

**Joint Energy Allocation for Sensing and Transmission in Rechargeable
Wireless Sensor Networks**

by

Shaobo Mao

B.E., Zhejiang University, China, 2010

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF APPLIED SCIENCE

in

THE FACULTY OF GRADUATE STUDIES

(Electrical and Computer Engineering)

THE UNIVERSITY OF BRITISH COLUMBIA
(Vancouver)

July 2012

© Shaobo Mao, 2012

Abstract

The energy management policy of a rechargeable wireless sensor network (WSN) needs to take into account the energy harvesting process, and is thus different from that of a traditional WSN powered by non-rechargeable batteries. In this thesis, we study the energy allocation for sensing and transmission in an energy harvesting sensor node with a rechargeable battery. The sensor aims to maximize the expected total amount of data transmitted subject to time-varying energy harvesting rate, energy availability in the battery, data availability in the data buffer, and channel fading. In this thesis, we first consider the energy allocation problem that assumes a fixed sensor lifetime. Then, we extend the energy allocation problem by taking into account the randomness of the sensor lifetime.

In the first part of this thesis, we study the joint energy allocation for sensing and transmission in an energy harvesting sensor node with a fixed sensor lifetime. We formulate the energy allocation problem as a finite-horizon Markov decision process (MDP) and propose an optimal energy allocation (OEA) algorithm using backward induction. We conduct simulations to compare the performance between our proposed OEA algorithm and the channel-aware energy allocation (CAEA) algorithm extended from [1]. Simulation results show that the OEA algorithm can transmit a much larger amount of data over a finite horizon than the CAEA algorithm under different settings.

In the second part of this thesis, we extend the joint energy allocation problem

by taking into account the randomness of the sensor lifetime, and formulate the problem as an infinite-horizon discounted MDP. We propose an optimal stationary energy allocation (OSEA) algorithm using the value iteration. We then consider a special case with infinite data backlog and prove that the optimal transmission energy allocation (OTEA) policy is monotone with respect to the amount of battery energy available. Finally, we conduct extensive simulations to compare the performance of the OSEA, OTEA, and CAEA algorithms. Results show that the OSEA algorithm transmits the largest amount of data, and the OTEA algorithm can achieve a near-optimal performance.

Preface

Hereby, I declare that I am the first author of this thesis. Chapters 2-3 are based on work that has been published or submitted for publication. The related publications were done in collaboration with Dr. Man Hon Cheung and Prof. Vincent Wong.

For all publications, I conducted the paper survey on related topics, performed the analysis, and carried out the simulations of the considered communication systems. Dr. Man Hon Cheung checked the analytical model and the simulation codes. Prof. Vincent Wong gave important suggestions on the presentation of the papers. The papers were originally prepared by me, and further revised by all the co-authors. The following publications are accomplished through this research.

Journal Paper

- Shaobo Mao, Man Hon Cheung, and Vincent W.S. Wong, “Joint energy allocation for sensing and transmission in rechargeable wireless sensor networks,” submitted.

Conference Paper

- Shaobo Mao, Man Hon Cheung, and Vincent W.S. Wong, “An optimal energy allocation algorithm for energy harvesting wireless sensor networks,” in *Proc. of IEEE International Conference on Communications (ICC)*, Ottawa, Canada, June, 2012.

Table of Contents

| | |
|---|-----|
| Abstract | ii |
| Preface | iv |
| Table of Contents | v |
| List of Figures | vii |
| List of Acronyms | ix |
| Acknowledgments | x |
| 1 Introduction | 1 |
| 1.1 Wireless Sensor Network | 1 |
| 1.2 Energy Harvesting Technology | 2 |
| 1.2.1 Energy Harvesting Methods | 2 |
| 1.2.2 Energy Harvesting Architectures | 3 |
| 1.3 Energy Management in Energy Harvesting WSNs | 5 |
| 1.4 Motivations | 7 |
| 1.5 Contributions | 7 |
| 1.6 List of Publications | 9 |
| 1.7 Structure of the Thesis | 9 |

Table of Contents

| | | |
|----------|--|-----------|
| 2 | Energy Allocation Algorithm Based on Finite-horizon MDP . . . | 11 |
| 2.1 | System Model | 11 |
| 2.2 | Problem Formulation | 14 |
| 2.3 | Finite-Horizon MDP | 18 |
| 2.4 | Performance Evaluation | 21 |
| | | |
| 3 | Energy Allocation Algorithms Based on Infinite-horizon MDP . | 27 |
| 3.1 | Problem Formulation | 27 |
| 3.2 | Energy Allocation Algorithms | 30 |
| 3.2.1 | General Case | 30 |
| 3.2.2 | Special Case: Infinite Data Backlog | 34 |
| 3.3 | Performance Evaluation | 41 |
| | | |
| 4 | Conclusions and Future Work | 49 |
| 4.1 | Conclusions | 49 |
| 4.2 | Future Work | 50 |
| | | |
| | Bibliography | 52 |

List of Figures

| | | |
|-----|---|----|
| 1.1 | Energy harvesting architectures with and without energy storage capability. | 4 |
| 2.1 | The system model of an energy harvesting wireless sensor node transmitting data to the receiver \mathbf{R}_x of the sink. | 12 |
| 2.2 | Timing diagram of a Markov decision process (MDP). | 14 |
| 2.3 | A three-state Markov chain for the channel gain, where “ B ”, “ N ”, and “ G ” represent the channel in the bad, normal, and good states, respectively. | 21 |
| 2.4 | The total amount of data transmitted of the two algorithms for different number of total time slots K | 23 |
| 2.5 | The total amount of data transmitted of the two algorithms for different average energy harvesting rates when $K = 30$ | 24 |
| 2.6 | The total amount of data transmitted of the two algorithms for different values of data-sensing efficiency parameter γ | 26 |
| 3.1 | The total amount of data transmitted of OTEA algorithm under different percentage of energy allocated for sensing p . Since the OSEA algorithm does not allocate a fixed amount of energy for sensing, its total amount of data transmitted is independent of p | 42 |

| | | |
|-----|--|----|
| 3.2 | The optimal percentage of energy allocated for sensing under different data-sensing efficiency γ for the OTEA algorithm. | 43 |
| 3.3 | The total amount of data transmitted of the three algorithms for different average energy harvesting rates \bar{H} | 44 |
| 3.4 | The total amount of data transmitted of the OSEA algorithm and the OTEA algorithm for different values of data-sensing efficiency parameter γ | 45 |
| 3.5 | The total amount of data transmitted of the OSEA algorithm and the OTEA algorithm for different battery storage capacity b_{\max} | 46 |
| 3.6 | The total amount of data transmitted of the OSEA algorithm and the OTEA algorithm for different data buffer size q_{\max} | 47 |
| 3.7 | The total amount of data transmitted of the OSEA algorithm and the OTEA algorithm for different values of discount factor ν | 48 |

List of Acronyms

| | |
|-------|--|
| AWGN | Additive White Gaussian Noise |
| CAEA | Channel-Aware Energy Allocation |
| CSI | Channel State Information |
| MAC | Medium Access Control |
| MDP | Markov Decision Process |
| OEA | Optimal Energy Allocation |
| OSEA | Optimal Stationary Energy Allocation |
| OTEA | Optimal Transmission Energy Allocation |
| POMDP | Partially Observable Markov Decision Process |
| SNR | Signal-to-Noise Ratio |
| WSN | Wireless Sensor Network |

Acknowledgments

Foremost, I would like to express my deepest gratitude to my supervisor, Prof. Vincent Wong, for his continuous support of my graduate study and research, for his patience, motivation, enthusiasm, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better advisor and mentor for my Master study. I would also like to thank Prof. Victor Leung and Prof. Sathish Gopalakrishnan for their invaluable comments on my thesis.

I am heartily thankful to my colleague, Dr. Man Hon Cheung, who has provided me with precious assistance during my graduate research. I would also like to thank my colleagues in Prof. Wong's research group: Dr. Vahid Shah-Mansouri, Dr. Taesoo Kwon, Dr. Keivan Ronasi, Pedram Samadi, Binglai Niu, Enxin Yao, Bojiang Ma, Suyang Duan, as well as my colleagues in the Communications Lab: Dr. Derrick Wing Kwan Ng, Dr. Hu Jin, Peiran Wu, Jun Zhu, who have provided constructive suggestions on my work. Finally, I would like to express my profound gratitude to my beloved parents for their understanding, support, and endless love to me. To them I dedicate this thesis.

This research is supported by the Natural Sciences and Engineering Research Council (NSERC) of Canada.

Chapter 1

Introduction

This chapter introduces some background of wireless sensor network (WSN), energy harvesting technology, and energy management in rechargeable WSNs. The scope of the thesis is given at the end of this chapter.

1.1 Wireless Sensor Network

A WSN consists of a large number of spatially distributed sensor nodes, which have the capabilities of sensing, data processing, and communicating [2]. It can be deployed for remote environmental monitoring and target tracking, for example, volcano monitoring [3], habitat monitoring [4], vehicle tracking [5], and structural monitoring [6].

A wireless sensor node is typically equipped with three basic components, a sensing module for data acquisition from the surrounding environment, a processing module for local data processing and storage, and a wireless communication module for data transmission. Besides, a battery with limited energy budget supplies the energy needed by the device to perform tasks. In addition, an actuator may also be incorporated in a sensor node, depending on the application and the type of sensors used.

In the design of a WSN, there are several constraints, such as limited amount of energy due to a finite battery capacity, short communication range, low bandwidth,

and limited processing and storage in each sensor node. Among all these design constraints, the major limitation is that the sensor node can only operate for a limited amount of time due to the finite capacity of the battery. However, a sensor network should have a lifetime long enough to fulfill the requirements of the application. In many cases, the sensor network may be required to perform the task for several months, or even years. Therefore, how to prolong the lifetime of a WSN is a crucial question.

A lot of research efforts have been dedicated to prolong the lifetime of a WSN by improving its energy efficiency. Some of these include power-aware storage, energy-aware medium access control (MAC) protocols [7],[8],[9], routing protocols [10] [11], and duty-cycling strategies [12]. While all the techniques above optimize the energy consumption so as to maximize the lifetime of the sensor network, the lifetime remains bounded and finite, and thus the energy-related inhibitions are not precluded.

1.2 Energy Harvesting Technology

Recently, the idea of *energy harvesting* was proposed to address the problem of finite lifetime in a WSN by enabling the sensor nodes to replenish energy from ambient sources, for example, by using solar panels to convert sunlight into electricity, by using vibration-based energy harvesting technology, or by utilizing thermoelectric generators [13] [14] [15].

1.2.1 Energy Harvesting Methods

There are mainly three energy harvesting methods, as mentioned above, including *photonic method*, *vibrational method*, and *thermal method*.

- *Photonic method*: Silicon solar cells exploit the photovoltaic effect to convert sunlight into electricity. When the photons of sunlight strike the silicon cell, their energy may be absorbed and transferred to electrons of the silicon, which are then able to escape from their normal positions in the silicon to become part of the current in an electrical circuit. This phenomenon is called the photovoltaic effect. Since solar energy is a convenient harvesting source, lots of implementations of solar energy harvesting sensor nodes have already existed, for example, Helimote [16], Everlast [17], Prometheus [18], and HydroWatch [19].
- *Vibrational method*: Vibrations can generate electric energy. There are mainly three methods to harvest vibrations, including piezoelectric materials, inductive systems, and capacitive systems [15].
- *Thermal method*: The thermoelectric effect is the direct conversion of temperature differences to electric voltage. Thermoelectric devices utilize this effect and can generate electricity when there exists a temperature gradient across the device. Compared with vibration-based devices, thermoelectric devices can function for a much longer duration due to the absence of any moving parts.

1.2.2 Energy Harvesting Architectures

In general, energy harvesting architectures for sensor nodes can be divided into two categories, *harvest-use* architecture and *harvest-store-use* architecture [14], as shown in Figure 1.1.

- *Harvest-use architecture*: As shown in Figure 1.1 (a), in harvest-use architecture, the energy harvesting system powers the sensor node directly. Therefore,

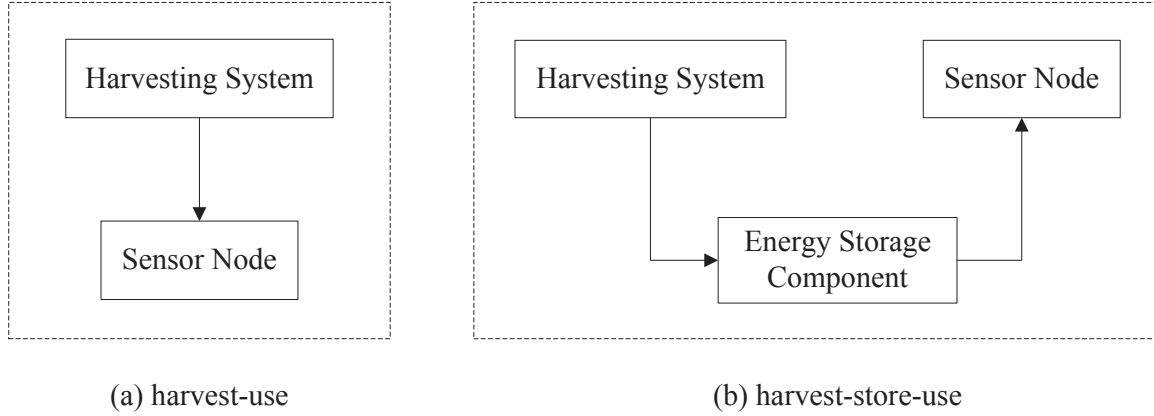


Figure 1.1: Energy harvesting architectures with and without energy storage capability.

in order to keep the sensor operational, the power output of the harvesting system must be continuously above the minimum operating point. Otherwise, the sensor node will be disabled.

- *Harvest-store-use architecture*: Figure 1.1 (b) depicts the harvest-store-use architecture, which has an additional energy storage component compared with the harvest-use architecture. The energy is harvested by the harvesting system and stored in the energy storage component. The energy storage would be quite useful when the harvested energy is more than the sensor's current need. The stored energy can either be used later when there is no harvesting opportunity or the energy usage of the sensor node has to be increased to improve the performance.

1.3 Energy Management in Energy Harvesting WSNs

The energy management of an energy harvesting WSN is different from a non-rechargeable battery powered WSN in many ways. First, with a potentially infinite amount of energy available to the sensor nodes, an energy harvesting WSN can remain functional for a long period of time. Hence, energy conservation is not the prime design issue. Second, the energy management strategy for an energy harvesting WSN needs to take into account the energy replenishment process. For example, an overly conservative energy expenditure may limit the amount of transmitted data by failing to take the full advantage of the energy harvesting process. On the other hand, an overly aggressive use of energy may result in an energy outage, which prevents some sensor nodes from functioning properly. Third, the energy availability constraint, which requires the energy consumption to be less than the energy stored in the battery, must be met at all time. This constraint complicates the design of an energy management policy, since the current energy consumption decision would affect the outcome in the future.

