University of Alberta

BEHAVIOR MODELING AND ANALYSIS IN MULTIMEDIA SHARING NETWORKS

by

Bo Hu

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Digital Signals and Image Processing

Department of Electrical and Computer Engineering

Spring 2013
Edmonton, Alberta

Dedication

To my family.

Abstract

In multimedia sharing networks such as YouTube, and Flicker, etc, users actively participate and interact with each other, which influences not only each individual but also the entire system performance. Successful deployments of multimedia sharing networks show that user cooperation helps provide efficient and highly scalable platforms for multimedia distribution. However, since users are selfish, their cooperation cannot be guaranteed. In this thesis, we aim to design incentive mechanisms to stimulate user cooperation and also optimize the system performance.

Without loss of generality, we use two multimedia applications as examples to show how to achieve our research goals. We first study a two-hop cooperative wireless multicast system, where after the base station broadcasts a packet, a relay node who receives the packet correctly helps forward it to the others. We model user interaction in this system as a multi-seller multi-buyer payment based game, where users pay to receive relay service and get paid if they help forward a packet.

We then study an interactive multiview video streaming (IMVS) system, where an user can select one out of many available views for observation and switch views frequently. With the advances of multiview video coding techniques, users can cooperatively download videos even if they are watching different views. We then model user interaction as an indirect reciprocity game and formulate users' decision making associated with their view switching as a Markov decision process.

In these two examples, our analysis shows that user behavior impacts the system performance significantly. Thus, we optimize our incentive mechanisms, which drive the games to desired stable equilibria, where users cooperate with each other and the system performance is maximized at the same time.

Acknowledgments

First and foremost, I would like to take this opportunity to express my sincere gratitude to my advisor, Prof. H. Vicky Zhao, for providing me this valuable opportunity to work with her. This thesis would not be possible without her patient guidance, constant encouragement and continuous support. Her vision, enthusiasm and desire for excellence not only have influenced my current research, but also set up a role model for my professional and personal development with lifetime.

I would like to thank Prof. Hai Jiang and Prof. Gene Cheung not only for their inspiring suggestions on my research, but also for the sharing of their research and career development experience. I would also like to thank Prof. Hai Jiang, Prof. Dileepan Joseph, Prof. Yindi Jing and Prof. Jie Liang for agreeing to serve in my thesis committee, and their time and effort spent on reviewing my thesis.

Thanks should also go to all the colleagues and friends during my Ph.D study, for their friendship, help and inspirations. I feel lucky to be their friends, and am also deeply grateful for the things I have learned from them. Special thanks go to Xingyu Li, Zhaorui Liu, Binglai Niu, Gencheng Guo, Gongpu Wang, Rongfei Fan, Zhi Liu, Yunlong Feng, Lihua Zhao, Wei Cai and Lei Zhong. The happy moments with them will be always in my memory.

Finally and most importantly, I give my heartfelt gratitude to my parents. Without their love, unconditional support and countless sacrifices, I could never accomplish so much and reach this milestone in my life. I dedicate this thesis to them.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

Acronyms	Definition
AF	amplify-and-forward
BS	base station
CCS	credit clearance service
CDMA	code-division multiple access
CRC	cyclic redundancy check
DF	decode-and-forward
DSC	distributed source coding
ESS	evolutionarily stable strategy
GOP	group of pictures
IMVS	interactive multiview video streaming
MDP	Markov decision process
P2P	peer-to-peer
PfC	pay-for-cooperation
RDSTC	randomized distributed space time coding
SNR	signal-to-noise ratio
SPNE	subgame perfect Nash Equilibrium
TDMA	time-division multiple access
WLAN	wireless local area network
WWAN	wireless wide area network

List of Symbols

Symbol[1]	Definition
$\mathcal{A}$	the action space
$\mathcal{R}$	the reputation space
$\bar{\mathcal{R}}$	the reputation set with reputations no less than $t_r - 1$
$\mathcal{S}$	the state space
$\mathcal{V}$	the view space
$\underline{\mathcal{V}}$	the low utility view set
$\bar{\mathcal{V}}$	the high utility view set
$\mathbf{Q}$	the reputation updating matrix
$\mathbf{T}$	the view transition matrix
$\mathbf{v}$	the steady state view distribution
$\mathbf{x}$	the reputation distribution
c	a user's cost to help others (in the homogeneous user case, where users have the same cost)
c_i	user i's cost to help others (in the heterogeneous user case, where users have different costs)
$[c_l, c_h]$	the range of users' cost in the heterogeneous user case, where users have different cost
$d^r(N_s)$	the reserve bid of the auction game with N_s bidders

[1] All caligraphic capital letters denote spaces or sets, all bold capital letters denote matrices, and all bold small letters denote vectors.

d^w	the payment that the winner of the auction gets paid if he/she relays a segment
d_i	the bid sent by user i
g	the gain of receiving a segment correctly in the wireless multicast system
g_v	The expected short-term gain if the view switching starts from view v and helpers always help
p_1	the probability of correctly receiving a segment from the BS
p_s	the probability that a user who does not participate in the auction game is a successful user
q^*	the optimal price that maximizes the relay portion throughput
q	the price selected by the local agent
r	a user's reputation level
t_r	the reputation threshold to differentiate beneficial users from non-beneficial users
$\breve{x}_f, \breve{x}_h$	the root of $f(x) = 0$, $h(x) = 0$
$\tilde{x}_f, \tilde{x}_h$	the root of $f'(x) = 0$, $h'(x) = 0$
x^*	evolutionarily stable strategy
y	the percentage of beneficial users
y_{in}	the percentage of selected users in the PfC scheme
$G_{r,v}$	the expected short-term gain from others' help with reputation r and view switching starting from view v
L	the average interval between two consecutive requests received by a user
M	the number of views

N	the number of users
N_s, N_b	the number of sellers, buyers
N_{su}	the number of successful users after the broadcast from the BS
P_a	the probability of switching to anchor views
$P^a_{r\to r'}$	the reputation transition probability from r to r' by action a
R	the highest reputation level
T_R	the relay portion throughput
$U^a_{r,v}$	The expected short-term utility by taking action a at reputation r and view v
$\bar{U}_B, \bar{V}_B$	the average utility to be a buyer in homogeneous, heterogeneous user case
$\bar{U}_{NB}, \bar{V}_{NB}$	the average utility not to be a buyer in homogeneous, heterogeneous user case
$W^{\pi}_{r,v}$	The lifetime utility at reputation r and view v by action policy π
$W^{a',\pi}_{r,v}$	The lifetime utility of the one-shot deviation to a' at reputation r and view v
η	the discounting factor quantifying how much users care about the future utility
γ	the discounting factor after receiving $(t_r - 1)$ requests
π	the action policy
$\phi_c(c)$	the probability density function of cost in the case that users have different cost

Chapter 1

Introduction

1.1 Motivation

In the past ten years, we witness the emergence of large-scale multimedia sharing networks. For example, users can upload and browse images and videos on Flickr [1] and YouTube [2], download music from Napster, and watch online videos through peer-to-peer (P2P) live streaming softwares such as PPLive [3], PPstream [4], and Sopcast [5], etc. In these networks, millions of users all over the world participate to create and share multimedia data with each other, which produces massive multimedia data for distribution. From the study in [1], in Feb. and Mar. 2012, the daily upload volume of Flickr has reached 1.8 million photos per day. In addition, with the emergence of high-speed cellular and WiFi networks and the increasing popularity of advanced mobile devices such as smart phones and tablets, users can easily create, share and browse multimedia data anywhere and anytime. From [2], the traffic downloaded to or uploaded from mobile devices was tripled at YouTube in 2011, and currently, more than 20% browse requests are sent from mobile devices. The massive production and frequent exchange of multimedia data pose great challenges to multimedia distribution over wired and wireless networks. Thus, technologies that can support efficient and reliable exchange of multimedia date are in demand.

Different from traditional multimedia systems, in multimedia sharing networks, users do not *passively* receive provided service, but *actively* participate in and contribute to the systems. A lot of studies (e.g. [6]–[10]) have shown that users play a key role in multimedia sharing networks. For example, in P2P file sharing systems, when downloading files, users will simultaneously upload the downloaded data to others. Thus, their average downloading time can be effectively shorten when compared with the traditional client-server based model. In addition, a higher level of user participation (e.g., more upload bandwidth they contribute), the system can provide more efficient service. Furthermore, users can learn from others or their own past experience in multimedia sharing networks. For example, in YouTube users can read the comments of a video wrote by others, and then decide whether to watch it. If users find most of videos published by a user having good ratings, they may even subscribe to this user for all future videos. However, people may also manipulate the recommendation service provided by multimedia sharing networks to promote their own content for profits. As shown in [10], software programs were developed for scammers, which can mimic legitimate YouTube traffic and provide positive feedbacks for any video they want to promote. All these examples show that users in multimedia sharing networks interact with each others, which significantly influences not only each individual's decision but also the entire system performance. Thus, for a better design of multimedia sharing networks, we need to take human factor into consideration, understand how users learn from and influence each other, and analyze how such user behavior dynamics affect the system performance. The ultimate goal of such investigation is to provide important guidelines for designing a multimedia sharing network with satisfactory, efficient and personalized service.

From the above discussion, user behavior dynamics introduce different issues to be addressed in multimedia sharing networks, such as incentive mechanisms for higher level of user participation, social learning, security and privacy issues, etc. In this thesis, we focus on designing incentive mechanisms to stimulate user cooperation. The studies in [6]–[8] show that user cooperation can help provide efficient and highly scalable platforms for multimedia exchange and distribution. For example, the work in [8] provides a measurement study of a real P2P live streaming software, PPLive, which provides live broadcast of hundreds of Chinese TV channels. On May 12th, 2010, one of these channels, HunanTV, was broadcasted via PPLive to over 1600 users over the Internet at a bit rate of 400 kbps, corresponding to an aggregated bit rate of more than 600 megabits per second. In this example, peers cooperatively download/upload video packets from/to each other so that everyone can receive a high quality video. Thus, cooperation allows users to access available resources in the entire network, and therefore, the system can achieve much higher throughput than the traditional client-server based model.

In multimedia sharing networks, users receive gains from accessing others' network resources. However, sharing their own network resources to help others' downloading may incur some cost. Since users are intelligent and rational, they have the ability to choose optimal actions towards maximizing their utilities,[1] and the optimal actions may not always be playing cooperatively. A study [12] on a P2P file sharing system, Gnutella, shows that 25% of users are free riders, who only download from other peers but do not share any file at all. This is because users are *selfish* [13], and only care about their own utilities. If free riding can result in a higher utility, they will tend to free ride rather than cooperate.

To address this problem, we need to study user interaction and design incentive mechanisms to stimulate user cooperation. Game theory [14], [15] provides the fundamental tools to model user interaction, and study their strategic decision making. In particular, we are interested in the Nash Equilibrium of a game, from which no user has incentive to deviate. When designing incentive mechanisms for multimedia sharing networks, there are several challenging issues that need to be addressed.

i) In the current state-of-arts, incentive mechanisms for multimedia sharing networks mainly study the point-to-point interaction, where a pair of users establish

[1] In practical multimedia sharing systems, users use softwares such as PPlive and Sopcost provided by the systems for video downloading and uploading. Most of such softwares force user cooperation in the networks. However, the work in [11] assumes that intelligent users can manipulate the software and develop their own protocols to interact with other users.

partnership and cooperate with each other. In such systems, a user can choose different strategies towards different partners. Thus, a rational user will only cooperate with cooperative users, and free-riders can be easily isolated. However, in reality, there is one-to-many interaction, especially in wireless communication. If one user transmits a message, due to the broadcast nature of wireless communication, all nearby users can hear it. In this case, a user cannot select different strategies towards different users, since even if he/she only wants to cooperate with one user, other users can overhear and free ride. In this scenario, free riding is much easier, and cooperation stimulation is a more challenging problem.

ii) In the literature, many incentive mechanisms consider the scenario where users face the same game every time they interact with each other. This is because they assume that user states, such as available network resources, do not change over time. However, this assumption may not be true. For example, in a mobile multimedia sharing system, mobile devices all have energy constraints. Since their remaining energy will change over time, their strategies may also change. Thus, understanding how user state change affects their strategies and the system performance is important in designing incentive mechanisms to stimulate user cooperation.

iii) In multimedia sharing networks, users many join and leave the system from time to time. User membership dynamics may also impact their strategies, and thus, the system performance. For example, given that a group of new users just join the network, and the existing users do not know how those new users will behave in the system, whether the existing users should continue their cooperation is a problem that should be addressed.

iv) Each user may have private information, which other users do not known, e.g., his/her own available network resources. Selfish users may lie about their private information if cheating can improve their performance. For example, a user may claim that he/she has very low upload bandwidth, so that to avoid contributing to the system. Thus, we should address users' cheating behavior to improve the system efficiency.

In this thesis, we focus on these challenging issues in designing incentive mechanisms for multimedia sharing networks.

1.2 Thesis Contribution and Outlines

From the above discussion, we know that user behavior has significant impacts on multimedia sharing networks, and also know the four important issues that should be addressed in designing incentive mechanisms. Without loss of generality, in this thesis, we study two multimedia applications, which are typical applications raising those four challenging issues. The first application is a two-hop cooperative wireless multicast system, where all users receive the same multimedia data for observing, and it raises issue **i)** and **iv)** to be addressed. The second application is an interactive multiview video streaming (IMVS) system, where users receive not exactly the same but correlated multimedia data for observing, and it raises issue **ii)** and **iii)** to be addressed. We use these two applications as examples to show how user behavior impacts the system performance, and how to design incentive mechanisms to address the four important issues. Our contributions are summarized as follows:

1.2.1 Incentive Analysis for Two-Hop Cooperative Wireless Multicast

In two-hop cooperative wireless multicast, all users receive the same multimedia data. After the base station (BS) broadcasts a video packet, successful users who receive the packet correctly can help forward the packet to the rest unsuccessful users. Due to the broadcast nature of wireless communication, when a user relays a packet, he/she actually helps all nearby users who can hear him/her. Thus, it is a typical one-to-many interactoin model in multimedia sharing networks.

In Chapter 3, we study the incentive mechanism to stimulate user cooperation in this system. We first model the interaction among users in this system as a multi-seller multi-buyer payment based game, where users pay to receive relay service and they will get paid if they help forward packets. We also derive the optimal price, which drives the game to the desired Nash Equilibrium, where unsuccessful users have low free-riding probability and the system throughput is maximized.

We then consider that users may have different cost to upload one packet, which is their private information. To address their cheating behavior, we design a second-price sealed-bid auction game. In this game, bidding the true cost is their dominant strategy, which everyone will choose.

1.2.2 Incentive Analysis for Cooperative Interactive Multiview Video Streaming

In recent years, Free Viewpoint Video [16] becomes popular, where the same scene is captured by a large array of cameras (e.g. more than 100 cameras in [17]) from different viewpoints, and an audience can interactively select one interested view for observing. In such systems, users are likely to watch different views. Thus, they do not receive exactly the same but correlated multimedia data. With the frame structure proposed in [18], users can cooperatively download packets even if they are watching different views. In this system, due to different popularity of views, we observe that users watching different views may receive different utilities from others' help, and thus, may take different actions accordingly. Since users switch views frequently, their actions may also change frequently. Thus, it is a typical example of user interaction with state change.

In Chapter 4, we model users' state transition and decision making as a Markov decision process (MDP) in cooperative IMVS. From the game analysis, we observe that users may cooperate at some views but not others. Furthermore, we observe that the game may have multiple Nash Equilibria corresponding to different cooperation levels, (e.g., in the full cooperation Nash Equilibrium, users cooperate at all views, while in the partial cooperation Nash Equilibrium, users only cooperate at certain views.). We then propose a Pay-for-Cooperation (PfC) scheme to drive the game to the desired full cooperation Nash Equilibrium to improve the system efficiency.

We then investigate the impact of user membership dynamics on user cooperation, and observe that as long as the percentage of new users is below a predetermined threshold, cooperation is still a dominant strategy. Otherwise, cooperation will be interrupted, and the PfC should be used to resume user cooperation.

1.2.3 Thesis Outline

The thesis is organized as follows. In Chapter 2, we first introduce the prior arts on cooperative wireless multicast, IMVS and related works. We then introduce the fundamental concepts of game theory. In Chapter 3, we propose an incentive mechanism for two-hop cooperative wireless multicast to stimulate one-to-many cooperation, and encourage users to tell their true cost. In Chapter 4, we introduce the game model for cooperative IMVS and show how to stimulate user cooperation with state change, and then study how user membership dynamics affect user cooperation. Conclusions and potential future research topics are drawn in Chapter 5.

Chapter 2

Background and Literature Survey

In this chapter, we will first introduce the background of cooperative wireless multicast, IMVS and related works. We then introduce the fundamental concepts of game theory.

2.1 Cooperative Communication in Wireless Networks

In wireless communication, the fading effect of wireless channels is a feature impediment that limits the channel capacity and transmission range. The recent advances in 3G/4G networks do have significantly improved the channel capacity, while it is still not sufficient for the even faster increase of consumers' demand (e.g., high bandwidth for the transmission of high quality videos). A lot of works have been proposed to address this issue, and cooperative wireless communication emerges as a promising approach. The work in [19] first proposes two-hop cooperative wireless communication with relay, where any pair of users have a direct link. The transmission of one packet takes two phases. In the first phase, the source node transmits to the destination node, and both the destination node and the relay node can hear the signal. In the second phase, the relay node forwards the received signal to the destination node. The destination node then combines the two signals for decoding. This scheme explores the spatial diversity to effectively improve the channel capacity. Multi-hop cooperative wireless communication is then proposed mainly for mobile ad hoc networks [20], [21], where the source node and the destination node do not necessarily have a direct link. The source node relies on intermediate relay nodes to form a multi-hop route to the destination for packet transmission. Thus, this scheme can effectively extend the transmission range.

In the literature, based on the number of intended recipients, cooptative wireless communication schemes can be divided into two categories: unicast and multicast. In unicast, a source node has a single destination node, and different source-destination pairs transmit different messages. In multicast, a source node has multiple destination nodes who receive the same messages. In the following, we summarize the previous works on cooperative wireless communication for both unicast and multicast .

2.1.1 Cooperative Wireless Unicast

In this section, we discuss both two-hop and multi-hop cooperative wireless unicast.

2.1.1.1 Two-Hop Cooperative Wireless Unicast

In the literature, two-hop cooperative unicast schemes all follow the framework in [19] as discussed earlier. The work in [22] studies the lower bound of the channel capacity with relay. The work in [23] proposes two cooperative protocols: amplify-and-forward (AF) and decode-and-forward (DF). With the AF strategy, the relay simply amplifies the received signal from the source node, and then forwards to the destination node. In the DF strategy, the relay decodes the received signal, re-encodes it and then forwards to the destination node.

The works in [24], [25] study the optimal relay selection for a source-destination pair among a group of potential relay nodes at different locations. The works in [26]–[29] study the scenario, where a single relay node serves multiple source-destination pairs of nodes. For example, the work in [26] studies the optimal resource allocation at the relay node to maximize the network capacity, which is the summation of all source-destination pairs' channel capacities.

2.1.1.2 Multi-Hop Cooperative Wireless Unicast

Multi-hop cooperative wireless unicast is mainly used in mobile ad hoc networks, which have many design issues to be addressed, such as routing, security and energy management, etc.

In ad hoc networks, nodes may move, which will change the network topologies. The works in [20], [21], [30], [31] focus on designing efficient routing protocols to adapt to dynamic network topologies. For example, the work in [31] proposes a routing on demand algorithm, where if a source transmits to a destination without existing route, the source will broadcast a route request to its neighbors. Those neighbor nodes will also forward this message until the destination is reached or a route is found to the destination.

Ad hoc networks are vulnerable to attacks. For example, an attacker may create a *black hole* [32], which attracts packets by transmitting faked routing information, and always drops packets without forwarding. The works in [33]–[35] are proposed to address security issues. Energy management is also a challenging issue due to the energy constraints at mobile nodes. The work in [36] proposes a balanced energy consumption scheme so that the network can maintain a reasonable lifetime for a certain task.

2.1.2 Cooperative Wireless Multicast

Due to the broadcast nature of wireless communication, wireless multicast is a very efficient way for media distribution to a group of users who want the same data. In this section, we review previous works for both two-hop and multi-hop cooperative wireless multicast.

2.1.2.1 Two-Hop Cooperative Wireless Multicast

Two-hop cooperative wireless multicast [37]–[44] also takes two phase for one packet transmission. In the first phase, the source node broadcasts the message. In the second phase, successful users who receive the packet correctly will forward it to the rest unsuccessful users.

The work in [37] proposes a time-division multiple access (TDMA) based two-hop cooperative wireless multicast system. In the second phase, the BS randomly selects several successful users to serve as relays, and they take turns to forward the packet to the rest users. To maximize the system throughput, the optimal time allocation between the two phases and the optimal number of relays are derived. Their simulation results show that the cooperative multicast scheme can significantly improve the system performance.

The work in [40] investigates cooperative wireless multicast with both distributed and genie-aided cooperation schemes. In the distributed model, all successful users serve as relays, and forward the packet simultaneously in the second phase. In the genie-aided model, only a fixed number of users at predetermined locations can be relays, and they will forward the packet if they receive it correctly. Given a total power constraint in the two stages, this work then derives the optimal power allocation between the BS broadcast and the relay forwarding to minimize the average outage probability, which is the probability that the received signal-to-noise ratio (SNR) is below a predetermined threshold at the users' side.

A similar work is proposed in [38], where the randomized distributed space time coding (RDSTC) technique is employed. In the second phase, successful users encode the received packet using RDSTC and forward the encoded packets simultaneously. This work also uses scalable video coding to provide differential service. In scalable video coding, the video is coded into the base layer and enhancement layers. With the base layer only, the video can be decoded with the lowest video quality. Enhancement layers can refine the base layer and improve the video quality. In this scheme, the cooperative multicast is used to transmit the base layer to ensure that all users can receive the base layer and reconstruct the video. The enhancement layers are broadcasted by the BS without cooperative multicast, and only users with good channel conditions can receive it correctly to improve the video quality.

Another work in [39] exploits the network coding technique. In this work, the packets are divided into groups. In the first phase, the BS broadcasts a group of packets to all users. In the second phase, each user encodes the correctly received packets into one repair packet using network coding, and then forwards this repair packet to others. For a user with packet loss in the first phase, a repair packet is novel to him/her if this repair packet is encoded by packets he/she is missing. As long as he/she can receive enough novel repair packets (more than the number of packets he/she is missing) from different users, he/she can decode all the missing packets in the packet group.

2.1.2.2 Multi-Hop Cooperative Wireless Multicast

The multi-hop cooperative wireless multicast is also mainly studied in ad hoc networks, where the challenge is how to find efficient and robust multicast route to reduce transmission redundance with dynamic topologies.

The works in [45], [46] propose tree-based routing protocal, where a multicast tree is constructed to deliver packets to each destination. In [45], a Shared Tree Ad-hoc Multicast Protocol is proposed, where the tree is rooted at the *core*, which is a special node to manage the tree structure. The core-rooted tree is shared by all source nodes, who transmit multicast packets to the core using the shortest path, and the core forwards packets along its tree. Though this structure is simple and efficient, it may suffer from topology dynamics, since after nodes move, multicast packets have to be dropped before the tree is reconstructed.

Mesh-based multicast protocols are proposed in [47], [48]. For example, the work in [47] proposes an On Demand Multicast Routing Protocol, where the source node initiates the path searching process by sending an initiation packet. The intermediate nodes help forward this message and find the route reaching all destination nodes. This protocol can cope well with network topology dynamics, while it may cause high control overhead.

There are also works [49]–[51] that take energy management into consideration. For example, the work in [49] proposes an energy-efficient multicast tree, which balances the energy consumption over the network to maintain a long network lifetime.

2.2 Cooperative Video Streaming

Cooperative video streaming has proven to be successful in the past ten years, where the system can access the network resources of participating users, and provide video streaming service with satisfactory quality to thousands of users simultaneously. Typical cooperative video streaming applications include P2P live streaming and P2P video on demand. In P2P live streaming, all users watch the live video with similar playback time, and they help each other upload/download packets. In P2P video on demand, users may join the video at different time instances. Since they are all likely to watch the video from the beginning, they usually have different playback time. In this case, only users who join earlier can share their downloaded video and help users who join late. Thus, the ways user cooperate in these two applications are different. In this section, we will focus on P2P live streaming, since it is more related to IMVS we study in this thesis. We then introduce IMVS, and its recent advances in video coding techniques that support cooperative multiview video distribution.

2.2.1 Single View Video P2P Live Streaming

In the literature, there are three types of P2P live streaming structures: tree-push based, mesh-pull based and pull-and-push hybrid structures.

2.2.1.1 Tree-Push Based P2P Live Streaming

Tree-Push [52] based structure is proposed in the early stage of P2P live streaming development. In this structure, the streaming server is the root of the tree, and all users are organized at different layers in this tree structure. Each node (including the root node) may have more than one child nodes depending on his/her upload bandwidth. The server then pushes the video stream from the root, and each intermediate node forwards the stream to his/her child nodes. Thus, the video stream is transmitted in a top-down manner.

In this system, once the tree is built, each video packet is forwarded by an intermediate node once he/she receives it. Thus, the playback latency (the playback time difference between the server and the bottom nodes) is low. However, as studied in [53], this structure has two major drawbacks, which significantly limit its application. First, this structure cannot handle user membership dynamics well. Once an intermediate node leaves the tree, all his/her child nodes have to find new parents. [53] shows that with frequent user churn, the tree structure has to be changed frequently and the system performance is reduced. In addition, the leaf nodes do not connect to any child node, and thus, their upload bandwidth cannot be utilized. However, the leaf nodes take a huge portion of the entire network.

2.2.1.2 Mesh-Pull Based P2P Live Streaming

In mesh-pull P2P live streaming [6], a compressed video stream is divided into small data chunks, all of which are available at the streaming server. When a peer joins the system, he/she fetches from the streaming server an initial list of peers, who are currently watching the video. Then, he/she can communicate with peers in the list and obtain additional peer lists in a gossip manner. Each user maintains a buffer (called "streaming buffer") to store received data chunks that have not been decoded. Each user also keeps a buffer map to record the indices of the received chunks. Users periodically exchange buffer map information with each other, so that they know who have which chunks. Then, each user can select missing chunks to request either from the server or from other peers who have those chunks. When a user receives chunk requests, he/she can either accept them and upload the requested chunks or reject the requests. In this way, users can cooperate with each other to spread the video content and everyone can enjoy the video at the same time. Note that when a user joins the network, he/she has to buffer enough continuous data chunks before launching the video player for rendering. Then, he/she periodically moves the received chunks in the streaming buffer with the earliest playback time to the video player for rendering.

In this system, users have the freedom to select any nodes in their lists to establish partnership and exchange video chunks. Users' arbitrary connections make the system form a mesh-like network structure, which can handle user churn and large scale P2P networks very well. Currently, most of the successful deployments, such as PPlive, and PPStream, etc, are based on this structure. However, this structure also has some drawbacks. First, since each user needs to send a request for each single chunk, this may cause high signaling traffic overhead. In addition, there is high playback latency between the streaming server and the nodes that are far away from the streaming server, due to the hop-to-hop accumulated latency.

2.2.1.3 Hybrid Pull-and-Push Based P2P Live Streaming

From the above discussion, the tree-push based and the mesh-pull based schemes have different advantages and drawbacks. The hybrid pull-and-push based P2P live streaming proposed in [54] combines the two schemes to achieve robustness to peer churn, low playback latency and low signaling traffic overhead.

The network structure is also mesh based to handle peer churn. The video chunks are divided into transmission groups with equal number of chunks per group. The first chunk in each group is called a pull chunk, while the rest are push chunks. When a user i requests a transmission group, he/she asks one of his/her neighbors, user j, who has the pull chunk of that group in the buffer. As long as user j agrees with user i's request and sends the pull chunk, user j also forwards every push chunk in this transmission group to user i, once user j receives that chunk. If there is packet loss due to network congestion, the lost chunks will be requested in a pull-based manner. Thus, in this system, most of chunks are forwarded without being requested, which results in low signaling traffic overhead and also helps reduce the playback time latency.

2.2.2 Cooperative IMVS

Free viewpoint video [56] is becoming popular in recent years, and it provides IMVS service, where an audience can select one viewpoint of the video to watch

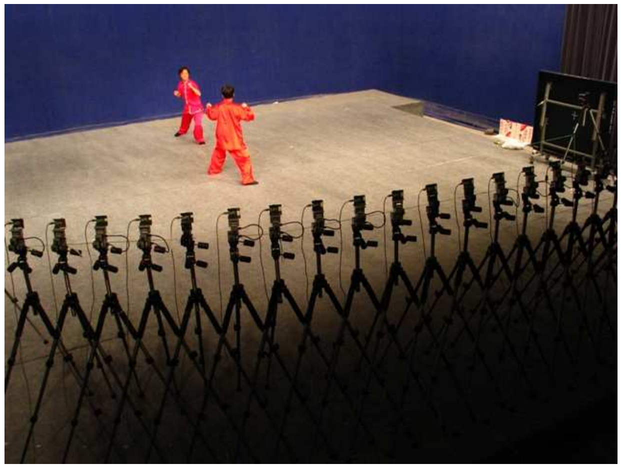


Fig. 2.1. A scene is captured by a large array of closely spaced cameras. This figure is from [55].

and switch views interactively and frequently. Thus, audiences have a 3D visual experience known as *motion parallax* [57]. To capture a 3D video, the traditional 3D modeling [58] uses 3D scanners to scan and process the surface of a 3D subject. It is very time consuming and computation intensive, which makes it difficult to achieve the real time video capture, compression and transmission. Alternatively, in most of the prototypes of free viewpoint video systems, a large array of closely spaced cameras are used to capture a scene from different angles, as shown in Fig. 2.1. For example, in [17], an array of more than 100 cameras are used. All the captured videos are collected by a server for compression and streaming, and a client can periodically select one out of many views for observation. In response, the server sends only pre-encoded data for the single requested view (rather than all the captured views) to reduce the streaming rate.

2.2.2.1 Advances in IMVS Using Distributed Source Coding

In IMVS, a straightforward video coding scheme is to encode and transmit each view independently. However, it will cause a large view switching delay, i.e., a user has to wait for a long time to switch to a different view. This is because video frames are divided into groups of pictures (GOP) [59], and each GOP is a video compression unit. In each GOP, the first frame (called I-frame) is encoded independently and can be directly reconstructed once being received. Each remaining frame is encoded with differential coding schemes [59] to achieve higher compression efficiency, and can be reconstructed only when it prior frame is decoded correctly. Thus, a user has to wait till the end of a GOP to switch to another view, and the average waiting time is the length of one GOP. Since a large GOP structure is usually required to achieve high video coding efficiency, users suffer from large view switching delay.

To address the tradeoff between view switching delay and coding efficiency, the works in [60]–[62] design frame structures using distributed source coding (DSC) [63], [64] for IMVS to achieve low bit rate video with low view switching delay. DSC states that several correlated information sources can be separately encoded at the encoder side and jointly decoded at the decoder side, while it can still achieve similar coding efficiency to the joint source encoding. To show how the DSC works, we consider a simple example with two correlated source symbols X and Y to be encoded, where each of them has 7 bits, and the Hamming distance between them is at most 1. At the encoder side, we encode them separately. First, Y can be transmitted to the decoder using 7 bits without compression. We then encode X. Since Y is available at the decoder side and we know that the Hamming distance between X and Y is at most 1, then, Y can be considered as a noisy version of X and a few parity bits is enough to retrieve X from Y. For example, if (7,4,3) Hamming code is used, 3 parity bits are generated as the encoding result of X. Finally, 10 bits are used to represent X and Y after DSC. To decode them, Y can be retrieved directly. Then, with Y and the 3 parity bits, X can also be correctly decoded.

The works in [60]–[62], [65] extend the idea of DSC to IMVS to support low bitrate video with low view switching delay. For example, [60] proposes a multiview video coding structure shown in Fig. 2.2 with $2N+1$ views, where the horizontal index is the view index, and the vertical index is the time index. The frames of each view is divided into segments with N_s frames per segment, and each audience is

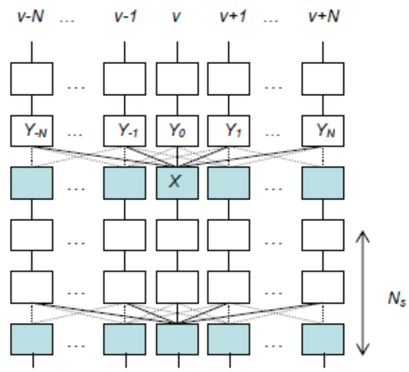


Fig. 2.2. Multiview coding structure that supports low delay view switch. This figure is from [60].

allowed to switch views at the end of each segment. Thus, N_s determines the view switching delay. To support low delay view switching (i.e., N_s is small) without significantly increasing the number of bits for the coded video, this work proposes to insert a DSC frame (the shaded rectangles in Fig. 2.2) for every N_s frames instead of an I-frame, and a DSC frame is much smaller than an I-frame. To understand how the DSC frame supports view switching, in Fig. 2.2, a DSC frame X is encoded using the prior frames Y_i from all views ($v - N \leq i \leq N + v$) as predictors. Then, frame X can be decoded with a prior frame from any view. Specifically, to encode X, each Y_i can be considered as a noisy version of X. Let Y_{max} denote the one prior frame with the largest difference from X. Then, the target of the DSC scheme is to encode X with enough parity bits, which can help retrieve X from Y_{max}. Thus, those parity bits are also enough to retrieve X from any other prior frame Y_i.

2.2.2.2 Cooperative Multiview Video Multicast in A Wireless Network

The work in [18] extends the above coding scheme to a cooperative multiview video wireless multicast system, which can help stop error propagation and improve the system reliability. In this system, video is also divided into segment with N_s frames per segment, where the first frame of a segment is a DSC frame. The BS multicasts all the views through different channels (frequency bands) of a wireless wide area network (WWAN) to all users at the same time. Each user can choose one view to decode and watch, and freely switch views at the end of a segment. Due to the fluctuation of wireless channels, a user i may receive the next segment correctly but not the current segment. However, without the last frame in the current segment, he/she cannot decode the next segment. To improve the system reliability, users who decode the current segment correctly share the last frame with others via a wireless local area network (WLAN). Then, user i can decode the next segment and the error propagation is stopped.

2.3 Incentive Mechanisms for Cooperative Wireless Communication and Cooperative Video Streaming

From the discussion in Chapter 2.1 and Chapter 2.2, user cooperation plays a fundamental role in improving the system performance in both cooperative wireless communication and cooperative video streaming. However, user cooperation cannot be guaranteed in such decentralized systems, since users are selfish and only care about their own performance. In the literature, many incentive mechanisms are proposed to simulate user cooperation. In this section, we first divide these mechanisms into two categories: incentive mechanisms without and with state change. For the incentive mechanisms without state change, we further divide them into two classes: incentive mechanisms for point-to-point interaction and for one-two-many interaction. In the following, we first discuss the two classes of incentive mechanisms without state change, and finally discuss incentive mechanisms with state change.

2.3.1 Incentive Mechanisms for Point-to-Point Interaction

In the literature, there is a big body of research on incentive mechanisms for cooperative wireless unicast and P2P live streaming that focus on point-to-point interaction. In this section, we classify those incentive mechanisms into three types.

• Direct reciprocity schemes in repeated games. In these schemes, it is assumed that a pair of users interact with each other for indefinite time duration, that is, they are not sure the exact time when they stop interacting with each other. Users who deviate from cooperation will be punished with a long term utility loss. Thus, this scheme is more effective if users expect to interact with each other for a longer time. Tit-for-tat [14] is a typical direct reciprocity scheme. Suppose that two users interact with each other and both of them cooperate at the beginning of the game. Then, they replicate the same action taken by his/her opponent in the last round. Thus, if either of the players deviates from cooperation for one round, this will result in non-cooperation and both of the users receive low utilities for the rest of the game.

Tit-for-tat based schemes are proprosed in [66]–[70] for cooperative wireless unicast. For example, the work in [66] proposes a tit-for-tat strategy for ad hoc networks, taking the network topology into consideration. If node i is on node j's route and can help node j forward messages, but node j is not on node i's route, node i will not help node j, since node i cannot receive any reciprocity in future interaction. Only when the pair of nodes can help each other and they both play tit-for-tat strategy, they will cooperate to forward messages.

The direct reciprocity mechanisms [11], [71]–[74] are also proposed for P2P live streaming. The work in [73] provides incentive using scalable video coding with a tit-for-tat strategy, where if user i can get a high download rate from a neighbor j, i will reciprocate j by providing a larger fraction of i's upload bandwidth. Therefore, a peer who contributes more upload bandwidth is more likely to obtain a larger share of neighbors' upload rates, thus receives more layers and has better video quality.

In P2P live streaming, packets may be dropped due to network congestion, and it is difficult to differentiate packet drop from intentional non-cooperation. To address this issue, the work in [11] proposes a credit-line mechanism. In this scheme, user i calculates the difference between the number of packets he/she uploaded to a neighbor j and the number of packets user j uploaded to him/her. User i continuous cooperate with user j as long as this difference does not exceed a predetermined credit line.

• Payment based schemes. In these schemes, virtual currency circulates in the network. Users need to pay to receive others' help, and users that help others will get paid to compensate their cost. When compared with the direct reciprocity game that is effective when each pair of users expect to interact with each other for a long time, the payment based scheme can stimulate user cooperation even if they know that they will not interact with each other from next round.

The works in [75]–[81] propose payment based schemes for cooperative wireless unicast. For example, in [76], the currency circulating in the network is called *nuglet*. Each user in the network is equipped with a tamper-resistant security hardware, called *nuglet counter* to record the nuglet each user possesses. If a user gets help from others, his/her nuglet counter is decreased. If he/she helps others, his/her nuglet counter is increased. Since everyone has to keep the nuglet counter non-negative, each user needs to help others to earn nuglet. However, this scheme requires the tamper-resistant hardware at each mobile device to track the transactions, which may not be satisfied in all mobile networks.

To address the above problem, the work in [75] proposes a purely software based payment scheme with the credit clearance service (CCS) provided by the central bank. In this work, after a user helps others forward packets, he/she keeps a receipt. When the channel condition to the central bank is good, he/she reports those receipts to the CCS. The CCS then processes the transactions and determines the payment or credit each user needs to pay or gets paid.

The payment based schemes are also proposed for P2P live streaming [82]–[84]. In [82], users can earn internal currency called *point* by uploading stream chunks to other users. Then, they use earned points to compete for connecting peers with high link capacities in a first price auction game, where the user with the highest bid will connect with the peer with the highest link capacity, and thus will receive the video with high quality. Furthermore, this system encourages off-session users (who are not receiving the streaming service or watching the video) to keep forwarding packets to accumulate their points for later use, which can effectively improve the system efficiency.

• Reputation based indirect reciprocity schemes. In these schemes, users help others to accumulate good reputations, and users with good reputations are likely to receive others' help. Therefore, node i helps node j with a good reputation is not because j helped i directly in previous interactions, but j helped someone else. Users' reputations can be updated in either centralized or distributed way. In the centralized system, a central authority will monitor user interaction, update their reputations and then broadcast to all players. In the distributed system, each user updates others' reputations based on both direct experience and indirect testimonies that he/she requests from neighbors.

When compared with the direct reciprocity game, the indirect reciprocity game is more suitable in the scenario where users change partners frequently (i.e., a pair of users expect to interact with each other for a short time.). When compared with the payment based game, it does not require tamper-resistant hardware at each user's side or the CCS provided by the central bank. However, the reputation based system requires extra singling traffic for reputation update.

The works in [85]–[89] propose reputation based schemes for cooperative wireless unicast. For example, the work in [87] proposes a mechanism called "CONFIDANT" in a mobile ad hoc network, which is a distributed reputation system. A scheme called *neighborhood watch* is used, where each user monitors and reports neighbors' behavior to other users. Each user gathers his/her direct experience and others' reports to identify misbehaving users.

The work in [89] proposes a centralized reputation updating scheme. A user will receive a high reputation if he/she cooperates with high reputation users or does not cooperate with low reputation users, and he/she will receive a low reputation otherwise. Thus, users will only cooperate with high reputation users.

The reputation based mechanisms are also proposed for P2P live streaming systems [90]–[93]. In [90], a rank based peer selection mechanism is proposed. In this scheme, a user who contributes more upload bandwidth is rewarded with a higher rank/priority to select peers, and thus, has better chance to connect to peers with high link capacities to receive a high quality video.

The work in [92] proposes an adaptive reputation updating scheme to stimulate users to keep cooperating, where for a user who has accumulated a high reputation, if he/she stops uploading, his/her reputation will drop quickly. Thus, users need to keep cooperating with others to maintain a high reputation.

2.3.2 Incentive Mechanisms for One-to-Many Interaction

In this section, we review incentive mechanisms for cooperative wireless multicast and P2P live streaming in a local area network (LAN), which are for one-to-many interaction.

The works in [78], [94] first propose a payment based method for cooperative wireless multicast in an ad hoc network, where a source node pays intermediate nodes for relaying messages to multiple destination nodes. In these works, each intermediate node claims the cost to relay a packet. Based on the claimed cost, the source calculates a multicast tree spanning all destination nodes with the minimum total payment. Then the source uses this multicast tree to deliver packets, and pays the intermediate nodes on the tree for relay service. To motivate intermediate nodes to report their true cost, a cheat-proof payment based scheme is proposed, where reporting the true cost is the dominant strategy for each intermediate node.

The above incentive mechanism describes a scenario where one user needs multiple users' help. In this thesis, we are more interested in another scenario, where multiple users need one user's help. For example, in two-hop cooperative wireless multicast [95], after the BS broadcasts a packet, one successful user is selected and he/she will decide whether to forward the packet as a relay. Once he/she relays the packet, the transmission can be overheard by all unsuccessful users. Thus, the relay cannot choose different actions towards different unsuccessful users.

To stimulate user cooperation in this scenario, the work in [95] proposes a direct reciprocity game with the worst behavior tit-for-tat strategy. In this system, time is divided into slots. In each slot, each user has the same probability to be a successful user and be selected to relay the packet. Each selected user chooses a transmission power to forward the packet. A higher transmission power gives a higher probability for others to receive the packet correctly, but introduces higher cost to the relay. Since he/she cannot choose with whom he/she cooperates, he/she uses the lowest observed transmission power. Specifically, each user monitors the forwarded packets by others, and estimates their transmission power. He/she discovers the lowest transmission power among all relays in previous rounds, and uses that power when being selected. Thus, if any user deviates to a low power, all others will use the same low power for a long time as penalty, which lowers everyone's payoff. Therefore, no one has incentive to deviate from the high transmission power.

The work in [96] proposes an incentive mechanism to study the one-to-many interaction for a P2P live streaming system in a LAN. In this system, each user can decide whether to be an agent, who requests the video data from the server, and the video data will be shared in the LAN. Thus, an agent uses his/her own network resources to request video, while all others can free ride. In this case, users tend to free ride rather than to be an agent. However, if there is no agent, everyone has no video to watch and thus receives a low utility. To address this issue, the work in [96] proposes an evolutionary game for users to learn how to address this tradeoff between lowering the cost and receiving the video, and derives a stable Nash Equilibrium, where even if players may sometimes deviate from this equilibrium, they will still move back, since users who take strategies in the equilibrium will always receive a higher utility.

2.3.3 Stochastic Game for User Interaction With State Change

In the literature, stochastic game is used to model user interaction with state change. Suppose that the game is currently at a state $s \in \mathcal{S}$ ($\mathcal{S}$ is the state set.). There are N players, and a player i selects action a_i from his/her action set $\mathcal{A}_i$. Based on current state and all players' actions, each player receives an immediate payoff $u_i : \mathcal{S} \times \mathcal{A}_1 \times ... \times \mathcal{A}_N \to (-\infty, +\infty)$, and the game will transit to another state with a state transition probability, $P : \mathcal{S} \times \mathcal{A}_1 \times ... \times \mathcal{A}_N \times \mathcal{S} \rightarrow [0, 1]$. Then, users take actions at the new state, and this process is repeated with finite horizon. For each player in the game, his/her goal is to find the optimal action policy (an action policy defines his/her action at each state) to maximize the expected lifetime utility $E\left[\sum_{j=0}^{+\infty} \eta u_{i,t+j}\right]$. Here, η is a discounting factor that values how users care the future payoffs, and $t + j$ is the time index. At the Nash Equilibrium of a stochastic game, no one has incentive to unilaterally deviate from his/her action policy at any state of the game.

For example, the work in [97] considers a cooperative communication scenario with three nodes, where two nodes transit packets to a common destination and they also help each other forward packets. Each of them also maintains a buffer to store packets that have not been transmitted. With the classical collision channel, only one node can transmit in a time slot. Thus, they need to decide who to transmit. In this system, the states are the numbers of packets in the two nodes' buffers, since with more packets in the buffer, they have higher desire and receive higher utility to transmit packets. With decisions made by the two nodes, their states change, which also affects their future decisions. The works in [97], [98] propose a stochastic game to formulate the above decision making process, and derive the optimal strategies for the two nodes to maximize the expected lifetime utilities.

The works in [99], [100] formulate stochastic games for a foresighted resource reciprocity problem in P2P live streaming. For example, in [99], a user's state is the bandwidth he/she gets allocated from the neighboring users. He/she then needs to take an action on how to allocate his/her own bandwidth to neighboring users for reciprocity. Since this action may change the neighboring users' responses, he/she will receive different bandwidth allocation and transit to another state. To estimate the state transition probability, he/she can learn from his/her past experience, and then the optimal reciprocity strategy is found to maximize his/her lifetime utility.

