Wireless Network Design and Optimization:

From Social Awareness to Security

by

Xiaowen Gong

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved April 2015 by the
Graduate Supervisory Committee:

Junshan Zhang, Chair
Douglas Cochran
Lei Ying
Yanchao Zhang

ARIZONA STATE UNIVERSITY

May 2015

ABSTRACT

A principal goal of this dissertation is to study wireless network design and optimization with the focus on two perspectives: 1) socially-aware mobile networking and computing; 2) security and privacy in wireless networking. Under this common theme, this dissertation can be broadly organized into three parts.

The first part studies socially-aware mobile networking and computing. First, it studies random access control and power control under a social group utility maximization (SGUM) framework. The socially-aware Nash equilibria (SNEs) are derived and analyzed. Then, it studies mobile crowdsensing under an incentive mechanism that exploits social trust assisted reciprocity (STAR). The efficacy of the STAR mechanism is thoroughly investigated. Next, it studies mobile users' data usage behaviors under the impact of social services and the wireless operator's pricing. Based on a two-stage Stackelberg game formulation, the user demand equilibrium (UDE) is analyzed in Stage II and the optimal pricing strategy is developed in Stage I. Last, it studies opportunistic cooperative networking under an optimal stopping framework with two-level decision-making. For both cases with or without dedicated relays, the optimal relaying strategies are derived and analyzed.

The second part studies radar sensor network coverage for physical security. First, it studies placement of bistatic radar (BR) sensor networks for barrier coverage. The optimality of line-based placement is analyzed, and the optimal placement of BRs on a line segment is characterized. Then, it studies the coverage of radar sensor networks that exploits the Doppler effect. Based on a Doppler coverage model, an efficient method is devised to characterize Doppler-covered regions and an algorithm is developed to find the minimum radar density required for Doppler coverage.

The third part studies cyber security and privacy in socially-aware networking

and computing. First, it studies random access control, cooperative jamming, and spectrum access under an extended SGUM framework that incorporates negative social ties. The SNEs are derived and analyzed. Then, it studies pseudonym change for personalized location privacy under the SGUM framework. The SNEs are analyzed and an efficient algorithm is developed to find an SNE with desirable properties.

To my parents.

ACKNOWLEDGEMENTS

I would like to express utmost gratitude to my advisor, Professor Junshan Zhang, for his constant support and encouragement. His sincere enthusiasm, strong dedication, and serious attitude towards research always motivate me and drive me to do the best of mine. Moreover, he is very kind and always ready to provide help to me in both work and life. He has been an excellent mentor for me during my PhD study. I am very fortunate to have him as my advisor.

I would like to express my sincere gratitude to Professor Douglas Cochran who has provided inspiring comments, useful advice and generous support for my research. I am very grateful to Professors Lei Ying and Yanchao Zhang for serving as my committee members. I also want to thank Professors Jesse Taylor, Vincent Poor, Tolga Duman, Cihan Tepedelenlioglu, and Oliver Kosut. I have greatly benefited from their diverse expertise, inspiring discussions, and helpful suggestions.

Many thanks to my current and former colleagues: Chandrashekhar Thejaswi, Shanshan Wang, Lei Yang, Dajun Qian, Miao He, Brian Proulx, Ling Tang, Mojgan Hedayati, David Ganger, Yang Cao, Xu Chen, Shibo He, Dong-Hoon Shin, Chuan Huang, Mengyuan Zhang, Zhengyu Zhang, Weina Wang, Dejun Yang, Rui Zhang, and Jingchao Sun for the pleasant and inspiring discussions. I also thank my collaborators: Lingjie Duan and Kai Xing for the insightful discussions I had with them.

Finally, I would like to thank my mother Guanglin Dong and my father Xiguang Gong for their endless love and support, without whom I cannot have come this far.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figure Page

Figure Page

Chapter 1

INTRODUCTION

1.1 Overview

Wireless networks have been widely recognized as an integral part of the next generation infrastructure in a wide range of fields, including communication, energy, transportation, manufacturing, and healthcare. Generally speaking, a wireless network nowadays is a system of wirelessly interconnected computational elements embedded in physical entities, such as mobile networks, smart grids, vehicular networks, and e-health systems. The ubiquitous integration of wireless connectivity and computational power provides enormous opportunity to "intelligently" engineer the physical systems in an unprecedented way. Recent technology advances have dramatically enhanced the functionality of individual nodes of wireless networks in various aspects such as sensing, computing, communication, and control. To fully exploit the potential benefits of wireless networks, a great deal of research have been conducted to make them towards more efficient, reliable, and robust systems of coordinated components working in concert. In particular, research in mobile networking and computing has become a key driving force to boost the performance of wireless networks. Different from earlier generations of physical engineering systems, nowadays wireless networks are characterized by complex interactions among the nodes due to the coupling in both cyber and physical domains. This emerging feature of wireless networks poses difficulties for their design and control. Another challenge rises from the growing needs for security and privacy in the applications of wireless networks. To overcome these challenges, this dissertation studies wireless network design and optimization with the focus on two perspectives: 1) socially-aware mobile networking and computing; 2) security and privacy in wireless networking.

1.1.1 Socially-aware Mobile Networking and Computing

With explosive growth of the Internet, online social networks (e.g., Facebook, Twitter) have gained widespread popularity over the past few years. In 2013, the number of online social network users worldwide has crossed 1.73 billion, nearly one quarter of the world's population [1]. With pervasive penetration of the Internet in people's daily life, online social services have dramatically facilitated people's social interactions with each other, and thereby tighten people's social relationships in an unprecedented way.

Mobile networks have been growing rapidly in the past few years and this trend will continue in the foreseeable future. In 2014, mobile phone shipments are projected to reach 1.9 billion units, which is about 7 times that of desktop and laptop combined [2]. The widespread popularity of mobile networks has been driven by continuing technology advances. On one hand, advanced wireless communication technologies (e.g., MIMO, OFDM) have drastically improved the communication efficiency in wireless networks (e.g., cellular networks, WLANs). On the other hand, advanced mobile devices (e.g. smartphones, tablets) equipped with powerful sensors (e.g., cameras) and great computing capability have enabled a wide range of applications on mobile platforms. As a result, mobile networks have nowadays become an indispensable infrastructure in people's everyday life.

As mobile devices are carried and used personally by mobile users, the users' behaviors on the mobile devices can be significantly influenced by their social relationships. This convergence of social networks and mobile networks gives rise to the interplay of these two traditionally disjoint domains. In particular, when mobile users take actions pertaining to their mobile devices, from physical layer parameter settings (e.g., transmit power selection) to application layer activities (e.g., data usage

behavior), they are aware of their social relationships with each other and would take into account the impacts of their actions on their social neighbors. The social aspect of mobile networking and computing is an emerging paradigm for wireless network design and optimization. A survey of mobile social networking can be found in [3].

In this dissertation, we study socially-aware mobile networking and computing with applications in mobile crowdsensing, wireless data pricing, and distributed opportunistic scheduling.

Mobile Crowdsensing

Mobile crowdsensing has recently emerged as a promising paradigm for a variety of applications, thanks to the pervasive penetration of mobile devices to people's daily lives. Indeed, with the development of 4G networks and powerful processors, smartphone sales crossed 1 billion units in 2013 [2]. As smartphones are equipped with advanced sensors such as accelerometers, compasses, gyroscopes, and cameras, they can collectively perform many sensing tasks, e.g., monitoring the environment. In a nutshell, by leveraging a crowd of mobile users, one can collect and process sensed data far beyond the scope of what was possible before.

Although the benefit of crowdsensing is pronounced, performing a sensing task typically incurs overhead for participating users, in terms of the users' resource consumption devoted to sensing, such as battery and computing power. Further, the participating users also take the risk of potential privacy loss by sharing their sensed data with others. As a result, a user may not participate in sensing without receiving adequate incentive. Therefore, effective incentive design is essential for realizing the benefit of crowdsensing.

Wireless Data Pricing

With explosive growth of the Internet and rapid advances of mobile device (e.g., smartphones, tablets), wireless data traffic has grown tremendously in the past few years. The proliferation of wireless data demand is largely driven by the popularity of bandwidth-intensive applications on mobile platforms such as online social services and video streaming. Indeed, mobile data traffic is predicted to increase by over 100 times in the next ten years [4]. This poses challenges for wireless operators (e.g., AT&T, Verizon) to consistently provide good quality services as heavy data traffic can result in serious congestion which degrade mobile user's experiences. On the other hand, it also brings opportunities for wireless operators to boost their revenue thanks to the increase of data usage. To address the challenges and exploit the opportunities, it is important to judiciously price wireless services in order to control mobile data usage in a desirable way.

Distributed Opportunistic Scheduling

There have been extensive research on channel-aware scheduling, which exploit the rich diversity inherent in wireless communications by scheduling links with favorable channel conditions [5, 6]. While most existing studies focus on centralized scheduling (see, e.g., [5–7]), a distributed opportunistic scheduling (DOS) framework is developed for ad hoc networks in [8], in which multi-user diversity and time diversity in wireless channels are exploited jointly in a distributed manner. Assuming perfect channel estimation, it is shown that the optimal opportunistic scheduling in such a scenario is intimately related to the fundamental tradeoff between throughput gain from better channel conditions and the cost of further probing.

1.1.2 Security and Privacy in Wireless Networking

Security has always been a critical issue for many important applications, as its failure can lead to enormous loss and cost. Physical security is concerned with security measures that are intended to deny unauthorized access to facilities, territories and resources, so as to protect them from damage or harm (e.g., espionage, theft, or attacks). Thanks to recent technology advances, sensor networks have shown great potential in a wide range of applications for physical security (e.g., intruder detection, security surveillance). To fully exploit this potential, it calls for strategic design of sensor networks that takes into account all possible attacks from adversaries.

With pervasive penetration of advanced wireless devices into every aspect of people's life, cyber security and privacy have raised broad and serious concerns over the past few years. Wireless networks are vulnerable by nature as wireless medium is shared resource that can be accessed by any wireless device. This security weakness can be potentially exploited by a number of attacks (e.g., eavesdropping, jamming). Another vulnerability comes from the fact that existing wireless devices are widely used by personal users for mobile communications and computing. As a result, personal users' private data are transferred and processed over the Internet through mobile networks, which can be compromised by possible adversaries. Therefore, wireless network design is expected to fulfill their primary goals while meeting various security and privacy needs arising from specific applications.

In this dissertation, we study security and privacy in wireless networking with applications in sensor network coverage and location privacy.

Sensor Network Coverage

Coverage, which defines how well the object of interest is monitored, is a critical performance metric for sensor networks [9]. While an individual sensor's coverage

depends on the sensing principles of different types of sensors, the collective coverage of a network of geographically distributed sensor nodes is of central importance for sensor network design. Various coverage problems arise depending on different network contexts and application objective. A typical class of coverage problems belongs to area coverage, where the objective is to cover the entire area of interest. Barrier coverage is another objective for coverage, which aims to build a barrier of sensors across an area of interest such that any object traversing the area can be detected.

Location Privacy

With rapid growth of mobile networking and computing, location-based services (LBSs) have become increasingly popular recently as locations are useful information for a wide range of applications (e.g., location-based navigation or recommendation). Most mobile devices nowadays are capable of localization (e.g., by GPS or wireless access points). Mobile users send their locations to a LBS provider for a certain LBS, and the LBS provider feedbacks the desired results to the users based on their reported locations. However, the providers of LBSs can not trustworthy, due to the risk of leaking users' location information to other parties (e.g., sell users' location data). As a result, mobile users are exposed to potential privacy threats when using a LBS.

1.2 Summary of Main Contributions

The main body of this dissertation can be organized into three parts. The first part (Chapters 2, 3, 4, 5) studies socially-aware mobile networking and computing with the focus on random access control, power control, mobile crowdsensing, wireless data pricing, and distributed opportunistic scheduling. Chapter 2 studies random access control and power control under a social group utility maximization framework. Chapter 3 studies mobile crowdsensing under an incentive mechanism that exploits

social trust assisted reciprocity. Chapter 4 studies mobile users' data usage behaviors under the impact of social services and the wireless operator's pricing. Chapter 5 studies opportunistic cooperative networking under an optimal stopping framework with two-level decision-making. The second part (Chapters 6, 7) studies physical security in wireless networking with the focus on radar sensor network coverage. Chapter 6 studies optimal placement of bistatic radar sensor networks for barrier coverage. Chapter 7 studies the coverage of radar sensor networks that exploits the Doppler effect. The third part (Chapters 8, 9) studies cyber security and privacy in socially-aware networking and computing with the focus on random access control, cooperative jamming, and location privacy. Chapters 8 studies random access control, cooperative jamming, and spectrum access under a social group utility maximization framework that incorporates negative social ties. Chapters 9 studies pseudonym change for personalized location privacy under a social group utility maximization framework. We summarize the main contributions of each chapter as follows.

Chapter 2 studies mobile users' altruistic behaviors based on their social ties. We develop a social group utility maximization (SGUM) framework which captures mobile users' diverse social ties and mobile devices' diverse physical relationships in a unified way. Specifically, instead of maximizing one's individual utility, each user aims to maximize its social group utility based on its social ties with other users. A primary merit of this framework is that it spans the continuum between non-cooperative game (NCG) and network utility maximization (NUM) - two traditionally disjoint paradigms for network optimization. Under the SGUM framework, we study two important applications in wireless networks: random access control and power control. We first derive socially-aware Nash equilibria (SNEs) for the SGUM based random access control and power control, respectively. Then we show that as social ties increase, each user's SNE strategy migrates monotonically from the Nash equilibrium

strategy for a standard NCG to the social optimal strategy for NUM, and the social welfare of the SNE also increases. This shows that the SGUM framework captures NCG and NUM as two special cases and spans the continuum in between.

Chapter 3 studies a socially-aware crowdsensing system which incentivizes mobile users to participate in sensing tasks by leveraging their social trust. For this system, we exploit social trust assisted reciprocity (STAR), a synergistic marriage of social trust and reciprocity, to design an incentive mechanism that stimulates users' participation. Given the social trust structure among users, we thoroughly investigate the efficacy of STAR for satisfying users' sensing requests. Specifically, we first show that all requests can be satisfied if and only if sufficient social credit can be "transferred" from users who request more sensing service than what they can provide to users who can provide more than what they request. Then we investigate utility maximization for sensing services, and show that it boils down to maximizing the utility of a circulation flow in the combined social graph and request graph. Accordingly, we develop an algorithm that iteratively cancels a cycle of positive weight in the residual graph, which finds the optimal solution efficiently, for both cases of divisible and indivisible sensing service. Extensive simulation results corroborate that STAR can significantly outperform the mechanisms using social trust only or reciprocity only.

Chapter 4 studies mobile users' data usage behaviors by jointly considering the network effect based on their social relationships in the social domain and the congestion effect in the physical wireless domain. Accordingly, we develop a Stackelberg game for problem formulation: In Stage I, a wireless provider first decides the data pricing to all users to maximize its revenue, and then in Stage II, users observe the price and decide data usage subject to mutual interactions under both network and congestion effects. We analyze the two-stage game using backward induction. For

Stage II, we first show the existence and uniqueness of a user demand equilibrium (UDE). Then we propose a distributed update algorithm for users to reach the UDE. We further investigate the impact of different parameters on the UDE. For Stage I, we develop an optimal pricing algorithm to maximize the wireless provider's revenue. We evaluate the performances of the developed algorithm by simulation results using real data.

Chapter 5 studies opportunistic cooperative networking (OCN) in wireless ad hoc networks, with a focus on characterizing the desired tradeoff between the probing cost for establishing cooperative relaying and hence higher throughput via opportunistic cooperative networking. Specifically, we treat opportunistic cooperative networking as an optimal stopping problem with two-levels of incomplete information. We consider the cases with or without dedicated relays, and we establish the existence of optimal strategies for both cases. Then we show that for the case with dedicated relays, the optimal strategy exhibits a threshold structure, in which it is optimal to probe the dedicated relay when the signal-to-noise ratio (SNR) of the source-relay link exceeds some threshold. For the case without dedicated relays, under more restrictive conditions, the optimal strategy is also threshold-based, in the sense that it is optimal to probe potential relays when the SNR of the source-destination link lies between two thresholds. Furthermore, these strategies can be implemented in a distributed manner.

Chapter 6 studies the coverage problem of a bistatic radar (BR) sensor network, which is very challenging due to the Cassini oval sensing region of a BR and the coupling of sensing regions across different BRs. In particular, we consider the problem of deploying a network of BRs in a region to maximize the worst-case intrusion detectability, which amounts to minimizing the vulnerability of a barrier. We show that it is optimal to place BRs on the shortest barrier if it is the shortest line segment

that connects the left and right boundary of the region. Based on this, we study the optimal placement of BRs on a line segment to minimize its vulnerability, which is a non-convex optimization problem. By exploiting certain specific structural properties pertaining to the problem (particularly an important structure of detectability), we characterize the optimal placement order and the optimal placement spacing of the BR nodes, both of which present elegant balanced structures. The findings provide valuable insights into the placement of BRs for barrier coverage.

Chapter 7 studies radar sensor networks where the Doppler effect is exploited to combat the effect of clutter on radar detection. We introduce the concept of Doppler coverage as a coverage metric for such radar sensor networks. Specifically, a target is said to be Doppler-covered if, regardless of its direction of motion, there exists some radar in the network whose signal-to-noise ratio (SNR) is sufficiently high and the DFS at that radar is sufficiently large. Based on the Doppler coverage model, we devise an efficient method to characterize Doppler-covered regions for arbitrarily deployed radars. Then we design an algorithm for deriving the minimum radar density required to achieve Doppler coverage in a region under any polygonal deployment pattern, and further apply it to investigate the regular triangle based deployment.

Chapter 8 studies an extended social group utility maximization (SGUM) framework that takes into account both "positive" and "negative" social ties. As a result, this extended SGUM framework captures the rich continuum from zero-sum game (ZSG) to non-cooperative game (NCG) to network utility maximization (NUM) - traditionally disjoint paradigms for network optimization. Under this SGUM framework, we study random access control, cooperative jamming, and spectrum access as three applications. For the SGUM based random access control, we derive the socially-aware Nash equilibrium. We show that as social ties increase, each user's

SNE strategy migrates monotonically from the NE strategy for a ZSG to that for a standard NCG, and then to the social optimal strategy for NUM, and the social welfare of the SNE also increases. For the SGUM based multi-channel cooperative jamming, we show that there exists a unique mixed strategy SNE. When the social tie between the legitimate user and the cooperative jammer exceeds a certain threshold, the cooperative jammer always jams the eavesdropper on some channel at the SNE, which improves the social welfare of the legitimate user and cooperative jammer. Then we consider the SGUM based spectrum access for social networks with structural balance. We characterize the SNE and show that it is equivalent to the SNE for the game where each player consists of the users in the same social coalition based on the balanced social structure.

Chapter 9 studies mobile users' pseudonym change to protect location privacy where users' social ties are leveraged to motivate them to participate. Drawing on the social group utility maximization (SGUM) framework developed in Chapter 2, we cast users' decision making of whether to change pseudonyms as a socially-aware pseudonym change game (PCG). The PCG further assumes a general anonymity model that allows a user to have its specific anonymity set for personalized location privacy. For the SGUMbased PCG, we show that there exists a socially-aware Nash equilibrium (SNE), and we quantify the system efficiency of SNEs with respect to the optimal social welfare. Then we develop a greedy algorithm that myopically determines users' strategies, based on the social group utility derived from only the users whose strategies have already been determined. We show that this algorithm can efficiently find an SNE that enjoys desirable properties: 1) it is socially-aware coalition-proof, and thus is also Pareto-optimal; 2) it achieves a larger social welfare than any SNE for the socially-oblivious PCG. We further quantify the system efficiency of the SNE with respect to the optimal social welfare. We also show that the

SNE can be achieved in a distributed manner. Numerical results corroborate that social welfare can be significantly improved by exploiting social ties.

Chapter 10 summarizes the dissertation and discusses possible directions for future research.

Chapter 2

SOCIAL GROUP UTILITY MAXIMIZATION FRAMEWORK WITH APPLICATIONS IN RANDOM ACCESS CONTROL AND POWER CONTROL

2.1 Introduction

Mobile networks have been growing rapidly in the past few years and this trend will continue in the foreseeable future. Indeed, mobile phone shipments are projected to reach 1.9 billion units in 2014, which is about 7 times that of desktop and laptop combined [2]. The widespread popularity of mobile networks has been driven by continuing advances of technologies. On one hand, advanced wireless communication technologies (e.g., MIMO, OFDM) have drastically improved the communication efficiency in existing wireless networks (e.g., cellular networks, WLANs). On the other hand, advanced mobile devices (e.g. smartphones) equipped with powerful sensors (e.g., cameras) and high computing capability have enabled a wide range of applications on mobile platforms. As a result, mobile networks have nowadays become an indispensable infrastructure in people's everyday life.

Different from other networks (e.g., sensor networks), a distinctive characteristic of mobile networks is that mobile devices are carried and operated by human beings. As a result, mobile users' interactions hinge heavily on human behavior. It is then natural to ask "How would mobile users' *social ties* influence their behaviors in mobile networks?" Social ties are built upon human social relationships (e.g., kinship, friendship, colleague relationship). Indeed, social ties play an unprecedented role in people's interactions with each other, mainly due to the explosive growth of online social networking services (e.g., Facebook, Twitter) in the past few years. In 2013, the number of online social network users worldwide has crossed 1.73 billion, nearly one quarter of the world's population [1]. With pervasive connectivity to the Internet

via mobile devices, mobile users can interact with each other much more readily than ever before via online social networking services.

One fundamental aspect of (positive) social ties is that people are altruistic to their social "friends" (friends, family, colleagues, etc.), as they care about the welfare of their social friends. As a result, a user would take into account the effect of its behavior on its social friends. It is then natural to ask "Is it possible to exploit users' social ties to stimulate their cooperative behaviors?" Indeed, altruistic behaviors are often observed among people with social ties. With this motivation, we view a wireless network as an overlay/underlay system (as illustrated in Fig. 2.1), where a "virtual social network" (social domain) overlays a physical communication network (physical domain). Wireless users are connected by social ties in the social domain, while their wireless devices are subject to physical relationships in the physical domain. It is important to observe that users generally have *diverse* social ties such that a user cares about others at different levels. For example, a user may care about her family members more than her friends, and cares about her friends more than an acquaintance of her. Similarly, it is clear that wireless devices also generally have *diverse* physical relationships. For example, depending on their physical locations, wireless devices can cause different levels of interference to each other. A primary goal here is to leverage the intrinsic diverse social tie structure among mobile users, which can be viewed as "hidden incentives" based on existing human relationships, to facilitate cooperative networking among their mobile devices subject to diverse physical relationships.

To this end, we advocate a social group utility maximization (SGUM) framework that takes into account both the diverse social coupling and diverse physical coupling among users. Specifically, we model the social coupling and physical coupling by a social graph and physical graph, respectively, and then we cast the distributed

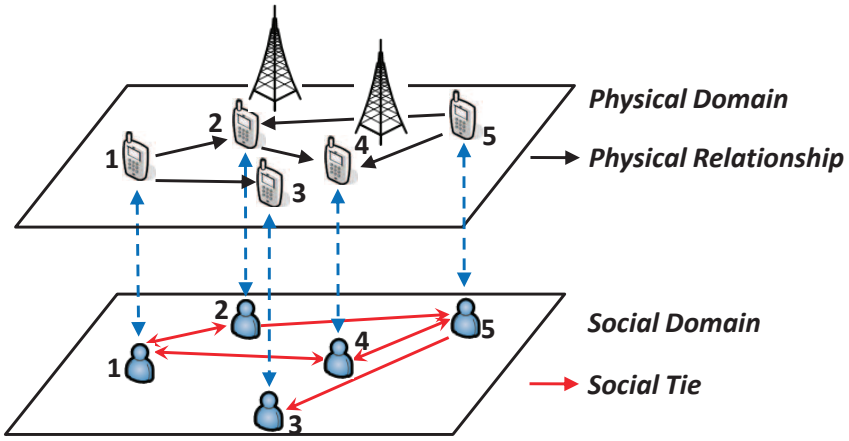


Figure 2.1: Illustration of social group utility maximization framework.

decision making problem among users as a SGUM game.

2.2 Related Work

For a network consisting of autonomous users (nodes) (e.g., ad hoc networks), each user may act in a selfish manner, in the sense that it only cares about its own benefit (e.g., utility) and does not care about the effect of its behavior on other users. In this case, the strategic interactions among users can be modeled by a *non-cooperative game* (NCG), where each user aims to maximize its payoff. NCG has been extensively studied for wireless networking applications [10]. Due to the selfish nature of users, the stable outcome of a non-cooperative game (e.g., a Nash equilibrium) may achieve a low social welfare (i.e., the total benefit of all users). In contrast to selfish users, for a network where nodes are controlled by a central authority (e.g., sensor networks), all nodes are fully cooperative and aim to achieve the same system-wide goal. In this case, a widely used objective is *network utility maximization* (NUM), which is to maximize the total utility of all nodes. NUM has been widely studied for resource allocation in wireless networks [11].

Although there exists a significant body of work on NCG and NUM, very

little attention has been paid to the continuum between these two extreme paradigms, especially in the context of mobile social networking. Recent work [12,13] have studied the impact of altruistic behavior in a routing game. [14] has recently investigated a random access game between two symmetrically altruistic players. We should note that the SGUM-based game is quite different from a *coalitional game* [15, 16], since each user in the latter aims to maximize its own welfare (although it is achieved by cooperating with other users). Furthermore, while a user in a coalitional game can only participate in *one* social group (coalition), a user in the SGUM-based game can be in *multiple* social groups associated with different users.

2.3 Social Group Utility Maximization (SGUM) Framework

2.3.1 Physical Network Graph Model

We consider a set of wireless users $\mathcal{N} = \{1, 2, ..., N\}$ where N is the total number of users. We denote the set of feasible strategies for each user $n \in \mathcal{N}$ as $\mathcal{X}_n$. For instance, a strategy $x \in \mathcal{X}_n$ can be choosing either a channel or a power level for wireless transmission. Subject to heterogeneous physical constraints, the strategy set $\mathcal{X}_n$ can be user-specific. For example, the strategy set $\mathcal{X}_n$ can be a set of feasible relay users that are in vicinity of user n for cooperative communication.

To capture the diverse physical coupling among the users in the physical domain, we introduce a *physical graph* $\mathcal{G}^p = \{\mathcal{N}, \mathcal{E}^p\}$ (see Fig. 2.1 for an example). Here the set of users $\mathcal{N}$ is the vertex set, and $\mathcal{E}^p \equiv \{(n, m) : e^p_{nm} = 1, \forall n, m \in \mathcal{N}\}$ is the edge set where $e^p_{nm} = 1$ if and only if users n and m have physical coupling (e.g., cause interference to each other). We also denote the set of users that have physical coupling with user n as $\mathcal{N}^p_n \equiv \{m \in \mathcal{N} : e^p_{nm} = 1\}$.

Let $\boldsymbol{x} = (x_1, ..., x_N) \in \prod_{n=1}^{N} \mathcal{X}_n$ be the strategy profile of all users. Given the strategy profile $\boldsymbol{x}$, the individual utility function of user n is denoted as $u_n(\boldsymbol{x})$, which represents the payoff of user n, accounting for the physical coupling among users. For

example, $u_n(\boldsymbol{x})$ can be the achieved data rate or the satisfaction of quality of service (QoS) requirement of user n under the strategy profile $\boldsymbol{x}$. Note that in general the underlying physical graph plays a critical role in determining the individual utility $u_n(\boldsymbol{x})$. For example, users' achieved data rates are determined by the interference graph and channel quality.

2.3.2 Social Network Graph Model

To capture the diverse social coupling among the users in the social domain, we introduce a *social graph* $\mathcal{G}^s = \{\mathcal{N}, \mathcal{E}^s\}$ to model their social ties. Here the edge set is given by $\mathcal{E}^s = \{(n,m) : e^s_{nm} = 1, \forall n, m \in \mathcal{N}\}$ where $e^s_{nm} = 1$ if and only if users n has a social tie with user m, which can be built on, e.g., the kinship, friendship, or colleague relationship between them. We denote the *social tie* from user n to user m as s_{nm}. We assume that each user n's social tie to itself is $s_{nn} = 1$, and we normalize user n's social tie to user $m \neq i$ as $s_{nm} \in (0, 1]$, which represents the extent to which user n cares about user m relative to user n cares about itself, with a greater value of s_{nm} indicating a stronger social tie. We also assume that $s_{nm} = 0$ if no social tie exists from user n to user m. We define user n's *social group* $\mathcal{N}^s_n$ as the set of users that have social ties with user n, i.e., $\mathcal{N}^s_n = \{m : e^s_{nm} = 1, \forall m \in \mathcal{N}\}$.

Based on the physical and social graph models described above, users are coupled in the physical domain due to the physical relationships, and are also coupled in the social domain due to their social ties. With this insight, we define the *social group utility* of each user n as

$$f_n(\boldsymbol{x}) = u_n(\boldsymbol{x}) + \sum_{m \in \mathcal{N}^s_n} s_{nm} u_m(\boldsymbol{x}). \tag{2.1}$$

It follows that the social group utility of each user consists of two parts: 1) its own individual utility and 2) the weighted sum of the individual utilities of other users having social tie with it. In a nutshell, the social group utility function captures that each user is socially-aware and cares about the users having social tie with it.

2.3.3 Social Group Utility Maximization Game

We consider the distributed decision making problem among the users for maximizing their social group utilities. Let $\boldsymbol{x}_{-n} = (x_1, ..., x_{n-1}, x_{n+1}, ..., x_N)$ be the set of strategies chosen by all other users except user n. Given the other users' strategies $\boldsymbol{x}_{-n}$, user n aims to choose a strategy $x_n \in \mathcal{X}_n$ that maximizes its social group utility, i.e.,

$$\max_{x_n \in \mathcal{X}_n} f_n(x_n, \boldsymbol{x}_{-n}), \forall n \in \mathcal{N}.$$

The distributed nature of the problem above naturally leads to a formulation based on game theory such that each user aims to maximize its social group utility. We thus formulate the decision making problem among the users as a strategic game $\Gamma = (\mathcal{N}, \{\mathcal{X}_n\}_{n\in\mathcal{N}}, \{f_n\}_{n\in\mathcal{N}})$, where the set of users $\mathcal{N}$ is the set of players, $\mathcal{X}_n$ is the set of strategies for each user n, and the social group utility function f_n of each user n is the payoff function of player n. In the sequel, we call the game Γ as the SGUM game. We next introduce the concept of *socially-aware Nash equilibrium* (SNE).

Definition 2.1. *A strategy profile $\boldsymbol{x}^* = (x_1^*, ..., x_N^*)$ is a socially-aware Nash equilibrium of the SGUM game if no player can improve its social group utility by unilaterally changing its strategy, i.e.,*

$$x_n^* = \arg\max_{x_n \in \mathcal{X}_n} f_n(x_n, \boldsymbol{x}_{-n}), \forall n \in \mathcal{N}.$$

It is worth noting that under different social graphs, the proposed SGUM game formulation can provide rich flexibility for modeling network optimization problems (as illustrated in Fig. 2.2). When the social graph consists of isolated nodes with $s_{nm} = 0$ for any $n, m \in \mathcal{N}$ (i.e., all users are socially-oblivious), the SGUM game degenerates to a standard non-cooperative game. When the social graph is fully meshed with edge weight $s_{nm} = 1$ for any $n, m \in \mathcal{N}$ (i.e., all users are fully altruistic), the SGUM game becomes a network utility maximization problem, which aims to

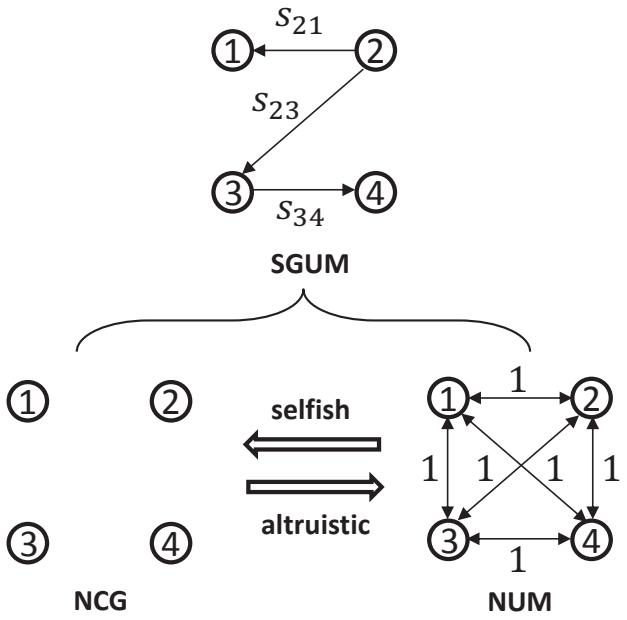


Figure 2.2: SGUM captures NCG and NUM as special cases.

maximize the system-wide utility. The SGUM framework can be applied with general social graphs and thus can bridge the gap between non-cooperative game and network utility maximization – two traditionally disjoint paradigms for network optimization (as illustrated in Fig. 2.3). These two paradigms are captured under the SGUM framework as two special cases where no social tie exists among users, and all users are connected by strongest social ties, respectively.

We note that the SGUM game is quite different from a *coalitional game* [17], since each user in the latter aims to maximize its individual benefit (although it is achieved by cooperating with other users). Furthermore, while each user in a coalitional game can only participate in *one* coalition, a user in the SGUM game can be in *multiple* social groups of different users.

To get a more concrete sense of our proposed social group utility maximization game model, in Section 2.4 and 2.5, we will study its applications in two scenarios: random access control and power control.

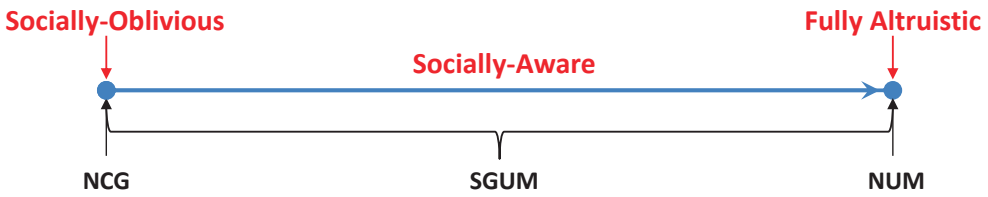


Figure 2.3: SGUM framework spans the continuum between NCG and NUM.

2.4 SGUM based Random Access Control

2.4.1 System Model

We consider a set of users under the protocol interference model, where each user i is a link consisting of transmitter T_i and receiver R_i. For example, in Fig. 2.4, T_1 interferes with R_2, T_2 interferes with R_1, T_3 interferes with R_1, where dashed circles define the interference ranges of transmitters. Let $\mathcal{I}_i^+$ denote the set of receivers that transmitter T_i causes interference to, and $\mathcal{I}_i^-$ denote the set of transmitters that causes interference to receiver R_i. In a time-slotted system, each user i contends for the opportunity of data transmission with probability $q_i \in [0, 1]$ in a time slot. If multiple interfering links contend in the same time slot, a collision occurs and no link can grab the transmission opportunity. Then the probability b_i that user i can grab the transmission opportunity is given by

$$b_i(q_i, \boldsymbol{q}_{-i}) = q_i \prod_{j \in \mathcal{I}_i^-} (1 - q_j). \tag{2.2}$$

We assume that the individual utility of user i is given by

$$u_i(q_i, \boldsymbol{q}_{-i}) = \log(\theta_i b_i) - c_i q_i \tag{2.3}$$

where $\theta_i > 0$ represents user i's efficiency of utilizing the transmission opportunity (e.g., transmission rate), and $c_i > 0$ represents user i's cost of contention. Note that the logarithmic function is widely used for modeling the utility of wireless users [18, 19]. Then, under the SGUM framework, we define the SGUM based random access

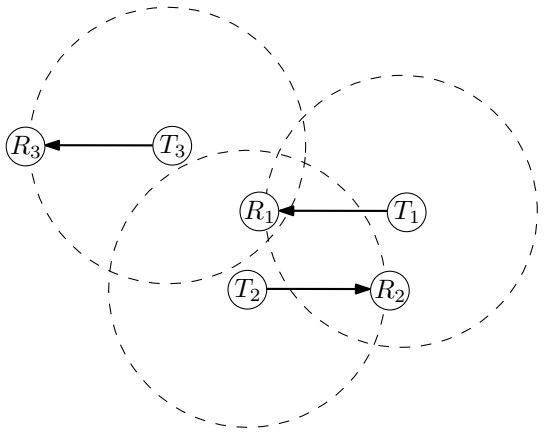


Figure 2.4: Illustration of protocol interference model.

control game as $G \triangleq (\mathcal{N}, \{q_i\}, \{f_i\})$, where

$$f_i(q_i, \boldsymbol{q}_{-i}) = \log\left(\theta_i q_i \prod_{j\in\mathcal{I}_i^-}(1-q_j)\right) - c_i q_i \sum_{j\neq i} s_{ij}\left[\log\left(\theta_j q_j \prod_{k\in\mathcal{I}_j^-}(1-p_k)\right) - c_j q_j\right]. \quad (2.4)$$

2.4.2 Game Analysis

We first have the following result.

Theorem 2.1. *For the SGUM based random access control game, there exists a unique SNE, which is*

$$q_i^{SNE} = \frac{\sum_{j\in\mathcal{I}_i^+} s_{ij} + 1 + c_i - \sqrt{(\sum_{j\in\mathcal{I}_i^+} s_{ij} + 1 + c_i)^2 - 4c_i}}{2c_i}, \forall i \in \mathcal{N}. \quad (2.5)$$

Proof: Using (2.4), setting the first-order derivative of $f_i(q_i, \boldsymbol{q}_{-i})$ to 0, we have

$$\frac{\partial f_i(q_i, \boldsymbol{q}_{-i})}{\partial q_i} = \frac{1}{q_i} - \sum_{j\in\mathcal{I}_i^+}\frac{s_{ij}}{1-q_i} - c_i = \frac{c_i q_i^2 - (\sum_{j\in\mathcal{I}_i^+} s_{ij} + 1 + c_i)q_i + 1}{q_i(1-q_i)} = 0. \quad (2.6)$$

Then we obtain the smaller root of equation (2.6) as

$$\frac{\sum_{j\in\mathcal{I}_i^+} s_{ij} + 1 + c_i - \sqrt{(\sum_{j\in\mathcal{I}_i^+} s_{ij} + 1 + c_i)^2 - 4c_i}}{2c_i} \leq \frac{1 + c_i - \sqrt{(1+c_i)^2 - 4c_i}}{2c_i} \leq 1 \tag{2.7}$$

where the first inequality follows from that the first-order derivative of the small root with respect to s_{ij} is

$$\frac{1}{2c_i}\left(1 - \frac{\sum_{j\in\mathcal{I}_i^+} s_{ij} + 1 + c_i}{\sqrt{(\sum_{j\in\mathcal{I}_i^+} s_{ij} + 1 + c_i)^2 - 4c_i}}\right) < 0. \tag{2.8}$$

We also obtain the larger root of equation (2.6) as

$$\frac{\sum_{j\in\mathcal{I}_i^+} s_{ij} + 1 + c_i + \sqrt{(\sum_{j\in\mathcal{I}_i^+} s_{ij} + 1 + c_i)^2 - 4c_i}}{2c_i} \geq \frac{1 + c_i + \sqrt{(1+c_i)^2 - 4c_i}}{2c_i} \geq 1.$$

Therefore, the SNE strategy q_i^{SNE} is unique and is the smaller root of equation (2.6). □

The result below directly follows from Theorem 2.1 and (2.8).

Corollary 2.1. *Each user's access probability at the SNE is decreasing as its social ties with others increase.*

Remark: We observe that each user's SNE strategy does not depend on other users' strategies (also known as a *dominant strategy*), but depends on the user's social ties with others. Clearly, when a user increases its access probability, it also increases the collision probabilities of the users within its interference range, and thus reduces their individual utilities. Therefore, a user would decrease its access probability when its social ties with those within its interference range get stronger (as illustrated in Fig. 2.5).

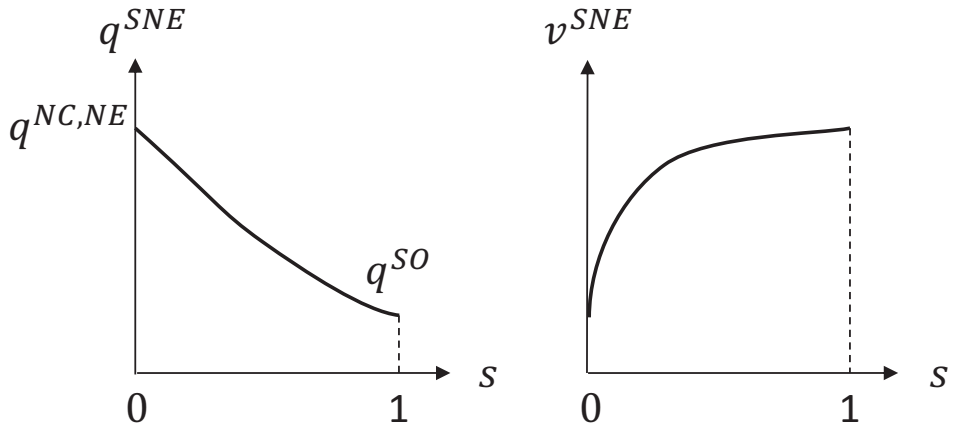


Figure 2.5: SNE for two-user SGUM based random access control.

Let $V(\boldsymbol{q})$ denote the social welfare of all users, i.e., the total individual utility of all users:

$$V(\boldsymbol{q}) \triangleq \sum_{i=1}^{N} \left[\log \left(\theta_i q_i \prod_{j \in \mathcal{I}_i^-} (1 - q_j) \right) - c_i q_i \right]. \tag{2.9}$$

Proposition 2.1. *The social welfare of the SNE is increasing as social ties increase, and reaches the social optimal point when all social ties are equal to 1.*

Proof: Using (2.9), setting the first-order derivative of $V(\boldsymbol{q})$ to 0, we have

$$\frac{\partial V(\boldsymbol{q})}{\partial q_i} = \frac{c_i q_i^2 - (|\mathcal{I}_i^+| + 1 + c_i) q_i + 1}{q_i(1 - q_i)} = 0. \tag{2.10}$$

Similar to the proof of Theorem 2.1, we obtain the social optimal strategy q_i^{SO} that maximizes $V(\boldsymbol{q})$ as the smaller root of equation (2.10), which is

$$q_i^{SO} = \frac{|\mathcal{I}_i^+| + 1 + c_i - \sqrt{(|\mathcal{I}_i^+| + 1 + c_i)^2 - 4c_i}}{2c_i}.$$

Since the larger root of equation (2.10) is

$$\frac{|\mathcal{I}_i^+| + 1 + c_i + \sqrt{(|\mathcal{I}_i^+| + 1 + c_i)^2 - 4c_i}}{2c_i} \geq \frac{1 + c_i + \sqrt{(1 + c_i)^2 - 4c_i}}{2c_i} \geq 1,$$

we have $\frac{\partial V(\boldsymbol{q})}{\partial q_i} < 0$ for $q_i \in [q_i^{SO}, 1]$, and thus $V(\boldsymbol{q})$ is decreasing in q_i when $q_i \in [q_i^{SO}, 1]$. Using Corollary 2.1, q_i^{SNE} is decreasing in s_{ij}, $\forall j \in \mathcal{I}_i^+, \forall i \in \mathcal{N}$, and hence $V(\boldsymbol{q}^{SNE})$ is increasing in s_{ij}, $\forall j \in \mathcal{I}_i^+, \forall i \in \mathcal{N}$. □

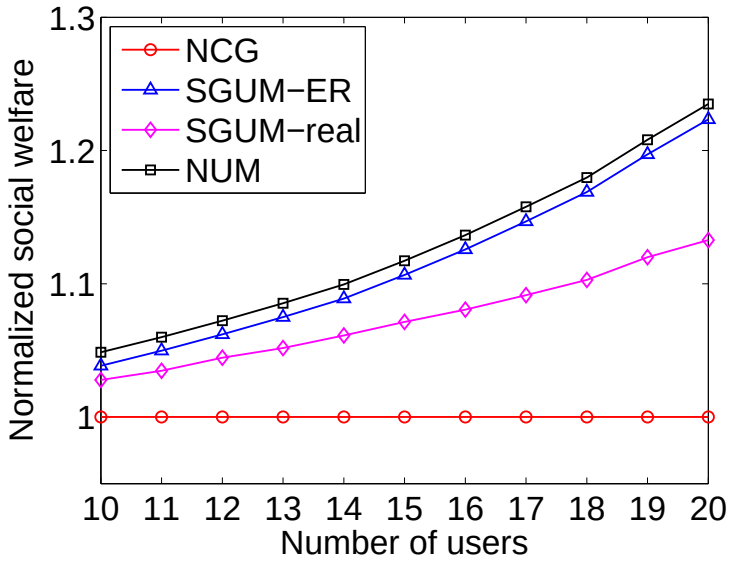


Figure 2.6: Impact of number of users.

Remark: Intuitively, since the social welfare is equal to users' individual utilities summed up with the same weight 1, a user's SNE strategy is closer to the social optimal strategy when other users weigh more in that user's social group utility (i.e., the social ties to them increase), and the social welfare increases. As social ties increase, a user's SNE strategy migrates from its NE strategy for a standard NCG to its social optimal strategy for NUM. For example, in Fig. 2.5, as the social tie $s \triangleq s_{12} = s_{21}$ increases from 0 to 1, each user's SNE strategy q^{SNE} migrates from its NE strategy $q^{NC,NE}$ for a standard NCG to its social optimal strategy q^{SO} for NUM, and the social welfare v^{SNE} of the SNE also migrates correspondingly. This demonstrates that the SGUM game framework spans the continuum between these traditionally disjoint paradigms.

2.4.3 Numerical Results

We consider N users each of which is a link consisting of a transmitter and a receiver. Each transmitter or receiver is randomly located in a square area with side length 500m. Under the protocol interference model, we assume that a link causes interfer-

ence to another link if the former link's transmitter is within 100m of the latter link's receiver. We simulate the social graph based on both the Erdos-Renyi (ER) model with link probability 0.5 and the real data trace of the friendship network Brightkite. We assume that the strength of a social tie is 1 if the social tie exists.

To illustrate the system efficiency of the SGUM solution, we compare it with the NCG solution where each user aims to maximize its individual utility, and the NUM solution where the total individual utility of all users is maximized. Fig. 2.6 depicts the social welfare of the SNE for SGUM and the social optimal solution for NUM normalized with respect to the NE for NCG, as the number of users increases. We can see that the SGUM solution for the ER model based social graph always dominates that of the NCG, with a substantial performance gain up to 22%. On the other hand, it performs almost as well as the NUM solution. This demonstrates that system efficiency can be significantly improved by exploiting social ties. We observe that the SGUM solution for the real data based social graph is worse than that for the ER model based social graph due to that social ties are weaker in the real data than in the ER graph with link probability 0.5. However, it still can achieve a performance gain up to 13% over that of the NCG solution.

2.5 SGUM based Power Control

2.5.1 System Model

We consider a set of users under the physical interference model, where each user i is a link consisting of a transmitter T_i and a receiver R_i. The channel gain of communication link i is h_i, and the channel gain of the interference link between transmitter T_i and receiver R_j is g_{ij} (as illustrated in Fig. 2.7). The noise at receiver R_i is n_i. Then the signal to interference and noise ratio (SINR) γ_i of link i is given by

$$\gamma_i(p_i, \boldsymbol{p}_{-i}) = \frac{h_i p_i}{n_i + \sum_{j=1}^{N} g_{ji} p_j}$$

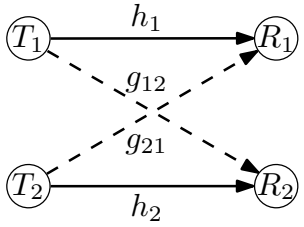


Figure 2.7: Illustration of physical interference model.

where p_i denotes the transmit power of T_i. We assume that the individual utility u_i of player i is given by

$$u_i(p_i, \boldsymbol{p}_{-i}) = \log(\gamma_j) - c_i p_i$$

where c_i denotes the cost of per unit power consumption. Similar to Section 2.4, we also use the logarithmic function to model the utility of a user. For example, $\log(\gamma_i)$ can be a good approximation for the channel capacity $\log(1+\gamma_i)$ under the high SINR regime. Also, $\log(\gamma_i)$ can be used to quantify the satisfaction of wireless users' requirements in terms of SINR. Then, under the SGUM framework, we define the SGUM game for power control as $G \triangleq (\mathcal{N}, \{p_i\}, \{f_i\})$, where

$$f_i(p_i, \boldsymbol{p}_{-i}) = \log\left(\frac{h_i p_i}{n_i + \sum_{j\neq i} g_{ji} p_j}\right) - c_i p_i + \sum_{k\neq i} s_{ik}\left(\log\left(\frac{h_k p_k}{n_k + \sum_{j\neq k} g_{jk} p_j}\right) - c_k p_k\right). \quad (2.11)$$

2.5.2 Game Analysis

We first have the following result.

Theorem 2.2. *The SGUM based power control game is a supermodular game, and thus there exists at least one SNE.*

Proof: Using (2.11), we have

$$\frac{\partial f_i(p_i, \boldsymbol{p}_{-i})}{\partial p_i} = \frac{1}{p_i} - \sum_{k\neq i} \frac{s_{ik} g_{ik}}{n_k + \sum_{j\neq k} g_{jk} p_j} - c_i.$$

Since each term in the above summation term is decreasing in p_j, $\forall j \in \mathcal{N} \setminus i$, it follows that

$$\frac{\partial^2 f_i(p_i, \boldsymbol{p}_{-i})}{\partial p_i \partial p_j} > 0, \forall j \in \mathcal{N} \setminus i$$

which implies that the social group utility function $f_i(p_i, \boldsymbol{p}_{-i})$ is supermodular. It follows from [20] that there exists at least one NE. □

Since the SGUM based power control game is a supermodular game, it follows from [21] that users can start from any strategies (e.g., $\boldsymbol{p} = (0, \cdots, 0)$) and use asynchronous *best response* updates such that their strategies will monotonically converge to a SNE.

For ease of exposition, in the rest of this section we will focus on the SGUM based power control game with *two* users, because the two-user case can shed light on the impact of social ties on users' strategies and social welfare. Furthermore, in general, the game with more than two users does not yield closed-form SNE strategies, and hence is much more difficult to quantify the impact.

Theorem 2.3. *For the two-user SGUM based power control game, there exists a unique SNE, which is*

$$p_1^{SNE} = \sqrt{\alpha_1^2 + \beta_1} - \alpha_1, \ p_2^{SNE} = \sqrt{\alpha_2^2 + \beta_2} - \alpha_2$$

where

$$\alpha_1 \equiv \frac{s_{12}g_{12} + c_1 n_2 - g_{12}}{2c_1 g_{12}}, \ \beta_1 \equiv \frac{n_2}{c_1 g_{12}}$$

and

$$\alpha_2 \equiv \frac{s_{21}g_{21} + c_2 n_1 - g_{21}}{2c_2 g_{21}}, \ \beta_2 \equiv \frac{n_1}{c_2 g_{21}}.$$

Proof: Since

$$u_1(p_1, p_2) = \log\left(\frac{h_1 p_1}{n_1 + g_{21}p_2}\right) - c_1 p_1 + s_{12}\log\left(\frac{h_2 p_2}{n_2 + g_{12}p_1}\right) - s_{12}c_2 p_2,$$

we have

$$\frac{\partial u_1(p_1,p_2)}{\partial p_1}=\frac{1}{p_1}-\frac{s_{12}g_{12}}{n_2+g_{12}p_1}-c_1.$$

Since

$$\lim_{p_1\to 0}(\frac{1}{p_1}-\frac{s_{12}g_{12}}{n_2+g_{12}p_1})\geq\lim_{p_1\to 0}(\frac{1}{p_1}-\frac{s_{12}}{p_1})=\infty$$

and

$$\lim_{p_1\to\infty}(\frac{1}{p_1}-\frac{s_{12}g_{12}}{n_2+g_{12}p_1})=0$$

and

$$\frac{\partial\left(\frac{1}{p_1}-\frac{s_{12}g_{12}}{n_2+g_{12}p_1}\right)}{\partial p_1}=-\frac{1}{p_1^2}+\frac{s_{12}g_{12}^2}{(n_2+g_{12}p_1)^2}=\frac{(s_{12}-1)g_{12}^2p_1^2-2n_2g_{12}p_1-n_2^2}{p_1^2(n_2+g_{12}p_1)^2}<0,$$

there exists a unique value of p_1 such that

$$\frac{1}{p_1}-\frac{s_{12}g_{12}}{n_2+g_{12}p_1}-c_1=0, \tag{2.12}$$

which is also the value of p_1^{SNE}. Solving (2.12), we obtain the desired result. Similarly, we can obtain p_2^{SNE}. □

Next we have the following result.

Corollary 2.2. *For the two-user SGUM based power control game, each user's transmit power at the SNE is decreasing as its social tie with the other increases.*

Proof: Since

$$p_1^{SNE}=\sqrt{\alpha_1^2+\beta_1}-\alpha_1$$

and

$$\alpha_1\equiv\frac{s_{12}g_{12}+c_1n_2-g_{12}}{2c_1g_{12}},\ \beta_1\equiv\frac{n_2}{c_1g_{12}}>0,$$

we have

$$\frac{\partial p_1^{SNE}}{\partial s_{12}}=\frac{\partial\left(\sqrt{\alpha_1^2+\beta_1}-\alpha_1\right)}{\partial\alpha_1}\frac{\partial\alpha_1}{\partial s_{12}}=\left(\frac{\alpha_1}{\sqrt{\alpha_1^2+\beta_1}}-1\right)\frac{1}{2c_1}<0.$$

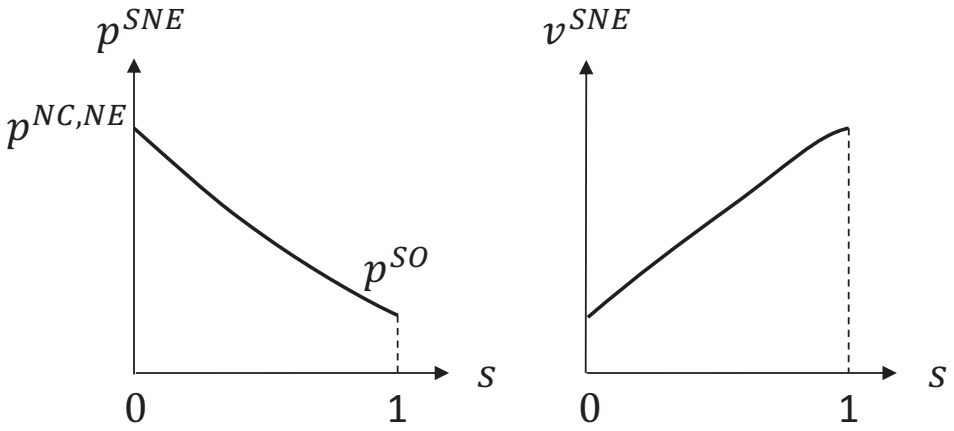


Figure 2.8: SNE for two-user SGUM based power control.

So p_1^{SNE} is decreasing in s_{12}. Similarly, we can show that p_2^{SNE} is decreasing in s_{21}. □

Proposition 2.2. *For the two-user SGUM based power control game, the social welfare of the SNE is increasing as social ties increase, and reaches the social optimal point when all social ties are equal to 1.*

Proof: Since

$$V(p_1, p_2) = \log\left(\frac{h_1 p_1}{n_1 + g_{21} p_2}\right) - c_1 p_1 + \log\left(\frac{h_2 p_2}{n_2 + g_{12} p_1}\right) - c_2 p_2$$

we have

$$\frac{\partial V(p_1, p_2)}{\partial p_1} = \frac{1}{p_1} - \frac{g_{12}}{n_2 + g_{12} p_1} - c_1.$$

Using the same argument as in the proof of Theorem 2.3, the optimal value p_1^{SO} of p_1 for $V(p_1, p_2)$ is the unique solution of

$$\frac{1}{p_1} - \frac{g_{12}}{n_2 + g_{12} p_1} - c_1 = 0.$$

In particular, we have $p_1^{SNE} \geq p_1^{SO}$. Since $\frac{\partial V(p_1, p_2)}{\partial p_1} < 0$ when $p_1 \geq p_1^{SO}$, $V(p_1, p_2)$ is decreasing in p_1 when $p_1 \geq p_1^{SO}$. Using Lemma 2.2, p_1^{SNE} is decreasing in s_{12}, and

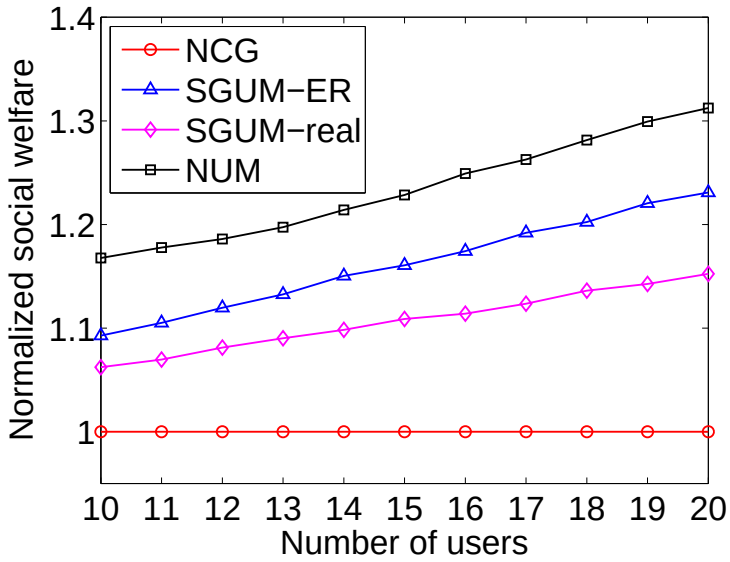


Figure 2.9: Impact of number of users.

hence $V(p_1^{SNE}, p_2^{SNE})$ is increasing in s_{12} since p_2^{SNE} is independent of s_{12}. Similarly, we can show that $V(p_1^{SNE}, p_2^{SNE})$ is increasing in s_{21}. □

Remark: Similar to the SGUM based random access control game, for the two-user SGUM based power control game, each user's strategy at the SNE is also a dominant strategy. As a user's social tie with the other increases, the user's transmit power at the SNE decreases, and the social welfare increases. Therefore, as the social tie increases, a user's SNE strategy migrates from its NE strategy for a standard NCG to its social optimal strategy for NUM. For example, in Fig. 2.8, as the social tie $s \triangleq s_{12} = s_{21}$ increases from 0 to 1, each user's SNE strategy p^{SNE} migrates from its NE strategy $p^{NC,NE}$ for a standard NCG to its social optimal strategy p^{SO} for NUM, and the social welfare v^{SNE} of the SNE also migrates correspondingly.

2.5.3 Numerical Results

We consider N users each of which is a link consisting of a transmitter and a receiver. Each transmitter or receiver is randomly located in a square area with side length 500m. Under the physical interference model, we assume that the channel condition

of a link (communication or interference link) only depends on the path loss effect with path loss factor 3. We assume that the transmit power of each link is 1W and the noise power at each receiver is 0.1W.

Fig. 2.9 shows the normalized social welfare for a varying number of users. We can see that the SGUM solution for the ER model based social graph can achieve a performance gain up to 23% over the NCG solution, and its performance loss from the NUM solution is at most 10%. The SGUM solution for the real data based social graph can achieve a performance gain up to 15%.

2.6 Conclusion

In this chapter, we have developed a general social group utility maximization game framework that bridges the gap between non-cooperative game and network utility maximization. In particular, we have studied two applications in mobile social networks under this framework: random access control and power control. Our findings provide useful insight into the impact of social ties on users' strategies and network efficiency. We believe that this work will open a new door to exploring the impact of social behavior on mobile networking.

Chapter 3

EXPLOITING SOCIAL TRUST ASSISTED RECIPROCITY (STAR) TOWARDS UTILITY-OPTIMAL CROWDSENSING

3.1 Introduction

Mobile crowdsensing has recently emerged as a promising paradigm for a variety of applications, thanks to the pervasive penetration of mobile devices to people's daily lives. Indeed, with the development of 4G networks and powerful processors, smartphone sales crossed 1 billion units in 2013 [2]. As smartphones are equipped with advanced sensors such as accelerometer, compass, gyroscope, and camera, they can collectively carry out many sensing tasks, e.g., monitoring the environment. In a nutshell, by leveraging a crowd of mobile users, one can collect and process sensed data far beyond the scope of what was possible before.

Although the benefit of crowdsensing is pronounced, performing a sensing task typically incurs *overhead* for the participating user, in terms of the user's resource consumption devoted to sensing, such as battery and computing power. Further, the participating user also incurs the risk of potential privacy loss by sharing its sensed data with others. In general, a user may not participate in sensing without receiving adequate incentive. Therefore, effective *incentive design* is essential for realizing the benefit of crowdsensing.

There have been some recent studies on incentive design for crowdsensing (see, e.g., [22–24]). Most of these work use monetary reward to stimulate users' participation, which rely on a global (virtual) currency system. However, enforcing the circulation of a global currency typically incurs a high implementation overhead, especially for large-scale networks, due to the need to, e.g., resolve disputes and punish counterfeiters. Therefore, it is appealing to design a crowdsensing system that can

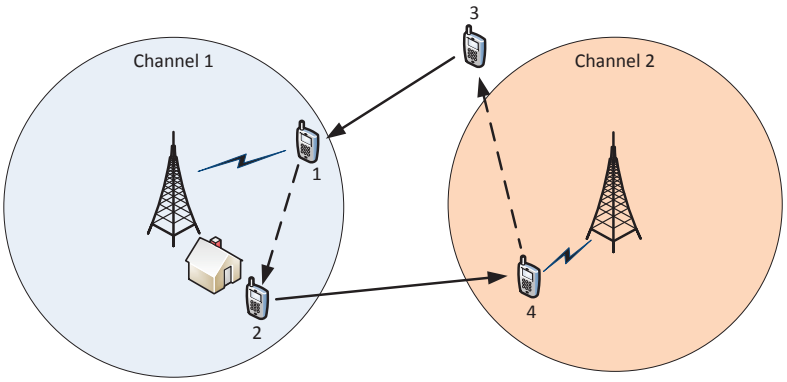


Figure 3.1: An example of social trust assisted reciprocity for spectrum crowdsensing.

motivate a large number of users to participate without using a global currency, which is a goal of this study.

Online social networks have been explosively growing over the past few years. In 2013, the number of online social network users worldwide reached 1.73 billion, nearly one quarter of the world's population [1]. As a result, social relationships increasingly influence people's behaviors in their interactions. In particular, as an important aspect of social relationships, *social trust* can be exploited to stimulate crowdsensing: if Alice has social trust in Bob, then Alice is willing to help Bob, since Alice can trust Bob in that Bob would help Alice in the future to return the favor.

In this chapter, we devise an incentive mechanism to stimulate users' participation in crowdsensing, by using *Social Trust Assisted Reciprocity* (STAR) - a synergistic marriage of social trust and reciprocity. The basic idea of STAR is that Alice is willing to help Bob if someone who trusts Bob can help someone trusted by Alice. This is because the overhead of Alice for helping Bob is compensated, as the one trusted by Alice will help Alice in the future to return the favor. We illustrate this idea by an example of spectrum crowdsensing in Fig. 3.1. Users 1 and 4 have social trust in user 3 and 2, respectively (denoted by solid edges); user 2 and 3 request user 1 and 4 to sense channel 1 and 2, respectively (denoted by dashed edges). User 1

is willing to help user 2 in exchange for that user 4 helps user 1's social friend, user 3, while user 4 is willing to help user 3 in exchange for that user 1 helps user 4's social friend, user 2.

By taking advantage of reciprocity ("synchronous exchange") with the assistance of social trust ("asynchronous exchange"), STAR can efficiently encourage users' participation in crowdsensing. In particular, STAR greatly enhances the chance that sensing requests are matched, since they can be matched through existing social trust among users. As illustrated in Fig. 3.1, without using either social trust or reciprocity, neither of CR1 and CR3 would be willing to help CR2 and CR4, respectively. If users are well connected in the social network, the number of requests that can be matched with the assistance of social trust can be significant. Furthermore, compared to traditional currency-based schemes, STAR can incur much lower implementation overhead due to the use of the already existing social trust. We will discuss the overhead of STAR and related work in Section 3.6.

The main thrust of this study is devoted to characterizing the fundamental performance of STAR, particularly for satisfying users' sensing requests given the social trust structure among them. Since sensing requests are mismatched in general and social trust levels are limited, it may not be possible to satisfy all requests. Therefore, a natural question is *"Can all requests be satisfied?"* The benefit of sensing service provided under STAR can be quantified by the utility of users who receive the service. In the case that not all requests can be satisfied, another important question arises: *What is the maximum utility that can be achieved by the provided service?* These two questions are similar in spirit to admission control and network utility maximization, respectively.

We summarize the main contributions of this chapter as follows.

- We propose a socially-aware crowdsensing system that stimulates users' participation by leveraging their social trust. The incurred overhead can be significantly lower than that of traditional currency-based schemes, since it uses *social credit* as a "local" currency enabled by social trust, rather than a global currency.

- For the proposed system, we design STAR, an incentive mechanism which stimulates users' participation by using *social trust assisted reciprocity*. We investigate thoroughly the efficacy of STAR for satisfying users' sensing requests, given the social trust structure among users. Specifically, we first show that all requests can be satisfied if and only if users who request more sensing service than they can provide can transfer sufficient social credit to users who can provide more than they request. Then we investigate utility maximization for sensing service, and show that this problem is equivalent to maximizing the utility of a *circulation flow* in the combined social graph and request graph. Based on this observation, we develop an algorithm that iteratively cancels the cycles of positive weights in the *residual graph*, and hence computes the optimal solution efficiently, for both cases of divisible and indivisible service.

- We evaluate the performance of STAR through extensive simulations for a random setting based on the Erdős-Rényi graph model, and for a practical setting based on real social data with application to spectrum crowdsensing. For both settings, simulation results demonstrate that STAR can achieve significantly better system efficiency and individual user performance than only using social trust or reciprocity.

The rest of this chapter is organized as follows. In Section 3.2, we propose a socially-aware crowdsensing system. In Section 3.3, we design an incentive mechanism

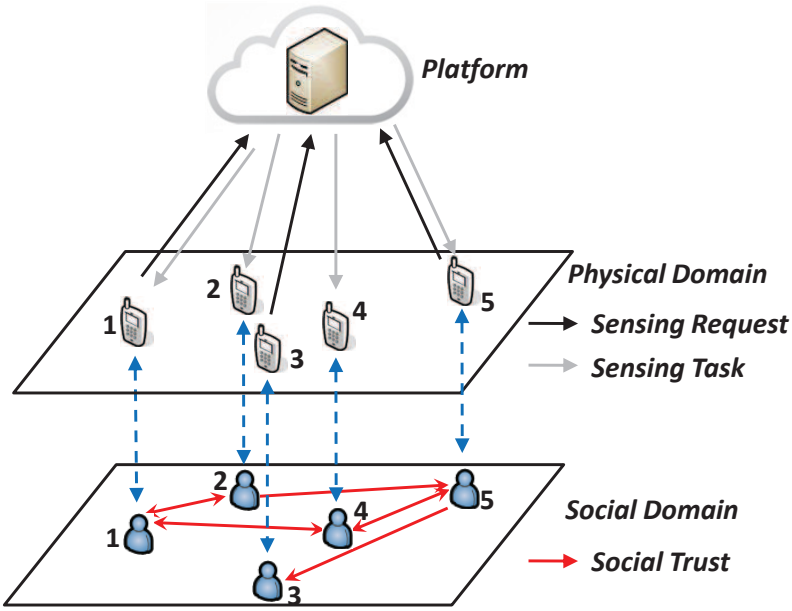


Figure 3.2: Illustration of socially-aware crowdsensing system.

based on social trust assisted reciprocity for the proposed system. Based on STAR, Section 3.4 investigates conditions for satisfying all sensing requests and the utility maximization for sensing service. Section 3.5 provides simulation results to illustrate the efficacy of STAR. Related work is reviewed in Section 3.6 and the chapter is concluded in Section 3.7.

3.2 Socially-aware Crowdsensing System

In this section, we describe a crowdsensing system that stimulates users' participation by leveraging their social trust.

3.2.1 Motivation

Social relationships play an increasingly important role in people's interactions with each other. One important aspect of the social relationship between two users is their *social trust*: one user has belief in and relies on the other user's behavior in the future. To stimulate users' participation in crowdsensing, social trust can be exploited in the form of *social credit*. Specifically, social credit is transferred between

two users with social trust if one user owes a favor to the other and commits to return the favor later. Therefore, a user is willing to participate if it receives social credit from another user that it has social trust in. This "asynchronous exchange" of favors via social credit is in the same spirit as a global currency. However, since this pairwise credit commitment is enabled by existing social trust between two users, social credit would incur a much lower overhead than a global currency.

Most existing crowdsensing systems assume that a platform announces sensing tasks and motivates users to participate in these tasks by providing monetary incentives [22, 24]. In contrast, we are interested in a system where sensing requests are generated by users. Indeed, a few crowdsensing systems that have been deployed are based on this model. For example, the Waze [25] system employs traffic monitoring data collected from a crowd of drivers to answer an individual driver's request (e.g., navigation to a specific destination).

3.2.2 System Description

We consider a crowdsensing system as illustrated in Fig. 3.2. The system consists of a *platform* that operates in the cloud, and a set of mobile *users* $V = \{1, \cdots, N\}$ connected to the platform via the cloud. Initially, each user registers at the platform and publishes its social information (e.g., Facebook account) such that users can identify their social relationships with each other. Then, each user declares to the platform a social credit *limit* for each other user that it has social trust in, based on the strength of their social relationship. The social credit limit quantifies the social trust level by specifying how much social credit one user is willing to accept from another. For example, a user typically has high social trust in a close friend, while it may have low social trust in an acquaintance. The system proceeds in rounds and the workflow in each round consists of four major components as depicted in Fig. 3.3. We describe each component in detail below.

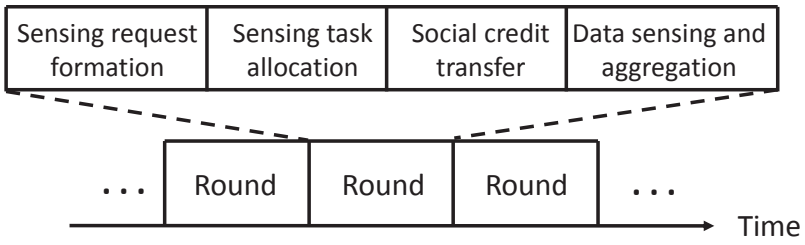


Figure 3.3: Workflow of socially-aware crowdsensing system.

- **Sensing request formation.** A user can submit to the platform a *sensing request* that describes the sensing service it needs. For example, a user may request to know if a licensed channel is available. Upon receiving a request, the platform can find a particular set of users who can serve the request, based on users' sensing capabilities, such as their physical locations and the functions of sensors on their devices. In this way, the platform determines the request relationships among users, i.e., which user requests service from which other users. For example, a user with a good sensing channel condition for a licensed channel can serve another user's request to sense that channel.

- **Sensing task allocation.** Based on the sensing requests, the platform allocates sensing tasks to users. A sensing task specifies how much sensing service is needed from that user. For example, a sensing task may require a user to sense a licensed channel for a period of time. A key challenge for the platform is to ensure that users have incentive to carry out their allocated tasks.

  As expected, a user who requests sensing service can also receive requests from others for service. Therefore, it is plausible to take advantage of *direct (bilateral)* or *indirect (multi-lateral) reciprocity* (as illustrated in Fig. 3.4(a),(b)): Alice is willing to help Bob if Bob simultaneously helps Alice. While this "synchronous exchange" of favors is appealing as it obviates the need for currency, a major drawback is that users' requests have to be simultaneously matched, which does

not hold in general. As illustrated by the example in Fig. 3.1, user 1 has a good sensing channel for channel 1 while user 2 does not. Therefore, user 2 needs user 1's help while user 1 does not need user 2's help.

- **Social credit transfer.** The platform can stimulate users' participation by using social credit. The platform maintains the social credit limit for each pair of users, and updates it for the next round to reflect the amount of social credit transferred between them in the current round. Besides the update performed by the platform, each user can also change its credit limit for another by reporting the new value to the platform. Section 3.4.3 will discuss credit limit setting in more detail.

- **Data sensing and aggregation.** Based on the transferred social credit, users have incentive to carry out their allocated sensing tasks. After collecting and processing the sensing data from users, the platform distributes the aggregated data to the intended users. Therefore, most of the communication and computation burdens shifts from the users to the platform.

Unlike most existing work, our proposed crowdsensing system exploits social trust to stimulate users' participation, which obviates the need of a global currency. For this system, one key challenge is to make the best use of social credit such that users have incentive to carry out sensing tasks, and more importantly, the system achieves good performance, which is the focus in the rest of this chapter.

3.3 STAR: Social Trust Assisted Reciprocity Based Incentive Mechanism

In this section, we design an incentive mechanism based on social trust assisted reciprocity.

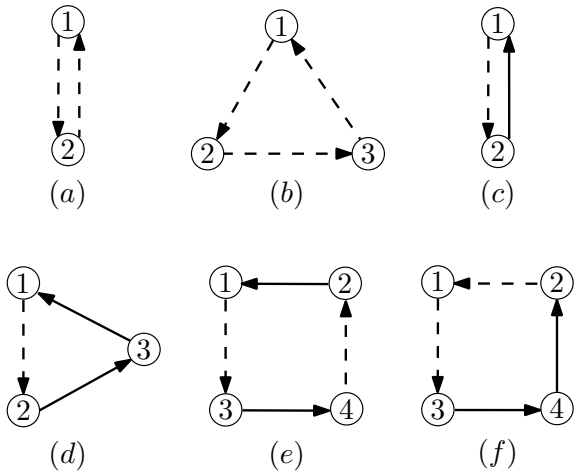


Figure 3.4: Examples of social trust assisted reciprocity cycles.

3.3.1 System Model

We model users' sensing requests by a *request graph* $G^R \triangleq (V, E^R)$, in which user i and user j are connected by a directed *request edge* $e_{ij}^R \in E^R$ if user j requests sensing service[1] from user i. The capacity $R_{ij} > 0$ of each request edge e_{ij}^R represents the *amount* of service requested by user j from user i[2]. The flow $f_{ij}^R > 0$ on the request edge e_{ij}^R represents the amount of service provided by user i to user j. Depending on the specific application, sensing service can be *divisible* (e.g., quantified by sensing time) such that R_{ij} and f_{ij}^R for each $e_{ij}^R \in E^R$ have continuous values, or *indivisible* (e.g., quantified by the number of sensing data samples) such that they have to be integers. In some situations, a user cannot provide all the service requested from it (e.g., due to its resource constraints). To take this into account, let C_i be the maximum amount of service that user i can provide and E_i^C be the set of outgoing request edges of user i. Then the following constraint applies:

$$\sum_{j:e_{ij}\in E_{i^+}^R} f_{ij}^R \leq C_i. \tag{3.1}$$

[1]For brevity, we use "sensing service" and "service" interchangeably throughout the chapter.

[2]Recall that users' request relationships are determined by the platform in the sensing request formation phase as described in Section 3.2.2.

We will discuss how to capture constraint (3.1) in our incentive mechanism in Section 3.4.

A user obtains utility from its requested sensing service, which depends on the amount of service provided by each user who is requested for that service. We assume that user j obtains a utility of value U_{ij} for each *unit* amount of service provided by user i. In general, user j can request different types of service from user i, which have utilities of different values. For example, a user may request sensing multiple channels from another user, whose sensing capability varies across different channels. In this case, there are multiple *parallel* request edges (in the same direction) from user i to user j, each with a specific utility of service[3]. In this chapter, we assume that there exists *at most one* request edge from one user to another. However, all the results obtained under the assumption can be directly extended to the case of parallel request edges. We further assume that a user's utility is equal to the total utility of the service provided to that user. More complex forms of utility will be studied in our future work.

We model the social trust structure among users by a *social graph* $G^S \triangleq (V, E^S)$, in which user i and user j are connected by a directed *social edge* $e_{ij}^S \in E^S$ if user j has social trust in user i. The capacity $S_{ij} > 0$ of each social edge e_{ij}^S represents the social credit limit, which specifies the maximum amount of social credit that can be transferred from user i to user j. The flow f_{ij}^S on the social edge e_{ij}^S represents the amount of social credit transferred between user i and user j. The social credit unit is the same as the sensing service unit, and is the same for all users. Note that $f_{ij}^S = -f_{ji}^S$ holds for each pair of social edges between two users, where $f_{ij}^S > 0$ (or $f_{ji}^S > 0$) indicates that a credit of f_{ij}^S (or f_{ji}^S, respectively) is transferred from user i to user j (or from user j to user i, respectively).

[3]For brevity, we say "utility of service" instead of "utility per unit service".

3.3.2 An Example of Spectrum Crowdsensing

As an illustrative example, we next discuss how the system model described above can be applied to spectrum crowdsensing.

Spectrum sensing is an important and challenging task in cognitive radio networks [26]. To access a licensed channel in a cognitive radio network, a user needs to sense the channel to ensure that the channel is not used by licensed transmitters. When a user's sensing channel condition is impaired by severe fading (e.g., path loss, shadowing), the user needs other users' help to sense the channel. Consider a cognitive radio network where each user intends to sense one or multiple licensed channels for access. A user's sensing capability for a channel depends on its sensing channel condition, which can vary across different users and different channels. If user i has a good sensing channel condition for a channel, user j may request user i to sense that channel. The overhead incurred by sensing that channel can be user i's resource consumption (e.g., device battery) for the sensing task. Therefore, the amount of sensing service f_{ij}^R provided by user i to user j can be quantified by user i's sensing time. The utility $U_{ij,k}$ of user j derived from the service provided by user i on channel k, can depend on user i's sensing capability on channel k as well as user j's utilization efficiency (e.g., transmission rate) of channel k.

