9

Degrees of Cooperation in Dynamic Spectrum Access for Distributed Cognitive Radios *

Zhu Han

Electrical and Computer Engineering Department,
Boise State University, Boise, ID 83725, USA
zhuhan@boisestate.edu

9.1 Introduction

The Federal Radio Act under Federal Communications Commission (FCC) allows predetermined users the right to transmit at a given frequency. Non-licensed users are regarded as "harmful interference" and not allowed to transmit in a certain frequency bands. As the demands for wireless communication become more and more pervasive, the wireless devices must find a way for the right to transmit at frequencies in the extremely limited radio band. However, there exist a large number of frequency bands that have considerable, and sometimes periodic, dormant time intervals. In the literature, those frequency bands refer to spectrum holes [1,2]. So there is a dilemma that on one hand the mobile users have no spectrum to transmit, while on the other hand some spectrums are not fully utilized.

In order to cope with the dilemma, the FCC has recently investigated the efficient spectrum usage for cognitive radios, which is a novel paradigm that improves the spectrum utilization by allowing secondary networks (users) to borrow unused radio spectrum from primary licensed networks (users) or to share the spectrum with the primary networks (users). As an intelligent wireless communication system, cognitive radios are aware of the radio frequency environment, select the communication parameters (such as carrier frequency, bandwidth and transmission power) to optimize the spectrum usage and adapt the transmission and reception accordingly. Cognitive radios can bring a variety of benefits: for a regulator, cognitive radios can significantly increase in spectrum availability for new and existing applications. For a license holder, cognitive radios can reduce the complexity of frequency planning, facilitate the secondary spectrum market agreements, increase system capacity through access to more spectrum and avoid interference. For equipment manufacturers, cognitive radios can increase demands for wireless devices. Finally, for a user, cognitive radios can bring more capacity per user, enhance inter-operability and bandwidth-on-demand and provide ubiquitous mobility with a single user device across disparate spectrum access environments.

The process for spectrum access is first to sense what the available spectrum is, then to get access to some of the available spectrum, next to use the available spectrum and finally to release the used spectrum. Significant research is necessary to investigate how to dynamically access the spectrum, which enables the opportunistic management of radio resources within a single access system or between different radio access systems. As a result, dynamic spectrum access can improve spectral efficiency, increased capacity and improve ease of access to the spectrum. In the literature, much work [8,9] has been done for dynamic spectrum access.

In this chapter, we classify some of the dynamic spectrum access techniques for cognitive radios, according to the degrees of cooperation. The relations between distributed cognitive radios ranges from complete autonomy and non-cooperation, to full obeyance to the centralized controller. Specifically, we will discuss the following techniques for different degrees of cooperation:

1. Non-cooperative competition (Sect. 9.2)
2. Learning for better equilibria (Sect. 9.3)
3. Referee mediation (Sect. 9.4)
4. Threat and punishment from repeated interactions (Sect. 9.5)
5. Spectrum auction (Sect. 9.6)
6. Mutual benefits via bargaining (Sect. 9.7)
7. Contract using cooperative game (Sect. 9.8)
8. centralized scheme (Sect. 9.9)

There are some tradeoffs for different types of approaches. For example, for non-cooperative competition, the transceiver is simple but the performance can be inferior due to the extensive non-cooperation. On the other hand, the centralized scheme can achieve the optimal solution, but it is necessary for extensive measuring of channels and signaling to exchange channel information. Our goals are to investigate those different approaches with different degrees of cooperation, study in which network scenarios the approaches fit most and understand the underlying tradeoffs for the wireless cognitive network design.

9.2 Non-cooperative Competition

In cognitive wireless networks, it is hard for an individual cognitive user to know the channel conditions of the other users. The cognitive users cannot cooperate with each other for spectrum usage. They act selfishly to maximize their own performances in a distributive fashion. Such a fact motivates us to adopt game theory. Dynamic spectrum access can be modeled as a game that deals largely with how rational and intelligent individuals interact with each other in an effort to achieve their own goals. In this game, each cognitive user is self-interested and trying to optimize its utility function, where the utility function represents the cognitive user's performance and controls the outcomes of the game. There are many advantages of applying game theory to dynamic spectrum usage for cognitive radios:

1. *Only local information and distributive implementation*: The individual cognitive user observes the outcome of the game and adjusts only its parameters in response to optimize its own benefit. As a result, there is no need for collecting all the information and conducting optimization in a centralized way.
2. *More robust outcome*: For the centralized optimization, if the information for optimization is not quite accurately obtained, the optimized results can be far away from optimality. In contrast, the local information is always accurate, so the outcome of the distributed game approaches is robust.
3. *Combinatorial nature*: For traditional optimization technique such as programming, it is hard to handle the combinatorial problems. For game theory, it is natural to discuss the problem in a discrete form. In the problems such as spectrum access, to analyze the combinatorial problems by game theory is considerably convenient.
4. *Rich mathematics for optimization*: There are many mathematical tools available to analyze the outcome of the game. Specifically, if the (non-cooperative) game is played once, the static game can be studied. If the game is played multiple times, dynamic game theory is employed. If some contracts and mutual benefits can be obtained, cooperative game explains how to divide the profits. Auction theory studies the behaviors of both seller and bidder. We will study some of those techniques in the following sections.

Next, we define some basic game concepts and study two ways to present a game. Then we give some properties of the game, such as dominance, Nash equilibrium, Pareto optimality and mixed strategies. Further, we discuss the low efficiency of the outcome for non-cooperative static games. Finally, some methods are briefly discussed to improve the game outcomes.

A game can be roughly defined as each user adjusts its strategy to optimize its own utility to compete with others. Strategy and utility can be defined as:

Definition 9.1. *A strategy σ is a complete contingent plan, or a decision rule, that defines the action an agent will select in every distinguishable state Ω of the world.*

Definition 9.2. *In any game, utility (payoff) u represents the motivations of players. A utility function for a given player assigns a number for every possible outcome of the game with the property that a higher (or lower) number implies that the outcome is more preferred.*

One of the most common assumptions made in game theory is rationality. Generally speaking, rationality implies that all players are motivated by maximizing their own utilities. In a stricter sense, it implies that every player always maximizes its utility, thus being able to perfectly calculate the probabilistic result of every action. A game can be defined as follows.

Definition 9.3. *A game G in the strategic form has three elements: the set of players $i \in \mathcal{I}$, which is a finite set $\{1, 2, ..., K\}$; the strategy space Ω_i for each player i; and* *utility function u_i, which measures the outcome of the ith user for each strategy profile $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, ..., \sigma_K)$. We define $\boldsymbol{\sigma}_{-i}$ as the strategies of player i's opponents, i.e., $\sigma_{-i} = (\sigma_1, ..., \sigma_{i-1}, \sigma_{i+1}, ..., \sigma_K)$. In static games, the interaction between users occurs only once, while in dynamic games the interaction occurs several times.*

One of the most simple games is the non-cooperative static game which can be presented by the strategic (normal) form.

Definition 9.4. *A non-cooperative game is one in which players are unable to make enforceable contracts outside of those specifically modeled in the game. Hence, it is not defined as games in which players do not cooperate, but as games in which any cooperation must be self-enforcing.*

Definition 9.5. *A static game is one in which all players make decisions (or select a strategy) simultaneously, without knowledge of the strategies that are being chosen by other players. Even though the decisions may be made at different points in time, the game is simultaneous because each player has no information about the decisions of others; thus, it is as if the decisions are made simultaneously.*

Definition 9.6. *The strategic (or normal) form is a matrix representation of a simultaneous game. For two players, one is the "row" player, and the other, the "column" player. Each row or column represents a strategy and each box represents the payoffs to each player for every combination of strategies.*

To analyze the outcome of the game, the Nash equilibrium is a well-known concept which states that in the equilibrium every agent will select a utility-maximizing strategy given the strategies of every other agent.

Definition 9.7. *Define a strategy vector $\sigma = [\sigma_1 \dots \sigma_K]$ and define the strategy vector of the ith player's opponents as $\sigma_i^{-1} = [\sigma_1 \dots \sigma_{i-1}\ \sigma_{i+1} \dots \sigma_K]$, where K is the number of users and σ_i is the ith user's strategy. u_i is the ith user's utility. Nash equilibrium point σ^* is defined as:*

$$u_i(\sigma_i^*, \sigma_i^{-1}) \geq u_i(\tilde{\sigma}_i, \sigma_i^{-1}),\ \forall i,\ \forall \tilde{\sigma}_i \in \Omega,\ \sigma_i^{-1} \in \Omega^{K-1}, \tag{9.1}$$

i.e., given the other users' resource allocations, no user can increase its utility alone by changing its own resource allocation.

In other words, a Nash equilibrium, named after John Nash, is a set of strategies, one for each player, such that no player has incentive to unilaterally change its action. Players are in an equilibrium if a change in strategies by any one of them will lead that player to earn less than if it remains with its current strategy.

Until now, we have only discussed the strategy that is deterministic, or pure strategy. A pure strategy defines a specific move or action that a player will follow in every possible attainable situation in a game. Such moves may not be random, or drawn from a distribution, as in the case of mixed strategies.

Definition 9.8. *Mixed Strategy: A strategy consisting of possible moves and a probability distribution (collection of weights) which corresponds to how frequently each move is about to play. A player will only use a mixed strategy when it is indifferent about several pure strategies. Moreover, if the opponent can benefit from knowing the next move, the mixed strategy is preferred since keeping the opponent guessing is desirable.*

There might be an infinite number of Nash equilibriums. Among all these equilibriums, we need to select the optimal one. There are many criteria by which to judge if the equilibrium is optimal or not. Among these criteria, Pareto optimality is one of the most important definitions.

Definition 9.9. ***Pareto optimal***: *Named after Vilfredo Pareto, Pareto optimality is a measure of efficiency. An outcome of a game is Pareto optimal if there is no other outcome that makes every player at least as well off and at least one player strictly better off. That is, a Pareto optimal outcome cannot be improved upon without hurting at least one player. Often, a Nash equilibrium is not Pareto optimal implying that the players' payoffs can all be increased.*

Since the individual user has no incentive to cooperate with the other users in the system and imposes harm to the other users, the outcome of the non-cooperative static game might not be optimal from the system point of view. To overcome this problem, *pricing (or taxation)* has been used as an effective tool both by economists and researchers in computer networks. The pricing technique is motivated by the following two objectives:

1. The revenue for the system is optimized.
2. The cooperation for resource usage is encouraged.

An efficient pricing mechanism can make the distributed decisions compatible with the system efficiency obtained by centralized control. A pricing policy is called *incentive compatible*, if pricing enforces a Nash equilibrium that achieves the system optimum. Specifically, the new utility with pricing is

$$u' = u - \alpha Q \tag{9.2}$$

where u is the original utility, α the price for user's resource Q and the price can be different for different users. It is known that the above utility function can achieve Pareto optimality, if the utility is quasi-convex or quasi-concave.

In cognitive radio literature, it is worth mentioning the following games that can be modeled for spectrum access technology. First, a game is considered a *potential game* if the incentive of all players to change their strategy can be expressed in one global function, the potential function. The potential function is a useful tool to analyze equilibrium properties of games, since the incentives of all players are mapped into one function, and the set of pure Nash equilibria can be found by simply locating the local optima of the potential function. In [10], a potential game was utilized for problems such as interference avoidance.

In [11], a *Cournot game* was used to model the spectrum sharing problem as an oligopoly market in which a few firms compete with each other in terms of amount of commodity supplied to the market to gain the maximum profit. In this case the secondary users are analogous to the firms who compete for the spectrum offered by the primary user and the cost of the spectrum is determined by using a pricing function. Both static and dynamic Cournot games were investigated.

In some cognitive scenarios, the primary users and secondary users can be formulated as the multiple level market game, so that both types of users can be satisfied with the shared spectrum size and the charge pricing. The available techniques are the demand-and-request functions [17], Stackleberge game [13] and so on. The multiple level game is non-cooperative game and Nash equilibria can be derived for spectrum usage.

9.3 Correlated Equilibrium Through Learning

One of the major design challenges for cognitive radios is to coordinate and cooperate in accessing the spectrum opportunistically among multiple distributive users with only local information. In this section, we study the behavior of an individual distributed secondary user to control its rate when the primary user is absent. Each secondary user seeks to maximize its rates over different channels. However, excessive transmissions can cause the collisions with the other secondary users. The collisions reduce not only the system throughput but also individual performances. We propose a new solution concept, the correlated equilibrium, which is better compared to the non-cooperative Nash equilibrium in terms of spectrum utilization efficiency and fairness among the distributive users. Using the correlated equilibrium concept, the distributive users adjust their transmission probabilities over the available channels, so that the collisions are avoided and the users' benefits are optimized. We exhibit the adaptive no-regret algorithm [24] to learn the correlated equilibrium in a distributed manner. We show that the proposed learning algorithm converges to a set of correlated equilibria with probability one.

For the system model, we consider the general models for dynamic opportunistic spectrum access for cognitive radios, in which there exist several primary users with a set of available channels and a large number of secondary users. The channel availability of secondary users inherently depends on the activities of the primary users. Moreover, the secondary users have to compete for the idle channels among the interfering secondary users. If collisions occur, there are some penalties in the forms of packet loss and power waste. This is the major focus here. We consider that there are N channels in the wireless network. Without loss of generality, each channel has a unit bandwidth. These channels are shared among M primary users and K secondary users seeking channel access opportunistically.