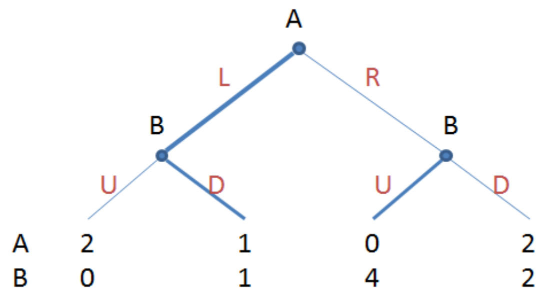


Fig. 2.3. An example of Stackelberg game. There are two players: A and B. A has two strategies, L and R. B has two strategies U and D.

2.4 Game Theory Review

Game theory [14], [15] provides important tools to study user behavior in dynamic networks. In a game, there are three essential components, *player set*, *strategy sets* and *utility functions*. Player set defines the players that play in the game. Each player i has a strategy set S_i, including the strategies he/she can use in the game. He/she chooses a strategy s_i from S_i to play with others, and all N players' strategies form a strategy profile $\xi = \{s_1, s_2, ..., s_N\}$. Each player has a utility function $U_i(\xi)$, which measures his/her payoff based on all players' strategies ξ.

To analyze a game, game theory provides a very important concept called *Nash Equilibrium*. A Nash Equilibrium is defined as a strategy profile ξ^*, where each player i's strategy s_i^* is the best response to the others' strategies in the profile, (i.e., $\xi_{-i}^* = \{s_1^* ..., s_{i-1}^*, s_{i+1}^* ..., s_N^*\}$) and we have

$$U_i(s_i^*, \xi_{-i}^*) \geq U_i(s_i, \xi_{-i}^*) \text{ for } \forall s_i \in S_i. \tag{2.1}$$

In other words, for a player i, if all other players keep their strategies ξ_{-i}^* unchanged, i will receive a lower utility when he/she unilaterally deviates from s_i^*. Thus, in a Nash Equilibrium, no rational player has incentive to change to another strategy.

2.4.1 Stackelberg Game and Subgame Perfect Nash Equilibrium

In practice, users in a game do not necessarily make decisions at the same time. For users who make decisions first, their actions may be observed by others, and thus, may affect other players decision making. In game theory, *dynamic game* can be used to model such interaction, where players make moves following a certain predetermined order and select their strategies sequentially. Stackelberg game is an example of dynamic game, which typically has two stages. Players who move at the first stage are called leaders. The other players are followers, who make decisions after observing leaders' actions. Figure 2.3 shows an example of a Stackelberg game with two players, A and B. A has two strategies, L and R, and makes a move first. B also has two strategies, U and D. After observing A's action, B selects his/her strategy. Their payoffs are enumerated at the bottom of the figure. For example, if A chooses L and B chooses U, A's payoff is 2, while B's payoff is 0.

To analyze the Stackelberg game, a concept called *Subgame Perfect Nash Equilibrium* (SPNE) is used. A strategy profile is a SPNE if a user cannot increase his/her payoff by unilaterally deviating to any other strategy from any stage. *Backward Induction* is used to find the SPNE of the Stackelberg game. It starts from the last stage of the game, which is player B's game in the above example, and finds player B's optimal strategy for each possible outcome in stage 1. In Fig. 2.3, if A selects L in stage 1, player B's optimal strategy is D that gives a higher payoff of 1. Similarly, if A chooses R in stage 1, player B should select U. Then, the game analysis moves one stage up and analyzes player A's strategy. If A selects L, from the previous analysis, player B will selects D, which gives player A a payoff of 1. If A selects R, player B will select U in stage 2, which results in a payoff of 0 for player A. Comparing these two, A will select L, and the strategy profile (L, D) is the SPNE of this game. The same idea can be used to analyze games with more than two stages.

2.4.2 Evolutionary Game Theory

In a game, it is possible that there are more than one Nash Equilibria, and to which equilibrium players will converge is an interesting problem. Furthermore, sometimes, players may only have limited information, and they may not know what the Nash Equilibrium is or how other players will play, which may result in non-rational behavior due to uncertainties. To address this problem, evolutionary game theory [101], [102] provides tools to study a stable Nash equilibrium, where players can learn from others or past experience, adjust their strategies towards higher payoffs and finally converge to the stable equilibrium.

In evolutionary game, an important concept, called *Evolutionarily Stable Strategy* (ESS) [102], states that under the condition that a strategy is prevalent (i.e., it is taken by most of players in the system), it is an ESS if it can resist a small group of mutant players with any other strategy (i.e., this small group of mutant players will finally be extinct during the evolution). Mathematically, we have the definition as follows:

Definition 2.1. *A strategy z^* is an ESS if and only if, $\forall z \neq z^*$, z^* satisfies*

- *equilibrium condition: $U(z, z^*) \leq U(z^*, z^*)$, and*
- *stability condition: if $U(z, z^*) = U(z^*, z^*)$, $U(z, z) < U(z^*, z)$.*

where $U(z_1, z_2)$ is a player's utility when he/she uses strategy z_1 and the other player uses strategy z_2.

From the equilibrium condition of Definition 1, we observe that an ESS has to be a Nash Equilibrium. From the stable condition, we observe an important property of ESS, i.e., even if at some time instance, some players deviate from the ESS, they will still come back to the ESS, since the one who uses ESS receives a higher payoff.

To derive the ESS, evolution game theory provides a very useful tool, called *replicator dynamics* [102]. Let S be a strategy set with size $|S|$. Let x_i be the population share playing strategy $s_i \in S$, where $x_i \in [0, 1]$. Let $x = \{x_1, x_2, ..., x_{|S|}\}$.

By replicator dynamics, the dynamic of x_i is given by the following differential equation:

$$\dot{x}_i = \eta(\bar{U}(s_i) - \bar{U}(x))x_i, \tag{2.2}$$

where η is a constant step size, $\bar{U}(s_i)$ is the average payoff of individuals using strategy s_i, and $\bar{U}(x)$ denotes the average payoff of the entire population. The intuition behind this differential equation is that if players using strategy s_i have a higher payoff than the average payoff of the entire population, the corresponding population share x_i should increase. At the stable state, this differential equation should be equal to 0. If there is only one non-zero item in x at the stable state, e.g., $x_i = 1$, this means all players finally take the pure strategy s_i, which is the ESS. If there are more than one non-zero item in x, x can be interpreted as a mixed strategy [102], where x_i in x denotes the probability that a player uses strategy s_i. Then, this mixed strategy x is the ESS.

2.4.3 Auction Game

In auction games, the auctioneer has a good to sell to a group of bidders, and the auctioneer decides the bidder who will buy the good and the price the bidder will pay following predetermined auction rules. In practice, auction has been proven to be effective in allocating a good to the bidder that values this good with the highest price, even though bidders' values about the good are their private information. In auction theory [103], there are four well known auctions: English auction, Dutch auction, first-price sealed-bid auction and second-price sealed-bid auction.

In English auction, the auctioneer increases the price for the good round by round. A bidder will stay in the auction until the price is too high for him/her. When the second last bidder leaves, the auction ends and the winner is the last bidder and he/she will pay the price when the auction ends. In the Dutch auction, the auctioneer reduces the price round by round until one bidder accepts the price. This bidder wins the auction and pays the price he/she accepted. In first-price sealed-bid auction, each bidder put his/her bid in a sealed envelop and then submits to the

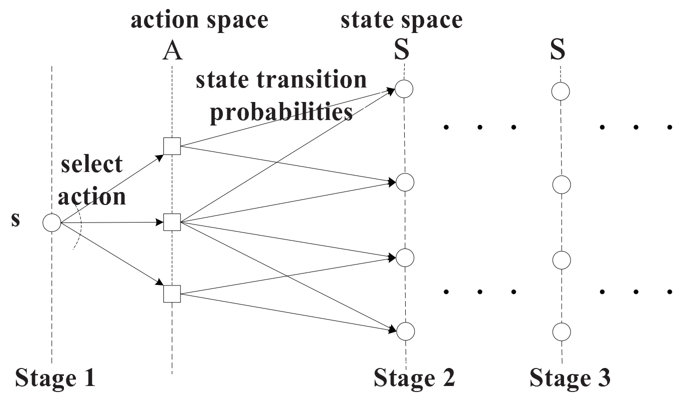


Fig. 2.4. An example of MDP.

auctioneer. Then, the bidder with the highest bid is the winner, and he/she will pay his/her bid. In second-price sealed-bid auction, the bid submitting is the same as the first-price sealed-bid auction. The winner is still the bidder with the highest bid, but he/she only needs to pay the second-highest bid.

[103] shows that English auction and the second-price sealed-bid auction are equivalent, while Dutch auction and the first-price sealed-bid auction are equivalent. When comparing the second-price and the first-price sealed bid auctions, the second-price sealed-bid auction is a true-telling auction, where each bidder bids the true value is a dominant strategy. This property is useful to reveal users' private information.

2.4.4 Stochastic Game and Markov Decision Process

As discussed earlier, stochastic game models players' interaction with state change. In a stochastic game, if a player's state transition and the utility function rely on his/her own action but not others',[1] his/her strategy selection can be simplified as an MDP, which is discussed as follows.

An MDP [104] is a mathematic model of decision making process. A user makes decisions at different states, and the impacts of each decision involve ran-

[1] This is true when we verify whether an strategy profile is an Nash Equilibrium, where we study a user's strategy selection given that others all take their strategies in the profile. Thus, from this user's perspective, given others' strategies are fixed, his/her state transition and the utility function depends on his/her own action only.

domness. The goal of an MDP is to find the optimal decision at each state to maximize the expected payoff of the entire process. Specifically, an MDP is defined as a four-tuple, $M = \langle \mathcal{S}, \mathcal{A}, P, U \rangle$, and we have

- $\mathcal{S}$, the state space.
- $\mathcal{A}$, the action space.
- $P^a_{s\to s'}$, the state transition function: $\mathcal{S} \times \mathcal{A} \times \mathcal{S} \to [0, 1]$, which calculates the probability of transiting to a state $s' \in \mathcal{S}$ when an action $a \in \mathcal{A}$ is taken at state $s \in \mathcal{S}$.
- U^a_s, the expected short-term payoff function: $\mathcal{S} \times \mathcal{A} \to (-\infty, +\infty)$, which calculates the expected short-term payoff for an action a taken at state s.

Fig. 2.4 shows an example of an MDP with infinite horizon, where each circle and square represent a state and an action, respectively. At stage 1, a player is at a state $s \in \mathcal{S}$. Suppose that he/she takes action $a \in \mathcal{A}$. He/she will receive an expected short-term payoff U^a_s and the state will randomly transit to another state s' in the next stage following the state transition probability $P^a_{s\to s'}$. Then, for each state he/she may transit to at stage 2, he/she needs to select an action and this process is repeated at each stage to infinity. In the MDP, an action policy is defined as $\pi = \{a_s | s \in \mathcal{S}\}$, which determines the action a_s at each state s. For each π, the corresponding lifetime utility can be written as a Bellman equation [104],

$$W_s(\pi) = U^{a_s}_s + \eta \sum_{s' \in \mathcal{S}} P^a_{s\to s'} W_{s'}(\pi), \tag{2.3}$$

where the second term is the expected lifetime utility since the next stage and η is the discounting factor. This equation recursively define the expected utility of the entire decision making process, and dynamic programming can be used to find the optimal π to maximize $W_s(\pi)$.

Chapter 3

Two-Hop Cooperative Wireless Multicast: Incentive Mechanism and Analysis

From the literature survey in Chapter 2, the issue of one-to-many interaction raised in wireless multicast systems is seldom addressed. The work in [95] proposes the worst behavior tit-for-tat strategy to address this problem, which is effective when users expect to interact with each other for a long time. However, in wireless multicast, users may frequently join and leave the multicast service, which makes this method impractical. Furthermore, the work in [95] assumes homogenous users who have the same cost to forward a packet. However, in reality, with different mobile devices, users may have heterogenous cost to forward a packet, which is their private information. They may cheat if cheating can help get a higher payment.

In this chapter, to address the one-to-many interaction, we model user interaction as a multi-seller multi-buyer payment based game, where unsuccessful users pay to receive relay service and successful users will get paid if they forward packets as relays.[1] Our game can stimulate user cooperation even if they know that they

[1] In wireless multicast systems, the BS can provide the service with different Quality-of-Service (QoS) depending on the agreement between the BS and intended receivers. One example is the BS charges the intended receivers for subscription fees, and accordingly, it has to provide reliable multicast service. In this case, the BS can pay a successful user to ensure his/her cooperation as

will leave the multicast service in the next round. Thus, our scheme works in dynamic wireless networks, where users can frequently join and leave the system. To address the problem where users have different cost and may cheat on such information, we formulate the sellers' game as a second-price sealed-bid auction game. It is a truth-telling auction, where each user are encouraged to bid their true cost, since bidding the true cost is a weakly dominant strategy.

In our payment based game, we model the buyers' game as an evolutionary game, and derive the ESS. It is a stable equilibrium, where even if some players may deviate from it at some time, they will still move back to the ESS, since using the ESS gives a higher utility. Unlike the work in [105], we further investigate how the price affects users' decisions and the system performance. We observe that at different prices, the buyers' game can converge to different ESS, where unsuccessful users have different probabilities to free ride (i.e., not buy but overhear the relay bought by others), resulting in different system throughput. From the system designer's point of view, we aim at selecting the optimal price to maximize the system throughput. For the simple scenario with homogeneous users who have the same cost, we derive the closed-form optimal price, under which unsuccessful users cannot free ride, and they will share the cost of the relay and pay together to afford the relay service, while the system throughput is maximized at the same time. For the scenario with heterogeneous users who have different cost, we propose an efficient algorithm to find the optimal price, under which unsuccessful users have very low probability to free ride and the system throughput is also maximized.

3.1 System Model

In this section, we will introduce the cooperative wireless multicast system and the multi-seller multi-buyer payment based game model.

---

a relay so that the BS can maintain a high QoS. Another example is that the BS does not charge from receivers, only provides the service with the best effort, and does not guarantee any QoS. In this case, users in the service has to decide by themselves on whether and how to cooperate. The problem in the later example is more complicated and more interesting than the first one, which we will study in this work.

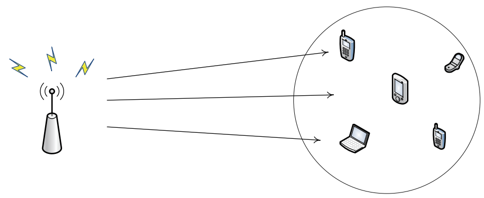


Fig. 3.1. System model.

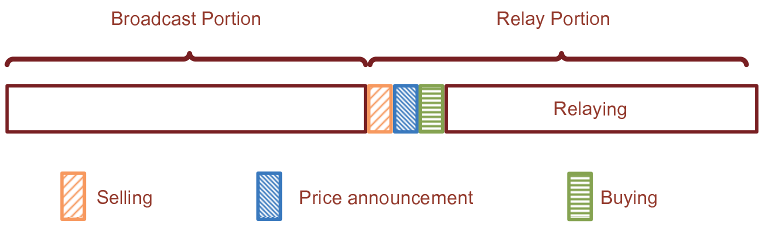


Fig. 3.2. The procedure of a segment transmission.

3.1.1 Two-Hop Cooperative Wireless Multicast

A BS provides multicast service to a group of users, who are close to each other in a circular area as shown in Fig. 3.1. We consider a dynamic network, where users frequently join and leave the multicast service. Let $N(t)$, or in its short form N (for presentation simplicity), be the number of users at time t. The data traffic is divided into segments, and for the transmission of each segment, we consider a two-portion wireless multicast as shown in Fig. 3.2. In the broadcast portion, the BS broadcasts a segment. Then, the users who receive the segment correctly are successful users, and they decide whether to provide relay service to unsuccessful users. In this work, we consider a simple scenario where at most one successful user forwards the segment in the relay portion. At the beginning of the relay portion, there are some information exchanges among the users, to be detailed in Chapter 3.1.2. Similar to the work in [39], we assume that all communications in the relay portion, including information exchanges and segment relaying, are on a different frequency band from the band used by the BS. Therefore, when the BS finishes broadcast of one segment, it can start broadcasting the next segment immediately.

In this work, it is assumed that the distances from the BS to the users are much larger than those between users. Therefore, each user has the same probability, denoted by p_1, to receive a segment from the BS successfully. Since users are close to each other, we assume that all information exchanges and segment relaying in the relay portion are received correctly with probability 1 by all users.

To evaluate the system performance, we define the relay portion throughput, T_R, as the average percentage of unsuccessful users who receive the segment correctly in the relay portion.

3.1.2 Payment Based Game Formulation

In this systems, relays use their own power to forward segments and help others, but they cannot benefit during this process. To stimulate user cooperation, in this work, we model users' interaction as a multi-seller multi-buyer payment based game, where each successful user decides whether to sell relay service, and each unsuccessful user decides whether to purchase it. To implement the billing process, we assume that there exists a trusted local agent, who listens to the data transmission in the relay portion, charges fees from buyers, and pays the relay.

Our multi-seller multi-buyer cooperative multicast game is a 4-stage Stackelberg game as described in details below.

Stage 1: The Sellers' Game.[2] After the broadcast portion, suppose that there are N_{su} successful users. Each of them decides whether to sell relay service. Let $\{S, NS\}$ denote their strategy set, including being a seller (S) and not being a seller (NS). Suppose that N_s $(\leq N_{su})$ successful users decide to be sellers. They will send feedbacks to the local agent attached with their IDs and the cyclic redundancy check (CRC)[3] bits of the received segment. Note that they send these messages simultaneously[4] in the selling part at the beginning of the relay portion as shown in Fig. 3.2, which can prevent a successful user from observing other sellers' messages and adjusting his/her own decision.

[2] In this work, we arrange the sellers' game in the first stage, since revealing the number of sellers in this stage can help in the later optimal price setting to maximize the system performance.

[3] Letting sellers attach the CRC bits of the received segment can prevent an unsuccessful user from pretending to be a successful user.

[4] This can be achieved by code-division multiple access (CDMA) technology. Each node is assigned a unique code. The code is used to spread the node's message. The local agent monitors codes of all the users.

Stage 2: The Price Setting Game. If $N > N_s > 0$,[5] the local agent selects a seller to provide relay service if there are more than one sellers, and selects a relay price q that will be charged to each buyer for the relay service, and then announces to all users the number N_s of sellers, the user ID of the selected seller, and the relay price q, in the price announcement part in Fig. 3.2. Details of seller and price selections are given in subsequent sections.

Stage 3: The Buyers' Game. In Stage 3, each unsuccessful user decides whether to purchase the relay service at price q. Let $\{B, NB\}$ denote their strategy set, including being a buyer (B) and not being a buyer (NB). All buyers broadcast their IDs simultaneously using CDMA technology in the buying part in Fig. 3.2. The selected seller hears the buyers' messages and knows the number of buyers, denoted N_b ($\leq (N-N_s)$), and thus knows the total payment, $N_b q$, that unsuccessful users provide for the relay service.

Stage 4: The Transaction Game. For the selected seller, if forwarding the segment is profitable, i.e., the selected seller can gain a non-negative net utility, then he/she will forward the segment in the relaying part in Fig. 3.2; otherwise, he/she will not forward. After the relaying, the local agent charges from the buyers and pays to the relay node.

This game is repeated for the transmission of all segments.

3.1.3 Utility Functions

For each user i, let g denote the utility gain of receiving a segment correctly, and c_i denote his/her cost to forward one segment. In the 4-stage Stackelberg game, if user i is a successful user, his/her utility function is his/her received payment minus c_i if he/she is the selected seller and forwards the segment, and 0 otherwise. If user i is an unsuccessful user, his/her utility function is $(g-q)I_{relay}$ if he/she is a buyer, and gI_{relay} otherwise (i.e., user i is a free-rider). Here I_{relay} is a binary value: $I_{relay} = 1$ if there is relay service in the relay portion, and $I_{relay} = 0$ otherwise. Note that we ignore the cost of information exchanges in the selling and buying parts, since the

[5] If $N_s = N$, all users are successful after the broadcast portion. If $N_s = 0$, there is no seller. In either scenario, there is no need for the following stages and the game ends.

amount of related information exchanges is small.

3.2 Game Analysis with Homogeneous Users

We start with a simple scenario where the users are homogeneous, i.e., they have the same cost of providing relay service with $c_i = c$ being a positive constant. In addition, when there are more than one sellers, the local agent will randomly select one to forward the segment, and all sellers have the same probability to be selected.

We use backward induction to find the SPNE of the game. Typically, backward induction first analyzes the last stage of the game, moves up stage by stage, and studies the first stage the last. However, the result of the transaction game can simplify the sellers' game and we can easily find the sellers' optimal strategy that belongs to the SPNE. Thus, we will study the sellers' game after the transaction game. The result of the analysis for the sellers' game can help reduce the number of possible outcomes of the sellers' game, and simplify the analysis of the price setting game and the buyers' game.

3.2.1 The Transaction Game

After buyers broadcast their decisions, the selected seller knows the amount of payment buyers offer, $N_b q$. The selected seller will forward the segment if $N_b q \geq c$, and will not forward otherwise. Therefore, the transaction game ensures that the selected seller will always receive a non-negative net utility gain in the game. After the relaying, the local agent charges price q from each buyer and pays $N_b q$ to the relay node.

3.2.2 The Sellers' Game

Since the selected seller will make a non-negative net utility gain in the transaction game, the sellers' game has an obvious solution belonging to the SPNE, i.e., all successful users take strategy S and become sellers. This is because by taking the strategy S and being a seller, a successful user's utility gain in the relay portion is no less than zero, while by taking the strategy NS and not being a seller, his/her utility gain in the relay portion is zero. Therefore, S is a weakly dominant strategy over NS, and every successful user should choose it. Thus, after the sellers' game we have $N_s = N_{su}$, i.e., the number of sellers equals the number of successful users. As the local agent will announce (in the price setting game) the number N_s of sellers, all unsuccessful users will know the value of $N_{su} = N_s$ before the buyers' game.

Given the above analysis on the transacting game and sellers' game, we then study the stage 2, the price setting game, and stage 3 the buyers' game. With backward induction, we first analyze stage 3 under any price q selected in stage 2. We then move upwards to stage 2 and study the optimal price selection.

3.2.3 The Buyers' Game

Given the relay price q decided by the local agent and the number N_{su} of sellers, unsuccessful users decide whether to purchase the relay service. Recall that when the total payment from all buyers $N_b q$ is no less than the cost c, the selected seller will relay the segment. Due to the broadcast nature of wireless communications, unsuccessful users who do not pay may overhear the segment forwarded by the relay and enjoy a free ride. Here, unsuccessful users face a dilemma: everyone wants to free ride the relay service bought by others and pay nothing, while there will be no relay service if there are not sufficient buyers, and every unsuccessful user will gain nothing. To solve this problem, [6] we model the buyers' game as an evolutionary game [102], and derive the ESS, which is a stable Nash Equilibrium. This means that, even if some players deviate from the ESS, they will still come back to the ESS, since using the ESS gives a higher payoff.

To derive the ESS, as discussed in Chapter 2.4.2, we use *replicator dynamics*. In our game, each unsuccessful user has two strategies: B or NB. For all unsuccessful users, let x be the population share playing strategy B, where $x \in [0, 1]$, and the rest $(1 - x)$ population share plays strategy NB. By replicator dynamics, we have

[6] One alternative solution is using encryption, where only the buyers can receive the correct key to decrypt the segment sent by the relay. However, this method requires sophisticated key management mechanisms to ensure that only the buyers can receive the correct key. Furthermore, since for different segment the intended recipients are different, the key should be updated for each segment transmission, which introduces a large amount of extra information exchange.

the following differential equation:

$$\begin{aligned}\dot{x} &= \eta(\bar{U}_B(x) - \bar{U}(x))x = \eta[\bar{U}_B(x) - x\bar{U}_B(x) - (1-x)\bar{U}_{NB}(x)]x \\ &= \eta x(1-x)f(x), \end{aligned} \tag{3.1}$$

where $\dot{x}$ is the population increase of strategy B, η is a constant step size, $\bar{U}_B(x)$ is the average payoff of using pure strategy B, $\bar{U}_{NB}(x)$ is the average payoff of using pure strategy NB, $\bar{U}(x) = x\bar{U}_B(x) + (1-x)\bar{U}_{NB}(x)$ denotes the average payoff of the population, and $f(x) = \bar{U}_B(x) - \bar{U}_{NB}(x)$. The intuition behind this differential equation is that if using pure strategy B introduces a higher payoff than the average payoff of the entire population, the population share of pure strategy B should increase. At the stable state x, this differential equation should be equal to 0. As discussed in [102], [105], [106], the population share x can be interpreted as a mixed strategy, which denotes the probability that players adopt pure strategy B. Since any unsuccessful user gets the same gain g if he/she correctly receives a segment, all unsuccessful users are symmetric and should have the same mixed strategy x, denoted x^*, when they reach the ESS. For presentation simplicity, we say ESS is x^*. In the following, given the number N_{su} of sellers, and for any relay price q selected by the local agent, we derive $\bar{U}_B(x)$ and $\bar{U}_{NB}(x)$ for unsuccessful user i, and then find the ESS x^*.[7]

3.2.3.1 Analysis of $\bar{U}_B(x)$ and $\bar{U}_{NB}(x)$

Given N_{su} sellers, for unsuccessful user i, let $\mathcal{X}_{-i}$ denote the set of all other unsuccessful users. So $|\mathcal{X}_{-i}| = l \triangleq N - N_{su} - 1$. Recall that each unsuccessful user purchases the relay service with probability x. Therefore, the number of buyers in

[7] In games with incomplete information, each user has private information, which is unknown to the others. Replicator dynamics can help solve games with incomplete information (e.g. [106]), where the game is repeated for multiple shots, and users learn from the interactions with others, adjust their strategies towards a higher payoff, and finally may reach the ESS. Unlike the game with incomplete information, our game is a one-shot game with complete information, where each user's gain g, the number of unsuccessful users $(N - N_{su})$, and the relay cost c are all public information. Thus, similar to [105], the ESS can be derived directly by solving (3.1), and there is no learning process involved in our game analysis.

$\mathcal{X}_{-i}$, denoted k, follows Binomial distribution $B(l, x)$.

In this context, if user i decides to be a buyer, the total number of buyers is $(k+1)$, and thus, the total payment from all buyers is $(k+1)q$. If $(k+1)q \geq c$, this payment can afford the relay service, and user i receives the segment correctly and pays the price q; otherwise, there is no relay service and user i's utility in the relay portion is 0. Therefore, in the relay portion, user i's average utility of strategy B is

$$\bar{U}_B(x) = (g - q) \sum_{k=0}^{l} \binom{l}{k} x^k (1-x)^{l-k} I\left[(k+1)q \geq c\right], \tag{3.2}$$

where $I[\cdot]$ is an indicator function. If user i decides not to be a buyer, the total number of buyers is k, and the total payment is kq. If $kq \geq c$, this payment can still afford a relay. After the relay portion, user i can overhear the relay and receive the segment correctly. Therefore, the average utility of the strategy NB is

$$\bar{U}_{NB}(x) = g \sum_{k=0}^{l} \binom{l}{k} x^k (1-x)^{l-k} I[kq \geq c]. \tag{3.3}$$

Then, we have

$$\begin{aligned} f(x) &= \bar{U}_B(x) - \bar{U}_{NB}(x) \\ &= \sum_{k=0}^{l} \binom{l}{k} x^k (1-x)^{(l-k)} \Big\{ \Big(I[(k+1)q \geq c] - I[kq \geq c] \Big) g - I[(k+1)q \geq c] q \Big\} \\ &= g \binom{l}{k^*} x^{k^*} (1-x)^{(l-k^*)} - q \sum_{k=k^*}^{l} \binom{l}{k} x^k (1-x)^{(l-k)}, \end{aligned} \tag{3.4}$$

where $k^* = \lceil c/q \rceil - 1$, and $\lceil \cdot \rceil$ is the ceiling function. Here, $\lceil c/q \rceil$ is the minimal number of buyers required to afford the relay service at price q.

3.2.3.2 The ESS Solution

From (3.1), at the stable state $\dot{x} = 0$, there are three possible solutions: $x = 0$, $x = 1$, and x that satisfies $f(x) = \bar{U}_B(x) - \bar{U}_{NB}(x) = 0$. In our game, the relay price q plays an important role in the unsuccessful users' decision-making process, and all those three solutions can be ESS x^*, which will be discussed as follows. The

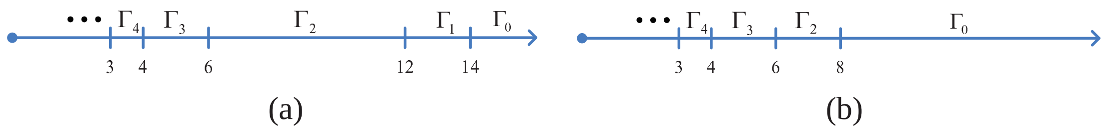


Fig. 3.3. Examples of price intervals. (a) $g = 14$ and $c = 12$. (b) $g = 8$, $c = 12$ and Γ_1 does not exist in this case.

analysis results are summarized in Theorem 3.1 following the analysis.

Define $\underline{j} = \lfloor c/g \rfloor + 1 \geq 1$, where $\lfloor \cdot \rfloor$ is the floor function. We partition the price range $[0, +\infty)$ into the following subintervals:

$$\Gamma_0 = [g, \infty)\,, \Gamma_{\underline{j}} = \left[\frac{c}{\underline{j}}, g\right)\,, \text{ and } \Gamma_j = \left[\frac{c}{j}, \frac{c}{j-1}\right) \text{ for } j > \underline{j}. \tag{3.5}$$

When the price q is in range Γ_j with $j \geq \underline{j}$, at least j buyers are needed to afford the relay service. Γ_0 is the range of the price that equals or exceeds users' utility gain of receiving a segment correctly. Fig. 3.3 shows examples of the price intervals when g and c take different values.

• Case 1, $q \in \Gamma_0$, i.e., $q \geq g$: From (3.4), for all $x \in [0, 1]$, we have

$$\begin{aligned} f(x) &= \binom{l}{k^*} x^{k^*} (1-x)^{(l-k^*)} g - \sum_{k=k^*}^{l} \binom{l}{k} x^k (1-x)^{(l-k)} q \\ &\leq \binom{l}{k^*} x^{k^*} (1-x)^{(l-k^*)} (g-q) \leq 0. \end{aligned} \tag{3.6}$$

Thus, the strategy NB always outperforms B and users will converge to $x^* = 0$, which is the ESS. This is because given $q \geq g$, the price is too high when compared to the utility gain from receiving the relay service. Thus, nobody will buy.

• Case 2, $q \in \Gamma_j$ with $j \in \{\underline{j}, \underline{j}+1, ..., N - N_{su} - 1\}$: In this case, we analyze the ESS when $j = \underline{j} = 1$ (which happens only when $c < g$) and when $1 < j < N - N_{su}$ separately.

When $j = \underline{j} = 1$, i.e., $q \in \Gamma_1 = [c, g)$, one buyer is sufficient to buy the relay service and $k^* = \lceil c/q \rceil - 1 = 0$ in (3.4). Therefore, $f(x)$ in (3.4) can be simplified as $f(x) = (1-x)^{N-N_s-1} g - q$ with $f(0) = g - q > 0$ and $f(1) = -q < 0$. In addition, $f'(x) = -g(N - N_s - 1)(1-x)^{N-N_s-2} < 0$ and thus, $f(x)$ is a decreasing

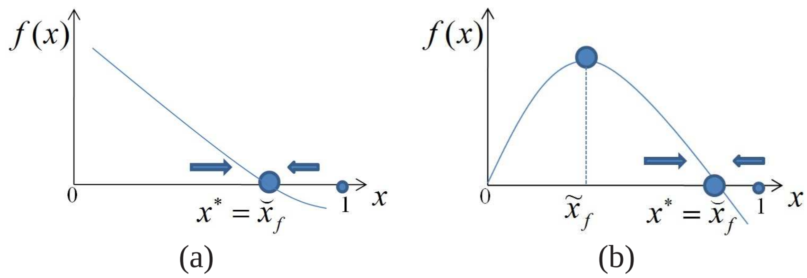


Fig. 3.4. (a) $q \in \Gamma_j$, where $j = \underline{j} = 1$. (b) $q \in \Gamma_j$, where $1 < j < N - N_{su}$.

function for $x \in (0,1)$, as shown in Fig. 3.4a. Thus, $f(x) = 0$ has a single root $\breve{x}_f = (1 - \sqrt[l]{\frac{q}{g}}) \in (0,1)$, which is the ESS. To understand this, let x deviate from $\breve{x}_f$. If $x \in [0, \breve{x}_f)$, we have $f(x) > 0$, which means strategy B can give a higher utility than NB. Therefore, users will increase the probability of using B and x will move towards $\breve{x}_f$. Similarly, if $x \in (\breve{x}_f, 1]$, we have $f(x) < 0$, which means strategy B will give a lower utility. Thus, users will reduce the probability of using strategy B and adjust their strategy towards $\breve{x}_f$. Thus, $x^* = \breve{x}_f$ is the ESS.

When $q \in \Gamma_j$ with $1 < j < N - N_{su}$, i.e., $q \in [\frac{c}{j}, \frac{c}{j-1})$, at least j buyers are required to afford the relay service, and $k^* = \lceil c/q \rceil - 1 > 0$. From (3.4), $f(0) = 0$ and $f(1) = -q < 0$. In addition, we prove in Appendix A that $f'(x) = 0$ has a single root $\tilde{x}_f$ in the range $(0,1)$, where $f'(x) > 0$ when $x \in (0, \tilde{x}_f)$ and $f'(x) < 0$ when $x \in (\tilde{x}_f, 1)$, as shown in Fig. 3.4b. Therefore, $f(\tilde{x}_f) > 0$, and $f(x) = 0$ has a single root $\breve{x}_f$ in the range $(\tilde{x}_f, 1)$. Same as the analysis in Fig. 3.4a, $x^* = \breve{x}_f$ is the ESS.

• Case 3, $q \in \Gamma_{N-N_{su}} = [\frac{c}{N-N_{su}}, \frac{c}{N-N_{su}-1})$: In this price range, the relay price requires at least $(N - N_{su})$ buyers, while there are $(N - N_{su})$ unsuccessful users and thus at most $(N - N_{su})$ buyers. Therefore, there is no chance to free ride, and all unsuccessful users will buy with $x^* = 1$. Mathematically, when $q \in \Gamma_{N-N_{su}}$, $k^* = \lceil c/q \rceil - 1 = N - N_{su} - 1 = l$ and $f(x)$ in (3.4) can be simplified as $f(x) = (g - q)x^{N-N_{su}-1} \geq 0$ for all $x \in [0,1]$. Therefore, the strategy B always outperforms strategy NB, and $x^* = 1$.

• Case 4, $q \in \Gamma_j$ with $j \geq N - N_{su} + 1$: In these price ranges, at least $j > N - N_{su}$ buyers are required to afford the relay service, while there are only $(N - N_{su})$

unsuccessful users. Therefore, there are not sufficient buyers to afford the relay service, and the game ends.

In summary, we have the following theorem.

Theorem 3.1. *Given N_{su} sellers in stage 1 and the relay price q,*

- *Case 1, when $q \geq g$ (i.e., $q \in \Gamma_0$), $x^* = 0$ is the ESS and no one buys;*
- *Case 2, when $\frac{c}{N-N_{su}-1} \leq q < g$ (i.e., $q \in \Gamma_{\underline{j}} \cup \Gamma_{\underline{j}+1} \cup ... \cup \Gamma_{N-N_{su}-1}$), for $x \in (0,1)$, $f(x) = 0$ has a single root $\breve{x}_f$, which is the ESS, i.e., $x^* = \breve{x}_f$;*
- *Case 3, when $\frac{c}{N-N_{su}} \leq q < \frac{c}{N-N_{su}-1}$ (i.e., $q \in \Gamma_{N-N_{su}}$), $x^* = 1$ is the ESS and all unsuccessful users buy;*
- *Case 4, when $q < \frac{c}{N-N_{su}}$ (i.e., $q \in \Gamma_{N-N_{su}+1} \cup \Gamma_{N-N_{su}+2} \cup ...$), there are not sufficient buyers, and the game ends.*

Note that in the above discussion and Theorem 3.1, we assume that $\frac{c}{N-N_{su}} < g$. When $\frac{c}{N-N_{su}} \geq g$, if $q \geq \frac{c}{N-N_s} \geq g$, following the discussion in Case 1, $x^* = 0$ and no one buys; while if $q < \frac{c}{N-N_s}$, following the discussion in Case 4, there are not sufficient buyers to afford the relay service. Therefore, with $\frac{c}{N-N_{su}} \geq g$, the game will end with the relay portion throughput being zero.

Fig. 3.5 shows an example of ESS x^* at different price q and with different number $(N - N_{su})$ of unsuccessful users. The total number of users in the network is $N = 12$. The cost to forward one segment is $c = 2$ and the gain of correctly receiving one segment is $g = 1$. We first study x^* at different price with a fixed number of unsuccessful users and use $(N - N_{su}) = 5$ as an example. We observe that when $q < \frac{c}{N-N_{su}} = 0.4$ (i.e., in price ranges Γ_6, Γ_7...), the number of buyers is not sufficient and the game ends with no relay service. If $q \in [0.4, 1)$, we observe that at a lower price, the game requires more buyers to pay the relay service, and thus, unsuccessful users have a smaller probability to free ride with a larger x^*. For $q \geq g = 1$, no user buys and the ESS is $x^* = 0$. Fig. 3.5 also shows the ESS with a different number of unsuccessful users, $(N - N_{su}) = 8$. It can be seen that for a given price q in the price range Γ_5 to Γ_3, more unsuccessful users

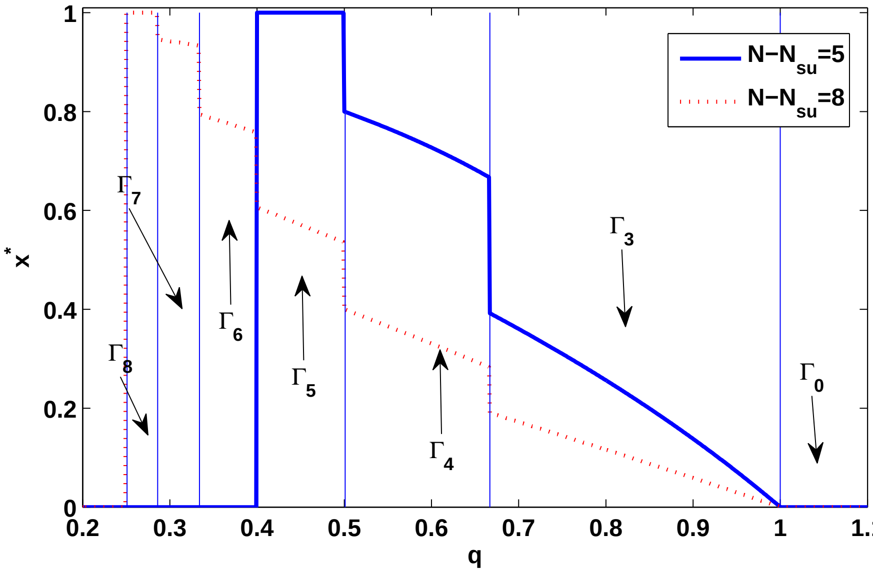


Fig. 3.5. An example of x^*.

give a smaller ESS x^*. This is because when the number of unsuccessful users is large, each unsuccessful user expects other unsuccessful users to purchase the relay service and he/she has a higher tendency to free ride.

3.2.4 Price Setting Game and Throughput Optimization

From the previous discussion, at different price q selected by the local agent, we may have different ESS, and thus different relay portion throughput. Therefore, x^* is a function of q.[8] In the following, we will analyze the optimal price that maximizes the system throughput.

Given $(N - N_{su})$ unsuccessful users and price q, each unsuccessful user follows the ESS x^* to play the buyers' game, and the relay portion throughput is

$$T_R(x^*|N_{su}) = \sum_{k=\lceil c/q \rceil}^{N-N_{su}} \binom{N-N_{su}}{k} (x^*)^k (1-x^*)^{N-N_{su}-k}, \tag{3.7}$$

where the summation term denotes the probability that there are sufficient buyers to afford the relay service. The local agent aims to find the optimal q^* that can

[8] For presentation simplicity, we use x^* to represent $x^*(q)$ in the sequel.

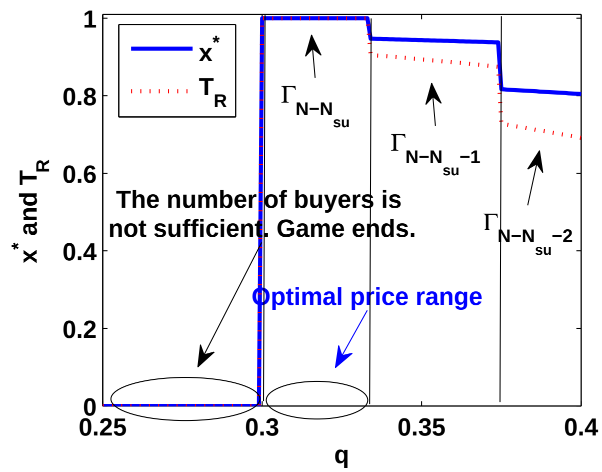


Fig. 3.6. An example of x^* and T_R.

maximize $T_R(x^*|N_{su})$,

$$q^* = \arg\max_q T_R(x^*|N_{su}). \tag{3.8}$$

From Theorem 3.1, when $q \geq g$, $x^* = 0$ and $T_R(x^*|N_{su}) = 0$. When $q < \frac{c}{N-N_{su}}$, the number of buyers is insufficient and the game ends also with zero relay portion throughput. Therefore, the optimal price q^* is in the range $\left[\frac{c}{N-N_{su}}, g\right)$.

From Theorem 3.1, when $\frac{c}{N-N_{su}} \leq q < g$, there are two possible ESS, $x^* = 1$ when $\frac{c}{N-N_{su}} \leq q < \frac{c}{N-N_{su}-1}$ and $x^* \in (0,1)$ when $\frac{c}{N-N_{su}-1} \leq q < g$. Comparing the relay portion throughput when $x^* = 1$ and $x^* \in (0,1)$, we have

$$T_R(x^*|N_{su})|_{x^*=1} = \sum_{k=\lceil c/q \rceil}^{N-N_{su}} \binom{N-N_{su}}{k} 1^k (1-1)^{N-N_{su}-k} = 1, \tag{3.9}$$

and

$$\begin{aligned} T_R(x^*|N_{su})|_{x^* \in (0,1)} &= \sum_{k=\lceil c/q \rceil}^{N-N_{su}} \binom{N-N_{su}}{k} (x^*)^k (1-x^*)^{N-N_{su}-k} \\ &< \sum_{k=0}^{N-N_{su}} \binom{N-N_{su}}{k} (x^*)^k (1-x^*)^{N-N_{su}-k} = 1. \end{aligned} \tag{3.10}$$

Therefore, $x^* = 1$ gives a higher relay portion throughput, and the optimal price q^* should be in the price range $\Gamma_{N-N_{su}}$ that gives $x^* = 1$.

Fig. 3.6 shows an example of x^* and $T_R(x^*|N_{su})$ at different price q with $N - N_{su} = 10$, $c = 3$, and $g = 1$. Following the previous discussion, the optimal price should lie in $\Gamma_{N-N_{su}}$, which corresponds to $[0.30, 0.333)$ in Fig. 3.6. In this price range, we observe that $x^* = 1$, and $T_R(x^*|N_{su}) = 1$, which is the maximum relay portion throughout. In this scenario, all unsuccessful users have to buy together to afford the relay service. Any free riding behavior will result in the failure of purchasing a relay, which lowers the utilities of all unsuccessful users. Thus, the optimal price drives the buyers' game to the equilibrium where all unsuccessful users buy and share the cost to afford the relay service.

3.2.5 Equilibrium of the Stackelberg Game

To summarize, the SPNE of the Stackelberg game is

- Stage 1, the sellers' game. S is a weakly dominant strategy, and all successful users decide to be sellers.

- Stage 2, the price setting game. Given $(N - N_{su})$ unsuccessful users, if $\frac{c}{N-N_{su}} < g$, the optimal price can be any value in the range $\Gamma_{N-N_{su}}$. Otherwise, the game ends.

- Stage 3, the buyers' game. At the optimal price q^* selected in Stage 2, all unsuccessful users decide to be buyers with probability $x^* = 1$.

- Stage 4, the transaction game. Since under the optimal price q^*, all unsuccessful users are buyers and $q^*(N - N_{su}) \geq c$, being a relay is profitable, and the selected seller forwards the segment, and gets payment $q^*(N - N_{su})$.

At this SPNE, the relay portion throughput is 1 if $\frac{c}{N-N_{su}} < g$, and 0 otherwise.

3.3 Wireless Multicast With Heterogeneous Users

In wireless multicast, users may use different mobile devices, whose cost to forward a segment are different. For example, the cost to forward a segment using a smart phone is much higher than that when using a laptop due to the limited power available. In this section, we will study cooperative wireless multicast with heterogenous users, who have different cost of forwarding a segment.

3.3.1 Game Model for Heterogeneous Users

In this work, we assume that each user's cost is his/her private information, which is independent and identically distributed following the same distribution $\phi_c(c)$ in the range $[c_l, c_h]$, where c_l and c_h are the lower and upper bounds of a user's cost. Thus, successful users may request different payments to forward a segment. Since the cost is their private information, they may lie to others if cheating can help improve their own utilities. For example, successful users may claim high cost so that to ask high payments for providing the relay service. However, if the asked payments are too high, unsuccessful users may not be able to afford, and thus, the system efficiency may be reduced. To encourage successful users to tell their true cost, we use the second-price sealed-bid auction, which is a truth-telling auction [103]. To help readers have the whole picture of the system, the detailed game model is illustrated below.

