem with constraints on both the stability of data queue and the probability of having an empty battery. Due to the coupling between the battery and data buffers it is unlikely that a stationary policy optimal, hence traditional resource allocation techniques do not directly apply. Further, dynamic programming based solutions result in prohibitively high complexity, even for the single-link scenario. Nonetheless, we are able to develop a simple joint rate control and power allocation algorithm that is provably efficient.
- Our solution structure can be readily extended to the multihop network scenario under the primary interference model [15].

II. SYSTEM MODEL

In this paper, we focus on the problem of controlling the energy used for data transmission as a single link system, as illustrated in Figure 1. We assume a time slotted system and in time slot t , the amount of data (in packets) available for the sensor node to sense is denoted by $A(t)$, which is upper bounded by $A_{\max} > 0$. In the same time slot, the actual amount of data the node senses and places in the data buffer is $R(t)$ packets. The data in the data buffer waits to be transmitted to a collection point over a wireless channel. Let $P(t)$ denote the amount of energy the node spends at time t for data transmission, which is bounded by a peak power constraint $P_{\text{peak}} > 0$. The achievable data rate (in packets/time slot) at power level $P(t)$ is $\mu(P(t))$, where we assume $\mu(\cdot)$ to be monotonically increasing, reversible and differentiable on the half real line $\mathbb{R}^+ \cup \{0\}$. The node has a finite battery of size B_b , where $r(t)$ is the instantaneous rate at time t of the (time-varying) energy that is harvested by the node. This harvested energy could be used for data transmission and/or replenishing the battery queue. The energy state of the battery and the data

state of the data buffer at time t are given by $q_b(t)$ and $q_d(t)$, respectively. We assume that the data buffer is infinite. Under this setting, we describe the objective in the following section.

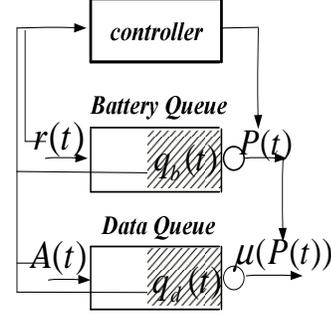


Fig. 1. System Model

A. Problem Formulation

Our general objective is to maximize the average long-term sensing rate, subject to the constraints on the stability of the data queue and the ‘‘frequency’’ of complete battery discharge due to the battery state hitting zero. Formally, we have:

$$\begin{aligned}
 (A) \quad & \max_{\vec{P}, \vec{R}} \liminf_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} R(t) \\
 \text{s.t.} \quad & q_d(t+1) = (q_d(t) - \mu(P(t)))^+ + R(t), \quad (a) \\
 & q_b(t+1) = \min[q_b(t) - P(t) + r(t), B_b], \quad (b) \\
 & 0 \leq R(t) \leq A(t), \quad (c) \\
 & 0 \leq P(t) \leq \min[q_b(t), P_{\text{peak}}], \quad (d) \\
 & \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} q_d(t) < \infty, \quad p_o \leq \eta_o \quad (e)
 \end{aligned}$$

where $(\cdot)^+ = \max[\cdot, 0]$, $\vec{R} = \{R(0), R(1), \dots, R(T-1), \dots\}$ is the actual sensing data vector, \vec{P} is the power allocation policy $\{P(0), P(1), \dots, P(T-1), \dots\}$ and

$$p_o = \limsup_{T \rightarrow \infty} \frac{\sum_{t=0}^{T-1} I_o(t)}{T} \quad (1)$$

is the frequency of visits to zero battery state with given threshold η_o , where

$$\begin{aligned}
 I_o(t) &= \mathbf{1}_{\text{battery state '0' is visited from higher states in slot } t} \\
 &= \begin{cases} 0 & \text{if } P(t) = 0 \text{ or } P(t) < q_b(t) \\ 1 & \text{otherwise} \end{cases} \quad (2)
 \end{aligned}$$

is the indicator that the battery discharges completely in time slot t . Note that, if the battery is completely discharged at time t and there is no replenishment at slot t , we do not consider a complete discharge event occurring at time $t+1$ as well. We do not assume ergodicity of the system parameters, but if they are ergodic, then p_o represents the actual probability of a complete discharge event as $t \rightarrow \infty$.