For adjacent secondary users, they can interference with each other. We use *interference matrix* $\mathbf{L}$ to depict the interference graph. The interference matrix has the dimension of K by K, and its elements are defined as

$$\mathbf{L}_{ij} = \begin{cases} 1, \text{ if } i \text{ and } j \text{ interfere with each other} \\ 0, \text{ otherwise.} \end{cases} \tag{9.3}$$

The interference matrix depends on the relative location of secondary users.

Next, we define *channel availability matrix* as a K by N matrix, $\boldsymbol{A}(t)$. Each user can transmit over a specific channel with a set of different rates. The elements of the matrix are defined as

$$\mathbf{A}_{in}(t) = \begin{cases} 1, \text{ if channel } n \text{ is available for secondary user } i \text{ at time } t \\ 0, \text{ otherwise.} \end{cases} \tag{9.4}$$

We note that the channel availability matrix $\boldsymbol{A}(t)$ varies over time. This matrix is the result of a sensing task done by secondary users and depends on the primary users' traffic, relative location between the secondary users and the primary users. Notice that each individual secondary user only knows its corresponding row of matrix $\mathbf{L}$ and $\mathbf{A}(t)$.

Define the set of secondary user i as I which is the finite set $\{1, 2, \ldots, K\}$. For each available channel, a secondary user can select $L + 1$ discrete rates $\Upsilon = \{0, \upsilon_1, \ldots, \upsilon_L\}$. The strategy space Ω_i for secondary user i is on the available channels and can be denoted as $\Omega_i = \prod_{n=1}^{N} \Upsilon^{\mathbf{A}_{in}}$. The action of user is $r_i^n = \upsilon_l$ representing secondary user i occupies channel n by rate υ_l. We define the strategy profile $\mathbf{r}^n = (r_1^n, r_2^n, \ldots, r_K^n)'$, and we define $\mathbf{r}_{-i}^n$ as the strategies of user i's opponents (interference neighbors defined in $\mathbf{L}$) for channel n. We also define $\mathbf{r}_i = (r_i^1, \ldots, r_i^N)'$ as the action of secondary users over all channels, and $\mathbf{r}_{-i}$ as the secondary user i's opponents' actions.

The utility function U_i measures the outcome of secondary user i for each strategy profile $\mathbf{r}^1, \ldots, \mathbf{r}^N$ over different channels. We define the utility function as the maximum achievable rate for the secondary users over all the available channels as:

$$U_i = \sum_{n=1}^{N} \mathbf{A}_{in} R_i(r_i^n, \mathbf{r}_{-i}^n) \tag{9.5}$$

where $R_i(r_i^n, \mathbf{r}_{-i}^n)$ is the outcome of resource competition for user i and the other users. Notice that the utility function represents the maximum achievable rate. In practice, the secondary users need not occupy all the available channels.

We consider un-slotted One-persistent CSMA as the random multiple access protocol for the secondary users. Since the channel can be occupied by the primary user again in the near future, each secondary user transmits whenever the channel is idle. From [14], we have

$$R_i(r_i^n, \mathbf{r}_{-i}^n) = \begin{cases} \frac{r_i^n S^n}{\sum_i r_i^n}, & \text{if } G \le G_0 \\ 0, & \text{otherwise} \end{cases} \tag{9.6}$$

where

$$S^n = \frac{G^n[1 + G^n + \tau G^n(1 + G^n + \tau G^n/2)]\mathrm{e}^{-G^n(1+2\tau)}}{G^n(1 + 2\tau) - (1 - \mathrm{e}^{-\tau G^n}) + (1 + \tau G^n)\mathrm{e}^{-G^n(1+\tau)}}, \tag{9.7}$$

$G^n = \sum_i r_i^n$, and τ is the propagation delay over packet transmission time. When the network payload increases, more collisions happen and consequently the average delay for each packet increases. For some types of payloads like multimedia services, the delayed packets can cause significant QoS loss. In [15], it has been shown that the average delay can be unbounded for a sufficiently large load. Moreover, for cognitive radios, since the primary users can reoccupy the channel in the near future, a certain delay can cause the second user to lose the opportunity for transmission entirely. So we define G_0 as the maximum network payload. Any network payload larger than G_0 will cause an unacceptable average delay. As a result, the utility function is zero.

In the following, we first propose a new solution concept, correlated equilibrium. Then, we investigate a linear programming method to calculate the optimal correlated equilibrium. Finally, we utilize a no-regret algorithm to learn the correlated equilibria in a distributed way.

To analyze the outcome of the game, Nash equilibrium is a well-known concept, which states that in the equilibrium every user will select a utility-maximizing strategy given the strategies of every other user. If a user follows an action in every possible attainable situation in a game, the action is called pure strategy, in which the probability of using action ν_l, $p(r_i^n = \nu_l)$, has only one non-zero value 1 for all l. In the case of mixed strategies, the user will follow a probability distribution over different possible action, i.e., different rate l.

In Table 9.1, we illustrate an example of two secondary users with different actions. In Table 9.1a, we list the utility function for two users taking action 0 and 1. We can see that when two users take action of 0, they have the best overall benefit. We can see this action as a cooperative action (in our case the users transmit less aggressively). But if any user plays more aggressively using action 1 while the other still plays action 0, the aggressive user has a better utility, but the other user has a lower utility and the overall benefit is reduced. In our case, the aggressive user can achieve a higher rate. However, if both users play aggressively using action 1, both users obtain very low utilities. This situation represents the congested network with low throughput of CSMA. In Table 9.1b, we show two Nash equilibria, where one of the users dominates the other. The dominating user has the utility of 6 and the dominated user has the utility of 3, which is unfair. In Table 9.1c, we show the mixed Nash equilibrium where two users have the probability 0.75 for action 0 and 0.25 for action 1, respectively. The utility for each user is 4.5.

Table 9.1. Example of two secondary users game (a) reward table (left most); (b) Nash equilibrium (middle left); (c) mixed Nash equilibrium (middle right); (d) correlated equilibrium (right most).

	0	1
0	(5,5)	(6,3)
1	(3,6)	(0,0)

	0	1
0	0	(0 or 1)
1	(1 or 0)	0

	0	1
0	9/16	3/16
1	3/16	1/16

	0	1
0	0.6	0.2
1	0.2	0

Next, we study a new concept of correlated equilibrium which is more general than Nash equilibrium and was first proposed by Nobel Prize winner, Robert J. Aumann [16], in 1974. The idea is that a strategy profile is chosen randomly according to a certain distribution. Given the recommended strategy, it is to the players' best interests to conform with this strategy. The distribution is called the correlated equilibrium.

We assume $N = 1$ and we omit the notation n. Define a finite K-user game in strategic form as $\mathcal{G} = \{K, (\Omega_i)_{i \in K}, (U_i)_{i \in K}\}$, where Ω_i is the strategy space for user i and U_i is the utility function for user i. Define Ω_{-i} as the strategy space for user i's opponents. Denote the action for user i and its opponents as $\mathbf{r}_i$ and $\mathbf{r}_{-i}$, respectively. Then, the correlated equilibrium is defined as:

Definition 9.10. *A probability distribution p is a correlated strategy of game $\mathcal{G}$, if and only if, for all $i \in K$, $\boldsymbol{r}_i \in \Omega_i$, and $\boldsymbol{r}_{-i} \in \Omega_{-i}$,*

$$\sum_{\boldsymbol{r}_{-i} \in \Omega_{-i}} p(\boldsymbol{r}_i, \boldsymbol{r}_{-i})[U_i(\boldsymbol{r}'_i, \boldsymbol{r}_{-i}) - U_i(\boldsymbol{r}_i, \boldsymbol{r}_{-i})] \leq 0, \forall \boldsymbol{r}'_i \in \Omega_i. \tag{9.8}$$

By dividing inequality in (9.8) with $p(\boldsymbol{r}_i) = \sum_{\boldsymbol{r}_{-i} \in \Omega_{-i}} p(\boldsymbol{r}_i, \boldsymbol{r}_{-i})$, we have

$$\sum_{\boldsymbol{r}_{-i} \in \Omega_{-i}} p(\boldsymbol{r}_{-i} | \boldsymbol{r}_i)[U_i(\boldsymbol{r}'_i, \boldsymbol{r}_{-i}) - U_i(\boldsymbol{r}_i, \boldsymbol{r}_{-i})] \leq 0, \forall \boldsymbol{r}'_i \in \Omega_i. \tag{9.9}$$

The inequality in (9.9) means that when the recommendation to user i is to choose action $\mathbf{r}_i$, then choosing action $\mathbf{r}'_i$ instead of $\mathbf{r}_i$ cannot obtain a higher expected payoff to i.

We note that the set of correlated equilibria is non-empty, closed and convex in every finite game. Moreover, it may include the distribution that is not in the convex hull of the Nash equilibrium distributions. In fact, every Nash equilibrium is a correlated equilibrium and Nash equilibria correspond to the special case where $p(\mathbf{r}_i, \mathbf{r}_{-i})$ is a product of each individual user's probability for different actions, i.e., the play of the different players is independent [16–18]. In Table 9.1b and c, the Nash equilibria and mixed Nash equilibria are all within the set of correlated equilibria. In Table 9.1d, we show an example where the correlated equilibrium is outside the convex hull of the Nash equilibrium. Notice that the joint distribution is not the product of two users' probability distributions, i.e., the two users' actions are not independent. Moreover, the utility for each user is 4.8, which is higher than that of the mixed strategy.

The characterization of the correlated equilibria set illustrates that there are solutions of correlated equilibria that achieve strictly better performance compared to the Nash equilibria in terms of the spectrum utilization efficiency and fairness. However, the correlated equilibrium defines a set of solutions which is better than Nash equilibrium, but it does not tell any more information regarding which correlated equilibrium is most suitable in practice. We propose two refinements. The first one is the maximum sum correlated equilibrium that maximizes the sum of utilities of the secondary users. The second is the maxmin fair correlated equilibrium that seeks to improve the worst-case situation. The problem can be formulated as a linear programming problem as:

$$\max_p \sum_{i\in K} E_p(U_i) \text{ or } \max_p \min_i E_p(U_i) \tag{9.10}$$

$$\text{s.t. } \begin{cases} p(\mathbf{r}_i, \mathbf{r}_{-i})[U_i(\mathbf{r}'_i, \mathbf{r}_{-i}) - U_i(\mathbf{r}_i, \mathbf{r}_{-i})] \le 0 \\ \forall \mathbf{r}_i, \mathbf{r}'_i \in \Omega_i, \forall i \in K \end{cases}$$

where $E_p(\cdot)$ is the expectation over p. The constraints guarantee the solution is within the correlated equilibrium set.

Next, we will exhibit a class of algorithm called regret-matching algorithm [18]. In particular, for any two distinct actions $\mathbf{r}_i \neq \mathbf{r}'_i$ in Ω_i and at every time T, the regret of user i at time T for not playing $\mathbf{r}'_i$ is

$$\mathcal{R}_i^T(\mathbf{r}_i, \mathbf{r}'_i) := \max\{D_i^T(\mathbf{r}_i, \mathbf{r}'_i), 0\} \tag{9.11}$$

where

$$D_i^T(\mathbf{r}_i, \mathbf{r}'_i) = \frac{1}{T}\sum_{t\le T}(U_i^t(\mathbf{r}'_i, \mathbf{r}_{-i}) - U_i^t(\mathbf{r}_i, \mathbf{r}_{-i})). \tag{9.12}$$

$D_i^T(\mathbf{r}_i, \mathbf{r}'_i)$ has the interpretation of average payoff that user i would have obtained, if it had played action $\mathbf{r}'_i$ every time in the past instead of choosing $\mathbf{r}_i$. The expression $\mathcal{R}_i^T(\mathbf{r}_i, \mathbf{r}'_i)$ can be viewed as a measure of the average regret. The probability $p_i(\mathbf{r}_i)$ for user i to take action $\mathbf{r}_i$ is a linear function of the regret. The algorithm was named regret-matching (no-regret) algorithm, because the stationary solution of the learning algorithm exhibits no regret and the play probabilities are proportional to the "regrets" for not having played other actions. The detail regret-matching algorithm is shown in Table 9.2. The complexity of the algorithm is $\mathrm{O}(L)$.

For every period T, let us define the relative frequency of users' action $\mathbf{r}$ played till T periods of time as follows

$$z_T(\mathbf{r}) = \frac{1}{T}\#\{t \le T : \mathbf{r}_t = \mathbf{r}\} \tag{9.13}$$

where $\#(\cdot)$ denotes the number of times the event inside the bracket happens and $\mathbf{r}_t$ is all users' action at time t. The following theorem guarantees that the adaptive learning algorithm shown in Table 9.2 has the property that z_T converges almost surely to a set of the correlated equilibria.

