Chapter 23

Network Selection and Handoff in Wireless Networks: A Game Theoretic Approach

Josephina Antoniou, Vicky Papadopoulou, Vasos Vassiliou, and Andreas Pitsillides

Contents

Convergence, that is, the integration of various access technologies combining their resources to best serve the increased user requirements, may be supported, through a system architecture where different access networks, terminals, and services coexist. Thus, a new, *user centric* communication paradigm is motivated, that is, the user is no longer bound to only one access network but may indirectly *select* the best available access network(s) to support a service session. Upon a new service request or even any dynamic change, for example, mobility, one (or a group) of the participating access networks needs to be selected in order to support the session. Thus, the converged system architecture must be equipped with a *network selection* mechanism to effectively assign the best access network(s) to handle a service session.

23.1 Introduction

Converged networks allow different access networks, terminals, and services to coexist, bringing forth a new communication paradigm, which is *user-centric* [1], that is, the user is no longer bound to only one access network but may indirectly *select* the *best* available access network(s) to support a service session [2]. Upon a new service request or even any dynamic change affecting the session, (e.g., mobility) one, two, or multiple of the participating access networks need to be selected in order to support the session.

Next-generation communication networks—controlled by an IP core network—need to be equipped with a *network selection mechanism* to assign the *best* access network(s) to handle service activation or any dynamic session change. Such decisions may result in one, two, or multiple access networks handling a service (the cooperation of multiple networks to handle a service is treated in [3]). This chapter studies the resulting interaction between a user and a network when a single network is selected and proposes a payment partition between two networks cooperating to support a session, seeking the best behavior for each entity such that their conflicting interests are overcome and satisfaction is achieved.

Since any communication network, such as a converged network [4] considered in this chapter, is a multi-entity system, decisions are taken by different system entities. The decision-making entities in a converged network are (a) the user and (b) the network operator. Both entities are motivated to make decisions that maximize their own *satisfaction* functions [5,6], which are based on each entity's criteria. The criteria for the user include quality of experience (QoE) and cost; on the other hand, the network operator makes decisions that are driven mainly by one criterion: the network's revenue maximization. Therefore, the network operator sets the *price per user session* for using network resources based on its own revenue-maximization criterion, and the user, given the price set by the network as well as some further parameters that comprise its own satisfaction function, makes the decision whether or not to participate in the particular network, always subject to the network's own admission policy, for example [7,8].

We provide a study of the interaction between a user and a network during network selection, and we define the selection process based on the conclusions obtained from the study of the user–network interaction. The selection decision can be modified to result in a list of prioritized available networks instead of a single network. Thus, for users requesting enhanced service with additional quality guarantees, for example, *premium* service, the two *best* networks on the priority list cooperate to provide these guarantees; such additional guarantees include, for example, service continuity during fast session handoffs that may be required during a session. This situation is considered in this chapter, and we utilize the notion of a Nash bargaining solution (NBS) to compute an optimal partition agreement between the two cooperative networks for the service payment.

Therefore, we consider first the interaction between one user and one access network in order to propose a model for the interaction of each user and each one of the available access networks for network selection. Then, our study turns toward the case of the interaction between two networks cooperating to support a *premium* service, to handle fast session handoffs and ensure service continuity throughout the duration of the session. Such interactions between entities with conflicting interests follow action plans designed by each entity in such a way as to achieve a particular selfish goal and are known as strategic interactions. *Game theory* is a theoretical framework that studies strategic interactions, by developing models that prescribe actions in order for the interacting entities to achieve satisfactory gains from the situation. In this chapter, we utilize game theory in order to model, analyze, and finally propose solutions for the selected interactions. The user–network interaction during network selection is modeled as a game between the user and the network, whereas the second game models the cooperation between two users to support premium service and the user is not involved in the game model. Figure 23.1 presents the relation between the two game models representing the earlier-mentioned situations.

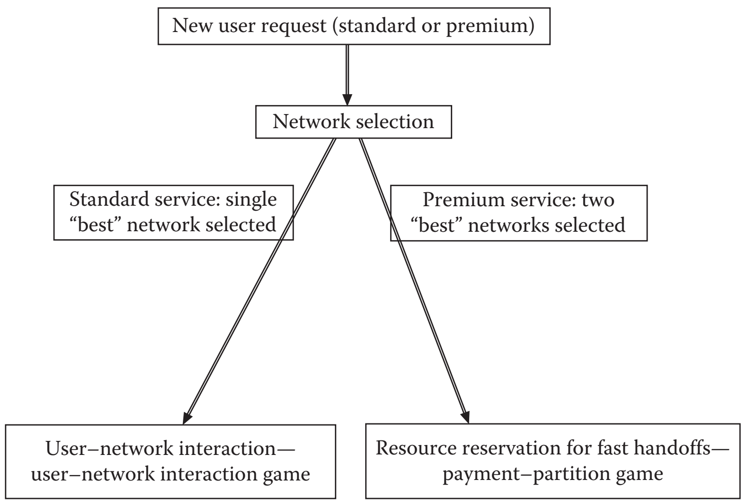

Figure 23.1 Relation between the user–network interaction game and the payment–partition game.

The chapter is organized as follows: Section 23.2 provides an overview of network selection–related research. Section 23.3 defines and analyzes models of the network selection problem when a single network is selected.* In Section 23.4, we introduce an enhanced game, which captures the case where the two best networks cooperate to enable fast session handoffs when the service continuation and the time to handoff are crucial for a specific service.† Next, Section 23.5 provides a simulative evaluation of the strategies for the models presented, and Sections 23.6 and 23.7 offer conclusions for the chapter and a discussion on open research challenges, respectively.

23.2 Related Work

The growing popularity of the next-generation communication networks, which promote technology convergence and allow the coexistence of heterogeneous constituents, for example, various access networks, has pushed toward the efficient resource management (*including network selection*) of the overall converged system.

Earlier works on network selection have explored varying approaches of describing and analyzing this new resource management mechanism. Such approaches included fuzzy logic [10,11], adaptive techniques [1,12], utility-based and game-theoretic models [13,14], technology-specific solutions, especially focusing on the interoperation of cellular systems with wireless LANs [15], as well as architectural models focusing on a more comprehensive architectural view [2]. Decision making in these works is either user controlled [11,15] or network controlled [2,12]; the interaction between them is not really considered and the selection decision mainly involves one entity, in some cases involving the other entity indirectly, that is, network decision considering some user preference [1,14] or user decision considering some network-specific rankings [10,13].

Further, in more recently published research, we observe a popular trend toward both dynamic/automated solutions that are either based on the user or the network as a decision-making entity [16,17], in some cases proposing improvements to the user-perceived quality during network selection [18,19]. Other than the general solutions for network selection, specific solutions handling particular services such as multicast have been recently proposed [20,21].

Given that game theory is a theoretical framework for strategical decision making, it has been a very popular approach among the most recently presented research works. These explore various game-theoretic models such as noncooperative games [22,23] and cooperation schemes where limited resources or need for quality guarantees exists [3,24]. In addition to network selection, game-theoretic models for admission and rate control exist where appropriate admission criteria guard against users leaving the selected network due to, for instance, unsatisfactory service [25]. The game-theoretic study of users having the option to leave their network, that is, churning, is further carried out, in particular, pricing schemes (both noncooperative and cooperative) that maximize the revenues of service providers [26]. Another topic of interest in our work is how to model bargaining situations using game theory. An interesting such work that dives into the two-player bargaining situation, considering a game of incomplete information where players are uncertain about each other's preferences and offer a utility-efficient solution, is found in [27].

AQ1

An issue that arise in game-theoretic models, which we will also address in this chapter, is that of truthfulness. An interesting and very promising way to guarantee truthfulness of the participating

* a detailed study for the case of single network selection is found in [9].

† The case of multiple networks cooperating to support a service request is explored in [3].

networks is through pricing mechanisms [28]. Such mechanisms could penalize a player who turned out to lie, when a revelation of the real value of the particular quantity is possible at a later stage of the procedure. Such a mechanism would enforce the players in a particular game to be truthful. In the worldwide literature, there is a whole research field that is focused on the development, limitations, and capabilities of such pricing mechanisms: the algorithmic mechanism design [28]. Successful paradigms in this context include (combinatorial) auctions [29] and task scheduling [30] using techniques such as the revelation principle [28], incentive compatibility [28], direct revelation [28], and Vickrey–Clarke–Groves mechanisms [31]. We believe that the algorithmic tools and theoretical knowledge that is already developed in the field of algorithmic mechanism design constitute a fruitful pool for extracting algorithmic tools for enforcing players to truthfulness, through pricing mechanisms.

The above-mentioned game-theoretic approaches have inspired our work and the current chapter to explore network selection through a game-theoretic modeling framework. In this chapter, we model aspects of network selection as cooperative game models between interacting entities, capturing the case where decision making is controlled by both the user and the network.

23.3 Single Network Selection

Network selection is an inherently complex decision because it involves considerations of user characteristics, individual access network characteristics, and overall network efficiency. Selecting the network that both satisfies the user and the network may become a challenging task. A careful examination of the user–network interaction prior to defining the proposed network selection model is necessary.

In this section, we consider the interaction between a user and a network, and we seek to investigate the outcomes of their interaction and indicate possible incentives to motivate the cooperation of the two. We first look at the interaction between a user and a network during network selection, to reach a decision that is both user-satisfying and network-satisfying. For a fixed quality level, the interaction may be viewed as an exchange between two entities: the user gives some *compensation* and the network gives a *promise* of the specific quality level. Primarily, we seek the incentives for each entity to select certain strategies, that is, sets of actions, that result in a cooperative selection decision, by which both entities are satisfied; such incentives are usually the payoffs to the entities involved. In the user–network interaction, the payoffs are the following:

- User's payoff: the difference between the *perceived satisfaction* and the compensation offered by the user to the network (for a given quality level).
- Network's payoff: the difference between the compensation received and the cost of supporting the session, for a given quality level. The cost is analogous to the requested quality by the user quality level.