In Problem (A), constraints (a) and (b) describe how the data and battery queues evolve, respectively. Constraint (c) bounds

the actual amount of sensed data $R(t)$ by the available amount of data $A(t)$ in slot t . Constraint (d) states that we cannot oversubscribe the energy that is unavailable in the battery nor can we exceed the peak power level. The constraints described in (e) contain requirement for data queue stability and a bound η_o on the battery discharge rate.

Problem (A) has an inventory control structure, and typically such a structure can be solved optimally using dynamic programming, albeit with high complexity. Furthermore, depending on the exogenous processes $\{A(t), t \geq 0\}$ and $\{r(t), t \geq 0\}$, Problem (A) may not have a feasible solution, i.e., there may exist no power allocation policy $\{P(t), t \geq 0\}$ that satisfies all constraints simultaneously. In this paper, our purpose is to develop a simple algorithm, which performs arbitrarily close to the performance of the optimal power allocation, whenever a feasible solution exists. To achieve that purpose, we do the following:

- we define a *virtual queue* for the battery buffer to avoid the difficulties involved in dealing with the discharge probability directly. We show that keeping the virtual battery queue stable ensures that the constraint on p_o is met.
- we design a power allocation scheme based on simple index policies and show that our scheme keeps the data queue and virtual battery queue stable, and at the same time performs arbitrarily close to the optimal performance.

B. Virtual Queue

We define \tilde{q}_b as the virtual battery queue. The virtual queue evolves according to the following Lindley's queue evolution equation:

$$\tilde{q}_b(t+1) = (\tilde{q}_b(t) - \eta_o)^+ + P(t) - r(t) + M(t) + I_o(t), \quad (3)$$

where $M(t) = (q_b(t) - P(t) + r(t) - B_b)^+$ is the amount of missed replenishment, and $I_o(t)$ is defined in Equation (2). Without loss of generality, the initial state $\tilde{q}_b(0)$ can be set to be zero. The virtual queue $\tilde{q}_b(t)$ can be interpreted as the number of visits to the zero battery state beyond what is tolerable. The following proposition shows that if the virtual queue $\tilde{q}_b(t)$ is strongly stable, p_o is guaranteed to meet its constraint.

Proposition 1: If the virtual battery queue $\tilde{q}_b(t)$ is strongly stable, i.e.,

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} \tilde{q}_b(t) < \infty,$$

then $p_o \leq \eta_o$.

The proof of this result can be found in Appendix A. Next, we present our scheme.

III. JOINT RATE CONTROL AND POWER ALLOCATION ALGORITHM AND ANALYSIS

A. Algorithm

The algorithm consists of two components: a *rate control* component and a *power allocation* component. Both compo-

nents are *index policies*, i.e., the solutions are memoryless and they depend only on the instantaneous values of the system variables.

Rate Control (RC):

We define $V \in \mathbb{R}^+$ to be the control parameter of our algorithm. If $q_d(t) \leq \frac{V}{2}$, the transmitting node chooses to sense all the available data packets, i.e., $R(t) = A(t)$. Otherwise, $R(t) = 0$.

Power Allocation (PA):

I) Calculation Phase:

Solve the following optimization problem:

$$\max_{P(t) \in \Pi(t)} q_d(t)\mu(P(t)) - \tilde{q}_b(t)P(t), \quad (4)$$

where $\Pi(t) = \{P(t) : 0 \leq P(t) \leq \min[q_b(t), P_{\text{peak}}]\}$ is a compact and nonempty set. Let $P^I(t)$ be its solution.

II) Modification Phase:

If $q_b(t) - P^I(t) + r(t) > B_b$, i.e., $P^I(t)$ will lead to wasted replenishment energy, then increase the transmit power to utilize all the current replenishment: choose transmit power $P^{II}(t) = \min\{r(t), P_{\text{peak}}\}$. Allocate $P^{II}(t)$ for transmission; Otherwise, allocate $P^I(t)$ for transmission.

The set, $\Pi(t)$ of possible power allocations guarantees Constraint (d) on $P(t)$ in Problem (A). If $\mu(\cdot)$ is concave, the objective function is a concave function of $P(t)$. Consequently, PA solves a simple convex optimization problem in each time slot. The positive term $q_d(t)\mu(P(t))$ can be viewed as a utility received from the allocated power $P(t)$ and the term $\tilde{q}_b(t)P(t)$ can be viewed as its associated cost. When the control parameter V is chosen to be large, $q_d(t)$ tends to be large according to RC, and PA tries to allocate higher values of $P(t)$ to increase the utility, whereas, when the virtual battery queue length $\tilde{q}_b(t)$ is large, PA avoids allocating a high amount of power to reduce cost. Thus, this index policy of PA can be viewed as a *greedy profit maximization* scheme. Note that (4) leads to a solution, which is oblivious to the battery state. This leads to the wastage of replenishment energy, if the battery state is too large to contain it. The modification Phase eliminates such inefficiencies by increasing the allocated power, whenever there is available data to be transmitted in the data buffer.