Theorem 9.1. [18] *If every player plays according to adaptive learning algorithm in Table 9.2, then the empirical distributions of play z_T converge almost surely to the set of correlated equilibrium distributions of the game G, as $T \to \infty$.*

In the simulations, we employ the maximal sum utility function as the objective. In Fig. 9.1a, we show the different equilibria as a function of G_0 for the three-user game. We show the results of the gain obtained by the greedy user in the Nash equilibrium point (NEP), the gain obtained by the victim of the greedy user in NEP,

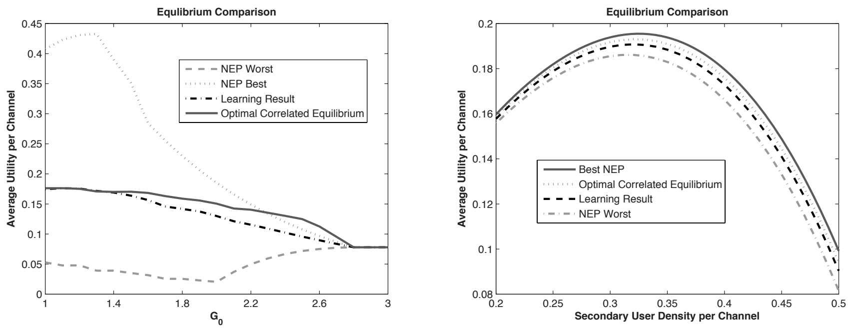
Table 9.2. The regret-matching learning algorithm © *2007 IEEE. Reprinted, with permission, from [4]*.

Initialize arbitrarily probability for taking action of user i, $p_i^1(\mathbf{r}_i), \forall i \in K$
for t = 1,2,3,... 1. Find $D_i^T(\mathbf{r}_i, \mathbf{r}_i')$ as in (9.12) 2. Find average regret $\mathcal{R}_i^T(\mathbf{r}_i, \mathbf{r}_i')$ as in (9.11) 3. Let $\mathbf{r}_i \in \Omega_i$ be the strategy last chosen by user i, i.e., $\mathbf{r}_i^t = \mathbf{r}_i$. Then probability distribution action for next period, p_i^{t+1} is defined as $p_i^{t+1}(\mathbf{r}_i') = \frac{1}{\mu}\mathcal{R}_i^T(\mathbf{r}_i, \mathbf{r}_i') \quad \forall \mathbf{r}_i' \neq \mathbf{r}_i$ $p_i^{t+1}(\mathbf{r}_i) = 1 - \sum_{\mathbf{r}_i' \neq \mathbf{r}_i} p_i^{t+1}(\mathbf{r}_i')$, where μ is a certain constant that is sufficiently large.

the learning result and the optimal correlated equilibrium calculated by linear programming. Here the action space is $[0.1, 0.2, \ldots, 1.5]$. When G_0 is large, there is less penalty for greedy behaviors. So all users tend to transmit as aggressively as possible. This results in the prisoners' dilemma [19], where all users suffer. When G_0 is less than 2.8, the greedy user can have better performance (NEP best) than that (NEP worst) of the cooperative user. Due to the less significant penalty if all users transmit aggressively, the game will not degrade to the prison dilemma. However, the performances are quite unfair for the greedy users with best NEP and the cooperative users with worst NEP. All users have the same utility in the correlated equilibrium and learning result. So fairness is better than that of the NEP. When G_0 is from 2.2 to 2.8, the correlated equilibrium has a better performance even than that of the greedy user (NEP best). When G_0 is from 1.4 to 2.8, the optimal correlated equilibrium has a better performance than that of the learning result. When G_0 is sufficiently small, most of the uncooperative strategies are eliminated by significant penalty. Consequently, the learning result has the same performance as that of the optimal correlated equilibrium.

In Fig. 9.1b, we show the network performance of the proposed algorithm. For simplicity, we assume the hidden terminal problem [14] has been solved. We show the average user utility per channel as a function of the network density. When the network density is small, the average utility increases since there is an increasing number of users occupying the channel. When the user density is sufficiently large, the utility begins to decrease due to the collisions. The best NEP and worst NEP are different while the correlated equilibrium and learning result achieve almost the same performance as the best NEP and 5–15% better than the worst NEP.

There are many other works for learning based on finite-state Markov decision process (MDP), such as the decentralized cognitive medium access based on partial observable Markov decision process (POMDP), which is presented in [20]. Some other learning schemes include reinforcement learning, Q-learning and so on. All these techniques can be utilized for the spectrum access for cognitive radios.



(a) (b)

Fig. 9.1. (**a**) Utility function versus. G_0 (for three users) and (**b**) network performances [© *2007 IEEE. Reprinted, with permission, from [4]*].

9.4 Referee for Mediation

In this section, a concept of virtual referee [5,21] is introduced to improve wireless resource usage of cognitive radios. We use this referee approach for an example application to conduct channel assignment, adaptive modulation and power control for multi-cluster cognitive networks. The goal is to minimize the overall transmitted power under the constraints that each cognitive user has the desired throughput and each cognitive user's power is bounded. Each cognitive user in the different clusters minimizes its own utility function, e.g., transmitted power, in a distributed and non-cooperative game by employing water-filling scheme [22]. We define the channel set that the ith cognitive user can allocate to its throughput R_i as transmission channel set S_i. Each channel can be occupied by more than one cognitive user but not necessarily by all users. Within the transmission channel set, the user would allocate the throughput to different channels by the algorithms, such as water-filling [22], so that the utility such as the power can be optimized. When the interferences are severe, the channel will be over crowded with users and consequently, the radio resource cannot be efficiently utilized. Under this condition, a virtual referee will be introduced to mediate the resource usage, so that the game outcome can be improved. This virtual referee can be the base station, access point or cluster head. This approach can significantly improve the network performance without adding much hardware to cognitive networks.

The K co-channel clusters are taken into consideration. Each cluster consists of one cognitive radio link. The total number of channels is L. The ith user's signal to interference noise ratio (SINR) at channel l can be expressed as:

$$\Gamma_i^l = \frac{P_i^l G_{ii}^l}{\sum_{k \neq i} P_k^l G_{ki}^l + N_0} \tag{9.14}$$

where P_k^l and G_{ki}^l is the transmitted power and propagation loss from the kth cognitive source to the ith cognitive destination in the lth channel, respectively, and N_0 is the thermal noise level.

Rate adaptation such as adaptive modulation provides each channel with the ability to match the effective bit rates, according to the interference and channel conditions. MQAM is a modulation method with high spectrum efficiency. In [23], for a desired rate r_i^l of MQAM, the BER of the lth channel of the ith user can be approximated as a function of the received SINR Γ_i^l by:

$$\mathrm{BER}_i^l \approx c_1 \mathrm{e}^{-c_2\left(\Gamma_i^l/2^{r_i^l}-1\right)} \tag{9.15}$$

where $c_1 \approx 0.2$ and $c_2 \approx 1.5$ with small BER_i^l. Rearranging (9.15), for a specific desired BER_i^l, the ith user's transmission rate of the lth channel for the SINR Γ_i^l and the desired BER_i^l can be expressed as:

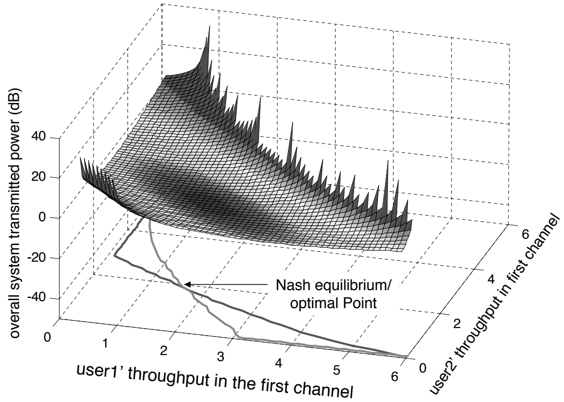
$$r_i^l = W \log_2(1 + c_3^i \Gamma_i^l). \tag{9.16}$$

In order to compare the Nash equilibriums (NEP) and the optimal solution for power minimization, a simple two-user two-channel example is illustrated as follows. The simulation setup is: BER $= 10^{-3}$, $N_0 = 10^{-3}$, the maximal power for each user over different channels is $P_{\max} = 10^4$ and channel gain matrices are

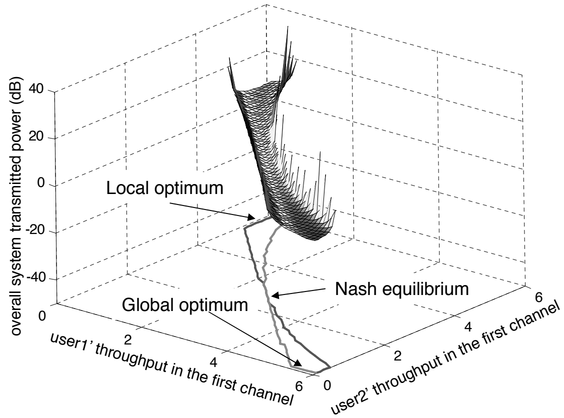
$$G^1 = \begin{bmatrix} 0.0631 \; 0.0100 \\ 0.0026 \; 0.2120 \end{bmatrix}, G^2 = \begin{bmatrix} 0.4984 \; 0.0067 \\ 0.0029 \; 0.9580 \end{bmatrix}.$$

Figure 9.a shows the overall power contour as a function of two users' rate allocations, where each user's minimal rate requirement $R_1 = R_2 = 6$. The two curves show the minimal locations for the two users' own power when the interference from the other user is fixed, respectively. Each user tries to minimize its power by adjusting its rate allocation so that the operating point is more close to the curve. Consequently, the cross is a Nash equilibrium, where no user can reduce its power alone. We can see that the Nash equilibrium under this setup is unique and optimal for the overall power. (It is worth mentioning that the feasible domain is not convex at all.) Figure 9.b shows the situation when $R_1 = R_2 = 8$. Because the rate is increased, the co-channel interferences are increased and the NEP is no longer the optimum. There exists more than one local optimum, and the global optimum occurs when user does not occupy the channel 1. Figure 9.c shows the situation when $R_1 = R_2 = 8.5$. The contour graph is not connected. There are two NEPs and two local optima. Under the above two conditions, we need to remove users from using the channels. If we further increase $R_1 = R_2 = 10$, there exists no feasible area, i.e., neither user can have a resource allocation that satisfies both power and rate constraints. In this case, the minimal rate requirement should be reduced.

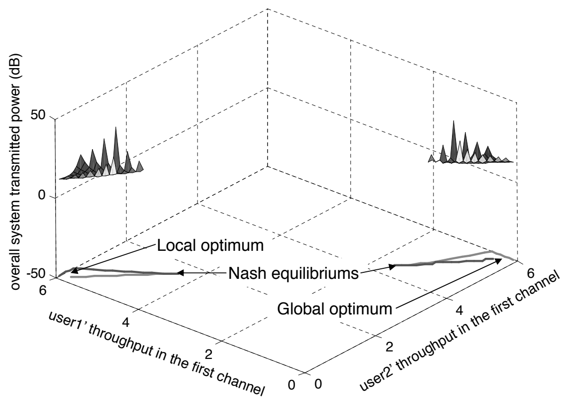
From the above observations, we can see that the behaviors of the optimal power minimization solution and NEP depend on how severe interferences are. In order to let NEP converge to the desired solution, we need to find a criterion to decide whether the users can make a good use of the channels like the situation in Figure 9.a.



(a) Unique optimal NEP



(b) Multiple local optima



(c) Multiple NEPs

Fig. 9.2. Two-user example [© *2007 IEEE. Reprinted, with permission, from [5]*].

If not, we should decide which user should be kicked out of using specific channels. The criterion is to check whether the KKT condition [24] is satisfied. Specifically, if the co-channel interferences are too severe, the constraints of throughput and maximal transmitted power are not satisfied. As a result, NEP is not a local optimum.

Before developing the proposed algorithm, we analyze two extreme cases. In the first case, the groups of channels are assigned to different clusters without overlapping such that there are no co-channel interferences among clusters. We call it the fixed channel assignment scheme. However, this extreme method has the disadvantage of low spectrum efficiency because of the low frequency re-usage. In the second extreme case, all cognitive users share all the channels. We call it pure water-filling scheme. From Fig. 9.b and Fig. 9.c, we can see that the system can be balanced at the undesired point, because of the severe inter-cluster co-channel interferences. So the facts motivate us to believe that the optimal resource allocation is between these two extreme cases, i.e., each channel can be shared by only a group of selected users for transmission.

In Fig. 9.3, we show the block diagram of the proposed algorithm from system point of view. We initially set S_i to have all channels. Then the non-cooperative competition for radio resources is employed. After the system is iteratively balanced by the water-filling among cognitive radio users, if the system is balanced in a desired solution, the water-filling is continuously employed. Otherwise, some users must remove some channels from the transmission group S_i. If the removal can make all users balanced in the desired NEP, the algorithm continues in the water-filling step. Otherwise, the user removal step is continued, until no user can be removed or the desired NEP is achieved. If no user can be removed and the desired NEP is still not achieved, the desired throughput requirement R_i has to be reduced.