Prior to modeling the user–network interaction, we make the following general assumptions:

1. The players in the game are heterogeneous, aiming at different payoffs.
2. The players make simultaneous decisions, without knowing the decision of their opponent at the time of their own decision.

3. There is a minimum and a maximum payment as well as a minimum and a maximum quality, considered in the user–network interaction model.
4. There is always a probability that quality degradation will be perceived by the user because of the dynamic nature of the network, even if the network decides to cooperate.
5. For a set quality q, it holds that $p(q) - \kappa > 0$, that is, the difference between the maximum satisfaction perceived by the user and the maximum compensation given by the user is greater than 0. Similarly, the difference between the minimum satisfaction and the minimum compensation is greater than 0, that is, $p(q') - \kappa' > 0$.
6. Both the user and the network have nonnegative payoff functions. This assumption is introduced in order to motivate the players to participate in the interaction (reflecting a selection of an appropriate access network during network selection).

The above assumptions result in the following specifications, which are imposed by the game model:

- Assumption 6 results in the requirement that the maximum compensation* offered by the user is less than or equal to the satisfaction corresponding to the minimum requested quality. Thus, given that compensation and satisfaction from perceived quality are measurable in comparable units, the maximum compensation offered by the user should be defined to be less than or equal to the satisfaction corresponding to the minimum requested quality. In this way, the user *plays* the game without risking to have a negative payoff, satisfying Assumption 6.
- Assumption 6 results in the requirement that the minimum compensation offered by the user is greater than or equal to the cost of the maximum requested quality. Thus, given that the compensation and cost of supporting a requested quality are measurable in comparable units, the minimum compensation offered by the user should be defined to be greater than or equal to the cost of the maximum requested quality. In this way, the network *plays* the game without risking to have a negative payoff satisfying Assumption 6.

23.3.1 One-Shot User–Network Interaction Game

We first consider the simpler case of the interaction, where the current actions of the involved entities do not affect future actions, nor are affected by previous ones. Since we aim to extract conclusions on the user–network cooperation, we model this simple scenario as a one-shot game. The game model, considering several quality levels, is illustrated in (Figure 23.2); the subsequent analysis considers the quality level requested to be fixed. The players play only once and in which the current outcomes do not affect any future outcomes of the game. In particular, we assume that the user and network interaction is as follows: the user decides on a specific quality level to request and proposes to the network a corresponding compromise for that particular quality level. Then the network decides to accept it or not. In case of acceptance,

* Regarding, setting the maximum compensation, if a telecommunications regulator exists, as currently in Europe, the regulator takes the role of setting this maximum compensation, otherwise the platform administrator will set the maximum compensation to encourage the user to remain a customer. The regulator or the administrator has information that prevents exploitation of the users by the operators and based on that information the maximum compensation is set.

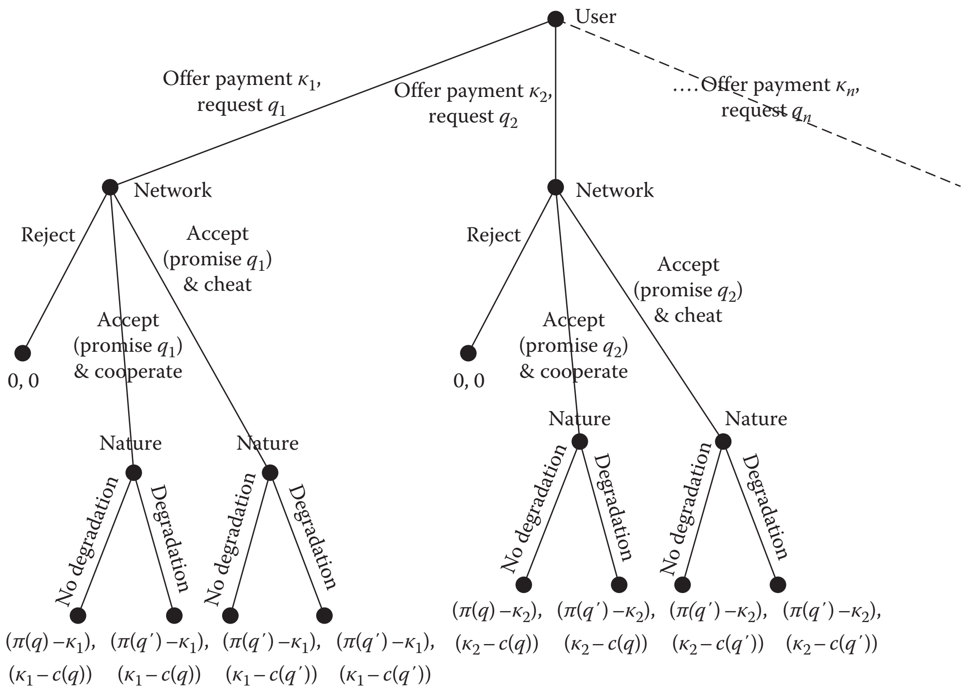

Figure 23.2 The one shot sequential moves game of user–network interaction in extensive form.

the network may then decide to provide the promised quality level or to cheat. Finally, we assume that there exists a random event of quality degradation provided by the user, even in the case of deciding not to cheat. More formally, the user–network interaction game is defined as follows:

Definition 23.1 (The user–network interaction game). Let the user choose an incentive compensation scheme $\kappa_i \in K$ to offer to the network, where $K = \{\kappa_1, \kappa_2, \ldots \kappa_n\}$, in order to encourage the network to make and keep a *promise* of offering $q_i \in Q$, where $Q = \{q_1, q_2, \ldots q_n\}$. The user has the capability to cheat and offer κ_i', such that $\forall\ i \in [n], \kappa_i' < \kappa_i$, but this will not be known to the network until the end of the game when payoffs are collected. The network makes a decision of whether to accept the user's compensation and promise to offer q_i or to reject it. If the network rejects the offer, the interaction terminates and the payoff to both players is 0. If the network accepts the offer, it can decide to keep the promise of supporting q_i or to cheat, by offering q_i', such that $\forall\ i \in [n], q_i' < q_i$. This will be known to the user also at the end of the game, when payoffs are collected.

Regardless of the network's decision, there is a random event, represented by *Nature*, that is, an event that is not controlled by neither the network nor the user. This random event has two outcomes: the quality offered by the network degrades (the user perceives a lower quality $q_i' < q_i$), or the perceived quality indicates no degradation (quality q_i perceived by the user). A quality degradation might be observed by the user, even when the network takes all the necessary measures to keep the promise for the requested quality. Whether the network cheats, or *nature's* event is

Table 23.1 User and Network Play the User–Network Interaction Game, Nature's Play: No Degradation

	Network Accepts and Cooperates	*Network Accepts and Cheats*	*Network Rejects*
User cooperates	$\pi(q) - \kappa, \kappa - c(q)$	$\pi(q') - \kappa, \kappa - c(q')$	0, 0
User cheats	$\pi(q) - \kappa', \kappa' - c(q)$	$\pi(q') - \kappa', \kappa' - c(q')$	0, 0

one of degradation, or whether both occur, the user perceives a quality q_i' without knowing what caused it.

In the rest of the section, we assume a fixed requested quality q and corresponding compensation κ, so we simplify κ_i to κ and q_i to q. Then, the payoffs of the two players are given in Table 23.1, assuming the *nature* outcome is no degradation; for the case of degradation caused by cheating, q is replaced with q'. Let $\pi(q)$ represent the satisfaction resulting from the perceived quality q and $c(q)$ represent the cost for supporting quality q. The payoffs are presented by order pairs: the first term is the user payoff and the second term is the network payoff.

23.3.1.1 Equivalence to Prisoner's Dilemma

Definition 23.2 (Prisoner's dilemma equivalent game) [32]. Consider a one-shot strategic game with two players in which each player has two possible actions: to cooperate with his opponent or to defect from cooperation. Furthermore, assume that the two following additional restrictions on the payoffs are satisfied:

1. The payoffs of the two players for each possible outcome of the game are shown in Table 23.2. For each player $j \in [2]$ and is such that $A_j > B_j > C_j > D_j$.
2. The reward for mutual cooperation is such that each player is not motivated to exploit his opponent or be exploited with the same probability, that is, for each player it holds that $B_j > (A_j + D_j)/2$.

Then, the game is said to be equivalent to a prisoner's dilemma type of game.

Consider now the following outcome for the user–network interaction game: the network accepts the user's offer and the outcome of nature's random event is *no degradation*. We prove that this outcome of the game is equivalent to a *prisoner's dilemma type of game*. In the subsequent

Table 23.2 General Payoffs for the Prisoner's Dilemma

	Player 2 Cooperates	*Player 2 Cheats*
Player 1 cooperates	B_1, B_2	D_1, A_2
Player 2 cheats	A_1, D_2	C_1, C_2

discussion, we recall Assumptions 5 and 6 from the general assumptions and we generate the following requirements for the user–network interaction game:

1. The difference between the maximum and minimum satisfaction, $\pi(q) - \pi(q')$, is greater than the diifference between the maximum and minimum compensation, $\kappa - \kappa'$, that may be offered by the user, since the user will always prefer that the maximum compensation offered is set as low as it is allowed by the platform administrator, whereas the satisfaction received from the maximum quality may be evaluated as high as possible, since no such constraints exist.
2. The network overall payoff (*compensation cost*) is always positive if accepting, that is, $\kappa - \kappa' > c(q) - c(q')$.

