Self-adaptive Spectrum Management in Partially Observable Environments

Contents

5.1 Introduction

Due to the recent and dramatic development of the wireless communication industry, the demand for wireless spectrum has been growing rapidly. Thus, the spectrum scarcity is becoming a challenge for several recent studies. Both academic and industry are recognizing that traditional fixed spectrum allocation is very inefficient, such that most of the time the bandwidth that was allocated is not optimally used and the corresponding channel is idle, which forms spectrum holes [8]. CR [1], which is a new paradigm for designing wireless communication systems, appeared in order to enhance the utilization of the radio frequency spectrum. It was considered as the key technology that enables SUs to access the licensed spectrum. Typically, SUs access opportunistically the spectrum when it is not used by PUs. The presence of several SUs in the same portion of spectrum band enhanced the need to efficiently share the spectrum. Indeed, the utilization of the radio spectrum is reduced due to collisions among SUs under decentralized channel selection schemes. In order to optimize the utilization of the scarce spectrum resources, DSA become a promising approach to increase the efficiency of spectrum usage and to solve the scarcity problem.

Surprisingly, the impact of the energy constraint, due to the limited mobile users' battery, and the capacity of CR to support additional QoS were somehow ignored and not sufficiently studied in the literature. In many wireless systems, it is very important to provide reliable communications while sustaining a certain level of QoS. However, challenges in providing the QoS assurances increase due to the fact that SUs operate under constraints on the licensed channels' occupancy, and competition between each other.

We investigate an important problem for determining the OSA mechanism, and we propose a general model that allows us to study the impact of energy consumption and expected delay on the OSA policy. The main novelty of our approach is to consider a POSG framework. The theory of POMDP was widely and successfully used, like in [80], [53] and [90], to model and build OSA mechanisms in CR networks. However, those works do not consider the competition between SUs. Very few works proposed to model such competition(see [94] and [95] for example). Moreover, those works do not have significant results. In fact, using a DP approach to solve a POMDP is possible by transforming it into a completely observable MDP over belief states [95]. It is very difficult to generalize this technique for POSG as the SUs may have different beliefs. This problem was alleviated by introducing the notion of generalized belief state in [41], however the optimal algorithm becomes intractable beyond a small horizon. In our work, we focus on the existence of an SNE between SUs. The SNE is solved using a Linear Program (LP). Second, we identify paradoxical behaviors of SUs. One of the observed paradoxes here is a kind of Braess paradox, a well-studied paradox in routing context [96]. Our paradox indicates that decreasing the spectrum occupancy may lead degradation of the performance in term of the average throughput for SUs. This observation is due to the increase of the aggressiveness of SUs when the spectrum availability increases. We look further for a network control mechanism in order to optimize the average throughput of SUs at the SNE. For this end, we consider a Stackelberg game formulation [97].Note that Stackelberg game formulations was already proposed in the CR literature (see for example [39], [40] and [98]), as the natural hierarchy between PUs and SUs is very similar to the hierarchy between leaders and followers. Nevertheless, it was not used in order to enhance the network usage. In the second part of this chapter, we propose a control mechanism, for the network manager using a Stackelberg game formulation, such that the total average throughput of the SUs is maximized in this partially observable environment.

Many works focused on the study of optimal OSA policies in CR networks. In [80], the authors studied decentralized MAC protocols such that SUs search for spectrum opportunities without a central controller. They considered a POMDP and proposed an analytical framework based on this mathematical tool. However, the authors consider neither energy consumption nor any QoS constraint in their OSA policy. The problem of maximizing the throughput of traffic subject to some constraints on its delay received the extensive attention of pioneering work [99]. Authors of [100] described linear programming solvers for MDP, which are able to handle finite and infinite horizon problems. Moreover, authors of [101] considered a problem similar to ours but in a queueing context. They used the linear programming in order to solve an MDP and to study the equilibria for N players scenario in a stochastic game context. Few works focused on how SUs should operate in order to satisfy some QoS requirements and energy constraints. Authors of [53] incorporated the energy constraint in the design of the optimal OSA policy, in a single user context, and formulated their problem as a POMDP. The major difference between this work and ours is that the authors do not considered the competition between SUs. In [102], the authors presented a queueing analysis of a CR with multiple SUs. They proposed an adaptive algorithm to find the optimal contention probability that minimizes the expected delay. Authors of [103], proposed an energy-efficient non-cooperative strategy for resource allocation in CR networks based on a game theoretical approach. In summary, the main contributions of the chapter are as follows:

- We model a non-cooperative sensing and access game as a POSG. We prove the existence and uniqueness of an SNE for this OSA game.
- In the non-saturated regime, we exhibit an optimal sensing policy where SUs may sense licensed channels, even if they do not have any packets to transmit. Indeed, by sensing the licensed channels, a SU gets information on the RF environment.
- We highlight an interesting paradox, which says that increasing the spectrum occupancy may increase the SUs' average throughput. Indeed, SUs become less aggressive, which induces a better utilization of the spectrum holes (less collisions).
- Finally, we propose a control mechanism for the network manager in order to increase the average total throughput of the network at the SNE. For this purpose, we formulate the hierarchical framework as a Stackelberg game, where the network manager acts as the leader and SUs act as followers.

The remainder of the chapter is organized as follows. In Section 5.2, we introduce our system model. The utility function and the NE analysis are presented in Section 5.3. We propose a Stackelberg-based mechanism for the network manager in order to optimize the licensed channels' utilization in Section 5.4. We present some simulation results in order to discuss the performance of the proposed model in Section 5.5, and we conclude the chapter in Section 5.6.

5.2 The model

We consider M time-varying channels licensed for PUs and N SUs accessing opportunistically the available channels. The occupancy of each channel $k \in \{1, \dots, M\}$ is modeled by a time-homogeneous discrete Markov process denoted s_k, where the state $s_k = 0$ (resp. $s_k = 1$) means that the channel is idle (resp. busy). The licensed channels' transition rates are illustrated in Figure 5.1, where β_k represents the probability that the licensed channel k becomes idle, such that it was occupied in the previous time slot, and α_k represents the probability that the licensed channel k becomes idle such that it was idle, in the previous time slot.

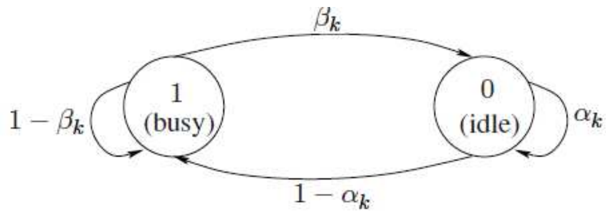

FIGURE 5.1: The discrete time Markov chain describing channel k occupation state.

The global system state, at each time slot t, is composed of the states of the M channels and is denoted by the vector $\mathbf{s}(t) = (s_1(t), ..., s_M(t))$. This global state is also called the Spectrum Occupancy State (SOS). The global state space is denoted by $\mathcal{S} = \{0, 1\}^M$.

We consider a slotted system, where SU access opportunistically the licensed channels when they are not used by PUs. Moreover, we consider a non-saturated regime such that the arrival of packets from upper layer to the transmission layer follows a Bernoulli process with parameter q_a. As long as the SU has a packet to transmit, a new packet is blocked and lost. The packet arrival processes for SUs are supposed to be independent and identically distributed. We further assume that a SU transmits, at most, one packet per time slot. Moreover, we consider an exclusive access to the licensed channels. In fact, when at least two SUs decide to transmit over the same channel, there is a collision and packets are lost (see Figure 5.2). This assumption is usual in CR networks problems related to the MAC layer (see [90] and [104]).

At each time slot t, we define the packet delay $l_i(t)$ for the SU i as the number of elapsed time slot from the arrival of the packet into the transmission buffer until the time slot

