DYNAMIC ALLOCATION OF RENEWABLE ENERGY THROUGH A STOCHASTIC KNAPSACK PROBLEM FORMULATION FOR AN ACCESS POINT ON THE MOVE

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

ELİF TUĞÇE CERAN

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
ELECTRICAL AND ELECTRONICS ENGINEERING

JUNE 2014

Approval of the thesis:

DYNAMIC ALLOCATION OF RENEWABLE ENERGY THROUGH A STOCHASTIC KNAPSACK PROBLEM FORMULATION FOR AN ACCESS POINT ON THE MOVE

submitted by ELİF TUĞÇE CERAN in partial fulfillment of the requirements for the degree of **Master of Science in Electrical and Electronics Engineering Department, Middle East Technical University** by,

Prof. Dr. Canan Özgen
Dean, Graduate School of **Natural and Applied Sciences** ______________

Prof. Dr. Gönül Turhan Sayan
Head of Department, **Electrical and Electronics Engineering** ______________

Prof. Dr. Elif Uysal Bıyıkoğlu
Supervisor, **Electrical and Electronics Eng. Dept., METU** ______________

Examining Committee Members:

Prof. Dr. Yalçın Tanık
Electrical and Electronics Engineering Department, METU ______________

Prof. Dr. Elif Uysal Bıyıkoğlu
Electrical and Electronics Engineering Department, METU ______________

Prof. Dr. Kemal Leblebicioğlu
Electrical and Electronics Engineering Department, METU ______________

Assoc. Prof. Dr. Melek Diker Yücel
Electrical and Electronics Engineering Department, METU ______________

Assoc. Prof. Dr. Tolga Girici
Electrical and Electronics Engineering, TOBB ETU ______________

Date: ______________

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name: ELİF TUĞÇE CERAN

Signature :

ABSTRACT

DYNAMIC ALLOCATION OF RENEWABLE ENERGY THROUGH A STOCHASTIC KNAPSACK PROBLEM FORMULATION FOR AN ACCESS POINT ON THE MOVE

ELİF TUĞÇE CERAN,
M.S., Department of Electrical and Electronics Engineering
Supervisor : Prof. Dr. Elif Uysal Bıyıkoğlu

June 2014, 67 pages

The problem studied in this thesis has been motivated by recent industry efforts toward providing Internet service in areas of the world devoid of regular telecommunications infrastructure via flying or floating platforms in the lower stratosphere. According to the abstraction in the thesis, the Access Point on the Move (APOM) having a renewable energy supply feature (solar, wind, etc.) must judiciously allocate this resource to provide service to users that demand service from it while it moves over an area. Within the problem setup, users with various stochastic characteristics (resource demands or rewards) appear in a sequential manner and the APOM must make online decisions whether or not to provide service to each appearing user. The objective of the APOM is to maximize a total utility (reward) provided to the encountered users. The problem is formulated as a 0/1 stochastic knapsack problem with stochastically growing dynamic capacity, solution of which is not available in previous literature. In this thesis, dynamic and stochastic policies are proposed considering the cases of both finite and infinite problem horizons. A threshold based policy based on dynamic programming approach is shown to be optimal under some conditions. Taking advantage of the structural characteristics of the optimal problem, promising suboptimal solutions that can adapt to short-time-scale dynamics are proposed and their performance are analysed in different scenarios. Kalman filtering based prediction of solar energy to inform online resource allocation policies is also considered.

Keywords: Access Point on the move, dynamic resource allocation, decision problem, threshold policy , energy harvesting, dynamic programming, Stochastic 0/1 knapsack problem, Knapsack with dynamic capacity, Markov Decision Process, Reinforcement Learning, Finite horizon, Infinite horizon, Discount factor

ÖZ

HAREKETLİ ERİŞİM NOKTALARI İÇİN OLASILIKSAL KNAPSACK PROBLEMİ FORMÜLASYONU İLE DİNAMİK YENİLENEBİLİR KAYNAK ATAMASI

ELİF TUĞÇE CERAN,
Yüksek Lisans, Elektrik ve Elektronik Mühendisliği Bölümü
Tez Yöneticisi : Prof. Dr. Elif Uysal Bıyıkoğlu

Haziran 2014 , 67 sayfa

Son yıllarda endüstriyel alanda telekommünikasyon altyapısı eksik bölgelere internet sağlamak amacıyla stratosferde uçan veya süzülen platformlar geliştirilmeye başlanmıştır. Bundan hareketle, bu tezde enerji hasatı kabiliyetli (güneş, rüzgar vb.) ve Hareketli Erişim Noktalarının (HEN) hareket sürecince karşılaşacağı kullanıcılara kaynak atama problemi incelenecektir. Problem kurulumunda, ard arda beliren farklı özelliklere sahip (fayda ve enerji talebi) kullanıcılar için HEN'in çevrimiçi bir şekilde servis verip vermeme kararı vermesi gereklidir. HEN'in nihai hedefi karşılaşılan kullanıcılardan elde edilecek fayda beklentisini maksimize etmek aynı zamanda enerji kapasitesini aşmamaktır. Problem 0/1 olasılıksal knapsack problemi olarak formüle edilebilir. Mevcut enerjinin olasılıksal enerji hasatı ile arttığı düşünüldüğünde, formülasyonda kullanılan knapsack problemi dinamik kapasiteye sahiptir. Bu tezde sonlu ve sonsuz ufuklu durumlar için dinamik ve olasılıksal yöntemler önerilmektedir. Dinamik programlama yaklaşımı kullanılarak belirli durumlarda eşik bazlı bir yöntemin ideal olduğu gösterilmektedir. İdeal çözümün yapısal özellikleri göz önünde bulundurularak kısa vadeli dinamiklere uyum sağlayabilen etkili en iyi altı çözümler önerilip farklı senaryolardaki performansları incelenmektedir. Ayrıca, kaynak atama yöntemleri ile birlikte kullanılmak üzere Kalman bazlı güneş enerjisi kestirim algoritması değerlendirilmektedir.

Anahtar Kelimeler: Hareketli erişim noktaları, dinamik kaynak atama, karar problemi, enerji hasatı, Olasılıksal 0/1 knapsack problemi, dinamik kapasiteli Knapsack, Markov Karar Süreci, takviyeli öğrenme, sonlu ufuk, sonsuz ufuk, kırdırım katsayısı

to my beloved family

ACKNOWLEDGMENTS

I would like to express my sincerest gratitude to my supervisor Prof. Dr. Elif Uysal Bıyıkoğlu for her excellent guidance, insight, patience and kindness motivated me to complete this thesis. All talks and discussions with her was a unique inspiration to me. Her contribution to this dissertation was truly invaluable. Moreover, I take this opportunity to thank Prof. Dr. Kemal Leblebicioğlu and Assoc. Prof. Dr. Tolga Girici who I had a chance to work with, for their invaluable support, constructive comments and attentiveness to weekly discussions through the development of this thesis.

I also would like to thank my friends and colleagues in the METU Communication Networks Research Group and Tuğçe Erkılıç for their generous help and precious friendship.

I would like to acknowledge the Scientific and Technological Research Council of Turkey(TÜBİTAK) for their financial support during my graduate study.

Last but not the least, I would like to thank to my beloved parents Ümran and Bilal and my brother Salih. I am indebted to them for their everlasting support, encouragement and sincere prayers. I also would like to express my heart-felt gratitude to my friends Nazan, Öykü, Günce and Serhat for their patience and love.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

LIST OF ABBREVIATIONS

ABBRV	Abbreviation
AP	Access Point
APOM	Access Point On the Move
DP	Dynamic Programming
DSKP	Dynamic and Stochastic Knapsack Problem
KP	Knapsack Problem
KSEP	Kalman Based Solar Energy Prediction
MDP	Markov Decision Process

CHAPTER 1

INTRODUCTION

In the era of communication, access to information via Internet is gradually becoming more and more crucial. However, a lot of people on the planet are completely devoid of connectivity because of a remote locale, a lack of infrastructure, a coverage gap and, in some cases, a combination of all tree. Motivated by this fact, recently, major industry players have been pushing for ubiquitous Internet access, because of the large potential increase to their businesses from including users from areas of the world that are still devoid of connectivity. Among all the possible solutions for increasing the pervasiveness of the Internet, the trend of putting mobile Internet service providers (ISP) in the Earth's atmosphere (e.g. Google Loon Project[33] and Facebook Drones [36]) is a novel and promising technology.

There have been a significant amount of studies regarding mobile sinks in conventional networks, not considering renewable energy sources [17],[42],[50]. Furthermore, the advantages of mobile service providers over fixed providers has been shown in previous studies [50], [56]. Some of the these studies focused on determining the optimal path in order to prolong network lifetime [30],[22]. Alkesh et al. [4] proposed a fuzzy logic strategy to select cluster heads of a network and tried to maximize the lifetime of wireless sensor networks with limited batteries.

In addition to the mobility advantage of Access Points (AP), to achieve self-sustainable and environmentally friendly development of service providers, AP's should be powered with renewable energy devices that harvest its own energy, e.g., solar or wind power. The recent advances in the area of energy scavenging enable communication devices with not only low carbon footprint, but also high performance give us the op-

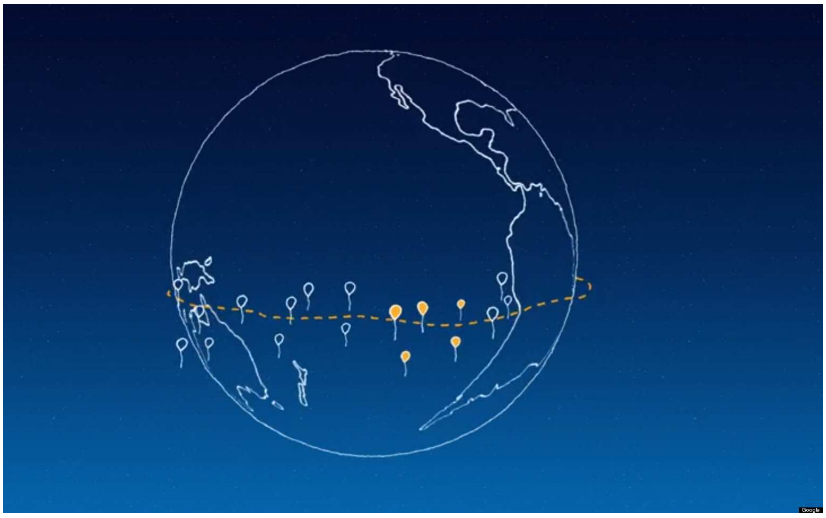

Figure 1.1: Recent industrial efforts toward putting ISP's in the Earth's atmosphere, e.g. Google Loon Project [1]

Figure 1.2: Recent industrial efforts toward putting ISP's in the Earth's atmosphere, e.g. Facebook Drone Project [24]

portunity to design such a system. Hence, Access Point on the Move (APOM) with the capability of energy harvesting from renewable energy sources (solar, wind etc.) is a viable solution to provide ubiquitous Internet to the users in different areas around the world.

However, an APOM with limited capacity of energy and high amount of service requests encounters with a well known problem of resource allocation. Resource management is very critical in Information and Communication Technologies (ICT), especially for the industrial applications like APOM [52]. Considering the case, where there are multiple Access Points (APs) working as routers and communicating with each other in the area of coverage, an AP should decide which users to be served by the current router or deferred to get service from the next router during its route. The deterministic version of the problem has been extensively analysed in [11] that proposes several online heuristics to maximize total utility regarding energy causality constraints in a deterministic environment.

On the other hand, exploiting the uncertainty of both the environment and user characteristics in the real life conditions, the problem is formulated as a dynamic and stochastic resource allocation problem in which utilities and resource demands associated with the appearing users and energy capacity changes in a probabilistic manner. The stochastic and dynamic setup of the problem is quite realistic since energy harvesting is actually a stochastic event [58], and channel conditions randomly change in time [7]. To the best of our knowledge, maximizing the total expected utility subject to energy capacity constraints in a mobile AP with energy harvesting is still a quite open problem. Hence, the main aim of this thesis is to introduce the problem of dynamic and stochastic resource allocation in AP's with movement and energy harvesting capability and to investigate some optimal and suboptimal solutions.

This thesis is divided into 6 chapters. In the following chapter, Chapter 2, recent developments in the resource allocation of mobile routers are reviewed and a detailed background on previously defined Knapsack problems and related solutions are given. Solutions from the literature as well as those proposed by us are introduced. The main contribution of this thesis are given in the Chapters 3, 4, 5. First, a Kalman filter based solar energy prediction method (KSEP) is introduced to be used with resource allocation algorithms in an online manner.

In Chapter 4, first, the system model and the problem statement of dynamic and stochastic resource allocation in APOM are stated. The problem is investigated as a finite horizon 0/1 Stochastic Knapsack problem. A Dynamic Programming (DP) approach is presented as a optimal dynamic solution and its structural properties are investigated. Several suboptimal policies are defined afterwards. The numerical and simulation results showing the performance of the low computational complexity sub-optimal policies with respect to optimal solutions are examined.

In Chapter 5, the problem is investigated as a infinite horizon 0/1 Stochastic Knapsack problem and a Markov Decision Process is submitted to state the behaviour of the model. Value iteration algorithm is presented for the optimal dynamic solution and several suboptimal policies are proposed following the structure of optimal solution.

Chapter 6 summarizes the work conducted in this thesis and outlines some important results. Moreover, future directions with application areas are also stated.

CHAPTER 2

BACKGROUND INFORMATION

A detailed background information on the subject of this thesis is outlined in this chapter. First, some previous studies about the resource allocation for both fixed and mobile service providers including the cases with both conventional and renewable energy sources is summarized. Then, a background information related to the general Knapsack problems which will be used to model the resource allocation problem in this thesis and their applications are provided.

2.1 A Review on the Efficient Resource Allocation in Energy Harvesting Wireless Networks

2.1.1 A Review on Energy Harvesting and Energy Harvesting Systems

Recently, employing energy harvesting (via ambient energy sources such as solar irradiation [9], vibrations [37], [45], and wind [57]) to power transmitters of wireless networks, such as APs has gained tremendous interest [2]. Today, especially solar energy is becoming widely used, due to its high power density compared to other sources of ambient energy [38]. With the development of high performance and low carbon-footprint energy scavenging devices, the main aim in the energy allocation to different tasks becomes to adapt the energy consumption of a transmitter to the energy harvesting characteristics. A conservative approach in energy allocation will cause the energy storage costs not to mention the low utility, on the other hand, a greedy approach will result in inefficient use of resources and reduces the total reward. Therefore, nowadays, the focus of the research community has been changed into providing energy neutrality in wireless networks rather than low energy consumption to design energy efficient communication systems.

Energy harvesting systems do not instantly consume the energy harvested from the environment and hence they need an energy storage device. Energy harvesting circuitry to efficiently convert and store renewable energy are available [41], [51]. This device can be a rechargeable battery or a supercapacitor [58]. In most of the theoretical works [7], [25], on the other hand, the energy is assumed to be stored in an energy buffer. The time slotted energy harvesting system where $e(k)$, $c(k)$ and $h(k)$ denotes the available energy, energy consumption and the harvested energy at slot k is defined as follows:

$$e(k+1) = e(k) - c(k) + h(k) \tag{2.1}$$

The effect of the mobility for outdoor solar energy harvesting was outlined in [7] and [44]. Although the harvested energy for a fixed transmitter with static energy harvester can be considered partially predictable and deterministic [52], it is rather stochastic on a mobile device like APOM. The majority of harvestable energy is stochastic in nature [58], considering solar harvesting where solar radiation can be randomly reduced by clouds passing over. To deal with this ambiguity, the stochastic models and analysis should be adopted to investigate the performance of energy harvesting communication systems and to design new methods that consider the randomness of the energy source while optimizing system performance.

The stochastic energy harvesting process is modelled in several ways. Some of the literature follows the assumption that at each slot in a time slotted system, the size of the energy buffer increases by one with a probability q [25]. This assumption is quite valid, considering short time horizons with the statistics of the harvest that do not change quickly. The other model includes the exponential interarrival times for the energy harvests, that is Poisson process [59]. The solar harvesting is further considered as a Markov process as in Figure 2.1 introducing the memory [7], [58] where h_1, h_2 are harvest states and q_{ij} stands for the transition probabilities from state i to j.

