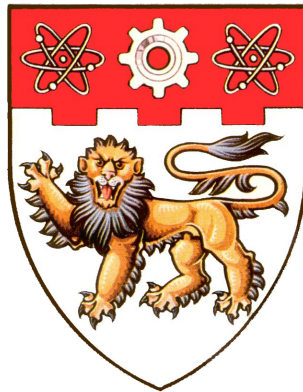


Game-Theoretic Modeling for Resource Allocation in Relay-Based Wireless Networks



Li Yifan

School of Computer Engineering

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

2013

Statement of Originality

I hereby certify that all the work embodied in this thesis is the result of original researchment done by myself and has not been submitted for a higher degree to any other University or Institute.

.....

Date

.....

Li Yifan

*To Dad and Mom,
for their unconditional love and encouragements.*

Abstract

Game theory, originally invented to investigate the complex behavior in economics, has lent itself to a plethora of areas such as politics, biology, sociology and computer science nowadays. With its effectiveness in modeling the complicated behavior among different players, it is envisioned that it will play a significant role in the field of wireless communications and networking. In this dissertation, the author focuses on investigating some challenging resource allocation problems in relay-based wireless networks, a special type of wireless networks in presence of relays. To properly model the mathematical relationship between different network components and achieve a stable state where all components are satisfied with the outcome, different game-theoretic models are tailored to cope with different network scenarios. The specific issues the author addresses in this dissertation are summed up as follows.

Firstly, the problem of price competition between different service providers (as owners of the relays) in wireless relay networks in presence of mobile users is considered. The payoff of a service provider is defined to be a function of both the relay price and the frequency that the relays are selected. To tackle the problem, the author designs a hierarchical framework via applying a Stackelberg-game model, where dynamic relay selection for mobile users and price competition for service providers are jointly taken into consideration. At the lower level, the author investigates the relay selection problem for the mobile users under given relay prices. It is formulated as a Markov decision process (MDP) problem with the objective to minimize the mobile user's long-term average cost, and to be solved by applying the linear programming (LP) technique. At the upper level, the author studies the game of setting relay prices for the service providers, with the knowledge that the mobile users will make relay

selections based on their given prices. Nash equilibrium is obtained as the solution. The obtained results can help to provide a guidance for service providers to compete for providing relay services.

Secondly, the author investigates the problem of power allocation for secondary users (SUs) in cognitive relay networks, where the SUs are involved as cooperative relays in a primary user's (PU) communication. Opportunistic channel access scheme is considered, which motivates the SUs to use some relay power to speed up the PU's transmissions. Consequently, the PU's buffer will be depleted faster, resulting in more transmission opportunities for the SUs. Nonetheless, due to the energy limitation, less power can be used for an SU's own transmissions if too much power is consumed on relaying. To address this tradeoff, the author designs power allocation strategies for both single-SU and multiple-SU cases. The former is formulated as a utility maximization problem, while the latter is modeled as a non-cooperative game. The existence and uniqueness of the Nash equilibrium (NE) are proved, indicating that a stable network state can be achieved where no SU can benefit more by unilaterally changing its power allocation strategy when the other SUs keep their strategies unaltered.

Thirdly, the author studies the problem of cooperation formation for downlink data transmission in mobile infostation networks, where content distribution is achieved by exploiting the opportunistic contact among mobile users and infostations. Focusing on such a network scenario, the author discovers that the cooperation among infostations and mobile users can improve the network performance in term of the expected delay of content distribution. More interestingly, jointly taking both cooperation among infostations and that among mobile users into account may achieve even better improvement in performance, which inspires the author to propose a hierarchical cooperation formation model to analyze the bi-level cooperation. Coalitional formation game model and network formation game model are properly applied to obtain the stable

structure of the bi-level cooperation, and an implementable distributed algorithm is proposed. Through extensive numerical experiments, the proposed algorithm is shown to be highly effective in obtaining the stable hierarchical cooperation structure.

Finally, the problem of content delivery in relay-based publish-subscribe (pub-sub) networks is discussed. For the sake of optimizing the content delivery process from the content provider (CP) to the subscribers via relays in between, the CP plays a vital role in the strategy design. Modeling the pub-sub network using a tree-structured topology, tandem queueing model and absorbing Markov chain model can be applied to derive the quality-of-service (QoS) received by the subscribers (in term of the delivery ratio within deadline). It is observed that the content delivery problem has two major characteristics: (1) the CP always has the capability to control the content delivery process and is pursuing the maximum profit by strategically designing the “rights” and “obligation” items for the subscribers; (2) the subscribers are also self-interested so that they will attempt to achieve the highest benefits of themselves, however, their actions are constrained by the possible choices provided by the CP. Based on the above problem characteristics, contract theory is adopted to reach an economically optimal solution. The numerical results verify the effectiveness of the contract-theoretic approach in maximizing the CP’s profit, and the capability to ensure the satisfaction of the heterogeneous subscribers.

Acknowledgements

First and foremost, I would like to express my deepest gratitude and respect to my supervisors, Prof. Wang Ping and Prof. Dusit Niyato, for their invaluable guidance and suggestions, and also for their constant patience, kindness, continuous encouragement and support during my whole PhD candidature. Prof. Wang Ping offers me the precious opportunity to do research here in NTU, her critical thoughts always help me to improve my work, and she always encourages me when having a hard time, her kindness and generosity always teaches me a better way to live my life. Prof. Dusit Niyato is also a great supervisor whom I feel lucky to work with, his wisdom and rich knowledge, his invaluable guidance and advices, have always offered great inspirations and help on my work. I enjoy working with them so much, and having the opportunity to learn to do research in their group is my greatest honor.

Also, I would like to thank Ms. Chua Poo Hua, Ms. Cindy and Ms. Siom Siew Ling for their kindness to us and diligence in maintaining the Center for Multimedia and Network Technology (CeMNet) an ideal place for doing research. It is my great pleasure to work in CeMNet, which is much more than a lab.

I would also express my great appreciation to all friends I met here, who brought me laugh and tears, making my life full of sunshine and happiness. I specifically thank to our research group members, Vamsi, Zhu Kun, Shimin, Wang Ran, Zhang Yang, Qiumin, Sudarshan, Hoang and John, who have always been nice and helpful for both my work and life. I also feel so lucky that I can be together with so many good friends, including my roommate Zuo Zhen, my friends Shou Wei, Zhixiang, Yuanyuan, Yiqun, Wang Yan, Li Xiang, Ren Zhou, Wenjie, Yang Ming, Haifeng, Xiang Liu, Fang Shuo, Yuming, Yang Huan, Shanshan, Wenbo, Wenjin, Xu Xiang, Sun Qian, Lijun,

Zhongyuan, Guan Xing, etc. I thank them all for all the endless encouragement and selfless support.

Last but not least, I would like to thank my parents, my grandmother and my cousins for the endless love and everlasting support. My family always provides me with courage to fight against difficulties and encourages me to face with everything from an optimistic perspective, without fear about loneliness and the unknown future. Thank you so much for always being with me together.

Contents

Abstract	i
Acknowledgements	iv
List of Figures	x
List of Abbreviations	xii
1 Introduction	1
1.1 Background and Research Objective	1
1.2 Motivation and Scope	4
1.2.1 Dynamic Relay Selection for Mobile Users and Price Competition among the Service Providers in Wireless Relay Networks	5
1.2.2 Power Allocation for Secondary Users in Cognitive Relay Networks	7
1.2.3 Cooperation Formation in Mobile Infostation Networks	9
1.2.4 Content Delivery in Relay-Based Publish-Subscribe Networks	12
1.3 Major Contributions	14
1.3.1 Dynamic Relay Selection for Mobile Users and Price Competition among the Service Providers in Wireless Relay Networks	15
1.3.2 Power Allocation for Secondary Users in Cognitive Relay Networks	16
1.3.3 Cooperation Formation in Mobile Infostation Networks	17
1.3.4 Content Delivery in Relay-Based Publish-Subscribe Networks	18
1.4 Dissertation Organization	20
2 Literature Review	22
2.1 Game-Theoretic Tools	22

2.2	Relay Selection Issue	29
2.3	Price Competition among Wireless Service Providers	32
2.4	Cooperative Relaying in Cognitive Relay Networks	34
2.5	Content Delivery in Mobile Infostation Networks	36
2.6	Content Delivery in Publish-Subscribe Mobile Social Networks	39
2.6.1	Delay Tolerant Networks	39
2.6.2	Mobile Social Networks	40
2.6.3	Publish-Subscribe Systems	41
2.6.4	Contract Theory Application	42
3	Dynamic Relay Selection for Mobile Users and Price Competition for Service Providers in Wireless Relay Networks	43
3.1	System Model	44
3.1.1	Networking Scenario	44
3.1.2	Mobility Model	46
3.1.3	Relay Transmission	47
3.2	Hierarchical Framework of Dynamic Relay Selection for Mobile Users and Price Competition for Service Providers	50
3.2.1	Dynamic Relay Selection	51
3.2.2	Game Formulation for Pricing	57
3.2.3	Architecture of the Proposed Algorithm	59
3.3	Performance Evaluation	61
3.3.1	Scenario and Parameter Settings	61
3.3.2	Performance Evaluation of the Relay Selection	62
3.3.3	Numerical Results of the Relay Pricing	67
3.4	Summary	71
4	Optimal Power Allocation for Secondary Users in Cognitive Relay Networks	72
4.1	System Model	73
4.1.1	Network Scenario	73
4.2	Optimal Power Allocation Scheme	74
4.2.1	Utility Function Design	74

4.2.2	Problem Formulation of Single-PU-Single-SU case	76
4.2.3	Problem Formulation of Single-PU-Multiple-SU Case	79
4.2.4	Architecture of the Proposed Algorithm	83
4.3	Numerical Results	83
4.3.1	Single-PU-Single-SU Case	83
4.3.2	Single-PU-Two-SU Case	86
4.4	Summary	90
5	Cooperation Formation for Downlink Data Transmission in Mobile Infostation Networks	92
5.1	System Model	93
5.2	Hierarchical Cooperation Formation	97
5.2.1	Cooperation among the Mobile Users	98
5.2.2	Cooperation among the Infostations	101
5.2.3	Stackelberg Equilibrium and Proposed Distributed Algorithm .	103
5.3	Performance Evaluation	108
5.4	Summary	114
6	Content Delivery in Relay-Based Publish-Subscribe Networks	116
6.1	System Model	117
6.1.1	Relay-Based Publish-Subscribe Network	117
6.1.2	Queueing Model	120
6.2	Content Provider/ Subscriber Model and Optimal Contract Design . .	129
6.2.1	Content Provider Side	129
6.2.2	Subscriber Side	132
6.2.3	Optimal Contract Design	133
6.2.4	Architecture of the Proposed Algorithm for Contract Design . .	140
6.3	Performance Evaluation	140
6.3.1	Performance Metrics Obtained from Queueing Model	140
6.3.2	Optimal Contract Design for the Content Provider	148
6.4	Summary	152
7	Conclusions and Future Work	154

7.1	Summary of Contributions	154
7.2	Future Research Directions	157
7.2.1	Future Work on Dynamic Relay Selection	157
7.2.2	Future Work on Power Allocation in Cognitive Relay Networks .	158
7.2.3	Future Work on Network Formation	158
7.2.4	Future Work on Contract Design	159
	Author's Publications	161
	Bibliography	163

List of Figures

3.1	The wireless relay network.	45
3.2	The proposed hierarchical framework of the Stackelberg-fashion problem.	51
3.3	Flowchart of the proposed algorithm to solve the hierarchical framework.	60
3.4	The effect of weight variation.	63
3.5	The number of selected RSs under different queue length.	65
3.6	Probability of selecting different RSs under different maximum queue size in both scenarios 1 and 2.	65
3.7	The demand of RSs 1 and 2 under different price of RS 1.	66
3.8	Comparison of the average user cost.	67
3.9	Illustration of service provider 2's best response strategy, given $p_1 = 9.5$	69
4.1	The cognitive relay network.	74
4.2	Flowchart of the proposed algorithm to solve the Power Allocation Problem.	84
4.3	Optimal power allocation strategy variation with weight in utility function.	85
4.4	Interpretation of the best-response strategy.	87
4.5	Investigation of different impact factors on NE.	89
5.1	An illustrative example scenario of mobile infostation network.	94
5.2	Flowchart of the proposed distributed hierarchical cooperation formation algorithm.	107
5.3	Stationary probability of the ISs' coalitional structures vs. the unit cost for IS cooperation	110
5.4	Comparison of the ISs' payoffs under different coalitional structures.	112
5.5	Stable MU cooperation network formation result.	113

5.6	Comparison of the MUs' payoffs with/without forming cooperation network.	114
6.1	An example scenario of relay-based publish-subscribe network.	118
6.2	The corresponding tree structure of relay-based publish-subscribe network.	119
6.3	Queue structure.	122
6.4	State-transition-rate diagram of one-dimensional Markovian queue with bulk service.	122
6.5	State-transition-rate diagram of 2D Markovian queue with bulk services I.	126
6.6	State-transition-rate Diagram of 2D Markovian queue with bulk services II.	127
6.7	Flowchart of the proposed algorithm to implement the Contract Design.	141
6.8	CDF of content delay distribution for different queue service rates μ in Case I.	142
6.9	Comparison of simulation result and analytical result when $\mu = 0.6$ (1000 content arrivals).	142
6.10	Comparison of simulation result and analytical result when $\mu = 0.6$ (10000 content arrivals).	143
6.11	Comparison of simulation result and analytical result when $\mu = 1$ (1000 content arrivals).	143
6.12	Comparison of simulation result and analytical result when $\mu = 1$ (10000 content arrivals).	144
6.13	Comparison of simulation result and analytical result of DRWD (1000 content arrivals).	144
6.14	Comparison of simulation result and analytical result of DRWD (10000 content arrivals).	145
6.15	CDF of content delay distribution for different bandwidth allocation r_1 in Case I.	146
6.16	Comparison of simulation result and analytical result when $r_1 = 2$	146
6.17	Comparison of simulation result and analytical result when $r_1 = 3$	147
6.18	DRWD vs. bandwidth allocation r_1 in Case I.	147
6.19	DRWD vs. queue service rates μ_1 and μ_2 in Case II.	149
6.20	Utility of content provider vs. queue service quality selection in Case I.	151
6.21	Utility of content provider vs. r_1 and r_2 for type- θ_1 subscriber in group \mathbb{G}_R^3 in Case II.	153

List of Abbreviations

ABBREVIATIONS

FULL EXPRESSIONS

SP	Service Provider
MIMO	Multiple-Input-Multiple-Output
AF	Amplify-and-Forward
DF	Decode-and-Forward
CC	Coded Cooperation
WRN	Wireless Relay Network
QoS	Quality-of-Service
PoA	price of anarchy
FCC	Federal Communications Commission
SINR	Signal-to-Interference-and-Noise Ratio
BS	Base Station
RS	Relay Station
MRC	Maximal Ratio Combining
SD	Source-to-Destination
CDF	Cumulative Distribution Function
AMC	Adaptive Modulation and Coding
MAR	Maximum Achievable Rate
QPSK	Quadrature Phase Shift Keying
QAM	Quadrature Amplitude Modulation
EER	Effective End-to-End Rate
MDP	Markov Decision Process
LP	Linear Programming
NE	Nash Equilibrium
PU	Primary User
SU	Secondary User
CRN	Cognitive Relay Network
PT	Primary Transmitter
PR	Primary Receiver
ST	Secondary Transmitter
SR	Secondary Receiver
TDMA	Time-Division-Multiple-Access

NPG	Non-cooperative Power Allocation Game
3G	Third Generation
IS	Infostation
MU	Mobile User
NTU	Non-Transferrable-Utility
MSN	Mobile Social Network
IoT	Internet of Things
CP	Content Provider
DRWD	Delivery Ratio within Deadline
IC	Incentive compatibility
IR	Individual Rationality
CTMC	Continuous-Time Markov Chain
pdf	probability density function
IR	Individual Rationality

Chapter 1

Introduction

1.1 Background and Research Objective

From the first milestone of the evolution of game theory, the publication of the book “Theory of Games and Economic Behavior” written by the mathematicians John von Neumann and Oskar Morgenstern in 1944 [1], and then contributed by John Nash’s seminal work on non-cooperative games and bargaining, game theory has considerably broadened its area far beyond the confines of mathematics. From the 1970s, the modern methodological framework in game theory has been widely applied to diverse areas including economics, politics, biology, psychology, computer science, etc.

Game theory is the formal study of interactions between two or more decision-makers with possibly conflicting or common interests [2], [3]. The decision makings of different individuals or agents are inter-related, and the decision makers are generally aware that their decisions have impact on the others’ and are affected by others’ at the same time. All such circumstances are considered as *games* in game theory, while the decision makers participating in the game are referred to as *players*, and the actions

they take in the game are referred to as *strategies*. Games are always defined in terms of their rules, describing how a single player's action impacts on payoff of itself, as well as the payoffs of the other players [4]. Generally, the rule of a game includes information about the identities of the players, their knowledge about the game, the possible moves they may take, and their payoffs can be achieved from playing the game. Payoff of a single player measures how well he does in the outcome of the game, which can be measured by monetary profit or satisfaction, depending on the game he plays. It is assumed that any player with rationality prefers higher payoff to lower one, thus he will always pursue the highest payoff due to selfishness so long as there are no controllers existed to guide his action with some compulsive power. The striking similarity between interactions among different components in wireless networks and game-theoretic models inspires the research objective of the author: attempting to apply proper game models to study non-cooperative or cooperative behavior of different network components in typical wireless network scenarios, seeking to achieving optimal resource allocation in such networks.

With rapid growth of telecommunication technologies, wireless networking, pushed by the dramatically increasing demand of the widespread mobile applications, has become ubiquitous nowadays. It is envisioned that future wireless networks will further revolutionize our daily lives, where a plethora of services such as high-speed Internet access, video conferencing, interactive media, automated freeways and smart homes will emerge and become popular. New development also raises new challenges, and one of the most critical issues to implement the future realm of wireless is the design for effective and efficient algorithms to be implanted into distributed wireless nodes and to achieve an optimal network resource allocation or a stable network operation. The wireless nodes (e.g., base stations, mobile phones, laptops, servers, etc.) can be made "intelligent" in this way, and can thus be considered as players in terminology

of game theory. As game theory can provide a formal framework to investigate the complicated behavior among multiple players with conflict or common interests, the plentiful mathematical tools included in the framework can be applied to model and analyze modern communication systems. For instance, to study the competition for radio resources among a group of wireless users, or the price competition among several wireless service providers, non-cooperative game can be utilized to analyze the non-cooperative behavior of the wireless users and wireless service providers [5]. While for problems such as how to fairly allocate transmission rates to a number of users accessing the same wireless channel [6], or how to spontaneously form a network among multiple nodes to facilitate routing traffic among themselves [7], cooperative game theory can be leveraged to search for the effective solutions.

Moreover, various solution concepts from game theory can be utilized in different network scenarios, where the network components can have different relationships in nature. For instance, Nash equilibrium solutions are highly appropriate for competitive scenarios, where the competitive players can arrive at a stable network state characterized by no one having incentive to deviate. Stackelberg equilibrium solutions are specifically designed for solving Stackelberg games where natural hierarchy exists among the players [5]. Correlated equilibrium [8], which can be viewed as a generalization of the Nash equilibrium, targets on situations where an arbitrator exists and is able to help the players to correlate their actions by sending public or private signals to them. While Wardrop equilibrium [9], which appeared originally from the context of road traffic, can act as a solution concept to capture the key characteristics of resource sharing among multiple individuals with selfish goals.

Impressed by the plentiful and powerful mathematical tools provided by game theory for modeling various practical problems, and inspired by the striking similarity between some wireless networking scenarios and game-theoretic models, the author aims

at making a targeted investigation into some typical wireless network scenarios, studying the competitive or cooperative behavior of the network components, and searching for solutions to achieve optimal resource allocation or stable network operation state. In order to solve different problems, the author attempts to exploit appropriate models from the rich landscape of game theory, which fit well into the considered network scenario and are capable of deriving the satisfied solutions.

1.2 Motivation and Scope

Within this dissertation, the author focuses on wireless networks in presence of relays. The relays can be cooperative relay stations involved in cooperative communications to improve the spatial diversity of data transmission, or mobile relays which can facilitate data transmission between nodes in a carry-and-forward manner, playing significant roles in delay-tolerant networks. In the following, the author mainly addresses four different problems: (1) in a wireless relay network (WRN), how will the service providers (SPs) strategically set prices for their relay services to gain the highest profit, where the demands for the relay services depend on the proposed dynamic relay selection (DRS) scheme; (2) in a cognitive relay network (with the secondary users (SUs) acting as relays), which is a special case of WRN, how should the SUs allocate their power to relaying for the primary user (PU) and to self-transmission, in order to achieve the best tradeoff between transmission time and transmission power; (3) in mobile infostation networks, where content distribution is achieved by exploiting the opportunistic contact among mobile users (MUs) and infostations (ISs), how to improve the network performance by allowing cooperation formation among both the ISs and the MUs, and (4) in publish-subscribe (pub-sub) networks where the content provider has full bargaining power over the subscribers, how to reach an economically optimal solution for

content delivery, for the sake of maximizing the CP's profit and meanwhile ensuring the satisfaction of the heterogeneous subscribers. Considering the different characteristics of the above mentioned problems, the author will choose appropriate mathematical tools to tackle them.

1.2.1 Dynamic Relay Selection for Mobile Users and Price Competition among the Service Providers in Wireless Relay Networks

It has been widely acknowledged that spatial diversity is of primary importance to combat channel fading in wireless communications [10]. As one of the most well-known spatial diversity schemes, multiple-input-multiple-output (MIMO) has been included in recent wireless system standards [11]. However, it may not be practical to apply MIMO in some wireless scenarios, especially for a wireless device where multiple antennas cannot be deployed due to size, cost or other hardware limitations [12]. To overcome this restriction, cooperative communication is proposed as a viable solution, in which some nodes can act as relays to help with the transmission from a source to a destination. In cooperative communication, a virtual MIMO can be formed for users with a single antenna, and thus the cooperative diversity can be achieved.

A challenging issue in cooperative communications is how to strategically select relays for the realization of cooperative diversity. Most research efforts have been devoted to investigating relay selection problems in a static scenario. In this work, the author takes user mobility into account, and thus DRS schemes are required to be designed for mobile users. For instance, when mobile users travel among different locations in a WRN, depending on locations, different relays can be selected to help with mobile users' transmission. Additionally, since the relays spend energy on relaying

the users' data, voluntary relaying as in [13] may not be valid in reality. Therefore, a natural solution is to make payment to the relay nodes, which provides an incentive for their cooperation. Generally, multiple-relay selection can improve the user's quality-of-service (QoS) performance, but will also result in a high price payment at the same time. Therefore, when the mobile users make decisions on relay selection, both the QoS performance (e.g., packet loss probability) can be achieved and the prices of selecting the relays should be taken into consideration.

Similar as a typical competitive hotspot environment such as a public coffee bar, where several Internet service providers may coexist to provide wireless access services (not for free) to the same group of wireless users, the author considers several SPs providing relay services to a group of mobile users requiring such service. With the purpose of pursuing the highest profits, the SPs can strategically set the prices of their relay services to attract users to select their services, resulting in a highly competitive market environment. Thus, with the capability to impact the users' decisions on relay selection and the revenue of the SPs, price competition among different SPs is the problem worthy of the author's attention. Pricing strategy of each SP has impact on the revenue of all SPs, and further affects pricing strategies of the other SPs. Motivated by the close resemblance between the problem of price competition among the SPs and the **non-cooperation game model**, the author formulates the price competition problem as a non-cooperative game with the SPs being as players and the price settings being as strategies. The most challenging issue is that, the utility function for an individual SP is naturally composed of both the relay prices and the frequency that the relays are chosen (which is affected by multiple factors such as the relay prices and the users' mobility patterns). To deal with this problem, the author first notices that given the pricing strategies of all the SPs, the frequency of selecting the relays can be known according to the proposed DRS policy. Then, the author observes that

with the knowledge of how the mobile users make relay selections based on the given relay prices, the SPs can strategically determine their pricing strategies for competition. According to such characteristics, the author designs a hierarchical framework using the **Stackelberg game model**, where the SPs are modeled as leaders while the mobile users are modeled as followers. The upper level in the hierarchical framework is a non-cooperative game, while the lower level is the DRS problem where Markov decision process (MDP) is applied in modeling, which will be specifically introduced in Chapter 3. The combinatorial strategy pair including the price settings of the SPs and the optimal DRS strategies comprises the Stackelberg equilibrium of the proposed Stackelberg game model.

1.2.2 Power Allocation for Secondary Users in Cognitive Relay Networks

The cognitive radio network is emerging as an attractive solution to address the issues of spectrum shortage and under-utilization. In a cognitive radio network, the unlicensed users (also called secondary users (SUs)) can opportunistically access the channels (i.e., frequency bands) in the spectrum allocated to the licensed users (also called primary users (PUs)) when they are found idle. In some cognitive radio networks, the licensed spectrum is treated as a tradable resource that can be leased to secondary users. In return, the SUs may offer some services to improve the PU's transmission. Considering a scenario where a PU is experiencing deep fading on its channels, suppose that the primary user is aware of the SUs in its vicinity, then it can require the neighboring SUs to relay its messages. In turn, the SUs may agree to participate in the cooperation with PU in exchange for some idle channels. With such arrangement, the PU benefits from the spatial diversity offered by the cooperating SUs whereas the SUs gain by

getting their transmission opportunities. This type of network model is referred to as a cognitive relay network (CRN).

Considering a scenario in a CRN that the transmission time of the SU is not arranged by PU directly, but depends on how long the PU's channel is idle. With the cooperation from the SUs, the PU's transmission rate will be increased. Consequently, the PU's buffer will be depleted faster, resulting in more channel idle time for the SUs to utilize. Nonetheless, due to the energy limitation, if SUs spend too much power on relaying, less power will be left for their own transmissions. Hence, a well-designed power allocation scheme is desired to address this tradeoff. For the single-SU case, the power allocation for the SU is an optimization problem with relay power and self-transmission power as the decision variables, by optimizing which the SU can achieve the highest benefit. However, when multiple SUs are included in the cooperative relaying for the PU, the problem becomes more challenging due to the following characteristics:

- The utility of an SU depends on both the idle time of the PU's channel and the self-transmission rate when it seizes the transmission opportunity. The channel idle time cannot be controlled by a single SU's relaying behavior, but depends on power allocation strategies of all the other cooperative SUs.
- When each SU makes power allocation strategy, the energy constraint has to be considered. In particular, the summation of the product of relay power and relay time, and the product of self-transmission power and self-transmission time cannot exceed an energy threshold, which is due to the battery life limit of a cognitive node. It can be observed that power allocation strategies of different SUs impact each other, since their cooperation behavior determines how long the PUs channel is idle, and thus how many transmission opportunities can be obtained by the SUs.

Note that although the SUs cooperate with each other to relay the PU's data, such cooperation are still based on selfish consideration. In other words, the reason for the SUs' spending relay power and helping with the PU's transmission together is based on the fact that they can achieve more transmission opportunities through the cooperation. The intrinsic nature of the power allocation problem in the context of multiple-SU case motivates the author to model it as a **non-cooperative game** with the SUs as the players, and the power allocation policies as the strategies. The purpose of the author is to exploit the game-theoretic tools to investigate the non-cooperative nature behind the cooperative behavior among the SUs, study how their strategies impact each other in detail, and finally achieve the Nash equilibrium as a solution to this game, where none of the SUs has incentive to deviate from the current power allocation strategies for higher benefits.

1.2.3 Cooperation Formation in Mobile Infostation Networks

With the dramatic growth in demand for broadband wireless access, wireless infrastructures become more and more widespread in our daily life. The well-known cellular systems are initially designed to provide users with "anytime anywhere" voice communications. However, due to users' increasing requirement for the quality and convenience of both work and entertainment, wireless data services become of equal importance to the traditional voice communication services. The third generation (3G) system has been proposed to efficiently support voice and data services. However, it is worthy to notice that under the charging standard for voice communications, wireless data service will not become inexpensive so long as it uses the same cellular architecture with voice traffic. Reexamining the cellular paradigm, it can be found that unlike voice systems, data services such as messaging and application downloading may have high tolerance on delay [14]. Therefore, if the application is delay-tolerant, mobility

of the users in the network can be exploited for the sake of data transmission, which inspires new network architectures to achieve a good tradeoff between the requirement for connectivity and economic investment. To offload delay-tolerant traffic from cellular network, the mobile infostation network [15] is proposed as a new network architecture with low cost. The mobile infostation network consists of an array of wireless ports (i.e., infostations (ISs)) which can be deployed along highways or placed in selected hot spots, acting as gateways to provide users with convenient access to high-speed wireless connections [14]. With proper deployment of the infostations, they can provide transmission rates higher than one megabit per second. This will allow efficient transfer of various categories of data (e.g., e-mail, fax, voice messages, game applications and movie clips). Note that unlike WiFi hot spots providing services for any general-purpose Internet access, the infostation network is more application specific.

Considering the issue of downlink data transmission in mobile infostation networks, a publish-subscribe (pub-sub) network infrastructure can be used to describe the scenario: there are one or more infostations playing the role as the data sources, while each mobile user subscribes to one infostation. Up to now, data transmission from the infostation to its subscribers is the only mechanism for content distribution. However, the performance of content distribution can be improved through some cooperation schemes. In particular, in addition to downloading data from the fixed infostations, mobile users can also act as mobile infostations and exchange information when they are in proximity. Hence, to achieve high network performances, the author considers a hierarchical network architecture for downlink data transmission in mobile infostation networks, in which the upper-level cooperation makes use of the infostation communication paradigm, while the lower-level cooperation utilizes the communications among the mobile users.

Based on the proposed hierarchical network architecture, the author observes that

the expected delay of data dissemination experienced by the users will be significantly reduced if the cooperation among the ISs is established. Assuming that there is a dedicated channel for communication among the ISs, then if two ISs form a coalition, each of them is responsible not only for transmitting data to its own subscribers, but also for providing data downloading service to the subscribers of its partner. It can be evidently seen that the diversity of the content circulating in the network (i.e., the number of pathways that a user can obtain its required content) is improved [16]. As a result, the delay performance is enhanced. Similarly, if cooperation among the mobile users is also formed, the high-quality channels between closely located users can be properly utilized to transmit required contents to each other. Therefore, the diversity of the content circulating in the network and thus the delay performance can be further improved. However, it is not always beneficial for an infostation or mobile user to cooperate with others since the cooperation also incurs cost. For the infostation cooperation, the cooperation cost is due to spending additional energy and bandwidth for transmitting contents for the cooperating infostations, while for the user cooperation, the cost comes from spending energy on forwarding contents for other cooperating mobile users. Till now, the author has observed the striking similarity between the cooperation among the ISs and the **coalitional formation game model**, which motivates the author to study the cooperative behavior among the ISs by borrowing ideas from the coalitional game theory. In the context of cooperation among the mobile users, the author observes that it is the relationship between each pair of users that determines their cooperative behavior, therefore, how the users link with each other is the major concern of the author. Note that **network formation game model** fits into this scenario exactly well, which inspires the author to model the cooperative behavior among the mobile users as a network formation process. The final objective of the author is to find a stable cooperation structure for both the ISs and

MUs. Specifically, the ISs are stably grouped into coalitions, and none of the ISs has an incentive to join a new coalition or break from its current coalition. Meanwhile, the MUs are stably connected with each other, indicating that no one can achieve a higher benefit by constructing new cooperation relationships with other users or breaking from existing cooperation relationships.

1.2.4 Content Delivery in Relay-Based Publish-Subscribe Networks

Recently, mobile social network (MSN) has been widely studied in the community, and applied to enable various communication activities amongst people with mobile devices (such as smart phones, tablets and vehicular devices). Unlike conventional communication models where mobile users have to access centralized base stations to transmit and receive data, in MSNs, mobile users can exchange their data simply by contacting with each other in an ad hoc manner. The technique of transmission by contacting is implemented and supported by the current Internet of Things (IoT) architecture and protocols, which eliminate the obstacles of different kinds of mobile devices and allow them to communicate in a unified communication infrastructure. An important issue in MSNs is the content distribution problem. Generally, there exists one or more content sources, i.e., the content providers (CPs), who have the capability to generate the contents such as daily news and digital magazines, with the purpose of gaining profits by delivering the contents to the subscribers using communication facilities such as base stations or access points. Due to the transmission reachability of a content provider, the content data should be relayed if the subscribers are not within the communication range. Therefore, as an effective solution to the content delivery problem, relay-based publish-subscribe network architecture is worthy of the author's consideration, where

content delivery paths can be constructed from the content provider to the far-away subscribers via relays in between.

Note that such a relay-based publish-subscribe infrastructure is intrinsically an asymmetric information system, i.e., the content provider has more information than the subscribers and always plays a dominant role in the content delivery process. Being in such a situation, the content provider is capable of designing “rights” and “obligation” items for the subscribers with egoistic motivation. Here the “rights” can be considered as being provided with the content delivery services, while the “obligation” can be the payment corresponding to the received service quality. As a promising technique for such an asymmetric scenario, **contract theory** is a suitable tool to tackle the content delivery problem, which motivates us to model the content provider as the principal with absolute power to make a “take-it-or-leave-it” offer to the subscribers. The author defines the delivery-ratio-within-deadline (DRWD) as the probability that a subscriber can get the content within the deadline. Using the DRWD as a measurement of the quality of the content delivery service, the author models the content delivery path from the content provider to the end subscriber as a tandem queue and the DRWD is derived from investigating the queueing model. It is observed that in order to achieve a certain DRWD, the content provider has to guarantee bandwidth allocation at itself, as well as that at the relays. In this case, the content provider has to pay the relays for providing the required bandwidth, thus, higher QoS requirement (in term of DRWD) also introduces higher cost.

Generally, different subscribers may have different requirements for the quality of the content delivery service. For instance, some subscribers would like to pay more for higher QoS, while others would rather not purchase the service if the payment cannot be tolerated. Subscriber type is used by the author to capture such heterogeneity of different subscribers, and the subscribers can be classified into different groups accord-

ing to their types. With such heterogeneity of different subscribers taken into account, the content provider can design suitable contract items for each subscriber type, in order to attract subscribers with various types to purchase its services with different qualities. By signing a contract, the content provider and the subscribers reach agreements on the content delivery service. In particular, the content provider can benefit itself most, while the subscribers have no incentives to deviate their choices of contract items.

1.3 Major Contributions

Although applying game models to studying resource allocation problems in wireless networks is not novel, and there are indeed a lot of literature working on this direction. However, this is such a generalized approach since there are so many game models and a wide range of problems in wireless networks can be solved by game-theoretic tools, and the most significant consideration is: given the considered problems, which mathematical tools can fit well into the problem frameworks and be applied to provide satisfied solutions to the problems. This is exactly the creative point the author intends to emphasize in this dissertation. Specifically, the major contributions of this dissertation (combined with the discussion of the specific problems) are summarized in the following:

1.3.1 Dynamic Relay Selection for Mobile Users and Price Competition among the Service Providers in Wireless Relay Networks

- Different from existing literature on relay-selection policy design, the author considers a challenging scenario in presence of user mobility, which requires the relay selection scheme to be adaptive to user mobility. Particularly in mobile communications, due to user mobility, relays selected at one location may no longer help at another location. As a result, a static relay selection scheme may fail to achieve the communication quality required by the mobile users. Based on user mobility pattern, a dynamic relay selection (DRS) scheme is proposed to minimize the user's long-term average cost. According to the characteristics of the relay selection problem, the author applies an optimization model based on Markov decision process (MDP) to achieve the optimal DRS policy. Impact of different factors on the DRS policy has been comprehensively studied in the simulation results, and performance comparison with the static relay selection schemes has validated the effectiveness of the proposed DRS policy.
- The author considers a challenging scenario involving several service providers (SPs) coexisted to provide relay services to the same group of mobile users, who compete with each other and attract the mobile users to select their relay services. On observing the relay prices, the mobile users will allocate their demands of the relays across the SPs. Such process will in turn impact the profits of the SPs. Considering the necessity for a joint consideration of the relay selection policy based on given prices and the price competition among the SPs, the author proposes a hierarchical framework with a Stackelberg game model utilized to investigate the interactions between components belonging to different hierar-

chies. In this framework, the SPs and the mobile users are modeled as leaders and followers, respectively. The upper-level problem is a non-cooperative game aiming at studying the competitive behavior among the SPs, while the lower-level problem is the DRS policy design based on the relay prices announced by the SPs. The obtained results can help to provide a guidance for SPs to compete for providing relay services.

1.3.2 Power Allocation for Secondary Users in Cognitive Relay Networks

- Different from most of the existing works studying the cooperative relaying of the SUs to improve the PU's transmission, where the PU decides how much time the SUs can utilize its channel, the author considers a challenging scenario in which the transmission time of the SU is not arranged by the PU directly, but depends on how long the PU's channel is idle. On one hand, the SUs can use some relay power to speed up the PU's transmissions. Consequently, the PU's buffer will be depleted faster, resulting in more channel idle time (i.e., more transmission opportunities for SUs). On the other hand, due to the energy limitation, less power can be used for the SUs' own transmissions if too much power is consumed on relaying. To address this tradeoff, the author proposes optimal power allocation strategies in the context of both single-SU and multiple-SU cases.
- Power allocation for the single-SU case is modeled as a utility maximization problem. The author investigates the relationship between relay power and self-transmission power at the optimal solution, which is concluded by important propositions leading to an effective power allocation of the SU.

- Based on the non-cooperative nature implicitly included in the cooperation behavior among the SUs, the power allocation problem involving multiple SUs is modeled as a non-cooperative game. After a clear definition of the Nash equilibrium for the power allocation game, the author makes a comprehensive study of the characteristics of the Nash equilibrium, including its existence and uniqueness. Moreover, the impacts of different system parameters such as the average packet arrival rate at the PU and the channel gain of different channels are investigated through numerical results.

1.3.3 Cooperation Formation in Mobile Infostation Networks

- The author observes that the cooperation among infostations and mobile users can lead to an obvious improvement of the network performance in term of the expected delay of content distribution experienced by the users. More interestingly, jointly taking both levels of cooperation into account may achieve even better improvement in performance. The author is inspired to propose a hierarchical cooperation formation model and analyze the bi-level cooperation in mobile infostation networks.
- After a comprehensive study of the cooperative behavior of the infostations and the mobile users, the author finds a striking resemblance between the problem of infostation cooperation and the coalitional formation game model, and also between the problem of mobile user cooperation and the network formation game model. The author has arrived at a stable cooperation structure for both the infostations and the mobile users, which is considered as the solution to the proposed hierarchical cooperation formation framework. Specifically, the infostations are stably grouped into coalitions, and none of the infostations has an incentive to

join a new coalition or break from its current coalition. Meanwhile, the mobile users are stably connected with each other, and no one can achieve a higher benefit by constructing new cooperation relationships with the other users or breaking from existing cooperation relationships.

- A distributed hierarchical cooperation formation algorithm is presented, including the essence of the merge-and-split schemes [17] targeted on guiding the infostations to form coalitions, and the key idea of Jackson and Watts's network formation method [18] exploited to lead the mobile users to a pair-wise stable network. Through extensive numerical experiments, the proposed algorithm is shown to be highly effective in obtaining the stable hierarchical cooperation structure. The proposed hierarchical cooperation formation model can also provide some insight into other network scenarios, where cooperation is needed to improve the network performance.

1.3.4 Content Delivery in Relay-Based Publish-Subscribe Networks

- To describe a relay-based publish-subscribe network (including components such as content providers, subscribers and relays), and study the content delivery from the content provider to the subscribers, the author constructs a tree-structured abstract model with the content provider as the root node, relay nodes as the intermediate nodes and subscribers as the leaf nodes. Based on the characteristic that the contents are generated at the root node, and finally received by some leaf nodes, the concept of content delivery path is proposed. Considering the possible content arrivals and departures at each node, the content delivery along a certain path is modeled as a tandem of queues. Studying the performance of the

proposed queueing model provides the author with insight on the effectiveness of the content delivery process.

- Based on the fact that contents always have their time-effectiveness, and the practical concern of what the mobile subscribers care most in the content delivery process, the author imports the concept of *delivery ratio within deadline* (DRWD), which is defined as the probability that a subscriber can obtain the content within a specific deadline, as a measurement of the content delivery service quality. A higher DRWD indicates a better content delivery service quality. Absorbing Markov chain model is applied to derive the DRWD based on the tandem queueing model.
- Considering the content delivery service as a kind of “commodity”, the relay-based publish-subscribe network can be viewed as a trading market from the perspective of economics. The content provider is referred to as the seller having overwhelming control over the content delivery process, while the subscribers are referred to as the buyers, which can only decide whether to purchase the content delivery service and which quality to buy. According to the problem characteristics, contract theory is properly applied to model the special interactive behavior between the content provider and the subscribers. Taken the heterogeneity of the subscribers (denoted by different subscriber types) into account, the author proposes an optimal contract design scheme for the content providers, with the employment of which the content provider’s profit can be maximized, while the subscribers are satisfied with the contract items designed for their particular types.

1.4 Dissertation Organization

The rest of this dissertation is organized as follows:

Chapter 2 provides a brief introduction to the game-theoretic tools which will be applied in this dissertation, and presents a comprehensive literature review of all the problems described in Section 1.2, where the advantages and limitations of the existing related research works are discussed.

Chapter 3 considers the environment of wireless relay networks, jointly takes the dynamic relay selection problem for mobile users and price competition among the service providers into consideration, and designs a Stackelberg-game framework to solve the hierarchical problem. The combinatorial strategy pair consisted of the pricing strategies of the service providers, and the optimal relay selection strategies of the mobile users, has been obtained.

Chapter 4 studies the power allocation problem for the cooperative secondary users in cognitive relay networks, in which the tradeoff of the intention to spending more relay power to speed up the clearance of the primary user's data buffer, and the desire for more self-transmission power to be used when having seized the transmission opportunity, has been addressed. The power allocation problem for single-SU case is formulated as a utility maximization problem, while that for multiple-SU case is modeled as a non-cooperative game. The existence and uniqueness of the Nash equilibrium as the power allocation solution are comprehensively investigated and proved.

Chapter 5 investigates the hierarchical cooperation formation within the architecture of mobile infostation networks, where both the cooperation among the ISs and among the MUs are discussed. In particular, the former is modeled as a coalitional formation game, while the latter is modeled as a network formation game. A distributed hierarchical cooperation formation algorithm is proposed to achieve the stable

cooperation structure for both the ISs and the MUs.

Chapter 6 discusses the problem of content delivery in relay-based publish-subscribe networks, where contract theory is utilized to model the interactions between the content provider and the subscribers. With the power of dominating the content delivery process, the content provider can strategically design the contract items and make a “take-it-or-leave-it” offer to the subscribers. The heterogeneity of the subscribers is implicitly taken into account in the optimal contract design. Numerical results are provided to verify that by signing a well-designed contract, the content provider can achieve the highest benefit while the subscribers are also satisfied with their choices of contract items.