We now show the equivalence between the user–network interaction game and the prisoner's dilemma game:

Proposition 23.1 Consider the user–network interaction game. Assume that the network accepts the user's offer, and the event of nature is no degradation. Then the game is equivalent to a prisoner's dilemma game.

Proof By Definition 23.1 we observe that

Observation 23.1 Both the network and the user have two possible actions: to cooperate and to cheat (i.e., defect from cooperation).

Observation 23.1 combined with Definition 23.2 imply that the actions of the players in the user–network interaction game are the same as the actions of the players of a prisoner's dilemma game. Furthermore, consider the mapping of each player's payoffs (for the cases that the network accepts the user's offer), shown in Table 23.1, to payoffs A_j, B_j, C_j, D_j, where $j \in [2]$, as defined in Definition 23.2. We proceed to prove: AQ2

LEMMA 23.1 Set A_j, B_j, C_j, D_j according to Table 23.3 in the user–network interaction game. Then condition 1 of Definition 23.2 is satisfied.

Table 23.3 The Mapping between the User and Network Payoffs and the Payoffs in the Prisoner's Dilemma Type of Game

	User Payoffs ($j = 1$)	*Network Payoffs* ($j = 2$)
A_j	$\pi(q) - \kappa'$	$\kappa - c(q')$
B_j	$\pi(q) - \kappa$	$\kappa - c(q)$
C_j	$\pi(q') - \kappa'$	$\kappa' - c(q')$
D_j	$\pi(q') - \kappa$	$\kappa' - c(q)$

Proof We prove the claim by showing that $A_j > B_j > C_j > D_j$ for each $j \in [2]$ as required by condition 1 of Definition 23.2, when the network accepts the user's offer and no degradation occurs.

Firstly, consider the user. We verify straightforward that $\pi(q) - \kappa' > \pi(q) - \kappa$; thus, $A_1 > B_1$, and that $\pi(q') - \kappa' > \pi(q') - \kappa$; thus, $C_1 > D_1$, since $\kappa > \kappa'$. Given that $\pi(q) - \pi(q') > \kappa - \kappa'$, it holds that $B_1 > C_1$. It follows that $A_1 > B_1 > C_1 > D_1$ as required by condition 1 of Definition 23.2.

Consider now the network. We verify straightforward that $\kappa - c(q') > \kappa - c(q)$; thus, $A_2 > B_2$, and that $\kappa' - c(q') > \kappa' - c(q)$, since $c(q) > c(q')$; thus, $C_2 > D_2$. Assuming that the network accepts to participate in the interaction in a riskless manner, that is, only if the range of possible compensations exceeds the range of possible costs, then $\kappa - \kappa' > c(q) - c(q')$, and $B_2 > C_2$. It follows that $A_2 > B_2 > C_2 > D_2$ as required by condition 1 of Definition 23.2. □

We now proceed to prove that

LEMMA 23.2 Set A_j, A_j, B_j, C_j, D_j for each $j \in [2]$ as in Table 23.3. Then, the user–network interaction game satisfies condition 2 of Definition 23.2.

Proof To prove the claim, we must prove that the reward for cooperation is greater than the payoff for the described situation, that is, for each player it must hold that $B_j > \frac{A_j + D_j}{2}$.

For the user,

$$\pi(q) - \kappa > \frac{\pi(q) - \kappa' + \pi(q') - \kappa}{2} > \frac{\pi(q) + \pi(q')}{2} - \frac{\kappa' + \kappa}{2} \tag{23.1}$$

since $\pi(q) > \pi(q')$ and $\kappa > \kappa'$.

For the network,

$$\kappa - c(q) > \frac{\kappa - c(q') + \kappa' - c(q)}{2} > \frac{\kappa + \kappa'}{2} - \frac{c(q) + c(q')}{2} \tag{23.2}$$

Setting A_2, B_2, D_2 as shown in Table 23.3, we get that $B_j > \frac{A_j + D_j}{2}$ as required by condition 2 of Definition 23.2. □

Observation 23.1, Lemmas 23.1, and 23.2 together complete the proof of Proposition 23.1. □

The decision of each player in the user–network interaction game is based on the following reasoning [32]: If the opponent cooperates, defect to maximize payoff (A_j *in Table 23.3*); if the opponent defects, defect (*payoff* C_j *instead of payoff* D_j). This reasoning immediately implies

---

COROLLARY 23.1 [32] In a one-shot prisoner's dilemma game, a best-response strategy of both players is to defect.

Proposition 23.1 combined with Corollary 23.1 immediately implies

---

COROLLARY 23.2 In the one-shot user–network interaction game, when the event of nature is one of no degradations, a best-response strategy of both players is to cheat.

Given this result, we investigate next how the behavior of both the user and the network changes when the interaction between the two entities takes into account the history of their interaction.

23.3.2 Repeated User–Network Interaction Game

Here, we capture the fact that the interactions between networks and users do not occur only once but are recurring. In such relationships, the players do not only seek the immediate maximization of payoffs but instead the long-run optimal payoff. Such situations are modeled in game theory by finite horizon and infinite horizon repeated game models [33, chapter 11]. We model the user–network interaction model as an infinite horizon repeated game, since the number of new session requests is not known.

Well-known strategies in repeated game are *trigger strategies* [36, chapter 6], that is, strategies that change in the presence of a predefined trigger. A popular trigger strategy is the *grim strategy* [33, chapter 11], which dictates that the player participates in the relationship, but if *dissatisfied*, leaves the relationship forever. Another popular strategy used to elicit cooperative performance from an opponent is for a player to mimic the actions of his opponent, giving him the incentive to play cooperatively, since in this way he will be rewarded with a similar mirroring behavior. This strategy is referred to as *tit-for-tat* strategy [33, chapter 11].

23.3.2.1 Present Value

In order to compare different sequences of payoffs in repeated games, we utilize the idea of the *present value of a payoff sequence* [33, chapter 11], and we refer to it as the *present value* (PV). PV is the sum that a player is willing to receive now as payoff instead of waiting for future payoffs. Consider the interaction of the network and the user, which takes into account their decisions in the previous periods of their interaction. Let r be the rate by which the current payoffs increase in the next period, for example, in terms of satisfaction since the longer the user and network cooperate the more the satisfaction is for both players. Therefore, if the user (*or the network*) were set to receive a payoff equal to 1 in the next period, today the current payoff the user (*or the network*) would be willing to accept would be equal to $1/(1+r)$.

Since in each game repetition it is possible that degradation may be perceived, there is always a probability p that the game will not continue in the next period. Then, the probability of no degradation is $1-p$, and the payoff the user (*or the network*) is willing to accept today, that is, its PV, would be equal to $(1-p)/(1+r)$. Let $\delta = (1-p)/(1+r)$, where $\delta \in [0,1]$ and often referred to as the *discount factor* in repeated games [33, chapter 11]. In order to determine whether cooperation is a better strategy in the repeated game for both the network and the user, we utilize PV and examine for which values of δ, cooperation is a player's best-response strategy to the other player's strategy.

23.3.2.2 Equilibria

We first provide some useful strategies for the repeated game.

Definition 23.3 (Cheat-and-leave strategy). When the user cheats in one period and leaves in the next period to avoid punishment, the strategy is called the cheat-and-leave strategy.

Definition 23.4 (Cheat-and-return strategy). When the network cheats in one period and returns to cooperation in the next period, the strategy is called the cheat-and-return strategy.

We are now ready to introduce a repeated game modeling the user–network interaction when the history is taken into account in the decisions of the entities:

Definition 23.5 (Repeated user–network interaction game). Consider a game with infinite repetitions of the one-shot user–network interaction game, where $p \in [0, 1]$ is the probability of degradation, and r is the rate of satisfaction gain of continuing cooperation in the next period. Let PVs of each player be calculated after a history, that is, record of all past actions that the players made [33, chapter 11], of cooperation from both players in terms of a discount factor $\delta = (1 - p)/(1 + r)$.

The user has a choice between the two following strategies: (i) the *grim* strategy, that is, offer a compensation κ but if degradation is perceived, then leave the relationship forever, and (ii) the *cheat-and-leave* strategy. The network has a choice between (a) the *tit-for-tat* strategy, that is, mimic the actions of its opponent, and (b) the *cheat-and-return* strategy.

The payoffs from each iteration are equal to the payoffs from the one-shot user–network interaction game. The cumulative payoffs (or PVs) for each player from the repeated interaction are equal to the sum of the player's payoffs in all periods; thus, the number of periods, for example, infinite, should be considered.

Then, the game is referred to as repeated user–network interaction game.

A sequence of strategies with which in each iteration a player plays its best-response strategy, that is, giving the player the highest payoff, to the opponent's strategy after every history is called a *subgame perfect strategy* for the player [33, chapter 11]. To have a subgame perfect strategy, we must show that for every possible iteration of the game, the current choice of each player results in the highest payoff, against all possible actions of the opponent player. In the repeated user–network interaction game, we consider a history of cooperation from both players. Hence, the best-response strategy of each player against all possible actions of the opponent, in the current period, is given in terms of PV assuming such a cooperation history. When both players play their *subgame perfect strategies*, the profile of the game is an equilibrium in the repeated game, known as a *subgame perfect equilibrium* [33, chapter 11].

When neither of the two players cheats, the sequence of profiles is defined more formally next:

Definition 23.6 (Conditional-cooperation profile). When the user employs the grim strategy and the network employs the tit-for-tat strategy, the profile of the repeated game is referred to as conditional-cooperation profile of the game.