B. Performance Analysis

Recall that $A(t)$ is the available amount of sensing data and $R(t)$ is the actual amount of data the transmitting node chooses to sense. Clearly, using the rate controller, we make sure that the data queue remains within a certain bound and this has a positive effect on the battery as well, since certain proportion of the data packets are not allowed into the transmitting node. The natural question one would ask here is, whether our rate controller rejects too many packets in the first place to *synthetically* meet the constraints. In the following theorem, we show that this is not the case. Indeed, if there exists a solution, $\lambda^* = \liminf_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} R^*(t)$ to Problem (A) for the arrival process $\{A(t), t \geq 0\}$ and the replenishment process $\{r(t), t \geq 0\}$, then the sensing rate associated with

RC and PA can be made closer to λ^* by increasing the control parameter V . We use the notation $y = O(x)$ to represent y going to 0 as x goes to 0.

Theorem 1: If

(1) $\mu(\cdot)$ is concave on $\mathbb{R}^+ \cup \{0\}$, and its slope at 0 satisfies¹ $0 \leq \beta = \mu'(0) < \infty$,

(2) $0 < r(t) \leq r_{\max} \leq P_{\text{peak}} \leq B_b, \forall t \geq 0$,

(3) A feasible solution to Problem (A) exists and the optimal instantaneous sensing rate is $R^*(t)$,

then the joint power allocation and admission control algorithm (with RC and PA) achieves:

$$q_d(t) \leq \frac{V}{2} + A_{\max}, \quad \forall t \geq 0 \quad (5)$$

$$\tilde{q}_b(t) \leq \beta\left(\frac{V}{2} + A_{\max}\right), \quad \forall t \geq 0 \quad (6)$$

$$\liminf_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} R(t) \geq \liminf_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} R^*(t) - \eta_o(\mu_{\max} + \beta) - O\left(\frac{1}{V}\right), \quad (7)$$

where $\mu_{\max} = \mu(P_{\text{peak}})$ is the upper bound for the transmission rate.

The proof of Theorem 1 can be found in Appendix B. Equation (5) and Equation (6) show that the data queue q_d and virtual battery queue \tilde{q}_b are both strongly stable and thus, by Proposition 1, $p_o \leq \eta_o$. In Equation (7), the term $\eta_o(\mu_{\max} + \beta)$ captures the influence of battery outages with our algorithm. This term is small, since the battery outage threshold η_o is usually chosen to be small. Thus, the gap between the average sensing rate with our algorithm and the optimal average sensing rate can be made arbitrarily small by choosing parameter V large and η_o small.

IV. NUMERICAL EXAMPLE

In this section we simulate our algorithm and numerically compare it with the optimal performance. In the simulation, the number of time slots is $T = 10^6$, and the duration of each slot is 10 secs. We use the rate power function $\mu(P) = 10 \log_2(1 + \frac{gP}{N})$ packets/slot, where the power spectral density of the background noise is $N = 1.6 \times 10^{-14}W$, and the channel gain is $g = 1.6 \times 10^{-13}$. The battery buffer size is set to be $B_b = 800J$. The number of arrivals $A(t), t \geq 0$, are independent Poisson random variables with mean 20 packets/slot. The replenishment process is periodic with Gaussian noise, as shown in Figure 2 (a) (the cycles imitate the daily solar cycles for a solar battery). We also set η_o , the threshold of battery outage probability to 0.03.

We run the simulation for different values of the control coefficient V and compare the results with the optimal value². In Figure 2 (b), one can observe that, as V increases, the average sensing rate approaches the optimal value, which is

¹For instance, consider the additive white Gaussian noise channel capacity, $\mu(P) = \log(1 + P/N_0)$, where N_0 is the two sided noise power spectral density.

²The optimal value can be obtained by using dynamic programming. The details are omitted here.