Chapter 7 provides the concluding remarks of this dissertation, and outlines potential directions of the future work.

Chapter 2

Literature Review

In this chapter, the author first gives a detailed introduction to some game-theoretic tools in Section 2.1, and then, a comprehensive survey of the literature on the specific research issues investigated in this dissertation will be provided. The relay selection problem and the price competition among wireless service providers are reviewed in Section 2.2 and Section 2.3, respectively. The study of cooperative relaying in cognitive relay networks are surveyed in Section 2.4. Content distribution in mobile infostation networks are reviewed in Section 2.5. In Section 2.6, the research works on content delivery in publish-subscribe (pub-sub) mobile social networks and on the application of contract theory are surveyed.

2.1 Game-Theoretic Tools

In the following, the author will introduce some fundamental game-theoretic tools such as: (1) non-cooperative game; (2) Stackelberg game; (3) coalitional formation game; (4) network formation game, and also make a brief introduction to contract theory, which can be considered as a special branch of game theory. According to different

characteristics of different issues, these game models can be properly applied to solve the problems considered by the author. Specifically, (1) and (2) are applied in Chapter 3 to solve the problem of price competition among the service providers (SPs). (2) is used to investigate the power allocation for multiple secondary users (SUs). Due to the advantages of (3) and (4) in studying the cooperative behavior of multiple decision-makers, they are utilized to analyze different levels of cooperation in mobile infostation networks, while contract theory is exploited to achieve an economically optimal content delivery in relay-based pub-sub networks.

Non-cooperative game, which is one of the most significant building blocks of game theory, aims at studying and analyzing the competitive decision-making among several players. It succeeds in proving an analytical framework for modeling the interactions among multiple players with partially or entirely conflicting expectations for the outcome of the game, which is affected by the actions of all the players. For instance, utility companies providing service to the same market have to compete over pricing strategies for the sake of maximizing their profits. A formal definition of a non-cooperative game in strategic (or normal) form is given in the following [5]:

Definition 2.1. *Non-cooperative game in strategic form is denoted by a triplet $\mathcal{G} = \{\mathcal{N}, (\mathcal{S}_i)_{i \in \mathcal{N}}, (\mathcal{U}_i)_{i \in \mathcal{N}}\}$, where:*

- $\mathcal{N} = \{1, \dots, N\}$ is a finite set of players.
- \mathcal{S}_i ($i = 1, \dots, N$) is the set of available strategies for player i .
- $\mathcal{U}_i : \mathcal{S} \rightarrow \mathcal{R}$ ($i = 1, \dots, N$) is the utility (payoff) function of player i , where \mathcal{S} is the Cartesian product of the strategy sets of all players, i.e., $\mathcal{S} = \mathcal{S}_1 \times \dots \times \mathcal{S}_N$.

For a non-cooperative game in strategic form, each player has to choose its strategy $s_i \in \mathcal{S}_i$, to optimize its utility function (i.e., maximize the profit or minimize the cost),

which does not only depend on the selection of s_i , but also on the strategies of the other players, denoted by $\mathbf{s}_{-i} = [s_j]_{j \in \mathcal{N}, j \neq i}$. $\mathbf{s} = [s_i, \mathbf{s}_{-i}]$ is referred to as a *strategy profile*. It should be emphasized that in non-cooperative game, cooperation among the players are not forbidden. However, all the possible cooperation are still due to selfish consideration, i.e., the choice of cooperating with others of any player must be self-enforcing without coordination and communication with the other players to achieve an interest collusion. The most accepted solution to a non-cooperative game is a well known concept named *Nash equilibrium* (proposed by John Nash), which is defined as follows [19]:

Definition 2.2. *A pure-strategy Nash equilibrium of a non-cooperation game $\mathcal{G} = \{\mathcal{N}, (\mathcal{S}_i)_{i \in \mathcal{N}}, (\mathcal{U}_i)_{i \in \mathcal{N}}\}$ is a strategy profile $\mathbf{s}^* \in \mathbf{S}$ satisfying the following condition:*

$$\mathcal{U}_i(s_i^*, \mathbf{s}_{-i}^*) \geq \mathcal{U}_i(\tilde{s}_i, \mathbf{s}_{-i}^*), \quad \forall \tilde{s}_i \in \mathcal{S}_i, \quad (2.1)$$

for all $i \in \mathcal{N}$.

The essence of the Nash equilibrium is that none of the players can further improve its utility by unilaterally changing its strategy, when the other players keep their current strategies fixed. In order to solve a non-cooperative game and achieve the Nash equilibrium, the concept of *best-response* plays a vital role, which is defined as follows:

Definition 2.3. *The best response $\mathcal{B}_i(\mathbf{s}_{-i})$ of a player i to the strategies of the other players (i.e., \mathbf{s}_{-i}) is a set of strategies satisfying the following condition:*

$$\mathcal{B}_i(\mathbf{s}_{-i}) = \{s_i \in \mathcal{S}_i | \mathcal{U}_i(s_i, \mathbf{s}_{-i}) \geq \mathcal{U}_i(\tilde{s}_i, \mathbf{s}_{-i}), \forall \tilde{s}_i \in \mathcal{S}_i\}. \quad (2.2)$$

The best response for player i indicates that given the strategies of the other players, player i has no incentive to choose the other available strategies in \mathcal{S}_i rather than

strategies in $\mathcal{B}_i(\mathbf{s}_{-i})$, for the sake of optimizing its utility. Given the definition of best response, the Nash equilibrium can have an alternative explanation as follows [3]:

Proposition 2.1. *A strategy profile $\mathbf{s}^* \in \mathcal{S}$ is a Nash equilibrium of a non-cooperative game $\mathcal{G} = \{\mathcal{N}, (\mathcal{S}_i)_{i \in \mathcal{N}}, (\mathcal{U}_i)_{i \in \mathcal{N}}\}$ if and only if each player has made the best response to the strategies of the other players, i.e., for each player i , we have*

$$\mathbf{s}_i^* \in \mathcal{B}_i(\mathbf{s}_{-i}^*). \quad (2.3)$$

Stackelberg games are a special class of non-cooperative games, characterized by the hierarchy naturally existed among the players. Specifically, one or more players can take actions ahead of the other players and announce their strategies, based on which the other players can make their decisions. In such a hierarchical framework, the players with the capability to enforce their own strategies on the others is referred to as the *leader*, while the players who can only react to the strategy announced by the *leader* are referred to as the *followers*. Taking a two-player non-cooperative game involving one leader and one follower as an example, and let \mathcal{S}_1 and \mathcal{S}_2 denote the strategy set of the leader and the follower, respectively. When the leader announces its strategy $s_1 \in \mathcal{S}_1$, the follower has to respond with its strategy $s_2 \in \mathcal{S}_2$. With the assumption that the follower is rational, it will always choose strategies to maximize its utility given the leader's strategy, which leads to the definition of optimal response in the following:

Definition 2.4. *In a two-person finite game, the optimal response of player 2 to the strategy of player 1, i.e., $s_1 \in \mathcal{S}_1$, is given by*

$$\mathcal{R}_2(s_1) = \{s_2 \in \mathcal{S}_2 | \mathcal{U}_2(s_1, s_2) \geq \mathcal{U}_2(s_1, \tilde{s}_2), \forall \tilde{s}_2 \in \mathcal{S}_2\}. \quad (2.4)$$

Note that the optimal response of player 2 to the strategy of player 1 can be considered as a special case of the best response defined for general non-cooperative games. Again, with the assumption that the leader is rational, it can strategically update its strategy $s_1 \in \mathcal{S}_1$ with the knowledge about how the follower reacts to given s_1 , in order to achieve the highest utility. Then comes the following definition:

Definition 2.5. *In a two-person finite game with player 1 as the leader, the Stackelberg strategy of player 1, i.e., s_1^* is defined as:*

$$s_1^* = \arg \max_{s_1 \in \mathcal{S}_1, s_2 \in \mathcal{R}_2(s_1)} \mathcal{U}_1(s_1, s_2). \quad (2.5)$$

Combining the optimal decision-making process of both the leader and the follower, a Stackelberg solution of the two-person game can be defined as follows:

Definition 2.6. *The strategy pair (s_1^*, s_2^*) is a Stackelberg solution to the two-person game with player 1 as the leader, if and only if the following conditions are satisfied:*

- $s_1^* \in \mathcal{S}_1$ is a Stackelberg strategy for the leader. i.e., player 1.
- $s_2^* \in \mathcal{R}_2(s_1^*)$.

Based on the Stackelberg solution (s_1^, s_2^*) , the utility pair $(\mathcal{U}_1(s_1^*, s_2^*), \mathcal{U}_2(s_1^*, s_2^*))$ is the corresponding Stackelberg equilibrium outcome.*

After the introduction of the non-cooperative game model, the author will introduce the coalitional game theory, which is a main building block of cooperative games. Coalitional game models have also been widely applied in various disciplines such as economics, politics, biology science and wireless communications [5]. For instance, if several mobile users attempt to download contents from an access point belonging to

a certain service provider, however, buying bandwidth separately may not be economically beneficial. In this case, they may form a coalition to buy a certain amount of bandwidth together, and negotiate on how to efficiently use the bandwidth among themselves, for the sake of spending less and obtaining more satisfaction. In the following, the author will focus on coalitional formation games and network formation games, which can be considered as a branch of coalitional graph games.

Considering the players as a set of nodes with intelligence, forming a coalition may bring the coalition members substantial benefit. For instance, consider the case in which nodes belonging to one coalition are responsible for helping with each other when necessary. Each node can achieve higher benefit with others' cooperation than simply acting alone, however, at the cost of using some self resources to cooperate with the others. In this case, several important questions need to be answered, such as: how will the nodes form coalitions? Can the coalitional structure they form be considered as stable? Will someone prefer acting alone to joining any coalition? and so on. Coalitional formation game model is mainly utilized to study the network coalitional structure and provide answers to the questions similar as mentioned above. Different from canonical games which is superadditive [5], coalitional formation game characterizes itself by the existence of cost for forming coalitions. Thus it is not always beneficial to form a grand coalition, where all nodes are included in the same coalition. To achieve a stable coalitional structure, several dynamic coalitional formation algorithms have been proposed such as heuristic methods [20], set-theory approaches [21], Markov chain-based schemes [22]. The so-called merge-and-split algorithm [21] has been proved to be an effective and efficient generic rule to derive application-specific coalitional formation, which will be introduced in detail in Chapter 5. Different stability criteria in terms of \mathbb{D}_c stability and \mathbb{D}_{hp} stability are provided in [5], which can be applied to judge whether the obtained coalitional structure is stable or not.

In contrast to the coalitional formation game, in which the payoffs of the players are fixed so long as the coalitional structure has been determined, in network formation game (based on the pioneering work of Myerson [23]), the particular connections of the players in each single coalition also has impact on the players' payoffs. Network formation game model can be considered as a combination of coalitional graph games and non-cooperative games (which also play a significant role in the process of forming a network). The objective of network formation game is to form a network graph and guarantee that it is in stable state. Both *myopic* and *far-sighted* approaches [18] can be applied to form the network graph, while the former method allows the players to choose strategies solely based on the current network state, while the latter scheme inspires the players to adapt their strategies through learning procedure and making prediction on the future strategies of the other players. Here, strategy of a player is to choose one or more connection links to form or break, *myopic best-response dynamics* can be exploited to solve the network formation game, where each player is able to select its own strategy to maximize its utility. The best-response dynamics can lead the network formation process to a Nash network under certain conditions, indicating that the players in the network formation game has arrived at a Nash equilibrium. Due to the drawback of pursuing a Nash network which may lead to trivial graphs such as an empty graph, a new network formation game has evolved as a solution, where new concepts of stability such as *conditional stability* and *pairwise stability* [18] are utilized. The concept *pairwise stability* will be applied as the rule to guide the network formation procedure in Chapter 5.

As for contract theory, there are different ways to define its confines such as a subset of game theory or an extension of price theory which studies how decision makers interact when being provided with the power to select prices [24]. However, the author would argue that it can be considered as a special branch of game theory,

where one player of the game (which is typically called as the principal) is provided with all bargaining power and thus is capable of making a “take-it-or-leave-it” offer to the other players. For many cases, the principal and the other players have conflict objectives, which proposes an important question to the principal that how to design a contract to maximize its own benefit, while the other players would like to accept the offered contract items. Generally, a contract item is in the form of a (*reward*, *cost*) pair, where *reward* specifies what the agent can gain, and *cost* indicates that how much the agent has to pay for the corresponding *reward*. Note that all the other players are also free decision makers, but are characterized by constrained actions of selecting contract items provided by the principal, or rejecting to accept any item. Contract theory targets on understanding organizations, institutions, and interest relationship between the principal and the other agents with possibly different personal preferences. Most of the techniques used in contract theory are also developed by game theorists aiming at investigating resource allocation schemes. Due to the characteristics of contract theory, it becomes an effective tool to design incentive compatible mechanisms in a monopoly market with incomplete information [25]. The essence of contract theory is to strategically design proper contract items so that all agents have the incentive to reveal their private information truthfully.

2.2 Relay Selection Issue

Having been incorporated into the 802.16j standard for mobile multihop relay systems [26], cooperative communication is envisioned to be a promising technology to effectively combat wireless channel fading and improve the performance of wireless systems, such as higher spacial diversity, higher throughput and lower delay, lower power consumption, extended coverage and adaptability to network conditions [12]. In the new

paradigm brought by cooperative communication, relay nodes are assigned to help with the source node's transmission, resulting in a virtual distributed antenna array, which can achieve transmit diversity without requirement for installing multiple antennas in mobile devices. In cooperative wireless communication, considering a wireless network which can be cellular or ad-hoc, when the source aims at transmitting its data to the destination, taking advantage of the information overheard by the neighboring nodes can bring the source a more robust communication, which is the essence of cooperative communication [27]. In particular, cooperative techniques utilize the broadcast nature of the wireless channels, where some intermediate nodes (relays) between the source and the destination may overhear the signals transmitted by the source and retransmit the signals to the destination. Before the retransmission, the relays can simply amplify the received signals, or decode and then re-encode them, which corresponds to two most widely-used relaying schemes, namely amplify-and-forward (AF) and decode-and-forward (DF) relaying, respectively. While a lot of efforts have been put on the research in cooperative communications, there are still a lot of important issues remained to be addressed, one of which is that how cooperative relays are assigned to the transmission source, or how can the source choose its partners in the cooperation? Such question is the major concern of the author's research on wireless relay networks.

Generally, the relay selection schemes can be classified into single relay selection and multiple relay selection schemes. Single relay selection has been extensively studied. In [28], the neighbor node with the maximum signal-to-interference-and-noise ratio (SINR) is selected as the best relay, while in [29] the node nearest to the base station is employed in cooperation. For dual-hop protocols [30], each relay is associated with both the source-to-relay channel and the relay-to-destination channel. The relay whose worse channel between source-to-relay and relay-to-destination is the best will be selected. For a similar case, another strategy named as the best harmonic mean

selection is proposed in [31]. In [32], optimal selection and transmission of a single relay among multiple candidates using AF relaying strategies is proposed to minimize the outage probability. A distributed relay selection scheme is proposed in [33], where the user transmits with the best end-to-end path among a set of relays, based on the instantaneous condition of channels. In [34], selective single-relay cooperative scheme combining selective-relay cooperative communication with physical-layer power control is proposed. The best relay is selected in a distributed fashion with minimum signaling overhead. The impact of the interference on the relay assignment problem within single-relay cooperative scenario is investigated in [35]. In [36], the problem of relay selection based on instantaneous channel conditions to account for user mobility is investigated.

Although single-relay selection is attractive due to its simplicity, it may fail to meet the QoS performance required by users due to the limited diversity gain. To enhance the service quality, multiple-relay selection may be favored to increase the cooperative diversity order. In [37], the idea of single-relay selection is extended to multiple-relay selection and two SINR-suboptimal schemes are proposed based on the concept of relay ordering and recursion, respectively. In [38], a threshold-based relay selection protocol is proposed, in which all qualified relays (with received SINRs higher than a pre-determined threshold) are allowed to cooperate. The authors in [39] further propose a multiple-relay selection scheme for the cooperative spatial multiplexing with multiple antennas at the receiver side. In [40], the authors propose a centralized method for grouping and partner selection to achieve the maximal network-wide diversity. In the proposed protocol, each transmitting node is assigned enough partners which can improve the probability of successfully decoding the message. In [41], a relay selection strategy for multi-user cooperative diversity systems is proposed, where only a subset of the available nodes can be activated to achieve a balance between the transmission

reliability and the total energy consumption. Based on the theory of combinatorial optimization and specifically on the knapsack problem, the authors in [41] propose a method of optimizing the relay selection among the candidates, particularly minimize the error probability while the constraints on total energy consumption have to be satisfied.

It can be noticed that, few of the aforementioned relay selection schemes has taken user mobility into account, which limits their applicability in mobile communications. And to the best of the author's knowledge, there is no relay selection scheme considering user mobility and meanwhile enforcing no constraint on the number of relays can be selected. In order to satisfy the different QoS requirements of the heterogeneous mobile users, it is necessary for the relay selection scheme to be dynamic to adapt to user mobility. Moreover, users with higher QoS requirements are allowed to select multiple relays when necessary, which is the motivation of the author to design a dynamic relay selection scheme in the research work.

2.3 Price Competition among Wireless Service Providers

Besides the resource competition among wireless users, the competition for revenue among wireless service providers is also a common problem in a non-cooperative environment [5], hence, studying the competitive behavior of the service providers plays a vital role in modeling the non-cooperative scenarios in wireless systems. For instance, several Internet service providers may coexist to provide Internet access service in a typical hotspot scenario such as a restaurant or a coffee bar, targeting on the same group of users. The scenario results in a highly competitive market environment for

the Internet-access service providers, competition is inevitable since each of them attempts to attract wireless users for the sake of gaining profits. To achieve this goal, the service providers need to strategically set prices for their services, with the QoS they can provide taken into consideration. From the perspective of the wireless users, they can allocate their demands across different Internet service providers on observing the service prices and the QoS can be obtained. The service selection strategies will in turn affect the profits of the service providers, motivating them to update their pricing strategies in the competition. It can be observed that with the capability to affect the demand of the users and the revenue of the service providers, pricing in wireless networks is considerably worthy of investigation [42].

General ideas on pricing can be found in [42]. In [43], the authors study and analyze the pricing game among wireless providers when the capacity is fixed. In [44], a pricing issue where the users are price-sensitive and congestion delay-sensitive are investigated, with tight bounds on the price of anarchy (PoA) [45] given. In [5], the duopoly competition involving two Internet service providers is comprehensively discussed, where the two service providers compete for attracting the same group of users by adjusting their pricing strategies. The service providers are pursuing their maximum profits, while the users make intelligent choices on service providers, the one offering the most beneficial combination of QoS and the corresponding price is always preferred. However, so far as we know, few of the existing research works has discussed the problem of price competition among wireless service providers offering relaying services. The author focuses on this issue based on the consideration that for most cases, wireless devices in cooperative networks do not belong to a single entity, but many independent profit-maximizing entities. In a wireless relay network, the relay nodes can be either wireless relay stations deployed by wireless service providers, or dual-mode mobile terminals which can relay messages voluntarily for other users

[13]. Pricing is an important issue for the former case. Voluntary cooperativeness assumption may not be valid in reality because relaying for other nodes consumes energy and other resources of the relay node. A natural solution is to make payment to the relay nodes, which provides an incentive for their cooperation. Therefore, the service providers are possible to gain profit by deploying relays and strategically set the prices of their relay services.

The pricing issue of relay service is considered in [46], in which each relay belongs to a service provider. In [47], a pricing-based system is proposed to stimulate the cooperation via payment to the relay nodes. In [48], a buyer/seller Stackelberg game is introduced, in which a single buyer attempts to buy service from multiple relays. The buyer announces its relay selection decision and the required transmission power, then the relays ask for proper prices to maximize their profits. In this work, the author considers a scenario that one service provider can deploy multiple relays in different locations. The author investigates how each service provider should set the price for its relay service to maximize its profit.

2.4 Cooperative Relaying in Cognitive Relay Networks

With the rapid deployment of emerging wireless devices and applications, a growing demand for wireless radio spectrum has been witnessed. However, it is observed that a large portion of the licensed spectrum, which has been assigned to licensed holders (namely the primary users) by the Federal Communications Commission (FCC), always remains under-utilized, resulting in an inefficient spectrum utilization [49]. Such disadvantage of the existing spectrum usage has motivated researchers from both academia

and industry to review the policy and intend to look for solutions for utilizing the limited spectrum resources in a more intelligent and flexible way. The innovative concept of cognitive radio has been proposed to resolve the problem of inefficient spectrum utilization, and has been receiving more and more attention recently. Cognitive radio differs from the traditional communication techniques since the devices are capable to adapt their operating parameters (such as transmission power and frequency) to the change in the surrounding environment [50]. With cognitive capability [51], the cognitive radio devices can be aware of the possible transmission opportunities and tune their transmitting frequencies to the available white space in spectrum. A typical cycle of cognitive radio consists of several major components, i.e., detecting spectrum not in use, choosing the best frequency bands, coordinating spectrum use with the other users, and leaving the frequency band when a licensed user comes back [49]. In order to support the cognitive radio cycle, there are some important issues required to be addressed, i.e., spectrum sensing and analysis, spectrum management and handoff, spectrum allocation and sharing, the last of which is the major concern in the author's research work on cognitive radio networks.

The author considers the cooperative relaying environment, where the broadcast nature wireless networks are utilized to improve the network performance by increasing spacial diversity [52], [53]. Cognitive relay network (CRN) is a perfect integration of wireless relay network and cognitive radio technology, where the secondary users can be permitted to use the under-utilized spectrum of the primary user, through relaying for the primary user to improve its transmission quality. Recently, several CRN models are proposed in the literature (i.e., [54], [55], [56], [57], [58]). In [54], a CRN is proposed in which the PU calculates the optimal time slot duration for leasing its channels to the SUs, to maximize its QoS. The SUs compete with each other for transmission within the leased time slot. Considering that transmission rate is no longer the focus for the

PU when it is already satisfied, [55] proposes an improved framework based on [54], in which the SUs also make a payment proportional to the channel access time allotted by the PU. In [56] [57] [58], along with the problem of leasing spectrum to the SUs, the issue of power allocation within the CRN has been considered. A novel dynamic spectrum leasing scheme in CRN via power control games involving both PU and SUs is proposed in [56]. With the objective to optimize the CRN performance while limiting the interference measured at the primary receiver, transmit power strategies among the cognitive relays are investigated in [57]. In [58], a cooperative Nash bargaining power-control game is formulated. In the above works, a common point is that, it is the PU who decides how much time the SUs can utilize its channel. Nonetheless, in this work, the author considers a different scenario in which the transmission time of the SUs are not arranged by the PU directly, but depends on how long the PU's channel is idle. The non-cooperative nature behind the cooperative behavior among the SUs makes the problem challenging, especially when considering the tradeoff of spending power on relaying and on self-transmission.

2.5 Content Delivery in Mobile Infostation Networks

Cellular systems has brought wireless users with “anytime anywhere” communications. On the contrary, the concept of infostation network involving an array of isolated wireless access points (namely infostations) has been proposed for facilitating frequent access to high bit-rate connections [14]. Such “many-time many-where” communication is suitable for a plethora of wireless data services such as messaging. Different from voice systems where the transmitted signal (i.e., voice) is extremely sensitive to delay, and thus constant bit-rate service is necessary, the messaging and information systems

are much more tolerative about delays [59]. Moreover, while voice systems are generally equal for uplink traffic and downlink traffic, the downlink traffic is predominant for messaging and data services. The characteristics of the problem of effectively providing data services lead to a wireless infrastructure with isolated infostations, which caters to the mobility of users as well as their indispensable requirements for information [14], [60]. Individual infostations in scenarios such as freeways and streets may function independently of each other, providing specialized gateways to accessing to the Internet or remote servers for data information [14]. Mobile users cannot access the network infrastructure when they are not within the transmission coverage of the infostations [61].

There is a rich body of literature on studying mobile infostation networks. While theoretical analysis of network delay, transmission range and capacity is provided in [62], [63], [64], many other works focus on the performance evaluations of protocols and applications [65], [66]. [14] can be considered as a pioneer work on infostation networks, which describes the challenges of infostation protocol design. A mathematical analysis of infostation information transfer is also presented, focusing on the optimal combination of transmission rate and coverage area. In [61], the authors derive an analytical framework for studying the conditions for stability in infostation systems, and the results can be a guidance for designing a stable network. There are also research works such as [67] focusing on the tradeoff between energy efficiency and performance experienced by mobile users. From the application perspective, [67] focuses on timely and energy-efficient communications in rural infostation systems. Analytical results are utilized to gain insights into the corresponding protocol design. In [68], the authors also consider a rural infostation system, and propose a cross-layer resource allocation scheme for efficient data dissemination. The numerical results obtained from studying a cross-layer design example verify that the performance of direct transmission can

be guaranteed for the relays and the QoS of content distribution services can be improved. Data exchanging is investigated for hierarchical infostation systems in [69], where a new network entity named as access controller is introduced. A middleware design is presented to be adaptable to the new network paradigm where the storage model and class of services can be implemented.

Note that the issue of downlink data transmission in mobile infostation networks is similar to data downloading through a roadside unit (RSU) in a vehicular ad hoc network (VANET), there are also a variety of research works discussing this topic. In [70], the authors focus on a content downloading system leveraging both infrastructure-to-vehicle and vehicle-to-vehicle communication. With the aim to maximize the system throughput, a max-flow problem is formulated, which accounts for several practical aspects such as channel contention and the data transfer paradigm. In [71], a scheduling algorithm is proposed to download data from RSU using multicast technique in order to serve multiple requests, with the priority of data taken into account. In [72], popular content distribution in VANETs is addressed, where a popular file is downloaded by a group of on-board units (OBUs) passing by a road-side unit (RSU). Since the OBUs may not finish downloading the entire file from the RSU due to high speed and channel fading, but only possess several content pieces, a cooperative method using coalitional graph game is proposed to establish a peer-to-peer network among the OBUs, where the content pieces can be exchanged to complete the file downloading procedure. Considering the problem of limited RSU coverage, the authors in [73] investigate the possibility of exploiting parked vehicles to extend the RSU service coverage. With similar objective, the authors in [74] propose a scheme where vehicles are selected as relay nodes to connect with each other distributively and thus a hierarchical structure is formed.

In recent years, social tie has also been exploited for content distribution in mobile

infostation networks, in which the cooperative relay selection is based on the social contacts. Nonetheless, almost all of these studies rely on the assumption that mobile users are willing to cooperate by default, which is unrealistic for many applications. One of the few examples is [16], in which the issue of noncooperation in the context of mobile infostation networks for movie downloading is discussed. Based on the fact that nodes are selfish, a social contract is constructed that two nodes will exchange files only when each of them has something that the other does not have but requires. The authors construct the content distribution model for the case in which all mobile nodes have common interest, and then, the common interest model is extended to a more general case where the nodes may have various interests while some of them may be overlapped. From the above related works discussing the data transmission in mobile infostation networks, it can be observed that few of them has taken both cooperation among the infostations and that among the mobile users into account, while a joint consideration of such bi-level cooperation is the main focus of the author's research work. The author designs a hierarchical cooperation formation model for downlink data transmission in mobile infostation networks, where game theoretic models are utilized to analyze the cooperative behavior of the infostations and the mobile users, as well as the stability of the achieved cooperation structure.

2.6 Content Delivery in Publish-Subscribe Mobile Social Networks

2.6.1 Delay Tolerant Networks

Delay-tolerant networks (DTNs) have the potential to interconnect devices in different regions that current networking technology cannot reach. To achieve this goal, routes

must be found over multiple intermittently-connected hops [75]. The authors in [75] design a metric estimating the average waiting time for each potential next hop, based on which a routing protocol that uses only observed information about the network. Taking the advantage of node mobility and end-to-end disconnections in DTNs, [76] extends the situation that data is forwarded to a single destination, and studies multicast in DTNs from the social network perspective. The difference between unicast and multicast in DTNs are investigated, and relay selections for multicast is formulated as a unified knapsack problem by exploiting node centrality and social community structures. In most of the existing routing algorithms for DTNs, it is assumed that nodes are willing to forward data for each other, the authors in [77] propose a Social Selfishness Aware Routing (SSAR) algorithm based on the realization that most people are socially selfish. Data forwarding process is formulated as a multiple knapsack problem with assignment restrictions (MKPAR) to satisfy user demands for selfishness and performance. In [78], human-to-human communications in DTNs, where human behavior exhibits the characteristics of networks by forming communities are investigated. Besides routing in DTNs, there are various other problems attracting researchers' interest. In [79], the authors propose efficient buffer management strategies for DTNs. In [80], data compression algorithms for energy-constrained devices in DTNs are investigated.

2.6.2 Mobile Social Networks

Recently, mobile social network (MSN) has been increasingly interested by the community. The general concept and possible architectures have been proposed that the MSN is a system mainly composed of heterogeneous types of mobile devices communicating with each other in an opportunistic manner. The applications of the MSN are various, including location-based services [81], data transmission [82], healthcare [83], etc. One of the important concerns of MSN is the content distribution in an MSN, e.g., in [84],

a set of wireless users request to obtain frequent update of contents from the content provider. The content provider in [84] allocates its downlink channel capacity to offer the users a relatively fair chance to receive “fresh” contents. Besides, the contact process and contact graph of an MSN are studied in [84]. The QoS issues are discussed in [84] and [82] that the contents in a content subscription system should be optimized to arrive at the content users on time. [82] also presents that the delay can be measured by the delivery ratio within the time constraint T .

2.6.3 Publish-Subscribe Systems

Publish-subscribe (pub-sub) is a typical paradigm for content dissemination in a distributed system with autonomous content users. As described in [85], a pub-sub system is driven by the events that the subscribers notify the subscription information to the publishers, and the publishers efficiently manage the subscriptions and make content deliveries. Also, a grand view of pub-sub system and its variants (e.g., topic-based, type-based and content-based pub-sub) are presented in [85]. Pub-sub system adopted to a mobile environment is studied in [86], where the contents are generated from the publishers and routed to a content broking system (a centralized server, or a cluster of distributed servers) for subscription management and deliveries. However, [86] does not further consider the system performance and incentive issues of the system components. [87] develops an efficient multicast algorithm for content-based pub-sub systems. A tree structure is established to describe the connections among publishers, content brokers and subscribers. However, [87] is based on general cases, so that the issues of mobile systems, e.g., contact rate and bandwidth, have not been discussed.

2.6.4 Contract Theory Application

Contract theory [25] studies how the two rational and self-interested parties in a system reach agreements in an economic manner. Recently, contract-theoretic approaches are employed to analyze resource distribution in wireless systems. For example, [88] applies contract theory model to optimally solve the spectrum resource allocation problem in cognitive radio systems. The spectrum owner (i.e., primary user) is the monopolist who designs contracts to sell the resources, while the secondary users purchase those contracts. To ensure the secondary users are willing to accept the contracts, the primary user makes the contracting decision subject to secondary users' incentive compatibility (IC) and individual rationality (IR) constraints. [89] also employs contract to investigate the cooperative spectrum sharing problem, where the primary user is the monopolist trading available spectrum with the secondary users in exchange of the secondary users' assistance to cooperatively transmit data. Besides the literature in cognitive radio, contract-theoretic approaches are also adopted in other areas, e.g., distributed computing. [89] uses contract theory for the client (Google, as the example in [90]) to offload portions of computing tasks to the smartphone users (as contract buyers). [91] discusses the application of contract theory to allocating bandwidth in a geo-location information brokerage system.

Chapter 3

Dynamic Relay Selection for Mobile Users and Price Competition for Service Providers in Wireless Relay Networks

In this chapter, the author proposes a hierarchical framework of dynamic relay selection for mobile users and price competition for service providers in wireless relay networks (WRNs), with the purpose of benefiting both the mobile users and the service providers. With detailed research background provided in Section 1.2, the author starts this chapter from introducing the system model in Section 3.1. The hierarchical framework of dynamic relay selection for mobile users and price competition for service providers is introduced in Section 3.2. The stackelberg-game formulation is decomposed into a lower-level problem and an upper-level problem, which are discussed in detail in this section. Section 3.3 presents the simulation results of the proposed dynamic relay selection policy and the comparison with other relay selection schemes.

This is the lower-level problem in the Stackelberg-game framework. The Nash equilibrium of the pricing game for the service providers is also shown and discussed in this section. This is the higher-level problem in the Stackelberg-game framework. The summary of this chapter is given in Section 3.4.

3.1 System Model

3.1.1 Networking Scenario

As shown in Fig. 3.1, the author considers a WRN which consists of a base station (BS) and K fixed relay stations (RSs), i.e., RS_1, RS_2, \dots, RS_K , geographically located in the coverage area of the BS. Each RS adopts the DF strategy, and is equipped with multiple channels. There are multiple mobile users in the WRN, acting as the traffic sources. Considering uplink transmission from mobile user to BS (note that the same model can be applied for downlink transmission), when an RS is selected by a mobile user, the RS will utilize one channel to help the mobile user relay packets so long as there is an available channel. Through this mechanism, multiple mobile users can transmit data simultaneously in different channels without interference. A mobile user who is trying to access the system when no idle channel exists will be rejected by an admission control mechanism in place. However, the admission control problem is beyond the scope of this work. Without loss of generality, it is assumed that the mobile users in the system are treated equally, and each of them can independently adopt the optimal relay selection policy. For presentation clarity, the author discusses the relay selection procedure for a single mobile user.

Locations of the mobile user are discretized in a service coverage area, and let $\mathbb{L} = \{L_1, L_2, \dots, L_M\}$ be the set of all possible locations of the mobile user, where

$M = |\mathbb{L}|$. Assume that the mobile user has the information of available RSs at each location. Here the available RSs are referred to those which have reliable source-to-relay links, and thus can correctly decode the signals received from the mobile user. Let $\mathbb{R}_m = \{RS_m^1, RS_m^2, \dots, RS_m^{K_m}\}$ denote the set of available RSs for location L_m , where $K_m = |\mathbb{R}_m|$. Based on such information, the mobile user can select the relay stations from \mathbb{R}_m to help its transmission. It is assumed that whenever an RS is selected, the RS will spend a fixed power on relaying the user's data, which introduces a fixed monetary cost to be paid by the user in return for obtaining the relay transmission. In practice, different RSs may belong to different service providers, and the price of selecting the RSs is determined by the corresponding service providers.

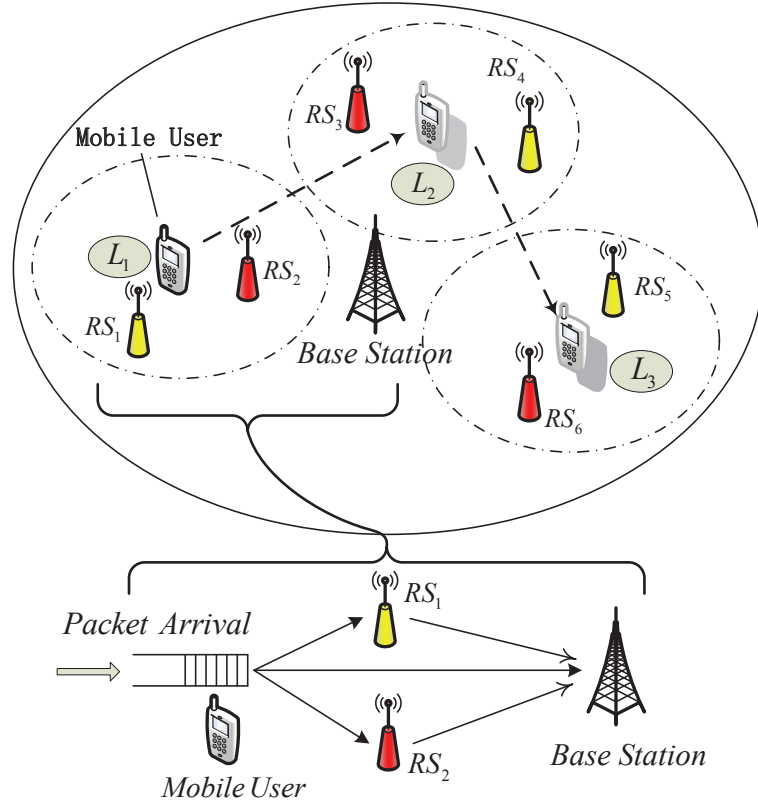


Figure 3.1: The wireless relay network.

3.1.2 Mobility Model

Consider a long observation time which is partitioned into multiple decision periods of fixed duration (e.g., time slot). At the beginning of each decision period, the mobile user makes a decision on relay selection. It is assumed that the user location may vary over different decision periods but remains unchanged during one decision period. Given its current location L_{m_1} , the mobile user will either stay at the same location or move to different location $L_{m_2} \in \mathcal{N}(L_{m_1})$ in the next decision period, where $\mathcal{N}(L_{m_1})$ denotes the set of neighboring locations of L_{m_1} . Assume that the time taken by the mobile user to move from one location to a neighboring location is negligible when compared with the length of one decision period. The user mobility can be modeled using an $M \times M$ transition matrix \mathbf{P}_m , which is shown as follows:

$$\mathbf{P}_m = \begin{bmatrix} p_{1,1} & p_{1,2} & 0 & \cdots & \cdots & 0 & p_{1,M} \\ p_{2,1} & p_{2,2} & p_{2,3} & 0 & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & 0 & \cdots & 0 \\ \vdots & & p_{m,m-1} & p_{m,m} & p_{m,m+1} & & 0 \\ \vdots & & & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & p_{M-1,M-2} & p_{M-1,M-1} & p_{M-1,M} \\ p_{M,1} & 0 & \cdots & \cdots & 0 & p_{M,M-1} & p_{M,M} \end{bmatrix} \quad (3.1)$$

In \mathbf{P}_m , the element at the m_1^{th} row and m_2^{th} column (i.e., p_{m_1,m_2}) denotes the probability for the mobile user to be at location L_{m_2} in the next decision period, under the condition that its location is L_{m_1} in the current decision period. When $L_{m_2} \notin \mathcal{N}(L_{m_1})$, the transition probability $p_{m_1,m_2} = 0$, note that we have $\sum_{m_2: L_{m_2} \in \mathcal{N}(L_{m_1})} p_{m_1,m_2} = 1$. Similar mobility model can be found in [92], [93].

3.1.3 Relay Transmission

For relay transmission, the protocol II in [94] is applied due to its advantage in battery life efficiency. In this protocol, the transmission is divided into two phases. In the first phase, the user transmits data, and this data is received by both the RSs and the BS. In the second phase, all the selected RSs forward the decoded data to the BS simultaneously, while the user remains silent. Consider a general case that K RSs are selected to relay for the mobile user, assuming that the maximal ratio combining (MRC) is applied at the BS, then the post-processing SINR at the destination is [95]

$$\gamma_p^{DF} = \gamma_{SD} + \sum_{k=1}^K \gamma_{R_k D} \quad (3.2)$$

where γ_{SD} and $\gamma_{R_k D}$ represent the instantaneous SINR of the source-to-destination (SD) channel and the channel between relay RS_k and the BS, respectively. Note that $K = 1$ in (3.2) corresponds to a special case that only one relay is selected. For a Rayleigh fading channel, the cumulative distribution function (CDF) of the post-processing SINR for no-relay (i.e., direct transmission) case is given by

$$F_d(\gamma) = 1 - e^{-\frac{\gamma}{\alpha}}, \quad \gamma \geq 0 \quad (3.3)$$

where $\alpha = \overline{\gamma_{SD}}$ is the average SINR of the SD channel. In the context of K relays, the CDF of the post-processing SINR is

$$F_m(\gamma) = \left(\prod_{j=1}^K \frac{\alpha}{\alpha - \beta_j} \right) (1 - e^{-\frac{\gamma}{\alpha}}) + \sum_{k=1}^K \left(\frac{\beta_k}{\beta_k - \alpha} \prod_{j=1, j \neq k}^K \frac{\beta_k}{\beta_k - \beta_j} \right) (1 - e^{-\frac{\gamma}{\beta_k}}) \quad (3.4)$$

where $\gamma \geq 0$, $K \geq 1$, and $\beta_k = \overline{\gamma_{R_k D}}$ ($k = 1, \dots, K$) are the corresponding average SINRs. The derivation of $F_m(\gamma)$ is provided in the following.

For K relays, the post processing SINR can be obtained from $\gamma_p^{\text{DF}} = \gamma_{SD} + \sum_{k=1}^K \gamma_{R_k D}$ when MRC is applied at the destination. Since it is assumed that the wireless channels suffer from Rayleigh fading, the amplitudes of the channel coefficients are random variables following a Rayleigh distribution. Then the square of the amplitudes (i.e., channel gain) has Chi-square distribution with two degrees of freedom, which is equivalent to an exponential distribution. Since the SINR is a linear function of channel gain (i.e., the randomness of SINR is only due to the randomness of channel gain), the SINR also has an exponential distribution.

It is known that if random variable $X = \sum_{k=1}^K X_k$, in which all random variables are *i.i.d.* and each random variable X_k is exponentially distributed with parameter α_k , then their sum will have a hypo-exponential distribution. The probability density function (pdf) of X can be represented as follows:

$$f_X(x) = \sum_{k=1}^K c_k(\alpha) f_{X_k}(x)$$

where

$$c_k(\alpha) = \prod_{j=1, j \neq k}^K \frac{\alpha_j}{\alpha_j - \alpha_k}, f_{X_k}(x) = \alpha_k e^{-\alpha_k x}.$$

Let $\gamma_0 = \gamma_{SD}$, $\gamma_k = \gamma_{R_k D}$ ($k = 1, \dots, K$), and $\alpha_k = \frac{1}{\gamma_k}$, then the pdf of γ_0 is given by

$$f_{\gamma_0}(\gamma) = \frac{1}{\gamma_{SD}} e^{-\frac{\gamma}{\gamma_{SD}}}. \quad (3.5)$$

The pdf of γ_k ($k = 1, \dots, K$) is

$$f_{\gamma_k}(\gamma) = \frac{1}{\gamma_{R_k D}} e^{-\frac{\gamma}{\gamma_{R_k D}}}. \quad (3.6)$$

Since $\gamma = \sum_{k=0}^{K_m} \gamma_k$ has a hypo-exponential distribution, the pdf of γ can be denoted as follows:

$$\begin{aligned} f_{\Gamma}(\gamma) &= \sum_{k=0}^K \left(\prod_{j=0, j \neq k}^K \frac{\frac{1}{\gamma_j}}{\frac{1}{\gamma_j} - \frac{1}{\gamma_k}} \right) f_{\gamma_k}(\gamma) \\ &= \left(\prod_{j=1}^K \frac{\overline{\gamma_{SD}}}{\overline{\gamma_{SD}} - \overline{\gamma_{R_j D}}} \right) \frac{1}{\overline{\gamma_{SD}}} e^{-\frac{\gamma}{\overline{\gamma_{SD}}}} + \sum_{k=1}^K \left(\left(\prod_{j=0, j \neq k}^K \frac{\overline{\gamma_{R_k D}}}{\overline{\gamma_{R_k D}} - \overline{\gamma_{R_j D}}} \right) \frac{1}{\overline{\gamma_{R_k D}}} e^{-\frac{\gamma}{\overline{\gamma_{R_k D}}}} \right) \end{aligned}$$

from which the CDF of the post processing SINR can be obtained by integrating the pdf calculated above.

