Computer Networks 54 (2010) 2537–2561

Contents lists available at ScienceDirect

Computer Networks

journal homepage: www.elsevier.com/locate/comnet

Game theory for cognitive radio networks: An overview

Beibei Wang *, Yongle Wu, K.J. Ray Liu

Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20742, USA

ARTICLE INFO

Article history:
Received 2 February 2010
Received in revised form 5 April 2010
Accepted 6 April 2010
Available online 13 April 2010
Responsible Editor: M.C. Vuran

Keywords:
Game theory
Wireless networks
Cognitive radio networks

ABSTRACT

Cognitive radio technology, a revolutionary communication paradigm that can utilize the existing wireless spectrum resources more efficiently, has been receiving a growing attention in recent years. As network users need to adapt their operating parameters to the dynamic environment, who may pursue different goals, traditional spectrum sharing approaches based on a fully cooperative, static, and centralized network environment are no longer applicable. Instead, game theory has been recognized as an important tool in studying, modeling, and analyzing the cognitive interaction process. In this tutorial survey, we introduce the most fundamental concepts of game theory, and explain in detail how these concepts can be leveraged in designing spectrum sharing protocols, with an emphasis on state-of-the-art research contributions in cognitive radio networking. Research challenges and future directions in game theoretic modeling approaches are also outlined. This tutorial survey provides a comprehensive treatment of game theory with important applications in cognitive radio networks, and will aid the design of efficient, self-enforcing, and distributed spectrum sharing schemes in future wireless networks.

1. Introduction

Cognitive radio technology [1] is emerging in recent years as a revolutionary communication paradigm, which can provide faster and more reliable wireless services by utilizing the existing spectrum band more efficiently [2,3]. A notable difference of a cognitive radio network from traditional wireless networks is that users need to be aware of the dynamic environment and adaptively adjust their operating parameters based on the interactions with the environment and other users in the network. Traditional spectrum sharing and management approaches, however, generally assume that all network users cooperate unconditionally in a static environment, and thus they are not applicable to a cognitive radio network.

In a cognitive radio network, users are intelligent and have the ability to observe, learn, and act to optimize their performance. If they belong to different authorities and pursue different goals, e.g., compete for an open unlicensed band, fully cooperative behaviors cannot be taken for granted. Instead, users will only cooperate with others if cooperation can bring them more benefit. Moreover, the surrounding radio environment keeps changing, due to the unreliable and broadcast nature of wireless channels, user mobility and dynamic topology, and traffic variations. In traditional spectrum sharing, even a small change in the radio environment will trigger the network controller to re-allocate the spectrum resources, which results in a lot of communication overhead. To tackle the above challenges, game theory has naturally become an important tool that is ideal and essential in studying, modeling, and analyzing the cognitive interaction process, and designing efficient, self-enforcing, distributed and scalable spectrum sharing schemes.

Game theory is a mathematical tool that analyzes the strategic interactions among multiple decision makers. Its history dates back to the publication of the 1944 book *Theory of Games and Economic Behavior* by J. von Neumann and O. Morgenstern, which included the method for finding mutually consistent solutions for two-person zero-sum

* Corresponding author. Tel.: +1 3014055823.
E-mail addresses: bebewang@umd.edu (B. Wang), wuyl@umd.edu (Y. Wu), kjrliu@umd.edu (K.J.R. Liu).

1389-1286/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.comnet.2010.04.004

games and laid the foundation of game theory. During the late 1940s, cooperative game theory had come into being, which analyzes optimal strategies for groups of individuals, assuming that they can enforce collaboration between them so as to jointly improve their positions in a game. In early 1950s, J. Nash developed a new criterion, known as Nash equilibrium, to characterize mutually consistent strategies of players. This concept is more general than the criterion proposed by von Neumann and Morgenstern, since it is applicable to non-zero-sum games, and marks a quantum leap forward in the development of non-cooperative game theory. During the 1950s, many important concepts of game theory were developed, such as the concepts of the core, the extensive-form games, repeated games, and the Shapley value. Refinement of Nash equilibriums and the concepts of complete information and Bayesian games were proposed in the 1960s. Application of game theory to biology, i.e., the evolutionary game theory, was introduced by J. M. Smith in the 1970s, during which time, the concepts of correlated equilibrium and common knowledge were introduced by R. Aumann. Starting from the 1960s, game theorists have started to investigate a new branch of game theory, mechanism design theory, focusing on the solution concepts for a class of private information games. In nowadays, game theory has been widely recognized as an important tool in many fields, such as social sciences, biology, engineering, political science, international relations, computer science, etc., for understanding cooperation and conflict between individuals.

In cognitive radio networks, network users make intelligent decisions on their spectrum usage and operating parameters based on the sensed spectrum dynamics and actions adopted by other users. Furthermore, users who compete for spectrum resources may have no incentive to cooperate with each other and instead behave selfishly. Therefore, it is natural to study the intelligent behaviors and interactions of selfish network users from a game theoretic perspective.

The importance of studying cognitive radio networks in a game theoretic framework is multifold. First, by modeling dynamic spectrum sharing among network users (primary and secondary users) as games, network users' behaviors and actions can be analyzed in a formalized game structure, by which the theoretical achievements in game theory can be fully utilized. Second, game theory equips us with various optimality criteria for the spectrum sharing problem. To be specific, the optimization of spectrum usage is generally a multi-objective optimization problem, which is very difficult to analyze and solve. Game theory provides us with well defined equilibrium criteria to measure game optimality under various game settings. Third, non-cooperative game theory, one of the most important branch of game theory, enables us to derive efficient distributed approaches for dynamic spectrum sharing using only local information. Such approaches become highly desirable when centralized control is not available or flexible self-organized approaches are necessary.

In this tutorial survey, we aim at providing a comprehensive treatment of game theory oriented towards their applications to cognitive radio networks in recent years. Considering game theory is still rarely taught in engineering or computer science curricula, we assume that the reader has very little background in this area. Therefore, we start each section by introducing the most basic game theoretic concepts, and then address how these concepts can be leveraged in designing efficient spectrum sharing schemes from a network designer's perspective. The organization of the tutorial survey is illustrated in Fig. 1, where the game theoretic spectrum sharing schemes are classified into four categories. We first discuss non-cooperative spectrum sharing games in Section 2, since networks users are mostly assumed to be selfish and only aim at maximizing their own spectrum usage. Then, we talk about the

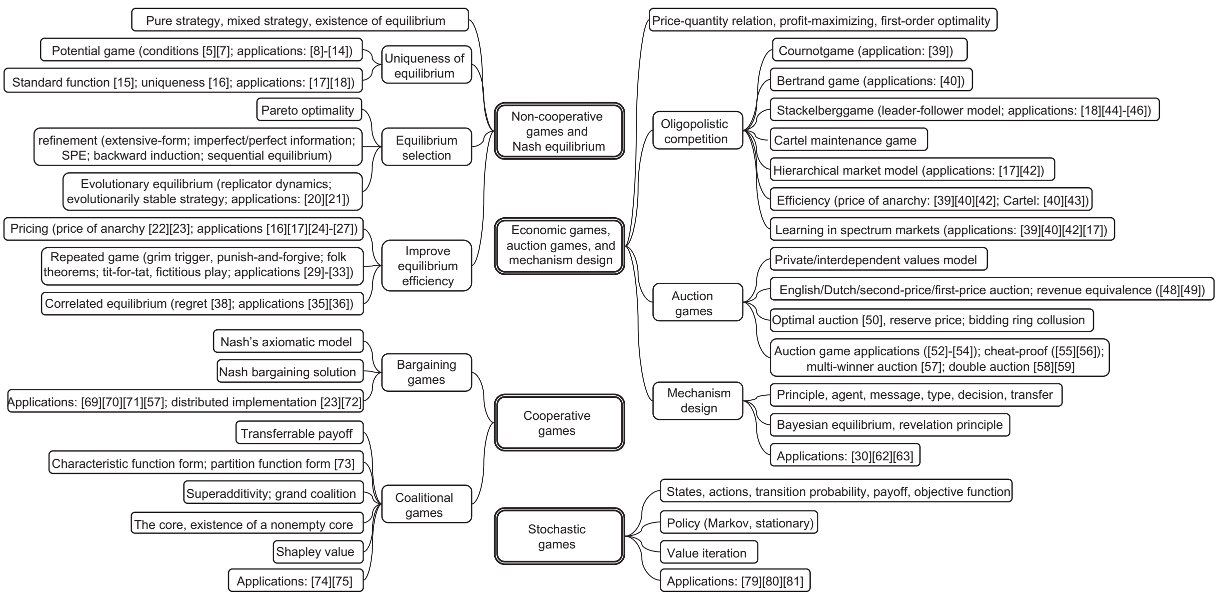

Fig. 1. Four categories of the game theoretic spectrum sharing approaches.

application of economic games and mechanism design to cognitive radio networks in Section 3, including spectrum pricing and auctions, where spectrum resources are traded like exchangeable goods in a spectrum market. Cooperative spectrum sharing games where network users have an agreement on how to utilize and distribute the spectrum resources are discussed in Section 4, and stochastic spectrum sharing games where network users adapt their strategies according the changing environment and other users' strategies are discussed in Section 5. Section 6 presents some research challenges and future directions about game theoretic spectrum sharing in cognitive radio networks.

2. Non-cooperative games and Nash equilibrium

Nash equilibrium is a key concept to understand non-cooperative game theory. Given a game where two or more players interactively make their decisions, it is natural to ask "*What will the outcome of a game be like?*" The answer is given by Nash equilibrium, which, informally speaking, is an equilibrium where everyone plays the best strategy when taking decision-making of others into account. Then, the next questions are "*Does a Nash equilibrium always exist in a game?*"and "*Is it unique?*". We will show in Section 2.1 that the existence of Nash equilibria is quite general, but the uniqueness has to be analyzed case by case.

Nash equilibrium tells us what the equilibrium outcome will be, but it does not answer the question "*How can we get to the equilibrium?*". This is more important in the context of cognitive radio networks, where players may lack the global information to directly predict the equilibrium. Instead, they may start from an arbitrary strategy, update their strategies according to certain rules, and hopefully converge to the equilibrium. Section 2.2 provides two specific conditions that guarantee the convergence to a unique Nash equilibrium.

When there exist multiple equilibria, one needs to select those equilibria that are superior to others. In Section 2.3, we discuss several equilibrium selection criteria. Pareto optimality is defined to compare multi-dimension payoff profiles, and an equilibrium not as good as others in the Pareto sense can be ignored. Moreover, some refinement can be used to narrow down the game outcomes, e.g., removing the ones with incredible actions or implausible beliefs. Evolutionary equilibrium is the one that is evolutionarily stable.

In general, Nash equilibrium often suffers from excessive competition among selfish players in a non-cooperative game, and the outcome of the game is inefficient. Hence, we are eager to know "*Can we go beyond a Nash equilibrium?*" In Section 2.4, three approaches, namely, usage of pricing, repeated game formulation, and correlated equilibrium, are discussed that can improve the efficiency of Nash equilibria.

2.1. Nash equilibrium

Game theory is a mathematical tool that analyzes the strategic interactions among multiple decision makers. Three major components in a strategic-form game model are:

- a finite set of **players**, denoted by N;
- a set of **actions**, denoted by A_i, for each player i; and
- **payoff/utility** function, denoted by $u_i : A \to \mathbb{R}$, which measures the outcome for player i determined by the actions of all players, $A = \times_{i \in N} A_i$.

Given the above definition and notations, a strategic game is often denoted by $\langle N, (A_i), (u_i) \rangle$.

Since users in a cognitive radio network may compete for the limited spectrum resources, we can model the interactions between them as a game, which is detailed as follows.

In a cognitive radio network, users who do not own a spectrum license are known as secondary users or unlicensed users, and the spectrum license holders are known as primary users or licensed users. According to the spectrum bands that secondary users are using, spectrum sharing and allocation schemes can be divided to two types. Spectrum sharing among the secondary users who access unlicensed spectrum bands is referred to as *open spectrum sharing*. In open spectrum sharing, since no users own spectrum licenses, they all have the same rights in using the unlicensed spectrum. Spectrum sharing among the secondary users and primary users in licensed spectrum bands is referred to as *hierarchical access model* or *licensed spectrum sharing*. Secondary users are allowed to access the licensed bands as long as they will not cause harmful interference to the primary users. For instance, in opportunistic spectrum access, secondary users will listen to the licensed spectrum before each transmission to make sure the primary users are inactive, and then they will choose proper operating parameters to optimize the performance or quality of service (QoS) from sharing the spectrum. In negotiation-based licensed spectrum sharing, the primary users will announce the available spectrum bands to the secondary users and distribute these bands through auction/pricing, where both primary and secondary users can maximize their profits by leasing the licensed bands.

Efficient spectrum sharing schemes are essential for improving spectrum utilization. However, since users in a cognitive radio network are intelligent and able to observe, learn, and act to optimize their performance, if they belong to different authorities and pursue different goals, fully cooperative behavior cannot be taken for granted. Instead, selfish users will compete for the limited spectrum resources, and only aim at maximizing their own benefit. As traditional spectrum sharing approaches only assume cooperative, static, and centralized network settings, new solutions based on game theoretic modeling are preferred, which can offer more flexibility in analyzing network users' strategic interactions and achieve efficient dynamic spectrum sharing. An example that explains the components of spectrum sharing games in cognitive radio networks is provided in Table 1.

In a non-cooperative spectrum sharing game with rational network users, each user only cares about his/her own benefit and chooses the optimal strategy that can maximize his/her payoff function. Such an outcome of

Table 1
Components of spectrum sharing games in cognitive radio networks.

	Open spectrum sharing	Licensed spectrum sharing (auction)
Players	Secondary users that compete for an unlicensed spectrum band	Both primary and secondary users
Actions	Transmission parameters, such as transmission power level, access rates, waveform, etc.	Secondary users: which licensed bands they want to rent and how much they would pay for leasing the licensed bands; primary users: which secondary users they will lease each unused band to and the charge
Payoff	Non-decreasing function of the quality of service (QoS) by utilizing the spectrum	Monetary gains, e.g., revenue minus cost, by leasing the licensed spectrum

the non-cooperative game is termed as Nash equilibrium (NE), which is the most commonly used solution concept in game theory.

Definition 2.1. A Nash equilibrium of a strategic game $\langle N,(A_i),(u_i)\rangle$ is a profile $a^* \in A$ of actions such that for every player $i \in N$ we have

$$u_i(a_i^*, a_{-i}^*) \geqslant u_i(a_i, a_{-i}^*) \tag{1}$$

for all $a_i \in A_i$, where a_i denotes the strategy of player i and a_{-i} denotes the strategies of all players other than player i.

The definition indicates that no player can improve his/her payoff by a unilateral deviation from the NE, given that the other players adopt the NE. In other words, NE defines the best-response strategy of each player, as stated below:

$$a_i^* \in B_i(a_{-i}^*) \quad \text{for all } i \in N, \tag{2}$$

with the set-valued function B_i defined as the **best-response function** of player i, i.e.,

$$B_i(a_{-i}) = \{a_i \in A_i : u_i(a_{-i}, a_i) \geqslant u_i(a_{-i}, a_i')\} \quad \text{for all } a_i' \in A_i. \tag{3}$$

Given the definition of NE, one is naturally interested in whether there exists an NE for a certain game so that we can study its properties. Based on the fixed point theorem, the following theorem has been shown [4].

Theorem 2.1. *A strategic game $\langle N,(A_i),(u_i)\rangle$ has a Nash equilibrium if, for all $i \in N$, the action set A_i of player i is a non-empty compact convex subset of a Euclidian space, and the payoff function u_i is continuous and quasi-concave on A_i.*

In the above definition and notation, it is implicitly assumed that players only take deterministic strategies, also known as pure strategies. More often, the players' strategies may not be deterministic and are regulated by probabilistic rules. Mixed strategy Nash equilibrium concept is then designed to describe such a scenario where players' strategies are non-deterministic.

Denote $\Delta(A_i)$ as the set of probability distributions over A_i, then each member of $\Delta(A_i)$ is a mixed strategy of player i. In general, the players take their mixed strategies independent of each other's decision. If we denote a strategy profile of player i by $(\alpha_i)_{i\in N}$ which represents the probability distribution over action set A_i, then the probability of the action profile $a = (a_i)_{i\in N}$ will be $\prod_{i\in N}\alpha_i(a_i)$, and player j's payoff under the strategy profile $(\alpha_i)_{i\in N}$ is $\sum_{a\in A}(\prod_{i\in N}\alpha_i(a_i))u_j(a)$, if each A_i is finite.

The NE defined for strategic games where players take pure strategies can then be naturally extended, and a **mixed strategy Nash equilibrium** of a strategic game is a Nash equilibrium where players in the game adopt mixed strategies, following the above extension. Without providing proof (interested readers can refer to [4]), we give the property about the existence of a mixed strategy NE in games where each player has a finite number of actions in the following theorem.