The following theorem identifies the *conditional-cooperation* profile as a subgame perfect equilibrium for the repeated user–network interaction game.

THEOREM 23.1 In the repeated user–network interaction game, assume $\delta > (c(q) - c(q'))/(\kappa - c(q'))$ and $\delta > \frac{\kappa-\kappa'}{\pi(q)-\kappa'}$. Then, the conditional-cooperation profile is a subgame perfect equilibrium for the game.

Proof We assume a history of cooperative moves in the past. We compute the PVs of both the user and the network, and after comparing them, we conclude that the conditional-cooperation profile is a subgame perfect equilibrium.

1. Assume first that the user plays the grim strategy. If the network plays the tit-for-tat strategy, it will cooperate in the current period, whereas, if it plays the cheat-and-return strategy, it may cheat.

If the network cooperates, then

$$PV_{coop}^{net} = \frac{\kappa - c(q)}{1 - \delta}$$

If the network cheats, then

$$PV_{cheat}^{net} = \kappa - c(q') + \frac{\delta \cdot 0}{1 - \delta}$$

For the network to be motivated to cooperate, its PV in case of cooperation must be preferable than its PV in case of cheating. Thus

$$PV_{coop}^{net} > PV_{cheat}^{net} = \frac{\kappa - c(q)}{1 - \delta} > \kappa - c(q') + \frac{\delta \cdot 0}{1 - \delta}$$

If the user plays the grim strategy, the network is motivated to cooperate when $\delta > (c(q) - c(q'))/(\kappa - c(q'))$.

2. Assume now that the user plays the cheat-and-leave strategy. Considering the network's possible strategies, it could either cooperate or cheat in the current period.

If the network cooperates, then

$$PV_{coop}^{net} = \kappa' - c(q) + \frac{\delta \cdot 0}{1 - \delta}$$

If the network cheats, then

$$PV_{cheat}^{net} = \kappa' - c(q') + \frac{\delta \cdot 0}{1 - \delta}$$

If the user plays the cheat-and-leave strategy, the network is not motivated to cooperate since $PV_{cheat}^{net} > PV_{coop}^{net}$

3. Assume now that the network plays the tit-for-tat strategy. If the user plays the grim strategy, it will cooperate in the current period, whereas, if it plays the cheat-and-leave strategy, it may cheat.

If the user cooperates, then

$$PV_{coop}^{user} = \frac{\pi(q) - \kappa}{1 - \delta}$$

If the user cheats, then

$$PV_{cheat}^{user} = \pi(q) - \kappa' + \frac{\delta \cdot 0}{1 - \delta}$$

For the user to be motivated to cooperate, its PV in case of cooperation must be preferable than its PV in case of cheating. Thus

$$PV^{user}_{coop} > PV^{user}_{cheat} = \frac{\pi(q) - \kappa}{1 - \delta} > \pi(q) - \kappa' + \frac{\delta \cdot 0}{1 - \delta}$$

If the network cooperates, the user is motivated to cooperate when $\delta > (\kappa - \kappa')/(\pi(q) - \kappa')$.

4. Assume finally that the network plays the cheat-and-return strategy. Considering the user's possible strategies, it could either cooperate or cheat in the current period.

If the user cooperates, then

$$PV^{user}_{coop} = \pi(q') - \kappa + \frac{\delta \cdot 0}{1 - \delta}$$

If the user cheats, then

$$PV^{user}_{cheat} = \pi(q') - \kappa' + \frac{\delta \cdot 0}{1 - \delta}$$

If the network plays the cheat-and-return strategy, the user is not motivated to cooperate since $PV^{user}_{cheat} > PV^{user}_{coop}$.

It follows that the conditional-cooperation profile is a subgame perfect equilibrium. □

23.3.3 Introducing Punishment in the Repeated Interaction

We now investigate the repeated user–network interaction game, where we introduce punishment for cheating behavior. The game is modified as follows: the user may employ a strategy such that the punishment imposed on the network for cheating lasts only for one period; namely, let the user be allowed to employ the *leave-and-return* strategy as defined next:

Definition 23.7 (Leave-and-return strategy). When the user cooperates as long as the network cooperates and leaves for one period in case the network cheats, returning in the subsequent period to cooperate again, the user's strategy is called leave-and-return strategy.

Based on the newly defined strategy, we introduce the *one-period punishment* profile of the game.

Definition 23.8 (One-period-punishment profile). When the user employs the leave-and-return strategy and the network employs the tit-for-tat strategy, the profile of the repeated game is called one-period-punishment profile of the game.

In [34] it was proved that the conditions to sustain cooperation with grim trigger strategies, which are the stricter strategies that may be employed in a repeated prisoner's dilemma, are necessary conditions for the feasibility of any form of conditional cooperation. That is, a grim trigger strategy can sustain cooperation in the iterated prisoner's dilemma under the least favorable circumstances of any strategy that can sustain cooperation.

Motivated by the result in [34], we show next that it is easier to impose cooperation in our game under the *conditional-cooperation* profile than under the *one-period-punishment* profile.

THEOREM 23.2 Assume that $\delta > (c(q) - c(q'))/(\kappa - c(q'))$ and $\delta > (\kappa - \kappa')/(\pi(q) - \kappa')$ in the repeated user–network interaction game. Then, the conditional-cooperation profile motivates cooperation of the players. The same conditions on δ are also necessary to motivate cooperation in the one-period-punishment profile.

Proof We first show that cooperation is motivated under the one-period punishment profile. Given a history of cooperation, we seek the values of δ that can motivate cooperation under the *one-period-punishment* profile.

Assume first that the user cooperates in the current period. Then, the network has two options: to cooperate or to cheat.

If the network cooperates, then

$$PV_{coop}^{net} = \kappa - c(q) + \delta \cdot (\kappa - c(q)) + \frac{\delta^2 \cdot (\kappa - c(q))}{1 - \delta}$$

If the network cheats, then

$$PV_{cheat}^{net} = \kappa - c(q') + \delta \cdot 0 + \frac{\delta^2 \cdot (\kappa - c(q))}{1 - \delta}$$

In order for cooperation to be motivated, it must be that

$$PV_{coop}^{net} > PV_{cheat}^{net} = \kappa - c(q) + \delta \cdot (\kappa - c(q)) + \frac{\delta^2 \cdot (\kappa - c(q))}{1 - \delta}$$
$$> \kappa - c(q') + \delta \cdot 0 + \frac{\delta^2 \cdot (\kappa - c(q))}{1 - \delta}$$

Simplifying, we get $\delta > (c(q) - c(q'))/(\kappa - c(q))$.

Now, assume that the network cooperates in the current period. Then, the user has two options: to cooperate or to cheat.

If the user cooperates, then

$$PV_{coop}^{user} = \pi(q) - \kappa + \delta \cdot (\pi(q) - \kappa) + \frac{\delta^2 \cdot (\pi(q) - \kappa)}{1 - \delta}$$

If the user cheats, then

$$PV_{cheat}^{user} = \pi(q) - \kappa' + \delta \cdot (\pi(q') - \kappa) + \frac{\delta^2 \cdot (\pi(q) - \kappa)}{1 - \delta}$$

Table 23.4 Cooperation Thresholds

	Conditional Cooperation	*One-Period Punishment*
Network cooperates if	$\delta_{cc}^{net} > \dfrac{c(q) - c(q')}{\kappa - c(q')}$	$\delta_{pun}^{net} > \dfrac{c(q) - c(q')}{\kappa - c(q)}$
User cooperates if	$\delta_{cc}^{user} > \dfrac{\kappa - \kappa'}{\pi(q) - \kappa'}$	$\delta_{pun}^{user} > \dfrac{\kappa - \kappa'}{\pi(q) - \pi(q')}$

For cooperation to be motivated, it must be that

$$PV_{coop}^{user} > PV_{cheat}^{user} = \pi(q) - \kappa + \delta \cdot (\pi(q) - \kappa) + \frac{\delta^2 \cdot (\pi(q) - \kappa)}{1 - \delta}$$
$$> \pi(q) - \kappa' + \delta \cdot (\pi(q') - \kappa) + \frac{\delta^2 \cdot (\pi(q) - \kappa)}{1 - \delta}$$

Simplifying, we get $\delta > (\kappa - \kappa')/(\pi(q) - \pi(q'))$.

The cooperation thresholds for both players are summarized in Table 23.4, where the conditional cooperation profile is indicated by the subscript *cc* and the one-period punishment profile is indicated by the subscript *pun*.

Remark 23.1 It holds that $\delta_{cc}^{net} < \delta_{pun}^{net}$ since $c(q') < c(q)$, and also that $\delta_{cc}^{user} < \delta_{pun}^{user}$ since $\pi(q') - \kappa' > 0$; hence, $\pi(q') > \kappa'$. Thus, both players are more motivated to cooperate under the conditional-cooperation profile.

The proof of Theorem 23.2 is now complete. □

23.3.4 User–Many Networks Interaction

Having examined the generic relationship between a single user and a single network through two different types of strategies for each of the players, we now propose a model for network selection in a converged environment, which is based on the knowledge obtained from the analysis of previous games investigated. The user–network interaction is modeled as a game between one user of the converged environment and the participating access networks that are available to the specific user: the networks play simultaneously as one player called *Networks*. The payoff for the player *Networks* is given as an array of payoffs corresponding to each of the individual access networks.

The situation we model is the following: the user plays first and offers a compensation to *Networks*, and *Networks* examine the compensation and make a decision concerning how many of them to accept and how many to reject the compensation. Any subset of *Networks* could accept or reject the proposed offer, including all accepting and all rejecting. In the latter case, the game terminates with zero payoff to the user and to *Networks*. If one or more networks accept the compensation, then for each network, the user estimates his own satisfaction, given by the function $\pi(q)$, and selects the network that is predicted to best satisfy that value (if only one network accepts, the selection is trivial).