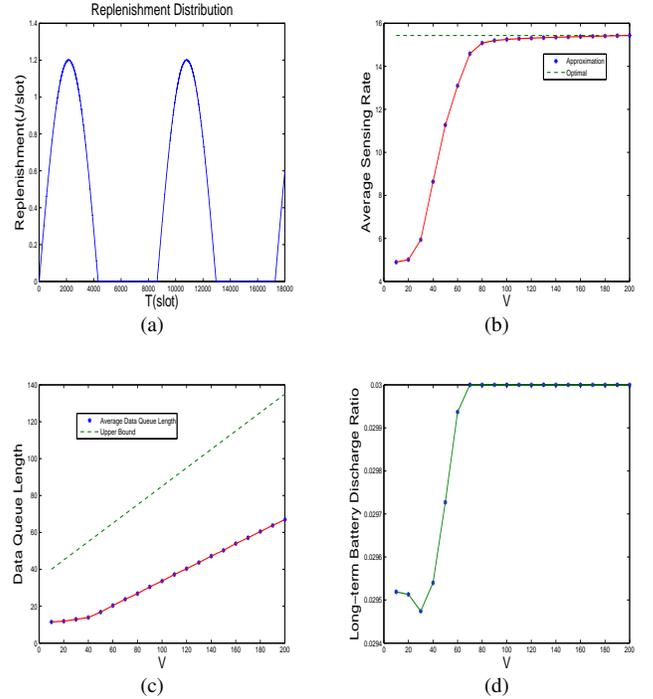


Fig. 2. Simulation Setting and Performance of the proposed algorithm. (a) Replenishment Process. Impact of the control parameter V on (b) the average sensing rate, (c) average data queue length, and (d) the battery discharge probability

consistent with Equation (7). Figure 2 (c) shows that, as V increases, the average data queue length increases but remains upper bounded as given in Equation (5). Figure 2 (d) illustrates that, the battery discharge probability increases to η_o as V increases. Thus, increasing V can increase the average sensing rate at the expense of the increased average data queue length and battery discharge frequency.

V. CONCLUSION

In this paper, we studied the problem of energy management in rechargeable wireless sensor networks. Our objective was to maximize the average sensing rate subject to constraints on the stability of the data queue and the frequency of battery discharge. We provided a simple index strategy for jointly controlling the rate and power and showed through both analysis and simulation that the performance of our strategy is close to that of the optimal solution. Our solution is readily extendable to the multihop scenario under the primary interference model. However, due to the space limitation, we omit the description here. For future work, we are interested in studying the performance and implementation of the algorithm for networks under the physical interference model.

REFERENCES

- [1] A. Mainwaring, J. Polastre, R. Szewczyk, D. Culler, and J. Anderson, "Wireless Sensor Networks for Habitat Monitoring," in *Proc. of ACM WSN/A*, September 2002, pp. 88–97.
- [2] N. Xu, S. Rangwala, K. Chintalapudi, D. Ganesan, A. Broad, R. Govindan, and D. Estrin, "A Wireless Sensor Network for Structural Monitoring," in *Proc. of SenSys*, 2004, pp. 13–24.