Theorem 2.2. *Every finite strategic game has a mixed strategy Nash equilibrium.*

We use the following example to explain how to derive the NE of a game. In Table 2, we list the payoff of a two-player game which can represent the competition between two users for the access of an open spectrum band, where action Dare (D) means access aggressively with a high rate and action Chicken out (C) means access moderately with a low rate. If one user accesses the spectrum aggressively while the other moderately, the former will gain more. If both access aggressively (no cooperation), neither of them will gain due to frequent collisions. If both access moderately, each will gain a much higher payoff than no cooperation.

In this game, if one user is going to dare, it is better for the other to chicken out. If one user is going to chicken out, it is better for the other to dare. Therefore, there are two pure strategy NEs in this game, namely, (D,C) and (C,D). To calculate the mixed strategy NE, we can assume that the probability distribution of player 1 (row player) over $A_i = (D,C)$ is $\alpha_1 = [x, 1-x]$, and that of player 2 (column player) is $\alpha_2 = [y, 1-y]$. Then, the expected payoff of player 1 is

$$\bar{u}_1 = 0\cdot x\cdot y + 3\cdot(1-x)\cdot y + 6\cdot x\cdot(1-y) + 5\cdot(1-x)\cdot(1-y). \tag{4}$$

According to the definition in (3), at equilibrium, player 1's expected payoff should satisfy $\frac{\partial \bar{u}_1}{\partial x} = 0$. Solving the equation, we obtain $y = 1/4$ and further get $x = 1/4$ in a similar way. Thus, there is a mixed strategy equilibrium where each user dares with probability 1/4.

Table 2
An example of multiple access game.

	D	C
D	0,0	6,3
C	3,6	5,5

2.2. Uniqueness of equilibrium

Besides existence, the uniqueness of an equilibrium is another desirable property. If we know there exists only one equilibrium, we can predict the equilibrium strategy of the players and the resulting performance of the cognitive radio network. By optimally tuning the design parameters of the game, it is possible to manipulate the behavior of the rational players towards efficient spectrum sharing at the equilibrium. Unlike the establishment of existence using the fixed point theorem, uniqueness of an equilibrium only holds for several special cases. For instance, if the payoff function of each player is strictly convex and the feasible region is also convex, then there exists a unique equilibrium in the game. In the following, we will discuss two other special cases that can guarantee the uniqueness of an equilibrium.

2.2.1. Potential games

The NE gives the best strategy given that all the other players stick to their equilibrium strategy too. However, the question is how to find the Nash equilibrium, especially when the system is implemented in a distributed manner. One approach is to let players adjust their strategies iteratively based on accumulated observations as the game unfolds, and hopefully the process could converge to some equilibrium point. Although not true in general, the iteration does converge and lead to the NE, when the game has certain special structures. For example, when the game can be modeled as a potential game, convergence to the NE is guaranteed.

The concept of potential games, proposed by [5], was first applied to cognitive radio networks in [6] and has been widely employed in the context of cognitive radio since then.

Definition 2.2. A game $\langle N,(A_i),(u_i)\rangle$ is a **potential game** if there is a potential function $P : A \to \mathbb{R}$ such that one of the following conditions holds. The game is an **exact potential game** if the first condition holds, and an **ordinal potential game** if the second condition holds.

(i) $P(a_i, a_{-i}) - P(a'_i, a_{-i}) = u_i(a_i, a_{-i}) - u_i(a'_i, a_{-i})$ for any $i \in N$, $a \in A$, and $a'_i \in A_i$.
(ii) $\mathrm{sgn}(P(a_i, a_{-i}) - P(a'_i, a_{-i})) = \mathrm{sgn}(u_i(a_i, a_{-i}) - u_i(a'_i, a_{-i}))$ for any $i \in N$, $a \in A$, and $a'_i \in A_i$, where $\mathrm{sgn}(\cdot)$ is the sign function.

From the definition, it is easy to see that any single player's individual interest is aligned with the group's interest (i.e., the potential function), and any player choosing a *better* strategy given all other players' current strategies will necessarily lead to improvement in the value of potential function. A potential game in which all players take better strategies sequentially will terminate in finite steps to an NE that maximizes the potential function.

Several useful conditions for potential games have been established in [5] and [7], and we summarize them in the following theorem. These conditions can be used to prove a game to be a potential game or guide the design of a potential game. The third condition is of particular interest, as it shows that a game is a potential game as long as payoff functions have some symmetric property.

Theorem 2.3. *A game $\langle N,(A_i),(u_i)\rangle$ is an exact potential game with a potential function $P(\cdot)$:*

(i) *if and only if*

$$\frac{\partial^2 u_i}{\partial a_i \partial a_j} = \frac{\partial^2 u_j}{\partial a_i \partial a_j} \quad \text{for all } i,j \in N, \tag{5}$$

provided that A_i is an interval of real numbers and u_i is continuously differentiable for all $i \in N$;

(ii) *if and only if there exist functions $P_0 : A \to \mathbb{R}$ and $P_i : A_{-i} \to \mathbb{R} (i \in N)$ such that*

$$u_i(a_i, a_{-i}) = P_0(a_i, a_{-i}) + P_i(a_{-i}) \quad \text{for all } i \in N, \tag{6}$$

where $P(a_i, a_{-i}) = P_0(a_i, a_{-i})$;

(iii) *if there exist functions $P_{ij} : A_i \times A_j \to \mathbb{R}$ and $P_i : A_i \to \mathbb{R}$ such that $P_{ij}(a_i, a_j) = P_{ji}(a_j, a_i)$ and*

$$u_i(a) = \sum_{j \in N \setminus \{i\}} P_{ij}(a_i, a_j) - P_i(a_i) \quad \text{for all } i,j \in N \text{ and } a \in A. \tag{7}$$

This is known as the bilateral symmetric game with $P(a) = \sum_{i \in N} \sum_{j=1}^{i-1} P_{ij}(a_i, a_j) - \sum_{i \in N} P_i(a_i)$.

Potential games have been widely used in cognitive radio networks, such as [8–14] and so on. For example, let us mention a few applications as follows.

- Waveform selection [8]. In this game, players distributively choose their signature waveform $a_i \in A_i$ to reduce correlation. The signal-to-interference-and-noise ratio (SINR) of player i is

$$\gamma_i(a_i, a_{-i}) = \frac{h_i p_i}{\sum_{j \neq i} h_{ji} p_j \rho(a_j, a_i) + n_i}, \tag{8}$$

  where h_i, p_i, n_i are the channel gain, power level, and noise variance for player i, h_{ji} is the cross channel gain from transmitter j to receiver i, and $\rho(a_i, a_j)$ is the correlation when player i and j choose waveform a_i and a_j, respectively. The payoff function is defined as some function of the SINR γ_i minus costs associated with the selected waveform,

$$u_i(a_i, a_{-i}) = f_i(\gamma_i(a_i, a_{-i})) - c_i(a_i). \tag{9}$$

  In [8], the game is claimed to be a bilateral symmetric game (Theorem 2.3, condition (iii)) when certain conditions hold.
- Power control [8]. This game is similar to the previous one except that the action space consists of all possible power levels and the cost is associated with power levels. The fixed waveforms result in correlation ρ_{ji} between player i and j. The payoff function can be written as

$$\begin{aligned} u_i(a_i, a_{-i}) &= f_i(\gamma_i(a_i, a_{-i})) - c_i(a_i) \\ &= f_{i,1}(h_i a_i) - f_{i,2}\left(\sum_{j \neq i} h_{ji} a_j \rho_{ji} + n_i\right) - c_i(a_i), \end{aligned} \tag{10}$$

when $f_i(\cdot)$ can be detached to one function of the numerator and another function of the denominator (e.g., if $f_i(\cdot)$ is in the form of a logarithm function). It is easy to show the first condition in Theorem 2.3 is satisfied for this game, and hence it is a potential game.

- Channel allocation [9]. In this game, a player's strategy is to select a channel from multiple channels for transmission, and players in the same band interfere with each other. In order to reduce mutual interference, the payoff function is defined as the total interference not only caused by other players but also causing to other players, i.e.,

$$u_i(a_i, a_{-i}) = -\sum_{j=1, j\neq i}^{N} p_j h_{ji} 1(a_j = a_i) - \sum_{j=1, j\neq i}^{N} p_i h_{ij} 1(a_i = a_j), \tag{11}$$

where the indicator function $1(a_j = a_i)$ implies player i and j have mutual interference only if they choose the same channel. Condition (iii) in Theorem 2.3 is satisfied when we define $P_{ij}(a_i, a_j) = -(p_j h_{ji} + p_i h_{ij}) 1(a_i = a_j)$ and $P_i(a_i) = 0$.

2.2.2. Standard function

Standard function is first introduced in [15] to aid the power control in cellular networks. Assume there are N cellular users, M base stations, and a common radio channel. Denote p_i as the transmitted power of user $i, h_{m,i}$ as the channel gain of user i to base station m, and n_m as the noise power at base station m. Treating the interference at a base station as noise, we can express the SINR of user i at base station m as

$$\mu_{m,i} = \frac{h_{m,i}}{\sum_{j\neq i} h_{m,j} p_j + n_m}. \tag{12}$$

Assume k_i is user i's assigned base station, to ensure an acceptable communication performance, the received power at the base station k_i should be no less than a certain level γ_i, i.e., $p_i \mu_{k_i,i}(\boldsymbol{p}) \geqslant \gamma_i$, where $\boldsymbol{p}$ denotes the power vector of the N users. The interference constraint can be rewritten as an interference function $\boldsymbol{I}(\boldsymbol{p}) = (I_1(\boldsymbol{p}), \ldots, I_N(\boldsymbol{p}))$, with each $I_i(\boldsymbol{p})$ defined below:

$$p_i \geqslant I_i(\boldsymbol{p}) = \frac{\gamma_i}{\mu_{k_i,i}}. \tag{13}$$

It has been shown that the interference function $\boldsymbol{I}(\boldsymbol{p})$ defined above and several other types of interference function are standard function, whose precise definition is given as follows [15].

Definition 2.3. A function $I(p)$ is *standard* if for all $p \geqslant 0$, the following properties are satisfied:

- *Positivity:* $I(p) > 0$,
- *Monotonicity:* If $p \geqslant p'$, then $I(p) \geqslant I(p')$,
- *Scalability:* For all $\alpha > 1$, $\alpha I(p) > I(\alpha p)$.

Due to the nice properties associated with the standard function, synchronous iterative power control algorithm based on standard function $\boldsymbol{p} = \boldsymbol{I}(\boldsymbol{p})$, also called *standard power control algorithm*, is proposed with proved convergence to a unique fixed point [15].

Theorem 2.4. *If $I(p)$ is feasible, then for any initial power vector p, the standard power control algorithm converges to a unique fixed point p^*.*

As the common radio channel is a shared medium, each user's transmission will cause interference to others, and the interference is getting more severe with higher transmitted power. On the other hand, the selfish users try to pursue high utility by increasing their transmitted power. Therefore, conventional power control can be cast into a non-cooperative power control game (NPG). If the best response strategy is a standard function of the variable that represents the user's action, then the NPG has a unique equilibrium [16]. The idea of standard function has been used in some previous works [17,18]. For instance, in [18] which considers a cooperative cognitive radio network, secondary users serve as cooperative relays for the primary users, so that they can have the opportunity to access the wireless channel. The secondary users target at maximizing the utility defined as a function of their achievable rate minus the payment, by selecting the proper payment in the non-cooperative game. By proving the best response payment is a standard function, it is shown that the non-cooperative payment selection game has a unique equilibrium.

2.3. Equilibrium selection

2.3.1. Pareto optimality

When there is more than one equilibrium in the game, it is natural to ask whether some outperform others, and whether there exists an optimal one. Because game theory solves multi-objective optimization problem, it is not easy to define the optimality in such scenarios. For example, when players have conflicting interests with each other, an increase in one player's payoff might decrease others' payoffs. In order to define the optimality, one possibility is to compare the weighted sum of the individual payoffs, which reduces the multi-dimension problem into a one-dimension one. A more popular alternative is the Pareto optimality, which, informally speaking, is a payoff profile that no strategy can make at least one player better off without making any other player worse off.

Definition 2.4. Let $U \subseteq \mathbb{R}^N$ be a set. Then $\mathbf{u} \in U$ is **Pareto efficient** if there is no $\mathbf{u}' \in U$ for which $u_i' > u_i$ for all $i \in N$; $\mathbf{u} \in U$ is **strongly Pareto efficient** if there is no $\mathbf{u}' \in U$ for which $u_i' \geqslant u_i$ for all $i \in N$ and $u_i' > u_i$ for some $i \in N$. The *Pareto frontier* is defined as the set of all $\mathbf{u} \in U$ that are Pareto efficient.

Pareto efficiency, or Pareto optimality, has been widely used in game theory, as well as economics, engineering and social sciences. If there are more than one equilibrium candidates, usually the optimal ones in the Pareto sense are preferred. For example, in the repeated game that we will discuss later in this tutorial, a lot of equilibria may exist if certain strategies have been applied. Out of many possible choices, the ones on the Pareto frontier are superior to others. In the bargaining game, which is also a topic in the tutorial, Pareto optimality has been used as an axiom to define the bargaining equilibrium in this game.

However, because of the selfish nature of players in a non-cooperative game, an NE may be Pareto inefficient when compared with payoff profiles of all possible outcomes. Several methods to improve an inefficient equilibrium will be introduced later.

2.3.2. Equilibrium refinement

The NE is an important concept, but it is a relatively weak criterion in some sense, especially when the game exhibits more complex structures. Hence, there may exist multiple equilibria according to the Nash criterion, but some of them may not be a desirable or reasonable outcome, and it is necessary to narrow down, or *refine*, the equilibrium solutions. One simple example is when the game has a symmetric structure, sometimes we may be more interested in symmetric equilibria where every player adopts the same strategy. Take the multiple access game (Table 2) as an example, we have known that there are two pure strategy NE and one mixed strategy NE, but if symmetry in the equilibrium is required, only the mixed strategy NE will be the outcome.

Things become more involved as the game becomes more complex. In the strategic-form game, it is assumed that all players move *simultaneously*; however, it is possible that players move *sequentially*, and they are informed of the previous moves. This is the **extensive-form game**, also known as the **multi-stage game** and the **dynamic game**. When all information of history is perfectly known to all players, it is called a game **with perfect information**; otherwise, it is a game **with imperfect information**.

Let us begin with a simple channel selection game illustrated by the tree in Fig. 2, which is an extensive-form game with perfect information. In this game, player 1 moves first to select one channel B_1 or B_2, and then player 2 chooses one channel after observing player 1's selection. Different bands yield different communication gains to players; for example, we assume player 1 gains 4 and 2 from exclusive usage of channel B_1 and B_2, respectively, and player 2 gains 6 and 4, respectively. However, if both players select the same channel, they can only receive half of the gain due to time sharing of the band. The payoff pair for each possible action history is given in the figure accordingly. At the beginning of the game, player 2 has four strategies, namely, B_1B_1, B_1B_2, B_2B_1, and B_2B_2, where each strategy specifies the two actions given player 1 chooses B_1 or B_2, respectively. For example, strategy B_2B_1 implies that player 2 will choose B_2 if player 1 chooses B_1, and choose B_1 if player 1 chooses B_2. The other strategies are defined in the same way. Then, this new game can be reformed to an equivalent strategic-form game given by Table 3, from which it is easy to verify that three Nash equilibria exist, i.e., $\{B_2,B_1B_1\}$,$\{B_1,B_2B_1\}$, and $\{B_1,B_2B_2\}$.

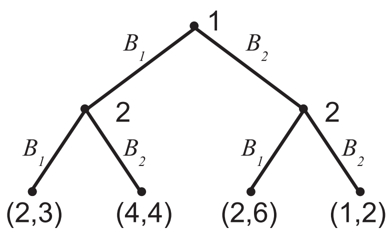

Fig. 2. Extensive-form game with perfect information.

Table 3
Equivalent strategic-form game.

	B_1B_1	B_1B_2	B_2B_1	B_2B_2
B_1	$\underline{2}$,3	$\underline{2}$,3	$\underline{4}$,$\underline{4}$	$\underline{4}$,$\underline{4}$
B_2	$\underline{2}$,$\underline{6}$	1,2	2,$\underline{6}$	1,2

Let us take a closer look at the equilibrium $\{B_2,B_1B_1\}$. This is an equilibrium because B_2 is the best response to a "threat" made by player 2, who pledges to always choose B_1. However, this is an *incredible threat*, because it is B_2 rather than B_1 that player 2 should choose for a higher payoff if player 1 has chosen B_1 for some reason. Therefore, although $\{B_2,B_1B_1\}$ is a Nash equilibrium viewed at the beginning of the game, as the game progresses, the strategy is not optimal any more for the subgame where player 1 has already chosen B_1. Similar analysis disqualifies $\{B_1,B_2B_2\}$ as a reasonable equilibrium. The remaining one, $\{B_1,B_2B_1\}$, guaranteeing the Nash equilibrium at any subgame of the original game, is called a **subgame perfect equilibrium**.

