ABSTRACT

Title of dissertation: EFFICIENT MEDIA ACCESS CONTROL AND DISTRIBUTED CHANNEL-AWARE SCHEDULING FOR WIRELESS AD-HOC NETWORKS

Hua Chen
Doctor of Philosophy, 2013

Dissertation directed by: Professor John S. Baras
Department of Electrical and Computer Engineering

We address the problem of channel-aware scheduling for wireless ad-hoc networks, where the channel state information (CSI) are utilized to improve the overall system performance instead of the individual link performance. In our framework, multiple links cooperate to schedule data transmission in a decentralized and opportunistic manner, where channel probing is adopted to resolve collisions in the wireless medium.

In the first part of the dissertation, we study this problem under the assumption that we know the channel statistics but not the instant CSI. In this problem, channel probing is followed by a transmission scheduling procedure executed independently within each link in the network. We study this problem for the popular block-fading channel model, where channel dependencies are inevitable between different time instances during the channel probing phase. We use optimal stopping theory to formulate this problem, but at carefully chosen time instances at which

effective decisions are made. The problem can then be solved by a new stopping rule problem where the observations are independent between different time instances. We first characterize the system performance assuming the stopping rule problem has infinite stages. We then develop a measure to check how well the problem can be analyzed as an infinite horizon problem, and characterize the achievable system performance if we ignore the finite horizon constraint and design stopping rules based on the infinite horizon analysis. We then analyze the problem using backward induction when the finite horizon constraint cannot be ignored. We develop one recursive approach to solve the problem and show that the computational complexity is linear with respect to network size. We present an improved protocol to reduce the probing costs which requires no additional cost.

Based on our analysis on single-channel networks, we extend the problem to adhoc networks where the wireless spectrum can be divided into multiple independent sub-channels for better efficiency. We start with a naive multi-channel protocol where the scheduling scheme is working independently within each sub-channel. We show that the naive protocol can only marginally improve the system performance. We then develop a protocol to jointly consider the opportunistic scheduling behavior across multiple sub-channels. We characterize the optimal stopping rule and present several bounds for the network throughputs of the multi-channel protocol. We show that by joint optimization of the scheduling scheme across multiple sub-channels, the proposed protocol improves the system performance considerably in contrast to that of single-channel systems.

In the second part of the dissertation, we study this problem under the assumption that neither the instant CSI nor the channel statistics are known. We formulate the channel-aware scheduling problem using multi-armed bandit (MAB). We first present a semi-distributed MAB protocol which serves as the baseline for performance comparison. We then propose two forms of distributed MAB protocols, where each link keeps a local copy of the observations and plays the MAB game independently. In Protocol I the MAB game is only played once within each block, while in Protocol II it can be played multiple times. We show that the proposed distributed protocols can be considered as a generalized MAB procedure and each link is able to update its local copy of the observations for infinitely many times. We analyze the evolution of the local observations and the regrets of the system. For Protocol I, we show by simulation results that the local observations that are held independently at each link converge to the true parameters and the regret is comparable to that of the semi-distributed protocol. For Protocol II, we prove the convergence of the local observations and show an upper bound of the regret.

EFFICIENT MEDIA ACCESS CONTROL AND DISTRIBUTED CHANNEL-AWARE SCHEDULING FOR WIRELESS AD-HOC NETWORKS

by

Hua Chen

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2013

Advisory Committee:
Professor John S. Baras, Chair/Advisor
Professor Richard J. La
Professor P.S. Krishnaprasad
Professor Sennur Ulukus
Professor S. Raghavan, Dean's Representative

Dedication

To my family.

Acknowledgments

I am grateful to all the people who have made this dissertation possible. First and foremost, I woule like to thank my advisor Prof. John S. Baras for his vision and his continues support, encouragement and mentorship during my graduate studies at the University of Maryland. His immense energy, enthusiasm, persistence and keen insight in various topics has always been an excellent source of motivation for me. I had a unique chance to work on various problems across the control and wireless networking area, thanks to the fact he develops an open research environment where a student is encouraged to explore different research areas and problems. This involvement helped me build a strong interdisciplinary background. Moreover, I also appreciate the constant attention, support and encouragement from Prof. Baras to his students. I am also grateful to other committee members, Profs. Richard J. La, P.S. Krishnaprasad, Sennur Ulukus and S. Raghavan for agreeing to serve on my committee.

As a graduate student I have to deal with administrative issues from time to time. I woule like to thank Kimberly Edwards for her efficiency in keeping all these details to a minimum level. I appreciate all her helps with conferences, project reviews, travel arrangements and lab purchasing. I would also like to thank the staff of both ECE and ISR for doing their best to help students with official matters.

I would also like to thank several colleagues from Prof. Baras' research group. A special thanks goes to Pedram Hovareshti for his close collaboration and very useful suggestions with my research work in my third and fourth year. I would also like to thank Vladimir Ivanov for his efficient yet amusing way of assistant whenever I need help with computers, lab equipments and so on. Last but not least, I would like to thank a number of colleagues who made the long hours in the office not boring at all: Vladimir Ivanov, Ion Matei, George Papageorgiou, Pedram Hovareshti, Senni Perumal, Kiran Somasundaram, Kaustubh Jain, Baobing (Brian) Wang, Shalabh Jain.

My graduate studies and my research were supported by the Department of Electrical and Computer Engineering at University of Maryland, College Park, by the U.S. Army Research Laboratory (ARL) award number W911NF-08-1-0238 to Ohio State University, by BAE Systems award number W911NF-08-2-0004 under the U.S. ARL Collaborative Technology Alliance (CTA) on Micro Autonomous Systems and Technology (MAST), by the U.S. Air Force Office of Scientific Research (AFOSR) under MURI grant award FA9550-09-1-0538 to Georgia Tech, by the Defence Advanced Research Projects Agency (DARPA) award number SA00007007 to the University of California - Berkeley and by the National Science Foundation (NSF) award number CNS-1018346 to the University of Maryland, College Park.

Table of Contents

List of Figures vii

List of Abbreviations ix

1 Introduction 1
1.1 Background 1
1.2 Opportunistic Scheduling 3
1.3 Contribution of the Dissertation 12

2 Distributed Opportunistic Communications for Collaborative Control Application 15
2.1 Introduction 15
2.2 System Model 17
2.3 The Collaborative Control Algorithm for Networked Vehicles 19
2.3.1 Potential Functions 20
2.3.2 Distributed Control Algorithm with Local Information 22
2.3.3 Distributed Control Algorithm with Collaborative Information 24
2.4 The Distributed Opportunistic Communication Algorithm for Networked Vehicles 25
2.4.1 Wireless Inter-Vehicle Networks 25
2.4.2 The Distributed Opportunistic Communication Algorithm . . 27
2.5 Simulation Results 31

3 Distributed Opportunistic Scheduling for Single-Channel Networks 39
3.1 Introduction 39
3.2 System Model and Motivation 43
3.3 The Optimal Stopping Problem Formulation 45
3.4 A Rigorous Performance Analysis: the Finite Horizon Approach . . . 49
3.5 An Approximation for Performance Analysis: the Infinite Horizon Approach 55
3.6 An Energy Efficient Improvement of the Protocol 61
3.7 The Constant Data Time Problem 68
3.8 Numerical Results 72

4 Distributed Opportunistic Scheduling for Multi-Channel Networks 80
4.1 Introduction 80
4.2 Motivation for the Multi-Channel Problem 83
4.3 The Multi-Channel Opportunistic Scheduling Algorithm 85
4.3.1 Protocol Description 85
4.3.2 Performance Analysis 87
4.4 Numerical Results 94

5 A Multi-Armed Bandit Approach for Distributed Channel-Aware Scheduling 97
5.1 Introduction . . . 97
5.2 Problem Formulation . . . 99
5.3 A Semi-Distributed MAB Protocol for Opportunistic Scheduling . . . 101
5.4 Distributed MAB Protocols for Channel-Aware Scheduling . . . 104
5.4.1 Distributed MAB Protocol A . . . 106
5.4.2 Distributed MAB Protocol B . . . 111
5.5 Distributed MAB Protocol II for Channel-Aware Scheduling . . . 118
5.5.1 Protocol Description . . . 118
5.5.2 Performance Analysis: $\epsilon = 0$ . . . 122
5.5.3 Performance Analysis: $\epsilon > 0$ . . . 124
5.6 Numerical Results . . . 133

6 Conclusions 141

A Optimal Stopping Theory 144
A.1 Finite Horizon Problem . . . 147
A.2 Infinite Horizon Problem . . . 147

B Proof of Theorems in Chapter 3 149
B.1 Proof of Theorem 3.7.2 . . . 149
B.2 Proof of Theorem 3.7.4 . . . 150

Bibliography 152

List of Figures

2.1 Neighboring potential function 22
2.2 The distributed control algorithm with only local information 23
2.3 The distributed control algorithm with collaborative information . . . 25
2.4 The opportunistic communication algorithm for networked vehicles . 30
2.5 Performance comparing of the non-opportunistic and opportunistic communication algorithms, assuming that the vehicles only exchange position information. (a) Trajectories of the vehicles for non-opportunistic algorithm; (b) Trajectories of the vehicles for opportunistic algorithm; (c) Average performance of the vehicles for non-opportunistic algorithm; (d) Average performance of the vehicles for opportunistic algorithm; (e) Performance of vehicle 1 for non-opportunistic algorithm; (f) Performance of vehicle 1 for opportunistic algorithm. 37
2.6 Performance comparing of the non-opportunistic and opportunistic communication algorithms, assuming that the vehicles exchange both position information and additional data traffic. (a) Trajectories of the vehicles for non-opportunistic algorithm; (b) Trajectories of the vehicles for opportunistic algorithm; (c) Average performance of the vehicles for non-opportunistic algorithm; (d) Average performance of the vehicles for opportunistic algorithm; (e) Performance of vehicle 1 for non-opportunistic algorithm; (f) Performance of vehicle 1 for opportunistic algorithm. 38

3.1 The distributed opportunistic scheduling protocol 47
3.2 The improved distributed opportunistic scheduling protocol 63
3.3 Numerical results for ad-hoc networks with M links, where the parameters are $p = 1/M$, $\tau = 0.01$, $\rho = -10$dB and $\sigma = 1$: (a) network throughputs; (b) $P[N^* > M]$; (c) energy savings in probing signals. . 73
3.4 Numerical results for ad-hoc networks with M links, where the default parameters are $p = 1/M$, $\tau = 0.01$, $\rho = -10$dB and $\sigma = 1$: (a) network throughputs with $p = 0.01$; (b) network throughputs with $p = 0.1$; (c) network throughputs with $\tau = 0.05$; (d) network throughputs with $\rho = 10$dB. 76
3.5 Numerical results for the CDT problem for ad-hoc networks with M links, where the parameters are $p = 1/M$, $\rho = -10$dB and $\sigma = 1$: (a) network throughputs with $\tau = 0.01$; (b) network throughputs with $\tau = 0.05$. 78

4.1 The channel probing within one block duration T for a network with 3 sub-channels. 86
4.2 The Distributed Opportunistic Scheduling Protocol for Multi-Channel Networks 88
4.3 System throughput with varying number of sub-channels in the network, where $\tau = 0.02$, $\rho = -10$dB and $\sigma = 1$. 95

4.4 System throughput with varying number of sub-channels in the network, where $\tau = 0.02$, $\rho = -10$dB and $\sigma = 1$. 96

5.1 CAT based system model 100
5.2 The Semi-Distributed MAB Protocol 103
5.3 The distributed MAB protocol A 107
5.4 The distributed MAB protocol B 113
5.5 The initialization of Protocol B in Figure 5.4 with: (a) homogeneous observations, (b) heterogeneous observations. 114
5.6 A sketch for the proof of Theorem 5.4.4. 116
5.7 The distributed MAB protocol II 121
5.8 Statistics of the local observations held independently among all links $j = 1, \ldots, M$, where $M = 20$ and $p = 1/M$. (a) RSD of $\hat{Y}_k^j(n)$ from Protocol B for the 8-th arm; (b) RSD of $\hat{Y}_k^j(n)$ from Protocol B at time $n = 1000$; (c) $\hat{Y}_k^A(n)$ from Protocol A at time $n = 1000$; (d) Average of $\hat{Y}_k^j(n)$ from Protocol B with homogeneous initialization; (e) Average of $\hat{Y}_k^j(n)$ from Protocol B with heterogeneous initialization; (f) Average of $\hat{Y}_k^j(n)$ from 100 independent simulation runs of Protocol B with homogeneous initialization. 135
5.9 Regrets of the semi-distributed MAB protocol vs. that of the distributed MAB protocols, where $M = 20$ and $p = 1/M$ unless explicitly indicated in the sub-figure. (a) Theoretical bound vs. simulation results for regret of the semi-distributed MAB protocol; (b) Regret of Protocol A; (c) Regret of Protocol B with homogeneous initialization; (d) Regret of Protocol B with heterogeneous initialization; (e) Regret of Protocol B with homogeneous initialization where $p = 0.1$; (f) Regret of Protocol B with homogeneous initialization where $M = 10$. 138
5.10 Regrets of the semi-distributed MAB protocol vs. that of the distributed MAB protocols, where there is a one-time data loss within $\mathcal{A}$ or $\mathcal{B}$ at time $n_I = 50$. (a) Regret of Protocol A; (b) Regret of Protocol B with homogeneous initialization; (c) Regret of Protocol B with heterogeneous initialization. 140

List of Abbreviations

RTS	Request to Send
CTS	Clear to Send
CSI	Channel State Information
PHY	Physics Layer
MAC	Media Access Control
DOS	Distributed Opportunistic Scheduling
CAT	Constant Access Time
CDT	Constant Data Time
MAB	Multi-Armed Bandit

Chapter 1

Introduction

In this chapter, we describe the background and motivation for the problem that is studied in this dissertation. We summarize related works for this topic and the contribution of this dissertation.

1.1 Background

There are several common factors that affect the performance of a wireless communication system, e.g. the time-varying channel fading and the co-channel interference from concurrent wireless transmissions within a neighborhood. In the traditional design of wireless systems, the loss due to these two factors are usually handled independently, i.e. the channel fading is considered at the physical (PHY) layer while the co-channel interference is addressed at the media access control (MAC) layer. The layered design philosophy enables us to overcome these two issues independently, reducing the complexity of the system design and analysis. The separation of point-to-point link reliability and multiple access functionality relies on the implicit assumption that the PHY layer can work perfectly to hide fading from the MAC layer. However, recent results show that there is a coupling between the time-scales of PHY fading and MAC. It is shown in [1,2] that channel fading and co-channel interference often occur on the same time scale. It is difficult to tell if a packet loss is due to PHY layer channel fading or MAC layer co-channel interference. Fading can often adversely affect the MAC layer protocols in many realistic systems. Hence a separate design for the PHY and MAC layers cannot achieve the best overall system performance. A unified PHY and MAC layer design for wireless ad-hoc network is desired in order to achieve better overall system performance compared to the separate system design approach case. On the other hand, the traditional PHY layer techniques are usually based on the assumption that intrinsic channel fluctuations (e.g. temporal and frequency variations) of fading channels are harmful to reliable communications. These conventional PHY layer techniques usually try to counteract the adverse effects of channel fluctuations to improve the system performance. This is true if the focus is the instantaneous performance of an individual wireless link. However, if the overall system performance is considered, it is possible to take advantage of the channel fluctuations. It is usually referred to as opportunistic communications in the literature [3–6]. One way to jointly consider these two issues is to schedule data transmissions at the MAC layer such that PHY layer channel information are utilized to exploit better opportunities for communications in the system. These works are usually known as as opportunistic scheduling or channel-aware scheduling [7–13].[1]

The concept of opportunistic scheduling comes from a system-wide point of view: Instead of treating the channel fading as a source of unreliability and try-

[1] There are also works on opportunistic communications focusing at the network layer, often known as opportunistic packet forwarding or opportunistic routing, see e.g. [14–17]. Works along this line are not discussed in this work.

ing to mitigate the channel fluctuations, fading can be exploited by transmitting information opportunistically when and where the channel is strong [3–6]. Opportunistic scheduling is particularly desirable for communications between multiple autonomous micro agents in adversarial environments, where the highly dynamic nature of the terrain blocks the line-of-sights (LOS) between the vehicles and results in reflection and scattering that make reliable point-to-point communications challenging [18]. We showed that in such circumstances it is impossible to maintain active communications at all times for each individual link even with collaborations between these micro agents [19]. On the other hand, energy is also strictly constrained in each agent due to its small size. In [20], we studied the problem of efficient communications between a group of autonomous vehicles with energy consumption and total operation time constraints. Our idea is to exploit communication opportunities at different positions along the trajectories of the nodes. We showed that opportunistic communications significantly improves the system performance, both in terms of the total operation time when the agents only transmit situational information and data throughput when additional data transmission is needed.

1.2 Opportunistic Scheduling

Generally speaking, opportunities for communication arise in the presence of multiple independently faded signal paths which could originate from different sources, e.g. time, frequency or multiuser diversity. As a result, channel fluctuations can be exploited in an opportunistic fashion focusing on different aspects of these sources in the system.

Many works on opportunistic scheduling has focused on exploiting channel fluctuations from multiple users. In short, in a multiuser network where the same wireless medium is shared between users, there always exists some user with better channel quality compared to others. Hence the shared wireless medium can be utilized more efficiently if the user with better channel quality can be chosen for data transmission. It is often referred to as multiuser diversity in literature [3–13]. Almost all work on exploiting multiuser diversity in wireless networks have their roots in [3], where a power control scheme is proposed to maximize the information theoretic capacity of the uplink of a single cell with time-varying channels. The problem of exploiting opportunism in multiuser diversity can be generally described as the following [4, 5]: Consider a single antenna downlink flat fading channel with M users

$$y_m[k] = h_m[k]x[k] + w_m[k], \quad m = 1, \ldots, M,$$

where $\{h_m[k]\}_k$ is the channel fading process of user m and is assumed to be *i.i.d.*. Compared to a single-user system, the gain from opportunism in multiuser diversity lies in that the effective channel gain at time k is improved from $|h_1[k]|^2$ to $\max_{1\leq m\leq M} |h_m[k]|^2$. The amount of multiuser diversity gain depends on the tail of the fading distribution $|h_m|^2$. With a heavy tail, it is more likely that there is a user with a very strong channel, and a large multiuser diversity gain can be achieved. In general, the more users available to choose from, the larger the performance gain can be achieved by the scheduler. Hence a large channel fluctuation is not a drawback, but is preferred in terms of opportunism from multiuser diversity [4, 5].

To fully exploit opportunism in multiuser diversity, it is crucial to dynamically schedule resources among the users as a function of the channel states $|h_m[k]|$ for every time instance k. Since the seminar work [3], many works have been developed to design scheduling algorithms for downlink transmissions in multiuser wireless cellular networks [7–13]. There is also industry implementation like Qualcomm's High Data Rate (HDR) system (1xEV-DO) [6]. Several scheduling algorithms have been developed to study the throughput-optimal performance under different rules, e.g. the revenue-based policy [7] or the exponential rule [8]. Fairness is also an important issue when exploiting multiuser diversity, since few users with excellent channel qualities can easily starve other users if only efficiency or throughput are considered. Therefore, many works in this area concentrates on designing schemes that tradeoff between overall performance (e.g. throughput) and quality-of-service (QoS) requirements (e.g. fairness) [9, 10]. A general framework for opportunistic scheduling is presented in [11] to exploit channel fluctuations and maximize system performance stochastically under a certain resource allocation fairness constraint. A utility based approach is used to solve the problem of determining which user should be scheduled to transmit at each time slot so that the network performance is optimized under fairness constraints. The optimality of the scheduling scheme is established and a practical procedure is proposed to implement the scheme. It is shown that their scheme results in performance improvements of 20%-150% compared with a scheduling algorithm that does not take into account channel conditions. Three heuristic time-fraction assignment schemes are proposed to balance the system performance and fairness among good and bad users. This problem is further extended to multichannel wireless networks in [12]. The authors developed a general methodology to exploit multiuser diversity over multiple wireless channels using an adaptive control framework. The selection of the best users and rates from a complex general optimization problem is transformed into a decoupled formulation, i.e. a multiuser scheduling problem that maximizes total system throughput, and a control-update problem that ensures long-term deterministic or probabilistic fairness constraints. These two sub-problems can be solved separately to simplify the design procedure. In solving the multichannel problem, their key technique is to jointly exploit the variations in the source consumption of multiple users to opportunistically select users with greater throughput potential while ensuring fairness constraints. Practical schedulers are also designed and evaluated to approximate these objectives. While [7–11] focus on the performance of opportunistic scheduling algorithms at the packet level for a static user population, this problem is studied at the flow level in a dynamic setting with random finite-size service demands [13]. The authors show that the user-level performance may be evaluated in certain cases by means of a multi-class processor-sharing model, where the total service rate varies with the total number of users. Various statistics, e.g. the distribution of the number of active users from different classes, the mean throughput, are discussed under this model. It is also shown that in the presence of channel variations, greedy or myopic strategies which maximize throughput in a static scenario, may result in sub-optimal throughput performance for a dynamic user configuration and cause instability effects.

Opportunistic scheduling for multicast applications is also studied under the name of opportunistic multicasting [21–25]. Conventional multicast scheduling exploits the multicasting gain by serving all users simultaneously. However, to prevent channel outage, the message has to be sent at a low rate which is constrained by the user with the poorest channel conditions, which in turn results in performance degradation. The opportunistic multicasting problem is first studied in [21] to jointly exploit the multiuser diversity gain and the multicast gain offered by the wireless medium. Their idea is to schedule transmission to a fraction of the users that have favorable channel conditions. A throughput-delay tradeoff is studied by varying the fraction of users targeted in each transmission. The asymptotic optimality of the throughput for a median user scheduler is established, where the best 50% users are served in each transmission. In [22], the optimal user selection ratio in static and dynamic opportunistic multicasting is investigated to maximize multicast throughput in a homogeneous network. The authors study the general order statistics of users' supportable data rates and derive its limiting distribution based on extreme value theory, and establish the optimal selection ratio by maximizing the average network throughput. A dynamic selection algorithm is proposed to adjust the user selection ratio adaptively in each transmission. It is shown that the opportunistic multicasting scheme with optimized static and dynamic selected ratios outperform the median-user static opportunistic multicasting scheme. This problem is further extended to heterogeneous networks where users are subject to different channel statistics [23]. The authors consider a single-cell wireless network with users uniformly distributed in a circular region around the base station. The key idea of [23] is that system performance may be predicted by the behavior of users in the outmost ring of the cell, which are approximately homogeneous. The multicast throughput maximization problem is also studied for opportunistic multicasting problem with erasure coding [24]. A linear gain for the multicast capacity is shown over *i.i.d.* Rayleigh fading channels with respect to the number of users for a large number of blocks. The analysis are extended to the case with shorter block lengths and the delay-capacity tradeoffs under a simple setting are quantified. The authors also show the achievable gains for non-*i.i.d.* channel conditions by modifying the proportional fair sharing (PFS) scheduling algorithm presented in [4]. It is shown that opportunistic multicasting with erasure coding can significantly improve the system performance under both *i.i.d.* and non-*i.i.d.* channel conditions. The opportunistic multicasting problem is also extended to multichannel networks in [25]. The problem is formulated as a system throughput maximization problem subject to the fairness constraints derived from each user's specific minimum QoS requirement. A priority metric is adopted which follows an exponential rule. This metric takes into account the CSI for opportunism from different subsets of the multicast users at different time slots and sub-channels. It also explicitly includes a term to compensate the deviations observed by individual users from their QoS targets.

In these work, timely channel information is required to enable effective opportunistic scheduling, which makes the scheduling algorithm essentially centralized. This kind of architecture usually happens in the downlink of a wireless cellular network, where the base station acts as a central controller and control channels are available for channel state feedback. However, frequent feedback of CSIs also imposes additional overhead for the system at the same time. There are few works in systematic characterization of the overhead due to opportunism in the context of multiuser diversity [26].

There are also works on opportunistic scheduling for the uplink in a many-to-one network, where a group of users communicate to a single user (e.g. a base station in a cellular network) [27–31]. Different from the downlink, the closed-loop channel feedback is not available and the sender has no centralized control. Channel-aware ALOHA [27, 28] first discusses this problem under a collision channel model. In channel-aware ALOHA, each user only knows its own channel quality and the statistics of other users' channel fading. Channel-aware ALOHA is first proposed in [27] as a modification to the ALOHA protocol, where a user only transmits when its channel gain is above a threshold H_0. It is shown that the multiuser diversity is achieved. In fact, the total system throughput increases at the same rate as in a system with a centralized scheduler, and asymptotically the fraction of throughput loss due to random access is $1/e$. Splitting algorithms are presented in [28] to resolve collisions over a sequence of mini-slots and determine the user with the best channel. The idea is to determine two thresholds (H_l and H_h) for each mini-slot such that at each time only users whose channel gains satisfying $H_l < h < H_h$ are allowed to transmit. It is shown that for a system with *i.i.d.* block fading and a fixed number of backlogged users, the average number of mini-slots required to find the user with the best channel is less than 2.5, which is independent of the number of users or the fading distribution. With the proposed splitting algorithms the throughput is improved and approaches the optimal value as the channel's coherence time increases. Channel-aware ALOHA is also studied under a reception channel model [29], where the reception is allowed to depend on the channel states of the transmitting users and it is also possible to model the simultaneous reception of multiple packets. In [29], the transmission probability is allowed to be a function of the CSI named transmission control, where both polulation-independent and polulation-dependent transmission controls are considered. It is shown that the effect of transmission control is equivalent to changing the probability distribution of the channel state. In [30], the authors derive sufficient conditions for system stability as well as upper bounds on average queue sizes using the dominant system approach for channel-aware with random arrival traffic. They prove the binary scheduling scheme of [27] maximizes the sum throughput for a homogeneous network, and maximizes the sum of the logs of the average throughputs while asymptotically guaranteeing fairness among users. Opportunistic ALOHA is also extended to the multichannel case for OFDMA wireless network [31], where the number of channels used and the corresponding channel conditions would affect a user's transmission rate, and on the other hand the number of channels used and the transmission probability would in turn determine the collision probability. A key step in [31] is to characterize the mapping from a user's channel conditions to the number of channels used and the transmission probability. The multichannel opportunistic ALOHA has a larger diversity gain compared to the single channel case, since the channel variation comes not only from the independence across users but also from the independence across frequency channels. It is showed that the multichannel opportunistic ALOHA approaches $1/e$ of the optimal centralized scheme, provided that the only performance loss is due to the contention from random access.

There have been few works to exploit opportunism in multiuser diversity for wireless ad-hoc networks due to the distributed nature of the network. Existing works are usually based on heuristics and focus on IEEE 802.11 like network based on the RTS/CTS handshaking. For example, the idea of multicast RTS and prioritized CTS have been proposed to exploit multiuser diversity [32, 33]. In [32], based on the multicast RTS probing, the sender chooses the neighboring nodes with channel quality better than a certain level to schedule the transmissions of packets in its queue. In the Contention-Based Prioritized Opportunistic (CBPO) MAC protocol [33], based on multicast RTS channel probing, all qualified users contend the channel with pulses of energy signals called black burst, and only the users with the highest priority will send CTS back to the sender. Opportunistic Medium Access and Auto Rate (OMAR) [34] considers a cluster based approach to exploit the multiuser diversity in ad hoc networks. In OMAR protocol, each node with a certain number of links is enabled to form a cluster and function as the clusterhead to coordinate multiuser communications locally. In each cycle, the cluster head initiates medium access with certain probability and then the cluster members distributedly make medium access decision based on the observed instantaneous channel conditions. A CDF-based K-ary opportunistic splitting algorithm and a distributed stochastic scheduling algorithm are proposed to resolve intra- and inter-cluster collisions. Fairness is formulated and solved in terms of social optimality within and across clusters. In [35], the authors proposed to study a distributed opportunistic scheduling (DOS) problem for ad-hoc networks, where M links contend the wireless medium and schedule data transmissions in a distributed fashion. In such networks, the transmitter has no knowledge of other links' channel conditions, and even its own channel condition is not available before a successful channel probing. The channel quality corresponding to one successful probing can either be good or poor due to channel fluctuations. In each round of channel probing, the winner makes a decision on whether or not to send data over the channel. If the winner gives up the current opportunity, all links re-contend again, hoping that some link with better channel condition can utilize the channel after re-contention. The goal is to optimize the overall system performance. The authors show that the decision on scheduling further channel probing or data transmission is based only on local channel conditions, and the optimal strategy is a threshold policy.

1.3 Contribution of the Dissertation

In this dissertation, we address the channel-aware scheduling problem for wireless ad-hoc networks where channel probing is adopted to property assist in accessing the wireless medium and resolving collisions in a decentralized fashion. The main contribution is that we consider this problem for channel fading with both parametric and non-parametric models, and we present a mathematical treatment to handle how channel dependencies affect the performance of the protocol.

In the first part of the dissertation, we consider the channel-aware scheduling problem assuming that the instant CSIs are not available but the channel statistics are known. In Chapter 3, we consider dependencies of the channel rates between different time instances during the channel probing phase and its impacts on the transmission scheduling. We use optimal stopping theory to model this problem, but at carefully chosen time instances when effective decisions are made. By merging repeated decisions, the new model transforms winners' channel rates into independent random variables, even though these channel rates are not independent under the block fading channel model. We characterize the system throughputs of the distributed opportunistic scheduling scheme assuming the problem has infinite stages. We then develop a measure to see how likely the problem can be analyzed as an infinite horizon optimal stopping problem. We characterize the actual system throughput if we ignore the finite horizon constraint and design stopping rule based on the infinite horizon analysis. We then analyze the problem using backward induction if the finite horizon constraint cannot be ignored. We develop one recursive approach to solve the problem and show that the computational complexity is linear with respect to the network size. We then present an improved protocol to reduce the probing costs which requires no additional cost and characterize its performance. In Chapter 4, we extended this problem to ad-hoc networks where the wireless spectrum can be divided into multiple independent sub-channels for better efficiency. We start with a naive multi-channel protocol where the scheduling scheme is working independently from sub-channel to sub-channel. We show that the naive protocol can only marginally improve the system throughput. We then develop a protocol to jointly consider the opportunistic scheduling behavior across multiple sub-channels. We characterize the optimal stopping rule and present several bounds for the system throughput of the multi-channel protocol. We show that by joint optimization of the scheduling scheme across multiple sub-channels, the proposed protocol improves the system throughput considerably in contrast to that of single-channel systems.

In the second part of the dissertation, we study this problem under the assumption that neither the instant CSI nor the channel statistics are known. In Chapter 5, we formulate and solve the opportunistic scheduling problem using multi-armed bandit (MAB). We first present a semi-distributed MAB protocol which serves as the baseline for performance comparison. We then propose two distributed MAB protocols, where each link keeps a copy of local observations and plays its own MAB game independently. We prove that these distributed protocols can be considered as a generalized MAB procedure, where each link can update its local observations for infinitely many times. By means of numerical results, we compare statistics of the local observations and the regrets under different parameters. Simulation results suggest that the local observations that are held independently at each link converge, and they converge to the true parameters if the game is played independently for multiple times. Simulations also show that the distributed protocols yields a regret which is greater than but still comparable to that of the semi-distributed protocol.

Chapter 2

Distributed Opportunistic Communications for Collaborative Control Application

2.1 Introduction

Collaborative control of groups of autonomous agents (robots, unmanned vehicles, etc.) has gained a growing amount of attention recently. Collaborative robotics, automated highway services, mobile sensor networks, and disaster relief operations are examples of applications in which advances in wireless and other technologies has led engineers to design groups of unmanned mobile vehicles [36]. In all of the above applications there is a strong incentive to come up with efficient decentralized control and decision-making schemes. Decentralization is preferred due to lack of expensive central coordination and robustness to single node failure. A challenging issue in design of collaborative swarms of autonomous vehicles is the need to implement efficient communication mechanisms. In many control theoretic studies, certain communication capabilities are implicitly assumed to hold [37–40].

In this chapter, we explicitly address the effects of communication on the performance of the networked system with emphasis on maintaining group connectivity. We study both the control and communication problems for a group of autonomous vehicles that are maneuvering with little or no direct human supervision in an adversarial environment. The mission is to explore the terrain, cover a target area while avoiding any possible obstacles or threats, and finally send information about features of the area to a command center. Building on our earlier work [40, 41], we use gradient flow based artificial potential methods for path planning [42, 43] in a kinematic setup. Despite their limitations, artificial potential based navigation functions have been found lots of applications in collaborative control [37, 38, 44, 45]. Hybrid stochastic methods have been proposed to overcome local minima problems [41]. We study the effects of communication between nodes on the group's path planning by comparing two schemes. In one scheme the vehicles only process their sensed local information whereas in the second scheme they collaborate by communicating among themselves. We study the performance of the wireless inter-vehicle network based on the IEEE 802.11 media access mechanism. Simulation results show that collaboration between vehicles results in better performance for path planning and wireless inter-vehicle communications. However, contention-based communication protocols are likely to fail due to severe channel conditions.

To solve this issue, we look at the joint control and communication problem from a different point of view. When there are other constraints besides maintaining inter-vehicle communications, it is not efficient to attempt reliable communication *regularly*. We address this problem under the constraints of energy consumption and total operation time to perform the mission. We propose an algorithm to seek communication opportunities based on the qualities of the wireless channels and make communication attempts accordingly. We compare our algorithm with a non-opportunistic algorithm, in which a vehicle makes a fixed number of communication attempts at a new position before moving to the next position. We show by simulation that the proposed algorithm reduces the total operation time when there are only position information to be exchanged, and also communicates more packets utilizing the same operation time when there are additional data traffic.

2.2 System Model

We consider a group of n autonomous ground vehicles maneuvering within an area $\mathcal{A} \subset \mathcal{R}^2$, e.g. a battlefield or a building with unknown potential dangers. Besides the boundary of $\mathcal{A}$, there is very limited knowledge available regarding the internal structure or topology of $\mathcal{A}$. The vehicles' mission is to explore $\mathcal{A}$ under little or no direct human supervision, cover a target area $\mathcal{T} \subset \mathcal{A}$ while avoiding any possible obstacles or threats, and finally send information about features of $\mathcal{A}$ to some server, e.g. a command center.[1] We assume there is only one common target $\mathcal{T}$ for all vehicles.

The main constraints of maneuvering the group of vehicles come from the obstacles and moving threats that are distributed in $\mathcal{A}$. An obstacle is a closed area that cannot be entered by any vehicle. A moving threat is an object that moves along an unpredictable trajectory with an unknown speed. A vehicle must keep at least a distance of R_e away from any moving threat, otherwise it will be destroyed. In addition to the obstacles and moving threats, vehicles should keep a safety distance from each other in order to avoid collisions while maintaining communications with

[1]We will use the term server or command center interchangeably throughout this paper.

nearby vehicles. We further assume that the size of an obstacle is much larger than that of a vehicle or a moving threat, hence we denote a vehicle or a moving threat as a point in $\mathcal{A}$ for simplicity.

Before starting to maneuver, each vehicle is given an initial position of the target $\mathcal{T}$. However, the position of the target can change during the maneuver. This is because either better motion planning results are available after collecting certain amount of information, or capturing a new target is required after the environment has changed considerably. In this paper, we assume the change of the target position can only be initiated from the server. The server sends the message of target update to one or several vehicles depending on links available, and the message is gradually spread to other vehicles via vehicle-to-vehicle communications. Since the environment in $\mathcal{A}$ is highly dynamic, we assume that there is no global information available about the positions of other vehicles, obstacles or moving threats. Instead, each vehicle is equipped with devices for short-range detection, i.e. a vehicle can discover another object (another vehicle, an obstacle or a moving threat) if it is within a distance of R_d.

