University of Alberta

Cooperative Wireless Multicast: Cooperation Strategy and Incentive Mechanism

by

Binglai Niu

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

Master of Science
in
Signal and Image Processing

Department of Electrical and Computer Engineering

Fall 2010
Edmonton, Alberta

Examining Committee

Dr. H. Vicky Zhao (co-supervisor), Electrical and Computer Engineering

Dr. Hai Jiang (co-supervisor), Electrical and Computer Engineering

Dr. Yindi Jing, Electrical and Computer Engineering

Dr. Ke-li Xu, Finance and Management Science

Abstract

Multicast is a bandwidth efficient mechanism to provide wireless services for a group of nodes. Providing reliable wireless multicast is challenging due to channel fading. This thesis investigates cooperation among receiving nodes to enhance the reliability of wireless multicast. A time division based cooperative multicast strategy is proposed, and the optimal scheduling scheme is found to maximize the system throughput. It is shown that the optimal relay number is bounded by a threshold, and the optimal time allocation can be found using an efficient algorithm. Numerical results show that the proposed strategy can enhance network performance when the average channel condition between receiving nodes is better than that of the direct link. To provide incentive for cooperation, this thesis further studies the interactions among selfish nodes using game theoretic approaches. The cooperative multicast process is modeled as a repeated game and the desired cooperation state which satisfies the absolute fairness and the Pareto optimality criteria is found. A *Worst Behavior Tit-for-Tat* incentive strategy is designed to enforce cooperation and its effectiveness is studied under both the perfect and the imperfect monitoring scenarios. To address the issue of imperfect monitoring, an interval based estimation method is proposed. Simulation results show that the proposed strategy can enforce cooperation efficiently even the monitoring is imperfect.

Acknowledgements

First of all, I would like to express my eternal gratitude to my supervisors, Dr. H.Vicky Zhao and Dr. Hai Jiang, for providing me this valuable opportunity to work under their guidance in every stage of my research. Their advice and systematic training help me build up a solid foundation in this research field, and prepare me for any challenge in the future study.

I shall extend my sincere thanks to my committee members, Dr. Yindi Jing and Dr. Ke-li Xu, for providing me valuable feedbacks and advice in improving this thesis.

I owe my deepest gratitude to my family. Their support, love and encouragement give me confidence to overcome any challenge in my life.

I am deeply grateful for my colleagues in the multimedia laboratory, Zhaorui Liu, Xingyu Li and Bo Hu, for providing me a comfortable study environment. Also, I would like to thank all my friends in Edmonton for their help.

Last but not least, I gratefully acknowledge the financial support of NSERC, the Department of Electrical and Computer Engineering at University of Alberta, and my supervisors.

Contents

List of Tables

List of Figures

Acronyms

Acronyms	Definition
AWGN	additive white Gaussian noise
CCS	credit clearance service
i.i.d.	independent and identically distributed
IPTV	Internet protocol television
MBMS	multimedia broadcast/multicast service
MLAB	maximum lifetime accumulative broadcast
NCBC	network coding based cooperation
NE	Nash equilibrium
R-DSTC	randomized distributed space time codes
SNR	signal to noise ratio
STCBC	space-time coding based cooperation
TDBC	time division based cooperation
3G	3rd generation
WBTFT	worst behavior tit-for-tat
WiMax	Worldwide Interoperability for Microwave Access

List of Symbols

Symbol	Definition
$\boldsymbol{A}$	action profile in a static game
$\mathcal{A}$	joint set of action space in a static game
$\mathcal{A}_i$	action space of node i in a static game
$\boldsymbol{A}^k$	action profile in stage k
$\boldsymbol{A}^*$	desired NE action profile
A_i	node i's action in a static game
A_i^k	node i's action in stage k
b_i^k	node i's behavior in stage k
$\widetilde{b}_{ij}^k$	node j's estimation of node i's behavior in stage k
$b_\dagger^k$	worst behavior in stage k
$b^\dagger$	desired behavior in the current stage according to WBTFT
b_T	behavior threshold in WBTFT strategy
$\hat{b}_l$	the lth behavior level
b_{jil}^k	node j's lth behavior level estimated by node i in stage k
$\mathscr{B}$	set of behavior levels
$\mathscr{B}_{ji}^k$	set of node j's behavior levels estimated by node i in stage k
c_0	cost for transmitting a packet with unit power
C_{si}	channel capacity between the source and node i
$\overline{D}$	average value of $N_0(2^{\gamma_r}-1)/\sigma_{ij}^2$

$\overline{D}_{ji}^{k}$	average value of $N_0(1-2^{\gamma_r})/\sigma_{ji}^2$ in stage k
h_{si}	channel gain between the source and node i
L	parameter for resuming cooperation in WBTFT strategy
L_p	packet size
m_i	number of relays to be selected when i nodes succeed in T_1
m^*	relay number threshold
m_i^{opt}	optimal relay number when i nodes succeed in T_1
M	number of time slots within a stage
n_{ij}^k	number of packets that node j received from node i in stage k
$\overline{n}$	expected number of packets that a node should transmit in a stage
N	total number of receiving nodes
$\mathcal{N}$	set of the nodes
N_0	noise power per unit bandwidth
p_s	packet delivery success probability from the source to a node
p_r	packet delivery success probability from the relay to a node
p_{ij}	packet delivery success probability from node i to node j
P_s	transmission power of the source node
P_r	transmission power of the relay
P_i	transmission power of node i
$P_{\max}$	maximum power that a node would like to offer for cooperation
$P^{\dagger}$	desired action in the current stage according to WBTFT strategy
P_i'	node i's deviating action in current stage
$\hat{P}_l$	the lth power level
$\mathscr{P}$	set of power levels
q_1	probability of a node being selected as a relay
q_2	probability that node i fails in T_1 and node j is selected in T_2
r_0	reward for correctly receiving a packet

s_i	node i's decision making strategy
s^*	WBTFT strategy
s'	deviation strategy
T	length of the whole time slot for transmitting one packet
T_1	length of the subslot for source transmission
T_2	length of the subslot for node cooperation
$\boldsymbol{v}^k$	expected payoff profile in stage k
v_i	node i's payoff in a static game
v_i^k	node i's expected payoff in stage k
V_i	node i's long term expected payoff
$\widetilde{V}_i$	node i's estimated long term payoff
$\mathcal{V}$	payoff profile in a static game
$\mathcal{V}^e$	enforceable payoff profile in a static game
$\mathcal{V}^f$	feasible payoff profile in a static game
y_{ij}^k	node j's observed information from node i in stage k
α	time allocation parameter
β_{jil}^k	left interval boundary of behavior level $\hat{b}_{jil}^k$
β_l	left interval boundary of behavior level $\hat{b}_l$
γ_i	transmission rate of the node i
γ_r	transmission rate of the relay
γ_s	transmission rate of the source
Γ_1	average received SNR between the source and a node
Γ_2	average received SNR between the relay and a node
μ	total number of successful nodes in T
μ_1	number of successful nodes in T_1
μ_2	number of successful nodes in T_2
$\overline{\tau}$	average value of $N_0/(P_r\sigma_{ij}^2)$

σ_r^2	variance of the channel gain between the relay and a node
σ_s^2	variance of the channel gain between the source and a node
δ	common discount factor
ε	interval partition parameter
ε_T	estimation error threshold in refined WBTFT strategy
η	bounded rational parameter
φ_i	transmission probability of node i

Chapter 1

Introduction

1.1 Motivation

In the past decade, the growing demand for group-oriented, high speed, high quality wireless communications and the scarcity of conventional frequency spectrum leverage the use of multicast technique, where information data are delivered to a group of nodes simultaneously [1–3]. Applications include the Internet Protocol Television (IPTV) over WiMax [4] and multimedia broadcast/multicast service (MBMS) in 3G networks [5]. However, the fading characteristic of wireless channels poses great challenge in wireless multicast. For instance, when a source tries to multicast to multiple wireless nodes, generally these nodes undergo different channel fading. To guarantee that nodes with bad channels receive the data correctly, the source may have to use a low transmission rate or even retransmissions, which greatly limits the network performance. Therefore, it is important to design reliable wireless multicast techniques.

Recently, cooperative multicast has emerged as an effective technique to combat fading, which exploits the benefits of spatial diversity among the receiving nodes [6]. In cooperative multicast networks, data packets are not only transmitted by the source, but also rebroadcasted by nodes who successfully receive them from the source. Therefore, a node can receive copies of the data from multiple paths. Although a packet might be lost in the direct link between the source and a node, it can still be successfully delivered through other paths with good channel conditions. In general, cooperation among receiving nodes can increase the packet delivery success probability and the network performance can be greatly improved.

Although several cooperative multicast strategies have been proposed and investigated in existing works, there are still some fundamental issues that need to be exploited. For instance, in delay-sensitive applications, where the time for delivering a packet is limited, how to allocate the time for source transmission and relay cooperation to maximize the system throughput? When a set of successful nodes are available, how to determine the number of relays to obtain the optimal performance? These issues have not been well studied in the literature, which inspires the first part of the research in this thesis.

Another challenging issue in cooperative multicast is the incentive mechanism design for cooperation stimulation. In most of the existing cooperative strategies, it is assumed that nodes will cooperate unconditionally and forward data whenever being selected as relays. However, in many applications, cooperative behavior may not be carried out since the wireless nodes tend to be "selfish". In wireless multicast networks, a selfish node is able to hear the broadcasted data from any relay and prefers to taking advantage of others but does not contribute at all, since relaying data would cost extra energy. Therefore, before the cooperative multicast strategy can be successfully deployed, it is critical to design incentive mechanisms to stimulate cooperation among selfish nodes.

In the literature, cooperation stimulation strategies have been proposed for unicast ad-hoc networks [7], where nodes communicate with each other through point-to-point links. The special features of wireless multicast networks bring new challenges for cooperation stimulation. First, in multicast networks, communication is point-to-multiple points. The heterogeneous behaviors of nodes make it challenging to design a general decision rule, since a node's decision may depend on the behaviors of all its neighbors. Second, due to the existence of noise and fading in wireless channels, packets may be dropped during transmission. Therefore, a cooperative behavior may not provide the expected service, and hence may not get any reward. Moreover, in multicast applications, a node may make decisions based on the monitoring results of all others' behaviors. Since the monitoring process is usually imperfect, the monitoring results may be erroneous, which may result in frequent undesired decisions and may discourage user cooperation. The above issues have not been discussed for wireless multicast networks, which motivates the second part of the research in this thesis.

1.2 Contributions and Thesis Outline

This thesis studies node cooperation and incentive strategies in wireless multicast networks. The major contributions lie in two aspects:

- A simple time division based cooperative multicast strategy is proposed and analyzed for delay-sensitive applications. The relay selection problem is studied and the optimal relay number is derived. It is shown that the optimal relaying strategy is a pure threshold policy, where the optimal relay number is bounded by a threshold. Moreover, to avoid the complexity in exhaustive search for the optimal time allocation between source and relay transmissions, a simple algorithm is proposed that can find the optimal time allocation quickly.
- An incentive strategy is designed to stimulate cooperation among selfish nodes in cooperative multicast networks. The cooperative multicast process is mod-

eled as a repeated game, and the desired cooperation state is analyzed. To stimulate cooperation, a *Worst Behavior Tit-for-Tat* incentive strategy is proposed, which can achieve cooperation efficiently. Moreover, an interval based estimation method is developed to address the issue of imperfect monitoring.

The rest of this thesis is organized as follows. In Chapter 2, the related works on cooperative multicast and incentive mechanism design are reviewed, and some background knowledge of game theory is provided. Chapter 3 studies a time division based cooperative multicast strategy in wireless networks. The optimal relay number is derived and a simple algorithm is proposed to find the optimal time allocation efficiently. In Chapter 4, a *Worst Behavior Tit-for-Tat* incentive strategy is proposed for cooperative multicast networks. Game theoretic approaches are employed and equilibrium of the strategy is analyzed under both the perfect and the imperfect monitoring scenarios. An interval based estimation method is also proposed to address the issue of imperfect monitoring. Finally, Chapter 5 concludes this thesis and discusses the future work.

Chapter 2

Literature Review

2.1 Cooperative Wireless Multicast

Wireless multicast is a bandwidth efficient communication technique, where the same data packets are delivered to multiple destinations simultaneously [1–3, 8, 9]. However, the fading characteristic of wireless channels poses great challenge in providing reliable multicast transmissions. Traditionally, to guarantee that nodes with bad channels receive the data correctly, the source may have to use a low transmission rate [10], or employ retransmissions until every node receives the packet successfully [11], which greatly limits the network performance. Recent advances on cooperative communication provide an alternative way to combat fading, which exploits the benefits of spatial diversity among receiving nodes [12–15]. Motivated by this idea, cooperative multicast has emerged as an efficient method to provide reliable wireless multicast services [6, 16–22].

2.1.1 Two-Phase Cooperative Multicast Model

Fig. 2.1 depicts a single-hop cooperative multicast system, where a source multicasts data packets to a group of wireless nodes. Generally, the cooperative multicast

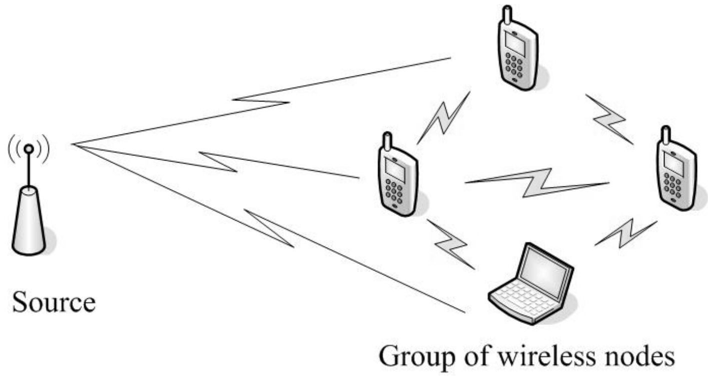

Fig. 2.1. Single-hop cooperative multicast system.

process contains two phases. In the first phase, the source multicasts data packets to the group, and nodes with good channel conditions can receive the packets successfully. In the second phase, some of the successful nodes who correctly receive the packets from the source are selected as relays to rebroadcast the packets in orthogonal channels. It can be seen that by employing cooperation, a node can receive multiple copies of the data from different paths. Although a packet might be lost in the direct link between the source and a node, it can still be successfully delivered through other paths with good channel conditions. Therefore, the reliability of wireless multicast can be greatly improved.

2.1.2 Cooperative Multicast Strategies

In the literature, a number of cooperative multicast strategies have been proposed to improve the network performance, which can be roughly classified into the following categories.

- Time division based cooperation (TDBC) [23, 24]: When multiple relays are selected, the second phase of the cooperative multicast process is divided into several subslots as shown in Fig. 2.2, and the selected relays take turns to re-

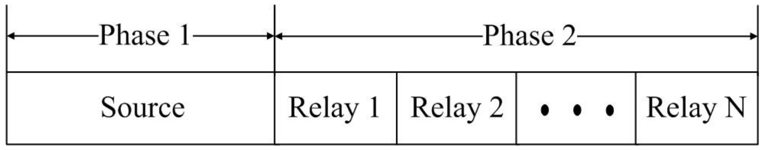

Fig. 2.2. Time division based cooperation.

broadcast the data, each utilizing a designated subslot. This kind of strategy does not require complex coding schemes and is easy to deploy. Moreover, the interference among relays is avoided due to the orthogonal transmission in different subslots. However, when the number of relays increases, the time duration for each subslot becomes smaller and the corresponding transmission rate becomes higher, which may decrease the success probability of each transmission and limit the performance.

- Space-time coding based cooperation (STCBC) [20, 21, 25]: In the second phase of cooperative multicast, instead of transmitting data in different time slots, the relays can employ space time codes to re-encode the received data packets and forward them simultaneously. For instance, in [25], a STCBC strategy is proposed for video multicast, where the relays employ randomized distributed space-time codes (R-DSTC) to re-encode the data and transmit them simultaneously. The whole duration of the cooperation phase is utilized for each relay and the system performance is improved, with the cost of computational complexity. In this strategy, the instantaneous channel state information between the relays and a receiving node is needed to decode the rebroadcasted data. Therefore, when the number of relay becomes larger, the overhead for channel estimation increases as well, which brings more complexity in deploying the cooperative strategy.

Phase 1		Phase 2
Source transmits X_1	Source transmits X_2	Relay transmits $X_1 \oplus X_2$

Fig. 2.3. Network coding based cooperation.

- Network coding based cooperation (NCBC) [17–19]: Network coding has the capability of increasing network throughput by exploiting the shared nature of wireless medium [26]. In the NCBC strategies proposed in [19], cooperation among nodes is triggered after a batch of packets have been sent from the source. An example is shown in Fig. 2.3. In this scheme, the source transmits two different symbols in the two slots in phase 1, then the relays may forward the combined symbol (which is generated using network coding) in phase 2 under ceratin scenarios. The nodes who already have one of the symbols can successfully decode the missing one from the combined symbol. It can be seen that this scheme saves almost $1/4$ time compared with the conventional TDBC scheme, and hence is more efficient. However, if both symbols are lost at a node, they cannot be recovered from the combined symbol. Therefore, the packet loss information is needed to schedule cooperation when deploying the NCBC strategy.

2.1.3 Optimal Scheduling in Cooperative Multicast

Scheduling of cooperation plays an important role in designing cooperative multicast strategies. Generally, it contains relay selection and time allocation between source transmission and relay cooperation. Recently, efforts have been made towards finding the optimal scheduling schemes for cooperative multicast under different performance metrics. Some of the research directions related to optimal scheduling are listed as follows.

- Optimal scheduling for TDBC: In TDBC, a relay's transmission time is directly related to the number of relays. On one hand, a larger relay number results in a shorter transmission time and a higher transmission rate, which decreases the success probability of each relay transmission. On the other hand, more relays can provide higher spatial diversity that may increase the chance of receiving successfully. Therefore, to maximize the system throughput, the optimal number of relays can be determined by solving an optimization problem. In some scenarios, the relays are selected based on topology information. For instance, when the receiving nodes are distributed within a large area that beyond the source's transmission range [23], the optimal relaying scheme considering path loss effect is to select the minimum number of relays so that their transmission ranges cover all the receiving nodes. The optimal time allocation between source multicast and relay cooperation can be found via numerical methods when the optimal relay selection is determined.

- Optimal scheduling under total power constraint: When the size of a network increases, the power consumption in wireless multicast will also increase. Therefore, some scheduling schemes are designed to find the optimal relay selection and power allocation under a total power constraint. Typically, numerical methods can be used in searching the optimal relay transmission order and corresponding power allocation according to the channel state information among the receiving nodes. It has been shown that such optimization problem can be solved within polynomial time [21, 27]. In the literature, power efficient scheduling in dense wireless networks has been studied and low complexity distributed power allocation protocols have been proposed [22].

- Optimal scheduling for maximum lifetime: In wireless sensor networks, lifetime is the critical concern when deploying multicast strategies. The network lifetime is usually defined as the time until the first node runs out of battery. Then the optimal scheduling scheme should be energy efficient such that every node lasts for a period as long as possible. A simple optimal solution–Maximum Lifetime Accumulative Broadcast (MLAB) algorithm has been proposed in [28] that specifies the optimal relay transmission order and corresponding power level. This scheme ensures that the lifetimes of relay nodes are the same and their batteries die simultaneously, and therefore the lifetime of the network is maximized.

This thesis focuses on studying the TDBC strategy and aims to find the optimal relay selection and time allocation schemes that maximize the system throughput, which is discussed in Chapter 3.

2.2 Incentive Mechanisms for Cooperation

In many applications, wireless users are selfish in nature, and incentive mechanisms are needed in order to achieve cooperation. Cooperation stimulation in wireless ad hoc networks and cognitive radio networks has been extensively studied in recent years. Many incentive mechanisms have been proposed, which can be briefly classified into three categories: Payment based mechanisms [29–32], reputation based mechanisms [7, 33–38] and punishment based mechanisms [39–41].

2.2.1 Payment Based Mechanism

Payment based mechanisms introduce virtual currency (or credit) as payment for receiving cooperative services. The main idea is that nodes who get services should be charged while nodes who help others should be remunerated. For instance, in [29], a type of virtual currency called “nuglets” is introduced. When a node wants others to help deliver data packets, it will load a number of nuglets into the packets and those who help forward the data will deduct a certain amount of nuglets. Another example is the credit system “Sprite” presented in [30]. In this system, a node keeps a receipt after forwarding data for others and reports the receipt to a credit clearance service (CCS) in order to get some payment. The CCS will charge the sender a certain amount of credits and pay them to the cooperative nodes. It has been shown that the above mechanisms can stimulate cooperation efficiently in mobile ad-hoc networks. However, this kind of mechanism either requires temper-proof hardware for each node to guarantee the deduction of the virtual currency or needs some central banking service to coordinate the credit exchange, which greatly limits its applications.

2.2.2 Reputation Based Mechanism

An alternative way to stimulate cooperation is to use reputation based mechanisms. In this kind of mechanism, nodes monitor each other’s behaviors and cooperate with those who maintain good reputations. For instance, in [33], a reputation based mechanism is proposed to stimulate cooperation, where each node launches a “watchdog” to monitor its neighbors’ transmissions. If a node does not forward others’ traffic, it will be identified as a misbehavior node and this reputation information is distributed throughout the network. Then, each node employs a “pathrater” to select the routes that avoid the misbehavior nodes. Following the same idea, some protocols such as “CORE” [34] and “CONFIDENT” [35] are proposed to enforce cooperation, which have been shown can detect and isolate the misbehavior nodes efficiently. The major challenge of designing reputation based mechanism for wireless networks is that the monitoring process may not be perfect due to noise and fading, where the obtained information may be inaccurate and undesired decisions may damage the cooperation state.

2.2.3 Punishment Based Mechanism

Another possibility to provide incentive is to use the punishment based mechanism, which is suitable for the scenarios where nodes compete for limited resources (e.g. bandwidth, power, etc.). The main idea is that nodes employ some punishment strategy to punish those who do not cooperate. In [40], a "punish-and-forgive" strategy is proposed to motivate cooperative spectrum sharing in cognitive radio networks. In this strategy, whenever a node deviates from cooperation, a punishment period will be triggered. For example, when a node occupies the whole spectrum to deliver its own data, the channel will be jammed. Nodes who sense the channel are aware of this misbehavior and they will perform non-cooperatively for a certain period, in which no one can get a good service. When the punishment period is long enough, nodes will forgive the misbehavior and resume cooperation again. It can be seen that by introducing punishment, selfish nodes do not have intention to deviate and cooperation can be enforced. However, this kind of mechanism also depends on the accuracy of monitoring or detecting the misbehavior. Although nodes have incentive to cooperate by following such strategies, the random error in wireless channels may result in false alarm that leads the cooperation into punishment period, which degrades the network performance.

2.3 Game Theoretic Approaches

Game theory is a branch of applied mathematics that models and analyzes the interactions among individual decision makers who have potentially conflicting interests [42]. It has been widely used in economics, political science, biology and engineering areas to study cooperation and competition. Recently, game theoretic approaches have been applied in studying cooperation stimulation problems in wireless networks [32, 37–41, 43–45]. The cooperative scenarios such as packet forwarding and spectrum sensing are modeled as games, and incentive strategies are designed and analyzed according to game theory. In the following, some fundamental concepts and models in game theory are introduced, and game theoretic approaches applied in wireless networks are reviewed.