Stage 1: The Sellers' Auction Game. After the broadcast portion, assume that there are N_{su} successful users. Each of them decides whether to sell relay service. Let $\{S, NS\}$ denote their strategy set, including being a seller (S) and not being a seller (NS). Assume that $N_s(\leq N_{su})$ successful users decide to be sellers. They will enter the auction game, where a seller, say user i, submits to the local agent his/her bid including his/her ID and the payment d_i he/she asks for. In our work, messages from the sellers are encrypted and then sent simultaneously using CDMA technology in the selling part in Fig. 3.2. So only the local agent can decrypt them, which can prevent others from overhearing the transmission and avoid potential leak of their bidding information.

Stage 2: The Price Setting Game. The local agent, as the host of the auction, decides the winner of the auction, and the winning bid, denoted by d^w, that the winner will get paid after relaying the segment. Following the second-price sealed-bid protocol, the winner is the seller with the lowest bid, and d^w is the second lowest bid or the reserve bid $d^r(N_s)$, whichever is less. Here $d^r(N_s)$ is a function of N_s and denotes the highest payment that buyers will accept. From the discussion in Chapter 3.2.3, no user will buy if $q \geq g$. Here q is still the relay price that will be charged to each buyer, and g is still the utility gain of correctly receiving a segment. Given that there are N_s sellers, the number of buyers is no more than $(N - N_s)$, and the highest total payment from all buyers is less than $(N - N_s)g$. Thus, the reserve bid should satisfy $d^r(N_s) < (N - N_s)g$. In this work, we set $d^r(N_s) = (N - N_s)g - \epsilon$, where ϵ is an arbitrarily small positive number. Note that if the bids of all the sellers are larger than $d^r(N_s)$, there is no winner of this auction game, and the game ends with no relay service.

The local agent also selects a relay price q that will be charged to each buyer for the relay service, and then announces (in the price announcement part in Fig. 3.2) to all users the number N_s of sellers, the winning bidder's ID, the winning bid d^w, and the relay price q. The local agent selects the optimal relay price q to maximize the system throughput. As will be shown in our analysis in Chapter 3.3.2.3, the optimal price q depends on the number of users who do not participate in the auction game (i.e., $(N - N_s)$) and users' cost distribution $\phi_c(c)$.

Stage 3: The Buyers' Game. Based on information announced by the local agent, unsuccessful users decide whether to be buyers. All buyers broadcast their IDs simultaneously in the buying part in Fig. 3.2. The winning bidder listens to buyers' messages, and knows the number N_b of buyers and the total payment $N_b q$ the buyers offer.

Stage 4: The Transaction Game. For the winning bidder, say user i, if the winning bid d^w is not less than his/her cost c_i, user i is willing to forward the segment. However, the local agent pays user i only when the total payment $N_b q$ from the buyers is not less than the winning bid d^w and user i forwards the segment. Thus, user i will forward the segment if $N_b q \geq d^w \geq c_i$, and the local agent charges q from each buyer and pays d^w to user i. Otherwise, user i will not relay the segment and the game ends.

Note that after the transaction, there might be extra unused payment of $(N_b q - d^w)$. The local agent will keep this unused payment and accumulate it from each segment transmission. Once the accumulated amount is larger than the winning bid d^w in one round, the local agent uses it to pay the relay service and all unsuccessful users enjoy a free segment forwarding in that round. In this work, we consider the scenario where users frequently join and leave the multicast service. They may leave before the next free relay service. Thus, we ignore the impact of free relay service on buyers' utility in our analysis.

3.3.2 SPNE Analysis

Note that the transaction game ensures that the relay will always receive a non-negative net utility gain in the game. This result can simplify the sellers' auction game. Thus, similar to the analysis in Chapter 3.2, next we first study the sellers' auction game, followed by the buyers' game and the price setting game.

3.3.2.1 The Sellers' Auction Game

In this stage, each successful user decides whether to bid and how to bid in the auction game. Note that in this stage, the reserve bid $d^r(N_s)$ is unknown (as the number N_s of sellers is unknown), and a successful user's decision will also affect $d^r(N_s)$, which should be taken into consideration when choosing his/her strategy. For a successful user i, we have the following proposition.

Proposion 3.1. *For a successful user i with cost c_i to forward a segment, if $c_i > d^r(1)$, he/she should not enter the auction game. Otherwise, he/she should participate in the auction game and bid $d_i = c_i$. This is a weakly dominant strategy.*

Proof: We first show that a successful user, say user i, should not enter the auction game if $c_i > d^r(1)$, and he/she should participate in the game otherwise.

Note that $d^r(1)$ is the highest possible reserve bid (i.e., the reserve bid when there is only one successful user), and thus, it is the highest payment a seller can receive. If user i's cost c_i is larger than $d^r(1)$, he/she cannot benefit from serving as a relay, and thus should not enter the auction game. When $c_i \leq d^r(1)$, if user i takes strategy NS and does not bid, his/her utility in the relay portion is 0. However, if he/she participates in the auction, he/she has a positive probability to win the auction and make a non-negative net utility gain by forwarding the segment. Thus, using strategy S, his/her expected utility in the relay portion is non-negative, and S is a weakly dominant strategy over NS. Therefore, user i should enter the auction game as long as $c_i \leq d^r(1)$.

In the following, we prove that when the successful user i decides to participate in the auction game, bidding his/her real cost is a weakly dominant strategy. To illustrate this, we define $\hat{d}_i = \min_{j \neq i} d_j$ as the smallest bid excluding d_i, and consider two different scenarios.

- $c_i \leq \min(d^r(N_s), \hat{d}_i)$**:** In this case, user i can win the auction by bidding any value in the range $\left(0, \min(d^r(N_s), \hat{d})\right]$, and this range includes his/her true cost c_i. Then, his/her winning bid is $d^w = \min(d^r(N_s), \hat{d}_i)$. Thus, he/she has the chance to make a net utility gain of $d^w - c_i \geq 0$. However, if he/she bids a price higher than $\min(d^r(N_s), \hat{d})$, he/she cannot win the auction, and his/her utility in the relay portion is zero. Thus, $d_i = c_i$ is a weakly dominant strategy.
- $c_i > \min(d^r(N_s), \hat{d}_i)$**:** In this case, user i cannot win the auction by bidding any value in the range $\left(\min(d^r(N_s), \hat{d}_i), +\infty\right)$, and this range includes his/her true cost c_i. Then, his/her payoff in the relay portion is zero. On the other hand, if he/she bids $d_i \leq \min(d^r(N_s), \hat{d}_i)$, he/she can win the auction, but will not relay, as the payment for relaying is $d^w = \min(d^r(N_s), \hat{d}_i)$ which is less than his/her cost c_i. Thus, his/her payoff is also zero. Therefore, in this scenario, for any d_i, user i's payoff is zero.

From the above discussions, the successful user i should bid $d_i = c_i$ if $c_i \leq d^r(1)$, and it is a weakly dominant strategy. ■

From Proposition 3.1, each seller bids his/her real cost in the auction game. Thus, if there is a winner of the auction (i.e., the lowest bid is no larger than the reserver bid, $d^r(N_s)$), the winning bidder is the bidder who has the lowest cost among all bidders, and the winning bid d^w is no less than the winner's cost. Thus, in the transaction game, as long as $N_b q \geq d^w$, the winning bidder will forward the segment.

3.3.2.2 The Buyers' Game

In the buyers' game, based on the number N_s of bidders, the winning bid d^w, and the price q determined by the local agent, each unsuccessful user decides whether to be a buyer. Similar to the game with homogeneous users, we also model unsuccessful users' interaction as an evolutionary game, while the analysis is more complicated. This is because given N_s bidders, the rest $(N - N_s)$ users include unsuccessful users who may or may not purchase the relay service as well as successful users who do not bid in the sellers' auction game. Unsuccessful users should take this into consideration when choosing their strategies. Specifically, given N_s bidders, for unsuccessful user i, let $\mathcal{Y}_{-i}$ denote the set of all other unsuccessful users and successful users who do not bid in the sellers' auction game. So $|\mathcal{Y}_{-i}| = l \triangleq N - N_s - 1$. Recall that a successful user will bid if his/her cost is below $d^r(1)$. Since each user's cost follows the probability distribution $\phi_c(c)$, the probability that a successful user does not bid is $1 - \Phi_c(d^r(1))$, where $\Phi_c(\cdot)$ is the cumulative distribution function of ϕ_c. Then for each user in $\mathcal{Y}_{-i}$, the probability that he/she is a successful user is

$$p_s = \frac{p_1\Big[1 - \Phi_c(d^r(1))\Big]}{p_1\Big[1 - \Phi_c(d^r(1))\Big] + (1 - p_1)}. \tag{3.11}$$

Let n be the number of unsuccessful users in $\mathcal{Y}_{-i}$, and it follows Binomial distribution $B(l, 1 - p_s)$.

Based on the above discussion on the number of unsuccessful users, following a similar analysis in Chapter 3.2.3, for each unsuccessful user i, we first derive his/her average utility $\bar{V}_B(x)$ and $\bar{V}_{NB}(x)$ by using the strategy B and NB, respectively, and then find the ESS x^*.[9]

[9] Similar to the homogeneous case, the evolutionary game in the heterogenous case is also a one-shot game with complete information, where each user's gain g, the probability distribution function of the number of unsuccessful users, and the winning bid d^w are all public information. Thus, we can directly derive the ESS by solving (3.1).

Recall that each unsuccessful user purchases the relay service with probability x. Thus, given n unsuccessful users in $\mathcal{Y}_{-i}$, the conditional number k of buyers in $\mathcal{Y}_{-i}$ follows Binomial distribution $B(n, x)$. In this context, if user i decides to be a buyer, the total number of buyers is $k+1$, and thus, the total payment from all buyers is $(k+1)q$. If $(k+1)q \geq d^w$, the winning bidder forwards the segment, and user i receives the segment correctly and pays the price q; otherwise, there is no relay service and user i's utility in the relay portion is 0. Therefore, in the relay portion, user i's average utility of strategy B is

$$\begin{aligned} \bar{V}_B(x) &= (g-q)\sum_{n=0}^{l}\binom{l}{n}(1-p_s)^n p_s^{l-n} \\ &\quad \times \left\{\sum_{k=0}^{n}\binom{n}{k}x^k(1-x)^{n-k}I[(k+1)q \geq d^w]\right\}. \end{aligned} \quad (3.12)$$

Similarly, if user i chooses the strategy NB, its average utility is

$$\begin{aligned} \bar{V}_{NB}(x) &= g\sum_{n=0}^{l}\binom{l}{n}(1-p_s)^n p_s^{l-n} \\ &\quad \times \left\{\sum_{k=0}^{n}\binom{n}{k}x^k(1-x)^{n-k}I[kq \geq d^w]\right\}. \end{aligned} \quad (3.13)$$

Let $h(x) = \bar{V}_B(x) - \bar{V}_{NB}(x)$, and Appendix B shows that

$$\begin{aligned} h(x) &= g\binom{l}{k^*}[x(1-p_s)]^{k^*}[1-x(1-p_s)]^{(l-k^*)} \\ &\quad -q\sum_{k=k^*}^{l}\binom{l}{k}[x(1-p_s)]^k[1-x(1-p_s)]^{(l-k)} \\ &= f\Big((1-p_s)x\Big). \end{aligned} \quad (3.14)$$

In (3.14), $k^* = \lceil d^w/q \rceil - 1$ where $\lceil d^w/q \rceil$ is the minimum number of buyers required to afford the relay service, and $f(x)$ is defined in (3.4).

Similar to the analysis in Chapter 3.2.3, at the stable state x, we have $\dot{x} = \eta x(1-x)h(x) = 0$. Thus, we have three possible solutions: $x = 0$, $x = 1$, and x satisfies $h(x) = 0$, all of which can be ESS x^*. To study the ESS x^* at different prices, similar to the analysis in Chapter 3.2.3.2, we let $\underline{j} = \lfloor \frac{d^w}{g} \rfloor + 1$ and partition the whole price range $[0, +\infty)$ into subintervals

$$\Gamma_0 = [g, +\infty), \Gamma_{\underline{j}} = \left[\frac{d^w}{\underline{j}}, g\right), \text{ and } \Gamma_j = \left[\frac{d^w}{j}, \frac{d^w}{j-1}\right) \text{ for } j > \underline{j}. \tag{3.15}$$

In price range Γ_j with $j \geq \underline{j}$, at least j buyers are required to afford the relay service.

• Case 1, $q \in \Gamma_0$, i.e., $q \geq g$: Similar to the analysis in the game with homogeneous users, no user buys and $x^* = 0$ is the ESS.

• Case 2, $q \in \Gamma_j$ with $\underline{j} \leq j \leq N - N_s - 1$: Similar to Chapter 3.2.3.2, we study the ESS when $j = \underline{j} = 1$ (which happens only when $d^w < g$) and when $1 < j < N - N_s$, separately.

When $j = \underline{j} = 1$, i.e, $q \in [d^w, g)$, we have $k^* = \lceil d^w/q \rceil - 1 = 0$. From the analysis in Chapter 3.2.3.2, when $k^* = 0$, $f(x)$ in (3.4) is a decreasing function of x for $0 \leq x \leq 1$, and $f(x) = 0$ has a single root $\breve{x}_f \in (0, 1)$. From $h(x) = f\Big((1 - p_s)x\Big)$, $h(x)$ is a decreasing function of x in the interval $\left(0, \frac{1}{1-p_s}\right)$, and $h(x) = 0$ has a single root in the interval $\left(0, \frac{1}{1-p_s}\right)$, given as $\breve{x}_h = \frac{\breve{x}_f}{1-p_s}$. When $\breve{x}_f \geq 1 - p_s$, or equivalently, $h(1) = f(1 - p_s) \geq 0$, as shown in Fig. 3.7a, we have $h(x) \geq 0$ for all $0 \leq x \leq 1$, and strategy B gives a higher utility than strategy NB. So all unsuccessful users buy with $x^* = 1$ being the ESS. When $\breve{x}_f < 1 - p_s$, that is, $h(1) = f(1 - p_s) < 0$, $\breve{x}_h$ is in the interval $(0, 1)$. Similar to the homogeneous user case, $x^* = \breve{x}_h$ is the ESS, as shown in Fig. 3.7b.

When $q \in \Gamma_j$ with $1 < j < N - N_s$, we have $k^* = \lceil d^w/q \rceil - 1 > 0$. From the analysis in Chapter 3.2.3.2, when $k^* > 0$, $f'(x) = 0$ has a single root $\tilde{x}_f \in (0, 1)$, where $f'(x) > 0$ when $0 < x < \tilde{x}_f$, and $f'(x) < 0$ when $x > \tilde{x}_f$. Also, $f(x) = 0$ has a single root $\breve{x}_f$ in the interval $(\tilde{x}_f, 1)$. Note that $h'(x) = (1 - p_s) f'\Big((1 - p_s)x\Big)$. Therefore, $h'(x) = 0$ has a single root $\tilde{x}_h = \frac{\tilde{x}_f}{1-p_s}$ in the interval $(0, \frac{1}{1-p_s})$, $h(x)$ is an increasing function of x when $0 < x < \tilde{x}_h$, $h(x)$ is a decreasing function when $\tilde{x}_h < x < \frac{1}{1-p_s}$, and $h(x) = 0$ has a single root $\breve{x}_h = \frac{\breve{x}_f}{1-p_s}$ in the interval $(\tilde{x}_h, \frac{1}{1-p_s})$. If $\breve{x}_f \geq 1 - p_s$, or equivalently, $h(1) = f(1 - p_s) \geq 0$, as shown in Fig. 3.8a, $h(x) \geq 0$ for all $0 \leq x \leq 1$ and $x^* = 1$ is the ESS since strategy B always

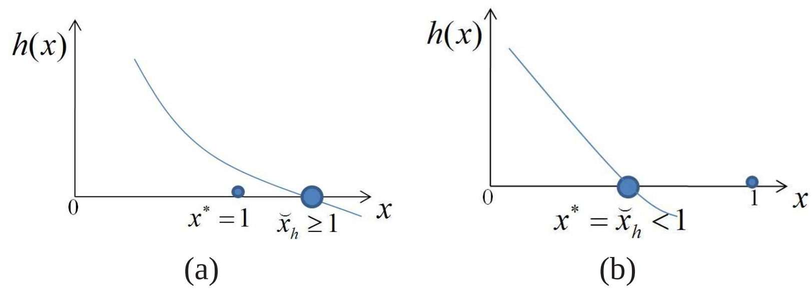


Fig. 3.7. $q \in \Gamma_1$. (a): $h(1) > 0$, (b): $h(1) < 0$.

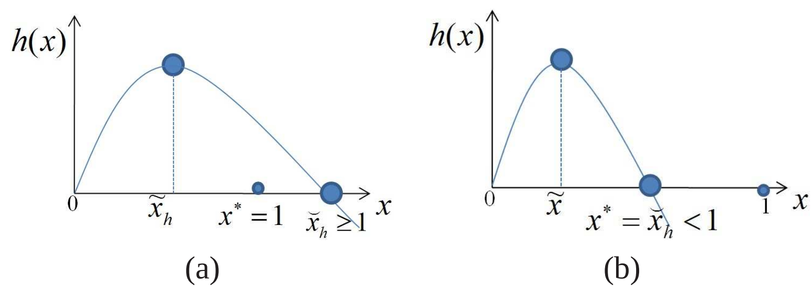


Fig. 3.8. $q \in \Gamma_j$, where $1 < j < N - N_s$. (a): $h(1) > 0$, (b): $h(1) < 0$.

outperforms strategy NB. If $\breve{x}_f < 1 - p_s$, or equivalently, $h(1) = f(1 - p_s) < 0$, as shown in Fig. 3.8b, $\breve{x}_h$ is in the interval $(0, 1)$. So similar to the homogeneous user case, $x^* = \breve{x}_h$ is the ESS.

• Case 3, $q \in \Gamma_{N-N_s} = [\frac{d^w}{N-N_s}, \frac{d^w}{N-N_s-1})$: In this price range, at least $(N - N_s)$ buyers are required to afford the relay service, while there are at most $(N - N_s)$ possible buyers. Therefore, there is no chance to free ride, and all unsuccessful users will buy with $x^* = 1$. Mathematically, when $q \in \Gamma_{N-N_s}$, $h(x) = f\Big(x(1 - p_s)\Big) = (g - q)[x(1 - p_s)]^l \geq 0$ for all $x \in [0, 1]$. Therefore, the strategy B always outperforms strategy NB, and $x^* = 1$.

• Case 4, $q \in \Gamma_j$ with $j \geq N - N_s + 1$: At least $(N - N_s + 1)$ buyers are required to afford the relay service. However, the number of total potential buyers is no more than $(N - N_s)$. Thus, there are not sufficient buyers to afford the relay service, and the game ends.

In summary, we have the following theorem.

Theorem 3.2. *Given N_s bidders in Stage 1, the winning bid d^w, and the relay price*

q,

- *Case 1, when* $q \geq g$ *(i.e.,* $q \in \Gamma_0$*),* $x^* = 0$ *is the ESS, and no unsuccessful user buys;*

- *Case 2, when* $\frac{d^w}{N-N_s-1} \leq q < g$ *(i.e.,* $q \in \Gamma_{\underline{j}} \cup \Gamma_{\underline{j}+1} \cup ... \cup \Gamma_{N-N_s-1}$*), if* $h(1) \geq 0$*,* $x^* = 1$ *is the ESS and all unsuccessful users buy. If* $h(1) < 0$*, for* $x \in (0,1)$*,* $h(x) = 0$ *has a single root* $0 < \breve{x}_h < 1$*, which is the ESS,* $x^* = \breve{x}_h$*;*

- *Case 3, when* $\frac{d^w}{N-N_s} \leq q < \frac{d^w}{N-N_s-1}$ *(i.e.,* $q \in \Gamma_{N-N_s}$*),* $x^* = 1$ *is the ESS and all unsuccessful users buy;*

- *Case 4, when* $q < \frac{d^w}{N-N_s}$ *(i.e.,* $q \in \Gamma_{N-N_s+1} \cup \Gamma_{N-N_s+2} \cup ...$*), there are not sufficient buyers to afford the relay service, and the game ends.*

3.3.2.3 The Price Setting Game

In the price setting game, the local agent finds the optimal price q^* to maximize the relay portion throughput. Given the number N_s of bidders and the winning bid d^w, with the ESS x^* (which is a function of relay price q) from Theorem 3.2, the relay portion throughput is

$$T_R(x^*|N_s, d^w) = \sum_{n=0}^{N-N_s} \binom{N-N_s}{n} (1-p_s)^n p_s^{(N-N_s-n)} \times \left\{ \sum_{k=\lceil d^w/q \rceil}^{n} \binom{n}{k} (x^*)^k (1-x^*)^{n-k} \right\}. \quad (3.16)$$

To maximize $T_R(x^*|N_s, d^w)$, based on Theorem 3.2, the optimal price q^* should be in the range $\Gamma_{\underline{j}} \cup \Gamma_{\underline{j}+1} \cup \cdots \cup \Gamma_{N-N_s}$ where $x^* > 0$.

Fig. 3.9 shows an example of x^* and the corresponding throughput T_R where there are a total of $N = 12$ users in the system, and $N_s = 2$ of them bid in Stage 1 with the winning bid $d^w = 2$. For each user, the probability to correctly receive the segment in the broadcast portion is $p_1 = 0.4$, and the gain of correctly receiving the segment is $g = 1$. Users' cost $\{c_i\}$ is uniformly distributed in the range $[c_l =$

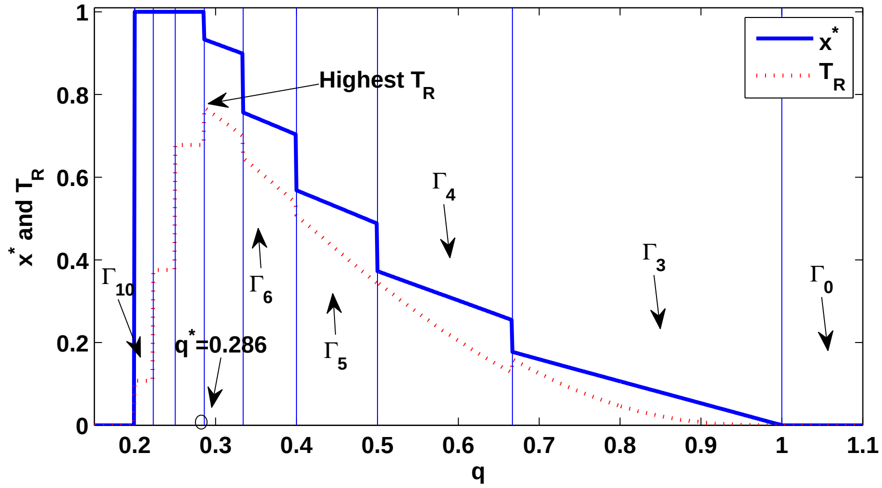


Fig. 3.9. An example of x^* and T_R.

$1, c_h = 17]$. From (3.11), the probability that a user who does not bid is a successful user is $p_s = 0.2$. From Fig. 3.9, when $q < d^w/(N - N_s) = 0.2$ (i.e., in price ranges $\Gamma_{11}, \Gamma_{12}, ...$), there are not sufficient buyers and the relay portion throughput is zero. When $q \in [0.2, 1.0)$ (i.e., in price ranges $\Gamma_{10}, ..., \Gamma_3$), we observe that at a low price (e.g., in Γ_{10}), the game requires a large number of buyers to afford the relay service. However, since there are some successful users who do not bid in the sellers' auction game, even if all unsuccessful users buy with $x^* = 1$, it is still possible that the total payment $N_b q$ is smaller than the winning bid d^w, and therefore, the relay portion throughput is small. At a high price, e.g., when $q \in \Gamma_3$, the minimum number of required buyers is small. Thus, unsuccessful users have a high tendency to free ride, and the probability that there are not sufficient buyers is high, which also results in a low relay portion throughput. Therefore, the optimal price should be appropriately selected to address this tradeoff, and T_R is maximized when $q = d^w/7 = 0.286$ in this example.

To efficiently find the optimal price, we have the following proposition, whose proof is in Appendix C.

Proposion 3.2. *In each price range Γ_j with $j \in \{\underline{j}, ..., N - N_s\}$, $T_R(x^*|N_s, d^w)$ is a non-increasing function of q.*

This can also be observed from the example in Fig. 3.9. In each price range

Γ_j with $3 \leq j \leq 10$, T_R is a non-increasing function of q and is maximized at the left boundary $q = d^w/j$. Based on this observation, we propose Algorithm 1 to efficiently find the global optimal price that maximizes the relay portion throughput. Specifically, Algorithm 1 compares $T_R(x^*|N_s, d^w)$ at $q = d^w/j$ when $j = \underline{j}, \cdots, N-N_s$, and chooses the optimal price q^* that maximizes $T_R(x^*|N_s, d^w)$.

Algorithm 1: Optimal Price Selection

1: $T_R^* = 0$, $q^* = 0$
2: **for** $j = \underline{j}$ to $(N - N_s)$ **do**
3: Set $q = d^w/j$ and use Theorem 3.2 to find x^*, and use (3.16) to calculate $T_R(x^*|N_s, d^w)$
4: **if** $T_R(x^*|N_s, d^w) > T_R^*$ **then**
5: $T_R^* = T_R(x^*|N_s, d^w)$ and $q^* = d^w/j$
6: **end if**
7: **end for**

We then discuss the properties of the optimal price q^*. We first study the ESS x^* at the optimal price q^* when $(N - N_s)$ takes different values. For the system in Fig. 3.9, we vary the value of $(N - N_s)$ from 5 to 11, and other parameters are the same as in Fig. 3.9. In Table 3.1, for different $(N - N_s)$, we list $(N - N_s)(1 - p_s)$, the average number of unsuccessful users among the $(N - N_s)$ users who do not bid, $\lceil d^w/q^* \rceil$, the minimal number of users required to afford the relay service, and the ESS x^* at the optimal price q^*. From Table 3.1, we observe that at the optimal price, $\lceil d^w/q^* \rceil$ has a similar value to $(N - N_s)(1 - p_s)$. It means that the optimal price is chosen carefully to let each unsuccessful user has a small or zero probability to free ride (i.e., x^* is close to 1 as shown in Table 3.1), while ensuring that the total payment from unsuccessful users is sufficient to pay the relay service.

We then study the optimal price q^* at different values of $(N-N_s)$ (the number of users who do not bid) and p_s (probability that a user who does not bid is a successful user). For the system in Fig. 3.9, we vary $(N - N_s)$ from 4 to 10, and vary p_1 so that p_s is 0.2, 0.4, or 0.6. Other parameters are the same. The optimal price q^* is shown in Fig. 3.10. We can see that given a fixed winning bid (i.e., $d^w = 2$ in the example), the optimal price increases when $(N - N_s)$ decreases. This is because,

TABLE 3.1

x^* AT THE OPTIMAL PRICE WITH DIFFERENT NUMBER N_s OF BIDDERS.

$(N-N_s)$	5	6	7	8	9	10	11
$(N-N_s)(1-p_s)$	4	4.8	5.6	6.4	7.2	8	8.8
$\lceil d^w/q^* \rceil$	4	5	5	6	7	7	8
x^*	1.0	1.0	0.90	0.99	1.0	0.93	0.99

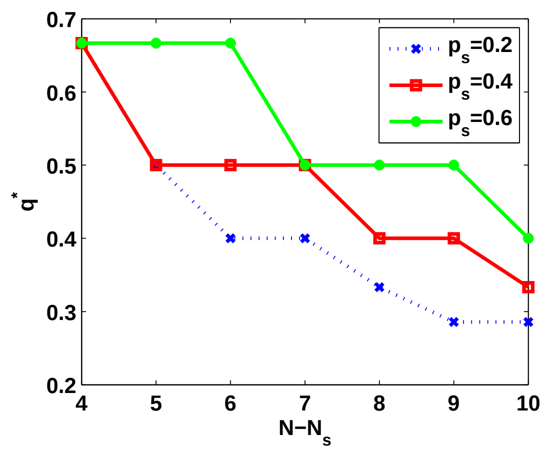


Fig. 3.10. An example of the optimal price q^*.

with a smaller $(N-N_s)$ and thus potentially fewer unsuccessful users, each buyer needs to pay more to purchase the relay service. Similarly, with a larger p_s, there are fewer unsuccessful users, and each buyer also has to pay a higher price to purchase the relay service. Note that, in Fig. 3.10, when $(N-N_s)$ and p_s vary, q^* takes values from a common finite set. This is because from Proposition 3.2, the optimal price can only take values in the finite set $\left\{\frac{d^w}{\underline{j}}, \frac{d^w}{\underline{j}+1}, \cdots, \frac{d^w}{N-N_s}\right\}$.

3.3.2.4 SPNE of the Stackelberg Game with Heterogeneous Users

To summarize, the SPNE of the multi-buyer multi-seller game with heterogeneous users is:

- Stage 1, the sellers' auction game. Successful users whose cost is no larger than $d^r(1)$ will enter the auction game, and bid their true cost.
- Stage 2, the price setting game. The local agent follows the second-price

TABLE 3.2

OVERALL THROUGHPUT COMPARISON WITH AND WITHOUT PROPOSED INCENTIVE MECHANISM.

λ	0.5	0.67	1
Incentive	81%	95%	99%
No incentive	37%	37%	37%

sealed-bid auction protocol to select the winner of the auction and the winning bid d^w. Then the local agent uses Algorithm 1 to select the optimal price q^*.

- Stage 3, the buyers' game. Based on the number N_s of sellers, the winning bid d^w, and the price q^* selected by the local agent, each unsuccessful user follows Theorem 3.2 to find the ESS x^*, and decides to be a buyer with probability x^*.

- Stage 4, the transaction game. Given the number N_b of buyers, if $q^* N_b \geq d^w$, the auction winner relays the segment, and receives a payment of d^w. Otherwise, there is no relay and the game ends.

3.4 Simulation Results

In our simulation, we consider a multicast network with a BS and a group of users who dynamically join and leave the multicast service. For each user, the probability p_1 of receiving a segment correctly from the BS is 0.37, and the utility gain of receiving a segment correctly is $g = 1$. The initial number of users is 10. Users join the multicast service according to a Poisson process with an average arrival rate of λ users per segment duration (i.e., the length of the broadcast portion in Fig. 3.2). The sojourning period of each user in the system follows an exponential distribution with an average of μ segments. In our simulation, we fix $\mu = 20$ and test the system when $\lambda = 0.5$, 0.67 and 1, which correspond to the average network size of $N = 9.8$, 13.6 and 20.2, respectively.

Table 3.2 first compares the overall throughput T_O with and without our incentive mechanism. Here, the overall throughput T_O is defined as the average percentage of users who receive the segment correctly after the broadcast and the relay

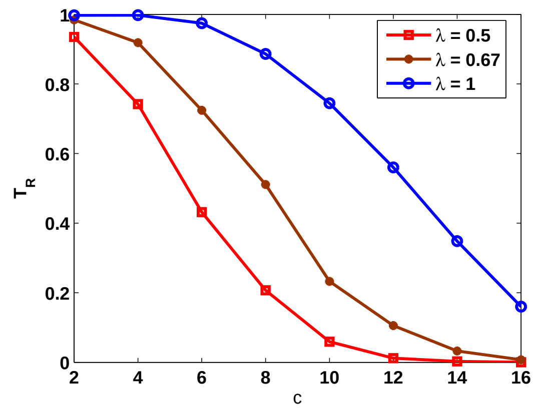


Fig. 3.11. Relay portion throughput with homogeneous users under different c and arrival rate λ.

portions. With our price based system, users follows the SPNE as discussed in Chapter 3.2.5. In the case without any incentive mechanisms applied, users who cooperate consumes their own transmission power only without receiving any reward. Thus, they do not have incentive to cooperate, and there is no relay in the relay portion. In this simulation, we fix the cost as 8, and test different λ. Table 3.2 shows the average results of 5000 segment transmission. We observe that with our payment based scheme, users will cooperate and relay segments, which can significantly improve the overall throughput, when compared with the case without incentive mechanisms, where there is not relay and the overall throughput is very low.

Fig. 3.11 shows the relay portion throughput with homogeneous users when we have different c and λ. For each segment transmission, users follow the SPNE as discussed in Chapter 3.2.5. Fig. 3.11 shows the average results for 5000 segment transmissions. From this figure, we observe that when c increases, the relay portion throughput decreases. This is because, when c increases, the probability that the condition $\frac{c}{N-N_{su}} < g$ is not satisfied increases. Recall that, if the condition $\frac{c}{N-N_{su}} < g$ is not satisfied, there is no relay service, as shown in Chapter 3.2.5. So a higher c leads to a lower relay portion throughput. In Fig. 3.11, we also observe that when λ increases, the relay portion throughput increases. This is because a larger λ will on average give a larger network size, and thus more unsuccessful

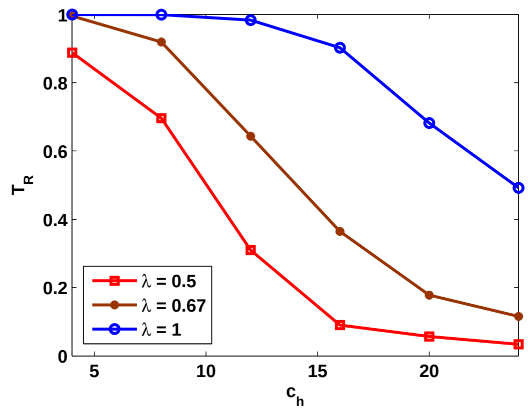


Fig. 3.12. Relay portion throughput with heterogeneous users under different c_h and arrival rate λ.

users. Then, the probability that the number of buyers is sufficient increases, which gives a higher relay portion throughput.

Fig. 3.12 shows the relay portion throughput with heterogeneous users with different c_h and λ. In this simulation, users' cost is uniformly distributed and randomly generated in $[c_l, c_h]$ with a fixed $c_l = 4$, and we test the system performance with different c_h. For each segment transmission, users follow the SPNE as discussed in Chapter 3.3.2.4, and we also test the system for 5000 segment transmissions. From this figure, we observe that when c_h increases, the relay portion throughput decreases. This is because when c_h increases, on average, each user has a higher cost to relay a segment. Therefore, the winning bid d^w increases, and it requires more buyers to afford the relay service. Thus, with a higher c_h, the probability that the number of buyers is sufficient decreases, causing a decrease in the relay portion throughput. Furthermore, we observe from Fig. 3.12 that when λ increases, the relay portion throughput increases. This is because a larger λ gives, on average, a larger network size N (thus a larger $d^r(1) = (N-1)g - \epsilon$) and more successful users (a larger N_{su}). From Proposition 3.1, a successful user will bid if its cost is lower than $d^r(1)$. So with a larger λ, more successful users will bid, resulting in a lower winning bid d^w. In addition, a larger λ gives on average more unsuccessful users, which, together with the fact that the winning bid d^w decreases, increases the probability that there are sufficient buyers to afford the relay service, and thus,

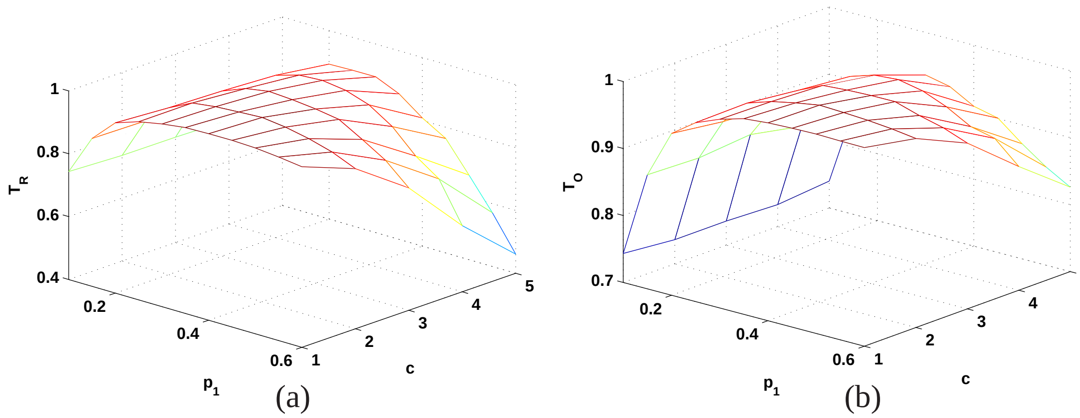


Fig. 3.13. System throughput with homogeneous users, when we have different p_1 and c. (a) The relay portion throughput. (b) The overall throughput.

increases the relay portion throughput.

In this payment based game, the numbers of buyers and sellers in each round are affected by the probability p_1. A larger p_1 will on average give a larger number of sellers but a smaller number of buyers, which may affect the system performance. Fig. 3.13 compares the relay portion and the overall throughput in homogeneous case when c and p_1 vary. The simulation setup is similar to that in Fig. 3.11. We fix $\lambda = 0.67$, and other parameters are the same. From this figure, when c is low, the relay portion and the overall throughput increases as p_1 increases. However, for a large c, for example, $c = 5$, when p_1 increases from 0.2 to 0.6, the relay portion throughput first increases and then decreases. This is because when p_1 is small, for example $p_1 = 0.2$, the probability that there are successful users who can provide relay service is small. As p_1 increases, the probability that successful users exist increases. Thus, the relay portion throughput increases, which also helps increase the overall throughput. However, as p_1 keeps increasing, the relay portion throughput starts to decrease. This is because when the cost is high, the required number of buyers is also large. When p_1 becomes large, there are fewer unsuccessful users. Thus, the probability that there are sufficient buyers decreases, and the relay portion throughput decreases. The overall throughput also decreases, which means that, for a larger p_1, the decrease in the relay portion throughput dominates the increase in the broadcast portion throughput. Therefore, if the BS consumes high power for a large p_1, it may not always be the optimal decision, especially when c is high. For a given range of p_1, the BS can search the optimal p_1^* in this range, which gives the maximal overall throughput while keeping the consumed transmission power of the BS low.

3.5 Summary

In a wireless multicast system, the cooperation among users can significantly improve the system performance. However, successful users may not be willing to help unsuccessful users, as forwarding costs their transmission power. In addition, due to the broadcast nature of wireless communications, unsuccessful users may prefer to free ride rather than buying the relay service. In this work, to stimulate user cooperation, we formulate the interaction among users as a multi-seller multi-buyer payment based game, where users pay to receive relay service and get paid if they forward their successfully received segments to others. In either homogeneous user case or heterogeneous user case, we derive the ESS of the buyers' game, and further derive the optimal price to maximize the relay portion throughput. It is shown that, under the optimal price, there is no chance for an unsuccessful user to free ride in the homogeneous user case, while there is a very small probability for an unsuccessful user to free ride in the heterogeneous user case. Therefore, our mechanisms have the merits of improving system efficiency and stimulating users to cooperate (i.e., successful users sell, and unsuccessful users buy).

Chapter 4

Incentive Analysis for Cooperative Interactive Multiview Video Streaming

From the literature survey in Chapter 2, there are very few works addressing the issue of user interaction with state change raised in an IMVS system, where users may switch views and change corresponding strategies frequently. Besides this issue, there are two more challenges in design incentive mechanisms for a high dimensional IMVS system. First, a small number of peers in a local area are likely watching different views among a large number of available views, making it difficult for a peer to find partners watching the exact same view to cooperate. To address this problem, following the frame structure discussed in Chapter 2.2.2.1, we use a DSC based multiview video coding structure to facilitate *cooperative view switching*, where users can help each other even if they are watching different views.

Second, since users switch views frequently and independently, it is hard for them to maintain partnership. In Chapter 2.3 we review the incentive mechanisms for single view video cooperative video streaming. The direct reciprocity schemes [11], [71]–[74] are not suitable for IMVS, since these schemes work only when users expect to interact with each other for a long time. The payment based game proposed in [82]–[84] works for the scenario where users change partners frequently. However, in such schemes, users make decisions based on short-term payoffs only, where a user will cooperate when the gain from others' payment is higher than the cost to help. In this chapter, similar to [90]–[93], we model user interaction as a reputation based indirect reciprocity game. This scheme also works for users who change partners frequently. We will show that in this scheme users make decisions taking future utility into consideration, and they may cooperate even if the short-term gain is lower than the cost. Therefore, this scheme is more effective in cooperation stimulation than the payment based scheme. To study how users' view switching and reputation updating affect their cooperation, we model users' action selection as an MDP. We then summarize the major contributions of this work as follows:

- To the best of our knowledge, this is the first work that provides theoretical analysis on how the multiview video affects user behavior in cooperative video streaming. In this work, we derive users' strategies in Nash Equilibria, and observe that users may cooperate at some views but not others. This is because peers can predict their future view navigation paths probabilistically, and thus, can estimate the probability that he/she needs others' help in the future. If a peer is at a view leading to a view-navigation path not requiring others' help, he/she also has less incentive to cooperate.
- We show that a large number of reputation levels provide higher incentive for user cooperation. We first observe that the 2-level reputation system is memoryless, and each user makes decisions only based on his/her expected short-term utility. Thus, if a user is at a view, where cooperation only results in a negative expected short-term utility, he/she will not cooperate. In the R-level reputation system with $R \geq 3$, a user needs to take his/her future utility into consideration, and may still cooperate even if the expected short-term payoff is negative. This is because cooperation helps him/her maintain a high reputation and get others' help in the future. If the future payoffs can compensate his/her current loss, he/she may still cooperate.
- We observe that the game may have multiple Nash Equilibria corresponding to different cooperation levels (e.g., users cooperate at all views in the full cooperation

equilibrium, while users only cooperate at certain views in the partial cooperation equilibrium.). The final equilibrium the game will converge to depends on the initial cooperation level of the game. To address this issue, We propose a PfC scheme at the beginning of the game to drive the game to the desired full cooperation equilibrium to improve the system efficiency.

- In addition, we also study the impact of user membership dynamics on user cooperation and system performance. From our theoretical analysis and simulations, we observe that as long as the percentage of new users is smaller than a predetermined threshold, full cooperation is a dominant strategy for all users, and they will all cooperate. Otherwise, user cooperation will be interrupted and the PfC scheme should be used to resume user cooperation.

4.1 System Model

In this work, we consider an IMVS system, where a scene is captured by a large one-dimensional array of M evenly spaced cameras. A server compresses video of each view into coding segments of K frames each, and provides IMVS service to a group of N users who are synchronized in playback time. Once a user selects a view, he/she remains in this view for one segment of K consecutive frames. At the end of this segment, he/she can switch to another view.

Based on this IMVS system, in the following of this section, we first describe an interaction model that captures users' view switching behavior, and a multiview video coding structure that facilitates cooperative view switching among peers. We then propose an indirect reciprocity game to stimulate user cooperation. Finally, we model users' optimal action selection as an MDP.

4.1.1 View Switching Model

Views are divided into two categories: *anchor views* and *normal views*. Suppose that there are n_a anchor views, which evenly divide normal views into (n_a+1) view sets of $n_n = (M - n_a)/(n_a + 1)$ views per set. When seeking interested views, a user first browses views coarsely through anchor views. Once he/she reaches an interested anchor view, he/she can switch to neighboring normal views to refine view selection. In this work, we assume that users switch interested views frequently. After finding an interested view and remaining for one segment, they will likely seek another interested view from the next segment. Thus, anchor views are more frequently selected (more popular) than normal views.

At each view, a user can only switch to his/her nearby anchor views with probability P_a, or nearby normal views with probability $(1 - P_a)$. Specifically, we model the view transition as a discrete time Markov chain, and construct a $M \times M$ transition matrix $\mathbf{T}$, where $\mathbf{T}(v, v')$ is the probability of a user selecting view v' in the next segment after viewing v, and it is defined as follows:

$$\mathbf{T}(v, v') = \begin{cases} P_a/|\mathcal{Z}_a| & \text{if } \ v' \in \mathcal{Z}_a, \\ (1 - P_a)/|\mathcal{Z}_n| & \text{if } \ v' \in \mathcal{Z}_n, \\ 0 & \text{otherwise}, \end{cases} \tag{4.1}$$

where $\mathcal{Z}_a$ is the set including v's nearby anchor views, and $\mathcal{Z}_n$ is the set including v's nearby normal views. Specifically, if v is an anchor view, $\mathcal{Z}_a$ includes v's left/right closest anchor views and v itself, and $\mathcal{Z}_n$ includes v's left/right adjacent normal view sets. If v is a normal view, $\mathcal{Z}_a$ only includes v's left/right closest anchor views, and $\mathcal{Z}_n$ is the normal view set where v belongs. Given the one-step transition matrix $\mathbf{T}$, the l-step transition matrix is $\mathbf{T}^l$ ($\mathbf{T}$ raised to the lth power), where $\mathbf{T}^l(v, v')$ is the probability to transit to view v' in l segments after viewing v. The steady state view probability distribution is $\mathbf{v}$ satisfying $\mathbf{vT} = \mathbf{v}$, where $\mathbf{v}(v)$ is the probability that a user is at view v at the steady state.

