Dynamic games and applications in wireless communication networks

Zhu, Kun

2012

Zhu, K. (2012). Dynamic games and applications in wireless communication networks.
Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/50663>

<https://doi.org/10.32657/10356/50663>

Dynamic Games and Applications in Wireless Communication Networks

Kun Zhu

**School of Computer Engineering
Nanyang Technological University**

A thesis submitted to Nanyang Technological University
in partial fulfillment of the requirement for the degree of
Doctor of Philosophy

2012

Acknowledgments

First and foremost, I would like to express the deepest gratitude to my supervisors, Prof. Dusit Niyato and Prof. Ping Wang, for their invaluable guidance, discussions, and suggestions throughout the doctoral program. They share their knowledge, experiences, and working style which would benefit me throughout my life. Prof. Dusit Niyato points out the novel research direction which leads me to my research road and gives me inspirations from detailed discussions. Prof. Ping Wang is rigorous and careful. She can always find the issues I ignored. I am also grateful for their help in improving my technical writing. I have greatly benefited from their supervision.

Also, I express my sincere thanks to Dr. Amitabha Das, my ex-advisor, for his valuable guidance and encouragement which laid the foundation for my research.

I would like to thank our research group members for creating a cheerful environment to carry out my research. I would also like to thank my lab colleagues and technicians, for their self-giving help and warmhearted assistance for research facilities and experiments. Special thanks to my friends for their support.

Finally, I wish to express my deepest gratitude to my parents and my wife Xu Yuanyuan, for their everlasting love, encouragement, and support in my life.

Contents

Acknowledgments	ii
Table of Contents	iii
List of Figures	vii
List of Tables	ix
Abstract	x
1 Introduction	1
1.1 Introduction to Game Theory and Its Applications in Wireless Communication Networks	1
1.2 Research Objective and Motivation	5
1.2.1 Research Objective	5
1.2.2 Motivation	5
1.3 Research Focus	7
1.3.1 Dynamic Network Selection in Heterogeneous Wireless Networks with Incomplete Information	7
1.3.2 Dynamic Spectrum Leasing and Service Selection in Spectrum Secondary Market of Cognitive Radio Networks	10
1.3.3 Dynamic Service Selection and Bandwidth Allocation in IEEE 802.16m Mobile Relay Networks	11
1.3.4 Connections among Research Issues	12
1.4 Organization of the Chapters	13

2	Background and Related Works	15
2.1	Background and Related Works on Network Selection in Heterogeneous Wireless Networks	15
2.2	Background and Related Works on Spectrum Leasing in Cognitive Radio Networks	17
2.3	Background and Related Works on IEEE 802.16j/m Mobile Relay Networks	20
2.4	Background and Related Works on Applications of Evolutionary Game and Differential Game	22
2.4.1	Introduction to Evolutionary Game	22
2.4.2	Introduction to Differential Game	23
2.4.3	Related works	24
3	Network Selection in Heterogeneous Wireless Networks: Evolution with Incomplete Information	25
3.1	Introduction	25
3.2	System Model and Assumptions	26
3.3	Bayesian Evolutionary Game Formulation of Decentralized Network Selection . .	27
3.3.1	Bayesian Network Selection Game	28
3.3.2	Payoff Function	28
3.3.3	Nash Equilibrium of Underlying Bayesian Network Selection Game	30
3.4	Dynamics of Network Selection Game	31
3.4.1	Bayesian Evolutionary Dynamics for Network Selection Game	32
3.4.2	Aggregate Dynamics for Network Selection Game	33
3.5	Numerical Analysis	34
3.5.1	Parameter Setting	34
3.5.2	Numerical Results	35
3.6	Conclusion	37
4	Dynamic Spectrum Leasing and Service Selection in Spectrum Secondary Market of Cognitive Radio Networks	40

4.1	Introduction	40
4.2	A Novel Hierarchical Dynamic Game Framework	42
4.2.1	Motivation	42
4.2.2	Links Between the Two Levels	43
4.2.3	Differences and Advantages	44
4.3	System Model	46
4.4	Modeling the Evolution of Service Selection	48
4.4.1	Formulation of the Evolutionary Game	48
4.4.2	Existence and Uniqueness of Solution under Control	50
4.4.3	Equilibrium and Stability Analysis	51
4.5	Dynamic Spectrum Leasing	51
4.5.1	Non-cooperative Dynamic Spectrum Leasing as a Differential Game	52
4.5.2	Open-loop and Closed-loop Nash Equilibrium	53
4.5.3	Calculation of Nash Equilibrium	54
4.5.4	An Illustrative Example	58
4.6	Dynamic Spectrum Leasing with Hierarchical Play	61
4.7	Performance Evaluation	63
4.7.1	Parameter Setting	63
4.7.2	Numerical Results	63
4.8	Conclusion	71
5	Dynamic Service Selection and Bandwidth Allocation in IEEE 802.16m Mobile Relay Networks	72
5.1	Introduction	72
5.2	System Model	73
5.2.1	IEEE 802.16m Relay Network	73
5.2.2	Queue Dynamics	75

5.3	Service Selection Evolutionary Game of Advanced Mobile Stations	76
5.3.1	Service Selection Dynamics	76
5.3.2	Existence and Uniqueness of Population State Evolution Trajectory under Bandwidth Allocation Control	78
5.3.3	Equilibrium and Evolutionary Stability Analysis	78
5.4	Dynamic Bandwidth Allocation in IEEE 802.16m Mobile Relay Networks	79
5.4.1	Dynamic Bandwidth Allocation as a Differential Game	80
5.4.2	Stochastic Optimal Control Approach	81
5.4.3	H_∞ Robust Optimal Control Approach	83
5.5	Performance Evaluation	86
5.5.1	Parameters Setting	86
5.5.2	Numerical Results	87
5.6	Conclusion	89
6	Conclusions and Future Work	92
6.1	Major Research Contributions	92
6.2	Future Work	93
6.2.1	Future Work on Network Selection	93
6.2.2	Future Work on Spectrum Leasing	94
6.2.3	Future Work on Bandwidth Allocation	94
	Bibliography	95
A	List of Abbreviations	107
B	List of Symbols	109
C	Publication List	113

List of Figures

1.1	A model of heterogeneous wireless environment.	8
1.2	Mobility management issues in heterogeneous wireless networks.	9
3.1	Network model of heterogeneous wireless network with network selection.	27
3.2	Phase portrait of the aggregate best response dynamics.	36
3.3	Trajectory of network selection strategy adaption.	37
3.4	Allocated bandwidth to each user at equilibrium under different number of users.	38
3.5	Impact of the handover cost on the dynamics.	38
3.6	Impact of the price of WLAN on the equilibrium distribution.	39
3.7	Impact of the learning rate on the dynamics.	39
4.1	Two-level dynamic game framework.	41
4.2	Different types of game models.	45
4.3	System model of a spectrum secondary market in a cognitive radio network.	47
4.4	Population distribution evolution of service provider 2 under different initial population shares.	64
4.5	Optimal control strategies of secondary service providers for finite horizon.	65
4.6	Optimal control strategies of secondary service provider 1 for infinite horizon.	65
4.7	Optimal control strategies of secondary service provider 2 for infinite horizon.	66
4.8	Best response functions and Nash equilibrium of the static non-cooperative game.	66
4.9	Population state evolution under different control strategies.	67

4.10	The impact of price to the cumulative profit of provider.	68
4.11	Comparison of cumulative profit of secondary service providers.	69
4.12	Comparison of instantaneous profit.	70
4.13	Comparison of utility and delay of secondary users under different control strategies.	70
5.1	System model of a competitive IEEE 802.16m mobile relay network with M advanced base stations and N advanced mobile stations.	74
5.2	Population state of the service selection evolutionary game under stochastic optimal control.	87
5.3	Service rate of different queues under stochastic optimal control.	88
5.4	Empirical mean and 95% confidence interval (CI) under stochastic optimal control.	88
5.5	Queue states of q_{11} under Gaussian noise.	89
5.6	Queue states of the q_{11} under stochastic disturbance with different mean and variance.	90
5.7	Queue states of q_{11} under non-Gaussian noise.	90

List of Tables

4.1	Control strategy implementation schemes for spectrum leasing.	58
5.1	Control strategy implementation schemes for bandwidth allocation.	86

Abstract

With the advances in telecommunication technologies and dramatic performance enhancement of communication equipments, the communication and computing are converging and autonomous distributed architectures will play more important roles in future wireless communication networks. Therefore, devising distributed and dynamic algorithms for ensuring a robust network operation in time-varying and heterogeneous environments becomes a critical issue. Game theory as a discipline studying the interactions of interdependent autonomous agents provides an ideal framework with a set of mathematical tools for this purpose. In this dissertation, we focus on the use of dynamic games to model, analyze, and design efficient distributed algorithms for the competitive resource management in wireless networks. The motivation for the use of dynamic games is from the consideration of dynamic nature of wireless environment and the wide existence of hierarchical structures in wireless networks modeling. The specific issues addressed in this dissertation are summarized as follows.

The first issue is the dynamic network selection in heterogeneous wireless networks with incomplete information. A network selection Bayesian game is formulated for this purpose. In general, the preference (i.e., utility) of a mobile user is private information. Therefore, each user has to make the decision of network selection optimally given only the partial information of the preferences of other users. To study the dynamics of such network selection, the Bayesian best response dynamics and aggregate best response dynamics are applied. Bayesian Nash equilibrium is considered to be the solution of this game, and there is a one-to-one mapping between the Bayesian Nash equilibrium and the equilibrium distribution of the aggregate dynamics. We show that even with incomplete information, the equilibrium of network selection decisions of mobile users can be reached.

The second issue is dynamic spectrum leasing and service selection in spectrum secondary market of cognitive radio networks. In spectrum secondary market, the secondary service providers lease spectrum from spectrum brokers to provide service to secondary users. The problem is challenging when the optimal decisions of both secondary service providers and secondary users are made

dynamically under competition. To address this problem, a two-level dynamic game framework is developed. Since the secondary users can adapt the service selection strategies according to the received service quality and price, the dynamic service selection is modeled as an evolutionary game at the lower level. The replicator dynamics is applied to model the service selection adaptation and evolutionary equilibrium is considered to be the solution. With dynamic service selection, competitive secondary providers can dynamically lease spectrum to provide services to secondary users. A spectrum leasing differential game is formulated to model this competition at the upper level. Both simultaneous play model and asynchronous play model are considered. Both open-loop and closed-loop Nash equilibria are obtained as the solution of dynamic control of the differential game. We show the advantages of dynamic control over static control in terms of profit and convergence speed.

The third issue is dynamic service selection and bandwidth allocation in IEEE 802.16m relay networks. Cooperative relay will be supported in IEEE 802.16m to improve the coverage and performance of mobile broadband wireless access services. The advanced mobile stations (AMSS) perform the selection of advanced base station (ABS) and transmission mode (i.e., direct transmission or relay-cooperation transmission) for a better service quality. The ABSs allocate the bandwidth for different transmission modes to maintain the desired queue level at base stations and user distribution for satisfying performance requirements. Due to the cyclic dependency, the problem of dynamic service selection and bandwidth allocation in IEEE 802.16m needs to be jointly considered. To address this problem, we extend the proposed two-level dynamic game framework by considering the fluctuation of traffic flow rate as disturbance. The robust bandwidth allocation strategy design is performed. Both stochastic optimal control and H_∞ optimal control approaches are adopted for average performance and worst-case performance design, respectively. We show the effectiveness of the proposed dynamic control scheme.

In summary, this dissertation investigates and demonstrates different applications of dynamic game theory in wireless communications and networking. Due to dynamic nature of wireless systems, evolutionary and differential games provide suitable mathematical tools to model and analyze the competitive dynamic decision making problems. Compared with static game formulations for such problems, the dynamic games provide better performance and higher efficiency.

Chapter 1

Introduction

In this chapter, we first introduce game theory and its applications in wireless communication networks. Then the research objective and motivation are discussed. A brief introductions to our specific research focus are also provided.

1.1 Introduction to Game Theory and Its Applications in Wireless Communication Networks

Game theory is a discipline aimed at providing a mathematical basis for modeling and analyzing interactive decision-making problems in which the interests of agents (i.e., players) may conflict with each other [1,2]. Different from optimization techniques in which the decision making process is not affected by mutual decisions, game theory is a collection of modeling tools that study the interactive decision problems and predict the possible outcomes of such problems. Game theory has been used primarily in economics to model the competition in market, and in recent years, the application of game theory has been extended to wide areas, e.g., politics, biology, and wireless communications [3,4].

A non-technical definition of a game is given as follow. A game is a process in which the agents choose certain strategies from their own strategy sets and obtain payoffs according to the strategies of all agents [1]. The choice of a strategy can be made both simultaneously and non-simultaneously. In addition, an agent may make decisions multiple times according to the game rule. A game

consists of a set of players, a set of available strategies to those players, and a specification of payoffs for each combination of strategies.

Different models in game theory may differ considerably in structure from many aspects, e.g., the number of players, the number of strategies, and payoffs. The number of players may vary in different games. If a game has only one player, the game becomes an optimization problem. We call a two person game or multiple person game if the game has two or more players, respectively. In different games, the number of strategies for players can be either finite (e.g., in a rock-scissor-paper game) or infinite (e.g., in a pricing game). The analyses of a finite strategy game and an infinite strategy game are different. The summation of payoffs of all players may also differ in different models. In general, this summation can be zero, a non-zero constant number, or any arbitrary value. The game process is an important aspect in the game structure. The players in a game may take actions simultaneously, in a certain order, or in a repeated fashion, according to which we call static game, dynamic game, and repeated game, respectively. Information is important for the decision making process. The information in game theory can be broadly categorized into information on payoffs and information on game process. For a game, the information on payoffs is said to be complete if the payoffs of all players are common knowledge. Otherwise the information is incomplete. Similarly, the information on game process is said to be perfect if the process of the game is clear to all players. Otherwise the information is imperfect. In addition, the assumptions of players' rationality are different. Most game theory models assume perfect rationality of players, while some models consider the more realistic situations in which the players are with limited rationality (i.e., bounded rationality).

According to the analysis of the structure of a game, game models can be classified based on different criteria. With the first criterion, game theory can be categorized into non-cooperative game theory and cooperative game theory. In the following, we only consider the classification within non-cooperative game framework. In the second criterion, non-cooperative game theory can be categorized into perfect rationality game and limited rationality game according to different rationality assumptions. Based on the process of the game, non-cooperative game can be categorized into static game and dynamic game with the third criterion. With the fourth criterion, according to whether the information on payoffs is complete and whether the information on the game process is perfect, a non-cooperative game can be categorized into complete information static game, incomplete information static game, complete and perfect information dynamic game, complete and imperfect information dynamic game, and incomplete information dynamic game.

As we have known, game theory studies the interactions of autonomous agents. In future wireless networks, more control schemes will be deployed in a distributed fashion. Also, the dramatic performance enhancement for communication equipments makes it possible for various network entities to make independent and selfish decisions. In this case, an entity running a distributed protocol makes its own decisions and the outcome depends on all decisions from other related entities. Naturally, game theory is a suitable tool for the study of such interactions in wireless networks. A number of works have explored the applications of game theory for the analysis and optimization of various issues in wireless systems, in most cases to solve resource allocation problems in a competitive environment. For example, the issues include the uplink power control in CDMA networks [5–7], the network selection in multi-access wireless networks [8–11], the MAC protocol design [12–15], random access control [16, 17], admission control [18–20] in wireless local area networks, the routing in multi-hop networks [21–24], and the cooperative spectrum sensing [25–27], spectrum leasing [28–32], and pricing [33–36] in cognitive radio networks. We provide brief description of some selected examples as follows:

1. *Pricing*: In heterogeneous wireless networks, different service providers coexist and compete for certain number of users. For users, the price always plays an important role in subscribing to which service provider. Similar to the pricing in economics, the pricing problem in heterogeneous wireless networks can also be analyzed using a game-theoretic framework. Much effort has been put in this field and different game models are formulated according to different competition relationships. We consider a game formulated between M service providers and N users as an example. Every service provider provides different service classes and offers prices for these classes accordingly to users. These prices are the strategies of service providers. For users, they have different quality of service (QoS) requirements and the strategies are the choices of service provider and service class. In this case, a Stackelberg game [1] in which the leader players (i.e., the service providers) move first and then the follower players (i.e., users) move subsequently can be formulated and the Stackelberg equilibrium is considered to be the solution of this game. When reaching the equilibrium, the profits of service providers are maximized given the price strategies of others and the QoS requirements of users can be satisfied [37–39].
2. *Medium access game*: Game theory can also be used in the design of distributed medium access protocols. We consider the simplest model where two wireless nodes compete for access to a common wireless channel. We assume that each node has one packet to send in each time slot and at the beginning of each time slot both nodes make decisions on whether

to transmit the packet or to wait. Only one node is allowed to transmit in each time slot and mutual interference which results in packet collision or transmission degradation occurs when the two nodes choose to transmit simultaneously. Each transmission incurs a certain cost and each successful transmission yields a certain gain. Naturally, this medium access problem can be modeled as a game in which the two nodes are players and the actions of transmit and wait are strategies of the two players [40–42].

3. *Power control:* Power control is important for users in a mobile cellular system to adapt to time varying radio channels not only for maintaining QoS but also for enhancing the battery life. In addition, the power control reduces the network interference and therefore maximizes the spatial reuse of channels. For power control, a distributed scheme is of special interest due to reduced complexity in infrastructure deployment. Non-cooperative game has been a natural tool for the analysis and design of distributed power control algorithms. Much work has been done in this area [43–45]. Basically, for a simple power control game considering only a single cell, the players are the active nodes in the cell and the strategies are the choices of power level. The payoff of a player is a function of his own transmit power and his signal to interference plus noise ratio (SINR), while the SINR is determined by the transmission power of all players. Nash equilibria are considered to be the solution of this game.

There are several techniques (e.g., optimization techniques and game theory) suitable for solving resource management problems in wireless networks. The choice of mathematical tools depends on the nature of the investigated problems as well as the system model. In general, the optimization techniques are used to study the optimal decision making of a single agent or multiple agents which have a common objective, while game theory techniques are used to study the optimal decision making of multiple agents the interests of which may conflict with each other (e.g., competing for the same resource) or the decision of an agent influences the payoff of other agents. The key difference is that game theory studies the multi-agent decision making problem with interaction while for optimization techniques the decision making process is not affected by mutual decisions. Therefore, game theory is a suitable mathematical modeling and analysis tool for interactive decision making problems especially in a competitive environment. Note that optimization techniques serve as the foundations for game theory. Specifically, for each player (i.e., agent), the calculation of best response is basically an optimization problem given the decisions of other players. Also, if there is only one player in the game, the problem becomes a pure optimization problem.

In wireless networks, the resource competition issues widely exist (e.g., the pricing competition

among different service providers and the medium access competition) for which game theory provides the most appropriate analysis tool. To summarize, optimization is mainly used for centralized system which wishes to optimize the whole system performance, while game theory can be used for the modeling and analysis of distributed systems where every autonomous distributed agent wishes to optimize its own performance.

1.2 Research Objective and Motivation

1.2.1 Research Objective

In this research, we aim to use non-cooperative dynamic games to study the competitive resource management issues in wireless communication networks considering the dynamic nature of the networks. In particular, this aim can be divided into three objectives. First, the dynamics of the competitive resource management issues need to be investigated. Second, a suitable dynamic game framework should be chosen or created to mathematically model the problem. Third, suitable algorithms should be developed to reach the equilibrium solutions. Specifically, in this dissertation, we study the issues of dynamic network selection in heterogeneous wireless networks, dynamic spectrum leasing in spectrum secondary market of cognitive radio networks, and dynamic bandwidth allocation in IEEE 802.16m mobile relay networks. The design and analysis of the dynamic game models for these problems should take the differences of various network entities (e.g., role and capability) into account. For example, in cellular networks, the base station (BS) acts as a service provider with the strong computation capability while the mobile station (MS) acts as a service subscriber with the limited computation capability. Therefore, different assumptions on the rationality of different agents (e.g., BS and MS) should be considered. Also, in the case that the assumption of complete information does not hold, the incomplete information model should be used. Besides, the learning behavior (i.e., strategy adaption) of agents should be taken into account.

1.2.2 Motivation

To illustrate the motivation of applying dynamic games to study various issues in wireless communication networks, we first give a brief introduction to the properties and differences of static game and dynamic game. A static game is one in which a single decision (time irrelevant but may contain

multiple actions) is made by each player, and each player has no knowledge of the decision made by the other players before making their own decisions. Decisions are made simultaneously (or their order is irrelevant). A game is dynamic if the order in which the decisions are made is important or the strategy itself is time dependent. For dynamic games, the dynamics can be abstracted from different aspects which lead to different types of dynamic games listed as follows:

- (1) *Dynamic nature in game's play order*: The dynamic nature in game's play (decision) order leads to the development of multi-stage game (e.g., Stackelberg game). In this case, the decisions are made asynchronously and the game's play order is important. The players move later can observe the decisions of the players who move first and then make the decisions accordingly. Note that if multiple players exist in one stage, the competition within this stage is usually formulated as a stage game.
- (2) *Dynamic nature in time dependency*: The dynamic nature in the time dependency leads to the development of differential game and evolutionary game. For differential game, the strategy of a player is time dependent (i.e., function of time t). That is, the player seeks a best response strategy considering the entire time horizon. For evolutionary game, the players adapt their strategies according to the time varying system state.

In wireless communication networks, especially for infrastructure-based networks, the competitive decision making problems with hierarchical structure exist extensively. For example, for the problem of service selection in a multi-access heterogeneous network, the service providers could be modeled in an upper level at which the service providers make decision (e.g., price) first. Then, the users could be modeled in a lower level at which the users make decision according to the decision of the providers. The decisions of providers are made considering the possible behavior of the users. Such cyclic dependency of pricing and service selection naturally requires the formulation of a dynamic (i.e., multi-stage) game. Also, the dynamic (i.e., time dependent) decision making is needed for many issues (e.g., the dynamic spectrum leasing in spectrum secondary market where the leasing is operated in a fine time scale). The importance for considering the time dependency strategy can be represented by the performance gain in terms of higher payoff and faster convergence speed than those of static strategy. Due to the lack of complete information and limited computation capability, many network entities (e.g., MSs) gradually learn and adapt their strategies to achieve better payoff which gives the motivation for the use of evolutionary game.

1.3 Research Focus

In this research, we focus on the application of dynamic games to analyze different issues in wireless communication networks. Specifically, the following research problems are undertaken.

1.3.1 Dynamic Network Selection in Heterogeneous Wireless Networks with Incomplete Information

One of the most significant features of the next generation wireless network is heterogeneity in which multiple wireless technologies are integrated to provide seamless mobility and achieve high-speed communications [46,47]. This concept of heterogeneous wireless network can be seen as the co-existing of radio access technologies (RATs) and network types. These different RATs complement each other in terms of service rate, QoS, coverage and cost so as to satisfy different service requirements. An example of such heterogeneous wireless environment is shown in Fig. 1.1. Wireless LANs (WLANs) [48] are suitable for high data rate with local area access, while for wide area communications, the third generation (3G) cellular networks [49] or mobile broadband wireless access (MBWA) based on IEEE 802.16 [50] may be used.

Similar to the always-connected concept in the second generation (2G) network, the concept of always-best-connected (ABC) [51] has been introduced for the heterogeneous wireless networks. For the purpose of ABC and reducing the deployment cost, integration of different networks is needed [52]. This integration introduces several challenging issues and one of which is the design of efficient network selection schemes [53]. Network selection aims to select the best set of access networks for a mobile user to achieve seamless mobility, to meet QoS requirements, and to support load balancing.

Network selection is also an important component of mobility management in heterogeneous wireless networks. As shown in Fig. 1.2, mobility management is composed of location management and handoff management [54]. Handoff management can be divided into the problems of “when to take handoff” and “how to take handoff”. “When to take handoff” is the decision made by the network selection algorithms. On the other hand, “how to take handoff” is the problem of designing efficient handoff scheme to minimize the handoff delay and handoff cost.

An efficient network selection algorithm should always select the set of best access networks that

Figure 1.1: A model of heterogeneous wireless environment.

most suit for current needs. Many factors need to be considered for the decision of network selection, e.g., user's preferences, network access cost, coverage area, system capacity, and bandwidth requirement. Naturally, three questions arise when the network selection algorithm is being designed, i.e., how to define the best, how to determine the best, and what information is needed [55].

The definition of the best network relates to the objective of the decision-making entity. A general approach for a decision making process is to formulate it as a mathematical optimization problem using utility function [56]. For the context of network selection in heterogeneous wireless networks, utility functions are widely applied. The satisfaction of using a wireless network can be quantified by a utility value. This utility is a function of the performance metrics and the cost. The service demand of a mobile user (i.e., QoS requirement) determines the type of the utility functions.

The information needed for network selection can be broadly categorized into policy-based and performance-metrics-based. The policy-based information includes user preferences, operator preferences, authentication mechanism, security mechanism, and cost. The performance-metrics-based information includes the access technology, signal strength, available bandwidth, utilization, packet delay, packet jitter, packet loss, and handover cost [57].

For choosing the best network, optimization techniques [58] and game theory [1] can be used. Optimization techniques are typically applied to the provider-driven network selection. The provider-

Figure 1.2: Mobility management issues in heterogeneous wireless networks.

driven network selection algorithm for integrated networks are controlled by the service provider with a single objective (e.g., maximize the resource utilization or maximize the profit). Therefore, optimization problem can be formulated given the objective and constraints of the service provider and network access. In contrast, game theory is typically applied to the user-driver network selection. The user-driver network selection algorithm is performed by each user independently and mostly in the distributed fashion. The users have their own objectives which may conflict with each other. Therefore, a game model can be formulated given the self-interest behavior of the user and the solution in terms of equilibrium can be obtained.

Due to the information dissemination constraints in current wireless networks, we focus our research on the network selection problem under incomplete information environment considering the dynamics of decision making process by the mobile users. Further details are provided in Chapter 3.

1.3.2 Dynamic Spectrum Leasing and Service Selection in Spectrum Secondary Market of Cognitive Radio Networks

Cognitive radio networks are envisaged as the solution for effective spectrum management for the future communications. The spectrum utilization can be improved via dynamic spectrum access (DSA) techniques [59] which enable dynamic spectrum sharing and access. Two main frameworks exist for spectrum sharing, i.e., cooperative DSA and non-cooperative DSA [60]. With cooperative DSA, some spectrum holders may lease portion of spectrum to secondary users in exchange for appropriate remuneration. In this case, two types of spectrum market can be established, i.e., primary market and secondary market [61]. The concept of spectrum primary market is well known and extensively studied in which the spectrum leasing is performed directly among primary users (i.e., the spectrum holders) and secondary users. The concept of spectrum secondary market is proposed in [29,61,62] for secondary users to dynamically trade the channel resources obtained in the primary market from spectrum holders among themselves.

In the spectrum primary market, the spectrum leasing is usually operated in a long-term basis (e.g., monthly auctions by legacy owners), while in the spectrum secondary market, the spectrum leasing is usually operated in a short-term basis (e.g., several minutes) among secondary users themselves. One important application scenario of the secondary market is the existence of secondary service providers (also termed as Mobile Virtual Network Operators (MVNO) [63]) which lease spectrum

resources from spectrum brokers to provide service to secondary users. Note that there are 645 active MVNOs worldwide as of May 2011 (around 400 as of February 2009 [61]). Though extensive studies have been performed for the primary market (e.g., leasing and pricing), the competitive spectrum leasing in a secondary market is largely overlooked. The importance for establishing spectrum secondary market can be summarized as follows:

- (1) *Practical consideration:* In a primary market, the spectrum is usually leased by companies (e.g., MVNO) or organizations (e.g., public safety units) from the spectrum holders (e.g., through monthly auctions) with long-term contracts. With an exclusive-use DSA model, it is not common for an individual user to lease spectrum directly from spectrum holders which usually prefer long-term lease.
- (2) *Advantages of spectrum secondary market:* The flexibility, spectrum utilization and user performance can be improved by establishing a spectrum secondary market.

In this research, we consider the dynamic spectrum leasing in the spectrum secondary market with competitive secondary service providers which compete to provide service to secondary users by leasing resource from spectrum brokers. The secondary users can select and churn to different secondary service providers according to the price and perceived service quality. In this case, the spectrum is traded dynamically depending on the demand of secondary providers which in turn depends on the number of subscribed secondary users. Note that, in this model, the primary market is not directly used for secondary users leasing spectrum from holders but is assumed to be used for spectrum brokers to lease spectrum from primary users (spectrum holders) in a long-term basis.

To address the issues of dynamic service selection and dynamic spectrum leasing, a hierarchical game framework with inherent dynamic nature is developed to jointly obtain the optimal strategies for both secondary users and providers given the strategies of others. Further details are provided in Chapter 4.