- [3] K. Martinez, R. Ong, and J. Hart, "Glacweb: a Sensor Network for Hostile Environments," in *Proc. of the 1st IEEE Communications Society Conference on Sensor and Ad Hoc Communications and Networks*, Santa Clara, CA, 2004, pp. 81–87.
- [4] J. Paradiso and M. Feldmeier, "A Compact, Wireless, Self-Powered Pushbutton Controller," in *Proc. of the 3rd International Conference on Ubiquitous Computing*, 2001, pp. 299–304.
- [5] W. Weber, "Ambient Intelligence: Industrial Research on A Visionary Concept," in *Proc. of the 2003 International Symposium on Low Power Electronics and Design*, 2003, pp. 247–251.
- [6] K. Kar, A. Krishnamurthy, and N. Jaggi, "Dynamic Node Activation in Networks of Rechargeable Sensors," *IEEE/ACM Transactions on Networking*, vol. 14, pp. 15–26, 2006.
- [7] K. Fan, Z. Zheng, and P. Sinha, "Steady and Fair Rate Allocation for Rechargeable Sensors in Perpetual Sensor Networks," in *Proc. of the 6th ACM Conference on Embedded Network Sensor Systems*, 2008, pp. 239–252.
- [8] L. Lin, N. Shroff, and R. Srikant, "Asymptotically Optimal Energy-Aware Routing for Multihop Wireless Networks with Renewable Energy Sources," *IEEE/ACM Transactions on Networking*, vol. 15, no. 5, pp. 1021–1034, October 2007.
- [9] A. Kansal, J. Hsu, S. Zahedi, and M. Srivastava, "Power Management in Energy Harvesting Sensor Networks," *ACM Transactions on Embedded Computing Systems*, vol. 6, no. 4, September 2007.
- [10] D. Niyato, E. Hossain, and A. Fallahi, "Sleep and Wakeup Strategies in Solar-Powered Wireless Sensor/Mesh Networks: Performance Analysis and Optimization," *IEEE Transactions on Mobile Computing*, vol. 6, no. 2, pp. 221–236, February 2007.
- [11] C. Vigorito, D. Ganesan, and A. Barto, "Adaptive Control of Duty Cycling in Energy-Harvesting Wireless Sensor Networks," in *Proc. of IEEE SECON*, 2007.
- [12] M. Neely, "Energy Optimal Control for Time Varying Wireless Networks," *IEEE Transactions on Information Theory*, vol. 52, no. 7, pp. 2915–2934, July 2006.
- [13] N. Salodkar, A. Borkar, A. Karandikar, and V. Borkar, "An On-Line Learning Algorithm for Energy Efficient Delay Constrained Scheduling over a Fading Channel," *IEEE Journal on Selected Areas in Communications*, May 2008.
- [14] N. Salodkar, A. Karandikar, and V. Borkar, "Power Efficient Scheduling under Delay Constraints over Multi-user Wireless Channels," *IEEE Transactions on Mobile Computing*, 2007.
- [15] Z. Mao, C. Koksal, and N. Shroff, "Joint queue and energy management for rechargeable sensor networks," Tech. Rep., 2010. [Online]. Available: http://www.archive.org/details/Tp_zhoujiaMao

APPENDIX

A. Proof of Proposition 1

We prove the proposition for a more general case. Consider a virtual battery queue that follows the following evolution:

$$\begin{aligned} \tilde{q}_b'(t+1) &= (\tilde{q}_b'(t) + C(t) - \eta_o)^+ + I_o(t) \\ &\quad + P(t) - r(t) + M(t) - C(t+1), \end{aligned} \quad (8)$$

for some process $C(t)$, $t \geq 0$, where $0 \leq C(t) < \infty$, $\forall t \geq 0$. The original virtual queue $\tilde{q}_b(t)$ as described in Equation (3) is a special case of $\tilde{q}_b'(t)$, where $C(t) = 0$, $\forall t \geq 0$. Thus, it is sufficient that the proposition holds for $\tilde{q}_b'(t)$.

Using the idea similar to [12], we have the fact that if any queue represented with $Q(t)$ is strongly stable, then $\limsup_{T \rightarrow \infty} \frac{Q(T)}{T} \leq 0$. Hence, if $\tilde{q}_b'(t)$ is strongly stable, $\limsup_{T \rightarrow \infty} \frac{\tilde{q}_b'(T)}{T} \leq 0$. From Equation (3), we have

$$\begin{aligned} \tilde{q}_b'(t+1) &\geq \tilde{q}_b'(t) + C(t) - \eta_o + I_o(t) \\ &\quad + P(t) - r(t) + M(t) - C(t+1). \end{aligned}$$

Note that $q_b(t+1) = q_b(t) - P(t) + r(t) - M(t)$. By adding from 0 to $T-1$, dividing by T and taking \limsup on both

sides, we have

$$\begin{aligned} \limsup_{T \rightarrow \infty} \frac{\tilde{q}_b'(T)}{T} &\geq \lim_{T \rightarrow \infty} \frac{\tilde{q}_b'(0)}{T} - \eta_o + \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} I_o(t) \\ &\quad + \lim_{T \rightarrow \infty} \frac{C(0) - C(T) + q_b(0) - q_b(T)}{T}. \end{aligned}$$

Since $\limsup_{T \rightarrow \infty} \frac{\tilde{q}_b'(T)}{T} \leq 0$, $\lim_{T \rightarrow \infty} \frac{\tilde{q}_b'(0)}{T} = \lim_{T \rightarrow \infty} \frac{C(0)}{T} = \lim_{T \rightarrow \infty} \frac{C(T)}{T} = \lim_{T \rightarrow \infty} \frac{q_b(0)}{T} = \lim_{T \rightarrow \infty} \frac{q_b(T)}{T} = 0$, so we get $p_o = \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} I_o(t) \leq \eta_o$. ■

B. Proof of Theorem 1

Proof of Equation (5): The rate allocation unit RC is chosen to satisfy Equation (5).