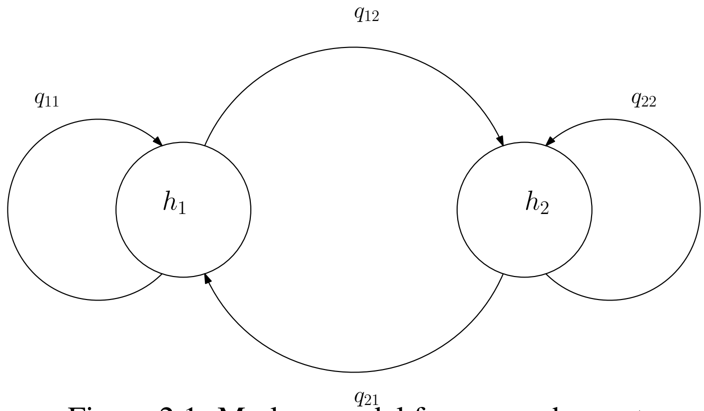

Figure 2.1: Markov model for energy harvests

In a slotted system, it is assumed that a communication device can only use the unconsumed energy left from the previous slot plus the energy harvested in the previous slot. The energy harvested during the slot of the transmission event is not taken into account. Transmission time, typically, is considerably smaller than the time required for harvesting sufficient energy for one event. This is reasonable considering the following illustration from [58]. For example, a MICAz mote requires $4.73mJ$ for transmitting a 132B packet and thus requires 2.37 seconds for a solar energy harvester with a conversion rate of $2mW$ per second (with a solar cell size of 18 cm^2) to generate sufficient energy for transmitting one packet. On the other hand, protocols such as IEEE 802.11 and IEEE 802.15.4 require only few tens of milliseconds for channel access even under saturated traffic conditions for moderate network sizes [58]. Therefore, the assumption about ignoring the harvesting during transmission does not result in a noticeable deviation in the results.

Although, some literature work take the imperfections of the energy harvester such as battery leakage into account [14], in most of the studies like this thesis it is ignored as being secondary effects not significant in the behaviour of the system of interest.

With the proliferation of energy harvesting devices, resource allocation problems of conventional sources need to be revised. As also stated in [28], conservative energy expenditure may lead to missed recharging opportunities if the battery is already full. On the other hand, aggressive usage of energy may result in reduced coverage or connectivity for certain time periods, that could make the BS temporarily unavailable to transfer time-sensitive data. In practical applications, this may sometimes create jeopardous situations and may lead to loss of production [39]. Hence, new resource allocation and scheduling schemes need to be developed to balance these contradictory goals, in order to maximize the network performance.

2.1.2 A Review on Dynamic Resource Allocation in Wireless Networks with Energy Harvesting

Resource allocation is very critical in wireless applications like APOM where operation is battery dependent. Among all the resources of a communication device such as memory, I/O bandwidth, CPU etc., mostly energy is the limiting case due to increasing amount of energy demand in all applications and its dynamically changing nature with the recent developments in energy harvesting. In this thesis, a router like APOM should use its energy efficiently to maximize the expected amount of utility which is actually the goal of this thesis.

B. T. Bacinoglu et al. [6], [7] propose some scheduling policies in order to efficiently allocate renewable energy resources over fading channels. In [19] and [54], authors also implement related scheduling policies on software define radio. Moreover, for the application to low energy Bluetooth devices, some duty cycle optimization methods in energy harvesting Wireless Sensor Networks (WSN) are also considered as in [3]. On the other hand, optimization methods for the feedback system in a multiuser miso downlink communication are also investigated in previous studies, considering the case where users have energy harvesting capability [46] and [47]. In [35], [21] and [20], the scheduling problem in single hop wireless networks with energy harvesting nodes is formulated through a restless multi-armed bandit problem and the performance is shown to behave very close to optimal.

The closely related work of stochastic environment but fixed locations is [7] and [25]. In [7], authors formulate an online discrete power selection scheme for stochastic energy arrivals and channel gains. Both optimal iterative solutions and suboptimal heuristics with lower time complexities are introduced in a finite time horizon. The concavity of power-rate relation is taken into account to determine optimal policies. Kashef et. al. [25] considers a binary decision problem to transmit or defer several tasks according to a stochastic model based on Gilbert-Elliot channel. In that problem setting, authors prove that a threshold based approach is optimal, however do not state any optimal or suboptimal threshold function due to its computational complexity.

Until recently, there has been no study related to the resource allocation of mobile Access Points with energy harvest capability. Several of the recent studies on the topic are concerned with finding optimal routing paths [43]. Xie et al. [55] address the problem of colocating the mobile service provider on the wireless charging machine with the objective of minimizing energy consumption. The closely related works of Ren and Liang [44], [31] consider a distributed time allocation method to maximize data collection in energy harvesting sensor networks while defining a scenario of a constrained path with all sensors having their own renewable energy sources.

Maximizing the total utility subject to some energy constraints in a mobile AP with energy harvesting is still a quite open problem. The deterministic version of the problem has been extensively analysed in [11] that proposes several online heuristics to maximize total utility regarding energy causality constraint in a deterministic environment.

Orthogonal to these existing works, this thesis will deal with a mobile Access Point which is powered through energy harvesting, aims to maximize the expected total utility to the users with probabilistic characteristics and appearing in a sequential manner. Differently from most existing works, the problem will be analysed for both finite and infinite horizon cases. To the best of our knowledge, there is no prior work addressing the objective of this thesis, that is, maximizing the total expected data service provided by a mobile service provider with energy harvesting capability, and employing dynamic heuristics in a stochastic manner.

2.2 A Review on Knapsack Problems

2.2.1 Knapsack Problems and NP-hardness

The knapsack problem is a well-known combinatorial optimization problem with numerical practical applications such as operational research, economics, computer science etc. The classical problem is defined as follows: several objects with given known weights (w_i) and values(v_i) must be packed in a 'knapsack' of given capacity in order to maximize the total value of selected items under the constraint of total capacity consumption (W). The problem definition of the classic knapsack problem is given as follows:

Problem 1.

$$\textit{Maximize: } \sum_{i=1}^{N} v_i x_i \tag{2.2}$$

$$\textit{subject to: } \sum_{i=1}^{N} w_i x_i \leq W, \tag{2.3}$$

$$x_i \in 0, 1, ..., c_i \tag{2.4}$$

When c_i goes to infinity, this is a an unbounded knapsack problem and if it is a bounded number it is called a bounded knapsack problem. The most common problem being solved is the 0/1 knapsack problem, a bounded special case, which limits the number of each kind of item to either zero or one, i.e. $x_i = \{0, 1\}$ [34].

A decision form of knapsack problem (Can a total value of more than a value V be achieved under the constraint of the total weight W?) is NP-complete, thus there is no possible solution to the knapsack problem defined in Problem 1 that works both correct and fast (polynomial-time) for all cases. Therefore, the classical Knapsack optimization problem stated above is an offline and deterministic problem which is proven to be NP-hard, that is, it is not possible to compute an optimal solution in polynomial time [18].

Although it has received less attention than the deterministic case, stochastic approaches are much more suited to the real life scenarios because of the uncertainty in the environment. In the dynamic and stochastic version of the knapsack problem items arrive sequentially over time and their value/weight combination is stochastic but becomes known to the designer at arrival times. Instantaneous decisions on each impatient user's arrival should be made to maximize the expected utility. The decisions are irreversible, that is users can not be recalled later, once a user is rejected value associated with that user will be lost. While in deterministic problem setting, an optimal rule of placing items in the knapsack is determined, an optimal policy is searched for the adverse.

2.2.2 Observations and Methods

There has been considerable amount of proposed solutions to the offline solution of the problem including the dynamic programming and the branch and bound methods [34]. Approximation ratios are often used to indicate the performance of approximation algorithms, that is, an approximation ratio of α corresponds to the algorithm that value obtained by optimal algorithm is less than the α times the value of that algorithm. In the offline and deterministic setting with problem size n, the lowest approximation ratio obtained by to date is $1+\epsilon$, where $0 < \epsilon < 1$ which takes the $O(n)$ number of computation size to achieve [29]. However, it has been shown that without making any assumptions, it is not possible to obtain a constant competitive ratio for online knapsack problems [60]. In an online problem, under some assumptions on item characteristics (values and weights), in [60], competitive online heuristics with efficient approximation ratios are achieved.

An application of stochastic knapsack problems extensively studied in the literature is stochastic scheduling problems, where job durations are modelled to be random variables with known probability distributions. In [13], a stochastic variant of the NP-hard 0/1 knapsack problem is considered, in which item values are deterministic and item sizes are independent random variables with known, arbitrary distributions. Items are placed in the knapsack sequentially, and the act of placing an item in the knapsack instantiates its size. Authors propose a solution policy that maximizes the expected value of selected items. Stochastic knapsack problems with random values but deterministic sizes have also been investigated by several authors such as Carraway et al. [10], and Steinberg and Parks [40], trying to compute a fixed set of items placed in the knapsack that provides maximum probability of achieving a target value and the constraint on the minimum probability of excessing capacity. Another somewhat more related study, namely stochastic and dynamic knapsack problem by Kleywegt and Papastavrou [26], Papastavrou et al. [27], involves items that arrive in an online manner following a stochastic process such that the exact characteristics of an item are not known until it arrives, at which point in time a controller must irreversibly decide to either accept the item and include to knapsack, or reject the item. However, their results are dependent on some assumed distributions for capacity requirements and utility, namely, that the conditional probability density functions for the capacity requirements given returns are concave. Thus, their method is not general and this requirement may not be met in real life.

Pak and Dekker [40] consider a decision model where an item is accepted if the value associated with the item is higher than the the expected future revenue lost by accepting the item according to a simulation of future demand occurrences. Although they have reported very satisfying results, the use of it was still too computation intensive to be used in online decision making at the time.

The Markov Decision Processes (MDP) formulation is frequently applied to model the stochastic knapsack problems using the available energy as the state and the values of appearing users as instantaneous rewards especially in the infinite horizon cases [27], [26], Han et. al. [23]. A one dimensional MDP relaxation is suitable with dynamic knapsack exploiting the fact that expected value in a Knapsack at each time only depend on the available capacity state and the characteristics of incoming users and the environment. Amaruchkul et. al. [5] uses a dynamic programming approach policy for optimizing the accept/reject decisions. Dynamic programming approach is indeed an optimal iterative solution to the Markov Decision Processes.

An overview of the previous observations related to knapsack problems have been given in the Table 2.1.

The problem studied in this thesis is a generalization of dynamic stochastic 0/1 knapsack problem with multiple constraints. In each energy harvest interval, the energy reserve is the newly harvested energy (with known probability distribution) plus the energy saved from previous intervals. Energy expenditure in an interval cannot be greater than the energy capacity, i.e. energy causality. The overall structure may be modelled as a dynamic and stochastic knapsack problem with randomly increasing capacity due to energy harvests of coming at each slot stochastically.

There has been several research in the last century related to optimal and suboptimal solutions of Knapsack problems. On the other hand, the dynamic and stochastic knapsack problem with randomized incremental capacity, which is of interest, is still quite open. To the best of our knowledge, a competitive solution has not appeared in the

Table 2.1: An Overview of Previous Studies

	Decision variables	Method used
Papavastrou et. al. (2001) [27]	Accepting/rejecting incoming items to maximize the expected utility or returns	Dynamic programming, deadlined knapsack problem with stochastic weights and returns
Pak and Dekker (2004) [40]	Bid-prices for weight and volume, knowing which fixes the accept/reject decision	Simulation, deterministic offline threshold generation
Amaruchkul et. al. (2007) [5]	Accept/reject decision for booking offers	Heuristics based on dynamic programming using bounds on some statistic relationships
D. Chakrabarty et. al. (2010) [60]	Accept/reject decision of items in an online knapsack problem	Threshold generation with respect to capacity fullness under some assumptions on item efficiencies
E.T. Ceran et. al. (2014) [11]	Accepting/rejecting incoming users as a dynamic resource allocation problem	Heuristics with threshold function with respect to capaticy fullness and competitive ratio analysis of threshold policies

literature to the Dynamic and Stochastic Knapsack problem with incremental capacity. Moreover, there has been no previous study relates the Stochastic and Dynamic Knapsack problem with mobile access points with randomized energy harvesting, which is the case in this thesis.

2.3 A review on the Deterministic Knapsack Problem with Incremental Capacity formulation of the Resource Allocation Problem in Wireless Access Point on the Move

In [11] and [15], the problem of online resource for an Access Point on the Move (APOM) scenario has been investigated in a deterministic manner, which is also applicable to available mobile sinks in WSN topologies nowadays. Energy harvesting in a deterministic manner have also been taken into account by developing heuristic solutions to the problem. Intuitively considered adaptive threshold based policies have been considered in an online scenario with low time complexity constraints. A user is served if its efficiency metric is better than a threshold, where the threshold also varies depending on availability of energy and closeness of an energy harvesting instant.

Some bounding assumptions on the user efficiency (v/w) have been made to achieve a constant competitive ratio. First, monotonic and piecewise monotonic threshold functions have been proposed, based on a previous work [60]. Under some assumptions in the energy harvest arrivals, the related competitive ratio is shown to be $ln(U/L) + 1$ where the user efficiency ratios (v/w) is assumed to be upper bounded by U and lower bounded by L and weights of each user is much smaller that capacity.

Then, as an original contribution, two different threshold functions using a Genetic Algorithm and Rule Based (Fuzzy) approaches have been devised. The competitive ratios of the algorithms were measured and compared using Monte Carlo simulations. Experimental results demonstrate that the proposed decision methods using different threshold functions for the resource allocation problem of the energy harvesting Access Point on the Move are efficient in satisfying a certain competitive ratio even in the worst cases [15].

The work stated in [11] has satisfying results for online admission problem defined as the classical knapsack problem. However, in real scenarios, energy harvests and user characteristics may not behave deterministically or be fully predictable. Therefore, the stochastic and dynamic version of the problem, which is the subject of this thesis, will be examined to fill this gap.

CHAPTER 3

PREDICTION BASED DYNAMIC RESOURCE ALLOCATION IN ENERGY HARVESTING WIRELESS NETWORKS WITH A FIXED TRANSMITTER

Resource allocation is a very common problem in Information and Communication Technologies (ICT), especially for energy limited applications like APOM. With the increasing demand of energy consumption and increase in the carbon footprint of ICT, there has been a significant amount of efforts to seek ways to energy efficiency.

Many efficient resource allocation algorithms for energy harvesting wireless networks depend on the a-priori knowledge of future energy harvests that arrives in different points in time. However, this kind of information in the offline manner may not be available in practical situations as also stated in Section 2.1.1.

To illustrate, the proportional-fair energy harvesting resource allocation problem first formulated in an offline manner in [53]. The problem was shown to be a *biconvex* problem which is nonconvex, and has multiple optima. The optimum off-line schedule developed in [53] (which assumes that the energy arrival profile at the transmitter is deterministic and known ahead of time in an off-line manner) Moreover, to reduce computational convexity A simple heuristic, called PTF (Proportional Time Fair), that can closely track the performance of the BCD solution was developed in [53]. However, PTF heuristic still needs a-priori knowledge on energy arrivals. To obtain a dynamic and proportional-fair energy harvesting resource allocation method, the offline solution must be combined with a promising prediction method for energy arrivals.

Thus, levaraging the PTF algorithm which is an original work of [53], an online PTF-On algorithm that operates two algorithms in tandem: A Kalman filter-based solar energy prediction algorithm, and a modified version of the PTF is proposed in [52].

This section mostly summarizes the work conducted in [52] based on the idea of solar energy harvest estimation and concatenating the prediction algorithm with any novel offline resource allocation method to obtain an online method.

3.1 Solar Energy Prediction based on Discrete Kalman Filter

An adaptive solar energy prediction algorithm is proposed considering the energy harvesting characteristics of solar harvester in [52]. Outdoor solar irradiation exhibits a daily periodicity. However, there are seasonal as well as short-term variations. In order to find a novel prediction method with low computational cost and high performance, a prediction algorithm based on discrete Kalman filer is proposed.

The illustration of the energy arrival process is given in Figure 3.1 [16].