1.3.3 Dynamic Service Selection and Bandwidth Allocation in IEEE 802.16m Mobile Relay Networks

IEEE 802.16m is proposed as an enhancement for the existing mobile WiMAX systems to meet the requirements of next-generation mobile broadband communication in the framework of IMT-Advanced [64]. Relay transmission as a cost-effective approach for extending the coverage area

and enhancing the overall throughput is specified in the IEEE 802.16m standard [65], [66]. Relay-cooperation transmission can be seen as a kind of cooperative communications, in which a relay station helps to forward data between mobile stations and base stations. In IEEE 802.16m, the advanced base station (ABS) frame is divided into access zone and relay zone for supporting both advanced base station - advanced relay station (ABS-ARS) communications and advanced base station - advanced mobile station (ABS-AMS) communications, and the bandwidth of ABS in IEEE 802.16m for both communications has to be allocated accordingly. Due to the asymmetry of downlink (DL) and uplink (UL) traffic load, the duration of the access and relay zones could be different.

The bandwidth allocation for ABS-ARS and ABS-AMS communications is challenging due to the fact that the bandwidth allocation strategies of ABSs and the service selection strategies of AMSs influence each other and the decisions are made dynamically under competition. In particular, an AMS can choose ABSs and service modes (i.e., direct transmission or relay-cooperation transmission) dynamically while the ABSs can competitively allocate the bandwidth in a dynamic manner such that their individual target system performance can be achieved. The former (defined as the service selection problem) should be performed by taking the time-varying user performance into account. Meanwhile the latter (defined as the bandwidth allocation problem) should be done by considering the dynamic behavior of AMSs to maintain the desired queue level at the base stations and user distribution for satisfying system performance requirements. The service selection and bandwidth allocation are intertwined in such a way that the bandwidth will determine the preference of the users in selecting the services. Also, the user service selection affects the performance of the network.

To jointly address the issues of dynamic service selection and dynamic bandwidth allocation in IEEE 802.16m mobile relay networks, we extend the proposed hierarchical (i.e., two-level) dynamic game framework to obtain the equilibrium strategies for both AMSs and ABSs. Further details are provided in Chapter 5.

1.3.4 Connections among Research Issues

In this dissertation, we aim to use dynamic games to study competitive resource management in wireless communication networks. Though the system models differ in each problem due to the unique features and requirements for the specific topic, some player (e.g., user) behaviors in these models show certain similarities and therefore gives the motivation for the use of similar mathemati-

cal modeling tools. In this case, the connections of each topic exhibit in the underlying mathematical model. The connections among the three aforementioned research issues in terms of dynamic game models used are outlined as follows. For the first research issue (i.e., Chapter 3), evolutionary game is used to analyze the dynamic service selection behavior of users. For the second research issue (i.e., Chapter 4), based on the analysis of dynamic service selection behavior of users using evolutionary game, we analyze the competitive behavior of service providers using differential game. In this case, a two-level dynamic game framework based on evolutionary game and differential game is proposed. For the third research issue (i.e., Chapter 5), we further extend the proposed two-level dynamic game framework by considering the stochastic disturbance to system. Generally, the outline follows the evolutionary use (e.g., from simple to complicated and from single level to multiple level) of dynamic game theory.

1.4 Organization of the Chapters

The rest of this dissertation is organized as follows:

Chapter 2 The background and related works on the specific research issues investigated in this dissertation are reviewed. Also, the evolutionary game and differential game as well as their applications in wireless networks are introduced.

Chapter 3 We model and analyze the dynamics of network selection under incomplete information environment where the users make the decision of network selection optimally given only the partial information of the preferences of other users. In particular, a Bayesian network selection game is formulated and the dynamic behavior of which is studied based on Bayesian best response dynamics. Bayesian Nash equilibrium is considered to be the solution of this game.

Chapter 4 The problem of dynamic spectrum leasing and service selection in a spectrum secondary market of cognitive radio networks is investigated where the secondary service providers lease spectrum from spectrum brokers to provide service to secondary users. To address this problem, a novel hierarchical dynamic game framework is proposed. A lower level evolutionary game is formulated to model the service selection competition among secondary users and an upper level differential game is formulated to model the competition among secondary service providers.

Chapter 5 The problem of dynamic service selection and bandwidth allocation in IEEE 802.16m mobile relay networks is jointly considered. Specifically, the advanced mobile stations (AMSs) perform the selection of advanced base station (ABS) and transmission mode (i.e., direct transmission or relay-cooperation transmission) for a better service quality. The ABSs allocate the bandwidth for different transmission modes to maintain the desired queue level at base stations and user distribution for satisfying performance requirements. This problem is addressed using the extension of the proposed dynamic game framework.

Chapter 6 We give the concluding remarks of this dissertation, summarize the major research contributions, and outline possible further work.

Chapter 2

Background and Related Works

In this chapter, the background and related works on the specific research issues investigated in this dissertation are reviewed. Also, the evolutionary game and differential game as well as their applications in wireless networks are introduced.

2.1 Background and Related Works on Network Selection in Heterogeneous Wireless Networks

Today's wireless networks are highly heterogeneous in terms of the coexistence of various radio access technologies (RATs) and network architectures. Many factors contribute to such heterogeneity. First, comparing with the fast development of new wireless technologies, the wide deployment of new technologies usually adopt an evolutionary roadmap mainly due to the consideration of cost and compatibility issues. In this case, the new wireless communication systems will co-exist with legacy systems in a long period, e.g., the coexistence of secondary generation (2G) and the third generation (3G) mobile telecommunication technologies. Second, the enormous existing and emerging wireless applications (e.g., radio broadcasting, TV broadcasting, satellite communication, text message, voice call, video call, video streaming, online game, and wireless payment) with considerable different characteristics have different technical requirements. No single wireless technology could satisfy all these requirements. Therefore, various wireless technologies are required which complement each other in terms of data rate provided, QoS support, coverage area, mobility and roaming support, and cost [67]. For example, the 2G, 3G, and 4G systems provide wide coverage area, full

mobility and roaming support which are crucial for providing services to mobile users, whereas the wireless local area network (WLAN) technology targets at providing high data rate services to users in public and private areas with limited coverage at low cost. Third, competitive technologies for the same/similar target wireless market are developed by different organizations/companies/countries due to technical, economical, security, and political reasons. For example, four 3G standards (i.e., W-CDMA, CDMA2000, TD-SCDMA, and WiMAX) coexist and compete with each other.

Enabling users to connect to the best available network, dynamic network selection scheme is important for satisfying various QoS requirements, achieving seamless mobility and load balancing in heterogeneous wireless networks. Network selection in heterogeneous wireless networks can be categorized into two approaches, i.e., network-driven and user-driven selections. With a network-driven approach, the selection decision is made from the network-side (i.e., service provider). Therefore, it is suitable for tightly integrated environment in which a central controller distributes the traffic flows among different networks. In contrast, with a user-driven approach, users make decisions to select the network in a distributed fashion. Therefore, it does not require any modification and coordination among different networks.

A few works have been done for the network selection problem. The concept of always best connected (ABC) was proposed in [51] which allows a user to choose the most suitable available access networks for current needs at any point and at any time. The great challenges and complexity are also indicated in the paper. In [68], a network selection scheme was developed for an integrated cellular/wireless LAN system. The scheme consists of two parts. First, analytic hierarchy process (AHP) was used to weight the evaluation criteria set according to user preferences and service applications. Then, grey relational analysis (GRA) was used to select the best network. The network selection scheme proposed in [57] is based on a unique decision process in which the compensatory and non-compensatory multi-attribute decision making algorithms are jointly used to assist the terminal in selecting the most suitable network. TOPSIS, a compensatory multi-attribute decision making (MADM) algorithm is described along with disjunctive and conjunctive type of non-compensatory MADM algorithm. In addition, various factors that would influence the selection of an optimal network are identified.

Most of existing schemes for network selection are designed for satisfying users' individual needs, e.g., the bandwidth requirement, connection fees, or the power consumption. In [69], a cost function based network selection strategy was proposed from the system perspective while at the same time considering the users' needs. The system performance of the proposed scheme is analyzed. How-

ever, the dynamics of network selection was not considered. In [8], evolutionary game approach based on replicator dynamics was used to study the dynamics of network selection with complete information. Groups of users in different service areas compete to share the limited bandwidth in the available wireless access networks. Such competition is formulated as a dynamics evolutionary game and the evolutionary equilibrium is considered as the solution of this game. Population evolution and reinforcement learning algorithms for network selection were also proposed for users to reach the equilibrium decisions. Evolutionary game was also used in [70] to model the user churning behavior in heterogeneous wireless networks. The connection arrival and departure process, as well as the rational and irrational churning behaviors are captured using a continuous-time Markov chain model. Besides, non-cooperative and cooperative pricing schemes are proposed based on the evolutionary game framework. In [71], a Markov Decision Process (MDP) based control scheme was proposed for flow assignment among different networks in conjunction with stochastic characterization of the available bit rate and delay of the networks. However, the incomplete information of utilities and the handover cost were not considered.

2.2 Background and Related Works on Spectrum Leasing in Cognitive Radio Networks

In recent years, with the rapid increasing wireless markets due to advances in wireless technologies, the demand for spectrum resource increases dramatically. The number of wireless devices is estimated to reach 100 billion by the year 2025 [72]. The scarce usable spectrum resource is not sufficient for satisfying the huge demand under the current static spectrum allocation policy. In the meanwhile, much of the statically allocated spectrum is wasted in the sense of heavy under-utilization. According to the report published by Federal Communications Commission (FCC), the traffic distribution across the spectrum band is heavily uneven (e.g., heavy on unlicensed spectrum while light on certain licensed spectrum) and much of the licensed spectrum is idle [73]. The unused allocated spectrum at a particular time and specific location is called spectrum holes or white spaces [74]. Such paradox of the current status of spectrum usage naturally requires an effective solution of efficient spectrum management for future wireless networks.

Cognitive radio [75] together with dynamic spectrum access (DSA) [76] is envisioned as a promising solution for efficient spectrum management. Cognitive radio is not only a new radio technology but an intelligent wireless communication system which has the potential to improve many aspects

of communications systems with capabilities of learning from and adapting to its surrounding environment [74]. With the development of cognitive radio technologies, the spectrum utilization can be improved via DSA techniques. The idea for the spectrum utilization improvement is to enable spectrum sharing and to exploit the existence of spectrum holes by making it possible for unlicensed users to access such holes. In cognitive radio networks, the licensed and unlicensed users are termed as primary users and secondary users, respectively. Basically, two main frameworks exist for spectrum sharing between primary users and secondary users, i.e., cooperative DSA and non-cooperative DSA [60]. With cooperative DSA, the primary users can share (e.g., sell or lease) the unused spectrum resource with secondary users in exchange for certain remuneration (e.g., rental fee). The dynamic exclusive use model is used in cooperative DSA. In this case, the spectrum market can be established which could generate new sources of revenue for primary users while satisfying the resource demand of secondary users. With non-cooperative DSA, the secondary users can access the spectrum resource of primary users without permission. The main approach for spectrum access in non-cooperative DSA is opportunistic spectrum access which is also termed as a spectrum overlay approach. The opportunistic spectrum access enables secondary users to exploit the availability of spectrum holes which vary both temporally and spatially without intruding the primary users. To this end, the functions of spectrum sensing (i.e., detect the spectrum holes) [77] in the physical layer and spectrum hole management (e.g., access the spectrum hole) [78] in the medium access layer (MAC) are required.

In this research, we mainly focus on the spectrum leasing under cooperative DSA framework with exclusive use model. A few works addressed issue of the spectrum leasing with different market models in cognitive radio networks. The competitive spectrum leasing in a primary market was analyzed using game theory in [30–32] and using auction methods in [79–82]. In [8], the problem of spectrum sharing among a primary user and multiple secondary users were considered. An oligopoly market was established and a non-cooperative game was formulated to obtain the spectrum allocation for secondary users. Nash equilibrium is considered as the solution. In [31], the problem of spectrum sharing among multiple primary users and multiple secondary users was investigated. In this case, the secondary users can choose to buy spectrum from different primary users by observing the variations in price and quality of spectrum offered. In [32], an implementation framework for the spectrum leasing was proposed and analyzed, where the primary user may lease the spectrum to secondary users in exchange for possible cooperation. In [80], a real time spectrum auction framework was proposed to allocate the spectrum among a set of secondary users. With such auction framework, the spectrum allocation is conflict free and both the auction rev-

enue and spectrum utilization can be maximized. In [79], a truthful and computationally efficient spectrum auction scheme termed as VERITAS was proposed which can maintain truthfulness while maximizing the spectrum utilization. In [82], another spectrum auction framework was proposed which consider the competition of multiple primary users. Specifically, a non-cooperative multiple primary user multiple secondary user auction game was formulated and analyzed. The efficiency of spectrum utilization in the secondary market operated in a fine time scale was investigated in [61]. The feasibility of fine time scale spectrum leasing operation and users' switching operation were discussed in [62], spectrum broker and service broker were assumed for this purpose. Comparing with primary market, only a few recent works studied the spectrum leasing in the secondary market. The competitive spectrum leasing in a secondary market was analyzed using multi-stage games (e.g., Stackelberg game) in [83, 84] and using auction methods in [62, 85]. In [83], a static multiple stage Stackelberg game was formulated for obtaining the optimal amounts sensing and leasing spectrum as well as the optimal retail price for the secondary service provider. In [84], a three layer spectrum market model was proposed and a two level Stackelberg game was formulated to investigate the interaction among spectrum holder, secondary service providers, and secondary users. In [62], an economic framework was proposed for the spectrum secondary market which can be used to guide the dynamic spectrum allocation process as well as the pricing of secondary providers. A knapsack based auction model was used to maximize the revenue and spectrum usage of secondary providers. Also, the competition among multiple secondary providers are captured using game theory. In [85], an optimal auction scheme based on virtual valuation using Vickrey-Clarke-Groves (VCG) mechanism was proposed to maximize the primary user's revenue. Also, some alternative auction formats are classified. However, comparing with spectrum primary market, the spectrum trading in spectrum secondary market is usually operated in a fine time scale (e.g., minutes [61]). In this case, time dependent dynamic decision making is needed. Due to the inherent cyclic dependency of spectrum leasing and service selection, some existing works also use the multi-stage dynamic game structure (e.g., two-level game framework) to analyze the competitions. However, the stage game itself in these works is static (i.e., time irrelevant). That is, the "dynamic" in these works focuses on the dynamic nature of game's play order (i.e., the multi-stage structure), while the dynamic nature in time dependency is not considered. For example, the upper level competition (i.e., competition among providers) is usually modeled as a static non-cooperative game (e.g., pricing game). None of existing works jointly consider the dynamic nature of time dependency as well as the play order for the spectrum leasing in spectrum secondary market.

2.3 Background and Related Works on IEEE 802.16j/m Mobile Relay Networks

The IEEE 802.16 family standards [86–89], commonly known as Worldwide Interoperability for Microwave Access (WiMAX) [90], provide orthogonal frequency-division multiplexing (OFDM) based point to multi-point (PMP) solutions for both fixed and mobile broadband wireless access over long-range transmission. In the earlier standard versions (i.e., 802.16-2004 and former versions), only fixed wireless access is supported by specifying a reliable, high data rate (maximum of 120 Mbps), broadband air interface operating in frequency ranges of 10-66 GHz and 2-11 GHz with large coverage area (maximum transmission range of 50 km) [91]. In this case, WiMAX is envisioned as an effective and economic solution for the last-mile connection in wireless metropolitan area networks (WMANs) especially in rural areas. WiMAX also provides the wireless solution for the backhaul connection in core networks. Since the approval of IEEE 802.16e (i.e., mobile WiMAX), mobility and roaming are supported and therefore mobile WiMAX has been accepted as a 3G standard. Besides the mobility support, several significant technical enhancements in terms of coverage increase, indoor penetration improvement, and frequency reuse are added in IEEE 802.16e [92]. In IEEE 802.16j, relay-based cooperative multi-hop communication is supported. With the deployment of low cost relays, the coverage can be improved in terms of extending the base station (BS) coverage range and addressing coverage hole. Also the capacity can be improved with the use of multiple links. Note that, IEEE 802.16j is backward compatible with IEEE 802.16e end terminals. IEEE 802.16m as next generation mobile WiMAX technology which supports high data rate communications (i.e., up to 100 Mbps and 1 Gbps for mobile and fixed users, respectively) is designed to meet the requirements of International Mobile Telecommunications-Advanced (IMT-Advanced) for 4G systems. Several enhancements and new capabilities with the use of innovative communications technologies (e.g., multi-user MIMO) in the physical and MAC layer of IEEE 802.16m increase the capacity, improve the reliability and user experience (e.g., high mobility support, seamless handover, and better QoS support) [93]. IEEE 802.16m keeps backward compatibility with legacy WiMAX systems. Also, the femto-cells [94] and relay-based cooperative communications are supported.

The performance of relay-based cooperative communications highly relates to the issues such as frame design, networking planning, relay selection, and bandwidth allocation. A few works have been done on these issues in IEEE 802.16j/m mobile relay networks. [95] proposed a frame structure for multihop communications in IEEE 802.16j with backward compatibility with the legacy IEEE

802.16 standards. [96], [97] jointly considered the problems of network planning and bandwidth allocation in IEEE 802.16j relay networks. In [96], a Markov decision process (MDP) based model was used to optimize the amount of bandwidth reserved from base station and admission control for the mobile stations using relay cooperative transmissions so that the utility of a relay station is maximized. Based on the optimal action of relay stations, the long term networking planning problem was addressed considering uncertainties in user mobility and traffic load. In [97], an optimization framework was proposed to exploit the benefits of cooperative relaying. Specifically, the problem of joint relay station placement and bandwidth allocation was formulated as a mixed integer program and a heuristic algorithm was proposed to solve the problem. [98], [99] addressed the relay selection issue, i.e., when to cooperate and whom to cooperate with. In [98], a relay selection scheme considering the possible channel state information (CSI) from the source was proposed. In this work, the source makes the decisions on the time (i.e., when) to cooperative with a relay as well as which relay (i.e., who) to cooperate. The criterion for the optimal relay selection is the maximum instantaneous harmonic mean function of the channel gains. In [99], a two-level game framework was proposed to stimulate cooperation and improve system performance in a multiuser cooperative communication networks. In the framework, the source nodes and relay nodes are considered to be buyers and sellers, respectively. The proposed scheme maximizes the utilities of sellers and help the buyers choose the best seller and decide the optimal amount of power to be bought. In [100], the problem of joint relay selection and link adaption in IEEE 802.16j/m was addressed by a proposed algorithm. Near optimal performance can be achieved using this algorithm. In [101], the selection of an optimal relay station for a vehicular subscriber station was studied. A nonlinear optimization problem was formulated which maximizes the end to end capacity. The problems of bandwidth allocation and frame partitioning in IEEE 802.16j were studied in [102–104]. In [102], the IEEE 802.16j was used to provide network access for vehicular networks. A dynamic bandwidth allocation algorithm was proposed for base stations in vehicular networks to allocate bandwidth to its serving relaying stations considering the requirements of different service types and different velocities. In [103], a dynamic frame partitioning scheme using a dynamic Markov model was proposed for IEEE 802.16 mesh and IEEE 802.16j multi-hop relay networks. With dynamic partition which is adapted according to the demand, the system can support more users and therefore more revenues can be obtained for service providers. However, none of the works jointly considered the problems of dynamic service selection of advanced mobile stations and dynamic bandwidth allocation of advanced base stations in IEEE 802.16m mobile relay networks. The key difference between our work and previous approaches is the dynamic robust optimal decision making for base stations and dynamic resource competition among mobile stations.

2.4 Background and Related Works on Applications of Evolutionary Game and Differential Game

2.4.1 Introduction to Evolutionary Game

Evolutionary game theory [105] is developed as the mathematical tool to study the interaction among bound-rational agents. In evolutionary game theory, an agent adapts the chosen strategy based on its fitness (i.e., payoff). In wireless networks, the agents could be the N users (flows, mobile stations) in certain area competing for the wireless resource which also constitute the population. Denote $n_l(t)$ the number of agents selecting pure strategy $l \in \mathcal{L}$ at time t , where $\mathcal{L}$ is the set of pure strategies. Then the normalized variable $x_l(t) = n_l(t)/N$ denotes the population share of strategy l and $x_l(t) \in [0, 1]$, $\sum_{l=1}^{|\mathcal{L}|} x_l(t) = 1$. The population shares of all $|\mathcal{L}|$ pure strategies constitute the population state which can be denoted by a vector $\mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_{|\mathcal{L}|}(t)]^T \in \mathbb{X}$, where $\mathbb{X}$ is the state space. Denote $\pi(l, \mathbf{x})$ the expected payoff of agents using pure strategy l given the current population state $\mathbf{x}$ which also represents the fitness. Accordingly, the average payoff of the population can be derived as $\pi(\mathbf{x}, \mathbf{x}) = \sum_{l=1}^{|\mathcal{L}|} x_l \pi(l, \mathbf{x})$. The dynamic strategy adaptation process (i.e., selection mechanism) in the evolutionary game can be modeled using replicator dynamics [106] described by the following differential equations:

$$\dot{x}_l = \delta x_l (\pi(l, \mathbf{x}) - \pi(\mathbf{x}, \mathbf{x})), \quad (2.1)$$

where δ is the learning rate of the population which controls the speed of change in selection. According to the replicator dynamics, the population shares of strategies with higher expected payoffs than the average will increase while the ones with lower payoffs will decrease. Evolutionary equilibrium will be reached when all population shares do not vary (i.e., $\dot{x}_l = 0$ for all $l \in \mathcal{L}$). Evolutionary equilibria are the fixed points along the trajectory of the dynamics. That is, when the population state evolves into such fixed points, the population state will remain there. The physical meaning of the evolutionary equilibria is that, at this equilibria, none of the individuals wants to change its strategy since its payoff is equal to the average payoff of the population. Though there may exist multiple evolutionary equilibria, not all of them are stable. An equilibrium population state $\mathbf{x}^* \in \mathbb{X}$ is evolutionary stable if for any population state $\mathbf{y} \in \mathbb{X}$ and $\mathbf{y} \neq \mathbf{x}^*$, there exists some $\varepsilon_{\mathbf{y}} \in (0, 1)$ such that for all $\varepsilon \in (0, \varepsilon_{\mathbf{y}})$, the inequality $\pi(\mathbf{x}^*, \varepsilon \mathbf{y} + (1 - \varepsilon) \mathbf{x}^*) > \pi(\mathbf{y}, \varepsilon \mathbf{y} + (1 - \varepsilon) \mathbf{x}^*)$ holds. An evolutionary stable strategy (ESS) [107] is robust against invaders, that is, all mutants in the population will eventually get extinct. For an ESS population distribution, the small perturbation of

the state does not influence the convergence to the equilibrium state. The physical meaning is that, the average utility of the perturbed state is smaller than the average utility of the equilibrium state. In this way, the ESS is robust against invaders (i.e., perturbation). Actually, in some cases, the replicator dynamics is globally asymptotically stable. In this case, any deviation from the equilibrium will converge back to the equilibrium.

2.4.2 Introduction to Differential Game

Differential games [108] belong to a subclass of dynamic games and are used to study the dynamic decision making problems in competitive systems that evolve in continuous time. The time varying system is abstracted in a vector of state variables the dynamics of which is described by a set of differential equations. In wireless networks, the players could be the M resource providers each of which competes to maximize (minimize) the integrated present value of its instantaneous payoff (cost) function $F_i(\mathbf{x}, u_i, \vec{u}_{-i}, t)$ derived over certain time horizon (either finite time interval $[0, T]$, or infinite time interval $[0, \infty)$), where $\mathbf{x}$ is the system state and $\{u_i, \vec{u}_{-i}\}$ is the strategy profile.

In differential games, the control strategy of each player influences the state evolution of the system as well as the objective function values of other players. Depending on different informational structures, the control strategies can be represented in different ways (i.e., open-loop and closed-loop [109]). An open-loop strategy $u_i = u_i^{ol}(t)$ is a function of time alone which does not need any feedback information (e.g., current state). With open-loop strategies, each player determines the action plan in advance and complies with the plan during the play and the output of the control needs not be measured. In comparison, a closed-loop strategy $u_i = u_i^{cl}(\mathbf{x}, t)$ can use feedback information to adjust the control process. The player adopting a closed-loop strategy commits to the control law instead of a particular control path. Closed-loop (Feedback) control has many advantages over open-loop control (e.g., stabilize unstable systems, handle disturbance). However, obtaining feedback control is more difficult than open-loop control. Also, feedback control needs to estimate the system state from the observed output which adds complexity to the system design. In general, for wireless networks the closed-loop control can be implemented in infrastructure-based networks with the central controller collecting and distributing the state information, while the open-loop control is the appropriate choice for ad hoc networks for which the state measurement and information distribution are costly.

For non-cooperative differential games, open-loop (closed-loop) Nash equilibrium is a reasonable

notion of optimality and is considered to be the solution. To obtain the Nash equilibrium of the differential game, each player needs to solve an optimal control problem given the controls of other players. In this case, both Pontryagin's maximum principle and Hamilton-Jacobi-Bellman equation [110] can be used. Pontryagin's maximum principle is used as necessary condition to find candidate optimal solutions and some other conditions (e.g., convexity of the maximized Hamiltonian function) are needed for proving the optimality. In this case, a two-point boundary value problem (TPBVP) needs to be solved. For Hamilton-Jacobi-Bellman method as sufficient condition, a partial differential equation needs to be solved. In most cases, it is difficult to obtain the analytical solution. However, for certain differential games with special structures, tractable solutions can be obtained.

2.4.3 Related works

Evolutionary game theory has been widely applied in wireless networks. [31, 111–116] addressed different issues (e.g., routing, access network selection, energy management, and congestion control) in wireless networks using evolutionary game theory. Differential game belongs to a subclass of dynamic games and it is used to study the dynamic decision making problems in competitive systems that evolve over continuous time. A linear state differential game was formulated to address the problem of dynamic bandwidth allocation in heterogeneous wireless networks in our work [117]. Optimal control and H_∞ robust control as the approaches for deriving the Nash equilibrium of differential game have wide applications (e.g., bandwidth allocation, rate control, routing, and power control) in wireless networks [118–121]. However, to the best of our knowledge, evolutionary and differential games have not been considered jointly and applied to solve radio resource management problem in wireless networks.

Chapter 3

Network Selection in Heterogeneous Wireless Networks: Evolution with Incomplete Information

3.1 Introduction

In next generation heterogeneous wireless networks, different wireless access technologies are integrated to complement each other. Due to the heterogeneity, dynamic network selection scheme is required not only to achieve seamless mobility, but also to support quality of service (QoS) enhancement and load balancing.

To achieve the best performance with minimum cost, mobile users in heterogeneous wireless networks can perform network selection iteratively in a distributed fashion. The decisions will evolve to the equilibrium point at which the payoff of every user is maximized given the decisions of others and no one can benefit by choosing other networks unilaterally. Compared with the perfect rationality assumption in traditional game theory, it is more realistic to consider the users to be with bounded rationality. We assume that the users are able to perform best response to the current state but these users lack the ability to predict the behaviors of others based on previous behaviors. Therefore, considering the dynamics of network selection and bounded rationality of users, evolutionary game approach is more suitable to investigate the decisions of users over time. In this chapter, we investigate the dynamics of network selection with incomplete information in heterogeneous

wireless networks. In particular, a Bayesian game is formulated by considering users with different bandwidth requirements. Since the preference (i.e., utility) of the mobile user is private information, each user has to make the decision of network selection optimally given only the distributions of the preferences of other users.

To analyze the dynamics of the Bayesian network selection game, the Bayesian best response dynamics and aggregate best response dynamics are applied. Bayesian Nash equilibrium is considered to be the solution of this game. This Bayesian Nash equilibrium can be obtained analytically by solving the aggregate best response dynamics as there is a one-to-one mapping between them. Extensive performance evaluation is performed to investigate the impact of system parameters (e.g., handover cost, price, and the number of users) on the network selection equilibrium distributions. The numerical results show the convergence of the aggregate best response dynamics for the Bayesian network selection game. This result ensures the reachability to the equilibrium decisions of mobile users under incomplete information environment. To the best of our knowledge, this is the first work which applies the technique in Bayesian evolutionary game to the network selection problem.

The rest of the chapter is organized as follows. Section 3.2 presents the system model and assumptions. The user-driven network selection in heterogeneous wireless networks is formulated as a Bayesian game in Section 3.3. The dynamics of the formulated Bayesian network selection game is studied in Section 3.4. Section 3.5 presents the numerical results and analysis. The summary of this chapter is given in Section 3.6.

3.2 System Model and Assumptions

We consider a particular service area in the coverage of a heterogeneous wireless environment consisting of multiple access networks with the same or different types (e.g., WLAN, WiMAX, and 3G). Without loss of generality, we consider service area a with three access networks as shown in Fig. 3.1. There are N selfish users in this area a who are only interested in maximizing their own benefits (i.e., payoffs). Each user periodically receives beacon signals from base stations or access points of the available access networks. Similar to [55], the set of candidate access networks of user i is denoted by $\mathcal{C}$. The user selects to connect to any network in this candidate set $\mathcal{C}$. We assume all users in area a have the same candidate access set which consists of $K = |\mathcal{C}|$ access networks where $|\mathcal{C}|$ is the cardinality of set $\mathcal{C}$.