For convenience of analysis, the author considers the data transmission from a source to a destination through both direct link and relay links as achieved on an equivalent virtual link. For the DF based relay transmission, the choice of adaptive modulation and coding (AMC) mode only depends on the post-processing SINR at the corresponding destination [95], which is calculated from (3.2). Given γ_p^{DF} , the maximum achievable rate (MAR) of the virtual link, which is denoted by r^* can be obtained (It needs to be emphasized here that in the model, the rate is measured by transmitted packets per decision period). With this r^* , the AMC mode can be chosen accordingly. For instance, if $r^* = 2$, the AMC of QPSK with code rate of 1 or 16QAM with code rate of 1/2 can be used. In this work, the author assumes that a mechanism for information exchange among mobile user, RSs, and BS is in place to facilitate the cooperative communication (the design of such a mechanism is beyond the scope of this work). With the communication mechanism, the destination is capable to feed back its SINR measurement to the source, based on which the source selects a proper AMC mode for transmission and informs the selected relays to use the same mode for transmission. Note that in the two-phase cooperative transmission, the source only transmits in the first phase and, hence, the effective end-to-end rate (EER) (i.e., the

actual throughput at the destination) is $r = \frac{r^*}{2}$.

Let $\mathfrak{r} = \{r_0, r_1, \dots, r_R\}$ denote the set of the EERs corresponding to different AMC modes. Without loss of generality, consider $r_0 < r_1 < \dots < r_R$. If the SINR of the channel is within the SINR interval $[\Gamma_{r_i}, \Gamma_{r_{i+1}})$ (the SINR thresholds Γ_{r_i} and $\Gamma_{r_{i+1}}$ are measured in dB), EER r_i can be used to transmit the user's data. Given the required SINR interval $[\Gamma_{r_i}, \Gamma_{r_{i+1}})$, the probability that EER is r_i when selecting K RSs, can be calculated by:

$$P_{r_i, K} = \begin{cases} F_d(\Gamma_{r_{i+1}}) - F_d(\Gamma_{r_i}), & K = 0 \\ F_m(\Gamma_{r_{i+1}}) - F_m(\Gamma_{r_i}), & K \geq 1 \end{cases} \quad (3.7)$$

where $\Gamma_{r_0} = -\infty$ and $\Gamma_{r_{R+1}} = \infty$.

3.2 Hierarchical Framework of Dynamic Relay Selection for Mobile Users and Price Competition for Service Providers

Since the author jointly investigates the relay selection problem for mobile users and the price competition problem for service providers, an effective solution method is of significant importance. In this section, the author proposes a hierarchical framework as shown in Fig. 3.2. This hierarchical framework is Stackelberg-game framework with two levels. At the upper level, the service providers, considered as leaders, can set the relay prices first, while the mobile users, considered as followers at the lower level, will determine their relay selection policies based on the observation of the relay prices. The author defines the *demand* for RS_k as the ratio of the number of times that the

mobile user requests for RS_k to the total number of decision periods. The information exchange between these two levels consists of the price of the relay services (from the upper level to the lower level) and demand of the relays (from the lower level to the upper level). As the leader, the service providers can strategically set relay prices to maximize their own profits, given the demand of relays. The lower-level problem and upper-level problem are discussed in Section 3.2.1 and Section 3.2.2, respectively.

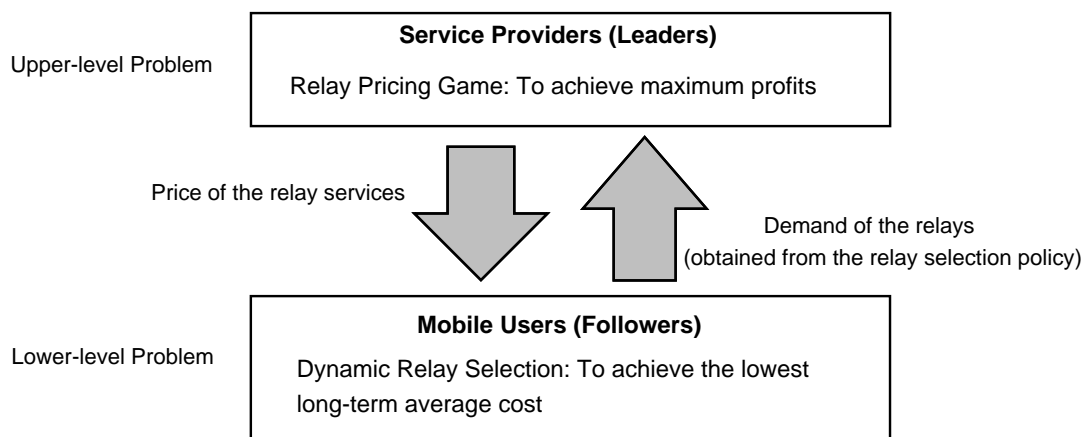


Figure 3.2: The proposed hierarchical framework of the Stackelberg-fashion problem.

3.2.1 Dynamic Relay Selection

First, the author investigates the lower-level problem, which can be described as follows: given the prices of all RSs, how can the mobile user select the most beneficial relays, such that its long-term average cost can be minimized. The significance of minimizing the user cost includes both the aspects of reducing payment to the relays and decreasing packet loss probability. The user cost is formally defined as follows. In one decision period, when the user makes relay selection, a monetary payment has to be made to the relay(s) in return for requiring the relay services. Such a cost is dependent on the price settings by the service providers. For instance, if the user selects a relay (with price p) of service provider 1, then service provider 1 charges the user p for using its

relay in this decision period. In the future periods, different relays may be selected and the user cost can be calculated accordingly. Also, the packet loss can be obtained at the end of one arbitrary decision period, when the user has made relay selection. As a result of the relay selection, the queue length of the user changes, which will affect a new round of relay selection in the next decision period. In this case, the user's decision on relay selection not only affects the current decision period, but also has impact on the future decision periods. Hence, the best relay selection in the current decision period may not be optimal in the long run. Thus, the author attempts to design the most beneficial policy considering a finite time horizon consisting of a large number of decision periods, rather than minimizing the cost for each decision period. In the following, the author proposes a dynamic relay selection scheme for mobile users by formulating an optimization model based on the Markov decision process (MDP), with the objective of minimizing the mobile user's long-term average cost. In the following, the state and action spaces, as well as the transition probability matrix of the MDP model, are defined. Then, the method to obtain the optimal policy of the MDP problem is presented.

State and Action Spaces

The state space of the mobile user is defined as $\Omega = \{(\mathcal{L}, \mathcal{Q})\}$, where $\mathcal{L} \in \{L_1, L_2, \dots, L_M\}$ and $\mathcal{Q} \in \{0, 1, \dots, Q_B\}$ represent the location of mobile user and the number of packets in the mobile user's buffer (i.e., queue length), respectively. Here Q_B is the buffer size of mobile user.

The action space of the mobile user represents the available choices of relay selection, and is denoted as

$$\mathcal{A} = \{\mathcal{A}_1, \dots, \mathcal{A}_m, \dots, \mathcal{A}_M\}$$

where \mathcal{A}_m is the available action set at location L_m . The total number of actions at location L_m is $|\mathcal{A}_m| = \sum_{j=0}^{K_m} \binom{K_m}{j} = 2^{K_m}$, which increases exponentially with the number of the available relays (i.e., K_m).

Transition Probability Matrix

As the state of the mobile user consists of both its location and its buffer occupancy, the state transition depends on location transition and/or queue length transition. Location transition can be described by transition probability matrix \mathbf{P}_m defined in Section 3.1.2. For the queue length transition, assume that packets arrive in batches and the probability for n_a packets arriving in one decision period is λ_{n_a} . Moreover, at most N_a packets can arrive in one decision period. For packet departures, let $\mathbb{N}_d = \{n_{d_0}, n_{d_1}, \dots, n_{d_R}\}$ denote the set of all possible numbers of departing packets in a decision period, where n_{d_i} corresponds to the EER rate $r_i \in \mathbf{r}$ (i.e., when the cooperative transmission achieves EER rate r_i , maximum n_{d_i} packets depart the queue of the mobile user in a decision period). Let $D^{n_d}(a)$ denote the probability of n_d packets departing from the data queue of the mobile user when action a is taken. It can be calculated as follows.

1. When no relay is selected (direct transmission),

$$D^{n_d}(a) = \begin{cases} P_{r_i,0}(\overline{\gamma_{SD}}), & n_d = n_{d_i} \in \mathbb{N}_d \\ 0, & \text{otherwise} \end{cases}$$

2. When K ($K \geq 1$) relays are selected,

$$D^{n_d}(a) = \begin{cases} P_{r_i, K}(\overline{\gamma_{SD}}, \overline{\gamma_{R_1D}}, \dots, \overline{\gamma_{R_KD}}), & n_d = n_{d_i} \in \mathbb{N}_d \\ 0, & \text{otherwise} \end{cases}$$

where the probabilities $P_{r_i, k}(\cdot)$, $k = 0$ or K , for all cases can be calculated as in (3.7). Denote the transition probability matrix for queue length as $\mathbf{P}_{Q_b}(a)$ whose element $P_{Q_b, Q_{b'}}(a)$ represents the probability for the data queue length to change from Q_b to $Q_{b'}$ given that action a is taken, which can be obtained by

$$P_{Q_b, Q_{b'}}(a) = \sum_{\{(n_a, n_d) | Q_b + n_a - n_d = Q_{b'}\}} \lambda_{n_a} D^{n_d}(a). \quad (3.8)$$

Based on the transition probability matrices \mathbf{P}_m and $\mathbf{P}_{Q_b}(a)$, the state transition probability matrix $\mathbf{P}_{s, s'}(a)$, consisting of the probabilities for the mobile user to change from state $s \in \Omega$ to $s' \in \Omega$ when action a is taken, can be obtained from

$$\mathbf{P}_{s, s'}(a) = \mathbf{P}_m \otimes \mathbf{P}_{Q_b}(a) \quad (3.9)$$

where \otimes denotes the Kronecker product.

MDP Problem Formulation

Consider two components in the mobile user's cost function (which needs to be minimized on a long term basis): (1) the payment to the selected relays, and (2) the QoS performance in terms of packet loss suffered by the mobile user. Define $\mathbf{X}(\pi) = [\mathcal{C}(\pi), \mathcal{P}_L(\pi)]^T$ as the performance vector which measures the performance of the relay selection policy in terms of the long-term average payment (i.e., $\mathcal{C}(\pi)$) and the probability of packet loss due to lack of buffer space (i.e., $\mathcal{P}_L(\pi)$). Here π

denotes the relay selection policy, which is a mapping of state $s \in \Omega$ to action $a \in \mathcal{A}$, i.e., $\pi = \{\Psi(s, a) | s \in \Omega, a \in \mathcal{A}\}$, where $\Psi(s, a)$ is the probability of taking action a when the mobile user is at state s . Consider a long observation time and let W denote the number of decision periods in the observation time. When mobile user is at state s_t ($t = 1, \dots, W$), it will take action a_t according to relay selection policy π . In the proposed problem formulation, the objective is to minimize the mobile user's long-term average cost given by:

$$\mathcal{U}(\pi) = \mathbf{w}^T \mathbf{X}(\pi) \quad (3.10)$$

where $\mathbf{w} = [w_c, w_l]^T$ is a weighting vector with positive elements w_c and w_l , which are the weights corresponding to the two elements of $\mathbf{X}(\pi)$. The long-term objective of the mobile user can be formulated as an MDP problem given by

$$\min_{\pi} \mathcal{U}(\pi) = \mathbf{w}^T \mathbf{X}(\pi) \quad (3.11)$$

$$= \lim_{W \rightarrow \infty} \sup \frac{1}{W} \sum_{t=1}^W (w_c \cdot E(\mathcal{C}(s_t, a_t)) + w_l \cdot E(\mathcal{P}_L(s_t, a_t))) \quad (3.12)$$

where $E(\cdot)$ denotes the expectation, $\mathcal{C}(s_t, a_t)$ is immediate payment to the selected relay(s) and $\mathcal{P}_L(s_t, a_t)$ is immediate packet loss probability in the t^{th} decision period. Note that $\mathcal{C}(s_t, a_t)$ comes from the prices of the relay services set by the service providers. In other words, the lower-level problem is dependent on the upper-level problem at this point. The expected packet loss probability $E(\mathcal{P}_L(s_t, a_t))$ can be calculated as follows:

$$E(\mathcal{P}_L(s_t, a_t)) = \frac{\sum_{n_{d,t} \in \mathbb{N}_d} \sum_{n_{a,t}=0}^{N_a} P_{L_t} \lambda_{n_{a,t}} D^{n_{d,t}}(a_t)}{\bar{\lambda}} \quad (3.13)$$

where $P_{L_t} = \max(q_{t-1} + n_{a,t} - n_{d,t} - Q_B, 0)$, q_{t-1} is the queue length at the end of the

$(t-1)^{th}$ decision period, $n_{a,t}$ and $n_{d,t}$ are the numbers of packet arrivals and departures in the t^{th} decision period, respectively; $\lambda_{n_{a,t}}$ and $D^{n_{d,t}}(a_t)$ denote the corresponding probabilities. Note that if $q_{t-1} + n_{a,t} - n_{d,t} > Q_B$, $P_{L_t} > 0$, thus, P_{L_t} is the number of dropped packets due to lack of buffer space, when $n_{a,t}$ packets arrive and $n_{d,t}$ packets depart in decision period t , given that the queue length at the end of the last decision period is q_{t-1} . The whole expression in the numerator denotes the expected packet loss, divided by $\bar{\lambda}$ in the denominator (which is the average packet arrival rate, and can be calculated as $\bar{\lambda} = \sum_{n_a=0}^{N_a} n_a \lambda_{n_a}$), the expected packet loss probability can be obtained.

The optimal policy π^* can be obtained by transforming the MDP formulation into an equivalent linear programming (LP) problem [96]. Let $\Phi(s, a)$ denote the probability of the mobile user being in state s and taking action a , the LP problem is given in the following:

$$\begin{aligned}
 \min_{\Phi(s,a)} \quad & \sum_{s \in \Omega} \sum_{a \in \mathcal{A}} [w_c \mathcal{C}(s, a) + w_l \mathcal{P}_L(s, a)] \Phi(s, a) & (3.14) \\
 \text{s. t.} \quad & \sum_{a \in \mathcal{A}} \Phi(s', a) = \sum_{s \in \Omega} \sum_{a \in \mathcal{A}} P(s'|s, a) \Phi(s, a) \\
 & \sum_{s \in \Omega} \sum_{a \in \mathcal{A}} \Phi(s, a) = 1 \\
 & \Phi(s, a) \geq 0, \quad \forall a \in \mathcal{A}
 \end{aligned}$$

where $P(s'|s, a)$ is an element of matrix $\mathbf{P}_{s,s'}(a)$ obtained from (3.9), indicating the probability of the mobile user changing to state $s' \in \Omega$ in the next decision period given the current state s when action a is taken. Denoting the optimal solution of the LP problem as $\Phi^*(s, a)$, the optimal policy of the MDP problem can be calculated as follows [92]:

$$\pi^* = \Psi^*(s, a) = \frac{\Phi^*(s, a)}{\Phi^*(s)} = \frac{\Phi^*(s, a)}{\sum_{a' \in \mathcal{A}} \Phi^*(s, a')}. \quad (3.15)$$

Here $\Psi^*(s, a)$ ($s \in \Omega, a \in \mathcal{A}$) indicates the optimal probability of taking action a when the mobile user is at state s . The optimal policy for all MU states $\Psi_{\pi^*}(\cdot, \cdot)$ can be considered as a lookup table with size $|\Omega| \times \Lambda$, where $\Lambda = \max_{m=1, \dots, M} \mathcal{A}_m$ denotes the maximum number of actions for one location. Note that some table values may naturally be zero since there may be some locations for which $|\mathcal{A}_m| < \Lambda$. Each table element $\Psi_{\pi^*}(s, a)$ ($s \in \Omega, a \in \mathcal{A}$) indicates the optimal probability of taking action a when the MU is at state s . At each decision epoch, the MU determines its state s according to its current location and queue length. Then it follows randomized policy to makes relay selection a according to the corresponding value $\Psi_{\pi^*}(s, a)$.

3.2.2 Game Formulation for Pricing

Next, the author investigates the upper-level problem, which can be described as follows: with the knowledge that how the mobile users will make relay selections based on given prices, how can the service providers set prices for their relay services to win the competition and thus achieve the maximum profits. In the considered model, the RSs owned by the same service provider are assumed to have the same price. A mobile user will always choose the most beneficial RSs. As a result, when the prices of the RSs are changed, the mobile users may make different selections which result in the variation of the RSs' demands, and further fluctuation of the service providers' profits. Therefore, the pricing problem for the service providers can be viewed from a game theoretic perspective.

For service provider i , the choice of the optimal price p_i^* cannot be considered solely by service provider i , since it is dependent on not only its RSs' channel qualities but also the prices set by the other service providers. In this case, the apparently expensive RSs may unlikely be selected, resulting in the benefit loss of the corresponding

service provider. On the other hand, if one service provider chooses too low price to gain the market share, it will also fail in achieving high utility. Overall, there is a tradeoff for setting the price for all service providers, and the service providers have to select the price such that a stable market share and the best utility is achieved. Let K_s and Z denote the number of the service providers and active mobile users in the WRN, respectively. Note that the upper-level problem links with the lower-level problem discussed in Section 3.2.1 in two aspects: (1) the prices of the RSs, and (2) the corresponding demands based on the given prices of the RSs. The utility of service provider i is defined as follows:

$$U_i(p_i, \mathbf{p}_{-i}) = p_i \cdot d_i(p_i, \mathbf{p}_{-i}) \quad (3.16)$$

where p_i denotes the price of the RSs belonging to service provider i , $\mathbf{p}_{-i} = [p_1, \dots, p_{i-1}, p_{i+1}, \dots, p_{K_s}]^T$ is a vector of prices of all service providers except service provider i , and $d_i(p_i, \mathbf{p}_{-i})$ denotes the sum of the demands of all RSs owned by service provider i from the Z mobile users. In the upper-level game, each service provider aims at maximizing its own utility. The optimization problem for service provider i can be formulated as follows:

$$\begin{aligned} \max_{p_i} \quad & U_i(p_i, \mathbf{p}_{-i}) \\ \text{s. t.} \quad & 0 \leq p_i \leq p_{max} \end{aligned} \quad (3.17)$$

where p_{max} is the maximum price acceptable by the mobile users.

Note that the most challenging issue in Problem (3.17) is that, besides price p_i , the utility function of service provider i also depends on its demand $d_i(p_i, \mathbf{p}_{-i})$, which is

obtained from the solutions of the lower-level MDP problem. Thus it has no closed-form expression for the solution. As an alternative, the author attempts to solve this problem numerically. Nash equilibrium is defined as the solution of the non-cooperative game, denoted by $\{p_1^*, p_2^*, \dots, p_{K_s}^*\}$, at which service provider i ($i \in \{1, 2, \dots, K_s\}$) fixes on a certain price p_i^* so that its utility is maximized. At this point, each service provider has made the best response according to the other service providers' price settings, i.e.,

$$p_i^* = \mathcal{B}_i(\mathbf{p}_{-i}^*), \forall i \in \{1, \dots, K_s\} \quad (3.18)$$

where $\mathbf{p}_{-i}^* = [p_1^*, \dots, p_{i-1}^*, p_{i+1}^*, \dots, p_{K_s}^*]^T$ denotes the vector of best responses of all service providers except service provider i . The best response of service provider i ($i \in \{1, \dots, K_s\}$) is defined as follows:

$$\mathcal{B}_i(\mathbf{p}_{-i}) = \arg \max_{p_i} U_i(p_i, \mathbf{p}_{-i}). \quad (3.19)$$

The Nash equilibrium is a set of strategies with the property that no service provider can improve its profit by increasing or decreasing the prices while the other service providers keep their strategies unchanged. To obtain the Nash equilibrium, an iterative algorithm similar to that in [97] can be applied, which is shown in the following:

3.2.3 Architecture of the Proposed Algorithm

In this section, the author proposes the architecture of the proposed algorithm to solve the hierarchical framework, which is shown in the following:

Note that the stable price settings refer to the situation that each service provider has made its best response, given the price settings of the other service providers. In other words, none of the service providers can improve its utility by updating its price

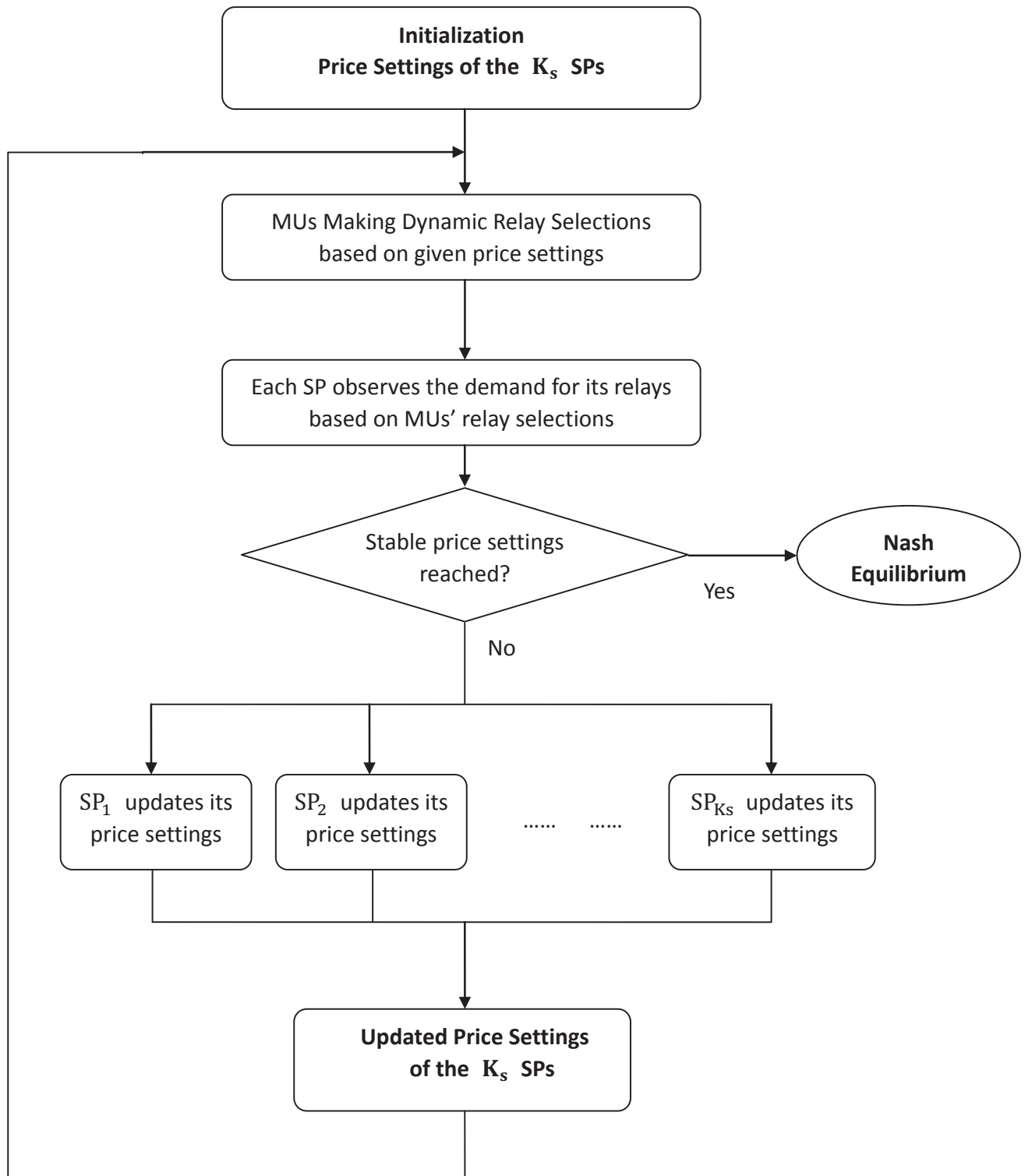


Figure 3.3: Flowchart of the proposed algorithm to solve the hierarchical framework.

Algorithm 1 Iterative algorithm to obtain Nash equilibrium of the price competition problem.

```

1: Initialize  $p_i(k)$ ,  $\forall i$  for  $k = 0$ 
2: loop
3:   repeat
4:      $k \leftarrow k + 1$ 
5:     for  $i = \{1, \dots, K_s\}$  do
6:        $p_i(k) \leftarrow \mathcal{B}_i(\mathbf{p}_{-i}(k))$ 
7:     end for
8:   until ( $\max_i |p_i(k) - p_i(k-1)| < \epsilon$ ) or ( $k > K_{max}$ )
9: end loop
10: return  $p_i^* \leftarrow p_i(k)$ 

```

settings unilaterally.

3.3 Performance Evaluation

3.3.1 Scenario and Parameter Settings

As illustrated in Fig. 3.1, here the author considers 3 locations, RS_1 and RS_2 are in location L_1 , RS_3 and RS_4 are in location L_2 , while RS_5 and RS_6 are in location L_3 . Among all RSs, RS_2 , RS_3 and RS_6 belong to service provider 1, while RS_1 , RS_4 and RS_5 are owned by service provider 2. To investigate the relay selection problem with given relay prices for the mobile users, let all RSs have the same price, which is set to be 10, unless otherwise mentioned. Consider the mobile user with a uniform mobility pattern. Specifically, when the mobile user is at location L_m ($m = 1, 2, 3$) in one decision period, the mobile user has equal probabilities to stay at location L_m or to move to the neighboring locations in the next decision period. The buffer size for the mobile user is 15 packets. There are four AMC modes with SINR thresholds $\Gamma_1 = 6.4$ dB, $\Gamma_2 = 9.4$ dB, $\Gamma_3 = 11.2$ dB, and $\Gamma_4 = 16.4$ dB [93], resulting in $\mathbb{N}_d = \{0, 2, 4, 6, 8\}$ packets/decision period. The packet arrival is assumed to be a

Poisson process with average arrival rate $\bar{\lambda} = 2.5$ packets/decision period. The author implements the proposed optimal relay selection policy in the simulation which consists of 200000 decision periods.

3.3.2 Performance Evaluation of the Relay Selection

The effect of weight variation

First, the author investigates the impact of different weight settings of user cost function on the performance metrics and relay selection results. Different settings reflect different emphases on packet loss performance and payment. For ease of discussion, w_c is fixed to be 1 and increase w_l from 10 to 120. The variations of average payment, average packet loss probability, and probability of selecting different RSs with w_l are shown in Figs. 3.4 (a), (b) and (c), respectively. It can be observed that when w_l is small (i.e., $w_l = 10$), direct transmission is always used and no relay service is needed at all, resulting in zero average payment. With the increase of w_l , the long-term average cost can be minimized by reducing the packet loss probability, therefore, the relay transmission becomes necessary, resulting in the payment increase. Also, all performance metrics and relay selection results change dramatically when w_l increases from 10 to 30. However, when w_l further increases (i.e., from 30 to 120), such variation becomes inconspicuous. The reason is that, when $w_l = 30$, the packet loss probability is already small (i.e., 0.02), and further increase of w_l cannot stimulate the mobile user to pursue even lower packet loss. Consequently, setting a large value of w_l does not have significant impact on relay selection strategy of the mobile user, as shown in Fig. 3.4 (c). According to this observation, the author fixes $w_l = 30$ in the rest of evaluation.

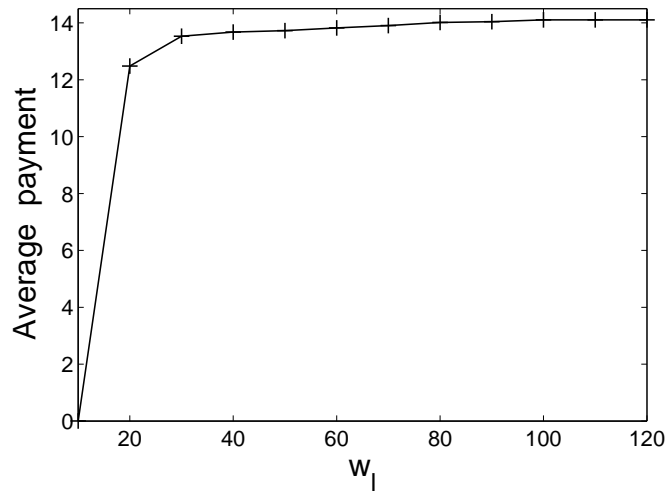
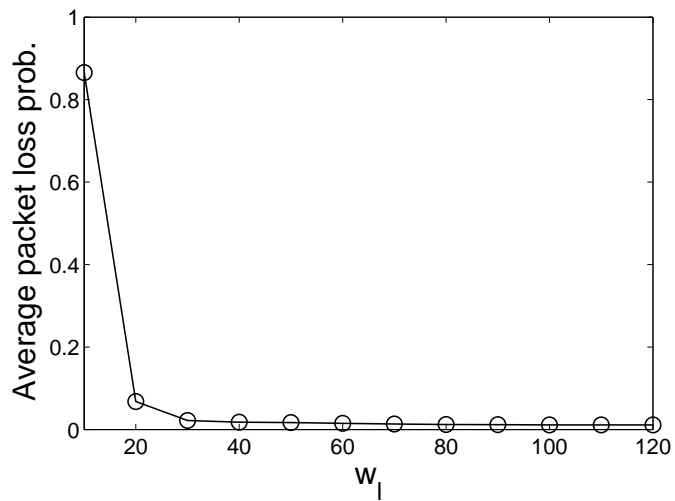
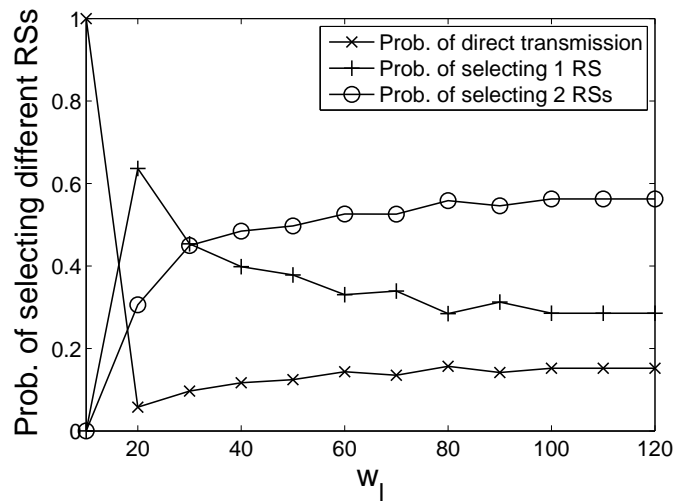
(a) Average payment given w_l (b) Average packet loss probability given w_l (c) Probability of selecting different RSs given w_l

Figure 3.4: The effect of weight variation.

The effect of queue length

To study the effect of queue length (i.e., the number of packets in queue) variation on the relay selection results, the author considers three average packet arrival rates: $\bar{\lambda} = 1, 2$ and 3 packets/decision period. Fig. 3.5 shows the variation in the number of the selected RSs in location L_1 with different queue length. Based on the result, the author has two observations:

1. For a certain value of $\bar{\lambda}$, RSs are rarely needed when the queue is nearly empty. However, the number of selected RSs increases when the queue length exceeds some threshold, and achieves the maximum value when the queue is nearly full as expected.
2. For a certain queue length, the number of selected RSs may increase as $\bar{\lambda}$ increases. For instance, when the queue length is 8 packets, direct transmission yields sufficient departure rate for $\bar{\lambda} = 1$ packet/decision period. However, for $\bar{\lambda} = 2$ and 3 packets/decision period, one RS and two RSs are needed, respectively, to meet the QoS of mobile user.

The effect of the maximum queue size

In this section, the author studies the effect of the maximum queue size variation on relay selection. $\bar{\lambda}$ is set to be 2.5 packets/decision period, and increase the queue buffer size from 5 to 50 packets. As shown in Fig. 3.6, the probability of selecting two RSs decreases when the maximum queue size becomes larger, while that of selecting one RS increases. The reason is that, as the maximum queue size increases, there is more available space to buffer the arriving packets. Therefore, it is not necessary to transmit them immediately to avoid unacceptably high packet loss.

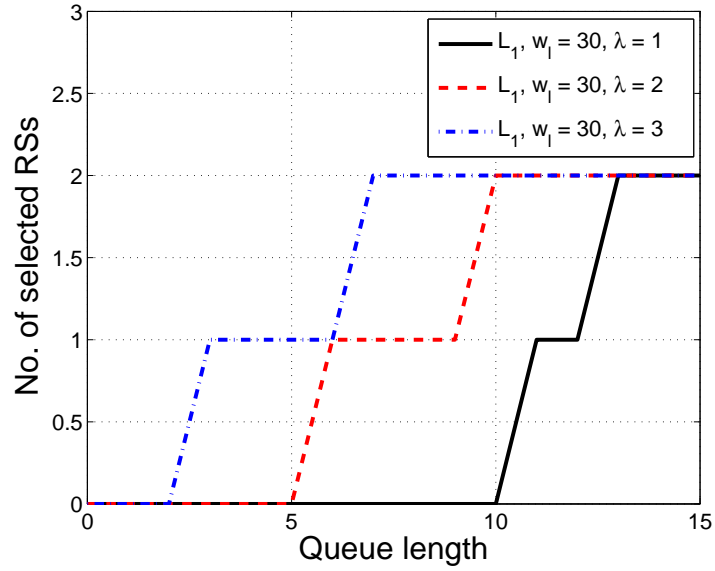


Figure 3.5: The number of selected RSs under different queue length.

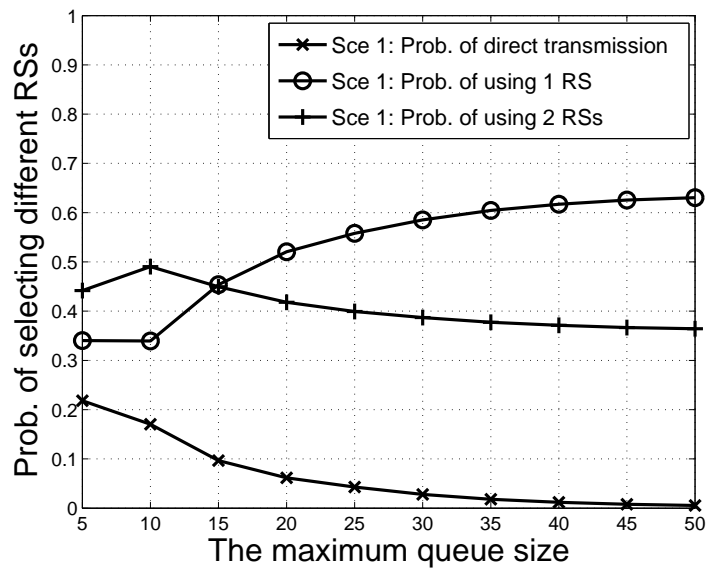


Figure 3.6: Probability of selecting different RSs under different maximum queue size in both scenarios 1 and 2.

The effect of price variation

The author varies the price of RS_1 from 0 to 20, and illustrates the variations of the demands of all RSs at location L_1 (i.e., RS_1 and RS_2) in Fig. 3.7. From Fig. 3.7, the author observes that RS_2 is always the preferred choice regardless of the price of RS_1 . This is because RS_2 has better SINR in location L_1 , thus can bring higher QoS improvement to the mobile user through cooperation. The demand of RS_1 drops gradually and becomes zero when the price increases, demonstrating that the proposed MDP scheme can effectively adapt the relay selection to the price variation.

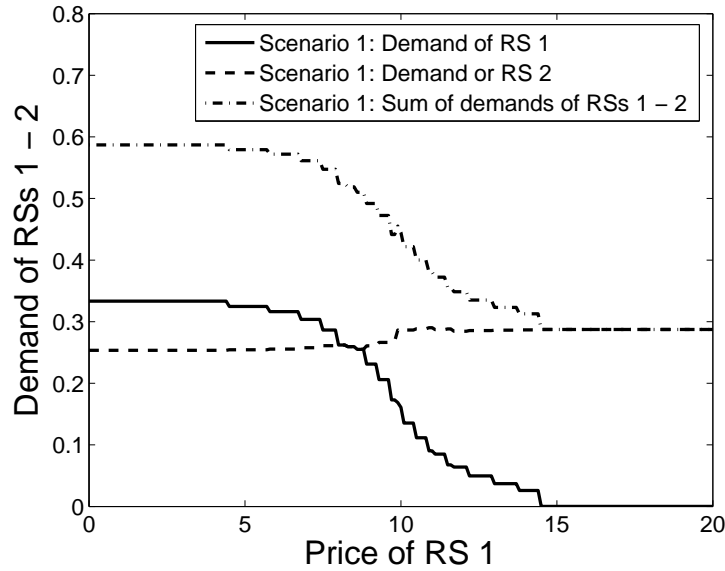


Figure 3.7: The demand of RSs 1 and 2 under different price of RS 1.

Performance Comparison

In this section, the author compares the performance of the proposed MDP scheme with that of other relay selection methods. Here the author chooses the following relay selection schemes for comparison: (1) direct transmission; (2) single-relay selection with the best SINR, and (3) always uses two relays. The performance metric in terms

of average user cost with the cases $w_l = 30$ and $w_l = 60$ is evaluated and compared in Fig. 3.8. It can be observed that the proposed MDP scheme can always achieve the lowest long-term average cost among the four relay selection schemes under both weight settings. Note that the proposed scheme will not sacrifice more price (measured by the payment) to achieve an over-satisfied QoS performance. These comparison results can demonstrate the effectiveness of the proposed scheme in minimizing the mobile user's long-term average cost.

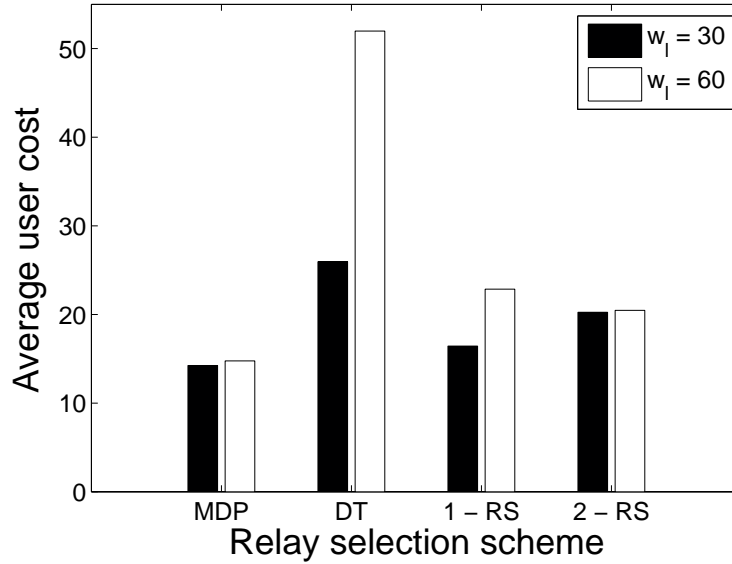


Figure 3.8: Comparison of the average user cost.

3.3.3 Numerical Results of the Relay Pricing

The Best Response Strategy

To reach the Nash equilibrium, each service provider makes the best response to the given prices of the other service providers. Let $w_l = 30$, suppose the decision of service provider 1 has been given such that $p_1 = 9.5$, the author investigates how

service provider 2 will determine an optimal price p_2^* based on p_1 . This is illustrated in Figs. 3.9(a) and (b) which show the variation of both demand and utility with the price of service provider 2, respectively.

As shown in Fig. 3.9(a), the demand of service provider 2 decreases from 10 to 0 when the price of service provider 2 increases from 0 to 20, while the corresponding utility of service provider 2 increases first and then decreases (Fig. 3.9(b)). Note that the turning point is at price 9.7 (i.e., the point marked by the small circle), which is the optimal price for service provider 2. If service provider 2 chooses a price higher than 9.7, the mobile users may prefer to select relay services of service provider 1. Alternatively, if service provider 2 attempts to set a lower price to attract more users, its utility also decreases. The best response strategy of service provider 1 given the decision of service provider 2 can be illustrated in the same way.

Nash Equilibrium

In this section, the Nash equilibrium of the relay pricing game for the service providers is presented. The maximum price is set to be 20. There are 10 mobile users in the wireless relay network. Without loss of generality, the author assumes that different mobile users have different mobility patterns which are shown in the following:

- **Mobility pattern 1** M_1 : Uniform mobility as stated before.
- **Mobility pattern 2** M_2 :

$$\mathbf{P}_m = \begin{bmatrix} 2/3 & 1/6 & 1/6 \\ 2/3 & 1/6 & 1/6 \\ 2/3 & 1/6 & 1/6 \end{bmatrix}$$



(a) Demand of service provider 2 versus the price



(b) Profit of service provider 2 versus the price

Figure 3.9: Illustration of service provider 2's best response strategy, given $p_1 = 9.5$.

- **Mobility pattern 3** M_3 :

$$\mathbf{P}_m = \begin{bmatrix} 0.5 & 0.3 & 0.2 \\ 0.1 & 0.6 & 0.3 \\ 0.4 & 0.1 & 0.5 \end{bmatrix}$$

To be representative, the author uses mobility pattern 1 to denote the case that some mobile users have uniform mobility. While mobility pattern 2 represents the case that some mobile users may have a larger probability to stay in one location, compared with that of staying in other locations. Mobility pattern 3 represents the case that some mobile users may have random mobility. Suppose that mobile users 1-3 have mobility pattern M_1 , while mobile users 4-6 and mobile users 7-10 have mobility patterns M_2 and M_3 , respectively. The author varies w_l from 20 to 50, and the Nash equilibria for the prices of service provider 1 and service provider 2 under different weight settings are obtained and compared in Table 3.1.

Note that under different weight settings, the price of service provider 1 is lower than that of service provider 2 at the Nash equilibrium, which is due to the better average channel qualities of the RSs belonging to service provider 2. As w_l increases from 20 to 50, the prices of both service provider 1 and service provider 2 at the Nash equilibrium increase, so are the utilities. The results are as expected since the increase in w_l indicates that the packet loss performance has more impact on the user utility, therefore it is necessary for the mobile users to select more RSs to improve the transmission quality. Based on this fact, the service providers can increase the prices of their RSs to achieve higher profit.