A lot of research efforts have been devoted recently to study the energy management and data transmission in energy harvesting WSNs. Kansal *et al.* in [20] proposed analytically tractable models to characterize the complex time varying nature of energy sources. Distributed algorithms were developed to utilize the harvested energy efficiently. Sharma *et al.* in [21] proposed energy management schemes for a single energy harvesting sensor node that achieves the maximum throughput and minimum mean delay. A greedy policy was shown to achieve both objectives in the low signal-to-noise ratio (SNR) regime. Gatzianas *et al.* in [22] presented an online

adaptive transmission scheme for wireless networks with rechargeable batteries that maximizes total system utility and stabilizes the data queue using Lyapunov techniques. Huang *et al.* in [23] proposed an online algorithm that achieves a close-to-optimal utility performance in finite capacity energy storage devices. The Lyapunov optimization techniques with weight perturbation were used. In [24], utility-optimal energy allocation algorithms were proposed for systems with predictable or stochastic energy availability.

References [25, 26] studied the transmission completion time minimization problem in energy harvesting wireless networks, and assumed that the energy harvesting times and harvested energy amounts were known before the transmission started. Yang *et al.* in [25] investigated two different scenarios of data arrivals and proposed optimal off-line scheduling policies. Antepi *et al.* in [26] considered the problem with an additive white Gaussian noise (AWGN) broadcast channel. The special structure in the problem was exploited, and an iterative off-line algorithm that minimizes the transmission completion time for the case of two-user broadcast channel was proposed.

Some of the recent works on energy harvesting WSNs have formulated the energy management problem as a Markov decision process (MDP) [27, 28]. Ho *et al.* in [1] proposed a throughput-optimal energy allocation algorithm for a time-slotted system under time-varying fading channel and energy source by using MDP. In [29], a throughput-optimal energy allocation policy was derived in a continuous time model and suboptimal online waterfilling schemes were proposed to address the dimensionality problem inherent in the MDP solution. Chen *et al.* in [30] studied the energy allocation problem of a single node using the shortest path approach. A simple distributed heuristic scheme was proposed that solves the joint energy allocation and

routing problem in a rechargeable WSN. Li *et al.* in [31] proposed energy efficient scheduling strategies for cooperative communications in energy harvesting WSNs to maximize the long-term utility. The scheduling problems under two different assumptions were formulated and solved using MDP and partially observable MDP (POMDP).

1.4 Motivations

Most of these results from [1, 29, 30, 21, 22, 23, 24] for energy management in energy harvesting WSNs only considered the special case that there is either an infinitely long data backlog or data buffer. Yet, it is more practical to consider a *finite data buffer*. Besides, the energy consumed in *data sensing* has always been overlooked in the literature. This motivates us to design an optimal energy allocation (OEA) algorithm for energy harvesting WSNs which takes into account both the data sensing energy consumption and the finite capacity of the data buffer. However, these considerations introduce new challenges. For instance, if the sensor node consumes an insufficient amount of energy for sensing but an excessive amount of energy for transmission, then the data buffer may be empty, which leads to a reduction in the total amount of data transmitted. Thus, the sensor node needs to maintain a good *balance* between the energy consumed for sensing and the energy for transmission.

1.5 Contributions

In this thesis, we consider the joint energy allocation algorithms design for sensing and transmission in energy harvesting WSNs. We consider a point-to-point wireless link between an energy harvesting sensor node and a sink. The channel and energy

harvesting rate may vary over time. The sensor node has a rechargeable battery and a data buffer with finite capacity. Our objective is to maximize the expected total amount of data transmitted. The sensor node needs to decide the amount of energy it should allocate for sensing and transmission in each time slot by taking into account the battery energy level, data buffer level, energy harvesting rate, and channel condition. In Chapter 2, we consider the case that the sensor lifetime is fixed. In Chapter 3, we take into account the randomness of the sensor lifetime.

The main contributions of this thesis are as follows:

- In Chapter 2, we study the energy allocation problem for sensing and transmission in an energy harvesting WSN over a finite horizon. The sensor lifetime is a fixed value. We formulate it as a finite-horizon MDP under channel fluctuations and energy variations in a time-slotted system. We obtain the optimal energy allocation policy and propose the OEA algorithm by using backward induction. We provide extensive simulation results to compare the performance of the OEA algorithm and the channel-aware energy allocation (CAEA) algorithm extended from [1]. The results show that the OEA algorithm can transmit a much larger amount of data over a finite horizon than the CAEA algorithm under different settings.
- In Chapter 3, we extend the joint energy allocation problem by taking into account the randomness of the sensor lifetime, and formulate the problem as an infinite-horizon discounted MDP. We obtain the optimal stationary energy allocation (OSEA) policy and propose the OSEA algorithm by using value iteration in MDP. We also study the transmission energy allocation problem under the assumption of infinite data backlog. We obtain structural results for the optimal transmission energy allocation (OTEA) policy, and prove that the OTEA

policy is a monotonically increasing function of the available battery energy. Finally, we provide extensive simulation results to compare the performance of the OSEA, OTEA, and CAEA algorithms. We study the impact of the average energy harvesting rate, the battery capacity, the data buffer size, the lifetime of the sensor node, and the data-sensing efficiency (i.e., the amount of data that the sensor can sense per unit energy) on the performance of total transmitted data. The results show that the OSEA algorithm transmits the maximum amount of total transmitted data among these three algorithms, and the OTEA algorithm can achieve a near-optimal performance.

1.6 List of Publications

The following publications have been completed based on the work in this thesis.

- Shaobo Mao, Man Hon Cheung, and Vincent W.S. Wong, “An optimal energy allocation algorithm for energy harvesting wireless sensor networks,” in *Proc. of IEEE International Conference on Communications (ICC)*, Ottawa, Canada, June, 2012.
- Shaobo Mao, Man Hon Cheung, and Vincent W.S. Wong, “Joint energy allocation for sensing and transmission in rechargeable wireless sensor networks,” submitted to *IEEE Transactions on Wireless Communications*, 2012.

1.7 Structure of the Thesis

The rest of the thesis is organized as follows: In Chapter 2, we present the energy allocation algorithm design in energy harvesting WSNs over a finite horizon. In

Chapter 3, we extend the energy allocation problem in Chapter 2 by taking into account the randomness of the sensor lifetime. Conclusions and future work are given in Chapter 4.

Chapter 2

Energy Allocation Algorithm Based on Finite-horizon MDP

In this chapter, we present the energy allocation algorithms design for sensing and transmission in an energy harvesting sensor node with a rechargeable battery and a finite data buffer. We formulate the energy allocation problem as a finite-horizon MDP and solve it by using backward induction.

2.1 System Model

As shown in Figure 2.1, we consider a single energy harvesting sensor node, which contains a rechargeable battery with capacity b_{max} Joule and a data buffer with size q_{max} Mbits. We assume that the system is time-slotted with K time slots and the duration of a time slot is τ sec. We let $k \in \mathcal{K} \triangleq \{0, 1, \dots, K-1\}$ be the time slot index. The sensor node performs sensing in the field, stores the sensed data in the buffer, and transmits the data to the receiver $\mathbf{R}\mathbf{x}$ of the sink over a wireless channel. We consider an AWGN channel with block flat fading. That is, the channel remains constant for the duration of each time slot, but may change at the slot boundaries. Let α_k be the channel gain in time slot k . We assume that the data transmission at every time slot is successful, which is reasonable since we can apply proper channel coding techniques.

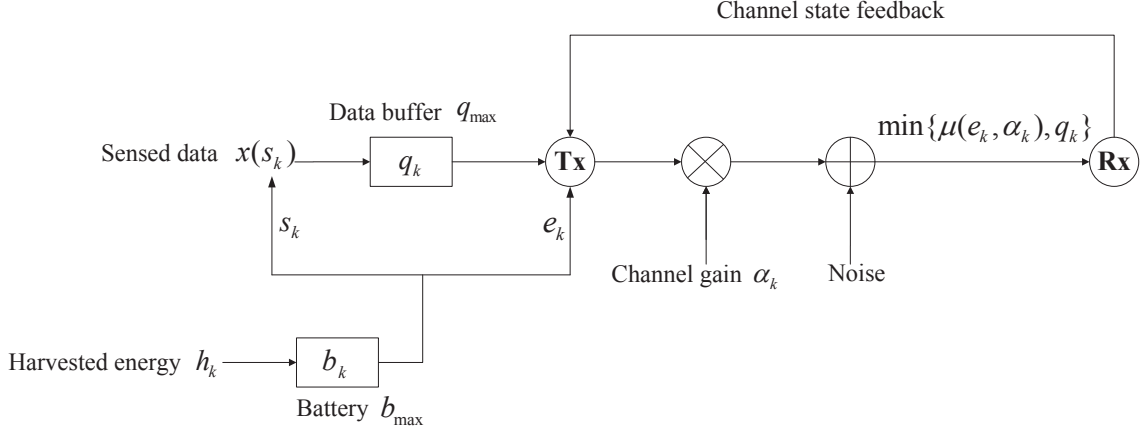


Figure 2.1: The system model of an energy harvesting wireless sensor node transmitting data to the receiver \mathbf{Rx} of the sink.

We assume that the sink sends *delayed* channel state information (CSI) of the previous time slot back to the sensor node. At the beginning of time slot k , the sensor node only knows the value of α_{k-1} , but not α_k . The stored battery level is b_k and the amount of stored data in the data buffer is q_k . During the whole time slot k , the sensor node is able to replenish energy by h_k , which can be used for sensing or transmission in time slot $k+1$ onward. As a result, the sensor node does not know the value of h_k until the beginning of the next time slot $k+1$. In other words, at the beginning of time slot k , the sensor node knows the value of h_{k-1} , but not h_k .

If the channel gain is α_k and the allocated transmission energy is e_k in time slot k , then the instantaneous transmission power is $\frac{e_k}{\tau}$. We consider that the sensor node is able to transmit $\mu(e_k, \alpha_k)$ bits of data in time slot k , where $\mu(e_k, \alpha_k)$ is a monotonically non-decreasing and concave function in e_k given α_k in general. One such function is given by [32, pp.172]:

$$\mu(e_k, \alpha_k) = \tau W \log_2 \left(1 + \frac{\alpha_k e_k}{N_0 W \tau} \right) \text{ bits}, \quad (2.1)$$

where N_0 is the power spectral density of the Gaussian noise, and W is the bandwidth of the channel.

For sensing in time slot k , we let $x(s_k)$ be the amount of data generated when s_k units of energy are used for sensing. In general, $x(s_k)$ is a monotonically non-decreasing and concave function in s_k . The data obtained by sensing in time slot k will be stored in the data buffer until they are transmitted in the subsequent time slots. Except for sensing and transmission, we assume that other energy consumptions in the sensor node, for example, the processing energy, the energy consumed for storing information in the memory, the energy for turning on and off the transmitter, the energy consumed for receiving feedback from the sink, the energy leakage from the battery, and battery relaxation effects [33], are negligible.

At time slot k , the sensor node needs to choose e_k and s_k , for all $k \in \mathcal{K}$ such that the expected total amount of data transmitted is maximized. To achieve this goal, the sensor has to maintain a good *tradeoff* between the energy allocation for e_k and s_k . Given a fixed energy budget in a time slot, if e_k is too small, then the transmitted data in time slot k will be small. However, if e_k is too large, then s_k will be small such that insufficient amount of sensing data is stored in the buffer for transmission in the next time slot, which leads to a reduction in the amount of data transmitted in future time slots. In addition, the total energy budget $e_k + s_k$ in time slot k should also be carefully controlled. If the energy management policy is overly aggressive such that the rate of energy consumption is greater than the rate of energy harvesting, the sensor node may stop functioning because of the energy outage. On the other hand, an overly conservative energy management policy would limit the amount of data transmitted in each time slot. Thus, it is a challenging problem to decide the values of e_k and s_k optimally in each time slot $k \in \mathcal{K}$.

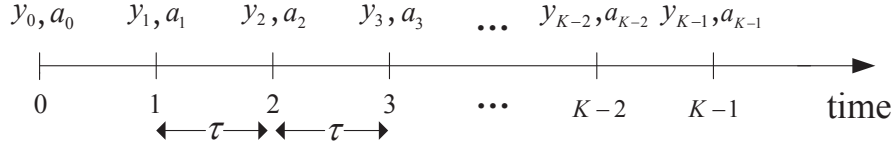


Figure 2.2: Timing diagram of a Markov decision process (MDP).

2.2 Problem Formulation

In this section, we formulate the problem of finding the optimal energy allocation for sensing and transmission as an MDP [27] [28], which consists of five elements: decision epochs, states, actions, state transition probabilities, and rewards. Referring to Figure 2.2, the *decision epochs* are $k \in \mathcal{K} = \{0, 1, \dots, K-1\}$. The *state* of the system is denoted as

$$\mathbf{y} = (b, q, h, \alpha),$$

which includes the battery energy state b and data buffer state q for the *current* time slot, as well as the energy harvesting state h and channel state α in the *previous* time slot. We denote the state space as $Y = \mathcal{B} \times \mathcal{Q} \times \mathcal{H} \times \mathcal{A}$, where \mathcal{B} is the set of battery energy states, \mathcal{Q} is the set of data buffer states, \mathcal{H} is the set of energy harvesting states, and \mathcal{A} is the set of channel states. Let \mathbf{y}_k denote the state of the system at time slot k , i.e.,

$$\mathbf{y}_k = (b_k, q_k, h_{k-1}, \alpha_{k-1}).$$

First, for the battery energy state in time slot k , the sensor node harvests h_k units of energy from the environment. On the other hand, it consumes e_k units of energy for data transmission and s_k units of energy for sensing. Since the battery has a finite capacity b_{\max} , the energy stored in the battery is updated as

$$b_{k+1} = \min\{b_k - (e_k + s_k) + h_k, b_{\max}\}, \quad \forall k \in \mathcal{K}, \quad (2.2)$$

such that the battery energy state transition probability is given by

$$P(b_{k+1} | b_k, h_k, e_k, s_k) = \begin{cases} 1, & \text{if (2.2) is satisfied,} \\ 0, & \text{otherwise.} \end{cases} \quad (2.3)$$

Equation (2.2) ensures that the maximum stored energy b_{\max} is not exceeded. We assume that the initial energy b_0 is known and satisfies the constraint $0 \leq b_0 \leq b_{\max}$. Moreover, the amount of energy consumed for sensing and transmission must be no more than the battery level:

$$e_k + s_k \leq b_k, \quad \forall k \in \mathcal{K}. \quad (2.4)$$

Second, for the data buffer state in time slot k , $x(s_k)$ amount of sensing data is generated and queued up in the data buffer if s_k units of energy are allocated for sensing. On the other hand, if the amount of data available in the data buffer for transmission at time slot k is q_k , and e_k units of energy are used for transmission, then the amount of data transmitted and removed from the data buffer at time slot k is given by $\min\{\mu(e_k, \alpha_k), q_k\}$. Since the data buffer is finite with capacity q_{\max} , the amount of data in the buffer is then updated as

$$q_{k+1} = \min\{[q_k - \mu(e_k, \alpha_k)]^+ + x(s_k), q_{\max}\}, \quad \forall k \in \mathcal{K}, \quad (2.5)$$

where $[z]^+ = \max\{z, 0\}$. The data buffer state transition probability is then given by

$$P(q_{k+1} | q_k, \alpha_k, e_k, s_k) = \begin{cases} 1, & \text{if (2.5) is satisfied,} \\ 0, & \text{otherwise.} \end{cases} \quad (2.6)$$

We assume that the initial amount of data in the data buffer q_0 is known and satisfies $0 \leq q_0 \leq q_{\max}$. Equation (2.5) implies that if the sensor allocates too much energy for transmission such that $\mu(e_k, \alpha_k) > q_k$, then energy will be wasted. On the other hand, if the sensor allocates too much energy for sensing so that $x(s_k) > q_{\max}$, then the data buffer will be overflowed and energy will be wasted too. Thus the sensor should make a proper energy allocation decision at each time slot. Third, since the energy harvesting rate and the current channel state information at time slot k is not known to the sensor, we use two independent first-order stationary Markovian models to model h_k and α_k [1, 31, 34]. The transition probability of these two independent random variables are denoted as $P(h_k | h_{k-1})$ and $P(\alpha_k | \alpha_{k-1})$.