3.3.3 Design Description

The basic incentive structure of the STAR mechanism is a *social trust assisted reciprocity cycle* (STAR) in which a set of users have incentive to provide service. It is defined in the combined social and request (social-request) graph $G \triangleq (V, E^S \cup E^R)$. Fig. 3.5 gives an example of the social-request graph in (c) constructed from the social graph in (a) and the request graph in (b).

Definition 3.1. *A social trust assisted reciprocity cycle is a directed cycle in the*

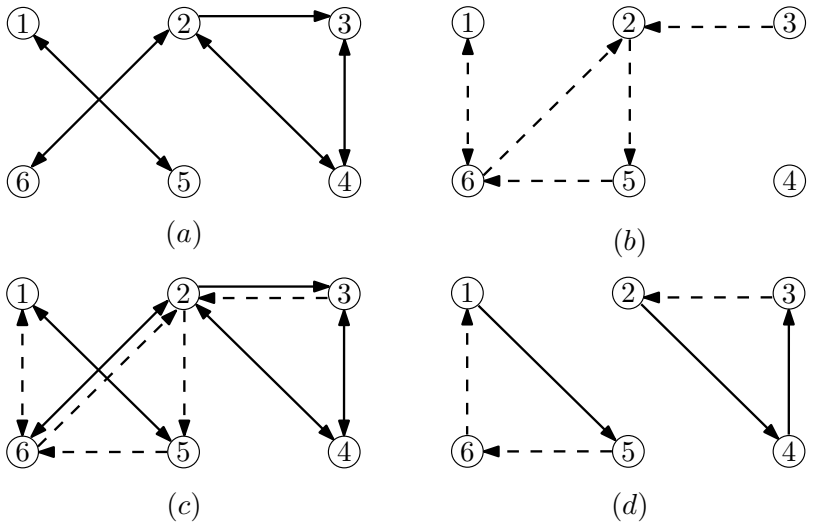


Figure 3.5: An example of combining social graph and request graph.

social-request graph G.

Fig. 3.5(d) illustrates two STAR cycles for the social-request graph in Fig. 3.5(c). Fig. 3.4 also gives some examples of social trust assisted reciprocity cycles. In particular, (a)-(d) are special cases: (a) is direct reciprocity cycle; (b) is indirect reciprocity cycle; (c) is direct social trust based cycle; (d) is indirect social trust based cycle. In a STAR cycle, a user is willing to provide service since *the overhead is compensated by receiving credit or service from another user in that cycle*. For example, user 1 in Fig. 3.4(e) is willing to provide service to user 3 since it receives credit from user 2; user 1 in Fig. 3.4(f) is willing to provide service to user 3 since it receives service from user 2. Note that a STAR cycle can involve *intermediate* users that only transfer credits with their social neighbors. For example, in Fig. 3.4(f), user 4 is an intermediate user.

For each user in a STAR cycle, the amount of service or credit it receives should be *equal* to that of service or credit it provides or spends, respectively. Let f_c denote a *balanced* flow along a STAR cycle c, which has the same flow value on

each edge in c. The flow on a social or request edge in the *aggregate* flow f of a set of balanced flows $\{f_c, c \in \mathcal{C}\}$ along cycles $\mathcal{C}$ is given by

$$f_{ij}^S = \sum_{c \in \mathcal{C}: e_{ij}^S \in c} f_c - \sum_{c \in \mathcal{C}: e_{ji}^S \in c} f_c, \quad f_{ij}^R = \sum_{c \in \mathcal{C}: e_{ij}^R \in c} f_c$$

respectively. Note that the credit transferred from user i to j (i.e., the flow on $e_{ij}^S \in E^S$) in the balanced flow along a STAR cycle can be partly or completely *canceled* by that from user j to i in another STAR cycle. Users can participate in a set of balanced flows along STAR cycles if and only if the aggregate flow satisfies the capacity constraints on request and social edges.

Definition 3.2. *A set of balanced flows along STAR cycles is feasible if the aggregate flow satisfies the following capacity constraints:*

$$-S_{ji} \leq f_{ij}^S \leq S_{ij},\ f_{ji}^S = -f_{ij}^S,\ \forall e_{ij} \in E^S \tag{3.2}$$

$$0 \leq f_{ij}^R \leq R_{ij},\ \forall e_{ij} \in E^R. \tag{3.3}$$

Recall that the amount of service a user can provide can be constrained (i.e., constraint (3.1)). To capture this constraint, we can modify the social-request graph G as follows. We first construct a virtual node $i' \in V$ and change all the outgoing request edges from node i to being from node i', and then we add a virtual edge $e_{ii'}^R \in E^R$ and set its capacity and utility as C_i and 0, respectively. For example, in Fig. 3.6, the social-request graph can be modified to capture the constraint $f_{12}^R + f_{13}^R \leq C_1$. Note that all other edges keep unchanged. It can be easily shown that it suffices to focus on the modified graph: any feasible set of balanced flows along STAR cycles in the modified graph has a *one-to-one correspondence* in the original graph that also satisfies constraint (3.1).

Under the STAR mechanism, all users are willing to participate in any feasible set of balanced flows along STAR cycles.

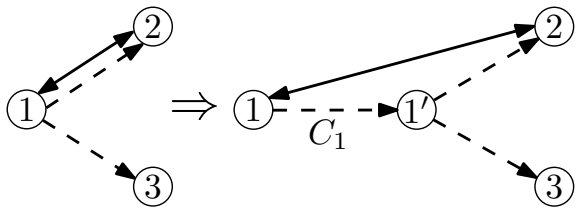


Figure 3.6: An example that shows how to capture constraint (3.1).

3.4 Exploiting STAR to Satisfy Sensing Requests

In this section, we characterize the efficacy of the STAR mechanism. We first investigate conditions under which all sensing requests can be satisfied. Then we study the maximum total utility that can be achieved by provided sensing service.

3.4.1 Satisfying All Sensing Requests

Based on the STAR cycles, we first show that it suffices to focus on *circulation* flows in the social-request graph defined as follows.

Definition 3.3. *A flow f in the social-request graph G is a circulation if f satisfies the capacity constraints* (3.2), (3.3), *and the flow conservation constraints*

$$\sum_{j:e_{ij}\in E^R} f_{ij}^R + \sum_{j:e_{ij}\in E^S} f_{ij}^S = \sum_{j:e_{ji}\in E^R} f_{ji}^R, \ \forall i \in V. \tag{3.4}$$

It is clear that the aggregate flow of any feasible set of balanced flows along STAR cycles is a circulation flow in G. The following lemma shows that the converse is also true.

Lemma 3.1. *Any circulation flow in the social-request graph amounts to the aggregate flow of a feasible set of balanced flows along STAR cycles.*

Proof: Consider a non-empty circulation flow f. We can find a node v_1 with a positive flow on an outgoing edge from v_1 and trace along this edge to another node

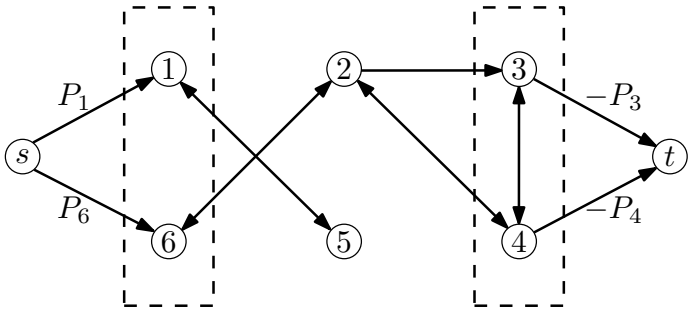


Figure 3.7: An example of extended social graph.

v_2. Due to the flow conservation constraint, we can find an outgoing edge from v_2 with a positive flow and trace along it to a node v_3. We continue this tracing process until we visit a node v_j that has been visited before, i.e., $v_i = v_j$ for some $i < j$, and hence we find a STAR cycle $v_i \to v_{i+1} \to \cdots \to v_j$. Then we subtract flow f by a balanced flow along this cycle with value equal to the minimum flow value on an edge in that cycle. Thus the remaining flow is still a circulation flow in which the number of edges with non-zero flows is reduced. We can repeat this argument to subtract the remaining flow by a balanced flow along a cycle until it is empty. This implies that flow f is the aggregate flow of the subtracted balanced flows along the cycles, which is also feasible. □

We define P_i as the total amount of service requested by user i deducted by the amount that user i can provide:

$$P_i \triangleq \sum_{j:e_{ji}\in E^R} R_{ji} - \sum_{j:e_{ij}\in E^R} R_{ij}.$$

Then we construct an extended social graph G^{S^+} from the social graph G^S by adding a directed edge with capacity P_i from a virtual source node s to each node i with $P_i > 0$, and adding a directed edge with capacity $-P_i$ from each node i with $P_i < 0$ to a virtual destination node t. Fig. 3.7 illustrates the extended social graph constructed from the social-request graph in Fig. 3.5(d) where $P_1 > 0$, $P_2 = 0$, $P_3 < 0$, $P_4 < 0$,

$P_5 = 0$, $P_6 > 0$. Let P be defined as

$$P \triangleq \sum_{i:P_i>0} P_i = -\sum_{i:P_i<0} P_i.$$

Theorem 3.1. *All sensing requests can be satisfied under STAR if and only if P is equal to the maximum flow value from s to t in the extended social graph G^{S^+}.*

Proof: By Lemma 3.1, all requests can be satisfied if and only if there is a circulation flow f in the social-request graph G that saturates all request edges (i.e., $f^R_{ij} = R_{ij}$, $\forall e^R_{ij} \in E^R$).

We first show the "if" part. Suppose S is equal to the value of the maximum flow f^{S^+} from s to t in G^{S^+}. Let f^S be the flow comprised of the flows on the social edges E^S in f^{S^+} (i.e., not including the edges from s and to t in G^{S^+}). Let f^R be the flow in the request graph G^R that saturates all request edges. Then we augment flow f^S in the social-request graph G with flow f^R to obtain a flow f in G. According to the construction of G^{S^+}, we have $\sum_{j:e^S_{ij}\in E^S} f^S_{ij} = P_i$ for each node $i \in V$, while we also have $\sum_{j:e^R_{ji}\in E^R} f^R_{ji} - \sum_{j:e^R_{ij}\in E^R} f^R_{ij} = P_i$. This shows that f is a circulation flow.

Next we show the "only if" part. Suppose f is a circulation flow in G that saturates all request edges. Let f^S be the flow comprised of the flows on the social edges E^S in f. Then we augment flow f^S with saturated flows on the edges from s and to t in G^{S^+} to obtain a flow f^{S^+} in G^{S^+}. According to the construction of G^{S^+}, f^{S^+} is a flow in G^{S^+} satisfying the capacity and flow conservation constraints, with a flow value of P from s to t. □

Remark: Theorem 3.1 provides a useful insight: all requests can be satisfied if and only if *users who request more service than they can provide can transfer sufficient social credit to users who can provide more than they request, to compensate their imbalance in requests*. Intuitively speaking, the social graph serves as a "buffer" to partially or completely "absorb" the mismatch among users' requests. It is worth

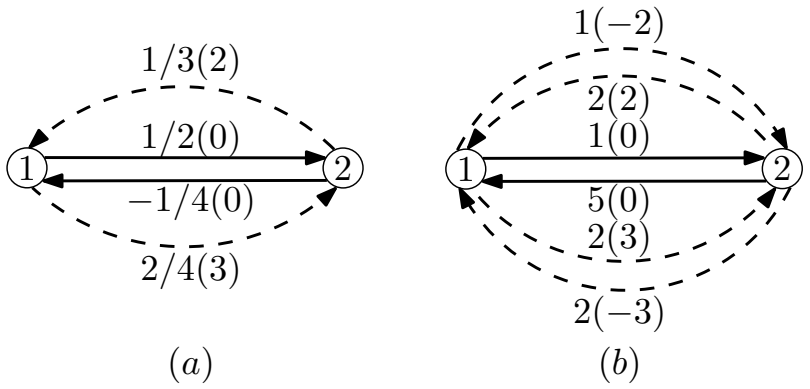


Figure 3.8: An example of social-request graph and the residual graph.

noting that the maximum amount of service provided under STAR is in general *not* equal to the maximum flow value from s to t in G^{S^+}.

Remark: We note that an important difference between [27] and our study is that the results in [27] is based on the assumption that *all users are connected in the social network*, whereas our model here does not have this assumption. This is essentially because that reciprocity is used in STAR but not in [27]. As illustrated in Fig. 3.1, without using reciprocity, user 1 and 4 are not willing to help user 2 and 3, respectively. In Section 3.5, simulation results demonstrate that the STAR mechanism can significantly outperform the mechanism in [27].

3.4.2 Utility Maximization for Sensing Service

Due to the mismatch of sensing service requests and social credit limits, it is possible that not all requests can be satisfied. In this case, a natural objective from the platform's view is to maximize the total utility of provided service. The next result follows from Lemma 3.1.

Theorem 3.2. *The maximum utility of sensing service provided under STAR is equal to the maximum utility of a circulation flow in the social-request graph.*

Note that the flow on a social edge does not generate any utility. By Theo-

rem 3.2, our problem can be written as

$$\begin{aligned} \underset{f_{ij}^S, f_{ij}^R}{\text{maximize}} \quad & \sum_{i,j:e_{ij}\in E^R} U_{ij} f_{ij}^R \\ \text{subject to} \quad & \text{constraints } (3.2), (3.3), (3.4). \end{aligned} \tag{3.5}$$

Note that we can *maximize the total amount of service provided under STAR by solving problem* (3.5) *with the utility* U_{ij} *set to 1 for each request edge.*

In the following, we will solve problem (3.5) using an algorithm inspired by the *cycle-canceling* algorithm for solving the minimum cost flow problem [28]. We should note that problem (3.5) is quite different from a typical network flow problem in that two nodes can be connected by multiple edges (request edges and social edges). Furthermore, *request edges and social edges carry different types of flows* (as illustrated in Fig. 3.8(a)): the flows on all request edges are *non-negative* and *independent* (as in constraint (3.3)), while the flows on social edges can be *negative* and must be *inverse* between a pair of users (as in constraint (3.2)).

We start with constructing a *residual graph* $G_f \triangleq (V, E_f^S \cup E_f^R)$ of the social-request graph G for a given flow f. Specifically, for each request edge $e_{ij}^R \in E^R$, we construct a pair of *forward* edge $\overrightarrow{e}_{ij}^R \in E_f^R$ and *backward* edge $\overleftarrow{e}_{ji}^R \in E_f^R$ with capacities

$$\overrightarrow{R}_{ij} = R_{ij} - f_{ij}^R, \ \overleftarrow{R}_{ij} = f_{ij}^R$$

respectively. For each *pair* of social edges $e_{ij}^S, e_{ji}^S \in E^S$, we construct a pair of edges $\overrightarrow{e}_{ij}^S, \overrightarrow{e}_{ji}^S \in E_f^S$ with capacities

$$\overrightarrow{S}_{ij} = S_{ij} - f_{ij}^S, \ \overrightarrow{S}_{ji} = S_{ji} - f_{ji}^S$$

respectively. We do *not* construct an edge in the residual graph if its capacity is zero. Then we set the *weights* of each forward edge $\overrightarrow{e}_{ij}^R \in E_f^R$ and each backward edge

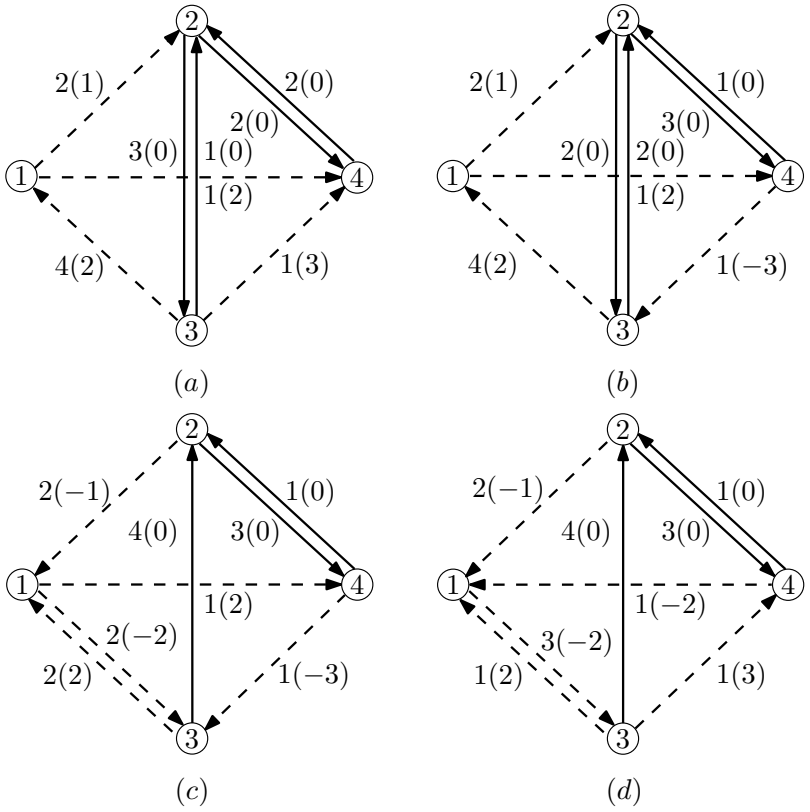


Figure 3.9: Illustration of Algorithm 1.

$\overleftarrow{e}^R_{ji} \in E^R_f$ as

$$\overrightarrow{W}^R_{ij} = U_{ij},\ \overleftarrow{W}^R_{ij} = -U_{ij}$$

respectively. The weights of each pair of edges $\overrightarrow{e}^S_{ij}, \overrightarrow{e}^S_{ji} \in E^S_f$ are set to

$$\overrightarrow{W}^S_{ij} = \overrightarrow{W}^S_{ji} = 0.$$

We show how to construct the residual graph by an illustrative example in Fig. 3.8. In particular, for each edge, the number before / is the flow value; the number before () is the capacity; the number in () is the weight. The following lemma establishes the optimality condition for solving problem (3.5).

Lemma 3.2. *A flow f is optimal for problem* (3.5) *if and only if there exists no cycle of positive weight in the residual graph G_f.*

Algorithm 1: Find the optimal flow for problem (3.5) in social-request graph G

input : Social-request graph G
output: The optimal flow for problem (3.5)

```
Initialize an empty flow f in G;
while There exists a cycle of positive weight in the residual graph G_f of flow f do
    Find a cycle c of positive weight in G_f;
    Compute the residual capacity r_c of cycle c;
    Augment flow f with a balanced flow of value r_c along cycle c;
end
return Flow f;
```

Proof: The "only if" part is easy to show: If there exists a cycle of positive weight in G_f, then we can augment the flow f with a balanced flow of value $\epsilon > 0$ along that cycle to construct a circulation flow with larger utility.

Next we show the "if" part. Suppose there exists no cycle of positive weight in G_f but there exists a circulation flow f' in G with larger utility than f. Similar to the residual graph G_f, we construct a graph $\overline{G} \triangleq (V, \overline{E}^S \cup \overline{E}^R)$ from G by constructing $\overrightarrow{e}^R_{ij}, \overleftarrow{e}^R_{ij} \in \overline{E}^R$ for each $e^R_{ij} \in E^R$ and $\overrightarrow{e}^S_{ij}, \overrightarrow{e}^S_{ji} \in \overline{E}^S$ for each pair of $e^S_{ij}, e^S_{ji} \in E^S$, and setting their weights the same as those in G_f. The difference between G_f and $\overline{G}$ is that all the edges are constructed in $\overline{G}$ (an edge is not constructed in G_f if its capacity is 0) and have unlimited capacities. Therefore, the edges in G_f is a subset of the edges in $\overline{G}$. Then we can define a flow g in $\overline{G}$ by defining the flows in g on the edges of $\overline{G}$ as

$$\overrightarrow{g}^R_{ij} = \max\{0, f'^R_{ij} - f^R_{ij}\},\ \forall \overrightarrow{e}^R_{ij} \in \overline{E}^R$$

$$\overleftarrow{g}^R_{ij} = \max\{0, f^R_{ij} - f'^R_{ij}\},\ \forall \overleftarrow{e}^R_{ij} \in \overline{E}^R$$

$$\overrightarrow{g}^S_{ij} = f'^S_{ij} - f^S_{ij},\ \forall \overrightarrow{e}^S_{ij} \in \overline{E}^S.$$

It follows from the definition that

$$\overrightarrow{g}^R_{ij} - \overleftarrow{g}^R_{ij} = f'^R_{ij} - f^R_{ij},\ \forall e^R_{ij} \in E^R.$$

Then the net flow value at each node $i \in V$ in flow g is

$$\begin{aligned}
&\sum_{j:\overrightarrow{e}^R_{ij}\in\overline{E}^R} \overrightarrow{g}^R_{ij} + \sum_{j:\overleftarrow{e}^R_{ji}\in\overline{E}^R} \overleftarrow{g}^R_{ji} + \sum_{j:\overrightarrow{e}^S_{ij}\in\overline{E}^R} \overrightarrow{g}^S_{ij} - \sum_{j:\overleftarrow{e}^R_{ij}\in\overline{E}^R} \overleftarrow{g}^R_{ij} - \sum_{j:\overrightarrow{e}^R_{ji}\in\overline{E}^R} \overrightarrow{g}^R_{ji} \\
&= \sum_{j:e^R_{ij}\in E^R} \left(f'^R_{ij} - f^R_{ij}\right) - \sum_{j:e^R_{ji}\in E^R} \left(f'^R_{ji} - f^R_{ji}\right) + \sum_{j:e^S_{ij}\in E^S} \left(f'^S_{ij} - f^S_{ij}\right) \\
&= \left(\sum_{j:e^R_{ij}\in E^R} f'^R_{ij} + \sum_{j:e^S_{ij}\in E^S} f'^S_{ij} - \sum_{j:e^R_{ji}\in E^R} f'^R_{ji} \right) - \left(\sum_{j:e^R_{ij}\in E^R} f^R_{ij} + \sum_{j:e^S_{ij}\in E^S} f^S_{ij} - \sum_{j:e^R_{ji}\in E^R} f^R_{ji} \right) = 0
\end{aligned}$$

where the last equality follows from that f' and f are circulation flows in G. Therefore, g is a circulation flow in $\overline{G}$. We observe that the flow on any edge $e \in \overline{E}^R \setminus E^R_f$ is zero in g because 1) if $e = \overrightarrow{e}^R_{ij}$, then we have $f^R_{ij} = R_{ij}$ and hence $\overrightarrow{g}^R_{ij} = 0$; 2) if $e = \overleftarrow{e}^R_{ij}$, then we have $f^R_{ij} = 0$ and hence $\overleftarrow{g}^R_{ij} = 0$. We further observe that $\overrightarrow{g}^S_{ij} \leq 0$ for any edge $\overrightarrow{e}^S_{ij} \in \overline{E}^S \setminus E^S_f$ since we have $f^S_{ij} = S_{ij}$. Since $\overrightarrow{W}^S_{ij} = 0, \forall \overrightarrow{e}^S_{ij} \in \overline{E}^S$, the weight of flow g in $\overline{G}$ is

$$\sum_{i,j:\overrightarrow{e}^R_{ij}\in\overline{E}^R} \left(\overrightarrow{W}^R_{ij}\overrightarrow{g}^R_{ij} + \overleftarrow{W}^R_{ij}\overleftarrow{g}^R_{ij}\right) = \sum_{i,j:e^R_{ij}\in E^R} U_{ij}\left(f'^R_{ij} - f^R_{ij}\right) = \sum_{i,j:e^R_{ij}\in E^R} U_{ij} f'^R_{ij} - \sum_{i,j:e^R_{ij}\in E^R} U_{ij} f^R_{ij} > 0$$

where the last inequality follows from the assumption that f' has larger utility than f in G. Since g only has positive flows on the edges in G_f, using a similar argument as in the proof of Lemma 3.1, g is the aggregate flow of balanced flows along cycles each comprised of edges in G_f. Then the total weight of these flows along the cycles in G_f is equal to the weight of flow g in $\overline{G}$, which is greater than 0. This implies that there must exist a cycle of positive weight in G_f, which is a contradiction to the previous assumption. This completes the proof. □

Using Lemma 3.2, we can develop an algorithm as described in Algorithm 1 to solve problem (3.5). The algorithm starts with the empty flow in the network. It iteratively finds a cycle of *positive weight* in the residual graph and cancels each cycle by augmenting the current flow in the graph with a balanced flow along that cycle, until no cycle of positive weight exists. In each iteration, the value of the flow

to augment with is set to be the *residual capacity* of the cycle, which is the minimum capacity of all edges in that cycle. We show how Algorithm 1 works by an illustrative example in Fig. 3.9. In particular, for each edge, the number before () is the capacity; the number in () is the weight. Fig. 3.9(a) shows the initial social-request graph with the empty flow; Fig. 3.9(b) shows the residual graph after augmenting with a flow of value 1 along cycle $2 \to 3 \to 4 \to 2$; Fig. 3.9(c) shows the residual graph after augmenting with a flow of value 2 along cycle $1 \to 2 \to 3 \to 1$; Fig. 3.9(d) shows the residual graph after augmenting with a flow of value 1 along cycle $1 \to 4 \to 3 \to 1$.

As for the step 2 in Algorithm 1, we can use an algorithm similar to the Bellman-Ford algorithm [29] to find a cycle of positive weight in the residual graph, if there exists one. In particular, the algorithm iteratively updates the maximum weight $M(t)$ from a source node $s \in V$ to each other node $t \in V \setminus \{s\}$. In each iteration, the algorithm checks each edge $e_{ij}^S \in E_f^S$ or $e_{ij}^R \in E_f^R$ once, and increases the maximum weight $M(j)$ to $M(i) + \overrightarrow{W}_{ij}^S$ if $M(i) + \overrightarrow{W}_{ij}^S > M(j)$. The algorithm runs for $|V| - 1$ iterations. When it terminates, if $M(t)$ for some $t \in V \setminus \{s\}$ can be further reduced by checking some edge, then there exists a cycle of positive weight in the graph. The algorithm has running time $O(|V|(|E^S| + |E^R|)$.

For ease of exposition, we will focus on problem (3.5) with *rational* parameters: the utilities and capacities of all social and request edges are rational numbers. This setting is of important interest in general, since the parameters of most practical problems are rational numbers. Then problem (3.5) with rational parameters can be equivalently converted to one with *integral* parameters by multiplying with a suitably large integer[4] K. The solution of the original problem (with rational parameters) is equal to the solution of the new problem (with integral parameters) divided by K.

[4]For example, it can be the least common multiple of the denominators in the fractional forms of the rational numbers.

For problem (3.5) with rational parameters, let $\overline{U}$ and $\overline{R}$ denote the maximum utility and maximum capacity of a request edge, respectively (i.e., $\overline{U} = \max_{e_{ij} \in E^R} U_{ij}$, $\overline{R} = \max_{e_{ij} \in E^R} R_{ij}$). The following theorem shows that Algorithm 1 is correct and efficient when the service divisible.

Theorem 3.3. *For problem* (3.5) *with divisible sensing service and rational parameters, Algorithm 1 computes the optimal flow and has running time* $O(|V||E^R|(|E^R| + |E^S|)\overline{RU}K^2)$.

Proof: As discussed earlier, we first equivalently convert the problem to one with integral parameters by multiplying them by an integer K.

Since the capacities of all edges in the graph are integral and the initial empty flow is integral, the residual capacity of the cycle found in the first iteration of the algorithm is integral, and hence the flow after augmentation is integral. Thus, by induction, the updated flow after each iteration is also integral. This shows that the algorithm finds an integral flow when it terminates, which is optimal by Lemma 3.2.

The utility of the initial empty flow is 0. The utility of any flow is upper bounded by the utility of the flow that saturates all request edges, which is $|E^R|\overline{RU}K^2$. Since the capacities of all edges are integral, the flow utility increases by an integer no less than one at each iteration of Algorithm 1. Therefore, it takes the algorithm at most $|E^R|\overline{RU}K^2$ iterations to terminate. Since each iteration has running time $O(|V|(|E^S| + |E^R|)$, the desired result follows. □

In many practical situations, sensing service is indivisible such that the optimization variables of problem (3.5) have to be integers. In this case, we can equivalently convert problem (3.5) with rational parameters to one with integer parameters by rounding the capacities of all social and request edges to their respective nearest integers below (i.e., taking the floor function) and multiplying the utilities of all request

edges by a suitably large integer K. Using a similar proof as that of Theorem 3.3, we have the following result.

Theorem 3.4. *For problem* (3.5) *with indivisible sensing service and rational parameters, Algorithm 1 computes the optimal flow and has running time* $O(|V||E^R|(|E^R|+|E^S|)\lfloor \overline{R} \rfloor \overline{U} K)$.

In Section 3.5, simulation results demonstrate that the running time of Algorithm 1 is much lower than the above bound.

Remark: The underlying rational of Algorithm 1 can be interpreted as follows. In each iteration of Algorithm 1, the flow in the social-request graph is augmented with a balanced flow along a cycle of positive weight in the residual graph. For each edge with positive weight in that cycle, the utility of flow on the corresponding request edge increases, while for each edge with negative weight in that cycle, that utility decreases. Since the total weight of the edges in the cycle is positive, the total utility of flow increases. In other words, a balanced flow along the positive weight cycle captures *the tradeoff between increasing the utilities on some request edges and decreasing the utilities on some other request edges such that the total utility increases*. Note that although the flows on social edges does not generate utility, the edges with zero weights in the residual graph, which are constructed from the social edges, contribute to forming a cycle, and hence the utility obtained on request edges.

Remark: It is worth noting that, when sensing service is indivisible, problem (3.5) is essentially an integer linear program (ILP), which is NP-hard to solve in general. However, using a network flow approach, we can capture and exploit the specific combinatorial structure of the problem, based on which a polynomial-time algorithm can be developed to solve it.

3.4.3 Further Discussions

It is a widely observed phenomenon in real social networks that people tend to be similar to their friends in a variety of ways (e.g., age, gender, affiliation), which is known as the principle of homophily [30]. As a result, social friends are also likely to be similar in physical locations (e.g., colleagues in the same workplace). Therefore, when a user's request can only be served by a physical neighbor, it is likely to be served by a social friend nearby under the STAR mechanism. However, even when the user has no social friends nearby, it can still be served by any physical neighbor if they are both in a STAR cycle (as illustrated in Fig. 3.4). Furthermore, depending on the specific context of a sensing service, a user's request is not necessarily served by its physical neighbors. For example, a user can request to know the traffic condition at a remote location that it will visit later, or the availability of a channel used by a licensed transmitter at a distant location. In other words, if two users are connected in the request graph, they are not necessarily close physically. Therefore, the effectiveness of the STAR mechanism does not rely on significant overlapping of a user's social friends and physical neighbors.

Although each social credit limit is set by a user individually, users can negotiate the credit limits for each other. After the initial declaration, credit limits are updated by the platform in each round according to the credit transfers among users in that round. In addition, users can also change the credit limits based on the dynamics of their social trust. For example, users can increase the credit limits to each other if their social trust improves.

3.5 Performance Evaluation

In this section, we provide simulation results to evaluate the performance of the STAR mechanism. We compare STAR with two incentive mechanisms as benchmarks, which

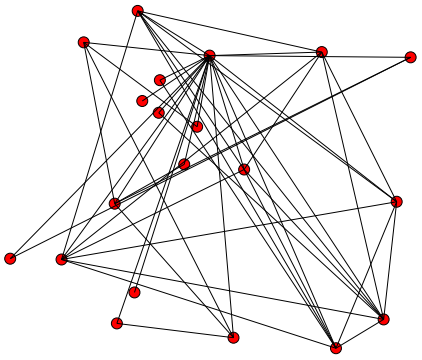


Figure 3.10: Social network structure in real dataset.

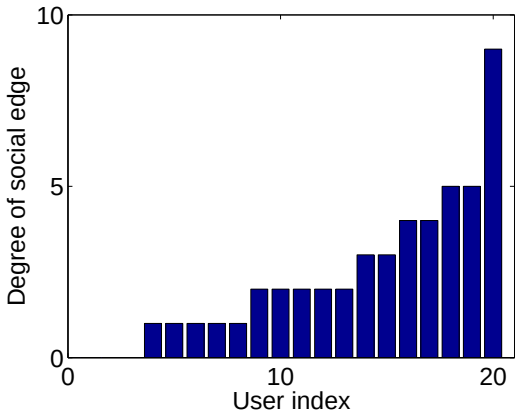


Figure 3.11: Degree of social edge in real dataset.

use social trust only and reciprocity only, respectively:

- *Social trust based mechanism* (ST): Under this mechanism, a user is willing to provide service to another if and only if it receives social credit from that user or an intermediate user. Therefore, in the social-request graph, this mechanism can use a cycle consisting of social edges and *exactly one* request edge (e.g., as illustrated in Fig. 3.4(c),(d));

- *Reciprocity based mechanism* (RP): Under this mechanism, a user is willing to provide service if and only if it also receives service from another. Therefore, in the social-request graph, this mechanism can use a cycle consisting of *only* request edges (e.g., as illustrated in Fig. 3.4(a),(b)).

We observe that each benchmark mechanism only uses a *subset* of the incentive structures (i.e., the cycles in the social-request graph) used in the STAR mechanism. Note that the incentive mechanism only using social trust is equivalent to that in [27].

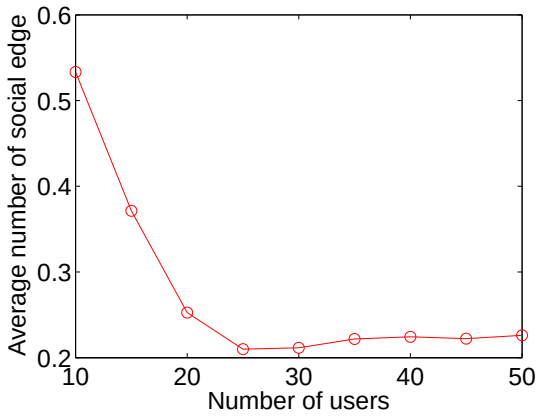


Figure 3.12: Average number of social edge in real dataset.

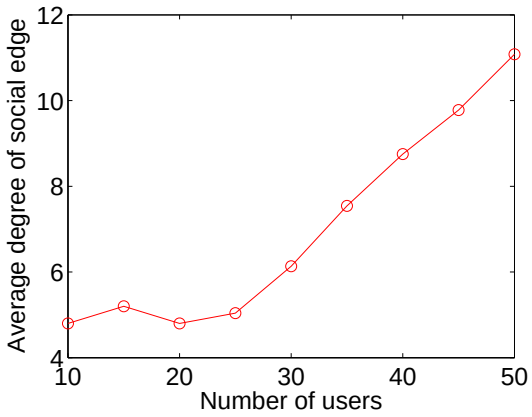


Figure 3.13: Average degree of social edge in real dataset.

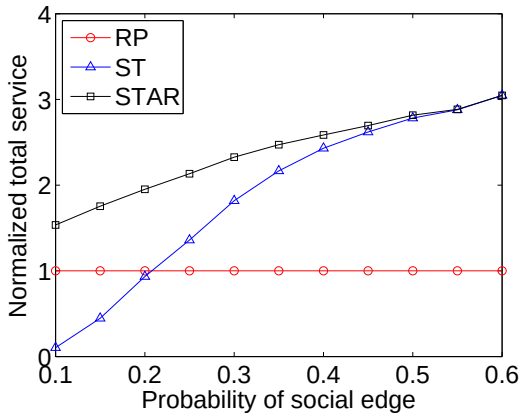


Figure 3.14: Impact of P_S for random setting.

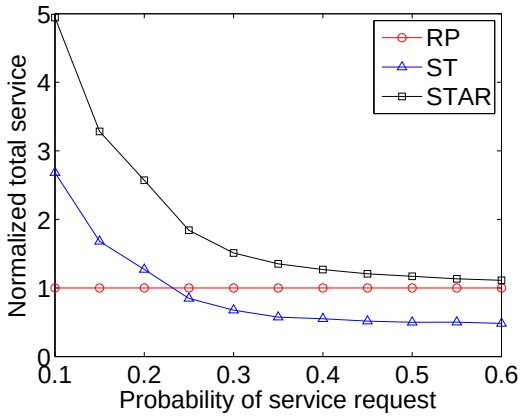


Figure 3.15: Impact of P_R for random setting.

3.5.1 Simulation Setup

To illustrate the impact of different parameters of the mobile social network on the performance, we consider a random setting as follows. We simulate the social graph G^S and the request graph G^R using the Erdős-Rényi (ER) graph model [31], where a social edge and a request edge exist from one node to another with probability P_S and P_R, respectively. We assume that service is divisible. If a social edge exists, its social credit limit follows a normal distribution $N(\mu_S, \sigma_S^2)$, where μ_S and σ_S^2 denote

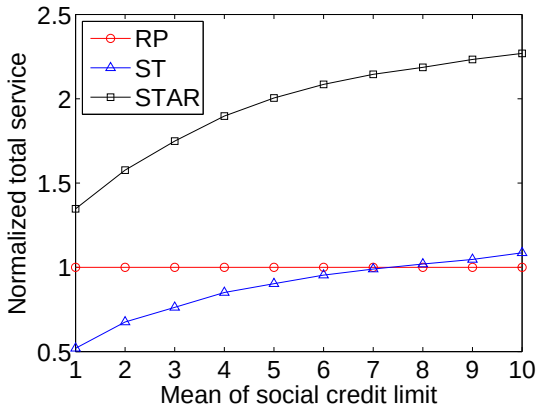


Figure 3.16: Impact of μ_S for random setting.

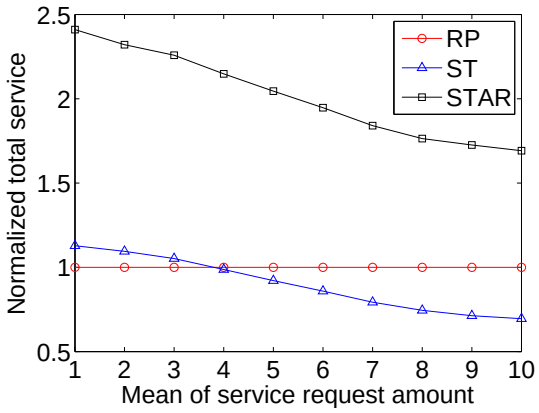


Figure 3.17: Impact of μ_R for random setting

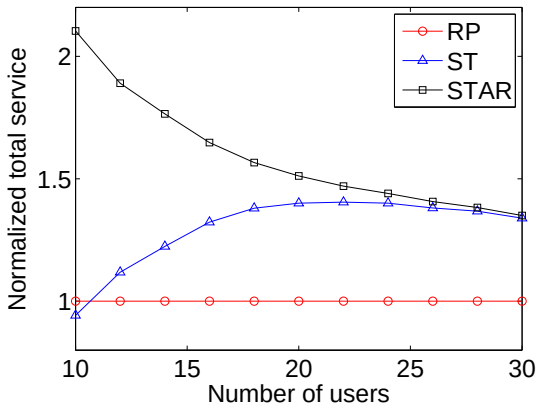


Figure 3.18: Impact of N for random setting.

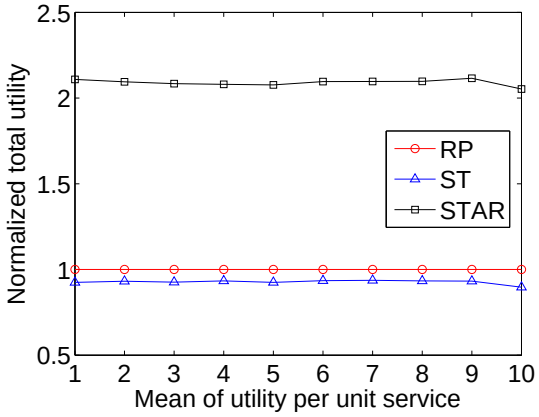


Figure 3.19: Impact of μ_U for random setting.

the mean and variance, respectively; if a physical edge exists, the amount of requested service and the utility per unit service follows a normal distribution $N(\mu_R, \sigma_R^2)$ and $N(\mu_U, \sigma_U^2)$ respectively. We set default parameter values as: $N = 10$, $P_S = 0.2$, $P_R = 0.2$, $\mu_S = \mu_R = 5$, $\sigma_S^2 = \sigma_R^2 = 1$, $\mu_U = 10$, $\sigma_U^2 = 2$.

To evaluate the performance of the STAR mechanism in practice, we also consider a practical setting. Specifically, we generate the social graph according to the real dataset from Brightkite [32]. Brightkite is a online social networking service based on mobile phones where users share their checking-in locations in an explicit

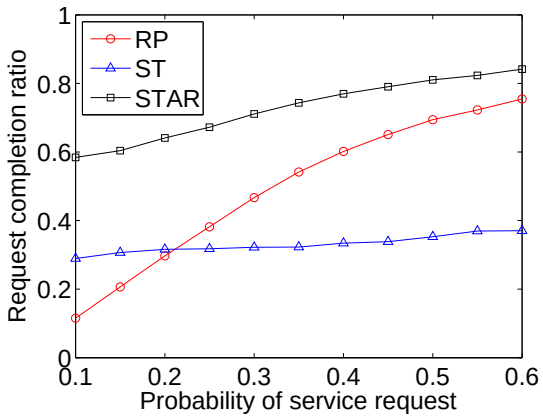


Figure 3.20: Impact of P_R for random setting.

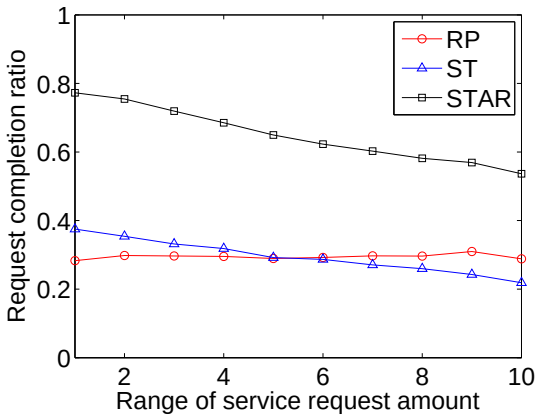


Figure 3.21: Impact of μ_R for random setting.

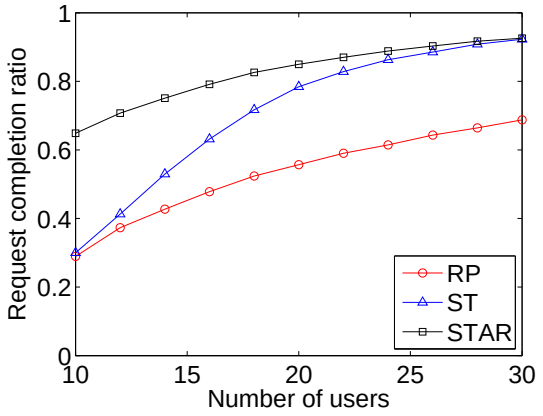


Figure 3.22: Impact of N for random setting.

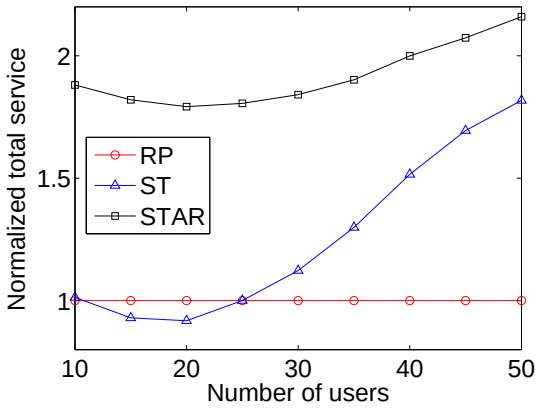


Figure 3.23: Impact of N for practical setting.

social network. For this dataset, we illustrate the social network structure of 20 users in Fig. 3.10 and the users' degree of social edge in Fig. 3.11. We also plot the average number of social edge between a pair of users (in analogy to the probability of social edge in the ER model) versus the number of users in Fig. 3.12, and plot the average degree of social edge in Fig. 3.13. We simulate the request graph based on the context of spectrum crowdsensing discussed in Section 3.3.2. We consider 5 licensed transmitters and N users randomly located in a $1000m \times 1000m$ area. The licensed transmitters operate on 5 orthogonal channels, respectively. We assume that the

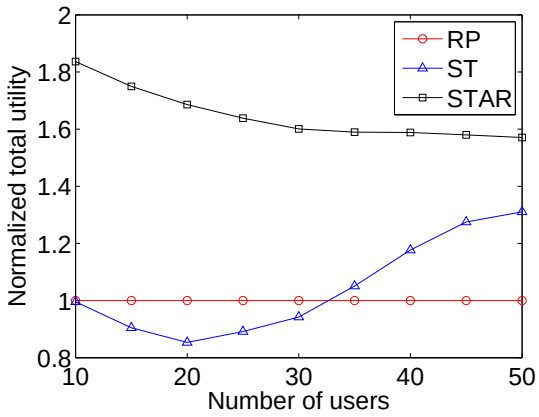


Figure 3.24: Impact of N for practical setting.

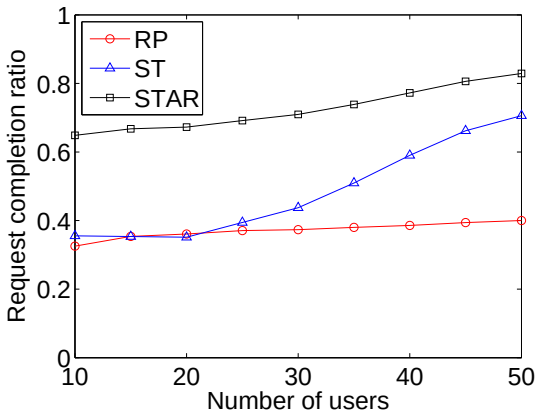


Figure 3.25: Impact of N for practical setting.

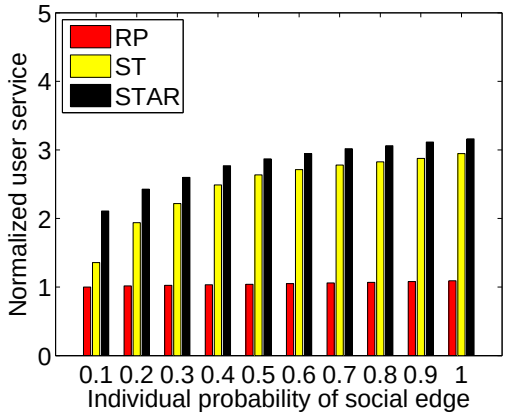


Figure 3.26: Impact of P_{S_i} for random setting.

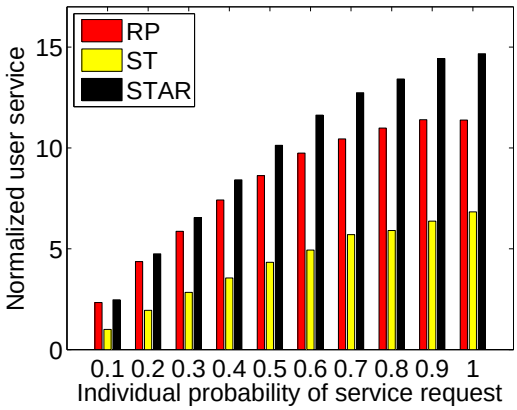


Figure 3.27: Impact of P_{R_i} for random setting.

utility of a user's sensing service for a channel is equal to the inverse of its distance from the licensed transmitter that operates on that channel. Each user randomly selects one channel, and requests sensing service for that channel from at most 3 users randomly selected from the other users who have better channel conditions than itself for that channel. We assume that the sensing service is indivisible. The social credit limit and the service request amount are randomly drawn from $\{1, \cdots, N_S\}$ and $\{1, \cdots, N_R\}$, respectively. We set $N_S = N_R = 5$ as default values.

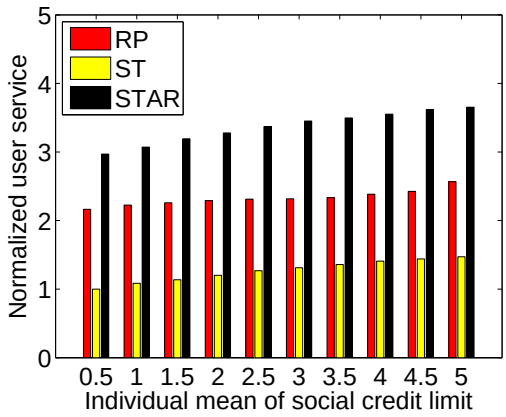


Figure 3.28: Impact of μ_{S_i} for random setting.

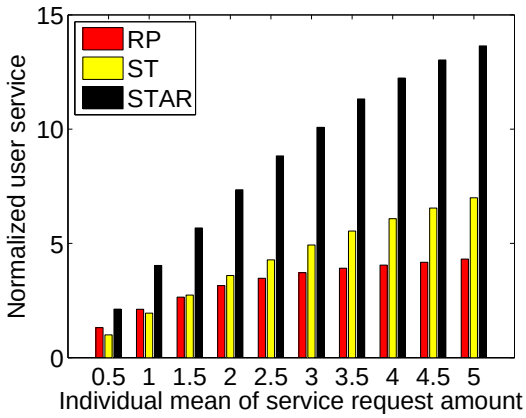


Figure 3.29: Impact of μ_{R_i} for random setting.

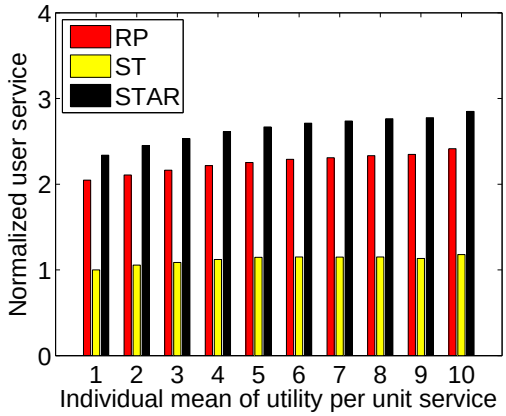


Figure 3.30: Impact of μ_{U_i} for random setting.

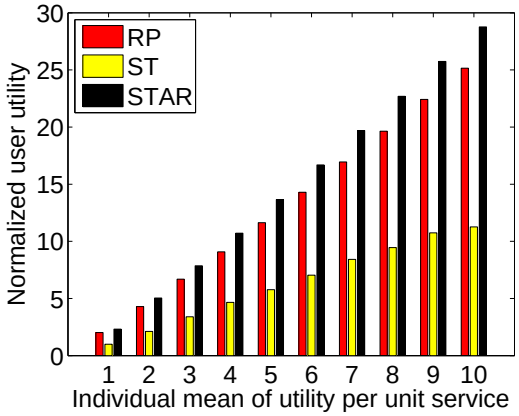


Figure 3.31: Impact of μ_{R_i} for random setting.

3.5.2 Simulation Results

3.5.2.1 System Efficiency

We compare the system performance of the STAR mechanism with the benchmark mechanisms RP and ST. We first evaluate the maximum total amount of service provided under different mechanisms. To highlight the performance comparison, we normalize the results of STAR and ST with respect to RP. We illustrate the impact of P_S, P_R, μ_S, μ_R, and N on the maximum total amount of provided service in Figs.

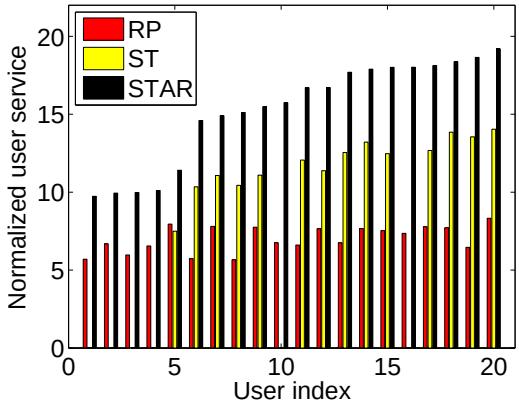


Figure 3.32: Individual user service amount for practical setting.

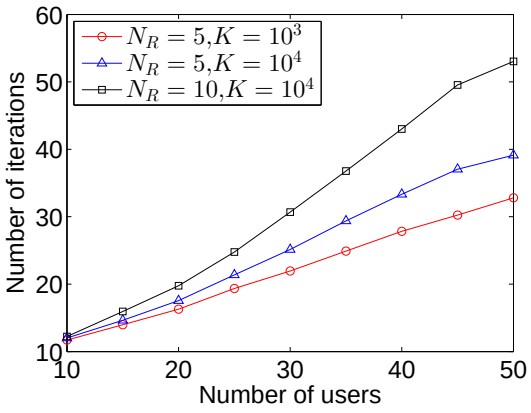


Figure 3.33: Number of cycle-canceling iterations in Algorithm 1 for practical setting.

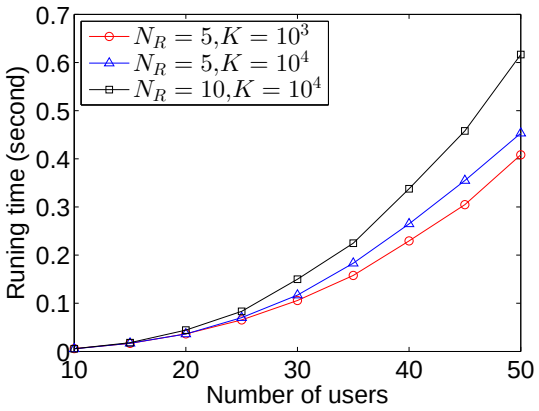


Figure 3.34: Running time of Algorithm 1 for practical setting.

3.14-3.18, respectively. As expected, the performance of STAR always dominates that of RP and ST, which is due to that STAR jointly exploits social trust and reciprocity. Figs. 3.14 and 3.16 show that STAR and ST perform better with respect to RP as P_S or μ_S increases. This is because that as social trust improves, more service can be provided using social trust under STAR and ST, while RP does not benefit from the improved social trust. On the other hand, Figs. 3.15 and 3.17 show that STAR and ST perform worse with respect to RP as P_R or μ_R increases. The reason is that as users have more service requests among each other, a significant part of the increment

in service request can be satisfied using reciprocity. We observe from Figs. 3.14 and 3.16 that the performance gap between STAR and ST decreases as P_S increases, while it remains almost the same as μ_S increases. This shows that the connectivity of the social network has a greater impact on the performance of ST than the social trust levels. Due to this reason, Fig. 3.18 shows that the performance gap between STAR and ST decreases as N increases, since the connectivity of the social network improves as the number of users increases. We also evaluate the maximum total utility of service provided under different mechanisms. We illustrate the impact of μ_U on the maximum total utility of provided service in Fig. 3.19. We observe that the performance gaps among RP, ST, and STAR remain almost the same as μ_U increases. This is as expected, since the utility per unit service acts as a "scaling" factor that has the same effect on the performance of different mechanisms.

Next we evaluate the *request completion ratio* under different mechanisms, which is defined as the ratio of the amount of provided service to the amount of requested service. We illustrate the impact of P_R, μ_R, and N on the request completion ratio in Figs. 3.20-3.22, respectively. We observe from Figs. 3.20 and 3.21 that for all mechanisms, the total amount of provided service increases faster than that of requested service as P_R increases, while it increases slower as μ_R increases. This shows that a large diversity of users' service requests is beneficial for system efficiency. Due to this reason, as illustrated in Fig. 3.22, the request completion ratio improves for all mechanisms as the number of users increases.

For the practical setting, Figs. 3.23-3.25 illustrate the total service amount, total service utility, and request completion ratio when the total utility of provided service is maximized, respectively, as N increases. We can see that STAR always significantly outperforms RP and ST, with a performance gain ranging from 14% to 82%, especially when the number of users is small.

3.5.2.2 Individual Performance

We evaluate individual users' performance under different mechanisms when the system efficiency is maximized. To demonstrate the impact of a particular parameter, we vary that parameter for different users, while keeping other parameters the same for all users. We also normalize the results to highlight the performance comparison. In Figs. 3.26-3.29, we illustrate the amount of received service of each user (i.e., the amount of satisfied service requests of each user) for a system of 10 users when the total amount of provided service is maximized, where users are different only in a user's probability of having a social edge from another user P_{S_i}, probability of having service request from another user P_{R_i}, mean of social credit limit from another user μ_{S_i}, and mean of service request amount from another user μ_{R_i}, respectively. In Figs. 3.30-3.31, we illustrate each user's received service amount and received service utility when the total utility of provided service is maximized, where users are different only in a user's utility per unit service μ_{U_i}. We observe that each user always performs better under STAR than under RP and ST. This shows that STAR can improve each individual user's performance while the system objective is to maximize system efficiency. We also observe that an individual user performs better than other users if it has a larger parameter value than others. This shows that STAR can achieve *service differentiation*, which is a desirable property for fairness: if a user has more social trust or service requests from others than other users have, then that user can also receive more service than others.

Fig. 3.32 illustrates each individual user's received service amount for the 20 users in the real dataset [32] with the social network structure as given in Fig. 3.10. We observe from Fig. 3.11 that the degree of social edge can be very different for different users in real social networks. Accordingly, Fig. 3.32 shows that users with

higher degrees (with larger user indices) receive more service than those with lower degrees.

3.5.2.3 Computational Complexity

We evaluate the computational complexity of using Algorithm 1 to find the maximum total utility of provided service under the STAR mechanism for the practical setting. We convert the service utility values into integers by multiplying them by a large integer K and rounding them down to the respective nearest integers. We run simulations on a Windows 7 desktop with 3.1GHz CPU and 8GB memory. We illustrate the number of cycle-canceling iterations (i.e., the iterations of the **while** loop) and the running time of executing Algorithm 1 as N increases for different values of N_R and K in Fig. 3.33 and Fig. 3.34, respectively. We assume that $N_S = N_R$. We observe that the number of iterations increases almost linearly in the number of users while the running time is increasing quadratically. This shows that Algorithm 1 is scalable for large systems in practice. As expected, we also observe that the computational complexity is higher when N_R or K is larger.

3.6 Related Work

There have been numerous studies on incentive design for stimulating user cooperation in networks. Existing literature on this subject can be broadly classified into three categories. One category of work makes use of reciprocity (also known as *barter*) [33–38]. Although a reciprocity-based approach is simple to implement, it is inefficient in general since synchronously matched requests are unusual. Another category is based on (virtual) currency [39–42], in which a user earns currency by providing service to others and spends currency to receive service from others. The use of currency as a medium of exchange overcomes the shortcoming of reciprocity-based approaches by enabling users to "asynchronously trade" service. However, a major

drawback of using currency is that it incurs a high implementation overhead, mainly due to the need to inhibit malicious manipulation among users without mutual trust. Consider, for example, *Bitcoin* [43] which has recently drawn widespread attention as a promising digital currency. The creation and transfer of bitcoins need to consume considerable computing resources so that they can be secured against potential cheating using cryptographic tools. Reputation-based approaches [44–46] constitute the third category. Since reputation score can be viewed as a form of currency, these approaches share the same advantages and disadvantages as the currency-based ones.

The social credit model used in this chapter falls into the class of *credit networks* [27, 47–50]. The credit is similar to a currency in that there is a need to keep track of the credit information between each pair of neighbor users in the credit network. However, since the credit is ensured by existing trust among users, it obviates the need to secure the credit against cheating, and therefore can reduce implementation overhead significantly.

Compared to the studies mentioned above, the STAR mechanism overcomes the inefficiency of only using reciprocity by using social credit as a "local" currency, while it also circumvents the high implementation overhead incurred by a currency-based approach since social credit is "secured" by existing social trust. Therefore, STAR can efficiently stimulate users to provide service in a cost-effective way.

Exploiting social aspect for mobile networking is an emerging paradigm for network design and optimization [16, 51–53]. Very few work have exploited both social trust and reciprocity for stimulating cooperation in networks. [16] has recently studied using social trust and reciprocity to stimulate cooperative communication based on a coalitional game. Our work is different from [16] in that each user in the latter can participate in *at most one* reciprocity cycle and social trust levels are *unlimited* therein.

3.7 Conclusion

In this chapter, we have proposed a socially-aware crowdsensing system that exploits social trust to stimulate users' participation. The incurred implementation overhead is low since it obviates the need of a global currency. For this system, we have designed STAR, an incentive mechanism using a synergistic marriage of social trust and reciprocity. Based on the STAR mechanism, we have shown that all sensing requests can be satisfied if and only if users who request more sensing service than they can provide can transfer sufficient social credit to users who can provide more than they request. We have also developed an efficient algorithm to maximize the utility of sensing service provided under STAR, for both cases of divisible and indivisible service. Extensive simulation results have confirmed that STAR can achieve significantly better efficacy than using social trust only or reciprocity only.

Chapter 4

EXPLOITING SOCIAL SERVICES TO BOOST DATA USAGE IN WIRELESS SERVICES

4.1 Introduction

The past few years have witnessed pervasive penetration of mobile devices in people's daily life, thanks to the wireless technology advances. Motivated by many social applications on mobile platforms (e.g., WeChat, WhatsApp [54, 55]), mobile users' data usage behaviors have been increasingly influenced by their social relationships. In 2014, the number of online social media users on mobile platforms has reached 1.6 billion, accounting for 44% of mobile users and 80% of online social media users [56].

The popularity of social services on mobile platforms also gives opportunities to wireless service providers who operate the mobile networks. Intuitively, social services can encourage mobile users to demand more data usage by stimulating their interactions with each other through these services (e.g., online social gaming and blogging). When a user increases its activity in a social service, its social friends would also increase their activities. Therefore, users' data usage levels for social services present social network effect to others [57]. This demand increase provides a great potential for wireless providers' revenue increase.

However, this potential benefit is subject to the limited wireless capacity in physical communication networks (e.g., spectrum). As users increase their data usage, they also experience more congestion (e.g., service delays), which discourages them to use more. The increasing congestion poses a significant challenge for wireless providers to increase their revenues.

As a result, mobile users' data usage behaviors are not only subject to congestion effect in the physical network, but also social network effect in the social

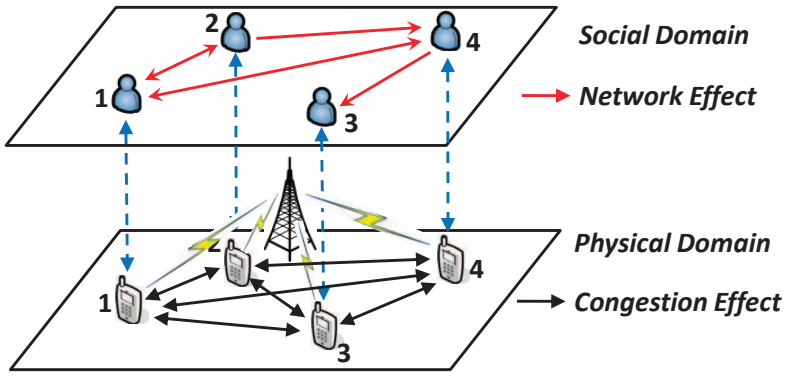


Figure 4.1: Illustration of socially-aware mobile users.

network (as illustrated in Fig. 4.1), which has been largely overlooked by traditional wireless providers. To fully exploit the potential benefit brought by social services, it is necessary to investigate users' data usage behaviors in these two domains, so that a wireless provider can take the best strategy in favor of its revenue. With this insight, we not only analyze users' interactions subject to both network and congestion effects, but also study the optimal pricing strategy for the wireless provider.

The main contributions of this chapter can be summarized as follows.

- *Stackelberg game formulation:* By jointly considering users' social relationships and the wireless network's congestion, we formulate the interaction between the wireless provider and mobile users as a Stackelberg game: In Stage I, the wireless provider chooses a price to maximize its revenue; in Stage II, mobile users choose their data usage levels based on the price to maximize their socially-aware payoffs.

- *Equilibrium analysis for user demands in Stage II:* We first give a general condition under which there exists a user demand equilibrium (UDE), and then we show that under a further general condition the game is a concave game and thus admits a unique UDE. We also propose a distributed algorithm for users

to achieve the UDE. Next we show that a user's usage can increase when price increases. We further show that if the social network is symmetric, the total usage always increases when a user's parameter (e.g., social tie) improves.

- *Provider's optimal pricing in Stage I:* Finally, by taking into account users' equilibrium demands, we develop an optimal pricing algorithm to maximize the revenue of the wireless provider. We evaluate the performance of total usage and revenue by simulations, and draws useful engineering insights for the wireless provider's operation.

The rest of this chapter is organized as follows. In Section 4.2, we formulate the Stackelberg game between the wireless provider and mobile users. Section 4.3 studies users' demand equilibrium in Stage II. Section 4.4 studies the provider's optimal pricing strategy in Stage I. Simulation results and discussions are given in Section 4.5. Related work are reviewed in Section 4.6. Section 4.7 concludes this chapter.

4.2 System Model

4.2.1 Socially-aware Wireless Service

Consider a set of users $\mathcal{N} \triangleq \{1, \ldots, N\}$ participating in a wireless data service provided by a wireless operator (e.g., AT&T). Each user $i \in \mathcal{N}$ consumes an amount of data usage in the wireless service, denoted by x_i where $x_i \in [0, \infty)$. Let $\mathbf{x} \triangleq (x_1, \ldots, x_N)$ denote the usage profile of all the users and $\mathbf{x}_{-i}$ denote the the usage profile without user i. Affected by the other users' usage subject to congestion effect due to the limited network resource, the payoff of user i by consuming data usage x_i is

$$v_i(x_i, \mathbf{x}_{-i}, p) = a_i x_i - \frac{1}{2} b_i x_i^2 - \frac{1}{2} c(\sum_{j \in \mathcal{N}} x_j)^2 - p x_i,$$

where $a_i > 0$ and $b_i > 0$ are *internal utility coefficients* that capture the intrinsic value of wireless service to user i, $c > 0$ is a *congestion coefficient* that is determined

by the resources constraints of the wireless network, and p is the usage-based price charged by the wireless provider[1]. As in [58], the quadratic form of the internal utility function not only allows for tractable analysis, but also serves as a good second-order approximation for a broad class of concave utility functions. In particular, a_i models the maximum internal demand rate, and b_i models the internal demand elasticity factor. For the congestion model, the quadratic sum form reflects that a user's congestion experience is affected by all the users, and the marginal cost of congestion increases as the total usage increases.

Traditional wireless providers' operation does not take into account the fact that social services encourage mobile users to demand more data usage. We thus account for this effect in our model. Then user i's payoff includes the addition of social utility, i.e.,

$$u_i(x_i, \mathbf{x}_{-i}, p) = a_i x_i - \frac{1}{2} b_i x_i^2 + \sum_{j \neq i} g_{ij} x_i x_j - \frac{1}{2} c (\sum_{j \in \mathcal{N}} x_j)^2 - p x_i \tag{4.1}$$

where $g_{ij} \geq 0$ is the *social tie* that quantifies the social influence from user j to user i. As in [58], the product form $g_{ij} x_i x_j$ of the social utility function captures that a user derives more utility by increasing its usage in social services, and the marginal gain of social utility increases as its social friends increase their usage. Therefore, the social services bring in social network effect among users and can increase their utilities.

4.2.2 Stackelberg Game Formulation

We model the interaction between the wireless provider and mobile users for the socially-aware wireless service as a two-stage Stackelberg game.

Definition 4.1 (Two-Stage Pricing-Usage Game)**.**

[1]Usage-based pricing is widely used in practice by wireless operators to control the demand. The price is the same here for all users to ensure fairness.

- Stage I (Pricing): *The wireless provider chooses price p to maximize its revenue:*

$$p^* = \arg \max_{p \in [0,\infty)} t(\mathbf{x}) p$$

where $t(\mathbf{x}) \triangleq \sum_{i \in \mathcal{N}} x_i$ denotes the total usage under strategy profile $\mathbf{x}$;

- Stage II (Usage): *Each user $i \in \mathcal{N}$ chooses its data usage level x_i to maximize its payoff given the price p and the usage levels of the other users $\mathbf{x}_{-i}$:*

$$x_i^* = \arg \max_{x_i \in [0,\infty)} u_i(x_i, \mathbf{x}_{-i}, p).$$

We study the two-stage pricing-usage game by backward induction [59]. For Stage II, given a price p chosen by the wireless provider in Stage I, we are interested in the existence of a stable outcome of users' interactions at which no user will deviate. This leads to the concept of user equilibrium.

Definition 4.2 (User Demand Equilibrium)**.** *For any price p given in Stage I, the user demand equilibrium (UDE) in Stage II is a strategy profile $\mathbf{x}^*$ such that no user can improve its payoff by unilaterally changing its usage, i.e.,*

$$x_i^* = \arg \max_{x_i \in [0,\infty)} u_i(x_i, \mathbf{x}_{-i}^*, p), \ \forall i.$$

Given the UDE in Stage II, we will study the optimal pricing strategy for the wireless provider in Stage I.

4.3 Stage II: User Demand Equilibrium

In this section, we study users' demands in Stage II.

Using the concave payoff function (4.1), by setting the derivative $\frac{\partial u_i(x_i, \mathbf{x}_{-i})}{\partial x_i} = 0$ as the first-order condition, we obtain the *best response function* of user i as

$$r_i(\mathbf{x}_{-i}) = \max \left\{ 0, \ \frac{a_i - p}{b_i + c} + \sum_{j \neq i} \frac{g_{ij} - c}{b_i + c} x_j \right\}. \tag{4.2}$$

According to (4.2), each user i's usage demand consists of two parts: *internal demand* $\frac{a_i-p}{b_i+c}$ that is independent of the other users, and *external demand* $\sum_{j\neq i}\frac{g_{ij}-c}{b_i+c}x_j$ that depends on the other users. The coefficient $\frac{g_{ij}-c}{b_i+c}$ represents the marginal increase or decrease of user i's demand as user j's usage increases: when $g_{ij} > c$, user i's impact from user j is dominated by social network effect; when $g_{ij} < c$, it is dominated by congestion effect.

4.3.1 Existence and Uniqueness of UDE

We first investigate the existence of UDE in Stage II. We make the following assumption.

Assumption 4.1. $\sum_{j\neq i}\frac{|g_{ij}-c|}{b_i+c} < 1,\ \forall i.$

This assumption is for analysis tractability. It is an important condition to guarantee the existence of UDE, as there can exist no UDE when it does not hold (as illustrated by an example in Fig. 4.2). According to the best response function (4.2), Assumption 1 implies that any user's absolute external demand $|\sum_{j\neq i}\frac{g_{ij}-c}{b_i+c}x_j|$ is less than the maximum usage $\max_{j\neq i} x_j$ among all the other users. This is a mild condition as the aggregate effect experienced by a user from all the other users would be less than the largest effect the user can experience from an individual of the other users. A similar assumption is made in [58] for similar considerations.

Now we can show that there always exists a UDE in Stage II.

Theorem 4.1. *Under Assumption 1, the Stage II game admits a UDE.*

The proof is given in Appendix and the main idea is to show that the game has an equivalent game which admits a UDE.

Next we give another general technical condition under which the game admits a unique UDE.

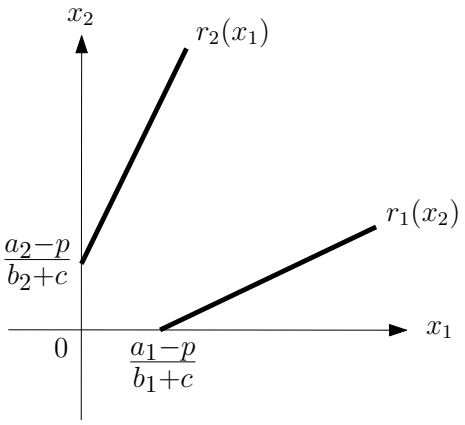


Figure 4.2: Illustration of Stage II for two users.

Theorem 4.2. *Under Assumption 1, the Stage II game admits a unique UDE if* $\sum_{j\neq i} \frac{|g_{ji}-c|}{b_i+c} < 1$, $\forall i$.

The proof is given in Appendix and the main idea is is to show that the game is a concave game [60], and thus admits a unique UDE.

Remark: According to Theorem 4.2, it is worth noting that the Stage II game admits a unique UDE when users' social ties are symmetric (i.e., $g_{ij} = g_{ji}$, $\forall i \neq j$). The symmetric setting of social networks is of great interests. Motivated by the idea of social reciprocity [33], a user's social behavior to another is likely to imitate the latter's behavior to the former. As a result, two users' social ties to each other tend to be the same.

4.3.2 Computing and Achieving UDE

As we have showed the existence of UDE, we then design an algorithm to compute the UDE, as described in Algorithm 2. The algorithm iteratively updates users' strategies based on their best response functions (4.2) and converges to the UDE.

Theorem 4.3. *Algorithm 2 computes the UDE in Stage II.*

The proof is given in Appendix and the main idea is to show that the best re-

Algorithm 2: Compute the UDE in Stage II

input: precision threshold ϵ;
$x_i^{(0)} \leftarrow 0, \forall i \in \mathcal{N}$; $t \leftarrow 1$;
repeat
  foreach $i \in \mathcal{N}$;
  do
    $x_i^{(t+1)} = \max\left\{0,\ \frac{a_i - p}{b_i + c} + \sum_{j \neq i} \frac{g_{ij} - c}{b_i + c} x_j^{(t)}\right\}$;
  end
  $t \leftarrow t + 1$;
until $\|\mathbf{x}^{(t)} - \mathbf{x}^{(t-1)}\| \leq \epsilon$;
return $\mathbf{x}^{(t)}$;

sponse updates in the algorithm result in a contraction mapping and hence converges to a fixed point.

It is desirable for users to reach the UDE in a distributed manner. We then propose a distributed update algorithm based on Algorithm 2, as described below.

Algorithm 3: Distributed algorithm to achieve the UDE in Stage II

each user $i \in \mathcal{N}$ **chooses** an initial usage $x_i^{(0)} \geq 0$;
loop at each time interval $t = 1, 2, \ldots$
each user $i \in \mathcal{N}$ in parallel:
updates its usage by $\max\left\{0,\ \frac{a_i - p}{b_i + c} + \frac{1}{b_i + c} \sum_{j \neq i, g_{ij} > 0} g_{ij} x_j^{(t)} - \frac{c}{b_i + c} \sum_{j \neq i} x_j^{(t)}\right\}$
end loop

Note that the usage update in Algorithm 3 is equivalent to the best response update in Algorithm 2. Each user i chooses its best response usage based on the usage of its social friends who have social influences to it (i.e., each user j with $g_{ij} > 0$), which can be obtained from the social friends, and the total usage of all users, which can be obtained from the wireless provider. The correctness of Algorithm 3 follows from that of Algorithm 2 and is thus omitted.

Proposition 4.1. *Algorithm 3 achieves the UDE in Stage II.*

4.3.3 Parameter Analysis of UDE

We first investigate the impact of price on the UDE. To draw clean insights, we start with the case for two users. Without loss of generality, assume that $a_1 \geq a_2$.

Proposition 4.2. *For Stage II for two users, there exists a price threshold* $p_{th} \in [0, a_1]$ *where*

$$p_{th} = \frac{a_2(b_1 + c) - a_1(c - g_{21})}{b_1 + g_{21}} \tag{4.3}$$

such that the UDE $\mathbf{x}^*$ *is given as follows, depending on the price* p*:*

- High price regime: *When* $p \geq a_1$, $x_1^* = x_2^* = 0$*;*
- Medium price regime: *When* $p_{th} \leq p < a_1$, $x_1^* = \frac{a_1 - p}{b_1 + c}$ *and* $x_2^* = 0$*;*
- Low price regime: *When* $0 \leq p < p_{th}$,
  $x_1^* = \frac{(a_1 - p)(b_2 + c) - (a_2 - p)(c - g_{12})}{b_1 b_2 - g_{12} g_{21} + c(b_1 + b_2 + g_{12} + g_{21})}$
  and $x_2^* = \frac{(a_2 - p)(b_1 + c) - (a_1 - p)(c - g_{21})}{b_1 b_2 - g_{12} g_{21} + c(b_1 + b_2 + g_{12} + g_{21})}$.

The proof is given in Appendix. According to (4.3), there are three cases of the threshold p_{th} depending on which effect dominates user 2's experience from user 1 (as illustrated in Fig.4.3).

1. *Neither effect:* When $g_{21} = c$, we have $p_{th} = a_2$. As social network effect and congestion effect cancel each other, user 2 experiences neither effect from user 1. Then user 2's usage demand is equal to its internal demand, and it reaches 0 when $p = a_2$.
2. *Congestion effect:* When $g_{21} < c$, we have $p_{th} < a_2$. As user 2 experiences congestion effect from user 1, even when p is less than a_2 such that user 2 has a positive internal demand, its external demand can be sufficiently negative such

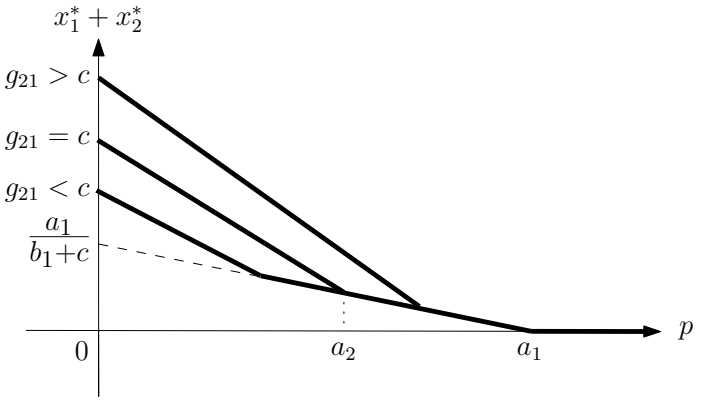


Figure 4.3: Total usage at the UDE in Stage II for two users.

that user 2's usage demand is negative. In particular, when $\frac{a_1}{a_2} \geq \frac{b_1+c}{c-g_{21}}$, user 2's usage is 0 even when the price is 0.