The vehicles also have other devices such as sensors or cameras to capture various features of the internal structure of $\mathcal{A}$, which are later delivered to the server. However, due to the highly dynamic nature of the environment, the server can only access limited number of vehicles at any time. From time to time, the server may change the vehicles from which data are pulled. Hence it is necessary that a vehicle can deliver its data to any other vehicle via vehicle-to-vehicle communications. On the other hand, since each vehicle only has information about positions of the neighboring objects from local detections, they need to exchange such information between themselves. These information are time-sensitive, since each vehicle can have a better trajectory if it collects more information regarding the position of the target, other vehicles, obstacles and moving threats. We denote these information as *position information*, and other information that are delivered to the server as additional *data traffic*.

It is well known that wireless channels used for communication in such settings are vulnerable to fading and interference. The mathematical modeling of the wireless channels for our application is very challenging due to the highly dynamic nature of the terrain which blocks the line-of-sights (LOS) between the vehicles and results in reflection and scattering among many other physical phenomena which affect the transmitted signals [18]. Interference happen when more than one pair of vehicles attempt to communicate simultaneously within a short distance and thus lead to confliction in the wireless medium. In this paper, we mainly consider the shadowing effects and model the path loss based on the Fresnel zone radius and the obstruction that lie in the first Fresnel zone [46].

2.3 The Collaborative Control Algorithm for Networked Vehicles

We consider the general high level kinematic path planning problem for the group of vehicles over a wireless network. The algorithm generates a sequence of waypoints to follow by each vehicle. The algorithm uses artificial potential navigation functions and is based on our previous work [40]. The potential functions are chosen to lead the vehicles towards the target while avoiding collision and moving threats.

2.3.1 Potential Functions

We assume time is slotted. At time t, let $\mathcal{V}(t)$ denote the set of vehicles that are alive, $\mathcal{O}(t)$ the set of obstacles, and $\mathcal{M}(t)$ the set of moving threats. Let $p_i(t) = (x_i(t), y_i(t))$ be the position of the i-th vehicle at time t. We let $\mathcal{N}_v^i(t)$ denote the set of vehicles known to the i-th vehicle at time t [2]

$$\mathcal{N}_v^i(t) = \{j \in \mathcal{V}(t) : j \neq i,\ i \text{ knows the position of } j\}. \tag{2.1}$$

Similarly, we define the set of the obstacles and moving threats known to the i-th vehicle at time t as

$$\mathcal{N}_o^i(t) = \{j \in \mathcal{O}(t) :\ i \text{ knows the position of obstacle } j\} \tag{2.2}$$

and

$$\mathcal{N}_m^i(t) = \{j \in \mathcal{M}(t) :\ i \text{ knows the position of threat } j\}, \tag{2.3}$$

respectively. We denote by $\mathcal{T}^i(t)$ the target area at time t as far as the i-th vehicle knows.

To maneuver the vehicles, the following optimization problem is solved locally

[2] We define $\mathcal{N}_v^i(t)$ in this way instead of a set of neighboring vehicles within the detection range R_d, i.e. $\mathcal{N}_v^i(t) = \{j \in \mathcal{V}(t) : j \neq i, \|p_i(t) - p_j(t)\| \leq R_d\}$, since the i-th vehicle knows the positions of some vehicles beyond R_d in Figure 2.3.

at time t at the i-th vehicle

$$\begin{aligned} \min_{p_i(t)} \quad & J_{i,t}(p_i(t)) \\ \text{s.t.} \quad & G_k(p_i(t)) \leq 0,\ \ k \in \mathcal{N}_o^i(t) \\ & \|p_i(t) - p_i(t-1)\| \leq \delta, \end{aligned} \tag{2.4}$$

where $G_k(p_i(t))$ is the nonlinear constraint corresponding to the k-th obstacle, and δ is the step size. A potential function is constructed for each vehicle consisting of several terms, each of which reflects a goal or a constraint. The potential function $J_{i,t}(p_i(t))$ for the i-th vehicle at time t is

$$J_{i,t}(p_i(t)) = \lambda_g J_t^g(p_i(t)) + \lambda_n J_{i,t}^n(p_i(t)) + \lambda_o J_t^o(p_i(t)) + \lambda_m J_t^m(p_i(t)), \tag{2.5}$$

where J_t^g, $J_{i,t}^n$, J_t^o and J_t^m are the component potential functions relating to the target, neighboring vehicles, obstacles and moving threats respectively, and λ_g, λ_n, λ_o and λ_m are weighting factors. The potentials are chosen such that they encode the intended behavior of the vehicles regarding obstacle avoidance, keeping distance from neighbors and target finding correctly. For example, the target potential function is $J_t^g(p_i) = f_g(\rho(p_i, \mathcal{T}^i(t)))$, where $\rho(p_i, \mathcal{T}^i(t)) = \inf_{a \in \mathcal{T}^i(t)} \|p_i - a\|$ is the smallest distance from p_i to the target area $\mathcal{T}^i(t)$. Here $f_g(\cdot)$ is a strictly increasing function with $f_g(0) = 0$. This function guarantees that the i-th vehicle will move toward the target $\mathcal{T}^i(t)$ in absence of other objects. We use $f_g(r) = r^2$ in our simulations. The threat and obstacle avoidance potentials are on the contrary strictly decreasing functions of the vehicles' distance to threats and obstacles, and tend to infinity as this distance approaches 0. The neighboring potential is more involved, since it is designed to make the vehicles maintain some optimal distance. Figure 2.1

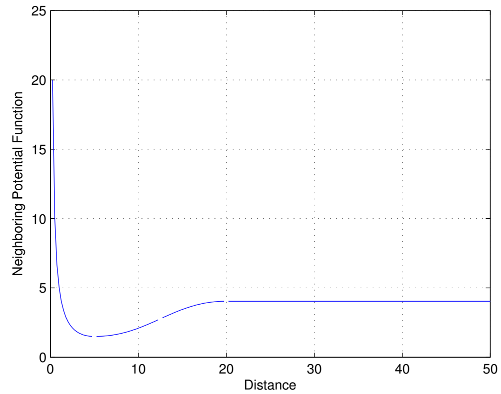


Figure 2.1: Neighboring potential function

provides the shape of the neighboring potential function that we use. For the detailed discussion of these components and the effects of the weights, we refer the reader to our earlier work [40]. The velocity of the i-th vehicle at time t is derived from the gradient descent equation:

$$\dot{p}_i(t) = -\frac{\partial J_{i,t}(p_i)}{\partial p_i} \tag{2.6}$$

In this chapter, we introduce two distributed algorithms to control the trajectories of the autonomous vehicles. It should be noted that in both cases the optimization is performed locally at each vehicle. However, we will show later that better performance can be achieved through collaboration even with the same optimization algorithm.

2.3.2 Distributed Control Algorithm with Local Information

We first introduce the distributed algorithm with only local information. In this case, the i-th vehicle performs a local detection and identifies all neighboring vehicles, obstacles, and moving threats within R_d and update its $\mathcal{N}_v^i(t)$, $\mathcal{N}_o^i(t)$ and

The initial position of the target $\mathcal{T}$ is loaded into each vehicle;
$t \leftarrow 0$;
while $\mathcal{V}(t)$ is not empty and some vehicle in $\mathcal{V}(t)$ is outside the target $\mathcal{T}(t)$ **do**
  for all vehicles in $\mathcal{V}(t)$ **do**
    The i-th vehicle identifies its local information set $\mathcal{N}_v^i(t)$, $\mathcal{N}_o^i(t)$ and $\mathcal{N}_m^i(t)$ through a local detection procedure;
    The i-th vehicle starts an optimization algorithm to minimize $J_{i,t}(p_i(t))$ and finds an optimal solution $p_i^*(t)$ based on $\mathcal{N}_v^i(t)$, $\mathcal{N}_o^i(t)$ and $\mathcal{N}_m^i(t)$;
    The i-th vehicle moves to the new position $p_i^*(t)$;
  end for
  $t \leftarrow t + 1$;
  Update the set of alive vehicles $\mathcal{V}(t)$;
end while

Figure 2.2: The distributed control algorithm with only local information

$\mathcal{N}_m^i(t)$. All other vehicles, obstacles or moving threats which are beyond the range R_d are completely unknown to the i-th vehicle, and thus are not included in $J_{i,t}(p_i(t))$. For the sake of simulations, we have considered discretization of the system with time steps small enough to preserve the stability of the original continuous algorithm. We also assume a prearranged synchronization scheme. The algorithm can be described as Figure 2.2.

From Figure 2.2 we can see the local information set $\mathcal{N}_v^i(t)$, $\mathcal{N}_o^i(t)$ and $\mathcal{N}_m^i(t)$ that can be obtained by the i-th vehicle highly depends on the detection range R_d, which is limited by the device and energy constraint. Hence the benefits of the application of Figure 2.2 to highly adversarial environments are limited, since local information may not be sufficient to provide the vehicles with appropriate maneuvering capabilities.

2.3.3 Distributed Control Algorithm with Collaborative Information

We notice if the i-th vehicle can access the position information of the objects that are beyond its detection range R_d, then a better performance is expected even with the same local optimization procedure to derive $p_i^*(t)$. This can possibly be done through local vehicle-to-vehicle communications. In our algorithm, at each time t, the vehicles exchange this control information before they start to transmit the data traffic. We note that the amount of data for control information is much smaller than that of the bulk data traffic. Hence the control information can spread rapidly among the vehicles, either by using the control channel when the wireless connections are established, or through other local, epidemic or gossip based protocols [47, 48]. The details of the algorithm are shown in Figure 2.3.

Figure 2.3 shows that not only a vehicle can get the position information of the objects beyond its detection range, but the server can also update the target position by communicating with only one or several vehicles. This is very useful when more accurate position estimates of the target is calculated at the server after collecting more information, or the target position must be changed due to discovery of hazardous objects nearby. Figure 2.3 can be used in more adversarial and highly dynamic situations.

The initial position of the target $\mathcal{T}$ is loaded into each vehicle;

$t \leftarrow 0$;

while $\mathcal{V}(t)$ is not empty and some vehicle in $\mathcal{V}(t)$ is outside the target $\mathcal{T}(t)$ **do**

  for all vehicles in $\mathcal{V}(t)$ **do**

    The i-th vehicle performs a local detection procedure and updates its local information set $\mathcal{N}_v^i(t)$, $\mathcal{N}_o^i(t)$ and $\mathcal{N}_m^i(t)$ accordingly;

    The i-th vehicle updates $\mathcal{T}^i(t)$ if notified;

    The i-th vehicle exchanges the local information set $\mathcal{T}^i(t)$, $\mathcal{N}_v^i(t)$, $\mathcal{N}_o^i(t)$ and $\mathcal{N}_m^i(t)$ with the vehicles that have connections between them, and updates the local information set accordingly;

    The i-th vehicle starts an optimization algorithm to minimize $J_{i,t}(p_i(t))$ and finds an optimal solution $p_i^*(t)$ based on $\mathcal{T}^i(t)$, $\mathcal{N}_v^i(t)$, $\mathcal{N}_o^i(t)$ and $\mathcal{N}_m^i(t)$;

    The i-th vehicle moves to the new position $p_i^*(t)$;

  end for

  $t \leftarrow t + 1$;

  Update the set of alive vehicles $\mathcal{V}(t)$;

end while

Figure 2.3: The distributed control algorithm with collaborative information

2.4 The Distributed Opportunistic Communication Algorithm for Networked Vehicles

2.4.1 Wireless Inter-Vehicle Networks

The communication module in our system is responsible for both inter-vehicle control information and data traffic transmission to and from the command center. Note that the control information is more time-sensitive but needs much less bandwidth compared to the bulk data traffic. In light of that, we assume that either the transmission of control information can be accomplished via control channels, or the bandwidth that is consumed by the control information is negligible compared to that of the bulk data traffic. Hence, in this paper we assume that the exchange of control information can be finished before the transmission of bulk data traffic in each time slot. With this assumption, the transmission of control information and data traffic can be well separated.

Hereafter we only consider the data transmission when discussing the performance of the wireless vehicle-to-vehicle network. We assume each vehicle always has data to transmit whenever communication opportunities are available.

Modeling the physical layer loss for wireless networks of moving vehicles is very challenging. The physical loss is highly environment dependent. Since the vehicles' motion in our scenarios are generally slow enough, we can simplify the problem by only considering the shadowing effects. The concept of Fresnel zone clearance has been used to analyze interference caused by obstacles near the path of a wireless transmission [46], where the first zone must be kept largely free from obstructions. We model the physical layer path loss by considering the obstructions occurring in the first Fresnel zone and the Fresnel zone radius. We use the IEEE 802.11 based medium access protocol. Under this assumption, the wireless medium is shared between vehicles using the CSMA/CA mechanism. We use an ad hoc routing protocol at the network layer, e.g. Dynamic Source Routing (DSR) routing [49]. We assume UDP protocol at the transport layer. This is because smaller delays are desirable for timely decision making at the server in our application, where certain level of packet transmission errors can be overcome by aggregating data traffic from all vehicles.

2.4.2 The Distributed Opportunistic Communication Algorithm

The wireless communication module in the networked system exchange messages for both position information and additional data traffic between vehicles. Note the position information is time sensitive but needs much less bandwidth compared to the bulk data traffic. In light of that, throughout this paper we assume whenever there is available bandwidth, the vehicles transmit data traffic only if there are currently no position information need to be sent.

We assume *the duration of one snapshot*, i.e. the time elapsed between a vehicle's presence in two consecutive positions, is $T_0 = T_s + T_c + T_m$, where T_s is the time used for local sensing, T_c is the time used for inter-vehicle communications, and T_m is the time used to move from the current position to the next position. When the resources are not strictly restrained, we can choose a large enough T_c such that the vehicles can exchange at least the position information within T_c. However, in other situations there are other constraints such as the total operation time to finish the mission which makes a choice of large T_c unacceptable. On the other hand, even without such constraints, it is difficult to achieve reliable inter-vehicle communications over the wireless medium *at all positions* along their trajectories, since the time-varying wireless links sometimes experience severe degradation due to obstacles, mobility of vehicles and radio interferences. In this case reliability for communications comes with a higher price of increased complexity of the communication algorithms and higher energy consumption. As a result, it is not efficient to make efforts for reliable communications equally at any positions, especially when the qualities of the wireless links are not good enough.

We now address the problem of efficient communications between vehicles under the constraints on energy consumption and total operation time to finish the mission. The proposed communication algorithm is based on opportunistic communications or channel-aware scheduling in some literature [3, 4, 11, 13, 27, 50–52]. The general idea is to communicate more when opportunities arise and less otherwise. There has been work on exploiting communication opportunities across time slots [50–52] or across multi-users [3, 4, 11, 13, 27]. In this paper, we explore another communication opportunity, i.e. we seek *communication opportunities across different positions* along the trajectories of the moving vehicles. At those positions where the wireless links are likely to fail, the vehicles proceed with their planned motion, and they attempt to make more communication attempts at positions where the qualities of the wireless links are better. As a result, the vehicles exchange information efficiently at positions with better link qualities and maneuver when the wireless links are severely degraded.

We use energy consumption and total operation time as the metrics for performance comparison. We assume most of the energy is consumed by local sensing, wireless communications and mechanical move.[3] Furthermore, we assume local sensing consumes much less energy compared to wireless communications and mechanical

[3]Here we ignore the energy consumption due to local computation and decision making.

move. If we assume energy consumption is proportional to the time duration, we can set $T_s \approx 0$. In order to compare the energy consumption and total operation time using *one single metric*, we further make the following assumptions: We choose a T such that a vehicle can communicate exactly one message with its peer within duration T. We then select a step size δ such that a vehicle spends the same amount of energy on communicating one message or moving a distance of δ, i.e. $T_m = T$.[4]

For the non-opportunistic algorithm, we assume that a vehicle makes K communication attempts in each snapshot regardless of the qualities of the wireless channels, i.e. $T_c = KT$. The duration of a snapshot is thus $T_1 = T_c + T_m = (K + 1)T$. For opportunistic communication algorithm, on the other hand, we can utilize the time slots with bad channel qualities to maneuver to some other positions with better channel qualities. Hence, unlike the non-opportunistic algorithm, a snapshot may have different number of slots for inter-vehicle communications, which depends on the qualities of the wireless channels. Here an important part is to estimate the qualities of the wireless channels in each position, based on which decisions on whether to communicate or not will be made. However, in practice it is difficult to estimate the wireless channel qualities accurately. Hence, we make the current decision on whether to communicate or not based on the result of the last communication attempt *at the same position*. To find the channel quality at a new position, a vehicle first makes one communication attempt after moving to a new position. Figure 2.4 implements this opportunistic approach.

In Figure 2.4, a vehicle first makes one communication attempt at a new po-

[4] Here a message can be a chunk of many packets.

$M \leftarrow$ maximum number of slots for communication without moving;
$t \leftarrow 0$, $\mathcal{V}(t) \leftarrow$ all vehicles that are initially alive;
Load the initial position of the target into each vehicle in $\mathcal{V}(t)$;
$M_i \leftarrow 0$ for each vehicle in $\mathcal{V}(t)$;
while $\mathcal{V}(t)$ is not empty and at least one vehicle in $\mathcal{V}(t)$ has not reached the target $\mathcal{T}(t)$ **do**
  for all vehicles in $\mathcal{V}(t)$ **do**
    The i-th vehicle performs a local detection procedure and updates its local set $\mathcal{N}_v^i(t)$, $\mathcal{N}_o^i(t)$, $\mathcal{N}_m^i(t)$ and $\mathcal{T}^i(t)$ accordingly;
    if the current position is new for the i-th vehicle **then**
      The i-th vehicle attempts to communicate one message;
      $M_i \leftarrow 1$;
    else
      if the last communication at the current position is successful and $M_i < M$ **then**
        The i-th vehicle attempts to communicate one message;
        $M_i \leftarrow M_i + 1$;
      else
        The i-th vehicle starts an optimization algorithm to minimize $J_{i,t}(p_i(t))$ and finds an optimal solution $p_i^*(t)$ based on $\mathcal{T}^i(t)$, $\mathcal{N}_v^i(t)$, $\mathcal{N}_o^i(t)$ and $\mathcal{N}_m^i(t)$;
        The i-th vehicle moves to the new position $p_i^*(t)$;
      end if
    end if
  end for
  $t \leftarrow t + 1$;
  Update the set of alive vehicles $\mathcal{V}(t)$;
end while

Figure 2.4: The opportunistic communication algorithm for networked vehicles

sition and the further decisions for the following slots are based on the previous outcomings. If the communication attempt is successful, the vehicle continues to communicate until a communication failure happens or the vehicle has communicated for M consecutive slots without moving; otherwise the vehicle moves to the next position. Hence the length of a snapshot is $T_2 = (X + 1)T$ where X is an integer-valued random variable depending on the channel qualities. Here M ensures the mission can be finished within a reasonable time. Although the duration of a snapshot can be different from time to time for different vehicles, in this paper we assume time is synchronized by the duration of slot T among different vehicles.

2.5 Simulation Results

Modeling the physical layer loss for wireless networks of moving vehicles is very challenging. The physical loss is highly environment dependent. Since the vehicles' motion in our scenarios are generally slow enough, we can simplify the problem by only considering the shadowing effects. The concept of Fresnel zone clearance has been used to analyze interference caused by obstacles near the path of a wireless transmission [46], where the first zone must be kept largely free from obstructions. We model the physical layer path loss by considering the obstructions occurring in the first Fresnel zone and the Fresnel zone radius.

We consider a group of autonomous vehicles in an 40m $\times$ 40m area $\mathcal{A}$ with 10 obstacles randomly distributed. We choose a scenario of 4 vehicles for illustration purposes, which are indexed from 0 to 3. The target area is a point whose position is

$(30, 30)$. There are 6 moving threats circling around to protect the target, where 4 of them are on a circle centered at the target $(30, 30)$, 1 of them is on a circle centered at $(28, 24)$, and 1 of them is on a circle centered at $(24, 28)$. The detection range is $R_d = 3$, and $R_e = \sqrt{2}/2$. The step size for maneuvering is $\delta = 0.5$. There are 4 wireless links in our simulation, where Flow 1 is from vehicle 0 to 3, Flow 2 is from vehicle 2 to 0, Flow 3 is from vehicle 1 to 2, and Flow 4 is from vehicle 3 to 1. In our simulation, we assume the wireless modules of the vehicles are full-duplex. We also assume these devices can detect communication success or failure. As an illustrative example, we set $K = 4$ for the non-opportunistic communication algorithm.

In our first simulation, we assume that there are only position information to be exchanged. Since there are no additional data traffic, we are interested in the saving of total operation time that can be achieved by the opportunistic communication algorithm. Hence we set $M = 4$, i.e. the vehicles have at most the same communication opportunities compared to the non-opportunistic algorithm. We run 100 independent simulations for the two algorithms respectively and show the results in Table 2.1 and Figure 2.5. We compare the means and standard deviations from the 100 simulations in Table 2.1. Table 2.1(a) and 2.1(b) show that the ratios of standard deviation to mean is at most 1.5% for the non-opportunistic algorithm and at most 9.9% for the opportunistic algorithm. We notice that the standard deviations of Table 2.1(b) are relatively larger, which is due to vehicles communicating opportunistically. We then show the means and standard deviations for the number of mechanical moves and messages exchanged until the 160-th slot in Table 2.1(c) and 2.1(d) respectively. The largest ratio of $\frac{\text{STD}}{\text{Mean}}$ is 0.8% in Table 2.1(c) and 7.7%

in Table 2.1(d). While running more independent simulations provides more reliable results, 100 simulations are enough for our illustration example. We randomly pick one from the 100 simulations and show the vehicles' trajectories for the two algorithms in Figure 2.5(a) and 2.5(b) respectively. We notice that the trajectories of the vehicles are slightly different due to the opportunistic way of communication in the latter case. We then compare the average performance of 4 vehicles from 100 simulations for the two algorithms in Figure 2.5(c) and 2.5(d). The average total operation time of the non-opportunistic algorithm is 428.8 slots in Figure 2.5(c). The total operation time of the opportunistic algorithm reduces to 223.3 slots in Figure 2.5(d), which is a time saving of 47.9%. On the other hand, the vehicles spend an average of 85.8 and 84.6 slots on actual mechanical move respectively, as shown in Figure 2.5(c) and 2.5(d). Hence there is no additional energy consumption on mechanical move for the opportunistic communication algorithm. Furthermore, the vehicles are able to exchange an average of 71.2 packets in the non-opportunistic algorithm as shown in Figure 2.5(c) and an average of 77.8 packets in the opportunistic algorithm as shown in Figure 2.5(d). Hence the opportunistic algorithm is able to exchange more position information even though the vehicles give up communication after failures at some positions. We also randomly pick one vehicle and compare the average performance from 100 simulations for the two algorithms. Figure 2.5(e) and 2.5(f) show the performance of vehicle 1 for the two algorithms. The average total operation time of this vehicle is 427.0 and 201.3 slots respectively for the two algorithms, i.e. using opportunistic algorithm the vehicle can save 52.9% of the total operation time compared to the non-opportunistic algorithm. Meanwhile,

Table 2.1: Mean and standard deviation from 100 independent simulations: (a) Total operation time for non-opportunistic algorithm; (b) Total operation time for opportunistic algorithm; (c) Number of mechanical moves until slot 160 for opportunistic algorithm; (d) Number of messages for position information exchanged until slot 160 for opportunistic algorithm.

Vehicle	0	1	2	3
Mean	407.2	427.0	447.6	433.5
STD	4.5	5.2	4.3	6.6
$\frac{\text{STD}}{\text{Mean}}$	1.1%	1.2%	1.0%	1.5%

(a)

Vehicle	0	1	2	3
Mean	236.9	201.3	259.6	195.4
STD	16.0	14.0	14.1	19.3
$\frac{\text{STD}}{\text{Mean}}$	6.8%	7.0%	5.4%	9.9%

(b)

Vehicle	0	1	2	3
Mean	63.3	66.6	66.9	71.7
STD	0.52	1.03	0.32	0.68
$\frac{\text{STD}}{\text{Mean}}$	0.8%	1.6%	0.5%	1.0%

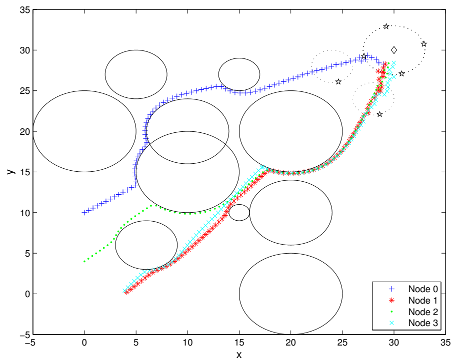
(c)

Vehicle	0	1	2	3
Mean	35.2	28.6	28.9	16.8
STD	1.2	2.0	0.8	1.3
$\frac{\text{STD}}{\text{Mean}}$	3.4%	7.0%	2.7%	7.7%

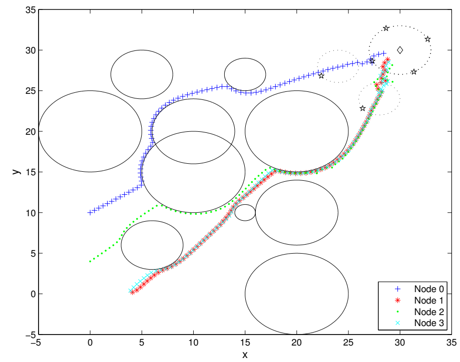
(d)

the average of the actual time used for mechanical move is 85.4 and 82.4 respectively in the two algorithms, and the average number of packets exchanged is 60.5 and 65.2 respectively, as shown in Figure 2.5(e) and 2.5(f). Hence, by utilizing opportunistic communications, the vehicles can maneuver to the target much earlier while there is no additional energy required for mechanical moving and no sacrifices to the amount of position information exchanged.

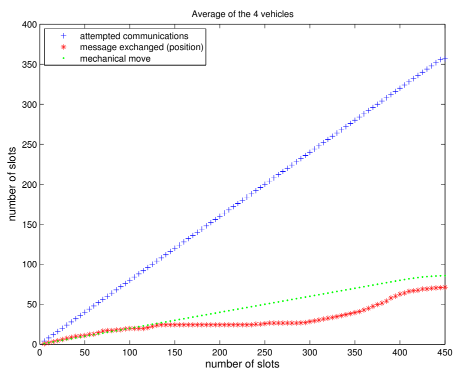
In the second simulation, we assume there is additional data traffic besides position information. In this case, we are interested in comparing the number of additional data packets successfully exchanged within the same operation time. Here we adjust the value of M until the total operation times for the two algorithms are close to each other. A value of $M = 12$ is finally used in our simulation. We also run 100 independent simulations. Again, Figure 2.6(a) and 2.6(b) show the trajectories of the vehicles from a randomly drawn simulation from the 100 independent simulations. We compare the average performance from 100 simulations for the two algorithms in Figure 2.6(c) and 2.6(d). The average total operation time is 428.7 slots for the non-opportunistic algorithm in Figure 2.6(c) and 406.5 slots for the opportunistic algorithm in Figure 2.6(d). Notice the total operation time for the opportunistic algorithm is slightly smaller than that of the non-opportunistic algorithm. The average of the actual time used for mechanical move is 85.7 slots for the non-opportunistic algorithm and 83.5 slots for the opportunistic algorithm. We then take a look at the number of packets exchanged. For position information, the vehicles exchange an average of 71.2 packets for the non-opportunistic algorithm and an average of 155.2 packets for the opportunistic algorithm. For additional data traffic, only an average of 94.2 packets is exchanged in Figure 2.6(c) and this number is increased to 148.8 in Figure 2.6(d). As a result, the vehicles can exchange more packets for both the position information and additional data traffic in the opportunistic communication algorithm even with a slightly smaller total operation time. Finally, following Figure 2.5, we also take a look at the average performance of vehicle 1 from 100 simulations for the two algorithms. The average total operation time of this vehicle is 426.7 slots in Figure 2.6(e) and 452.0 slots in Figure 2.6(f), which are close to each other. The average of the actual time used for mechanical move is 85.3 and 81.0 slots respectively. Meanwhile, for position information, the vehicle is able to exchange 60.4 packets for the non-opportunistic algorithm in Figure 2.6(e), and this number increases to 89.5 in Figure 2.6(f). The number of packets for additional data traffic is 87.9 for the non-opportunistic algorithm in Figure 2.6(e) and 239.4 for the opportunistic algorithm in Figure 2.6(f). Hence by using approximately the same total operation time and the actual time used for mechanical move, there is a considerable increase in the number of packets exchanged for both position information and additional data traffic.



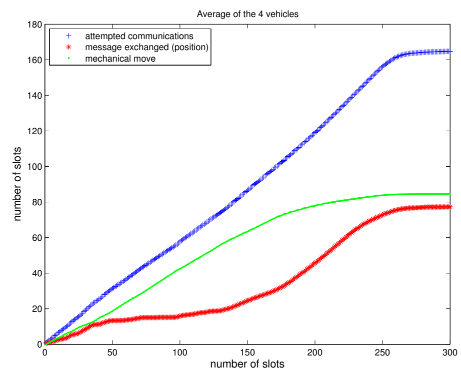
(a)



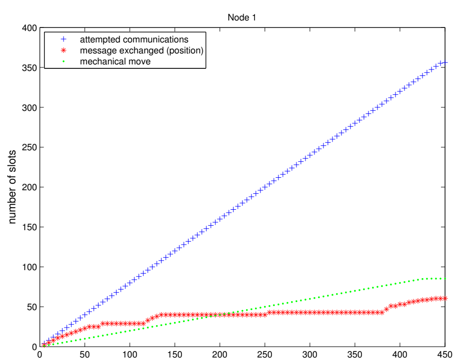
(b)



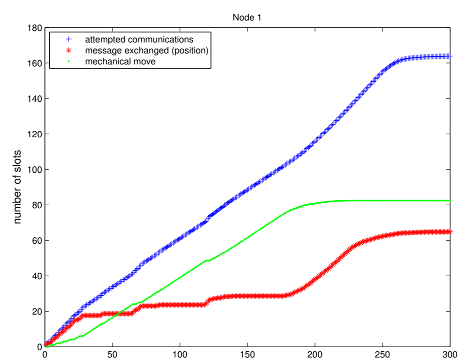
(c)



(d)



(e)



(f)

Figure 2.5: Performance comparing of the non-opportunistic and opportunistic communication algorithms, assuming that the vehicles only exchange position information. (a) Trajectories of the vehicles for non-opportunistic algorithm; (b) Trajectories of the vehicles for opportunistic algorithm; (c) Average performance of the vehicles for non-opportunistic algorithm; (d) Average performance of the vehicles for opportunistic algorithm; (e) Performance of vehicle 1 for non-opportunistic algorithm; (f) Performance of vehicle 1 for opportunistic algorithm.

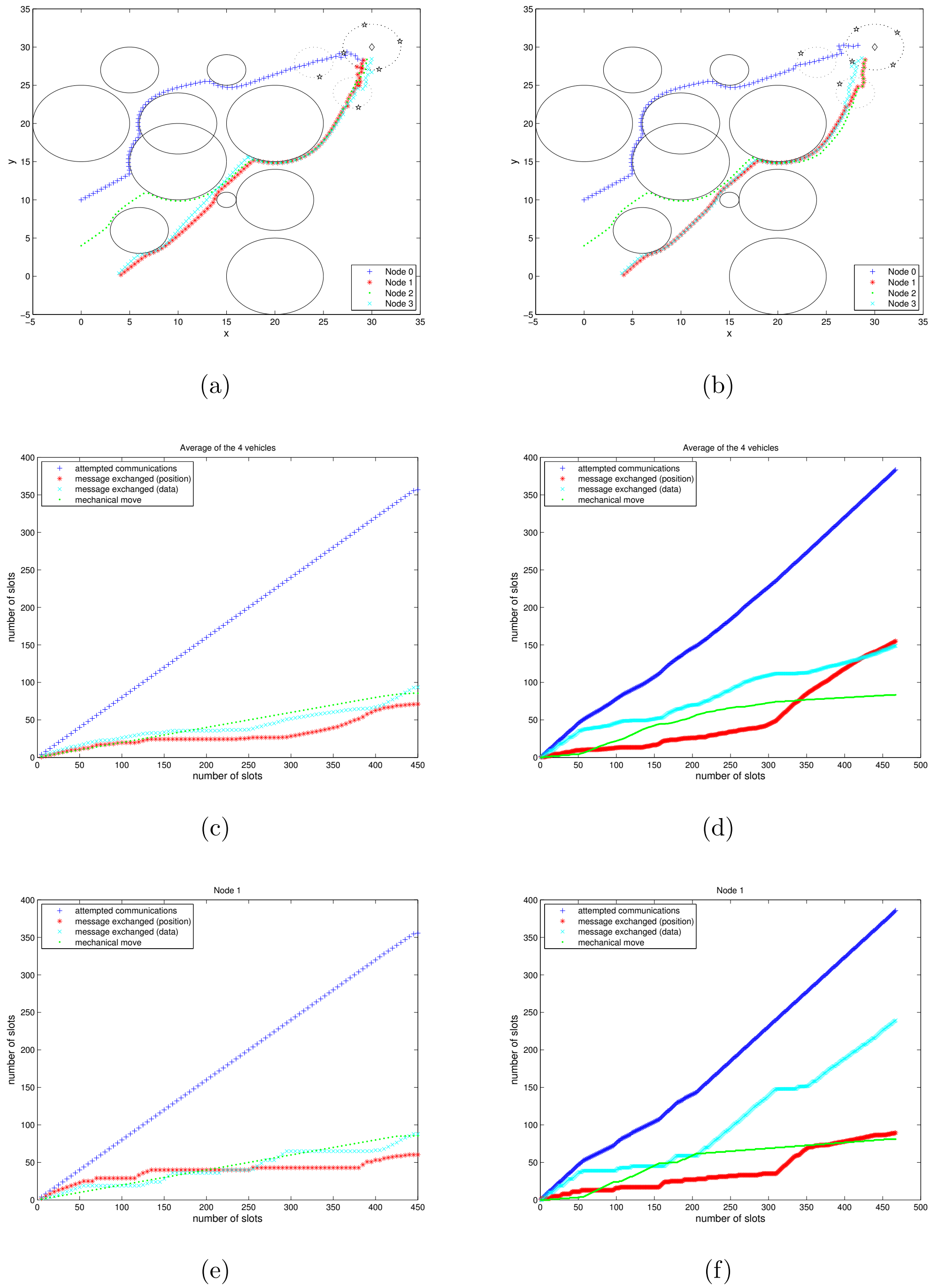


Figure 2.6: Performance comparing of the non-opportunistic and opportunistic communication algorithms, assuming that the vehicles exchange both position information and additional data traffic. (a) Trajectories of the vehicles for non-opportunistic algorithm; (b) Trajectories of the vehicles for opportunistic algorithm; (c) Average performance of the vehicles for non-opportunistic algorithm; (d) Average performance of the vehicles for opportunistic algorithm; (e) Performance of vehicle 1 for non-opportunistic algorithm; (f) Performance of vehicle 1 for opportunistic algorithm.