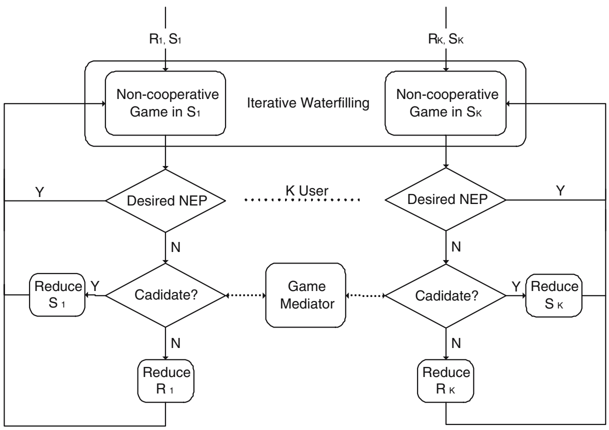


Fig. 9.3. Proposed distributed referee approach [© *2007 IEEE. Reprinted, with permission, from [5]*].

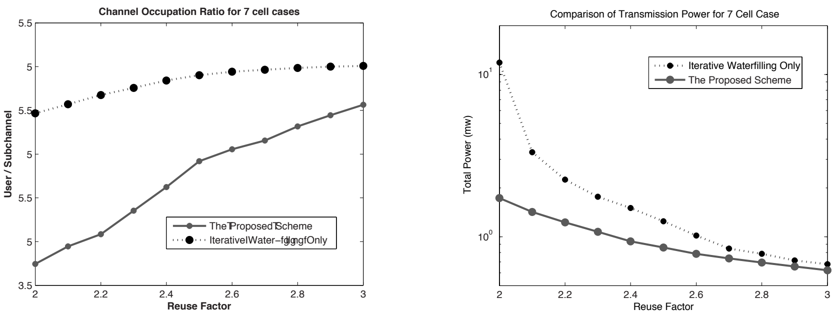


Fig. 9.4. (**a**) User per channel and (**b**) power saving [© *2007 IEEE. Reprinted, with permission, from [5]*].

The complexity of the proposed referee-based scheme is $\mathrm{O}(N \log N)$, where N is the number of channels. The convergence speed of the non-cooperative competition is similar to that of closed-loop power control proposed in [25,26]. The overhead for the proposed scheme occurs only when the system cannot be balanced in a good Nash equilibrium. Under this condition, a referee needs to collect information from all the co-channel interfered clusters. The frequency for this overhead is much lower than that of the non-cooperative competition. The collected information includes power value, channel gain value and noise-plus-interference variance value over all channels. Since all these values are consistently obtained by all the distributed users at any time, there is no need for extra measurement. The amount of this information is also small and can be exchanged among the cells with few packets. So the overhead is negligible. In summary, this referee-based scheme imposes little burden on wireless sensor network implementation.

We consider the simulations with 32 channels and seven cognitive radio links. The overall bandwidth is 6.4 MHz. The receiver thermal noise is -70 dBm. The required BER of the transmitted symbols is 10^{-3} for every subchannel and user. We define the reuse factor R_u as the distance between two base stations D over the cell radius r which is set as 100 m, which is one of the main factors to affect the severeness of co-channel interference. The rate constraint is set as 10 Mbits for each user. In Fig. 9.4a, we show the average number of users per channel. In Fig. 9.4b, we show the overall transmitted power versus reuse distance R_{u} for the pure water-filling algorithm and the proposed algorithm. The smaller the reuse distance R_{u} is, the higher the co-channel interference. We can see that the proposed algorithm can reduce the overall power about 90% when the co-channel interferences are severe ($R_{\mathrm{u}} = 2$), because more users are kicked out in this case. When R_{u} increases, the co-channel interferences reduce. Consequently, water-filling and proposed schemes yield the same overall transmitted power.

The referee-based approach creates a virtual referee to mediate the network performances. If the autonomous cognitive users cannot share the network resources efficiently, the referee will make some mandatory changes for resource usage so as to improve the system performances and game outcomes. There is no need to add additional hardware, while the performances can be greatly improved.

9.5 Threat and Punishment Using Repeated Interactions

In some types of the autonomous and distributed wireless cognitive radio networks, tasks need to be performed cooperatively while greediness might lead to the performance breakdown. The individual user may act cooperatively such that the overall system performance is high, or they may act non-cooperatively where everybody suffers low efficiency. However, if only one user deviates from the cooperative agreement, it can get benefits. In order to prevent users from greediness, repeated interaction, such as repeated game, is proposed to enforce cooperation among cognitive users. The basic rationale is to punish the user that deviates by playing non-cooperatively in the near future, such that the benefits obtained in a short-term deviation will be eliminated by a long-term punishment. In this section, we outline the punishment approach and give two examples.

The basic idea of the threat and punishment using repeated game comes from the concept of Cartel in the economics literature [7]. Cartel means the combination of independent commercial or industrial enterprises designed to limit competition. The soul of Cartel maintenance is to construct contracts among independent individuals for cooperative benefits and non-cooperative punishment, so as to limit inefficient competition. Next, we combine the idea with the repeated game theory, so that the new approach will punish anyone who deviates from cooperation.

To analyze the outcome of a game, the *Nash equilibrium* is a well-known concept, which states that in the equilibrium every agent selects a utility-maximizing strategy given the strategies of other agents. However, one problem with an NEP is that it is not necessarily very efficient in performances. If the users can play cooperatively, the performances can be greatly improved. Thus, the question arises as to how to enforce the greedy users to cooperate with each other. The repeated game provides us possible mechanisms to enforce the users to cooperate by considering long-term scenarios. In the repeated games, the players face the same static game in every period, and the player's overall payoff is a weighted average of the payoffs in each stage over time. In the repeated game, the players can observe some information reflecting their opponents' past play. Hence, they are able to condition their future plays on the observed information in history to obtain better equilibriums.

Definition 9.11. *Let G be a static game and β be a discount factor. The $\mathcal{T}$-period repeated game, denoted as $G(\mathcal{T},\beta)$, consists of game G repeated $\mathcal{T}$ times. The payoff for such a game is given by*

$$V_i = \sum_{t=1}^{\mathcal{T}} \beta^{t-1} u_i^t \tag{9.17}$$

where u_i^t denotes the payoff to player i in period t. If $\mathcal{T}$ goes to infinity, then $G(\infty,\beta)$ is referred as the infinitely repeated game. In the following, we use infinitely repeated game.

Now the question is whether cooperation among users can be enforced by the repeated games to generate better performances. The Folk's theorem [19] for infinitely repeated games asserts that if the player's discount factor β approaches 1, any feasible, individually rational payoff can be enforced by an equilibrium. This equilibrium can yield better performances than those of static game NEP. We need to further develop the game rule for enforcing cooperation among users to achieve this better equilibrium.

The basic idea for the proposed Cartel maintenance repeated game framework is to provide enough threat to greedy users so as to prevent them from deviating from cooperation. First the cooperative point is obtained so that all users have better performances than those of non-cooperative NEP. However, if any user deviates from cooperation while others still play cooperatively, this deviating user has a better utility, while others have relatively worse utilities. If no rule is employed, the cooperative users will also have incentives to deviate. Consequently, the network deteriorates to non-cooperation with inefficient performances. The proposed framework provides a mechanism so that the current defecting gains of the selfish user will be outweighed by future punishment strategies from other users. For any rational user, this threat of punishment prevents them from deviation. So cooperation is enforced.

To implement the mechanism, we propose a trigger strategy to introduce punishment on the defecting users. In the trigger strategy, the players start with cooperation. Assume that each user can observe the public information (e.g., the outcome of the game), P_t at time t. Examples of this public information can be the successful transmission rate, network throughput, etc. Notice that such public information is mostly imperfect or simply partial information about the users' strategies. Here we assume a larger P_t stands for a higher cooperative level, resulting in higher performances for all users. Let the cooperative strategies be $\bar{\lambda} = [\lambda_1, \lambda_2, ..., \lambda_K]^{\mathrm{T}}$ and the non-cooperative strategies be $\bar{s} = [s_1, s_2, ..., s_K]^{\mathrm{T}}$, respectively. The trigger-punishment game rule is characterized by three parameters: the optimal punishment time T, trigger threshold P^* and the cooperative strategy $\bar{\lambda}$. Trigger punishment strategy $(\bar{\lambda}, P^*, T)$ for distributed user i is given as follows:

(a) User i plays the strategy of the cooperative phase, $\bar{\lambda}$, in period 0.
(b) If the cooperative phase is played in period t and $P_t > P^*$, user i plays the cooperative phase in period $t + 1$.
(c) If the cooperative phase is played in period t and $P_t < P^*$, user i switches to a punishment phase for $T - 1$ periods, in which the players play a static Nash equilibrium $\bar{s}$ regardless of the realized outcomes. At the Tth period, play returns to the cooperative phase.

Note that $\bar{s}$ generates the non-cooperative outcome, which is much worse than that generated by the cooperative strategy $\bar{\lambda}$. Therefore, the selfish users that deviate will have much lower utilities in the punishment phase. Moreover, the punishment time T is designed to be long enough to let all cheating gains of the selfish users be outweighed by the punishment. So the users have no incentive to deviate from cooperation, since the users aim to maximize the long-run payoffs over time.

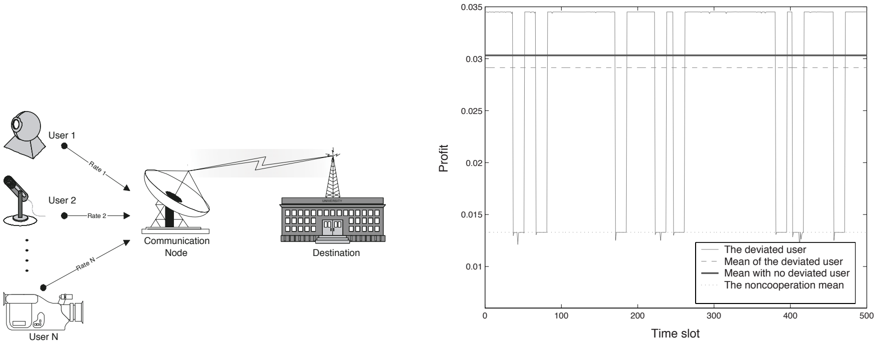


Fig. 9.5. (**a**) Wireless network block diagram, (**b**) Punishment for deviation [© *2004 IEEE. Reprinted, with permission, from [7]*].

Next, we study two examples using the proposed framework. The first one is for cognitive radio multiple access networks. The second example further investigates the learning schemes if the cognitive radio users do not know how to cooperate.

In the first example, we employ the proposed framework to a multi-user network shown in Fig. 9.5a. There are many distributed users and one communication node (e.g., cluster head). Each user can transmit its data packets to the communication node by using the multiple access protocols such as Aloha, CSMA, etc. The communication node has the ability to transmit the data packets to the remote destination via a wireless link. We assume that there is a reliable feedback channel. So, the system can be described as multiple users sharing a communication link. Each user can control its transmission rate. The users need to compete with each other for the communication link which is fluctuating due to the wireless channel conditions. Thus one user's rate can affect the performances of other users and the whole system. So it is necessary to find a rate control algorithm such that the system can operate at the optimal point. Moreover, it is hard to have communication channels among cognitive users. Therefore, a distributed algorithm is required for rate control.

For distributed users in the network, there are costs to transmit their packets and benefits if their packets are successfully transmitted. Each user's profit is defined as the benefits minus the cost. The users are able to adapt their packet transmission rates for the cooperation or punishment play. They can observe their successful packet transmission probability, and correspondingly play cooperation or non-cooperation. Based on the proposed framework, we derive the optimal parameters of the packet transmission rate, punishment time and trigger threshold for the distributed greedy users. In Fig. 9.5 b, we show how the scheme punishes the cheating user. We assume one user deviates from the cooperative rate λ^* and transmit at the higher rate s, while others transmit at λ^*. We show that the profit of this deviating user fluctuates over time. For comparison, we also show the average profits (as the straight lines) when the user transmits at optimal rate from overall system point of view, cooperative rate λ^* and non-cooperative rate. We can see that at first the user does get more profit than the mean without the deviated user by diverging from λ^*. However, this deviation is soon detected by others' and the punishment phase is performed by other users. The non-cooperation mean is much lower than that during cooperation. The mean of this deviated user is lower than the mean without the deviated user, because the deviation gain is eliminated by others over time. This shows the reason why the proposed scheme can enforce cooperation among users by threatening punishment.

In the second example, we further investigate the combination of learning schemes. In some ad hoc cognitive networks, cognitive users need to forward others' packet so as to communicate with each other. Forwarding the others' packets consumes the user's own limited battery resource. Therefore, it may not be of the autonomous user's best interest to forward all the arriving packets. In fact, it is reasonable to assume that the users are selfishly maximizing their own benefits by dropping others' packets. However, not forwarding others' packets will severely affect the network connectivity and the proper functionality of the network, which in turn impairs the users' own benefits as well. The non-cooperation usually causes very low system and users' performances. Therefore, it is very crucial to design a mechanism to enforce cooperation among greedy users. Moreover, even though the users would like to cooperate, they might not know how to cooperate. So it is important to develop self-learning algorithms so that the cooperative points can be studied distributively in the autonomous users.