3. *Social network effect:* When $g_{21} > c$, we have $p_{th} > a_2$. As user 2 experiences social network effect from user 1, even when p is greater than a_2 such that user 2 has a negative internal demand, its external demand can be sufficiently positive such that user 2's usage demand is positive.

Next we study the general case with multiple users. For convenience, let us define

$$B = \begin{bmatrix} b_1 + c & c & \cdots & c \\ c & b_2 + c & \cdots & c \\ \vdots & \vdots & \ddots & \vdots \\ c & c & \cdots & b_N + c \end{bmatrix}, G = \begin{bmatrix} 0 & g_{12} & \cdots & g_{1N} \\ g_{21} & 0 & \cdots & g_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ g_{N1} & g_{N2} & \cdots & 0 \end{bmatrix}.$$

Also define C as the $N \times N$ matrix with each entry being c. For a UDE $\mathbf{x}^*$, let $\mathcal{S}$ be the set of users with positive usage in $\mathbf{x}^*$ (i.e., $x_i^* > 0$, $\forall i \in \mathcal{S}$ and $x_i^* = 0$, $\forall i \notin \mathcal{S}$). For convenience, let $\mathbf{v}_{\mathcal{S}}$ denote the $|\mathcal{S}| \times 1$ vector comprised of the entries of a vector $\mathbf{v}$ with indices in $\mathcal{S}$, $M_{\mathcal{S}}$ denote the $|\mathcal{S}| \times |\mathcal{S}|$ matrix comprised of the entries of a matrix M with indices in $\mathcal{S} \times \mathcal{S}$, and $[M]_{i,\mathcal{S}}$ denote the $1 \times |\mathcal{S}|$ vector comprised of the entries of the ith row of a matrix M with column indices in $\mathcal{S}$. According to the

best response function (4.2), $\mathbf{x}^*_{\mathcal{S}}$ is the solution to the system of equations

$$B_{\mathcal{S}}\mathbf{x}_{\mathcal{S}} = \mathbf{a}_{\mathcal{S}} - p\mathbf{1}_{\mathcal{S}} + G_{\mathcal{S}}\mathbf{x}_{\mathcal{S}}$$

where $\mathbf{1}$ denotes the $N \times 1$ vector of 1s. We need the following lemma:

Lemma 4.1. *$(B_{\mathcal{S}} - G_{\mathcal{S}})$ is invertible for any set $\mathcal{S} \subseteq \mathcal{N}$.*

The proof of Lemma 4.1 is given in our anonymous online technical report [61]. Thus we have

$$\mathbf{x}^*_{\mathcal{S}} = (B_{\mathcal{S}} - G_{\mathcal{S}})^{-1}(\mathbf{a}_{\mathcal{S}} - p\mathbf{1}_{\mathcal{S}}). \tag{4.4}$$

When users have the same internal coefficients a_i, we can show that the same set of users have positive equilibrium usage at different prices.

Proposition 4.3. *For Stage II, when $a_i = a$, $\forall i$, the UDE $\mathbf{x}^*$ is given as follows, depending on the price p:*

- *when $p > a$, $x^*_i = 0$, $\forall i$;*
- *when $0 \leq p \leq a$, there exists a set $\mathcal{S} \subseteq \mathcal{N}$ such that for any $p \in [0, a)$,*

$$x^*_i = [(B_{\mathcal{S}} - G_{\mathcal{S}})^{-1}(\mathbf{a}_{\mathcal{S}} - p\mathbf{1}_{\mathcal{S}})]_i > 0, \forall i \in \mathcal{S},$$

and $x^*_i = 0$, $\forall i \notin \mathcal{S}$, where $[M]_i$ denotes the ith row of matrix M.

The proof is given in Appendix. Proposition 4.3 shows that the set of users with positive equilibrium usage (if they exist) does not change with price, and each user's positive usage decreases when price increases.

We then show by a counterexample that if users have different internal coefficients a_i, a user's equilibrium usage can increase when price increases. Consider

a three-user game where $a_1 = 2$, $a_2 = a_3 = 1.5$, $b_i = 3$, $\forall i$, $c = 2$, $p = 0.4$, and $g_{23} = g_{32} = 4$, $g_{ij} = 0$, $\forall \{i, j\} \neq \{2, 3\}$[2]. We can show that there exists a unique UDE $\mathbf{x}^*$ and it is the solution to the system of equations:

$$5x_1 + 2x_2 + 2x_3 = 1.6$$

$$5x_2 + 2x_1 - 2x_3 = 1.1$$

$$5x_3 - 2x_2 + 2x_1 = 1.1.$$

Solving these equations, we have $x_1^* = 0.0571$, $x_2^* = 0.3286$, $x_3^* = 0.3286$. When price p increases to 0.5, the new UDE is the solution to

$$5x_1 + 2x_2 + 2x_3 = 1.5$$

$$5x_2 + 2x_1 - 2x_3 = 1$$

$$5x_3 - 2x_2 + 2x_1 = 1$$

which is $x_1^* = 0.0714 > 0.0571$, $x_2^* = 0.2857$, $x_3^* = 0.2857$. Thus the usage of user 1 increases.

Remark: Intuitively, when the price increases, the usage of both user 2 and 3 decrease and the internal demand of user 1 decreases. However, as user 1 experiences strong congestion effect from both user 2 and 3, user 1's external demand increases due to the decrease of congestion effect, and it increases faster than the decrease of user 1's internal demand as price increases, such that the total of internal and external demand increases. In addition, a larger internal coefficient a_1 of user 1 than that of user 2 and 3 allows user 1 to have a positive equilibrium usage $x_1^* = 0.0714$ even when its external demand is negative due to the strong congestion effect. Indeed, if user 1 has the same internal coefficient $a_1 = 1.5$ as user 2 and 3, then we can show that its equilibrium usage is 0.

[2]Note that Assumption 1 holds under this setting.

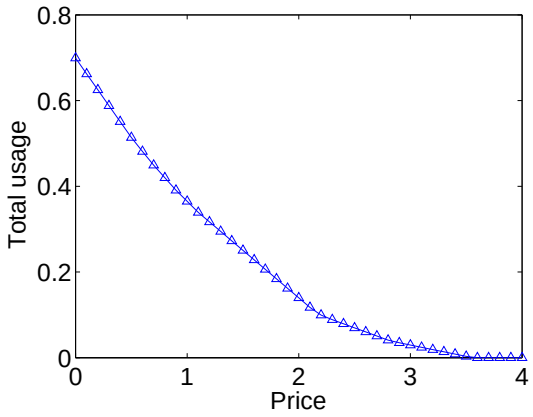


Figure 4.4: Total usage at the UDE is a piece-wise linear function of price.

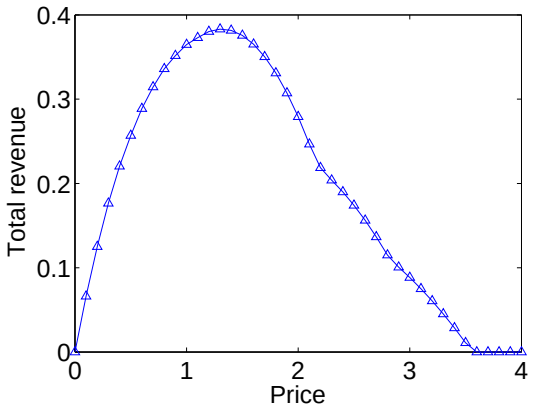


Figure 4.5: Total revenue at the UDE is a piece-wise quadratic function of price.

Furthermore, when users have different internal coefficients a_i, we have the following result.

Proposition 4.4. *For Stage II, the UDE* $\mathbf{x}^*$ *is given as follows, depending on the price* p*:*

- *when* $p > \max_{i\in\mathcal{N}} a_i$, $x_i^* = 0$, $\forall i$;
- *when* $0 \le p \le \max_{i\in\mathcal{N}} a_i$, *there is a set of prices* $p_0 \triangleq 0 < p_1 < \cdots < p_M < p_{M+1} \triangleq \max_{i\in\mathcal{N}} a_i$, *and for each* $k \in \{0,\ldots,M\}$, *there exists a set* $\mathcal{S}_k \subseteq \mathcal{N}$ *such that for any* $p \in [p_k, p_{k+1}]$,

$$x_i^* = [(B_{\mathcal{S}_k} - G_{\mathcal{S}_k})^{-1}(\mathbf{a}_{\mathcal{S}_k} - p\mathbf{1}_{\mathcal{S}_k})]_i > 0, \forall i \in \mathcal{S}_k$$

and $x_i^* = 0$, $\forall i \notin \mathcal{S}_k$.

The proof is given in Appendix. Proposition 4.4 shows that each user's equilibrium usage is a piece-wise linear function of price: within each interval of price $[p_k, p_{k+1}]$, the equilibrium usage is a linear function of the price p.

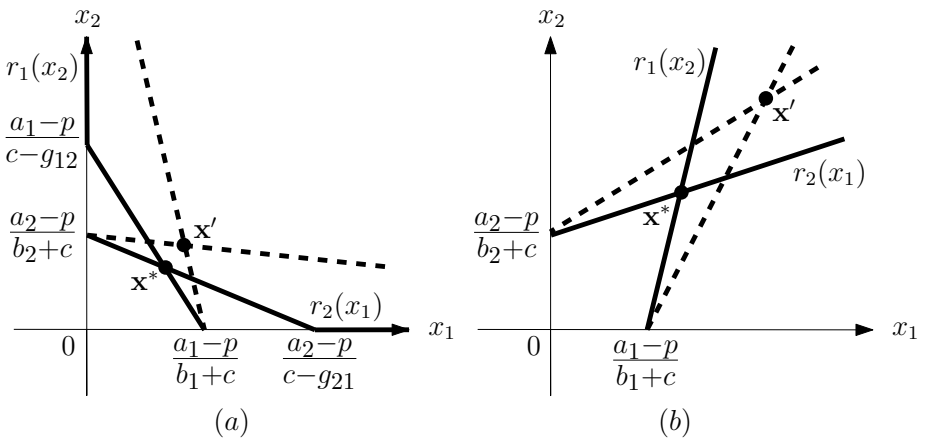


Figure 4.6: An example of Stage II for two users.

Next we investigate the impacts of other parameters on the UDE. We show that total usage always increases when a user's parameter improves, under the condition that users' social network is symmetric. For tractable analysis, we assume that users have the same internal coefficient a_i[3].

Proposition 4.5. *For Stage II, when $a_i = a$, $\forall i$ and social ties are symmetric (i.e., $g_{ij} = g_{ji}$, $\forall i \neq j$), the total equilibrium usage increases when a or any g_{ij} increases, or any b_i or c decreases.*

The proof is given in Appendix. We illustrate Proposition 4.5 by an example in Fig. 4.6. In particular, the unique UDE $\mathbf{x}^*$ is achieved at the intersection of the best response functions (bold lines): (a) $\frac{g_{12}-c}{b_1+c} \in (-1, 0)$, $\frac{g_{21}-c}{b_2+c} \in (-1, 0)$; (b) $\frac{g_{12}-c}{b_1+c} \in (0, 1)$, $\frac{g_{21}-c}{b_2+c} \in (0, 1)$. When $g_{12} = g_{21}$ increases, the UDE $\mathbf{x}^*$ moves to $\mathbf{x}'$. As mentioned before, users' social ties tend to be symmetric in practice due to social reciprocity [33]. In Section 4.5, simulation results will show that the performance under asymmetric social ties is very close to that under symmetric social ties.

[3]Users can still have different b_i.

4.4 Stage I: Optimal Pricing

In the previous section, we have investigated UDE in Stage II given a price chosen by the wireless provider. In this section, we study the optimal pricing of the provider in the Stage I game.

We first observe from Proposition 4.4 that the total usage is a piece-wise linear function of price (as illustrated in Fig. 4.5(a)). As a result, the total revenue is a piece-wise quadratic function of price (as illustrated in Fig. 4.5(b)). Based on this observation, we develop an algorithm that computes the optimal price to maximize the provider's revenue, as described in Algorithm 4. The basic idea is to first determine the price intervals that characterize the piece-wise structure, such that within each price interval, the set of users with positive usage is the same at any price. Then we find the optimal price within each interval that maximizes the revenue. Thus we can find the optimal price with the maximum revenue among all the intervals.

In particular, Algorithm 4 starts with computing the set of users $\mathcal{S}$ with positive usage at price 0 by using Algorithm 2. Then this set $\mathcal{S}$ serves as the initial condition for the following steps. As the price p increases from 0 to $\max_{i\in\mathcal{N}} a_i$ (which is the largest possible value of the optimal price according to Theorem 4.4), it iteratively finds the critical prices at which the set $\mathcal{S}$ changes. In each iteration, given the current critical price $\underline{p}$, the next critical price $\overline{p}$ is the minimum price greater than $\underline{p}$ at which some user $i \in \mathcal{S}$ with positive usage decreases its usage to 0, or some user $i \notin \mathcal{S}$ with usage 0 increases its usage to a positive value[4]. Within each price interval $[\underline{p}, \overline{p}]$, as the revenue R is a quadratic function of price p, the optimal price in $[p, \overline{p}]$ that maximizes the revenue is the price $\hat{p}$ such that $\frac{\partial R(p)}{\partial p}|_{p=\hat{p}} = 0$, if $\hat{p}$ is in

[4]Recall that some user's equilibrium usage can increase when price increases as illustrated by the example in Section 4.3.

Algorithm 4: Compute the optimal price to maximize revenue in Stage I

compute the UDE $\mathbf{x}^*$ at price 0 using Algorithm 2;
find the set of users $\mathcal{S}$ with positive usage at the UDE $\mathbf{x}^*$;
$\underline{p} \leftarrow 0$; $p^* \leftarrow 0$; $r^* \leftarrow 0$;
while $\underline{p} \leq \max_{i \in \mathcal{N}} a_i$ *and* $\mathcal{S} \neq \varnothing$ **do**
  $\mathcal{S}' \leftarrow \varnothing$; $\mathcal{S}'' \leftarrow \varnothing$;
  foreach $i \in \mathcal{S}$ **do**
    if $[(B_{\mathcal{S}} - G_{\mathcal{S}})^{-1}]_i \mathbf{1}_{\mathcal{S}} > 0$ **then**
      $\mathcal{S}' \leftarrow \mathcal{S}' \cup \{i\}$; $\tilde{p}_i \leftarrow \frac{[(B_{\mathcal{S}} - G_{\mathcal{S}})^{-1}]_i \mathbf{a}_{\mathcal{S}}}{[(B_{\mathcal{S}} - G_{\mathcal{S}})^{-1}]_i \mathbf{1}_{\mathcal{S}}}$;
    end
  end
  foreach $i \notin \mathcal{S}$ **do**
    if $[G - C]_{i,\mathcal{S}} (B_{\mathcal{S}} - G_{\mathcal{S}})^{-1} \mathbf{1}_{\mathcal{S}} < -1$ **then**
      $\mathcal{S}'' \leftarrow \mathcal{S}'' \cup \{i\}$; $\tilde{p}_i \leftarrow \frac{[G-C]_{i,\mathcal{S}} (B_{\mathcal{S}} - G_{\mathcal{S}})^{-1} \mathbf{a}_{\mathcal{S}} + a_i}{[G-C]_{i,\mathcal{S}} (B_{\mathcal{S}} - G_{\mathcal{S}})^{-1} \mathbf{1}_{\mathcal{S}} + 1}$;
    end
  end
  $\overline{p} \leftarrow \min_{i \in \mathcal{S}' \cup \mathcal{S}''} \tilde{p}_i$; $k \leftarrow \arg\min_{i \in \mathcal{S}' \cup \mathcal{S}''} \tilde{p}_i$; $\hat{p} \leftarrow \frac{\mathbf{1}_{\mathcal{S}}^T (B_{\mathcal{S}} - G_{\mathcal{S}})^{-1} \mathbf{a}_{\mathcal{S}}}{2 \mathbf{1}_{\mathcal{S}}^T (B_{\mathcal{S}} - G_{\mathcal{S}})^{-1} \mathbf{1}_{\mathcal{S}}}$;
  if $\hat{p} \in [\underline{p}, \overline{p}]$ **then**
    $p' \leftarrow \hat{p}$;
    else
      if $\hat{p} < \underline{p}$ **then**
        $p' \leftarrow \underline{p}$;
        else
          $p' \leftarrow \overline{p}$;
        end
      end
    end
  end
  $r' \leftarrow p' \mathbf{1}_{\mathcal{S}}^T (B_{\mathcal{S}} - G_{\mathcal{S}})^{-1} (\mathbf{a}_{\mathcal{S}} - p' \mathbf{1}_{\mathcal{S}})$;
  if $r' > r^*$ **then**
    $p^* \leftarrow p'$; $r^* \leftarrow r'$;
  end
  $\underline{p} \leftarrow \overline{p}$;
  if $k \in \mathcal{S}$ **then**
    $\mathcal{S} \leftarrow \mathcal{S} \setminus \{k\}$;
    else
      $\mathcal{S} \leftarrow \mathcal{S} \cup \{k\}$;
    end
  end
end
return p^*, r^*;

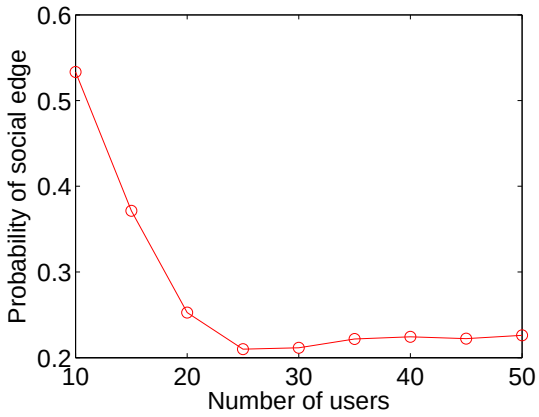


Figure 4.7: Probability of social edge in real data trace.

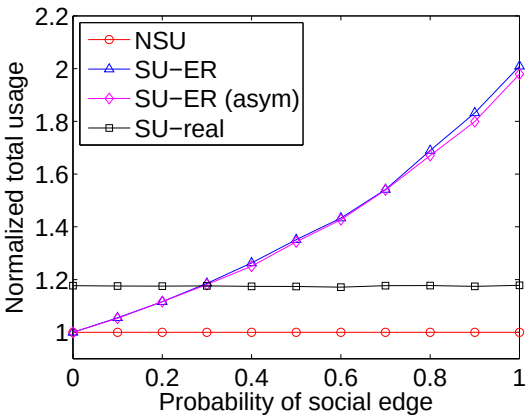


Figure 4.8: Impact of P_S.

$[\underline{p}, \overline{p}]$; otherwise, the optimal price is one of the endpoints $\underline{p}$ and $\overline{p}$. By comparing the maximum revenues at the optimal prices for all the price intervals, the algorithm finds the optimal price in the entire range of price.

Theorem 4.4. *Algorithm 4 computes the optimal price in Stage I.*

The proof is given in Appendix. In the next section, numerical results will show that the computational complexity of Algorithm 4 is linear in the number of users.

4.5 Performance Evaluation

In this section, we first use simulation results to evaluate the performance of the two-stage game between the mobile users and wireless provider. Then we discuss the engineering insights that can be drawn from the simulation results.

4.5.1 Simulation Setup

To illustrate the impacts of different parameters of mobile social networks on the performance, we consider a random setting as follows. We simulate the social graph G using the Erdős-Rényi (ER) graph model [31], where a social edge exists between

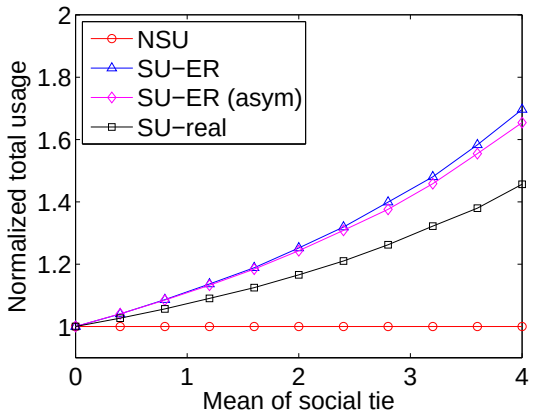


Figure 4.9: Impact of μ_G.

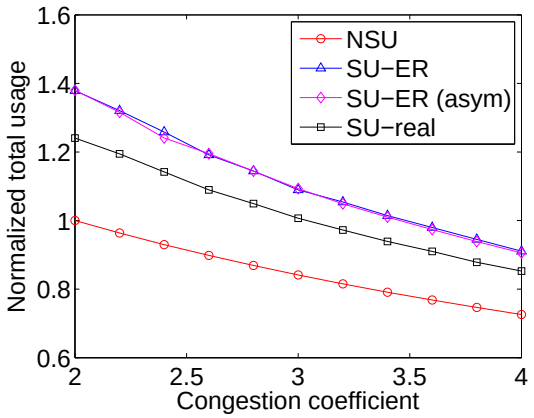


Figure 4.10: Impact of c.

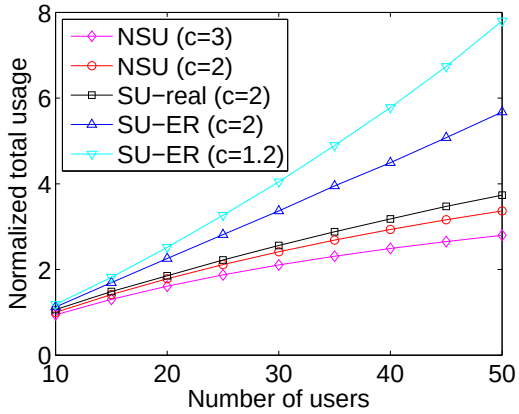


Figure 4.11: Impact of N.

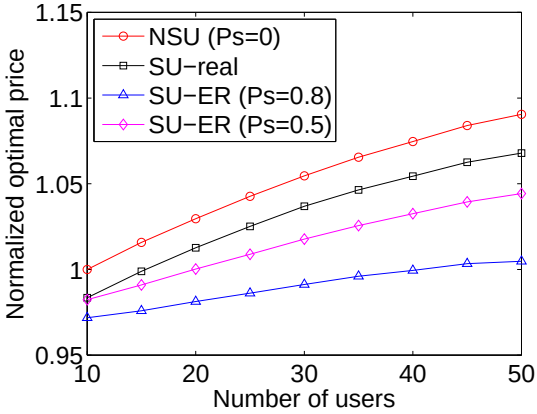


Figure 4.12: Impact of N.

each pair of users with probability P_S. If a social edge exists, the social tie follows a normal distribution $N(\mu_G, 2)$ (with mean μ_G and variance 2). We assume that each a_i follows a normal distribution $N(\mu_A, 2)$, and each b_i follows a normal distribution $N(\mu_B, 2)$. We set default parameter values as follows: $N = 10$, $P_s = 0.8$, $\mu_A = 4$, $\mu_B = 10$, $\mu_G = 4$, $c = 4$. To evaluate the performance in practice, we also simulate the social graph according to the real data trace from Brightkite [32], which is a social friendship network based on mobile phones. For this data trace, we plot the average number of social ties between two users versus the number of users in Fig. 4.7. If a social edge exists based on the real data, the social tie also follows a normal

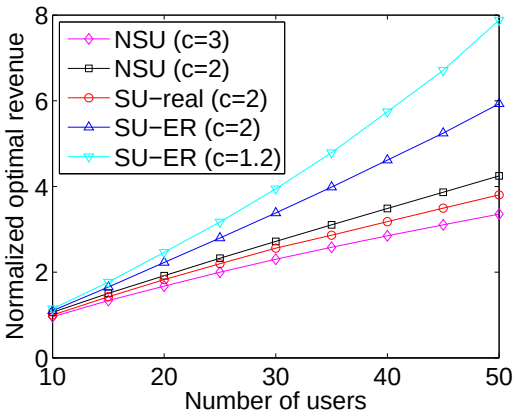


Figure 4.13: Impact of N.

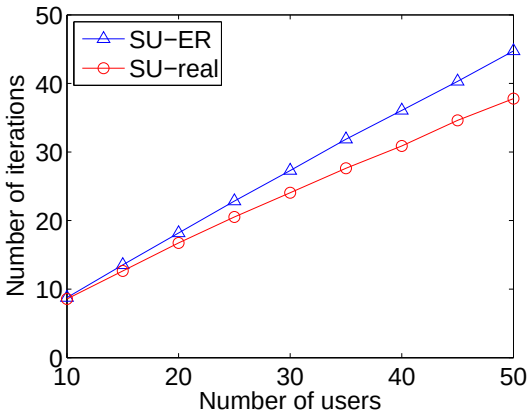


Figure 4.14: Computational complexity of Algorithm 4.

distribution $N(\mu_G, 2)$.

As a benchmark, we evaluate the performance when users demand non-socially-aware usage (NSU) in comparison to our proposed socially-aware usage (SU). Since NSU is a special case of SU with all social ties being 0, the UDE and optimal pricing for NSU can be computed as for SU. To highlight the performance comparison, we normalize the results with respect to NSU. We also compare the performance under SU with ER model based social graph (SU-ER) and with real data based social graph (SU-real).

4.5.2 Simulation Results

4.5.2.1 Total Usage in Stage II

We first evaluate the performance of total usage in Stage II.

We illustrate the impacts of P_S, μ_G, c on total usage in Figs. 4.8-4.10, respectively. As expected, we observe from all these figures that SU always dominates NSU, and can perform significantly better than NSU. From Figs. 4.8-4.9, we can see that the performance gain of SU over NSU increases as P_S or μ_G increases, and the marginal gain is also increasing. Similarly, we can see from Fig. 4.10 that the performance gain of SU over NSU increases as congestion coefficient c decreases, and the

marginal gain is also increasing. We also evaluate the performance under SU with ER model based the asymmetric social graph. We observe that its performance is very close to that with the symmetric social graph.

Fig. 4.11 illustrates the impact of N on total usage. As expected, we observe that the total usage always increases with the number of users. However, for the case of NSU and SU-real, the marginal gain of total usage decreases with the number of users, while for the case of SU-ER, the marginal gain increases. Intuitively, in the former case, when a new user joins the network, as the new user's social ties with the existing users are weak, the congestion effect between the new user and the existing users outweighs the social network effect between them. Furthermore, as more users exist in the network, the weight difference between the congestion effect and the social network effect increases, and thus the marginal gain of total usage by adding a user decreases. In the latter case, as the new user's social ties with the existing users are strong, the roles of the congestion effect and social network effect are switched.

4.5.2.2 Optimal Price in Stage I

Next we evaluate the performance of the optimal price and optimal revenue in Stage I.

Fig. 4.12 illustrates the optimal price as the number of users increases. We observe that the optimal price always decreases with the number of users. Intuitively, this is because that as the number of users increases, more users have a higher internal demand, so that increasing the price does not result in significant decrease in total usage. Comparing different curves, we can also see that the optimal price decreases as P_S increases from 0 to 0.3 and then to 0.8. Intuitively, this is because that when social network effect is strong, a low price is desirable, since it encourages users' internal usage which further stimulate significantly more usage by social network effect; when congestion effect is strong, a high price is desirable, since decreasing the price cannot significantly encourage users' usage due to the congestion effect.

Fig. 4.13 illustrates the optimal revenue achieved at the optimal price as the number of users increases. As expected, we can have similar observations as for Fig. 4.11: when social network effect dominates congestion effect, the marginal gain of optimal revenue by taking in more users is increasing; otherwise, the marginal gain is decreasing.

Fig. 4.14 illustrates the computational complexity of Algorithm 4 as the number of users increases. The number of iterations is equal to the number of price intervals that determine the piece-wise structure of total usage and revenue as a function of price. We observe that the complexity is $O(N)$.

4.5.3 Further Discussions

Based on the simulation results, we can draw the following engineering insights for the operation of wireless providers.

- The observations from Figs. 4.8-4.10 suggest that as users' social ties become stronger (which can be promoted by social services), the wireless provider can receive an increasing total usage and thus revenue, and also an increasing marginal gain. In addition, the wireless provider can also receive an increasing marginal return by incorporating more resources for the wireless service to mitigate congestion.

- The observations from Figs. 4.11 and 4.13 suggest that the wireless provider should be aware of whether the social network effect determined by users' social ties dominates the congestion effect. If the social network effect dominates, it receives an increasing marginal gain by taking in more users; otherwise, the marginal gain is decreasing and the total usage will saturate when the number of users is sufficiently large.

- The observations from Fig. 4.12 suggest that the wireless provider should set a low price when users' social ties are strong (evidenced by the popularity of social services), as the decrease of price will be outweighed by the increase of total usage resulted from the social network effect, so that the total revenue increases. Otherwise, the wireless provider should set a high price, as cutting the price cannot stimulate sufficiently more usage due to the congestion effect to compensate the price decrease.

4.6 Related Work

There have been many studies on users' behaviors and the provider's pricing strategy when either network effect (also known as positive externality) or congestion effect is present, respectively [58,62,63]. In [58], different pricing strategies of a provider have been studied where users' behaviors are only subject to network effect. When users experience both network effect and congestion effect as considered in this chapter, the coupling among users is very different and more complex than when only network effect is present as in [58]. Very few work have studied the case where both network effect and congestion effect are present. [64] has studied users' behaviors when they experience both network effect and congestion effect. However, it assumes that the network effect is the same for all users, which does not capture the fact that users experience different levels of network effect based on their diverse social ties as considered in this chapter.

The social aspect of mobile networking is an emerging paradigm for network design and optimization. Social contact patterns have been exploited for efficient data forwarding and dissemination in delay tolerant networks [65, 66]. Social trust and social reciprocity have been leveraged in [16] to enhance cooperative D2D communication based on a coalitional game. A social group utility maximization (SGUM)

framework has been recently studied in [51,52,67], which captures the impact of mobile users' diverse social ties on the interactions of their mobile devices subject to diverse physical relationships.

4.7 Conclusion

In this chapter, we have formulated the interaction between mobile users and a wireless provider as a Stackelberg game, by jointly considering the social effect in the social domain and the congestion effect in the physical wireless domain. For Stage II, we have analyzed users' demand equilibrium given a price chosen by the wireless provider. For Stage I, we have developed an algorithm to compute the optimal price to maximize the wireless provider's revenue. We have also conducted simulations using real data to evaluate the performance, and drawn useful engineering insights for the operation of wireless providers.

Appendix

Proof of Theorem 4.1

To show the existence of UDE, we make use of the following lemma, which shows that the Stage II game with unbounded usage range is equivalent to that with bounded usage range.

Lemma 4.2. *Under Assumption 1, the Stage II game* $\mathcal{G} \triangleq \{\mathcal{N}, \{u_i\}_{i\in\mathcal{N}}, [0,\infty)^N\}$ *admits the same set of UDEs as the game* $\mathcal{G}' \triangleq \{\mathcal{N}, \{u_i\}_{i\in\mathcal{N}}, [0,\bar{x}]^N\}$*, where* $\bar{x}$ *is any number that satisfies* $\bar{x} > \max_{i\in\mathcal{N}} |a_i - p|/(b_i + c - \sum_{j\neq i} |g_{ij} - c|)$.

Proof: Let $\mathbf{x}^*$ be any UDE of game $\mathcal{G}$ and x_i^* be the largest in $\mathbf{x}^*$, i.e., $x_i^* \geq x_j^*$, $\forall i \neq j$. If $x_i^* > 0$, using the best response function (4.2), we have

$$x_i^* = \frac{a_i - p}{b_i + c} + \sum_{j\neq i} \frac{g_{ij} - c}{b_i + c} x_j^* \leq \frac{|a_i - p|}{b_i + c} + \sum_{j\neq i} \frac{|g_{ij} - c|}{b_i + c} x_i^*. \tag{4.5}$$

Using Assumption 1, it follows from (4.5) that

$$x_i^* \leq |a_i - p|/(b_i + c - \sum_{j \neq i} |g_{ij} - c|) < \bar{x}.$$

Since x_i^* is the largest in $\mathbf{x}^*$, we have $x_j^* \in [0, \bar{x}]$, $\forall j \in \mathcal{N}$, and thus $\mathbf{x}^* \in [0, \bar{x}]^N$. Therefore, as game $\mathcal{G}$ and game $\mathcal{G}'$ have the same set of payoff functions and the strategy spaces in both games contain $[0, \bar{x}]^N$, they have the same set of UDEs. $\square$

Using a celebrated result in [68–70], the infinite game $\mathcal{G}'$ admits a UDE if the strategy space $[0, \bar{x}]^N$ is compact and convex, the payoff function $u_i(x_i, \mathbf{x}_{-i})$ is continuous in x_i and $\mathbf{x}_{-i}$, and the payoff function $u_i(x_i, \mathbf{x}_{-i})$ is concave in x_i. It is easy to check that all these conditions hold, and thus the game $\mathcal{G}'$ admits a UDE. Then it follows from Lemma 4.2 that the Stage II game $\mathcal{G}$ admits a UDE.

Proof of Theorem 4.2

We will show that the UDE is unique by showing that the game $\mathcal{G}'$ defined in Lemma 4.2 is a concave game. The Jacobian matrix $\nabla u(\mathbf{x})$ of the payoff function profile $u(\mathbf{x}) \triangleq (u_1(\mathbf{x}), \ldots, u_N(\mathbf{x}))$ of game $\mathcal{G}'$ is given by

$$\begin{aligned}
\nabla u(\mathbf{x}) &= \begin{bmatrix} \frac{\partial^2 u_1(\mathbf{x})}{\partial x_1^2} & \frac{\partial^2 u_1(\mathbf{x})}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 u_1(\mathbf{x})}{\partial x_1 \partial x_N} \\ \frac{\partial^2 u_2(\mathbf{x})}{\partial x_2 \partial x_1} & \frac{\partial^2 u_2(\mathbf{x})}{\partial x_2^2} & \cdots & \frac{\partial^2 u_2(\mathbf{x})}{\partial x_2 \partial x_N} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 u_N(\mathbf{x})}{\partial x_N \partial x_1} & \frac{\partial^2 u_N(\mathbf{x})}{\partial x_N \partial x_2} & \cdots & \frac{\partial^2 u_N(\mathbf{x})}{\partial x_N^2} \end{bmatrix} \\
&= \begin{bmatrix} -b_1 - c & g_{12} - c & \cdots & g_{1N} - c \\ g_{21} - c & -b_2 - c & \cdots & g_{2N} - c \\ \vdots & \vdots & \ddots & \vdots \\ g_{N1} - c & g_{N2} - c & \cdots & -b_N - c \end{bmatrix} \\
&= -(B - G).
\end{aligned}$$

Using Assumption 1, it follows that

$$[B-G]_{ii} \geq \sum_{j\neq i} |[B-G]_{ij}|, \ \forall i$$

where $[M]_{ij}$ denotes the entry in the ith row and jth column of matrix M. Therefore, $B-G$ is strictly diagonal dominant [71]. It follows from the condition $\sum_{j\neq i} \frac{|g_{ji}-c|}{b_i+c} < 1$, $\forall i$ that $(B-G)^T$ is also strictly diagonal dominant. Then we have that

$$\nabla u(\mathbf{x}) + \nabla u(\mathbf{x})^T = -(B-G) - (B-G)^T$$

is strictly diagonal dominant. Also observe that it is symmetric. It is known that a symmetric matrix that is strictly diagonally dominant with real nonnegative diagonal entries is positive definite [71]. Therefore, $\nabla u(\mathbf{x}) + \nabla u(\mathbf{x})^T$ is negative definite. It follows from [60, Theorem 6] that $u(\mathbf{x})$ is diagonally strictly concave. Therefore, using [60, Theorem 2], game $\mathcal{G}'$ has a unique UDE.

Proof of Theorem 4.3

Let $\Delta x_i^{(t)} \triangleq x_i^{(t)} - x_i^*, \forall i$. For any $i \in \mathcal{N}$, according to step 6 in Algorithm 2, we have

$$|\Delta x_i^{(t+1)}| \leq |\sum_{j\neq i} \frac{g_{ij}-c}{b_i+c} \Delta x_j^{(t)}| \leq \sum_{j\neq i} \frac{|g_{ij}-c|}{b_i+c} |\Delta x_j^{(t)}|. \tag{4.6}$$

Let $\|\Delta \mathbf{x}^{(t)}\|_\infty$ be the l_∞-norm of vector $(\Delta x_i^{(t)}, \ldots, \Delta x_N^{(t)})$, i.e.,

$$\|\Delta \mathbf{x}^{(t)}\|_\infty \triangleq \max_{i\in\mathcal{N}} |\Delta x_i^{(t)}|.$$

Then, using Assumption 1 and (4.6), we have

$$\begin{aligned}
\|\Delta \mathbf{x}^{(t+1)}\|_\infty &\leq \max_{i\in\mathcal{N}} \left(\sum_{j\neq i} \frac{|g_{ij}-c|}{b_i+c} |\Delta x_j^{(t)}| \right) \\
&\leq \left(\max_{i\in\mathcal{N}} \sum_{j\neq i} \frac{|g_{ij}-c|}{b_i+c} \right) \left(\max_{i\in\mathcal{N}} |\Delta x_j^{(t)}| \right) \\
&= \left(\max_{i\in\mathcal{N}} \sum_{j\neq i} \frac{|g_{ij}-c|}{b_i+c} \right) \|\Delta \mathbf{x}^{(t)}\|_\infty.
\end{aligned}$$

According to Assumption 1, we have $\left(\max_{i\in\mathcal{N}} \sum_{j\neq i} \frac{|g_{ij}-c|}{b_i+c}\right)$ < 1. Then it follows that the algorithm results in a contraction mapping of $|\Delta x_i^{(t)}|$, and thus converges to the UDE.

Proof of Proposition 4.2

If the UDE is positive, i.e., $x_1^* > 0$ and $x_2^* > 0$, according to 4.2, we have $\mathbf{x}^* > 0$ is the solution to

$$x_1 = \frac{a_1 - p}{b_1 + c} + \frac{g_{12} - c}{b_1 + c} x_2, \; x_2 = \frac{a_2 - p}{b_2 + c} + \frac{g_{21} - c}{b_2 + c} x_1.$$

Solving it, we have the expression given in the low price regime. Then observe that x_1^* and x_2^* are both positive when $p = 0$, and decrease when p increases. Also observe that $x_1^* = 0$ when $p = p_1 \triangleq \frac{a_1(b_2+c)-a_2(c-g_{12})}{b_2+g_{12}}$, and $x_2^* = 0$ when $p = p_2 \triangleq \frac{a_2(b_1+c)-a_1(c-g_{21})}{b_1+g_{21}}$. We can check that $p_1 \geq p_2$. Therefore, when $p > p_2 = p_{th}$, we have $x_1^* > 0$ and $x_2^* = 0$. Thus $x_1^* = \frac{a_1-p}{b_1+c}$ according to (4.2). Then we further observe that $x_1^* = x_2^* = 0$ when $p > a_1$.

Proof of Lemma 4.1

We only prove the case when $\mathcal{S} = \mathcal{N}$, since the case when $\mathcal{S} \subset \mathcal{N}$ can be proved similarly. Let

$$\bar{B} = \begin{bmatrix} b_1 + c & 0 & \cdots & 0 \\ 0 & b_2 + c & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & b_N + c \end{bmatrix},$$

$$\bar{G} = \begin{bmatrix} 0 & g_{12} - c & \cdots & g_{1N} - c \\ g_{21} - c & 0 & \cdots & g_{2N} - c \\ \vdots & \vdots & \ddots & \vdots \\ g_{N1} - c & g_{N2} - c & \cdots & 0 \end{bmatrix}.$$

Since $\bar{B}$ is a diagonal matrix with positive diagonal entries, it is invertible. Let λ be any eigenvalue of $\bar{B}^{-1}\bar{G}$ with v being the corresponding eigenvector. Let v_i be the largest entry of v in absolute value, i.e., $|v_i| \geq |v_j|$, $\forall j$. Since $(\bar{B}^{-1}\bar{G})v = \lambda v$, it

follows that

$$|\lambda v_i| = |[\bar{B}^{-1}\bar{G}]_i v| \leq \sum_{j\in\mathcal{N}} |[\bar{B}^{-1}\bar{G}]_{ij}||v_j| \leq |v_i| \sum_{j\in\mathcal{N}} \frac{|g_{ij}-c|}{b_i+c} < |v_i|$$

where the last inequality follows from Assumption 1. It follows that the spectral radius of $\bar{B}^{-1}\bar{G}$ is strictly less than 1. Since each eigenvalue of $I - \bar{B}^{-1}\bar{G}$ is equal to $1-\lambda$ where λ is an eigenvalue of $\bar{B}^{-1}\bar{G}$, where I denotes the $N \times N$ identity matrix, it follows that $I - \bar{B}^{-1}\bar{G}$ has no eigenvalue of 0, and thus is invertible. Thus $B - G = \bar{B} - \bar{G} = \bar{B}(I - \bar{B}^{-1}\bar{G})$ is also invertible.

Proof of Proposition 4.3

We first show part 1). Suppose $p > a$ and $x_i^* > 0$ is the largest in $\mathbf{x}^*$, i.e., $x_i^* \geq x_j^*$, $\forall i \neq j$. Using the best response function (4.2), we have

$$x_i^* = \frac{a-p}{b_i+c} + \sum_{j\neq i} \frac{g_{ij}-c}{b_i+c} x_j^* \leq \sum_{j\neq i} \frac{|g_{ij}-c|}{b_i+c} x_i^* < x_i^*$$

where the last inequality follows from Assumption 1. This shows a contradiction. Thus we have $x_i^* = 0$, $\forall i$.

Next we show part 2). Let $\mathcal{S}$ be the set of users with positive usage in $\mathbf{x}^*$ at price 0. For any $i \notin \mathcal{S}$, using (4.2), we have

$$x_i^* = 0 \geq \frac{a}{b_i+c} + \frac{a}{b_i+c}[G-C]_{i,\mathcal{S}}(B_{\mathcal{S}} - G_{\mathcal{S}})^{-1}\mathbf{1}_{\mathcal{S}}. \tag{4.7}$$

For any $p \in (0, a]$, we next show that $\mathbf{x}'$ with $\mathbf{x}'_{\mathcal{S}} = (a-p)(B_{\mathcal{S}} - G_{\mathcal{S}})^{-1}\mathbf{1}_{\mathcal{S}}$ and $x_i' = 0$, $\forall i \notin \mathcal{S}$ is the UDE at price p. We observe that for any $i \in \mathcal{S}$, x_i' is its best response at $\mathbf{x}'$. For any $i \notin \mathcal{S}$, using (4.7), we have

$$\frac{a-p}{b_i+c} + \frac{a-p}{b_i+c}[G-C]_{i,\mathcal{S}}(B_{\mathcal{S}} - G_{\mathcal{S}})^{-1}\mathbf{1}_{\mathcal{S}} \leq 0 = x_i^*,$$

and thus is user i's best response at $\mathbf{x}'$.

Proof of Proposition 4.4

The proof of part 1) is the same as the proof of part 1) of Proposition 4.3 except that a should change to $\max_{i\in\mathcal{N}} a_i$. Now we show part 2). For any price $p \in [0, \max_{i\in\mathcal{N}} a_i]$, the usage of the set of users $\mathcal{S}$ with positive usage at the UDE (if they exist) is given by (4.4). Observe that the usage demand $\frac{a_i-p}{b_i+c} + \sum_{j\neq i} \frac{g_{ij}-c}{b_i+c} x_j^*$ of any user i at the UDE is a continuous function of price p and other users' usage x_j^*. Therefore, when the price p increases by a sufficiently small amount to p', the set of user with positive usage at the UDE is still the set $\mathcal{S}$, and thus their usage is still given by (4.4) except with p replaced by p'. Therefore, the set of user with positive usage is the same at any price in a continuous price interval. Then the desired result follows.

Proof of Proposition 4.5

We only prove the case when any g_{ij} increases, since the cases when a increases, any b_i decreases, or c increases can be proved similarly. Then it suffices to prove the case when any g_{ij} increases by any small amount. Let G be a symmetric matrix. Let $\mathbf{x}^*$ be the UDE under G and $\mathcal{S}$ be the set of users with positive usage in $\mathbf{x}^*$. It is easy to check that the UDE is a continuous function of the matrix G. Then we can always find a symmetric matrix G' with $[G']_{ij} \geq [G]_{ij}$, $\forall i, j$ and at least one strict inequality, such that the set of users with positive usage at the UDE $\mathbf{x}'$ under G' is also $\mathcal{S}$. Therefore, using the best response functions (4.2), we have

$$B_{\mathcal{S}}\mathbf{x}^*_{\mathcal{S}} = (a-p)\mathbf{1}_{\mathcal{S}} + G_{\mathcal{S}}\mathbf{x}^*_{\mathcal{S}} \tag{4.8}$$

$$B_{\mathcal{S}}\mathbf{x}'_{\mathcal{S}} = (a-p)\mathbf{1}_{\mathcal{S}} + G'_{\mathcal{S}}\mathbf{x}'_{\mathcal{S}}. \tag{4.9}$$

Subtracting (4.8) from (4.9), we have

$$B_{\mathcal{S}}(\mathbf{x}'_{\mathcal{S}} - \mathbf{x}^*_{\mathcal{S}}) = G_{\mathcal{S}}(\mathbf{x}'_{\mathcal{S}} - \mathbf{x}^*_{\mathcal{S}}) + \Delta G_{\mathcal{S}}\mathbf{x}'_{\mathcal{S}} \tag{4.10}$$

where $\Delta G_{\mathcal{S}} \triangleq G'_{\mathcal{S}} - G_{\mathcal{S}}$. According to Lemma 4.1, $B_{\mathcal{S}} - G_{\mathcal{S}}$ is invertible. Then it follows from (4.10) that

$$\mathbf{x}'_{\mathcal{S}} - \mathbf{x}^*_{\mathcal{S}} = (B_{\mathcal{S}} - G_{\mathcal{S}})^{-1} \Delta G_{\mathcal{S}} \mathbf{x}'_{\mathcal{S}}. \tag{4.11}$$

On the other hand, it follows from (4.8) that

$$\mathbf{x}^*_{\mathcal{S}} = (a - p)(B_{\mathcal{S}} - G_{\mathcal{S}})^{-1} \mathbf{1}_{\mathcal{S}}. \tag{4.12}$$

Using (4.11) and (4.12), we have

$$\begin{aligned} t(\mathbf{x}') - t(\mathbf{x}^*) &= \mathbf{1}_{\mathcal{S}}^T (\mathbf{x}'_{\mathcal{S}} - \mathbf{x}^*_{\mathcal{S}}) \\ &= \mathbf{1}_{\mathcal{S}}^T (B_{\mathcal{S}} - G_{\mathcal{S}})^{-1} \Delta G_{\mathcal{S}} \mathbf{x}'_{\mathcal{S}} \\ &= [(B_{\mathcal{S}} - G_{\mathcal{S}})^{-1} \mathbf{1}_{\mathcal{S}}]^T \Delta G_{\mathcal{S}} \mathbf{x}'_{\mathcal{S}} \\ &= \frac{1}{a - p} (\mathbf{x}^*_{\mathcal{S}})^T \Delta G_{\mathcal{S}} \mathbf{x}'_{\mathcal{S}} \end{aligned}$$

where the third equality is due to the fact that $(B_{\mathcal{S}} - G_{\mathcal{S}})^{-1}$ is symmetric since $B_{\mathcal{S}} - G_{\mathcal{S}}$ is symmetric. Since $a > p$ and $\mathbf{x}^*_{\mathcal{S}}$, ΔG, $\mathbf{x}'_{\mathcal{S}}$ only have nonnegative entries, it follows that $t(\mathbf{x}') \geq t(\mathbf{x}^*)$.

Proof of Theorem 4.4

We will show that given the current critical price $\underline{p}$ and the set of users $\mathcal{S}$ with positive usage at the UDE at the price $\underline{p}$, each iteration from step 4 to step 39 finds the next critical price $\overline{p}$, and the optimal price and revenue in the price interval $[\underline{p}, \overline{p}]$. For any $i \in \mathcal{S}$ with $x_i^* > 0$, it follows from (4.4) that

$$x_i^* = [(B_{\mathcal{S}} - G_{\mathcal{S}})^{-1}]_i \mathbf{a}_{\mathcal{S}} - p[(B_{\mathcal{S}} - G_{\mathcal{S}})^{-1}]_i \mathbf{1}_{\mathcal{S}} > 0. \tag{4.13}$$

Therefore, x_i^* decreases when the price p increases if

$$[(B_{\mathcal{S}} - G_{\mathcal{S}})^{-1}]_i \mathbf{1}_{\mathcal{S}} > 0. \tag{4.14}$$

For any $i \notin \mathcal{S}$ with $x_i^* = 0$, it follows from (4.2) that the usage demand of user i is no greater than 0 such that

$$\begin{aligned} 0 &\geq [G-C]_{i,\mathcal{S}}(B_{\mathcal{S}} - G_{\mathcal{S}})^{-1}(\mathbf{a}_{\mathcal{S}} - p\mathbf{1}_{\mathcal{S}}) + a_i - p \\ &= [G-C]_{i,\mathcal{S}}(B_{\mathcal{S}} - G_{\mathcal{S}})^{-1}\mathbf{a}_{\mathcal{S}} + a_i - p([G-C]_{i,\mathcal{S}}(B_{\mathcal{S}} - G_{\mathcal{S}})^{-1}\mathbf{1}_{\mathcal{S}} + 1) \end{aligned} \tag{4.15}$$

Therefore, the usage demand of user $i \notin \mathcal{S}$ increases when the price p increases if

$$[G-C]_{i,\mathcal{S}}(B_{\mathcal{S}} - G_{\mathcal{S}})^{-1}\mathbf{1}_{\mathcal{S}} < -1. \tag{4.16}$$

Using (4.13), the price at which a user $i \in \mathcal{S}$ changes its usage from positive to 0 is

$$\tilde{p_i} \triangleq \frac{[(B_{\mathcal{S}} - G_{\mathcal{S}})^{-1}]_i \mathbf{a}_{\mathcal{S}}}{[(B_{\mathcal{S}} - G_{\mathcal{S}})^{-1}]_i \mathbf{1}_{\mathcal{S}}}$$

where $\mathcal{S}'$ is the set of users such that (4.14) holds. Using (4.15), the price at which a user $i \notin \mathcal{S}$ changes its usage from 0 to positive is

$$\tilde{p_i} \triangleq \frac{[G-C]_{i,\mathcal{S}}(B_{\mathcal{S}} - G_{\mathcal{S}})^{-1}\mathbf{a}_{\mathcal{S}} + a_i}{[G-C]_{i,\mathcal{S}}(B_{\mathcal{S}} - G_{\mathcal{S}})^{-1}\mathbf{1}_{\mathcal{S}} + 1}$$

where $\mathcal{S}''$ is the set of users such that (4.16) holds. Therefore, the next critical price $\overline{p}$ is

$$\overline{p} = \min_{i \in \mathcal{S}' \cup \mathcal{S}''} \tilde{p}_i$$

Using (4.4), the revenue R is given by

$$R(p) = p\mathbf{1}_{\mathcal{S}}^T \mathbf{x}_{\mathcal{S}}^* = p\mathbf{1}_{\mathcal{S}}^T (B_{\mathcal{S}} - G_{\mathcal{S}})^{-1}(\mathbf{a}_{\mathcal{S}} - p\mathbf{1}_{\mathcal{S}})$$

which is a concave quadratic function of p. By setting $\frac{\partial R(p)}{\partial p} = 0$, we obtain that the optimal price p' in the price interval $[\underline{p}, \overline{p}]$ that maximizes the revenue is

$$p' = \hat{p} \triangleq \frac{\mathbf{1}_{\mathcal{S}}^T (B_{\mathcal{S}} - G_{\mathcal{S}})^{-1}\mathbf{a}_{\mathcal{S}}}{2\mathbf{1}_{\mathcal{S}}^T (B_{\mathcal{S}} - G_{\mathcal{S}})^{-1}\mathbf{1}_{\mathcal{S}}} \tag{4.17}$$

if $\hat{p} \in [\underline{p}, \overline{p}]$. If $\hat{p} \notin [\underline{p}, \overline{p}]$, the optimal price is $p' = \underline{p}$ if $\hat{p} < \underline{p}$, or $p' = \overline{p}$ if $\hat{p} > \overline{p}$. Thus the optimal revenue r' in the price interval $[\underline{p}, \overline{p}]$ is

$$r' = p'\mathbf{1}_{\mathcal{S}}^T (B_{\mathcal{S}} - G_{\mathcal{S}})^{-1}(\mathbf{a}_{\mathcal{S}} - p'\mathbf{1}_{\mathcal{S}}).$$

Then the optimal price and revenue in the entire range $[0, \max_{i \in \mathcal{N}} a_i]$ is found by comparing the optimal revenue for all the iterations from step 4 to step 39.

Chapter 5

DISTRIBUTED OPPORTUNISTIC SCHEDULING FOR COOPERATIVE NETWORKING

5.1 Introduction

There has been extensive research on channel-aware scheduling, which exploits the rich diversities inherent in wireless communications by scheduling links with favorable channel conditions [5, 6]. While most existing studies focus on centralized scheduling (see, e.g., [5–7]), a distributed opportunistic scheduling (DOS) framework is developed for ad hoc networks in [8], in which multi-user diversity and time diversity in wireless channels are exploited jointly in a distributed manner. Assuming perfect channel estimation, it is shown that the optimal opportunistic scheduling in such a scenario is intimately related to the fundamental tradeoff between throughput gain from better channel conditions and the cost of further probing.

Along a different avenue of research, cooperative communication has recently been studied as a promising technique for improving communication efficiency, because it can achieve spatial diversity without equipping individual nodes with multiple antennas. For instance, the severe fading in the direct link between a source and a destination can be overcome by employing a relay to assist the transmission. Two popular mechanisms among the existing relaying techniques are decode-and-forward (DF) and compress-and-forward (CF) [72–74]. Simply put, with DF relaying, the relay decodes and re-encodes the received signal from the source before forwarding it to the destination. In contrast, with CF relaying, the relay first compresses the received signal from the source and then forwards it to the destination. While most of the existing studies on cooperative communication focus on the physical (PHY) layer, notably a few works have explored cooperation at the medium access control (MAC)

layer and above (see, e.g., [75–77]). Nevertheless, cross-layer design for cooperative networking is still not well understood.

Clearly, cooperative relaying can be of great value for enhancing ad hoc communications via exploiting spatial diversity. Thus motivated, a primary goal of this study is to obtain a rigorous understanding of cooperative networking for ad hoc communications from a holistic perspective. More specifically, in an ad hoc network based on random access, even if a user with successful contention observes a poor condition in the direct link (source-destination link), it is possible that the relay channel, consisting of source-destination (S-D), source-relay (S-R) and relay-destination (R-D) links, can support a much higher transmission rate. What remains to be quantified in this line of inquiry, however, is the additional overhead needed to probe a relay for possible cooperative transmission. From the point view of opportunistic scheduling, channel state information (CSI) of the relay channel is crucial for the successful user to know the achievable rate of cooperative transmission, which then serves as the basis for it to decide if it should pursue cooperative transmission. Intuitively speaking, if the overhead for probing the relay is overwhelming, the rate gain by conducting cooperative transmission may not be worthwhile. In this chapter, we shall study the tradeoff between throughput gain from cooperative relaying and the probing cost for establishing cooperative relaying. In particular, we seek to answer the following questions: How should the successful user probe for cooperative relaying? How should it decide whether to probe for cooperative relaying?

To this end, we cast the problem of opportunistic cooperative networking (OCN) as an optimal stopping problem with two-levels of incomplete information [78]. Specifically, we consider two cases: i) OCN with dedicated relays (OCN-DR), and ii) OCN without dedicated relays (OCN-NDR). We show that the optimal strategies exist for both cases. In particular, our investigation yields the following interesting

findings: For OCN-DR, it is optimal for the successful user to probe the relay when the signal-to-noise ratio (SNR) of the S-R link is higher than some threshold; for OCN-NDR, under more restrictive conditions (to be clarified in Section 5.5), it is optimal for the successful user to probe potential relays when the SNR of the S-D link lies between two thresholds. Furthermore, the thresholds in the optimal strategy can be computed based on statistical information, making it amenable to distributed implementation.

Before proceeding further, we emphasize the main contributions distinguishing this work from other existing related works. It is clear that this work goes beyond [8] by opening up the possibility of cooperative networking under the DOS framework. Recently, DOS with imperfect channel estimation has been investigated in [79]. Despite one common feature that both this work and [79] cast DOS as an optimal stopping problem with incomplete information, this work concentrates on exploiting cooperative networking to reap multi-user diversity from a global perspective, whereas [79] aims to improve transmission rates by refining local channel estimation.

The rest of this chapter is organized as follows. Section 5.2 gives the background and the system model for opportunistic cooperative networking. The problem of opportunistic cooperative networking is formulated in Section 5.3. We then study, in Section 5.4 and Section 5.5, the optimal strategies for OCN-DR and OCN-NDR. Section 5.7 concludes this chapter.

5.2 Background and System Model

5.2.1 Opportunistic Cooperative Networking

Consider a single-hop ad hoc network in which L users (source-destination node pairs) contend for a channel using random access. A slotted contention model is assumed for the random access, in which each user sends a probing packet with a pre-defined contention probability in a slot. At the end of a contention slot, each user learns from

its received packet if there was a successful contention, a collision, or the channel was idle. A channel contention of a user is said to be successful if no other users contend in the same slot. Then, the overall successful channel contention probability p_s is given by $p_s = \sum_{l=1}^{L}(p_l \prod_{i \neq l}(1 - p_i))$, where p_l is the probability that user l contends for the channel. The random duration of achieving one successful channel contention is called one channel probing round for brevity. It can be seen that the number of slots in a channel probing round, denoted as K, is a geometric random variable with parameter p_s. Let τ denote the duration of one slot. Then the random duration of a channel probing round is $K\tau$, the expectation of which is τ/p_s.

Since the transmission rate of a successful user depends on the time-varying wireless channel condition(s), it is random. We assume that after a channel probing round, the channel is available to the successful user for a fixed duration of T, during which the channel condition(s) remains constant. Depending on the observed CSI, the successful user may decide to transmit data via the direct link (S-D link) for a duration of T, or may skip the transmission opportunity and let all users recontend. For ease of exposition, we assume that the CSI obtained by channel probing is error free.

With the potential rate gain due to cooperative communication, the successful user may seek the help of a relay, in the hope of achieving a higher rate with cooperative relaying (consisting of the S-D, S-R and R-D links), besides the options of transmitting directly or skipping transmission. However, in order to carry out cooperative communications, the user has to know the full CSI of the relay channel, which would not be available after probing the direct link only. That is to say, to obtain the full CSI, the user has to probe the relay channel, say for a duration of τ', which would however reduce the duration for data transmission to $T - \tau'$. It is worth noting that even if the cooperative transmission achieves a higher rate than the direct

transmission, it is still possible for the user to skip the transmission opportunity.

For the successful user after the nth round of channel probing, let $R_{d,n}$ denote its direct transmission rate, and $R_{m,n}$ denote the maximum of its direct transmission rate and cooperative transmission rate(s), which can only be achieved after probing for cooperative relaying. The successful user makes a *first-level* decision from the following three options:

1) skip transmission and let all users recontend; or

2) probe for cooperative relaying; or

3) transmit data directly (via the S-D link only) at rate $R_{d,n}$ for a duration of T.

Further, under the option 2, the user makes a *second-level* decision from the following two options: transmit data (either directly or with cooperative relaying) at rate $R_{m,n}$ for a duration of $T - \tau'$; or skip transmission and let all users recontend.

5.2.2 Channel Model and Relaying Techniques

We now discuss the channel model briefly. Let ρ denote the normalized SNR, h_{sd} be the channel coefficient and $\gamma_{sd} = \rho|h_{sd}|^2$ be the SNR of the S-D link. For the relay channel, let h_{sr}, and h_{rd} denote the channel coefficients of the S-R, and R-D link, respectively. Also, let γ_{sr} and γ_{rd} denote the corresponding link SNRs, denoted by $\gamma_{sr} = \rho|h_{sr}|^2$ and $\gamma_{rd} = \rho|h_{rd}|^2$. We assume that all links are subject to independent random fading and additive white Gaussian noise (AWGN). For ease of exposition, we assume that all links have identical channel fading statistics with finite means and variances (e.g. Rayleigh fading). Let $G(\cdot)$ and $g(\cdot)$ denote the cumulative distribution function (CDF) and probability distribution function (PDF) of a link channel gain (e.g., $|h_{sd}|^2$, $|h_{sr}|^2$, $|h_{rd}|^2$), respectively. Also let $\overline{G}(\cdot)$ denote the complementary CDF of $G(\cdot)$. When CSI is available to the user, the instantaneous achievable rate of direct

transmission is given by

$$R_d = \log(1 + \gamma_{sd}). \tag{5.1}$$

We briefly discuss two relaying techniques as follows.

Decode-and-forward relaying: Under full duplex DF relaying, the following instantaneous rate is achievable with full CSI [72]:

$$R_r^{\mathrm{DF}} = \max_{0 \le \beta \le 1} \min\{\log(1 + (1-\beta)\gamma_{sr}), \log(1 + \gamma_{sd} + \gamma_{rd} + 2\sqrt{\beta\gamma_{sd}\gamma_{rd}})\}. \tag{5.2}$$

Note that the rate of DF relaying can be less than that of the direct transmission since $R_r^{\mathrm{DF}} < R_d$ when $\gamma_{sr} < \gamma_{sd}$. The rate in (5.2) corresponds to the case of synchronous DF relaying, where it is assumed that the source, the relay and the destination are perfectly phase-synchronized. However, in many practical situations, this assumption may not hold, leading to the case of asynchronous DF relaying [73]. Under asynchronous DF relaying[1], the instantaneous achievable rate is given by

$$R_r^{\mathrm{DF}-} = \min\{\log(1 + \gamma_{sr}), \log(1 + \gamma_{sd} + \gamma_{rd})\} \tag{5.3}$$

which is equal to (5.2) if and only if $\gamma_{sr} \le \gamma_{sd} + \gamma_{rd}$. In light of its practicality, we shall focus on asynchronous DF relaying in Section 5.5.

Compress-and-forward relaying: If full duplex CF relaying is adopted, it yields the following instantaneous rate with full CSI [72]:

$$R_r^{\mathrm{CF}} = \log\left(1 + \gamma_{sd} + \frac{\gamma_{sr}\gamma_{rd}}{1 + \gamma_{sd} + \gamma_{sr} + \gamma_{rd}}\right). \tag{5.4}$$

Note that the rate of CF relaying is always greater than that of the direct transmission.

In this chapter, we focus on how to establish opportunistic cooperative relaying on the fly and on quantifying when it is worthwhile to pursue cooperative relaying. Specifically, we adopt full-duplex relaying which fully exploits the rate gain over direct

[1]In the sequel, we will use "DF relaying" to imply "synchronous DF relaying" for brevity.

transmission. However, all the approaches for the case of full-duplex relaying can be directly carried over to study the case of half-duplex relaying.

5.3 Problem Formulation

In this section, we cast the problem of opportunistic cooperative networking as an optimal stopping problem with incomplete information.

Let ϕ_n and θ_n denote the successful user's first-level and second-level decisions, respectively, after the nth round of channel probing. In particular, $\phi_n = 1$, 2 or 3 refers to the option 1, 2 or 3 described in the last section, correspondingly; under the option 2, $\theta_n = 1$ or 2 means transmitting or skipping transmission, respectively. Let R_n and $T_{d,n}$ denote the transmission rate and transmission duration of the successful user after the nth round of channel probing, respectively, which are then given by $R_n = \mathbf{1}(\phi_n = 3)R_{d,n} + \mathbf{1}(\phi_n = 2)\mathbf{1}(\theta_n = 1)R_{m,n}$[2] and $T_{d,n} = 1 - \mathbf{1}(\phi_n = 2)\mathbf{1}(\theta_n = 1)\tau'$ with $\mathbf{1}(\cdot)$ being the indicator function. Let $T_n = \sum_{i=1}^{n} K_i\tau + 1$ denote the total time of n rounds of channel probing and a transmission duration. The decision sequences ϕ_n and θ_n determine a stopping rule N given by $N = \inf\{n \geq 1 | \phi_n = 3, \text{ or } \phi_n = 2 \text{ and } \theta_n = 1\}$. Our main objective is to find a stopping rule N (i.e., decision sequences) that maximizes the throughput λ of the network. The optimal throughput λ^* and the corresponding optimal stopping rule N^* are defined by

$$\lambda^* \triangleq \sup_{N \in Q} \frac{E[R_N T_{d,N}]}{E[T_N]} \text{ and } N^* \triangleq \arg\sup_{N \in Q} \frac{E[R_N T_{d,N}]}{E[T_N]}$$

where $Q \triangleq \{N : N \geq 1, E[T_N] < \infty\}$.

Let λ denote the cost per unit system time. The optimal expected net reward, given a particular λ, is defined as

$$r(\lambda) \triangleq E[R_N T_{d,N} - \lambda T_N]. \tag{5.5}$$

[2]In the sequel, we will omit the index n in $R_{d,n}$ and $R_{m,n}$ for brevity.

The following lemma relates the optimal throughput λ^* to the optimal expected return.

Lemma 5.1. *The optimal strategy for OCN exists, with*

$$r(\lambda^*) = \sup_{N \in Q} E[R_N T_{d,N} - \lambda^* T_N] = 0$$

$$N^* = arg \sup_{N \in Q} E[R_N T_{d,N} - \lambda^* T_N].$$

The proof is given in Appendix. Let X denote the CSI observed by the successful user based on which it makes the first-level decision. (The particular channel condition(s) that X represents will be specified in Section 5.4 and Section 5.5.) The expected net reward corresponding to the three options 1, 2, and 3 of the first-level decision can be expressed as $r(\lambda)$, $V(X|\lambda, r(\lambda))$, and $R_d - \lambda$, respectively, where $V(X|\lambda, r(\lambda))$ represents the expected reward of probing for cooperative relaying as a function of the observed CSI X given the system cost λ, and is given by

$$V(X|\lambda, r(\lambda)) = r(\lambda) F_{R_m}(\frac{r(\lambda)}{1-\tau'} + \lambda | X) + (1-\tau') \int_{\frac{r(\lambda)}{1-\tau'}+\lambda}^{\infty} (z-\lambda) f_{R_m}(z|X) dz - \lambda\tau' \quad (5.6)$$

with $F_{R_m}(\cdot|X)$ and $f_{R_m}(\cdot|X)$ denoting the CDF and PDF of R_m conditioned on the knowledge of X, respectively. The first and second terms on the right-hand side (RHS) of (5.6) stand for the expected reward of transmitting or skipping transmission after probing for cooperative relaying, respectively.

It follows that the Bellman optimality equation for (5.5) is given by

$$E_X[\max\{r(\lambda), V(X|\lambda, r(\lambda)), R_d(X) - \lambda\}] - \frac{\lambda\tau}{p_s} = r(\lambda) \quad (5.7)$$

where the expectation is taken over X. Using Lemma 1, we have $r(\lambda^*) = 0$, and hence (5.7) can be rewritten as

$$E_X[\max\{0, V(X|\lambda^*, 0), R_d(X) - \lambda^*\}] = \frac{\lambda^*\tau}{p_s}. \quad (5.8)$$

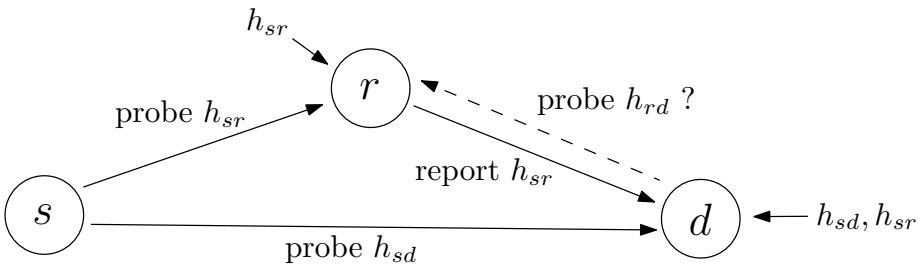


Figure 5.1: Network model for OCN-DR.

In order to find the optimal throughput λ^*, it suffices to characterize the expected reward associated with each option of the first-level decision for any $\lambda > 0$, and then solve the optimality equation (5.8) for $\lambda = \lambda^*$.

5.4 Opportunistic Cooperative Networking with Dedicated Relays

In this section, we consider ad hoc networks in which relays are pre-deployed to assist the transmissions of all users, and furthermore, we assume that each user has one dedicated relay.

5.4.1 Probing the Dedicated Relay

After a successful channel contention, the destination node d of the successful user acquires h_{sd}, based on the probing packet from its source node s. For d to make a well-informed first-level decision, the knowledge of either h_{sr} or h_{rd} is crucial. However, observe that, due to the broadcast nature of the wireless medium, the dedicated relay r overhears the probing packet, and hence it can acquire h_{sr} and report h_{sr} to d, the overhead of which is negligible. Then d can make the first-level decision based on the knowledge of (h_{sd}, h_{sr}). The network model for OCN-DR is illustrated in Fig. 5.1.

When it is optimal to probe the dedicated relay, the successful user needs to probe h_{rd} to make the second-level decision. The user then performs the probing for a fixed duration of τ_1: d sends a probing packet to r to acquire h_{rd}, assuming that the channel is reciprocal. After that, using the full CSI (h_{sd}, h_{sr}, h_{rd}), d computes R_m

and then makes the second-level decision based on R_m.

5.4.2 Optimal Strategy

The successful user makes the first-level decision based on the knowledge of $X = (\gamma_{sd}, \gamma_{sr})$, which is obtained directly from (h_{sd}, h_{sr}). We can see that the transmission rate after probing the dedicated relay is $R_m = \max\{R_d, R_r\}$, where R_r depends on the relaying technique adopted. Let $V_0(\gamma_{sd}, \gamma_{sr}|\lambda) \triangleq V(\gamma_{sd}, \gamma_{sr}|\lambda, 0)$ for brevity.

Then, for the case with DF relaying, $V_0(\gamma_{sd}, \gamma_{sr}|\lambda)$ can be expressed as

$V_0(\gamma_{sd}, \gamma_{sr}|\lambda) =$

$$\begin{cases} (1-\tau_1)[\log(1+\gamma_{sd}) - \lambda]^+ - \lambda\tau_1, & \gamma_{sd} \geq \gamma_{sr} \quad (5.9a) \\ (1-\tau_1)\displaystyle\int_{\log(1+\gamma_{sd})}^{\log(1+\gamma_{sr})} \overline{G}(\gamma_{rd}^{\mathrm{DF}}(\gamma_{sd}, \gamma_{sr}, z)/\rho)dz & \\ \qquad +(1-\tau_1)(\log(1+\gamma_{sd}) - \lambda) - \lambda\tau_1, & \\ & e^{\lambda} - 1 \leq \gamma_{sd} < \gamma_{sr} \quad (5.9b) \\ (1-\tau_1)\displaystyle\int_{\lambda}^{\log(1+\gamma_{sr})} \overline{G}(\gamma_{rd}^{\mathrm{DF}}(\gamma_{sd}, \gamma_{sr}, z)/\rho)dz - \lambda\tau_1, & \\ & \gamma_{sd} < e^{\lambda} - 1 < \gamma_{sr} \quad (5.9c) \\ -\lambda\tau_1, & \gamma_{sd} < \gamma_{sr} \leq e^{\lambda} - 1 \quad (5.9d) \end{cases}$$

where $\gamma_{rd}^{\mathrm{DF}}(\gamma_{sd}, \gamma_{sr}, z)$ is given by

$$\gamma_{rd}^{\mathrm{DF}}(\gamma_{sd}, \gamma_{sr}, z) = \left(\sqrt{(e^z - 1)\left(1 - \frac{\gamma_{sd}}{\gamma_{sr}}\right)} - \sqrt{\left(1 - \frac{e^z - 1}{\gamma_{sr}}\right)\gamma_{sd}}\right)^2 \quad (5.10)$$

and, for the case with CF relaying, $V_0(\gamma_{sd}, \gamma_{sr}|\lambda)$ can be expressed as

$V_0(\gamma_{sd}, \gamma_{sr}|\lambda) =$

$$\begin{cases} (1-\tau_1)\int_{\log(1+\gamma_{sd})}^{\log(1+\gamma_{sd}+\gamma_{sr})} \overline{G}(\gamma_{rd}^{\mathrm{CF}}(\gamma_{sd},\gamma_{sr},z)/\rho)dz \\ \qquad\qquad +(1-\tau_1)(\log(1+\gamma_{sd})-\lambda)-\lambda\tau_1, \\ \qquad\qquad\qquad\qquad e^{\lambda}-1 \le \gamma_{sd} & \text{(5.11a)} \\ (1-\tau_1)\int_{\lambda}^{\log(1+\gamma_{sd}+\gamma_{sr})} \overline{G}(\gamma_{rd}^{\mathrm{CF}}(\gamma_{sd},\gamma_{sr},z)/\rho)dz-\lambda\tau_1, \\ \qquad\qquad\qquad\qquad \gamma_{sd} < e^{\lambda}-1 < \gamma_{sd}+\gamma_{sr} & \text{(5.11b)} \\ -\lambda\tau_1, \qquad\qquad\qquad\qquad \gamma_{sd}+\gamma_{sr} \le e^{\lambda}-1 & \text{(5.11c)} \end{cases}$$

where $\gamma_{rd}^{\mathrm{CF}}(\gamma_{sd},\gamma_{sr},z)$ is given by

$$\gamma_{rd}^{\mathrm{CF}}(\gamma_{sd},\gamma_{sr},z)=\frac{(e^z-\gamma_{sd}-1)(\gamma_{sd}+\gamma_{sr}+1)}{\gamma_{sd}+\gamma_{sr}+1-e^z}. \tag{5.12}$$

The details for deriving (5.9) and (5.11) are given in Appendix.

Next, we examine the structure of the optimal strategy. For convenience, define $W(\gamma_{sd},\gamma_{sr}|\lambda) \triangleq R_d(\gamma_{sd})-\lambda-V_0(\gamma_{sd},\gamma_{sr}|\lambda)$. Let $A_1(\lambda) \triangleq \{(\gamma_{sd},\gamma_{sr}) \in \mathbb{R}_+^2 : R_d(\gamma_{sd})-\lambda \le 0 \text{ and } V_0(\gamma_{sd},\gamma_{sr}|\lambda) \le 0\}$, $A_2(\lambda) \triangleq \{(\gamma_{sd},\gamma_{sr}) \in \mathbb{R}_+^2 : V_0(\gamma_{sd},\gamma_{sr}|\lambda) \ge 0 \text{ and } W(\gamma_{sd},\gamma_{sr}|\lambda) \le 0\}$, and $A_3(\lambda) \triangleq \{(\gamma_{sd},\gamma_{sr}) \in \mathbb{R}_+^2 : R_d(\gamma_{sd})-\lambda \ge 0 \text{ and } W(\gamma_{sd},\gamma_{sr}|\lambda) \ge 0\}$ denote the optimal decision regions of the three options, each standing for the set of channel conditions $(\gamma_{sd},\gamma_{sr})$ under which an option (the option 1, 2, and 3, respectively) achieve the maximum expected reward among the three. Also, define $L_{1,2}(\lambda) \triangleq \{(\gamma_{sd},\gamma_{sr}) \in \mathbb{R}_+^2 : V_0(\gamma_{sd},\gamma_{sr}|\lambda)=0\}$, $L_{1,3}(\lambda) \triangleq \{(\gamma_{sd},\gamma_{sr}) \in \mathbb{R}_+^2 : R_d(\gamma_{sd})-\lambda=0\}$ and $L_{2,3}(\lambda) \triangleq \{(\gamma_{sd},\gamma_{sr}) \in \mathbb{R}_+^2 : W(\gamma_{sd},\gamma_{sr}|\lambda)=0\}$, each standing for the set of channel conditions $(\gamma_{sd},\gamma_{sr})$ under which a pair of options (the option 2 and 3, 1 and 3, 1 and 2, respectively) achieve the same expected reward.

Before we present the main result of this section, we need the following lemmas to characterize the optimal decision regions.

Lemma 5.2. *For any $\lambda > 0$, the optimal decision regions $A_1(\lambda)$, $A_2(\lambda)$, and $A_3(\lambda)$ are nonempty connected sets, and are uniquely characterized by curves $L_{1,2}(\lambda)$, $L_{1,3}(\lambda)$, and $L_{2,3}(\lambda)$ in $\mathbb{R}_+^2$.*

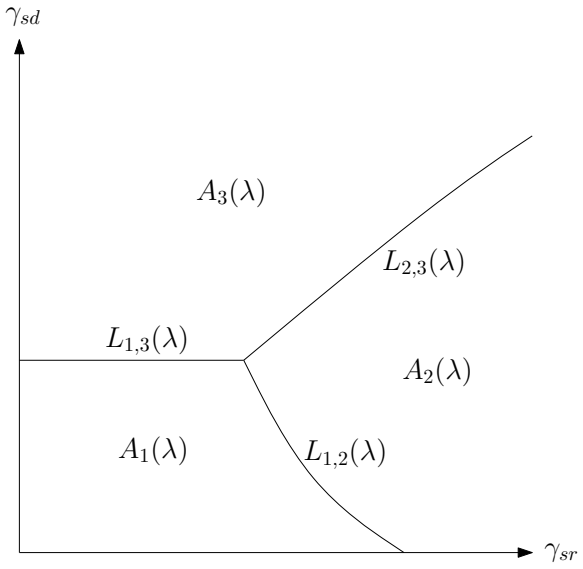


Figure 5.2: Optimal decision regions for OCN-DR.

The proof is given in Appendix. The structures of the optimal decision regions described in Lemma 5.2 are depicted in Fig. 5.2.