Figure 3.1: Network model of heterogeneous wireless network with network selection.

Fixed connection fee is assumed for the pricing of each network access, i.e., P_k per connection per unit of time for network k . For the bandwidth allocation, we assume that all users accessing the same network share the available bandwidth equally. The bandwidth of user i received from network k is denoted as τ_i^k where $\tau_i^k = B_k/N_k$, B_k is the available bandwidth of network k , and N_k is the total number of users choosing network k in service area a for $N = \sum_{k \in \mathcal{C}} N_k$.

3.3 Bayesian Evolutionary Game Formulation of Decentralized Network Selection

Utility of one user is a function of allocated bandwidth which depends on the decisions of all users. In this section, Bayesian evolutionary game model is formulated for the decentralized user-driven network selection in heterogeneous wireless networks. Different users have different minimum bandwidth requirements. In this Bayesian network selection game, we use the minimum bandwidth requirement to represent the type of a user which is private information, and the uncertainty of the minimum bandwidth requirement will be taken into account.

3.3.1 Bayesian Network Selection Game

The Bayesian network selection game in heterogeneous wireless networks can be described as follows:

- *Players*: N active users in service area a .
- *Actions*: Action of a player is the selection of access network k from the candidate access set $\mathcal{C}$. Let $\Delta = \{\mathbf{y} = [y_1 \ \cdots \ y_k \ \cdots \ y_K]^T \in R_+^K : \sum_{k \in \mathcal{C}} y_k = 1\}$ denote the set of probability distribution over actions, where y_k represents the probability of choosing network k .
- *Types*: Type of player i is the minimum bandwidth requirement $b_i \in \Gamma$, where Γ is the type space. We assume all users have the same probability distribution of type and the probability density function of which is denoted by $f(b_i)$.
- *Strategies*: Strategy of player i , $s_i : \Gamma \rightarrow \Delta$, is a mapping from the type space to the action distribution set. $s_i(b_i) = [s_i^1(b_i) \ \cdots \ s_i^k(b_i) \ \cdots \ s_i^K(b_i)]^T$ represents the probability distribution over actions given the Bayesian strategy s_i and the minimum bandwidth requirement b_i , where $s_i^k(b_i)$ is equal to y_k . To simplify the presentation, in the following we will use $s_i(b_i)$ and s_i interchangeably. The set of all Bayesian network selection game strategies is denoted as Ω .
- *Payoffs*: For the underlying Bayesian network selection game, we let $\bar{\pi}_i$ denote the expected payoff of player i which is the bandwidth utility minus connection fee. For the evolutionary process, the handover cost needs to be considered, and the instantaneous payoff of user i at decision epoch m is denoted by $\pi_i(m)$.

3.3.2 Payoff Function

The instantaneous payoff (i.e., utility) of user i selecting network k can be expressed as follows:

$$\pi_i^k = \begin{cases} \mathcal{U}(\tau_i^k) - P_k, & \tau_i^k \geq b_i, \\ -P_k, & \tau_i^k < b_i, \end{cases} \quad (3.1)$$

for $i \in \{1, 2, \dots, N\}$ and $k \in \mathcal{C}$ where $\mathcal{U}(\tau_i^k) = \alpha \log(1 + \beta \tau_i^k)$. In particular, $\mathcal{U}(\tau_i^k)$ is a concave function representing the bandwidth utility of user i given its allocated bandwidth τ_i^k from

network k and P_k is the price charged by network k (i.e., connection fee). Two properties are required for such concave bandwidth utility function. First, if the allocated bandwidth is zero, the utility is zero. Second, the increasing rate of the utility decreases with respect to the allocated bandwidth. This is known as the law of diminishing returns. The logarithm function satisfies the above two requirements and connects the user satisfaction level with QoS parameters (e.g., bandwidth) well and therefore is used in many existing works [8, 125]. The use of two parameters α and β in the logarithm utility function provides flexibility to fit the function with empirical data for different types of user's applications. The adjustment of these two parameters within constraints does not change the qualitative properties of the investigated problem and solution. For a particular application, these parameters can be obtained from the surveyed quality of experience (QoE) data by data fitting. If the received bandwidth is less than threshold b_i (i.e., the minimum bandwidth requirement of the user cannot be met), the utility of user i is negative value of price. Otherwise, the utility monotonically increases as the allocated bandwidth increases. This utility function is applicable to many applications in the Internet (e.g., elastic services using transmission control protocol (TCP)) [55].

Let $\delta = \{s_1, s_2, \dots, s_N\}$ denote the strategy profile in Bayesian network selection game which is the set of strategies adopted by N players. To ease the presentation, the strategy profile can be represented as $\delta = \{s_i, \mathbf{s}_{-i}\}$, where s_i is the strategy of user i and $\mathbf{s}_{-i}$ is a vector of strategies of all users except user i . Similarly, the set of types of all users can be denoted as $\{b_i, \mathbf{b}_{-i}\}$ where $\mathbf{b}_{-i}$ is a vector of types of all users except user i .

The expected number of users choosing network k given all other users' strategies $\mathbf{s}_{-i}$ and types $\mathbf{b}_{-i}$ can be obtained from

$$\mathcal{J}_k(\mathbf{s}_{-i}, \mathbf{b}_{-i}) = \sum_{j=1, j \neq i}^N s_j^k(b_j). \quad (3.2)$$

And the expected number of users choosing network k containing all possible type combinations is expressed as follows:

$$\mathcal{J}_k(\mathbf{s}_{-i}) = \int_{b_1} \dots \int_{b_j} \dots \int_{b_N} \mathcal{J}_k(\mathbf{s}_{-i}, \mathbf{b}_{-i}) \prod_{j=1}^N f(b_j) db_N \dots db_j \dots db_1 \quad (3.3)$$

for $j \neq i$.

Therefore, if user i chooses to access network k , the total expected number of users choosing net-

work k becomes

$$\mathcal{L}(N_k) = 1 + \mathcal{J}_k(\mathbf{s}_{-i}). \quad (3.4)$$

Given all other users' strategies, the bandwidth allocated to user i by network k is

$$\tau_i^k(\mathbf{s}_{-i}) = \frac{B_k}{\mathcal{L}(N_k)}. \quad (3.5)$$

Let $\Phi_i^k(\mathbf{s}_{-i}, b_i)$ denote the probability of satisfying the minimum bandwidth requirement of user i by choosing network k given all other users' strategies $\mathbf{s}_{-i}$. $\Phi_i^k(\mathbf{s}_{-i}, b_i)$ can be defined as follows:

$$\Phi_i^k(\mathbf{s}_{-i}, b_i) = \text{Prob}[\tau_i^k(\mathbf{s}_{-i}) > b_i]. \quad (3.6)$$

If user i chooses network k , the expected payoff of user i is expressed as

$$\bar{\pi}_i^k(\mathbf{s}_{-i}, b_i) = \Phi_i^k(\mathbf{s}_{-i}, b_i) \left(\mathcal{U}(\tau_i^k(\mathbf{s}_{-i})) - P_k \right) - [1 - \Phi_i^k(\mathbf{s}_{-i}, b_i)] P_k. \quad (3.7)$$

We have obtained the expected payoff of users for the underlying static Bayesian network selection game. For the dynamics of network selection which is performed iteratively, we consider the cost of handover (e.g., due to delay and loss). The handover cost occurs when the user makes different network selection decision in successive decision epoches (e.g., previously connect to WLAN and now connect to 3G). In this case, at decision epoch $m - 1$, if the user decides to switch the network from $k(m - 1)$ to $k(m)$ at decision epoch m where $k(m) \neq k(m - 1)$, the cost H_i incurs to user i . Therefore, the instantaneous payoff of user i at decision epoch m can be expressed as follows:

$$\pi_i(m) = \begin{cases} \bar{\pi}_i^{k(m)}, & k(m) = k(m - 1) \\ \bar{\pi}_i^{k(m)} - H_i, & k(m) \neq k(m - 1). \end{cases} \quad (3.8)$$

3.3.3 Nash Equilibrium of Underlying Bayesian Network Selection Game

For the Nash equilibrium of the underlying Bayesian network selection game, the expected payoff of user i considering the action distribution $\mathbf{y}$ and the strategy s_i are derived. Based on (3.7),

$\bar{\pi}_i(\mathbf{y}, \mathbf{s}_{-i}, b_i)$ can be obtained from

$$\bar{\pi}_i(\mathbf{y}, \mathbf{s}_{-i}, b_i) = \sum_{k \in \mathcal{C}} \bar{\pi}_i^k(\mathbf{s}_{-i}, b_i) y_k. \quad (3.9)$$

According to (3.9), $\bar{\pi}_i(s_i, \mathbf{s}_{-i}, b_i)$ can be expressed as

$$\bar{\pi}_i(s_i, \mathbf{s}_{-i}, b_i) = \sum_{k \in \mathcal{C}} s_i^k(b_i) \bar{\pi}_i^k(\mathbf{s}_{-i}, b_i). \quad (3.10)$$

Therefore, the expected payoff of user i given strategy profile $\{s_i, \mathbf{s}_{-i}\}$ can be expressed as follows:

$$\begin{aligned} \bar{\pi}_i(s_i, \mathbf{s}_{-i}) &= \int_{\Gamma} \bar{\pi}_i(s_i, \mathbf{s}_{-i}, b_i) f(b_i) db_i \\ &= \int_{\Gamma} \sum_{k \in \mathcal{C}} s_i^k(b_i) \{ \Phi_i^k(\mathbf{s}_{-i}, b_i) \left(\mathcal{U}(\tau_i^k(\mathbf{s}_{-i})) - P_k \right) - \left(1 - \Phi_i^k(\mathbf{s}_{-i}, b_i) \right) P_k \} f(b_i) db_i. \end{aligned} \quad (3.11)$$

Let $\mathcal{R}(\mathbf{s}_{-i})$ denote the best response of user i given other users' strategies $\mathbf{s}_{-i}$. For every type of user i , the best response can be obtained from

$$\mathcal{R}_i(\mathbf{s}_{-i}, b_i) = \arg \max_{\mathbf{y} \in \Delta} \bar{\pi}_i(\mathbf{y}, \mathbf{s}_{-i}, b_i). \quad (3.12)$$

The Bayesian strategy profile δ^* is a Bayesian Nash equilibrium if and only if no user can benefit by unilaterally changing his strategy even just an action under a certain type [1]. Therefore, the Bayesian Nash equilibrium of the network selection game is defined as follows:

Definition 1: A strategy profile $\delta^* = \{s_i^*, \mathbf{s}_{-i}^*\}$ is a Nash equilibrium if and only if $\forall s_i \in \Omega$, $\bar{\pi}_i(s_i^*, \mathbf{s}_{-i}^*) > \bar{\pi}_i(s_i, \mathbf{s}_{-i}^*)$ for all $i \in \{1, 2, \dots, N\}$, and for every i and b_i , $s_i^*(b_i) = \mathcal{R}_i(\mathbf{s}_{-i}^*, b_i)$.

3.4 Dynamics of Network Selection Game

In this section we study the strategy evolution based on the Bayesian best response dynamics which models the dynamic behavior of Bayesian games. Also, this dynamics is used to obtain the Bayesian Nash equilibrium.

3.4.1 Bayesian Evolutionary Dynamics for Network Selection Game

Bayesian best response dynamics provides a method to study the evolution in Bayesian games [122]. Compared with best response dynamics, the state of Bayesian best response dynamics is a set of Bayesian strategies rather than action distributions. In the following, the Bayesian best response dynamics for the network selection game will be introduced.

First, two important operators $\mathcal{E} : \Omega \rightarrow \Delta$ and $\mathcal{B} : \Delta \rightarrow \Omega$ in Bayesian best response dynamics are defined for the context of network selection game. Notice that these two operations are not specific to any user i (e.g., to represent the strategy of a user, s will be used rather than s_i) since this network selection game is symmetric (i.e., the action set and type distribution are identical to every user in the same area).

Definition 2: Let $\mathcal{E}(s)$ denote the aggregate network selection distribution induced by Bayesian strategy $s \in \Omega$. This aggregate network selection distribution can be expressed as follows:

$$\mathcal{E}(s) = (\mathcal{E}(s)_1, \mathcal{E}(s)_2, \dots, \mathcal{E}(s)_K) \quad (3.13)$$

where $\mathcal{E}(s)_k = \int_{\Gamma} s^k(b) f(b) db$, $k \in \mathcal{C}$ denotes the proportion of users in a service area choosing network k under strategy s .

Definition 3: Let $\mathcal{B}(\mathbf{x})$ denote the best response correspondence to the social aggregate network selection distribution $\mathbf{x} = \begin{bmatrix} x_1 & \dots & x_k & \dots & x_K \end{bmatrix}^T$, where x_k represents the aggregate proportion of users choosing network k . This best response correspondence can be expressed as follows:

$$\mathcal{B}(\mathbf{x}) = \arg \max_{\mathbf{y} \in \Delta} \pi(\mathbf{y}, \mathbf{x}, b) \quad (3.14)$$

where $\pi(\mathbf{y}, \mathbf{x}, b)$ is the obtained payoff under selection distribution $\mathbf{y}$, social aggregate distribution $\mathbf{x}$, and minimum bandwidth requirement b .

According to Definition 2, the aggregate distribution $\mathbf{x}$ can be induced by certain Bayesian strategies. Therefore, $\bar{\pi}_i(\mathbf{y}, \mathbf{s}_{-i}, b_i)$ is equivalent (i.e., after operation $\mathcal{E}$) to $\pi(\mathbf{y}, \mathbf{x}, b)$ in the static underlying game. And for the dynamics (i.e., considering the handover cost), $\pi(\mathbf{y}, \mathbf{x}, b)$ can be obtained jointly from (3.7), (3.8), and (3.9). In fact, it is also easy to show the equivalence between the best response $\mathcal{R}(\mathbf{s}_{-i})$ defined in (3.12) and the best response correspondence $\mathcal{B}(\mathbf{x})$ in the static underlying game.

The best response correspondence in complete information game is an inclusion rather than a function since it may contain multiple best responses. However, in Bayesian best response dynamics, if the type distribution is sufficiently diverse and smooth, $\mathcal{B}(\mathbf{x})$ returns a single value and hence the inclusion becomes a function. In the following analysis, we assume $\mathcal{B}(\mathbf{x})$ yields single value.

According to Definitions 2 and 3, each Bayesian strategy s induces the network selection distribution $\mathcal{E}(s)$, and the best response to the distribution can be expressed as $\mathcal{B}(\mathcal{E}(s))$. Based on [122], the definition of Bayesian best response dynamics for network selection game is given as follows.

Definition 4: The Bayesian best response dynamics is described by the law of motion on the space of Bayesian strategies as follows:

$$\dot{s} = \mathcal{B}(\mathcal{E}(s)) - s. \quad (3.15)$$

For continuity of (3.15), L^1 norm is used to measure the distances of Bayesian strategies. The rest points of the Bayesian best response dynamics form the set of Bayesian Nash equilibria. As shown in [122], $\mathcal{E}(s)$ and $\mathcal{B}(\mathbf{x})$ has the Lipschitz continuous property. Therefore, the Bayesian best response dynamics is Lipschitz continuous which guarantees the existence and uniqueness of solutions to the dynamics.

3.4.2 Aggregate Dynamics for Network Selection Game

Due to the complexity in analyzing the Bayesian best response dynamics in the L^1 space, aggregate best response dynamics is applied. According to [122], the definition of aggregate best response dynamics for Bayesian network selection game is given as follows:

$$\dot{\mathbf{x}}_t = \gamma(\mathcal{E}(\mathcal{B}(\mathbf{x}_t)) - \mathbf{x}_t) \quad (3.16)$$

where $\mathbf{x}_t$ is the aggregate network selection distribution at time t and γ is the learning rate which represents the proportion of users adjusting their strategies towards best response to the current network selection distribution at each selection epoch. Operators $\mathcal{E}(\cdot)$ and $\mathcal{B}(\cdot)$ have the same definitions with those in Bayesian best response dynamics, i.e., (3.13) and (3.14), respectively. Notice that $\mathbf{x}(m)$ is different from $\mathbf{x}_t$. $\mathbf{x}(m)$ is a network selection distribution point which is the weighted best response (i.e., considering the learning rate) to $\mathbf{x}(m-1)$, while $\mathbf{x}_t$ describes the path from $\mathbf{x}(m-1)$ to $\mathbf{x}(m)$.

Similar to the Bayesian best response dynamics, if the type distribution is sufficiently diverse and smooth, $\mathcal{B}(\mathbf{x})$ is single valued and Lipschitz continuous. Therefore, the solution to the aggregate best response dynamics exists and is unique. The rest points of the aggregate dynamics for Bayesian network selection game form the set of equilibrium network selection distributions. A one-to-one correspondence is established between the Bayesian Nash equilibrium and the equilibrium distribution of aggregate best response dynamics [122]. Therefore, we can analyze the Bayesian dynamics through the aggregate dynamics.

For a given initial aggregate network selection distribution $\mathbf{x}(0)$, the best response Bayesian strategy $\mathcal{B}(\mathbf{x}(0))$ can be obtained from (3.14), and $\mathcal{E}(\mathcal{B}(\mathbf{x}(0)))$ which is the expected network selection distribution induced by $\mathcal{B}(\mathbf{x}(0))$ can be calculated according to Definition 2. To reach the equilibrium distribution, many iterations of network selection need to be performed to construct the convergence trajectory. Within an epoch, $\mathbf{x}_t$ describes the path from the initial distribution state to its best response distribution. And at next selection epoch, this best response distribution is considered as the initial state. Therefore, the network selection distribution at network selection epoch m can be obtained from

$$\mathbf{x}(m) = \gamma \mathcal{E}(\mathcal{B}(\mathbf{x}(m-1))) + (1 - \gamma) \mathbf{x}(m-1) \quad (3.17)$$

where γ is the learning rate. The impact of system parameters (e.g., learning rate γ and handover cost H_i) on the equilibrium distributions will be analyzed analytically in next section.

3.5 Numerical Analysis

3.5.1 Parameter Setting

We consider the coverage area a of a IEEE 802.11b access point which is also in the cover of a IEEE 802.16 cell and a CDMA-based cellular network cell. Area a is totally overlapped as shown in Fig. 1. For the IEEE 802.11b-based WLAN, we assume the maximum saturation throughput is 7 Mbps [123]. For the IEEE 802.16-based access network, the transmission bandwidth is assumed to be 20 MHz and the signal-to-noise ratio (SNR) at the receiver is assumed to be 10dB. The spectral efficiency is assumed to be 1.5 bit/s/Hz. Therefore, the transmission rate of IEEE 802.16 access network is 30 Mbps in a single cell. Considering the bandwidth usage by other users in the same cell of IEEE 802.16 and cellular access network, we assume the available bandwidth of IEEE 802.16 access network and cellular access network for area a are 5 Mbps and 2 Mbps, respectively. We

consider fixed connection fees for IEEE 802.11b, IEEE 802.16, and cellular which are 0.3, 0.2, and 0.4, respectively.

For the bandwidth utility function, we set $\alpha = 1$ and $\beta = 1$. The number of users in area a is assumed to be 20. For Bayesian best response dynamics, we set the learning rate to be $\gamma = 0.2$. We set an identical handover cost to be $H_i = 0.1$. We assume users' minimum bandwidth requirements follow the uniform distribution. We set the initial proportion of users choosing IEEE 802.11b, IEEE 802.16, and cellular access network to be $x_1 = 0.3$, $x_2 = 0.5$, and $x_3 = 0.2$, respectively.

3.5.2 Numerical Results

We first investigate the phase portrait and the convergence property of the aggregate best response dynamics for our network selection game in Fig. 3.2. In this case, the network selection distribution states are mapped from the three-dimensional space to a triangle in the two-dimensional space. For example, the three vertexes A, B, and C represent the selection distribution $[x_1 = 1, x_2 = 0, x_3 = 0]$, $[x_1 = 0, x_2 = 1, x_3 = 0]$, and $[x_1 = 0, x_2 = 0, x_3 = 1]$, respectively. The phase portrait shows the solution trajectories of the aggregate dynamics. As shown in Fig. 3.2, the dynamics converges to equilibrium distributions from different initial states. For example, the dynamics from an initial state D follows the trajectory which composes of linear orbits pointing towards the best responses to reach the equilibrium network selection distribution. The equilibrium distributions correspond to the cyclically stable set (CSS) [124] which may have a single or multiple points.

The strategy adaption trajectory from initial point $[x_1 = 0.5, x_2 = 0.3, x_3 = 0.2]$ is shown in Fig. 3.3. The proportions of users choosing WLAN and WiMAX access networks gradually increase. This increase is due to the larger bandwidth available in WLAN and WiMAX in which the probability of satisfying user's minimum bandwidth requirement is higher. While the proportion of users choosing cellular network decreases due to the limited bandwidth and the high price. We observe that the aggregate distribution converges to the equilibrium point $[x_1 = 0.4785, x_2 = 0.4688, x_3 = 0.0526]$.

Fig. 3.4 shows the amount of allocated bandwidth at equilibrium distribution under different number of users. Since the proportion of users choosing WLAN and WiMAX networks is high, the allocated bandwidth to each user by these two networks decreases with the increase of the number of users. In contrast, for the cellular network, the allocated bandwidth to each user has only small change since the cellular network is only chosen by a small proportion of users under various equilibrium

Figure 3.2: Phase portrait of the aggregate best response dynamics.

distributions.

Fig. 3.5 shows the adaptation of network selection under different handover cost. This handover cost is due to the handover delay or packet loss. When the handover cost is small (e.g., $H_i = 0.05$), users are more willing to churn to another network if the payoff is higher (e.g., due to larger allocated bandwidth or cheaper price). Using WiMAX network as an example, when the handover cost is $H_i = 0.05$, the proportion of users choosing WiMAX access network is fluctuated. When the handover cost is large (e.g., $H_i = 0.25$), it may not be worth to switch the network even though the allocated bandwidth is larger or the price is lower. Therefore, the proportion of users choosing to churn to other networks is much lower (Fig. 3.5).

Fig. 3.6 shows the equilibrium distributions under different price of WLAN. As the price of WLAN increases, the proportion of users choosing WiMAX and cellular networks increases. In this case, the proportion of users choosing cellular access network is smaller than that of WiMAX due to the smaller capacity of cellular network. It is worth noting that even though the price of WLAN is 0, not all users will choose WLAN. Since the WLAN network can become congested and some users are willing to select different networks even though their costs are higher.

The learning rate of the users is varied and the proportion of users selecting WiMAX network is shown in Fig. 3.7 as an example. When this rate is small, the impact of the adjusted strategies on the aggregate network selection distribution is small and the variation of the solution trajectory is small.

Figure 3.3: Trajectory of network selection strategy adaption.

On the contrary, the fluctuation is observed in the trajectory before converging to the equilibrium when the learning rate is large.

3.6 Conclusion

The network selection in heterogeneous wireless networks has been formulated as the Bayesian evolutionary game to study the equilibrium decisions of the users under incomplete information about the preference (i.e., utility). The dynamics of this network selection game has been analyzed using Bayesian best response dynamics and aggregate best response dynamics. The rest points of this aggregate dynamics determine the equilibrium distributions which correspond to the Bayesian Nash equilibria. Numerical results have shown the convergence of aggregate best response dynamics and the analysis has shown the impact of system parameters (e.g., price, handover cost, and the number of users) on the equilibrium distributions.

Figure 3.4: Allocated bandwidth to each user at equilibrium under different number of users.

Figure 3.5: Impact of the handover cost on the dynamics.

Figure 3.6: Impact of the price of WLAN on the equilibrium distribution.

Figure 3.7: Impact of the learning rate on the dynamics.

Chapter 4

Dynamic Spectrum Leasing and Service Selection in Spectrum Secondary Market of Cognitive Radio Networks

4.1 Introduction

Cognitive radio networks aim at improving spectrum utilization via dynamic spectrum access (DSA) techniques [59]. Two main frameworks exist for spectrum sharing, i.e., cooperative DSA and non-cooperative DSA [60]. With cooperative DSA, some spectrum holders may lease portion of spectrum to secondary users in exchange for appropriate remuneration. In this case, two types of spectrum market can be established, i.e., primary market and secondary market [61]. In a primary market, the spectrum leasing is performed directly among primary users (i.e., the spectrum holders) and secondary users mainly in a long-term manner (e.g., weeks or months). While in a secondary market, the spectrum can be traded among secondary users themselves in a finer time scale (e.g., several minutes [61]) to adapt to the time-varying demand and channel condition. Though extensive studies have been performed for the primary market, the competitive spectrum leasing in secondary market is largely overlooked. One important application scenario of the secondary market is the existence of secondary service providers¹ which lease spectrum resources from spectrum brokers to provide service to secondary users. In this case, the spectrum is traded dynamically depending on

¹Also termed as mobile virtual network operators the number of which is over 400 worldwide as of February 2009 [29].

the demand of secondary providers² which in turn depends on the number of subscribed secondary users. The objective of this spectrum trading is to maximize the profit of secondary providers while maximizing the performance satisfaction of secondary users. Naturally, two issues arise. First, the secondary users compete for the service and dynamically adapt the service selection strategies according to the time-varying perceived performance. Second, according to the dynamic service selection distribution of users, the competitive secondary providers dynamically lease spectrum to obtain the maximum profits.

To address the issues of dynamic service selection and dynamic spectrum leasing, we proposed a hierarchical (i.e., two-level) game framework with inherent dynamic nature (Fig. 4.1) to jointly obtain the optimal strategies for both secondary users and providers given the strategies of others. The dynamic competition of service selection among secondary users is modeled as an evolutionary game. The strategy adaptation of this evolutionary game is subject to the control of providers in terms of spectrum size leased to provide service for the users. Also, the providers observe the service selection distribution of the secondary users and decide the optimal spectrum size leased dynamically. This dynamic spectrum leasing is modeled as a differential game and the state of which is the service selection distribution.

Figure 4.1: Two-level dynamic game framework.

The rest of this chapter is organized as follows: Section 4.2 introduces the motivation, differences, and advantages of the proposed novel hierarchical dynamic game framework. Section 4.3 presents

²For the rest of this chapter, “secondary user” and “user” as well as “secondary service provider” and “provider” are used interchangeably.

the system model. The evolutionary game formulation for strategy adaption of the secondary users is presented in Section 4.4. The competition of dynamic spectrum leasing among providers is formulated as a differential game in Section 4.5. The calculation and implementation schemes of Nash equilibrium are also presented in this section. Section 4.6 presents the Stackelberg game formulation for the dynamic spectrum leasing with hierarchical play. The numerical results and analysis are presented in Section 4.7. Section 4.8 draws the conclusions.

4.2 A Novel Hierarchical Dynamic Game Framework

4.2.1 Motivation

In this chapter, a novel hierarchical (i.e., two-level) dynamic game framework with inherent dynamic nature based on evolutionary game [105] and differential game [108] is proposed to address the joint dynamic decision making problems in competitive environments with hierarchical structure. This hierarchical structure exists widely in wireless networks especially in infrastructure based networks. The competition exists at each level and is naturally dynamic. For example, in cognitive radio networks, the primary users (i.e., spectrum providers) competitively lease portion of spectrum to secondary users in exchange for appropriate remuneration, while secondary users compete for the available spectrum resource. The secondary users can select and churn to different spectrum providers depending on cost, geographic availability, and spectrum availability. According to the selection distribution of secondary users, the primary users dynamically decide the optimal amount of spectrum to be leased as well as the price charged.

The competition for wireless resources among users (e.g., mobile stations, flows) is modeled as an evolutionary game in the lower level. To illustrate the motivation for the use of evolutionary game, we first describe the four main characteristics of this lower-level resource competition. First, the number of users involved in this competition is commonly very large which is difficult to handle in traditional non-cooperative game theory. While for evolutionary game, the introduction of the concept of population makes the problem readily to solve. Second, this resource competition is naturally dynamic due to the characteristics of wireless networks. Evolutionary game has inherent dynamic nature through the use of differential equations (stochastic differential equations for stochastic models) describing the population state evolution. Third, the users in this level usually lack of complete information and computation capability. In this case, the rational player assump-

tion in traditional non-cooperative game theory is rarely held. Also, due to the lack of complete information, all users gradually learn during the strategy adaption process. Evolutionary game considering bound-rational users is the suitable tool for modeling this learning process. Last but not least, fairness is an important issue of resource allocation in competitively environment. For evolutionary game model using replicator dynamics, all users in the same population receive the same individual utility as the average at the equilibrium. Therefore, the fairness can be guaranteed.

Different from the lower-level users, the players in the upper-level are usually the central controllers (e.g., base stations) of the networks and are considered to be rational. The upper-level players competitively and dynamically make the optimal decision to maximize their own objective functions. This dynamic optimal decision making process in competitive environment can be formulated as a differential game. The main motivation for the use of differential game stems from the fact that the upper-level players need to make the optimal decision dynamically according to the time varying service selection distribution of the lower-level users. Traditional non-cooperative game theory has been developed mostly for static analysis. Although the dynamics of the decision-making process can be modeled in the extensive form, this extensive form becomes intractable for most game settings with reasonable complexity. Also for modeling the competition in the upper level, some works only consider the steady state of the lower-level dynamics based on which static games are formulated. However, in this case the performance behaved in the transit period which may be long is largely ignored. Though repeated game is often used for dynamic decision making in competitive environment, the convergence of sub-stage game cannot be guaranteed within a small time period. Therefore, differential game with inherent dynamic nature which uses the lower-level dynamics as optimization state constraints and considers the entire time period is the suitable tool for this task.