2.3.1 Individual rationality

One of the fundamental assumptions in game theory is that decision makers are rational, which means they make decisions in order to achieve some objectives [46]. The rationality in game theory can be explained from two aspects. First, a rational individual makes decisions consistently. In a game, each player has a personal preference or an objective, which is referred as payoff (or utility). Once the payoff is defined, the player makes decisions to maximize his/her payoff. Second, a rational individual is aware of others' rationality when making a decision. This means a player assumes that others are as smart as he/she is, and makes decisions considering others' potential reactions. In wireless networks, decision makers are either mobile users or network operators who control the wireless devices, and they can program the devices and let them follow certain decision rules. Therefore, in many wireless applications, the rationality assumption holds and game theory can be applied to model and analyze the interactions among wireless users.

2.3.2 Static games

Typically, a game consists of the following three elements: a set of rational players, a set of actions (or strategies) available to each player, and a payoff function for each player that specifies the outcome for a combination of actions. If the players only have one move as a strategy, the game is called *Static Game* or *Single-Stage*

Game. Mathematically, we can represent the static game as $\mathcal{G} = \langle \mathcal{N}, \mathcal{A}, \mathcal{V} \rangle$, where $\mathcal{N} = \{1, 2, ..., N\}$ is the set of players, $\mathcal{A} = \{\mathcal{A}_1, \mathcal{A}_2, ..., \mathcal{A}_N\}$ is the joint set of action spaces, and $\mathcal{V} = \{v_1, v_2, ..., v_N\}$ is the set of payoff functions. When the game is played, player i chooses an action A_i from its action spaces $\mathcal{A}_i$, and all players' actions constitute an action profile $\boldsymbol{A}$. After the actions are carried out, player i gets a payoff $v_i(\boldsymbol{A})$.

In such a static game, a player tends to maximize his/her own payoff, which depends on the whole action profile. Therefore, in order to find the optimal solution, it is important to analyze the outcome of the game with different action profiles. In game theory, Nash equilibrium (NE) is an important concept to measure the outcome of the game, which characterizes the steady state that no player has incentive to change his/her action [42]. Denote $\boldsymbol{A} = \{A_i, A_{-i}\}$, where A_{-i} is the action profile of players except i. Then NE of the static game can be defined as follows.

Definition 1. *The action profile $\boldsymbol{A}^*$ constitutes a Nash equilibrium if , for each player i,*

$$v_i(A_i^*, A_{-i}^*) \geq v_i(A_i, A_{-i}^*), \ \forall \ A_i \in \mathcal{A}_i. \tag{2.1}$$

At NE state, none of the player can increase his/her payoff by unilaterally changing to other actions. Therefore, in a static game, players will take equilibrium actions as the optimal decision.

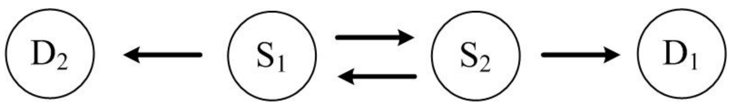

Fig. 2.4. Two-node packet forwarding process.

TABLE 2.1

Two-node packet forwarding game.

	node 2 cooperates	node 2 defects
node 1 cooperates	$(r_1 - c_1, r_2 - c_2)$	$(-c_1, r_2)$
node 1 defects	$(r_1, -c_2)$	(0, 0)

Many scenarios in wireless communication can be modeled as static games. For instance, Fig. 2.4 shows the two-node cooperative packet forwarding process, where a source node's packet needs to be delivered via the other one. S_1 and S_2 represent the two source nodes, and D_1 and D_2 are their destinations, respectively. This process can be modeled as a static game if each source node only transmits one packet. In this game, each source node can be viewed as a player, and their actions are either "cooperate" or "defect". The payoff of a player can be defined as the reward for successfully delivering its own packet minus the cost to help the other source node forward a packet. The game can be expressed in matrix form as shown in Table 2.1. The elements inside the matrix represent the payoff profile for each combination of actions, where r_1 and r_2 represent the reward and c_1 and c_2 represent the cost. In this static game, it is easy to check that the only NE is "defect" for both nodes. Therefore, mutual cooperation cannot be achieved if the cooperative forwarding game is played only once.

2.3.3 Repeated games

In many situations, players interact several times, and the static game is played repeatedly. This type of game is referred as the repeated game, and each interaction is called a stage. Unlike static games, repeated games have the following features.

- First, a player can observe others' past behaviors and hence can condition his/her action in the future. The past moves of all players are referred as the history of the game. In a repeated game, at the beginning of stage k, player i makes a decision A_i^k based on a period of history according to some strategy s_i. Note that there is no history in the first stage. Therefore, the strategy should define the initial action.

- Second, instead of maximizing the payoff in one stage, players make decisions in order to maximize their long term payoffs. Specifically, if players do not know the end of the game, the game can be viewed as an infinite repeated game, and the long term payoff of player i can be represented as $V_i = \sum_{t=0}^{+\infty}(\delta)^t v_i^t$, where $\delta \in (0, 1)$ is a common discount factor that characterizes how much the player cares about the future payoffs.

- Moreover, since the action is determined by the history and behavior strategy, mutual cooperation becomes possible. For instance, in the two-node packet forwarding scenario, if both nodes keep transmitting packets and they adopt the same strategy where the decision is to cooperate as long as the other one cooperates, then cooperation state can be maintained.

In repeated games, NE can be analyzed according to the behavior strategy.

Definition 2. *A strategy profile $\mathcal{S}^* = \{s_i^*, s_{-i}^*\}$ constitutes an NE if, for each player i,*

$$V_i(s_i^*, s_{-i}^*) \geq V_i(s_i', s_{-i}^*), \; \forall \; s_i' \neq s_i^*. \tag{2.2}$$

If a strategy employed by all players constitutes an NE starting from any stage of the game, then it is called a subgame perfect equilibrium strategy [47]. Obviously, when designing incentive mechanisms, a subgame perfect equilibrium strategy is preferred, since by employing such strategy, no one has incentive to deviate at any time. In the literature, *One-Shot Deviation Principle* is proposed to study the subgame perfection of a strategy, which states that a strategy is subgame perfect if no player can increase his/her profit by deviating to another strategy for one stage and then come back to follow the original strategy again in the rest stages.

Another aspect of the equilibrium analysis is what payoffs can be achieved at the NE point. The Folk theorem is one of the most important principles to study the possible NE points in repeated games. Before introducing the theorem, the following terms are defined.

Definition 3. *The payoff profile $\mathcal{V}^f$ is feasible if, for each player i, the corresponding action $A_i^f \in \mathcal{A}_i$.*

Definition 4. *The payoff profile $\mathcal{V}^e$ is enforceable if, for each player i,*

$$v_i(A_i^e, A_{-i}^e) \geq \min_{A_{-i} \in \mathcal{A}_{-i}} \max_{A_i \in \mathcal{A}_i} v_i(A_i, A_{-i}). \tag{2.3}$$

It can be seen that *feasible* means the payoff profile can be realized by some actions, and *enforceable* means a player's payoff is no less than his/her worst payoff when he/she plays the best response against the others. Based on the above definition, the Folk theorem can be described as follows.

Theorem 1. *The Folk Theorem : For any feasible and enforceable payoff profile $\mathcal{V}$, there exists a discount factor $\underline{\delta} < 1$, such that for all $\delta \in (\underline{\delta}, 1)$, there is a Nash equilibrium with payoff profile $\mathcal{V}$.*

The Folk theorem implies that there might be several NE points in a repeated game, and enforcing cooperation is possible.

In some games, players may have different cooperation actions and there might be different cooperation states. For instance, in the two-node packet forwarding game, nodes can forward the packet with different power. Then a question arises: which cooperation state should be enforced as an NE when designing the incentive mechanism? One method in selecting the desired NE is to compare the corresponding action profile according to the concept of Pareto optimality [42], which is defined as follows.

Definition 5. *The action profile $\mathbf{A}^{po}$ is Pareto optimal if, for each player i,*

$$v_i(A_i^{po}, A_{-i}^{po}) \geq v_i(A_i, A_{-i}), \; \forall \; A_i \in \mathcal{A}_i. \tag{2.4}$$

From the above definition, it can be seen that Pareto optimality is a multi-objective optimization criterion considering each node's payoff. An alternative explanation of Pareto optimality is that no one can increase his/her own payoff without decreasing any other node's payoff, which can be proved according to the above definition. Another important criterion in equilibrium selection is the absolute fairness criterion, which requires that each player obtains the same payoff. This is natural in strategy design for wireless networks since users should be treated fairly in many applications.

2.3.4 Game Theoretic Approaches for Cooperation Stimulation

Game theoretic approaches have been incorporated in designing and analyzing cooperation stimulation strategies for wireless networks [32, 38, 39, 44]. Generally, such approaches consists of the following procedures: game modeling, equilibrium analysis and strategy design.

- Game modeling: In many cooperative communication scenarios, the interactions among wireless nodes can be modeled as repeated games. Each wireless node can be viewed as a player, and it takes actions to maximize a certain payoff. The action of a node can be defined as cooperate or not, or cooperate at a certain level. The corresponding payoff is usually defined as the reward from others minus the cost incurred by offering help. After the game is formulated, game theory can be applied to analyze the outcome of nodes' interactions, and incentive mechanisms can be designed accordingly.

- Equilibrium analysis: Before designing incentive mechanisms, it is important to know whether cooperation can be enforced, which is equivalent to analyze the NE of the repeated game. According to the Folk theorem, sometimes there might be multiple NE points. To stimulate cooperation efficiently, a proper NE should be selected according to some criteria, such as the Pareto optimality and the absolute fairness criteria.

- Strategy design: In repeated games, to provide incentive is to design an equilibrium strategy such that nodes will benefit by following the strategy rather than deviating from it. In the literature, there are several strategies that can enforce cooperation, such as the *Tit-for-Tat* and *Grim Trigger*. In *Tit-for-Tat* strategy, a player imitates the others' past behaviors. For instance, in the two-node packet forwarding game, a node who follows the *Tit-for-Tat* strategy will act exactly the same as the other node did in the previous stage. The *Grim Trigger* strategy introduces more severe punishment, where a node cooperates as long as the other one cooperates and defects forever once the other defects. It has been shown that both strategies are subgame perfect equilibrium strategies. When designing the incentive mechanisms, these strategies can be modified to fit the specific application, and sometimes can be applied together to form a more efficient strategy.

The above game theoretic approaches have been widely used for cooperation stimulation in ad-hoc networks, and are also incorporated in this thesis to study the incentive mechanism design in cooperative multicast networks.

Chapter 3

A Time Division Based Cooperative Multicast Strategy

Wireless multicast is a bandwidth efficient communication technique that allows a source to deliver the same information to multiple destinations simultaneously. However, the existence of noise and fading in wireless channels poses great challenge in providing reliable multicast transmissions. Recently, cooperative multicast has emerged as an efficient technique to combat fading in wireless channels, which exploits node cooperation for information relaying. Although different cooperation strategies have been developed to improve the network performance, the optimal scheduling problem, such as relay selection and time allocation, has not been well addressed for delay-sensitive applications.

This chapter studies node cooperation in wireless multicast networks for delay-sensitive applications, and proposes a time division based cooperative multicast strategy. The major contribution is two-fold. First, the optimal relaying strategy that maximizes the system throughput is shown to be a pure threshold policy, where the optimal relay number is bounded by a threshold. Second, to avoid the complexity in exhaustive search, a simple algorithm is proposed to find the optimal time allocation

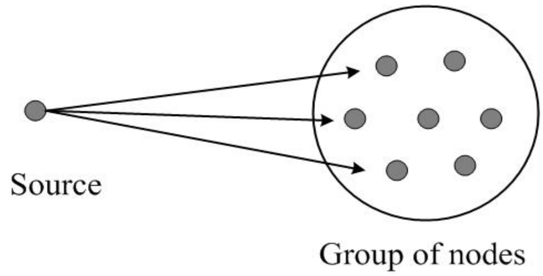

Fig. 3.1. System model.

between source multicast and relay cooperation.

The rest of this chapter is organized as follows. Section 3.1 describes the system model. Section 3.2 gives a detail discussion of the optimum relay selection scheme. Section 3.3 presents a simplified numerical method to find the optimal time allocation. Section 3.4 provides performance analysis of the proposed strategy. Section 3.5 gives some further discussion, followed by conclusions in Section 3.6.

3.1 System model

3.1.1 Network Description and Channel Model

The considered model is a single-hop wireless multicast network as shown in Fig. 3.1, where a source multicasts data packets to a group of N nodes that are close to each other. The distance between the source and each node is much larger than the distance between any pair of nodes, so that the average channel condition between any two nodes is better than that of the direct link from the source to a node. The live multicast applications are considered, which have stringent delay constraint and require that a packet of length L_p must be delivered to all nodes within time duration T. All the wireless channels are assumed to undergo path loss and Rayleigh fading.

The channel gain between the source and any node is assumed to be independent and identically distributed (i.i.d.). Denote x_s as the symbols broadcasted by the source, and the received symbols at node i, denoted as y_{si}, can be represented as

$$y_{si} = \sqrt{P_s} h_{si} x_s + n_{si}, \tag{3.1}$$

where P_s is the source's transmission power, h_{si} is the channel gain from the source to node i, which can be modeled as a complex Gaussian random variable with $\mathcal{CN}(0, \sigma_s^2)$, and n_{si} is the additive white Gaussian noise (AWGN) with $\mathcal{CN}(0, N_0)$. According to Shannon theory, the corresponding channel capacity is given by

$$C_{si} = \log\left(1 + \frac{P_s |h_{si}|^2}{N_0}\right). \tag{3.2}$$

The block-fading channel model is considered, where the channel gain remains unchanged within a time slot (with length T) and is i.i.d. in different time slots. It is also assumed that a data packet can be successfully delivered if the source's transmission rate γ_s is no larger than the corresponding channel capacity, and the packet delivery success probability from the source is given by

$$p_s(\gamma_s, P_s) = \Pr\{\gamma_s \leq C_{si}\} = \exp\left(\frac{N_0(1 - 2^{\gamma_s})}{P_s \sigma_s^2}\right), \tag{3.3}$$

The channel gain between any pair of nodes is assumed to be i.i.d. with $\mathcal{CN}(0, \sigma_r^2)$. The corresponding packet delivery success probability from node i to node j can be similarly derived as $p_r(\gamma_i, P_i) = \exp\left(\frac{N_0(1-2^{\gamma_i})}{P_i \sigma_r^2}\right)$, where γ_i is node i's transmission rate and P_i is the transmission power.

3.1.2 Time Division Based Cooperative Multicast Strategy

Fig. 3.2 shows the considered time division based cooperative multicast strategy. It contains two phases that have time duration $T_1 = (1 - \alpha)T$ and $T_2 = \alpha T$, respectively, where $0 \leq \alpha < 1$. In the first phase, the source transmits the packet at rate

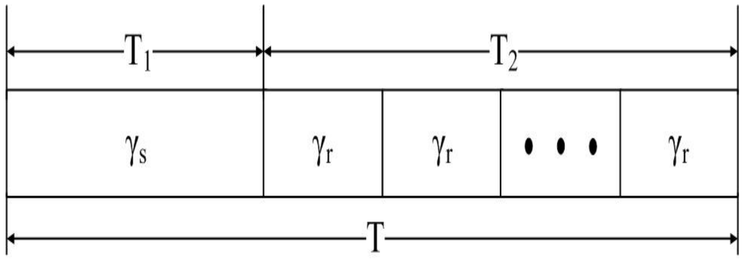

Fig. 3.2. Time division based cooperative multicast strategy.

$\gamma_s = L_p/T_1$ using power P_s. According to the channel model, the probability that a node successfully receives the packet from the source is $p_s = \exp\left(\frac{(1-2^{\gamma_s})N_0}{P_s\sigma_s^2}\right) = \exp\left(\frac{1-2^{\gamma_s}}{\Gamma_1}\right)$, where $\Gamma_1 = \frac{P_s\sigma_s^2}{N_0}$. At the end of Phase 1, the successful nodes send feedbacks to the source indicating their success in receiving the packet, and the feedback time is assumed to be negligible.

In Phase 2, nodes who successfully receive the packet from the source help forward the packet to others. Given that i nodes receive the packet correctly in the first phase, since all channel gains between nodes are i.i.d., the source randomly selects $m_i \leq i$ of them to serve as relays, and the selected relays take turns to broadcast the packet in Phase 2. Note that the number of selected relay, m_i, depends on the number of successful nodes in Phase 1. The time duration T_2 in Phase 2 is divided into m_i intervals of equal length, each for a relay who transmits the packet at the same rate $\gamma_r = \frac{L_p}{T_2/m_i}$ using the same power P_r. The corresponding probability that a node successfully receives from a relay is $p_r = \exp\left(\frac{(1-2^{\gamma_r})N_0}{P_r\sigma_r^2}\right) = \exp\left(\frac{1-2^{\gamma_r}}{\Gamma_2}\right)$, where $\Gamma_2 = \frac{P_r\sigma_r^2}{N_0}$. Note that, when all nodes correctly receive the packet from the source, there is no need for nodes to forward the packet in Phase 2 and $m_N = 0$. For a node who fails in Phase 1, it still can succeed in Phase 2 if it correctly receives the packet from at least one of the m_i relays, which happens with probability $\big(1-(1-p_r)^{m_i}\big)$.

Let μ_1 be the number of nodes that successfully receive the packet in Phase 1, and μ_2 be the number of nodes who fail in phase 1 but successfully receive the packet in Phase 2. Given $\mu_1 = i$ and m_i relays help forward the packet, μ_2's conditional mean is $\mathbb{E}[\mu_2|i, m_i] = (N-i)\big(1-(1-p_r)^{m_i}\big)$. Define $\mathbf{m} \stackrel{\triangle}{=} (m_1, m_2, \cdots, m_N)$, and let $\mu(\alpha, \mathbf{m})$ be the total number of nodes who successfully receive the packet within T (either from the source or from one of the relays). Following the above analysis, the system throughput is defined as the mean of $\mu(\alpha, \mathbf{m})$, which can be calculated using

$$\begin{aligned}
& \mathbb{E}[\mu(\alpha, \mathbf{m})] \\
= & \sum_{i=1}^{N} \left[i + (N-i) \cdot \big(1 - (1-p_r)^{m_i}\big)\right] \Pr\{\mu_1 = i\} \\
= & \sum_{i=1}^{N} \left\{ \left[i + (N-i) \cdot \big(1 - (1-p_r)^{m_i}\big)\right] \cdot \binom{N}{i} p_s^i (1-p_s)^{N-i} \right\}. \quad (3.4)
\end{aligned}$$

3.2 Optimal Relay Selection Scheme

This section studies the optimal relay selection scheme for the cooperative multicast strategy. Specifically, the time allocation parameter α is fixed and the optimal relay number $\mathbf{m}$ that maximizes the system throughput is derived. From (3.4), for a particular number of successful nodes in phase 1, $\mu_1 = i$, the relay selection parameter m_i only affects the term $\big(1 - (1-p_r)^{m_i}\big)$. Therefore, for a fixed α, maximization of the throughput $\mathbb{E}[\mu(\alpha, \mathbf{m})]$ is equivalent to maximization of $\big(1 - (1-p_r)^{m_i}\big)$, or equivalently, minimization of $(1-p_r)^{m_i}$ for all $1 \leq i \leq N-1$. Define $\omega \stackrel{\triangle}{=} 1/\Gamma_2$ and $\rho \stackrel{\triangle}{=} \gamma_r/m_i = L_p/(\alpha T)$, and $p(\alpha, m_i) \stackrel{\triangle}{=} \exp\left(\omega(1-2^{\rho m_i})\right)$. Thus, $p_r = \exp\left(\frac{1-2^{\gamma_r}}{\Gamma_2}\right) = p(\alpha, m_i)$, and the optimal relay selection problem becomes the following optimization problem: for $1 \leq i \leq N-1$,

$$\text{minimize}_{m_i} \quad H(\alpha, m_i) \stackrel{\triangle}{=} \big(1 - p(\alpha, m_i)\big)^{m_i} \quad \text{subject to} \quad 0 \leq m_i \leq i. \quad (3.5)$$

Note that in (3.5), m_i is an integer. To gain some insights into the optimal relay selection problem, it is important to analyze the properties of $H(\alpha, x) \overset{\triangle}{=} \big(1 - p(\alpha, x)\big)^x$, where $p(\alpha, x) \overset{\triangle}{=} \exp\big(\omega(1 - 2^{\rho x})\big)$ and $x \in \mathbb{R}^+$ is a positive real number. The following result can be obtained.

Lemma 1. *For a fixed* α, $H(\alpha, x)$ *achieves its minimum value at* $x = x_0$ *where* $\frac{\partial H(\alpha,x)}{\partial x}\Big|_{x=x_0} = 0$. *In addition,* $H(\alpha, x)$ *is a monotonically decreasing function of* x *when* $0 < x \le x_0$, *and it monotonically increases when* $x > x_0$.

Proof. Take the first derivative of $\ln H(\alpha, x) = \ln\left\{\big(1 - p(\alpha, x)\big)^x\right\}$ with respect to x, it can be obtained that

$$
\begin{aligned}
\frac{\partial \ln H(\alpha, x)}{\partial x} &= \frac{1}{H(\alpha, x)} \frac{\partial H(\alpha, x)}{\partial x} \\
&= \ln\big(1 - p(\alpha, x)\big) + \frac{(-1) \cdot x}{1 - p(\alpha, x)} \frac{\partial p(\alpha, x)}{\partial x}
\end{aligned}
$$

where

$$
\begin{aligned}
\frac{\partial p(\alpha, x)}{\partial x} &= \exp\big(\omega(1 - 2^{\rho x})\big) \cdot \omega \cdot (-1) \cdot \ln 2 \cdot \rho \cdot 2^{\rho x} \\
&= -\rho\omega(\ln 2) p(\alpha, x) 2^{\rho x}.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\frac{\partial H(\alpha, x)}{\partial x} &= H(\alpha, x)\left\{\ln(1 - p(\alpha, x)) + \frac{(-1) \cdot x}{1 - p(\alpha, x)} \frac{\partial p(\alpha, x)}{\partial x}\right\} \\
&= \frac{H(\alpha, x) p(\alpha, x)}{1 - p(\alpha, x)} \left\{\frac{1 - p(\alpha, x)}{p(\alpha, x)} \ln\big(1 - p(\alpha, x)\big) + \rho\omega(\ln 2) x 2^{\rho x}\right\} \\
&= \frac{H(\alpha, x) p(\alpha, x)}{1 - p(\alpha, x)} f(\alpha, x) \qquad (3.6)
\end{aligned}
$$

where

$$
f(\alpha, x) \overset{\triangle}{=} \frac{1 - p(\alpha, x)}{p(\alpha, x)} \ln\big(1 - p(\alpha, x)\big) + \rho\omega(\ln 2) x 2^{\rho x}. \qquad (3.7)
$$

Appendix A shows that $f(\alpha, x)$ has the following properties:

i)

$$
\lim_{x \to 0} f(\alpha, x) = 0, \qquad \lim_{x \to +\infty} f(\alpha, x) = +\infty. \qquad (3.8)
$$

ii) Define x^* as the root of $\frac{\partial f(\alpha,x)}{\partial x} = 0$, then

$$\begin{cases} \frac{\partial f(\alpha,x)}{\partial x} < 0 & \text{if } 0 < x < x^*; \\ \frac{\partial f(\alpha,x)}{\partial x} = 0 & \text{if } x = x^*; \\ \frac{\partial f(\alpha,x)}{\partial x} > 0 & \text{if } x > x^*. \end{cases} \tag{3.9}$$

From the two properties it can be seen that $f(\alpha, x) = 0$ has a single root $x_0 \in (0, +\infty)$ and

$$\begin{cases} f(\alpha,x) < 0 & \text{if } 0 < x < x_0; \\ f(\alpha,x) = 0 & \text{if } x = x_0; \\ f(\alpha,x) > 0 & \text{if } x > x_0. \end{cases} \tag{3.10}$$

For $x \in (0, +\infty)$, it is obvious that $p(\alpha, x) \in (0, 1)$, $H(\alpha, x) \in (0, 1)$, and thus, $\frac{H(\alpha,x)p(\alpha,x)}{1-p(\alpha,x)} > 0$. Based on (3.6), (3.8), and (3.10), it can be obtained that

$$\begin{cases} \frac{\partial H(\alpha,x)}{\partial x} < 0 & \text{if } 0 < x < x_0; \\ \frac{\partial H(\alpha,x)}{\partial x} = 0 & \text{if } x = x_0; \\ \frac{\partial H(\alpha,x)}{\partial x} > 0 & \text{if } x > x_0. \end{cases} \tag{3.11}$$

Therefore, $H(\alpha, x)$ has one global minimum at $x = x_0$, and it is a monotonically decreasing function of x if $x \in (0, x_0)$, and monotonically increases as x increases from x_0. This completes the proof. □

For a given α, let $x_0(\alpha)$ be the root of the equation $\frac{\partial H(\alpha,x)}{\partial x} = 0$, and define

$$m^*(\alpha) = \begin{cases} \lfloor x_0(\alpha) \rfloor & \text{if } H(\alpha, \lfloor x_0(\alpha) \rfloor) \leq H(\alpha, \lceil x_0(\alpha) \rceil); \\ \lceil x_0(\alpha) \rceil & \text{otherwise.} \end{cases} \tag{3.12}$$

Here, $\lfloor \cdot \rfloor$ and $\lceil \cdot \rceil$ are the floor and the ceiling functions, respectively. Based on Lemma 1, for integer m_i, $H(\alpha, m_i)$ is a monotonically decreasing function of m_i when $m_i \leq m^*(\alpha)$, and $H(\alpha, m_i)$ monotonically increases if $m_i > m^*(\alpha)$. Note that the number of relays should be no larger than the number of successful nodes in phase 1. Therefore, based on the property of $H(\alpha, m_i)$, given $\mu_1 = i$ nodes successfully received the packet in phase 1, to minimize $H(\alpha, m_i)$ and to maximize the network throughput, if i is smaller than $m^*(\alpha)$, then all these successful nodes should be selected as relays; otherwise, the source should randomly select only $m^*(\alpha)$ of the i successful nodes to serve as relays in phase 2.