Based on the current state \mathbf{y}_k at time slot k , the sensor will choose to consume e_k units of energy for data transmission and s_k units of energy for sensing. That is, an *action* $\mathbf{a}_k = (e_k, s_k)$ is taken for transmission and sensing energy allocation from its feasible set $U(\mathbf{y}_k)$. We have

$$\mathbf{a}_k \in U(\mathbf{y}_k) = \{(e, s) \mid e + s \leq b_k, e \geq 0, s \geq 0\}, \quad (2.7)$$

where $U(\mathbf{y}_k)$ represents the feasible set of the action \mathbf{a}_k given the current state \mathbf{y}_k at time slot k . The constraint $e_k + s_k \leq b_k, \forall k \in \mathcal{K}$ ensures that the amount of energy consumed for sensing and transmission must be no more than the battery level. In addition, it is possible to impose additional constraints on \mathbf{a}_k . For example, a constraint on the minimum amount of energy for sensing or transmission to ensure a minimum amount of sensed data or transmitted data for each time slot, respectively. Also, a maximum transmission power constraint can be imposed.

The *state transition probability* $P(\mathbf{y}_{k+1} | \mathbf{y}_k, \mathbf{a}_k)$ is the probability that the system will go into state \mathbf{y}_{k+1} if action \mathbf{a}_k is taken at state \mathbf{y}_k at time slot k . Due to the

independence between (b_{k+1}, h_k) and (q_{k+1}, α_k) for all $k \in \mathcal{K}$, we can simplify the state transition probability as

$$\begin{aligned}
 P(\mathbf{y}_{k+1} | \mathbf{y}_k, \mathbf{a}_k) &= P(b_{k+1}, q_{k+1}, h_k, \alpha_k | b_k, q_k, h_{k-1}, \alpha_{k-1}, e_k, s_k) \\
 &= P(b_{k+1}, h_k | b_k, h_{k-1}, e_k, s_k) P(q_{k+1}, \alpha_k | q_k, \alpha_{k-1}, e_k, s_k) \\
 &= P(b_{k+1} | b_k, h_k, e_k, s_k) P(h_k | h_{k-1}) P(q_{k+1} | q_k, \alpha_k, e_k, s_k) P(\alpha_k | \alpha_{k-1}),
 \end{aligned} \tag{2.8}$$

where $P(b_{k+1} | b_k, h_k, e_k, s_k)$ and $P(q_{k+1} | q_k, \alpha_k, e_k, s_k)$ are defined in (2.3) and (2.6), respectively.

Given the current state \mathbf{y}_k and the action \mathbf{a}_k , $\mathbb{E}_{\alpha_k}[\mu(e_k, \alpha_k) | \alpha_{k-1}]$ is the expected amount of data that can be transmitted when e_k units of energy are used for transmission. However, since the data available in the data buffer for transmission at time slot k are q_k , the expected amount of data transmitted at time slot k is given by $\mathbb{E}_{\alpha_k}[\min\{\mu(e_k, \alpha_k), q_k\} | \alpha_{k-1}]$. We define the *reward* at time slot k , $r(\mathbf{y}_k, \mathbf{a}_k)$ to be the expected amount of data transmitted at time slot k . That is,

$$r(\mathbf{y}_k, \mathbf{a}_k) = \mathbb{E}_{\alpha_k}[\min\{\mu(e_k, \alpha_k), q_k\} | \alpha_{k-1}]. \tag{2.9}$$

A *decision rule* prescribes a procedure for action selection in each state at a specified time slot. We denote the deterministic Markovian decision rule at time slot k as δ_k , i.e., $\mathbf{a}_k = \delta_k(\mathbf{y}_k)$, which specifies the action choice \mathbf{a}_k when the system occupies state \mathbf{y}_k at time slot k . A *policy* $\pi = (\delta_0, \delta_1, \dots, \delta_{K-1})$ is a sequence of decision rules to be used at all the time slots. A *feasible* policy should satisfy (2.7) at all the time slots. Let Π be the feasible set of π . The sensor node aims to find an optimal and feasible sensing and transmit energy allocation policy π^* that maximizes the *expected total reward* for all the K time slots. That is, for any given initial state

$\mathbf{y}_0 = (b_0, q_0, h_{-1}, \alpha_{-1})$ at the first time slot, the *optimal expected total reward* is given by

$$\begin{aligned}\mathcal{T}^* &= \max_{\pi \in \Pi} \sum_{k=0}^{K-1} \mathbb{E} \left\{ r(\mathbf{y}_k, \mathbf{a}_k) \mid \mathbf{y}_0, \pi \right\} \\ &= \max_{\pi \in \Pi} \sum_{k=0}^{K-1} \mathbb{E} \left\{ \min \{ q_k, \mu(e_k, \alpha_k) \} \mid \mathbf{y}_0, \pi \right\},\end{aligned}\tag{2.10}$$

where $\mathbb{E}\{\cdot\}$ denotes the statistical expectation taken over all relevant random variables given initial state \mathbf{y}_0 and policy π . It should be noted that with a different policy π and initial state \mathbf{y}_0 , a different action will be chosen in each time slot in general, which results in a different state transition probability when the expectation $\mathbb{E}\{\cdot\}$ is computed.

2.3 Finite-Horizon MDP

In this section, we solve problem (2.10) by using finite-horizon MDP. An OEA algorithm is proposed that can transmit the maximal total amount of data in problem (2.10).

Let $V_k(b_k, q_k, h_{k-1}, \alpha_{k-1})$ be the maximum expected amount of data transmitted from time slot k to $K-1$, given that the system is in state $(b_k, q_k, h_{k-1}, \alpha_{k-1})$ immediately before the decision at time slot k . The *Bellman's equations* are given by the

following recursive equations starting from $k = K - 1$ to $k = 0$.

For $k = K - 1$, we have

$$V_{K-1}(b_{K-1}, q_{K-1}, h_{K-2}, \alpha_{K-2}) = \max_{\mathbf{a}_{K-1} \in U(\mathbf{y}_{K-1})} \mathbb{E}_{\alpha_{K-1}} \{ \min\{\mu(e_{K-1}, \alpha_{K-1}), q_{K-1}\} \mid \alpha_{K-2} \}. \quad (2.11a)$$

For $k = K - 2, \dots, 0$, we have

$$\begin{aligned} V_k(b_k, q_k, h_{k-1}, \alpha_{k-1}) = & \max_{\mathbf{a}_k \in U(\mathbf{y}_k)} \left\{ \mathbb{E}_{\alpha_k} \{ \min\{\mu(e_k, \alpha_k), q_k\} \mid \alpha_{k-1} \} \right. \\ & \left. + \mathbb{E}_{h_k, \alpha_k} \{ V_{k+1}(b_{k+1}, q_{k+1}, h_k, \alpha_k) \mid h_{k-1}, \alpha_{k-1} \} \right\}, \end{aligned} \quad (2.11b)$$

where b_{k+1} and q_{k+1} are updated as in (2.2) and (2.5), respectively. Notice that if the feasible set of \mathbf{a}_k is $U(\mathbf{y}_k)$ as defined in (2.7), then (2.11a) can be simplified as

$$V_{K-1}(b_{K-1}, q_{K-1}, h_{K-2}, \alpha_{K-2}) = \mathbb{E}_{\alpha_{K-1}} \{ \min\{\mu(b_{K-1}, \alpha_{K-1}), q_{K-1}\} \mid \alpha_{K-2} \}. \quad (2.12)$$

That is, we use all the available energy for transmission in the final time slot. Thus the optimal energy allocation for the final time slot $K - 1$ is $(e_{K-1}^*, s_{K-1}^*) = (b_{K-1}, 0)$. For (2.11b), the first and second terms on the right hand side represent, respectively, the expected *immediate* reward for time slot k and the expected total *future* rewards for time slot $k + 1$ to $K - 1$ if action \mathbf{a}_k is chosen. Hence, the equation in (2.11b) describes the *tradeoff* between the current reward and the future rewards.

Theorem 2.1. *The optimal policy of problem (2.10) is $\pi^* = \{\mathbf{a}_k^*(\mathbf{y}_k), \forall \mathbf{y}_k, k \in \mathcal{K}\}$, where*

$$\begin{aligned} \mathbf{a}_k^*(\mathbf{y}_k) = & \arg \max_{\mathbf{a}_k \in U(\mathbf{y}_k)} \left\{ \mathbb{E}_{\alpha_k} \{ \min\{\mu(e_k, \alpha_k), q_k\} \mid \alpha_{k-1} \} \right. \\ & \left. + \mathbb{E}_{h_k, \alpha_k} \{ V_{k+1}(b_{k+1}, q_{k+1}, h_k, \alpha_k) \mid h_{k-1}, \alpha_{k-1} \} \right\}. \end{aligned} \quad (2.13)$$

Moreover, for every initial state $\mathbf{y}_0 = (b_0, q_0, h_{-1}, \alpha_{-1})$, the maximum amount of transmitted data \mathcal{T}^* is given by $V_0(b_0, q_0, h_{-1}, \alpha_{-1})$.

Proof. The proof follows by applying the Bellman's equations and backward induction [27] and using (2.2) and (2.5). \square

Algorithm 1 Optimal Energy Allocation (OEA) Algorithm for Energy Harvesting Sensor Node.

- 1: Planning Phase:
 - 2: Set $V_{K-1}(b_{K-1}, q_{K-1}, h_{K-2}, \alpha_{K-2}), \forall b_{K-1}, \forall q_{K-1}, \forall h_{K-2}, \forall \alpha_{K-2}$, using (2.11a).
 - 3: Set $k := K - 2$.
 - 4: **while** $k \geq 0$ **do**
 - 5: Calculate $V_k(b_k, q_k, h_{k-1}, \alpha_{k-1}), \forall b_k, \forall q_k, \forall h_{k-1}, \forall \alpha_{k-1}$, using (2.11b).
 - 6: Find the optimal action $\mathbf{a}_k^*(\mathbf{y}_k) := (e_k^*(\mathbf{y}_k), s_k^*(\mathbf{y}_k))$, using (2.13).
 - 7: Set $k := k - 1$.
 - 8: **end while**
 - 9: Sensing and Transmission Phase:
 - 10: Set $k := 0$.
 - 11: **while** $k \leq K - 1$ **do**
 - 12: Track the energy harvesting rate of the previous time slot h_{k-1} .
 - 13: Track the energy available for use in the battery b_k .
 - 14: Track the amount of data in the buffer q_k .
 - 15: Obtain the channel feedback α_{k-1} from the sink.
 - 16: Set $\mathbf{y}_k := (b_k, q_k, h_{k-1}, \alpha_{k-1})$.
 - 17: Obtain $\mathbf{a}_k^*(\mathbf{y}_k) := (e_k^*(\mathbf{y}_k), s_k^*(\mathbf{y}_k))$ based on optimal policy π^* .
 - 18: Consume $e_k^*(\mathbf{y}_k)$ amount of energy for transmission and $s_k^*(\mathbf{y}_k)$ amount of energy for sensing.
 - 19: Update battery energy b_{k+1} by using (2.2) and the amount of data in the buffer q_{k+1} by using (2.5).
 - 20: Set $k := k + 1$.
 - 21: **end while**
-

We then propose our OEA algorithm in Algorithm 1. In the planning phase, the sensor solves for the optimal policy π^* and records it as a look-up table. In the sensing and transmission phase, the sensor first tracks the energy harvesting rate of the previous time slot h_{k-1} , the battery energy level b_k , the amount of data in the buffer q_k , and obtains the channel feedback α_{k-1} from the sink. Then, the sensor chooses the action $\mathbf{a}_k^* = (e_k^*, s_k^*)$ based on current system state \mathbf{y}_k and the optimal

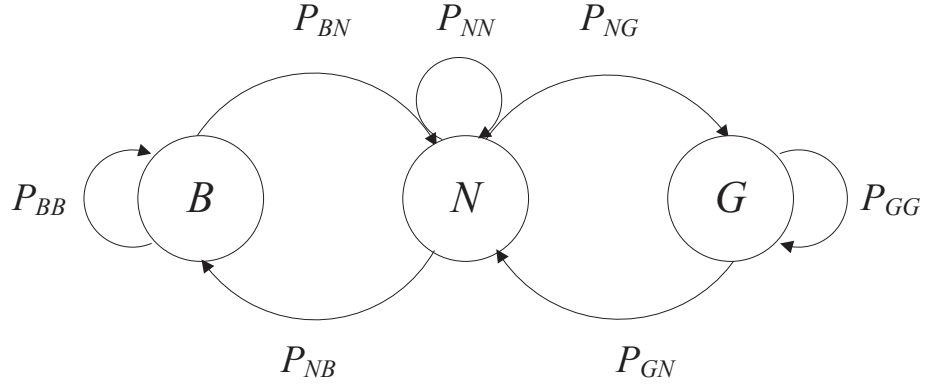


Figure 2.3: A three-state Markov chain for the channel gain, where “B”, “N”, and “G” represent the channel in the bad, normal, and good states, respectively.

policy π^* . That is, it consumes e_k^* and s_k^* amount of energy for transmission and sensing, respectively.