Table 3.1: Nash equilibrium comparison of the SPs under different w_l

Nash equilibrium	$w_l = 20$	$w_l = 30$	$w_l = 40$	$w_l = 50$
service provider 1	$p_1^* = 6.4$	$p_1^* = 9.5$	$p_1^* = 12.8$	$p_1^* = 16.0$
	$U_1^* = 60.99$	$U_1^* = 90.83$	$U_1^* = 121.99$	$U_1^* = 153.04$
service provider 2	$p_2^* = 6.5$	$p_2^* = 9.7$	$p_2^* = 13$	$p_2^* = 16.4$
	$U_2^* = 63.17$	$U_2^* = 94.20$	$U_2^* = 126.51$	$U_2^* = 158.63$

3.4 Summary

A hierarchical framework of dynamic relay selection for mobile users and price competition for service providers has been proposed in this work. To address the joint optimization, the author decomposes the Stackelberg-game formulation into a lower-level problem and an upper-level problem. At the lower-level, the dynamic relay selection problem with given relay prices from the upper-level problem is investigated, taking user's mobility into account. An MDP based optimization model has been formulated and solved to obtain the optimal relay selection policy, aiming at minimizing the user's long-term average cost. The simulation results have clearly shown the effectiveness and flexibility of the proposed scheme in balancing the payment and required QoS performance, as compared with several other relay selection schemes. At the upper-level, techniques from non-cooperative game theory are applied to design pricing strategies for the service providers to maximize their profits, which depend on both the prices of the relay services and demands of the mobile users from the lower-level problem. Nash equilibrium has been obtained as a solution, which can provide a guidance for service providers to compete for providing relay services.

Chapter 4

Optimal Power Allocation for Secondary Users in Cognitive Relay Networks

In this chapter, the author focuses on cognitive relay networks where the secondary users (SUs) are involved as cooperative relays in a primary user (PU)'s data transmission, and propose optimal power allocation schemes for both single-SU and multiple-SU cases, aiming at providing a guidance to the SUs to strategically allocate their relay power and transmission power to achieve the highest benefit. The author starts this chapter from describing the proposed system model in Section 4.1. The power allocation schemes for both single-SU and multiple-SU cases are discussed in Section 4.2, and the corresponding analysis is also presented. Section 4.3 presents the numerical results, followed by Section 4.4, where the summary of this chapter is provided.

4.1 System Model

4.1.1 Network Scenario

As shown in Fig. 4.1, the author considers a cognitive relay network consisting of one primary transmitter (PT) - primary receiver (PR) pair and multiple secondary transmitter (ST) - secondary receiver (SR) pairs equipped with CRs (Note that throughout this paper, PU and PT are used interchangeably, so is the case with SU and ST). The SUs can access the channel of the PU only when the PU has no data to transmit (i.e., the corresponding data queue is empty). In this case, a rational SU should expect the PU's queue to be empty as soon as possible, which may encourage SU to use some power on relaying PU's data. Let s and d denote the PT and PR, respectively, and r_i represent SU_i for ease of presentation. Quasi-static channels are considered in the proposed model. The channels between nodes are assumed to be independent proper complex Gaussian random variables [54], which are considered to be constant within a period of time, but generally varying over different time slots. $h_{s,d}$ denotes the complex channel coefficient between PT and PR. h_{s,r_i} and $h_{r_i,d}$ denote the channel coefficient between PT and SU_i , SU_i and PR, respectively. $h_{r_i,i}$ denotes the channel coefficient between ST_i and SR_i .

In the proposed model, each SU operates in a Decode-and-Forward (DF) mode. For DF transmission, if the SINR is sufficiently high for r_i to decode the source transmission, then r_i can serve as a relay for PU s . It can be mathematically denoted as $r_i \in \mathcal{D}(s)$, where $\mathcal{D}(s)$ is the decoding set of s [98]. Here the author only considers the SUs which can successfully decode the message of s . Therefore, $\mathcal{D}(s) = \{1, 2, \dots, N\}$ denotes the set of these SUs, where $N = |\mathcal{D}(s)|$. The SUs equally share the idle channel using time-division-multiple-access (TDMA), so that each SU can transmit without others' interference during its transmission period.

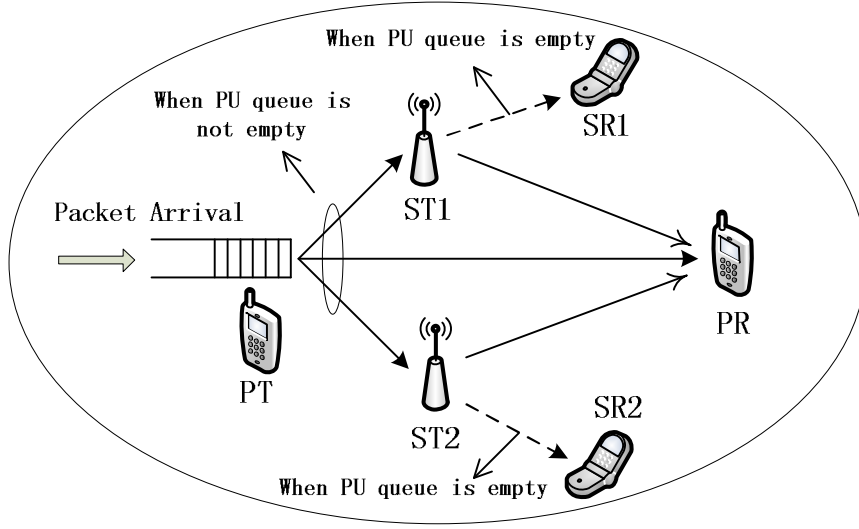


Figure 4.1: The cognitive relay network.

4.2 Optimal Power Allocation Scheme

4.2.1 Utility Function Design

When the PU has data to transmit, Protocol II described in [94] is applied, in which PU s communicates with the SUs in $\mathcal{D}(s)$ and PR over the first time slot, while in the second time slot, the SUs forward the received packet to PR. The maximum achievable rate for this protocol in the DF mode is given by [94]:

$$\mu_s = \min\{R_{sr}^{\max}, R_{\text{coop}}^{\max}\} \quad (4.1)$$

where R_{sr}^{\max} and R_{coop}^{\max} are the achievable rates of the PU-SU link and the cooperative link (including both PT-PR link and ST-PR links), respectively. When all the N SUs in the decoding set relay for PU s , it is shown in [54] that R_{sr}^{\max} is dominated by the worst channel. Let σ^2 denote the average power of the background thermal noise, then

we have

$$R_{\text{sr}}^{\max} = \log_2 \left(1 + \frac{\min_{i \in \mathcal{D}(s)} P_s |h_{s,r_i}|^2}{\sigma^2} \right), \quad (4.2)$$

where P_s is the transmission power of the PU. According to [94], we have

$$R_{\text{coop}}^{\max} = \log_2 \left(1 + \frac{P_s |h_{s,d}|^2}{\sigma^2} + \sum_{i \in \mathcal{D}(s)} \frac{P_i^{(r)} |h_{r_i,d}|^2}{\sigma^2} \right), \quad (4.3)$$

where $P_i^{(r)}$ is the power spent on relaying PU's data by SU_i .

The average packet arrival rate of the PU is λ_s . Packet departure rate μ_s can be obtained from (4.1). Let $\rho_s = \frac{\lambda_s}{\mu_s}$ denote the traffic intensity of PU s , then the probability for PU s 's queue to be empty can be derived as [99]:

$$P(0) = 1 - \rho_s = 1 - \frac{\lambda_s}{\mu_s}. \quad (4.4)$$

The maximum transmission rate of SU_i when utilizing the channel of PU can be obtained from

$$R_i = \log_2 \left(1 + \frac{P_i^{(t)} |h_{r_i,i}|^2}{\sigma^2} \right). \quad (4.5)$$

where $P_i^{(t)}$ is the power SU_i used to transmit its own data through the channel of PU.

The utility function of SU_i is defined as follows:

$$U_i = w_1 \mu_s + w_2 R_i, \quad (4.6)$$

where w_1 and w_2 are the corresponding weights of the transmission rate of PU and SU_i , respectively. The relay power of SU_i can affect μ_s (which further impacts the queue empty probability of PU), while the self-transmission power has direct influence on R_i . The author's objective is to find an optimal power allocation to maximize U_i of SU_i .

4.2.2 Problem Formulation of Single-PU-Single-SU case

First, the author considers a case of single PU and single SU (namely SU_1), i.e., $N = 1$.

The optimization problem for SU_1 can be formulated as follows:

$$\max_{P_1^{(r)}, P_1^{(t)}} U_1 = w_1 \mu_s + w_2 R_1, \quad (4.7)$$

$$\text{s.t.} \quad \lambda_s < \mu_s, \quad (4.8)$$

$$\frac{\lambda_s}{\mu_s} P_1^{(r)} + \left(1 - \frac{\lambda_s}{\mu_s}\right) P_1^{(t)} \leq E_{\text{th}}, \quad (4.9)$$

$$0 \leq P_1^{(r)} \leq P_{\text{max}}, \quad (4.10)$$

$$0 \leq P_1^{(t)} \leq P_{\text{max}}, \quad (4.11)$$

where R_1 is obtained from (4.5). (4.8) is the basic constraint for a queue to be stable [100]. Constraint (4.9) specifies that the energy consumption of SU_1 in unit time duration should be upper bounded by the threshold E_{th} , which is due to the battery life limit of a cognitive node. Constraints (4.10) and (4.11) limit the power SU_1 to be within an upper threshold P_{max} .

Note that the average achievable transmission rate of PU s (i.e., μ_s) depends on the relationship between $R_{\text{sr}}^{\text{max}}$ and $R_{\text{coop}}^{\text{max}}$. Thus, U_1 has different expressions in following cases:

- When SU_1 uses a relay power which results in $R_{\text{coop}}^{\text{max}} \geq R_{\text{sr}}^{\text{max}}$, μ_s becomes a constant equal to $R_{\text{sr}}^{\text{max}}$. Consequently, U_1 is only dependent on the self-transmission power of SU_1 .
- When SU_1 uses a relay power resulting in $R_{\text{coop}}^{\text{max}} \leq R_{\text{sr}}^{\text{max}}$, μ_s is determined by the relay power and U_1 varies with both relay power and self-transmission power of SU_1 .

Based on the above analysis, and with the assumption that $\lambda_s < R_{\text{sr}}^{\max}$ and $P_1^{(r)} \leq P_{\max}$ is always guaranteed to assure that the feasible set of Problem (4.7)-(4.11) is not empty, the author has the following propositions:

Proposition 4.1. For the optimal power allocation strategy, (i.e., the solution of Problem (4.7)-(4.11)):

$$P_1^{(t)} = \left[\frac{E_{\text{th}} - \lambda_s/\mu_s \cdot P_1^{(r)}}{1 - \lambda_s/\mu_s} \right]_0^{P_{\max}}, \quad (4.12)$$

where $[\cdot]_l^h = \min\{\max\{\cdot, l\}, h\}$.

Proof: It is obvious that U_1 is continuous with $P_1^{(t)}$. By taking partial derivative of U_1 with respect to (w.r.t.) $P_1^{(t)}$, it can be shown that the first order derivative of U_1 w.r.t $P_1^{(t)}$ is always greater than 0, while the second derivative is always less than 0, therefore, U_1 is a concave increasing function of $P_1^{(t)}$. According to constraint (4.9), we have

$$P_1^{(t)} \leq \frac{E_{\text{th}} - \lambda_s/\mu_s \cdot P_1^{(r)}}{1 - \lambda_s/\mu_s}. \quad (4.13)$$

Taking constraint (4.11) into account, we have $P_1^{(t)} = \left[\frac{E_{\text{th}} - \lambda_s/\mu_s \cdot P_1^{(r)}}{1 - \lambda_s/\mu_s} \right]_0^{P_{\max}}$. \square

Proposition 4.2. Assuming that SU_1 is rational, it will not choose a relay power resulting in $R_{\text{coop}}^{\max} > R_{\text{sr}}^{\max}$, if the following condition is satisfied.

$$\max f(P_1^{(r)}) \leq P_{\max}. \quad (4.14)$$

where $f(P_1^{(r)}) = (E_{\text{th}} - \lambda_s/\mu_s \cdot P_1^{(r)}) / (1 - \lambda_s/\mu_s)$.

Proof: Proof by contradiction is applied to prove this proposition. Assume that for the optimal power allocation strategy, SU_1 uses relay power $\hat{P}_1^{(r)}$ resulting in $R_{\text{coop}}^{\text{max}} > R_{\text{sr}}^{\text{max}}$, then we have $U_1 = w_1 \cdot R_{\text{sr}}^{\text{max}} + w_2 \cdot R_1$. According to Proposition 1, $P_1^{(t)} = \left[f(P_1^{(r)}) \right]_0^{P_{\text{max}}}$, which is equal to $f(P_1^{(r)})$ when condition (4.14) is satisfied. Since $\mu_s = R_{\text{sr}}^{\text{max}}$ when $R_{\text{coop}}^{\text{max}} > R_{\text{sr}}^{\text{max}}$, which is a constant, $f(P_1^{(r)})$ becomes a decreasing function of $P_1^{(r)}$. Thus, its maximum should be achieved when $P_1^{(r)}$ takes its minimum feasible value, i.e., the smallest value satisfying $R_{\text{coop}}^{\text{max}} > R_{\text{sr}}^{\text{max}}$. Based on the assumption of the author, $\hat{P}_1^{(r)}$ should be the minimum feasible value of $P_1^{(r)}$.

Note that when condition (4.14) is satisfied, according to Proposition 1, $P_1^{(t)}$ will always be bounded by $f(P_1^{(r)})$ rather than P_{max} . Consider that if SU_1 uses relay power $\tilde{P}_1^{(r)}$ which results in $R_{\text{coop}}^{\text{max}} = R_{\text{sr}}^{\text{max}}$, it is straightforward to have $f(\tilde{P}_1^{(r)}) > f(\hat{P}_1^{(r)})$ since $\tilde{P}_1^{(r)} < \hat{P}_1^{(r)}$. Hence, the author arrives at the conclusion that using relay power $\tilde{P}_1^{(r)}$ instead of $\hat{P}_1^{(r)}$ can achieve higher utility for SU_1 , which contradicts with the previous assumption that $\hat{P}_1^{(r)}$ is the relay power for optimal strategy of SU_1 . Consequently, SU_1 will not choose a relay power resulting in $R_{\text{coop}}^{\text{max}} > R_{\text{sr}}^{\text{max}}$ when condition (4.14) is satisfied. \square

Note that Proposition 4.2 can be extended to multi-SU case, the proof by contradiction can be applied similarly. Based on Proposition 4.1 and Proposition 4.2, it can be found that $P_1^{(t)}$ can be obtained from $P_1^{(r)}$ by $P_1^{(t)} = \left[f(P_1^{(r)}) \right]_0^{P_{\text{max}}}$. Thus U_1 is only dependent on $P_1^{(r)}$, and the maximum value of U_1 can be obtained by taking its derivative w.r.t. $P_1^{(r)}$ and equating it to 0. The corresponding optimal power allocation strategies including $P_1^{(r)}$ and $P_1^{(t)}$ can be obtained.

4.2.3 Problem Formulation of Single-PU-Multiple-SU Case

Next, the author considers a more general case of multiple SUs. The power allocation problem for SU_i can be formulated as:

$$\max_{P_i^{(r)}, P_i^{(t)}} U_i = w_1 \cdot \mu_s + w_2 \cdot R_i, \quad (4.15)$$

$$\text{s.t.} \quad \lambda_s < \mu_s, \quad (4.16)$$

$$\frac{\lambda_s}{\mu_s} P_i^{(r)} + \frac{1}{N} \left(1 - \frac{\lambda_s}{\mu_s} \right) P_i^{(t)} \leq E_{\text{th}}, \quad (4.17)$$

$$0 \leq P_i^{(r)} \leq P_{\max}, \quad (4.18)$$

$$0 \leq P_i^{(t)} \leq P_{\max}, \quad (4.19)$$

where the factor $\frac{1}{N}$ is due to equal share of the idle channel among N cooperating SUs. For simplicity, it is assumed that all SUs have the same P_{\max} value. Different from Problem (4.7), the utility U_i in Problem (4.15) depends on not only the strategy of SU_i (i.e., the relay power $P_i^{(r)}$ and self-transmission power $P_i^{(t)}$), but also the other SUs' strategies, denoted by $(P_{-i}^{(r)}, P_{-i}^{(t)})$. Strategies of different SUs are also coupled in constraint (4.17). Thus, problem (4.15)-(4.19) is not a pure optimization problem but a non-cooperative power allocation game (NPG). The players are the SUs in $\mathcal{D}(s)$, and each of which attempts to optimally allocate $P_i^{(r)}$ and $P_i^{(t)}$ for maximizing its own utility U_i . The strategy space can be denoted as $\mathcal{P} = [P_i]_{i \in \mathcal{D}(s)}$, where $P_i = (P_i^{(r)}, P_i^{(t)})$. The utility set is $\mathcal{U} = [U_i]_{i \in \mathcal{D}(s)}$. Nash equilibrium (NE) is considered to be the solution of this NPG, which is defined as follows:

Definition 4.1. A strategy vector $P^* = (P_1^*, \dots, P_N^*)$, is a Nash equilibrium (NE) of the NPG $G = [\mathcal{D}(s), \{\mathcal{P}\}, \{\mathcal{U}\}]$ if, for every $i \in \mathcal{D}(s)$, $U_i(P_i^*, P_{-i}^*) \geq U_i(\hat{P}_i, P_{-i}^*)$ for all $\hat{P}_i \in \mathcal{P}$, where $U_i(P_i^*, P_{-i}^*)$ denotes the utility of SU_i when taking strategies P_i^* , given the other players' strategies P_{-i}^* .

The essence of NE is that, at this point, no player can achieve higher benefit by unilaterally changing its own strategy when the other players keep their strategies unaltered. After investigating this NPG, the author has the following result:

Proposition 4.3. Under the assumption that all SUs in \mathcal{S} are rational, for the optimal power allocation strategy:

$$P_i^{(t)} = \left[\frac{(E_{\text{th}} - \lambda_s/\mu_s \cdot P_i^{(r)}) \cdot N}{(1 - \lambda_s/\mu_s)} \right]_0^{P_{\max}}. \quad (4.20)$$

The proof follows the same way as derived for Proposition 4.1. In the following, the author analyzes the existence and uniqueness of NE in the NPG.

Theorem 4.1. An NE exists in the NPG $G = [\mathcal{D}(s), \{\mathcal{P}\}, \{\mathcal{U}\}]$ defined in (4.15)-(4.19), if the following condition is satisfied:

$$\max f(P_i^{(r)}) \leq P_{\max}, \quad \forall i \in \mathcal{D}(s). \quad (4.21)$$

Proof: According to [101], a concave game is defined as a game in which player i selects an action $x_i \in \mathfrak{X}^{m_i}$ so that the strategy vector $x = (x_1, \dots, x_n) \in S$, where S is a closed bounded convex set. Let $\phi_i(x)$ denote the payoff function of player i , it should be continuous in x and concave in x_i , for $x \in S$.

For the considered problem, it is easy to verify that the strategy set \mathcal{P} is a closed bounded convex set, and the utility function U_i of SU_i is continuous in $P = (P_1, \dots, P_N)$. Now the author proves the concavity of U_i w.r.t. P_i . Recall that the condition in Proposition 4.2 can be extended to multi-SU case with similar proof, specifying that $P_i^{(t)}$ will always be bounded by $f(P_i^{(r)})$ rather than P_{\max} . Hence, when condition (4.21) is satisfied, we have $\mu_s = R_{\text{coop}}^{\max}$. With the definition of μ_s and R_i , U_i is obviously continuous in P . Now, the author proves its concavity w.r.t. P_i by

checking the corresponding Hessian matrix. To simplify the notations, let $x = P_i^{(r)}$, $y = P_i^{(t)}$, $a = \frac{P_s |h_{s,d}|^2}{\sigma^2}$, $b = \frac{|h_{r_i,d}|^2}{\sigma^2}$, $c = \sum_{j \in \mathcal{D}(s), j \neq i} \frac{P_j^{(r)} |h_{r_j,d}|^2}{\sigma^2}$, and $d = \frac{|h_{r_i,i}|^2}{\sigma^2}$, then U_i can be denoted as follows:

$$U_i = w'_1 \cdot \ln(1 + a + c + b \cdot x) + w'_2 \cdot \ln(1 + d \cdot y) \quad (4.22)$$

where $w'_1 = \frac{w_1}{\ln(2)}$ and $w'_2 = \frac{w_2}{\ln(2)}$. The Hessian matrix of U_i can be obtained as follows:

$$\mathbf{H}_{U_i} = \begin{bmatrix} -\frac{b^2 w'_1}{(1+a+c+b \cdot x)^2} & 0 \\ 0 & -\frac{d^2 w'_2}{(1+d \cdot y)^2} \end{bmatrix}. \quad (4.23)$$

It can be found that \mathbf{H}_{U_i} is a square matrix, since the first-order principal minor determinant of \mathbf{H}_{U_i} , i.e., the first element of \mathbf{H}_{U_i} , is obviously less than 0, and the second-order principal minor determinant of \mathbf{H}_{U_i} is:

$$\det(\mathbf{H}_{U_i}) = \frac{b^2 d^2 w'_1 w'_2}{(1 + a + c + b \cdot x)^2 (1 + d \cdot y)^2} > 0.$$

According to the definition of *negative definiteness* for a square symmetric matrix [101], \mathbf{H}_{U_i} is negative definite. Thus, $U_i(P_i)$ is concave in P_i . Consequently, the NPG is a concave game. According to Rosen's Theorem that a concave game always has at least one NE [102], the author can arrive at the conclusion that an NE exists in the NPG $G = [\mathcal{S}, \{P_i\}, \{U_i\}]$ under condition (4.21). \square

Next, the author investigates the uniqueness of NE. The key aspect of the uniqueness proof is to verify that the game model is *diagonally strictly concave* on its strategy space \mathcal{P} , i.e., for any two strategy vectors $P(0) \neq P(1)$ with $P(k) = [P_1(k), \dots, P_N(k)]^T$

for $k = 0, 1$, where $P_i(k) = (P_i^{(r)}(k), P_i^{(t)}(k))$, and for $r = [r_1, \dots, r_N]^T$ with $r_i > 0$ ($i = 1, \dots, N$), the following inequality holds [102]:

$$\Delta P(1, 0)g(P(0), r) + \Delta P(0, 1)g(P(1), r) > 0 \quad (4.24)$$

where $\Delta P(1, 0)$ is a simplified denotation of $(P(1) - P(0))^T$, and $\Delta P(0, 1)$ represents $(P(0) - P(1))^T$. The function $g(P, r)$ is defined as follows:

$$g(P, r) = \left[r_1 \frac{\partial U_1}{\partial P_1}, \dots, r_N \frac{\partial U_N}{\partial P_N} \right]^T. \quad (4.25)$$

The following theorem addresses the uniqueness of the NE of the NPG.

Theorem 4.2. The NE of the NPG defined in (4.15)-(4.19) is unique, when condition (4.21) is satisfied.

Proof: For Condition (4.24), we have

$$\begin{aligned} & (P(1) - P(0))^T [g(P(0), r) - g(P(1), r)] \quad (4.26) \\ &= (P(1) - P(0))^T \times \\ & \quad \left[r_1 \left(\frac{\partial U_1}{\partial P_1(0)} - \frac{\partial U_1}{\partial P_1(1)} \right), \dots, r_N \left(\frac{\partial U_N}{\partial P_N(0)} - \frac{\partial U_N}{\partial P_N(1)} \right) \right]^T \\ &= \sum_{i=1}^N r_i (P_i(1) - P_i(0)) \left(\frac{\partial U_i}{\partial P_i(0)} - \frac{\partial U_i}{\partial P_i(1)} \right) \\ &= \sum_{i=1}^N r_i \begin{bmatrix} P_i^{(r)}(1) - P_i^{(r)}(0) \\ P_i^{(t)}(1) - P_i^{(t)}(0) \end{bmatrix}^T \begin{bmatrix} \frac{\partial U_i}{\partial P_i^{(r)}(0)} - \frac{\partial U_i}{\partial P_i^{(r)}(1)} \\ \frac{\partial U_i}{\partial P_i^{(t)}(0)} - \frac{\partial U_i}{\partial P_i^{(t)}(1)} \end{bmatrix}. \end{aligned}$$

Let $\delta_i = (P_i^{(r)}(1) - P_i^{(r)}(0))(\frac{\partial U_i}{\partial P_i^{(r)}(0)} - \frac{\partial U_i}{\partial P_i^{(r)}(1)})$, and $\varphi_i = (P_i^{(t)}(1) - P_i^{(t)}(0))(\frac{\partial U_i}{\partial P_i^{(t)}(0)} - \frac{\partial U_i}{\partial P_i^{(t)}(1)})$. According to the first and second order derivatives of U_i w.r.t. $P_i^{(r)}$ and $P_i^{(t)}$, it can be found that when condition (4.21) is satisfied, $\frac{\partial U_i}{\partial P_i^{(r)}}$ and $\frac{\partial U_i}{\partial P_i^{(t)}}$ are monotonically decreasing w.r.t. $P_i^{(r)}$ and $P_i^{(t)}$, respectively. Thus, if $P_i^{(r)}(1) > P_i^{(r)}(0)$, we have $\partial U_i / \partial P_i^{(r)}(0) - \partial U_i / \partial P_i^{(r)}(1) > 0$, and, hence, $\delta_i > 0$. Similarly, we also have $\delta_i > 0$ for $P_i^{(r)}(1) < P_i^{(r)}(0)$. Using the same method, the author can verify that $\varphi_i > 0$. Consequently, $\sum_{i=1}^N r_i(\delta_i + \varphi_i) > 0$, and thus the author arrives at the conclusion that the proposed NPG has a unique NE under condition (4.21). \square

To obtain the NE, the best response strategy is applied. Specifically, given the other SUs' strategies P_{-i} , the optimal strategy of SU_i can be obtained from

$$P_i^* = \mathcal{B}(P_{-i}) = \arg \max_{P_i} U_i(P_i, P_{-i}). \quad (4.27)$$

4.2.4 Architecture of the Proposed Algorithm

In this section, the author proposes the architecture of the proposed algorithm to solve the power allocation problem, which is shown in the following:

4.3 Numerical Results

4.3.1 Single-PU-Single-SU Case

In this section, the author considers the single-PU-single-SU case and the parameters are set as follows: $\lambda_s = 1$, $h_{s,d} = 0.5$, $h_{s,r_1} = 1$, $h_{r_1,d} = 0.6$, $h_{r_1,1} = 0.7$, $\sigma^2 = 1$, and $P_s = 40$. The maximum power for an SU is $P_{\max} = 100$ and $E_{\text{th}} = 20$. With

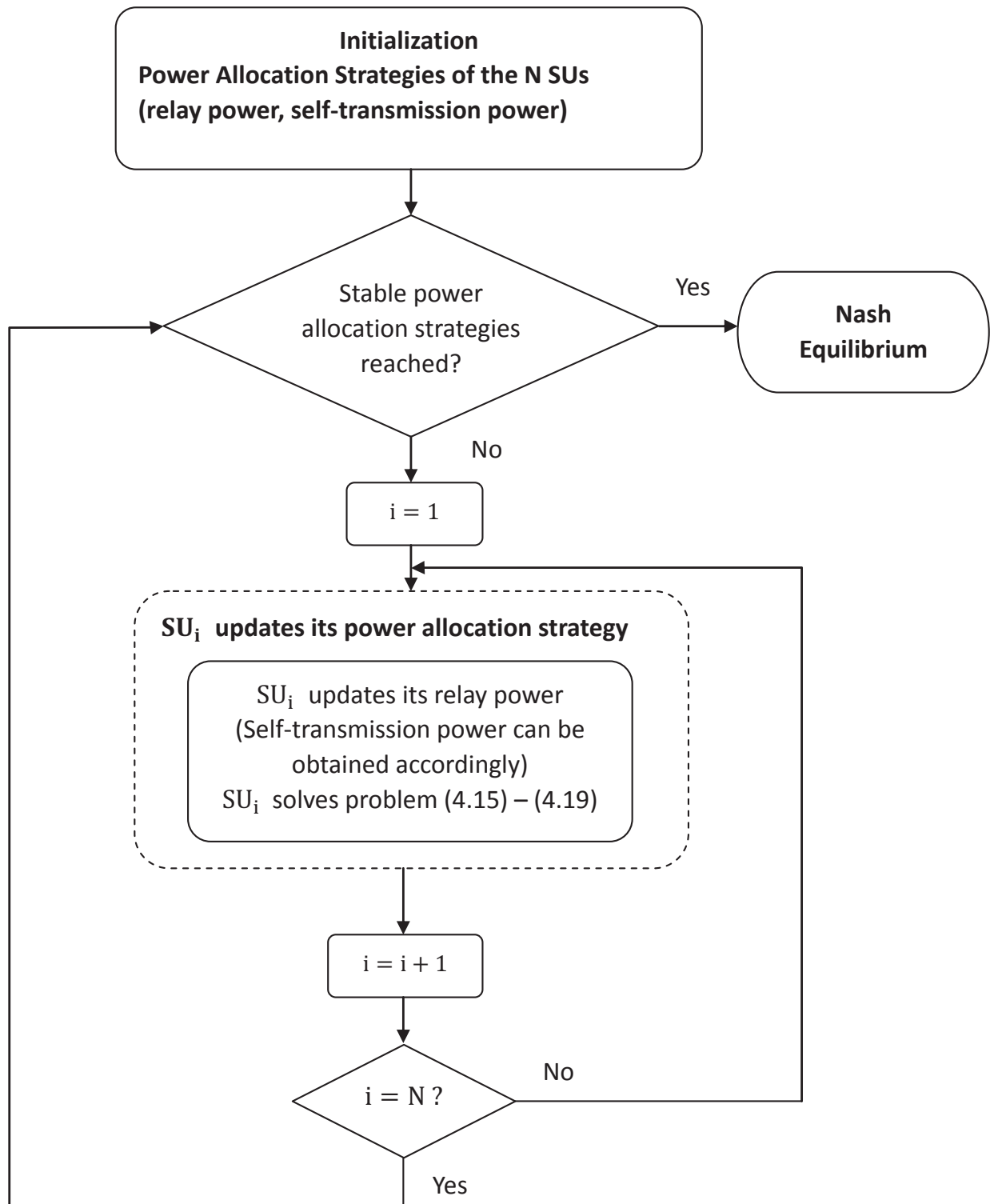


Figure 4.2: Flowchart of the proposed algorithm to solve the Power Allocation Problem.

these parameters, $R_{\text{sr}}^{\text{max}} = 5.3576$. First, the author investigates the impact of different weight settings of utility function on the power allocation. Note that different weight settings can reflect the emphasis on either PU or SU. The author keeps w_2 fixed to be 1 and increases w_1 from 0 to 4. The variation of relay power and self-transmission power of SU_1 for the optimal power allocation strategy is shown in Fig. 4.3.

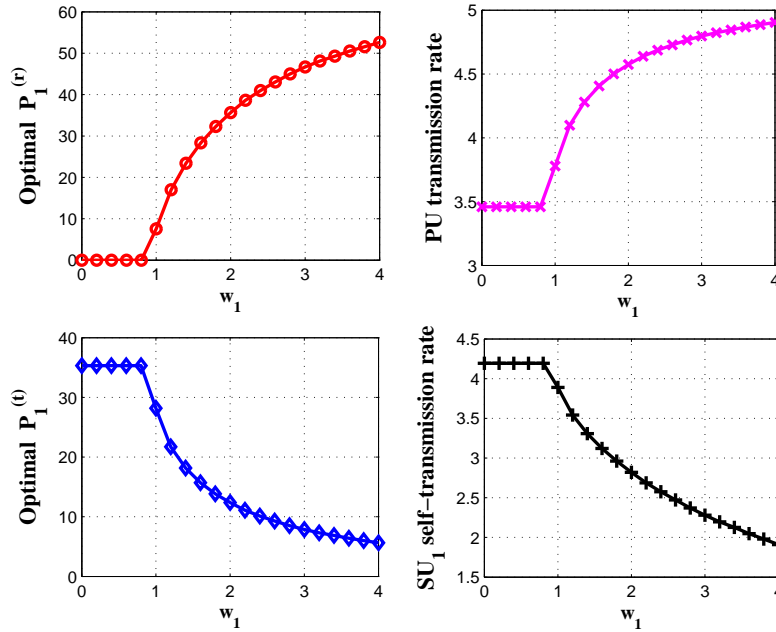


Figure 4.3: Optimal power allocation strategy variation with weight in utility function.

In Fig. 4.3, as w_1 varies from 0 to 4, the optimal relay power $P_1^{(r)}$ increases from 0 to 52.61, and the PU transmission rate increases from 3.4594 to 4.9040. The self-transmission power $P_1^{(t)}$ decreases from 35.31 to 5.63, which results in the decrease of SU self-transmission rate from 4.1939 to 1.9102. The reason is that, when w_1 is small compared with w_2 , the PU transmission rate does not contribute much to the SU_1 's utility, thus $P_1^{(r)} = 0$ for the optimal strategy at first. This indicates that the SU will not waste power on relaying but to concentrate on improving its self-transmission rate to maximize the utility. However, as w_1 keeps increasing with w_2 unchanged,

the PU transmission rate gradually becomes a major concern for the SU to achieve higher utility. Hence, it is beneficial to use higher relay power for increasing PU transmission rate, although its self-transmission rate decreases at the same time, it will not deteriorate the utility significantly due to the effect of weight.

4.3.2 Single-PU-Two-SU Case

The Best Response Strategy

The author considers the single-PU-multiple-SU case where $N = 2$. The two-SU case is taken as an example since the NE can be graphically shown by the intersection of best response curves. The parameters are set as follows: $\lambda_s = 1.5$, $h_{s,d} = 0.5$, $h_{s,r_1} = h_{s,r_2} = 1$, $h_{r_1,d} = h_{r_2,d} = 0.6$, and $h_{r_1,1} = h_{r_2,2} = 0.7$. The weights w_1 and w_2 are set as 2 and 1, respectively.

According to the definition of Nash equilibrium, each SU has made best response to the given power allocation of the other SU. To specify how the best response strategy behaves, suppose the power allocation of SU_1 have been given such that $P_1^{(r)} = 10$, $P_1^{(t)} = 20$, SU_2 will set its relay power $P_2^{(r)}$ and self-transmission power $P_2^{(t)}$ to 31.51 and 29.2, respectively. This is illustrated in Fig. 4.4 which shows the variation of SU_2 's utility with the relay power that SU_2 may select.

As shown in Fig. 4.4(a), when SU_1 keeps its power allocation fixed, as the relay power of SU_2 increases, the PU transmission rate μ_s also increases, which indicates that the PU can achieve a higher transmission rate and its buffer will be depleted faster. In Fig. 4.4(b), it can be found that the self-transmission rate of SU_2 gradually decreases. The reason is that as the relay power increases, the self-transmission power decreases simultaneously due to the energy constraint. Since the utility of SU_2 depends on both PU transmission rate and self-transmission rate of SU_2 , it increases first and

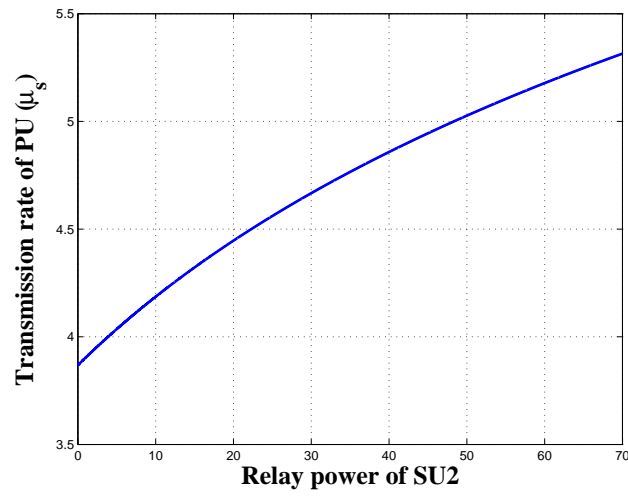
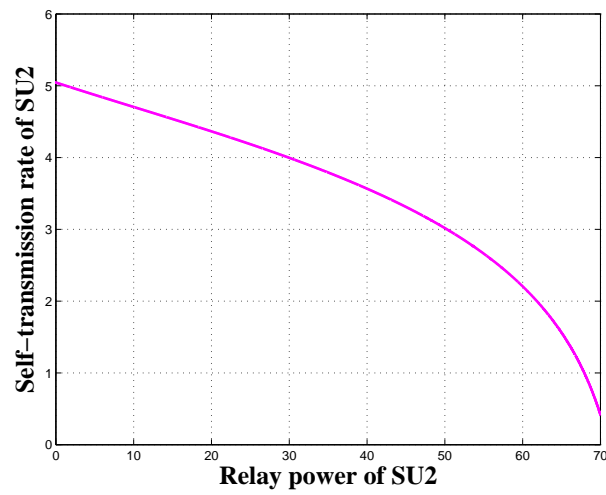
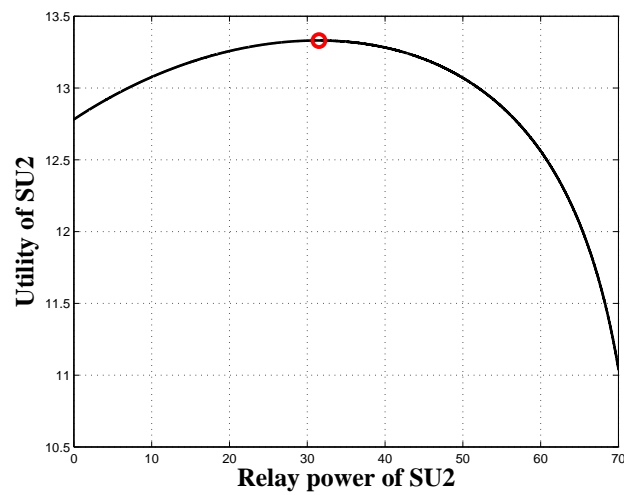
(a) Transmission rate of PU vs. $P_2^{(r)}$ (b) Transmission rate of SU_2 vs. $P_2^{(r)}$ (c) Utility of SU_2 vs. $P_2^{(r)}$

Figure 4.4: Interpretation of the best-response strategy.

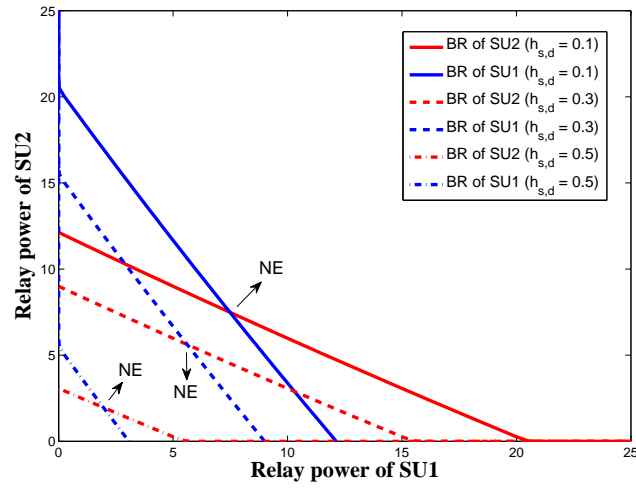
then decreases, achieving the maximum of 13.33 when $P_2^{(r)} = 31.48$ and $P_2^{(t)} = 29.17$, as shown in Fig. 4.4(c). The best response strategy aims at searching for this turning point at which the utility is maximized. In other words, each SU will always make the optimal decision to maximize its own benefit, given the decisions of the other SUs.

The Effect of $h_{s,d}$ on the NE

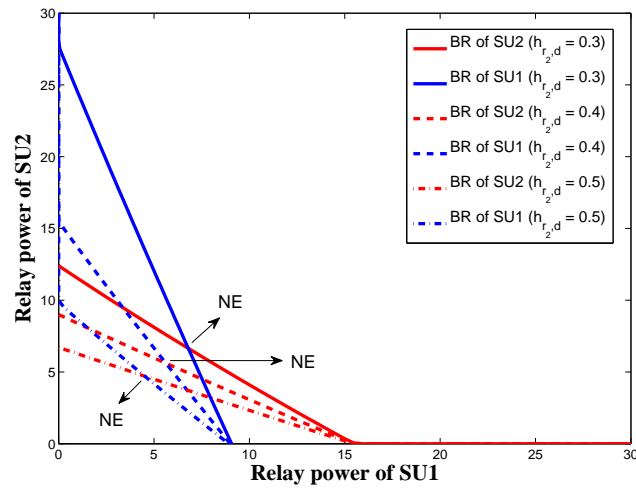
To study the effect of $h_{s,d}$ on NE, the author sets $\lambda_s = 1.5$, $P_s = 10$, $E_{\text{th}} = 15$, $h_{r_1,d} = h_{r_2,d} = 0.4$ and vary $h_{s,d}$. The weights are set as $w_1 = w_2 = 1$, and remain the same for the rest of simulations. Fig. 4.5 shows the best response (denoted as BR in the figure) curves of SU_1 and SU_2 under different parameter settings. As shown in Fig. 4.5(a), at the NE (corresponding to the intersection of the two best response curves for the same parameter setting), the relay powers of SU_1 and SU_2 are 7.49 when $h_{s,d} = 0.1$, while those of SU_1 and SU_2 when $h_{s,d} = 0.3$ and $h_{s,d} = 0.5$ are 5.62 and 1.96, respectively. It can be observed that when the direct channel between the PT and PR is in good condition, both SUs can use less power on relaying, while more power can be utilized for their own transmissions.

The Effect of $h_{r,d}$ on the NE

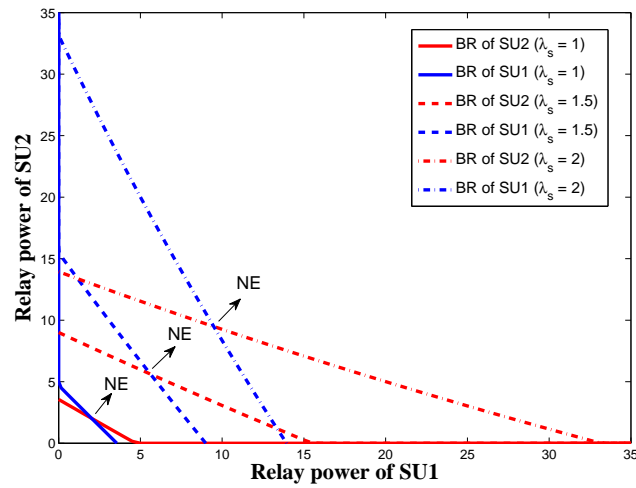
To study the effect of $h_{r,d}$ on NE, the author sets $\lambda_s = 1.5$, $h_{s,d} = 0.3$, $h_{r_1,d} = 0.4$ and vary $h_{r_2,d}$. As shown in Fig. 4.5(b), at the NE, the relay powers of SU_1 and SU_2 are 6.8 and 6.64, respectively, when $h_{r_2,d} = 0.3$. The relay powers of SU_1 and SU_2 when $h_{r_2,d} = 0.4$ are both 5.62. When $h_{r_2,d} = 0.5$, at the NE, the relay powers of SU_1 and SU_2 are 4.52 and 4.71, respectively. From the results, the author observes that when the ST-PR channel gain is improved, SUs can use less power on relaying. Moreover, the SU with a better channel condition to PR (compared with others) will use more relay



(a) NE variation with $h_{s,d}$



(b) NE variation with $h_{r,d}$



(c) NE variation with λ_s

Figure 4.5: Investigation of different impact factors on NE.

power. The reason is that, with the same amount of power, the SU with a better relay channel can contribute more to the PU transmission rate, which has a more impact on the improvement of both SUs' utilities.