Chapter 3

Distributed Opportunistic Scheduling for Single-Channel Networks

3.1 Introduction

In general, opportunistic scheduling requires the CSIs at the PHY layer to schedule transmissions at the MAC layer. The difficulty involved with obtaining CSIs in a distributed manner makes many works on opportunistic scheduling only considers the centralized scenario. For example, many existing works assume a cellular like system model where a central scheduler tries to optimize the overall system performance by selecting the *on-peak* user for data transmissions [3, 4, 7, 10–12]. In contrast, in ad-hoc networks it is necessary to access the wireless medium and schedule data transmissions in a distributed fashion. So far few existing works have studied this problem in distributed scenarios. Such examples include rate adaptation with MAC design based on the RTS/CTS handshaking for IEEE 802.11 networks [50–52] and channel-aware ALOHA for uplink communications [27–29]. However, rate adaptation focuses on protocols to exploit temporal opportunities while leaving the distributed medium access issue to the RTS/CTS mechanism. On the other hand, channel-aware ALOHA associates the probability to access the uplink with channel quality under the assumption that each user knows its own channel state information (CSI) from the uplink. These schemes ignore the overhead due to the distributed nature of ad-hoc networks when considering the medium access and scheduling problem. In fact, these costs should be counted into the protocol design in order to fully exploit the channel fluctuations in the network.

In [35], the authors proposed to study a distributed opportunistic scheduling (DOS) problem for ad-hoc networks, where M links contend the wireless medium and schedule data transmissions in a distributed fashion. In such networks, the transmitter has no knowledge of other links' channel conditions, and even its own channel condition is not available before a successful channel probing. The channel quality corresponding to one successful probing can either be good or poor due to channel fluctuations. In each round of channel probing, the winner makes a decision on whether or not to send data over the channel. If the winner gives up the current opportunity, all links re-contend again, hoping that some link with better channel condition can utilize the channel after re-contention. The goal is to optimize the overall system performance. The authors show that the decision on further channel probing or data transmission is based only on local channel conditions, and the optimal strategy is a threshold policy.

One key issue in the design and analysis of opportunistic scheduling protocols for wireless ad-hoc networks is to seek an optimal trade-off between the costs to obtain the CSIs and the opportunities that can be exploited based on these information. When channel probing is adopted for this purpose, the problem reduces to the tradeoff between the durations elapsed for channel probing and those remaining for data transmissions. The authors in [35] consider the constant data time (CDT) model, where a fixed duration of T is available for data transmission regardless of the time consumed for channel probing. To further understand this tradeoff and its impact on the system performance, we consider the constant access time (CAT) model, where the total time duration available is a fixed amount T and the protocol needs to decide how to split T between channel probing and data transmissions in order to improve the system performance. On the other hand, in [35] the winners' channel rates were explicitly assumed to be independent during the channel probing phase, which is an ideal assumption. As we will explain in Section 3.3, there are inevitable dependencies between the winners' rates at different time instances during the channel probing phase. In our earlier conference paper [53], we analyzed the distributed opportunistic scheduling problem for the CAT problem under the ideal assumption that the winners' channel rates are independent during the channel probing phase. In a later technical report [54], we further investigate this problem under the popular block fading channel model. We explicitly consider how such dependencies could impact the transmission scheduling and hence the system performance. We use optimal stopping theory [55–57] to describe this problem, where we only choose the time instances when an effective decision is taken to make our mathematical analysis tractable. The contributions of this chapter include:

1) We study a distributed opportunistic scheduling problem under the popular block fading channel model where there are inevitable dependencies between the winners' channel rates during the channel probing phase. To the best of our knowledge, this problem has not been studied in the literature.

2) We present a concept named "effective observation points", where we only take observations at time instances when effective decisions are made. In this ap-

proach, repeated decisions by the same link are properly treated as a single decision. This approach makes our mathematical analysis tractable, where winners' channel rates in the probing phase are not independent in the first place.

3) We characterize the optimal stopping rules and network throughputs for networks at different scales. We show that the finite horizon analysis is necessary for networks whose sizes are not large enough, otherwise the actual achievable network throughputs may deviate a lot from the infinite horizon analysis results.

4) We propose a modified protocol to reduce the probing costs, which requires no additional overhead for protocol design. By analytical and numerical results, we show that the new protocol improves the system performance, in particular for scenarios when the network size is not large or the network is "over-probed". Furthermore, we show that the new protocol can reduce the energy consumed in the channel probing phase considerably. This makes the improved protocol of particular interest for networks whose nodes have limited battery life.

This chapter is organized as follows. In Section 3.2 we describe our system model for the distributed opportunistic scheduling problem. In Section 3.3 we formulate the problem as an optimal stopping problem and present our concept of effective observation points for analyzing the problem. We first present a rigorous analysis for the CAT problem based on the finite horizon approach in Section 3.4. Due to its computational complexity, in Section 3.5 we introduce an approximate approach to characterize the system performance. In Section 3.6 we present a modified protocol to reduce the probing costs, and analyze the performance improvement in network throughputs and energy savings in the channel probing phase. In Section 3.7 we introduce the results for the CDT problem and a performance comparison to the CAT problem. We show our numerical results in Section 3.8.

3.2 System Model and Motivation

In this section, we introduce our system model for the distributed opportunistic scheduling problem. Similar to the problem discussed in [35], we assume M links share the wireless medium without any centralized coordinator in an ad-hoc network. To access the wireless medium, all links have to probe first. Suppose the links adopt a fixed probing duration τ. A collision channel model is assumed, where a link wins the channel if and only if no other links are probing simultaneously. If link m probes the channel with probability $p^{(m)}$, the duration of the n-th round of channel probing is $T_n = \tau K_n$, where K_n is the number of probings before the channel is won by some link. Hence K_n has a geometric distribution $\text{Geom}(p_s)$ with parameter p_s, where

$$p_s = \sum_{m=1}^{M} p^{(m)} \prod_{j \neq m} \left(1 - p^{(j)}\right) \tag{3.1}$$

is the successful probing probability. Throughout this paper, we use superscript (m) to denote variables related to the m-th link, and subscript n to denote variables related to the winner in the n-th round of channel probing. We also use the terms "n-th round" of channel probing and "time n" interchangeably. At the end of the n-th round, winner s_n has an option to send data through the channel at the current available rate R_n or to give up this opportunity. Based on the current rate R_n, s_n makes a decision on whether or not to utilize the channel for data transmission in order to optimize the overall network throughput. If s_n gives up the opportunity, all links re-contend again. This procedure repeats until some link finally utilizes the channel. The goal is that all links cooperate indirectly to make the channel accessible by some link with a good enough channel quality.

The performance analysis in [35] relies on an important assumption: the winners' channel rates R_n are independent with respect to time n in the channel probing phase but can be locked for a constant duration T in the data transmission phase. It should be noted that the independence of $R^{(m)}$ within one block does not necessarily imply the independence of the winners' rates R_n. In fact, possible dependencies do exist between the winners' channel rates R_n, since some link $\tilde{m}$ might win the channel for multiple times within one block. This assumption can generally hold when the network size (i.e. the number of links in the network) is infinitely large. It is not necessarily true for a network with a finite size M. On the other hand, although opportunistic scheduling has been shown to improve the system performance dramatically for large networks [5, 12, 35], there are other factors we need to consider in the design of such systems. For example, we could take a look at the average waiting time for any link to access the medium [58]. Suppose the channel fading are *i.i.d.* for all M links in the network. Then based on the distributed opportunistic scheduling scheme [35, 53, 54], any link m is able to access the current block with a probability $\frac{1}{M}$. Hence it takes roughly M blocks before link m is able to send data over the wireless channel. This will lead to a long delay for large networks. Hence for such kind of systems, one practical approach is to consider multi-cell or multi-channel schemes [59–61] to trade-off several design goals (e.g. throughput, delays). In line with that, we argue that it is important to consider this problem for a network with a finite size M, which is the basis for a more complex multi-cell or multi-channel system.

To investigate how the dependencies of the winners' channel rates in the channel probing phase affect the system performance, we study this distributed opportunistic scheduling problem for the popular block fading channel model. We assume the channel rates are flat fading within one block. Hence the channel rate $R^{(m)}$ for any link m does not change within one block. The total block length T_s is separated into two parts as $T_s = T_p + T_d$, where T_p is for channel probing and T_d is for data transmission. At the end of the n-th round of channel probing, the total time duration for channel probing is $T_p = \sum_{i=1}^{n} T_i$. We consider the CAT model [53,54,59,60], where the transmitter has a fixed duration $T_s = T$ in total, leaving the available duration for data transmission as $T_d = T - \sum_{i=1}^{n} T_i$. If we decide to send data at the end of the n-th round, the normalized network throughput is

$$Y_n = \frac{R_n \cdot (T - \sum_{i=1}^{n} T_i)}{T}. \tag{3.2}$$

3.3 The Optimal Stopping Problem Formulation

In this section, we formulate the distributed opportunistic scheduling problem as an optimal stopping problem. In particular, we present the concept of *effective observation points* to facilitate the mathematical treatment of our problem.

The theory of optimal stopping [55–57] is about the problem of choosing a time to take a given action based on sequentially observed random variables in order to maximize an expected payoff. The stopping rule problem is defined by a sequence of random variables $\mathbf{X}_1, \mathbf{X}_2, \ldots$ whose joint distribution is known and a sequence of real-valued reward functions $Y_0, Y_1(\mathbf{x}_1), \ldots$. Let $(\Omega, \mathcal{B}, P)$ be the probability space, and $\mathcal{F}_n$ be the sub-σ-field of $\mathcal{B}$ generated by $\mathbf{X}_1, \ldots, \mathbf{X}_n$. We have a sequence of σ-fields as $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \ldots \subset \mathcal{F}_n \subset \ldots \subset \mathcal{B}$. A stopping rule is defined as a random variable $N \in \{0, 1, \ldots, \infty\}$ such that the event $\{N = n\}$ is in $\mathcal{F}_n$. Our goal is to choose a stopping rule N^* to maximize the expected reward $E[Y_N]$. If there is no bound on the number of stages at which one has to stop, this is an infinite horizon problem and the optimal return can be computed via the optimality equation. When there is a known upper bound on the number of stages, it is a finite horizon problem and the optimal return can be solved by backward induction. A short summary of optimal stopping theory can be found in Appendix A, and details on this topic can be found in [55–57].

At the end of the n-th round, winner s_n observes the probing duration T_n and the available channel rate R_n. Recalling that $T_n = \tau K_n$ and the fact that τ is a constant, we denote the observations at time n as a random vector $\mathbf{X}_n = (R_n, K_n)$ and one realization of $\mathbf{X}_n$ as $\mathbf{x}_n = (r_n, k_n)$. The σ-fields can be denoted as

$$\mathcal{F}_n = \{\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n\} = \{R_1, K_1;\, R_2, K_2;\, \ldots;\, R_n, K_n\}\,. \tag{3.3}$$

Then s_n makes a decision on whether or not to stop based on $\mathcal{F}_n$, to maximize the overall network throughput (3.2). Here a decision to "stop" means that s_n decides to utilize the remaining time duration for data transmissions. A decision to "continue" means that s_n decides to give up the current opportunity. Another round

for each link m **do**
  m probes the channel with a fixed probability $p^{(m)}$;
  if m wins the channel **then**
    m makes a decision on whether or not to send data over the channel;
    if m decides to utilize the channel **then**
      m sends data through the channel for a duration of $T - \sum_{i=1}^{n} T_i$ (CAT) or T (CDT), where n is the current index of channel probing;
    end if
  end if
end for

Figure 3.1: The distributed opportunistic scheduling protocol

of channel probing and decision making then begins. This probing and decision behavior continues within this block until winner s_N finally utilizes the channel for data transmissions, where N is the stopping time. It could be easily sensed and detected by all other links at this point. Hence all links don't send probing signals anymore until the beginning of the next block. If this procedure is repeated for I blocks independently, the decision making process can be described as

$$Y^* = \max_{N \in \mathcal{F}_n} \frac{E\left[R_N \cdot (T - \tau \sum_{i=1}^{N} K_i)\right]}{T}, \tag{3.4}$$

where N is the stopping time. This procedure can be described as in Figure 3.1.

Now the problem is to find an optimal rule N^* to maximize the overall network throughput. To do this, we need to characterize the joint distribution of R_n and

K_n. We notice that R_n and K_n are independent of each other, and K_n are also independent with respect to time n. However, the winners' channel rates R_n are not independent due to the block fading assumption. The dependencies of R_n make the mathematical analysis of this problem intractable. In this paper, we tackle this problem by using effective observation points instead of the original observation points in (3.3). The whole idea is motivated from the following fact: at time n, if the winner s_n decides to give up the opportunity, the same decision will be repeated for all future $\tilde{n} > n$ in this block when the channel is won by s_n again at time $\tilde{n}$. This is because utilizing the channel at time $\tilde{n}$ will only yield a smaller reward, i.e.

$$Y_{\tilde{n}} = \frac{R_{\tilde{n}}(T - \tau \sum_{i=1}^{\tilde{n}} K_i)}{T} < \frac{R_n(T - \tau \sum_{i=1}^{n} K_i)}{T} = Y_n, \tag{3.5}$$

where we used the fact that $R_{\tilde{n}} = R^{(s_{\tilde{n}})} = R^{(s_n)} = R_n$. It implies that an *effective* decision is always made at the time instances when a link wins the channel *for the first time*. If we only take observations at these time instances, the channel rates $\tilde{R}_n$ are independent. We denote the σ-fields at these time instances as

$$\tilde{\mathcal{F}}_n = \left\{\tilde{R}_1,\ \tilde{K}_1;\ \tilde{R}_2,\ \tilde{K}_2;\ \ldots;\ \tilde{R}_n,\ \tilde{K}_n\right\}. \tag{3.6}$$

Lemma 3.3.1. *The solutions to the optimal stopping problem based on $\tilde{\mathcal{F}}_n$ and $\mathcal{F}_n$ have different distributions for the stopping time N. However, both solutions have the same network throughputs and distributions for the elapsed probing durations $L = \sum_{i=1}^{N} K_i$.*

Hence if we do not care how many times a given link m has given up its opportunity upon winning the wireless medium, the problem is equivalent to analyzing the problem using $\tilde{\mathcal{F}}_n$ instead of $\mathcal{F}_n$. For the rest of this paper, we *always refer to the σ-fields at the effective observation points unless noted otherwise*. Hence we use the notations $\mathcal{F}_n$, T_n, K_n instead of $\tilde{\mathcal{F}}_n$, $\tilde{T}_n$, $\tilde{K}_n$ for short for the rest of the paper.

3.4 A Rigorous Performance Analysis: the Finite Horizon Approach

In this section, we characterize the optimal stopping rules and the network throughputs. We analyze the protocol using σ-fields (3.6) recorded at those effective observation points. By this notation, the number of effective probing links is *monotonically* decreasing as time n moves on, even though physically all links are still probing the wireless medium as in Figure 3.1. On the other hand, since no recall is allowed, if link m gives up its opportunity at some point, link m cannot reclaim it at a later time. As a result, the "last" winner must utilize the wireless medium for data transmission, otherwise the channel will be completely wasted for this block. Hence the stopping rule problem always has an implicit horizon at M, where M is the network size. The problem should be treated as a finite horizon problem and be solved by the backward induction approach [55–57].

We denote the optimal expected reward based on observations until the n-th round of channel probing as $\lambda_n^* = \lambda_n^*(\mathbf{x}_1, \ldots, \mathbf{x}_n)$. We will use the term the n-th round of channel probing, "time n" or "stage n" interchangeably in this section. The backward induction procedure can be described as

$$\lambda_M^*(\mathbf{x}_1, \ldots, \mathbf{x}_M) = Y_M(\mathbf{x}_1, \ldots, \mathbf{x}_M), \tag{3.7}$$

$$\lambda_n^*(\mathbf{x}_1, \ldots, \mathbf{x}_n) = \max \Big\{ Y_n(\mathbf{x}_1, \ldots, \mathbf{x}_n),$$
$$E\left[\lambda_{n+1}^*(\mathbf{X}_1, \ldots, \mathbf{X}_n, \mathbf{X}_{n+1}) | \mathbf{X}_1 = \mathbf{x}_1, \ldots, \mathbf{X}_n = \mathbf{x}_n\right] \Big\}, \tag{3.8}$$

where $n = 0, 1, \ldots, M-1$, and $Y_n(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ is the instant reward based on $\mathcal{F}_n$. At stage n, it is optimal to stop if $Y_n(\mathbf{x}_1, \ldots, \mathbf{x}_n) \geq \lambda_n^*(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ and to continue otherwise. The optimal return at stage n is the instant payoff if the decision is to stop and the expected payoff if the decision is to continue. The optimal network throughput is λ_0^*, i.e. the optimal expected reward before taking any observations.

However, it is not practical to directly solve this problem using (3.7) and (3.8) for two reasons. First, the channel rates r_n are generally continuous variables. We have to discretize r_n to use (3.7) and (3.8). Second, the instant observation $\mathbf{x}_n$ at time n is a two dimensional vector. To directly apply the backward induction procedure on $\mathbf{x}_n$, there will be too many states in the state space. The overwhelming computational complexity will restrict us to solve problems only with a small M. In this paper, we develop one approach to reduce the computational complexity for this procedure. First we note that the last item in (3.8) only depends on $\mathbf{x}_1, \ldots, \mathbf{x}_n$ since the expectation is taken with respect to $\mathbf{X}_{n+1}$. Hence we can denote it as

$$w_n(\mathbf{x}_1, \ldots, \mathbf{x}_n) = E\left[\lambda_{n+1}^*(\mathbf{X}_1, \ldots, \mathbf{X}_n, \mathbf{X}_{n+1}) | \mathbf{X}_1 = \mathbf{x}_1, \ldots, \mathbf{X}_n = \mathbf{x}_n\right] \tag{3.9}$$

for short. Now the problem in (3.7) and (3.8) reduces to the calculation of $w_n(\mathbf{x}_1, \ldots, \mathbf{x}_n)$. Next, we show that the calculation of $w_n(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ does not need all of these observations $\mathbf{x}_1, \ldots, \mathbf{x}_n$. To show this, we define the total number of probings up to time n as $L_n = \sum_{i=1}^n K_i$. Note that L_n is a random variable. We denote one realization of L_n as l_n.

Lemma 3.4.1. *Suppose the network size is $M \geq 2$, the expected reward at time n can be characterized as*

$$w_n(\mathbf{x}_1, \ldots, \mathbf{x}_n) = \begin{cases} w_M(r_M, l_M) & \text{for } n = M, \\ w_n(l_n) & \text{for } n = M-1, \ldots, 1, \\ w_0 = \lambda_0^* & \text{for } n = 0. \end{cases} \tag{3.10}$$

Proof. Since the network size is M, the backward induction procedure has a horizon at stage M. The reward at stage M is $w_M(\mathbf{x}_1, \ldots, \mathbf{x}_M) = \max\left\{0, \frac{r_M(T-\tau l_M)}{T}\right\}$. Hence $w_M(\mathbf{x}_1, \ldots, \mathbf{x}_M)$ only depends on r_M and l_M, and it can be denoted as $w_M(r_M, l_M)$ for short.

Now we let $n = M - 1$ in (3.9). We can see that the expectation in (3.9) is taken with respect to $\mathbf{X}_M$, i.e. R_M and K_M. We have showed that $w_M(\mathbf{x}_1, \ldots, \mathbf{x}_M)$ only depends on r_M and l_M, but is independent of r_{M-1}. Hence after taking the expectation, $w_{M-1}(\mathbf{x}_1, \ldots, \mathbf{x}_{M-1})$ is still independent of r_{M-1}. On the other hand, $w_{M-1}(\mathbf{x}_1, \ldots, \mathbf{x}_{M-1})$ does depend on l_{M-1}, since l_{M-1} remains in the expression after taking expectations with respect to K_M, where $L_M = L_{M-1} + K_M$. Hence w_{M-1} only depends on l_{M-1}. We can iterate this procedure from $n = M - 2$ to $n = 1$. As a result, for $n = M-1, \ldots, 1$, $w_n(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ can be denoted as $w_n(l_n)$ for short.

Finally, the network throughput is the optimal expected reward before taking any observations. That is to say $n = 0$. In this case, l_n can only be 0. Hence we can write it as $w_0 = \lambda_0^*$ for short. □

Following Lemma 3.4.1, we can use l_n as the only state for the backward induction procedure. The problem is reduced to a one-dimensional problem. To calculate w_0, we need to calculate $w_1(l_1)$ for all possible l_1, and then $w_2(l_2)$ for all possible l_2, and so on until stage M. Hence the problem is to compute $w_n(l_n)$ for $n = 1, \ldots, M-1$ and $w_M(r_M, l_M)$.

Theorem 3.4.1. *The optimal stopping rule for the distributed opportunistic scheduling problem is*

$$N^* = \min \left\{ n \geq 1 : R_n \geq \lambda_n^* \cdot \frac{T}{T - \tau \sum_{i=1}^{n} K_i} \right\}. \tag{3.11}$$

The optimal network throughput is $w_0 = \lambda_0^*$. Suppose the network size is M. The finite horizon analysis reduces to the calculation of w_0, which eventually iterates all $w_n(l_n)$ for $n = 1, \ldots, M-1$ and $w_M(r_M, l_M)$. The expected reward can be calculated recursively as

$$w_{n-1}(l_{n-1}) = \sum_{k \in \Omega_n(l_{n-1})} (1 - p_{s,n})^{k-1} p_{s,n} \cdot q_n(l_{n-1}, k) \tag{3.12}$$

$$q_n(l_{n-1}, k) = P_n(k) \cdot E_n(k) + [1 - P_n(k)] \cdot w_n(l_{n-1} + k), \tag{3.13}$$

where $q_n(l_{n-1}, k)$ is the conditional expected reward given $K_n = k$. K_n can take values in

$$\Omega_n(l_{n-1}) = \{ k \,|\, \tau \cdot (l_{n-1} + k) < T,\ k \in \mathbb{N} \},$$

and $P_n(k)$ and $E_n(k)$ can be calculated as

$$P_n(k) = P \left[R_n > \frac{w_n(l_{n-1} + k) \cdot T}{T - \tau(l_{n-1} + k)} \right]$$

$$E_n(k) = E \left[R_n \middle| R_n > \frac{w_n(l_{n-1} + k) \cdot T}{T - \tau(l_{n-1} + k)} \right] \cdot \frac{T - \tau(l_{n-1} + k)}{T}.$$

Proof. To calculate $w_{n-1}(l_{n-1})$, we use n to substitute $n-1$ in (3.9) and take expectation on both sides of (3.8) as

$$w_{n-1}(l_{n-1}) = E\left[\max\{Y_n(\mathbf{x}_1,\ldots,\mathbf{x}_{n-1},\mathbf{X}_n), w_n(\mathbf{x}_1,\ldots,\mathbf{x}_{n-1},\mathbf{X}_n)\}\right], \qquad (3.14)$$

where the expectation is taken with respect to $\mathbf{X}_n$. We further take its conditional expectation with respect to K_n and write it as

$$w_{n-1}(l_{n-1}) = \sum_{k\in\Omega_n(l_{n-1})} P[K_n = k]\cdot q_n(l_{n-1},k),$$

where $q_n(l_{n-1},k)$ is the conditional expectation of (3.14) given $K_n = k$. As we showed in Section 3.5, K_n has a geometric distribution with parameter $p_{s,n}$. Hence we have $P[K_n = k] = (1-p_{s,n})^{k-1}p_{s,n}$. On the other hand, combing (3.2) and Lemma 3.4.1, we have

$$q_{n-1}(l_{n-1},k) = E\left[\max\left\{\frac{R_n\left(T-\tau(l_{n-1}+k)\right)}{T}, w_n(l_{n-1}+k)\right\}\right].$$

Now if we take its conditional expectation with respect to the following event

$$\left\{\frac{R_n\left(T-\tau(l_{n-1}+k)\right)}{T} > w_n(l_{n-1}+k)\right\},$$

we can immediately have (3.13). This proves the theorem. □

We can also bound the computational complexity of the procedure described in Theorem 3.4.1.

Proposition 3.4.1. *To calculate the optimal network throughput w_0 and the expected reward based on a set of given observations $\{r_1,k_1;\ldots;r_M,k_M\}$ with a relative error less than ϵ where $0<\epsilon\ll 1$, the computational complexity of the procedure in*

Theorem 3.4.1 is

$$\min\left\{M\left\lceil T/\tau\right\rceil, \sum_{n=1}^{M} n\left\lceil \frac{\log\frac{\epsilon}{1+\epsilon}}{\log(1-p_{s,n})}\right\rceil\right\}. \tag{3.15}$$

Proof. For a network with size M, the backward induction procedure in Theorem 3.4.1 has M stages. In the n-th stage, the procedure involves calculation of all possible $w_n(l_n)$. For the CAT problem, l_n can simply be bounded as $1 \leq l_n \leq \lceil T/\tau\rceil$. Hence the computational complexity in the n-th stage is at most $\lceil T/\tau\rceil$, and the total computational complexity of the backward induction procedure is at most $M\lceil T/\tau\rceil$.

On the other hand, since $q_n(l_{n-1}, k)$ is the conditional expected reward if the probing duration is $K_n = k$ at time n, $q_n(l_{n-1}, k)$ is a decreasing function of k. For a given integer k_ϵ, we have

$$\frac{\sum_{k>k_\epsilon} P[K_n = k]\cdot q_n(l_{n-1}, k)}{\sum_{k\leq k_\epsilon} P[K_n = k]\cdot q_n(l_{n-1}, k)} < \frac{\sum_{k>k_\epsilon} P[K_n = k]\cdot q_n(l_{n-1}, k_\epsilon)}{\sum_{k\leq k_\epsilon} P[K_n = k]\cdot q_n(l_{n-1}, k_\epsilon)} = \frac{1}{1-(1-p_{s,n})^{k_\epsilon}} - 1, \tag{3.16}$$

where we used the fact that K_n has a geometric distribution Geom($p_{s,n}$) with parameter $p_{s,n}$. To ensure the relative error in the calculation of $w_{n-1}(l_{n-1})$ is less than ϵ, we let the right hand side of (3.16) be less than ϵ. After some manipulation, we have $k_\epsilon \geq \frac{\log\frac{\epsilon}{1+\epsilon}}{\log(1-p_{s,n})}$. Hence we only need to iterate $\left\lceil\frac{\log\frac{\epsilon}{1+\epsilon}}{\log(1-p_{s,n})}\right\rceil$ items in the n-th stage. Iterating this procedure from the top level $n = 0$ to $n = M$ and noticing that $l_0 = 0$, we immediately have our conclusion. □

3.5 An Approximation for Performance Analysis: the Infinite Horizon Approach

As we can see in Section 3.4, the computational complexity of backward induction can quickly become overwhelming as M increases. In contrast, the infinite horizon analysis based on the optimality equation [55–57] has a much smaller computational complexity. Hence we would like to see if the performance analysis in Section 3.4 can be approximated using the infinite horizon approach. In this section, we analyze the protocol using the infinite horizon approach and develop a metric as a guideline to choose the appropriate approach for a given network.

Lemma 3.5.1. *For the same stopping rule problem described in Section 3.3, the infinite horizon analysis yields an optimal network throughput slightly larger than that from the finite horizon analysis. The gap decreases to* 0 *as the network size* $M \to \infty$.

If the network size M is large enough, this problem does not have a finite horizon and can be analyzed using the optimality equation [55–57]. We make the following assumptions:

[A1] The total number of links M in the network is large enough;

[A2] The channel rates only take values in $(0, +\infty)$;

[A3] Each link m probes the wireless medium with probability $p^{(m)} = p$;

[A4] The channel rates for all links have the same cumulative distribution function (CDF) $F_R(r)$.

Here [A1] ensures the problem does not have a finite horizon, and [A2]-[A4] make our mathematical analysis tractable. To analyze this problem, we first characterize the distribution of K_n. By the n-th round, $n-1$ links in total have given up their opportunities in previous rounds. Hence only the rest of the $M-n+1$ links can contribute to an effective channel probing. If we ignore the events when the channel is won by any of these $n-1$ links, K_n has a geometric distribution Geom($p_{s,n}$) with parameter $p_{s,n}$, where

$$p_{s,n} = (M-n+1)\cdot p(1-p)^{M-1} \tag{3.17}$$

is the successful probing probability in the n-th round. To better explain our results, we introduce some notations that will be used frequently in our proof. We define a sequence of parameters $f_n \triangleq \frac{p_{s,n}}{p_{s,1}} = \frac{M-n+1}{M}$ and a sequence of random variables $\tilde{K}_n = f_n K_n$. Since f_n is a constant, $\tilde{K}_n$ also has a geometric distribution with mean $E[\tilde{K}_n] = f_n E[K_n] = E[K_1]$. Hence $\tilde{K}_n$ and K_1 can be considered *equal in distribution* [62, 63].

Theorem 3.5.1. *The average network throughput λ_O^* of Figure 3.1 is the solution of*

$$E\left[1 + \frac{M(M+1)}{(M+0.5)^2}\cdot\frac{\tau\tilde{K}_1}{T} - \frac{\lambda}{R_0}\right]^+ = \frac{M(M+1)}{(M+0.5)^2}\cdot\frac{\tau}{Tp_{s,1}}, \tag{3.18}$$

where $E[\cdot]^+$ is defined as $E[X]^+ = E[\max(X,0)]$. The optimal stopping rule is

$$N^* = \min\left\{n \geq 1 : R_n \geq \lambda_n^* \cdot \frac{T}{T - \tau\sum_{i=1}^{n} K_i}\right\}, \tag{3.19}$$

where λ_n^* is the solution of

$$E\left[1-\frac{\tau}{T}\left\{\sum_{i=1}^{n}K_i-\frac{M(M+n+1)}{(M+0.5)^2}\tilde{K}_1\right\}-\frac{\lambda}{R_n}\right]^+=\frac{M(M+n+1)}{(M+0.5)^2}\cdot\frac{\tau}{Tp_{s,1}}. \tag{3.20}$$

Proof. By [A2], we can rewrite the network throughput (3.2) at time n as $Y_n = \frac{T-\tau\sum_{i=1}^{n}K_i}{T/R_n}$. This problem can be solved as a maximal rate of return problem. For a fixed rate λ, we define a new reward function at time n as

$$V_n(\lambda) = T - \tau\sum_{i=1}^{n}K_i - \frac{\lambda T}{R_n}. \tag{3.21}$$

The problem is then to characterize the optimal rate λ_n^* and the stopping rule to achieve λ_n^*. First, we need to show the existence of the optimal stopping rule. We notice that $E\{\sup_n V_n(\lambda)\} < T < \infty$. On the other hand, we can easily see that $\limsup_{n\to\infty} V_n(\lambda) \to -\infty$ and $V_n(\lambda) \to -\infty$ *a.s.*. Putting them together leads to $\limsup_{n\to\infty} V_n(\lambda) \to V_\infty(\lambda)$ *a.s.*. Hence an optimal stopping rule exists and can be described by the optimality equation. By the definition of $\tilde{K}_n$, we notice that $K_i = \frac{M}{M-i+1}\tilde{K}_i$. Substituting it into (3.21) and using the *i.i.d.* property of $\tilde{K}_i$, we have

$$V_n(\lambda) = T - \tau\sum_{i=1}^{n}\frac{M}{M-i+1}\tilde{K}_i - \frac{\lambda T}{R_n} = T - \tau\tilde{K}_1\sum_{i=1}^{n}\frac{M}{M-i+1} - \frac{\lambda T}{R_n}.$$

Note that the above equation holds *in distribution*. Since the network size M is large enough and the problem can be solved as an infinite-horizon problem, the number of rounds n is usually much smaller compared to M. To calculate the above summation, we approximate $\frac{M}{M-i+1} + \frac{M}{M-(n+1-i)+1}$ as $\frac{2M}{M-n/2+0.5}$. By repeating this procedure for all $i \leq n/2$, we can approximate $V_n(\lambda)$ as

$$V_n(\lambda) \approx T - \tau \tilde{K}_1 \cdot \frac{Mn}{M - n/2 + 0.5} - \frac{\lambda T}{R_n}.$$

Similarly, the payoff at time $n+1$ can be written as

$$V_{n+1}(\lambda) \approx T - \tau \tilde{K}_1 \cdot \frac{M(n+1)}{M - (n+1)/2 + 0.5} - \frac{\lambda T}{R_{n+1}}.$$

Meanwhile, note that R_n are *i.i.d.* by [A4]. Hence in the sense of distribution the difference between $V_n(\lambda)$ and $V_{n+1}(\lambda)$ can be written as

$$\Delta V_n(\lambda) = V_{n+1}(\lambda) - V_n(\lambda) = -\tau \tilde{K}_1 \cdot \frac{M}{M+0.5} \left[\frac{n+1}{1 - \frac{(n+1)/2}{M+0.5}} - \frac{n}{1 - \frac{n/2}{M+0.5}} \right].$$

By [A1], we can approximate the item in the above square bracket as

$$(n+1)\left\{1 + \frac{\frac{n+1}{2}}{M+0.5}\right\} - n\left\{1 + \frac{\frac{n}{2}}{M+0.5}\right\} = \frac{M+n+1}{M+0.5}. \tag{3.22}$$

Substituting it into the optimality equation $V_n^* = \max\{Y_n, E(V_{n+1}^* | \mathcal{F}_n)\}$ [55–57], we have

$$V_n^*(\lambda) = E\left[\max\left\{T - \tau \sum_{i=1}^{n} K_i - \frac{\lambda T}{R_n}, V_n^*(\lambda) - \tau \tilde{K}_1 \cdot \frac{M(M+n+1)}{(M+0.5)^2}\right\}\right].$$

According to optimal stopping theory [55–57], the optimal rate λ_n^* that maximizes the rate of return should yield 0 for (3.21). If we substitute $V_n^*(\lambda_n^*) = 0$ into the above equation and note that $E[\tilde{K}_1] = 1/p_{s,1}$, we can rewrite the equation as (3.20). The uniqueness of λ_n^* can be easily verified. The optimal stopping rule can be written as

$$N^* = \min\left\{n \geq 1 : T - \tau \sum_{i=1}^{n} K_i - \frac{\lambda_n^* T}{R_n} \geq V_n^*(\lambda_n^*) = 0\right\},$$

which leads to (3.19) after some manipulation. The optimal network throughput is the expected rate of return before taking any observations. Hence we get the optimal network throughput λ_O^* if we let $n = 0$ in (3.20), which immediately yields (3.18). □

The optimal network throughput (3.18) can be further simplified under certain conditions.

Proposition 3.5.1. *Assume $\tau \ll T$, the network throughput λ_O^* of Figure 3.1 can be approximated as the solution of*

$$E\left[1 - \frac{\lambda}{R_0}\right]^+ = \frac{M(M+1)}{(M+0.5)^2} \cdot \frac{\tau}{Tp_{s,1}}. \tag{3.23}$$

Proof. By [A1], we have $\frac{M(M+1)}{(M+0.5)^2} \approx 1$. Since $\frac{\tau}{T} \ll 1$, the term $\frac{\tau}{T} \cdot \frac{M(M+1)}{(M+0.5)^2}\tilde{K}_1$ can be ignored compared to 1 on the left hand of (3.18). This completes the proof. □

An immediate question following Lemma 3.5.1 and Theorem 3.5.1 is: how good is the approximation compared to the rigorous analysis in Section 3.4, in particular for networks at a finite size M? In fact, we prefer to design the stopping rule based on the analytical results from Theorem 3.5.1 due to their low computational complexity even for a finite network size M. What will the actual achievable network throughput be like?