2.4 Performance Evaluation

In this section, we simulate the performance of our OEA and CAEA algorithms in terms of the total amount of data transmitted in Matlab, where the CAEA algorithm is extended based on the algorithm proposed in [1]. We consider a band-limited AWGN channel, where the channel bandwidth is $W = 100$ KHz and the noise power spectral density is $N_0 = 10^{-18}$ W/Hz. The channel state can be “G = Good”, “N = Normal”, or “B = Bad”. It evolves according to the three-state Markov chain as shown in Figure 2.3 [35] with the transition matrix of the Markov chain given by

$$P_\alpha = \begin{bmatrix} P_{BB} & P_{BN} & P_{BG} \\ P_{NB} & P_{NN} & P_{NG} \\ P_{GB} & P_{GN} & P_{GG} \end{bmatrix} = \begin{bmatrix} 0.3 & 0.7 & 0 \\ 0.25 & 0.5 & 0.25 \\ 0 & 0.7 & 0.3 \end{bmatrix}, \quad (2.14)$$

where P_{XZ} represents the probability of the channel state going from state X to

state Z , X and $Z \in \{B, N, G\}$. The channel gain α is 0.5×10^{-13} , 1×10^{-13} , and 1.5×10^{-13} when the channel state is “Bad”, “Normal”, and “Good”, respectively. The battery buffer size b_{max} is set to be 100 Joules, and the data buffer size q_{max} is set to be 1 Mbits. For tractability, we assume that the energy harvesting state h_k takes values from the finite set $\mathcal{H} = \{H_1, H_2, H_3, H_4\}$ and evolves according to the four-state *Markov chain* with the state transition probability given by

$$P_h = \begin{bmatrix} P_{H_1H_1} & P_{H_1H_2} & P_{H_1H_3} & P_{H_1H_4} \\ P_{H_2H_1} & P_{H_2H_2} & P_{H_2H_3} & P_{H_2H_4} \\ P_{H_3H_1} & P_{H_3H_2} & P_{H_3H_3} & P_{H_3H_4} \\ P_{H_4H_1} & P_{H_4H_2} & P_{H_4H_3} & P_{H_4H_4} \end{bmatrix} = \begin{bmatrix} 0.3 & 0.7 & 0 & 0 \\ 0.25 & 0.5 & 0.25 & 0 \\ 0 & 0.25 & 0.5 & 0.25 \\ 0 & 0 & 0.7 & 0.3 \end{bmatrix}, \quad (2.15)$$

where $P_{H_iH_j}$ represents the probability of the energy harvesting state going from state H_i to state H_j , $\forall i, j \in \{1, 2, 3, 4\}$. The steady state probability is then given by $[P_{H_1} P_{H_2} P_{H_3} P_{H_4}] = [0.13 \ 0.37 \ 0.37 \ 0.13]$. $x(s_k)$ is assumed to be a linear function of s_k [30], given by

$$x(s_k) = \gamma s_k, \quad (2.16)$$

where γ is the data-sensing efficiency parameter (i.e., the amount of data that the sensor can sense per unit energy). Unless specified otherwise, we assume that γ is equal to be 0.02 Mbits/J.

The CAEA algorithm in [1] assumed infinite backlogged data and neglected the sensing energy. For fair comparison, we modify the CAEA algorithm by allowing the data buffer to be finite with size q_{max} . We assume that the sensor allocates a fixed percentage of energy available in the battery for sensing in each time slot, and optimizes the energy allocated for transmission to achieve the maximal amount of total transmitted data.

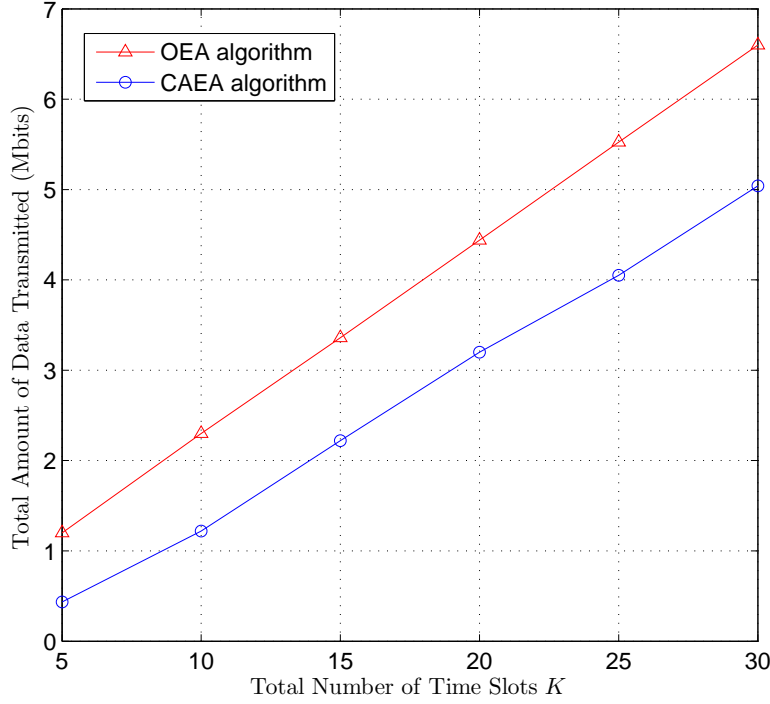


Figure 2.4: The total amount of data transmitted of the two algorithms for different number of total time slots K .

We start by examining the total amount of transmitted data of the OEA algorithm and the CAEA algorithm with different number of total time slots K . We set the fixed percentage of energy for sensing to be 10% in the CAEA algorithm, which is reasonable in WSNs. The value of energy harvesting rate is taken from the set $\mathcal{H} = \{H_1, H_2, H_3, H_4\} = \{6, 12, 18, 24\}$ J/time slot. As shown in Figure 2.4, our proposed OEA algorithm outperforms the CAEA algorithm in terms of the achieved amount of transmitted data. The reason is that in the CAEA algorithm, the sensor just optimally controls the energy for transmission, while the sensing energy is fixed. However, in our OEA algorithm, both the sensing and transmitting energy is optimally allocated, which results in a better performance than the CAEA algorithm.

Next, we consider the performance of the two algorithms under different average

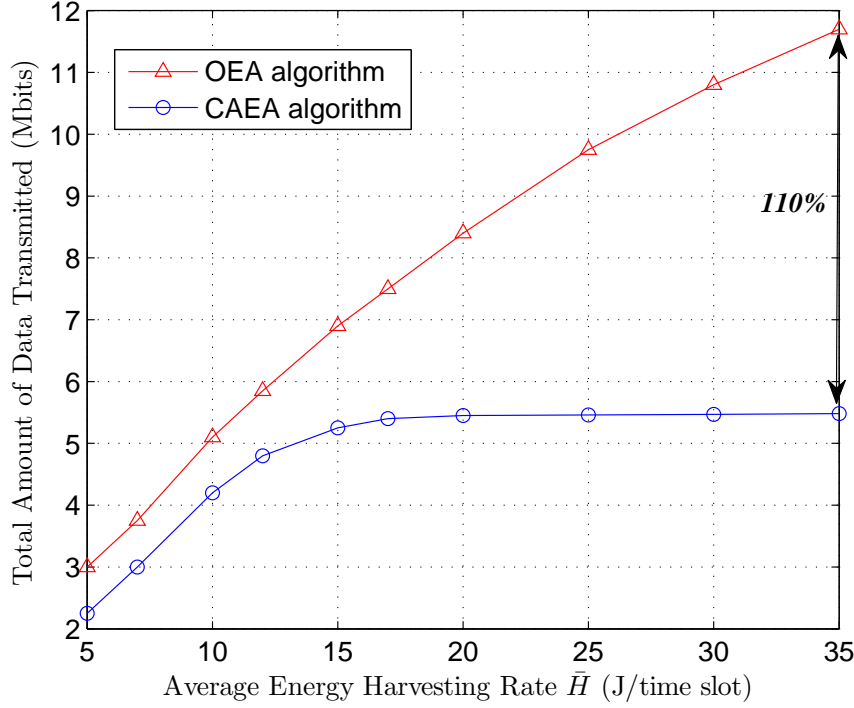


Figure 2.5: The total amount of data transmitted of the two algorithms for different average energy harvesting rates when $K = 30$.

energy harvesting rates \bar{H} , where $\bar{H} = \sum_{i=1}^4 H_i P_{H_i}$. In Figure 2.5, we plot the total amount of transmitted data against the average energy harvesting rate when the total number of time slots $K = 30$. We observe that our OEA algorithm performs much better than the CAEA algorithm, especially when the average energy harvesting rate \bar{H} is high. As shown in Figure 2.5, our OEA algorithm achieves 110% larger amount of transmitted data than the CAEA algorithm when the average energy harvesting rate $\bar{H} = 35$ J/time slot. Moreover, the performance of the CAEA algorithm saturates very quickly as the average harvesting rate is increased. It is because the harvested energy cannot be accommodated, and more and more energy is lost due to the overflow of the battery energy. However, in our algorithm, energy wastage will not occur as long as the harvesting rate is less than b_{max} and the data buffer is large

enough. The reason is that under the OEA algorithm, the sensor node maintains a good balance between the energy allocated for sensing and transmission, and thus achieves a better performance.

Finally, we study the impact of the data-sensing efficiency on the amount of transmitted data. We fix K to be 30 and examine the amount of total transmitted data under different values of γ . A larger value of γ corresponds to a higher data-sensing efficiency, since the sensor node spends less energy for sensing the same amount of data. As shown in Figure 2.6, when γ is increased, the amount of transmitted data increases as well, because more energy is available for data transmission. However, the performance saturates as γ is increased beyond a certain value. When γ approaches infinity, it corresponds to the case where the sensing is extremely efficient. The throughput of this case provides an upper bound for the performance of the OEA algorithm for sensor nodes with different sensing efficiency.

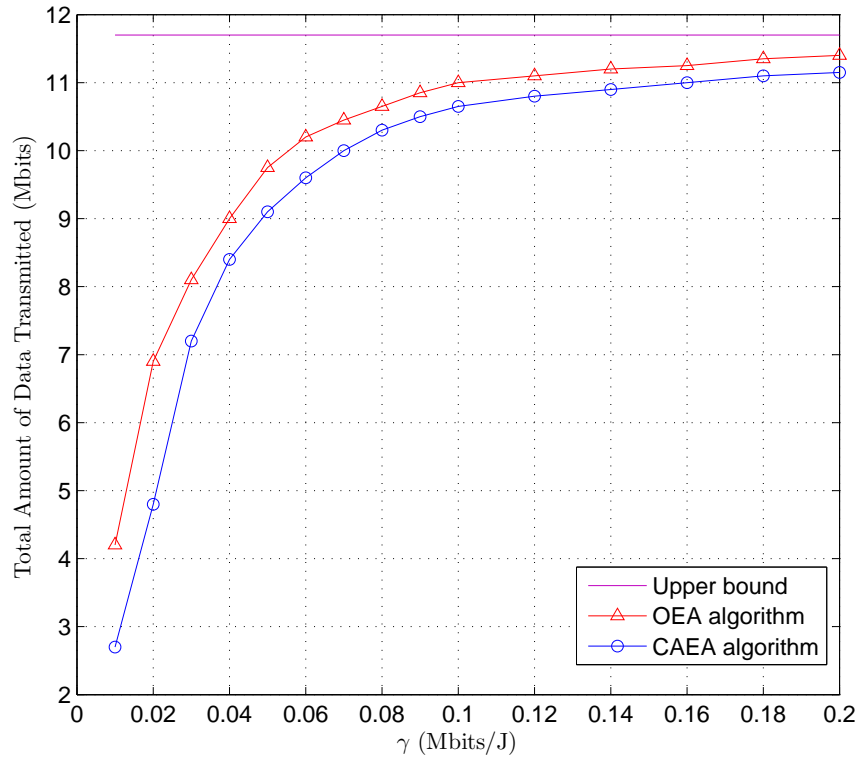


Figure 2.6: The total amount of data transmitted of the two algorithms for different values of data-sensing efficiency parameter γ .

Chapter 3

Energy Allocation Algorithms Based on Infinite-horizon MDP

In Chapter 2, we considered the case that the lifetime of the sensor node is fixed. The joint energy allocation problem was formulated as a finite-horizon MDP, and was solved by using backward induction. In this chapter, we consider a different setting, where the lifetime of the sensor node is a random variable, and formulate the energy allocation problem as an infinite-horizon discounted MDP. We solve the infinite-horizon MDP by using the value iteration algorithm.

3.1 Problem Formulation

The system model in this chapter is the same as the model in Chapter 2, as shown in Figure 2.1. The sensor node can function for K time slots and will stop functioning after that time. We take into account the randomness of the sensor node lifetime K and assume that K is geometrically distributed with mean $1/(1-\nu)$, where $\nu \in [0, 1)$. We apply the same notations as in Chapter 2. The sensor aims to maximize the total amount of data transmitted from the first time slot to the time that the sensor stops functioning. The setting of the sensor lifetime K and the objective to maximize the total amount of transmitted data are reasonable in many applications. For example, on the battle field, a camera with a wireless transceiver in the enemy territory, which

is used to monitor the action of the enemy, can be detected by the enemy at any time. Thus, its objective is to sense and transmit as much information as possible before it is detected and destroyed by the enemy. Another application is that nowadays in Europe, sensor nodes are placed in the forest to monitor the forest fire. These sensors are aimed to gather and transmit as much information as possible before they are damaged by the fire.

Considering that the lifetime of the sensor node is a random variable, for any given state $\mathbf{y}_0 = (b_0, q_0, h_{-1}, \alpha_{-1})$ at the first time slot, the *expected total reward* between the first time slot till the sensor stops functioning with policy $\pi \in \Pi$ is given by

$$J^\pi(\mathbf{y}_0) = \mathbb{E} \left\{ \mathbb{E}_K \left\{ \sum_{k=0}^{K-1} r(\mathbf{y}_k, \mathbf{a}_k) \right\} \middle| \mathbf{y}_0, \pi \right\}, \quad (3.1)$$

where $\mathbb{E}\{\cdot\}$ denotes the statistical expectation taken over all relevant random variables given initial state \mathbf{y}_0 and policy π . $\mathbb{E}_K\{\cdot\}$ denotes the expectation with respect to the random variable K , which is the lifetime of the sensor node.

Lemma 3.1. *Based on the geometric distribution of the lifetime K of the sensor node with mean $1/(1 - \nu)$, equation (3.1) is equivalent to the objective function of infinite-horizon MDP with discounted reward given by*

$$J^\pi(\mathbf{y}_0) = \mathbb{E} \left\{ \sum_{k=0}^{\infty} \nu^k r(\mathbf{y}_k, \mathbf{a}_k) \middle| \mathbf{y}_0, \pi \right\}. \quad (3.2)$$

Proof. Since K is distributed as

$$P(K = m) = \nu^{m-1}(1 - \nu), \quad m = 1, 2, 3, \dots,$$

equation (3.1) can be written as

$$\begin{aligned}
 J^\pi(\mathbf{y}_0) &= \mathbb{E} \left\{ \sum_{K=1}^{\infty} \sum_{k=0}^{K-1} r(\mathbf{y}_k, \mathbf{a}_k) \nu^{K-1} (1 - \nu) \middle| \mathbf{y}_0, \pi \right\} \\
 &= \mathbb{E} \left\{ \sum_{k=0}^{\infty} \sum_{K=k+1}^{\infty} r(\mathbf{y}_k, \mathbf{a}_k) \nu^{K-1} (1 - \nu) \middle| \mathbf{y}_0, \pi \right\} \\
 &= \mathbb{E} \left\{ \sum_{k=0}^{\infty} \nu^k r(\mathbf{y}_k, \mathbf{a}_k) \middle| \mathbf{y}_0, \pi \right\}.
 \end{aligned}$$

□

Here, we can interpret ν as the *discount factor* of the model. Since the sensor node will stop functioning at some time in the future, the reward at time slot k is discounted by a factor ν^k .

Lemma 3.2. $J^\pi(\mathbf{y}_0)$ defined in (3.2) is finite. That is $|J^\pi(\mathbf{y}_0)| < \infty$.