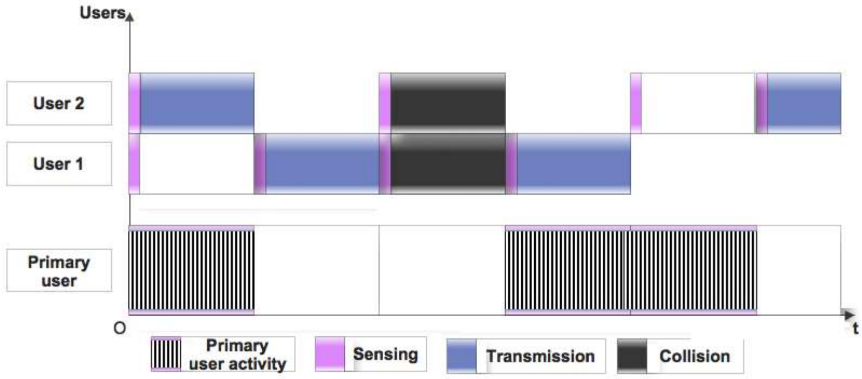

FIGURE 5.2: SUs transmissions

t. Therefore, $l_i(t) = 0$ means that the SU has no packet to transmit at the time slot t. At the beginning of each time slot, the SU i has a perfect knowledge about the current packet delay $l_i(t)$, but ignores the SOS that can not be directly observed due to the partial spectrum sensing. Then, SUs have a partial observation of the global system state. Specifically, we study our problem using a POSG formulation.

A POSG is defined as a tuple $(\mathcal{N}, \mathcal{S}, b^0, \{\mathcal{A}_i\}, \{\mathcal{O}_i\}, \mathcal{P}, \{\mathcal{R}_i\})$, described as follows:

- $\mathcal{N}$ a finite set of SUs indexed $\{1, \ldots, N\}$,
- $\mathcal{S}$ a finite set of states, $|\mathcal{S}| = M$
- b^0 the initial state distribution,
- $\mathcal{A}_i$ the finite set of actions for SU i (we define by $\mathcal{A} = \mathcal{A}_1 \times \ldots \times \mathcal{A}_N$ the joint action set),
- $\mathcal{O}_i$ the finite set of observations for SU i (we define by $\mathcal{O} = \mathcal{O}_1 \times \ldots \times \mathcal{O}_N$ the joint observation set),
- $\mathcal{P}$ a set of state transition and observation probabilities, i.e. $\mathcal{P}(s', o|s, a)$ is the probability that taking action a in state s results in observing o and a transition to state s',
- $\mathcal{R}_i : \mathcal{S} \times \mathcal{A} \rightarrow I\!R$ the reward function for SU i.

System state: We denote the state of the users by $\mathbf{x}(t) = (x_1(t), \ldots, x_N(t))$, where $x_i(t) = (\lambda^i(t), l_i(t))$ represents the state of SU i, and $\mathbf{x_{-i}}(t)$ denotes the state of SUs other than i. Since the M channels are independent, it was proved in [80] that we can consider the following simpler belief vector:

$$\lambda^i(t) = (\lambda^i_1(t), .., \lambda^i_M(t)),$$

where $\lambda_k^i(t)$ is the conditional probability for the SU i that the channel k is available at the time slot t. The state space of SU i is referred to as $\mathcal{X}_i$, and $\mathcal{X} = \cup_i \mathcal{X}_i$ represents the set of all possible joint state of SUs.

Belief: Each SU senses at most one licensed channel in order to get information about the SOS. We denote by $\theta(t) = (\theta_1(t), \cdots, \theta_N(t))$ the set of observations of all the SUs, where $\theta_i(t) = 0$ means that the SU i has sensed the licensed channel as idle. If $\theta_i(t) = 1$, then the licensed channel was sensed as occupied. The observation space is denoted by $\mathcal{O} = \{0, 1\}$. Each SU i updates its belief vector $\lambda^i(t)$ based on its observation outcome $\theta_i(t)$. Define the observation probability $P_i(\theta_i(t) = \theta')$, the probability that the SU i observes θ' at the time slot t. For each licensed channel k, the conditional probability $\lambda_k(t+1)$ depends not only on the observation of the SU, but also on its action. We denote by $\Omega(.|a_i(t), \theta_i(t))$ the update operator of the belief vector for each licensed channel.