The Effect of λ_s on the NE

To study the effect of λ_s on NE, the author sets $h_{s,d} = 0.3$, $h_{r_1,d} = h_{r_2,d} = 0.4$ and vary λ_s . As shown in Fig. 4.5(c), at the NE, the relay powers of SU_1 and SU_2 are both 2.02, when $\lambda_s = 1$. The relay powers of SU_1 and SU_2 when $\lambda_s = 1.5$ are 5.62. When $\lambda_s = 2$, the relay powers of SU_1 and SU_2 are 9.49. From the results, the author concludes that as λ_s increases, the SUs should use more power on relaying for the PU to satisfy the constraint $\mu_s > \lambda_s$. The results also accord with the author's expectation, since when the packet arrival rate at the PU increases, it is rational for the SUs to use more relay power on cooperation for emptying the PU's queue faster. Consequently, more self-transmission opportunities can be obtained. Otherwise, the PU always has data to transmit, thus the SUs can get seldom chances to access the PU's channel.

4.4 Summary

In this chapter, the author has considered cooperation in cognitive relay networks from a new perspective. In the proposed scheme, SUs are allowed to choose their relay power and self-transmission power separately. For single-SU case, the power allocation problem has been formulated as an optimization problem, and the method to solve the optimization problem has been provided. In the case of multiple SUs, a non-cooperative game model has been formulated and the Nash equilibrium (NE) has been achieved. The existence as well as the uniqueness of the NE under certain conditions have been proved. The effect of different weight settings on the utility function has been

discussed. Moreover, the variation in NE with respect to different system parameters has also been presented.

Chapter 5

Cooperation Formation for Downlink Data Transmission in Mobile Infostation Networks

In this chapter, the author proposes a hierarchical cooperation formation model for downlink data transmission in mobile infostation networks, aiming at finding a stable cooperation structure for both the infostations and mobile users. A distributed hierarchical cooperation formation algorithm is proposed to achieve this goal. As the research background described in detail in Section 1.2, the author starts this chapter from describing the proposed system model in Section 5.1, followed by Section 5.2, where the hierarchical cooperation formation framework for downlink data transmission is comprehensively investigated. Performance evaluation of the proposed framework is presented in Section 5.3. Finally, the summary of this chapter is drawn in Section 5.4.

5.1 System Model

The author considers a mobile infostation network that consists of N infostations (ISs) and M mobile users (MUs). Let $\mathbb{N} = \{1, \dots, N\}$ and $\mathbb{M} = \{1, \dots, M\}$ be the sets of the ISs and MUs, respectively. Each infostation can provide a certain type of content downloading service to a group of mobile users. Each user can subscribe to a certain infostation¹. The entire group of mobile users can be partitioned into multiple subsets, according to different infostations which they subscribe to. Let \mathbb{H}_i denote the set of subscribers of IS_i , thus $\bigcup_{i \in \mathbb{N}} \mathbb{H}_i = \mathbb{M}$. In practice, each infostation needs to maintain a user database to keep track the status of its subscribers. For example, when a mobile user enters the transmission range of the infostation, the infostation should have the ability to decide whether the user is its subscriber and whether data service should be provided to the user or not.

Fig. 5.1 shows an illustrative example of the proposed network model with three infostations (i.e., $IS_1 - IS_3$) and six mobile users (i.e., $MU_1 - MU_6$). In this example, MU_1, MU_2 subscribe to IS_1 , while MU_3, MU_4 subscribe to IS_2 , and MU_5, MU_6 subscribe to IS_3 . As the mobile users move, the inter-encounter interval between MU_m and its subscribed infostation IS_i is defined as the time period between two consecutive meetings of MU_m and IS_i , which is assumed to be exponentially distributed with mean $1/\lambda_{mi}$ minutes [82], where λ_{mi} is the mean contact rate between MU_m and IS_i (as shown in Fig. 5.1). The inter-encounter interval between two mobile users (i.e., MU_m and MU_n) can be defined accordingly. It is also assumed to be exponentially distributed, with the mean of $1/\mu_{mn}$ minutes, where μ_{mn} is the mean contact rate between MU_m and MU_n . Here, we have the relationship that μ_{mn} is equal to μ_{nm} . The history on contact events can be analyzed to obtain the knowledge of inter-encounter

¹Please note that the proposed model can be extended to a more general case which allows one user to subscribe to multiple infostations, according to its multiple interests. In this case, the calculation of a user's payoff is different, since delays of different contents have to be jointly taken into account.

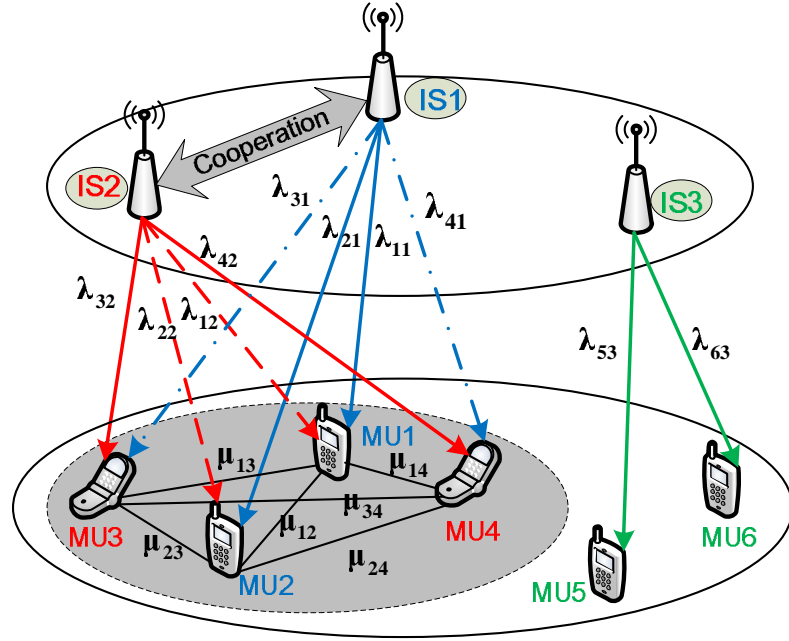


Figure 5.1: An illustrative example scenario of mobile infostation network.

interval among mobile users and infostation. Note that both λ_{mi} and μ_{mn} may vary over time. However, they change with a much larger time scale than that needed to disseminate data content, and thus it is assumed that the values of λ_{mi} and μ_{mn} keep constant in the following discussion.

With a certain coalitional structure of the infostations, the author assumes that the infostations in a coalition will impose some policy to stimulate the cooperation among their subscribers. Therefore, only the subscribers of the infostations in the same coalition are possibly willing to help each other in the proposed model. It is assumed that the contents are available at each infostation (e.g., due to wired connection among infostation). As a result, the time used for content transmission between infostations is excluded from the expected delay of the user. Moreover, the content transmission time between a pair of mobile users is considered to be much smaller when compared with the inter-encounter interval, and hence it is also ignored in the proposed model. Let $d_m(\mathcal{S}_I, \mathcal{G}_M)$ denote the expected delay of MU_m , given the coalitional structure of the

ISs (i.e., \mathcal{S}_I) and the cooperation formation of the MUs (i.e., \mathcal{G}_M). Note that the formal definitions of \mathcal{S}_I and \mathcal{G}_M will be given in Section 5.2.2 and Section 5.2.1, respectively. The payoff of MU_m is defined as a function of $d_m(\mathcal{S}_I, \mathcal{G}_M)$ as follows:

$$u_m(\mathcal{S}_I, \mathcal{G}_M) = 1 - \frac{d_m(\mathcal{S}_I, \mathcal{G}_M)}{d_m^{noc}} - C_i^M(\mathcal{G}_M, \eta) \quad (5.1)$$

where d_m^{noc} is the expected delay of MU_m when neither the ISs nor the MUs cooperate. The author uses $C_i^M(\mathcal{G}_M, \eta)$ to denote the cost of MU_m for cooperating with other MUs, (e.g., to store and forward the content). Assuming that the cost of forming cooperation between a pair of MUs is fixed, which is denoted by η , and hence if MU_m has cooperation with more users, the cost to maintain the cooperation relationship will increase. Note that if the cooperation does not exist among either the ISs or the MUs, $d_m(\mathcal{S}_I, \mathcal{G}_M)$ will be equal to d_m^{noc} , resulting in a zero utility of MU_m . As more users cooperate with MU_m , $d_m(\mathcal{S}_I, \mathcal{G}_M)$ will be lower, resulting in the increase of the benefit term $(1 - \frac{d_m(\mathcal{S}_I, \mathcal{G}_M)}{d_m^{noc}})$, which will be closer to 1. However, the cost term will increase at the same time, and thus a tradeoff exists here. The payoff from cooperation of IS_i is defined as the sum of payoffs of its subscribers, while the cost of forming such coalition will be subtracted, which is denoted by C_i^I , i.e.,

$$U_{IS_i}(\mathcal{S}_I, \mathcal{G}_M) = \sum_{m \in \mathbb{H}_i} u_m(\mathcal{S}_I, \mathcal{G}_M) - C_i^I(\mathcal{S}_I, \sigma) \quad (5.2)$$

Similarly, σ is used to denote the cost for forming coalition with another IS. It is also assumed that the cost for an infostation to form coalition with another one is fixed, and hence the larger the coalition size, the more cost each infostation has to undertake for maintaining such a coalition.

Next, the author describes how to obtain the expected delay with the knowledge

of the cooperation relationship among the infostations and mobile users. Consider the example shown in Fig. 5.1, IS_1 and IS_2 have formed a coalition while IS_3 acts independently. As a result, a cooperation network will be formed among $MU_1 - MU_4$. Suppose MU_1 and MU_2 form a cooperation link, then there are four paths for MU_1 to obtain the desired content: (1) when MU_1 meets IS_1 (corresponding to expected delay T_1); (2) when MU_1 meets IS_2 (corresponding to expected delay T_2); (3) when MU_1 meets MU_2 and at that time, MU_2 has already downloaded the content for MU_1 from IS_1 (corresponding to expected delay T_3), and (4) when MU_1 meets MU_2 and at that time, MU_2 has already downloaded the content for MU_1 from IS_2 (corresponding to expected delay T_4). In this case, no matter which of the four events occurs, MU_1 can successfully obtain the required content. Therefore, the expected delay of MU_1 to obtain the content should be the minimum of T_1, T_2, T_3 and T_4 , which can be expressed as $d_1 = \min\{T_1, T_2, T_3, T_4\}$. It can be evidently seen that the diversity of the content circulating in the network is improved in this way. It can also be foreseen that the expected delay for MU_m to get the required content will be reduced. The higher the content circulating diversity is, the lower the expected delay will be, this is the reason why the MUs can benefit from the cooperation among the ISs.

Let X_{mi} denote the time needed for MU_m to meet IS_i . Let B_{mn} denote the time needed for MU_m to meet MU_n . Then $Y_1 = X_{21} + B_{12}$ and $Y_2 = X_{22} + B_{12}$ represent the expected delay of MU_1 to meet MU_2 with the content previously obtained from IS_1 and IS_2 , respectively (which are corresponding to T_3 and T_4 in the example). Note that probability density function $f_{X_{mi}}(t) = \lambda_{mi}e^{-\lambda_{mi}t}$ ($m = 1, 2; i = 1, 2$), $f_{B_{12}}(t) = \mu_{12}e^{-\mu_{12}t}$. Without loss of generality, it is assumed that the mean contact rate differentiates from one to another. Thus both Y_1 and Y_2 are hypo-exponentially distributed variables. Then the expected delay of MU_1 should be the minimum value of random variables X_{11}, X_{12}, Y_1 , and Y_2 . To obtain $d_m(\mathcal{S}_I, \mathcal{G}_M)$, the author introduces

the following theorem and propositions.

Theorem 5.1. Let $Z = \min\{X_1, X_2, \dots, X_n, Y_1, Y_2, \dots, Y_m\}$, where $X_i (1 \leq i \leq n)$ are exponentially distributed variables and n is the number of such variables, $Y_j (1 \leq j \leq m)$ are hypo-exponentially distributed variables and m is the number of such variables. Then the cumulative distribution function (CDF) of Z can be calculated as:

$$F_Z(t) = P(Z \leq t) = 1 - \prod_{j=1}^m (1 - F_{Y_j}(t)) e^{-(\sum_{i=1}^n \lambda_i)t}. \quad (5.3)$$

where λ_i is the mean of random variable X_i , and $F_{Y_j}(t)$ is the CDF of random variable Y_j .

The theorem can be proved by applying the mathematical induction techniques. Note that for different types of the two-level cooperation structure, the expected delay of the users can be obtained accordingly.

5.2 Hierarchical Cooperation Formation

In this section, the author proposes a hierarchical cooperation formation model for data transmission in mobile infostation networks based on the aforementioned system model. The structure of the hierarchical cooperation formation framework is similar to the Stackelberg game [55]. The infostations and the mobile users are considered as the leaders and followers, respectively.

- At the lower level, the mobile users strategically establish a cooperation network to achieve the highest individual payoffs, given the coalitional structure of the infostations. The author aims at achieving a stable network in which none of the users can further improve its payoff by unilaterally changing its own cooperation

strategy, while the strategies of the other users are unaltered. This lower-level problem is solved by utilizing the network formation process [17].

- At the upper level, with the knowledge of how the mobile users will form cooperation network under the given coalitional structure of the infostations, the infostations will intelligently decide how to form the coalition to maximize their individual payoffs. This higher-level problem is solved by applying the coalition formation game theory [18].

Note that in the cooperation formation of the infostations, it is assumed that the fixed infostations have complete knowledge (e.g., the location information, subscriber information, and contact rates between subscribers of themselves and other infostations) about each other, and each infostation can make its cooperation strategy such as whether to join in a new coalition or leave the current coalition. Different from this situation, the number of mobile users in the network is generally much larger than that of infostations, and the mobility makes the cooperation formation among the mobile users more challenging. Also, the mobile user may not have full knowledge about the other mobile users. Hence, the author utilizes the network formation process to study the pairwise cooperation formation among the mobile users instead of the coalition formation game. The cooperation formations among the mobile users and infostations are introduced in Section 5.2.1 and Section 5.2.2, respectively. The author also proposes a distributed algorithm for the infostations and mobile users to reach the stable hierarchical cooperation structure, which is described in Section 5.2.3.

5.2.1 Cooperation among the Mobile Users

First, the author discusses the lower-level cooperation formation. The aim of the author is to design a low complexity distributed scheme for the mobile users who wish

to build a stable cooperation network structure among the mobile users, specifying the set of MUs that $MU_m (m = 1, \dots, M)$ is willing to help, as well as that of MUs which can benefit from helping $MU_m (m = 1, \dots, M)$. The cooperation formation among the MUs is formulated as a network formation process, the participants of which are the MUs in \mathbb{M} . Note that a mapping can be constructed from an arbitrary cooperation network structure \mathcal{G}_M to a vector of payoffs that the users can achieve under network structure \mathcal{G}_M , which is denoted by $V_M(\mathcal{G}_M)$ in this work. In fact, with the objective of calculating payoffs, it is unrealistic and unnecessary for an individual user to obtain the knowledge of the whole structure of the cooperation formation network. Instead, the only information that a user needs to know is who have cooperation relationship with it, which can be simply implemented in reality.

For ease of presentation, \mathcal{G}_M can be represented by an undirected graph, where a link between two nodes implies the cooperation relationship between them. For instance, if MU_m and MU_n have formed a link, denoted by $l_{m,n} \in \mathcal{G}_M$, they are willing to help each other to forward contents. Let $\mathcal{G}_M + l_{m,n}$ denote the graph obtained by adding link $l_{m,n}$ to the existing graph \mathcal{G}_M , where $l_{m,n} \notin \mathcal{G}_M$ and $\mathcal{G}_M - l_{m,n}$ denote the graph obtained by removing link $l_{m,n}$ from the existing graph \mathcal{G}_M where $l_{m,n} \in \mathcal{G}_M$, i.e., $\mathcal{G}_M + l_{m,n} = \mathcal{G}_M \cup \{l_{m,n}\}$ and $\mathcal{G}_M - l_{m,n} = \mathcal{G}_M \setminus \{l_{m,n}\}$ [18]. A cooperation link can be established only if the two end nodes of the link agree to collaborate with each other. However, any MU is assumed to have the discretion to unilaterally terminate cooperation relationships that they are currently involved in, so long as it can benefit more from breaking up with such relationships. Note that the cooperation relationship defined here is symmetric but not transitive. For instance, it is possible that MU_1 is willing to cooperate with MU_2 , MU_2 chooses to cooperate with MU_1 and MU_3 simultaneously, while MU_1 and MU_3 may have no cooperation relationship.

With the notions defined above, the author applies the notion of pairwise stability

proposed by Jackson and Wolinsky [18] to define what a stable network is. The network formed by the mobile users, \mathcal{G}_M , is pairwise stable with respect to utility function u_m if both of the following conditions are satisfied:

1. None of the users has an incentive to remove a link that it is involved in.
2. For any link not existed in the pairwise-stable network, if one user can benefit from the addition of the link, the other user involved will definitely suffer from adding the link.

To achieve a pairwise stable network, dynamic models are applied [18]. Specifically, the network formation process that the author defines here is time-dependent, i.e., the payoffs received by the users vary over time. User makes decision of cooperation formation periodically at iteration $t = 1, 2, \dots$, resulting in varying network structures. Hence, for a more proper notation, $\mathcal{G}_M(t)$ is used to denote the cooperation network structure at time t . Network $\mathcal{G}_M(t+1)$ is considered to be adjacent to network $\mathcal{G}_M(t)$ if $\mathcal{G}_M(t+1)$ can be formed from $\mathcal{G}_M(t)$ by adding link $l_{m,n}$ or removing link $l_{m,n}$, for some link $l_{m,n}$ [18]. Following Watts' idea, the author also defines the notion of defeat that, network $\mathcal{G}_M(t+1)$ defeats network $\mathcal{G}_M(t)$ if either adding link $l_{m,n}$ or removing link $l_{m,n}$ to network $\mathcal{G}_M(t)$ will not reduce the payoffs of both node m and node n , while at least one of these two nodes can achieve higher payoff in network $\mathcal{G}_M(t+1)$. Based on this notion, improving path is introduced as follows: an improving path is a series of intermediate networks $\{\mathcal{G}_M(1), \mathcal{G}_M(2), \dots, \mathcal{G}_M(T)\}$, where each network $\mathcal{G}_M(t+1)$ (adjacent to the previous one $\mathcal{G}_M(t)$) defeats $\mathcal{G}_M(t)$. For an implementation purpose, the author applies the distributed improving path method proposed by Jackson and Watts to obtain the pairwise-stable network [18].

5.2.2 Cooperation among the Infostations

Next, the author investigates the upper-level cooperation formation, with the objective to form the most beneficial coalitional structure for the infostations. The cooperation formation among the infostations is formulated as a *non-transferrable-utility* (NTU) coalitional game, which is denoted by $G_I = (\mathbb{N}; V_I)$, where \mathbb{N} is the set of players and V_I is a mapping from an arbitrary coalitional structure \mathcal{S}_I to $V_I(\mathcal{S}_I)$, which is a vector of payoffs that the infostations can achieve under coalitional structure \mathcal{S}_I . A formal definition is given as follows:

$$\begin{aligned} \max_{\mathcal{S}_I} \quad & U_{IS_i}(\mathcal{S}_I, \mathcal{G}_M) \\ \text{s. t.} \quad & \bigcup_{j=1}^J \mathcal{S}_I^j = \mathbb{N}, \quad \mathcal{S}_I^j \cap \mathcal{S}_I^{j'} = \emptyset \quad (j \neq j', 1 \leq j, j' \leq J) \end{aligned}$$

where $\mathcal{S}_I = \{\mathcal{S}_I^1, \dots, \mathcal{S}_I^j, \dots, \mathcal{S}_I^J\}$ is the coalitional structure of the ISs, e.g., if $\mathcal{S}_I = \{\{IS_1, IS_2\}, \{IS_3\}\}$, then $\mathcal{S}_I^1 = \{IS_1, IS_2\}$, and $\mathcal{S}_I^2 = \{IS_3\}$. The number of possible coalitional structures for N ISs is given by the Bell number D_N , where

$$D_i = \sum_{j=0}^{i-1} \binom{i-1}{j} D_j, \quad 1 \leq i \leq N, \quad \text{and} \quad D_0 = 1. \quad (5.4)$$

The author uses $\mathcal{S}_{I,x}$ to represent a certain coalitional structure of the ISs. Consider the example shown in Fig. 5.1, the coalitional structures of three ISs are defined as follows: $\mathcal{S}_{I,1} = (\{IS_1\}, \{IS_2\}, \{IS_3\})$, $\mathcal{S}_{I,2} = (\{IS_1, IS_2\}, \{IS_3\})$, $\mathcal{S}_{I,3} = (\{IS_1, IS_3\}, \{IS_2\})$, $\mathcal{S}_{I,4} = (\{IS_1\}, \{IS_2, IS_3\})$ and $\mathcal{S}_{I,5} = (\{IS_1, IS_2, IS_3\})$. Note that the number of possible coalitional structures will increase dramatically when the number of ISs becomes large. To address this complexity problem, the approximation and reduction method proposed in [17] can be applied, which is not discussed in this work.

Assuming that the infostations are rational to maximize their individual payoffs, the coalitional structure \mathcal{S}_I satisfying the conditions for both internal stability and external stability is considered to be the stable solution of the proposed coalitional game [103].

- *Internal stability:* A coalition \mathcal{S}_I is internally stable if no infostation is able to improve its payoff by splitting from its current coalition and acting independently, i.e, no splitting process can occur.
- *External stability:* A coalition \mathcal{S}_I is externally stable if at least one infostation will be harmed from merging one coalition with another coalition, i.e., no merging process can occur.

The decision of an infostation to join or leave a coalition depends on how the payoff may vary. Given coalitional structure \mathcal{S}_I , the actions of IS_i can be as follows [103]:

- *Splitting:* As a result of splitting process, one coalition \mathcal{S}_I^j in \mathcal{S}_I can be partitioned into multiple coalitions $\mathcal{S}_I^{j'\dagger}$, resulting in a new coalitional structure \mathcal{S}_I^\dagger . This operation can take place only if all the affected infostations will not achieve lower payoffs than before, and at least one of them can achieve higher payoff, i.e.,

$$U_{IS_i}(\mathcal{S}_I^\dagger) \geq U_{IS_i}(\mathcal{S}_I), \quad \forall i \in \mathcal{S}_I^j = \bigcup \mathcal{S}_I^{j'\dagger}. \quad (5.5)$$

where at least one of the inequalities should not be met with the equal sign.

- *Merging:* As a consequence of merging process, multiple coalitions \mathcal{S}_I^j in \mathcal{S}_I may jointly agree to form a new single coalition $\mathcal{S}_I^{j'\dagger}$, resulting in a new coalitional structure \mathcal{S}_I^\dagger . This operation can happen only if all affected infostations will not achieve lower payoffs than before, and at least one of them can achieve higher

payoff, i.e.,

$$U_{IS_i}(\mathcal{S}_I^\dagger) \geq U_{IS_i}(\mathcal{S}_I), \quad \forall i \in \mathcal{S}_I^{j'\dagger} = \bigcup \mathcal{S}_I^j, \quad (5.6)$$

where at least one of the inequalities should not be met with the equal sign.

5.2.3 Stackelberg Equilibrium and Proposed Distributed Algorithm

Combining the lower-level and upper-level cooperation formation processes, the author proposes a notion named as the Stackelberg equilibrium to define the solution of the hierarchical cooperation formation game. The Stackelberg equilibrium is defined as the solution of the hierarchical cooperation formation game, denoted by $\{\mathcal{S}_I^*, \mathcal{G}_M^*\}$, at which all ISs stick to coalitional structure \mathcal{S}_I^* , so that their payoffs are maximized. Meanwhile, under the coalitional structure of the infostations \mathcal{S}_I^* , the mobile users can also benefit most by forming cooperation relationships described by a pairwise stable network \mathcal{G}_M^* .

To achieve the Stackelberg equilibrium of the hierarchical cooperation formation game, a distributed algorithm based on the merge and split mechanisms and Jackson and Watts's network formation method is adopted in the proposed algorithm. Note that the network formation process always starts from an empty network, where no link is formed. A common problem during such process is that, one might not be able to escape from an initial state, i.e., the empty network. Consider a situation that for an arbitrary pair of the nodes, forming a cooperation link cannot benefit both of them simultaneously, then none of the nodes would like to cooperate with each other due to their rationality, resulting in an empty network as the steady state. To solve

this problem, some random perturbations in the procedure of network formation is introduced, i.e., the nodes are allowed to make some irrational choices with a small probability. With such randomness, the network formation process is able to visit any category of network, some of which are more likely to be formed (driven by the intension of the nodes), while some can just be reached through a consequence of errors.

The distributed hierarchical cooperation formation algorithm of infostations and mobile users is presented in Algorithm 2. For ease of illustration, the author also makes a flowchart according to Algorithm 2, which is described in Fig. 6.7. The network formation process of the users and merge-and-split process of the infostations are explained in the following. For the network formation process of the mobile users, the author first randomly identifies a link $l_{m,n}$. If the selected link $l_{m,n}$ does not exist in the current network $\mathcal{G}_M(\tau)$, the author evaluates whether adding this link will bring higher payoffs to the two end nodes (MU_m and MU_n) simultaneously. If this is true, link $l_{m,n}$ is added to $\mathcal{G}_M(\tau)$, which leads to a new network $\mathcal{G}_M(\tau + 1)$. Otherwise, link $l_{m,n}$ may also be added if $rand() \leq \varepsilon$. Note that $0 \leq rand() \leq 1$ is a random number generator function, and ε denotes the small probability of making irrational choice of the nodes. On the other hand, if the selected link $l_{m,n}$ already exists in the current network $\mathcal{G}_M(\tau)$, the author evaluates whether removing this link will achieve higher payoffs to both MU_m and MU_n or not. If this is true, link $l_{m,n}$ is removed from $\mathcal{G}_M(\tau)$, which leads to a new network $\mathcal{G}_M(\tau + 1)$. Otherwise, again link $l_{m,n}$ may also be removed if $rand() \leq \varepsilon$. For the merge-and-split process of the infostations, if the merging process will achieve higher payoffs to the infostations affected by this operation, the coalitions are merged. On the other hand, if splitting process will achieve higher payoffs to the infostations affected by this operation, the coalition is split. Note that the coalitions may also be merged/split if $rand() \leq \varepsilon$, representing the case of making irrational choices. After a series of merge and split operations, the

Algorithm 2 Distributed hierarchical cooperation formation algorithm of ISs and MUs in mobile infostation networks.

```

1: Initialize iteration  $t = 0$  and the coalitional structure of ISs  $\mathcal{S}_I(t) = \{\dots, \mathcal{S}_I^j(t), \dots\}$ 
2: loop
3:   repeat
4:     Initialize  $\tau = 0$ , and cooperation network structure of MUs  $\mathcal{G}_M(\tau) = \{\dots, l_{m,n}(\tau), \dots\}$ 
5:     for  $m \in \mathcal{M}$  do
6:       MU  $m$  computes payoff  $u_m$  from Eq. (5.1)
7:     end for
8:     Randomly identify a link  $l_{m,n}$ 
9:     if  $l_{m,n} \notin \mathcal{G}_M(\tau)$  then
10:      if  $(u_h(\mathcal{S}_I(t), \mathcal{G}_M(\tau) + l_{m,n}) > u_h(\mathcal{S}_I(t), \mathcal{G}_M(\tau))$  for both  $h \in \{m, n\}$ ) OR  $(rand() \leq \varepsilon)$  then
11:         $\mathcal{G}_M(\tau + 1) \leftarrow \mathcal{G}_M(\tau) + l_{m,n}$ 
12:      end if
13:    end if
14:    if  $l_{m,n} \in \mathcal{G}_M(\tau)$  then
15:      if  $(u_h(\mathcal{S}_I(t), \mathcal{G}_M(\tau) - l_{m,n}) > u_h(\mathcal{S}_I(t), \mathcal{G}_M(\tau))$  for both  $h \in \{m, n\}$ ) OR  $(rand() \leq \varepsilon)$  then
16:         $\mathcal{G}_M(\tau + 1) \leftarrow \mathcal{G}_M(\tau) - l_{m,n}$ 
17:      end if
18:    end if
19:     $\tau = \tau + 1$ 
20:  until Steady cooperation structure of MUs is reached
21:  for  $i \in \mathcal{N}$  do
22:     $IS_i$  computes its payoff  $U_{IS_i}$  from Eq. (5.2)
23:  end for
24:  Merge mechanism of the ISs
25:  if (Condition in Eq. (5.6) is satisfied) OR  $(rand() \leq \varepsilon)$  then
26:    Merge coalitions and  $\mathcal{S}_I(t + 1) \leftarrow \mathcal{S}_I^\dagger$ 
27:  end if
28:  Split mechanism of the ISs
29:  if (Condition in Eq. (5.5) is satisfied) OR  $(rand() \leq \varepsilon)$  then
30:    Split the coalition and  $\mathcal{S}_I(t + 1) \leftarrow \mathcal{S}_I^\ddagger$ 
31:  end if
32:   $t = t + 1$ 
33: end loop

```

set of infostations converges to a partition of disjoint coalitions, in which none of the coalitions has an incentive to perform further merge or split process. In this case, the partition is considered to be *merge-and-split proof*. However, due to the randomness existed in the proposed algorithm, running the algorithm different times may result in different stable coalitional structures. In this case, the author runs the algorithm for a large amount (e.g., 10^5) of times, and all possible stable coalitional structures and their corresponding probabilities (namely stationary probabilities) can be obtained. The stationary probability distribution among the possible stable coalitional structures is an important metric to analyze the result of the proposed algorithm.

A problem rises up in the proposed algorithm when the number of the participating nodes becomes large, since there can be a great number of possible links need to be evaluated, resulting in the inefficiency of the algorithm. However, this problem can be relieved if it is assumed that only those users having high contact rates are possible to cooperate with each other. In other words, the number of possible links need to be evaluated can be significantly reduced by proposing a notion called “cooperation threshold”, such that if the contact rate between two users falls below this threshold, it is assumed that they will never build up a cooperation link, and only if their contact rate is higher than the threshold, they are allowed to become possible partners. This assumption also coincides with our intuition that only friends who meet frequently are likely to forward content for each other.

Note that the convergence of any iteration composed of successive merge and split operations, as presented in the proposed algorithm description, is guaranteed as shown in [104]. Another worth noting point, the proposed hierarchical cooperation formation framework actually consists of two stages. During the first stage, the infostations and mobile users form cooperation structures through an iteration of the proposed algorithm until termination. During the second stage, the content downloading procedure

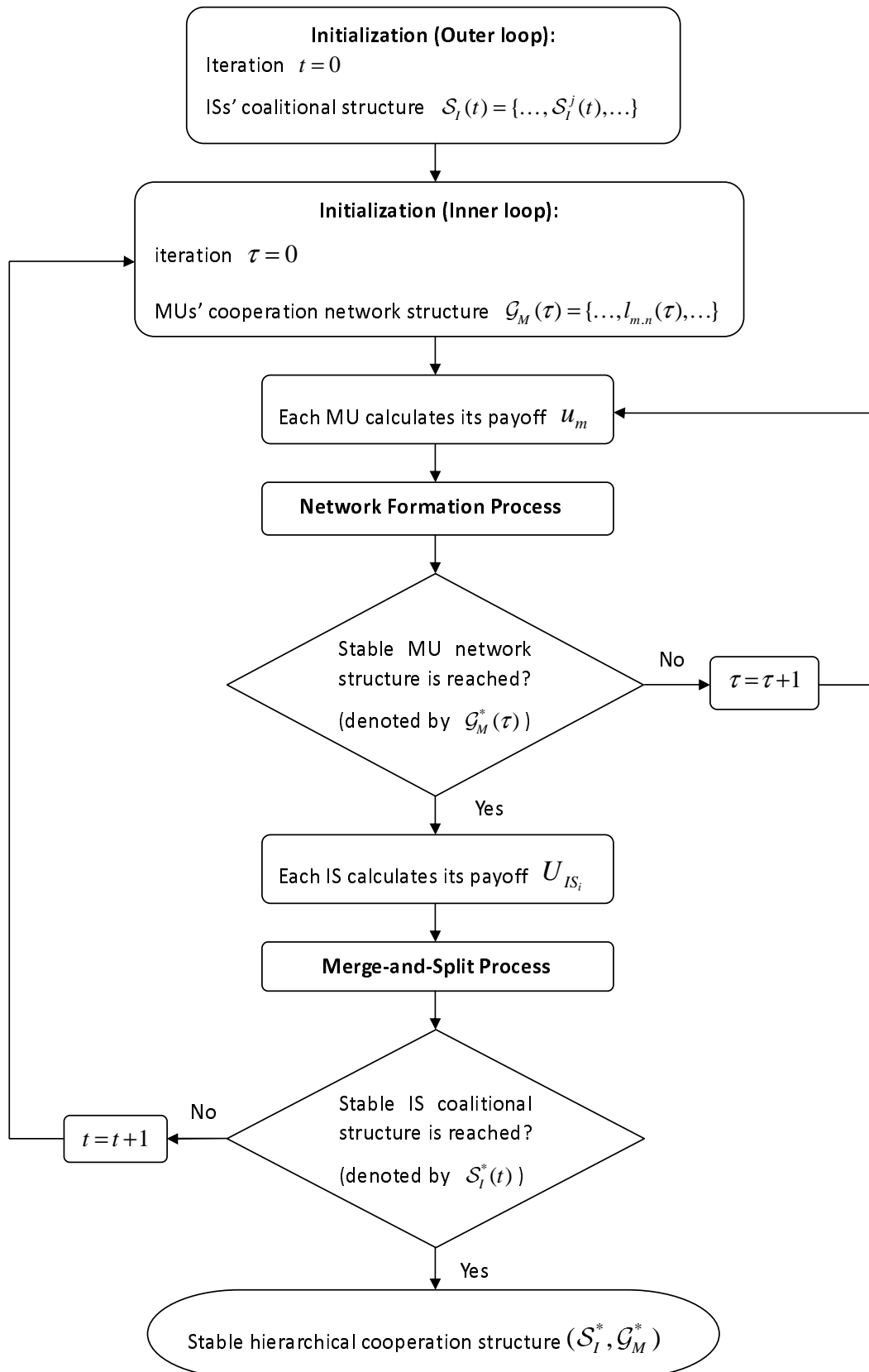


Figure 5.2: Flowchart of the proposed distributed hierarchical cooperation formation algorithm.

takes place under the stable cooperation structure obtained in the first stage. Finally, the proposed algorithm can also enable the infostations or mobile users to make decisions distributively to adapt their cooperation formation to environment changes such as the variations of the contact rates between them. In this case, the proposed algorithm can be repeated when the system parameters change, permitting the infostations and mobile users making dynamic decisions on forming cooperation relationships to maximize their benefits in different situations.

5.3 Performance Evaluation

In this section, the author evaluates the performance of the proposed algorithm and investigate the impact of different system parameters on the stable cooperation formation result. The numerical results are based on the network scenario shown in Fig. 5.1. For the parameter settings, ε is set to be 10^{-3} , the unit cost for cooperation among infostations and mobile users are set to be 0.2 and 0.05, respectively.

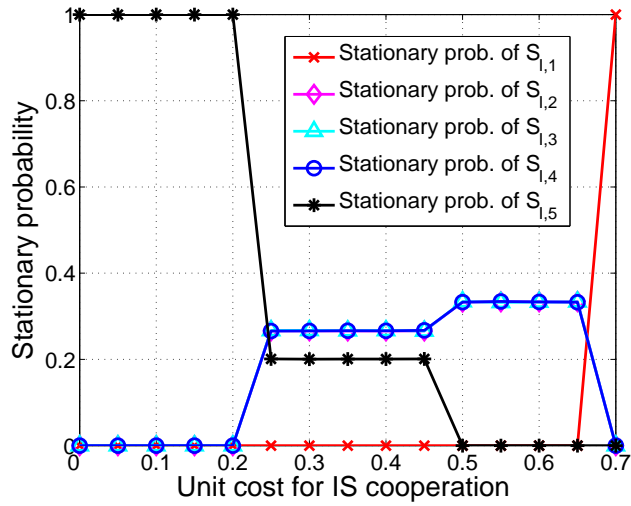
Firstly, the cooperation formation of the infostations obtained from using the proposed algorithm is discussed. As an example, the author sets $\mu_{12} = \mu_{13} = \mu_{34} = \mu_{45} = \mu_{56} = \mu_{26} = 0.7$. Here the other contact rates are not considered since they are set to be below a certain cooperation threshold. As a result, no cooperation can be formed. With the fixed user-user contact rates, the author considers two scenarios with different user-infostation contact rates shown as follows:

- Scenario 1: The contact rates $\lambda_{mi} = 0.6$ ($m = 1, \dots, M, i = 1, \dots, N$). To be representative, this scenario describes a uniform case that the mobile users have the same mean contact rate with all infostations.
- Scenario 2: $\lambda_{11} = \lambda_{31} = \lambda_{22} = \lambda_{42} = 0.8, \lambda_{12} = \lambda_{21} = \lambda_{32} = \lambda_{41} = 0.4,$

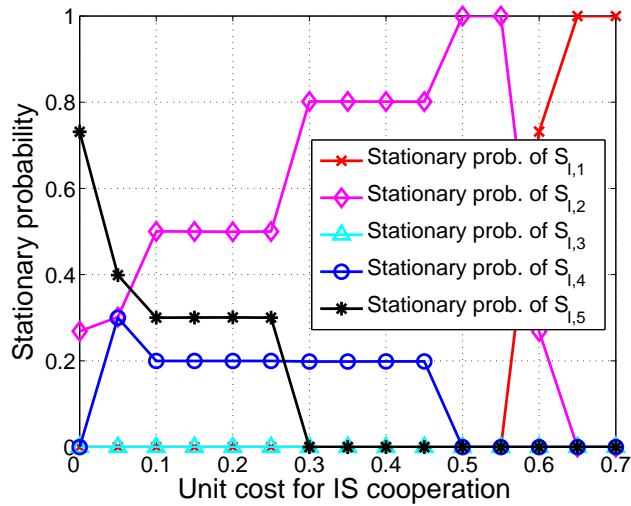
$\lambda_{53} = \lambda_{63} = 0.8$, $\lambda_{52} = \lambda_{62} = 0.4$. In this scenario, subscribers of IS_1 may contact frequently with IS_2 and those of IS_2 may also contact with IS_1 frequently. Subscribers of IS_3 have high contact rates with IS_3 . As a result, they may seldom need cooperation from IS_1 or IS_2 .

For a more comprehensive discussion, the author varies the unit cost of forming coalition among infostations and investigate the stationary probability distribution of the coalitional structures under different costs. Note that in the implementation, it is assumed that at most two coalitions can merge into a single coalition at one time. Similarly, one coalition can be split into at most two smaller coalitions at one time. This assumption is due to the consideration for reducing the complexity, the author intends to remove this constraint and investigate the numerical results in the future work. As shown in Fig. 5.3(a), it can be seen that in scenario 1, all infostations will cooperate (i.e., the grand coalition $\mathcal{S}_{I,5}$ is formed) to provide the content downloading service to their subscribers, when there is no cost for forming coalition or the cost is relatively low, i.e., below 0.2 in the example. However, as the unit cost increases, the grand coalition is no longer the most beneficial choice because of the high cost for maintaining it. Alternatively, each infostation may only form coalition with one other infostation at most. Since a uniform scenario is considered in scenario 1, it is expected that the variations of the stationary probabilities of coalitional structures $\mathcal{S}_{I,2} - \mathcal{S}_{I,4}$ are nearly the same, as shown by the overlapping curves in Fig. 5.3(a). As the unit cost continually increases to 0.7, the benefit of forming coalition among infostations cannot compensate the cost that has to be paid. Therefore, all infostations will choose to act independently.

In scenario 2, besides similar conclusions obtained for scenario 1, there are some important observations different from scenario 1, as shown in Fig. 5.3(b): (1) since subscribers of IS_1 (IS_3) have no regular contact with IS_3 (IS_1), the payoff improvement



(a) Stationary probability of the ISs' coalitional structures in Scenario 1 vs. the unit cost for IS cooperation



(b) Stationary probability of the ISs' coalition structures in Scenario 2 vs. the unit cost for IS cooperation

Figure 5.3: Stationary probability of the ISs' coalitional structures vs. the unit cost for IS cooperation

brought by the cooperation between these two infostations may be relatively limited. As a result, the stationary probability (calculated by the statistical result, which is obtained from a large number of times of running the algorithm) of $\mathcal{S}_{I,3}$ is always 0. This reason can also explain why a grand coalition is not always most beneficial even when forming coalition introduces no cost; (2) according to the mean contact rate information in scenario 2, the author may intuitively suppose that IS_1 and IS_2 are helpful to each other, while IS_3 should prefer to act independently or occasionally cooperate with IS_2 . The simulation result perfectly coincides with such expectation, note that coalitional structure $\mathcal{S}_{I,2}$ always dominates the infostations' decisions on forming coalitions when the unit cost for infostation cooperation exceeds 0.1. The advantage of $\mathcal{S}_{I,2}$ can also be verified in Fig. 5.4. It is observed that both IS_1 and IS_2 can have the significant payoff improvement when they form a coalition, while IS_3 is left to act independently. Although IS_3 can achieve higher payoff when forming coalition with IS_2 , IS_2 's payoff will be harmed at the same time. Therefore, IS_2 will intelligently choose IS_1 rather than IS_3 to form a coalition, and coalitional structure $\mathcal{S}_{I,2}$ is stable since none of the infostations can achieve higher payoff by unilaterally switching its strategy, with the payoffs of the other infostations unaffected.

Next, the cooperation network formation of the mobile users obtained from using the proposed algorithm is discussed. As shown in Fig. 5.5, for scenario 1, the grand coalition of the infostations is taken as an example to show how the users will form cooperation networks under a certain coalitional structure of the infostations. For scenario 2, take $\mathcal{S}_{I,2}$ is taken as an instance, which is the dominating coalitional structure. Since the unit cost of user cooperation will evidently affect the formation of the cooperation network, the author compares the network formation results when the unit cost for user cooperation is 0.05 and 0.1 for both cases.