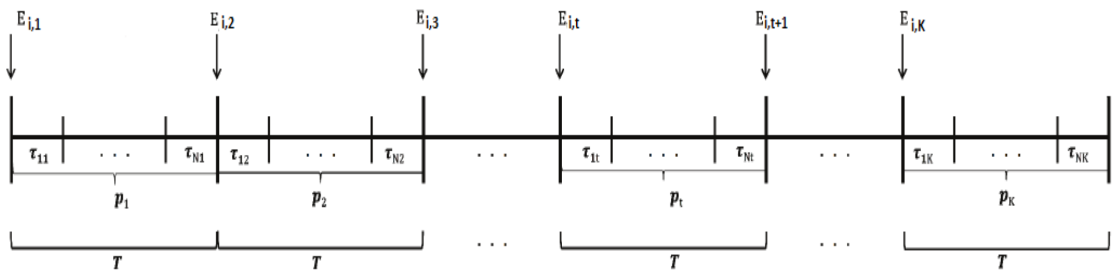

Figure 3.1: One of the multiple frames in a timeline. The highlighted frame, frame i (of 24 hours), includes K energy arivals. The time between consecutive arrivals is allocated to N users [16].

Note that, energy arrivals are modelled to be periodic. Not all generality is lost, since harvest amounts are arbitrary and the absence of a harvest in a certain duration can be expressed with a harvest of amount zero for the respective slot. The amount of energy harvested from the environment at the beginning of time slot t of frame i is $E_{i,t}$, as illustrated in Fig. 3.1. The BS chooses a power level p_t and a time allocation vector $\tau_t = (\tau_{1t}, ..., \tau_{Nt})$, for each time slot t of the frame, where $p_{nt} = p_t$ is the transmission power for gateway n during slot t and, τ_{nt} is the time allocated for transmission to gateway n during slot t.

Kalman filter is introduced to forecast the energy arrivals within a frame, for a BS powered with solar panel. Solar power for fixed locations is shown to be piecewise stationary over half an hour periods [49], which may be called as slots. Consider sub-hourly prediction of the energy arrivals for a frame of 24 hours (facilitating the daily predictions) as an example, and, formulate the Kalman filter for the following state and measurement models:

$$x(k+1) = \alpha_1 x(k) + \alpha_2 x(k-47) + \beta_1 y(k) + w(k) \tag{3.1}$$

$$z(k) = x(k) + v(k) \tag{3.2}$$

where x and z represent the state (energy level) and the measurement respectively. This model is mainly based on the idea that; due to the diurnal cycle of a day, the amount of energy that will be harvested in the $(k+1)^{th}$ sub-hour of an arbitrary day, $x(k+1)$, should be related to the energy harvested in the k^{th} sub-hour of the same day, $x(k)$, the solar irradiation received in the k^{th} sub-hour of the same day, $y(k)$, and, the energy harvested in the $(k+1)^{th}$ sub-hour of the previous day (the energy that was harvested 48 sub-hours ago: $x((k+1)-48) = x(k-47)$), $x(k-47)$. In (3.1), $w(k)$ is a modelling error, which represents the effects of the uncontrolled events on the harvested energy (such as shadowing caused by clouds passing through, disturbance to the solar panel, or damage due to malicious act, etc.). In this thesis, it is modelled as Gaussian i.i.d. with zero mean and variance σ_w^2. The parameters α_1,α_2 and β_1 represent the weights assigned to emphasize the importance of the parameters that will be used for prediction. In the measurement model, v denotes the measurement noise and it is also modelled as IID Gaussian with zero mean and variance σ_v^2.

By considering that there are 48 half-hours in a day, the overall state equations can be re-stated in matrix form as in (3.6).

Now, an augmented state vector, $\overline{\xi_k}$, is defined which contains the energy amounts harvested today:

$$\begin{bmatrix} x(k+1) \\ x(k) \\ x(k-1) \\ \vdots \\ x(k-45) \\ x(k-46) \end{bmatrix} = \mathbf{A} \begin{bmatrix} x(k) \\ x(k-1) \\ x(k-2) \\ \vdots \\ x(k-46) \\ x(k-47) \end{bmatrix} + \overline{\beta} y(k) + \overline{\Gamma} w(k) \tag{3.6}$$

$$\overline{\xi_k} = \begin{bmatrix} x(k) & x(k-1) & \dots & x(k-46) & x(k-47) \end{bmatrix}' \tag{3.3}$$

Moreover a new matrix is defined: **A**, column vectors $\overline{B}$, and $\overline{\Gamma}$ as follows:

$$\mathbf{A} = \begin{bmatrix} \alpha_1 & 0 & 0 & \dots & 0 & 0 & \alpha_2 \\ 1 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 & 0 \\ \vdots & & & \ddots & & & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 & 0 \end{bmatrix} \tag{3.4}$$

$$\overline{B} = \begin{bmatrix} \beta_1 & 0 & \dots & 0 & 0 \end{bmatrix}' \qquad \overline{\Gamma} = \begin{bmatrix} 1 & 0 & \dots & 0 & 0 \end{bmatrix}' \tag{3.5}$$

Thus, the state model in (3.6), and the measurement model in (3.2) reduce to

$$\overline{\xi_{k+1}} = \mathbf{A}\overline{\xi_k} + \overline{B}y(k) + \overline{\Gamma}w(k) \tag{3.7}$$

$$z(k) = x(k) + v(k) \tag{3.8}$$

which is structurally equivalent to the “truth” model described in (5.27) (in page 252) of [12]. Thus, by applying the Discrete-Time Linear Kalman Filter described in [12], it is possible to predict the amount of energy arrival in the next sub-hour by only using the amount of energy arrival in this sub-hour, the solar irradiation received in this sub-hour and, the arrival in the previous day’s next sub-hour. Please note that, in order to compute the best weights α_1, α_2 and β_1 that will be used for simulations, a data fitting method can be described as follows: By using 18 days’ data (real power measurements belonging to 01.10.2009-18.10.2009 for Amherst, Massachusetts, USA) provided by Navin Sharma [48], a Newton algorithm is designed to minimize the Mean Squared Error (MSE) between the data obtained from real measurements and the estimated data according to the state and measurement models in 3.1 and 3.2 .

Thus, the objective function to be minimized by the Newton algorithm is described below:

$$\frac{1}{N}\sum_{k=1}^{N}(z(k) - z_m(k))^2 \tag{3.9}$$

where z denotes the data obtained from actual measurements and z_m denotes the estimated data obtained from the models, in (3.1) and (3.2). Note there are 48 subhours (slots) for a day (frame), and need the past day's data at the same subhour for the prediction of a subhour's solar irradiation. For 17 days (17 days= 816 half-hours) data [48], the objective function can be stated as:

$$\frac{1}{816}\sum_{k=48}^{863}(z(k+1) - (\alpha_1 x(k) + \alpha_2 x(k-47) + \beta_1 y(k)))^2 \tag{3.10}$$

The simulation results, provided in Section 3.3, show that the best values for defined weights, α_1,α_2,β_1 are 0.7184, 0.1439, and, 0.0063 respectively, when the $x(k)$'s are in terms of kilojoules and the initial values for the data fitting operation of α_1,α_2,β_1 are taken as 0.9, 0.1 and 0.01 respectively.

Furthermore, considering the equivalence of the "truth" model in [12], the prediction and update equations of the Kalman estimator can be stated as follows:

$$\overline{\xi_{k+1}^-} = \mathbf{A}\overline{\xi_k^+} + \overline{B}y(k) \tag{3.11}$$

$$\overline{\xi_k^+} = \overline{\xi_k^-} + \mathbf{K}_k[z(k) - \overline{\xi_k^-}] \tag{3.12}$$

where ξ_k^- and ξ_k^+ denotes the pre-measurement and post-measurement states respectively. Moreover, K, R , I, P are the Kalman gain function, measurement noise matrix, identity matrix and error covariance matrix respectively and defined as:

$$K_k = P_k^- - [P_k^- + R] \tag{3.13}$$

$$P_k^+ = [I - \mathbf{K}_k]P_k^- \tag{3.14}$$

$$P_{k+1}^- = \mathbf{A}P_k^+\mathbf{A}^T + \overline{\Gamma}\sigma_\omega^2\overline{\Gamma}^T \tag{3.15}$$

Similar to the notation of states ξ_k^- and ξ_k^+, P_k^+ and P_k^- denotes the pre-measurement and post measurement error covariance matrices. Thus, KSEP with the state and measurement models is obtained and given in (3.1) and (3.2).

The obtained solar prediction algorithm can be used with any viable offline-deterministic resource allocation algorithm, results in a competitive online heuristic. In the next section, as an example of efficient resource allocation algortihmsa based on apriori knowledge of energy arrivals, a proportional fair resource allocation problem [53] and the solution with solar prediction algorithm are detailed.

3.2 Proportional Fair Resource Allocation Algorithm with Offline Knowledge of Energy Harvest

In this section, for completeness, Proportional Fair Resource Allocation Algorithm and the Power-Time-Fair (PTF) heuristic operates in an offline setting is restated by following the work submitted in [52] (the times and amounts of energy harvests are known at the beginning of the frame).

The total achievable rate for user n (the number of bits transmitted to user n) within a frame, R_n. The goal is to maximize a total utility, i.e., the log-sum of the user rates $\sum_{n=1}^{N} log_2(R_n)$, which is known to result in proportional fairness [32]. Lets define $R_n = \sum_{t=1}^{K} \tau_{nt} W \log_2 \left(1 + \frac{p_t g_n}{N_o W}\right)$. Thus, the constrained optimization problem is obtained, Problem 2 in [52], where (3.16) represents the nonnegativity constraints for $t = 1, ..., K$, $n = 1, ..., N$. The formulations in (3.17), called time constraints, ensure that the total time allocated to users does not exceed the slot length, and, every user gets a non-zero time allocation during the frame. Finally, (3.18), called energy causality constraints, ensure no energy is consumed before becoming available.

Problem 2.

$$\textit{Maximize: } U(\overline{\tau},\overline{p}) = \sum_{n=1}^{N} \log_2 \left(\sum_{t=1}^{K} \tau_{nt} W \log_2 \left(1 + \frac{g_n p_t}{N_o W} \right) \right)$$

$$\textit{subject to: } \tau_{nt} \geq 0 \ , \ p_t \geq 0 \tag{3.16}$$

$$\sum_{n=1}^{N} \tau_{nt} = T_t \ , \ \sum_{t=1}^{K} \tau_{nt} \geq \epsilon \tag{3.17}$$

$$\sum_{i=1}^{t} p_i T_i \leq \sum_{i=1}^{t} E_i \tag{3.18}$$

An sub-optimal heuristic with low computational complexity is defined as in the work of [53]:

1. **For Power Allocation:** Assign nondecreasing powers through the slots, as follows:
   (a) From a slot, say i, to the next one $i+1$: If harvested energy decreases, defer a Δ amount of energy from slot i to slot $i+1$ to equalize the power levels. Do this until all powers are nondecreasing, and, form a virtual nondecreasing harvest order. Note that energy causality is still maintained with this operation.
   (b) By using the virtual harvest order, assign nondecreasing powers through the slots, i.e., in each slot, spend what you virtually harvested at the beginning of that slot.
2. **For Time Allocation:** For the allocation found in 1), let, $B_{nt} = R_{nt}T$ be the number of bits that would be sent to user n if the whole slot (of length T) was allocated to that user. Assign the first slot to the user who has the maximum rate, R_{nt}, in that slot. For the other slots, apply the following: At the beginning of each slot, $t \in \{2, \dots, K\}$, determine the user with the maximum β where, $\beta_n = \frac{B_{nt}}{\sum_{i=1}^{t-1} B_{ni}}$. Then, assign the whole slot to that user. If multiple users share the same β, then, allocate the slot to the user with the best channel.

The problem may be solved in an online manner by first executing the solar prediction algorithm defined in Section 3.1 at each slot. Once the energy harvest at all

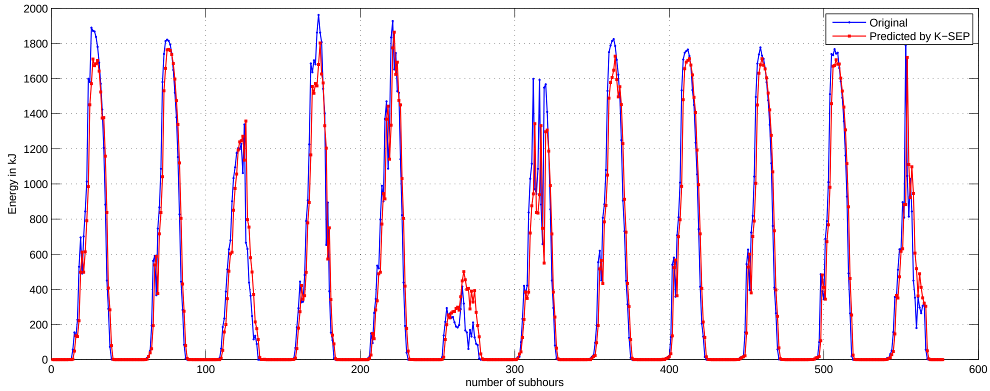

Figure 3.2: Performance of K-SEP with respect to the real measurements provided by the University of Oregon Solar Radiation Laboratory; belonging to 07.05.2009-19.05.2009 for Salem, MA, USA.

future slots are calculated, the problem can be solved by PTF that closeness to optimality previously proved [53]. After restating the PTF algorithm as an example of efficient resource allocation algorithms in wireless networks, PTF is combined with the Kalman Based Prediction Algorithm (KSEP) and the performance results are provided in the next.

3.3 Performance Evaluation and Numerical Results

First the simulations conducted on standalone K-SEP algorithm without combining with any resource allocation algorithm to measure how well the K-SEP predicts the solar energy comparing with real datas. To achieve this, the performance of K-SEP is tested with the solar irradiation measurements obtained from the University of Oregon Solar Radiation Laboratory. Obtained results related to the performance of K-SEP algorithm can be seen from Figures 3.2 and 3.3 for 12 days belonging the dates between 7-19 May 2009 and 1-13 January 2011 arbitrarily. The solar data is obtained from the station at the Oregon Department of Energy in Salem, MA, USA. The performance of the K-SEP algorithm is very close to the original data which the all energy harvests are known in an offline manner.

To further test the performance and robustness of K-SEP, simulations in [16] are also

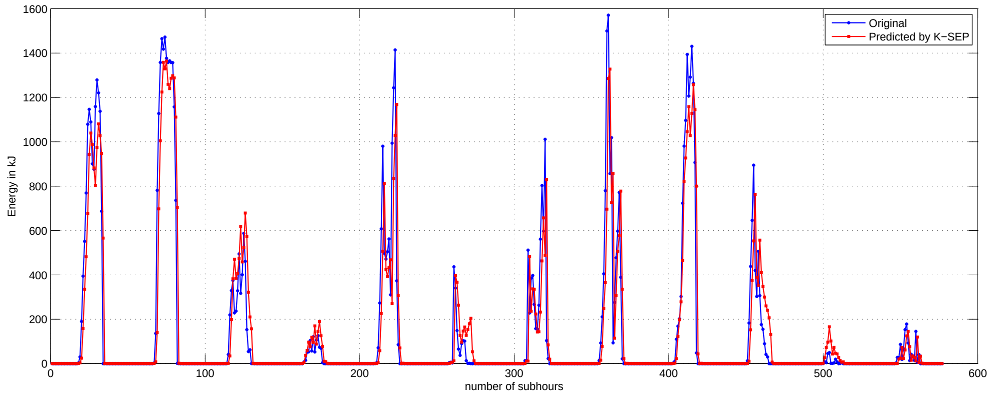

Figure 3.3: Performance of K-SEP with respect to the real measurements provided by the University of Oregon Solar Radiation Laboratory; belonging to 01.01.2011-13.01.2011 for Salem, MA, USA.

shown on a different solar irradiation data obtained from [48]. To be able to compare with simpler prediction methods, a Simple Solar Energy Predictor (S-SEP) is adopted, which does not use the data that was harvested in the previous slots (mainly predicts the amount of energy that will be harvested in today's k^{th} sub-hour as the average of the energy arrival amounts of the past two days' k^{th} sub-hours), to predict the next arrival. Figures 3.4, 3.5, 3.6 illustrate the performances of the two predictors for three days in which S-SEP performs the best, the second best, and the worst in its 16 days's performance. As it can be seen from the figure, K-SEP outperforms S-SEP at all instances. However, even S-SEP as a simple prediction method of solar energy harvests provide some usefulness for the solution of online fashion 2.

In addition, it is more important to note that harvesting energies predicted by K-SEP algorithm always follow the original energies obtained from real measurements as shown in Figures 3.2, 3.3, 3.4, 3.6 and 3.5. By considering the numerical and simulation results conducted with K-SEP and S-SEP, two main conclusions can be derived: the advantage of using a prediction method for the estimation of solar energy harvesting with an offline allocation algorithm (PTF) in tandem and the novelty of K-SEP which performs very close to optimal situation where energy harvests are known a priori.