The optimal throughput is the unique solution $\lambda = \lambda^*$ to the following equation:

$$E_{(\gamma_{sd},\gamma_{sr})}[V_0(\gamma_{sd},\gamma_{sr}|\lambda);(\gamma_{sd},\gamma_{sr}) \in A_2(\lambda)] + E_{(\gamma_{sd},\gamma_{sr})}[R_d(\gamma_{sd}) - \lambda;(\gamma_{sd},\gamma_{sr}) \in A_3(\lambda)] = \frac{\lambda\tau}{p_s}. \quad (5.13)$$

Indeed, we can see that the left hand side (LHS) of (5.13) is continuously decreasing in λ and is greater than 0 when $\lambda = 0$ and approaches to 0 as $\lambda \to \infty$, while the RHS of (5.13) is continuously increasing in λ and is equal to 0 when $\lambda = 0$. Thus, we conclude that (5.13) has a unique solution.

We can see from the proof of Lemma 5.2 that, for a particular γ_{sd}, if $\gamma_{sd} \geq e^{\lambda^*} - 1$, there is a unique $\gamma^*_{sr,1}(\gamma_{sd})$ such that $(\gamma_{sd}, \gamma^*_{sr,1}(\gamma_{sd})) \in L_{2,3}(\lambda^*)$; if $\gamma_{sd} < e^{\lambda^*} - 1$, there is a unique $\gamma^*_{sr,3}(\gamma_{sd})$ such that $(\gamma_{sd}, \gamma^*_{sr,3}(\gamma_{sd})) \in L_{1,2}(\lambda^*)$. We have the following theorem regarding the structural properties of the optimal strategy.

Theorem 5.1. For OCN-DR with DF relaying or with CF relaying, the optimal strategy takes the following form: after observing γ_{sd} and γ_{sr}, it is optimal for the successful user to

1) skip transmission and let all users recontend if $\gamma_{sd} < e^{\lambda^*} - 1$ and $\gamma_{sr} \leq \gamma^*_{sr,3}(\gamma_{sd})$; or

2) probe the dedicated relay if $\gamma_{sd} \geq e^{\lambda^*} - 1$ and $\gamma_{sr} > \gamma^*_{sr,1}(\gamma_{sd})$, or if $\gamma_{sd} < e^{\lambda^*} - 1$ and $\gamma_{sr} > \gamma^*_{sr,3}(\gamma_{sd})$, and after the probing and obtaining R_m, proceed to transmit if $R_m > \lambda^*$, or skip transmission and let all users recontend otherwise; or

3) transmit directly otherwise.

Proof: This result follows directly from Lemma 5.2 and its proof when $\lambda = \lambda^*$. □

Remarks:

- The above results reveals that the successful user's decision on whether to probe the dedicated relay hinges heavily on γ_{sr}. In particular, it is optimal to probe the dedicated relay if γ_{sr} is above some threshold, where the value of the threshold depends on γ_{sd}. If γ_{sr} is below that threshold, the optimal strategy then depends on γ_{sd}: if γ_{sd} is above a threshold, it is optimal to transmit directly; otherwise, it is optimal to skip transmission. Our intuition is that the S-R link condition determines the potential advantage of probing the dedicated relay over the other options, while the S-D link condition determines the decision between transmitting directly and skipping transmission.

- It is clear that, for the case with DF relaying, if $\gamma_{sr} \leq \gamma_{sd}$, R_m must be no greater than R_d, and hence, in this case, it is never optimal to probe the dedicated relay.

We note that the optimal strategy holds a threshold structure and is amenable to distributed implementation. Indeed, using the knowledge of channel statistics, each user can independently compute off-line λ^* by numerically solving the equation (5.13), and then compute off-line curves $L_{1,2}(\lambda^*)$, $L_{1,3}(\lambda^*)$, and $L_{2,3}(\lambda^*)$ by numerically solving their defining equations. These curves determine the thresholds of the optimal strategy.

5.5 Opportunistic Cooperative Networking without Dedicated Relays

In this section, we consider another interesting network scenario in which users have no dedicated relays and the successful user needs to look for cooperative relaying from other neighboring nodes.

5.5.1 Probing Potential Relays

Suppose that there are $M > 1$ nodes, denoted by $r_1, \cdots, r_M$, each of which can serve as a relay for the successful user upon request. For fair comparison with the case with dedicated relays, we assume that the successful user can select *only one* among the potential relays $r_1, \cdots, r_M$ for cooperation. After a successful channel contention, the destination node d of the successful user acquires h_{sd} using the probing packet sent from its source node s. Also, $r_1, \cdots, r_M$ overhear the probing packet, and acquire the local CSI $h_{sr_1}, \cdots, h_{sr_M}$, respectively. Note that different from the case with dedicated relays, d cannot acquire all of $h_{sr_1}, \cdots, h_{sr_M}$, because the overhead of message passing by a large number of potential relays can be too costly. Hence, d has to make the first-level decision based only on h_{sd}. The network model for OCN-NDR is illustrated in Fig. 5.3.

When probing the potential relays is preferable, the successful user needs to perform the probing with minimum possible overhead. Under the assumption that only one relay is employed in cooperative transmission, the relay CSI of only one

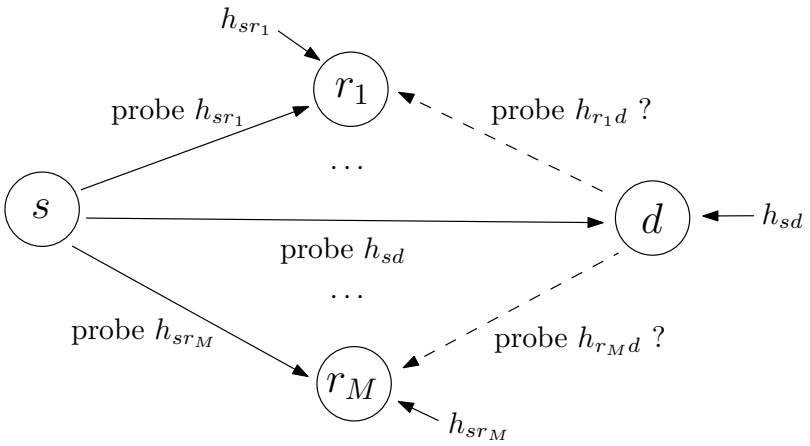


Figure 5.3: Network model for OCN-NDR.

potential relay is required for the user to make the second-level decision. In light of this, we adopt a distributed approach in which only the "best" potential relay feeds back its local CSI to node d. To this end, d first broadcasts h_{sd} and also a probing packet to all potential relays such that each of $\{r_1, \cdots, r_M\}$ acquires h_{sd} and the local observation $h_{r_1d}, \cdots, h_{r_Md}$, respectively. Next, based on the CSI of the relay channel, each potential relay independently estimates its rate of cooperative transmission, denoted by $R_{r_1}, \cdots, R_{r_M}$, respectively. Then the potential relay with the highest estimated rate can report its local CSI to d, e.g., via a backoff-based carrier sensing protocol [80]. Specifically, in this protocol, each potential relay chooses a backoff time based on a pre-determined backoff function that maps an estimated rate to a backoff time, and then it listens to the channel for a duration of a pre-determined maximum backoff time. A potential relay would transmit a beacon signal delayed by its chosen backoff time if and only if no one else transmits before its backoff time expires. A properly designed backoff function ensures that only the potential relay with the highest estimated rate transmits, and then it is detected by node d (see [80] for more details). We denote by τ_2 a *fixed time* for probing the potential relays, which is chosen conservatively a priori such that the destination node is guaranteed to hear the response of the best relay. Clearly, this probing overhead is larger than that in

the case with dedicated relays, i.e., $\tau_2 > \tau_1$. After the probing, d can compute R_m and then make the second-level decision based on R_m.

We emphasize that τ_2 is a fixed time chosen conservatively a priori. If τ_2 is assumed to be random depending on the response time of the relay node with the best channel conditions, there could be some performance gain (mainly) due to the reduced overhead. However, the subsequent analysis would be coupled with the specific design of the backoff function, which would unnecessarily complicate the analysis.

It is worth noting that a relay assists the successful user at the cost of consuming its own resources, and hence requires an incentive for cooperation. Therefore, we introduce $c > 0$ to be the price paid by the user to the selected relay. A plausible price has the same unit as rate such that the cost of employing a relay is proportional to the duration of cooperative relaying. Then the effective rate of cooperative transmission using relay r_i, $i \in \{1, \cdots, M\}$, is $R_{r_i} - c$.

5.5.2 Optimal Strategy

Note that the first-level decision is based only on the knowledge of $X = \gamma_{sd}$, which is obtained directly from h_{sd}. The transmission rate after probing potential relays is $R_m = \max\{R_d, R_{r_1} - c, R_{r_2} - c, \cdots, R_{r_M} - c\}$, where R_{r_i}, $i \in \{1, \cdots, M\}$, depends on the relaying technique adopted. Let $V_0(\gamma_{sd}|\lambda) \triangleq V(\gamma_{sd}|\lambda, 0)$ for brevity.

Then, $V_0(\gamma_{sd}|\lambda)$ can be expressed as

$$V_0(\gamma_{sd}|\lambda) =$$

$$\begin{cases} (1-\tau_2)\displaystyle\int_{\log(1+\gamma_{sd})}^{\infty}(1-F_{R_{r_1}}^{M}(z+c|\gamma_{sd}))dz & \\ \qquad +(1-\tau_2)(\log(1+\gamma_{sd})-\lambda)-\lambda\tau_2, & e^{\lambda}-1 \le \gamma_{sd} \qquad (5.14a) \\ (1-\tau_2)\displaystyle\int_{\lambda}^{\infty}(1-F_{R_{r_1}}^{M}(z+c|\gamma_{sd}))dz-\lambda\tau_2, & \gamma_{sd} < e^{\lambda}-1 \qquad (5.14b) \end{cases}$$

where $F_{R_{r_1}}(z+c|\gamma_{sd})$ for the case with synchronous DF relaying is given by

$$F_{R_{r_1}}(z+c|\gamma_{sd}) = G((e^{z+c}-1)/\rho) + \int_{(e^{z+c}-1)/\rho}^{\infty} g(v)G(\gamma_{rd}^{\mathrm{DF}}(\gamma_{sd},\rho v, z+c)/\rho)dv \quad (5.15)$$

with $\gamma_{rd}^{\mathrm{DF}}(\gamma_{sd},\rho v, z+c)$ given by (5.10), and $F_{R_{r_1}}(z+c|\gamma_{sd})$ for the case with asynchronous DF relaying is given by

$$F_{R_{r_1}}(z+c|\gamma_{sd}) = G((e^{z+c}-1)/\rho) + \overline{G}((e^{z+c}-1)/\rho)G((e^{z+c}-\gamma_{sd}-1)/\rho) \quad (5.16)$$

and $F_{R_{r_1}}(z+c|\gamma_{sd})$ for the case with CF relaying is given by

$$F_{R_{r_1}}(z+c|\gamma_{sd}) = \int_0^{\infty} g(v)G(\gamma_{rd}^{\mathrm{CF}}(\gamma_{sd},\rho v, z+c)/\rho)dv \quad (5.17)$$

with $\gamma_{rd}^{\mathrm{CF}}(\gamma_{sd},\rho v, z+c)$ given by (5.12). The details for deriving (5.14) are given in Appendix.

Next, we investigate the structural properties of the optimal strategy. The optimal strategy for the case with synchronous DF relaying turns out to be difficult for analysis due to the complex expression in (5.2). However, for the cases with asynchronous DF relaying or with CF relaying, the optimal strategy admits a threshold-based structure. In what follows, we present our results for these two cases.

For convenience, define $W(\gamma_{sd}|\lambda) \triangleq R_d(\gamma_{sd}) - \lambda - V_0(\gamma_{sd}|\lambda)$. Let $B_1(\lambda) \triangleq \{\gamma_{sd} \in \mathbb{R}_+ : R_d(\gamma_{sd}) - \lambda \leq 0 \text{ and } V_0(\gamma_{sd},\gamma_{sr}|\lambda) \leq 0\}$, $B_2(\lambda) \triangleq \{\gamma_{sd} \in \mathbb{R}_+ : V_0(\gamma_{sd}|\lambda) \geq 0 \text{ and } W(\gamma_{sd}|\lambda) \leq 0\}$, and $B_3(\lambda) \triangleq \{\gamma_{sd} \in \mathbb{R}_+ : R_d(\gamma_{sd}) - \lambda \geq 0 \text{ and } W(\gamma_{sd}|\lambda) \geq 0\}$ denote the optimal decision regions of the three options, each standing for the set of channel conditions γ_{sd} under which an option (the option 1, 2, and 3, respectively) achieve the maximum expected reward among the three. Also, let $\gamma_{sd,1}(\lambda)$ denote some $\gamma_{sd} \geq e^{\lambda} - 1$ such that $W(\gamma_{sd}|\lambda) = 0$, and $\gamma_{sd,2}(\lambda)$ and $\gamma_{sd,3}(\lambda)$ denote some

$\gamma_{sd} > 0$ such that $R_d(\gamma_{sd}) - \lambda = 0$ and $V_0(\gamma_{sd}|\lambda) = 0$, respectively, each standing for the channel condition γ_{sd} under which a pair of options (the option 2 and 3, 1 and 3, 1 and 2, respectively) achieve the same expected reward.

Before presenting the main result of this section, we give the following lemmas to characterize the optimal decision regions.

Lemma 5.3. *Consider the case with asynchronous DF relaying or with CF relaying. For any $\lambda > 0$, $\gamma_{sd,2}(\lambda)$ exists and is unique; $\gamma_{sd,1}(\lambda)$ and $\gamma_{sd,3}(\lambda)$ may or may not exist, and are unique if they exist. Furthermore,*

I) if $\gamma_{sd,3}(\lambda)$ exists and $\gamma_{sd,2}(\lambda) \leq \gamma_{sd,3}(\lambda)$, then $\gamma_{sd,1}(\lambda)$ does not exist and $B_1(\lambda) = [0, \gamma_{sd,2}(\lambda))$, $B_2(\lambda) = \varnothing$, $B_3(\lambda) = [\gamma_{sd,2}(\lambda), \infty)$;

II) if $\gamma_{sd,3}(\lambda)$ exists and $\gamma_{sd,3}(\lambda) < \gamma_{sd,2}(\lambda)$, then $\gamma_{sd,1}(\lambda)$ exists and $B_1(\lambda) = [0, \gamma_{sd,3}(\lambda))$, $B_2(\lambda) = [\gamma_{sd,3}(\lambda), \gamma_{sd,1}(\lambda))$, $B_3(\lambda) = [\gamma_{sd,1}(\lambda), \infty)$;

III) if $\gamma_{sd,3}(\lambda)$ does not exist, then $\gamma_{sd,1}(\lambda)$ exists and $B_1(\lambda) = \varnothing$, $B_2(\lambda) = [0, \gamma_{sd,1}(\lambda))$, $B_3(\lambda) = [\gamma_{sd,1}(\lambda), \infty)$.

The proof is given in Appendix. The structures of the optimal decision regions described in Lemma 5.3 are depicted in Fig. 5.4, Fig. 5.5 and Fig. 5.6.

The optimal throughput is the unique solution of $\lambda = \lambda^*$ to the following equation:

$$E_{\gamma_{sd}}[V_0(\gamma_{sd}|\lambda); \gamma_{sd} \in B_2(\lambda)] + E_{\gamma_{sd}}[R_d(\gamma_{sd}) - \lambda; \gamma_{sd} \in B_3(\lambda)] = \frac{\lambda\tau}{p_s}. \tag{5.18}$$

Indeed, similarly to (5.13), we can see that there is a unique solution to (5.18).

For notational simplicity, let $\gamma^*_{sd,1} \triangleq \gamma_{sd,1}(\lambda^*)$ and $\gamma^*_{sd,3} \triangleq \gamma_{sd,3}(\lambda^*)$ if they exist. We have the following theorem regarding the structural properties of the optimal strategy.

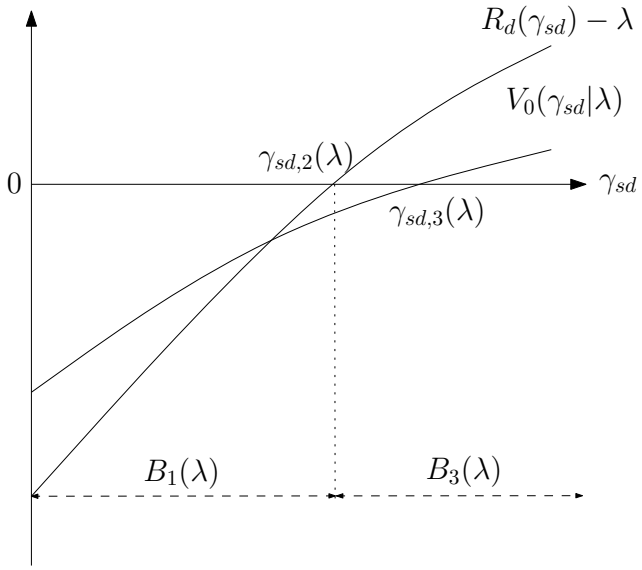


Figure 5.4: Optimal decision regions for OCN-NDR: Case I.

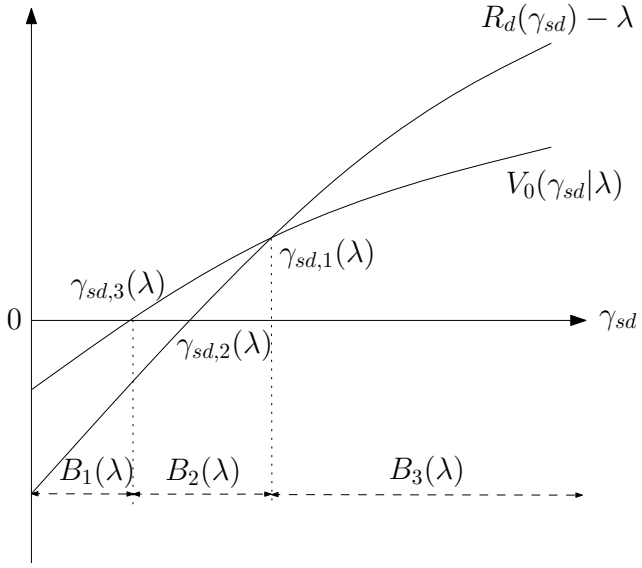


Figure 5.5: Optimal decision regions for OCN-NDR: Case II.

Theorem 5.2. For OCN-NDR with asynchronous DF relaying or with CF relaying, the optimal strategy takes one of the following three forms:

I) If $\gamma^*_{sd,1}$ does not exist and $\gamma^*_{sd,3}$ exists, after observing γ_{sd}, it is optimal for the successful user to transmit directly if $\gamma_{sd} \geq e^{\lambda^*} - 1$; or skip transmission and let all users recontend otherwise.

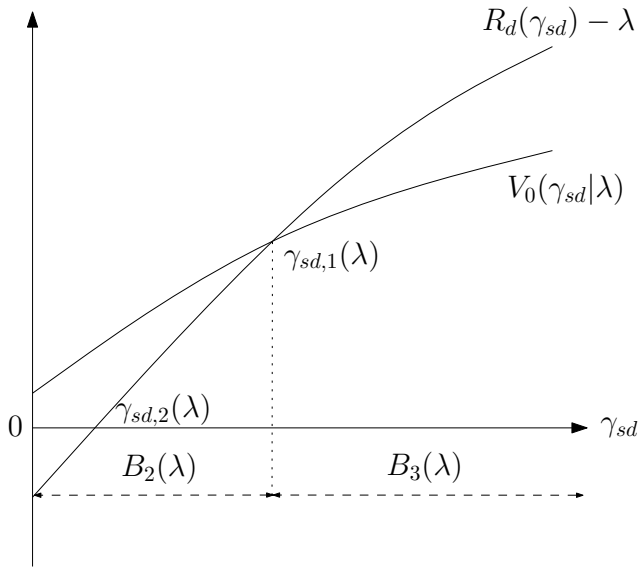


Figure 5.6: Optimal decision regions for OCN-NDR: Case III.

II) If both $\gamma^*_{sd,1}$ and $\gamma^*_{sd,3}$ exist, after observing γ_{sd}, it is optimal for the successful user to transmit directly if $\gamma_{sd} \geq \gamma^*_{sd,1}$; or skip transmission and let all users recontend if $\gamma_{sd} < \gamma^*_{sd,3}$; or probe potential relays otherwise, and after the probing and obtaining R_m, proceed to transmit if $R_m > \lambda^*$; or skip transmission and let all users recontend otherwise.

III) If $\gamma^*_{sd,1}$ exists and $\gamma^*_{sd,3}$ does not exist, after observing γ_{sd}, it is optimal for the successful user to transmit directly if $\gamma_{sd} \geq \gamma^*_{sd,1}$; or probe potential relays otherwise, and after the probing and obtaining R_m, proceed to transmit if $R_m > \lambda^*$; or skip transmission and let all users recontend otherwise.

Proof: This result follows directly from Lemma 5.3 and its proof when $\lambda = \lambda^*$. □

Remarks:

- The above results indicate that for each form of the optimal strategy, if γ_{sd} is above some threshold, it is optimal for the successful user to transmit directly.

The underlying rationale behind this is two-fold: first, the rate gain of R_d with the increase of γ_{sd} is larger than that of $R_r^{\mathrm{DF}-}$ or R_r^{CF}; second, the transmission duration reduces after probing potential relays. These combined makes transmitting directly a better option than probing potential relays when γ_{sd} is large enough.

- We note that the form taken by the optimal strategy hinges on the price of relaying c and the number of potential relays M. Specifically, if the conditions are unfavorable for cooperative relaying, i.e., c is high and M is small, the optimal strategy may take Form I, in which probing potential relays is never optimal; in contrast, if the conditions are favorable for cooperative relaying, i.e., c is low and M is large, the optimal strategy may take Form III, in which skipping transmission is never optimal, even when the S-D link condition is very unfavorable.

We note that the optimal strategy is threshold-based and is amenable to distributed implementation. Indeed, using the knowledge of channel statistics, each user can independently compute off-line λ^* by numerically solving the equation (5.18), and then compute off-line $\gamma_{sd,1}^*$, $\gamma_{sd,2}^*$ and $\gamma_{sd,3}^*$ by numerically solving their defining equations, which are the thresholds of the optimal strategy.

5.6 Numerical Results

In this section, we provide some numerical examples to illustrate the effectiveness of our proposed schemes for opportunistic cooperative networking. Specifically, we compare the performance of OCN-DR and OCN-NDR with that of DOS [8]. We focus on the throughput gain defined as

$$\Gamma \triangleq (\lambda^* - \lambda_0^*)/\lambda_0^*$$

where λ_0^* denotes the optimal throughput in DOS. We adopt the DF relaying technique and set $p_s = 0.1$ and $\tau = 0.1$. We assume that path loss fading effects are reflected in the normalized SNRs. Fig. 5.7 depicts the throughput gain of OCN-DR over DOS where we set $\tau_1 = 0.1$ and $T = 1$. It is clear that OCN-DR substantially outperforms DOS. In particular, the performance improvement is more significant when ρ is small. This shows that cooperative relaying is more efficient for improving the rate when SNR is low. Moreover, we observe that the throughput gain significantly increases as ρ_{sr} improves. This illustrates that the S-R link condition is crucial for the cooperative relaying rate. In Fig. 5.8, we present the performance comparison of OCN-NDR with DOS where we set $\tau_2 = 0.2$ and $T = 1$. Fig. 5.8 also illustrates higher throughput gains when ρ is small. Furthermore, we observe that the throughput gain degrades when ρ becomes too small. This implies that the price of relaying offsets the rate gain provided by cooperative relaying when the rate is small. In addition, it can be seen that the throughput gain is higher for smaller c and larger M. This confirms our intuition that cooperative relaying is more favorable when the relaying price is low or the number of potential relays is large. In Fig. 5.9, we examine the impact of τ_1, τ_2, and T on the performance of OCN-DR and OCN-NDR where we set $c = 0.1$ and $M = 5$ for OCN-NDR. We can see that the throughput gain decreases as τ_1 and τ_2 increases, indicating that the benefit of cooperative relaying is degraded when the time overhead of probing for cooperative relaying is large. Moreover, we observe that the throughput gain degrades faster for a smaller T. This implies that the impact of probing cost for cooperative relaying is more significant when the channel available duration is shorter.

5.7 Conclusion and Discussions

We have considered distributed opportunistic scheduling for ad hoc networks in which a user with successful contention may look for cooperative communication from relays.

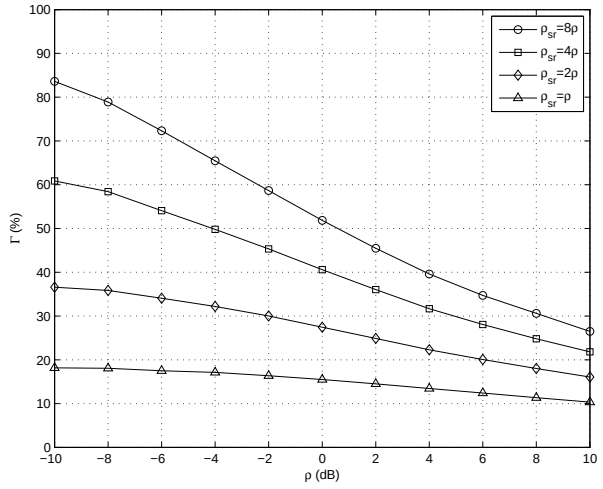


Figure 5.7: Γ vs ρ for OCN-DR: $\rho_{sd} = \rho_{rd} = \rho$.

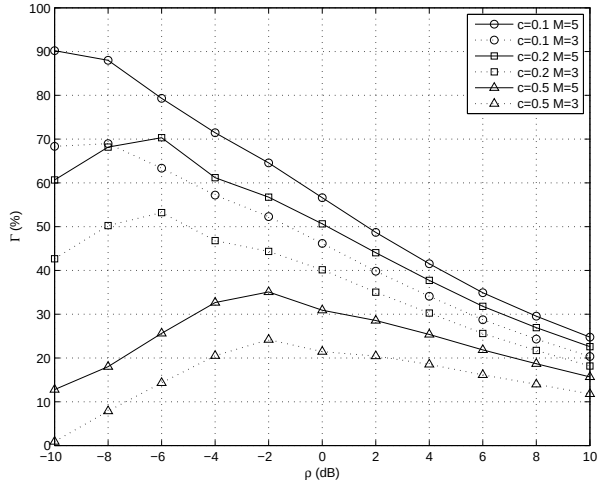


Figure 5.8: Γ vs ρ for OCN-NDR: $\rho_{sr} = 4\rho$, $\rho_{sd} = \rho_{rd} = \rho$.

Specifically, we have investigated opportunistic cooperative networking by characterizing and optimizing the tradeoff between throughput gain via cooperative communication and the probing cost of establishing cooperative relaying. We have studied two cases: OCN with dedicated relays (OCN-DR) and OCN without dedicated relays (OCN-NDR). Casting opportunistic cooperative networking as an optimal stopping problem with two-levels of incomplete information, we have shown that the optimal

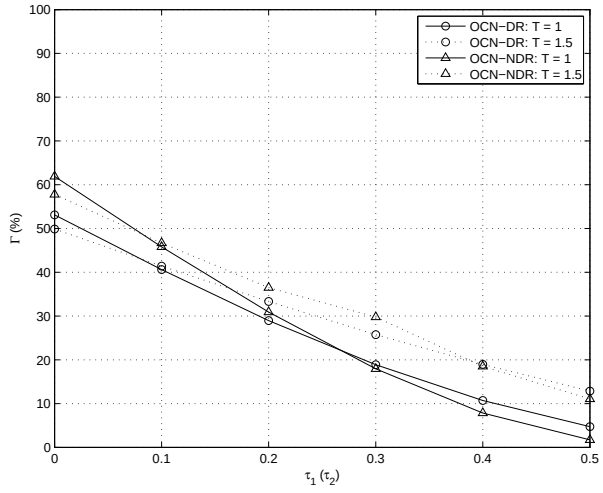


Figure 5.9: Γ vs τ_1 (τ_2) for OCN-DR (OCN-NDR): $\rho_{sr} = 4\rho$, $\rho_{sd} = \rho_{rd} = \rho = 0$ (dB).

strategies exist for both cases. In particular, for OCN-DR, it is optimal to probe the dedicated relay when the SNR of the S-R link is above some threshold; for OCN-NDR, it is optimal to probe potential relays when the SNR of the S-D link is between two thresholds. Furthermore, the optimal strategies are amenable to distributed implementation.

We offer some comments on comparing the optimal strategies of OCN-DR and OCN-NDR. Recall that for OCN-DR, the decision on whether to probe for cooperative relaying is based on the knowledge of both S-D and S-R link conditions, while for OCN-NDR, only the S-D link condition is available for making the decision. Thus, intuitively, the optimal strategy of OCN-DR is more favorable for achieving a higher throughput in the sense that it makes a more educated decision. Indeed, as we discussed earlier, the knowledge of the S-R link is crucial for evaluating the potential reward of cooperative relaying. Furthermore, OCN-DR is more advantageous in that cooperative relaying entails a smaller probing overhead and no price for the relay. Generally speaking, the existence of dedicated relays facilitates establishing and conducting cooperative relaying. However, it is worth noting that OCN-NDR benefits

from multi-user diversity for cooperative relaying through the best relay selection.

5.8 Appendix

5.8.1 Proof of Lemma 5.1

Using [81, Ch.6,Th.1], it suffices to show that there exist a stopping rule $N(\lambda)$ for a given $\lambda > 0$ such that

$$N(\lambda) = \arg \sup_{N \in Q} E[R_N T_{d,N} - \lambda T_N].$$

Let $Z_n \triangleq R_n T_{d,n} - \lambda T_n$ for brevity. It then follows from [81, Ch.3,Th.1] that $N(\lambda)$ exists if the following two conditions are satisfied:

$$\text{(A1) } E[\sup_n Z_n] < \infty \text{ and } \text{(A2) } \limsup_{n \to \infty} Z_n = -\infty \text{ a.s.}$$

Since all link channel gains have finite means and variances, it can be shown that the transmission rates R_d, R_r^{DF}, $R_r^{\mathrm{DF}-}$ and R_r^{CF} also have finite means and variances. Therefore, we have $E[R_n] < \infty$ and $E[{R_n}^2] < \infty$. It then follows from [81, Ch.4,Th.1] that (A2) holds. Also, observing that

$$E[\sup_n Z_n] \le E\left[\sup_n \left\{ R_n T - n\lambda\tau \left(\frac{1}{p_s} - \epsilon \right) \right\}\right] + E\left[\sup_n \sum_{j=1}^{n} \lambda\tau \left(\frac{1}{p_s} - \epsilon - K_j \right)\right] - \lambda T \quad (5.19)$$

where ϵ is chosen to satisfy $0 < \epsilon < 1/p_s$, it follows from [81, Ch.4,Th.1 and 2] that the first two terms on the RHS of (5.19) are finite, and hence (A1) holds.

5.8.2 Derivation of (5.9) and (5.11)

Applying integration by parts to (5.6), $V_0(\gamma_{sd}, \gamma_{sr}|\lambda)$ can be written as

$$V_0(\gamma_{sd}, \gamma_{sr}|\lambda) = (1 - \tau_1) \int_{\lambda}^{\infty} (1 - F_{R_m}(z|\gamma_{sd}, \gamma_{sr}))dz - \lambda\tau_1. \quad (5.20)$$

Case with DF relaying:

Consider the case $\gamma_{sd} \geq \gamma_{sr}$. Observe from (5.2) that $R_r^{\mathrm{DF}} \leq \log(1+\gamma_{sr}) \leq \log(1+\gamma_{sd}) = R_d$, and hence $R_m = R_d$. Then we have

$$F_{R_m}(z|\gamma_{sd},\gamma_{sr}) = \begin{cases} 1, & z \geq \log(1+\gamma_{sd}) \\ 0, & z < \log(1+\gamma_{sd}) \end{cases}$$

which, upon substituting into (5.20), yields (5.9a).

Consider the case $\gamma_{sd} \geq \gamma_{sr}$. It follows from (5.2) that $R_r^{\mathrm{DF}} \geq \log(1+\gamma_{sd}) = R_d$, and hence $R_m = R_r^{\mathrm{DF}}$. Also observe from (5.2) that, for given γ_{sd} and γ_{sr}, we have $\log(1+\gamma_{sd}) < R_r^{\mathrm{DF}} \leq \log(1+\gamma_{sr})$. Further, if the strict inequality $\log(1+\gamma_{sd}) < R_r^{\mathrm{DF}} < \log(1+\gamma_{sr})$ holds, from (5.2) it follows that $R_r^{\mathrm{DF}} = \log(1+(1-\beta^*)\gamma_{sr}) = \log(1+\gamma_{sd}+\gamma_{rd}+2\sqrt{\beta^*\gamma_{sd}\gamma_{rd}})$, where β^* is the optimal β in (5.2). Solving this equation for γ_{rd} with $z = R_r^{\mathrm{DF}}$, we obtain $\gamma_{rd}^{\mathrm{DF}}(\gamma_{sd},\gamma_{sr},z)$ which is given by (5.10). Note that $\gamma_{rd}^{\mathrm{DF}}(\gamma_{sd},\gamma_{sr},z)$ is increasing in z. From the above observations we have

$$F_{R_m}(z|\gamma_{sd},\gamma_{sr}) = \begin{cases} 1, & z \geq \log(1+\gamma_{sd}) \\ G(\gamma_{rd}^{\mathrm{DF}}(\gamma_{sd},\gamma_{sr},z)/\rho), & \log(1+\gamma_{sd}) < z < \log(1+\gamma_{sr}) \\ 0, & z \leq \log(1+\gamma_{sd}) \end{cases}$$

Substituting $F_{R_m}(z|\gamma_{sd},\gamma_{sr})$ into (5.20) yields (5.9b)-(5.9d).

Case with CF relaying:

From (5.4) we note that $R_r^{\mathrm{CF}} \geq \log(1+\gamma_{sd}) = R_d$, and hence $R_m = R_r^{\mathrm{CF}}$. Also observe from (5.4) that, for any given γ_{sd} and γ_{sr}, we have $\log(1+\gamma_{sd}) < R_r^{\mathrm{CF}} < \log(1+\gamma_{sd}+\gamma_{sr})$. Solving $z = \log(1+\gamma_{sd}+\frac{\gamma_{sr}\gamma_{rd}}{1+\gamma_{sd}+\gamma_{sr}+\gamma_{rd}})$ for γ_{rd}, yields $\gamma_{rd}^{\mathrm{CF}}(\gamma_{sd},\gamma_{sr},z)$ which is given by (5.12). Observe that $\gamma_{rd}^{\mathrm{CF}}(\gamma_{sd},\gamma_{sr},z)$ is increasing in z. Thus, we have

$$F_{R_m}(z|\gamma_{sd},\gamma_{sr}) = \begin{cases} 1, & z \geq \log(1+\gamma_{sd}+\gamma_{sr}) \\ G(\gamma_{rd}^{\mathrm{CF}}(\gamma_{sd},\gamma_{sr},z)/\rho), & \log(1+\gamma_{sd}) < z < \log(1+\gamma_{sd}+\gamma_{sr}) \\ 0, & z \leq \log(1+\gamma_{sd}) \end{cases}$$

Substituting $F_{R_m}(z|\gamma_{sd},\gamma_{sr})$ in (5.20) yields (5.11a)-(5.11c).

5.8.3 Proof of Lemma 5.2

We first show that $L_{2,3}(\lambda)$, $L_{1,3}(\lambda)$ and $L_{1,2}(\lambda)$ are curves in $\mathbb{R}_+^2$, and each curve divides $\mathbb{R}_+^2$ into optimal decision regions for the pair of options associated with the curve.

It is clear that $L_{1,3}(\lambda)$ is the curve given by $\{(\gamma_{sd}, \gamma_{sr}) : \gamma_{sd} = e^{\lambda} - 1, \gamma_{sr} \geq 0\}$. Furthermore, since $R_d(\gamma_{sd}) - \lambda$ is increasing in γ_{sd}, we have $R_d(\gamma_{sd}) - \lambda > 0$ on one side of $L_{1,3}(\lambda)$ and $R_d(\gamma_{sd}) - \lambda < 0$ on the other side.

Next, we consider $L_{1,2}(\lambda)$ for the case with DF relaying, that is, the set $\{(\gamma_{sd}, \gamma_{sr}) \in \mathbb{R}_+^2\}$ such that $V_0(\gamma_{sd}, \gamma_{sr}|\lambda) = 0$. Clearly, in the case of (5.9a), $L_{1,2}(\lambda)$ is the curve given by $\{(\gamma_{sd}, \gamma_{sr}) : \gamma_{sd} = e^{\lambda/(1-\tau_1)} - 1, 0 \leq \gamma_{sr} \leq e^{\lambda/(1-\tau_1)} - 1\}$. It is also clear that $L_{1,2}(\lambda)$ does not exist in the case of (5.9d). Consider now the cases of (5.9b) and (5.9c). Note that $V_0(\gamma_{sd}, \gamma_{sr}|\lambda) > 0$ for all $\{(\gamma_{sd}, \gamma_{sr}) : \gamma_{sr} > \gamma_{sd} \geq e^{\lambda/(1-\tau_1)} - 1\}$. Also observe that for any γ_{sd} with $0 \leq \gamma_{sd} < e^{\lambda/(1-\tau_1)} - 1$, we have $\lim_{\gamma_{sr} \to \max\{e^{\lambda}-1, \gamma_{sd}\}} V_0(\gamma_{sd}, \gamma_{sr}|\lambda) < 0$ and $\lim_{\gamma_{sr} \to \infty} V_0(\gamma_{sd}, \gamma_{sr}|\lambda) = \infty$. Then, since $V_0(\gamma_{sd}, \gamma_{sr}|\lambda)$ is increasing in γ_{sr}, for any $\gamma_{sd} \in [0, e^{\lambda/(1-\tau_1)} - 1)$, there exists a unique $\gamma_{sr,3}(\gamma_{sd})$ such that $V_0(\gamma_{sd}, \gamma_{sr,3}(\gamma_{sd})|\lambda) = 0$, and hence $(\gamma_{sd}, \gamma_{sr,3}(\gamma_{sd})) \in L_{1,2}(\lambda)$. Thus, it follows from the continuity of $V_0(\gamma_{sd}, \gamma_{sr}|\lambda)$ that $L_{1,2}(\lambda)$ is a curve. Furthermore, it is easy to see from the above discussions that we have $V_0(\gamma_{sd}, \gamma_{sr}|\lambda) > 0$ on one side of $L_{1,2}(\lambda)$ and $V_0(\gamma_{sd}, \gamma_{sr}|\lambda) < 0$ on the other side. Using similar arguments, we can establish the proof of $L_{1,2}(\lambda)$ for the case with CF relaying.

Then, consider $L_{2,3}(\lambda)$ for the case with DF relaying, that is, the set $\{(\gamma_{sd}, \gamma_{sr}) \in \mathbb{R}_+^2\}$ such that $W(\gamma_{sd}, \gamma_{sr}|\lambda) = R_d(\gamma_{sd}) - \lambda - V_0(\gamma_{sd}, \gamma_{sr}|\lambda) = 0$. It is clear that in the case of (5.9a) and (5.9d), $L_{2,3}(\lambda)$ is the curve given by $\{(\gamma_{sd}, \gamma_{sr}) : \gamma_{sd} = e^{\lambda(1-\tau_1)} - 1, 0 \leq \gamma_{sr} \leq e^{\lambda} - 1\}$. Consider now the cases of (5.9b) and (5.9c). We note that for any $\gamma_{sd} > e^{\lambda(1-\tau_1)} - 1$, we have $\lim_{\gamma_{sr} \to \max\{e^{\lambda}-1, \gamma_{sd}\}} W(\gamma_{sd}, \gamma_{sr}|\lambda) > 0$ and

$\lim_{\gamma_{sr}\to\infty} W(\gamma_{sd},\gamma_{sr}|\lambda) = -\infty$. Then, since $W(\gamma_{sd},\gamma_{sr}|\lambda)$ is decreasing in γ_{sr}, for any $\gamma_{sd} \in (e^{\lambda(1-\tau_1)}-1,\infty)$, there exists a unique $\gamma_{sr,1}(\gamma_{sd})$ such that $W(\gamma_{sd},\gamma_{sr,1}(\gamma_{sd})|\lambda) = 0$, and hence $(\gamma_{sd},\gamma_{sr,1}(\gamma_{sd})) \in L_{2,3}(\lambda)$. Thus, from the continuity of $W(\gamma_{sd},\gamma_{sr}|\lambda)$ it follows that $L_{2,3}(\lambda)$ is a curve. Furthermore, from the above discussions we can easily see that we have $W(\gamma_{sd},\gamma_{sr}|\lambda) > 0$ on one side of $L_{2,3}(\lambda)$ and $W(\gamma_{sd},\gamma_{sr}|\lambda) < 0$ on the other side. We can use similar arguments to establish the proof of $L_{2,3}(\lambda)$ for the case with CF relaying.

It can be seen from the above discussions that $L_{1,3}(\lambda)$ has a unique intersection point with $L_{2,3}(\lambda)$ and $L_{1,2}(\lambda)$, respectively. Also, it is clear from the definitions of $L_{2,3}(\lambda)$, $L_{1,3}(\lambda)$ and $L_{1,2}(\lambda)$ that any intersection point of two of them is an intersection point of the three. We thus conclude that $L_{2,3}(\lambda)$, $L_{1,3}(\lambda)$ and $L_{1,2}(\lambda)$ intersect at a unique point.

Based on the above results, now we can see from Fig. 5.2 that $\mathbb{R}_+^2$ is divided by one section of $L_{2,3}(\lambda)$, one section of $L_{1,3}(\lambda)$ and one section of $L_{1,2}(\lambda)$ into three nonempty and connected sets, which are exactly the optimal decision regions $A_1(\lambda)$, $A_2(\lambda)$ and $A_3(\lambda)$. This completes the proof.

5.8.4 Derivation of (5.14)

Applying integration by parts to (5.6), $V_0(\gamma_{sd}|\lambda)$ can be written as

$$V_0(\gamma_{sd}|\lambda) = (1-\tau_2)\int_\lambda^\infty (1 - F_{R_m}(z|\gamma_{sd}))dz - \lambda\tau_2. \tag{5.21}$$

Since R_m is bounded by $R_d = \log(1+\gamma_{sd}) \le R_m$ for a given γ_{sd}, we have $F_{R_m}(z|\gamma_{sd}) = 0$ when $e^z - 1 < \gamma_{sd}$. When $\gamma_{sd} \ge e^z - 1$, it follows that $F_{R_m}(z|\gamma_{sd}) = \prod_{i=1}^{M} F_{R_{r_i}}(z+c|\gamma_{sd})$, where $F_{R_{r_i}}(\cdot|\gamma_{sd})$ denotes the CDF of R_{r_i} conditioned on γ_{sd}. Since the links are homogeneous, we have $F_{R_{r_1}}(\cdot|\gamma_{sd}) = \cdots = F_{R_{r_M}}(\cdot|\gamma_{sd})$, and hence $F_{R_m}(z|\gamma_{sd}) = F_{R_{r_1}}^M(z+c|\gamma_{sd})$. Then (5.14) follows.

Since R_{r_1} is bounded by $\log(1+\gamma_{sd}) \leq R_{r_1}$ for a given γ_{sd}, we have $F_{R_{r_1}}(z+c|\gamma_{sd}) = 0$ when $e^{z+c}-1 < \gamma_{sd}$. Consider the case $\gamma_{sd} \leq e^{z+c}-1$. For the case with synchronous DF relaying, since $R_{r_1}^{\mathrm{DF}} \leq \log(1+\gamma_{sr_1})$ and $R_{r_1}^{\mathrm{DF}} \leq \log(1+(1-\beta^*)\gamma_{sr_1}) = \log(1+\gamma_{sd}+\gamma_{r_1d}+2\sqrt{\beta^*\gamma_{sd}\gamma_{r_1d}})$, we have (5.15). For the case with asynchronous DF relaying, it follows from $R_{r_1}^{\mathrm{DF}-} \leq \log(1+\gamma_{sr_1})$ and $R_{r_1}^{\mathrm{DF}-} \leq \log(1+\gamma_{sd}+\gamma_{r_1d})$ that we have (5.16). For the case with CF relaying, we have (5.17).

5.8.5 Proof of Lemma 5.3

It is clear from the definition of $\gamma_{sd,2}(\lambda)$ that $\gamma_{sd,2}(\lambda) = e^{\lambda}-1$. Further, since $R_d(\gamma_{sd})-\lambda$ is increasing in γ_{sd}, we have $R_d(\gamma_{sd})-\lambda < 0$ when $\gamma_{sd} \in [0, e^{\lambda}-1)$, and $R_d(\gamma_{sd})-\lambda > 0$ when $\gamma_{sd} \in (e^{\lambda}-1, \infty)$.

Consider then $\gamma_{sd,3}(\lambda)$, i.e., some $\gamma_{sd} > 0$ such that $V_0(\gamma_{sd}|\lambda) = 0$. Observe from (5.16) and (5.17) that $F_{R_{r_1}}(R_m + c|\gamma_{sd})$ is decreasing in γ_{sd}, then it follows from (5.14a) and (5.14b) that $V_0(\gamma_{sd}|\lambda)$ is increasing in γ_{sd}. It is clear from (5.14a) that $V_0(\gamma_{sd}|\lambda) > 0$ when $\gamma_{sd} > e^{\lambda/(1-\tau_2)}-1$. We also note from (5.14b) that $V_0(0|\lambda)$ can be greater or less than 0, depending on system parameters such as c and M. Thus, if $V_0(0|\lambda) \leq 0$, $\gamma_{sd,3}(\lambda)$ exists and is unique, and we have $V_0(\gamma_{sd}|\lambda) < 0$ when $\gamma_{sd} \in [0, \gamma_{sd,3}(\lambda))$, and $V_0(\gamma_{sd}|\lambda) > 0$ when $\gamma_{sd} \in (\gamma_{sd,3}(\lambda), \infty)$; if $V_0(0|\lambda) > 0$, $\gamma_{sd,3}(\lambda)$ does not exist, and we have $V_0(\gamma_{sd}|\lambda) > 0$ for any $\gamma_{sd} \in [0, \infty)$.

Next, consider $\gamma_{sd,1}(\lambda)$, i.e., some $\gamma_{sd} \geq e^{\lambda}-1$ such that $W(\gamma_{sd}|\lambda) = R_d(\gamma_{sd})-\lambda-V_0(\gamma_{sd}|\lambda) = 0$. Recall from (5.3) and (5.4) that we can write the rate of cooperative transmission as a function $R_r(\gamma_{sd}, \gamma_{sr}, \gamma_{rd})$. Using (5.6), for any $\gamma_{sd} \geq e^{\lambda}-1$, we can

write $V_0(\gamma_{sd}|\lambda)$ as

$$\begin{aligned} V_0(\gamma_{sd}|\lambda) &= (1-\tau_2)\int_{\lambda}^{\infty}(z-\lambda)F_{R_m}(z|\gamma_{sd})dz - \lambda\tau_2 \\ &\overset{(a)}{=} (1-\tau_2)\int_S (R_m(\gamma_{sd},\mathbf{x})-\lambda)f_\Gamma(\mathbf{x})d\mathbf{x} - \lambda\tau_2 \\ &= (1-\tau_2)\int_S (\max\{R_d(\gamma_{sd}), R_{r_1}(\gamma_{sd},x_1,x_2)-c, \\ &\cdots, R_{r_M}(\gamma_{sd},x_{2M-1},x_{2M})-c\}-\lambda)f_\Gamma(\mathbf{x})d\mathbf{x} - \lambda\tau_2 \end{aligned} \tag{5.22}$$

where $S \triangleq \{(\gamma_{sr_1},\gamma_{r_1d},\cdots,\gamma_{sr_M},\gamma_{r_Md}) : \gamma_{sr_1}>0,\gamma_{r_1d}>0,\cdots,\gamma_{sr_M}>0,\gamma_{r_Md}>0\}$, $\mathbf{x} \triangleq (x_1,x_2,\cdots,x_{2M-1},x_{2M}) \in \mathbb{R}^+_{2M}$, $f_\Gamma(\cdot)$ denotes the joint PDF of $\Gamma \triangleq (\gamma_{sr_1},\gamma_{r_1d},\cdots,\gamma_{sr_M},\gamma_{r_Md})$, and (a) follows from the fact that, given $\gamma_{sd} \geq e^\lambda - 1$, $R_m(\gamma_{sd},\mathbf{x}) \geq \lambda$ for any $\mathbf{x} \in S$. Observe from (5.3) and (5.4) that, for any $i \in \{1,\cdots,M\}$, $\gamma_{sr_i}>0$ and $\gamma_{r_id}>0$, we have

$$\frac{\partial R_{r_i}^{\mathrm{DF}-}(\gamma_{sd},\gamma_{sr_i},\gamma_{r_id})}{\partial\gamma_{sd}} \leq \frac{1}{1+\gamma_{sd}+\gamma_{r_id}} < \frac{1}{1+\gamma_{sd}} \tag{5.23}$$

for the case with asynchronous DF relaying, and

$$\frac{\partial R_{r_i}^{\mathrm{CF}}(\gamma_{sd},\gamma_{sr_i},\gamma_{r_id})}{\partial\gamma_{sd}} = \frac{1-\frac{\gamma_{sr_i}\gamma_{r_id}}{(1+\gamma_{sd}+\gamma_{sr_i}+\gamma_{r_id})^2}}{1+\gamma_{sd}+\frac{\gamma_{sr_i}\gamma_{r_id}}{1+\gamma_{sd}+\gamma_{sr_i}+\gamma_{r_id}}} < \frac{1}{1+\gamma_{sd}} \tag{5.24}$$

for the case with CF relaying. Then, for any $\gamma_{sd} \geq e^\lambda - 1$, from (5.22) we have

$$\begin{aligned} \frac{\partial V_0(\gamma_{sd}|\lambda)}{\partial\gamma_{sd}} &\leq (1-\tau_2)\int_S \max\{\frac{\partial R_d(\gamma_{sd})}{\partial\gamma_{sd}}, \frac{\partial R_{r_1}(\gamma_{sd},x_1,x_2)}{\partial\gamma_{sd}}, \\ &\qquad \cdots, \frac{\partial R_{r_M}(\gamma_{sd},x_{2M-1},x_{2M})}{\partial\gamma_{sd}}\} f_\Gamma(\mathbf{x})d\mathbf{x} \\ &\overset{(b)}{\leq} (1-\tau_2)\int_S \frac{1}{1+\gamma_{sd}} f_\Gamma(\mathbf{x})d\mathbf{x} \\ &= (1-\tau_2)\frac{1}{1+\gamma_{sd}} \\ &< \frac{1}{1+\gamma_{sd}} = \frac{\partial(R_d(\gamma_{sd})-\lambda)}{\partial\gamma_{sd}} \end{aligned} \tag{5.25}$$

where (b) follows from (5.23) and (5.24). Thus, $W(\gamma_{sd}|\lambda)$ is increasing in γ_{sd} for $\gamma_{sd} \in [e^\lambda - 1, \infty)$. We also observe that for any $\gamma_{sd} \geq e^\lambda - 1$,

$$\begin{aligned} W(\gamma_{sd}|\lambda) &= W(e^\lambda - 1|\lambda) + \int_{e^\lambda - 1}^{\gamma_{sd}} \frac{\partial W(x|\lambda)}{\partial x} dx \\ &\overset{(c)}{\geq} W(e^\lambda - 1|\lambda) + \int_{e^\lambda - 1}^{\gamma_{sd}} \frac{\tau_2}{1+x} dx \\ &= W(e^\lambda - 1|\lambda) + \tau_2 \log(1+x) \Big|_{e^\lambda - 1}^{\gamma_{sd}} \end{aligned}$$

where (c) follows from (5.25). Hence, we have $\lim_{\gamma_{sd} \to \infty} W(\gamma_{sd}|\lambda) = \infty$. Recall that $V_0(\gamma_{sd}|\lambda)$ is increasing in γ_{sd}. Then we note that if $e^\lambda - 1 \leq \gamma_{sd,3}(\lambda)$, we have $W(e^\lambda - 1|\lambda) = -V_0(e^\lambda - 1|\lambda) \geq -V_0(\gamma_{sd,3}(\lambda)|\lambda) = 0$, and hence $\gamma_{sd,1}(\lambda)$ does not exist, and we further have $W(\gamma_{sd}|\lambda) > 0$ for any $\gamma_{sd} \in (e^\lambda - 1, \infty)$; if $\gamma_{sd,3}(\lambda) < e^\lambda - 1$ or $\gamma_{sd,3}(\lambda)$ does not exist, we have $W(e^\lambda - 1|\lambda) = -V_0(e^\lambda - 1|\lambda) < -V_0(\gamma_{sd,3}(\lambda)|\lambda) = 0$ or $W(e^\lambda - 1|\lambda) = -V_0(e^\lambda - 1|\lambda) < -V_0(0|\lambda) < 0$, and hence $\gamma_{sd,1}(\lambda)$ exists and is unique, and we further have $W(\gamma_{sd}|\lambda) < 0$ when $\gamma_{sd} \in (e^\lambda - 1, \gamma_{sd,1}(\lambda))$, and $W(\gamma_{sd}|\lambda) > 0$ when $\gamma_{sd} \in (\gamma_{sd,1}(\lambda), \infty)$.

Lemma 5.3 directly follows from the above discussions.

Chapter 6

OPTIMAL PLACEMENT FOR BARRIER COVERAGE IN BISTATIC RADAR SENSOR NETWORKS

6.1 Introduction

Wireless sensor networks have received tremendous attention over the past decade. Typically, it is assumed that a sensor network is composed of *passive sensors* (e.g., thermal, acoustics, optic sensors) which detect radiation that is emitted or reflected by an object. In contrast, an active *radar* (RAdio Detection And Ranging) purposefully emits radio waves with the objective of collecting echoes. The ability to design the structure and power of the transmitted radio signal imbues active radars with performance advantages over passive sensors in many application scenarios, though this is typically at the expense of additional system complexity.

Thanks to recent technological advances, radars are becoming less expensive and more compact, making it feasible to deploy a network of radars working in concert. Indeed, the application scale and scope of networked radar sensors[1] are expected to expand significantly. Due to the advantages of radars over traditional passive sensors, radar networks have great potential for many applications, such as border security [82] and traffic monitoring [83]. Nevertheless, to fully exploit this potential, radar networks should be judiciously designed.

Coverage, which defines how well the object of interest is monitored, is a critical performance metric for sensor networks. *Barrier coverage* has recently emerged as an efficient coverage strategy for numerous sensor network applications centered around *intruder detection*, such as border monitoring and drug interdiction, and has drawn a surge of research interest [12,84–86]. Despite tremendous research efforts on

[1]For brevity, we use "radar" and "radar sensor" interchangeably.

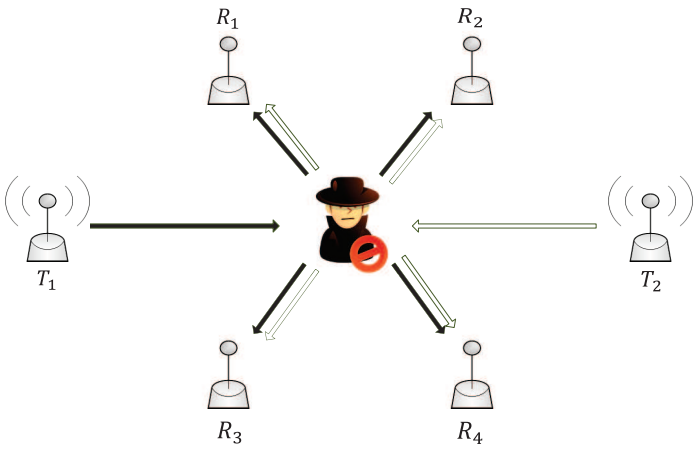


Figure 6.1: Illustration of bistatic radar network.

coverage problems for sensor networks [9], those pertaining to radar sensors remain largely unexplored, and this is the main subject of this study.

In this chapter, we consider the problem of deploying a network of bistatic radars (BRs) for intrusion detection. Due to the flexibility to deploy the radar transmitter and receiver separately, a BR is more favorable than a monostatic radar (MR) for coverage. Our goal is to *build a fundamental understanding of a bistatic radar network (BRN) for coverage*. In particular, a central question we ask here is: *Where should the BRs be placed to achieve the optimal coverage quality?*

The coverage problem of a BRN is dramatically different and more challenging than that of a network of traditional passive sensors, because 1) departing from the disk sensing region of a passive sensor, the sensing region of a BR depends on the locations of both the transmitter and receiver, and is characterized by a Cassini oval. Formally, a Cassini oval is a locus of points for which the distances to two fixed points (foci) have a constant product. Fig. 6.2 illustrates Cassini ovals with foci at BR transmitter T and receiver R for different distance products: $c_1 < c_2 < c_3 < c_4$; 2) the sensing regions of different BRs are coupled with each other, since each

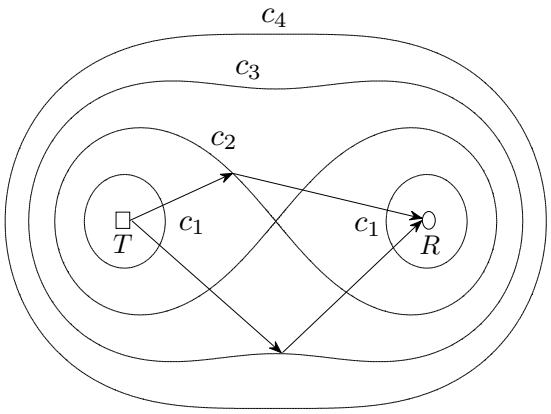


Figure 6.2: Bistatic radar SNR contours as Cassini ovals.

BR transmitter[2] (or receiver) can potentially pair with different BR receivers (or transmitters, respectively) to form multiple BRs such that its location would impact multiple BRs.

Next we summarize the main contributions of this chapter.

- We consider the problem of deploying a network of BRs in a region to maximize the *worst-case intrusion detectability*, which is equivalent to minimizing the *vulnerability* of the optimal *barrier* in the region. We show that it is optimal to place BRs on the *shortest barrier* if it is the shortest line segment that connects the left and right boundary of the region.

- The main thrust of this study is devoted to characterizing the optimal placement of BRs on a line segment to minimize its vulnerability, *which is a highly non-trivial optimization problem due to its non-convexity.* To tackle the challenges herein, we recast the problem as finding the optimal *placement order* of BR nodes with the optimal *placement spacing*. Based on an important structure

[2]For brevity, we use "transmitter" and "BR transmitter", "receiver" and "BR receiver" interchangeably, respectively.

of detectability, we characterize *balanced* placement spacing and show that it is optimal. Using the optimal placement spacing, we then characterize the optimal placement orders, which also present balanced structures. These findings provide valuable insights into the placement of BRs for barrier coverage.

Although it is somewhat idealized, the Cassini oval sensing model (see SNR equation (7.1)) used in this chapter can capture the essential feature of a BR, compared to a passive sensor or MR. Furthermore, the coverage problem of a BRN corresponding to the Cassini oval sensing model gives rise to significant technical difficulties (as will be seen later). Needless to say, future work is needed to generalize this study to more complex and realistic situations. In short, we believe that this study will open a new door to explore radar sensor networks.

The rest of this chapter is organized as follows. Section 6.2 introduces the model of bistatic radar network and the worst-case coverage, and defines the optimal placement problem. In Section 6.3, we address the optimal placement problem based on the barrier coverage strategy. We study the optimal placement of BRs on a line segment in Section 6.4. Numerical results are provided in Section 6.5 and related work is reviewed in Section 6.6. Section 6.7 concludes this chapter and discusses future work.

6.2 Model and Problem Definition

In this section, we first describe the model of bistatic radar network and the worst-case coverage, and then define the optimal placement problem.

6.2.1 Bistatic Radar Network

The radar transmitter and receiver of a BR are at different locations, whereas they are co-located for a MR. Intuitively, a BR can achieve better coverage than a MR by appropriate placement of the transmitter and receiver, such that the target is more

likely to be physically closer to either the transmitter or receiver, and thus attains a high signal-to-noise ratio (SNR). This advantage of BR will be illustrated by an example in Section 6.3.

One fundamental metric of target detection for a BR is its received SNR: the strength of the received radar signal indicates how likely the target is present. Let $\|AB\|$ and $\overline{AB}$ denote the (Euclidean) distance and the line segment between point A and B, respectively. For convenience, we also use T_i or R_j to denote the location (point) of a transmitter node T_i or receiver node R_j, respectively. For a BR T_i-R_j, the received SNR from the target located at a point X is given by [87]:

$$\text{SNR} = \frac{K}{\|T_iX\|^2\|R_jX\|^2} \tag{6.1}$$

where K denotes a *bistatic radar constant* that reflects physical characteristics of the BR, such as transmit power, *radar cross section*[3], and antenna power gains. The SNR contours of a BR are characterized by the Cassini ovals with foci at the transmitter and receiver of the BR.

For a network of BRs, we assume that all transmitters operate on orthogonal radio resources (e.g., by using orthogonal waveforms [88–91]) to avoid mutual interference at a receiver. While multiple receivers can pair with the same transmitter to form multiple BRs, a receiver can also pair with multiple transmitters. Typically, a BRN has more receivers than transmitters, mainly because that a transmitter incurs higher cost than a receiver (e.g., since signal transmission consumes more energy than other sensor activities such as signal reception and processing). In addition, the number of transmitters can also be limited by the available radio resources (e.g., the number of orthogonal waveforms).

[3]Radar cross section measures the amount of radar signal energy reflected by an object depending on its physical characteristics (e.g., shape, material).

We consider the deployment of a BRN consisting of M transmitters $T_i \in \mathcal{T}$, $i \in \mathcal{M} \triangleq \{1, \cdots, M\}$ and N receivers $R_j \in \mathcal{R}$, $j \in \mathcal{N} \triangleq \{1, \cdots, N\}$. We assume that transmitters and receivers, respectively, have homogeneous physical characteristics such that all BRs have the same bistatic radar constant. We assume that a receiver can potentially pair with all transmitters to form multiple BRs. However, in Section 6.4 we will show that it suffices for a receiver to pair with *at most two* transmitters. We also assume that transmitters and receivers are *omnidirectional*. We further assume that the radar signals reflected by the target are *omnidirectional*[4].

6.2.2 Worst-case Coverage

The BRN is deployed in a 2D geographical *region of interest* F to detect an intruder that traverses through the region. The region F is defined by an entrance side, a destination side, a left boundary F_l, and a right boundary F_r (as illustrated in Fig. 6.3). The intruder can choose any *intrusion path* P in region F that connects the entrance to the destination.

Existing studies on sensor network coverage [93–95] use the distance from a point to its *closest* sensor to measure the coverage of the point (also known as the *closest sensor observability*). In the same spirit, we measure the coverage of a point by the *highest* SNR received by a BR among all BRs, when the target is present at the point. In Section 6.4.4, we will discuss the case where data fusion is used such that the coverage depends on the SNRs received by multiple BRs. Considering (7.1), we have the following definition.

Definition 6.1 (*Detectability*)**.** *The* detectability[5] *of a point X, denoted by $I(X)$, is*

[4]The reflected radar signals may not be omnidirectional. For the sake of tractability, most of the literature on bistatic radar (e.g., [82, 87, 92]) assumes that they are omnidirectional.

[5]With a little abuse of notation, we use $I(X)$ to denote the detectability of X, while the detectability of X changes inversely with the value of $I(X)$.

the minimum distance product of X with respect to a BR among all BRs:

$$I(X) \triangleq \min_{T_i \in \mathcal{T}, R_j \in \mathcal{R}} \|T_i X\| \|R_j X\|. \tag{6.2}$$

In other words, the detectability of a point is determined by the *closest* BR to the point, which consists of its closest transmitter and closest receiver. Similar to [93–95], we use the *worst-case intrusion* to quantify the coverage of the intruder.

Definition 6.2 (*Worst-case Intrusion [93]*)**.** *The* worst-case intrusion path*, denoted by P^*, is the intrusion path with the minimum detectability among all possible intrusion paths:*

$$P^* \triangleq arg \max_{P \in \mathcal{P}} D(P) \tag{6.3}$$

where $\mathcal{P}$ denotes the set of all possible intrusion paths, and $D(P)$ denotes the detectability of intrusion path P, which is the maximum detectability of a point among all the points in P:

$$D(P) \triangleq \min_{X \in P} I(X). \tag{6.4}$$

6.2.3 Problem Definition

We are interested in finding the optimal placement of the BRN (i.e., the optimal locations of M transmitters and N receivers) in region F that maximizes the worst-case intrusion detectability:

$$\underset{T_i \in F, R_j \in F}{\text{minimize}} \quad D(P^*). \tag{6.5}$$

Based on the notion of worst-case coverage, problem (6.5) is of great interest for the intruder detection problem. In particular, solving problem (6.5) provides the answer to an important question: How many transmitters and receivers are needed, and where should they be placed to provide the required coverage quality such that *at least one* BR will receive an SNR above a predefined threshold, regardless of the intruder's path?

Note that problem (6.5) is difficult to solve in general (even for sensors with disk sensing regions). This is because the shape of region F can be arbitrary, and the feasible solution space contains infinitely many placements in region F.

6.3 Placement for Barrier Coverage

In this section, we address problem (6.5) using the approach of barrier coverage. We show that under certain conditions it is optimal to place BRs on the shortest barrier in the region, in which case it is a line segment. We also investigate the placement on the shortest barrier that is an arbitrary curve. All the proofs of this chapter are given in Appendix.

6.3.1 Optimality Condition for Shortest Barrier based Placement

Similar to [12,84–86], we define a *barrier* as a curve in region F such that *any* intrusion path intersects with the curve. We use the following concept as the coverage metric of a barrier.

Definition 6.3 (*Vulnerability*)**.** *The vulnerability $V(U)$ of a barrier U is the minimum detectability of a point among all the points in U:*

$$V(U) \triangleq \max_{X \in U} I(X). \tag{6.6}$$

The rationale of using vulnerability as the coverage metric is that, since a barrier intersects with any possible intrusion path, the vulnerability of a barrier serves as an *upper bound* on the worst-case intrusion detectability. Furthermore, this bound is *tight* when all the barriers are taken into account, such that

$$D(P^*) = \min_{U \in \mathcal{U}} V(U) \tag{6.7}$$

where $\mathcal{U}$ denotes the set of all the barriers in region F.

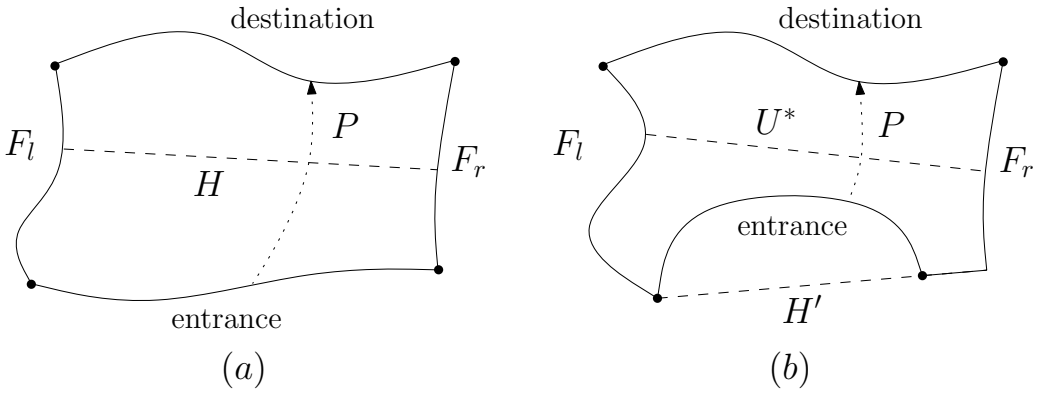


Figure 6.3: Examples of region of interest F (a) with and (b) without the shortcut barrier.

Using (6.7), problem (6.5) boils down to finding the *optimal barrier*, which is the optimal solution of the following problem:

$$\underset{U \in \mathcal{U}}{\text{minimize}} \;\; V^*(U) \tag{6.8}$$

where $V^*(U)$ denotes the minimum vulnerability of U, which is the optimal value of the following problem:

$$\underset{T_i \in F, R_j \in F}{\text{minimize}} \;\; V(U). \tag{6.9}$$

It is plausible that the optimal barrier for problem (6.8) should be the *shortest barrier*, denoted by U^*, which is the barrier with the minimum length among all the barriers. However, this strategy is not optimal in general, because $V^*(U^*)$ can be greater than $V^*(U)$ for a barrier U with a greater length than U^*. We give an illustrative example in Fig. 6.4. For line segment $\overline{AB}$ with $\|AB\| = 2\sqrt{2}$ in Fig. 6.4 (a), it is clear that the optimal placement of a BR T_1-R_1 that minimizes $V(\overline{AB})$ is to set $\|AT_1\| = \|R_1B\| = \sqrt{2} - 1$ such that $V^*(\overline{AB}) = \|AT_1\|\|AR_1\| = 1$. For curve $\widetilde{CF}$[6] in Fig. 6.4 (b) with $\|CD\| = \|DE\| = \|EF\| = 1$ and $\overline{CD}, \overline{EF} \perp \overline{DE}$, it has a greater length than $\overline{AB}$, while we have $V^*(\widetilde{CF}) \leq V(\widetilde{CF}) = \|TE\|\|ER\| = \sqrt{5}/4 <$

[6]We use $\widetilde{PQ}$ to denote a curve with endpoints P, Q, and $\|\widetilde{PQ}\|$ to denote its length.

$1 = V^*(\overline{AB})$, where $V(\widetilde{CF})$ denotes the vulnerability of $\widetilde{CF}$ when T_1 and R_1 are placed at the midpoint of $\overline{CD}$ and $\overline{EF}$, respectively. Therefore, if $\overline{AB}$ is the shortest barrier in region F while $\widetilde{CF}$ is another barrier in F (which is possible), the optimal barrier cannot be $\overline{AB}$.

Before proceeding further, we use a simple example to illustrate the advantage of a BR over a MR for barrier coverage. If we place a MR (which consists of a pair of co-located radar transmitter and receiver) on $\overline{AB}$ in Fig. 6.4 (a) to minimize $V(\overline{AB})$, the optimal placement location is clearly at the midpoint Y_{AB} of $\overline{AB}$ such that $V(\overline{AB}) = \|AY_{AB}\|^2 = 2$. This is greater than $V^*(\overline{AB}) = 1$ which is achieved under the optimal placement of a BR T_1-R_1.

Although it is in general not optimal to place BRs on the shortest barrier U^*, this strategy is optimal if U^* is also the *shortcut barrier* defined as follows.

Definition 6.4 (*Shortcut Barrier*)**.** *The shortcut barrier, denoted by H, exists and is the shortest barrier U^* if and only if U^* is the shortest line segment that connects left boundary F_l and right boundary F_r (i.e., the length of U^* is the minimum distance between a point in F_l and a point in F_r).*

Although the shortest line segment that connects F_l and F_r always exists, it may not be in region F, and therefore is not a barrier. As illustrated in Fig. 6.3 (b), H' is the shortest line segment that connects F_l and F_r but is not a barrier, and thus the shortcut barrier does not exist. The shortcut barrier exists for a large class of shapes of region F (e.g., any convex region). Note that if the shortest barrier U^* is also the shortcut barrier, then U^* must be a line segment; otherwise, U^* may or may not be a line segment. As illustrated in Fig. 6.3 (b), the shortest barrier U^* is a line segment but is not the shortcut barrier. We next show that the existence of the shortcut barrier is the optimality condition.

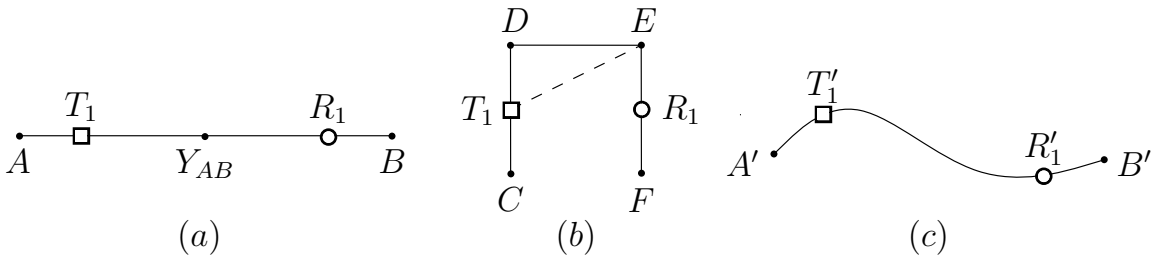


Figure 6.4: An example that shows the optimal barrier may not be the shortest barrier.

Theorem 6.1. *If the shortcut barrier H exists, then H is the optimal barrier for problem* (6.8)*. As a result, it suffices to solve problem* (6.9) *for H in order to solve problem* (6.5)*.*

Note that Theorem 6.1 provides a *sufficient* condition under which the optimal barrier is the shortest barrier. With regard to the optimal placement of BRs for a line segment, we have the following result.

Proposition 6.1. *For a line segment $\overline{AB}$, the optimal placement for problem* (6.9) *for $\overline{AB}$ is on $\overline{AB}$, and the optimal value of problem* (6.9) *for $\overline{AB}$ increases as its length $\|AB\|$ increases.*

The proof is based on a similar argument as in the proof of Theorem 6.1, and is thus omitted. By Theorem 6.1 and Proposition 6.1, as it is optimal to place BRs on the shortcut barrier H, which is a line segment, in Section 6.4 we will focus on finding the optimal placement of BRs on a line segment that minimize its vulnerability.

6.3.2 Placement on Curved Shortest Barrier

If the shortest barrier U^* is not the shortcut barrier, it can be an arbitrary curve and may not be the optimal barrier for problem (6.8). Furthermore, it is in general difficult to find the optimal placement for problem (6.9) for an arbitrary curve (even for sensors with disk sensing regions). In this case, we can find a placement

$\{\mathcal{T}', \mathcal{R}'\}$ on U^* that *imitates* the optimal placement $\{\mathcal{T}, \mathcal{R}\}$ for problem (6.9) for a line segment $\overline{U^*}$ that has the same length as U^*, such that $\|AT_1\| = \|\widetilde{A'T_1'}\|$, $\|T_1T_2\| = \|\widetilde{T_1'T_2'}\|$, $\cdots$, $\|T_MB\| = \|\widetilde{T_M'B'}\|$ and $\|AR_1\| = \|\widetilde{A'R_1'}\|$, $\|R_1R_2\| = \|\widetilde{R_1'R_2'}\|$, $\cdots$, $\|R_MB\| = \|\widetilde{R_M'B'}\|$, where A, B are the endpoints of $\overline{U^*}$ and A', B' are the endpoints U^*. For example, in Fig. 6.4 (a) and (c), the placement $\{T_1', R_1'\}$ on $\widetilde{A'B'}$ imitates the placement $\{T_1, R_1\}$ on $\overline{AB}$ with $\|AB\| = \|\widetilde{A'B'}\| = 2\sqrt{2}$, such that $\|\widetilde{A'T_1'}\| = \|\widetilde{R_1'B'}\| = \sqrt{2} - 1$.

It has an appealing property as stated in the following result.

Proposition 6.2. *For the optimal placement $\{\mathcal{T}, \mathcal{R}\}$ on the line segment $\overline{U^*}$ and the imitation placement $\{\mathcal{T}', \mathcal{R}'\}$ on the shortest barrier U^*, we have $V^*(\overline{U^*}) \geq V'(U^*)$.*

The proof is based on a similar argument as in the proof of Theorem 6.1, and is thus omitted. By Proposition 6.2, $V^*(\overline{U^*})$ is no less than $V(U^*)$ under the imitation placement, and thus $V^*(\overline{U^*})$ serves as an upper bound for the worst-case intrusion detectability $D(P^*)$, which is the objective that we aim to minimize (i.e., the objective value of problem (6.5)). Since the shortest barrier U^* has the minimum length among all the barriers, according to Proposition 6.1, the imitation placement on U^* gives the *minimum* upper bound $V^*(\overline{U^*})$ for $D(P^*)$ among all the barriers.

6.4 Optimal Placement on A Line Segment

In this section, we study the optimal placement of BRs on a line segment, say the shortcut barrier H, to minimize its vulnerability $V(H)$.

6.4.1 Problem Recast

Let H_l and H_r be the left and right endpoint of H, respectively, and h be the length of H. Also let $t_i \triangleq \|H_lT_i\|$ and $r_j \triangleq \|H_lR_j\|$. Mathematically, our problem can be

written as

$$\begin{aligned} \underset{t_i, r_j}{\text{minimize}} \quad & \max_{0 \le x \le h} \min_{i \in \mathcal{M}, j \in \mathcal{N}} |x - t_i||x - r_j| \\ \text{subject to} \quad & 0 \le t_i \le h, \forall i \in \mathcal{M} \\ & 0 \le r_j \le h, \forall j \in \mathcal{N} \end{aligned} \tag{6.10}$$

where $\min_{i \in \mathcal{M}, j \in \mathcal{N}} |x - t_i||x - r_j|$ represents the detectability of a point $X \in H$ with $\|H_l X\| = x$, and $\max_{0 \le x \le h} \min_{i \in \mathcal{M}, j \in \mathcal{N}} |x - t_i||x - r_j|$ represents the vulnerability of H. In general, we can show that problem (6.10) is *non-convex*. Therefore, standard optimization methods would not work well here.

To gain useful insight into problem (6.10), we view the line segment based placement in an intuitive way as follows. First we treat H_l and H_r as two (virtual) nodes and relax the constraint $\|H_l H_r\| = h$. Then we place all the BR nodes $\mathcal{T}$ and $\mathcal{R}$ as well as H_l and H_r on a horizontal line subject to the constraint that H_l and H_r are the *leftmost* and *rightmost* node, respectively. We can use the following concepts to characterize any placement on a line.