For example, for a $M = 3$ views with a single anchor view in the middle, the

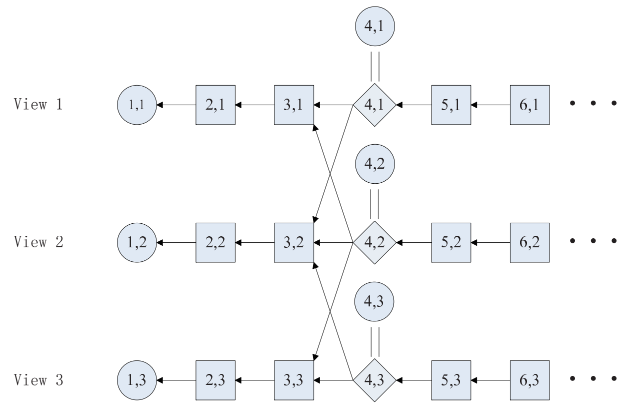


Fig. 4.1. Example of our multiview video coding structure for $M = 3$ views, segment size $K = 3$. Circles, squares and diamonds denote I-, P- and DSC frames, respectively. Each frame $F_{\tau,v}$ is labeled by its frame index τ and view v.

one-step and two-step transition matrices are

$$\mathbf{T} = \begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array}\!\!\begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{pmatrix} 1-P_a & P_a & 0 \\ (1-P_a)/2 & P_a & (1-P_a)/2 \\ 0 & P_a & (1-P_a) \end{pmatrix} \end{array} \quad \text{and}$$

$$\mathbf{T}^2 = \begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array}\!\!\begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{pmatrix} (1-P_a)(1-\frac{P_a}{2}) & P_a & P_a\frac{1-P_a}{2} \\ \frac{1-P_a}{2} & P_a & \frac{1-P_a}{2} \\ P_a\frac{1-P_a}{2} & P_a & (1-P_a)(1-\frac{P_a}{2}) \end{pmatrix} \end{array} \tag{4.2}$$

respectively, and the steady state view distribution is $\mathbf{v} = \left[\frac{1-P_a}{2}, P_a, \frac{1-P_a}{2}\right]$.

4.1.2 Multiview Video Coding Structure and Cooperative View Switching

To address the issue that users have difficulty in establishing partnership for cooperation in a high dimensional IMVS, we use a frame structure similar to that in [18]. It supports cooperative view switching, where users may cooperate with each other even they are in different views. Fig. 4.1 shows an example of the frame structure used in this work. Each view is encoded into segments of K frames. We encode the first segment using an intra-coded I-frame with $K-1$ trailing P-frames. For the next segment, for view switching we encode the first frame $F_{K+1,v}$ into two versions. The first version is an intra-coded I-frame, which can be decoded independently. The second version is a *DSC frame* [107]. To encode the DSC frame, we use the I-frame of the same frame as target, and use at most three decoded P-frames $F_{K,max(1,v-1)}, \ldots, F_{K,min(M,v+1)}$ as predictors (i.e., if the DSC frame is in view 1, it only has two predictors: $F_{K,1}$ and $F_{K,2}$. Similarly, if the DSC frame is in view M, it also only has two predictors: $F_{K,M-1}$ and $F_{K,M}$. If the DSC frame is in view v that is other than 1 and M, it has three predictors: $F_{K,v-1}$, $F_{K,v}$ and $F_{K,v+1}$). As long as one of the predictor frames is available at the decoder buffer, the DSC frame can be correctly decoded, and the decoded frame is bit-by-bit equivalent to the frame decoded from the I-frame. Frame $F_{K+1,v}$ is followed by $K-1$ trailing P-frames. The following segments have the same structure. Averagely, an I-frame is much larger than a DSC frame and a DSC frame is larger than a P-frame.

This structure can support cooperative view switching. Using Fig. 4.1 as an example, suppose that a peer i switches from view 1 to view 3 after the first segment. If another peer watches view 2 in the first segment and would like to share the reconstructed frame $F_{3,2}$, then i only needs to ask the server for the DSC frame of $F_{4,3}$ and the following $(K-1)$ trailing P-frames to reconstruct the video in view 3. If no one helps user i (either no user watches view 2 or view 3 in the first segment, or the users who can help are not willing to help), user i has to request the I-frame of $F_{4,3}$ from the server.

As studied in [6], [53], user cooperation plays the key role to make the streaming service be able to scale to large networks with thousands of users. In this work, to motivate users to seek others' help, we assume that the server's upload bandwidth is limited and expensive. Thus, it charges virtual currency from peers that pull video data as subscription fees to compensate its cost, and α denotes the price for the transmission of each single bit from the server. As discussed above, when a peer switches to a non-adjacent view, if he/she can get help from others, he/she will download the last reconstructed frame in the previous segment from the helper for free, and will only download a DSC frame from the server instead of an I-frame. Thus, he/she can receive a gain of $\alpha(size_I - size_{DSC})$ for paying less to the server, where $size_I$ and $size_{DSC}$ denote the number of bits of one I-frame and one DSC frame, respectively. However, uploading a reconstructed frame will incur a cost to the helper due to the consumed bandwidth, CPU time, etc. In this work, we consider the scenario with homogeneous users with the same cost to upload a frame. In the following discussion, without loss of generality, we normalize the gain of receiving a reconstructed frame to 1, and let c denote the normalized cost to upload a reconstructed frame to a peer.

4.1.3 Indirect Reciprocity Game

Since users are selfish, they want to receive others' help, but do not want to help others. To address the issue that users change partners frequently, we design a reputation-based mechanism to stimulate user cooperation, where peers who keep helping others will keep high reputations, and peers with high reputations also tend to receive others' help. In this mechanism, peer i helping peer j is not because j directly helped i previously, but j helped someone else. Thus, it is an indirect reciprocity game.

4.1.3.1 Peer Reputation and Interaction

In this system, each peer i is assigned a discrete reputation $r_i \in \mathcal{R} = \{1, 2, ..., R\}$, where a larger r_i indicates a higher reputation and peer i is more likely to receive others' help. Users' reputations change as they interact with each other. When a peer needs help, he/she first needs view information of other peers to find a suitable helper. To implement this, we can either let peers exchange their view information and seek help in a distributed way, or have a central controller that tracks peers' up-to-date view information and assigns helpers to peers that need help. For simplicity, we assume that there is a trustworthy local agent close to the N peers, who tracks peers' view switching, helps each peer find helpers, observes their interactions, and updates their reputations. Specifically, in this centralized system, when peer j needs help, the local agent randomly selects peer i from peers that can help, and sends a request. Upon receiving a request, peer i takes an threshold-based action $a_i \in \mathcal{A} = \{1, 2, ..., R+1\}$, with $\mathcal{A}$ as the action space. Here, the action a_i is not a direct answer of whether to help or not, but a threshold on reputation of peers whom peer i is willing to help, i.e., user i is willing to help users whose reputation is at least a_i. If $a_i = R+1$, user i will not cooperate with anyone, and $a_i = 1$ means user i is willing to help all users. After i sends the action a_i back to the local agent, the local agent compares j's reputation r_j with a_i. If $a_i \leq r_j$, the local agent informs i to upload the requested data to j. Otherwise, the local agent informs j to pull the I-frame from the streaming server. The rationale of this threshold-based action is that if user i is willing to help others with reputation a_i, user i should also be willing to help others with reputation higher than a_i. With the threshold based action, user i does not have to know from whom he/she receives this request, and thus, it is more suitable to our system, since as discussed later, each user needs to select actions for future interactions, and he/she does not know with whom he/she will interact at a later time.

4.1.3.2 Social Norm and Reputation Update

Based on the previous observed interaction between peer i who receives the request and j who sends the request, the local agent updates i's reputation following the pre-determined social norm that defines reputation update rules, while j's reputation remains the same. In this work, we use the social norm similar to that in [93], since it is effective in user cooperation stimulation.

In this reputation system, we have a pre-determined threshold $1 < t_r \leq R$. If user i has reputation $r_i \geq t_r$, he/she has high reputation and is likely to get others' help. Thus, he/she is called a beneficial user. Otherwise, he/she has low reputation and is not likely to get others' help. Thus, he/she is a non-beneficial user. If $r_i \geq t_r - 1$ (i.e., user i is a beneficial user or may become a beneficial user after this interaction), i's reputation is updated following the social norm,

$$\mathbf{Q} = \begin{array}{l} \\ a_i \leq r_j, \text{ uploading} \\ a_i > r_j, \text{ not uploading} \end{array} \begin{array}{c} \begin{array}{cc} \quad r_j \geq t_r \quad & \quad r_j < t_r \quad \end{array} \\ \begin{pmatrix} \min\{r_i+1, R\} & 1 \\ 1 & \min\{r_i+1, R\} \end{pmatrix} \end{array}. \tag{4.3}$$

From (4.3), if user i cooperates with a beneficial user or defects with a non-beneficial user, i's behavior complies with the social norm, and he/she is rewarded by one-step increase of his/her reputation. Otherwise, his/her behavior does not comply with the social norm, and he/she is punished by lowering his/her reputation to 1. Therefore, with this social norm, peers are encouraged to help beneficial users, but discouraged to help non-beneficial users.

If user i has reputation $r_i < t_r - 1$ (i.e. he/she is a non-beneficial user and cannot become beneficial after this interaction), his/her reputation will be increased by one-step to $(r_i + 1)$ no matter how he/she responses to user j's request. This is because the reputation $r_i < t_r - 1$ means that user i did not comply with the social norm in a previous interaction and was punished with the reputation being updated to 1. Similar to [108], the system takes time to forgive his/her misbehavior. During the forgiveness period, he/she has reputation $r_i < t_r - 1$ and hardly receives others' help, which may result in loss of utility. Here, t_r determines the duration of the forgiveness period and thus determines the punishment level of the reputation system, since user i needs to receive $t_r - 2$ requests to let his/her reputation climb from 1 to $t_r - 1$. At $t_r - 1$, he/she is forgiven by the system, and has the chance to become a beneficial user again after one interaction. Therefore, a larger t_r means it takes a longer time for the system to forgive a misbehavior and thus gives a harsher punishment. Note that when users make decisions, in addition to the social norm, they also take other factors into consideration, which will be discussed in details in Chapter 4.2, 4.3 and 4.4.

In this work, we consider the scenario that when users make decisions on the current requests, they also take the future interactions into consideration. Since they do not know with whom they will interact at a later time, the information of peers' reputation distribution $\mathbf{x}$ helps in their decision making, where $\mathbf{x}(r \in R)$ denotes the probability that a user has reputation r. Since we consider homogeneous users in this work, all users should follow the same $\mathbf{x}$ at the stationary state. Given that the local agent has the record of all peers' reputations at different time instances, it can estimate $\mathbf{x}$ using

$$\mathbf{x}(r) \;=\; \frac{\sum_{t=1}^{T_c} \sum_{i=1}^{N} I[r_i^t = r]}{NT_c}, \; \forall\, r \in R \tag{4.4}$$

where T_c is the current segment index, and $I[\cdot]$ is the indicator function. The local agent broadcasts $\mathbf{x}$ to all peers periodically to assist their decision making.

4.1.4 Optimal Action Selection with Markov Decision Process (MDP)

In our cooperative IMVS system, users may frequently switch views and their reputations may also change from time to time. Thus they may take different actions at different views and reputations. To address this issue of dynamic environment, we use Markov Decision Process (MDP) [104] to study user cooperation, where the game is played in a sequence of stages. In our IMVS, a stage represents an instance when a user receives a request and needs to make a decision, and there are $L \geq 1$ segments of video playback between two neighboring stages. Fig. 4.2 shows an example where a user receives a request at segment t_1 and will receive another request two segments later at $t_2 = t_1 + 2$ with $L = 2$. Following the work in [93], to simplify the analysis, we let L be the average interval between two consecutive requests received by a user in our work.

An MDP is defined as a four-tuple: the state space $\mathcal{S}$, the action pace $\mathcal{A}$, the state transition function $\mathcal{P}$ and the expected short-term utility function U. In our cooperative IMVS, a state $s = (r, v)$ represents a user's reputation r and view v when he/she receives a request. In the following sections, we will interchangeably use s and (r, v) to denote a state. Hence, the state space is denoted as $\mathcal{S} = \mathcal{R} \times \mathcal{V}$, where $\mathcal{V} = \{1, ..., M\}$ is the view space and $\mathcal{R} = \{1, 2, ..., R\}$

is the reputation space. At each state (r, v), a user can select action $a_{r,v}$ from the action space $\mathcal{A} = \{1, \cdots, R+1\}$. In the example in Fig. 4.2, there are $M = 3$ views and $R = 3$ reputation levels with $t_r = 3$. The state space includes a total of 9 states $\{(r, v)\}_{1 \leq r \leq 3, 1 \leq v \leq 3}$, and the action space includes 4 possible actions $\mathcal{A} = \{1, 2, 3, 4\}$.

A user receives a request at time t and he/she is at state (r, v). He/she takes action $a_{r,v}$ and transits to another state (r', v') with state transition probability $P^{a_{r,v}}_{(r,v)\to(r',v')}$ when he/she receives another request after L segments of video playback in the next stage. By taking action $a_{r,v}$, the user receives an expected short-term utility $U^{a_{r,v}}_{r,v}$, which contains two parts. First, this action may result in a frame upload to another peer, which incurs an expected cost $C^{a_{r,v}}$ immediately at time t. In addition, this action $a_{r,v}$ results in the update of the user's reputation to r' at time $t+1$, and he/she keeps reputation r' from time $t+1$ to $t+L$ until he/she receives another request. This updated reputation affects whether others are willing to help him/her in the following L segments (i.e., from time $t+1$ to $t+L$), and thus his/her gain in these L segments. Given the updated reputation r' and the view v that he/she is watching at time t, let $\theta(t+l)$ be the expected gain he/she receives at time $t+l$ for $1 \leq l \leq L$, and define $G_{r',v} = \sum_{l=1}^{L} \eta^l \theta(t+l)$ as the expected short-term gain, where $\eta \in (0, 1)$ is the discounting factor that quantifies how much users care about their future payoffs. Then, the expected short-term utility function is $U^{a_{r,v}}_{r,v} = G_{r',v} - C^{a_{r,v}}$. In the example in Fig. 4.2, when the user receives a request at time t_1, he/she is at state $(r = 3, v = 1)$. He/she selects an action $a \in \{1, 2, 3, 4\}$, receives an expected short-term utility $U^{a}_{3,1}$, and transits to another state $s' = (r', v')$ with probability $P^{a}_{(3,1)\to(r',v')}$ when he/she receives another request after $L = 2$ segments in the next stage at time t_2. This process is repeated until the end of the game.

The action policy in MDP is defined as $\pi = \{a_{r,v} \in \mathcal{A} | (r, v) \in \mathcal{S}\}$ that defines the action $a_{r,v}$ at each state (r, v). The goal of MDP is to find the optimal action policy that maximizes the expected lifetime utility, which is recursively defined as

$$W^{\pi}_{r,v} = U^{a_{r,v}}_{r,v} + \eta^L \sum_{(r',v')\in\mathcal{S}} P^{a}_{(r,v)\to(r',v')} W^{\pi}_{r',v'} \quad \forall (r, v) \in \mathcal{S}, \tag{4.5}$$

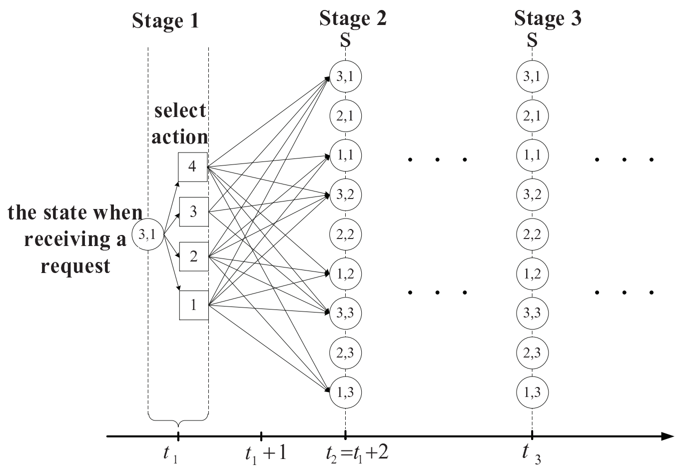


Fig. 4.2. An example of MDP with $M = 3$ views and $R = 3$ levels in the reputation system. All circles represent states, while all squares represent actions. The action space is $\mathcal{A} = \{1, 2, 3, 4\}$, and the average interval between two consecutive requests is $L = 2$.

where the second term denotes the user's lifetime utility since the next stage. In our cooperative IMVS, we consider the scenario with homogenous users, and thus their optimal action policies at the Nash Equilibrium are the same. To exam whether a policy π gives a Nash Equilibrium, for each user, we assume other users all take π, and if π also maximizes his/her lifetime utility, he/she has no incentive to deviate and π is an equilibrium policy.

4.2 MDP Analysis and Equilibrium Action Policy Discussion

In this section, we first analyze the state transition probability, the expected short-term utility and the lifetime utility of the MDP. We then discuss the equilibrium action policy that maximizes each user's lifetime utility.

4.2.1 MDP Analysis

4.2.1.1 State Transition Probability

We first analyze the probability that a user i transits from state (r, v) to (r', v') after L segments of video playback in the next stage. Note that in our IMVS, the view and reputation transition probabilities are independent. Given the one-step view transition matrix $\mathbf{T}$ in (4.1), the probability that user i transits from view v to v' in L segments is $\mathbf{T}^L(v, v')$. In the example in Fig. 4.2 with $L = 2$, user i is at view 1 at time t_1. From (4.2), after $L = 2$ segments, he/she will transit to view 1, 2 and 3 with probabilities $\mathbf{T}^2(1,1) = (1 - P_a)(1 - \frac{P_a}{2})$, $\mathbf{T}^2(1,2) = P_a$ and $\mathbf{T}^2(1,3) = P_a\frac{1-P_a}{2}$, respectively.

To find the reputation transition probability, suppose that user i at state (r, v) takes action $a_{r,v}$ as the response to a requester j, and user i's reputation is updated to r'. From Chapter 4.1.3.2, if user i's reputation is $r < t_r - 1$, then his/her reputation is always increased by 1, and we have $P^{a_{r,v}}_{r\to r+1} = 1$ and $P^{a_{r,v}}_{r\to r'} = 0$ for $r' \neq r+1$. When $r \geq t_r - 1$, user i's reputation is updated using social norm in (4.3) and the updated reputation is either $\min\{r + 1, R\}$ or 1. The updated reputation is 1 when i defects with j who is a beneficial user (i.e., $a_{r,v} > r_j \geq t_r$) or i cooperates with j who is a non-beneficial user (i.e., $a_{r,v} \leq r_j < t_r$). In addition, $P^{a_{r,v}}_{r\to\min\{r+1,R\}} = 1 - P^{a_{r,v}}_{r\to 1}$ and $P_{r\to r'} = 0$ for $r' \neq 1$ and $r' \neq \min\{r + 1, R\}$.

Since user i does not know user j's reputation r_j, i assumes that r_j follows the reputation distribution $\mathbf{x}$ and calculates the reputation transition probability using

$$
\begin{aligned}
P^{a_{r,v}}_{r\to 1} &= \begin{cases} \sum_{r_j=a_{r,v}}^{t_r-1} \mathbf{x}(r_j) & r \geq t_r - 1, a_{r,v} < t_r, \\ \sum_{r_j=t_r}^{a_{r,v}-1} \mathbf{x}(r_j) & r \geq t_r - 1, a_{r,v} > t_r, \\ 0 & \text{otherwise}, \end{cases} \\
P^{a_{r,v}}_{r\to\min\{r+1,R\}} &= 1 - P^{a_{r,v}}_{r\to 1}, \text{and } P^{a_{r,v}}_{r\to r'} = 0 \ \forall r' \neq 1, r' \neq \min\{r + 1, R\}. \qquad (4.6)
\end{aligned}
$$

In summary, the state transition probability is $P^{a_{r,v}}_{(r,v)\to(r',v')} = \mathbf{T}^L(v, v') \cdot P^{a_{r,v}}_{r\to r'}$. In the example in Fig. 4.2 (with $t_r = 3$), if the user takes action 1 at state $(r = 3, v =$

$1)$ (with $r > t_r - 1$) and cooperates with all users, he/she will help a non-beneficial user and his/her reputation will be lowered to 1 with probability $(\mathbf{x}(1) + \mathbf{x}(2))$. Thus, he/she will transit to state $(r' = 1, v' = 3)$ in the next stage with probability $P_a \frac{1-P_a}{2}(\mathbf{x}(1) + \mathbf{x}(2))$, and the probability to transit to other states can be calculated in the same way.

4.2.1.2 Expected Short-term Utility and Lifetime Utility

We now analyze the expected short-term and the lifetime utility functions in (4.5), and start with the expected short-term utility.

From the discussion in Chapter 4.1.4, the expected short-term utility $U_{r,v}^{a_{r,v}}$ contains two parts: the expected immediate cost $C^{a_{r,v}}$ and the expected short-term gain $G_{r',v}$. If user i chooses action $a_{r,v}$, assuming the requester j's reputation follows the distribution $\mathbf{x}$, the probability that he/she uploads the frame and thus incurs an immediate cost of c at time t is $\sum_{r_j=a_{r,v}}^{R} \mathbf{x}(r_j)$.[1] Therefore, his/her expected cost is $C^{a_{r,v}} = c\sum_{r_j=a_{r,v}}^{R} \mathbf{x}(r_j)$.

To analyze $G_{r',v}$, note that taking action $a_{r,v}$ makes user i's reputation updated to r' at time $t+1$ and he/she keeps r' for the following L segments (i.e., from time $t+1$ to time $t+L$). We then derive the gain he/she receives at each time $t+l$ (with $1 \le l \le L$) given that he/she watches view v at time t. At time $t+l$, user i receives a positive normalized gain 1 if and only if he/she switches to a non-adjacent view (i.e., he/she needs help) at time $t+l$ and there is a user who can and is willing to help him. Otherwise, his/her gain is 0. Let $P_{r',v}(t+l)$ denote the probability that user i switches to a non-adjacent view at time $t+l$ and there is a user who can and is willing to help him/her. Thus, we have $G_{r',v} = \sum_{l=1}^{L} \eta^l P_{r',v}(t+l)$. We then derive $P_{r',v}(t+l)$ step by step.

Let $v_i(t+l)$ denote the view that user i watches at time $t+l$. For a given view v', let $\mathcal{V}_{v'} \triangleq \{\max(v'-1,1), v', \min(M, v'+1)\}$ be the set including all adjacent views of v'. In the example in Fig. 4.2, $\mathcal{V}_1 = \{1,2\}$, $\mathcal{V}_2 = \{1,2,3\}$ and $\mathcal{V}_3 = \{2,3\}$.

[1] $\sum_{r_j=a_{r,v}}^{R} \mathbf{x}(r_j) = 0$ if $a_{r,v} = R+1$.

Then, we have

$$P_{r',v}(t+l) = \sum_{v'=1}^{M} \Big\{ P\left[\mathbb{H}_h | \mathbb{H}_1(v')\right] P\left[\mathbb{H}_1(v')\right] \times P\left[(v_i(t+l) = v', v_i(t+l-1) \notin \mathcal{V}_{v'} | v_i(t) = v\right] \Big\}, \quad (4.7)$$

where $P\left[v_i(t+l) = v', v_i(t+l-1) \notin \mathcal{V}_{v'} | v_i(t) = v\right]$ is the probability that given that user i is at view v at time t, he/she switches to view v' at time $t+l$ from a non-adjacent view and needs help, $\mathbb{H}_1(v')$ is the event that there is at least one helper who can help user i switch to view v' at time $t+l$ (i.e., there is at least one user who is watching a neighboring view of v' at time $t+l-1$), and $\mathbb{H}_h$ is the event that the selected helper is willing to help. Note that

$$\begin{aligned} & P\left[v_i(t+l) = v', v_i(t+l-1) \notin \mathcal{V}_{v'} | v_i(t) = v\right] \\ = & \sum_{v'' \notin \mathcal{V}_{v'}} P\left[v_i(t+l) = v' | v_i(t+L-1) = v''\right] P\left[v_i(t+l-1) = v'' | v_i(t) = v\right] \\ = & \sum_{v'' \notin \mathcal{V}_{v'}} \mathbf{T}(v'', v') \mathbf{T}^{l-1}(v, v''). \end{aligned} \quad (4.8)$$

To find $P\left[\mathbb{H}_1(v')\right]$, given the stationary view distribution $\mathbf{v}$, we have

$$P\left[\mathbb{H}_1(v')\right] = 1 - \left(1 - \sum_{v'' \in \mathcal{V}_{v'}} \mathbf{v}(v'')\right)^{N-1}. \quad (4.9)$$

To find the probability that the selected helper k is willing to help, helper k will help user i if user i's current reputation r' is larger than or equal to helper k's decision $a_{r_k, v_k(t+1)}$, which depends on helper k's reputation r_k and view $v_k(t+l)$ at time $t+l$ when k receives the request. Therefore, we have

$$\begin{aligned} P\left[\mathbb{H}_h | \mathbb{H}_1(v')\right] &= \sum_{r_k=1}^{R} \sum_{v_k(t+l)=1}^{M} \mathbf{x}(r_k) \mathbf{p}_{v'}(v_k(t+l)) I\left[a_{r_k, v_k(t+l)} \leq r'\right], \\ \text{where } \mathbf{p}_{v'}(v_k(t+l)) &= P\left[v_k(t+l) | v_k(t+l-1) \in \mathcal{V}_{v'}\right] \\ &= \sum_{v'' \in \mathcal{V}_{v'}} \Big\{ P\left[v_k(t+l) | v_k(t+l-1) = v''\right] \\ &\quad \times P\left[v_k(t+l-1) = v'' | v_k(t+l-1) \in \mathcal{V}_{v'}\right] \Big\} \end{aligned}$$

$$= \sum_{v'' \in \mathcal{V}_{v'}} \mathbf{T}(v'', v_k(t+l)) \frac{\mathbf{v}(v'')}{\sum_{\tilde{v} \in \mathcal{V}_{v'}} \mathbf{v}(\tilde{v})}. \quad (4.10)$$

Here, $\mathbf{x}$ is user's reputation distribution, $\mathbf{v}$ is the steady state view distribution, and $\mathbf{p}_{v'}(v_k(t+l))$ is the probability that given helper k watches a neighboring view of v' at time $t+l-1$, he/she switches to view $v_k(t+l)$ at time $t+l$. Therefore, we have

$$\begin{aligned} G_{r',v} &= \sum_{l=1}^{L} \eta^l \sum_{v'=1}^{M} \Big\{ P\left[\mathbb{H}_h | \mathbb{H}_1(v')\right] P\left[\mathbb{H}_1(v')\right] \\ &\quad \times P\left[v_i(t+l) = v', v_i(t+l-1) \notin \mathcal{V}_{v'} | v_i(t) = v\right] \Big\}. \end{aligned} \quad (4.11)$$

Based on the above discussion, $G_{r',v}$ is affected by user i's view navigation path in the next L segments starting from view v. If he/she has a low probability to switch to non-adjacent views during the next L segments, he/she will also tend to have a small $G_{r',v}$. It is also easy to observe that $G_{r',v}$ is an increasing function of r', since a higher r' gives a higher probability to get others' help. Note that in (4.11), if $P\left[\mathbb{H}_h | \mathbb{H}_1(v')\right]$ is always 1, that is, helpers are always willing to help user i, then $G_{r',v}$ becomes

$$G_{r',v} = g_v \triangleq \sum_{l=1}^{L} \eta^l \sum_{v'=1}^{M} P[\mathbb{H}_1(v')] P[v_i(t+l) = v', v_i(t+l-1) \notin \mathcal{V}_{v'} | v_i(t) = v] \quad (4.12)$$

Thus, g_v is a user's maximum expected short-term gain if he/she starts view switching from v, where he/she always receives help when needed. Based on the above analysis, together with the fact that the updated r' can only be $\min\{r+1, R\}$ or 1, we can derive user i's expected short-term utility after taking action $a_{r,v}$ as

$$U_{r,v}^{a_{r,v}} = -c \sum_{r_j = a_{r,v}}^{R} \mathbf{x}(r_j) + (1 - P_{r \to 1}^{a_{r,v}}) G_{\min(r+1,R),v} + P_{r \to 1}^{a_{r,v}} G_{1,v}. \quad (4.13)$$

Thus, following (4.5), we have the lifetime utility $W_{r,v}^{\pi}$ following the action policy

π,

$$W^{\pi}_{r,v} = U^{a_{r,v}}_{r,v} + \eta^L \sum_{v'=1}^{M} \mathbf{T}^L(v,v') \left[(1 - P^{a_{r,v}}_{r\to 1}) W^{\pi}_{\min(r+1,R),v'} + P^{a_{r,v}}_{r\to 1} W^{\pi}_{1,v'}\right]. \quad (4.14)$$

4.2.2 Discussion on the Equilibrium Policies

4.2.2.1 Elimination of Non-Equilibrium Policies

In this work, we aim to derive the equilibrium action policy π, from which no one has incentive to unilaterally deviate. In our MDP, the size of the state space is $|\mathcal{R}||\mathcal{V}| = RM$, and we have $(R+1)^{RM}$ possible action policies. To avoid examining all these policies, we need to first eliminate non-equilibrium ones using Theorem 4.1, which we will discuss one by one in the following.

Theorem 4.1. *In an equilibrium action policy* π*,*

a) For all $r < t_r - 1$ *and all* $v \in \mathcal{V}$*,* $a_{r,v} = R + 1$*.*

b) For all $r \geq t_r - 1$ *and all* $v \in \mathcal{V}$*,* $a_{r,v} \in \{t_r, R + 1\}$*.*

c) For any view v*, a user will take the same action with all reputations* $r \geq t_r - 1$*, i.e.,* $a_{t_r-1,v} = a_{t_r,v} = ... = a_{R,v}$*.*

4.2.2.1.1 Proof of Theorem 4.1a) Theorem 4.1a) says if a user is not a beneficial user and cannot become a beneficial user after this decision, he/she will not cooperate no matter which view he/she is watching. This is because when $r < t_r - 1$, no matter what action he/she takes, his/her reputation will be always increased by one, while $a_{r,v} = R + 1$ gives him/her zero cost since he/she will not help anyone. Thus, $a_{r,v} = R + 1$ dominates the other actions if $r < t_r - 1$.

4.2.2.1.2 Proof of Theorem 4.1b) Theorem 4.1b) says if a user is a beneficial user or may become a beneficial user after this decision, he/she will either cooperate with beneficial users (i.e., $a = t_r$) or do not cooperate with anyone (i.e., $a = R+1$). It takes two steps to prove this. We will first show that the action $a_{r,v} = t_r$ dominates all actions $a_{r,v} < t_r$. We then show that any action policy with action $t_r+1 \leq a_{r,v} \leq R$ cannot be an equilibrium action policy. From these two results, $a_{r,v}$ can only be t_r or $R+1$.

We first compare the action $a_{r,v} = t_r$ with $a_{r,v} < t_r$ in terms of the incurred cost $C^{a_{r,v}}$ and the updated reputation r'. First, with action $a_{r,v} = t_r$, the expected cost to upload is $C^{t_r} = c\sum_{r_j=t_r}^{R} \mathbf{x}(r_j)$, and with $a_{r,v} < t_r$, the cost to upload is $C^{a_{r,v}<t_r} = c\sum_{r_j=a_{r,v}}^{R} \mathbf{x}(r_j) \geq C^{t_r}$. Second, with action $a_{r,v} = t_r$, from (4.6), the user's reputation is rewarded with one-step increase with probability 1. However, with $a_{r,v} < t_r$, his/her reputation is rewarded with one-step increase with probability $1 - (\sum_{r_j=a_{r,v}}^{t_r-1} \mathbf{x}(r_j)) \leq 1$. Therefore, $a_{r,v} = t_r$ introduces a lower cost, but gives a higher probability to be rewarded with one-step increase of a user's reputation. Thus, $a_{r,v} = t_r$ is a dominant strategy over all $a_{r,v} < t_r$, and in the equilibrium, we should have $a_{r,v} \geq t_r$.

Then, to show that action $t_r + 1 \leq a_{r,v} \leq R$ cannot be in an equilibrium policy, we use the *One-shot Deviation Principle* [109], which says a strategy profile (including all users' action policies) is an equilibrium if and only if no one can gain by one-shot deviation when others keep their strategies unchanged. Here, one-shot deviation of a given action policy means that a user takes a different action rather than the one defined in the action policy only for the current response to a request, but still follows the given action policy in the future responses. We have the following proposition and its proof is in Appendix D.

Proposion 4.1. *For a policy π with $t_r+1 \leq a_{r,v} \leq R$ for the reputation $r \geq t_r - 1$, one-shot deviation to either $a'_{r,v} = t_r$ or $a'_{r,v} = R+1$ will give a higher lifetime utility and thus this π cannot be an equilibrium policy.*

Thus, from Proposition 4.1 and the fact that $a_{r,v} = t_r$ dominates all action $a_{r,v} < t_r$, $a_{r,v}$ can only be either t_r or $R+1$ in an equilibrium policy, i.e., either cooperate with all beneficial users or do not cooperate with anyone.

4.2.2.1.3 Proof of Theorem 4.1c) We first define a reputation subspace $\bar{\mathcal{R}} = \{r | t_r - 1 \leq r \leq R\}$ including all reputations no less than $t_r - 1$. Then, we define a state subspace $\mathcal{S}_{\bar{\mathcal{R}},v} = \{(r,v)|r \in \bar{\mathcal{R}}, v \in \mathcal{V}\}$ that includes all states with view v and reputations no less than $t_r - 1$. Theorem 4.1c) says that for all states in $\mathcal{S}_{\bar{\mathcal{R}},v}$, a user should take the same action (either cooperating with beneficial users with $a_{r,v} = t_r$ or not cooperating with anyone with $a_{r,v} = R+1$), i.e., for any view v, we have $a_{t_r-1,v} = a_{t_r,v} = ... = a_{R,v}$ in an equilibrium policy. To prove this, we use the concept of *Bisimilarity* [110], which is defined bellow.

Definition 4.1. *(*Bisimilarity*) In an MDP, suppose that the state space is divided into m non-overlapping subspaces: $\mathcal{S} = \mathcal{S}_1 \bigcup \mathcal{S}_2 \bigcup ... \bigcup \mathcal{S}_m$. For any $\mathcal{S}_i$ ($1 \leq i \leq m$), and any two states $s \in \mathcal{S}_i$ and $s' \in \mathcal{S}_i$ ($s \neq s'$), if for any action a, we have i) $\sum_{s'' \in \mathcal{S}_j} P^a_{s \to s''} = \sum_{s'' \in \mathcal{S}_j} P^a_{s' \to s''}$ for all $1 \leq j \leq m$ (i.e., with the same action a, the two states s and s' have the same probability to transit to another state subspace $\mathcal{S}_j$); and ii) $U^a_s = U^a_{s'}$, (i.e., with the same action a, the two states s and s' have the same expected short-term utility), then all states in the same state subspace $\mathcal{S}_i$ have the bisimilarity relationship, i.e., they are equivalent and can be aggregated as one state ξ_i.*

To study states with bisimilarity relationship in our MDP, we first divide the state space $\mathcal{S}$ into subspaces. For any view $v \in \mathcal{V}$, we have $S_{\bar{\mathcal{R}},v}$ defined earlier, which includes all states with view v and reputations no less than $t_r - 1$. For any view $v \in \mathcal{V}$ and reputation $1 \leq r \leq t_r - 2$, the state (r,v) forms a state subspace $\{(r,v)\}$ with a single element. All these subspaces are non-overlapping and we have $\bigcup_{v \in \mathcal{V}} \left(\{(1,v)\} \bigcup ... \bigcup \{(t_r-2,v)\} \bigcup \mathcal{S}_{\bar{\mathcal{R}},v}\right) = \mathcal{S}$. For the example in Fig. 4.2 with 3 views and 3-level reputation system (where $t_r = 3$), Fig. 4.3a shows the corresponding states partition with 6 subspaces $\{(1,1)\}, \mathcal{S}_{\bar{\mathcal{R}},1} = \{(2,1),(3,1)\}, \{(1,2)\}, \mathcal{S}_{\bar{\mathcal{R}},2} = \{(2,2),(3,2)\}, \{(1,3)\}$, and $\mathcal{S}_{\bar{\mathcal{R}},3} = \{(2,,3),(3,3)\}$. We then have the following proposition, and the proof is in Appendix E.

Proposion 4.2. *Following the above state partition, all states in $\mathcal{S}_{\bar{\mathcal{R}},v}$ have bisimilarity relationship, and can be aggregated as one state.*

In the example in Fig. 4.3a, after state aggregation, there are 6 aggregated states: $(1,1)$, $(\bar{\mathcal{R}},1)$, $(1,2)$, $(\bar{\mathcal{R}},2)$, $(1,3)$ and $(\bar{\mathcal{R}},3)$, and the MDP in Fig. 4.2 becomes

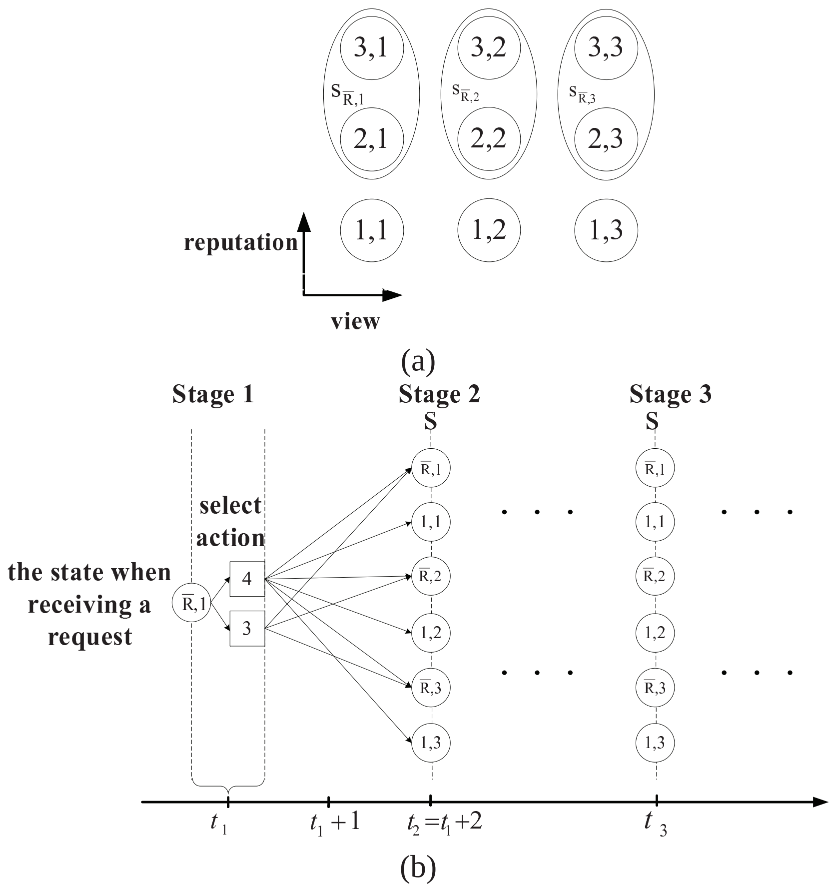


Fig. 4.3. Example of the state classification and aggregation with 3 views and 3-level reputation system ($t_r = 3$). (a) The state classification, where the states $(3, v)$ and $(2, v)$ forms a subspace $\mathcal{S}_{\bar{\mathcal{R}},v}$ and state $(1, v)$ forms a subspace $\{(1, v)\}$ with a single element. (b) The MDP after aggregating the state space $\mathcal{S}_{\bar{\mathcal{R}},v}$ as one state $(\bar{\mathcal{R}}, v)$. The action can only be selected from $\{t_r, R + 1\} = \{3, 4\}$.

the one shown in Fig. 4.3b.

The next step is to find the transition probability and the expected short-term utility function for the updated MDP. Following Definition 4.1, given the aggregated states $\{\xi_1, \cdots, \xi_m\}$, by taking action a, a user transits from state ξ_i to state ξ_j with probability $P^a_{\xi_i \to \xi_j} = \sum_{s' \in \mathcal{S}_j} P^a_{s \to s'}$ for any $s \in \mathcal{S}_i$, and the expected short-term utility at the aggregated state ξ_i is $U^a_{\xi_i} = U^a_s$ for any $s \in \mathcal{S}_i$. In our MDP, the view and the reputation transition are independent, and the state aggregation here affects the reputation transition probability only. Therefore, we need to first find the updated reputation transition probability $P^a_{\bar{\mathcal{R}} \to \bar{\mathcal{R}}}$, $P^a_{\bar{\mathcal{R}} \to r'}$ and $P^a_{r' \to \bar{\mathcal{R}}}$ for $r' < t_r - 1$. First, if $r' = t_r - 2$, since the reputation will always be increased by one step to $t_r - 1 \in \bar{\mathcal{R}}$ regardless the action a, thus, we have $P^a_{r' \to \bar{\mathcal{R}}} = 1$. For $r' \leq t_r - 3$, we have $P^a_{r' \to \bar{\mathcal{R}}} = 0$. As discussed in Chapter 4.1.3.2, with a reputation $r \in \bar{\mathcal{R}}$, the updated reputation r' can only be 1 or $\min\{r+1, R\} \in \bar{\mathcal{R}}$. Therefore, we have $P^a_{\bar{\mathcal{R}} \to \bar{\mathcal{R}}} = 1 - P^a_{\bar{\mathcal{R}} \to 1}$ and $P^a_{\bar{\mathcal{R}} \to r'} = 0$ for $2 \leq r' \leq t_r - 2$. The proof of Proposition 4.2 shows that $P^a_{r \to 1}$ is the same for any $r \geq t_r - 1$. Thus, we have $P^a_{\bar{\mathcal{R}} \to 1} = P^a_{r \to 1}$ for any $r \in \bar{\mathcal{R}}$. Then, we can find the updated state transition probability $P^a_{(r,v) \to (r',v')} = P^a_{r \to r'} \mathbf{T}^L(v, v')$ for all $r, r' \in \{1, \cdots, t_r - 1, \bar{\mathcal{R}}\}$.

In the MDP in Fig. 4.3b, when a user is at the aggregated state $(\bar{\mathcal{R}}, v)$, from Theorem 4.1b), he/she will take either action $a = t_r = 3$ or action $a = R + 1 = 4$. With $a = 3$, he/she conforms with the social norm in (4.3) and his/her reputation will be updated to 3 with probability 1, that is, $P^{a=t_r}_{\bar{\mathcal{R}} \to \bar{\mathcal{R}}} = P^{a=3}_{2 \to 3} = P^{a=3}_{3 \to 3} = 1$, and $P^{a=t_r}_{\bar{\mathcal{R}} \to 1} = P^{a=3}_{2 \to 1} = P^{a=3}_{3 \to 1} = 0$. Thus, we can calculate the updated state transition probabilities $P^{a=3}_{(\bar{\mathcal{R}},v) \to (\bar{\mathcal{R}},v')} = P^{a=3}_{\bar{\mathcal{R}} \to \bar{\mathcal{R}}} \mathbf{T}^L(v, v') = \mathbf{T}^L(v, v')$ and $P^{a=3}_{(\bar{\mathcal{R}},v) \to (1,v')} = P^{a=3}_{\bar{\mathcal{R}} \to 1} \mathbf{T}^L(v, v') = 0$. Similarly, we can find the state transition probabilities with action $a = 4$ for the MDP in Fig. 4.3b.

The last step is to update the expected short-term utility function $U^a_{\bar{\mathcal{R}},v}$. Proposition 4.2 shows that with the same action a, $U^a_{r,v}$ is the same for all $r \in \bar{\mathcal{R}}$ and thus, $U^a_{\bar{\mathcal{R}},v} = U^a_{r,v}$ for any $r \in \bar{\mathcal{R}}$.

4.2.2.2 Lifetime Utility Functions

In the following, we will study how the state aggregation affects the lifetime utility functions. From Theorem 4.1, when a user's reputation is lower than $t_r - 1$, his/her only equilibrium action is $a = R+1$ (does not cooperate with any one). Therefore, we will focus on the action selection for reputation $r \in \bar{\mathcal{R}} = \{t_r - 1, ..., R\}$ in the following section. In addition, from Theorem 4.1c, when $r \in \bar{\mathcal{R}}$, the equilibrium actions are the same for all $r \geq t_r - 1$ and depend on view v only. Thus, to simplify the notation, we omit the reputation index r in the action and the action policy becomes $\pi = \{a_1, a_2, ...a_M\}$, where $a_v \in \{t_r, R+1\}$ is the action at view $v \in \mathcal{V}$ with $r \in \bar{\mathcal{R}}$.