To show (6) and (7), we use the more general virtual queue $\tilde{q}_b'(t)$ as given in (8) and choose $C(t)$ carefully. One can observe that, to obtain $\tilde{q}_b'(t)$ from $\tilde{q}_b(t)$, we subtract $C(t)$ from $\tilde{q}_b(t)$ at time t and add it back to it with a unit time delay. Using $\tilde{q}_b'(t)$, the modification phase of PA can be rewritten as follows:

If $q_b(t) - P^I(t) + r(t) > B_b$, let $\tilde{q}_b''(t)$ be such that $\arg \max_{P(t) \in \Pi(t)} q_d(t)\mu(P(t)) - \tilde{q}_b''(t)P(t) = \min[r(t), P_{\text{peak}}]$. Let $C(t) = \tilde{q}_b'(t) - \tilde{q}_b''(t)$, update $\tilde{q}_b'(t)$ according to $\tilde{q}_b'(t) = (\tilde{q}_b'(t-1) + C(t-1) - \eta_o)^+ + I_o(t-1) + P(t) - r(t) + M(t) - C(t)$. With the updated $\tilde{q}_b'(t)$, allocate $P(t)$ amount of energy for transmission, where $P(t)$ is the solution of

$$\max_{P(t) \in \Pi(t)} q_d(t)\mu(P(t)) - \tilde{q}_b'(t)P(t).$$

Otherwise, allocate $P^I(t)$ amount of energy for transmission. Finally, update $\tilde{q}_b'(t+1)$ according to Equation (8) with $C(t+1) = 0$. Thus, the modification phase also follows the same optimization structure as in Equation (4).

Proof of Equation (6): Since $\mu(\cdot)$ is concave on $\mathbb{R}^+ \cup \{0\}$, we have $\mu(P(t)) \leq \mu(0) + \beta P(t)$ for $P(t) \in \Pi(t)$, $\forall t \geq 0$, where $0 \leq \beta = \mu'(0) < \infty$. Then, $q_d(t)\mu(P(t)) - \tilde{q}_b'(t)P(t) \leq q_d(t)\mu(0) + \beta q_d(t)P(t) - \tilde{q}_b'(t)P(t)$ where $P(t)$ is the solution of PA .

If $\beta q_d(t)P(t) - \tilde{q}_b'(t)P(t) < 0$, then we get $q_d(t)\mu(P(t)) - \tilde{q}_b'(t)P(t) < q_d(t)\mu(0)$. However, PA chooses $P(t)$ that maximizes $q_d(t)\mu(P(t)) - \tilde{q}_b'(t)P(t)$ which means $q_d(t)\mu(P(t)) - \tilde{q}_b'(t)P(t) \geq q_d(t)\mu(0)$ since $0 \in \Pi(t)$. Then, we must have $\beta q_d(t)P(t) - \tilde{q}_b'(t)P(t) \geq 0$, i.e., $\beta q_d(t)P(t) \geq \tilde{q}_b'(t)P(t)$. Thus, if $P(t) > 0$, $\tilde{q}_b'(t) \leq \beta q_d(t) \leq \beta(\frac{V}{2} + A_{\text{max}})$; if $P(t) = 0$, by Equation (2), $I_o(t) = 0$, and $M(t) \leq r(t)$, then $\tilde{q}_b'(t)$ does not increase anyway. Therefore, $\tilde{q}_b(t) \leq \tilde{q}_b'(t) \leq \beta(\frac{V}{2} + A_{\text{max}})$ which is Equation (6).

Proof of Equation (7): We define the Lyapunov function $L(q_d(t), \tilde{q}_b'(t)) = q_d^2(t) + (\tilde{q}_b'(t))^2$, and $\Delta(q_d(t), \tilde{q}_b'(t)) = L(q_d(t+1), \tilde{q}_b'(t+1)) - L(q_d(t), \tilde{q}_b'(t))$. From Equation (3), we have $(\tilde{q}_b'(t+1))^2 + 2\tilde{q}_b'(t+1)C(t+1) + C^2(t+1) \leq (\tilde{q}_b'(t) + C(t) - \eta_o)^2 + (I_o(t) + P(t) - r(t) + M(t))^2 + 2(\tilde{q}_b'(t) + C(t) - \eta_o)^+(I_o(t) + P(t) - r(t) + M(t))$. Also from the data