To summarize, the optimal relay number m_i can be found by

$$m_i^{opt} = \begin{cases} \min\big(i, m^*(\alpha)\big) & \text{if } i < N, \\ 0 & \text{if } i = N \end{cases} \tag{3.13}$$

From (3.13), it can be seen that $m^*(\alpha)$ works as a threshold in determining the optimal number of relays and it can be found numerically according to (3.12).

With the optimal relay selection algorithm, the network throughput $\mathbb{E}[\mu(\alpha, \mathbf{m})]$ in (3.4) becomes

$$\begin{aligned} \mathbb{E}[\mu(\alpha)] \;\; = \;\; & \sum_{i=1}^{N} \Bigg\{ \Big[i + (N-i) \cdot \Big(1 - \big(1 - p_r(\alpha)\big)^{m_i^{opt}} \Big) \Big] \\ & \cdot \binom{N}{i} p_s(\alpha)^i \big(1 - p_s(\alpha)\big)^{N-i} \Bigg\}. \end{aligned} \tag{3.14}$$

3.3 Optimal Time Allocation Scheme

In this section, based on the optimal relay selection scheme in the previous discussion, the optimal time allocation is determined. Specifically, a simple numerical method is proposed to jointly determine the time allocation parameter α and the relay number $\mathbf{m}$ that maximize the system throughput. Note that the optimal relay number threshold $m^*(\alpha)$ in (3.12) is a function of α. Therefore, to determine the optimal time allocation, it is necessary to find $m^*(\alpha)$ for all $\alpha \in [0, 1)$, which is computationally expensive.

For any given $\alpha \in [0, 1)$, $m^*(\alpha)$ can be obtained based on (3.12), where it is related to the root of the differential equation $\frac{\partial H(\alpha, x)}{\partial x} = 0$, $x_0(\alpha)$. Instead of solving the above differential equation to obtain $m^*(\alpha)$, an efficient algorithm is proposed to find $m^*(\alpha)$ with less complexity. The following lemmas are introduced to facilitate the algorithm.

Lemma 2. *$x_0(\alpha)$ linearly scales with α. That is, given $\alpha_1 \neq \alpha_2$, let $x_0(\alpha_1)$ and $x_0(\alpha_2)$ be the solutions to $\frac{\partial H(\alpha_1,x)}{\partial x} = 0$ and $\frac{\partial H(\alpha_2,x)}{\partial x} = 0$, respectively. Then, $\frac{x_0(\alpha_1)}{x_0(\alpha_2)} = \frac{\alpha_1}{\alpha_2}$.*

Proof. From the proof of Lemma 1, $\frac{\partial H(\alpha,x)}{\partial x} = 0$ has a single root $x_0(\alpha)$ which is also the single root of equation

$$\frac{1-p\big(\alpha,x\big)}{p\big(\alpha,x\big)} \ln\Big(1-p\big(\alpha,x\big)\Big) + \omega(\ln 2)\frac{L_p}{T}\frac{x}{\alpha}2^{\frac{L_p}{T}\frac{x}{\alpha}} = 0 \tag{3.15}$$

where $p\big(\alpha,x\big) = \exp\left(\omega\left(1-2^{\frac{L_p}{T}\frac{x}{\alpha}}\right)\right)$. Denote $y = \frac{x}{\alpha}$. Thus, finding $x_0(\alpha)$ is equivalent to finding the single root $y = y_0$ of the following equation

$$\frac{1-p(y)}{p(y)} \ln\Big(1-p(y)\Big) + \omega(\ln 2)\frac{L_p y}{T}2^{\frac{L_p y}{T}} = 0 \tag{3.16}$$

where $p(y) = \exp\left(\omega\left(1-2^{\frac{L_p y}{T}}\right)\right)$, and then letting $x_0(\alpha) = \alpha \cdot y_0$. It can be seen that the solution to (3.16), $y_0 = \frac{x_0(\alpha)}{\alpha}$, does not depend on α. Therefore, $x_0(\alpha)$ scales linearly with α. $\square$

Lemma 2 means that, if $x_0(\alpha_0)$ is known for a specific value α_0, then for any other $\alpha \in (0,1)$, the corresponding $x_0(\alpha)$ can be determined as $x_0(\alpha) = x_0(\alpha_0) \cdot \alpha/\alpha_0$.

Lemma 3. *For an integer n, assume $a, b \in (0,1)$ satisfy $x_0(a) = n$ and $x_0(b) = n+1$. If $m^*(\alpha_0) = n+1$ for a specific $\alpha_0 \in (a,b)$, then $m^*(\alpha) = n+1$ for any $\alpha \in [\alpha_0, b]$.*

Proof. To prove Lemma 3, the first step is to prove that for an infinitely small positive value ϵ, $m^*(\alpha_0+\epsilon) = n+1$.

First, note that for $\alpha_0 \in (a,b)$, $m^*(\alpha_0) = n+1$ means that $x_0(\alpha_0) \in (n, n+1]$ and $H(\alpha_0, n) > H(\alpha_0, n+1)$. (When $H(\alpha_0,n) = H(\alpha_0,n+1)$, that is, n and n+1 relays give the same system throughput, $m^*(\alpha) = n$ is selected in order to reduce the total energy consumed by relays.) From Lemma 1, $H(\alpha_0,x)$ decreases when x increases from n to $x_0(\alpha_0)$, and increases when x increases from $x_0(\alpha_0)$

to $n+1$. Since $H(\alpha_0, n) > H(\alpha_0, n+1)$, there exists $z \in (n, x_0(\alpha_0))$ such that $H(\alpha_0, z) = H(\alpha_0, n+1)$.

Taking the first derivative of $H(\alpha, x) = \big(1 - p(\alpha, x)\big)^x$ with respect to α, then

$$\begin{aligned}
\frac{\partial H(\alpha, x)}{\partial \alpha} &= x\big(1 - p(\alpha, x)\big)^{x-1} \cdot (-1) \cdot \frac{\partial p(\alpha, x)}{\partial \alpha} \\
&= \ln 2 \cdot \omega \cdot 2^{\frac{L_p x}{\alpha T}} x^2 \big(1 - p(\alpha, x)\big)^{x-1} p(\alpha, x) \frac{L_p}{T} (-\frac{1}{\alpha^2}) \\
&= \ln 2 \cdot \omega \cdot 2^{\frac{L_p x}{\alpha T}} x^2 \frac{H(\alpha, x) p(\alpha, x)}{1 - p(\alpha, x)} \frac{L_p}{T} \left(-\frac{1}{\alpha^2}\right) < 0 \qquad (3.17)
\end{aligned}$$

which means that $H(\alpha, x)$ decreases as α increases. In addition, as proved in Appendix B, for any given $\alpha \in [0, 1)$, if $H(\alpha, z) = H(\alpha, n+1)$ and $0 < z < x_0(\alpha) < n+1$, then

$$\Big|\frac{\partial H(\alpha, x)}{\partial \alpha}|_{x=z}\Big| < \Big|\frac{\partial H(\alpha, x)}{\partial \alpha}|_{x=n+1}\Big|. \qquad (3.18)$$

Since $H(\alpha_0, z) = H(\alpha_0, n+1)$ and ϵ is an infinitely small positive value, then

$$\begin{aligned}
& H(\alpha_0 + \epsilon, z) - H(\alpha_0 + \epsilon, n+1) \\
= \;& \epsilon \left(\frac{\partial H(\alpha, x)}{\partial \alpha}\Big|_{x=z, \alpha=\alpha_0} - \frac{\partial H(\alpha, x)}{\partial \alpha}\Big|_{x=n+1, \alpha=\alpha_0} \right) + O(\epsilon^2) > 0. \; (3.19)
\end{aligned}$$

Furthermore, $z \in (n, x_0(\alpha))$ and $x_0(\alpha_0 + \epsilon) = x_0(\alpha_0)(\alpha_0 + \epsilon)/\alpha_0 > x_0(\alpha_0)$ implies that $z \in (n, x_0(\alpha_0 + \epsilon))$. Note that from Lemma 1, $H(\alpha_0 + \epsilon, x)$ decreases when x changes from n to $x_0(\alpha_0 + \epsilon)$. Consequently,

$$H(\alpha_0 + \epsilon, n) > H(\alpha_0 + \epsilon, z) \overset{\text{from}(3.19)}{>} H(\alpha_0 + \epsilon, n+1) \qquad (3.20)$$

and therefore, $m^*(\alpha_0 + \epsilon) = n+1$. From $m^*(\alpha_0 + \epsilon) = n+1$, it can be seen that $m^*(\alpha_0 + 2\epsilon) = m^*\big((\alpha_0 + \epsilon) + \epsilon\big) = n+1$, using the same proof as above. Keeping adding ϵ, eventually, it can be concluded that $m^*(\alpha) = n+1$ for any $\alpha \in [\alpha_0, b]$. This completes the proof. □

For interval $[a, b]$ where $x_0(a) = n$ and $x_0(b) = n+1$ (n is an integer), based on the property of $H(\alpha, x)$, it can be obtained that $H(a, n) < H(a, n+1)$ and $H(b, n) > H(b, n+1)$. Therefore, it is obvious that when α increases from a to

b, there is an α_n that satisfies $H(\alpha_n, n) = H(\alpha_n, n+1)$. From Lemma 3, it can be seen that $m^*(\alpha) = n$ if $\alpha \in [a, \alpha_n]$, and $m^*(\alpha) = n+1$ for any $\alpha \in (\alpha_n, b]$. Therefore, the relay threshold $m^*(\alpha)$ increases with α in a stair case manner. Based on this observation, the following simplified algorithm can be used to efficiently find the values of $m^*(\alpha)$ for any $\alpha \in [0, 1)$. An illustration is shown in Fig. 3.3, in which $x_0(\alpha)$ and $m^*(\alpha)$ are represented by the dashed-dotted line and solid line, respectively.

Algorithm 1 Determination of $m^*(\alpha)$ $(0 \leq \alpha < 1)$.

1: let $\alpha = 1$ and solve $\frac{\partial H(\alpha=1,x)}{\partial x} = 0$ to get $x_0(\alpha)|_{\alpha=1}$.

2: draw the figure of $x_0(\alpha) = \alpha \cdot x_0(\alpha)|_{\alpha=1}$.

3: draw lines $x_0(\alpha) = 1, 2, ...,$ which will partition the line of $x_0(\alpha) = \alpha \cdot x_0(\alpha)|_{\alpha=1}$ into K regions with index $0, 1, ..., K-1$. Assume region k is between lines $x_0(\alpha) = k$ and $x_0(\alpha) = k+1$.

4: in region $k = 0$, if $\alpha = 0$, then $m^*(\alpha) = 0$; Otherwise $m^*(\alpha) = 1$.

5: **for** $k = 1 : K-1$ **do**

6: in region k, use numerical methods such as bi-section search to find the point of $\alpha = \alpha_k$ such that $H(\alpha_k, k) = H(\alpha_k, k+1)$.

7: in region k, $m^*(\alpha) = k$ if $\alpha \in [j_k, \alpha_k]$ and $m^*(\alpha) = k+1$ if $\alpha \in (\alpha_k, j_{k+1}]$, where $j_k = \frac{k}{x_0(\alpha)|_{\alpha=1}}$.

8: **end for**

Step 3 of the algorithm is to partition the line $x_0(\alpha)$ into several regions. In region k, at the left end, $x_0(\alpha) = k$, while at the right end, $x_0(\alpha) = k+1$. Therefore, in region k, $m^*(\alpha)$ is either k or $k+1$. From Lemma 3, when α moves from the left end to the right end in region k, once $m^*(\alpha) = k+1$ at one particular value of α, the value of $m^*(\alpha)$ will keep at value $k+1$ when α further moves from the particular value to the right end of the region. Therefore, in Step 6, bi-section search method is used to find the particular value of α in region k. And in the region, $m^*(\alpha) = k$

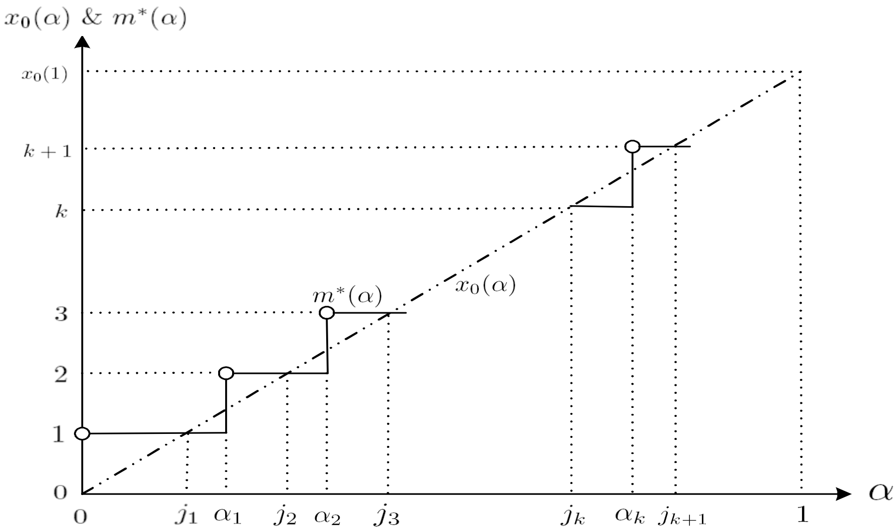

Fig. 3.3. Illustration of Algorithm 1.

and $m^*(\alpha) = k + 1$ at the left and right side of the particular value, respectively, as shown in Step 7.

The major advantage of the above algorithm is that the differential equation $\frac{\partial H(\alpha,x)}{\partial x} = 0$ only need to be solved once, and then all values of $m^*(\alpha)$ can be found by fast search. Note that when the algorithm is not adopted, $\frac{\partial H(\alpha,x)}{\partial x} = 0$ has to be solved for each value of α within $0 \leq \alpha < 1$, which is computationally complex.

Given the above efficient method to determine $m^*(\alpha)$ for all $\alpha \in [0, 1)$, the optimal time allocation can be found as follows. First, α is uniformly sampled in the range of $[0, 1)$ with a step size Δ ($\ll$1). Then $m^*(\alpha)$ can be found for all sampled α values according to Algorithm 1 and the corresponding network throughput $\mathbb{E}[\mu(\alpha)]$ can be calculated using (3.14). Finally the value of α that maximizes $\mathbb{E}[\mu(\alpha)]$ is selected as the optimal time allocation parameter. Note that α=0 corresponds to direct transmission without node cooperation.

3.4 Performance Evaluation

In this section, the performance of proposed cooperative multicast strategy is evaluated and compared with that of the direct multicast scheme without cooperation. Consider a network with $N = 100$ nodes. The packet size L_p is 8 Kbits, which needs to be received within $T = 10$ ms over a wireless channel with 1 MHz bandwidth. Other parameters are $\Gamma_1 = 0$ dB and $\Gamma_2 = 10$ dB. Fig. 3.4 shows $\mathbb{E}[\mu(\alpha)]$, the average number of nodes that successfully receive the packet during T with different values of the time allocation parameter α. It is observed that $\mathbb{E}[\mu(\alpha)]$ decreases with α at first and then increases to a maximum point and then decreases again. This can be explained as follows. When the time duration for relay session (i.e., Phase 2) is very short (α is close to 0), the success probability p_r in phase 2 is very small such that almost no node can receive successfully from a relay. Thus, a larger α means more waste of time, and degrades the performance. As α increases to a certain level, the success probability p_r becomes larger, and cooperation in phase 2 takes effect. So the overall throughput increases. As α further increases (close to 1), the time for source transmission, T_1, is very small (close to 0), which means the success probability p_s is very small and almost no node can receive the packet correctly in phase 1. As a result, the throughput decreases dramatically.

Next the cooperation strategy is compared with direct multicast. The cooperation strategy is designed with the optimal time allocation and optimal relay selection. The system throughput is derived numerically for different values of Γ_1 and Γ_2, where the difference between them is kept at $\Gamma_2 - \Gamma_1 = 10$ dB. Note that Γ_1 and Γ_2 are the average SNR for the channel from the source to the nodes, and between any two nodes, respectively. Fig. 3.5 shows the average number of successful nodes versus Γ_1. It can be observed that when Γ_1 is smaller than -6.4 dB, optimal cooperation strategy is just direct multicast. The reason is as follows. When Γ_1 is very small (which means that the channel from the source to each node is poor), it is not beneficial if the duration T is partitioned into two phases and let the source transmit in the first phase, because almost no node will successfully receive the packet in the first phase due to poor channel quality. As the value of Γ_1 increases beyond a certain level, cooperation achieves a better performance. Particularly, in Γ_1's region from -4 dB to 4 dB, the optimal cooperation strategy achieves significant performance gain, compared to the direct multicast. When Γ_1 further increases (more than 8 dB), both the cooperation strategy and direct multicast can achieve high throughput, and the difference between them becomes smaller and smaller.

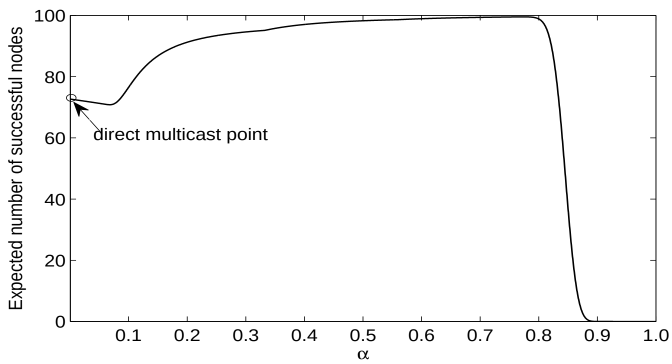

Fig. 3.4. Throughput versus α.

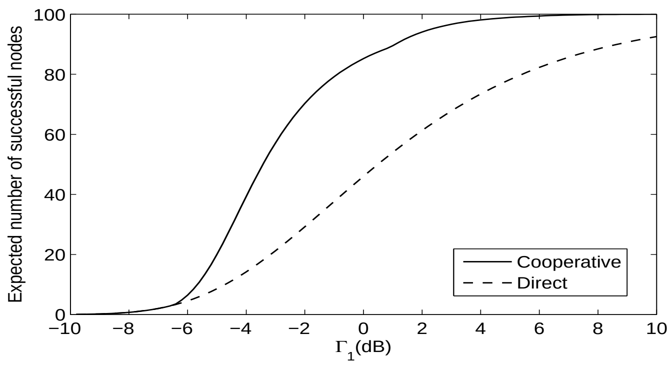

Fig. 3.5. Comparison of cooperative multicast and direct multicast.

3.5 Further Discussion

The analysis in preceding sections is based on the assumption that the channel gain between any pair of nodes is i.i.d., referred to as *i.i.d. case*. It is essential to investigate the impact if the channel gain between nodes is not identically distributed. Consider the following non-i.i.d. case. The nodes are uniformly distributed within a circle with radius d_0. The channel between any pair of nodes undergoes path loss with exponent 2 and Rayleigh fading. In the non-i.i.d. case, the probability p_r in preceding sections is not the same anymore for each pair of nodes, and is expressed as $p_r(\tau_{ij}, \gamma_r) = \exp(\tau_{ij}(1 - 2^{\gamma_r}))$, where $\tau_{ij} = \frac{N_0}{P_r \sigma_{ij}^2}$ depends on the statistics of the channel gain between node i and node j. To deal with the heterogeneity of p_r, an approximation is used, which can be described as follows. For the N nodes uniformly distributed within the circle, the average value of τ_{ij}, denoted as $\overline{\tau}$, can be calculated using numerical methods, and the corresponding value of p_r, denoted $\overline{p_r} = p_r(\overline{\tau}, \gamma_r)$ can be found for a particular value of γ_r. By using $\overline{p_r}$ to approximate the success probability between any pair of nodes, the system throughput can be approximated according to (3.4), by replacing p_r with $\overline{p_r}$. By employing this approximation, the optimal relay selection and optimal time allocation can be derived as well.

Simulations are used to evaluate the effectiveness of the above approximation. Consider $N = 10$ nodes uniformly distributed within a circle with unit radius. The wireless channel bandwidth is 1 MHz. The length of the whole time slot is $T = 20$ ms, and the packet size is L_p = 8 Kbits. The average received SNR for a node from the source is $\Gamma_1 = 0$ dB, and the average received SNR between two nodes is 10 dB if the distance between the two nodes is unit. 1000 topologies are randomly generated based on the node distribution. First, the time allocation parameter α is fixed at 0.5, where the whole time slot is equally partitioned. When $\mu_1(\in \{1, 2, ..., N\})$ nodes successfully receive the packet from source in T_1, by

TABLE 3.1

Optimal number of relays and maximal system throughput versus μ_1

μ_1		1	2	3	4	5	6	7	8	9
Optimal number of relays	Exhaustive search	1	2	2	2	2	2	2	2	2
	Approximation	1	2	2	2	2	2	2	2	2
Maximal system throughput	Exhaustive search	9.37	9.74	9.78	9.81	9.84	9.87	9.90	9.94	9.97
	Approximation	9.36	9.73	9.76	9.80	9.83	9.86	9.90	9.93	9.97

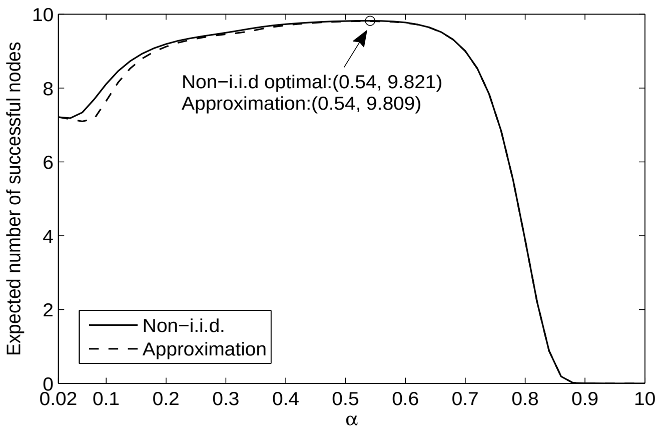

Fig. 3.6. Optimal time allocation for non i.i.d. case.

exhaustive search (on how many relays are selected and which relays are selected) the optimal number of relays can be found, which achieves the maximal average (note that here the average is for the 1000 topologies) system throughput for the 1000 topologies, as shown in Table 3.1. The approximation results are also given in Table 3.1. Then, the optimal time allocation scheme is also compared by changing α, based on the optimal number of relays that determined using exhaustive search. The results are shown in Fig. 3.6. It can be seen that the approximation results match well with the exhaustive search simulation results, which demonstrates the accuracy of the approximation for the non-i.i.d. case.

