UNIVERSITY OF CALIFORNIA
Los Angeles

System and Incentive Design in Socio-technical Networks

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Electrical Engineering

by

Yu Zhang

2013

Abstract of the Dissertation

System and Incentive Design in Socio-technical Networks

by

Yu Zhang

Doctor of Philosophy in Electrical Engineering

University of California, Los Angeles, 2013

Professor Mihaela van der Schaar, Chair

Socio-technical networks (e.g. social networking services, peer-to-peer systems, etc.) provide a popular, cost-effective and scalable framework for sharing user-generated resources or services. Achieving resource sharing efficiency in socio-technical networks is a challenging problem, because the information available about the various resources is decentralized and it is changing dynamically; the agents may be heterogeneous and have different learning abilities; the agents may make proactive decisions on link formation; and most importantly, the agents may be self-interested, i.e. they take actions which maximize their individual utilities rather than the collective social welfare and thus choose to free-ride rather than share their resources.

The overarching goal of my dissertation is to develop a rigorous and unified paradigm for the joint design of efficient incentive mechanisms and resource management schemes in socio-technical networks. It can be generally divided into two parts.

The first part focuses on the efficient resource sharing in socio-technical networks. Existing distributed network optimization techniques that enable efficient resource allocation when agents are obedient or cooperative are no longer suitable in socio-technical networks which are formed by self-interested agents. The strategic interactions of such self-interested agents lead in numerous socio-technical networks to (Nash) equilibria

that are highly inefficient from a social perspective. To achieve social efficiency, incentives need to be provided to agents such that they find in their own self-interest to cooperate and thus act in a socially-optimal way. I propose a general methodology for the design and analysis of rating protocols and associated multi-agent learning algorithms to sustain cooperation in socio-technical networks. Under a rating protocol, an agent is rated based on its behavior. The rating affects the agent's rewards received in the network, which are typically determined according to a differential resource management scheme: compliant agents receive higher ratings and are rewarded by gaining more access to resources compared to non-compliant agents. This preferential treatment thus provides an incentive for agents to cooperate. I rigorously formalize and study the design of rating protocols to optimize the social resource sharing efficiency while encompassing various unique features of socio-technical networks, including the anonymity of agents, asymmetry of interests between different parties in the network, imperfect monitoring, dynamics in the agent population, and white-washing effects (i.e., an individual agent creating multiple identities in the network).

Different from the first part where the underlying network topology is exogenously determined, the second part of my dissertation augments the proposed rating protocols by investigating the endogenous formation of network topologies by the strategic, self-interested agents who produce, disseminate or collect resources. I propose a novel game-theoretic framework to model and analyze the trade-offs (of each individual agent) between the costs and benefits of producing resources personally and forming links to acquire and disseminate resources. A central point of my analysis, which departs from the existing literature on social network formation, is the assumption that the strategic agents are heterogeneous and that agents value this heterogeneity. The heterogeneity of agents and the ability of agents to strategically produce, disseminate or collect resources have striking consequences on the endogenously emerging topology, which provide important guidelines for the design of effective incentive mechanisms and resource management schemes in endogenous socio-technical networks. I first show that the network topology that emerges (at equilibrium) necessarily displays a core-periphery type: hub agents (at the core of the network) produce most of the resources and also create and maintain links to the agents at the periphery, while spoke agents (at the periphery of the network) derive most of their resources from hub agents, producing little of it themselves. As the population becomes larger, the number of hub agents and the total amount of resources produced grow in proportion to the total population. I then show that the networks that emerge at equilibrium are frequently minimally connected and have short network diameters. These “scale-free” conclusions had been conjectured for many networks, such as the “small-world” phenomenon in the World-Wide-Web, but not derived in any formal framework, and are in stark contradiction to the “law of the few” that had been established in previous work, under the assumption that agents solely benefit by forming links for resource acquisition, while resources are homogeneous and part of the endowment of agents, rather than heterogeneous and produced.

The dissertation of Yu Zhang is approved.

Jason L. Speyer

Ichiro Obara

Alan J. Laub

Panagiotis Christofides

Mihaela van der Schaar, Committee Chair

University of California, Los Angeles

2013

To my parents and my wife

Table of Contents

LIST OF FIGURES

List of Tables

Acknowledgments

Foremost, I would like to express my sincere appreciation and deepest gratitude to my advisor Professor Mihaela van der Schaar for her generous support and invaluable guidance throughout the past five years on both research and life. I really appreciate her comprehensive vision of research and her endeavor for high-quality work, and also have benefited tremendously from her unique blend of enthusiasm, energy, and technical insights. I would also like to thank the other members of my dissertation committee, Professor Panagiotis Christofides, Professor Alan J. Laub, Professor Ichiro Obara, and Professor Jason L. Speyer for their time and efforts in evaluating my work.

I am grateful to Dr. Deepak S. Turaga and Dr. Daby M. Sow for being my mentors during my internship at the Exploratory Stream Analytics department in IBM T. J. Watson Research Center, where I spent wonderful three months doing very interesting research with them.

I have also had the great fortune to learn from and work with so many incredible individuals at UCLA who made my five years here an unforgettable experience. I would like to thank all my friends for the great time we had together. In particular, I would like to thank all my amazing colleagues in the Networks, Economics, Communication Systems, Informatics and Multimedia Research Lab for countless inspiring discussions, comments, and advices, from every one of which I learned something new: Fangwen Fu, Hsien-po Shiang, Yi Su, Hyunggon Park, Jaeok Park, Nicholas Mastronarde, Shaolei Ren, Yuanzhao Xiao, Jie Xu, Linqi Song, Luca Canzian, Cem Tekin. My dearest friends outside UCLA, Yihong Wu, Bo Jiang, Shuo Guo, Yue Zhao, are always there for me during both joyful and stressful times, to whom I will always be thankful.

Finally, I wish to express my deepest gratitude to my family for their love, endless support, and encouragement throughout my Ph.D. study. Most importantly, I am incredibly grateful for the love and support from my wife, Yumo Chen, with whom I have been through all the ups and downs, in happiness and in sorrow. Words can-

not describe my immense gratitude for my family whose kindness and inspiration have helped me pursue my life long dreams.

VITA

1984	Born in Zhoushan, China.
2006	B.E. in Electrical Engineering Department Tsinghua University, Beijing, China
2008	M.S. in Electrical Engineering Department Tsinghua University, Beijing, China
2008–2013	Research Assistant, Electrical Engineering Department UCLA, Los Angeles, California
2011	Research Intern, IBM T. J. Watson Research Center Yorktown Heights, New York

PUBLICATIONS

S. Won, I. Cho, K. Sudusinghe, J. Xu, Yu Zhang, M. van der Schaar, S. S. Bhattacharyya, "A Design Methodology for Distributed Adaptive Stream Mining Systems," *International Conf. on Computational Science*, 2013.

Yu Zhang, M. van der Schaar, "Strategic Information Dissemination and Link Formation in Social Networks," *IEEE ICASSP*, 2013.

J. Xu, Yu Zhang, M. van der Schaar, "Rating Systems for Enhanced Cyber-security Investments," *IEEE ICASSP*, 2013.

Y. Xiao, Yu Zhang, M. van der Schaar, “Socially-Optimal Design of Crowdsourcing Platforms With Reputation Update Errors,” *IEEE ICASSP*, 2013.

Yu Zhang, D. Sow, D. Turaga, M. van der Schaar, “A Fast Online Learning Algorithm for Distributed Mining of BigData,” *SIGMETRICS Workshop*, 2013.

Yu Zhang, M. van der Schaar, “Strategic Networks: Information Dissemination and Link Formation Among Self-interested Agents,” *IEEE Journal on Selected Areas in Communications*, *Special Issue on Network Science*, vol. 31, no. 6, to be published in June 2013.

Yu Zhang, M. van der Schaar, “Strategic Learning and Robust Reputation Protocol Design for Online Communities: A Stochastic Stability Analysis,” *IEEE Journal of Selected Topics in Signal Processing, Special Issue on Learning-based Decision Making in Dynamic Systems under Uncertainty*, accepted in 2013.

Yu Zhang, J. Park, M. van der Schaar, “Rating Protocols for Online Communities,” *ACM Transactions on Economics and Computation*, accepted in 2012.

Yu Zhang, M. van der Schaar, “Incentive Provision and Service Allocation in Social Cloud Systems,” *IEEE Journal on Selected Areas in Communications, Special Issue on Emerging Technology in Communications*, accepted in 2012.

Yu Zhang, M. van der Schaar, “Information Production and Link Formation in Social Computing Systems,” *IEEE Journal on Selected Areas in Communications, Special Issue on Economics of Communication Networks and Systems*, vol. 30, no. 11, pp. 2136-2145, 2012.

Yu Zhang, M. van der Schaar, “Peer-to-Peer Multimedia Sharing based on Social Norms,” *Elesevier Journal on Signal Processing: Image Communication, Special Issue on Advances in Video Streaming over P2P Networks*, vol. 27, no. 5, pp. 383-400, 2012.

Yu Zhang, F. Fu, M. van der Schaar, “On-line Learning and Optimization for Wireless Video Transmission,” *IEEE Transactions on Signal Processing*, vol. 58, no. 6, pp. 3108-3124, 2010.

Yu Zhang, M. van der Schaar, “Reputation-based Protocols in Crowdsourcing Applications,” *IEEE INFOCOM*, 2012.

Yu Zhang, M. van der Schaar, “Collective Rating for Online Labor Markets,” *Allerton Conf.*, 2012.

C. J. Ho, Yu Zhang, J. Wortman Vaughan, M. van der Schaar, “Towards Social Norm Design for Crowdsourcing Markets,” *HCOMP*, 2012.

Yu Zhang, M. van der Schaar, “User Adaptation and Long-run Evolution in Online Communities,” *IEEE CDC*, 2011.

Yu Zhang, M. van der Schaar, “Influencing the Long-Term Evolution of Online Communities Using Social Norms,” *Allerton Conf.*, 2011.

Yu Zhang, J. Park, M. van der Schaar, “Designing Social Norm Based Incentive Schemes to Sustain Cooperation in a Large Community,” *Gamenets*, 2011.

Yu Zhang, M. van der Schaar, “Social Norm and Long-run Learning in Peer-to-Peer Networks,” *IEEE ICASSP*, 2011.

Yu Zhang, J. Park, M. van der Schaar, “Social Norm based Incentive Mechanisms for Peer-to-Peer Networks,” *IEEE ICASSP*, 2011.

Yu Zhang, M. van der Schaar, “Designing Incentives for P2P Multimedia Sharing,” *IEEE Globecom*, 2011.

Yu Zhang, F. Fu, M. van der Schaar, “On-line Learning for Wireless Video Transmission with Limited Information,” *Packet Video Workshop*, 2009.

CHAPTER 1

Introduction

1.1 Motivation

People turn to the web to exchange services, data, and resource, as evidenced by the popularity of online social networking services like Facebook and Twitter, peer production sites like Wikipedia, peer-to-peer (P2P) systems like BitTorrent, local business review sites like Yelp, and online labor markets like Amazon Mechanical Turk. While these online systems differ in many ways, they share a common vulnerability to selfish behavior and free-riding. A worker on Amazon Mechanical Turk may attempt to complete jobs with as little effort as possible while still being paid. Similarly, a user in a peer-to-peer system may wish to download files from others without using its bandwidth to upload files to others. In view of this fact, it is of critical importance to properly motivate the self-interested participants to contribute their resource, knowledge or services in order for these online systems to thrive.

Systems like these, which I refer to as socio-technical networks, have been attracting unprecedented amount of attention. In order to fully realize the benefits of such systems, tremendous technological efforts have been dedicated in the past decades to enhance the system performance in terms of resource sharing efficiency, fairness, reliability, derived revenues etc. Among the efforts, a major category of the existing research in engineering has been devoted to proposing optimization techniques to enable the efficient allocation of resource that largely relies on system-wide centralized/decentralized management over obedient or cooperative agents. Nevertheless, this research typically ignores the potentially devastating impact of selfish behavior and free-riding from self-interested agents when they are not correctly motivated to contribute and thus solely have incentives to optimize their individual utilities rather than the collective social welfare. The strategic interactions of such self-interested agents lead in numerous socio-technical networks to (Nash) equilibria that are highly inefficient from a social perspective. Only recently have engineers started to investigate incentive issues in systems formed by self-interested agents. The existing mechanisms to combat free-riding problems rely on some well-studied economic concepts such as pricing and direct reciprocity. These mechanisms, though appropriate in some settings, do not make sense for most socio-technical networks due to their unique technical characteristics (e.g. agents are not willing or not capable to pay money for services, they are anonymous and do not maintain long-lasting relationships, the underlying network topology is formed endogenously due to agents' proactive creation and dissemination of communication links, no credible monitoring technology to track agents' behavior, various accidents in operation, etc.).

In this dissertation, I develop a rigorous and unified framework for the analysis as well as the design of efficient incentive mechanisms and resource management schemes for socio-technical networks, which take into account the unique technical characteristics of the considered application, the environment in which it is deployed, and the information available to both the agents and the designer (e.g. the entity running the system or the distributed protocol designer). In the development of this framework, I specifically seek answers to the following two questions: (i) how the platform designer can configure and operate socio-technical networks at the highest level of efficiency and robustness by designing optimal incentive mechanisms and associated resource management rules; and (ii) how the agents' self-interested nature, decentralized and heterogeneous knowledge, and their learning and link formation abilities will shape their interactions and impact the optimal incentive mechanism design.

Importantly, the proposed framework can serve as a guideline for platform designers when modeling, analyzing, comparing, and selecting incentive mechanisms and resource management schemes for novel social-technical networks, and has the potential to influence a wide variety of related domains that include (but are not limited to) social networking services, social sharing services, cloud/social computing, and cyber-security.

1.2 Key Challenges

The key design challenges brought forward by socio-technical networks are:

- **Asymmetry of interests.** In socio-technical networks, agents usually have asymmetric interests on a variety of resources. As an example, consider a knowledge-sharing website where users with different areas of expertise share knowledge with each other. It will be rarely the case that a pair of users has a mutual interest in the expertise of each other simultaneously. In the design of incentive mechanisms and resource allocation schemes, I allow the possibility of asymmetric interests by modeling the interaction between a pair of users as a gift-giving game, instead of a prisoner's dilemma game which is commonly adopted in the existing literature and assumes mutual interests between a pair of users.

- **Imperfect monitoring.** The design of effective incentive mechanisms in the existing literature relies on the perfect monitoring over agents' resource sharing behavior. Nevertheless, the monitoring technologies implemented in socio-technical networks can never be perfect in practice. This makes the design problem more complicated. If monitoring was perfect, the platform designer can employ a strong punishment on an agent's selfish behavior (e.g. a trigger strategy): isolating an agent from the system if it is detected not sharing its resources in a cooperative manner. Such approach provides the strongest incentive for individual agents to cooperate. However, when monitoring is imperfect, its results cannot be guaranteed to be accurate and thus a strong punishment by the incentive mechanism would therefore lead to very low sharing efficiency due to the false punishments on cooperative agents. Instead, choosing milder punishments becomes more efficient. However, the punishment should also not be too mild such that sufficient incentives are provided to agents to cooperate. Therefore, an optimal incentive mechanism needs to be designed as a tradeoff between sharing efficiency and incentive-compatibility.

- **Whitewashing**. Another distinctive characteristic of socio-technical networks is that the agents are anonymous (i.e. their online identities are not directly associated with their real-world identities). Meanwhile, an agent can create multiple online identities by repeatedly entering the system. As a result, an agent who has been detected to be non-cooperative and is subject to punishments may attempt to switch its online identity by leaving and rejoining the community as a new member to avoid the punishments imposed by the system upon its old identity, a behavior commonly known as "whitewashing". The incentive mechanisms proposed in this dissertation consider this possibility and are designed to be "whitewashing-proof".

- **Distributed information availability and strategic learning.** Due to the distributed availability of information in socio-technical networks, agents do not have complete knowledge about the system and other agents. Thus, it is of particular importance in this regard to investigate how to approach the performance bounds of incentive mechanisms when agents can proactively adapt their strategies based on limited, heterogeneous and evolving knowledge and reasoning about the system, the environment, and the strategies of other agents. It should be also noted that taking into account the agents' strategic learning over time will couple the platform designer's decision with the agents' self-interested decisions, thereby significantly complicating the formal analysis and design of incentive mechanisms.

- **Strategic link formation.** One of the defining features that distinguish socio-technical networks from traditional communication networks is that agents make strategic and independent decisions regarding forming links and exchanging the resources with other agents. For example, a user on Facebook or Twitter can strategically determine its friendship connections; a peer node in P2P networks makes active decisions on which peers it should connect to and download content from. Hence, the resource management schemes proposed in the existing literature, which assume that the network topology in the system is exogenously determined and fixed, are no longer applicable to socio-technical networks that are usually formed endogenously by self-interested agents who determine their link formation strategies in a way to maximize their individual utilities.

1.3 Contribution of the Dissertation

The overarching goal of my dissertation is to develop a rigorous and unified paradigm for the joint design of efficient incentive mechanisms and resource management schemes for socio-technical networks. It can be generally divided into two parts.

The first part (Chapter 2-3) focuses on the design of efficient resource sharing schemes in socio-technical networks. I propose a general methodology for the design and analysis of rating protocols and associated multi-agent learning algorithms to sustain cooperation in socio-technical networks. Formulating the agents' interaction as non-cooperative repeated games, I rigorously model and study the design of rating protocols to optimize the resource sharing efficiency among agents while encompassing various unique features of socio-technical networks. I also propose practical benchmarks and criteria to evaluate the performance of incentive mechanisms both in terms of economic efficiency and implementation overhead for a variety of applications.

The second part of my dissertation (Chapter 4-5) augments the proposed rating protocols by explicitly considering the strategic resource production and link formation of agents, where the resource represents “information” (e.g. files in P2P networks, news on social networks, traffic/road conditions in vehicular networks, etc.). I propose a novel game-theoretic framework to model and analyze the trade-offs (of each individual agent) between the costs and benefits of producing information personally and forming links to collect and disseminate resources. The analysis has implications for the topology that emerges endogenously. For large populations, the implication is that the topology is necessarily of a core-periphery type. Meanwhile as the population grows, the agents’ degree distribution converges to a “power-law” distribution and the network becomes “scale-free”. My conclusions had been conjectured by numerous empirical studies on distributed networking systems and social computing systems but had not been previously derived formally.

1.3.1 Chapter 2: A General Framework of Rating Protocol Design for Online Communities

In Chapter 2 of the dissertation, I propose a general framework that enables the rigorous design of rating protocols for service and resource sharing in online communities. I model the agents’ interactions as anonymous random matching games, in which each agent is matched with different partners over time to exchange services. The goal of each agent is to maximize its discounted long-term utility. My proposed rating protocols rely on the idea of social reciprocation. Under a rating protocol, an agent is rated based on its behavior. The rating affects the agent’s rewards received in the community, which are typically determined according to a differential resource management scheme: compliant agents receive higher ratings and are rewarded by gaining more access to resources compared to non-compliant agents. This preferential treatment thus provides an incentive for agents to cooperate. I derive conditions for sustainable rating protocols, under which no agent gains by deviating from the socially-optimal resource sharing strategy. Given these sustainability conditions, I formulate and solve the problem of designing an optimal rating protocol and characterize the optimal social welfare. As special cases, I also analyze the one-sided rating protocol which only utilizes the rating score of one party in the stage-game. I show that when only the rating scores of clients (i.e. agents who request resources) are utilized, the optimal one-sided rating protocol preserves a simple structure with two-level rating scores, whereas when only the rating scores of servers (i.e. agents who share resources) are utilized, no sustainable one-sided rating protocol can be designed. Under this framework, I also study the impacts of punishment lengths and whitewashing possibility on the design and performance of optimal rating protocols, identifying a trade-off between efficiency and incentives. Lastly, I present numerical results to illustrate the impacts of the discount factor, the turnover rate, and the probability of report errors on the performance of optimal rating protocols.

1.3.2 Chapter 3: Strategic Learning in Online Communities

Chapter 2 assumes that the community starts to operate directly at a selected ("best") equilibrium, i.e. agents start by directly adopting a desirable equilibrium strategy when joining the community, and rigorously designs rating protocols that ensure agents will not deviate from this equilibrium strategy and that the community will indeed operate at the selected equilibrium. Nevertheless, in practice, there are often multiple equilibria in a community and agents can adopt arbitrary (non-equilibrium) strategies when they join the community and learn to adapt their strategies based on the accrued knowledge about the community and other agents.

Chapter 3 augments the design in Chapter 2 by specifically investigating how the strategic learning of self-interested agents impacts the design of optimal rating protocols. To optimize their individual long-term performance, users learn the environment and adapt their strategies by solving individual stochastic control problems. The users' learning and adaptation catalyze a stochastic dynamic process, in which the strategies of users in the community evolve over time. I first characterize the structural properties of the users' best response strategies. Subsequently, using these structural results I design rating protocols for governing the online communities, which not only sustains the best equilibrium, but also ensures the convergence of the community towards this equilibrium in the long-run. I prove that by appropriately penalizing and rewarding users based on their behavior in the community, such incentive mechanisms can eliminate free-riding and ensure that the community converges to a desirable equilibrium selected by the community designer such that social welfare is maximized and in which users learn that it is in their self-interest to cooperate with each other.

1.3.3 Chapter 4: Information Production and Link Formation in Social Computing Systems

In Chapter 4, I focus on the analysis of strategic link formation behavior of self-interested agents in social computing systems. This analysis is conducted from several points of view. First, I analyze the trade-offs (of each individual agent) between the costs and benefits of producing information personally and forming links to collect information (from other agents), and the strategic implications of these trade-offs. A central point of the analysis is that information are assumed to be heterogeneous (rather than homogeneous as in previous analyses), which is formulated by the well-known Dixit-Stiglitz utility function, and agents value this heterogeneity. The analysis has implications for the topology that emerges endogenously. For large populations, the implication is that the topology is necessarily of a core-periphery type: hub agents (at the core of the network) produce and share most of the information, while spoke agents (at the periphery of the network) derive most of their information from hub agents, producing little of it themselves. As the population becomes larger, the number of hub agents and the total amount of information produced grow in proportion to the total population. These conclusions had been conjectured for many social computing systems but it has not been previously derived in any formal framework, and are in stark contradiction to the “law of the few” that had been established in previous work, under the assumption that information is homogeneous and part of the endowment of agents, rather than heterogeneous and produced.

1.3.4 Chapter 5: Information Dissemination in Socio-technical Networks

Chapter 5 analyzes an alternative problem in strategic link formation, in which the self-interested agents are benefit from disseminating its own information (e.g. knowledge, data, advertisement) instead of collecting the information produced by other agents. I formulate the agents' interactions as a non-cooperative game, where each agent proactively determines its own information production and link formation strategies in order to maximize its individual utility from disseminating its produced information. The strategic production and dissemination of information have striking consequences. I show first that the network structure that emerges (in equilibrium) typically displays a core-periphery structure, with the few agents at the core playing the role of "connectors", creating and maintaining links to the agents at the periphery. I then determine conditions under which the networks that emerge are minimally connected and have short network diameters (properties that are important for efficiency). Finally, I show that the number of agents who produce information and the total amount of information produced in the network grow at the same rate as the agent population. That is, the "law of the few" also does not in socio-technical networks where agents benefit from information dissemination.

1.3.5 Chapter 6: Conclusion

Chapter 6 concludes the dissertation and includes a discussion about future research directions.

CHAPTER 2

A General Framework of Rating Protocol Design for Online Communities

2.1 Introduction

Recent developments in technology have expanded the boundaries of communities in which individuals interact with each other. For example, nowadays individuals can obtain valuable information or content from remotely located individuals in an online community formed through online networking services [1]-[7]. However, a large population and the anonymity of individuals in such an online community make it difficult to sustain cooperative behavior among self-interested individuals [8][9]. For example, it has been reported that "free-riding" is widely observed in peer-to-peer networks [10][11]. Hence, incentive schemes are needed to cultivate cooperative behavior in online communities.

A variety of incentive schemes have been explored to induce cooperation in such online communities. The most popular incentives are based on pricing schemes and differential service provision. Pricing schemes use payments to reward and punish individuals for their behavior, which in principle can induce self-interested individuals to cooperate with each other to attain the social optimum by internalizing their external effects (see, for example, [12][13]). The main challenge with the pricing scheme is that the resources/services being exchanged need to be accurately priced in order to make the punishment scheme effective. However, this prerequisite is difficult to fulfill in many online communities, where agents interact frequently with each other and the services being exchanged between them are not real goods but rather solutions to small tasks or small amounts of resources which are difficult to price. One example of such applications is the online question and answer forum [1], where the service represents a small "favor" in answering the questions posted by other agents, i.e. knowledge is what is being exchanged here. The difficulty in pricing such "small" services (i.e. the knowledge and resources being exchanged) in these applications prevents the pricing scheme to be effective. Therefore, it is more effective to incentivizing agents to provide services by rewarding them in the form of rating scores/virtual credits, which allow them to obtain the same type of services in the future from the community as a returning favor. Also, a pricing scheme often requires a complex accounting infrastructure, which introduces substantial communication and computation overheads [14]. Hence, it is impractical for the pricing scheme and the corresponding infrastructure to be implemented in large-scale online communities, e.g. peer-to-peer systems [17], mobile networks [2], etc., where agents have limited computing and communication capabilities and are interacting frequently with each other for small services (i.e. frequent transactions). For example, there have been some works that design monetary-based incentive mechanisms for peer-to-peer systems where each peer maintains a bank account and use real money to purchase resources. However, it has been measured that significant overheads are introduced by the deployment of a payment infrastructure [14] and hence, such pricing schemes were never actually implemented in such online communities. Finally, pricing the services in online communities might discourage agents from participating into the community.

Differential service schemes, on the other hand, reward and punish individuals by providing differential services depending on their behavior instead of using monetary rewards [15]-[30]. Differential services can be provided by community operators or by community members. Community operators can treat individuals differentially (for example, by varying the quality or scope of services) based on the information about the behavior of individuals. Incentive provision by a central entity can offer a robust method to sustain cooperation [15]. However, such an approach is impractical in a large community because the burden of a central entity to monitor individuals' behavior and provide differential services for them becomes prohibitively heavy as the population size grows. Alternatively, more distributed incentive schemes exist where community members monitor the behavior of each other and provide differential services based on their observations [16]-[30]. Such incentive schemes are based on the principle of reciprocity and can be classified into personal reciprocation (or direct reciprocity) [16]-[18] and social reciprocation (or indirect reciprocity) [19]-[30]. In personal reciprocation schemes, individuals can identify each other, and behavior toward an individual is based on their personal experience with that individual. Personal reciprocation is effective in sustaining cooperation in a small community where individuals can identify each other and interact frequently with fixed opponents, but it loses its power in a large community where individuals have asymmetric interests and can freely and frequently change the opponents they interact with [26]. In social reciprocation schemes, individuals obtain some information about other individuals (for example, rating) and decide their actions toward an individual based on this available information. Hence, an individual can be rewarded or punished by other individuals in the online community who have not had past interactions with it [26][27]. Therefore, social reciprocation has a potential to form a basis of successful incentive schemes for online communities. As such, this chapter is devoted to the study of incentive schemes based on social reciprocation.

Sustaining cooperation using social reciprocation has been investigated in the literature using the framework of anonymous random matching games, in which each individual is repeatedly matched with different partners over time for service exchange and tries to maximize its discounted long-term utility. To implement social reciprocation, it is important for the community to share enough information about past interactions such that the community members know how to reward or punish others. This existing literature makes different assumptions on the information revealed to community members about other members. In [28] each community member observes the entire history of the past plays of its current partner. In [29][30], community members are informed about the outcomes of the matches in which they have been directly involved. Rating protocols have been proposed in [26] and [27], where each community member is attached a rating score indicating its social status, which takes a value from a finite set and records its past plays, and community members with different rating scores are treated differently by other individuals they interact with. For online communities, maintaining direct records of individuals' past plays which are used in [28]-[30] are not appropriate, because the communication and storage cost for revealing the entire history of the past plays of an individual grow unbounded with time. Since the use of rating score as a summary record requires significantly less amount of information to be maintained, we will design incentive schemes based on rating protocols for online communities. A rating-based incentive scheme can be easily implemented in online communities that deploy entities (e.g., a tracker in P2P networks [17], or a web portal in web-based applications [1]) who can collect, process, and deliver information about individuals' play history to generate rating scores.

Cooperation among community members can be sustained in all the above works on anonymous random matching games. However, all of them have focused on obtaining the Folk Theorem by characterizing the set of equilibrium payoffs that can be achieved when the discount factor of individuals is sufficiently close to 1. Our work, on the contrary, addresses the problem of designing a rating-based incentive scheme given a discount factor and other parameters arising from practical considerations, which are not fully considered in the existing literature on anonymous random matching games. Specifically, our work takes into account the following features of online communities:

- *Asymmetry of interests.* As an example, consider a community where individuals with different areas of expertise share knowledge with each other. It will be rarely the case that a pair of individuals has a mutual interest in the expertise of each other simultaneously. We allow the possibility of asymmetric interests by modeling the interaction between a pair of individuals as a gift-giving game, instead of a prisoner's dilemma game which assumes mutual interests between a pair of individuals [17][26]-[28]. It should be noted that the "asymmetry of interests" in this chapter specifically refers to the fact that two agents in one stage game do not have mutual interests in the resources or services possessed by each other simultaneously. Such "asymmetry of interests" exists in various applications. For example, in an online question and answer forum such as Yahoo! Answers, a user i who answers a question of another user j does not necessarily have to propose questions to user j as well [1].

- *Report errors.* In an incentive scheme based on a rating protocol, it is possible that the rating score (or label) of a specific individual is updated incorrectly because of errors in the reports of its partners (i.e. other individuals it interacts with). Our model incorporates the possibility of report errors, which allows us to analyze its impact on design and performance, whereas most existing works on rating schemes (e.g. [26][27]) adopt an idealized assumption that rating scores are always updated correctly.

- *Whitewashing.* Whitewashing refers to the behavior of an individual creating multiple identities by repeatedly entering to an anonymous online community. In an online community, individuals with bad rating scores may attempt to whitewash their rating scores by leaving and rejoining the community as new members to avoid the punishments imposed by the system upon their old identities [17]. We consider this possibility and study the design of "whitewash-proof" rating protocols and their performance.

Note that our model and analysis also differ significantly from most existing works on reputation systems [23]-[25]. First, the models in [23]-[25] assume that individuals assume fixed roles in the community (i.e. seller or buyer), which is common in applications where the groups of sellers and buyers are separated and usually do not overlap [21]. Nevertheless, in online communities such as P2P networks, online labor markets,

Table 2.1: Comparison with the existing literature

	[28]-[30][35]	[26][27]	[23][24]	[25]	Our work
Incentive device	Differential services	Differential services	Monetary rewards	Monetary rewards	Differential services
Asymmetry of interests	N/A	No	No	No	Yes
Report errors	N/A	No	Yes	Yes	Yes
Information requirement	Entire history of stage game outcomes	Individual rating	Individual rating	Individual rating	Individual rating
Discount factor	Sufficiently close to 1	Sufficiently close to 1	Sufficiently large	Arbitrary	Arbitrary
Number of long-lived players	Multiple	Multiple	One	One	Multiple
Protocol design	No	No	Yes	Yes	Yes
Optimization criterion	Individual long-term utility	Individual long-term utility	Individual long-term utility	Individual long-term utility	Sum utility of all players

etc., each agent can be both the provider and the receiver of services. Second, the reputation systems in [23]-[25] rely on differential pricing schemes to incentivize sellers to cooperate. We have already mentioned that the services being exchanged and shared in online communities are difficult to price, thereby preventing such pricing-based reputation systems to be effectively deployed. Finally and most importantly, [23]-[25] consider the repeated game between a unique long-lived seller and many short-lived buyers, and the design principle there is to maximize the expected (discounted) long-term utility of the individual long-lived seller. In contrast, we consider in this chapter the interplay among a large number of long-lived individuals, and aim to maximize the social welfare (i.e. the sum utility) of the entire community, which makes the designs in [23]-[25] inapplicable here. A more detailed and in-depth comparison between our work and [23]-[25] is provided in Section 2.7. The differences between our work and the existing literature on social reciprocity are summarized in Table 2.1 in order to highlight our contribution and novelty.

The remainder of this chapter is organized as follows. In Section 2.2, we describe the repeated anonymous matching game and incentive schemes based on a rating protocol. In Section 2.3, we formulate the problem of designing an optimal rating protocol. In Section 2.4, we provide analytical results about optimal rating protocols. In Section 2.5, we extend our model to address the impacts of variable punishment lengths, whitewashing possibility, and one-sided rating. We provide simulation results in Section 2.6, discuss the related works in Section 2.7, and conclude the chapter in Section 2.8.

2.2 Model

2.2.1 Repeated Matching Game

We consider a community where each member, or agent, can offer a valuable service to other agents. Examples of services are expert knowledge, customer reviews, job information, multimedia files, storage space, and computing power. We consider an infinite-horizon discrete time model with a continuum of agents [17] to highlight our focus on online communities with large agent populations. Such continuum population model is commonly adopted in the analysis for large-scale dynamic networks, e.g. peer-to-peer systems [3][17], grid networks [4], social sharing websites [1][5], etc. In a period, each agent generates a service request [37], which is sent to another agent that can provide the requested service [1]. We model the request generation and agent selection process using uniform random matching: each agent receives exactly one request in every period and each agent is equally likely to receive the request of an agent, and the matching is independent across periods [2]. Such model well approximates the matching process between agents in large-scale online communities where agents interact with others in an ad-hoc fashion and the interactions between agents are constructed randomly over time. For example, in a mobile relay network [2] where agents (e.g. mobile devices)

[1]It should be noted that our analysis can be readily extended to the case where each agent generates a service request with a probability $\lambda < 1$. We assume $\lambda = 1$ in this chapter only for the simplicity of illustrations.

[2]The impact of matching schemes on the incentive of agents and the performance of online communities falls out the scope of this chapter, but serves as an important next step in this line of research.

within a certain area are able to relay traffic for each other through unlicensed spectrum (e.g. WLAN) to the destination (e.g. nearby cellular base stations), the relay node that each mobile agent encounters at each moment could be approximately assumed to be random, since this mobile agent is moving around the area randomly over time. In other words, each idle agent in the same area has approximately the same probability to be chosen as the relay node.

In a pair of matched agents, the agent that requests a service is called a client while the agent that receives a service request is called a server. In every period, each agent in the community is involved in two matches, one as a client and the other as a server. Note that the agent with whom an agent interacts as a client can be different from that with whom it interacts as a server, reflecting asymmetric interests between a pair of agents at a given instant.

We model the interaction between a pair of matched agents as a gift-giving game [32]. In a gift-giving game, the server has the binary choice of whether to fulfill or decline the request, while the client has no choice. The server's action determines the payoffs of both agents. If the server fulfills the client's request, the client receives a service benefit of $b > 0$ while the server suffers a service cost of $c > 0$. We assume that $b > c$ so that the service of an agent creates a positive net social benefit. If the server declines the request, both agents receive zero payoffs. The set of actions for the server is denoted by $\mathcal{A} = \{F, D\}$, where F stands for "fulfill" and D for "decline". The payoff matrix of the gift-giving game is presented in Table 2.2. An agent plays the gift-giving game repeatedly with changing partners until it leaves the community. We assume that at the end of each period a fraction $\alpha \in [0, 1]$ of agents in the current population leave and the same amount of new agents join the community. We refer to α as the turnover rate [17].

Social welfare in a time period is measured by the average payoff of the agents in that period. Since $b > c$, social welfare is maximized when all the servers choose action F in the gift-giving games they play, which yields payoff $b - c$ to every agent. On the contrary, action D is the dominant strategy for the server in the gift-giving game, which constitutes a Nash equilibrium of the gift-giving game. When every server chooses its action to maximize its current payoff myopically, an inefficient outcome arises where every agent receives zero payoff.

Table 2.2: Payoff matrix of a gift-giving game

	Server F	Server D
Client	$b, -c$	$0, 0$

2.2.2 Incentive Schemes Based on a Rating Protocol

In order to improve the efficiency of the myopic equilibrium, we use incentive schemes based on rating protocols. A rating protocol is defined as the rules that a community uses to regulate the behavior of its members. These rules indicate the established and approved ways of "operating" (e.g., exchanging services) in the community: adherence to these rules is positively rewarded, while failure to follow these rules results in (possibly severe) punishments [36]. This gives rating protocols a potential to provide incentives for cooperation. We consider a rating protocol that consists of a rating scheme and a recommended strategy, as in [26] and [27]. A rating scheme determines the ratings of agents depending on their past actions as a server, while a recommended strategy prescribes the actions that servers should take depending on the ratings of the matched agents.

Formally, a rating scheme is represented by three parameters (Θ, K, τ): Θ denotes the set of rating scores that an agent can hold, $K \in \Theta$ denotes the initial rating score attached to newly joining agents, and τ is the rating update rule. After a server takes an action, the client sends a report (or feedback) about the action of the server to the third-party device or infrastructure that manages the rating scores of agents, but the report is subject to errors with a small probability ε. That is, with probability ε, D is reported when the server takes action F, and vice versa. Assuming a binary set of reports, it is without loss of generality to restrict ε in $[0, 1/2]$. When $\varepsilon = 1/2$, reports are completely random and do not contain any meaningful information about the actions of servers. We consider a rating scheme that updates the rating score of a server based only on the rating scores of matched agents and the reported action of the server. Then, a rating scheme can be represented by a mapping $\tau : \Theta \times \Theta \times \mathcal{A} \to \Theta$, where $\tau(\theta, \tilde{\theta}, a_R)$ is the new rating score for a server with current rating score θ when it is matched with a client with rating score $\tilde{\theta}$ and its action is reported as a_R. A recommended strategy is represented by a mapping $\sigma : \Theta \times \Theta \to \mathcal{A}$, where $\sigma(\theta, \tilde{\theta})$ is the approved action for a server with rating score θ that is matched with a client with rating score $\tilde{\theta}$ [3].

To simplify our analysis, we initially impose the following restrictions on rating schemes [4].

(1) Θ is a nonempty finite set, i.e., $\Theta = \{0, 1, \ldots, L\}$ for some nonnegative integer L.

(2) $K = L$.

(3) τ is defined by

$$\tau(\theta, \tilde{\theta}, a_R) = \begin{cases} \min\{\theta + 1, L\} & \text{if } a_R = \sigma(\theta, \tilde{\theta}) \\ 0 & \text{if } a_R \neq \sigma(\theta, \tilde{\theta}) \end{cases} . \tag{2.1}$$

Note that with the above three restrictions a nonnegative integer L completely describes a rating scheme, and thus a rating protocol can be represented by a pair $\kappa = (L, \sigma)$. We call the rating scheme determined by L the maximal punishment rating scheme (MPRS) with punishment length L. In the MPRS with punishment length L, there are $L + 1$ rating scores, and the initial rating score is specified as L. If the reported action of the server is the same as that specified by the recommended strategy σ, the server's rating score is increased by 1 while not exceeding L. Otherwise, the

[3] The strategies in the existing rating mechanisms [26][27] determine the server's action based solely on the client's rating score, and thus can be considered as a special case of the recommended strategies proposed in this chapter.

[4] We will relax the second and third restrictions in Section 2.5.

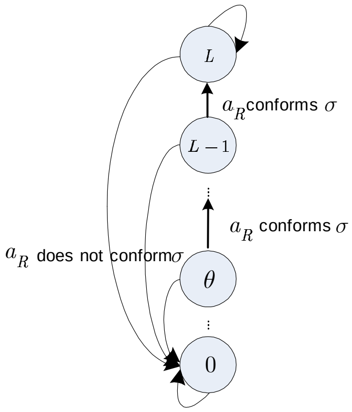


Figure 2.1: Schematic representation of a maximal punishment rating scheme

server's rating score is set as 0. A schematic representation of an MPRS is provided in Figure 2.1.

Below we summarize the sequence of events in a time period:

(1) Agents generate service requests and are matched.

(2) Each server observes the rating of its client and then determines its action.

(3) Each client reports the action of its server.

(4) The rating scores of agents are updated, and each agent observes its new rating score for the next period.

(5) A fraction of agents leave the community, and the same amount of new agents join the community.