Definition 6.5 (*Placement Order and Spacing*)**.** *A* placement order *(referred to as "order" for brevity)* $\mathbf{S}$ *is an order of all the nodes on the line from left to right:*

$$\mathbf{S} \triangleq (H_l, S_1, \cdots, S_J, H_r)$$

where $J \triangleq M + N$ *and* $(S_1, \cdots, S_J)$ *is a permutation of the BR nodes such that* $\|H_l H_l\| \le \|H_l S_1\| \le \cdots \le \|H_l S_J\| \le \|H_l H_r\|$. *The* placement spacing *(referred to as "spacing" for brevity)* $\mathbf{D_S}$ *of a placement order* $\mathbf{S}$ *consists of the distances each between a pair of neighbor nodes in* $\mathbf{S}$*:*

$$\mathbf{D_S} \triangleq (\|H_l S_1\|, \cdots, \|S_J H_r\|).$$

Table 6.1: Frequently used notation

Notation	Description
$\overline{AB}$	line segment between point A and B
$\|AB\|$	distance between point A and B (length of $\overline{AB}$)
Y_{AB}	midpoint of $\overline{AB}$
M, N	total number of BR transmitters, BR receivers
$\mathcal{T}$, $\mathcal{R}$	set of all BR transmitters, BR receivers
T_i, R_j	BR transmitter i, BR receiver j
T_i-R_j	BR consisting of T_i and R_j
$I(X)$	detectability of X
$V(\overline{AB})$	vulnerability of $\overline{AB}$
H, h	shortcut barrier and its length
H_l, H_r	left and right endpoint (node) of H
$\mathcal{S}$	placement of $\mathcal{T}$, $\mathcal{R}$, H_l, and H_r on a line
$\mathbf{S}$, $\mathbf{D_S}$	placement order and its placement spacing
$\mathbf{S}_i$, $\mathbf{D}_{\mathbf{S}_i}$	independent local placement order and its local placement spacing
$Z_{\mathbf{S}_i}$, $L_{\mathbf{S}_i}$	local zone of $\mathbf{S}_i$ and its length
e_c^i	parameter that characterizes balanced spacing

A local placement order (referred to as "local order" for brevity) $(S_{i+1}, \cdots, S_{i+j})$ *is an order of a set of neighbor nodes in* $\mathbf{S}$*, and its placement spacing is*

$$\mathbf{D}_{(S_{i+1}\cdots,S_{i+j})} \triangleq (\|S_{i+1}S_{i+2}\|, \cdots, \|S_{i+j-1}S_{i+j}\|).$$

We can see that any order $\mathbf{S}$ with any spacing $\mathbf{D_S}$ characterize a unique placement of BRs on a line segment with length $\|H_lH_r\|$; conversely, any placement of BRs on the line segment H can be uniquely characterized by some order $\mathbf{S}$ with some spacing $\mathbf{D_S}$ that satisfies $\|H_lH_r\| = h$. Therefore, our problem (6.10) can be recast as

$$\begin{aligned} \underset{\mathbf{S},\mathbf{D_S}}{\text{minimize}} \quad & V(\overline{H_lH_r}) \\ \text{subject to} \quad & \|H_lH_r\| = h. \end{aligned} \tag{6.11}$$

It is clear that the optimal value of problem (6.11) increases as h increases.

Algorithm 5: Compute the optimal placement for problem (6.11)

input : line segment length h, precision threshold ϵ

output: optimal order $\mathbf{S}^*$, optimal spacing $\mathbf{D}^*_{\mathbf{S}^*}$, optimal value c^*

$c_1 \leftarrow 0$, $c_2 \leftarrow h^2$, $c \leftarrow \frac{c_1+c_2}{2}$;
repeat
  Compute the optimal order $\mathbf{S}^*$ and the optimal spacing $\mathbf{D}^*_{\mathbf{S}^*}$ for problem (6.12) subject to $V(\overline{H_l H_r}) \leq c$;
  if $l_c > h + \epsilon$ **then**
    $c_2 \leftarrow c$; $c \leftarrow \frac{c_1+c_2}{2}$;
  end
  if $l_c < h - \epsilon$ **then**
    $c_1 \leftarrow c$; $c \leftarrow \frac{c_1+c_2}{2}$;
  end
until $|c - c_1| \leq \epsilon$;
return $\mathbf{S}^*$, $\mathbf{D}^*_{\mathbf{S}^*}$, $c^* \leftarrow c$;

As a result, we can formulate a problem relevant to problem (6.11) as

$$\begin{aligned} \underset{\mathbf{S},\mathbf{D_S}}{\text{maximize}} \quad & \|H_l H_r\| \\ \text{subject to} \quad & V(\overline{H_l H_r}) \leq c. \end{aligned} \tag{6.12}$$

Let l_c denote the optimal value of problem (6.12) under the constraint $V(\overline{H_l H_r}) \leq c$. It is also clear that l_c is increasing in c and, in particular, $l_c \to 0$ when $c \to 0$ and $l_c \to \infty$ when $c \to \infty$. Therefore, if we can solve problem (6.12) for any $c > 0$, we can also solve problem (6.11) by a *bisection search* as described in Algorithm 5. Specifically, Algorithm 5 keeps track of an interval $[c_1, c_2]$ that must contain the optimal value c^* of problem (6.11), and reduce the interval by half at each step, until the interval is sufficiently small such that the difference between c^* and the endpoint c_1 or c_2 is upper bounded by a predefined precision threshold ϵ. Since the initial interval is set to $[0, h^2]$, the number of steps for running Algorithm 5 is upper bounded by $O(\log(h^2/\epsilon))$.

We make two observations regarding any placement of BRs. First, since all BRs are homogeneous, *swapping* the locations of any pair of transmitters (or receivers,

respectively) results in an equivalent placement. Second, transmitters and receivers are *reciprocal* in the sense that replacing all transmitters by receivers while replacing all receivers by transmitters results in an equivalent placement. As a result, for ease of exposition, in the rest of this chapter we assume that $M \leq N$. All the analysis hereafter can directly apply to the case $M > N$ by treating transmitters as receivers while treating receivers as transmitters.

6.4.2 Optimal Placement Order and Spacing

In this subsection, our goal is to characterize the optimal order and the optimal spacing for problem (6.12). We outline the major steps to achieve this goal as follows.

1) We show an important structure of detectability for any placement on a line (Lemma 6.1), based on which we define *balanced spacings* and *independent local orders*.

2) We characterize the balanced local spacing for an independent local order (Lemma 6.2), and show that it is optimal (Lemma 6.3).

3) We show that the balanced spacing for a *dividable* order consists of balanced local spacings for independent local orders, and it is optimal (Theorem 6.2).

4) We show that there exist optimal orders in the class of dividable orders (Lemma 6.4), based on which we characterize the optimal orders (Theorem 6.3).

We start with the observation that the optimal order $\mathbf{S}^*$ for problem (6.12) is equivalent to the optimal order of the following problem:

$$\underset{\mathbf{S}}{\text{maximize}}\ f_c^{\mathbf{S}} \tag{6.13}$$

where $f_c^{\mathbf{S}}$ denotes the optimal value of the following problem for order $\mathbf{S}$:

$$\begin{aligned} \underset{\mathbf{D_S}}{\text{maximize}} \quad & \|H_l H_r\| \\ \text{subject to} \quad & V(\overline{H_l H_r}) \leq c. \end{aligned} \tag{6.14}$$

Therefore, the optimal spacing $\mathbf{D}^*_{\mathbf{S}^*}$ for problem (6.12) is equivalent to the optimal spacing for problem (6.14) for the optimal order $\mathbf{S}^*$.

The following lemma presents an important structure of detectability for any placement on a line. We use Y_{AB} to denote the midpoint of a line segment $\overline{AB}$.

Lemma 6.1. *For any order* $\mathbf{S}$ *with any spacing* $\mathbf{D_S}$*, the detectability on* $\overline{H_l H_r}$ *attains local maximums at the end nodes and at the midpoint of each pair of neighbor BR nodes (as illustrated in Fig. 6.5):*

$$\arg \max_{X \in \overline{H_l S_1}} I(X) = H_l, \quad \arg \max_{X \in \overline{S_J H_r}} I(X) = H_r$$

$$\arg \max_{X \in \overline{S_i S_{i+1}}} I(X) = Y_{S_i S_{i+1}}, \quad \forall i \in \{1, \cdots, J-1\}.$$

Definition 6.6 (*Local Vulnerable Point*)**.** *A local vulnerable point is a local maximum point of detectability on* $\overline{H_l H_r}$*, and a* local vulnerable value *is its detectability.*

By Lemma 6.1, it suffices to examine the local vulnerable values to determine the vulnerability of a line segment. Based on Lemma 6.1, we define the following concept.

Definition 6.7 (*Independent Local Order*)**.** *A local order* $\mathbf{S}_i$ *is an* independent local order *if it has any of the following* types*:*

$$(T, R), (R, T)$$

$$(T, R^k, H_r), (R, T^k, H_r), (H_l, R^k, T), (H_l, T^k, R), k \geq 1$$

$$(T, R^k, T), (R, T^k, R), k \geq 1$$

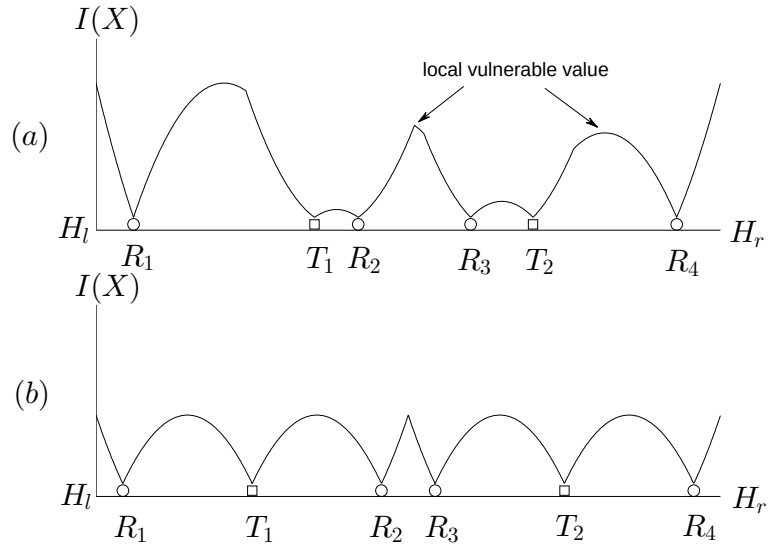


Figure 6.5: Local vulnerable values under (a) arbitrary and (b) balanced placement spacing.

where T^k and R^k denote k consecutive transmitters and receivers, respectively. The independent local zone $Z_{\mathbf{S}_i}$ *of an independent local order* $\mathbf{S}_i$ *is the line segment between the two end nodes in* $\mathbf{S}_i$*, with its length denoted by* $L_{\mathbf{S}_i}$.

For any local placement with any independent local order $\mathbf{S}_i$, the *closest* BR for any local vulnerable point on the independent local zone $Z_{\mathbf{S}_i}$ consists of the nodes in $\mathbf{S}_i$. For example, for $\mathbf{S}_i = (T_1, R_1)$, the closest BR for $Y_{T_1R_1}$ is T_1-R_1; for $\mathbf{S}_i = (T_1, R_1, \cdots, R_k, H_r)$, the closest BR for any of $Y_{T_1R_1}$, $\cdots$, $Y_{R_{k-1}R_k}$, and H_r, consists of transmitter T_1 and some receiver among R_1, $\cdots$, R_k. Therefore, *all the local vulnerable values on* $Z_{\mathbf{S}_i}$*, and hence the vulnerability* $V(Z_{\mathbf{S}_i})$*, are determined by the spacing* $\mathbf{D}_{\mathbf{S}_i}$ *(i.e., independent of any distance not in* $\mathbf{D}_{\mathbf{S}_i}$*)*. Based on this property, the following concept is well-defined for an independent local order.

Definition 6.8 (*Balanced Spacing*)**.** *The spacing* $\mathbf{D}_{\mathbf{S}}$ *(or local spacing* $\mathbf{D}_{\mathbf{S}_i}$*) of an order* $\mathbf{S}$ *(or an independent local order* $\mathbf{S}_i$*, respectively) is* balanced *if all the local vulnerable values on* $\overline{H_lH_r}$ *(or the independent local zone* $Z_{\mathbf{S}_i}$*, respectively) are equal*

(as illustrated in Fig. 6.5 (b)).

The next lemma characterizes the balanced local spacing for an independent local order.

Lemma 6.2. *For any $c > 0$, define $e_c^0 \triangleq 2\sqrt{c}$ and let e_c^j denote the unique positive value of x such that*

$$\left(\sum_{i=0}^{j-1} e_c^i + \frac{x}{2}\right)\frac{x}{2} = c \tag{6.15}$$

for each $j \in \mathbb{N}^+$. For any $c > 0$ and any independent local order $\mathbf{S}_i$, there exists a unique balanced local spacing $\mathbf{D}_{\mathbf{S}_i}$ such that $V(Z_{\mathbf{S}_i}) = c$, and furthermore, it can be characterized by e_c^i, $i \in \mathbb{N}$ as follows: for $\mathbf{S}_i$ with type (T, R) or (R, T),

$$\mathbf{D}_{\mathbf{S}_i} = (e_c^0);$$

for $\mathbf{S}_i$ with type (T, R^k, H_r) or (R, T^k, H_r),

$$\mathbf{D}_{\mathbf{S}_i} = (e_c^0, e_c^1, \cdots, e_c^{k-1}, \frac{e_c^k}{2});$$

for $\mathbf{S}_i$ with type (H_l, R^k, T) or (H_l, T^k, R),

$$\mathbf{D}_{\mathbf{S}_i} = (\frac{e_c^k}{2}, e_c^{k-1}, \cdots, e_c^1, e_c^0);$$

for $\mathbf{S}_i$ with type (T, R^k, T) or (R, T^k, R), if k is even,

$$\mathbf{D}_{\mathbf{S}_i} = (e_c^0, e_c^1, \cdots, e_c^{\frac{k}{2}-1}, e_c^{\frac{k}{2}}, e_c^{\frac{k}{2}-1}, \cdots, e_c^1, e_c^0);$$

if k is odd,

$$\mathbf{D}_{\mathbf{S}_i} = (e_c^0, e_c^1, \cdots, e_c^{\frac{k-1}{2}}, e_c^{\frac{k-1}{2}}, \cdots, e_c^1, e_c^0).$$

By definition, given c, the value of e_c^i, $i \in \mathbb{N}^+$ can be found iteratively using (6.15), which decreases as i increases (as shown in Table 6.2).

Table 6.2: Values of balanced spacing

c	e_c^0	e_c^1	e_c^2	e_c^3	e_c^4
1	2.0000	0.8284	0.6357	0.5359	0.4721
5	4.4721	1.8524	1.4214	1.1983	1.0557
10	6.3246	2.6197	2.0102	1.6947	1.4930
20	8.9443	3.7048	2.8428	2.3966	2.1115

Based on the independent property, we can cast a problem in the same spirit as problem (6.14) but for an independent local order $\mathbf{S}_i$ as

$$\begin{aligned} &\underset{\mathbf{D}_{\mathbf{S}_i}}{\text{maximize}} \quad L_{\mathbf{S}_i} \\ &\text{subject to} \quad V(Z_{\mathbf{S}_i}) \leq c. \end{aligned} \tag{6.16}$$

The next lemma shows that the balanced local spacing is optimal for problem (6.16).

Lemma 6.3. *For any $c > 0$ and any independent local order $\mathbf{S}_i$, the balanced local spacing $\mathbf{D}_{\mathbf{S}_i}$ such that $V(Z_{\mathbf{S}_i}) = c$ is the optimal local spacing for problem* (6.16).

The next definition presents a useful structure of a class of orders.

Definition 6.9 (*Dividable Order*)**.** *An order* $\mathbf{S}$ *is* dividable *if it can be decomposed into independent local orders $\mathbf{S}_1, \cdots, \mathbf{S}_m$ such that 1) each node in $\mathbf{S}$ is included in some $\mathbf{S}_i$; 2) the last node of $\mathbf{S}_i$ is the first node of $\mathbf{S}_{i+1}$ for each $i = 1, \cdots, m-1$. Therefore, the spacing $\mathbf{D}_{\mathbf{S}}$ consists of* disjoint *independent local spacings $\mathbf{D}_{\mathbf{S}_1}$, $\cdots$, $\mathbf{D}_{\mathbf{S}_m}$. For example, the following order is dividable:*

$$\mathbf{S} = (\overbrace{H_l, R_1, R_2, \underbrace{T_1}_{\mathbf{S}_2}}^{\mathbf{S}_1}, R_3, R_4, R_5, \overbrace{\underbrace{T_2}_{\mathbf{S}_2}, \underbrace{R_6}_{\mathbf{S}_4}}^{\mathbf{S}_3}, T_3, T_4, H_r). \tag{6.17}$$

For any placement with any dividable order $\mathbf{S}$, the local vulnerable points on $\overline{H_l H_r}$ consists of *disjoint* sets of local vulnerable points on the independent local zones $Z_{\mathbf{S}_1}$, $\cdots$, $Z_{\mathbf{S}_m}$. Therefore, *problem* (6.14) *for* $\mathbf{S}$ *can be decomposed into independent subproblems, each of which is an instance of problem* (6.16) *for* $\mathbf{S}_1, \cdots, \mathbf{S}_m$, with the

optimal local spacing given by Lemma 6.3. The next theorem follows from the above observation.

Theorem 6.2. *For any $c > 0$ and any dividable order $\mathbf{S}$, where $\mathbf{S}$ can be decomposed into independent local orders $\mathbf{S}_1, \cdots, \mathbf{S}_m$, the balanced spacing $\mathbf{D}_{\mathbf{S}}$ such that $V(\overline{H_l H_r}) = c$ exists and consists of the balanced local spacings $\mathbf{D}_{\mathbf{S}_1}, \cdots, \mathbf{D}_{\mathbf{S}_m}$. Furthermore, it is the optimal spacing for problem* (6.14) *for* $\mathbf{S}$.

Next we show that there exist optimal orders in the class of dividable orders. Since all transmitters are homogeneous, we index the transmitters from left to right such that $0 \leq \|H_l T_1\| \leq \cdots \leq \|H_l T_M\| \leq \|H_l H_r\|$. Define

$$\mathbf{N}_{\mathbf{S}} \triangleq (n_1, n_2, \cdots, n_M, n_{M+1})$$

where n_i, n_1, n_{M+1} denote the number of receivers in $\mathbf{S}$ between T_{i-1} and T_i for $i \in \{2, \cdots, M\}$, between H_l and T_1, between T_M and H_r, respectively. Since all transmitters and all receivers are homogeneous, respectively, any order $\mathbf{S}$ can be uniquely characterized by $\mathbf{N}_{\mathbf{S}}$.

Lemma 6.4. *There exists an optimal order $\mathbf{S}^*$ that satisfies the following conditions:*

$$\nexists i, j \in \{1, \cdots, M\} \text{ such that } n_i \geq 2,\ n_j = 0 \tag{6.18a}$$

$$n_2 \neq 0,\ n_M \neq 0. \tag{6.18b}$$

Furthermore, any order $\mathbf{S}$ that satisfies the above conditions is dividable.

We should note that a non-optimal order (e.g., the order in (6.17)) can also be dividable.

Based on Lemma 6.4, the following theorem provides a *sufficient* condition for the optimal order.

Theorem 6.3. *An order* $\mathbf{S}$ *is optimal if it satisfies the following conditions:*

$$|n_i - n_j| \leq 1, \forall i, j \in \{2, \cdots, M\} \tag{6.19a}$$

$$|n_i - 2n_1| \leq 1, |n_i - 2n_{M+1}| \leq 1, \forall i \in \{2, \cdots, M\}. \tag{6.19b}$$

Using Theorem 6.3, we characterize the optimal order $\mathbf{S}^*$ as follows. Let two integers q and r be the quotient and remainder of $\frac{N}{M}$, respectively. If q is even (e.g., as in Fig.6.6 (a) and (b)), we have

$$\mathbf{N}_{\mathbf{S}^*} = (\frac{q}{2}, \overbrace{q+1, \cdots, q+1}^{r}, \overbrace{q, \cdots, q}^{M-1-r}, \frac{q}{2});$$

if q is odd and $r = 0$ (e.g., as in Fig.6.6 (c)), we have

$$\mathbf{N}_{\mathbf{S}^*} = (\frac{q+1}{2}, \overbrace{q, \cdots, q}^{M-1}, \frac{q-1}{2});$$

if q is odd and $r \geq 1$, we have

$$\mathbf{N}_{\mathbf{S}^*} = (\frac{q+1}{2}, \overbrace{q+1, \cdots, q+1}^{r-1}, \overbrace{q, \cdots, q}^{M-r}, \frac{q+1}{2}).$$

In addition, for any $\mathbf{N}_{\mathbf{S}^*}$, if we swap the values of n_1 and n_{M+1}, or the values of n_i and n_j for $i, j \in \{2, \cdots, M\}$, it also satisfies (6.19a) and (6.19b), and hence is optimal.

Given the above optimal order $\mathbf{S}^*$ for problem (6.13), which is dividable, we can characterize the optimal spacing $\mathbf{D}^*_{\mathbf{S}^*}$ for problem (6.14) using Theorem 6.2.

6.4.3 Remarks

Regarding the analysis and results in the last subsection, we have the following remarks.

Remark 1: The detectability structure given in Lemma 6.1 plays a fundamental role in our analysis, based on which the concepts of local vulnerable point, independent local order, balanced spacing, and dividable order can be defined thereafter. This structure is mainly due to that 1) all BRs are homogeneous and 2) for

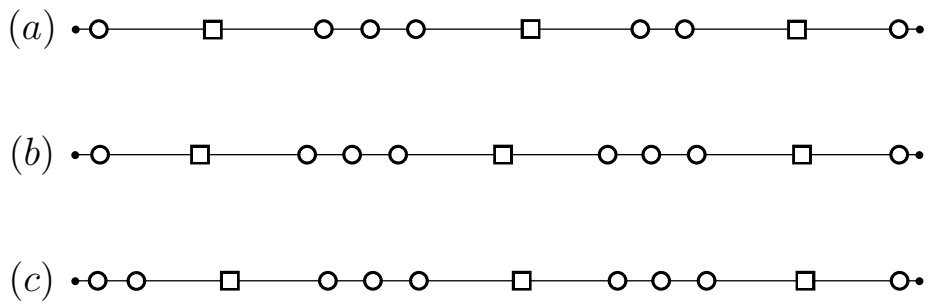


Figure 6.6: Optimal placement of BRs for $M = 3$ and (a) $N = 7$; (b) $N = 8$; (c) $N = 9$.

any BR, say T_1-R_1, the received SNR from a point on the line segment $\overline{T_1R_1}$ achieves local minimum at the midpoint $Y_{T_1R_1}$.

Remark 2: The optimal spacings are *balanced* in the sense that all the local vulnerable values are *equal* under the balanced spacing. The optimality of this balanced structure is due to that minimizing the vulnerability is equivalent to minimizing the maximum local vulnerable value.

The balanced spacing is in general *non-uniform* in the sense that the distance between each pair of neighbor BR nodes is not the same. For the balanced spacing of an independent local order $\mathbf{S}_i$ with type (T, R^k, T), (H_l, R^k, T), or (T, R^k, H_r), the distance between two neighbor receivers decreases as the distance to their closest transmitter increases (i.e., e_c^i decreases as i increases). This non-uniform structure is essentially due to that the detectability of a point is *jointly* determined by its closest transmitter and closest receiver, while the number of transmitters is *unbalanced* with that of receivers.

Remark 3: The optimal orders are also *balanced* but in a more subtle sense: for the optimal order that satisfies (6.19), n_i for $i = 2, \cdots, M$ are *as equal as possible*, while each n_i is *as equal as possible to two times* n_1 and n_{M+1}, respectively (as illustrated in Fig. 6.6). The optimality of this balanced structure is mainly because

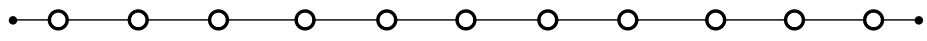


Figure 6.7: Optimal placement of disk sensors.

that the optimal value of problem (6.16) for an independent local order $\mathbf{S}_i$ with type (T, R^k, T), (H_l, R^k, T), or (T, R^k, H_r) increases as k increases, while the *marginal increment* decreases as k increases. Therefore, to maximize the sum of the optimal values of problem (6.16) for independent local orders $(H_l, \cdots, T_1)$, $(T_1, \cdots, T_2)$, $\cdots$, $(T_{M-1}, \cdots, T_M)$, $(T_M, \cdots, H_r)$, the optimal order must have that balanced structure.

Remark 4: We can gain useful insights by comparing the placement of BRs on a line segment to that of sensors with disk sensing regions (we assume that they are homogeneous and refer them as "disk sensors" for brevity). For disk sensors, we define the detectability of a point as the distance to its closest sensor. Interestingly, for any placement of disk sensors on a line segment, we can observe the *same* structure as in Lemma 6.1: the detectability on the line segment also achieves local maximums at the midpoint of each pair of neighbor disk sensors and the endpoints. While this detectability structure is clear for disk sensors (which are homogeneous), it is not obvious for BRs.

Based on the detectability structure, we can show that the balanced spacing is also optimal for disk sensors. However, as illustrated in Fig. 6.7, the balanced spacing for disk sensors is *uniform* (i.e., each pair of neighbor disk sensors have the same distance), in contrast to that it is non-uniform in general for BRs. We should also note that while the balanced spacing is an *equivalent* condition for the optimal placement of disk sensors, it is only a *necessary* condition for the optimal placement of BRs. This is because that different orders of BR nodes can be non-equivalent.

Remark 5: The optimal placement results imply that one assumption made in Section 6.2 can be relaxed without losing the optimality. Since the detectability

of a point is determined by its closest BR, it suffices for a receiver to pair with a transmitter only if they form the closest BR for some point on the line segment. Therefore, for the optimal placement of a BRN consisting of more receivers than transmitters (e.g., as in Fig. 6.6), we can see that a receiver only needs to pair with its *closest* transmitter(s), the number of which is *one or two*.

6.4.4 Discussions

In this study, we assume that there is no data fusion among different BRs for target detection. If data fusion is used, the metric of target detection would depend on the SNRs received by multiple BRs (e.g., the sum of the k highest SNRs where $k > 1$) rather than only the highest SNR. As a result, the detectability of a point would depend on its distant product with respect to multiple BRs rather than only the closest BR. As expected, it is very difficult to find the optimal placement of BRs for problem (6.10) in this case. However, with data fusion, we can show that all the results in Section 6.3 (including Theorem 6.1, Proposition 6.1 and 6.2) still hold via similar analysis, as long as the data fusion model satisfies that target detection improves as the SNR of any BR involved in the data fusion increases (which is typically true). In addition, for the optimal placement on a line segment H under the model without data fusion, we can analyze the coverage quality of this placement under a data fusion model. Specifically, using the detectability of a point as a function of its distances to certain BR nodes (which depends on the data fusion model), we can find the local maximums of the detectability on H (as we do in Lemma 6.1), and thus the vulnerability of H. Clearly, the vulnerability of H without data fusion serves as an upper bound for that with data fusion.

6.5 Numerical Results

In this section, we provide numerical results to illustrate the advantage of the optimal placement of BRs on a line segment H.

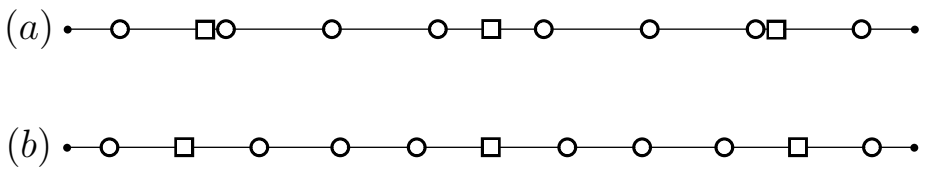


Figure 6.8: Heuristic placement of BRs for $M = 3$ and $N = 8$: (a) HEU-1; (b) HEU-2.

6.5.1 Comparison with Heuristic Placement

As no existing work has studied the placement of BRs for barrier coverage, we compare the optimal placement strategy (OPT) with two heuristic strategies. The heuristics are motivated by the rationale of the optimal placement strategy for a network of homogeneous disk sensors, which is to minimize the maximum distance from a point to its closest sensor among all the points on H.

The first heuristic (HEU-1) is to place transmitters (or receivers, respectively) with uniform spacing such that the maximum distance from a point on H to its closest transmitter (or receiver, respectively) is minimized (as illustrated in Fig. 6.8 (a)):

$$2\|H_l T_1\| = \|T_1 T_2\| = \cdots = \|T_{M-1} T_M\| = 2\|T_M H_r\|$$

$$2\|H_l R_1\| = \|R_1 R_2\| = \cdots = \|R_{N-1} R_N\| = 2\|R_N H_r\|.$$

Comparing Fig. 6.6 (b) with Fig. 6.8 (a) (under the same setting of $M = 3$ and $N = 8$), we can see that neither the placement of transmitters nor receivers in OPT is the same as that in HEU-1. Compared to OPT, the main drawback of HEU-1 is that it places transmitters and receivers *independently*.

The second heuristic (HEU-2) is to place transmitters and receivers according to the optimal order $\mathbf{S}^*$, but with uniform spacing such that the maximum distance from a point on H to its closest BR node (either transmitter or receiver) is minimized

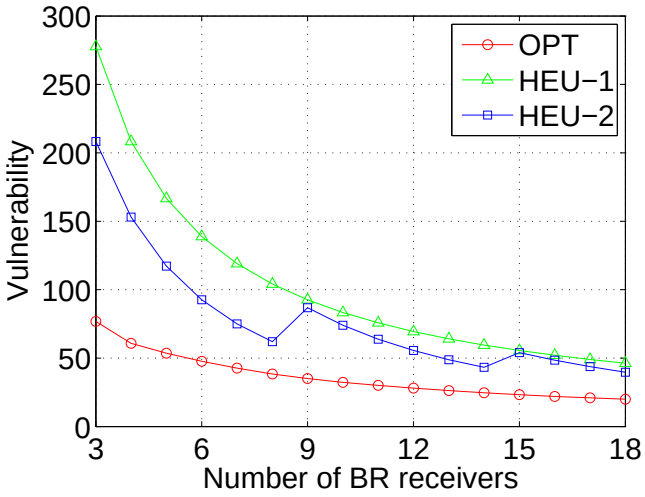


Figure 6.9: Impact of N for $M = 3$.

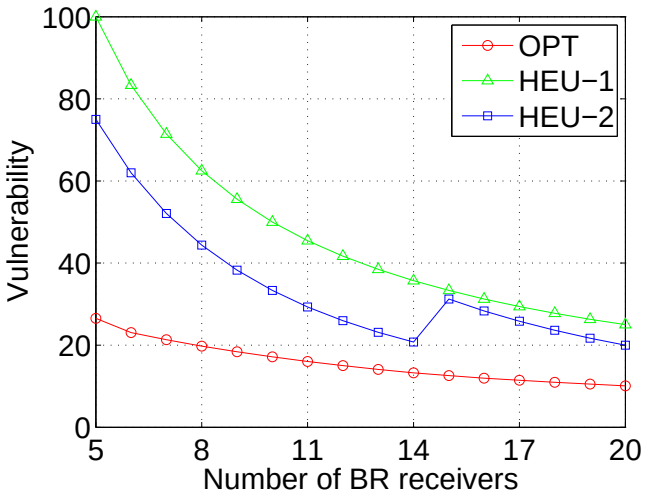


Figure 6.10: Impact of N for $M = 5$.

(as illustrated in Fig. 6.8 (b)):

$$2\|H_l S_1\| = \|S_1 S_2\| = \cdots = \|S_{M-1} S_M\| = 2\|S_M H_r\|.$$

Although HEU-2 follows the optimal order, its main drawback is that it treats transmitters and receivers *equivalently*.

Figs. 6.9-6.11 depict the vulnerability of H under OPT, HEU-1, and HEU-2 for a varying number of receivers and 3, 5, 10 transmitters, respectively. We set

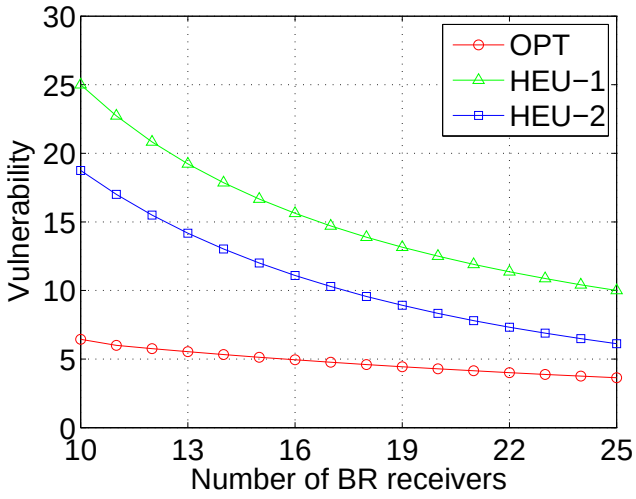


Figure 6.11: Impact of N for $M = 10$.

the length of H to 100m. We observe that HEU-2 results in considerably lower vulnerability than HEU-1, and OPT further outperforms HEU-2 significantly. This shows that OPT is highly advantageous for improving the barrier coverage, which is essentially due to that the design rationale for a BRN under the Cassini oval sensing model is quite different from that for a network of passive sensors or MRs under the disk sensing model. Therefore, the optimal placement of a BRN requires judicious design of transmitters and receivers as we do in this chapter.

6.5.2 Comparison with Monostatic Radar Network

In Fig. 6.12, we compare the vulnerability of H under the optimal placement of a BRN to that of a monostatic radar network (MRN) for a varying number of transmitters and receivers. The optimal placement strategy for a MRN is to minimize the maximum distance from a point on H to its closest MR. For fair comparison, we set the number of transmitters in the BRN equal to that of receivers, and also equal to the number of MRs in the MRN. We observe that the advantage of a BRN is significant, which demonstrates that the flexibility to place transmitters and receivers separately is highly beneficial for barrier coverage.

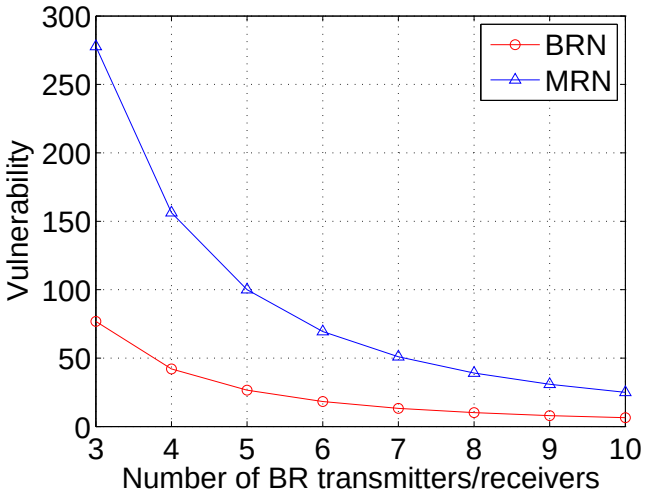


Figure 6.12: Impact of M for $M = N$ under optimal placement: BRN vs MRN.

6.6 Related Work

Radar has been extensively studied for decades [96]. However, radar sensor networks have garnered attention only in the past few years, largely driven by the emergence of cheaper and more compact radar sensors in place of conventionally expensive and bulky radar systems. For example, in [97], a platform has been successfully designed and built to integrate ultrawideband radars with mote-class sensor devices. The existing literature has studied different problems for radar sensor networks, including waveform design and diversity [98], radar scheduling [99], data management [100], for a variety of objectives, such as target detection [101] and localization [92]. In particular, BRs have also been considered in [92]. However, coverage problems of a radar sensor network have received very little attention. Recently, a novel Doppler coverage model has been introduced in [102] for a radar sensor network that exploits the Doppler effect. To our best knowledge, this work is the first to explore the barrier coverage of a network of BRs.

Numerous studies on sensor network coverage can be found in the literature [9].

Worst-case intrusion was first introduced in [93]. [93,94,103] have studied how to find the worst-case intrusion path for arbitrarily deployed sensors. [104,105] have considered adding sensors to improve the coverage of the worst-case intrusion path. Along another avenue, barrier coverage was first introduced in [84] and has attracted much research interests recently. [84,85] have studied the critical sensor density for barrier coverage under random deployment. The coverage of a barrier has been investigated using a quantitative metric in [12]. Barrier coverage of sensors with mobility have been considered in [86,106]. Barrier coverage for camera sensor networks has also been studied recently based on a novel full-view coverage model [107,108].

While most aforementioned studies are concerned with how to find the worst-case intrusion path or a barrier covered by sensors (if such a barrier exists) under an *existing* deployment of sensors, our work focuses on *where* to deploy sensors to cover a barrier such that the worst-case intrusion detectability is maximized. More importantly, the existing sensing models (particularly the widely used disk sensing model) are quite different from the Cassini oval sensing model of a BR, and the latter is further complicated by the coupling of sensing regions across multiple BRs.

6.7 Conclusion

Radar sensor networks have great potential in many applications, such as border surveillance and traffic monitoring. In this chapter, we studied the problem of deploying a network of BRs in a region for intruder detection. The optimal placement of BRs is highly non-trivial, since 1) the coverage region of a BR is characterized by a Cassini oval that presents complex geometry; 2) the coverage regions of different BRs are coupled and the network coverage is intimately related to the locations of all BR nodes. We showed that it is optimal to place BRs on the shortest barrier if it is also the shortest line segment that connects the left and right boundary of the

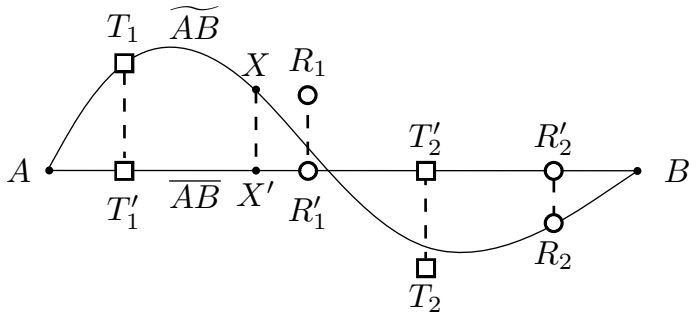


Figure 6.13: A snapshot of proof of Theorem 6.1.

region. Furthermore, we characterized the optimal placement order and spacing of BR nodes on a line segment, both of which present elegant balanced structures.

6.8 Appendix

6.8.1 Proof of Theorem 6.1

For any barrier $\widetilde{AB}$ and the optimal placement $\{\mathcal{T}, \mathcal{R}\}$ that minimizes $V(\widetilde{AB})$, we can construct a placement $\{\mathcal{T}', \mathcal{R}'\}$ for $\overline{AB}$ by moving each $T_i \in \mathcal{T}$ and each $R_j \in \mathcal{R}$ to their respective projections $T_i' \in \mathcal{T}'$ and $R_j' \in \mathcal{R}'$ on the line passing through A and B, respectively, as illustrated in Fig. 6.13. Then for any point $X' \in \overline{AB}$, there exists a point $X \in \widetilde{AB}$ whose projection on $\overline{AB}$ is X'. We can observe that $\|T_iX\| \geq \|T_i'X'\|$, $\forall i \in \mathcal{M}$ and $\|R_jX\| \geq \|R_j'X'\|$, $\forall j \in \mathcal{N}$. Then it follows that

$$I(X) = \min_{i\in\mathcal{M}, j\in\mathcal{N}} \|T_iX\|\|R_jX\| \geq \min_{i\in\mathcal{M}, j\in\mathcal{N}} \|T_i'X'\|\|R_j'X'\| = I'(X')$$

where we use $I'(X')$ to denote the detectability of X' under placement $\{\mathcal{T}', \mathcal{R}'\}$. Hence we have

$$V^*(\overline{AB}) \leq V'(\overline{AB}) = \max_{X'\in\overline{AB}} I'(X') \leq \max_{X\in\widetilde{AB}} I(X) = V^*(\widetilde{AB}) \tag{6.20}$$

where we use $V'(\overline{AB})$ to denote the vulnerability of $\overline{AB}$ under placement $\{\mathcal{T}', \mathcal{R}'\}$. Since $V^*(U)$ for a line barrier U increases as the length of U increases, and the shortcut barrier H is not longer than $\overline{AB}$, using (6.20) we have $V^*(H) \leq V^*(\overline{AB}) \leq V^*(\widetilde{AB})$.

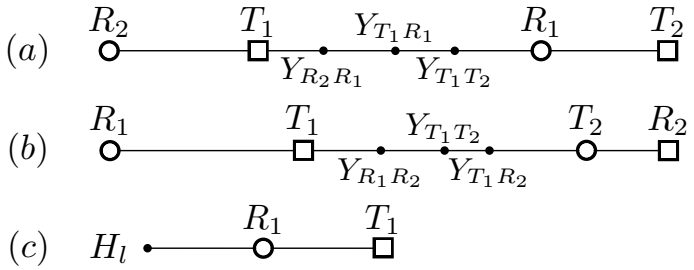


Figure 6.14: Snapshots of proof of Lemma 6.1.

6.8.2 Proof of Lemma 6.1

The main idea of the proof is to divide the line segment between each pair of neighbor nodes into intervals such that all the points on each interval have the *same* closest BR, and then we examine the detectability on each interval. We consider three possible cases of two neighbor nodes as follows.

Case 1: T_1 and R_1 (two nodes of different types)

As illustrated in Fig. 6.14 (a), for any point $X \in \overline{T_1R_1}$, its closest transmitter must be either T_1 or the leftmost transmitter T_2 on $\overline{R_1H_r}$, and its closest receiver must be either R_1 or the rightmost receiver R_2 on $\overline{H_lT_1}$. Suppose the closest BR for a point on $\overline{T_1Y_{R_2R_1}}$, $\overline{Y_{R_2R_1}Y_{T_1T_2}}$, $\overline{Y_{T_1T_2}R_1}$ is T_1-R_2, T_1-R_1, T_2-R_1, respectively. Then we observe that for $X \in \overline{T_1Y_{R_2R_1}}$, $I(X) = \|R_2X\|\|T_1X\|$ increases as X moves closer to $Y_{R_2R_1}$; for $X \in \overline{Y_{R_2R_1}Y_{T_1T_2}}$, $I(X) = \|T_1X\|\|XR_1\|$ increases as X moves closer to $Y_{T_1R_1}$; for $X \in \overline{Y_{T_1T_2}R_1}$, $I(X) = \|XR_1\|\|XT_2\|$ decreases as X moves closer to R_1. Therefore, $I(X)$ attains maximum on $\overline{T_1R_1}$ when $X = Y_{T_1R_1}$.

Case 2: T_1 and T_2 (two nodes of the same type)

As illustrated in Fig. 6.14 (b), for any point $X \in \overline{T_1T_2}$, its closest transmitter must be either T_1 or T_2, and its closest receiver must be either the rightmost receiver R_1 on $\overline{H_lT_1}$, or the leftmost receiver R_2 on $\overline{T_2H_r}$. Suppose the closest BR for a point

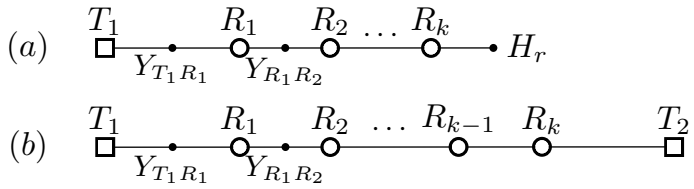


Figure 6.15: Snapshots of proof of Lemma 6.2.

on $\overline{T_1Y_{R_1R_2}}$, $\overline{Y_{R_1R_2}Y_{T_1T_2}}$, $\overline{Y_{T_1T_2}T_2}$ is T_1-R_1, T_1-R_2, T_2-R_2, respectively. Then we observe that for $X \in \overline{T_1Y_{R_1R_2}}$, $I(X) = \|T_1X\|\|R_1X\|$ increases as X moves closer to $Y_{R_1R_2}$; for $X \in \overline{Y_{R_1R_2}Y_{T_1T_2}}$, since X is on the left side of $Y_{T_1R_2}$ $I(X) = \|T_1X\|\|XR_2\|$ increases as X moves closer to $Y_{T_1T_2}$; for $X \in \overline{Y_{T_1T_2}T_2}$, $I(X) = \|XT_2\|\|XR_2\|$ decreases as X moves closer to T_2. Therefore, $I(X)$ attains maximum on $\overline{T_1T_2}$ when $X = Y_{T_1T_2}$.

Case 3: H_l and T_1 (two nodes including an end node)

As illustrated in Fig. 6.14 (c), for any point $X \in \overline{H_lT_1}$, its closest transmitter must be T_1, and its closest receiver must be the leftmost receiver R_1 on $\overline{T_1H_r}$. Hence the closest BR for a point on $\overline{H_lT_1}$ is T_1-R_1. Then we observe that for $X \in \overline{H_lT_1}$, $I(X) = \|XT_1\|\|XR_1\|$ increases as X moves closer to H_l, and hence attains maximum when $X = H_l$.

6.8.3 Proof of Lemma 6.2

The main idea of the proof is to sequentially determine the distance between each pair of neighbor nodes. In the following, we consider all three cases of an independent local order $\mathbf{S}_i$ with balanced spacing $\mathbf{D}_{\mathbf{S}_i}$ such that $V(\overline{Z_{\mathbf{S}_i}}) = c$.

Case 1: $\mathbf{S}_i = (T_1, R_1)$

Since the closest BR for $Y_{T_1R_1}$ is T_1-R_1, we have $I(Y_{T_1R_1}) = (\|T_1R_1\|/2)^2 = c$. Then it follows that $\|T_1R_1\| = 2\sqrt{c} = e_c^0$.

Case 2: $\mathbf{S}_i = (T_1, R_1, \cdots, R_k, H_r)$

Similar to case 1, we can show that $\|T_1R_1\| = 2\sqrt{c} = e_c^0$, as illustrated in

Fig. 6.15 (a). Since the closest BR for $Y_{R_1R_2}$ is T_1-R_1 or T_1-R_2, we have

$$I(Y_{R_1R_2}) = \|T_1Y_{R_1R_2}\|\|Y_{R_1R_2}R_2\| = \frac{\|T_1R_1\| + \|R_1R_2\|}{2}\frac{\|R_1R_2\|}{2} = c. \tag{6.21}$$

Then using (6.21) and $\|T_1R_1\| = 2\sqrt{c}$, we obtain a unique value $\|R_1R_2\| = e_c^1$. Following this argument recursively, using the values of $\|T_1R_1\|$, $\|R_1R_2\|$, $\cdots$, $\|R_{i-1}R_i\|$, we can obtain a unique value $\|R_iR_{i+1}\| = e_c^i$ such that $I(Y_{R_iR_{i+1}}) = c$ for $i = 2, \cdots, k-1$, and at last, using the values of $\|T_1R_1\|$, $\|R_1R_2\|$, $\cdots$, $\|R_{k-1}R_k\|$, we can obtain a unique value $\|R_kH_r\| = e_c^k/2$ such that $I(H_r) = c$.

Case 3: $\mathbf{S}_i = (T_1, R_1, \cdots, R_k, T_2)$

Similar to case 1, we can show that $\|T_1R_1\| = \|R_kT_2\| = e_c^0$, as illustrated in Fig. 6.15 (b). Then the closest BR for $Y_{R_1R_2}$ must be T_1-R_1 or T_1-R_2. Using the same argument as in case 2, we can obtain a unique value $\|R_1R_2\| = e_c^1$ such that $I(Y_{R_1R_2}) = c$. Then the closest BR for $Y_{R_{k-1}R_k}$ must be T_2-R_{k-1} or T_2-R_k. Following the above argument recursively, we can obtain that $\|R_{k-1}R_k\| = e_c^1$, then that $\|R_2R_3\| = e_c^2$ ... until that $\|R_{\frac{k}{2}}R_{\frac{k}{2}+1}\| = e_c^{\frac{k}{2}}$ when k is even, or $\|R_{\frac{k+1}{2}}R_{\frac{k+3}{2}}\| = e_c^{\frac{k-1}{2}}$ when k is odd.

6.8.4 Proof of Lemma 6.3

The proof is based on contradiction. Suppose there exists another placement $\mathbf{S}_i'$ with spacing $\mathbf{D}_{\mathbf{S}_i'}$ such that $V(Z_{\mathbf{S}_i'}) \le c$ and $Z_{\mathbf{S}_i} < Z_{\mathbf{S}_i'}$. We consider all three cases of an independent local order $\mathbf{S}_i$ as follows.

Case 1: $\mathbf{S}_i = (T_1, R_1)$

Since $I(Y_{T_1'R_1'}) = (\|T_1'R_1'\|/2)^2 \le c = I(Y_{T_1R_1}) = (\|T_1R_1\|/2)^2$, we have $\|T_1'R_1'\| \le \|T_1R_1\|$, which is a contradiction.

Case 2: $\mathbf{S}_i = (T_1, R_1, \cdots, R_k, H_r)$

Similar to case 1, we can show that $\|T_1'R_1'\| \le \|T_1R_1\|$. Using this and

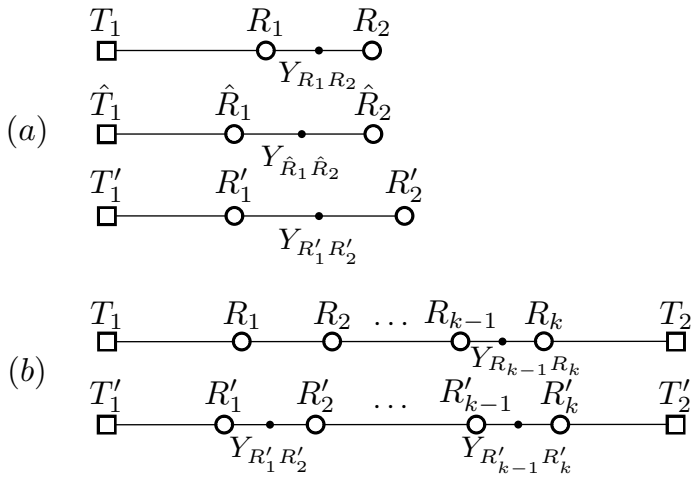


Figure 6.16: Snapshots of proof of Lemma 6.3.

$I'(Y_{R_1'R_2'}) \leq c = I(Y_{R_1R_2})$, we next show that $\|T_1'R_2'\| \leq \|T_1R_2\|$. Suppose $\|T_1'R_2'\| > \|T_1R_2\|$, as illustrated in Fig. 6.16 (a). Then we can find a placement $\{\hat{T}_1, \hat{R}_1, \hat{R}_2\}$ such that $\|\hat{T}_1\hat{R}_1\| = \|T_1'R_1'\|$ and $\|\hat{T}_1\hat{R}_2\| = \|T_1R_2\|$. We observe that $\|R_1'R_2'\| > \|\hat{R}_1\hat{R}_2\|$, and hence $\|T_1'Y_{R_1'R_2'}\| > \|\hat{T}_1Y_{\hat{R}_1\hat{R}_2}\|$ and $\|Y_{R_1'R_2'}R_2'\| > \|Y_{\hat{R}_1\hat{R}_2}\hat{R}_2\|$. We also observe that $\|\hat{T}_1\hat{R}_2\|/2 > \|Y_{\hat{R}_1\hat{R}_2}\hat{R}_2\| > \|Y_{R_1R_2}R_2\|$. Using these observations, we have

$$\begin{aligned} I'(Y_{R_1'R_2'}) &= \|T_1'Y_{R_1'R_2'}\|\|Y_{R_1'R_2'}R_2'\| \\ &> \|\hat{T}_1Y_{\hat{R}_1\hat{R}_2}\|\|Y_{\hat{R}_1\hat{R}_2}\hat{R}_2\| = \hat{I}(Y_{\hat{R}_1\hat{R}_2}) \\ &> \|T_1Y_{R_1R_2}\|\|Y_{R_1R_2}R_2\| = I(Y_{R_1R_2}) \end{aligned} \tag{6.22}$$

where $\hat{I}(Y_{\hat{R}_1\hat{R}_2})$ denotes the detectability of $Y_{\hat{R}_1\hat{R}_2}$ under placement $\{\hat{T}_1, \hat{R}_1, \hat{R}_2\}$. Then (6.22) contradicts that $I'(Y_{R_1'R_2'}) \leq I(Y_{R_1R_2})$. Thus we show that $\|T_1'R_2'\| \leq \|T_1R_2\|$.

Following the above argument recursively, using $\|T_1'R_{i-1}'\| \leq \|T_1R_{i-1}\|$ and $I(Y_{R_{i-1}R_i}) \geq I'(Y_{R_{i-1}'R_i'})$, we can show that $\|T_1'R_i'\| \leq \|T_1R_i\|$ for $i = 2, \cdots, k$, and at last, we can show that $\|T_1'H_r'\| \leq \|T_1H_r\|$, which is a contradiction.

Case 3: $\mathbf{S}_i = (T_1, R_1, \cdots, R_k, T_2)$

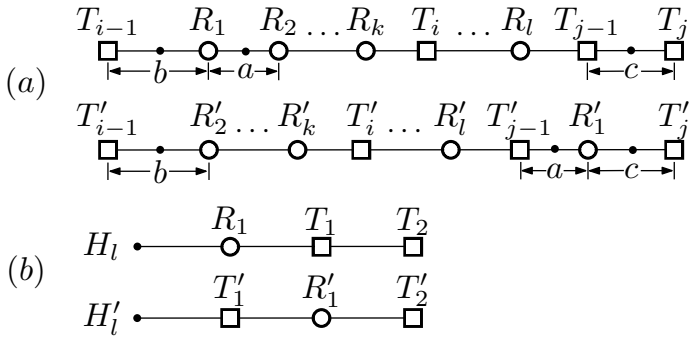


Figure 6.17: Snapshots of proof of Lemma 6.4.

Similar to case 1, we can show that $\|T'_1R'_1\| \leq \|T_1R_1\|$ and $\|R'_kT'_2\| \leq \|R_kT_2\|$. If the closest transmitter for $Y_{R'_1R'_2}$ is T'_1, similar to case 2, we can show that $\|T'_1R'_2\| \leq \|T_1R_2\|$. Next we show that the closest transmitter for $Y_{R'_1R'_2}$ cannot be T'_2. Suppose the closest transmitter for $Y_{R'_1R'_2}$ is T'_2. Then we have $\|R'_2T'_2\| \leq \|T'_1R'_1\| \leq \|T_1R_1\| = \|R_kT_2\|$. Following a similar argument as in case 2, using $\|R'_2T'_2\| \leq \|R_kT_2\|$ and $I'(Y_{R'_1R'_2}) \leq I(Y_{R_{k-1}R_k})$, we can show that $\|R'_1T'_2\| \leq \|R_{k-1}T_2\|$. Then it follows that $\|T'_1T'_2\| = \|T'_1R'_1\| + \|R'_1T'_2\| \leq \|T_1R_1\| + \|R_{k-1}T_2\| \leq \|T_1T_2\|$, which is a contradiction. Thus we show that the closest transmitter for $Y_{R'_1R'_2}$ cannot be T'_2, and hence must be T'_1.

Using the above argument, we can show that the closest transmitter for $Y_{R'_{k-1}R'_k}$ must be T'_2, and then similar to case 2, we can show that $\|R'_{k-1}T'_2\| \leq \|R_{k-1}T_2\|$. Following this argument recursively, we can show that $\|T'_1R'_3\| \leq \|T_1R_3\|$, and then that $\|R'_{k-2}T'_2\| \leq \|R_{k-2}T_2\| \cdots$ until we can show that $\|T'_1R'_j\| \leq \|T_1R_j\|$ and $\|R'_jT'_2\| \leq \|R_jT_2\|$ for $j = \frac{k+1}{2}$ when k is odd or $j = \frac{k}{2}+1$ when k is even. Then we have $\|T'_1T'_2\| = \|T'_1R'_j\| + \|R'_jT'_2\| \leq \|T_1R_j\| + \|R_jT_2\| = \|T_1T_2\|$, which is a contradiction.

6.8.5 Proof of Lemma 6.4

The main idea of this proof is as follows. For any placement $\mathcal{S}$ with an order $\mathbf{S}$ that does not satisfy the condition, we can find another placement $\mathcal{S}'$ with an order $\mathbf{S}'$ that satisfies the condition, such that any local vulnerable value under placement $\mathcal{S}'$,

and hence the vulnerability, must be *no greater than* that under placement $\mathcal{S}$. This implies that order $\mathbf{S}'$ is at least as good as order $\mathbf{S}$ (i.e., $f_c^{\mathbf{S}'} \geq f_c^{\mathbf{S}}$, $\forall c > 0$).

First we show that there exists an optimal order that satisfies (6.18a). Consider a placement $\mathcal{S}$ with an order $\mathbf{S}$ that does not satisfy (6.18a). Then we can always find some $i \neq j$ such that $n_i \geq 2$, $n_j = 0$, and $n_{j-1} > 0$ or $n_{j+1} > 0$. Without loss of generality, in the following we assume that $n_{j-1} > 0$, $i \neq 1, M+1$, and $j \neq 1, M+1$, while the other possible cases can be proved using a similar argument. Then $\mathbf{S}$ includes a local order $(T_{i-1}, R_1, R_2 \cdots, R_k, T_i)$ where $k \geq 2$ and a local order (R_l, T_{j-1}, T_j) (as illustrated in Fig. 6.17 (a)). We can find another placement $\mathcal{S}'$ with an order $\mathbf{S}'$ constructed from $\mathbf{S}$ by moving R_1 to being between T_{j-1} and T_j. We construct the spacing $\mathbf{D}_{\mathbf{S}'}$ from $\mathbf{D}_{\mathbf{S}}$ by setting $\|T'_{i-1}R'_2\| = \|T_{i-1}R_1\|$, $\|T'_{j-1}R'_1\| = \|R_1R_2\|$, $\|R'_1T'_j\| = \|T_{j-1}T_j\|$, while keeping the distance between each other two neighbor nodes unchanged, such that $\|H_lH_r\| = \|H'_lH'_r\|$. Then we have $I(Y_{R_1R_2}) \geq I'(Y_{T'_{j-1}R'_1})$, $I(Y_{T_{i-1}R_1}) \geq I'(Y_{T'_{i-1}R'_2})$, and $I(T_{j-1}T_j) \geq I'(R'_1T'_j)$. Furthermore, we can observe that any other local vulnerable value under placement $\mathcal{S}'$ must be no greater than that under placement $\mathcal{S}$. This implies that $\mathbf{S}'$ is at least as good as $\mathbf{S}$. Repeating the above construction, we can always find a order $\mathbf{S}'$ that satisfies (6.18a), which shows that there must exist an optimal order that satisfies (6.18a).

Next we show that among all the orders that satisfy (6.18a), there exist an optimal order that also satisfies (6.18b), and it is dividable. We consider two cases of an order $\mathbf{S}$ that satisfies (6.18a) as follows.

$$\mathbf{S} = (\overbrace{H_l, R_1, \underbrace{T_1}_{\mathbf{S}_2}, R_2, T_2}^{\mathbf{S}_1}, \cdots, \overbrace{T_{k-2}, \underbrace{R_{k-1}}_{\mathbf{S}_k}, T_{k-1}, T_k, \overbrace{R_k}^{\mathbf{S}_{k+1}}, T_{k+1}}^{\mathbf{S}_{k-1}}, \cdots, \overbrace{T_{M-1}, R_{M-1}, \underbrace{T_M}_{\mathbf{S}_{M+1}}, R_M, H_r}^{\mathbf{S}_M}) \tag{6.25}$$

Case 1: $M < N$

Since $\mathbf{S}$ satisfies (6.18a), we must have $n_i > 0$ for each $i = 1, \cdots, M+1$, and hence (6.18b) is also satisfied. Then $\mathbf{S}$ can be decomposed into independent local orders as

$$\mathbf{S} = (\overbrace{H_l, \cdots, \underbrace{T_1}_{\mathbf{S}_2}, \cdots, T_2}^{\mathbf{S}_1}, \cdots, \overbrace{T_{M-1}, \cdots, \underbrace{T_M}_{\mathbf{S}_{M+1}}, \cdots, H_r}^{\mathbf{S}_M}) \tag{6.23}$$

and hence is dividable.

Case 2: $M = N$

Suppose order $\mathbf{S}$ does not satisfy (6.18b). Then we must have $n_2 = 0$ and $n_i = 1$, $\forall i \neq 2$, or $n_M = 0$ and $n_i = 1$, $\forall i \neq M$. Suppose the former case holds with out loss of generality. Then for any placement $\mathcal{S}$ with order $\mathbf{S}$, we can find another placement $\mathcal{S}'$ constructed from $\mathcal{S}$ by swapping the locations of node R_1 and node T_1, such that $\mathcal{S}'$ has an order $\mathbf{S}'$ that satisfies (6.18b) (as illustrated in Fig. 6.17 (b)). Then we observe that any local vulnerable value under placement $\mathcal{S}'$ must be no greater than that under placement $\mathcal{S}$, which implies that order $\mathbf{S}'$ is at least as good as order $\mathbf{S}$. Thus there must exist an optimal order that satisfies both (6.18a) and (6.18b).

For an order $\mathbf{S}$ that satisfies (6.18), we consider two cases: 1) if $n_1 = 0$ and $n_i = 1$, $\forall i \neq 1$, then $\mathbf{S}$ can be decomposed into independent local orders as

$$\mathbf{S} = (\overbrace{H_l, T_1, \underbrace{R_1}_{\mathbf{S}_2}}^{\mathbf{S}_1}, \overbrace{T_2, R_2, T_3}^{\mathbf{S}_3}, \cdots, \overbrace{T_{M-1}, R_{M-1}, \underbrace{T_M}_{\mathbf{S}_{M+1}}, R_M, H_r}^{\mathbf{S}_M}) \tag{6.24}$$

and hence is dividable. Similarly, $\mathbf{S}$ is also dividable if $n_{M+1} = 0$ and $n_i = 1$, $\forall i \neq M+1$; 2) if $n_k = 0$ where $k \neq 1, 2, M, M+1$ and $n_i = 1$, $\forall i \neq k$, then $\mathbf{S}$ can be decomposed into independent local orders as in (6.25), and hence is dividable.

6.8.6 Proof of Theorem 6.3

By Lemma 6.4, there exists an optimal order among all the orders that satisfy (6.18), which are dividable. We observe that any order that satisfies (6.19) also satisfies (6.18). Therefore, it suffices to show that for any two orders $\mathbf{S}$ and $\mathbf{S}'$ that satisfy (6.18): 1) if $\mathbf{S}$ satisfies (6.19) and $\mathbf{S}'$ does not satisfy (6.19), then $f_c^{\mathbf{S}} \geq f_c^{\mathbf{S}'}$; 2) if $\mathbf{S}$ and $\mathbf{S}'$ both satisfy (6.19), then $f_c^{\mathbf{S}} = f_c^{\mathbf{S}'}$.

Let $g_c(n)$ denote the optimal value of problem (6.16) for an independent local order $\mathbf{S}_i$ with type (T, R^n, T) or (R, T^n, R) under the constraint $V(Z_{\mathbf{S}_i}) \leq c$. By Lemma 6.2 and Lemma 6.3, we have

$$g_c(n) = \begin{cases} 2\sum_{i=0}^{\frac{n}{2}-1} e_c^i + e_c^{\frac{n}{2}}, & \text{if } n \text{ is even} \\ 2\sum_{i=0}^{\frac{n+1}{2}-1} e_c^i, & \text{if } n \text{ is odd.} \end{cases} \tag{6.26}$$

Then it follows from (6.26) that

$$g_c(n+1) - g_c(n) = \begin{cases} e_c^{\frac{n}{2}}, & \text{if } n \text{ is even} \\ e_c^{\frac{n+1}{2}}, & \text{if } n \text{ is odd.} \end{cases} \tag{6.27}$$

Next we consider two cases of an order $\mathbf{S}$ that satisfies (6.18).

Case 1: $M < N$

We have shown in the proof of Lemma 6.4 that $\mathbf{S}$ can be decomposed into independent local orders as (6.23). Using Theorem 6.2 and (6.23), we have

$$f_c^{\mathbf{S}} = \sum_{i=2}^{M} g_c(n_i) + \frac{g_c(2n_1)}{2} + \frac{g_c(2n_{M+1})}{2}. \tag{6.28}$$

Without loss of generality, suppose $\mathbf{S}$ does not satisfy (6.19a). The case where $\mathbf{S}$ does not satisfy (6.19b) can be proved using a similar argument. Then there exist $i, j \in \{2, \cdots, M\}$ such that $n_i \geq n_j + 2$. We can find another order $\mathbf{S}'$ with

$\mathbf{N}_{\mathbf{S}'} = \{n'_1, \cdots, n'_{M+1}\}$ constructed from $\mathbf{N}_{\mathbf{S}}$ by setting $n'_i = n_i - 1$, $n'_j = n_j + 1$, and $n'_k = n_k$, $\forall k \neq i, j$. Using (6.28) and (6.27), we have

$$\begin{aligned} f_c^{\mathbf{S}} - f_c^{\mathbf{S}'} &= g_c(n_i) - g_c(n_i - 1) + g_c(n_j) - g_c(n_j + 1) \\ &= \begin{cases} e_c^{\frac{n_i}{2}}, & \text{if } n_i \text{ is even} \\ e_c^{\frac{n_i-1}{2}}, & \text{if } n_i \text{ is odd} \end{cases} - \begin{cases} e_c^{\frac{n_j}{2}}, & \text{if } n_j \text{ is even} \\ e_c^{\frac{n_j+1}{2}}, & \text{if } n_j \text{ is odd} \end{cases} \\ &\leq 0, \end{aligned}$$

which shows that $\mathbf{S}$ is not optimal. Furthermore, for any two orders $\mathbf{S}$ and $\mathbf{S}'$ that both satisfy (6.19), we can show that $f_c^{\mathbf{S}} = f_c^{\mathbf{S}'}$. Therefore, any order that satisfies (6.19) is optimal.

Case 2: $M = N$

Suppose $\mathbf{S}$ does not satisfy (6.19). Then we have $n_k = 0$ where $k \neq 1, 2, M, M+1$ and $n_i = 1$, $\forall i \neq k$. Using (6.25), we have

$$f_c^{\mathbf{S}} = (M-4)g_c(1) + 2g_c(2) + 2e_c^0. \tag{6.29}$$

Suppose another order $\mathbf{S}'$ satisfies (6.19) with $\mathbf{N}_{\mathbf{S}'} = \{n'_1, \cdots, n'_{M+1}\}$ such that $n'_1 = 0$ and $n'_i = 1$, $\forall i \neq 1$. Using (6.24), we have

$$f_c^{\mathbf{S}'} = (M-2)g_c(1) + g_c(2) + e_c^0. \tag{6.30}$$

It follows from (6.29) and (6.30) that

$$f_c^{\mathbf{S}} - f_c^{\mathbf{S}'} = g_c(2) + e_c^0 - 2g_c(1) = e_c^1 - e_c^0 < 0, \tag{6.31}$$

which shows that $\mathbf{S}$ is not optimal. Furthermore, for any two orders $\mathbf{S}$ and $\mathbf{S}'$ that both satisfy (6.19), we can show that $f_c^{\mathbf{S}} = f_c^{\mathbf{S}'}$. Therefore, any order that satisfies (6.19) is optimal.

Chapter 7

EXPLOITING DOPPLER EFFECT FOR COVERAGE IN RADAR SENSOR NETWORKS

7.1 Introduction

Wireless sensor networks have potential in a wide range of applications, such as border control, homeland security, and intruder detection. Traditional sensors in such applications operate by passively collecting natural radiation emitted or reflected by an object. In contrast to such passive sensors, a *radar* (RAdio Detection And Ranging) performs detection by actively emitting radio waves and receiving echoes reflected by objects. Due to this salient feature of active sensing, radars provide a number of benefits over traditional passive sensors. For example, an ultra-wideband (UWB) radar has a larger sensing range than either infrared or magnetic sensors. With the emergence of cheap and compact radar devices, it is becoming feasible to deploy a network of radar sensors working in concert[1] [97]. Indeed, the radar network has been recognized as a promising paradigm for numerous applications, including road traffic monitoring [83] and earthquake sensing [109].

One important issue for sensor networks is how well an object of interest (target) is monitored, referred to as the coverage problem [9]. Previous studies have mostly assumed that a sensor's capability of detecting a target is quantified by its distance to the target. As a result, the coverage region of a sensor is typically modeled as a disk or a sector. However, this distance-based coverage model is barely valid for radar sensors in the presence of *clutter*. Simply put, clutter refers to echoes reflected by undesired objects in the monitored environment, such as rocks, trees, and walls. The clutter can be by many orders stronger than the echo from a target, making the

[1]For brevity, we may use "radar", "sensor", and "radar sensor" interchangeably throughout this chapter.

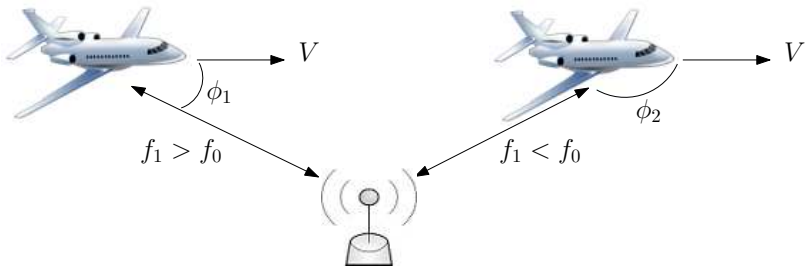


Figure 7.1: Illustration of Doppler frequency shift.

target detection very difficult, if not impossible.

Perhaps the most widely used method for detecting a target in the midst of strong clutter is by taking advantage of the Doppler effect. In many applications, the target of interest is a moving object such as an aircraft, ship, ground vehicle, or person, while most of the objects that cause clutter are stationary or moving slowly compared to the target. The *Doppler frequency shift* (DFS) is the frequency difference between the emitted and received radar signals due to the relative velocity between a radar and a moving target. As illustrated in Fig. 7.1, compared to the emitted frequency, the received frequency f_1 is higher than the emitted frequency f_0 when the target is approaching, and is lower f_0 when it is receding. Given the speed of a moving target V, $|f_1 - f_0|$ increases as angle ϕ_1 ($< 90°$) decreases, or as angle ϕ_2 ($> 90°$) increases. Therefore, *the amount of DFS is intimately related to the target's moving direction with respect to the radar*. Specifically, it is maximized when the target is moving directly toward or away from the radar, and diminishes as the angle between the moving direction and the direction toward the radar increases, until achieving zero when the target is moving at right angles to the radar. By exploiting the DFS, a radar can distinguish the target signal from clutter in a Doppler perspective.

As none of the existing sensor coverage models take into account the Doppler effect for moving target detection, we develop a novel *Doppler coverage* model to capture this feature of radar sensors. A target is Doppler-covered if and only if, for any

moving direction, the target is within some radar's sensing range, and furthermore, *the angle between the moving direction and the direction from the target to that radar is close enough to 0 or* π. In other words, if the target is Doppler-covered, then no matter which direction it is moving, there always exists some radar that can observe both a sufficiently high SNR and a sufficiently large DFS from its received target signal.

This Doppler coverage model differs significantly from the traditional distance-based coverage models, and gives rise to the following new challenges. First, the Doppler coverage of a target depends on the angular positions of all the radars (whose sensing ranges include the target) with respect to the target. This coupling in both distance and angle complicates the analysis of Doppler coverage. Second, a radar can contribute to the Doppler coverage of a target in two ways: *up-Doppler* coverage with positive DFS, and *down-Doppler* coverage with negative DFS. In particular, as illustrated in Fig. 7.2, the moving directions that are Doppler-covered by a radar constitute *two opposite angular ranges*. This geometric structure introduces additional complexity into the Doppler coverage problem.

With the proposed Doppler coverage model, it is natural to ask a fundamental question: *How to characterize the Doppler-covered regions for arbitrarily deployed radar sensors?* The answer to this question can be used to evaluate the coverage of sensors under random deployment. In practice, one efficient sensor deployment strategy is deterministic deployment, in which we can deploy sensors at desirable locations. Then another important problem is to *find the minimum possible number of radar sensors in a region such that the entire region is Doppler-covered*. The main thrust of this chapter is devoted to answering these two questions.

We summarize the main contributions of this chapter as follows. First, we develop a novel Doppler coverage model for moving target detection in radar sensor

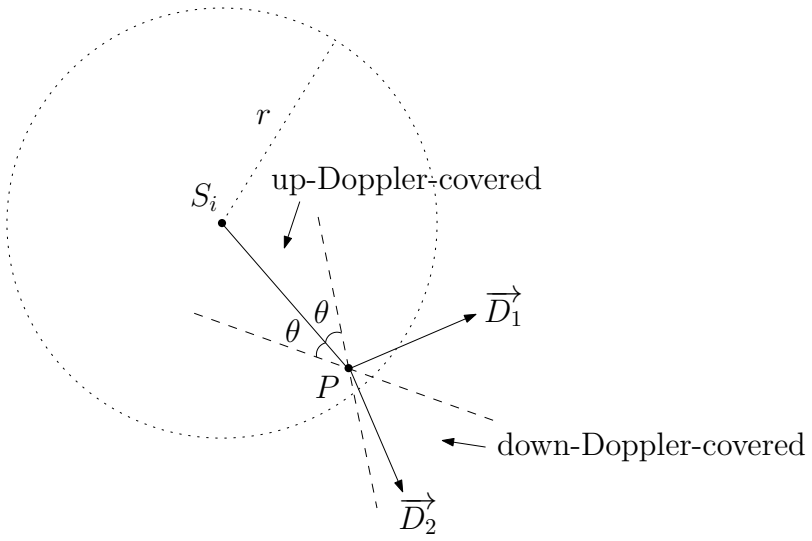


Figure 7.2: Illustration of Doppler coverage model.

networks by exploiting the Doppler effect. Next, we propose an efficient method to characterize the Doppler-covered regions for arbitrarily deployed radar sensors. Then we design an algorithm CSR for computing the minimum sensor density required to achieve Doppler coverage in a region under any polygonal deployment pattern, and further apply it to investigate the regular triangle deployment. To the best of our knowledge, this work is the first to exploit the Doppler effect for coverage in radar networks.

The rest of this chapter is organized as follows. Section 7.2 introduces the model of Doppler coverage. In Section 7.3, we characterize the Doppler-covered regions for arbitrary deployed radar sensors. We investigate in Section 7.4 the minimum radar density required to achieve Doppler coverage in a region under deterministic deployment. Section 7.5 provides some numerical results and Section 7.6 discusses related work. The chapter is concluded in Section 7.7.

7.2 Doppler Coverage Model

We consider a set of radar[2] sensors $\mathcal{S} \triangleq \{S_1, \cdots, S_n\}$ deployed in a *region of interest* F for detecting a target. For convenience, we also use S_i to denote the location (point) of a sensor S_i. We denote $\|UV\|$ as the (Euclidean) distance between two points U and V. We denote the angle between two vectors (directions) $\vec{v_1}$ and $\vec{v_2}$ as $\alpha(\vec{v_1}, \vec{v_2})$.

For a radar S_i, the received SNR from a target located at a point P is given by [96]

$$\text{SNR} = \frac{K}{\|S_i P\|^4} \tag{7.1}$$

where K denotes *radar constant* which reflects physical layer characteristics of the radar (e.g., transmit power, *radar cross section*[3], transmitter/receiver antenna power gains). The sensing range r of a radar can be determined based on (7.1) such that the SNR received by the radar is higher than certain threshold when the target is within its sensing range. We say that a point P is *covered* by a sensor S_i (in the sense of SNR or distance) if it is within the sensing range of S_i, i.e., $\|S_i P\| \leq r$.

Let V be the speed of a moving target, and ϕ be the angle at which the target is moving with respect to a radar. Then the DFS observed by the radar, denoted by Δf, can be expressed as

$$\Delta f = \frac{2V \cos\phi}{c} f_0 \tag{7.2}$$

where f_0 is the frequency of the emitted radar signal and c is the speed of light. We assume that the target's speed V is known (or is lower bounded by some known

[2]In this chapter, we consider monostatic radars each comprised of co-located radar transmitter and receiver, which is the typical radar configuration.

[3]Radar cross section measures the amount of radar signal energy reflected by an object due to its physical characteristics such as shape, size, and material.

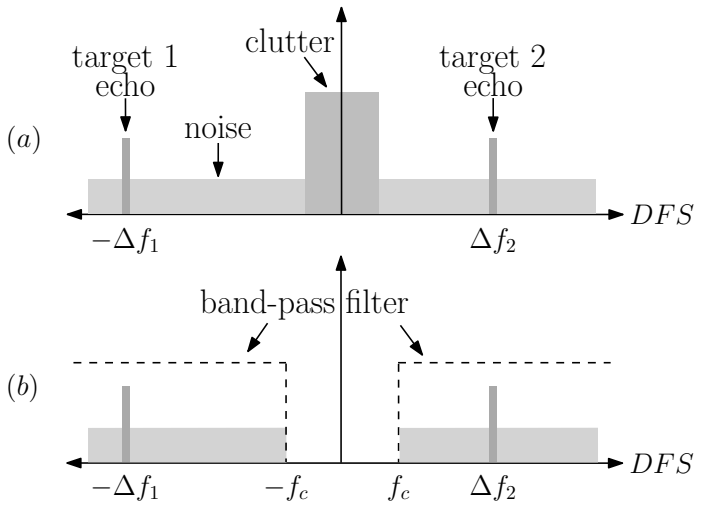


Figure 7.3: Illustration of clutter filter by Doppler frequency shift.

value). This information can be obtained from some general knowledge about the target (e.g., knowing that the target is a ground vehicle or a aircraft).