With the above simplification of notations, we then study the lifetime utility with the aggregated states, and (4.14) can be rewritten as

$$W^{\pi}_{\bar{\mathcal{R}},v} = U^{a_v}_{\bar{\mathcal{R}},v} + \eta^L \sum_{v'=1}^{M} \mathbf{T}^L(v,v') \left[(1 - P^{a_v}_{\bar{\mathcal{R}}\to 1})W^{\pi}_{\bar{\mathcal{R}},v'} + P^{a_v}_{\bar{\mathcal{R}}\to 1} W^{\pi}_{1,v'}\right]. \quad (4.15)$$

Note that (4.15) has a recursive term $W_{1,v}(\pi)$ that we need to solve first. Since a user with reputation $r < t_r - 1$ always uses action $R+1$ and does not help anyone, his/her expected immediate cost is zero. In addition, from Theorem 4.1, with the equilibrium policy, no one helps users with reputation smaller than t_r. Therefore, he/she does not receive any expected short-term gain from others' help with $G_{r',v} = 0$. Thus, his/her expected short-term utility is always zero. Furthermore, his/her reputation always increases by 1 every time he/she receives a request, until his/her reputation climbs to $t_r - 1$, i.e., to $\bar{\mathcal{R}}$. Therefore, $W_{1,v}(\pi)$ can be expanded following (4.14) as

$$\begin{aligned}
W^{\pi}_{1,v} &= 0 + \eta^L \sum_{v'=1}^{M} \mathbf{T}^L(v,v') W^{\pi}_{2,v'} \\
&= \eta^L \sum_{v'=1}^{M} \mathbf{T}^L(v,v') \left[0 + \eta^L \sum_{v''=1}^{M} \mathbf{T}^L(v',v'') W^{\pi}_{3,v''}\right] \\
&= \eta^{2L} \sum_{v''=1}^{M} \mathbf{T}^{2L}(v,v'') W^{\pi}_{3,v''} = ... = \eta^{(t_r-2)L} \sum_{v'=1}^{M} \mathbf{T}^{(t_r-2)L}(v,v') W^{\pi}_{t_r-1,v'}
\end{aligned}$$

$$= \quad \eta^{(t_r-2)L}\sum_{v'=1}^{M}\mathbf{T}^{(t_r-2)L}(v,v')W^{\pi}_{\bar{\mathcal{R}},v'}. \tag{4.16}$$

Note that in (4.15) and (4.16), $\bar{\mathcal{R}}$ is a common reputation index in the subscripts of $W^{\pi}_{\bar{\mathcal{R}},v}$ and $U^{a_v}_{\bar{\mathcal{R}},v}$. We can further simplify the notation by omitting this common reputation index, substitute (4.16) into (4.15), and rewrite (4.15) as

$$W^{\pi}_{v}= U^{a_v}_{v}+ \eta^L(1-P^{a_v}_{\bar{\mathcal{R}}\to 1})\sum_{v'=1}^{M}\mathbf{T}^L(v,v')W^{\pi}_{v'} + \gamma P^{a_v}_{\bar{\mathcal{R}}\to 1}\sum_{v'=1}^{M}\mathbf{T}^{(t_r-1)L}(v,v')W^{\pi}_{v'}. \tag{4.17}$$

where $\gamma = \eta^{L(t_r-1)}$ is the discounting factor after receiving $t_r - 1$ requests.

To determine if a policy π is a Nash Equilibrium, we need to show that it can resist any one-shot deviation, where the user takes action a'_v other than the action a_v defined in π only for the current response to a request, and he/she will follow π in all later responses. The lifetime utility with the one-shot deviation to action a'_v is

$$W^{a'_v,\pi}_{v}=U^{a'_v}_{v}+ \eta^L(1-P^{a'_v}_{\bar{\mathcal{R}}\to 1})\sum_{v'=1}^{M}\mathbf{T}^L(v,v')W^{\pi}_{v'}+\gamma P^{a'_v}_{\bar{\mathcal{R}}\to 1}\sum_{v'=1}^{M}\mathbf{T}^{(t_r-1)L}(v,v')W^{\pi}_{v'}. \tag{4.18}$$

Comparing (4.17) and (4.18), one-shot deviation to a'_v gives a different expected short-term utility $U^{a'_v}_{v}$ and a different reputation transition probability $P^{a'_v}_{\bar{\mathcal{R}}\to 1}$. We then use the one-shot deviation principle to examine whether a policy π is an equilibrium policy, that is, from a user's perspective, given that other users all take π unchanged, π is an equilibrium policy if and only if $W^{\pi}_{v} \geq W^{a'_v,\pi}_{v}$ for any v and a'_v.

4.2.3 Stationary Reputation Distribution

From the previous analysis, users' reputation distribution $\mathbf{x}$, where $\mathbf{x}(r)$ being the probability that a user has reputation r, affects the state transition probability and users' expected short-term utilities. Thus, it affects users' decision making. If an equilibrium exists in the game, in the simple scenario with homogeneous users, all users should use the same strategy in the equilibrium, and it is expected that the reputation distribution should also converge to a stationary state. In the following, given a policy π adopted by all users, we determine whether there exists a stationary reputation distribution $\mathbf{x}$.

We first let $y = \sum_{r=t_r}^{R} \mathbf{x}(r)$ be the probability that a user is a beneficial user. Assume that user i receives a request from user j, and both of their reputation distribution follows $\mathbf{x}$. Then, following the social norm in Chapter 4.1.3.2, if user i's current reputation is $r \leq t_r - 2$ (with probability $\mathbf{x}(r)$), his/her reputation will be increased to $r+1$ for any action he/she takes. Thus, in the updated reputation distribution $\mathbf{x}'$, we should have $\mathbf{x}'(r) = \mathbf{x}(r-1)$ for $2 \leq r \leq t_r - 1$. If the stationary state exists, the reputation distribution should remain the same and $\mathbf{x}'(r) = \mathbf{x}(r)$. Therefore, we have $\mathbf{x}(1) = \mathbf{x}(2) = ... = \mathbf{x}(t_r - 1)$. In addition, given $y + \sum_{r=1}^{t_r-1} \mathbf{x}(r) = 1$, we have $\mathbf{x}(1) = \cdots = \mathbf{x}(t_r - 1) = (1-y)/(t_r - 1)$.

If user i's reputation is $r \in \bar{\mathcal{R}}$ (which happens with probability $y + \mathbf{x}(t_r - 1)$), his/her action $a_v = \{t_r, R+1\}$ only depends on his/her view v when receiving a request, and at the stationary state he/she is at view v with probability $\mathbf{v}(v)$. i's reputation will then be updated to either 1 or $\min\{r+1, R\}$. Given his/her possible action $a_v = \{t_r, R+1\}$ and from (4.6), his/her reputation is updated to 1 if and only if he/she takes action $R+1$ and user j who sends the request is a beneficial user (i.e., user j has reputation no less than t_r). Therefore, given user i is at view v when he/she receives the request, his/her reputation is reduced to 1 with probability $P^{a}_{\bar{\mathcal{R}}\to 1} = I[a_v = R+1]y$, and he/she becomes (or remains) a beneficial user with probability $1 - P^{a}_{\bar{\mathcal{R}}\to 1}$. Therefore, after the reputation update, user i is a beneficial user with probability

$$\begin{aligned} y' &= [y + \mathbf{x}(t_r - 1)] \sum_{v\in\mathcal{V}} \mathbf{v}(v)(1 - P^{a_v}_{\bar{\mathcal{R}}\to 1}) \\ &= [y + \mathbf{x}(t_r - 1)] \sum_{v\in\mathcal{V}} \mathbf{v}(v) \left(1 - yI[a_v = R+1]\right). \end{aligned} \tag{4.19}$$

For a given policy $\pi = \{a_v\}$, if the stationary state exists, we should have $y' = y$, and y should satisfy

$$y' - y = \left(y + \frac{1-y}{t_r - 1}\right) \sum_{v\in\mathcal{V}} \mathbf{v}(v) \left\{1 - yI[a_v = R+1]\right\} - y = 0. \tag{4.20}$$

We first observe that given π, the left hand side (LHS) of (4.20) is a quadrature function of y. When $y = 1$, $LHS = \sum_{v \in \mathcal{V}} \mathbf{v}(v) \{1 - I[a_v = R + 1]\} - 1 \leq \sum_{v \in \mathcal{V}} \mathbf{v}(v) - 1 = 0$, and when $y = 0$, $LHS = \frac{1}{t_r - 1} \sum_{v \in \mathcal{V}} \mathbf{v}(v) > 0$. Thus, (4.20) has a single root in the range $[0, 1]$. Therefore, given π, there exists a unique stationary reputation distribution $\mathbf{x}$. To find $\mathbf{x}$ for a given policy π, we first solve (4.20) and find y, and then calculate $\mathbf{x}(1) = ... = \mathbf{x}(t_r - 1) = \frac{1-y}{t_r - 1}$.

4.3 Equilibrium Action Policy Derivation

In this section, we analytically derive the equilibrium action policy of the game. We first consider a simple scenario with a single anchor view and analyze the equilibrium policy in Chapter 4.3.1. We then extend our analysis to the general case with multiple anchor views in Chapter 4.3.2.

4.3.1 Game Analysis with A Single Anchor View

4.3.1.1 View Switching Model with A Single Anchor View

Following the view switching model described in Chapter 4.1.1, with a single anchor view as shown in Fig. 4.4, this anchor view is in the middle, and partitions the rest $M - 1$ normal views into two normal view sets with $(M - 1)/2$ views per set. (Here, we assume M is an odd number.) Let $\sigma = (M + 1)/2$ denote the anchor view index. Following Chapter 4.1.1, the one-step view transition matrix is

$$\mathbf{T} = \begin{array}{cc} & \begin{array}{ccccccc} 1 & \cdots & \sigma - 1 & \sigma & \sigma + 1 & \cdots & M \end{array} \\ \begin{array}{c} 1 \\ \vdots \\ \sigma - 1 \\ \sigma \\ \sigma + 1 \\ \vdots \\ 1 \end{array} & \begin{pmatrix} \frac{2(1-P_a)}{M-1} & \cdots & \frac{2(1-P_a)}{M-1} & P_a & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{2(1-P_a)}{M-1} & \cdots & \frac{2(1-P_a)}{M-1} & P_a & 0 & \cdots & 0 \\ \frac{(1-P_a)}{M-1} & \cdots & \frac{(1-P_a)}{M-1} & P_a & \frac{(1-P_a)}{M-1} & \cdots & \frac{(1-P_a)}{M-1} \\ 0 & \cdots & 0 & P_a & \frac{2(1-P_a)}{M-1} & \cdots & \frac{2(1-P_a)}{M-1} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & P_a & \frac{2(1-P_a)}{M-1} & \cdots & \frac{2(1-P_a)}{M-1} \end{pmatrix} \end{array}. \quad (4.21)$$

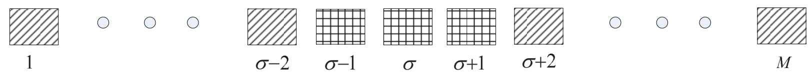


Fig. 4.4. $\underline{\mathcal{V}} = \{\sigma - 1, \sigma, \sigma + 1\}$, which denotes the view set including the anchor view σ and its left and right adjacent views. $\bar{\mathcal{V}} = \mathcal{V} \setminus \underline{\mathcal{V}}$, which denotes the view set including the rest views.

With the above one-step view transition matrix $\mathbf{T}$, it is easy to show that the steady-state view distribution is $\mathbf{v}(\sigma) = P_a$ and $\mathbf{v}(v) = (1 - P_a)/(M - 1)$ for all other views.

4.3.1.2 Expected Short-term Gain with A Single Anchor View

For IMVS with a single anchor view, we first study the expected short-term gain $G_{r',v}$ for different views. In Appendix F, we show that with a single anchor view, $G_{r',v}$ in (4.11) can be rewritten as

$$G_{r',v} = \left(\sum_{r_k = t_r - 1}^{R} \mathbf{x}(r_k) \right) \left(\sum_{v_k = 1}^{M} \mathbf{v}(v_k) I[a_{v_k} \leq r'] \right) g_v, \qquad (4.22)$$

where g_v defined in (4.12) is the maximum expected short-term gain at view v when helpers always help, and it is the only term in (4.22) that is affected by view v. In the following, we compare g_v with different v's.

We first divide the view space $\mathcal{V}$ into two sets $\underline{\mathcal{V}} = \{\sigma - 1, \sigma, \sigma + 1\}$ and $\bar{\mathcal{V}} = \mathcal{V} \setminus \underline{\mathcal{V}}$ as shown in Fig. 4.4. $\underline{\mathcal{V}}$ includes the anchor view and its left and right adjacent views, and $\bar{\mathcal{V}}$ includes the rest views. Then, we have the following Proposition 4.3, and the proof is in Appendix G.

Proposion 4.3. *In a high dimensional IMVS (where the total number of views, M, is large, e.g., $M \geq 30$) with a single anchor view, all views in $\underline{\mathcal{V}}$ have approximately the same g_v's, and for views in $\bar{\mathcal{V}}$, their g_v's are also approximately the same.*

We first let $g_{\underline{\mathcal{V}}} \triangleq \sum_{v \in \underline{\mathcal{V}}} g_v / |\underline{\mathcal{V}}|$ denote the average g_v of all $v \in \underline{\mathcal{V}}$, and similarly, we let $g_{\bar{\mathcal{V}}} \triangleq \sum_{v \in \bar{\mathcal{V}}} g_v / |\bar{\mathcal{V}}|$. We then let $\delta_{\bar{\mathcal{V}}} \triangleq \frac{\max_{v \in \bar{\mathcal{V}}} g_v - \min_{v \in \bar{\mathcal{V}}} g_v}{g_{\bar{\mathcal{V}}}}$ denote the relative maximum difference of g_v in $\bar{\mathcal{V}}$ with respect to the average $g_{\bar{\mathcal{V}}}$. Similarly, we also

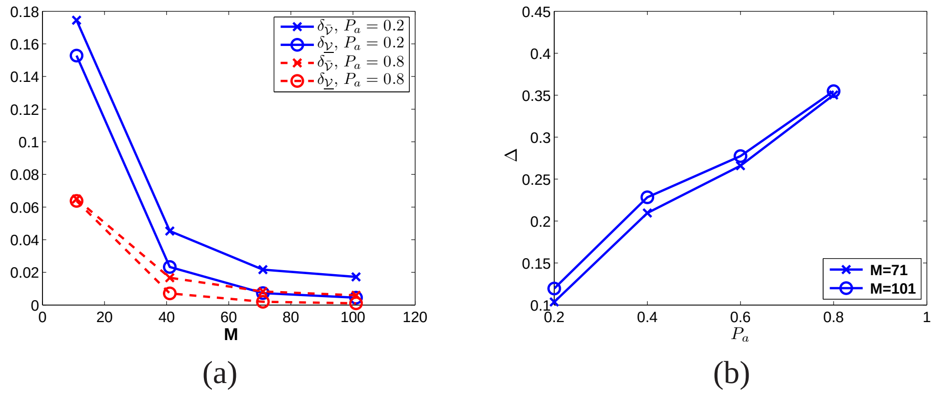


Fig. 4.5. (a) $\delta_{\underline{\mathcal{V}}}$ and $\delta_{\bar{\mathcal{V}}}$ with different M. (b) Δ with different P_a.

define $\delta_{\underline{\mathcal{V}}} \triangleq \frac{\max_{v\in\underline{\mathcal{V}}} g_v - \min_{v\in\underline{\mathcal{V}}} g_v}{g_{\underline{\mathcal{V}}}}$. Fig. 4.5 plot $\delta_{\underline{\mathcal{V}}}$ and $\delta_{\bar{\mathcal{V}}}$ with different M from 11 to 101. In this figure, we have $N = 10$ users, the forgetting factor is $\eta = 0.95$. We test the probability to switch to the anchor view $P_a = 0.2$ and 0.8. We observe that for both $P_a = 0.2$ and 0.8, when the number of views M increases, $\delta_{\underline{\mathcal{V}}}$ and $\delta_{\bar{\mathcal{V}}}$ decreases. For example, with $P_a = 0.8$ and $M = 101$, we have $\delta_{\underline{\mathcal{V}}} = 0.001$ and $\delta_{\bar{\mathcal{V}}} = 0.006$. In the following analysis, we consider the scenario where M is large and the difference of g_v in the same set $\underline{\mathcal{V}}$ (or $\bar{\mathcal{V}}$) is very small and can be ignored. Thus, $g_{\bar{\mathcal{V}}}$ and $g_{\underline{\mathcal{V}}}$ can be used to denote the g_v from $\underline{\mathcal{V}}$ and $\bar{\mathcal{V}}$, respectively. We then define $\Delta \triangleq \frac{g_{\bar{\mathcal{V}}} - g_{\underline{\mathcal{V}}}}{g_{\bar{\mathcal{V}}} + g_{\underline{\mathcal{V}}}}$ as the relative difference of $g_{\bar{\mathcal{V}}}$ and $g_{\underline{\mathcal{V}}}$ with respect to $g_{\bar{\mathcal{V}}} + g_{\underline{\mathcal{V}}}$. Fig. 4.5b shows Δ with different P_a. We observe that for a larger P_a, the difference between $g_{\bar{\mathcal{V}}}$ and $g_{\underline{\mathcal{V}}}$ is even larger. This is because when P_a is larger, users at views in $\underline{\mathcal{V}}$ have a higher probability to switch to the anchor view, which does not require others' help and thus results in a lower expected short-term gain, $g_{\underline{\mathcal{V}}}$. Meanwhile, with a larger P_a, users at views in $\bar{\mathcal{V}}$ have a higher probability to switch to the non-adjacent anchor view, and also have a higher probability to find a helper, since other users are more likely at the anchor view. Thus, it gives a higher expected short-term gain, $g_{\bar{\mathcal{V}}}$. Therefore, a larger P_a gives a larger Δ.

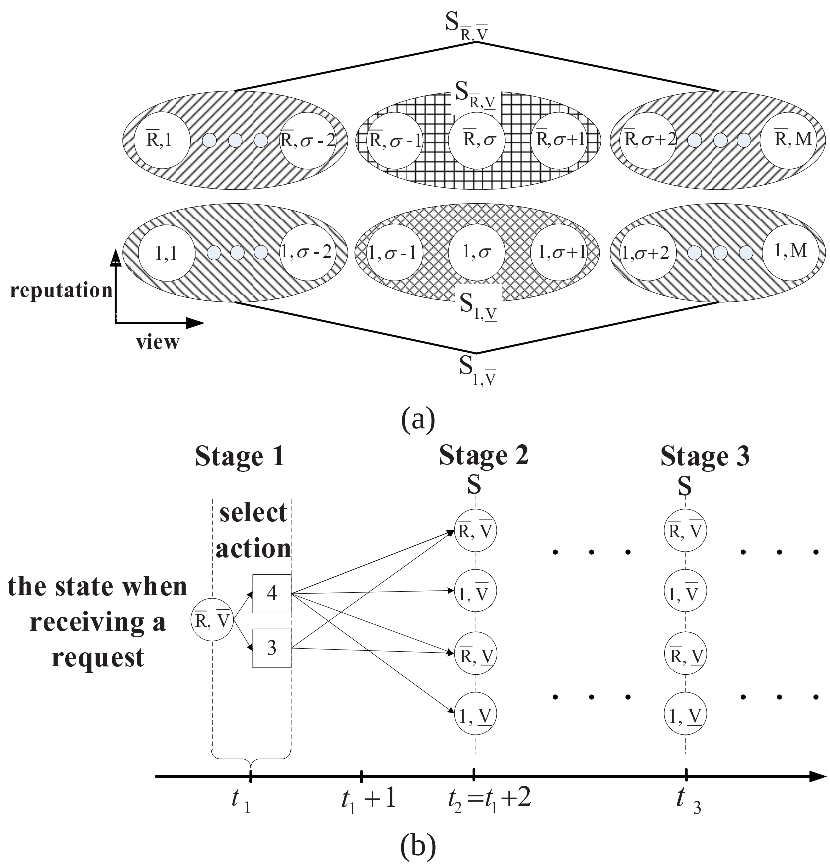


Fig. 4.6. Example of the state classification and aggregation with M views and 3-level reputation system ($t_r = 3$). (a) The state classification, where we have 4 state subspace $\mathcal{S}_{1,\underline{\mathcal{V}}}$, $\mathcal{S}_{1,\bar{\mathcal{V}}}$, $\mathcal{S}_{\bar{\mathcal{R}},\underline{\mathcal{V}}}$ and $\mathcal{S}_{\bar{\mathcal{R}},\bar{\mathcal{V}}}$. (b) The MDP after state aggregation, where we have only 4 states in the state space. The action can only be selected from $\{t_r, R+1\} = \{3, 4\}$.

4.3.1.3 State Aggregation

With the above observation, we can aggregate more states in the MDP to further simplify the analysis. We first classify the state space. For any $r \leq t_r - 2$, we define $\mathcal{S}_{r,\underline{\mathcal{V}}} = \{(r, v)|v \in \underline{\mathcal{V}}\}$ and $\mathcal{S}_{r,\bar{\mathcal{V}}} = \{(r, v)|v \in \bar{\mathcal{V}}\}$. We then define $\mathcal{S}_{\bar{\mathcal{R}},\underline{\mathcal{V}}} = \{(\bar{\mathcal{R}}, v)|v \in \underline{\mathcal{V}}\}$ and $\mathcal{S}_{\bar{\mathcal{R}},\bar{\mathcal{V}}} = \{(\bar{\mathcal{R}}, v)|v \in \bar{\mathcal{V}}\}$. Those state subspaces are non-overlapping, and $\mathcal{S}_{\bar{\mathcal{R}},\underline{\mathcal{V}}} \bigcup \mathcal{S}_{\bar{\mathcal{R}},\bar{\mathcal{V}}} \bigcup_{r \leq t_r - 2} \left(\mathcal{S}_{r,\underline{\mathcal{V}}} \bigcup \mathcal{S}_{r,\bar{\mathcal{V}}}\right) = \mathcal{S}$. Fig. 4.6a shows an example of the state classification with M views and 3-level reputation system, where $t_r = 3$ and $\bar{\mathcal{R}} = \{2, 3\}$. In Figure 4.6a, there are four non-overlapping state subspaces $\mathcal{S}_{1,\underline{\mathcal{V}}}$, $\mathcal{S}_{1,\bar{\mathcal{V}}}$, $\mathcal{S}_{\bar{\mathcal{R}},\underline{\mathcal{V}}}$ and $\mathcal{S}_{\bar{\mathcal{R}},\bar{\mathcal{V}}}$.

We then have the following proposition with proof in Appendix H.

Proposion 4.4. *With the above state classification, states in each subspace have bisimilarity relationship and can be aggregated as one state.*

From Proposition 4.4, all states in the same subspace can be aggregated into one state. Thus, for the example in Fig. 4.6a, there are four aggregated states denoted as $(1, \underline{\mathcal{V}})$, $(1, \bar{\mathcal{V}})$, $(\bar{\mathcal{R}}, \underline{\mathcal{V}})$ and $(\bar{\mathcal{R}}, \bar{\mathcal{V}})$, and Fig. 4.6b shows the updated MDP after state aggregation. From Theorem 4.1 and the discussion in Chapter 4.2.2, for the aggregated state with reputation $r < t_r - 1$, users will always take action $a = R + 1$ and do not cooperate with anyone. Therefore, we only need to consider the aggregated states, $(\bar{\mathcal{R}}, \underline{\mathcal{V}})$ and $(\bar{\mathcal{R}}, \bar{\mathcal{V}})$, and let $a_{\underline{\mathcal{V}}}$ and $a_{\bar{\mathcal{V}}}$ denote actions taken at these two aggregated states, respectively.

The next step is to study the state transition probability for the aggregated states. Note that the reputation and view transition probabilities are independent, and the reputation transition probabilities are the same as in Chapter 4.2.2.1. Therefore, we only need to analyze the updated view transition probabilities. Note that given the one-step view transition matrix in (4.21), starting from any view $v \in \mathcal{V}$, after one segment, it will transit to views in $\underline{\mathcal{V}}$ with the same probability $\sum_{v' \in \underline{\mathcal{V}}} \mathbf{T}(v, v') = P_a + \frac{2(1-P_a)}{M-1}$, and to views in $\bar{\mathcal{V}}$ with the same probability $\sum_{v' \in \bar{\mathcal{V}}} \mathbf{T}(v, v') = 1 - P_a - \frac{2(1-P_a)}{M-1}$. Therefore, with the aggregated states, the one-step view transition probability is denoted as $\mathbf{T}\left(\underline{\mathcal{V}}, \underline{\mathcal{V}}\right) = P_a + \frac{2(1-P_a)}{M-1}$, which is the probability that a user transits from $\underline{\mathcal{V}}$ (i.e., from any view in $\underline{\mathcal{V}}$) to views in $\underline{\mathcal{V}}$. Similarly, we also have $\mathbf{T}\left(\bar{\mathcal{V}}, \underline{\mathcal{V}}\right) = P_a + \frac{2(1-P_a)}{M-1}$ and $\mathbf{T}\left(\underline{\mathcal{V}}, \bar{\mathcal{V}}\right) = \mathbf{T}\left(\bar{\mathcal{V}}, \bar{\mathcal{V}}\right) = 1 - P_a - \frac{2(1-P_a)}{M-1}$. Also, with the aggregated states, the steady-state view distribution is $\mathbf{v}\left(\underline{\mathcal{V}}\right) = \mathbf{T}\left(v, \underline{\mathcal{V}}\right) = P_a + \frac{2(1-P_a)}{M-1}$ and $\mathbf{v}\left(\bar{\mathcal{V}}\right) = \mathbf{T}\left(v, \bar{\mathcal{V}}\right) = 1 - P_a - \frac{2(1-P_a)}{M-1}$ for any $v \in \{\underline{\mathcal{V}}, \bar{\mathcal{V}}\}$.[2] Therefore, with the aggregated states, the state transition probability is $P^a_{(r,v)\to(r',v')} = P^a_{r\to r'}\mathbf{v}(v')$ for all $r, r' \in \{1, \cdots, t_r - 2, \bar{\mathcal{R}}\}$ and $v, v' \in \{\underline{\mathcal{V}}, \bar{\mathcal{V}}\}$.

In the example in Fig. 4.6b, given the current state $(\bar{\mathcal{R}}, \bar{\mathcal{V}})$ at stage 1, from Theorem 4.1, the possible equilibrium actions are $a = t_r = 3$ and $a = t_r + 1 = 4$. When taking action $a = 3$, the user follows the social norm, and his/her reputation stays at $\bar{\mathcal{R}}$ with probability $P^{a=3}_{\bar{\mathcal{R}}\to\bar{\mathcal{R}}} = 1$, and the state transition probabilities are $P^{a=3}_{(\bar{\mathcal{R}},\bar{\mathcal{V}})\to(\bar{\mathcal{R}},\underline{\mathcal{V}})} = \mathbf{v}(\underline{\mathcal{V}})$, $P^{a=3}_{(\bar{\mathcal{R}},\bar{\mathcal{V}})\to(\bar{\mathcal{R}},\bar{\mathcal{V}})} = \mathbf{v}(\bar{\mathcal{V}})$, and $P^{a=3}_{(\bar{\mathcal{R}},\bar{\mathcal{V}})\to(1,v')} = 0$ for any $v' \in \{\underline{\mathcal{V}}, \bar{\mathcal{V}}\}$. Similarly, we can derive the rest of the state transition probabilities.

The next step is to derive the expected short-term utility of the updated MDP after state aggregation. From the proof of Proposition 4.4, all views in the same view set $\underline{\mathcal{V}}$ (or $\bar{\mathcal{V}}$) give the same expected short-term utility. Define $U^{a_{\underline{\mathcal{V}}}}_{\underline{\mathcal{V}}} \triangleq U^{a_{\underline{\mathcal{V}}}}_v$ for any $v \in \underline{\mathcal{V}}$, and $U^{a_{\bar{\mathcal{V}}}}_{\bar{\mathcal{V}}} \triangleq U^{a_{\bar{\mathcal{V}}}}_v$ for any $v \in \bar{\mathcal{V}}$.

After this state aggregation, the lifetime utility in (4.17) can be written as

$$\begin{cases} W^{\pi}_{\underline{\mathcal{V}}} = U^{a_{\underline{\mathcal{V}}}}_{\underline{\mathcal{V}}} + \left[\eta^L(1 - P^{a_{\underline{\mathcal{V}}}}_{\bar{\mathcal{R}}\to1}) + \gamma P^{a_{\underline{\mathcal{V}}}}_{\bar{\mathcal{R}}\to1}\right]\left[\mathbf{v}(\underline{\mathcal{V}})W^{\pi}_{\underline{\mathcal{V}}} + \mathbf{v}(\bar{\mathcal{V}})W^{\pi}_{\bar{\mathcal{V}}}\right], \\ W^{\pi}_{\bar{\mathcal{V}}} = U^{a_{\bar{\mathcal{V}}}}_{\bar{\mathcal{V}}} + \left[\eta^L(1 - P^{a_{\bar{\mathcal{V}}}}_{\bar{\mathcal{R}}\to1}) + \gamma P^{a_{\bar{\mathcal{V}}}}_{\bar{\mathcal{R}}\to1}\right]\left[\mathbf{v}(\underline{\mathcal{V}})W^{\pi}_{\underline{\mathcal{V}}} + \mathbf{v}(\bar{\mathcal{V}})W^{\pi}_{\bar{\mathcal{V}}}\right], \end{cases} \tag{4.23}$$

and we only need to study the action policy $\pi = \{a_{\underline{\mathcal{V}}}, a_{\bar{\mathcal{V}}}\}$ for $r \in \bar{\mathcal{R}}$.

4.3.1.4 Equilibrium Analysis with 2-Level Reputation System

In this section, we consider a simple scenario with a 2-level reputation system (i.e., $R = 2$), and derive the equilibrium action policies of the game. With $R = 2$, if $t_r = 1$, users with the lowest reputation can also get help, which discourages users to cooperate with each other. Therefore, $1 < t_r \leq R = 2$ and t_r can only be 2. Note that the 2-level reputation system is memoryless. This is because if a

[2] Here, we still use $\mathbf{v}$ to denote the steady state view distribution over those two view sets for notation simplicity.

user's behavior complies with the social norm, his/her reputation is updated to 2. Otherwise, his/her reputation is updated to 1 regardless of his/her past reputation.

Note that with $t_r = R = 2$, $\bar{\mathcal{R}} = \{r \geq t_r - 1\} = \{1, 2\} = \mathcal{R}$. Therefore, after state aggregation in Chapter 4.3.1.3, there are only two aggregated state $(\bar{\mathcal{R}}, \underline{\mathcal{V}})$ and $(\bar{\mathcal{R}}, \bar{\mathcal{V}})$ with corresponding action $a_{\underline{\mathcal{V}}}$ and $a_{\bar{\mathcal{V}}}$, respectively. Here, $a_{\underline{\mathcal{V}}}, a_{\bar{\mathcal{V}}} \in \{t_r, R+1\} = \{2, 3\}$, where $a = 2$ means cooperation with beneficial users, and $a = 3$ means no cooperation with anyone. Thus, we have 4 possible action policies $\{a_{\underline{\mathcal{V}}} = 2, a_{\bar{\mathcal{V}}} = 2\}$, $\{a_{\underline{\mathcal{V}}} = 2, a_{\bar{\mathcal{V}}} = 3\}$, $\{a_{\underline{\mathcal{V}}} = 3, a_{\bar{\mathcal{V}}} = 2\}$ and $\{a_{\underline{\mathcal{V}}} = 3, a_{\bar{\mathcal{V}}} = 3\}$. By examining each of them, we have the following Proposition 4.5, and the proof is in Appendix I.

Proposion 4.5. *For an IMVS with a single anchor view and 2-level reputation system,*

a) *If $g_{\underline{\mathcal{V}}} \geq c$, $\{a_{\underline{\mathcal{V}}}, a_{\bar{\mathcal{V}}}\} = \{2, 2\}$ is an equilibrium policy, where users cooperate at all views (*full cooperation*).*

b) *If $g_{\bar{\mathcal{V}}}\mathbf{v}(\bar{\mathcal{V}}) \geq c \geq g_{\underline{\mathcal{V}}}\mathbf{v}(\bar{\mathcal{V}})$, $\{a_{\underline{\mathcal{V}}}, a_{\bar{\mathcal{V}}}\} = \{3, 2\}$ is an equilibrium policy, where users only cooperate at views in $\bar{\mathcal{V}}$ with high expected short-term gains but not at views in $\underline{\mathcal{V}}$ with low expected short-term gains (*partial cooperation*).*

c) *$\{a_{\underline{\mathcal{V}}}, a_{\bar{\mathcal{V}}}\} = \{3, 3\}$ is always an equilibrium policy, where users do not cooperate at all (*no cooperation*).*

d) *$\{a_{\underline{\mathcal{V}}}, a_{\bar{\mathcal{V}}}\} = \{2, 3\}$ is not an equilibrium policy.*

From Proposition 4.5, there are multiple Nash Equilibriums coexisting. In addition, from Proposition 4.5.a, with a 2-level reputation system, users cooperate at views in $\underline{\mathcal{V}}$ only when $g_{\underline{\mathcal{V}}} \geq c$. This is because the 2-level reputation system is memoryless, and users decide their actions only based on the expected short-term utility. If cooperation at views in $\underline{\mathcal{V}}$ gives a negative expected short-term utility ($g_{\underline{\mathcal{V}}} < c$), users will not cooperate.

4.3.1.5 Equilibrium Analysis with R-level ($R \geq 3$) Reputation System

The R-level ($R \geq 3$) reputation system is non-memoryless, and users need to take their future utilities into consideration. As discussed in the previous section, we only need to study the policy $\{a_{\underline{\mathcal{V}}}, a_{\bar{\mathcal{V}}}\}$ for reputation $r \geq t_r - 1$, where $a_{\underline{\mathcal{V}}}, a_{\bar{\mathcal{V}}} \in \{t_r, R+1\}$. Thus, we also have 4 possible policies $\{a_{\underline{\mathcal{V}}} = t_r, a_{\bar{\mathcal{V}}} = t_r\}$, $\{a_{\underline{\mathcal{V}}} = t_r, a_{\bar{\mathcal{V}}} = R+1\}$, $\{a_{\underline{\mathcal{V}}} = R+1, a_{\bar{\mathcal{V}}} = t_r\}$ and $\{a_{\underline{\mathcal{V}}} = R+1, a_{\bar{\mathcal{V}}} = R+1\}$. By examining each of them, we have the following proposition.

Proposion 4.6. *For an IMVS with a single anchor view and R-level reputation system where $R \geq 3$,*

a) If $\bar{c}_1 \triangleq \dfrac{(\eta^L - \gamma)(g_{\bar{\mathcal{V}}} - g_{\underline{\mathcal{V}}})\mathbf{v}(\bar{\mathcal{V}})}{1-\gamma} + g_{\underline{\mathcal{V}}} \geq c$, $\{a_{\underline{\mathcal{V}}}, a_{\bar{\mathcal{V}}}\} = \{t_r, t_r\}$ *is an equilibrium policy, where users cooperate at all views (*full cooperation*).*

b) If

$$\begin{aligned} & \frac{[y + \mathbf{x}(t_r - 1)]\mathbf{v}(\bar{\mathcal{V}}) \left[(1-\gamma)g_{\bar{\mathcal{V}}} - (\eta^L - \gamma)\mathbf{v}(\underline{\mathcal{V}})(1-y)(g_{\bar{\mathcal{V}}} - g_{\underline{\mathcal{V}}})\right]}{1 - \eta^L + \eta^L y - y\gamma} \\ \geq \;& c \\ \geq \;& \frac{[y + \mathbf{x}(t_r - 1)]\mathbf{v}(\bar{\mathcal{V}}) \left[\mathbf{v}(\bar{\mathcal{V}})(\eta^L - \gamma)(g_{\bar{\mathcal{V}}} - g_{\underline{\mathcal{V}}}) + (1-\gamma)g_{\underline{\mathcal{V}}}\right]}{1 - \eta^L + \eta^L y - y\gamma}, \end{aligned} \tag{4.24}$$

$\{a_{\underline{\mathcal{V}}}, a_{\bar{\mathcal{V}}}\} = \{R+1, t_r\}$ *is an equilibrium policy, where users cooperate at views in $\bar{\mathcal{V}}$ with high expected short-term gains but not at views in $\underline{\mathcal{V}}$ with low expected short-term gains (*partial cooperation*).*

c) $\{a_{\underline{\mathcal{V}}}, a_{\bar{\mathcal{V}}}\} = \{R+1, R+1\}$ *is always an equilibrium policy, where users do not cooperate at all (*no cooperation*).*

d) $\{a_{\underline{\mathcal{V}}}, a_{\bar{\mathcal{V}}}\} = \{t_r, R+1\}$ *is not an equilibrium policy.*

Proof: In the following, we will prove Proposition 4.6.a, and the rest of the proof is in Appendix J.

For the policy $\{a_{\underline{\mathcal{V}}}, a_{\bar{\mathcal{V}}}\} = \{t_r, t_r\}$, we determine when it is an equilibrium. To do this, we first assume that all users use this policy and study the corresponding stationary reputation distribution $\mathbf{x}$ following the discussion in Chapter 4.2.3. Then, using the one-shot deviation principle, we exam whether a user has incentives to unilaterally deviate to any one-shot deviation at any view.

As discussed in Chapter 4.2.3, by solving (4.20), we have $y = 1$ and $\mathbf{x}(1) = \cdots = \mathbf{x}(t_r - 1) = 0$. This is because all users keep cooperating with beneficial users and thus have the highest reputation R.

We then exam the one-shot deviation principle. First, the given policy is $a_{\underline{\mathcal{V}}} = a_{\bar{\mathcal{V}}} = t_r$. When a user receives a request at $v \in \{\underline{\mathcal{V}}, \bar{\mathcal{V}}\}$, by taking action $a_v = t_r$ following the given policy, he/she will upload the requested frame with probability 1, since all other users have reputation R. Thus, the expected immediate cost is $C^{a=t_r} = c$. In addition $a_v = t_r$ complies with the social norm in (4.3), the user's reputation will be lowered to 1 with probability $P^{t_r}_{\bar{\mathcal{R}}\to 1} = 0$, and he/she is a beneficial user with probability 1. Since others also take policy $a_{\underline{\mathcal{V}}} = a_{\bar{\mathcal{V}}} = t_r$, he/she will always receive others' help and have the maximum expected short-term gain g_v for $v \in \{\underline{\mathcal{V}}, \bar{\mathcal{V}}\}$. Therefore, with action $a_v = t_r$, his/her expected short-term utility is $U_v^{a=t_r} = -c + g_v$ for $v \in \{\underline{\mathcal{V}}, \bar{\mathcal{V}}\}$. Thus, with the policy $a_{\underline{\mathcal{V}}} = a_{\bar{\mathcal{V}}} = t_r$, the lifetime utility (4.23) becomes

$$\begin{cases} W^{\pi}_{\underline{\mathcal{V}}} = -c + g_{\underline{\mathcal{V}}} + \eta^L \left[\mathbf{v}(\underline{\mathcal{V}})W^{\pi}_{\underline{\mathcal{V}}} + \mathbf{v}(\bar{\mathcal{V}})W^{\pi}_{\bar{\mathcal{V}}}\right], \\ W^{\pi}_{\bar{\mathcal{V}}} = -c + g_{\bar{\mathcal{V}}} + \eta^L \left[\mathbf{v}(\underline{\mathcal{V}})W^{\pi}_{\underline{\mathcal{V}}} + \mathbf{v}(\bar{\mathcal{V}})W^{\pi}_{\bar{\mathcal{V}}}\right]. \end{cases} \tag{4.25}$$

Note that (4.25) is a linear system with two unknowns $W^{\pi}_{\underline{\mathcal{V}}}$ and $W^{\pi}_{\bar{\mathcal{V}}}$, which can be solved easily,

$$\begin{cases} W^{\pi}_{\underline{\mathcal{V}}} = \frac{g_{\underline{\mathcal{V}}} - c + \eta^L \mathbf{v}(\bar{\mathcal{V}})(g_{\bar{\mathcal{V}}} - g_{\underline{\mathcal{V}}})}{1-\eta^L}, \\ W^{\pi}_{\bar{\mathcal{V}}} = \frac{g_{\bar{\mathcal{V}}} - c - \eta^L \mathbf{v}(\underline{\mathcal{V}})(g_{\bar{\mathcal{V}}} - g_{\underline{\mathcal{V}}})}{1-\eta^L}, \end{cases} \tag{4.26}$$

Now we examine the user's lifetime utility if he/she takes one-shot deviation. As discussed in Chapter 4.2.2, action t_r and $R + 1$ dominate other strategies, and thus, we only need to exam the one-shot deviation to $R + 1$. First, with $a'_v = R + 1$ with $v \in \{\underline{\mathcal{V}}, \bar{\mathcal{V}}\}$, this user does not help anyone and the immediate cost is

0. Since all other users have reputation R, the action $a'_v = R+1$ makes his/her reputation lowered to 1 with probability $P_{\bar{r}\to 1}^{a'_v=R+1} = 1$. Thus, he/she cannot receive others' help in the following L segments, and the expected short-term gain is $G_{1,v} = 0$. Therefore, the expected short-term utility by one-shot deviation to $R+1$ is $U_v^{a'_v=R+1} = 0$. Thus, with one-shot deviation to $R+1$, the lifetime utility in (4.18) can be rewritten as

$$\begin{cases} W_{\underline{\mathcal{V}}}^{a'_v=R+1,\pi} = \gamma\left[\mathbf{v}(\underline{\mathcal{V}})W_{\underline{\mathcal{V}}}^{\pi} + \mathbf{v}(\bar{\mathcal{V}})W_{\bar{\mathcal{V}}}^{\pi}\right], \\ W_{\bar{\mathcal{V}}}^{a'_v=R+1,\pi} = \gamma\left[\mathbf{v}(\underline{\mathcal{V}})W_{\underline{\mathcal{V}}}^{\pi} + \mathbf{v}(\bar{\mathcal{V}})W_{\bar{\mathcal{V}}}^{\pi}\right] \end{cases} \tag{4.27}$$

Substitute (4.26) into (4.27) and compare W_v^{π} with $W_v^{a'_v=R+1,\pi}$ for $v \in \{\underline{\mathcal{V}}, \bar{\mathcal{V}}\}$. We have

$$\begin{aligned} W_{\bar{\mathcal{V}}}^{\pi} - W_{\bar{\mathcal{V}}}^{a'_v=R+1,\pi} &= \frac{(\eta^L-\gamma)(g_{\bar{\mathcal{V}}}-g_{\underline{\mathcal{V}}})\mathbf{v}(\bar{\mathcal{V}}) - (c-g_{\underline{\mathcal{V}}})(1-\gamma)}{1-\eta^L} + (g_{\bar{\mathcal{V}}}-g_{\underline{\mathcal{V}}}), \\ W_{\underline{\mathcal{V}}}^{\pi} - W_{\underline{\mathcal{V}}}^{a'_v=R+1,\pi} &= \frac{(\eta^L-\gamma)(g_{\bar{\mathcal{V}}}-g_{\underline{\mathcal{V}}})\mathbf{v}(\bar{\mathcal{V}}) - (c-g_{\underline{\mathcal{V}}})(1-\gamma)}{1-\eta^L}. \end{aligned} \tag{4.28}$$

It is easy to observe that $W_{\bar{\mathcal{V}}}^{\pi} - W_{\bar{\mathcal{V}}}^{a'_v=R+1,\pi} > W_{\underline{\mathcal{V}}}^{\pi} - W_{\underline{\mathcal{V}}}^{a'_v=R+1,\pi}$. Thus, as long as $W_{\underline{\mathcal{V}}}^{\pi} - W_{\underline{\mathcal{V}}}^{a'_v=R+1,\pi} = \frac{(\eta^L-\gamma)(g_{\bar{\mathcal{V}}}-g_{\underline{\mathcal{V}}})\mathbf{v}(\bar{\mathcal{V}}) - (c-g_{\underline{\mathcal{V}}})(1-\gamma)}{1-\eta^L} \geq 0$, i.e., $\bar{c}_1 \triangleq \frac{(\eta^L-\gamma)(g_{\bar{\mathcal{V}}}-g_{\underline{\mathcal{V}}})\mathbf{v}(\bar{\mathcal{V}})}{1-\gamma} + g_{\underline{\mathcal{V}}} \geq c$, we have $W_{\bar{\mathcal{V}}}^{\pi} - W_{\bar{\mathcal{V}}}^{a'_v=R+1,\pi} > W_{\underline{\mathcal{V}}}^{\pi} - W_{\underline{\mathcal{V}}}^{a'_v=R+1,\pi} \geq 0$, and $\{a_{\underline{\mathcal{V}}}, a_{\bar{\mathcal{V}}}\} = \{t_r, t_r\}$ is an equilibrium policy. This completes the proof of Proposition 4.6.a. ■

From Proposition 4.6, same as the single anchor view IMVS system with a 2-level reputation system, when $R \geq 3$, there are multiple Nash Equilibria coexisting. Also, comparing Proposition 4.6a and Proposition 4.5a for the conditions for full cooperation, we observe that in a R-level reputation system with $R \geq 3$, a user may still cooperate at views in $\underline{\mathcal{V}}$ when $g_{\underline{\mathcal{V}}} < c \leq \bar{c}_1$, where cooperation gives him/her a negative expected short-term utility. This is because the R-level reputation system is non-memoryless, and a user needs to consider his/her future utilities when making a decision. Although cooperation at $\underline{\mathcal{V}}$ gives a negative expected short-term utility, this cooperation help him/her maintain a high reputation and keep receiving others' help in future view switching. As long as the expected future gain can compensate his/her current loss, he/she will still cooperate.

From Proposition 4.5a, $\bar{c}_1$ plays an important role in cooperation stimulation, and a larger $\bar{c}_1$ allows a larger range of cost for users to have full cooperation as an equilibrium, i.e., provides more incentive for user cooperation. In $\bar{c}_1$, we have the term $\gamma = \eta^{(t_r-1)L}$, where t_r reflects the punishment a user will receive if he/she deviates from cooperation with a beneficial user, and t_r is determined by the reputation system. It is easy to show that $\partial\bar{c}_1/\partial t_r > 0$. Thus, the reputation system should select the highest $t_r = R$ (that gives the harshest punishment), to provide the most incentive and give the largest $\bar{c}$ to have full cooperation as an equilibrium.

4.3.2 Game Analysis with Multiple Anchor Views

For the general IMVS with multiple anchor view and R-level ($R \geq 3$) reputation system, similar to the analysis in Chapter 4.3.1.5, non-cooperation at all views is still an equilibrium, and partial cooperation and full cooperation may be equilibrium policies in certain scenarios. With a large view space $\mathcal{V}$ and different views with different g_v's, we have many partial cooperation policies, and the analysis for each partial cooperation policy is also complicated. Note that from the system designer's perspective, the full cooperation equilibrium makes all users cooperate whenever possible, minimizes the consumed upload bandwidth at the server's side, and thus is the desired equilibrium policy. In the following, we will derive the conditions for full cooperation to be an equilibrium policy in an IMVS with multiple anchor views.