2.3 Problem Formulation

2.3.1 Stationary Distribution of Rating Scores

As time passes, the rating scores of agents are updated and agents leave and join the community. Thus, the distribution of rating scores in the community evolves over time. Let $\eta^t(\theta)$ be the fraction of θ-agents in the total population at the beginning of an arbitrary period t, where a θ-agent means an agent with rating θ. Suppose that all the agents in the community follow a given recommended strategy σ. Then the transition from $\{\eta^t(\theta)\}_{\theta=0}^{L}$ to $\{\eta^{t+1}(\theta)\}_{\theta=0}^{L}$ is determined by the rating scheme, taking into account the turnover rate α and the error probability ε, as shown in the following expressions:

$$\begin{array}{rcl} \eta^{t+1}(0) & = & (1-\alpha)\varepsilon \\ \eta^{t+1}(\theta) & = & (1-\alpha)(1-\varepsilon)\eta^t(\theta-1) \text{ for } 1 \leqslant \theta \leqslant L-1 \\ \eta^{t+1}(L) & = & (1-\alpha)(1-\varepsilon)\{\eta^t(L)+\eta^t(L-1)\}+\alpha \end{array} . \tag{2.2}$$

Since we are interested in the long-term payoffs of the agents, we study the distribution of rating scores in the long run.

Definition 1 (Stationary distribution). $\{\eta(\theta)\}$ is a stationary distribution of rating scores under the dynamics defined by (2.2) if it satisfies $\sum_{\theta=0}^{L}\eta(\theta)=1$, $\eta(\theta)\geqslant 0, \forall\theta$, and

$$\begin{array}{rcl} \eta(0) & = & (1-\alpha)\varepsilon \\ \eta(\theta) & = & (1-\alpha)(1-\varepsilon)\eta(\theta-1) \text{ for } 1 \leqslant \theta \leqslant L-1 \\ \eta(L) & = & (1-\alpha)(1-\varepsilon)\{\eta(L)+\eta(L-1)\}+\alpha \end{array} . \tag{2.3}$$

The following lemma shows the existence of and convergence to a unique stationary distribution.

Lemma 1. For any $\varepsilon \in [0,1/2]$ and $\alpha \in [0,1]$, there exists a unique stationary distribution $\{\eta(\theta)\}$ whose expression is given by

$$\begin{array}{rcl} \eta(\theta) & = & (1-\alpha)^{\theta+1}(1-\varepsilon)^{\theta}\varepsilon, \; for \; 0 \leqslant \theta \leqslant L-1 \\ \eta(L) & = & \begin{cases} 1 & \text{if } \alpha=\varepsilon=0 \\ \frac{(1-\alpha)^{L+1}(1-\varepsilon)^{L}\varepsilon+\alpha}{1-(1-\alpha)(1-\varepsilon)} & \text{otherwise} \end{cases} \end{array} . \tag{2.4}$$

Moreover, the stationary distribution $\{\eta(\theta)\}$ is reached within $(L+1)$ periods starting from any initial distribution.

Proof: Suppose that $\alpha > 0$ or $\varepsilon > 0$. Then (2.3) has a unique solution

$$\begin{array}{rcl} \eta(\theta) & = & (1-\alpha)^{\theta+1}(1-\varepsilon)^{\theta}\varepsilon, \; for \; 0 \leqslant \theta \leqslant L-1 \\ \eta(L) & = & \frac{(1-\alpha)^{L+1}(1-\varepsilon)^{L}\varepsilon+\alpha}{1-(1-\alpha)(1-\varepsilon)} \end{array} , \tag{2.5}$$

which satisfies $\sum_{\theta=0}^{L} \eta(\theta) = 1$. Suppose that $\alpha = 0$ and $\varepsilon = 0$. Then solving (2.3) together with $\sum_{\theta=0}^{L} \eta(\theta) = 1$ yields a unique solution $\eta(\theta) = 0$ for $0 \leqslant \theta \leqslant L-1$ and $\eta(L) = 1$. It is easy to see from the expressions in (2.2) that $\eta(\theta)$ is reached within $(\theta+1)$ periods, for all θ, starting from any initial distribution. ■

Since the coefficients in the equations that define a stationary distribution are independent of the recommended strategy that the agents follow, the stationary distribution is also independent of the recommended strategy, as can be seen in (2.4). Thus, we will write the stationary distribution as $\{\eta_L(\theta)\}$ to emphasize its dependence on the rating scheme, which is represented by L.

2.3.2 Sustainable Rating Protocols

We now investigate the incentive of agents to follow a prescribed recommended strategy. For simplicity, we check the incentive of agents at the stationary distribution of rating scores, as in [27] and [31]. Since we consider a non-cooperative scenario, we need to check whether an agent can improve its long-term payoff by a unilateral deviation. Note that any unilateral deviation from an individual agent would not affect the evolution of rating scores and thus the stationary distribution, because we consider a continuum of agents [5].

Let $c_\sigma(\theta, \tilde{\theta})$ be the cost suffered by a server with rating score θ that is matched with a client with rating score $\tilde{\theta}$ and follows a recommended strategy σ, i.e., $c_\sigma(\theta, \tilde{\theta}) = c$ if $\sigma(\theta, \tilde{\theta}) = F$ and $c_\sigma(\theta, \tilde{\theta}) = 0$ if $\sigma(\theta, \tilde{\theta}) = D$. Similarly, let $b_\sigma(\theta, \tilde{\theta})$ be the benefit received by a client with rating score $\tilde{\theta}$ that is matched with a server with rating score θ following a recommended strategy σ, i.e., $b_\sigma(\theta, \tilde{\theta}) = b$ if $\sigma(\theta, \tilde{\theta}) = F$ and $b_\sigma(\theta, \tilde{\theta}) = 0$ if $\sigma(\theta, \tilde{\theta}) = D$. Since we consider uniform random matching, the expected period payoff of a θ-agent under rating protocol κ before it is matched is given by

$$v_\kappa(\theta) = \sum_{\tilde{\theta}\in\Theta} \eta_L(\tilde{\theta}) b_\sigma(\tilde{\theta}, \theta) - \sum_{\tilde{\theta}\in\Theta} \eta_L(\tilde{\theta}) c_\sigma(\tilde{\theta}, \theta). \tag{2.6}$$

[5] This is true for any deviation by agents of measure zero.

To evaluate the long-term payoff of an agent, we use the discounted sum criterion in which the long-term payoff of an agent is given by the expected value of the sum of discounted period payoffs from the current period. Let $p_\kappa(\theta'|\theta)$ be the transition probability that a θ-agent becomes a θ'-agent in the next period under rating protocol κ. Under MPRS, $p_\kappa(\theta'|\theta)$ can be expressed as

$$p_\kappa(\theta'|\theta) = \begin{cases} 1-\varepsilon & \text{if } \theta' = \min\{\theta+1, L\} \\ \varepsilon & \text{if} \theta' = 0 \\ 0 & \text{otherwise} \end{cases} . \tag{2.7}$$

Then we can compute the long-term payoff of an agent from the current period (before it is matched) by solving the following recursive equations

$$v_\kappa^\infty(\theta) = v_\kappa(\theta) + \delta \sum_{\theta' \in \Theta} p_\kappa(\theta'|\theta) v_\kappa^\infty(\theta') \forall \theta \in \Theta, \tag{2.8}$$

where $\delta = \beta(1-\alpha)$ is the weight that an agent puts on its future payoff. Because an agent leaves the community with probability α at the end of the current period, the expected future payoff of a θ-agent is given by $(1-\alpha)\sum_{\theta' \in \Theta} p_\kappa(\theta'|\theta) v_\kappa^\infty(\theta')$, assuming that an agent receives zero payoff once it leaves the community. The expected future payoff is multiplied by a common discount factor $\beta \in [0,1)$, which reflects the time preference, or patience, of agents.

Now suppose that an agent deviates and uses a strategy σ' under rating protocol κ. Because the deviation of a single agent does not affect the stationary distribution, the expected period payoff of a deviating θ-agent is given by

$$v_{\kappa,\sigma'}(\theta) = \sum_{\tilde{\theta} \in \Theta} \eta_L(\tilde{\theta}) b_\sigma(\tilde{\theta}, \theta) + \sum_{\tilde{\theta} \in \Theta} \eta_L(\tilde{\theta}) c_{\sigma'}(\theta, \tilde{\theta}). \tag{2.9}$$

Let $p_{\kappa,\sigma'}(\theta'|\theta, \tilde{\theta})$ be the transition probability that a θ-agent using the strategy σ' becomes a θ'-agent in the next period under rating protocol κ, when it is matched with a client with rating score $\tilde{\theta}$. For each θ, $\theta' = \min\{\theta+1, L\}$ with probability $(1-\varepsilon)$ and $\theta' = 0$ with probability ε if $\sigma(\theta, \tilde{\theta}) = \sigma'(\theta, \tilde{\theta})$ while the probabilities are reversed otherwise. Then $p_{\kappa,\sigma'}(\theta'|\theta) = \sum_{\tilde{\theta} \in \Theta} \eta_L(\tilde{\theta}) p_{\kappa,\sigma'}(\theta'|\theta, \tilde{\theta})$ gives the transition probability of a θ-agent before knowing the rating score of its client, and the long-term payoff of a deviating agent from the current period (before it is matched) can be computed by solving

$$v^{\infty}_{\kappa,\sigma'}(\theta) = v_{\kappa,\sigma'}(\theta) + \delta \sum_{\theta' \in \Theta} p_{\kappa,\sigma'}(\theta'|\theta) v^{\infty}_{\kappa,\sigma'}(\theta') \forall \theta \in \Theta. \tag{2.10}$$

In our model, a server decides whether to provide a service or not after it is matched with a client and observes the rating score of the client. Hence, we check the incentive for a server to follow a recommended strategy at the point when it knows the rating score of the client. Suppose that a server with rating score θ is matched with a client with rating score $\tilde{\theta}$. When the server follows the recommended strategy σ prescribed by rating protocol κ, it receives the long-term payoff $-c_\sigma(\theta,\tilde{\theta}) + \delta \sum_{\theta'} p_\kappa(\theta'|\theta) v^{\infty}_{\kappa}(\theta')$, excluding the possible benefit as a client in the current period. On the contrary, when the server deviates to a recommended strategy σ', it receives the long-term payoff $-c_{\sigma'}(\theta,\tilde{\theta}) + \delta \sum_{\theta'} p_{\kappa,\sigma'}(\theta'|\theta,\tilde{\theta}) v^{\infty}_{\kappa,\sigma'}(\theta')$, again excluding the possible benefit as a client. By comparing these two payoffs, we can check whether a θ-agent has an incentive to deviate to σ' when it is matched with a client with rating score $\tilde{\theta}$.

Definition 2 (Sustainable rating protocols). A rating protocol κ is *sustainable* if

$$-c_\sigma(\theta,\tilde{\theta}) + \delta \sum_{\theta'} p_\kappa(\theta'|\theta) v^{\infty}_{\kappa}(\theta') \geqslant -c_{\sigma'}(\theta,\tilde{\theta}) + \delta \sum_{\theta'} p_{\kappa,\sigma'}(\theta'|\theta,\tilde{\theta}) v^{\infty}_{\kappa,\sigma'}(\theta') \tag{2.11}$$

for all σ', for all $(\theta,\tilde{\theta})$.

In words, a rating protocol $\kappa = (L,\sigma)$ is sustainable if no agent can gain from a unilateral deviation regardless of the rating score of the client it is matched with when every other agent follows the recommended strategy σ and the rating scores are determined by the MPRS with punishment length L. Thus, under a sustainable rating protocol, agents follow the prescribed recommended strategy in their self-interest. Checking whether a rating protocol is sustainable using the above definition requires computing deviation gains from all possible recommended strategies, whose computation complexity can be quite high for moderate values of L. By employing the criterion of unimprovability in Markov decision theory [38], we establish the one-shot deviation principle for sustainable rating protocols, which provides simpler conditions. For notation, let c_a be the cost suffered by a server that takes action a, and let $p_{\kappa,a}(\theta'|\theta,\tilde{\theta})$ be the transition probability that a θ-agent becomes a θ'-agent in the next period under rating protocol κ when it takes action a to a client with rating score $\tilde{\theta}$. The values of $p_{\kappa,a}(\theta'|\theta,\tilde{\theta})$ can be obtained in a similar way to obtain $p_{\kappa,\sigma'}(\theta'|\theta,\tilde{\theta})$, by comparing a with $\sigma(\theta,\tilde{\theta})$.

Lemma 2 (One-shot Deviation Principle). A rating protocol κ is sustainable if and only if

$$c_\sigma(\theta,\tilde{\theta}) - c_a \leqslant \delta \left[\sum_{\theta'} \{p_\kappa(\theta'|\theta) - p_{\kappa,a}(\theta'|\theta,\tilde{\theta})\} v_\kappa^\infty(\theta') \right] \tag{2.12}$$

for all $a \neq \sigma(\theta,\tilde{\theta})$, for all $(\theta,\tilde{\theta})$.

Proof: If rating protocol κ is sustainable, then clearly there are no profitable one-shot deviations. We can prove the converse by showing that, if κ is not sustainable, there is at least one profitable one-shot deviation. Since $c_\sigma(\theta,\tilde{\theta})$ and c_a are bounded, this is true by the unimprovability property in Markov decision theory [33][34]. ■

Lemma 2 shows that if an agent cannot gain by unilaterally deviating from σ only in the current period and following σ afterwards, it cannot gain by switching to any other recommended strategy σ' either, and vice versa. The left-hand side of (2.12) can be interpreted as the current gain from choosing a, while the right-hand side of (2.12) represents the discounted expected future loss due to the different transition probabilities induced by choosing a. Using the one-shot deviation principle, we can derive incentive constraints that characterize sustainable rating protocols.

First, consider a pair of rating scores $(\theta,\tilde{\theta})$ such that $\sigma(\theta,\tilde{\theta}) = F$. If the server with rating score θ serves the client, it suffers the service cost of c in the current period while its rating score in the next period becomes $\min\{\theta+1, L\}$ with probability $(1-\varepsilon)$ and 0 with probability ε. Thus, the expected long-term payoff of a θ-agent when it provides a service is given by

$$V_\theta(F;F) = -c + \delta[(1-\varepsilon)v_\kappa^\infty(\min\{\theta+1, L\}) + \varepsilon v_\kappa^\infty(0)]. \tag{2.13}$$

On the contrary, if a θ-agent deviates and declines the service request, it avoids the cost of c in the current period while its rating score in the next period becomes 0 with probability $(1-\varepsilon)$ and $\min\{\theta+1, L\}$ with probability ε. The expected long-term payoff of a θ-agent when it does not provide a service is given by

$$V_\theta(D;F) = \delta[(1-\varepsilon)v_\kappa^\infty(0) + \varepsilon v_\kappa^\infty(\min\{\theta+1, L\})]. \tag{2.14}$$

The incentive constraint that a θ-agent does not gain from a one-shot deviation is given by $V_\theta(F;F) \geqslant V_\theta(D;F)$, which can be expressed as

$$\delta(1-2\varepsilon)[v_\kappa^\infty(\min\{\theta+1, L\}) - v_\kappa^\infty(0)] \geqslant c. \tag{2.15}$$

Now, consider a pair of rating scores $(\theta, \tilde{\theta})$ such that $\sigma(\theta, \tilde{\theta}) = D$. Using a similar argument as above, we can show that the incentive constraint that a θ-agent does not gain from a one-shot deviation can be expressed as

$$\delta(1-2\varepsilon)[v_\kappa^\infty(\min\{\theta+1, L\}) - v_\kappa^\infty(0)] \geqslant -c. \tag{2.16}$$

Note that (2.15) implies (2.16), and thus for θ such that $\sigma(\theta, \tilde{\theta}) = F$ for some $\tilde{\theta}$, we only have to check the first incentive constraint (2.15). Therefore, a rating protocol κ is sustainable if and only if (2.15) holds for all θ such that $\sigma(\theta, \tilde{\theta}) = F$ for some $\tilde{\theta}$ and (2.16) holds for all θ such that $\sigma(\theta, \tilde{\theta}) = D$ for all $\tilde{\theta}$. The left-hand side of the incentive constraints (2.15) and (2.16) can be interpreted as the loss from punishment that rating protocol κ applies to a θ-agent for not following the recommended strategy. In order to induce a θ-agent to provide a service to some clients, the left-hand side should be at least as large as the service cost c, which can be interpreted as the deviation gain. We use $\min_{\theta\in\Theta}\{\delta(1-2\varepsilon)[v_\kappa^\infty(\min\{\theta+1, L\}) - v_\kappa^\infty(0)]\}$ to measure the strength of the incentive for cooperation under the rating protocol κ, where cooperation means providing the requested service in our context.

2.3.3 Rating Protocol Design Problem

Since we assume that the community operates at the stationary distribution of rating scores, social welfare under rating protocol κ can be computed by

$$U_\kappa = \sum_\theta \eta_L(\theta) v_\kappa(\theta). \tag{2.17}$$

The community operator aims to choose a rating protocol that maximizes social welfare among sustainable rating protocols. Then the problem of designing a rating protocol can be formally expressed as

$$\begin{array}{ll} \max\limits_{(L,\sigma)} & U_\kappa = \sum\limits_\theta \eta_L(\theta) v_\kappa(\theta) \\ \text{subject to} & \delta(1-2\varepsilon)[v_\kappa^\infty(\min\{\theta+1, L\}) - v_\kappa^\infty(0)] \geqslant c, \ \ \exists \tilde{\theta} \ \text{ with } \sigma(\theta, \tilde{\theta}) = F \\ & \delta(1-2\varepsilon)[v_\kappa^\infty(\min\{\theta+1, L\}) - v_\kappa^\infty(0)] \geqslant -c, \ \ \sigma(\theta, \tilde{\theta}) = D, \text{ for all } \tilde{\theta} \end{array} \ . \tag{2.18}$$

A rating protocol that solves the design problem (2.18) is called an *optimal rating protocol*.

2.4 Analysis of Optimal Rating Protocols

2.4.1 Optimal Value of the Design Problem

We first investigate whether there exists a sustainable rating protocol, i.e., whether the design problem (2.18s) has a feasible solution. Fix the punishment length L and consider a recommended strategy σ_L^D defined by $\sigma_L^D(\theta, \tilde{\theta}) = D$ for all $(\theta, \tilde{\theta})$. Since there is no service provided in the community when all the agents follow σ_L^D, we have

$$v^\infty_{(L,\sigma_L^D)}(\theta) = 0$$

for all θ. Hence, the relevant incentive constraint (2.16) is satisfied for all θ, and the rating protocol (L, σ_L^D) is sustainable. This shows that the design problem (2.18) always has a feasible solution.

Assuming that an optimal rating protocol exists, let U^* be the optimal value of the design problem (2.18). In the following proposition, we study the properties of U^*.

Proposition 1. The optimal value of the design problem (2.18) satisfies the following properties:

(i) $0 \leqslant U^* \leqslant b - \frac{1-\varepsilon}{1-2\varepsilon}c$.

(ii) $U^* = 0$ if $\frac{c}{b} > \frac{\beta(1-\alpha)(1-2\varepsilon)}{1-\beta(1-\alpha)(2-3\varepsilon)}$.

(iii) $U^* \geqslant [1-(1-\alpha)\varepsilon](b-c)$ if $\frac{c}{b} \leqslant \beta(1-\alpha)(1-2\varepsilon)$.

(iv) $U^* = b - c$ if $\varepsilon = 0$ and $\frac{c}{b} \leqslant \beta(1-\alpha)$.

(v) $U^* = b - c$ only if $\varepsilon = 0$ and $\frac{c}{b} \leqslant \frac{\beta(1-\alpha)}{1-\beta(1-\alpha)}$.

Proof: See Appendix A. ■

Proposition 1(i) proves that the optimal social welfare cannot be negative but is always strictly bounded away from $b - c$, which is the social welfare when all agents cooperates, when $\varepsilon > 0$. Hence full cooperation cannot be achieved in this scenario. Since we obtain zero social welfare at myopic equilibrium, without using a rating protocol, we are interested in whether we can sustain a rating protocol in which agents cooperate in a positive proportion of matches. In other words, we look for conditions on the parameters $(b, c, \beta, \alpha, \varepsilon)$ that yield $U^* > 0$. From Proposition 1(ii) and (iii), we can regard $c/b \leqslant [\beta(1-\alpha)(1-2\varepsilon)]/[1-\beta(1-\alpha)(2-3\varepsilon)]$ and $c/b \leqslant \beta(1-\alpha)(1-2\varepsilon)$ as necessary and sufficient conditions for $U^* > 0$, respectively. Moreover, when there are no report errors (i.e., $\varepsilon = 0$), we can interpret $c/b \leqslant \beta(1-\alpha)/[1-\beta(1-\alpha)]$ and $c/b \leqslant \beta(1-\alpha)$ as necessary and sufficient conditions to achieve the maximum social welfare $U^* = b - c$, respectively. As a corollary of Proposition 1, we obtain the following results in the limit.

Corollary 1. For any (b, c) such that $b > c$, (i) U^* converges to $b - c$ as $\beta \to 1$, $\alpha \to 0$, and $\varepsilon \to 0$, and (ii) U^* converges to 0 as $\beta \to 0$, $\alpha \to 1$, or $\varepsilon \to 1/2$. ■

Corollary 1 shows that we can design a sustainable rating protocol that achieves near efficiency (i.e., U^* close to $b - c$) when the community conditions are good (i.e.,

β is close to 1, and α and ε are close to 0). Moreover, it suffices to use only two ratings (i.e., $L = 1$) for the design of such a rating protocol. On the contrary, no cooperation can be sustained (i.e., $U^* = 0$) when the community conditions are bad (i.e., β is close to 0, α is close to 1, or ε is close to 1/2), as implied by Proposition 1(ii).

2.4.2 Optimal Recommended Strategies Given a Punishment Length

In order to obtain analytical results, we consider the design problem (2.18) with a fixed punishment length L, denoted DP_L. Note that DP_L has a feasible solution, namely σ_L^D, for any L and that there are a finite number (total $2^{(L+1)^2}$) of possible recommended strategies given L. Therefore, DP_L has an optimal solution for any L. We use U_L^* and σ_L^* to denote the optimal value and the optimal recommended strategy of DP_L, respectively. We first show that increasing the punishment length cannot decrease the optimal value.

Proposition 2. $U_L^* \geqslant U_{L'}^*$ for all L and L' such that $L \geqslant L'$.

Proof: See Appendix B. ■

Proposition 2 shows that U_L^* is non-decreasing in L. Since $U_L^* < b - c$ when $\varepsilon > 0$, we have $U^* = \lim_{L\to\infty} U_L^* = \sup_L U_L^*$. It may be the case that the incentive constraints eventually prevent the optimal value from increasing with L so that the supremum is attained by some finite L. This conjecture is verified in Figure 2.2, where U_L^* stops increasing when $L \geqslant 5$. Hence, it is plausible for the protocol designer to set an upper bound on L in practical designs with little efficiency loss incurred. Now we analyze the structure of optimal recommended strategies given a punishment length. The properties characterized in the following proposition can effectively reduce the design space of the optimal recommended strategy given L and thus reduces the computation complexity of the optimal rating protocol design.

Proposition 3. If we have that $\varepsilon > 0$ and $\alpha < 1$, the optimal rating protocol exhibits the following structures:

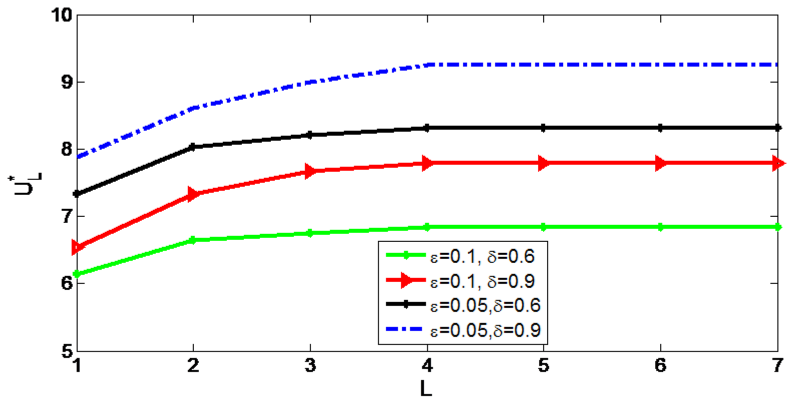


Figure 2.2: Optimal performance giving the punishment length L.

(i) A 0-agent does not receive service from some agents, i.e., $\sigma_L^*(\theta, 0) = D$, $\exists \theta \in \Theta$.

(ii) If $\sigma_L^*(0, \hat{\theta}) = F$ for some $\hat{\theta}$, then agents with sufficiently high rating scores always receive service from 0-agents, i.e., $\sigma_L^*(0, \tilde{\theta}) = F$ for all $\tilde{\theta} \geqslant \min\{\ln \frac{c}{b} / \ln \beta, L\}$.

(iii) L- agents receive service from other agents whose rating scores are sufficiently high, i.e., if $\theta \in \{1, \ldots, L-1\}$ satisfies $\theta \geqslant L - (\ln \frac{c}{b} - \ln Y(\alpha, \varepsilon, L)) / \ln \beta$, where

$$Y(\alpha, \varepsilon, L) = \frac{(1-\alpha)^{L+1}(1-\varepsilon)^{L}\varepsilon - (1-\alpha)^{L+2}(1-\varepsilon)^{L+1}\varepsilon}{(1-\alpha)^{L+1}(1-\varepsilon)^{L}\varepsilon + \alpha}, \tag{2.19}$$

then $\sigma_L^*(\theta, L) = F$.

(iv) L-agents always provide service to other L-agents, i.e., $\sigma_L^*(L, L) = F$.

Proof: See Appendix C. ■

As Proposition 3 shows, to construct an optimal rating protocol, sufficient punishment should be provided to agents with low rating scores while sufficient rewards should be provided to agents with high rating scores.

2.4.3 Illustration with $L = 1$ and $L = 2$

We can represent a recommended strategy σ_L as an $(L+1) \times (L+1)$ matrix whose (i, j)-entry is given by $\sigma_L(i-1, j-1)$. Proposition 3 provides some structures of an optimal recommended strategy σ_L^* in the first column and the last row of the matrix representation, but it does not fully characterize the solution of DP_L. Here we aim to obtain the solution of DP_L for $L = 1$ and 2 and analyze how it changes with the parameters. We first begin with the case of two ratings, i.e., $L = 1$. In this case, if $\sigma_1(\theta, \tilde{\theta}) = F$ for some $(\theta, \tilde{\theta})$, the relevant incentive constraint to sustain $\kappa = (1, \sigma_1)$ is $\delta(1-2\varepsilon)[v_\kappa^\infty(1) - v_\kappa^\infty(0)] \geqslant c$. By Proposition 3(ii) and (iv), if $\sigma_1^*(\theta, \tilde{\theta}) = F$ for some $(\theta, \tilde{\theta})$, then $\sigma_1^*(0,1) = \sigma_1^*(1,1) = F$, provided that $\varepsilon > 0$ and $\alpha < 1$. Hence, among the total of 16 possible recommended strategies, only four can be optimal recommended strategies. These four recommended strategies are

$$\sigma_1^1 = \begin{bmatrix} D & F \\ F & F \end{bmatrix}, \ \sigma_1^2 = \begin{bmatrix} F & F \\ D & F \end{bmatrix}, \ \sigma_1^3 = \begin{bmatrix} D & F \\ D & F \end{bmatrix}, \ \sigma_1^4 = \sigma_1^D = \begin{bmatrix} D & D \\ D & D \end{bmatrix}. \quad (2.20)$$

For notational convenience, we define a recommended strategy σ_L^{D0} by $\sigma_L^{D0}(\theta, 0) = D$ for all θ and $\sigma_L^{D0}(\theta, \tilde{\theta}) = F$ for all θ and all $\tilde{\theta} > 0$. In (2.20), we have $\sigma_1^3 = \sigma_1^{D0}$. The following proposition specifies the optimal recommended strategy given the parameters.

Proposition 4. Suppose that $0 < (1-\alpha)\varepsilon < 1/2$. Then

$$\sigma_1^* = \begin{cases} \sigma_1^1 & \text{if } 0 < \frac{c}{b} \leqslant \frac{\beta(1-\alpha)^2(1-2\varepsilon)\varepsilon}{1+\beta(1-\alpha)^2(1-2\varepsilon)\varepsilon} \\ \sigma_1^2 & \text{if } \frac{\beta(1-\alpha)^2(1-2\varepsilon)\varepsilon}{1+\beta(1-\alpha)^2(1-2\varepsilon)\varepsilon} < \frac{c}{b} \leqslant \frac{\beta(1-\alpha)(1-2\varepsilon)[1-(1-\alpha)\varepsilon]}{1-\beta(1-\alpha)^2(1-2\varepsilon)\varepsilon} \\ \sigma_1^3 & \text{if } \frac{\beta(1-\alpha)(1-2\varepsilon)[1-(1-\alpha)\varepsilon]}{1-\beta(1-\alpha)^2(1-2\varepsilon)\varepsilon} < \frac{c}{b} \leqslant \beta(1-\alpha)(1-2\varepsilon) \\ \sigma_1^4 & \text{if } \beta(1-\alpha)(1-2\varepsilon) < \frac{c}{b} < 1 \end{cases}. \quad (2.21)$$

Proof: Let $\kappa^i = (1, \sigma_1^i)$, for $i = 1, 2, 3, 4$. We obtain that

$$\begin{aligned} U_{\kappa^1} &= (1-\eta_1(0)^2)(b-c), \quad & U_{\kappa^2} &= (1-\eta_1(0)\eta_1(1))(b-c) \\ U_{\kappa^3} &= (1-\eta_1(0))(b-c), \quad & U_{\kappa^4} &= 0 \end{aligned}. \quad (2.22)$$

Since $0 < (1-\alpha)\varepsilon < 1/2$, we have $\eta_1(0) < \eta_1(1)$. Thus, we have $U_{\kappa^1} > U_{\kappa^2} > U_{\kappa^3} >$

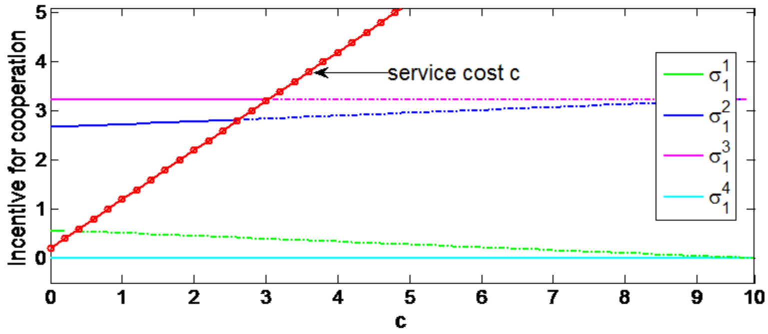
U_{κ^4}. Also, we obtain that

$$\begin{array}{rclrcl} v_{\kappa^1}^{\infty}(1) - v_{\kappa^1}^{\infty}(0) & = & \eta_1(0)(b-c), & v_{\kappa^2}^{\infty}(1) - v_{\kappa^2}^{\infty}(0) & = & b - \eta_1(0)(b-c) \\ v_{\kappa^3}^{\infty}(1) - v_{\kappa^3}^{\infty}(0) & = & b, & v_{\kappa^4}^{\infty}(1) - v_{\kappa^4}^{\infty}(0) & = & 0 \end{array}. \tag{2.23}$$

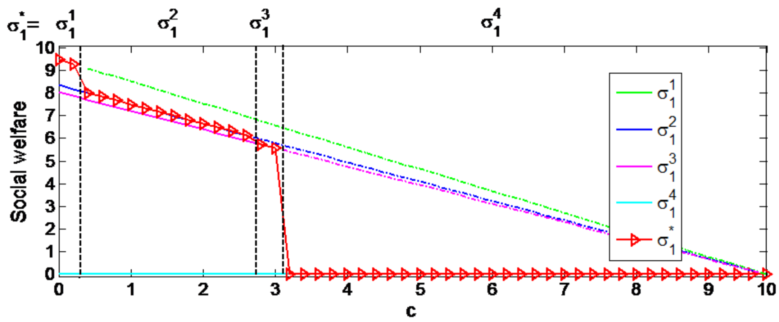
Thus, we have $v_{\kappa^3}^{\infty}(1) - v_{\kappa^3}^{\infty}(0) > v_{\kappa^2}^{\infty}(1) - v_{\kappa^2}^{\infty}(0) > v_{\kappa^1}^{\infty}(1) - v_{\kappa^1}^{\infty}(0) > v_{\kappa^4}^{\infty}(1) - v_{\kappa^4}^{\infty}(0)$. By choosing the recommended strategy that yields the highest social welfare among sustainable ones, we obtain the result. ■

Proposition 4 shows that the optimal recommended strategy is determined by the service cost-to-benefit ratio c/b. When c/b is sufficiently small, the recommended strategy σ_1^1 can be sustained, yielding the highest social welfare among the four candidate recommended strategies. As c/b increases, the optimal recommended strategy changes from σ_1^1 to σ_1^2 to σ_1^3 and eventually to σ_1^4. Figure 2.3 shows the optimal recommended strategies with $L = 1$ as c varies. The parameters we use to obtain the results in the figures of this chapter are set as follows unless otherwise stated: $\beta = 0.8$, $\alpha = 0.1$, $\varepsilon = 0.2$, and $b = 10$. Figure 2.3(a) plots the incentive for cooperation of the four recommended strategies. We can find the region of c in which each strategy is sustained by comparing the incentive for cooperation with the service cost c for σ_1^1, σ_1^2, and σ_1^3, and with $-c$ for σ_1^4. The solid portion of the lines indicates that the strategy is sustained while the dashed portion indicates that the strategy is not sustained. Figure 2.3(b) plots the social welfare of the four candidate strategies, with solid and dashed portions having the same meanings. The triangle-marked line represents the optimal value, which takes the maximum of the social welfare of all sustained strategies.

Next, we analyze the case of three ratings, i.e., $L = 2$. In order to provide a partial characterization of the optimal recommended strategy σ_2^*, we introduce the following notation. Let $\sigma_2^{\#}$ be the recommended strategy with $L = 2$ that maximizes $\min\{v_{\kappa}^{\infty}(1) - v_{\kappa}^{\infty}(0), v_{\kappa}^{\infty}(2) - v_{\kappa}^{\infty}(0)\}$ among all the recommended strategies with $L = 2$. Let $\gamma \triangleq \delta(1-\varepsilon)$ as defined in Appendix A, and define a recommended strategy σ_L^B by $\sigma_L^B(L-1, 0) = D$ and $\sigma_L^B(\theta, \tilde{\theta}) = F$ for all $(\theta, \tilde{\theta}) \neq (L-1, 0)$. We have



(a) Incentive for cooperation



(b) Social welfare and the optimal recommended strategy

Figure 2.3: Performance of the four candidate recommended strategies when $L = 1$.

the following conclusion about σ_2^* and $\sigma_2^\#$.

Proposition 5. Suppose that $\varepsilon > 0$, $\alpha < 1$, and

$$\frac{c}{b} < \frac{\eta_2(2)}{\eta_2(1)} \frac{1-\gamma}{\gamma} < \frac{b}{c}, \tag{2.24}$$

(i) $\sigma_2^\# = \sigma_2^{D0}$; (ii) if $\eta_2(0) < \eta_2(2)$, then $\sigma_2^* = \sigma_2^B$.

Proof: (i) Let $\kappa = (2, \sigma_2^{D0})$. Then $v_\kappa^\infty(1) - v_\kappa^\infty(0) = v_\kappa^\infty(2) - v_\kappa^\infty(0) = b$. We can show that, under the given conditions, any change from σ_2^{D0} results in a decrease in the value of $v_\kappa^\infty(1) - v_\kappa^\infty(0)$, which proves that σ_2^{D0} maximizes $\min\{v_\kappa^\infty(1) - v_\kappa^\infty(0), v_\kappa^\infty(2) - v_\kappa^\infty(0)\}$.

(ii) Since $\varepsilon > 0$ and $\alpha < 1$, we have $\eta_2(\theta) > 0$ for all $\theta = 0, 1, 2$, and thus replacing D with F in an element of a recommended strategy always improves social welfare. Hence, we first consider the recommended strategy σ_L^F defined by $\sigma_L^F(\theta, \tilde{\theta}) = F$ for all $(\theta, \tilde{\theta})$. σ_2^F maximizes social welfare U_κ among all the recommended strategies with $L = 2$, but $v_\kappa^\infty(1) - v_\kappa^\infty(0) = v_\kappa^\infty(2) - v_\kappa^\infty(0) = 0$. Thus, we cannot find parameters such that σ_2^F satisfies the incentive constraints, and thus $\sigma_2^F \neq \sigma_2^*$. Now consider recommended strategies in which there is exactly one D element. We can show that, under the given conditions, having $\sigma_2(\theta, \tilde{\theta}) = D$ at $(\theta, \tilde{\theta})$ such that $\tilde{\theta} > 0$ yields $v_\kappa^\infty(1) - v_\kappa^\infty(0) < 0$, whereas having $\sigma_2(\theta, \tilde{\theta}) = D$ at $(\theta, \tilde{\theta})$ such that $\tilde{\theta} = 0$ yields both $v_\kappa^\infty(1) - v_\kappa^\infty(0) > 0$ and $v_\kappa^\infty(2) - v_\kappa^\infty(0) > 0$. Thus, for any recommended strategy having the only D element at $(\theta, \tilde{\theta})$ such that $\tilde{\theta} > 0$, there do not exist parameters in the considered parameter space with which the incentive constraint for 0-agents, $\delta(1 - 2\varepsilon)\left[v_\kappa^\infty(1) - v_\kappa^\infty(0)\right] \geqslant c$, is satisfied. On the other hand, for any recommended strategy having the only D element at $(\theta, \tilde{\theta})$ such that $\tilde{\theta} = 0$, we can satisfy both incentive constraints by choosing $\beta > 0$, $\alpha < 1$, $\varepsilon < 1/2$, and c sufficiently close to 0. This shows that, among the recommended strategies having exactly one D element, only those having D in the first column are possibly sustainable. S-ince $\eta_2(1) < \eta_2(0) < \eta_2(2)$, σ_2^B achieves the highest social welfare among the three candidate recommended strategies. ■

Let us try to better understand now what Proposition 5 means. Proposition 5(i) implies that the maximum incentive for cooperation that can be achieved with three ratings is $\beta(1-\alpha)(1-2\varepsilon)b$. Hence, cooperation can be sustained with $L = 2$ if and only if $\beta(1-\alpha)(1-2\varepsilon)b \geqslant c$. That is, if $c/b > \beta(1-\alpha)(1-2\varepsilon)$, then σ_2^D is the only sustainable recommended strategy and thus $U_2^* = 0$. Therefore, when we increase c while holding other parameters fixed, we can expect that σ_2^* changes from σ_2^{D0} to σ_2^D around $c = \beta(1-\alpha)(1-2\varepsilon)b$. Note that the same is observed with $L = 1$ in Proposition 4. We can see that $[\eta_\tau(2)/\eta_\tau(1)][(1-\gamma)/\gamma]$ converges to 1 as α goes to 0 and β goes to 1. Hence, for given values of b, c, and ε, the condition (2.24) is satisfied and thus some cooperation can be sustained if α and β are sufficiently close to 0 and 1, respectively.

Consider a rating protocol $\kappa = (2, \sigma_2^B)$. We obtain that

$$\min\{v_\kappa^\infty(1) - v_\kappa^\infty(0), v_\kappa^\infty(2) - v_\kappa^\infty(0)\} = v_\kappa^\infty(2) - v_\kappa^\infty(0) = (1-\alpha)^2(1-\varepsilon)\varepsilon(b - \beta c) \tag{2.25}$$

and $U_\kappa = (1-(1-\alpha)^3(1-\varepsilon)\varepsilon^2)(b-c)$. Proposition 5(ii) is stating that $\sigma_2^* = \sigma_2^B$ when the community conditions are "favorable". More precisely, we have $\sigma_2^* = \sigma_2^B$ if $(1-\alpha)^2(1-\varepsilon)\varepsilon(b-\beta c) \geqslant c$, or

$$\frac{c}{b} \leqslant \frac{\beta(1-\alpha)^3(1-2\varepsilon)(1-\varepsilon)\varepsilon}{1+\beta^2(1-\alpha)^3(1-2\varepsilon)(1-\varepsilon)\varepsilon}. \tag{2.26}$$

Also, Proposition 5(ii) implies that $U_2^* \leqslant (1-(1-\alpha)^3(1-\varepsilon)\varepsilon^2)(b-c)$ always holds.