4.2.2 Links Between the Two Levels

In the proposed two-level framework, the population state of the lower-level evolutionary game describes the state of the upper-level differential game. This state describes the links between the two levels in two folds. On one hand, the population state evolution of the evolutionary game is governed by the control strategies of the upper-level players. On the other hand, in the upper-level differential game, the optimization of payoff functions of players need to take this state evolution into account as constraints.

4.2.3 Differences and Advantages

As discussed in Chapter 1, for dynamic games the dynamic nature can be presented from game's play order and time dependency. Many existing works also proposed various hierarchical dynamic game frameworks for the modeling in wireless networks (e.g., [29, 79]). However, the "dynamic" in these existing works focuses on the dynamic nature of game's play order (i.e., the multi-stage structure), while the "dynamic" in our work jointly considers the dynamic nature of time dependence as well as the game's play order.

To illustrate the differences between our work and previous works, the time correlations of different game models are shown in Fig. 4.2 and described as follows:

- *Static game*: As shown in Fig. 4.2 a, the static game only models the competition at a particular time.
- *Repeated game*: As shown in Fig. 4.2 b, the repeated game partly considers the time dependency by repeating the game at different time points along the time horizon. However, the time intervals between successive time points are ignored.
- *Dynamic game considering time dependency*: As shown in Fig. 4.2 c, the dynamic game fully considering time dependency (e.g., differential game) obtains the optimal dynamic strategy over the entire time horizon.
- *Multi-stage game with static stage game*: As shown in Fig. 4.2 d, the multi-stage game with static stage game only models the competition at a particular time. The dynamic nature is reflected only in the game's play order.
- *Multi-stage game with dynamic stage game*: As shown in Fig. 4.2 e, the multi-stage game with dynamic stage game considers both the dynamic nature in play order (i.e., multi-stage structure) and the dynamic nature in time dependency (i.e., considering the whole time horizon).

The key difference of our work from the previous works is the consideration of dynamic nature in time dependency at each stage which was not studied in previous works. In most previous works, the competition of service providers and the competition of users are formulated as static games at corresponding stages (i.e., the stage game itself is static.). Though the entire game framework is dynamic in terms of a game's play order, the dynamic nature in time dependency is not considered.

Figure 4.2: Different types of game models.

The importance for considering the time dependency can be represented by the performance gain in terms of payoff for the upper level game and convergence speed for the lower level game. That is, with time dependent strategy (i.e., dynamic game strategy) the players gain larger payoff than that using time irrelevant strategy (i.e., static game strategy). The reason is due to the fact that for a multi-stage game with static stage games, backward induction is used and the Nash equilibrium of the upper level static game is obtained only considering the equilibrium state of the lower level game. In other words, the impact of the transient states (the states before reaching the equilibrium state) of the lower level game to the upper level game was ignored. That is, the optimization (i.e., best response) in the upper level does not consider the transient states. Such transient states appear before the convergence to the equilibrium state in the lower level game. For example, for the practical implementation of Nash equilibrium of a static game, the algorithms (e.g., best response dynamics) are needed by players to iteratively update their strategies to reach the equilibrium states.

Though repeated game partly consider the time dependency by repeating the same game finite or infinite times along the time horizon. The one-shot game itself is static, and the same analysis above can be applied.

4.3 System Model

We consider a particular service area a in a cognitive radio system consisting of M secondary service providers and N active secondary users as shown in Fig. 4.3. A secondary market of spectrum is assumed in which the secondary service providers lease the spectrum resources from spectrum brokers to provide service to secondary users. Cooperative dynamic spectrum access (DSA) is used for spectrum sharing and the interference caused to primary users by the secondary system is not considered here (i.e., this is based on exclusive-use DSA model). In this case, the spectrum brokers first broadcast the spectrum availability, then the providers lease certain amount of available spectrum by sending the request messages to the spectrum brokers. Once the spectrum resources are allocated, the providers broadcast the service to the secondary users. Note that in this model the primary market is not directly used for secondary users leasing spectrum from holders but is assumed to be used for spectrum brokers to lease spectrum from primary users (spectrum holders) in a long-term basis.

For differentiating services among providers, Paris Metro Pricing (PMP) model [126] is used, the idea of which is to charge different prices for identical transport service (e.g., best effort service).

Figure 4.3: System model of a spectrum secondary market in a cognitive radio network.

In this case, the services with higher prices would attract less traffic, and therefore provide better service [127]. Different secondary providers can lease the spectrum from the same or different spectrum brokers depending on cost, geographic availability, and spectrum availability. Also, the service time duration for a secondary provider in area a can be either finite or infinite which are denoted as T and ∞ , respectively.

The secondary users can select and churn to different providers independently and dynamically according to the perceived instantaneous utility which depends on the allocated bandwidth and price. Here, we assume that a secondary user can access the service from only one provider at a time. Denote $c_i(t) \in [0, F_i]$ the spectrum size that provider i leases from spectrum brokers at time t , where F_i is the maximum spectrum size that provider i can lease. Accordingly, the bandwidth (i.e., transmission rate) that provider i can provide is $c_i(t)k_i = c_i(t) \log_2(1 + \gamma_i \bar{h}_i)$, where k_i is the average spectral efficiency in bits/s/Hz, γ_i is the signal to interference plus noise ratio (SINR), $\bar{h}_i$ is the average channel gain (i.e., due to user mobility in the service area) [128]. For the resource

allocation among secondary users, we consider an un-weighted round-robin scheme where all users subscribed to the same provider share the available bandwidth equally. Therefore, the bandwidth of a secondary user allocated from provider i is denoted as $\tau_i(t) = c_i(t)k_i/n_i(t)$, where $n_i(t)$ represents the total number of secondary users choosing provider i at time t , and $\sum_{i=1}^M n_i(t) = N$.

4.4 Modeling the Evolution of Service Selection

The bound-rational secondary users in service area a compete for the service by selecting different candidate secondary providers. The objective of this service selection is to maximize the individual satisfaction (i.e., utility) jointly considering performance and cost. Initially, all secondary users choose one of the providers randomly. To obtain better utility, each secondary user can adapt the service selection decision periodically according to the price and the time-varying observed service quality which depends on both the size of available spectrum and the current congestion condition.

4.4.1 Formulation of the Evolutionary Game

The dynamic competition of service selection among secondary users is formulated as an evolutionary game. This is the lower-level game in the proposed two-level framework. The players are the N active secondary users in area a which also constitute the population in the context of evolutionary game. In practice, users can be grouped into different populations according to geographical location, application type, and service quality requirements. The population members have similar channel condition and service requirements. Also, the providers with service differentiation typically allocate different resources for different application types. Accordingly, users in the same population typically have the same competing objective, while the objectives differ among populations. Therefore, it is possible to model the competition among population members independent of other populations. In this case, the complex competition among users in the different populations is independent, and the analysis can be performed for each population separately. Without loss of generality, we focus on the analysis of one population (i.e., best effort users), in which the service requirement is elastic. The same analysis can be applied to other populations (e.g., guaranteed service users) as well.

The strategy of each player is the selection of a particular provider. The payoff of a player is the utility that quantifies the performance satisfaction level. As a function of allocated bandwidth

$\tau_i(t)$ and charged price p_i (i.e., price charged by secondary service provider i per user per unit of time), the utility is assumed to be equivalent to the rate utility per unit price, which is defined as $u_i(t) = \mathcal{U}(\tau_i(t))/p_i$, where $\mathcal{U}(\tau_i(t))$ is the rate utility for which we consider a linear function $\mathcal{U}(\tau_i(t)) = \alpha\tau_i(t)$, where α is an adjustable constant to provide flexibility to fit the function with empirical data for different types of user applications. Note that α can be different for different populations. Let $x_i(t) = n_i(t)/N$ denote the proportion of secondary users in area a choosing provider i at time t which is also referred to as the population share. The population shares of all M providers is the population state denoted by a vector $\vec{x}(t) = [x_1(t) \ \cdots \ x_i(t) \ \cdots \ x_M(t)]^T$, where $x_i(t) \in [0, 1]$ and $\sum_{i=1}^M x_i(t) = 1$ for all $t \geq 0$. Accordingly, the expected payoff of a secondary user selecting provider i at time t can be defined as $\pi_i(t) = u_i(t)$. The average payoff of the population can be derived as $\bar{\pi}(t) = \sum_{i=1}^M x_i(t)\pi_i(t)$.

In the service selection evolutionary game, the strategy with higher payoff (e.g., choosing a provider with larger offered bandwidth and less number of accessing users) than the average can be replicated by other secondary users in the sense of evolution. For a small period of time, the rate of service selection strategy change can be modeled by replicator dynamics which is defined as follows:

$$\frac{\partial x_i(t)}{\partial t} = \dot{x}_i(t) = \delta x_i(t) (\pi_i(t) - \bar{\pi}(t)), \quad (4.1)$$

with initial condition $\vec{x}(0) = \vec{x}_0 \in \mathbb{X}$, for all $i \in \{1, 2, \dots, M\}$, where δ is the learning rate of the population which controls the speed of change in service selection, $\vec{x}_0$ is the initial population state, and $\mathbb{X} \subseteq \mathbb{R}^M$ is the state space which contains all possible population distributions. In contrast to static non-cooperative game, all players are programmed to use only pure strategy in replicator dynamics. For a particular user, it can randomly switch to provider which gives a higher utility. The switching probabilities to different providers are proportional to the difference of the observed utility and average utility. Details of implementation algorithm (e.g., users' switching operation) can be found in [8].

Note that, in the current system model, we mainly consider the population of users with best effort applications whose service requirement is elastic (i.e., there is no specific performance guarantee). For the population of users running applications with performance requirements, either of the following two approaches can be used to support QoS guarantee.

- *Admission control mechanism:* Admission control mechanism can be used to limit the number of ongoing users. In particular, the service provider admits a service request only when the

minimum performance requirements can be satisfied. This admission control mechanism will be done first (i.e., to admit or reject the request), and then, the service selection of the users will be performed. From the user perspective, the decision of admission control (i.e., accepted or rejected) has an impact on the user utility function.

- *Withdraw strategy:* If the performance requirement of a user cannot be met, user can choose to withdraw the request for service. Given a population distribution, the selecting of each strategy associates with different utility. In this case, the user can compare the utilities of all strategies and make a decision accordingly.

4.4.2 Existence and Uniqueness of Solution under Control

To investigate the solution trajectory of this replicator dynamics, we first substitute $\pi_i(t)$ and $\bar{\pi}(t)$ into the right hand side of (5.5), which is denoted by $f_i(x_i(t), c_i(t), t)$ for simplicity. We can obtain $f_i(x_i(t), c_i(t), t) = \delta\alpha \left(c_i(t)k_i/p_i - x_i(t) \sum_{l=1}^M c_l(t)k_l \right) / N$, for all $i \in \{1, 2, \dots, M\}$. We observe that the control variable $c_i(t)$ of provider i influences the evolution of the population state. The existence and uniqueness of the solution to this dynamical system under control are discussed in the following theorem:

Theorem 1 *For the dynamical system $\dot{x}_i(t) = f_i(x_i(t), c_i(t), t)$, $i \in \{1, 2, \dots, M\}$, with initial condition $\vec{x}(0) = \vec{x}_0$, if $c_i(t)$ is a measurable function on $[0, \infty)$, then there exists a unique solution $\vec{x}(t)$ defined for all $t \in [0, \infty)$.*

Proof: For fixed t , the partial derivative of $f_i(x_i(t), c_i(t), t)$ with respect to $x_i(t)$ is continuous, and if $c_i(t)$ is measurable on $[0, \infty)$, then $f_i(x_i(t), c_i(t), t)$ is measurable for fixed $x_i(t)$ on the same interval. Moreover, for any given closed bounded set $K \in \mathbb{X}$ and interval $[a, b] \in [0, \infty)$, there always exists a positive number m ($m \geq |x|$ for all $x \in K$) to construct an integrable function $m_i(t) = |\delta\alpha c_i k_i / p_i N| + m \left| \delta\alpha \sum_{l=1}^M c_l k_l / N \right|$ such that $|f_i(x_i(t), c_i(t), t)| \leq m_i(t)$ and $|\partial f_i(x_i(t), c_i(t), t) / \partial x_i(t)| \leq m_i(t)$ for all $(x_i, t) \in K_i \times [a, b]$. Finally, $|f_i(x(t), c_i(t), t) - f_i(y(t), c_i(t), t)| = \delta\alpha \sum_{l=1}^M c_l k_l |x(t) - y(t)| / N$. Denote $L = \max\{\delta\alpha \sum_{l=1}^M c_l k_l / N\}$, we can obtain $|f_i(x(t), c_i(t), t) - f_i(y(t), c_i(t), t)| \leq L|x(t) - y(t)|$ which implies that $f_i(x_i(t), c_i(t), t)$ satisfies the global Lipschitz condition. Based on the above arguments, we state that the solution to this dynamical system under control is unique and exists globally [129]. ■

4.4.3 Equilibrium and Stability Analysis

The replicator dynamics describes the variation rate of all population shares. The proportion of secondary users choosing certain provider increases if the corresponding payoff is above the average. Evolutionary equilibrium which is defined as the stable fixed point of the replicator dynamics is considered to be the solution of the service selection evolutionary game. The evolutionary equilibrium can be obtained by solving the system of algebraic equation $\dot{x}_i(t) = 0, \forall i$. Both boundary equilibria (i.e., there exists $x_i = 1$) and interior equilibrium (i.e., $x_i \in (0, 1), \forall i$) exist. The boundary equilibria are not stable in the sense that any small perturbation will make the system deviate from the equilibrium state. In contrast, we will show below that the interior equilibrium is asymptotically stable.

We state that the dynamical system defined in (5.5) is stable and asymptotically converges to the evolutionary equilibrium point from any initial interior population state $\vec{x}_0$. We rewrite (5.5) in a vector form as $\dot{\vec{x}} = \mathbf{A}\vec{x} + \mathbf{b}$, where $\mathbf{b} = \left[\delta\alpha c_1(t)k_1/p_1N \quad \cdots \quad \delta\alpha c_M(t)k_M/p_MN \right]^T$, and $\mathbf{A}$ is a $M \times M$ dimensional matrix, the characteristic function of which can be expressed as follows:

$$\det(\gamma\mathbf{I} - \mathbf{A}) = \left(\gamma + \frac{\delta\alpha}{N} \sum_{l=1}^M c_l(t)k_l \right)^M = 0, \quad (4.2)$$

where $\mathbf{I}$ is an identity matrix. It is straightforward that $\mathbf{A}$ always has M repeated negative eigenvalues which shows the asymptotically stable property of the system [130].

4.5 Dynamic Spectrum Leasing

With the dynamic behavior of secondary users, the providers can optimally control the spectrum size leased from the spectrum brokers. Increasing the leased spectrum size is a natural way to improve the performance and also to attract more secondary users. However, the increased cost for leasing more spectrum may result in a reduced total profit of providers. In this case, we formulate the competition of dynamic spectrum leasing among providers as a non-cooperative differential game. This is the upper-level game in the proposed two-level framework (Fig. 4.1) which takes the dynamic service selection of secondary users into account. As common, we first consider the simultaneous play model. The asynchronous play model will be studied in section V.

4.5.1 Non-cooperative Dynamic Spectrum Leasing as a Differential Game

In the non-cooperative simultaneous play differential game, each of the M providers competes to maximize the present value of its objective function derived over certain time horizon (either bounded time interval $[0, T]$, or unbounded time interval $[0, \infty)$) by controlling the leased spectrum size. For simplicity of notation, $[0, T)$ is used in both cases, where $[0, T) = [0, T]$ for finite time horizon and $[0, T) = [0, \infty)$ for infinite time horizon.

The set of players is composed of all available providers in area a . For provider i , the strategy is the dynamic control of the spectrum size $c_i(t)$ leased from the spectrum brokers at time t . Similar to the common notation used in game theory, $\Phi = \{c_i(t), \vec{c}_{-i}(t)\}$ denotes the strategy profile of this differential game, where $\vec{c}_{-i}(t)$ is a vector of strategies of all players except player i .

Depending on different informational structures, the dynamic control strategies can be represented in different ways (i.e., open-loop and closed-loop). An open-loop strategy is a function of time alone which does not need any feedback information (e.g., current state) from the system. With an open-loop strategy, a player determines the action plan in advance and complies with the plan during the play. Also, the output of the control process does not need to be measured. In comparison, a closed-loop strategy can use feedback information to adjust the control process. The player adopting a closed-loop strategy commits to the control law instead of a particular path. The closed-loop control can be implemented in infrastructure-based networks with the central controller collecting and distributing the state information, while the open-loop control is the appropriate choice for ad hoc networks for which the state measurement and information distribution are costly.

In this chapter, we consider both the open-loop strategy and the closed-loop strategy for applicability to different network structures (i.e., infrastructure-based and ad hoc network). The open-loop and closed-loop strategy spaces of provider i are defined as $\mathcal{C}_i^{OL} = \{c_i(t) | c_i(t) \text{ is measurable on } [0, T) \text{ and } F_i \geq c_i(t) \geq 0 \text{ for all } t \in [0, T)\}$ and $\mathcal{C}_i^{CL} = \{c_i(\vec{x}(t), t) | c_i(\vec{x}(t), t) \text{ is measurable on } [0, T) \text{ and } F_i \geq c_i(\vec{x}(t), t) \geq 0 \text{ for all } (t, x) \in [0, T) \times [0, 1]\}$, respectively. For simplicity of notation, c_i and $\mathcal{C}_i$ are used where $c_i = c_i(t)$, $\mathcal{C}_i = \mathcal{C}_i^{OL}$ for open-loop strategy and $c_i = c_i(\vec{x}(t), t)$, $\mathcal{C}_i = \mathcal{C}_i^{CL}$ for closed-loop strategy. Also, as common in control theory, the time parameter will be omitted whenever no ambiguity occurs. The state of this upper level differential game is the population state $\vec{x}$ of the underlying service selection game. The replicator dynamics differential equations in (5.5) describe how the current state and the providers' controls influence the rate of change of the state.

In the spectrum leasing differential game, the control strategies of all providers (i.e., players) influence the evolution of the state as well as their own and their opponents' profit functions. For each of M providers, the maximization of profit under dynamic competition becomes an optimal control problem subject to the constraints (e.g., state evolution differential equations) given the control strategies of other providers. The instantaneous profit of provider i choosing control strategy c_i is expressed as $J_{\text{ins}}^i(c_i, \vec{c}_{-i}) = p_i N x_i - \beta_i c_i^2$, where $p_i N x_i$ is the instant revenue of provider i at time t , $\beta_i c_i^2$ is the cost of provider i leasing spectrum of size c_i , β_i is a weighted cost factor depending on which spectrum broker provider i leases spectrum from. The convex cost term is used to prevent aggressive behavior of certain provider which could result in a monopoly market. Specifically, all providers can request certain amount of spectrum from spectrum brokers with reasonable cost; however, the cost for leasing extra spectrum increases rapidly. In this case, the aggressive behavior of providers is prevented. Note that we assume that the prices (i.e., p_i and β_i) are fixed, and the pricing issue is out of scope of this chapter.

With the definition of strategy, time horizon, instantaneous profit, for each rational provider $i \in \{1, 2, \dots, M\}$, the formulation of the optimal control can be expressed as follows:

$$\max_{c_i} J^i(c_i, \vec{c}_{-i}) = \int_0^T e^{-\rho t} J_{\text{ins}}^i(c_i, \vec{c}_{-i}) dt, \quad (4.3)$$

$$\text{s.t. } \dot{x}_i = \delta x_i (\pi_i - \bar{\pi}), \quad \vec{x}(0) = \vec{x}_0, \text{ and } c_i \in \mathcal{C}_i,$$

where ρ is the discount rate influencing the discounted present value of future profit.

4.5.2 Open-loop and Closed-loop Nash Equilibrium

For the dynamic spectrum leasing differential game, we assume that no binding agreements can be made among providers. In this case, Nash equilibrium is a reasonable notion of optimality and is considered to be the solution. At the equilibrium point, every provider obtains maximized profit given the controls of other providers. The definition of an optimal (i.e., best response) spectrum control strategy under competition is given as follows:

Definition 1: A spectrum control strategy c_i^* is optimal for provider i if the inequality condition $J^i(c_i^*, \vec{c}_{-i}) \geq J^i(c_i, \vec{c}_{-i})$ holds for all feasible control paths c_i in the spectrum leasing differential game.

According to the Definition 1, the definitions of both open-loop Nash equilibrium and closed-loop Nash equilibrium for the spectrum leasing differential game are given as follows:

Definition 2: Denote $c_i \in \mathcal{C}_i$ the open-loop (closed-loop) strategy of provider i . The strategy profile $\{c_i^*, \vec{c}_{-i}^*\}$ is an open-loop (closed-loop) Nash equilibrium if for each provider $i \in \{1, 2, \dots, M\}$, c_i^* is an optimal control strategy given other providers' control strategies $\vec{c}_{-i}^*$.

4.5.3 Calculation of Nash Equilibrium

(1) Open-loop Nash Equilibrium

To obtain the open-loop Nash equilibrium, each provider needs to solve an optimal control problem (i.e., given controls of competitors). In this case, Pontryagin's maximum principle can be used as the necessary conditions to find the candidate optimal strategies. Then, the sufficient conditions are given for proving their optimality. To this end, the present value Hamiltonian function of provider i is denoted as $H_i(\vec{x}, c_i, \vec{c}_{-i}, \vec{\lambda}_i, t) = J_{\text{ins}}^i(c_i, \vec{c}_{-i}) + \sum_{j=1}^M \lambda_{ij} \dot{x}_j$, where $\vec{\lambda}_i = [\lambda_{i1} \ \dots \ \lambda_{ij} \ \dots \ \lambda_{iM}]^T$ and λ_{ij} is the costate variable of provider i associated with state x_j . This costate variable can be regarded as the shadow value representing the variation in the total profit induced by the evolution of population state. With the definition of Hamiltonian function, the corresponding maximized Hamiltonian function H^* is defined as

$$H_i^*(\vec{x}, \vec{\lambda}_i, t) = \max_{c_i} \{H_i(\vec{x}, c_i, \vec{c}_{-i}, \vec{\lambda}_i, t) | c_i \geq 0\}. \quad (4.4)$$

As proven in optimal control theory, the optimal control strategy of problem (5.11) must also maximize the corresponding Hamiltonian function given any triple $(\vec{x}, \vec{\lambda}_i, t)$. Therefore, all candidate optimal strategies for open-loop Nash equilibrium have to satisfy the following necessary optimality conditions:

$$c_i^* = \arg \max_{c_i} H_i(\vec{x}, c_i, \vec{c}_{-i}, \vec{\lambda}_i, t), \quad (4.5)$$

$$\dot{\lambda}_{ij} = \rho \lambda_{ij} - \frac{\partial H_i^*(\vec{x}, \vec{\lambda}_i, t)}{\partial x_j}, \quad (4.6)$$

$$\dot{x}_i = \delta x_i (\pi_i - \bar{\pi}), \quad (4.7)$$

for all $i, j \in \{1, 2, \dots, M\}$, where (4.6) is the adjoint equation for describing the costate variable and $\lambda_{ij}(T) = 0, \lim_{T \rightarrow \infty} \lambda_{ij}(T) = 0$ are the transversality conditions for finite horizon and infinite

horizon, respectively. In the case that the strategy space $\mathcal{C}_i$ does not depend on the system state $\vec{x}$, the maximized Hamiltonian $H_i^*(\vec{x}, \vec{\lambda}_i, t)$ in the right hand side of (4.6) can be replaced by the Hamiltonian $H_i(\vec{x}, c_i, \vec{c}_{-i}, \vec{\lambda}_i, t)$ [131].

Setting the first order partial derivative $\partial H_i(\vec{x}, c_i, \vec{c}_{-i}, \vec{\lambda}_i, t) / \partial c_i = 0$, we obtain the candidate optimal strategy for provider i as follows:

$$c_i^* = \max \left(0, \frac{1}{2\beta_i} \left(\frac{\delta\alpha}{p_i N} k_i \lambda_{ii} (1 - x_i) - \sum_{j=1, j \neq i}^M \frac{\delta\alpha}{N} k_i \lambda_{ij} x_j \right) \right). \quad (4.8)$$

Also, after derivation, (4.6) can be expressed as follows:

$$\begin{cases} \dot{\lambda}_{ii} = \rho \lambda_{ii} - p_i N + \lambda_{ii} \sum_{l=1}^M c_l k_l, \\ \dot{\lambda}_{ij} = \rho \lambda_{ij} + \lambda_{ij} \sum_{l=1}^M c_l k_l, \quad i \neq j. \end{cases} \quad (4.9)$$

From (4.9), we observe that if $\lambda_{ii} < 0$, then $\dot{\lambda}_{ii} < 0$. In this case, the transversality condition $\lambda_{ij}(T) = 0$ ($\lim_{T \rightarrow \infty} \lambda_{ij}(T) = 0$) cannot be satisfied. Therefore, we state that $\lambda_{ii} > 0$ for all $t \in [0, T]$. Similarly, we can obtain $\lambda_{ij} < 0, i \neq j$. This result proves the intuition argument that the increase of provider i 's population share x_i increases the value of its objective function, while the increase of population share of other providers will decrease this value. In this case, (4.8) can be reduced to

$$c_i^* = \frac{1}{2\beta_i} \left(\frac{\delta\alpha}{p_i N} k_i \lambda_{ii}^* (1 - x_i^*) - \sum_{j=1, j \neq i}^M \frac{\delta\alpha}{N} k_i \lambda_{ij}^* x_j^* \right). \quad (4.10)$$

Substituting (4.10) into (4.7) and (4.9), we can obtain a system of differential equations with respect to state $\vec{x}$ and costate $\vec{\lambda}_i$ with initial conditions and transversality conditions. Solving for the solution for this system is a two point boundary value problem (TPBVP). Note that the method for obtaining the solution differs for finite horizon model and infinite horizon model. For finite horizon modeling, the TPBVP with fixed T can be solved directly (e.g., numerically), while for infinite horizon modeling, the common method is to consider the solutions that converge to the steady state. Denote x_i^* and λ_{ij}^* for all $i, j \in \{1, 2, \dots, M\}$ the solution to the above system. Then, $c_i^* = \left(\delta\alpha k_i \lambda_{ii}^* (1 - x_i^*) / p_i - \sum_{j=1, j \neq i}^M \delta\alpha k_i \lambda_{ij}^* x_j^* \right) / (2N\beta_i)$ is the candidate optimal strategy of provider i for the open-loop optimal control problem and $\Phi^* = \{c_i^*, \vec{c}_{-i}^*\}$ is the candidate solution

for the Nash equilibrium. The following theorem shows that the candidate solution Φ^* is indeed the Nash equilibrium.

Theorem 2 *For the formulated spectrum leasing differential game, the strategy profile $\Phi^* = \{c_i^*, \vec{c}_{-i}^*\}$ is an open-loop Nash equilibrium (i.e., Φ^* satisfies the Definition 2).*

Proof: It is straightforward that the construction of strategy profile Φ^* satisfies all the necessary conditions. The following arguments constitute the sufficient conditions for optimality. First, the state space $\mathbb{X}$ is a convex compact set, and Theorem 1 has proven the existence and uniqueness of solution to the state evolution dynamics (5.5) under control. Also, for all $t \in [0, T]$, the Hamiltonian function H is concave and continuously differentiable with respect to $\vec{x}$. According to these conditions, we can state that the obtained strategy profile Φ^* is an open-loop Nash equilibrium [131]. ■

(2) Closed-loop Nash Equilibrium

An open-loop strategy is a degenerate form of a closed-loop strategy. Naturally, using a non-degenerate Markovian representation of an open-loop strategy is one way to construct the closed-loop strategy. In this case, we can obtain a closed-loop strategy

$$c_i(\vec{x}, t) = \left(\delta \alpha k_i \lambda_{ii}^* (1 - x_i) / p_i - \sum_{j=1, j \neq i}^M \delta \alpha k_i \lambda_{ij}^* x_j \right) / (2N\beta_i) \quad (4.11)$$

for provider i simply by replacing the computed optimal state path $\vec{x}^*$ by the measured current state $\vec{x}$. Accordingly, we can obtain a closed-loop Nash equilibrium strategy profile $\varphi = \{c_i(\vec{x}, t), \vec{c}_{-i}(\vec{x}, t)\}$. However, the equilibrium constructed in this way is not Markov perfect (i.e., not a subgame perfect closed-loop Nash equilibrium [131]) which implies that φ is not necessarily optimal for the states that are not in the optimal path. Note that φ and Φ^* generate exactly the same state trajectory and control path if no disturbance and deviation occur.