We try to propose a distributed self-learning repeated game framework to enforce cooperation in performing packet-forwarding tasks as shown in Fig. 9.6a. The framework has two major schemes: first, an adaptive repeated game scheme ensures cooperation among ad hoc cognitive users, which maintains the current cooperative packet-forwarding probabilities. The repeated game scheme provides the users with a mechanism that any deviating user would be punished enough by others in the future, so that no user has incentive to deviate. Second, a self-learning scheme tries to find the better cooperative probabilities that are feasible and benefit all users. Starting from non-cooperation, the above two proposed schemes are employed iteratively. Better cooperation is discovered and maintained over iterations, until convergence to some close optimal solution.

In Fig. 9.6b, we show the simulation results of the proposed framework for utility and packet forwarding probability over time. Initially, packet-forwarding probability $\alpha = 0$, because of the non-cooperative transmission. Then the system tries to find a better packet transmission rate. When it finds a better solution, all users adapt its α to the value. However, since the punishment period T is not adjusted to an optimal value, the deviation can have benefits. So there exists a period that the utility and α switch from cooperation to non-cooperation. In this period, T is increased until every user realizes that there is no benefit for deviation because of the long period of punishment. If the system is stable for a period of time, a new α is determined to see whether the performance can be improved. If so, the new value is adopted, otherwise the original value is restored. So the packet-forwarding probability is adjusted until the optimal solution is found, and the learned utility function is a non-decreasing function.

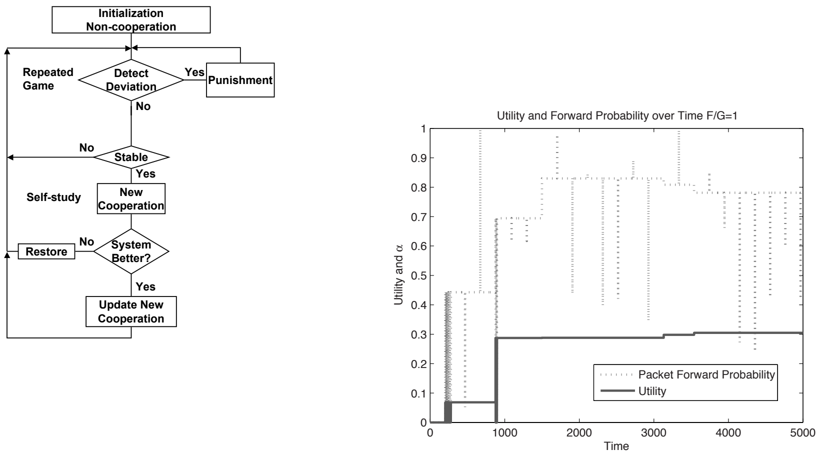


Fig. 9.6. (**a**) Proposed self-learning repeated game and (**b**) self-learning curve [© *2005 IEEE. Reprinted, with permission, from [6]*].

9.6 Spectrum Auction

In this section, we first discuss the basics of auction theory. Then we investigate the mechanism design for auctions. Finally, we use an example to explain how to utilize the auction theory for spectrum usage in cognitive radios.

Auction theory is important for practical, empirical and theoretical reasons. First, a large amount of wireless networking and resource allocation problems can be formulated as auction theory. For example, the routing problem for self-interested users is studied in [27]. Second, the auction theory has a simple game setup, and many theoretical results are available for analysis. The definition of auction is as follows.

Definition 9.12. *A market mechanism in which an object, service or set of objects, is exchanged on the basis of bids submitted by participants. Auction provides a specific set of rules that will govern the sale or purchase (procurement auction) of an object to the submitter of the most favorable bid.*

The interactions and outcome of an auction are determined by the *rules*, which include four components:

- *Information*: what the auctioneer and bidders know before the auction starts.
- *Bids*: what the bidders submit to the auctioneer to express their interests in the good.
- *Allocation*: how the good is allocated among the bidders as a function of the bids.
- *Payments*: how the bidders pay the auctioneer as functions of the bids and allocation.

To implement auction theory in wireless networking and resource allocation, the credit-based system is usually proposed. The individual user can select to pay for some kind of services such as a route. The payment can be implemented via a certain central "bank" system. However, this requires more control than the other game theory approaches, such as non-cooperative games. Moreover, in order to achieve different design goals such as the network total benefit, the auction method shall be designed according to different available information. Mechanism design is the tool for game and auction design.

Mechanism design is the subfield of microeconomics and game theory that considers how to implement good system-wide solutions to problems that involve multiple self-interested agents, each with private information about their preferences. The goal is to achieve a social choice function implemented in distributed systems with private information and rational agents. The design criteria can be different as follows:

1. *Efficiency*: select the outcome that maximizes total utility.
2. *Fairness*: select the outcome that minimizes the variance in utility.
3. *Revenue maximization*: select the outcome that maximizes revenue to a seller (or more generally, utility to one of the agents).
4. *Budget-balance*: implement outcomes that have balanced transfers across agents.
5. Pareto optimality.

One well-known auction mechanism that achieves the efficient allocation is the Vickery–Clarke–Groves (VCG) auction [28]. In a VCG auction, the bidders are asked to reveal their bids simultaneously, from which the auctioneer determines the efficient allocation. The auctioneer then asks each bidder i to pay for the "performance loss" of other bidders due to bidder i's participation in the auction, which involves solving one additional optimization problem for each bidder. It is well known that it is a (weakly) dominant strategy for the bidders to bid truths in the VCG auction, i.e., revealing their true rate increase functions. As a result, the VCG auction achieves the efficient allocation.

The limitations of the VCG auction for cognitive radio users are as follows. First, the users (bidders) need to submit the complete information to the central control unit serving as the auctioneer, which involves revealing users' complete private information. This might be overheard by other users and so can lead to security problems. Also, accurately specifying the information requires much signaling overhead and communication bandwidth, which may significantly reduce the network performance. Furthermore, it is usually computationally expensive for solving the optimization problems.

Due to these concerns, in [29], two simpler share auctions are proposed for cognitive radios. First, we discuss the system model. Suppose K user-CDMA is utilized with processing gain B. The received SINR is given by

$$\Gamma_i = \frac{P_i G_{ii}}{N_0 + \frac{1}{B}(\sum_{j \neq i} P_j G_{ji})} \tag{9.18}$$

where P_i is the transmit power, G_{ij} is the channel gain and N_0 is the thermal noise power. User i receives a strictly concave increasing utility as $U_i(\Gamma_i, \theta_i)$, where θ_i is user-dependent priority parameter.

Next, we discuss the two share auctions, namely the *SNR auction* and the *power auction*. The main advantages of the two auctions are the simplicities of bids and allocation. The rules of the two auctions are described below, with the only difference being in payment determination.

9.6.1 Share Auction

- *Information*: The auctioneer (which can be a cluster head) announces a positive *reserve bid* $\zeta > 0$ and a unit *price* $\pi > 0$ to all users before the auction starts. Here ζ ensures unique outcome, and π is for unit SINR or received power.
- *Bids*: User i submits $b_i \geq 0$ to the auctioneer.
- *Allocation*: The auctioneer allocates transmit power according to

$$P_i G_{ii} = \frac{b_i}{\sum_{j=1}^{N} b_j + \zeta} P \tag{9.19}$$

  where P is the overall allowable power.
- *Payments*: In an SNR auction, cognitive user i pays the auctioneer

$$C_i = \pi \bigtriangleup \mathtt{SNR}_i. \tag{9.20}$$

  In a power auction, source i pays the relay

$$C_i = \pi P_i G_{ii}. \tag{9.21}$$

A bidding profile is defined as the vector containing the users' bids, $\mathbf{b} = (b_1, ..., b_K)$. The bidding profile of user i's opponents is defined as $b_{-i} = (b_1, ..., b_{i-1}, b_{i+1}, ..., b_K)$, so that $\mathbf{b} = (b_i; b_{-i})$. User i chooses b_i to maximize its payoff $U_i(b_i; b_{-i}, \pi)$. The desirable outcome of an auction is called a *Nash equilibrium* (NE), which is a bidding profile $\mathbf{b}^*$ such that no user wants to deviate unilaterally, i.e.,

$$U_i\left(b_i^*; b_{-i}^*, \pi\right) \geq U_i\left(b_i; b_{-i}^*, \pi\right), \forall i \in 1, \ldots, K, \forall b_i \geq 0. \tag{9.22}$$

Define user i's *best response* (for fixed b_{-i} and price π) as

$$\mathcal{B}_i\left(b_{-i}, \pi\right) = \{b_i | b_i = \arg\max_{\tilde{b}_i \geq 0} U_i(\tilde{b}_i; b_{-i}, \pi)\} \tag{9.23}$$

which in general could be a set. An NE is also a fixed point solution of all users' best responses. In [29], the following four questions for both auctions are answered. First, an NE does exist, and in some mild conditions, the NE is unique. This NE can be converged by using a distributed iterative algorithm with some partial information that is private and local. The SNR auction with log utility can achieve weighted max–min fairness, while the power auction can achieve social optimum for large bandwidth.

9.7 Mutual Benefits Through Bargaining

In order for the distributed cognitive users to cooperate with each other, one method is to give the individuals mutual benefits for cooperative behavior. In most existing literature, the benefit incentive approach is performed in a framework of "pricing anarchy," where a price is announced by the system, so that the distributed users have to pay high price for non-cooperation and the cooperative behaviors will be rewarded. However, there are many potential design challenges for the pricing technique to be employed in cognitive networks. First, the price itself may not represent the true benefits of the cognitive users. Instead, the price might be artificial so that autonomous users may just ignore it. Furthermore, pricing technique needs a lot of computation power and signaling to calculate the optimal price. This is especially hard to implement in cognitive networks. In addition, if the utility of each user is not convex, there might be many local optima for the pricing methods. Finally, for the heterogeneous networks and for the resource allocation with integer/combinatorial optimization, the pricing techniques are hard to be effective. Because of the above reasons, we need to have novel perspective and find new approaches to give users mutual benefits to cooperate.

In daily life, a market is served as a central gathering point, where people can exchange goods and negotiate transactions, so that people will be satisfied through bargaining. Similarly, in wireless cognitive networks, there exist some nodes, like cluster heads, that can serve as a function of the market. The distributed cognitive users can negotiate via these nodes to cooperate in making the decisions on the resource usage, such that each of them will operate at its optimum and joint agreements are made about their operating points. Such a fact motivates us to employ the cooperative game theory [3,30,31], which can achieve the crucial notion of fairness and maximize the overall system performances. The idea is to negotiate among users so that the mutual benefits can be obtained, which enlightens us with the new perspective on how to provide incentive for cooperation. In the following, we list one possible problem formulation, a basic illustration of the proposed approaches and some simulation results.

We have proposed the cooperative game theory approaches for resource allocation in multiple-user multiple-channel scenario within a cluster of cognitive networks. The problem can be formulated in the following example. There are K users and a total of N channels. Each channel can be occupied by only one user so as to avoid severe co-channel interferences and maintain the basic link quality. Since a channel condition for a specific channel may be good for more than one user, there is a competition among users for their transmissions over these good channels. So this is where the game concept comes in. Moreover there are some other practical constraints. For example, the maximal transmitted power for each user is bounded by the maximal transmitted power $P_{\max}$, and each user has a minimal rate requirement $R^i_{\min}$. To formulate the problem, we define $a_{ij} = [\mathbf{A}]_{ij} = 1$, if the $i\text{th}$ user occupies the $j\text{th}$ channel; $a_{ij} = 0$, otherwise, and $[\mathbf{P}]_{ij}$ as the corresponding power. One example of the optimization goal is to determine different users' channel assignment matrix $\mathbf{A}$ and power matrix $\mathbf{P}$ such that the network objective function U will be maximized, i.e.,

$$\max_{\mathbf{A},\mathbf{P}} U \tag{9.24}$$

$$\text{subject to} \begin{cases} \text{Assignment: } \sum_{i=1}^{K} a_{ij} = 1, \forall j \\ \text{Minimal rate: } R_i \geq R_{\min}^i, \forall i \\ \text{Maximal power: } \sum_{j=1}^{N} P_{ij} \leq P_{\max}, \forall i \end{cases}$$

where R_i is the ith user's rate and U can have different definitions for network objectives such as:

- Maximal rate: $U = \sum_{i=1}^{N} R_i$.
- Max–min fairness: $U = \min R_i$.
- Nash bargaining solutions: $U = \prod_{i=1}^{K} \left(R_i - R_{\min}^i\right)$.

The first two network objectives are widely studied in the literature. In [3], we proposed the concept of Nash bargaining solution (NBS), because of the following two reasons: first, it can be shown that this network objective will ensure NBS fairness of allocation in the sense that this NBS fairness is a generalized proportional fairness. From the simulation results, this NBS fairness ensures that users' allocated resources are not affected by other users' situations. Second, cooperative game theories prove that there exists a unique and efficient solution under the six axioms shown in [19]. The intelligent merit of this NBS solution is that it can provide a special new tradeoff between the fairness and efficiency, which is widely researched recently in academia and industry.