To answer these questions, we present one metric which serves as a guideline when we decide whether or not we could use the infinite horizon analysis. For a given network, if the problem can be treated as in Section 3.5, in a probabilistic sense the optimal stopping time N^* should be much smaller than the network size

M. Hence one necessary condition is that the probability $P[N^* > M]$ should be small enough.

Theorem 3.5.2. *For a network with size M, suppose the infinite horizon analysis in Theorem 3.5.1 yields a sequence of optimal expected network throughputs λ_n^* for Figure 3.1. If $\tau \ll T$, the probability $P[N^* > M]$ can be approximated as*

$$P[N^* > M] \approx \prod_{n=1}^{M} F_R(\lambda_n^*). \tag{3.24}$$

If this probability is not small enough, it is not recommended to design stopping rules based on the infinite horizon analysis.

Proof. For a given integer $k > 0$, we have

$$P\left[N^* > k\right] = P\left[\min\left\{n \geq 1 : R_n \geq \lambda_n^* \cdot \frac{T}{T - \tau \sum_{i=1}^{n} K_i}\right\} > k\right]. \tag{3.25}$$

Since $\tau \ll T$ and the optimal stopping time N^* is much smaller than M, we consider $\frac{\tau}{T} \cdot \sum_{i=1}^{n} K_i \ll 1$ for approximation. Substituting it into (3.25), we have

$$\begin{aligned} P\left[N^* > k\right] &\approx P\left[\min\left\{n \geq 1 : R_n \geq \lambda_n^*\right\} > k\right] \\ &= \prod_{n \leq k} P\left[R_n < \lambda_n^*\right], \end{aligned} \tag{3.26}$$

where we used the fact that R_n are *i.i.d.*. To get (3.25), simply let $k = M$ in (3.26). $\square$

On the other hand, if $P[N^* > M]$ is not small enough, it implies that the stopping rule problem cannot be treated as an infinite horizon problem. In this case, if we design a stopping rule based on Theorem 3.5.1 nevertheless, the procedure will quickly reach the last stage and be forced to stop then. In this case, the actually achieved network throughput is generally not optimal.

Theorem 3.5.3. *Suppose the infinite horizon analysis yields a sequence of λ_n^* for a network with size M. Suppose we design a stopping rule $\hat{N}$ based on these rates and (3.19). If $\tau \ll T$, the achievable network throughput based on $\hat{N}$ is*

$$\hat{\lambda}^* = \sum_{n=1}^{M} E\left[R_n | R_n \geq \lambda_n^*\right] \frac{T - \tau \sum_{i=1}^{n} 1/p_{s,i}}{T} \times [1 - F_R(\lambda_n^*)] \prod_{i=1}^{n-1} F_R(\lambda_i^*). \tag{3.27}$$

Proof. According to the stopping rule (3.19), the expected reward can be written as

$$\hat{\lambda}^* = \sum_{n=1}^{M} E\left[Y_n(\mathbf{X}_1, \ldots, \mathbf{X}_n) \cdot P(N = n | \mathbf{X}_1, \ldots, \mathbf{X}_n)\right]. \tag{3.28}$$

The condition to stop at time n is $R_n \geq \lambda_n^* \cdot \frac{T}{T - \tau \sum_{i=1}^{n} K_i}$. When $\tau \ll T$, this condition can be simplified as $R_n \geq \lambda_n^*$. Hence the expected reward at time n can be written as a conditional expectation, i.e.

$$E\left[R_n \cdot \frac{T - \tau \sum_{i=1}^{n} K_i}{T} \middle| R_n \geq \lambda_n^*\right] = E\left[R_n | R_n \geq \lambda_n^*\right] \cdot \frac{T - \tau \sum_{i=1}^{n} 1/p_{s,i}}{T},$$

where we used the fact that K_i has a geometric distribution $\mathrm{Geom}(p_{s,i})$ and is independent from R_n. On the other hand, by (3.26) the probability to stop at time n can be approximated as

$$P(N = n | \mathbf{X}_1, \ldots, \mathbf{X}_n) = \prod_{i=1}^{n-1} F_R(\lambda_i^*) - \prod_{i=1}^{n} F_R(\lambda_i^*) = [1 - F_R(\lambda_n^*)] \prod_{i=1}^{n-1} F_R(\lambda_i^*). \tag{3.29}$$

Substituting it into (3.28) together with the expected reward at time n, we have (3.27). □

3.6 An Energy Efficient Improvement of the Protocol

In this section, we present an improved distributed opportunistic scheduling protocol, which is directly motivated by the concept of effective observation points introduced in Section 3.3.

According to (3.23), the network throughput λ_O^* decreases as the successful probing probability $p_{s,1}$ decreases. Hence to improve the network throughputs, we need to improve $p_{s,1}$. For a given network with size M, we can first tune the parameter p to maximize $p_{s,1}$. To do this, we let $n = 1$ in (3.17) and take the first-order derivative as $\frac{\partial p_{s,1}}{\partial p} = M(1-p)^{M-2}(1-Mp) = 0$. The non-trivial solution in $(0, 1)$ is $p^* = 1/M$. Hence to maximize $p_{s,1}$, on average there is exactly $Mp^* = 1$ link probing the channel. The maximal successful probing probability is $p_{s,1} = \frac{1}{1-\frac{1}{M}} \cdot \left(1 - \frac{1}{M}\right)^M$ for $M \geq 2$, which is a decreasing function of M. Hence the optimal throughput λ_O^* is a decreasing function as M increases. From the perspective of channel probing costs, a smaller M is preferred for better system performance. On the other hand, from (3.5) we can see that if link m ever gives up the current opportunity, m will always repeat the same decision in the current block. Hence if link m ever decides to send data in the current block, it should happen when m wins the channel for the first time. If after that m still contends the medium, it would not lead to an effective decision, and meanwhile it lowers the successful probability $p_{s,n}$. Based on this observation, we have an improved protocol as shown in Figure 3.2 [53, 54].

Suppose at time n the set of active probing links is $\mathcal{M}_n$. This is the set of links whose current state is TRUE in Figure 3.2. Denoting its cardinality as $M_n \triangleq \|\mathcal{M}_n\|$, we have $M_n = M - n + 1$ following line 9 of Figure 3.2. We can see that M_n is decreasing as time n moves on. The shrinking of $\mathcal{M}_n$ is an important feature of the improved protocol. It not only reduces the probing costs, but also ensures the winner s_n is different at each time n. Hence the winners' rates R_n are independent

Link m sets its state as TRUE, where $m = 1, \ldots, M$;

for each link m whose state is TRUE **do**

  m probes the channel with a fixed probability $p^{(m)}$;

  if m wins the channel **then**

    m makes a decision on whether or not to send data over the channel;

    if m decides to utilize the channel **then**

      m sends data through the channel for a duration of $T - \sum_{i=1}^{n} T_i$ (CAT) or T (CDT), where n is the current index of channel probing;

    else

      m sets its state as FALSE;

    end if

  end if

end for

Figure 3.2: The improved distributed opportunistic scheduling protocol

in Figure 3.2. At time n, the successful probing probability can be written as

$$p_{s,n} = M_n p(1-p)^{M_n - 1}. \tag{3.30}$$

We now characterize the performance of the improved protocol shown in Figure 3.2. First of all, the finite horizon analyses described in Section 3.4 can be applied in a similar way here. The computational complexity can also be estimated similarly. The only difference is that the successful probing probability $p_{s,n}$ in (3.12) should be calculated according to (3.30). Now we analyze this problem assuming that it can be treated as an infinite horizon problem. By [A1], we can approximate the successful probing probability (3.30) as

$$p_{s,n} \approx Mp(1-p)^{M_n-1} = Mp(1-p)^{M-n}. \tag{3.31}$$

We can see that $p_{s,1} < p_{s,2} < \ldots < p_{s,n}$. Similar to Theorem 3.5.1, we introduce a sequence of parameters $g_n \triangleq \frac{p_{s,n}}{p_{s,1}} = (1-p)^{-(n-1)}$ and a sequence of random variables $\tilde{K}_n = g_n K_n$. It is easy to verify that $\tilde{K}_n$ and K_1 can be considered *equal in distribution* and thus $\{\tilde{K}_n\}$ are *i.i.d.*.

Theorem 3.6.1. *The network throughput λ_P^* of Figure 3.2 is the solution of*

$$E\left[1 + \frac{\tau}{T} \cdot (1-p)^2 \tilde{K}_1 - \frac{\lambda}{R_0}\right]^+ = (1-p)^2 \frac{\tau}{T p_{s,1}}. \tag{3.32}$$

The optimal stopping rule is

$$N^* = \min\left\{n \geq 1 : R_n \geq \lambda_n^* \cdot \frac{T}{T - \tau \sum_{i=1}^n K_i}\right\}, \tag{3.33}$$

where λ_n^* is the solution of

$$E\left[1 - \frac{\tau}{T}\left\{\sum_{i=1}^n K_i - (1-p)^{n+1}\tilde{K}_{n+1}\right\} - \frac{\lambda}{R_n}\right]^+ = (1-p)^{n+1}\frac{\tau}{T p_{s,1}}. \tag{3.34}$$

Proof. We use V_n defined in (3.21) in our proof. The existence of the optimal stopping rule can be verified in the same way as in Theorem 3.5.1. To compute the optimal reward V_n^*, we take a look at the reward after l steps since time n. By the definition of g_n, we can write $K_n = (1-p)^{n-1}\tilde{K}_n$. Substituting it into (3.21), we have

$$V_{n+l}(\lambda) = T - \tau \sum_{i=1}^n (1-p)^{i-1}\tilde{K}_i - \left[\tau \sum_{i=n+1}^{n+l} (1-p)^{i-1}\tilde{K}_i + \frac{\lambda T}{R_{n+l}}\right].$$

If we start from time $n+1$, the reward after l rounds is

$$V_{n+l+1}(\lambda) = T - \tau \sum_{i=1}^{n} (1-p)^{i-1} \tilde{K}_i - \tau (1-p)^n \tilde{K}_{n+1} - \left[\tau \sum_{i=n+2}^{n+l+1} (1-p)^{i-1} \tilde{K}_i + \frac{\lambda T}{R_{n+l+1}} \right].$$

The item in the above square bracket is the recursive part for l rounds of observations since time $n+1$. We can rewrite it as

$$(1-p) \left\{ \tau \sum_{i=n+1}^{n+l} (1-p)^{i-1} \tilde{K}_{i+1} + \frac{\lambda T}{R_{n+l+1}} \right\} + p \cdot \frac{\lambda T}{R_{n+l+1}}.$$

By [A1], p should be reasonably small; otherwise the average number of probing links Mp will be much larger than 1, leading to increased probing costs. Hence we can ignore the last term and write the optimality equation as

$$V_n^*(\lambda) = E \left[\max \left\{ T - \tau \sum_{i=1}^{n} K_i - \frac{\lambda T}{R_n}, (1-p) \left(V_n^*(\lambda) - \tau K_{n+1} \right) \right\} \right].$$

Again, the optimal reward λ_n^* that maximizes the rate of return must satisfy $V_n^*(\lambda_n^*) = 0$. We substitute this into the optimality equation and rewrite it as

$$E \left[1 - \frac{\tau}{T} \left\{ \sum_{i=1}^{n} K_i - (1-p) K_{n+1} \right\} - \frac{\lambda_n^*}{R_n} \right]^+ = (1-p) \cdot \frac{\tau}{T} E[K_{n+1}].$$

If we further notice that $K_{n+1} = 1/g_{n+1} \tilde{K}_{n+1} = (1-p)^n \tilde{K}_{n+1}$ and that $\tilde{K}_{n+1}$ and K_1 are *i.i.d.*, we can rewrite the above equation as (3.34). The optimal stopping rule N^* can be derived in the same way as in Theorem 3.5.1. To get the optimal system throughput λ_P^*, we let $n = 0$ in (3.34) and rewrite the equation as (3.32). □

Similar to Section 3.5, we further simplify the network throughput as Proposition 3.6.1 if $\tau \ll T$. Based on this, we show that the modified protocol improves the network throughput as in Proposition 3.6.2. The proofs are straight forward and are skipped due to space limitations.

Proposition 3.6.1. *If $\tau \ll T$, the network throughput λ_P^* can be approximated as the solution of*

$$E\left[1 - \frac{\lambda}{R_0}\right]^+ = (1-p)^2 \cdot \frac{\tau}{T p_{s,1}}. \tag{3.35}$$

Proposition 3.6.2. *The improved protocol in Figure 3.2 yields a higher network throughput compared to the protocol in Figure 3.1, i.e. $\lambda_P^* > \lambda_O^*$.*

In the improved protocol any link who decides to give up the current opportunity for data transmission will not probe the channel anymore until the beginning of the next block. Hence these links can temporarily switch to a sleep mode until the beginning of the next block and reduce the energy used for channel probing. This could be very useful for mobile ad-hoc or sensor networks where most of their mobile nodes have limited battery life.

Similar to the analyses for throughputs, we focus on the total energy savings for all links in the channel probing phase, not for a specific link. Suppose each probing signal consumes roughly a constant energy of c. Then the energy consumed during the channel probing phase can be written as $c\sum_{i=1}^N Z_i$, where Z_i is the number of probing signals sent during the i-th round of channel probing, and N is the stopping time associated with the stopping rule. Hence the average energy spent during the channel probing phase is $z = cE\left[\sum_{i=1}^N Z_i\right]$. Using the law of total expectation, we can write

$$z = cE\left[E\left[\sum_{i=1}^N Z_i \middle| N\right]\right] = c\sum_n P[N=n]\sum_{i=1}^n E[Z_i]. \tag{3.36}$$

Theorem 3.6.2. *The average energy consumed for the channel probing of Figure*

3.1 can be written as

$$z_O = c \sum_n P[N_O^* = n] \cdot \frac{1}{(1-p)^{M-1}} \sum_{i=1}^{n} \frac{M}{M-i+1}, \tag{3.37}$$

where N_O^* is the optimal stopping rule for Figure 3.1, and the average probing energy of Figure 3.2 can be written as

$$z_P = c \sum_n P[N_P^* = n] \cdot \frac{1}{(1-p)^{M-1}} \frac{1-(1-p)^n}{1-(1-p)}, \tag{3.38}$$

where N_P^* is the optimal stopping rule for Figure 3.2.

Proof. As we mentioned in Section 3.3, we will use the notation of $\tilde{\mathcal{F}}_n$ in the proof. For the protocol in Figure 3.1, in the i-th round there are a total of $\tilde{K}_i$ probings, each of which has a duration of τ and on average Mp links sending probing signals. Hence there are on average $E[Z_i] = Mp \cdot E[\tilde{K}_i]$ probing signals sent in the i-th round. Hence we can write

$$E[Z_i] = Mp \cdot \frac{1}{(M-i+1)p(1-p)^{M-1}} = \frac{1}{(1-p)^{M-1}} \cdot \frac{M}{M-i+1}.$$

Substituting the above equation into (3.36), we can immediately have (3.37).

On the other hand, for the improved protocol in Figure 3.2, in the i-th round there are a total of $\tilde{K}_i$ probings, and each of them has on average $(M-i+1)p$ links sending probing signals. This is because in the improved protocol once a link gives up its opportunity, he would not probe again until the beginning of the next block. Hence we can write

$$\sum_{i=1}^{n} E[Z_i] = \sum_{i=1}^{n} (M-i+1)p \cdot \frac{1}{(M-i+1)p(1-p)^{M-i}} = \frac{1}{(1-p)^{M-1}} \cdot \frac{1-(1-p)^n}{1-(1-p)}.$$

Combining the above equation with (3.36), we have (3.38). □

In Theorem 3.6.2, the probability of $P[N^* = n]$ can be approximated in the same way as (3.29).

3.7 The Constant Data Time Problem

Our analyses in Section 3.4, 3.5 and 3.6 can be applied to the CDT problem in a similar way. In the CDT problem [35, 59, 60], the transmitter has a fixed duration $T_d = T$ for data transmission, regardless of the duration T_p elapsed for channel probing. The normalized network throughput to utilize the channel at the end of the n-th round is

$$Y_n = \frac{R_n \cdot T}{T + \sum_{i=1}^{n} T_i}. \tag{3.39}$$

We list the analytical results for the CDT problem in this section and compare its numerical results to that of the CAT problem in Section 3.8.

First of all, due to the block fading assumption, the CDT problem also has an implicit horizon at M. Hence the CDT problem for the original protocol in Figure 3.1 or the improved protocol in Figure 3.2 should be treated as a finite horizon problem.

Theorem 3.7.1. *The network throughput of the CDT problem based on backward induction is $w_0 = \lambda_0^*$, and the optimal stopping rule is*

$$N^* = \min\left\{ n \geq 1 : R_n \geq \lambda_n^* \cdot \left(1 + \frac{\tau}{T}\sum_{i=1}^{n} K_i\right)\right\}. \tag{3.40}$$

The finite horizon analysis reduces to the calculation of w_0, which eventually iterates all $w_n(l_n)$ for $n = 1, \ldots, M-1$ and $w_M(r_M, l_M)$. The expected reward can be *calculated recursively as*

$$w_{n-1}(l_{n-1}) = \sum_{k \in \mathbb{N}} (1 - p_{s,n})^{k-1} p_{s,n} \cdot q_n(l_{n-1}, k) \tag{3.41}$$

$$q_n(l_{n-1}, k) = P_n(k) \cdot E_n(k) + [1 - P_n(k)] \cdot w_n(l_{n-1} + k), \tag{3.42}$$

where $q_n(l_{n-1}, k)$ *is the conditional expected reward given* $K_n = k$*, and* $P_n(k)$ *and* $E_n(k)$ *can be calculated as*

$$\begin{aligned} P_n(k) &= P\left[R_n > w_n(l_{n-1} + k) \cdot \frac{T + \tau(l_{n-1} + k)}{T}\right] \\ E_n(k) &= E\left[R_n \middle| R_n > w_n(l_{n-1} + k) \cdot \frac{T + \tau(l_{n-1} + k)}{T}\right] \times \frac{T}{T + \tau(l_{n-1} + k)}. \end{aligned}$$

Proposition 3.7.1. *For the CDT problem, to calculate the optimal network throughput* w_0 *and the expected reward for given observations* $\{r_1, k_1; \ldots; r_M, k_M\}$ *with a relative error less than* ϵ *where* $0 < \epsilon \ll 1$*, the computational complexity of the procedure in Theorem 3.7.1 is at most* $\sum_{n=1}^{M} n \left\lceil \frac{\log \frac{\epsilon}{1+\epsilon}}{\log(1-p_{s,n})} \right\rceil$.

Similar to Section 3.5, we prefer to analyze the CDT problem using the infinite horizon approach when it can yield a good approximation to the finite horizon approach. The proof of Theorem 3.7.2 can be found in Appendix B.1.

Theorem 3.7.2. *The average network throughput* λ_O^* *of the CDT problem is the solution of*

$$E\left[\frac{R_0}{\lambda} + \frac{M(M+1)}{(M+0.5)^2} \cdot \frac{\tau \tilde{K}_1}{T} - 1\right]^+ = \frac{M(M+1)}{(M+0.5)^2} \cdot \frac{\tau}{T p_{s,1}}. \tag{3.43}$$

The optimal stopping rule N^* *is*

$$N^* = \min\left\{n \geq 1 : R_n \geq \lambda_n^* \cdot \left(1 + \frac{\tau}{T} \sum_{i=1}^{n} K_i\right)\right\}, \tag{3.44}$$

and λ_n^* is the solution of

$$E\left[\frac{R_n}{\lambda} - \frac{\tau}{T}\left\{\sum_{i=1}^{n} K_i - \frac{M(M+n+1)}{(M+0.5)^2}\tilde{K}_1\right\} - 1\right]^+ = \frac{M(M+n+1)}{(M+0.5)^2} \cdot \frac{\tau}{Tp_{s,1}}. \tag{3.45}$$

Proposition 3.7.2. *Assume $\tau \ll T$, the network throughput λ_O^* for the CDT problem can be approximated as the solution of*

$$E\left[\frac{R_0}{\lambda} - 1\right]^+ = \frac{M(M+1)}{(M+0.5)^2} \cdot \frac{\tau}{Tp_{s,1}}. \tag{3.46}$$

In one block, the available duration for data transmission is $T - \tau\sum_{i=1}^{n} K_i$ for the CAT problem and T for the CDT problem respectively. Hence intuitively the protocol in Figure 3.1 should yield a higher network throughput for the CDT model.

Proposition 3.7.3. *Denote λ_{CAT}^* and λ_{CDT}^* as the optimal network throughput for the CAT and CDT problem respectively, we have $\lambda_{CAT}^* < \lambda_{CDT}^*$.*

Proof. For any $r > \lambda_{CAT}^*$, we have $1 - \frac{\lambda *_{CAT}}{r} < \frac{r}{\lambda_{CAT}^*} - 1$. By taking integration on both sides of the inequality, we have

$$E\left[\frac{R_0}{\lambda_{CAT}^*} - 1\right]^+ > E\left[1 - \frac{\lambda_{CAT}^*}{R_0}\right]^+ = \frac{M(M+1)}{(M+0.5)^2} \cdot \frac{\tau}{Tp_{s,1}} = E\left[\frac{R_0}{\lambda_{CDT}^*} - 1\right]^+,$$

where the first and second equality is from Proposition 3.5.1 and Proposition 3.7.2 respectively. If we compare the first and last item in the above inequality, we have $\lambda_{CAT}^* < \lambda_{CDT}^*$. □

Theorem 3.7.3. *Suppose the infinite horizon analysis yields a sequence of rates λ_n^* for a network of size M. If $\tau \ll T$, the probability $P[N^* > M]$ can be approximated* *as*

$$P[N^* > M] \approx \prod_{n=1}^{M} F_R(\lambda_n^*). \tag{3.47}$$

If this probability is not small enough, it is not recommended to design stopping rules based on the infinite horizon analysis. Otherwise if we use the stopping rule based on these rates and (3.44), the achievable network throughput is

$$\hat{\lambda}^* = \sum_{n=1}^{M} E\left[R_n | R_n \geq \lambda_n^*\right] \frac{T}{T + \tau \sum_{i=1}^{n} 1/p_{s,i}} \times [1 - F_R(\lambda_n^*)] \prod_{i=1}^{n-1} F_R(\lambda_i^*). \tag{3.48}$$

For the improved protocol shown in Figure 3.2, the performance for the CDT problem can be shown in Theorem 3.7.4. The proof of Theorem 3.7.4 can be found in Appendix B.2.

Theorem 3.7.4. *The network throughput λ_P^* of Figure 3.2 for the CDT problem is the solution of*

$$E\left[\frac{R_0}{\lambda} + \frac{\tau}{T} \cdot (1-p)^2 \tilde{K}_1 - 1\right]^+ = (1-p)^2 \frac{\tau}{T p_{s,1}}. \tag{3.49}$$

The optimal stopping rule N^* is

$$N^* = \min\left\{n \geq 1 : R_n \geq \lambda_n^* \cdot \left(1 + \frac{\tau}{T} \sum_{i=1}^{n} K_i\right)\right\}, \tag{3.50}$$

where λ_n^* is the solution of

$$E\left[\frac{R_n}{\lambda} - \frac{\tau}{T}\left\{\sum_{i=1}^{n} K_i - (1-p)^{n+1} \tilde{K}_{n+1}\right\} - 1\right]^+ = (1-p)^{n+1} \frac{\tau}{T p_{s,1}}. \tag{3.51}$$

Proposition 3.7.4. *If $\tau \ll T$, we can approximate the network throughput λ_P^* as the solution of*

$$E\left[\frac{R_0}{\lambda} - 1\right]^+ = (1-p)^2 \frac{\tau}{T p_{s,1}}. \tag{3.52}$$

3.8 Numerical Results

In this section, we show numerical results based on our discussions from Section 3.4 to Section 3.7. We consider an ad-hoc network where the wireless medium is Rayleigh fading within each block. The channel rate can be written as

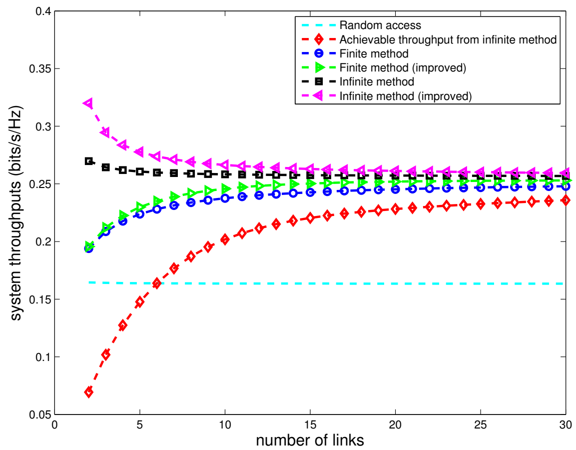
$$R(h) = \log_2(1 + \rho h) \qquad \text{bits/s/Hz},$$

where ρ is the average signal-to-noise ratio (SNR), and h is the channel gain corresponding to Rayleigh fading. Hence the probability density function (PDF) of h can be written as

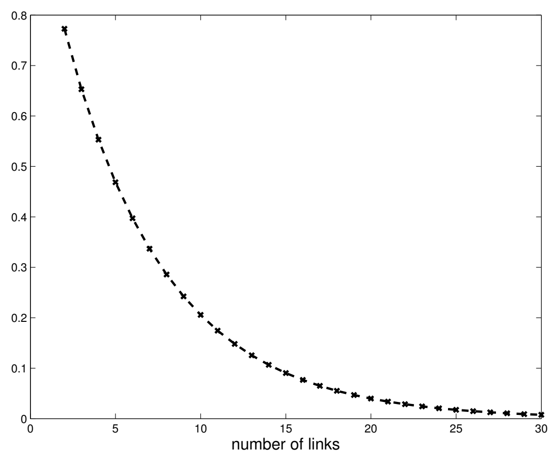
$$f(h) = \frac{h}{\sigma^2} e^{-\frac{h^2}{2\sigma^2}}, \quad h \geq 0.$$

We assume $T = 1$ fixed throughout all simulations in this section. We compare numerical results from the finite horizon and the infinite horizon analyses with various settings of the parameters M, p, τ and ρ. For performance comparison purposes, we also show network throughputs from a pure random access approach, where the *first* winner of the wireless medium always utilizes the channel for data transmission, regardless of the available channel rates.

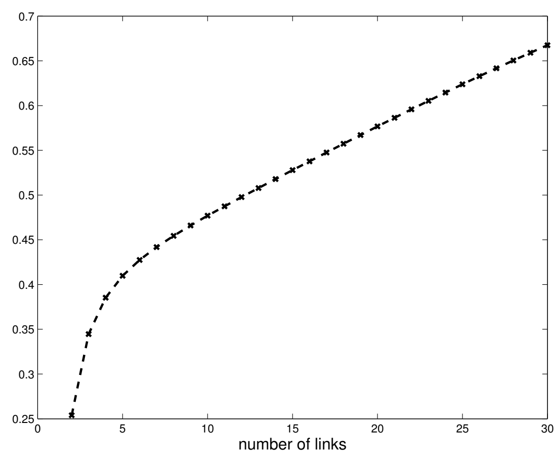
In Figure 3.3, we show numerical results from both the infinite horizon and finite horizon analysis, where the network size is M and other parameters are $p = 1/M$, $\tau = 0.01$, $\rho = -10$dB and $\sigma = 1$. In Figure 3.3(a), the dashed line shows the network throughputs for the pure random access scheme. Clearly we can see that the distributed opportunistic scheduling schemes show a considerable performance improvement, e.g. 57% improvement at a network size $M = 30$ in Figure 3.3(a).



(a)



(b)



(c)

Figure 3.3: Numerical results for ad-hoc networks with M links, where the parameters are $p = 1/M$, $\tau = 0.01$, $\rho = -10$dB and $\sigma = 1$: (a) network throughputs; (b) $P[N^* > M]$; (c) energy savings in probing signals.

On the other hand, we notice that the finite horizon and infinite horizon analyses yield quite different network throughputs, especially when the network size M is not large enough. Figure 3.3(a) shows the network throughputs for the distributed opportunistic scheduling protocol described in Figure 3.1, where the line with "∘" is from the finite horizon analysis and the line with "□" is from the infinite horizon analysis. The network throughputs show opposite trends as the network size M increases in Figure 3.3(a). The network throughput from the infinite horizon analysis decreases while the network throughput from the finite horizon analysis increases. This is because in the infinite horizon analysis, there is enough multiuser diversity to be exploited. In the finite horizon analysis, there is not enough multiuser diversity to be exploited when the network size M is small, which is constrained by the finite horizon. Hence the infinite horizon analysis always shows a larger network throughput than the finite horizon analysis, and the gap between these two lines gradually decreases to 0 as the network size M increases. For example, the two lines show a gap of 8.7% at $M = 10$, and the gap drops to 4.9% at $M = 20$. In Figure 3.3(b), we show the estimated probability $P[N^* > M]$ in Theorem 3.5.2. We can see that $P[N^* > M]$ is as high as 20% at $M = 10$, but drops quickly to 5% at $M = 20$. Hence for a given network, the estimated $P[N^* > M]$ serves as a measure of how well the problem can be treated as an infinite horizon problem. In line with this guideline, the line with "$\diamond$" in Figure 3.3(a) shows the actual achievable rewards based on Theorem 3.5.3 if the stopping rule is designed based on the results from the infinite horizon analysis. To our surprise, the actual reward is much smaller than the one from the infinite horizon analysis. This gap is pretty large when the network size M is not large enough, say $M \leq 20$ in Figure 3.3(a). This observation agrees with the trend of $P[N^* > M]$ in Figure 3.3(b). Hence if the problem is not suitable to be treated as an infinite horizon problem, it is not recommended to design stopping rules based on the infinite horizon analysis; otherwise the actual achievable rewards may deviate a lot from the infinite horizon analysis results for small and medium-size networks.

In addition, Figure 3.3(a) shows the network throughputs for the improved protocol described in Figure 3.2, where the line with "$\triangleright$" is from the finite horizon analysis and the line with "$\triangleleft$" is from the infinite horizon analysis respectively. We can see that the improved protocol always yields a slightly better performance. For example, the line with "$\triangleright$" steadily shows a 2% performance improvement over the line with "$\circ$" based on the finite horizon analysis. This coincides with our theoretical result in Proposition 3.6.2. Even though the performance improvement is not significant, it is still worth mentioning since there is no additional cost in the protocol design of Figure 3.2. This performance improvement can be considered as a "free ride" based on the concept of effective observation points. On the other hand, in Figure 3.3(c) we show the energy savings in probing signals that can be achieved by the improved protocol, where the y-axis is z_P/z_O for each M. We can see that the improved protocol can considerably reduce the total number of probing signals sent in the network. For example, at $M = 30$ the improved protocol only needs 67% of the probing signals sent in the original protocol in Figure 3.1. This results in 33% energy savings for probing signals. Hence with only a simple modification, the improved protocol can slightly improve the network throughputs while saving considerably energy used for probing signals. This is of particular interest for mobile ad-hoc networks or sensor networks where many nodes in the network have limited battery life.

In Figure 3.4(a)-(d), we compare network throughputs with different parameters, where we vary one parameter at a time from the default parameter settings.

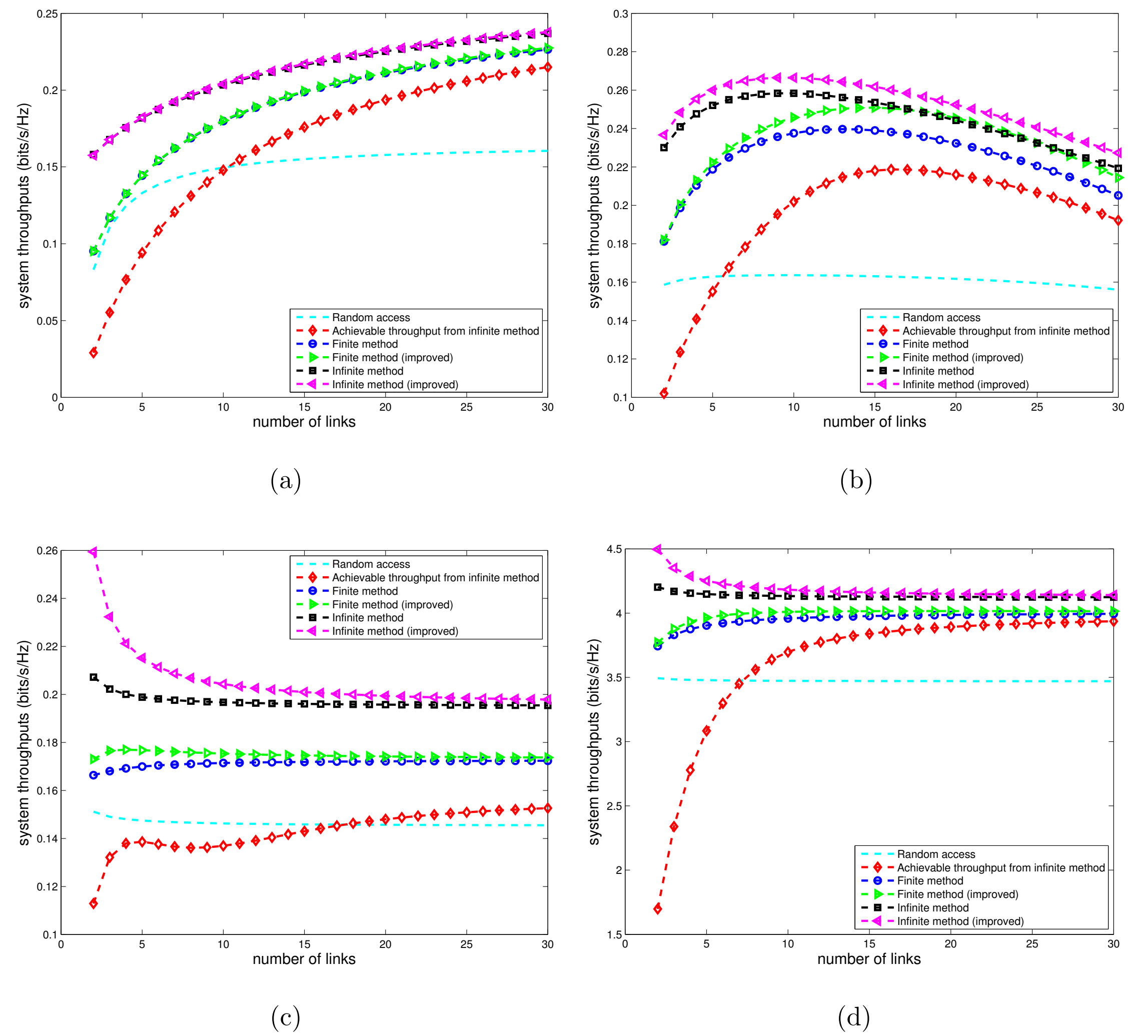
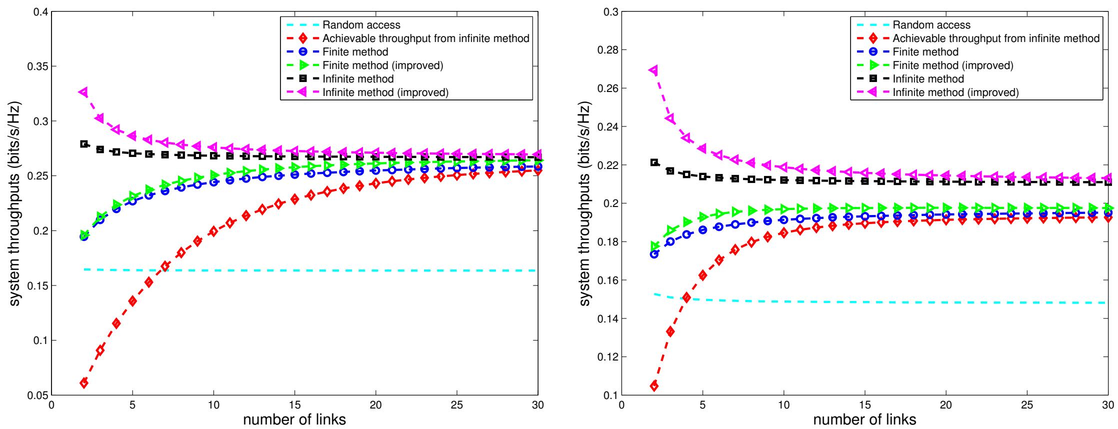


Figure 3.4: Numerical results for ad-hoc networks with M links, where the default parameters are $p = 1/M$, $\tau = 0.01$, $\rho = -10$dB and $\sigma = 1$: (a) network throughputs with $p = 0.01$; (b) network throughputs with $p = 0.1$; (c) network throughputs with $\tau = 0.05$; (d) network throughputs with $\rho = 10$dB.