Similar to the proof in Proposition 4.6.a, we first assume that all users take the full cooperation policy $\pi = \{a_1, a_2, ..., a_M\} = \{t_r, t_r, ..., t_r\}$ and derive the corresponding reputation distribution $\mathbf{x}$. We then exam whether π can resist the one-shot deviation to $a'_v = R + 1$ for any $v \in \mathcal{V}$.

If all users cooperate with the policy π, they will keep the highest reputation R, and the reputation distribution is $y = 1$ and $\mathbf{x}(1) = \cdots \mathbf{x}(t_r - 1) = 0$. For a user receiving a request at view $v \in \mathcal{V}$, he/she will help upload with probability 1 by following π. Thus, the expected immediate cost is c. In addition, with $P_{\mathcal{R}\to 1}^{t_r} = 0$, his/her reputation remains to be R. Therefore, he/she always receives others' help, and therefore receives the maximum expected short-term gain g_v. Thus, his/her expected short-term utility is $U_v^{a_v=t_r} = -c + g_v$, and his/her lifetime utility is

$$W_v^\pi = -c + g_v + \eta^L \sum_{v'=1}^{M} \mathbf{T}^L(v,v') W_{v'}^\pi, \forall v \in \mathcal{V}. \tag{4.29}$$

Here, the only difference between (4.29) and (4.25) is that the summation term in (4.29) is over all M views instead of 2 aggregated view sets in (4.25). To solve (4.29), we expand the recursive term $W_{v'}^\pi$ at the right side of (4.29) and have

$$\begin{aligned}
W_v^\pi &= -c + g_v + \eta^L \sum_{v'=1}^{M} \mathbf{T}^L(v,v')(-c + g_{v'}) \\
&\quad + \eta^{2L} \sum_{v'=1}^{M} \sum_{v''=1}^{M} \mathbf{T}^L(v,v')\mathbf{T}^L(v',v'') W_{v''}^\pi \\
&= -c + g_v + \eta^L \sum_{v'=1}^{M} \mathbf{T}^L(v,v')(-c + g_{v'}) + \eta^{2L} \sum_{v''=1}^{M} \mathbf{T}^{2L}(v,v'') W_{v''}^\pi \\
&= \cdots = -c + g_v + \sum_{n=1}^{\infty} \eta^{nL} \sum_{v'=1}^{M} \mathbf{T}^{nL}(v,v')(-c + g_{v'}) \\
&= \underbrace{g_v + \sum_{n=1}^{\infty} \eta^{nL} \sum_{v'=1}^{M} \mathbf{T}^{nL}(v,v') g_{v'}}_{\triangleq \mathbb{G}_v} - c - c \sum_{n=1}^{\infty} \eta^{nL} \left(\sum_{v'=1}^{M} \mathbf{T}^{nL}(v,v') \right) \\
&= \mathbb{G}_v - \frac{c}{1-\eta^L}.
\end{aligned} \tag{4.30}$$

In (4.30), $\mathbb{G}_v$ is the maximum lifetime gain a user can receive (when helpers always help him/her) if he/she starts view switching from view v, and $\frac{c}{1-\eta^L}$ is his/her lifetime cost to help others and upload frames whenever asked.[3] From (4.30), a necessary condition for the full cooperation policy $\pi = (t_r, \cdots, t_r)$ to be an equilibrium is to enable a non-negative lifetime utility with $W_v^\pi = \mathbb{G}_v - \frac{c}{1-\eta^L} \geq 0$ for all views, that is, $c \leq (1-\eta^L) \min_v \{\mathbb{G}_v\}$. Otherwise, users have no incentive to cooperate.

Similarly, we also derive the lifetime utility with one-shot deviation to $a'_v =$

[3] In (4.30), $\mathbb{G}_v$ includes an infinite series. Since $\eta < 1$, it is easy to show that this series converges, and $\mathbb{G}_v$ is finite, which users can calculate offline.

$R+1$

$$W_v^{a'_v=R+1,\pi} = \gamma \sum_{v'=1}^{M} \mathbf{T}^{L(t_r-1)}(v,v') W_{v'}^{\pi}. \tag{4.31}$$

We then substitute (4.30) into (4.31), compare W_v^{π} and $W_v^{a'_v=R+1,\pi}$, and have

$$\begin{aligned}
& W_v^{\pi} - W_v^{a'_v=R+1,\pi} \\
= \ & \mathbb{G}_v - \frac{c}{1-\eta^L} - \gamma \sum_{v'=1}^{M} \mathbf{T}^{L(t_r-1)}(v,v') \left(\mathbb{G}_{v'} - \frac{c}{1-\eta^L} \right) \\
= \ & \mathbb{G}_v - \gamma \sum_{v'=1}^{M} \mathbf{T}^{L(t_r-1)}(v,v') \mathbb{G}_{v'} - \frac{c}{1-\eta^L} + \gamma \sum_{v'=1}^{M} \mathbf{T}^{L(t_r-1)}(v,v') \frac{c}{1-\eta^L} \\
= \ & \mathbb{G}_v - \gamma \sum_{v'=1}^{M} \mathbf{T}^{L(t_r-1)}(v,v') \mathbb{G}_{v'} - \frac{1-\gamma}{1-\eta^L} c.
\end{aligned} \tag{4.32}$$

Therefore, if $c \le \frac{1-\eta^L}{1-\gamma} \min_v \left\{ \mathbb{G}_v - \gamma \sum_{v'=1}^{M} \mathbf{T}^{L(t_r-1)}(v,v') \mathbb{G}_{v'} \right\} \triangleq \bar{c}_2$ (i.e., $W_v^{\pi} - W_v^{a'_v=R+1,\pi} \ge 0$ for all v's and the one-shot deviation always gives a lower lifetime utility), together with the condition $c \le (1-\eta^L) \min_v \{\mathbb{G}_v\}$ being satisfied, the full cooperation policy $\pi = \{t_r, t_r, ..., t_r\}$ can resist any one-shot deviation and is an equilibrium policy.

4.4 Reputation System Optimization and Cooperation Initiation

In this section, we first study the optimal parameter selection for the reputation system to stimulate user cooperation as much as possible and to optimize the system performance with the minimum consumed bandwidth at the server's side. Given that there are more than one equilibrium policies in the game and the initial state of MDP determines the final equilibrium to which the game converges, we then propose a Pay-for-Cooperation (PfC) scheme to drive the game to the desired full cooperation equilibrium.

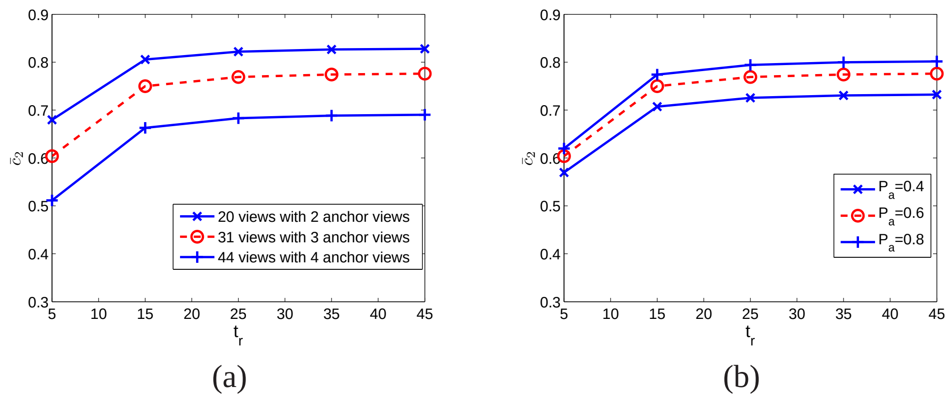


Fig. 4.7. $\bar{c}_2$ with t_r under different system setup. $R = 45$ level reputation system, $N = 10$ users and $\eta = 0.95$. (a) $P_a = 0.6$. M varies from 20 to 44 and the number of anchor views varies from 2 to 4. (b) $M = 31$ with 3 anchor views. P_a varies from 0.4 to 0.8.

4.4.1 Optimal t_r and R of The Reputation System

For single anchor view IMVS, as discussed in Chapter 4.3.1.5, a larger t_r punishes the non-cooperative behavior by a larger amount and provides more incentive for users to fully cooperate with each other. For the multi-anchor view IMVS, from the analysis in Chapter 4.3.2, the first necessary condition for full cooperation $c \leq (1 - \eta^L) \min_v \{\mathbb{G}_v\}$ does not depend on t_r or R. In the second condition, $\bar{c}_2$ is a function of t_r, which affects not only the term γ but also the summation term $\sum_{v'=1}^{M} \mathbf{T}^{L(t_r-1)}(v, v')\mathbb{G}_{v'}$, which makes the analysis difficult. In Fig. 4.7, we show the numerical results of $\bar{c}_2$ with t_r under different system setup. In this simulation, we have a $R = 45$ level reputation system, $N = 10$ users and the discounting factor is $\eta = 0.95$. In Fig. 4.7a, users switch to anchor views with a fixed probability $P_a = 0.6$, and we exam $\bar{c}_2$ with t_r under different numbers of views M and different numbers of anchor views. For example, if we have $M = 31$ views and 3 anchor views, the 3 anchor views are view 8, 16 and 24. Fig 4.7a shows $\bar{c}_2$ always increases with t_r under different M and different number of anchor views. In Fig. 4.7b, we have $M = 31$ views with 3 anchor views, and we exam $\bar{c}_2$ with t_r under different P_a. We also observe that $\bar{c}_2$ always increases with t_r for different P_a. Thus, similar to the IMVS with a single anchor view, a larger t_r also provides more incentive

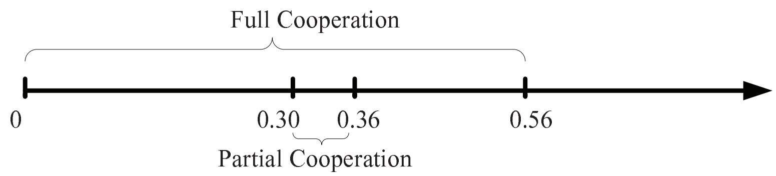


Fig. 4.8. An example of the coexistence of multiple equilibrium policies for a single anchor view IMVS.

for user cooperation and gives a larger $\bar{c}_2$ for full cooperation being an equilibrium policy in the multi-anchor view IMVS. Since t_r is no larger than R, the optimal t_r is R. Given that t_r is constrained by R, we should select the largest R, which is infinite. However, for a practical reputation system, R cannot and also does not have to be infinite. This is because from Fig. 4.7, we can observe that after $t_r = R = 25$, the increase of $\bar{c}_2$ for each step increase of $t_r = R$ is very small. For example, in Fig. 4.7b, when $P_a = 0.6$, the increase of $\bar{c}_2$ is 0.01 when $t_r = R$ increases from 25 to 45. Thus, a finite reputation system, e.g., $t_r = R = 25$ in this example, provides almost the same level of incentive for user cooperation when compared to the $t_r = R$ approaching infinite. Furthermore, if we set $t_r = R$ approaching infinite, a user who deviate once from the social norm will never have a chance to be a beneficial user again. However, if $t_r = R$ is a reasonably large number (e.g., 25 in Fig. 4.7), the system will punish a user harshly for his/her deviation from the social norm, while he/she still have the chance to recover his/her reputation and cooperate with others. Thus, this can accommodate the scenario, where the network error happens and the reputation system may punish users mistakenly. Based on the above analysis, a reasonably large $t_r = R$ provides enough high incentive for user cooperation and can also resist network errors, and therefore, should be selected.

4.4.2 Full Cooperation Initiation

From the discussion in the previous sections, for both single- and multi- anchor view IMVS, we may observe multiple equilibrium policies. Fig. 4.8 shows an example for a single-anchor view IMVS with $M = 101$ views, $N = 10$ users, $R = 10$ levels of reputations with $t_r = 10$. In the view switching model, users switch to the single anchor view with probability $P_a = 0.5$. The discounting factor is $\eta = 0.95$, and the expected short-term gains for the view set $\bar{\mathcal{V}}$ and $\underline{\mathcal{V}}$ are $g_{\bar{\mathcal{V}}} = 0.82$ and $g_{\underline{\mathcal{V}}} = 0.33$, respectively. From Fig. 4.8, when $c \in [0, 0.56]$, the full cooperation policy is an equilibrium. When $c \in [0.30, 0.36]$, the partial cooperation policy is an equilibrium. The non-cooperation policy is always an equilibrium for all $c \geq 0$. Thus, we have three equilibrium policies when $c \in [0.30, 0.36]$, and we have two equilibrium policies when $c \in [0, 0.30) \bigcup (0.36, 0.56]$.

We observe that the MDP's initial state is critical on the equilibrium to which the game will converge. For example, if no user cooperates, then a user who unilaterally cooperate receives a negative utility due to the cost of frame upload, and thus, is unwilling to cooperate. In this work, to initiate user cooperation, we propose a PfC scheme at the beginning of the game.

Specifically, *first*, the local agent assigns each user a reputation R at the beginning of the game. This is because users may cooperate only when they have reputation $r \geq t_r - 1$, and assigning each user the highest reputation R makes him/her have cooperation as an option.

Second, the local agent randomly selects y_{in} percentage of users, and pays them for their cooperation with other beneficial users. Here, for each frame upload, the payment is at least their cost c, and thus, cooperation with beneficial users becomes the weakly dominant strategy for the selected users. The local agent also announces y_{in} to the other unselected users, who are not paid for cooperation, to assist their decision making. As the game goes, users interact with each other and their reputations are updated using the social norm in (4.3). Once the local agent observes that all unselected users have started to cooperate, we call that cooperation is *initiated*. Then, the local agent will gradually stop paying the selected users one by one. Note that once a selected user is stopped from being paid, he/she has to estimate other users' actions and makes his/her own decision on whether to continue cooperation. If the local agent stops paying all selected users at the same time, each selected user may have different estimation of other selected users' actions. Thus, their behavior may be unpredictable, which may also affect the unselected users' cooperation.

The proposed strategy where the local agent stops paying the selected users one by one can avoid this problem. This is because if at a time only one selected user is stopped from being paid, he/she considers that the other users either still get paid for cooperation or have started to cooperate. In such a case, continuing cooperation is a dominant strategy for him/her.

In the proposed PfC scheme, the local agent wants to select a large enough y_{in} to initiate user cooperation, while it also tries to keep y_{in} as low as possible to minimize its payment in the initial period. In the following analysis, we study each unselected user's action selection at the beginning of the game, and derive a sufficient condition to guarantee cooperation initiation.

For an unselected user, since he/she does not know how the other unselected users behave at the beginning of the game, we consider the worst case scenario where he/she assumes that all other unselected users do not cooperate at all. Thus, an unselected user makes his/her decision based on the assumption that all selected users take the full cooperation policy $\pi_c = \{t_r, t_r, ..., t_r\}$ and all others take the non-cooperation policy $\pi_n = \{R+1, R+1, ..., R+1\}$. We then use the one-shot deviation principle to exam his/her action selection and study whether the cooperation policy π_c is his/her optimal policy.

Since all users have reputation R at the beginning, by taking action t_r, he/she will cooperate and upload a frame with probability 1 when requested and his/her reputation is lowered to 1 with probability $P^{t_r}_{\bar{\mathcal{R}}\to 1} = 0$. Furthermore, since he/she assumes that only the selected users cooperate, his/her expected short-term gain is $g_v y_{in}$. Thus, his/her expected short-term utility is $U_v^{t_r} = -c + g_v y_{in}$, and his/her lifetime utility using policy π_c is

$$W_v^{\pi_c} \quad = \quad -c + g_v y_{in} + \eta^L \sum_{v'=1}^{M} \mathbf{T}^L(v,v') W_{v'}^{\pi_c}. \tag{4.33}$$

Same as in (4.30), we also expand the recursive term $W_{v'}^{\pi_c}$ in (4.33), and have

$$W_v^{\pi_c} = -c + g_v y_{in} + \sum_{n=1}^{\infty} \eta^{nL} \sum_{v'=1}^{M} \mathbf{T}^{nL}(v,v')(-c + g_{v'} y_{in}) = y_{in}\mathbb{G}_v - \frac{c}{1-\eta^L}. \tag{4.34}$$

where $\mathbb{G}_v$ is defined in (4.30).

Similar to the proof of Proposition 4.6a, when studying the lifetime utility with one-shot deviation, we only need to study the one-shot deviation to $a_v' = R+1$. By taking action $R+1$, he/she will upload with probability 0 and his/her reputation is lowered to 1 with probability $P_{\mathcal{R}\to 1}^{t_r} = 1$. Thus, he/she cannot receive others' help in the next L segments and the expected short-term gain is 0. Therefore, his/her expected short term utility is also zero. Since his/her reputation is lowered to one, it takes $(t_r-1)L$ segments for his/her reputation to clime to $t_r - 1$ again. Following the same analysis in Chapter 4.2.2.2, during these $(t_r-1)L$ segments, he/she always receives zero expected short-term utilities. Thus, the lifetime utility with one-shot deviation to $a_v' = R+1$ is

$$W_v^{a_v'=R+1,\pi_c} = \gamma \sum_{v'=1}^{M} \mathbf{T}^{L(t_r-1)}(v,v') W_{v'}^{\pi_c}. \tag{4.35}$$

We then substitute (4.34) into (4.35), compare (4.34) and (4.35), and have

$$\begin{aligned} & W_v^{\pi_c} - W_v^{a_v'=R+1,\pi_c} \\ = \; & y_{in}\mathbb{G}_v - \frac{c}{1-\eta^L} - \gamma \sum_{v'=1}^{M} \mathbf{T}^{L(t_r-1)}(v,v') \left[y_{in}\mathbb{G}_{v'} - \frac{c}{1-\eta^L} \right] \\ = \; & y_{in}\mathbb{G}_v - \gamma \sum_{v'=1}^{M} \mathbf{T}^{L(t_r-1)}(v,v') y_{in}\mathbb{G}_{v'} - \frac{1-\gamma}{1-\eta^L} c. \end{aligned} \tag{4.36}$$

We can show that (4.36) is an increasing function of y_{in}. By solving $W_v^{\pi_c} - W_v^{a_v'=R+1,\pi_c} \geq 0$ for all views, we have

$$y_{in} \geq \frac{(1-\gamma)c}{(1-\eta^L)\min_{v\in\mathcal{V}} \left[\mathbb{G}_v - \gamma \sum_{v'=1}^{M} \mathbf{T}^{L(t_r-1)}(v,v')\mathbb{G}_{v'} \right]}. \tag{4.37}$$

If (4.37) is satisfied, π_c is the optimal strategy for all unselected users, and they will start cooperation immediately after the game begins. Then, the local agent can stop paying the selected users one by one. Since any y_{in} satisfying (4.37) can initiate user cooperation, we should select the smallest y_{in} satisfying (4.37) to minimize the local agent's payment. In (4.37), we observe that its right hand side (RHS) is an increasing function of c, and with a higher cost c, we need to select and pay more users at the beginning of the game. This is because for each unselected user, with a higher cost c, he/she requires more selected user to cooperate with him/her to compensate his/her cost for cooperation.

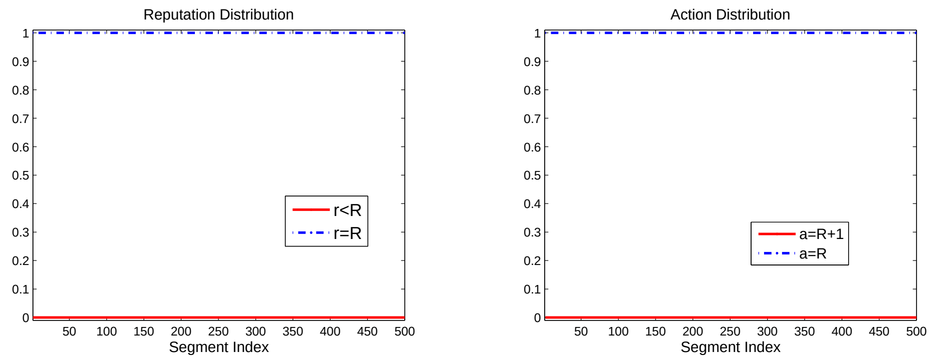
4.5 Simulation Results

This section evaluates the system performance by simulations. In the simulation setup, we have a $R = 10$ level reputation system and select the optimal $t_r = R$ as discussed in Chapter 4.4.1. The server provides IMVS with $M = 31$ views to a group of $N = 10$ users, and each user is assigned reputation 10 at the beginning of the game. In the view transition model, users switch to anchor views with probability $P_a = 0.5$. The discounting factor is $\eta = 0.95$. Since an IMVS with a single anchor view is a special case of that with multiple anchor views, in this section, we only show the results with multiple anchor views, and let view 8, 16 and 24 be the three anchor views.

4.5.1 Cooperation Initiation Verification

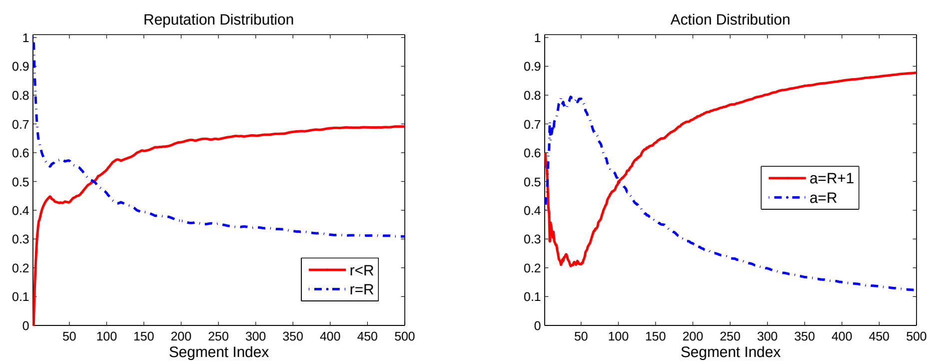
In this simulation, following the discussion in Chapter 4.3.2, we find that the condition for full cooperation to be an equilibrium policy is $c \leq 0.7$. We then select $c = 0.65 \leq 0.7$ as an example where full cooperation is an equilibrium policy. Following (4.37), the sufficient condition to initiate user cooperation is $y_{in} \geq 0.72$. In the following, we test different y_{in} to verify our theoretical analysis. In our experiments, once the local agent observes all unselected users have started to cooperate, it will stop paying selected users one by one, and the local agent will stop paying after the 50th segment in all scenarios.

Fig. 4.9 shows the simulation results when $y_{in} = 0.8$. Since $y_{in} = 0.8 > 0.72$ satisfies the condition (4.37), it can initiate user cooperation. Fig. 4.9b shows the percentage of users who use action $a = t_r = R$ (i.e. cooperation) when their reputations are no less than $t_r - 1$. We observe that all users cooperate, and thus they all have reputation $R = 10$ as shown in Fig. 4.9a. We then test $y_{in} = 0.5 < 0.72$,



(a) (b)

Fig. 4.9. The reputation and action distribution in the network, when $y_{in} = 0.8$. (a) The reputation distribution of the network. (b) The percentage of users that use action R (cooperation) and $R + 1$ (non-cooperation), respectively, when their reputation is no less than $t_r - 1$.



(a) (b)

Fig. 4.10. The reputation and action distribution in the network, when $y_{in} = 0.5$. (a) The reputation distribution of the network. (b) The percentage of users that use action R (cooperation) and $R + 1$ (non-cooperation), respectively, when their reputation is no less than $t_r - 1$.

and the results are in Fig. 4.10. We observe that it cannot initiate user cooperation. From Fig. 4.10b, we observe that after the 50th segment, the selected users are stopped from being paid, and percentage of users who take cooperative action starts to decrease. Thus, the probability of a user's reputation being R also drops as shown in Fig. 4.10a.

4.5.2 User Membership Dynamics

In the last section, we consider a fixed group of users interacting with each other. Once cooperation is initiated, they will fully cooperate until the end of this game. However, in a real video streaming system, users may join and leave the system from time to time, which may also affect their cooperation.

Consider a scenario where a group of existing users have been fully cooperating with each other and they all have reputation R. Then, some existing users may leave, while several new users join the system with assigned reputation R. Let $y_e \in [0, 1]$ denote the percentage of existing users after this membership change. Since existing users have established cooperative partnership, each of them assume other existing users still use π_c and keep cooperating. However, since they do not know how the new users will behave, we assume that they consider the worst case scenario, where new users use policy π_n and do not cooperate at all. We assume that each new user has the same assumption about existing users and other new users' behavior. We then exam whether an existing user will continue cooperation and whether a new user will start to cooperate. In fact, this problem is similar to that in Chapter 4.4.2. Here, all users have reputation R, and y_e percentage of users take policy π_c and the rest users take policy π_n. Following the same analysis, if y_e satisfies (4.37) for all views, then all users will cooperate. Let y_e^{min} denote the minimum value that can satisfy (4.37) for all views. If $y_e < y_e^{min}$ users may not cooperate after the membership update.

In this section, we test how user membership dynamics affect user cooperation. In this simulation, we let $c = 0.25$ so that full cooperation is an equilibrium policy. At the beginning of the game, we select $y_{in} = 0.8$ that is high enough to initiate user cooperation. For the membership dynamics, the initial number of users is 10. Users arrive the IMVS according to a Poisson process with an average arrival rate of λ users per segment duration. The sojourning period of each user follows an exponential distribution with an average of μ segments. Thus, a higher λ and a smaller μ result in more frequent membership update. In our simulations, we use batch join where new users can only join the streaming service at periodic moments, called batch moments. All new users coming between two neighboring batch moments will join and start receiving the streaming service at the same batch moment. In our simulations, the interval between neighboring batch moments is 30 segments, corresponding to a maximum of 10 seconds waiting time for a newly arrival user. For existing users in the streaming service, they can leave at any time instance. At each batch moment, the local agent will update the number of users and the percentage of existing users y_e, and broadcast to everyone.

From the previous analysis, at each batch moment, as long as $y_e \geq y_e^{min}$, users will still cooperate. Note that (4.37) includes the term $\mathbb{G}_v$, which is affected by the number of users. Thus, with user membership dynamics, y_e^{min} also changes. In the following, we test different λ and μ, and study how y_e impacts user cooperation at each batch moment.

Fig. 4.11 shows the simulation results with less frequent membership update with $\lambda = 0.1$ and $\mu = 100$. Fig. 4.11a gives the number of users at each time instance. Fig. 4.11b shows y_e^{min} and y_e at each batch moment. Since y_e is always higher than y_e^{min}, user cooperation will not be affected at each batch moment. Thus, from Fig. 4.11d users will always cooperate, and they will also maintain the reputation R as shown in Fig. 4.11c.

Fig. 4.12 shows the simulation results with more frequent membership update with $\lambda = 0.33$ and $\mu = 30$. Comparing Fig. 4.12a with Fig. 4.11a, we observe $\lambda = 0.33$ and $\mu = 30$ result in much more frequent membership update. From Fig. 4.12b, we observe that at the 90th segment, y_e is much smaller than y_e^{min}. Thus, user cooperation is interrupted after this batch moment. From Fig. 4.12d, we observe that users start to play non-cooperatively after the 90th segment, and

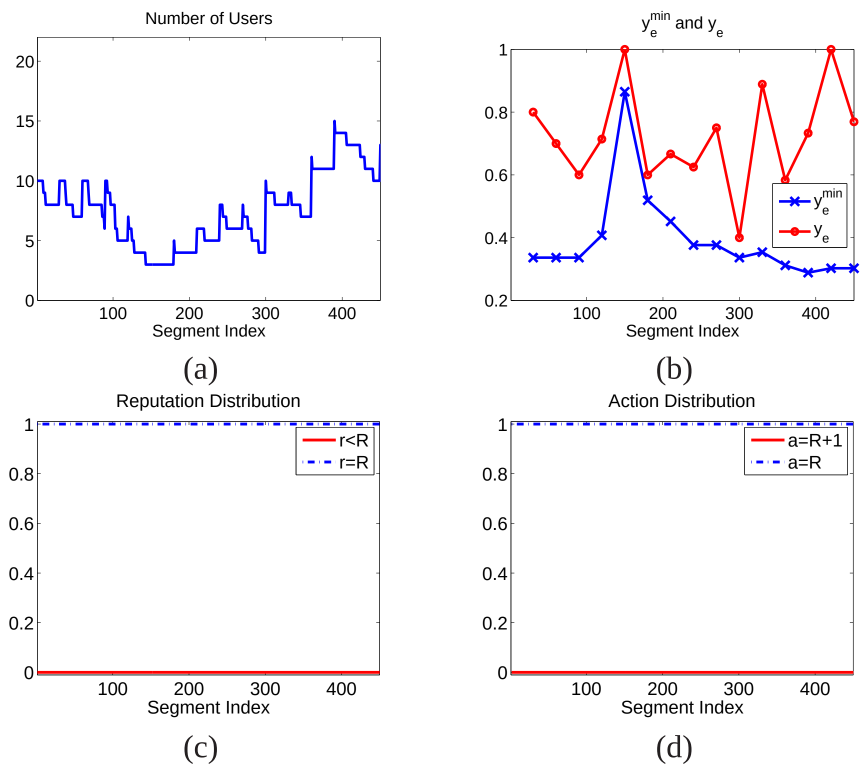


Fig. 4.11. The simulation results with low frequent membership update. $\lambda = 0.1$ and $\mu = 100$. (a) The number of users in the network. (b) y_e^{min} and y_e at each batch moment. (c) The reputation distribution. (d) The action distribution for users with reputation no less than $t_r - 1$.

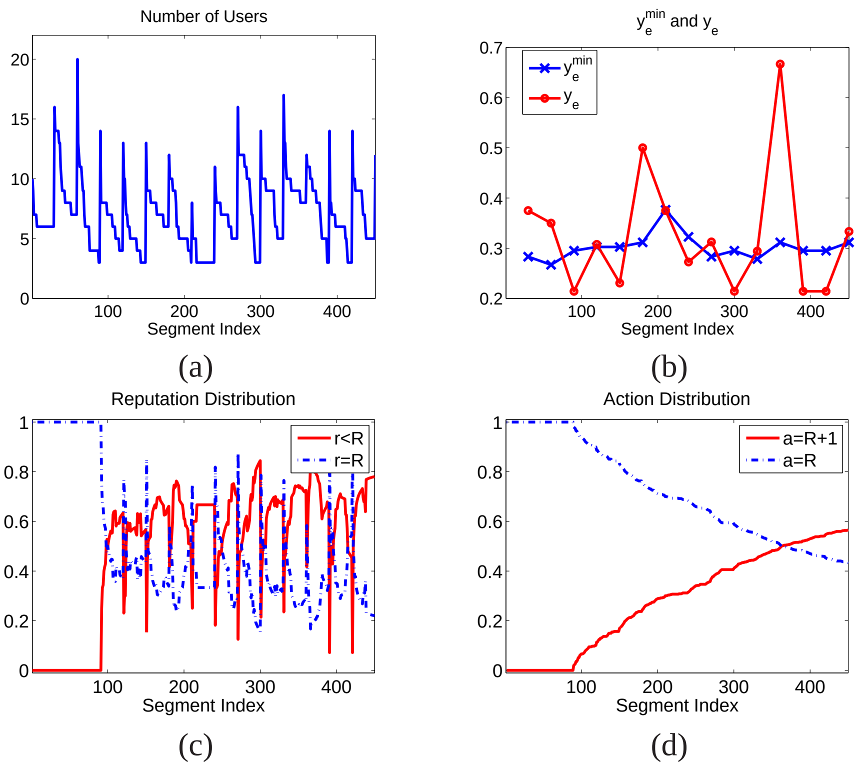


Fig. 4.12. The simulation results with high frequent membership update. $\lambda = 0.33$ and $\mu = 30$. (a) The number of users in the network. (b) y_e^{min} and y_e at each batch moment. (c) The reputation distribution. (d) The action distribution for users with reputation no less than $t_r - 1$.

the probability of a user's reputation being R also drops as shown in Fig. 4.12c. Since user cooperation has been interrupted, even if we have $y_e > y_e^{min}$ at a later batch time, users still do not cooperate. Here, the reputation fluctuation is due to the new users who are assigned with the highest reputation $R = 10$ when joining the system. Since they do not cooperate, the probability of their reputations being R decreases rapidly. To overcome the cooperation interruption due to the membership dynamics, the local agent should resume the PfC scheme, where it first reset all users reputation as R, and randomly selects y_{in} percentage users to pay for their cooperation with y_{in} satisfying (4.37), and announce y_{in} to other unselected users to assist their decision making.

4.6 Summary

In this work, we propose an IMVS system that supports cooperative view switching. To stimulate user cooperation, we model user interaction as an indirect reciprocity game. From the game analysis, we observe that users cooperate at some views but not others. Since peers can predict their future view navigation paths probabilistically, a peer likely to enter a view switching path not requiring others' help will receive low utility from cooperation, and thus has less incentive to cooperate. Furthermore, we observe that a larger number of reputation levels provides more incentive for user cooperation, and thus should be used. In addition, we observe that the game may have multiple equilibria with different cooperation levels. To initiate user cooperation, we propose a PfC scheme. Finally, we study how user membership dynamics affect user cooperation. We observe that as long as the percentage of new users is smaller than a predetermined threshold, users will continue cooperation. Otherwise, the PfC scheme should be used to resume user cooperation.

Chapter 5

Conclusions and Future Work

5.1 Conclusions

Multimedia sharing networks attract thousands of users all over the world to create and share multimedia data, and user behavior significantly impacts the system performance. Thus, the understanding of human factor provides important guidelines for designing a multimedia sharing network with satisfactory and efficient service. In this thesis, we model user interaction in such systems as games, and focus on designing incentive mechanisms to stimulate user cooperation. In this thesis, we investigate four challenging issues in the design of incentive mechanisms. i) One-to-many interaction, where users tend to free ride rather than cooperate since free riding is so easy. ii) User interaction with state change, where users take different strategies at different states, which affects the system performance. iii) Membership dynamics, which affect existing users' cooperation, since they do not know how the new users behave in the game. iv) Cheating on private information, which may reduce the system efficiency.

Specifically, we first study the incentive mechanism for a two-hop cooperative wireless multicast network. It is a typical example of one-to-many interaction, where one successful user relaying can help multiple unsuccessful users at the same time. We then model their interaction as a multi-seller multi-buyer payment based game, where unsuccessful users pay to receive relay service and the selected successful user will get paid if he/she helps forward packets. From the game analysis, we observe that at different prices, the game may converge to different Nash Equilibria, where unsuccessful users have different probabilities to free ride, resulting in different system throughput. In this work, we also study the optimal price selection, which drives the game to the desired Nash Equilibrium, where unsuccessful users have low free-riding probability and the system throughput is maximized. Therefore, the price is a very powerful tool in the payment based scheme, which can be exploited by the system designer to achieve desirable system performance.

We also address the issue that users have different cost to forward packets, and it is their private information. They may cheat if cheating can help them gain a higher payment. We then design a second-price sealed-bid auction game, which is a truth-telling auction, and users' dominant strategy is to bid their true cost.

We then investigate an IMVS supporting cooperative view switching. Since users may take different actions when they switch to different views, and they switch views frequently, it is a typical example to study user interaction with state change. We then analyze the game based on an MDP formulism, and observe that users cooperate at some views but not others. This is because peers can predict their future view navigation paths probabilistically. A peer likely to enter a view switching path not requiring others' help will receive less gain from cooperation, and thus, has less incentive to cooperate. To stimulate user cooperation at all views, we show that more reputation levels provide higher incentive for user cooperation, and thus, should be used. Furthermore, we observe that the game may have multiple Nash Equilibria corresponding to different cooperation levels. The final equilibrium the game will converge to depends on the initial cooperation level of the game. We then propose a PfC scheme to drive the game to the full cooperation equilibrium to improve the system efficiency.

In addition, we also study the impact of user membership dynamics on user cooperation and system performance. From our theoretical analysis and simulations, we observe that as long as the percentage of new users is smaller than a predetermined threshold, full cooperation is a dominant strategy for all users, and they will all cooperate. Otherwise, PfC should be used to stimulate user cooperation.

5.2 Future Work

Behavior modeling and analysis is still at its young age and there are many important and interesting problems that require further investigation.

In Chapter 3, we study user cooperation in a small circular region as shown in Fig. 3.1, where a moderate number of users (e.g. 25 users) help each other to forward video segments. This is because users close to each other have wireless channels with high capacity for cooperation. In cellular networks, a BS covers a big cellular area with hundreds of users, and it is not beneficial to let a pair of users who are far away from each other to cooperate. Thus, to apply our cooperative scheme in the entire cellular area, we need first find a mechanism to divide the cellular area into small circles, and apply our method in each circle. Furthermore, consider the case that a user is close to the boundary of his/her circle. He/she may overhear messages from neighboring circles as well, which may affect his/her cooperative behavior. Thus, it will be an interesting problem to study user behavior dynamics due to the boundary effect. Also, in Chapter 3, we assume this is local agent to facilitate the billing service and the auction. In our future work, we will also investigate how to select such a local agent in the cooperative network.

In Chapter 4, we study user cooperative behavior in an IMVS system. As introduced in [111], our IMVS system only supports *dynamic view switching*, where users are all synchronized in playback time. [111] proposes another IMVS system supporting *static view switching*, where each user can pause the video in time, and browse different views at the same time instance to gain better 3D visual experience. In this system, users are probably watching the video at different views and different time instances, which makes it even harder for user cooperation. To make users be able to cooperate with each other, each user can equip a buffer to store popular frames to assist others' downloading. Payment and reputation based incentive mechanisms may be applied to simulate user cooperation since peers change partners frequently. Also, in Chapter 4 we use a centralized reputation system, while in the literature, distributed reputation systems with more complicated reputation update rules are proposed. In our future work, we would like to investigate how other reputation systems affect the system performance.

In this thesis, we assume users are intelligent and selfish, who always try to maximize their own utilities. However, in reality, users may also take actions emotionally and irrationally, and some users may even would like to contribute voluntarily. In our future work, we also plan to investigate how such user behavior impact our system.

In online media sharing networks, such as YouTube and Flicker, users rely on the attached information of multimedia data, such as tags, comments and ratings, etc, to search and retrieve desired content for browsing. However, the noisy and spam information wildly existing in such systems may reduce the performance of the searching service. The work in [112] shows that only a half of tags in Flicker provide truly related information for images. Among all the imprecise tags, some of them are *unintentionally* made by careless users, while the rest are *intentionally* made by scammers. Usually, scammers make imprecise tags much more frequently than careless users. User trust modeling [113] may take advantage of this observation to identify scammers and reduce spam to improve the searching service.

References

[1] [Online], “http://www.flickr.com.”

[2] ——, “http://www.youtube.com.”

[3] ——, “http://www.pptv.com.”

[4] ——, “http://www.pps.tv.”

[5] ——, “http://www.sopcast.com.”

[6] X. Hei, C. Liang, J. Liang, Y. Liu, and K. W. Ross, “Insights into PPlive: A measurement study of a large-scale P2P iptv system,” *IPTV workshop in conjunction with WWW2006*, May 2006.

[7] J. L. Y. L. X. Hei, C. Liang and K. W. Ross, “A measurement study of a large-scale P2P iptv system,” *IEEE Transactions on Multimedia*, vol. 9, no. 8, pp. 1672–1687, Dec. 2007.

[8] R. Li, G. Gao, W. Xiao, and Z. Xu, “Measurement study on PPlive based on channel popularity,” *Communication Networks and Services Research Conference*, pp. 18 –25, May 2011.

[9] J. Liang, R. Kumar, Y. Xi, and K. Ross, “Pollution in P2P file sharing systems,” *in Proceedings of IEEE INFOCOM*, pp. 1174 – 1185, 2005.

[10] M. Hines, “Scammers gaming youtube ratings for profit,” *InfoWorld*, pp. 1–7, May 2007.

[11] W. Lin, H. V. Zhao, and K. J. R. Liu, “Incentive cooperation strategies for Peer-to-Peer live streaming social networks,” *IEEE Transactions on Multimedia*, vol. 11, no. 3, pp. 396–412, April 2009.

[12] S. Saroiu, P. K. Gummadi, , and S. D. Gribble, “A measurement study of peer-to-peer file sharing systems,” *in Proceedings of Multimedia Computing and Networking*, vol. 4673, Jan. 2002.

[13] “Understanding peer behavior and designing incentive mechanism in peer-to-peer networks: An analytical model based on game theory,” *in Proceedings of International Conference on Algorithms and Architectures for Parallel Processing*, 2007.

[14] G. Owen, *Game Theory*, 3rd ed. Academic Press, 1995.

[15] M. Osborne and A. Rubinste, *A Course in Game Theory*. The MIT Press, 1994.

[16] A. Kubota, A. Smolic, M. Magnor, M. Tanimoto, T. Chen, and C. Zhang, "Multi-view imaging and 3dtv," *IEEE Signal Processing Magazine*, vol. 24, no. 6, pp. 10–21, Nov. 2007.

[17] T. Fujii, K. Mori, K. Takeda, K. Mase, M. Tanimoto, and Y. Suenaga, "Multipoint measuring system for video and sound—100 camera and microphone system," *IEEE International Conference on Multimedia and Expo*, July 2006.

[18] Z. Liu, G. Cheung, and Y. Ji, "Distributed source coding for wwan multiview video multicast with cooperative peer-to-peer repair," *IEEE International Conference on Communications, Kyoto, Japan*, June 2011.

[19] E. van der Meulen, "Three terminal communication channels," *Advances in Applied Probability*, vol. 3, pp. 120–154, 1971.

[20] X. Hong, K. Xu, and M. Gerla, "Scalable routing protocols for mobile ad hoc networks," *IEEE Network*, vol. 16, no. 4, pp. 11–21, July 2002.

[21] M. Abolhasan, T. Wysocki, and E. Dutkiewicz, "A review of routing protocols for mobile ad hoc networks," *Ad Hoc Networks*, vol. 2, no. 1, pp. 1–22, Jan. 2004.

[22] T. M. Cover and A. E. Gamal, "Capacity theorems for the relay channel," *IEEE Transaction on Information Theory*, vol. 25, no. 5, pp. 572–584, Sept. 1979.

[23] J. N. Laneman, D. Tse, and G. Wornell, "Cooperative diversity in wireless networks: efficient protocols and outage behavior," *IEEE Transaction on Information Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.

[24] J. Cai, S. Shen, J. W. Mark, and A. S. Alfa, "Semi-distributed user relaying algorithm for amplify-and-forward wireless relay networks," *IEEE Transaction on Wireless Communications*, vol. 7, no. 4, pp. 1348–1357, April 2008.

[25] Y. Zhao, R. S. Adve, , and T. J. Lim, "Improving amplify-and-forward relay networks: optimal power allocation versus selection," *IEEE Transaction on Wireless Communications*, vol. 6, no. 8, pp. 3114–3123, Aug. 2007.

[26] S. Serbetli and A. Yener, "Relay assisted F/TDMA ad hoc networks: Node classification, power allocation and relaying strategies," *IEEE Transaction on Communications*, vol. 56, no. 6, pp. 937–947, June 2008.

[27] N. Zhou, X. Zhu, Y. Huang, and H. Lin, "Adaptive resource allocation for multi-destination relay systems based on OFDM modulation," *in Proceedings of IEEE International Conference on Communications, Dresden, Germany*, pp. 1–5, June 2009.

[28] G. A. Sidhu and F. Gao, "Resource allocation for relay aided uplink multiuser OFDMA system," *in Proceedings of IEEE Wireless Communications and Networking Conference*, pp. 1–5, April 2010.

[29] M. Chen and A. Yener, "Power allocation for F/TDMA multiuser two-way relay networks," *IEEE Transaction on Wireless Communications*, vol. 9, no. 2, pp. 546–551, Feb. 2010.

[30] M. Mauve, A. Widmer, and H. Hartenstein, "A survey on position-based routing in mobile ad hoc networks," *IEEE Network*, vol. 15, no. 6, pp. 30–39, Nov./Dec. 2001.

[31] C. Perkins, E. Belding-Royer, and S. Das, "Ad hoc on-demand distance vector (AODV) routing." RFC Editor, 2003.

[32] Y. Hu and A. Perrig, "A survey of secure wireless ad hoc routing," *IEEE Security Privacy*, vol. 2, no. 3, pp. 28–39, May-June 2004.

[33] Y. Hu, A. Perrig, and D. B. Johnson, "Rushing attacks and defense in wireless ad hoc network routing protocols," *in Proceedings of the 2nd ACM workshop on Wireless security*, pp. 30–40, 2003.

[34] ——, "Packet leashes: a defense against wormhole attacks in wireless networks," *in Proceedings of INFOCOM 2003*, vol. 3, pp. 1976–1986, Mar. 2003.

[35] S. Marti, T. J. Giuli, K. Lai, and M. Baker, "Mitigating routing misbehavior in mobile ad hoc networks," *in Proceedings of the 6th annual international conference on Mobile computing and networking*, pp. 255–265, 2000.

[36] G. Anastasi, M. Conti, M. Di Francesco, and A. Passarella, "Energy conservation in wireless sensor networks: A survey," *Ad Hoc Network*, vol. 7, no. 3, pp. 537–568, May 2009.

[37] B. Niu, H. Jiang, and H. V. Zhao, "A cooperative multicast strategy in wireless networks," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 6, pp. 3136–3143, July 2010.

[38] O. Alay, P. Liu, Z. Guo, Y. Wang, E. Erkip, and S. Panwar, "Cooperative layered video multicast using randomized distributed space time codes," *in Proceedings of IEEE INFOCOM*, pp. 176–181, 2009.