The difficulty to solve (9.24) by traditional methods lies in the fact that the problem itself is a constrained combinatorial problem and the constraints are nonlinear. Thus the complexities of the traditional schemes are high especially with a large number of users. Moreover, distributed algorithms are desired for cognitive networks, while centralized schemes are dominant in the literature. To develop algorithms that can be easily deployed in distributed cognitive networks, we outline the ideas of the proposed approaches as follows:

Bargaining for two-user case: Due to the facts that in social life most negotiations are taken between two parties, we first consider the case in which the number of users $K = 2$ and we will develop a fast two-user bargaining solution. Since different users might have different gains over the same channel, the intuitive idea is to allow two users to negotiate and exchange their occupied channels such that mutual benefits will be obtained. The difficulty is to determine how to optimally exchange channels, which is a complex integer programming problem. An interesting low complexity algorithm was given in [22]. The idea is to sort the order of channels first and then to use a simple two-band partition for the channel assignment. When signal to noise ratio (SNR) is high, the two-band partition for two-user channel assignment can be near optimal for the optimization goal. The possible solution has the complexity of $O(N^2)$ and can be further improved by using a binary search algorithm with a complexity of only $O(N \log N)$.

Multiple users using coalitions: For the case in which the number of users is larger than two, the computational complexity is very high with respect to the number of channels. Here, we propose a two-step iterative scheme: first, users are

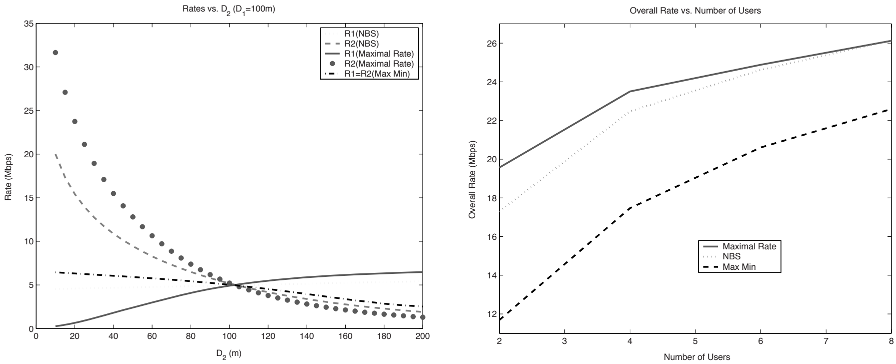


Fig. 9.7. (**a**) Each user's rate versuss. D_2 and (**b**) overall rate versus. number of users [© *2005 IEEE. Reprinted, with permission, from [3]*].

grouped into pairs, which are called *coalitions*. Then with each coalition, the two-user solution is employed for two users to negotiate and improve their performances by exchanging channel sets. Further, the users are regrouped and then renegotiated again. The above regrouping-and-negotiation iteration is repeated until convergence. By using this scheme, the computational cost can be greatly reduced. Some algorithms such as Hungarian method [32] can be utilized to find the optimal coalition pairs in each round. These minimal optimization efforts can be performed in the central point, such as the cluster head, while lower implementation costs are imposed on the distributed less-sophisticated users. Moreover, the above-mentioned approach can also be generalized to other formulated problems dealing with multi-user communications with different optimization goals and constraints.

To demonstrate the effectiveness of the proposed scheme, a simulation is conducted for a multiple-cognitive-user cluster with 32 channels. In Fig. 9.7a, a two-user case is studied. The rates of both users for the NBS, maximal rate and max–min schemes are shown versus the second user's distance from base station D_2. Here the first user's location is fixed at 100 m ($D_1 = 100$). For the maximal rate scheme, the user closer to the base station has a higher rate, and the rate difference is very large when D_1 and D_2 are different. For the max–min scheme, both users have the same rate which is reduced when D_2 increases. This is because the system has to accommodate the user with the worst channel condition. While for the NBS scheme, user 1's rate is almost the same regardless of D_2 and user 2's rate is reduced when D_2 increases. This shows that the NBS solution is fair in the sense that the user's rate is determined only by its channel condition and not by other interfering users' conditions.

In Fig. 9.7b, we show the sum of all users' rates versus the number of users in the system for three schemes. We can see that all three schemes have better performances when the number of users increases. This is because of multi-user diversity, provided by the independent varying channels across the different users. The performance improvement saturates gradually. The NBS scheme has a similar performance to that of the maximal rate scheme and has a much better performance than that of the max–min scheme. The performance gap between the maximal rate scheme and the NBS scheme reduces when the number of users is large. This is because more bargain pair choices are available to increase the system performance. The simulation results show that the proposed NBS scheme achieves a good tradeoff between fairness and efficiency.

We propose the idea of mutual benefits using bargaining for ensuring cooperation of resource allocation. By using cooperative game theory such as Nash bargaining solution and coalition, users' performances can be improved by locally exchanging the resources. The wireless cognitive network performance can be significantly improved and fairness among distributed users can be ensured in a self-organized way. Many other works are based on the proposed idea. In [33], a dynamic spectrum access scheme was proposed for ad hoc networks using the bargaining scheme. In [34], the idea was extended to cognitive radios. In [35], mesh networks were investigated, and in [36], multimedia source coding was also considered.

9.8 Contract Using Cooperative Game

Until now, we have discussed how to play the cooperative game and obtain mutual benefits by bargaining. To further analyze the benefits and rewards, we investigate a game coalition that describes how much collective payoff a set of nodes can gain and how to divide the payoff. The associated analysis concepts include core, Shapley function and nucleolus. In the following, we will explain these concepts and explain how to use them in the cognitive radio networks.

Definition 9.13. *A coalition S is defined to be a subset of the total set of player K, $S \in K$. The users in a coalition try to cooperate with each other. The coalition form of a game is given by the pair (K, v), where v is a real value function, called characteristic function. $v(S)$ is the value of the cooperation for coalition S with the following properties:*

1. $v(\emptyset) = 0$
2. *(Superadditivity) if S and T are disjoint coalitions ($S \cap T = \emptyset$), then $v(S) + v(T) \leq v(S \cup T)$.*

The coalition states the benefit obtained via cooperation agreement. But we still need to study how to divide the benefit to the cooperative users. One of the possible properties of an agreement is mutual benefit. The agreement is stable since no coalition shall have the incentive and power to upset the cooperative agreement. The set of such division of v is called the core defined in the following definitions.

Definition 9.14. *A payoff vector $\boldsymbol{x} = (x_1, \ldots, x_K)$ is said to be group rational or efficient if $\sum_{i=1}^{K} x_i = v(K)$. A payoff vector x is said to be individually rational if the user can obtain the benefit no less than acting alone, i.e., $x_i \geq v(\{i\})$, $\forall i$. An imputation is a payoff vector satisfying the above two conditions.*

Definition 9.15. *An imputation **x** is said to be unstable through a coalition S if $v(S) > \sum_{i \in S} x_i$, i.e., the users have incentive for coalition S and upset the proposed **x**. The set C of a stable imputation is called the core, i.e.,*

$$C = \{\boldsymbol{x} : \sum_{i \in K} x_i = v(K) \text{ and } \sum_{i \in S} x_i \geq v(S), \ \forall S \subset K\}. \tag{9.25}$$

Core gives a reasonable set of possible shares. A combination of shares is in a core if there exists no subcoalition in which its members may gain a higher total outcome than the share of concern. If the share is not in a core, some members may be frustrated and may think of leaving the whole group with some other members and form a smaller group.

To illustrate the idea of core, we give the following example. Suppose the game with the following characteristic functions:

$$\begin{aligned} v(\emptyset) = 0, v(\{1\}) = 1, v(\{2\}) = 0, v(\{3\}) = 1, \\ v(\{1,2\}) = 4, v(\{1,3\}) = 3, v(\{2,3\}) = 5, v(\{1,2,3\}) = 8. \end{aligned} \tag{9.26}$$

By using $v(\{2,3\}) = 5$, we can eliminate the payoff vector (such as $(4,3,1)$), since user 2 and user 3 can achieve better payoff by forming coalition themselves. Using the same analysis, the final core of the game is (3,4,1), (3,3,2), (3,2,3), (3,1,4), (2,5,1), (2,4,2), (2,3,3), (2,2,4), (1,5,2), (1,4,3) and (1,3,4).

Core concept defines the stability of an allocation of payoff. However, it does not define how to allocate the utility. Next, we study each individual player's power in the coalition by defining a value called Shapley function.

Definition 9.16. *A Shapley function ϕ is a function that assigns to each possible characteristic function v a real number, i.e.,*

$$\phi(v) = (\phi_1(v), \phi_2(v), \ldots, \phi_K(v)) \tag{9.27}$$

where $\phi_i(v)$ represents the worth or value of player i in the game. The Shapley axioms for $\phi(v)$ is

1. *Efficiency: $\sum_{i \in K} \phi_i(v) = v(K)$.*
2. *Symmetry: If i and j are such that $v(S \cup \{i\}) = v(S \cup \{j\})$ for every coalition S not containing i and j, then $\phi_i(v) = \phi_j(v)$.*
3. *Dummy Axiom: If i is such that $v(S) = v(S \cup \{i\})$ for every coalition S not containing i, then $\phi_i(v) = 0$.*
4. *Additivity: If u and v are characteristic functions, then $\phi(u + v) = \phi(v + u) = \phi(u) + \phi(v)$.*

It can be proved that there exists a unique function ϕ satisfying the Shapley axioms. To calculate the Shapley function, suppose we form the grand coalition by entering the players into this coalition one at a time. As each player enters the coalition, he receives the amount by which his entry increases the value of the coalition he enters. The amount a player receives by this scheme depends on the order in which the players are entered. The Shapley value is just the average payoff to the players if the players are entered in completely random order, i.e.,

$$\phi_i(v) = \sum_{S \subset K, i \in S} \frac{(|S|-1)!(K-|S|)!}{K!}[v(S) - v(S - \{i\})]. \tag{9.28}$$

For the example in (9.26), it can be shown that the Shapley value is $\phi = (14/6, 17/6, 17/6)$.

Another concept for multiple cooperative games is nucleolus. For a fixed characteristic function, an imputation **x** is found such that the worst inequity is minimized, i.e., for each coalition S and its associated dissatisfaction, an optimal imputation is calculated to minimize the maximum dissatisfaction. First we define the concept of excess which measures the dissatisfactions.

Definition 9.17. *The measure of the inequity of an imputation* **x** *for a coalition S is defined as the excess:*

$$e(\boldsymbol{x}, S) = v(S) - \sum_{j \in S} x_j. \tag{9.29}$$

Obviously, any imputation **x** is in the core, if and only if all its excesses are negative or zero.

Among all allocation, kernel is a fair allocation, defined as in the following

Definition 9.18. *A kernel of v is the set of all allocations* **x** *such that*

$$\max_{S \subseteq K-j, i \in S} e(\boldsymbol{x}, S) = \max_{T \subseteq K-i, j \in T} e(\boldsymbol{x}, T). \tag{9.30}$$

If players i and j are in the same coalition, then the highest excess that i can make in a coalition without j is equal to the highest excess that j can make in a coalition without i.

Finally, we define nucleolus as follows.

Definition 9.19. *Nucleolus is the allocation* **x** *which minimizes the maximum excess.*

$$\boldsymbol{x} = \arg\min_{\boldsymbol{x}}(\max e(\boldsymbol{x}, S),\ \forall S). \tag{9.31}$$

The nucleolus has the following property: the nucleolus of a game in coalitional form exists and is unique. The nucleolus is group rational, individually rational and satisfies the symmetry axiom and the dummy axiom. If the core is not empty, the nucleolus is in the core and kernel. In other words, the nucleolus is the best allocation with the min–max criteria.

To utilize the cooperative game in dynamic spectrum allocation for cognitive networks, the cognitive users sign a contract for spectrum usage before accessing the spectrum. This contract ensures that the benefits of cooperation are greater than those of the individual actions. The core concepts can test whether or not the cooperation is stable. Then if the average fairness is considered, Shapley values can allocate different cognitive users their share of cooperation benefits. On the other hand, if the max–min fairness is considered, the concepts of excess, kernel and nucleolus define the allocation. Overall, the cognitive users seek the contracts for resource usage that can benefit all.

9.9 Centralized Optimization

In this section, we discuss how to formulate the centralized optimization for resource allocation of cognitive radios. Specifically, we study what the resources are, what the parameters are, what the practical constraints are and what the optimized performances across the different layers are. In addition, we address how to perform resource allocation in multi-user scenarios. The tradeoffs between the different optimization goals and different users' interests are also investigated. This centralized optimization can serve as the performance upper bound for the other approaches and can also provide insights for the design of other schemes.

Many resource allocation problems for cognitive radios can be formulated as constrained optimization problems, which can be optimized from the network point of view or from the individual point of view. The general formulation can be written as:

$$\min_{\mathbf{x}\in\Omega} f(\mathbf{x}) \tag{9.32}$$
$$\text{s.t.} \begin{cases} g_i(\mathbf{x}) \leq 0, & \text{for } i = 1, \ldots, m \\ h_j(\mathbf{x}) = 0, & \text{for } j = 1, \ldots, l \end{cases}$$

where $\mathbf{x}$ is the parameter vector for optimizing the resource allocation, Ω is the feasible range for the parameter vector and $f(\mathbf{x})$ is the optimization goal matrix, objective goal or utility function that represents the performance or cost. Here, $g_i(\mathbf{x})$ and $h_j(\mathbf{x})$ are the inequality and equality constraints, respectively, for the parameter vector. The optimization process finds the solution $\bar{\mathbf{x}}$ that satisfies all the inequality and equality constraints. For the optimal solution, $f(\bar{\mathbf{x}}) \leq f(\mathbf{x}), \forall \mathbf{x} \in \Omega$.