Actions and strategies: Each SU has two actions to take sequentially, as illustrated in Figure 5.2. The first action, called *sensing-action*, is taken at the beginning of each time slot. This action determines whether the SU senses or not the licensed channels, based on the belief vector and the current packet delay. This sensing action induces an observation θ_i. Then, the SU takes a second action, called *access-action*, which determines if it transmits its packet using the licensed channel or not. Certainly, this action has to be taken only if there are free licensed channels, and the SU has a packet to transmit. The joint action of all SUs is denoted by $\mathbf{a}(t) = (a_1(t), \cdots, a_N(t))$, where $a_i(t)$ denotes the action of SU i and $\mathbf{a_{-i}}(t) = (a_1(t), \cdots, a_{i-1}, a_{i+1} \cdots, a_N(t))$ denotes the joint action set of SUs other than i. For notations convenience, we consider that the SU has only 3 possible actions:

- The action $a_i = 0$: the SU chooses to be inactive during the time slot. If the SU has a packet in its buffer, then the delay of the packet increases.
- The action $a_i = 1$: the SU chooses to sense licensed channels and not to transmit. Note that sensing licensed channels allows the SU to get more information that may improve the future rewards. If the SU has a packet in its buffer, then the delay of the packet increases.
- The action $a_i = 2$: the SU chooses to sense licensed channels and to transmit if idle. This action is possible only if the SU has a packet in its buffer.

Let us denote by $A_i(x_i)$ the action space of SU i, when it is in the state x_i, and by $A = \cup_i A_i$ the set of possible joint actions of SUs. Note that the action space for a SU depends on its state. For example, a SU that has no packet in its buffer $(l_i(t) = 0)$

cannot choose the action 2, i.e. $A_i = \{0, 1\}$. However, a SU having a packet to transmit chooses any action, i.e. $A_i = \{0, 1, 2\}$.

Based on the SU's action a_i and its observation θ_i, we have the following belief update, which comes from the Markov process. For all licensed channels $n \in \{1, \cdots, M\}$, the belief is updated as follows:

$$\lambda_n(t+1) := \Omega(\lambda_n(t)|a_i(t), \theta_i(t)) = \begin{cases} \beta_n + (\alpha_n - \beta_n)\lambda_n(t) & \text{if} \quad a_i(t) = 0; \\ \beta_n & \text{if} \quad a_i(t) \neq 0 \text{ and } \quad \theta_i(t) = 1; \\ \alpha_n & \text{if} \quad a_i(t) \neq 0 \text{ and } \quad \theta_i(t) = 0. \end{cases}$$

The strategy of SUs is defined by the probability of choosing a given action depending on its state $x_i(t) = (\lambda^i(t), l_i(t))$. We call a strategy for the SU i, a function $\mathbf{u_i}$ as a vector $[u_i(1), u_i(2), \ldots]$, where $u_i(t) : \mathcal{X}_i \times A_i \to [0, 1]$ is a mapping from a state $x_i(t)$ and an action $a_i(t)$ to a probability of taking the action $a_i(t)$ in the state $x_i(t)$. We denote by $\mathbf{u} := (u_1, \cdots, u_N)$ the multi-policy of all SUs (whose ith element is $u_i = [u_i(1), u_i(2), \ldots])$, and $\mathbf{u_{-i}}$ is the set of strategies of all SUs other than i. The set of all possible strategies is denoted $\mathcal{U}$.

Instantaneous reward: We denote by c_s the energy spent for sensing and c_t the energy spent for transmission. For each SU i, a natural definition of the instantaneous reward $r_i(t)$ is a composition of the throughput Φ and the energy costs. We introduce an additional cost, $f(l_i(t))$, in order to penalize the current packet delay. The instantaneous reward of a SU depends explicitly not only on its action $a_i(t)$, but also on the actions of all other SUs, denoted by $\mathbf{a_{-i}}(t)$. Furthermore, it depends on the state and the observation of SU i, x_i and θ_i. The instantaneous reward of the SU i at the time slot t is defined by:

$$r_i(x_i(t), \mathbf{a}(t), \theta_i(t)) = \begin{cases} \Phi - c_s - c_t, & \text{if } a_i(t) = 2,\ \theta_i(t) = 0 \text{ and } \forall j \neq i, a_j(t) \neq 2; \\ -c_s - c_t, & \text{if } a_i(t) = 2,\ \theta_i(t) = 0 \\ & \text{and } \exists j \neq i,\ a_j(t) = 2 \text{ (collision)}; \\ -f(l_i(t)) - c_s, & \text{if } a_i(t) = 1 \text{ or } a_i(t) = 2 \text{ and } \theta_i(t) = 1; \\ -f(l_i(t)), & \text{if } a_i(t) = 0. \end{cases} \tag{5.1}$$

where $\mathbf{a}(t) = [a_i(t)|\mathbf{a_{-i}}(t)]$, and $x_i(t) = (\lambda_i(t), l_i(t))$.