queue dynamics, we have $q_d^2(t+1) \leq q_d^2(t) + \mu^2(P(t)) + R^2(t) + 2q_d(t)R(t) - 2q_d(t)\mu(P(t))$, then

$$\begin{aligned} \Delta &= \Delta(q_d(t), \tilde{q}'_b(t)) \\ &\leq \mu^2(P(t)) + R^2(t) + 2q_d(t)R(t) - 2q_d(t)\mu(P(t)) + (1 + P_{\text{peak}})^2 + \eta_o^2 + \eta_o r_{\text{max}} + 2\tilde{q}'_b(t)(I_o(t) + P(t) - r(t) + M(t)) + 2C(t)I_o(t) + 2C(t)(P(t) - r(t) + M(t)) + 2\tilde{q}'_b(t)C(t) - 2\tilde{q}'_b(t+1)C(t+1) + C^2(t) - C^2(t+1) \\ &\leq (1 + P_{\text{peak}})^2 + \eta_o^2 + \eta_o r_{\text{max}} + \mu^2(P(t)) + R^2(t) + VR(t) + 2[q_d(t) - V/2]R(t) - 2[q_d(t)\mu(P(t)) - \tilde{q}'_b(t)P(t)] + 2\tilde{q}'_b(t)(I_o(t) - r(t) + M(t)) + 2C(t)I_o(t) + 2C(t)(P(t) - r(t) + M(t)) + 2\tilde{q}'_b(t)C(t) - 2\tilde{q}'_b(t+1)C(t+1) + C^2(t) - C^2(t+1) \end{aligned}$$

It is apparent that RC is trying to minimize the term $[q_d(t) - V/2]R(t)$, and PA is trying to maximize the value of the term $[q_d(t)\mu(P(t)) - \tilde{q}'_b(t)P(t)]$. Since the optimal solution for Problem (A) may not be unique, we let \mathcal{P}^* be the optimal solution set and $P^* \in \mathcal{P}^*$ be any optimal solution, for Problem (A). In time slot t , let $P_m(t)$ be the value that maximize the unconstrained objective function $q_d(t)\mu(P(t)) - \tilde{q}'_b(t)P(t)$. Since $\Pi(t) = \{P(t) : 0 \leq P(t) \leq \min[q_b(t), P_{\text{peak}}]\}$, only when $q_b(t) \leq P_{\text{peak}}$ and $P_m(t), P^*(t) \notin \Pi(t)$, we may have $q_d(t)\mu(P(t)) - \tilde{q}'_b(t)P(t) \leq q_d(t)\mu(P^*(t)) - \tilde{q}'_b(t)P^*(t)$. Thus, we have

$$\begin{aligned} \Delta &\leq (1 + P_{\text{peak}})^2 + \eta_o^2 + \eta_o r_{\text{max}} + \mu_{\text{max}}^2 + A_{\text{max}}^2 + VR(t) - VR^*(t) + 2q_d(t)[R^*(t) - \mu(P^*(t))] + 2\tilde{q}'_b(t)[P^*(t) - r(t) + M(t)] + 2\tilde{q}'_b(t)I_o(t) + 2C(t)I_o(t) + 2C(t)(P(t) - r(t) + M(t)) + C^2(t) - C^2(t+1) + 2\tilde{q}'_b(t)C(t) + 2[q_d(t)(\mu(P^*(t)) - \mu(P(t))) + \tilde{q}'_b(t)(P(t) - P^*(t))] \cdot I_{[P_m(t), P^*(t)] \notin \Pi(t)} \cap [q_b(t) \leq P_{\text{peak}}] - 2\tilde{q}'_b(t+1)C(t+1). \end{aligned}$$