For an extensive-form game with finite stages, **backward induction** can be employed to obtain the subgame perfect equilibria. For the previous example, player 2's credible strategy is B_2B_1 at stage 2, which reduces the possible outcomes of game to (4,4) and (2,6). Then, player 1's best response is B_1, and $\{B_1,B_2B_1\}$ is the subgame perfect equilibrium.

When the game has imperfect information, even the concept of subgame perfect equilibria is not strong enough. The tree in Fig. 3 presents another channel selection game, where player 1 has an additional option to select a dedicated channel B_0 with gain 3, and player 2 is unable to observe whether player 1 has chosen B_1 or B_2, as indicated by a dashed line in the figure. In this case, $\{B_1,B_2\}$ and $\{B_0,B_1\}$ are two subgame perfect equilibria, but the latter is not preferable because it is based on an *implausible belief*. When making a decision, player 2 is uncertain about whether player 1 has selected B_1 or B_2. To pursue a higher expected payoff, it is easy to show that player 2 should choose B_1 only if he/she believes that player 1 selected B_2 with a probability larger than 1/5. However, from player 1's perspective, channel B_2 is always dominated by B_0. If player 1 did not choose B_0, he/she must have chosen B_1 with probability 1, which makes player 2's belief implausible. On the contrary, the other equilibrium $\{B_1,B_2\}$ passes

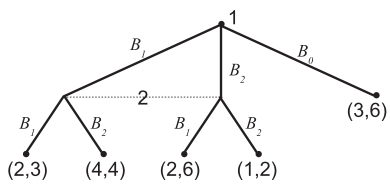

Fig. 3. Extensive-form game with imperfect information.

the plausible-belief criterion, and is called a **sequential equilibrium**.

Furthermore, when stability is an issue, other stronger refinements, such as **trembling hand equilibria** and **proper equilibria**, are needed. Interested readers are referred to [4] for more details.

2.3.3. Evolutionary equilibrium

In the above, we have discussed the concept of Pareto optimality, under which no player can improve his/her utility individually without hurting the others' benefit. The establishment of Pareto optimality usually relies on the assumption that players are fully aware of the game they are playing and others' actions in the past, and that players are rational and willing to cooperate in their moves. More often, however, players only have limited information about the other players' strategies, or they are even unaware of the game being played. In addition, not all players are rational and always following their optimal strategy. Under all these circumstances, will there exist an NE? If so, how many? If there is more than one equilibrium, which one will be selected, if some players play out-of-equilibrium strategy?

These questions can be answered by the evolutionary game theory (EGT), with evolutionary equilibrium as the core concept. The idea of evolutionary games was inspired by the study of ecological biology. EGT differs from classical game theory by focusing on the dynamics of strategy change more than the properties of strategy equilibria. It can tell us how a rational player should behave to approach a best strategy against a small number of players who do not follow the best strategy, and thus EGT can better handle the unpredictable behavior of players.

More specifically, assume the game is played by a set of homogeneous players, who have the same form of utility function $u_i(\cdot) = u(\cdot)$ and action space [19]. Assume that the players are programmed to play pure strategies at each time. At time t, the number of players taking pure strategy a_i is $p_{a_i}(t)$, then the population size is $p(t) = \sum_{a_i \in A_i} p_{a_i}(t)$, and the population share that is taking strategy a_i becomes $x_{a_i}(t) = p_{a_i}(t)/p(t)$. Replicator dynamics in continuous time format can be written by

$$\dot{x}_{a_i} = [u(a_i, x_{-a_i}) - \bar{u}(x)]x_{a_i}, \tag{14}$$

where $u(a_i, x_{-a_i})$ denotes the average payoff of players using a_i, and $\bar{u}(x)$ denotes the average payoff of the entire population. Eq. (14) means that the higher payoff a strategy a_i achieves, the population share using a_i will grow, and the growth rate is proportional to the difference between a_i's average payoff and the average payoff in the entire population.

Using the replicator dynamics, players can adapt their strategy and converge to the evolutionarily stable strategy (ESS). We provide the definition of ESS for a two-player game as follows.

Definition 2.5. In a symmetric strategic game with two players $G = \langle\{1,2\},(A,A),(u_i)\rangle$, where $u_1(a,a') = u_2(a',a) = u(a,a')$ for some utility function u, an evolutionarily stable strategy (ESS) of G is an action $a^* \in A$ for which $u(a,a^*) \leqslant u(a^*,a^*)$, and $u(a,a) < u(a^*,a)$ for every best response $a \in A$ with $a \neq a^*$.

These conditions ensure that as long as the fraction of mutants who play a is not too large, the average payoff of a will fall short of a^*. Since strategies with a higher payoff value are expected to propagate faster, evolution will cause the players not to use mutation strategy a, and instead use ESS. Therefore, in a game with a few players taking out-of-equilibrium strategies, the equilibrium after the convergence of the replicator dynamics is the ESS. Actually, the first condition in Definition 2.5 says that the ESS must first be a Nash equilibrium, and the second condition can be viewed as a selection criterion that ensures the stability of the equilibrium under strategy mutation.

Since the EGT with the replicator dynamics characterizes the change of the population sizes, we can apply EGT to cognitive radio networking, which can provide guidelines for upgrading existing networking protocols and determining operating parameters related to new protocols, and thus achieve reconfigurability to the time-varying radio environment with stability guarantee. In [20], an evolutionary game modeling for cooperative spectrum sensing is proposed, where selfish users tend to overhear the others' sensing results and contribute less to the common task. The behavior dynamics of secondary users are studied using the replicator dynamics equations. In the distributed implementation derived from the replicator dynamics, users update their strategies by exploring different actions at each time, adaptively learning during the strategic interaction, and approaching the best response strategy. Another evolutionary game theoretic approach in cognitive radio networking is considered in [21], where sensor nodes act as players and interact in randomly drawn pairs in an impulse radio UWB sensor network. Each player adapts the pulse repetition frequency value upon the observation about the bit error rate of the other player in the interactive pair. It is shown that through the interaction-learning process, a certain QoS can be guaranteed.

2.4. How to improve the inefficient NE

2.4.1. Pricing

From a network designer's point of view, he/she would like to have a satisfying social welfare, which can be defined as maximizing the sum of all users' payoff values (utilitarian type), or maximizing the minimum payoff value among all users' payoffs (egalitarian type). However, this contradicts with users' selfish nature, if they do not work towards a common goal. In a cognitive radio network consisting of selfish users competing for spectrum resources, the social optimum is usually not achieved at the NE, since selfish users are only interested in their own benefit.

In order to study the optimality of the non-cooperative game outcomes, *Price of Anarchy* is an important measure, which is the ratio between the worst possible NE and social optimum that can be achieved only if a central authority is available. By studying the bounds on the price of anarchy, we can gain better understanding about the NE of non-cooperative games in cognitive radio networks.

In [22], the price of anarchy is extensively studied for non-cooperative spectrum sharing games, in which the channel assignment for the access points (APs) is studied for WiFi networks. The price of anarchy in this scenario represents the ratio between the number of APs assigned spectrum channels in the worst NE and the optimal number of covered APs if a central authority assigns the channels. The analysis of the NE in spectrum sharing games is performed by considering it as a maximal coloring problem. The theoretical bounds on the price of anarchy are derived for the scenarios of different number of spectrum buyers and sellers. One interesting finding is that the price of anarchy is unbounded in general spectrum sharing games unless certain constraints are applied such as the distribution of the users. Similarly, in [23], the price of anarchy is studied for spectrum assignment in a local-bargaining scenario.

To improve the efficiency of the NE of non-cooperative games in cognitive radio networks, pricing can be introduced when designing the non-cooperative game, since selfish network users will be guided to a more efficient operating point [16]. Intuitively, pricing can be viewed as the cost of the services or resources a network user receives, or the cost of harm the user imposes on other users, in terms of performance degradation, revenue deduction, or interference. As the selfish network users only optimize their own performance, their aggressive behavior will degrade the performance or QoS of all the other users in the network, and hence deteriorate the system efficiency. By adopting an efficient pricing mechanism, selfish users will be aware of the inefficient NE, encouraged to compete for the network resources more moderately and efficiently, bringing more benefit for all network users and a higher revenue for the entire network.

Linear pricing which increases monotonically with the transmit power of a user has been widely adopted [17,24], because of its implementation simplicity and a reasonable physical meaning. In [17], a network-user hierarchy model consisting of a spectrum manager, service provider and end users for dynamic spectrum leasing is proposed for joint power control and spectrum allocation. When optimizing their payoff, the end users trade off the achievable data rate and the spectrum cost through transmission power control. With a proper pricing term, which is defined as a linear function of the spectrum access cost and transmission power, efficient power control can be achieved which alleviates interference between end users; moreover, the revenue of the service provider is maximized. In [24], the service provider charges each user a certain amount of payment for each unit of the transmitting power on the uplink channel in wide-band cognitive radio networks for revenue maximization, while ensuring incentive compatibility for the users. In [25], the authors further point out that most existing pricing techniques, e.g., a linear pricing function with a fixed pricing factor for all users, can usually improve the equilibrium by pushing it closer to the Pareto optimal frontier. However, they may not be (Pareto) optimal, and not suitable for distributed implementation, as they require global information. Therefore, a user-dependent linear pricing function which drives the NE close to the Pareto optimal frontier is proposed [25], through analysis of the Karush–Kuhn–Tucker conditions. The optimal pricing factor for a link only depends on its neighborhood information, so the proposed spectrum management can be implemented in a distributed way.

More sophisticated nonlinear pricing function can also be used, according to the specific problem setting and requirements. In an underlay spectrum sharing problem [26] where secondary users transmit in the licensed spectrum concurrently primary users, secondary users' transmission is constrained by the interference temperature limit (ITL),

$$\sum_{j=1}^{N} p_j h_{mj} \leqslant Q_m^{\max}, \tag{15}$$

where p_j denotes secondary user j's transmit power, h_{mj} denotes the channel gain from secondary user j to primary user m, and $Q_m^{\max}$ denotes the ITL of primary user m. Thus, an exponential part, e^{ω_m}, with $\omega_m = \delta\left(\frac{\sum_{j=1}^{N} p_j h_{mj} - Q_m^{\max}}{Q_m^{\max}}\right)$, is introduced as a pricing factor into the pricing function, as well as another part representing the interference to other secondary users. When the ITL is violated, the utility function will decrease dramatically. This way, efficient secondary spectrum sharing will be achieved, with sufficient protection of primary transmission. In the spectrum sharing problem considered in [27], each wireless transmitter selects a single channel from multiple available channels and the transmission power. To mitigate the effects of interference externalities, users should exchange information that can reflect interference levels. Such information is defined by interference "prices",

$$\pi_k^{\varphi(k)} = \left| \frac{\partial u_k(\gamma_k^{\varphi(k)}(\boldsymbol{p}^{\varphi(k)}))}{\partial\left(\sum_{j\neq k} p_j^{\varphi(k)} h_{jk}^{\varphi(k)}\right)} \right|, \tag{16}$$

where $u_k(\gamma_k^{\varphi(k)}(\boldsymbol{p}^{\varphi(k)}))$ represents users' utility function that is a concave and increasing function of the received SINR $\gamma_k^{\varphi(k)}(\boldsymbol{p}^{\varphi(k)})$, and $\sum_{j\neq k} p_j^{\varphi(k)} h_{jk}^{\varphi(k)}$ represents the interference at receiver k. So the interference price in (16) indicates the marginal loss/increase in user k's utility if its received interference is increased/decreased by one unit. With the definition of interference price, user k's new utility becomes the net benefit

$$u_k(\gamma_k^{\varphi(k)}(\boldsymbol{p}^{\varphi(k)})) - p_k^{\varphi(k)} \sum_{j\neq k} \pi_j^{\varphi(k)} h_{jk}^{\varphi(k)}. \tag{17}$$

It is shown [27] that the proposed algorithm considering interference price always outperforms the heuristic algorithm where each user only picks the best channel without exchanging interference prices, and the iterative water-filling algorithm where users do not exchange any information.

2.4.2. Repeated game and folk theorems

In order to model and analyze long-term interactions among players, the repeated game model is used where the game is played for multiple rounds. A repeated game is a special form of an extensive-form game in which each stage is a repetition of the same strategic-form game. The number of rounds may be finite or infinite, but usually the infinite case is more interesting. Because players care about not only the current payoff but also the future payoffs, and a player's current behavior can affect the other players' future behavior, cooperation and mutual trust among players can be established.

Definition 2.6. Let $\langle N,(A_i),(u_i)\rangle$ be a strategic game. A δ-**discounted infinitely repeated game** of $\langle N,(A_i),(u_i)\rangle$ $(0 < \delta < 1)$ is an extensive-form game with perfect information and simultaneous moves in which

- the set of players is N;
- for every value of t, the chosen action a^t may depend on the history $(a^1,a^2,\ldots,a^{t-1})$;
- the set of actions available to any player i is A_i, regardless of any history;
- the payoff function for player i is the discounted average of immediate payoffs from each round of the repeated game, i.e., $u_i(a^1,a^2,\ldots,a^t,\ldots) = (1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}u_i(a^t)$.

Note that the discount factor δ measures how much the players value the future payoff over the current payoff. The larger the value is, the more patient the players are. There are alternative ways to define an infinitely repeated game without the use of the discount factor, such as the **limited of means infinitely repeated game** and the **overtaking infinitely repeated game**, but they are rarely applied to cognitive radio networks.

In order to stimulate cooperation among selfish players, the so-called "grim trigger" strategy is a common approach. In the beginning, all players are in the cooperative stage, and they continue to cooperate with each other until someone deviates from cooperation. Then, the game jumps to the punishment stage where the deviating player will be punished by other peers, and there will be no cooperation forever. A less harsh alternative, also known as the "punish-and-forgive" strategy, is similar except for the limited punishment where deviation is forgiven and cooperation resumes after long enough punishment. Because cooperation is often more beneficial, the threat of punishment will prevent players from deviation, and hence cooperation is maintained. This is formally established by folk theorems, a family of theorems characterizing equilibria in repeated games. To begin with, we give some definitions.

Definition 2.7. Player i's **minmax payoff** in a strategic game $\langle N,(A_i),(u_i)\rangle$ is

$$\min_{a_{-i}\in A_{-i}}\max_{a_i\in A_i} u_i(a_i,a_{-i}). \tag{18}$$

The minmax payoff is the lowest payoff that the other players can force upon player i, and can be used as the threat of punishment. Another option is to threat to use the NE as punishment. It is known that the NE gives at least the minmax payoff to any player, so the Nash threat is usually weaker than the minmax threat.

Definition 2.8. A vector $v\in\mathbb{R}^N$ is a **payoff profile of** $\langle N,(A_i),(u_i)\rangle$ if there is an outcome $a\in A = \times_{i\in N}A_i$ such that $v = u(a)$. A vector $v\in\mathbb{R}^N$ is referred to as a **feasible payoff profile** of game $\langle N,(A_i),(u_i)\rangle$ if it is a convex combination of payoff profiles of outcomes in A. Denote the minmax payoff of player i as v_i^0. A payoff profile v is (**strictly**) **individually rational** if $v_i > v_i^0$ for all $i\in N$.

Depending on the type of the equilibrium (the NE or the subgame perfect equilibrium), the length of punishment (grim-trigger or punish-and-forgive), the punishment payoff (the minmax threat or the Nash threat), and the criterion of the infinitely repeated game (δ-discounted or others), folk theorems vary slightly from case to case. Here, we only pick one of them to present, that is, the perfect grim-trigger-strategy folk theorem with Nash threats for the discounting criterion. The proof and other variants can be found in [28] and [4].

Theorem 2.5. *For any feasible and strictly payoff profile v such that $v_i > v_i^N$, for all $i\in N$, and v_i^N being the payoff of the stage-game Nash equilibrium, there exists $\underline{\delta}\in(0,1)$, such that for all $\delta\in[\underline{\delta},1)$, there exists a repeated-game strategy profile which is a subgame perfect equilibrium of the repeated game and yields the expected payoff profile v.*

In Fig. 4, we illustrate the feasible utility region of a repeated game with two players that can be envisioned in a simplified power control problem for Gaussian interference channel with two power levels. Point D can represent that both users transmit with very high power levels and suffer from severe interference, point A or C represents that one user transmits with high power while the other uses low transmit power, and point B represents that the two users cooperate by transmitting with low power levels to alleviate interference and improve utility. If the game is only played for only one round, the NE will correspond to D and thus is very inefficient; however, if the game is played for multiple rounds, any point lying in the convex hull (shaded area in Fig. 4) can be achievable, according to the folk theorems. In other words, the efficiency of the game can be greatly improved.