If the optimization goal, the inequality constraints, and the equality constraints are all linear functions of the parameter vector $\mathbf{x}$, then the problem in (9.32) is called a *linear program*. One important characteristic of a linear program problem is that there is a global optimal point that is very easy to obtain by linear programming. But on the other hand, one major drawback of linear program is that most of the practical problems in wireless networking and resource allocation are non-linear. Therefore, it is hard to model these practical problems as linear programs. If either the optimization goal or the constraint functions are non-linear, the problem in (9.32) is a *non-linear program*. In general, there are multiple local optima in a non-linear program, and to find the global optimum is not an easy task. Furthermore, if the feasible set Ω contains some integer sets, the problem in (9.32) is an *integer program*. Most integer programs are NP-hard problems which cannot be solved by polynomial time.

One special kind of non-linear program is a convex optimization problem in which the feasible set Ω is a convex set, and the optimization goal and the constraints are convex/concave/linear functions. A convex set is defined as follows.

Definition 9.20. *A set Ω is a convex set if for any $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \Omega$ and any θ with $0 \leq \theta \leq 1$, we have $\theta\boldsymbol{x}_1 + (1-\theta)\boldsymbol{x}_2 \in \Omega$.*

A convex function f is defined as follows.

Definition 9.21. *A function f is a convex function in $\boldsymbol{x}$, if the feasible range Ω of parameter vector $\boldsymbol{x}$ is a convex set, and if for all $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \Omega$ and $0 \leq \theta \leq 1$,*

$f(\theta \boldsymbol{x}_1 + (1-\theta)\boldsymbol{x}_2) \leq \theta f(\boldsymbol{x}_1) + (1-\theta) f(\boldsymbol{x}_2)$.

A function f is strictly convex if the strict inequality holds whenever $\boldsymbol{x}_1 \neq \boldsymbol{x}_2$ and $0 < \theta < 1$. A function f is called concave if $-f$ is convex.

If function f is differentiable, and if either the following two conditions hold, then f is a convex function.

First order condition: $f(\mathbf{x}_2) \geq f(\mathbf{x}_1) + \nabla f(\mathbf{x}_1)^T(\mathbf{x}_2 - \mathbf{x}_1)$.

Second order condition: $\nabla^2 f(\mathbf{x}) \succeq 0$.

One important application of the convex function is Jensen's inequality. Suppose function f is convex and the parameter $\mathbf{x}$ has any arbitrary random distribution over Ω then the following equality holds

$f(E(\mathbf{x})) \leq E(f(\mathbf{x}))$

where E denotes expectation.

The advantages of convex optimization for wireless-networking-and-resource-allocation problems are shown as follows:

- There are a variety of applications such as automatic control systems, estimation and signal processing, communications and networks, electronic circuit design, data analysis and modeling and statistics.
- Computation time is usually quadrature. Problems can then be solved, very reliably and efficiently, using interior-point methods or other special methods for convex optimization.
- Solution methods are reliable enough to be embedded in a computer-aided design or analysis tool, or even a real-time reactive or automatic control system.
- There are also theoretical or conceptual advantages of formulating a problem as a convex optimization problem.

The challenges of the convex optimization are to recognize and model the problem as a convex optimization. Moreover, there are many tricks for transforming problems into convex forms.

We have discussed the basics for constrained optimization problems. Next we will see how the problem can be formulated. In resource allocation for cognitive networks, the parameters, functions and constraints in (9.32) can have the following physical meaning:

- Parameters
  1. *Physical layer*: transmitted power, modulation level, channel coding rate, channel/code selection and others.
  2. *MAC layer*: transmission time/frequency, service rate, priorities for transmission and others.
  3. *Network layer*: route selection, routing cost and others.
  4. *Application layer*: source-coding rate, buffer priority, packet arrival rate and others.
- Optimization Goals

1. *Physical layer*: minimal overall power, maximal throughput, maximal rate per joule, minimal bit error rate, and others.
2. *MAC layer*: maximal overall throughput, minimal buffer overflow probability, minimal delay and others.
3. *Network layer*: minimal cost, maximal profit and others.
4. *Application layer*: minimal distortion, minimal delay and others.

- Constraints
  1. *Primary user*: channel occupancy, interference level and others.
  2. *Physical layer*: maximal mobile transmitted power, available modulation constellation, available channel coding rate, limited energy and others.
  3. *MAC layer*: contentions, limited time/frequency slot, limited information about other mobiles and others.
  4. *Network layer*: maximal hops, security concerns and others.
  5. *Application layer*: the base layer transmission, limited source rate, strict delay requirement, security and others.

After formulating the constrained optimization problem for resource allocation over cognitive networks, we need to find solutions. In general for centralized optimization, we classify the different approaches as the following categories.

- *Closed-form solution*: One of the most important methods used to find a closed form solution for constrained optimization is the Lagrangian method, which has the following steps
  1. Rewrite (9.32) as a Lagrangian multiplier function J as
  $$J = f(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i g_i(\mathbf{x}) + \sum_{j=1}^{l} \mu_j h_j(\mathbf{x}) \tag{9.33}$$
  where λ_i and μ_j are Lagrangian multipliers.
  2. Differentiate J over $\mathbf{x}$ and set to zero as
  $$\frac{\partial J}{\partial \mathbf{x}} = 0. \tag{9.34}$$
  3. From (9.34), solve λ_i and μ_j.
  4. Replace λ_i and μ_j in the constraints to get optimal $\mathbf{x}$.

  Notice that the difficulty in the Lagrangian method is Step (3) and Step (4), where the closed form solution is obtained for the Lagrangian multipliers. Some approximations and mathematical tricks are necessary to obtain the closed form solutions.
- *Mathematical programming*: If the optimization problem is used to find the best objective function within a constrained feasible region, such a formulation is sometimes called a mathematical program. Many real-world and theoretical problems can be modeled in this general framework. There are the four major subfields of the mathematical programming:
  1. Linear programming studies the case in which the objective function is linear and the feasible set is specified using only linear equalities and inequalities.

2. Convex programming studies the case where the constraints and the optimization goals are all convex or linear.
3. Non-linear programming studies the general case in which the objective function or the constraints or both contain non-linear parts.
4. Dynamic programming studies the case in which the optimization strategy is based on splitting the problem into smaller subproblems, or considers the optimization problems over time.

- *Integer/combinatorial optimization*: The discrete optimization is the problem in which the decision variables assume discrete values from a specified set. The combinatorial optimization problems, on the other hand, are problems of choosing the best combination out of all possible combinations. Most combinatorial problems can be formulated as integer programs. In cognitive radio resource allocation, many variables have only integer values such as the modulation rate, and other variables such as the channel allocation have a combinatorial nature. Integer optimization is the process of finding one or more best (optimal) solutions in a well-defined discrete problem space. The major difficulty with these problems is that we do not have any optimality conditions to check if a given (feasible) solution is optimal or not. There are several possible solutions such as relaxation and decomposition, enumeration, cutting planes and the knapsack problem.

Overall, the centralized scheme has the best performance but needs considerable signaling and overheard. The centralized scheme can fit the network scenarios where the topology is simple, or can be served as a performance upper bound to compare with other more practical schemes.

9.10 Degrees of Cooperation

In this section, we conclude this chapter by discussing the degrees of cooperation for the different approaches. As we have mentioned previously, the non-cooperation among cognitive radios can significantly reduce the network performances and in turn the users' own benefits. The cooperation can bring mutual benefits to cognitive radio users. However, these benefits do not come for free. Some network infrastructure is needed to build up these mutual benefits, which cause some design issues. In the previous sections, we have already discussed the different approaches. Next, we concentrate on the design issues such as signaling and complexity. Then the pros and cons are investigated. We further study the best network scenarios under which a certain approach fits best. By understanding the above issues, finally, we compare the different approaches.

In cognitive networks, in order to obtain the information such as channel conditions, signaling is performed so that resource allocation can be conducted in an optimal way. However, signaling incur considerable communication overhead. Most of the current wireless networks have more than 20% of overhead. Reducing the overhead can greatly enhance the spectrum utilization, increase the number of users and improve the network performance. One of the possible ways to reduce overhead is to conduct resource optimization using only local information. This is very important especially if the system topology is distributed as in cognitive radio networks.

Since the cognitive radios are usually equipped with simple transceivers, the complexity issue has to be considered. There are two concerns for optimization complexity. First one is how complex the optimization algorithm is, and the second one is where the optimization is performed. A large number of optimization problems especially for those with integer nature are NP-hard. To solve the problem, some suboptimal simple solutions should be developed. Currently, the common hardware and software can solve the problems with the complexity up to $\mathrm{O}N^2$ where N is the bottleneck parameter. The complexity for the distributed cognitive radio users should be even lower.

The next important problem for design of wireless cognitive networks is mobility. Due to the topology changes and channel variation, the optimization needs to be performed in a timely fashion. This requirement casts a significant challenge for the iterative solutions and demand for the information without delay. The convergence speed for the iterative algorithms should be at least as fast as the variation caused by mobility. For example, in a 3G UMTS system, the closed-loop power control signal is performed 1500 times per second. This fast update for the iteration can improve the convergence speed, but on the other hand cause additional overhead for the signaling. For a non-iterative algorithm, the information must be accurate without delay. Otherwise, the optimization results will become obsolete and generate inferior performances.

Different approaches have their own pros and cons, and there is no one "elixir" that can handle all design problems for all types of networks. In addition, there are some other design issues that need to be paid attention to. So we need to understand the strength and weakness for different schemes, so that we can select the one that fits the network scenario best. In the sequel, we discuss and compare all types of schemes discussed in this chapter. Table 9.3 summarizes some of the discussions.

- *Non-cooperative competition*: The cognitive radio users have their own autonomy and they access the spectrum in a fully distributed way. The cognitive users utilize only local information for resource allocation, and no signaling or overheard is necessary. The complexity of the non-cooperative competition algorithms is usually low, due to the commonly used convex (or concave) utility function. This type of approaches can fully adapt to the user mobility, since the users can simply change their strategies for better payoff if the situations change. All the above factors are the advantages of non-cooperative competition. However, the significant problem for such approaches is the possible low performance, due to the severe non-cooperation. Even though the problem can be improved using techniques such as pricing, the solutions do not come for free. For example, calculating the optimal price is a difficult problem and might need considerable signaling, which counteracts the advantages for such approaches. So the best network scenarios for non-cooperative competition are those where the Nash equilibria have similar performances to those of the optima. Specifically for the interference avoidance, if the clusters are located sufficiently far away, the

Table 9.3. Degrees of cooperation.

Types	Cooperation	Signaling	Pros	Cons
Non-cooperative competition	Nash equilibria	None	No overhead, simple	Less efficient
Correlated equilibrium and learning	Outside convex hull of Nash equilibria	None	No overhead, better performance	Convergence slow, little mobility
Referee mediation	Only good Nash equilibria	Some for referee	Better performance	Bad stability, low mobility
Repeated interaction	Any feasible solution better than Nash equilibria	Perfectly observed public information	Local information, better performance	No mobility, need mutual dependency, false war
Spectrum auction	Nash equilibria	Some for auctioneer	Simple, fair	No mobility, signaling
Bargaining	Fair Pareto optimum	Only to partners	Simple, mobility	Monopoly
Cooperative game	Fair mutual benefits	Global information before participation	Stable, fair, autonomous	Signaling before contract, no mobility
Centralized optimization	Global optimum	Global information	Optimum, mobility	Overhead, estimator errors

non-cooperative competition has good performance due to less co-channel interference.

- *Learning for better equilibria*: Nash equilibria might not be the best equilibria for distributed cognitive radio users. Learning scheme can achieve the better equilibria using only the past history and without requiring more signaling and overhead. The complexity of learning algorithms can be relatively high. Moreover, there is a tradeoff between the convergence speed and complexity. To achieve the fast convergence speed, the complexity of the learning algorithms can be high. Some simple learning algorithms have been proved to converge to the optimal solution with sufficiently long learning time. However, the long learning time causes a problem for mobility. If the users move frequently, before the learning algorithms converge, the situations such as network topologies and channel conditions may change. This is similar to slope overload distortion in ADPCM or delta modulation. Moreover, if the non-cooperative competition is too severe, the learning algorithms might converge too slowly, fluctuate or become very sensitive to randomness. So the learning schemes fit the situation in which the non-cooperative competition is not so severe; there is an achievable gap between Nash equilibria and the optimal solutions; and the network mobility is sufficiently low.