3.6 Summary

In this chapter, a time division based cooperative multicast strategy is proposed for transmissions of delay-sensitive traffic. The optimum relay selection and time allocation schemes are analyzed. It has been shown that the optimal relay number is bounded by a threshold value which can be found via numerical methods. An efficient algorithm is also proposed to find the optimal time allocation for source transmission and node cooperation. Numerical results show that the proposed strategy can enhance the network performance.

Chapter 4

Cooperation Stimulation

In most existing cooperative wireless multicast strategies, it is assumed that nodes will cooperate unconditionally, and always forward data when selected as relays. However, in many applications, nodes are selfish and prefer receiving data than offering help, since relaying data costs extra energy. Therefore, it is critical to design incentive mechanisms to stimulate cooperation among selfish nodes.

This chapter employs game theoretic approaches to study the cooperation stimulation problem in wireless multicast networks. The cooperation among selfish nodes is modeled as a repeated game and an incentive strategy is proposed according to game theory. The rest of this chapter is organized as follows. Section 4.1 describes the system model and Section 4.2 formulates the repeated game. Equilibrium of the game is studied in Section 4.3. The incentive strategy is proposed and analyzed under both perfect and imperfect monitoring scenarios in Section 4.4. Performance evaluation is provided in Section 4.5 and finally conclusion is drawn in Section 4.6.

4.1 System Model

The general system model is the same as the one considered in Chapter 3, where a source multicasts data packets to a group of N nodes that are close to each other. In this chapter, the mobile scenario is considered where nodes are randomly moving within a circular area. The distance between the source and each node is much larger than the average distance between any pair of nodes. To enhance the system performance, the previously discussed two-phase cooperative multicast strategy is employed, where the source transmits in the first phase, and some successful nodes are selected as relays to forward the data in the second phase. For simplicity, this chapter studies the single relay case, where only one relay is selected to rebroadcast the data in the second phase.

Similar to Chapter 3, the wireless channel between the source and a node experiences Rayleigh fading and AWGN noise (with distribution $\mathcal{CN}(0, N_0)$), and the corresponding channel gain is assumed to be i.i.d. with $\mathcal{CN}(0, \sigma_s^2)$. The probability that a node correctly receives a packet from the source is $p_s = \exp\left(\frac{N_0(1-2^{\gamma_s})}{P_s\sigma_s^2}\right)$, where P_s is the source's transmission power and γ_s is the transmission rate. The channel gain between any two nodes i and j is modeled as $\mathcal{CN}(0, \sigma_{ij}^2)$ and the corresponding probability that node j successfully receives a packet from node i is $p_{ij} = \exp\left(\frac{N_0(1-2^{\gamma_r})}{P_i\sigma_{ij}^2}\right)$, where P_i is node i's transmission power and γ_r is the corresponding transmission rate. Since nodes are mobile, the success probability p_{ij} changes from time to time. To simplify the analysis, the following approximation is used. For any pair of mobile nodes i and j, the average value of $\frac{N_0(2^{\gamma_r}-1)}{\sigma_{ij}^2}$ can be calculated according to the mobility model and the statistics of the channel gain, which is denoted as $\overline{D}$, and the overall probability that node j successfully receives a packet from node i is approximated as $p_{ij} \approx p_r(P_i) = \exp\left(\frac{-\overline{D}}{P_i}\right)$, $\forall\, i, j \in \{1, 2, ...N\}, i \neq j$. The impact of the approximation on system performance will be discussed in Section 4.4.6.

4.2 The Repeated Game Model

In the cooperative multicast strategy discussed in Chapter 3, it is assumed that the selected relay nodes always help forward the data with the same designated power. However, in real applications, nodes may not follow the strategy and can decide the transmission power by themselves. Note that forwarding data benefits the unsuccessful nodes, but incurs some cost to the relay, such as extra energy consumption. Therefore, a selfish node would expect others to cooperate but would not cooperate itself. Consider the scenario where nodes are all selfish and aim to maximize their own profits. Then, the cooperative multicast process can be modeled as a game, where each node is a player. When selected as a relay, a node can decide the transmission power to maximize its payoff, which is defined as the reward for receiving packets (either from the source or from the relays) minus the cost to forward packets. Since the multicast session usually lasts for a long time, nodes may perform the decision-making process repeatedly. Therefore, the interactions between nodes can be formulated as a repeated game. Denote the reward for correctly receiving a packet as $r_0 > 0$ and the cost for transmitting a packet with unit power as $c_0 > 0$, where r_0 and c_0 are application dependent constants. Consider M time slots as a stage, in which nodes are assumed not to change their transmission power. Then the cooperative multicast game can be defined as follows:

Definition 6. *The Cooperative Multicast Game is the game $\mathcal{G}=\langle \mathcal{N}, \boldsymbol{A}^k, \boldsymbol{v}^k \rangle$, where*

- $\mathcal{N}=\{1,2,...,N\}$ *is the set of nodes (players),*
- $\boldsymbol{A}^k = [A_1^k, A_2^k, ..., A_N^k]$. $A_i^k = P_i^k \in [0, P_{\max}]$ *is node i's action in stage k, which means that if node i is selected as a relay, it transmits the packet with power P_i^k. $P_{\max}$ is the maximum power that a node is willing to use for cooperation during a long-period multicast program.*
- $\boldsymbol{v}^k = [v_1^k, v_2^k, ..., v_N^k]$. $v_i^k = r_0(n_{si}^k + n_{ri}^k) - c_0 n_i^k P_i^k$ *is node i's expected payoff*

in stage k, where n_{si}^k is the expected number of packets correctly received from the source, n_{ri}^k is the expected number of packets not correctly received from the source but correctly received from the relays, and n_i^k is the expected number of packets that node i should forward to others.

Consider the scenario where nodes stay in the multicast session for a long time and no one knows exactly when others will leave the multicast service and when the game will end. Then the cooperative multicast process can be viewed as an N-player infinite repeated game. In such a repeated game, at the beginning of stage k, node i makes a decision of A_i^k based on other nodes' past behaviors according to strategy s_i. Assume that any node i is rational, and it makes decisions to maximize its long term expected payoff

$$V_i = \sum_{k=0}^{\infty} (\delta)^k v_i^k, \tag{4.1}$$

where v_i^0 is node i's expected payoff in the current stage, and $\delta \in (0, 1)$ is a common discount factor that characterizes how much nodes care about future payoffs. From the above game model, it can be seen that in order to motivate nodes to work cooperatively, a proper incentive strategy should be designed such that nodes will get more benefit by following the strategy rather than deviating from it.

4.3 Equilibrium Analysis of Cooperative Multicast Game

In the previous section, the cooperative multicast process is modeled as an N-player infinite repeated game. Next, before designing the incentive strategy, it is important to analyze the Nash Equilibria (NE) of the game. NE is a steady state of the game where players do not change their actions (or strategies). Obviously, the incentive strategy should enforce cooperation at an NE from which no one has intention to deviate. According to the Folk theorem [42], any *feasible* and *enforceable* payoff profile of a stage game is possible to be an NE of the corresponding infinite repeated game. In this game model, *feasible* means the expected payoff profile can be realized by taking certain actions (where $A_i^k \in [0, P_{\max}]\ \forall i \in \mathcal{N}$), and *enforceable* means each node's expected payoff should be no worse than that without cooperation. Then the question is which NE is the desired one that should be enforced. Several issues need to be considered when choosing the NE. First, in many applications, nodes should be treated fairly. Second, nodes tend to maximize their expected payoffs. Thus the enforced NE should achieve high payoff for all nodes. Based on the above discussion, the desired NE can be determined according to the following criteria:

(i) *absolute fairness*, which requires that all nodes have the same expected payoff;

(ii) *maximum payoff*, which means the desired NE achieves the maximum payoff profile under the fairness criterion.

According to the above criteria, the desired NE can be found as follows. Denote $\boldsymbol{A} = [P_1, P_2, ..., P_N]$ as the action profile and $\boldsymbol{v}(\boldsymbol{A}) = [v_1(\boldsymbol{A}), v_2(\boldsymbol{A}), ..., v_N(\boldsymbol{A})]$ as the corresponding expected payoff profile. Then finding the desired action profile $\boldsymbol{A}^*$ is equivalent to solving the following optimization problem:

$$\max_{\boldsymbol{A}} v_i(\boldsymbol{A}) \quad \text{subject to} \quad v_i(\boldsymbol{A}) = v_j(\boldsymbol{A})\ \forall\, i, j \in \mathcal{N}, i \neq j. \tag{4.2}$$

Since the channel gain between the source and a node is i.i.d. and the relay is randomly selected, each node has the same chance to be selected as a relay, and that probability is

$$\begin{aligned} q_1 &= \sum_{\mu_1=1}^{N-1} \Pr\{i \text{ is selected}|\mu_1 \text{ nodes succeed in Phase } 1\} \\ &\quad \times \Pr\{\mu_1 \text{ nodes succeed in Phase } 1\} \\ &= \sum_{\mu_1=1}^{N-1} \frac{1}{N}\binom{N}{\mu_1} p_s^{\mu_1}(1-p_s)^{N-\mu_1} = \frac{1}{N}\left(1 - p_s^N - (1-p_s)^N\right). \end{aligned} \tag{4.3}$$

Moreover, the probability that node i does not receive the packet correctly from the source while node j is selected as a relay is

$$\begin{aligned}
q_2 &= \sum_{\mu_1=1}^{N-1} \Pr\{j \text{ is selected}|i \text{ does not succeed and } \mu_1 \text{ nodes succeed in Phase } 1\} \\
&\quad \times \Pr\{i \text{ does not succeed and } \mu_1 \text{ nodes succeed in Phase } 1\} \\
&= \sum_{\mu_1=1}^{N-1} \frac{1}{N-1}\binom{N-1}{\mu_1} p_s^{\mu_1}(1-p_s)^{N-\mu_1} \\
&= \frac{1}{N-1}\left(1-p_s-(1-p_s)^N\right). \qquad (4.4)
\end{aligned}$$

In a stage, the expected number of packets that node i correctly receives from the source is Mp_s, the expected number of packets that node i receives from relay node j when i does not receive them correctly from the source is either $Mq_2 \cdot p_r(P_j)$ (if $P_j > 0$) or 0 (if $P_j = 0$), and the expected number of packets that node i should transmit is Mq_1. In a stage, let $r(x)$ denote the reward that a node gets from another node whose transmission power is x, and $c(x)$ denotes the cost of a node if it uses power x to cooperate, then

$$r(x) = \begin{cases} 0 & \text{if } x = 0; \\ Mq_2 r_0 \cdot p_r(x) & \text{otherwise,} \end{cases} \quad \text{and} \quad c(x) = Mq_1 c_0 \cdot x. \qquad (4.5)$$

Then the expected payoff of node i in one stage is

$$v_i(\boldsymbol{A}) = Mp_s r_0 + \sum_{j\in\mathcal{N}, j\neq i} r(P_j) - c(P_i). \qquad (4.6)$$

Before solving (4.2), the following lemma can be proved.

Lemma 4. $v_1(\boldsymbol{A}) = v_2(\boldsymbol{A}) = ... = v_N(\boldsymbol{A})$ *if and only if* $P_1 = P_2 = ... = P_N$.

Proof. First, if $P_i = P_j\ \forall i, j \in \mathcal{N}$, $i \neq j$, based on (4.6), it is obvious that $v_i(\boldsymbol{A}) = v_j(\boldsymbol{A})$, $\forall i, j \in \mathcal{N}$, $i \neq j$. Next, the following steps will show that given $v_i(\boldsymbol{A}) = v_j(\boldsymbol{A})\ \forall i, j \in \mathcal{N}$, $i \neq j$, then $P_i = P_j$, $\forall i, j \in \mathcal{N}$, $i \neq j$. For any two nodes $i, j \in \mathcal{N}$, according to (4.6), it can be obtained that

$$\begin{aligned}
v_i(\boldsymbol{A}) - v_j(\boldsymbol{A}) &= \left\{\sum_{t\in\mathcal{N}, t\neq i} r(P_t) - c(P_i)\right\} - \left\{\sum_{t\in\mathcal{N}, t\neq j} r(P_t) - c(P_j)\right\} \\
&= \{r(P_j) + c(P_j)\} - \{r(P_i) + c(P_i)\} = h(P_j) - h(P_i) \qquad (4.7)
\end{aligned}$$

where $h(x)\triangleq r(x)+c(x)$. For $x\in(0,+\infty)$, the following can be obtained:

$$\frac{dh(x)}{dx}=\frac{dr(x)}{dx}+\frac{dc(x)}{dx}=Mq_2r_0\left\{\exp\left(-\frac{\overline{D}}{x}\right)\cdot\overline{D}x^{-2}\right\}+Mq_1c_0>0. \quad (4.8)$$

Therefore, $h(x)$ is a monotonically increasing function of x. Since $v_i(\boldsymbol{A})=v_j(\boldsymbol{A})$, it is equivalent to $v_i(\boldsymbol{A})-v_j(\boldsymbol{A})=h(P_j)-h(P_i)=0$. Based on the monotonically increasing property of $h(x)$, it can be concluded that $v_i(\boldsymbol{A})=v_j(\boldsymbol{A})$ if and only if $P_i=P_j$. This completes the proof. □

The above lemma indicates that to achieve the same expected payoff, all nodes' transmission power should be the same. Thus, the expected payoff for node i under the fairness constraint can be represented as:

$$v_i(P_i)=Mp_sr_0+(N-1)r(P_i)-c(P_i)=Mp_sr_0+g(P_i). \quad (4.9)$$

where $g(x)\triangleq(N-1)r(x)-c(x)$. Note that in order to enforce cooperation, the expected payoff of each node at the desired NE should be larger than that without cooperation. That is $v_i(P_i)>Mp_sr_0$, or equivalently, $g(P_i)>0$. Assume there exists a $P_i\in(0,P_{\max}]$ such that $g(P_i)>0$. Then solving (4.2) is equivalent to solving

$$\max_x g(x), \quad \text{subject to} \quad x\in(0,P_{\max}] \text{ and } g(x)>0, \quad (4.10)$$

and the following result can be obtained.

Lemma 5. *If there exists an $x\in(0,P_{\max}]$ such that $g(x)>0$, the following can be obtained:*

$$\operatorname*{arg\,max}_{x\in(0,P_{\max}],g(x)>0} g(x)=\begin{cases}\arg\max_{x>0}g(x) & \textit{if } P_{\max}\geq\arg\max_{x>0}g(x);\\ P_{\max} & \textit{if } P_{\max}<\arg\max_{x>0}g(x).\end{cases} \quad (4.11)$$

Proof. Let $R\triangleq(N-1)Mq_2r_0$, $C\triangleq Mq_1c_0$, then $g(x)=Rp_r(x)-Cx$. Hence,

$$\frac{dg(x)}{dx}=R\left\{\exp\left(-\frac{\overline{D}}{x}\right)\cdot\overline{D}x^{-2}\right\}-C=R\overline{D}\frac{p_r(x)}{x^2}-C. \quad (4.12)$$

and

$$\frac{d^2g(x)}{dx^2} = R\overline{D}\frac{p_r(x)}{x^4}(\overline{D} - 2x). \tag{4.13}$$

From (4.12) and (4.13), it can be seen that $\frac{dg(x)}{dx}$ is monotonically increasing in $(0, \overline{D}/2)$ and monotonically decreasing in $(\overline{D}/2, +\infty)$ with the maximum value achieved at $x = \overline{D}/2$. Moreover, it is easy to prove that $\lim_{x\to 0^+} \frac{dg(x)}{dx} = -C < 0$ and $\lim_{x\to 0^+} g(x) = 0$, which implies that $g(x)$ first decreases and takes negative values when x increases from 0. Based on the above property, if there exists an x such that $g(x) > 0$, then there exists at least one x that satisfies $\frac{dg(x)}{dx} > 0$. Since $\frac{dg(x)}{dx}$ achieves maximum value at $x = \overline{D}/2$, then $\frac{dg(x)}{dx}|_{x=\overline{D}/2} > 0$. Therefore, $\frac{dg(x)}{dx} = 0$ has two roots x_{g1} and $x_{g2}(0 < x_{g1} < x_{g2})$, where $\frac{dg(x)}{dx} > 0$ for $x \in (x_{g1}, x_{g2})$ and $\frac{dg(x)}{dx} < 0$ for $x \in (0, x_{g1}) \cup (x_{g2}, +\infty)$. It implies that $g(x)$ monotonically increases in (x_{g1}, x_{g2}) and decreases in other intervals, and $x_{g2} = \arg\max_{x>0} g(x)$. Then, it can be seen that $g(x_{g1}) < 0$ and $g(x_{g2}) > 0$, Based on the increasing property of $g(x)$ in (x_{g1}, x_{g2}), there exists a point $x_g^* \in (x_{g1}, x_{g2})$ such that $g(x_g^*) = 0$ and $g(x) > 0$ in (x_g^*, x_{g2}). Since there exists $P_i \in [0, P_{max}]$ such that $g(P_i) > 0$, then $P_{\max} > x_g^*$. Then the solution to (4.10) is

$$\underset{x\in(0,P_{\max}],g(x)>0}{\arg\max} \quad g(x) = \begin{cases} x_{g2} & \text{if } P_{\max} \geq x_{g2}; \\ P_{\max} & \text{if } P_{\max} < x_{g2}. \end{cases} \tag{4.14}$$

where $x_{g2} = \arg\max_{x>0} g(x)$. This completes the proof. $\square$

Based on the above lemma, the desired power vector that satisfies the absolute fairness and the maximum payoff criteria can be found as $\boldsymbol{A}^* = [P^*, ..., P^*]$, where

$$P^* = \begin{cases} \arg\max_{x>0} g(x) & \text{if } P_{\max} \geq \arg\max_{x>0} g(x); \\ P_{\max} & \text{if } P_{\max} < \arg\max_{x>0} g(x). \end{cases} \tag{4.15}$$

Interestingly, the NE with action profile $\boldsymbol{A}^*$ also has the following property.

Lemma 6. *The NE with $\boldsymbol{A}^* = [P^*, ..., P^*]$, is Pareto optimal, that is, no one can increase its expected payoff without decreasing any other node's expected payoff by taking a different action.*

Proof. Take the summation of all nodes' expected payoff in a stage, then

$$\begin{aligned}\sum_{i=1}^{N} v_i(\boldsymbol{A}) &= NMp_s r_0 + \sum_{i=1}^{N} \sum_{j\in\mathcal{N}, j\neq i} r(P_j) - \sum_{i=1}^{N} c(P_i) \\ &= NMp_s r_0 + \sum_{i=1}^{N} g(P_i). \end{aligned} \tag{4.16}$$

By applying action profile $\boldsymbol{A}^* = [P^*, ..., P^*]$, it can be obtained that

$$\sum_{i=1}^{N} v_i(\boldsymbol{A}^*) = NMp_s r_0 + Ng(P^*). \tag{4.17}$$

Since P^* maximizes $g(x)$ in $(0, P_{\max}]$, according to (4.16), it is obvious that

$$\max \sum_{i=1}^{N} v_i(\boldsymbol{A}) = \sum_{i=1}^{N} v_i(\boldsymbol{A}^*). \tag{4.18}$$

For any other action profile $\boldsymbol{A}^{'}$, if $v_i(\boldsymbol{A}^{'}) > v_i(\boldsymbol{A}^*) > 0$, then

$$\begin{aligned}\sum_{j\in\mathcal{N}, j\neq i} v_j(\boldsymbol{A}^{'}) &= \sum_{i=1}^{N} v_i(\boldsymbol{A}^{'}) - v_i(\boldsymbol{A}^{'}) < \sum_{i=1}^{N} v_i(\boldsymbol{A}^*) - v_i(\boldsymbol{A}^*) \\ &= \sum_{j\in\mathcal{N}, j\neq i} v_j(\boldsymbol{A}^*) \end{aligned} \tag{4.19}$$

This indicates that a node cannot increase its expected payoff by taking any other action without decreasing any other node's expected payoff. Thus the NE point $\boldsymbol{A}^*$ is Pareto optimal. This completes the proof. □

So far, the NE that achieves absolute fairness and Pareto optimality is found. Note that $P_{\max}$ is the maximum power that a node is willing to offer for cooperation during the long-period multicast program, and it is usually small in real applications, since nodes may preserve energy for other applications. In this thesis, the scenario that $P^* = P_{\max}$ is considered, that is, $P_{\max} < \arg\max_{x>0} g(x)$. In the following sections, an incentive strategy is proposed to stimulate cooperation at the desired NE point $\boldsymbol{A}^* = [P_{\max}, ..., P_{\max}]$.

4.4 Cooperation Stimulation Strategy

4.4.1 Worst Behavior Tit-for-Tat Incentive Strategy

In this section, a *Worst Behavior Tit-for-Tat* incentive strategy is proposed to stimulate cooperation at the desired equilibrium state. As mentioned before, in the cooperative multicast game, a node makes decisions based on others' past behaviors. And the incentive strategy is a general decision rule such that everyone can benefit by following it rather than deviating from it. Before introducing the strategy, the behavior of a node is defined as follows:

Definition 7. *The behavior of node i observed by node j in stage k, denoted as b_{ij}^k, is defined as the probability that node j successfully receives a packet from node i.*

According to this definition, the observed behavior is a function of a node's action. Note that based on the approximation of p_{ij} as $p_r(P_i)$ in Section 4.1, for any node i, its behavior observed by all other nodes is the same, which can be further simplified as $b_i^k = p_r(P_i^k) = \exp(-\overline{D}/P_i^k)$. Although the transmission power of a node is private information which may not be observed, the packet delivery success probability can be estimated and monitored based on the number of successfully delivered packets. During the multicast process, a node can monitor others' behaviors and adjust its decision accordingly. Then the proposed *Worst Behavior Tit-for-Tat* strategy s^* is as follows:

(i) At the beginning stage, all nodes cooperate with action $A^* = P_{\max}$.

(ii) In each stage, all nodes monitor others' behaviors.

(iii) In stage k, if the worst behavior among node i's observations in the previous stage, denoted as $b_{\dagger}^{k-1}$, is greater than a threshold b_T, then node i takes an action that gives the same behavior as $b_{\dagger}^{k-1}$. Otherwise, node i do not cooperate. Mathematically, from the definition of behavior, it can be obtained that

$$A_i^k = P_i^k = \begin{cases} \frac{-\overline{D}}{\ln b_\dagger^{k-1}}, & \text{if } b_\dagger^{k-1} > b_T; \\ 0, & \text{otherwise.} \end{cases} \tag{4.20}$$

(iv) If node i's observations of all other nodes' behaviors are the same as its own behavior in all the previous L consecutive stages, then node i should use $A^* = P_{\max}$ to resume cooperation in stage k.