The estimation of $\pi(q)$ by the user is based on user measures of network context and is different for each network. The user's decision to select one of the networks induces the specific network to start interacting with the user, while the rest of the networks do not interact any further with the user in subsequent iterations of the game. From then on, the interaction between the user and the network is as previously described. The payoffs for the networks that are not selected are zero, while the payoffs for the user and the selected network are as previously discussed.

Similarly, the estimation of $\pi(q)$ for selecting a single network by the user–many networks interaction may be used to prioritize the available networks by the user, in order to be able to select more than one networks, for example, two networks. Thus, the two networks ranked first, for example, the two *best* networks, in this list could be the ones chosen to interact in order to support fast session handoffs throughout a session. The subsequent section considers the case that $\pi(q)$ is used to prioritize all available networks and further that the two *best* networks are selected to cooperate for serving a user *premium* service request, that is providing additional quality guarantees including handoffs throughout a session.

23.4 Support for Fast Session Handoffs

In a converged network, there may exist multiple network operators, each one of them interacting with the participating users indirectly, through an IP-based common management platform, operated by a platform administrator, to achieve user participation and consequently revenue maximization. In this section, we consider the case where enhanced quality demands may require the cooperation of two networks in advance. The two networks are selected by the use of a prioritized list based on the estimation of the satisfaction to be received by the user, $\pi(q)$, as this is calculated by the user. Once the selection is completed, the platform administrator takes over and instructs the networks to cooperate, enforcing a payment partition configuration on the two cooperating networks, to avoid either of the two networks gaining bargaining advantage by handling the partition, and furthermore to ensure that the whole process is transparent to the user, obeying the user-centric paradigm followed in converged, next-generation communication networks.

In particular, the cooperation between the two networks occurs if delay constraints are critical for a service session, and then the second best network also reserves resources, in order to act as immediate backup in case quality degradation is detected and session handoff is necessary. The cooperation of the two networks does not take place only during one handover but further continues throughout the whole duration of the session to act as a "backup," ensuring service continuity, in case that quality degradation is observed in the current network, or that the user performs more than one handoffs. The support for faster session handoffs, enabled by the cooperation of the two best networks, enhances service quality in terms of service continuity and handoff delay, which is a crucial aspect for real-time, critical services (e.g., medical video).

Since a network's satisfaction is represented by its revenue gain, and since two networks must cooperate for a single service, the payment for supporting the service needs to be partitioned between them, in order for the networks to have an incentive to cooperate. Moreover, the partitioning must be done in a way that is satisfying to both networks. This reasoning motivated our research toward modeling this situation through a cooperative game presented next.

We model the *payment partition* as a game of bargaining between the two networks. Firstly, we define the payment partition as a game between the two networks and we show that this is equivalent to the well-known Rubinstein bargaining game [35, chapter 3], when the agreement is reached in the first negotiation period. Given this equivalence, an optimal solution to the Rubinstein bargaining game would also constitute an optimal solution to the payment partition game. To reach the optimal solution, we utilize the well-known NBS [35, chapter 2], which applies to Rubinstein bargaining games when the agreement is reached in the first negotiation period and therefore to the payment–partition game.

23.4.1 Cooperative Bargaining Model

Let $q \in Q$ be the quality level for which the two networks negotiate. Consider the payment partition scenario, where two networks want to partition a service payment $\pi_i(q)$ set by the converged platform administrator. Let $c_i(q)$ be the resource reservation cost of network i. Given the cost characteristics of network i, each network seeks a portion:

$$\pi_i(q) = c_i(q) + \phi_i(q) \tag{23.3}$$

where $\phi_i(q)$ is the actual profit of network i, such that

$$\pi_1(q) + \pi_2(q) = \Pi(q) \tag{23.4}$$

where $\Pi(q)$ is the total payment announced by the converged platform administrator. The networks' goal is to find the payment partition, which will maximize the value of $\phi_i(q)$, given the values of $\Pi(q)$ and $c_i(q)$. Definition 23.9 defines the bargaining game between the two networks:

Definition 23.9 (Payment–partition game). Fix a specific quality level $q \in Q$ such that a fixed payment Π is received. Consider a one-shot strategic game with two players corresponding to the two networks. The profiles of the game, that is, the strategy sets of the two players, are all possible pairs (π_1,π_2), where $\pi_1, \pi_2 \in [0, \Pi]$ such that $\pi_1 + \pi_2(q) = \Pi$. All such pairs are called agreement profiles and define set S^a. So, $S^a = \pi_1 \times \pi_2$. In addition, there exists a so-called disagreement pair $\{s_1^d, s_2^d\}$, which corresponds to the case where the two players do not reach an agreement. So, the strategy set of the game is given by $\mathcal{S} = S^a \bigcup \{s_1^d, s_2^d\}$. For any agreement point $s \in S^a$, the payoff $U_i(s)$, for player $i \in [2]$, is defined as follows:

$$U_i(s) = \pi_i(q) - c_i(q) \tag{23.5}$$

Otherwise,

$$U_1(s_1^d) = U_2(s_2^d) = 0 \tag{23.6}$$

This game is referred to as the payment–partition game.

Fact 23.1 Let $s^* = (\pi_1^*, \pi_2^*)$ be an optimal solution of the payment–partition game. Then $\phi_i(q) = \pi_i^*(q) - c_i(q)$, where $i \in [2]$, comprises an optimal solution of the payment partition scenario.

23.4.1.1 Equivalence to a Rubinstein Bargaining Game

Initially, we show the equivalence between the *payment–partition* game and a Rubinstein bargaining game, a.k.a., the basic alternating-offer game defined next according to [29]:

Definition 23.10 (Rubinstein bargaining game). Assume a game of offers and counteroffers between two players, π_{it}^r, where $i \in \{1, 2\}$ and t indicates the time of the offer, for the partition of a cake, of initial size of Π^r. The offers continue until either agreement is reached or disagreement stops the bargaining process.

At the end of each period without agreement, the cake is decreased by a factor of δ_i. If the bargaining procedure *times out*, the payoff to each player is 0. Offers can be made at time slot $t \in \mathcal{N}_0$. If the two players reach an agreement at time $t > 0$, payment $U_i(t)$ of player i a share $\pi_{it}^r \cdot t \cdot \delta_i$, where $\delta_i \in [0, 1]$ is a player's discount factor for each negotiation period that passes without agreement being reached. The following equation gives the payment partitions of the two players: AQ3

$$\pi_1^r(t) \cdot t \cdot \delta_1 = \Pi^r - \pi_2^r(t) \cdot t \cdot \delta_2 \tag{23.7}$$

So, if agreement is reached in the first negotiation period, the payment partition is as follows:

$$\pi_1^r(t) = \Pi^r - \pi_2^r(t) \tag{23.8}$$

The payoff U_i^r of the players $i, j \in [2]$ if the agreement is reached in iteration t is the following:

$$U_i^r(t) = \pi_i^r(t) = \Pi^r - \pi_j^r(t) \tag{23.9}$$

Such a game is called a Rubinstein bargaining game.

Proposition 23.2 Fix a specific quality q. Then, the payment–partition game is equivalent to the Rubinstein bargaining game, when the agreement is reached in the first negotiation period.

Proof Assuming that an agreement in the Rubinstein bargaining game is reached in the first negotiation period $t = 1$, then the game satisfies the following:

$$U_1^r(1) + U_2^r(1) = \Pi^r(1),$$

which is a constant.

In the payment–partition game, assuming an agreement profile s, we have

$$U_1(s) + U_2(s) = \pi_1(q) - c_1(q) + \pi_2(q) - c_2(q) = \Pi(q) - c_1(q) - c_2(q),$$

since $\Pi(q) = \pi_1(q) + \pi_2(q)$ and $c_1(q), c_2(q)$ are constants for a fixed quality level. It follows that $U_1(s) + U_2(s)$ is also constant. It follows that the *Rubinstein bargaining* game and the *Payment–partition* game are equivalent. □

We define

Definition 23.11 (Optimal payment partition). The optimal partition is when bargaining ends in an agreement profile that gives the highest possible payoff to each player given all possible actions taken by the opponent.

Proposition 23.2 immediately implies:

COROLLARY 23.3 Assume that agreement in a Rubinstein bargaining game is reached in the first negotiation period, that $\Pi^r = \Pi$, and that the corresponding profile s^* is an optimal partition for the Rubinstein game. Then, s^* is also an optimal partition for the payment–partition game.

23.4.2 Payment Partition Based on the Nash Bargaining Solution

Since the Nash bargaining game and the payment–partition game are equivalent when agreement is reached in the first period, we utilize the solution of a Nash bargaining game in order to compute an optimal solution, that is, a partition, which is satisfactory for the two networks in terms of payoffs from the payment–partition game. The solution of the Nash bargaining game, known as the NBS, captures such configuration. Therefore, since disagreement results in payoffs of 0, we are looking for an agreement profile $s = (\pi_1, \pi_2)$ such that the corresponding partition of the players is an optimal payment partition, that is, the partition that best satisfies both players' objectives.

The next theorem proves the existence of an optimal partition of the payment between the two players, given each network's cost c_i.