Problem statement: The objective of the SU i is to maximize the average expected reward, given the initial condition $x_i(0) = x_0$. Usually, in OSA problems modeled using a POMDP formulation, the objective function is the expected total discounted reward like in [80], [105], [106] and [107]. In our context, we observe that decisions have to be taken frequently, at each time slot, which leads to a discount rate very close to 1 (see [86]). Thus, it is natural to consider policies on the basis of their average expected reward. Therefore, the SU i seeks for the optimal strategy u_i that maximize:

$$R_i(u_i, \mathbf{u_{-i}}) = \lim_{T \to \infty} \frac{1}{T} \mathbb{E}_{\mathbf{u}} \left(\sum_{t=1}^{T} r_i(x_i(t), \mathbf{a}(t), \theta_i(t)) | x_0 \right). \tag{5.2}$$

We study the OSA problem in a non-cooperative setting, where each SU has its own state information and tries to maximize its average expected reward. Then, our problem will be studied in the following section through the concept of NE. Indeed, the SUs interact themselves through collisions when several SUs transmit over the same idle licensed channel. For simplicity reasons, and to get a deep theoretical analysis for the non-cooperative game between SUs, we consider only the set of stationary policies. A stationary policy is a mapping from a state x_i and action a_i to a probability $u_i(x_i, a_i)$, which does not depend on the time slot t. In the next section, we propose an analysis of the non-cooperative game. Our goal is to compute the set of all best responses strategies for a SU against a stationary multi-policy of all other SUs. Furthermore, we use a LP technique, which gives us a description of the NE for our non-cooperative game.

5.3 Nash equilibrium

In this section, we consider one licensed channel ($M = 1$), and N SUs trying to access it. Note that SUs decide, solely, whether to access or not this licensed channel. Each SU looks for maximizing its average expected reward defined in Equation (5.2). Before analyzing the NE and its properties, we define, in the next section, the Best Response (BR) strategy, a standard concept in game theory (see [108]).

5.3.1 The best response function

In game theory, the best response is defined to be the strategy (or strategies) that produces the most favorable outcome for a player, given others' strategies. The concept of best response is central to John Nash's best-known contribution, the Nash equilibrium.

Definition 5.1. The best response strategy BR(.) is defined as follows:

$$\forall i \in \{1, \cdots, N\}, \quad BR_i(\mathbf{u_{-i}}) = arg \max_{u_i} R_i(u_i, \mathbf{u_{-i}}). \tag{5.3}$$

Note that the average expected reward function $R_i(u_i, \mathbf{u_{-i}})$ can be expressed as follows:

$$
\begin{aligned}
R_i(u_i, \mathbf{u_{-i}}) &= \sum_{\mathbf{x}\in\mathcal{X}} \sum_{\mathbf{a}\in A} \sum_{\theta'=0}^{1} \prod_{j\neq i} \pi_j^{u_j}(x_j) u_j(x_j, a_j) r_i(x_i, \mathbf{a}, \theta') \pi_i^{u_i}(x_i) u_i(x_i, a_i) P_i(\theta_i = \theta') \\
&= \sum_{\mathbf{x}\in\mathcal{X}} \sum_{\mathbf{a}\in A} \pi_i^{u_i}(x_i) u_i(x_i, a_i) \prod_{j\neq i} \sum_{\theta'=0}^{1} P_i(\theta_i = \theta') \pi_j^{u_j}(x_j) u_j(x_j, a_j) r_i(x_i, \mathbf{a}, \theta') \\
&= \sum_{x_i\in\mathcal{X}} \sum_{a_i\in A_i} \pi_i^{u_i}(x_i) u_i(x_i, a_i) \sum_{\mathbf{x_{-i}}} \sum_{\mathbf{a_{-i}}} \sum_{\theta'=0}^{1} \qquad (5.4) \\
&\quad \prod_{j\neq i} P_i(\theta_i = \theta') \pi_j^{u_j}(x_j) u_j(x_j, a_j) r_i(x_i, \mathbf{a}, \theta'),
\end{aligned}
$$

where $\pi_i^{u_i}(x_i)$ is the stationary probability that the state of the SU i is x_i, which depends on the strategy u_i of the SU. The following lemma gives us a simpler expression of the average expected reward.