Other than the grim trigger strategy and the punish-and-forgive strategy, "tit-for-tat" and "fictitious play" are also popular strategies in a repeated game. Both of them involve learning from opponents. When using the "tit-for-tat" strategy, a player chooses an action based on the outcome of the very last stage of the game, for example, he/she decides to cooperate only when all the other players cooperated in the last time. If the "fictitious play" strat-

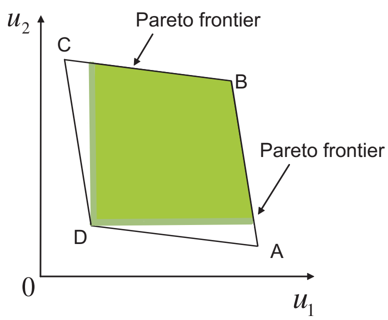

Fig. 4. The feasible utility region of a repeated two-player game.

egy is used, a player learns the empirical frequency of each action of the other players from all history outcomes, and then chooses the best strategy accordingly assuming the opponents are playing stationary strategies.

Examples of cooperation enforcement and adaptive learning in repeated games can be found in context of cognitive radio networks, e.g., [29–33] and so on. For instance, it is shown in [29] that any achievable rate in a Gaussian interference channel, in which multiple unlicensed users share the same band, can be obtained by piece-wise constant power allocations in that band with the number of segments at most twice of the number of users. The paper also shows that only the pure strategy Nash equilibria exist, and under certain circumstances, spreading power evenly over the whole band is the unique NE that is often inefficient. Then, a set of Pareto optimal operating points are made possible by repeated game modeling and punishment-based strategies. Extending this work to time-varying channels, [30] also applies the repeated game framework to achieve better performance than the inefficient NE. In this paper, users exchange instantaneous channel state information and cooperatively share the spectrum using the "punish-and-forgive" strategy.

2.4.3. Correlated equilibrium

In deriving the NE of a game, players are assumed to take their strategies independent of the others' decisions. When they no longer do so, for instance, following the recommendation of a third party [34–37], the efficiency of the game outcome can be significantly improved. In the following, we will discuss the concept of correlated equilibrium, if the recommended strategy further satisfies a certain property.

Correlated equilibrium is a more general equilibrium concept than the NE, where players can observe the value of a public signal and choose their actions accordingly. When no player would deviate from the recommended strategy, given that the others also adopt the recommendation, the resulting outcome of the game is a correlated equilibrium.

Take the multiple access game shown in Table 2 as an example. Assume now both players observe a public signal from a third party, who draws one of three cards labeled (D,C), (C,D), and (C,C) uniformly (with probability 1/3 for each card). The players choose their actions according to the one assigned from the card drawn by the third party. If a player is assigned D, he/she will not deviate, since the payoff of 6 is the highest, assuming the other player also played the assigned action. If action C is assigned, both cards (C,D) and (C,C) can be drawn, with equal probability. Thus, the expected payoff of action D is $0 \cdot (\frac{1}{2}) + 6 \cdot (\frac{1}{2}) = 3$, while the expected payoff of action C is $3 \cdot (\frac{1}{2}) + 5 \cdot (\frac{1}{2}) = 4$. So the player will not deviate from action C, either. Clearly, no player has an incentive to deviate, and the resulting outcome forms a correlated equilibrium. Note that in this equilibrium, the expected payoff of a player is $\frac{14}{3}$, greater than that of the pure-strategy NE or mixed-strategy NE, as can be calculated in Section 2.1. Hence, correlated equilibrium can be more system efficient than NE.

After explaining the multiple access game example, we now give the definition of correlated equilibrium [4].

Definition 2.9. A **correlated equilibrium** of a strategic game $\langle N,(A_i),(u_i)\rangle$ consists of

- a finite probability space (Ω,π), where Ω is a set of states and π is a probability measure on Ω,
- an information partition $\mathscr{P}_i$ of Ω, $\forall i \in N$, and
- a function $\sigma_i : \Omega \to A_i$, which represents player i's strategy and maps an observed state to an action, with $\sigma_i(\omega) = \sigma_i(\omega')$, $\omega, \omega' \in P_i$, for some $P_i \in \mathscr{P}_i$,

such that

$$\sum_{\omega\in\Omega} \pi(\omega)u_i(\sigma_{-i}(\omega), \sigma_i(\omega)) \geqslant \sum_{\omega\in\Omega} \pi(\omega)u_i(\sigma_{-i}(\omega), \tau_i(\omega)), \quad (19)$$

for all $i \in N$, and any strategy function $\tau_i(\cdot)$.

From this definition, we know that action $\sigma_i(\omega)$ in equilibrium, where state ω occurs with a positive probability, is optimal to player i, given the other players' strategies also following the correlated equilibrium. Since the equilibrium points are defined by a set of linear constraints in (19), there can exist multiple correlated equilibrium points of a game. For instance, in the multiple access game we considered above, drawing the three cards (D,C), (C,D), and (C,C) with probability 0.2, 0.2, and 0.6, respectively, is another equilibrium. Moreover, the set of mixed strategy NE is a subset of the set of correlated equilibria.

In order to adjust their strategies and converge to the set of correlated equilibria, players can track a set of "regret" values for strategy update [38]. The regret value can be defined by

$$\mathscr{R}_i^T(a_i, a_i') = \max\left\{\frac{1}{T}\sum_{t\leqslant T}[u_i^t(a_i', a_{-i}) - u_i^t(a_i, a_{-i})], 0\right\}, \quad (20)$$

where $u_i^t(a_i', a_{-i})$ represents the payoff of player i at time t by taking action a_i' against the other players taking action a_{-i}. Therefore, the regret value denotes the average payoff that player i would have obtained, if he/she had adopted action a_i' every time in the past instead of action a_i. If the regret value is smaller than 0, meaning adopting action a_i brings a higher average payoff, then there will be no regret, and thus the regret value is lower bounded by 0. According to the regret value, players can update their strategies by adjusting the probability of taking different actions. If player i's action at time t is $a_i^t = a_i$, then the probability of taking action $a_i' \neq a_i$ at time $t+1$ is updated by

$$p_i^{t+1}(a_i') = \frac{1}{\mu}\mathscr{R}_i^t(a_i, a_i'), \quad (21)$$

otherwise, it is updated by

$$p_i^{t+1}(a_i) = 1 - \sum_{a_i' \neq a_i} p_i^{t+1}(a_i'). \quad (22)$$

If all players learn their strategies according to (21) and (22), then their strategies will converge to the set of correlated equilibria almost surely, as time goes to infinity [38].

The concept of correlated equilibrium and regret learning has been used to design dynamic spectrum access protocols [35,36], where a set of secondary users compete for the access of spectrum white space. Users' utility function is defined as the average throughput in [35], and in [36] a term representing performance degradation due to excess access and collisions is further included. Since the common history observed by all users can serve as a natural coordination device, users can pick their actions based on the observation about the past actions and payoff values, and achieve better coordination with a higher performance.

3. Economic games, auction games, and mechanism design

As game theory studies interaction between rational and intelligent players, it can be applied to the economic world to deal with how people interact with each other in the market. The marriage of game theory and economic models yields interesting games and fruitful theoretic results in microeconomics and auction theory. On one hand, they can be regarded as applied branches of game theory which builds on top of key game theoretic concepts such as rationality and equilibria. Often, players are sellers and buyers in the market (e.g., firms, individuals, and so on), payoff functions are defined as the utility or revenue that players want to maximize, and equilibrium strategies are of considerable interest. On the other hand, they are distinguished from fundamental game theory, not only because additional market constraints such as supply and demand curves and auction rules give insight on market structures, but also because they are fully-developed with their own research concerns. In fact, the research on the Cournot model, one of market equilibria, dates back much earlier than game theory literally exists as a unique field. Hence, we make a separate section to address those economic games, so as to respect the distinction of these games and to highlight their intensive use in cognitive radio networks.

The application of games in economy into cognitive radio networks has the following reasons. First, economic models are suitable for the scenario of the secondary spectrum market where primary users are allowed to sell unused spectrum rights to secondary users. Primary users, as sellers, have the incentive to trade temporarily unused spectrum for monetary gains, while secondary users, as buyers, may want to pay for spectrum resources for data transmissions. The deal is made through pricing, auctions, or other means. Second, these games in economy do not confine themselves to the scenario with explicit buyers and sellers, and the ideas behind can be extended to some cognitive radio scenarios other than secondary spectrum markets. One example is that the Stackelberg game, originally describing an economic model, has been generalized to a strategic game consisting of a leader and a follower. More details and other examples will be discussed in this section. Third, as cognitive radio goes far beyond technology and its success will highly rely on the combination of technology, policy, and markets, it is of extreme importance to understand cognitive radio networks from the economic perspective and develop effective procedures (e.g., auction mechanisms) to regulate the spectrum market.

To highlight the underlining economic features in these games, we will use p and q to refer to prices and quantities in this section. In general, p and q are interrelated given a certain market, and their relation can be modeled by the demand curve and the supply curve. For example, at the given market price p, the amount of a good that buyers are willing to buy is $q = \mathscr{D}(p)$, whereas the amount that sellers are willing to sell is $q = \mathscr{S}(p)$. Functions $\mathscr{D}(\cdot)$ and $\mathscr{S}(\cdot)$ are known as the demand function and the supply function. Moreover, if the quantity q is fixed in the market, the price that buyers are willing to pay can be derived by the inverse demand function, i.e., $p = \mathscr{D}^{-1}(q)$, and similarly, the price charged by sellers is given by the inverse supply function, i.e., $p = \mathscr{S}^{-1}(q)$. Often, the demand function is a non-increasing function of p and the supply function is a non-decreasing function of p.

3.1. Oligopolistic competition

When the market is fully competitive, the market equilibrium, denoted by (p^*, q^*), is the intersection of the demand curve and the supply curve,

$$q^* = \mathscr{D}(p^*) \text{ and } q^* = \mathscr{S}(p^*). \tag{23}$$

The other extreme is monopoly, when only one firm controls all over the market of one product. Assume the cost associated with the quantity q is $\mathscr{C}(q)$, the firm can maximize the profit which is revenue minus cost,

$$u(q) = q\mathscr{D}^{-1}(q) - \mathscr{C}(q), \tag{24}$$

by applying the first-order condition

$$\frac{\partial u(q)}{\partial q} = \mathscr{D}^{-1}(q) + q\frac{\partial \mathscr{D}^{-1}(q)}{\partial q} - \frac{\partial \mathscr{C}(q)}{\partial q} = 0. \tag{25}$$

Lying between the full competition and no competition (monopoly), oligopoly is more complicated and interesting, which is defined as a market with only a few firms and with substantial barriers to entry in economics. Because the number of firms is limited, each one can influence the price and hence affect other firms; for example, their strategies are to decide quantity or the price of goods supplied to the market. The interaction and competition between different firms can be well modeled by game theory, and several models have been proposed long before. These models share common attributes including price–quantity relations, profit-maximizing goals, and first-order optimality, but they are different in actions (quantities vs. prices), structures (simultaneous moves vs. sequential moves), or forms (competition vs. cooperation). In what follows, we give a brief summary of these games. We assume there are only two competing firms (i.e., a duopoly) for convenience, but it is straightforward to generalize to the scenario with multiple firms.

In the *Cournot game*, oligopoly firms choose their quantities q_1, q_2 independently and simultaneously. Because the market price depends on the total quantity, each firm's action directly affects others' profits. The market price is determined by the inverse demand function $p = \mathscr{D}^{-1}(q_1 + q_2)$. Assume the cost associated with a production quantity q_i is $\mathscr{C}_i(q_i),\ i = 1, 2,$ for the two firms. The utility function of each firm is revenue minus cost,

$$u_i(q_i) = q_i\mathscr{D}^{-1}(q_1 + q_2) - \mathscr{C}_i(q_i), \quad i = 1, 2. \tag{26}$$

Hence, the equilibrium of this game (q_1^*, q_2^*) is the solution to the following equations derived from first-order conditions:

$$\begin{cases} \frac{\partial u_1(q_1)}{\partial q_1} = \mathscr{D}^{-1}(q_1+q_2) + q_1 \frac{\partial \mathscr{D}^{-1}(q_1+q_2)}{\partial q_1} - \frac{\partial \mathscr{C}_1(q_1)}{\partial q_1} = 0, \\ \frac{\partial u_2(q_2)}{\partial q_2} = \mathscr{D}^{-1}(q_1+q_2) + q_2 \frac{\partial \mathscr{D}^{-1}(q_1+q_2)}{\partial q_2} - \frac{\partial \mathscr{C}_2(q_2)}{\partial q_2} = 0. \end{cases} \tag{27}$$

In the *Bertrand game*, firms also decide their actions independently and simultaneously, but their decisions are prices p_1, p_2 and their produce capacity is unlimited. Although it looks like the Cournot game, the outcome is significantly different. Since the firm with the lower price will occupy the entire market, firms will try to reduce their price until hitting the bottom line with zero profit. Hence, the equilibrium of this game is trivial. A modification of the game is to assume each firm produces a somewhat differentiated product. The demand function of product 1 is $\mathscr{D}_1(p_1, p_2)$, a decreasing function of p_1 and often an increasing function of p_2. Similarly, we can define $\mathscr{D}_2(p_1, p_2)$. Then, the equilibrium price can be found through first-order conditions that maximize the profit given by

$$u_i(p_i) = p_i \mathscr{D}_i(p_1, p_2) - \mathscr{C}_i(\mathscr{D}_i(p_1, p_2)), \quad i = 1, 2. \tag{28}$$

In the *Stackelberg game*, firms still choose their quantities q_1, q_2 as in the Cournot game, but the two firms make decisions sequentially rather than simultaneously. The firm that moves first is called the *leader*, and the other is called the *follower*. Without loss of generality, we assume firm 1 is the leader in this game. Because firm 2 takes action after firm 1 announces the quantity q_1, the best response of firm 2 can be derived from

$$\frac{\partial u_2(q_2)}{\partial q_2} = \mathscr{D}^{-1}(q_1+q_2) + q_2 \frac{\partial \mathscr{D}^{-1}(q_1+q_2)}{\partial q_2} - \frac{\partial \mathscr{C}_2(q_2)}{\partial q_2} = 0, \tag{29}$$

which is essentially a function of q_1. We denote it by $q_2^*(q_1)$ to emphasize that it is firm 2's best response to the announced quantity q_1. Knowing that firm 2 will choose the quantity $q_2^*(q_1)$, firm 1 can maximize the profit by setting q_1 according to the first-order condition,

$$\frac{\partial u_1(q_1)}{\partial q_1} = \frac{\partial (q_1 \mathscr{D}^{-1}(q_1 + q_2^*(q_1)) - \mathscr{C}_1(q_1))}{\partial q_1} = 0. \tag{30}$$

This process is known as the *backward induction*. If firm 1 chooses the Cournot equilibrium quantity, the best response of firm 2 will also be the Cournot equilibrium quantity. Because the optimal quantity from (30) works better than (at least equal to) the Cournot equilibrium, the leader takes an advantage from the asymmetric structure.

In the *Cartel maintenance game*, things are quite different because firms no longer compete with each other but cooperate with each other. In general, they can reduce the output, which leads to higher prices and higher profits for each firm. One example is the OPEC that manipulates the stability of international oil price. In order to enforce cooperation among selfish firms, the Cartel maintenance can be modeled as a repeated game that has been introduced earlier. From the firms' perspective, cooperation in the form of Cartel reduces competition and improves their profits, but in reality, it is harmful to economic systems and hence is forbidden by antitrust laws in many countries.

In what follows, we will show some examples on how these microeconomic concepts inspire research in cognitive radio networks. Depending on the assumptions and structures of spectrum markets, different models can be applied.