Proof. Since

$$\sup_{\mathbf{a} \in U(\mathbf{y}), \mathbf{y} \in Y} |r(\mathbf{y}, \mathbf{a})| = \max_{\alpha \in \mathcal{A}} \left\{ \mathbb{E}_{\alpha'} [\min\{\mu(b_{\max}, \alpha'), q_{\max}\} \mid \alpha] \right\} < \infty, \quad (3.3)$$

the objective function $J^\pi(\mathbf{y}_0)$ of the infinite-horizon MDP converges to a finite value [28, pp. 121].

□

The sensor node aims to find an optimal sensing and transmit energy allocation policy that maximizes the expected total discounted reward given in (3.2). That is, given the initial state \mathbf{y}_0 , the sensor aims to obtain the *optimal expected total discounted reward* $J(\mathbf{y}_0)$ and the *optimal policy* π^* defined as

$$J(\mathbf{y}_0) = \max_{\pi \in \Pi} J^\pi(\mathbf{y}_0) \quad \text{and} \quad \pi^* = \arg \max_{\pi \in \Pi} J^\pi(\mathbf{y}_0). \quad (3.4)$$

A policy is said to be *stationary* if $\delta_k = \delta$ for all $k \in \mathcal{K}$ such that $\pi = (\delta, \delta, \dots)$. For the rest of this chapter, a general policy is denoted by π , while a stationary policy is denoted by δ . For an infinite-horizon MDP, the only case of interest is when a stationary optimal policy exists. Thus our objective is to find an *optimal stationary deterministic policy* δ^* , which maximizes the expected total discounted reward in (3.2).

3.2 Energy Allocation Algorithms

In this section, we obtain the *optimal stationary policies* for energy allocation. First, we consider a general case that takes into account a finite data buffer and the energy allocated for sensing. Next, we study a special case where we assume that there is an infinite data backlog.

3.2.1 General Case

In this subsection, we obtain the *optimal stationary policy* for the general case. An OSEA algorithm that achieves the maximum expected total discounted reward in (3.4) is proposed based on the *value iteration* algorithm [28, pp. 161].

The optimal expected total discounted reward $J(\mathbf{y})$ given current state \mathbf{y} satisfies the *Bellman's equation of optimality* [28]:

$$J(\mathbf{y}) = \max_{\mathbf{a} \in U(\mathbf{y})} \left\{ r(\mathbf{y}, \mathbf{a}) + \nu \sum_{\mathbf{y}' \in Y} P(\mathbf{y}' | \mathbf{y}, \mathbf{a}) J(\mathbf{y}') \right\}. \quad (3.5)$$

In equation (3.5), the first and second terms on the right hand side represent, respectively, the *immediate reward* at the current time slot and the *expected total discounted*

future reward if action \mathbf{a} is chosen. Hence, equation (3.5) describes the *tradeoff* between the current reward and the expected future reward.

Theorem 3.1. *There exists an optimal stationary deterministic policy δ^* that maximizes the right hand side of (3.5), given by*

$$\delta^*(\mathbf{y}) = \arg \max_{\mathbf{a} \in U(\mathbf{y})} \left\{ r(\mathbf{y}, \mathbf{a}) + \nu \sum_{\mathbf{y}' \in Y} P(\mathbf{y}' | \mathbf{y}, \mathbf{a}) J(\mathbf{y}') \right\}. \quad (3.6)$$

Proof. Notice that the system state space Y is countable and discrete, and $U(\mathbf{y})$ is finite for each $\mathbf{y} \in Y$. From [28, Theorem 6.2.10], an optimal stationary deterministic policy exists. \square

We then propose the OSEA algorithm in Algorithm 2. In the planning phase, the sensor solves for the optimal stationary policy δ^* based on value iteration algorithm, and records it as a look-up table. Specifically, in line 2, we initialize $J_0(\mathbf{y})$ for all $\mathbf{y} \in Y$ arbitrarily, specify the error bound ε , and set the iteration sequence n to be 0. In line 3, we compute $J_{n+1}(\mathbf{y})$ for each $\mathbf{y} \in Y$ based on the knowledge of $J_n(\mathbf{y})$. In line 4, we first check whether $\|\mathbf{J}_{n+1} - \mathbf{J}_n\| < \frac{\varepsilon(1-\nu)}{2\nu}$ holds where $\mathbf{J}_{n+1} = (J_{n+1}(\mathbf{y}), \forall \mathbf{y} \in Y)$ and $\mathbf{J}_n = (J_n(\mathbf{y}), \forall \mathbf{y} \in Y)$, and the norm function is defined to be $\|\mathbf{J}\| = \max |J(\mathbf{y})|$ for $\mathbf{y} \in Y$. If the inequality holds, which means that the value iteration algorithm has converged, then we proceed to obtain the optimal stationary policy δ^* in line 5 and stop. Otherwise, we go back to line 3 and continue to iterate. In the sensing and transmission phase, the sensor node first tracks the energy harvesting rate of the previous time slot h_{k-1} , the battery energy level b_k , the amount of data in the buffer q_k , and obtains the channel feedback α_{k-1} from the sink in line 9 to 12. Then, the sensor node chooses the action $\delta^*(\mathbf{y}) = (e^*(\mathbf{y}), s^*(\mathbf{y}))$ based on current system state \mathbf{y} and the optimal stationary policy δ^* in line 14. That is, it consumes $e^*(\mathbf{y})$ and

Algorithm 2 Optimal Stationary Energy Allocation (OSEA) Algorithm for Energy Harvesting Sensor Node.

- 1: Planning Phase:
- 2: Arbitrarily select $J_0(\mathbf{y})$ for each $\mathbf{y} \in Y$, specify $\varepsilon > 0$, and set $n := 0$.
- 3: For each $\mathbf{y} \in Y$, compute $J_{n+1}(\mathbf{y})$ by

$$J_{n+1}(\mathbf{y}) := \max_{\mathbf{a} \in U(\mathbf{y})} \left\{ r(\mathbf{y}, \mathbf{a}) + \nu \sum_{\mathbf{y}' \in Y} P(\mathbf{y}' | \mathbf{y}, \mathbf{a}) J_n(\mathbf{y}') \right\}. \quad (3.7)$$

- 4: If $\|\mathbf{J}_{n+1} - \mathbf{J}_n\| < \frac{\varepsilon(1-\nu)}{2\nu}$, go to line 5. Otherwise increment n by 1 and go to line 3.
- 5: For each $\mathbf{y} \in Y$, choose stationary ε -optimal policy

$$\delta^*(\mathbf{y}) := \arg \max_{\mathbf{a} \in U(\mathbf{y})} \left\{ r(\mathbf{y}, \mathbf{a}) + \nu \sum_{\mathbf{y}' \in Y} P(\mathbf{y}' | \mathbf{y}, \mathbf{a}) J_{n+1}(\mathbf{y}') \right\}, \quad (3.8)$$

and stop.

- 6: Sensing and Transmission Phase:
 - 7: Set $k := 0$.
 - 8: **while** $k \leq K - 1$ **do**
 - 9: Track the energy harvesting rate of the previous time slot h_{k-1} .
 - 10: Track the energy available for use in the battery b_k .
 - 11: Track the amount of data in the buffer q_k .
 - 12: Obtain the channel feedback α_{k-1} from the sink.
 - 13: Set $\mathbf{y} := (b_k, q_k, h_{k-1}, \alpha_{k-1})$.
 - 14: Obtain action $\delta^*(\mathbf{y}) := (e^*(\mathbf{y}), s^*(\mathbf{y}))$ based on the optimal policy.
 - 15: Consume $e^*(\mathbf{y})$ amount of energy for transmission and $s^*(\mathbf{y})$ amount of energy for sensing.
 - 16: Update battery energy b_{k+1} using (2.2) and the amount of data in the buffer q_{k+1} using (2.5).
 - 17: Set $k := k + 1$.
 - 18: **end while**
-

$s^*(\mathbf{y})$ amount of energy for transmission and sensing, respectively.

For the *convergence*, $J_n(\mathbf{y})$ generated in line 3 converges in norm to $J(\mathbf{y})$ for all $\mathbf{y} \in Y$. The stationary policy δ^* defined in line 5 is ε -optimal, and whenever the convergence criterion $\|\mathbf{J}_{n+1} - \mathbf{J}_n\| < \frac{\varepsilon(1-\nu)}{2\nu}$ is satisfied, $\|\mathbf{J}_{n+1} - \mathbf{J}\| < \varepsilon/2$ holds, where $\mathbf{J} = (J(\mathbf{y}), \forall \mathbf{y} \in Y)$ is the vector of *optimal* expected total discounted reward defined in (3.5). Besides, the convergence is linear at rate ν . In practice, choosing ε small enough ensures to obtain a policy that is very close to optimal.

Lemma 3.3. (a) $J(b, q, h, \alpha)$ is increasing in battery state b for any given data buffer state q , energy harvesting state h , and channel state α . (b) $J(b, q, h, \alpha)$ is increasing in q for any given b , h and α .

Proof. We prove it by mathematical induction. In order to show that $J(b, q, h, \alpha)$ is increasing in b and q , we aim to prove that $J_n(b, q, h, \alpha)$ generated by equation (3.7) in Algorithm 2 is increasing in b and q for all n . Since for any initialization $J_0(b, q, h, \alpha)$, $J_n(b, q, h, \alpha)$ converges to the same optimal expected total discounted reward $J(b, q, h, \alpha)$ [28], we can select $J_0(b, q, h, \alpha)$ which is increasing in b and q . Assume $J_n(b, q, h, \alpha)$ is increasing in b and q . We expand (3.7) as

$$J_{n+1}(b, q, h, \alpha) = \max_{\alpha \in U(\mathbf{y})} \left\{ \mathbb{E}_{\alpha'} [\min\{\mu(e, \alpha'), q\} \mid \alpha] + \right. \\ \left. \nu \mathbb{E}_{h', \alpha'} \left[J_n \left(\min\{b - (e + s) + h', b_{\max}\}, \min\{[q - \mu(e, \alpha')]^+ + x(s), q_{\max}\}, h', \alpha' \right) \mid h, \alpha \right] \right\}. \quad (3.9)$$

Note that the first term on the right hand side of equation (3.9) is independent of b and increasing in q , and the second term is increasing in b and q based on the assumption that $J_n(b, q, h, \alpha)$ is increasing in b and q . Therefore, $J_{n+1}(b, q, h, \alpha)$ is increasing in b and q . By induction, $J_n(b, q, h, \alpha)$ is increasing in b and q for all n . Thus, $J(b, q, h, \alpha) = J_\infty(b, q, h, \alpha)$ is increasing in b for any given q , h and α .

Moreover, it is increasing in q for any given b , h and α . \square

This property is intuitive. If more energy is available in the battery (i.e., a larger b), we can allocate more energy for sensing and transmission so that the total reward J increases. Similarly, if more data are available in the data buffer for transmission (i.e., a larger q), we can then allocate less energy for sensing and more energy for transmission, which would result in a larger total reward J .

3.2.2 Special Case: Infinite Data Backlog

In this subsection, we consider a special case where the sensor has an infinite data backlog. As a result, we do not need to consider the sensing energy s and the data buffer state q . So the system state is left with three elements: the battery energy b for current time slot, the energy harvesting rate h , and the channel state α for the previous time slot. Based on the current system state, the sensor will choose e units of energy for transmission. We denote the expected optimal total discounted reward as $\hat{J}(b, h, \alpha)$, which satisfies the following Bellman's equation of optimality:

$$\hat{J}(b, h, \alpha) = \max_{0 \leq e \leq b} \{ \mathbb{E}_{\alpha'} [\mu(e, \alpha') | \alpha] + \nu \bar{J}(b - e, h, \alpha) \}, \quad (3.10)$$

where

$$\bar{J}(\hat{b}, h, \alpha) = \mathbb{E}_{h', \alpha'} [\hat{J}(\min\{\hat{b} + h', b_{\max}\}, h', \alpha') | h, \alpha]. \quad (3.11)$$

The first term on the right hand side of equation (3.10) represents the immediate reward for allocating e units of energy for transmission, and the second term represents the total future discounted reward. Equation (3.10) can be solved via the value iteration algorithm as in Section 3.2.1. However, we can prove some properties related to

$\hat{J}(b, h, \alpha)$ and $\bar{J}(\hat{b}, h, \alpha)$ in Lemmas 3.4 and 3.5, which leads to the *monotone policy* [28] in Theorem 3.2.

Lemma 3.4. $\hat{J}(b, h, \alpha)$ is increasing in b for any given h and α .

Proof. We prove it by mathematical induction. The optimal discounted reward $\hat{J}(b, h, \alpha)$ is obtained by the value iteration algorithm, given by

$$\hat{J}_{n+1}(b, h, \alpha) = \max_{0 \leq e \leq b} \left\{ \mathbb{E}_{\alpha'}[\mu(e, \alpha') \mid \alpha] + \nu \bar{J}_n(b - e, h, \alpha) \right\}, \quad (3.12)$$

where

$$\bar{J}_n(\hat{b}, h, \alpha) = \mathbb{E}_{h', \alpha'}[\hat{J}_n(\min\{\hat{b} + h', b_{\max}\}, h', \alpha') \mid h, \alpha]. \quad (3.13)$$

In order to show that $\hat{J}(b, h, \alpha)$ is increasing in b , we aim to prove that $\hat{J}_n(b, h, \alpha)$ generated by equation (3.12) is increasing in b for all n . Since for any initialization $\hat{J}_0(b, h, \alpha)$, $\hat{J}_n(b, h, \alpha)$ converges to the same optimal expected total discounted reward $\hat{J}(b, h, \alpha)$, we can select $\hat{J}_0(b, h, \alpha)$ which is increasing in b . Assume $\hat{J}_n(b, h, \alpha)$ is increasing in b , which indicates that $\bar{J}_n(\hat{b}, h, \alpha)$ is increasing in \hat{b} . Let $b' > b$, and we have

$$\hat{J}_{n+1}(b', h, \alpha) = \max_{0 \leq e \leq b'} \left\{ \mathbb{E}_{\alpha'}[\mu(e, \alpha') \mid \alpha] + \nu \bar{J}_n(b' - e, h, \alpha) \right\}, \quad (3.14a)$$

$$\hat{J}_{n+1}(b, h, \alpha) = \max_{0 \leq e \leq b} \left\{ \mathbb{E}_{\alpha'}[\mu(e, \alpha') \mid \alpha] + \nu \bar{J}_n(b - e, h, \alpha) \right\}. \quad (3.14b)$$

For any chosen action e , the first term on the right hand side of (3.14a) is the same as that of (3.14b) and the second term on the right hand side of (3.14a) is greater than or equal to that of (3.14b). Besides, the action set at the right hand side of (3.14a), $\{e \mid 0 \leq e \leq b'\}$, is larger than the action set at the right hand side of (3.14b),

$\{e \mid 0 \leq e \leq b\}$. Thus, we have

$$\hat{J}_{n+1}(b', h, \alpha) \geq \hat{J}_{n+1}(b, h, \alpha). \quad (3.15)$$

Thus, by induction, $\hat{J}(b, h, \alpha) = \hat{J}_\infty(b, h, \alpha)$ is increasing in b for any given h and α . \square

Lemma 3.5. (a) $\hat{J}(b, h, \alpha)$ is concave in b for any given h and α . (b) $\bar{J}(\hat{b}, h, \alpha)$ is concave in \hat{b} for any given h and α .