Till now, performance of KSEP algorithm has been shown with the real datas taken

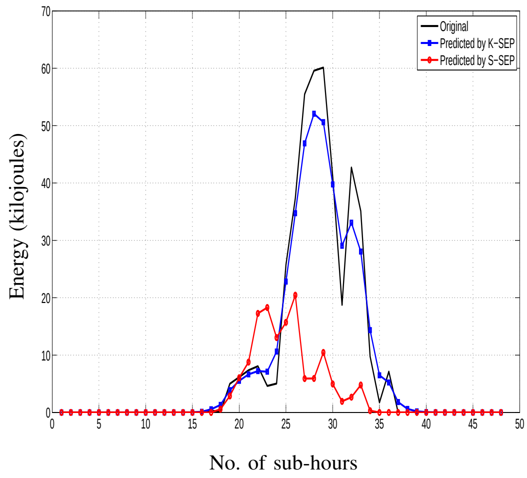

Figure 3.4: Performances of K-SEP and S-SEP compared with the real power measurements provided in [48]; belonging to 16.10.2009 for Amherst, MA, USA [16].

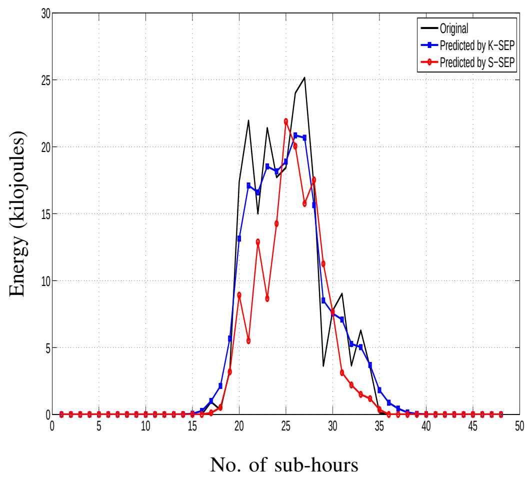

Figure 3.5: Performances of K-SEP and S-SEP compared with the real power measurements provided in [48]; belonging to 10.10.2009 for Amherst, MA, USA [16].

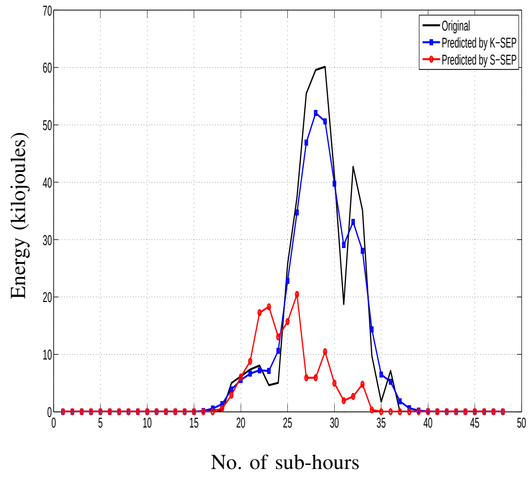

Figure 3.6: Performances of K-SEP and S-SEP compared with the real power measurements provided in [48]; belonging to 04.10.2009 for Amherst, MA, USA [16].

Table 3.1: Average Over 14 Frames of 3 Gateways' Throughputs (Gigabytes/day)

Algorithm	Gateway 1	Gateway 2	Gateway 3	Total
SG+TDMA	188.0338	133.7222	105.8971	427.6531
PTF	458.8113	340.0307	269.5249	1068.3669
PTF-ON	473.5146	326.1724	244.9720	1044.6595

from two different sources. Next, the numerical and simulation results related to the proposed online heuristic, PTF with energy prediction are proposed to further test the KSEP method. Throughout simulations, the following setup has been used: $W = 10MHz$, $N_o = 10^{-19}W/Hz$. For the sake of an example, we suppose that there are three sensor networks. The path loss of the gateways are 78, 92, and, 100 dB respectively. In Figures 3.7 and 3.8 and Table 3.1 the performance of the proposed algorithm with the performance of the "Spend What You Get" policy (where the amount of energy harvested at the beginning of a slot is completely spent during that slot) combined with TDMA time allocation, and with the performance of the offline PTF heuristic that is proved to operate very close-to-optimal are compared.

It is proven by numerical evaluations that the joint prediction and resource allocation algorithm that proposed performs very close to the optimal offline resource allocation. The developed framework can be applied to various other scenarios, including various types of energy harvesting (such wind, vibration etc.) and various types of applications and utility functions (such as delay, reliability etc.) However, in the case

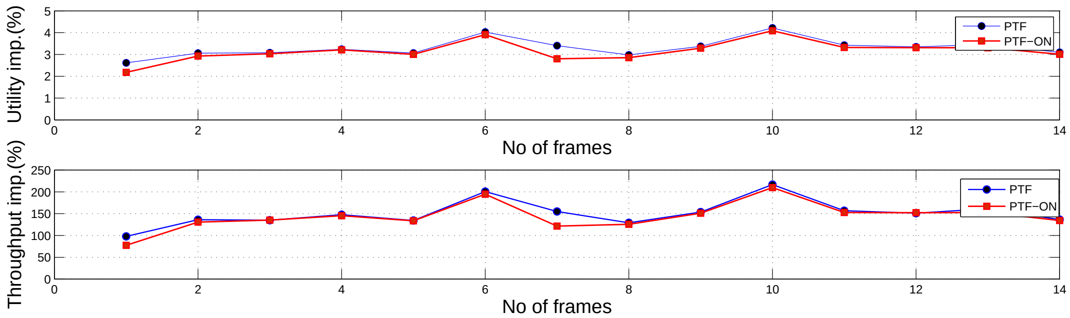

Figure 3.7: Utility and Throughput Improvements of PTF and PTF-On over SG+TDMA on Amherst, MA solar irradiation data.

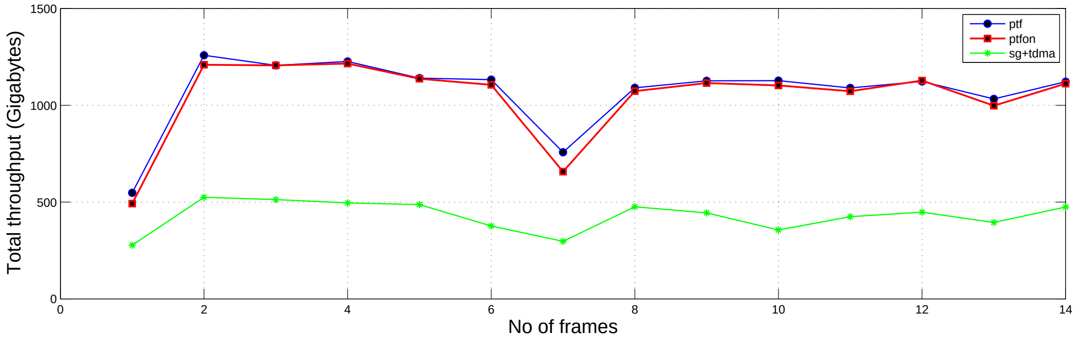

Figure 3.8: Total Throughput of the three gateways (in Gigabytes) over 14 Frames on Amherst, MA solar irradiation data.

of mobile APs with unpredictable and random energy arrivals, a promising resource allocation algorithm calls for a dynamic and stochastic policy. In the next chapters, dynamic problem of expected utility maximization will be investigated to address this problem.

CHAPTER 4

FINITE HORIZON DYNAMIC EXPECTED UTILITY MAXIMIZATION UNDER ENERGY CONSTRAINTS

4.1 System Model and Problem Statement

4.1.1 System Model

Consider an Access Point (AP) moving on a path to provide service to the different users in a large geographical area. It is powered by devices capable of energy harvesting, that is, renewable energy sources such as solar, wind, vibration etc. are employed to meet the energy expenditure of the downlink communication system. Most probable source of energy will be the solar irradiation exploiting the fact that high potential of energy harvesting and assumption of mobile outdoor system above the level of clouds as stated in Section 2.1. The overall system description can be seen explicitly from Figure 4.1.

Following a widely adopted assumption about energy replenishment, rate of energy replenishment is lower than the energy consumption rate. Therefore, a controller should carefully select the best to serve among possible users to maximize the utility considering the known user characteristics and future energy harvests in an offline problem setting. However, in real scenarios, this kind of information about the future is not available. In the online setting, users appear one by one and the APOM makes a binary decision whether to serve the user or not. Once the decision on a user has been made, there will be no re-evaluation on that user. Decision on whether to serve a user or not is made upon the reward and the weight associated with that user. The reward

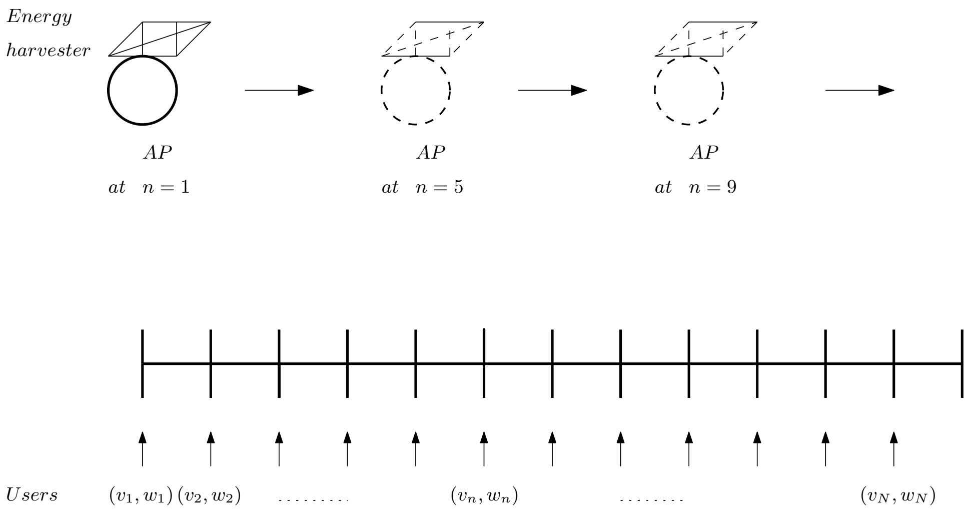

Figure 4.1: Overall system model of APOM with renewable energy sources

could be the utility of providing service to that user, and the cost may be related to the power required to serve this user. If there was no energy limitation at the AP, it would serve all available users to achieve maximum utility. However, when energy replenishment rate is slower than power it would require to serve all users, the APOM has to refuse a number of users.

Hence the problem is an online decision problem with the goal of picking a set of users to maximize the total utility under the energy constraints regarding the causality of this energy, and the causality of the information gathered about users.

A dynamic scenario is considered where users associated with random values and weights appear to the APOM sequentially during its route. The characteristics of the users (value and weight) are not known deterministically, but they become known to the agent following the related probability distributions. The value of a user can be defined as the instantaneous utility gained by providing service to that user. The weight may also be related to required energy consumption of a user. Associated value and weight distributions depend both on the requirements of a user and the channel gain between the APOM and the user (when the user has a poor link to the APOM, for example, the cost of serving it will be high or changes in variations of the link quality may result in a higher variance in the rate that can be provided to it).

A (finite) sequence of users $\sigma = 1, ..., N$ appearing on the APOM's route to be provided with service is defined. In this simplified model, one user is observed per time slot, which corresponds to a problem horizon of N slots. Each user arrives randomly according to a probability distribution and is identified by a utility achieved by serving it (i.e. amount of transmitted data), and the energy consumption it requires. For each user n, APOM chooses whether to transmit to it or not based on that user's individual characteristics and the probability of future possible users and energy harvests. In the rest, each user will be classified by a "value" and "weight" pair: (v_i, w_i) for the i^{th} user, where the value corresponds to the utility gained by serving this user, and the *weight* corresponds to the power consumption required to serve it. The user characteristics are not available to the APOM in advance but become known in time of arrival, so the (v_i, w_i) information of the i^{th} user appears at the time it sends a request. The users that APOM encounters are classified as regards their utilization v_i value which is a random variable

A finite number of discrete states for energy requirement, instantaneous utility and energy harvesting will be assumed, which are not fully unrealistic considering some practical correspondents to these limitations exist: in practice, APs admit only users in a certain area of coverage, which automatically limits the power consumption. Rewards from a user (rate, pricing, etc.) are inherently bounded. Finally, solar irradiation is quite stationary on an hourly basis [49].

In the dynamic and stochastic version of the knapsack problem items arrive sequentially over time and their value/weight combinations are stochastic but become known to the scheduler at arrival times. An instantaneous decision needs to be made on each impatient user's arrival. The decisions are irreversible, that is, users can not be recalled later. Once a user is rejected, value associated with that user will be lost.

Following knapsack terminology, the AP is characterized by its "capacity" to serve, which corresponds in our case to the amount of energy it has stored in its battery. The problem is to collect the expected reward over N users while ensuring that the total weight does not exceed the service capacity. Stated this way, the problem is dynamic stochastic knapsack problem as in Section 2.2. However, in accordance with energy harvesting, capacity replenishment is allowed. At the beginning of the problem horizon, there is a certain amount of energy, assumed to be stored in the battery of the AP.

Indeed, it never makes sense to allocate an insufficient resource to a user, because individually rational user will provide zero utility is its resource requirement not fully be satisfied. On the other hand, allocating more energy capacity than the reported demand is useless as well since such allocations do not further increase utility gained more than the requested value.

Using this setup, the problem can be stated in terms of x_n's $x_n = \{0, 1\}$, which indicate the decision to either serve user n or pass it up. In terms of horizon over which the APOM will act; the problem statement approach can be classified into two parts: finite and infinite time horizons.

4.1.2 Problem Statement: Finite Horizon

The problem in finite horizon is to collect the expected reward (utility) over N slots while ensuring that the total weight (energy consumption) does not exceed the energy capacity considering the random user characteristics and energy harvests in a certain time horizon. If the problem were defined with conventional energy sources rather than renewable energy sources, the problem formulation given in Problem 3 is a generalization of dynamic stochastic knapsack problem where the objective in maximizing expected value.

Problem 3.

$$\textit{Maximize: } E(\sum_{n=1}^{N} v_n x_n) \tag{4.1}$$

$$\textit{subject to: } \sum_{n=1}^{N} w_n x_n \leq B, \tag{4.2}$$

$$x_n \in \{0, 1\} \tag{4.3}$$

Then, considering the energy replenishments, available energy is the newly harvested energy plus the energy left over from previous intervals. Energy expenditure in an interval cannot exceed this amount. The overall structure may be modelled as an extension of the stochastic knapsack problem with increasing capacity due to energy harvests prevailing at N_k intervals as stated in Problem 4. The energy arrivals B_k's are modelled as IID random processes to appear at the beginning of each time slot where $B_i \in 0, e$. Thus, energy is replenished with probability q at each interval such as right after the N_1^{th} slot, the N_2^{th} slot, and so on, up to some $N_k = N$.

Problem 4.

$$\textit{Maximize: } E(\sum_{n=1}^{N_k} v_n x_n) \tag{4.4}$$

$$\textit{subject to: } \sum_{n=1}^{N_1} w_n x_n \leq B_1,$$

$$\sum_{n=1}^{N_2} w_n x_n \leq B_1 + B_2, ..., \sum_{n=1}^{N_k} w_n x_n \leq \sum_{j=1}^{k} B_j \tag{4.5}$$

$$x_n \in \{0, 1\} \tag{4.6}$$

Equations (4.4) and (4.5) indicates the objective of the problem, which is to maximize the expected total value in a finite time horizon N under the constraints of energy consumption. (4.5) holds for energy causality, i.e., the energy expenditure in any interval can not be bigger than the energy production in that interval. In here, the setup of the problem differs from the most usual knapsack problems. The assumption that only one user can get service at each slot without preemption is given in (4.6).

Theorem 1. *The utility maximization problem in energy harvesting AP stated in Problem 4 is at least NP-hard.*

Proof. The claim can be shown by a reduction from a well known NP-hard problem: 0/1 knapsack problem where proposed utilities and energy consumption demand of each user corresponds to values and the weights of items in the knapsack. A special case of Problem 4 is equivalent to the 0/1 knapsack problem where there are no energy harvest or the energy harvest interval is infinitely long. The simpler version of the Problem 4 can be stated in a same way of Problem 3 which is known to be NP-hard [34]. Therefore Problem 4 is at least an NP-hard problem. □

Following Theorem 1 and motivated by the results obtained in similar works, dynamic programming method offers an optimal solution to the the dynamic problem in pseudo-polynomial time.