We first show the network throughputs under two different scenarios for p in Figure 3.4(a) and Figure 3.4(b) respectively. Figure 3.4(a) shows the network throughputs for $p = 0.01$, which represents an "under-probed" scenario since $Mp < 1$. We can see that the protocols yield smaller throughputs compared to Figure 3.4(a). On the other hand, the improved protocol has almost the same performance as the original protocol. In this case, it would not help to reduce the probing costs since the system is already under-probed. Figure 3.4(b) shows the opposite scenario with $p = 0.1$ where the medium is "over-probed" since $Mp > 1$. The network throughputs are also smaller compared to Figure 3.4(a). However, the improved protocol shows a 5% performance improvement compared to the original protocol. Recall that this quantity is 2% in Figure 3.4(a). In this case, it helps to reduce the probing costs since the network is over-probed. In Figure 3.4(c), we show the network throughputs with a larger probing cost $\tau = 0.05$. With larger probing costs the protocols yield smaller network throughputs. Meanwhile, there is a larger gap between the finite horizon and infinite horizon analyses results. This is because with larger τ/T, a smaller horizon is imposed for the CAT problem, which makes it less likely to be treated as an infinite horizon problem. Figure 3.4(d) shows the network throughputs with $\rho = 10$dB. With higher SNR, the protocols have much better network throughputs. However, compared to the random access scheme, the performance gain from opportunistic scheduling is only 13%. This shows that the opportunistic scheduling scheme is particularly useful at lower SNR regions, where the random access scheme does not perform well in the first place.

In comparison, Figure 3.5 shows numerical results for the CDT problem with the same default parameters. Similar to the CAT problem, in Figure 3.5(a) we can see the infinite horizon analysis always yields larger network throughputs than the finite horizon analysis. The gap of the network throughputs between them is more than 30%, but eventually decreases to 0 as the network size M becomes



(a) (b)

Figure 3.5: Numerical results for the CDT problem for ad-hoc networks with M links, where the parameters are $p = 1/M$, $\rho = -10$dB and $\sigma = 1$: (a) network throughputs with $\tau = 0.01$; (b) network throughputs with $\tau = 0.05$.

large enough. On the other hand, with the same parameters the CDT problem in Figure 3.5(a) yields slightly larger network throughputs than the CAT problem in Figure 3.3(a). This coincides with our theoretical result in Proposition 3.7.3. On the other hand, we can see that the line with "$\diamond$" in Figure 3.5(a) approaches the finite horizon analysis faster than that of Figure 3.3(a). It implies that the CDT problem requires a smaller network size M than the CAT problem for using the infinite horizon analysis. In addition, Figure 3.5(a) shows the network throughputs from the improved protocol. Similar to Figure 3.3(a), the improved protocol always yields a slightly better performance from both analyses. Finally Figure 3.5(b) shows the network throughputs for a larger probing cost $\tau = 0.05$. We can see that the gap in the network throughputs between the two analyses is 10.6%, while this gap for the CAT problem is 14.7% in Figure 3.3(c). It implies that for the same network the CAT problem shows a smaller horizon compared to the CDT problem. This coincides with our earlier observation: to safely use infinite horizon analysis, the CAT problem generally requires a larger network size M. Furthermore, comparing both lines with "$\diamond$" in Figure 3.3(c) and Figure 3.5(b), we can see that the real rewards that can be achieved by the stopping rules from the infinite horizon analysis are very different. For the CAT problem, the expected real reward has a huge gap from the result based on the finite horizon analysis. For the CDT problem, the expected real reward approximates the result based on finite horizon analysis pretty well when the network size M is large enough, say $M = 15$. This implies that when the probing cost is high, it is particularly not recommended to design stopping rules based on the infinite horizon analysis for the CAT problem. The source of this difference lies in that there is always a constant duration of T for data transmission in the CDT problem.

Chapter 4

Distributed Opportunistic Scheduling for Multi-Channel Networks

4.1 Introduction

Many wireless systems now provide multiple channels for data transmission, where a channel can be treated as a frequency in a frequency division multiple access (FDMA) network, a code in a code division multiple access (CDMA) network, or an antenna or its polarization state in a multiple-input multiple-output (MIMO) network [59–61, 64–67]. For example, IEEE 802.11a has 8 channels for indoor use and 4 channels for outdoor use in the 5GHz band [68], and IEEE 802.11b has 3 channels in the 2.4GHz band [69]. Moreover, software defined radio (SDR) [70] and cognitive radio (CR) [71] systems also provide multiple channels, e.g. tunable frequency bands through programmable hardware that are controlled by software. In a multi-channel system, when the channel separation is greater than the coherent bandwidth, different channels experience independent channel fluctuations. Hence the availability of multiple channels substantially enhances the probability of the existence of at least one channel with acceptable channel quality. In such systems, the presence of multiple channels is a source of diversity which can be exploited opportunistically to enhance the system throughput. In general, to be able to exploit such kind of diversity, there should be enough vacant channels for the scheduler to decide which channels to choose. Hence it is not practical to exploit opportunism in multi-channel diversity if the wireless network is heavily loaded. However, it is shown that almost all wireless networks experience extended periods of low activity or hot spots [72], during which some or all access points can try to exploit the available spectrum opportunism.

In a multi-channel network, it is possible to exploit channel fluctuations from multi-channel diversity by sending data at a higher rate through one channel that is carefully chosen, which results in an enhanced system throughput. It requires the user to obtain the current states of all channels for decision making at the scheduler. However, to learn the instantaneous states of the channels, a user needs to probe the channel in many ad-hoc wireless networks, which in turn consumes both additional energy and time. Hence a user needs to not only optimally select the channel based on available information but also optimally determine the amount of information it should acquire about the instantaneous states of its available channels. In general, this problem is a joint optimization of the rewards obtained from informed selections and the cost incurred in acquiring the required information. To exploit opportunism in multi-channel diversity, many works have focused on how to derive an optimal strategy to determine which channels to probe, in what sequence, and which channel to use for data transmission [59–61, 64–67]. A heuristic based approach is first presented in [64], where the authors take channel fluctuations into consideration in the MAC design to exploit channel variations across multiple frequency channels for IEEE 802.11 networks. The multi-channel opportunistic auto rate (MOAR) [59, 65] considers the optimal opportunistic scheduling problem for IEEE 802.11 wireless networks with statistically identical channels and equal probing costs. It is assumed that unprobed channels cannot be used for transmission and a channel can only be used immediately after probing, i.e. no recall of previous channels. MOAR allows users to opportunistically find the channels with the best channel quality and seeks to optimally balance the throughput gain with measurement overhead. The key idea of MOAR is that if the quality of the current channel is not favorable, users can opportunistically skip to better quality channels for possible data transmission at a higher rate. An optimal skipping rule for MOAR which maps the channel conditions at the PHY layer to a MAC rule is presented to allow nodes to limit the number of times they skip in search for a better channel. It is shown that opportunistic channel skipping is most beneficial in low signal to noise regions, which are typically the cases when the node throughput in single-channel system is the minimum. The opportunistic spectrum access (OSA) [60, 66] studies this problem for a general wireless network where a channel can be in one of multiple states and the channel statistics are not necessarily identical. The authors consider both the case where the number of channel state is finite and the case where they can take an uncountably infinite number of states. Moreover, both recall of previous channel probing and transmission in unprobed channels are allowed. Both the constant data time (CDT) and the constant access time (CAT) problem are studied. The authors derive key properties of the optimal strategy and show that the optimal strategy has a threshold property and can only take one of a few structural forms. Based on the key structural properties of the optimal strategy, the authors show the optimal channel probing scheme for a number of special cases of practical interest as well as an algorithm that computes the optimal strategy in a finite number of steps even when the channel has an uncountably infinite state space. This problem is further extended in [61] to allow channels with different distributions of transmission qualities and different probing costs. The authors show that for an arbitrary number of states the optimal net gain can be approximated within a factor of 1/2 using a simple approximation algorithm. This approximation ratio can be improved to 2/3 when the number of channel states is 3. In [67], this problem is studied for the case where all channels have equal probing costs but potentially different distributions for the channel states. A joint channel probing and selection scheme is proposed to approximate a utility function that captures both the cost and value of information. The approximation can be made arbitrarily close to the optimum with an increasing computation time of the solution.

4.2 Motivation for the Multi-Channel Problem

We consider a wireless ad-hoc network with available bandwidth W. There are a total of M links competing the medium in a cooperative and opportunistic manner. The whole spectrum can be directly used as one single channel, using the distributed opportunistic scheduling protocols described in [53]. We assume a homogeneous network where the channel statistics are identical for different links. Now we are interested in better efficiency by dividing the whole bandwidth into J sub-channels, where $J < M$. We assume each sub-channel has a bandwidth of $\frac{W}{J}$, and the sub-channels are orthogonal and hence their channel fadings are independent from each other. This is a common assumption in literature. For example, wireless networks with independent sub-channels have been discussed in [59, 61, 66]. We denote the time as t and the number of active probing links on the j-th sub-channel at time t as $M_t^{(j)}$. For simplicity we use the scenario $M_t^{(j)} = \frac{M}{J}$ to illustrate our idea.

We first take a look at the average waiting time for any given link to access the medium. For a single-channel network, a given link m is able to access the current block with a probability $\frac{1}{M}$. Since the procedure is independent from block to block, the average waiting time before link m can send data through the wireless medium is MT, i.e. M blocks. For a multi-channel system, link m is able to access the current block with a probability $\frac{1}{M_t^{(j)}} = \frac{J}{M}$. Hence the average delay for link m to access the medium is $\frac{MT}{J}$. This is only $\frac{1}{J}$ of that of a single-channel network. Hence multi-channel protocols can considerably reduce the average waiting time for any given link to access the medium.

Now we consider the system throughput. Intuitively speaking, the system throughput is determined by how likely a "good" link can be found in the network. For a single-channel network, we assume the probability that the current captured channel rate for a given link being "good" is P_g. Hence the probability that the wireless medium will be utilized by a "good" link is MP_g. For the multi-channel network, on the other hand, we need to find J "good" links. This is because the bandwidth of each sub-channel is only $\frac{W}{J}$. Suppose the probability that the current captured transmission rate being "good" on a given sub-channel is $\tilde{P}_g$. If we consider the scheduling is independent between sub-channels, this probability is $J \cdot \frac{M}{J}\tilde{P}_g = M\tilde{P}_g$. There is no big difference compared to the single-channel scenario, since we can treat $P_g \approx \tilde{P}_g$ if the bandwidth W is evenly allocated to each sub-channel. On the other hand, if the distributed opportunistic scheduling is jointly designed across all sub-channels, this probability becomes $\binom{M}{J}\tilde{P}_g$. Hence this probability will be improved considerably when there are enough number of sub-channels J. It gives us some hint on the benefit from joint optimization across multiple sub-channels. However, it is tricky to design protocols that can work in a distributed scenario to achieve the opportunism introduced by these sub-channels.

4.3 The Multi-Channel Opportunistic Scheduling Algorithm

In this chapter, we study this problem under the constant access time (CAT) model [53,66], where the total duration of the channel probing and data transmission is a constant, i.e. $T_p + T_d = T$. We adopt this model so that the beginning of each block T on different sub-channels can easily be synchronized in a multi-channel network. Note that the duration of channel probing is a random variable depending on the stopping time N, i.e. $T_{p,N} = \sum_{i=1}^{N} T_i$.

4.3.1 Protocol Description

Similar to the single-channel scenario, we consider a collision model for each sub-channel where channel probing is required before accessing any sub-channel. The channel probing is still independent from sub-channel to sub-channel due to lack of centralized coordinator. Suppose at time t there are $M_t^{(j)}$ links actively probing the j-th sub-channel with a fixed probability p. The j-th sub-channel is

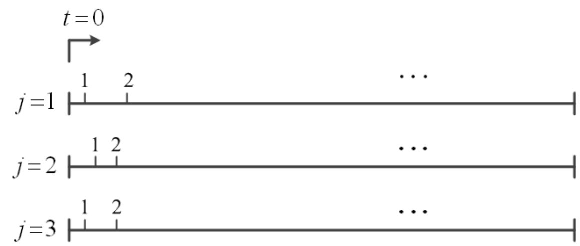


Figure 4.1: The channel probing within one block duration T for a network with 3 sub-channels.

won by some link after a duration of $\tau K_{n_j}^{(j)}$ and is captured at a transmission rate of $R_{n_j}^{(j)}$. Here $n_j = n_j(t)$ is the index for the round of successful channel probing on the j-th sub-channel. We can see $K_{n_j}^{(j)}$ has a geometric distribution with parameter $p_{s,t}^{(j)}$, where $p_{s,t}^{(j)}$ is the successful probing probability on the j-th sub-channel at time t and it depends on $M_t^{(j)}$.

Since the channel probing is independent for different sub-channels, $n_j(t)$ is generally asynchronous for different j. This is illustrated in Figure 4.1, where the numbers above each sub-channel indicate n_j at different time t. We can see that sub-channel 2 has its first winner link later than sub-channel 1 and 3, while sub-channel 1 has its second winner link later than sub-channel 2 and 3 respectively. Whenever any sub-channel is won by some link that is actively probing that sub-channel, we say one *event* happens in the system. Now we take a look at the whole procedure from time $t = 0$. We already know that it takes a duration of $\tau K_{n_j}^{(j)}$ for the n_j-th event to happen on the j-th sub-channel. Hence it takes a duration of $\tau \min K_{n_j}^{(j)}$ for the first event *ever* to happen in this network. Similarly, starting from the first event, it takes a duration of $\tau \min K_{n_j}^{(j)}$ for the second event to appear in the system, and so on. We denote the minimum duration across different sub-channels as

$$\tilde{K}_n = \min_{j \in \mathcal{J}_t} K_{n_j}^{(j)}, \tag{4.1}$$

where $n = n(t)$ is the index for the round of successful probing in the system, and $\mathcal{J}_t$ is the set of sub-channels that have not been utilized for data transmission until time t. We can see that $\tilde{K}_n$ is the shortest time interval between any two events (not necessarily originated from the same sub-channel) in the system.

Hence for a multi-channel network, it is not necessary to trace all events on a specific sub-channel and design optimal stopping rule for that sub-channel. Instead the protocol could make a decision as soon as there is a new event available in the system, no matter from which sub-channel this event is originated. As a result, one major difference of the multi-channel protocol is that the decision making at different time instances could be based on observations from different sub-channels. The full protocol can be described in Figure 4.2.

Note that in Figure 4.2, there are generally multiple winners on different sub-channels at the same time and hence multiple decision makings based on different instant transmission rates.

4.3.2 Performance Analysis

In this section, we analyze the proposed multi-channel protocol and characterize its system throughput. We present lower and upper bounds on the system throughput under various constraints.

In this chapter, we only characterize system performances for *homogeneous*

Each link m picks one sub-channel;
$\mathcal{J}_t \leftarrow \{1, 2, \ldots, J\}$;
while $\mathcal{J}_t \neq \emptyset$ **do**
  for each $j \in \mathcal{J}_t$ **do**
    links probe the j-th sub-channel;
  end for
  if link m wins some sub-channel j **then**
    m makes a decision on whether to send data on j or not;
    if m decides to utilize sub-channel j **then**
      m sends data over sub-channel j until the end of this block;
      sub-channel j is deleted from $\mathcal{J}_t$;
    end if
  end if
end while

Figure 4.2: The Distributed Opportunistic Scheduling Protocol for Multi-Channel Networks

networks, where the distributions of the transmission rates are identical with respect to different links or sub-channels. To facilitate our performance analysis, we make some additional assumptions as we did in Chapter 3:

[A1] The channel rates can only take values in $(0, +\infty)$;

[A2] The probing duration τ is much smaller compared to the block length, i.e. $\tau \ll T$.

We take a look at the channel probing and decision making procedure described by Figure 4.2. Suppose the j-th sub-channel is won by some link after a duration of $\tilde{K}_n$, which makes it the n-th round of successful channel probing for the multi-channel system. Suppose sub-channel j is then captured by the winner s_n at rate $\tilde{R}_n$.[1] Hence the reward is

$$Y_n = \frac{\tilde{R}_n \cdot (T - \tau \sum_{i=1}^{n} \tilde{K}_n)}{T} \tag{4.2}$$

if the winner s_n decides to utilize the channel, and is 0 otherwise. We can rewrite it as

$$Y_n = \frac{T - \tau \sum_{i=1}^{n} \tilde{K}_n}{T/\tilde{R}_n},$$

and the optimization is now reduced to maximize the rate of return [55, 56]. To do this, we need to characterize the probabilistic distribution of $\tilde{K}_n$.

Lemma 4.3.1. *$\tilde{K}_n$ has a geometric distribution* $\mathrm{Geom}(\tilde{p}_{s,t})$ *with parameter*

$$\tilde{p}_{s,t} = 1 - \prod_{j \in \mathcal{J}_t} \left[1 - p_{s,t}^{(j)}\right], \tag{4.3}$$

where $p_{s,t}^{(j)}$ is the successful probing probability on the j-th sub-channel at time t.

Proof. It is easy to compute the CDF of $\tilde{K}_n$ based on (4.1) if we notice that $K_{n_j}^{(j)}$ are independent geometric distributions with parameter $p_{s,t}^{(j)}$. □

Now that $\tilde{K}_n$ has a geometric distribution $\mathrm{Geom}(\tilde{p}_{s,t})$, we can apply a similar procedure as in [53] to characterize the optimal stopping rule.

Lemma 4.3.2. *Suppose at time t, the set of sub-channels that have not been utilized for data transmission is $\mathcal{J}_t$. Then the optimal stopping rule is*

$$N^* = \min\left\{ n \geq 1 : \tilde{R}_n \geq \lambda_n^* \cdot \frac{T}{T - \tau \sum_{i=1}^{n} \tilde{K}_i} \right\}, \tag{4.4}$$

[1] Strictly speaking, they should be denoted as $s_n^{(j)}$ and $\tilde{R}_n^{(j)}$ respectively, since there might be multiple winners on different sub-channels at time t. Here we ignore the superscript when discussing one of these winners.

where λ_n^* *is the solution to*

$$E\left[1 - \frac{\tau}{T}\left(\sum_{i=1}^{n} \tilde{K}_i - \tilde{K}_{n+1}\right) - \frac{\lambda}{\tilde{R}_n}\right]^+ = \frac{\tau}{T \cdot \tilde{p}_{s,t}}. \tag{4.5}$$

The optimal system throughput λ^* *is the solution to*

$$E\left[1 - \frac{\lambda}{\tilde{R}_n}\right]^+ = \frac{\tau}{T \cdot \tilde{p}_{s,t}}. \tag{4.6}$$

Here the successful probing probability $\tilde{p}_{s,t}$ *is defined in (4.3).*

Proof. The proof can be obtained in a similar way as the proof of Theorem 1 in [53]. $\square$

From Lemma 4.3.2, we can see the optimal reward and stopping rule only depend on the cumulative durations $\tau \sum_{i=1}^{n} \tilde{K}_n$ for channel probing and the captured instant channel rate $\tilde{R}_n$. Both of them are readily available for the winners' decision makings in a *distributed* setting, even though physically these events might be originated from different sub-channels.

We now take a look at the whole decision making procedure. Once a sub-channel is utilized for data transmission, it will not be involved in the channel probing until the beginning of the next block. Hence the cardinality of $\mathcal{J}_t$ (denoted as $J_t = \|\mathcal{J}_t\|$) decreases whenever there is a decision to stop. All sub-channels will be eventually utilized for data transmission. Hence there should be J decisions that are to stop in the end. The decreasing of J_t will affect the successful probing probability (4.3) for the multi-channel system and hence the system throughput.

We first characterize the optimal reward when the successful probing probability $\tilde{p}_{s,t}$ is varying as the procedure moves on.

Lemma 4.3.3. *Suppose the successful probing probability* $\tilde{p}_{s,t}$ *in Lemma 4.3.2 is varying as the procedure moves on. Suppose before one winner decides to stop, the minimum and maximum of* $\tilde{p}_{s,t}$ *are* $\tilde{p}_{s,min}$ *and* $\tilde{p}_{s,max}$ *respectively. Then the optimal system throughput* λ^* *for this decision can be bounded as*

$$\lambda^*_{min} \leq \lambda^* \leq \lambda^*_{max}, \tag{4.7}$$

where λ^*_{min} *is the system throughput if the successful probing probability is always* $\tilde{p}_{s,min}$*, and* λ^*_{max} *is the system throughput if the successful probing probability is always* $\tilde{p}_{s,max}$*.*

Proof. The proof is straight-forward if we notice that the optimal reward λ^* in (4.6) monotonically increases as $\tilde{p}_{s,t}$ increases. □

To calculate the system throughput, note that the successful probing probability $\tilde{p}_{s,t}$ increases as J_t increases. Hence $\tilde{p}_{s,t}$ reaches its maximal value in the beginning when $J_t = J$. Based on this we can get an upper bound on the system throughput. To simplify our notation, we further make the following assumptions:

[A3] Each sub-channel has the same number of links, i.e. $M_t^{(j)} = M_t^{(1)}$ for $j = 1, \ldots, J$;

[A4] All links are probing with the same probability, i.e. $p_{(m)} = p$ for $m = 1, \ldots, M$.

Thus all sub-channels have the same successful probing probabilities, i.e. $p_{s,t}^{(j)} = p_{s,t}^{(1)}$ for $j = 1, \ldots, J$.

Theorem 4.3.1. *The system throughput of Figure 4.2 is at most $J\lambda_0^*$, where λ_0^* is the solution to*

$$E\left[1-\frac{\lambda}{\tilde{R}_n}\right]^+ = \frac{\tau/T}{1-\left[1-p_{s,t}^{(1)}\right]^J}. \tag{4.8}$$

We can also have a lower bound on the system throughput if J_t is available for optimal decision making through some means.

Theorem 4.3.2. *If $J_t = \|\mathcal{J}_t\|$ is available for decision making in Figure 4.2, the system throughput is at least $\sum_{j=1}^{J}\gamma_j^*$, where γ_j^* is the solution to*

$$E\left[1-\frac{\gamma}{\tilde{R}_n}\right]^+ = \frac{\tau/T}{1-\left[1-p_{s,t}^{(1)}\right]^j}. \tag{4.9}$$

Proof. We can see when $J_t = j$, the optimal system throughput is γ_j^*. Hence if there is at most one sub-channel that is decided to be utilized for data transmission at any time t, the total network throughput will be exactly $\sum_{j=1}^{J}\gamma_j^*$.

Now suppose $J_t = j$. At this time there are still j sub-channels involved in active channel probing and decision making. Suppose Δ sub-channels are decided to be utilized for data transmission at some point. Then the reward from these sub-channels are $\Delta \cdot \gamma_j^*$. We can easily see that

$$\Delta \cdot \gamma_j^* > \sum_{i=j-\Delta+1}^{j} \gamma_i^*. \tag{4.10}$$

To bound the system throughput, iterate j from the very beginning $j = J$ and apply (4.10) when multiple sub-channels are decided to be utilized for data transmission at the same time. $\square$

Unfortunately, in ad-hoc networks J_t is not readily available for decision making. Starting from $J_t = J$, more and more sub-channels will be eventually utilized for data transmission as the procedure moves on. At any time more than one decisions over multiple sub-channels might be made to stop. Hence J_t is a random process which depends on the channel probing and decision making behavior. One solution to this problem is to conservatively use a *fixed* small J_0 as the true J_t for decision making. We can get a lower bound on the system throughput if the protocol works in this way.

Theorem 4.3.3. *If a fixed J_0 is used in Figure 4.2 to replace J_t when computing the optimal stopping rule, the system throughput is at least $(J - J_0 + 1)\zeta^*$, where ζ^* is the solution to*

$$E\left[1 - \frac{\zeta}{\tilde{R}_n}\right]^+ = \frac{\tau/T}{1 - \left[1 - p_{s,t}^{(1)}\right]^{J_0}}. \tag{4.11}$$

Proof. To characterize the throughput from each decision, we divide the whole procedure into two phases:

- $J_t \geq J_0$: The decision rule is more conservative as it is using a smaller $\tilde{p}_{s,t}$. Hence the decision making will stop earlier and result in a reward $\hat{\zeta}$. Apparently we have $\hat{\zeta} \geq \zeta^*$. The first $J - (J_0 - 1)$ sub-channels that are decided to be utilized for data transmission fall into this category.

- $J_t < J_0$: The decision rule is more optimistic compared to the true situation. There is a chance that it will never stop properly since a larger threshold is used here. The worst case is that we get a total reward of 0 for these $J_0 - 1$ sub-channels.

Now combine these two cases, we get a total throughput which is at least $(J - J_0 + 1)\zeta^*$. $\square$

4.4 Numerical Results

We consider a wireless network with a total bandwidth W. Without loss of generality, we assume the bandwidth is 1 in certain units, e.g. $W = 1$MHz. We assume the wireless medium is Rayleigh fading within each block $T = 1$. Hence if the whole spectrum is used as a single wireless channel, its channel rate can be written as

$$R(h) = \log(1 + \rho h)$$

in Mbits/s/Hz, where ρ is the average signal-to-noise ratio (SNR), and h is the channel gain corresponding to Rayleigh fading. We write the probability density function (pdf) of h as

$$f(h;\sigma) = \frac{h}{\sigma^2} e^{-\frac{h^2}{2\sigma^2}}, \quad h > 0.$$

There are a total of $M = 400$ links accessing the wireless medium with distributed opportunistic scheduling protocols.

For a multi-channel network, we split the total bandwidth evenly as $\frac{W}{J} = \frac{1}{J}$, where J is the number of sub-channels in the system. Accordingly the rate for each sub-channel can be written as

$$R^{(j)}(h) = \frac{1}{J}\log(1 + \rho h)$$

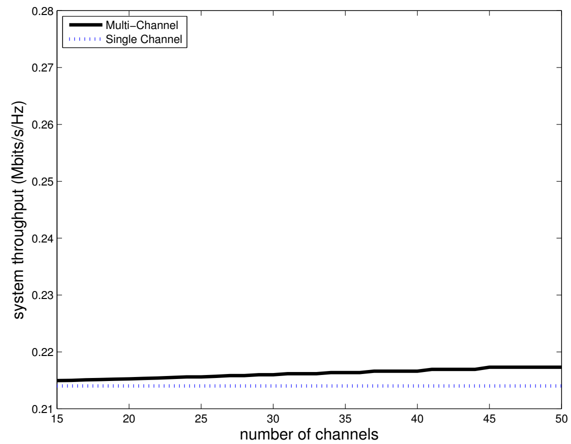


Figure 4.3: System throughput with varying number of sub-channels in the network, where $\tau = 0.02$, $\rho = -10$dB and $\sigma = 1$.

in Mbits/s/Hz, where $j = 1, \ldots, J$.

We first show that it only marginally improves the system throughput if the opportunistic scheduling is working independently on each sub-channel. Figure 4.3 shows the system throughput for this case, with parameters $\tau = 0.02$, $\rho = -10$dB and $\sigma = 1$. The number of sub-channels J is varying from $J = 15$ to $J = 50$. For comparison, the dotted line shows the system throughput for the single-channel network. For the single-channel system, the distributed opportunistic scheduling protocol is running where all links probe with probability $p = \frac{1}{M}$. For the multi-channel system, each sub-channel is running the single-channel protocol independently with $p = 1/\lfloor \frac{M}{J} \rfloor$. We can see it only yields a performance improvement of roughly 1.5% with $J = 50$ sub-channels.

In Figure 4.4, we show system throughput of the multi-channel opportunistic scheduling protocols based on various bounds discussed in Section 4.3.2. Similarly the dotted line shows the system throughput for the single-channel network. The dashdott line shows the upper bound of the system throughput described in Theorem

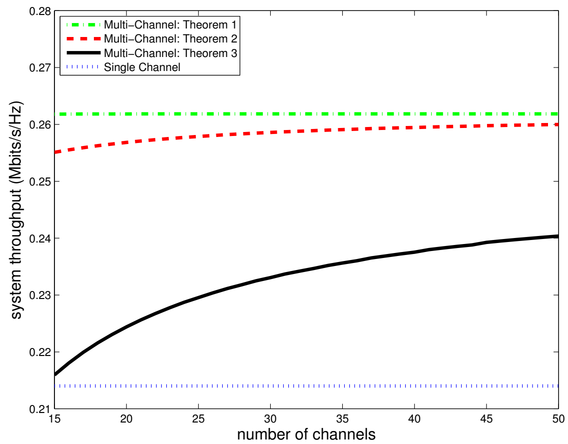


Figure 4.4: System throughput with varying number of sub-channels in the network, where $\tau = 0.02$, $\rho = -10$dB and $\sigma = 1$.

4.3.1. We can see the system throughput quickly reaches a maximal value at a relatively medium J. It shows an increase of almost 22.4% in network throughput. The dashed line shows the lower bound of the network throughput shown in Theorem 4.3.2, where J_t is available through other means for decision making. We can see that as J increases, it increases slower than the upper bound. For a large enough J, say $J > 25$, it shows an increase of 21.4% in network throughput. Finally, the solid line shows the network throughput of Figure 4.2 if we simply use $J_0 = 3$ in the decision making procedure. We can see that the network throughput increases much slower compared to the dashdot line. For $J = 50$, it shows an increase of 12.3% in the network throughput compared to the single-channel scenario. Hence even without any additional information, the distributed version of Figure 4.2 can still improve the system throughput considerably.

Chapter 5

A Multi-Armed Bandit Approach for Distributed Channel-Aware Scheduling

5.1 Introduction

In Chapter 3 and Chapter 4, we studied the distributed channel-aware scheduling problem where the instant CSIs are unknown but the channel statistics are known. In this chapter, we study one different distributed opportunistic scheduling problem where the channel statistics of the links are fixed but unknown. The goal is to learn these parameters in a distributed manner and minimize regret of the learning problem. We assume all links in the network have to probe the shared wireless medium before sending any data over the channel. We formulate this distributed learning problem using multi-armed bandit.

One of the earliest MAB papers solved classic non-Bayesian infinite horizon MAB problem with one single play [73]. Assume K independent arms and the rewards are *i.i.d.* over time from a given distribution with an unknown parameter, an order optimal policy is presented to provide expected regret that is $O(K \log n)$. This work is later extended to the case when multiple simultaneous plays are allowed [74]. Easier to compute policies based on the sample means are discussed in [75] that also has asymptotically logarithmic regret. In particular, the work in [76] considered arms with non-negative rewards that are *i.i.d.* over time with an arbitrary distribution, where the only restriction is that the distribution should have a finite support. A simple policy named UCB1 is proposed to achieve logarithmic regret uniformly over time [76].

Recently the MAB framework has been popularly used to formulate learning problems in wireless ad-hoc networks, e.g. cognitive radio networks. For example, a combinatorial MAB problem is considered in cognitive radio networks where a channel offers independent Bernoulli rewards with unknown means for different users [77]. A centralized policy based on full information exchange and cooperation among users is presented that achieves logarithmic order of the regret growth rate. In [78], the problem of secondary users selecting channels is formulated as a decentralized MAB problem, and a policy is presented to achieve asymptotically logarithmic regret with respect to time. The work in [79] extends the index-type single-user policy in [76] and proposes two order-optimal distributed policies, where users are orthogonalized to different channels.

Unlike other MAB problems, here all the rewards for different links are from the only shared wireless channel in the network. For performance comparison purpose, we first introduce a semi-distributed MAB protocol, which serves as our performance baseline due to its ideal assumption. Based on the semi-distributed protocol, we remove its ideal assumption and propose two distributed MAB protocols, where each link holds a set of local observations and plays the MAB game independently. We show that these distributed protocols can be considered as a generalized MAB protocol, where each link can update its local observations for infinitely many times.

We run simulations under different parameters and compare statistics of the local observations and regrets.

5.2 Problem Formulation

We use the similar channel model in Chapter 3 to formulate our problem. Here we use the constant access time (CAT) model [66] as an example to formulate our problem. Notice this problem can also be formulated using the constant data time (CDT) model in a similar way.

We assume M links are sharing the wireless medium in an ad-hoc network without any centralized coordinator. The channel has a collision model, where a link can successfully send data if and only if no other links are transmitting simultaneously. Hence to avoid collisions among themselves, links have to probe the medium first. A link wins the channel if and only if no other links are probing at the same time. Suppose the duration of a mini-slot for channel probing is fixed as τ. If link m probes the wireless medium with probability p_m, the duration of the channel probing phase is

$$T_s = \tau \cdot K_s, \tag{5.1}$$

where K_s is the number of mini-slots elapsed until some link finally wins the channel. Hence K_s has a geometric distribution Geom(p_s), where p_s is the successful probing probability

$$p_s = \sum_{m=1}^{M} p_m \prod_{j \neq m} (1 - p_j). \tag{5.2}$$

A sketch of our CAT based model is shown in Figure 5.1, where the total

Figure 5.1: CAT based system model

duration for channel probing and data transmission is a constant T. Furthermore, we assume the channel has a block fading with length T and the channel rate R_m is *i.i.d.* over time for the m-th link. Hence if link m wins the medium in some block, the available channel rate in this block is R_m and the duration available for data transmission is $T - \tau K_s$, which yields a total reward of $R_m(T - \tau K_s)$. We assume these links have no knowledge on the distribution of the channel rates $\{R_m\}$ except that they have a finite support. Without loss of generality, we assume $\{R_m\}$ are properly normalized to a finite support $[0, 1]$. We assume that R_m has a mean μ_m that is unknown to the links, even link m itself. We denote the set of the means as $\vec{\mu} = \{\mu_m\}$.