For ease of exposition, we assume that all radars have homogeneous physical characteristics such that they have the same radar constant. We also assume that radars share a set of orthogonal radio resources (e.g., different frequencies or time slots, orthogonal waveforms) through proper spatial reuse so as to avoid mutual interference. We further assume that the transmitted and reflected radar signals are *omni-directional*. Therefore, a radar's coverage region is a disk with radius r. However, our results in this chapter can be directly extended to the case of sector-based coverage region.

Every radar employs a filtering method to suppress clutter in the received signal (This method is also known as "moving target indication (MTI)" in radar literature). In particular, a band-pass filter with cutoff frequency f_c is applied to the received signal such that all the frequency components below f_c are removed while the rest are maintained (as illustrated in Fig. 7.3). We assume that the maximum moving speed of an object that causes clutter is known. Based on this knowledge,

the cutoff frequency f_c can be determined such that the clutter is guaranteed to be filtered out completely. Consequently, depending on the DFS Δf, the target signal may or may not exist in the filtered signal: it exists if and only if $\Delta f \geq f_c$. After the filtering, a typical signal detection method (e.g., energy detection or matched filter detection) is used to detect the target signal in the filtered signal. We assume that the target signal can be detected with certain required probability subject to certain false alarm probability, if and only if it exists in the filtered signal and the SNR is higher than certain threshold.

Definition 7.1. *(Doppler Coverage) A direction $\overrightarrow{D}$ from a point P is Doppler-covered by a sensor S_i if $\|S_iP\| \leq r$ (i.e., P is covered by S_i) and $\alpha(\overrightarrow{D}, \overrightarrow{PS_i}) \leq \theta$ or $\pi - \theta \leq \alpha(\overrightarrow{D}, \overrightarrow{PS_i}) \leq \pi$ where $\theta \in [0, \pi/2)$ is a pre-determined parameter called effective Doppler angle. A point P is Doppler-covered if any direction from P is Doppler-covered by some sensor.*

Given the target's speed V (or a lower bound of V), the effective Doppler angle θ can be determined based on (7.2) such that the target signal can be detected by a radar with certain required probability (after filtering) when the target's moving direction is Doppler-covered by that radar. For convenience, we say that a direction $\overrightarrow{D}$ from P *can be* Doppler-covered by a sensor S_i if it is Doppler-covered by S_i conditioned on that P is covered by S_i. For example, as illustrated in Fig. 7.2, $\overrightarrow{D_1}$ is not Doppler-covered by S_i while $\overrightarrow{D_2}$ is Doppler-covered by S_i. Similarly, we say that a point P can be Doppler-covered by a set of sensors $\mathcal{S}'$ if any direction from P can be Doppler-covered by some sensor in $\mathcal{S}'$. We say that a region is Doppler-covered if every point in it is Doppler-covered.

The Doppler coverage model is related to but much more complicated than traditional distance-based coverage models. In particular, Doppler coverage requires

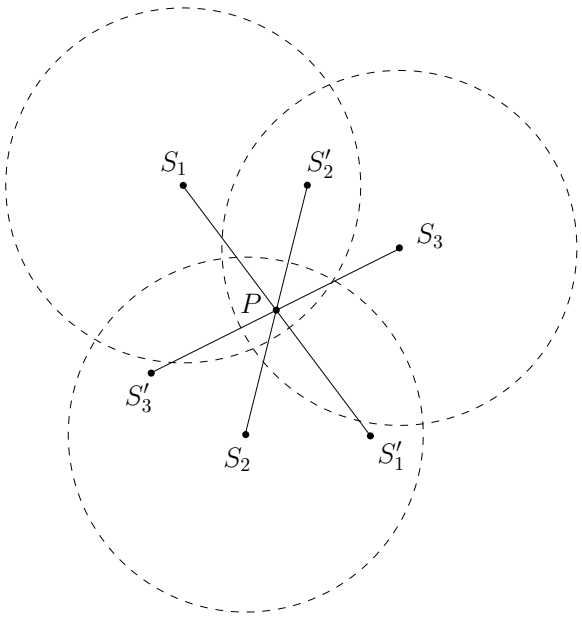


Figure 7.4: Illustration of constructing a coverage list.

not only that a point is covered by multiple sensors in the sense of distance, but also that any direction from that point is Doppler-covered by some of those sensors. As will be shown later, this complication gives rise to significant challenges for the analysis of Doppler coverage.

7.3 Characterization of Doppler Coverage

In this section, we propose an efficient method to characterize Doppler-covered regions for arbitrarily deployed sensors. Our results in this section can be used to evaluate the Doppler coverage of sensors under random deployment.

We denote a set of sensors that cover a point P as $\mathcal{S}_P$. We also use $\mathcal{S}_A$ to denote a set of sensors that cover a region A if each sensor in $\mathcal{S}_A$ covers all the points in A. For any point P, we can define a coverage list constructed from $\mathcal{S}_P = \{S_1, \ldots, S_m\}$ as follows. First, we construct an *image* point S_i' for each $S_i \in \mathcal{S}_P$ with respect to P such that $\overrightarrow{PS_i'}$ is opposite to $\overrightarrow{PS_i}$. We pick any $S_i \in \mathcal{S}_P$ (or its image S_i') and place it in the coverage list. Then we rotate $\overrightarrow{PS_i}$ around P in the counter-clockwise

(or clockwise) direction until it points to the same direction as $\overrightarrow{PS_j}$ (or $\overrightarrow{PS'_j}$) for some $S_j \in \mathcal{S}_P$, and place S_j (or S'_j, respectively) in the list next to S_i. We continue rotating until all the sensors in $\mathcal{S}_P$ and their images are placed in the list sequentially, denoted by $\mathbf{L}_P$. One can easily check that $\mathbf{L}_P$ is a circular list and can be written as $\mathbf{L}_P = (s_1, \ldots, s_m, s'_1, \ldots, s'_m)$, where s_i represents some sensor $S_j \in \mathcal{S}_P$ or its image S'_j, and s'_i represents the counterpart of s_i, i.e., the image S'_j or the sensor S_j, respectively. For example, in Fig. 7.4, the coverage list $\mathbf{L}_P = (S_1, S'_2, S_3, S'_1, S_2, S'_3)$ of a point P is constructed from the set of sensors $\mathcal{S}_P = \{S_1, S_2, S_3\}$ that cover P based on their angular positions with respect to P. For convenience, we may write $\mathbf{L}_P$ in a compact form without ambiguity as $\mathbf{L}_P = (s_1, \ldots, s_m)$ by including only half consecutive points and let $s_{n+1} = s'_1$. For example, $(S_1, S'_2, S_3, S'_1, S_2, S'_3)$ can be written for short as (S_1, S'_2, S_3), or (S'_2, S_3, S'_1), or (S'_3, S_1, S'_2). Also for convenience, we use $S(s_i)$ to denote the sensor represented by s_i if s_i represents a sensor, or the sensor whose image is represented by s_i if s_i represents an image. For example, $S(S_1) = S(S'_1) = S_1$.

Lemma 7.1. *A point P is Doppler-covered by $\mathcal{S}_P$ if and only if $\angle s_i P s_{i+1}$ is less than or equal to 2θ for any pair of neighbors s_i and s_{i+1} in the coverage list $\mathbf{L}_P$ constructed from $\mathcal{S}_P$.*

Proof: Suppose the condition holds. Then for any direction $\overrightarrow{D}$ from P, there exists a sensor S_i or an image S'_i such that $\alpha(\overrightarrow{PS_i}, \overrightarrow{D}) \leq \theta$ or $\pi - \alpha(\overrightarrow{PS_i}, \overrightarrow{D}) = \alpha(\overrightarrow{PS'_i}, \overrightarrow{D}) \leq \theta$. Thus P is Doppler-covered.

Suppose S_i and S_j are neighbors in $\mathbf{L}_P$ and $\angle S_i P S_j > 2\theta$. Then there exists a direction $\overrightarrow{D}$ from P which is parallel to the bisector of $\angle S_i P S_j$ such that $\alpha(\overrightarrow{PS_i}, \overrightarrow{D}) = \alpha(\overrightarrow{PS_j}, \overrightarrow{D}) > \theta$. It follows that $\alpha(\overrightarrow{PS_k}, \overrightarrow{D}) > \theta$ and $\pi - \alpha(\overrightarrow{PS_k}, \overrightarrow{D}) = \alpha(\overrightarrow{PS'_k}, \overrightarrow{D}) > \theta$ for any $S_k \in \mathcal{S}_P$. Thus P is not Doppler-covered. □

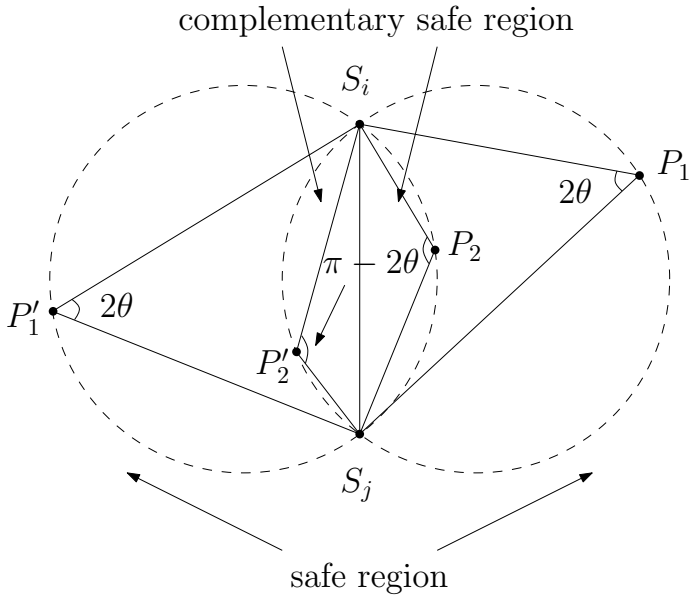


Figure 7.5: Illustration of safe region and complementary safe region.

Lemma 7.1 presents a sufficient and necessary condition for a point to be Doppler-covered. Our next goal is to characterize all the points in region F for which this condition holds. To this end, we introduce the following concepts. For two sensors S_i and S_j, we define *safe region* to be the area in which $\angle S_iPS_j \leq 2\theta$ for any point P, and *complementary safe region* to be the area in which $\pi - \angle S_iPS_j \leq 2\theta$ for any point P. Accordingly, the areas complementary to the safe region and complementary safe region are defined as *unsafe region* and *complementary unsafe region*, respectively.

Lemma 7.2. *For two sensors S_i and S_j, the safe region is the closed region complementary to the open region bounded by two arcs $\widehat{S_iS_j}$ and $\widehat{S_iS_j}'$ such that $\angle S_iPS_j = 2\theta$ for any $P \in \widehat{S_iS_j} \cup \widehat{S_iS_j}'$; the complementary safe region is the closed region bounded by two arcs $\widehat{S_iS_j}$ and $\widehat{S_iS_j}'$ such that $\angle S_iPS_j = \pi - 2\theta$ for any $P \in \widehat{S_iS_j} \cup \widehat{S_iS_j}'$.*

Proof: As illustrated in Fig. 7.5, the two arcs $\widehat{S_iS_j}$ and $\widehat{S_iS_j}'$ are segments of the circumscribed circles of triangles $\triangle S_iPS_j$ and $\triangle S_iP'S_j$ where $\angle S_iPS_j = \angle S_iP'S_j = 2\theta$ (or $\pi - 2\theta$), and P and P' are two points on different sides of the line segment

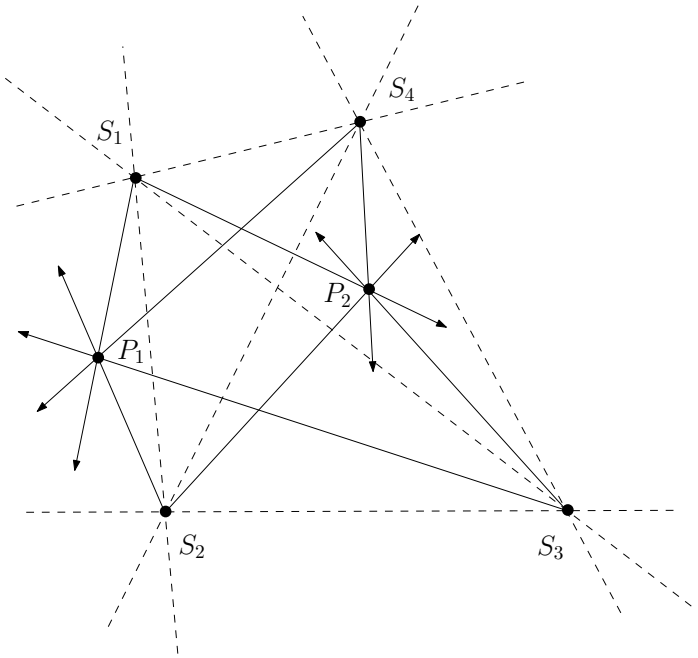


Figure 7.6: An example of two points covered by the same sensors but with different coverage lists.

$\overline{S_iS_j}$. The proof of the desired property follows from elementary geometry. □

For convenience, if s_i and s_j both represent sensors or both represent images, we denote by $R(s_i, s_j)$ the safe region of two sensors $S(s_i)$ and $S(s_j)$; otherwise, if s_i and s_j represent a sensor and an image, we denote by $R(s_i, s_j)$ the complementary safe region of $S(s_i)$ and $S(s_j)$. For example, $R(S_1, S_2)$ and $R(S_1', S_2')$ both denote the safe region of S_1 and S_2; $R(S_1', S_2)$ and $R(S_1, S_2')$ both denote the complementary safe region of S_1 and S_2. We use $R^c(s_i, s_j)$ to denote the region complementary to $R(s_i, s_j)$, i.e., the unsafe or complementary unsafe region of $S(s_i)$ and $S(s_j)$.

Lemma 7.3. *For a region A covered by $\mathcal{S}_A$, A is partitioned by all the lines each passing through a pair of sensors in $\mathcal{S}_A$ into a set of sub-regions such that all the points in a sub-region have the same coverage list constructed from $\mathcal{S}_A$.*

Proof: For any point $P \in A$, depending on which side of the line l pass-

ing through S_i and S_j in $\mathcal{S}_A$ that P is located at, the relative orders of S_i, S_j, S_i' and S_j' in the coverage list $\mathbf{L}_P$ constructed from $\mathcal{S}_A$ has either the form $\mathbf{L}_P = (\ldots, S_i, \ldots, S_j, \ldots, S_i', \ldots, S_j', \ldots)$ or $\mathbf{L}_P = (\ldots, S_i, \ldots, S_j', \ldots, S_i', \ldots, S_j, \ldots)$. Suppose A_i is a partition of A divided by all the lines each passing through a pair of sensors in $\mathcal{S}_A$. Then any two points P_1 and P_2 in A_i are located at the same side of each line, and hence the relative orders of any pair of sensors in $\mathcal{S}_A$ and their images are the same in $\mathbf{L}_{P_1}$ and $\mathbf{L}_{P_2}$. One can easily verify that there exists a unique coverage list which satisfies all these relative orders. Thus all the points in A_i have the same coverage list. □

Algorithm 6: Characterize Doppler-covered regions

input : region F, set of deployed sensors $\mathcal{S}$, sensing range r, effective Doppler angle θ

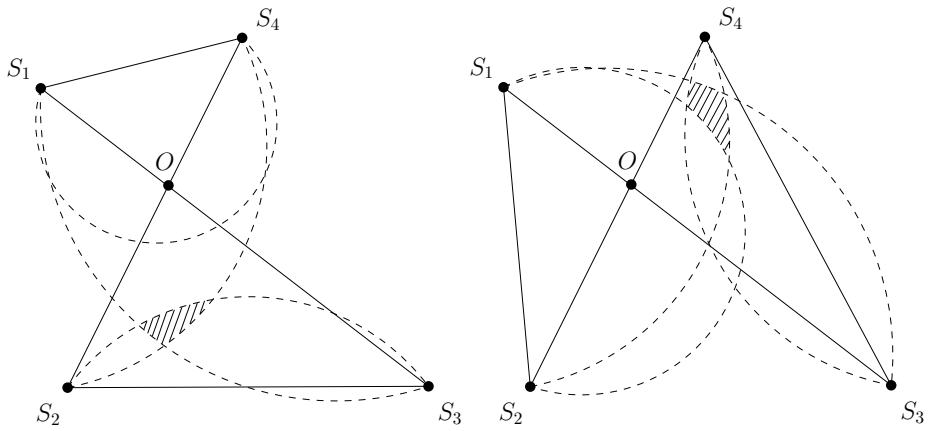
output: Doppler-covered regions in F

1 Partition F into a set of sub-regions $\mathcal{A}$ such that each sub-region $A_i \in \mathcal{A}$ is covered by the same set of sensors $\mathcal{S}_{A_i}$;
2 **foreach** *sub-region* $A_i \in \mathcal{A}$ **do**
3 Partition A_i into a set of sub-regions $\mathcal{A}_i$ such that all the points in a sub-region A_i^j have the same coverage list $\mathbf{L}_{A_i^j}$ constructed from $\mathcal{S}_{A_i}$;
4 **foreach** *sub-region* $A_i^j \in \mathcal{A}_i$ **do**
5 **return** *a Doppler-covered region* $A_i^j \cap (\cap_{k=1}^{n} R(s_k, s_{k+1}))$ *where* $\mathbf{L}_{A_i^j} = (s_1, \cdots, s_m)$;
6 **end**
7 **end**

Theorem 7.1. *The Doppler-covered regions in region F can be found by Algorithm 6.*

Proof: Since $\angle s_i P s_{i+1} = \angle s_i' P s_{i+1}'$ for any $i \in \{1, \ldots, n\}$, by Lemma 7.1, it suffices to check if $\angle s_i P s_{i+1}$ for any $i \in \{1, \ldots, n\}$ is less than or equal to 2θ. Then by Lemma 7.2, the intersection of the safe regions and complementary safe regions constructed in Algorithm 6 within the sub-region A_i^j is a Doppler-covered region. □

Following a similar argument based on Euler's formula as in [110], the number



(a) (b)

Figure 7.7: Illustration of Algorithm 6.

of sub-regions in $\mathcal{A}$ is $O(|\mathcal{S}|^4)$. The number of sub-regions in $\mathcal{A}_i$ is $O(|\mathcal{S}_{A_i}|^4)$ and the reason is as follows. The number of lines passing through a pair of points in a set of n points is at most $n(n-1)/2$. By induction, one can easily show that n lines can partition the plane into at most $n(n+1)/2+1$ regions. Therefore, all the lines passing through a pair of sensors in $|\mathcal{A}_i|$ sensors partition $\mathcal{S}_{A_i}$ into $O(|\mathcal{S}_{A_i}|^4)$ sub-regions.

Fig. 7.7 gives an example to illustrate the idea of line 5 in Algorithm 6. In Fig. 7.7 (a), we find the Doppler-covered region in a sub-region ΔS_2OS_3 where the coverage list is $\mathbf{L}_P = (S_1, S_3', S_2, S_4')$ for any $P \in \Delta S_2OS_3$. Hence, we construct the safe region for S_1 and S_4, the complementary safe region for S_1 and S_3, S_2 and S_3, S_2 and S_4, respectively. For brevity, we only depict half of the boundary of each constructed region. The intersection area (shaded) of the constructed safe regions and complementary safe regions within ΔS_2OS_3 is the Doppler-covered region within ΔS_2OS_3. Similarly, in Fig. 7.7 (b), we find the Doppler-covered region within a sub-region ΔS_3OS_4 with coverage list $\mathbf{L}_P = (S_1, S_2, S_4', S_3)$ for any $P \in \Delta S_3OS_4$. In this case, we construct the safe region for S_1 and S_2, the complementary safe region for S_1 and S_3, S_2 and S_4, S_3 and S_4, respectively.

7.4 Critical Sensor Density for Doppler Coverage under Deterministic Deployment

In this section, we consider deterministic deployment where sensors can be precisely deployed at any locations in the region F. We design an algorithm to compute the minimum sensor density required for the entire region to be Doppler-covered under a particular deployment pattern.

7.4.1 Problem Formulation

We consider a *polygonal* deployment pattern in which sensors are deployed on the vertices of the polygons of the same shape and size (e.g., triangles, rectangles, hexagons). Under such a deployment, there exists a *unit region* $U \subset F$ such that *U is Doppler-covered if and only if the entire region F is Doppler-covered*. For example, as illustrated in Fig. 7.8, a unit region for the regular triangle pattern can be a regular (equilateral) triangle, or can be just a partition of a regular triangle divided by the perpendicular bisectors of its three sides. This is because for any point $P \in F$ covered by any set of sensors $\mathcal{S}_P$, there exists a point $P' \in U$ covered by some set of sensors $\mathcal{S}_{P'}$ such that the relative positions of sensors in $\mathcal{S}_{P'}$ with respect to P' are the same as those in $\mathcal{S}_P$ with respect to P, and hence the Doppler coverage on P is equivalent to that on P'. As a result, it suffices to focus on the Doppler coverage in a unit region. In this chapter, as the optimal deployment pattern is difficult to find in general, we focus on any given polygonal deployment pattern.

We assume that region F is sufficiently large such that the boundary effect can be ignored. For a polygonal deployment pattern, the sensor density (or equivalently the number of sensors) is determined by the *size* of the polygons, which is a single-valued parameter denoted by l (e.g., the side length of a regular triangle in the regular triangle deployment). Clearly, the sensor density decreases as the size of the polygons increases. Given a polygonal deployment pattern, an effective Doppler angle θ, and

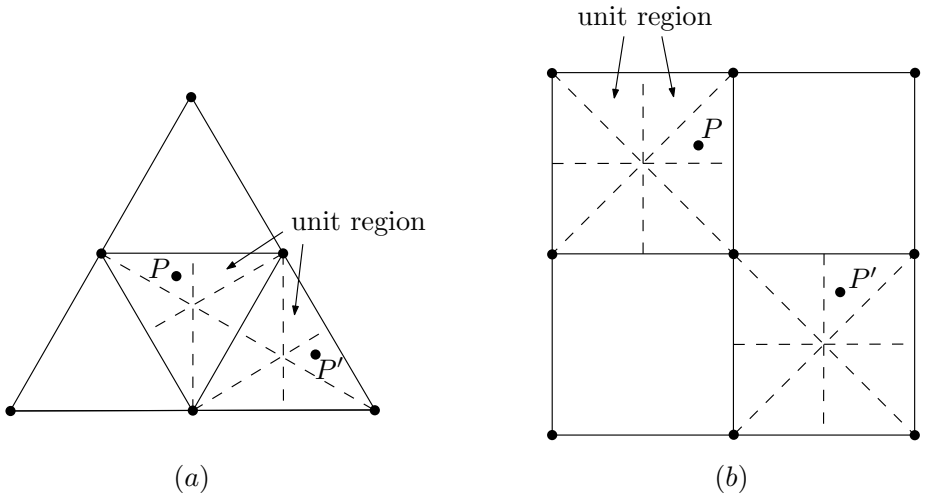


Figure 7.8: (a) Regular triangle deployment and (b) square deployment.

a sensing range r, our goal is to find the minimum sensor density, or equivalently the maximum size $l^*(\theta, r)$ of the polygons, such that F is Doppler-covered. It can be easily shown that, given any θ, $l^*(\theta, r)$ is a one-to-one increasing function of the sensing range r. Therefore, it suffices to find the minimum sensing range $r^*(\theta, l)$ required to Doppler-cover F as a function of θ and l, from which we can obtain $l^*(\theta, r)$ as the inverse function $(r^*)^{-1}(\theta, r)$ where θ is a parameter. For brevity, we will use r^* instead of $r^*(\theta, l)$.

We should note that it is highly non-trivial to find the critical sensing range r^* based on the Doppler coverage model. An intuitive approach is to first find the minimum sensing range required to Doppler-cover each point P in U, and then r^* can be found as the maximum of those minimum sensing ranges for all the points in U. However, since there are infinitely many points in U, this approach is not computationally feasible. In the next subsection, we will design an algorithm that can find the exact value of r^* analytically with finite computational complexity.

Algorithm 7: CSR (Compute the Critical Sensing Range)

input : polygonal deployment pattern, effective Doppler angle θ
output: critical sensing range r^*

Choose a unit region U;
// Phase 1: greedy search;
$\widetilde{\mathcal{S}} \leftarrow \varnothing$;
repeat
  $S^+ \leftarrow \arg\min_{S_i \in \mathcal{S} \setminus \widetilde{\mathcal{S}}} d_{S_i}^{\max}(U)$;
  $\widetilde{\mathcal{S}} \leftarrow \widetilde{\mathcal{S}} \cup S^+$;
until *U can be Doppler-covered by $\widetilde{\mathcal{S}}$*;
$\overline{r} \leftarrow d_{S^+}^{\max}(U)$;
$\mathcal{S}' \leftarrow \{S_i \in \mathcal{S} \setminus \widetilde{\mathcal{S}} | d_{S_i}^{\min}(U) \leq \overline{r}\}$;
$\underline{r} \leftarrow \min_{S_i \in \mathcal{S}' \cup S^+} d_{S_i}^{\min}(U)$;
$\overline{\mathcal{S}} \leftarrow \widetilde{\mathcal{S}} \cup \mathcal{S}'$;
$\underline{\mathcal{S}} \leftarrow \{S_i \in \widetilde{\mathcal{S}} | d_{S_i}^{\max}(U) \leq \underline{r}\}$;
$r^* \leftarrow \underline{r}$;
// Phase 2: divide and conquer;
Partition U into a set of sub-regions $\mathcal{U}$ such that all the points in each sub-region $U_i \in \mathcal{U}$ have the same coverage list $\mathbf{L}_{U_i}$ constructed from $\overline{\mathcal{S}}$;
foreach $U_i \in \mathcal{U}$ **do**
  Split $\mathbf{L}_{U_i}$ into partial coverage lists $\mathbf{L}_{U_i}^1, \cdots, \mathbf{L}_{U_i}^k$ such that $\mathbf{L}_{U_i}^j = (s_j, \cdots, s_{j+1})$ for any $j \in \{1, \cdots, k\}$ where $S(s_j) \in \underline{\mathcal{S}}$ for any $j \in \{1, \cdots, k\}$ and $k = |\underline{\mathcal{S}}|$;
  foreach $j \in \{1, \cdots, k\}$ **do**
    $r^* \leftarrow \max\{r^*, \text{CSRP}\ (U_i, \mathbf{L}_{U_i}^j)\}$;
  end
end
return r^*;

7.4.2 Algorithm Design

In this subsection, we design an algorithm named CSR for computing the Critical Sensing Range r^*. As described in Algorithm 7, CSR consists of two phases. In the first phase, CSR finds a lower bound and an upper bound of r^* using a greedy search. In the second phase, CSR takes a *divide and conquer* approach: it first partitions the unit region into sub-regions based on the bounds found in the first phase; then for each sub-region, it finds the minimum sensing range required to Doppler-cover a

particular range of directions from any point in that sub-region. In the following, we describe the design of CSR in details.

In the first phase, CSR constructs a set of sensors $\widetilde{\mathcal{S}}$ by iteratively including the sensor that requires the minimum sensing range to cover *all* the points in a unit region U, until U *can be* Doppler-covered by $\widetilde{\mathcal{S}}$. Recall that region U can be Doppler-covered by $\widetilde{\mathcal{S}}$ if it is Doppler-covered by $\widetilde{\mathcal{S}}$ conditioned on that U is covered by $\widetilde{\mathcal{S}}$ (i.e., when the sensing range r is sufficiently large). We use $d_{S_i}^{\max}(A) \triangleq \max_{P\in A} \|S_iP\|$ to denote the minimum sensing range required for sensor S_i to cover all the points in a region A. To check if region U can be Doppler-covered by $\widetilde{\mathcal{S}}$, we use a method similar to Algorithm 6: we first partition U into a set of sub-regions $\mathcal{U}$ such that all the points in each sub-region $U_i \in \mathcal{U}$ have the same coverage list $\mathbf{L}_{U_i}$ constructed from $\widetilde{\mathcal{S}}$, and then U can be Doppler-covered by $\widetilde{\mathcal{S}}$ if and only if $U_i \cap R^c(s_j, s_{j+1}) = \varnothing$ for each pair of neighbors s_j and s_{j+1} in each $\mathbf{L}_{U_i}$.

Using $\widetilde{\mathcal{S}}$ found above, CSR finds a lower bound $\underline{r}$ and an upper bound $\overline{r}$ of r^* (i.e., $\underline{r} \le r^* \le \overline{r}$). Using these bounds, CSR further finds $\underline{\mathcal{S}}$ as a set of sensors that *must* cover *all* points in U when the sensing range r is equal to r^* (i.e., $r = r^*$), and $\overline{\mathcal{S}}$ as a set of sensors such that any sensor *not* included in $\overline{\mathcal{S}}$ *must not* cover *any* point in U when $r = r^*$. In other words, $\overline{\mathcal{S}}$ consists of *all* the sensors that *may* cover *some* points in U when $r = r^*$ (i.e., may contribute to the Doppler coverage of some points in U).

In the second phase, based on $\underline{\mathcal{S}}$ and $\overline{\mathcal{S}}$ obtained in the first phase, U is partitioned into a set of sub-regions $\mathcal{U}$ such that all the points in each sub-region $U_i \in \mathcal{U}$ have the same coverage list $\mathbf{L}_{U_i}$ constructed from $\overline{\mathcal{S}}$. Then each coverage list $\mathbf{L}_{U_i}$ is split into k *partial* coverage lists $\mathbf{L}_{U_i}^1, \cdots, \mathbf{L}_{U_i}^k$ (which are not circular), where k is the number of sensors in $\underline{\mathcal{S}}$, such that for any $j \in \{1, \cdots, k\}$, $\mathbf{L}_{U_i}^j$ starts from some s_j with $S(s_j) \in \underline{\mathcal{S}}$ and ends at the *next* s_{j+1} with $S(s_{j+1}) \in \underline{\mathcal{S}}$ along

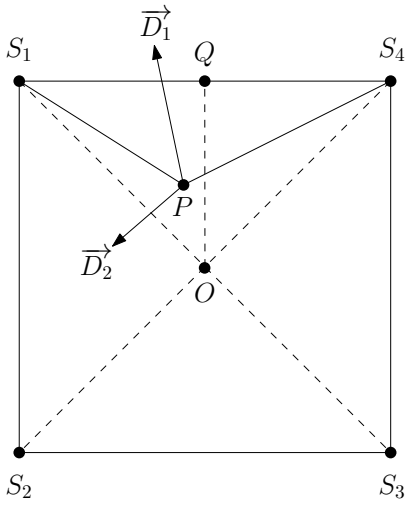


Figure 7.9: Illustration of Algorithm 8.

the list $\mathbf{L}_{U_i}$. For example, if $\overline{\mathcal{S}} = \{S_1, S_2, \cdots, S_{10}\}$, $\underline{\mathcal{S}} = \{S_2, S_5, S_6, S_9\}$, and $\mathbf{L}_{U_1} = (S_3, S_4', S_2, S_8, S_6', S_1, S_7', S_5, S_{10}, S_9')$, then $\mathbf{L}_{U_i}$ can be split into $\mathbf{L}_{U_1}^1 = (S_2, S_8, S_6')$, $\mathbf{L}_{U_1}^2 = (S_6', S_1, S_7', S_5)$, $\mathbf{L}_{U_1}^3 = (S_5, S_{10}, S_9')$, $\mathbf{L}_{U_1}^4 = (S_9', S_3', S_4, S_2')$.

Input with a region A and a partial coverage list $(s_1, \cdots, s_m)$, a subroutine named CSRP is used to find the minimum sensing range required to Doppler-cover any direction (if there is any) from any point in A that is *between* $\overrightarrow{Ps_j}$ and $\overrightarrow{Ps_{j+1}}$ and that *cannot* be Doppler-covered by $S(s_1)$ or $S(s_m)$. For example, as illustrated in Fig. 7.9, CSRP$[\Delta S_1 O S_4, (S_1, S_3', S_2', S_4)]$ outputs the minimum sensing range under which any direction from any point $P \in \Delta S_1 O S_4$ that is between $\overrightarrow{PS_1}$ and $\overrightarrow{PS_4}$ and that cannot be Doppler-covered by S_1 or S_4 (e.g., $\overrightarrow{D_1}$ but *not* $\overrightarrow{D_2}$), is Doppler-covered by S_2 or S_3. Then CSR outputs the maximum among $\underline{r}$ and the outputs of executing CSRP for each sub-region in $\mathcal{U}$ and the corresponding partial coverage list.

As described in Algorithm 8, CSRP also follows a divide and conquer approach. CSRP first checks if any direction from any $P \in A$ that is between $\overrightarrow{Ps_1}$ and $\overrightarrow{Ps_m}$ can be Doppler-covered by $S(s_1)$ or $S(s_m)$, which holds if and only if $A \cap R^c(s_1, s_m) = \varnothing$. If it holds, CSRP outputs 0; otherwise, CSRP partitions A (if necessary) to a set of

sub-regions $\mathcal{A}$ such that for each sub-region $A_i \in \mathcal{A}$, there exists some sensor $S(s_k)$ closest to each point in A_i among the sensors $S(s_2), \cdots, S(s_{n-1})$. This partition can be done by constructing the perpendicular bisectors of all the line segments (if necessary) each connecting a pair of sensors in $S(s_2), \cdots, S(s_{n-1})$. For example, as illustrated in Fig. 7.9, $\Delta S_1 O S_4$ is partitioned by the perpendicular bisector of $\overline{S_2 S_3}$ (which is the line passing through O and Q) such that any point in $\Delta S_1 O Q$ is closer to S_2 than to S_3. After the partition, CSRP outputs the maximum among the outputs of executing CSRP for each sub-region in $\mathcal{A}$ and the same partial coverage list.

If the input region A does not need to be partitioned, then CSRP outputs the maximum of $d^{\max}_{S(s_k)}(A \cap R^c(s_1, s_m))$ and the outputs of CSRP$[A \cap R^c(s_1, s_m), (s_1, \cdots, s_k)]$ and CSRP$[A \cap R^c(s_1, s_m), (s_k, \cdots, s_m)]$. Recall that $S(s_k)$ is the closest sensor to any point in A among $S(s_2), \cdots, S(s_{n-1})$. This is to guarantee that the sensing range is sufficiently large such that any direction from any point $P \in A \cap R^c(s_1, s_m)$ that is between $\overrightarrow{Ps_1}$ and $\overrightarrow{Ps_m}$ and that cannot be Doppler-covered by $S(s_1)$ or $S(s_m)$, is Doppler-covered by some sensor in $S(s_2), \cdots, S(s_{n-1})$.

Next we prove the correctness of CSR.

Theorem 7.2. *The critical sensing range r^* can be found by CSR.*

Proof: First we show that CSR can correctly compute $\underline{r}$, $\overline{r}$, $\underline{\mathcal{S}}$, and $\overline{\mathcal{S}}$ as defined. Due to the construction of $\widetilde{\mathcal{S}}$ in line 3 to 6, when the sensing range r is equal to $d^{\max}_{S^+}(U)$ where S^+ is the last sensor included in $\widetilde{\mathcal{S}}$ (i.e., the S^+ obtained in line 5 during the last iteration of the **repeat** loop), U is Doppler-covered by $\widetilde{\mathcal{S}}$. This implies that $r^* \leq d^{\max}_{S^+}(U) = \overline{r}$. Due to the value assignments of $\mathcal{S}'$ and $\overline{\mathcal{S}}$ in line 9 and 11, we observe that any S_i for which $d^{\min}_{S_i}(U) \leq \overline{r}$ must be in $\overline{\mathcal{S}}$. Therefore, any $S_i \in \mathcal{S} \setminus \overline{\mathcal{S}}$ must satisfy $d^{\min}_{S_i}(U) > \overline{r} \geq r^*$, and hence does not cover any point in U when $r = r^*$. Let $\mathcal{S}'' = \{S_i | d^{\min}_{S_i}(U) < \min_{S_j \in \mathcal{S}' \cup S^+} d^{\min}_{S_j}(U)\}$ be the set of all

Algorithm 8: CSRP (Compute the Critical Sensing Range for a Partial coverage list)

input : region A, partial coverage list $\mathbf{L} = (s_1, \cdots, s_m)$
output: critical sensing range r' for A and $\mathbf{L}_A$

```
1  r' ← 0;
2  if A ∩ R^c(s_1, s_m) ≠ ∅ then
3      Ŝ ← {S(s_2), ··· , S(s_{n−1})};
4      if ∄S(s_k) ∈ Ŝ such that ‖PS(s_k)‖ ≤ ‖PS(s_i)‖ for any P ∈ A and any s_i ∈ Ŝ
       then
5          Partition A into a set of sub-regions 𝒜 such that for any sub-region
           A_i ∈ 𝒜, ∃S(s_k) ∈ Ŝ such that ‖PS(s_k)‖ ≤ ‖PS(s_i)‖ for any P ∈ A_i and any
           S(s_i) ∈ Ŝ;
6          foreach A_i ∈ 𝒜 do
7              r' ← max{r', CSRP(A_i, L_A)};
8          end
9      end
10     Split L into two partial coverage lists L^1_A = (s_1, ··· , s_k), L^2_A = (s_k, ··· , s_m);
11     r' ← max{d^max_{S(s_k)}(A ∩ R^c(s_1, s_m)),
       CSRP(A ∩ R^c(s_1, s_m), L^1_A),CSRP(A ∩ R^c(s_1, s_m), L^2_A)};
12 end
13 return r';
```

the sensors that cover some points in U when $r = \min_{S_j \in \mathcal{S}' \cup S^+} d^{\min}_{S_j}(U)$. We observe that $\mathcal{S}'' \subseteq \widetilde{\mathcal{S}} \setminus \{S^+\}$. Then U cannot be Doppler-covered by $\mathcal{S}''$ since it cannot be Doppler-covered by $\widetilde{\mathcal{S}} \setminus \{S^+\}$. This implies that $r^* \geq \min_{S_j \in \mathcal{S}' \cup S^+} d^{\min}_{S_j}(U) = \underline{r}$. Then it follows from the value assignment of $\underline{\mathcal{S}}$ in line 12 that $\underline{\mathcal{S}}$ is a set of sensors that must cover all the points in U when $r = r^* \geq \underline{r}$.

Let $\tilde{r}$ be the actual output of CSR. To show that $\tilde{r} = r^*$, it suffices to show two parts: 1) U is Doppler-covered when $r = \tilde{r}$; 2) U is not Doppler-covered when $r < \tilde{r}$.

Part 1) can be shown as follows. First we observe that CSR can always terminate. The only possible case in which CSR does not terminate is when CSRP is input with a region A and a partial coverage list (s_1, s_2) of length 2 while we have that

$A \cap R^c(s_1, s_2) \neq \varnothing$. However, this cannot occur since s_1 and s_2 must be neighbors in the coverage list of any point in A constructed from $\overline{\mathcal{S}}$, and $A \subseteq U$ can be Doppler-covered by $\overline{\mathcal{S}}$. Next we observe that the output r' of any execution of CSRP is set to guarantee that all the points in $A \cap R^c(s_1, s_m)$ are covered by $S(s_k)$ when $r = r'$. Since the recursive executions of CSRP are input with region $A \cap R^c(s_1, s_m)$ and partial coverage lists $(s_1, \cdots, s_k)$ and $(s_k, \cdots, s_m)$, by induction, for any execution of CSRP input with region A and $(s_1, \cdots, s_m)$, all the points in A are covered by both $S(s_1)$ and $S(s_m)$ when $r = r'$. According to the design of CSR and CSRP, any direction $\overrightarrow{D}$ from any point $P \in U$ is found to be Doppler-covered by $S(s_1)$ or $S(s_m)$ in some execution of CSRP. Therefore, $\overrightarrow{D}$ must be Doppler-covered when $r = \tilde{r}$.

Next we show part 2). Suppose $r < \tilde{r}$. If $\tilde{r} = \underline{r}$, the desired result directly follows from $r < \tilde{r} = \underline{r}$. Otherwise, we have $\tilde{r} > \underline{r}$, and $\tilde{r}$ must be equal to $d^{\max}_{S(s_k)}(A \cap R^c(s_1, s_m))$ due to some execution of CSRP$[A, (s_1, \cdots, s_m)]$. Then there must exist a point $P \in (A \cap R^c(s_1, s_m))$ with $\|PS(s_k)\| = d^{\max}_{S(s_k)}(A \cap R^c(s_1, s_m)) = \tilde{r}$, and a direction $\overrightarrow{D}$ from P that is between $\overrightarrow{Ps_1}$ and $\overrightarrow{Ps_m}$ and that cannot be Doppler-covered by $S(s_1)$ or $S(s_m)$. We can see that $\overrightarrow{D}$ also cannot be Doppler-covered by any $S_i \in \overline{\mathcal{S}} \setminus \widehat{\mathcal{S}}$. Due to the partition operation in CSRP, we must have $r < \tilde{r} = \|PS(s_k)\| \leq \|PS_i\|$ for any $S_i \in \widehat{\mathcal{S}}$. Thus P is not covered by any $S_i \in \widehat{\mathcal{S}}$ and hence $\overrightarrow{D}$ is not Doppler-covered by any $S_i \in \widehat{\mathcal{S}}$. Therefore, $\overrightarrow{D}$ is not Doppler-covered by any $S_i \in \overline{\mathcal{S}}$. However, since $P \in A \subseteq U$ can be Doppler-covered by $\overline{\mathcal{S}}$, there must exist some $S_j \in \widehat{\mathcal{S}}$ such that P is covered by S_j and $\overrightarrow{D}$ is Doppler-covered by S_j when $r = r^*$. Then we have that $r^* \geq \|PS_j\| \geq \|PS(s_k)\| = \tilde{r} > r$, which implies that P is not covered by any $S_i \in \mathcal{S} \setminus \overline{\mathcal{S}}$ and hence $\overrightarrow{D}$ is not Doppler-covered by any $S_i \in \mathcal{S} \setminus \overline{\mathcal{S}}$. Therefore, $\overrightarrow{D}$ is not Doppler-covered by any sensor, which shows that U is not Doppler-covered. □

We have a few remarks on the complexity of CSR. Following a similar argument

as in Section 7.3, the complexity of checking if U can be Doppler-covered by $\widetilde{\mathcal{S}}$ once in line 7 is $O(|\widetilde{\mathcal{S}}|^5)$, and hence the complexity of the **repeat** loop is $|\widetilde{\mathcal{S}}| \times O(|\widetilde{\mathcal{S}}|^5) = O(|\widetilde{\mathcal{S}}|^6)$. Since we have $|\mathcal{U}| = O(|\overline{\mathcal{S}}|^4)$, the total number of executions of CSRP in line 19 is $|\mathcal{U}| \times |\underline{\mathcal{S}}| = O(|\overline{\mathcal{S}}|^4|\underline{\mathcal{S}}|)$. We can see that the complexity of CSRP can be upper bounded by some function of the length of the input partial coverage list, which determines the maximum possible number of recursive executions of CSRP involved in one execution of CSRP. As the input θ decreases, both $|\underline{\mathcal{S}}|$ and $|\overline{\mathcal{S}}|$ will increase, and the length of a partial coverage list obtained in line 14 is on average $|\overline{\mathcal{S}}|/|\underline{\mathcal{S}}| + 1$.

7.4.3 Case Study: Regular Triangle Deployment

In this subsection, as an illustrative example to show how CSR works, we use CSR to find the critical sensing range r^* under the regular triangle deployment. As discussed earlier, it suffices to consider the Doppler coverage in a unit region $U = \triangle COQ$ as shown in Fig. 7.10. For any $\theta \in [0, \pi/2)$, we observe that $l \leq r^*$. This is because that when $r < l$, there always exists a point P sufficiently close to a sensor such that P is only covered by that sensor, and hence is not Doppler-covered. In the following, for ease of exposition, we investigate the case $\theta \in [\pi/6, \pi/2)$.

Consider the case $\theta \in [\pi/3, \pi/2)$ as illustrated in Fig. 7.10(a). In this case, r^* can be found as follows without using CSR. When $r = l$, U is covered by $\{A, B, C\}$ with coverage list (A, B', C). Since the complementary safe regions of A and B, B and C, A and C, respectively, all include U, U is Doppler-covered. This implies that $r^* \leq l$. Then, using that $l \leq r^*$, we conclude that $r^* = l$.

Consider the case $\theta \in [\pi/4, \pi/3)$ as illustrated in Fig. 7.10(b). Using CSR, we obtain that $\underline{r} = l, \underline{\mathcal{S}} = \{A, B, C\}, \overline{\mathcal{S}} = \{A, B, C, E, F, G, H\}, \mathbf{L}_U = (A, F', E, B', G, C, H)$. It follows from $U \subseteq R(C, A')$ and $U \subseteq R(B', C)$ that $\mathsf{CSRP}[U, (C, H, A')] = 0$ and $\mathsf{CSRP}[U, (B', G, C)] = 0$. Since $\|EP\| \leq \|FP\|$ for all $P \in U$ and $U \subseteq R(A, E)$, $U \subseteq R(E, B')$, we have (7.3), where $\widehat{AB}$ intersects

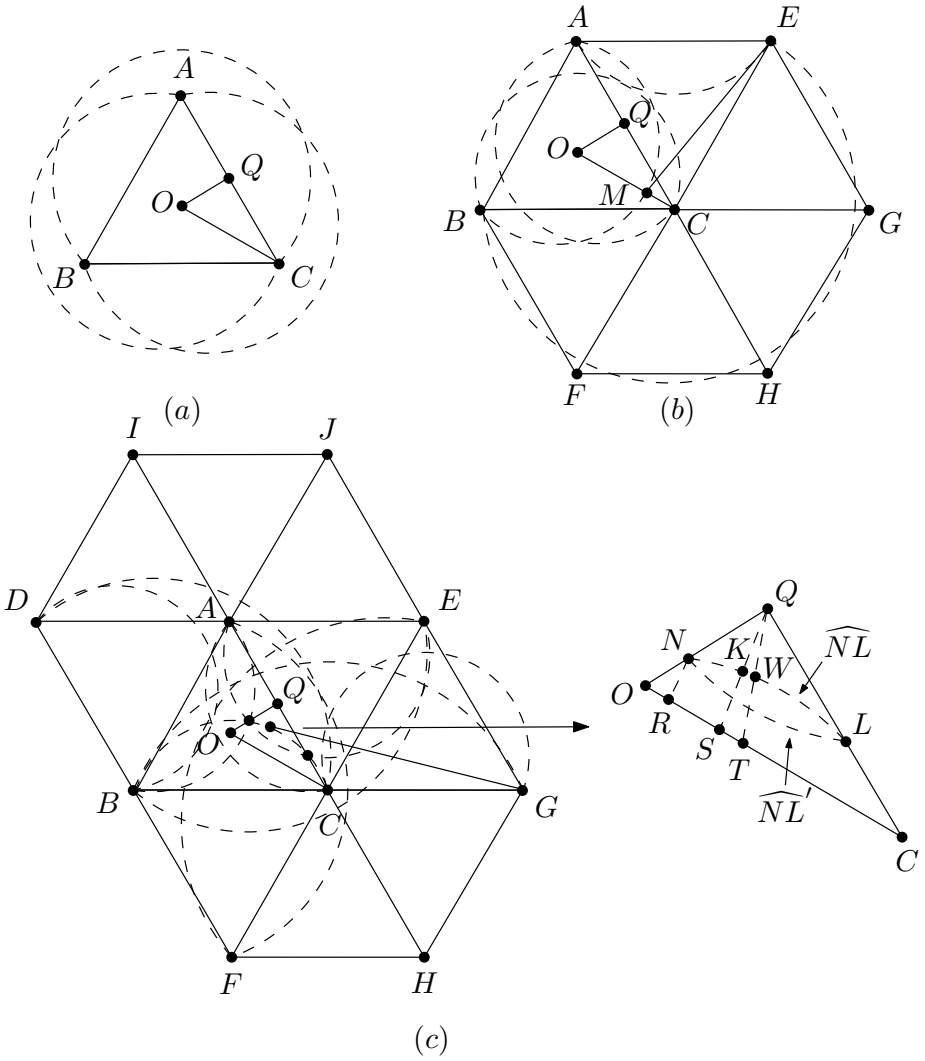


Figure 7.10: Illustration of applying Algorithm 7 for regular triangle deployment.

$\overline{OC}$ at M and $\|EM\| = d_E^{\max}(U \cap R^c(A, B'))$. Therefore, the output of CSR is $r^* = \max\{\underline{r}, 0, 0, \|EM\|\} = \|EM\|$. By geometric calculation, the closed form expression for r^* can be found as

$$r^* = l\sqrt{\left(\frac{\sqrt{3}}{2} - \frac{1}{\sin 2\theta} - \frac{1}{\tan 2\theta}\right)^2 + 1}.$$

Consider the case $\theta \in [\pi/6, \pi/4)$ as illustrated in Fig. 7.10(c). Using CSR, we obtain that $\underline{r} = l$, $\underline{\mathcal{S}} = \{A, B, C\}$, $\overline{\mathcal{S}} = \{A, B, C, D, E, F, G, H, I, J\}$, $\mathbf{L}_{U_1} = (A, F', J, E, B', G, D', C, H, I')$, $\mathbf{L}_{U_2} = (A, J, F', E, B', G, D', C, H, I')$, where U is par-

$$\begin{aligned}\mathsf{CSRP}[U,(A,F',E,B')] &= \max\{\|EM\|, \mathsf{CSRP}[U,(A,F',E)], \mathsf{CSRP}[U,(E,B')]\} \\ &= \max\{\|EM\|, 0, 0\} = \|EM\| \end{aligned} \tag{7.3}$$

$$\begin{aligned}\mathsf{CSRP}[U_1,(A,F',J,E,B')] &= \max\{\|ER\|, \mathsf{CSRP}[U_1^1,(A,F',J,E)], \mathsf{CSRP}[U_1^1,(E,B')]\} \\ &= \max\{\|ER\|, \max\{\|FQ\|, \mathsf{CSRP}[U_1^2,(A,F')], \\ &\qquad\qquad \mathsf{CSRP}[U_1^2,(F',J,E)]\}, 0\} \\ &= \max\{\|ER\|, \max\{\|FQ\|, 0, 0\}, 0\} = \|FQ\| \end{aligned} \tag{7.4}$$

$$\begin{aligned}\mathsf{CSRP}[U_1,(B',G,D',C)] &= \max\{\mathsf{CSRP}[U_1^3,(B',G,D',C)], \mathsf{CSRP}[U_1^4,(B',G,D',C)]\} \\ &= \max\{\max\{\|DK\|, \mathsf{CSRP}[U_1^5,(B',G,D')], \mathsf{CSRP}[U_1^5,(D',C)]\}, \\ &\qquad \max\{\|GK\|, \mathsf{CSRP}[U_1^6,(B',G)], \mathsf{CSRP}[U_1^6,(G,D',C)]\}\} \\ &= \max\{\max\{\|DK\|, 0, 0\}, \max\{\|GK\|, 0, 0\}\} = \|GK\| \end{aligned} \tag{7.5}$$

titioned by $\overline{JF}$ into U_1 and U_2. Regarding U_1, we have $\mathsf{CSRP}[U_1,(C,H,I',A')] = 0$, and also (7.4), (7.5), where all the input regions are described in Fig. 7.10(c). Therefore, the maximum value output by CSRP executed for U_1 is $\max\{\|FQ\|, \|GK\|, 0\} = \|GK\|$. On the other hand, following similar steps (omitted), the maximum value output by CSRP executed for U_2 is $\|GW\|$. Therefore, CSR outputs $r^* = \max\{\underline{r}, \|GK\|, \|GW\|\} = \|GK\|$. The closed form expression for r^* is lengthy and omitted here for brevity.

7.5 Numerical Results

In this section, we provide numerical results to evaluate the performance of Doppler coverage. We demonstrate the impact of different parameters on Doppler coverage under both random deployment and deterministic deployment.

We consider a region of interest F with an area of 100×100. We deploy sensors in an enlarged area of $(100 + 2r) \times (100 + 2r)$ to overcome the boundary effect. We use the percentage of Doppler coverage as the performance metric, which

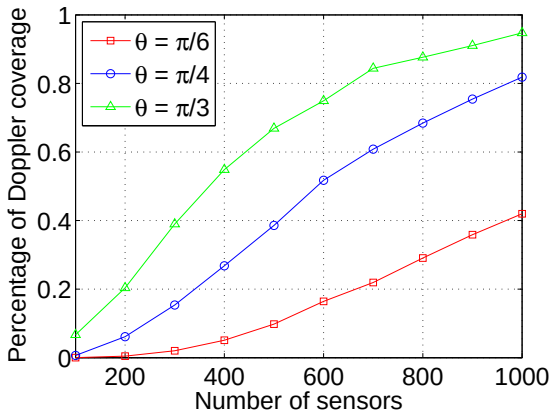


Figure 7.11: Impact of n under random deployment: $r = 5$.

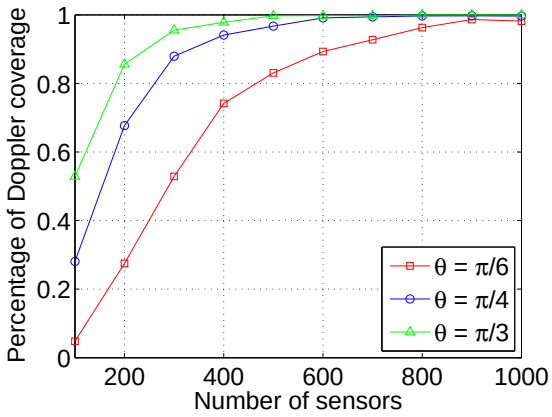


Figure 7.12: Impact of n under random deployment: $r = 10$.

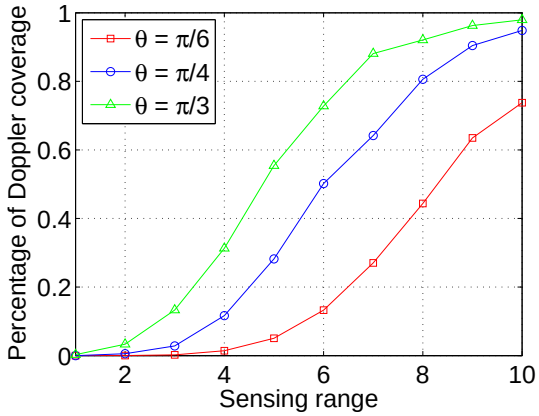


Figure 7.13: Impact of r under random deployment: $n = 400$.

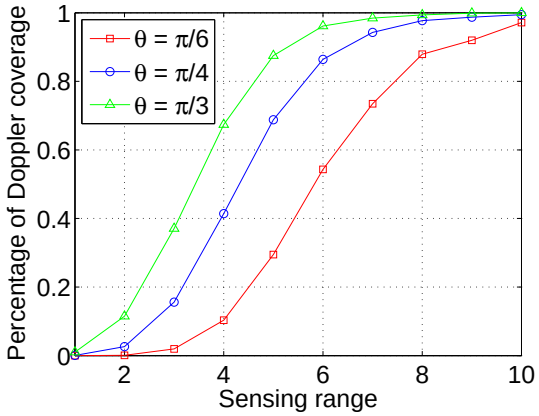


Figure 7.14: Impact of r under random deployment: $n = 800$.

is the percentage of Doppler-covered areas in region F. Following the methodology in [111], we use a grid of discrete points to approximate region F, and we set the side length of the grid to 0.1. As a result, given any deployed sensors, we calculate its percentage of Doppler coverage as the percentage of Doppler-covered grid points.

7.5.1 Random Deployment

In this subsection, we study random deployment under which sensors are deployed in region F according to uniform distribution. For each setting of parameters, we run the simulation for 100 times and take the averaged results.

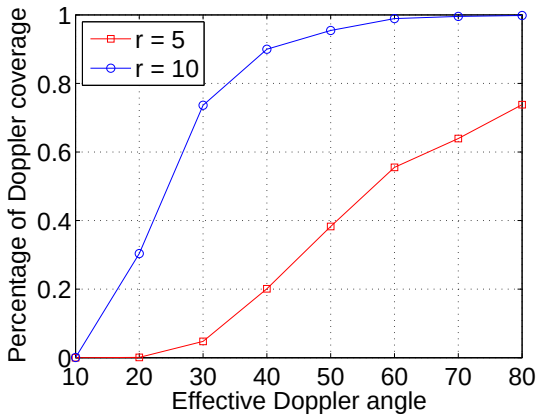


Figure 7.15: Impact of θ under random deployment: $n = 400$.

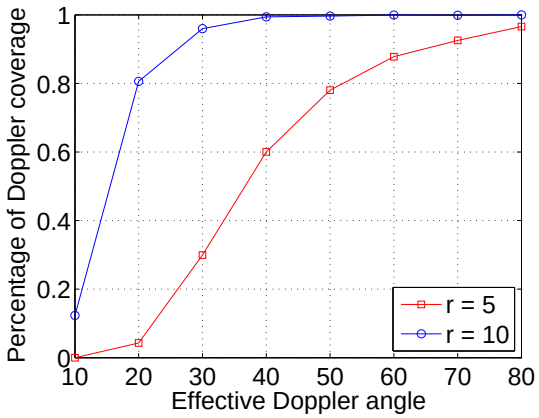


Figure 7.16: Impact of θ under random deployment: $n = 800$.

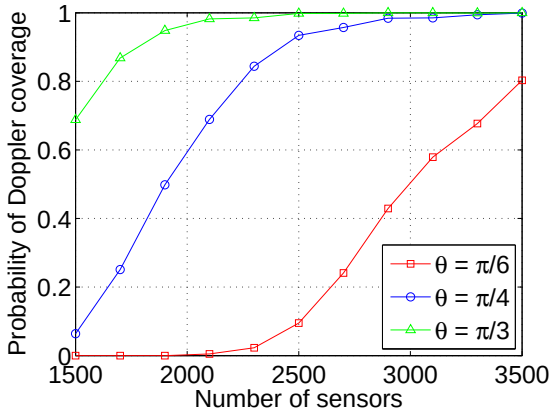


Figure 7.17: Impact of n under random deployment: $r = 5$.

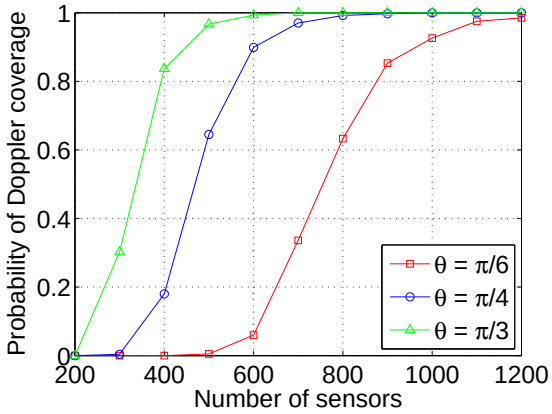


Figure 7.18: Impact of n under random deployment: $r = 10$.

Fig. 7.11 and 7.12 show the impact of sensor number when the sensing range is set to 5 and 10, respectively, for three settings of effective Doppler angle: $\theta = \pi/6(30°), \pi/4(45°), \pi/3(60°)$. We observe that the Doppler-covered percentage increases as the number of sensors grows, which is due to that on average more sensors contribute to the Doppler coverage of a point. Fig. 7.13 and 7.14 show the impact of sensing range when the sensor number is 400 and 800, respectively, for $\theta = \pi/6, \pi/4, \pi/3$. We can observe that the Doppler-covered percentage increases as the sensing range increases. This is because that a sensor with larger sensing range

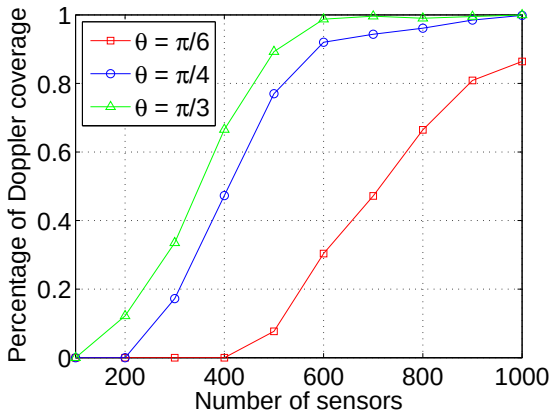


Figure 7.19: Impact of n under regular triangle deployment: $r = 5$.

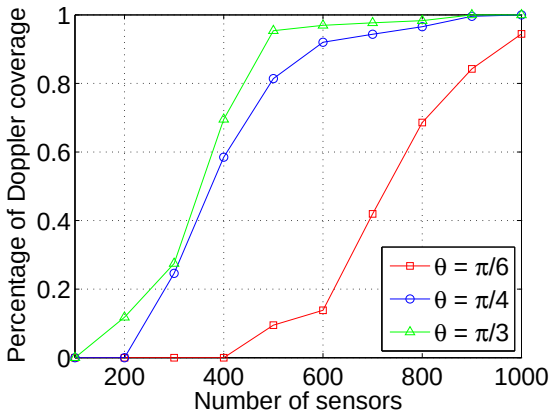


Figure 7.20: Impact of n under square deployment: $r = 5$.

can cover more points and hence contribute to the Doppler coverage of more points. Fig. 7.15 and 7.16 show the impact of effective Doppler angle when the sensor number is 400 and 800, respectively, for two settings of sensing range: $r = 5, 10$. We also observe that the Doppler-covered percentage increases as the effective Doppler angle increases. The reason is that a sensor with larger effective Doppler angle can Doppler-cover a wider range of directions from a point and hence contribute more to the Doppler coverage of that point.

Next we take the probability of Doppler coverage as the performance metric, which is the probability that all the grid points in region F are Doppler-covered. We calculate it as the ratio between the number of times that this event occurs and the total number of simulation runs, which is 100. Fig. 7.17 and 7.18 show the impact of sensor number on the probability of Doppler coverage when the sensing range is 5 and 10, respectively, for $\theta = \pi/6, \pi/4, \pi/3$. Similar to the Doppler-covered percentage, we observe that the probability of Doppler coverage also increases as the sensor number, sensing range, or effective Doppler angle increases. Compared to Fig. 7.11 and 7.12, we observe that many more sensors are needed to achieve a high probability of Doppler coverage than a high Doppler-covered percentage. For example, when $r = 10$ and

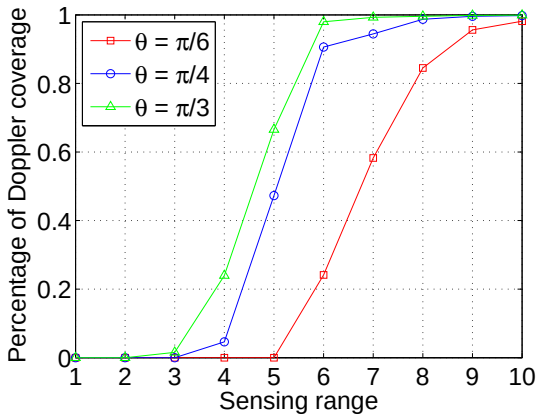


Figure 7.21: Impact of r under regular triangle deployment: $n = 400$.

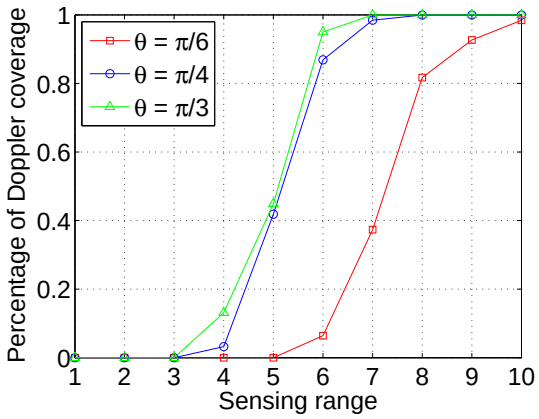


Figure 7.22: Impact of r under square deployment: $n = 400$.

$\theta = \pi/4$, the probability of Doppler coverage achieves 0.9 when the sensor number is 600, while the Doppler-covered percentage is about 0.9 when the sensor number is 300.

7.5.2 Deterministic Deployment

In this subsection, we study deterministic deployment of two deployment patterns: the regular triangle deployment and the square deployment.

In Fig. 7.19, 7.20, 7.21, 7.22, 7.23, 7.24, we show the impact of sensor number, sensing range, and effective Doppler angle, respectively, on the Doppler-covered percentage under regular triangle and square deployment, respectively. Similar to the case of random deployment, we observe that the Doppler-covered percentage increases as the sensor number, sensing range, or effective Doppler angle increases. We also observe that while the regular triangle deployment outperforms the square deployment under most parameter settings, there are cases where the square deployment performs better. For example, comparing Fig. 7.19 and 7.20, the Doppler-covered percentage under regular triangle deployment is lower than the square deployment when $n = 1000$ and $\theta = \pi/6$; comparing Fig. 7.21 and 7.22, the square deployment outperforms the regular triangle deployment when $r = 7$ and $\theta = \pi/4$. The above

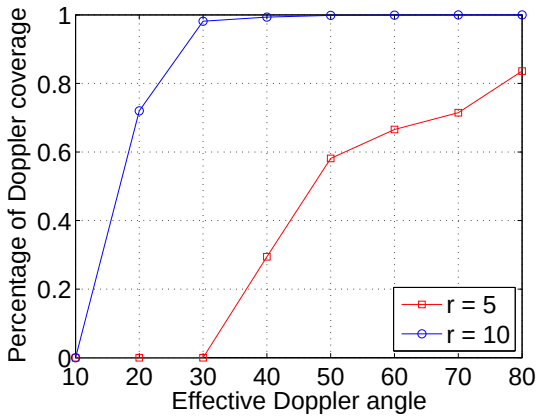


Figure 7.23: Impact of θ under regular triangle deployment: $n = 400$.

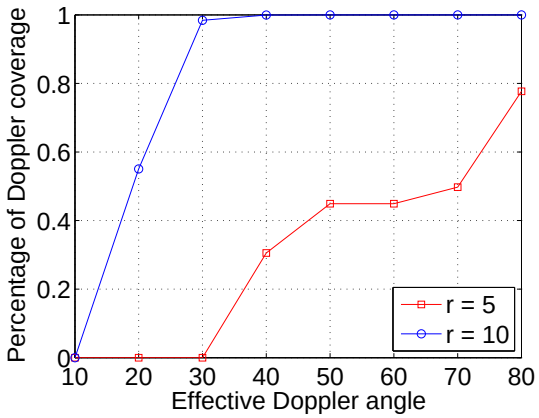


Figure 7.24: Impact of θ under square deployment: $n = 400$.

observation implies that the regular triangle deployment is in general not always a better deployment pattern than the square deployment.

We further observe that deterministic deployment performs better than random deployment when the parameters are favorable for Doppler coverage. For example, when $n = 600$, $r = 5$, and $\theta = \pi/3$, the percentage of Doppler coverage is less than 0.8 under random deployment as illustrated in Fig. 7.11, while it is very close to 1 under both the regular triangle and square deployments as illustrated in Fig. 7.19 and 7.20; when $n = 400$, $r = 7$, and $\theta = \pi/3$, the percentage of Doppler coverage is almost 1 under regular triangle and square deployment as illustrated in Fig. 7.21 and 7.22, while it is less than 0.9 under random deployment as illustrated in Fig. 7.13. This shows that deterministic deployment is more efficient than random deployment for achieving Doppler coverage.

7.6 Related Work

Radar technologies have continued to advance ever since its invention in the early 20th century [96]. However, radar networks have begun to draw research interests only in the past few years, largely driven by the emergence of cheap, efficient, and

compact radar sensors in place of conventionally expensive and bulky radar systems. For example, in [97], a platform has been successfully designed and built to integrate ultrawideband radars with mote-class sensor devices. The existing literature have studied different problems arising from radar sensor networks, including waveform design and diversity [88,98], radar scheduling [99], data management [100], for a variety of objectives, such as target detection [101], localization [92], and tracking [112,113]. In particular, Doppler effect has also been considered in several studies [112, 113]. However, very little attention has been paid to coverage issues specific to radar sensor networks, especially by exploring the Doppler effect. Our work fills this void by taking into account the Doppler effect for moving target detection.

There have been vast literature on sensor network coverage. A recent and comprehensive survey can be found in [9]. Many existing works focus on area coverage problems, where the subject to be covered is the entire area of a region. It has been proved in [114] that the regular triangle lattice is the optimal deterministic deployment pattern in terms of requiring the minimum number of sensors to cover a region. The optimal deployment for achieving both network coverage and connectivity has been studied in a line of works [115]. Our work in this chapter also falls into the category of area coverage. As the Doppler coverage model is very different from traditional distance-based models, the existing results cannot be applied to our problem.

A full-view coverage model has been recently introduced in [116] for camera sensor networks. This model shares a similar flavor with the Doppler coverage model proposed in our study: an object is full-view covered if its facing direction is sufficiently close to some camera's viewing direction. However, our work is quite different from [116] due to several reasons. First, our model is motivated by the Doppler effect, which is an utterly different physical phenomenon. Second, a direction is Doppler-

covered if it is sufficiently close to *either the direction towards or the direction opposite to* a radar, rather than only sufficiently close to a camera's viewing direction as in the full-view coverage model. Due to this complication, the Doppler coverage problem studied here is more complex and challenging. Third, while the critical sensor density for full-view coverage is derived in [116] specifically for the regular triangle deployment, the algorithm designed in our work can apply for any polygon deployment pattern. Therefore, our algorithm can be modified to apply for the full-view coverage model. Based on the full-view coverage model, barrier coverage for camera sensors has been studied in [110] and [108]. [117] has investigated the full-view coverage for heterogeneous camera sensors under random deployment.

7.7 Conclusion

The radar network has emerged as a promising paradigm for many applications. Based on a key observation that the existing sensor coverage models cannot capture the Doppler effect, which can otherwise be employed by a radar to distinguish a moving target from stationary or slow-moving clutter, we introduce a novel Doppler coverage model. A point P is said to be Doppler-covered if for any direction $\overrightarrow{D}$ from P, there exists some sensor S_i such that P is within S_i's sensing range, and the angle between $\overrightarrow{D}$ and $\overrightarrow{PS_i}$ is no greater than θ or no less than $\pi - \theta$, where θ is a parameter that depends on the maximum moving speed of an object that causes clutter. Based on this Doppler coverage model, we propose an efficient method to characterize Doppler-covered regions for arbitrary deployed radar sensors. We also design an algorithm CSR to compute the minimum sensor density required to Doppler-cover a region under deterministic deployment of any polygonal pattern, and further apply it to investigate the regular triangle deployment. Our results can be used to evaluate the coverage of any radar sensor network that exploits the Doppler effect for target detection, and also to estimate the number of sensors needed for coverage.

Chapter 8

SOCIAL GROUP UTILITY MAXIMIZATION WITH NEGATIVE SOCIAL TIES

8.1 Introduction

In Chapter 2, we develop a *social group utility maximization* (SGUM) game framework that takes into account mobile users' social ties and physical relationships. Under this framework, each user aims to maximize its social group utility, defined as the weighted sum of the individual utilities of the users that it has social ties with. Depending on the nature of social relationship, the social tie between two users can be "positive" (e.g., between family members or friends) such that one user cares about the welfare of the other, or it can be "negative" (e.g., against opponent or enemy) such that one user intends to damage the other's welfare. With this insight, we extend the SGUM framework to capture both positive and negative social ties. As a result, it *spans the rich continuum from zero-sum game (ZSG) to standard non-cooperative game (NCG) to network utility maximization (NUM)* - traditionally disjoint paradigms for network optimization (as illustrated in Fig. 8.1). These paradigms are encompassed under the SGUM framework as special cases.

Building on the extended SGUM framework, we study the SGUM based game for two applications: random access control and multi-channel cooperative jamming. For the SGUM based random access control game, we show that there exists a unique *social-aware Nash equilibrium* (SNE). As social ties increase, each user's access probability at the SNE migrates from its NE strategy for a ZSG to that for a standard NCG, and then to its social optimal strategy for NUM, while the social welfare of all users at the SNE improves gradually. We then turn our attention to the SGUM based multi-channel cooperative jamming game, which is always a ZSG, and show that there exists a unique mixed strategy SNE. When the social tie between the le-

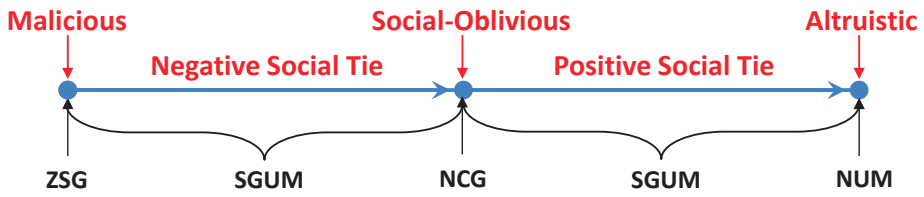


Figure 8.1: SGUM spans the continuum from ZSG to NCG to NUM.

gitimate user and the cooperative jammer exceeds certain threshold, the cooperative jammer always jams the eavesdropper on some channel at the SNE, and it improves the social welfare of the legitimate user and cooperative jammer.

The rest of this chapter is organized as follows. In Section 8.2, we introduce the social group utility maximization game framework. Section 8.3 studies the SGUM based random access control game. Section 8.5 studies the multi-channel cooperative jamming game. Section 8.6 concludes this chapter.

8.2 SGUM with Negative Social Ties

We consider a mobile network consisting of a set of users $\mathcal{N} \triangleq \{1, \cdots, N\}$. We model the social ties among the users by a social graph $(\mathcal{N}, \mathcal{E}^S)$, where user i is connected by a directed social edge $e_{ij}^S \in \mathcal{E}^S$ to user j if user i has a social tie with user j. The social tie from user i to user j is denoted by s_{ij}. We assume that the social tie from user i to j is 0 if no social ties exists from user i to j. For convenience, let s_{ii} denote user i's social tie to itself. We model the physical coupling among the users by a physical graph $(\mathcal{N}, \mathcal{E}^P)$, where user j is connected by a directed physical edge $e_{ji}^P \in \mathcal{E}^P$ to user i if user j's behavior affects user i's welfare. Let u_i denote the *individual utility* of user i. To take into account the social ties among users, each user i aims to maximize its *social group utility* f_i, which is its payoff function in the SGUM game, defined as

$$f_i(a_i, \boldsymbol{a}_{-i}) \triangleq \sum_{j=1}^{N} s_{ij} u_i(a_i, \boldsymbol{a}_{-i})$$

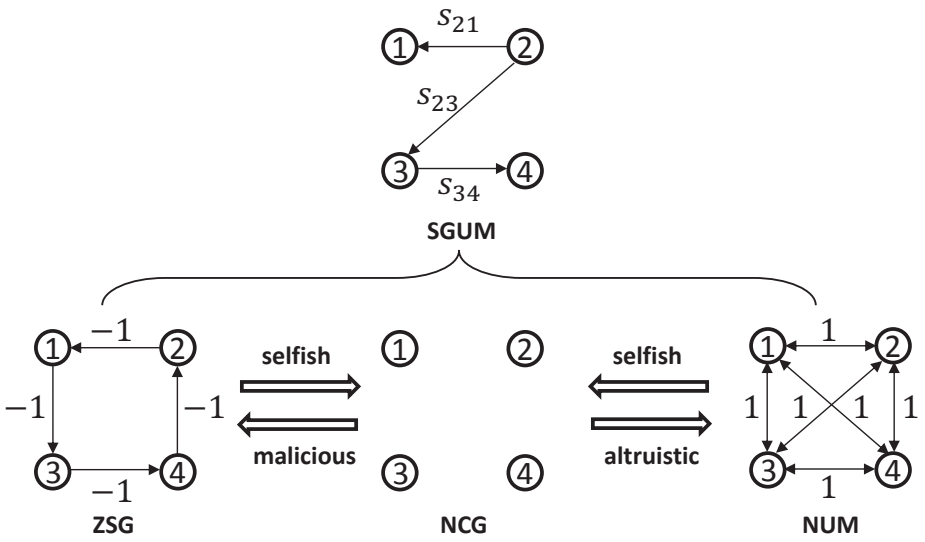


Figure 8.2: Extended SGUM captures ZSG, NCG, and NUM as special cases.

where a_i denotes user i's strategy and $\boldsymbol{a}_{-i}$ denotes the vector consisting of all users' strategies except i.

We assume that each user's social tie to itself is normalized as 1 (i.e., $s_{ii} = 1$, $\forall i \in \mathcal{N}$). We further assume that $s_{ij} \in (-\infty, 1]$: When $s_{ij} \in (0, 1]$, it represents the extent to which user i cares about user j's utility, and it reaches the highest when $s_{ij} = 1$ (i.e., user i cares about user j's utility as much as its own utility); when $s_{ij} \in (-\infty, 0)$, it pinpoints to how much user i intends to damage user j's utility, and reaches the extreme as s_{ij} goes to $-\infty$ (i.e., user i would sacrifice its utility to damage user j's utility). A strategy profile is a *social-aware Nash equilibrium* (SNE) if no user can improve its social group utility by unilaterally changing its strategy.