Lemma 5.2. *The average expected reward $R_i(u_i, \mathbf{u_{-i}})$ of the SU i is expressed as follows:*

$$
\begin{aligned}
R_i(u_i, \mathbf{u_{-i}}) &= \sum_{x_i\in\mathcal{X}_i} \sum_{a_i=0}^{1} \pi_i^{u_i}(x_i) u_i(x_i, a_i) r_i(x_i, \mathbf{a}, \theta_i) + [\Phi(1 - \bar{P}_{tr}(\mathbf{u_{-i}}))\Pi(0) \\
&\quad -(1 - \Pi(0)) f(l_i) - c_s - \Pi(0) c_t] u_i(x_i, 2), \qquad (5.5)
\end{aligned}
$$

where $\Pi(0)$ is the stationary probability that the licensed channel is idle, and $\bar{P}_{tr}(\mathbf{u_{-i}})$ represents the probability that at least one SU $j \neq i$ transmits over the licensed channel during the current time slot.

Proof. The average reward function, that a SU is trying to maximize, is expressed by:

$$
R_i(u_i, \mathbf{u_{-i}}) = \sum_{\mathbf{x}} \sum_{\mathbf{a}} \sum_{\theta'=0}^{1} P_i(\theta_i = \theta') \prod_{j\neq i} \pi_j^{u_j}(x_j) u_j(x_j, a_j) r_i(x_i, \mathbf{a}, \theta_i) \pi_i^{u_i}(x_i) u_i(x_i, a_i).
$$

Let us define the set $A^*_{-i} = \{a_{-i} | \exists j \neq i \; s.t. \; a_j = 2\}$. The expected reward can be expressed by:

$$\begin{aligned}
R_i(u_i, u_{-i}) &= \sum_{x_i}\sum_{x_{-i}}\sum_{a_i=0}^{1}\sum_{a_{-i}}\sum_{\theta'=0}^{1}\prod_{j\neq i} P_i(\theta_i=\theta')\pi_j^{u_j}(x_j)u_j(x_j,a_j)r_i(x_i,\mathbf{a},\theta_i)\pi_i^{u_i}(x_i)u_i(x_i,a_i) \\
&+ \sum_{x_i}\sum_{x_{-i}}\sum_{a_{-i}\in A^*_{-i}}\prod_{j\neq i} P_i(\theta_i=0)\pi_j^{u_j}(x_j)u_j(x_j,a_j)[-c_s-c_t]\pi_i^{u_i}(x_i)u_i(x_i,2) \\
&+ \sum_{x_i}\sum_{x_{-i}}\sum_{a_{-i}\in A^*_{-i}}\prod_{j\neq i} P_i(\theta_i=1)\pi_j^{u_j}(x_j)u_j(x_j,a_j)[-c_s-f(l_i)]\pi_i^{u_i}(x_i)u_i(x_i,2) \\
&+ \sum_{x_i}\sum_{x_{-i}}\sum_{a_{-i}\in A/A^*_{-i}}\prod_{j\neq i} P_i(\theta_i=0)\pi_j^{u_j}(x_j)u_j(x_j,a_j)[\Phi-c_s-c_t]\pi_i^{u_i}(x_i)u_i(x_i,2) \\
&+ \sum_{x_i}\sum_{x_{-i}}\sum_{a_{-i}\in A/A^*_{-i}}\prod_{j\neq i} P_i(\theta_i=1)\pi_j^{u_j}(x_j)u_j(x_j,a_j)[-c_s-f(l_i)]\pi_i^{u_i}(x_i)u_i(x_i,2) \\
R_i(u_i, u_{-i}) &= \sum_{x_i}\sum_{a_i=0}^{1}\pi_i^{u_i}(x_i)u_i(x_i,a_i)r_i(x_i,\mathbf{a},\theta_i) \\
&+ \sum_{x_i}\sum_{x_{-i}}\sum_{a_{-i}}\prod_{j\neq i} P_i(\theta_i=0)\pi_j^{u_j}(x_j)u_j(x_j,a_j)[-c_s-c_t]\pi_i^{u_i}(x_i)u_i(x_i,2) \\
&+ \sum_{x_i}\sum_{x_{-i}}\sum_{a_{-i}}\prod_{j\neq i} P_i(\theta_i=1)\pi_j^{u_j}(x_j)u_j(x_j,a_j)[-c_s-f(l_i)]\pi_i^{u_i}(x_i)u_i(x_i,2)
\end{aligned}$$