To this end, we formulate this problem using a non-Bayesian MAB [73, 74, 76]. Our work is mostly inspired by the work in [76], which can be applied to arms with rewards that are *i.i.d.* over time with an arbitrary unparameterized distribution. Since the channel can be occupied by at most one link at any time n, we define the k-th arm as the medium access configuration that *only* the k-th link is selected to send data over the wireless channel, where $k = 1, \ldots, M$. Suppose the reward from pulling the k-th arm at time n is $Y_k(n)$. We define *regret*, i.e. the difference between the expected reward that could be obtained by a genie who can pick the best arm at each time, and that obtained by a given policy π, i.e.

$$\mathcal{R}_n^{\pi}(\vec{\mu}) = n \cdot \max_k E[Y_k] - E_{\pi}[\sum_{t=1}^{n} Y_{\pi(t)}(t)]. \tag{5.3}$$

In the following sections, we design protocols to implement the selection of arms in a decentralized way. We first present a semi-distributed MAB protocol with ideal assumption in Section 5.3, then distributed MAB protocols with weak assumption in Section 5.4.

5.3 A Semi-Distributed MAB Protocol for Opportunistic Scheduling

We first present a semi-distributed MAB protocol for our opportunistic scheduling problem and give a theoretical bound on its regret. We use a relatively strong assumption for the semi-distributed MAB protocol presented here. As a result, this protocol cannot be directly used in ad-hoc networks. The main purpose of this semi-distributed MAB protocol is to serve as a performance baseline for the distributed MAB protocols in Section 5.4.

We assume all links in the network get some additional assistant from a third party, which is denoted as $\mathcal{S}$ in this paper. $\mathcal{S}$ is a node in the network with limited memory space and limited communication capability. $\mathcal{S}$ keeps a copy of the observations for the MAB game played in the network, i.e. $\{\hat{Y}_k, N_k\}$. Meanwhile, its limited communication capability only allows it to contact one link at a time. This could be an access point (AP) in wireless communications or an information fusion center in wireless sensor networks.

At time n (i.e. the beginning of the n-th block), $\mathcal{S}$ plays the MAB game and picks one link, say link k. $\mathcal{S}$ then notifies link k and link k will probe the medium with probability 1 instead of p_k. On the other hand, all other links are still probing the medium with probability p_m where $m \neq k$, since $\mathcal{S}$ cannot contact any more link besides k. Hence the instant reward that can be obtained by link k is

$$Y_k = R_k \cdot (T - \tau K_k), \tag{5.4}$$

where R_k is the channel rate for link k and K_k is the number of mini-slots elapsed by channel probing before the k-th link actually wins the wireless medium. Combine (5.2) and the fact that link k is probing with probability 1, K_k has a geometric distribution $\text{Geom}(p_{s,k})$ with

$$p_{s,k} = \prod_{m \neq k} (1 - p_m). \tag{5.5}$$

From Section 5.2, the channel rate R_k has a finite support $[0, 1]$. We also properly normalize the duration of a block and a mini-slot so that $T = 1$. Hence based on (5.4), the reward Y_k also has a finite support $[0, 1]$. Since the MAB game is played only by $\mathcal{S}$ who holds a copy of all observations on rewards, we design a protocol based on the UCB1 policy in [76]. The semi-distributed MAB protocol is described in Figure 5.2.

In Figure 5.2, $\hat{Y}_k(n)$ is the mean reward from pulling the k-th arm up to the current time n, and $N_k(n)$ is the number of times the k-th arm has been pulled up to time n.

Similar to the UCB1 policy in [76], we can have a bound on the regret.

Theorem 5.3.1. *If $p_k = p$ for all $k = 1, \ldots, M$, the expected regret of the semi-*

for each link k **do**

  pull arm k once and update $\hat{Y}_k$, N_k accordingly;

end for

$n \leftarrow M + 1$;

while 1 **do**

  pull arm k that maximizes $\hat{Y}_k + \sqrt{\frac{2 \ln n}{N_k}}$;

  for each link j **do**

    probe the medium with probability 1 if $j = k$, otherwise with probability p_j;

  end for

  link k sends data over the channel after winning the medium;

  update $\hat{Y}_k$ and N_k according to the current reward y_k;

  $n \leftarrow n + 1$;

end while

Figure 5.2: The Semi-Distributed MAB Protocol

distributed protocol is at most

$$\left[\frac{8}{T - \frac{\tau}{p_{s,1}}} \sum_{k:\, \mu_k < \mu^*} \frac{\ln n}{\delta_k}\right] + \left(1 + \frac{\pi^2}{3}\right)\left(T - \frac{\tau}{p_{s,1}}\right)\left(\sum_{k:\, \mu_k < \mu^*} \delta_k\right), \tag{5.6}$$

where $\mu^* = \max_k \mu_k$ *and* $\delta_k = \mu^* - \mu_k$.

Proof. We define the mean reward from pulling arm k as

$$\begin{aligned}\theta_k &= E[Y_k] = E[R_k] \cdot E[T - \tau K_k] \\ &= \mu_k (T - \frac{\tau}{p_{s,k}}).\end{aligned}$$

From Theorem 1 in [76], regret of Figure 5.2 is at most

$$\left[8 \sum_{k:\theta_k < \theta^*} (\frac{\ln n}{\Delta_k})\right] + (1 + \frac{\pi^2}{3}) \left(\sum_{k:\theta_k < \theta^*} \Delta_k\right), \tag{5.7}$$

where $\theta^* = \max_k \theta_k$ and $\Delta_k = \theta^* - \theta_k$. Notice in (5.5) we have $p_{s,k} = p_{s,1}$ when $p_k = p$ for $k = 1, \ldots, M$. Substituting $\theta_k = \mu_k(T - \frac{\tau}{p_{s,1}})$ into (5.7) immediately yields (5.6). $\square$

5.4 Distributed MAB Protocols for Channel-Aware Scheduling

In this section, we present our distributed MAB protocols for the opportunistic scheduling problem. Unlike the semi-distributed protocol in Section 5.3, the MAB game is played independently in each link j in both protocols.

In ad-hoc networks, the existence of $\mathcal{S}$ in Section 5.3 is often questionable due to the distributed nature of the network. One possible idea is to let each link j hold a copy of its own observations and play the MAB game independently at each time n. We denote these observations that are held locally at link j as $\{\hat{Y}_k^j, \hat{N}_k^j\}_k$. Similar to Section 5.3, link j probes the wireless medium with probability 1 if the j-th arm is pulled in its local MAB game based on $\{\hat{Y}_k^j, \hat{N}_k^j\}_k$, otherwise with probability p_j. However, one unique feature here is at any time n, at most one link j is able to win the wireless medium and hence update its local observations. This is due to the collision channel model we used. As a result, the observations held at link j can only be updated when link j wins the medium, which only brings *new* observations to $\hat{Y}_j^j$ and $\hat{N}_j^j$. In other words, observations on other arms $\{\hat{Y}_k^j, \hat{N}_k^j\}_{k \neq j}$ can never be updated at link j as in other MAB problems.

To properly update these observations, we introduce the following two operations:

[Type-T update] when the k-th arm is pulled and a reward y_k is obtained, update the observations as:

$$\begin{aligned} \hat{Y}_k(n+1) &\leftarrow \frac{\hat{Y}_k(n) \cdot \hat{N}_k(n) + y_k}{\hat{N}_k(n) + 1} \\ \hat{N}_k(n+1) &\leftarrow \hat{N}_k(n) + 1 \end{aligned} \tag{5.8}$$

[Type-F update] when the k-th arm is pulled but no reward is able to be observed, update the observations as:

$$\begin{aligned} \hat{Y}_k(n+1) &\leftarrow \hat{Y}_k(n) \\ \hat{N}_k(n+1) &\leftarrow \hat{N}_k(n) + 1 \end{aligned} \tag{5.9}$$

Notice a type-F update only increases $\hat{N}_k$ by 1 to reflect the fact that the k-th arm has been pulled at time n. In fact, we have implicitly used the assumption that the instant reward y_k at time n is the up-to-date observation $\hat{Y}_k(n)$, i.e.

$$\hat{Y}_k(n+1) \leftarrow \frac{\hat{Y}_k(n) \cdot \hat{N}_k(n) + \hat{Y}_k(n)}{\hat{N}_k(n) + 1} = \hat{Y}_k(n).$$

This is because we cannot observe a new reward y_k at this time. Hence to update the average reward $\hat{Y}_k^j$, link j must be able to exchange information with link k in a certain way, since the average reward on the k-th arm can only be updated by the k-th link. In this paper, we present distributed MAB protocols based on some additional assistance from a different type of third party $\mathcal{W}$. Unlike $\mathcal{S}$ in Section 5.2, $\mathcal{W}$ is a node in the network who cannot initiate communications to any link j. Instead, the sole functionality of $\mathcal{W}$ is to provide a common memory space which can only be visited by at most one link at a time. We assume that only the link who wins the wireless medium can visit $\mathcal{W}$ and access the data kept at $\mathcal{W}$. Depending on the reliability of $\mathcal{W}$, we present two different protocols:[1]

Protocol A when $\mathcal{W}$ is highly reliable (denoted as $\mathcal{A}$);

Protocol B when $\mathcal{W}$ is not reliable (denoted as $\mathcal{B}$).

5.4.1 Distributed MAB Protocol A

In this section, we assume there is a third party $\mathcal{A}$ who is able to provide a *reliable* memory space to be visited by the winner link in the network at any time n.[2] Due to its high reliability, any data kept at $\mathcal{A}$ is unlikely to be destroyed. As a result, we keep a copy of the observations based on all type-T updates in the network at $\mathcal{A}$, which is denoted as $\{\hat{Y}_k^A, \hat{N}_k^A\}_k$ in this paper. The protocol works as follows: each link j starts with observations $\{\hat{Y}_k^A, \hat{N}_k^A\}_k$, and plays a local MAB game at each time n. If the j-th arm is pulled, link j probes the wireless medium with probability 1, otherwise with probability p_j. If link j does not win the wireless medium in this block, a type-F update is performed. If link j successfully wins the wireless medium, link j will access the observations at $\mathcal{A}$ and have a type-T update based on the instant reward y_j in this block. Then link j will update its local observations $\{\hat{Y}_k^j, \hat{N}_k^j\}_k$ and $\mathcal{A}$'s observations $\{\hat{Y}_k^A, \hat{N}_k^A\}_k$ using these new quantities. The detail of Protocol A can be described as Figure 5.3.

We show that the distributed MAB Protocol A with type-F update works as a generalized MAB procedure, where the average operation in a conventional MAB

[1] We need to explain clearly about the physical meaning of $\mathcal{W}$.

[2] Similarly, we need to clarify more about the physical meaning of $\mathcal{A}$.

for each arm k **do**
  pull arm k once and get reward y_k according to (5.4);
  set initial observations as $\hat{Y}_k^j \leftarrow y_k$ and $\hat{N}_k^j \leftarrow 1$ for $j = \mathcal{A}$ and $j = 1, \ldots, M$;
end for
$n \leftarrow M + 1$;
while 1 **do**
  for each link j **do**
    pull the k-th arm that maximizes $\hat{Y}_k^j + \sqrt{\frac{2 \ln n}{\hat{N}_k^j}}$;
    if $k = j$ **then**
      link j probes the medium with probability 1;
      if consistent collision detected **then**
        have a type-F update on $\{\hat{Y}_j^j, \hat{N}_j^j\}$;
      end if
    else
      link j probes the medium with probability p_j;
      have a type-F update on $\{\hat{Y}_k^j, \hat{N}_k^j\}$;
    end if
    if link j wins the medium **then**
      send data over the channel until the end of this block;
      have a type-T update on $\{\hat{Y}_j^A, \hat{N}_j^A\}$ based on the current reward y_j;
      $\hat{Y}_k^j(n+1) \leftarrow \hat{Y}_k^A(n+1)$ and $\hat{N}_k^j(n+1) \leftarrow \hat{N}_k^A(n+1)$ for all k;
    end if
  end for
  $n \leftarrow n + 1$;
end while

Figure 5.3: The distributed MAB protocol A

is replaced by a weighted average in Protocol A.

Theorem 5.4.1. *For any link j, Protocol A updates $\hat{Y}_k^j(n)$ as a weighted average based on all $y_k(t)$ that are available to link j by time n, where $y_k(t)$ is the instant reward when link k wins the wireless medium at time t.*

Proof. We prove Theorem 5.4.1 by mathematical induction. According to the initialization in Figure 5.3, at time $n = M+1$, we have $\hat{Y}_k^j(n) = y_k(k)$. Suppose the statement holds at time n, we can write $\hat{Y}_k^j(n)$ as

$$\hat{Y}_k^j(n) = \frac{\sum_{t\in\Gamma_k(n)} \hat{c}_k^j(t)\cdot y_k(t)}{\hat{N}_k^j(n)},$$

where $\Gamma_k(n)$ is the set of time t when link k wins the wireless medium and obtains a reward $y_k(t)$, and $y_k(t)$ has already been spread to link j, and $\hat{c}_k^j(t)$ is a coefficient. We notice that $\hat{N}_k^j(n) = \sum_{t\in\Gamma_k(n)} \hat{c}_k^j(t)$.

At time $n+1$, we can write the type-F update as

$$\begin{aligned}
\hat{Y}_k^j(n+1) = \hat{Y}_k^j(n) &= \frac{[\hat{N}_k^j(n)+1]\cdot \frac{\sum_{t\in\Gamma_k(n)} \hat{c}_k^j(t)\cdot y_k(t)}{\hat{N}_k^j(n)}}{\hat{N}_k^j(n)+1} \\
&= \frac{\sum_{t\in\Gamma_k(n)} \tilde{c}_k^j(t)\cdot y_k(t)}{\hat{N}_k^j(n)+1} \\
&= \frac{\sum_{t\in\Gamma_k(n+1)} \tilde{c}_k^j(t)\cdot y_k(t)}{\hat{N}_k^j(n+1)},
\end{aligned}$$

where $\hat{N}_k^j(n+1) = \hat{N}_k^j(n)+1$, and

$$\tilde{c}_k^j(t) = \frac{\hat{N}_k^j(n)+1}{\hat{N}_k^j(n)}\cdot \hat{c}_k^j(t),$$

and $\Gamma_k(n+1) = \Gamma_k(n)$ since this is a type-F update. Hence the statement also holds at time $n+1$. By mathematical induction this proves the theorem. $\square$

We show that under Protocol A, each link j can win the wireless medium and update its local observations $\{\hat{Y}_k^j, \hat{N}_k^j\}_k$ for infinitely many times. We use the following notations in our proof.

$$\begin{aligned} \mathcal{U} = \{j \,|\, &\text{link } j \text{ wins the wireless medium for an} \\ &\text{infinite number of times}\} \end{aligned} \tag{5.10}$$

$$\begin{aligned} \mathcal{V} = \{j \,|\, &\text{link } j \text{ wins the wireless medium for a} \\ &\text{finite number of times}\} \end{aligned} \tag{5.11}$$

We first show that at least one link j is able to win the wireless medium and update its local observations for infinitely many times.

Lemma 5.4.1. *Under Protocol A, $\mathcal{U} \neq \emptyset$.*

Proof. We prove it by contradiction. We assume that Lemma 5.4.1 doesn't hold, i.e. $\mathcal{U} = \emptyset$, which immediately leads to

$$\mathcal{V} = \{1, 2, \ldots, M\}.$$

Hence any link j can only win the wireless medium and update its local observations for a finite number of times. Then we can always find a large enough integer n_0 such that after time n_0, no links can win the wireless medium again.

Consider the MAB game played at link j. We notice that since link j never wins the wireless medium for $n > n_0$, $\hat{Y}_k^j(n)$ is a constant for any k. Hence only $\hat{N}_k^j(n)$ is updated when the k-th arm is pulled in link j's local MAB game. This means that for $n > n_0$, the pulling of the arms is solely determined by

$$\sqrt{\frac{2\ln(n)}{\hat{N}_k^j(n)}}.$$

Furthermore, since $\ln(n)$ is the same for any k, the fact that which arm will be pulled at when is decided by $\hat{N}_k^j(n)$ in a deterministic fashion. Hence we can always find some $n_1 > n_0$ such that after $n > n_1$, any arm k will be pulled for exactly once during time interval $[n, n+M]$ in any link j's local MAB game.

We take a look at this time interval. For any $n \leq t \leq n+M$, at least two links will have a type-T update. As a result, from $t = n$ to $t = n + M$, there will be at least $2 \cdot M$ type-T updates. On the other hand, we notice any link j can have exactly 1 type-T update during this time interval. Hence these type-T updates require at least $2M$ links. This leads to a contradiction since there are only M links in the network. $\square$

We then show all links j are able to win the wireless medium and update its local observations for infinitely many times.

Theorem 5.4.2. *Under Protocol A, any link l will win the wireless medium for an infinite number of times, i.e. $\mathcal{V} = \emptyset$.*

Proof. We prove Theorem 5.4.2 by contradiction. Suppose that $\mathcal{V} \neq \emptyset$. On the other hand, we know $\mathcal{U} \neq \emptyset$ from Lemma 5.4.1.

We can always find a large enough constant n_1 such that for $n \geq n_1$, no links in $\mathcal{V}$ will win the wireless medium again. We then consider the following

$$\hat{Y}_j^i(n) + \sqrt{\frac{2\ln(n)}{\hat{N}_j^i(n)}}$$

and

$$\hat{Y}_k^i(n) + \sqrt{\frac{2\ln(n)}{\hat{N}_k^i(n)}},$$

where $j \in \mathcal{V}$, and $i, k \in \mathcal{U}$. We notice that whenever link i wins the medium and update its local observations according to $\{\hat{Y}_k^A(n), \hat{N}_k^A(n)\}_k$, $\hat{N}_j^i(n)$ will be updated to the true number of times that link j has won the medium. Since after time n_1, j never wins the wireless medium again, $\hat{N}_j^i(n)$ is a constant whenever it is updated by link i. Meanwhile, for $k \in \mathcal{U}$, $\hat{N}_k^i(n)$ will increase since link k will win the wireless medium for infinitely many times. Hence we can always find a large enough constant n_2 such that for immediately after $n \geq n_2$,

$$\hat{N}_j^i(n) \ll \hat{N}_k^i(n).$$

Hence after that i must pull some $k \in \mathcal{V}$ for at least M consecutive times, since

$$\hat{Y}_j^i(n) + \sqrt{\frac{2\ln(n)}{\hat{N}_j^i(n)}} > \hat{Y}_k^i(n) + \sqrt{\frac{2\ln(n)}{\hat{N}_k^i(n)}}.$$

This means a pure random access of the wireless medium, which leads to a contradiction with the assumption that some $l \in \mathcal{V}$ cannot win the wireless medium again after time n_1. This proves Theorem 5.4.2. □

5.4.2 Distributed MAB Protocol B

Notice in Section 5.4.1, it is still a strong assumption that the local data kept at $\mathcal{A}$ are reliable. In this section, we remove this strong condition. We assume a different third party $\mathcal{B}$ similar to $\mathcal{A}$ except that the memory space provided by $\mathcal{B}$ is not reliable.[3] Denote the observations that are held at $\mathcal{B}$ as $\{\hat{Y}_k^B, \hat{N}_k^B\}_k$. As a result, $\{\hat{Y}_k^B, \hat{N}_k^B\}_k$ might be destroyed at any time n. As a result each link j must keep

[3] Similarly, we need to clarify more about the physical meaning of $\mathcal{B}$.

its local observations separately. Each link j starts with its own local observations $\{\hat{Y}_k^j, \hat{N}_k^j\}_k$, and plays a local MAB game at each time n. Link j probes the wireless medium with probability 1 if the j-th arm is pulled, otherwise with probability p_j. If link j does not win the wireless medium in this block, a type-F update is carried out. If link j successfully wins the wireless medium, link j will first have a type-T update on $\{\hat{Y}_j^j, \hat{N}_j^j\}$, and then exchange information with $\mathcal{B}$ and update them as

$$\begin{aligned} \hat{Y}_k^j(n+1) = \hat{Y}_k^B(n+1) = \frac{\hat{Y}_k^j(n) + \hat{Y}_k^B(n)}{2} \\ \hat{N}_k^j(n+1) = \hat{N}_k^B(n+1) = \frac{\hat{N}_k^j(n) + \hat{N}_k^B(n)}{2} \end{aligned} \tag{5.12}$$

for all $k = 1, \ldots, M$. Notice here $\hat{Y}_k^j(n)$ is no longer an integer due to the average carried out in (5.12). The detail of the protocol is shown in Figure 5.4.

One remaining question in Protocol B is the initialization of the observations $\hat{Y}_k^j$ for all links $j = 1, \ldots, M$. We discuss two different methods in this paper. In the first approach, we assume all links j have exactly the same initial observations, as shown as Figure 5.5(a). This is a relatively strong assumption for ad-hoc networks.

For the second approach, we assume all links j have different initial observations. The detail is shown in Figure 5.5(b). Notice in this case, each link j is required to obtain a separate initial observation $\hat{Y}_k^j$ for the k-th arm through a certain way.

To show the effectiveness of Protocol B, we show similar conclusions as Section 5.4.1.

Theorem 5.4.3. *For any link j, Protocol B updates $\hat{Y}_k^j(n)$ as a weighted average based on all $y_k(t)$ that are available to link j by time n, where $y_k(t)$ is the instant reward when link k wins the wireless medium at time t.*

initialization with methods in Figure 5.5;
$n \leftarrow M + 1$;
while 1 **do**
  for each link j **do**
    pull the k-th arm that maximizes $\hat{Y}_k^j + \sqrt{\frac{2 \ln n}{\hat{N}_k^j}}$;
    if $k = j$ **then**
      link j probes the medium with probability 1;
      if consistent collision detected **then**
        have a type-F update on $\{\hat{Y}_j^j, \hat{N}_j^j\}$;
      end if
    else
      link j probes the medium with probability p_j;
      have a type-F update on $\{\hat{Y}_k^j, \hat{N}_k^j\}$;
    end if
    if link j wins the medium **then**
      send data over the channel until the end of this block;
      have a type-T update on $\{\hat{Y}_j^j, \hat{N}_j^j\}$ based on the current reward y_j;
      exchange observations with $\mathcal{B}$ and update them according to (5.12);
    end if
  end for
  $n \leftarrow n + 1$;
end while

Figure 5.4: The distributed MAB protocol B

Proof. The proof can follow the proof of Theorem 5.4.1, and notice that (5.12) in Protocol B is also a weighted average. □

Lemma 5.4.2. *Under Protocol B, we also have* $\mathcal{U} \neq \emptyset$.

The proof of this lemma can directly follow the proof in Section 5.4.1. We

for each arm k **do**
  pull arm k once and get reward y_k according to (5.4);
  update initial observations as $\hat{Y}_k^j \leftarrow y_k$ and $\hat{N}_k^j \leftarrow 1$ for all links $j = 1, \dots, M$;
end for

(a)

for each link j **do**
  for each arm k **do**
    pull arm k once and get reward y_k according to (5.4);
    update initial observation as $\hat{Y}_k^j \leftarrow y_k$ and $\hat{N}_k^j \leftarrow 1$;
  end for
end for

(b)

Figure 5.5: The initialization of Protocol B in Figure 5.4 with: (a) homogeneous observations, (b) heterogeneous observations.

then show that all links j are able to win the wireless medium and update its local observations for infinitely many times.

Theorem 5.4.4. *For any link $l \in \mathcal{M}$, l will win the wireless medium for an infinite number of times, i.e. $\mathcal{V} = \emptyset$.*

Proof. We prove Theorem 5.4.4 by contradiction. We assume that $\mathcal{V} \neq \emptyset$ and its cardinality as

$$\|\mathcal{V}\| = b,$$

where b is an integer such that $1 \leq b \leq M$. From Lemma 5.4.2, we know that the cardinality of $\mathcal{U}$ is also a positive integer, i.e. $\|\mathcal{U}\| = M - b > 0$.

From the proof of Lemma 5.4.1, we know that for any link $j \in \mathcal{U}$, the pulling of the arms in j's local MAB game is deterministic with a period M after $n > n_1$. If we look at the system time from $t = n_1$ to $t = n_1 + M$, and we pay special attention to the timing where link j performs a type-T update. Without loss of generality, we can divide the whole duration $[n_1, n_1 + M]$ into three sub-intervals:

- Time interval $t \in [n_1, n_1 + \tilde{b}_1]$: there are at least two links $j \in \mathcal{V}$ who are having type-T updates at time t;
- Time interval $t \in [n_1 + \tilde{b}_1, n_1 + \tilde{b}_1 + \tilde{b}_2]$: there is *exactly* one link $j \in \mathcal{V}$ who is having a type-T update at time t, and one link $i \in \mathcal{U}$ who is having a type-T update at time t;
- Time interval $t \in [n_1 + \tilde{b}_1 + \tilde{b}_2, n_1 + M]$: all links $j \in \mathcal{V}$ are having type-F updates, and at leat one link $i \in \mathcal{U}$ is having a type-T update at time t;

Here $\tilde{b}_1$ and $\tilde{b}_2$ are non-negative integers. Define $\tilde{b} = \tilde{b}_1 + \tilde{b}_2$, we have the following:

$$\begin{aligned} 0 \le \tilde{b}_1 &\le \tilde{b}, \\ 0 \le \tilde{b}_2 &\le \tilde{b}, \\ \tilde{b}_1 + \tilde{b}_2 &= \tilde{b}, \\ \bar{N}_1 \cdot \tilde{b}_1 + \tilde{b}_2 &= b, \end{aligned}$$

where $\bar{N}_1$ is the average number of links $j \in \mathcal{V}$ that have type-T updates at any time $t \in [n_1, n_1 + \tilde{b}_1]$. It must be $\bar{N}_1 \ge 2$, since there is a collision in the wireless medium for any $t \in [n_1, n_1 + \tilde{b}_1]$. This procedure can be explained in Figure 5.6.

Now consider all links $i \in \mathcal{U}$. During time interval $[n_1, n_1 + M]$, in total $\tilde{b}_2 + (M - \tilde{b})$ type-T updates have been carried out from links $i \in \mathcal{U}$. Hence on

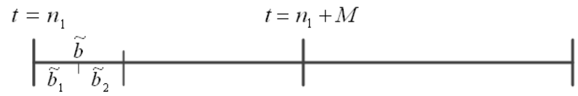


Figure 5.6: A sketch for the proof of Theorem 5.4.4.

average, links from $\mathcal{U}$ have type-T updates with a rate

$$\frac{\tilde{b}_2 + M - \tilde{b}}{M(M-b)}. \tag{5.13}$$

It is easy to show that this rate is greater than $1/M$. First of all, we notice that $\tilde{b} \leq b$ and thus $M - \tilde{b} \geq M - b$. If $\tilde{b}_2 > 0$, we can easily show that

$$\frac{\tilde{b}_2 + M - \tilde{b}}{M(M-b)} > \frac{M - \tilde{b}}{M(M-b)} \geq \frac{M-b}{M(M-b)} = 1/M.$$

On the other hand, if $\tilde{b}_2 = 0$, we have

$$\frac{\tilde{b}_2 + M - \tilde{b}}{M(M-b)} = \frac{M - \tilde{b}}{M(M-b)} = \frac{M - \tilde{b}_1}{M(M - \bar{N}_1 \cdot \tilde{b}_1)} > 1/M.$$

The last inequality is because $\tilde{N}_1 \geq 2$.

Since the average rate of type-T updates for all links from $\mathcal{U}$ is larger than $1/M$, there must be at least one link $i \in \mathcal{U}$ such that the rate of type-T updates in link i's local MAB game is larger than $1/M$. Denote this link as i^*. Now consider the local MAB game played in link i^*. The i^*-th arm is pulled with a rate larger than $1/M$, which is the average rate for any arm l being pulled. Hence there must be some arm k such that the rate of its being pulled is smaller than $1/M$. Hence in link i^*'s local MAB game, $\hat{N}_{i^*}^{i^*}$ increases faster than $\hat{N}_k^{i^*}$. Now consider

$$\hat{Y}_{i^*}^{i^*}(n) - \hat{Y}_k^{i^*}(n) + \sqrt{\frac{2\ln(n)}{\hat{N}_{i^*}^{i^*}(n)}}$$

and

$$\sqrt{\frac{2\ln(n)}{\hat{N}_k^{i^*}(n)}}.$$

Since link i^* wins the wireless medium for an infinite number of times, its local observations $\{\hat{Y}_k^{i^*}, \hat{N}_k^{i^*}\}$ will be updated for infinitely many times. Hence for any small positive ϵ_{i^*}, we can always find some $N(\epsilon_{i^*})$ such that for $n > N(\epsilon_{i^*})$, we can bound $\hat{Y}_{i^*}^{i^*}(n)$ as

$$|\hat{Y}_{i^*}^{i^*}(n) - \mu_{i^*}| \leq \epsilon_{i^*}$$

with high probability. Similarly, we can also bound $\hat{Y}_k^{i^*}(n)$ as

$$|\hat{Y}_k^{i^*}(n) - \mu_k| \leq \epsilon_k$$

with high probability. We notice that $\hat{N}_{i^*}^{i^*}(n)$ increases faster than $\hat{N}_k^{i^*}(n)$ as time n moves on. Hence for any integer L, we can always find some n_L such that

$$\begin{aligned}\sqrt{\frac{2\ln(n)}{\hat{N}_k^{i*}(n) + L}} &> (\mu_{i^*} + \epsilon_{i^*}) - (\mu_k - \epsilon_k) + \sqrt{\frac{2\ln(n)}{\hat{N}_{i^*}^{i*}(n)}} \\ &\geq \hat{Y}_{i^*}^{i^*}(n) - \hat{Y}_k^{i^*}(n) + \sqrt{\frac{2\ln(n)}{\hat{N}_{i^*}^{i*}(n)}}.\end{aligned}$$

Hence from $t = n_L$ to $t = n_L + L$, the following inequality always holds

$$\hat{Y}_k^{i^*}(t) + \sqrt{\frac{2\ln(t)}{\hat{N}_k^{i*}(t)}} > \hat{Y}_{i^*}^{i^*}(t) + \sqrt{\frac{2\ln(t)}{\hat{N}_{i^*}^{i*}(t)}}.$$

This means link i^* cannot have a type-T update in the time interval $[n_L, n_L + L]$. For any L, we can always find such a n_L. It contradicts with our assumption that starting from any time n, link i^* can win the wireless medium and update its local observations within a finite time duration. Hence the assumption $\mathcal{V} \neq \emptyset$ cannot hold, which proves that $\mathcal{V} = \emptyset$. $\square$

5.5 Distributed MAB Protocol II for Channel-Aware Scheduling

In this section, we present another distributed MAB protocol for the channel-aware scheduling problem. In contrast to Section 5.4, we use a different approach to exchange the local information sets among different links in the network.

In Section 5.4, we introduced a third party $\mathcal{W}$ as a mean to exchange information between different links in the network. Even though we considered the possibility that the local information set that are held at $\mathcal{W}$ are not reliable, sometimes it is still difficult to implement such a third party $\mathcal{W}$ in an ad-hoc network. In this section, we use a different approach to solve this problem, where the information exchange between different links are achieved through broadcasting at a carefully selected time interval at the end of one block. We explicitly consider how the packet loss during the broadcasting phase affects the performance of the protocol.

5.5.1 Protocol Description

We consider a channel model similar to the one used in Section 5.4. M links are sharing the wireless medium in an ad-hoc network without any centralized coordinator. The channel has a collision model, where a link can successfully send data if and only if no other links are transmitting simultaneously. Hence to avoid collisions among themselves, links have to probe the medium first. A link wins the channel if and only if no other links are probing at the same time. Suppose the duration of a mini-slot for channel probing is fixed as τ. If link m probes the wireless medium with probability p_m, the duration of the channel probing phase is p_s shown in (5.2). We still use the CAT model to formulate our problem. The total duration for channel probing and data transmission is a constant T. Furthermore, we assume the channel has a block fading with a block length T and the channel rate R_m is *i.i.d.* over time for the m-th link.

As we explained in Section 5.4, in an ad-hoc network, each link j has to independently maintain a local copy of the observations on different arms, denoted as $\{\hat{Y}_k^j, \hat{N}_k^j\}_k$. The problem is, the observations held at link j can only be updated when link j wins the medium, which only brings *new* observations to $\hat{Y}_j^j$ and $\hat{N}_j^j$. In other words, observations on other arms $\{\hat{Y}_k^j, \hat{N}_k^j\}_{k \neq j}$ can never be updated at link j. Hence these links need to exchange their local observations $\{\hat{Y}_k^j, \hat{N}_k^j\}_k$ in a distributed fashion. In this section, we accomplish this by allowing links to listen at a specifically designated time interval within one block T.

To do this, we notice that all links in the network can "listen" to the medium provided that the starting and ending time of the listening period are well synchronized. It should be noted that the starting and ending point of each block is fixed since the block length is a constant T. Hence we utilize a small time interval τ_0 at the end of each block T to exchange the local information set at each link via broadcasting. Suppose within this block the winner link is j. Then at time $T - \tau_0$, link j will perform a type-T update on its local observations $\{\hat{Y}_j^j, \hat{N}_j^j\}$ on the j-th arm and broadcast it to other links. It should be noted that the local observations on other arms at link j, e.g. $\{\hat{Y}_k^j, \hat{N}_k^j\}_{k \neq j}$, are not broadcasted to other links. This is because these observations might not be the most up-to-date ones, in particular when we consider possible packet loss during the broadcasting phase. This issue will be explained in details later in this chapter. Since each time only $\{\hat{Y}_j^j, \hat{N}_j^j\}$ needs to be sent out, the broadcasting can be done within a short time duration τ_0 where $\tau_0 \ll T$. The remaining question is that if no link utilizes the channel within the current block, this broadcasting will not happen and hence if all other links are listening, this duration τ_0 will be wasted. To fix this issue and make the protocol more efficient, we use a different scheme for channel probing.

At the beginning of each block T, all links probe the shared wireless medium according to the outcomings of its local MAB games. For a given link j, j first plays its local MAB game. If the j-th arm is pulled in its local MAB game, link j will probe the wireless medium for a duration of τ; otherwise if the k-th arm is pulled where $k \neq j$, link j will not probe the medium for a duration of τ. In other words, in the beginning of one block T, for each mini-slot τ, link j will probe the wireless medium if and only if the j-th arm is pulled in its local MAB game. If more than one link probes the wireless medium, collisions will happen and this procedure repeats after a duration of τ. If no links probe the wireless medium, the current mini-slot is wasted and the same procedure will repeat for the next mini-slot. If there is only one link probing the wireless medium, this link will utilize the channel for data transmission until a duration of τ_0 immediately before the end of this block. As a result, if the channel rate for link j is R_j in this block, link j then utilizes the channel for data transmission for a total duration of $T - \tau K_j - \tau_0$, which yields a total reward of $R_j(T - \tau K_j - \tau_0)$. Here K_j is the number of mini-slots elapsed for channel probing. It should be noted that K_j is not a geometric distribution anymore since whether link j probes or not depends on its local MAB game. The

for each arm k **do**
  pull arm k once and get reward y_k according to (5.4);
  update initial observations as $\hat{Y}_k^j \leftarrow y_k$ and $\hat{N}_k^j \leftarrow 1$ for all links j;
end for
$n \leftarrow M + 1$;
while 1 **do**
  while no data transmission is detected **do**
    for each link j **do**
      pull the k-th arm that maximizes $\hat{Y}_k^j + \sqrt{\frac{2 \ln n}{\hat{N}_k^j}}$;
      if $k = j$ **then**
        link j probes the medium for a duration of τ;
      else
        link j has a type-F update on $\{\hat{Y}_k^j, \hat{N}_k^j\}$;
      end if
    end for
    if collision detected **then**
      for all links j that probed for the previous duration of τ **do**
        have a type-F update on $\{\hat{Y}_j^j, \hat{N}_j^j\}$;
      end for
    end if
  end while
  the winner link j sends data over the channel until time $T - \tau_0$ within the block, performs a type-T update on $\{\hat{Y}_j^j, \hat{N}_j^j\}$ based on y_j, and broadcasts $\{\hat{Y}_j^j, \hat{N}_j^j\}$ to other links using a duration of τ_0;
  all other links update their local observations on the j-th arm;
  $n \leftarrow n + 1$;
end while

Figure 5.7: The distributed MAB protocol II

whole procedure can be described as Figure 5.7.