In scenario 1, it can be seen that each user will choose a partner to cooperate

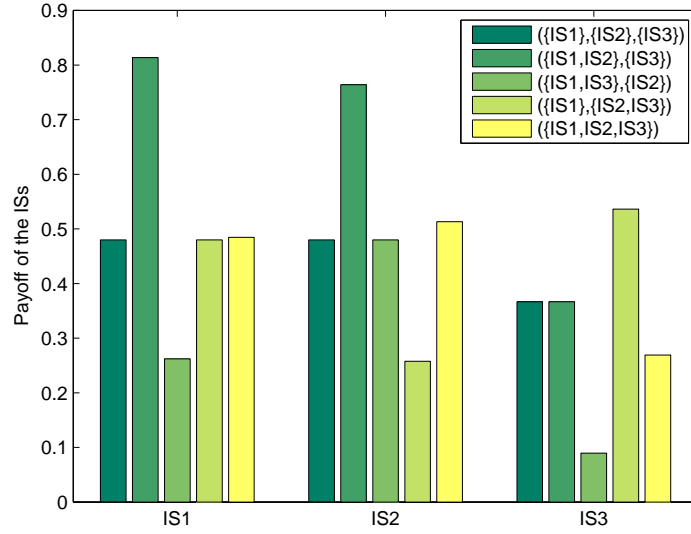


Figure 5.4: Comparison of the ISs' payoffs under different coalitional structures.

with at stable state when the unit cost for user cooperation is 0.05, as described by forming a link with another user in the network graph as shown in Fig. 5.5. Note that in the parameter settings, users seldom choose to form multiple links because of the additional cost for maintaining more cooperation relationships. When the unit cost for user cooperation is 0.1, none of the users will cooperate with each other, resulting in an empty network. This is because users have already gained enough benefit from the infostations' grand coalition, and forming cooperation network among themselves will not provide further payoff improvement because of the high cost introduced by cooperation. In scenario 2, the author has similar observations except some differences. Note that even when the unit cost for user cooperation is 0.1, the link between MU_5 and MU_6 always exists since IS_3 does not form coalition with either IS_1 or IS_2 , which may stipulate its subscribers to persist in cooperating with each other for higher individual payoffs. When the unit cost for user cooperation is 0.05, the improvement of forming cooperation network among the users based on the coalitional structure of

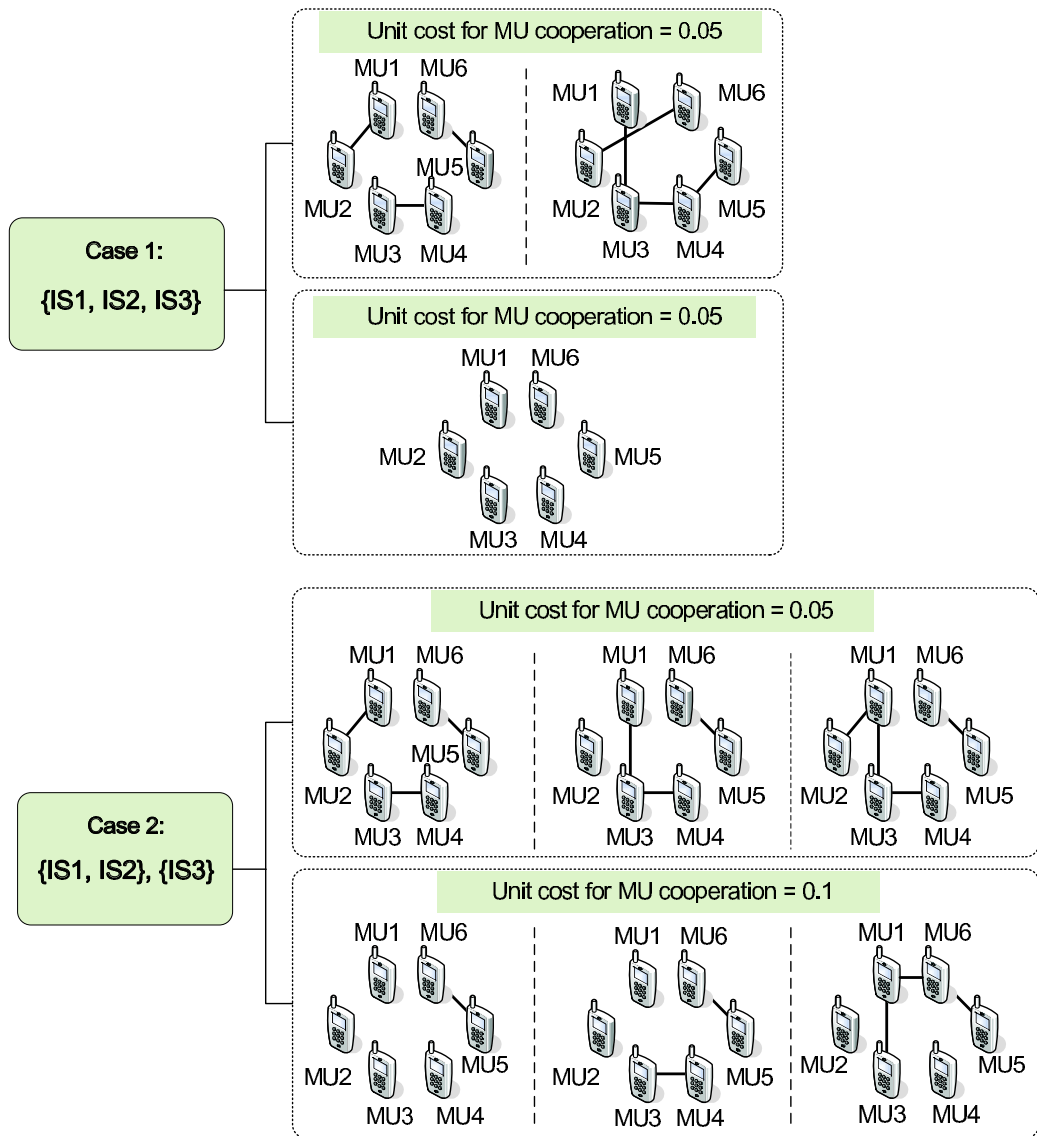


Figure 5.5: Stable MU cooperation network formation result.

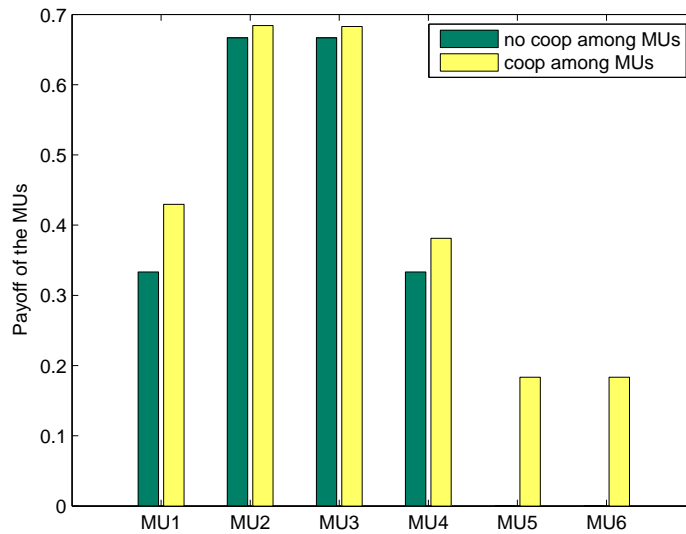


Figure 5.6: Comparison of the MUs' payoffs with/without forming cooperation network.

the infostations can be illustrated by Fig. 5.6. It can be seen that on the basis of the improvement brought by the coalition formation of the infostations, the cooperation network formed among the mobile users can indeed achieve further improvement of the individual payoffs of the mobile users, so long as the cost needed to forming such a cooperation network does not overwhelm the benefit.

5.4 Summary

In this chapter, the author has proposed a hierarchical cooperation formation model for downlink data transmission in mobile infostation networks. At the lower level, given the coalitional structure of the infostations, the mobile users can intelligently form a stable cooperation network to maximize their individual payoffs. At the upper level, the infostations can make full use of the knowledge about the mobile users' network formation strategies and intelligently decide on forming coalitions to achieve the highest

benefits for themselves. A distributed hierarchical cooperation formation algorithm has been proposed to solve this Stackelberg-fashion problem and to obtain the stable cooperation structure for both infostations and mobile users. The effectiveness of the algorithm has been extensively verified by the numerical results. Moreover, the impact of different system parameters on the stable cooperation structure has also been investigated.

Chapter 6

Content Delivery in Relay-Based Publish-Subscribe Networks

In this chapter, the author considers the problem of content delivery in relay-based publish-subscribe (pub-sub) networks, and proposes contract theoretic model to investigate the interaction between the content provider and the mobile subscribers, with the purpose of guiding the two interest parties to achieve an agreement which can benefit them both. The author starts this chapter from describing the proposed system model in Section 6.1, where the network model of the relay-based pub-sub networks is presented, and the queueing model of the content delivery process in the network is introduced. The content provider/subscriber model, along with the optimal contract design for the content provider are provided in Section 6.2. In Section 6.3, the author presents the performance evaluation of the proposed queueing model and numerical results on the optimal contract design. Finally, Section 6.4 summarizes this chapter.

6.1 System Model

In this section, the system model of content delivery in a relay-based publish-subscribe network is presented. After describing the example network scenario, the queueing model constructed among the network components is introduced to analyze the service quality of content delivery in the proposed network model.

6.1.1 Relay-Based Publish-Subscribe Network

Consider the scenario that a content provider (CP) needs to deliver contents to users scattered around it, including users nearby and users far away. To achieve this goal, a relay-based publish-subscribe model is adopted. With the assistance of relays such as mobile vehicular carriers equipped with wireless transmission and receiving devices, content distribution from the data source to the users at different locations becomes possible.

As shown in Fig. 6.1, the author considers a relay-based publish-subscribe network, which consists of the following components:

- Content provider: In the proposed model, the content provider is the data source, publishing desired contents (e.g., news updates) to its subscribed mobile users. It can transmit the contents directly to the users in the transmission range of the base station.
- Mobile users (Subscribers): The mobile users subscribe contents from the content provider, some may contact often with the content provider, while others may seldom contact with the content provider and experience long periods of disconnection. Note that in the proposed model, a contact between a pair of nodes is considered to be occurred when a wireless node enters the transmission range of

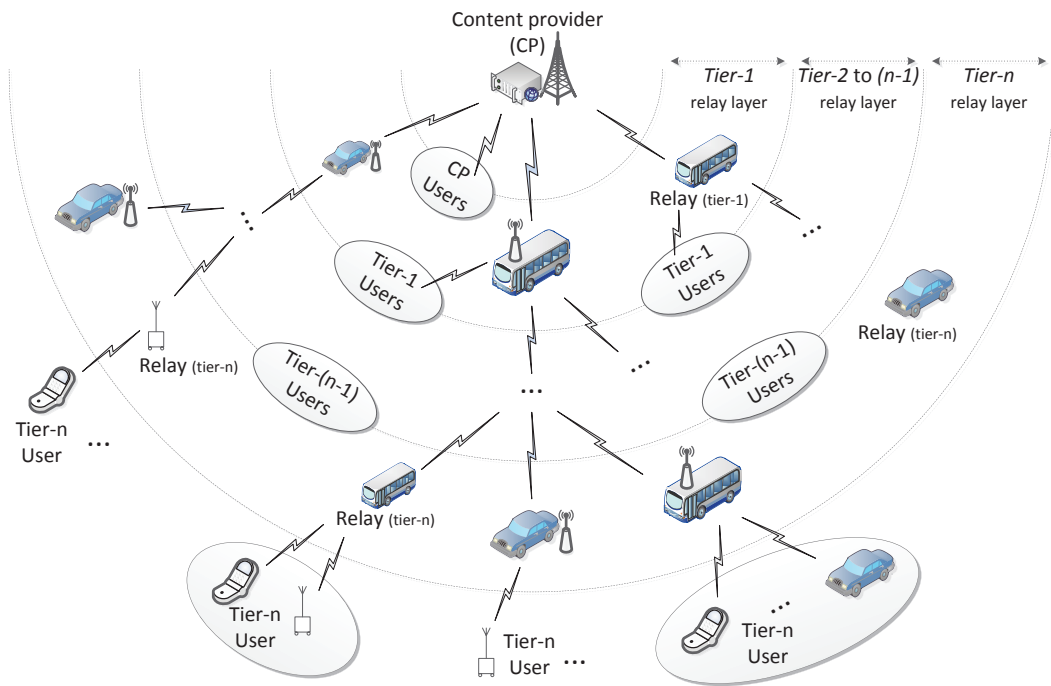


Figure 6.1: An example scenario of relay-based publish-subscribe network.

another wireless node.

- Relays: The content provider faces a problem when it needs to deliver contents to the remote users, and thus cannot get the contents in time through direct transmission. To tackle this problem, the content provider can employ vehicles or particular mobile devices equipped with short-range wireless devices and storage to carry the data, and let them help with the content delivery, i.e., they may enable indirect connectivity among subscribers at different areas. Such intermediary wireless nodes bridging the content provider and the mobile users are recognized as relays.

The distance between the content provider and the subscribers could be large that multiple relays are required to relay contents. In this case, the relays can be marked by different tiers according to the sequence they receive the content. For instance, the relays directly contacting with the content provider are defined as

Tier-1 relays, and the relays which relay by carrying-and-forwarding the data for the Tier-1 relays are Tier-2 relays, so on and so forth. In the last hop, the Tier- n relays transmit the data to the subscribers furthest from the content provider. The users directly connected with the content provider are defined as the CP users, while the users connected with Tier- n relays are defined as Tier- n users.

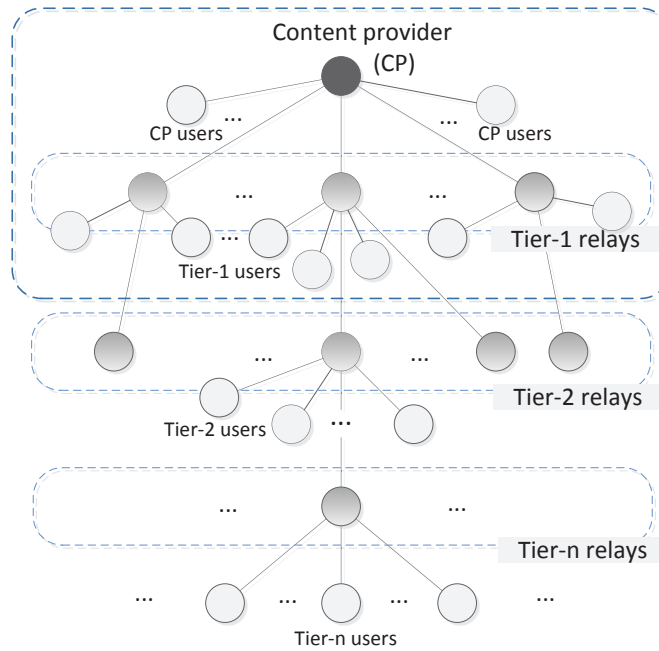


Figure 6.2: The corresponding tree structure of relay-based publish-subscribe network.

For clearer presentation, a tree structure shown in Fig. 6.2 is used to describe the relay-based publish-subscribe network. The content source, the intermediate nodes, and the destination consist of a content delivery path. Since each content has a life cycle, indicating that if the content delay exceeds some deadline, it will be useless, or the value of the content will decrease to a quite low value. The author uses the delivery ratio within deadline (DRWD) to measure the quality of the content delivery service, which is defined as:

$$\eta = P(t \leq T). \quad (6.1)$$

where T is the deadline of the content, and $P(t \leq T)$ denotes the probability that the content can be delivered to the user within deadline T . In the following section, the author applies queueing model to analyze the DRWD of the content. In this system model, the bandwidth between a transmitter and receiver is allocated in such a way that with a unit bandwidth, one content can be transmitted through one contact between the transmitter and the receiver. Thus, the number of contents can be transmitted simultaneously through one contact grows linearly with the bandwidth allocated by the transmitter. Note that given the contact rates, in order to provide service quality η to the subscriber, the content provider has to guarantee a certain bandwidth provision to the subscriber directly if the subscriber is a CP user (as shown in Fig. 6.2). Otherwise, it has to guarantee a certain bandwidth provision at its base station and each relay on the content delivery path from the content provider to the subscriber. Here the content provider has to pay for the required bandwidth at its base station and at the intermediate relays. Due to the cost introduced by bandwidth provision, the content provider has to charge the subscribers for requiring a certain service quality η .

6.1.2 Queueing Model

Considering the content arrival and departure at each node, the content delivery along a certain path can be modeled as a tandem of queues. For the simplicity of presentation, the author investigates the tandem queueing model for the content delivery from the content provider to the Tier-1 mobile users. The model can be extended to content delivery paths from the content provider to Tier-n mobile users straightforwardly.

As shown in Fig. 6.3, Q_1 is the queue of contents at the content provider, while Q_2 is the queue of contents at the Tier-1 relay R_1 . We take a typical content delivery path Q_1-Q_2 as an instance. The contents are injected into the queue at the content provider.

It is assumed that the content generation is a Poisson process with average rate λ [105]. Note that the departure of contents from Q_1 occurs only when the Tier-1 relay R_1 contacts with the content provider. Generally, the time period between two consecutive meetings of the Tier-1 relay and the content provider is assumed to be exponentially distributed [105]. In the example, let μ_1 denote the contact rate between the Tier-1 relay R_1 and the content provider, then the time period between two consecutive contacts is an exponentially distributed variable with mean $1/\mu_1$. Therefore, R_1 can be considered as the server of Q_1 with service rate μ_1 if given a unit bandwidth. Similarly, subscriber S_1 can be considered as the server of Q_2 with service rate μ_2 if given a unit bandwidth, where μ_2 denotes the contact rate between subscriber S_1 and Tier-1 relay R_1 .

It needs to be emphasized here that although there are multiple relays, they cannot be considered as multiple servers in the queueing model, since when each relay contacts with the content provider and obtains the contents, it is actually not taking contents away from the queue of the content provider, since the contents are still available for other relays. For the simplicity of representation, it can be considered that the content provider has multiple virtual queues corresponding to different relays, each of which takes content away from its associated queue. Similar reason can be used to explain the content delivery from a particular Tier-1 relay to multiple connected mobile users. In this case, the content delivery from the content provider to different Tier-1 relays can be considered as independent, as well as that from a particular Tier-1 relay to the Tier-1 mobile users.

In the example scenario, there are two cases of the content delivery service: (1) content delivery to subscribers directly connected to the content provider (referred to as Case I), and (2) content delivery to the Tier-1 subscribers (referred to as Case II). Thus the author derives the DRWD for these two categories of services, respectively.

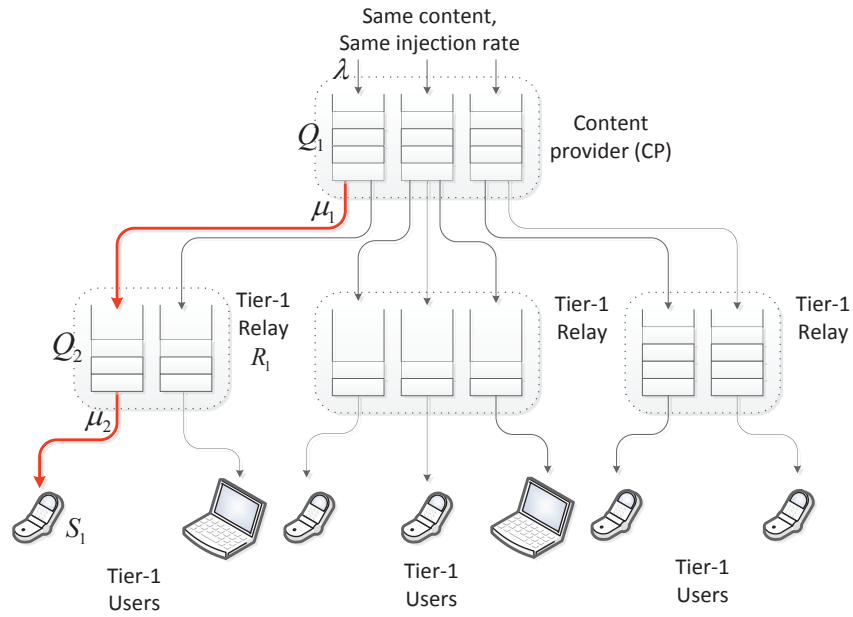


Figure 6.3: Queue structure.

- Case I:

For the content delivery service from the content provider to the subscribers directly connected to it, the state transition of the corresponding Markovian queue with bulk services is shown in Fig. 6.4. In this figure, r_1 is the bulk size, indicating that r_1 units of bandwidth is allocated by the content provider. In other words, the subscriber can obtain up to r_1 contents through one contact with the content provider.

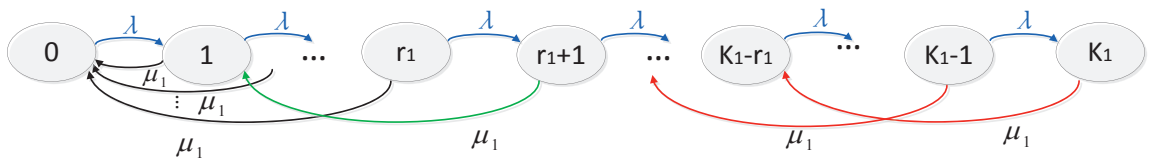


Figure 6.4: State-transition-rate diagram of one-dimensional Markovian queue with bulk service.

Let $\pi(k)$ denote the stationary probability that the queue contains k contents in its buffer space. According to the state-transition-rate diagram shown in Fig. 6.4,

the balancing equations can be obtained as follows:

$$\pi(0) \cdot \lambda = \sum_{k=1}^{r_1} \pi(k) \cdot \mu_1, \quad (6.2)$$

$$\pi(k) \cdot (\lambda + \mu_1) = \pi(k-1) \cdot \lambda + \pi(k+r_1) \cdot \mu_1, \quad (1 \leq k \leq K_1 - r_1) \quad (6.3)$$

$$\pi(k) \cdot (\lambda + \mu_1) = \pi(k-1) \cdot \lambda, \quad (K_1 - r_1 < k < K_1) \quad (6.4)$$

$$\pi(K_1) \cdot \mu_1 = \pi(K_1 - 1) \cdot \lambda. \quad (6.5)$$

By introducing the normalizing equation $\sum_{k=0}^{K_1} \pi(k) = 1$, the above balancing equations can be solved. Consequently, the stationary probabilities $\pi(k)$ ($0 \leq k \leq K_1$) can be obtained.

Next, the author derives the time required for a particular content to be transmitted to a subscriber, i.e., the content delay. A content has to wait until all contents ahead of it have been transmitted, i.e., the number of contents ahead of it becomes zero. The author models an absorbing Markov chain with 0 (denoting the situation that there is no content ahead of the concerned one) being the absorbing state and other states (state k ($k > 0$) denotes that there are k contents ahead of the tagged one) being the transient states. The content delay is the absorption time of the considered absorbing Markov chain. Take $r_1 = 2$ as an example, then based on the proposed queueing model, the transition rate matrix describing the transition rates among all states (including the absorbing

state and the transient states) is shown in the following:

$$\tilde{\mathbf{S}} = \begin{bmatrix} 1 & 0 & 0 & \cdots & \cdots & \cdots & 0 \\ \mu_1 & -\mu_1 & 0 & \cdots & \cdots & \cdots & 0 \\ \mu_1 & 0 & -\mu_1 & \cdots & \cdots & \cdots & 0 \\ 0 & \mu_1 & 0 & -\mu_1 & \cdots & \cdots & 0 \\ 0 & 0 & \mu_1 & 0 & -\mu_1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \ddots & \cdots & \ddots & \vdots \\ 0 & 0 & \cdots & \cdots & \mu_1 & 0 & -\mu_1 \end{bmatrix} \quad (6.6)$$

where the element in the i^{th} row, j^{th} column in matrix $\tilde{\mathbf{S}}$ denotes the transition rate from state i ($i = 1$ corresponds to the absorbing state) to state j ($j = 1$ corresponds to the absorbing state). Thus, the transition rate matrix describing the transition rates among all the transient states \mathbf{S} can be obtained by removing the first row and the first column of $\tilde{\mathbf{S}}$.

Let α denote the initial probabilities of starting in any state (i.e., phase), which can be obtained from the stationary probabilities $\pi(k)$ as mentioned above. Let t denote the time required to reach the absorbing state (i.e., the content delay), then t follows a *phase-type distribution* denoted by $PH(\alpha, \mathbf{S})$. The cumulative distribution function (CDF) of t ($t > 0$) is given by [106]

$$F(t) = 1 - \alpha \exp(\mathbf{S}t)\mathbf{1}, \quad (6.7)$$

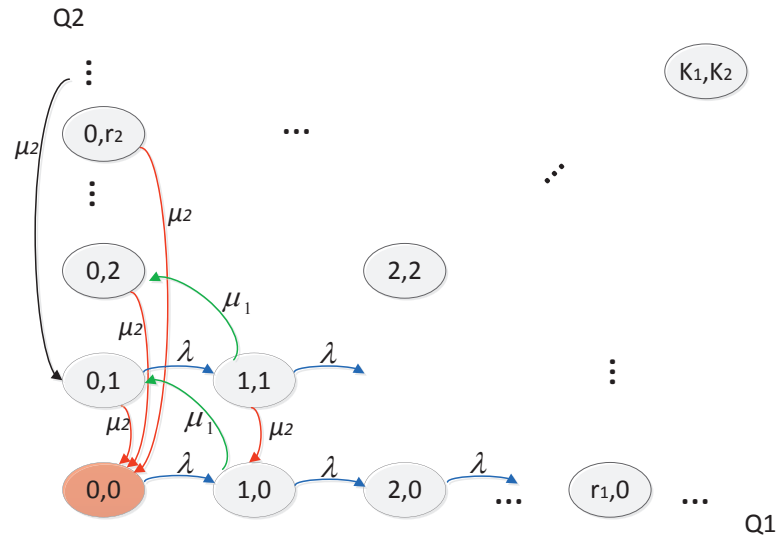
from which the DRWD can be obtained.

- Case II:

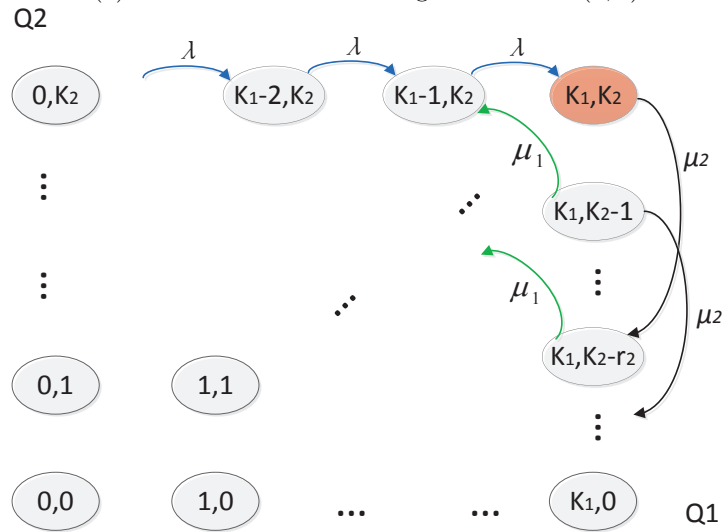
As for the content delivery service from the content provider to a Tier-1 subscriber

(i.e., S_1) directly connected to some Tier-1 relay (i.e., R_1), the state transition of the corresponding Markovian queue with bulk services is shown in Fig. 6.5 and Fig. 6.6. Let K_1 and K_2 denote the buffer size of Q_1 and Q_2 , respectively. (k_1, k_2) is the state variable, where k_1 ($0 \leq k_1 \leq K_1$) and k_2 ($0 \leq k_2 \leq K_2$) are the number of contents in Q_1 and Q_2 , respectively. For ease of presentation, the author only exhibits the state transition at state $(0, 0)$ and (K_1, K_2) , as well as the state transition at typical state on the edge of the state space and that in the center of the state space. In Fig. 6.5 and Fig. 6.6, r_1 and r_2 are the bulk size of Q_1 and Q_2 , respectively, indicating that r_1 units of bandwidth are allocated by the content provider, while r_2 units of bandwidth are allocated by the Tier-1 relay R_1 .

Let $\pi(k_1, k_2)$ denote the stationary probability that the queue contains k_1 and k_2 contents in the buffer space of Q_1 and Q_2 , respectively. From the state-transition-rate diagram shown in Fig. 6.5 and Fig. 6.6, the balancing equations can be obtained accordingly. Due to the page limit, the author only shows some

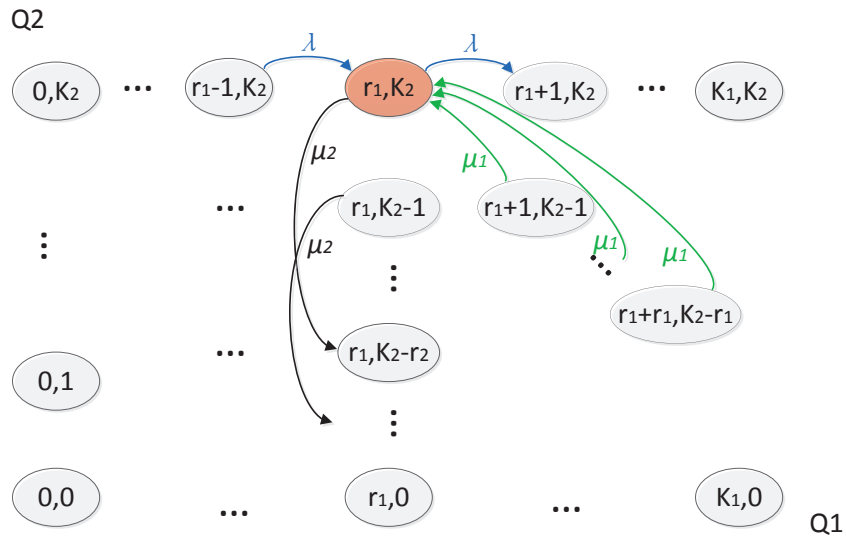


(a) State-transition-rate diagram at state $(0, 0)$

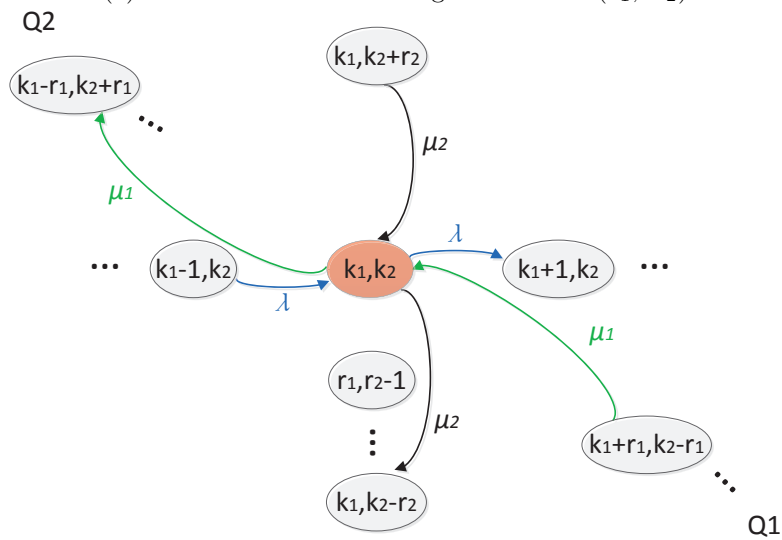


(a) State-transition-rate diagram at state (K_1, K_2)

Figure 6.5: State-transition-rate diagram of 2D Markovian queue with bulk services I.



(a) State-transition-rate Diagram at state (r_1, K_2)



(a) State-transition-rate diagram at state (r_1, r_2)

Figure 6.6: State-transition-rate Diagram of 2D Markovian queue with bulk services II.

sample equations for typical states in the following:

$$\pi(0, 0) \cdot \lambda = \sum_{k=1}^{r_2} \pi(0, k) \cdot \mu_2, \quad (6.8)$$

$$\begin{aligned} \pi(0, k_2) \cdot (\lambda + \mu_2) &= \sum_{k=1}^{k_2} \pi(k, k_2 - k) \cdot \mu_1 + \pi(0, k_2 + r_2) \cdot \mu_2, \\ &(0 < k_2 \leq r_1) \end{aligned} \quad (6.9)$$

$$\begin{aligned} \pi(0, k_2) \cdot (\lambda + \mu_2) &= \sum_{k=1}^{r_1} \pi(k, k_2 - k) \cdot \mu_1 + \pi(0, k_2 + r_2) \cdot \mu_2, \\ &(r_1 < k_2 \leq K_2 - r_2) \end{aligned} \quad (6.10)$$

$$\dots \quad \dots \quad \dots \quad \dots \quad \dots \quad (6.11)$$

$$\begin{aligned} \pi(0, k_2) \cdot (\lambda + \mu_2) &= \sum_{k=1}^{r_1} \pi(k, k_2 - k) \cdot \mu_1, \\ &(K_2 - r_2 < k_2 \leq K_2) \end{aligned} \quad (6.12)$$

$$\begin{aligned} \pi(k_1, k_2) \cdot (\lambda + \mu_1 + \mu_2) &= \pi(k_1 - 1, k_2) \cdot \lambda + \pi(k_1, k_2 + r_2) \cdot \mu_2, \\ &(0 < k_1 < K_1, 0 < k_2 < r_1) \end{aligned} \quad (6.13)$$

$$\begin{aligned} \pi(k_1, k_2) \cdot (\lambda + \mu_1 + \mu_2) &= \pi(k_1 - 1, k_2) \cdot \lambda + \pi(k_1 + r_1, k_2 - r_1) \cdot \mu_1 \\ &+ \pi(k_1, k_2 + r_2) \cdot \mu_2, \\ &(0 < k_1 \leq K_1 - r_1, r_1 \leq k_2 \leq K_2 - r_2) \end{aligned} \quad (6.14)$$

$$\dots \quad \dots \quad \dots \quad \dots \quad \dots \quad (6.15)$$

$$\begin{aligned} \pi(k_1, K_2) \cdot (\lambda + \mu_2) &= \pi(k_1 - 1, K_2) \cdot \lambda + \sum_{k=1}^{r_1} \pi(k_1 + k, K_2 - k) \cdot \mu_1, \\ &(0 < k_1 \leq K_1 - r_1) \end{aligned} \quad (6.16)$$

$$\begin{aligned} \pi(k_1, K_2) \cdot (\lambda + \mu_2) &= \pi(k_1 - 1, K_2) \cdot \lambda + \sum_{k=1}^{K_1 - k_1} \pi(k_1 + k, K_2 - k) \cdot \mu_1, \\ &(K_1 - r_1 < k_1 < K_1) \end{aligned} \quad (6.17)$$

$$\pi(K_1, 0) \cdot \mu_1 = \pi(K_1 - 1, 0) \cdot \lambda + \sum_{k=1}^{r_2} \pi(K_1, k) \cdot \mu_2 \quad (6.18)$$

$$\dots \quad \dots \quad \dots \quad \dots \quad \dots \quad (6.19)$$

$$\pi(K_1, K_2) \cdot \mu_2 = \pi(K_1 - 1, k_2) \cdot \lambda. \quad (6.20)$$

By introducing the normalizing equation $\sum_{(k_1, k_2)=(0,0)}^{(K_1, K_2)} \pi(k_1, k_2) = 1$, the above balancing equations can be solved, and thus the vector of the initial probabilities can be obtained. Similar as Case I, an absorbing Markov chain can also be constructed for Case II with absorbing state $(0, 0)$ and the other states being the transient states. Again, the CDF of the content delay can be obtained from Eq.(6.7).

Note that the derived DRWD in this section will be used in the contract design in the following section, as a metric to quantify the quality of the content delivery service.

6.2 Content Provider/ Subscriber Model and Optimal Contract Design

6.2.1 Content Provider Side

Considering the content delivery service as a kind of “commodity”, then the relay-based publish-subscribe network can be modeled as a trading market with the content provider being the seller and the subscribers being the consumers (i.e., buyers). The content provider sets the service qualities it can provide with, and the corresponding prices for that service. Intuitively, higher price has to be paid for higher service quality offered. Notified by the quality and price information, the subscribers can optimally decide which qualities to purchase.

As in Section 6.1.2, both Case I and Case II are considered in the proposed model. In Case I, note that different subscribers can have different contact rates with the content provider. Let \mathbb{S}_{CP} denote the set of subscribers connected to the content provider directly, which can be divided into different groups according to different contact rates

with the content provider. Let $\mathbb{G}_{CP} = \{\mathbb{G}_{CP}^i\}$ ($i = 1, \dots, I$) be the set of subscriber groups in Case I, where I is the number of such groups. Note that $\mathbb{S}_{CP} = \bigcup_{i=1, \dots, I} \mathbb{G}_{CP}^i$. The subscribers in the same group \mathbb{G}_{CP}^i are considered to have the same contact rate (denoted by ρ_i) with the content provider.

As for Case II, $\mathbb{R} = \{R_1, \dots, R_K\}$ is used to denote the set of the Tier-1 relays, where K is the number of the Tier-1 relays. Let \mathbb{S}_k denote the set of subscribers connected to Tier-1 relay R_k ($k = 1, \dots, K$), and $\mathbb{S}_R = \bigcup \mathbb{S}_k$ denotes the set of all subscribers connected to the Tier-1 relays. Similarly, the Tier-1 subscribers \mathbb{S}_R can be classified into different groups by contact rate information. Let $\mathbb{G}_R = \{\mathbb{G}_R^j\}$ ($j = 1, \dots, J$) be the set of subscriber groups in Case II, where J is the number of such groups. In a certain group \mathbb{G}_R^j , the subscribers are considered to have the same contact rates (denoted by ρ_j) with the Tier-1 relays they connect to. Also, the Tier-1 relays they connect to have the same contact rates (denoted by ν_j) with the content provider.

Considering one single subscriber group \mathbb{G} ($\mathbb{G} = \mathbb{G}_{CP}^i$ or \mathbb{G}_R^j), let Γ and Π denote the set of all service qualities and the set of all prices, respectively. Within group \mathbb{G} , each quality of service $\eta \in \Gamma$ corresponds to a price $p \in \Pi$. As the content delivery along different paths is independent, the content provider can solve the utility maximization problem separately, for each content delivery path. The utility of the content provider (obtained by selling a service quality η to a subscriber in group \mathbb{G} ($\mathbb{G} = \mathbb{G}_{CP}^i$ or \mathbb{G}_R^j)) is defined as the difference between the monetary gain collected from the subscriber in \mathbb{G} and the cost of providing service quality η (e.g., due to bandwidth allocation), which is denoted by:

$$U_{CP}^{\mathbb{G}}(\eta) = p^{\mathbb{G}}(\eta) - C^{\mathbb{G}}(\eta) \quad (6.21)$$

Note that from the queueing analysis in Section 6.1.2, there is a one-to-one mapping

between the bandwidth allocation W (which determines the bulk size in the queueing model) and the service quality η . Thus, η can be considered as a function of W , and $C^{\mathbb{G}}(\eta(W))$ can be written as $C^{\mathbb{G}}(W)$ for simplicity. Similar as [88], in this work, the cost function is defined as:

$$C^{\mathbb{G}}(W) = \delta \cdot W^{\beta} + C_0(p_c^b, p_c^r) \quad (6.22)$$

where δ is a weighting factor, $0 < \beta < 1$ is a parameter controlling the behavior of the cost function. p_c^b and p_c^r denote the power consumption of the base station and relay nodes due to content transmission, respectively, and $C_0(p_c^b, p_c^r)$ is the corresponding power consumption cost. $C_0(p_c^b, p_c^r)$ can be defined as $c_1 \cdot p_c^b + c_2 \cdot p_c^r$ for simplicity, where c_1 and c_2 are constants denoting the cost of consuming one unit of power for the base station and relay nodes, respectively. Thus, the cost function for Case I can be obtained from:

$$C^{\mathbb{G}} = \delta_1 \cdot (W_{CP}^{\mathbb{G}})^{\beta_1} + c_1 \cdot p_c^b, \quad \mathbb{G} = \mathbb{G}_{CP}^i, \quad (6.23)$$

where $W_{CP}^{\mathbb{G}}$ is the bandwidth required to provide service quality η to the subscriber in group \mathbb{G}_{CP}^i by the content provider. The cost function for Case II can be obtained from:

$$C^{\mathbb{G}} = \delta_1 \cdot (W_{CP}^{\mathbb{G}})^{\beta_1} + \delta_2 \cdot (W_R^{\mathbb{G}})^{\beta_2} + c_1 \cdot p_c^b + c_2 \cdot p_c^r, \quad \mathbb{G} = \mathbb{G}_R^j, \quad (6.24)$$

where $W_{CP}^{\mathbb{G}}$ and $W_R^{\mathbb{G}}$ are the bandwidth required to be allocated at the content provider and the Tier-1 relay, respectively, in order to provide service quality η to the subscriber in group \mathbb{G}_R^j . The aim of the content provider is to maximize its utility by strategically designing the service quality it can provide, as well as the corresponding cost. As a rational content provider, negative utility from providing a service can never be accepted, and thus it will always set $p^{\mathbb{G}}(\eta) \geq C^{\mathbb{G}}(\eta)$, $\forall \eta \in \Gamma$ for the subscriber in group

\mathbb{G} .

6.2.2 Subscriber Side

In practice, different subscribers may have different requirements for the content delivery quality, thus the valuation of the same service quality η can be different for different subscribers. The valuation of service quality η to subscriber S_i is defined as:

$$V_i(\eta) = \kappa \cdot \log(1 + \vartheta_i \cdot \eta) \quad (6.25)$$

where $\kappa > 0$ is a weighting factor, while $\vartheta_i > 0$ can be described as the time-sensitivity of subscriber S_i . A higher ϑ_i refers to a higher requirement for the service quality, i.e., less tolerance on content delivery delay. According to different preference for the quality of the content delivery service, the subscribers in group \mathbb{G} are classified into different types. A subscriber is referred to as a type- θ subscriber if $\vartheta_i = \theta$.

Considering subscriber group \mathbb{G}_{CP}^i , the author denotes the set of all subscriber types as Θ_{CP}^i . While considering subscriber group \mathbb{G}_R^j , the set of all subscriber types is denoted as Θ_R^j . Note that the subscriber type is private information (capturing the heterogeneity in preference for the quality of the content delivery service for each subscriber) since the content provider cannot know the exact time-sensitivity of a particular subscriber. However, it is assumed that the content provider can obtain some statistical information about the subscriber types, such as the probability distribution of the subscriber type. According to Eq. (6.25), for a service quality η , the valuation of a subscriber with type θ is calculated by

$$V_S(\theta, \eta) = \kappa \cdot \log(1 + \theta \cdot \eta), \quad \theta \in \Theta_{CP}^i, \text{ or } \Theta_R^j. \quad (6.26)$$

The author defines the utility of a subscriber with type θ in group \mathbb{G} by purchasing a service quality η as follows:

$$U_S^{\mathbb{G}}(\theta, \eta) = V_S(\theta, \eta) - p^{\mathbb{G}}(\eta), \quad \mathbb{G} = \mathbb{G}_{CP}^i, \text{ or } \mathbb{G}_R^j. \quad (6.27)$$

Assuming that each subscriber is selfish and rational, who always tries to achieve the maximum utility, then the optimal strategy for a subscriber with type θ in group \mathbb{G} can be obtained from:

$$\mathcal{A}^{\mathbb{G}}(\theta) = \arg \max_{\eta \in \Gamma} U_S^{\mathbb{G}}(\theta, \eta), \quad \mathbb{G} = \mathbb{G}_{CP}^i, \text{ or } \mathbb{G}_R^j. \quad (6.28)$$

6.2.3 Optimal Contract Design

With the relay-based publish-subscribe network being modeled as a trading market and the content delivery service quality modeled as commodity, both the utility maximization problem for the content provider and that for the subscribers can be considered from an economic perspective. According to Eqs. (6.21) and (6.27), it can be observed that the content provider wants to sell the service at a higher price to maximize its utility, while the subscriber prefers to purchase the required service quality with lower price, resulting in conflict objectives between them. To resolve the problem due to conflicting objectives in presence of incomplete information (the subscriber type can be hidden to the content provider), a contract-theoretic approach is applied in the following. With optimally designed service quality-price contract, which is offered by the content provider, the revenue of the content provider can be maximized, and meanwhile, the satisfactions of the subscribers can be achieved.