The spectrum market in [39] consists of one primary user and multiple secondary users who compete for spectrum resources. Secondary user i requests a quantity q_i for the allocated spectrum size, and the price is determined by the inverse supply function $\mathscr{S}^{-1}(\sum_{i\in N} q_i)$. This is essentially a Cournot game, but the players in the game are buyers instead of sellers in the original setting. With the inverse supply function in the paper specified as

$$\mathscr{S}^{-1}\left(\sum_{i\in N} q_i\right) = c_1 + c_2 \left(\sum_{i\in N} q_i\right)^{c_3}, \tag{31}$$

where c_1, c_2, c_3 are non-negative constants and $c_3 \geqslant 1$, the payoff is defined as

$$u_i(q_i) = u_i^0 q_i - q_i \left(c_1 + c_2 \left(\sum_{j\in N} q_j\right)^{c_3}\right), \tag{32}$$

where u_i^0 is the effective revenue per unit bandwidth for user i. The equilibrium can be derived from the first-order condition, i.e.,

$$u_i^0 - c_1 - c_2 \left(\sum_{j\in N} q_j^*\right)^{c_3} - c_2 c_3 q_i^* \left(\sum_{j\in N} q_j^*\right)^{c_3 - 1} = 0. \tag{33}$$

Another spectrum market proposed in [40] consists of multiple competing primary users and one secondary user network. This game falls into the category of Bertrand games, as primary users adjust the price of spectrum resources. To avoid the triviality that we mentioned in the introduction of the Bertrand game, the spectrum resources cannot be identical, and the authors adopt a commonly used quadratic utility function [41] for the secondary user network

$$u = \sum_{i\in N} u_i^0 q_i - \frac{1}{2}\left(\sum_{i\in N} q_i^2 + 2\nu \sum_{i\neq j} q_i q_j\right) - \sum_{i\in N} p_i q_i, \tag{34}$$

leading to linear demand functions

$$\mathscr{D}(p) = \frac{(1 + (N-2)\nu)(u_i^0 - p_i) - \nu \sum_{i\neq j}(u_j^0 - p_j)}{(1-\nu)(1+(N-1)\nu)}. \tag{35}$$

Here, p_i and q_i are the price and quantity purchased from primary user i, u_i^0 is the effective revenue per unit bandwidth, and the parameter $\nu(-1 \leqslant \nu \leqslant 1)$ reflects the *cross elasticity of demand* among different spectrum resources. Specifically, $\nu > 0$ implies substitute products, that is, one spectrum band can be used in place of another, while $\nu < 0$ implies complementary products, that is, one band has to be used together with another (like uplink and downlink). The value of ν measures the degrees of substitution or complement. In this model, the revenue is defined as the sum of monetary gains collected from the secondary network and the transmission gains of primary services, whereas the cost is defined as the performance loss to primary services due to spectrum transactions. Then, the equilibrium pricing is derived from the first-order condition.

The structure of spectrum markets could be more complicated. For instance, [42] proposes a hierarchical model in which there are two levels of markets: in the upper level, a few wireless service providers buy some spectrum bands from spectrum holders, and in the lower level, they sell these bands to end users. Wireless service providers are the players in this game who not only decide the quantity bought from spectrum holders but also the price charged to end users, and therefore, it is actually a combination of the Cournot game in the upper level with the inverse supply function (31) and the Bertrand game in the lower level with the demand function (35). The two levels are coupled in that the quantity sold to end users cannot exceed the quantity bought from license holders. The authors discuss four possible cases in the lower-level game due to quantity limitation, and conclude that only one equilibrium exists in the whole game. Another hierarchical market is proposed in [17] which considers both channel allocation and power allocation. In this model, the spectrum holder takes control of the upper-level market and hence the market fits in the monopoly model. In the lower-level game, service providers adjust the price of resources in the market, but the demand from end users comes from the equilibrium of a non-cooperative power-control game.

Just like other non-cooperative games, the Nash equilibria in these games are often inefficient due to competition among players. The difference between the equilibrium utility and the ideal maximum utility is known as the *price of anarchy*, which has been introduced in Section 2.4.1. In [39,40] and [42], the price of anarchy for the proposed spectrum market has been analyzed through theoretical derivation or demonstrated by simulation results. In addition, [40,43] shows that the efficiency can be improved by enforcing cooperation among users, that is, establishing a Cartel.

In game models, it is common to assume all players have full knowledge about each other. However, it is not always true in realistic setting such as in a cognitive radio network. For instance, one player may know nothing about other players' profits or current strategies. Therefore, to make those games implementable in spectrum markets, it is crucial to involve learning processes. The learning processes in [39,40,42] and [17] can be roughly classified into two categories. When the information of strategies is available, players always update their strategies with the best response against other players' current strategies

$$a_i(t+1) = B(a_{-i}(t)), \tag{36}$$

where action a may refer to the quantity or the price depending on the market model. When only local information is available, a gradient-based update rule can be applied, i.e.,

$$a_i(t+1) = a_i(t) + \varepsilon \frac{\partial u_i(a(t))}{\partial a_i}, \tag{37}$$

where ε is the learning rate and the partial derivative can be approximated by local observations. The convergence of the learning process has been analyzed using the Jacobian matrix, e.g., see [40].

Although originally a game between two sellers of the same product, the Stackelberg game in a broad sense can refer to any two-stage game where one player moves after the other has made a decision. The problem can be formulated as

$$\begin{aligned} &\max_{a_1\in A_1, a_2\in A_2} \quad u_1(a_1,a_2), \\ &\text{s.t. } a_2 \in \arg\max_{a_2'\in A_2} u_2(a_1,a_2') \end{aligned} \tag{38}$$

where player 1 is the leader and player 2 is the follower. Similar to the Stackelberg game in an oligopoly market, the general Stackelberg game can also be solved using the backward induction. A few applications on cognitive radio networks can be found in [44,18,45,46] and so on.

For instance, in [44], the Stackelberg game is employed to model and analyze the cooperation between a primary user and several secondary users where the primary user trade some spectrum usage to some secondary users for cooperative communications. Specifically, the primary user can choose to transmit the entire time slot on its own, or choose to ask for secondary users' cooperation by dividing one time slot to three fractions with two parameters τ_1, τ_2 $(0 \leqslant \tau_1, \tau_2 \leqslant 1)$. In the first $(1-\tau_1)$ fraction of the slot, the primary transmitter sends data to secondary users, and then they form a distributive antenna array and cooperatively transmit information to the primary receiver in the following $\tau_1\tau_2$ fraction of the slot. As rewards, the secondary users involved in the cooperative communications are granted with the spectrum rights in the rest $\tau_1(1-\tau_2)$ fraction of the slot. The primary user chooses the strategy including τ_1, τ_2, and the set of secondary users for cooperation, and then the selected secondary users will choose powers for transmission according to the primary user's strategy. As the leader of the game, the primary user is aware of secondary users' best response to any given strategy, and hence is able to choose the optimal strategy that maximizes the payoff. The cooperation structure in [18] is similar, where the major difference is that the secondary users pay for spectrum opportunities in addition to cooperative transmissions for the primary user. The implementation protocol and utility functions change, but the underlying Stackelberg game remains the same.

3.2. Auction games

Auction theory is an applied branch of game theory which analyzes interactions in auction markets and researches the game theoretic properties of auction markets. An auction, conducted by an **auctioneer**, is a process of buying and selling products by eliciting **bids** from potential buyers (i.e., **bidders**) and deciding the auction outcome based on the bids and auction rules. The rules of auction, or **auction mechanisms**, determine whom the goods are allocated to (i.e., the allocation rule) and how much price they have to pay (i.e., the payment rule).

As efficient and important means of resource allocation, auctions have quite a long history and have been widely used for a variety of objects, including antiques, real properties, bonds, spectrum resources, and so on. For example, the Federal Communications Commission (FCC) has used auctions to award spectrum since 1994, and the United States 700 MHz FCC wireless spectrum auction held in 2008, also known as Auction 73, generated 19.1 billion dollars in revenue by selling licenses in the 698–806 MHz band [47]. The spectrum allocation problem in cognitive radio networks, although micro-scaled and short-termed compared with the FCC auctions, can also be settled by auctions.

Auctions are used precisely because the seller is uncertain about the values that bidders attach to the product. Depending on the scenario, the values of different bidders to the same product may be independent (the **private values** model) or dependent (the **interdependent values** model). Almost all the existing literature on auctions in cognitive radio networks assumes private values. Moreover, if the distribution of values is identical to all bidders, the bidders are **symmetric**. Last, it is common to assume a **risk neutral** model, where the bidders only care about the expected payoff, regardless of the variance (risk) of the payoff.

There are a lot of ways to classify auctions. We start with the four simple auctions:

- **English auction**: a sequential auction where price increases round by round from a low starting price until only one bidder is left, who wins the product and pays his/her bid.
- **Dutch auction**: a sequential auction where price decreases round by round from a high starting price until one bidder accepts the price, who wins the product and pays the price at acceptance.
- **Second-price (sealed-bid) auction**: an auction where each bidder submits a bid in a sealed envelope simultaneously, and the highest bidder wins the product with payment equal to the second highest bid.
- **First-price (sealed-bid) auction**: an auction where each bidder submits a bid in a sealed envelope simultaneously, and the highest bidder wins the product with payment equal to his/her own bid.

Interestingly, the four simple auctions, albeit quite different at first glance, are indeed equivalent in some sense under certain conditions. The main idea was established in the seminal work [48] by William Vickrey, a Nobel laureate in Economics. This is summarized in the following theorem; interested readers are referred to [49] for more details.

Theorem 3.1.

(1) *The Dutch auction is strategically equivalent to the first-price sealed-bid auction, that is, for every strategy in the first-price auction, there is an equivalent strategy in the Dutch auction, and vice versa.*
(2) *Given the assumption of private values, the English auction is equivalent to the second-price sealed-bid auction.*
(3) *Given symmetric and risk-neutral bidders and private values, all four auctions yield the same expected revenue of the seller. This is a special case of a more general* **revenue equivalence theorem**.

If the assumption in Theorem 3.1 holds, it will suffice to study or adopt only one kind of auction out of the four basic forms. Usually, the second-price auction is a favorite candidate, because the procedure is simple, and more importantly, the mechanism enforces bidders to bid their true values, as stated in Theorem 3.2. In a second-price auction, bidder i whose value of the product is v_i submits a sealed bid b_i to the auctioneer. Then, the winner of the auction is $\arg\max_{j\in N} b_j$, and payoffs are

$$u_i = \begin{cases} v_i - \max\limits_{j\neq i} b_j & \text{if } i = \arg\max\limits_{j\in N} b_j, \\ 0 & \text{otherwise}. \end{cases} \tag{39}$$

Theorem 3.2. *In a second-price sealed-bid auction, it is a weakly dominant strategy to bid truthfully, i.e., $b_i = v_i$ for all $i \in N$.*

The seller plays a passive role in the auctions so far, because his/her benefit has not been taken into consideration. When the seller wants to design an auction game that has the NE with the highest possible expected revenue, it is called the **optimal auction** [50]. Assume that all bidders' values of the product are drawn from i.i.d. random variables with the same probability distribution, whose probability density function and probability distribution function are denoted by $f(v)$ and $F(v)$, respectively. Then, an optimal auction may be constructed by adding a **reserve price** on top of a second-price auction. In this case, the seller reserves the right not to sell the product to any bidder if the highest bid is lower than the reserve price.

Theorem 3.3. *Suppose the values of all bidders are private and symmetric, and the function $T(v) = v - \frac{1-F(v)}{f(v)}$ is increasing. Then setting a reserve price equal to*

$$b_0 = T^{-1}(v_0) \tag{40}$$

in a second-price auction represents an optimal auction. The function $T^{-1}(\cdot)$ is the inverse function of $T(\cdot)$, and v_0 is the seller's value of the product.

In addition, setting a reserve price is also an effective measure against **bidding ring collusion**, where some or even all of the bidders collude not to overbid each other and hence the price is kept low.

An auction becomes more involved when more than one item are simultaneously sold and bidders bid for "packages" of products instead of individual products. This is known as the **combinatorial auction** [51]. The second-price mechanism can be generalized to the Vickrey-Clarke-Groves (VCG) mechanism, which maintains the incentive to bid truthfully. The basic idea is that the allocation of products maximizes the social welfare and each winner in the auction pays the opportunity cost that their presence introduces to all the other bidders.

Beyond the basic types of auctions, there are other forms of auctions such as the clock auction, the proxy auction, the double auction, and so on. Furthermore, there are a lot of practical concerns and variants in the real-world auctions. We will not go into the details of these issues;

instead, we will focus on the auction games in cognitive radio networks in what follows.

In [52], SINR auctions and power auction mechanisms are studied subjected to a constraint on the accumulated interference power, or the so-called "interference temperature", at a measurement point, which must be below the tolerable amount of the primary system. In this auction game, the resource to sell is not the spectrum band; instead, users compete for the portion of interference that they may cause to the primary system, because the interference is the "limited resources" in this auction. Another kind of auctions has been used in [53], where spectrum sensing effort, rather than monetary payment, is the price to pay for the spectrum opportunities. The auction still follows the form of first-price and second-price sealed-bid auctions.

In the auction framework proposed in [54], users bid for a fraction of the band and the auction outcome has to satisfy the interference constraint. In this auction, each user has a piece-wise linear demand curve, and it is assumed that all users reveal demand curves to the auctioneer truthfully. Because the corresponding revenue is a piece-wise quadratic function, the auctioneer can find the revenue-maximizing point under the constraint that the allocation is conflict-free.

The cheat-proof property is a major concern in auction design, and we have mentioned that the VCG mechanism is capable of enforcing truth-telling. However, the VCG mechanism sometimes suffers from high complexity and vulnerability to collusive attacks. In [55] and [56], system efficiency is traded for low complexity using the greedy algorithm, while the authors carefully design the mechanism to guarantee that truth-telling is still a dominant strategy in this auction game.

Because of the unique feature in wireless communications that spectrum can be reused by users geographically far apart, spectrum resources are quite different from other commercial commodities in that they are often *interference-limited* rather than *quantity-limited*. From this point of view, [57] establishes a new auction not existing in economic literature. In this auction game, one spectrum band is simultaneously awarded to multiple users without interference, and the number of winners highly depends on the mutual interference among secondary users. For this so-called "*multi-winner auction*", proper auction mechanisms are developed to eliminate user collusion and improve revenue, and near-optimal algorithms are further applied to reduce the complexity.

When there are multiple sellers who also compete in selling the spectrum, the scenario can be modeled as a double auction. A truth-telling enforced double auction mechanism has been proposed in [58], and an anti-collusion double auction mechanism has been developed in [59] where history observations are employed to estimate users' private values.

3.3. Mechanism design

Auction is one of the many possible ways of selling products. If striping off any particular selling format (e.g., an auction format), we arrive at a fundamental question: what is the best way to allocate a product? This generalized allocation problem falls into the category of **mechanism design**, a field of game theory on a class of private information games. The 2007 Nobel Memorial Prize in Economic Sciences was awarded to Leonid Hurwicz, Eric Maskin, and Roger Myerson as the founders of mechanism design theory.

The distinguishing feature of mechanism design is that the game structure is "designed" by a game designer called a "**principal**" who wants to choose a mechanism for his/her own interest. Like in an auction, the players, called the "**agents**", hold some information that is not known by the others, and the principal asks the agents for some **messages** (like the bids in an auction) to elicit the private information. Hence, this is a game of incomplete information where each agent's private information, formally known as the "**type**", is denoted by θ_i, a value drawn from a set Θ_i, for $i \in N$. Based on the messages from agents, the principal makes an allocation **decision** $d \in D$, where D is the set of all potential decisions on how resources are allocated. Because agents are not necessarily honest, incentives have to be given in terms of monetary gains, known as **transfers**. The transfer may be negative values (as if paying tax) or positive values (as if receiving compensation). Then, agent i's utility is the benefit from the decision d plus a transfer, i.e., $u_i = v_i(d, \theta_i) + t_i$, which may provide agents with incentives to reveal the information truthfully. In summary, the basic insight of mechanism design is that both resource constraints and incentive constraints are coequally considered in an allocation problem with private information [60].

Definition 3.1. A **mechanism** defines a message space M_i for each agent $i \in N$ and an allocation function $(d, t_{i,i\in N}) : \times_{i\in N} M_i \to D \times \mathbb{R}^N$. For a vector of messages $\mathbf{m} \in \times_{i\in N} M_i$, $d(\mathbf{m})$ is the decision while $t_i(\mathbf{m})$ is agent i's transfer.

For a given mechanism, the agents' strategy is mapping the individual type to a message, i.e., $m : \Theta_i \to M_i$, being aware that their own utilities depend on all the reported messages. Each agent only has the prior distribution or beliefs of others' types but not the exact type; hence, this is a **Bayesian game** because players have incomplete information of others. As a result, each agent has to maximize the expected payoff which averages out the unknowns. Just as the NE presents an equilibrium for games with complete information, the **Bayesian equilibrium**, with payoffs in the NE replaced by expected payoffs, defines the equilibrium for Bayesian games. For this message game, the strategy profile, $\{m_i^*(\theta_i), i \in N\}$ is a Bayesian equilibrium if for all $i \in N$

$$\begin{aligned} &E(v_i(d(m^*(\theta)), \theta_i) + t_i(m^*(\theta))) \\ &\quad \geqslant E(v_i(d(m_i(\theta_i), m_{-i}^*(\theta_{-i})), \theta_i) \\ &\qquad + t_i(m_i(\theta_i), m_{-i}^*(\theta_{-i}))) \quad \forall m_i(\theta_i) \in M_i. \end{aligned} \tag{41}$$

Notice that the expectation is taken over the prior distribution given that agent i's type is θ_i.

Because there are unlimited possibilities of choosing message spaces and allocation functions, analyzing the equilibrium and designing the mechanism seem to be extremely challenging. However, thanks to the equivalence established in the "**revelation principle**" in Theorem 3.4, as shown below, we can restrict attention to only **"direct" mechanisms** in which the message space coincides with the type space, i.e., $M_i = \Theta_i$, and all agents will truthfully announce their types [61].

Theorem 3.4. *For any Bayesian equilibrium supported by any general mechanism, there exists an equivalent direct mechanism with the same allocation and an equilibrium* $m_i(\theta_i) = \theta_i\ \forall i \in N$.