In Figure 2.4, we show the optimal value and the optimal recommended strategy of DP_2 as we vary c. The optimal recommended strategy σ_2^* changes in the following order before becoming σ_2^D as c increases:

$$\sigma_2^1 = \begin{bmatrix} F & F & F \\ D & F & F \\ F & F & F \end{bmatrix}, \sigma_2^2 = \begin{bmatrix} D & F & F \\ F & F & F \\ F & F & F \end{bmatrix}, \sigma_2^3 = \begin{bmatrix} D & F & F \\ D & F & F \\ F & F & F \end{bmatrix},$$

$$\sigma_2^4 = \begin{bmatrix} F & F & F \\ F & F & F \\ D & F & F \end{bmatrix}, \sigma_2^5 = \begin{bmatrix} F & F & F \\ D & F & F \\ D & F & F \end{bmatrix}, \sigma_2^6 = \begin{bmatrix} D & F & F \\ F & F & F \\ D & F & F \end{bmatrix}, \sigma_2^7 = \begin{bmatrix} D & F & F \\ D & F & F \\ D & F & F \end{bmatrix}. \tag{2.27}$$

Note that $\sigma_2^1 = \sigma_2^B$ for small c and $\sigma_2^7 = \sigma_2^{D0}$ for large c (but not too large to sustain cooperation), which are consistent with the discussion about Proposition 5. For the intermediate values of c, only the elements in the first column change in order to increase the incentive for cooperation. We find that the order of the optimal recommended strategies between $\sigma_2^1 = \sigma_2^B$ and $\sigma_2^7 = \sigma_2^{D0}$ depends on the community's parameters $(b, c, \beta, \alpha, \varepsilon)$.

2.5 Extensions

2.5.1 Rating Schemes with Shorter Punishment Length

So far we have focused on MPRS under which any deviation in reported actions results in a rating score of 0. Although this class of rating schemes is simple in that a rating

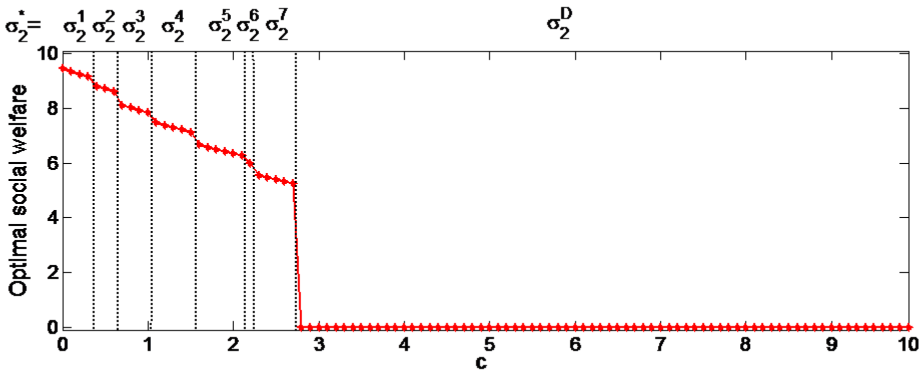


Figure 2.4: Optimal social welfare and the optimal recommended strategy of DP_2.

scheme can be identified with the number of rating scores, it may not yield the highest social welfare among all possible rating schemes when there are report errors. When there is no report error, i.e., $\varepsilon = 0$, an agent maintains rating score L as long as it follows the prescribed recommended strategy. Thus, in this case, punishment exists only as a threat and it does not result in an efficiency loss. On the contrary, when $\varepsilon > 0$ and $\alpha < 1$, there exist a positive proportion of agents with any rating from 0 to $L - 1$ in the stationary distribution even if all the agents follow the recommended strategy. Thus, there is a tension between efficiency and incentive. In order to sustain a rating protocol, we need to provide a strong punishment so that agents do not gain by deviation. At the same time, too severe a punishment reduces social welfare. This observation suggests that, in the presence of report errors, it is optimal to provide incentives just enough to prevent deviations. If we can provide a weaker punishment while sustaining the same recommended strategy, it will improve social welfare. One way to provide a weaker punishment is to use a random punishment. For example, we can consider a rating scheme under which the rating score of a θ-agent becomes 0 in the next period with probability $q_\theta \in (0, 1]$ and remains the same with probability $1 - q_\theta$ when it reportedly deviates from the recommended strategy. By varying the punishment probability q_θ for θ-agents, we can adjust the severity of the punishment applied to θ-agents. This class of rating schemes can be identified by $(L, \{q_\theta\})$. MPRS can be considered as a special case where $q_\theta = 1$ for all θ.

Another way to provide a weaker punishment is to use a smaller punishment length, denoted M. Under the rating scheme with $(L+1)$ rating scores and punishment length M, rating scores are updated by

$$\tau(\theta, \tilde{\theta}, a_R) = \begin{cases} \min\{\theta+1, L\} & \text{if } a_R = \sigma(\theta, \tilde{\theta}) \\ \max\{\theta - M, 0\} & \text{if } a_R \neq \sigma(\theta, \tilde{\theta}) \end{cases}. \tag{2.28}$$

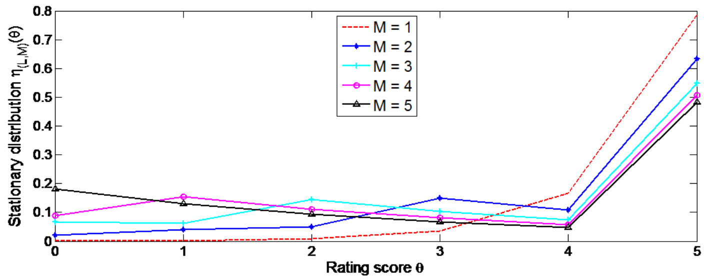
When a θ-agent reportedly deviates from the recommended strategy, its rating score is reduced by M in the next period if $\theta \geqslant M$ and becomes 0 otherwise. Note that this rating scheme is analogous to real-world rating schemes for credit rating and auto insurance risk rating. This class of rating schemes can be identified by (L, M) with $1 \leqslant M \leqslant L$ [6]. MPRS can be considered as a special case where $M = L$.

In this chapter, we focus on the second approach to investigate the impacts of the punishment length on the social welfare U_κ and the incentive for cooperation $\min\limits_{\theta}\{\delta(1-2\varepsilon)[v_\kappa^\infty(\min\{\theta+1, L\}) - v_\kappa^\infty(\max\{\theta - M, 0\})]\}$ of a rating protocol κ, which is now defined as (L, M, σ). The punishment length M affects the evolution of the rating distribution, and the stationary distribution of rating scores with the rating scheme (L, M), $\{\eta_{(L,M)}(\theta)\}_{\theta=0}^{L}$, satisfies the following equations:

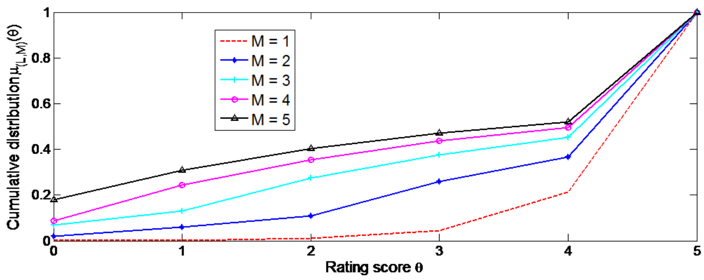
$$\begin{array}{l} \eta_{(L,M)}(0) = (1-\alpha)\varepsilon \sum\limits_{\theta=0}^{M} \eta_{(L,M)}(\theta) \\ \eta_{(L,M)}(\theta \in \{1, L-M\}) = (1-\alpha)(1-\varepsilon)\eta_{(L,M)}(\theta-1) + (1-\alpha)\varepsilon\eta_{(L,M)}(\theta+M) \\ \eta_{(L,M)}(\theta \in \{L-M+1, L-1\}) = (1-\alpha)(1-\varepsilon)\eta_{(L,M)}(\theta-1) \\ \eta_{(L,M)}(L) = (1-\alpha)(1-\varepsilon)\{\eta_{(L,M)}(L) + \eta_{(L,M)}(L-1)\} + \alpha \end{array}. \tag{2.29}$$

Let $\mu_{(L,M)}(\theta) = \sum\limits_{k=0}^{\theta} \eta_{(L,M)}(k)$ for $\theta = 0, \ldots, L$. Figure 2.5 plots the stationary distribution $\{\eta_{(L,M)}(\theta)\}_{\theta=1}^{L}$ and its cumulative distribution $\{\mu_{(L,M)}(\theta)\}_{\theta=1}^{L}$ for $L = 5$ and $M = 1, \ldots, 5$. We can see that the cumulative distribution monotonically decreases with M, i.e., $\mu_{(L,M_1)}(\theta) \leqslant \mu_{(L,M_2)}(\theta)$ for all θ if $M_1 > M_2$. This shows that, as

[6] We can further generalize this class by having the punishment length depend on the rating. That is, when a θ-agent reportedly deviates from the recommended strategy, its rating is reduced to $\theta - M_\theta$ in the next period for some $M_\theta \leqslant \theta$.



(a)



(b)

Figure 2.5: (a) Stationary distribution of rating scores and (b) the cumulative distribution when $L = 5$.

the punishment length increases, there are more agents holding a lower rating score. As a result, when the community adopts a recommended strategy that treats an agent with a higher rating score better, increasing the punishment length reduces social welfare while it increases the incentive for cooperation. This trade-off is illustrated in Figure 2.6, which plots social welfare and the incentive for cooperation under a rating protocol $(3, M, \sigma_3^C)$ for $M = 1, 2, 3$, where the recommended strategy σ_L^C is defined by $\sigma_L^C(\theta, \tilde{\theta}) = F$ if and only if $\tilde{\theta} \geqslant \theta$, for all θ.

In general, the recommended strategy adopted in the community is determined together with the rating scheme in order to maximize social welfare while satisfying the incentive constraints. The design problem with variable punishment lengths can be formulated as follows. First, note that the expected period payoff of a θ-agent, $v_\kappa(\theta)$, can be computed by (2.6), with the modification of the stationary distribution to $\{\eta_{(L,M)}(\theta)\}_{\theta=1}^{L}$. Agents' long-term payoffs can be obtained by solving (2.8), with the transition probabilities now given by

$$p_\kappa(\theta'|\theta) = \begin{cases} 1-\varepsilon & \text{if } \theta' = \min\{\theta+1, L\} \\ \varepsilon & \text{if } \theta' = \max\{\theta - M, 0\} \\ 0 & \text{otherwise} \end{cases}. \tag{2.30}$$

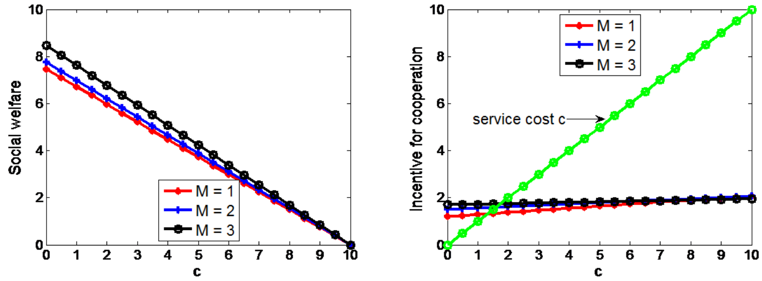
Finally, the design problem can be written as

$$\begin{aligned} \max_{(L,M,\sigma)} \quad & U_\kappa = \sum_\theta \eta_{(L,M)}(\theta) v_\kappa(\theta) \\ \text{subject to} \quad & \delta(1-2\varepsilon)[v_\kappa^\infty(\min\{\theta+1,L\}) - v_\kappa^\infty(\max\{\theta-M,0\})] \geqslant c \\ & \forall\theta \; such \; that \; \exists\tilde{\theta} \; such \; that \; \sigma(\theta,\tilde{\theta}) = F, \\ & \delta(1-2\varepsilon)[v_\kappa^\infty(\min\{\theta+1,L\}) - v_\kappa^\infty(\max\{\theta-M,0\})] \geqslant -c, \\ & \forall\theta \; such \; that \; \sigma(\theta,\tilde{\theta}) = D \; \forall\tilde{\theta} \end{aligned} . \tag{2.31}$$

We find the optimal recommended strategy given a rating scheme (L, M) for $L = 3$ and $M = 1, 2, 3$, and plot the social welfare and the incentive for cooperation of the optimal recommended strategies in Figure 2.7. Since different values of M induce different optimal recommended strategies given the value of L, there are no monotonic relationships between the punishment length and social welfare as well as the incentive for cooperation, unlike in Figure 2.6. The optimal punishment length given L can be obtained by taking the punishment length that yields the highest social welfare, which is plotted in Figure 2.8. We can see that, as the service cost c increases, the optimal punishment length increases from 1 to 2 to 3 before cooperation becomes no longer sustainable. This result is intuitive in that larger c requires a stronger incentive for cooperation, which can be achieved by having a larger punishment length.

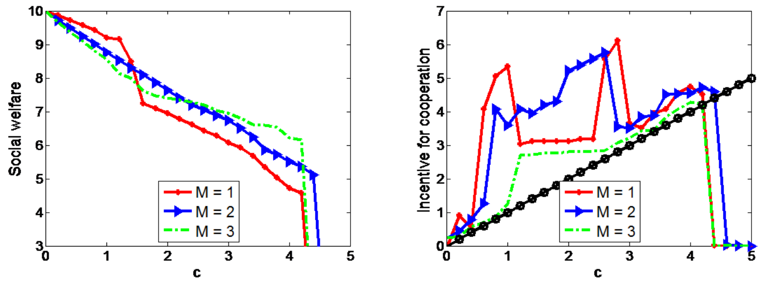
2.5.2 White-washing-Proof Rating Protocols

So far we have restricted our attention to rating schemes where newly joining agents are endowed with the highest rating score, i.e., $K = L$, without worrying about the possibility of whitewashing. We now make the initial rating score K as a choice variable of



(a) (b)

Figure 2.6: (a) Social welfare and (b) the incentive for cooperation under recommended strategy σ_L^C when $L = 3$.



(a) (b)

Figure 2.7: (a) Social welfare and (b) the incentive for cooperation under the optimal recommended strategy when $L = 3$.

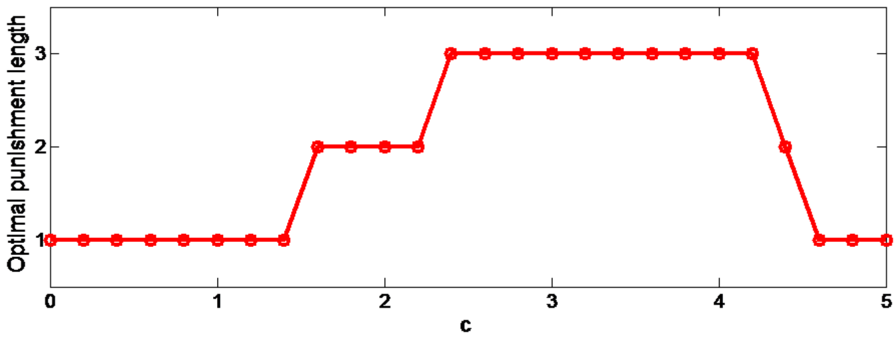


Figure 2.8: Optimal punishment length when $L = 3$.

the design problem while assuming that agents can whitewash their rating scores in order to obtain rating score K [17]. At the end of each period, agents can decide whether to whitewash their rating scores or not after observing their rating scores for the next period. If an agent chooses to whitewash its rating score, then it leaves and re-joins the community with α fraction of agents and receives initial rating score K. The cost of whitewashing is denoted by $c_w \geqslant 0$.

The incentive constraints in the design problem (2.18) are aimed at preventing agents from deviating from the prescribed recommended strategy. In the presence of potential whitewashing attempts, we need additional incentive constraints to prevent agents from whitewashing their rating scores. A rating protocol κ is whitewash-proof if and only if $v_\kappa^\infty(K) - v_\kappa^\infty(\theta) \leqslant c_w$ for all $\theta = 0, \ldots, L$ [7]. Note that $v_\kappa^\infty(K) - v_\kappa^\infty(\theta)$ is the gain from whitewashing for an agent whose rating score is updated as θ. If $v_\kappa^\infty(K) - v_\kappa^\infty(\theta) \leqslant c_w$, there is no net gain from whitewashing for a θ-agent. We measure the incentive for whitewashing under a rating protocol κ by $\max_{\theta \in \Theta}\{v_\kappa^\infty(K) - v_\kappa^\infty(\theta)\}$. A rating protocol is more effective in preventing whitewashing as the incentive for whitewashing is smaller.

To simplify our analysis, we fix the punishment length at $M = L$ so that a rating scheme is represented by (L, K) with $0 \leqslant K \leqslant L$. Let $\{\eta_{(L,K)}(\theta)\}_{\theta=1}^{L}$ be the stationary distribution of rating scores under rating scheme (L, K). Then the design problem is modified as follows (it should be noted here that similar to Section 5.A, both $\{v_\kappa(\theta)\}$ and $\{v_\kappa^\infty(\theta)\}$ in this section are computed using a stationary distribution different than

[7] This condition assumes that an agent can whitewash its rating only once in its lifespan in the community. More generally, we can consider the case where an agent can whitewash its rating multiple times. For example, an agent can use a deterministic stationary decision rule for whitewashing, which can be represented by a function $w : \Theta \to \{0, 1\}$, where $w(\theta) = 1$ (resp. $w(\theta) = 0$) means that the agent whitewashes (resp. does not whitewash) its rating if it holds rating θ in the next period. This will yield a different expression for the gain from whitewashing.

(2.4), which depends on the value of K)

$$
\begin{array}{ll}
\max\limits_{(L,K,\sigma)} & U_\kappa = \sum\limits_\theta \eta_{(L,K)}(\theta) v_\kappa(\theta) \\
\text{subject to} & \delta(1-2\varepsilon)[v_\kappa^\infty(\min\{\theta+1,L\}) - v_\kappa^\infty(0)] \geqslant c,\ \ \exists\tilde{\theta} \text{ with } \sigma(\theta,\tilde{\theta}) = F \\
& \delta(1-2\varepsilon)[v_\kappa^\infty(\min\{\theta+1,L\}) - v_\kappa^\infty(0)] \geqslant -c,\ \ \sigma(\theta,\tilde{\theta}) = D, \text{ for all } \tilde{\theta} \\
& v_\kappa^\infty(K) - v_\kappa^\infty(\theta) \leqslant c_w,\ \forall\theta
\end{array}
. \quad (2.32)
$$

Now an optimal rating protocol is the one that maximizes social welfare among sustainable and whitewash-proof rating protocols. Note that the design problem (2.32) always has a feasible solution for any $c_w \geqslant 0$ since (L, K, σ_L^D) is sustainable and whitewash-proof for all (L, K). Nevertheless, (L, K, σ_L^D) is trivial since no service takes place in the community as a consequence. Next, we show that given the existence of sustainable rating protocols which deliver a positive level of cooperation, i.e., when U^* solved by (2.18) is positive, whitewash-proof rating protocols also exist.

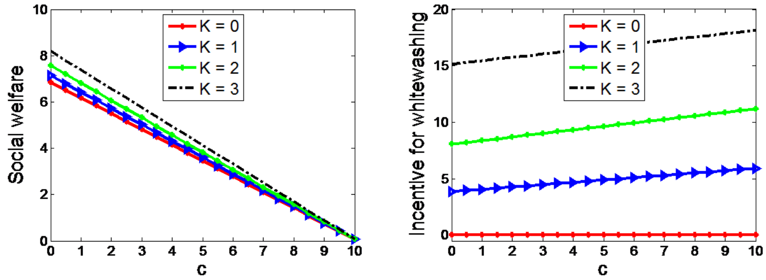
Lemma 3. If a rating protocol (L, K, σ) is sustainable, then the rating protocol $(L, 0, \sigma)$ is also sustainable. ■

Lemma 3 shows that it never reduces agents' incentive of cooperation by assigning the newly joined agents the lowest ratings. With this result, we prove the existence of whitewash-proof rating protocols.

Proposition 6. If $U^* > 0$, then whitewash-proof rating protocols always exist.

Proof: If $U^* > 0$, then sustainable rating protocols that stimulate positive levels of cooperation always exist. According to Lemma 3, if a protocol κ with $K = \theta > 0$ is sustainable, then a protocol κ' with the same rating scheme and recommended strategy and $K = 0$ is also sustainable. Meanwhile, it can be verified that κ' is whitewash-proof since an agent cannot get any benefit by leaving and rejoining the community while also suffering the whitewashing cost. Hence, Proposition 6 follows. ■

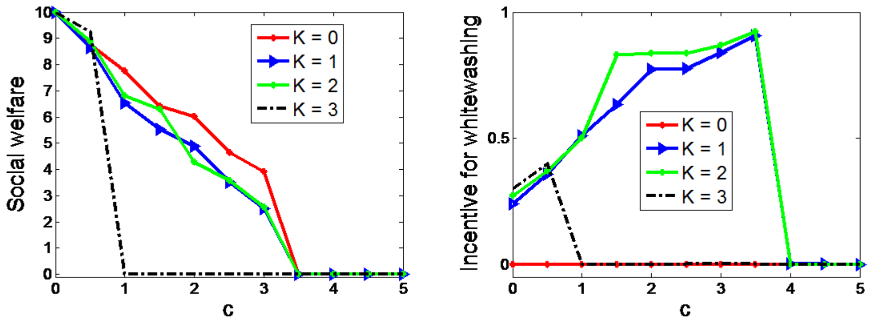
Now we investigate the impacts of the initial rating score K on social welfare and the incentive for whitewashing. We first consider the case where the recommended



(a) (b)

Figure 2.9: (a) Social welfare and (b) the incentive for whitewashing under recommended strategy σ_L^C when $L = 3$ and $c_w = 1$.

strategy is fixed. Figure 2.9 plots social welfare and the incentive for whitewashing under a rating protocol $(3, K, \sigma_3^C)$ for $K = 0, \ldots, 3$, where σ_L^C is defined by $\sigma_L^C(\theta, \tilde{\theta}) = F$ if and only if $\tilde{\theta} \geqslant \theta$, for all θ, as in the above section. We can see that larger K yields higher social welfare and at the same time a larger incentive for whitewashing since new agents are treated better. Hence, there is a trade-off between efficiency and whitewash-proofness as we increase K while fixing the recommended strategy. Next we consider the optimal recommended strategy given a rating scheme (L, K). Figure 2.10 plots social welfare and the incentive for whitewashing under the optimal recommended s-trategy for $L = 3$ and $K = 0, \ldots, 3$. We can see that giving the highest rating score to new agents ($K = 3$) yields the highest social welfare but it can prevent whitewashing only for small values of c. With our parameter specification, choosing $K = 3$ is optimal only for small c, and optimal K drops to 0 for other values of c with which some cooperation can be sustained. Figure 2.11 plots the optimal initial rating K^* as we vary the whitewashing cost c_w, for $c = 1, 2, 3$. As c_w increases, the incentive constraints for whitewashing becomes less binding, and thus $K^*$is non-decreasing in c_w. On the other hand, as c increases, it becomes more difficult to sustain cooperation while the difference between $v_\kappa^\infty(0)$ and $v_\kappa^\infty(\min\{\theta + 1, L\})$ increases for all θ such that $\sigma(\theta, \tilde{\theta}) = F$ for some $\tilde{\theta}$. As a result, K^* is non-increasing in c.



(a) (b)

Figure 2.10: (a) Social welfare and (b) the incentive for whitewashing under the optimal recommended strategy when $L = 3$ and $c_w = 1$.

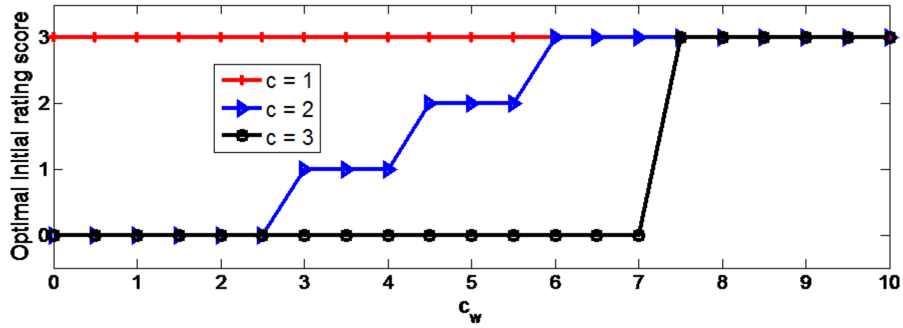


Figure 2.11: Optimal initial rating score when $L = 3$.

2.5.3 One-sided Rating Protocols

The above discussion focuses on the design of optimal rating protocols where the recommended strategy utilizes both the rating scores of the client and the server in order to determine the server's action. We refer to such recommended strategies as *two-sided recommended strategies* since they involve the rating scores of both players involved in the stage game.

In this section, we discuss the design of optimal rating protocols with a simple class of recommended strategies that only utilize one-sided rating scores, which we refer to as one-sided recommended strategies. Particularly, a one-sided recommended strategy determines an agent's serving action solely based on either the agent's own rating score or the rating score of its client. To differentiate it with the previously discussed two-sided recommended strategies, we denote a one-sided recommended strategy by φ, which can be represented by a mapping $\varphi : \Theta \to \mathcal{A}$, and the corresponding rating protocol, which is called as a one-sided rating protocol, by π. It should be noted that for a one-sided recommended strategy φ that utilizes the clients' rating scores, it is equivalent to a two-sided recommended strategy σ if $\sigma(\theta, \tilde{\theta}) = \varphi(\tilde{\theta})$, $\forall\theta \in \Theta$ and $\forall\tilde{\theta} \in \Theta$. Similarly, for a one-sided recommended strategy φ' that utilizes the servers' rating scores, it is equivalent to a two-sided recommended strategy σ' if $\sigma'(\theta, \tilde{\theta}) = \varphi'(\theta)$, $\forall\theta \in \Theta$ and $\forall\tilde{\theta} \in \Theta$. Therefore, the one-sided recommended strategies represent a subset of the class of two-sided recommended strategies. In this section, we investigate the emerging protocol designs which can be found using such simpler strategies and the corresponding efficiency loss compared to the optimal performance obtained in Section 4.

We first analyze one-side recommended strategies utilizing the clients' rating s-cores. Given a one-sided recommended strategy φ, the expected period payoff of a θ-agent before it is matched, which is still denoted as $v_\kappa(\theta)$ with slight abuse of notation, is given as follows

$$v_\pi(\theta) = bI(\varphi(\theta) = F) - \sum_{\tilde{\theta}\in\Theta} \eta_L(\tilde{\theta})cI(\varphi(\tilde{\theta}) = F), \tag{2.33}$$

where $I(x)$ is an indicator function which takes value of 1 when $x = 1$. The corresponding social welfare, which is denoted as W_π, can be computed by

$$W_\pi = \sum_{\theta\in\Theta} \eta_L(\theta)v_\pi(\theta) = \sum_{\theta\in\Theta} \eta_L(\theta)(b - c)I(\varphi(\theta) = F). \tag{2.34}$$

The following proposition characterizes the general designing rule of the optimal rating protocol, which is denoted as $\pi^* = (L^*, \varphi^*)$, with the corresponding optimal social welfare denoted as W^*.

Proposition 7. With one-sided recommended strategies utilizing the clients' rating scores, the optimal rating protocol π^* that maximizes (34) satisfies the following conditions: when $c/b \leqslant \delta(1-2\varepsilon)$, $W^* = (1-\varepsilon+\varepsilon\alpha)(b-c)$ with $L^* = 1$, $\varphi^*(0) = D$, and $\varphi^*(1) = F$; when $c/b > \delta(1-2\varepsilon)$, $W^* = 0$ with $\varphi^*(\theta) = D$, $\forall\theta \in \Theta$.

Proof: See Appendix D. ■

The optimal rating protocol from Proposition 7 is surprisingly simple and intuitive. When the cost-to-benefit ratio is sufficiently small such that a positive level of cooperation can be sustained in the community, the optimal rating protocol contains only two different rating values. Modulo the effects of noise, agents who comply with the recommended strategy in the previous period have rating score 1, while agents who deviate from the recommended strategy have rating score 0. The recommended strategy then says that agents should play a tit-for-tat-like strategy, providing services to agents with rating score 1 and punishing those with rating score 0. On the other hand, when the cost-to-benefit ratio is sufficiently large, no cooperation can be sustained and $W^* = 0$. Hence, when the recommended strategy is one-sided and solely utilizes the clients' rating scores, there is no need to construct complicated rating protocols that are difficult for agents to understand, or to heavily optimize parameters of the rating protocol based on the properties of the community. Regarding the fact that the optimal social welfare U^* achieved by two-sided recommended strategies is upper-bounded by $b-c$, we have the following corollary.

Corollary 2. When $c/b \leqslant \delta(1-2\varepsilon)$, $U^* - W^* < \varepsilon(1-\alpha)(b-c)$. ■

Therefore, the efficiency loss introduced by one-sided recommended strategies monotonically decreases and approaches 0 when $\varepsilon \to 0$ or $\alpha \to 1$.

We then study one-sided recommended strategies utilizing only the servers' rating scores. It is shown in the following proposition that no cooperation can be sustained in this case, which always yields an optimal social welfare $W^* = 0$. Hence, one-sided recommended strategies utilizing only the servers' rating scores can never correctly incentivize service provisions and prevent agents from free-riding.

Proposition 8. With one-sided recommended strategies utilizing the servers' rating scores, the optimal rating protocol π^* always delivers an optimal social welfare $W^* = 0$ with $\varphi^*(\theta) = D, \forall\theta \in \Theta$.

Proof: See Appendix D. ■

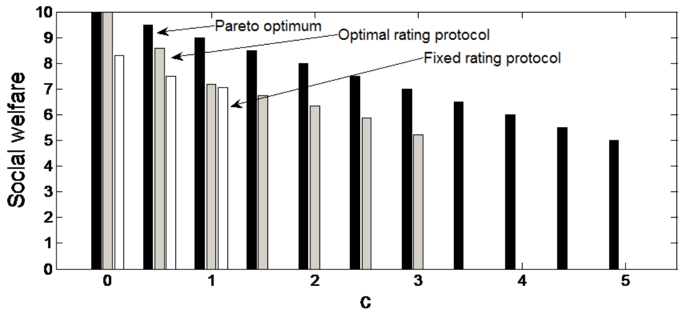
2.6 Illustrative Examples

In this section, we present numerical results to illustrate in detail the performance of optimal rating protocols. Unless stated otherwise, the setting of the community is as follows: the benefit per service ($b = 10$), the cost per service ($c = 1$), the discount factor ($\beta = 0.8$), the turnover rate ($\alpha = 0.1$), the report error ($\varepsilon = 0.2$), the punishment step ($M = L$), and the initial rating score ($K = L$). Since the number of all possible recommended strategies given a punishment length L increases exponentially with L, it takes a long time to compute the optimal recommended strategy even for a moderate value of L. Hence, we consider rating protocols $\kappa = (L, \sigma_L^*)$ for $L = 1, 2, 3$.

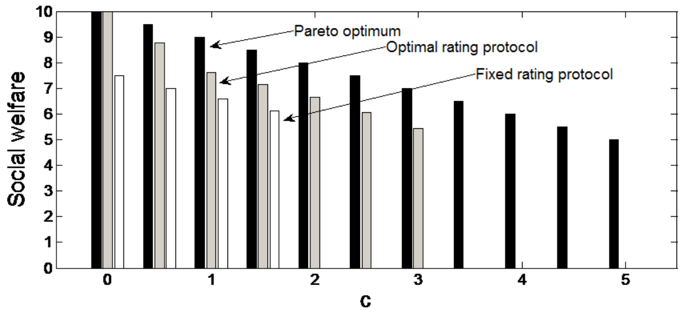
We first compare the performances of the optimal rating protocol and the fixed rating protocol for $L = 1, 2, 3$. For each L, we use (L, σ_L^C) as the fixed rating protocol. Figure 2.12 illustrates the results, with the black bar representing the pareto-optimal value $b - c$, i.e., the highest social welfare that can be possibly sustained by a rating protocol, the gray bar representing the social welfare of the optimal rating protocol, and the white bar representing the social welfare of (L, σ_L^C). As it shows, the optimal rating protocol (L, σ_L^*) outperforms (L, σ_L^C). When c is small, (L, σ_L^*) delivers higher social welfare than (L, σ_L^C). When c is sufficiently large such that no cooperation can be sustained under (L, σ_L^C) (the height of the white bar becomes 0), a positive level of cooperation can still be sustained under (L, σ_L^*).

Next, we analyze the impacts of the community's parameters on the performance of optimal rating protocols.

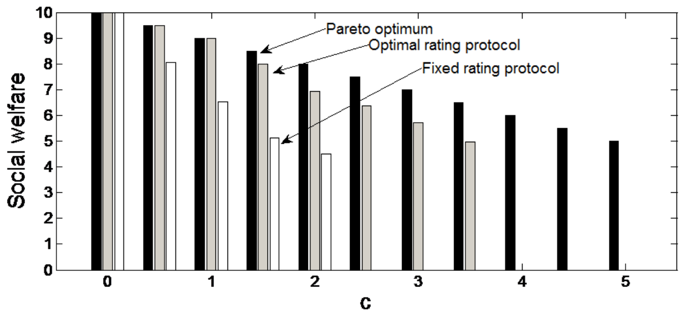
Impact of the Discount Factor: We discuss the impact of the discount factor β on the performance of optimal rating protocols. As β increases, an agent puts a higher



(a) $L = 1$



(b) $L = 2$

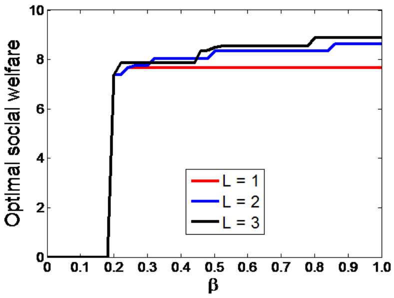


(c) $L = 3$

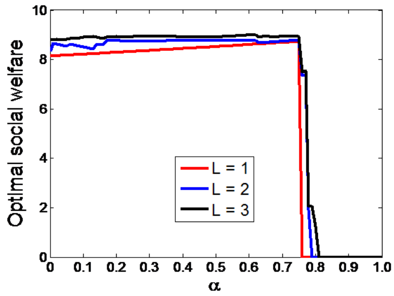
Figure 2.12: Performances of the optimal rating protocol (L, σ_L^*) and the fixed rating protocol (L, σ_L^C).

weight on its future payoff relative to its instant payoff. Hence, with larger β, it is easier to sustain cooperation using future reward and punishment through a rating protocol. This is illustrated in Figure 2.13(a), which shows that social welfare is non-decreasing in β.

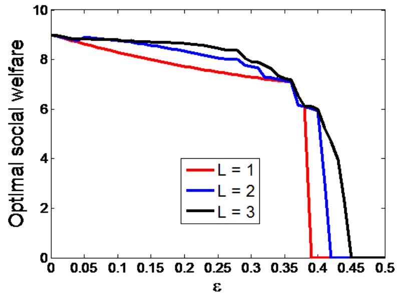
Impact of the Turnover Rate: Increasing α has two opposite effects on social wel-



(a)



(b)



(c)

Figure 2.13: Optimal social welfare given L as (a) β, (b) α, and (c) ε vary.

fare. As α increases, the weight on the future payoffs, $\delta = \beta(1-\alpha)$, decreases, and thus it becomes more difficult to sustain cooperation. On the other hand, as α increases, there are more agents holding the highest rating score given the restriction $K = L$. In general, agents with the highest rating score are treated well under optimal recommended strategies, which implies a positive effect of increasing α on social welfare. From Figure 2.13(b), we can see that, when α is large, the first effect is dominant, making cooperation not sustainable. We can also see that the second 0effect is dominant for the values of α with which cooperation can be sustained, yielding an increasing tendency of social welfare with respect to α.

Impact of the Report Errors: As ε increases, it becomes more difficult to sustain cooperation because reward and punishment provided by a rating protocol becomes more random. At the same time, larger ε increases the fraction of 0-agents in the stationary distribution, which usually receive the lowest long-term payoff among all rating scores. Therefore, we can expect that optimal social welfare has a non-increasing tendency with respect to ε, as illustrated in Figure 2.13(c). When ε is sufficiently close to 1/2, σ_L^D is the only sustainable recommended strategy and social welfare falls to . On the other direction, as ε approaches 0, social welfare converges to its upper bound $b-c$, regardless of the punishment length, as can be seen from Proposition 1(iii). We can also observe from Figure 13 that the regions of α and ε where some cooperation can be sustained (i.e., $U_L^* > 0$) become wider as L increases, whereas that of β is independent of L.

2.7 Comparison with Existing Works on Repeated Games with Imperfect Monitoring

In this section, we compare our work with the existing literature on repeated games with imperfect monitoring, including the works on reputation systems [23]-[25] and the seminal work of Fudenberg, Levine and Maskin [35]. Importantly, our work exhibits significant technical differences from the existing literature. We would like to point out that although the results derived in this work exhibit some structural similarity to [23]-[25], they are derived under completely different settings and using different analytical methods. Also, the methodology in [35] cannot be applied in our work.

We first compare our work with [23]-[25]. The models in [23]-[25] assume that agents have fixed roles in the community (i.e. sellers and buyers), which is common in applications where the groups of sellers and buyers are separated and usually do not overlap. However, in our work, agents are symmetric in the sense that each of them can play both roles of server and of client. Our model is more appropriate to characterize resource/knowledge sharing online communities.

More importantly, the objective in the protocol design of our work is also different from [23]-[25]. The design objective in [23]-[25] is to maximize the expected discounted long-term payoff of the seller starting from the best reputation and a clean history (e.g. Propositions 2 and 3 in [23]). Translating this into the mathematical representation adopted in this chapter, the objective function in [23]-[25] is

$$\max_{(L,\sigma)} U_\kappa = v_\kappa^\infty(L). \tag{2.35}$$

Such objective function is reasonable in the setting of [23]-[25] because they assume that there is a unique long-lived player (seller) in the game who is foresighted. In each period, the seller selects one buyer to interact with. Therefore, it makes sense to focus on the life-time discounted utility of this seller starting from the moment it joins the community. Also, it should be pointed out that the selection of this objective function (2.35) is the main reason why the optimal design can be achieved in [23]-[25] using a two-level reputation set with $L = 1$. Since the future utility is discounted, the optimum of (2.35) can be achieved using a simple grim-trigger strategy, i.e. the seller cooperates as long as its reputation remains at "good" and does not cooperate when its reputation falls to "bad".

In contrast, our work assumes multiple long-lived players coexisting in the community (i.e. each agent is long-lived and foresighted). Meanwhile, in each period, there are multiple interactions between different players. Hence, it is more reasonable to maximize the average social welfare of all agents in the long-run as defined in (2.17), which can be proved to be equivalent to the following objective function:

$$\underset{(L,\sigma)}{\text{maximize}} \; U_\kappa = \sum_\theta \eta_L(\theta) v_\kappa^\infty(\theta). \tag{2.36}$$

It is easy to observe that given the grim-trigger strategy designed in [23], the long-run timing-average payoff of the seller is actually 0, even though its expected discounted long-term payoff is maximized. Therefore, the design in [23]-[25] is never optimal from a social welfare perspective, given $\varepsilon > 0$.

Finally, the works in [23] and [24] focus on deriving the rating schemes that can achieve the upper bound of (2.35) under the condition when the seller is sufficiently patient with its discount factor δ (i.e. β in this chapter) close to 1 or when the payment-to-cost ratio ρ (i.e. b/c in this chapter) is sufficiently large. They did not provide much insight on how to derive the optimal rating scheme and what is the optimal expected long-term utility that can be achieved when δ and ρ are small (They simply state that the optimal expected long-term utility is below the upper bound of (2.35) in this case). In contrast, our work tries to characterize the optimal rating scheme design for the entire region of the parameters $(\beta, b, c, \varepsilon)$, but not only the scenario when these parameters are ideal. Next, we compare our work with [35]. The model in [35] does not consider the anonymity and random matching among agents. Also, it assumes that the entire history of public signals (i.e. the outcome of each stage game) is revealed during the repeated game. It should be noted that the proof of Proposition 1 in [23] also relies on this assumption on the information structure in order to obtain the upper bound on the efficiency. Nevertheless, under the rating protocol proposed in our work, each agent only observes a limited set of past L signals from the past periods and hence, the assumption that all past signals are revealed no longer holds here.

Another difference between our work and [35] is that the objective in [35] is also to maximize the expected long-term utility of players starting from the beginning of the game (as described in (2.35)), which is different from our objective function (2.17).

Regarding all the above reasons, we would like to point out that the methodology in [35] cannot be applied to our work.

2.8 Conclusions

In this chapter, we used the idea of rating protocols to establish a rigorous framework for the design and analysis of a class of incentive schemes to sustain cooperation in online communities. We derived conditions for sustainable rating protocols, under which no agent gains by deviating from the prescribed recommended strategy. We formulated the problem of designing an optimal rating protocol and characterized optimal social welfare and optimal recommended strategies given parameters. As special cases, we analyzed the one-sided rating protocol which only utilizes the rating score of one party in the stage-game. It was shown that when only the clients' rating scores are utilized, the optimal one-sided rating protocol preserves a simple structure with two-level rating scores, whereas when only the servers' rating scores are utilized, no sustainable one-sided rating protocol can be designed. We also discussed the impacts of punishment lengths and whitewashing possibility on the design and performance of optimal rating protocols, identifying a trade-off between efficiency and incentives. Lastly, we presented numerical results to illustrate the impacts of the discount factor, the turnover rate, and the probability of report errors on the performance of optimal rating protocols. Our framework provides a foundation for designing incentive schemes which can be deployed in real-world communities populated by anonymous, self-interested individuals.