Proof. We prove it by mathematical induction. Since for any initialization $\hat{J}_0(b, h, \alpha)$, the sequence $\hat{J}_n(b, h, \alpha)$ generated by equation (3.12) converges to the optimal discounted reward $\hat{J}(b, h, \alpha)$, we can choose such $\hat{J}_0(b, h, \alpha)$ that is concave in b for any given h and α . Assume $\hat{J}_n(b, h, \alpha)$ is concave in b for any given h and α . We denote the optimal action that achieves $\hat{J}_{n+1}(b_1, h, \alpha)$ by e_1 , and the optimal action that achieves $\hat{J}_{n+1}(b_2, h, \alpha)$ by e_2 . Then, we have

$$\hat{J}_{n+1}(b_1, h, \alpha) = \mathbb{E}_{\alpha'}[\mu(e_1, \alpha') \mid \alpha] + \nu \bar{J}_n(b_1 - e_1, h, \alpha), \quad (3.16)$$

$$\hat{J}_{n+1}(b_2, h, \alpha) = \mathbb{E}_{\alpha'}[\mu(e_2, \alpha') \mid \alpha] + \nu \bar{J}_n(b_2 - e_2, h, \alpha). \quad (3.17)$$

Since $\mu(e, \alpha')$ is concave in e for any given α' , $\mathbb{E}_{\alpha'}[\mu(e, \alpha') \mid \alpha]$ is also concave in e because it is a weighted sum of concave functions. We then prove that $\bar{J}_n(\hat{b}, h, \alpha)$ is concave in \hat{b} . From the procedure of the proof of Lemma 3.4, we have $\hat{J}_n(b', h', \alpha')$ is increasing in b' for given h' and α' for all n . We already assume at the beginning of the proof that $\hat{J}_n(b', h', \alpha')$ is concave in b' for given h' and α' . And $b' = \min\{\hat{b} + h', b_{\max}\}$ is a concave function in \hat{b} [36]. Thus, by applying the results of composition [36, (3.10)], we can conclude that $\hat{J}_n(\min\{\hat{b} + h', b_{\max}\}, h', \alpha')$ is concave in \hat{b} for given h' and α' , which indicates that $\bar{J}_n(\hat{b}, h, \alpha)$ is concave in \hat{b} , since it is a weighted sum of

concave functions.

Now combining equation (3.16) and equation (3.17), and using the concavity of $\mathbb{E}_{\alpha'}[\mu(e, \alpha') | \alpha]$ and $\bar{J}_n(\hat{b}, h, \alpha)$, we have

$$\lambda \hat{J}_{n+1}(b_1, h, \alpha) + (1 - \lambda) \hat{J}_{n+1}(b_2, h, \alpha) \leq \mathbb{E}_{\alpha'}[\mu(e_\lambda, \alpha') | \alpha] + \nu \bar{J}_n(b_\lambda - e_\lambda, h, \alpha), \quad (3.18)$$

where $e_\lambda = \lambda e_1 + (1 - \lambda)e_2$, $b_\lambda = \lambda b_1 + (1 - \lambda)b_2$. Since $0 \leq e_1 \leq b_1$ and $0 \leq e_2 \leq b_2$, we have $0 \leq e_\lambda \leq b_\lambda$. By applying the definition of maximum and $\hat{J}_{n+1}(b, h, \alpha)$ in equation (3.12), we have

$$\begin{aligned} \mathbb{E}_{\alpha'}[\mu(e_\lambda, \alpha') | \alpha] + \nu \bar{J}_n(b_\lambda - e_\lambda, h, \alpha) &\leq \max_{0 \leq e \leq b_\lambda} \{ \mathbb{E}_{\alpha'}[\mu(e, \alpha') | \alpha] + \nu \bar{J}_n(b_\lambda - e, h, \alpha) \} \\ &= \hat{J}_{n+1}(b_\lambda, h, \alpha). \end{aligned} \quad (3.19)$$

Combining inequalities (3.18) and (3.19), we have

$$\lambda \hat{J}_{n+1}(b_1, h, \alpha) + (1 - \lambda) \hat{J}_{n+1}(b_2, h, \alpha) \leq \hat{J}_{n+1}(\lambda b_1 + (1 - \lambda)b_2, h, \alpha). \quad (3.20)$$

Inequality (3.20) shows that $\hat{J}_{n+1}(b, h, \alpha)$ is concave in b for given h and α . By induction, we can conclude that $\hat{J}_n(b, h, \alpha)$ is concave in b for given h and α for all n . Also, $\bar{J}_n(\hat{b}, h, \alpha)$ is concave in \hat{b} for all n . Hence, $\hat{J}(b, h, \alpha) = \hat{J}_\infty(b, h, \alpha)$ is concave in b for given h and α , and $\bar{J}(\hat{b}, h, \alpha) = \bar{J}_\infty(\hat{b}, h, \alpha)$ is concave in \hat{b} for given h and α .

□

Since $\mu(e, \alpha')$ is concave in e , $\mathbb{E}_{\alpha'}[\mu(e, \alpha') | \alpha]$ is also concave in e . By applying Lemma 3.5(b), $\nu \bar{J}(b - e, h, \alpha)$ is concave in $(b - e)$. Thus, the concavities of the two terms in (3.10) translate into a diminishing marginal reward for consuming energy at the current time slot, and saving energy for the future time slots, respectively.

Balancing these two terms properly results in an optimal policy.

Theorem 3.2. *For the optimal stationary policy*

$$\hat{e}^*(b, h, \alpha) = \min \left\{ e' \in \arg \max_{0 \leq e \leq b} \left\{ \mathbb{E}_{\alpha'}[\mu(e, \alpha') \mid \alpha] + \nu \bar{J}(b - e, h, \alpha) \right\} \right\}, \quad (3.21)$$

it is monotone increasing in b for any given h and α . That is, for any $b' \geq b$, we have

$$\hat{e}^*(b', h, \alpha) \geq \hat{e}^*(b, h, \alpha), \quad \forall h \in \mathcal{H}, \forall \alpha \in \mathcal{A}. \quad (3.22)$$

Proof. We prove Theorem 3.2 by applying [37, Theorem 2]. We aim to prove that $\hat{e}_{n+1}(b, h, \alpha)$, which is defined as

$$\hat{e}_{n+1}(b, h, \alpha) = \min \left\{ e' \in \arg \max_{0 \leq e \leq b} \left\{ \mathbb{E}_{\alpha'}[\mu(e, \alpha') \mid \alpha] + \nu \bar{J}_n(b - e, h, \alpha) \right\} \right\}, \quad (3.23)$$

is increasing in b for given h and α for all n . We drop the arguments of h and α from all functions. We denote $f(e) = \mathbb{E}_{\alpha'}[\mu(e, \alpha') \mid \alpha]$, and $g_n(\hat{b}) = \nu \bar{J}_n(\hat{b}, h, \alpha)$. Then, equation (3.12) can be written as

$$\hat{J}_{n+1}(b) = \max_{0 \leq e \leq b} \{f(e) + g_n(b - e)\}. \quad (3.24)$$

Define $e_l(b)$ and $e_u(b)$ to be the lower bound and upper bound of the set of feasible actions e , respectively, when the available energy in the battery is b . In equation (3.24), we have $e_l(b) = 0$ and $e_u(b) = b$, which are both increasing in b . To apply [37, Theorem 2], it is sufficient to show that $f(e) + g_n(b - e)$ has increasing difference in

(b, e) , that is, for any $b' \geq b, e' \geq e$,

$$(f(e') + g_n(b' - e')) - (f(e') + g_n(b - e')) \geq (f(e) + g_n(b' - e)) - (f(e) + g_n(b - e)). \quad (3.25)$$

Inequality (3.25) can be simplified to

$$g_n(b' - e') - g_n(b - e') \geq g_n(b' - e) - g_n(b - e), \quad \forall b' \geq b, e' \geq e. \quad (3.26)$$

From the proof of Lemma 3.5, $g_n(\hat{b}) = \nu \bar{J}_n(\hat{b}, h, \alpha)$ is concave in \hat{b} for all n . By applying the property of concave functions, we have

$$g_n(w + \Delta) - g_n(w) \geq g_n(v + \Delta) - g_n(v), \quad \forall w \leq v, \Delta \geq 0. \quad (3.27)$$

Substituting $w = b - e', v = b - e, \Delta = b' - b$, we obtain (3.26). Now, by applying the conclusion of [37, Theorem 2], we prove that $\hat{e}_{n+1}(b, h, \alpha)$ is increasing in b for any given h and α for all n . Thus, $\hat{e}^*(b, h, \alpha) = \hat{e}_\infty(b, h, \alpha)$ is increasing in b for given h and α . \square

With this monotone structure, we can significantly reduce the computational complexity of the value iteration algorithm, and propose our OTEA algorithm in Algorithm 3. The planning phase of the OTEA algorithm and OSEA algorithm (i.e., Algorithm 2) are similar. The main difference is the procedure in computing $\hat{J}_{n+1}(b, h, \alpha)$ in (3.28) in Algorithm 3, which has a lower complexity than that of the OSEA algorithm. Specifically, in line 6 of Algorithm 3, for any given $h \in \mathcal{H}$ and $\alpha \in \mathcal{A}$, we have $\hat{e}_{n+1}(b + \Delta_b, h, \alpha) \geq \hat{e}_{n+1}(b, h, \alpha)$ from the proof of Theorem 3.2, where Δ_b is the quantization resolution of battery energy. When we compute

Algorithm 3 Optimal Transmission Energy Allocation (OTEA) Algorithm for Energy Harvesting Sensor Node.

- 1: Planning Phase:
- 2: Arbitrarily select $\hat{J}_0(b, h, \alpha)$ for each $b \in \mathcal{B}$, $h \in \mathcal{H}$, $\alpha \in \mathcal{A}$, specify $\varepsilon > 0$, $\Delta_b > 0$, and set $n := 0$.
- 3: **for** each $h \in \mathcal{H}$, $\alpha \in \mathcal{A}$ **do**
- 4: Set $b := 0$ and $l := 0$.
- 5: **while** $b \leq b_{\max}$ **do**
- 6: Compute

$$\hat{J}_{n+1}(b, h, \alpha) := \max_{l \leq e \leq b} \left\{ \mathbb{E}_{\alpha'}[\mu(e, \alpha') | \alpha] + \nu \mathbb{E}_{h', \alpha'}[\hat{J}_n(\min\{b - e + h', b_{\max}\}, h', \alpha') | h, \alpha] \right\}, \quad (3.28)$$

$$\hat{e}_{n+1}(b, h, \alpha) := \min \left\{ e' \in \arg \max_{l \leq e \leq b} \left\{ \mathbb{E}_{\alpha'}[\mu(e, \alpha') | \alpha] + \nu \mathbb{E}_{h', \alpha'}[\hat{J}_n(\min\{b - e + h', b_{\max}\}, h', \alpha') | h, \alpha] \right\} \right\}.$$
- 7: Set $b := b + \Delta_b$, $l := \hat{e}_{n+1}(b, h, \alpha)$.
- 8: **end while**
- 9: **end for**
- 10: If $\|\hat{J}_{n+1} - \hat{J}_n\| < \frac{\varepsilon(1-\nu)}{2\nu}$, go to line 11. Otherwise increment n by 1 and go to line 3.
- 11: For each $b \in \mathcal{B}$, $h \in \mathcal{H}$, $\alpha \in \mathcal{A}$, choose stationary ε -optimal policy

$$\hat{\delta}^*(b, h, \alpha) := \arg \max_{0 \leq e \leq b} \left\{ \mathbb{E}_{\alpha'}[\mu(e, \alpha') | \alpha] + \nu \mathbb{E}_{h', \alpha'}[\hat{J}_{n+1}(\min\{b - e + h', b_{\max}\}, h', \alpha') | h, \alpha] \right\}, \quad (3.29)$$

and stop.

- 12: Sensing and Transmission Phase:
 - 13: Set $k := 0$.
 - 14: **while** $k \leq K - 1$ **do**
 - 15: Track the energy harvesting rate of the previous time slot h_{k-1} .
 - 16: Track the energy available for use in the battery b_k .
 - 17: Track the amount of data in the buffer q_k .
 - 18: Obtain the channel feedback α_{k-1} from the sink.
 - 19: Choose the amount of energy for sensing to be $\hat{s}^* := pb_k$, where p is the fixed percentage of energy for sensing.
 - 20: Choose the amount of energy for transmission to be $\hat{e}^* := \hat{\delta}^*((1-p)b_k, h_{k-1}, \alpha_{k-1})$ based on the optimal policy.
 - 21: Consume \hat{e}^* amount of energy for transmission and \hat{s}^* amount of energy for sensing.
 - 22: Update battery energy b_{k+1} using (2.2) and the amount of data in the buffer q_{k+1} using (2.5).
 - 23: Set $k := k + 1$.
 - 24: **end while**
-

$\hat{J}_{n+1}(b + \Delta_b, h, \alpha)$ and search for $\hat{e}_{n+1}(b + \Delta_b, h, \alpha)$, we can find the optimal solution in the interval of $[\hat{e}_{n+1}(b, h, \alpha), b + \Delta_b]$ instead of the longer interval $[0, b + \Delta_b]$. In the sensing and transmission phase, when we apply our OTEA algorithm to a practical system, we still need to take into account the energy for sensing. In this way, we fix the percentage of energy allocated for sensing to be p in line 19. The other operations in the sensing and transmission phase are the same as that in the OSEA algorithm.

3.3 Performance Evaluation

In this section, we simulate the performance of the OSEA, OTEA, and CAEA algorithms in terms of the total amount of data transmitted in Matlab. We consider the same wireless channel as in Chapter 2, which evolves according to the three-state Markov chain as shown in Figure 2.3 with the transition matrix of the Markov chain given by (2.14).