The main idea of this strategy is that if a node deviates from the desired action (A^*), all other nodes would behave the same as the deviating node, or even do not cooperate when the deviating behavior is below a threshold. Note that in this strategy, nodes make decisions mainly based on the one-stage observation, and a deviating behavior will result in a reduced reward immediately in the next stage. Then, intuitively, a selfish and rational node will not deviate if the benefit by deviating is less than the corresponding decrease of reward. If a node happens to make a mistake (i.e. due to imperfect monitoring), the desired cooperation state will be damaged since other nodes take the misbehavior as a deviation. To address this issue, the proposed strategy allows nodes to resume cooperation at the desired state when everyone takes the same behavior for a consecutive of L stages.

4.4.2 Analysis of WBTFT Strategy with Perfect Monitoring

According to the proposed strategy, a node makes a decision based on its monitoring results of others' behaviors. Thus, the action of a node is greatly affected by the accuracy of the monitoring technique. To gain some insight, it is worth analyzing an ideal scenario, where the monitoring process is perfect and everyone knows others' actual behaviors (or actions) in the previous stage. The proposed strategy can be analyzed from two aspects: whether the desired NE can be achieved if everyone follows the proposed strategy, and whether the strategy is a subgame perfect equilibrium strategy, where a node cannot get more benefits by deviating to any other strategy from any stage if other nodes follow the proposed strategy.

First, consider the scenario that all nodes follow the proposed strategy from the starting stage. Since the monitoring process is perfect, all nodes' observed behaviors are the same in the first stage. According to the proposed strategy, nodes will cooperate with the desired power $P_{\max}$ in the second stage. Following the same idea, it can be seen that in the rest stages, nodes will always cooperate with $P_{\max}$. Hence, the desired cooperation state can be achieved if all nodes follow the proposed strategy.

The next question is whether the proposed strategy is a subgame perfect equilibrium strategy. In the literature, *One-Shot Deviation Principle* is used to analyze the subgame perfection of a strategy [47], which states that a strategy is subgame perfect if a player cannot get more benefits by deviating for one stage and then coming back to follow it again. Based on this principle, the WBTFT strategy can be analyzed as follows.

Denote $P^\dagger$ as the desired action for the current stage according to the proposed strategy s^*. Without loss of generality, it is assumed $P^\dagger$ has been used by all nodes in the previous K stages, where $K < L$. Then if everyone follows s^*, node i's long term expected payoff is given by

$$V_i(s^*) = \sum_{k=0}^{L-K-1} (\delta)^k v_i(\boldsymbol{A}^k) + \sum_{k=L-K}^{\infty} (\delta)^k v_i(\boldsymbol{A}^*), \tag{4.21}$$

where $A_i^k = P^\dagger$ when $k < L - K$, $\forall i \in \mathcal{N}$. Note that with perfect monitoring, any one-shot deviation behavior of node i will be recognized by other nodes, who will react accordingly in the next L stages. Assume node i employs a one-shot deviation strategy s' in the current stage that gives $A_i^0 = P_i' < P^\dagger$. Then the corresponding long term expected payoff is given by

$$V_i(s') = v_i(\boldsymbol{A}'^0) + \sum_{k=1}^{L} (\delta)^k v_i(\boldsymbol{A}'^k) + \sum_{k=L+1}^{\infty} (\delta)^k v_i(\boldsymbol{A}^*), \tag{4.22}$$

where $A_j'^0 = P^\dagger$, $\forall j \in \mathcal{N}, j \neq i$ and $A_j'^k = P_i'$ ($\forall j \in \mathscr{N}$, $0 < k < L+1$). According to the *One-Shot Deviation Principle*, s^* is a subgame perfect equilibrium strategy if $V_i(s^*) \geq V_i(s')$. The following result can be obtained.

Lemma 7. *For any $P_i' \in (0, P^\dagger]$, $V_i(s^*) \geq V_i(s')$ can be obtained if $G(P^\dagger) \geq G(P_i')$, where $G(x) = \frac{\sum_{k=1}^{L}(\delta)^k}{\sum_{k=0}^{L}(\delta)^k}(N-1)r(x) - c(x)$.*

Proof. First, note that for the one-shot deviating node i, it will not deviate to an action $P_i' > 0$ in the current stage such that $0 < b_i^0 < b_T$. Because in this case, all other nodes who follow the proposed strategy will not cooperate during the next L stages, and the reward for node i is the same as that when choosing $P_i' = 0$, but the cost is larger. Therefore, node i's deviating behavior b_i^0 will either be 0 or $b_i^0 > b_T$, and all other nodes will employ the same behavior during the next L stages. Then from (4.9), (4.21) and (4.22), it can be obtained that

$$V_i(s^*) = \sum_{k=0}^{L-K-1}(\delta)^k\{Mp_s r_0 + g(P^\dagger)\} + \sum_{k=L-K}^{\infty}(\delta)^k\{Mp_s r_0 + g(P_{\max})\} \quad (4.23)$$

where $g(x) = (N-1)r(x) - c(x)$, and

$$\begin{aligned} V_i(s') &= Mp_s r_0 + (N-1)r(P^\dagger) - c(P_i') + \sum_{k=1}^{L}(\delta)^k\{Mp_s r_0 + g(P_i')\} \\ &\quad + \sum_{k=L+1}^{\infty}(\delta)^k\{Mp_s r_0 + g(P_{\max})\} \end{aligned} \quad (4.24)$$

According to the discussion in Section 4.3, $P^* = P_{\max}$ maximizes $g(x)$ in $(0, P_{\max}]$, then $g(P_{\max}) \geq g(P^\dagger)$. Therefore,

$$\begin{aligned} & V_i(s^*) - V_i(s') \\ = & \sum_{k=1}^{L-K-1}(\delta)^k g(P^\dagger) - c(P^\dagger) + \sum_{k=L-K}^{L}(\delta)^k g(P_{\max}) - \left\{\sum_{k=1}^{L}(\delta)^k g(P_i') - c(P_i')\right\} \\ \geq & \sum_{k=1}^{L}(\delta)^k g(P^\dagger) - c(P^\dagger) - \left\{\sum_{k=1}^{L}(\delta)^k g(P_i') - c(P_i')\right\} \\ = & \sum_{k=0}^{L}(\delta)^k\{G(P^\dagger) - G(P_i')\} \end{aligned} \quad (4.25)$$

where $G(x) \triangleq \left(\sum_{k=1}^{L}(\delta)^k / \sum_{k=0}^{L}(\delta)^k\right)(N-1)r(x) - c(x)$. The inequality in (4.25) is obtained by replacing $g(P_{\max})$ with $g(P^{\dagger})$ in the term $\sum_{k=L-K}^{L}(\delta)^k g(P_{\max})$. If $G(P^{\dagger}) \geq G(P_i')$, then from (4.25), it can be concluded that $V_i(s^*) \geq V_i(s')$, this completes the proof. $\square$

Based on the above lemma, it can be seen that the proposed strategy constitutes a subgame perfect equilibrium under certain conditions. The above condition can be further extended, as described in the following proposition.

Proposition 1. *In the cooperative multicast game with perfect monitoring, the Worst Behavior Tit-For-Tat strategy is a subgame perfect equilibrium strategy if the following conditions are satisfied:*
(a) $G(x) > 0$ *for* $x \in (-\overline{D}/\ln b_T, P_{\max}]$, *and* (b) $P_{\max} \leq \arg\max_{x>0} G(x)$.

Proof. Note that $G(x)$ has the same structure as $g(x)$ except that there is a constant factor $\left(\sum_{k=1}^{L}(\delta)^k / \sum_{k=0}^{L}(\delta)^k\right)$. From the proof of lemma 5, it can be seen that the constant factor does not affect the analysis. Following similar procedures, it can be obtained that if there exists an $x \in (0, P_{\max}]$ such that $G(x) > 0$, $G(x)$ monotonically increases in $(\min\{x|G(x) = 0, x > 0\}, \arg\max_{x>0} G(x))$. From the above property, if condition (a) holds and $P_{\max} \leq \arg\max_{x>0} G(x)$, then $G(x) > 0$ and it monotonically increases in $(-\overline{D}/\ln(b_T), P_{\max}]$. Then, given $-\overline{D}/\ln(b_T) \leq P_i' \leq P^{\dagger} \leq P_{\max}$, it can be obtained that $G(P^{\dagger}) \geq G(P_i')$, which indicates the condition in lemma 7 is satisfied and the proposed strategy is a subgame perfect equilibrium strategy. This completes the proof. $\square$

Given that δ is close to 1, b_T and L can be properly selected such that the above two conditions are satisfied. Hence, cooperation can be stimulated using the proposed WBTFT strategy.

4.4.3 Refined Strategy for Imperfect Monitoring

The previous analysis is based on the assumption that the monitoring process is perfect. In this section, the analysis is extended to a more realistic case, where nodes' monitoring results may be erroneous. In a distributed scenario, a node can estimate others' behaviors based on the number of successfully received packets. However, due to packet loss, the monitoring results may be erroneous, and undesired actions may be carried out. Therefore, it is critical to design an estimation method that can achieve high accuracy. In the following, an interval based estimation method is proposed to address the issue of imperfect monitoring.

Typically, the estimation is based on some observed information, such as the number of successfully received packets. In this game model, since the behavior is defined in terms of probability, then any node j's observed information from another node i can be defined as the proportion of packets that node j receives correctly from node i among all the packets that node i should transmit. Denote $y_{ij}^k = \frac{n_{ij}^k}{\overline{n}}$ as the information that node j observes from node i in stage k, where n_{ij}^k is the number of packets that node j receives correctly from node i, and $\overline{n}$ is the average number of packets that a node should transmit in a stage. Note that if node i transmits $\overline{n}$ packets with certain power that results in a behavior b_i^k, then statistically n_{ij}^k follows a binomial distribution $B(\overline{n}, b_i^k)$, and therefore the mean of the observed information y_{ij}^k is b_i^k. Based on this fact, an interval based estimation method can be designed, which is described as follows.

The range [0,1] is divided into $m+1$ intervals, $[0, \beta_1], (\beta_1, \beta_2], ...(\beta_m, 1]$, within the lth interval ($0 \leq l \leq m$), a certain behavior $\hat{b}_l$ is selected to represent the lth behavior level. Then node j's estimated behavior of node i is

$$\widetilde{b}_{ij}^k = \begin{cases} \hat{b}_0 & \text{if } y_{ij}^k \in [0, \beta_1]; \\ \hat{b}_l & \text{if } y_{ij}^k \in (\beta_l, \beta_{l+1}]; \\ \hat{b}_m & \text{if } y_{ij}^k \in (\beta_m, 1], \end{cases} \tag{4.26}$$

where $\hat{b}_0 = \beta_0 = 0$. The interval boundaries ($\{\beta_l\}$) and behavior levels ($\{\hat{b}_l\}$) are designed according to a parameter ε such that if node i transmits $\overline{n}$ packets with behavior $\hat{b}_l$, then any other node's estimation error probability (i.e. the probability that the estimation of node i's behavior is not $\hat{b}_l$) is no larger than a small value 2ε. The design procedures are shown in Algorithm 2 for $1 \leq l \leq m$.

Algorithm 2 Determination of $\hat{b}_l$ and β_l ($l = 1, 2, ..., m$).

1: Select a index value $m = 100$ and let $\hat{b}_m = b^* \triangleq \exp(-D/P_{\max})$;
2: Find $\beta_m \in \{\frac{1}{\overline{n}}, \frac{2}{\overline{n}}, ..., \frac{\overline{n}-1}{\overline{n}}, 1\}$ such that $\Pr\{y_{ij}^k \leq \beta_m | b_i^k = \hat{b}_m, n_i^k = \overline{n}\} = F(\overline{n}\beta_m; \overline{n}, \hat{b}_m) = \varepsilon$;
3: Set $l = m$;
4: **while** $\beta_l > b_T$ **do**
5:     **for** $\beta_{l-1} = \beta_l - \frac{1}{\overline{n}}$ to 0 with step $\frac{1}{\overline{n}}$ **do**
6:         Find $\hat{b}_{l-1}$ that satisfies $F(\overline{n}\beta_{l-1}; \overline{n}, \hat{b}_{l-1}) = \varepsilon$;
7:         **if** $1 - F(\overline{n}\beta_l; \overline{n}, \hat{b}_{l-1}) \leq \varepsilon$ **then**
8:             break and go to step 11,
9:         **end if**
10:     **end for**
11:     Set $l = l - 1$;
12: **end while**
13: Adjust the index values $l \to 1$, $l + 1 \to 2$,..., $m \to m - l + 1$.

In this algorithm, $F(x; n, p)$ is the cumulative distribution function of a binomial random variable $X \sim B(n, p)$. And β_l is selected from the set $\{\frac{1}{\overline{n}}, \frac{2}{\overline{n}}, ..., 1\}$. Note that the number of intervals $m + 1$ is affected by the system parameters (i.e. $P_{\max}$, $\overline{D}$, $\overline{n}$, b_T and ε), and its value is unknown until all the intervals are determined. Since the algorithm starts searching the intervals from the highest level, then the index m is initialized with an arbitrarily chosen large number (i.e. $m = 100$) in the first step, and in the last step, after all intervals are determined, m and all other

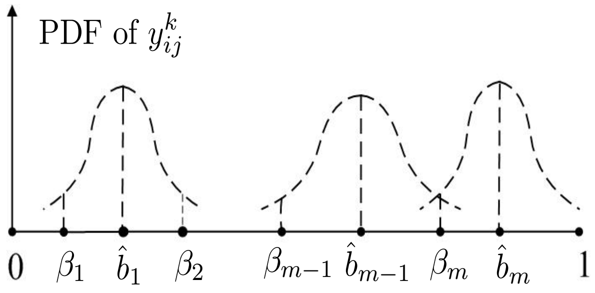

Fig. 4.1. Probability density function (PDF) of y_{ij}^k.

indices are adjusted accordingly based on the number of determined intervals. The details of Algorithm 2 are explained as follows.

In step 1, the desired cooperation behavior b^* is selected as the highest behavior level $\hat{b}_m$. In step 2, β_m is determined according to the design parameter ε, such that if node i takes behavior $\hat{b}_m$, the probability that y_{ij}^k falls in $[0, \beta_m]$ is ε. Here, the design parameter ε is selected from the set $\{F(1; \overline{n}, b^*), F(2; \overline{n}, b^*), ..., 1\}$ to guarantee the existence of β_m which satisfy the above condition, and the determination of ε is discussed in Section 4.5. Then from step 4 to step 12, the boundary points and the corresponding behavior levels for other intervals are found one by one in a repeated manner until β_l is smaller than the threshold b_T. For example, given the right boundary of the $(m-1)$th interval β_m, the algorithm searches the left boundary β_{m-1} from $\beta_m - \frac{1}{\overline{n}}$ to 0 with step size $\frac{1}{\overline{n}}$. For each possible value of β_{m-1}, the corresponding $\hat{b}_{m-1}$ is found such that if node i takes behavior $\hat{b}_{m-1}$, the probability that y_{ij}^k falls in $[0, \beta_{m-1}]$ is $\Pr\{y_{ij}^k \in [0, \beta_{m-1}] | b_i^k = \hat{b}_{m-1}, n_i^k = \overline{n}\} = F(\overline{n}\beta_{m-1}; \overline{n}, \hat{b}_{m-1}) = \varepsilon$ (i.e. the probability in the left tail in the PDF figure of y_{ij}^k, as shown in Fig. 4.1, is ε). Then the desired searching result is largest value of β_{m-1} such that its corresponding $\hat{b}_{m-1}$ and itself satisfy the constraint that

$\Pr\{y_{ij}^k \in (\beta_m, 1] | b_i^k = \hat{b}_{m-1}, n_i^k = \overline{n}\} = 1 - F(\overline{n}\beta_m; \overline{n}, \hat{b}_{m-1}) \leq \varepsilon$. Note that the algorithm searches β_{m-1} from a discrete set, it may not find a pair of β_{m-1} and $\hat{b}_{m-1}$ such that both $\Pr\{y_{ij}^k \in [0, \beta_{m-1}]\}$ and $\Pr\{y_{ij}^k \in (\beta_m, 1]\}$ are equal to ε. To address this issue and to be consistent with the determination of β_m, the searching constraint for the left boundary of the interval is set to be $\Pr\{y_{ij}^k \in [0, \beta_{m-1}]\} = \varepsilon$ and the constraint for the right boundary is $\Pr\{y_{ij}^k \in (\beta_m, 1]\} \leq \varepsilon$, such that when node i cooperates with $\hat{b}_{m-1}$, another node j's estimation error probability is no larger than 2ε.

According to the interval based estimation method, if nodes take behaviors from the set $\mathscr{B} = \{0, \hat{b}_1, ..., \hat{b}_m\}$, the monitoring results will be more accurate when ε approaches zero. Denote $\mathscr{P} = \{0, \hat{P}_1, ..., \hat{P}_m\}$ as the set of power levels that is associated with the behavior levels in $\mathscr{B}$, where $\hat{P}_l = -\overline{D}/\ln \hat{b}_l$, ($\hat{b}_l \in \mathscr{B}$, $0 < l \leq m$). Then, the WBTFT strategy can be refined as follows. In step (ii), nodes employ the above interval based estimation method to estimate others' behaviors. Moreover, a communication period is added at the end of each stage where nodes exchange the worst observed behavior level index, and it is assumed that the information exchange is perfect. In step (iii), at the beginning of a stage, nodes choose their transmission power from the set $\mathscr{P}$ according to the smallest exchanged behavior level index obtained in the previous stage.

4.4.4 Analysis of Refined Strategy with Imperfect Monitoring

This subsection studies the equilibrium conditions of the refined strategy under imperfect monitoring. With imperfect monitoring, a node is uncertain about others' behaviors in the previous stage, and it is also not sure how others will react in the next stage if it takes a certain action. Therefore, it is difficult to find a closed form formulation of the long term expected payoff due to the infinite number of possible future paths, which makes it difficult to check the equilibrium condition. Note that in reality, a node may be bounded rational [48] when making a decision. That is, when the decision making process is complex and the optimal solution is difficult to reach, a node tends to simplify its analysis and only care those outcomes with large probabilities rather than considering all possible scenarios. Based on this fact, the refined strategy can be analyzed under the bounded rational assumption.

In the following analysis, a bounded rational node i is assumed to have the following characteristics. When node i estimates other nodes' behaviors and estimates its own expected payoff, if node z takes an action that gives the behavior b_z^k in stage k, node i will consider the scenario where another node j's observed information y_{zj}^k falls in the interval $(u_1(b_z^k), u_2(b_z^k))$, which satisfies $\Pr\{y_{zj}^k \leq u_1(b_z^k)|b_z^k\} \leq \eta$ and $\Pr\{y_{zj}^k \geq u_2(b_z^k)|b_z^k\} \leq \eta$, where η is a small value close to zero; and node i will ignore the small-probability scenarios where y_{zj}^k falls outside the interval $(u_1(b_z^k), u_2(b_z^k))$. It can be seen that when estimating the long term payoff, instead of considering all possible outcomes, the bounded rational nodes will focus on the monitoring results within a range with a large coverage probability (equal to or greater than $1 - 2\eta$). Intuitively, when η becomes smaller, the estimated payoff is closer to the expected payoff. On the other hand, a smaller η may result in higher computational complexity that is difficult to handle. Note that when all nodes follow the refined strategy, they will choose behaviors in the set $\mathscr{B}$. In this case, according to the interval based estimation, it is easy to check that a bounded rational node i with $\eta \geq \varepsilon$ will only consider the scenario that its observed information from another node j with behavior $b_j^k = \hat{b}_l$ falls in the interval $(u_1(\hat{b}_l), u_2(\hat{b}_l))$, where $\Pr\{y_{ji}^k \leq u_1(\hat{b}_l)|\hat{b}_l\} \leq \varepsilon \leq \eta$ and $\Pr\{y_{ji}^k \geq u_2(\hat{b}_l)|\hat{b}_l\} \leq \varepsilon \leq \eta$. From the interval based estimation, it can be shown that $u_1(\hat{b}_l) \geq \beta_l$ and $u_2(\hat{b}_l) \leq \beta_{l+1}$, and node i's estimation result of node j's behavior is exactly $\widetilde{b}_{ji}^k = \hat{b}_l = b_j^k$. Similarly, when estimating its long term payoff, node i will only consider the scenario that all nodes' estimations in the future stages are correct, which greatly simplifies the calculation of the estimated payoff. In this work, we use $\eta = \varepsilon$ when analyzing the equilibrium.

Similar to the analysis for the perfect monitoring scenario, the one-shot deviation principle is also employed to study the equilibrium condition for the refined strategy. Assume all nodes follow the refined strategy in the past stages, and node i decides to take a one-shot deviation in the current stage. If node i takes an action from the power set $\mathscr{P}$, according to the bounded rational assumption and interval based estimation, it will consider that everyone's estimation is accurate. Then the equilibrium analysis is the same to that under the perfect monitoring scenario, and the condition in proposition 1 is sufficient for the refined strategy being an equilibrium strategy. The next question is whether node i can benefit more by deviating to an action other than those in $\mathscr{P}$, which is analyzed as follows.

Denote $\hat{P}_{l^\dagger}$ as the desired action for the current stage according to the refined strategy, where $l^\dagger$ represents the power level index. Without loss of generality, it is assumed that if all nodes follow the strategy, they will cooperate with $P_{l^\dagger}$ in the next L stages and then resume cooperation in stage $L+1$. Assume node i decides to take a one-shot deviation in the current stage with $P_i' < \hat{P}_{l^\dagger}$, where $P_i' \in (\hat{P}_l, \hat{P}_{l+1}]$, and its corresponding behavior is $b_i^0 \in (\hat{b}_l, \hat{b}_{l+1}]$, where $\hat{b}_{l+1} \leq \hat{b}_{l^\dagger}$. Based on the bounded rational assumption with $\eta = \varepsilon$, node i will believe that any other node j's observation y_{ij}^0 falls in the interval $(u_1(b_i^0), u_2(b_i^0))$, where $\Pr\{y_{ij}^0 \leq u_1(b_i^0)|b_i^0\} \leq \varepsilon$ and $\Pr\{y_{ij}^0 \geq u_2(b_i^0)|b_i^0\} \leq \varepsilon$. From the interval based estimation, it can be shown that $u_1(b_i^0) \geq \beta_l$ and $u_2(b_i^0) \leq \beta_{l+2}$. Therefore, node i will believe that j's estimated behavior $\widetilde{b}_{ij}^0$ is either $\hat{b}_l$ or $\hat{b}_{l+1}$. Denote $w_{ij}(\hat{b}_{l+1}|b_i^0)$ as node i's estimated probability that node j's estimation is $\widetilde{b}_{ij}^0 = \hat{b}_{l+1}$ given $b_i^0 \in (\hat{b}_l, \hat{b}_{l+1}]$, which can be calculated as

$$w_{ij}(\hat{b}_{l+1}|b_i^0) = \frac{\Pr\{y_{ij}^0 \in (\beta_{l+1}, u_2(b_i^0))|b_i^0\}}{\Pr\{y_{ij}^0 \in (u_1(b_i^0), u_2(b_i^0))|b_i^0\}}. \tag{4.27}$$

Here, node i's estimated probability of an event means the probability that node i estimates the event happens. Since nodes exchange their worst behavior levels at the end of each stage, node i's estimated probability that the worst behavior level in current stage is $\hat{b}_{l+1}$, denoted as $w_i(\hat{b}_{l+1}|b_i^0)$, is

$$w_i(\hat{b}_{l+1}|b_i^0) = \prod_{j\in\mathcal{N}, j\neq i} w_{ij}(\hat{b}_{l+1}|b_i^0). \tag{4.28}$$

Then, node i's estimated probability that the worst behavior level in the current stage is $\hat{b}_l$ is $w_i(\hat{b}_l|b_i^0) = 1 - w_i(\hat{b}_{l+1}|b_i^0)$. Hence, node i's estimated long term payoff is

$$\begin{aligned}\widetilde{V}_i(P_i') &= Mp_s r_0 + (N-1)r(\hat{P}_{l^\dagger}) - c(P_i') \\ &\quad + \sum_{k=1}^{L}(\delta)^k \left\{ v_i(\hat{P}_l)w_i(\hat{b}_l|b_i^0) + v_i(\hat{P}_{l+1})w_i(\hat{b}_{l+1}|b_i^0) \right\} \\ &\quad + \sum_{k=L+1}^{\infty}(\delta)^k v_i(P_{\max}).\end{aligned} \tag{4.29}$$

Based on the above discussion, the following result can be obtained.