Specifically, the problem for the content provider is to choose a set of service qualities in terms of DRWD and a price for each service quality for subscribers in each

group. Considering a specific group \mathbb{G} ($\mathbb{G} = \mathbb{G}_{CP}^i$ or \mathbb{G}_R^j), since only one service quality will be chosen by each type of subscribers, the content provider will assign a service quality $\eta^{\mathbb{G}}(\theta)$ and a corresponding price $p^{\mathbb{G}}(\eta^{\mathbb{G}}(\theta))$ to each subscriber type $\theta \in \Theta_{\mathbb{G}}$. For simplicity, $p^{\mathbb{G}}(\eta^{\mathbb{G}}(\theta))$ is written as $p^{\mathbb{G}}(\theta)$, since $\eta^{\mathbb{G}}(\theta)$ is a single value function. Therefore, it is enough to design a contract consisting of M contract items, where M is the number of the subscriber types. The contract designed for the considered problem is defined as follows:

Definition 6.1. *CP-G Contract:* The CP-G contract \mathcal{C}_{CP-G} proposed by the content provider to group \mathbb{G} ($\mathbb{G} = \mathbb{G}_{CP}^i$ or \mathbb{G}_R^j) of subscribers has the form of $\mathcal{C}_{CP-G} = (\eta^{\mathbb{G}}(\theta), p^{\mathbb{G}}(\theta))$, $\forall \theta \in \Theta_{\mathbb{G}}$, where $\eta^{\mathbb{G}}(\theta)$ is the content delivery service quality (in terms of DRWD) provided by the content provider to the subscribers with type θ in group \mathbb{G} , and $p^{\mathbb{G}}(\theta)$ is the price required to be paid to the content provider by the subscribers with type θ in group \mathbb{G} , in order to get service quality $\eta^{\mathbb{G}}(\theta)$.

Definition 6.2. *Feasible CP-G Contract:* A feasible CP-G contract is a set of service quality-price combinations, where for each subscriber type $\theta \in \Theta_{\mathbb{G}}$, purchasing the service with quality $\eta^{\mathbb{G}}(\theta)$ at price $p^{\mathbb{G}}(\theta)$ is always the best choice for a subscriber with type θ . In other words, subscriber with type θ can achieve the highest utility by choosing the contract item $(\eta^{\mathbb{G}}(\theta), p^{\mathbb{G}}(\theta))$, compared with that can be obtained if purchasing other items or nor purchasing at all.

Specifically, for a feasible CP-G contract, two properties have to be satisfied, namely the incentive compatibility (IC) and the individual rationality (IR). In the considered problem, IC property indicates that each subscriber prefers to buy the content delivery service aligned to his type, i.e.,

$$V_S(\theta, \eta^{\mathbb{G}}(\theta)) - p^{\mathbb{G}}(\theta) \geq V_S(\theta, \eta^{\mathbb{G}}(\tilde{\theta})) - p^{\mathbb{G}}(\tilde{\theta}), \quad \forall \tilde{\theta} \neq \theta, \quad (6.29)$$

for all type $\theta \in \Theta_{\mathbb{G}}$. IR means that a subscriber will choose the contract item designed for his type only when this choice can provide him with a non-negative utility, i.e.,

$$V_S(\theta, \eta^{\mathbb{G}}(\theta)) - p^{\mathbb{G}}(\theta) \geq 0, \quad \forall \theta \in \Theta_{\mathbb{G}}. \quad (6.30)$$

Note that for a feasible contract $\mathcal{C}_{CP-G}^{\mathbb{G}}$, the overall utility of the content provider obtained from selling content delivery services to subscribers in group \mathbb{G} is calculated by:

$$U_{CP}^{\mathbb{G}} = \sum_{\theta \in \Theta_{\mathbb{G}}} N_{\theta}^{\mathbb{G}} \left[p^{\mathbb{G}}(\theta) - C^{\mathbb{G}}(\eta^{\mathbb{G}}(\theta)) \right], \quad \forall \theta \in \Theta_{\mathbb{G}}, \quad (6.31)$$

where $N_{\theta}^{\mathbb{G}}$ is the number of type- θ subscribers in group \mathbb{G} . The optimal contract for the content provider, denoted by $\mathcal{C}_{CP-G}^{\mathbb{G},*}$, is a feasible contract maximizing its overall utility, i.e.,

$$\mathcal{C}_{CP-G}^{\mathbb{G},*} = \{(\eta^{\mathbb{G},*}(\theta), p^{\mathbb{G},*}(\theta))\} = \arg \max_{\{\eta^{\mathbb{G}}(\theta)\}, \{p^{\mathbb{G}}(\theta)\}} \sum_{\theta \in \Theta_{\mathbb{G}}} N_{\theta}^{\mathbb{G}} \left[p^{\mathbb{G}}(\theta) - C^{\mathbb{G}}(\eta^{\mathbb{G}}(\theta)) \right], \quad \forall \theta \in \Theta_{\mathbb{G}}, \quad (6.32)$$

subject to the corresponding IC and IR constraints, as mentioned in Eqs. (6.29) and (6.30).

In the proposed model, the author considers a finite number of subscriber types, denoted by $\theta_1, \theta_2, \dots, \theta_M$. Without loss of generality, it is assumed that $\theta_1 < \theta_2 \dots < \theta_M$. Then, Eq. (6.33) can be written as:

$$U_{CP}^{\mathbb{G}} = \sum_{m=1}^M N_{\theta_m}^{\mathbb{G}} \left[p^{\mathbb{G}}(\theta_m) - C^{\mathbb{G}}(\eta^{\mathbb{G}}(\theta_m)) \right]. \quad (6.33)$$

Note that with the purpose of maximizing $U_{CP}^{\mathbb{G}}$, both variables $\eta^{\mathbb{G}}(\theta_m)$ and $p^{\mathbb{G}}(\theta_m)$ have to be decided simultaneously, for $m = 1, \dots, M$. The author defines a quality

assignment $\eta^{\mathbb{G}} = \{\eta^{\mathbb{G}}(\theta_1), \eta^{\mathbb{G}}(\theta_2), \dots, \eta^{\mathbb{G}}(\theta_M)\}$ as a feasible quality assignment, if the constraints $\eta^{\mathbb{G}}(\theta_1) \leq \eta^{\mathbb{G}}(\theta_2) \leq \dots \leq \eta^{\mathbb{G}}(\theta_M)$. To design the optimal contract for the considered problem, similar method in [88] (i.e., sequential optimization technique) can be applied. Specifically, in the first step, the author derives the best prices under a fixed feasible service quality assignment, while for the second step, the best service quality assignment is derived.

Let $U_{CP}^{\mathbb{G},*}(\{\eta^{\mathbb{G}}\})$ denote the maximum utility the content provider can achieve from selling services to subscriber group \mathbb{G} under a given feasible quality assignment $\{\eta^{\mathbb{G}}\}$, which can be obtained from:

$$U_{CP}^{\mathbb{G},*}(\eta^{\mathbb{G}}) = \max_{p^{\mathbb{G}}(\theta_m)} \sum_{m=1}^M N_{\theta_m}^{\mathbb{G}} \left[p^{\mathbb{G}}(\theta_m) - C^{\mathbb{G}}(\eta^{\mathbb{G}}(\theta_m)) \right] \quad (6.34)$$

subject to the price constraints in the following:

- $0 \leq p^{\mathbb{G}}(\theta_1) \leq V_S(\theta_1, \eta^{\mathbb{G}}(\theta_1))$
- $p^{\mathbb{G}}(\theta_m) \geq p^{\mathbb{G}}(\theta_{m-1}) + V_S(\theta_{m-1}, \eta^{\mathbb{G}}(\theta_m)) - V_S(\theta_{m-1}, \eta^{\mathbb{G}}(\theta_{m-1})), \quad 2 \leq m \leq M.$
- $p^{\mathbb{G}}(\theta_m) \leq p^{\mathbb{G}}(\theta_{m-1}) + V_S(\theta_m, \eta^{\mathbb{G}}(\theta_m)) - V_S(\theta_m, \eta^{\mathbb{G}}(\theta_{m-1})), \quad 2 \leq m \leq M.$

Note that the second price constraint can be written as:

$$V_S(\theta_{m-1}, \eta^{\mathbb{G}}(\theta_m)) - p^{\mathbb{G}}(\theta_m) \leq V_S(\theta_{m-1}, \eta^{\mathbb{G}}(\theta_{m-1})) - p^{\mathbb{G}}(\theta_{m-1}) \quad (6.35)$$

which can be further written as:

$$V_S(\theta_m, \eta^{\mathbb{G}}(\theta_{m+1})) - p^{\mathbb{G}}(\theta_{m+1}) \leq V_S(\theta_m, \eta^{\mathbb{G}}(\theta_m)) - p^{\mathbb{G}}(\theta_m) \quad (6.36)$$

which implies that type θ_m will not choose quality $\eta^{\mathbb{G}}(\theta_{m+1})$ designed for type θ_{m+1} ,

since the utility of type θ_m on quality $\eta^{\mathbb{G}}(\theta_{m+1})$ is smaller than that on quality $\eta^{\mathbb{G}}(\theta_m)$.

Similarly, the third price constraint can be written as:

$$V_S(\theta_m, \eta^{\mathbb{G}}(\theta_{m-1})) - p^{\mathbb{G}}(\theta_{m-1}) \leq V_S(\theta_m, \eta^{\mathbb{G}}(\theta_m)) - p^{\mathbb{G}}(\theta_m) \quad (6.37)$$

which can be illustrated as that type θ_m will not choose quality $\eta^{\mathbb{G}}(\theta_{m-1})$ designed for type θ_{m-1} , since the utility of type θ_m on quality $\eta^{\mathbb{G}}(\theta_{m-1})$ is smaller than that on quality $\eta^{\mathbb{G}}(\theta_m)$. The price constraints are sufficient and necessary conditions for the contract to be feasible, which can be proved by mathematical induction as in [88]. For the problem defined in Eq. (6.34), the optimal price assignment with given service quality assignment, which is denoted by $\mathbf{p}^{\mathbb{G},*} = \{p^{\mathbb{G},*}(\theta_1), p^{\mathbb{G},*}(\theta_2), \dots, p^{\mathbb{G},*}(\theta_M)\}$, can be obtained from:

$$p^{\mathbb{G},*}(\theta_1) = V_S(\theta_1, \eta^{\mathbb{G}}(\theta_1)) \quad (6.38)$$

$$p^{\mathbb{G},*}(\theta_m) = p^{\mathbb{G},*}(\theta_{m-1}) + V_S(\theta_m, \eta^{\mathbb{G}}(\theta_m)) - V_S(\theta_m, \eta^{\mathbb{G}}(\theta_{m-1})), 2 \leq m \leq M \quad (6.39)$$

Note that the optimality of the price assignment and the uniqueness can be proved similarly as in [88]. For simplicity of denotations, let $\Delta_1 = 0$, and $\Delta_m = V_S(\theta_m, \eta^{\mathbb{G}}(\theta_m)) - V_S(\theta_m, \eta^{\mathbb{G}}(\theta_{m-1}))$, $\forall m = 2, \dots, M$. Note that Δ_m represents the gap between the valuation that can be achieved by a θ_m subscriber from choosing service quality designed for type- θ_m subscribers, and that can be achieved from choosing service quality designed for the subscribers with the nearest lower type, i.e., type θ_{m-1} . Then, the optimal prices denoted by Eq. (6.39) can be written as:

$$p^{\mathbb{G},*}(\theta_m) = V_S(\theta_1, \eta^{\mathbb{G}}(\theta_1)) + \sum_{i=1}^m \Delta_i, \quad m = 1, \dots, M. \quad (6.40)$$

By substituting the optimal price assignment $p^{\mathbb{G},*}(\theta_m)$ under fixed service quality assignment into Eq. (6.34), we have

$$U_{CP}^{\mathbb{G},*}(\eta^{\mathbb{G}}) = \sum_{m=1}^M N_{\theta_m}^{\mathbb{G}} \left[V_S(\theta_1, \eta^{\mathbb{G}}(\theta_1)) + \sum_{i=1}^m \Delta_i - C^{\mathbb{G}}(\eta^{\mathbb{G}}(\theta_m)) \right] \quad (6.41)$$

which can be further written as:

$$U_{CP}^{\mathbb{G},*}(\eta^{\mathbb{G}}) = \sum_{m=1}^M N_{\theta_m}^{\mathbb{G}} \left[V_S(\theta_m, \eta^{\mathbb{G}}(\theta_m)) + \sum_{i=1}^{m-1} (V_S(\theta_i, \eta^{\mathbb{G}}(\theta_i)) - V_S(\theta_{i+1}, \eta^{\mathbb{G}}(\theta_i))) - C^{\mathbb{G}}(\eta^{\mathbb{G}}(\theta_m)) \right] \quad (6.42)$$

Compared with Eq. (6.33) containing two variables, i.e., the service quality assignment and the price assignment, there is only one variable, i.e., the service quality assignment, in Eq. (6.42). The optimal service quality assignment can be denoted as

$$\eta^{\mathbb{G},*} = \arg \max_{\eta^{\mathbb{G}}} U_{CP}^{\mathbb{G},*}(\eta^{\mathbb{G}}), \quad (6.43)$$

which can be derived as follows. Considering that the subscribers enter group \mathbb{G} in sequence, from the lowest type θ_1 to the highest type θ_M . Thus the first element of the optimal service quality assignment, i.e., $\eta^{\mathbb{G},*}(\theta_1)$, can be obtained from $\eta^{\mathbb{G},*}(\theta_1) = \arg \max_{\eta^{\mathbb{G}}(\theta_1)} N_{\theta_1}^{\mathbb{G}} (V_S(\theta_1, \eta^{\mathbb{G}}(\theta_1)) - C^{\mathbb{G}}(\eta^{\mathbb{G}}(\theta_1)))$. Based on $\eta^{\mathbb{G},*}(\theta_1)$, when the subscribers of type θ_2 enter group \mathbb{G} , the second element of the optimal service quality assignment can be obtained from

$$\eta^{\mathbb{G},*}(\theta_2) = \arg \max_{\eta^{\mathbb{G}}(\theta_2)} \left[N_{\theta_1}^{\mathbb{G}} (V_S(\theta_1, \eta^{\mathbb{G}}(\theta_1)) - C^{\mathbb{G}}(\eta^{\mathbb{G}}(\theta_1))) + N_{\theta_2}^{\mathbb{G}} \Phi(\theta_2, \eta^{\mathbb{G}}(\theta_2)) \right] \quad (6.44)$$

where $\Phi(\theta_2, \eta^{\mathbb{G}}(\theta_2)) = V_S(\theta_2, \eta^{\mathbb{G}}(\theta_2)) + V_S(\theta_1, \eta^{\mathbb{G}}(\theta_1)) - V_S(\theta_2, \eta^{\mathbb{G}}(\theta_1)) - C^{\mathbb{G}}(\eta^{\mathbb{G}}(\theta_2))$. By

applying this scheme, the m^{th} element (i.e., $\eta^{\mathbb{G},*}(\theta_m)$) of the service quality assignment can be obtained based on all elements determined before, i.e, from the first element to the $(m - 1)^{\text{th}}$ element. Finally, the optimal service quality assignment $\eta^{\mathbb{G}}$ can be obtained. Combined with the corresponding price assignment given by Eq. (6.40), all contract items of the CP-G contract for the considered problem are optimally designed.

Note that the principle behind the sequential derivation of the optimal contract items comes from mathematical induction. First, it can be proved that the contract designed for subscriber group containing only one type (denoted by $\mathcal{C}_{CP-G}(\theta_1)$) is feasible. Second, it can be proved that so long as the contract designed for subscriber group containing m types (denoted by $\mathcal{C}_{CP-G}(\theta_1, \dots, \theta_m)$) is feasible, it will be still feasible when type- $(m + 1)$ subscribers are involved. To make sure that the first condition holds, it only needs to be shown that the contract $\mathcal{C}_{CP-G}(\theta_1)$ satisfies the IR constraint in Eq. (6.30). While to guarantee that contract $\mathcal{C}_{CP-G}(\theta_1, \dots, \theta_{m+1})$ is feasible based on the prerequisite that contract $\mathcal{C}_{CP-G}(\theta_1, \dots, \theta_m)$ is feasible, the following conditions have to be met:

- For the new coming subscribers with type θ_{m+1} , the IC and IR constraints (shown in Eq. (6.29) and Eq. (6.30)) are satisfied.
- For the existing type- θ_1 to type- θ_m subscribers, the IC constraints are still satisfied, in other words, they will not be affected by the presence of the subscribers with type θ_{m+1} .

Similar method as presented in [88] can be applied to prove. In Section 6.3, the author will show that with the optimal contract design, the author's objective of maximizing the revenue of the content provider while satisfying the requirements for service quality of the subscribers (with their heterogeneity being taken into account) can be achieved.

6.2.4 Architecture of the Proposed Algorithm for Contract Design

In this section, the author proposes the architecture of the proposed algorithm to implement the contract design, which is shown in the following:

6.3 Performance Evaluation

6.3.1 Performance Metrics Obtained from Queueing Model

In this section, the author presents the analytical results as well as simulation results of the performance metrics obtained from the proposed queueing model, as shown in Figs. 6.8-6.19. The CDF of content delay with different queue service rate (i.e., μ) is shown in Fig. 6.8. It can be observed that as μ increases, the CDF curve of content delay converges with a higher speed. In order to verify that validity of the analytical results, the author has done simulation of the queueing model using MATLAB. Some simulation results are shown as examples in Figs. 6.9-6.12, it can be found that the simulation results match the analytical results very well, especially when the number of content arrivals is large enough. Note that the queue service rate is dependent on the contact rate between the subscriber and the content provider, it can be observed from Fig. 6.13 that as μ increases, the DRWD keeps increasing, which accords with the expectation that when the subscriber has a frequent contact with the content provider, it can always get the latest content more easily, resulting in a higher DRWD. From Figs. 6.13-6.14, it can also be observed that the simulation results match the analytical results very well, they can be nearly overlapped when the number of content arrivals becomes large.

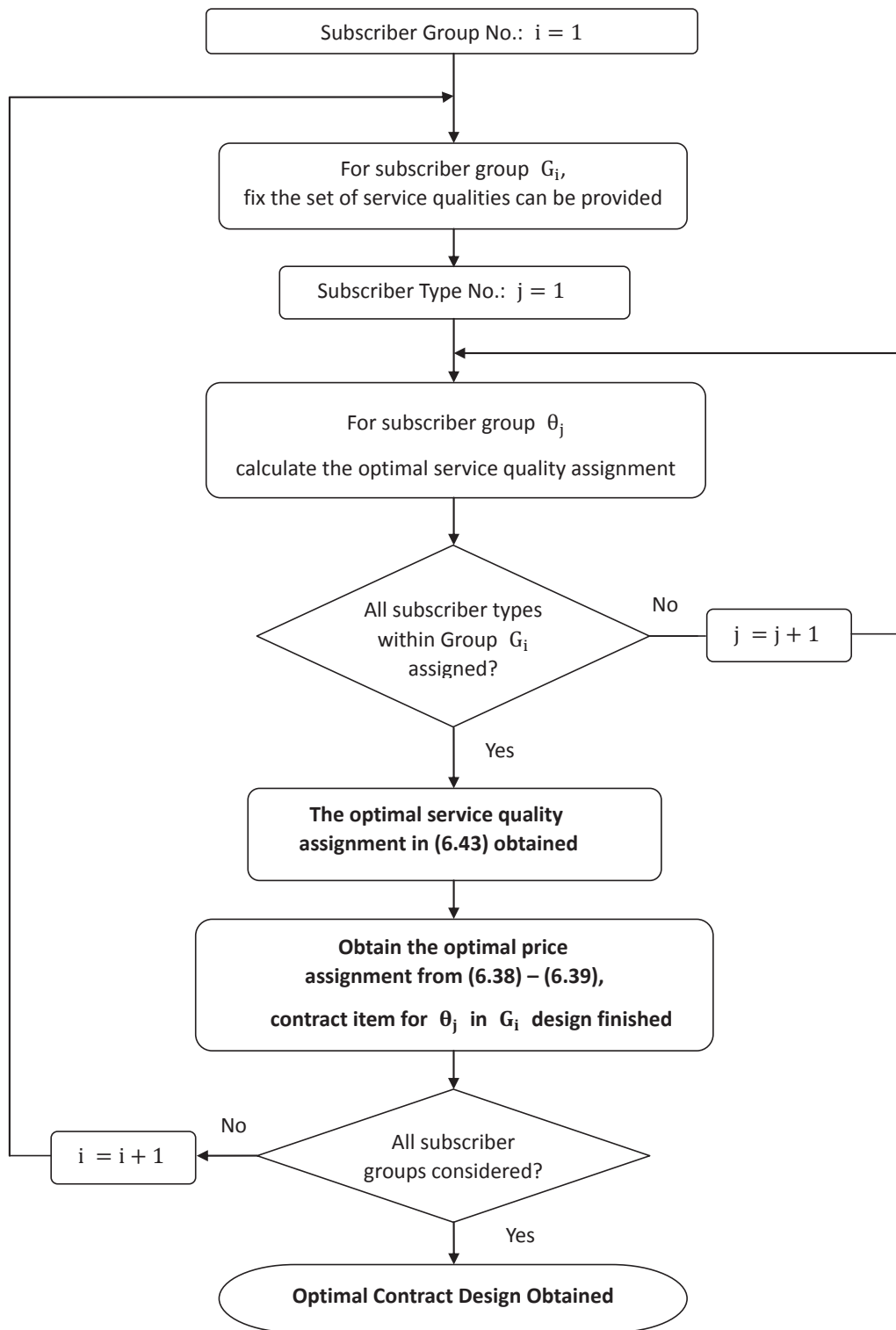


Figure 6.7: Flowchart of the proposed algorithm to implement the Contract Design.

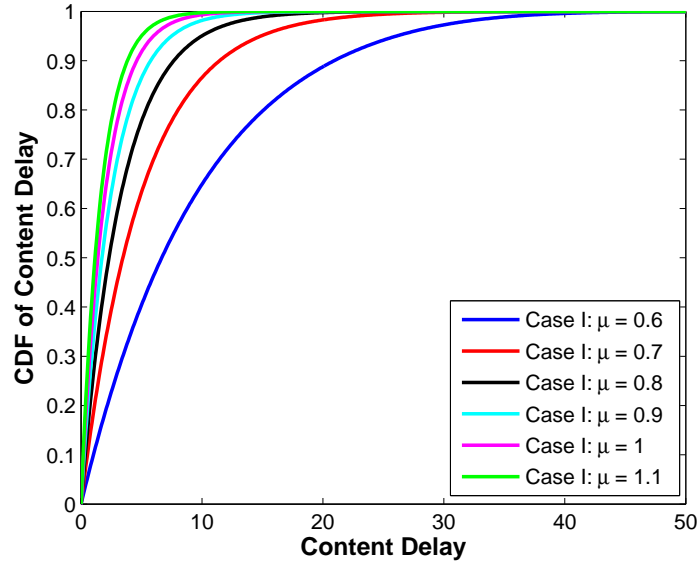


Figure 6.8: CDF of content delay distribution for different queue service rates μ in Case I.

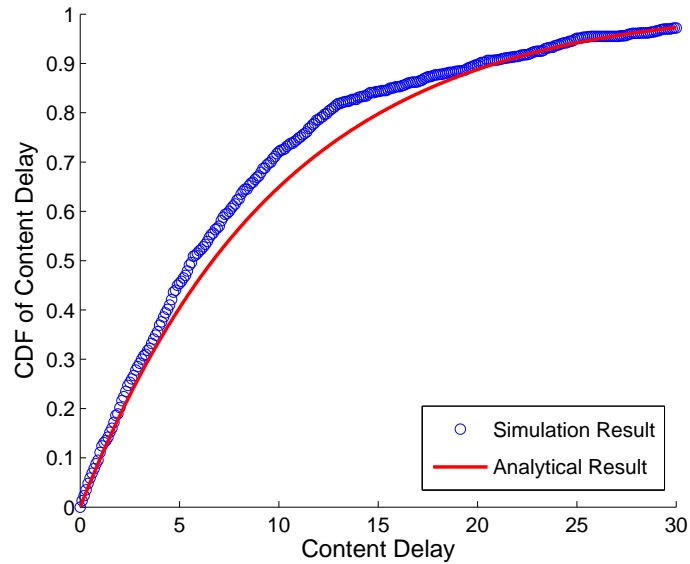


Figure 6.9: Comparison of simulation result and analytical result when $\mu = 0.6$ (1000 content arrivals).

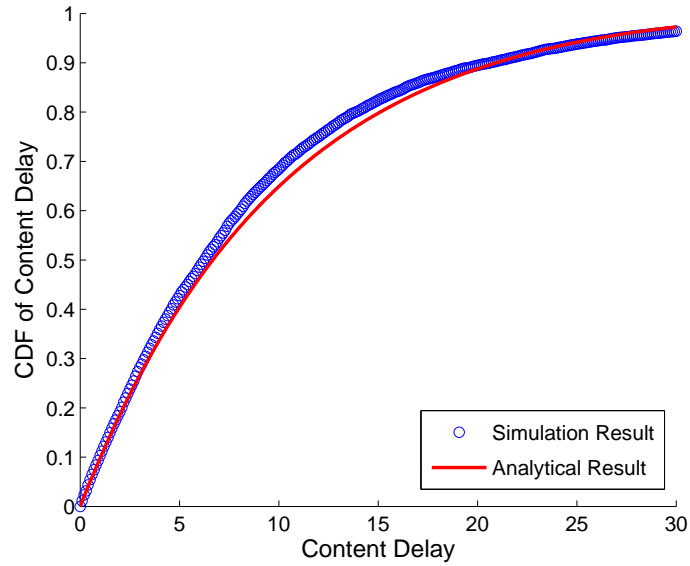


Figure 6.10: Comparison of simulation result and analytical result when $\mu = 0.6$ (10000 content arrivals).

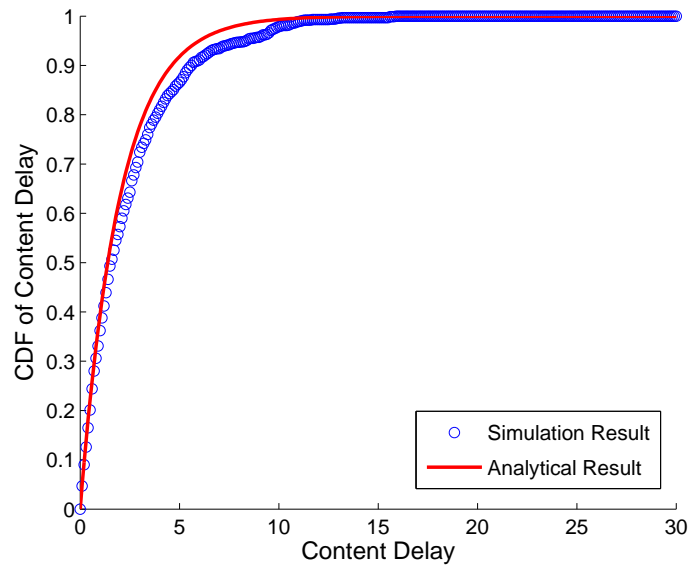


Figure 6.11: Comparison of simulation result and analytical result when $\mu = 1$ (1000 content arrivals).

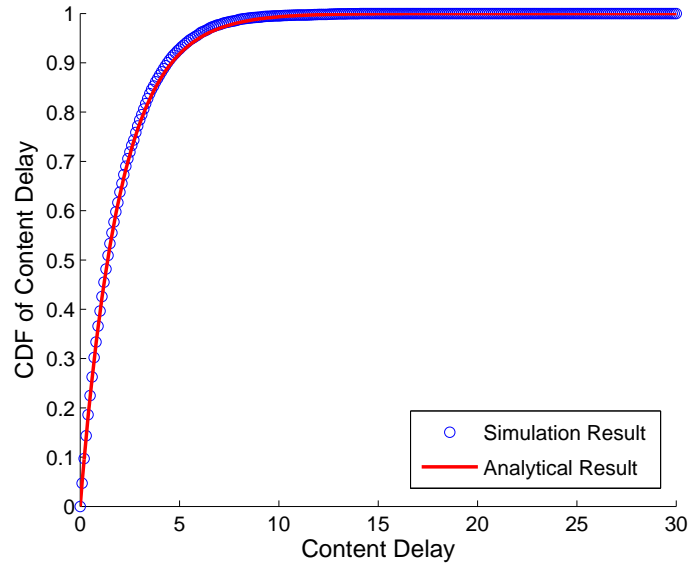


Figure 6.12: Comparison of simulation result and analytical result when $\mu = 1$ (10000 content arrivals).

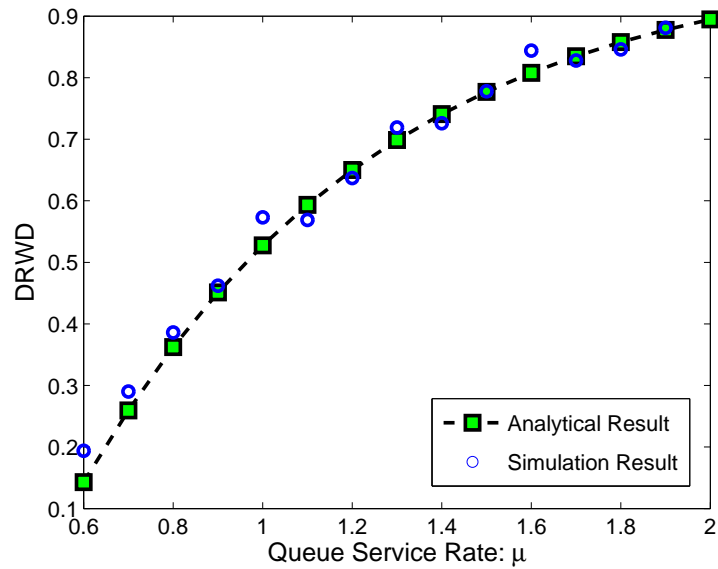


Figure 6.13: Comparison of simulation result and analytical result of DRWD (1000 content arrivals).

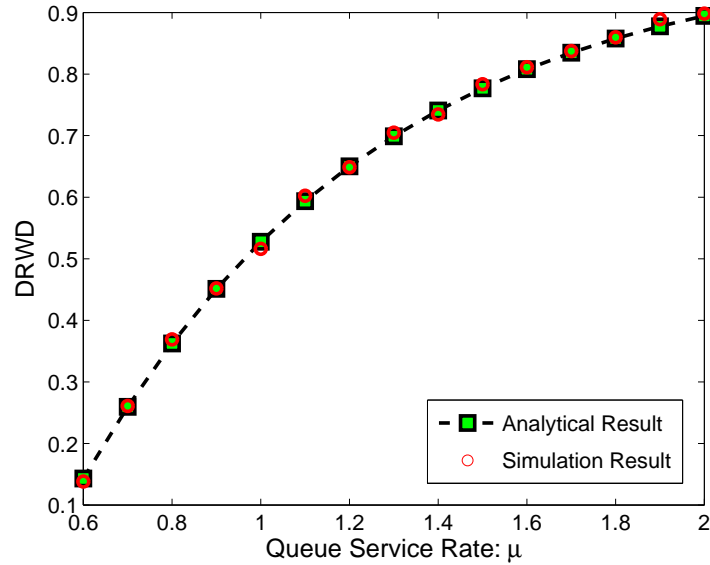


Figure 6.14: Comparison of simulation result and analytical result of DRWD (10000 content arrivals).

The CDF of content delay is shown for different bandwidth allocation at the content provider (i.e., r_1) is shown in Fig. 6.15, while sampled simulation results are shown in Figs. 6.16-6.17. The variation of the DRWD with r_1 is shown in Fig. 6.18. It can be observed that the DRWD increases as r_1 increases, however, the increasing speed becomes obviously slower when $r_1 \geq 3$, indicating that increasing bandwidth allocation will not always bring apparent improvement of the content delivery service. This observation can also be verified by the more and more indistinguishable difference in the CDF of content delay as r_1 increases, as shown in Fig. 6.15. Therefore, when the improvement of the service quality cannot compensate the cost has to be paid, the content provider has no motivation to increase the bandwidth allocation.

In Figs. 6.19(a) and (b), the variation of DRWD with queue service rate μ_1 and μ_2 is illustrated, respectively. From Fig. 6.19(a), it can be observed that with given bandwidth allocation at the Tier-1 relay (i.e., r_2), increasing the bandwidth allocation at the content provider (i.e., r_1) will result in higher content delivery service quality

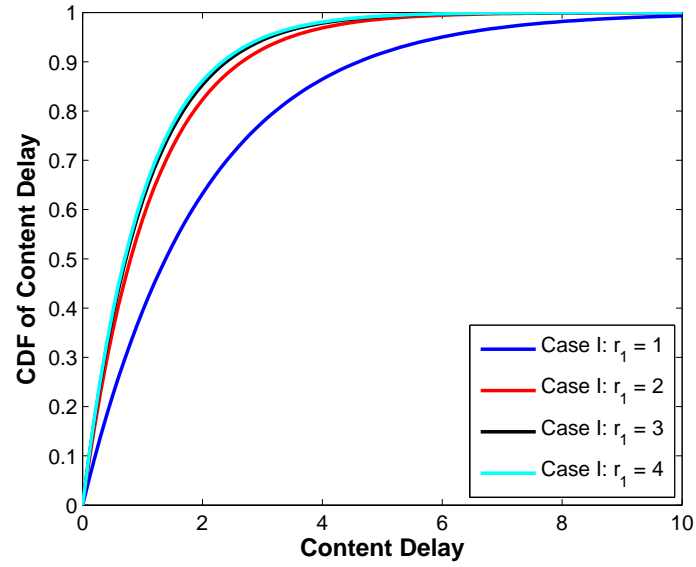


Figure 6.15: CDF of content delay distribution for different bandwidth allocation r_1 in Case I.

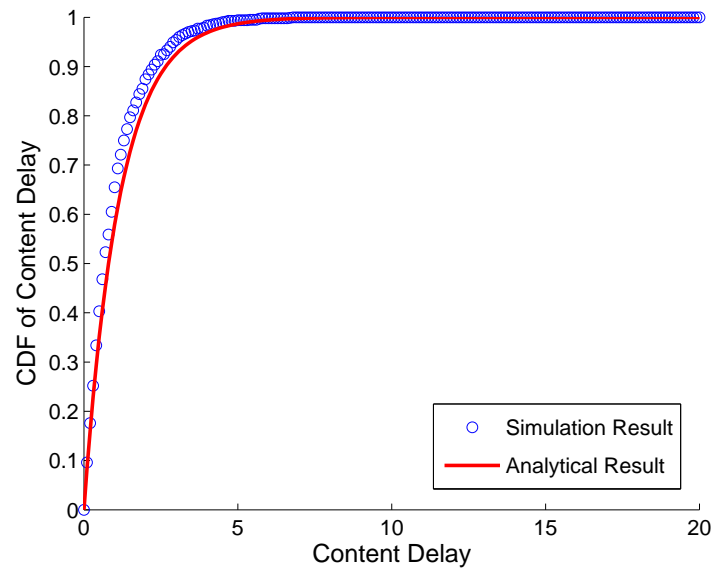


Figure 6.16: Comparison of simulation result and analytical result when $r_1 = 2$.

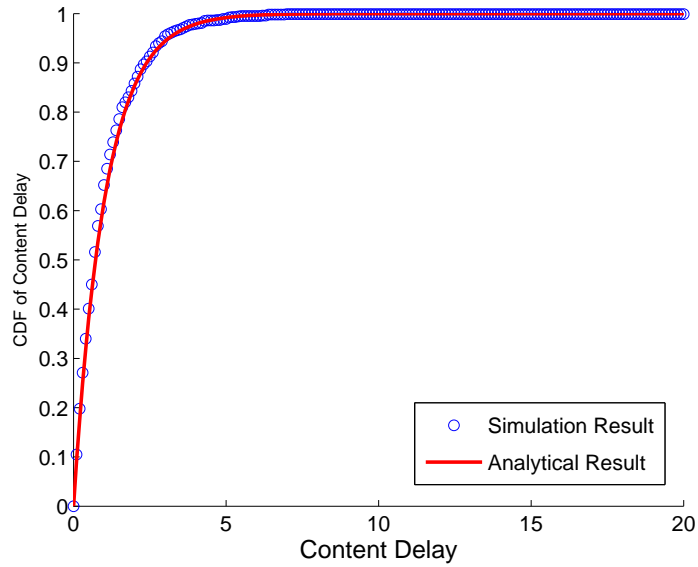


Figure 6.17: Comparison of simulation result and analytical result when $r_1 = 3$.

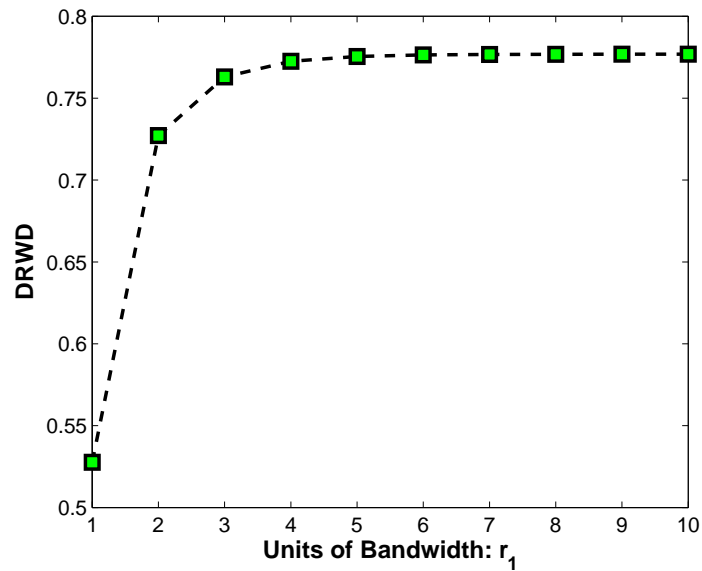


Figure 6.18: DRWD vs. bandwidth allocation r_1 in Case I.

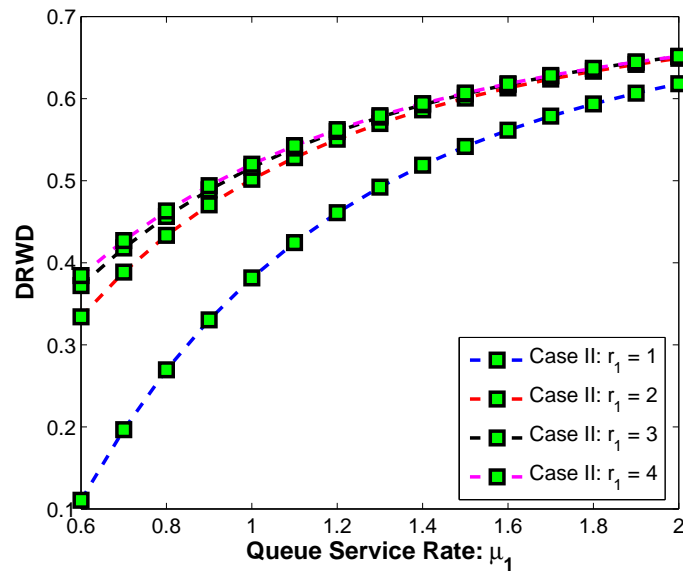
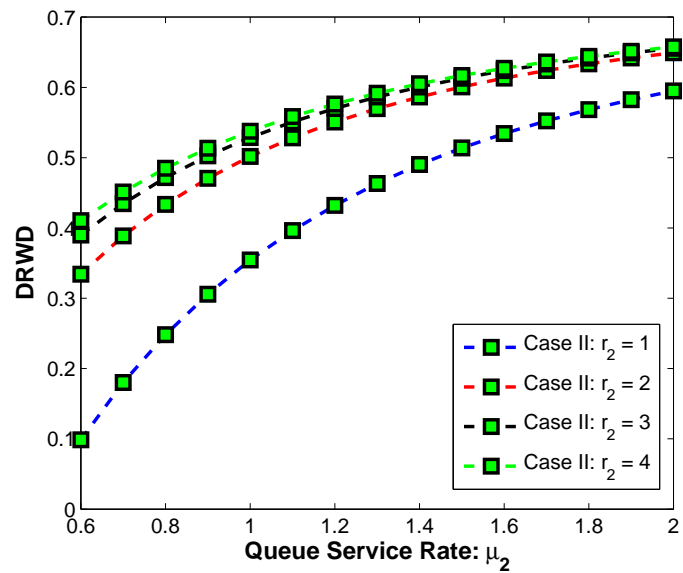
measured by DRWD. The improvement will be less obvious when r_1 increases, which accords with the author's observation from Fig. 6.18. Moreover, with a fixed r_1 , the DRWD will increase as μ_1 increases, indicating that the Tier-1 subscriber can achieve a higher service quality when the Tier-1 relay contacts with the content provider at a higher rate. Similar conclusion can be obtained from Fig. 6.19(b), where r_1 is fixed. The author observes that the Tier-1 subscriber can also achieve a higher service quality through more frequent contact with the Tier-1 relay connected with it.

6.3.2 Optimal Contract Design for the Content Provider

In this section, the author implements the proposed contract \mathcal{C}_{CP-G} for the content provider in both Case I and Case II, illustrating that the contract design is optimal from the perspective of both the content provider and the subscribers. The content arrival rate λ is 0.5 contents per hour, and the life cycle time of the content is set to be 1.5 hour in Case I, and 2.5 hour in Case II, respectively. The parameters of the cost function shown in Eqs. (6.23)-(6.24) are set to be: $\delta_1 = 0.02$, $\beta_1 = 0.02$, and $\delta_2 = 0.01$, $\beta_2 = 0.01$.