Unless specified otherwise, we assume that the battery buffer size $b_{\max} = 30$ J, and the data buffer size $q_{\max} = 0.5$ Mbits. The initial amount of energy in the battery is 10 J and the initial amount of data in the buffer is 0.1 Mbits. For tractability, we assume that the energy harvesting rate h_k takes values from the finite set $\mathcal{H} = \{H_1, H_2, H_3\} = \{4, 8, 12\}$ J/time slot, and evolves according to the three-state *Markov chain* with the state transition probability given by

$$P_h = \begin{bmatrix} P_{H_1H_1} & P_{H_1H_2} & P_{H_1H_3} \\ P_{H_2H_1} & P_{H_2H_2} & P_{H_2H_3} \\ P_{H_3H_1} & P_{H_3H_2} & P_{H_3H_3} \end{bmatrix} = \begin{bmatrix} 0.5 & 0.5 & 0 \\ 0.25 & 0.5 & 0.25 \\ 0 & 0.5 & 0.5 \end{bmatrix}. \quad (3.30)$$

The steady state probability is then given by $[P_{H_1} \ P_{H_2} \ P_{H_3}] = [0.25 \ 0.5 \ 0.25]$. $x(s_k)$ is assumed to be a linear function of s_k given by $x(s_k) = \gamma s_k$. We adopt $\gamma = 0.08$

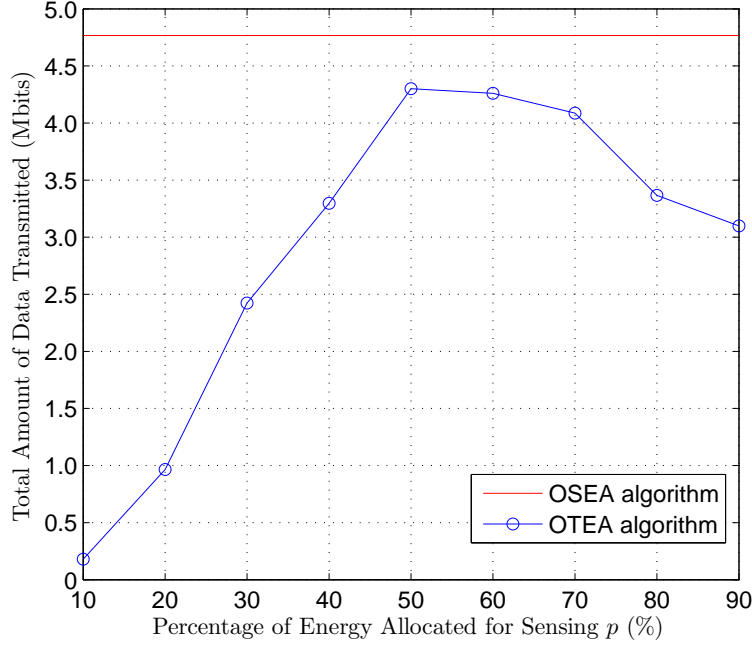


Figure 3.1: The total amount of data transmitted of OTEA algorithm under different percentage of energy allocated for sensing p . Since the OSEA algorithm does not allocate a fixed amount of energy for sensing, its total amount of data transmitted is independent of p .

Mbits/J. For the value iteration algorithm, we choose ε to be 10^{-3} and the discount factor ν to be 0.95. Since the CAEA algorithm considers the energy allocation over a finite horizon, where the lifetime of the sensor node is known, we fix the sensor lifetime in the CAEA algorithm to be equal to the mean of the sensor lifetime in the OSEA and OTEA algorithms.

Since the performance of the OTEA algorithm is related to the fixed amount of energy allocated for sensing, we examine the total amount of data transmitted by the OTEA algorithm under different percentage of energy allocated for sensing, and compare with the OSEA algorithm. As shown in Figure 3.1, with around 50% of the available battery energy allocated for sensing, the OTEA algorithm transmits the largest amount of data, and it is close to that of the OSEA algorithm. This

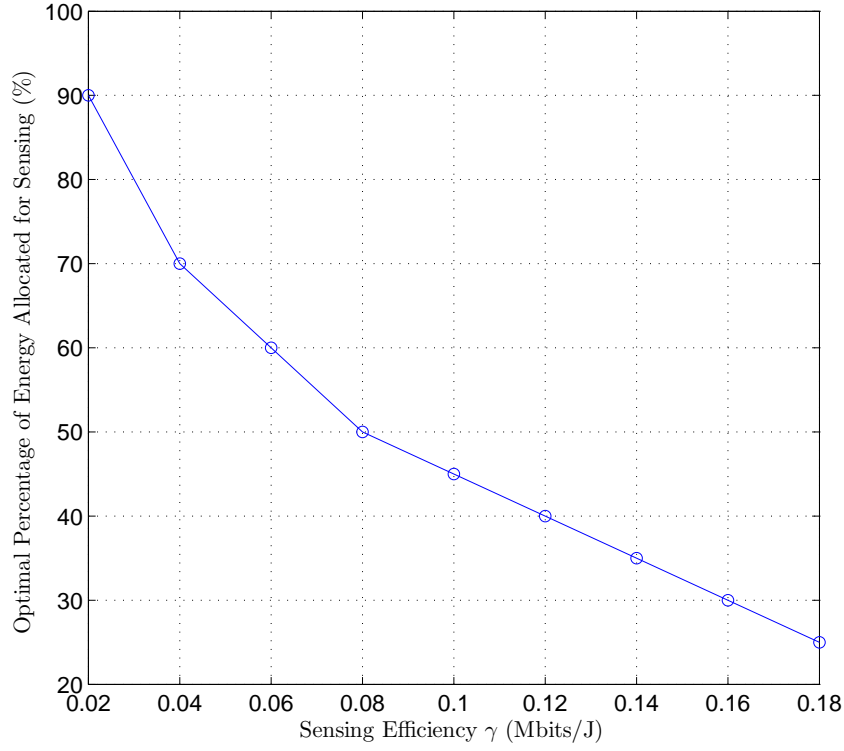


Figure 3.2: The optimal percentage of energy allocated for sensing under different data-sensing efficiency γ for the OTEA algorithm.

implies that we can apply the OTEA algorithm, which has a lower complexity than the OSEA algorithm, and choose the optimal fixed percentage of energy for sensing to achieve a near-optimal performance.

Moreover, the optimal percentage of energy allocated for sensing for the OTEA algorithm depends on the data-sensing efficiency γ . With a larger γ , the sensor can sense more data using the same amount of energy. Figure 3.2 shows that as γ increases, the optimal percentage of energy for sensing decreases.

We then examine the total amount of data transmitted by the OSEA algorithm, the OTEA algorithm, and the CAEA algorithm under different average energy harvesting rates \bar{H} , where $\bar{H} = \sum_{i=1}^3 H_i P_{H_i}$. For the OTEA and CAEA algorithms,

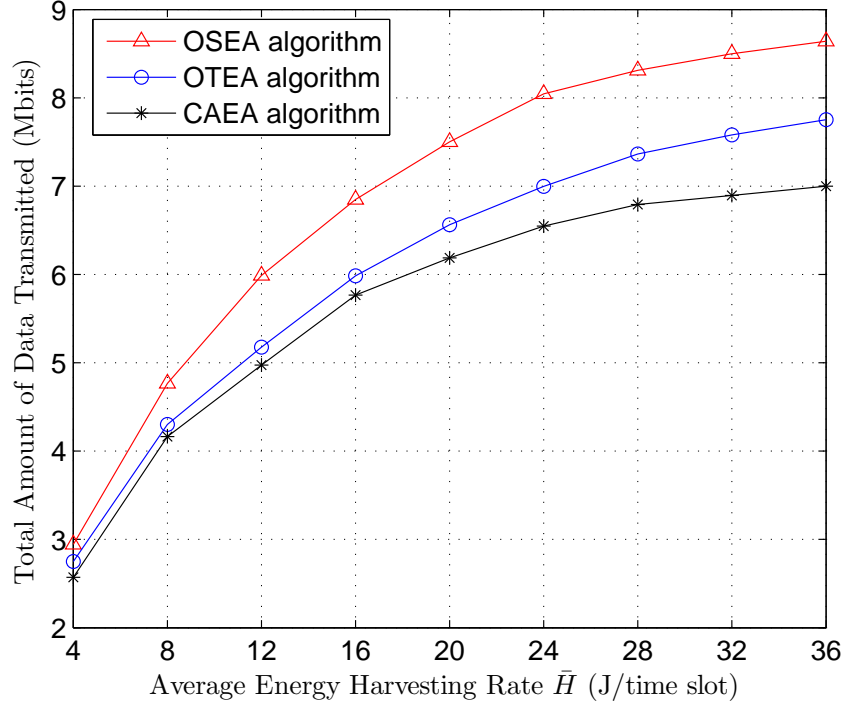


Figure 3.3: The total amount of data transmitted of the three algorithms for different average energy harvesting rates \bar{H} .

the percentage of energy allocated for sensing is fixed to be 50%. In Figure 3.3, we plot the total amount of data transmitted against different average energy harvesting rate for these three algorithms. We observe that the OSEA algorithm performs better than the OTEA algorithm and the CAEA algorithm, since the OSEA algorithm achieves the optimal performance by solving the problem (3.4). Moreover, the OTEA algorithm has a better performance than the CAEA algorithm. It is because the OTEA algorithm takes into account the randomness of the lifetime of the sensor node, while the CAEA algorithm just considers the lifetime of the sensor node to be its mean value. Besides, the total amount of data transmitted by these three algorithms saturates as the average harvesting rate is increased beyond a certain level. It is because when the energy harvesting rate is larger than the battery capacity, part

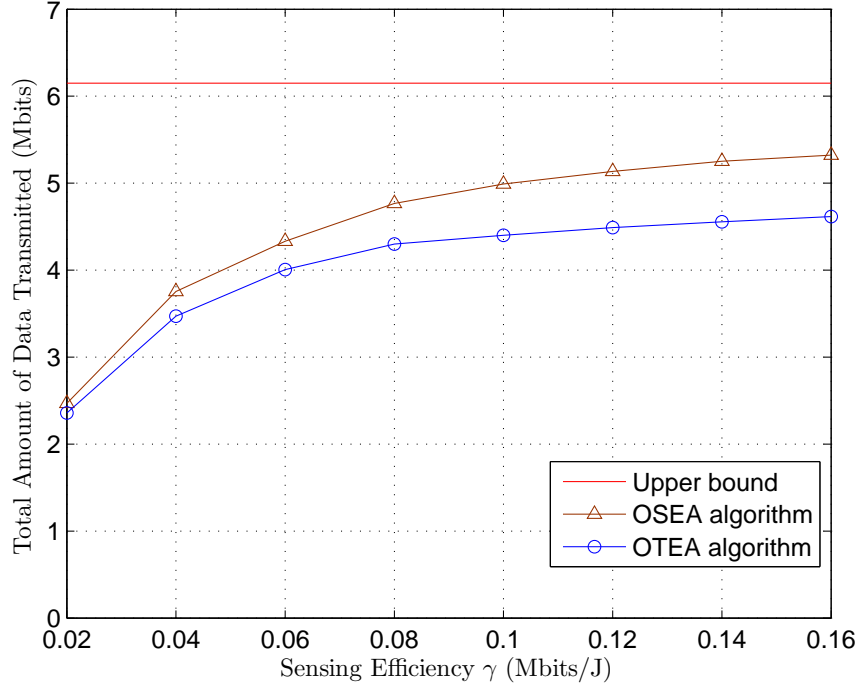


Figure 3.4: The total amount of data transmitted of the OSEA algorithm and the OTEA algorithm for different values of data-sensing efficiency parameter γ .

of the harvested energy cannot be accommodated, and is lost due to the overflow of the battery energy.

Next, we examine the total amount of data transmitted of the OSEA and OTEA algorithms under different data-sensing efficiency γ . As shown in Figure 3.4, when γ is increased, the total amount of data transmitted increases as well, because more energy is left for data transmission. However, the performance saturates as γ is increased beyond a certain value. To the extreme when γ approaches infinity, it corresponds to the case where the sensing is extremely efficient. The total amount of data transmitted in this case provides an upper bound for the performance of the OSEA algorithm for the sensor node with different sensing efficiency.

Figure 3.5 shows the impact of the battery storage capacity b_{\max} on the total

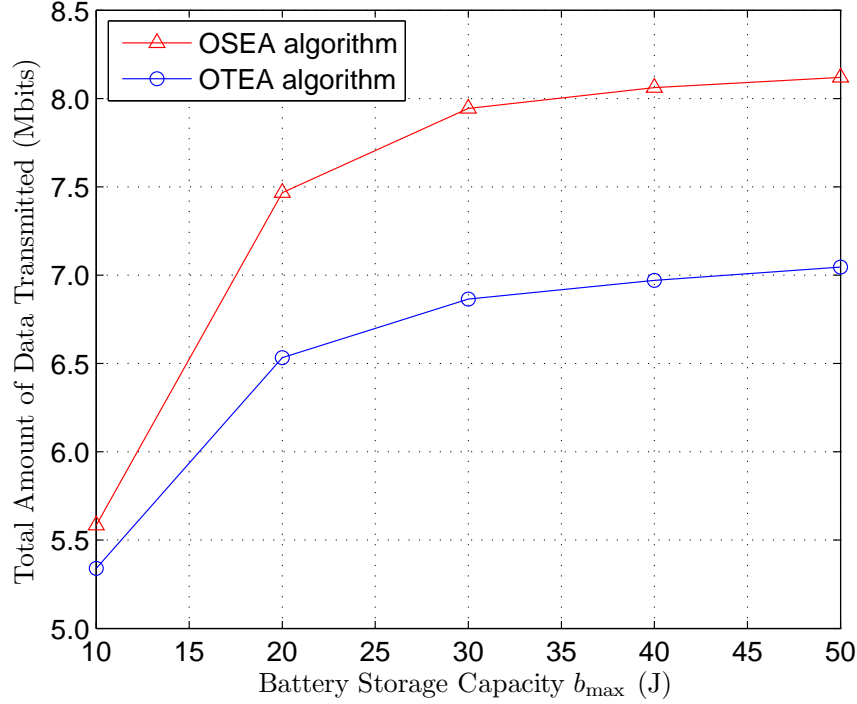


Figure 3.5: The total amount of data transmitted of the OSEA algorithm and the OTEA algorithm for different battery storage capacity b_{\max} .

amount of data transmitted. We consider that the value of h is taken from the set $\mathcal{H} = \{20, 24, 28\}$ J/time slot. As shown in Figure 3.5, the total amount of data transmitted increases as the battery storage capacity b_{\max} increases. It is because with a larger battery storage capacity b_{\max} , the sensor node can manage the harvested energy better since the sensor can save more energy for future use if necessary. In other words, the sensor has more freedom to manage the incoming energy when b_{\max} is larger. The total amount of data transmitted saturates when b_{\max} goes to large values because under current energy harvesting rates, the battery energy level never exceeds some value, thus for all the battery capacities b_{\max} that are larger than that certain value, the sensor has the same performance.

In Figure 3.6, we study the impact of the data buffer size q_{\max} on the total amount

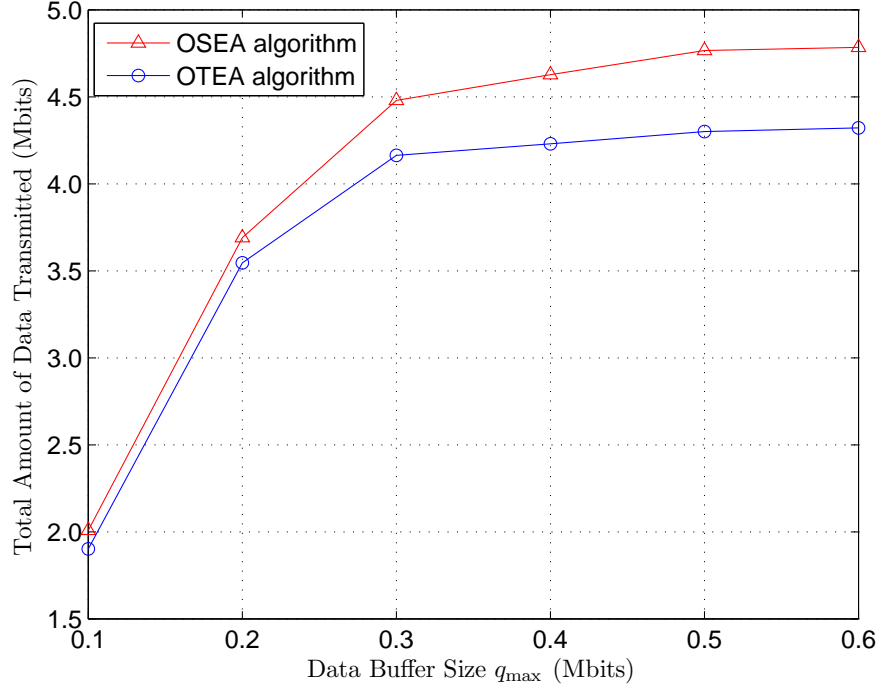


Figure 3.6: The total amount of data transmitted of the OSEA algorithm and the OTEA algorithm for different data buffer size q_{\max} .

of data transmitted by the OSEA and OTEA algorithms. We can observe that the total amount of transmitted data increases when q_{\max} increases. The performance saturates when q_{\max} is increased to a certain large value. This means that the amount of data in the buffer never exceeds a certain level under the optimal energy allocation policy. Otherwise, we should have observed that the total amount of transmitted data would continue to increase with the data buffer size q_{\max} .