2.9 Appendices

2.9.1 Appendix A: Proof of Proposition 1

(i) $U^* \geqslant 0$ follows by noting that (L, σ_L^D) is sustainable. It is shown in [35] that in a repeated game with public signal, the maximum long-run player sequential equilibrium payoff when the signal has full support is the solution of the following linear programming problem:

$$\begin{array}{ll} \max\limits_{a} & u^\infty \\ \text{subject to} & u^\infty = u - c + \delta(1-\varepsilon)v^\infty(+) + \delta\varepsilon v^\infty(-),\ for\ a = F \\ & u^\infty \geqslant u - c + \delta(1-\varepsilon)v^\infty(+) + \delta\varepsilon v^\infty(-),\ for\ a = D\ \cdot \\ & u^\infty \geqslant u + \delta\varepsilon v^\infty(+) + \delta(1-\varepsilon)v^\infty(-),\ for\ a = F \\ & u^\infty = u + \delta\varepsilon v^\infty(+) + \delta(1-\varepsilon)v^\infty(-),\ for\ a = D \end{array} \tag{2.37}$$

Here, u is the benefit that the player can receive as a client in a stage game, and $v^\infty(+)$ is the expected long-term utility of this player if his rating score is increased and $v^\infty(-)$ is the expected long-term utility when his rating score is decreased. With simple computation, the maximum solution of (2.37) can be written as $\frac{1}{1-\delta}(b - \frac{1-\varepsilon}{1-2\varepsilon}c)$. Since no player can achieve a long-run payoff higher than this, it can be concluded that the social welfare is also upper-bounded by $b - \frac{1-\varepsilon}{1-2\varepsilon}c$.

(ii) By (2.8), we can express $v_\kappa^\infty(1) - v_\kappa^\infty(0)$ as

$$\begin{aligned} & v_\kappa^\infty(1) - v_\kappa^\infty(0) \\ = \;& v_\kappa(1) + \delta[(1-\varepsilon)v_\kappa^\infty(2) + \varepsilon v_\kappa^\infty(0)] - v_\kappa(0) - \delta[(1-\varepsilon)v_\kappa^\infty(1) + \varepsilon v_\kappa^\infty(0)] \\ = \;& v_\kappa(1) - v_\kappa(0) + \delta(1-\varepsilon)[v_\kappa^\infty(2) - v_\kappa^\infty(1)] \end{aligned} \cdot \tag{2.38}$$

Similarly, we have

$$\begin{array}{l} v_\kappa^\infty(2) - v_\kappa^\infty(1) = v_\kappa(2) - v_\kappa(1) + \delta(1-\varepsilon)[v_\kappa^\infty(3) - v_\kappa^\infty(2)] \\ \vdots \\ v_\kappa^\infty(L) - v_\kappa^\infty(L-1) = v_\kappa(L) - v_\kappa(L-1) \end{array} . \tag{2.39}$$

In general, for $\theta = 1, \ldots, L$,

$$v_\kappa^\infty(\theta) - v_\kappa^\infty(\theta-1) = \sum_{l=0}^{L-\theta} \gamma^l [v_\kappa(\theta+l) - v_\kappa(\theta+l-1)], \tag{2.40}$$

where we define $\gamma = \delta(1-\varepsilon)$. Thus, we obtain

$$\begin{aligned} & v_\kappa^\infty(\theta) - v_\kappa^\infty(0) \\ = \;& v_\kappa(\theta) - v_\kappa(0) + \gamma[v_\kappa(\theta+1) - v_\kappa(1)] + \cdots + \gamma^{L-\theta}[v_\kappa(L) - v_\kappa(L-\theta)] \\ & + \gamma^{L-\theta+1}[v_\kappa(L) - v_\kappa(L-\theta+1)] + \cdots + \gamma^{L-1}[v_\kappa(L) - v_\kappa(L-1)] \\ = \;& \sum_{l=0}^{L-1} \gamma^l [v_\kappa(\min\{\theta+l, L\}) - v_\kappa(l)] \end{aligned} \tag{2.41}$$

for $\theta = 1, \ldots, L$.

Since $-c \leqslant v_\kappa(\theta) \leqslant b$ for all θ, we have $v_\kappa(\theta) - v_\kappa(\tilde{\theta}) \leqslant b + c$ for all $(\theta, \tilde{\theta})$. Hence, by (2.41),

$$v_\kappa^\infty(\theta) - v_\kappa^\infty(0) \leqslant \frac{1-\gamma^L}{1-\gamma}(b+c) \leqslant \frac{b+c}{1-\gamma} \tag{2.42}$$

for all $\theta = 1, \ldots, L$, for all $\kappa = (L, \sigma)$. Therefore, if $\delta(1-2\varepsilon)[(b+c)/(1-\gamma)] < c$, or equivalently, $c/b > [\beta(1-\alpha)(1-2\varepsilon)]/[1 - \beta(1-\alpha)(2-3\varepsilon)]$, then the incentive constraint (2.15) cannot be satisfied for any θ, for any rating protocol (L, σ). This implies that any recommended strategy σ such that $\sigma(\theta, \tilde{\theta}) = F$ for some $(\theta, \tilde{\theta})$ is not sustainable, and thus $U^* = 0$.

(iii) For any L, define a recommended strategy σ_L^{D0} by $\sigma_L^{D0}(\theta, \tilde{\theta}) = D$ for $\tilde{\theta} = 0$ and $\sigma_L^{D0}(\theta, \tilde{\theta}) = F$ for all $\tilde{\theta} > 0$, for all θ. In other words, with σ_L^{D0} each agent declines the service request of 0-agents while providing a service to other agents. Consider a rating protocol $\kappa = (1, \sigma_1^{D0})$. Then $v_\kappa(0) = -\eta_1(1)c$ and $v_\kappa(1) = b - \eta_1(1)c$. Hence, $U_\kappa = [1-(1-\alpha)\varepsilon](b-c)$ and $v_\kappa^\infty(1) - v_\kappa^\infty(0) = b$, and thus the incentive constraint $\delta(1-2\varepsilon)(v_\kappa^\infty(1) - v_\kappa^\infty(0)) \geqslant c$ is satisfied by the hypothesis $c/b \leqslant \beta(1-\alpha)(1-2\varepsilon)$. Since there exists a feasible solution that achieves $U_\kappa = [1-(1-\alpha)\varepsilon](b-c)$, we have $U^* \geqslant [1-(1-\alpha)\varepsilon](b-c)$.

(iv) The result can be obtained by combining (i) and (iii).

(v) Suppose that $U^* = b - c$, and let (L, σ) be an optimal rating protocol that achieves $U^* = b - c$. It is easy to obtain that $\varepsilon = 0$. Then by (2.4), $\eta_L(\theta) = 0$ for all $0 \leqslant \theta \leqslant L-1$ and $\eta_L(L) = 1$. Hence, σ should have $\sigma(L, L) = F$ in order to attain $U^* = b - c$. Since $v_\kappa(L) = b - c$ and $v_\kappa(\theta) \geqslant -c$ for all $0 \leqslant \theta \leqslant L-1$, we have $v_\kappa^\infty(L) - v_\kappa^\infty(0) \leqslant b/(1-\gamma)$ by (2.41). If $\delta b/(1-\delta) < c$, then the incentive constraint for L-agents, $\delta[v_\kappa^\infty(L) - v_\kappa^\infty(0)] \geqslant c$, cannot be satisfied. Therefore, we obtain $c/b \leqslant \delta/(1-\delta)$. ■

2.9.2 Appendix B: Proof of Proposition 2

Choose an arbitrary L. To prove the result, we will construct a recommended strategy σ_{L+1} using punishment length $L+1$ that is sustainable and achieves U_L^*. Define σ_{L+1} by

$$\sigma_{L+1}(\theta, \tilde{\theta}) = \begin{cases} \sigma_L^*(\theta, \tilde{\theta} & \text{for } \theta \leqslant L \text{ and } \tilde{\theta} \leqslant L \\ \sigma_L^*(L, \tilde{\theta}) & \text{for } \theta = L+1 \text{ and } \tilde{\theta} \leqslant L \\ \sigma_L^*(\theta, L) & \text{for } \theta \leqslant L \text{ and } \tilde{\theta} = L+1 \\ \sigma_L^*(L, L) & \text{for } \theta = L+1 \text{ and } \tilde{\theta} = L+1 \end{cases} . \tag{2.43}$$

Let $\kappa = (L, \sigma_L^*)$ and $\kappa' = (L+1, \sigma_{L+1})$. From (2.4), we have $\eta_{L+1}(\theta) = \eta_L(\theta)$ for $\theta = 0, \ldots, L-1$ and $\eta_{L+1}(L) + \eta_{L+1}(L+1) = \eta_L(L)$. Using this and (2.6), it is straightforward to see that $v_{\kappa'}(\theta) = v_\kappa(\theta)$ for all $\theta = 0, \ldots, L$ and $v_{\kappa'}(L+1) = v_\kappa(L)$. Hence, we have that

$$\begin{aligned} U_{\kappa'} &= \sum_{\theta=0}^{L+1} \eta_{L+1}(\theta) v_{\kappa'}(\theta) = \sum_{\theta=0}^{L-1} \eta_{L+1}(\theta) v_{\kappa'}(\theta) + \sum_{\theta=L}^{L+1} \eta_{L+1}(\theta) v_{\kappa'}(\theta) \\ &= \sum_{\theta=0}^{L-1} \eta_L(\theta) v_\kappa(\theta) + \sum_{\theta=L}^{L+1} \eta_{L+1}(\theta) v_\kappa(L) \\ &= \sum_{\theta=0}^{L-1} \eta_L(\theta) v_\kappa(\theta) + \eta_L(L) v_\kappa(L) = U_\kappa = U_L^* \end{aligned} . \tag{2.44}$$

Using (2.44), we can show that $v_{\kappa'}^\infty(\theta) - v_{\kappa'}^\infty(0) = v_\kappa^\infty(\theta) - v_\kappa^\infty(0)$ for all $\theta = 1, \ldots, L$and $v_{\kappa'}^\infty(L+1) - v_{\kappa'}^\infty(0) = v_\kappa^\infty(L) - v_\kappa^\infty(0)$. By the definition of σ_{L+1}, the right-hand side of the relevant incentive constraint (i.e., c or $-c$) for each $\theta = 0, \ldots, L$ is the same both under σ_L^* and under σ_{L+1}. Also, under σ_{L+1}, the right-hand side of the relevant incentive constraint for $\theta = L+1$ is the same as that for $\theta = L$. Therefore, σ_{L+1} satisfies the relevant incentive constraints for all $\theta = 0, \ldots, L+1$. ■

2.9.3 Appendix C: Proof of Proposition 3

To facilitate the proof, we define $u_\kappa^\infty(\theta)$ by

$$u_\kappa^\infty(\theta) = \sum_{l=0}^{\infty} \gamma^l v_\kappa(\min\{\theta + l, L\}) \tag{2.45}$$

for $\theta = 0, \ldots, L$. Then, by (2.41), we have $v_\kappa^\infty(\theta) - v_\kappa^\infty(0) = u_\kappa^\infty(\theta) - u_\kappa^\infty(0)$ for all $\theta = 1, \ldots, L$. Thus, we can use $u_\kappa^\infty(\theta) - u_\kappa^\infty(0)$ instead of $v_\kappa^\infty(\theta) - v_\kappa^\infty(0)$ in the incentive constraints of DP_L.

Suppose that $\sigma_L^*(0, \hat{\theta}) = F$ for some $\hat{\theta}$. Then the relevant incentive constraint for a 0-agent is $\delta(1 - 2\varepsilon)[u_\kappa^\infty(1) - u_\kappa^\infty(0)] \geqslant c$. Suppose that $\sigma_L^*(0, \bar{\theta}) = D$ for some $\bar{\theta} \in \{1, \ldots, L-1\}$ such that $\bar{\theta} \geqslant \ln \frac{c}{b} / \ln \beta$. Consider a recommended strategy σ_L' defined by

$$\sigma_L'(\theta, \tilde{\theta}) = \begin{cases} \sigma_L^*(\theta, \tilde{\theta}) & \text{for } (\theta, \tilde{\theta}) \neq (0, \bar{\theta}) \\ F & \text{for } (\theta, \tilde{\theta}) = (0, \bar{\theta}) \end{cases} . \tag{2.46}$$

That is, σ'_L is the recommended strategy that differs from σ^*_L only at $(0, \bar{\theta})$. Let $\kappa = (L, \sigma^*_L)$ and $\kappa' = (L, \sigma'_L)$. Note that $v_{\kappa'}(0) - v_\kappa(0) = -\eta_\tau(\bar{\theta})c < 0$ and $v_{\kappa'}(\bar{\theta}) - v_\kappa(\bar{\theta}) = \eta_\tau(0)b > 0$ since $\varepsilon > 0$ and $\alpha < 1$. Thus, $U_{\kappa'} - U_\kappa = \eta_L(0)\eta_L(\bar{\theta})(b-c) > 0$. Also,

$$u^\infty_{\kappa'}(\theta) - u^\infty_\kappa(\theta) = \begin{cases} (1-\alpha)^{\bar{\theta}+1}(1-\varepsilon)^{\bar{\theta}}\varepsilon[\beta^{\bar{\theta}}b - c] & \text{for } \theta = 0, \\ \gamma^{\bar{\theta}-\theta}[v_{\kappa'}(\bar{\theta}) - v_\kappa(\bar{\theta})] & \text{for } \theta = 1, \ldots, \bar{\theta}, \\ 0 \quad \text{for } \theta = \bar{\theta}+1, \ldots, L \end{cases} . \tag{2.47}$$

Since $\bar{\theta} \geqslant \ln\frac{c}{b}/\ln\beta$, we have $u^\infty_{\kappa'}(0) - u^\infty_\kappa(0) \leqslant 0$. Thus, $u^\infty_{\kappa'}(\theta) - u^\infty_{\kappa'}(0) \geqslant u^\infty_\kappa(\theta) - u^\infty_\kappa(0)$ for all $\theta = 1, \ldots, L$. Since $\sigma^*_L(0, \hat{\theta}) = F$ for some $\hat{\theta}$, the relevant incentive constraint for a θ-agent is the same both under σ^*_L and under σ'_L, for all θ. Hence, σ'_L satisfies the incentive constraints of DP_L, which contradicts the optimality of σ^*_L. This proves that $\sigma^*_L(0, \tilde{\theta}) = F$ for all $\tilde{\theta} \geqslant \ln\frac{c}{b}/\ln\beta$. Similar approaches can be used to prove $\sigma^*_L(0, L) = F$, (i), and (iii). ■

2.9.4 Appendix D: Proof of Proposition 7

To prove this proposition, we first show that the recommended strategy in the optimal rating protocol is always threshold based. That is, there is an integer h such that $\varphi^*(\theta) = D$ for all $\theta < h$ and $\varphi^*(\theta) = F$ for all $\theta \geqslant h$. We use a contradiction to verify this. Suppose $\varphi^*(\tilde{\theta}) = F$ and $\varphi^*(\tilde{\theta}+1) = D$ for some θ. Consider a strategy φ' satisfying $\varphi'(\theta) = \varphi^*(\theta)$ for all $\theta \neq \tilde{\theta}+1$ and $\varphi'(\tilde{\theta}) = \varphi^*(\tilde{\theta})$. Since φ^* is sustainable, then according to (2.8), it is easy to verify that φ' is also sustainable. Meanwhile, φ' delivers a higher social welfare than φ^*. Hence, the fact that φ^* is optimal is contradicted and we can conclude that the optimal recommended strategy is always threshold-based. According to (2.34), it is obvious that the social welfare upon agents' compliance monotonically decreases with the threshold h. Next, we analyse how h affects the sustainability of the recommended strategy.

Under a threshold-based recommended strategy with threshold h, the expected long-term utility can be recursively represented as

$$
\begin{array}{rcl}
v_\kappa^\infty(\theta \geqslant h) & = & b - \sum\limits_{\theta \geqslant h} \eta_L(\theta) c \\
& & +\delta(1-\varepsilon) v_\kappa^\infty(\min\{\theta+1, L\}) + \delta\varepsilon v_\kappa^\infty(\max\{\theta-1, 0\}) \\
v_\kappa^\infty(\theta < h) & = & - \sum\limits_{\theta \geqslant h} \eta_L(\theta) c \\
& & +\delta(1-\varepsilon) v_\kappa^\infty(\min\{\theta+1, L\}) + \delta\varepsilon v_\kappa^\infty(\max\{\theta-1, 0\})
\end{array}. \quad (2.48)
$$

Hence, it can be shown that $v_\kappa^\infty(1) - v_\kappa^\infty(0) = \min\limits_{\theta}\{v_\kappa^\infty(\theta) - v_\kappa^\infty(0)\}$. Meanwhile, $v_\kappa^\infty(1) - v_\kappa^\infty(0)$ monotonically decreases against h. Therefore, if a recommended s- trategy with threshold h is sustainable, then the recommended strategy with threshold $h-1$ is also sustainable. However, according to Proposition 3, it should be noted that the recommended strategy with $h=0$ can never be sustainable. Therefore, to sum up, for a given punishment length L, when sustainable one-sided rating protocols exist, the threshold of the optimal recommended strategy φ^* is always $h=1$. It is easy to compute that any threshold-based recommended strategy φ with $h=1$ is sustainable if and only if $c/b \leqslant \delta(1-2\varepsilon)$ with the resulting social welfare to be $W = (1-\varepsilon+\varepsilon\alpha)(b-c)$. Hence, Proposition 7 follows. ■

2.9.5 Appendix E: Proof of Proposition 8

The proof of this proposition is similar to that of Proposition 7. First, it can be proved that in the optimal rating protocol that uses servers' rating scores, the recommended strategy is always threshold-based. That is, there is an integer h such that $\varphi^*(\theta) = F$ for all $\theta < h$ and $\varphi^*(\theta) = D$ for all $\theta \geqslant h$. However, under such threshold-based recommended strategy, $v_\kappa^\infty(\theta) - v_\kappa^\infty(0) \leqslant c$ hold for all $\theta \in \Theta$ and hence, (2.12) is never satisfied and Proposition 8 follows. ■

CHAPTER 3

Strategic Learning in Online Communities

3.1 Introduction

Online communities play an increasingly central role in enabling individual users to remotely share resources or crowdsource their service needs. However, they are vulnerable to intrinsic incentive problems which lead to prevalent free-riding behaviors among users that sacrifice the collective social welfare of the community [39][40].

Various incentive mechanisms have been proposed to encourage cooperation in online communities [40][41], with a large body of them relying on the idea of reciprocity, in which users monitor the behavior of each other and provide differential services based on their observations. Such reciprocity-based mechanisms can be classified into direct reciprocity and indirect reciprocity. In direct reciprocity mechanisms [40][41][43], users can identify each other, and determine whether they should supply services to a specific user based on their personal history of interactions with that specific user. Such mechanisms rely on the players' ability to maintain private histories of other users' play which involve frequent interactions among identifiable users and hence are not effective in online communities with large populations of anonymous users, who are randomly matched and interact infrequently [39].

Indirect reciprocity mechanisms [23]-[25][45], in contrast, deploy reputation protocols that allow the users in a community to share information and observations to form collective opinions about other users (e.g. assign reputations). In a reputation protocol, a reputation score is assigned to each user upon entering the community and is updated based on the reports (which may be erroneous) provided by other users who he interacts with. Hence, an individual user can be rewarded or punished (based on his reputation score) by other users who have not had past interactions with him. To encourage users to contribute their services to the community, differential services are provided to the users based on their reputations, i.e. users with higher reputations will receive better services from other users in return. Since indirect reciprocity mechanisms require neither observable identities (i.e. only a user's reputation but not his identity needs to be revealed in order to determine the differential services towards him), nor frequent interactions, they form a suitable basis for designing incentive mechanisms for online communities with large populations.

The reputation protocols designed in [23]-[25][45] are able to sustain the selected ("best") equilibrium for various online communities. They assume that the community starts to operate "directly" at the selected equilibrium, i.e. users start by directly adopting a desirable equilibrium strategy when joining the community, and design reputation protocols that ensure users will not deviate from this equilibrium strategy. Nevertheless in practice, there are often multiple equilibria in a community and users can adopt arbitrary (non-equilibrium) strategies when they join the community. In this case, all the above works fail to determine 1) how the community evolves if it starts from an arbitrary (non-equilibrium) state and 2) whether in the long-run it can converge to the best equilibrium given the deployed reputation protocol. Moreover, none of the prior works investigated the robustness of equilibria, i.e. to determine if the community deviates from an equilibrium due to stochastic perturbations, e.g. noises, errors, whether it can converge to this equilibrium again in the long run.

In summary, this chapter generalizes [23]-[25][45] and focuses on the design of reputation protocols that do not only sustain the best equilibrium, but also ensure the convergence of the community towards this equilibrium in the long-run. To design such a protocol, it is of critical importance to consider the following design aspects, which were neglected in [23]-[25][45]:

(1) how users learn, i.e. how they update their knowledge about the community, during their interactions and how they adapt their strategies based their limited (imperfect) knowledge;

(2) how the community evolves if users learn and adapt their strategies over time;

(3) how the designed reputation protocol impacts the evolution of the community and its convergence.

Some of the above three questions have been investigated in the literature on prescriptive multi-agent learning [47]-[49], which studies how protocol designers should construct protocols that induct users to learn certain strategies to achieve certain goals. However, this strand of research mainly focuses on cooperative games where the interests of the protocol designers and the interests of the learning users are aligned [48][49], or users are compliant (obedient) and have no freedom to deviate from the prescribed learning rules [47].

Our work brings a different perspective. Unlike [47]-[49], the interests of the protocol designer (e.g. the designer of the community) and that of the self-interested users in our work are in conflict with each other. Therefore, the protocol designer has to carefully trade-off his goal to maximize the social welfare against the need for providing sufficient incentives for users to comply with the protocol. Instead of assuming that users are obedient and restricting the users' learning and adaptation behavior, we assume that users form their own beliefs about the community based on their past interactions and, using this knowledge, adapt their service strategies to maximize their individual long-term utilities. By considering the users' best response dynamics, we focus on the design of effective reputation protocols which teach users to learn and cooperate with each other in the long-run, by contributing their services to the community.

The remainder of this chapter is organized as follows. In Section 2.2, we formalize the user interaction in an online community as a repeated random matching game and introduce our proposed indirect reciprocity framework for the design of reputation protocols. Section 2.3 formulates the decision problem of individual users and then discusses the resulting evolution of the community: whether it will converge to a stationary state in the long run when time goes to infinity and to which state it will converge. To characterize the long run evolution, we introduce the concept of stochastically stable equilibrium (SSE) (Definition 3), which is used to describe the states that the community will converge to in the long run. Generally speaking, an SSE characterizes a stationary state of the community in which users are playing stationary strategies that are best responses against each other. Meanwhile, different from the concept of evolutionary stable equilibrium used in evolutionary game theory, an SSE is also stochastically stable, indicating that if the community deviates from an SSE due to stochastic perturbations, it can converge to this SSE again with probability 1 in the long-run. Understanding how a community evolves over time enables us to define the design problem of the optimal reputation protocol (Definition 4). To solve this optimal design problem, Section 3.4 studies the structures of user's best response dynamics (Proposition 1 and 2), i.e. their optimal service strategies. By utilizing these structures, Section 3.5 then solves the optimal design problem (Theorem 2). Section 3.6 presents the numerical results and Section 3.7 concludes the chapter and discusses future research directions.

3.2 System Model

3.2.1 Repeated Matching Game

We consider an online community consisting of N users, each of whom indexed by an integer $i \in \{1, \ldots, N\}$ and each possessing some resources or services which are valuable to others. The community is modeled as a discrete-time system with time divided into periods. In each period, each user generates one service request and is randomly matched with another user who can provide the requested services [40]. We model the request generation and user selection process using uniform random matching: each user receives exactly one request in every period and each user is equally

Table 3.1: Utility matrix of a gift-giving game

		Server	
		$z = 1$	$z = 0$
Client		$b, -c$	$0, 0$

likely to receive the request of a user, and the matching is independent across periods [1]. Such model well approximates the matching process between users in large-scale online communities where agents interact in an ad-hoc fashion and the interactions among users are constructed randomly over time, e.g. peer-to-peer networks [39] and online crowdsourcing platforms [42]. The matched users play a two-player stage game in which the user requesting service is called the client and the user who is being requested is called the server. Given the uniform random matching, each individual user in the community is involved in two stage games during a period, one as a server and one as a client. Also, the opponents in these two stage games are in general different. That is, a user does not request services from and supply services to the same opponent in the same period.

The stage game played by a pair of matched users is formalized as an asymmetric gift-giving game [32] where the server is the only strategic part who chooses his action z from a binary set $\mathcal{Z} = \{0, 1\}$. The utility matrix of a gift-giving game is presented in Table 3.1. If the server supplies the service (chooses $z = 1$), it incurs a cost $c > 0$ to him and the client receives a benefit $b > 0$. On the other hand, if the server refuses to supply the service (chooses $z = 0$), both users receive a utility of 0. Here we assume that $b > c$ such that the service exchange is socially valuable and the values of b and c are the same for all users. Hence, suppose a user i is matched with user j_1 in the stage game where user i is the server and is matched with user j_2 where user i is the requester, then the utility received by user i in a period t can be represented as

$$u_i^{(t)} = z_{j_2}^{(t)} b - z_i^{(t)} c, \tag{3.1}$$

[1] The impact of matching schemes on the incentive of users and the performance of online communities falls out of this chapter, but serves as an important next step in this line of research.

where $z_i^{(t)}$ and $z_{j_2}^{(t)}$ are the actions adopted by user i and user j_2 in period t, respectively. The social welfare of the community is thus defined as

$$\bar{U} = \lim_{T\to\infty} \frac{1}{TN} \sum_{t=1}^{T} \sum_{i=1}^{N} u_i^{(t)}. \tag{3.2}$$

It is obvious that $\bar{U}$ is maximized at the value $b-c$, which is called the social optimum, when each user supplies services in each period. Nevertheless, $z = 0$ is the dominant action for the server in the gift-giving game, which constitutes a Nash equilibrium. When every server chooses $z = 0$ to myopically maximize his utility in a period, an inefficient social welfare arises where every user receives a utility of 0 in each period. In the rest of this chapter, we focus on designing a reputation protocol by exploiting the repeated nature of the gift-giving game, in order to incentivize users to supply services within their self-interest and achieve the optimal social welfare $b - c$.

3.2.2 Reputation Protocols

In the repeated matching game, we consider each user to be self-interested and far-sighted. Hence, he chooses his service strategy to maximize his individual long-term utility, which is the discounted sum of his utility in the current period and that from the future periods. The decision problem of an individual user will be formally presented in Section 3.3.1. In this section, we first discuss the design of the reputation protocol in the repeated matching game. A reputation protocol regulates the behavior of a self-interested user by affecting his social status and thus, the service he receives from other users.

A user i's social status in the repeated game is represented with a reputation score $\theta_i \in \Theta$, which records the history of the user's past plays and takes values from a non-empty and finite set Θ. Since any finite history can be represented by a finite set of reputation scores and the elements in any finite set can be indexed using a finite group of integer numbers, we denote $\Theta = \{0, 1, 2, \cdots, L\}$ without loss of generality, where L represents the highest reputation score and 0 represents the lowest reputation score.

The service strategy adopted by a user i describes his choice of actions in each stage game he involves as the server and can be represented by a mapping $\sigma_i : \Theta \times \Theta \to \mathcal{Z}$. That is, in each stage game where user i's reputation is θ and his client's reputation is $\tilde{\theta}$, user i chooses the action $\sigma_i(\theta, \tilde{\theta})$. Since Θ and $\mathcal{Z}$ are finite sets, the space of service strategies that a user can choose is also finite and is denoted as $\Gamma = \{\sigma | \sigma : \Theta \times \Theta \to \mathcal{Z}\}$ with $|\Gamma| = 2^{L^2}$.

Formally, a reputation protocol κ, which is designed by the protocol/community designer, consists of a reputation set Θ, a social rule ϕ and a reputation scheme τ.

The social rule $\phi \in \Gamma$ is a service strategy that specifies the recommended behavior for users within the community. ϕ is broadcasted by the community operator (e.g. the website portal, or the tracker in P2P networks) to all users. However, it is a user's own strategic decision whether or not to comply with the recommended ϕ. In this chapter, we restrict the protocol design to threshold-based social rules, because they are simple to deploy and have been shown in [46] to be nearly optimal. A threshold-based social rule can be formalized as follows:

$$\phi(\theta, \tilde{\theta}) = \begin{cases} 1 & \text{if } \theta \geqslant H \text{ and } \tilde{\theta} \geqslant H \\ 1 & \text{if } \theta < H \\ 0 & \text{otherwise} \end{cases} . \tag{3.3}$$

Each social rule has a social threshold $H \in \{0, 1, ..., L+1\}$. It instructs a user whose reputation is no less than H, i.e. a "good user", to play discriminatively and supply services only to clients whose reputations are at least H. On the contrary, if a user has a reputation less than H, i.e. a "bad user", he has to supply services to all clients regardless of their reputations. According to the social rule, a good user also can always receive services from others upon his requests, while a bad user gets served only when he requests services from other bad users. Differential services are thus provided to users based on their reputations.

A reputation scheme τ updates a user's reputation based on his past behavior of providing services. It works in the following manner: after a server takes an action (i.e. whether or not to supply services), his client reports his action to the community operator. The operator then uses this report to update the server's reputation according to τ. Since each user is involved as the server in one stage game during each period, his reputation is also updated once per period. Formally, a reputation scheme is represented as a mapping $\tau : \Theta \times \Theta \times \Theta \times \mathcal{Z} \rightarrow [0, 1]$, where $\tau(\theta, \tilde{\theta}, \theta', z)$ is the probability that a server of reputation θ is assigned with a new reputation θ' in the next period, when the reputation of his client is $\tilde{\theta}$ and the reported action is z. To keep the protocol design simple, we consider the following reputation scheme in this chapter:

$$\tau(\theta, \tilde{\theta}, \theta', z) = \begin{cases} 1 & \text{if } z \geqslant \phi(\theta, \tilde{\theta}) \text{ and } \theta' = \min(\theta + 1, L) \\ \beta & z < \phi(\theta, \tilde{\theta}) \text{ and } \theta' = 0 \\ 1 - \beta & z < \phi(\theta, \tilde{\theta}) \text{ and } \theta' = \theta \\ 0 & \text{otherwise} \end{cases} . \tag{3.4}$$

Given this scheme, when a server is reported to deviate from the social rule ϕ by supplying less service than what is required by ϕ in a stage game, his reputation is decreased to 0 with a probability β and remains unchanged with a probability $1 - \beta$. Otherwise, his reputation is increased by 1 while not exceeding the maximum reputation L. The protocol designer can thus adjust the strength of punishment imposed upon user deviations by designing the value of β, where a higher β indicates a stronger punishment. We assume that the reputation scheme τ (including β) is also broadcasted or posted by the community operator and known by all users. A schematic representation of the reputation protocol is presented in Figure 3.1.

During its operation, we assume that the community experiences stochastic shocks, which we model as small stochastic perturbations [44]. In practice, such perturbations arise due to various types of operation errors. For example, a client is unsatisfied with the received service (due to some network errors) even if the server plays $z = 1$. To model such errors we denote $Prob(z_{rpt}|z)$ as the conditional probability that the client reports z_{rpt} while the server plays z and we assume that a server's reported action is inconsistent with his actual action with a probability ε, i.e. $Prob(z_{rpt} = 0|z = 1) = \varepsilon$

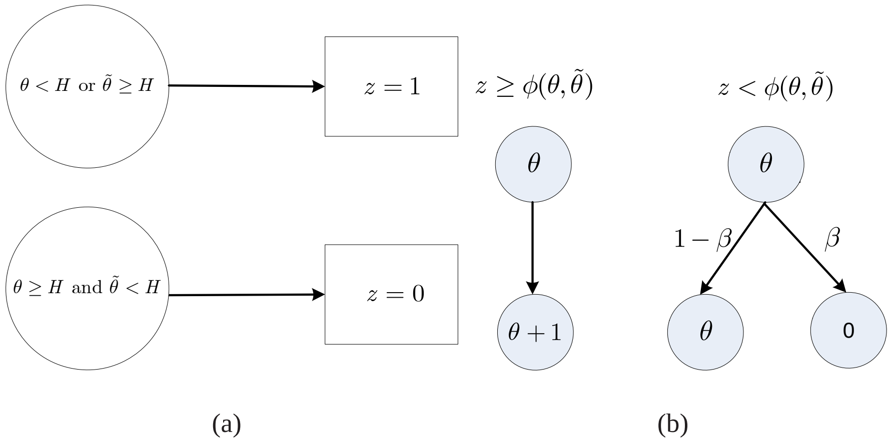


Figure 3.1: The schematic representation of the social rule and the reputation scheme

and $Prob(z_{rpt} = 1|z = 0) = \varepsilon$ [2]. Hence, a server is reported to comply with the social rule with probability ε when he deviates, and vice versa. Since the stochastic perturbations take place rarely in the repeated game and thus the value of ε is small and difficult to be accurately measured, we focus on the extreme case of $\varepsilon \to 0$ in the design of reputation protocols, which will be formalized in Section 3.3.2.

To summarize, the repeated matching game in each period can be characterized by a strategy profile $\boldsymbol{\sigma} = (\sigma_i)_{i=1}^N$ representing the service strategies adopted by all users and a reputation profile $\boldsymbol{\theta} = (\theta_i)_{i=1}^N$ representing the reputations of all users in each period. The reputation protocol considered in this chapter can be parameterized as $\kappa = (L, H, \beta)$, where L represents the length of history recording a server's past behavior, H indicates how to differentiate users, and β determines the strength of punishment imposed upon the user deviations. In the rest of this chapter, we first discuss the evolution of $\boldsymbol{\sigma}$ and $\boldsymbol{\theta}$ in the repeated matching game and formulate the design problem of the optimal reputation protocol in Section 3.3, which induces all users to cooperate, i.e. find it in their self-interest to supply services in each period. In Section 3.4, we characterize the structure of users' best response (i.e. the optimal service strategies adopted in each period). Then in Section 3.5, we solve the optimal protocol design, i.e.

[2] It should be noted that our analysis straightforwardly applies to the scenario where error probabilities are asymmetric, i.e. $Prob(z_{rpt} = 0|z = 1) \neq Prob(z_{rpt} = 1|z = 0)$. We assume symmetric error probabilities only for the simple illustrations.

to determine the optimal reputation protocol κ that induces the full cooperation among users in the long run.

3.3 Community Evolution and Stochastic Stable Equilibrium

In this section, we formalize the decision problem of individual self-interested users, and then analyze the evolution of the community in the long-run. Finally, we formalize the considered protocol design problem.

3.3.1 User's Decision Problem

We begin by investigating a typical user i's decision problem in the repeated matching game under the reputation protocol κ. Users are assumed to be far-sighted. Each of them adapts his strategy in order to maximize his expected long-term utility, i.e. his expected sum utilities from the current period to the infinite future. By analyzing the expected one-period and long-term utilities in each period, we are able to formalize an individual users decision problem.

A user i's expected one-period utility is the sum of his expected utilities received in the stage games in which he is involved during one period [3]. Due to the uniform random matching, the expected one-period utility depends on the user's own reputation and strategy as well as the distribution of other users' reputations and strategies. In the rest of this subsection, we first discuss each user's knowledge and belief about other users' reputations and strategies, and then construct the user's utility function and his decision problem.

The reputation distribution of the community in each period is referred to as the community state of this period and denoted by a vector $\boldsymbol{\mu} = (\mu(\theta))_{\theta=0}^{L}$, where $\mu(\theta)$ represents the number of users of reputation θ. Given the reputation profile $\boldsymbol{\theta} = (\theta_i)_{i=1}^{N}$,

[3] It should be noted that a user's expected one-period utility is the utility he expects to receive at the beginning of a period and thus is different from his realized one-period utility (3.1).

the corresponding community state is derived as $\mu(\theta) = \sum_{i=0}^{N} I(\theta_i = \theta),\ \forall\theta$. For better illustration, we use $\boldsymbol{\mu}^{\boldsymbol{\theta}}$ to denote the community state that is derived from the reputation profile $\boldsymbol{\theta}$. It should be noted that there are multiple reputation profiles that can lead to the same community state. Hence, we also define $\Lambda_\mu = \{\boldsymbol{\theta}|\boldsymbol{\mu}^{\boldsymbol{\theta}} = \boldsymbol{\mu}\}$ as the set of reputation profiles that lead to the community state $\boldsymbol{\mu}$. We assume that the community state is broadcasted by the community operator at the beginning of each period and hence is known by all users. Given the community state, a user i then calculates the reputation distribution of all users other than himself, which is referred to as his opponent state and denoted as a vector $\boldsymbol{\eta}_i = \{\eta_i(\theta')\}_{\theta'=0}^{L}$, which can be derived as $\eta_i(\theta') = \mu(\theta')$ for all $\theta' \neq \theta_i$ and $\eta_i(\theta_i) = \mu(\theta_i) - 1$ [4]. We also use $\boldsymbol{\eta}_i^{\theta}$ to denote the user i's opponent state under the reputation profile $\boldsymbol{\theta}$.

Different from the reputation, the strategy adopted by an individual user is private information and cannot be perfectly observed by either the community operator or other users. Hence, each user i has to maintain a belief on other users' strategies. In this chapter, user i forms a simple belief represented by a value $\rho_i(\theta_i)$, which denotes the probability that he expects to receive services in one period. Given his opponent state, the belief is calculated as:

$$\rho_i(\theta_i) = \begin{cases} \sum\limits_{\theta'=1}^{H-1} \eta_i(\theta')/(N-1) & \text{if } \theta_i < H \\ 1 - \eta_i(0)/(N-1) & \text{if } \theta_i \geqslant H \end{cases}. \tag{3.5}$$

Equivalently, we are assuming that user i believes that any user of reputation 0 will not supply services while any user of reputation $\theta > 0$ complies with the social rule.

Remark: This belief model is based on the assumption that each user implicitly conjectures that all other users are playing stationary strategies which do not change over time. This assumption is commonly adopted in the related literature (e.g. the

[4]It should be noted that there is a deterministic mapping from the community state and the opponent state and they can be derived from each other. The reason why we introduce the opponent state is for the better illustration, since a user can never be matched with himself and hence it is easier to express a user's expected one period utility with his opponent state.

frequency-based belief update model in fictitious play [51] as well as in stochastic evolutionary games [44][50]). Hence, the current strategy that is observed as being played by a user is believed to be played by him in the future. Starting from this presumption, each user believes that a user of reputation 0, which indicates his deviation in the previous period, will deviate again in the current period. For a user of reputation $\theta > 0$, it is unknown whether this user deviated in the previous period or not (due to the existence of random punishment in (3.4)). Therefore, other users simply believe that this user complied with the social rule in the previous period and will continue his compliance in the current period.

Based on this belief, the expected utility of a user i in one period can be expressed as a function of his current strategy σ_i, his current reputation θ_i, and his current opponent state $\boldsymbol{\eta}_i$:

$$v_i(\theta_i|\sigma_i, \boldsymbol{\eta}_i) = \rho_i(\theta_i)b - \sum_{\tilde{\theta}=0}^{L} \frac{\eta_i(\tilde{\theta})}{N-1} c(\theta_i, \tilde{\theta}, \sigma(\theta_i, \tilde{\theta})). \tag{3.6}$$

Here $c(\theta_i, \tilde{\theta}, z)$ is the one-period cost incurred by user i when he plays action z and his matched client has a reputation $\tilde{\theta}$.

A user's expected long-term utility at any period is evaluated as the discounted sum of his expected one-period utilities starting from this period over the infinite time horizon. When computing his expected long-term utility, each user i assumes that his opponent state $\boldsymbol{\eta}_i$ in the current period does not change in the future. In other words, each user does not try to learn the evolution of his opponent state while making his decision in each period. This assumption is reasonable given the fact that users in the online community normally have bounded rationality [5].