$$\begin{aligned}
&+ \sum_{x_i}\sum_{x_{-i}}\sum_{a_{-i}\in A/A^*_{-i}}\prod_{j\neq i} P_i(\theta_i=0)\pi_j^{u_j}(x_j)u_j(x_j,a_j)\Phi\pi_i^{u_i}(x_i)u_i(x_i,2) \\
R_i(u_i, u_{-i}) &= \sum_{x_i}\sum_{a_i=0}^{1}\pi_i^{u_i}(x_i)u_i(x_i,a_i)r_i(x_i,\mathbf{a},\theta_i) \\
&- \sum_{x_i}\sum_{x_{-i}}\sum_{a_{-i}}\prod_{j\neq i}\pi_j^{u_j}(x_j)u_j(x_j,a_j)c_s\pi_i^{u_i}(x_i)u_i(x_i,2) \\
&- \sum_{x_i}\sum_{x_{-i}}\sum_{a_{-i}}\prod_{j\neq i} P_i(\theta_i=0)\pi_j^{u_j}(x_j)u_j(x_j,a_j)c_t\pi_i^{u_i}(x_i)u_i(x_i,2) \\
&- \sum_{x_i}\sum_{x_{-i}}\sum_{a_{-i}}\prod_{j\neq i} P_i(\theta_i=1)\pi_j^{u_j}(x_j)u_j(x_j,a_j)f(l_i)\pi_i^{u_i}(x_i)u_i(x_i,2) \\
&+ \sum_{x_i}\sum_{x_{-i}}\sum_{a_{-i}\in A/A^*_{-i}}\prod_{j\neq i} P_i(\theta_i=0)\pi_j^{u_j}(x_j)u_j(x_j,a_j)\Phi\pi_i^{u_i}(x_i)u_i(x_i,2) \\
R_i(u_i, u_{-i}) &= \sum_{x_i}\sum_{a_i=0}^{1}\pi_i^{u_i}(x_i)u_i(x_i,a_i)r_i(x_i,\mathbf{a},\theta_i(t)) - c_s\pi_i^{u_i}(x_i)u_i(x_i,2) \\
&- \sum_{x_i} f(l)(1-\Pi(0))\pi_i^{u_i}(x_i)u_i(x_i,2) \\
&- \sum_{x_i}\Pi(0)c_t\pi_i^{u_i}(x_i)u_i(x_i,2) \\
&+ \sum_{x_i}\Phi(1-\bar{P}^*)\Pi(0)\pi_i^{u_i}(x_i)u_i(x_i,2).
\end{aligned}$$

□

Note that $\bar{P}_{tr}(\mathbf{u_{-i}})$ can be expressed as follows:

$$\bar{P}_{tr}(\mathbf{u_{-i}}) = 1 - \prod_{j \neq i} \sum_{x_j \in \mathcal{X}_j} \sum_{a_j=0}^{1} \pi_j^{u_j}(x_j) u(x_j, a_j). \tag{5.6}$$