In [62], mechanism design has been applied to multimedia resource allocation problem in cognitive radio networks. For the multimedia transmission, the utility function is defined as the expected distortion reduction resulting from using the channels. Since the system designer wants to maximize the system utility, mechanism-based resource allocation is used to enforce users to represent their private parameters truthfully. A cheat-proof strategy for open spectrum sharing has been proposed based on the Bayesian mechanism design [30,63]. In this work, a cooperative sharing is maintained by repeated game modeling, and users share the spectrum based on their channel state information. In order to provide users an incentive to reveal their private information honestly, mechanism design has been employed to determine proper transfer functions.

4. Cooperative games

In this section, we discuss two important types of cooperative spectrum sharing games, bargaining games and coalitional games, where network users have an agreement on how to fairly and efficiently share the available spectrum resources.

4.1. Bargaining games

The bargaining game is one interesting kind of cooperative games in which individuals have the opportunity to reach a mutually beneficial agreement. In this game, individual players have conflicts of interest, and no agreement may be imposed on any individual without his/her approval. Despite there are other models such as the strategic approach with a specified bargaining procedure [64], we will focus on Nash's axiomatic model which has been established in Nash's seminal paper [65], because it has been widely applied to cognitive radio networks.

For convenience, we consider the two-player bargaining game $N = \{1,2\}$, which can be extended to more players straightforwardly. For a certain agreement, player 1 receives utility u_1 and player 2 receives utility u_2. If players fail to reach any agreement, they receive utilities u_1^0 and u_2^0, respectively. The set of all possible utility pairs is the feasible set denoted by U.

Definition 4.1. A two-player **bargaining problem** is a pair $\langle U, (u_1^0, u_2^0)\rangle$, where $U \subset \mathbb{R}^2$ is a compact and convex set, and there exists at least one utility pair $(u_1, u_2) \in U$ such that $u_1 > u_1^0$ and $u_2 > u_2^0$. A **bargaining solution** is a function $(u_1^*, u_2^*) = f(U, u_1^0, u_2^0)$ that assigns a bargaining problem $\langle U, (u_1^0, u_2^0)\rangle$ to a unique element of U.

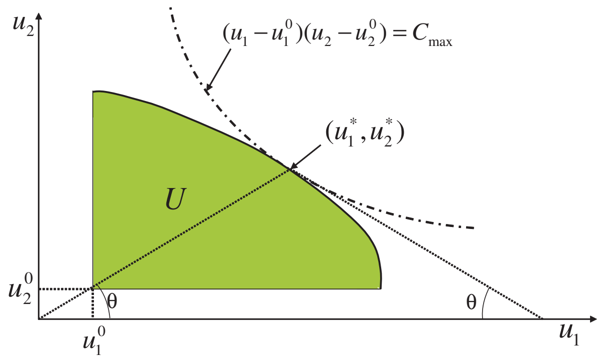

Fig. 5. NBS of a two-player game.

The axioms imposed on the bargaining solution (u_1^*, u_2^*) are listed as follows [66]:

- *Individual rationality.* $u_1^* > u_1^0$ and $u_2^* > u_2^0$.
- *Feasibility.* $(u_1^*, u_2^*) \in U$.
- *Pareto efficiency.* If $(u_1, u_2),\ (u_1', u_2') \in U,\ u_1 < u_1'$, and $u_2 < u_2'$, then $f(U, u_1^0, u_2^0) \neq (u_1, u_2)$.
- *Symmetry.* Suppose a bargaining problem is symmetric, i.e., $(u_1, u_2) \in S \Leftrightarrow (u_2, u_1) \in S$ and $u_1^0 = u_2^0$. Then, $u_1^* = u_2^*$.
- *Independence of irrelevant alternatives.* If $(u_1^*, u_2^*) \in U' \subset U$, then $f(U', u_1^0, u_2^0) = f(U, u_1^0, u_2^0) = (u_1^*, u_2^*)$.
- *Independence of linear transformations.* Let U' be obtained from U by the linear transformation $u_1' = c_1 u_1 + c_2$ and $u_2' = c_3 u_2 + c_4$ with $c_1, c_3 > 0$. Then, $f(U', c_1 u_1^0 + c_2, c_3 u_2^0 + c_4) = (c_1 u_1^* + c_2, c_3 u_2^* + c_4)$.

Theorem 4.1. *There is a unique bargaining solution satisfying all the axioms above, which is given by*

$$(u_1^*, u_2^*) = \arg \max_{(u_1,u_2)\in U, u_1 > u_1^0, u_2 > u_2^0} (u_1 - u_1^0)(u_2 - u_2^0). \tag{42}$$

This is called the Nash bargaining solution (NBS).

Fig. 5 illustrates the feasible utility region of a two-player game. The shaded area U represents the feasible range of u_1 and u_2, and the NBS corresponds to point (u_1^*, u_2^*) in the figure, where $C_{\max}$ is the largest value of $(u_1 - u_1^0)(u_2 - u_2^0)$ for the feasible set U. The meaning of the NBS is that after the players are assigned with the minimal utility, the remaining welfare are divided between them in a ratio equal to the rate at which the utility can be transferred [67].

Remarks.

(1) The NBS is well-defined. Since the function $(u_1 - u_1^0)(u_2 - u_2^0)$ is strictly quasi-concave on $\{(u_1, u_2) \in U : u_1 > u_1^0, u_2 > u_2^0\}$, a non-empty compact and convex set guaranteed by the definition of the bargaining problem, there exists a unique maximizer for this maximization problem.
(2) The proof of the theorem can be found, e.g., in [68]. The idea is first to show the NBS satisfies these

axioms, and then to show that it is the only choice satisfying all axioms.

(3) Moreover, a detailed discussion has been included in [68] to show that no axioms are superfluous. In other words, removing any of the axioms will not guarantee the uniqueness of the bargaining solution.

(4) It is ready to extend Theorem 4.1 to the bargaining problem with more than two players. Specifically, the NBS for the bargaining problem with the player set N is the solution to the following optimization problem,

$$\arg \max_{(u_1,u_2,\ldots)\in U, u_k > u_k^0, \forall k \in N} \prod_{k\in N}(u_k - u_k^0). \tag{43}$$

(5) As shown in [67], when $u_k^0 = 0$, $\forall k \in N$, the NBS coincides with the proportional fairness which is one of the widely-used criteria for resource allocation. In simpler words, the NBS achieves some degree of fairness among cooperative players through bargaining.

In what follows, we give a brief summary on how bargaining games have been applied to cognitive radio networks where cooperation among players is possible and fairness is an important concern.

In [69], the NBS is directly applied to allocate frequency-time units in an efficient and fair way, after a learning process is first applied to find the payoffs with disagreement.

The symmetry axiom implies that all players are equal in the bargaining game; however, sometimes it is not true because some players have priority over others. To accommodate this situation, a variant of the NBS is to offset the disagreement point to some other payoff vectors that implicitly incorporate the asymmetry among players. An alternative approach is to modify the objective function to $\prod_{k\in N}(u_k - u_k^0)^{w_k}$ with weights w_k reflecting the priority of players. For instance, in the power allocation game consisting of primary users and secondary users [70], different values of u_k^0 are set to primary users and secondary users because primary users have the priority to use spectrum resources in cognitive radio networks. In [71] with heterogenous wireless users, the disagreement point in the NBS objective function is replaced by the threat made by individual players.

Moreover, finding the NBS needs global information which is not always available. A distributed implementation is proposed in [23] where users adapt their spectrum assignment to approximate the optimal assignment through bargaining within local groups. Although not explicitly stated, it actually falls into the category of the NBS, because the objective is to maximize the total logarithmic user throughput which is equivalent to maximizing the product of user payoffs. In this work, neighboring players adjust spectrum band assignment for better system performance through one-to-one or one-to-many bargaining. In addition, a theoretic lower bound is derived to guide the bargaining process.

A similar approach is conducted in [72] which iteratively updates the power allocation strategy using only local information. In this game, players allocate power to channels and their payoffs are the corresponding capacity. Given the assumption that players far away from each other have negligible interference, from a particular player's perspective, the global objective is detached to two parts: the product of faraway players' payoffs and the product of neighboring players' payoffs. Because the player's power allocation strategy only affects the second term, maximizing the second term is equivalent to maximizing the global objective. Each player sequentially adjusts the strategy, and it is proved that the iterative process is convergent. Although it is not sure whether it converges to the NBS, simulation results show that the convergence point is close to the true NBS.

The concept of the NBS can also be applied to scenarios without explicit bargaining. For example, the NBS is employed to determine how to split payment among several users in a cognitive spectrum auction in [57], where the auctioneer directly set the NBS as the price to each player, and they will be ready to accept because the NBS is an equilibrium. In this paper, the objective function is defined as the product of individual payoffs which is similar to the NBS, but additional constraints have been introduced to eliminate collusive behavior in the auction.

4.2. Coalitional games

Coalitional game is another type of cooperative game. It describes how a set of players can cooperate with others by forming cooperating groups and thus improve their payoff in a game.

Denote the set of players by N, and a non-empty subset of N, i.e., a coalition, by S. Since the players in coalition S have agreed to cooperate together, they can be viewed as one entity and is associated with a value $v(S)$ which represents the worth of coalition S. Then, a coalitional game is determined by N and $v(S)$. This kind of coalitional games is known as games *with transferrable payoff*, since the value $v(S)$ is the total payoff that can be distributed in any way among the members of S, e.g., using an appropriate fairness rule.

However, in some coalitional games, it is difficult to assign a single real-number value to a coalition. Instead, each coalition S is characterized by an arbitrary set $V(S)$ of consequences. Such games are known as coalitional games *without transferrable payoff*. In these games, the payoff that each player in S receives depends on the joint actions of the members of S, and $V(S)$ becomes a set of payoff vectors, where each element x_i of a vector $x \in V(S)$ represents player i's payoff as a member of S.

In coalitional games with or without transferrable payoff values, the value of a coalition S only depends on the members of S, while not affected by how the players outside coalition S are partitioned. We call these coalition games are in *characteristic function form*. Sometimes, the value of S is also affected by how the players in $N\setminus S$ are partitioned into various coalitions, and we call those coalitional games are of the *partition function form* [73]. In partition form games, the coalitional structure is denoted by a partition $\mathscr{S}$ of N, where $\mathscr{S} = \{S_1, \ldots, S_K\}$, $S_i \bigcap S_j = \emptyset$, for $i \neq j$, and $\bigcup_{i=1}^K S_i = N$. The value of $S_i \in \mathscr{S}$ depends on the coalitional structure $\mathscr{S}$, and can be denoted by $v(S, \mathscr{S})$.

In characteristic function form coalitional games, cooperation by forming larger coalitions is beneficial for players in terms of a higher payoff. This property is referred to as *superadditivity*. For instance in games with transferrable payoff, superadditivity means

$$v(S_1 \bigcup S_2) \geqslant v(S_1) + v(S_2), \forall S_1, S_2 \subset N, S_1 \bigcap S_2 = \emptyset. \tag{44}$$

Therefore, forming larger coalitions from disjoint (smaller) coalitions can bring at least a payoff that can be obtained from the disjoint coalitions individually. Due to this property, it is always beneficial for players in a superadditive game to form a coalition that contains all the players, i.e., the *grand coalition*.

As the grand coalition provides the highest total payoff for the players, it is the optimal solution that is preferred by rational players. Naturally, one may wonder: is the grand coalition always achievable and stable? To answer this question, we first introduce the solution concept for coalitional games, *the core* [4]. The idea behind the core is similar to that behind a Nash equilibrium of a non-cooperative game: a strategy profile where no player would deviate unilaterally to obtain a higher payoff. In a coalitional game, an outcome is stable if no coalition is willing to deviate and obtain an outcome that is better for all its members. Let $\langle N, v\rangle$ denote a coalitional game. For any payoff profile $(x_i)_{i\in N}$ of real numbers and any coalition S, let $x(S) = \sum_{i\in S} x_i$. A vector $(x_i)_{i\in S}$ is an *S-feasible payoff vector* if $x(S) = v(S)$.

Definition 4.2. The core of the coalitional game $\langle N, v\rangle$ is the set of feasible payoff profile $(x_i)_{i\in N}$ for which there is no coalition S and S-feasible payoff vector $(y_i)_{i\in S}$, such that $y_i > x_i$ for all $i \in S$.

In other words, no coalition $S \subset N$ has an incentive to reject the proposed payoff profile in the core, deviate from the grand coalition, and form coalition S instead. Therefore, the definition of the core is equivalent to

$$\mathscr{C} = \{(x_i) : \sum_{i\in N} x_i = v(N) \quad \text{and} \quad \sum_{i\in S} x_i \geqslant v(S), \forall S \subseteq N\}. \tag{45}$$

As long as one can find a payoff allocation (x_i) that lies in the core, the grand coalition is a stable and optimal solution for the coalitional game.

It can be seen that the core is the set of payoff profiles that satisfy a system of weak linear inequalities, and thus is closed and convex. Moreover, we can find the core by solving a linear program

$$\min_{(x_i)_{i\in N}} \sum_{i\in N} x_i, \ \text{s.t.} \ \sum_{i\in N} x_i = v(N), \ \sum_{i\in S} x_i \geqslant v(S) \quad \forall S \subseteq N. \tag{46}$$

The existence of the core depends on the feasibility of the linear program and is related to the *balanced* property of a game. A coalitional game with transferrable payoff is called *balanced* if and only if the following inequality:

$$\sum_{S\subseteq N} \lambda_S v(S) \leqslant v(N), \tag{47}$$

holds for all non-negative weight collections $\lambda = (\lambda_S)_{S\subseteq N}$, where the collection $(\lambda_S)_{S\in\mathscr{S}}$ of numbers in [0,1] denotes a balanced collection of weights, and the sum of λ_S over all the coalitions that contain player i is $\sum_{S\supseteq i} \lambda_S = 1$. If we assume any player $i \in N$ has a single unit of time for distribution among all the coalitions in which he/she is a member, in order for a coalition S to be active for a fraction of time λ_S, all members of S must be active in S during λ_S, and the resulting payoff is $\lambda_S v(S)$. Then, the balanced property means that there exists no feasible allocation of time that can yield a total payoff higher than that of the grand coalition $v(N)$, and thus the grand coalition is optimal, indicating there may exist a non-empty core. Without giving the detailed proof (interested reader can refer to [4]), we present the result about the existence of a non-empty core in the following theorem [4].

Theorem 4.2. *A coalitional game with transferrable payoff has a non-empty core if and only if it is balanced.*

If the balanced property of a game does not hold, the core will be empty, and one will have trouble in finding a suitable solution of a coalitional game. Thus, an alternative solution concept that always exists in a coalitional game is in need. Shapley proposed a solution concept, known as the **Shapley value** ψ, to assign a unique payoff value to each player in the game. In the following, we provide an axiomatic characterization of the Shapley value, where ψ_i denotes the payoff assigned to player i according to the Shapley value.

- (Symmetry) If player i and player j are interchangeable in v, i.e., $v(S \bigcup \{i\}) = v(S \bigcup \{j\})$ for every coalition S that does not contain player i or j, then $\psi_i(v) = \psi_j(v)$.
- (Dummy player) If player i is a dummy in v, i.e., $v(S) = v(S \bigcup \{i\})$ for every coalition S, then $\psi_i(v) = 0$.
- (Additivity) For any two games u and v, define the game $u + v$ by $(u + v)(S) = u(S) + v(S)$, then $\psi_i(u + v) = \psi_i(u) + \psi_i(v)$ for all $i \in N$.
- (Efficiency) $\sum_{i\in N} \psi_i(v) = v(N)$.

The Shapley value is the only value that satisfies all the above axioms, and is usually calculated as the expected marginal contribution of player i when joining the grand coalition given by

$$\psi_i(v) = \sum_{S\subseteq |N|\setminus\{i\}} \frac{|S|!(|N| - |S| - 1)!}{|N|!} [v(S \cup \{i\}) - v(S)]. \tag{48}$$

In a cognitive radio network, cooperation among rational users can generally improve the network performance due to the multiuser diversity and spatial diversity in a wireless environment. Thus, coalitional game theory has been used to study user cooperation and design optimal, fair, and efficient collaboration strategies. In [74], spectrum sharing through receiver cooperation is studied under a coalitional game framework. The authors model the receiver cooperation in a Gaussian interference channel as a coalitional game with transferrable payoff, where the value of the game is defined as the sum-rate achieved by jointly decoding all users in the coalition. It is shown that the grand coalition that maximizes the sum-rate is stable, and the rate allocation to members of a coalition is solved by a bargaining game modeling. Receiver cooperation by forming a linear multiuser detector is modeled as a game without transferrable payoff, where the payoff of each player is the received SINR. At high SINR regime, the grand coalition is proved to be stable and sum-rate maximizing. The work in [75] has modeled cooperative spectrum sensing among secondary users as a coalition game without transferrable payoff, and a distributed algorithm is proposed for coalition formation through merge and split. It is shown that the secondary users can self-organize into disjoint independent coalitions, and the detection probability is maximized while maintaining a certain false alarm level.