When $P_m(t) \notin \Pi(t)$ and $q_b(t) \leq P_{\text{peak}}$, PA will allocate $q_b(t)$ amount of energy for transmission in order to maximize $q_d(t)\mu(P(t)) - \tilde{q}'_b(t)P(t)$ within $\Pi(t)$. Under this situation, we must have $I_o(t) = 1$ since $r(t-1) > 0$. Further, $P^*(t) \notin \Pi(t)$ means $P^*(t) \geq P(t)$. Moreover, when $I_o(t) = 1$, $P(t) = q_b(t)$ so that there is no missed replenishment since $r_{\text{max}} \leq B_b$, i.e., $C(t) = 0$, then $C(t)I_o(t) = 0$. Thus,

$$\begin{aligned} \Delta &\leq (1 + P_{\text{peak}})^2 + \eta_o^2 + \eta_o r_{\text{max}} + \mu_{\text{max}}^2 + A_{\text{max}}^2 + VR(t) - VR^*(t) + 2q_d(t)[R^*(t) - \mu(P^*(t))] + 2\tilde{q}'_b(t)[P^*(t) - r(t) + M^*(t)] + 2\tilde{q}'_b(t)M(t) + 2\beta\left(\frac{V}{2} + A_{\text{max}}\right)I_o(t) + 2C(t)(P(t) - r(t) + M(t)) + (V + 2A_{\text{max}})\mu_{\text{max}}I_o(t) + 2\tilde{q}'_b(t)C(t) - 2\tilde{q}'_b(t+1)C(t+1) + C^2(t) - C^2(t+1). \end{aligned} \quad (9)$$

Note that $q_b(t+1) = q_b(t) - P(t) + r(t) - M(t)$, we then have $C(t)(P(t) - r(t) + M(t)) = C(t)(q_b(t) - q_b(t+1))$. If $C(t) > 0$, then in time slot t , there is missed replenishment and PA goes to the modification part, and after adjustment,

$q_b(t+1) = q_b(t)$. Thus,

$$\begin{aligned} &C(t)(P(t) - r(t) + M(t)) \\ &= C(t)(q_b(t) - q_b(t+1)) = 0. \end{aligned} \quad (10)$$

Similarly, by $q_b^*(t+1) = q_b^*(t) - P^*(t) + r(t) - M^*(t)$ for the optimal policy, we have $\tilde{q}'_b(t)[P^*(t) - r(t) + M^*(t)] = \tilde{q}'_b(t)[q_b^*(t) - q_b^*(t+1)]$. Since the battery increment under the optimal policy is highly uncorrelated to the virtual battery queue state under our algorithm, we have $\liminf_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} \tilde{q}'_b(t)[q_b^*(t) - q_b^*(t+1)] = \liminf_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} \tilde{q}'_b(t) \liminf_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} [q_b^*(t) - q_b^*(t+1)] = O(\frac{1}{V})$. By summing from 0 to $T-1$, dividing by T and taking $\liminf_{T \rightarrow \infty}$, we have

$$\begin{aligned} &\liminf_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} \tilde{q}'_b(t)[P^*(t) - r(t) + M^*(t)] \\ &= O(\frac{1}{V}). \end{aligned} \quad (11)$$

Since $r_{\text{max}} \leq P_{\text{peak}}$, by the modification phase of PA , we have

$$\liminf_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} \tilde{q}'_b(t)M(t) = 0 \quad (12)$$

Note that $q_d^*(t+1) = (q_d^*(t) - \mu(P^*(t)))^+ + R^*(t) \geq q_d^*(t) - \mu(P^*(t)) + R^*(t)$. By multiplying both sides with $q_d(t)$ and rearranging terms, we obtain $q_d(t)(R^*(t) - \mu(P^*(t))) \leq q_d(t)(q_d^*(t+1) - q_d^*(t))$. With the optimal policy, q_d^* is strongly stable. By summing from 0 to $T-1$, dividing by T and taking $\liminf_{T \rightarrow \infty}$, we have

$$\begin{aligned} &\liminf_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} q_d(t)[R^*(t) - \mu(P^*(t))] \\ &\leq \liminf_{T \rightarrow \infty} \frac{q_d(T)q_d^*(T) - q_d(0)q_d^*(0)}{T} \\ &\quad + \liminf_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T q_d^*(t)(q_d(t-1) - q_d(t)) \\ &\leq A_{\text{max}} \liminf_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T q_d^*(t), \end{aligned} \quad (13)$$

which is a finite constant that is not related to V .

By summing from 0 to $T-1$, dividing by T and V , taking $\liminf_{T \rightarrow \infty}$ over Equation (9), combined with Equation (10), Equation (11), Equation (12) and Equation (13), we get

$$\begin{aligned} \liminf_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} R(t) &\geq \liminf_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} R^*(t) - \eta_o(\mu_{\text{max}} + \beta) \\ &\quad - O(\frac{1}{V}). \quad \blacksquare \end{aligned}$$