[39] X. Liu, G. Cheung, and C. Chuah, "Structured network coding and cooperative wireless ad hoc peer-to-peer repair for WWAN video broadcast," *IEEE Transactions on Multimedia*, vol. 11, no. 4, pp. 730–741, June 2009.

[40] H. V. Zhao and W. Su, "Cooperative wireless multicast: performance analysis and power/location optimization," *IEEE Transactions on Wireless Communications*, vol. 9, no. 6, pp. 2088–2100, June 2010.

[41] I. Güvenc, U. C. Kozat, M. R. Jeong, F. Watanabe, and C. C. Chong, "Reliable multicast and broadcast services in relay-based emergency communications," *IEEE Wireless Communications*, vol. 15, no. 3, pp. 40–47, Jun. 2008.

[42] S. Raza, D. Li, C. Chuah, and G. Cheung, "Cooperative peer-to-peer repair for wireless multimedia broadcast," *IEEE International Conference on Multimedia and Expo*, pp. 1075–1078, Jul. 2007.

[43] J. Si, Z. Li, Z. Liu, and X. Chen, "Energy efficient cooperative broadcasting in wireless networks," *IEEE International Conference on Communications*, pp. 5235–5240, 2009.

[44] P. Fan, C. Zhi, C. Wei, and K. B. Letaief, "Reliable relay assisted wireless multicast using network coding," *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 5, pp. 749–762, June 2009.

[45] L. Canourgues, J. Lephay, L. Soyer, and A. L. Beylot, "Stamp: Shared-tree ad hoc multicast protocol," *IEEE Military Communications Conference*, pp. 1–7, Oct. 2006.

[46] B. Kaliaperumal, A. Ebenezer, and Jeyakumar, "Adaptive core based scalable multicasting networks," *Annual IEEE INDICON*, pp. 198–202, Dec. 2005.

[47] S. Lee and C. Kim, "Neighbor supporting ad hoc multicast routing protocol," *in Proceedings of the 1st ACM international symposium on Mobile ad hoc networking and computing*, pp. 37–44, 2000.

[48] S. Y. Oh, J. S. Park, and M. Gerla, "E-odmrp: enhanced odmrp with motion adaptive refresh," *2nd International Symposium on Wireless Communication Systems*, pp. 130–134, Sept. 2005.

[49] J. Li, D. Cordes, and J. Zhang, "Power-aware routing protocols in ad hoc wireless networks," *IEEE Wireless Communications*, vol. 12, no. 6, pp. 69–81, Dec. 2005.

[50] W. Cheng, C. Wen, and K. Feng, "Power-controlled hybrid multicast routing protocol for mobile ad hoc networks," *IEEE 63rd Vehicular Technology Conference*, vol. 3, pp. 1087–1091, May 2006.

[51] M. X. Cheng, J. Sun, M. Min, Y. Li, and W. Wu, "Energy-efficient broadcast and multicast routing in ad hoc wireless networks," *in Proceedings of the IEEE International Conference on Performance, Computing, and Communications Conference*, pp. 87–94, 2003.

[52] J. Jannotti, D. K. Gifford, K. L. Johnson, M. F. Kaashoek, and J. W. O'Toole, "Overcast: reliable multicasting with an overlay network," *in Proceedings of operating systems design and implementation*, pp. 197–212, 2000.

[53] Y. Liu, Y. Guo, and C. Liang, "A survey on Peer-to-Peer video streaming systems," *Journal of Peer-to-Peer Networking and Applications, by Springer New York*, Feb. 2008.

[54] Q. Z. M. Zhang and S. Yang, “Understanding the power of pull-based streaming protocol: Can we do better?” *IEEE Journal on Selected Areas in Communications, special issue on Advances in Peer-to-Peer Streaming Systems*, vol. 25, no. 8, pp. 1678–1694, Dec. 2007.

[55] [Online], “http://research.microsoft.com/en-us/projects/imv/.”

[56] M. Tanimoto, M. P. Tehrani, T. Fujii, and T. Yendo, “Free-viewpoint TV,” in *IEEE Signal Processing Magazine*, vol. 28, no.1, January 2011.

[57] C. Zhang, Z. Yin, and D. Florencio, “Improving depth perception with motion parallax and its application in teleconferencing,” in *IEEE International Workshop on Multimedia Signal Processing*, Rio de Jeneiro, Brazil, October 2009.

[58] F. Bernardini and H. Rushmeier, “The 3d model acquisition pipeline,” *Computer Graphics Forum*, vol. 21, no. 2, pp. 149–172, 2002.

[59] D. Le Gall, “Mpeg: a video compression standard for multimedia applications,” *Communications of the ACM*, vol. 34, no. 4, pp. 46–58, 1991.

[60] N. Cheung and A. Ortega, “Distributed source coding application to low-delay free viewpoint switching in multiview video compression,” *in Proceedings of Picture Coding Symposium (PCS)*, 2007.

[61] N. Cheung, A. Ortega, and G. Cheung, “Distributed source coding techniques for interactive multiview video streaming,” *in Proceedings of Picture Coding Symposium (PCS)*, 2009.

[62] G. Cheung, A. Ortega, and N. Cheung, “Interactive streaming of stored multiview video using redundant frame structures,” *IEEE Transactions on Image Processing*, vol. 20, no. 3, pp. 744–761, March 2011.

[63] J. D. Slepian and J. K. Wolf, “Noiseless coding of correlated information sources,” *IEEE Transactions on Information Theory*, vol. 19, pp. 471–480, Jul. 1973.

[64] A. D. Wyner and J. Ziv, “The rate-distortion function for source coding with side information at the decoder,” *IEEE Transactions on Information Theory*, vol. 22, pp. 1–10, 1976.

[65] A. Aaron, P. Ramanathan, and B. Girod, “Wyner-Ziv coding of light fields for random access,” *IEEE International Workshop on Multimedia Signal Processing*, Sept. 2004.

[66] M. Felegyhazi, J. P. Hubaux, and L. Buttyan, “Nash equilibria of packet forwarding strategies in wireless ad hoc networks,” *IEEE Transactions on Mobile Computing*, vol. 5, no. 5, pp. 463–476, May 2006.

[67] V. Srinivasan, P. Nuggehalli, C. F. Chiasserini, and R. R. Rao, “Cooperation in wireless ad hoc networks,” *in Proceedings of IEEE INFOCOM*, vol. 2, pp. 808–817, April 2003.

[68] M. Felegyhazi, L. Buttyan, and J. P. Hubaux, "Equilibrium analysis of packet forwarding strategies in wireless ad hoc networks - the static case," *in Proceedings of Personal Wireless Communications (PWC), Venice, Italy*, Mar. 2003.

[69] E. Altman, A. A. Kherani, P. Michiardi, and R. Molva, "Non-cooperative forwarding in ad hoc networks," *in Proceedings of the 4th IFIP-TC6 international conference on Networking Technologies, Services, and Protocols*, pp. 486–498, 2005.

[70] Z. Han, P. Charles, and K. J. R. Liu, "A self-learning repeated game framework for optimizing packet forwarding networks," *in Proceedings of IEEE Wireless Communications and Networking Conference*, pp. 2131–2136, 2005.

[71] W. Yu and K. J. R. Liu, "Game theoretic analysis of cooperation stimulation and security in autonomous mobile Ad-hoc networks," *IEEE Transactions on Mobile Computing*, vol. 6, no. 5, pp. 507–521, May 2007.

[72] W. S. Lin, H. V. Zhao, and K. J. R. Liu, "Cooperation stimulation strategies for Peer-to-Peer wireless live video-sharing social networks," *IEEE Transactions on Image Processing*, vol. 19, no. 7, pp. 1768–1784, July 2010.

[73] Z. Liu, Y. Shen, S. Panwar, K. Ross, and Y. Wang, "Using layered video to provide incentives in P2P live streaming," *ACM SigComm Workshop on P2P Streaming and IP-TV*, pp. 311–316, Aug. 2007.

[74] Z. Liu, Y. Shen, K. W. Ross, S. S. Panwar, and Y. Wang, "Substream trading: Towards an open p2p live streaming system," *IEEE International Conference on Network Protocols*, pp. 94–103, Oct. 2008.

[75] S. Zhong, J. Chen, and Y. R. Yang, "Sprite: a simple, cheat-proof, credit-based system for mobile Ad-hoc networks," *in Proceedings of IEEE INFOCOM*, pp. 1987–1997, 2003.

[76] L. Buttyán and J. P. Hubaux, "Stimulating cooperation in self-organizing mobile Ad-hoc networks," *ACM/Kluwer Mobile Network Applications*, vol. 8, no. 5, pp. 579–592, 2003.

[77] Z. Ji, W. Yu, and K. J. R. Liu, "A game theoretical framework for dynamic pricing-based routing in self-organized manets," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 7, pp. 1204–1217, Sept. 2008.

[78] W. Wang and X. Y. Li, "Low-cost truthful multicast in selfish and rational wireless Ad-hoc networks," *2004 IEEE International Conference on Mobile Ad-hoc and Sensor Systems*, pp. 534 – 536, Oct. 2004.

[79] S. Eidenbenz, G. Resta, and P. Santi, "Commit: a sender-centric truthful and energy-efficient routing protocol for Ad-hoc networks with selfish nodes," *in Proceedings of 19th IEEE International Parallel and Distributed Processing Symposium*, Apr. 2005.

[80] M. Neely, "Optimal pricing in a free market wireless network," *in Proceedings of IEEE INFOCOM*, pp. 213–221, 2007.

[81] D. Niyato, E. Hossain, and Z. Han, "Dynamics of multiple-seller and multiple-buyer spectrum trading in cognitive radio networks: A game-theoretic modeling approach," *IEEE Transactions on Mobile Computing*, vol. 8, no. 8, pp. 1009–1022, Aug. 2009.

[82] G. Tan and S. A. Jarvis, "A payment-based incentive and service differentiation mechanism for peer-to-peer streaming broadcast," *In Proceedings of International Workshop on Quality of Service (IWQoS)*, June 2006.

[83] D. Guo and Y. Kwok, "A new auction based approach to efficient p2p live streaming," *2011 IEEE 17th International Conference on Parallel and Distributed Systems,*, pp. 573–580, Dec. 2011.

[84] H. V. Zhao and G. Cheung, "Game theoretical analysis of wireless multiview video multicast using cooperative peer-to-peer repair," *IEEE Workshop on Streaming and Media Communications (in conjunction with ICME 2011),Barcelona, Spain*, July 2011.

[85] F. Milan, J. Jaramillo, and R. Srikant, "Performance analysis of reputation-based mechanisms for multi-hop wireless networks," *in Proceedings of 40th Annual Conference on Information Sciences and Systems*, pp. 12–17, Mar. 2006.

[86] P. Michiardi and R. Molva, "Core: A collaborative reputation mechanism to enforce node cooperation in mobile Ad-hoc networks," *in Proceedings of 6th IFIP TC6/TC11 Joint Working Conference Communication Multimedia Security*, pp. 107–121, 2001.

[87] S. Buchegger and J.-Y. L. Boudec, "Performance analysis of the CONFIDANT protocol," *in Proceedings of 3rd ACM International Symposium Mobile Ad-Hoc Networking and Computing (MobiHoc)*, pp. 226–236, 2002.

[88] J. J. Jaramillo and R. Srikant, "A game theory based reputation mechanism to incentivize cooperation in wireless ad hoc networks," *Ad-Hoc Networks*, vol. 8, no. 4, pp. 416–429, 2010.

[89] Y. Chen and K. J. R. Liu, "Indirect reciprocity game modelling for cooperation stimulation in cognitive networks," *IEEE Transaction on Communications*, vol. 59, no. 1, pp. 159–168, Jan. 2011.

[90] A. Habib and J. Chuang, “Incentive mechanism for peer-to-peer media streaming,” *International Workshop on Quality of Service (IWQoS)*, pp. 171–180, June 2004.

[91] P. K. Hoong and H. Matsuo, “Push-pull incentive-based P2P live media streaming system,” *WSEAS Transaction on Communications*, vol. 7, no. 2, pp. 33–42, Feb. 2008.

[92] T. Qiu, I. Nikolaidis, and F. Li, “On the design of incentiv-aware P2P streaming,” *Journal of Internet Engineering*, vol. 1, no. 2, pp. 61–71, Oct. 2007.

[93] Y. Zhang and M. van der Schaar, “Peer-to-Peer multimedia sharing based on social norms,” *Image Communications*, vol. 27, no. 5, pp. 383–400, May 2012.

[94] W. Wang, X. Li, and Y. Wang, “Truthful multicast in selfish wireless networks,” *in Proceedings of the 10th ACM Annual International Conference on Mobile Computing and Networking (MobiCom)*, pp. 245–259, 2004.

[95] B. Niu, H. V. Zhao, and H. Jiang, “A cooperation stimulation strategy in wireless multicast networks,” *IEEE Transactions on Signal Processing*, vol. 59, no. 5, pp. 2355–2369, May 2011.

[96] Y. Chen, B. Wang, W. S. Lin, Y. Wu, and K. J. R. Liu, “Cooperative peer-to-peer streaming: An evolutionary game-theoretic approach,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 20, no. 10, pp. 1346–1357, Oct. 2010.

[97] Y. E. Sagduyu and A. Ephremides, “A game-theoretic look at simple relay channel,” *Journal on Wireless Networks*, vol. 12, no. 5, pp. 545–560, Sept. 2006.

[98] F. Afghah and A. Abedi, “Stochastic game theoretical model for packet forwarding in reliable relay networks,” *in Proceedings of the 2010 International Conference on Wireless Networks, Las Vegas, Nevada*, pp. 227–233, July 2010.

[99] H. Park and M. van der Schaar, “A framework for foresighted resource reciprocation in P2P networks,” *IEEE Transactions on Multimedia*, vol. 11, no. 1, pp. 101–116, Jan. 2009.

[100] E. Maani and A. K. Katsaggelos, “A game theoretic approach to video streaming over peer-to-peer networks,” *2010 17th IEEE International Conference on Image Processing*, pp. 2909–2912, Sept. 2010.

[101] J. M. Smith, *Evolutionary and the theory of games*. Cambridege University Press, 1982.

[102] J. W. Weibull, *Evolutionary Game Theory*. MIT Press, 1995.

[103] V. Krishna, *Auction Theory*, 2nd ed. Academic Press, 2009.

[104] M. L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley and Sons, 1994.

[105] Y. Chen, B. Wang, W. S. Lin, Y. Wu, and K. J. R. Liu, "Cooperative Peer-to-Peer streaming: An evolutionary game-theoretic approach," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 20, no. 10, pp. 1346–1357, Oct. 2010.

[106] B. Wang, K. J. R. Liu, and T. C. Clancy, "Evolutionary cooperative spectrum sensing game: how to collaborate?" *IEEE Transactions on Communications*, vol. 58, no. 3, pp. 890–900, March 2010.

[107] N.-M. Cheung, A. Ortega, and G. Cheung, "Distributed source coding techniques for interactive multiview video streaming," *27th Picture Coding Symposium*, May 2009.

[108] Y. Sun, Z. Han, and K. J. R. Liu, "Defense of trust management vulnerabilities in distributed network," *IEEE Comm. Magazine*, vol. 46, no. 2, pp. 112–119, Feb. 2008.

[109] D. Blackwell, "Discounted dynamic programming," *Annals of Mathematical Statistics*, vol. 36, no. 1, pp. 226–235, Feb. 1965.

[110] R. Givan, T. Dean, and M. Greig, "Equivalence notions and model minimization in Markov decision processes," *Artificial Intelligence*, vol. 147, no. 1-2, pp. 163–223, 2003.

[111] H. Huang, B. Zhang, G. Chan, G. Cheung, and P. Frossard, "Coding and replication codesign for interactive multiview video streaming," *in Proceedings of IEEE INFOCOM*, pp. 2791–2795, 2012.

[112] L. S. Kennedy, S. Chang, and I. V. Kozintsev, "To search or to label?: predicting the performance of search-based automatic image classifiers," *in Proceedings of the 8th ACM International Workshop on Multimedia Information Retrieval*, pp. 249–258, 2006.

[113] F. Benevenuto, T. Rodrigues, V. Almeida, J. Almeida, and M. Goncalves, "Detecting spammers and content promoters in online video social networks," *in Proceedings of the 32nd International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 620–627, 2009.

Appendix A

For $q \in \Gamma_j$ with $1 < j < N - N_{su}$, $f'(x) = 0$ has a single root $\tilde{x}_f \in \left(0, \frac{j-1}{N-N_{su}-1}\right)$, and $f'(x) > 0$ when $x \in (0, \tilde{x}_f)$ and $f'(x) < 0$ when $x \in (\tilde{x}_f, 1)$.

Proof: From (3.4), we have

$$f(x) = \underbrace{g\binom{l}{k^*}x^{k^*}(1-x)^{(l-k^*)}}_{\triangleq X(x)} - \underbrace{q\sum_{k=k^*}^{l}\binom{l}{k}x^{k}(1-x)^{(l-k)}}_{\triangleq Y(x)} = X(x) - Y(x), \tag{A.1}$$

where $l = N - N_{su} - 1$ and $k^* = \lceil c/q \rceil - 1 = j - 1$ with $1 \leq k^* \leq N - N_{su} - 2$. We then study the monotonicity of $X(x)$ and $Y(x)$, respectively. First, we have $X'(x) = g\binom{l}{k^*}x^{(k^*-1)}(1-x)^{(l-k^*-1)}(k^* - lx)$, and thus, $X'(x) > 0$ when $x \in (0, k^*/l)$, and $X'(x) \leq 0$ when $x \in [k^*/l, 1)$. Similarly, we have $Y'(x) = q\sum_{k=k^*}^{l}\binom{l}{k}x^{(k-1)}(1-x)^{(l-k-1)}(k-lx)$. When $x \in (0, k^*/l)$, we have $(k-lx) > 0$, and thus $Y'(x) > 0$. When $x \in [k^*/l, 1)$, since $lx > 0, 1, ..., k^* - 1$, we have

$$\begin{aligned} Y'(x) &= q\sum_{k=k^*}^{l}\binom{l}{k}x^{(k-1)}(1-x)^{(l-k-1)}(k-lx) \\ &> q\sum_{k=0}^{l}\binom{l}{k}x^{(k-1)}(1-x)^{(l-k-1)}(k-lx) \\ &= q\frac{d}{dx}\left[\sum_{k=0}^{l}\binom{l}{k}x^{k}(1-x)^{(l-k)}\right] = q\frac{d(1)}{dx} = 0. \end{aligned} \tag{A.2}$$

Therefore, we have $Y'(x) > 0$ for $x \in (0, 1)$.

Based on the above analysis, when $x \in [k^*/l, 1)$, we have $X'(x) \leq 0$ and $Y'(x) > 0$, and therefore, $f'(x) = X'(x) - Y'(x) < 0$. When $x \in (0, k^*/l)$, we have $X'(x) > 0$ and $Y'(x) > 0$. Thus, we need to further investigate $f'(x)$, and have

$$\begin{aligned} f'(x) &= g\binom{l}{k^*}x^{(k^*-1)}(1-x)^{(l-k^*-1)}(k^*-lx) - q\sum_{k=k^*}^{l}\binom{l}{k}x^{(k-1)}(1-x)^{(l-k-1)}(k-lx) \\ &= gx^{(k^*-1)}(1-x)^{(l-k^*-1)}(k^*-lx)\underbrace{\left\{\binom{l}{k^*} - \frac{q}{g}\sum_{k=k^*}^{l}\binom{l}{k}\left(\frac{x}{1-x}\right)^{(k-k^*)}\frac{k-lx}{k^*-lx}\right\}}_{\triangleq W(x)} \\ &= gx^{(k^*-1)}(1-x)^{(l-k^*-1)}(k^*-lx)W(x). \end{aligned} \tag{A.3}$$

In the last line of (A.3), when $x \in (0, k^*/l)$, except for the term $W(x)$, all the other terms are larger than zero. To study $W(x)$, we have

$$\begin{aligned} W'(x) &= -\frac{q}{g}\sum_{k=k^*}^{l}\binom{l}{k}\left\{\left(\frac{x}{1-x}\right)^{(k-k^*-1)}\frac{k-k^*}{(1-x)^2}\frac{k-lx}{k^*-lx}\right. \\ &+ \left.\left(\frac{x}{1-x}\right)^{(k-k^*)}\frac{l(k-k^*)}{(k^*-lx)^2}\right\}. \end{aligned} \tag{A.4}$$

Since $k \geq k^*$ and $x \in (0, k^*/l)$, we have $W'(x) < 0$. We then study the function value of $W(x)$ when x approaches 0 and k^*/l, respectively. Since $q \in \Gamma_j$ with $2 \leq j \leq N - N_{su} - 1$, we have $q < g$. Then we have

$$\lim_{x\to 0} W(x) = \binom{l}{k^*} - \frac{q}{g}\binom{l}{k^*}\frac{k^*-0}{k^*-0} = \binom{l}{k^*}\left(1-\frac{q}{g}\right) > 0, \tag{A.5}$$

$$\lim_{x\to k^*/l} W(x) = \binom{l}{k^*} - \infty = -\infty. \tag{A.6}$$

Therefore, $W(x) = 0$ has a single root, $\tilde{x}_f$, in the range $(0, k^*/l)$. From (A.3), $\tilde{x}_f$ is also the single root of $f'(x) = 0$. Thus, when $x \in (0, \tilde{x}_f)$, $W(x) > 0$, and $f'(x) > 0$. When $x \in (\tilde{x}_f, k^*/l)$, $W(x) < 0$, and $f'(x) < 0$.

Based on the above analysis, $f'(x) = 0$ has a single root, $\tilde{x}_f$, in the range $(0, k^*/l)$ ($k^* = j - 1$ and $l = N - N_{su} - 1$), and $f(x) > 0$ when $x \in (0, \tilde{x}_f)$ and $f(x) < 0$ when $x \in (\tilde{x}_f, 1)$. ■

Appendix B

Proof of (3.14)

Proof:

$$\begin{aligned}
h(x) &= g\binom{l}{k^*}[x(1-p_s)]^{k^*}[1-x(1-p_s)]^{(l-k^*)} \\
&- q\sum_{k=k^*}^{l}\binom{l}{k}[x(1-p_s)]^{k}[1-x(1-p_s)]^{(l-k)} \\
&= f\Big((1-p_s)x\Big). \qquad \text{(B.1)}
\end{aligned}$$

Proof: $h(x) = \bar{V}_B(x) - \bar{V}_{NB}(x)$. Let $k^* = \lceil d^w/q \rceil - 1$, and we have

$$\begin{aligned}
h(x) &= \underbrace{\sum_{n=k^*}^{l}\binom{l}{n}(1-p_s)^n p_s^{l-n}\binom{n}{k^*}x^{k^*}(1-x)^{n-k^*}g}_{\triangleq G(x)} \\
&- \underbrace{\sum_{n=k^*}^{l}\binom{l}{n}(1-p_s)^n p_s^{l-n}\left\{\sum_{k=k^*}^{n}\binom{n}{k}x^{k}(1-x)^{(n-k)}q\right\}}_{\triangleq H(x)} \\
&= G(x) - H(x). \qquad \text{(B.2)}
\end{aligned}$$

$$\begin{aligned}
G(x) &= g\sum_{n=k^*}^{l}\frac{l!}{n!(l-n)!}(1-p_s)^n p_s^{l-n}\frac{n!}{k^*!(n-k^*)!}x^{k^*}(1-x)^{n-k^*} \\
&= g\frac{l!}{k^*!(l-k^*)!}(1-p_s)^{k^*}x^{k^*}\sum_{n=k^*}^{l}\frac{(l-k^*)!}{(n-k^*)!(l-n)!}[(1-p_s)(1-x)]^{(n-k^*)}p_s^{l-n} \\
&\overset{m\triangleq n-k^*}{=\!=\!=} g\binom{l}{k^*}[x(1-p_s)]^{k^*}\sum_{m=0}^{l-k^*}\frac{(l-k^*)!}{m!(l-k^*-m)!}[(1-p_s)(1-x)]^m p_s^{l-k^*-m}
\end{aligned}$$

$$= \quad g\binom{l}{k^*}[x(1-p_s)]^{k^*}[(1-p_s)(1-x)+p_s]^{(l-k^*)}. \tag{B.3}$$

Similarly, $H(x) = q\sum_{k=k^*}^{l}\left\{\binom{l}{k}[(1-p_s)x]^k[(1-p_s)(1-x)+p_s]^{(l-k)}\right\}$. Thus,

$$\begin{aligned} h(x) &= g\binom{l}{k^*}[x(1-p_s)]^{k^*}[1-x(1-p_s)]^{(l-k^*)} \\ &- q\sum_{k=k^*}^{l}\binom{l}{k}[(1-p_s)x]^k[1-x(1-p_s)]^{(l-k)}. \end{aligned} \tag{B.4}$$

When comparing $h(x)$ with $f(x)$ in (3.4), it is easy to observe that $h(x) = f\Big((1-p_s)x\Big)$. ■

Appendix C

Proof of Proposition 3.2

Proof: For $q \in \Gamma_j$ with $j \in \{\underline{j}, ..., N - N_s\}$, we first prove that, $\frac{\partial T_R}{\partial x^*} \geq 0$, and then prove that $\frac{\partial x^*}{\partial q} \leq 0$. Therefore, we have $\frac{\partial T_R}{\partial q} = \frac{\partial T_R}{\partial x^*}\frac{\partial x^*}{\partial q} \leq 0$ and T_R is a non-increasing function of q.

To prove that $\partial T_R/\partial x^* \geq 0$, we first define $J(n, x^*) \overset{\triangle}{=} \sum_{k=\lceil d^w/q \rceil}^{n} \binom{n}{k} (x^*)^k (1 - x^*)^{n-k}$ and rewrite (3.16) as $T_R(x^*|N_s, d^w) = \sum_{n=0}^{N-N_s} \binom{N-N_s}{n} (1-p_s)^n p_s^{(N-N_s-n)} J(n, x^*)$. The first derivative of $J(n, x^*)$ over x^* is,

$$\frac{\partial J(n, x^*)}{\partial x^*} = \sum_{k=\lceil d^w/q \rceil}^{n} \binom{n}{k} (x^*)^{(k-1)} (1 - x^*)^{(n-k-1)} (k - nx^*). \quad \text{(C.1)}$$

Similar to (A.2), we can prove that $\partial J(n, x^*)/\partial x^* \geq 0$ for $x^* \in [0, 1]$. Therefore, we have

$$\frac{\partial T_R}{\partial x^*} = \sum_{n=0}^{N-N_s} \binom{N - N_s}{n} (1 - p_s)^n p_s^{(N-N_s-n)} \frac{\partial J(n, x^*)}{\partial x^*} \geq 0. \quad \text{(C.2)}$$

We then prove for $q \in \Gamma_j$ with $j \in \{\underline{j}, ..., N - N_s\}$, $\frac{\partial x^*}{\partial q} \leq 0$. First, if $q \in \Gamma_{N-N_s}$ (Case 3 in Theorem 3.2), $x^* = 1$ for any $q \in \Gamma_{N-N_s}$. Therefore, when $q \in \Gamma_{N-N_s}$, we have $\partial x^*/\partial q = 0$.

If $q \in \Gamma_j$ with $j \in \{\underline{j}, ..., N - N_s - 1\}$ (Case 2 in Theorem 3.2), from Theorem 3.2, it can be seen that x^* is the non-zero root, denoted $\breve{x}_h$, of $h(x) = 0$ if $\breve{x}_h < 1$, and $x^* = 1$ otherwise. When the price q increases within Γ_j, k^* keeps the same, and thus, $h(x)$ decreases based on (3.14). From Fig. 3.7 and Fig. 3.8, it can be seen that, if $h(x)$ decreases, the non-zero root of $h(x) = 0$ decreases. Therefore, when q increases, the ESS x^* either keeps at $x^* = 1$, or decreases. This means $\partial x^*/\partial q \leq 0$.

■

Appendix D

Proof of Proposition 4.1

Proof: As discussed in Chapter 4.2.2, one-shot deviation of a given action policy means that a user takes a different action rather than the one defined in the action policy only for the current response to a request, but still follows the given action policy in the future responses. Suppose that a user receives a request at reputation r and view v, with the action policy π, the lifetime utility is defined in (4.14). If the user takes action $a'_{r,v}$ rather than $a_{r,v}$ defined in π only for response of this request, but still follows π in future responses, we have the lifetime utility,

$$W_{r,v}^{a'_{r,v},\pi} = U_{r,v}^{a'_{r,v}} + \eta^L \sum_{v'=1}^{M} \mathbf{T}^L(v,v') \left[(1 - P_{r\to 1}^{a'_{r,v}}) W_{\min(r+1,R),v'}^{\pi} + P_{r\to 1}^{a'_{r,v}} W_{1,v'}^{\pi} \right] \quad \text{(D.1)}$$

Comparing (4.14) and (D.1), one-shot deviation to $a'_{r,v}$ gives a different expected short-term utility $U_{r,v}^{a'_{r,v}}$ and a different reputation transition probability $P_{r\to 1}^{a'_{r,v}}$. We then use one-shot deviation principle to prove this proposition.

Assume that π is a policy including action $t_r + 1 \leq a_{r,v} \leq R$ for $r \geq t_r - 1$ and any $v \in \mathcal{V}$. We then exam whether it can resist the one-shot deviation to $a'_{r,v} = t_r$ or $a'_{r,v} = R + 1$. Following the discussion of Chapter 4.2, for a user at view v with reputation $r \geq t_r - 1$, by taking action $a_{r,v}$, he/she has expected immediate cost $C^{a_{r,v}} = c \sum_{r_j = a_{r,v}}^{R} \mathbf{x}(r_j)$, while following (4.6), the reputation transits to 1 with probability $P_{r\to 1}^{a_{r,v}} = \sum_{r_j = t_r}^{a_{r,v}-1} \mathbf{x}(r_j)$. Thus, the expected short-term payoff is $U_{r,v}^{a_{r,v}} = -c \sum_{r_j = a_{r,v}}^{R} \mathbf{x}(r_j) + (1 - \sum_{r_j = t_r}^{a_{r,v}-1} \mathbf{x}(r_j)) G_{\min\{r+1,R\},v} + \sum_{r_j = t_r}^{a_{r,v}-1} \mathbf{x}(r_j) G_{1,v}$.

Substitute $U_{r,v}^{a_{r,v}}$ and $P_{r\to 1}^{a_{r,v}}$ into (4.14), we can get the lifetime utility

$$\begin{aligned} W_{r,v}^{\pi} = & -c \sum_{r_j=a_{r,v}}^{R} \mathbf{x}(r_j) + \left(1 - \sum_{r_j=t_r}^{a_{r,v}-1} \mathbf{x}(r_j)\right) G_{\min\{r+1,R\},v} + \sum_{r_j=t_r}^{a_{r,v}-1} \mathbf{x}(r_j) G_{1,v} \\ & + \eta^L \left(1 - \sum_{r_j=t_r}^{a_{r,v}-1} \mathbf{x}(r_j)\right) \sum_{v'=1}^{M} \mathbf{T}^L(v,v') W_{\min\{r+1,R\},v'}^{\pi} \\ & + \eta^L \left(\sum_{r_j=t_r}^{a_{r,v}-1} \mathbf{x}(r_j)\right) \sum_{v'=1}^{M} \mathbf{T}^L(v,v') W_{1,v'}^{\pi}. \end{aligned}$$

However, by the one-shot deviation to $a'_{r,v} = t_r$, the expected immediate cost is $C^{t_r} = c\sum_{r_j=t_r}^{R} \mathbf{x}(r_j)$ and his/her reputation falls to 1 with probability $P_{r\to 1}^{t_r} = 0$. Thus, the expected short-term payoff becomes $U_{r,v}^{t_r} = -c\sum_{r_j=t_r}^{R} \mathbf{x}(r_j) + G_{\min\{r+1,R\},v}$. Thus, following (D.1), we have

$$W_{r,v}^{a'_{r,v}=t_r,\pi} \;=\; -c \sum_{r_j=t_r}^{R} \mathbf{x}(r_j) + G_{\min\{r+1,R\},v} + \eta^L \sum_{v'=1}^{M} \mathbf{T}^L(v,v') W_{\min\{r+1,R\},v'}^{\pi}.$$

Similarly, by the one-shot deviation to $a'_{r,v} = R+1$, he/she will help upload with probability 0, which gives zero cost, and his/her reputation falls to 1 with probability $P_{r\to 1}^{R+1} = \sum_{r_j=t_r}^{R} \mathbf{x}(r_j)$. Thus, the expected short-term payoff is $U_{r,v}^{R+1} = (1 - \sum_{r_j=t_r}^{R} \mathbf{x}(r_j))G_{\min\{r+1,R\},v} + \sum_{r_j=t_r}^{R} \mathbf{x}(r_j) G(1,v)$. Thus, following (D.1), we have

$$\begin{aligned} W_{r,v}^{a'_{r,v}=R+1,\pi} = & \left(1 - \sum_{r_j=t_r}^{R} \mathbf{x}(r_j)\right) G_{\min\{r+1,R\},v} + \sum_{r_j=t_r}^{R} \mathbf{x}(r_j) G_{1,v} \\ & + \eta^L \left(1 - \sum_{r_j=t_r}^{R} \mathbf{x}(r_j)\right) \sum_{v'=1}^{M} \mathbf{T}^L(v,v') W_{\min\{r+1,R\},v'}^{\pi} \\ & + \eta^L \left(\sum_{r_j=t_r}^{R} \mathbf{x}(r_j)\right) \sum_{v'=1}^{M} \mathbf{T}^L(v,v') W_{1,v'}^{\pi}. \end{aligned}$$

We then compare $W_{r,v}^{\pi}$ with $W_{r,v}^{a'_{r,v}=t_r,\pi}$ and $W_{r,v}^{a'_{r,v}=R+1,\pi}$, and observe that

$$W_{r,v}^{\pi} - W_{r,v}^{a'_{r,v}=t_r,\pi} = \sum_{r_j=t_r}^{a_{r,v}-1} \mathbf{x}(r_j) \Big\{ c - G_{\min\{r+1,R\},v} + G_{1,v}$$

$$- \eta^L \sum_{v'=1}^{M} \mathbf{T}^L(v,v') \left[W^{\pi}_{\min\{r+1,R\},v'} - W^{\pi}_{1,v'}\right] \Bigg\}. \quad \text{(D.2)}$$

$$W^{\pi}_{r,v} - W^{a'_{r,v}=R,\pi}_{r,v} = - \sum_{r_j=a_{r,v}}^{R} \mathbf{x}(r_j) \Bigg\{ c - G_{\min\{r+1,R\},v} + G_{1,v} - \eta^L \sum_{v'=1}^{M} \mathbf{T}^L(v,v') \left[W^{\pi}_{\min\{r+1,R\},v'} - W^{\pi}_{1,v'}\right] \Bigg\}. \quad \text{(D.3)}$$

Thus, we have

$$\sum_{r_j=a_{r,v}}^{R} \mathbf{x}(r_j) \left[W^{\pi}_{r,v} - W^{a'_{r,v}=t_r,\pi}_{r,v}\right] = - \sum_{r_j=t_r}^{a_{r,v}-1} \mathbf{x}(r_j) \left[W^{\pi}_{r,v} - W^{a'_{r,v}=R+1,\pi}_{r,v}\right]. \quad \text{(D.4)}$$

Given that $\sum_{r_j=a_{r,v}}^{R} \mathbf{x}(r_j) \geq 0$ and $\sum_{r_j=t_r}^{a_{r,v}-1} \mathbf{x}(r_j) \geq 0$, we either have $W^{\pi}_{r,v} - W^{a'_{r,v}=t_r,\pi}_{r,v} \leq 0$, or have $W^{\pi}_{r,v} - W^{a'_{r,v}=R+1,\pi}_{r,v} \leq 0$. Therefore, π cannot resist the one-shot deviation to either $a'_{r,v} = t_r$ or $a'_{r,v} = R+1$, and thus cannot be an equilibrium policy. ■

Appendix E

Proof of Proposition 4.2

Proof: This proof takes two steps. We first prove the part on state transition probabilities, and then we prove $U^a_{r,v} = U^a_{r',v}$.

• For any action a and two states (r, v) and $(r', v) \in \mathcal{S}_{\bar{\mathcal{R}},v}$ ($r \neq r'$), we have $P^a_{(r,v)\to(r'',v')} = P^a_{(r',v)\to(r'',v')}$ for $r'' < t_r - 1$, and $\sum_{s\in\mathcal{S}_{\bar{\mathcal{R}},v'}} P^a_{(r,v)\to s} = \sum_{s\in\mathcal{S}_{\bar{\mathcal{R}},v'}} P^a_{(r',v)\to s}$.

We first prove $P^a_{(r,v)\to(r'',v')} = P^a_{(r',v)\to(r'',v')}$ for $r'' < t_r - 1$. Following the discussion in Chapter 4.2, the reputation and view transition probabilities are independent, i.e., $P^a_{(r,v)\to(r'',v')} = P^a_{r\to r''}\mathbf{T}^L(v, v')$. Thus, we only need to prove $P^a_{r\to r''} = P^a_{r'\to r''}$ for $r'' < t_r - 1$. Since the reputation $r \geq t_r - 1$ can only be updated to either 1 or $\min\{r + 1, R\} \in \bar{\mathcal{R}}$, thus, we have $P^a_{r\to r''} = P^a_{r'\to r''} = 0$ for $2 \leq r'' \leq t_r - 2$. Furthermore, from (4.6), for a user with reputation no less than $t_r - 1$, the probability that his/her reputation is lowered to 1 depends on the action a only. Thus, we also have $P^a_{r\to 1} = P^a_{r'\to 1}$. Based on the above analysis, we have $P^a_{(r,v)\to(r'',v')} = P^a_{(r',v)\to(r'',v')}$ for $r'' < t_r - 1$.

We then prove $\sum_{s\in\mathcal{S}_{\bar{\mathcal{R}},v'}} P^a_{(r,v)\to s} = \sum_{s\in\mathcal{S}_{\bar{\mathcal{R}},v'}} P^a_{(r',v)\to s}$. Similarly, we also only need to prove $\sum_{r''\in\bar{\mathcal{R}}} P^a_{r\to r''} = \sum_{r''\in\bar{\mathcal{R}}} P^a_{r'\to r''}$. Since a reputation $r \geq t_r - 1$ can only be updated to either 1 or $\min\{r + 1, R\} \in \bar{\mathcal{R}}$, thus, we have $\sum_{r''\in\bar{\mathcal{R}}} P^a_{r\to r''} = P^a_{r\to\min\{r+1,R\}}$. Therefore, we only need to prove $P^a_{r\to\min\{r+1,R\}} = P^a_{r'\to\min\{r'+1,R\}}$. Since we have $P^a_{r\to 1} = P^a_{r'\to 1}$, and we also have $P^a_{r\to\min\{r+1,R\}} = 1 - P^a_{r\to 1}$ and $P^a_{r'\to\min\{r'+1,R\}} = 1 - P^a_{r'\to 1}$, thus, we have $P^a_{r\to\min\{r+1,R\}} = P^a_{r'\to\min\{r'+1,R\}}$. Based on the above analysis, we have $\sum_{s\in\mathcal{S}_{\bar{\mathcal{R}},v'}} P^a_{(r,v)\to s} = \sum_{s\in\mathcal{S}_{\bar{\mathcal{R}},v'}} P^a_{(r',v)\to s}$.

- $U^a_{r,v} = U^a_{r',v}$.

Following (4.13), $U^a_{r,v} = -c\sum_{r=a}^{R}\mathbf{x}(r) + (1 - P^a_{r\to 1})G_{\min(r+1,R),v} + P^a_{r\to 1}G_{1,v}$. First, the probability of helping to upload, $\sum_{r=a}^{R}\mathbf{x}(r)$, depends on a only. Thus, the same action taken at (r,v) and (r',v) introduces the same cost. Second, we observe that $G_{\min(r+1,R),v} = G_{\min(r'+1,R),v}$. This is because users select action only from $\{t_r, R+1\}$. Since we have $t_r \leq \min(r+1,R) \leq R$ and $t_r \leq \min(r'+1,R) \leq R$, the helpers' action (either t_r or $R+1$) does not differentiate the reputation $\min(r+1,R)$ from $\min(r'+1,R)$. Thus, $G_{\min(r+1,R),v} = G_{\min(r'+1,R),v}$. Since we also have $P^a_{r\to 1} = P^a_{r'\to 1}$, it is easy to show $U^a_{r,v} = U^a_{r',v}$. ■

Appendix F

Proof of (4.22)

$$G_{r',v} \quad = \quad \left(\sum_{r_k=t_r-1}^{R} \mathbf{x}(r_k)\right)\left(\sum_{v_k=1}^{M} \mathbf{v}(v_k) I[a_{v_k} \leq r']\right) g_v.$$

Proof: From the discussion in Chapter 4.2, if the helper k has reputation $r_k < t_r - 1$, he/she takes action $R+1$ and does not cooperate. When he/she has reputation $r_k \geq t_r - 1$, his/her action depends on his/her view only. Thus, we can first rewrite (4.11) as

$$G_{r',v} = \sum_{l=1}^{L} \eta^l \sum_{v'=1}^{M} \left\{ \left(\sum_{r_k=t_r-1}^{R} \mathbf{x}(r_k)\right)\left(\sum_{v_k(t+l)=1}^{M} \mathbf{p}_{v'}(v_k(t+l)) I\left[a_{v_k(t+l)} \leq r'\right]\right) \right.$$
$$\left. \times P\left[\mathbb{H}_1(v')\right] P\left[v_i(t+l) = v', v_i(t+l-1) \notin \mathcal{V}_{v'} | v_i(t) = v\right] \right\}, \tag{F.1}$$

where $v_i(t)$ and $v_k(t)$ are user i and the helper k's view at time t, respectively. We then focus on the term $\left(\sum_{v_k(t+l)=1}^{M} \mathbf{p}_{v'}(v_k(t+l)) I\left[a_{v_k(t+l)} \leq r'\right]\right)$, and prove that the summation is the same for different v' in the case with a single anchor view. Following (4.10), we first have

$$\sum_{v_k(t+l)=1}^{M} \mathbf{p}_{v'}(v_k(t+l)) I\left[a_{v_k(t+l)} \leq r'\right]$$

$$
\begin{aligned}
&= \sum_{v_k(t+l)=1}^{M} \sum_{v'' \in \mathcal{V}_{v'}} \mathbf{T}(v'', v_k(t+l)) \frac{\mathbf{v}(v'')}{\sum_{\tilde{v} \in \mathcal{V}_{v'}} \mathbf{v}(\tilde{v})} I\left[a_{v_k(t+l)} \leq r'\right] \\
&= \sum_{v'' \in \mathcal{V}_{v'}} \frac{\mathbf{v}(v'')}{\sum_{\tilde{v} \in \mathcal{V}_{v'}} \mathbf{v}(\tilde{v})} \left(\sum_{v_k(t+l)=1}^{M} \mathbf{T}(v'', v_k(t+l)) I\left[a_{v_k(t+l)} \leq r'\right] \right) \qquad \text{(F.2)}
\end{aligned}
$$

For the term $\left(\sum_{v_k(t+l)=1}^{M} \mathbf{T}(v'', v_k(t+l)) I\left[a_{v_k(t+l)} \leq r'\right]\right)$ in (F.2), following the discussion in Chapter 4.3, if v'' is the anchor view, we have $\mathbf{T}(v'', v_k(t+l)) = \mathbf{v}(v_k(t+l))$, where $\mathbf{v}$ is the steady state view distribution. Thus,

$$
\begin{aligned}
&\left(\sum_{v_k(t+l)=1}^{M} \mathbf{T}(v'', v_k(t+l)) I\left[a_{v_k(t+l)} \leq r'\right] \right) \\
&= \sum_{v_k(t+l)=1}^{M} \mathbf{v}(v_k(t+1)) I[a_{v_k(t+l)} \leq r']. \qquad \text{(F.3)}
\end{aligned}
$$

If v'' is a normal view, for example, a normal view at the left side, the help k will only transit to the anchor view with probability P_a and to each normal view at the left side normal view set with probability $\frac{2(1-P_a)}{M-1}$. Thus, we have

$$
\begin{aligned}
&\left(\sum_{v_k(t+l)=1}^{M} \mathbf{T}(v'', v_k(t+l)) I\left[a_{v_k(t+l)} \leq r'\right] \right) \\
&= \sum_{v_k(t+l)=1}^{\sigma-1} \frac{2(1-P_a)}{M-1} I[a_{v_k(t+l)} \leq r'] + P_a I[a_\sigma \leq r_i'], \qquad \text{(F.4)}
\end{aligned}
$$

where σ is the anchor view index. Note that the views are symmetric with respect to the anchor view, and a user at views with symmetric locations should take the same action, i.e., $a_1 = a_M$, $a_2 = a_{M-1}$,..., $a_{\sigma-1} = a_{\sigma+1}$. Thus, (F.4) can be written as

$$
\begin{aligned}
&\left(\sum_{v_k(t+l)=1}^{M} \mathbf{T}(v'', v_k(t+l)) I\left[a_{v_k(t+l)} \leq r'\right] \right) \\
&= \sum_{v_k(t+l)=1}^{\sigma-1} \frac{(1-P_a)}{M-1} I[a_{v_k(t+l)} \leq r'] + \sum_{v_k(t+l)=\sigma+1}^{M} \frac{(1-P_a)}{M-1} I[a_{v_k(t+l)} \leq r'] + P_a I[a_\sigma \leq r'] \\
&= \sum_{v_k(t+l)=1}^{M} \mathbf{v}(v_k(t+l)) I[a_{v_k(t+l)} \leq r'], \qquad \text{(F.5)}
\end{aligned}
$$

which is the same as (F.3). The analysis is the same if v'' is a normal view in the right side normal view set. Based on the above analysis, we have

$$\sum_{v_k(t+l)=1}^{M} \mathbf{p}_{v'}(v_k(t+l))I\left[a_{v_k(t+l)} \leq r'\right] = \sum_{v_k(t+l)=1}^{M} \mathbf{v}(v_k(t+l))I\left[a_{v_k(t+l)} \leq r'\right], \quad \text{(F.6)}$$

which is not related to v'. Thus, (F.1) can be rewritten as

$$\begin{aligned} G_{r',v} = & \left(\sum_{r_k=t_r-1}^{R} \mathbf{x}(r_k)\right)\left(\sum_{v_k=1}^{M} \mathbf{v}(v_k)I\left[a_{v_k} \leq r'\right]\right) \\ & \times \sum_{l=1}^{L} \eta^l \sum_{v'=1}^{M} P\left[\mathbb{H}_1(v')\right] P\left[v_i(t+l) = v', v_i(t+l-1) \notin \mathcal{V}_{v'} | v_i(t) = v\right] \\ = & \left(\sum_{r_k=t_r-1}^{R} \mathbf{x}(r_k)\right)\left(\sum_{v_k=1}^{M} \mathbf{v}(v_k)I\left[a_{v_k} \leq r'\right]\right) g_v. \end{aligned} \quad \text{(F.7)}$$

where g_v is defined in (4.12). ■

Appendix G

Proof of Proposition 4.3

Proof: In this proof, following (4.12), we first let

$$z(v,l) \triangleq \eta^l \sum_{v'=1}^{M} P\left[\mathbb{H}_1(v')\right] P\left[v_i(t+l)=v', v_i(t+l-1) \notin \mathcal{V}_{v'} | v_i(t)=v\right] \quad \text{(G.1)}$$

and g_v can be rewritten as $g_v = \sum_{l=1}^{L} z(v,l)$, where $z(v,l)$ is the expected gain received at the lth segment after the view switching from view v, if helpers always help. Since the views are symmetric with respect to the anchor view, switching from views with symmetric positions, users should receive the same gain, i.e., $z(v,l)$ and g_v are both symmetric with respect to the anchor view (e.g., $z(1,l) = z(M,l)$ and $g_1 = g_M$). In this proof, we first show that for all $l \geq 2$, $z(v,l)$ is the same for all $v \in \mathcal{V}$. Therefore, for g_v of different view v, the difference is caused by $z(v,1)$ (i.e., $l = 1$). In the second step, we then show that views in the same subset $\underline{\mathcal{V}}$ (or $\bar{\mathcal{V}}$), the corresponding $z(v,1)$'s are approximately the same.