4.2 Optimal Online Solution with Dynamic Programming

Instantaneous requests are accepted by the APOM along its path as it encounters different users. A decision has to be made as the APOM encounters a new user. Using a threshold based approach as a decision mechanism is desirable in terms of computational complexity. Hence, we shall mainly develop a threshold based scheme after the optimal solution for the APOM scenario is derived via stochastic dynamic programming.

As stated in Section 4.1.2, there are K types of users based on the utility they provide such as $v_n(k) \in \{v_n(1), v_n(2)\}$ for a two user type system where users appear with the corresponding statistics $P(v_i(1))$ and $P(v_n(2))$ at each slot. The costs of the users to system $w_n(k)$'s are considered to be all equal to w without loss of all generality. If a user of type k appears at any slot n, then a value of $v(k)$ and weight of $w(k)$ are obtained. For the following analysis, a state for the n^{th} user is to be defined in terms of the available energy e and user type k. Each user class denoted by k appears as an IID process with probability $p(k)$ such that $\sum_{k=1}^{K} p(k) = 1$. $V(e, k, n)$ state denotes the expected total value from slot n till the end of time horizon N, that is, $V(e, k, n) = E(\sum_{i=n}^{N} v_i x_i)$.

At first, future harvests are not taken into account where $B_1 = E$ is the initial amount of energy which is also the available sack capacity. Let the action taken for n^{th} user be $d_n = \{T, D\}$ where T denotes $x_n = 1$ and D implies that $x_n = 0$. The action space for N slots is denoted with the following set: $d = [d_1, ..., d_N]$. The maximum expected total value at slot n till the deadline$V^*(e, k, n)$ can be formulated using the dynamic programming equation as follows:

$$V^*(e, k, n) = \max_{d} \{V_d(e, k, n)\}, n \geq 1, where \tag{4.7}$$

$$V_{d_n}(e, k, n) = v(k)x_n + \sum_{k'=1}^{K} p(k')V^*(e - w(k)x_n, k', n + 1) \tag{4.8}$$

A threshold based approach is revealed after the backward induction of the dynamic programming equation revealed that for smaller utility type users, call $v(1)$, the function adopts a conservative attitude at first and turns to a Greedy form as it gets closer to the end of the horizon, N. Whilst, the threshold function for a user type with higher utility, of value $v(2)$, always has a priority. The APOM attempts to serve user with higher utility as long as the energy constraints are satisfied, which implies that the residual energy of the APOM should be greater than the weight of the corresponding user.

As an extension to (4.7), the energy replenishment statistics are considered where energy arrivals may occur with probability q at each slot while incrementing the capacity by 0 or 1 since $B_n \in \{0, 1\}$. In the next, lets denote the harvest process with a random variable Q which is equal to 1 with probability q and $e = [1, ..., E]$ implies the available energy at any time slot. Then, we define π as the decision policy for a horizon that spans N slots. At each slot; a decision of Transmit or Deny is made according to the comparison of the instantaneous reward gained by serving the current user with the total expected reward by saving that energy for the next slots. A Stochastic Dynamic Programming Solution using 'divide and conquer' method is proposed. Action $d_n = \{T, D\}$ corresponding $x_n = \{1, 0\}$, an action space for N slots till the end of tie horizon is denoted with the following set: $d = [d_1, ..., d_N]$. Using Bellman's equation for the expected value maximization dynamic programming equation is constructed as:

$$V^*(e, k, n) = \max_d \{V_d(e, k, n)\} \tag{4.9}$$

$$d \in \{T, D\}^N \tag{4.10}$$

The problem formulation using a dynamic programming equation is constructed as follows: Start with

$$V^*(e, k, N) = v(k), \forall e \geq 1 \tag{4.11}$$

and go backwards using

$$V_{d_n}(e,k,n) = v(k)x_n + \mathbb{E}_{k',Q}\{V^*(e - w(k)x_n + Q, k', n+1)\} \tag{4.12}$$

$$V^*(e,k,n) = \max_{d_n}\{V_{d_n}(e,k,n)\} \tag{4.13}$$

$$\dots$$

$$V_{d_1}(e,k,1) = v(k)x_1 + \mathbb{E}_{k',Q}\{V^*(e - w(k)x_1 + Q, k', 2)\} \tag{4.14}$$

$$V^*(e,k,1) = \max_{d_1}\{V_{d_1}(e,k,1)\} \tag{4.15}$$

$$V^*(e,k) = V^*(e,k,1) \tag{4.16}$$

For a current user n, the expected value achieved by transmitting to that user is denoted as $V_T(e,k,n)$ whereas the expected value after denying that user is represented as $V_D(e,k,n)$. Comparing these quantities, the optimal expected value may be stated as:

$$V^*(e,k,n) = \max\{V_T(e,k,n), V_D(e,k,n)\} \tag{4.17}$$

When IID harvest probability (q) and K type users appearing with probability $p(k)$ are assumed without loss of all generality, the explicit form of the functions are given as:

(4.18)

$$V_T(e,k,n) = v(k) + (1-q)\sum_{k'=1}^{K} p(k')V^*(e - w(k), k', n+1)$$

$$+q\sum_{k'=1}^{K} p(k')V^*(e + 1 - w(k), k', n+1) \tag{4.19}$$

$$V_D(e,k,n) = (1-q)\sum_{k'=1}^{K} p(k')V^*(e, k', n+1)$$

$$+q\sum_{k'=1}^{K} p(k')V^*(e+1, k', n+1) \tag{4.20}$$

Using the results of the model stated in (4.17), it is intuitive to expect that an optimal solution turns out to be more greedy at the end while adopting a more conservative approach at the beginning.

To define DP relaxation of the problem explicitly, the pseudo code of solution is given as follows:

Algorithm 1 DP solution to the problem for finite horizon

---

for $e = 0 to E$ **do**
  for $k = 1 to K$ **do**
    $V(e, k, N + 1) = 0$ {Initialization step}
  end for
end for
for $n = N to 1$ **do**
  for $e = 0 to E$ **do**
    for $k = 1 to K$ **do**
      if $w(k) > e$ **then**
        $V(e, k, n) = E_{(k',Q)}\{V^*(e + Q, k', n + 1)\}$
      else
        $V(e, k, n) = \max\{\mathbb{E}_{(k',Q)}\{V^*(e+Q, k', n+1)\}, v(k) + \mathbb{E}_{(k',Q)}\{V^*(e - w(k) + Q, k', n + 1\}\}$ {Recursive equation}
      end if
      return $V(e, k, n)$
    end for
  end for
end for

Theorem 2. *The dynamic programming approach provides an optimal solution to the problem defined in Section 4.1.2, where the objective function is expected utility maximization with energy harvesting.*

Proof. As a principle of DP, the problem is divided into subproblems. Moreover, Bellman's optimality condition states that optimization of the future does not depend on the decisions in the past. Therefore, solutions in each instance is optimal for a subproblem since the decision ($\{D, T\}$) is made according to $V^*(e, k, n) = max\{V_T(e, k, n), V_D(e, k, n)\}$. Once the decisions in each sub-problem is optimal, Theorem can be proved by induction by following the principle of optimality defined by Bertsekas [8]. □

4.3 Structure of the Optimal Policy

After exhibiting the stochastic dynamic programming solution, structure of the optimal policy may be obtained based on the DP relaxation.

Given any finite time horizon N, intuitively, dynamic programming optimal solution behaves more conservatively at the beginning, on the other hand, as the end of horizon gets closer, it becomes a greedy policy. It is also intuitive to expect that an optimal policy becomes more lenient in deciding to serve a given user as the available energy increases and vice versa.

Theorem 3. *For the dynamic stochastic Knapsack problem with increasing capacity on a probabilistic model, the optimal online policy that maximizes the total expected reward is of the form:*

$$x_n(k) = \begin{cases} 1 & : v_n(k) \geq \beta \\ 0 & : v_n(k) < \beta \end{cases} \tag{4.21}$$

where

$$\beta \triangleq \mathbb{E}_{(k',Q)}\{V^*(e+Q, k', n+1)\} - \mathbb{E}_{(k',Q)}\{V^*(e - w_n(k) + Q, k', n+1)\} \tag{4.22}$$

Proof. A valid policy makes a decision to either serve or not to serve a user at each slot. Dynamic programming formulation implies that $\arg\max\{v_n(k)+\mathbb{E}_{(k',Q)}\{V^*(e-w_n(k)+Q, k', n+1)\}, \mathbb{E}_{(k',Q)}\{V^*(e+Q, k', n+1)\}$. $V(e-w_n, k, n+1)$ is the expected future cost when the user is transmitted whereas $V(e, k, n+1)$ denotes the expected future cost when the decision is 'Deny'. Considering Bellman's principle of optimality, the optimal policy makes a decision based on a comparison of the value that current users offering exchange of a capacity consumption with the expected value that can be obtained with this energy expenditure in the remaining slots to maximize the total expected utility. □

The result of Theorem 3 may be considered to be straightforward to show regard to DP relaxation. However, structural properties of the optimal solution and whether there exists a threshold based policy on the available energy depending on problem state need to be examined in detail.

4.3.1 Existence of Threshold

Lemma 1. *For a given k and n, the state defined as expected total reward* $V(e,k,n)$ *is super-modular in available energy and decision pair (e,d), that is* $V_T(e+1,k,n) + V_D(e,k,n) \geq V_D(e+1,k,n) + V_T(e,k,n)$ *where* $1 \leq n \leq N$ *and* $d \in \{T, D\}$.

Proof. Stated super-modularity corresponds the statement:

$$V_T(e+1,k,n) - V_D(e+1,k,n) \geq V_T(e,k,n) - V_D(e,k,n), 1 \leq n \leq N \quad (4.23)$$

which can be shown to hold by the following argument:

$$\begin{aligned} V_T(e,k,n) - V_D(e,k,n) = v(k) &+ \sum_{k'}^{K} qV(e,k',n+1) + (1-q)V(e-1,k',n+1) \\ &- \sum_{k'}^{K} qV(e+1,k',n+1) + (1-q)V(e,k',n+1) \end{aligned} \quad (4.24)$$

$$\begin{aligned} = v(k) &+ \sum_{k'=1}^{K} q(V(e,k',n+1) - V(e+1,k',n+1)) \\ &+(1-q)(V(e-1,k',n+1) - V(e,k',n+1)) \end{aligned} \quad (4.25)$$

and,

$$\begin{aligned} V_T(e+1,k,n) - V_D(e+1,k,n) = v(k) &+ \sum_{k'=1}^{K} q(V(e+1,k',n+1) - V(e+2,k',n+1)) \\ &+(1-q)(V(e,k',n+1) - V(e+1,k',n+1)) \end{aligned} \quad (4.26)$$

By subtracting (4.25) from (4.26), a sufficient condition for (4.23) to hold becomes:

$$V(e,k,n) - V(e-1,k,n) \geq V(e+1,k,n) - V(e,k,n) \quad \forall n \geq 1 \quad (4.27)$$

Then, the condition of (4.27) is proved by induction. First, the condition is satisfied when $n = 1$, that is both sides of the equation are equal to 0. Second, if it is true for some $n - 1$ then it also holds for n.

$$V(e, k, n) - V(e - 1, k, n) \geq V(e + 1, k, n) - V(e, k, n) \tag{4.28}$$

We will examine the three cases corresponding 3 energy states ($e + 1, e, e - 1$) and the three decisions ($d_1, d_2, d_3 \in \{T, D\}$) respectively.

$$V_{d1}(e + 1, k, n) - V_{d2}(e, k, n) - (V_{d2}(e, k, n) - V_{d3}(e - 1, k, n)) \leq 0 \tag{4.29}$$

$$\begin{aligned} V_{d1}(e + 1, k, n) - V_{d1}(e, k, n) + V_{d1}(e, k, n) - V_{d2}(e, k, n) \\ -(V_{d2}(e, k, n) - V_{d3}(e - 1, k, n)) - V_{d3}(e, k, n) + V_{d3}(e, k, n) \leq 0 \end{aligned} \tag{4.30}$$

By optimality of d_2 for energy state e, $V_{d1}(e, k, n) - V_{d2}(e, k, n)$ statement is already smaller and equal to 0. Same property holds for the $V_{d3}(e, k, n) - V_{d2}(e, k, n)$ statement. Therefore, we should only consider the remaining terms. For each possible cases of $[d_1, d_3] \in \{T, D\}^2$, the inequality in (4.30) is shown to be satisfied. For example, lets examine the case where $d_1 = T$ and $d_2 = T$:

$$\begin{aligned} V_T(e + 1, k, n) - V_T(e, k, n) - (V_T(e, k, n) - V_T(e - 1, k, n)) \\ = \sum_{k'=1}^{K} p(k)q(V(e + 1, k, n - 1) - V(e, k, n - 1) \\ -V(e, k, n - 1) + V(e - 1, k, n - 1)) \\ +(1 - q)(V(e, k, n - 1) - V(e - 1, k, n - 1) \\ -V(e - 1, k, n - 1) + V(e - 2, k, n - 1)) \leq 0 \end{aligned} \tag{4.31}$$

The above inequality holds since the difference is assumed to be non increasing in e. Similar steps may be followed for all combinations of d_1 and d_3 where $[d_1, d_3] \in \{T, D\}^2$. Hence, the total expected reward is a super-modular function in (d,e). □

Lemma 2. *Suppose decision is 'T' when state is* (e_1, k, n)*, then the decision is 'T' for the state* (e_2, k, n) *for all* $e_2 \geq e_1$ *such that* $\{e_1, e_2\} \in E^2$.

Proof. Following Lemma 1 namely the super-modularity property, it can be stated that:

$$V_T(e + 1, k, n) - V_D(e + 1, k, n) \geq V_T(e, k, n) - V_D(e, k, n) \tag{4.32}$$

when the optimal decision is $'T'$ for a state (e, k, n), by Bellman's equation:

$$V_T(e, k, n) \geq V_D(e, k, n) \tag{4.33}$$

then, combining (4.32) with (4.33), we have:

$$V_T(e + 1, k, n) \geq V_D(e + 1, k, n) \tag{4.34}$$

as a result, optimal solution is $'T'$ for the energy state $e + 1$ if it is $'T'$ for state e, that is, if $d(e) = T$ then $d(e + 1) = T$. □

Theorem 4. *The optimal policy is a threshold type policy in the available energy at each slot n and there is a threshold η defined as:* $x_n(k) = \begin{cases} 1 & : e \geq \eta(n, k) \\ 0 & : e < \eta(n, k) \end{cases}$

Proof. Let $\{e_1, e_2, e_3\}$ be the available energies at three decision instants such that $e_1 < e_2 < e_3$. For the sake of reaching contradiction, let's assume that the optimal policy is not a threshold based policy. Then, there exists an optimal policy which chooses to $'T'$ at energy levels e_1 and e_3 while denying the user at the energy level e_2. This contradicts with Lemma 2; hence, the assumed policy was not optimal. Therefore, the crossover from $'D'$ to $'T'$ happens only one as e is increased (holding all other parameters constant). This implies the existence of threshold. □

4.3.2 Monotonicity of Threshold

First, the following obvious result is recorded.

Proposition 1. *When users $k = 1, 2, ..., K$ with identical resource demands $w(k)$'s are given such that $v(1) \leq v(2) \leq ... \leq v(K)$, then $V(e, k, n) \leq V(e, k + 1, n)$.*

Proof. Since $V(e, k, n) = max\{v(k)+V(e-w(k), n+1), V(e, n+1)\}$ and $V(e, k-1, n) = max\{v(k - 1) + V(e - w(k - 1), k, n + 1), V(e, k, n + 1)\}$, the RHS of two equations differ only in the first terms inside the bracket. Considering $v(k) \leq v(k + 1)$, the statement in the proposition is shown to hold. □

Proposition 2. *$V(e, k)$ is a non-decreasing function of e.*

Proof. Proof follows by inspection of the value function in 4.2. □

Lemma 3. *Expected total reward '$V(e,k,n)'$ is super-modular in slot number and decision pair (n,d), that is* $V_T(e,k,n+1)+V_D(e,k,n) \geq V_D(e,k,n+1)+V_T(e,k,n)$.