5. Stochastic games

In the above, we have discussed the various aspects of game theory with their applications to cognitive radio networking, from non-cooperative spectrum competition, to spectrum trading using market equilibrium concepts, and to cooperative spectrum sharing games. Generally speaking, in these games the players are assumed to face the same stage game at each time, meaning the game and the players' strategies are not depending on the current state of the network. However, this is not true for a cognitive radio network where the spectrum opportunities and the surrounding radio environment keep changing over time. In order to study the cooperation and competition behaviors of cognitive users in a dynamic environment, the theory of stochastic games is a better fit.

A stochastic game [76] is an extension of Markov Decision Process (MDP) [77] by considering the interactive competition among different agents. In a stochastic game $\mathbb{G}$, there is a set of states, denoted by $\mathcal{S}$, and a collection of action sets, $A_1,\ldots,A_{|N|}$, one for each player in the game. The game is played in a sequence of stages. At the beginning of each stage the game is in some state. After the players select and execute their actions, the game then moves to a new random state with transition probability determined by the current state and one action from each player: $T:\mathcal{S}\times A_1\times\cdots\times A_{|N|}\mapsto PD(\mathcal{S})$. Meanwhile, at each stage each player receives a payoff $u_i:\mathcal{S}\times A_1\times\cdots\times A_{|N|}\mapsto\mathbb{R}$, which also depends on the current state and the chosen actions. The game is played continually for a number of stages, and each player attempts to maximize an objective function. The objective function can be defined as the expected sum of discounted payoffs in an infinite horizon, $E\{\sum_{j=0}^{\infty}\gamma^j u_{i,t+j}\}$, where $u_{i,t+j}$ is the reward received j steps into the future by player i and γ is the discount factor. It can also be defined as the expected sum of discounted payoffs over a finite time horizon, or the limit of the average reward. Since in a cognitive radio network, data transmission is usually assumed to last for a sufficiently long time and be sensitive to time delay (e.g., multimedia content), the most widely adopted form of objective function is the expected sum of discounted payoffs over an infinite horizon.

The solution, also called a *policy* of a stochastic game is defined as a probability distribution over the action set at any state, $\pi_i:\mathcal{S}\to PD(A_i)$, for all $i\in N$. Given the current state s^t at time t, if player i's policy π_i^t at time t is independent of the states and actions in all previous time slots, the policy π_i is said to be *Markov*. If the policy is further independent of time, it is said to be *stationary*.

The stationary policy of the players in a stochastic game, i.e., their optimal strategies, can be obtained by value iteration according to Bellman's optimality condition. For example, in a two-player stochastic game with opposite objectives, let us denote $V(s)$ as the expected reward (of player 1) for the optimal policy starting from state s, and $Q(s,a_1,a_2)$ as the expected reward of player 1 for taking action a_1 against player 2 taking action a_2 from state s and continuing optimally thereafter [78]. Then, the optimal strategy for player 1 can be obtained from the following iterations:

$$V(s)=\max_{\pi\in\alpha(A_1)}\min_{a_2\in A_2}\sum_{a_1\in A_1}Q(s,a_1,a_2)\pi_{a_1}, \tag{49}$$

$$Q(s,a_1,a_2)=u_1(s,a_1,a_2)+\gamma\sum_{s'\in\mathcal{S}}T(s,a_1,a_2,s')V(s'), \tag{50}$$

where π_{a_1} denotes player 1's strategy profile, and $T(s,a_1,a_2,s^{'})$ denotes the transition probability from state s to $s^{'}$, when player 1 takes a_1 and player 2 takes a_2.

For several special kinds of stochastic games, linear program can also formulated to obtain the optimal policy, which is defined as the probability profile $(\pi_i(s,a_1,\ldots,a_{|N|}))_{i\in N}$, for all $s\in\mathcal{S}$, and $a_i\in A_i$. Examples include single-controller discounted games, separable reward state independent transition discounted games, and switching controller discounted games. Interested readers are referred to [77] for more details.

In the following, we use several example applications of stochastic game theory to cognitive radio networking to illustrate how to formulate a stochastic game for different problems and how to solve the game.

- *Spectrum auction* [79]: At each time slot, a central spectrum moderator auctions the currently available spectrum resources, and a set of secondary users strategically bid for the resources. As secondary users need to cope with uncertainties from both the environment (e.g., channel availability and quality variations, packet arrivals from the source) and interactions with the other secondary users (e.g., resource allocation from the auction), the state of the stochastic game is composed of the buffer state and channel state, where the buffer state is dependent on the current spectrum allocation status. The transition probability of the game can be derived, since the packet arrival is assumed to be a Poisson process and the channel state transition is modeled as a Markov chain. Strategic secondary users want to maximize the number of transmitted packets by choosing the optimal bidding strategy. To this end, an interactive learning algorithm is proposed, where the high dimensional state space is decomposed and reduced to a simpler expression, based on the conjecture from previous spectrum allocations, and state transition probabilities are further estimated using past observations on transitions between different states. In this way, secondary users can approximate the future reward and approach the optimal policy through value iteration.

- *Transmission control* [80]: The secondary users' rate adaptation problem is formulated as a constrained zero-sum stochastic game. Under TDMA assumption, the system state transition probabilities are only dependent on the user who is transmitting, and thus the game falls into the category of switching controller game. The state of the transmission control stochastic game comprises channel state, secondary users' buffer state, and incoming traffic state; and the action of each user is the transmission rate. The cost that the users try to minimize is composed of a transmission cost, which is a function related to channel quality and transmission rate, and a holding cost, which is a function related to the buffer state. It is shown that there exist NE in the transmission control game, since it is a zero-sum game; moreover, a stochastic approximation algorithm is proposed to search for the NE.
- *Anti-jamming defense* [81]: Cognitive attackers may exist in a cognitive radio network, who can adapt their attacking strategy to the time-varying spectrum opportunities and secondary users' strategy. To alleviate the damage caused by cognitive attackers, a dynamic security mechanism is investigated in [81] by a stochastic game modeling. The state of the anti-jamming game includes the spectrum availability, channel quality, and the status of jammed channels observed at the current time slot. The action of the secondary users reflects how many channels they should reserve for transmitting control and data messages and how to switch between the different channels. Since the secondary users and attackers have opposite objectives, the anti-jamming game is a zero-sum game, and the optimal policy of the secondary users is obtained by the minimax-Q learning algorithm based on (49) and (50).

6. Research challenges and future directions

Although game theory has been extensively used and has offered a lot of benefits in studying, modeling, and analyzing the strategic interaction process among various users in cognitive radio networks, there are still some research challenges that require the research community to pay close attention to. In this section, we discuss some major challenges and future directions in this area, which we hope could inspire researchers' interests for the development of game theoretic modeling approaches for cognitive radio networks.

6.1. Defining a proper payoff function

The payoff function reflects the objective that a player wants to achieve from playing a game. For games in cognitive radio networks, different payoff functions have been chosen for different problem settings: in auction games or spectrum trading, the payoff function is usually defined as the net profit, i.e., the gain from using the spectrum minus the cost of holding the spectrum band. In other works that are not directly related to spectrum trading, the most commonly adopted payoff function is usually a simple (often concave) function containing one or more QoS metrics, such as throughput/rates/capacity, delay, or error probability. Although choosing a simple payoff function may simplify the analysis of equilibrium and other properties of the game, a lot of real-world constraints and situations are largely ignored. For instance, several works (e.g., [17,24]) have chosen the cost term in the payoff function as a linear function of the transmit power. However, the cost of the transmit power may depend on the specific device/user and the remaining power level, and thus is probably not linear in the transmit power. Even a linear function can roughly capture the cost, if the gain term (e.g., a function of some QoS metric) in the payoff function is in another unit rather than power unit, how to choose the weight of the linear function to balance the gain and the cost still remains a problem. Therefore, it is important to choose a meaningful payoff function that can precisely characterize players' objectives, instead of choosing an easy-to-analyze one while over-simplifying the practical problem.

6.2. Learning in games

The focus of most game theoretic spectrum allocation approaches is on solving the equilibrium and analyzing its properties, without further considering how players should interact to approach the equilibrium. These works implicitly assume that players have complete knowledge about the game being played, such as the action space and payoff functions of each other, and hence NE can be achieved without exploring the action space. However, the benefit of selfish players who compete for spectrum resources may get impaired if they reveal their private information to other players, since others tend to take advantage of such information to improve their own payoffs. Thus, complete knowledge about the game cannot be taken for granted, and equilibrium may not be achieved within a small number of stages. Instead, the game is usually played in a series of stages, where players gradually improve their payoffs according to the observation about the others' actions and the payoff received in the past, if such observation is available.

Learning is involved in this process, and a good learning algorithm will enable players to choose the right strategy that converges to a desirable equilibrium. While network researchers can refer to existing learning approaches introduced by game theorists, e.g., [82,83], to help network design, practical learning algorithms that are suitable for communication networks are worth further research, especially for games with multiple players which are often encountered in spectrum allocation. First, in decentralized networks, information that can be gathered by different users may be asymmetric. With different levels of side information and different objectives, users can adopt various learning approaches, and it cannot be anticipated where the outcome of the game will converge to. Even if learning algorithms can be programmed in cognitive radio devices, when those devices are attacked and compromised by malicious attackers, the learning algorithms can still be tampered. Thus, we need to study the users' behavior dynamics and see how their strategies converge in establishing an equilibrium. In addition, due to the limited processing capability, no devices can monitor and process every piece of information at any time. Therefore, effective learning with limited observation is needed so that the improvement of network performance can be speeded up.

6.3. Efficiency of equilibrium

When selfish players compete for spectrum resources, the resulting NE is usually not efficient in terms of social welfare. Thus, how to stimulate and enforce cooperation becomes critical for further improving the overall spectrum efficiency. As discussed in Section 2.4, pricing and repeated game have been proposed to prevent the tragedy of commons; by introducing recommendation of a third party, selfish players' strategies will be led into the set of correlated equilibria with probably a more efficient outcome. However, not all of these approaches will be effective in suppressing over-competition under some circumstances. For instance, most repeated game modeling approaches assume that the game is played for infinitely many times and the discount factors of all players are very close to 1. This way, no player will deviate from mutual cooperation since the gain from deviation will be negated by potential punishment from the other players. However, if the discount factors of some players are much smaller than those of the others, e.g., when some players have much shorter data sequences to transmit and thus will access the spectrum band for a shorter time period, they may take advantage of players with larger discount factors, and desirable cooperation and system efficiency cannot be guaranteed. Even within the set of correlated equilibria, not all operating points are efficient, such as those coincide with the NE points. Therefore, cooperation stimulation still remains a challenging task for efficient spectrum utilization.

6.4. Issues in mechanism design

Mechanism design can be viewed as another way to improve the efficiency of equilibrium by motivating players to compete honestly, so that the spectrum resources are allocated to players who value them most by means of transfers. Transfers must be collected and redistributed by a trusted entity. In spectrum auction games, a specific form of mechanism design, a spectrum auctioneer collects bids and allocates spectrum bands according to some distribution rule, e.g., using a second-price auction, and the honesty of players in bidding can be assured. In general spectrum allocation games without monetary gain explicitly involved, however, transfers can no longer be chosen as money, and alternative forms of transfers (e.g., "virtual credits" or "virtual currency") must be properly defined, and a trusted management point (e.g., a "virtual central bank") is required in the game. If there exists no trustful entity, it will be difficult to have honest competition and effective mechanisms.

6.5. Issues in stochastic games

The theory of stochastic games, in which players can update their strategy for each observed state by policy iteration or value iteration, has been recently used to effectively model players' interactions and derive their optimal strategies in dynamic spectrum access. How much information players can observe during the game will heavily affect their choice on adjusting strategy, especially in a stochastic spectrum access game that requires players to closely observe the states of the environment for correct decision making. Usually, the states of a stochastic game must be fully observable by all players. When the state information of the game is only partially observable, players need to either define new state variables that can be observed or estimate and build a belief of the original state based on the available observation. It is difficult to have a good estimate if there is little information available or if the dimension of state space is very large which requires approximation to further reduce complexity. In addition, the estimation will become much more complicated if the state set is a continuous set. Usually when the state or action set is a continuous set, players can "quantize" the continuous variable(s) to discrete values. However, the granularity of discretization will influence the policy, and better performance can only be achieved at the cost of higher complexity. Hence, how to obtain an optimal policy when the state information is only partially observable and when the state set is continuous is worth further investigation.

6.6. Security

Due to the intrinsic feature of dynamic spectrum access, a cognitive radio network is very vulnerable to malicious attacks. First, in the opportunity-based spectrum access, secondary users do not own the spectrum band, and hence their access to that band cannot be protected from adversaries. Second, the spectrum availability is highly dynamic in nature, and the traditional security enhancing mechanisms are not directly applicable, since they only fit in a static spectrum environment. Moreover, some cognitive radio networks may work in a distributed fashion, which makes it more difficult to fight against malicious attacks than in a centralized system. Last but not least, as cognitive radio networks benefit from technology evolution to be capable of utilizing spectrum adaptively and intelligently, the same technologies can also be exploited by malicious attackers to launch more complicated and unpredictable attacks with even greater damage. Therefore, ensuring security is critical for the wide deployment of cognitive radio networks.

Game theory is a natural tool for the design of defense, since the attackers will try every means to prevent legitimate users from efficiently utilizing the spectrum resources, and the interactions between the mutually distrustful parties can be modeled as a non-cooperative game. Dynamic games, e.g., stochastic games, can be used to help derive the optimal defense strategies that accommodate both the environment dynamics and the adaptive attacking strategy. Pursuit-evasion modeling can be used for attacker tracking in multi-band networks. Evolutionary game theory can help legitimate users find stable equilibrium strategies that will overwhelm possible attacks, if the attackers are viewed as mutants. Coalitional games, graphical games, and network formation games can help legitimate users form optimal partitions or networks to defend against multiple distributed attackers in a large area. Establishing reputation systems can identify malicious attackers who provide false feedback, create feedback from fake identities (sybil attack), acquire new identities and start over with a clear reputation (whitewashing), and etc. The games we mentioned here are just a few examples that are potential solutions to security enhancement in cognitive radio networks. They can also be used to design dynamic spectrum allocation schemes in an environment without malicious attackers.

7. Conclusions

In this tutorial survey, we provided a comprehensive overview of game theory, and its applications to the research on cognitive radio networks. To this end, we classify state-of-the-art game theoretic research contributions on cognitive radio networking into four categories, non-cooperative spectrum sharing, spectrum trading and mechanism design, cooperative spectrum sharing, and stochastic spectrum sharing games. For each category, we explained the fundamental concepts and properties, and provided a detailed discussion about the methodologies on how to apply these games in spectrum sharing protocol design. We also discussed some research challenges and suggested future research directions related to game theoretic modeling in cognitive radio networks. This article provides a tutorial survey on game theoretic approaches for dynamic spectrum sharing in cognitive radio networks, by in-depth theoretic analysis and an overview of the most recent practical implementations, and thus fills a void in the existing literature of cognitive radio communications and networking.