Given all the knowledge and beliefs, a user i's expected long-term utility is given

[5] If a user i has the capability to infer the evolution of his opponent state $\boldsymbol{\eta}_i$ according to his belief $\rho_i(\theta_i)$ and incorporates such inference into his expected long-term utility, it can be proven that the computation complexity of his optimal strategy in each period is at the order of $O(N^L)$, which quickly becomes intractable even for a moderate value of N.

by

$$V_i(\theta_i|\sigma_i, \boldsymbol{\eta}_i, \kappa) = v_i(\theta_i|\sigma_i, \boldsymbol{\eta}_i) + \delta \sum_{\theta'} p_i(\theta'|\theta_i, \sigma_i, \boldsymbol{\eta}_i, \kappa) V_i(\theta'|\sigma_i, \boldsymbol{\eta}_i, \kappa), \tag{3.7}$$

where $\delta \in [0,1)$ is the discount factor and $\{p_i(\theta'|\theta, \sigma_i, \boldsymbol{\eta}_i, \kappa)\}$ is the probability that user i expects his reputation changes from θ to θ' if he plays the strategy σ_i, while the reputation protocol is κ and the opponent state is $\boldsymbol{\eta}_i$ [6].

User i's best-response play in each period is to select an optimal strategy that maximizes his expected long-term utility $V_i(\theta_i|\sigma_i, \boldsymbol{\eta}_i, \kappa)$. Due to the recursive structure of the expected long-term utility (3.7), the optimal strategy in each period can be obtained by solving the following optimization problem

$$\max_{\sigma_i \in \Gamma} V_i(\theta|\sigma_i, \boldsymbol{\eta}_i, \kappa), \forall \theta \in \Theta \tag{3.8}$$

by using well-known dynamic programming algorithms such as value iteration [7]. It should be noted that user i has to optimize his expected long-term utilities starting from any reputation θ, even if his current reputation $\theta_i \neq \theta$. This is because of the recursive structure of the expected long-term utilities and the resulting fact that $V_i(\theta_i|\sigma_i, \boldsymbol{\eta}_i, \kappa)$ will be influenced by $\{V_i(\theta|\sigma_i, \boldsymbol{\eta}_i, \kappa)\}_{\theta \neq \theta_i}$ and thus $\{\sigma_i(\theta, \cdot)\}_{\theta \neq \theta_i}$. To show the dependency of the solution of (3.8) on the reputation protocol and the opponent state, it is denoted as $\sigma_i^{\kappa, \boldsymbol{\eta}_i}$, with the corresponding optimal long-term utilities being represented as $\{V_i^{\kappa, \boldsymbol{\eta}_i}(\theta)\}_{\theta=0}^{L}$. Importantly, it should be noted that because users having the same reputation observe the same opponent state, they have the same optimal strategy and optimal long-term utilities, i.e. $\sigma_i^{\kappa, \boldsymbol{\eta}_i} = \sigma_j^{\kappa, \boldsymbol{\eta}_j}$ and $V_i^{\kappa, \boldsymbol{\eta}_i}(\theta) = V_j^{\kappa, \boldsymbol{\eta}_j}(\theta), \forall \theta$ if $\theta_i = \theta_j$. In Section 3.4, we will unravel the structural properties of $\sigma_i^{\kappa, \boldsymbol{\eta}_i}$.

[6] Here we assume that users are not aware the existence of stochastic perturbations and do not incorporate it into the calculation of $p_i(\theta'|\theta, \sigma_i, \boldsymbol{\eta}_i, \kappa)$. One explanation to this assumption is that the error probability is negligible to users when $\varepsilon \to 0$ since stochastic perturbations take place infrequently in the community.

[7] To break the tie when the expected long-term utilities by playing $z = 0$ and $z = 1$ are the same, we assume that the user always chooses $z = 1$ in this case.

3.3.2 Long-run Evolution and Stochastic Stable Equilibrium

In this section, we examine how the community state $\boldsymbol{\mu}$ evolves over time, under the users' best response dynamics introduced by (3.8), and then formalize the design problem of the optimal reputation protocol.

The evolution of $\boldsymbol{\mu}$ can be characterized by the following nonlinear stochastic difference equation:

$$\mu^{(t+1)}(\theta) = \begin{cases} x_+^{(t)}(\theta-1) + x_{un}^{(t)}(\theta) + x_+^{(t)}(\theta) & \text{if } \theta = L \\ x_+^{(t)}(\theta-1) + x_{un}^{(t)}(\theta) & \text{if } \theta \in \{1, ...L-1\} \\ x_{un}^{(t)}(\theta) + \sum_{\theta'=0}^{L} x_-^{(t)}(\theta') & \text{if } \theta = 0 \end{cases} . \tag{3.9}$$

Here t is the time index. $x_+^{(t)}(\theta)$ and $x_-^{(t)}(\theta)$ are random variables representing the numbers of users whose reputations change, in period t, from θ to $\theta+1$ and $\theta-1$, respectively; while $x_{un}^{(t)}(\theta)$ representing the number of users whose reputations remains unchanged. All of these variables follow binomial distributions:

$$\begin{array}{lcl} x_+^{(t)}(\theta) & - & Bin(\mu^{(t)}(\theta), q_{reward}^{(t)}(\theta)) \\ x_-^{(t)}(\theta) & - & Bin(\mu^{(t)}(\theta), \beta q_{punish}^{(t)}(\theta)) \\ x_{un}^{(t)}(\theta) & - & Bin(\mu^{(t)}(\theta), (1-\beta) q_{punish}^{(t)}(\theta)) \end{array} . \tag{3.10}$$

where $q_{reward}^{(t)}(\theta)$ and $q_{punish}^{(t)}(\theta)$ are the probabilities that a user of reputation θ is reported to be complying with and deviating from the social rule respectively, when he adopts his optimal strategy in period t. According to (3.8), a user's optimal strategy in each period is implicitly determined by the community state $\boldsymbol{\mu}$ and his reputation θ and hence, both $q_{reward}^{(t)}(\theta)$ and $q_{punish}^{(t)}(\theta)$ are also determined by $\boldsymbol{\mu}$ and θ. Also, $q_{reward}^{(t)}(\theta)$ and $q_{punish}^{(t)}(\theta)$ are implicitly influenced by the error probability ε.

The dynamic system (3.9) defines a Markov chain on the finite space

$$\mathcal{U} = \left\{ \boldsymbol{\mu} = (\mu(\theta))_{\theta=0}^{L} \middle| \mu(\theta) \in \mathbb{N} \; for \; 0 \leqslant \theta \leqslant L, \sum_{\theta=0}^{L} \mu(\theta) = N \right\}, \tag{3.11}$$

with the transition probability matrix being denoted as $P_{\varepsilon,\kappa} = [p_{\varepsilon,\kappa}(\boldsymbol{\mu}'|\boldsymbol{\mu})]_{\boldsymbol{\mu},\boldsymbol{\mu}' \in \mathcal{U}}$. Note that with $\varepsilon > 0$, this Markov chain is communicating. Hence, it is well-known from

[54] that this Markov chain is also irreducible and aperiodic, which introduces a unique stationary distribution over $\mathcal{U}$. Let

$$\Delta_{|\mathcal{U}|} \triangleq \left\{ \boldsymbol{\omega} = (\omega(\boldsymbol{\mu}))_{\boldsymbol{\mu}\in\mathcal{U}} \in \mathcal{R}^{|\mathcal{U}|} | \omega(\boldsymbol{\mu}) \geqslant 0 \; for \; \forall \boldsymbol{\mu} \in \mathcal{U} \; and \; \sum_{\boldsymbol{\mu}\in\mathcal{U}} \omega(\boldsymbol{\mu}) = 1 \right\} \tag{3.12}$$

be the $|\mathcal{U}|$-dimensional simplex, a stationary distribution is a row vector $\boldsymbol{\omega}_{\varepsilon,\kappa}$ satisfying $\boldsymbol{\omega}_{\varepsilon,\kappa} = \boldsymbol{\omega}_{\varepsilon,\kappa} P_{\varepsilon,\kappa}$. When this Markov chain in communicating, irreducible and aperiodic, $\boldsymbol{\omega}_{\varepsilon,\kappa}$ satisfies the following properties.

Lemma 1. $\boldsymbol{\omega}_{\varepsilon,\kappa}$ preserves the following properties:

(1) Stability: For any $\boldsymbol{\omega} \in \Delta_{|\mathcal{U}|}$, we have that $\lim\limits_{t\to\infty} \boldsymbol{\omega} P_{\varepsilon,\kappa}^t \to \boldsymbol{\omega}_{\varepsilon,\kappa}$.

(2) Ergodicity: Starting from an arbitrary community state $\boldsymbol{\mu}^{(0)}$, the fraction of time that the community stays at a particular community state $\boldsymbol{\mu}$ in the long run is $\boldsymbol{\omega}_{\varepsilon,\kappa}(\boldsymbol{\mu})$, i.e. $\lim\limits_{T\to\infty} \frac{1}{T} \sum\limits_{t=1}^{T} I(\boldsymbol{\mu}^{(t)} = \boldsymbol{\mu}) = \boldsymbol{\omega}_{\varepsilon,\kappa}(\boldsymbol{\mu}),\ \forall \boldsymbol{\mu} \in \mathcal{U}$, where $I(\cdot)$ is the indicator function that takes the value 1 when $\boldsymbol{\mu}^{(t)} = \boldsymbol{\mu}$.

Proof: This lemma is proved in [44] and is omitted here. ■

The stability shows that independent of the initial community state, the asymptotic distribution of the community state in the long run is always given by $\boldsymbol{\omega}_{\varepsilon,\kappa}$. The ergodicity indicates that the proportion of time that the community spends in each state is equivalent to the corresponding term of this state in the stationary distribution $\boldsymbol{\omega}_{\varepsilon,\kappa}$.

When $\varepsilon \to 0$, the resulting stationary distribution is referred to as the limit distribution and is defined as follows.

Definition 1 (Limit Distribution). The limit distribution is defined by

$$\bar{\boldsymbol{\omega}}_{\kappa} = \lim_{\varepsilon\to 0} \boldsymbol{\omega}_{\varepsilon,\kappa}. \tag{3.13}$$

Due to the fact that each element in $P_{\varepsilon,\kappa}$ is a polynomial of ε and is thus continuous on ε, it could be concluded that the limit distribution also satisfies the following condition

$$\bar{\boldsymbol{\omega}}_{\kappa} P_{0,\kappa} = \bar{\boldsymbol{\omega}}_{\kappa}. \tag{3.14}$$

Therefore, the limit distribution is also a stationary distribution over $\mathcal{U}$ when $\varepsilon = 0$ [8].

The limiting distribution characterizes the asymptotic stationary distribution of the community state over the space $\mathcal{U}$ when the error probability is sufficiently small. The existence and the uniqueness of $\bar{\boldsymbol{\omega}}_\kappa$ has been proven in [44]. A community state that receives a positive probability in $\bar{\boldsymbol{\omega}}_\kappa$ are called a stochastically stable state under the reputation protocol κ, which is defined as follows.

Definition 2 (Stochastically Stable State). A community state $\boldsymbol{\mu}$ is a stochastically stable state under a reputation protocol κ if and only if $\bar{\boldsymbol{\omega}}_\kappa(\boldsymbol{\mu}) > 0$.

According to Definition 2, given a reputation protocol κ, the community stays at a stochastically stable state with a positive fraction of time in the long run. On the contrary, if a community state is not stochastically stable, the fraction of time that the community stays at this state approaches to 0 when $\varepsilon \to 0$. Let $C(\boldsymbol{\omega})$ be the carrier of a probability distribution $\boldsymbol{\omega} \in \Delta_{|\mathcal{U}|}$, i.e. $C(\boldsymbol{\omega}) = \{\boldsymbol{\mu} \in \mathcal{U} | \boldsymbol{\omega}(\boldsymbol{\mu}) > 0\}$, the set of stochastically stable states under the reputation protocol κ can be denoted as $C(\bar{\boldsymbol{\omega}}_\kappa)$.

Therefore, starting from an arbitrary initial state, the community converges with probability 1 to states that belong to $C(\bar{\boldsymbol{\omega}}_\kappa)$ in the long run.

We then define the equilibrium concept in this chapter, which is referred to as s-tochastically stable equilibrium (SSE) [44].

Definition 3 (Stochastically Stable Equilibrium). A strategy profile $\boldsymbol{\sigma}^\# = (\sigma_i^\#)_{i=1}^N$ and a reputation profile $\boldsymbol{\theta}^\# = (\theta_i^\#)_{i=1}^N$ form an SSE under a reputation protocol κ if and only if

(1) The community state under $\boldsymbol{\theta}^\#$ belongs to $C(\bar{\boldsymbol{\omega}}_\kappa)$, i.e. $\boldsymbol{\mu}^{\boldsymbol{\theta}^\#} \in C(\bar{\boldsymbol{\omega}}_\kappa)$.

(2) For each user i, $\sigma_i^\#$ is his best response, i.e. $\sigma_i^\# = \sigma_i^{\kappa, \boldsymbol{\eta}_i^{\boldsymbol{\theta}^\#}}$.

According to this definition, each SSE introduces a community state which is s-

[8] It is worth to be noted that there might be multiple stationary distributions under the same reputation protocol when $\varepsilon = 0$, while the limit distribution is always unique.

tochastically stable, and conversely, for each stochastically stable state $\boldsymbol{\mu} \in C(\bar{\boldsymbol{\omega}}_\kappa)$, there is at least one SSE that introduces it. Let $E = \{(\boldsymbol{\sigma}^\#, \boldsymbol{\theta}^\#)|(\boldsymbol{\sigma}^\#, \boldsymbol{\theta}^\#)\ is\ an\ SSE\}$ denote the set of stochastically stable equilibria, we have the following lemma.

Lemma 2. (1) For each $(\boldsymbol{\sigma}^\#, \boldsymbol{\theta}^\#) \in E$, $\boldsymbol{\mu}^{\boldsymbol{\theta}^\#} \in C(\bar{\boldsymbol{\omega}}_\kappa)$; (2) $C(\bar{\boldsymbol{\omega}}_\kappa) = \{\boldsymbol{\mu}^{\boldsymbol{\theta}^\#}|(\boldsymbol{\sigma}^\#, \boldsymbol{\theta}^\#) \in E,\ \exists \boldsymbol{\sigma}^\#\}$.

Proof: This lemma is straightforward according to Definition 3. ■

The protocol designer's problem is to design a reputation protocol that induces users to cooperate in all SSE, which can be formally defined as follows:

Definition 4 (Protocol Design Problem). The protocol designer's goal is to select a protocol κ such that for any SSE $(\boldsymbol{\sigma}^\#, \boldsymbol{\theta}^\#) \in E$, the following is satisfied

$$\sigma_i^\#(\boldsymbol{\theta}_i^\#, \boldsymbol{\theta}_j^\#) = 1,\ \forall i, j \in \{1, ..., N\}\ and\ i \neq j. \tag{3.15}$$

Since the community stays at stochastically stable states with probability 1 in the long run, full cooperation among users is induced in the community and hence the community operates at its highest efficiency with the optimal social welfare $\bar{U} = b - c$.

3.4 Optimal User Strategy

In this section, we investigate the structure of a typical user i's optimal strategy, i.e. $\sigma_i^{\kappa,\boldsymbol{\eta}_i}$. Since users are homogeneous with respect to the stage game benefit b and cost c as well as the discount factor δ. The results derived for user i in this section is applicable to the optimal strategy of any user in the community. The proofs in this section heavily rely on the one-shot deviation principle of discounted dynamic programming, which is proven in [50]. Below we briefly explain the basic idea of the one-shot deviation principle [9].

[9]There is a little abuse of notations here since the terms "state" and "action" are used to describe the Markov decision process. It should be noted that the state and action here are generic and are different from the community state and user action in this chapter.

Given a discrete-time Markov decision process with the state space S, the action space A, the state transition probability $p(s'|s,a)$, the policy $\pi(s,a)$, the one-period reward function $r(s,a)$ and the value function $V^{\pi}(s)$, suppose a user is optimal policy is π^* with the corresponding optimal value function V^{π^*}. Starting from any period t when user i is at state $s^{(t)}$, then user i cannot increase his long-term utility by deviating from π^* in this period and continues to comply with π^* in the future, i.e.

$$V^{\pi^*}(s^{(t)}) \geqslant r(s^{(t)},a) + \sum_{s' \in S} p(s'|s,a) V^{\pi^*}(s'),\ \forall t \text{ and } \forall a \neq \pi^*(s). \tag{3.16}$$

Mapping the general MDP formulation above into our problem, we have $s = \theta$, $a = \sigma(\theta,\cdot)$ and $\pi = \sigma$ for each individual user. Hence, the optimization (3.8) is equivalent to the derivation of the optimal policy of an MDP problem for each individual user. Using this idea, we have the following lemma.

Lemma 3. For any $\theta, \tilde{\theta} \in \Theta$, $\sigma_i^{\kappa,\boldsymbol{\eta}_i}(\theta,\tilde{\theta}) = 0$ if $\phi(\theta,\tilde{\theta}) = 0$.

Proof: First, it is easy to observe that given the social rule (3.3) and the user belief (3.5), we have $V_i(\theta|\sigma_i,\boldsymbol{\eta}_i,\kappa) \leqslant V_i(\theta+1|\sigma_i,\boldsymbol{\eta}_i,\kappa), \forall \theta \in \{0,...,L-1\}$. Suppose that for a user i, we have $\sigma_i^{\kappa,\boldsymbol{\eta}_i}(\theta',\tilde{\theta}') = 1$ for some $(\theta',\tilde{\theta}')$ with $\phi(\theta',\tilde{\theta}') = 0$. If user i chooses to play $z = 0$ at reputation θ' when he meets a client of reputation $\tilde{\theta}'$ (i.e. adopting a one-shot deviation from $\sigma_i^{\kappa,\boldsymbol{\eta}_i}$), he receives a higher utility for the current stage game because he saves the cost of supplying services in this stage game, while his future utility remaining unchanged because $\phi(\theta',\tilde{\theta}') = 0$ and he will not be punished in this period regardless of his choice on z. Therefore, $V_i(\theta'|\sigma_i,\boldsymbol{\eta}_i,\kappa)$ increases if user i chooses $\sigma_i(\theta',\tilde{\theta}') = 0$. According to the one-shot deviation principle, $\sigma_i^{\kappa,\boldsymbol{\eta}_i}$ thus cannot be optimal. Hence, we should have $\sigma_i^{\kappa,\boldsymbol{\eta}_i}(\theta',\tilde{\theta}') = 0$. ■

The basic idea of Lemma 3 is that by supplying more services than required by the social rule, user i's expected cost in the current period increases while his expected future utility does not change. Hence, his expected long-term utility is always maximized by choosing $\sigma_i^{\kappa,\boldsymbol{\eta}_i}(\theta,\tilde{\theta}) = 0$.

Using Lemma 3, we first prove that $\sigma_i^{\kappa,\boldsymbol{\eta}_i}$ is always threshold-based for any κ and

$\boldsymbol{\eta}_i$.

Proposition 1. Given κ, $\boldsymbol{\eta}_i$, there exists a set of service thresholds $(h_{\kappa,\boldsymbol{\eta}_i}(\theta))_{\theta=0}^{L} \in \mathbb{N}^{L+1}$ such that $\sigma_i^{\kappa,\boldsymbol{\eta}_i}$ preserves the following property for any $\theta \in \Theta$:

$$\sigma_i^{\kappa,\boldsymbol{\eta}_i}(\theta,\tilde{\theta}) = \begin{cases} 1 & \text{if } \tilde{\theta} \geqslant h_{\kappa,\boldsymbol{\eta}_i}(\theta) \\ 0 & \tilde{\theta} < h_{\kappa,\boldsymbol{\eta}_i}(\theta) \end{cases}. \tag{3.17}$$

Proof: See Appendix A. ■

Remark: Proposition 1 proves that the optimal strategy always instructs user i, at any reputation θ, to supply services only to clients of reputations higher than or equal to a threshold $h_{\kappa,\boldsymbol{\eta}_i}(\theta)$. This is because if user i supplies services to a client of reputation $\tilde{\theta}$ in his optimal strategy, i.e. $\sigma_i^{\kappa,\boldsymbol{\eta}_i}(\theta,\tilde{\theta}) = 1$, we should also have $\sigma_i^{\kappa,\boldsymbol{\eta}_i}(\theta,\tilde{\theta}') = 1$ for any $\tilde{\theta}' > \tilde{\theta}$, since serving a client with a higher reputation incurs the same cost as serving a client with a lower reputation but always brings a user higher future benefit.

In the following proposition, we further analyse the property of the service thresholds $\{h_{\kappa,\boldsymbol{\eta}_i}(\theta)\}_{\theta=0}^{L}$ that are derived in Proposition 1.

Proposition 2. For any κ and $\boldsymbol{\eta}_i$, (1) $h_{\kappa,\boldsymbol{\eta}_i}(\theta) \in \{H, L+1\}$, if $\theta \geqslant H$; (2) $h_{\kappa,\boldsymbol{\eta}_i}(\theta) \in \{0, L+1\}$, if $\theta < H$.

Proof: See Appendix B. ■

Remark: Proposition 2 proves that the optimal choice of each user at any reputation θ is binary: he either does not provide any service with $\sigma_i^{\kappa,\boldsymbol{\eta}_i}(\theta,\tilde{\theta}) = 0,\ \forall\tilde{\theta} \in \Theta$, or fully complies with the social rule, i.e. $\sigma_i^{\kappa,\boldsymbol{\eta}_i}(\theta,\tilde{\theta}) = \phi(\theta,\tilde{\theta}),\ \forall\tilde{\theta} \in \Theta$. Hence, Proposition 1 and 2 significantly simplify the set of strategies that can possibly serve as a user's optimal strategy, thereby facilitating the design of the optimal reputation protocol in the next section.

3.5 Optimal Reputation Protocol

This section employs the structural results obtained in Section 3.4 to solve the problem of optimal reputation protocol design. According to Lemma 2, the design problem (3.15) could be decomposed into two steps: (1) determine the set of stochastically stable states $C(\bar{\boldsymbol{\omega}}_\kappa)$; (2) determine whether full cooperation is induced under each $\mu \in C(\bar{\boldsymbol{\omega}}_\kappa)$, i.e. whether $\{\sigma_i^{\kappa,\boldsymbol{\eta}_i^\theta}\}_{i=1}^N$ satisfies (3.15) for all $\theta \in \Lambda_{\boldsymbol{\mu}}$.

We first prove that each stochastically stable state is an invariant state, which is defined as follows.

Definition 5 (Invariant State). A community state $\boldsymbol{\mu}$ is a invariant state under a reputation protocol κ if and only if $p_{0,\kappa}(\boldsymbol{\mu}|\boldsymbol{\mu}) = 1$.

Hence, an invariant state is an absorbing state in the Markov chain induced by users' best response dynamics when the error probability $\varepsilon = 0$. When $\varepsilon \to 0$, the community will stay at an invariant state once it enters this state until stochastic perturbations occur. It is important to note that in general, a stochastically stable state does not necessarily to be invariant according to (3.14). A simple example would be a cycle consisting of two states $\boldsymbol{\mu}$ and $\boldsymbol{\mu}'$ such that $p_{0,\kappa}(\boldsymbol{\mu}'|\boldsymbol{\mu}) = 1$ and $p_{0,\kappa}(\boldsymbol{\mu}|\boldsymbol{\mu}') = 1$. These two states are not invariant but could be stochastically stable. Such cycle is defined as an absorbing class as follows.

Definition 6 (Absorbing Class). A absorbing class G is a set of community states such that $p_{0,\kappa}(\boldsymbol{\mu}'|\boldsymbol{\mu}) > 0, \forall \boldsymbol{\mu}, \boldsymbol{\mu}' \in G$ and $p_{0,\kappa}(\boldsymbol{\mu}'|\boldsymbol{\mu}) = 0, \forall \boldsymbol{\mu} \in G$ and $\forall \boldsymbol{\mu}' \notin G$.

Nevertheless, the following theorem proves that in the repeated random matching game discussed in this chapter, each stochastically stable state is also invariant. Hence, there will be no cycle of states existing in the limit distribution carrier $C(\bar{\boldsymbol{\omega}}_\kappa)$. The main idea to prove this is to show that for any stationary distribution $\boldsymbol{\omega}$ such that $\boldsymbol{\omega} P_{0,\kappa} = \boldsymbol{\omega}$, each state $\boldsymbol{\mu} \in C(\boldsymbol{\omega})$ is invariant. Since we have $\bar{\boldsymbol{\omega}}_\kappa P_{0,\kappa} = \bar{\boldsymbol{\omega}}_\kappa$, each stochastically stable state is thus also invariant.

Theorem 1. Each $\boldsymbol{\mu} \in C(\bar{\boldsymbol{\omega}}_\kappa)$ is an invariant state.

Proof: See the online appendix [56]. ■

As a by-product of Theorem 1, we have the following corollary.

Corollary 1. A community state $\boldsymbol{\mu} = \{\boldsymbol{\mu}(0), \cdots, \boldsymbol{\mu}(L)\}$ that is stochastically stable preserves the following property

$$\boldsymbol{\mu}(\theta) = 0, \forall\theta \in \{1, ..., L-1\}. \tag{3.18}$$

Proof: Any state $\boldsymbol{\mu}'$ that does not satisfy (3.18) cannot be invariant. Hence, it is also not stochastically stable according to Theorem 1. ■

Corollary 1 shows that there are only $N+1$ states that could possibly be stochastically stable. For better illustration, we define a set $\mathcal{U}_B = \{\boldsymbol{\mu}|\boldsymbol{\mu}\ satisfies\ (3.18)\}$ and index these $N+1$ states in $\mathcal{U}_B$ according to the value $\boldsymbol{\mu}(L)$. A state $\boldsymbol{\mu}_B^k$ has $\boldsymbol{\mu}_B^k(L) = k$ and $\boldsymbol{\mu}_B^k(0) = N-k$ with $k \in \{0, ..., N\}$. In the next corollary, we prove that full cooperation cannot be induced in any state $\boldsymbol{\mu}_B^k$ with $k < N$.

Corollary 2. For each $\boldsymbol{\mu} \in C(\bar{\boldsymbol{\omega}}_\kappa)$ and for each reputation profile $\boldsymbol{\theta} \in \Lambda_{\boldsymbol{\mu}}$, the following properties hold for $\{\sigma_i^{\kappa,\boldsymbol{\eta}_i^{\boldsymbol{\theta}}}\}_{i=1}^N$:

$$h_{\kappa,\boldsymbol{\eta}_i^{\boldsymbol{\theta}}}(\theta_i) = L+1,\ if\ \theta_i = 0, \tag{3.19}$$

$$h_{\kappa,\boldsymbol{\eta}_i^{\boldsymbol{\theta}}}(\theta_i) = H,\ if\ \theta_i = L. \tag{3.20}$$

Proof: If (3.19) is not satisfied for some user i. Then there is always a positive probability that θ_i changes from 0 to 1 after one period and hence, $\boldsymbol{\mu}$ is not invariant. Similarly, if (3.20) is not satisfied for some user j, $\boldsymbol{\mu}$ is not invariant. ■

Using Corollary 2, the design problem (3.15) can be "translated" into the design of κ with which $C(\bar{\boldsymbol{\omega}}_\kappa)$ contains a unique state $\boldsymbol{\mu}_B^N$ with $\mu_B^N(\theta) = 0$, $\forall\theta < L$ and $\mu_B^N(L) = N$. This is solved in the next theorem.

Theorem 2. (1) If $\delta \geqslant \frac{c}{b}$, the optimal reputation protocol that solves the design problem (3.15) is $L = H = 1$, and $\beta \in [\frac{c}{b\delta}, \beta_{critical})$, where $\beta_{critical}$ is the solution of the following equation:

$$\frac{(1-\delta)c}{\delta(b-\beta_{critical}c)} + \frac{(1-\delta\beta_{critical})c}{\delta\beta_{critical}(b-c)} = 1. \tag{3.21}$$

(2) If $\delta < \frac{c}{b}$, there is no reputation protocol that can induce full cooperation among users in the long run.

Proof: See the online appendix [56]. ■

Remark: Theorem 2 provides various insights into the optimal protocol design. First, a two-level reputation protocol with $L = H = 1$ is sufficient to achieve the optimality and solve the design problem (3.15) [10]. When both L and H are larger than 1, we would always have $\boldsymbol{\mu}_N^0$ being a stochastically stable state, since it takes only one stochastic perturbation for a good user to become bad and be punished, whereas it takes multiple stochastic perturbations (proportional to H) for a bad user to become a good user again. Another important insight is that the punishment strength β cannot be too small or too large. When $\beta < c/b\delta$, a good user does not have sufficient incentives to comply with the social rule because the enforced punishment (i.e. the probability that his reputation is degraded to 0 upon his deviation from the social rule) is small. In this case, the state $\boldsymbol{\mu}_B^N$ is not invariant and stochastically stable. When $\beta \geqslant \beta_{critical}$, the enforced punishment is too severe and makes the state $\boldsymbol{\mu}_N^0$ stochastically stable and full cooperation among users cannot be sustained in the long run.

Remark: It should also be noted that the optimal design of β is a continuous region instead of a particular value. This is because we focus on designing the reputation protocol under which $\boldsymbol{\mu}_B^N$ is the unique stochastically stable state and as a result, there can be a range of β which could satisfy this requirement.

[10] However it should be noted that the reputation protocol with $L = H = 1$ is not the only optimal protocol. In fact according to the proof of Theorem 2, if $\kappa = \{1, 1, \beta\}$ is optimal, any protocol $\kappa' = \{L, 1, \beta\}$ is also optimal.

3.6 Numerical Examples

In our experiments, we illustrate the long run performance of the reputation protocol designed in Theorem 2. We deploy a number of $N = 1000$ users in the community and use a reputation set $\Theta = \{0, 1\}$.

3.6.1 Performance Comparison

In the first experiment, we compare the performance of the reputation protocol designed in Theorem 2 with those of other incentive protocols from the existing literature.

Particularly, three protocols are considered. Protocol 1 is the reputation protocol from the optimal design in Theorem 2. Protocol 2 is the reputation protocol designed in our prior work [45]. Protocol 3 is the BitTorrent protocol which implements the tit-for-tat direct reciprocation mechanism [41]. In tit-for-tat, each user records his personal interactions with other users, and determines his reciprocation behavior with others based on this personal record. Particularly, each user i supplies services to another user j if and only if user j has provided services to him in their last stage game. We run an experiment in which the users in the community interact for 10^8 periods under each protocol. In the experiment, we assume that each user has a discount factor $\delta = 0.8$. Table 3.2 presents the average social welfare received in these 10^8 periods which is normalized by the social optimum $b - c$, i.e. $\frac{\sum_{t=1}^{T}\sum_{i=1}^{N} u_i^{(t)}}{TN(b-c)}$ with $T = 10^8$ and $N = 1000$, where $u_i^{(t)}$ is defined in (1). The result shows that the protocol designed in this chapter (Protocol 1) delivers significantly better performance than other protocols. The main reason for these gains is that Protocols 2 and 3 do not explicitly consider the long-run evolution of the community in their designs. For example, Protocol 2 optimizes the social welfare when the community is sustained at the equilibrium under which all users comply with the social rule and are cooperative with each other. Nevertheless, when the community deviates from this equilibrium (i.e. some users no longer comply with the social rule) due to stochastic perturbations, Protocol 2 cannot ensure that the

Table 3.2: Average social welfare under different protocols

	$c = 1, b = 3,$ $\varepsilon = 0.01$	$c = 1, b = 5,$ $\varepsilon = 0.01$	$c = 1, b = 3,$ $\varepsilon = 0.05$	$c = 1, b = 5,$ $\varepsilon = 0.05$
Protocol 1	0.9328	0.9428	0.8536	0.8810
Protocol 2	0.6628	0.8234	0.2426	0.3072
Protocol 3	0.4372	0.6782	0.1728	0.2876

community converges back to this particular equilibrium and hence, the community stays at non-equilibrium states in the long run which results to a low average social welfare over time.

3.6.2 Convergence Speed

Theorem 2 prescribes "how to design the reputation protocol such that the community converges to a fully cooperative equilibrium in the long run". In this section, we analyze how fast the community will converge to this equilibrium, i.e. starting from an arbitrary community state, what is the expected waiting time for the community to enter a stochastically stable equilibrium. In the experiment, we set $\delta = 0.8$, $b = 5$ and $c = 1$. According to Theorem 2, it can be calculated that $\boldsymbol{\mu}_B^N$ is the unique stochastically stable state when $\beta \geqslant 0.25$. We run the experiment with different selections of (ε, β) and different initial community states that are randomly sampled from the set $\{\boldsymbol{\mu}_B^0, \ldots, \boldsymbol{\mu}_B^{N-1}\}$. Table 3.3 presents the number of periods (i.e. the waiting time) for the community to converge to the state $\boldsymbol{\mu}_B^N$. A first observation is that the waiting time is smaller with larger ε, which is intuitive since it is easier for the community to leave a state that is invariant but not stochastically stable when more stochastic perturbations take place in each period. Another interesting observation from Table 3.3 is that the waiting time monotonically increases with the punishment strength β. This is due to the fact that when β decreases, it is more likely for a user not to be punished upon his deviation. Hence, it becomes more difficult for the community to enter invariant states that contain a large portion of bad users, e.g. $\boldsymbol{\mu}_B^0$, and easier to enter invariant states

Table 3.3: The convergence time to the stochastically stable equilibrium

ε	β	t (10^3 periods)
0.05	0.4	36
0.05	0.6	105
0.05	0.8	890
0.05	1.0	1049
0.1	0.4	23
0.1	0.6	98
0.1	0.8	650
0.1	1.0	820

that contain a large portion of good users, e.g. $\boldsymbol{\mu}_B^N$.

3.7 Conclusions and Future Research

We have studied the problem of designing reputation protocols for online communities in which users adapt their behaviors and learn over time. In such communities, protocol designers can determine the long-run evolution of the community of users and use this knowledge to induce cooperation among users in the long run. Our framework can be extended in several directions, among which we mention three. First, users in the community do not necessarily need to be homogeneous with respect to their stage-game benefits and costs. Different users can have different benefits and costs for the service received/provided. Also, they can choose different discount factors δ when evaluating the long-term utility. Second, instead of always reporting truthfully, clients can use more complicated decision rules while reporting the servers' actions to the community operator in order to maximize their own long-term utility. Third, users adopt a simple belief model in this chapter. However, more sophisticated belief models can be introduced into our framework. For example, the formation of user beliefs and opinions in social networks are extensively studied in [52] and [53]. Understanding how the evolutions of users' beliefs and users' strategies will impact each other will also form an important future research direction.

3.8 Appendices

3.8.1 Appendix A: Proof of Proposition 1

To prove this proposition, it is equivalent to showing the correctness of the following statement for any $\tilde{\theta} > 0$: if $\sigma_i^{\kappa,\boldsymbol{\eta}_i}(\theta,\tilde{\theta}) = 0$, then $\sigma_i^{\kappa,\boldsymbol{\eta}_i}(\theta,\tilde{\theta}-1) = 0$ always holds.

We use contradictions to prove the above statement. Suppose that there are $\theta \in \Theta$ and $\tilde{\theta} > 0$ such that $\sigma_i^{\kappa,\boldsymbol{\eta}_i}(\theta,\tilde{\theta}) = 0$ and $\sigma_i^{\kappa,\boldsymbol{\eta}_i}(\theta,\tilde{\theta}-1) = 1$. According to Lemma 3, we have $\phi(\theta,\tilde{\theta}-1) = 1$ and $\phi(\theta,\tilde{\theta}) = 1$. Now consider an alternative strategy σ' such that $\sigma_i^{\kappa,\boldsymbol{\eta}_i}(\theta',\theta'') = \sigma'(\theta',\theta''), \forall(\theta',\theta'') \neq (\theta,\tilde{\theta})$ and $\sigma'(\theta,\tilde{\theta}) = 1$. That is, σ' and $\sigma_i^{\kappa,\boldsymbol{\eta}_i}$ differ only at the reputation pair $(\theta,\tilde{\theta})$. Then according to the one-shot deviation principle, the expected long-term utility of user i by adopting $\sigma_i^{\kappa,\boldsymbol{\eta}_i}$ at reputation θ, i.e. $V_i^{\kappa,\boldsymbol{\eta}_i}(\theta)$, should be larger than the expected long-term utility when he adopts σ' in the current period and continues to comply with $\sigma_i^{\kappa,\boldsymbol{\eta}_i}$ in the future, which gives the following inequality:

$$\begin{array}{rl} & v_i(\theta|\sigma_i^{\kappa,\boldsymbol{\eta}_i},\boldsymbol{\eta}_i) + \delta\sum\limits_{\theta'} p_i(\theta'|\theta,\sigma_i^{\kappa,\boldsymbol{\eta}_i},\boldsymbol{\eta}_i,\kappa)V_i(\theta'|\sigma_i^{\kappa,\boldsymbol{\eta}_i},\boldsymbol{\eta}_i,\kappa) \\ > & v_i(\theta|\sigma_i^{\kappa,\boldsymbol{\eta}_i},\boldsymbol{\eta}_i) - \frac{\eta_i(\tilde{\theta})}{N-1}c + \delta\sum\limits_{\theta'} p_i(\theta'|\theta,\sigma',\boldsymbol{\eta}_i,\kappa)V_i(\theta'|\sigma_i^{\kappa,\boldsymbol{\eta}_i},\boldsymbol{\eta}_i,\kappa). \end{array} \tag{3.22}$$

By manipulating (3.22) , we have

$$\frac{\eta_i(\tilde{\theta})}{N-1}c > \delta\sum_{\theta'}\begin{pmatrix} p_i(\theta'|\theta,\sigma',\boldsymbol{\eta}_i,\kappa) \\ -p_i(\theta'|\theta,\sigma_i^{\kappa,\boldsymbol{\eta}_i},\boldsymbol{\eta}_i,\kappa)\end{pmatrix} V_i(\theta'|\sigma_i^{\kappa,\boldsymbol{\eta}_i},\boldsymbol{\eta}_i,\kappa). \tag{3.23}$$

By solving its RHS, (3.23) can be transformed as follows

$$\frac{\eta_i(\tilde{\theta})}{N-1}c > \frac{\eta_i(\tilde{\theta})}{N-1}\delta\begin{pmatrix} V(\theta+1|\sigma_i^{\kappa,\boldsymbol{\eta}_i},\boldsymbol{\eta}_i,\kappa) - \beta V(0|\sigma_i^{\kappa,\boldsymbol{\eta}_i},\boldsymbol{\eta}_i,\kappa) \\ -(1-\beta)V(\theta|\sigma_i^{\kappa,\boldsymbol{\eta}_i},\boldsymbol{\eta}_i,\kappa)\end{pmatrix}. \tag{3.24}$$

The LHS of (3.24) is the extra immediate cost that is incurred to user i when it is at reputation θ and provides services to a client of reputation $\tilde{\theta}$, and the RHS of (3.24) is the extra future utility that is brought by doing so. Since the immediate cost is higher than the extra future utility, user i chooses $\sigma_i^{\kappa,\boldsymbol{\eta}_i}(\theta,\tilde{\theta})=0$ in his optimal strategy. Using the same idea, the following condition should be met when $\sigma_i^{\kappa,\boldsymbol{\eta}_i}(\theta,\tilde{\theta}-1)=1$:

$$\frac{\eta_i(\tilde{\theta}-1)}{N-1}c \leqslant \delta\frac{\eta_i(\tilde{\theta}-1)}{N-1}\begin{pmatrix} V(\theta+1|\sigma_i^{\kappa,\eta_i},\kappa,\eta_i)-\beta V(0|\sigma_i^{\kappa,\boldsymbol{\eta}_i},\kappa,\boldsymbol{\eta}_i) \\ -(1-\beta)V(\theta|\sigma_i^{\kappa,\boldsymbol{\eta}_i},\kappa,\boldsymbol{\eta}_i) \end{pmatrix}. \tag{3.25}$$

It is easy to observe that (3.24) and (3.25) cannot hold at the same time and hence, we have a contradiction. Therefore, the statement at the beginning of this proposition always holds for $\sigma_i^{\kappa,\boldsymbol{\eta}_i}$ and Proposition 1 follows. ■

3.8.2 Appendix B: Proof of Proposition 2

According to Lemma 3, $h_{\kappa,\boldsymbol{\eta}_i}(\theta)\in[H,L+1]$ if $\theta\geqslant H$. Similarly, we also have $h_{\kappa,\boldsymbol{\eta}_i}(\theta)\in[0,L+1]$ if $0\leqslant\theta<H$.