THEOREM 23.3 There exists an optimal solution for the payment–partition game and is given by the following: $\pi_1(q) = \frac{1}{2}(\Pi(q) + c_1(q) - c_2(q))$ and $\pi_2(q) = \frac{1}{2}(\Pi(q) + c_2(q) - c_1(q))$

Proof We consider only agreement profiles and thus refer to the partition $\pi_i(q)$ assigned to player i. In any such profile, it holds that $\pi_1(q) + \pi_2(q) = \Pi(q)$. Assuming a disagreement implies that cooperation fails between the two networks and the payoff gained by player i equals to $U_i(s^d) = 0$. Since in any such profile, it holds that $U_i(s^a) > 0$, it follows that the disagreement point is not an optimal solution. Since the payment–partition game is equivalent to the Nash bargaining game (Proposition 23.2), an NBS of the bargaining game is an optimal solution of the payment–partition game between two players, that is, a partition $(\pi_1^*(q), \pi_2^*(q))$ of an amount of goods (such as the payment). According to the NBS properties, it holds that

$$
\begin{aligned}
NBS &= (U_1(\pi_1(q)^*) - U_1(s^d))(U_2(\pi_2(q)^*) - U_2(s^d)) \\
&= \max(U_1(\pi_1(q)) - U_1(s^d))(U_2(\pi_2(q)) - U_2(s^d)) \\
&\quad 0 \le \pi_1(q) \le \Pi(q), \quad \pi_2(q) = \Pi(q) - \pi_1(q)
\end{aligned}
$$

Since $U_1(s^d) = 0$, $U_2(s^d) = 0$, and $\pi_2(q) = \Pi(q) - \pi_1(q)$,

$$\max(\pi_1(q) - c_1(q))(\Pi(q) - \pi_1(q) - c_2(q)) = (-2\pi_1(q) + \Pi(q) - c_2(q) + c_1(q)) = 0$$

Therefore,

$$\pi_1(q) = \frac{1}{2}(\pi(q) - c_2(q) + c_1(q)), \quad \pi_2(q) = \frac{1}{2}(\pi(q) + c_2(q) - c_1(q)) \tag{23.10}$$

□

Remark 23.2 Note that the probability of degradation does not affect the optimal partition, but it is still part of the networks' payoff functions; therefore, a network with a high estimated probability of degradation will receive much less of a payoff than its actual payment partition.

We proceed to investigate how the solution to the payment partition behaves when we consider the existence of a constant set by the converged platform administrator representing the probability of degradation.

THEOREM 23.4 Assume that the converged platform administrator assigns a constant value p_i^f to network $i \in [2]$ representing the expected quality degradation, based on the particular service and current network conditions. Then, the value of the optimal solution is the same as in Theorem 23.3.

Proof Our game has the same strategy set as before. Concerning the utility functions of the players in case of disagreement, we have also $U_1(s^d) = 0$, $U_2(s^d) = 0$, and $\pi_2(q) = \Pi(q) - \pi_1(q)$ as before. In case of agreement, we have in addition a constant probability p_i^f in the payoff function of each network:

$$U_i(s) = (1 - p_i^f)(\pi_i(q) - c_i(q))$$

Therefore,

$$\begin{aligned}\max(1 - p_1^f)(\pi_1(q) - c_1(q))(1 - p_2^f)(\Pi(q) - \pi_1(q) - c_2(q))\\ = (1 - p_1^f)(1 - p_2^f)(-2\pi_1(q) + \Pi(q) - c_2(q) + c_1(q)) = 0\end{aligned}$$

The optimization removes the probabilities and the optimal solution is the same as Equation 23.10:

$$\pi_1(q) = \frac{1}{2}(\pi(q) - c_2(q) + c_1(q)), \quad \pi_2(q) = \frac{1}{2}(\pi(q) + c_2(q) - c_1(q)) \tag{23.11}$$

□

23.4.3 A Bayesian Form of the Payment–Partition Game

Since the partitioning is based on each network's cost, it is required that the networks are truthful about their costs. Truthfulness is a very important consideration in cooperative situations, especially in bargaining games. The question that arises is whether it would be wise for a player to lie, considering that the player cannot be aware of who the other player is from the original set of available networks and thus cannot guess whether the other player has more or less cost, thus not being able to correctly assess the risk of such an action.

A Bayesian game [36, chapter 5] is a strategic form game with incomplete information attempting to model a player's knowledge of private information, such as privately observed costs, that the other player does not know. Therefore, in a Bayesian game, each player may have several types of behavior (with a probability of behaving according to one of these types during the game). We use the Bayesian form for the *payment–partition* game, in order to investigate the outcomes of the game, given that each network does not know whether the cost of its opponent is lower or higher than its own.

A Bayesian player begins with a prior knowledge, which is not always precise, and is expressed as the distribution on the "type" parameters of the opponent. The consistency shows whether the updated knowledge on these "type" parameters becomes more accurate as "evidence" data are collected; in fact it is believed that two Bayesian players will ultimately have very close predictive distributions [37]. We allow for both players to have the same "type" parameters and we investigate their behavior according to belief for these types. We study both players in a similar manner adopting the above statement, that both networks will ultimately have very close predictive distributions.

Let each network in the *payment–partition* game have two types: the *lower-cost* type (including networks of equal cost) and the *higher-cost* type. Suppose that each of the two networks has incomplete information about the other player, that is does not know the other player's type. Furthermore, each of the two networks assigns a probability to each of the opponent's types according to own beliefs and evaluations. Let p_i^l be the probability according to which network i believes that the opponent is likely to be of type *lower cost*, and $p_i^h = (1 - p_i^l)$ be the probability according to which network i believes that the opponent is likely to be of type *higher cost*.

Since the two players are identical, that is they have the same two types and the same choice of two actions, we will only analyze network i; conclusions also hold for network j, where $i, j \in [2], i \neq j$. Therefore, network i believes that network j is of type *lower cost* with probability p_i^l and of type *higher cost* with probability $1 - p_i^l$. Each network has a choice between two possible actions: to declare its own real costs (D) or to cheat (C), that is, declare higher costs $c_i'(q) > c_i(q)$. The possible payoffs for network 1 are given in Tables 23.5 and 23.6.

Table 23.5 Network *i* Payoffs when Opponent Is of Type *Lower Cost*

	Network j Strategies	
	D	*C*
Network *i* Strategies		
D	$\frac{1}{2}(\Pi(q) + c_i(q) - c_j(q))$	$\frac{1}{2}(\Pi(q) + c_i(q) - c_j'(q))$
C	$\frac{1}{2}(\Pi(q) + c_i'(q) - c_j(q))$	$\frac{1}{2}(\Pi(q) + c_i'(q) - c_j'(q))$

Table 23.6 Network *i* Payoffs when Opponent Is of Type *Higher Cost*

	Network j Strategies	
	D	*C*
Network *i* Strategies		
D	$\frac{1}{2}(\Pi(q) + c_j(q) - c_i(q))$	$\frac{1}{2}(\Pi(q) + c'_j(q) - c_i(q))$
C	$\frac{1}{2}(\Pi(q) + c_j(q) - c'_i(q))$	$\frac{1}{2}(\Pi(q) + c'_j(q) - c'_i(q))$

LEMMA 23.3 If $p_i^l > p_i^h$, that network j is believed by network i to be of lower cost, then it is more preferable for network i to lie, where $i, j \in [2]$, $i \neq j$.

Proof In Table 23.5, network i has higher or equal costs to network j since network j is of type *lower cost*; thus, $c_i(q) \geq c_j(q)$. When both players play D, that is, they both declare their real costs, and an equal or greater piece of the payment is assigned to network i, since the partition of the payment is directly proportional to the networks' costs. If network i plays C, that is, cheats, while network j plays D, then $c'_i(q) > c_i(q) > c_j(q)$, a profitable strategy for network i, since an even greater piece of the payment will be received. For the cases that network j decides to play C, then the payment partition may or may not favor network j (it depends on the actual amount of cheating and the action of network i). If network i plays C, then it is more likely that $c'_i(q) > c'_j(q)$, and network i will get a greater piece, than it would if it plays D. □

LEMMA 23.4 If $p_i^h > p_i^l$, that network j is believed by network i to be of higher cost, then it is more preferable for network i to lie, where $i, j \in [2]$, $i \neq j$.

Proof In Table 23.6, network i has lower costs compared to network j; thus, $c_i(q) < c_j(q)$. When both players play D, that is, they both declare their real costs, and an equal or greater piece of the payment is assigned to network j. If network i plays C, that is, cheats, then $c'_i(q) > c_i(q)$, so playing C will end up in a higher payoff for network i, and in case network j plays D, i may even get the bigger piece of the partition. If network j plays C, it is still better for network i to play C, since this will end up in network i receiving a greater piece than it would if it plays D when network j plays C, although, more likely, not the greater of the two pieces. □

23.4.3.1 Motivating Truthfulness

The earlier results motivated us to seek suitable conditions in order to motivate networks to be truthful. Next, we enforce truthfulness through the penalty functions of the networks. First note:

COROLLARY 23.5 Two networks playing the Bayesian form of the *payment–partition* game are not motivated to declare their real costs but instead they are motivated to cheat and declare higher costs, that is, $c'_i(q) > c_i(q)$, $1 \in [2]$, in order to get greater payoffs.

Proof Straightforward by Lemmas 23.3 and 23.4. □

In order to motivate the two networks to declare their real costs, there must exist a mechanism that can penalize a player who turns out to lie on its real cost, assuming that it is detectable whether a player has lied or not; we refer to such mechanisms as pricing mechanisms [30]. Let the converged platform administrator be able to detect after the service session has terminated, whether either of the participating networks has lied about its costs. In order to motivate the networks to declare their real costs, we introduce a *pricing mechanism*, that is, a new variable that tunes the resulting payoffs, in the payoff function of each player. The pricing mechanism is a postgame punishment, that is, cheating in a game does not affect the game in which a network cheats but subsequent games. Thus, a state of history of a player's behavior in similar interactions must be kept.