Proof. Following the similar steps in the proof of lemma 1, supermodularity corresponds to the statement:

$$V_T(e,k,n+1) - V_D(e,k,n+1) \geq V_T(e,k,n) - V_D(e,k,n) \tag{4.35}$$

$$\begin{aligned} V_T(e,k,n+1) - V_D(e,k,n+1) = v(k)+ \\ \sum_{k'=1}^{K} q(V(e,k',n+2) - V(e+1,k',n+2)) \\ +(1-q)(V(e-1,k',n+2) - V(e,k',n+2)) \end{aligned} \tag{4.36}$$

$$\begin{aligned} V_T(e,k,n) - V_D(e,k,n) = v(k)+ \\ \sum_{k'=1}^{K} q(V(e,k',n+1) - V(e+1,k',n+1)) \\ +(1-q)(V(e-1,k',n+1) - V(e,k',n+1)) \end{aligned} \tag{4.37}$$

$$\tag{4.38}$$

by subtracting (4.36) from (4.37), the condition becomes:

$$V(e-1,k,n+1) - V(e,k,n+1) \leq V(e-1,k,n+2) - V(e,k,n+2) \tag{4.39}$$

which is equivalent to:

$$V(e,k,n+1) - V(e,k,n+2) \geq V(e-1,k,n+1) - V(e-1,k,n+2) \tag{4.40}$$

Expected total value is a non-increasing function of slot-number (both sides of the inequality have nonnegative values) and a non-decreasing function of available energy, therefore inequality in (4.40) holds for all e, k, n variables. □

Lemma 4. *Suppose decision is 'T' when state is (e,k,n), then the decision is 'T' for the state $(e,k,n+1)$ for all n such that $n \in \{1,...,N\}$.*

Proof. Lemma holds when the state defined as expected total reward '$V(e,k,n)'$ is super-modular in (n,x), that is $V_T(e,k,n+1) + V_D(e,k,n) \geq V_D(e,k,n+1) +$

$V_T(e,k,n)$. When super-modularity holds: $V_T(e,k,n+1) - V_D(e,k,n+1) \geq V_T(e,k,n) - V_D(e,k,n)$. Thus, considering the Bellman's decision formulation it is straightforward that the decision is $'T'$ for the state $(e,k,n+1)$ if it is $'T'$ when state is (e,k,n). □

Theorem 5. *The threshold function on the available energy to serve a user, $\eta(n,k)$ defined in Theorem 4 is monotonically non-increasing with slot number n. Moreover, the threshold function on the users' value, $\beta(e,n)$ defined in (4.22) is monotonically non-increasing with slot number n and monotonically non-decreasing with energy level e.*

Proof. The proof follows the Lemma 4. Theorem may further be proved by contradiction. Let $\{n, n+1\}$ be the available slot numbers such that $n \geq 1$. Let's assume that the optimal threshold value is increasing with slot number. Then, there exists an optimal policy which chooses to transmit at slot n while denying a user of type k at the slot $n+1$. Following the formulation given in Theorem 3 results of Lemmas 3 and 4, it is feasible but not a solution of the Bellman equation. Therefore a valid threshold function is monotonically non-increasing with slot number n. Monotonicity with respect to energy level e can be shown following the similar steps stated for slot number n. □

4.3.3 Curse of dimensionality

Dynamic programming method provides an optimal solution for 0/1 dynamic and stochastic resource allocation problem with randomly growing capacity; however as in most optimization problems it suffers from 'curse of dimensionality'. Indeed, the result is consistent with the NP-hardness definition such that there is no algorithm that works both optimal and fast.

Both computational and storage requirements grow exponentially as the number of state and control variables increases. The control variables in this problem have only two states $\{T, D\}$, but the state variables quickly explode with increasing number of possible user types (K), energy states (E) and length of time horizon (N). Its computational complexity is $O(2^{(NEK)})$ which might cause low performance and unnecessary delays in practical situations with high dimension. In the next, some methods and heuristics will be proposed to deal with this 'curse of dimensionality'.

4.4 Suboptimal Solution: Expected Threshold Policy

The curse of dimensionality may be overcomed by using pre-calculated tables that include optimal expected values and the decision might be done upon them. However, it still needs a high amount of memory by exchanging time complexity with space complexity. Moreover, in real life scenarios with dynamically changing environment, this is not feasible. To deal with the cases where real time computation is necessary; some suboptimal approaches with low complexity is needed. In this section, a computationally cost effective suboptimal policy called "Expected Threshold" is derived.

Considering the structure of optimal online policy, decision of transmitting to a user or not is made mostly upon the available energy on that time. In the computationally cost effective suboptimal solution approach, we set some energy boundaries on each user type that AP should have in order to serve that user. In the process of constituting an energy threshold, AP should also regard the potential energy harvests till the end and potential user characteristics.

Each user class denoted by k appears as an IID process with probability $p(k)$ such that $\sum_{k=1}^{K} p(k) = 1$. $V(e, k, n)$ state denotes the expected total value from slot n till the end of time horizon N. Considering the causality constraint on energy, the amount of depleted energy in any interval can not be bigger than the sum of available energy at the beginning of that interval plus the energy renewed, we define the following bound on the expectation of energy depletion (RHS of (4.41)) at slot n if the energy is e and expected harvest amount from slot n till the end of time horizon N if is denoted as $\sum_{m=n}^{N-1} \mathbb{E}\{Q_m|Q_1^{n-1}\}$.

$$e + \sum_{m=n}^{N-1} \mathbb{E}\{Q_m|Q_1^{n-1}\} \geq \sum_{m=n+1}^{N} \mathbb{E}\{w_m x_m\} \tag{4.41}$$

If the weights are all equal and harvest process is IID then, the inequality 4.41 becomes:

$$e + \sum_{m=n}^{(N-1)} q \geq (N - n + 1)w \tag{4.42}$$

After stating a bound of the expected energy consumption from slot n till the end of time horizon in (4.42), a computationally cost effective suboptimal policy called "Expected Threshold" is proposed as follows in (4.43):

$$x_n(k) = \begin{cases} 1 & : e \geq \eta \\ 0 & : e < \eta \end{cases} \tag{4.43}$$

where

$$\eta = (N - n + 1)(\sum_{k'=k+1}^{K} p'_k - q) \tag{4.44}$$

As it can be seen from (4.43) and (4.44, APOM makes a decision to serve a user of type k appearing in slot n if the available energy e at slot n is greater or equal to threshold level η. η is stated as the difference between the expected energy consumption for users with higher value and expected energy replenishment from slot n till the end of time horizon N.

It can be devised from (4.44) that threshold function follows the structure of the optimal policy defined in Sec. (4.3), i.e. monotonically decreases with slot number and harvest rate depending on the probability distribution of users.

Please note that, without loss of generality, a suboptimal threshold function can be derived for general distribution of harvest event as in (4.45) and (4.46) following the bound in (4.41).

$$x_n(k) = \begin{cases} 1 & : e \geq \widehat{\eta} \\ 0 & : e < \widehat{\eta} \end{cases} \tag{4.45}$$

where

$$\widehat{\eta} = \sum_{m=n+1}^{N} \mathbb{E}\{w_m x_m\} - \sum_{m=n}^{N-1} \mathbb{E}\{Q_m | Q_1^{n-1}\} \tag{4.46}$$

To examine the performance of the expected threshold policy in the next section (4.5), two more different heuristics are also defined as: greedy policy (Definition 1) and conservative policy (Definition 2).

Definition 1. *Greedy policy is a policy that decides to serve an appearing user whenever there is available energy to serve it.*

Definition 2. *Conservative policy is a policy that serves only the best user when there is available energy to serve it.*

4.5 Performance Evaluation and Numerical Results

It appears that an optimal policy becomes more lenient in accepting demands as the deadline approaches, the remaining amount of resource increases; in other words, the optimal reward threshold is decreasing in time for any given resource requirement and increasing in the expected amount of energy incoming.

The two observations given above have been proved in previous section. The sub-optimal policy is also shown to exhibit the same structural characteristics defined in the section. The statement can further be justified by considering the sub-optimal threshold function.

Adopting the suboptimal threshold policy defined in (4.43), the results are obtained and compared with the optimal policy over 10000 Monte Carlo Simulations as illustrated in Figures 4.2 and 4.3 for two different levels of initial energy stored in the battery and the two different energy harvest probabilities. It can be deduced from the figures that, initial stored energy effect on the performance is insignificant.

It is observed that both the optimal and suboptimal policies behaves more conservatively at the beginning, then as the end of horizon gets closer, it becomes a greedy policy given any finite time horizon N. As illustrated in Figures 4.4 and 4.5, the Expected Threshold Policy is shown to perform very close to optimal policy. However

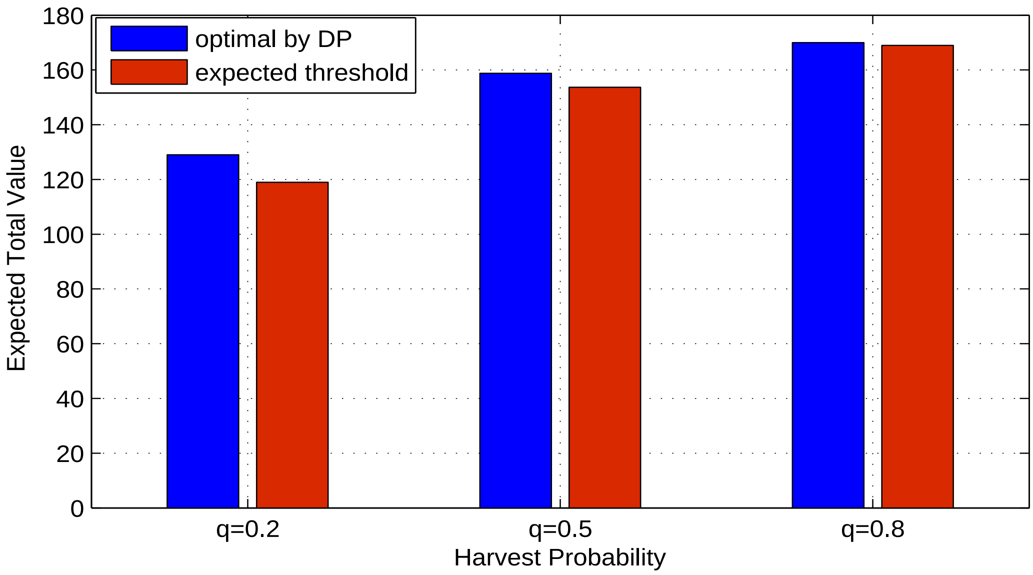

Figure 4.2: The performance evaluation of Expected Energy Threshold Heuristic wrt. Optimal Policy when available energy=10, $N = 20$, $K = 2$ for two different user types with efficiency ratios 10 and 5

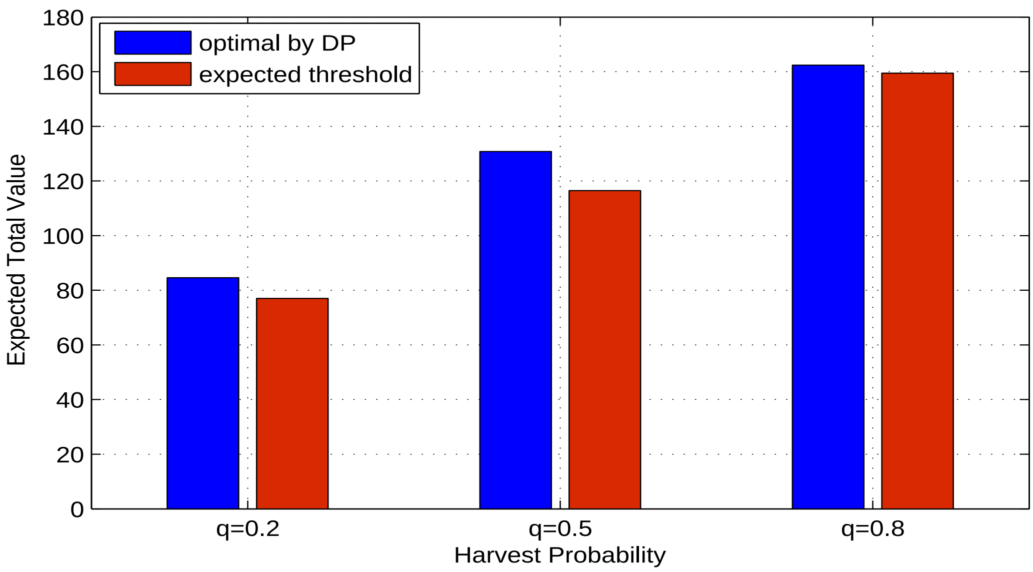

Figure 4.3: The performance evaluation of Expected Energy Threshold Heuristic wrt. Optimal Policy when available energy=5, $N = 20$, $K = 2$ for two different user types with efficiency ratios 10 and 5

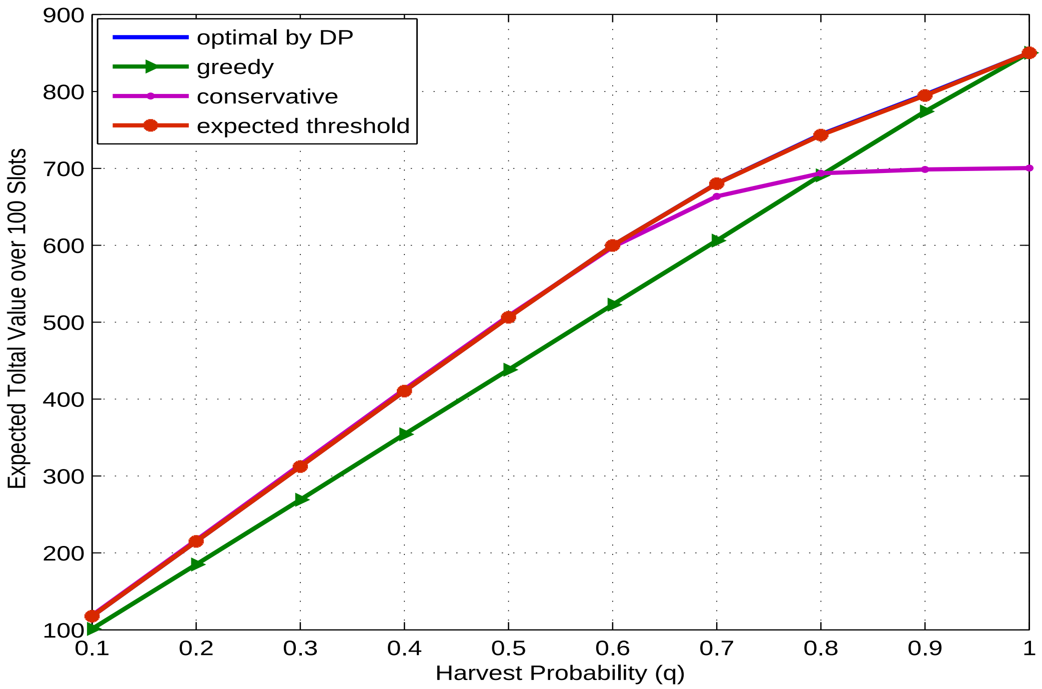

Figure 4.4: The comparison of the performances for Expected Energy Threshold Policy, Greedy Policy and Conservative Policy with respect to Optimal Policy when available energy=5, $N = 100$, $K = 2$ for two different user types with efficiency ratios 10 and 5 (best users appearing with high probability e.g. 0.7)

the greedy and conservative policies perform less efficiently compared with the proposed heuristic. It can also be derived from Figure 4.4 that when the efficient users appear with high probability, conservative policies outperforms greedy policies. On the other hand, when the inefficient users appear with low probability, greedy policies are more advantageous than the conservative approaches. However, Expected Threshold Policy proposed in this thesis is robust to variations in user distributions.

In addition, to investigate more general scenarios number of user types (K) is increased to 5 and simulations has been conducted on the users with both the equal weight (energy demand) in Figure 4.6 and different weights in Figure 4.7. In Figure 4.7, simulations have been conducted for five different user types with efficiency (value/weight) ratios given as [10/1, 5/1, 8/4, 5/8, 2/6 1/5] with probabilities [0.3 0.15 0.15 0.3 0.1]. As it can be seen from Figures 4.6 and 4.7, performance of the Expected Threshold Policy is very close to optimal.