The rest of this proof follows the same argument as in Proposition 1. Suppose $h_{\kappa,\boldsymbol{\eta}_i}(\theta)\in(0,L+1)$ for some $\theta<H$, i.e. $\sigma_i^{\kappa,\boldsymbol{\eta}_i}(\theta,h_{\kappa,\boldsymbol{\eta}_i}(\theta))=1$ and $\sigma_i^{\kappa,\boldsymbol{\eta}_i}(\theta,h_{\kappa,\boldsymbol{\eta}_i}(\theta)-1)=0$, then according to the one-shot deviation principle, we have that

$$\begin{array}{rcl} \frac{\eta_i(h_{\kappa,\boldsymbol{\eta}_i}(\theta))}{N-1}c & > & \delta\frac{\eta_i(h_{\kappa,\boldsymbol{\eta}_i}(\theta))}{N-1}\begin{pmatrix} V_i^{\kappa,\boldsymbol{\eta}_i}(\theta+1)-\beta V_i^{\kappa,\boldsymbol{\eta}_i}(0) \\ -(1-\beta)V_i^{\kappa,\boldsymbol{\eta}_i}(\theta) \end{pmatrix} \\ \frac{\eta_i(h_{\kappa,\boldsymbol{\eta}_i}(\theta)-1)}{N-1}c & \leqslant & \delta\frac{\eta_i(h_{\kappa,\boldsymbol{\eta}_i}(\theta)-1)}{N-1}\begin{pmatrix} V_i^{\kappa,\boldsymbol{\eta}_i}(\theta+1)-\beta V_i^{\kappa,\boldsymbol{\eta}_i}(0) \\ -(1-\beta)V_i^{\kappa,\boldsymbol{\eta}_i}(\theta) \end{pmatrix} \end{array}, \tag{3.26}$$

which introduces a contradiction. Hence, $h_{\kappa,\boldsymbol{\eta}_i}(\theta)\notin(0,L+1)$ and we have $h_{\kappa,\boldsymbol{\eta}_i}(\theta)\in\{0,L+1\}$ for any $\theta<H$. Similar arguments apply to good users and thus $h_{\kappa,\boldsymbol{\eta}_i}(\theta)\in\{H,L+1\}$ for any $\theta\geqslant H$. ■

CHAPTER 4

Information Production and Link Formation in Social Computing Systems

4.1 Introduction

Social computing is booming nowadays [5][42][57] and has recently led to a new trend: the integration of social computing techniques with communication networks to enable individuals with similar interests to connect to each other over a mobile or low-cost network infrastructure to share and disseminate user-generated resources, such as multi-modal content, geographical information, event-related information, personal sensory information, etc. For instance, [58] and [59] propose mobile content distribution architectures relying on social relationship and providing economic incentives to offload the bandwidth-intensive task of content dissemination from service providers to the mobile subscribers in mobile and wireless networks. Distributed spectrum sensing in peer-to-peer cognitive networks is proposed in [60] to efficiently map the spectrum opportunities based on contributions from individual cognitive terminals. The work of [61] formulates the location-aware information sharing and link formation in vehicular networks as a coalition game to improve the efficiency of road information dissemination among vehicles.

In all the aforementioned social computing applications, the decentralized strategic interactions of agents have decisive impact on the emerging network topologies as well as the stability and the efficiency of emerging networks and distributed systems. In this chapter, we are focusing on two key features that characterize the strategic behavior of agents: resource production and link formation. Resource production refers to the agents' decisions to personally generating resources. Here we use the abstract concept of "resources" to represent any knowledge, multi-modal content, and other kinds of data and service that can be exchanged and shared in social computing systems. Several examples on resources production would be the upload and creation of blogs, videos and photos on online platforms with user-generated content [5], the download and purchase of content from service provider in peer-to-peer networks [57] and content distribution networks [58], and the creation of local spectrum information sensed by individual cognitive terminals in peer-to-peer cognitive networks [60]. Link formation refers to all kinds of agents' decisions regarding exchanging the produced resources with other agents, i.e. how agents create and dissolve links with each other to acquire resources. For example, a peer node in P2P networks makes active decisions on which peers it should connect to and download content from; in peer-to-peer spectrum sensing, a cognitive terminal can request and exchange information about spectrum opportunities which have already been discovered by other terminals [60][62]. Most of the existing works assume that agents decisions on these actions are exogenously determined by a system designer aiming to maximize the overall network utility. Hence, agents do not proactively make decisions but rather obediently follow the actions prescribed by the designer. While this assumption is realistic for traditional communication networks composed of obedient agents (e.g. sensor nodes), it does no longer hold in social computing systems, where networks are usually formed in an ad-hoc fashion by self-interested agents determining their own strategies in order to maximize their own utilities. Hence, a novel framework on studying the incentive of agents in the joint decision on resources production and link formation according to their own self-interest is necessary. Such a framework does not only help to understand the agents' behavior in social computing systems but also facilitates the design of incentive protocols which encourage the agents to cooperate with each other in a way to optimize the system efficiency. In the rest of this chapter, we use "social computing system" and "network" interchangeably.

Self-interested link formation in social computing systems has been studied by microeconomics and network science researchers (see e.g. [63]), who analyze how the agents' self-interests lead to strategic link creation in a network and determine what network topologies arise. A simple model is proposed in [64], where the problem of link formation is formulated as a non-cooperative game among strategic agents. The agents can select which links to create with other agents in order to individually maximize their own utilities by trading off the potential rewards obtained from forming a link (e.g. resources acquisition) against the incurred link creation cost (e.g. payment and maintenance costs for links). Analysis of this model can lead to useful predictions of what "equilibrium" topologies emerge. This basic model has subsequently been extended in various directions. For example, [65] studies the bilateral network formation in which the link creation requires mutual consent and the cost is two-sided. The network formation problem for the scenario where agents are heterogeneous is analyzed in [66], which shows that a strict equilibrium network is minimal and, conversely, every minimal network is a strict equilibrium for suitable costs and benefits [1][2].

Importantly, there are two key drawbacks that prevent these existing models from being successfully applied to social computing systems. First, most of these works assume that agents are endowed with exogenous amounts of resources and focus solely on the strategic aspect of link formation. The fact that agents usually have the capability to self-produce resources is neglected. Second, these works treat resources collected from different agents as homogeneous, i.e. being equally valued and perfectly substitutable, while in social computing systems, resources from different agents is heterogeneous with respect to location, devices used for gathering this resources, content contained, etc., and thus it cannot be perfectly substitutable with each other. Meanwhile, given the information heterogeneity, an agent's benefit from information consumption no longer

[1]A strict equilibrium is a strategy profile in which each agent's utility is strictly larger than the resulting utility by choosing any other strategy.

[2]A minimal network is a topology where there is a unique path between any two agents in the network.

only depends on the total amount of information it consumes, but also on how many types of information and what amount of each type it acquires. Therefore, an agent normally has certain appreciation for information variety by valuing information produced by numerous sources higher than information produced by a single source [67].

Based on these characteristics of the agents, our work proposes a novel Information Production and Link Formation game (IPLF game) for social computing systems to jointly study agents' strategic behavior on information production and link formation. To deal with the agents' appreciation for information variety, we deploy the well-known Dixit-Stiglitz utility function [70] to capture the impact of information heterogeneity on the agents utilities. Using this formalism, we study what asymptotic equilibria emerge in networks of large sizes. Different from the idealized works in economics, e.g. [69], which assume perfectly substitutable information (i.e. agents have no appreciation for variety) and predict the occurrence of "the law of the few" [69] [3], our analysis shows that when the size (population) of a network is sufficiently large, every (strict) non-cooperative equilibrium of the IPLF game consists of a hierarchical core-periphery structure with all agents belonging to one of the two types: hub agents producing large amounts of information have a large number of connections and serve as the major source of information sharing; and spoke agents produce and share limited amounts of information and mainly consume information acquired from hub agents. Importantly, as the network size grows, the population of hub agents as well as the total amount of information produced in the network grows proportionally to the network size. This shows that in social computing systems where agents appreciate information variety, the "law of the few" does not exist with the production and sharing no longer being dominated by a small group of powerful agents but rather being more distributed, and the network structure remaining flat but not scale-free [4].

[3] "The law of the few" [69] states that under perfect information substitutability, only a small amount of agents in a network will produce information with the fraction of information producers in the total population goes to 0. Meanwhile, the total amount of information produced in the network is upper-bounded.

[4] A scale-free network is a network whose degree distribution follows a power law asymptotically. If

Following the characterization on the equilibria in the IPLF game, we then analyze the efficiency of such equilibrium topologies and compare it with the social optimum by providing bounds on the Price of Stability and the Price of Anarchy. Stemming from this basic model, we further propose and investigate the IPLF game which incorporates indirect information sharing, which is a common feature in social computing systems that an agent cannot only consume and access information produced by its neighbors but also that produced by the neighbors of its neighbors.

The remainder of this chapter is organized as follows. In Section 4.2, we describe our basic model of the IPLF game. In Section 4.3, we characterize the emerging non-cooperative equilibria. In Section 4.4, the model of the IPLF game with indirect information sharing is presented and discussed. We conclude in Section 4.5 and also outline future research topics.

4.2 Basic Model

We consider information sharing in a social computing system where agents have the capability to produce information personally as well as form links with others in order to acquire information that they produce. The definition of "agent" in this chapter is very general. It can represent individual nodes, devices, peers, and also autonomous sub-networks or groups of nodes cooperating with each other to form an integrated entity participating in the IPLF game. Examples of information being shared are multi-modal content, environmental sensory data in sensor networks, road information in vehicular networks, geographical information in location-based services, etc.

In all these different applications, the precise formulation of actions and utilities of the participating agents will depend on details of the application. For example, precisely how agents can form links with others, whether there is any budget or practical

a network is scale-free, the fraction of agents with large numbers of connections in the entire population decays fast to 0 as the network size grows [71].

constraints on connections such as physical locations and bandwidth limits, how does an agent benefit/cost from the presence of a link, etc. In this work, we use a stylized model to formulate the IPLF game without delving into the idiosyncrasies of any particular application, in order to capture the basic trade-offs and draw qualitative insights about the effects of self-interested behavior on the network structure of social computing systems. However, it should be noted that various alternative models are ready to be extended from this basic model for the analyses of particular applications, as discussed in Section 4.5.

Let $N = \{1, 2, \ldots, n\}$ denote the set of agents in the system, where i and j represent typical members. Each agent i determines the amount of information it produces, which is called as its production level and is denoted by $x_i \in \mathbb{R}^+$. We assume a linear cost on information production as in [66] and hence, a cost of $c \in \mathbb{R}^+$ incurs to an agent for each unit of information it produces. Examples of the cost could be the energy consumption on spectrum sensing in P2P cognitive networks, subscription fees paid to service providers for content download in content distribution networks, etc. Besides the self-production, each agent also determines whether to create links with other peer agents in order to acquire information from them. Here we consider unilateral link formation as in [73], where links are created by the unilateral actions of agents and link costs are one-sided. An agent i's link formation action to another agent j is represented by a binary variable $g_{ij} = \{0, 1\}$. We set $g_{ij} = 1$ if agent i forms a link to agent $j \neq i$ and $g_{ij} = 0$ otherwise. For convenience, we set $g_{ii} = 0$ for all $i \in N$. For illustration purposes, we refer to the agent who forms a link and the agent with whom the link is formed as the creator and recipient of the link, respectively. Forming a link incurs a cost k, which abstracts all the costs and payments for creating and maintaining the link, as well as costs incurred in information exchange and transmission (e.g. uploading and downloading cost, energy consumption in transmission, etc.). Here we assume the value of k to be constant for each link. Meanwhile, since a link is formed unilaterally without the consent of the recipient, we assume that the creator undertakes the entire cost k (i.e. if there is some cost incurred for the recipient in receiving a link, the creator will also compensate the receiver by paying for it). The link formation decision by agent i is then represented by a row vector $\mathbf{g}_i = (g_{ij})_{j \in N, j \neq i} \in \{0,1\}^{n-1}$. We assume that the information exchange across a link is undirected and thus, the connection status between agents i and j can be represented by a variable $\bar{g}_{ij} = max\{g_{ij}, g_{ji}\}$ which takes the value of 1 when they are connected and called as "neighbors" of each other. We also assume that $\bar{g}_{ii} = 1$. Since the link is undirected and we assume that each link has sufficient capacity to support the information transmission, neighboring agents can mutually access all information produced by each other.

We consider a simultaneous move game where agents make decisions on information production and link formation at the same time. The set of strategies of agent i is thus denoted by $S_i = \mathbb{R}^+ \times \{0,1\}^{n-1}$. A strategy profile in the IPLF game is written as $\mathbf{s} = (\mathbf{x}, \mathbf{g})$, where $\mathbf{x} = (x_i)_{i=1}^n$ and $\mathbf{g} = (\mathbf{g}_i)_{i=1}^n$ denote the information production and link formation decisions of all agents, respectively. Let $\bar{\mathbf{g}} \triangleq [\bar{g}_{ij}]_{i,j \in N}$ denote the connectivity graph of the network, we also define $N_i(\mathbf{g}) \triangleq \{j | g_{ij} = 1\}$ as the set of agents to whom agent i forms a link and $N_i(\bar{\mathbf{g}}) \triangleq \{j | \bar{g}_{ij} = 1\}$ as the set of agents to whom agent i is connected, i.e. the neighbors of agent i.

Agents in the IPLF game benefit from consuming information available to them. In the basic model, we assume pure local externalities [69]: each agent only consumes information personally produced by itself and its neighbors [5], and cannot consume information produced by another agent who is more than one hop away from it on the connectivity graph $\bar{\mathbf{g}}$. As mention previously, existing works on strategic link formation assume that information produced by different agents is perfectly substitutable [64][69], i.e. the total amount of information that an agent consumes fully determines its benefit. Therefore, an agent's benefit from information consumption is not affected by where the information comes from as long as the total amount of information consumed remains constant. This assumption, however, fails to capture agents' interests

[5] An analysis about indirect information sharing can be found in Section 4.4

and benefit from consuming diverse information in social computing systems. For example, the spectrum information collected by different cognitive terminals at different locations should have different values in consumption and cannot be fully replaced. Hence, instead of assuming the perfect substitutability, we use the Dixit-Stiglitz preference model from [70] to capture agents' appreciation for information variety, under which an agent i's benefit from information consumption is given by

$$f_i(\mathbf{x}, \mathbf{g}) = v((x_i^\rho + \sum_{j \in N_i(\bar{\mathbf{g}})} x_j^\rho)^{1/\rho}) \tag{4.1}$$

for some function $v : \mathbb{R}^+ \to \mathbb{R}^+$ and some $\rho \in (0, 1)$. ρ measures an agent's appreciation for information variety. When $\rho < 1$, an agent obtains a higher benefit when it consumes information from a more diverse bundle of agents. It should be noted that when $\rho \to 1$, the agent's appreciation for information variety disappears and hence, the function (4.1) could also be used to model an agent's benefit when information is perfectly substitutable as in [64] and [69] by setting $\rho = 1$.

For the illustration purpose, we define $X_i \triangleq (x_i^\rho + \sum_{j \in N_i(\bar{\mathbf{g}})} x_j^\rho)^{1/\rho}$ as the amount of agent i's effective information. This could be interpreted as the amount of information that agent i really consumes after aggregating and processing all the information it acquires. The following two assumptions are imposed on the benefit function v:

Assumption 1: $v(\cdot)$ is a twice continuously differentiable, increasing, and strictly concave function.

Assumption 2: $v(\cdot)$ satisfies $v(0) = 0$, $v'(0) \triangleq \lim_{x \to 0^+} v(x) < \infty$, $v'(0) > \alpha$ where $\alpha > 0$ is a constant, and $lim_{x \to \infty} v'(x) = 0$.

These two assumptions capture the diminishing marginal benefit in an agent's information consumption with respect to the amount of its effective information. That is, an agent's benefit increases with the amount of its effective information, while the rate of this increase decreases and approaches to 0. Hence, there always exists an upper bound on the amount of information that an individual agent is willing to acquire given the cost of information production and link formation.

Next, we analyze the effect of ρ on an agent's information consumption.

Lemma 1. Given a strategy profile $\mathbf{s} = (\mathbf{x}, \mathbf{g})$, $(x_i^\rho + \sum_{j \in N_i(\bar{\mathbf{g}})} x_j^\rho)^{1/\rho} > (x_i^{\rho'} + \sum_{j \in N_i(\bar{\mathbf{g}})} x_j^{\rho'})^{1/\rho'}$ always holds for any $i \in N$ and any $0 < \rho < \rho' \leq 1$. ■

Remark: For the proofs that are omitted in this chapter, the complete proofs and computations can always be found in the online appendix [75].

This lemma can be proven by first showing that $(y^\rho + z^\rho)^{1/\rho} > (y^{\rho'} + z^{\rho'})^{1/\rho'}$ always holds for any $y, z > 0$ and $0 < \rho < \rho' \leq 1$ through the first-order partial derivatives of $(y^\rho + z^\rho)^{1/\rho}$ over ρ, which is always negative when $0 < \rho \leq 1$. Then this lemma follows as a straightforward extension.

Lemma 1 shows that given a strategy profile, an agent's level of appreciation for information variety increases as ρ becomes smaller, which is reflected in the increase on the amount of its effective information and the resulting benefit from information consumption.

With the above two assumptions, $f_i(\mathbf{x}, \mathbf{g})$ is also provably twice continuously differentiable, increasing and strictly concave in x_i, as shown in the following lemma.

Lemma 2. $f_i(\mathbf{x}, \mathbf{g}) = v((x_i^\rho + \sum_{j \in N_i(\bar{\mathbf{g}})} x_j^\rho)^{1/\rho})$ is twice continuously differentiable, increasing, and strictly concave in x_i on $\mathbb{R}^+$. ■

Since $f_i(\mathbf{x}, \mathbf{g})$ is twice continuously differentiable, we further assume it to be submodular with the following assumption.

Assumption 3: $\frac{\partial^2 f_i}{\partial x_i \partial x_j}|_{x_i, x_j \in \mathbb{R}^+} < 0$, $\forall i,\ j \in N_i(\bar{\mathbf{g}})$.

Assumption 3 formalizes the substitutability among agents' information. Agent i's marginal benefit of production decreases against the amount of information produced by its neighbors. That is, the more information i acquires, the less incentive i has to produce information by itself.

To sum up, the utility of agent i is given by its benefit minus all incurred costs:

$$u_i(\mathbf{x}, \mathbf{g}) = v((x_i^\rho + \sum_{j \in N_i(\bar{\mathbf{g}})} x_j^\rho)^{1/\rho}) - cx_i - k|N_i(\mathbf{g})|. \tag{4.2}$$

We analyze the case of homogeneous agents in that v, c, k, and ρ are the same for all agents. We assume that $\alpha > c$ to ensure the network is socially valuable.

4.3 Equilibrium in the IPLF Game

4.3.1 Equilibrium analysis for the basic model

We analyze the IPLF game as a non-cooperative one-shot game and consider pure s-trategies. Each agent maximizes its own utility given the strategies of others. A Nash equilibrium of the IPLF game is a strategy profile $\mathbf{s}^* = (\mathbf{x}^*, \mathbf{g}^*)$ such that

$$u_i(\mathbf{s}_i^*, \mathbf{s}_{-i}^*) \geq u_i(\mathbf{s}_i, \mathbf{s}_{-i}^*), \forall \mathbf{s}_i \in \{0,1\}^{n-1}, \forall i \in N. \tag{4.3}$$

Here, we use the convention that $\mathbf{s}_i$ and $\mathbf{s}_{-i}$ represents the strategies of agent i and all agents other than i, respectively. This section analyzes the equilibrium production and link formation behavior of agents. Given the definition of equilibrium (4.3), an equilibrium production profile $\mathbf{x}^*$ satisfies the following equality for each agent:

$$v'(X_i^*)(X_i^*/x_i^*)^{1-\rho} = c, \ \forall i, \tag{4.4}$$

where $X_i^* = ((x_i^*)^\rho + \sum_{j \in N_i(\bar{\mathbf{g}})} (x_j^*)^\rho)^{1/\rho}$. That is, the marginal benefit of production should equal to the marginal cost and thus, an agent i has no incentive to produce more than x_i^* when the amount of its effective information is X_i^*.

We first derive the basic properties of agents' equilibrium behavior on information production and link formation. This is summarized in the following lemma. Although simple, these properties are important in characterizing the emerging equilibrium later.

Lemma 3. In any equilibrium $\mathbf{s}^* = (\mathbf{x}^*, \mathbf{g}^*)$ of the IPLF game, (i) $g_{ij}^* g_{ji}^* = 0$ for all $i, j \in N$; (ii) $x_i^* > 0$ for all $i \in N$; (iii) $x_i^* \leq \bar{x}$ for all $i \in N$ where $\bar{x}$ is the unique solution of the equation $v'(\bar{x}) = c$. ■

We briefly explain Lemma 3.

(i) Since the information exchange across a link is undirected, if an agent i forms a link to another agent j it has already connected with, it incurs an additional cost of link formation but does not receive any extra information from j and hence, its utility strictly decreases by doing so. Hence in an equilibrium, there should be no redundant investment on link formation.

(ii) If an agent has no neighbors (i.e. no connections), it has a positive production level as $\alpha > c$. On the contrary, if an agent has neighbors, it also has a positive production level according to (4.4). Hence, an agent always has a positive production level and the information it acquires from neighbors can never fully replace its self-produced information.

(iii) According to Assumption 3, an agent has the largest marginal utility from production when it has no neighbors (i.e. it acquires no information from others). Hence, the highest production level could also be achieved in this case and is upper bounded due to the concavity of the benefit function (4.1).

The dependence of the value of $\bar{x}$ on c and v is neglected in the notation when no confusion is brought in. In the rest of this chapter, $\bar{x}$ is referred as the maximum equilibrium production level.

The maximum equilibrium production level is critical in characterizing the relationship between the link formation cost and emerging equilibria of the IPLF game, as shown in the following theorem on the existence of the Nash equilibrium.

Theorem 1 (The Existence of Equilibrium). Pure strategy Nash equilibria always exist in the IPLF game and each equilibrium belongs to one of the following two types:

(i) each agent personally produces an amount $\bar{x}$ of information and no one forms any link;

(ii) each agent personally produces an amount strictly smaller than $\bar{x}$ and is connected with at least one other agent.

Proof Sketch: We first prove the existence of Nash equilibrium. In general, it is difficult to show the existence of pure Nash equilibrium in network formation games directly. Hence, in this proof, we first consider the IPLF game where agents play mixed strategy, which is called as IPLFM. Particularly, in IPLFM, the link formation choice between two agents is not binary, but continuous. That is, the link formation strategy of an agent i now becomes a vector $\mathbf{p}_i = \{p_{i1}, \ldots, p_{in}\}$, where $p_{ij} \in [0, 1]$ and $p_{ii} = 0$. We define the strength of a link to be $\bar{g}_{ij} = \bar{g}_{ji} = max\{p_{ij}, p_{ji}\}$. Since each agent plays a mixed strategy on both information production and link formation, it is always true that the IPLFM game has at least one equilibrium. We then show that each equilibrium of the IPLFM game has $p_{ij} \in \{0, 1\}$ for any $i, j \in N$, and thus is also an equilibrium of the IPLF game where the link formation choice is binary.

To prove the second part of the theorem, we classify all strategy profiles into two classes. The first class S^A contains strategy profiles with which there is no link in the network. The second class $S^B = S/S^A$ contains all other strategy profiles. Here S is the set of all strategy profiles. It can be shown that an equilibrium in S^A belongs to type (i) and an equilibrium in S^B should contain no isolated agent and belongs to type (ii). Otherwise, if a strategy profile in S^B contains isolated agents, then due to the fact that an isolated agent always has the largest production level, non-isolated agents will be attracted to connect with it, and which contradicts the definition of an equilibrium. ■

Theorem 1 shows that in an equilibrium of the IPLF game, the network either is empty with all agents being isolated, or has no agent being isolated. In the rest of this section, we analyze the non-trivial case of strict equilibria with which the inequality (4.3) is strict, and we use equilibrium to mean strict pure strategy equilibrium when no confusion occurs. Without loss of generality, agents are ordered by their production levels in a strategy profile $\mathbf{s} = (\mathbf{x}, \mathbf{g})$, i.e. $x_1 \geq x_2 \geq \cdots \geq x_n$. Given such ordering, there is always a positive integer $n_h(\mathbf{s}) \leq n$, such that $x_i = x_{n_h(\mathbf{s})}$ for all $i \leq n_h(\mathbf{s})$ and $x_j < x_{n_h(\mathbf{s})}$ for all $j > n_h(\mathbf{s})$. We call an agent $i \leq n_h(\mathbf{s})$ to be a high producer and an agent $j > n_h(\mathbf{s})$ to be a low producer. According to Lemma 3(iii), the production level of each high producer will never exceed $\bar{x}$ and is denoted as $\tilde{x}(\mathbf{x})$. To avoid the trivial case that the set of low producers is empty and all agents have the same production level, we assume that $n_h(\mathbf{s}) < n$ in this chapter. A detailed analysis for strategy profiles where all agents produce the same amount of information can be found in the technical report [76].

By separating agents into two types, we are able to characterize in the following lemma the interactions between high and low producers at equilibrium.

Lemma 4. In an equilibrium $\mathbf{s}^* = (\mathbf{x}^*, \mathbf{g}^*)$, the following properties hold for $\mathbf{g}^*$: (i) For each $i \leq n_h(\mathbf{s}^*)$, $g^*_{ij} = 0,\ \forall j > n_h(\mathbf{s}^*)$; (ii) For each $i \leq n_h(\mathbf{s}^*)$, $\bar{g}^*_{i,i'} = 0$ for some $i' \leq n_h(\mathbf{s}^*)$ and $i \neq i'$; (iii) For each $j > n_h(\mathbf{s}^*)$, $g^*_{ji} = 1$, for some $i \leq n_h(\mathbf{s}^*)$.

Proof: The proof in Lemma 4 heavily relies on the use of contradictions. Here we only provide the proof of statement (i) due to the space limitation. The proofs for the other two statements follow similar ideas and can be found in the online appendix [75].

Suppose that there is an agent $j > n_h(\mathbf{s}^*)$ such that $g^*_{ij} = 1$. It is always true that i is connected with all other high producers, i.e. $\bar{g}_{ii'} = 1,\ \forall i' \leq n_h(\mathbf{s}^*)$ and $i' \neq i$. Otherwise, i could always strictly increase its payoff by switching its link with j to some other high producer that it does not connect with. If i is also connected with all low producers, then according to Assumption 3, x^*_i is the smallest among all agents in the network, which contradicts the fact that i is a high producer. Hence, there is an agent $j' > n_h(\mathbf{s}^*)$ such that $\bar{g}^*_{ij'}$. Clearly, $g^*_{j'j''} = 0$ for all $j'' > n_h(\mathbf{s}^*)$. Otherwise, j' can strictly increase its utility by switching the link from j'' to i.

Now we prove that each neighbor of agent j' is also a neighbor of agent i. Suppose there is an agent $j'' > n_h(\mathbf{s}^*)$ such that $g^*_{j''j'} = 1$. It can be concluded that $\bar{g}^*_{j''i} = 1$ also holds. Otherwise, j'' can strictly increase its utility by switching the link from j' to i . Regarding the fact that i is connected with all high producers, it implies that every agent who is a neighbor of agent j' is also a neighbor of agent i. Therefore, the amount of information that agent j' receives from its neighbors is no more than what is received by agent i. According to Assumption 3, we have that $x_{j'}^* \geq x_i^*$, which contradicts the fact that $x_{j'}^* < x_i^*$. Statement (i) follows. ■

Lemma 4 provides important insights on how agents will interact with each other at equilibrium. Statement (i) and (ii) jointly characterize the redundancy on produced information at equilibrium. With a non-zero cost of link formation, the total amount of information produced by high producers will always be more than what each agent needs. Hence, each high producer neither has a full connection with all other high producers nor forms links with any low producer. Statement (iii) proves that a low producer always forms links with some high producers.

Given the agents' link formation behavior at equilibrium and the information redundancy, we can further determine, in the next lemma, that if a low producer forms links to other low producers in an equilibrium, it should have formed links to all high producers.

Lemma 5. In an equilibrium $\mathbf{s}^* = (\mathbf{x}^*, \mathbf{g}^*)$, $g_{jj'}^* = 1$, for some $j, j' > n_h(\mathbf{s}^*)$ only if $g_{ji}^* = 1,\ \forall i \leq n_h(\mathbf{s}^*)$. ■

With Lemma 5, low producers can be further classified into two types in order to facilitate our analysis: (i) a low producer forming links to all high producers and to some of the low producers; (ii) a low producer forming links with some (or all) high producers and no low producer.

4.3.2 Asymptotic equilibria in the IPLF game

By analyzing the equilibrium properties of agents, we are then able to characterize the asymptotic equilibria when the network size grows. We prove in the following theorem that when the network size is sufficiently large, low producers of type (i) will disappear in any equilibrium $\mathbf{s}^* = (\mathbf{x}^*, \mathbf{g}^*)$, and there are only high producers and low producers of type (ii) left in the network. Given this, we can further characterize the equilibria topologies. Particularly, we show that each high producer produces an amount $\tilde{x}(\mathbf{x}^*)$

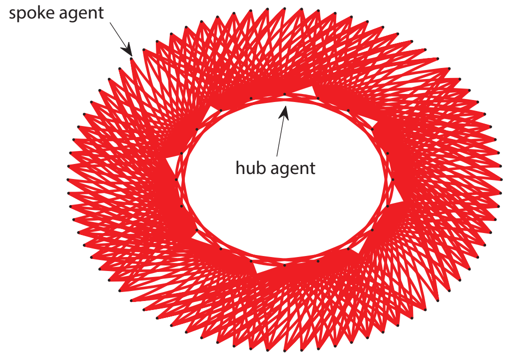
and each low producer (type (ii)) produces a smaller amount of information, denoted as $\underset{\sim}{x}(\mathbf{x}^*)$. Meanwhile, each low producer forms links with $q(\mathbf{g}^*)$ high producers and does not form any link to low producers. Therefore, in a network at equilibrium, all links are formed towards high producers, which play the role of hubs for information sharing in the network. The network then exhibits a "core-periphery" structure with high producers form the core and low producers stay at the periphery.

Theorem 2 (Asymptotic Equilibria). Given c, k, ρ, and v, and when the network size n goes to infinity: (i) only two types of agents exist in any equilibrium $\mathbf{s}^* = (\mathbf{x}^*, \mathbf{g}^*)$: a hub agent is a high producer who produces an amount $\tilde{x}(\mathbf{x}^*)$ and form links only with other hub agents; a spoke agent is a low producer who produces an amount $\underset{\sim}{x}(\mathbf{x}^*)$ and forms links with $q(\mathbf{g}^*)$ hub agents and no link to other spoke agent;

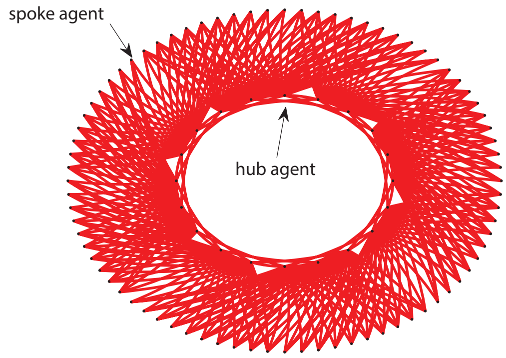
(ii) $\lim_{n\to\infty} \inf_{\mathbf{s}^* \in S_n^*}\{n_h(\mathbf{s}^*)\} = \infty$ and $\lim_{n\to\infty} \inf_{\mathbf{s}^* \in S_n^*}\{n_l(\mathbf{s}^*)\} = \infty$ where $n_l(\mathbf{s}^*) \triangleq n - n_h(\mathbf{s}^*)$ and S_n^* denotes the set of equilibrium strategy profiles when the network size is n. ■

A key observation in proving Theorem 2 is that because of the cost of link formation, an agent has to produce a sufficiently large amount of information, which is lower-bounded away from 0, in order to attract others to form links to it. Meanwhile, an agent has an upper bound on the amount of information that it would like to consume due to the concavity of the benefit function (4.1). As a result, there is an upper bound on the number of links that an agent would like to maintain with others. Therefore, when $n \to \infty$, the number of hub agents (i.e. high producers) goes to infinity as well to provide sufficient information for the consumption of all agents in the network, and low producers will become spoke agents by not mutually connecting with each other but only connecting with hub agents.

Theorem 2 proves that the core-periphery structure is a necessary condition for an equilibrium when n is sufficiently large, which is shown in Figure 4.1 with two examples of equilibrium topologies. Throughout this chapter, we use $v(x) = ln(1 + x^\rho)^{1/\rho}$ as an exemplary utility function in the experiments. Both topologies exhibit the core-



(a) $k = 2$



(b) $k = 5$

Figure 4.1: Exemplary equilibria in a network with $n = 100$ agents ($c = 2,\ \rho = 0.8$)

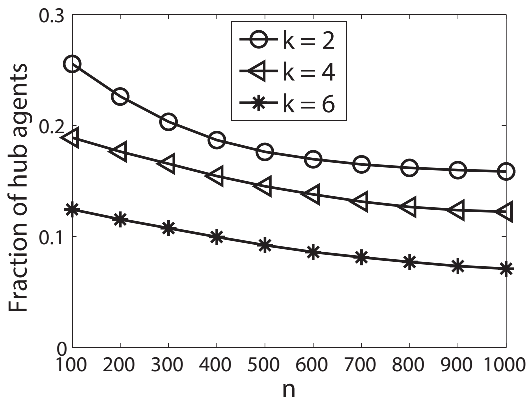
periphery structure composed of two rings. The inner ring (core) represents hub agents and the outer ring (periphery) represents spoke agents. Hence, our results are consistent with the empirical observations that a majority of users in social sharing services get most of their information from a relatively small group of hub agents [72]. By comparing Figure 4.1(a) and Figure 4.1(b), which portray equilibrium topologies with k changing from 2 to 5, we can see that the network becomes sparse as the link formation cost increases, when, instead of forming links with multiple hub agents, spoke agents get information only from hub agents in their vicinity and the network is virtually divided into many small sub-networks where each hub agent takes charge of the information provision to its local agents.

The core-periphery structure exhibits some similarity to the small-world networks [68], which is also predicted in [69] as the "law of the few". Nevertheless, different from the scale-free property where the fraction of hub agents diminishes to 0 as the network size grows to infinity, we show that under agents' appreciation for information variety, the population of hub nodes will grow proportionally to the network size, and their fraction in the total population is lower-bounded away from 0. Similar results have also been illustrated in [72], which reveals, through empirical measurement, that the probability of users who are followed by very large numbers of users on Twitter is above what a scale-free distribution would predict. This indicates that as more agents join the network, information production which is dominated by a small number of powerful producers can no longer satisfy the agents' desire for diverse information. As a result, new hub agents with different varieties of information will emerge to provide the information and stabilize the network.

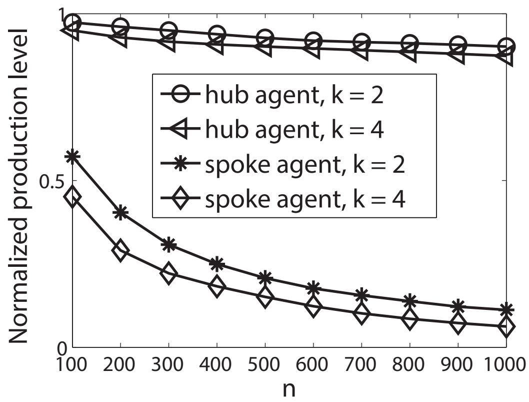
Theorem 3. The number of hub agents at equilibrium grows at the same order as the entire population, i.e. $\inf_{\mathbf{s}^* \in S_n^*}\{n_h(\mathbf{s}^*)\}$ is $\Theta(n)$. More specifically, there exists a constant $\eta > 0$ such that

$$\lim_{n\to\infty} \inf_{\mathbf{s}^* \in S_n^*} \{n_h(\mathbf{s}^*)\}/n > \eta. \tag{4.5}$$

Proof: It has been proved in Theorem 2 that we can always find a sufficiently large value T such that spoke agents will not mutually form links to each other in any equilibrium when the network size $n > T$. It is obvious that we can find a constant value $\mu \in (0, 1)$ such that $\inf_{\mathbf{s}^* \in S_n^*}\{n_h(\mathbf{s}^*)\}/n > \mu,\ \forall n < T$. Now look at the case when $n > T$. Due to the concavity of the benefit function $v(\cdot)$ and the fixed cost of link formation k, a spoke agent cannot connect to more than L_h hub agents, which upper-bounds the amount of information it receives from others at $L_h\bar{x}$. Hence, there exists a constant $\underline{x}$ such that the production level of a low producer $\underset{\sim}{x}(\mathbf{x}^*)$ is no less than $\underline{x}$ in any equilibrium $\mathbf{s}^*$ when $n > T$. The total number of links that each high producer receives from low producers is thus also upper-bounded by a constant, denoted as H_l. Given the population of high producers, i.e. $n_h(\mathbf{s}^*)$, the population of low producers should be no



(a)



(b)

Figure 4.2: (a) The fraction of hub agents changing against n; (b) The normalized production levels of agents changing against n ($c = 2,\ \rho = 0.8$)

more than $H_l n_h(\mathbf{s}^*)$, which further delivers a lower bound on $\inf_{\mathbf{s}^* \in S_n^*}\{n_h(\mathbf{s}^*)\}/n$ as $1/(1+H_l),\ \forall n < T$. We take $\eta = \min(\mu,\ 1/(1+H_l))$ and hence Theorem 3 follows. ■

Figure 4.2 plots how the fraction of high productions as well as $\tilde{x}(\mathbf{x}^*)$ and $\underset{\sim}{x}(\mathbf{x}^*)$ at equilibrium change as n increases [6]. Figure4.2(a) illustrates Theorem 3. The number

[6] It should be noted that for a given set of the network parameters k, c, ρ, and v, there are multiple equilibria. In the experiment, we run the experiment multiple rounds. At the beginning of each round, we assign a randomly chosen initial strategy profile $\mathrm{s}^{(0)}$ to all agents in the network. Then in each step t, agents sequentially update their individual strategies using best response dynamics with some inertia probability γ and the resulting strategy profile being $\mathbf{s}^{(t)}$ (i.e. an agent adapts its strategy to its best response with probability $1 - \gamma$ and remains to use its strategy adopted at step $t - 1$ with probability

of hub agents grows at a slower speed than n at the beginning and hence, $n_h(\mathbf{s}^*)/n$ monotonically decreases. However, when n is sufficiently large, the number of spoke agents that a hub agent can support reaches its upper bound. Hence, to support an equilibrium, more hub agents will emerge and $n_h(\mathbf{s}^*)/n$ stops decreasing, which indicates that $n_h(\mathbf{s}^*)$ starts to grow proportionally to n. Figure 4.2(b) shows how the production levels of agents change against n. For a better illustration, we plot the normalized production levels which are compared with their values at $n = 100$ on each curve. As it shows, the normalized production level of spoke agents drops more drastically than that of hub agents as n grows, since a spoke agent mainly relies on information acquired from others rather than self-production and will be more significantly influenced by the change on the network size.

As a direct result of Theorem 3, we quantify in the following corollary the amount of information generated in the network.

Corollary 1. The total amount of information produced in the network at equilibrium, i.e. $\sum_{i \in N} x_i^*$, grows at the same order as the population size n. That is, $\inf_{\mathbf{s}^* \in S_n^*}\{\sum_{i \in N} x_i^*\}$ and $\sup_{\mathbf{s}^* \in S_n^*}\{\sum_{i \in N} x_i^*\}$ are $\Theta(n)$. ■

Therefore, as the network grows, it will become informationally richer with more agent-generated information available.

4.3.3 Equilibrium efficiency

In this section, we analyze the social welfare of the IPLF game to quantitatively study the efficiency of equilibria. The social welfare of the IPLF game is defined to be the sum of agents' individual utilities. For a strategy profile $\mathbf{s} = (\mathbf{x}, \mathbf{g})$, the social welfare is given by $W(\mathbf{x}, \mathbf{g}) \triangleq \sum_{i \in N} u_i(\mathbf{x}, \mathbf{g})$. A strategy profile $\mathbf{s}^\# = (\mathbf{x}^\#, \mathbf{g}^\#)$ achieves the social optimum, denoted as $W^\#$, if and only if

$$W^\# \triangleq W(\mathbf{x}^\#, \mathbf{g}^\#) \geq W(\mathbf{x}, \mathbf{g}), \ \forall (\mathbf{x}, \mathbf{g}). \tag{4.6}$$

γ). Each round terminates when the strategy profile converges to some fixed point $\mathbf{s}$. The result plot in Figure 4.2 is averaged over all rounds of experiments.

In the following proposition, an upper bound on $W^\#$ is given.