Finally, we study the total amount of transmitted data of the OSEA algorithm and the OTEA algorithm under different discount factors ν . Since $1/(1-\nu)$ represents the average lifetime of the sensor node, a larger ν corresponds to a longer lifetime, which leads to a larger total amount of data transmitted. In Figure 3.7, the total amount of data transmitted increases as the discount factor ν increases. When ν approaches

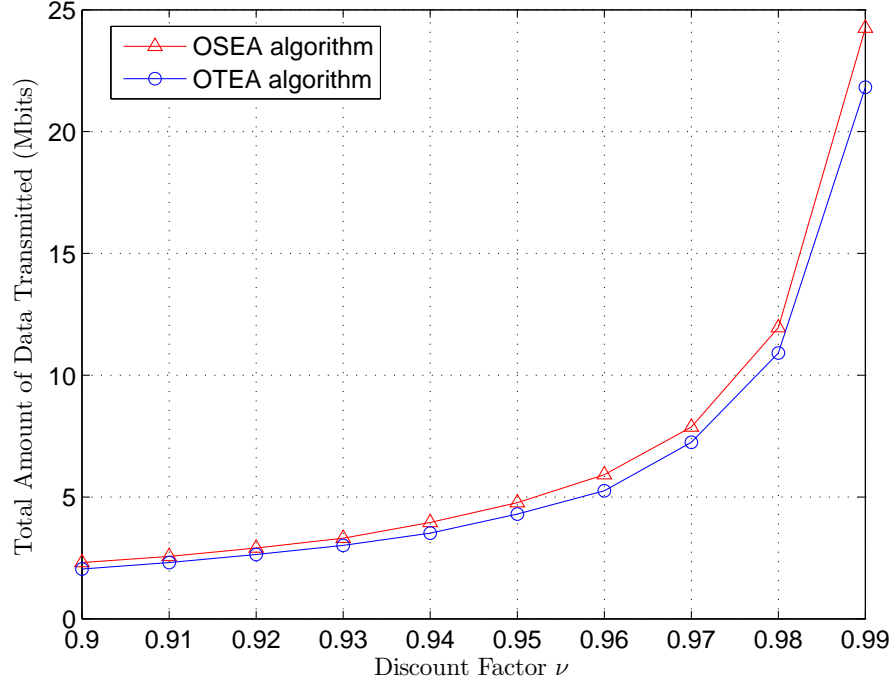


Figure 3.7: The total amount of data transmitted of the OSEA algorithm and the OTEA algorithm for different values of discount factor ν .

1, where the lifetime of the sensor node approaches infinity, the total amount of transmitted data goes to infinity. Besides, the number of iterations required for the value iteration algorithm to converge depends on ν . With a larger ν , a larger number of iterations are required.

Chapter 4

Conclusions and Future Work

In this chapter, we conclude the thesis by summarizing our contributions. We also suggest topic for further research.

4.1 Conclusions

In this thesis, we studied the problem of maximizing the expected total amount of data transmitted for an energy harvesting sensor node under energy harvesting rate variations and channel fluctuations in a time-slotted system. A finite data buffer and the energy consumed for sensing data were considered for the first time. In this case, the sensor should achieve a good tradeoff between the energy consumed for sensing and transmission so as to achieve a large amount of total transmitted data. We considered two cases of the sensor lifetime, a fixed value and a random variable, in Chapters 2 and 3, respectively.

- In Chapter 2, we studied the energy allocation problem for sensing and transmission in a energy harvesting WSN over a finite horizon. The sensor lifetime is a fixed value. We formulated it as a finite-horizon MDP under channel fluctuations and energy variations in a time-slotted system. We obtained the optimal energy allocation policy and propose the OEA algorithm by using backward induction. We also provided extensive simulation results to compare the performance of the OEA algorithm and the CAEA algorithm. The results showed

that the OEA algorithm could transmit much more amount of data over a finite horizon than the CAEA algorithm under different settings.

- In Chapter 3, we extended the joint energy allocation problem by taking into account the randomness of the sensor lifetime. Since the lifetime of the sensor node is a random variable with geometric distribution, we formulated the problem as an infinite-horizon MDP. We obtained the optimal stationary energy allocation policy and proposed an OSEA algorithm based on value iteration in MDP. We also studied the transmission energy allocation problem under the assumption that there was infinite data backlog. We obtained structural results for the OTEA policy and proved that the OTEA policy was a monotonically increasing function of the available battery energy. Finally, we provided extensive simulation results to compare the performances of the OSEA algorithm, the OTEA algorithm, and the CAEA algorithm, and studied the impact of the average energy harvesting rate, the data-sensing efficiency, the battery capacity, the data buffer size and the lifetime of the sensor node on the total amount of data transmitted. The results showed that the OSEA algorithm transmitted the largest amount of data among the three algorithms, and the OTEA algorithm could achieve a near-optimal performance when the fixed percentage of energy for sensing was chosen properly.

4.2 Future Work

In terms of future work, we can consider several potential extensions of current work.

Maximizing the average throughput for an energy harvesting wireless sensor node over an infinite horizon. We consider maximizing the average

throughput of a sensor node over an infinite horizon instead of maximizing the total amount of transmitted data. In this case, we can formulate the problem as an infinite-horizon MDP with average reward, and solve it by using value iteration, policy iteration or linear programming.

Joint optimal energy allocation, scheduling, and routing policy in a multi-hop energy harvesting wireless sensor network. We consider extending the single hop scenario to a multi-hop setting for data transmission. In the multi-hop case, the sensor nodes sense the field and transmit data to a fusion node. In each time slot, the sensor node can harvest energy from the environment and store it in a rechargeable battery for future use. The sensor is in either sleep mode or active mode in any time slot. In the sleep mode, the sensor nodes can only harvest energy, but can not sense or transmit data. In the active mode, the sensor nodes can harvest energy, sense, process, and transmit data. Each node can transmit data to any other nodes in the network. For such a model, we can aim to develop a joint optimal energy allocation, scheduling and routing policy that ensures a fair utilization of the network resources. We can also take into consideration the fact that the energy harvesting process is unpredictable and stochastic in nature, and develop adaptive routing and scheduling algorithms that are able to dynamically learn and adapt to time variations in the energy harvesting process and network environments.

Bibliography

- [1] C. K. Ho and R. Zhang, “Optimal energy allocation for wireless communications powered by energy harvesters,” in *Proc. of IEEE Int’l. Symp. on Inform. Theory (ISIT)*, Austin, TX, Jun. 2010.
- [2] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, “Wireless sensor networks: A survey,” *Computer Networks*, vol. 38, no. 4, pp. 393–422, Mar. 2002.
- [3] G. Werner-Allen, K. Lorincz, M. Ruiz, O. Marcillo, J. Johnson, J. Lees, and M. Welsh, “Deploying a wireless sensor network on an active volcano,” *IEEE Internet Comput.*, vol. 10, no. 2, pp. 18–25, Mar. 2006.
- [4] A. Mainwaring, D. Culler, J. Polastre, R. Szewczyk, and J. Anderson, “Wireless sensor networks for habitat monitoring,” in *Proc. of the 1st ACM International Workshop on Wireless Sensor Networks and Applications*, Atlanta, GA, Sept. 2002.
- [5] M. Karpiriski, A. Senart, and V. Cahill, “Sensor networks for smart roads,” in *Proc. of IEEE International Conference on Pervasive Computing and Communications Workshops*, Pisa, Italy, Mar. 2006.
- [6] K. Chebrolu, B. Raman, N. Mishra, P. K. Valiveti, and R. Kumar, “Brimon: A sensor network system for railway bridge monitoring,” in *Proc. of ACM International Conference on Mobile Systems, Applications, and Services*, Breckenridge, CO, Jun. 2008.
- [7] L. van Hoesel, T. Nieberg, H. Kip, and P. Havinga, “Advantages of a TDMA based, energy-efficient, self-organizing MAC protocol for WSNs,” in *Proc. of IEEE Vehicular Technology Conference*, Los Angeles, CA, Sept. 2004.
- [8] Y. Kim, H. Shin, and H. Cha, “Y-MAC: An energy-efficient multi-channel MAC protocol for dense wireless sensor networks,” in *Proc. of the 7th International Conference on Information Processing in Sensor Networks*, St. Louis, MO, Apr. 2008.
- [9] Q. Ren and Q. Liang, “An energy-efficient MAC protocol for wireless sensor networks,” in *Proc. of IEEE Globecom*, St. Louis, MO, Nov. 2005.

- [10] D. Ganesan, R. Govindan, S. Shenker, and D. Estrin, “Highly-resilient, energy-efficient multipath routing in wireless sensor networks,” *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 5, no. 4, pp. 11–25, Oct. 2001.
- [11] S. D. Muruganathan, D. C. F. Ma, R. I. Bhasin, and A. O. Fapojuwo, “A centralized energy-efficient routing protocol for wireless sensor networks,” *IEEE Communications Magazine*, vol. 43, no. 3, pp. S8–S13, Mar. 2005.
- [12] F. Wang and J. Liu, “Duty-cycle-aware broadcast in wireless sensor networks,” in *Proc. of IEEE INFOCOM*, Rio de Janeiro, Brazil, Apr. 2009.
- [13] D. Niyato, E. Hossain, M. M. Rashid, and V. K. Bhargava, “Wireless sensor networks with energy harvesting technologies: A game-theoretic approach to optimal energy management,” *IEEE Wireless Communications*, vol. 14, no. 4, pp. 90–96, Aug. 2007.
- [14] S. Sudevalayam and P. Kulkarni, “Energy harvesting sensor nodes: Survey and implications,” *IEEE Communications Surveys and Tutorials*, vol. 13, no. 3, pp. 443–461, Third Quarter 2011.
- [15] S. Priya and D. J. Inman, *Energy Harvesting Technologies*. Springer, 2009.
- [16] V. Raghunathan, A. Kansal, J. Hsu, J. Friedman, and M. Srivastava, “Design considerations for solar energy harvesting wireless embedded systems,” in *Proc. of the 4th International Symposium on Information Processing in Sensor Networks*, Los Angeles, CA, Apr. 2005.
- [17] F. Simjee and P. H. Chou, “Everlast: Long-life, supercapacitor-operated wireless sensor node,” in *Proc. of the 2006 International Symposium on Low Power Electronics and Design*, Tegernsee, Germany, Oct. 2006.
- [18] X. Jiang, J. Polastre, and D. Culler, “Perpetual environmentally powered sensor networks,” in *Proc. of the 4th International Symposium on Information Processing in Sensor Networks*, Los Angeles, CA, Apr. 2005.
- [19] J. Taneja, J. Jeong, and D. Culler, “Design, modeling, and capacity planning for micro-solar power sensor networks,” in *Proc. of the 7th International Conference on Information Processing in Sensor Networks*, St. Louis, MO, Apr. 2008.
- [20] A. Kansal, J. Hsu, S. Zahedi, and M. B. Srivastava, “Power management in energy harvesting sensor networks,” *ACM Trans. on Embedded Computing Systems*, vol. 6, no. 4, Sept. 2007.
- [21] V. Sharma, U. Mukherji, V. Joseph, and S. Gupta, “Optimal energy management policies for energy harvesting sensor nodes,” *IEEE Trans. Wireless Communications*, vol. 9, no. 4, pp. 1326–1336, Apr. 2010.

- [22] M. Gatzianas, L. Georgiadis, and L. Tassiulas, "Control of wireless networks with rechargeable batteries," *IEEE Trans. on Wireless Communications*, vol. 9, no. 2, pp. 581–593, Feb. 2010.
- [23] L. Huang and M. J. Neely, "Utility optimal scheduling in energy harvesting networks," in *Proc. of ACM MobiHoc*, Paris, France, May 2011.
- [24] M. Gorlatova, A. Wallwater, and G. Zussman, "Networking low-power energy harvesting devices: Measurements and algorithms," in *Proc. of IEEE INFOCOM*, Shanghai, China, Apr. 2011.
- [25] J. Yang and S. Ulukus, "Optimal packet scheduling in an energy harvesting communication system," *IEEE Trans. on Communications*, vol. 60, no. 1, pp. 220–230, Jan. 2012.
- [26] M. A. Antepi, E. Uysal-Biyikoglu, and H. Erkal, "Optimal packet scheduling on an energy harvesting broadcast link," *IEEE J. Sel. Areas Communications*, vol. 29, no. 8, pp. 1721–1731, Sept. 2011.
- [27] D. P. Bertsekas, *Dynamic Programming and Optimal Control: Volume 1, 2nd ed.* Athena Scientific, 2000.
- [28] M. L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. New York, NY: John Wiley and Sons, 2005.
- [29] O. Ozel, K. Tutuncuoglu, J. Yang, S. Ulukus, and A. Yener, "Adaptive transmission policies for energy harvesting wireless nodes in fading channels," in *Proc. of IEEE Information Sciences and Systems (CISS)*, Baltimore, MD, Mar. 2011.
- [30] S. Chen, P. Sinha, N. B. Shroff, and C. Joo, "Finite-horizon energy allocation and routing scheme in rechargeable sensor networks," in *Proc. of IEEE INFOCOM*, Shanghai, China, Apr. 2011.
- [31] H. Li, N. Jaggi, and B. Sikdar, "Relay scheduling for cooperative communications in sensor networks with energy harvesting," *IEEE Trans. on Wireless Communications*, vol. 10, no. 9, pp. 2918–2928, Sept. 2011.
- [32] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.
- [33] C.-F. Chiasserini and R. R. Rao, "Improving battery performance by using traffic shaping techniques," *IEEE J. Sel. Areas Communications*, vol. 19, no. 7, pp. 1385–1394, Jul. 2001.
- [34] J. Lei, R. Yates, and L. Greenstein, "A generic model for optimizing single-hop transmission policy of replenishable sensors," *IEEE Trans. on Wireless Communications*, vol. 8, no. 2, pp. 547–551, Feb. 2009.

- [35] Q. Zhang and S. A. Kassam, “Finite-state Markov model for Rayleigh fading channels,” *IEEE Trans. on Communications*, vol. 47, no. 11, pp. 1688–1692, Nov. 1999.
- [36] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
- [37] R. Amir, “Supermodularity and complementarity in economics: An elementary survey,” *Southern Economic Journal*, vol. 71, no. 3, pp. 636–660, Jan. 2005.