Lemma 8. *Assume the conditions in proposition 1 are satisfied, for any one-shot deviation action $P_i' \in (\hat{P}_l, \hat{P}_{l+1})$ that gives behavior b_i^0, $\widetilde{V}_i(\hat{P}_{l+1}) > \widetilde{V}_i(P_i')$ can be obtained if $\widetilde{G}(b_i^0, l) = c(\frac{-\overline{D}}{\ln b_i^0}) - \left[w_i(\hat{b}_l|b_i^0)c(\hat{P}_l) + w_i(\hat{b}_{l+1}|b_i^0)c(\hat{P}_{l+1}) \right] > 0$.*

Proof. If node i takes action $\hat{P}_{l+1}$ in the current stage, then according to the bounded rational assumption and interval based estimation, $w_i(\hat{b}_{l+1}|b_i^0) = 1$. From (4.29), it can be obtained that

$$\begin{aligned}&\widetilde{V}_i(\hat{P}_{l+1}) \\ =\ & Mp_s r_0 + (N-1)r(\hat{P}_{l^\dagger}) - c(\hat{P}_{l+1}) \\ & + \sum_{k=1}^{L}(\delta)^k\{Mp_s r_0 + g(\hat{P}_{l+1})\} + \sum_{k=L+1}^{\infty}(\delta)^k\{Mp_s r_0 + g(P_{\max})\} \\ =\ & \sum_{k=0}^{\infty}(\delta)^k Mp_s r_0 + (N-1)r(\hat{P}_{l^\dagger}) \\ & + \sum_{k=0}^{L}(\delta)^k G(\hat{P}_{l+1}) + \sum_{k=L+1}^{\infty}(\delta)^k g(P_{\max})\end{aligned} \tag{4.30}$$

and

$$
\begin{aligned}
& \widetilde{V}_i(P_i') \\
= & \sum_{k=0}^{\infty}(\delta)^k M p_s r_0 + (N-1) r(\hat{P}_{l^\dagger}) - c(P_i') \\
& + \sum_{k=1}^{L}(\delta)^k[w_i(\hat{b}_l|b_i^0) g(\hat{P}_l) + w_i(\hat{b}_{l+1}|b_i^0) g(\hat{P}_{l+1})] + \sum_{k=L+1}^{+\infty}(\delta)^k g(P_{\max}) \\
= & \sum_{k=0}^{\infty}(\delta)^k M p_s r_0 + (N-1) r(\hat{P}_{l^\dagger}) + \sum_{k=L+1}^{+\infty}(\delta)^k g(P_{\max}) \\
& - c(P_i') + w_i(\hat{b}_l|b_i^0) c(\hat{P}_l) + w_i(\hat{b}_{l+1}|b_i^0) c(\hat{P}_{l+1}) \\
& + \sum_{k=0}^{L}(\delta)^k[w_i(\hat{b}_l|b_i^0) G(\hat{P}_l) + w_i(\hat{b}_{l+1}|b_i^0) G(\hat{P}_{l+1})]
\end{aligned} \tag{4.31}
$$

From the definition of behavior, it can be obtained that $P_i' = \frac{-\overline{D}}{\ln b_i^0}$. Since $w_i(\hat{b}_l|b_i^0) + w_i(\hat{b}_{l+1}|b_i^0) = 1$, it can be obtained that

$$
\widetilde{V}_i(\hat{P}_{l+1}) - \widetilde{V}_i(P_i') = \widetilde{G}(b_i^0, l) + \sum_{k=0}^{L} w_i(\hat{b}_l|b_i^0)\{G(\hat{P}_{l+1}) - G(\hat{P}_l)\}. \tag{4.32}
$$

According to the proof of proposition 1, $G(\hat{P}_{l+1}) > G(\hat{P}_l)$ for $\hat{P}_{l+1} > \hat{P}_l$. Then, if $\widetilde{G}(b_i^0, l) > 0$ is satisfied, it can be concluded that $\widetilde{V}_i(\hat{P}_{l+1}) > \widetilde{V}_i(P_i')$, this completes the proof. □

Based on the above result, it can be seen that a node will not deviate to an action other than those in $\mathscr{P}$ under certain conditions. To summarize, the equilibrium conditions for the refined strategy can be stated as follows.

Proposition 2. *In the cooperative multicast game with bounded rational nodes, the refined WBTFT strategy is an equilibrium strategy under imperfect monitoring if the following conditions are satisfied:*

(a) *Conditions in Proposition 1 are satisfied.*

(b) $\widetilde{G}(x, l) = c(\frac{-\overline{D}}{\ln x}) - \left[w_i(\hat{b}_l|x) c(\hat{P}_l) + w_i(\hat{b}_{l+1}|x) c(\hat{P}_{l+1})\right] > 0, \ \forall\, 0 \le l < m$ *where* $x \in (\hat{b}_l, \hat{b}_{l+1})$.

It can be seen that in the imperfect monitoring scenario with interval based estimation, besides the equilibrium conditions in Proposition 1, an additional condition is required to guarantee that any deviation to behaviors other than those in $\mathscr{B}$ is not beneficial. The above conditions can be satisfied by appropriately choosing the parameters ε and L, and hence the refined strategy maintains an equilibrium with imperfect monitoring.

4.4.5 WBTFT Strategy with Optimal Power Allocation

In the previous game model, it is the assumed that nodes use the same transmission power for all time slots during one stage. However, nodes may employ different power levels for different time slots or only transmit a portion of packets when selected as relays in a stage. Therefore, it is essential to investigate the optimal power allocation scheme within a stage.

First, consider the scenario that nodes vary the transmission power within a stage. Suppose node i transmits n_i packets in a stage, the power allocated for the tth packet is denoted as P_t, $t = 1, 2, ..., n_i$, and the total cost in this stage is $c_i = \sum_{t=1}^{n_i} c_0 P_t$. In this scenario, the behavior of node i observed by any other node can be represented as $\overline{b}_i = \frac{1}{n_i} \sum_{t=1}^{n_i} p_r(P_t)$. Obviously, a node prefers to provide large $\overline{b}_i$ with low cost. Then given a desired behavior, the optimal power allocation scheme should achieve the lowest cost. Based on this idea, the following lemma can be proved.

Lemma 9. *To obtain the same $\overline{b}_i$, where $\overline{b}_i > \frac{1}{e}$, the optimal power allocation scheme that gives the lowest cost c_i is $P_t = \overline{P} = \frac{-\overline{D}}{\ln(\overline{b}_i)}$ $\forall\, t = 1, 2, ..., n_i$.*

Proof. Denote $s^{\ddagger} = [P_1, P_2, ...P_{n_i}]$ as a power allocation scheme which is different with the equal power allocation scheme $\overline{s} = [\overline{P}, ..., \overline{P}]$. Let $b_t = \exp(-\overline{D}/P_t)$ for $P_t \in s^{\ddagger}$ and $\overline{b} = \exp(-\overline{D}/\overline{P})$, assume $s^{\ddagger}$ and $\overline{s}$ obtain the same behavior $\overline{b}_i$, then

$\bar{b}_i = \frac{1}{n_i}\sum_{t=1}^{n_i} b_t = \bar{b}$. The corresponding cost for each scheme is given by

$$c_i(s^{\ddagger}) = \sum_{t=1}^{n_i} c_0 \frac{-\overline{D}}{\ln b_t}, \qquad c_i(\bar{s}) = n_i c_0 \frac{-\overline{D}}{\ln \bar{b}}. \tag{4.33}$$

Denote $a(x) = -1/\ln x$, then for $\bar{b} = \bar{b}_i > 1/e$, it is proved in Appendix C that $a(x)$ satisfies the following property,

$$a(\bar{b}) = a\left(\frac{b_1 + b_2 + ... + b_{n_i}}{n_i}\right) \leq \frac{a(b_1) + a(b_2) + ... + a(b_{n_i})}{n_i}. \tag{4.34}$$

where the equal sign is achieved when $b_1 = ... = b_{n_i} = \bar{b}$. From (4.33) and (4.34), it can be concluded that $c_i(\bar{s}) \leq c_i(s^{\ddagger})$, this completes the proof. $\square$

Lemma 9 implies that using the same power to transmit the packets can provide others fixed observed information with the smallest cost. According to the WBTFT strategy, nodes make decisions based on the observed information. Therefore, to gain more profit, they will keep the transmission power unchanged within a stage.

The next question is whether a node will transmit all the time when selected as a relay in a stage. Suppose node i should transmit $\overline{n}$ packets in a stage, and its desired behavior is denoted as $b^{\dagger}$, which is assumed to be greater than $1/e$. Lemma 9 implies that node i will transmit packets with the same power. However, it can decide to transmit with a certain probability (or equivalently, only transmit a portion of $\overline{n}$ packets). Denote φ_i as the probability (or the proportion of $\overline{n}$) that node i decides to transmit with fixed power $\overline{P}_i$ when selected as a relay, then the expected number of packets that any other node j can successfully receive is $n_{ij} = \overline{n}\varphi_i p_r(\overline{P}_i)$. Equivalently, the behavior that any other node should observe in the stage can be represented as $\bar{b}_i = \varphi_i p_r(\overline{P}_i)$, and the corresponding cost in the stage is given by $c_i = c_0 \overline{n}\varphi_i \overline{P}_i$. Since nodes are selfish and tend to maximize their payoff, then determining the optimal φ_i and $\overline{P}_i$ is equivalent to solving the following optimization problem:

$$\min \quad \varphi_i \overline{P}_i \quad \text{subject to} \quad \bar{b}_i = \varphi_i p_r(\overline{P}_i) = b^{\dagger}. \tag{4.35}$$

Let x represent $\overline{P}_i$, according to the constraint, it can be obtained that $\varphi_i = b^\dagger/p_r(x)$. Define $f(x) \triangleq \varphi_i \overline{P}_i = b^\dagger x/p_r(x)$, then $\frac{df(x)}{dx} = \frac{b^\dagger(1-\overline{D}/x)}{p_r(x)}$. It can be seen that $f(x)$ has the following property:

$$\begin{cases} \frac{df(x)}{dx} < 0 & \text{if } 0 < x < \overline{D}; \\ \frac{df(x)}{dx} = 0 & \text{if } x = \overline{D}; \\ \frac{df(x)}{dx} > 0 & \text{if } x > \overline{D}. \end{cases} \tag{4.36}$$

Note that $\varphi_i \in (0, 1]$, it can be obtained that $p_r(x) \geq b^\dagger$, which leads to $x \geq \frac{-\overline{D}}{\ln b^\dagger}$. It can be seen that if $\frac{-\overline{D}}{\ln b^\dagger} < \overline{D}$ which is $b^\dagger < \frac{1}{e}$ then $\min\{f(x)\} = f(\overline{D})$, if $\frac{-\overline{D}}{\ln b^\dagger} \geq \overline{D}$, then $\min\{f(x)\} = f(\frac{-\overline{D}}{\ln b^\dagger})$. Therefore, the solution to (4.35) can be summarized as

$$\begin{cases} \varphi_i = eb^\dagger, \overline{P}_i = \overline{D}, & \text{if } b^\dagger < \frac{1}{e}; \\ \varphi_i = 1, \overline{P}_i = \frac{-\overline{D}}{\ln b^\dagger}, & \text{if } b^\dagger \geq \frac{1}{e}. \end{cases} \tag{4.37}$$

It can be seen that if the desired behavior $b^\dagger$ is no less than $1/e$, a node will always transmit with fixed power when it is selected as a relay. In addition, based on the above discussion, in the WBTFT strategy, if $b_T \geq 1/e$, then the optimal power allocation within a stage can also be achieved.

4.4.6 Further Discussion of Designing Issues

In the previous section, the WBTFT strategy is refined with an interval based estimation method to address the issue of imperfect monitoring. However, the design of intervals is based on the approximation that the packet delivery success probability between node i and node j at each stage is $p_{ij} \approx \exp(\frac{-D}{P_i})$, where D depends on the long term average channel condition between node i and node j. In real applications, given the mobility model, the average channel condition between two mobile nodes may fluctuate from stage to stage, which may result in large estimation errors when using the intervals designed according to D. To address this problem, the following adjustment of the interval based estimation is made. First, a general behavior level set $\mathscr{B} = \{0, \hat{b}_1, ..., \hat{b}_m\}$ (with interval boundaries $\{\beta_l\}$) is designed according to Algorithm 2 based on D, and its corresponding power set $\mathscr{P}$ is derived. Note that in the refined strategy, nodes choose transmission power from the same power set $\mathscr{P}$ based on the worst behavior level index. Next, in stage k, node j estimates the current average channel condition between node i and itself (either by exchanging topology information or using pilot signals), and calculates the average value of $\frac{N_0(1-2^{\gamma_r})}{\sigma_{ij}^2}$, denoted as D_{ij}^k. Then, node j adjusts its estimation of node i's behavior level as follows. In stage k, given $\mathscr{P}$ and D_{ij}^k, if node i chooses power $\hat{P}_l \in \mathscr{P}$ $(1 \leq l \leq m)$, node j's observed behavior of node i is $\hat{b}_{ijl}^k = \exp\left(-D_{ij}^k/\hat{P}_l\right) = (\hat{b}_l)^{D_{ij}^k/D}$, and node j's observed behavior level set of node i is updated as $\mathscr{B}_{ij}^k = \{0, \hat{b}_{ij1}^k, \hat{b}_{ij2}^k, ..., \hat{b}_{ijm}^k\}$. Then, the interval boundaries for each behavior level in $\mathscr{B}_{ij}^k$ are updated, by using $\beta_{ijl}^k = (\beta_l)^{\overline{D}_{ij}^k/\overline{D}}$. Then node j estimates node i's behavior level according to the new intervals at the end of stage k. That is, if node j's observed information of node i, y_{ij}^k, falls in the lth interval $(\beta_{ijl}^k, \beta_{ij(l+1)}^k]$, then node j's estimation result of node i's behavior level is $\hat{b}_{ijl}^k$. Note that for the original behavior levels in $\mathscr{B}$, the intervals are designed such that the estimation error probability for each behavior level is no larger than 2ε (i.e., $\Pr\left\{y_{ij}^k \in [0, \beta_l] \cup y_{ij}^k \in (\beta_{l+1}, 1] | b_i^k = \hat{b}_l, n_i^k = \overline{n}\right\} \leq 2\varepsilon$). Unfortunately, this feature cannot be maintained for the adjusted intervals. The estimation error probability decreases with the behavior level index when $D_{ij}^k > D$ and increases when $D_{ij}^k < D$. An example is shown in Fig. 4.2. To guarantee the accuracy of estimation, a threshold value ε_T is set. In stage k, if node j's observed information from node i, y_{ij}^k, falls in the lth adjusted interval, then node j first calculates $p_e = \Pr\left\{y_{ij}^k \in [0, \beta_{ijl}^k] \cup y_{ij}^k \in (\beta_{ij(l+1)}^k, 1] | b_i^k = \hat{b}_{ijl}^k, n_i^k = \overline{n}\right\}$ and compares p_e with ε_T. If $p_e > \varepsilon_T$, then the estimation result is unreliable and will be discarded. At the end of each stage, nodes only exchange the worst behavior level index from their reliable estimation results and make decisions accordingly.

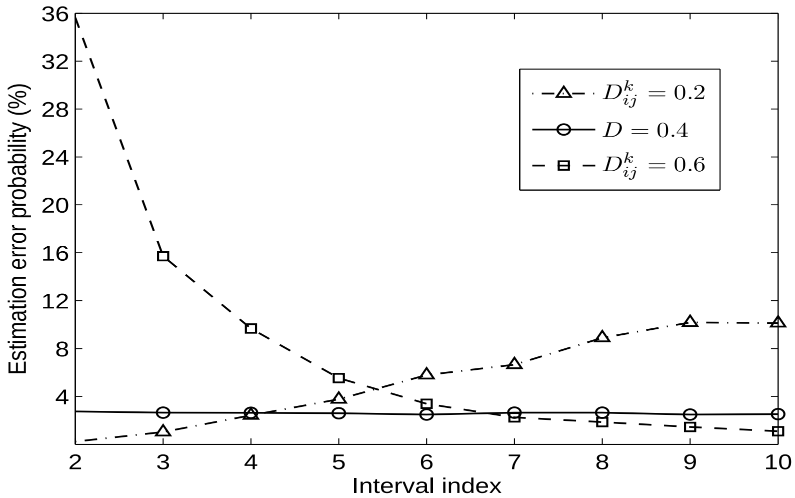

Fig. 4.2. Estimation error probability versus interval index.

4.5 Performance Evaluation

In this section, the performance of the proposed strategy is evaluated via simulation. Before showing the simulation results, it is worth investigating how to select the design parameters to satisfy the equilibrium conditions.

4.5.1 Evaluation of the NE Conditions

First, consider the conditions in Proposition 1. As shown in the proof of Proposition 1, $G(x)$ has the same property as $g(x)$. That is, if there exists an x such that $G(x) > 0$, then $G(x)$ monotonically increases in $(\min\{x|G(x) = 0, x > 0\}, \arg\max_{x>0} G(x)]$. Then, the conditions become $\min\{x|G(x) = 0, x > 0\} \leq -\overline{D}/\ln(b_T)$ and $P_{\max} \leq \arg\max_{x>0} G(x)$. Consider a system setup with $N = 10$, $p_s = 0.4$, $\overline{D} = 0.4$, $\delta = 0.99$ and $r_0/c_0 = 100$. Fig. 4.3 shows $\min\{x|G(x) = 0, x > 0\}$ and $\arg\max_{x>0} G(x)$ versus different L. It can be seen that as L increases, $\arg\max_{x>0} G(x)$ becomes larger, while $\min\{x|G(x) = 0, x > 0\}$ is close to zero, which implies the conditions are easier to satisfy when L becomes larger. Note that to achieve optimal power allocation in a stage, b_T should be no less than

$1/e$. Therefore, b_T can be selected such that $b_T \geq \max\{\exp(-\overline{D}/\min\{x|G(x) = 0, x > 0\}), 1/e\}$. In this system setup, if the unit power is 10 mW, and $P_{\max}$ is less than 107 mW, then $L = 1$ and $b_T = 1/e$ can satisfy the required conditions.

Next, consider the conditions in Proposition 2. It can be seen that condition (b) of Proposition 2 is related to the behavior intervals, which are determined based on the parameter ε. According to the interval based estimation, a small ε is preferred because the estimation result becomes more and more accurate when ε approaches zero. However, a small ε may violate the equilibrium conditions. Thus, given a network setup, the smallest ε that satisfies the conditions in Proposition 2 can be found via numerical methods. Some values of the minimum ε under different $P_{\max}$ are given in Table 4.1. The design parameter $\overline{n}$ is 1000, and other parameters are the same as the previous setup in Fig. 4.3. It can be seen that the minimum ε are all smaller than 0.02, which is acceptable.

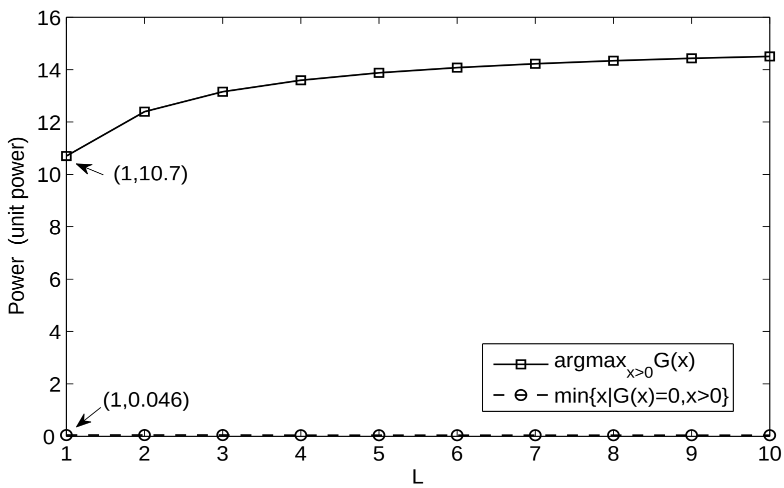

Fig. 4.3. Condition in Proposition 1.

TABLE 4.1

Minimum ε vs $P_{\max}$.

$P_{\max}$	1	2	3	4	5	6	7	8	9	10
Minimum ε	0.0119	0.0113	0.0106	0.0126	0.0158	0.0090	0.0121	0.0132	0.0181	0.0133

4.5.2 Simulation Results

In this part, simulation results are shown to demonstrate the effectiveness of the proposed incentive strategy. The simulated multicast network consists 10 mobile nodes that are randomly deployed within a circular area with radius $d_0 = 25$m. Each node is moving according to the random waypoint model: a node randomly chooses a destination and moves forward to the destination with a velocity uniformly chosen in [0.5m/s, 2.5m/s]. When arriving at the destination, the node will choose a new location and a new speed to move on. The source broadcasts a packet every 10 ms with rate 1 Mbps over a wireless channel with 1MHz bandwidth. The maximum power that a node offers for cooperation is 40 mW. When selected as a relay, a node can choose the power between 0 and 40 mW to rebroadcast the packet with rate 1 Mbps. All the wireless channels undergo path loss with exponent 2 and Rayleigh fading, and the average received SNR from the source to a receiving node is 0 dB. The average received SNR between a pair of receiving nodes with distance d_0 and unit transmission power 8 mW is 2 dB. The parameter $\overline{D}$ is calculated by averaging results from simulations of the above mobility model in 10^7 time slots and is approximated by $\overline{D} = 0.40$. Other system parameters are set as $r_0 = 100$, $c_0 = 2$ and $\delta = 0.99$.

Fig. 4.4 shows the average payoff and throughput of all nodes in different stages when everyone follows the proposed strategy. The intervals are designed according to $\overline{n} = 1000$ and $\varepsilon = 0.0158$. The corresponding stage length is $M = 10103$ time slots. Other parameters are $L = 1$, $b_T = 1/e$ and $\varepsilon_T = 0.1$. This simulation contains 50 runs, each with 50 stages, and the average payoff and the throughput of all nodes are plotted. The results with perfect monitoring process where all nodes cooperate at the desired power level are also plotted as a benchmark. It can be seen that the WBTFT strategy with interval based estimation can achieve cooperation efficiently at a level with average payoff 92.08 per time slot, which is close to the desired cooperation state with average payoff 93.37 per time slot.

In Fig. 4.5, the payoff and throughput of a node with two different strategies, the proposed strategy and a deviation strategy, are compared. The deviation strategy is characterized by a deviating ratio θ, which is defined as the ratio between the deviating behavior and the desired behavior according to the WBTFT strategy. Specifically, if the desired behavior in stage k is b^k for all nodes, then a node with a deviating ratio θ will take an action P_i^k such that $b_i^k = \exp(-\overline{D}/P_i^k) = \theta \cdot b^k$. In this simulation, node 1 deviates with $\theta = 0.5$ in stage 4,7,10,..., 49, and in other stages, it follows the proposed strategy. The design parameters are the same as those in Fig. 4.4. It can be seen that whenever node 1 deviates, its misbehavior will be identified and the corresponding payoff is reduced. Therefore, the proposed strategy is able to punish the deviating node. The average payoff with deviation strategy is 89.35 per time slot, which is less than that by following the proposed strategy, hence the proposed strategy can motivate cooperation efficiently.