Proposition 1. There is a value $\bar{k}$ such that the social optimum $W^\# \in (n[v(\bar{x}) - c\bar{x}],\ n[v(n^{1/\rho}\hat{x}_n) - c\hat{x}_n] - nk/2)$ when $k < \bar{k}$ and $W^\# = n[v(\bar{x}) - c\bar{x}]$ when $k > \bar{k}$. Here $\bar{x}$ is defined in Lemma 3(iii) and $\hat{x}_n$ is the solution of the equation $v'(n^{1/\rho}\hat{x}_n) = c/n^{(1-\rho)/\rho}$. ■

When the social optimum is achieved in an empty network, there is no cost for link formation and thus, we have $W^\# = n[v(\bar{x}) - c\bar{x}]$. On the other hand, if the social optimum is achieved in a non-empty network, it is proven that each agent has at least one connection and hence, the minimum total link cost in the network is $nk/2$ while the maximum total utility from information consumption and production is $n[v(n^{1/\rho}\hat{x}_n) - c\hat{x}_n]$, which gives the corresponding upper bound of the social optimum.

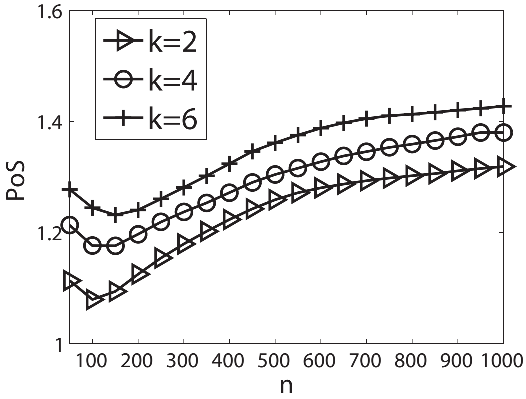
In order to prove the existence of the threshold $\bar{k}$, it is sufficient to show that if the social optimum is achieved in a non-empty network when $k = k_1$, then it is also achieved in a non-empty network when $k < k_1$. Similarly, if the social optimum is achieved in an empty network when $k = k_2$, then it is also achieved in an empty network when $k > k_2$.

Next, we characterize the efficiency of the equilibrium strategy profile. The following proposition provides a lower bound on the social welfare that can be achieved by an equilibrium.

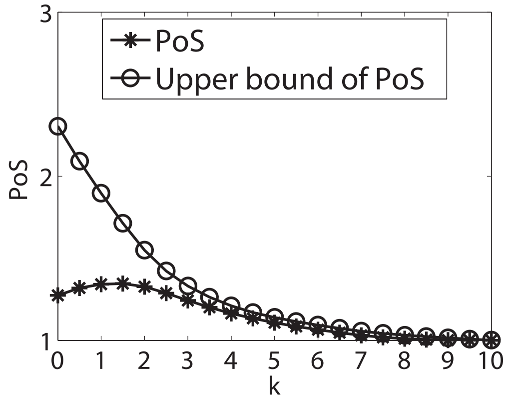
Proposition 2. The social welfare of any equilibrium strategy $\mathbf{s}^*$ is lower bounded by $n[v(\bar{x}) - c\bar{x}]$.

Proof: Since each agent can ensure a utility of $v(\bar{x}) - c\bar{x}$ by choosing not to form any link and produce an amount of $\bar{x}$, the utility it receives in an equilibrium is always no less than this. ■

By combining Proposition 1 and 2, the following theorem provides an upper bound for the Price of Stability (PoS) in the IPLF game. It should be noted that since the lower bound in Proposition 2 applies to all equilibria in the IPLF game, this result also applies



(a)



(b)

Figure 4.3: (a) The PoS against the network size n; (b) The PoS and its lower bound against the link formation cost k ($c = 2,\ \rho = 0.8,\ n = 500$)

to the Price of Anarchy (PoA).

Theorem 4 (Upper Bound for PoS). The Price of Stability satisfies that $PoS \in (1,\ ([v(n^{1/\rho}\hat{x}_n) - c\hat{x}_n] - k/2)/[v(\bar{x}) - c\bar{x}])$ when $k < \bar{k}$ and $PoS = 1$ when $k > \bar{k}$. ■

Figure 4.3 plots the PoS against n and k. As Figure 4.3(a) shows, the PoS monotonically increases with n when it is sufficiently large. This is due to the fact that the set of available strategy profiles in the IPLF game is enlarged along with n, whereas the set of equilibrium strategy profiles remains limited as each hub agent in an equilibrium can only support a limited number of spoke agents. However, the rate for such increase

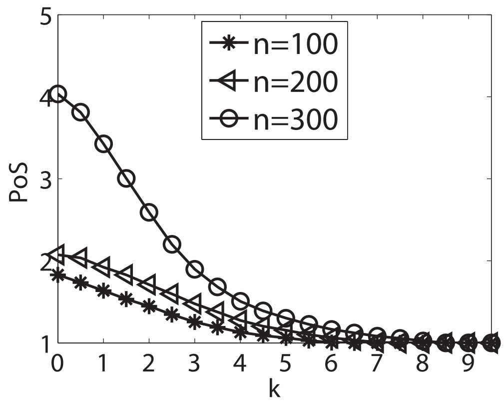


Figure 4.4: The Price of Stability with indirect information sharing ($c = 1,\ \rho = 0.7$)

slows down as n becomes larger. In other words, the PoS saturates when n is sufficiently large. Figure 4.3(b) compares the PoS and our derived upper bound in Theorem 4. The upper bound is not tight when k is small and it is optimal to form as many links as possible in order to increase the information sharing efficiency. When k increases, the upper bound becomes tight, as the topologies produced by both the socially optimal profile and the equilibrium profile become sparse with their social welfare approaching $n[v(\bar{x}) - c\bar{x}]$.

4.4 Indirect Information Sharing

In the basic model, an agent can only consume information produced by itself and its neighbors (i.e. agents it connects via links). In many practical scenarios, the information sharing in social computing applications could be multi-hop. That is, an agent can also consume information which its neighbors acquired from other agents. In this section, we consider such indirect information sharing.

Several concepts are defined below before we formalize the IPLF game with indirect information sharing. Given a connectivity graph $\bar{\mathrm{g}}$, we say that there is a path in g between two agents i and j if either $\bar{g}_{ij} = 1$ or there exist agents $j_1, \ldots, j_m$ distinct from i and j such that $\{\bar{g}_{ij_1} = \bar{g}_{j_1 j_2} = \cdots = \bar{g}_{j_m j} = 1\}$.

Definition 1 (Connected Network). A network is connected if there is a path between every pair of agents in its connectivity graph. A connected network is also called as a component.

Definition 2 (Minimally Connected Network). A network is minimally connected if there is a unique path between every pair of agents on its connectivity graph.

Given any two agents i and j, the distance between them, denoted as $d_{ij}(\bar{\mathbf{g}})$ is defined as the number of hops (links) on the shortest path (the path contains the smallest number of links) between i and j in $\bar{\mathbf{g}}$. If there is no path between i and j, then $d_{ij}(\bar{\mathbf{g}}) = \infty$.

In the IPLF game with indirect information sharing, we assume that each agent can acquire the effective information from its neighbors, which already included the information that the neighbors acquire from other agents. Therefore, it is equivalent that each agent can acquire information produced by all agents who are connected with him via a path. However, if an agent receives certain information produced by another agent from multiple paths, it will automatically remove the redundancy and only keep one copy of the information in consumption. Let $N_i^l(\bar{\mathbf{g}})$ denote the set of agents whose distance to agent i is l, the utility of agent i then can be formulated as follows:

$$w_i(\mathbf{x}, \mathbf{g}) = v((x_i^\rho + \sum_{l=1}^{n-1} \sum_{j \in N_i^l(\bar{\mathbf{g}})} x_j^\rho)^{1/\rho}) - cx_i - k|N_i(\mathbf{g})|. \tag{4.7}$$

In this section, we will use all the terminologies (e.g. equilibrium, social optimum, etc.) that we defined for the IPLF game with direct information sharing in Section 4.3 wherever they apply. With indirect information sharing, it can be proved that the IPLF game always has at least one equilibrium. The resulting equilibria are characterized in the following theorem.

Theorem 5. In the presence of indirect information sharing and when an agent's utility is given by (4.7), there are two values k_{max} and k_{min} for any given n, c, ρ, and v, such that

(i) When $k < k_{max}$, there exists a unique equilibrium where each agent has a production level $\bar{x}$ and no agent forms links;

(ii) When $k < k_{min}$, then each equilibrium preserves the following properties: (a) each agent produces an amount x_n which is the solution of $v(n^{1/\rho}x_n) = c$; (b) the network is minimally connected; (c) each agent has at least one connection to other agents;

(iii) When $k_{min} < k < k_{max}$, there are multiple equilibria each of which contains a minimally connected component with the rest of agents being isolated.

Proof Sketch: First, it is easy to show that for a component of size b in an equilibrium strategy profile $\mathbf{s}^*$, it is always minimally connected with each agent in it having the same production level. Then, we show that at any equilibrium, an agent's utility from information consumption and production monotonically increases with the size of component it is in. With this result, we prove that at any equilibrium, there is at most one component whose size is larger than 1. Suppose in an equilibrium strategy profile $\mathbf{s}^*$ where there are two components C_1 and C_2 of size b_1 and b_2. Without loss of generality, we assume that $1 < b_1 \leq b_2$. Then for an agent i in C_1 who forms a link, it is always beneficial to switching this link to any agents in C_2, which leads to a contradiction to the fact that $\mathbf{s}^*$ is an equilibrium. Now consider another equilibrium where there is an isolated agent j and a component C whose size is $b > 1$. With a little abuse of notation, this equilibrium is also denoted as $\mathbf{s}^*$. It can be shown that agent j has the incentive to connect with the component C if and only if the link formation cost k is smaller than a value γ_b which is determined by the size b, which is proved to monotonically increases with b.

Summarizing all the above, it can be concluded that if $k < \gamma_1$, there is a unique equilibrium which contains a unique component and there is a path between any two agents; if $\gamma_1 < k < \gamma_{n-1}$, there are two equilibria where all agents are either isolated with each other or fully connected in a unique component; if $k > \gamma_{n-1}$, the network has a unique equilibrium where all agents are isolated with each other. Hence, we have

$k_{max} = \gamma_{n-1}$ and $k_{min} = \gamma_1$. ■

Corollary 2. Given c, ρ, and $v(\cdot)$, k_{max} monotonically increases with n, while k_{min} is constant. ■

It should be noted that different from the core-periphery structure in the case of direct information sharing, where high producers are also the main source of information sharing and plays the role of hubs in the network, here the agents with the most connections form the hubs of the network and take the responsibility to share information. Next, we analyze the social optimum of the IPLF game with indirect information sharing. The proof is omitted due to its similarity to Theorem 5.

Theorem 6. In the presence of indirect information sharing and when an agent's utility is given by (4.7),

(i) There is a value $k^{opt}_{max} > k_{max}$ for any given n, c, ρ, and v, such that the social optimum is achieved in an empty network when $k > k^{opt}_{max}$;

(ii) There is a value $k^{opt}_{min} > k_{min}$ for any given n, c, ρ, and v, such that the social optimum is achieved in a minimally connected network when $k < k^{opt}_{min}$;

(iii) When $k^{opt}_{min} < k < k^{opt}_{max}$,the social optimum is achieved in a network which contains a minimally connected component with the rest of agents being isolated. ■

Hence, the Price of Stability in the IPLF game with indirect information sharing, which is plotted in Figure 4.4, achieves 1 when $k > k^{opt}_{max}$ and is strictly larger than 1 when $k < k^{opt}_{max}$. With indirect information sharing, the topologies of the network are similar under both the socially optimal strategy profile and an equilibrium profile. Hence, the difference on the social welfare is highly influenced by how much information the agents produce in total. When k is small, the discrepancy between the total information productions of the socially optimal strategy profile and an equilibrium profile is higher than that when k is large, which gives a PoS monotonically decreasing against k.

4.5 Conclusion

In this chapter, we investigate the problem of information production and network formation in social computing systems. Different from the existing literature, the agents' incentives for producing information themselves and for forming links to consume the information of others are jointly considered. Moreover, we determine rigorously how the agents' appreciation for information variety impacts their interactions and the emerging connectivity between them. We then analyze the efficiency of the emerging equilibrium topologies and compare their performances with that of the social optimum. We also study the IPLF game which incorporates indirect information sharing. Our analysis can be extended in several directions, among which we mention three. First, we assume in this model that agents are homogeneous in terms of their benefit functions and costs. The analysis on heterogeneous agents would represent an important future research direction which can hopefully provide new insights into the IPLF game, e.g. agents' heterogeneity could lead to the selection of equilibria and reduce the set of possible equilibrium topology as predicted in [66]. Second, alternative formulations on information transmission and link formation can be extended from our current model to encompass the features of various applications. Several examples include the bilateral link formation which requires the mutual consent of agents, the unilateral information transmission with which the information flow over a link is one-sided, and non-constant link formation cost with which the cost of establishing a link depends on the characteristics of the creator and the recipient as well as the amount of information (i.e. traffic) across this link. Finally, the design of effective incentive protocols (e.g. pricing schemes to subsidize or tax link formation) to stimulate the information production and sharing in order to reduce the Price of Stability and achieve the social optimum also forms an important future research direction.

CHAPTER 5

Information Dissemination in Socio-technical Networks

5.1 Introduction

The traditional analysis of communication networks assumes that the topology of the network is fixed exogenously or determined by the central designer and that the actions of users are obedient to the wishes of the designer. (It is typically assumed that the objective of the designer is to maximize social welfare [77][78], but other objectives might be considered as well.) However, these assumptions do not apply at all to social networks such as Facebook [79] and Twitter [72], expert networks such as Amazon Mechanical Turk [42], vehicular networks [61], social mobile networks [59], peer-to-peer overlay routing systems [80], etc. To the contrary, the topologies of such networks are determined endogenously by the actions of self-interested and strategic users (which leads us to use the term "strategic networks"). Aspects of strategic networks that are of particular interests include the topology that emerges, the efficiency/inefficiency of behavior and especially the protocols that the designer might implement to promote social welfare even in the face of self-interested behavior by users.

A central aspect of the strategic networks we study here is that links in the network can be created and maintained by individual users. Creating and maintaining links are costly, so will only be carried out if they provide sufficient benefits for these individual users. In some earlier work, especially in the economics literature, the benefits of links are that they permit the acquisition of information (files in P2P networks, news on social networks, traffic/road conditions in vehicular networks, etc.); our point of departure in this chapter is that we emphasize that links also permit the dissemination of information (advertising/marketing in social networks [81][82][88], routing traffic in overlay routing systems [83], etc.). The central message of this chapter is that the emphasis on the dissemination of information leads to very different conclusions about the structure of networks and the behavior of users in those networks.

To be precise, we consider the behavior of a group of self-interested and strategic agents/users who may create and maintain links to other users, and produce and disseminate information. We formalize the strategic interactions among agents as a non-cooperative game, the Information Dissemination Game (IDG). Because we emphasize information dissemination, we assume that link formation is unilateral: the decision to create a link from one agent to another is made unilaterally by the first agent and the cost of creating that link is borne entirely by the agent who creates it; this is a reasonable description of behavior and cost when the benefit of creating a link is disseminating information (advertising) rather than gathering information (see e.g. [63][64][86]). We consider a setting in which agents and information are both heterogeneous: agents differ in terms of their locations, access to devices, information and link production capabilities and costs. Agents are self-interested: each intends to maximize its own benefit from information dissemination net of the cost of the links it forms. Our notion of solution in the Information Dissemination Game is a non-cooperative equilibrium.

We prove first that the typical network that emerges from the self-interested behavior of agents displays a core-periphery structure, with a smaller number of agents at the core (center) of the network and a larger number of agents at the periphery (edges) of the network. Agents in the core create many links and communicate with many other agents; agents in the periphery create few (or no) links and communicate mostly (or entirely) with agents in the core. We go on to show that the typical networks that emerge are minimally connected and have short network diameters, which are independent of the size of the network. When agents' strategic behavior incorporates both link formation and information production (with the objective of maximizing information dissemination), we show that the number of agents who produce information and the total amount of information produced grow with the size of the network; this is in sharp contrast with the "law of the few" which has been demonstrated in [64] with settings where the purpose of forming links is the acquisition of information.

Our analysis is important for a number of reasons. At the theoretical level: small diameters tend to make information dissemination efficient and minimal connectivity tends to minimize the total cost of constructing the network. At the empirical level: they are consistent with the findings of numerous empirical investigations. More generally, our analysis provides guidance and tools for network designers to create protocols providing incentives for agents to take actions that are consistent with self-interest and still promote social welfare.

The remainder of this chapter is organized as follows. Section 5.2 describes our basic model of the IDG. Section 5.3 characterizes the non-cooperative equilibria that emerge in the basic model. Section 5.4 analyzes the IDG with strategic information production. Section 5.5 discusses the related literature and Section 5.6 concludes.

5.2 System Model

5.2.1 Settings

In this section, we propose a basic model to formulate the IDG, in order to capture the fundamental trade-offs between agents' benefit and cost from strategic information dissemination. Although simple, our formulation already provides qualitative insights on how the incentives of self-interested agents impact the network structure, and can be applied to numerous network applications (with slight modifications).

Let $N = \{1, 2, \ldots, n\}$ be the set of agents in the system with $n \geqslant 3$ and let i and j denote typical agents. Each agent i possesses some information in the amount $x_i \in \mathbb{R}^+$, which it finds in its own benefit to disseminate to other agents. We consider a non-cooperative game where each agent strategically determines whether to create links with other agents in order to disseminate its information. As in e.g. [64][73], links are created by the unilateral actions of an agent who bears the entire cost [1]. Thus, the mutual consent of two agents is not required in order to create a link between them. The link formation strategy adopted by an agent i is denoted by a tuple $\mathbf{g}_i = (g_{ij})_{j\in\{1,\ldots,n\}/\{i\}} \in \{0,1\}^{n-1}$; $g_{ij} = 1$ if agent i forms a link with agent j and $g_{ij} = 0$ otherwise. The creation of a link incurs a cost to the creator and hence, the decision to form a link involves trading-off the benefit received from disseminating information using this link and the incurred cost. A strategy profile in the information dissemination game is defined as $\mathbf{g} \triangleq (\mathbf{g}_i)_{i=1}^n \in \mathbf{G}$, where $\mathbf{G}$ is a finite space.

The information flow across a link is assumed to be undirected. That is, given a link between any two agents, the information can be transmitted in both directions (i.e. from the creator to the recipient and vice versa) across this link. We thus define the *topology* of the network as $E_{\mathbf{g}} = \{(i,j) \in N \times N | i \neq j \; and \; \max\{g_{ij}, g_{ji}\} = 1\}$. In the rest of this chapter, we will use the terms "topology" and "network" interchangeably. Given a topology $E_{\mathbf{g}}$, a path between two agents i and j is a sequence $path_{ij} = \{(i,j_1),(j_1,j_2),\ldots,(j_m,j)\}$ for some $m \geqslant 0$ such that $path_{ij} \subseteq E_{\mathbf{g}}$. Agent i can *reach* an agent j in a topology $E_{\mathbf{g}}$, denoted $i \to j$, if and only if there is at least one path from agent i to j in $E_{\mathbf{g}}$, otherwise i cannot reach agent j, denoted $i \nrightarrow j$. We assume that an agent i can disseminate its information to every agent j whom it can reach. Given this, the utility of an agent i in the IDG can be expressed as:

$$u_i(\mathbf{g}) = f\left(x_i | N_i(E_{\mathbf{g}})|\right) - \sum\nolimits_{j\in N_i(\mathbf{g})} k_{ij}. \tag{5.1}$$

Here $N_i(E_{\mathbf{g}}) \triangleq \{j | i \to j\}$ is the set of agents whom agent i can reach, and $N_i(\mathbf{g}) \triangleq \{j | g_{ij} = 1\}$ is the set of agents with whom agent i forms links. $f\left(x_i | N_i(E_{\mathbf{g}})|\right)$ thus

[1]The precise formulations of link formation and information flow among participating agents in an IDG depend on details of the considered application. Due to the infeasibility of enumerating all possible models, we use a stylized model in this work as an example to formulate the IDG. Our current formulation has the great merit of being simple to work with and can be applied to most existing applications to date with slight modifications, e.g. the telephone networks [101] and the Voice over IP applications such as Skype [102].

represents the total benefit that agent i receives from information dissemination, which depends on the amount of information it disseminates, i.e. x_i, as well as the total number of agents it can reach, i.e. $|N_i(E_{\mathbf{g}})|$. We assume that $f(\cdot)$ is twice continuously differentiable, increasing and concave with $f(0) = 0$. Hence, an agent's benefit increases, while the marginal benefit decreases, with x_i and $|N_i(E_{\mathbf{g}})|$. $\sum_{j \in N_i(\mathbf{g})} k_{ij}$ represents the total link formation cost of agent i, where $k_{ij} \in \mathbb{R}^+$ denotes the cost for agent i to form a link with agent j.

We assume that an agent cannot benefit from disseminating duplicated copies of its information to any other agent. That is, when there are multiple paths from agent i to agent j and multiple copies of agent i's information arrive at agent j, agent i receives a fixed benefit regardless of the number of copies that agent j receives. We assume that each agent benefits only from disseminating its own information, and forwarding the information that is received from other agents does not bring it any benefit.

5.2.2 Equilibrium and social welfare

We consider pure (not mixed) link formation strategies. Each agent maximizes its own utility given the strategies of others. A Nash equilibrium (NE) is defined as a strategy profile $\mathbf{g}^*$ such that the strategy of each agent i is a best response to the strategies of others:

$$u_i(\mathbf{g}_i^*, \mathbf{g}_{-i}^*) \geqslant u_i(\mathbf{g}_i, \mathbf{g}_{-i}^*), \forall \mathbf{g}_i \in \{0,1\}^{n-1}, \ \forall i \in N. \tag{5.2}$$

Here $\mathbf{g}_{-i}$ represents the strategies of all agents other than agent i. The set of NE is defined as $\mathbf{G}^* = \{\mathbf{g}^* | \mathbf{g}^* \ satisfies \ (2)\}$. A strict NE is an NE such that the strategy of each agent i is a strict best response to the strategies of others (with the inequality in (2) being strict whenever $\mathbf{g}_i^* \neq \mathbf{g}_i$). It is shown in the online appendix [23] that a network will always converge to a strict NE in a dynamic link formation process. Therefore, a strict NE characterizes a steady state in the dynamic link formation process. Note that strict NE are NE and thus, the results below on NE also apply to strict NE.

The social welfare of the IDG is defined to be the sum of agents' individual utilities. For a strategy profile g, the social welfare is given by $U(\mathbf{g}) \triangleq \sum_{i \in N} u_i(\mathbf{g})$. A strategy profile $\mathrm{g}^{\#}$ is called socially optimal if it achieves the social optimum, denoted by $U^{\#}$, i.e.

$$U^{\#} \triangleq U(\mathbf{g}^{\#}) \geqslant U(\mathbf{g}),\ \forall \mathbf{g} \in \mathbf{G}. \tag{5.3}$$

5.3 Equilibrium and Efficiency Analysis of the IDG

This section studies the IDG described in Section 5.2. First, we analyze the equilibrium link formation strategies of individual self-interested agents. Next, we explicitly compare the equilibrium social welfare of the IDG to the social optimum. The results provide important insights on the efficiency loss occurred due to the self-interested behavior of the agents in the IDG as compared to the case when the agents obediently follow the link formation actions dictated by some central designer.

5.3.1 Equilibrium analysis

Given a strategy profile g, a *component* C is a set of agents such that $i \rightarrow j,\ \forall i, j \in C$ and $i \nrightarrow j',\ \forall i \in C$ and $\forall j' \notin C$. Hence, each component defines a connected sub-network in a network $E_{\mathbf{g}}$: any two agents in this component can mutually reach each other, whereas no agent in the component can reach any other agent outside the component. An agent who is not connected with any other agents in the network (i.e. an isolated agent) forms a component by itself, called a *singleton* component; a component that is not singleton is called a *non-singleton* component. A component C is called *minimal* if and only if there is only one path in $E_{\mathbf{g}}$ from any agent $i \in C$ to any other agent $j \in C$. The *shortest path* from agent i to j is the path that contains the minimum number of links. The *distance* $d_{\mathbf{g}}(i, j)$ between i and j is the number of links on a shortest path between them. By convention, $d_{\mathbf{g}}(i, j) \triangleq \infty$ when $i \nrightarrow j$. The *diameter* of a component C is defined as the largest distance between any two agents in it, which is denoted as $D_C \triangleq \max\limits_{i,j \in C} d_{\mathbf{g}}(i,j)$. The diameter of a singleton component is defined to be 0. The diameter of the network is defined to be the largest diameter of all components it contains.

It should be noted that the strategy space for each agent in the IDG is compact and convex. Meanwhile, an agent's utility is quasi-concave over its link formation strategy. Hence, it has been shown in [90] that pure NE always exists in the IDG. We first derive some basic properties of the equilibria in the IDG. Although simple, these properties are important for characterizing the emerging equilibria later.

Proposition 1. Under an NE $\mathbf{g}^*$ of the IDG, each component is minimal.

Proof: Suppose that there is a component C such that there are two agents i and j who are connected by two paths $path_{ij}$, $path'_{ij} \subseteq E_{\mathbf{g}^*}$ with $path_{ij}/path'_{ij} \neq \phi$. Here $path_{ij}/path'_{ij}$ represents the relative complement of $path'_{ij}$ in $path_{ij}$ and contains all elements that belong to $path_{ij}$ but do not belong to $path'_{ij}$. Then there are always two agent i' and j' in C who satisfy: (1) $g^*_{i'j'} = 1$; and (2) there is a $path_{i'j'} \subseteq E_{\mathbf{g}^*}$ such that $path_{i'j'} \neq ((i', j'))$. Therefore, by setting $g_{i'j'} = 0$, agent i' always receives a strictly higher utility compared to what it can receive in $\mathbf{g}^*$, which contradicts the fact that $\mathbf{g}^*$ is an NE. Hence, this proposition follows. ■

Proposition 1 shows that in an equilibrium of the IDG, each connected sub-network (component) is minimal with no cycles in it. As we will show in Section 5.3.2, the social optimum in the IDG is always achieved by networks consisting of minimal components and hence, the equilibria in the IDG can frequently achieve the social optimum (i.e. being efficient).

Proposition 1 characterizes individual components in the equilibrium network, but it does not characterize the connectedness of the network, i.e. whether the network will be composed of a unique component where all agents are connecting with (and can disseminate information to) each other or several components that are isolated from each other. The following proposition provides a sufficient condition under which the network is connected at equilibrium.

Proposition 2. The network in each NE is always minimally connected if there is an agent i such that $f(x_i(|N|-1)) - f(x_i(|N|-2)) > \max_{j \in N}\{k_{ij}\}$.

Proof: Suppose there is a NE $\mathbf{g}^*$ which contains more than one component. We consider two components C_1 and C_2. Suppose agent i is in C_1, then it can always increase its utility by forming a link to any other agent in C_2 since $f(x_i(|N|-1)) - f(x_i(|N|-2)) > \max\{k_{ij}\}_{i,j \in N}$, which contradicts the fact that $\mathbf{g}^*$ is an NE. This proposition thus follows. ■

Proposition 2 shows that the network will be connected at equilibrium when the benefit from information dissemination is sufficiently large (i.e. x_i is sufficiently large) with respect to the link formation cost. The properties of the network topology at equilibrium (i.e. the shape and diameter of the network) depends on the specific values of $\{x_i\}_{i \in N}$ and $\{k_{ij}\}_{i,j \in N}$. In the rest of this section, we analyze two exemplary networks with particular structures in order to obtain further insights on the equilibrium topology.

5.3.1.1 Networks with recipient-dependent costs

In the first example, we consider the network where the cost of forming a link is exclusively recipient specific. In particular, we have $k_{ij} = k_j, \forall i \in N/\{j\}$. This can capture the practical networks in which the link formation cost only depends on the type of the recipient and there are some agents to which it is easier to connect with than other agents (i.e. with smaller costs to form links with). For example, in networks where the link formation cost represents the subscription fee that the creator sends to the recipient, each agent charges the same price to any agent who wants to form a link with it. In the following theorem, we show that if the link formation cost in the network is not arbitrary but only takes values from a finite set $\{k^1, ..., k^L\}$, i.e. there are L different types of link formation costs and $k_i \in \{k^1, ..., k^L\}, \forall i \in N$ (e.g. the subscription fee is quantized to several discrete levels depending on the agents' types but not takes arbitrary values), then the diameter of the network at each strict equilibrium should be no more than $2L + 2$.

Theorem 1. Suppose that there are L different types of link formation costs, e.g. $\{k^1, ..., k^L\}$, such that $k_i \in \{k^1, ..., k^L\}, \forall i \in N$, then under any strict NE $\mathbf{g}^*$ of the IDG, the diameter of the network is at most $2L + 2$.

Proof: Consider a strict NE $\mathbf{g}^*$ and a non-singleton component in it, there is at least one agent i in the component such that $g^*_{ij} = 1$ for some j. Consider a path $path_{ij_d} = ((i, j), (j, j_1), (j_1, j_2), \dots, (j_{d-1}, j_d))$. Since $g^*_{ij} = 1$, we should have $\{k_j < k_{j_1}, k_j < k_{j_2}, ..., k_j < k_{j_d}\}$. Now there are two cases: $g^*_{j_1 j} = 1$ or $g^*_{j j_1} = 1$.

In the first case, suppose $g^*_{j_2 j_1} = 1$, then we have $k_{j_1} < k_j$, which leads to a contradiction to the fact that $k_j < k_{j_1}$. Hence, we have $g^*_{j_1 j_2} = 1$, which gives $k_{j_2} < k_{j_3}$. Using the same arguments, we have $g^*_{j_l j_{l+1}} = 1$ and $k_{j_{l+1}} < k_{j_{l+2}}$, for all $l \in \{1, ..., d - 2\}$. Therefore, along the path $path_{jj_d} = ((j, j_1), (j_1, j_2), \dots, (j_{d-1}, j_d))$, there are at least d different link formation costs. We thus have $d \leqslant L$ and the length of path $path_{ij_d}$ is smaller than $L + 1$.

Now consider the second case where $g^*_{jj_1} = 1$. Then we have $k_{j_1} < k_{j_2}$. Similar to that for the first case, we have that on $path_{jj_d} = ((j, j_1), \dots, (j_{d-1}, j_d))$, there are at least d different link formation costs and hence the length of path $path_{ij_d}$ is still no more than $L + 1$.

Consider the longest path in this component, which is denoted as

$$((b_0, b_1), (b_1, b_2), \dots, (b_{T-1}, b_T)). \tag{5.4}$$

Suppose $T > 2(L+1)$ and consider the agents $b_{\lfloor T/2 \rfloor}$ and $b_{\lfloor T/2 \rfloor + 1}$. If $g^*_{b_{\lfloor T/2 \rfloor} b_{\lfloor T/2 \rfloor + 1}} = 1$, then the path $((b_{\lfloor T/2 \rfloor}, b_{\lfloor T/2 \rfloor + 1}), \dots, (b_{T-1}, b_T))$ has a length longer than $(L + 1)$. If $g^*_{b_{\lfloor T/2 \rfloor + 1} b_{\lfloor T/2 \rfloor}} = 1$, then the path $((b_{\lfloor T/2 \rfloor + 1}, b_{\lfloor T/2 \rfloor}), \dots, (b_1, b_0))$ has a length longer than $(L + 1)$. Both scenarios contradict our argument above. Therefore, we can conclude that $T \leqslant 2(L + 1)$ always holds and this theorem follows. ■

The link formation cost thus plays an important role in shaping the equilibrium network in the IDG. As shown in Theorem 1, if there are only a finite number of different link formation costs in the network, then the size of each component (a connected sub-network) cannot be arbitrarily large but is upper-bounded by some constant value, which is independent of the population size but proportional to the number of different link formation costs. Based on Proposition 1 and Theorem 1, the "minimally connected" and "short diameter" properties of the equilibria in strategic networks are thus proven.

As a special case of Theorem 1, we prove in the following corollary that when the link formation cost is the same for all agents, each component in a strict NE forms a star topology, regardless of the values $\{x_i\}_{i \in N}$.

Corollary 1. If $k_{ij} = k, \forall i, j \in N$, then under a strict NE g*, each non-singleton component forms a star topology.

Proof: See Appendix. ■

Hence when the link formation cost is the same for all agents, each component at equilibrium preserves the "core-periphery" property with one single agent staying at the center of it and playing the role of the "connector" who connects (maintains links) with all other agents to support their information dissemination.

5.3.1.2 Networks with groups

We discuss a network where agents are divided into groups and agents within the same group have the same type. The cost of forming links within a group (i.e. between agents of the same type) is lower than the cost of forming links across groups (i.e. between agents of different types). Examples of strategic networks where such groups exist are users of close social relationships or close interests in a social network [92], devices or processing nodes located in the same area [93], etc.

Formally, we consider that all agents are divided into Z different groups $N_1, \ldots, N_Z$ with $|N_z| \geqslant 2$ for all $1 \leqslant z \leqslant Z$, such that $N = \cup\{N_z\}_{z=1}^{Z}$ and $N_z \cap N_{z'} = \phi$ for any $1 \leqslant z < z' \leqslant Z$. For two agents from the same group, the cost of forming a link between them is $\underline{k}$, while for two agents from different groups, the cost of forming a link between them is $\bar{k} > \underline{k}$. Here we assume that $x_i = x, \forall i \in N$ to make our analysis tractable. The following theorem characterizes the strict equilibria with the presence of groups and proves that each non-empty strict equilibrium preserves the "core-periphery" property.

Theorem 2. In the presence of groups, the Nash equilibria can be characterized as follows:

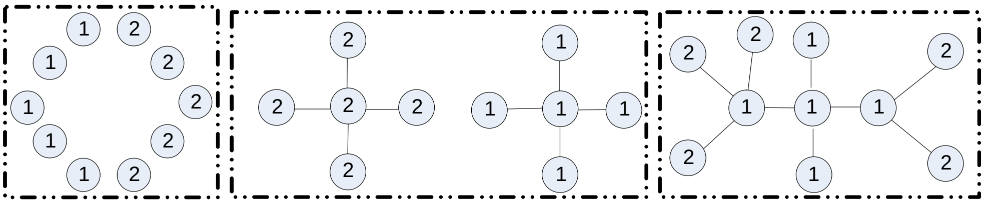
(i) When $f(x) < \underline{k}$, the unique strict NE $\mathbf{g}^*$ satisfies $g^*_{ij} = 0, \forall i, j$;

(ii) When $f(x) \in (\underline{k}, \bar{k})$, the unique strict NE consists of Z components, where each component contains only agents from the same group and the topology of each component is a star;

(iii) When $f(x) > \bar{k}$, in each strict NE $\mathbf{g}^*$, there is a group N_z and an agent $i \in N_z$ such that $g^*_{ij} = 1, \forall j \in N_z/\{i\}$. Also for each agent $j' \notin N_z$, there is an agent $j \in N_z$ such that $g^*_{jj'} = 1$.

Proof: See Appendix. ■

Several examples of the equilibrium topologies discussed in Theorem 2 are illustrated in Figure 5.1 in a network of $n = 10$ agents who are divided into 2 groups. The number on each node represents the group to which each agent belongs. Theorem 2 provides several important insights. First, in a strict equilibrium, agents from the same group always belong to the same component (i.e. are connected with each other). Second, each non-singleton component exhibits the "core-periphery" property. The agents that form the core are from the same group, while agents from other groups access the network via links maintained by the core. This analytical finding is reflected in numerous real world examples. For instance, in a large-scale overlay routing network [83], it is usually the case that a group of nodes who can inter-connect at a lower cost form the backbone of the network, while all other nodes connect to the network via this back-



(a) $f(x) < \underline{k}$ (b) $f(x) \in (\underline{k}, \bar{k})$ (c) $f(x) > \bar{k}$

Figure 5.1: The exemplary Nash equilibria in the network with groups

bone. Third, in each component, there is always a central agent and all paths within this component initiate from this agent. Also, the distance from the central agent to any periphery agent is no more than 2. Hence, the diameter of the network is no more than 4, which is also independent of the population size in the network.

5.3.2 Equilibrium efficiency of the IDG

In this section, we analyze the efficiency (social welfare) of the IDG. Because there are multiple equilibria, we use two metrics to measure the equilibrium efficiency: (i) the Price of Stability (PoS) is defined as the ratio between the social optimum and the highest social welfare that is achieved at equilibrium and measures the efficiency of the "best" equilibrium in the IDG, i.e. $PoS = U^{\#}/\max_{\mathbf{g}^* \in \mathbf{G}^*} U(\mathbf{g}^*)$; (ii) the Price of Anarchy (PoA) is defined as the ratio between the social optimum and the lowest social welfare that is achieved at equilibrium and measures the efficiency of the "worst" equilibrium in the IDG, i.e. $PoA = U^{\#}/\min_{\mathbf{g}^* \in \mathbf{G}^*} U(\mathbf{g}^*)$. In the rest of this section, we quantify the PoS and PoA in the IDG and show with multiple examples that the equilibria in the IDG frequently achieve the social optimum.

We first characterize the socially optimal strategy profiles. As with NE, it can be proven that the minimal property still holds in the network under any socially optimal strategy profile.

Proposition 3. In the IDG, each component under a socially optimal profile is always minimally connected.

Proof: This can be proven using the same idea as Proposition 1. ■

With the minimal property, we prove in the next theorem that when the link formation cost is recipient-dependent, i.e. $k_{ij} = k_j, \forall i \in N/\{j\}$, there is always an NE that can achieve the social optimum when the link formation cost is sufficiently small: the PoS of the IDG is always 1.

Theorem 3. If $k_{ij} = k_j, \forall i \in N/\{j\}$ and $\min\limits_{i\in N} f(x_i) \geqslant \min\limits_{i\in N} k_i$, the PoS of the IDG is always 1.

Proof: Let $k_{i_0} = \min\limits_{i\in N} k_i$ and consider a periphery-sponsored star $\mathbf{g}$ with $g_{ji_0} = 1, \forall j \in N/\{i_0\}$ and $g_{jj'} = 0, \forall j, j' \in N/\{i_0\}$. It is obvious that $\mathbf{g}$ is both social optimal and an NE. Hence, this theorem is proven. ■

However, the PoA of the IDG is not necessarily 1 in this case, i.e. there are some NE that incur positive efficiency loss. This is quantified in the next proposition.

Proposition 4. If $k_{ij} = k_j, \forall i \in N/\{j\}$ and $\min\limits_{i\in N} f(x_i) \geqslant \min\limits_{i\in N} k_i$, the PoA of the IDG is upper-bounded by $\max_{i,j\in N} k_i/k_j$.

Proof: We consider an arbitrary NE $\mathbf{g}^*$. Let $k_{i_0} = \min\limits_{i\in N} k_i$ and consider a component C_1 that contains i_0. Now consider another component C_2. If C_2 is a singleton component which contains a unique agent j, then j can always increase its utility by forming a link with i_0. If C_2 is a non-singleton component, then there is always an agent $j' \in C_2$ such that $g^*_{j'l} = 1, \exists l \in C_2$ with $\max\{g^*_{ll'}, g^*_{l'l}\} = 0, \forall l' \in N/\{l\}$. In this case, j' can also increase its utility by switching its link from l to i_0. Therefore, it can be concluded that $\mathbf{g}^*$ forms a connected network. From Proposition 3, we know that the network formed by $\mathbf{g}^*$ is also minimal which contains $|N| - 1$ links and hence, $U^{\#}/U(\mathbf{g}^*) \leqslant \max_{i,j\in N} k_i/k_j$. Since this conclusion applies to any NE $\mathbf{g}^*$, Proposition 4 thus follows. ■

Proposition 4 shows that the upper bound of PoA depends on how the link formation cost varies among agents. In the special case where the link formation cost is the same for all agents, i.e. $k_{ij} = k, \forall i, j \in N$, each NE $\mathbf{g}^*$ achieves the social optimum: PoA is

1.

Corollary 2. If $k_{ij} = k, \forall i, j \in N$ and $\min_{i \in N} f(x_i) \geqslant k$, every NE $\mathbf{g}^*$ achieves the social optimum and the PoA of the IDG is 1.

Proof: This can be proven straightforwardly using Proposition 4. ■

5.4 IDG with Information Production

In the IDG discussed so far, we assumed that the information possessed by the each agent is exogenously determined and fixed during the game. Nevertheless in practical networks, it is usually the case that each agent i can proactively determine the amount of information that it wants to disseminate throughout the network, i.e. the value of x_i. In this section, we consider such IDG with strategic information production from individual agents.