References

[1] J. Mitola, Cognitive radio: an integrated agent architecture for software defined radio, Ph.D. Dissertation, KTH Royal Institute of Technology, Stockholm, Sweden, 2000.
[2] S. Haykin, Cognitive radio: brain-empowered wireless communications, IEEE Journal on Selected Areas in Communications 23 (2) (2005) 201–220.
[3] I. Akyildiz, W. Lee, M. Vuran, S. Mohanty, NeXt generation/dynamic spectrum access/cognitive radio wireless networks: a survey, Computer Networks 50 (13) (2006) 2127–2159.
[4] M. Osborne, A. Rubinstein, A Course in Game Theory, MIT press, Cambridge, MA, 1999.
[5] D. Monderer, L. Shapley, Potential games, Games and Economic Behavior 14 (1) (1996) 124–143.
[6] J. Neel, R. Buehrer, J. Reed, R. Gilles, Game theoretic analysis of a network of cognitive radios, in: IEEE Midwest Symposium on Circuits and Systems, vol. 45, 2002.
[7] T. Ui, A Shapley value representation of potential games, Games and Economic Behavior 31 (1) (2000) 121–135.
[8] J. Neel, J. Reed, R. Gilles, The role of game theory in the analysis of software radio networks, in: SDR Forum Technical Conference, November, 2002.
[9] N. Nie, C. Comaniciu, Adaptive channel allocation spectrum etiquette for cognitive radio networks, Mobile Networks and Applications 11 (6) (2006) 779–797.
[10] R. Menon, A. MacKenzie, R. Buehrer, J. Reed, A game-theoretic framework for interference avoidance in ad hoc networks, in: Proceedings of the IEEE GLOBECOM 2006, San Francisco, CA, USA, 2006.
[11] Y. Xing, C. Mathur, M. Haleem, R. Chandramouli, K. Subbalakshmi, Dynamic spectrum access with QoS and interference temperature constraints, IEEE Transactions on Mobile Computing 6 (4) (2007) 423–433.
[12] J.O. Neel, R. Menon, A.B. MacKenzie, J.H. Reed, R.P. Gilles, Interference reducing networks, in: Proceedings of the 2nd International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CrownCom), August 2007, pp. 96–104.
[13] R.W. Thomas, R.S. Komali, A.B. MacKenzie, L.A. DaSilva, Joint power and channel minimization in topology control: a cognitive network approach, in: IEEE International Conference on Communications, 2007, pp. 6538–6543.
[14] L. Giupponi, C. Ibars, Distributed cooperation in cognitive radio networks: overlay versus underlay paradigm, in: IEEE 69th Vehicular Technology Conference (VTC09-Spring), April 2009.
[15] R. Yates, A framework for uplink power control in cellular radio systems, IEEE Journal on Selected Areas in Communications 13 (7) (1995) 1341–1348.
[16] C.U. Saraydar, N.B. Mandayam, D.J. Goodman, Efficient power control via pricing in wireless data networks, IEEE Transactions on Communications 50 (2) (2002) 291–303.
[17] J. Mwangoka, K. Letaief, Z. Cao, Joint power control and spectrum allocation for cognitive radio networks via pricing, Physical Communication 2 (1-2) (2009) 103–115.
[18] J. Zhang, Q. Zhang, Stackelberg game for utility-based cooperative cognitive radio networks, in: Proceedings of the 10th ACM International Symposium on Mobile Ad Hoc Networking and Computing, ACM, New York, NY, USA, 2009, pp. 23–32.
[19] J.W. Weibull, Evolutionary Game Theory, MIT Press, Cambridge, MA, 1995.
[20] B. Wang, K.J.R. Liu, T.C. Clancy, Evolutionary cooperative spectrum sensing game: how to collaborate?, IEEE Transactions on Communications 58 (3) (2010) 890–900
[21] M. Perez-Guirao, R. Luebben, T. Kaiser, K. Jobmann, Evolutionary game theoretical approach for IR-UWB sensor networks, in: IEEE International Conference on Communications Workshops, 2008, pp. 107–111.
[22] M. Halldórsson, J. Halpern, L. Li, V. Mirrokni, On spectrum sharing games, in: The 23rd Annual ACM Symposium on Principles of Distributed Computing, 2004, pp. 107–114.
[23] L. Cao, H. Zheng, Distributed spectrum allocation via local bargaining, in: IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON), Santa Clara, CA, Sep. 2005, pp. 475–486.
[24] A. Al Daoud, T. Alpcan, S. Agarwal, M. Alanyali, A Stackelberg game for pricing uplink power in wide-band cognitive radio networks, in: Proceedings of 47th IEEE Conference on Decision and Control, 2008, pp. 1422–1427.
[25] F. Wang, M. Krunz, S. Cui, Price-based spectrum management in cognitive radio networks, IEEE Journal of Selected Topics in Signal Processing 2 (1) (2008) 74–87.
[26] W. Wang, Y. Cui, T. Peng, W. Wang, Noncooperative power control game with exponential pricing for cognitive radio network, in: IEEE 65th Vehicular Technology Conference (VTC2007-Spring), 2007, pp. 3125–3129.
[27] J. Huang, R. Berry, M. Honig, Spectrum sharing with distributed interference compensation, in: Proceedings of IEEE DySPAN, 2005, pp. 88–93.
[28] J. Ratliff, A folk theorem sampler, Lecture Notes, 1996. Available from: <http://www.virtualperfection.com/gametheory/5.3.FolkTheoremSampler.1.0.pdf>.
[29] R. Etkin, A. Parekh, D. Tse, Spectrum sharing for unlicensed bands, IEEE Journal on Selected Areas in Communications 25 (3) (2007) 517–528.
[30] Y. Wu, B. Wang, K.J.R. Liu, T.C. Clancy, Repeated open spectrum sharing game with cheat-proof strategies, IEEE Transactions on Wireless Communications 8 (4) (2009) 1922–1933.
[31] M. van der Schaar, F. Fu, Spectrum access games and strategic learning in cognitive radio networks for delay-critical applications, Proc. IEEE 97 (4) (2009) 720–740.
[32] C. Song, Q. Zhang, Achieving cooperative spectrum sensing in wireless cognitive radio networks, ACM SIGMOBILE Mobile Computing and Communications Review 13 (2) (2009) 14–25.
[33] B. Wang, Z. Ji, K.J.R. Liu, Self-learning repeated game framework for distributed primary-prioritized dynamic spectrum access, in: Proceedings of the IEEE SECON, 2007, pp. 631–638.
[34] Z. Han, Z. Ji, K.J.R. Liu, Non-cooperative resource competition game by virtual referee in multi-cell OFDMA networks, IEEE Journal on Selected Areas in Communications 25 (6) (2007) 1079–1090.
[35] Z. Han, C. Pandana, K.J.R. Liu, Distributive opportunistic spectrum access for cognitive radio using correlated equilibrium and no-regret learning, in: IEEE Wireless Communications and Networking Conference (WCNC), 2007, pp. 11–15.

[36] M. Maskery, V. Krishnamurthy, Q. Zhao, Decentralized dynamic spectrum access for cognitive radios: Cooperative design of a non-cooperative game, IEEE Transactions on Communications 57 (2) (2009) 459–469.
[37] Z. Han, K.J.R. Liu, Noncooperative power-control game and throughput game over wireless networks, IEEE Transactions on Communications 53 (10) (2005) 1625–1629.
[38] S. Hart, A. Mas-Colell, A simple adaptive procedure leading to correlated equilibrium, Econometrica 68 (5) (2000) 1127–1150.
[39] D. Niyato, E. Hossain, Competitive spectrum sharing in cognitive radio networks: a dynamic game approach, IEEE Transactions on Wireless Communications 7 (7) (2008) 2651–2660.
[40] D. Niyato, E. Hossain, Competitive pricing for spectrum sharing in cognitive radio networks: dynamic game, inefficiency of Nash equilibrium, and collusion, IEEE Journal on Selected Areas in Communications 26 (1) (2008) 192–202.
[41] N. Singh, X. Vives, Price and quantity competition in a differentiated duopoly, The RAND Journal of Economics 15 (4) (1984) 546–554.
[42] J. Jia, Q. Zhang, Competitions and dynamics of duopoly wireless service providers in dynamic spectrum market, in: Proceedings of the 9th ACM International Symposium on Mobile Ad Hoc Networking and Computing, 2008, pp. 313–322.
[43] Z. Han, Z. Ji, K.J.R. Liu, A Cartel maintenance framework to enforce cooperation in wireless networks with selfish users, IEEE Transactions on Wireless Communications 7 (5) (2008) 1889–1899.
[44] O. Simeone, I. Stanojev, S. Savazzi, Y. Bar-Ness, U. Spagnolini, R. Pickholtz, Spectrum leasing to cooperating secondary ad hoc networks, IEEE Journal on Selected Areas in Communications 26 (1) (2008) 203–213.
[45] A. Ercan, J. Lee, S. Pollin, J. Rabaey, A revenue enhancing Stackelberg game for owners in opportunistic spectrum access, in: Proceedings of Dyspan, 2008.
[46] M. Bloem, T. Alpcan, T. Başar, A Stackelberg game for power control and channel allocation in cognitive radio networks, in: Proceedings of the 2nd International Conference on Performance Evaluation Methodologies and Tools, 2007.
[47] P. Bajari, J. Yeo, Auction design and tacit collusion in FCC spectrum auctions, Information Economics and Policy 21 (2) (2009) 90–100.
[48] W. Vickrey, Counterspeculation, auctions, and competitive sealed tenders, Journal of Finance 16 (1) (1961) 8–37.
[49] V. Krishna, Auction Theory, Academic press, New York, 2009.
[50] R. Myerson, Optimal auction design, Mathematics of Operations Research 6 (1) (1981) 58–73.
[51] P. Cramton, Y. Shoham, R. Steinberg, Combinatorial Auctions, MIT Press, Cambridge, MA, 2006.
[52] J. Huang, R.A. Berry, M.L. Honig, Auction-based spectrum sharing, ACM/Springer Mobile Networks and Applications Journal (MONET) 11 (3) (2006) 405–418.
[53] Y. Chen, G. Yu, Z. Zhang, H. Chen, P. Qiu, On cognitive radio networks with opportunistic power control strategies in fading channels, IEEE Transactions on Wireless Communications 7 (7) (2008) 2752–2761.
[54] S. Gandhi, C. Buragohain, L. Cao, H. Zheng, S. Suri, A general framework for wireless spectrum auctions, in: Proceedings of the 2nd IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN'07), 2007, pp. 22–33.
[55] X. Zhou, S. Gandhi, S. Suri, H. Zheng, eBay in the sky: strategy-proof wireless spectrum auctions, in: Proceedings of the 14th ACM International Conference on Mobile Computing and Networking, 2008, pp. 2–13.
[56] J. Jia, Q. Zhang, Q. Zhang, M. Liu, Revenue generation for truthful spectrum auction in dynamic spectrum access, in: Proceedings of the 10th ACM International Symposium on Mobile Ad Hoc Networking and Computing, 2009, pp. 3–12.
[57] Y. Wu, B. Wang, K.J.R. Liu, T.C. Clancy, A scalable collusion-resistant multi-winner cognitive spectrum auction game, IEEE Transactions on Communications 57 (12) (2009) 3805–3816.
[58] X. Zhou, H. Zheng, TRUST: a general framework for truthful double spectrum auctions, in: Proceedings of the IEEE INFOCOM, 2009.
[59] Z. Ji, K.J.R. Liu, Multi-stage pricing game for collusion-resistant dynamic spectrum allocation, IEEE Journal on Selected Areas in Communications 26 (1) (2008) 182–191.
[60] R.B. Myerson, Mechanism design, in: S.N. Durlauf, L.E. Blume (Eds.), The New Palgrave Dictionary of Economics, Palgrave Macmillan, Basingstoke, 2008.
[61] D. Fudenberg, J. Tirole, Game Theory, MIT Press, Cambridge, MA, 1991.
[62] A. Fattahi, F. Fu, M. van der Schaar, F. Paganini, Mechanism-based resource allocation for multimedia transmission over spectrum agile wireless networks, IEEE Journal on Selected Areas in Communications 25 (3) (2007) 601–612.
[63] B. Wang, Y. Wu, Z. Ji, K.J.R. Liu, T.C. Clancy, Game theoretical mechanism design methods: suppressing cheating in cognitive radio networks, IEEE Signal Processing Magazine 25 (6) (2008) 74–84.
[64] A. Rubinstein, Perfect equilibrium in a bargaining model, Econometrica 50 (1) (1982) 97–109.
[65] J. Nash Jr., The bargaining problem, Econometrica 18 (2) (1950) 155–162.
[66] G. Owen, Game Theory, third ed., Academic Press, London, 1995.
[67] Z. Han, Z. Ji, K.J.R. Liu, Fair multiuser channel allocation for OFDMA networks using Nash bargaining solutions and coalitions, IEEE Transactions on Communications 53 (8) (2005) 1366–1376.
[68] M.J. Osborne, A. Rubinstein, Bargaining and Markets, Academic Press, London, 1990.
[69] K. Han, J. Li, P. Zhu, and X. Wang, The frequency-time pre-allocation in unlicensed spectrum based on the games learning, in: Proceedings of the 2nd International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CrownCom), 2007, pp. 79–84.
[70] A. Attar, M. Nakhai, A. Aghvami, Cognitive radio game for secondary spectrum access problem, IEEE Transactions on Wireless Communications 8 (4) (2009) 2121–2131.
[71] H. Pham, J. Xiang, Y. Zhang, T. Skeie, QoS-aware channel selection in cognitive radio networks: a game-theoretic approach, in: IEEE Global Telecommunications Conference, 2008, pp. 1–7.
[72] J. Suris, L. DaSilva, Z. Han, A. MacKenzie, Cooperative game theory for distributed spectrum sharing, in: Proceedings of the IEEE International Conference on Communications, 2007, pp. 5282–5287.
[73] R. Thrall, W. Lucas, N-person games in partition function form, Naval Research Logistics Quarterly 10 (1) (1963) 281–298.
[74] S. Mathur, L. Sankaranarayanan, N. Mandayam, Coalitional games in Gaussian interference channels, in: Proceedings of the IEEE ISIT, 2006, pp. 2210–2214.
[75] W. Saad, Z. Han, M. Debbah, A. Hjørungnes, T. Basar, Coalitional games for distributed collaborative spectrum sensing in cognitive radio networks, in: Proceedings of IEEE INFOCOM, 2009.
[76] L. Shapley, Stochastic games, Proceedings of the National Academy of Sciences of the United States of America 39 (10) (1953) 1095–1100 .
[77] J. Filar, K. Vrieze, O. Vrieze, Competitive Markov Decision Processes, Springer, Berlin, 1997.
[78] M. Littman, Markov games as a framework for multi-agent reinforcement learning, in: Proceedings of the 11th International Conference on Machine Learning, vol. 157, 1994, pp. 163–169.
[79] F. Fu, M. van der Schaar, Learning to compete for resources in wireless stochastic games, IEEE Transactions on Vehicular Technology 58 (4) (2009) 1904–1919.
[80] J. Huang V. Krishnamurthy, Transmission control in cognitive radio systems with latency constraints as a switching control dynamic game, in: Proceedings of the 47th IEEE Conference on Decision and Control, 2008, pp. 3823–3828.
[81] B. Wang, Y. Wu, K.J.R. Liu, An anti-jamming stochastic game for cognitive radio networks, IEEE Journal on Selected Areas in Communications (submitted for publication). Available from: <http://www.ece.umd.edu/~bebewang/jsac09_bw.pdf>.
[82] D. Fudenberg, D.K. Levine, The Theory of Learning in Games, MIT Press, Cambridge, MA, 1998.
[83] N. Cesa-Bianchi, G. Lugosi, Prediction, Learning, and Games, Cambridge University Press, Cambridge, MA, 2006.

Beibei Wang (S'07) received the B.S. degree in electrical engineering (with the highest honor) from the University of Science and Technology of China, Hefei, in 2004, and the Ph.D. degree in electrical engineering from the University of Maryland, College Park in 2009, where she is currently a research associate. Her research interests include dynamic spectrum allocation and management in cognitive radio systems, wireless communications and networking, game theory, multimedia communications, and network security. Dr. Wang was the recipient of the Graduate School Fellowship, the Future Faculty Fellowship, and the Dean's Doctoral Research Award from the University of Maryland, College Park.

Yongle Wu (S'08) received the B.S. (with highest honor) and M.S. degrees in electronic engineering from Tsinghua University, Beijing, China, in 2003 and 2006, respectively. He is currently working towards the Ph.D. degree in the Department of Electrical and Computer Engineering, University of Maryland, College Park. His current research interests are in the areas of wireless communications and networks, including cognitive radio techniques, dynamic spectrum access, and network security. Mr. Wu received the Graduate School Fellowship from the University of Maryland in 2006, and the Future Faculty Fellowship from A. James Clark School of Engineering, University of Maryland in 2009.

K.J. Ray Liu (F'03) is a Distinguished Scholar-Teacher of University of Maryland, College Park. He is Associate Chair of Graduate Studies and Research of Electrical and Computer Engineering Department and leads the Maryland Signals and Information Group conducting research encompassing broad aspects of information technology including communications and networking, information forensics and security, multimedia signal processing, and biomedical technology.

Dr. Liu is the recipient of numerous honors and awards including best paper awards from IEEE Signal Processing Society, IEEE Vehicular Technology Society, and EURASIP; IEEE Signal Processing Society Distinguished Lecturer, EURASIP Meritorious Service Award, and National Science Foundation Young Investigator Award. He also received various teaching and research recognitions from University of Maryland including university-level Invention of the Year Award; and Poole and Kent Senior Faculty Teaching Award and Outstanding Faculty Research Award, both from A. James Clark School of Engineering. Dr. Liu is a Fellow of IEEE and AAAS.

Dr. Liu was Vice President – Publications and will serve as President-Elect from 2010. He was the Editor-in-Chief of IEEE SIGNAL PROCESSING MAGAZINE and the founding Editor-in-Chief of EURASIP JOURNAL ON APPLIED SIGNAL PROCESSING. His recent books include *Cooperative Communications and Networking* (Cambridge University Press, 2008); *Resource Allocation for Wireless Networks: Basics, Techniques, and Applications* (Cambridge University Press, 2008); *Ultra-Wideband Communication Systems: The Multiband OFDM Approach* (IEEE-Wiley, 2007); *Network-Aware Security for Group Communications* (Springer, 2007); *Multimedia Fingerprinting Forensics for Traitor Tracing* (Hindawi, 2005); *Handbook on Array Processing and Sensor Networks* (IEEE-Wiley, 2009).