To obtain a Markov perfect Nash equilibrium, Hamilton Jacobi Bellman (HJB) equation can be used. Denote $V^i(\vec{x}, t)$ the optimal value function of provider i for problem (5.11). $V_t^i(\vec{x}, t)$ and $V_{x_i}^i(\vec{x}, t)$ are the derivatives of optimal value with respect to t and x_i , respectively. Then, the HJB

equation for the formulated spectrum leasing differential game is expressed as follows:

$$\rho V^i(\vec{x}, t) - V_t^i(\vec{x}, t) = \max_{c_i} \left(J_{\text{ins}}^i(c_i(\vec{x}, t), \vec{c}_{-i}(\vec{x}, t)) + \sum_{j=1}^M V_{x_j}^i(\vec{x}, t) \dot{x}_j \right). \quad (4.12)$$

Maximizing the right hand side of (4.12) with respect to $c_i(\vec{x}, t)$ and assuming an interior solution yield $c_i^*(\vec{x}, t) = \left(\delta \alpha k_i V_{x_i}^i(\vec{x}, t) (1 - x_i) / p_i - \sum_{j=1, j \neq i}^M \delta \alpha k_i V_{x_j}^i(\vec{x}, t) x_j \right) / (2N\beta_i)$, which is similar to (4.10) if we interpret the costate variable as the shadow value of the optimal value function. Though similar in the structure, the solutions are different. Substituting $c_i^*(\vec{x}, t)$ into (4.12), we can obtain a system of partial differential equations as follows:

$$\rho V^i(\vec{x}, t) - V_t^i(\vec{x}, t) = J_{\text{ins}}^i(c_i^*(\vec{x}, t), \vec{c}_{-i}^*(\vec{x}, t)) + \sum_{j=1}^M V_{x_j}^i(\vec{x}, t) \dot{x}_j, \quad (4.13)$$

together with the boundary condition $V^i(\vec{x}, T) = 0$ for all $i \in \{1, 2, \dots, M\}$. Generally, it is difficult to find a closed form solution of this partial differential system. However, as a standard method for autonomous dynamic games, the problem can be simplified for infinite horizon models, where the common interests are confined to find stationary Markov perfect equilibrium. In this case, the control strategy and optimal value function of provider i are independent of time t and can be represented by $c_i(\vec{x})$ and $V^i(\vec{x})$ instead of $c_i(\vec{x}, t)$ and $V^i(\vec{x}, t)$, respectively. Accordingly, $V_t^i(\vec{x}) = 0$. Therefore, (4.13) can be rewritten as:

$$\rho V^i(\vec{x}) = J_{\text{ins}}^i(c_i^*(\vec{x}), \vec{c}_{-i}^*(\vec{x})) + \sum_{j=1}^M V_{x_j}^i(\vec{x}) \dot{x}_j. \quad (4.14)$$

Denote $V^{i*}(\vec{x})$ the solution to (4.14). Then,

$$c_i^*(\vec{x}) = \left(\delta \alpha k_i V_{x_i}^{i*}(\vec{x}) (1 - x_i) / p_i - \sum_{j=1, j \neq i}^M \delta \alpha k_i V_{x_j}^{i*}(\vec{x}) x_j \right) / (2N\beta_i) \quad (4.15)$$

is the optimal feedback strategy of provider i for spectrum control under dynamic competition. Accordingly, the closed-loop strategy profile $\varphi^* = \{c_i^*(\vec{x}), \vec{c}_{-i}^*(\vec{x})\}$ constitutes the Markov perfect Nash equilibrium for the spectrum leasing differential game. Note that different from Pontryagin's maximum principle method as the necessary conditions, Hamilton Jacobi Bellman (HJB) method provides sufficient conditions for optimality. Therefore, the strategy profile φ^* obtained using HJB method is not only a candidate but indeed the Nash equilibrium.

Table 4.1: Control strategy implementation schemes for spectrum leasing.

Control strategies	Implementation schemes
Open-loop Nash equilibrium control	For each provider i do , Compute the TPBVP (4.6), (4.7) and obtain the optimal control c_i using (4.10) offline. Store c_i in a table; Look up the table, update the leased spectrum size at each decision point; End for
Stackelberg Nash equilibrium control	For each follower provider j do , Collect the strategies information of all leader providers; Compute the TPBVP $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \vec{\mathbf{c}}_L^*, \vec{\lambda}_j)$, $\dot{\lambda}_j = \mathbf{h}(\mathbf{x}, \vec{\mathbf{c}}_L^*, \vec{\lambda}_j)$ and obtain the optimal control using $c_j = g_j(\mathbf{x}, \vec{\lambda}_j)$ offline, store c_j in a table; Look up the table, update the leased spectrum size at each decision point; End for For each leader provider i do , Compute the TPBVP (4.30) and obtain the optimal control using $c_i = g_i(\mathbf{x}, \vec{\lambda}_i, \vec{\lambda}_j, \vec{\mu})$ offline, store c_i ; Look up the table, update the leased spectrum size at each decision point; End for
Markov perfect Nash equilibrium control	For each provider i do , Measure the current system state $\mathbf{x}$; Update the leased spectrum size c_i using the solution of (4.26) at each decision point; End for

(3) Implementation Scheme

We consider discretized schemes for the implementation of optimal control strategies for providers. In this case, the time horizon T is partitioned into time frames with equal length. At the beginning of each time frame (i.e., decision point), all providers update the leased spectrum size and all secondary users update the service selection strategy. For the open-loop control, the providers compute the optimal control path offline and update the control actions at each decision point. While for closed-loop control strategies, the providers also need to measure the current system state to update the control actions. The detailed implementation schemes are provided in table 4.1. The service selection adaption algorithm for users is the same as that proposed in [8] and therefore is not included here.

4.5.4 An Illustrative Example

We illustrate the non-cooperative spectrum leasing differential game with a specific system model with two providers (i.e., $M = 2$). In this case, we can use x and $1 - x$ interchangeable with x_1 and x_2 to represent the proportions of secondary users choosing providers 1 and 2, respectively.

(1) Open-loop Solution

Following the aforementioned Pontryagin's maximum principle method, it is not difficult to obtain the optimal candidate strategies for both providers as follows:

$$c_i^* = (-1)^{i+1} \frac{\alpha \delta k_i \lambda_i^* (1 - x_i^*)}{2 p_i N \beta_i}, \quad i = 1, 2, \quad (4.16)$$

where λ_i^* and x^* are the solutions of the following system of differential equations with respect to system state and costate for two providers case, i.e.,

$$\begin{cases} \dot{x} = \frac{\alpha^2 \delta^2}{2 \beta_1 p_1 N^2} \left(k_1^2 \lambda_1 - 2 k_1^2 \lambda_1 x + k_1^2 \lambda_1 x^2 + \frac{\beta_1 p_1 k_2^2 \lambda_2 x^2}{\beta_2 p_2} \right) \\ \dot{\lambda}_1 = \rho \lambda_1 - N p_1 + \frac{\alpha^2 \delta^2}{2 \beta_1 p_1 N^2} \left(k_1^2 \lambda_1^2 - k_1^2 \lambda_1^2 x - \frac{\beta_1 p_1 k_2^2 \lambda_1 \lambda_2 x}{\beta_2 p_2} \right) \\ \dot{\lambda}_2 = \rho \lambda_2 + N p_2 + \frac{\alpha^2 \delta^2}{2 \beta_1 p_1 N^2} \left(k_1^2 \lambda_1 \lambda_2 - k_1^2 \lambda_1 \lambda_2 x - \frac{\beta_1 p_1 k_2^2 \lambda_2^2 x}{\beta_2 p_2} \right), \end{cases} \quad (4.17)$$

with boundary conditions $\vec{x}(0) = \vec{x}_0$, and transversality conditions $\lambda_i(T) = 0$ for finite horizon and $\lim_{T \rightarrow \infty} \lambda_i(T) = 0$ for infinite horizon, where $i = 1, 2$.

For finite horizon case, this two point boundary value problem can be numerically solved directly. For infinite horizon modeling, we consider the solution that converges to the steady state at which $\dot{x} = \dot{\lambda}_i = 0$. To analyze the optimal control strategy for secondary service providers in infinite horizon cases, we first rewrite (4.16) as $\lambda_i = (-1)^{i+1} 2 \beta_i p_i N c_i / \alpha \delta k_i (1 - x_i)$, $i = 1, 2$. Accordingly, the derivative of λ_i , $i = 1, 2$ with respect to t can be obtained as follows:

$$\dot{\lambda}_i = (-1)^{i+1} \frac{2 \beta_i p_i N}{\alpha \delta k_i} \left(\frac{\dot{c}_i}{1 - x_i} + \frac{c_i \dot{x}_i}{(1 - x_i)^2} \right). \quad (4.18)$$

Equal the right hand side of (4.18) and (4.17) yields the following dynamic relationships:

$$\dot{c}_i = c_i \left(\rho + \frac{\alpha \delta k_j c_j}{N(1 - x_i)} - \frac{p_i \alpha \delta k_i (1 - x_i)}{2 \beta_i c_i} \right), \quad i, j = 1, 2, i \neq j.$$

Let $\dot{x}_i = 0$, we can obtain $x_i/x_j = k_i c_i p_j / (k_j c_j p_i)$ and accordingly $x_i = k_i c_i p_j / (k_i c_i p_j + k_j c_j p_i)$ for $i, j = 1, 2, i \neq j$. Setting $\dot{c}_i = 0$, $i = 1, 2$, it is straightforward that

$$\rho + \frac{\alpha \delta k_j c_j}{N(1 - x_i)} - \frac{p_i \alpha \delta k_i (1 - x_i)}{2 \beta_i c_i} = 0. \quad (4.19)$$

Substituting x_i into (4.19) and after some derivations, we can obtain

$$c_i = \frac{2 \beta_j c_j}{p_j k_i k_j N} (k_i c_i + k_j c_j) \left(k_i c_i + k_j c_j + \frac{\rho N}{\alpha \delta} \right), \quad i, j = 1, 2, i \neq j. \quad (4.20)$$

According to (4.20), we can obtain the following relationships:

$$\frac{x_i}{x_j} = \frac{k_i c_i p_j}{k_j c_j p_i}, \quad \frac{c_i}{c_j} = \sqrt{\frac{p_i \beta_j}{p_j \beta_i}}, \quad i, j = 1, 2. \quad (4.21)$$

That is, the ratio of population share of the two providers at the steady state depends on the offered bandwidth, and the ratio of spectrum size leased by the two providers directly depends on the price and cost. With these relationships, numerical method can be used to obtain the control paths and the state trajectories for both providers.

(2) Closed-loop Solution

To obtain the Markov perfect solution, the HJB method provided above is used. Denote $V^i(x)$ the stationary optimal value function for provider i , and $V_x^i = dV_i(x)/dx$ for $i = 1, 2$. Then the corresponding HJB equations for providers $i, j = 1, 2, i \neq j$ are described as

$$\rho V^i(x) = \max \left(p_i N x_i - \beta_i c_i^2 + (-1)^{i+1} V_x^i \frac{\alpha \delta}{N} [k_i c_i x_j / p_i - k_j c_j x_i] \right). \quad (4.22)$$

Maximizing the right hand side of (4.22) with respect to control variable c_i yields

$$c_i = (-1)^{i+1} \alpha \delta k_i V_x^i (1 - x_i) / 2 \beta_i p_i N \quad (4.23)$$

for $i = 1, 2$. For the purpose of obtaining a system of equations not involving the optimal value functions, we can rewrite c_i as $V_x^i = (-1)^{i+1} 2 \beta_i p_i N c_i / \alpha \delta k_i (1 - x_i)$ the integration of which are

$$V^i(x) = (-1)^{i+1} \frac{2 \beta_i p_i N}{\alpha \delta k_i} \int_0^x \frac{c_i}{1 - s_i} ds_i + d_i, \quad (4.24)$$

where s_i are integration variables. Substituting $V_x^i, V^i(x)$ into (4.22) and collecting items yield

$$p_i N x_i + \beta_i c_i^2 - \frac{2 \beta_i k_j x_i c_i c_j}{k_i (1 - x_i)} - \rho \left((-1)^{i+1} \frac{2 \beta_i p_i N}{\alpha \delta k_i} \int_0^x \frac{c_i}{1 - s_i} ds_i + d_i \right) = 0. \quad (4.25)$$

For small positive discount ratio in infinite horizon case, the above equations can be simplified and the closed form solution can be obtained by solving the unique positive solution of the following equations.

$$c_i^2 - \frac{2 k_j x_j c_i c_j}{k_i (1 - x_i)} + \frac{p_i N x_i}{\beta_i} = 0, \quad i, j = 1, 2, i \neq j. \quad (4.26)$$

4.6 Dynamic Spectrum Leasing with Hierarchical Play

In Section 4.5, the dynamic spectrum leasing is formulated as a differential game in which all providers make their moves simultaneously. In reality, the simultaneous play of all providers cannot always be guaranteed or some providers may have a priority to make their moves before others. For example, some providers with close relationship (e.g., partner) with spectrum broker may be able to lease spectrum and offer service before other providers (i.e., followers). In this case, we formulate the dynamic spectrum leasing with hierarchical play as a Stackelberg differential game and an open-loop Stackelberg equilibrium is considered to be the solution.

For the Stackelberg spectrum leasing differential game, the players, strategy spaces, state dynamics, instantaneous profits, and the objective functions are the same with those of the simultaneous play game. Differently, in Stackelberg game, the players are categorized into two groups (i.e., the leader group and the follower group). Denote $\mathcal{L}$ and $\mathcal{F}$ the set of leader providers and follower providers, respectively, where $\mathcal{L} \cap \mathcal{F} = \emptyset$ and $|\mathcal{L}| + |\mathcal{F}| = M$. The leader providers make their spectrum leasing decisions first, and the follower providers can observe these decisions and decide the optimal leased spectrum size accordingly. This hierarchical information structure and the best responses of follower providers to any leaders' strategies have been considered by leader providers when the optimal strategies are obtained. Comparing with simultaneous play model, it is always true that the leader can obtain a higher (if not equal) payoff in the hierarchical setting since it can at least choose the Nash equilibrium strategy.

Different from simultaneous play model, the maximization of profit for leader providers in Stackelberg game is also subject to the costate constraints (e.g., equation (4.9)) of all followers besides the system state constraints. To obtain the Stackelberg equilibrium, Pontryagin's maximum principle can also be used as the necessary conditions. However, the added constraints make the Hamiltonian functions for the leader providers much more complicated than the simultaneous play counterpart. Denote H_i^L and H_j^F the Hamiltonian functions for leader provider $i \in \mathcal{L}$ and follower provider $j \in \mathcal{F}$, respectively. Specifically,

$$H_j^F(\vec{x}, \vec{c}_L^*, \vec{c}_F, \vec{\lambda}_j, t) = J_{\text{ins}}^j(\vec{c}_L^*, \vec{c}_F) + \sum_{k=1}^M \lambda_{jk} \dot{x}_k, \quad (4.27)$$

for all $j \in \mathcal{F}$, where $\vec{c}_L^*$ is the vector of observed leaders' strategies, $\vec{c}_F$ is the vector of followers'

strategies, and λ_{jk} is the costate variable with respect to state x_k for follower provider j , and

$$H_i^L(\vec{x}, \vec{c}_L, \vec{c}_F, \vec{\lambda}_i, \vec{\lambda}_j, \vec{\mu}_i, t) = J_{\text{ins}}^i(\vec{c}_L, \vec{c}_F) + \sum_{l=1}^M \lambda_{il} \dot{x}_l + \sum_{j \in \mathcal{F}} \sum_{k=1}^M \mu_{jk}^i \dot{\lambda}_{jk}, \quad (4.28)$$

for all $i \in \mathcal{L}$, where λ_{il} is leader i 's costate variable of state x_l and μ_{jk}^i is leader i 's costate variable with respect to the costate λ_{jk} of follower j (i.e., the costate of costate).

For a Stackelberg game, backward induction method is usually used to obtain the solution. In this case, we first apply the first order condition to the followers' Hamiltonian functions (4.27) to obtain the candidate optimal strategies of follower providers which have the same form as that in (4.10). It is straightforward that the candidate optimal strategy c_j^{F*} for follower j is a function of state $\vec{x}$ and costate $\vec{\lambda}_j$. For simplicity of notation, we let $c_j^{F*} = g_j(\vec{x}, \vec{\lambda}_j)$. Similarly, we can obtain the adjoint equation of follower j as $\dot{\lambda}_{jk} = \rho \lambda_{jk} - \partial H_j^F / \partial x_k = h_{jk}(\vec{x}, \vec{c}_L^*, \vec{c}_F, \vec{\lambda}_j)$, where $k \in \{1, 2, \dots, M\}$. Substituting $c_j^{F*} = g_j(\vec{x}, \vec{\lambda}_j)$ into the state dynamics and adjoint equations yields $\dot{x}_k = f_k(\vec{x}, \vec{c}_L^*, \vec{\lambda}_j)$ and $\dot{\lambda}_{jk} = h_{jk}(\vec{x}, \vec{c}_L^*, \vec{\lambda}_j)$, which are not explicit with followers' strategies. As mentioned, the adjoint equations of $\dot{\lambda}_{jk}$ are also the constraints of leaders' objective functions. In this case, the Hamiltonian function of leader i can be rewritten as follows:

$$H_i^L(\vec{x}, \vec{c}_L, \vec{c}_F, \vec{\lambda}_i, \vec{\lambda}_j, \vec{\mu}, t) = J_{\text{ins}}^i(\vec{c}_L, \vec{c}_F) + \sum_{l=1}^M \lambda_{il} f_l(\vec{x}, \vec{c}_L^*, \vec{\lambda}_j) + \sum_{j \in \mathcal{F}} \sum_{k=1}^M \mu_{jk}^i h_{jk}(\vec{x}, \vec{c}_L^*, \vec{\lambda}_j) \quad (4.29)$$

With a similar method, we can obtain the candidate optimal strategy of leader provider i as $c_i^{L*} = g_i(\vec{x}, \vec{\lambda}_i, \vec{\lambda}_j, \vec{\mu})$, the adjoint equations $\dot{\lambda}_{il} = \rho \lambda_{il} - \partial H_i^L / \partial x_l = h_{il}(\vec{x}, \vec{c}_L^*, \vec{\lambda}_i, \vec{\lambda}_j, \vec{\mu})$, and the adjoint equations $\dot{\mu}_{jk}^i = \rho \mu_{jk}^i - \partial H_i^L / \partial \lambda_{jk} = v_{jk}^i(\vec{x}, \vec{c}_L^*, \vec{\lambda}_i, \vec{\lambda}_j, \vec{\mu})$. Substituting $c_i^{L*} = g_i(\vec{x}, \vec{\lambda}_i, \vec{\lambda}_j, \vec{\mu})$ into $\dot{x}_k$, $\dot{\lambda}_{jk}$, $\dot{\lambda}_{il}$, and $\dot{\mu}_{jk}^i$ yields the following system of differential equations:

$$\begin{cases} \dot{x}_k = f_k(\vec{x}, \vec{\lambda}_i, \vec{\lambda}_j, \vec{\mu}), & \dot{\lambda}_{jk} = h_{jk}(\vec{x}, \vec{\lambda}_i, \vec{\lambda}_j, \vec{\mu}), \\ \dot{\lambda}_{il} = h_{il}(\vec{x}, \vec{\lambda}_i, \vec{\lambda}_j, \vec{\mu}), & \dot{\mu}_{jk}^i = v_{jk}^i(\vec{x}, \vec{\lambda}_i, \vec{\lambda}_j, \vec{\mu}), \end{cases} \quad (4.30)$$

for all $i \in \mathcal{L}, j \in \mathcal{F}, k, l \in \{1, 2, \dots, M\}$, with initial condition $\vec{x}(0) = \vec{x}_0$ and transversality condition $\vec{\lambda}(T) = 0, \lim_{T \rightarrow \infty} \vec{\lambda}(T) = 0$ for finite horizon and infinite horizon, respectively. From $\dot{\lambda}_{jk} = h_{jk}(\vec{x}, \vec{c}_L^*, \vec{\lambda}_j)$, $j \in \mathcal{F}$, we can observe that the dynamics of followers' costate variables λ_{jk} are dependent on the leaders' control strategies $\vec{c}_L^*$, and, hence, the initial value of the costate variables λ_{jk} is said to be controllable. Accordingly, we have another transversality condition $\vec{\mu}(0) = 0$ [131].

With the two point boundary value problem defined in (4.30), the steps for obtaining the candidate optimal strategies for both leader providers and follower providers are the same as that of the simultaneous play model. Also, the same sufficient conditions can be applied for proving optimality and therefore the obtained candidate optimal strategies of leaders and followers constitute the Stackelberg equilibrium. The implementation scheme for the open-loop Stackelberg equilibrium control is shown in table 4.1.

4.7 Performance Evaluation

4.7.1 Parameter Setting

We consider a cognitive radio network with two secondary providers (i.e., $M = 2$) providing service to 30 active secondary users (i.e., $N = 30$) in a particular service area. The average spectral efficiency is assumed to be 1.5 bits/s/Hz for both providers (i.e., $k_1 = k_2 = 1.5$). Fixed connection prices for two providers are set to be $p_1 = 0.3$ and $p_2 = 0.2$, respectively. For the replicator dynamics, we set the learning rate to be $\delta = 0.5$. The discount rate and cost factors for the objective functions of providers are set to be $\rho = 0.01$, and $\beta_1 = \beta_2 = 0.01$, respectively. The initial population share of providers 1 and 2 are assumed to be $x_1(0) = 0.3$, and $x_2(0) = 0.7$, respectively.

4.7.2 Numerical Results

(1) Dynamic behavior of secondary users and evolutionary equilibrium

We first study the dynamic behavior of secondary users by analyzing the direction field of the replicator dynamics as shown in Fig. 4.4. The secondary users can adapt their service selection strategies at each decision point which induces the variation of population share of all providers (i.e., population state). Evolutionary equilibrium will be reached when all secondary users receive equal utility with the average, and, hence, no user is willing to change its selection. This direction field shows the evolution directions to the equilibrium and the solution trajectories of the replicator dynamics under control. When the proportion of secondary users selecting certain provider is higher than the equilibrium proportion, users in this group receive less utility than the average, some users will churn to other providers for better performance. In this case, the evolution direction is represented by a line segment with negative slope. We observe that from any initial state, there exists a unique

trajectory to the equilibrium state. This observation complies with our analysis on the existence, uniqueness, and stability of the solution in Theorem 1.

Figure 4.4: Population distribution evolution of service provider 2 under different initial population shares.

(2) Optimal control strategies of the spectrum leasing differential game

For the dynamic competition among providers, we investigate both the open-loop Nash equilibrium (OLNE) strategy and the Markov perfect Nash equilibrium (MPNE) strategy for the simultaneous play game and also the open-loop Stackelberg equilibrium (OLSE) strategy for the Stackelberg game. For the Stackelberg game, providers 1 and 2 are assumed to be the leader and the follower, respectively. We consider both finite and infinite horizon.

Fig. 4.5 shows the finite horizon open-loop equilibrium strategies. With a large initial spectrum size leased but a small initial population share, provider 1 gradually reduces the leased spectrum resource. In contrast, provider 2 increases the leased spectrum size to accommodate more secondary users. Both providers reduce the leased spectrum size to zero at the boundary of the finite horizon (i.e., $T = 20$). This corresponds to the case that the provider stops service. Therefore, the optimal strategy for a provider is not to lease spectrum for this area once the service is terminated.

The optimal control strategies for providers 1 and 2 in infinite horizon modeling are shown in Fig. 4.6 and Fig. 4.7, respectively. For infinite horizon cases, the open-loop Nash equilibrium strategy, the open-loop Stackelberg equilibrium strategy, and the Markov perfect Nash equilibrium

Figure 4.5: Optimal control strategies of secondary service providers for finite horizon.

Figure 4.6: Optimal control strategies of secondary service provider 1 for infinite horizon.

strategy all converge to the corresponding steady state solution. Similar to the finite horizon model, provider 1 gradually reduces its leased spectrum and finally reaches a steady state solution. In contrast, provider 2 gradually increases its leased spectrum to the steady state solution.

(3) Nash equilibrium of the static game

For comparison purpose, we also formulate a static non-cooperative spectrum leasing game and

Figure 4.7: Optimal control strategies of secondary service provider 2 for infinite horizon.

Nash equilibrium (NE) is considered to be the solution. Fig. 4.8 shows the best response functions and the Nash equilibria for this static game with different cost. Note that derivation and analysis of static non-cooperative game and its Nash equilibrium are omitted for brevity.

Figure 4.8: Best response functions and Nash equilibrium of the static non-cooperative game.

(4) Population state evolution under different control strategies

The dynamics of population state and the evolution trajectory from initial state (i.e., $x(0) = 0.3$)

under different control strategies are shown in Fig. 4.9. Under all four control strategies (i.e., OLNE, OLSE, MPNE, and NE), the proportion of secondary users selecting provider 1 increases from the initial state to the equilibrium state. The increase is due to the higher utility for selecting provider 1 which leases a larger spectrum size. The open-loop Nash equilibrium strategy yields the same equilibrium population share with the Nash equilibrium strategy. However, the convergence of evolution path under the Nash equilibrium control is much slower than that under the open-loop Nash equilibrium control. With open-loop Stackelberg equilibrium strategy, though leasing less absolute amount of spectrum, the larger ratio c_1/c_2 makes the leader (i.e., provider 1) obtain a higher equilibrium population share than using open-loop Nash equilibrium strategy. Markov perfect Nash equilibrium control strategy yields the least equilibrium population for provider 1 due to the smallest ratio c_1/c_2 .

Figure 4.9: Population state evolution under different control strategies.

(5) Impact of price and initial population share

The impact of different prices of providers to the equilibrium population distribution and spectrum leasing amount can be analyzed through the analytical results. Accordingly, the impact of price to the profits can be analyzed. For service providers, increasing the price is a natural way to increase the profit. However, the decreased number of subscribed users may result in a lower cumulative profit. Such impact of price to the cumulative profit of provider is shown in Fig. 4.10.

Figure 4.10: The impact of price to the cumulative profit of provider.

Due to the asymptotical convergence property, the population state converges to the same equilibrium distribution from different initial distributions. Therefore, different initial population share will affect the instantaneous profit of providers before reaching the steady state, while it will not affect the instantaneous profit after reaching the steady state. In this way, different initial population share affects the cumulative profit. That is, the cumulative profit increases with the increase of the initial population share due to the larger weight of instantaneous profit of early period in the cumulative profit. This implies the importance for providers to attract initial users which can be verified by Fig. 4.6 where the provider with lower initial population share lease more spectrum to attract more users in the early period.

(6) Comparison

For the comparison of the four control strategies, we first investigate the yielded individual profit of providers 1 and 2 discounted and cumulated over infinite horizon as shown in Fig. 4.11. For both providers, all dynamic control strategies (i.e., OLNE, OLSE, MPNE) yield higher cumulative profit than that of the static control strategy (i.e., NE). The higher individual profit and the faster convergence speed to the equilibrium state give an important motivation to use the dynamic control rather than the static control. For the comparison within the three dynamic control strategies, it is always true that the Stackelberg equilibrium strategy can yield a higher (if not the same) profit for

the leader (i.e., provider 1) than that can be achieved by using open-loop Nash equilibrium strategy. The same argument does not apply to the follower. Also, though it is not a general case in differential games, here the Markov perfect Nash equilibrium strategy yields a higher profit than the open-loop Nash equilibrium for both providers.

Figure 4.11: Comparison of cumulative profit of secondary service providers.

The instantaneous profit of provider 1 is shown in Fig. 4.12 as an example. At the steady state, the Stackelberg equilibrium strategy yields the highest instantaneous total profit, following by Markov perfect Nash equilibrium strategy and open-loop Nash equilibrium strategy. The Nash equilibrium strategy yields the least instantaneous profit. We also investigate the impacts of using different control strategies to the mean utility and the mean transmission queueing delay of secondary users. For measuring the mean delay, the M/D/1 queueing model is used. Note that this M/D/1 model is used just for the mean delay performance evaluation, while the user arrival behavior is not considered in the system model. As shown in Fig. 4.13, the open-loop Nash equilibrium strategy provides the secondary users the highest mean utility and the least mean delay due to the largest spectrum size leased. In comparison, the Markov perfect Nash equilibrium strategy provides the secondary users the least mean utility and the highest mean delay.

(7) Summary

The properties of different dynamic control strategies for both finite horizon and infinite horizon

Figure 4.12: Comparison of instantaneous profit.

Figure 4.13: Comparison of utility and delay of secondary users under different control strategies.

cases are investigated. The convergence of the secondary users' population state evolution path under different dynamic controls of providers is shown. The numerical comparison shows the advantages of dynamic controls over static control in terms of both higher profit and faster convergence speed.