In the IDG with information production, the strategy of an agent i can be represented as $(x_i, \mathbf{g}_i)$, and the agent jointly maximizes its decisions on the information production and link formation in order to maximize its overall utility from information dissemination. A strategy profile of the IDG with information production is written as $\mathbf{s} = (\mathbf{x}, \mathbf{g})$, where $\mathbf{x} = (x_i)_{i=1}^n$ denotes the information production decisions of all agents. Given a strategy profile, the utility of agent i is expressed as:

$$u_i(\mathbf{x}, \mathbf{g}) = f(x_i | N_i(E_{\mathbf{g}})|) - cx_i - \sum\nolimits_{j \in N_i(\mathbf{g})} k_{ij}. \tag{5.5}$$

Here cx_i represents the cost of producing an amount x_i of information, where c is the unit production cost.

A Nash equilibrium of the IDG with information production is a strategy profile $\mathbf{s}^* = (\mathbf{x}^*, \mathbf{g}^*)$ such that

$$u_i(\mathbf{s}_i^*, \mathbf{s}_{-i}^*) \geqslant u_i(\mathbf{s}_i, \mathbf{s}_{-i}^*), \forall \mathbf{s}_i \in \mathbb{R}^+ \times \{0, 1\}^{n-1}, \forall i \in N. \tag{5.6}$$

When information production is a strategic choice, central questions are how many agents will produce information at equilibrium and how the total amount of information produced in the network changes with the population size. The seminal work in [69] analyzes the network formation game with information production where agents benefit from acquiring and consuming the information produced by other agents. It predicts the occurrence of the "law of the few" at equilibrium. That is, in each equilibrium there are only a small number of agents in the network who produce a positive amount of information (i.e. being information producers). As the population size grows to infinity, the fraction of information producers in the agent population goes to 0. Based on the "law of the few", [69] also predicts that the total amount of information that is produced (by all agents) in the network remains constant at equilibrium, which is independent of the population size.

The reason for the emergence of the "law of the few" in [69] is that each agent benefits solely from information consumption and hence its utility is not affected by how many agents it connects with and with whom it is connected so long as the total amount of its acquired information remains constant. Also, when there is a sufficient amount of information that has been acquired by an agent, it will stop producing information personally. Therefore, the information production at equilibrium is always dominated by a small fraction of information producers who produce all the information to be consumed by all agents.

In the rest of this section, we study the asymptotic information production behavior of agents in the IDG when the population size grows. It should be noted that in the IDG, the benefit of an individual agent is jointly determined by the amount of its own production, i.e. x_i, as well as the number of agents it connects with, i.e. $|N_i(E_{\mathbf{g}})|$, whereas the information produced by other agents has no influence on its information production decision. This makes the resulting asymptotic information production behavior at equilibrium exhibits significant differences to that in [69]. Importantly, we prove in the following theorem that the "law of the few" does not hold in the IDG. To illustrate this theorem, we define several auxiliary variables: S_N^* represents the set of equilibrium strategy profiles when the population is N, $I_N(\mathbf{s}) = \{i | i \in N \; and \; x_i^* > 0\}$ represents the set of information producers under the strategy profile $\mathbf{s}$, and $\bar{x}$ represents the solution of the equation $f'(x) = c$.

Theorem 4. In the IDG with information dissemination, when $\max\limits_{i,j \in N} k_{ij} < c\bar{x}$, (i) $\inf\limits_{\mathbf{s}^* \in S_N^*} \{I_N(\mathbf{s}^*)\}/|N| = 1$; (ii) the total amount of information produced in the network at equilibrium, i.e. $\sum\limits_{i \in N} x_i^*$, grows to infinity when the population size $|N| \to \infty$, i.e. $\lim\limits_{|N| \to \infty} \inf\limits_{\mathbf{s}^* \in S_N^*} \{\sum\limits_{i \in N} x_i^*\} \to \infty$.

Proof: To prove Statement (i), it is sufficient to see that each agent i will connect to at least one other agent in any NE and thus have $x_i^* > 0$ when $\max_{ij} k_{ij} < c\bar{x}$, which is independent of the population size $|N|$.

By taking the first order derivative of (4), we have that $|N_i(E_{\mathbf{g}^*})| f'(|N_i(E_{\mathbf{g}^*})| x_i^*) = c$ and thus $|N_i(E_{\mathbf{g}^*})| x_i^* \geqslant \bar{x}$. Also, for any two agents i, j within the same component, we have $x_i^* = x_j^*$. Therefore for any component C, the total amount of information produced by agents within this component at equilibrium is $\sum\limits_{i \in C} x_i^*$, which satisfies $f'(\sum\limits_{i \in C} x_i^*) = c/|C|$ and $\sum\limits_{i \in C} x_i^* > \bar{x}$. Suppose that there is a sufficiently large constant W such that for any N we have $\inf\limits_{\mathbf{s}^* \in S_N^*} \{\sum\limits_{i \in N} x_i^*\} < W$. Select $\mathbf{s}_N^* = \arg \inf\limits_{s^* \in S_N^*} \{\sum\limits_{i \in N} x_i^*\}$. Due to the concavity of $f(\cdot)$, we have that $f'(\sum\limits_{i \in C} x_i^*) = c/|C| \geqslant f'(W)$ for any component C under $\mathbf{s}_N^*$. Hence, $|C| \leqslant c/f'(W)$, and we have $\sum\limits_{i \in N} x_i^* \geqslant |N| f'(W) \bar{x}/c$. This shows that there is always a sufficiently large $|N|$ such that $|N| f'(W) \bar{x}/c > W$ which contradicts the assumption that $\inf\limits_{\mathbf{s}^* \in S_N^*} \{\sum\limits_{i \in N} x_i^*\} < W$ for any N. Therefore, we have a contradiction and Statement (ii) follows. ■

Theorem 4 shows that when agents benefit from information dissemination instead of information consumption, both the number of information producers and the total amount of information produced in the network grow at least at a linear order of the population size at equilibrium. Therefore, the information production at equilibrium is no longer dominated by a small number of information producers and the "law of the few" predicted in [69] no longer holds.

5.5 Related Works

There is a broad literature studying the information dissemination in social networks [94]-[100], which focuses on explaining how the information (e.g. epidemics, job openings, etc.) is propagated in social networks and how the agents' actions (e.g. becoming "infected" or not, buying products or not, etc.) are influenced by the disseminated information. However, the analysis in these works is based on the assumption that the underlying topologies of the social networks are exogenously determined and none of them explicitly considers the strategic link formation of self-interested agents.

There are also numerous works in network science investigating the evolution of social and information networks whose topologies are formed endogenously by agents' self-interested actions [87]-[89]. These works focus on empirical measurements of existing social networks and they fail to provide theoretical foundations which can explain and emulate the relationship between agents' incentives to form links based on their own self-interest and the emerging network topologies.

Theoretical study of network formation in social and economic networks has been conducted by micro-economists as well as computer scientists (see e.g. [63][64][69][84]-[86]), who analyze how the agents' self-interest in acquiring information from other agents leads to strategic link formation and particular network topologies. However, these works focus on the scenario in which agents benefit solely from consuming acquired information. In the rest of this section, we discuss the relationship and differences between our proposed information dissemination game and the existing models on network formation games with information acquisition.

5.5.1 Differences in agents' utilities and incentives

First, note that analysis of the IDG and that of the network formation game with information acquisition (which is referred to as the Information Acquisition Game (IAG) below) exhibit significant differences in both the agents' utilities and the problem formulation. To illustrate the differences, we write the utility function in the IDG below as well as an exemplary utility function in the IAG. To make the key differences even clearer, we assume that $k_{ij} = k, \forall i, j \in N$ throughout the analysis:

$$u_i^{IDG}(\mathbf{g}) = f(x_i|N_i(E_{\mathbf{g}})|) - k|N_i(\mathbf{g})|, \tag{5.7}$$

$$u_i^{IAG}(\mathbf{g}) = f(x_i + \sum\nolimits_{j \in N_i(E_{\mathbf{g}})} x_j) - k|N_i(\mathbf{g})|. \tag{5.8}$$

From these two utility functions, it can be observed that in the IAG, the benefit of an individual agent is determined by the total amount of information which it acquires, i.e. $x_i + \sum_{j \in N_i(E_{\mathbf{g}})} x_j$, and agents lose the incentive to form links as long as they are able to acquire sufficient information from the existing links, regardless of with how many agents they are connected and from which agents was the information acquired (i.e. the variety of agents). In contrast, in the IDG, the benefit of an individual agent is jointly determined by the amount of its own information, i.e. x_i, as well as the number of agents with whom it is connected, i.e. $|N_i(E_{\mathbf{g}})|$, while the information possessed by other agents has no influence on its link formation decision. The number and variety of agents that each agent is connected with thus form the most important factor that shapes its incentives.

It is important to note that in the IDG and the IAG, the amount of an agent's own information have opposite impacts on its incentive to form links with others: in the IAG, the more information an agent possesses, i.e. the larger x_i is, the smaller incentive it has to form links with other agents; whereas in the IDG, an agent with a larger x_i has a larger incentive to form links with others. As a result, the existing models used for IAG are not suitable to analyze the trade-off between the benefits and costs of information dissemination and link formation as well as the mutual impact between agents' strategic link formation decisions in an IDG. In fact, the existing IAG models can be applied to analyze the IDG only if agents are homogeneous with $x_i = x, \forall i$, so that the total amount of an agent's acquired information is proportional to the number of agents it connects with.

5.5.2 Differences in agents' equilibrium link formation behavior

A simple example show that these differences lead to highly different link formation behaviors thereby resulting in significant differences of the equilibrium topologies as opposed to those in the IAG. This point is further illustrated below using a simple example. For a fair comparison, we assume in the example that $f(y) = y^{\lambda}$ with $\lambda \in (0, 1)$. Suppose that there are n agents in the network. There is one agent i possessing an amount x_i of information with $x_i > \ln k/\lambda$ while all the other agents possess no information. Then in the IAG, it is easy to show that agent i forms no link at equilibrium, i.e. $g^*_{ij} = 0, \forall j \neq i$, and each agent $j \in N/\{i\}$ forms exactly one link with agent i, i.e. $g^*_{ji} = 1, \forall j \neq i$ and $g^*_{jj'} = 0, \forall j \neq i, j' \neq i$. Hence, the unique equilibrium in the IAG is a periphery-sponsored star. Nevertheless in the IDG, agent i forms at least one link with some other agent at equilibrium, i.e. $g^*_{ij} = 1, \exists j \neq i$, and each agent $j \in N/\{i\}$ forms no link with any other agent, i.e. $g^*_{jj'} = 0, \forall j \neq i, j' \in N$. Meanwhile, the larger x_i is, the more links that agent i forms in the IDG. We can show that when $x_i > \frac{1}{\lambda} \ln \frac{k}{(N-1)^{\lambda} - (N-2)^{\lambda}}$, agent i forms links with all other agents in the network, with the unique equilibrium being a center-sponsored star.

This example provides two important insights: (1) agents' link formation behaviors at equilibrium exhibit significant differences when they are playing the IAG or the IDG, even if they possess the same amount of information and incur the same link formation cost; (2) although the resulting equilibrium topologies in the IAG and IDG may exhibit some similarity with respect to their shapes (e.g. both IAG and IDG have the star topology as the unique equilibrium when $x_i > \frac{1}{\lambda} \ln \frac{k}{(N-1)^{\lambda} - (N-2)^{\lambda}}$), they may have completely different underlying structures which lead to different properties in practice, e.g. the center-sponsored star formed in the IDG has all its links supported by its center node and is more vulnerable to single-node failures than the periphery-sponsored star formed in the IAG.

5.5.3 Differences in emerging equilibrium topologies

The differences between the IDG and the IAG become even more distinct in the scenario where agents self-produce information. For instance, the "law of the few", which is shown as a robust feature at equilibrium in the IAG [69], no longer holds in the IDG as shown by Theorem 4.

5.6 Conclusion

In this chapter, we investigated the problem of information dissemination and link formation in strategic networks. We rigorously determined how the agents' desire to disseminate their own information throughout the network impacts their interactions and the emerging connectivity/topology among them. Our analysis proved several important properties of the strategic networks (arising from the agents' strategic link formation) at equilibria, such as "core-periphery", "minimally connected", "short diameter". These properties are important because they characterize the efficiency and robustness of the resulting equilibrium networks. We also studied the strategic information production by individual agents and its impact on the equilibria of the information dissemination games. Importantly, we showed that when agents benefit from information dissemination, the information production at equilibrium is no longer dominated by a small number of information producers and hence, the "law of the few" derived for traditional network formation games where agents benefit from information consumption no longer holds.

5.7 Appendices

5.7.1 Proof of Corollary 1

This can be proved using the same idea as Theorem 1. We first prove the following claim.

Claim 1. Given a strict NE $\mathbf{g}^*$ and when $k_{ij} = k, \forall i, j \in N$, if $g^*_{ij} = 1$ for some $i, j \in N$, then $\max\{g^*_{jj'}, g^*_{j'j}\} = 0$ for any $j' \neq i$ and $j' \neq j$.

Proof of Claim 1: Suppose, in contrast, $g^*_{ij} = 1$ and $\max\{g^*_{jj'}, g^*_{j'j}\} = 1$ for some $j' \neq i$ and $j' \neq j$. By deleting its link with j and forming a new link with j', i receives the same utility as what it receives in $\mathbf{g}^*$, which contradicts the fact that $\mathbf{g}^*$ is an (strict) equilibrium and hence this claim follows. ■

In the next step, we show that for each non-singleton component always has a star topology in a strict NE.

Without loss of generality, we select two agents $i, j \in C$ where C is a component in $E_{\mathbf{g}^*}$, such that $g^*_{ij} = 1$. According to Claim 1, we have that $\max\{g^*_{jj'}, g^*_{j'j}\} = 0$ for any $j' \in C$ and $j' \notin \{i, j\}$. According to Proposition 1, we should also have $g^*_{ji} = 0$, since agent j can strictly increase its utility otherwise by removing the link it forms to agent i.

Now suppose that $g^*_{j'i} = 1$ for some $j' \in C$ and $j' \notin \{i, j\}$. It is obvious that agent j' can switch its link from agent i to agent j without decreasing its utility, which gives a contradiction. Therefore, we can conclude that $g^*_{ij} = 1$, $\forall j \in C$ and $j \neq i$. Meanwhile, $g^*_{jj'} = 0$, $\forall j, j' \in C$ and $j, j' \neq i$. In other words, C has a star topology where agent i stays in the center and forms links with all other agents who stay in the periphery, while all the other agents do not form links mutually. This corollary thus follows. ■

5.7.2 Proof of Theorem 2

(i) When $f(x) < \underline{k}$, suppose that there is an equilibrium g* which contains a non-singleton component C. Let i and j be two agents in C such that $g^*_{ij} = 1$, it is obvious that agent i can strictly increase its utility by setting $g_{ij} = 0$. Hence, there is a contradiction and this statement follows.

(ii) When $f(x) \in (\underline{k}, \bar{k})$, consider a component C and one of its periphery agents i such that $\max\{g^*_{ij}, g^*_{ji}\} = 1, \exists j$ and $\max\{g^*_{ij'}, g^*_{j'i}\} = 0, \forall j' \neq j$.

Suppose $g^*_{ij} = 1$: If $k_{ij} = \bar{k}$ and $|C| > 2$, agent i can always switch its link to some other agent $j' \in C$ without decreasing its utility, If $k_{ij} = \bar{k}$ and $|C| = 2$, agent i can always increase its utility by switching its link to some other agent $j' \notin C$. Both cases contradict the fact that g* is a strict NE. Hence, we have $k_{ij} = \underline{k}$ and $g^*_{i'j} = 0, \forall i' \in C/\{i, j\}$ (otherwise i' can switch its link from j to i without decreasing its utility). Since i is a periphery agent, we have $g^*_{ji'} = 1, \exists i' \in C/\{i, j\}$. If $|C| = 3$, then $k_{ji'} = \underline{k}$ and agent i can switch its link from j to i' without decreasing its utility. Therefore, we have $|C| > 3$ and $\bar{g}^*_{i'i''} = 1, \exists i'' \in C/\{i, j, i'\}$. If $k_{ji'} = \bar{k}$, agent j can switch its link from i' to i'' without decreasing its utility, whereas if $k_{ji'} = \underline{k}$, agent i can switch its link from j to i' without decreasing its utility. Both cases contradict the fact that g* is a strict NE. It can be thus concluded that $g^*_{ij} = 1$ cannot hold in g* and we have $g^*_{ji} = 1$. As a result, j should belong to the same group as i with $k_{ij} = \underline{k}$.

Now consider another agent $j' \in C/\{i, j\}$. If $g^*_{j'j} = 1$, then j' can switch its link from j to i without decreasing its utility, which leads to a contradiction. Therefore, we have $g^*_{jj'} = 1, \forall j' \in C/\{j\}$ and the component forms a star topology. Also, if there is an agent $j'' \in C$ who is not from the same group as j, then j can always increase its utility by removing its link with j'' since $f(x) < \bar{k}$. Hence, this statement follows.

(iii) When $f(x) > \bar{k}$, it is still true that agents from the same group belong to the same component. Also, the network should be connected with a unique component existing under g*. It is always true that we can find two agents i and i' from one group

N_z such that $g^*_{ii'} = 1$. Using the same argument as that in statement (ii), it is easy to show that $g^*_{ii''} = 1, \forall i'' \in N_z/\{i\}$. Now consider an agent $j \notin N_z$. We have $g^*_{ji'} = 0, \forall i' \in N_z$ (otherwise the condition of a strict NE is violated). Now consider a path $path_{i'j} = ((i', j_1), (j_1, j_2), ..., (j_m, j))$ with $j_1, ..., j_m, j \notin N_z$ and $i' \in N_z$. Obviously, we have $g^*_{i'j_1} = 1$. Hence, i' can switch its link from j_1 to j without decreasing its utility, which again violates the fact that $\mathbf{g}^*$ is a strict NE. It can be thus concluded that for each $j \notin N_z$, $g^*_{i'j} = 1, \exists i' \in N_z$ and $g^*_{jj'} = 0, \forall j' \notin N_z$. Therefore, this statement follows. ■

CHAPTER 6

Conclusion

In this dissertation, I developed a rigorous and formal framework that integrates incentive mechanisms design with system resource management and is applicable to a wide range of socio-technical networks including online social networking platforms, online trading markets, social computing systems, and online user-generated content platforms. This framework encompasses various features that are common to a wide variety of socio-technical networks, including the anonymity and asymmetric interests of agents, the agents' strategic learning and adaptation, the agents' ability to make proactive link formation decisions, and the imperfect monitoring on the agents' behavior. A central point of my analysis, which departs from the existing literature on distributed optimization for large-scale systems, is the assumption that the agents are selfish and only interested in maximizing their individual utilities, instead of the collective social welfare of the network.

As the first part of the proposed framework, I established in Chapter 2 a rigorous framework for the design and analysis of a class of rating protocols to sustain cooperation in online communities. I rigorously formulated the platform designer's problem of designing an optimal rating protocol and derived conditions for sustainable rating protocols, under which no agent gains by deviating from the prescribed recommended strategy. Then, I continued to study in Chapter 3 the problem of designing rating protocols under which agents adapt their behaviors and learn over time. In such communities, the platform designer can determine the long-run evolution of the community of agents. With this knowledge, the goal of the platform designer is not only to design a rating protocol that is sustainable, but also to induces the community's convergence towards the (desirable) equilibrium in the long run. I explicitly showed that a simple two-label rating protocol is sufficient to induce full cooperation among agents and achieve the optimal performance in terms of the collective social welfare. Next, I turned to another important problem in socio-technical networks: the strategic information production and link formation of the self-interested agents. Chapter 4 and 5 provide complementary analyses for both the information acquisition and dissemination. I rigorously determined how the trade-off of producing information personally and forming links to acquire/disseminate information impacts the interactions between agents and thus the network topology that endogenously emerges. The key finding here is that the network topology that emerges (at equilibrium) necessarily displays a core-periphery type. Meanwhile, as the population becomes larger, the number of hub agents and the total amount of information produced grow in proportion to the total population. Importantly, this conclusion holds for both networks with information acquisition and information dissemination, which indicates that this "core-periphery" structure is a robust feature of endogenously formed socio-technical networks. I then showed that the networks that emerge at equilibrium are frequently minimally connected and have short network diameters. My conclusions had been conjectured for many networks, such as the "small-world" phenomenon in the World-Wide-Web, but not derived in any formal framework.

My framework can be extended in several directions, among which I mention three. First, the effectiveness of the incentive mechanisms proposed in this dissertation relies on the truthful reporting from agents. For instance, the update of a server's rating depends on the report of its client. In an extension of my framework, I will consider the scenario in which clients can use more complicated decision rules while reporting the servers' actions to the community operator in order to maximize their own long-term utility, instead of always reporting truthfully. Second, agents considered in this dissertation adopt simple beliefs about the system and the strategies of other agents. More sophisticated belief models can be introduced into my framework. Understanding how the evolutions of users' beliefs and users' strategies will impact each other form an important future research direction. Third, alternative formulations on information transmission and link formation can be extended from my current model to encompass the features of various applications. Several examples include the bilateral link formation which requires the mutual consent of agents, the unilateral information transmission with which the information flow over a link is one-sided, and non-constant link formation cost with which the cost of establishing a link depends on the characteristics of the creator and the recipient as well as the amount of information (i.e. traffic) across this link.

REFERENCES

[1] L. Adamic, J. Zhang, E. Bakshy, and M. Ackerman, “Knowledge Sharing and Yahoo Answers: Everyone Knows Something,” *Int’l Conf. on World Wide Web (WWW)*, 2008.

[2] W. Wang, V. Srinivasan, and K. Chua, “Using Mobile Relays to Prolong the Lifetime of Wireless Sensor Networks,” *Annual Int’l Conf. on Mobile Computing and Networking*, 2005.

[3] B. Zhao, J. Lui, and D. Chiu, “A Mathematical Framework for Analyzing Adaptive Incentive Protocols in P2P Networks,” *IEEE/ACM Trans. on Networking*, vol. 20, no. 2, pp. 367-380, 2012.

[4] K. Ranganathan, M. Pipeanu, A. Sarin, and I. Foster, “Incentive Mechanisms for Large Collaborative Resource Sharing,” *IEEE Int’l Symposium on Cluster Computing and the Grid*, 2004.

[5] M. Cha, H. Kwak, P. Rodriguez, Y. Ahn, and S. Moon, “I Tube, You Tube, Everybody Tubes: Analyzing the Worlds Largest User Generated Content Video System,” *ACM SIGCOMM*, 2007.

[6] A. Blanc, Y. Liu, and A. Vahdat, “Designing Incentives for Peer-to-Peer Routing,” *IEEE INFOCOM*, 2005.

[7] N. Miller, P. Resnick, and R. Zeckhauser, “Eliciting Honest Feedback: The Peer-Prediction Method,” *Management Science*, vol. 51, no. 9, pp. 1359-1373, 2005.

[8] N. Hanaki, A. Peterhansl, P. Dodds, and D. Watts, “Cooperation in Evolving Social Networks,” *Management Science*, vol. 53, no. 7, pp. 1036-1050, 2007.

[9] B. Awerbuch, B. Patt-Shamir, D. Peleg, and M. Tuttle, “Collaboration of Untrusting Peers with Changing Interests,” *ACM Conf. on Electronic Commerce*, 2004.

[10] S. Saroiu, P. K. Gummadi, and S. D. Gribble, “A Measurement Study of Peer-to-Peer File Sharing Systems,” *Multimedia Computing and Networking*, 2002.

[11] D. Lopez-Pintado, “The Spread of Free-Riding Behavior in a Social Network,” *Eastern Economic Journal*, vol. 34, pp. 464-479, 2007.

[12] D. Bergemann and D. Ozman, “Optimal Pricing with Recommender Systems,” *ACM Conf. on Electronic Commerce*, 2006.

[13] J. K. MacKie-Mason and H. R. Varian, “Pricing Congestible Network Resources,” *IEEE J. Sel. Areas Commun.*, vol. 13, no. 7, pp. 1141-1149, 1995.

[14] V. Vishnumurthy, S. Chandrakumar, and E. Gun Sirer, “KARMA: A Secure Economic Framework for P2P Resourcing Sharing,” *Workshop on Economics of P2P Systems*, 2003.

[15] J. Park and M. van der Schaar, “Medium Access Control Protocols with Memory,” *IEEE/ACM Trans. on Networking*, vol. 18, no. 6, pp. 1921-1934, 2010.

[16] R. Ma, S. Lee, J. Lui, and D. Yau, “Incentive and Service Differentiation in P2P Networks: a Game Theoretic Approach,” *IEEE/ACM Trans. on Networking*, vol. 14, no. 5, pp. 978-991, 2005.

[17] M. Feldman, K. Lai, I. Stoica, and J. Chuang, “Robust Incentive Techniques for Peer-to-Peer Networks,” *ACM Conf. on Electronic Commerce*, 2004.

[18] A. Habib and J. Chuang, “Service Differentiated Peer Selection: an Incentive Mechanism for Peer-to-Peer Media Streaming,” *IEEE Trans. on Multimedia*, vol. 8, no. 3, pp. 610-621, 2006.

[19] P. Resnick, R. Zeckhauser, E. Friedman, and K. Kuwabara, “Reputation System,” *Communications of the ACM*, vol. 43, no. 12, pp. 45-48, 2000.

[20] E. Friedman and P. Resnick, “The Social Cost of Cheap Pseudonyms,” *Journal of Economics and Management Strategy*, vol. 10, no. 2, pp. 173-199, 2000.

[21] P. Resnick and R. Zeckhauser, “Trust Among Strangers in Internet Transactions: Empirical Analysis of eBay’s Reputation System,” *Advances in Applied Microeconomics*, vol. 11, pp. 127-157, 2002.

[22] Y. Zhang and M. van der Schaar, “Strategic Learning and Robust Protocol Design for Online Communities with Selfish Users”, *UCLA Technical Report*, http://arxiv.org/abs/1108.5514, 2011.

[23] C. Dellarocas, “Reputation Mechanism Design in Online Trading Environments with Pure Moral Hazard,” *Information Systems Research*, vol. 16, no. 2, pp. 209-230, 2005.

[24] C. Dellarocas, “How Often Should Reputation Mechanisms Update a Trader’s Reputation Profile,” *Information Systems Research*, vol. 17, no. 3, pp. 271-285, 2006.

[25] R. Jurca, “Truthful Reputation Mechanisms for Online Systems,” *Int’l Conf. on Autonomous Agents and Multiagent Systems*, 2007.

[26] M. Kandori, “Social Norms and Community Enforcement,” *Rev. Economic Studies*, vol. 59, no. 1, pp. 63-80, 1992.

[27] M. Okuno-Fujiwara and A. Postlewaite, “Social Norms and Random Matching Games,” *Games Econ. Behavior*, vol. 9, no. 1, pp. 79-109, 1995.

[28] S. Takahashi, "Community Enforcement when Players Observe Partners' Past Play," *Journal of Economic Theory*, vol. 145, no. 1, pp. 42-62, 2010.

[29] G. Ellison, "Cooperation in the Prisoner's Dilemma with Anonymous Random Matching," *The Review of Economic Studies*, vol. 61, no. 3, pp. 567-588, 1994.

[30] J. Deb, "Cooperation and Community Responsibility: A Folk Theorem for Random Matching Games with Names," *memeo*, 2007.

[31] S. Adlakha, R. Johari, G. Y. Weinstraub, and A. Goldsmith, "Oblivious Equilibrium for Large-Scale Stochastic Games with Unbounded Costs," *IEEE Conf. on Decision and Control*, 2008.

[32] P. Johnson, D. Levine, and W. Pesendorfer, "Evolution and Information in a Gift-Giving Game", *Journal of Economic Theory*, vol. 100, no. 1, pp. 1-21, 2000.

[33] D. Kreps, "Decision Problems with Expected Utility Criteria, I: Upper and Lower Convergent Utility," *Mathematics of Operations Research*, vol. 2, no. 1, pp. 45-53, 1977.

[34] D. Kreps, "Decision Problems with Expected Utility Criteria, II: Stationary," *Mathematics of Operations Research*, vol. 2, no. 3, pp. 266-274, 1977.

[35] D. Fudenberg, D. Levin, and E. Maskin, "The Folk Theorem with Imperfect Public Information," *Econometrica*, vol. 62, no. 5, pp. 997-1039, 1994.

[36] G. J. Mailath and L. Samuelson, "Repeated Games and Reputations: Long-Run Relationships," *Oxford University Press*, 2006.

[37] L. Massoulie and M. Vojnovic, "Coupon Replication Systems," *IEEE/ACM Trans. on Networking*, vol. 16, no. 3, pp. 603-616, 2005.

[38] P. Whittle, "Optimization Over Time," *New York: Wiley*, 1983.

[39] E. Adar and B. A. Huberman, "Free Riding on Gnutella," *First Monday*, vol. 5, no. 10, 2000.

[40] M. Feldman, K. Lai, I. Stoica, and J. Chuang, "Robust Incentive Techniques for User-to-User Networks," *ACM Conf. on Elec. Commerce*, 2004.

[41] B. Cohen, "Incentives Building Robustness in BitTorrent," *P2P Econ. Workshop*, 2003.

[42] A. Kittur, E. Chi and B. Suh, "Crowdsourcing User Studies with Mechanical Turk," *Annual SIGCHI Conf. on Human Factors in Computing Systems*, 2008.

[43] H. Park and M. van der Schaar, "Evolution of Resource Reciprocation Strategies in P2P Networks," *IEEE Trans. on Signal Process.*, vol. 58, no. 3, pp. 1205-1218, 2009.

[44] M. Kandori, G. Mailath, and R. Rob, "Learning, Mutation, and Long-run Equilibria in Games," *Econometrica*, vol. 61, no. 1, pp. 29-56, 1993.

[45] Y. Zhang and M. van der Schaar, "Peer-to-Peer Multimedia Sharing based on Social Norms," *Elsevier J. on Signal Process.*, vol. 27, no. 5, pp. 383-400, 2012.

[46] Y. Zhang, J. Park, and M. van der Schaar, "Rating Protocols for Online Communities," *ACM Trans. on Economics and Computation*, to appear.

[47] Y. Su and M. van der Schaar, "Dynamic Conjectures in Random Access Networks Using Bio-inspired Learning," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 4, pp. 587-601, 2010.

[48] C. Camerer, T. Ho, and J. Chong, "Sophisticated EWA Learning and Strategic Teaching in Repeated Games," *Journal of Economic Theory*, vol. 104, pp. 137-188, 2002.

[49] J. Shamma and G. Arslan, "Dimensions of Cooperative Control," *Cooperative Control of Distributed Multiagent Systems*, 2003.

[50] D. Foster and P. Young, "Stochastic Evolutionary Game Dynamics," *Theo. Population Biology*, vol. 38, no. 2, pp. 219-232, 1990.

[51] D. Fudenberg and D. Levine, "The Theory of Learning in Games," *MIT Press*, 1998.

[52] D. Acemoglu and A. Ozdaglar, "Opinion Dynamics and Learning in Social Networks," *Dynamic Games and Applications*, vol. 1, no. 1, pp. 3-49, 2011.

[53] V. Krishnamurthy, "Quickest Time Herding and Detection for Optimal Social Learning," *arXiv:1003.4972 [cs.IT]*.

[54] J. Norris, "Markov Chains," *Cambridge University Press*, 1998.

[55] M. Freidlin and A. Wentzell, "Random Perturbations of Dynamical Systems," *New York: Springer Verlag*, 1984.

[56] Y. Zhang and M. van der Schaar, "Online Appendix: Strategic Learning and Robust Protocol Design, A Stochastic Stability Analysis," `http://medianetlab.ee.ucla.edu/papers/stochastic_learning_appendix`.

[57] J. Chakareski and P. Frossard, "Context-adaptive Information Flow Allocation and Media Delivery in Online Social Networks," *IEEE Journal of Sel. Topics in Signal Process.*, vol. 4, no. 4, 2010.

[58] M. Goemans, L. Li and M. Thottan, "Market Sharing Games Applied to Content Distribution in Ad-hoc Networks," *IEEE J. on Sel. Areas in Commun.*, vol. 24, no. 5, pp. 1020-1033, 2006.

[59] S. Ioannidis, A. Chaintreau and L. Massoulie, “Optimal and Scalable Distribution of Content Updates over a Mobile Social Network,” *Proc. IEEE INFOCOM 2009*, pp. 1422-1430, 2009.

[60] M. Gandetto and C. Regazzoni, “Spectrum Sensing: a Distributed Approach for Cognitive Terminals,” *IEEE J. on Sel. Areas in Commun.*, vol. 25, no. 3, pp. 546-557, 2007.

[61] W. Saad, Z. Han, A. Hjorungnes, D. Niyato and E. Hossain, “Coalition Formation Games for Distributed Cooperation among Roadside Units in Vehicular Networks,” *IEEE Journal on Sel. Area in Commun.*, vol. 29, no. 1, pp. 48-60, 2010.

[62] T. Shu and M. Krunz, “Coordinated Channel Access in Cognitive Radio Networks: a Multi-level Spectrum Opportunity Perspective,” *INFOCOM 2009*, pp. 2976-2980, 2009.

[63] S. Goyal, “Connections: an Introduction to the Economics of Networks,” *Princeton University Press*, New Jersey, 2007.

[64] V. Bala and S. Goyal, “A Non-cooperative Model of Network Formation,” *Econometrica*, vol. 68, no. 5, pp. 1181-1231, 2000.

[65] J. Corbo and D. Parkes, “The Price of Selfish Behavior in Bilateral Network Formation,” *Proc. 24th Annual Symposium on Principles of Distributed Computing (PODC)*, 2005.

[66] A. Galeotti and S. Goyal, “Network Formation with Heterogeneous Players,” *Games and Economic Behavior*, vol. 54, no. 2, pp. 353-372, 2006.

[67] J. Musachhio, G. Schwartz and J. Walrand, “A Two-sided Market Analysis of Provider Investment Incentives with an Application to the Net-neutrality Issue,” *Review of Network Economics*, vol. 8, no. 1, pp. 3-25, 2009.

[68] D. Watts and S. Strogatz, “Collective Dynamics of ’Small-world’ Networks,” *Nature*, vol. 393, no. 6684, pp. 440-442, 1998.

[69] A. Galeotti and S. Goyal, “The Law of the Few,” *The American Economic Review*, vol. 100, no. 4, pp. 1468-1492, 2010.

[70] A. Dixit and J. Stiglitz, “Monopolistic Competition and Optimum Product Diversity,” *The American Economic Review*, vol. 67, no. 3, pp. 297-308, 1977.

[71] A. Barabasi and R. Albert, “Emergence of Scaling in Random Networks,” *Science*, vol. 286, no. 5439, pp. 509-512, 1999.

[72] H. Kwak, C. Lee, H. Park and S. Moon, “What is Twitter, a Social Network or a News Media?” *Proc. 19th Int’l Conf. on World Wide Web (WWW)*, 2010.

[73] A. Fabrikant, A. Luthra, E. Maneva, C. Papadimitriou and S. Shenker, "On a Network Creation Game," *Proc. 22nd Annual Symposium on Principles of Distributed Computing (PODC)*, 2003.

[74] J. Nash, "Non-cooperative Games," *The Annals of Mathematics*, vol. 54, no. 2, pp. 286-295, 1951.

[75] Y. Zhang and M. van der Schaar, "Online Appendix: Information Production and Link Formation in Social Computing Systems," *Online Appendix*, `http://medianetlab.ee.ucla.edu/papers/link_formation_jsac_appendix.pdf`.

[76] Y. Zhang, J. Park and M. van der Schaar, "Content Production and Link Formation Games," *UCLA Technical Report*, `http://arxiv.org/abs/1109.4201, 2011`.

[77] T. Melodia, D. Pompili and I. Akyildiz, "On the Interdependence of Distributed Topology Control and Geographical Routing in Ad-hoc and Sensor Networks," *IEEE J. on Sel. Area in Commun.*, vol. 23, no. 3, pp. 520-532, 2005.

[78] M. Papadopouli and H. Schulzrinne, "Effects of Power Conservation, Wireless Coverage and Cooperation on Data Disseminataion among Mobile Devices," *MobiHoc*, pp. 117-127, 2001.

[79] E. Bakshy, I. Rosenn, C. Marlow and L. Adamic, "The Role of Social Networks in Information Diffusion," *WWW 2012*, pp. 519 -528, 2012.

[80] A. Blanc, Y. Liu and A. Vahdat, "Design Incentives for Peer-to-Peer Routing," *IEEE INFOCOM*, pp. 374-385, 2005.

[81] R. Bulander, M. Decker, G. Schiefer and B. Kolmel, "Comparison of Different Approaches for Mobile Advertising," *IEEE Int'l Workshop on Mobile Commerce and Services*, pp. 174-182, 2005.

[82] F. Provost, B. Dalessandro, R. Hook, X. Zhang and A. Murray, "Audience Selection for Online Brand Advertising: Privacy-friendly Social Network Targeting," *ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining (KDD)*, 2009.

[83] B. Chun, R. Fonseca, I. Stoica and J. Kubiatowicz, "Characterizing Selfishly Constructed Overlay Routing Networks," *INFOCOM*, 2004.

[84] M. Jackson and A. Wolinsky, "A Strategic Model of Social and Economic Networks," *J. of Economic Theory*, vol. 71, no. 3, pp. 44-74, 1996.

[85] L. Blume, D. Easley, J. Kleinberg, R. Kleinberg and E. Tardos, "Network Formation in the Presense of Contagious Risk," *ACM Conf. on Elec. Commerce*, pp. 1-10, 2011.

[86] J. Kleinberg, S. Suri, E. Tardos and T. Wexler, "Stratetic Network Formation with Structure Holes," *ACM SIGecom Exchanges*, vol. 7, no. 3, 2008.

[87] C. Gong, A. Livne, C. Brunneti and L. Adamic, "Coevolution of Network Strcture and Content," *WebSci*, 2011.

[88] M. Gomez-Rodriguez, J. Leskovec and A. Krause, "Inferring Networks of Diffusion and Influence," *Trans. on Knowledge Discovery from Data*, vol. 5, no. 4, 2012.

[89] G. Kossinets, J. Kleinberg and D. Watts, "The Structure of Information Pathways in a Social Communication Network," *ACM SIGKDD*, pp. 435-443, 2008.

[90] P. Reny, "On the Existence of Pure and Mixed Strategy Nash Equilibria in Discountinuous Games," *Econometrica*, vol. 67, no. 5, pp. 1029-1056, 1999.

[91] Y. Zhang and M. van der Schaar, "Online Appendix: Strategic Networks: Information Dissemination and Link Formation among Self-interested Agents," *Online Appendix*, `http://medianetlab.ee.ucla.edu/papers/information_dissemination_appendix`.

[92] A. Mislove, M. Marcon, K. Gummadi, P. Druschel and B. Bhattacharjee, "Measurement and Analysis of Online Social Networks," *SIGCOMM Conf. on Int'l Measurement*, pp. 29-42, 2007.

[93] Y. Suh, H. Shin and D. Kwon, "An Efficient Multicast Routing Protocol in Wireless Mobile Networks," *Wireless Networks*, vol. 7, no. 5, pp. 443-453, 2001.

[94] S. Currarini, M. Jackson and P. Pin, "An Economic Model of Friendship: Homophily, Minorities, and Segregation," *Econometrica*, vol. 77, no. 4, pp. 1003-1045, 2009.

[95] V. Colizza, A. Barrat, M. Barthelemy and A. Vespignani, "The Role of the Airline Transportation Network in the Prediction and Predictability of Global Epidemics," *Proc. of the National Academy of Sciences*, vol. 103, pp. 2015-2020, 2006.

[96] M. Chwe, "Communication and Coordination in Social Networks," *Review of Economic Studies*, vol. 67, pp. 1-16, 2000.

[97] E. Glaeser, B. Sacerdote and A. Scheinkman, "Crime and Social Interactions," *Quarterly Journal of Economics*, vol. 111, no. 2, pp. 507-548, 1996.

[98] Y. Ioannides and L. Datcher-Loury, "Job Information Networks, Neighborhood Effects and Inequality," *J. of Economic Literature*, vol. 424, pp. 1056-1093, 2004.

[99] M. Jackson and A. Watts, "On the Formation of Interaction Networks in Social Coordination Games," *Games and Economic Behavior*, vol. 41, pp. 265-291, 2002.

[100] M. Jackson and L. Yariv, “Diffusion on Social Networks,” *Economie Publique*, vol. 16, no. 1, pp. 69-82, 2005.

[101] D. Medhi and K. Ramasamy, “Network Routing: Algorithms, Protocols, and Architectures,” *Morgan Kaufmann*, 2007.

[102] S. Baset and H. Schulzrinne, “An Analysis of the Skype Peer-to-Peer Internet Telephony Protocol,” *arXiv preprint cs/041207*, 2004.