Fig. 4.6 shows the impact of L on the performance. Other parameters are the same as the previous figures and there is no deviating node. It can be seen that as L increases, the performance slightly degrades. This is mainly because when estimation error occurs, a smaller L can help resume cooperation faster. Therefore, when designing the incentive strategy for real applications, a small L is preferred so that cooperation can be easily recovered when undesired actions are carried out. Fig. 4.7 shows a similar trend to Fig. 4.6, where the performance degrades as ε_T increases. The reason is that when ε_T increases, nodes will exchange unreliable information that contains large estimation errors. Therefore, the desired cooperation behavior might be estimated as a deviation and nodes may not cooperate at the desired level even though no one intentionally deviates.

Fig. 4.8 shows the performance with different b_T. Other parameters are the same as those in Fig. 4.4. In this simulation, the average payoff and throughput of node 1 with two different deviating ratios are compared. Note that $\theta = 1$ means no one deviates. It can be seen that the changing of b_T does not affect the performance when all nodes follow the proposed strategy. However, as b_T increases, the deviating node's average payoff decreases. The reason is that when b_T becomes larger, the punishment becomes more severe, since no one will cooperate when a misbehavior that falls below b_T is detected.

Fig. 4.9 shows the average payoff and throughput for a deviating node with different deviating ratios. The incentive strategy is designed according to $\overline{n} = 100, 500$ and 1000, and the corresponding minimum ε is $0.0210, 0.0173$ and 0.0158, respectively. Other parameters are the same with Fig. 4.4. It can be seen that among the three designs, the one with $\overline{n} = 1000$ and that with $\overline{n} = 500$ achieve relatively close performance, with $\overline{n} = 1000$ being slightly better, and both performances are better than that with $\overline{n} = 100$. This can be explained as follows. First, according to the value of ε, the estimation error probability of $\overline{n} = 1000$ is the smallest, and a larger $\overline{n}$ also means a longer stage, in which nodes can collect more data to estimate each other's transmission probability. Second, the approximation of p_{ij} as $p_r(P_i)$ in the system model is more accurate when the stage is longer, which further reduces the estimation errors. Moreover, according to Algorithm 2, a larger $\overline{n}$ will result in more behavior levels and smaller interval ranges for each level. The behavior levels for $\overline{n} = 100, 500$ and 1000 are listed in Table 4.2. For $\overline{n} = 1000$ and $\overline{n} = 500$, the adjacent behavior levels are close to each other, while for $\overline{n} = 100$, the difference between two adjacent behavior levels is much larger. Therefore, estimation errors have greater impact on the performance when $\overline{n} = 100$. Based on the above discussion, when designing the incentive strategy, a large $\overline{n}$ is profitable.

4.6 Summary

In this chapter, cooperation stimulation among selfish nodes in wireless multicast networks is studied. The cooperative multicast process is formulated as a repeated game and the desired cooperation state that satisfies the absolute fairness and the Pareto optimality criteria is found. A *Worst Behavior Tit-for-Tat* incentive strategy is proposed to stimulate cooperation, and the equilibrium conditions are derived under perfect and imperfect monitoring scenarios. To address the issue of imperfect monitoring, an interval based estimation method is developed. Simulation results show that even with imperfect monitoring, the proposed strategy can efficiently enforce cooperation, and its performance is close to that of the desired cooperation state where all nodes fully cooperate with perfect monitoring.

TABLE 4.2

Behavior levels for $\overline{n} = 100, 500, 1000$.

l	2	3	4	5	6	7	8	9	10	11
Behavior for $\overline{n} = 100$	0.382	0.596	0.782	0.923						
Behavior for $\overline{n} = 500$	0.339	0.432	0.526	0.621	0.710	0.792	0.864	0.923		
Behavior for $\overline{n} = 1000$	0.390	0.457	0.526	0.593	0.659	0.722	0.781	0.835	0.883	0.923

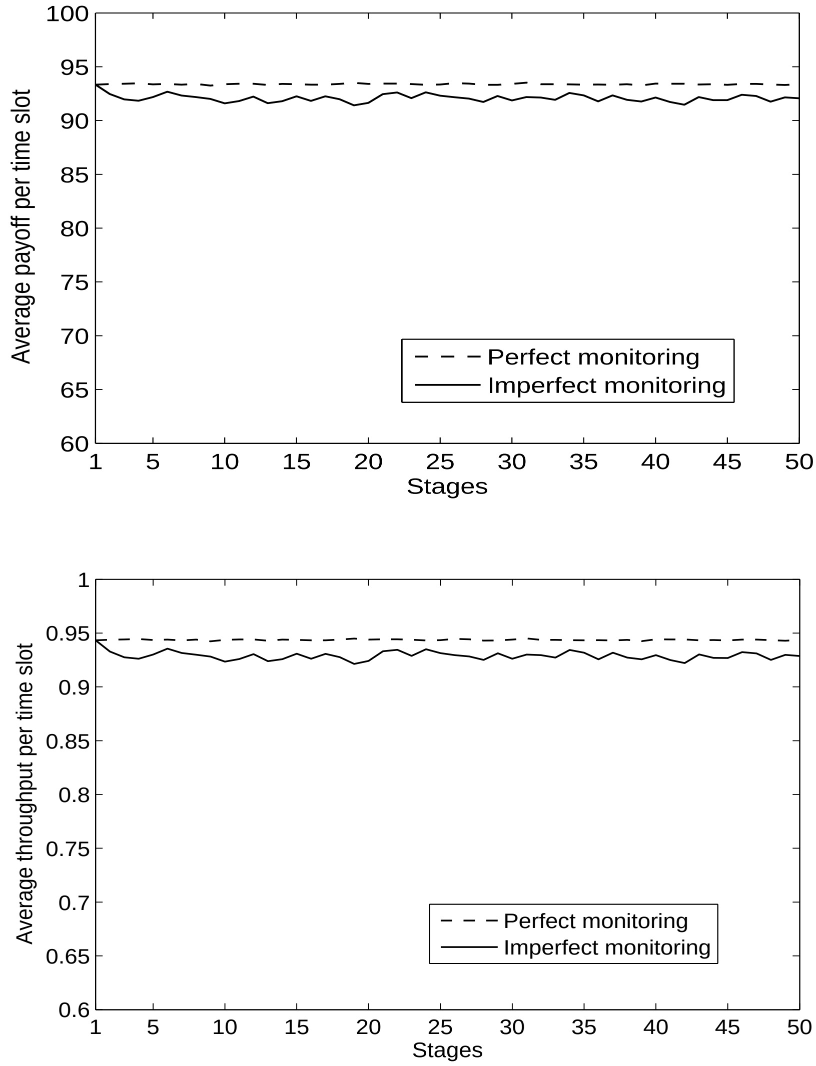

Fig. 4.4. Performance comparison under perfect and imperfect monitoring

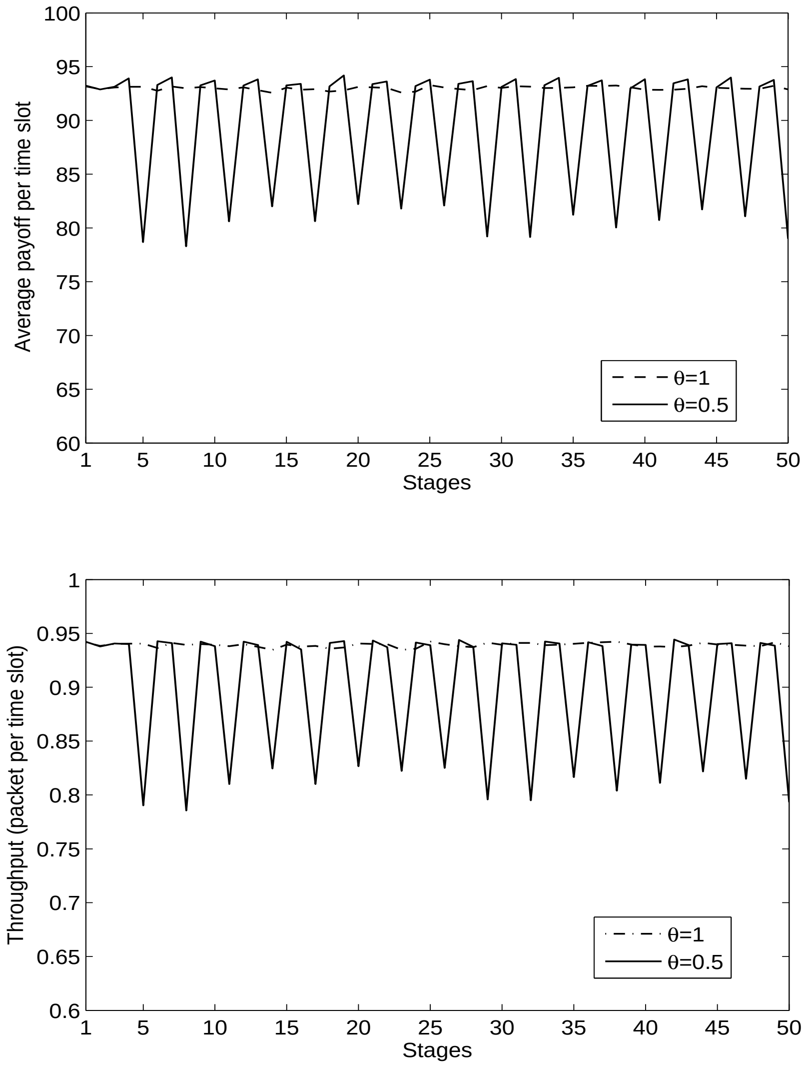

Fig. 4.5. Performance analysis with deviating node

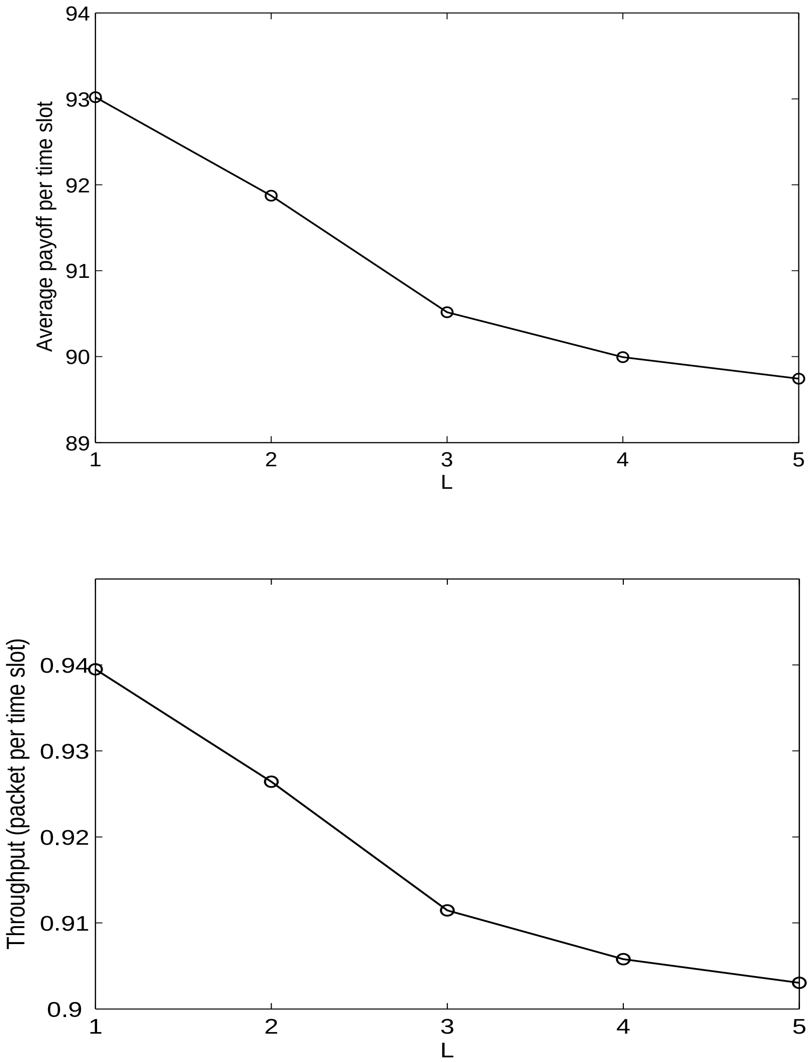

Fig. 4.6. Performance analysis under different L

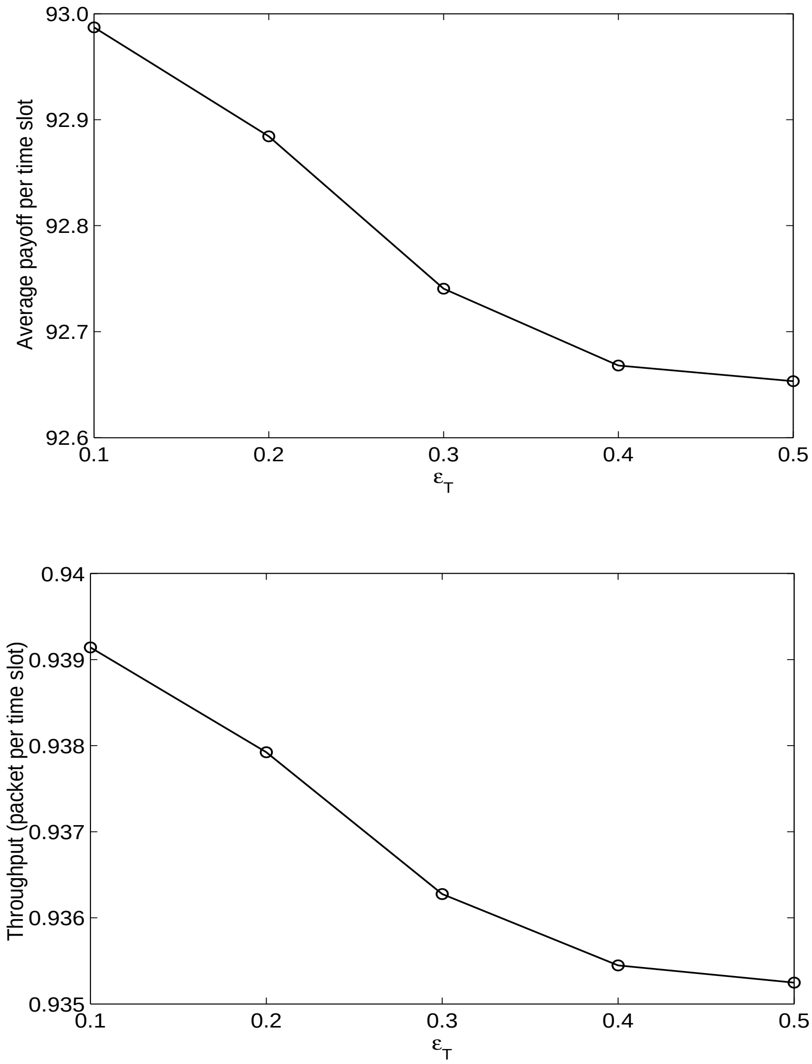

Fig. 4.7. Performance analysis under different ε_T

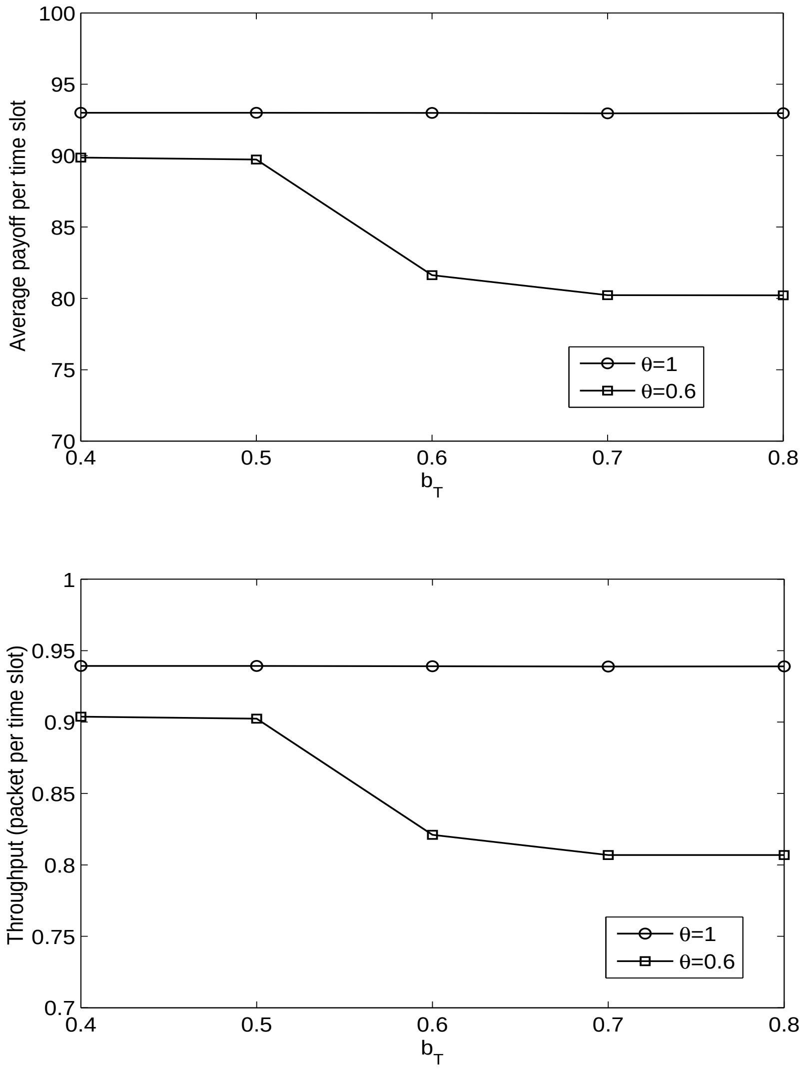

Fig. 4.8. Performance analysis under different b_T

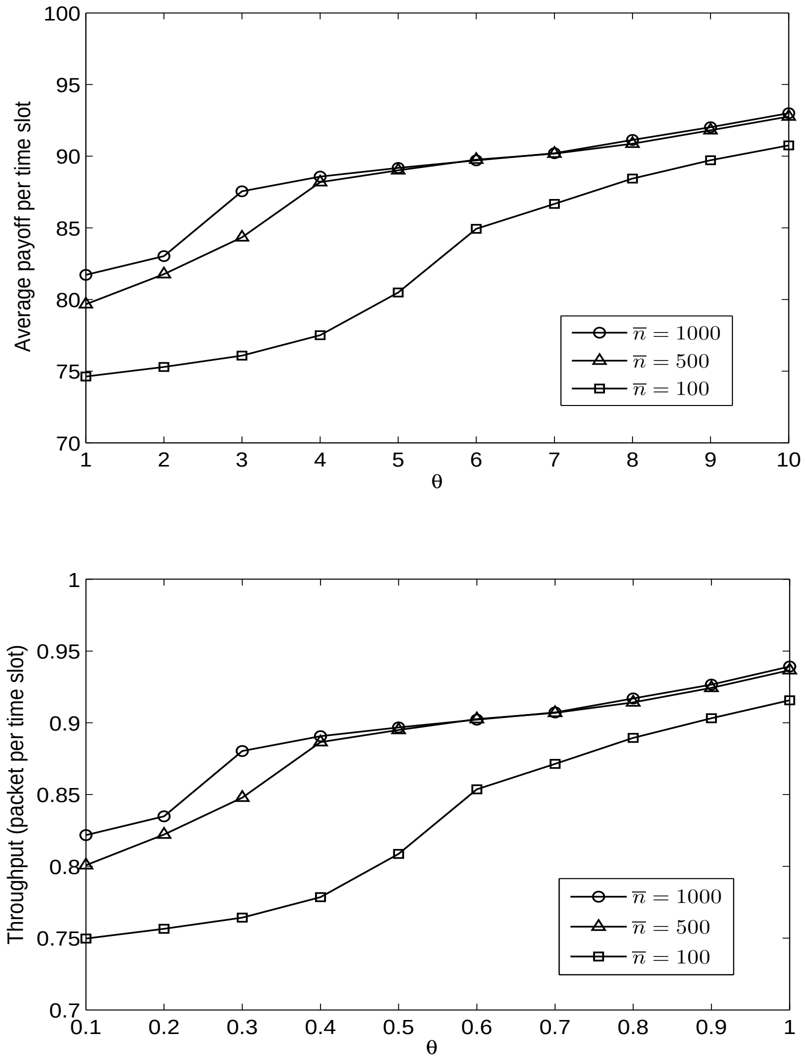

Fig. 4.9. Performance analysis under different $\overline{n}$

Chapter 5

Conclusions and Future Work

5.1 Conclusions

This thesis first proposes a time division based cooperative multicast strategy for delay-sensitive applications. In this strategy, some successful receiving nodes are selected as relays to rebroadcast the data after source transmission. The optimal relay selection and time allocation schemes are analyzed. It has been shown that the optimal relay number is bounded by a threshold value which can be found via numerical methods. An efficient algorithm is also proposed to find the optimal time allocation for source transmission and node cooperation. Numerical results show that the proposed strategy can enhance network performance significantly when the average channel condition between receiving nodes is better than that of the direct link.

After designing the cooperative multicast strategy, this thesis studies cooperation stimulation among mobile selfish nodes in wireless multicast networks using game theoretic approaches. The cooperative multicast process is formulated as an infinite repeated game. The Nash equilibrium of the game is analyzed and the desired cooperation state which satisfies the absolute fairness and the Pareto optimality criteria is found. A *Worst Behavior Tit-for-Tat* incentive strategy is proposed to stimulate cooperation, and the equilibrium conditions are derived under both the perfect and the imperfect monitoring scenarios. It is shown that the proposed strategy achieves optimal power allocation for a selfish node when the designing parameters are appropriately chosen. Moreover, to address the issue of imperfect monitoring, an interval based estimation method is developed. Simulation results show that the proposed strategy can enforce cooperation efficiently even the monitoring is imperfect.

5.2 Future Work

Although this thesis has thoroughly addressed several critical issues in designing cooperative multicast strategies, there still exist many issues that need further investigation.

In the cooperative multicast strategy design, coding scheme is not considered. It is possible to incorporate some coding techniques, to further improve the network performance. For example, this thesis considers that the relays transmit in a time division fashion, where each time only one relay is allowed to rebroadcast data. If space-time coding is employed, multiple relays can transmit simultaneously, and the performance can be further improved. Another possible direction is to jointly consider the power constraint, such as total power or maximum lifetime of the network, and design power efficient cooperation schemes.

In the incentive mechanism design, this thesis only considers the cooperation stimulation among selfish mobile nodes. In static scenarios, the packet delivery success probability between any pair of nodes cannot be approximated using the same method as in this thesis due to the fixed topology. Therefore, some refinement needs to be made for the WBTFT strategy to deal with the heterogeneity of channel conditions. Besides, this thesis only considers selfish nodes who aim to maximize their own profits. From a security perspective, it is also worth investigating cooperation stimulation when there are attackers, whose purpose is to destroy the system. In the proposed strategy, nodes make decisions not only based on its own observation, but also according to the information that others claim. Therefore, an attacker can exchange false information and destroy the cooperation state. A possible solution is to design some criteria to identify the attackers and then remove it from the multicast group by using encryption, which will be considered in the future work.