We assume these links have no knowledge on the distribution of the channel rates $\{R_m\}$ except that they have a finite support. Without loss of generality, we assume $\{R_m\}$ are properly normalized to a finite support $[0, 1]$. We assume that R_m has a mean μ_m that is unknown to the links, even link m itself. We denote the set of the means as $\vec{\mu} = \{\mu_m\}$.

In the following sections, we characterize the performance of the protocol with respect to the packet loss during the broadcasting phase. We assume the packet loss rate in the broadcasting phase is ϵ. We will discuss this problem for $\epsilon = 0$ and $\epsilon > 0$ respectively.

5.5.2 Performance Analysis: $\epsilon = 0$

If there is no packet loss during the broadcasting phase, for any $j = 1, \ldots, M$, the latest local observations on the j-th arm must be from the j-th link. This is because the local observations $\{\hat{Y}_k^j, \hat{N}_k^j\}$ on the j-th arm cannot be properly updated at link k if $k \neq j$. Hence after the winner of this block j broadcasts $\{\hat{Y}_j^j, \hat{N}_j^j\}$ to other links, the local observations on the j-th arm $\{\hat{Y}_j^j, \hat{N}_j^j\}$ can be updated at all other links as

$$\begin{aligned} \hat{Y}_k^j &= \hat{Y}_j^j \\ \hat{N}_k^j &= \hat{N}_j^j, \end{aligned} \tag{5.14}$$

where $k \neq j$. Hence if all links have the same initial observations at the beginning of the protocol, the local observations at different links are always synchronized by the end of each block duration T as the procedure moves on.

Theorem 5.5.1. *If there is no packet loss during the broadcasting phase, the local observations at different links in the network are always synchronized by the end of each block. The expected regret of the protocol is at most*

$$\left[\frac{8}{T-\tau-\tau_0}\sum_{k:\,\mu_k<\mu^*}\frac{\ln n}{\delta_k}\right]+\left(1+\frac{\pi^2}{3}\right)(T-\tau-\tau_0)\left(\sum_{k:\,\mu_k<\mu^*}\delta_k\right), \tag{5.15}$$

where $\mu^* = \max_k \mu_k$ *and* $\delta_k = \mu^* - \mu_k$.

Proof. We have showed that the local observations at different links are always synchronized by the end of each block. As a result, at the beginning of the next block, all links are playing its own MAB game based on exactly the same local observations $\{\hat{Y}_k^j, \hat{N}_k^j\}_k$. Hence the result of the MAB game is the same, i.e. the same arm will be pulled in all links' local MAB games. We can see that it takes exactly one duration of τ to have some link win the wireless medium.

Suppose the k-th link wins the wireless medium in the current block. The mean reward from pulling the k-th arm is

$$\theta_k = E[Y_k] = E[R_k]\cdot E[T-\tau-\tau_0] \tag{5.16}$$

$$= \mu_k(T-\tau-\tau_0). \tag{5.17}$$

Based on Theorem 1 in [76], the regret of the protocol can be upper bounded by

$$\left[8\sum_{k:\theta_k<\theta^*}\left(\frac{\ln n}{\Delta_k}\right)\right]+(1+\frac{\pi^2}{3})\left(\sum_{k:\theta_k<\theta^*}\Delta_k\right),$$

where $\theta^* = \max_k \theta_k$ and $\Delta_k = \theta^* - \theta_k$. Substitute (5.16) into the above upper bound, we can immediately have our conclusion. □

5.5.3 Performance Analysis: $\epsilon > 0$

The performance analysis in the previous section relies on an important assumption: there is no packet loss in the broadcasting phase when the latest local observations are sent out to all other links from the winner link. To do this, these messages have to be sent via a reliable channel, e.g. a separate signaling channel. This assumption might be a problem for an ad-hoc network due to the lack of resources and infrastructure. In this case, these local observations have to be broadcasted via the shared data channel, and we have to consider the effect from possible packet loss during the broadcasting phase.

Following the system model used in previous sections, we assume that each link is experiencing independent packet losses. We assume that the average packet loss is ϵ during the broadcasting phase where $\epsilon > 0$. It should be noted that the average packet loss rate is different from the winner link's channel quality. It can be explained as follows: during the broadcasting phase, the sender is the winner link which has better channel quality due to the channel-aware scheduling; on the other hand, the receivers of the broadcasted messages are the rest of the links in the network, who are experiencing independent packet loss.

In this section, we mainly address this problem by explicitly considering the packet loss rate ϵ during the broadcasting phase. Another related question is that each link may hold different initial observations in the beginning of the protocol. It also affects the performance of the protocol. We don't explicitly discuss this problem in this chapter. However, this issue can be considered as a special case, where the discrepancy in initial observations can be considered as from independent packet loss during the broadcasting phase.

To characterize the performance of the protocol, we start from several important observations. First of all, if link j wins the wireless medium in this block, any link $k \neq j$ can have its local observations on the j-th arm updated correctly with a probability ϵ. Hence all local observations on the j-th arm within the network can correctly be updated with a probability $(1-\epsilon)^{M-1}$.

Next, we show that any link j has its fair chance to win the medium for infinitely many times. We break the whole proof for this statement into two steps, i.e. Lemma 5.5.1 and Lemma 5.5.2.

Lemma 5.5.1. *For the protocol described in Figure 5.7, there is at least one link j that is able to win the wireless medium for infinitely many times, i.e. $\mathcal{U} \neq \emptyset$.*

Proof. We prove this lemma by contradiction. Suppose the lemma is not true, then any link in the network can only successfully capture the wireless medium for a finite number of times. Hence there exists some integer n_0 such that for $n \geq n_0$, the wireless medium will never be captured by any link anymore.

On the other hand, notice that at the beginning of each block, the wireless medium is successfully captured if and only if there is exactly one link who is probing the wireless medium during that mini-slot; otherwise all links will repeat the same procedure in the next mini-slot. This part of the protocol will repeat until the wireless medium is finally captured. On the other hand, since the duration of one mini-slot τ is much shorter compared to the total block length T, each block will be eventually captured by some link in one block. This contradicts with our assumption. Hence the statement in the Lemma is true. This completes the proof.

□

Lemma 5.5.2. *For the protocol described in Figure 5.7, any link j where $j = 1, \ldots, M$ will win the wireless medium for an infinite number of times, i.e. $\mathcal{V} = \emptyset$.*

Proof. We prove Lemma 5.5.2 by contradiction. Suppose the statement in the lemma is not true, hence we have $\mathcal{V} \neq \emptyset$. On the other hand, we know $\mathcal{U} \neq \emptyset$ from Lemma 5.5.1. We can always find a large enough constant n_1 such that for $n \geq n_1$, no links in $\mathcal{V}$ will win the wireless medium again.

We consider any link $j \in \mathcal{V}$ and any link $k \in \mathcal{U}$. We take a look at the local observations at link j after time $n \geq n_1$. $\hat{Y}_j^j$ will be a constant while only type-F update can be applied to the local observation $\hat{N}_j^j$. On the other hand, $\hat{Y}_k^j$ and $\hat{N}_k^j$ can be updated to $\hat{Y}_k^k$ and $\hat{N}_k^k$ respectively when link k wins the medium and there is no packet loss during the broadcasting phase. Now consider a long enough time duration n_2 starting from $n = n_1$. Since $\mathcal{V} \neq \emptyset$, there will be collisions when the j-th arm is pulled in link j's local MAB game. Hence on average more than one probing will happen in one block before the medium is captured by some link. We take a look at the local MAB game within link j. On average within one block the MAB game will be played multiple times before the medium is captured by some link. The difference is how $\hat{N}_k^j$ is updated over I blocks. If $i \in \mathcal{V}$, whenever the i-th arm is pulled in j's local MAB game, $\hat{N}_i^j$ is increased by 1. However if $k \in \mathcal{U}$, $\hat{N}_k^j$ is only increased by 1 when link k wins the medium. And this type of update can happen I times in total for all $k \in \mathcal{U}$ over the I blocks. As a result, we can see on average $\hat{N}_j^j$ increases faster than $\hat{N}_k^j$.

On the other hand, we notice that $\hat{Y}_j^j$ is a constant. Hence if $\hat{N}_j^j$ is larger than other arms $\hat{N}_k^j$, the following inequality

$$\hat{Y}_j^j + \sqrt{\frac{2\ln n}{\hat{N}_j^j}} > \hat{Y}_k^j + \sqrt{\frac{2\ln n}{\hat{N}_k^j}}$$

cannot hold for many times. This leads to a contradiction. Hence the assumption $\mathcal{V} \neq \emptyset$ cannot be true. □

We have showed that each link j has its fair chance to win the wireless medium and update its local observations properly for infinitely many times. Based on this, we show that the local MAB games at all links are synchronized in many cases except a few exceptions.

Lemma 5.5.3. *The local MAB games at all links are synchronized after sufficient number of blocks in many cases except a few exceptions.*

Proof. To prove this lemma, we need to show that for a given $k = 1, \ldots, M$, the local observations on this arm at different links will converge when the number of blocks n is large enough. We take a look at two different links i and j where $i \neq j$. Since the local MAB game is played at i and j independently, the following two quantities

$$\hat{Y}_k^i + \sqrt{\frac{2\ln n}{\hat{N}_k^i}}$$
$$\hat{Y}_k^j + \sqrt{\frac{2\ln n}{\hat{N}_k^j}}$$

will be checked for decision-making respectively. To show this quantity converges with respect to the index of the links, we take a look at the difference of the above quantity at two different links i and j as

$$\left(\hat{Y}_k^i + \sqrt{\frac{2\ln n}{\hat{N}_k^i}}\right) - \left(\hat{Y}_k^j + \sqrt{\frac{2\ln n}{\hat{N}_k^j}}\right). \tag{5.18}$$

The above equation can be rewritten as

$$\begin{aligned} & \left(\hat{Y}_k^i - \hat{Y}_k^j\right) + \left(\sqrt{\frac{2\ln n}{\hat{N}_k^i}} - \sqrt{\frac{2\ln n}{\hat{N}_k^j}}\right) \\ = & \left(\hat{Y}_k^i - \hat{Y}_k^j\right) + \frac{\sqrt{2\ln n}}{\sqrt{\hat{N}_k^i} + \sqrt{\hat{N}_k^j}} \cdot \frac{1}{\sqrt{\hat{N}_k^i} \cdot \sqrt{\hat{N}_k^j}} \cdot \left(\hat{N}_k^j - \hat{N}_k^i\right). \end{aligned} \tag{5.19}$$

We compare the first and second item in the above summation, focusing on how fast they decay with respect to n. We can roughly approximate them using the following approach.

First of all, we have showed that each link j has its fair chance to win the wireless medium and update its local observations for infinitely many times. Hence by the n-th block, the average number of times that each arm being pulled should be roughly proportional to n. As a result, we can approximate them as

$$\begin{aligned} \hat{N}_k^i &\approx \alpha \cdot \frac{n}{M}, \\ \hat{N}_k^j &\approx \beta \cdot \frac{n}{M}. \end{aligned}$$

Furthermore, for the same arm k, whenever the k-th link wins the wireless medium, both $\hat{N}_k^j$ and $\hat{N}_k^i$ have a chance to update itself to the latest value with a probability $1-\epsilon$. Hence both of them can be bounded as

$$\left\|\hat{N}_k^j - \hat{N}_k^k\right\| \le C_1,$$
$$\left\|\hat{N}_k^i - \hat{N}_k^k\right\| \le C_1.$$

Substituting them into the second item of (5.19), we can approximate the quantity as

$$\begin{aligned}\frac{\sqrt{2\ln n}}{(\alpha+\beta)\sqrt{\frac{n}{M}}}\cdot\frac{1}{\sqrt{\alpha\beta\cdot\frac{n}{M}}}\cdot\left(\hat{N}_k^j-\hat{N}_k^i\right) &\le \frac{\sqrt{2\ln n}}{(\alpha+\beta)\sqrt{\frac{n}{M}}}\cdot\frac{1}{\sqrt{\alpha\beta\cdot\frac{n}{M}}}\cdot 2C_1\\ &= \frac{\sqrt{2}}{(\alpha+\beta)\sqrt{\alpha\beta}}\cdot M\sqrt{M}\cdot\sqrt{\frac{\ln n}{n}}\cdot\frac{1}{n}\cdot C_1.\end{aligned} \tag{5.20}$$

On the other hand, since both $\hat{Y}_k^j$ and $\hat{Y}_k^i$ are the average of the local observations on the k-th arm so far, we can approximate their difference as

$$\hat{Y}_k^j - \hat{Y}_k^i \approx \frac{1}{n}\cdot C_2, \tag{5.21}$$

where C_2 is a constant.

Hence we can see that when n is large, the second item in (5.19) decays much faster than the first item. Hence (5.19) is dominated by the first item. On the other hand, for a given k, both $\hat{Y}_k^j$ and $\hat{Y}_k^i$ are the average of the original observations on the k-th arm, and the number of observations available wouldn't differ too much, i.e. $\left\|\hat{N}_k^i - \hat{N}_k^j\right\| \le 2C_1$. Hence $\hat{Y}_k^j$ and $\hat{Y}_k^i$ will converge to the true mean of $\hat{Y}_k^k$. As a result, the local MAB game at all links are synchronized in many cases except a few exceptions. □

Based on all the results above, we can put them together to get an upper bound of the regret of the distributed MAB protocol.

Theorem 5.5.2. *If there is an average packet loss rate ϵ during the broadcasting phase of the protocol shown in Figure 5.7, the local observations at different links in the network are synchronized in many cases with few exceptions after a sufficient number of blocks n. The expected regret of the protocol is at most*

$$\begin{aligned} A + &\left[\frac{8}{T-(1+2M\epsilon)\tau-\tau_0}\sum_{k:\,\mu_k<\mu^*}\frac{\ln(n-n_0)}{\delta_k}\right] \\ &+\left(1+\frac{\pi^2}{3}\right)[T-(1+2M\epsilon)\tau-\tau_0]\left(\sum_{k:\,\mu_k<\mu^*}\delta_k\right), \end{aligned} \tag{5.22}$$

where $\mu^* = \max_k \mu_k$ and $\delta_k = \mu^* - \mu_k$.

Proof. In Lemma 5.5.3, we have showed that after the protocol has been running for a sufficient number of blocks n, the local observations at different links are synchronized in many cases with few exceptions. Hence there exists some integer $n_0 > 0$ such that for $n \geq n_0$, the local observations at different links will converge to $\{\hat{Y}_k, \hat{N}_k\}_k$ for any arm $k = 1, \ldots, M$.

Now starting from this point $n = n_0$, we can view the protocol as a centralized MAB game, where the MAB is based on the converged observations $\{\hat{Y}_k, \hat{N}_k\}_k$. It should be noted that $\{\hat{Y}_k, \hat{N}_k\}_k$ are the converged versions of the local observations, which do not physically exist within any link j.

To estimate the regret of the protocol for $n > n_0$, we can split it into two parts. The first part comes from the time interval between $n = 0$ and $n = n_0$. Since n_0 is a fixed finite number, we can assume that at time n_0 the regret of the system is $\mathcal{R}(n_0) = A$. The second part comes from the time interval after n_0. If we consider time $n = n_0$ as the new starting point, based on Theorem 1 in [76] the regret of the protocol in this part can be upper bouded by

$$\begin{aligned}
&\left[8 \sum_{k:\, \tilde{\theta}_k < \tilde{\theta}^*} \left(\frac{\ln \tilde{n}}{\tilde{\Delta}_k}\right)\right] + (1 + \frac{\pi^2}{3}) \left(\sum_{k:\, \tilde{\theta}_k < \tilde{\theta}^*} \tilde{\Delta}_k\right) \\
=& \left[8 \sum_{k:\, \tilde{\theta}_k < \tilde{\theta}^*} \left(\frac{\ln(n - n_0)}{\tilde{\Delta}_k}\right)\right] + (1 + \frac{\pi^2}{3}) \left(\sum_{k:\, \tilde{\theta}_k < \tilde{\theta}^*} \tilde{\Delta}_k\right). \qquad (5.23)
\end{aligned}$$

The rest of our job is to characterize those constants in the above bound, i.e. $\tilde{\theta}_k$ and $\tilde{\Delta}_k$. It should be noted that these constants are different from those in Theorem 5.5.1, since the local observations at different links will not always be the same as the converged observations $\{\hat{Y}_k, \hat{N}_k\}_k$. This type of random unsynchronization will lead to a higher regret compared to the ideal case in Theorem 5.5.1.

To estimate these constants, we consider the following worst case scenario. Assume that at some point, there are u links whose local observations have lost synchronization compared to the converged versions. Notice that if $u = 0$, i.e. the local observations at all links are synchronized, then it takes only one mini-slot for channel probing. Now suppose $u \geq 1$, we are interested in knowing the number of mini-slots it takes before the wireless medium is successfully captured by some link. First of all, the rest of the $M - u$ links will always yield the same decisions during their local MAB games. Suppose in the current mini-slot the k-th arm is pulled in these $M - u$ links. The k-th link will capture the wireless medium successfully unless one of the following happens: 1) some other link k_1 also probes the wireless medium according to its own MAB game based on its unsynchronized local observations, hence a collision in the medium; 2) link k is not among these u links, hence a vacancy in the medium. It should be noted that no $\hat{Y}_k^j$ will be updated in this block before the wireless medium is actually captured by some link. Hence this type of behaviors are deterministic within this block. Now consider the worst case scenario: starting from the first mini-slot, every time either a collision or a vacancy occurs on the wireless medium. It will last for at most $2u + 1$ mini-slots.

On the other hand, in Lemma 5.5.3 we have showed that the local observations at all links will converge after the protocol has been running for a long enough time. From this point any link j will lose its synchronization if and only if there is a packet loss during the last broadcasting phase. All links in the network have an independent packet loss rate ϵ except the current winner of the wireless medium. Hence the probability that any link will have its local observations unsynchronized is also ϵ, and this probability is independent from link to link. As a result, the probability that u links out of M links have lost their synchronizations is

$$P(u) = \binom{M}{u} \epsilon^u (1-\epsilon)^{M-u}. \tag{5.24}$$

Hence, if we iterate u from $u = 0$ to $u = M$, we can add them together and get an upper bound of the elapsed number of time slots for channel probing as

$$\overline{K}_s = \sum_{u=0}^{M} (2u+1) \cdot \binom{M}{u} \epsilon^u (1-\epsilon)^{M-u} \tag{5.25}$$

$$= 1 + 2M\epsilon \tag{5.26}$$

Now we characterize those constants $\tilde{\theta}_k$ and $\tilde{\Delta}_k$. Suppose the k-th link is the winner of the wireless medium in the current block. Due to the randomness in the MAB game and corresponding channel probing, it may take any number of probings before the wireless medium is won by the k-th link in this block. Suppose it takes

$\tilde{K}_k$ probings before the wireless medium is won by the k-th link, and the available channel rate for the k-th link in the current block is $\tilde{R}_k$. The reward from sending data in this block is $\tilde{R}_k(T - \tau\tilde{K}_k - \tau_0)$. To characterize an upper bound on the regret, we need to calculate the average reward $\tilde{\theta}_k$ from pulling the k-th arm. By taking its mathematical expectation, we have

$$\begin{aligned}\tilde{\theta}_k &= E\left[\tilde{R}_k \cdot (T - \tau\tilde{K}_k - \tau_0)\right] \\ &= E\left[\tilde{R}_k\right] \cdot \left(T - \tau E[\tilde{K}_k] - \tau_0\right), \end{aligned} \tag{5.27}$$

where we have used the independence of the elapsed probing durations $\tilde{K}_k$ and the available channel rate $\tilde{R}_k$. Note that $\tilde{R}_k = \mu_k$ since μ_k is the average channel rate for the k-th link. On the other hand, $E[\tilde{K}_k]$ is the average number of probings elapsed before the k-th link wins the channel. For the upper bound of the regret, we can use the worst case scenario $\overline{K}_s$ instead. Put them together, we have

$$\begin{aligned}\tilde{\theta}_k &= \mu_k(T - \tau\overline{K}_s - \tau_0) \\ &= \mu_k\left[T - (1 + 2M\epsilon)\tau - \tau_0\right]. \end{aligned} \tag{5.28}$$

Substituting the above equation into (5.23), we can immediately have our conclusion.

□

5.6 Numerical Results

In this section, we show numerical results for our proposed distributed MAB protocols. We use theoretical bound (5.6) and numerical results of the semi-distributed MAB protocol as our performance baseline. This is because we have used ideal assumptions in the semi-distributed MAB protocol.

We consider an ad-hoc network with M links. The mean value of the channel rate for link j is

$$\delta_j = 0.3 + \frac{0.4}{M-1} \cdot (j-1) \quad \text{bits/s/Hz},$$

where $j = 1, \ldots, M$. Since the UCB1 policy [76] works for arbitrary reward distribution with a finite support, without loss of generality we assume that the channel rate R_j has a Beta distribution $\text{Be}(\delta_j, 1-\delta_j)$, which yields exactly $E[R_j] = \delta_j$. Throughout all simulations in this section, we use $T = 1$ and $\tau = 0.05$. We compare performance of the distributed MAB protocols with varying M and p. The default set of parameters used in our simulations is $M = 20$ and $p = 1/M$.

We first take a look at the statistics of the local observations that are held at link $j = 1, \ldots, M$.

First of all, we need to show that in Protocol B the local observations $\hat{Y}_k^j(n)$ converge with respect to j. This is because unlike Protocol A, each link j in Protocol B has to keep its own local observations separately. In this paper, we compare the relative standard deviations (RSD) in Figure 5.8(a) and Figure 5.8(b). The RSD is defined as the ratio of the standard deviation to the mean of $\hat{Y}_k^j(n)$ with respect to j. In Figure 5.8(a) we show the RSD of $\hat{Y}_k^j(n)$ for the 8-th arm. We can see the RSDs are decreasing over time both with homogeneous and heterogeneous initializations. This shows that the mean reward observed for a given arm will quickly converge across different links. On the other hand, we can see with homogeneous initializations the RSD decreases much faster compared to that with heterogeneous initializations. For

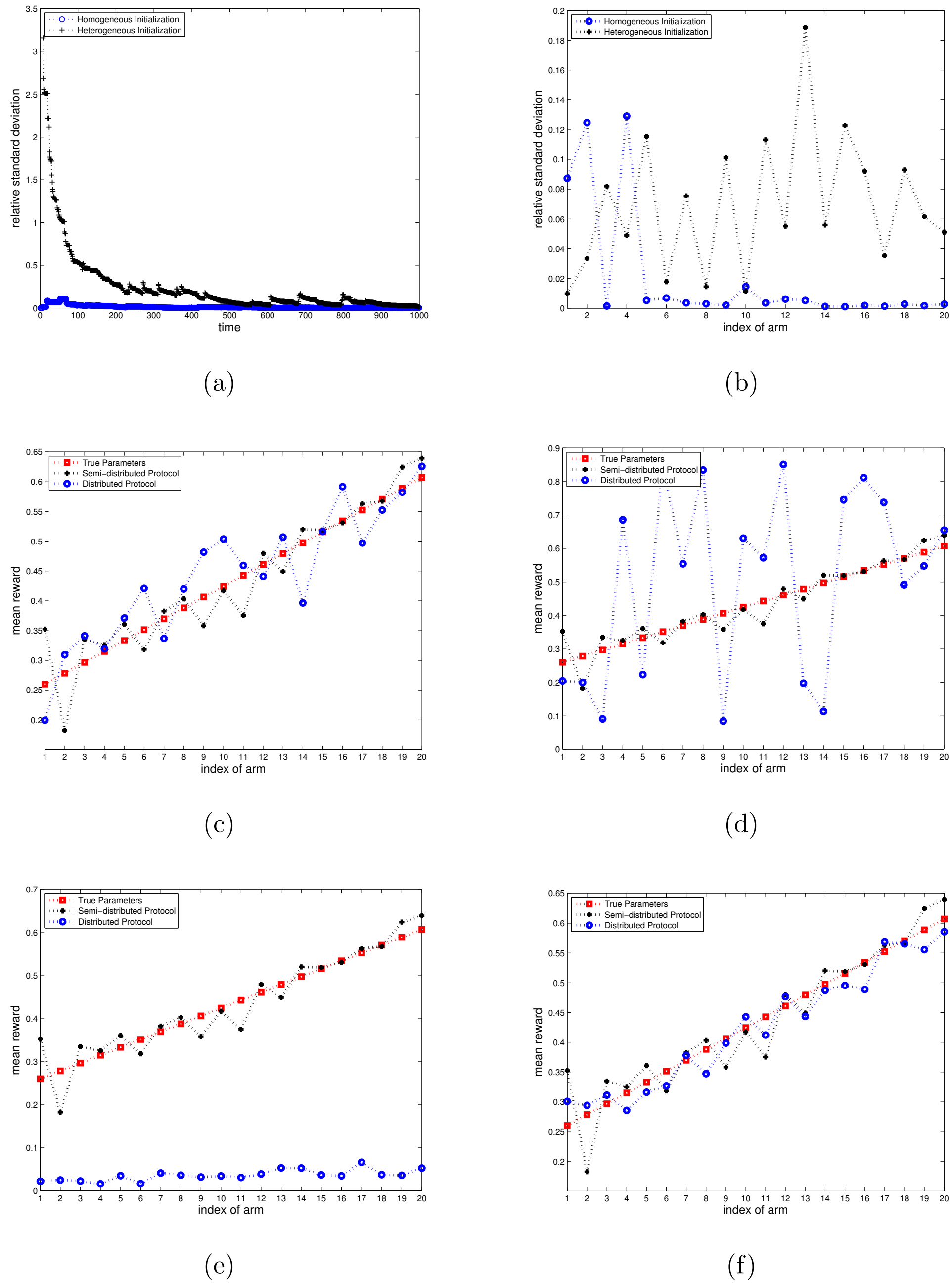


Figure 5.8: Statistics of the local observations held independently among all links $j = 1, \ldots, M$, where $M = 20$ and $p = 1/M$. (a) RSD of $\hat{Y}_k^j(n)$ from Protocol B for the 8-th arm; (b) RSD of $\hat{Y}_k^j(n)$ from Protocol B at time $n = 1000$; (c) $\hat{Y}_k^A(n)$ from Protocol A at time $n = 1000$; (d) Average of $\hat{Y}_k^j(n)$ from Protocol B with homogeneous initialization; (e) Average of $\hat{Y}_k^j(n)$ from Protocol B with heterogeneous initialization; (f) Average of $\hat{Y}_k^j(n)$ from 100 independent simulation runs of Protocol B with homogeneous initialization.

example, with homogeneous initialization, the RSD converges as early as $n = 100$, while with heterogeneous initialization the RSD starts to converge at $n = 500$. In Figure 5.8(b) we show the RSD of $\hat{Y}_k^j(n)$ at time $n = 1000$. We can see the RSDs stay below 20% with both initializations. In fact, with homogeneous initialization, the RSD of $\hat{Y}_k^j(n)$ stays below 2% for most arms, except for three arms which goes up to 13%. With heterogeneous initialization, the RSD stays below 12% for most arms, except for the 13-th arm. Again we notice that with homogeneous initialization, Protocol B yields a smaller RSD compared to that with heterogeneous initialization.

We then compare average of the local observations $\hat{Y}_k^j(n)$ to that from the semi-distributed MAB protocol as well as the true parameters. In Figure 5.8(c) we show $\hat{Y}_k^A(n)$ at time $n = 1000$, since this is the local observation kept at $\mathcal{A}$. We can see the rewards $\hat{Y}_k(n)$ from the semi-distributed MAB protocol at $n = 1000$ show a correct trend and they are very close to the true parameters used in our simulation. The observations $\hat{Y}_k^A(n)$ of Protocol A at time $n = 1000$ still show a quite consistent trend except some deviations that are slightly greater than that of the semi-distributed protocol. In Figure 5.8(d) - Figure 5.8(f) we show average of the local observations $\hat{Y}_k^j(n)$ of Protocol B with respect to all links j at time $n = 1000$.[4] From Figure 5.8(d), we can see that even with homogeneous initialization the average of the local observations $\hat{Y}_k^j(n)$ show quite a large jitter at some points, e.g. the 10-th arm. This suggests that average of the local observations from only one simulation run might not reflect the true parameters correctly even after they converge. In

[4]One review pointed out that the convergence in Figure 5.8(d) is not convicing. I will need to check numerical results under more different parameter settings.

contrast, Figure 5.8(f) shows the results from averaging 100 independent simulation runs of Protocol B. We can see that the average observations are very close to that of the semi-distributed MAB protocol. On the other hand, Figure 5.8(e) shows average of $\hat{Y}_k^j(n)$ from Protocol B with heterogeneous initialization at time $n = 1000$. We can see the curves show a much smoother trend compared to Figure 5.8(d), even though the average are not even close to the true parameters. This is because with heterogeneous initialization, each link j holds its own observations to start with, which brings robustness to Protocol B.

We then take a look at the regrets of the semi-distributed and our proposed distributed MAB protocols.

In Figure 5.9(a) we compare theoretical bound (5.6) and simulation result for the regret of the semi-distributed MAB protocol. We can see the theoretical bound does show a $\ln n$ trend. However the theoretical bound is quite loose compared to our simulation results.

In Figure 5.9(b) - Figure 5.9(f), we compare regret of the semi-distributed MAB protocol to that of our proposed distributed MAB protocols. We first show regret of Protocol A in Figure 5.9(b). We can see Protocol A yields a regret larger than but with similar trend compared to that of the semi-distributed MAB protocol. For example, at time $n = 1000$, the regret of Protocol A is almost 50% higher than that of the semi-distributed MAB protocol. Protocol B with homogeneous initialization yields a regret close to that of Protocol A in Figure 5.9(c). In contrast, with heterogeneous initialization Protocol B yields a much higher regret in Figure 5.9(d). For example, at time $n = 1000$, the regret of Protocol B is 160% higher

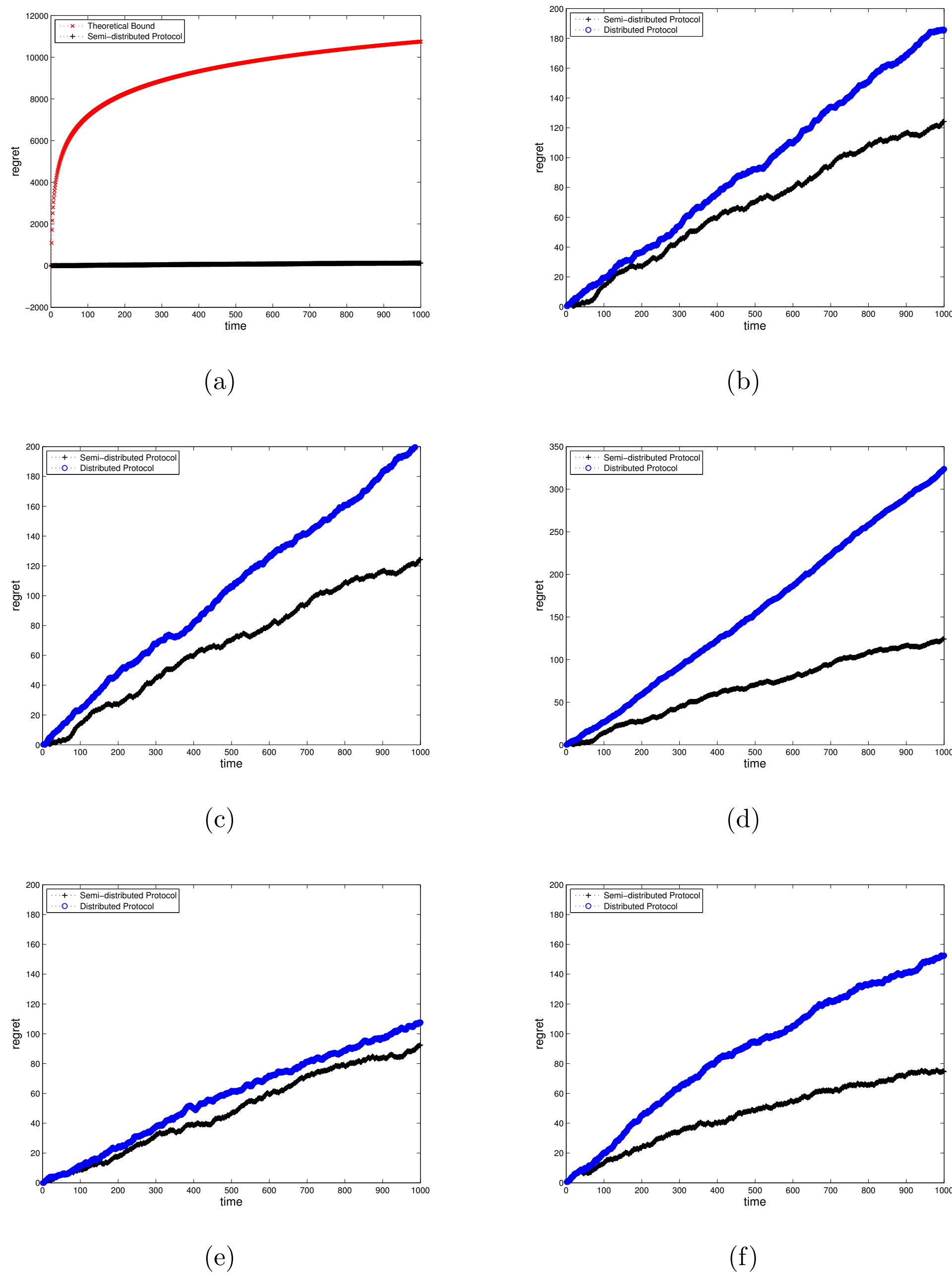
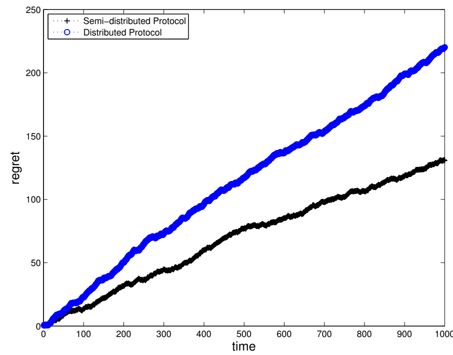


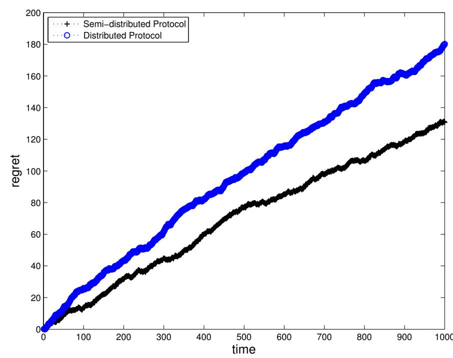
Figure 5.9: Regrets of the semi-distributed MAB protocol vs. that of the distributed MAB protocols, where $M = 20$ and $p = 1/M$ unless explicitly indicated in the sub-figure. (a) Theoretical bound vs. simulation results for regret of the semi-distributed MAB protocol; (b) Regret of Protocol A; (c) Regret of Protocol B with homogeneous initialization; (d) Regret of Protocol B with heterogeneous initialization; (e) Regret of Protocol B with homogeneous initialization where $p = 0.1$; (f) Regret of Protocol B with homogeneous initialization where $M = 10$.

than that of the semi-distributed MAB protocol. Notice all the above simulations use $M = 20$ and $p = 1/M$ so far. In Figure 5.9(e) and Figure 5.9(f), we show different parameters affected regrets of Protocol B with homogeneous initialization. Figure 5.9(e) shows regret of Protocol B with a different probing probability $p = 0.1$. We can see both the semi-distributed protocol and Protocol B yield a much smaller regret. Furthermore, the gap between these two is reduced to around 20%. This shows that the distributed MAB protocol is able to reduce contention in the wireless medium greatly by trying to let selected links access the medium with higher probability. Figure 5.9(f) shows regret of Protocol B with a different number of links $M = 10$. We can see with a smaller M, both regrets are reduced. This is due to reduced contention in the wireless medium.