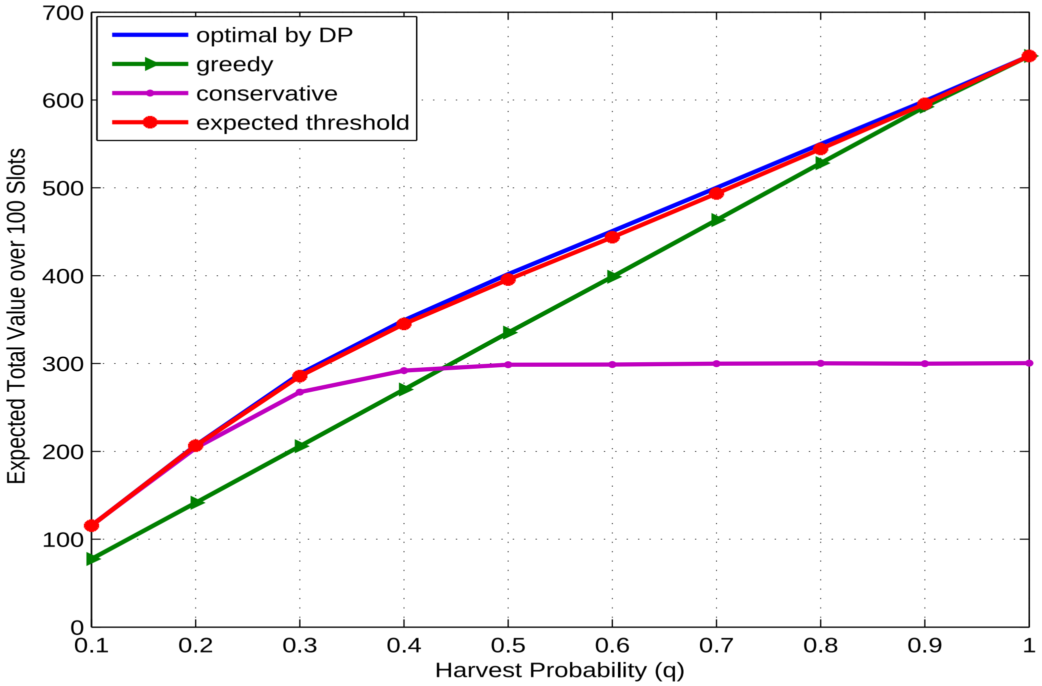

Figure 4.5: The comparison of the performances for Expected Energy Threshold Policy, Greedy Policy and Conservative Policy with respect to Optimal Policy when available energy=5, $N = 100$, $K = 2$ for two different user types with efficiency ratios 10 and 5 (worst users appearing with high probability e.g. 0.7)

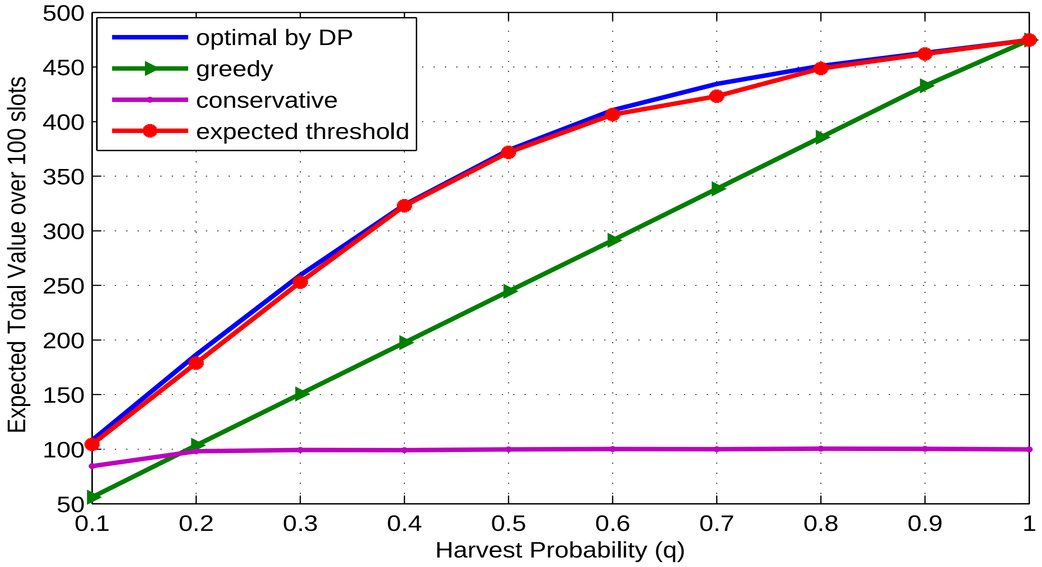

Figure 4.6: The comparison of the performances for Expected Energy Threshold Policy, Greedy Policy and Conservative Policy with respect to Optimal Policy when available energy=5 at the beginning, $N = 100$, $K = 5$ for five different user types with equal weights

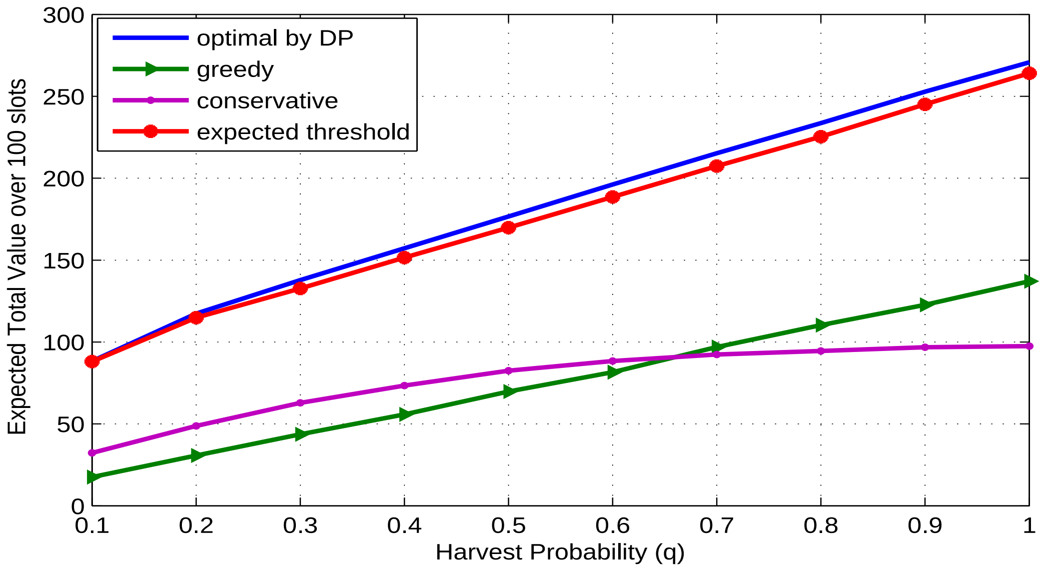

Figure 4.7: The comparison of the performances for Expected Energy Threshold Policy, Greedy Policy and Conservative Policy with respect to Optimal Policy when available energy=5 at the beginning, $N = 100$, $K = 5$ for five different user types with different efficiency (value/weight) ratios and different weights

CHAPTER 5

INFINITE HORIZON DYNAMIC EXPECTED UTILITY MAXIMIZATION UNDER ENERGY CONSTRAINTS

5.1 System Model and Problem Statement

When we would like to look infinitely far into the future, we need to maximize the total expected reward not only in a bounded time horizon, but also the expected total reward in an infinite time horizon. Please note that, an effective problem statement should be assumed to satisfy convergence in infinite horizon optimal policy. Two achieve this, one can consider the mean of expected reward infinite horizon or introduce a discount factor to the problem defined in Chapter 4. In this chapter, infinite horizon problem will be investigated with these two approaches.

5.1.1 System Model

The explicit model is given in the Figure 5.1 where a constant amount of energy (E_0) is replenished at random times (T_i). Each (T_n) interval with constant amount of capacity is called as an 'epoch'. The user distribution in all epochs is assumed to be same with rates $\lambda = \{\lambda_1, \lambda_2, ..., \lambda_K\}$. Again in the previous chapters, assume that each user associated with a instantaneous utility (v_n) and energy demand (w_n) is appearing sequentially and request service.

Following a widely adopted assumption about stochastic energy replenishment modelling [58], the harvest rate is assumed as a Poisson process with exponential inter-arrival times T_n. The goal is to maximize the expected utility while regarding the

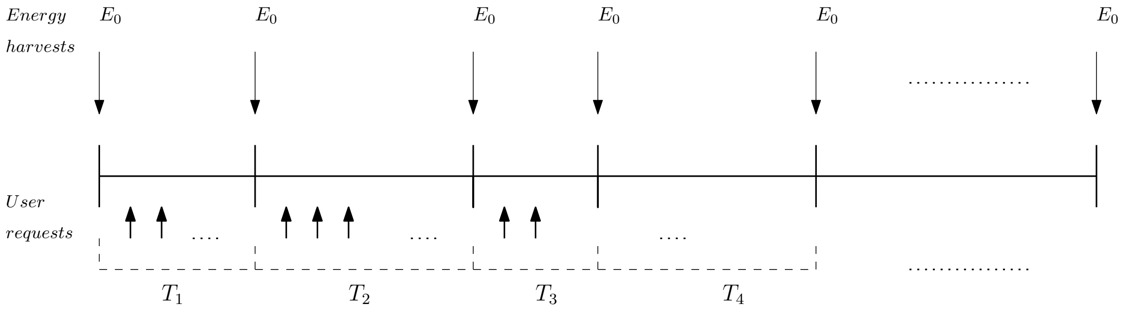

Figure 5.1: The overall system model for APOM with random energy harvesting intervals

energy arrival and user arrival statistics.

5.1.2 Problem Statement of Maximizing Mean Reward

The statement of the mean expected utility maximization problem over infinite epochs is stated in 5.

Problem 5.

$$\textit{Maximize:} \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} E\{V_n\} \tag{5.1}$$

$$\textit{subject to:} \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} E\{W_n\} \leq E_0 \tag{5.2}$$

where

$$V_n = \sum_{i=1}^{T_n} v_i x_i \tag{5.3}$$

$$W_n = \sum_{i=1}^{T_n} w_i x_i \tag{5.4}$$

$$\forall n \in 1, ..., N \quad as \quad N \to \infty \tag{5.5}$$

In (5.1), the mean of expected values in each epoch (n)is tried to be maximized where V_n stands for the sum of values obtained from selected items in that epoch. (5.1) represents the energy limitation based on the energy replenishment rate. W_n denotes the sum of energy consumption of selected users and (5.4) stands for the energy causality constraint between replenishments.

5.1.3 Problem Statement of Maximizing Discounted Expected Reward

When one would like to look infinitely far into the future, one will try to maximize the total expected reward not only in a bounded time horizon, but also the expected discounted reward in an infinite time horizon.

Here, we need to introduce a discount factor γ, to state the fundamental intuition in many areas such as economics, machine learning, operations research: "Money today is worth more than money tomorrow". If we do not have a discount factor for the future rewards, the obvious behaviour of the optimal algorithm in infinite horizon will be to wait for the user with highest value since we do not have any delay constraints here. Please note that, the infinite horizon optimal policy is stationary, i.e. optimal decision at a state s is the same decision at all times.

Problem 6.

$$\textit{Maximize: } E\{\lim_{N\to\infty}\sum_{n=1}^{N}\gamma^n v_n x_n\} \tag{5.6}$$

$$\textit{subject to: } \sum_{n=1}^{N_1} w_n x_n \le B_1, \tag{5.7}$$

$$\sum_{i=1}^{N_2} w_n x_n \le B_1 + B_2, ..., \lim_{N\to\infty}\sum_{i=1}^{N} w_n x_n \le \lim_{M\to\infty}\sum_{j=1}^{M} B_j \tag{5.8}$$

$$x_i \in \{0, 1\} \tag{5.9}$$

Despite the fact that the expectation is computed over an infinite sum, it is guaranteed to converge since value is defined over a finite state space, and $\gamma \in (0, 1)$.

5.2 Optimal Online Solution with Dynamic Programming

5.2.1 Markov Decision Process

In the previous sections, we have analysed the APOM resource allocation problem as a stochastic knapsack problem in infinite horizon. Next, we will propose a Markov Decision Model (MDP) to APOM resource allocation problem in infinite horizon with discounted reward.

To define a MDP, the problem statement may be written explicitly as follows: For a given set of states S and a set of decisions $'D'$, $'T'$ denotes the transition probabilities and R is the reward at each slot. State s is defined as the available energy e at the current slot. The reward function, R and the transition probability matrix, T depend on the harvest and user characteristics. T matrix represent the probability of next state s' if the current state is s and action d is taken. The instantaneous reward, R is equal to utility provided by that user if the decision is to serve and it is zero if the decision is not to serve. The main goal is to maximize the expected sum of discounted rewards.

$$V^*(s) = \max_d \sum_{s'} T(s, d, s')[R(s, d, s') + \gamma V^*(s')] \tag{5.10}$$

V^* is an optimal value function computed over the discounted infinite horizon which satisfies Bellman equations. The MDP may be solved with value or policy iteration policies. Iteration continues until the value function converges. Hence, following the assumption $\|V^*(e,k,n) - V^*(e,k,n+1)\| \leq \epsilon$ where $\epsilon \ll 1$, we may state $V^*(e,k) = V^*(e,k,n)$ for all n for infinite discounted problem statement.

The pseudo code based on DP method that summarizes the value iteration for the calculation of expected optimal value and the related policy are stated as the following.

Algorithm 2 Value iteration for infinite horizon

```
for each state (e,k) do
  while ||V*(e,k,n) − V*(e,k,n+1)|| ≤ ϵ do
    for each state (e,k) do
      if w(n) > e then
        V*(e,k) = γE_(k',Q){V*(e,k')}
      else
        V*(e,k) = max{γE_(k',Q){V*(e,k')}, v(k) + γE_(k',Q){V*(e − w(k), k')}}
      end if
    end for
  end while
end for
```

5.3 Suboptimal Online Policy

Again the optimal solution suffers from the "curse of dimensionality" growth of the state space and the fact that DP conducts a search over this space. The situation is consistent with the NP-hardness of knapsack problems.

5.3.1 Suboptimal Online Solution: Renewal Modelling

While trying to maximize the mean value over an infinite horizon, the optimal policy is stationary [8] if the underlying processes are stationary , i.e. the optimal action will be taken in any epoch should be equivalent.

The average rate of energy consumption up to any time t should be smaller or equal to average rate of replenishment. Lets denote the energy replenishment rate with P which is actually a random variable. The causality constraint may defined as follows:

$$\sum_{k=1}^{K} \mu_k \lambda_k w(k) \leq P \tag{5.11}$$

Assume that user types are arranged such that $v(1)/w(1) \leq v(2)/w(2) \leq ... \leq v(K-1)/w(K-1) \leq v(K)/w(K)$. When the acceptance rate for each user k is denoted as μ_k, the suboptimal policy defined as:

$$\mu_k = \max\{0, \min\{\frac{P - \sum_{k'=k+1}^{K} \mu_{k'} \lambda'_k w(k')}{\lambda_k w(k)}, 1\}\} \tag{5.12}$$

The time complexity of the randomized suboptimal policy is $O(K)$, since the acceptance rate calculation is repeated for all user types.

5.3.2 Suboptimal Online Solution: L-Level Lookahead Policy

Please note that, the infinite horizon optimal policy is stationary, i.e. optimal decision on a state s is the same decision at all times. Despite the fact that the expectation is computed over an infinite sum, it is guaranteed to converge since value is defined over a finite state space where $\gamma \in (0, 1)$. An approximate policy can be derived by ignoring the effects of slots after L^{th} slot in the future. The solution will compute L-level DP results by reducing computational complexity to $O(2^L)$ at each decision time.