Appendix A

Proof of (3.8) :

Note that $p(\alpha, x)$ is monotonically decreasing with x and

$$\lim_{x\to 0} p(\alpha, x) = 1, \qquad \lim_{x\to +\infty} p(\alpha, x) = 0.$$

Therefore, it can be obtained that

$$\begin{aligned}\lim_{x\to 0} f(\alpha, x) &= \lim_{p(\alpha,x)\to 1} \left\{ \frac{1 - p(\alpha, x)}{p(\alpha, x)} \ln\big(1 - p(\alpha, x)\big) \right\} \\ &\quad + \lim_{x\to 0} \{\rho\omega(\ln 2) x 2^{\rho x}\} = 0\end{aligned}$$

and

$$\begin{aligned}\lim_{x\to +\infty} f(\alpha, x) &= \lim_{p(\alpha,x)\to 0} \left\{ \frac{1 - p(\alpha, x)}{p(\alpha, x)} \ln\big(1 - p(\alpha, x)\big) \right\} \\ &\quad + \lim_{x\to +\infty} \{\rho\omega(\ln 2) x 2^{\rho x}\} = +\infty.\end{aligned}$$

Proof of (3.9) :

From the proof of lemma 1 it can be obtained that

$$\begin{aligned}\frac{\partial f(\alpha, x)}{\partial x} &= \frac{\partial \left\{ \frac{1-p(\alpha,x)}{p(\alpha,x)} \ln\big(1 - p(\alpha, x)\big) \right\}}{\partial x} + \frac{\partial \{\rho\omega(\ln 2) x 2^{\rho x}\}}{\partial x} \\ &= \rho\omega(\ln 2) 2^{\rho x} \left\{ \frac{1}{p(\alpha, x)} \ln\big(1 - p(\alpha, x)\big) + 2 + \rho(\ln 2) x \right\}. \text{(A-1)}\end{aligned}$$

Let $U_1(\alpha, x) \triangleq \frac{1}{p(\alpha,x)} \ln\big(1 - p(\alpha, x)\big)$. Then it can be shown that

$$\begin{aligned}\frac{\partial U_1(\alpha, x)}{\partial x} &= \left\{\frac{-1}{p^2(\alpha, x)} \ln\big(1 - p(\alpha, x)\big) - \frac{1}{p(\alpha, x)(1 - p(\alpha, x))}\right\} \cdot \frac{\partial p(\alpha, x)}{\partial x} \\ &= \frac{\rho\omega(\ln 2)2^{\rho x}}{p(\alpha, x)\big(1 - p(\alpha, x)\big)}\Big\{\big(1 - p(\alpha, x)\big)\ln\big(1 - p(\alpha, x)\big) + p(\alpha, x)\Big\}.\end{aligned}$$

Let $U_2(p) \triangleq (1-p)\ln(1-p)+p$, then $\frac{\partial U_2(p)}{\partial p} = -\ln(1-p) > 0$ for $p \in (0, 1)$. Therefore, $U_2(p)$ is monotonically increasing when $p \in (0, 1)$. Since $\lim_{p\to 0} U_2(p) = 0$, it indicates $U_2(p) > 0$ for $p \in (0, 1)$. Then it is obvious that $\frac{\partial U_1(\alpha,x)}{\partial x} > 0$ with $p(\alpha, x) \in (0, 1)$ where $x \in (0, +\infty)$. Thus, $U_1(\alpha, x)$ is a monotonically increasing function of x for $x \in (0, +\infty)$, and so is the function $W(\alpha, x) \triangleq U_1(\alpha, x) + 2 + \rho \ln 2 \cdot x$.

It can be shown that

$$\lim_{x\to 0} W(\alpha, x) = \lim_{p(\alpha,x)\to 1} \left\{\frac{1}{p(\alpha, x)} \ln\big(1 - p(\alpha, x)\big)\right\} + 2 = -\infty$$

and

$$\begin{aligned}\lim_{x\to +\infty} W(\alpha, x) &= \lim_{p(\alpha,x)\to 0} \left\{\frac{1}{p(\alpha, x)} \ln\big(1 - p(\alpha, x)\big)\right\} + 2 \\ &\quad + \lim_{x\to +\infty} \{\rho \ln 2 \cdot x\} = +\infty.\end{aligned}$$

Based on the monotonically increasing property, $W(\alpha, x) = 0$ has only one positive root denoted as $x = x^*$. Thus,

$$\begin{cases} W(\alpha, x) < 0 & \text{if } 0 < x < x^*; \\ W(\alpha, x) = 0 & \text{if } x = x^*; \\ W(\alpha, x) > 0 & \text{if } x > x^*. \end{cases}$$

Since $\rho\omega(\ln 2)2^{\rho x} > 0$ and $\frac{\partial f(\alpha,x)}{\partial x} = \rho\omega(\ln 2)2^{\rho x}W(\alpha, x)$ (from (A-1)), it can be obtained that

$$\begin{cases} \frac{\partial f(\alpha,x)}{\partial x} < 0 & \text{if } 0 < x < x^*; \\ \frac{\partial f(\alpha,x)}{\partial x} = 0 & \text{if } x = x^*; \\ \frac{\partial f(\alpha,x)}{\partial x} > 0 & \text{if } x > x^*. \end{cases}$$

This completes the proof.

Appendix B

Proof of (3.18)

Proof. For any given $\alpha \in [0, 1)$, from (3.17) and the fact $H(\alpha, z) = H(\alpha, n+1)$, it can be obtained that

$$\frac{\left|\frac{\partial H(\alpha,x)}{\partial \alpha}|_{x=z}\right|}{\left|\frac{\partial H(\alpha,x)}{\partial \alpha}|_{x=n+1}\right|} = \frac{\frac{1-p(\alpha,x)}{\rho\omega(\ln 2)x^2 2^{\rho x} p(\alpha,x)}\Big|_{x=n+1}}{\frac{1-p(\alpha,x)}{\rho\omega(\ln 2)x^2 2^{\rho x} p(\alpha,x)}\Big|_{x=z}} \tag{B-1}$$

where $\rho = \frac{L}{\alpha T}$. Since $H(\alpha, z) = H(\alpha, n+1)$, then $(1 - p(\alpha, z))^z = (1 - p(\alpha, n+1))^{n+1}$, which leads to $\frac{z}{n+1} = \frac{\ln(1-p(\alpha,n+1))}{\ln(1-p(\alpha,z))}$. Thus, (B-1) can be rewritten as

$$\frac{\left|\frac{\partial H(\alpha,x)}{\partial \alpha}|_{x=z}\right|}{\left|\frac{\partial H(\alpha,x)}{\partial \alpha}|_{x=n+1}\right|} = \frac{\frac{(1-p(\alpha,x))\ln(1-p(\alpha,x))}{\rho\omega(\ln 2)x 2^{\rho x} p(\alpha,x)}\Big|_{x=n+1}}{\frac{(1-p(\alpha,x))\ln(1-p(\alpha,x))}{\rho\omega(\ln 2)x 2^{\rho x} p(\alpha,x)}\Big|_{x=z}} = \frac{\left|\frac{A(\alpha,n+1)}{B(\alpha,n+1)}\right|}{\left|\frac{A(\alpha,z)}{B(\alpha,z)}\right|} \tag{B-2}$$

where $A(\alpha, x) \triangleq \frac{\left(1-p(\alpha,x)\right)\ln\left(1-p(\alpha,x)\right)}{p(\alpha,x)}$ and $B(\alpha, x) \triangleq \rho\omega(\ln 2)x2^{\rho x}$.

It can be seen that the $f(\alpha, x)$ in (3.7) can be expressed as $f(\alpha, x) = A(\alpha, x) + B(\alpha, x)$. In the proof of Lemma 1, it has been shown that $f(\alpha, x)$ has the following property:

$$\begin{cases} f(\alpha, x) < 0 & \text{if } 0 < x < x_0(\alpha); \\ f(\alpha, x) = 0 & \text{if } x = x_0(\alpha); \\ f(\alpha, x) > 0 & \text{if } x > x_0(\alpha). \end{cases}$$

For $x \in (0, +\infty)$, it is obvious that $p(\alpha, x) \in (0, 1)$, $A(\alpha, x) < 0$, and $B(\alpha, x) > 0$.

Thus, from the above property of $f(\alpha, x)$, it can be shown that

$$\begin{cases} |A(\alpha,x)| > |B(\alpha,x)| & \text{if } 0 < x < x_0(\alpha); \\ |A(\alpha,x)| = |B(\alpha,x)| & \text{if } x = x_0(\alpha); \\ |A(\alpha,x)| < |B(\alpha,x)| & \text{if } x > x_0(\alpha). \end{cases}$$

For $0 < z < x_0(\alpha) < n+1$, the following can be obtained:

$$\left|\frac{A(\alpha,z)}{B(\alpha,z)}\right| > 1 > \left|\frac{A(\alpha,n+1)}{B(\alpha,n+1)}\right|.$$

Together with (B-2) it can be concluded that

$$\Big|\frac{\partial H(\alpha,x)}{\partial \alpha}|_{x=z}\Big| < \Big|\frac{\partial H(\alpha,x)}{\partial \alpha}|_{x=n+1}\Big|.$$

This completes the proof. $\square$

Appendix C

Proof of property (4.34)

Proof. To prove (4.34), the following property of $a(x)$ is obtained first.

(1) For $x \in (0,1)$, $\frac{da(x)}{dx} = \frac{1}{x(\ln x)^2} > 0$, then $a(x)$ is monotonically increasing;

(2) $a(x)$ is concave in $(0, \frac{1}{e^2})$ and is convex in $(\frac{1}{e^2}, 1)$;

(3) $a(\frac{1}{e}) = 1$ and $\frac{da(x)}{dx}|_{x=1/e} = e$, then the tangent of $a(x)$ at the point $(\frac{1}{e}, 1)$ is $Z_2(x) = ex$;

(4) $\lim_{x\to 0} a(x) = \lim_{x\to 0} \frac{-1}{\ln x} = 0$.

(5) For any x_1, x_2 and λ that satisfy $x_1 < \frac{1}{e} < x_2$, $0 < \lambda < 1$, and $\lambda x_1 + (1-\lambda)x_2 > \frac{1}{e}$, the following holds: $a(\lambda x_1 + (1-\lambda)x_2) < \lambda a(x_1) + (1-\lambda)a(x_2)$.

Property (5) can be proved as follows. Let $x_1 \in (0, \frac{1}{e})$, $x_2 \in (\frac{1}{e}, 1)$ and $x_\lambda = \lambda x_1 + (1-\lambda)x_2$ where $\lambda \in (0,1)$. It can be seen that if $x_\lambda > \frac{1}{e}$, then $x_\lambda \in (\frac{1}{e}, x_2)$. Denote the line determined by the two points $(x_1, a(x_1))$ and $(x_2, a(x_2))$ as $Z_1(x)$, where $Z_1(x_1) = a(x_1)$ and $Z_1(x_2) = a(x_2)$. Then it can be obtained that

$$\frac{Z_1(x_\lambda) - Z_1(x_1)}{Z_1(x_\lambda) - Z_1(x_2)} = \frac{x_\lambda - x_1}{x_\lambda - x_2} = \frac{\lambda - 1}{\lambda} \tag{C-1}$$

Then, it can be seen that $Z_1(x_\lambda) = \lambda Z_1(x_1) + (1-\lambda)Z_1(x_2) = \lambda a(x_1) + (1-\lambda)a(x_2)$. To prove property (5) is equivalent to prove $a(x) < Z_1(x)$ for all $x \in (\frac{1}{e}, x_2)$. Based on the property (1)-(4), it is easy to check that $Z_1(x_1) = a(x_1) > Z_2(x_1)$ and $Z_1(x_2) = a(x_2) > Z_2(x_2)$, which indicates line $Z_1(x)$ is above line $Z_2(x)$ in (x_1, x_2). Since $x_1 < \frac{1}{e} < x_2$, it is obvious that $Z_1(\frac{1}{e}) > Z_2(\frac{1}{e}) = a(\frac{1}{e})$.

According to the monotonically increasing property of $a(x)$, it can be obtained that $Z_1(x_2) = a(x_2) > a(x_1) = Z_1(x_1)$, then $Z_1(x)$ is also an increasing function. Since $a(x)$ is a convex and increasing function in $(\frac{1}{e^2}, 1)$, from $Z_1(\frac{1}{e}) > a(\frac{1}{e})$ and $Z_1(x_2) = a(x_2)$, based on the monotonically increasing property of $Z_1(x)$, it can be concluded that $a(x) < Z_1(x)$ within $(\frac{1}{e}, x_2)$.

Then, based on the above properties, (4.34) can be proved using induction. First, according to property (2) and Jensen's inequality, it can be seen that (4.34) holds if $b_t > 1/e^2 \; \forall \, t = 1, 2, ..., n_i$. Next, consider the case that $b_t < 1/e^2$ for some t.

For $n_i = 2$, let $x_1 = b_1 < 1/e^2 < 1/e$, $x_2 = b_2 > 1/e$ and $\lambda = 1/2$, then $\lambda x_1 + (1-\lambda)x_2 = \frac{1}{2}(b_1 + b_2) = \bar{b} > 1/e$. From property (5), it can be obtained that

$$a(\bar{b}) = a\left(\frac{b_1 + b_2}{2}\right) < \frac{a(b_1) + a(b_2)}{2}. \tag{C-2}$$

Assume for $n_i = n$, the following holds:

$$a\left(\frac{\sum_{t=1}^{n} b_t}{n}\right) < \frac{\sum_{t=1}^{n} a(b_t)}{n}. \tag{C-3}$$

For $n_i = n + 1$, let $b_{n+1} < 1/e^2 < 1/e$, it can be shown that

$$a(\bar{b}) = a\left(\frac{\sum_{t=1}^{n+1} b_t}{n+1}\right) = a\left(\frac{1}{n+1} b_{n+1} + \frac{n}{n+1} \cdot \frac{\sum_{t=1}^{n} b_t}{n}\right). \tag{C-4}$$

Since $\bar{b} > 1/e$, then $(n+1)\bar{b} = \sum_{t=1}^{n+1} b_t > n/e + b_{n+1}$, which leads to $\sum_{t=1}^{n} b_t/n > 1/e$. Let $x_1 = b_{n+1}$, $x_2 = \sum_{t=1}^{n} \hat{b}_t/n$, and $\lambda = 1/(n+1)$, from property (5), it can be obtained that

$$a\left(\frac{1}{n+1} b_{n+1} + \frac{n}{n+1} \cdot \frac{\sum_{t=1}^{n} b_t}{n}\right) < \frac{1}{n+1} a(b_{n+1}) + \frac{n}{n+1} a\left(\frac{\sum_{t=1}^{n} b_t}{n}\right). \tag{C-5}$$

Since $a\left(\frac{\sum_{t=1}^{n} b_t}{n}\right) < \frac{\sum_{t=1}^{n} a(b_t)}{n}$, it can be obtained that

$$a(\bar{b}) = a\left(\frac{\sum_{t=1}^{n+1} b_t}{n+1}\right) < \frac{1}{n+1} a(b_{n+1}) + \frac{n}{n+1} \cdot \frac{\sum_{t=1}^{n} a(b_t)}{n} = \frac{\sum_{t=1}^{n+1} a(b_t)}{n+1}. \tag{C-6}$$

Thus, (4.34) also holds for $n_i = n + 1$. Based on induction, it can be concluded that (4.34) holds for all $n_i > 0$. This completes the proof. □

References

[1] U. Varshney, "Multicast over wireless networks," *Commun. ACM*, vol. 45, no. 12, pp. 31–37, Dec. 2002.

[2] de Morais, H. Gossain, and D. P. Agrawal, "Multicast over wireless mobile ad hoc networks: present and future directions," *IEEE Network*, vol. 17, no. 1, pp. 52–59, Jan.-Feb. 2003.

[3] R. Bhatia, L. Li, H. Luo, and R. Ramjee, "ICAM: Integrated cellular and ad hoc multicast," *IEEE Trans. Mobile Comput.*, vol. 5, no. 8, pp. 1004–1015, Aug. 2006.

[4] J. She, P. Ho, and L. Xie, "IPTV over WiMax: Key success factors, challenges, and solutins," *IEEE Commun. Mag.*, vol. 45, no. 8, pp. 87–93, Aug. 2007.

[5] E. H. S. Kota, Y. Qian and R. Ganesh, "Advances in mobile multimedia networking and QoS," *IEEE Commun. Mag.*, vol. 45, no. 8, pp. 52–53, Aug. 2007.

[6] 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 Commun.*, vol. 15, no. 3, pp. 40–47, June 2008.

[7] V. Srinivasan, P. Nuggehalli, C. F. Chiasserini, and R. R. Rao, "Cooperation in wireless ad hoc networks," *in Proc. IEEE INFOCOM'03*, pp. 808–817, 2003.

[8] A. Majumdar, D. Sachs, K. R. I. Kozintsev, and M. Yeung, "Multicast and unicast real-time video streaming over wireless LANs," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, no. 6, pp. 524–534, June 2002.

[9] W. Ge, J. Zhang, and X. Shen, "A cross-layer design approach to multicast in wireless networks," *IEEE Trans. Wireless Commun.*, vol. 6, no. 3, pp. 1063–1071, Mar. 2007.

[10] P. Chaporkar and S. Sarkar, "Wireless multicast: theory and approaches," *IEEE Trans. Info. Theory*, vol. 51, no. 6, pp. 1954–1972, June 2005.

[11] J. Metzner, "An improved broadcast retransmission protocol," *IEEE Trans. Commun.*, vol. 32, no. 6, pp. 679–683, June 1984.

[12] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity-Part I: System description," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1927–1938, Nov. 2003.

[13] ——, "User cooperation diversity-Part II: Implementation aspects and performance analysis," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1939–1948, Nov. 2003.

[14] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Info. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.

[15] P. Liu, Z. Tao, Z. Lin, E. Erkip, and S. Panwar, "Cooperative wireless communications: A cross-layer approach," *IEEE Wireless Commun.*, vol. 13, no. 4, pp. 84–92, Aug. 2006.

[16] S. Raza, D. Li, C.-N. Chuah, and G. Cheung, "Cooperative peer-to-peer repair for wireless multimedia broadcast," in *Proc. IEEE International Conference on Multimedia and Expo (ICME)*, pp. 1075–1078, 2007.

[17] X. Liu, S. Raza, C.-N. Chuah, and G. Cheungs, "Network coding based cooperative peer-to-peer repair in wireless ad-hoc networks," in *Proc. IEEE International Conference on Communications (ICC)*, pp. 2153–2158, 2008.

[18] J. Jin and B. Li, "Cooperative multicast scheduling with random network coding in WiMAX," in *Proc. 17th International Workshop on Quality of Service (IWQoS)*, 2009.

[19] P. Fan, C. Zhi, C. Wei, and K. B. Letaief, "Reliable relay assisted wireless multicast using network coding," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 5, pp. 749–762, June 2009.

[20] A. del Coso, O. Simeone, Y. Bar-ness, and C. Ibars, "Outage capacity of two-phase space-time coded cooperative multicasting," in *Proc. 40th Asilomar Conf. Signals, Systems and Computers (ACSSC)*, pp. 666–670, 2006.

[21] G. Jakllari, S. V. Krishnamurthy, M. Faloutsos, and P. V. Krishnamurthy, "On broadcasting with cooperative diversity in multi-hop wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 2, pp. 484–496, Feb. 2007.

[22] B. Sirkeci-Mergen and A. Scaglione, "On the power efficiency of cooperative broadcast in dense wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 2, pp. 497–507, Feb. 2007.

[23] O. Alay, T. Korakis, Y. Wang, E. Erkip, and S. Panwar, "Layered wireless video multicast using omni-directional relays," in *Proc. IEEE International Conference on Acoustics, Speach and Signal Processing (ICASSP)*, pp. 2149–2152, 2008.

[24] Q. Zhang, F. H. P. Fitzek, and V. B. Iversen, "Design and performance evaluation of cooperative retransmission scheme for reliable multicast services in cellular controlled p2p networks," in *Proc. 21st Annual IEEE International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*, 2007.

[25] O. Alay, P. Liu, Z.Guo, Y. Wang, E. Erkip, and S. Panwar, "Cooperative layered video multicast using randomized distributed space time codes," in *Proc. IEEE INFOCOM'09*, pp. 176–181, 2009.

[26] D. Nguyen, T. Nguyen, and B. Bose, "Wireless broadcast using network coding," *IEEE Trans. Veh. Technol.*, vol. 58, no. 2, pp. 914–925, Feb. 2009.

[27] Y. W. Hong and A. Scaglione, "Energy-efficient broadcasting with cooperative transmissions in wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 5, no. 10, pp. 2844–2855, Oct. 2006.

[28] I. Maric and R. D. Yates, "Cooperative multicast for maximum network lifetime," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 1, pp. 127–135, Jan. 2005.

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

[30] S. Zhong, J. Chen, and Y. R. Yang, "Sprite: A simple, cheat-proof, credit-based system for mobile ad-hoc networks," in *Proc. IEEE INFOCOM'03*, pp. 1987–1997, 2003.

[31] L. Anderegg and S. Eidenbenz, "Ad hoc-VCG: A truthful and cost-efficient routing protocol for mobile ad hoc networks with selfish agents," in *Proc. 9th Annual International Conference on Mobile Computing and Networking (MobiCom)*, pp. 245–259, 2003.

[32] N. R. Suri and Y. Narahari, "Design of an optimal bayesian incentive compatible broadcast protocol for ad hoc networks with rational nodes," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 7, pp. 1138–1148, Sept. 2008.

[33] S. Marti, T. J. Giuli, K. Lai, and M. Baker, "Mitigating routing misbehavior in mobile ad hoc networks," in *Proc. 6th Annual International Conference on Mobile Computing and Networking (MobiCom)*, pp. 255–265, 2000.

[34] P. Michiardi and R. Molva, "Core: A collaborative reputation mechanism to enforce node cooperation in mobile ad hoc networks," in *Proc. 6th IFIP TC6/TC11 Joint Working Conference on Communications and Multimedia Security*, pp. 107–121, 2002.

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

[36] P. Michiardi and R. Molva, "A game theoretical approach to evaluate cooperation enforcement mechanisms in mobile ad hoc networks," in *Proc. Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt)*, 2003.

[37] W. Yu and K. J. R. Liu, "Game theoretic analysis of cooperation and security in autonomous mobile ad hoc networks," *IEEE Trans. Mob. Comput.*, vol. 6, no. 5, pp. 459–473, May 2007.

[38] ——, "Secure cooperation in autonomous mobile ad-hoc networks under noise and imperfect monitoring: Strategies and game theoretic analysis," *IEEE Trans. Information Forensics and Security*, vol. 2, no. 2, pp. 240–253, June 2007.

[39] Z. Han, Z. Ji, and K. J. R. Liu, "A cartel maintenance framework to enforce cooperation in wireless networks with selfish users," *IEEE Trans. Wireless Commun.*, vol. 7, no. 5, pp. 1889–1899, May 2008.

[40] Y. Wu, B. Wang, K. J. R. Liu, and T. C. Clancy, "Repeated open spectrum sharing game with cheat-proof strategies," *IEEE Trans. Wireless Commun.*, vol. 8, no. 4, pp. 1922–1933, Apr. 2009.

[41] C. Song and Q. Zhang, "Achieving cooperative spectrum sensing in wireless cognitive radio networks," *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 13, no. 2, pp. 14–25, Apr. 2009.

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

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

[44] W. Lin, H. V. Zhao, and K. J. R. Liu, "Incentive cooperation strategies for peer-to-peer live streaming social networks, special section on communities and media computing," *IEEE Trans. Multimedia*, vol. 11, no. 3, pp. 890–900, Apr. 2009.

[45] J. Yang, A. G. Klein, and D. R. Brown, “Natural cooperation in wireless networks,” *IEEE Signal Processing Mag.*, vol. 26, no. 5, pp. 98–106, Sept. 2009.

[46] A. R. Mele and P. Rawling, *The Oxford handbook of rationality*. Oxford University Press, 2004.

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

[48] P. Jehiel, “Bounded rationality and imperfect learning: Game theory vs AI,” in *Transfer of Knowledge in Economic Decision Making*, D. J. Zizzo, Ed. Palgrave Macmillan, 2005.