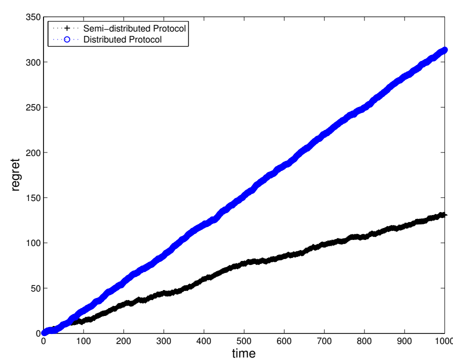
In all these simulations, we can see that Protocol A yields a better performance compared to Protocol B. This is because the local observations $\hat{Y}_k^A(n)$ is reliable. Now we assume that at some time n_I, when link j wins the wireless medium but finds out that the observations at $\mathcal{A}$ or $\mathcal{B}$ is lost. Hence link j copies its own observations to $\mathcal{A}$ or $\mathcal{B}$ and then continues the procedure. The regrets are shown in Figure 5.10 for $n_I = 50$. We can see that with $\hat{Y}_k^A(n)$ lost at $n_I = 50$, Protocol A yields a much larger regret in Figure 5.10(a). On the other hand, Protocol B still yields almost the same regrets in Figure 5.10(b) and Figure 5.10(c).



(a)



(b)



(c)

Figure 5.10: Regrets of the semi-distributed MAB protocol vs. that of the distributed MAB protocols, where there is a one-time data loss within $\mathcal{A}$ or $\mathcal{B}$ at time $n_I = 50$. (a) Regret of Protocol A; (b) Regret of Protocol B with homogeneous initialization; (c) Regret of Protocol B with heterogeneous initialization.

Chapter 6

Conclusions

In Chapter 3, we studied a distributed opportunistic scheduling problem for wireless ad-hoc networks under the popular block fading model. In this problem, we considered the inevitable dependencies between winners' channel rates at different time instances during the channel probing phase and their impact on the transmission scheduling. We formulated this problem using optimal stopping theory, but at carefully chosen time instances when effective decisions are made by merging repeated decisions. We studied this problem for both the CAT and CDT models. We first analyzed the problem assuming it has infinite stages, and then developed a measure to check how well the problem can be treated as an infinite horizon problem. We estimated the achievable network throughput if we ignore the finite horizon constraint and use the stopping rule based on the infinite horizon analysis nevertheless. If the finite horizon constraint cannot be ignored, we characterized its performance using backward induction. We presented one recursive approach to reduce its computational overhead and derived an upper bound for its computational complexity. We also presented an improved protocol to reduce the probing costs which requires no additional design cost. We showed numerical results for networks with various sizes under different settings of the parameters.

In Chapter 4, we extended this problem to ad-hoc networks where the wireless spectrum can be divided into multiple independent sub-channels for better efficiency. We showed that a naive protocol where the opportunistic scheduling is designed independently within each sub-channel can only slightly improve the system throughput. We then came up with the idea of opportunistic scheduling across multiple sub-channels. We developed a multi-channel protocol for ad-hoc networks and analyzed its performance. We characterized the optimal decision rule and the system throughput. Through numerical results we showed that by joint optimization of the scheduling schemes across multiple sub-channels, the proposed protocol improves the network throughput considerably.

In Chapter 5, we revisited the channel-aware scheduling problem under the assumption that neither the instant CSI nor the channel statistics are known. We formulated the problem using multi-armed bandit (MAB). We first presented a semi-distributed MAB protocol which serves as the baseline for performance comparison. We then proposed two forms of distributed MAB protocols, where each link keeps a local copy of the observations and plays the MAB game independently. In Protocol I the MAB game is only played once within each block, while in Protocol II it can be played multiple times. We showed that the proposed distributed protocols can be considered as a generalized MAB procedure and each link is able to update its local copy of the observations for infinitely many times. We characterized the evolution of the local observations and the regrets of the system. For Protocol I, we showed by simulation results that the local observations that are held independently at each link converge to the true parameters and the regret is comparable to that of the semi-distributed protocol. For Protocol II, we proved the convergence of the local observations and show an upper bound of the regret.

Appendix A

Optimal Stopping Theory

In this chapter, we briefly introduce some results used in this work from optimal stopping theory. We follow notations used in [55]. A detailed introduction can be found in [55–57, 80].

The theory of optimal stopping is concerned with the problem of choosing a time to take a given action based on sequentially observed random variables in order to maximize an expected payoff. An optimal stopping rule is a strategy for deciding when to take a given action based on the past events in order to maximize the average return, where the return is the net gain, i.e. the difference between the reward and the cost. Generally speaking, a stopping rule problem is defined by two objects:

(a) a sequence of random variables $X_1, X_2, \ldots$ whose joint distribution is assumed known, and

(b) a sequence of real-valued reward functions,

$$y_0, y_1(x_1), \ldots, y_\infty(x_1, x_2, \ldots). \tag{A.1}$$

Suppose the process starts at time 0. If we do not take any observations, we receive a constant amount y_0. At time n, after observing $X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n$, we may stop and receive a known reward $y_n(x_1, \ldots, x_n)$, or we may continue and observe X_{n+1}. If we never stop, we receive $y_\infty(x_1, x_2, \ldots)$. Hence the problem is to choose a time n^* to stop to maximize the expected reward. We may use randomized decisions, i.e. given observations up to stage n as $X_1 = x_1, \ldots, X_n = x_n$, we choose a probability of stopping $\phi_n(x_1, \ldots, x_n)$. Accordingly, we denote the probability mass function (PMF) of N given $\mathbf{X} = \mathbf{x} = (x_1, x_2, \ldots)$ by $\psi = (\psi_0, \psi_1, \ldots, \psi_\infty)$, i.e.

$$\psi_n(x_1, \ldots, x_n) = P(N = n|\mathbf{X} = \mathbf{x}), \tag{A.2}$$

$$\psi_\infty(x_1, x_2, \ldots) = P(N = \infty|\mathbf{X} = \mathbf{x}). \tag{A.3}$$

The problem is then to choose a stopping rule ϕ to maximize the expected return defined as

$$\begin{aligned} V(\phi) &= E_{y_N}(X_1, \ldots, X_N) \\ &= \sum_{j=0}^{\infty} \psi_j(X_1, \ldots, X_j) \cdot y_j(X_1, \ldots, X_j) \end{aligned} \tag{A.4}$$

In some applications, the reward sequence is described as a sequence of random variables $Y_0, Y_1, \ldots, Y_\infty$ whose joint distribution with observations $X_1, X_2, \ldots, X_\infty$ is known. The actual value of Y_n may not be known precisely at time n when the decision to stop or continue must be made. Allowing returns to be random does not represent a gain in general because since the decision to stop at time n may depend on $X_1, \ldots, X_n$, we may replace the sequence of random rewards Y_n by the sequence of reward functions $y_n(x_1, \ldots, x_n)$ for $n = 0, 1, \ldots, \infty$ as

$$y_n(x_1, \ldots, x_n) = E\{Y_n|X_1 = x_1, \ldots, X_n = x_n\}. \tag{A.5}$$

The increasing sequence of σ-field approach is a simpler, more widely used notation to model optimal stopping problems. Let $(\Omega, \mathcal{B}, P)$ denote the probability space on which our random variables are defined, and let $\mathcal{F}_n$ denote the sub-σ-field of $\mathcal{B}$ generated by $X_1, \ldots, X_n$, i.e. the smallest σ-field containing the sets $\{X_1 \leq x_1, \ldots, X_n \leq x_n\}$ for all $x_1, \ldots, x_n$. Hence we have $\mathcal{F}_0 = \{\Omega, \emptyset\}$ and $\mathcal{F}_\infty$ which is equivalent to the σ-field generated by $\cup \mathcal{F}_n$, where

$$\mathcal{F}_0 \subset \mathcal{F}_1 \subset \ldots \subset \mathcal{F}_n \subset \ldots \subset \mathcal{F}_\infty \subset \mathcal{B} \tag{A.6}$$

is an increasing sequence of σ-fields. For an arbitrary random variable Z, the conditional expectation of Z given $X_1, \ldots, X_n$ may be denoted by

$$E\{Z|\mathcal{F}_n\} = E\{Z|X_1, \ldots, X_n\}.$$

The stopping rule problem may be stated in terms of the sequence of σ-fields (A.6) instead of the random variables $X_1, X_2, \ldots$ as being defined by the following

(a') the increasing sequence of σ-fields (A.6), and

(b') a sequence of reward random variables $Y_0, Y_1, \ldots, Y_\infty$.

Using σ-fields, a stopping rule is defined to be a random variable N taking values in $\{0, 1, \ldots, \infty\}$ such that the event $\{N = n\}$ is in $\mathcal{F}_n$. So the decision to stop at time n depends only on $X_1, \ldots, X_n$ and is independent of future observations. Hence the problem is to choose a stopping rule N to maximize the expected return $E\{Y_N\}$. This is a more general approach since there exist σ-fields that are not generated by any sequence of random variables.[1]

There are two different types of optimal stopping problems.

[1]Some generality is lost here since the stopping rules defined by the σ-fields are non-randomized. However, we may restrict our attention to non-randomized stopping rules. To see this, we attach

A.1 Finite Horizon Problem

If there is a known upper bound on the number of stages at which one may stop, the stopping rule problem is a finite horizon problem. If stopping is required after observing $X_1, \ldots, X_T$, we say the problem has horizon T, which can be solved by the method of *backward induction*. Define $V_T^{(T)}(x_1, \ldots, x_T) = y_T(x_1, \ldots, x_T)$ and then inductively for $j = T - 1$, backward to $j = 0$,

$$V_j^{(T)}(x_1, \ldots, x_j) = \max \Big\{ y_j(x_1, \ldots, x_j), \\ E\left[V_{j+1}^{(T)}(x_1, \ldots, x_j, X_{j+1}) | X_1 = x_1, \ldots, X_j = x_j \right] \Big\}, \tag{A.7}$$

where $V_j^{(T)}(x_1, \ldots, x_j)$ represents the maximum return one can obtain starting from stage j having observed $X_1 = x_1, \ldots, X_j = x_j$. The optimal return is therefore the maximum of these two quantities, and it is optimal to stop at j if $V_j^{(T)}(x_1, \ldots, x_j) = y_j(x_1, \ldots, x_j)$, and to continue otherwise. The optimal value of the stopping rule problem is then $V_0^{(T)}$.

A.2 Infinite Horizon Problem

If there is no bound on the number of stages, i.e. we consider a stopping rule problem with observations $X_1, X_2, \ldots$ and rewards $Y_0, Y_1, \ldots, Y_\infty$ where $Y_n =$

an independent uniform $(0, 1)$ random variable U_j to each X_j. For a given stopping rule ϕ we can form an equivalent non-randomized stopping rule by stopping at time j when we reach it if $U_j < \phi_j(X_1, \ldots, X_j)$.

$y_n(X_1, \ldots, X_n)$. Under the following two assumptions

$$\text{A1.} \quad E\{\sup_n Y_n\} < \infty, \tag{A.8}$$

$$\text{A2.} \quad \limsup_{n\to\infty} Y_n \leq Y_\infty \text{ a.s.}, \tag{A.9}$$

we have the principle of optimality

$$V_n^*(x_1, \ldots, x_n) = \sup_{N\geq n} E\{Y_N | X_1 = x_1, \ldots, X_n = x_n\}, \tag{A.10}$$

where $\sup_{N\geq n}$ means supremum over the set of all stopping rules N such that $P(N \geq n) = 1$. It is optimal to stop at stage n having observed $X_1 = x_1, \ldots, X_n = x_n$ if and only if $y_n(x_1, \ldots, x_n) = V_n^*(x_1, \ldots, x_n)$.

The optimal return can be computed by the optimality equation

$$V_n^* = \max\{Y_n, E(V_{n+1}^* | \mathcal{F}_n)\}, \tag{A.11}$$

where the optimal return is re-defined using the concept of *essential supremum* as

$$V_n^* = \operatorname{ess\,sup}_{N\geq n} E\{Y_N | \mathcal{F}_n\}, \tag{A.12}$$

so that it can also handle problems with an uncountable collection of stopping rules.

Appendix B

Proof of Theorems in Chapter 3

B.1 Proof of Theorem 3.7.2

Similar to the CAT problem, we solve this problem as a maximal rate of return problem. For a fixed rate $\lambda > 0$, we define a new payoff at time n as

$$V_n(\lambda) = R_n T - \lambda \left(T + \tau \sum_{i=1}^{n} K_i \right). \tag{B.1}$$

To show the existence of the optimal rule, we first notice that $E\{\sup_n V_n\} < \infty$. On the other hand, we can see that $\limsup_{n\to\infty} V_n \to -\infty$ and $V_n \to -\infty$ *a.s.*. Putting them together leads to $\limsup_{n\to\infty} V_n \to V_\infty$ *a.s.*. Hence an optimal stopping rule exists and can be given by the optimality equation. Note that we used the equation $K_i = \frac{M}{M-i+1}\tilde{K}_i$ in the proof of Theorem 3.5.1. If we substitute it into (B.1) and notice the *i.i.d.* property of $\tilde{K}_i$, we can rewrite (B.1) as

$$V_n(\lambda) = R_n T - \lambda T - \lambda \tau \tilde{K}_1 \sum_{i=1}^{n} \frac{M}{M-i+1}.$$

The above equation should be understood *in distribution*. By taking the average of M and $M-n+1$, we approximate $V_n(\lambda)$ as

$$V_n(\lambda) \approx R_n T - \lambda T - \lambda \tau \tilde{K}_1 \cdot \frac{Mn}{M - n/2 + 0.5}.$$

Similarly, the payoff at time $n+1$ can be written as

$$V_{n+1}(\lambda) \approx R_{n+1} T - \lambda T - \lambda \tau \tilde{K}_1 \cdot \frac{M(n+1)}{M - (n+1)/2 + 0.5}.$$

Meanwhile, note that R_n are *i.i.d.* according to [A4]. Hence in the sense of distribution the difference between $V_n(\lambda)$ and $V_{n+1}(\lambda)$ can be written as

$$\lambda\tau\tilde{K}_1 \cdot \frac{M}{M+0.5}\left[\frac{n+1}{1-\frac{(n+1)/2}{M+0.5}} - \frac{n}{1-\frac{n/2}{M+0.5}}\right].$$

The item in the above square bracket has been calculated as (3.22) in the proof of Theorem 3.5.1. If we substitute it into the optimality equation, we have

$$V_n^*(\lambda) = E\left[\max\left\{R_nT - \lambda T - \lambda\tau\sum_{i=1}^{n} K_i, V_n^*(\lambda) - \frac{M(M+n+1)}{(M+0.5)^2}\cdot\lambda\tau\tilde{K}_1\right\}\right].$$

The optimal rate λ_n^* that maximizes the rate of return should yield $V_n^*(\lambda_n^*) = 0$. If we substitute it into the optimality equation and notice $E[\tilde{K}_1] = 1/p_{s,1}$, we immediately have (3.45). The uniqueness of λ_n^* can be verified easily. The optimal stopping rule can be written as

$$N^* = \min\left\{n \geq 1 : R_nT - \lambda_n^*T - \lambda_n^*\tau\sum_{i=1}^{n} K_i \geq V_n^*(\lambda_n^*) = 0\right\},$$

which immediately leads to (3.44). If we let $n = 0$ in (3.45), we get (3.43). The solution of (3.43) is the optimal system throughput λ_O^*.

B.2 Proof of Theorem 3.7.4

We use V_n defined in (B.1) in our proof. The existence of the optimal stopping rule can be verified in the same way as Theorem 3.7.2. To compute the optimal payoff V_n^*, we take a look at the payoff after l steps since time n. Note that we have used the equation $K_n = (1-p)^{n-1}\tilde{K}_n$ in the proof of Theorem 3.6.1. If we substitute it into (B.1), we have

$$V_{n+l}(\lambda) = -\lambda T - \lambda\tau\sum_{i=1}^{n}(1-p)^{i-1}\tilde{K}_i + \left[R_{n+l}T - \lambda\tau\sum_{i=n+1}^{n+l}(1-p)^{i-1}\tilde{K}_i\right].$$

If we start from time $n+1$, the payoff after l rounds is

$$
\begin{aligned}
V_{n+l+1}(\lambda) = & -\lambda T - \lambda\tau \sum_{i=1}^{n} (1-p)^{i-1}\tilde{K}_i - \lambda\tau(1-p)^n \tilde{K}_{n+1} \\
& + \left[R_{n+l+1}T - \lambda\tau \sum_{i=n+2}^{n+l+1} (1-p)^{i-1}\tilde{K}_i \right].
\end{aligned}
$$

The item in the above square bracket is the recursive part for l rounds of observations since time $n+1$. We can rewrite it as

$$
(1-p)\left\{ R_{n+l+1}T - \lambda\tau \sum_{i=n+1}^{n+l} (1-p)^{i-1}\tilde{K}_{i+1} \right\} + p \cdot R_{n+l+1}T.
$$

By [A1], p should be reasonably small; otherwise the average number of probing links Mp are much larger than 1, leading to increased probing costs. Hence we can ignore the last term and write the optimality equation as

$$
V_n^*(\lambda) = E\left[\max\left\{ R_nT - \lambda T - \lambda\tau \sum_{i=1}^{n} K_i, (1-p)\left(V_n^*(\lambda) - \tau K_{n+1}\right) \right\} \right].
$$

Again, the optimal payoff λ_n^* that maximizes the rate of return must satisfy $V_n^*(\lambda_n^*) = 0$. We substitute it into the optimality equation and rewrite it as

$$
E\left[\frac{R_n}{\lambda_n^*} - \frac{\tau}{T}\left\{ \sum_{i=1}^{n} K_i - (1-p)K_{n+1} \right\} - 1 \right]^+ = (1-p) \cdot \frac{\tau}{T} E[K_{n+1}].
$$

If we further notice that $K_{n+1} = 1/g_{n+1}\tilde{K}_{n+1} = (1-p)^n \tilde{K}_{n+1}$ and $\tilde{K}_{n+1}$ and K_1 are *i.i.d.*, we can rewrite the above equation as (3.51). The optimal stopping rule N^* can be derived in the same way as in Theorem 3.7.2. To get the overall optimal system throughput λ_P^*, we let $n = 0$ in (3.51) and rewrite the equation as (3.49).

Bibliography

[1] Daniel Aguayo, John Bicket, Sanjit Biswas, Glenn Judd, and Robert Morris. Link-level measurements from an 802.11b mesh network. In *Proceedings of ACM SIGCOMM*, pages 121–132, 2004.

[2] Min Cao, Vivek Raghunathan, and P. Kumar. Cross-layer exploitation of MAC layer diversity in wireless networks. In *Proceedings of IEEE ICNP*, pages 332–341, 2006.

[3] R. Knopp and P.A. Humblet. Information capacity and power control in single-cell multiuser communications. In *Proc. IEEE ICC*, pages 331–335, June 1995.

[4] P. Viswanath, D.N.C. Tse, and R. Laroia. Opportunistic beamforming using dumb antennas. *IEEE Transactions on Information Theory*, 48(6):1277–1294, June 2002.

[5] David Tse and Pramod Viswanath. *Fundamentals of wireless communication*. Cambridge University Press, New York, NY, USA, 2005.

[6] P. Bender, P. Black, M. Grob, R. Padovani, N. Sindhushyana, and S. Viterbi. CDMA/HDR: a bandwidth efficient high speed wireless data service for nomadic users. *IEEE Communications Magazine*, 38(7):70–77, July 2000.

[7] S. Borst and P. Whiting. Dynamic rate control algorithms for HDR throughput optimization. In *Proc. IEEE INFOCOM*, pages 976–985, 2001.

[8] S. Shakkottai, R. Srikant, and A. Stolyar. Pathwise optimality and state space collapse for the exponential rule. In *Proceedings of IEEE International Symposium on Information Theory*, page 379, 2002.

[9] M. Andrews, K. Kumaran, K. Ramanan, A. Stolyar, P. Whiting, and R. Vijayakumar. Providing quality of service over a shared wireless link. *IEEE Communications Magazine*, 39(2):150–154, February 2001.

[10] R. Agrawal, A. Bedekar, R. J. La, R. Pazhyannur, and V. Subramanian. Class and channel condition based scheduler for EDGE/GPRS. In *Proc. SPIE*, pages 59–68, 2001.

[11] Xin Liu, E.K.P. Chong, and N.B. Shroff. Transmission scheduling for efficient wireless utilization. In *Proc. IEEE INFOCOM*, pages 776–785, 2001.

[12] Y. Liu and E. Knightly. Opportunistic fair scheduling over multiple wireless channels. In *Proc. IEEE INFOCOM*, pages 1106–1115, 2003.

[13] S. Borst. User-level performance of channel-aware scheduling algorithms in wireless data networks. In *Proc. IEEE INFOCOM*, pages 321–331, 2003.

[14] Sushant Jain, Kevin Fall, and Rabin Patra. Routing in a delay tolerant network. In *Proceedings of ACM SIGCOMM*, pages 145–158, 2004.

[15] Zhensheng Zhang. Routing in intermittently connected mobile ad hoc networks and delay tolerant networks: overview and challenges. *IEEE Communications Surveys Tutorials*, 8(1):24–37, 2006.

[16] L. Pelusi, A. Passarella, and M. Conti. Opportunistic networking: data forwarding in disconnected mobile ad hoc networks. *IEEE Communications Magazine*, 44(11):134–141, November 2006.

[17] L. Pelusi, A. Passarella, and M. Conti. Beyond MANETs: Dissertation on opportunistic networking. *IIT-CNR Tech. Rep.*, May 2006.

[18] Y. Mostofi, A. Gonzales-Ruiz, A. Ghaffarkhah, and D. Li. Characterization and modeling of wireless channels for networked robotic and control systems - a comprehensive overview. In *Proceedings of IEEE International Conference on Intelligent Robots and Systems*, pages 4849–4854, October 2009.

[19] Hua Chen, Pedram Hovareshti, and John S. Baras. Distributed collaborative controlled autonomous vehicle systems over wireless networks. In *Proceedings of 18th Mediterranean Conference on Control Automation*, pages 1695–1700, June 2010.

[20] Hua Chen, P. Hovareshti, and J.S. Baras. Opportunistic communications for networked controlled systems of autonomous vehicles. In *Proceedings of Military Communications Conference*, pages 1430–1435, 2010.

[21] Praveen Kumar Gopala and H.E. Gamal. Opportunistic multicasting. In *Proceedings of IEEE Asilomar Conference on Signals, Systems and Computers*, pages 845–849, November 2004.

[22] Tze-Ping Low, Man-On Pun, and C.-C.J. Kuo. Optimized opportunistic multicast scheduling over cellular networks. In *Proceedings of IEEE Global Telecommunications Conference*, pages 1–5, 2008.

[23] Tze-Ping Low, Man-On Pun, Y.-W.P. Hong, and C.-C. Jay Kuo. Optimized opportunistic multicast scheduling (oms) over heterogeneous cellular networks. In *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 2545–2548, April 2009.

[24] U.C. Kozat. On the throughput capacity of opportunistic multicasting with erasure codes. In *Proceedings of IEEE INFOCOM*, pages 520–528, April 2008.

[25] Qi Qu and U.C. Kozat. On the opportunistic multicasting in OFDM-based cellular networks. In *Proceedings of IEEE International Conference on Communications*, pages 3708–3714, May 2008.

[26] D. Gesbert and M.-S. Alouini. How much feedback is multi-user diversity really worth? In *Proceedings of IEEE ICC*, volume 1, pages 234–238, June 2004.

[27] Xiangping Qin and Randall Berry. Exploiting multiuser diversity for medium access control in wireless networks. In *Proc. IEEE INFOCOM*, pages 1084–1094, 2003.

[28] Xiangping Qin and R. Berry. Opportunistic splitting algorithms for wireless networks. In *Proc. IEEE INFOCOM*, pages 1662–1672, March 2004.

[29] S. Adireddy and L. Tong. Exploiting decentralized channel state information for random access. *IEEE Transactions on Information Theory*, 51(2):537–561, February 2005.

[30] Y. Yu and G.B. Giannakis. Opportunistic medium access for wireless networking adapted to decentralized CSI. *IEEE Transactions on Wireless Communications*, 5(6):1445–1455, June 2006.

[31] K. Bai and Junshan Zhang. Opportunistic multichannel Aloha: distributed multiaccess control scheme for OFDMA wireless networks. *IEEE Transactions on Vehicular Technology*, 55(3):848–855, May 2006.

[32] Jianfeng Wang, Hongqiang Zhai, Yuguang Fang, and M.C. Yuang. Opportunistic media access control and rate adaptation for wireless ad hoc networks. In *Proceedings of IEEE International Conference on Communications*, pages 154–158, June 2004.

[33] Miao Zhao, Huiling Zhu, Wenjian Shao, V.O.K. Li, and Yuanyuan Yang. Contention-based prioritized opportunistic medium access control in wireless lans. In *Proceedings of IEEE International Conference on Communications*, volume 8, pages 3820–3825, June 2006.

[34] Jianfeng Wang, Hongqiang Zhai, Yuguang Fang, J.M. Shea, and Dapeng Wu. OMAR: Utilizing multiuser diversity in wireless ad hoc networks. *IEEE Transactions on Mobile Computing*, 5(12):1764–1779, December 2006.

[35] Dong Zheng, Weiyan Ge, and Junshan Zhang. Distributed opportunistic scheduling for ad hoc networks with random access: An optimal stopping approach. *IEEE Transactions on Information Theory*, 55(1):205–222, January 2009.

[36] D.A. Schoenwald. AUVs: In space, air, water, and on the ground. *IEEE Control Systems Magazine*, 20(6):15–18, Dec. 2000.

[37] R. Bachmayer and N.E. Leonard. Vehicle networks for gradient descent in a sampled environment. In *Proceedings of the 41st IEEE Conference on Decision and Control*, pages 112–117, Dec. 2002.

[38] Reza Olfati-Saber and Richard M. Murray. Distributed cooperative control of multiple vehicle formations using structural potential functions. In *Proceedings of the IFAC World Congress*, 2002.

[39] J.P. Desai, J.P. Ostrowski, and V. Kumar. Modeling and control of formations of nonholonomic mobile robots. *IEEE Transactions on Robotics and Automation*, 17(6):905–908, Dec. 2001.

[40] J. S. Baras, Xiaobo Tan, and P. Hovareshti. Decentralized control of autonomous vehicles. In *Proceedings of the 42nd IEEE Conference on Decision and Control*, pages 1532–1537, Dec. 2003.

[41] Wei Xi, Xiaobo Tan, and John S. Baras. Gibbs sampler-based coordination of autonomous swarms. *Automatica*, 42(7):1107–1119, 2006.

[42] O. Khatib. Real-time obstacle avoidance for manipulators and mobile robots. *The International Journal of Robotics Research*, 5(1):90–98, 1986.

[43] E. Rimon and D.E. Koditschek. Exact robot navigation using artificial potential functions. *IEEE Transactions on Robotics and Automation*, 8(5):501–518, Oct 1992.

[44] N.E. Leonard and E. Fiorelli. Virtual leaders, artificial potentials and coordinated control of groups. In *Proceedings of the 40th IEEE Conference on Decision and Control*, pages 2968–2973, Dec 2001.

[45] H. G. Tanner, A. Jadbabaie, and G. J. Pappas. Flocking in fixed and switching networks. *IEEE Transactions on Automatic Control*, 52(5):863–868, May 2007.

[46] K. Bullington. Radio propagation for vehicular communications. *IEEE Transactions on Vehicular Technology*, 26(4):295–308, Nov. 1977.

[47] M. Jelasity, A. Montresor, and O. Babaoglu. Gossip-based aggregation in large dynamic networks. *ACM Trans. on Computer Systems*, 23(3):219–252, Aug. 2005.

[48] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah. Randomized gossip algorithms. *IEEE/ACM Transactions on Networking*, 14(6):2508–2530, Jun. 2006.

[49] D. Johnson, Y. Hu, and D. Maltz. RFC 4728: The dynamic source routing protocol (DSR) for mobile ad hoc networks for IPv4. http://tools.ietf.org/html/rfc4728, February 2007.

[50] Gavin Holland, Nitin Vaidya, and Paramvir Bahl. A rate-adaptive MAC protocol for multi-hop wireless networks. In *Proc. ACM MobiCom*, pages 236–251, 2001.

[51] B. Sadeghi, V. Kanodia, A. Sabharwal, and E. Knightly. Opportunistic media access for multirate ad hoc networks. In *Proc. ACM MobiCom*, pages 24–35, 2002.

[52] Zhengrong Ji, Yi Yang, Junlan Zhou, Mineo Takai, and Rajive Bagrodia. Exploiting medium access diversity in rate adaptive wireless LANs. In *Proc. ACM MobiCom*, pages 345–359, 2004.

[53] Hua Chen, Pedram Hovareshti, and John S. Baras. Distributed medium access and opportunistic scheduling for ad-hoc networks: an analysis of the constant access time problem. In *Proc. IEEE GLOBECOM*, pages 1–6, December 2011.

[54] Hua Chen and John S. Baras. Distributed opportunistic scheduling for wireless ad-hoc networks with block-fading model. Technical report, Institute for Systems Research, 2013.

[55] Thomas S. Ferguson. Optimal stopping and applications, 2006.

[56] A. Shiryayev. *Optimal Stopping Rules*. Springer-Verlag, 1978.

[57] Dimitri P. Bertsekas. *Dynamic Programming and Optimal Control*. Athena Scientific, 2nd edition, 2000.

[58] Hua Chen and John S. Baras. A distributed opportunistic scheduling protocol for multi-channel wireless ad-hoc networks. In *Proc. IEEE GLOBECOM*, pages 292–297, December 2012.

[59] A. Sabharwal, A. Khoshnevis, and E. Knightly. Opportunistic spectral usage: Bounds and a multi-band CSMA/CA protocol. *IEEE/ACM Transactions on Networking*, 15(3):533–545, June 2007.

[60] N. B. Chang and Mingyan Liu. Optimal channel probing and transmission scheduling for opportunistic spectrum access. *IEEE/ACM Transactions on Networking*, 17(6):1805–1818, December 2009.

[61] S. Guha, K. Munagala, and S. Sarkar. Jointly optimal transmission and probing strategies for multichannel wireless systems. In *Proc. Annual Conference on Information Sciences and Systems*, pages 955–960, March 2006.

[62] John A. Rice. *Mathematical Statistics and Data Analysis*. Duxbury Press, 1999.

[63] Athanasios Papoulis and S. Unnikrishna Pillai. *Probability, Random Variables and Stochastic Processes*. McGraw Hill, 4th edition, 2002.

[64] Dong Zheng and Junshan Zhang. Protocol design and throughput analysis of frequency-agile multi-channel medium access control. *IEEE Transactions on Wireless Communications*, 5(10):2887–2895, October 2006.

[65] V. Kanodia, A. Sabharwal, and E. Knightly. MOAR: a multi-channel opportunistic auto-rate media access protocol for ad hoc networks. In *Proceedings of IEEE International Conference on Broadband Communications, Networks, Systems (BroadNets)*, pages 600–610, October 2004.

[66] Nicholas B. Chang and Mingyan Liu. Optimal channel probing and transmission scheduling for opportunistic spectrum access. In *Proc. ACM MobiCom*, pages 27–38, 2007.

[67] S. Guha, K. Munagala, and S. Sarkar. Approximation schemes for information acquisition and exploitation in multichannel wireless networks. In *Proceedings Annu. Allerton Conf. Commun., Control, Comput.*, pages 85–90, September 2006.

[68] IEEE 802.11a Working Group. Wireless LAN Medium Acces Control (MAC) and Physical Layer (PHY) Specifications: High speed Physical Layer (PHY) in the 5GHz Band, September 1999.

[69] IEEE 802.11b Working Group. Wireless LAN Medium Acces Control (MAC) and Physical Layer (PHY) Specifications: High speed Physical Layer (PHY) Extension in the 2.4GHz Band, September 1999.

[70] J. Kennedy and M.C. Sullivan. Direction finding and 'smart antennas' using software radio architectures. *IEEE Communications Magazine*, 33(5):62–68, May 1995.

[71] Yunxia Chen, Qing Zhao, and A. Swami. Joint design and separation principle for opportunistic spectrum access. In *Proceedings of IEEE Asilomar Conference on Signals, Systems and Computers*, pages 696–700, 2006.

[72] Tristan Henderson, David Kotz, and Ilya Abyzov. The changing usage of a mature campus-wide wireless network. In *Proceedings of ACM MobiCom*, pages 187–201, 2004.

[73] T. Lai and H. Robbins. Asymptotically efficient adaptive allocation rules. *Advances in Applied Mathematics*, 6(1), 1985.

[74] V. Anantharam, P. Varaiya, and J. Walrand. Asymptotically efficient allocation rules for the multiarmed bandit problem with multiple plays-part I: I.I.D. rewards. *IEEE Transactions on Automatic Control*, 32(11):968–976, November 1987.

[75] R. Agrawal. Sample mean based index policies with o(log n) regret for the multi-armed bandit problem. *Advances in Applied Probability*, 27:1054–1078, 1995.

[76] Peter Auer, Nicolò Cesa-Bianchi, and Paul Fischer. Finite-time analysis of the multiarmed bandit problem. *Machine Learning*, 47:235–256, May 2002.

[77] Yi Gai, B. Krishnamachari, and R. Jain. Learning multiuser channel allocations in cognitive radio networks: A combinatorial multi-armed bandit formulation. In *IEEE Symposium on New Frontiers in Dynamic Spectrum (DySPAN)*, pages 1–9, April 2010.

[78] Keqin Liu and Qing Zhao. Decentralized multi-armed bandit with multiple distributed players. In *Information Theory and Applications Workshop (ITA)*, pages 1–10, 2010.

[79] A. Anandkumar, N. Michael, and Ao Tang. Opportunistic spectrum access with multiple users: Learning under competition. In *Proc. IEEE INFOCOM*, pages 1–9, March 2010.

[80] Y. S. Chow, H. Robbins, and D. Siegmund. *Great Expectations: Theory of Optimal Stopping*. Houghton Mifflin, Boston, MA, USA, 1971.