We define a pricing mechanism consisting of variable $\beta_i \in [0, 1]$, which represents the probability of being truthful, and it may adaptively modify the payoffs of a player, according to the player's history of actions. In particular, it is a ratio of the number of times network i has been caught lying, over a selected finite number of periods representing the window of history that the converged platform administrator monitors for each network.*

The less frequently a player cheats, the closer to 1 its β_i is. Thus, the administrator sets the payoff of network i to be

$$\pi_i(q) = \frac{1}{2}(\Pi(q) + \beta_i \cdot c_i(q) - \beta_j \cdot c_j(q)), \tag{23.12}$$

where $i, j \in [2], i \neq j$. The players are motivated to declare their real costs, since any cheating would decrease β_i, affecting any future payoffs from such procedure. An evaluation of the the Bayesian form of the *payment–partition* game including β_i in the players' payoffs and allowing a history of behavior to be collected from repeating the game is given in Section 23.5.

23.5 Evaluating Solutions through Simulations

23.5.1 Evaluation of User–Network Interaction in Single Network Selection

The evaluation presented here is based on a MATLAB® [38] implementation of an iterated user–network interaction game, where user and network strategies are played against each other repeatedly, in order to evaluate each strategy in terms of payoff. The implementation of the user–network interaction game was based on a publicly available MATLAB implementation of the iterated prisoner's dilemma game [39], which has been extended to include all the strategies examined in Section 23.3. In each simulation run, both players play their strategies and get payoffs accordingly; we use numbers that follow the relationships of the payoffs as described in their general case in the repeated game model (Table 23.3). A randomly generated perceived quality and a fixed threshold of expected quality are also implemented in all simulation runs.

For each simulation, a random number of iterations were run to get cumulative user and network payoffs for each combination of a user strategy playing against a network strategy. The user payoffs per strategy and the network payoffs per strategy are eventually added to give the most profitable user and network strategies, respectively, for the total number of iterations of a simulation run; then the average cumulative payoffs from all simulation runs are calculated. Although the number

* We consider that a revelation of the real costs of the two access networks is always possible at a later stage of the procedure (e.g., after session termination).

Table 23.7 User Payoffs from All Strategy Combinations

User Payoffs	*Network Strategies*	
User Strategies	*Tit-for-Tat*	*Cheat and Return*
Grim	793.14	4.42
Cheat and leave	6.62	4.65
Leave and return	793.14	252.92

Table 23.8 Network Payoffs from All Strategy Combinations

Network Payoffs	*Network Strategies*	
User Strategies	*Tit-for-Tat*	*Cheat and Return*
Grim	793.14	7.42
Cheat and leave	3.82	4.23
Leave and return	793.14	617.21

of iterations for each simulation is randomly generated, we run the simulation process 100 times, that is, by randomly generating 100 different numbers of iterations, in order to include behaviors when the number of iterations is both small and large.

The payoffs are calculated for an average of 264.38 iterations per simulation run* and presented in Table 23.7 (user payoffs) and in Table 23.8 (network payoffs). The numbers we see are the average payoffs from the 100 simulation runs; however, for each simulation run, a cumulative payoff is given for all the iterations run, because it is important to capture the continuation of behavior in the repeated game by adding the payoffs for all iterations in the same simulation run.

In both tables, we see a score for each strategy combination. The score corresponds to either a user payoff (Table 23.7) or a network payoff (Table 23.8). The results show that the most profitable user strategies, when the network plays with the *tit-for-tat* strategy, are the *grim* and the *leave-and-return* strategies. The most profitable user strategy, when the network employs the *cheat-and-return* strategy, is the *leave-and-return* strategy. The reason is that the *grim* strategy involves the action of leaving forever from the interaction upon detection of cheating from the network, so it receives no more payoff from the interaction upon leaving. On the other hand, the most profitable network strategy when the user employs either the *grim* strategy or the *leave-and-return* strategy is the *tit-for-tat* strategy. When the user plays the *cheat-and-leave* strategy, the results generated a slightly increased payoff for the network's *cheat-and-return* strategy. AQ4

The *tit-for-tat* strategy generates better payoffs when the user strategy used involves cooperation as long as the network cooperates, no matter what the punishment is for cheating because, according to the mirroring behavior guided by the *tit-for-tat* strategy, the network keeps cooperating until the end of the game. This is seen in the identical payoffs for both the user and the network generated for the cases when the *tit-for-tat* strategy plays against (a) the *grim* strategy and (b) the *leave-and-return* strategy.

* minimum iterations generated: 8, maximum iterations generated: 1259.

The minimum punishment applied by the *leave-and-return* user strategy appears to have a more positive effect on the payoffs compared to the rest of the user strategies for both network strategies evaluated. However, we must note that the user strategy that motivates cooperation the most, according to the theoretical results, is the *grim* strategy; thus, it is more likely for network to cooperate when the *grim* strategy is employed by the user. The more mild punishment employed by the *leave-and-return* strategy may be less motivating, but for the case that the network cheats, it will return to the interaction after leaving for one period and thus collect more cumulative payoffs. Thus, the *one-period-punishment* profile of the game, that is, when the user plays the *leave-and-return* strategy and the network plays the *tit-for-tat* strategy, is the most profitable.

23.5.2 Evaluation of Truthfulness in Supporting Fast Session Handoffs

Next, we evaluate the Bayesian form of the *payment–partition* game between two networks cooperating to support a service session with strict QoE constraints. The payoffs of the game are based on the NBS as this has been demonstrated in Section 23.4.2 and further include the term β_i (Equation 23.12), in order to motivate the networks to cooperate. The numerical values used for $\Pi(q)$, $c_i(q)$, $c_{-i}(q)$ in case the networks are truthful or lying obey the payoff relations as these are given in Table 23.5 and Table 23.6 and the overall model of the payment–partition game as described in Section 23.4.1 and resolved in Section 23.4.2.

The player types are 1 = *lower cost* or 2 = *higher cost* as these are explained in Section 23.4.3. Each network has four strategies per simulation run: (1) always declare its real costs—indicated as *Only D*, (2) always cheat—indicated as *Only C*, (3) randomly declare real costs or cheat with a probability of 0.5 for each option—indicated as 50% C/D, and (4) adapt to the value of β_i, that is, cheat only if $\beta_i \geq 0.9$ else declare real costs—indicated as *Adapt*. Each player's initial value of β_i in the simulations is randomly generated and then adapted to the player's play according to the selected strategy. Overall, we have run 100 simulation runs with different initial values for β_i and with a random number of iterations, indicated as *R*. Table 23.9 presents the results from the 100 simulation runs as cumulative payoffs for the number of iterations, for each of the two networks partitioning the service payment. We observe that in all simulation runs, the second strategy, that is, always to cheat, is the least profitable strategy for both networks, regardless of their types and initial values of β_i. This is because of the presence of β_i in the payoffs, which punishes the choice of cheating, by detecting such an action after any iteration. On the other hand, the more profitable

Table 23.9 Average Network 1 Payoffs from Payment–Partition Game

Avg. R = 288.59	*Min. R* = 4	*Max. R* = 1749			
Network 1 Payoffs	*Network 2 Strategies*	*Only D*	*Only C*	*50% C/D*	*Adapt*
Network 1 Strategies					
Only D		1462.19	1735.50	1563.25	1481.17
Only C		1179.02	1471.58	1308.93	1212.26
50% C/D		1354.71	1600.62	1457.92	1377.08
Adapt		1441.80	1700.63	1539.63	1460.79

strategies include actions of declaring the real costs, that is, being truthful; specifically, the first strategy of always being truthful is in most cases the most profitable, with the fourth strategy of adapting to β_i generating comparable payoffs.

23.6 Conclusions

This chapter has provided game-theoretic studies for the interactions during the network selection mechanism. In particular, we present the interaction between a user and its serving network, as well as for handoff during a session, specifically for the payment–partition procedure between two cooperating networks. The game-theoretic tools used for these studies included the prisoner's dilemma game model, one-shot and repeated game models, bargaining game models, and bayesian game models with their corresponding equilibrium notions.

Primarily, we have modeled the interaction between a single user and its serving access network as a one-shot prisoner's dilemma and have concluded that if the actions available to the user and the access network were to either cooperate or to cheat, then both players would be motivated to cheat. Since this was not a desirable behavior, we moved on to model the interaction as a repeated game, which ended up motivating cooperation between the two players. Moreover, it has been shown that cooperation is motivated more easily if the strategies employed by the players incorporate threats for harsh punishments in case the opponent cheats. Based on the study of user–network interaction, a network selection model was provided.

The second game-theoretic study concerns the case where network cooperation is required to support a session handoff with service continuity and minimum handoff delays. The particular aspect we have investigated is the optimization of the payment–partition process between the two cooperating networks. It has been shown that this process is equivalent to the well-known Rubinstein bargaining game when the agreement is reached during the first negotiation period as well as to a Nash bargaining game. Therefore, an optimal payment partition has been derived using the NBS, which indicated the cost of resource reservation for each participating network to be the deciding factor for the optimal payment partition. Furthermore, it has been shown that any constant probability for the expected quality degradation does not affect the optimal payment partition. The payment partition game raises the issue of truthfulness, that is, whether the networks are motivated to be truthful about their resource reservation costs, since lying could in some cases increase their overall partition. Simple reasoning illustrates that lying is worth the risk, so networks are motivated to lie in the absence of pricing mechanisms, that is, mechanisms that punish players for cheating behavior. Hence, to motivate truthfulness, we have proposed such pricing mechanism in the players' payoff, through a Bayesian game model.