4.8 Conclusion

We have presented a two-level game framework based on differential game and evolutionary game for modeling the competitive dynamic spectrum leasing in a cognitive radio network. The evolution and dynamic service selection behavior of secondary users have been modeled as an evolutionary game and the strategy adaption process has been analyzed using the replicator dynamics. The dynamic competition among providers considering secondary users' service selection has been formulated as a non-cooperative differential game. The service selection distribution describes the state of the differential game. Both open-loop Nash equilibrium and Markov perfect Nash equilibrium are obtained which are applicable to different network structures. Besides the simultaneous play model, a Stackelberg differential game has been formulated to model the hierarchical play situation and Stackelberg equilibrium is considered to be the solution. Numerical studies have been performed to investigate the properties of solutions and to compare the impacts of using different control strategies to the profit of providers and utility, delay of secondary users.

Chapter 5

Dynamic Service Selection and Bandwidth Allocation in IEEE 802.16m Mobile Relay Networks

5.1 Introduction

IEEE 802.16m is proposed as an enhancement for the existing mobile WiMAX systems to meet the requirements of next-generation mobile broadband communication in the framework of IMT-Advanced [64]. The bandwidth can be divided to support both direct transmission service and relay-cooperative transmission service. Also, the advanced mobile stations (AMSs) can adapt the service selection decisions according to perceived performance. As discussed in Chapter 1, the service selection and bandwidth allocation are intertwined due to the cyclic dependency. To jointly address the issues of dynamic service selection and dynamic bandwidth allocation in IEEE 802.16m mobile relay networks, the proposed hierarchical (i.e., two-level) dynamic game framework based on evolutionary game and differential game is used to obtain the equilibrium strategies for both AMSs and advanced base stations (ABSs). At the lower level, the dynamic competition for resource access among AMSs is modeled as an evolutionary game since the AMSs can adapt their service selection strategies according to the received service quality (i.e., performance). The service selection strategy adaptation of this evolutionary game is subject to the control of ABSs in terms of allocated bandwidth, and is modeled by replicator dynamics [106]. Evolutionary equilibrium is considered

to be the solution of this lower level evolutionary game. At the upper level, the ABSs observe the service selection distribution of the AMSs and the queue states at the ABSs which jointly describe the system state, and then decide the optimal bandwidth allocation strategies dynamically. This dynamic bandwidth allocation is formulated as a differential game and a closed-loop Nash equilibrium is considered to be the solution. Since the rate of traffic input flows can fluctuate over time, the robust bandwidth allocation strategy design is conducted by viewing the rate fluctuation as a disturbance. In this case, both stochastic optimal control approach and H_∞ optimal control approach are applied for deriving the robust optimal strategies depending on different disturbance characteristics. Specifically, the stochastic optimal control is adopted for the average performance design when the statistics of the disturbance appear as a white noise Gaussian process, while the H_∞ optimal control is adopted for the worst-case performance design under the general finite energy disturbances.

The rest of this chapter is organized as follows: Section 5.2 presents the system model. The evolutionary game formulation for a strategy adaption of the AMSs is presented in Section 5.3. The competitive bandwidth allocation considering the dynamic service selection is formulated as a differential game in Section 5.4. Section 5.5 presents the numerical results. The conclusions are drawn in Section 5.6.

5.2 System Model

5.2.1 IEEE 802.16m Relay Network

We consider the following system of IEEE 802.16m relay network to address the problem of dynamic service selection jointly with bandwidth allocation. There is a service area a in an IEEE 802.16m mobile relay network with M competitive access providers and N advanced mobile stations (AMSs) as shown in Fig. 5.1. Without loss of generality, each access provider deploys one advanced base station (ABS), and hence, there are totally M ABSs providing service in this service area. Access provider also deploys advanced relay stations (ARSs) in this area to compete for providing an access service to AMSs. The current IEEE 802.16m specification only includes a subset of relay categories of IEEE 802.16j and only fixed ARS is considered [66]. Also, the ARSs in IEEE 802.16m are operated in a non-transparent mode and only two-hop communication is considered for relay-cooperation transmission. Both time-division duplex (TDD) and frequency-division duplex (FDD) modes are supported. We assume that the ARSs associated with the ABS

of the same provider are deployed densely and evenly in the area to make the access indifferent to the roaming AMSs. Two transmission modes (i.e., direct transmission denoted by ABS-AMS and relay-cooperation transmission denoted by ABS-ARS-AMS) can be used for AMSs. In this case, an AMS can choose to access the service either from an ABS directly or with the cooperation of an ARS. Also, the AMSs can select and churn to different access providers (i.e., ABSs) freely and dynamically. In the following, the two transmission modes are regarded as two service types provided by the ABS from the AMS perspective. Therefore, there are in total $2M$ choices for service selection of an AMS.

Figure 5.1: System model of a competitive IEEE 802.16m mobile relay network with M advanced base stations and N advanced mobile stations.

Consider downlink transmission, let C_i denote the total downlink bandwidth of ABS i . In IEEE 802.16m, the downlink (DL) subframe is divided into zones (i.e., access zone and relay zone), and the total downlink bandwidth of the ABS will be partitioned to support both ABS-AMS and ABS-ARS-AMS communications (Fig. 5.1). Let $\mu_{ij}(t)$ denote the allocated bandwidth of ABS i at time t for different transmission mode where $i \in \{1, 2, \dots, M\}$, $j = 1$ for direct and $j = 2$ for relay-cooperation transmissions. For the resource allocation among AMSs, we consider an unweighted round-robin scheme where all AMSs choosing the same service type of an ABS share the available bandwidth equally. Therefore, the bandwidth of an AMS allocated from ABS i is denoted by $b_{ij} = \mu_{ij}(t)/n_{ij}(t)$ where $n_{ij}(t)$ represents the number of AMSs choosing direct ($j = 1$) and relay-cooperation transmission ($j = 2$) of ABS i at time t .

The selection of ABS (i.e., access provider) as well as transmission mode of an AMS depends on the received performance which is dependent of the transmission rate and channel quality. Let τ_{ij} and k_{ij} denote the average downlink transmission rate and average spectral efficiency (in bits/s/Hz) for ABS i with transmission mode j , respectively. The spectral efficiency can be obtained based on the channel estimation. The average downlink transmission rate of an AMS in area a with direct communication with ABS i is obtained as $\tau_{i1} = b_{i1}k_{i1}$, where b_{i1} is the allocated bandwidth from ABS i and k_{i1} is the average spectral efficiency for direct communication. The average downlink transmission rate for an AMS with relay-cooperation communication with ABS i is obtained as $\tau_{i2} = b_{i2} \min\{k_{i2}, k_{\text{rm}}^i\}$, where b_{i2} is the allocated bandwidth from ABS i for relay-cooperation transmission, k_{i2} and k_{rm}^i (subscript rm stands for ARS-AMS) are the average spectral efficiency for ABS-ARS and ARS-AMS communication from ABS i , respectively. Basically, the rate of relay-cooperation transmission is the minimum rate between ABS-ARS and ARS-AMS [96]. For the performance satisfaction level of an AMS, utility function is used. Specifically, we consider a linear function $u_{ij} = \alpha b_{ij} k_{ij} (1 - p_{ij})$, where α is an adjustable constant to provide flexibility to fit the function with empirical data for different types of user applications, and p_{ij} is the average packet error rate (PER) for abstracting the channel quality. The estimation of PER can be found in [132].

5.2.2 Queue Dynamics

For the downlink data transmission, we assume that there is one queue associated with each service type of the ABS at a link layer (Fig. 5.1). Denote $q_{ij}(t)$ the queue length of aggregated traffic flow of service type j at ABS i at time t , where $j \in \{1, 2\}$ for direct and relay-cooperation transmissions, respectively. The vector of the queue states of all ABSs is then given by $\mathbf{q} = [q_{11}(t), q_{12}(t), \dots, q_{M2}(t)]^T$. A linear fluid model¹ is considered for modeling the queue dynamics, and the queue length evolves according to

$$\dot{q}_{ij}(t) = n_{ij}(t)r(t) - (1 - p_{ij})\mu_{ij}(t)k_{ij}, \quad (5.1)$$

for all $i \in \{1, 2, \dots, M\}, j \in \{1, 2\}$, where $\dot{q}_{ij}(t)$ is the derivative of $q_{ij}(t)$ with respect to t , $r(t)$ is the input traffic flow rate to the subscribed AMS², $n_{ij}(t)r(t)$ is the aggregate downlink flow rate at ABS i , and $\mu_{ij}(t)k_{ij}$ is the queue depletion rate (assigned bandwidth for draining the queue). The

¹This model is a continuous time version of the models used in [133].

²For simplicity, we assume one downlink flow for each AMS and the source rate characteristics are the same for all flows.

initial state $q_{ij}(0)$ represents the initial size of backlogged data of the queue. The input rate $r(t)$ can be fluctuated over time depending on the source behavior (e.g., variable bit rate (VBR) traffic) and can be viewed as the disturbance to the system. Specifically, we denote

$$r(t) = r_n + \xi(t), \quad (5.2)$$

where r_n is a nominal value of the input rate and $\xi(t)$ is a disturbance which can be either stochastic (e.g., white noise Gaussian process) or deterministic (e.g., impulse traffic load). This disturbance can occur due to the randomness of the packet arrival from the applications. It has been shown that the aggregated traffic (e.g., from users and applications) can be accurately modeled as the Gaussian process [134, 135]. From (5.1), we can see that both the bandwidth allocation strategies of ABSs and the service selection decisions of AMSs affect the queue states evolution. One objective of the dynamic bandwidth allocation is to maintain the queue length at the target level for satisfying system performance requirements.

5.3 Service Selection Evolutionary Game of Advanced Mobile Stations

We consider a large population of advanced mobile stations (AMSs) in service area a competing for the downlink bandwidth resources by selecting different service types (i.e., transmission modes) and different advanced base stations (ABSs). The objective of this selection is to maximize the individual satisfaction (i.e., utility) in terms of performance. In this section, first the evolutionary game formulation and service selection dynamics of AMSs are presented. The important properties (i.e., existence and uniqueness of the population state evolution) are discussed. Then, the equilibrium and evolutionary stability are analyzed.

5.3.1 Service Selection Dynamics

The players of a lower-level evolutionary game are the N AMSs in area a . The strategy of each player is the selection of ABS and the service type. Let the normalized variable $x_{ij}(t) = n_{ij}(t)/N$ denote the proportion of AMSs in area a choosing service type j from ABS i at time t , where

$i \in \{1, 2, \dots, M\}$ and $j \in \{1, 2\}$. The population state is denoted by a vector

$$\mathbf{x}(t) = [x_{11}(t), x_{12}(t), \dots, x_{M2}(t)]^T \in \mathbb{X}, \quad (5.3)$$

where $x_{ij}(t) \in [0, 1]$, $\sum_{i=1}^M \sum_{j=1}^2 x_{ij}(t) = 1$ for all $t \geq 0$, and $\mathbb{X}$ is the state space. The payoff (i.e., fitness) of AMS as a player is the utility that quantifies the performance satisfaction level according to the transmission rate and channel quality. Denote

$$\boldsymbol{\mu}(t) = [\mu_{11}(t), \mu_{12}(t), \dots, \mu_{ij}(t), \dots, \mu_{M1}(t), \mu_{M2}(t)]^T \quad (5.4)$$

the bandwidth allocation control vector. The expected payoff of an AMS selecting service type j of ABS i can be defined as $\pi_{ij}(\mathbf{x}, \boldsymbol{\mu}) = u(\tau_{ij}(\mathbf{x}, \boldsymbol{\mu}))$.³ The weighted average payoff of the population can be derived as $\bar{\pi}(\mathbf{x}, \boldsymbol{\mu}) = \sum_{i=1}^M \sum_{j=1}^2 c_{ij} x_{ij}(t) \pi_{ij}(\mathbf{x}, \boldsymbol{\mu})$, where c_{ij} is the weighting factor inverse proportional to the cost (e.g., power) for obtaining current service quality. For simplicity, we consider $c_{ij} = \beta/k_{ij}(1 - p_{ij})$, where β is a constant coefficient.

To model the strategy adaption process, replicator dynamics represented by a set of differential equations are used. In the replicator dynamics, the variation rate of the population share of a particular strategy is proportional to the difference between the payoff of that strategy and the average payoff of the population. The replicator dynamics [105, 106] for service selection are defined as follows:

$$\dot{x}_{ij}(t) = f_{ij}(\mathbf{x}, \boldsymbol{\mu}) = \delta x_{ij}(t) (\pi_{ij}(\mathbf{x}, \boldsymbol{\mu}) - \bar{\pi}(\mathbf{x}, \boldsymbol{\mu})), \quad (5.5)$$

with initial condition $\mathbf{x}(0) = \mathbf{x}_0 \in \mathbb{X}$, for all $i \in \{1, 2, \dots, M\}$, $j \in \{1, 2\}$, where $\dot{x}_{ij}(t)$ is the derivative of $x_{ij}(t)$ with respect to t , and δ is the learning rate of the population which controls the speed of change in service selection.

The constraints $\sum_{i=1}^M \sum_{j=1}^2 x_{ij}(t) = 1$ and $x_{ij}(t) \in [0, 1]$ can be guaranteed for all $t \in [0, \infty)$ under initial condition $\sum_{i=1}^M \sum_{j=1}^2 x_{ij}(0) = 1$. To verify this, we first substitute $\pi_{ij}(\mathbf{x}, \boldsymbol{\mu})$ and $\bar{\pi}(\mathbf{x}, \boldsymbol{\mu})$ into (5.5) which yields $f_{ij}(\mathbf{x}, \boldsymbol{\mu}) = \delta \alpha \left(\mu_{ij}(t) k_{ij}(1 - p_{ij}) - \beta x_{ij}(t) \sum_{i=1}^M C_i \right) / N$. Accordingly, we can obtain $\sum_{i=1}^M \sum_{j=1}^2 f_{ij}(\mathbf{x}, \boldsymbol{\mu}) = 0$ which indicates that the summation of all population shares does not vary with time. Also, we can obtain $f_{ij}(\mathbf{x}, \boldsymbol{\mu}) \geq 0$ with $x_{ij}(t) = 0$ and $f_{ij}(\mathbf{x}, \boldsymbol{\mu}) \leq 0$ with $x_{ij}(t) = 1$. Therefore, $x_{ij}(t) \in [0, 1]$ can be guaranteed.

³For the simplicity of presentation, the time parameter is omitted whenever no ambiguity occurs.

5.3.2 Existence and Uniqueness of Population State Evolution Trajectory under Bandwidth Allocation Control

We observe that the vector of bandwidth allocation strategies μ controls the evolution of the population state. The existence and uniqueness of the solution trajectory to the service selection dynamical system under bandwidth allocation control are discussed in the following theorem.

Theorem 3 *For the dynamical system $\dot{x}_{ij}(t) = f_{ij}(\mathbf{x}, \mu)$, $i \in \{1, 2, \dots, M\}$, $j \in \{1, 2\}$, with initial condition $\mathbf{x}(0) = \mathbf{x}_0$, if every element in μ is a measurable function on $[0, \infty)$, then there exists a unique solution trajectory $\mathbf{x}(t)$ defined for all $t \in [0, \infty)$.*

Proof: For fixed t , the partial derivative of $f_{ij}(\mathbf{x}, \mu)$ with respect to $x_{ij}(t)$ is continuous, and if $\mu_{ij}(t)$ is measurable on $[0, \infty)$, then $f_{ij}(\mathbf{x}, \mu)$ is measurable for fixed $x_{ij}(t)$ on the same interval. Moreover, for any given closed bounded set $K \in \mathbb{X}$ and interval $[a, b] \in [0, \infty)$, there always exists a positive number m ($m \geq |x_{ij}(t)|$ for all $\mathbf{x} \in K$) to construct an integrable function $g_{ij}(t) = |\delta\alpha\mu_{ij}(t)k_{ij}(1 - p_{ij})/N| + m \left| \delta\alpha\beta \sum_{l=1}^M C_l/N \right|$ such that $|f_{ij}(\mathbf{x}, \mu)| \leq g_{ij}(t)$ and $|\partial f_{ij}(\mathbf{x}, \mu)/\partial x_{ij}(t)| \leq g_{ij}(t)$ for all $(\mathbf{x}, t) \in K \times [a, b]$. Finally, $|f_{ij}(\mathbf{x}, \mu) - f_{ij}(\mathbf{y}, \mu)| = \delta\alpha\beta \sum_{l=1}^M C_l |x_{ij}(t) - y_{ij}(t)|/N$. Denote $L = \delta\alpha\beta \sum_{l=1}^M C_l/N + 1$, we can obtain $|f_{ij}(\mathbf{x}, \mu) - f_{ij}(\mathbf{y}, \mu)| \leq L|x_{ij}(t) - y_{ij}(t)|$ which implies that $f_{ij}(\mathbf{x}, \mu)$ satisfies the global Lipschitz condition. Based on the above arguments, we state that the solution to this dynamical system under control is unique and exists globally [129]. ■

5.3.3 Equilibrium and Evolutionary Stability Analysis

The replicator dynamics describe the variation rate of all population shares. The proportion of AMSs choosing a certain service type provided by an ABS increases if the corresponding payoff is above the average. Evolutionary equilibrium which is defined as the fixed point of replicator dynamics (i.e., $\dot{x}_{ij}(t) = 0, \forall i, j$) is considered to be the solution of this service selection evolutionary game. The evolutionary equilibrium can be obtained by solving algebraic equations $\dot{x}_{ij}(t) = 0, \forall i, j$.

At the equilibrium point, all AMSs in area a receive the same payoff and none of them has an incentive to change their strategy. Both boundary equilibrium (i.e., $\exists i, j, x_{ij} = 1$) and interior equilibrium (i.e., $x_{ij} \in (0, 1), \forall i, j$) exist. For the service selection evolutionary game, the boundary equilibrium is not stable in the sense that any small perturbation will make the system deviate from

the equilibrium state. In contrast, we will show below that the interior equilibrium is stable.

We first show that the service selection evolutionary game is a stable population game since the following condition is satisfied for all $\mathbf{x} \neq \mathbf{y}$:

$$\sum_{i=1}^M \sum_{j=1}^2 (x_{ij}(t) - y_{ij}(t))(\pi_{ij}(\mathbf{x}, \boldsymbol{\mu}) - \pi_{ij}(\mathbf{y}, \boldsymbol{\mu})) \leq 0. \quad (5.6)$$

Therefore, this evolutionary game possesses the property that the equilibrium is neutrally stable [136].

Further, we show that the equilibrium is asymptotically stable. We rewrite (5.5) in a vector form as $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{b}$, where

$$\mathbf{b} = [\delta\alpha\mu_{11}k_{11}(1 - p_{11})/N, \delta\alpha\mu_{12}k_{12}(1 - p_{12})/N, \dots, \delta\alpha C_M\mu_{M2}k_{M2}(1 - p_{M2})/N]^T, \quad (5.7)$$

and $\mathbf{A}$ is a $2M \times 2M$ dimensional matrix, the characteristic function of which can be expressed as follows:

$$\det(\gamma\mathbf{I} - \mathbf{A}) = \left(\gamma + \frac{\delta\alpha\beta}{N} \sum_{l=1}^M C_l \right)^{2M} = 0, \quad (5.8)$$

where $\mathbf{I}$ is an identity matrix. It is straightforward that $\mathbf{A}$ always has $2M$ repeated negative eigenvalues which verifies the asymptotically stable property of the service selection dynamics. Therefore, this dynamical system asymptotically converges to the evolutionary equilibrium point from any initial interior population state $\mathbf{x}_0$. According to [111], a strategy profile is evolutionary stable if and only if it is asymptotically stable. Therefore, the interior equilibrium is an evolutionary stable strategy (ESS) profile.

5.4 Dynamic Bandwidth Allocation in IEEE 802.16m Mobile Relay Networks

With the dynamic service selection behavior of AMSs and queue dynamics, the ABSs competitively allocate the downlink bandwidth for direct transmission and relay-cooperation transmission to maintain the desired queue level and user distribution for satisfying their own system performance requirements. Allocating small bandwidth for a relay-cooperation transmission will waste part of

the bandwidth of ARSs while allocating large bandwidth for a relay-cooperation transmission may result in performance degradation of direct transmission. Therefore, the bandwidth allocation needs to be performed optimally under competition. In this section, we formulate the dynamic bandwidth allocation in IEEE 802.16m mobile relay networks as a non-cooperative differential game. This is the upper-level game in the proposed two-level game framework which takes the dynamic service selection of AMSs and the queue state dynamics at the ABSs into account.

5.4.1 Dynamic Bandwidth Allocation as a Differential Game

We first formulate the competitive dynamic bandwidth allocation in IEEE 802.16m mobile relay networks as a differential game. In this case, each of the M ABSs competes to minimize its own cost function by controlling the bandwidth allocation strategy. The players are the M ABSs. The strategy of ABS i is the dynamic control of the downlink bandwidth allocated for the direct transmission and relay-cooperation transmission which can be denoted by $\boldsymbol{\mu}_i(t) = [\mu_{i1}(t), \mu_{i2}(t)]^T$. Similar to the common notation used in game theory, $\Phi = \{\boldsymbol{\mu}_i(t), \boldsymbol{\mu}_{-i}(t)\}$ denotes the strategy profile of this differential game, where $\boldsymbol{\mu}_{-i}(t)$ is a vector of strategies of all ABSs except ABS i .

In the bandwidth allocation differential game, the control strategy of each ABS influences the state evolution of the system. The population state $\mathbf{x}$ of the underlying service selection evolutionary game of AMSs and the queue state $\mathbf{q}$ of the ABSs jointly describe the system state of the upper-level differential game. With $n_{ij} = Nx_{ij}$ and $r = r_n + \xi$, the queue dynamics defined in (5.1) can be rewritten as follows:

$$\dot{q}_{ij} = Nx_{ij}r_n - (1 - p_{ij})\mu_{ij}k_{ij} + Nx_{ij}\xi, \quad (5.9)$$

for all $i \in \{1, 2, \dots, M\}$ and $j \in \{1, 2\}$. The replicator dynamics (5.5) and queue dynamics (5.9) describe how the current population state $\mathbf{x}$, queue state $\mathbf{q}$, bandwidth allocation strategies Φ , and the disturbance ξ influence the rate of change of the system state. Note that the multiplicative disturbance appearing in the queue dynamics indicates that the more subscribed users, the larger uncertainties introduced to the system by the disturbance.

The objective of the dynamic bandwidth allocation can be described from a system performance point of view. First, each ABS tries to maintain the queue length as close as possible to the target level for satisfying delay requirements. Second, each ABS tries to drive the AMS population state to the target distribution to avoid wasting bandwidth as well as overloading system. Third, each

ABS tries to prevent large disparities between aggregated traffic input flow rate and the allocated bandwidth. For measuring the performance, an instantaneous cost function is considered as follows:

$$J_{ins}^i(\boldsymbol{\mu}_i, \boldsymbol{\mu}_{-i}) = \sum_{j=1}^2 [w_{ij1} \tilde{q}_{ij}^2 + w_{ij2} \tilde{x}_{ij}^2 + w_{ij3} \tilde{u}_{ij}^2] e^{-\rho t}, \quad (5.10)$$

where $\tilde{q}_{ij} = (q_{ij} - Q_{ij}^{tar})$, $\tilde{x}_{ij} = (x_{ij} - N_{ij}^{tar}/N)$, and $\tilde{u}_{ij} = (Nx_{ij}r - (1 - p_{ij})\mu_{ij}k_{ij})$. Q_{ij}^{tar} and N_{ij}^{tar} are the target queue length and target number of AMSs, respectively. The first two terms in the cost function describe the penalties for deviating from the target queue length and target population distribution, respectively. The third term measures the disparities among the aggregated traffic input flow rate and the allocated system bandwidth. w_{ij1} , w_{ij2} , and w_{ij3} are positive weighting coefficients reflecting the relative importance of the requirements as well as the service priority. ρ is the discount factor influencing the present value of future cost.

For the formulated non-cooperative bandwidth allocation differential game, the closed-loop Nash equilibrium is a reasonable notion of optimality and is considered to be the solution. The definition of the closed-loop Nash equilibrium is given as follows:

Definition 1: Denote $J_i(\boldsymbol{\mu}_i, \boldsymbol{\mu}_{-i})$ the cost function of ABS i . The closed-loop strategy profile $\{\boldsymbol{\mu}_i^*, \boldsymbol{\mu}_{-i}^*\}$ is a closed-loop Nash equilibrium if for each ABS $i \in \{1, 2, \dots, M\}$, $J_i(\boldsymbol{\mu}_i^*, \boldsymbol{\mu}_{-i}^*) \leq J_i(\boldsymbol{\mu}_i, \boldsymbol{\mu}_{-i}^*)$ holds for all feasible control paths $\boldsymbol{\mu}_i$ in the bandwidth allocation differential game given other ABSs' control strategies $\boldsymbol{\mu}_{-i}^*$.

To obtain the closed-loop Nash equilibrium of this differential game, each ABS needs to solve an optimal control problem under disturbance. In this case, two approaches (i.e., stochastic optimal control and H_∞ robust optimal control) can be used depending on the characteristics of the disturbance ξ . If ξ enters the system dynamics in the form of white noise Gaussian process, the stochastic optimal control approach can be used to design a controller optimizing the average performance. If ξ is energy bounded, while the statistics are unknown, the H_∞ robust optimal control approach can be used for the worst-case performance control design.

5.4.2 Stochastic Optimal Control Approach

We first consider the stochastic optimal control approach with the assumption that the disturbance ξ is a white noise Gaussian process. In this case, the system state evolution is subject to a white noise

disturbance and described by stochastic differential equations. For each of M ABSs, the system performance optimization becomes a stochastic optimal control problem subject to the constraints (i.e., stochastic state evolution differential equations and physical constraints) given the control strategies of other ABSs. The stochastic optimal control problem can be expressed as follows:

$$\min_{\boldsymbol{\mu}_i} \lim_{T \rightarrow \infty} \frac{1}{T} E \left\{ \int_0^T J_{ins}^i(\boldsymbol{\mu}_i, \boldsymbol{\mu}_{-i}) dt \right\}, \quad (5.11)$$

$$s.t. \quad \begin{cases} \dot{x}_{ij} = \delta x_{ij} (\pi_{ij}(\mathbf{x}, \boldsymbol{\mu}) - \bar{\pi}(\mathbf{x}, \boldsymbol{\mu})), \\ \dot{q}_{ij} = Nx_{ij}r_n - (1 - p_{ij})\mu_{ij}k_{ij} + Nx_{ij}\xi, \end{cases} \quad (5.12)$$

for all $i \in \{1, 2, \dots, M\}, j \in \{1, 2\}$. Note that the state evolution cannot be fully determined given the current state and control path since the dynamics are stochastic. Therefore, the minimization is taken over the expectation value over all possible future realizations of the white noise process.

We first define $\tilde{\mathbf{y}}_i = e^{-\frac{1}{2}\rho t}[\tilde{x}_{i1}, \tilde{x}_{i2}, \tilde{q}_{i1}, \tilde{q}_{i2}]^T$ and $\tilde{\mathbf{u}}_i = e^{-\frac{1}{2}\rho t}[\tilde{u}_{i1}, \tilde{u}_{i2}]^T$ which serve as the new vectors of system state and control strategies, respectively. In this case, the system dynamics can be written in a vector form as follows:

$$d\tilde{\mathbf{y}}_i = [\tilde{\mathbf{A}}_i\tilde{\mathbf{y}}_i + \tilde{\mathbf{B}}_i\tilde{\mathbf{u}}_i + \tilde{\mathbf{b}}_i]dt + [\tilde{\mathbf{E}}_i\tilde{\mathbf{y}}_i + \tilde{\mathbf{d}}_i]d\xi, \quad (5.13)$$

where the coefficient matrices $\tilde{\mathbf{A}}_i$, $\tilde{\mathbf{B}}_i$, $\tilde{\mathbf{E}}_i$, $\tilde{\mathbf{b}}_i$, and $\tilde{\mathbf{d}}_i$ with suitable dimensions can be derived accordingly. The instantaneous cost function for provider i can be written as follows:

$$J_{ins}^i(\tilde{\mathbf{u}}_i, \tilde{\mathbf{y}}_i) = \tilde{\mathbf{y}}_i^T \tilde{\mathbf{Q}}_i \tilde{\mathbf{y}}_i + \tilde{\mathbf{u}}_i^T \tilde{\mathbf{R}}_i \tilde{\mathbf{u}}_i, \quad (5.14)$$

where $\tilde{\mathbf{Q}}_i$ and $\tilde{\mathbf{R}}_i$ are positive definite matrices which can be derived from (5.10). We observe that the cost function and state dynamics is independent of other ABSs, therefore, the optimization can be done separately for each ABS.