In Case I, the author considers four subscriber groups with contact rate with the content provider to be 0.6, 0.8, 1 and 1.2 contact per hour, respectively, which are denoted by group \mathbb{G}_{CP}^1 , group \mathbb{G}_{CP}^2 , group \mathbb{G}_{CP}^3 and group \mathbb{G}_{CP}^4 , as shown in Table 6.1, in each of which there are 20 subscribers. In each subscriber group, the author considers discrete types $\theta_1 = 1$ and $\theta_2 = 10$, indicating that some subscribers have lower requirements for service quality (type θ_1), while others have higher requirements for service quality (type θ_2). The number of θ_m ($m = 1, 2$) subscribers is 10. From Table 6.1, the author has two observations as shown in the following:

- Within one subscriber group, as type θ increases, the service quality designed for

(a) DRWD vs. queue service rate μ_1 in Case II(b) DRWD vs. queue service rate μ_2 in Case IIFigure 6.19: DRWD vs. queue service rates μ_1 and μ_2 in Case II.

type- θ subscribers will increase, as well as the corresponding price they have to pay.

- Comparing between different subscriber groups, the subscribers with higher contact rate with the content provider will receive higher service qualities than those in the other groups.

Combined with Table 6.1, the optimality of the design of the contract items in Case I can be verified by Fig. 6.20, where group \mathbb{G}_{CP}^1 and \mathbb{G}_{CP}^4 are taken as examples. As shown in Fig. 6.20, it can be observed that the content provider will always select the service quality maximizing its utility. Note that the IC constraints are implicitly satisfied in the contract design process, therefore, subscribers with type θ will benefit most by selecting the contract item designed for type θ . As a result, the author can verify that the proposed contract design is optimal, with the benefits of both the content provider and the subscribers taken into account.

Table 6.1: Contract Design of the Content Provider: Case I.

Contract Design of the Content Provider		
Subscriber Group	Contract Item $(\eta^{\mathbb{G}_{CP}}(\theta), p^{\mathbb{G}_{CP}}(\theta))$	
Group \mathbb{G}_{CP}^1	Type θ_1	(0.5919, 0.4650)
	Type θ_2	(0.5931, 0.4667)
Group \mathbb{G}_{CP}^2	Type θ_1	(0.6976, 0.5292)
	Type θ_2	(0.6984, 0.5301)
Group \mathbb{G}_{CP}^3	Type θ_1	(0.7755, 0.5741)
	Type θ_2	(0.7764, 0.5751)
Group \mathbb{G}_{CP}^4	Type θ_1	(0.8324, 0.6056)
	Type θ_2	(0.8340, 0.6074)

In Table 6.2, the optimal contract design for the content provider in Case II is presented. For instance, the author considers two Tier-1 relays having different contact rates with the content provider, which are set to be 0.8 and 1 contacts per hour, respectively. For each Tier-1 relay, there are 40 subscribers connected with it, where

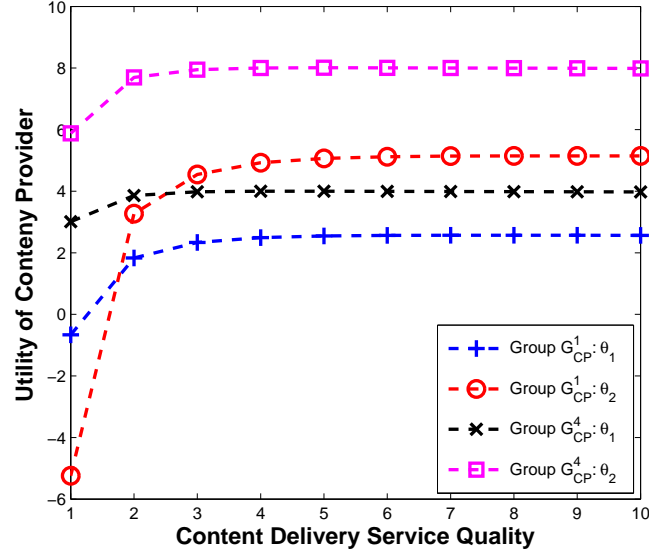


Figure 6.20: Utility of content provider vs. queue service quality selection in Case I.

20 of them contact with it at rate 0.6, while the other 20 contact with it at rate 1. Let (μ_1, μ_2) denote the full contact rate information between a subscriber and the content provider, where μ_1 is the contact rate between the Tier-1 relay and the content provider, and μ_2 is the contact rate between the subscriber and the Tier-1 relay. In this case, the entire group of Tier-1 subscribers can be further divided into four groups, namely $\mathbb{G}_R^1 (0.8, 0.6)$, $\mathbb{G}_R^2 (0.8, 1)$, $\mathbb{G}_R^3 (1, 0.6)$ and $\mathbb{G}_R^4 (1, 1)$, according to different contact rate information. Same as in Case I, the author considers discrete types $\theta_1 = 1$ and $\theta_2 = 10$ in each subscriber group. The number of θ_m ($m = 1, 2$) subscribers is 10. From Table 6.2, the author has two observations shown in the following:

- Comparing \mathbb{G}_R^1 and \mathbb{G}_R^2 (or \mathbb{G}_R^3 and \mathbb{G}_R^4), it can be observed that given μ_1 , as μ_2 increases, the subscribers will be guided to choose contract items with higher service qualities at higher price correspondingly.
- Comparing \mathbb{G}_R^1 and \mathbb{G}_R^3 (or \mathbb{G}_R^2 and \mathbb{G}_R^4), it can be observed that given μ_2 , the content provider will design contract items with higher service qualities for those

subscribers with higher μ_1 .

The optimality of the proposed contract design in Case II can be verified by Fig. 6.21, where group \mathbb{G}_R^3 is taken as an example. It can be observed that the content provider will strategically allocate bandwidth at itself as well as at the Tier-1 relays for the sake of maximizing its utility. Again, since the highest benefit of the subscribers are implicitly guaranteed during the contract design process, the author can demonstrate that the proposed contract design is optimal. Both the content provider and the subscribers can benefit from signing such contract, resulting in an effective content delivery procedure in relay-based publish-subscribe networks.

Table 6.2: Contract Design of the Content Provider: Case II .

Contract Design of the Content Provider		
Subscriber Group	Contract Item $(\eta^{\mathbb{G}_R}(\theta), p^{\mathbb{G}_R}(\theta))$	
Group \mathbb{G}_R^1	Type θ_1	(0.6224, 0.4839)
	Type θ_2	(0.6235, 0.4854)
Group \mathbb{G}_R^2	Type θ_1	(0.7431, 0.5557)
	Type θ_2	(0.7433, 0.5559)
Group \mathbb{G}_R^3	Type θ_1	(0.6637, 0.5091)
	Type θ_2	(0.6652, 0.5110)
Group \mathbb{G}_R^4	Type θ_1	(0.7937, 0.6637)
	Type θ_2	(0.7939, 0.6652)

6.4 Summary

In this chapter, the content distribution problem as a critical issue in relay-based publish-subscribe networks are discussed. With merely short-distance local transmission channels existing, the content provider has to transmit the contents to the end subscribers only with the help of relays. As the subscribers are heterogeneous, they may have different requirements on QoS of transmission. In this work, the example scenario has been formulated using a mobile social network over IoT architecture with

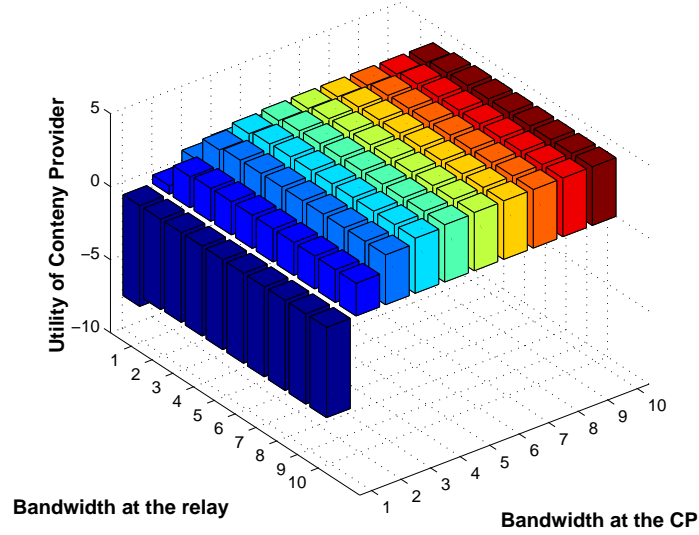


Figure 6.21: Utility of content provider vs. r_1 and r_2 for type- θ_1 subscriber in group \mathbb{G}_R^3 in Case II.

opportunistic contact and local wireless transmission. The author has introduced a contract-theoretic approach to optimize the content delivery strategies with respect to the QoS requirements. To obtain the correct information to make contracting decisions in the proposed system and evaluation the service quality requirements, the concept of delivery ratio within deadline (DRWD) has been adopted as a measurement. Then, to get the distribution of content delay, a tandem queueing model is utilized, which has been solved thereafter by employing absorbing Markov chain model. The numerical results have shown that the proposed contract-theoretic approach is capable of maximizing the profit of the content provider, meanwhile, it can be ensured that heterogeneous subscribers are satisfied with their choices on the contract items, specifying the DRWD they can get and the corresponding cost they have to undertake.

Chapter 7

Conclusions and Future Work

In this chapter, the author summarizes the major research contributions, followed by discussion on the possible future research directions.

7.1 Summary of Contributions

- The author has addressed the tradeoff between improvement (e.g., throughput increase) and corresponding cost (e.g., energy consumption) of cooperative communication, focusing on relay selection in a challenging scenario in the presence of mobile users. According to the user mobility pattern, a DRS scheme based on MDP optimization model has been proposed to minimize the mobile user's long-term average cost. The existence of several SPs providing relay services in the same network area results in price competition among the SPs, since they have to compete for attracting the same group of users. Through in-depth investigation of the interactions between the SPs and MUs, the author has realized the necessity for a joint consideration of the DRS policy design and the price competition among the SPs, and has proposed a hierarchical framework with solution to be

a combinatorial strategy pair including the SP's pricing strategies and the MUs' DRS policies, where both the SPs and the MUs benefit.

- The author has considered a challenging scenario in cognitive relay networks, where the transmission time of the SUs depends on how long the PU's channel is idle. Thus the SUs have incentives to cooperate with the PU's transmission to speed up the PU's transmission and get more transmission opportunities. With the cooperation from the SUs, the PU's transmissions rate will be increased. Consequently, the PU's buffer will be depleted faster, resulting in more channel idle times for the SUs to utilize. Nonetheless, due to the energy limit, if SUs use too much power on relaying, less power will be left for their own transmissions. The author has designed optimal power allocation schemes for both single-SU case and multiple-SU case to address this tradeoff. The power allocation for single-SU case has been formulated as a utility maximization problem, and important propositions on relationship between relay power and self-transmission power have been proposed to achieve effective power allocation of the SU. While in the context of multiple-SU case, a non-cooperative game model has been formulated to achieve the optimal power allocation strategies of the SUs, characterized by the Nash equilibrium. The existence and the uniqueness of the Nash equilibrium have been comprehensively discussed and proved.
- Based on the observation that a joint consideration of cooperation among infostations and mobile users can bring obvious improvement of the network performance in mobile infostation networks, the author has proposed a hierarchical cooperation formation model to analyze the bi-level cooperation. Due to the striking similarity between the cooperation formation problem and cooperative game theory, coalitional formation game and network formation game have been utilized to model the cooperation among the infostations and that among the

mobile users, respectively. A distributed hierarchical cooperation formation algorithm has been proposed to achieve the stable cooperation structure of both the infostations and the mobile users. In the stable cooperation structure, the infostations are stably grouped into coalitions with no incentives to deviate from the current structure. Meanwhile, the mobile users are stably connected with each other and no one can achieve a higher benefit by establishing new cooperation links or breaking from existing cooperation relationships. Verified by the extensive numerical results, the proposed algorithm has exhibited its effectiveness in achieving the stable hierarchical cooperation structure.

- The author has constructed a tree-structured topology to model the connection relationships among components in relay-based publish-subscribe networks, based on which a tandem queueing model has been applied to model the content delivery process along a path from the content provider to a particular subscriber. Capturing the subscribers' major concern on the quality of the content delivery service, the concept of delivery ratio within deadline (DRWD) has been exploited as a measurement of the service quality, which has been derived with the help of absorbing Markov chain analysis. Considering the content delivery service as a special "commodity", content provider as the seller and subscribers as buyers, the relay-based publish-subscribe network can be viewed as a trading market. Due to the power of dominant control on the content delivery process, the content provider is capable to design the market rule in the format of a contract, specifying the service qualities can be provided as well as the corresponding cost. Contract theory has been applied to analyze the behavior of the content provider and the heterogeneous subscribers, and optimal contract design has been proposed to benefit both.

7.2 Future Research Directions

In this dissertation, several challenging problems in relay-based wireless networks have been addressed, where proper game-theoretic models have been utilized to analyze the competitive or cooperative behavior among the network components, leading us to the satisfied solutions. In the following, the author discusses some potential research directions in the future.

7.2.1 Future Work on Dynamic Relay Selection

Note that in the design of the DRS scheme in Chapter 3, the mobile users can either select one relay or multiple relays to help with transmission. Multiple-relay selection can improve the QoS performance, but will bring higher cost as well. The cost can be due to a variety of factors, such as payment to the service providers as the author has considered in Chapter 3, and communication overhead among the users and the relays, as well as that among the relays themselves, which have not been discussed. Due to the consideration of practical implementation, the exploration of such communication overhead is significant, which will be integrated into the proposed DRS scheme. The admission control problem mentioned in Chapter 3 is another issue to be investigated in the future work. The admission control problem will play an important role to guarantee that the proposed DRS scheme works well. In Chapter 3, the author has implicitly assumed that the relays use fixed power in the cooperative communication process, however, the qualities of wireless channels are always fluctuating, which motivates the relays to improve relay power to gain more opportunities to be selected when the channel condition is poor. Based on this consideration, the author will discuss the power control problems of the relays with the purpose of designing an effective scheme to guide the relays to adjust their transmission power to combat with channel fading.

7.2.2 Future Work on Power Allocation in Cognitive Relay Networks

In Chapter 4, the author has made comprehensive study of the power allocation problem for SUs in cognitive relay networks with one PU being considered. In the future, the author plans to generalize the problem to involve multiple PUs, resulting in multiple channels possibly available for the SUs to access. The problem becomes more challenging when an SU is capable of selecting PU to cooperate with, while how many transmission opportunities can be obtained still depends on the strategies of the other SUs (in terms of which PU to choose and how much relay power to spend on relaying). Heterogeneity of the SUs will also be taken into account when designing the power allocation scheme. In particular, queueing model can be applied to analyze the traffic pattern at each SU, which generally differs from each other. Additionally, the author would further investigate the modeling of the possible cooperative behavior among the SUs, where some SUs may form a coalition for efficient channel use and thus higher benefit.

7.2.3 Future Work on Network Formation

In the modeling of the cooperation among the mobile users in Chapter 5, network formation game model has been applied. For forming the network graph, there exists a wide range of approaches which can be grouped into two categories: *myopic* and *far-sighted* [5]. In myopic schemes, the players make their decisions based on the current network state, while far-sighted approaches stimulate the players to adapt their strategies through learning and predicting about the future strategies of the others. As myopic network formation scheme has been applied to the problem discussed in Chapter 5, the author is interested in the possible outcome of the network formation

game where far-sighted algorithms are utilized. With the expectation that far-sighted algorithms may lead us to a more effective and robust network formation, the author plans to include this direction in the author's future work.

In addition, as the author has introduced in Section 2.1, aiming at addressing the drawback of pursuing a Nash network which may lead to trivial graphs, a new network formation game has evolved as a solution, where new concepts of stability such as *conditional stability* and *pairwise stability* [18] are utilized. In Chapter 5, the concept *pairwise stability* has been applied as the rule to guide the network formation procedure, where only deviations by a pair of players in the network formation game are considered. As for coalitional (strong) stability, deviations by groups of players are permitted. Note that the reason that pairwise stability is used as the measurement stability is due to its appealing simplicity in implementation, however, pairwise-stable network may not always exist in network formation games. Therefore, the author intends to consider conditional stability, as a strong refinement of the concept of pairwise stability, in the author's future work on network formation.

7.2.4 Future Work on Contract Design

In Chapter 6, contract-theoretic model has been exploited to model the content delivery in relay-based publish-subscribe networks, in which the content provider signs contract with the subscribers to guarantee that they can obtain certain service qualities at corresponding prices. In particular, in order to achieve a certain service quality, it is required that the bandwidth allocation at the content provider and at the relays should be guaranteed. However, it can be anticipated that the communication overhead between the content providers and the relays will become larger and the content delivery process will become more complicated when the subscribers are far from the

content provider, resulting in the requirement for a series of relays involved in content forwarding. To address this problem and attempt to achieve an effective and efficient content delivery scheme not deteriorated by multiple hierarchies, the author plans to borrow some ideas from delegation, where the delegation and contracting hierarchies are introduced in [107], [108].

Delegation is a common phenomenon in companies, financial institutions and procurement contracting. In particular, owners delegate the power of management to the top level managers, who can in turn delegate management to lower level managers. Taking advantage of the delegation concept, decision making can get efficiently dispersed among a great number of agents, which is apparently superior to the traditional contracting with a central authority required. As one of the potential research directions in the future, the author plans to apply the idea of delegation into the current relay-based publish-subscribe network scenario, where the content provider (the principal) can delegate the management and control of the content delivery process to relays nearby, which can in turn make delegations to the relays further from the content provider and are responsible for relaying content to far-away subscribers. Contract theory is capable to provide an analytical framework for investigating the advantages and disadvantages of delegation, therefore, the author intends to make a more comprehensive study of the contract theory and try to find a satisfied solution to the problem.

Author's Publications

Journals

- **Yifan Li**, Ping Wang, Dusit Niyato, Yang Zhang “Contract-Theoretic Modeling for Content Delivery in Relay-Based Publish-Subscribe Networks”, prepare to submit to IEEE Transactions on Vehicular Technology, 2013.
- **Yifan Li**, Ping Wang, Dusit Niyato, Zhu Han, “A Hierarchical Cooperation Formation Model for Downlink Data Transmission in Mobile Infostation Networks”, IEEE Wireless Communications Magazine, 2012.
- **Yifan Li**, Ping Wang, Dusit Niyato, Weihua Zhuang, “A Hierarchical Framework of Dynamic Relay Selection for Mobile Users and Profit Maximization for Service Providers in Wireless Relay Networks”, Wireless Communications and Mobile Computing, 2012.

Conferences

- **Yifan Li**, Ping Wang, Dusit Niyato, Yang Zhang, “Contract-Theoretic Modeling for Content Delivery in Relay-Based Publish-Subscribe Networks”, submitted to IEEE ICC, 2014.
- Ran Wang, **Yifan Li**, Ping Wang, Dusit Niyato, Walid Saad, Zhu Han, “Design of a V2G Aggregator to Optimize PHEV Charging and Frequency Regulation Control”, IEEE SmartGridComm, 2013.
- **Yifan Li**, Ran Wang, Ping Wang, Dusit Niyato, Walid Saad, Zhu Han, “Resilient PHEV charging policies under price information attacks”, IEEE Smart-GridComm, Tainan City, Taiwan, Nov. 2012.
- **Yifan Li**, R. Kaewpuang, Ping Wang, Dusit Niyato, “An energy efficient solution: Integrating PHEV in smart grid with renewable energy”, IEEE INFOCOM Workshop on Communications and Control for Sustainable Energy Systems: Green Networking and Smart Grids (INFOCOM WS - CCSES), Orlando, FL, USA, Mar. 2012.
- **Yifan Li**, Ping Wang, Dusit Niyato, “On hierarchical cooperation formation in mobile infostation networks”, IEEE Globecom, Houston, TX. USA, Dec. 2011.

-
- Dusit Niyato, X. Zhou, A. Hjrungnes, Ping Wang, **Yifan Li**, “Hierarchical coalition formation game of relay transmission in IEEE 802.16m”, International Conference on Game Theory for Networks (GameNets), Shanghai, China, Apr. 2011 (invited paper).
 - **Yifan Li**, Ping Wang, Dusit Niyato, “Optimal power allocation for secondary users in cognitive relay networks”, IEEE WCNC, Cancun, Mexico, Mar. 2011.
 - **Yifan Li**, Ping Wang, Dusit Niyato, Weihua Zhuang “A dynamic relay selection scheme for mobile users in wireless relay networks”, IEEE INFOCOM (mini-conference), Shanghai, China, Apr., 2011(acceptance ratio = 23%).
 - Dusit Niyato, Ping Wang, Ekram Hossain, **Yifan Li**, “Optimal content transmission policy in publish-subscribe mobile social networks”, IEEE Globecom, Miami, FL. USA, Dec. 2010.

Bibliography

- [1] J. Von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior (Commemorative Edition)*. Princeton university press, 2007.
- [2] M. J. Osborne, *An Introduction to Game Theory*. Oxford University Press, 2003.
- [3] D. Fudenberg and J. Tirole, *Game Theory*. The MIT Press,, 1991.
- [4] F. Carmichael, *A guide to game theory*. Pearson Education, 2005.
- [5] W. S. T. B. Z. Han, D. Niyato and A. Hj?rungenes, *Game Theory in Wireless and Communication Networks: Theory, Models, and Applications*. Cambridge University Press, 2011.
- [6] R. J. La and V. Anantharam, “A game-theoretic look at the gaussian multiaccess channel,” *DIMACS series in discrete mathematics and theoretical computer science*, vol. 66, pp. 87–106, 2004.
- [7] E. Arcaute, R. Johari, and S. Mannor, “Network formation: Bilateral contracting and myopic dynamics,” *Internet and Network Economics*, pp. 191–207, 2007.
- [8] R. J. Aumann, “Correlated equilibrium as an expression of bayesian rationality,” *Econometrica: Journal of the Econometric Society*, pp. 1–18, 1987.
- [9] J. G. Wardrop, “Some theoretical aspects of road traffic research,” in *Inst Civil Engineers Proc London/UK/*, 1900.
- [10] A. Sendonaris, E. Erkip, and B. Aazhang, “User cooperation diversity. part i. system description,” *IEEE Trans. Communications*, vol. 51, pp. 1927–1938, 2003.
- [11] Y. Wei, F. Yu, and M. Song, “Distributed optimal relay selection in wireless cooperative networks with finite-state markov channels,” *IEEE Trans. Vehicular Technology*, vol. 59, no. 5, pp. 2149 –2158, Jun 2010.

-
- [12] A. Nosratinia, T. Hunter, and A. Hedayat, "Cooperative communication in wireless networks," *IEEE Communications Magazine*, vol. 42, no. 10, pp. 74–80, Oct. 2004.
- [13] H. Yu Wei and R. Gitlin, "Two-hop-relay architecture for next-generation wwan/wlan integration," *Wireless Communications, IEEE*, vol. 11, no. 2, pp. 24 – 30, apr. 2004.
- [14] D. Goodman, J. Borras, N. Mandayam, and R. Yates, "Infostations: a new system model for data and messaging services," in *Proc. of IEEE Vehicular Technology Conference (VTC)*, 1997.
- [15] P. Gupta and P. Kumar, "The capacity of wireless networks," *IEEE Trans. Information Theory*, vol. 46, no. 2, pp. 388 –404, Mar. 2000.
- [16] W. H. Yuen, S.-C. Mau, and R. Yates, "Existence of data and multiuser diversities in noncooperative mobile infostation networks," *IEEE Trans. Mobile Computing*, vol. 8, no. 8, pp. 1117 –1131, August 2009.
- [17] W. Saad, Z. Han, M. Debbah, A. Hjørungnes, and T. Basar, "Coalitional game theory for communication networks," *IEEE Signal Processing Magazine*, vol. 26, pp. 77 –97, 2009.
- [18] M. O. Jackson, "A survey of models of network formation: Stability and efficiency," EconWPA, Game Theory and Information, 2003.
- [19] T. Basar and G.J.Olsder, *Dynamic Non-cooperative Game Theory*. SIAM Series in Classics in Applied Mathematics, Society for Industrial and Applied Mathematics, 1999.
- [20] C. St Jean and B. Jabbari, "Bayesian game-theoretic modeling of transmit power determination in a self-organizing cdma wireless network," in *Vehicular Technology Conference, 2004. VTC2004-Fall. 2004 IEEE 60th*, vol. 5. IEEE, 2004, pp. 3496–3500.
- [21] R. KRZYSZTOF and A. Witzel, "A generic approach to coalition formation," *International Game Theory Review*, vol. 11, no. 03, pp. 347–367, 2009.
- [22] D. Ray, *A Game-Theoretic Perspective on Coalition Formation*. Oxford University Press, 2007.
- [23] R. B. Myerson, "Graphs and cooperation in games," *Mathematics of operations research*, vol. 2, no. 3, pp. 225–229, 1977.

-
- [24] L. Stole, "Lectures on the theory of contracts and organizations," *Manuscript, University of Chicago Graduate School of Business*, 2001.
- [25] P. Bolton and M. Dewatripont, *Contract theory*. MIT press, 2005.
- [26] W. Ni, G. Shen, S. Jin, T. Fahldieck, and R. Muenzner, "Cooperative relay in IEEE 802.16 j mmr," *IEEE C802. 16j-06/006r1, IEEE 802.16 Broadband Wireless Access Working Group*, 2006.
- [27] P. Liu, Z. Tao, Z. Lin, E. Erkip, and S. Panwar, "Cooperative wireless communications: a cross-layer approach," *Wireless Communications, IEEE*, vol. 13, no. 4, pp. 84–92, Aug. 2006.
- [28] Y. Zhao, R. Adve, and T. Lim, "Improving amplify-and-forward relay networks: optimal power allocation versus selection," *IEEE Trans. Wireless Communications*, vol. 6, no. 8, pp. 3114–3123, August 2007.
- [29] A. Sadek, Z. Han, and K. Liu, "A distributed relay-assignment algorithm for cooperative communications in wireless networks," in *Proc. of IEEE International Conference on Communications (ICC)*, Istanbul, Turkey, 2006, pp. 1592–1597.
- [30] A. Bletsas, A. Khisti, D. Reed, and A. Lippman, "A simple cooperative diversity method based on network path selection," *IEEE Select. Areas Communications*, vol. 24, no. 3, pp. 659–672, March 2006.
- [31] A. Ribeiro, X. Cai, and G. Giannakis, "Symbol error probabilities for general cooperative links," *IEEE Trans. Wireless Communications*, vol. 4, no. 3, pp. 1264–1273, May 2005.
- [32] A. Bletsas, H. Shin, and M. Win, "Outage optimality of opportunistic amplify-and-forward relaying," *Communications Letters, IEEE*, vol. 11, no. 3, pp. 261–263, March 2007.
- [33] A. Bletsas, A. Lippman, and D. Reed, "A simple distributed method for relay selection in cooperative diversity wireless networks, based on reciprocity and channel measurements," in *Proc. of IEEE VTC*, 2005.
- [34] Z. Zhou, S. Zhou, J.-H. Cui, and S. Cui, "Energy-efficient cooperative communication based on power control and selective single-relay in wireless sensor networks," *Wireless Communications, IEEE Transactions on*, vol. 7, no. 8, pp. 3066–3078, 2008.

-
- [35] P. Zhang, Z. Xu, F. Wang, X. Xie, and L. Tu, "A relay assignment algorithm with interference mitigation for cooperative communication," in *Wireless Communications and Networking Conference, 2009. WCNC 2009. IEEE*, 2009, pp. 1–6.
- [36] H. Shan, H. T. Cheng, and W. Zhuang, "Cross-layer cooperative mac protocol in distributed wireless networks," *IEEE Transactions on Wireless Communications*, vol. 10, no. 8, pp. 2603–2615, Aug. 2011.
- [37] Y. Jing and H. Jafarkhani, "Single and multiple relay selection schemes and their achievable diversity orders," *IEEE Trans. Wireless Communications*, vol. 8, no. 3, pp. 1414–1423, March 2009.
- [38] F. Onat, Y. Fan, H. Yanikomeroglu, and H. Poor, "Threshold based relay selection in cooperative wireless networks," in *Proc. of IEEE GLOBECOM*, nov. 2008, pp. 1–5.
- [39] S. Zhang and V. Lau, "Design and analysis of multi-relay selection for cooperative spatial multiplexing," in *Proc. of ICC*, 2008.
- [40] A. Nosratinia and T. E. Hunter, "Grouping and partner selection in cooperative wireless networks," *Selected Areas in Communications, IEEE Journal on*, vol. 25, no. 2, pp. 369–378, February 2007.
- [41] D. Michalopoulos, G. Karagiannidis, T. Tsiftsis, and R. Mallik, "Wlc41-1: An optimized user selection method for cooperative diversity systems," in *Global Telecommunications Conference, 2006. GLOBECOM '06. IEEE*, 2006, pp. 1–6.
- [42] I.M.Weinstein, "Real world convergence for conferencing: Considerations for hosting ip video on the enterprise data network," *Wainhouse Research Issue Paper*, July 2005.
- [43] D. Charilas, O. Markaki, D. Nikitopoulos, and M. Theologou, "Packet-switched network selection with the highest qos in 4g networks," *Computer Networks*, vol. 52, no. 1, pp. 248–258, 2008.
- [44] A. Iera, A. Molinaro, C. Campolo, and M. Amadeo, "An access network selection algorithm dynamically adapted to user needs and preferences," in *Personal, Indoor and Mobile Radio Communications, 2006 IEEE 17th International Symposium on*. IEEE, 2006, pp. 1–5.
- [45] H. Lin, M. Chatterjee, S. Das, and K. Basu, "Arc: an integrated admission and rate control framework for competitive wireless cdma data networks using

-
- noncooperative games,” *Mobile Computing, IEEE Transactions on*, vol. 4, no. 3, pp. 243–258, 2005.
- [46] Y. Xi and E. Yeh, “Pricing, competition, and routing for selfish and strategic nodes in multi-hop relay networks,” in *Proc. of IEEE INFOCOM*, 2008.
- [47] N. Shastri and R. Adve, “Stimulating cooperative diversity in wireless ad hoc networks through pricing,” in *Proc. of IEEE ICC*, 2006.
- [48] B. Wang, Z. Han, and K. Liu, “Distributed relay selection and power control for multiuser cooperative communication networks using buyer/seller game,” in *INFOCOM 2007. 26th IEEE International Conference on Computer Communications. IEEE*, May 2007, pp. 544–552.
- [49] K. R. Liu and B. Wang, *Cognitive radio networking and security: A game-theoretic view*. Cambridge University Press, 2010.
- [50] F. C. Commission *et al.*, “Spectrum policy task force report, fcc 02-155,” 2002.
- [51] S. Haykin, “Cognitive radio: brain-empowered wireless communications,” *Selected Areas in Communications, IEEE Journal on*, vol. 23, no. 2, pp. 201–220, 2005.
- [52] K. R. Liu, A. K. Sadek, W. Su, and A. Kwasinski, *Cooperative communications and networking*. Cambridge University Press, 2009.
- [53] J. Laneman, D. Tse, and G. Wornell, “Cooperative diversity in wireless networks: Efficient protocols and outage behavior,” *Information Theory, IEEE Transactions on*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.
- [54] O. Simeone, I. Stanojev, S. Savazzi, Y. Bar-Ness, U. Spagnolini, and R. Pickholtz, “Spectrum leasing to cooperating secondary ad hoc networks,” *Selected Areas in Communications, IEEE Journal on*, vol. 26, no. 1, pp. 203–213, jan. 2008.
- [55] J. Zhang and Q. Zhang, “Stackelberg game for utility-based cooperative cognitive radio networks,” in *Proc. of ACM international symposium on Mobile ad hoc networking and computing (MobiHoc)*, 2009.
- [56] S. Jayaweera and T. Li, “Dynamic spectrum leasing in cognitive radio networks via primary-secondary user power control games,” *Wireless Communications, IEEE Transactions on*, vol. 8, no. 6, pp. 3300–3310, 2009.
- [57] J. Mietzner, L. Lampe, and R. Schober, “Distributed transmit power allocation for multihop cognitive-radio systems,” *Wireless Communications, IEEE Transactions on*, vol. 8, no. 10, p. 5187, 2009.

-
- [58] C. Yang, J. Li, and Z. Tian, "Optimal power control for cognitive radio networks under coupled interference constraints: A cooperative game-theoretic perspective," *Vehicular Technology, IEEE Transactions on*, vol. 59, no. 4, 2010.
- [59] G. Rajappan, J. Acharya, H. Liu, N. Mandayam, I. Seskar, and R. Yates, "Mobile infostation network technology," in *Proceedings of SPIE*, vol. 6248, 2006, pp. 62 480M1–62 480M9.
- [60] A. L. Iacono and C. Rose, "Mine, mine, mine: information theory, infostation networks, and resource sharing," in *Wireless Communications and Networking Conference, 2000. WCNC. 2000 IEEE*, vol. 3. IEEE, 2000, pp. 1541–1546.
- [61] L. Galluccio, G. Morabito, A. Moustakas, and S. Palazzo, "Opportunistic communications in infostation systems: Delay and stability analysis," in *Global Telecommunications Conference (GLOBECOM 2010), 2010 IEEE*, 2010, pp. 1–5.
- [62] M. Grossglauser and D. Tse, "Mobility increases the capacity of ad-hoc wireless networks," in *Proc. of IEEE International Conference on Computer Communications (INFOCOM)*, 2001.
- [63] W. H. Yuen and R. D. Yates, "Poster: Optimum transmit range and capacity of mobile infostation networks."
- [64] M. Neely and E. Modiano, "Capacity and delay tradeoffs for ad hoc mobile networks," *IEEE Trans. Information Theory*, vol. 51, no. 6, pp. 1917 – 1937, June 2005.
- [65] T. Small and Z. J. Haas, "The shared wireless infostation model: a new ad hoc networking paradigm," in *Proc. of ACM international symposium on Mobile ad hoc networking and computing (MobiHoc)*, 2003.
- [66] M. Papadopouli and H. Schulzrinne, "Effects of power conservation, wireless coverage and cooperation on data dissemination among mobile devices," in *Proc. of ACM international symposium on Mobile ad hoc networking and computing (MobiHoc)*, 2001.
- [67] L. Galluccio, A. Leonardi, G. Morabito, and S. Palazzo, "Timely and energy-efficient communications in rural infostation systems," *IEEE Trans. Wireless Communications*, vol. 15, no. 3, pp. 48 –53, June 2008.
- [68] H. Liang and W. Zhuang, "Cross-layer resource allocation for efficient message dissemination in rural infostation systems," in *Proc. of IEEE Global Telecommunications Conference (GLOBECOM)*, 2009.

-
- [69] M. Tlais and F. Weis, "Data exchanging in hierarchical infostation systems," in *Parallel Processing Workshops, 2007. ICPPW 2007. International Conference on*. IEEE, 2007, pp. 10–10.
- [70] F. Malandrino, C. Casetti, C. Chiasserini, and M. Fiore, "Content downloading in vehicular networks: What really matters," in *INFOCOM, 2011 Proceedings IEEE*, 2011, pp. 426–430.
- [71] S. Sahebgharani and M. Shahverdy, "A scheduling algorithm for downloading data from rsu using multicast technique," in *Information Technology: New Generations (ITNG), 2012 Ninth International Conference on*, 2012, pp. 809–814.
- [72] T. Wang, L. Song, and Z. Han, "Collaborative data dissemination in cognitive vanets with sensing-throughput tradeoff," in *Communications in China (ICC), 2012 1st IEEE International Conference on*, 2012, pp. 41–45.
- [73] F. Malandrino, C. Casetti, C. F. Chiasserini, C. Sommer, and F. Dressler, "Content downloading in vehicular networks: Bringing parked cars into the picture," in *Personal Indoor and Mobile Radio Communications (PIMRC), 2012 IEEE 23rd International Symposium on*, 2012, pp. 1534–1539.
- [74] W.-H. Kuo, Y.-S. Tung, and S.-H. Fang, "A node management scheme for r2v connections in rsu-supported vehicular adhoc networks," in *Computing, Networking and Communications (ICNC), 2013 International Conference on*, 2013, pp. 768–772.
- [75] E. P. Jones, L. Li, J. K. Schmidtke, and P. A. Ward, "Practical routing in delay-tolerant networks," *Mobile Computing, IEEE Transactions on*, vol. 6, no. 8, pp. 943–959, 2007.
- [76] W. Zhao, M. Ammar, and E. Zegura, "Multicasting in delay tolerant networks: semantic models and routing algorithms," in *Proceedings of the 2005 ACM SIGCOMM workshop on Delay-tolerant networking*. ACM, 2005, pp. 268–275.
- [77] Q. Li, S. Zhu, and G. Cao, "Routing in socially selfish delay tolerant networks," in *INFOCOM, 2010 Proceedings IEEE*. IEEE, 2010, pp. 1–9.
- [78] E. Yoneki, P. Hui, S. Chan, and J. Crowcroft, "A socio-aware overlay for publish/subscribe communication in delay tolerant networks," in *Proceedings of the 10th ACM Symposium on Modeling, analysis, and simulation of wireless and mobile systems*. ACM, 2007, pp. 225–234.
- [79] A. Krifa, C. Baraka, and T. Spyropoulos, "Optimal buffer management policies for delay tolerant networks," in *Sensor, Mesh and Ad Hoc Communications and*

-
- Networks, 2008. SECON'08. 5th Annual IEEE Communications Society Conference on.* IEEE, 2008, pp. 260–268.
- [80] C. M. Sadler and M. Martonosi, “Data compression algorithms for energy-constrained devices in delay tolerant networks,” in *Proceedings of the 4th international conference on Embedded networked sensor systems.* ACM, 2006, pp. 265–278.
- [81] H. Yin, “Location based service,” in *T-109.551 Research Seminar on Location Business II*, 2002.
- [82] W. Gao, Q. Li, B. Zhao, and G. Cao, “Multicasting in delay tolerant networks: a social network perspective,” in *Proceedings of the tenth ACM international symposium on Mobile ad hoc networking and computing.* ACM, 2009, pp. 299–308.
- [83] W. Yu and A. Siddiqui, “Towards a wireless mobile social network system design in healthcare,” in *Multimedia and Ubiquitous Engineering, 2009. MUE '09. Third International Conference on*, 2009, pp. 429–436.
- [84] S. Ioannidis, A. Chaintreau, and L. Massoulie, “Optimal and scalable distribution of content updates over a mobile social network,” in *INFOCOM 2009, IEEE*, 2009, pp. 1422–1430.
- [85] P. T. Eugster, P. A. Felber, R. Guerraoui, and A.-M. Kermarrec, “The many faces of publish/subscribe,” *ACM Computing Surveys (CSUR)*, vol. 35, no. 2, pp. 114–131, 2003.
- [86] Y. Huang and H. Garcia-Molina, “Publish/subscribe in a mobile environment,” *Wireless Networks*, vol. 10, no. 6, pp. 643–652, 2004.
- [87] G. Banavar, T. Chandra, B. Mukherjee, J. Nagarajao, R. Strom, and D. Sturman, “An efficient multicast protocol for content-based publish-subscribe systems,” in *Distributed Computing Systems, 1999. Proceedings. 19th IEEE International Conference on*, 1999, pp. 262–272.
- [88] L. Gao, X. Wang, Y. Xu, and Q. Zhang, “Spectrum trading in cognitive radio networks: A contract-theoretic modeling approach,” *Selected Areas in Communications, IEEE Journal on*, vol. 29, no. 4, pp. 843–855, 2011.
- [89] L. Duan, L. Gao, and J. Huang, “Contract-based cooperative spectrum sharing,” in *New Frontiers in Dynamic Spectrum Access Networks (DySPAN), 2011 IEEE Symposium on.* IEEE, 2011, pp. 399–407.

-
- [90] L. Duan, T. Kubo, K. Sugiyama, J. Huang, T. Hasegawa, and J. Walrand, "Incentive mechanisms for smartphone collaboration in data acquisition and distributed computing," in *INFOCOM, 2012 Proceedings IEEE*, 2012, pp. 1701–1709.
- [91] Y. Luo, L. Gao, and J. Huang, "Spectrum broker by geo-location database," in *IEEE GlobeCom*, 2012.
- [92] D. Niyato and P. Wang, "Optimization of the mobile router and traffic sources in vehicular delay-tolerant network," *IEEE Trans. Vehicular Technology*, vol. 58, no. 9, pp. 5095–5104, Nov. 2009.
- [93] D. Niyato and P. Wang, "Credit-based spectrum sharing for cognitive mobile multihop relay networks," in *Proc. of IEEE WCNC*, 2010.
- [94] R. Nabar, H. Bolcskei, and F. Kneubuhler, "Fading relay channels: performance limits and space-time signal design," *Selected Areas in Communications, IEEE Journal on*, vol. 22, no. 6, pp. 1099–1109, Aug. 2004.
- [95] B. Can, H. Yomo, and E. Carvalho, "Link adaptation and selection method for ofdm based wireless relay networks," *Journal of Communications and Networks*, 2007.
- [96] M. Puterman, "Markov decision processes: Discrete stochastic dynamic programming," *IMA Journal of Management Mathematics*, 1994.
- [97] D. Niyato, S. Chaisiri, and B.-S. Lee, "Economic analysis of resource market in cloud computing environment," in *Proc. of IEEE APSCC*, 2009.
- [98] J. Laneman and G. Wornell, "Distributed space-time-coded protocols for exploiting cooperative diversity in wireless networks," *Information Theory, IEEE Transactions on*, vol. 49, no. 10, pp. 2415–2425, Oct. 2003.
- [99] R. Cooper, "Introduction to queuing theory," *Edward Arnold Limited*, 1981.
- [100] R. Loynes, "The stability of a queue with non-independent inter-arrival and service times," in *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 58, no. 03. Cambridge University Press, 2008, pp. 497–520.
- [101] J. Friedman, *Game theory with applications to economics*. Oxford University Press, USA, 1990.
- [102] J. Rosen, "Existence and uniqueness of equilibrium points for concave n-person games," *Econometrica: Journal of the Econometric Society*, vol. 33, no. 3, pp. 520–534, 1965.

- [103] D. Niyato, P. Wang, E. Hossain, W. Saad, and A. Hjørungnes, “Exploiting mobility diversity in sharing wireless access: A game theoretic approach,” *IEEE Trans. Wireless Communications*, vol. 9, no. 12, pp. 3866–3877, December 2010.
- [104] W. Saad, Z. Han, M. Debbah, and A. Hjørungnes, “A distributed merge and split algorithm for fair cooperation in wireless networks,” in *Proc. of IEEE International Conference on Communications (ICC) Communications Workshops*, 2008.
- [105] D. Niyato, P. Wang, E. Hossain, and Y. Li, “Optimal content transmission policy in publish-subscribe mobile social networks,” in *Global Telecommunications Conference (GLOBECOM 2010), 2010 IEEE*, 2010, pp. 1–5.
- [106] S. Drekcic, “Phase-type distribution,” *Encyclopedia of Quantitative Finance*, 2006.
- [107] D. Mookherjee, “Delegation and contracting hierarchies: An overview,” *Boston University, manuscript*, 2003.
- [108] C. Choe and I.-U. Park, “Delegated contracting and corporate hierarchies,” *Available at SSRN 470382*, 2003.