- *Referee mediation*: To overcome the challenges for the learning schemes, a virtual referee can improve the outcome of non-cooperative competition by intervening in the game rules. The virtual referee needs to collect the information so as to improve the equilibria. However, this information exchange burden is not severe since it is only necessary when the networks are balanced to undesired equilibria. The complexity for this type of approaches is relatively low. Mobility is not an issue, if the network changes can be handled mostly by the non-cooperative competitions and the frequency for virtual referee's mediation is not too high. However, the referee mediation approaches require the assumptions that all the cognitive radios are able to follow the instructions to change the game rules. So the cognitive radio users are not fully autonomous. Moreover, too much intervention by the virtual referee can cause a network stability problem. This type of approach fits the similar scenarios like the learning schemes except that the cognitive users can have a certain extent of mobility.
- *Threat and punishment from repeated interactions*: If the cognitive radio users belong to different authorities, they will not listen to the virtual referee. Under this condition, threat and punishment from repeated interactions can be utilized to enforce user cooperation. There is public information that needs to be received accurately by all cognitive users who use this information to determine if any other user deviates from cooperation. Because of this reason, if this public information is not accurate, some "false war" can happen among distributed cognitive users. To a certain extent, the network can deteriorate to total non-cooperation. The complexity of such an approach is not high, since only detection, cooperation and non-cooperation need to be performed. This approach can hardly handle the mobility, since the deviating users can move from cluster to cluster to escape future punishment, or equivalently saying that mobile users might not care too much about future punishment so that they would rather behave non-cooperatively now. In addition, if some cognitive users have less dependency on other users, the other users can arbitrarily play non-cooperatively with these users without worrying about revenge. So this type of approach fits the network scenarios where the cognitive users have less mobility, have mutual dependency, and can access public information accurately.
- *Spectrum auction*: Similar to an auction in real life, a spectrum auction requires an auctioneer who can handle the bidding and resulting resource allocations. The information exchange requires the signaling of bidding and allocation results, which can be relatively trivial. The complexity of auction algorithms can be very high, for example the VCG auction. But the computation burden is for the auctioneer only. The spectrum auction cannot handle mobility. If the mobile users move, a new auction needs to be implemented. Similar to the referee case, the cognitive users are required to follow the instructions for resource usage from the auctioneer. The spectrum auction fits the network scenario without mobility, and there should be some semi-centralized nodes, such as cluster heads, that can serve as auctioneers.
- *Mutual benefits via bargaining*: The bargaining approach can provide the local mutual benefits to the adjacent cognitive radios. The cognitive users can exchange

the information locally to bargain on spectrum usage. The overhead is limited to local users only and the complexity of algorithms is usually low. This type of approach can handle the mobility, since the bargaining can take place whenever new mutual benefits appear. Moveover, this bargaining process fits the situations with integer and combinatorial optimization well. However, if one user occupies most of the spectrum, it is less efficient for the other cognitive users to negotiate with this monopolist.

- *Contract using cooperative game*: This approach is similar to a spectrum auction, except that there is no need for an auctioneer. Instead, all participant users "put their cards on the table" and figure out the best strategies for coalitions. The resulting mutual benefits are divided to cognitive radios according to different fairness criteria. A lot of information signaling is necessary before the contract is agreed by all users, but no signaling is needed after that. The complexity of coalition formation can be high. The mobility is required to be limited, otherwise the contract becomes obsolete too quickly. The cooperative game fits better if the users are located densely, so that the information exchange can be easy.
- *Centralized scheme*: The cognitive radios are the slave type, which means the users fully cooperate and follow the instructions from the centralized node. The optimization requires the accurate channel information without delay. For the scenario of multiple cognitive radio users talking to one common destination such as a base station, centralized control can be utilized since the channel information is constantly collected by the destination to maintain the links. On the other hand, for the network scenarios like the ad hoc case or the multiple cluster case, it is very difficult for the channel information to be exchanged over different destinations. In this situation, the centralized control can hardly be implemented but can serve as a performance upper bound for the other distributed schemes. The complexity of the centralized schemes are usually high, due to the non-linear, non-convex and probably integer or dynamic nature of the optimization problem. However, the optimization is usually performed in the destination where the computation ability is relatively high. For mobility, if the channel information is prompt, the centralized scheme is robust with the channel variation. However, if the channel information needs to be feeded back or sent via signaling, the delay can significantly degrade the performance of the centralized scheme.

References

1. S. Haykin, "Cognitive radio: Brain-empowered wireless communications," *IEEE J. Select. Areas Commun.*, vol. 23, no. 2, pp. 201–220, Feb. 2005.
2. I. F. Akyildiz, W. Y. Lee, M. C. Vuran, and S. Mohanty, "Next generation/dynamic spectrum access/cognitive radio wireless networks: A survey," *Comput. Netw. Int. J. Comput. Telecommun. Netw.*, vol. 50, no. 13, pp. 2127–2159, Sept. 2006.
3. Z. Han, Z. Ji, and K. J. Ray Liu, "Fair multiuser channel allocation for OFDMA networks using Nash bargaining solutions and coalitions," *IEEE Trans. Commun.*, vol. 53, no. 8, pp. 1366–1376, 2005.

4. Z. Han, C. Pandana, and K. J. Ray Liu, "Distributive opportunistic spectrum access for cognitive radio using correlated equilibrium and no-regret learning," in *Proc. of IEEE Wireless Communications and Networking Conference* (Hong Kong, China), Mar. 2007.
5. Z. Han, Z. Ji, and K. J. Ray Liu, "A referee-based distributed scheme of resource competition game in multi-cell multi-user OFDMA networks," *IEEE J. Select. Areas Commun.*, Special Issue on Non-cooperative Behavior in Networking, 2nd Quarter, 2007.
6. Z. Han, C. Pandana, and K. J. R. Liu, "A self-learning repeated game framework for optimizating packet forwarding networks," in *Proc. of IEEE Wireless Communications and Networking Conference* (New Orleans, LA), pp. 2131–2136, Mar. 2005.
7. Z. Han, Z. Ji, and K. J. Ray Liu, "Dynamic distributed rate control for wireless networks by optimal cartel maintenance strategy," in *Proc. of IEEE Global Telecommunications Conference*, vol. 6 (Dallas, TX), pp. 3454–3458, Nov. 2004.
8. X. Liu and W. Wang, "On the characteristics of spectrum-agile communication networks," in *Proc. IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks, 2005 (DySPAN 2005)* (Baltimore, MD), pp. 214–223, Nov. 2005.
9. Z. Han and K. J. Ray Liu, "Non-cooperative power control game and throughput game over wireless networks," *IEEE Trans. Commun.*, vol. 53, no. 10, pp. 1625–1629 Oct. 2005.
10. A. B. MacKenzie and L. A. DaSilva, *Game theory for wireless engineers.* Morgan and Claypool Publishers, 2006.
11. D. Niyato and E. Hossain, "A game-theoretic approach to competitive spectrum sharing in cognitive radio networks," in *Proc. of IEEE Wireless Communication and Networking Conference* (Hong Kong), March 2007.
12. D. Niyato and E. Hossain, "Hierarchical spectrum sharing in cognitive radio: A microeconomic approach," in *Proc. of IEEE Wireless Communication and Networking Conference* (Hong Kong), March 2007.
13. B. Wang, Z. Han, and K. J. Ray Liu, "Distributed relay selection and power control for multiuser cooperative communication networks using buyer/seller game," in *Proc. of Annual IEEE Conference on Computer Communications, INFOCOM* (Anchorage, Alaska), May 2007.
14. D. P. Bertsekas and R. G. Gallager, *Data networks*, 2nd ed. Prentice-Hall, 1992.
15. Y. Yang and T. S. P. Yum, "Delay distributions of slotted ALOHA and CSMA," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1846–1857, Nov. 2003.
16. R. J. Aumann, "Subjectivity and correlation in randomized strategy," *J. Math. Econ.*, vol. 1, no. 1, pp. 67–96, 1974.
17. R. J. Aumann, "Correlated equilibrium as an expression of Bayesian rationality," *Econometrica*, vol. 55, no. 1, pp. 1–18, Jan. 1987.
18. S. Hart and A. Mas-Colell, "A simple adaptive procedure leading to correlated equilibrium," *Econometrica*, vol. 68, no. 5, pp. 1127–1150, Sept. 2000.
19. R. B. Myerson, *Game theory: Analysis of conflict,* 5th ed. Harvard University Press 2002.
20. Q. Zhao, L. Tong, and A. Swami, "Decentralized cognitive MAC for dynamic spectrum access," in *Proc. of IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks, 2005 (DySPAN 2005)* (Baltimore, MD), pp. 224–232, Nov. 2005.
21. Z. Han, Z. Ji, and K. J. Ray Liu, "Power minimization for multi-cell OFDM networks using distributed non-cooperative game approach," in *Proc. of IEEE Global Telecommunications Conference* (Dallas, TX), Nov. 2004.
22. S. T. Chung, S. J. Kim, and J. M. Cioffi, "A game-theoretic approach to power allocation in frequency-selective Gaussian interference channels," in *Proc. of IEEE International Symposium on Information Theory* (Pacifico Yokohama, Kanagawa, Japan), June 2003.

23. S. T. Chung and A. J. Goldsmith, "Degrees of freedom in adaptive modulation: A unified view," *IEEE Trans. Commun.*, vol. 49, no. 9, pp. 1561–1571, Sept. 2001.
24. D. P. Bertsekas, *Nonlinear programming*, 2nd ed. Athena Scienific, 1999.
25. R. Yates, "A framework for uplink power control in cellular radio systems," *IEEE J. Select. Areas Commun.*, vol. 13, no. 7, pp. 1341–1348, Sept. 1995.
26. G. J. Foschini and Z. Miljanic, "A simple distributed autonomous power control algorithms and its convergence," *IEEE Trans. Veh. Technol.*, vol. 40, no. 4, pp. 641–646, Nov. 1993.
27. L. Anderegg and S. Eidenbenz, "Ad hoc-VCG: A truthful and cost-efficient routing protocol for mobile ad hoc networks with selfish agents," in *Proc. of ACM Ninth Annual International Conference on Mobile Computing and Networking (MobiCom)* (San Diego, CA), Sept. 2003.
28. V. Krishna, *Auction theory*. Academic Press, 2002.
29. J. Huang, R. Berry, and M. L. Honig, "Auction-based spectrum sharing," *ACM/Springer Mobile Netw. Appl. J. (MONET)*, vol. 11, no. 3, pp. 405–418, June 2006.
30. H. Yaiche, R. R. Mazumdar, and C. Rosenberg, "A game theoretic framework for bandwidth allocation and pricing in broadband networks," *IEEE/ACM Trans. Netw.*, vol. 8, no. 5, pp. 667–678, Oct. 2000.
31. D. Grosu, A. T. Chronopoulos, M. Y. Leung, "Load balancing in distributed systems: An approach using cooperative games," in *Proc. of IPDPS 2002*, pp. 52–61, 2002.
32. H. W. Kuhn, "The Hungarian method for the assignment problem," *Naval Res. Log.*, Quarterly 2, 1955.
33. J. E. Suris, L. DaSilva, Z. Han, and A. MacKenzie, "Cooperative game theory approach for distributed spectrum sharing," in *Proc. of IEEE International Conference on Communications* (Glasgow, Scotland), June 2007.
34. H. Zheng and L. Cao, "Device-centric spectrum management," in *Proc. of IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks, 2005 (DySPAN 2005)* (Baltimore, MD), pp. 56–65, Nov. 2005.
35. K.-D. Lee and V. C. M. Leung, "Fair allocation of subcarrier and power in an OFDMA wireless mesh network," *IEEE J. Select. Areas Commun.*, vol. 24, no. 11, pp. 2051–2060, Nov. 2006.
36. F. Fu, A. Fattahi, and M. van der Schaar, "Game-theoretic paradigm for resource management in spectrum agile wireless networks," in *Proc. of IEEE International Conference on Multimedia and Expo* (Toronto, Canada), July 2006.

Additional Reading

1. X. Liu and S. Shankar, "Sensing-based opportunistic channel access," *ACM MONET*, vol. 11, no. 4, pp. 577–591, Aug. 2006.
2. J. Huang, R. A. Berry, and M. L. Honig, "Spectrum sharing with distributed interference compensation," in *Proc. of IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks, 2005 (DySPAN 2005)* (Baltimore, MD), pp. 88–93, Nov. 2005.
3. V. Srivastava, J. Neel, A. MacKenzie, R. Menon, L. A. DaSilva, J. Hicks, J. H. Reed, and R. Gilles, "Using game theory to analyze wireless ad hoc networks," *IEEE Commun. Surv. Tutor.*, vol. 7, no. 4, pp. 46–56, 4th quarter 2005.
4. R. Etkin, A. Parekh, and D. Tse, "Spectrum sharing for unlicenced bands," in *Proc. of IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks, 2005 (DySPAN 2005)* (Baltimore, MD), pp. 251–258, Nov. 2005.

5. N. Nie and C. Comaniciu, “Adaptive channel allocation spectrum etiquette for cognitive radio networks,” *ACM MONET (Mobile Networks and Applications)*, Special Issue on Reconfigurable Radio Technologies in Support of Ubiquitous Seamless Computing, 2006.
6. J. Neel, J. Reed, and R. Gilles, “Game models for cognitive radio algorithm analysis,” in *Proc. of SDR Forum Technical Conference* (Phoenix, Arizona), pp. 15–18, Nov. 2004.
7. C. Saraydar, N. Mandayam, and D. Goodman, “Efficient power control via pricing in wireless data networks,” *IEEE Trans. Commun.*, vol. 50, no. 2, pp. 291–303, Feb. 2002.
8. C. Pandana and K. J. Ray Liu, “Near-optimal reinforcement learning framework for energy-aware sensor communications,” *IEEE J. Select. Areas Commun.*, vol. 23, pp. 259–268, Apr. 2005.