To obtain a closed-loop Nash equilibrium, Hamilton Jacobi Bellman (HJB) method [110] can be used. As common for autonomous dynamic games with infinite time horizon, the aim is to find the stationary Nash equilibrium. In this case, the optimal cost function and optimal control strategy are implicit of time t . Specifically, denote $V_i(\tilde{\mathbf{y}}_i)$ the optimal cost function for ABS i . In our case, the state dynamics are linear and the cost is quadratic. Therefore, the optimal cost function is also a

quadratic form denoted by

$$V_i(\tilde{\mathbf{y}}_i) = \frac{1}{2} \tilde{\mathbf{y}}_i^T \tilde{\mathbf{P}}_i \tilde{\mathbf{y}}_i + \tilde{\boldsymbol{\varphi}}_i^T \tilde{\mathbf{y}}_i + \tilde{\beta}_i, \quad (5.15)$$

where $\tilde{\mathbf{P}}_i$, $\tilde{\boldsymbol{\varphi}}_i$, and $\tilde{\beta}_i$ are coefficient matrices needed to be determined. And the optimal control strategy $\tilde{\mathbf{u}}_i$ can be expressed as follows:

$$\tilde{\mathbf{u}}_i^* = -\tilde{\boldsymbol{\Psi}}_i \tilde{\mathbf{y}}_i - \tilde{\boldsymbol{\psi}}_i, \quad (5.16)$$

where $\tilde{\boldsymbol{\Psi}}_i = \tilde{\mathbf{R}}_i^{-1} \tilde{\mathbf{B}}_i^T \tilde{\mathbf{P}}_i$ and $\tilde{\boldsymbol{\psi}}_i = \tilde{\mathbf{R}}_i^{-1} \tilde{\mathbf{B}}_i^T \tilde{\boldsymbol{\varphi}}_i$. $\tilde{\mathbf{P}}_i$, $\tilde{\boldsymbol{\varphi}}_i$ can be determined by solving the following algebraic equations (the derivation is omitted for brevity):

$$\begin{aligned} \tilde{\mathbf{P}}_i \tilde{\mathbf{A}}_i + \tilde{\mathbf{A}}_i^T \tilde{\mathbf{P}}_i + \tilde{\mathbf{E}}_i^T \tilde{\mathbf{P}}_i \tilde{\mathbf{E}}_i + \tilde{\mathbf{Q}}_i - (\tilde{\mathbf{B}}_i^T \tilde{\mathbf{P}}_i)^T \tilde{\mathbf{R}}_i^{-1} (\tilde{\mathbf{B}}_i^T \tilde{\mathbf{P}}_i) &= 0, \\ \left[\tilde{\mathbf{A}}_i - \tilde{\mathbf{B}}_i \tilde{\mathbf{R}}_i^{-1} \tilde{\mathbf{B}}_i^T \tilde{\mathbf{P}}_i \right]^T \tilde{\boldsymbol{\varphi}}_i + \tilde{\mathbf{E}}_i^T \tilde{\mathbf{P}}_i \tilde{\mathbf{d}}_i + \tilde{\mathbf{P}}_i \tilde{\mathbf{b}}_i &= 0, \\ \frac{1}{2} \left(\sqrt{\tilde{\mathbf{R}}_i} \tilde{\boldsymbol{\psi}}_i \right)^2 - \tilde{\boldsymbol{\varphi}}_i^T \tilde{\mathbf{b}}_i - \frac{1}{2} \tilde{\mathbf{d}}_i^T \tilde{\mathbf{P}}_i \tilde{\mathbf{d}}_i &= 0. \end{aligned}$$

and then $\tilde{\boldsymbol{\Psi}}_i$, and $\tilde{\boldsymbol{\psi}}_i$ can be obtained accordingly.

We have obtained the optimal value function as well as the optimal bandwidth allocation strategy $\tilde{\mathbf{u}}_i^*$ for ABS i . Accordingly, $\boldsymbol{\mu}_i^*$ can be obtained. Similarly, the optimal control strategies for all other ABSs can be obtained. The strategy profile $\Phi^* = \{\boldsymbol{\mu}_i^*, \boldsymbol{\mu}_{-i}^*\}$ constitutes the closed-loop Nash equilibrium for the bandwidth allocation differential game under white noise traffic input flow rate stochastic disturbance.

5.4.3 H_∞ Robust Optimal Control Approach

With stochastic optimal control approach, we have the assumption that the traffic input flow rate disturbance, though unpredictable, is not completely unknown. The statistics of the disturbance are known as a priori knowledge. For the general cases where the disturbance statistics cannot be adequately described, H_∞ robust optimal control as an alternative approach can be used to provide the worst-case performance guarantee which makes no assumption on the disturbance ξ other than it is constrained to have finite energy in terms of bounded L_2 norm (e.g., continuous deterministic damped traffic input flow rate disturbance). As one of the robust control techniques, H_∞ optimal control explicitly takes the system model (or the noise model) uncertainties into account. The

objective is to obtain a controller design that minimizes the performance cost function under the worst possible disturbances.

For the derivation of H_∞ robust optimal controller for the dynamic bandwidth allocation which minimizes the H_∞ norm of the system, we first regulate the system state dynamics and instantaneous objective function into standard form as follows: $\dot{\mathbf{y}}_i = \hat{\mathbf{A}}_i \mathbf{y}_i + \hat{\mathbf{B}}_i \hat{\mathbf{u}}_i + \hat{\mathbf{E}}_i \xi$, and $J_{ins}^i(\hat{\mathbf{u}}_i, \mathbf{y}_i) = \mathbf{y}_i^T \hat{\mathbf{Q}}_i \mathbf{y}_i + \hat{\mathbf{u}}_i^T \hat{\mathbf{R}}_i \hat{\mathbf{u}}_i$, where $\hat{\mathbf{u}}_i = \tilde{\mathbf{u}}_i$, $\mathbf{y}_i$, $\hat{\mathbf{A}}_i$, $\hat{\mathbf{B}}_i$, $\hat{\mathbf{C}}_i$, $\hat{\mathbf{Q}}_i$, and $\hat{\mathbf{R}}_i$ can be constructed accordingly. Since the state variable (e.g., queue length) can be internally obtained from the ABS, this system is classified as continuous-time with perfect state measurements.

In the following, the cost function as dynamic bandwidth allocation performance index in the context of H_∞ optimal control will be introduced. First, the controlled output $\mathbf{z}_i$ of ABS i is defined as follows: $\mathbf{z}_i = \hat{\mathbf{H}}_i \mathbf{y}_i + \hat{\mathbf{G}}_i \hat{\mathbf{u}}_i$, where matrices $\hat{\mathbf{H}}_i$ and $\hat{\mathbf{G}}_i$ represent the weights of cost for deviating from target state (e.g., queue length) and actual aggregated downlink traffic flow rate, respectively. No cross terms exist between the state $\mathbf{y}_i$ and control $\hat{\mathbf{u}}_i$ in the cost function of ABS i , i.e., $\hat{\mathbf{H}}_i^T \hat{\mathbf{G}}_i = \mathbf{0}$. Next, the L_2 norm of the controlled output $\mathbf{z}_i$ is defined as follows:

$$\|\mathbf{z}_i\|_2 = \left(\int_0^\infty \mathbf{z}_i^T \mathbf{z}_i dt \right)^{\frac{1}{2}}. \quad (5.17)$$

Accordingly, the H_∞ norm of system which reflects the worst-case L_2 gain of the system due to the exogenous signal ξ is defined as follows:

$$\|\mathbf{L}_i\|_\infty = \sup_{\xi} \frac{\|\mathbf{z}_i\|_2}{\|\xi\|_2}. \quad (5.18)$$

With the definition of the H_∞ norm, the robust optimal controller design problem for the dynamic bandwidth allocation of ABSs is expressed as a mathematical optimization problem which is concerned with finding a feasible state feedback controller $\hat{\mathbf{u}}_i^* = \hat{\mathbf{K}}_i \mathbf{y}_i$ such that the H_∞ norm of the system is minimized, i.e.,

$$\min_{\hat{\mathbf{u}}_i} \sup_{\xi} \frac{\|\mathbf{z}_i\|_2}{\|\xi\|_2}. \quad (5.19)$$

From (5.19), we can see that the H_∞ optimal control problem is in fact a minimax optimization problem. A natural way to solve this problem is to formulate a zero-sum sub-differential game in which the controller can be viewed as the minimizing player and the disturbance as the maximizing player [137]. The H_∞ norm cannot be used as the cost function directly for the sub-differential

game since the solution of which is intractable in general. Instead, as common in H_∞ optimal control theory, an augmented parameterized cost function is used and defined as follows: $J = \|\mathbf{z}_i\|^2 - \gamma^2 \|\xi\|^2$, where $\gamma > 0$ is a bound on the H_∞ norm defined as follows:

$$\|\mathbf{L}_i\|_\infty = \sup_{\xi} \frac{\|\mathbf{z}_i\|_2}{\|\xi\|_2} < \gamma. \quad (5.20)$$

Note that this augmented cost function is introduced purely for technical reasons. The scalar γ is a prescribed value interpreted as the performance bound which measures the ability of the system to mitigate the effect of the disturbance ξ on the output. A controller that satisfies this bound is called the suboptimal solution to the H_∞ optimal control problem. The performance of the controller relates to the value of γ . In particular, larger γ results in more conservative controller.

For the sub-differential game, the controller seeks to minimize the cost J while the disturbance viewed as another player seeks to maximize the cost J . The solution of this differential game is necessarily a saddle point of the cost function, which is defined by the following inequalities

$$J(\hat{\mathbf{u}}_i^*, \xi) \leq J(\hat{\mathbf{u}}_i^*, \xi^*) \leq J(\hat{\mathbf{u}}_i, \xi^*). \quad (5.21)$$

The suboptimal control can be obtained as $\hat{\mathbf{u}}_i^* = -\hat{\mathbf{B}}_i \hat{\mathbf{P}}_i \hat{\mathbf{y}}_i$, where $\hat{\mathbf{P}}_i$ is the minimal semi-definite solution of the following algebraic Riccati equation

$$\hat{\mathbf{P}}_i \hat{\mathbf{A}}_i + \hat{\mathbf{A}}_i^T \hat{\mathbf{P}}_i + \hat{\mathbf{H}}_i^T \hat{\mathbf{H}}_i - \hat{\mathbf{P}}_i (\hat{\mathbf{B}}_i \hat{\mathbf{B}}_i^T - \gamma^{-2} \hat{\mathbf{E}}_i \hat{\mathbf{E}}_i^T) \hat{\mathbf{P}}_i = 0.$$

The H_∞ optimal controller can be obtained with the minimal γ and it can be approximated numerically by decreasing the performance bound until the feasible solution of the above Riccati equation no longer exists.

With the obtained robust optimal bandwidth allocation strategies $\hat{\mathbf{u}}_i^*$ and accordingly μ_i^* for all ABSs, the strategy profile $\Phi^* = \{\mu_i^*, \mu_{-i}^*\}$ constitutes the closed-loop Nash equilibrium for the bandwidth allocation differential game under general energy bounded input rate disturbance. The detailed implementation schemes for the stochastic optimal controller and H_∞ optimal controller are provided in Table 5.1.

Note that the detailed design of a specific protocol for the game framework is out of the scope of this chapter. However, the information required for the implementation of algorithms can be

Table 5.1: Control strategy implementation schemes for bandwidth allocation.

Control strategies	Implementation schemes
Stochastic optimal controller	For each ABS i do , Measure the population distribution $\mathbf{x}$ and queue state $\mathbf{q}_i$; Update $\hat{\mathbf{u}}_i$ using (15); Update the bandwidth allocation strategy μ_i at each decision point; End for
H_∞ optimal controller	For each ABS i do , Measure the population distribution $\mathbf{x}$ and queue state $\mathbf{q}_i$; Update $\hat{\mathbf{u}}_i$ using $\hat{\mathbf{u}}_i^* = \hat{\mathbf{K}}_i \hat{\mathbf{y}}_i$; Update the bandwidth allocation strategy μ_i at each decision point; End for

analyzed. Specifically, each ABS needs to compute the average population payoff according to the user distribution and then distributes it to the subscribed AMSs. Besides, for the measurement of population distribution, each ABS needs to exchange the population distribution information with other ABSs.

5.5 Performance Evaluation

5.5.1 Parameters Setting

We consider a competitive IEEE 802.16m mobile relay network with two ABSs (i.e., $M = 2$) providing access service to 150 active AMSs (i.e., $N = 150$) in a particular service area. The total downlink bandwidth of the ABSs 1 and 2 for this area are $C_1 = 15$ MHz, and $C_2 = 10$ MHz, respectively. The average spectrum efficiency is assumed to be 1.2 bits/s/Hz. The nominal traffic input flow rate is assumed to be $r_n = 180$ kbps. We consider equal importance of different target requirements as well as service priority and we set $w_{ij1} = w_{ij2} = w_{ij3} = 1 \forall i, j$. The initial proportion of AMSs choosing two service types of two ABSs is assumed to be $x_{11}(0) = 0.2$, $x_{12}(0) = 0.3$, $x_{21}(0) = 0.1$, and $x_{22}(0) = 0.4$. The initial backlogged data are assumed to be $q_{11}(0) = 150\text{kb}$, $q_{12}(0) = 130\text{kb}$, $q_{21}(0) = 120\text{kb}$, and $q_{22}(0) = 100\text{kb}$. The target queue lengths are set to be $Q_{11}^{tar} = Q_{12}^{tar} = Q_{21}^{tar} = Q_{22}^{tar} = 40\text{kb}$. The target user distributions are set to be $N_{11}^{tar} = 60$, $N_{12}^{tar} = 30$, $N_{21}^{tar} = 40$, and $N_{22}^{tar} = 20$.

5.5.2 Numerical Results

The dynamic service selection behavior of AMSs subject to the bandwidth allocation control of ABSs under Gaussian white noise is investigated and the sample population share trajectories from the initial selection distribution are shown in Fig. 5.2. When the proportion of AMSs selecting a certain transmission mode of an ABS is higher than the equilibrium proportion, the AMSs in this group receive less utility than the average. As a result, some AMSs will switch to other strategies to achieve better performance. Fig. 5.2 shows that the state evolution paths converge to a certain population distribution at which all AMSs receive the same average utility of the population. To accommodate the traffic input flow rate fluctuation, the bandwidth allocated by closed-loop controls is dynamic and fluctuated (Fig. 5.3) which results in a fluctuated service selection distribution. In this case, the actual state evolution path fluctuates around the mean path but never remains at the steady state. Fig. 5.4 shows the empirical mean and 95% confidence interval of the population share evolution path x_{11} . The interval provides a precise bound for the fluctuated population share due to disturbance.

Figure 5.2: Population state of the service selection evolutionary game under stochastic optimal control.

We obtain the closed-loop Nash equilibrium control strategies of the bandwidth allocation differential game based on stochastic optimal control and H_∞ optimal control for optimizing average performance and worst-case performance, respectively. The performance of different control strategies for maintaining the queue length is evaluated in two cases. In Case I, the traffic input flow rate is random and follows Gaussian distribution $\mathcal{N}(0, 20^2)$. In Case II, energy bounded deterministic continuous disturbance and jitters (impulse traffic) occur on the traffic input flow rate. Specifically,

Figure 5.3: Service rate of different queues under stochastic optimal control.

Figure 5.4: Empirical mean and 95% confidence interval (CI) under stochastic optimal control.

we assume

$$R = \begin{cases} r_n + 50e^{-0.01t} + 50, & 200 < t < 400, 1000 < t < 1200 \\ r_n + 50e^{-0.01t}, & \text{otherwise.} \end{cases} \quad (5.22)$$

First, the queue dynamics under different control strategies of Case I is investigated. The queue

evolution of q_{11} is shown in Fig. 5.5 as an example. Both strategies control the queues from initial states to a level around target length. The stochastic optimal control approach performs better in terms of convergence speed since it is optimal in the sense that it minimizes the expected value of the performance. Also, due to continuously stochastic disturbance, the queue dynamics fluctuate around certain queue level.

Figure 5.5: Queue states of q_{11} under Gaussian noise.

Next, we vary the system parameters and investigate the impact on system performance. As shown in Fig. 5.6, the larger disturbance variance, the larger fluctuation to the queue dynamics is introduced. Also, with the increasing mean of the disturbance, the aggregated data rate increases and the network becomes overloaded. Therefore, the queue dynamics converge to a level above the target level.

The queue dynamics of q_{11} under different control strategies of Case II are investigated in Fig. 5.7. In this case, the disturbance on the traffic input flow rate is norm-bounded and the Gaussian white noise assumption is no longer valid. As shown in Fig. 5.7, the dynamic control strategies respond to the continuous disturbance and the traffic flow rate jitters. The queue length increases when the jitter occurs. In Case II, H_∞ optimal control approach performs better than the stochastic optimal control approach due to its robustness of the worst-case design.

5.6 Conclusion

In this chapter, we have developed a two-level dynamic game framework to jointly address the problem of dynamic service selection and dynamic bandwidth allocation in IEEE 802.16m mobile relay

Figure 5.6: Queue states of the q_{11} under stochastic disturbance with different mean and variance.

Figure 5.7: Queue states of q_{11} under non-Gaussian noise.

networks. The evolution and dynamic service selection behavior of advanced mobile stations have been modeled as an evolutionary game and the strategy adaption process has been analyzed using the replicator dynamics. The dynamic competition among providers considering mobile stations' service selection and queue states of the advanced base stations has been formulated as a differential game. The service selection distribution and the queue states jointly describe the state of the differential game. Both stochastic optimal control and H_∞ robust optimal control are used to obtain the closed-loop Nash equilibrium which is considered to be the solution of the differential game depending on the different disturbance types. Numerical studies have been performed to investigate the effectiveness of dynamic controls and to compare the system performance under different

control strategies.

Chapter 6

Conclusions and Future Work

In this chapter, we summarize the major research contributions and discuss the possible future research directions.

6.1 Major Research Contributions

This research aims to use dynamic games to model and investigate different issues (i.e., network selection, spectrum leasing, and bandwidth allocation) in wireless communication networks. Specifically, the major contributions of this research are summarized as follows:

- (1) A Bayesian evolutionary game model based on Bayesian best response dynamics is firstly used to investigate the dynamics of network selection with incomplete information in heterogeneous wireless networks. Specifically, the network selection in heterogeneous wireless networks has been formulated as the Bayesian evolutionary game to study the equilibrium decisions of the users under incomplete information about the preference (i.e., utility). The dynamics of this network selection game has been analyzed using Bayesian best response dynamics and aggregate best response dynamics.
- (2) A novel two-level dynamic game framework has been proposed based on evolutionary game and differential game with inherent dynamic nature. Both deterministic model and stochastic model could be incorporated. The proposed two-level framework has several favorable properties. First, the inherent dynamic nature makes this framework applicable for the com-

petitive dynamic decision making problems with hierarchical structure which are common in wireless networks. Second, since all lower-level agents receive the same individual utility at the equilibrium of the evolutionary game by using the replicator dynamics, the fairness can be guaranteed from the performance measure and the monetary cost perspectives. Third, comparing with static strategy, the higher efficiencies in terms of higher payoff and the faster convergence speed are the main advantages of using the dynamic strategies.

- (3) The dynamic spectrum leasing and service selection in spectrum secondary market of cognitive radio networks have been addressed using the proposed two-level dynamic game framework. Both open-loop Nash equilibrium and Markov perfect Nash equilibrium are obtained. Besides the simultaneous play model, a Stackelberg differential game has been formulated to model the hierarchical play situation and Stackelberg equilibrium is obtained as the solution.
- (4) The dynamic service selection and bandwidth allocation in IEEE 802.16m relay networks are addressed using the proposed two-level dynamic game framework considering stochastic disturbance. Depending on different disturbance types, both stochastic optimal control and H_∞ robust optimal control are used to obtain the closed-loop Nash equilibrium of the differential game. Both average performance and worst-case performance are analyzed.

6.2 Future Work

6.2.1 Future Work on Network Selection

In the current model, the user is assumed to access one network at a time instant. However, to improve the network utilization and also user's performance, the data can be transferred by multiple networks simultaneously. Two major issues arise with this simultaneous access in heterogeneous wireless network, i.e., multiplexing and rate allocation. Multiplexing scheme is used to distribute the traffic among different network accesses to balance the network utilization. The rate allocation scheme is used to decide the transmission rate on each network access. We will consider the design of rate allocation scheme in a distributed fashion. The distributed rate allocation scheme will be performed to meet performance objectives (e.g., maximum utilization of the available bandwidth) and prevent excessive rate fluctuations leading to instabilities.

In addition, in a wireless environment, the channel quality fluctuates due to fading and shadowing. Also, interference can affect the channel quality. Therefore, the actual network capacity, transmis-

sion delay, and packet loss rate can be time-varying which incur perturbations to the utilities of users. The impact of utility perturbations will also be considered in the design of network selection algorithm.

6.2.2 Future Work on Spectrum Leasing

Pricing is an important issue in a spectrum secondary market. In the current model, price is taken into account for the decision of a user to select the service provider. However, this price is assumed to be given and fixed. The different choices of the price will result in different equilibrium population states, different spectrum leasing amount, and therefore different profits for each provider. The optimal pricing strategy of provider in a competitive environment can be investigated. Note that the pricing issues in the spectrum secondary market can exist in multiple levels (e.g., pricing of spectrum holders, pricing of spectrum broker, and pricing of wireless service provider) which differ from the pricing issue (i.e., pricing of spectrum holders only) in a spectrum primary market. These pricing issues covering the whole spectrum secondary market can be jointly investigated.

6.2.3 Future Work on Bandwidth Allocation

We have studied the dynamic bandwidth allocation for the downlink transmission in IEEE 802.16m relay networks. For uplink transmission, due to the control signal delay in the uplink direction, a time-delayed bandwidth allocation framework needs to be studied. Also, in the current model, we have considered non-cooperative setting in which the players act selfishly to maximize individual payoffs. For the future work, the cooperation among different providers can be considered to maximize the social payoff. In this case, the fair and efficient allocation of social payoff among providers is important for motivating and maintaining the cooperation. Also, due to the time-varying wireless environment, some system parameters (e.g., channel condition and the number of users) vary from time to time. In this case, a non-autonomous system (i.e., a time-variant system the parameters of which depend on time) can be investigated.

Bibliography

- [1] M. J. Osborne, *An Introduction to Game Theory*, Oxford University Press, 2003.
- [2] D. Fudenberg and J. Tirole, *Game Theory*, The MIT Press, 1991.
- [3] M. Felegyhazi and J. P. Hubaux, *Game theory in wireless networks: A tutorial*, EPFL Technical report, 2007.
- [4] Z. Han, D. Niyato, W. Saad, T. Başar, and A. Hjørungnes, *Game Theory in Wireless and Communication Networks: Theory, Models, and Applications*, Cambridge University Press, 2011.
- [5] R. D. Yates, “A framework for uplink power control in cellular radio systems,” *IEEE Journal on Selected Areas in Communications*, vol. 13, no. 7, pp. 1341-1347, 1995.
- [6] D. Goodman and N. Mandayam, “Power control for wireless data,” *IEEE Journal on Selected Areas in Communications*, vol. 7, pp. 48-54, 2000.
- [7] Z. Han and K. J. R. Liu, *Resource Allocation for Wireless Networks: Basics, Techniques, and Applications*, Cambridge University Press, 2008.
- [8] D. Niyato and E. Hossain, “Dynamics of networks selection in heterogeneous wireless networks: An evolutionary game approach,” *IEEE Transactions on Vehicular Technology*, vol. 58, no. 4, pp. 2008-2017, May 2009.
- [9] K. Zhu, D. Niyato and P. Wang, “Network selection in heterogeneous wireless networks: Evolution with incomplete information,” in *Proc. of IEEE WCNC*, April 2010.
- [10] C.-J. Chang, T.-L. Tsai, and Y.-Han. Chen, “Utility and game-theory based network selection scheme in heterogeneous wireless networks,” in *Proc. of IEEE WCNC*, 2009.

- [11] D. Charilas, O. Markaki, and E. Tragos, "A theoretical scheme for applying game theory and network selection mechanisms in access admission control," in *Proc. of ISWPC*, 2008.
- [12] M. Cagalj, S. Ganeriwal, I. Aad, and J. P. Hubaux, "On selfish behavior in CSMA/CA networks," in *Proc. of IEEE INFOCOM*, 2005.
- [13] J. Konorski, "A game-theoretic study of CSMA/CA under a backoff attack," *IEEE/ACM Transactions on Networking*, vol. 14, no. 6, pp. 1167-1178, 2006.
- [14] P. Kyasanur and N. H. Vaidya, "Selfish MAC layer misbehavior in wireless networks," *IEEE Transactions on Mobile Computing*, vol. 4, no. 5, pp. 502-286, 2005.
- [15] Y. Xiao, X. Shan, and Y. Ren, "Game theory models for IEEE 802.11 DCF in wireless ad hoc networks," *IEEE Communications Magazine*, vol. 43, no. 3, pp. 22-26, 2005.
- [16] R. T. B. Ma, V. Misra, and D. Rubenstein, "An analysis of generalized slotted aloha with selfish users and perfect information," *IEEE/ACM Transactions on Networking*, vol. 17, no. 3, pp. 936-949, 2009.
- [17] H. Inaltekin and S. B. Wicker, "The analysis of Nash equilibrium of the one-shot random-access game for wireless networks and the behavior of selfish nodes," *IEEE/ACM Transactions on Networking*, vol. 16, no. 5, pp. 1094-1107, 2008.
- [18] Y. Kuo, C. Lu, E. H. K. Wu, and G. Chen, "An admission control strategy for differentiated services in IEEE 802.11," in *Proc. GLOBECOM*, 2003.
- [19] D. Niyato and E. Hossain, "Radio resource management games in wireless networks: An approach to bandwidth allocation and admission control for polling service in IEEE 802.16," *IEEE Wireless Communications Magazine*, vol. 14, no. 1, pp. 27-35, February 2007.
- [20] D. Niyato and E. Hossain, "QoS-aware bandwidth allocation and admission control in IEEE 802.16 broadband wireless access networks: a noncooperative game theoretic approach," *Elsevier Computer Networks*, vol. 51, no. 11, pp. 3305-3321, August 2007.
- [21] R. J. La and V. Anantharam, "Optimal routing control: Repeated game approach," *IEEE Transactions on Automatic Control*, vol. 47, no. 3, pp. 437-450, 2002.
- [22] V. Srivastava, J. Neel, A. B. Mackenzie, R. Menon, L. A. Dasilva, J. E. Hicks, J. H. Reed, and R. P. Gilles, "Using game theory to analyze wireless ad hoc networks," *IEEE Communications Surveys & Tutorials*, vol. 7, no. 4, pp. 46-56, February 2006.

- [23] S. Zhong, L. E. Li, Y. G. Liu, and Y. R. Yang, "On designing incentive-compatible routing and forwarding protocols in wireless ad-hoc networks: an integrated approach using game theoretic and cryptographic techniques," *Wireless Networks* vol. 13. no. 6, pp. 799-816, December 2007.
- [24] M. P. Anastasopoulos, P. D. M. Arapoglou, R. Kannan, and P. G. Cottis, "Adaptive Routing Strategies in IEEE 802.16 Multi-Hop Wireless Backhaul Networks Based On Evolutionary Game Theory", *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 7, pp. 1218-1225, 2008.
- [25] W. Saad, Z. Han, M. Debbah, A. Hjørungnes, and T. Başar, "Coalitional games for distributed collaborative spectrum sensing in cognitive-radio networks," in *Proc. IEEE INFOCOM*, 2009.
- [26] B. Wang, K. J. R. Liu, and T. C. Clancy, "Evolutionary cooperative spectrum sensing game: how to collaborate?," *IEEE Transactions on Communications*, vol. 58, no. 3, pp. 890-900, March 2010.
- [27] C. Song and Q. Zhang, "Achieving cooperative spectrum sensing in wireless cognitive radio networks," *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 13. no. 2, pp. 14-25, April 2009.
- [28] K. Zhu, D. Niyato, P. Wang, and Z. Han, "Dynamic spectrum leasing and service selection in spectrum secondary market of cognitive radio networks", *IEEE Transactions on Wireless Communications*, to appear.
- [29] L. Duan, J. Huang, and B. Shou, "Competition with dynamic spectrum leasing," in *Proc. IEEE DySPAN*, 2010.
- [30] D. Niyato and E. Hossain, "Competitive spectrum sharing in cognitive radio networks: a dynamic game approach," *IEEE Transactions on Wireless Communications*, vol. 7, no. 7, pp. 2651-2660, July 2008.
- [31] D. Niyato, E. Hossain, and Z. Han, "Dynamics of multiple-seller and multiple-buyer spectrum trading in cognitive radio networks: a game-theoretic modeling approach," *IEEE Transactions on Mobile Computing*, vol. 8, no. 8, pp. 1009-1022, August 2009.
- [32] O. Simeone, I. Stanojev, S. Savazzi, Y. Bar-Ness, U. Spagnolini, and R. Pickholtz, "Spec-

- trum leasing to cooperating secondary ad hoc networks,” *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 1, pp. 203-213, January 2008.
- [33] D. Niyato and E. Hossain, “Competitive pricing for spectrum sharing in cognitive radio networks: Dynamic game, inefficiency of Nash equilibrium, and collusion,” *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 1, pp. 192-202, January 2008.
- [34] D. Niyato and E. Hossain, “Market-equilibrium, competitive, and cooperative pricing for spectrum sharing in cognitive radio networks: Analysis and comparison,” *IEEE Transactions on Wireless Communications*, vol. 7, no. 11, pp. 4273-4283, November 2008.
- [35] J. Sun, E. Modiano, and L. Zheng, “Wireless channel allocation using an auction algorithm,” *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 5, pp. 1085-1096, May 2006.
- [36] Y. Xing, R. Chandramouli, and C. M. Cordeiro, “Price dynamics in competitive agile spectrum access markets,” *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 3, pp. 613-621, April 2007.
- [37] D. Niyato and E. Hossain, “Competitive pricing in heterogeneous wireless access networks: issues and approaches,” *IEEE Network*, vol. 22, no. 6, pp. 4-11, November 2008.
- [38] D. Niyato and E. Hossain, “A game theoretic analysis of service competition and pricing in heterogeneous wireless access networks,” *IEEE Transactions on Wireless Communications*, vol. 7, no. 12, pp. 5150-5155, December 2008.
- [39] P. Maillea and B. Tuffinb, “Price war in heterogeneous wireless networks,” *Computer Networks*, vol. 54, no. 13, pp. 2281-2292, September 2010.
- [40] L. Chen, S. H. Low, and J. C. Doyle, “Random access game and medium access control design,” *IEEE Transactions on Networking*, vol. 18, no. 4, pp. 1303-1316, August 2010.
- [41] T. Cui, L. Chen, and S. Low, “A game-theoretic framework for medium access control,” *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 7, pp. 1116-1127, August 2008.
- [42] V. S. Borkar and A. A. Kherani, “Random access in wireless ad hoc networks as a distributed game,” *Proc. WiOpt*, 2004.
- [43] L. Chen and J. Leneutre, “A game theoretic framework of distributed power and rate control in IEEE 802.11 WLANs,” *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 7, pp. 1128-1137, September 2008.