The social group utility maximization (SGUM) game framework provides rich modeling flexibility, and encompasses traditional network optimization paradigms as special cases. In particular, if no social tie exists (i.e., $s_{ij} = 0$, $\forall i \neq j$), the SGUM game degenerates to a standard non-cooperative game where each user is selfish and

only cares about its individual utility, i.e.,

$$f_i(a_i, \boldsymbol{a}_{-i}) = u_i(a_i, \boldsymbol{a}_{-i}),\ \forall i \in \mathcal{N}.$$

In this case, a strategy profile is a *Nash equilibrium* (NE) if no user can improve its individual utility by unilaterally changing its strategy. If all users are connected by social ties of 1 (i.e., $s_{ij} = 1, \forall i \neq j$), the SGUM game degenerates to network utility maximization where each user is altruistic and cares about the individual utilities of all other users as much as its own individual utility, i.e.,

$$f_i(a_i, \boldsymbol{a}_{-i}) = \sum_{j=1}^{N} u_j(a_j, \boldsymbol{a}_{-j}),\ \forall i \in \mathcal{N}.$$

In this case, a strategy profile is *social optimal* (SO) if it maximizes the social welfare among all strategy profiles, where the social welfare v is defined as the total individual utility of all users:

$$v(\boldsymbol{a}) \triangleq \sum_{i=1}^{N} u_i(\boldsymbol{a})$$

with $\boldsymbol{a} \triangleq (a_1, \cdots, a_N)$. In a zero-sum game, all users' payoff functions are summed up to 0. Therefore, in an SGUM game or its equivalent game, if each user's positive social ties "cancel out" its negative social ties, i.e.,

$$\sum_{j \in \mathcal{N}} s_{ji} = 0,\ \forall i \in \mathcal{N},$$

the SGUM game degenerates to a zero-sum game, where each user views the total gain of other users as its loss. For example, an SGUM game of two users with $f_1 = u_1 - u_2$ and $f_2 = u_2 - u_1$, or $f_1 = u_1$ and $f_2 = -u_1$, is a zero-sum game. Note that we obtain an equivalent game if a user's payoff function is multiplied by a number or added by a function independent of that user's strategy. For example, an SGUM game of two users with $f_1 = u_1$ and $f_2 = u_2 - s_{21}u_1$ where $s_{21} \to \infty$ is equivalent to that with $f_1 = u_1$ and $f_2 = -u_1$, and therefore is a zero-sum game.

To get a more concrete sense of the SGUM game framework, in Section 8.3 and 8.5, we will study the SGUM based game for two applications: random access control and multi-channel cooperative jamming.

8.3 SGUM based Random Access Control

We consider a set of mobile users under the protocol interference model, where user i is a link consisting of transmitter T_i and receiver R_i. Let $\mathcal{I}_i^+$ denote the set of receivers that transmitter T_i cause interference to, and $\mathcal{I}_i^-$ denote the set of transmitters that causes interference to receiver R_i. Each user i contends for the opportunity of data transmission with probability q_i in a time slot where $0 \leq q_i \leq 1$. If multiple interfering links contend in the same time slot, a collision occurs and no link can grab the transmission opportunity. Then the probability b_i that user i can grab the transmission opportunity is given by

$$b_i(q_i, \boldsymbol{q}_{-i}) = q_i \prod_{j \in \mathcal{I}_i^-} (1 - q_j).$$

We assume that the individual utility of user i is given by

$$u_i(q_i, \boldsymbol{q}_{-i}) = \log(\theta_i b_i)$$

where θ_i denotes user i's efficiency of utilizing the transmission opportunity (e.g., transmission rate). Note that the logarithmic function is widely used for modeling the utility of wireless users [18, 19].

8.3.1 Game Analysis

For the SGUM based random access control game, we have the following result.

Theorem 8.1. *For the SGUM based random access control game, there exists a unique SNE, where*

$$q_i = \frac{1}{1 + \max\left(\sum_{j \in \mathcal{I}_i^+} s_{ij}, 0\right)}, \forall i \in \mathcal{N}. \tag{8.1}$$

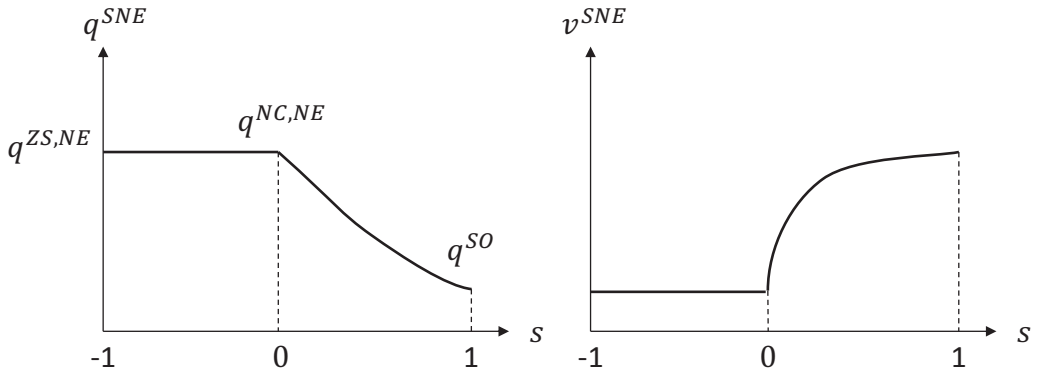


Figure 8.3: SNE for two-user SGUM based random access control.

Proof: Since

$$f_i = \log\left(\theta_i q_i \prod_{j\in\mathcal{I}_i^-}(1-q_j)\right) + \sum_{j\neq i} s_{ij}\log\left(\theta_j q_j \prod_{k\in\mathcal{I}_j^-}(1-q_k)\right),$$

it follows that

$$\frac{\partial f_i}{\partial q_i} = \frac{1}{q_i} - \sum_{j\in\mathcal{I}_i^+}\frac{s_{ij}}{1-q_i}.$$

When $\sum_{j\in\mathcal{I}_i^+} s_{ij} \leq 0$, we have $\frac{\partial f_i}{\partial q_i} > 0$ for $q_i \in [0,1]$, and therefore the SNE strategy is $q_i = 1$. When $\sum_{j\in\mathcal{I}_i^+} s_{ij} > 0$, we have

$$\lim_{q_i\to 0}\frac{1}{q_i} - \sum_{j\in\mathcal{I}_i^+}\frac{s_{ij}}{1-q_i} = \infty,\ \lim_{q_i\to 1}\frac{1}{q_i} - \sum_{j\in\mathcal{I}_i^+}\frac{s_{ij}}{1-q_i} = -\infty$$

and

$$\frac{\partial\left(\frac{1}{q_i} - \sum_{j\in\mathcal{I}_i^+}\frac{s_{ij}}{1-q_i}\right)}{\partial q_i} = -\frac{1}{q_i^2} - \frac{\sum_{j\in\mathcal{I}_i^+} s_{ij}}{(1-q_i)^2} < 0.$$

Hence, there exists a unique value of q_i such that

$$\frac{1}{q_i} - \sum_{j\in\mathcal{I}_i^+}\frac{s_{ij}}{1-q_i} = 0, \tag{8.2}$$

which is the value of the SNE strategy. Solving (8.2) yields $q_i = \frac{1}{1+\sum_{j\in\mathcal{I}_i^+} s_{ij}}$. Combining the above two cases we have (8.1). □

The result below directly follows from Theorem 8.1.

Corollary 8.1. *Each user's access probability at the SNE is non-increasing as social ties get stronger.*

Remark: Each user's SNE strategy is independent of other users' strategies (also known as a *dominant strategy*), but depends on the user's social ties with others. When a user's access probability increases, it raises the collision probability of the users within its interference range, and hence reduces their individual utilities. Therefore, a user would reduce its access probability when its social ties with those within its interference range get stronger (as illustrated in Fig. 8.3).

Proposition 8.1. *The social welfare of the SNE is non-decreasing as social ties get stronger, and reaches the social optimal point when all social ties are equal to 1.*

Proof: Since

$$v = \sum_{i=1}^{N} \log \left(\theta_i q_i \prod_{j \in \mathcal{I}_i^-} (1 - q_j) \right),$$

we have

$$\frac{\partial v}{\partial q_i} = \frac{1}{q_i} - \sum_{j \in \mathcal{I}_i^+} \frac{1}{1 - q_i} = \frac{1}{q_i} - \frac{|\mathcal{I}_i^+|}{1 - q_i}.$$

Using the same argument as in the proof of Theorem 8.1, the social optimal value q_i^{SO} of q_i that maximizes v is the unique solution of

$$\frac{1}{q_i} - \frac{|\mathcal{I}_i^+|}{1 - q_i} = 0,$$

which is

$$q_i^{SO} = \frac{1}{1 + |\mathcal{I}_i^+|}.$$

In particular, we have $q_i^{SNE} \geq q_i^{SO}$. Since $\frac{\partial v}{\partial q_i} < 0$ when $q_i \geq q_i^{SO}$, v is decreasing in q_i when $q_i \geq q_i^{SO}$. Using Lemma 8.1, q_i^{SNE} is non-increasing in s_{ij}, $\forall j \in \mathcal{I}_i^+, \forall i \in \mathcal{N}$, and hence $v(q_1^{SNE}, \cdots, q_N^{SNE})$ is non-decreasing in s_{ij}, $\forall j \in \mathcal{I}_i^+, \forall i \in \mathcal{N}$. □

Remark: Since users' individual utilities are added up with equal weights in the social welfare, a user's SNE strategy gets closer to its social optimal strategy when other users weigh more (i.e., the social ties get stronger) in that user's social group utility, and the social welfare is non-decreasing. As social ties get stronger, a user's SNE strategy migrates from its NE strategy for a ZSG to its NE strategy for a standard NCG, and then to its social optimal strategy for NUM. For example, in Fig. 8.3, as the social tie $s_{12} = s_{21} \triangleq s$ increases from -1 to 0 and then to 1, each user's SNE strategy q^{SNE} migrates from its NE strategy for a zero-sum game $q^{ZS,NE}$ to its NE strategy for a standard non-cooperative game $q^{NC,NE}$, and then to its social optimal strategy q^{SO}. The social welfare of the SNE v^{SNE} also migrates correspondingly. This demonstrates that the SGUM game framework spans the continuum between these traditionally disjoint paradigms.

8.4 SGUM based Multi-Channel Cooperative Jamming

8.4.1 System Model

For ease of exposition, we consider a wireless network consisting of three users: a legitimate user, a cooperative jammer, and an eavesdropper. The legitimate user (referred to as "user") can transmit data on one of a set of channels $\mathcal{M} \triangleq \{1, \cdots, M\}$. The eavesdropper can eavesdrop the data transmission on one of the M channels. As illustrated in Fig. 8.4, to assist the legitimate user's transmission from its transmitter (T) to its receiver (R), the cooperative jammer (J) can jam the eavesdropper (E) by transmitting jamming signal on one of the M channels. The desired user's utility of transmitting data on a channel depends on whether the eavesdropper is eavesdropping on that channel and whether the helper is jamming that channel. For example, the user's utility can be defined as the secrecy capacity [118]. We assume that the channels are homogeneous, and let u, u^e, u^j, $u^{j,e}$ denote the user's utility of transmitting data on a channel when the same channel is being accessed by neither the

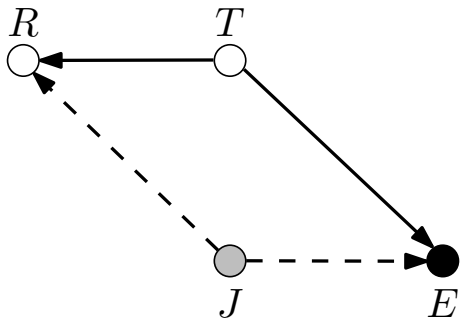


Figure 8.4: Illustration of cooperative jamming model.

jammer nor eavesdropper, only the eavesdropper, only the jammer, both the jammer and eavesdropper, respectively. We assume that $u > u^j > u^{j,e} > u^e$ to avoid the trivial case. We use p_i, p_i^j, p_i^e to denote the probability that the user, the jammer, the eavesdropper access channel i, respectively. We assume that $\sum_{i\in\mathcal{M}} p_i = 1$ and $\sum_{i\in\mathcal{M}} p_i^e = 1$ meaning that the user and eavesdropper always access some channel(s). We also assume that $\sum_{i\in\mathcal{M}} p_i^j \leq 1$ meaning that the jammer can choose to keep silent.

Under the SGUM framework, the user's expected payoff $\overline{f}$ is equal to its expected individual utility $\overline{u}$, and is given by

$$\overline{f} = \overline{u} = \sum_{i\in\mathcal{M}} p_i \left[u(1-p_i^j)(1-p_i^e) + u^j p_i^j (1-p_i^e) + u^{j,e} p_i^j p_i^e + u^e (1-p_i^j) p_i^e\right].$$

The jammer's expected payoff is

$$\overline{f^j} = \overline{u^j} + s\overline{u},$$

where $s \in [0,1]$ is the jammer's social tie with the user, and $\overline{u^j}$ is the jammer's expected individual utility, which is its expected jamming cost given by

$$\overline{u^j} = -c \sum_{i\in\mathcal{M}} p_i^j$$

where $c > 0$ is the jamming cost. The eavesdropper's expected payoff is given by

$$\overline{f^e} = -\overline{u}.$$

Since the eavesdropper's payoff function is equivalent to $\overline{f^e} = -(1+s)\overline{u} - \overline{u^j}$, the SGUM game is always a zero-sum game.

8.4.2 Game Analysis

For the SGUM based multi-channel cooperative jamming game, in general there does not exist a pure strategy SNE. Therefore, we aim to find a mixed strategy SNE. A mixed strategy profile is a SNE if no one can improve its expected payoff by unilaterally changing its mixed strategy. For convenience, define

$$M_0 \triangleq \frac{u^{j,e} - u^e}{u - u^j} + 1$$

and

$$s_0 \triangleq \frac{cM^2}{(u^{j,e} - u^e) - (M-1)(u - u^j)}.$$

Theorem 8.2. *For the SGUM based multi-channel cooperative jamming game, there exists a unique mixed strategy SNE. If $M < M_0$ and $s > s_0$, the SNE is given by*

$$p_i = p_i^j = p_i^e = \frac{1}{M}, \quad \forall i \in \mathcal{M};$$

otherwise, the SNE is

$$p_i = p_i^e = \frac{1}{M}, \quad p_i^j = 0, \quad \forall i \in \mathcal{M}.$$

Proof: We first show that $p_i > 0$ and $p_i^e > 0$ for each $i \in \mathcal{M}$. To this end, we show the following claims: 1) if $p_i = 0$ for some $i \in \mathcal{M}$, then $p_i^e = 0$; 2) $p_i > 0$ for each $i \in \mathcal{M}$; 3) $p_i^e > 0$ for any $i \in \mathcal{M}$ with $p_i > 0$. Claim 1 holds because if $p_i^e > 0$, we can decrease p_i^e by a sufficiently small Δ while increasing p_k^e by Δ for some $k \in \mathcal{M}$ with $p_k > 0$ such that $\overline{f^e}$ is increased. Claim 2 holds because if $p_i = 0$ for some $i \in \mathcal{M}$, we have $p_i^e = 0$, and also $p_i^j = 0$, since otherwise we can decrease p_i^j such that $\overline{f^j}$ is increased. Then we can decrease p_k by a sufficiently small Δ for some $k \in \mathcal{M}$ with $p_k > 0$ and $p_k^e > 0$ while increasing p_i by Δ such that $\overline{f}$ is increased. Claim 3 holds

because if $p_i^e = 0$, we must have $p_i^j = 0$, since otherwise we can decrease p_i^j such that $\overline{f^j}$ is increased. Then we can increase p_i by a sufficiently small Δ while decreasing p_k by Δ for some $k \in \mathcal{M}$ with $p_k^e > 0$ such that $\overline{f}$ is increased. Using Claim 1, 2, and 3, we must have $p_i > 0$ and $p_i^e > 0$ for each $i \in \mathcal{M}$.

Next we show that $p_i = p_k$, $\forall i \neq k$. To this end, define

$$\overline{f_i} \triangleq p_i \left[u(1-p_i^j)(1-p_i^e) + u^j p_i^j (1-p_i^e) + u^{j,e} p_i^j p_i^e + u^e (1-p_i^j) p_i^e \right]. \tag{8.3}$$

Then we have

$$\frac{\partial \overline{f_i}}{\partial p_i} = u(1-p_i^j)(1-p_i^e) + u^j p_i^j (1-p_i^e) + u^{j,e} p_i^j p_i^e + u^e (1-p_i^j) p_i^e. \tag{8.4}$$

Since we have shown that $p_i > 0$ for each $i \in \mathcal{M}$, we must have $\frac{\partial \overline{f_i}}{\partial p_i} = \frac{\partial \overline{f_k}}{\partial p_k}$, $\forall i \neq k$. This is because if $\frac{\partial \overline{f_i}}{\partial p_i} > \frac{\partial \overline{f_k}}{\partial p_k}$ for some $i \neq k$, we can increase p_i by a sufficiently small Δ while decreasing p_k by Δ such that $\overline{f_i} + \overline{f_k}$ is increased, and hence $\overline{f}$ is increased. It follows from (8.3) that

$$\frac{\partial \overline{f_i}}{\partial p_i^e} = p_i \left[(u - u^j + u^{j,e} - u^e) p_i^j - (u - u^e) \right]. \tag{8.5}$$

Since $u > u^j > u^{j,e} > u^e$, we have

$$0 < u - u^j + u^{j,e} - u^e < u - u^e,$$

and then it follows from (8.5) that $\frac{\partial \overline{f_i}}{\partial p_i^e} < 0$, $\forall i \in \mathcal{M}$. Since we have shown that $p_i^e > 0$ for each $i \in \mathcal{M}$, we must have $\frac{\partial \overline{f_i}}{\partial p_i^e} = \frac{\partial \overline{f_k}}{\partial p_k^e}$, $\forall i \neq k$. This is because if $\frac{\partial \overline{f_i}}{\partial p_i^e} < \frac{\partial \overline{f_k}}{\partial p_k^e}$, we can increase p_i^e by a sufficiently small Δ while decreasing p_k^e by Δ such that $\overline{f_i} + \overline{f_k}$ is decreased, and hence $\overline{f^e}$ is increased.

Suppose $p_i > p_k$ for some $i \neq k$. Since $\frac{\partial \overline{f_i}}{\partial p_i^e} = \frac{\partial \overline{f_k}}{\partial p_k^e} < 0$, it follows from (8.5) that $p_i^j > p_k^j$. It follows from (8.3) that

$$\frac{\partial \overline{f_i}}{\partial p_i^j} = p_i \left[(u - u^j + u^{j,e} - u^e) p_i^e - (u - u^j) \right]. \tag{8.6}$$

We consider four cases as follows.

Case 1: $\frac{\partial \overline{f_i}}{\partial p_i^j} \leq 0$ and $\frac{\partial \overline{f_k}}{\partial p_k^j} \leq 0$

In this case, we must have $p_i^j = p_k^j = 0$, since otherwise we can increase p_i^j or p_k^j such that $\overline{f^j}$ is increased. This contradicts to $p_i^j > p_k^j$.

Case 2: $\frac{\partial \overline{f_i}}{\partial p_i^j} \leq 0$ and $\frac{\partial \overline{f_k}}{\partial p_k^j} > 0$

Similar to Case 1, we must have $p_i^j = 0$. Then it follows that $0 = p_i^j > p_k^j \geq 0$, which is a contradiction.

Case 3: $\frac{\partial \overline{f_i}}{\partial p_i^j} > 0$ and $\frac{\partial \overline{f_k}}{\partial p_k^j} \leq 0$

In this case, using (8.6) we have

$$(u - u^j + u^{j,e} - u^e)p_i^e - (u - u^j) \geq 0$$

and

$$(u - u^j + u^{j,e} - u^e)p_k^e - (u - u^j) < 0,$$

and therefore we must have $p_i^e > p_k^e$. It follows from (8.4) that

$$\frac{\partial \overline{f_i}}{\partial p_i \partial p_i^e} = (u - u^j + u^{j,e} - u^e)p_i^j - (u - u^e) < 0. \tag{8.7}$$

Since $\frac{\partial \overline{f_k}}{\partial p_k^j} \leq 0$, we have

$$\frac{\partial \overline{f_k}}{\partial p_k \partial p_k^j} = (u - u^j + u^{j,e} - u^e)p_k^e - (u - u^j) = \frac{\partial \overline{f_k}}{p_k \partial p_k^j} \leq 0. \tag{8.8}$$

Using (8.4), we have

$$\frac{\partial \overline{f_i}}{\partial p_i} < u(1 - p_i^j)(1 - p_k^e) + u^j p_i^j (1 - p_k^e) + u^{j,e} p_i^j p_k^e + u^e (1 - p_i^j) p_k^e \leq \frac{\partial \overline{f_k}}{\partial p_k},$$

where the first inequality is due to $p_i^e > p_k^e$ and (8.7), the second inequality is due to $p_i^j > p_k^j$ and (8.8). This contradicts to $\frac{\partial \overline{f_i}}{\partial p_i} = \frac{\partial \overline{f_k}}{\partial p_k}$.

Case 4: $\frac{\partial \overline{f_i}}{\partial p_i^j} > 0$ and $\frac{\partial \overline{f_k}}{\partial p_k^j} > 0$

Since $\frac{\partial \overline{f_i}}{\partial p_i^j} = \frac{\partial \overline{f_k}}{\partial p_k^j} > 0$ and $p_i > p_k$, using (8.6) we have $p_i^e < p_k^e$. Since $\frac{\partial \overline{f_i}}{\partial p_i^j} > 0$, we have

$$\frac{\partial \overline{f_i}}{\partial p_i \partial p_i^j} = (u - u^j + u^{j,e} - u^e)p_i^e - (u - u^j) = \frac{\partial \overline{f_i}}{p_i \partial p_i^j} > 0. \tag{8.9}$$

Using (8.4), we have

$$\frac{\partial \overline{f_i}}{\partial p_i} > u(1-p_i^j)(1-p_k^e) + u^j p_i^j (1-p_k^e) + u^{j,e} p_i^j p_k^e + u^e (1-p_i^j) p_k^e > \frac{\partial \overline{f_k}}{\partial p_k},$$

where the first inequality is due to $p_i^e < p_k^e$ and (8.7), the second inequality is due to $p_i^j > p_k^j$ and (8.9). This contradicts to $\frac{\partial \overline{f_i}}{\partial p_i} = \frac{\partial \overline{f_k}}{\partial p_k}$. Therefore, we conclude that $p_i = p_k$, $\forall i \neq k$.

Since $p_i = p_k$ and $\frac{\partial \overline{f_i}}{\partial p_i^e} = \frac{\partial \overline{f_k}}{\partial p_k^e}$, $\forall i \neq k$, using (8.6) we have $p_i^j = p_k^j$, $\forall i \neq k$. Then since $\frac{\partial \overline{f_i}}{\partial p_i} = \frac{\partial \overline{f_k}}{\partial p_k}$, $\forall i \neq k$, using (8.4) we have $p_i^e = p_k^e$, $\forall i \neq k$. Therefore, we have $p_i = p_i^e = \frac{1}{M}$, $\forall i \in \mathcal{M}$.

Since $p_i^e = \frac{1}{M}$, $\forall i \in \mathcal{M}$, using (8.6) we have that $p_i^j > 0$ only if $(u - u^j + u^{j,e} - u^e)\frac{1}{M} - (u - u^j) > 0$, or equivalently $M < M_0$. Using $p_i = p_i^e = \frac{1}{M}$, $\forall i \in \mathcal{M}$, and $p_i^j = p_k^j$, $\forall i \neq k$, we have

$$\overline{f^j} = s\left[u(1-p_i^j)(1-\frac{1}{M}) + u^j p_i^j (1-\frac{1}{M}) + u^{j,e} p_i^j \frac{1}{M} + u^e (1-p_i^j)\frac{1}{M}\right] - cMp_i^j.$$

Then we have

$$\frac{\partial \overline{f^j}}{\partial p_i^j} = s\left[\frac{1}{M}(u^{j,e} - u^e) - (1-\frac{1}{M})(u - u^j)\right] - cM.$$

Since $\sum_{i \in \mathcal{M}} p_i^j \leq 1$ and $p_i^j = p_k^j$, $\forall i \neq k$, we have $p_i^j = \frac{1}{M}$ if and only if $\frac{\partial \overline{f^j}}{\partial p_i^j} > 0$, or equivalently $s > s_0$ and $M < M_0$. □

Remark: Intuitively, since the channels are homogeneous, if the user accesses a channel with a higher probability than other channels, then the eavesdropper should always access that channel so as to minimize the user's payoff; if the eavesdropper

accesses a channel with a lower probability than other channels, the user should always access that channel so as to maximize its payoff. Therefore, both the user and eavesdropper access all channels with equal probabilities at the SNE.

Remark: The jammer can affect the user's utility in two inverse ways: when the user accesses the same channel as the eavesdropper, the jammer can improve the user's utility by jamming that channel; when the user and eavesdropper access different channels, the jammer reduces the the user's utility if it jams the user's channel. The tradeoff between these inverse effects is captured by M_0: when the user's utilities under the jammer's interference (i.e., u^j and $u^{j,e}$) are sufficiently large, the jammer's "positive" effect outweighs its "negative" effect (i.e., $M < M_0$), and it should always jam some channel; otherwise, it should always keep silent.

Remark: The impact of social tie is captured by s_0: when the social tie s exceeds the threshold s_0, the jammer has incentive to assist the user, and the social welfare of the user and jammer (i.e., $\overline{u} + \overline{u^j}$) improves. Therefore, the positive social tie between the user and jammer mitigates the damage from the eavesdropper who has negative social ties with them.

Using Theorem 8.2, when $M < M_0$ and $s > s_0$, the user's payoff at the SNE is

$$\overline{f} = u(1 - \frac{1}{M})^2 + u^j(1 - \frac{1}{M})\frac{1}{M} + u^{j,e}\frac{1}{M^2} + u^e(1 - \frac{1}{M})\frac{1}{M};$$

when $M \geq M_0$ or $s \leq s_0$, the user's payoff at the SNE is

$$\overline{f'} = u(1 - \frac{1}{M}) + u^e\frac{1}{M}.$$

Therefore we have

$$\overline{f} - \overline{f'} = \frac{1}{M}\left[\frac{1}{M}(u^{j,e} - u^e) - (1 - \frac{1}{M})(u - u^j)\right] > 0.$$

This shows that the jammer's help improves the user's utility at the SNE. We also observe that

$$\frac{\partial \overline{f}}{\partial M} = \frac{1}{M^2}(1 - \frac{1}{M})(2u - u^j - u^e) + \frac{1}{M^3}(u^j - u^{j,e} + u^e) > 0$$

and

$$\frac{\partial \overline{f'}}{\partial M} = \frac{1}{M^2}(u - u^e) > 0.$$

This shows that the user's utility at the SNE improves when the number of channels increases, regardless of whether the jammer helps the user or not.

Remark: Intuitively, with a larger number of channels, the user has a better chance to avoid accessing the same channel as the eavesdropper, and therefore achieves a higher payoff. Furthermore, as the user's chance of not being eavesdropped increases, the jammer's "negative" effect grows while its "positive" effect diminishes. As a result, when the number of channels M is sufficiently large, the jammer should always keep silent (i.e., $M < M_0$).

8.5 SGUM based Spectrum Access for Balanced Social Networks

8.5.1 Social Network with Structural Balance

We consider a social network of a set of users where each pair of users know each other and they are either friends or enemies - no two users are indifferent to one another or unaware of each other. Such a network can be represented by a complete social graph where each edge is labeled with a positive + or negative − sign. The principles underlying structural balance are based on theories in social psychology [30]. The key idea is as follows: For any set of three users, only three social structures among them are possible, as illustrated in Fig. 8.5. If a social network is weakly balanced, it has been shown that a global structural property holds as follows: The users can be divided into groups in such a way that every two users belonging to the same group are friends, and every two users belonging to different groups are enemies. Therefore,

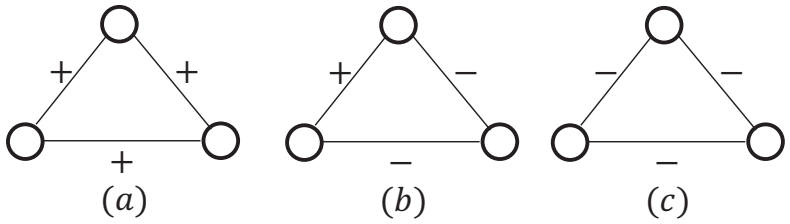


Figure 8.5: Illustration of structural balance: A social network is weakly balanced if only the social structures in (a), (b), and (c) exist for any set of three users; it is strongly balanced if only (a) and (b) exist.

the social network can be partitioned into a number of social coalitions, where all the users in the same coalition are mutual friends while any two users from different coalitions are enemies.

8.5.2 System Model

We consider a set of users $\mathcal{N} = \{1, 2, ..., N\}$ where N is the total number of users. We denote the set of available channels as $\mathcal{M} = \{1, 2, ..., M\}$. Each user i chooses a channel $a_i \in \mathcal{M}$ for data transmission. To capture the physical coupling, we construct an interference graph $G^P \triangleq \{\mathcal{N}, \mathcal{E}^P\}$ based on the users' interference relationships. In particular, the set of users $\mathcal{N}$ is the vertex set, and $\mathcal{E}^P \triangleq \{(i, j) : e_{ij}^P = 1, \forall i, j \in \mathcal{N}\}$ is the edge set where $e_{ij}^P = 1$ if and only if users i and j can generate significant interference that affect the data transmissions of each other. For example, we can construct the interference graph G^P based on users' spatial relationships [119]. Let δ denote the transmission range of each user. We then have $e_{ij}^P = 1$ if and only if the distance d_{ij} between user i and j is not greater than the threshold δ, i.e., $d_{ij} \leq \delta$. Each edge e_{ij}^P is associated with a weight W_{ij} which represents the interference generated by user i to user $j \neq i$ is when they access the same channel (i.e., $a_i = a_j$). We assume that the interference graph is symmetric such that $W_{ij} = W_{ji}$, $\forall i \neq j$.

Let $\mathbf{a} = (a_1, ..., a_N) \in \prod_{i=1}^{N} \mathcal{M}$ be the channel selection profile of all users.

Given the channel selection profile $\mathbf{a}$, the interference received by user i is given by

$$\gamma_i(\mathbf{a}) = \sum_{j \in \mathcal{N}, j \neq i} W_{ji} I_{\{a_i = a_j\}} \tag{8.10}$$

where $I_{\{A\}}$ is the indicator function with $I_{\{A\}} = 1$ indicating that event A is true and $I_{\{A\}} = 0$ otherwise. Then we define the individual utility function $u_i(\mathbf{a})$ as

$$u_i(\mathbf{a}) = -\gamma_n(\mathbf{a}) = -\sum_{j \in \mathcal{N}, j \neq i} W_{ji} I_{\{a_i = a_j\}} \tag{8.11}$$

where the negative sign comes from the convention that utility functions are typically the objectives for maximization. The individual utility reflects that each user is interested in reducing its received interference.

To capture that users with positive social ties intend to mitigate the interference to each other and users with negative social ties intend to enhance the interference to each other, we apply the SGUM framework to spectrum access. Then the social group utility of user i is given by

$$f_i(\mathbf{a}) = \sum_{j \in \mathcal{N}} s_{ij} u_j(\mathbf{a}) = -\sum_{j \in \mathcal{N}} s_{ij} \sum_{k \in \mathcal{N}, k \neq j} W_{kj} I_{\{a_j = a_k\}}. \tag{8.12}$$

8.5.3 Game Analysis

For the SGUM based spectrum access game, we have the following result for weakly balanced social networks.

Theorem 8.3. *For a weakly balanced social network, when the physical graph is homogeneous (i.e., $e_{ij}^P = 1$, $\forall i \neq j$ and $W \triangleq W_{ij} = W_{kl}$, $\forall i \neq j, k \neq l$), there exist SNEs for the SGUM based spectrum access game. Furthermore, at each SNE, for each social coalition $k \in \mathcal{K}$, there exist $M - q_k$ channels such that each channel is accessed by p_k users from social coalition k, and q_k channels such that each channel is accessed by $p_k + 1$ users from social coalition k, where $p_k \triangleq \lfloor N_k/M \rfloor$ and $q_k \triangleq (N_k \bmod M)$, $\forall k \in \mathcal{K}$.*

Proof: Let $n_k^{a_i}(\mathbf{a})$ denote the number of users from social coalition k that access channel a_i under strategy profile $\mathbf{a}$. Using the structural balance of the social network and (8.12), for any user i in any social coalition k, its social group utility can be expressed as

$$\begin{aligned} f_i(a_i, \mathbf{a}_{-i}) &= \sum_{l\in\mathcal{K}, l\neq k} \sum_{j\in\mathcal{N}_l} WI_{\{a_i=a_j\}} - \sum_{j\in\mathcal{N}_k, j\neq i} WI_{\{a_i=a_j\}} - \sum_{j\in\mathcal{N}, j\neq i} WI_{\{a_i=a_j\}} + g(\mathbf{a}_{-i}) \\ &= -\sum_{j\in\mathcal{N}_k, j\neq i} WI_{\{a_i=a_j\}} - \sum_{j\in\mathcal{N}_k, j\neq i} WI_{\{a_i=a_j\}} + g(\mathbf{a}_{-i}) \\ &= -2\sum_{j\in\mathcal{N}_k, j\neq i} WI_{\{a_i=a_j\}} + g(\mathbf{a}_{-i}) \\ &= -2(n_k^{a_i}(\mathbf{a}) - 1) + g(\mathbf{a}_{-i}) \end{aligned} \tag{8.13}$$

where $g(\mathbf{a}_{-i})$ is a function that is independent of user i's strategy a_i. Using (8.13), at the SNE $\mathbf{a}^*$, for any $a_i \neq a_i^*$ and any $i \in \mathcal{N}_k$, we have

$$f_i(a_i^*, \mathbf{a}_{-i}^*) = -2(n_k^{a_i^*}(\mathbf{a}^*) - 1) + g(\mathbf{a}_{-i}^*) \geq f_i(a_i, \mathbf{a}_{-i}^*) = -2n_k^{a_i}(\mathbf{a}^*) + g(\mathbf{a}_{-i}^*) \tag{8.14}$$

and it follows that $n_k^{a_i^*}(\mathbf{a}^*) \leq n_k^{a_i}(\mathbf{a}^*) + 1$. Therefore the desired result follows. □

Next we consider a game $\mathcal{G}' \triangleq \{\mathcal{K}, \{h_i\}_{i\in\mathcal{K}}, \{\mathcal{M}^{N_i}\}_{i\in\mathcal{K}}\}$ where each player $i \in \mathcal{K}$ consists of the users in social coalition i with strategy $x_i \triangleq (a_{n_i^1}, \cdots, a_{n_i^{N_i}})$ and payoff function $h_i(\mathbf{x}) \triangleq f_j(\mathbf{a})$ for any $j \in \mathcal{N}_i$. In other words, the users from the same coalition coordinate their strategies such that they behave as a single player. We can show that the NEs for the game $\mathcal{G}'$ coincide with the SNEs for the SGUM game $\mathcal{G} \triangleq \{\mathcal{N}, \{f_i\}_{i\in\mathcal{N}}, \{\mathcal{M}\}_{i\in\mathcal{N}}\}$.

Theorem 8.4. *A strategy profile $\mathbf{x}^*$ is an NE for the game $\mathcal{G}'$ if and only if the corresponding $\mathbf{a}^*$ is an SNE for the game $\mathcal{G}$.*

Proof: For any NE $\mathbf{x}^*$ for the game $\mathcal{G}'$, we must have $n_k^{a_j^*}(\mathbf{a}^*) \leq n_k^{a_j}(\mathbf{a}^*) + 1$, $\forall a_j \neq a_j^*$, $\forall j \in \mathcal{N}_i$, and $\forall i \in \mathcal{K}$. This is because otherwise we can find $x_i =$

$(a^*_{n_i^1}, \cdots, a_j, \cdots, a^*_{n_i^{N_i}})$ such that $h_i(x_i^*, \mathbf{x}^*_{-i}) = f_j(a_j^*, \mathbf{a}^*_{-j}) < f_j(a_j, \mathbf{a}^*_{-j}) = h_i(x_i, \mathbf{x}^*_{-i})$ using the proof of Theorem 8.3, which is a contradiction. Therefore, $\mathbf{a}^*$ is an SNE for the game $\mathcal{G}$. Then the desired result follows. □

8.6 Conclusion

In this chapter, we have developed a social group utility maximization game framework that spans the continuum from zero-sum game to standard non-cooperative game to network utility maximization. Then, we have studied two applications under this framework: random access control and multi-channel cooperative jamming. Our results shed light on the impact of social ties on users' strategies and system performance. We believe that this work will open a new door to exploring the impact of social behavior on mobile networking.

Chapter 9

SOCIAL GROUP UTILITY MAXIMIZATION FOR PERSONALIZED LOCATION PRIVACY IN MOBILE NETWORKS

9.1 Introduction

The proliferation of mobile devices is predicted to continue in the foreseeable future. In 2014, mobile phone shipments are projected to reach 1.9 billion units, about 7 times that of desktop and laptop combined [2]. With rapid growth of mobile networks, location-based services (LBS) have become increasingly popular recently (e.g., location-based navigation and recommendation). However, the providers of LBSs are often considered not trustworthy, due to the risk of leaking users' location information to other parties (e.g., sell users' location data). As a result, mobile users are exposed to potential privacy threats when using a LBS. Although a user can use a pseudonym for the LBS, an adversary can infer the user's real identity from its location traces (e.g., from the user's home and work addresses). To protect location privacy, an effective approach is to "confuse" the adversary using the notion of *anonymity* [120]: mobile users in physical proximity can change their pseudonyms simultaneously to form an anonymity set, so that the adversary cannot distinguish any of them from the others.

Clearly, as mobile devices are carried and operated by human beings, pseudonym change hinges heavily on human behavior. In particular, altruistic behaviors are widely observed among people with social ties[1]. It is then natural to ask "Is it possible to leverage social ties for pseudonym change to enhance location privacy?" The past few years have witnessed explosive growth of online social networks. In 2013, the number of online social network users worldwide has crossed 1.73 billion, nearly one quarter of the world's population [1]. As a result, social relationships influence

[1]In this chapter, "social tie" refers to "positive social tie" for brevity.

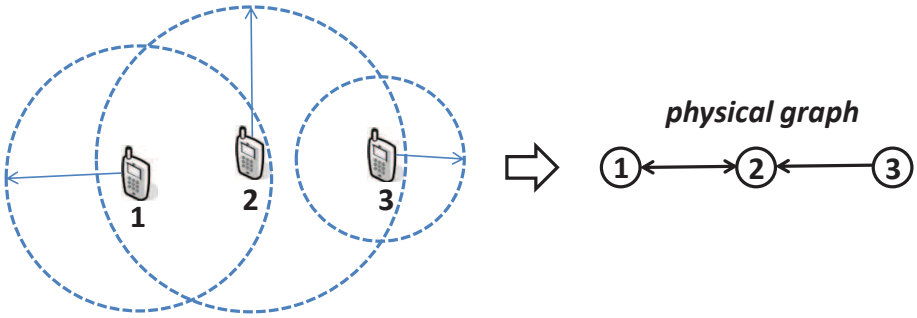


Figure 9.1: Illustration of anonymity model for personalized location privacy.

people's interactions with each other in an unprecedented manner. This motivates us to exploit social ties among users for pseudonym change to improve their location privacy. Since pseudonym change typically incurs considerable overhead (e.g., service interruption, resource consumption [121]), users need strong incentives (e.g., adequate privacy gain) to participate in pseudonym change. We caution that secure protocols are needed to hide users' real identities when social information is used (which will be elaborated further in Section 9.4.5).

A basic assumption commonly used in existing studies [121–123] is that all users participating in pseudonym change have the *same* anonymity set. However, from an individual user's perspective, the set of users that can obfuscate its pseudonym (i.e., its anonymity set) can be different from that of another user, depending on users' physical locations. For example, a user with a higher level of privacy sensitivity can have a smaller anonymity set than others. It is thus desirable to meet users' needs for *personalized location privacy*. To this end, we consider a general anonymity model where a user can define its specific anonymity set different from others'. For example, in Fig. 9.1, each user specifies its anonymity set for personalized location privacy by defining an anonymity range, which is a disk centered at the user's location. Then user 1 and 2 are out of user 3's anonymity range and thus are not in user 3's anonymity set (represented by no direct edge from user 1 or 2 to user 3); user

1 and 3 are within user 2's anonymity range and thus are in user 2's anonymity set (represented by directed edges from user 1 and 3 to user 2).

In this chapter, we leverage mobile users' social tie structure to motivate them to participate in pseudonym change. Drawing on a social group utility maximization (SGUM) framework recently developed in [51], we cast users' decision making of whether to participate in pseudonym change as a socially-aware pseudonym change game (PCG). The SGUM framework captures the impact of users' diverse social ties on the interactions of their mobile devices subject to diverse physical relationships. The PCG is based on a general anonymity model that allows each user to have its specific anonymity set. In the SGUM-based PCG, each user aims to maximize its social group utility, defined as the sum of its individual utility and the weighted sum of its social friends' individual utilities. For the SGUM-based PCG, we are interested in answering the following important questions: Does the game admit a socially-aware Nash equilibrium (SNE)? What is the system efficiency of SNEs? How can we efficiently find an SNE with desirable properties?

The main contributions of this chapter can be summarized as follows.

- We propose a framework where mobile users' social tie structure is leveraged to motivate them to participate in pseudonym change, based on a general anonymity model that allows each user to have its specific anonymity set for personalized location privacy. Taking a social group utility approach, we cast users' decision making of whether to change pseudonyms as a pseudonym change game.

- For the SGUM-based PCG, we first show that there always exists a socially-aware Nash equilibrium. Then we quantify the system efficiency of the SNE by bounding the gap between the optimal social welfare and the social welfare of

the "best" SNE, which is the SNE that achieves the maximum social welfare among all SNEs. We observe that the best SNE is difficult to compute in general, and the often used best response updates would converge to an SNE with a lower social welfare. Then we develop a greedy algorithm that myopically determines users' strategies, based on the social group utility derived from only the users whose strategies have already been determined. We first show that the algorithm can efficiently find an SNE that is socially-aware coalition proof, and thus is also Pareto-optimal. We then show that the SNE has a social welfare no less than that of the best SNE for the socially-oblivious PCG. Next we show that the social welfare of the SNE increases when social ties increase, and its performance gap with respect to the optimal social welfare is bounded above. We further show that the Pareto-optimal SNE can be achieved in a distributed manner.

- We evaluate the performance of the SNE by extensive simulation results for Erdos-Renyi model and real dataset based social networks. We demonstrate the impact of various parameters on system efficiency, individual users' performance, and computational complexity of the proposed algorithm. Numerical results corroborate that social welfare can be significantly improved by exploiting social ties.

The rest of this chapter is organized as follows. Related work are reviewed in Section 9.2. In Section 9.3, we formulate the socially-aware pseudonym change game based on a general anonymity model for personalized location privacy, under the social group utility maximization framework. Section 9.4 focuses on the analysis of the SGUM-based PCG. Numerical results are presented in Section 9.5. Section 9.6 concludes this chapter.

9.2 Related Work

With growing concerns for location privacy arising from pervasive mobile communication and computing, a great deal of research have been done to protect mobile users' location privacy. This work falls in the category of anonymity-based approaches [120–124]. Earlier studies [120, 125] show that an adversary can infer the real identity of a mobile user by analyzing the spatial-temporal correlation of its location traces. To overcome this vulnerability, pseudonyms should not only be changed over time but also be obfuscated across space to prevent inference attacks. Inspired by the notion of k-anonymity, Beresford and Stajano [120] introduced the notion of mix zone. By changing pseudonyms within a mix zone, users can make their new pseudonyms undistinguishable to the adversary. While the mix zone model assumes that all users have the same anonymity set, the general anonymity model proposed in this chapter allows each user to define its specific anonymity set different from others'. A few work have studied users' interactions in pseudonym change based on game-theoretic models. The mix zone based pseudonym change has been studied in [121] as a non-cooperative game with complete or incomplete information. Auction-based mechanisms have been designed in [123] to incentive users to participate in pseudonym change. To our best knowledge, this chapter is the first to exploit social relationships to improve location privacy by pseudonym change.

There have been much work on incentive design for stimulating user cooperation in networks. Most of the existing approaches are based on (virtual) currency [39–42], in which a user earns currency by providing cooperation to others and spends currency to receive cooperation from others. However, a major drawback of using currency is that it relies on a centralized authority (e.g., a bank) and typically incurs a high implementation overhead, due to the need to inhibit malicious manip-

ulation of the currency among users without mutual trust. Therefore, it is appealing to motivate users to cooperate without using currency.

The social aspect of mobile networking is an emerging paradigm for network design and optimization [126]. Social contact patterns have been exploited for efficient data forwarding and dissemination in delay tolerant networks [65, 66]. Social trust and social reciprocity have been leveraged in [16] to enhance cooperative D2D communication based on a coalitional game. Recently, a social group utility maximization (SGUM) framework is developed in [51–53], which captures the impact of users' diverse social ties on the interactions of their devices subject to diverse physical coupling. A primary merit of this framework is that it provides rich modeling flexibility and spans the continuum between non-cooperative game and network utility maximization, two traditionally disjoint paradigms for network optimization.

9.3 Model and Problem Formulation

9.3.1 Privacy Threat in Location-based Services

We consider a mobile network where users obtain their locations via mobile devices that are capable of localization (e.g., based on GPS or wireless access points). Users send their locations to a LBS provider for a certain LBS (e.g., location-based navigation or recommendation), and the LBS provider feedbacks the desired results to the users based on their reported locations. To protect privacy, each user uses a pseudonym as its identity for the LBS.

As in [120,123,124], we assume that the LBS provider is untrusted, i.e., it may leak users' location traces to an adversary. For example, the adversary may steal the location data by hacking into the LBS system. The adversary aims to learn the real identity of a user by analyzing the locations visited by the user's pseudonym. We also assume that users are honest-but-curious in the sense that each user honestly follows the protocols with others (which will be discussed in Section 9.4.5), but is curious

about others' private information. We further assume that the adversary may collude with a limited number of users to gain useful information for inferring a user's real identity.

The use of pseudonym allows short-term reference to a user (e.g., one pseudonym can be used for the navigation of an entire trip between two locations), which is useful for many LBSs and does not disclose private information. However, long-term linking among a user's locations should be prevented, as it may reveal sufficient information for inferring the user's real identity [125,127,128]. Although a user may hide explicit linking among its locations by changing its pseudonym, the adversary can still link different pseudonyms of the user by exploiting the spatial-temporal correlation of its locations. For example, consider a user that visits location l_1 with pseudonym Alice at time t_1, and then visits location l_2 which is close to location l_1 with pseudonym Bob at time t_2. If the adversary observes from the location traces that no other user changes its pseudonym between time t_1 and t_2, or there exists such a user but it does not visit any location close to location l_1 or l_2, then the adversary can infer that pseudonym Alice and Bob must refer to the same user, since only the same user can visit both location l_1 and l_2 within the limited period between time t_1 and t_2.

9.3.2 Pseudonym Change for Personalized Location Privacy

To protect location privacy from inference attacks, an effective approach is based on the notion of anonymity: users in physical proximity can coordinate their pseudonym changes to happen simultaneously [120], so that the adversary cannot link their pseudonyms before the changes to their respective pseudonyms after the changes. Existing studies [121–123] assume that all users participating in pseudonym change have the *same* anonymity set. However, based on an individual user's belief of the adversary's power against its location privacy (e.g., the adversary's side information about that user), the set of users that it believes can obfuscate its pseudonym (i.e.,

its anonymity set) can be different from that of another user. Thus motivated, we consider a general anonymity model that can meet users' needs for personalized location privacy, depending on users' physical locations. In particular, each user specifies an *anonymity range* (a physical area) such that the set of users within the anonymity range constitute that user's *potential anonymity set*. For example, a user's anonymity range can be a disk centered at the user's location, with a large radius indicating a low level of privacy sensitivity (as illustrated in Fig. 9.1). Note that for two users at different locations, their anonymity ranges are different even when they have the same shape (e.g., two disks with the same radius but different centers), and thus their potential anonymity sets can be different.

Formally, consider a set of users $\mathcal{N} \triangleq \{1, \cdots, N\}$ where each user i makes a decision a_i on whether or not to participate in pseudonym change, denoted by $a_i = 1$ and $a_i = 0$, respectively. Based on users' physical locations, the privacy gain perceived by a user participating in pseudonym change depends on *which* users also participate. Each user i incurs a cost $c_i > 0$ to participate in pseudonym change. This cost is due to a number of factors, e.g., the participating users should stop using the LBS for a period of time. Based on the general anonymity model, the physical coupling among users can be captured by a physical graph $(\mathcal{N}, \mathcal{E}^P)$, where user j is connected by a directed edge $e^P_{ji} \in \mathcal{E}^P$ to user i if user j is in user i's potential anonymity set, denoted by $\mathcal{N}_{i-}$ (i.e., $j \in \mathcal{N}_{i-}$). Note that the physical coupling between two users can be asymmetric. The privacy gain perceived by a participating user i is defined as its *anonymity set size*, i.e., the number of participating users in $\mathcal{N}_{i-}$. Note that the anonymity set size is a widely adopted privacy metric[2] for anonymity-based approaches. For example, k-anonymity is used as the privacy metric in [123, 124],

[2]Another privacy metric is the entropy of the adversary's uncertainty of a user's pseudonym. However, it is usually difficult to compute since it requires probability distribution which is difficult to attain.

where a user achieves location privacy if its pseudonym cannot be distinguished among k users. Then the individual utility of user i, denoted by u_i, is given by

$$u_i(a_i, \boldsymbol{a}_{-i}) \triangleq a_i \left(\sum_{j \in \mathcal{N}_{i-}} a_j - c_i \right) \tag{9.1}$$

where $\boldsymbol{a}_{-i}$ denotes the vector of the strategies of all users except user i. If a user participates, its individual utility is its privacy gain minus its participation cost; otherwise, it is zero. Note that c_i is a relative cost compared to privacy gain.

9.3.3 Social Group Utility Maximization (SGUM)

Social relationships play an increasingly important role in people's interactions with each other. One important attribute of social relationship is that people are altruistic to their social "friends" (including friends, family, colleagues, etc.), as they care about their social friends' welfare. As a result, a user would take into account the effect of its behavior on its social friends. Recently, a social group utility maximization framework has been developed in [51], which captures the impact of mobile users' diverse social ties on the interactions of their mobile devices subject to diverse physical relationships.

Appealing to the SGUM framework [51], we model the social tie structure among the users in $\mathcal{N}$ by a social graph $(\mathcal{N}, \mathcal{E}^S)$, where user i is connected by a directed edge $e_{ij}^S \in \mathcal{E}^S$ to user j if user i has a *social tie* with user j, denoted by s_{ij}. We assume that each user i's social tie to itself is $s_{ii} = 1$, and we normalize user i's social tie to user $j \neq i$ as $s_{ij} \in (0, 1]$, which quantifies the extent to which user i cares about user j relative to user i cares about itself. We also assume that $s_{ij} = 0$ if no social tie exists from user i to user j.

To take into account the social ties among users, each user i aims to maximize its *social group utility*, defined as

$$f_i(a_i, \boldsymbol{a}_{-i}) \triangleq u_i(a_i, \boldsymbol{a}_{-i}) + \sum_{j \in \mathcal{N}_{i+}} s_{ij} u_j(a_j, \boldsymbol{a}_{-j}) \tag{9.2}$$

where $\mathcal{N}_{i+}$ denotes the set of users whose potential anonymity sets include user i. Note that a user's social group utility consists of its own individual utility and the sum of the individual utilities of the other users weighted by social ties. Therefore, the social group utility captures both physical coupling and social coupling among users in a unified way. Also note that a user does not need to know the individual utilities of its social friends (which may be their private information) to make the decision (as will be shown in equation (9.3)). In Section 9.4.5, we will discuss how social information can be used while preserving the privacy of users' real identities with respect to each other.

9.3.4 SGUM based Pseudonym Change Game

Based on the SGUM framework, users' socially-aware decision making for pseudonym change boils down to a social group utility maximization game. Specifically, each user $i \in \mathcal{N}$ is a *player* and its *strategy*[3] is $a_i \in \{0, 1\}$. Let $\boldsymbol{a} = (a_1, \cdots, a_n)$ denote the *strategy profile* consisting of all users' strategies. The *payoff* function of a user is defined as its social group utility function. Given the strategies of other users, each user i aims to choose the *best response* strategy that maximizes its social group utility:

$$\underset{a_i}{\text{maximize}} \quad f_i(a_i, \boldsymbol{a}_{-i}), \ \forall i \in \mathcal{N}.$$

Similar in spirit to the Nash equilibrium [17] of a standard non-cooperative game, the following concept applies to the SGUM game.

Definition 9.1 (Socially-aware Nash Equilibrium [51])**.** *A strategy profile* $\boldsymbol{a}^{sne} = (a_1^{sne}, \cdots, a_n^{sne})$ *is a* socially-aware Nash equilibrium *(SNE) of the SGUM-based PCG if no user can improve its social group utility by unilaterally changing its strategy, i.e.,*

$$a_i^{sne} = \arg \max_{a_i \in \{0,1\}} \quad f_i(a_i, \boldsymbol{a}_{-i}), \ \forall i \in \mathcal{N}.$$

[3]As we focus on pure strategies in this work, we use "strategy" and "action" interchangeably.

Due to the rational and autonomous nature of users, an SNE is a stable outcome which is acceptable for all users.

For the sake of system efficiency, a natural objective is to maximize the *social welfare* of the system, which is the total individual utility of all users denoted by $v(\boldsymbol{a}) \triangleq \sum_{i\in\mathcal{N}} u_i(\boldsymbol{a})$.

Definition 9.2 (Social Optimal [17])**.** *A strategy profile* $\boldsymbol{a}^* = (a_1^*, \cdots, a_n^*)$ *is* social optimal *if it achieves the maximum social welfare among all strategy profiles, i.e.,*

$$v(\boldsymbol{a}^*) \geq v(\boldsymbol{a}), \forall \boldsymbol{a}.$$

Although the social optimal profile is the best outcome in terms of system efficiency, it is often not acceptable for all users. Then, it is desirable to achieve the "best" SNE, i.e., the SNE that achieves the maximum social welfare among all SNEs (referred to as "the best SNE").

Another desirable property for system efficiency is given below.

Definition 9.3 (Pareto-Optimal [17])**.** *A strategy profile* $\boldsymbol{a}^{po} = (a_1^{po}, \cdots, a_n^{po})$ *is* Pareto-optimal *if there does not exist a* Pareto-superior *profile* $\boldsymbol{a}' = (a_1', \cdots, a_n')$ *such that no user achieves a worse individual utility while at least one user achieves a better individual utility, i.e.,*

$$u_i(a_i', \boldsymbol{a}_{-i}') \geq u_i(a_i^{po}, \boldsymbol{a}_{-i}^{po}),\ \forall i \in \mathcal{N}$$

with at least one strict inequality.

For the SGUM-based PCG, we are interested in answering the following important questions:

- Does the game admit any SNE? If yes, what is the system efficiency of the best SNE?

- How can we efficiently find an SNE? What is the system efficiency of this SNE? What are the desirable properties of this SNE if any?

To answer these questions, in Section 9.4 we will focus on the analysis of the SGUM-based PCG.

9.4 SGUM based Pseudonym Change Game

9.4.1 Benchmark: Socially-oblivious Pseudonym Change Game

As a benchmark, we start with a basic case of the PCG: the PCG for socially-oblivious users (SO-PCG), i.e., $s_{ij} = 0$, $\forall i \neq j$. In this case, each user is selfish and the social group utility degenerates to the individual utility.

For SO-PCG, there can exist multiple SNEs[4] with different values of social welfare. For system efficiency, it is desirable to achieve the best SNE (i.e., the SNE that achieves the maximum social welfare among all SNEs). Let $\mathcal{N}_1(\boldsymbol{a}) \triangleq \{i \in \mathcal{N} | a_i = 1\}$ and $\mathcal{N}_0(\boldsymbol{a}) \triangleq \{i \in \mathcal{N} | a_i = 0\}$ denote the set of participating users and non-participating users under strategy profile $\boldsymbol{a}$, respectively. For example, in Fig. 9.2, there are three SNEs $\boldsymbol{a}^1$, $\boldsymbol{a}^2$, and $\boldsymbol{a}^3$ with $\mathcal{N}_1(\boldsymbol{a}^1) = \{1, 2, 4, 5\}$, $\mathcal{N}_1(\boldsymbol{a}^2) = \{4, 5\}$, and $\mathcal{N}_1(\boldsymbol{a}^3) = \varnothing$, respectively (the number beside a user is its cost). We can see that $\boldsymbol{a}^1$ is Pareto-superior to $\boldsymbol{a}^2$, $\boldsymbol{a}^3$, and hence is the best SNE. To find the best SNE, we can use best response updates as described in Algorithm 9: with all users' actions initially set to 1, each user asynchronously updates (i.e., no two users update at the same time) its action from 1 to 0 if it increases its individual utility. For the example in Fig. 9.2, using Algorithm 9, we have $u_6 = 2 - 2.2 < 0 \Rightarrow a_6 = 0 \Rightarrow u_3 = 0 - 0.5 < 0 \Rightarrow a_3 = 0 \Rightarrow \boldsymbol{a} = (1, 1, 0, 1, 1, 0)$ is an SNE. We show that the algorithm indeed finds the best SNE as follows.

[4]For SO-PCG, an SNE is equivalent to a NE for a standard non-cooperative game. For consistency of terminology, we still call it "SNE" in this case.

Algorithm 9: Compute the best SNE for SO-PCG

1 $\boldsymbol{a} \leftarrow (1, \cdots, 1)$;
2 **while** $\exists i \in \mathcal{N}$ *such that* $a_i = 1$ *and* $u_i(1, \boldsymbol{a}_{-i}) < 0$ **do**
3 $\quad a_i \leftarrow 0$;
4 **end**
5 **return** $\boldsymbol{a}^o \leftarrow \boldsymbol{a}$;

Proposition 9.1. *For SO-PCG, Algorithm 9 computes the best SNE which achieves the maximum social welfare among all SNEs.*

Proof: For any user i with $a_i^o = 0$, let $\boldsymbol{a}'$ be the strategy profile right before user i's action is changed from 1 to 0 during the execution of Algorithm 9. Since $\boldsymbol{a}^o_{-i} \leq \boldsymbol{a}'_{-i}$, we have $u_i(1, \boldsymbol{a}^o_{-i}) \leq u_i(1, \boldsymbol{a}'_{-i}) < 0$ due to the condition in line 2. Therefore, $a_i^o = 0$ is the best response strategy for user i. According to the condition in line 2, $a_j^o = 1$ is the best response strategy for any user j with $a_j^o = 1$. Thus the strategy profile $\boldsymbol{a}^o$ is an SNE.

Next we show that $\boldsymbol{a}^o$ achieves the maximum social welfare among all SNEs. It suffices to show that $\boldsymbol{a}^o$ is Pareto-superior to any other SNE. To this end, we first show that a profile $\boldsymbol{a}'$ is not an SNE if $\mathcal{N}_1(\boldsymbol{a}') \setminus \mathcal{N}_1(\boldsymbol{a}^o) \neq \varnothing$. Suppose such $\boldsymbol{a}'$ is an SNE. Let $i \in \mathcal{N}_1(\boldsymbol{a}') \setminus \mathcal{N}_1(\boldsymbol{a}^o)$ be the first user among $\mathcal{N}_1(\boldsymbol{a}') \setminus \mathcal{N}_1(\boldsymbol{a}^o)$ whose action is changed to 0, and $\bar{\boldsymbol{a}}$ be the profile right before that change. Since $\boldsymbol{a}' \leq \bar{\boldsymbol{a}}$, we have $u_i(\boldsymbol{a}') \leq u_i(\bar{\boldsymbol{a}}) < 0 = u_i(0, \boldsymbol{a}'_{i-})$ due to that 0 is the best response strategy. This shows that $\boldsymbol{a}'$ is not an SNE. Therefore, for any SNE $\boldsymbol{a}'$ other than $\boldsymbol{a}^o$, we must have $\boldsymbol{a}' < \boldsymbol{a}^o$. Then for each $i \in \mathcal{N}_1(\boldsymbol{a}')$, we have $u_i(\boldsymbol{a}') \leq u_i(\boldsymbol{a}^o)$. For each $i \in \mathcal{N}_0(\boldsymbol{a}')$, since $\boldsymbol{a}^o$ is an SNE, we have $u_i(\boldsymbol{a}') = 0 = u_i(0, \boldsymbol{a}^o_{-i}) \leq u_i(\boldsymbol{a}^o)$. Therefore $\boldsymbol{a}^o$ is Pareto-superior to $\boldsymbol{a}'$. Thus we show that $\boldsymbol{a}^o$ is the best SNE. □

As the best SNE achieves the maximum system efficiency among all SNEs, we will use the best SNE for SO-PCG as the benchmark for the general case of the

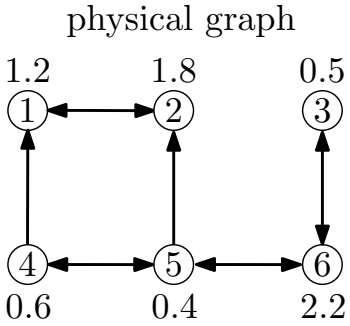


Figure 9.2: An example of SO-PCG.

PCG: the PCG for socially-aware users (SA-PCG), i.e., $\exists i \neq j$ such that $s_{ij} > 0$.

9.4.2 Existence and Efficiency of SNE

In this subsection we study the existence and efficiency of SNE for SA-PCG. We first establish the existence of SNE. Using (9.1) and (9.2), we have

$$\begin{aligned} & f_i(1, \boldsymbol{a}_{-i}) - f_i(0, \boldsymbol{a}_{-i}) \\ = & \, u_i(1, \boldsymbol{a}_{-i}) - u_i(0, \boldsymbol{a}_{-i}) + \sum_{j \in \mathcal{N}_{i+}} s_{ij} \left(u_j(1, \boldsymbol{a}_{-i}) - u_j(0, \boldsymbol{a}_{-i}) \right) \\ = & \sum_{j \in \mathcal{N}_{i-}} a_j - c_i + \sum_{j \in \mathcal{N}_{i+}} s_{ij} a_j. \end{aligned} \tag{9.3}$$

It is clear from (9.3) that no user participating is always an SNE. Then we have the following result.

Theorem 9.1. *For SA-PCG, there exists at least one SNE.*

Next we show an important property of the social group utility function. It follows from (9.3) that

$$\begin{aligned} & f_i(1, \boldsymbol{a}_{-i}) - f_i(0, \boldsymbol{a}_{-i}) - \left(f_i(1, \boldsymbol{a}'_{-i}) - f_i(0, \boldsymbol{a}'_{-i}) \right) \\ = & \sum_{j \in \mathcal{N}_{i-}} a_j - c_i + \sum_{j \in \mathcal{N}_{i+}} s_{ij} a_j - \left(\sum_{j \in \mathcal{N}_{i-}} a'_j - c_i + \sum_{j \in \mathcal{N}_{i+}} s_{ij} a'_j \right) \\ = & \sum_{j \in \mathcal{N}_{i-}} (a_j - a'_j) + \sum_{j \in \mathcal{N}_{i+}} s_{ij} (a_j - a'_j). \end{aligned} \tag{9.4}$$

Let $\boldsymbol{a} \leq \boldsymbol{a}'$ denote entry-wise inequality (i.e., $a_i \leq a'_i$, $\forall i \in \mathcal{N}$). Using (9.4), we have the following result.

Lemma 9.1. *If* $\boldsymbol{a}_{-i} \leq \boldsymbol{a}'_{-i}$, *then* $f_i(1, \boldsymbol{a}_{-i}) - f_i(0, \boldsymbol{a}_{-i}) \leq f_i(1, \boldsymbol{a}'_{-i}) - f_i(0, \boldsymbol{a}'_{-i})$.

Remark: Pseudonym change is a behavior with *network effect* such that each participating user benefits more when more users participate. As a result, Lemma 9.1 shows that a user's social group utility is a *supermodular* function: the marginal gain of social group utility by participating increases when more users participate. This implies that if a user's best response strategy is to participate, then it remains so if more users participate; if a user's best response strategy is to not participate, then it remains the so if less users participate.

To quantify the system efficiency of the SNE, we provide a bound of the gap between the social welfare of the best SNE and the optimal social welfare in the following result.

Theorem 9.2. *The performance gap between the maximum social welfare among all SNEs and the optimal social welfare is upper bounded by* $\sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{N}_{i+}} (1 - s_{ij})$.

Proof: We first show that we can construct an SNE from the social optimal profile $\boldsymbol{a}^*$ using best response updates: with all users' actions initially set according to the social optimal profile, a user's action is changed from 1 to 0 if that can improve its social group utility. To this end, we first show that this algorithm can terminate. Without loss of generality, we assume that there does not exist $\boldsymbol{a}'$ with $\boldsymbol{a}' > \boldsymbol{a}^*$ such that $v(\boldsymbol{a}') \geq v(\boldsymbol{a}^*)$. Suppose there exists $i \in \mathcal{N}_0(\boldsymbol{a}^*)$ such that $f_i(1, \boldsymbol{a}^*_{-i}) \geq f_i(0, \boldsymbol{a}^*_{-i})$.

Then we have

$$v(1, \boldsymbol{a}^*_{-i}) - v(0, \boldsymbol{a}^*_{-i}) = u_i(1, \boldsymbol{a}^*_{-i}) + \sum_{j \in \mathcal{N}_{i+}} a_j$$

$$\geq u_i(1, \boldsymbol{a}^*_{-i}) + \sum_{j \in \mathcal{N}_{i+}} s_{ij} a_j = f_i(1, \boldsymbol{a}^*_{-i}) - f_i(0, \boldsymbol{a}^*_{-i}) \geq 0$$

where the first equality follows a similar manipulation as in (9.3), and the second equality follows from (9.3). This contradicts the previous assumption. Therefore we must have $f_i(1, \boldsymbol{a}^*_{-i}) < f_i(0, \boldsymbol{a}^*_{-i})$ for each $i \in \mathcal{N}_0(\boldsymbol{a}^*)$. Then, according to Lemma 9.1, the algorithm must terminate and results in a profile $\boldsymbol{a}^b$, which is an SNE and satisfies that $\boldsymbol{a}^* \geq \boldsymbol{a}^b$.

Next we show an upper bound on $v(\boldsymbol{a}^*) - v(\boldsymbol{a}^b)$. For any $i \in \mathcal{N}_1(\boldsymbol{a}^*) \setminus \mathcal{N}_1(\boldsymbol{a}^b)$, let $\bar{\boldsymbol{a}}$ be the profile right before a_i is changed to 0 in the algorithm. Then we have

$$v_i^\Delta \triangleq v(1, \bar{\boldsymbol{a}}_{-i}) - v(0, \bar{\boldsymbol{a}}_{-i}) = u_i(1, \bar{\boldsymbol{a}}_{-i}) + \sum_{j \in \mathcal{N}_{i+}} \bar{a}_j$$

$$= u_i(1, \bar{\boldsymbol{a}}_{-i}) + \sum_{j \in \mathcal{N}_{i+}} s_{ij} \bar{a}_j + \sum_{j \in \mathcal{N}_{i+}} (1 - s_{ij}) \bar{a}_j$$

$$= f_i(1, \bar{\boldsymbol{a}}_{-i}) - f_i(0, \bar{\boldsymbol{a}}_{-i}) + \sum_{j \in \mathcal{N}_{i+}} (1 - s_{ij}) \bar{a}_j$$

$$< \sum_{j \in \mathcal{N}_{i+}} (1 - s_{ij}) \bar{a}_j \leq \sum_{j \in \mathcal{N}_{i+}} (1 - s_{ij})$$

where the last equality follows from (9.3), and the first inequality is due to that 0 is the best response strategy. Therefore we have

$$v(\boldsymbol{a}^*) - v(\boldsymbol{a}^b) = \sum_{i \in \mathcal{N}_1(\boldsymbol{a}^*) \setminus \mathcal{N}_1(\boldsymbol{a}^b)} v_i^\Delta \leq \sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{N}_{i+}} (1 - s_{ij}).$$

□

Remark: Theorem 9.2 shows that the performance gap decreases as social ties increase. In particular, when users are socially-oblivious (i.e., $s_{ij} = 0, \forall i \neq j$), the performance gap reaches the maximum, and the best SNE for SA-PCG degenerates

to the best SNE for SO-PCG; when users are fully altruistic (i.e., $s_{ij} = 1$, $\forall i \neq j$), the performance gap becomes 0, and the best SNE degenerates to the social optimal strategy profile. This demonstrates that the best SNE spans the continuum between a NE for a standard non-cooperative game and the optimal solution for network utility maximization, two traditionally disjoint paradigms for network optimization.

9.4.3 Computing SNE

In this subsection, we turn our attention to finding an SNE with desirable properties.

For the PCG for fully altruistic users (i.e., $s_{ij} = 1$, $\forall i \neq j$), we can see that the social optimal strategy profile $\boldsymbol{a}^*$ is an SNE, and is the solution to the following problem:

$$\begin{aligned} \underset{\boldsymbol{a}}{\text{maximize}} \quad & \sum_{i \in \mathcal{N}} a_i \left(\sum_{j \in \mathcal{N}_{i-}} a_j - c_i \right) \\ \text{subject to} \quad & a_i \in \{0, 1\}, \ \forall i \in \mathcal{N}. \end{aligned} \tag{9.5}$$

Note that problem (9.5) is an integer quadratic programming, which is in general difficult to solve[5]. Since the PCG for fully altruistic users is a special case of SA-PCG, it is also difficult to compute the best SNE for SA-PCG. Based on this observation, our objective below is to efficiently compute an SNE with other desirable properties.

To compute an SNE for SA-PCG, a plausible approach is to use best response updates in a similar way as Algorithm 9 for SO-PCG: with all users' actions initially set to 1, each user asynchronously updates its action from 1 to 0 if it increases its social group utility. Using Lemma 9.1, we can show that such best response updates always converge to an SNE. However, it has drawbacks: the SNE may not be Pareto-optimal and its social welfare may be worse than that of an SNE for SO-PCG. For example, in Fig. 9.3, using best response updates, we have $f_1(1,1) - f_1(0,1) = 1 - 1.5 + 0.8 > 0$, $f_2(1,1) - f_2(1,0) = 1 - 1.5 + 0.8 > 0$, and hence $\boldsymbol{a}^1$ with $\mathcal{N}_1(\boldsymbol{a}^1) = \{1, 2\}$ is an SNE.

[5]We conjecture that problem (9.5) is an NP-hard problem.

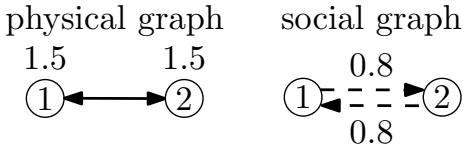


Figure 9.3: An example of SA-PCG where best response updates is not desirable.

However, it is not Pareto-optimal, since it is Pareto inferior to $\boldsymbol{a}^2$ with $\mathcal{N}_1(\boldsymbol{a}^2) = \varnothing$ as $u_1(0,0) = u_2(0,0) = 0 > 1 - 1.5 = u_1(1,1) = u_2(1,1)$. Furthermore, the social welfare of $\boldsymbol{a}^1$ is less than that of $\boldsymbol{a}^2$ as $v(1,1) = -1 < 0 = v(0,0)$, where $\boldsymbol{a}^2$ is also an SNE for SO-PCG. Thus motivated, our objective below is to efficiently find an SNE such that 1) it is Pareto-optimal and 2) its social welfare is no less than that of the best SNE for SO-PCG, which is the benchmark.

9.4.3.1 Algorithm Design

To this end, we design an algorithm as described in Algorithm 10. The main idea of the algorithm is to greedily determine users' strategies, depending on the social group utility derived from *the users whose strategies have been determined* (referred to as "determined users"), denoted by

$$f_i'(a_i, \boldsymbol{a}_{-i}) \triangleq u_i(a_i, \boldsymbol{a}_{-i}) + \sum_{j \in \mathcal{N}_{i+} \setminus \overline{\mathcal{N}}} s_{ij} u_j(a_i, \boldsymbol{a}_{-i})$$

where $\overline{\mathcal{N}}$ denotes the set of users whose strategies have not been determined (referred to as "undetermined users"). An undetermined user's action is fixed once it becomes determined.

Specifically, the algorithm proceeds in rounds and each round consists of phase I and phase II. In phase I, with all undetermined users' actions initially set to 1, an undetermined user's action is changed from 1 to 0 if it increases its social group utility derived from the determined users, i.e.,

$$f_i'(1, \boldsymbol{a}_{-i}) - f_i'(0, \boldsymbol{a}_{-i}) = u_i(1, \boldsymbol{a}_{-i}) + \sum_{j \in \mathcal{N}_{i+} \setminus \overline{\mathcal{N}}} s_{ij} a_j < 0$$

Algorithm 10: Compute the SNE for SA-PCG

$\overline{\mathcal{N}} \leftarrow \mathcal{N}$;
repeat
  // Phase I;
  $\boldsymbol{a} \leftarrow (1, \cdots, 1)$, $\mathcal{N}_I \leftarrow \overline{\mathcal{N}}$;
  while $\exists i \in \overline{\mathcal{N}}$ *such that* $u_i(1, \boldsymbol{a}_{-i}) + \sum_{j \in \mathcal{N}_{i+} \setminus \overline{\mathcal{N}}} s_{ij} a_j < 0$ **do**
    $a_i \leftarrow 0$, $\mathcal{N}_I \leftarrow \mathcal{N}_I \setminus \{i\}$;
  end
  // Phase II;
  $\overline{\mathcal{N}} \leftarrow \overline{\mathcal{N}} \setminus \mathcal{N}_I$, $\mathcal{N}_{II} \leftarrow \varnothing$;
  while $\exists i \in \overline{\mathcal{N}}$ *such that* $u_i(1, \boldsymbol{a}_{-i}) + \sum_{j \in \mathcal{N}_{i+} \setminus \overline{\mathcal{N}}} s_{ij} a_j \geq 0$ **do**
    $a_i \leftarrow 1$, $\overline{\mathcal{N}} \leftarrow \overline{\mathcal{N}} \setminus \{i\}$, $\mathcal{N}_{II} \leftarrow \mathcal{N}_{II} \cup \{i\}$;
  end
until $\mathcal{N}_I \cup \mathcal{N}_{II} = \varnothing$;
return $\boldsymbol{a}^e \leftarrow \boldsymbol{a}$;

until no such user exists. Then the undetermined users whose actions remain 1 become determined and their actions are fixed to 1. In phase II, with all undetermined users' actions initially set to 0, an undetermined user becomes determined and its action is fixed to 1 if it increases its social group utility derived from the determined users, until no such user exists. The algorithm terminates when no undetermined user becomes determined during either phase I or phase II of a round.

We use an example in Fig. 9.4 to illustrate how Algorithm 10 works and outline the steps as follows (the number beside a user is its cost; the number beside a social edge is its social tie).

- Phase I of 1st round: $u_1 = 1 - 1.2 < 0 \Rightarrow a_1 = 0$; $u_5 = 1 - 1.5 < 0 \Rightarrow a_5 = 0$; $u_4 = 0 - 1.2 < 0 \Rightarrow a_4 = 0 \Rightarrow u_3 = 2 - 2.5 < 0 \Rightarrow a_3 = 0$; $u_7 = 2 - 2.2 < 0 \Rightarrow a_7 = 0 \Rightarrow u_8 = 0 - 0.8 < 0 \Rightarrow a_8 = 0$; $u_2 = 1 - 0.6 > 0$; $u_6 = 1 - 0.8 > 0$; $\mathcal{N}_I = \{2, 6\}$.
- Phase II of 1st round: $u_5 + s_{56} = 1 - 1.5 + 0.6 > 0 \Rightarrow a_5 = 1 \Rightarrow u_1 + s_{12} = 1 - 1.2 + 0.5 > 0 \Rightarrow a_1 = 1$; $u_3 + s_{32} = 1 - 2.5 + 0.8 < 0$; $u_7 + s_{76} = 1 - 2.2 + 0.5 < 0$;

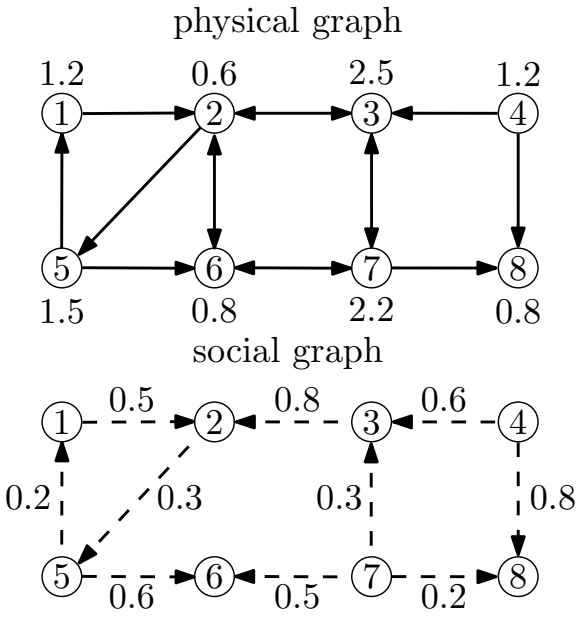


Figure 9.4: An example that illustrates how Algorithm 10 works.

$u_4 = 0 - 1.2 < 0$; $u_8 = 0 - 0.8 < 0$; $\mathcal{N}_{II} = \{1, 5\}$.

- Phase I of 2nd round: $u_4 = 0 - 1.2 < 0 \Rightarrow a_4 = 0 \Rightarrow u_3 + s_{32} = 2 - 2.5 + 0.8 > 0$; $u_7 + s_{76} = 2 - 2.2 + 0.5 > 0$; $u_8 = 1 - 0.8 > 0$; $\mathcal{N}_I = \{3, 7, 8\}$.

- Phase II of 2nd round: $u_4 + s_{43} + s_{48} = 0 - 1.2 + 0.6 + 0.8 > 0 \Rightarrow a_4 = 1$; $\mathcal{N}_{II} = \{4\}$.

Since the size of the set of undetermined users $\overline{\mathcal{N}}$ is upper bounded by N, the computational complexity of phase I and phase II of a round is bounded by $O(N^2)$. Since at least one user is determined during a round, the algorithm must terminate within N rounds. Therefore, the running time of the algorithm is bounded by $O(N^3)$. In Section 9.5, numerical results will demonstrate that the computational complexity of Algorithm 10 is nearly a quadratic function of the number of users. In Section 9.4.4, we will propose a distributed version of Algorithm 10.