The problem of maximizing discounted utility subject to energy constraints is introduced in Section 5.1.3.

$$V^* = max_\pi V^\pi \tag{5.13}$$

The discounted value obtained under policy π can be formulated as the following.

$$V^\pi = E\{\sum_{n=1}^{\infty} \gamma^n g_n(e_n, k)\} \tag{5.14}$$

where $g_n(e_n, k)$ is defined as the reward obtained at time n when available energy is e_n and k type users appears. (5.14) can be restated as:

$$V^\pi = E\{\sum_{n=1}^{L} \gamma^n g_n(e_n, k)\} + E\{\sum_{n=L+1}^{\infty} \gamma^n g_n(e_n, k)\} \tag{5.15}$$

Proposition 3. *The reward obtained at time n, $g_n(e_n, k)$, is upperbounded by the reward obtained from the user with best value, i.e, v_K.*

Proof. Since only one users can be served in a time slot n, the reward will be equal to values offered by a user admitted in that slot. The lemma holds because of the assumption in the finite set of users with finite set of values. □

Lemma 5. *The L-level Lookahead Policy, the optimal value is approximated by $\frac{\gamma^L v_K}{1-\gamma}$ depending on the discount factor and user characteristics.*

Proof. Following the Proposition 3, the discounted value of the slots after L is upperbounded as:

Table 5.1: Performance Evaluation of L-level Heuristic with Respect to Optimal Based on DP for Different Values of Discount Factor, γ

Methods	γ =0.1	γ =0.3	γ =0.5	γ =0.9
Optimal with DP	8.48	12.14	16.85	57.81
3-Level Heuristic	8.46	12.02	15.90	55..42

$$E\{\sum_{n=L+1}^{\infty} \gamma^n g_n(e_n, k)\} = E\{\sum_{n'=L}^{\infty} \gamma^{n'} g_{n'}(e_{n'}, k)\} \quad (5.16)$$

$$E\{\sum_{n'=L}^{\infty} \gamma^{n'} g_{n'}(e_{n'}, k)\} \leq \frac{\gamma^L v_K}{1-\gamma} \quad (5.17)$$

Therefore, in the L-level Lookahead policy, the optimal value is approximated by $\frac{\gamma^L v_K}{1-\gamma}$ depending on the discount factor and user characteristics. □

The performance will also be shown through simulations in Section 5.4.

5.4 Performance Evaluation and Numerical Results

To illustrate the performance of L-level heuristic in Section 5.3.2, a simulation has been conducted with 1000 Monte Carlo trials for 3-level heuristic case where $q = 0.3$, users of two type ($v_1 = 10$ and $v_2 = 5$) appear with probability (0.3 and 0.7) respectively. The average over obtained results are very close DP solution as can be seen from Table 5.1.

The results are also consistent with the approximation factor found in Section 5.3.2, that is $\frac{\gamma^L v_K}{1-\gamma}$.

CHAPTER 6

CONCLUSION

Throughout this thesis, the problem of dynamic user admission for an Access Point on the Move (APOM) scenario is addressed in a stochastic manner. At first, a detailed background on the resource allocation problems in energy harvesting wireless networks and on knapsack problems have been provided. Observations and performance results related to the possible approaches to related problems have been examined. Then, exploiting the inherent properties of renewable solar energy, a Kalman filter based solar prediction algorithm as novel method to adopt a valid offline resource allocation policy to online situations have been investigated.

Then, a finite-horizon dynamic utility-maximizing allocation problem with a discrete set of user requests has been formulated as a dynamic and stochastic 0/1 knapsack problem with randomized incremental capacity. The structure of the optimal solution of this problem has been studied through stochastic dynamic programming relaxation. It is proved that a threshold based policy is optimal for the problem and threshold is monotonic with respect to both available energy and slot number, where each user is admitted if its utility to weight ratio exceeds a certain threshold. Moreover, an efficient suboptimal policy to exhibit the structural properties of the optimal policy with low computational cost has been proposed. Experimental results demonstrate that the proposed decision methods using different threshold policies for the stochastic resource allocation problem of the energy harvesting Access Point on the Move are efficient in achieving close to optimal performances.

Likewise, the problem of expected discounted utility maximization problem has been formulated in infinite horizon. The problem is formulated as a Markov decision problem and solved with value iteration method. Stationary of the infinite horizon policy has been shown to hold following the idea of convergence in discounted reward. Again, a suboptimal policy with low computational complexity has been proposed and performance results have been analysed.

The APOM defined in this thesis is powered by renewable energy sources that have the capability of energy harvesting in a stochastic manner, which corresponds to incremental capacity. Although the results have been obtained for the dynamic user admission problem, they are also applicable to other instances of dynamic and stochastic knapsack problems with randomized incremental capacity. The insights gained in this study can be a basis for a more generic consideration of dynamic knapsack problems with incremental capacity that have a lot of real life applications in operational research, economics, computer science, satellite and ad-hoc networks etc.

REFERENCES

[1] Loon for all, May 2014.

[2] M. Ajmone Marsan and M. Meo. Green wireless networking: Three questions. *2011 The 10th IFIP Annual Mediterranean Ad Hoc Networking Workshop*, pages 41–44, June 2011.

[3] B. Akgün. Duty cycle optimization in energy harvesting sensor networks with application to bluetooth low energy. Master's thesis, METU, June 2014.

[4] A. Alkesh, A. Singh, and N. Purohit. A moving base station strategy using fuzzy logic for lifetime enhancement in wireless sensor network. In *Communication Systems and Network Technologies (CSNT), 2011 International Conference on*, pages 198–202, June 2011.

[5] K. Amaruchkul, W. L. Cooper, and D. Gupta. Single-leg air-cargo revenue management. *Transportation Science*, 41(4):457–469, 2007.

[6] B. T. Bacinoglu and E. Uysal-Biyikoglu. Finite horizon online lazy scheduling with energy harvesting transmitters over fading channels. *CoRR*, abs/1312.4798, 2013.

[7] B. T. Bacinoglu and E. Uysal-Biyikoglu. Finite-horizon online transmission rate and power adaptation on a communication link with markovian energy harvesting. *CoRR*, abs/1305.4558, 2013.

[8] D. P. Bertsekas. *Dynamic Programming and Optimal Control, Vol. II, 2nd Ed.* Athena Scientific, Belmont, MA, 2001.

[9] R. Bogue. Solar-powered sensors: a review of products and applications. *Sensor Review*, 32(2):95–100, 2012.

[10] R. L. Carraway, R. L. Schmidt, and L. R. Weatherford. An algorithm for maximizing target achievement in the stochastic knapsack problem with normal returns. *Naval Research Logistics (NRL)*, 40(2):161–173, 1993.

[11] E. T. Ceran, T. Erkilic, E. Uysal-Biyikoglu, T. Girici, and K. Leblebicioglu. Wireless access point on the move: Dynamic knapsack with incremental capacity. In *Globecom 2014 - Symposium on Selected Areas in Communications: GC14 SAC Green Communication Systems and Networks (GC14 SAC Green Communication Systems and Networks)*, Austin, USA, Dec. 2014. submitted.

[12] J. L. Crassidis and J. L. Junkins. *Optimal Estimation of Dynamic Systems*. Chapman and Hall, Abingdon, 2004.

[13] B. C. Dean, M. X. Goemans, and J. Vondrák. Approximating the stochastic knapsack problem: The benefit of adaptivity. *Mathematics of Operations Research*, 33(4):945–964, 2008.

[14] B. Devillers and D. Gunduz. Energy harvesting communication system with battery constraint and leakage. In *GLOBECOM Workshops (GC Wkshps), 2011 IEEE*, pages 383–388, Dec 2011.

[15] T. Erkılıç. Optimizing the service policy of a mobile service provider through competitive online solutions to the 0/1 knapsack problem with dynamic capacity. Master's thesis, METU, June 2014.

[16] N. T. Ersoy. *Efficient Resource Allocation in Energy Harvesting Wireless Networks*. PhD thesis, METU, 2013.

[17] S. Gao, H. Zhang, and S. K. Das. Efficient data collection in wireless sensor networks with path-constrained mobile sinks. *IEEE Transactions on Mobile Computing*, 10(4):592–608, Apr. 2011.

[18] M. R. Garey and D. S. Johnson. " strong " np-completeness results: Motivation, examples, and implications. *J. ACM*, 25(3):499–508, July 1978.

[19] B. T. B. E. U. B. Goksel Uctu, Omer Melih Gul. Implementation of energy efficient scheduling policies on software defined radio. In *Globecom 2014 - Wireless Networking Symposium (GC14 Wireless Networking Symposium)*, Austin, USA, Dec. 2014. submitted.

[20] O. M. Gul and E. U. Biyikoglu. Achieving nearly 100% throughput without feedback in energy harvesting wireless networks. In *ISIT 2014 – International Symposium on Information Theory*, Hawaii, USA, 2014. accepted.

[21] O. M. Gul and E. U. Biyikoglu. A randomized scheduling algorithm for energy harvesting wireless sensor networks achieving nearly 100% throughput. In *WCNC 2014 - Wireless Communication and Networking Conference*, Istanbul, Turkey, Apr. 2014. pp. 2492-2497.

[22] S. Guo and Y. Yang. A distributed optimal framework for mobile data gathering with concurrent data uploading in wireless sensor networks. In *INFOCOM, 2012 Proceedings IEEE*, pages 1305–1313, March 2012.

[23] X. Han, Y. Kawase, and K. Makino. Randomized algorithms for removable online knapsack problems. In M. Fellows, X. Tan, and B. Zhu, editors, *Frontiers in Algorithmics and Algorithmic Aspects in Information and Management*, volume 7924 of *Lecture Notes in Computer Science*, pages 60–71. Springer Berlin Heidelberg, 2013.

[24] C. Harress. Facebook (fb) to buy solar-powered drone company that can beam internet from the sky, 2014.

[25] M. Kashef and A. Ephremides. Optimal packet scheduling for energy harvesting sources on time varying wireless channels. *Communications and Networks, Journal of*, 14(2):121–129, April 2012.

[26] A. J. Kleywegt and J. D. Papastavrou. The dynamic and stochastic knapsack problem. *Oper. Res.*, 46(1):17–35, Jan. 1998.

[27] A. J. Kleywegt and J. D. Papastavrou. The dynamic and stochastic knapsack problem with random sized items. *Oper. Res.*, 49(1):26–41, Jan. 2001.

[28] C. E. Koksal and N. B. Shroff. Near Optimal Power and Rate Control of Multi-Hop Sensor Networks With Energy Replenishment: Basic Limitations With Finite Energy and Data Storage. *IEEE Transactions on Automatic Control*, 57(4):815–829, Apr. 2012.

[29] E. Lawler. Fast approximation algorithms for knapsack problems. In *Foundations of Computer Science, 1977., 18th Annual Symposium on*, pages 206–213, Oct 1977.

[30] W. Liang, J. Luo, and X. Xu. Prolonging network lifetime via a controlled mobile sink in wireless sensor networks. In *Global Telecommunications Conference (GLOBECOM 2010), 2010 IEEE*, pages 1–6, Dec 2010.

[31] W. Liang and X. Ren. The use of a mobile sink for quality data collection in energy harvesting sensor networks. In *Wireless Communications and Networking Conference (WCNC), 2013 IEEE*, pages 1145–1150, April 2013.

[32] Z. Mao, C. E. Koksal, and N. B. Shroff. Resource allocation in sensor networks with renewable energy. In *IEEE ICCCN*, pages 1–6, 2010.

[33] E. Mark. "meet google's 'project loon:' balloon-powered 'net access", June 2013. CNET.

[34] S. Martello and P. Toth. *Knapsack Problems: Algorithms and Computer Implementations*. John Wiley & Sons, Inc., New York, NY, USA, 1990.

[35] Ömer Melih Gül. A low-complexity, near-optimal scheduling policy for solving a restless multi-armed bandit problem occurring in a single-hop wireless network. Master's thesis, METU, June 2014.

[36] C. Metz. " facebook will build drones and satellites to beam internet around the world", 2014.

[37] M. I. Mohamed, Y. W. Wu, and M. Moniri. Power harvesting for smart sensor networks in monitoring water distribution system. In *IEEE International Conference on Networking, Sensing and Control (ICNSC)*, pages 393–398, 2011.

[38] D. K. Noh and K. Kang. A practical flow control scheme considering optimal energy allocation in solar-powered wsns. In *Proceedings of 18th Internatonal Conference on Computer Communications and Networks (ICCCN)*, pages 1–6, 2009.

[39] M. Paavola and K. Leiviska. Wireless sensor networks in industrial automation. In *Factory Automation*. InTech, 2010.

[40] K. Pak and R. Dekker. Cargo revenue management: Bid-prices for a 0-1 multi knapsack problem. Technical Report ERS-2004-055-LIS, E, Aug. 2004.

[41] C. Park and P. Chou. Ambimax: Autonomous energy harvesting platform for multi-supply wireless sensor nodes. In *Proceedings of the Sensor and Ad Hoc Communications and Networks (SECON)*, pages 168–177, 2006.

[42] J.-H. Park. Wireless internet access of the visited mobile isp subscriber on gprs/umts network. *Consumer Electronics, IEEE Transactions on*, 49(1):100–106, Feb 2003.

[43] X. Ren and W. Liang. Delay-tolerant data gathering in energy harvesting sensor networks with a mobile sink. In *Global Communications Conference (GLOBECOM), 2012 IEEE*, pages 93–99, Dec 2012.

[44] X. Ren, W. Liang, and W. Xu. Use of a mobile sink for maximizing data collection in energy harvesting sensor networks. In *Parallel Processing (ICPP), 2013 42nd International Conference on*, pages 439–448, Oct 2013.

[45] S. C. O. Z. E. U.-B. S. Baghaee, H. Ulusan and H. Kulah. Demonstration of energy-neutral operation on a wsn testbed using vibration energy harvesting. In *European Wireless 2014 (EW2014)*, Barcelona, Spain, May 2014.

[46] M. Shakiba-Herfeh. Optimization of feedback in a multiuser miso communication downlink with energy harvesting users. Master's thesis, METU, June 2014.

[47] M. Shakiba-Herfeh and E. Uysal-Biyikoglu. Optimization of feedback in a miso downlink with energy harvesting users. In *The 20th European Wireless 2014 , spain*, May 2014.

[48] N. Sharma, J. Gummeson, D. Irwin, and P. J. Shenoy. Cloudy computing: Leveraging weather forecasts in energy harvesting sensor systems. In *SECON'10*, pages 136–144, 2010.

[49] V. Sharma, U. Mukherji, and V. Joseph. Efficient Energy Management Policies for Networks with Energy Harvesting Sensor Nodes.

[50] Y. Shi and Y. Hou. Theoretical results on base station movement problem for sensor network. In *INFOCOM 2008. The 27th Conference on Computer Communications. IEEE*, pages –, April 2008.

[51] F. Simjee and P. H. Chou. Everlast: long-life, supercapacitor-operated wireless sensor node. In *International symposium on Low power electronics and design (ISLPED)*, pages 197–202, 2006.

[52] N. Tekbiyik, E. T. Ceran, K. Leblebicioglu, T. Girici, and E. U. Biyikoglu. Prediction based proportional fair resource allocation for industrial wireless sensor networks. *IEEE Trans. Industrial Informatics*, 2014. submitted.

[53] N. Tekbiyik, T. Girici, E. Uysal-Biyikoglu, and K. Leblebicioglu. Proportional fair resource allocation on an energy harvesting downlink. *IEEE Transactions on Wireless Communications*, 12(4):1699–1711, 2013.

[54] G. Uçtu. Optimal transmission scheduling for energy harvesting systems and implementation of energy efficient scheduling algorithms on software defined radio. Master's thesis, METU, June 2014.

[55] L. Xie, Y. Shi, Y. Hou, W. Lou, H. Sherali, and S. Midkiff. Bundling mobile base station and wireless energy transfer: Modeling and optimization. In *INFOCOM, 2013 Proceedings IEEE*, pages 1636–1644, April 2013.

[56] G. Xing, M. Li, T. Wang, W. Jia, and J. Huang. Efficient rendezvous algorithms for mobility-enabled wireless sensor networks. *Mobile Computing, IEEE Transactions on*, 11(1):47–60, Jan 2012.

[57] K. T. Yen and S. K. Panda. Energy harvesting for autonomous wireless sensor networks. *IEEE Transactions on Power Electronics*, 26(1):38–50, 2011.

[58] S. Zhang. *Modeling, Analysis and Design of Energy Harvesting Communication Systems*. PhD thesis, University of Rochester, Rochester, New Yorkk, 2013.

[59] S. Zhang and A. Seyedi. Harvesting resource allocation in energy harvesting wireless sensor networks. *CoRR*, abs/1306.4997, 2013.

[60] Y. Zhou, D. Chakrabarty, and R. Lukose. Budget constrained bidding in keyword auctions and online knapsack problems. In C. Papadimitriou and S. Zhang, editors, *Internet and Network Economics*, volume 5385 of *Lecture Notes in Computer Science*, pages 566–576. Springer Berlin Heidelberg, 2008.