Finally, numerical results for both the user-interaction game and the Bayesian model of the payment–partition game have been provided (based on MATLAB simulations). In both games, the numerical results provide cumulative payoffs of the players when different strategies are played iteratively. The numerical results verify the conclusions drawn from the theoretical studies. AQ5

23.7 Open Issues

We wish to explore further the repeated nature of the interaction between the networks and the user. In particular, we will seek to find optimal solutions of the repeated game and identify players' behaviors as well as the overall system performance. Furthermore, we will try to develop other tools that enable us to capture the effect of the previous outcomes of the interaction between the entities, in any future decisions.

Furthermore, we plan to investigate additional parameters for network selection that can also improve handoffs, such as consideration of reserved resources as an incentive for a user to select a particular network. This could be defined differently for each network and would be an additional factor for the user to consider, prior to evaluating the satisfaction estimation function for each network. Moreover, we plan to perform further studies on the way compensation is set to avoid user exploitation by the network, in the case that the user valuates this privately and such private valuation of service is significantly higher than network-incurring costs.

In this study, we utilized the notion of the discount factor. We would like to analyze further this notion but moreover to discover other such tools that may be able to capture better the interaction of the players. We also plan to further investigate the user–network interaction when more than one networks are involved, modeling it as a repeated game between multiple players. For this enhanced game, we plan to find the necessary and sufficient conditions enforcing the players into behaviors that maximize both local (individual) and global (system) objectives.

Finally, we plan to investigate mechanisms that will enforce the modeled interactions to converge into desired outcomes. Toward this goal, we plan to explore tools from the mechanism design. In particular, we will try to find suitable utilities for the players that the converged network will impose on the involved entities, such that trying to maximize their individual gain will end up in desired behaviors, for example, being truthful.

AQ6

References

1. A. Iera, A. Molinaro, C. Campolo, and M. Amadeo, An access network selection algorithm dynamically adapted to user needs and preferences, *Proceedings of the IEEE International Symposium on Personal, Indoor and Mobile Radio Communications* (PIMRC '06), Helsinki, Finland, pp. 1–5, IEEE, September 2006.
2. AQ7 F. Bari and V. C. M. Leung, Service delivery over heterogeneous wireless networks: Network selection aspects, *Proceedings of the ACM IWCMC 2006*, Vancouver, Canada, ACM, July 2006.
3. J. Antoniou, I. Koukoutsidis, E. Jaho, A. Pitsillides, and I. Stavrakakis, Access network synthesis game in next generation networks, *Elsevier Computer Networks Journal*, 53, 15, 2716–2726, Oct. 2009. AQ8
4. I. M. Weinstein, Real World Convergence for conferencing: Considerations for hosting IP video on the enterprise data network, Wainhouse Research Issue Paper, July 2005.
5. J. S. Dyer, P. C. Fishburn, R. Estever, J. Wallenius, and S. Zionts, Multiple-criteria decision-making: Multiattribute utility theory—The next ten years, *Management Science* 38, 5, 645–653, April 1992.
6. X. Xiaochun and W. Xiaoyan, A simple rate control algorithm for maximizing total user utility, *Proceedings of the International Conference on Communication Technology* 2003, pp. 135–138, IEEE, April 2003.
7. S. E. Elayoubi, T. Chahed, and G. Hebuterne, Admission control in UMTS in the presence of shared channels, *Computer Communications* 27, 11, 1115–1126, July 2004.
8. D. Gao, J. Cai, and K. N. Ngan, Admission control in IEEE 802.11e wireless LANs, *IEEE Network* 19, 4, 6–13, July 2005.
9. J. Antoniou, V. Papadopoulou, V. Vassiliou, and A. Pitsillides, Cooperative user network interactions in next generation communication networks, Submitted to Elsevier *Computer Networks Journal*, May 2009 (under 2nd review).
10. S. Kher, A. K. Somani, and R. Gupta, Network selection using fuzzy logic, *Proceedings of the 2nd International Conference on Broadband Networks*, Boston, MA, pp.876–885, IEEE, October 2005.

11. W. Zhang, Handover decision using fuzzy MADM in heterogeneous networks, *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC '04)*, Atlanta, GA, pp. 653–658, IEEE, March 2004.
12. Q. Song and A. Jamalipour, An adaptive quality-of-service network selection mechanism for heterogeneous mobile networks, *Wireless Communications and Mobile Computing*, 5, 6, 697–708, 2005.
13. H. Chan, P. Fan, and Z. Cao, A utility-based network selection scheme for multiple services in heterogeneous networks, *Proceedings of the IEEE Wireless Networks, Communications and Mobile Computing*, pp. 1175–1180, IEEE, June 2005.
14. J. Antoniou and A. Pitsillides, Radio access network selection scheme in next generation heterogeneous environments, *Proceedings of the IST Mobile Summit 2006*, Mykonos, Greece, IEEE, June 2006.
15. V. Gazis, N. Alonistioti, and L. Merakos, Towards a generic "Always Best Connected" capability in integrated WLAN/UMTS cellular mobile networks, *IEEE Wireless Communications Magazine* 12, 22–29, 2005.
16. X. Cai, L. Chen, R. Sofia, and Y. Wu, Dynamic and user-centric network selection in heterogeneous networks, *Proceedings of the IEEE International Conference on Performance, Computing and Communications*, New Orleans, LA, pp.538–544, IEEE, April 2007.
17. F. Bari and V. C. M. Leung, Automated network selection in a heterogeneous wireless network enviroment, *IEEE Network*, 21, 1, 34–40, January 2007.
18. O. Markaki, D. Charilas, and D. Nikitopoulos, Enhancing quality of experience in next generation networks through network selection mechanisms, *Proceedings of the B3G Environment Workshop, PIMRC 2007*, Athens, Greece, IEEE, September 2007.
19. D. Charilas, O. Markaki, D. Nikitopoulos, and M. Theologou, Packet switched network selection with the highest qos in 4G networks, *Computer Networks*, 52, 1, pp. 248–258, 2008.
20. J. Antoniou, M. Stylianou, and A. Pitsillides, RAN selection for converged networks supporting IMS signalling and MBMS service, *Proceedings of the 18th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC '07)*, Athens, Greece, June 2008.
21. J. Kim, J. H. Lee, T. Kwon, and Y. Choi, Wireless access network selection for live streaming multicast in future internet, *Proceedings of the 3rd International Conference on Future Internet Technologies*, Seoul, South Korea, June 2008.
22. D. Charilas, O. Markaki, and E. Tragos, A theoretical scheme for applying game theory and network selection mechanisms in access admission control, *Proceedings of the 3rd International Symposium on Wireless Pervasive Computing*, Santorini, Greece, pp.303–307, May 2008.
23. J. Antoniou and A. Pitsillides, 4G converged environment: Modeling network selection as a game, *Proceedings of the 16th IST Mobile and Wireless Communications Summit 2007*, Budapest, Hungary, IEEE, July 2007.
24. D. Niyato and E. Hossain, A cooperative game framework for bandwidth allocation in 4G heterogeneous wireless networks, *Proceedings of the IEEE International Conference on Communications 2006 (ICC '06)*, Istanbul, Turkey, pp.4357–4362, IEEE, June 2006.
25. H. Lin and M. Chatterjee, ARC: An integrated admission and rate control framework for competitive wireless CDMA data networks using noncooperative games, *IEEE Transactions on Mobile Computing*, 4, 3, 243–258, 2005.
26. D. Niyato and E. Hossain, Modeling user churning behavior in wireless networks using evolutionary game theory, *Proceedings of IEEE WCNC 2008*, Las Vegas, NV, 31 March–3 April 2008.
27. Roger B. Myerson, Two-person bargaining problems with incomplete information, *Econometrica*, 52, 2, 461–487, 1984.
28. N. Nisan and A. Ronen, Algorithmic mechanism design, *Games and Economic Behavior*, 35, 1–2, 166–196, April 2001.

29. P. Cramton, Y. Shoham, and R. Steinberg, editors. *Combinatorial Auctions*. MIT Press, Cambridge, MA, 2006.
30. W. E. Walsh and M. P.Wellman, A market protocol for decentralized task allocation: Extended version. *Proceedings of the Third International Conference on Multi-Agent Systems (ICMAS-98)*, Paris, France, 1998.
31. T. Groves. Incentives in teams. *Econometrica*, 41, 4, 617–631, July 1973.
32. G. Kendall, X. Yao, and S. Y. Chong, *The Iterated Prisoner's Dilemma: 20 Years On*, World Scientific Publishing Co., Singapore, Advances in Natural Computation Book Series, vol. 4, 2009.
33. A. Dixit and S. Skeath, *Games of Strategy*, W.W.Norton & Company, New York, 1999.
34. R. Hegselmann and A. Flache, Rational and adaptive playing: A comparative analysis for all possible prisoner's dilemmas, *Analyse and Kritik*, 22, 1, 75–97, 2000.
35. A. Muthoo, *Bargaining Theory with Applications*, Cambridge University Press, Cambridge, U.K., 2002.
36. H. Gintis, *Game Theory Evolving: A Problem-Centered Introduction to Modeling Strategic Interaction*, Princeton University Press, Princeton, NJ, 2000.
37. S. Ghosal, A review of consistency and convergence of posterior distribution. *Proceedings of Varanashi Symposium in Bayesian Inference*, Banaras Hindu University, 1996.
38. *MATLAB: The Language of Technical Computing*, The Mathworks, version 7.6.0.324 (R2008a), Natick, MA, February 2008.
39. G. Taylor, Iterated Prisoner's Dilemma in MATLAB, Archive for the Game Theory Category http://maths.straylight.co.uk/archives/category/game-theory, March 2007.