• For $l \geq 2$, $z(v,l)$ is the same for all $v \in \mathcal{V}$.

Substitute (4.8) in Chapter 4.2 into (G.1), we first rewrite $z(v,l)$ ($l \geq 2$) as

$$\begin{aligned} z(v,l) &\triangleq \eta^l \sum_{v'=1}^{M} P\left[\mathbb{H}_1(v')\right] \left(\sum_{v'' \notin \mathcal{V}_{v'}} \mathbf{T}^{l-1}(v,v'')\mathbf{T}(v'',v') \right) \\ &= \eta^l \sum_{v'=1}^{M} P\left[\mathbb{H}_1(v')\right] \left(\sum_{v'' \notin \mathcal{V}_{v'}} \sum_{\tilde{v}=1}^{M} \mathbf{T}(v,\tilde{v})\mathbf{T}^{l-2}(\tilde{v},v'')\mathbf{T}(v'',v') \right) \end{aligned}$$

$$
\begin{aligned}
&= \eta \sum_{\tilde{v}=1}^{M} \mathbf{T}(v,\tilde{v}) \left\{ \eta^{(l-1)} \sum_{v'=1}^{M} P\left[\mathbb{H}_1(v')\right] \left(\sum_{v'' \notin \mathcal{V}_{v'}} \mathbf{T}^{l-2}(\tilde{v},v'')\mathbf{T}(v'',v') \right) \right\} \\
&= \eta \sum_{\tilde{v}=1}^{M} \mathbf{T}(v,\tilde{v}) z(\tilde{v}, l-1).
\end{aligned} \tag{G.2}
$$

Therefore, if v is the anchor view, we have $\mathbf{T}(v,\tilde{v}) = \mathbf{v}(\tilde{v})$ and thus,

$$
z(\sigma, l) = \eta \sum_{\tilde{v}=1}^{M} \mathbf{v}(\tilde{v}) z(\tilde{v}, l-1). \tag{G.3}
$$

If v is a normal view, e.g., a normal view in the left side normal view set, after one segment, a user will only transit to the anchor view with probability P_a and to each normal view in the left side normal view set with probability $\frac{2(1-P_a)}{M-1}$. Thus, we have

$$
z(v,l) = \sum_{\tilde{v}=1}^{\sigma-1} \frac{2(1-P_a)}{M-1} z(\tilde{v}, l-1) + P_a z(\sigma, l-1). \tag{G.4}
$$

As discussed earlier, $z(\tilde{v}, l-1)$ is symmetric with respect to the anchor view. Therefore, (G.4) can be rewritten as,

$$
z(v,l) = \sum_{\tilde{v}=1}^{\sigma-1} \frac{(1-P_a)}{M-1} z(\tilde{v}, l-1) + \sum_{\tilde{v}=\sigma+1}^{M} \frac{(1-P_a)}{M-1} z(\tilde{v}, l-1) + P_a z(\sigma, l-1). \tag{G.5}
$$

which is the same as (G.3) when v is the anchor view. The analysis is the same if v is a normal view in the right side normal view set. Thus, for $l \geq 2$, $z(v,l)$ is the same for all $v \in \mathcal{V}$.

• For views in $\underline{\mathcal{V}}$, the corresponding $z(v,1)$'s are approximately the same.

With $\underline{\mathcal{V}} = \{\sigma-1, \sigma, \sigma+1\}$, we first prove that $z(\sigma,1) \approx z(\sigma-1,1)$. When $l=1$, we first have $z(v,1) = \eta \sum_{v' \notin \mathcal{V}_v} P\left[\mathbb{H}_1(v')\right] \mathbf{T}(v,v')$. Therefore, we have

$$
z(\sigma,1) = \eta \sum_{v'=1}^{\sigma-2} \frac{(1-P_a)}{M-1} P\left[\mathbb{H}_1(v')\right] + \eta \sum_{v'=\sigma+2}^{M} \frac{(1-P_a)}{M-1} P\left[\mathbb{H}_1(v')\right]. \tag{G.6}
$$

Similarly, we also have

$$
z(\sigma-1,1) = \eta \sum_{v'=1}^{\sigma-3} \frac{2(1-P_a)}{M-1} P\left[\mathbb{H}_1(v')\right]. \tag{G.7}
$$

Here, $P[\mathbb{H}_1(v')]$ is the probability of finding a helper to help the view switching to view v', which is also symmetric with respect to the anchor view. Thus, (G.7) can be rewritten as

$$z(\sigma-1,1) = \eta \sum_{v'=1}^{\sigma-3} \frac{(1-P_a)}{M-1} P[\mathbb{H}_1(v')] + \eta \sum_{v'=\sigma+3}^{M} \frac{(1-P_a)}{M-1} P[\mathbb{H}_1(v')]. \quad \text{(G.8)}$$

Therefore, we have

$$z(\sigma,1) - z(\sigma-1,1) = \eta \frac{(1-P_a)}{M-1} \{P[\mathbb{H}_1(\sigma-2)] + P[\mathbb{H}_1(\sigma+2)]\}. \quad \text{(G.9)}$$

Since in this work we consider a high dimension multiview video system, e.g., $M \geq 30$, $1/(M-1)$ approximates to zero, and thus, $z(\sigma,1) - z(\sigma-1,1) \approx 0$, and therefore, we have $g_\sigma - g_{\sigma-1} \approx 0$. Since $g_{\sigma-1} = g_{\sigma+1}$, thus, for views in $\underline{\mathcal{V}}$, the corresponding g_v's are approximately the same.

• g_v's are approximately the same for views in $\bar{\mathcal{V}}$. This proof can be done in the same way as the proof for $\underline{\mathcal{V}}$, which is omitted. ■

Appendix H

Proof of Proposition 4.4

Proof: This proof takes two steps to show that the two conditions of bisimilarity defined in Definition 4.1 are satisfied.

• The first condition of Definition 4.1 is satisfied, i.e., for any (r, v) and (r, v') from $\mathcal{S}_{r,\underline{\mathcal{V}}}$ (or $\mathcal{S}_{r,\bar{\mathcal{V}}}$) with $v \neq v'$, we have $\sum_{s\in\mathcal{S}_{r',\underline{\mathcal{V}}}} P^a_{(r,v)\to s} = \sum_{s\in\mathcal{S}_{r',\underline{\mathcal{V}}}} P^a_{(r,v')\to s}$ and $\sum_{s\in\mathcal{S}_{r',\bar{\mathcal{V}}}} P^a_{(r,v)\to s} = \sum_{s\in\mathcal{S}_{r',\bar{\mathcal{V}}}} P^a_{(r,v')\to s}$ where r, $r' \in \{1, 2, ..., t_r - 2, \bar{\mathcal{R}}\}$.

From Chapter 4.2, the reputation and view transition probabilities are independently. Thus, we have $\sum_{s\in\mathcal{S}_{r',\underline{\mathcal{V}}}} P^a_{(r,v)\to s} = P^a_{r\to r'} \sum_{v''\in\underline{\mathcal{V}}} \mathbf{T}^L(v, v'') = P^a_{r\to r'}(P_a + \frac{2(1-P_a)}{M-1})$. Similarly, we also have $\sum_{s\in\mathcal{S}_{r',\underline{\mathcal{V}}}} P^a_{(r,v')\to s} = P^a_{r\to r'}(P_a + \frac{2(1-P_a)}{M-1})$. Therefore, we have $\sum_{s\in\mathcal{S}_{r',\underline{\mathcal{V}}}} P^a_{(r,v)\to s} = \sum_{s\in\mathcal{S}_{r',\underline{\mathcal{V}}}} P^a_{(r,v')\to s}$. We can prove $\sum_{s\in\mathcal{S}_{r',\bar{\mathcal{V}}}} P^a_{(r,v)\to s} = \sum_{s\in\mathcal{S}_{r',\bar{\mathcal{V}}}} P^a_{(r,v')\to s}$ in the same way.

• The second condition of Definition 4.1 is satisfied, i.e., for any (r, v) and (r, v') from $\mathcal{S}_{r,\underline{\mathcal{V}}}$ (or $\mathcal{S}_{r,\bar{\mathcal{V}}}$) with $v \neq v'$, we have $U^a_{r,v} = U^a_{r,v'}$ for any $a \in \mathcal{A}$.

From the discussion in Chapter 4.2.2.2, when a user has reputation $r < t_r - 1$, he/she does not help anyone and also no one helps him/her. Thus, we have $U^a_{r,v} = U^a_{r,v'} = 0$. When $r \geq t_r - 1$, the short-term utility can be written as $U^a_v = -c\sum_{r_j=a}^{R} \mathbf{x}(r_j) + (1 - P^a_{\bar{\mathcal{R}}\to 1})G_{\min(r+1,R),v} + P^a_{\bar{\mathcal{R}}\to 1}G_{1,v}$. First, the probability of helping to upload depends on a only. Thus, the same action taken at (r, v) and (r, v') introduces the same cost. Second, we observe that $G_{\min(r+1,R),v} = G_{\min(r+1,R),v'}$. This is because based on (4.22), the only difference between $G_{\min(r+1,R),v}$ and $G_{\min(r+1,R),v'}$ is the term of g_v and $g_{v'}$. However, since $g_v = g_{v'}$ for v and v' from the same view set $\underline{\mathcal{V}}$ (or $\bar{\mathcal{V}}$). Thus, we have $G_{\min(r+1,R),v} = G_{\min(r+1,R),v'}$. We also have $G_{1,v} = G_{1,v'} = 0$. Based on the above analysis, we have $U^a_{r,v} = U^a_{r,v'}$. ■

Appendix I

Proof of Proposition 4.5

Proof: In this proof, we will exam each given policy one by one. For each policy π, we first assume that all users use this policy and study the corresponding stationary reputation distribution $\mathbf{x}$. Then, following the one-shot deviation principle, we exam whether a user has incentive to unilaterally deviate to any one-shot deviation. Note that in the 2-level reputation system, we have $t_r = 2$ (i.e., all users have reputation belonging to $\bar{\mathcal{R}} = \{1, 2\}$), and $\gamma = \eta^{(t_r-1)L} = \eta^L$. Substitute them into (4.23), we have the lifetime utility of policy π for $v \in \{\underline{\mathcal{V}}, \bar{\mathcal{V}}\}$,

$$\begin{aligned} W_v^\pi &= U_v^{a_v} + \left[\eta^L(1 - P_{\bar{\mathcal{R}}\to 1}^{a_v}) + \eta^L P_{\bar{\mathcal{R}}\to 1}^{a_v}\right]\left[\mathbf{v}(\underline{\mathcal{V}})W_{\underline{\mathcal{V}}}^\pi + \mathbf{v}(\bar{\mathcal{V}})W_{\bar{\mathcal{V}}}^\pi\right] \\ &= U_v^{a_v} + \eta^L\left[\mathbf{v}(\underline{\mathcal{V}})W_{\underline{\mathcal{V}}}^\pi + \mathbf{v}(\bar{\mathcal{V}})W_{\bar{\mathcal{V}}}^\pi\right]. \end{aligned} \tag{I.1}$$

Similarly, with the one-shot deviation to a_v', the lifetime utility becomes,

$$W_v^{a_v',\pi} = U_v^{a_v'} + \eta^L\left[\mathbf{v}(\underline{\mathcal{V}})W_{\underline{\mathcal{V}}}^\pi + \mathbf{v}(\bar{\mathcal{V}})W_{\bar{\mathcal{V}}}^\pi\right]. \tag{I.2}$$

Therefore, $W_v^\pi - W_v^{a_v',\pi} = U_v^{a_v} - U_v^{a_v'}$, and we only need to compare the expected short-term payoffs when exam each policy using the one-shot deviation principle.

• $\{a_{\underline{\mathcal{V}}}, a_{\bar{\mathcal{V}}}\} = \{2, 2\}$: By solving (4.20), we have $y = \mathbf{x}(2) = 1$ and $\mathbf{x}(1) = 0$. At view $v \in \{\underline{\mathcal{V}}, \bar{\mathcal{V}}\}$, with action 2, a user will help upload with probability 1. His/her reputation will be lowered to 1 with probability $P_{\bar{\mathcal{R}}\to 1}^{2} = 0$, and he/she is a beneficial user with probability 1. Since other users all have reputation no less than $t_r - 1 = 1$, and use the policy $\{2,2\}$, he/she can receive others' help whenever he/she needs in the next L segment, and $G_{2,v} = g_v$. Thus, the expected immediate payoff is $U_v^{a_v=2} = -c + g_v$.

As discussed in Chapter 4.2, action $t_r = 2$ dominates action 1, we only need to exam the one-shot deviation to $R+1 = 3$. By taking action 3, he/she will help to upload with probability 0, and his/her reputation falls to 1 with probability $P^3_{\mathcal{R}\to 1} = 1$. Thus, he/she cannot receive others' help, and the expected immediate gain 0. Therefore, the expected short-term payoff by taking action $R+1$ is $U_v^{a'_v=3} = 0$.

Compare $U_v^{a_v=2}$ and $U_v^{a'_v=3}$, and we have $U_v^{a_v=2} - U_v^{a'_v=3} = g_v - c$. Given that $g_{\bar{\mathcal{V}}} > g_{\underline{\mathcal{V}}}$, if we have $g_{\underline{\mathcal{V}}} \geq c$, then $U_v^{a_v=2} - U_v^{a'_v=3} \geq 0$ for all $v \in \{\underline{\mathcal{V}}, \bar{\mathcal{V}}\}$, and thus, $\{a_{\underline{\mathcal{V}}}, a_{\bar{\mathcal{V}}}\} = \{2,2\}$ is an equilibrium policy.

- $\{a_{\underline{\mathcal{V}}}, a_{\bar{\mathcal{V}}}\} = \{3,2\}$: Similar to the above analysis, by solving (4.20), we have $y = \mathbf{x}(2) = \frac{1}{1+\mathbf{v}(\underline{\mathcal{V}})}$ and $\mathbf{x}(1) = \frac{\mathbf{v}(\underline{\mathcal{V}})}{1+\mathbf{v}(\underline{\mathcal{V}})}$. We then first exam the one-shot deviation principle at view $\underline{\mathcal{V}}$. By taking action $a_{\underline{\mathcal{V}}} = 3$, he/she will help upload with probability 0. His/her reputation falls to 1 with probability $P^3_{\mathcal{R}\to 1} = \frac{1}{1+\mathbf{v}(\underline{\mathcal{V}})}$ and he/she is a beneficial user with probability $\frac{\mathbf{v}(\underline{\mathcal{V}})}{1+\mathbf{v}(\underline{\mathcal{V}})}$. Since other users all have reputation no less than $t_r - 1 = 1$ and they only cooperate at $\bar{\mathcal{V}}$ with probability $\mathbf{v}(\bar{\mathcal{V}})$, thus, his/her expected short-term gain is $\frac{\mathbf{v}(\underline{\mathcal{V}})}{1+\mathbf{v}(\underline{\mathcal{V}})} G_{2,\underline{\mathcal{V}}} = \frac{\mathbf{v}(\underline{\mathcal{V}})}{1+\mathbf{v}(\underline{\mathcal{V}})} \mathbf{v}(\bar{\mathcal{V}}) g_{\underline{\mathcal{V}}}$. Then, his/her expected short-term payoff is $U_{\underline{\mathcal{V}}}^{a_{\underline{\mathcal{V}}}=3} = \frac{\mathbf{v}(\underline{\mathcal{V}})}{1+\mathbf{v}(\underline{\mathcal{V}})} \mathbf{v}(\bar{\mathcal{V}}) g_{\underline{\mathcal{V}}}$.

Since action $t_r = 2$ dominates action 1, we only need to study the one-shot deviation to $a'_{\underline{\mathcal{V}}} = 2$. By taking action $a'_{\underline{\mathcal{V}}} = 2$, he/she will help upload with probability $\frac{1}{1+\mathbf{v}(\underline{\mathcal{V}})}$. His/her reputation will be lowered to 1 with probability $P^2_{\mathcal{R}\to 1} = 0$, and he/she is a beneficial user with probability 1. Thus, he/she will receive the expected short-term gain $g_{\underline{\mathcal{V}}} \mathbf{v}(\bar{\mathcal{V}})$, and his/her expected immediate payoff is $U_{\underline{\mathcal{V}}}^{a'_{\underline{\mathcal{V}}}=2} = -\frac{c}{1+\mathbf{v}(\underline{\mathcal{V}})} + g_{\underline{\mathcal{V}}} \mathbf{v}(\bar{\mathcal{V}})$. Compare $U_{\underline{\mathcal{V}}}^{a_{\underline{\mathcal{V}}}=3}$ and $U_{\underline{\mathcal{V}}}^{a'_{\underline{\mathcal{V}}}=2}$, and we have

$$\begin{aligned} U_{\underline{\mathcal{V}}}^{a_{\underline{\mathcal{V}}}=3} - U_{\underline{\mathcal{V}}}^{a'_{\underline{\mathcal{V}}}=2} &= \frac{\mathbf{v}(\underline{\mathcal{V}})}{1+\mathbf{v}(\underline{\mathcal{V}})} g_{\underline{\mathcal{V}}} \mathbf{v}(\bar{\mathcal{V}}) - \left(-\frac{c}{1+\mathbf{v}(\underline{\mathcal{V}})} + g_{\underline{\mathcal{V}}} \mathbf{v}(\bar{\mathcal{V}}) \right) \\ &= \frac{1}{1+\mathbf{v}(\underline{\mathcal{V}})} (c - g_{\underline{\mathcal{V}}} \mathbf{v}(\bar{\mathcal{V}})). \end{aligned} \tag{I.3}$$

Thus, to resist one-shot deviation at $\underline{\mathcal{V}}$, we should have $(c - g_{\underline{\mathcal{V}}}\mathbf{v}(\bar{\mathcal{V}})) \geq 0$.

We then exam the one-shot deviation principle at view $\bar{\mathcal{V}}$. Following a similar procedure for the analysis at view $\underline{\mathcal{V}}$, we can derive $U_{\bar{\mathcal{V}}}^{a_{\bar{\mathcal{V}}}=2}$ and $U_{\bar{\mathcal{V}}}^{a'_{\bar{\mathcal{V}}}=3}$, and compare them as

$$U_{\bar{\mathcal{V}}}^{a_{\bar{\mathcal{V}}}=2} - U_{\bar{\mathcal{V}}}^{a'_{\bar{\mathcal{V}}}=3} = \frac{1}{1+\mathbf{v}(\underline{\mathcal{V}})}(g_{\bar{\mathcal{V}}}\mathbf{v}(\bar{\mathcal{V}}) - c). \tag{I.4}$$

Thus, to resist one-shot deviation at $\bar{\mathcal{V}}$, we should have $(g_{\bar{\mathcal{V}}}\mathbf{v}(\bar{\mathcal{V}}) - c) \geq 0$. Thus, based on the above analysis, only when $g_{\bar{\mathcal{V}}}\mathbf{v}(\bar{\mathcal{V}}) \geq c \geq g_{\underline{\mathcal{V}}}\mathbf{v}(\bar{\mathcal{V}})$, we have both $U_{\underline{\mathcal{V}}}^{a_{\underline{\mathcal{V}}}=3} - U_{\underline{\mathcal{V}}}^{a'_{\underline{\mathcal{V}}}=2} \geq 0$ and $U_{\bar{\mathcal{V}}}^{a_{\bar{\mathcal{V}}}=2} - U_{\bar{\mathcal{V}}}^{a'_{\bar{\mathcal{V}}}=3} \geq 0$, and $\{a_{\underline{\mathcal{V}}}, a_{\bar{\mathcal{V}}}\} = \{3, 2\}$ is an equilibrium policy.

• $\{a_{\underline{\mathcal{V}}}, a_{\bar{\mathcal{V}}}\} = \{3, 3\}$: In this case, $y = \mathbf{x}(2) = 0.5$ and $\mathbf{x}(1) = 0.5$. Since no user cooperates, no user can gain from others' help and $G_{2,v} = G_{1,v} = 0$, while taking action $t_r = 2$ and cooperating with beneficial users only introduces a cost due to helping upload with probability 0.5. Thus, using action $R + 1 = 3$ and playing non-cooperatively is a dominant strategy. Therefore, $\{a_{\underline{\mathcal{V}}}, a_{\bar{\mathcal{V}}}\} = \{3, 3\}$ is always an equilibrium policy.

• $\{a_{\underline{\mathcal{V}}}, a_{\bar{\mathcal{V}}}\} = \{2, 3\}$: This action policy is symmetric with $\{a_{\underline{\mathcal{V}}}, a_{\bar{\mathcal{V}}}\} = \{3, 2\}$ that we discussed earlier. Thus, the cost range for $\{a_{\underline{\mathcal{V}}}, a_{\bar{\mathcal{V}}}\} = \{2, 3\}$ being an equilibrium policy can be symmetrically written as $g_{\underline{\mathcal{V}}}\mathbf{v}(\underline{\mathcal{V}}) \geq c \geq g_{\bar{\mathcal{V}}}\mathbf{v}(\underline{\mathcal{V}})$. However, since $g_{\bar{\mathcal{V}}} > g_{\underline{\mathcal{V}}}$, this range is empty, and thus, $\{a_{\underline{\mathcal{V}}}, a_{\bar{\mathcal{V}}}\} = \{2, 3\}$ cannot be an equilibrium policy. ■

Appendix J

Proof of Proposition 4.6

Proof: In Chapter 4.3.1.5, we focus on the proof of (a. In this proof, we prove all these four statements in this proposition. Thus, similar to the proof for Proposition 4.5, we first give the general analysis, and then exam each policy one by one using one-shot deviation principle.

(4.23) first gives the lifetime utility with aggregated views. Similarly, the lifetime utility of a one-shot deviation to a'_v at $v \in \{\underline{\mathcal{V}}, \bar{\mathcal{V}}\}$ can be written as

$$W_v^{a'_v,\pi} = U_{\underline{\mathcal{V}}}^{a'_v} + \left[\eta^L(1 - P_{\bar{\mathcal{R}}\to 1}^{a'_v}) + \gamma P_{\bar{\mathcal{R}}\to 1}^{a'_v}\right]\left[\mathbf{v}(\underline{\mathcal{V}})W_{\underline{\mathcal{V}}}^{\pi} + \mathbf{v}(\bar{\mathcal{V}})W_{\bar{\mathcal{V}}}^{\pi}\right]. \tag{J.1}$$

To compare W_v^{π} with its one-shot deviation $W_v^{a'_v,\pi}$, we have

$$\begin{aligned} W_v^{\pi} - W_v^{a'_v,\pi} =& U_{\underline{\mathcal{V}}}^{a_v} - U_{\underline{\mathcal{V}}}^{a'_v} + \left[\mathbf{v}(\underline{\mathcal{V}})W_{\underline{\mathcal{V}}}^{\pi} + \mathbf{v}(\bar{\mathcal{V}})W_{\bar{\mathcal{V}}}^{\pi}\right] \\ &\times\left\{\left[\eta^L(1 - P_{\bar{\mathcal{R}}\to 1}^{a_v}) + \gamma P_{\bar{\mathcal{R}}\to 1}^{a_v}\right] - \left[\eta^L(1 - P_{\bar{\mathcal{R}}\to 1}^{a'_v}) + \gamma P_{\bar{\mathcal{R}}\to 1}^{a'_v}\right]\right\}. \end{aligned} \tag{J.2}$$

Note that (J.2) has the recursive term $W_{\underline{\mathcal{V}}}^{\pi}$ and $W_{\bar{\mathcal{V}}}^{\pi}$ at the right side, which we should solve first. Note that (4.23) can be viewed as a linear system, and we can solve $W_{\underline{\mathcal{V}}}^{\pi}$

and $W^{\pi}_{\bar{\mathcal{V}}}$ from it as

$$\begin{cases} W^{\pi}_{\underline{\mathcal{V}}} = \dfrac{U^{a_{\underline{\mathcal{V}}}}_{\underline{\mathcal{V}}} - \mathbf{v}(\bar{\mathcal{V}})\left[\left((1-P^{a_{\bar{\mathcal{V}}}}_{\bar{\mathcal{R}}\to 1})\eta - \gamma P^{a_{\bar{\mathcal{V}}}}_{\bar{\mathcal{R}}\to 1}\right)U^{a_{\underline{\mathcal{V}}}}_{\underline{\mathcal{V}}} - \left((1-P^{a_{\underline{\mathcal{V}}}}_{\bar{\mathcal{R}}\to 1})\eta - \gamma P^{a_{\underline{\mathcal{V}}}}_{\bar{\mathcal{R}}\to 1}\right)U^{a_{\bar{\mathcal{V}}}}_{\bar{\mathcal{V}}}\right]}{1-\left((1-P^{a_{\bar{\mathcal{V}}}}_{\bar{\mathcal{R}}\to 1})\eta - \gamma P^{a_{\bar{\mathcal{V}}}}_{\bar{\mathcal{R}}\to 1}\right)\mathbf{v}(\bar{\mathcal{V}}) - \left((1-P^{a_{\underline{\mathcal{V}}}}_{\bar{\mathcal{R}}\to 1})\eta - \gamma P^{a_{\underline{\mathcal{V}}}}_{\bar{\mathcal{R}}\to 1}\right)\mathbf{v}(\underline{\mathcal{V}})}, \\ W^{\pi}_{\bar{\mathcal{V}}} = \dfrac{U^{a_{\bar{\mathcal{V}}}}_{\bar{\mathcal{V}}} + \mathbf{v}(\underline{\mathcal{V}})\left[\left((1-P^{a_{\bar{\mathcal{V}}}}_{\bar{\mathcal{R}}\to 1})\eta - \gamma P^{a_{\bar{\mathcal{V}}}}_{\bar{\mathcal{R}}\to 1}\right)U^{a_{\underline{\mathcal{V}}}}_{\underline{\mathcal{V}}} - \left((1-P^{a_{\underline{\mathcal{V}}}}_{\bar{\mathcal{R}}\to 1})\eta - \gamma P^{a_{\underline{\mathcal{V}}}}_{\bar{\mathcal{R}}\to 1}\right)U^{a_{\bar{\mathcal{V}}}}_{\bar{\mathcal{V}}}\right]}{1-\left((1-P^{a_{\bar{\mathcal{V}}}}_{\bar{\mathcal{R}}\to 1})\eta - \gamma P^{a_{\bar{\mathcal{V}}}}_{\bar{\mathcal{R}}\to 1}\right)\mathbf{v}(\bar{\mathcal{V}}) - \left((1-P^{a_{\underline{\mathcal{V}}}}_{\bar{\mathcal{R}}\to 1})\eta - \gamma P^{a_{\underline{\mathcal{V}}}}_{\bar{\mathcal{R}}\to 1}\right)\mathbf{v}(\underline{\mathcal{V}})}. \end{cases} \tag{J.3}$$

We then substitute (J.3) into (J.2) and finish the comparison of W^{π}_{v} and $W^{a'_v,\pi}_{v}$. The above analysis explain the general approach how we exam each policy with one-shot deviation principle. We then exam each policy one by one. Note that since action t_r and $R+1$ dominate all other actions, we only need to study the one-shot deviation to t_r or $R+1$.

• $\{a_{\underline{\mathcal{V}}}, a_{\bar{\mathcal{V}}}\} = \{t_r, t_r\}$: By solving (4.20), we first have $y = 1$ and $\mathbf{x}(r) = 0$ for $1 \leq r \leq t_r - 1$. Therefore, for a user with reputation no less than $t_r - 1$ and at view $v \in \{\underline{\mathcal{V}}, \bar{\mathcal{V}}\}$, by taking action t_r, he/she will help to upload with probability 1, his/her reputation will be lowered to 1 with probability $P^{t_r}_{\bar{\mathcal{R}}\to 1} = 0$, and he/she is a beneficial user with probability 1. Since others are all beneficial users and use policy $\{t_r, t_r\}$, he/she will always receive others' help in the next L segment and have the expected short-term gain g_v. Therefore, his/her expected short-term payoff by taking action t_r is $U^{a_v = t_r}_{v} = -c + g_v$.

We only need to exam the one-shot deviation to $a'_v = R+1$. With $a'_v = R+1$, he/she will help to upload with probability 0, and his/her reputation falls to 1 with probability $P^{R+1}_{\bar{\mathcal{R}}\to 1} = 1$. Thus, he/she cannot receive others' help, and the expected short-term gain is 0. Therefore, the expected short-term utility is $U^{a'_v = R+1}_{v} = 0$. We then substitute $U^{a_v = t_r}_{v}$, $P^{t_r}_{\bar{\mathcal{R}}\to 1}$, $U^{a'_v = R+1}_{v}$ and $P^{R+1}_{\bar{\mathcal{R}}\to 1}$ into (J.3) and compare $W_v(\pi)$ and $W^{a'_v,\pi}_{v}$ for both $\underline{\mathcal{V}}$ and $\bar{\mathcal{V}}$. We then have

$$\begin{aligned} W^{\pi}_{\bar{\mathcal{V}}} - W^{a'_{\bar{\mathcal{V}}} = R+1,\pi}_{\bar{\mathcal{V}}} &= \frac{(\eta^L - \gamma)(g_{\bar{\mathcal{V}}} - g_{\underline{\mathcal{V}}})\mathbf{v}(\bar{\mathcal{V}}) - (c - g_{\underline{\mathcal{V}}})(1-\gamma)}{1-\eta^L} + (g_{\bar{\mathcal{V}}} - g_{\underline{\mathcal{V}}}), \\ W^{\pi}_{\underline{\mathcal{V}}} - W^{a_{\underline{\mathcal{V}}} = R+1,\pi}_{\underline{\mathcal{V}}} &= \frac{(\eta^L - \gamma)(g_{\bar{\mathcal{V}}} - g_{\underline{\mathcal{V}}})\mathbf{v}(\bar{\mathcal{V}}) - (c - g_{\underline{\mathcal{V}}})(1-\gamma)}{1-\eta^L}. \end{aligned} \tag{J.4}$$

It is easy to observe that $W^{\pi}_{\bar{\mathcal{V}}} - W^{a_{\bar{\mathcal{V}}} = R+1,\pi}_{\bar{\mathcal{V}}} > W^{\pi}_{\underline{\mathcal{V}}} - W^{a_{\underline{\mathcal{V}}} = R+1,\pi}_{\underline{\mathcal{V}}}$. Thus, as long as $W_{\underline{\mathcal{V}}}^{\pi} - W_{\underline{\mathcal{V}}}^{a_{\underline{\mathcal{V}}}=R+1,\pi} = \frac{(\eta^L-\gamma)(g_{\bar{\mathcal{V}}}-g_{\underline{\mathcal{V}}})\mathbf{v}(\bar{\mathcal{V}}) - (c-g_{\underline{\mathcal{V}}})(1-\gamma)}{1-\eta^L} \geq 0$, in other words, $\bar{c}_1 \triangleq \frac{(\eta^L-\gamma)(g_{\bar{\mathcal{V}}}-g_{\underline{\mathcal{V}}})\mathbf{v}(\bar{\mathcal{V}})}{1-\gamma} + g_{\underline{\mathcal{V}}} \geq c$, we have $W_{\bar{\mathcal{V}}}^{\pi} - W_{\bar{\mathcal{V}}}^{a_{\bar{\mathcal{V}}}=R+1,\pi} > W_{\underline{\mathcal{V}}}^{\pi} - W_{\underline{\mathcal{V}}}^{a_{\underline{\mathcal{V}}}=R+1,\pi} \geq 0$, and $\{a_{\underline{\mathcal{V}}}, a_{\bar{\mathcal{V}}}\} = \{t_r, t_r\}$ is an equilibrium policy.

- $\{a_{\underline{\mathcal{V}}}, a_{\bar{\mathcal{V}}}\} = \{R+1, t_r\}$: We first have $y = \frac{\sqrt{(1+\mathbf{v}(\underline{\mathcal{V}}))^2+4(t_r-2)\mathbf{v}(\underline{\mathcal{V}})}-(1+\mathbf{v}(\underline{\mathcal{V}}))}{2(t_r-2)\mathbf{v}(\underline{\mathcal{V}})}$, and $\mathbf{x}(r) = \frac{1-y}{t_r-1}$ for $1 \leq r \leq t_r - 1$. Therefore, for a user with reputation no less than $t_r - 1$ and at view $\underline{\mathcal{V}}$, by taking action $a_{\underline{\mathcal{V}}} = R+1$, he/she will help upload with probability 0, and his/her reputation falls to 1 with probability $P_{\bar{\mathcal{R}}\to 1}^{R+1} = y$, and he is a beneficial user with probability $(1-y)$. Thus, he/she receives expected short-term gain $(1-y)G_{\min(r+1,R),\underline{\mathcal{V}}}$. To derive $G_{\min(r+1,R),\underline{\mathcal{V}}}$, since other users only cooperate when they have reputation no less than $t_r - 1$ and are at view $\bar{\mathcal{V}}$, we have $G_{\min(r+1,R),\underline{\mathcal{V}}} = [y + \mathbf{x}(t_r-1)]\mathbf{v}(\bar{\mathcal{V}})g_{\underline{\mathcal{V}}}$. Thus, his/her expected short-term payoff is $U_{\underline{\mathcal{V}}}^{a_{\underline{\mathcal{V}}}=R+1} = (1-y)g_{\underline{\mathcal{V}}}[y+\mathbf{x}(t_r-1)]\mathbf{v}(\bar{\mathcal{V}})$. When he/she is at view $\bar{\mathcal{V}}$, by taking action $a_{\bar{\mathcal{V}}} = t_r$ defined in the policy, he/she will help upload with probability y and his/her reputation falls to 1 with probability $P_{\bar{\mathcal{R}}\to 1}^{t_r} = 0$. He/she is a beneficial user with probability 1. Thus, he/she receives expected short-term gain $G_{\min(r+1,R),\underline{\mathcal{V}}} = g_{\bar{\mathcal{V}}}[y+\mathbf{x}(t_r-1)]\mathbf{v}(\bar{\mathcal{V}})$, and his/her expected short-term payoff is $U_{\bar{\mathcal{V}}}^{a_{\bar{\mathcal{V}}}=t_r} = -yc + g_{\bar{\mathcal{V}}}[y+\mathbf{x}(t_r-1)]\mathbf{v}(\bar{\mathcal{V}})$.

Then, for the one-shot deviation at view $\underline{\mathcal{V}}$, he/she can only deviate to $a'_{\underline{\mathcal{V}}} = t_r$. By taking action t_r, he/she will help to upload with probability y and the probability that his/her reputation falls to 1 is $P_{\bar{\mathcal{R}}\to 1}^{t_r} = 0$. Thus, he/she is a beneficial user with probability 1, and receives expected short-term payoff $U_{\underline{\mathcal{V}}}^{a'_{\underline{\mathcal{V}}}=t_r} = -yc + g_{\underline{\mathcal{V}}}[y + \mathbf{x}(t_r-1)]\mathbf{v}(\bar{\mathcal{V}})$. Substitute $U_{\underline{\mathcal{V}}}^{a'_{\underline{\mathcal{V}}}=t_r}$, $P_{\bar{\mathcal{R}}\to 1}^{t_r}$, $U_{\underline{\mathcal{V}}}^{a_{\underline{\mathcal{V}}}=R+1}$ and $P_{\bar{\mathcal{R}}\to 1}^{R+1}$ into (J.3) and (J.2), and we compare $W_{\underline{\mathcal{V}}}^{\pi}$ and $W_{\underline{\mathcal{V}}}^{a'_{\underline{\mathcal{V}}}=t_r,\pi}$ as

$$W_{\underline{\mathcal{V}}}^{\pi} - W_{\underline{\mathcal{V}}}^{a'_{\underline{\mathcal{V}}}=t_r,\pi} = y\left\{c - \frac{[y+\mathbf{x}(t_r-1)]\mathbf{v}(\bar{\mathcal{V}})D_{\underline{\mathcal{V}}}}{1-\eta^L+\eta^L y - y\gamma}\right\}, \tag{J.5}$$

where $D_{\underline{\mathcal{V}}} = \left\{\mathbf{v}(\bar{\mathcal{V}})(\eta^L-\gamma)g_{\bar{\mathcal{V}}} + [1-\gamma-(\eta^L-\gamma)\mathbf{v}(\bar{\mathcal{V}})]g_{\underline{\mathcal{V}}}\right\}$. Thus, to resist the one-shot deviation at view $\underline{\mathcal{V}}$, we need $\left\{c - \frac{[b+\mathbf{x}(t_r-1)]\mathbf{v}(\bar{\mathcal{V}})D_{\underline{\mathcal{V}}}}{1-\eta^L+\eta^L b - b\gamma}\right\} \geq 0$.

For the one-shot deviation at view $\bar{\mathcal{V}}$, he/she can only deviate to $a'_{\bar{\mathcal{V}}} = R+1$.

We can compare $W_{\bar{\mathcal{V}}}^{\pi}$ and $W_{\bar{\mathcal{V}}}^{a'_{\bar{\mathcal{V}}}=R+1,\pi}$ following the same procedure, and have

$$W_{\bar{\mathcal{V}}}^{\pi} - W_{\bar{\mathcal{V}}}^{a'_{\bar{\mathcal{V}}}=R+1,\pi} = y\left\{\frac{[y+\mathbf{x}(t_r-1)]\mathbf{v}(\bar{\mathcal{V}})D_{\bar{\mathcal{V}}}}{1-\eta^L+\eta^L y - y\gamma} - c\right\}, \tag{J.6}$$

where $D_{\bar{\mathcal{V}}} = \Big\{[1-\gamma-(\eta^L-\gamma)\mathbf{v}(\underline{\mathcal{V}})(1-y)]g_{\bar{\mathcal{V}}} + (\eta^L-\gamma)\mathbf{v}(\underline{\mathcal{V}})(1-y)g_{\underline{\mathcal{V}}}\Big\}$. Therefore, to resist the one-shot deviation at view $\bar{\mathcal{V}}$, we need $\left\{\frac{[y+\mathbf{x}(t_r-1)]\mathbf{v}(\bar{\mathcal{V}})D_{\bar{\mathcal{V}}}}{1-\eta^L+\eta^L y-y\gamma} - c\right\} \geq 0$. We then compare $D_{\bar{\mathcal{V}}}$ and $D_{\underline{\mathcal{V}}}$, and have

$$\begin{aligned} D_{\bar{\mathcal{V}}} - D_{\underline{\mathcal{V}}} &= \Big\{(1-\gamma) - (\eta^L-\gamma)[(1-y)+y\mathbf{v}(\bar{\mathcal{V}})]\Big\}(g_{\bar{\mathcal{V}}} - g_{\underline{\mathcal{V}}}) \\ &> (\eta^L-\gamma)[y - y\mathbf{v}(\bar{\mathcal{V}})](g_{\bar{\mathcal{V}}} - g_{\underline{\mathcal{V}}}) > 0. \end{aligned} \tag{J.7}$$

Thus, if we have the cost c in the range

$$\frac{[y+\mathbf{x}(t_r-1)]\mathbf{v}(\bar{\mathcal{V}})D_{\bar{\mathcal{V}}}}{1-\eta^L+\eta^L y-y\gamma} \geq c \geq \frac{[y+\mathbf{x}(t_r-1)]\mathbf{v}(\bar{\mathcal{V}})D_{\underline{\mathcal{V}}}}{1-\eta^L+\eta^L y-y\gamma}, \tag{J.8}$$

we have both $W_{\underline{\mathcal{V}}}^{\pi} - W_{\underline{\mathcal{V}}}^{a'_{\underline{\mathcal{V}}}=t_r,\pi} \geq 0$ and $W_{\bar{\mathcal{V}}}^{\pi} - W_{\bar{\mathcal{V}}}^{a'_{\bar{\mathcal{V}}}=R+1,\pi} \geq 0$, and therefore, $\{a_{\underline{\mathcal{V}}}, a_{\bar{\mathcal{V}}}\} = \{R+1, t_r\}$ is an equilibrium policy.

- $\{a_{\underline{\mathcal{V}}}, a_{\bar{\mathcal{V}}}\} = \{R+1, R+1\}$: In this case, if $t_r = 2$, we have $y = 0.5$ and $\mathbf{x}(1) = 0.5$. If $t_r \geq 3$, we have $y = \frac{\sqrt{t_r-1}-1}{t_r-2}$ and $\mathbf{x}(r) = \frac{1-y}{t_r-1}$ for $1 \leq r \leq t_r - 1$. Similar to the proof of Proposition 4.5c, since no user cooperates, no user can gain from others' help and $G_{r,v} = 0$ for all r and v. However, playing cooperatively with action t_r only introduces a cost due to helping upload with probability y. Thus, using action $R+1$ and playing non-cooperatively is a dominant strategy, from which no one will deviate.
- $\{a_{\underline{\mathcal{V}}}, a_{\bar{\mathcal{V}}}\} = \{t_r, R+1\}$: This policy is symmetric with the policy $\{a_{\underline{\mathcal{V}}}, a_{\bar{\mathcal{V}}}\} = \{R+1, t_r\}$ that we discussed earlier. Thus, the cost range for $\{a_{\underline{\mathcal{V}}}, a_{\bar{\mathcal{V}}}\} = \{t_r, R+1\}$ being an equilibrium policy can be symmetrically written as

$$\frac{[y+\mathbf{x}(t_r-1)]\mathbf{v}(\underline{\mathcal{V}})E_{\underline{\mathcal{V}}}}{1-\eta^L+\eta^L y-y\gamma} \geq c \geq \frac{[y+\mathbf{x}(t_r-1)]\mathbf{v}(\underline{\mathcal{V}})E_{\bar{\mathcal{V}}}}{1-\eta^L+\eta^L y-y\gamma}, \tag{J.9}$$

where we have $E_{\underline{\mathcal{V}}} = \Big\{[1-\gamma-(\eta^L-\gamma)\mathbf{v}(\bar{\mathcal{V}})(1-y)]g_{\underline{\mathcal{V}}} + (\eta^L-\gamma)\mathbf{v}(\bar{\mathcal{V}})(1-y)g_{\bar{\mathcal{V}}}\Big\}$ and $E_{\bar{\mathcal{V}}} = \Big\{\mathbf{v}(\underline{\mathcal{V}})(\eta^L-\gamma)g_{\underline{\mathcal{V}}} + [1-\gamma-(\eta^L-\gamma)\mathbf{v}(\underline{\mathcal{V}})]g_{\bar{\mathcal{V}}}\Big\}$. Compare $E_{\underline{\mathcal{V}}}$ and $E_{\bar{\mathcal{V}}}$, and we have

$$\begin{aligned} E_{\bar{\mathcal{V}}} - E_{\underline{\mathcal{V}}} &= \Big\{(1-\gamma) - (\eta^L-\gamma)[(1-y) + y\mathbf{v}(\underline{\mathcal{V}})]\Big\}(g_{\bar{\mathcal{V}}} - g_{\underline{\mathcal{V}}}) \\ &> (\eta^L-\gamma)[y - y\mathbf{v}(\underline{\mathcal{V}})](g_{\bar{\mathcal{V}}} - g_{\underline{\mathcal{V}}}) > 0. \end{aligned} \tag{J.10}$$

Thus, the cost range in (J.9) is empty, and $\{a_{\underline{\mathcal{V}}}, a_{\bar{\mathcal{V}}}\} = \{R+1, t_r\}$ is not an equilibrium policy. ■