- [44] Y. Xing and R. Chandramouli, "Stochastic learning solution for distributed discrete power control game in wireless data networks," *IEEE Transactions on Networking*, vol. 16, no. 4, pp. 932-944, August 2008.
- [45] E. E. Tsiropoulou, G. K. Katsinis, and S. Papavassiliou, "Distributed uplink power control in multiservice wireless networks via a game theoretic approach with convex pricing," *IEEE Transactions on Parallel and Distributed Systems*, vol. 23, no. 1, pp. 61-68, January 2012.
- [46] I. F. Akyildiz, S. Mohanty, and J. Xie, "A ubiquitous mobile communication architecture for next-generation heterogeneous wireless systems," *IEEE Communications Magazine*, vol. 43, no. 6, pp. 29-36, June 2005.
- [47] D. Cavalcanti, D. Agrawal, C. Cordeiro, B. Xie, and A. Kumar, "Issues in integrating cellular networks WLANs, AND MANETs: a futuristic heterogeneous wireless network" *IEEE Wireless Communications*, vol. 12, no. 3, pp. 30-41, June 2005.
- [48] R. V. Nee, G. Awater, M. Morikura, H. Takanashi, M. Webster, and K. W. Halford, "New high-rate wireless LAN standards," *IEEE Communications Magazine*, vol. 37, no. 12, pp. 82-88, August 2002.
- [49] C. Smith, *3G Wireless Networks, Second Edition*, McGraw-Hill, 2007.
- [50] C. Eklund, R. B. Marks, K. L. Stanwood, AND S. Wang, "IEEE standard 802.16: a technical overview of the WirelessMANTM air interface for broadband wireless access," *IEEE Communications Magazine*, vol. 40, no. 6, pp. 98-107, June 2002.
- [51] E. Gustafsson and A. Jonsson, "Always best connected," *IEEE Wireless Communications*, vol. 10, no. 1, pp. 49-55, February 2003.
- [52] G. Fodor, A. Eriksson, and A. Tuoriniemi, "Providing quality of service in always best connected networks," *IEEE Communications Magazine*, vol. 41, no. 7, pp. 154-163, 2003.
- [53] B. Xie, A. Kumar, and D. P. Agrawal, "Enabling multiservice on 3g and beyond: challenges and future directions," *IEEE Communications Magazine*, vol. 15, no. 3, pp. 66-72, 2008.
- [54] I. F. Akyildiz, J. Xie, and S. Mohanty, "A survey of mobility management in next-generation all-ip-based wireless systems," *IEEE Wireless Communications Magazine*, vol. 11, no. 4, pp. 16-28, 2004.

- [55] J. Sachs and P. J. Gebert, "Multi-access management in heterogeneous networks, *Wireless Personal Communications*, vol. 48, no. 1, pp. 7-32, January 2009.
- [56] C. S. Fishburn, *Utility Theory for Decision Making*, Wiley, 1970.
- [57] F. Bari and V. C. M. Leung, "Automated network selection in a heterogeneous wireless network environment, *IEEE Network*, vol. 21, no. 1, pp. 34-40, January 2007.
- [58] E. K. P. Chong and S. H. Zak, *An Introduction to Optimization, Third Edition*, Wiley, 2008.
- [59] Q. Zhao and B. M. Sadler, "A survey of dynamic spectrum access," *IEEE Signal Process Magazine*, vol. 24, no. 3, pp. 79-89, May 2007.
- [60] J. M. Chapin and W. H. Lehr, "The path to market success for dynamic spectrum access technology," *IEEE Communications Magazine*, vol. 45, no. 5, pp. 96-103, May 2007.
- [61] H. Xu, J. Jin, and B. C. Li, "A secondary market for spectrum," in *Proc. IEEE INFOCOM*, 2010.
- [62] S. Sengupta and M. Chatterjee, "An economic framework for dynamic spectrum access and service pricing," *IEEE Transactions on Networking*, vol. 17, no. 4, pp. 1200-1213, August 2009.
- [63] S. Ulset, "Mobile virtual network operators: a strategic transaction cost analysis of preliminary experiences," *Telecommunications Policy*, vol. 26, no. 9, pp. 537-549, October 2002.
- [64] I. Papapanagiotou, D. Toumpakaris, J. Lee, and M. Devetsikiotis, "A survey on next generation mobile WiMAX networks: objectives, features and technical challenges," *IEEE Communications Surveys & Tutorials*, vol. 11, no. 4, pp. 3-18, December 2009.
- [65] Y. Yang, H. Hu, J. Xu, and G. Mao, "Relay technologies for WiMAX and LTE-advanced mobile systems," *IEEE Communications Magazine*, vol. 47, no. 10, pp. 100-105, October 2009.
- [66] K. Loa, C. C. Wu, S. T. Sheu, Y. Yuan, M. Chion, D. Huo, and L. Xu, "IMT-advanced relay standards," *IEEE Communications Magazine*, vol. 48, no. 8, pp. 40-48, August 2010.
- [67] E. Hossian, *Heterogeneous wireless access networks*, Springer, 2009.
- [68] Q. Y. Song and A. Jamalipour, "Network selection in an integrated wireless LAN and UMTS environment using mathematical modeling and computing techniques," *IEEE Wireless Communications*, vol. 12, no. 3, pp. 42-48, June 2005.

- [69] W. Shen and Q. A. Zeng, "Cost-function-based network selection strategy in integrated wireless and mobile networks," *IEEE Transactions on Vehicular Technology*, vol. 57, no. 6, pp. 3778-3788, November 2008.
- [70] D. Niyato and E. Hossain, "Modeling user churning behavior in wireless networks using evolutionary game theory," in *Proc. IEEE WCNC*, April 2008, pp. 2793-2797.
- [71] J. P. Singh, T. Alpcan, P. Agrawal, and V. Sharma, "An optimal flow assignment framework for heterogeneous network access," in *Proc. WoWMoM*, June 2007.
- [72] P. Steenkiste, D. Sicker, G. Minden, and D. Raychaudhuri, "Future directions in cognitive radio network research," *NSF Workshop Report*, March 2009.
- [73] FCC Spectrum Policy Task Force, "Report of the spectrum efficiency working group, November, 2002.
- [74] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 2, February 2005.
- [75] J. Mitola, G. Q. Maguire, "Cognitive radio: making software radios more personal," *IEEE Personal Communications*, vol. 6, no. 4, August 1999.
- [76] I. F. Akyildiz, W.-Y. Lee, M. C. Vuran, and S. Mohanty, "NeXt generation/dynamic spectrum access/cognitive radio wireless networks: A survey," *Computer Networks*, vol. 50, no. 13, September 2009.
- [77] W. Zhang, R. Mallik, and K. Letaief, "Optimization of cooperative spectrum sensing with energy detection in cognitive radio networks," *IEEE Transactions on Wireless Communications*, vol. 8, no. 12, pp. 5761-5766, December 2009.
- [78] I. F. Akyildiz, W.-Y. Lee, M. C. Vuran, and S. Mohanty, "A survey on spectrum management in cognitive radio networks," *IEEE Communications Magazine*, vol. 46, no. 4, pp. 40-48, April 2008.
- [79] X. Zhou, S. Gandhi, S. Suri, and H. Zheng, "eBay in the sky: strategy-proof wireless spectrum auctions," in *Proc. ACM MobiCom*, 2008.
- [80] S. Gandhi, C. Buragohain, L. Cao, H. Zheng, and S. Suri, "A general framework for wireless spectrum auctions," in *Proc. DySPAN*, 2007.

- [81] X. Zhou and H. Zheng, "Trust: a general framework for truthful double spectrum auctions," in *Proc. IEEE INFOCOM*, 2009.
- [82] L. Chen, S. Iellamo, M. Coupechoux, and P. Godlewski, "An auction framework for spectrum allocation with interference constraint in cognitive radio networks," in *Proc. IEEE INFOCOM*, 2010.
- [83] L. Duan, J. Huang, and B. Shou, "Cognitive mobile virtual network operator: investment and pricing with supply uncertainty," in *Proc. IEEE INFOCOM*, 2010.
- [84] J. Jia and Q. Zhang, "Competitions and dynamics of duopoly wireless service providers in dynamic spectrum market," in *Proc. ACM MobiHoc*, 2008.
- [85] J. Jia, Q. Zhang, Q. Zhang, and M. Liu, "Revenue Generation for Truthful Spectrum Auction in Dynamic Spectrum Access," in *Proc. ACM MobiHoc*, 2009.
- [86] IEEE P802.16d, "IEEE standard for local and metropolitan area networks, part 16: air interface for fixed broadband wireless access systems, October 2004.
- [87] IEEE Std 802.16e, "IEEE standard for local and metropolitan area networks - part 16: air interface for fixed and mobile broadband wireless access systems," February 2006.
- [88] IEEE Standard 802.16j-2009, "Air interface for broadband wireless access systems amendment 1: multiple relay specification, June 2009.
- [89] IEEE Std 802.16m-2011, "Air interface for broadband wireless access systems: advanced air interface," May 2011.
- [90] J. G. Andrews, A. Ghosh, and R. Muhamed, *Fundamentals of WiMAX*, Prentice-Hall, 2007.
- [91] A. Ghosh and D. R. Wolter, "Broadband wireless access with WiMax/802.16: current performance benchmarks and future potential," *IEEE Communications Magazine*, vol. 43, no. 2, pp. 129-136, February 2005.
- [92] B. Li, Y. Qin, and C. P. Low, "A survey on mobile WiMAX," *IEEE Communication Magazine*, vol. 45, no. 12, pp. 70-75, December 2007.
- [93] S. Ahmadi, "An overview of next-generation mobile WiMAX technology," *IEEE Communications Magazine*, vol. 47, no. 6, pp. 84-98, June 2009.

- [94] R. Y. Kim and J. S. Kwah, "WiMAX femtocell: requirements, challenges, and solutions," *IEEE Communications Magazine*, vol. 47, no. 9, pp. 84-91, September 2009.
- [95] Z. Tao, A. Li, K. H. Teo, J. Zhang, "Frame structure design for IEEE 802.16j mobile multihop relay (mmr) networks," in *Proc. IEEE GLOBECOM*, 2007.
- [96] D. Niyato, E. Hossain, D. I. Kim, and Z. Han, "Relay-centric radio resource management and networking planning in IEEE 802.16j mobile multihop relay networks," *IEEE Transactions on Wireless Communications*, vol. 8, no. 12, pp. 6115-6125, December 2009.
- [97] B. Lin, P. H. Ho, L. L. Xie, X. Shen, and J. Tapolcai, "Optimal relay station placement in broadband wireless access networks," *IEEE Transactions on Mobile Computing*, vol. 9, no. 2, pp. 259-269, February 2010.
- [98] A. S. Ibrahim, A. K. Sadek, W. Su, K. J. R. Liu, "Cooperative communications with relay-selection: when to cooperate and whom to cooperate with?," *IEEE Transactions on Wireless Communications*, vol. 7, no. 7, pp. 2814-2827, July 2008.
- [99] B. Wang, Z. Han, K. J. R. Liu, "Distributed relay selection and power control for multiuser cooperative communication networks using stackelberg game," *IEEE Transactions on Mobile Computing*, vol. 8, no. 7, pp. 975-990, July 2009.
- [100] C. Y. Hong and A. C. Pang, "3-Approximation algorithm for joint routing and link scheduling in wireless relay networks" *IEEE Transactions on Wireless Communications*, vol. 8, no. 2, pp. 856-861, February 2009.
- [101] Y. Ge, S. Wen, Y. H. Ang, and Y. C. Liang, "Optimal relay selection in IEEE 802.16j multihop relay vehicular networks," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 5, pp. 2198-2206, June 2010.
- [102] R. Fei, K. Yang, and S. Ou, "A QoS-aware dynamic bandwidth allocation algorithm for relay stations in IEEE 802.16j-based vehicular networks," in *Proc. IEEE WCNC*, 2010.
- [103] Q. Albluwi, N. A. Ali, and H. Hassanein, "A dynamic frame partitioning scheme for IEEE 802.16 mesh and multihop relay networks," in *Proc. IEEE ICC*, 2009.
- [104] I. K. Chan and W. Liao, "Adaptive bandwidth allocation for TCP traffic in IEEE 802.16j wireless networks with transparent relay stations," in *Proc. IEEE PIMRC*, 2008.
- [105] J. Weibull, *Evolutionary Game Theory*. M.I.T. Press, 1995.

- [106] J. Hofbauer and K. Sigmund, "Evolutionary game dynamics," *Bull. Amer. Math. Soc.*, vol. 40, no. 4, pp. 479-519, 2003.
- [107] P. D. Taylor and L. B. Jonker, "Evolutionary stable strategies and game dynamics," *Mathematical Biosciences*, vol. 40, no. 1, pp. 145-156, July 1978.
- [108] T. Başar and G. J. Olsder, *Dynamic Noncooperative Game Theory*. Society for Industrial and Applied Mathematics, January 1999.
- [109] K. Ogata, *Modern Control Engineering, Fifth Edition*, Pearson Custom Publishing, 2009.
- [110] D. P. Bertsekas, *Dynamic Programming and Optimal Control*, Athena Scientific, November 2000.
- [111] M. P. Anastasopoulos, P.-D. M. Arapoglou, R. Kannan, and P. G. Cottis, "Adaptive routing strategies in IEEE 802.16 multi-hop wireless backhaul networks based on evolutionary game theory," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 7, pp. 1218-1225, September 2008.
- [112] S. Shakkottai, E. Altman, and A. Kumar, "Multihoming of users to access points in WLANs: a population game perspective," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 6, pp. 1207-1215, August 2007.
- [113] E. Altman and Y. Hayel, "A stochastic evolutionary game of energy management in a distributed aloha network," in *Proc. IEEE INFOCOM*, 2008.
- [114] E. H. Watanable, D. S. Menasché, E. S. Silva, and R. M. M. Leão, "Modeling resource sharing dynamics of VoIP users over a WLAN using a game-theoretic approach," in *Proc. IEEE INFOCOM*, 2008.
- [115] P. Coucheney, C. Touati, and B. Gaujal, "Fair and efficient user-network association algorithm for multi-technology wireless networks," in *Proc. IEEE INFOCOM*, 2009.
- [116] E. Altman, R. Elazouzi, Y. Hayel, and H. Tembine, "An evolutionary game approach for the design of congestion control protocols in wireless networks," in *Proc. IEEE WiOPT*, 2008.
- [117] K. Zhu, D. Niyato, and P. Wang, "Optimal bandwidth allocation with dynamic service selection in heterogeneous wireless networks," in *Proc. IEEE GLOBECOM*, 2010.
- [118] R. Secchi, P. Barsocchi, and F. Davoli, "Linear quadratic control of service rate allocation in a satellite network," *IET Communications*, vol. 4, no. 13, pp. 1580-1593, September 2010.

- [119] T. Alpcan, J. P. Singh, and T. Başar, "Robust rate control for heterogeneous network access in multihomed environments," *IEEE Transactions on Mobile Computing*, vol. 8, no. 1, pp. 41-51, January 2009.
- [120] F. Abdollahi and K. Khorasani, "A novel H_∞ control strategy for design of a robust dynamic routing algorithm in traffic networks," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 4, pp. 706-718, May 2008.
- [121] B. K. Lee, Y. H. Chen, and B. S. Chen, "Robust H_∞ power control for CDMA cellular communication systems," *IEEE Transactions on Signal Processing*, vol. 54, no. 10, pp. 3947-3956, October 2006.
- [122] J. C. Ely and W. H. Sandholm, "Evolution in Bayesian games I: Theory," *Games and Economic Behavior*, vol. 53, no. 1, pp. 83-109, Oct. 2005.
- [123] J. Choi, J. Yang, C. Kim and S. Choi, "EBA: An enhancement of the IEEE 802.11 DCF via distributed reservation," *IEEE Transactions on Mobile Computing*, vol. 4, no. 4, pp. 378-390, July 2005.
- [124] A. Matsui, "Best response dynamics and socially stable strategies," *Journal of Economic Theory*, vol. 57, no. 2, pp. 343-362, August 1992.
- [125] Z. Jiang, Y. Ge, and Li. Y, "Max-utility wireless resource management for best-effort traffic," *IEEE Transactions on Wireless Communications*, vol. 4, no. 1, pp. 100-111, Jan. 2005.
- [126] A. Odlyzko, "Paris metro pricing for the internet," in *Proc. ACM EC*, 1999.
- [127] S. Shakkottai, R. Srikant, A. Ozdaglar, and D. Acemoglu, "The price of smiplicity ," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 7 , pp. 1269 - 1276, September 2008.
- [128] A. J. Goldsmith and S. G. Chua, "Variable rate variable power MQAM for fading channels," *IEEE Transactions on Communications*, vol. 45, no. 10, pp. 1218-1230, October 1997.
- [129] J. C. Engwerda, *LQ Dynamic Optimization and Differential Games*, Wiley, June 2005.
- [130] K. Ogata, *Modern Control Engineering*. Prentice Hall, November 2001.
- [131] E. J. Dockner, S. Jørgensen, N. V. Long, and G. Sorger, *Differential Games in Economics and Management Science*, Cambridge Univ. Press, November 2000.

- [132] Q. Liu, S. Zhou, and G. B. Giannakis, "Cross-Layer combining of adaptive modulation and coding with truncated ARQ over wireless links," *IEEE Transactions on Wireless Communications*, vol. 3, no. 5, pp. 1746-1755, September 2004.
- [133] G. Boggia, P. Camarda, L. A. Grieco, and S. Mascolo, "Feedback-based control for providing real-time services with the 802.11e MAC," *IEEE Transactions on Networking*, vol. 15, no. 2, pp. 323-333, April 2007.
- [134] E. W. Knightly and N. B. Shroff, "Admission control for statistical QoS: Theory and practice," *IEEE Network*, vol. 13, no. 2, pp. 20-29, 1999.
- [135] D.-Y. Eun and N. B. Shroff, "A measurement-analytic approach for QoS estimation in a network based on the dominant time scale," *IEEE/ACM Transactions on Networking*, vol. 11, pp. 222-235, April 2003.
- [136] H. Tembine, E. Altman, R. ElAzouzi, and Y. Hayel, "Stable networking games," in *Proc. 46th Annual Allerton Conference on Communication, Control, and Computing*, 2008.
- [137] T. Başar, P. Bernhard, *H_∞ Optimal Control and Related Minmax Design Problems*, Birkhäuser Boston, October, 1995.

Appendix A

List of Abbreviations

Abbreviation	Description
2G	Second Generation Mobile Telecommunication System
3G	Third Generation Mobile Telecommunication System
4G	Fourth Generation Wireless System
ABC	Always Best Connected
ABS	Advanced Base Station
AC	Access Category
AMS	Advanced Mobile Station
ARS	Advanced Relay Station
BE	Best Effort
BS	Base Station
DL	DownLink
DSA	Dynamic Spectrum Access
ESS	Evolutionary Stable Strategy
FCC	Federal Communications Commission
FDD	Frequency Division Duplex

Abbreviation	Description
GRA	Grey Relational Analysis
HJB	Hamilton Jacobi Bellman
MAC	Media Access Control
MADM	Multi Attribute Decision Making
MBWA	Mobile Broadband Wireless Access
MDP	Markov Decision Process
MPNE	Markov Perfect Nash Equilibrium
MVNO	Mobile Virtual Network Operator
MS	Mobile Station
NE	Nash Equilibrium
OFDM	Orthogonal Frequency Division Multiplexing
OLNE	Open Loop Nash Equilibrium
OLSE	Open Loop Stackelberg Equilibrium
PMP	Paris Metro Pricing
QoS	Quality of Service
RAT	Radio Access Technology
SINR	Signal to Interference plus Noise Ratio
TDD	Time Division Duplex
TPBVP	Two Point Boundary Value Problem
UL	UpLink
WLAN	Wireless Local Area Network
WiMAX	Worldwide Interoperability for Microwave Access
WMAN	Wireless Metropolitan Area Networks

Appendix B

List of Symbols

Symbol	Description
T	Time interval.
a	Service area.
N	The number of users in service area a .
$\mathcal{C}$	The set of candidate access networks.
K	The number of candidate access network.
P_k	Fixed connection fee per connection per unit of time for network k .
τ_i^k	Bandwidth of user i received from network k .
B_k	The available bandwidth of network k .
N_k	The total number of users choosing network k in service area a .
Δ	The set of probability distribution over actions.
b_i	The minimum bandwidth requirement which is the type of user i .
Γ	The type space.
$f(b_i)$	probability density function of type.
s_i	The strategy of player i in the Bayesian game.
s_{-i}	The vector of strategies of all users except user i .
Ω	The set of all Bayesian network selection game strategies.

Symbol	Description
π_i^k	The instantaneous payoff of user i selecting network k .
$\mathcal{U}(\tau_i^k)$	Rate utility.
$\mathcal{J}_k(\mathbf{s}_{-i})$	The expected number of users choosing network k containing all possible type combinations.
$\Phi_i^k(\mathbf{s}_{-i}, b_i)$	The probability of satisfying the minimum bandwidth requirement of user i by choosing network k given all other users' strategies $\mathbf{s}_{-i}$.
H_i	The handover cost.
$\bar{\pi}_i(s_i, \mathbf{s}_{-i})$	The expected payoff of user i given strategy profile $\{s_i, \mathbf{s}_{-i}\}$.
$\mathcal{R}(\mathbf{s}_{-i})$	The best response of user i given other users' strategies $\mathbf{s}_{-i}$.
$\mathcal{E}(s)$	The aggregate network selection distribution induced by Bayesian strategy $s \in \Omega$.
$\mathcal{B}(\mathbf{x})$	The best response correspondence to the social aggregate network selection distribution.
$c_i(t)$	Spectrum size that provider i leases from spectrum brokers at time t .
$\mathbf{c}_{-i}$	The strategy of all providers except i .
$c_i^*(t)$	Optimal control strategy of provider i .
M	The number of providers.
F_i	The maximum spectrum size that provider i can lease.
k_i	The average spectral efficiency in bits/s/Hz.
γ_i	The signal to interference plus noise ratio (SINR).
$\bar{h}_i$	The average channel gain.
$\tau_i(t)$	The bandwidth of a secondary user allocated from provider i at time t .
$n_i(t)$	Total number of secondary users choosing provider i at time t .
p_i	Charged price by provider i .
$u_i(t)$	The utility of secondary user i at time t .
α	Coefficient of rate utility.
$x_i(t)$	Population share of provider i at time t .
$\mathbf{x}$	Population distribution.
$\mathbf{x}^*$	Equilibrium population distribution.
$\mathbb{X}$	State space.

Symbol	Description
$\pi_t(t)$	The expected payoff of a secondary user selecting provider i at time t .
$\bar{\pi}(t)$	Average payoff of the population.
δ	Learning rate of the population.
$\mathbf{x}_0$	Initial population distribution.
$\mathbf{A}$	$M \times M$ dimensional coefficient matrix.
$\mathbf{I}$	Identity matrix.
$\mathcal{C}_i^{OL}$	The open-loop strategy spaces.
$\mathcal{C}_i^{CL}$	The closed-loop strategy spaces.
$J^i(c_i, \vec{c}_{-i})$	The instantaneous profit of provider i choosing control strategy c_i .
β_i	Weighted cost factor.
ρ	Discount rate influencing the discounted present value of future profit.
$H_i(\vec{x}, c_i, \vec{c}_{-i}, \vec{\lambda}_i, t)$	The present value Hamiltonian function of provider i .
λ_{ij}	The costate variable of provider i associated with state x_j .
$H_i^*(\vec{x}, \vec{\lambda}_i, t)$	Maximized Hamiltonian function.
Φ^*	Open-loop Nash equilibrium.
$V^i(\vec{x}, t)$	The optimal value function of provider i .
$V_t^i(\vec{x}, t)$	Derivative of optimal value with respect to t .
$V_{x_i}^i(\vec{x}, t)$	Derivative of optimal value with respect to x_i .
φ^*	The Markov perfect Nash equilibrium.
$\mathcal{L}$	The set of leader providers.
$\mathcal{F}$	The set of follower providers.
H_i^L	The Hamiltonian functions for leader provider i .
H_j^F	The Hamiltonian functions for follower provider j .
$\mu_{ij}(t)$	The total allocated bandwidth of ABS i at time t for transmission mode j .
$b_{ij}(t)$	The bandwidth of an AMS allocated from ABS i .
$n_{ij}(t)$	The number of AMSs choosing ABS i and transmission mode j .
$u_{ij}(t)$	Utility of AMSs choosing ABS i and transmission mode j .

Symbol	Description
p_{ij}	Average packet error rate (PER).
$q_{ij}(t)$	Queue length of aggregated traffic flow of service type j at ABS i at time t .
$\mathbf{q}$	The vector of the queue states of all ABSs.
$r(t)$	The input traffic flow rate to the subscribed AMS.
r_n	A nominal value of the input rate.
$\xi(t)$	Disturbance.
$x_{ij}(t)$	Population distribution of transmission mode j of ABS i .
$\boldsymbol{\mu}(t)$	The bandwidth allocation control vector.
$\pi_{ij}(\mathbf{x}, \boldsymbol{\mu})$	The expected payoff of an AMS selecting service type j of ABS i .
Φ	The strategy profile of the differential game.
Φ^*	The closed-loop Nash equilibrium.
$\boldsymbol{\mu}_i(t)$	Strategy of ABS i .
$\boldsymbol{\mu}_{-i}(t)$	Vector of strategies of all ABSs except ABS i .
Q_{ij}^{tar}	Target queue length.
N_{ij}^{tar}	Target number of AMSs.

Appendix C

Publication List

1. **K. Zhu**, D. Niyato, P. Wang, and Z. Han, “Dynamic spectrum leasing and service selection in spectrum secondary market of cognitive radio networks,” *IEEE Transactions on Wireless Communications*, accepted.
2. **K. Zhu**, D. Niyato, and P. Wang, “Dynamic service selection and bandwidth allocation in IEEE 802.16m mobile relay networks,” *IEEE Journal on Selected Areas in Communications* special issue on Cooperative Networking Challenges and Applications, accepted.
3. **K. Zhu**, D. Niyato, P. Wang, E. Hossain, and D. I. Kim, “Mobility and handoff management in vehicular networks: A survey,” *Wireless Communications and Mobile Computing*, vol. 11, no. 4, pp. 459-476, April 2011.
4. **K. Zhu**, D. Niyato, and P. Wang, “Dynamic bandwidth allocation under uncertainty in cognitive radio networks,” to be presented in *IEEE GLOBECOM*, Houston, TX USA, 5-9 December 2011.
5. D. Niyato **K. Zhu**, and P. Wang, “Cooperative virtual machine management for multi-organization cloud computing environment,” in *Proceedings of International Workshop on Game Theory in Communication Networks (Gamecomm)*, Ecole Normale Suprieure de Cachan, Paris, France, 16 May 2011.
6. D. Niyato, **K. Zhu**, and A. V. Vasilakos, “Resource and revenue sharing with coalition formation of cloud providers: Game theoretic approach,” in *Proceedings of IEEE International Symposium on Cluster Computing and the Grid (CCGrid)*, Newport Beach, CA, USA, 23-26 May 2011.

7. **K. Zhu**, D. Niyato, and P. Wang, "Optimal bandwidth allocation with dynamic service selection in heterogeneous wireless networks," in Proceedings of IEEE GLOBECOM'10, Miami FL USA, 6-10 December 2010.
8. **K. Zhu**, D. Niyato and P. Wang, "Network selection in heterogeneous wireless networks: Evolution with incomplete information," in Proceedings of IEEE WCNC, Sydney, Australia, 18-21 April 2010.