9.4.3.2 Properties of the SNE

We first show that Algorithm 10 can find an SNE.

Theorem 9.3. *For SA-PCG, Algorithm 10 computes an SNE.*

Proof: We consider three cases of each user i as follows.

Case 1: $i \in \mathcal{N}_1(\boldsymbol{a}^e)$ and $i \in \mathcal{N}_{I,k}$

Let $\boldsymbol{a}'$ be the profile right after phase I during which i remains in $\mathcal{N}_I$. Since $\boldsymbol{a}^e \geq \boldsymbol{a}'$, using (9.3) we have

$$f_i(1, \boldsymbol{a}^e_{-i}) - f_i(0, \boldsymbol{a}^e_{-i}) \geq u_i(1, \boldsymbol{a}'_{-i}) + \sum_{j \in \mathcal{N}_{i+} \setminus \overline{\mathcal{N}}} s_{ij} a'_j \geq 0$$

where the second inequality is due to the condition in line 5.

Case 2: $i \in \mathcal{N}_1(\boldsymbol{a}^e)$ and $i \in \mathcal{N}_{II,k}$

Let $\boldsymbol{a}'$ be the profile right after i becomes determined in phase II. Since $\boldsymbol{a}^e \geq \boldsymbol{a}'$, using (9.3) we have

$$f_i(1, \boldsymbol{a}^e_{-i}) - f_i(0, \boldsymbol{a}^e_{-i}) \geq u_i(1, \boldsymbol{a}'_{-i}) + \sum_{j \in \mathcal{N}_{i+} \setminus \overline{\mathcal{N}}} s_{ij} a'_j \geq 0$$

where the second inequality is due to the condition in line 10.

Case 3: $i \in \mathcal{N}_0(\boldsymbol{a}^e)$

Since i is not included in $\mathcal{N}_{II}$ in phase II of the last round, using (9.3) we have

$$f_i(1, \boldsymbol{a}^e_{-i}) - f_i(0, \boldsymbol{a}^e_{-i}) = u_i(1, \boldsymbol{a}'_{-i}) + \sum_{j \in \mathcal{N}_{i+} \setminus \overline{\mathcal{N}}} s_{ij} a'_j < 0$$

where the inequality is due to the condition in line 10. □

As the strategy profile computed by Algorithm 10 is an SNE, it is acceptable for all users. We give another desirable property as follows.

Definition 9.4 (Socially-aware Coalition-Proof)**.** *A strategy profile* $\boldsymbol{a}^{scp} = \{a_1^{scp}, \cdots, a_n^{scp}\}$ *is* socially-aware coalition-proof *if no set of users* $\mathcal{N}' \subseteq \mathcal{N}$ *can change their strategies such that no user* $i \in \mathcal{N}'$ *or its social friend* $j \in \mathcal{N}_{i+}$ *with* $s_{ij} > 0$ *achieves a*

worse individual utility while at least one of them achieves a better individual utility under the strategy profile $\boldsymbol{a}'$ *after the change, i.e.,*

$$u_i(a'_i, \boldsymbol{a}'_{-i}) \geq u_i(a_i^{scp}, \boldsymbol{a}_{-i}^{scp}),\ \forall i \in \mathcal{N}'$$

and

$$u_j(a'_j, \boldsymbol{a}'_{-j}) \geq u_j(a_j^{scp}, \boldsymbol{a}_{-j}^{scp}),\ \forall j \in \{\mathcal{N}_{i+} : s_{ij} > 0\}, i \in \mathcal{N}'$$

with at least one strict inequality.

Note that the concept of socially-aware coalition-proof for a SGUM game is in the same spirit as strong Nash equilibrium for a non-cooperative game. Also note that it is different from the core for a cooperative game [17]. In a cooperative game (also known as coalitional game), a player's payoff only depends on the strategies of the players in the same coalition, while in a SGUM game, a player's payoff depends on the strategies of all other players, and furthermore, a set of players are concerned with the payoffs of themselves and their social friends.

Next we show that the above property holds for the SNE $\boldsymbol{a}^e$.

Theorem 9.4. *For SA-PCG, the SNE* $\boldsymbol{a}^e$ *is socially-aware coalition-proof.*

Proof: Suppose a set of users $\mathcal{N}' \subseteq \mathcal{N}$ change their actions (some users in $\mathcal{N}'$ may not change their actions but at least one user in $\mathcal{N}'$ changes its action) and the strategy profile changes from $\boldsymbol{a}^e$ to $\boldsymbol{a}'$ such that $u_i(\boldsymbol{a}^e) \leq u_i(\boldsymbol{a}')$, $\forall i \in \mathcal{N}'$, and $u_j(\boldsymbol{a}^e) \leq u_j(\boldsymbol{a}')$, $\forall j \in \{\mathcal{N}_{i+} : s_{ij} > 0\}, i \in \mathcal{N}'$, with at least one strict inequality. It suffices to show that 1) $\mathcal{N}_1(\boldsymbol{a}') \setminus \mathcal{N}_1(\boldsymbol{a}^e) = \varnothing$ and 2) $\mathcal{N}_1(\boldsymbol{a}^e) \setminus \mathcal{N}_1(\boldsymbol{a}') = \varnothing$ so that $\boldsymbol{a}^e = \boldsymbol{a}'$.

We first show part 1). Suppose $\mathcal{N}_1(\boldsymbol{a}') \setminus \mathcal{N}_1(\boldsymbol{a}^e) \neq \varnothing$. We observe that $i \in \mathcal{N}'$ for each $i \in \mathcal{N}_1(\boldsymbol{a}') \setminus \mathcal{N}_1(\boldsymbol{a}^e)$. Let i be the first user in $\mathcal{N}_1(\boldsymbol{a}') \setminus \mathcal{N}_1(\boldsymbol{a}^e)$ whose action

is set to 0 during phase I of the last round, and $\bar{\boldsymbol{a}}$ be the profile right before $a_i = 0$ is set. Since $\bar{\boldsymbol{a}}_{-i} \geq \boldsymbol{a}'_{-i}$, we have

$$u_i(1, \bar{\boldsymbol{a}}_{-i}) + \sum_{j \in \mathcal{N}_{i+} \setminus \overline{\mathcal{N}}} s_{ij} \bar{a}_j \geq u_i(1, \bar{\boldsymbol{a}}_{-i}) \geq u_i(1, \boldsymbol{a}'_{-i})$$
$$= u_i(\boldsymbol{a}') \geq u_i(\boldsymbol{a}^e) = u_i(0, \boldsymbol{a}^e_{-i}) \geq 0$$

where the second inequality follows from that $\bar{\boldsymbol{a}} \geq \boldsymbol{a}'$ and the third inequality follows from that $i \in \mathcal{N}'$. This contradicts the condition in line 5.

Next we show part 2). Since we have shown part 1), we must have $\mathcal{N}_1(\boldsymbol{a}') \subseteq \mathcal{N}_1(\boldsymbol{a}^e)$ (i.e., $\boldsymbol{a}' \leq \boldsymbol{a}^e$). Suppose $\mathcal{N}_1(\boldsymbol{a}^e) \setminus \mathcal{N}_1(\boldsymbol{a}') \neq \varnothing$. According to the assumption, there exists $i \in \mathcal{N}$ such that $u_i(\boldsymbol{a}^e) < u_i(\boldsymbol{a}')$. We consider three cases of i as follows.

Case 1: $i \in \mathcal{N}_1(\boldsymbol{a}^e) \setminus \mathcal{N}_1(\boldsymbol{a}')$

We observe that $i \in \mathcal{N}'$ and $u_i(\boldsymbol{a}^e) < u_i(\boldsymbol{a}') = u_i(0, \boldsymbol{a}'_{-i}) = 0$. Let $\hat{\boldsymbol{a}}$ be the profile right before user i becomes determined. Since $\boldsymbol{a}^e_{-i} \geq \hat{\boldsymbol{a}}_{-i}$, we have $u_i(1, \hat{\boldsymbol{a}}_{-i}) \leq u_i(1, \boldsymbol{a}^e_{-i}) = u_i(\boldsymbol{a}^e) < 0$. According to the condition in line 5, we have

$$u_i(1, \hat{\boldsymbol{a}}_{-i}) + \sum_{j \in \mathcal{N}_{i+} \setminus \overline{\mathcal{N}}} s_{ij} \hat{a}_j \geq 0.$$

Therefore, there must exist some $j \in \{\mathcal{N}_{i+} : s_{ij} > 0\}$ such that $\hat{a}_j = 1$. Then we observe that $a^e_j = 1$ and j must become determined before i. We consider two cases of j as follows.

Case 1.1: $j \in \mathcal{N}_1(\boldsymbol{a}')$

In this case, we have

$$u_j(\boldsymbol{a}^e) - u_j(\boldsymbol{a}') = u_j(1, \boldsymbol{a}^e_{-i}) - u_j(1, \boldsymbol{a}'_{-i}) = \sum_{l \in \mathcal{N}_{j-}} a^e_l - \sum_{l \in \mathcal{N}_{j-}} a'_l \geq a^e_i - a'_i = 1 > 0$$

where the first inequality is due to the fact that $\boldsymbol{a}^e_{-j} \geq \boldsymbol{a}'_{-j}$ and $j \in \mathcal{N}_{i+}$. This contradicts the assumption that $u_j(\boldsymbol{a}^e) \leq u_j(\boldsymbol{a}')$.

Case 1.2: $j \in \mathcal{N}_1(\boldsymbol{a}^e) \setminus \mathcal{N}_1(\boldsymbol{a}')$

In this case, we observe that $j \in \mathcal{N}'$. Let $\tilde{\boldsymbol{a}}$ be the profile right before user j becomes determined. Since j becomes determined before i, we have $\tilde{a}_i = 0$. Then we have

$$u_j(1, \boldsymbol{a}^e_{-j}) - u_j(1, \tilde{\boldsymbol{a}}_{-j}) = \sum_{l \in \mathcal{N}_{j-}} a^e_l - \sum_{l \in \mathcal{N}_{j-}} \tilde{a}_l \geq a^e_i - \tilde{a}_i = 1 > 0$$

where the first inequality is due to the fact that $\boldsymbol{a}^e_{-j} \geq \tilde{\boldsymbol{a}}_{-j}$. Then it follows that

$$u_j(1, \tilde{\boldsymbol{a}}_{-j}) < u_j(1, \boldsymbol{a}^e_{-j}) \leq u_j(\boldsymbol{a}^e) \leq u_j(\boldsymbol{a}') = u_j(0, \boldsymbol{a}'_{-j}) = 0.$$

According to the condition in line 5, we have

$$u_j(1, \tilde{\boldsymbol{a}}_{-j}) + \sum_{l \in \mathcal{N}_{j+} \setminus \overline{\mathcal{N}}} s_{jl} \tilde{a}_l \geq 0.$$

Therefore, there must exist some $k \in \{\mathcal{N}_{j+} : s_{jk} > 0\}$ such that $\tilde{a}_k = 1$. Similar to the arguments for j, we can consider two cases of k as Case 1.1 or Case 1.2 for j. If Case 1.2 holds, we can recursively apply the arguments until Case 1.1 holds, which leads to a contradiction.

Case 2: $i \in \mathcal{N}_1(\boldsymbol{a}') \subset \mathcal{N}_1(\boldsymbol{a}^e)$

Since $\boldsymbol{a}^e \geq \boldsymbol{a}'$, we have

$$u_i(\boldsymbol{a}^e) = u_i(1, \boldsymbol{a}^e_{-i}) \geq u_i(1, \boldsymbol{a}'_{-i}) = u_i(\boldsymbol{a}').$$

This contradicts the assumption that $u_i(\boldsymbol{a}^e) < u_i(\boldsymbol{a}')$.

Case 3: $i \in \mathcal{N}_0(\boldsymbol{a}^e) \subset \mathcal{N}_0(\boldsymbol{a}')$

We have

$$u_i(\boldsymbol{a}^e) = u_i(0, \boldsymbol{a}^e_{-i}) = 0 = u_i(0, \boldsymbol{a}'_{-i}) = u_i(\boldsymbol{a}')$$

which contradicts the assumption that $u_i(\boldsymbol{a}^e) < u_i(\boldsymbol{a}')$. □

Remark: Theorem 9.4 gives a desirable property that is particularly appealing when users are socially-aware: no group of users can collectively change their strategies to improve the individual utility of at least one of them or their social friends without reducing the individual utility of any of them or their social friends. As users are aware of their social ties, some users may have incentive to form a coalition based on their social ties and deviate from the SNE in the hope of improving the individual utilities of them or their social friends. The property of socially-aware coalition-proof eliminates this possibility so that users are willing to accept the SNE as their strategies.

As a special case of Theorem 9.4, the set of all users $\mathcal{N}$ cannot improve the individual utility of at least one of them without reducing that of another by collectively changing their strategies. This leads to the following result.

Corollary 9.1. *For SA-PCG, the SNE $\boldsymbol{a}^e$ is Pareto-optimal.*

Next we show that the social welfare of the SNE $\boldsymbol{a}^e$ is no less than that of the best SNE for SO-PCG. To this end, we first show that the SNE $\boldsymbol{a}^e$ is monotonically "expanding" with respect to social ties.

Proposition 9.2. *For SA-PCG, when social ties increase (i.e., $s'_{ij} \geq s_{ij}$, $\forall i \neq j$), the set of participating users at the SNE $\boldsymbol{a}^e$ grows (i.e., $\mathcal{N}_1(\boldsymbol{a}^{e'}) \supseteq \mathcal{N}_1(\boldsymbol{a}^e)$) and the social welfare of the SNE $\boldsymbol{a}^e$ increases (i.e., $v(\boldsymbol{a}^{e'}) \geq v(\boldsymbol{a}^e)$).*

Proof: Let $\mathcal{N}'_{I,k}$ be the set of users in $\mathcal{N}_{I,k}$ during the execution that computes $\boldsymbol{a}^{e'}$. For each $i \in \mathcal{N}_{I,1}$, we have

$$u_i(1, \boldsymbol{a}'_{-i}) + \sum_{j \in \mathcal{N}'_{i+} \setminus \overline{\mathcal{N}}'} s'_{ij} a_j \geq u_i(1, \boldsymbol{a}_{-i}) + \sum_{j \in \mathcal{N}_{i+} \setminus \overline{\mathcal{N}}} s_{ij} a_j \geq 0.$$

Therefore we must have $\mathcal{N}_{I,1} \subseteq \mathcal{N}'_{I,1}$. Similarly, we can show that for any $i \in \mathcal{N}_{II,1} \setminus \mathcal{N}'_{I,1}$, we must have $i \in \mathcal{N}'_{II,1}$. Using this argument sequentially, we can show that

$\cup_{i=1}^{k}(\mathcal{N}_{I,i} \cup \mathcal{N}_{II,i}) \subseteq \cup_{i=1}^{k}\left(\mathcal{N}'_{I,i} \cup \mathcal{N}'_{II,i}\right)$ for any k, and therefore $\boldsymbol{a}^e \leq \boldsymbol{a}^{e'}$. When a user becomes determined with action 1, the increment of social welfare of determined users by changing its action from 0 to 1 is no less than the increment of its social group utility involving only determined users, which is non-negative. Therefore we can see that $v(\boldsymbol{a}^e) \leq v(\boldsymbol{a}^{e'})$. □

Remark: Intuitively, a user with larger social ties with other users is more likely to participate in favor of its social group utility, even at the cost of obtaining a negative individual utility. Proposition 9.2 confirms that when social ties become larger, more users participate at the SNE. Furthermore, as each additional participating user increases the social welfare, the social welfare of the SNE also increases.

If Algorithm 10 is used for SO-PCG (i.e., $s_{ij} = 0$, $\forall i \neq j$), we can see that it is equivalent to Algorithm 9 which is used to find the best SNE for SO-PCG, so that they compute the same strategy profile. Based on this observation, using Theorem 9.2, we have the following result.

Corollary 9.2. *The social welfare of the SNE $\boldsymbol{a}^e$ for SA-PCG is no less than that of the best SNE for SO-PCG.*

Corollary 9.2 guarantees that the social welfare of the Pareto-optimal SNE is no less than the benchmark SNE for SO-PCG. In Section 9.5, numerical results will demonstrate that the Pareto-optimal SNE is efficient, with a performance gain up to 20% over the benchmark.

9.4.3.3 Efficiency of the SNE

Next we investigate the social welfare of the SNE compared to the optimal social welfare. To this end, we first show that the set of participating users at the SNE $\boldsymbol{a}^e$ is a subset of that at the social optimal strategy profile $\boldsymbol{a}^*$.

Lemma 9.2. *For SA-PCG, the SNE $\boldsymbol{a}^e$ satisfies that $\boldsymbol{a}^e \leq \boldsymbol{a}^*$.*

Proof: Suppose $\mathcal{N}_1(\boldsymbol{a}^e) \setminus \mathcal{N}_1(\boldsymbol{a}^*) \neq \varnothing$. Without loss of generality, we assume that $\boldsymbol{a}^*$ is unique. Suppose the first participating user among $\mathcal{N}_1(\boldsymbol{a}^e) \setminus \mathcal{N}_1(\boldsymbol{a}^*)$ is in $\widehat{\mathcal{N}} \triangleq \mathcal{N}_{I,k} \setminus \mathcal{N}_1(\boldsymbol{a}^*)$. Let $\overline{\mathcal{N}}_{I,k}$ be the set of users in $\overline{\mathcal{N}}$ right before phase I of round k and $\bar{\boldsymbol{a}}$ be the profile right after that phase. Define $\boldsymbol{a}' \triangleq \boldsymbol{a}^* \vee \bar{\boldsymbol{a}}$ where $\vee$ denotes entry-wise "or" operation such that $\mathcal{N}_1(\boldsymbol{a}') = \mathcal{N}_1(\boldsymbol{a}^*) \cup \widehat{\mathcal{N}}$. Then we have

$$\begin{aligned} v(\boldsymbol{a}') - v(\boldsymbol{a}^*) &= \sum_{i \in \widehat{\mathcal{N}}} \left(u_i(1, \boldsymbol{a}'_{-i}) + \sum_{j \in \mathcal{N}_{i+} \setminus \widehat{\mathcal{N}}} a'_j \right) \\ &\geq \sum_{i \in \widehat{\mathcal{N}}} \left(u_i(1, \bar{\boldsymbol{a}}_{-i}) + \sum_{j \in \mathcal{N}_{i+} \setminus \widehat{\mathcal{N}}} \bar{a}_j \right) \\ &\geq \sum_{i \in \widehat{\mathcal{N}}} \left(u_i(1, \bar{\boldsymbol{a}}_{-i}) + \sum_{j \in \mathcal{N}_{i+} \setminus \overline{\mathcal{N}}_{I,k}} s_{ij} \bar{a}_j \right) \geq 0 \end{aligned}$$

where the first inequality follows from $\bar{\boldsymbol{a}} \leq \boldsymbol{a}'$, and the last inequality follows from $\widehat{\mathcal{N}} \subseteq \mathcal{N}_{I,k} \subseteq \overline{\mathcal{N}}_{I,k}$ and the condition in line 5. This contradicts that $\boldsymbol{a}^*$ is unique. Similarly, if the first participating user among $\mathcal{N}_1(\boldsymbol{a}^e) \setminus \mathcal{N}_1(\boldsymbol{a}^*)$ participates in phase II of some round, we can also show a contradiction. Therefore we must have $\mathcal{N}_1(\boldsymbol{a}^e) \setminus \mathcal{N}_1(\boldsymbol{a}^*) = \varnothing$. □

Let $\mathcal{N}_\Delta \triangleq \mathcal{N}_1(\boldsymbol{a}^*) \setminus \mathcal{N}_1(\boldsymbol{a}^e)$ denote the set of users that participate under the social optimal strategy profile $\boldsymbol{a}^*$ but not under the SNE $\boldsymbol{a}^e$. Using Lemma 9.2, we have the following result.

Theorem 9.5. *The performance gap between the social welfare of the Pareto-optimal SNE and the optimal social welfare is upper bounded by $\sum_{i \in \mathcal{N}_\Delta} \sum_{j \in \mathcal{N}_{i-} \cap \mathcal{N}_\Delta} 1 + \sum_{i \in \mathcal{N}_\Delta} \sum_{j \in \mathcal{N}_{i+} \setminus \mathcal{N}_\Delta} (1 - s_{ij})$. Furthermore, this bound decreases when social ties increase.*

Proof: From Lemma 9.2 we have $\boldsymbol{a}^* \geq \boldsymbol{a}^e$. Then we have

$$\begin{aligned}
v(\boldsymbol{a}^*) - v(\boldsymbol{a}^e) &= \sum_{i\in\mathcal{N}_\Delta} \left(u_i(1, \boldsymbol{a}^*_{-i}) + \sum_{j\in\mathcal{N}_{i+}\setminus\mathcal{N}_\Delta} a^*_j \right) \\
&= \sum_{i\in\mathcal{N}_\Delta} \left(u_i(1, \boldsymbol{a}^e_{-i}) + \sum_{j\in\mathcal{N}_{i-}\cap\mathcal{N}_\Delta} a^*_j + \sum_{j\in\mathcal{N}_{i+}\setminus\mathcal{N}_\Delta} a^e_j \right) \\
&\leq \sum_{i\in\mathcal{N}_\Delta} \left(u_i(1, \boldsymbol{a}^e_{-i}) + \sum_{j\in\mathcal{N}_{i+}} s_{ij} a^e_j + \sum_{j\in\mathcal{N}_{i-}\cap\mathcal{N}_\Delta} a^*_j + \sum_{j\in\mathcal{N}_{i+}\setminus\mathcal{N}_\Delta} (1 - s_{ij}) a^e_j \right) \\
&= \sum_{i\in\mathcal{N}_\Delta} \left(f_i(1, \boldsymbol{a}^e_{-i}) - f_i(0, \boldsymbol{a}^e_{-i}) + \sum_{j\in\mathcal{N}_{i-}\cap\mathcal{N}_\Delta} a^*_j + \sum_{j\in\mathcal{N}_{i+}\setminus\mathcal{N}_\Delta} (1 - s_{ij}) a^e_j \right) \\
&\leq \sum_{i\in\mathcal{N}_\Delta} \left(\sum_{j\in\mathcal{N}_{i-}\cap\mathcal{N}_\Delta} a^*_j + \sum_{j\in\mathcal{N}_{i+}\setminus\mathcal{N}_\Delta} (1 - s_{ij}) a^e_j \right) \\
&\leq \sum_{i\in\mathcal{N}_\Delta} \sum_{j\in\mathcal{N}_{i-}\cap\mathcal{N}_\Delta} 1 + \sum_{i\in\mathcal{N}_\Delta} \sum_{j\in\mathcal{N}_{i+}\setminus\mathcal{N}_\Delta} (1 - s_{ij}).
\end{aligned}$$

Next we show that the above bound decreases when social ties increase. Define $\mathcal{N}'_\Delta \triangleq \mathcal{N}_1(\boldsymbol{a}^*) \setminus \mathcal{N}_1(\boldsymbol{a}^{e'})$ where $\boldsymbol{a}^{e'}$ is the SNE when social ties increase (i.e., $s'_{ij} \geq s_{ij}$, $\forall i, j \in \mathcal{N}$). By Theorem 9.2, we have $\mathcal{N}'_\Delta \subseteq \mathcal{N}_\Delta$. Define $\hat{\mathcal{N}}_\Delta \triangleq \mathcal{N}_\Delta \setminus \mathcal{N}'_\Delta$. We observe that

$$\begin{aligned}
\sum_{i\in\mathcal{N}_\Delta} \sum_{j\in\mathcal{N}_{i-}\cap\mathcal{N}_\Delta} 1 &= \sum_{i\in\widehat{\mathcal{N}}_\Delta} \sum_{j\in\mathcal{N}_{i-}\cap\widehat{\mathcal{N}}_\Delta} 1 + \sum_{i\in\mathcal{N}'_\Delta} \sum_{j\in\mathcal{N}_{i-}\cap\mathcal{N}'_\Delta} 1 \\
&\quad + \sum_{i\in\mathcal{N}'_\Delta} \sum_{j\in\widehat{\mathcal{N}}_{i-}\cap\mathcal{N}_\Delta} 1 + \sum_{i\in\widehat{\mathcal{N}}_\Delta} \sum_{j\in\mathcal{N}_{i-}\cap\mathcal{N}'_\Delta} 1 \\
&\geq \sum_{i\in\mathcal{N}'_\Delta} \sum_{j\in\mathcal{N}_{i-}\cap\mathcal{N}'_\Delta} 1 + \sum_{i\in\widehat{\mathcal{N}}_\Delta} \sum_{j\in\mathcal{N}_{i-}\cap\mathcal{N}'_\Delta} 1 \\
&= \sum_{i\in\mathcal{N}'_\Delta} \sum_{j\in\mathcal{N}_{i-}\cap\mathcal{N}'_\Delta} 1 + \sum_{i\in\mathcal{N}'_\Delta} \sum_{j\in\mathcal{N}_{i+}\cap\widehat{\mathcal{N}}_\Delta} 1. \qquad (9.6)
\end{aligned}$$

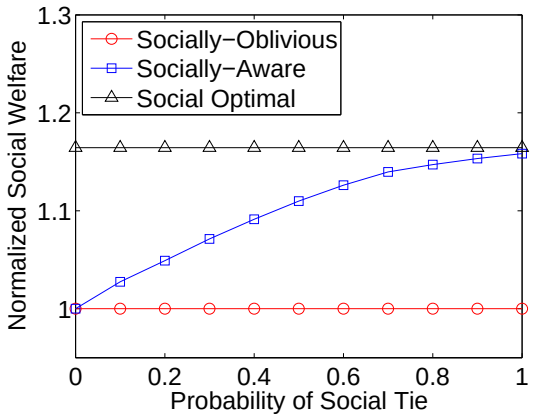


Figure 9.5: Impact of P_S for ER model based social network.

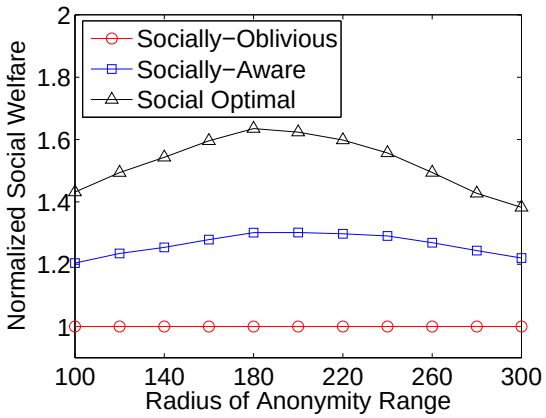


Figure 9.6: Impact of R for ER model based social network.

Then we have

$$\begin{aligned}
&\sum_{i\in\mathcal{N}_\Delta}\sum_{j\in\mathcal{N}_{i-}\cap\mathcal{N}_\Delta} 1+\sum_{i\in\mathcal{N}_\Delta}\sum_{j\in\mathcal{N}_{i+}\setminus\mathcal{N}_\Delta}(1-s_{ij})\\
&\geq \sum_{i\in\mathcal{N}'_\Delta}\sum_{j\in\mathcal{N}_{i-}\cap\mathcal{N}'_\Delta} 1+\sum_{i\in\mathcal{N}'_\Delta}\sum_{j\in\mathcal{N}_{i+}\cap\widehat{\mathcal{N}}_\Delta} 1+\sum_{i\in\mathcal{N}_\Delta}\sum_{j\in\mathcal{N}_{i+}\setminus\mathcal{N}_\Delta}(1-s_{ij})\\
&\geq \sum_{i\in\mathcal{N}'_\Delta}\sum_{j\in\mathcal{N}_{i-}\cap\mathcal{N}'_\Delta} 1+\sum_{i\in\mathcal{N}'_\Delta}\sum_{j\in\mathcal{N}_{i+}\cap\widehat{\mathcal{N}}_\Delta}(1-s_{ij})+\sum_{i\in\mathcal{N}_\Delta}\sum_{j\in\mathcal{N}_{i+}\setminus\mathcal{N}_\Delta}(1-s_{ij})\\
&= \sum_{i\in\mathcal{N}'_\Delta}\sum_{j\in\mathcal{N}_{i-}\cap\mathcal{N}'_\Delta} 1+\sum_{i\in\mathcal{N}'_\Delta}\sum_{j\in\mathcal{N}_{i+}\setminus\mathcal{N}'_\Delta}(1-s_{ij})
\end{aligned}$$

where the first inequality follows from (9.6). This shows the desired result. □

Remark: Theorem 9.5 shows that the performance gap decreases when social ties increase. The first part of the performance bound $\sum_{i\in\mathcal{N}_\Delta}\sum_{j\in\mathcal{N}_{i-}\cap\mathcal{N}_\Delta} 1$ is due to the fact that we trade the optimality of the SNE in terms of social welfare (among all SNEs) for computational tractability. As a result, when users are fully altruistic (i.e., $s_{ij}=1$, $\forall i\neq j$), the second part $\sum_{i\in\mathcal{N}_\Delta}\sum_{j\in\mathcal{N}_{i+}\setminus\mathcal{N}_\Delta}(1-s_{ij})$ becomes 0, while the first part reaches the minimum but can be greater than 0. In Section 9.5, numerical results will demonstrate that the SNE is efficient, with a performance gap less than 5% on average compared to the optimal social welfare.

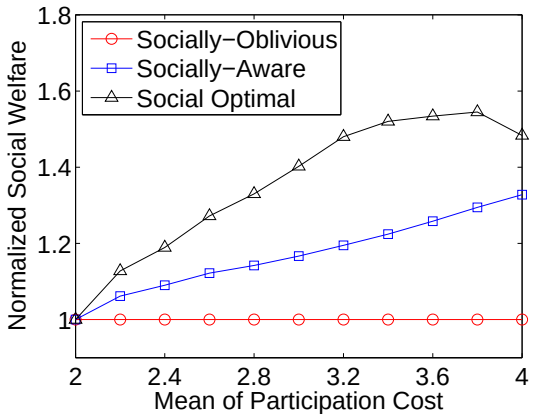


Figure 9.7: Impact of μ_C for ER model based social network.

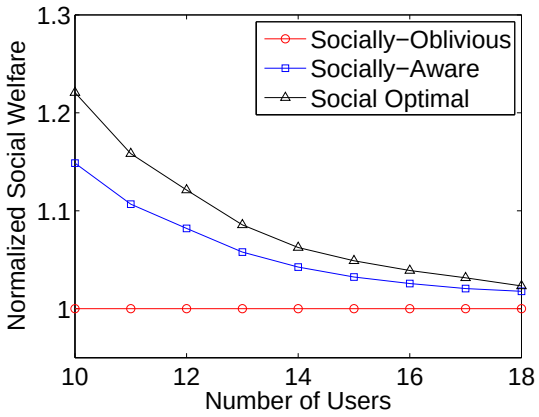


Figure 9.8: Impact of N for ER model based social network.

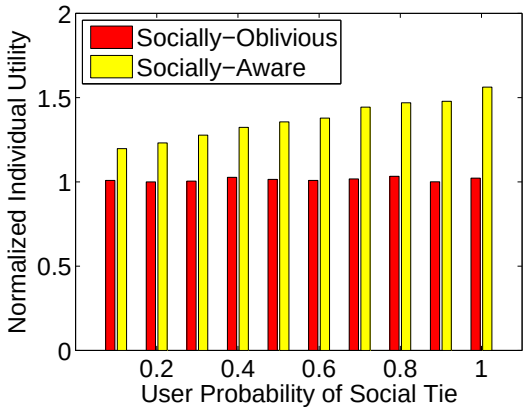


Figure 9.9: Impact of P_{S_i} for ER model based social network.

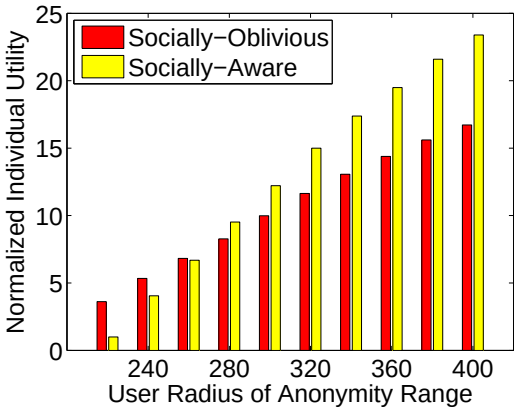


Figure 9.10: Impact of R_i for ER model based social network.

9.4.4 Distributed Computation of the SNE

The SNE computed by Algorithm 10 can be achieved in a distributed manner. To this end, each user first obtains its potential anonymity set and its social ties with others (which will be discussed in Section 9.4.5). Following Algorithm 10, each user checks if it should change its strategy according to the condition in line 5 or 10 based on other users' strategies, and if yes, announces the change to all users. With time divided into slots, a random backoff mechanism can be used so that at most one user announces a

change of strategy in a time slot. If no user announces a change, it indicates the end of phase I or phase II of a round. Therefore, all users keep track of the current state of the algorithm as it proceeds, and thus can act correctly according to the algorithm. The computational complexity of the distributed version of Algorithm 10 is almost the same as the centralized version, and is upper bounded by $O(N^3)$. Note that each user only knows the strategies of other users during the execution the algorithm, and thus users' privacy is preserved. After reaching the SNE, the users who decide to change their pseudonyms implement their pseudonym changes.

9.4.5 Implementation Issues

We assume that there is a third party platform where users interact with each other to make pseudonym change decisions and coordinate their pseudonym changes. The platform only serves to allow users to exchange information. We assume that the platform is honest-but-curious in the sense that it honestly delivers messages among users, but is curious about users' private information. To protect privacy, each user also uses a pseudonym as its identity on the platform (which can be different from that used for the LBS). To make a socially-aware pseudonym change decision, each users needs to know its potential anonymity set and its social ties with others. This can be achieved in a privacy-preserving manner using secure protocols as discussed below. Note that the platform is not involved in the computing tasks of these protocols.

A user can learn whether another user is within its anonymity range using a certain private proximity detection protocol [41, 129]. For example, the protocol proposed in [41] can be used if the anonymity range is a disk. Specifically, the protocol involves several message exchanges between the two users, including one message that contains encrypted values that are functions of a user' location or the radius of the anonymity range. The protocol guarantees that both users can only learn the binary result of whether or not one is in another's anonymity range, and neither user

can learn the other's location or anonymity range. In addition, since location information is encrypted in the messages, the platform cannot learn any user's location information. Similarly, the protocol in [129] can be used if the anonymity range is a convex polygon. Therefore, each user can learn its potential anonymity set without revealing its location information.

A user can also learn its social tie with another user without disclosing one's real identity to the other. To this end, each user keeps a social profile consisting of the social communities that it belongs to (e.g., a community of colleagues at the same workplace), and sets a *single* social tie level for each community based on its social relationships with those in the community. Each community is identified by a predefined key that is only known to the community's members. Using a certain private matching protocol such as [130, 131], two users can learn whether they have a community in common, and if yes, which community[6] it is. In particular, the protocol involves several message exchanges between the two users, including one message that contains encrypted values that are functions of the keys of a user's social communities. The protocol ensures that both users can only know the community they have in common (if it exists), and neither user can learn any additional social information of the other, or pretend to have a community in common with the other. Since a community typically has many members, neither user can know the other's real identity even when they know the community they both belong to. In addition, since social information is encrypted in the messages, the platform cannot learn any user's social information. Therefore, each user can learn its social ties with those in its potential anonymity set while keeping their real identities private. Note that although the adversary might collude with multiple users, it is almost infeasible for the adversary to find a sufficient number of colluding users who have social ties with

[6]To protect privacy, only one community in common is revealed if they have multiple communities in common.

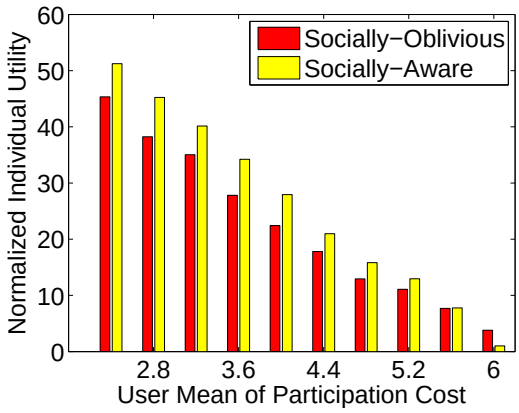


Figure 9.11: Impact of μ_{C_i} for ER model based social network.

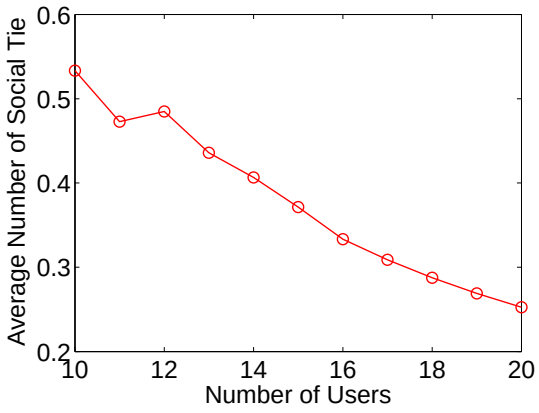


Figure 9.12: Average number of social ties per user for real dataset based social network.

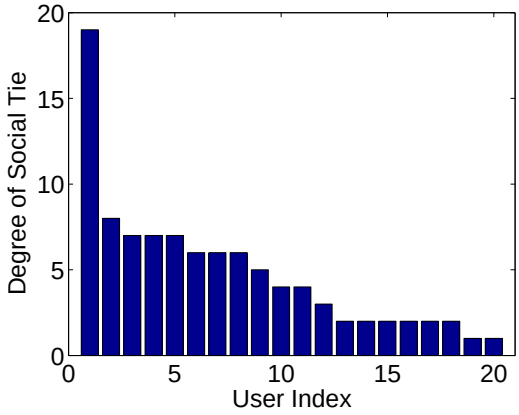


Figure 9.13: Degree of social tie for real dataset based social network.

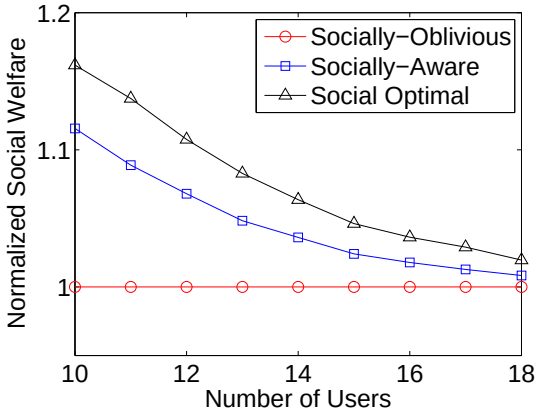


Figure 9.14: Impact of N for real dataset based social network.

a specific user, in order to infer the user's real identity.

9.5 Simulation Results

In this section, we provide numerical results to evaluate the performance of the SNE $\boldsymbol{a}^e$ for SA-PCG computed by Algorithm 10. We compare the SNE with two benchmarks: the best SNE $\boldsymbol{a}^o$ for SO-PCG which is computed by Algorithm 9, and the social optimal strategy profile $\boldsymbol{a}^*$, which is the optimal solution of problem (9.5) and is found by exhaustive search.

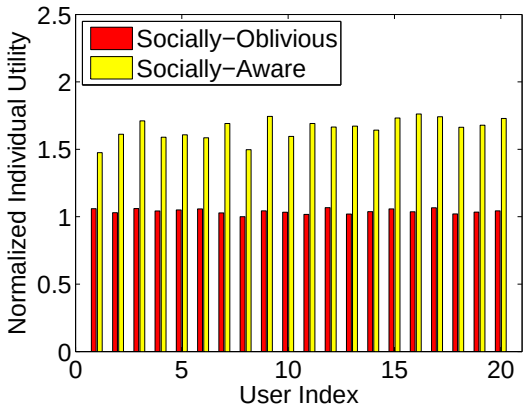


Figure 9.15: Impact of user degree for real dataset based social network.

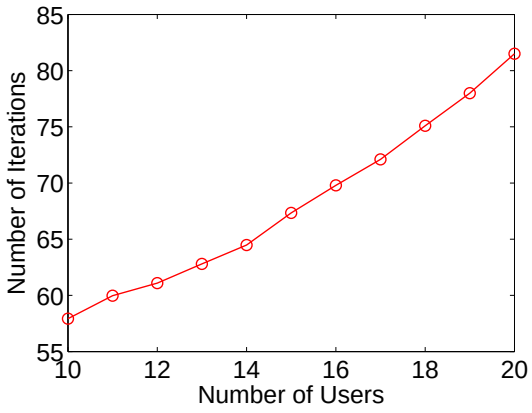


Figure 9.16: Computational complexity for real dataset based social network.

9.5.1 Simulation Setup

To illustrate the impact of different parameters of the mobile social network on the performance, we consider a synthetic social network as follows. We simulate the social graph G^S based on the Erdos-Renyi (ER) graph model [31], where a social tie exists between any pair of users with probability P_S. We assume that the social tie is 1 if it exists. We set $N = 10$, $P_S = 0.5$ as default values.

To evaluate the performance of the socially-aware pseudonym change in practice, we also consider an empirical social network. Specifically, we generate the social graph according to the real dataset from Brightkite [32], which is an online social networking service based on mobile phones. To illustrate the social network structure of this dataset, we plot the average number of social tie between a pair of users (in analogy to the probability of social tie in the ER model) versus the number of users in Fig. 9.12, and the users' degrees in the social network in Fig. 9.13.

We simulate the physical graph based on a practical setting as follows. We consider N mobile users randomly located in a square area with side length 500 m. We assume that the anonymity range of each user is a disk centered at the user's

location with radius R. Based on users' physical locations and anonymity ranges, there exists an edge from user i to user j in the physical graph if user i is in the anonymity range of user j. We assume that each user's participation cost follows a normal distribution with mean μ_C and variance σ_C^2. We set $\mu_C = 3$, $\sigma_C^2 = 1$ as default values.

9.5.2 Simulation Results

9.5.2.1 System Efficiency

We first evaluate the system efficiency of the SNE for SA-PCG. To highlight the performance comparison, we normalize the results of the SNE for SA-PCG and the social optimal solution with respect to the SNE for SO-PCG. We illustrate the impact of P_S, R, μ_C, and N on the social welfare for the ER model based social network in Figs. 9.5, 9.6, 9.7, and 9.8, respectively. We observe from Figs. 9.5-9.8 that socially-aware users significantly outperform socially-oblivious users, especially when P_S or μ_C is large, or N is small, or R is moderate. This is because more users participate when they are socially-aware, which improves the social welfare. Fig. 9.5 shows that the performance of socially-aware users improves when P_S increases, with a performance gain up to 16% over socially-oblivious users and a performance gap less than 10% on average from the optimal social welfare. This is because larger social ties encourage more users to participate. Figs. 9.7 and 9.8 show that the performance gap from socially-oblivious users to socially-aware users and the social optimal solution is small when μ_C is small or N is large. This is because with a small participation cost, or a large privacy gain of participation due to a large number of users, many users already participate even when they are socially-oblivious. Therefore, the performance gap is small as it depends on the users that participate only when they are socially-aware.

We illustrate the impact of N on the social welfare for the real dataset based social network in Fig. 9.14. We observe that the performance gain of socially-aware

users can achieve up to 11% over socially-aware users, and its performance gap from the optimal social welfare is less than 5% on average. This demonstrates the effectiveness of exploiting social ties for improving location privacy in practice.

9.5.2.2 Individual Performance

Next we evaluate individual users' performance at the SNE for SA-PCG. To demonstrate the impact of a particular parameter, we vary this parameter for different users and set other parameters the same for all users. We also normalize all the results to highlight the performance comparison. We illustrate the impact of P_{S_i} on users' individual utilities in Fig. 9.9 for 10 users, where we vary a user's probability P_{S_i} of having a social tie from another user from 0.1 to 1 with an increment of 0.1. We observe that users' individual utilities are almost the same when they are socially-oblivious, while a user with a higher probability P_{S_i} has a larger individual utility when users are socially-aware. This shows that socially-aware pseudonym change can achieve *service differentiation*, which is a desirable property for fairness: if a user has more or larger social ties from others than other users have, then that user can also achieve better individual utility than others. We also observe that each user has a larger individual utility in the socially-aware case than in the socially-oblivious case. This shows that each user can benefit from socially-aware behaviors of others in the long run on average.

Fig. 9.10 illustrates users' individual utilities when we vary a user's radius of anonymity range R_i from 220 to 400 with an increment of 20. We observe that for both socially-oblivious and socially-aware users, respectively, a user with a larger anonymity range has a larger individual utility as expected. We also observe that a user with a small anonymity range (e.g., with the radius of 220 or 240) can have a smaller individual utility in the socially-aware case than in the socially-oblivious case. This is because the user incurs a larger loss of its individual utility for its social group

utility than the gain it obtains from its social friends. Similarly, a user with a large anonymity range can have a larger individual utility when users are socially-aware.

Fig. 9.11 illustrates users' individual utilities when we vary a user's mean of participation cost μ_{C_i} from 2.4 to 6 with an increment of 0.4. We observe that for both socially-oblivious and socially-aware users, respectively, a user with a larger participation cost has a smaller individual utility as expected. We can also make a similar observation as in Fig. 9.10 due to similar reasons: a user with a large participation cost (e.g., with the mean of 6) can have a smaller individual utility in the socially-aware case than in the socially-oblivious case.

Fig. 9.15 illustrates users' individual utilities for the real dataset based social network. We observe that users' individual utilities are not significantly different from each other for socially-aware users as in Fig. 9.9, although users' degrees of social tie are significantly different as illustrated in Fig. 9.13. This is because of the fact that users' social ties are *symmetric* in the real dataset based social network. As a result, a user with a large degree of social tie (e.g., with the degree of 19) is likely to not only receive help from its social friends but also give help to them, such that these two effects largely cancel each other.

9.5.2.3 Computational Complexity

We evaluate the computational complexity of Algorithm 10 for finding the SNE for SA-PCG. We plot the number of iterations for running Algorithm 10 versus N in Fig. 9.16. We observe that the computational complexity increases nearly quadratically as the number of users increases. This shows that the algorithm is scalable for a large number of users.

9.6 Conclusion

In this chapter, we have studied a socially-aware pseudonym change game for personalized location privacy, based on a general anonymity model with user-specific anonymity sets. The game is based on a social group utility maximization framework that captures mobile users' diverse social ties and diverse physical relationships. For the SGUM-based PCG, we show that there exists a socially-aware Nash equilibrium, and quantify the performance gap of the SNE with respect to the optimal social welfare. Then we develop a greedy algorithm that can efficiently find an SNE that is socially-aware coalition-proof and Pareto-optimal. The SNE also achieves a social welfare larger than the social welfare of any SNE for the socially-oblivious PCG. We further quantify the performance gap of the SNE with respect to the optimal social welfare is bounded above. Numerical results demonstrate that social welfare can be significantly improved by exploiting users' social ties.

Chapter 10

CONCLUSION AND FUTURE WORK

In this dissertation, we have studied design and optimization of wireless networks with the focus on two perspectives: 1) socially-aware mobile networking and computing; 2) security and privacy in wireless networking. The dissertation can be broadly organized into three parts under this common theme. The first part (Chapters 2-5) studies socially-aware mobile networking and computing with the focus on random access control, power control, mobile crowdsensing, wireless data pricing, and distributed opportunistic scheduling. The second part (Chapters 6-7) studies physical security in wireless networking with the focus on radar sensor network coverage. The third part (Chapters 8-9) studies cyber security and privacy in socially-aware networking and computing with the focus on random access control, cooperative jamming, spectrum access, and location privacy. In the following, we summarize the main contributions and discuss possible directions for future work.

Chapter 2 studies mobile users' altruistic behaviors based on their social ties. We develop a social group utility maximization (SGUM) framework which captures diverse social ties among mobile users and diverse physical relationships among their mobile devices in a unified way. Specifically, instead of maximizing one's individual utility, each user aims to maximize its social group utility based on its social ties with other users. A primary merit of this framework is that it spans the continuum between non-cooperative game (NCG) and network utility maximization (NUM) - two traditionally disjoint paradigms for network optimization. Under the SGUM framework, we study two important applications in wireless networks: random access control and power control. We first derive socially-aware Nash equilibria (SNEs) for the SGUM based random access control and power control, respectively. Then we

show that as social ties increase, each user's SNE strategy migrates monotonically from the Nash equilibrium strategy for a standard NCG to the social optimal strategy for NUM, and the social welfare of the SNE also increases. This shows that the SGUM framework captures NCG and NUM as two special cases and spans the continuum in between. An interesting direction for future research is to consider more general utility functions for random access control and power control, and other wireless networking applications under the SGUM framework.

Chapter 3 studies a socially-aware crowdsensing system in which a cloud-based platform incentivizes mobile users to participate in sensing tasks by leveraging their social trust. For this system, we exploit social trust assisted reciprocity (STAR) - a synergistic marriage of social trust and reciprocity, to design an incentive mechanism that stimulates users' participation. Given the social trust structure among users, we thoroughly investigate the efficacy of STAR for satisfying users' sensing requests. Specifically, we first show that all requests can be satisfied if and only if sufficient social credit can be "transferred" from users who request more sensing service than what they can provide to users who can provide more than what they request. Then we investigate utility maximization for sensing services, and show that it boils down to maximizing the utility of a circulation flow in the combined social graph and request graph. Accordingly, we develop an algorithm that iteratively cancels a cycle of positive weight in the residual graph, which finds the optimal solution efficiently, for both cases of divisible and indivisible sensing service. Extensive simulation results corroborate that STAR can significantly outperform the mechanisms using social trust only or reciprocity only. For future work, it is interesting to study a game-theoretic setting where each user makes a distributed decision of how much social credit and sensing service to provide (or receive) to (or from, respectively) other users.

Chapter 4 studies mobile users' data usage behaviors by jointly considering the

network effect based on their social relationships in the social domain and the congestion effect in the physical wireless domain. Accordingly, we develop a Stackelberg game for problem formulation: In Stage I, a wireless provider first decides the data pricing to all users to maximize its revenue, and then in Stage II, users observe the price and decide data usage subject to mutual interactions under both network and congestion effects. We analyze the two-stage game using backward induction. For Stage II, we first show the existence and uniqueness of a user demand equilibrium (UDE). Then we propose a distributed update algorithm for users to reach the UDE. We further investigate the impact of different parameters on the UDE. For Stage I, we develop an optimal pricing algorithm to maximize the wireless provider's revenue. We evaluate the performances of the developed algorithm by simulation results using real data. For future work, we can examine other utility functions, e.g., a logarithmic function for internal utility, yet the major engineering insights should remain the same. Another interesting direction is to study the provider's pricing strategy when it is allowed to differentiate the price for different users. In this case, the price offered to each user will depend on its social influences to others based on the social network.

Chapter 5 studies opportunistic cooperative networking (OCN) in wireless ad hoc networks, with a focus on characterizing the desired tradeoff between the probing cost for establishing cooperative relaying and hence higher throughput via opportunistic cooperative networking. Specifically, we treat opportunistic cooperative networking as an optimal stopping problem with two-levels of incomplete information. We consider the cases with or without dedicated relays, and we establish the existence of optimal strategies for both cases. Then we show that for the case with dedicated relays, the optimal strategy exhibits a threshold structure, in which it is optimal to probe the dedicated relay when the signal-to-noise ratio (SNR) of the source-relay link exceeds some threshold. For the case without dedicated relays, under more restrictive

conditions, the optimal strategy is also threshold-based, in the sense that it is optimal to probe potential relays when the SNR of the source-destination link lies between two thresholds. Furthermore, these strategies can be implemented in a distributed manner. In this study, we have considered a fixed number of potential relays from which the best one is selected. It is of great interest to consider cooperative networking in the scenario where the number of potential relays is time-varying and unknown. In this scenario, a dynamic relay selection approach is required.

Chapter 6 studies the coverage problem of a bistatic radar (BR) sensor network, which is very challenging due to the Cassini oval sensing region of a BR and the coupling of sensing regions across different BRs. In particular, we consider the problem of deploying a network of BRs in a region to maximize the worst-case intrusion detectability, which amounts to minimizing the vulnerability of a barrier. We show that it is optimal to place BRs on the shortest barrier if it is the shortest line segment that connects the left and right boundary of the region. Based on this, we study the optimal placement of BRs on a line segment to minimize its vulnerability, which is a non-convex optimization problem. By exploiting certain specific structural properties pertaining to the problem (particularly an important structure of detectability), we characterize the optimal placement order and the optimal placement spacing of the BR nodes, both of which present elegant balanced structures. The findings provide valuable insights into the placement of BRs for barrier coverage. While in this study we assume that all BRs are homogeneous, a possible future direction is to consider heterogeneous BRs. It is also of interest to take into account the synchronization issue among BR transmitters and BR receivers, which is quite different from that of monostatic radars.

Chapter 7 studies radar sensor networks where the Doppler effect is exploited to combat the effect of clutter on radar detection. We introduce the concept of

Doppler coverage as a coverage metric for such radar sensor networks. Specifically, a target is said to be Doppler-covered if, regardless of its direction of motion, there exists some radar in the network whose signal-to-noise ratio (SNR) is sufficiently high and the DFS at that radar is sufficiently large. Based on the Doppler coverage model, we devise an efficient method to characterize Doppler-covered regions for arbitrarily deployed radars. Then we design an algorithm for deriving the minimum radar density required to achieve Doppler coverage in a region under any polygonal deployment pattern, and further apply it to investigate the regular triangle based deployment. An interesting direction for future work is to study the sensor density required for a region to be Doppler-covered under random deployment.

Chapter 8 studies an extended social group utility maximization (SGUM) framework that takes into account both "positive" and "negative" social ties. As a result, this extended SGUM framework captures the rich continuum from zero-sum game (ZSG) to non-cooperative game (NCG) to network utility maximization (NUM) - traditionally disjoint paradigms for network optimization. Under this SGUM framework, we study random access control, cooperative jamming, and spectrum access as three applications. For the SGUM based random access control, we derive the socially-aware Nash equilibrium. We show that as social ties increase, each user's SNE strategy migrates monotonically from the NE strategy for a ZSG to that for a standard NCG, and then to the social optimal strategy for NUM, and the social welfare of the SNE also increases. For the SGUM based multi-channel cooperative jamming, we show that there exists a unique mixed strategy SNE. When the social tie between the legitimate user and the cooperative jammer exceeds a certain threshold, the cooperative jammer always jams the eavesdropper on some channel at the SNE, which improves the social welfare of the legitimate user and cooperative jammer. Then we consider the SGUM based spectrum access for social networks with

structural balance. We characterize the SNE and show that it is equivalent to the SNE for the game where each player consists of the users in the same social coalition based on the balanced social structure. For future work, it is interesting to consider more general utility functions for random access control and cooperative jamming, and other wireless networking applications under the extended SGUM framework.

Chapter 9 studies mobile users' pseudonym change to protect location privacy where users' social ties are leveraged to motivate them to participate. Drawing on the social group utility maximization (SGUM) framework developed in Chapter 2, we cast users' decision making of whether to change pseudonyms as a socially-aware pseudonym change game (PCG). The PCG further assumes a general anonymity model that allows a user to have its specific anonymity set for personalized location privacy. For the SGUMbased PCG, we show that there exists a socially-aware Nash equilibrium (SNE), and we quantify the system efficiency of SNEs with respect to the optimal social welfare. Then we develop a greedy algorithm that myopically determines users' strategies, based on the social group utility derived from only the users whose strategies have already been determined. We show that this algorithm can efficiently find an SNE that enjoys desirable properties: 1) it is socially-aware coalition-proof, and thus is also Pareto-optimal; 2) it achieves a larger social welfare than any SNE for the socially-oblivious PCG. We further quantify the system efficiency of the SNE with respect to the optimal social welfare. We also show that the SNE can be achieved in a distributed manner. Numerical results corroborate that social welfare can be significantly improved by exploiting social ties. An interesting direction for future research is to study users' behaviors in pseudonym change under the extended SGUM framework that captures both "positive" and "negative" social ties.

REFERENCES

[1] "eMarketer: Social networking reaches nearly one in four around the world." [Online]. Available: http://www.emarketer.com/Article/Social-Networking-Reaches-Nearly-One-Four-Around-World/1009976

[2] "Gartner: Worldwide PC, tablet and mobile phone shipments to grow 4.5 percent in 2013." [Online]. Available: http://www.gartner.com/newsroom/id/2610015

[3] N. Kayastha, D. Niyato, P. Wang, and E. Hossain, "Applications, architectures, and protocol design issues for mobile social networks: A survey," *Proceedings of the IEEE*, vol. 99, no. 12, pp. 2130–2158, 2011.

[4] "Cisco white paper: Global mobile data traffic data forecast update 2013-2018."

[5] X. Qin and R. Berry, "Exploiting multiuser diversity for medium access control in wireless networks," in *Proc. IEEE INFOCOM*, San Francisco, CA, Apr. 2003.

[6] P. Viswanath, D. N. C. Tse, and R. Laroia, "Opportunistic beamforming using dumb antennas," *IEEE Trans. Information Theory*, vol. 48, pp. 1277–1294, Jun. 2002.

[7] X. Liu, E. K. Chong, and N. B. Shroff, "A framework for opportunistic scheduling in wireless networks," *Computer Networks*, vol. 41, pp. 451–474, Mar. 2003.

[8] D. Zheng, W. Ge, and J. Zhang, "Distributed opportunistic scheduling for ad hoc networks with random access: An optimal stopping approach," *IEEE Trans. Information Theory*, vol. 55, pp. 205–222, Jan. 2009.

[9] B. Wang, "Coverage problems in sensor networks: A survey," *ACM Computing Surveys*, vol. 43, Oct. 2011.

[10] E. Altman, T. Boulogne, R. El-Azouzi, T. Jimenez, and L. Wynter, "A survey on networking games in telecommunications," *Computer Operation Research*, vol. 33, pp. 286–311, 2006.

[11] D. P. Palomar and M. Chiang, "A tutorial on decomposition methods for network utility maximization," *IEEE Journal on Selected Areas in Communications*, vol. 24, pp. 1439–1451, Aug. 2006.

[12] A. Chen, T.-H. Lai, and D. Xuan, "Measuring and guaranteeing quality of barrier-coverage in wireless sensor networks," in *ACM MOBIHOC 2008*.

[13] M. Hoefer and A. Skopalik, "Altruism in atomic congestion games," in *European Symposium on Algorithms 2009*.

[14] G. Kesidis, Y. Jin, A. Azad, and E. Altman, "Stable nash equilibria of ALOHA medium access games under symmetric, socially altruistic behavior," in *IEEE CDC 2010*.

[15] W. Saad, Z. Han, M. Debbah, A. Hjorungnes, and T. Basar, "Coalitional game theory for communication networks," *IEEE Signal Processing Magazine*, vol. 26, pp. 77–97, Sept. 2009.

[16] X. Chen, B. Proulx, X. Gong, and J. Zhang, "Social trust and social reciprocity based cooperative D2D communications," in *Proc. ACM MOBIHOC 2013*.

[17] M. Osborne and A. Rubinstein, *A course in game theory*. MIT Press, 1994.

[18] G. Scutari, S. Barbarossa, and D. P. Palomar, "Potential games: A framework for vector power control problems with coupled constraints," in *IEEE ICASSP 2006*.

[19] U. O. Candogan, I. Menache, A. Ozdaglar, and P. A. Parrilo, "Near-optimal power control in wireless networks: a potential game approach," in *IEEE INFOCOM 2010*.

[20] D. M. Topkis, *Supermodularity and Complementarity*. Princeton University Press, 1998.

[21] E. Altman and Z. Altman, "S-modular games and power control in wireless networks," *IEEE Communications Surveys & Tutorials*, vol. 48, p. 839842, May 2003.

[22] D. Yang, G. Xue, X. Fang, and J. Tang, "Crowdsourcing to smartphones: incentive mechanism design for mobile phone sensing," in *Proc. ACM MOBICOM 2012*.

[23] T. Luo and C.-K. Tham, "Fairness and social welfare in incentivizing participatory sensing," in *Proc. IEEE SECON 2012*.

[24] I. Koutsopoulos, "Optimal incentive-driven design of participatory sensing systems," in *Proc. IEEE INFOCOM 2013*.

[25] "Waze: outsmarting traffic, together." [Online]. Available: https://www.waze.com/

[26] A. Sahai, S. M. Mishra, R. Tandra, and K. A. Woyach, "Cognitive radios for spectrum sharing," *IEEE Signal Processing Magazine*, vol. 26, pp. 140–145, Jan. 2009.

[27] Z. Liu, H. Hu, Y. Liu, K. W. Ross, Y. Wang, and M. Mobius, "P2P trading in social networks: The value of staying connected," in *Proc. IEEE INFOCOM 2010*.

[28] R. Ahuja, T. Magnanti, and J. Orlin, *Network flows: Theory, algorithms, and applications*. Prentice Hall, 1993.

[29] X. Huang, "Negative-weight cycle algorithms," in *Proc. 2006 International Conference on Foundations of Computer Science*.

[30] D. Easley and J. Kleinberg, *Networks, crowds, and markets: Reasoning about a highly connected world*. Cambridge University Press, 2010.

[31] P. Erdos and A. Renyi, "On the evolution of random graphs," *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, pp. 17–61, 1960.

[32] "SNAP: Network datasets: Brightkite." [Online]. Available: http://snap.stanford.edu/data/loc-brightkite.html

[33] F. Ernst and S. Gächter, "Fairness and retaliation: The economics of reciprocity," *Journal of Economic Perspectives*, vol. 14, no. 3, pp. 159–181, 2000.

[34] B. Cohen, "Incentives build robustness in bittorrent," in *Workshop on Economics of Peer-to-Peer Systems, 2003*.

[35] V. Srinivasan, P. Nuggehalli, C.-F. Chiasserini, and R. R. Rao, "Cooperation in wireless ad hoc networks," in *Proc. IEEE INFOCOM 2003*.

[36] W. H. Yuen, R. D. Yates, and S.-C. Mau, "Exploiting data diversity and multiuser diversity in non-cooperative mobile infostation networks," in *Proc. IEEE INFOCOM 2003*.

[37] P. Ganesan and M. Seshadri, "On cooperative content distribution and the price of barter," in *Proc. IEEE ICDCS 2005*.

[38] U. Shevade, H. H. Song, L. Qiu, and Y. Zhang, "Incentive-aware routing in DTNs," in *Proc. IEEE ICNP 2008*.

[39] S. Zhong, J. Chen, and Y. R. Yang, "Incentive-aware routing in DTNs," in *Proc. IEEE INFOCOM 2003*.

[40] W. Wang, S. Eidenbenz, Y. Wang, and X.-Y. Li, "OURS: optimal unicast routing systems in non-cooperative wireless networks," in *Proc. ACM MOBICOM 2006*.

[41] S. Zhong and F. Wu, "On designing collusion-resistant routing schemes for non-cooperative wireless ad hoc networks," in *Proc. ACM MOBICOM 2007*.

[42] C. Zhang, X. Zhu, Y. Song, and Y. Fang, "C4: A new paradigm for providing incentives in multi-hop wireless networks," in *Proc. IEEE INFOCOM 2011*.

[43] "Bitcoin: An open source P2P digital currency." [Online]. Available: http://bitcoin.org/en/

[44] S. Marti, T. J. Giuli, K. Lai, and M. Baker, "Mitigating routing misbehavior in mobile ad hoc networks," in *Proc. ACM MOBICOM 2000*.

[45] S. Buchegger and J.-Y. L. Boudec, "Performance analysis of the CONFIDANT protocol," in *Proc. ACM MOBIHOC 2002*.

[46] J. Jaramillo and R. Srikant, "DARWIN: Distributed and adaptive reputation mechanism for wireless networks," in *Proc. ACM MOBICOM 2007*.

[47] D. B. DeFigueiredo and E. T. Barr, "Trustdavis: A non-exploitable online reputation system," in *Proc. IEEE E-Commerce Technology 2005*.

[48] P. Dandekar, A. Goel, R. Govindan, and I. Post, "Liquidity in credit networks: A little trust goes a long way," in *Proc. ACM EC 2011*.

[49] P. Dandekar, A. Goel, M. P. Wellman, and B. Wiedenbeck, "Strategic formation of credit networks," in *Proc. ACM WWW 2012*.

[50] A. Mislove, A. Post, P. Druschel, and K. P. Gummadi, "Ostra: Leveraging trust to thwart unwanted communication," in *Usenix NSDI 2008*.

[51] X. Chen, X. Gong, L. Yang, and J. Zhang, "A social group utility maximization framework with applications in database assisted spectrum access," in *Proc. IEEE INFOCOM 2014*.

[52] X. Gong, J. Zhang, and D. Cochran, "When target motion matters: Doppler coverage in radar sensor networks," in *IEEE INFOCOM 2013*.

[53] X. Gong, X. Chen, and J. Zhang, "Social group utility maximization in mobile networks: From altruistic to malicious behavior," in *Proc. IEEE Conference on Information Sciences and Systems 2014*.

[54] "WeChat: The new way to connect." [Online]. Available: http://www.wechat.com/

[55] "WhatsApp: Simple. Personal. Real time messaging." [Online]. Available: http://www.whatsapp.com/

[56] "Global Digital Statistics." [Online]. Available: http://wearesocial.net/tag/sdmw/

[57] B. Briscoe, A. Odlyzko, and B. Tilly, "Metcalfe's law is wrong-communications networks increase in value as they add members-but by how much?" *IEEE Spectrum*, vol. 43, no. 7, pp. 34–39, 2006.

[58] O. Candogan, K. Bimpikis, and A. Ozdaglar, "Optimal pricing in networks with externalities," *INFORMS Operation Research*, vol. 60, no. 4, pp. 883–905, 2012.

[59] D. Fudenberg and J. Tirole, *Game Theory*. MIT Press, 1991.

[60] J. B. Rosen, "Existence and uniqueness of equilibrium points for concave N-person games," *Econometrica*, vol. 33, no. 3, pp. 520–534, 1965.

[61] "Personalized location privacy in mobile networks: A social group utility approach," Technical Report. [Online]. Available: http://informationnet.asu.edu/pub/SGUM-infocom15-TR.pdf

[62] J. Nair, A. Wierman, and B. Zwart, "Exploiting network effects in the provisioning of large scale systems," in *Proc. of IFIP Performance 2011*.

[63] W. Wu, T. B. Ma, and C. S. Lui, "Exploring bundling sale strategy in online service markets with network effects," in *Proc. of IEEE INFOCOM 2014*.

[64] R. Johari and S. Kumar, "Congestible services and network effects," in *Proc. ACM Conference on Electronic Commerce 2010*.

[65] P. Costa, C. Mascolo, M. Musolesi, and G. P. Picco, "Socially-aware routing for publish-subscribe in delay-tolerant mobile ad hoc networks," *IEEE J. Sel. Area. Comm.*, vol. 26, no. 5, pp. 748–760, June 2008.

[66] W. Gao, Q. Li, B. Zhao, and G. Cao, "Multicasting in delay tolerant networks: A social network perspective," in *ACM MOBIHOC 2009*.

[67] X. Gong, X. Chen, K. Xing, D. Shin, M. Zhang, and J. Zhang, "Personalized location privacy in mobile networks: A social group utility approach," in *IEEE INFOCOM 2015*.

[68] B. Debreu, "A social equilibrium existence theorem," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 38, no. 10, pp. 886–893, 1952.

[69] I. Glicksberg, "A further generalization of the Kakutani fixed point theorem, with application to Nash equilibrium," *Proceedings of the American Mathematical Society*, vol. 3, no. 1, pp. 170–174, 1952.

[70] K. Fan, "Fixed-point and minimax theorems in locally convex topological linear spaces," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 38, no. 2, pp. 121–126, 1952.

[71] R. Horn and C. Johnson, *Matrix Analysis*. Cambridge University Press, 1985.

[72] T. Cover and A. E. Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Information Theory*, vol. 25, pp. 572–584, Sept. 1979.

[73] A. Host-Madsen and J. Zhang, "Capacity bounds and power allocation for wireless relay channels," *IEEE Trans. Information Theory*, vol. 51, pp. 2020–2040, Jun. 2005.

[74] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: efficient protocols and outage behavior," *IEEE Trans. Information Theory*, vol. 50, pp. 3062–3080, Dec. 2004.

[75] A. K. Sadek, K. J. R. Liu, and A. Ephremides, "Cognitive multiple access via cooperation: Protocol design and performance analysis," *IEEE Trans. Information Theory*, vol. 53, pp. 3677–3696, Oct. 2007.

[76] B. Rong and A. Ephremides, "Protocol-level cooperation in wireless networks: stable throughput and delay analysis," in *Proc. the 7th International Conference on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks*, Seoul, Korea, 2009, pp. 151–160.

[77] ——, "Cooperation above the physical layer: the case of a simple network," in *Proc. IEEE ISIT*, Seoul, Korea, 2009, pp. 1789–1793.

[78] W. Stadje, "An optimal stopping problem with two levels of incomplete information," *Mathematical Methods of Operations Research*, vol. 45, pp. 119–131, 1997.

[79] C. Thejaswi P. S., J. Zhang, M.-O. Pun, H. V. Poor, and D. Zheng, "Distributed opportunistic scheduling with two-level probing," *IEEE/ACM Trans. Networking*, vol. 18, pp. 1464–1477, Oct. 2010.

[80] Q. Zhao and L. Tong, "Opportunistic carrier sensing for energy-efficient information retrieval in sensor networks," *EURASIP J. Wireless Communications and Networking*, pp. 231–241, 2005.

[81] T. Ferguson, *Optimal Stopping and Applications*. available at http://www.math.ucla.edu/∼tom/Stopping/Contents.html, 2006.

[82] C. J. Baker and H. D. Griffiths, "Bistatic and multistatic radar sensors for homeland security," *Advances in Sensing with Security Applications*, vol. 2, pp. 1–22, Feb. 2006.

[83] B. Donovan, D. J. McLaughlin, and J. Kurose, "Principles and design considerations for short-range energy balanced radar networks," in *IGARSS 2005*.

[84] S. Kumar, T.-H. Lai, and A. Arora, "Barrier coverage with wireless sensors," in *ACM MOBICOM 2005*.

[85] B. Liu, O. Dousse, J. Wang, and A. Saipulla, "Strong barrier coverage of wireless sensor networks," in *ACM MOBIHOC 2008*.

[86] A. Saipulla, B. Liu, G. Xing, X. Fu, and J. Wang, "Barrier coverage with sensors of limited mobility," in *ACM MOBIHOC 2010*.

[87] N. Willis, *Bistatic Radar*. SciTech Publishing, 2005.

[88] J. Liang and Q. Liang, "Orthogonal waveform design and performance analysis in radar sensor networks," in *IEEE MILCOM 2006*.

[89] H.-D. Ly and Q. Liang, "Collaborative multi-target detection in radar sensor networks," in *IEEE MILCOM 2007*.

[90] J. Liang and Q. Liang, "Design and analysis of distributed radar sensor networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 27, pp. 1926–1933, Nov. 2011.

[91] H. Deng, "Orthogonal netted radar systems," *IEEE Aerospace and Electronic Systems Magazine*, vol. 27, pp. 28–35, May 2012.

[92] E. Paolini, A. Giorgetti, M. Chiani, R. Minutolo, and M. Montanari, "Localization capability of cooperative anti-intruder radar systems," *EURASIP Journal on Advances in Signal Processing*, 2008.

[93] S. Meguerdichian, F. Koushanfar, M. Potkonjak, and M. Srivastava, "Coverage problems in wireless ad-hoc sensor networks," in *IEEE INFOCOM 2001*.

[94] S. Meguerdichian, S. Slijepcevic, V. Karayan, and M. Potkonjak, "Localized algorithms in wireless ad-hoc networks: Location discovery and sensor exposure," in *ACM MOBIHOC 2001*.

[95] X.-Y. Li, P.-J. Wan, and O. Frieder, "Coverage in wireless ad hoc sensor networks," *IEEE Transactions on Computers*, vol. 52, pp. 753–763, Jun. 2003.

[96] M. Skolnik, *Introduction to radar systems*. McGraw-Hill, 2002.

[97] A. A. Dutta, P. and S. Bibyk, "Towards radar-enabled sensor networks," in *ACM/IEEE IPSN 2006*.

[98] Q. Liang, "Waveform design and diversity in radar sensor networks: Theoretical analysis and application to automatic target recognition," in *IEEE SECON 2006*.

[99] T. Hanselmann, M. Morelande, B. Moran, and P. Sarunic, "Constrained multi-object Markov decision scheduling with application to radar resource management," in *Infomation Fusion 2010*.

[100] M. Li, T. Yan, D. Ganesan, E. Lyons, P. Shenoy, A. Venkataramani, and M. Zink, "Multi-user data sharing in radar sensor networks," in *ACM SenSys 2007*.

[101] S. Bartoletti, S. Conti, and A. Giorgetti, "Analysis of UWB radar sensor networks," in *IEEE ICC 2010*.

[102] X. Gong, J. Zhang, and D. Cochran, "When target motion matters: Doppler coverage in radar sensor networks," in *IEEE INFOCOM 2013*.

[103] S. Meguerdichian, F. Koushanfar, G. Qu, and M. Potkonjak, "Exposure in wireless ad-hoc sensor networks," in *ACM MOBICOM 2001*.

[104] R.-H. Gau and Y.-Y. Peng, "A dual approach for the worst-case-coverage deployment problem in ad-hoc wireless sensor networks," in *IEEE MASS 2006*.

[105] C. Lee, D. Shin, S.-W. Bae, and S. Choi, "Best and worst-case coverage problems for arbitrary paths in wireless sensor networks," in *IEEE MASS 2010*.

[106] G.-Y. Keung, B. Li, and Q. Zhang, "The intrusion detection in mobile sensor network," in *ACM MOBIHOC 2010*.

[107] Y. Wang and G. Cao, "Barrier coverage in camera sensor networks," in *ACM MOBIHOC 2011*.

[108] H. Ma, Y. Meng, D. Li, Y. Hong, and W. Chen, "Minimum camera barrier coverage in wireless camera sensor networks," in *IEEE INFOCOM 2012*.

[109] "Earthscope: Exploring the structure and evolution of the north american continent." [Online]. Available: http://www.earthscope.org

[110] Y. Wang and G. Cao, "Barrier coverage in camera sensor networks," in *ACM MOBIHOC 2011*.

[111] S. Kumar, T.-H. Lai, and J. Balogh, "On k-coverage in a mostly sleeping sensor network," in *ACM MOBICOM 2004*.

[112] B. Kusy, A. Ledeczi, and X. Koutsoukos, "Tracking mobile nodes using RF Doppler shifts," in *ACM SenSys 2007*.

[113] J.-H. Lim, I.-J. Wang, and A. Terzis, "Tracking a non-cooperative mobile target using low-power pulsed Doppler radars," in *IEEE Local Computer Networks (LCN) 2010*.

[114] R. Kershner, "The number of circles covering a set," *American Journal of Mathematics*, vol. 61, p. 665671, 1939.

[115] X. Bai, Z. Yun, D. Xuan, T.-H. Lai, and W. Jia, "Optimal patterns for four-connectivity and full coverage in wireless sensor networks," *IEEE Trans. Mobile Computing*, vol. 9, p. 435448, 2010.

[116] Y. Wang and G. Cao, "On full-view coverage in camera sensor networks," in *IEEE INFOCOM 2011*.

[117] Y. Wu and X. Wang, "Achieving full view coverage with randomly-deployed heterogeneous camera sensors," in *IEEE ICDCS 2012*.

[118] R. Bassily, E. Ekrem, X. He, E. Tekin, J. Xie, M. R. Bloch, S. Ulukus, and A. Yener, "Cooperative security at the physical layer," *IEEE Signal Processing Magazine*, vol. 30, pp. 16–28, Sept. 2013.

[119] K. N. Ramachandran, E. M. Belding-Royer, K. C. Almeroth, and M. M. Buddhikot, "Interference-aware channel assignment in multi-radio wireless mesh networks." in *INFOCOM*, vol. 6, 2006, pp. 1–12.

[120] A. Beresford and F. Stajano, "Location privacy in pervasive computing," in *IEEE Pervasive Computing 2003*.

[121] J. Freudiger, M. H. Manshaei, J.-P. Hubaux, and D. C. Parkes, "On non-cooperative location privacy: A game-theoretic analysis," in *ACM CCS 2009*.

[122] X. Liu, H. Zhao, M. Pan, H. Yue, X. Li, and Y. Fang, "Traffic-aware multiple mix zone placement for protecting location privacy," in *IEEE INFOCOM 2012*.

[123] D. Yang, X. Fang, and G. Xue, "Truthful incentive mechanisms for k-anonymity location privacy," in *IEEE INFOCOM 2013*.

[124] X. Liu, K. Liu, L. Guo, X. Li, and Y. Fang, "A game-theoretic approach for achieving k-anonymity in location based services," in *IEEE INFOCOM 2013*.

[125] B. Hoh and M. Gruteser, "Protecting location privacy through path confusion," in *SECURECOMM 2005*.

[126] K.-C. Chen, M. Chiang, and H. Poor, "From technological networks to social networks," *IEEE J. Sel. Area. Comm.*, vol. 31, no. 9, pp. 548–572, Sept. 2013.

[127] B. Hoh, M. Gruteser, H. Xiong, and A. Alrabady, "Enhancing security and privacy in traffic-monitoring systems," in *IEEE Pervasive Computing 2006*.

[128] J. Krumm, "Inference attacks on location tracks," in *IEEE Pervasive Computing 2007*.

[129] B. Mu and S. Bakiras, "Private proximity detection via computational geometric approaches," in *ACM MobiDE 2013*.

[130] R. Zhang, Y. Zhang, J. Sun, and G. Yan, "Fine-grained private matching for proximity-based mobile social networking," in *IEEE INFOCOM 2012*.

[131] L. Zhang, X.-Y. Li, and Y. Liu, "Message in a sealed bottle: Privacy preserving friending in social networks," in *IEEE ICDCS 2013*.