Note that the interaction between the SU i and other SUs is summarized in the probability $\bar{P}_{tr}(\mathbf{u_{-i}})$. We are able now to define the expected instantaneous reward $\bar{r}_i$ for SU i as follows:

$$\bar{r}_i(x_i, a_i, \mathbf{u_{-i}}) = \sum_{\theta'=0}^{1} I\!E_{\mathbf{u}}[r_i(x_i, \mathbf{a}, \theta_i)] P_i(\theta_i = \theta') \tag{5.7}$$

$$\bar{r}_i(x_i, a_i, \mathbf{u_{-i}}) = \begin{cases} (\Phi(1 - \bar{P}_{tr}(\mathbf{u_{-i}})) + f(l) - c_t)\Pi(0) - f(l) - c_s, & \text{if } a_i = 2, \\ -f(l_i) - c_s, & \text{if } a_i = 1, \\ -f(l_i), & \text{if } a_i = 0. \end{cases}$$

Note that $\bar{r}_i(x_i, a_i, \mathbf{u_{-i}})$ represents the instantaneous reward that the SU i expect when taking the action a_i in the state x_i, and the multi-policy of all other SUs is $\mathbf{u_{-i}}$. Thus, the average expected reward $R_i(u_i, \mathbf{u_{-i}})$, given by Lemma 5.2, can be rewritten as follows:

$$R_i(u_i, \mathbf{u_{-i}}) = \sum_{x_i} \sum_{a_i} \pi_i^{u_i}(x_i) u_i(x_i, a_i) \bar{r}_i(x_i, a_i, \mathbf{u_{-i}}). \tag{5.8}$$

The set of best response strategies for a SU, given fixed strategies for all other SUs, can be computed using a LP, as proposed in [100]. In the following, we present such a LP, which determines the set of all best response strategies for player i against a stationary policy $\mathbf{u_{-i}}$ of all its opponents. We denote by $z_{i,u_i}(x_i, a_i) = \pi_i^{u_i}(x_i) u_i(x_i, a_i)$, the steady state probability that the system state of SU i is $x_i \in \mathcal{X}$, and that the action $a_i \in \mathcal{A}_i$ is chosen. The following LP gives us the best response policies, for all SUs $i \in \{1, \cdots, N\}$, and for all multi-policy $\mathbf{u} \in \mathcal{U}$.

LP(i,u): Find $z^*_{i,u_i}(x_i, a_i)$, where $(x_i, a_i) \in \mathcal{X}_i \times \mathcal{A}_i$, that maximizes:

$$\sum_{x_i} \sum_{a_i} z_{i,u_i}(x_i, a_i) \bar{r}_i(x_i, a_i, \mathbf{u_{-i}}),$$

subject to:

$$\begin{aligned} \sum_{a_j} z_{i,u_i}(r, a_j) - \sum_{x_i} \sum_{a_i} z_{i,u_i}(x_i, a_i) p_{x_i a_i r} &= 0, \forall r \in \mathcal{X}, \\ \sum_{x_i} \sum_{a_i} z_{i,u_i}(x_i, a_i) &= 1, \\ z_{i,u_i}(x_i, a_i) &\geq 0, \end{aligned}$$

where p_{xay} is the probability that the system switches from state x to state y by taking the action a.

Let $M_1(A)$ denote the set of probabilities measures over a set A, and let us define $\Gamma_i(\mathbf{u})$ as the set of optimal solutions of $\mathbf{LP(i,u)}$. A point to set mapping $\gamma_i(z_i)$, given a non-negative real numbers $z_i = \{z_i(\mathbf{x}, \mathbf{a}), (x_i, a_i) \in \mathcal{X}_i \times \mathcal{A}_i\}$, is defined as follows:

- if $\sum_{a_i} z_i(x_i, a_i) \neq 0$ then $\gamma_i(x_i, a_i, z_i) := \{\frac{z_i(x_i,a_i)}{\sum_{a_i'} z_i(x_i,a_i')}\}$ is a singleton. Note that $\gamma_i(x_i, z_i) = \{\gamma_i(x_i, a_i, z_i) : a_i \in A_i(x_i)\}$ is a point in $M_1(A_i(x_i))$.
- Otherwise, $\gamma_i(x_i, z_i) := M_1(A_i(x_i))$, the set of all probabilities measures over $A_i(x_i)$.

Define $g_i(z_i)$ as the set of stationary policies that choose the action a_i in the state x_i with a probability in $\gamma_i(a_i, x_i, z_i)$. Moreover, we define the occupancy measures $f(x_0, \mathbf{u})$ for a multi-policy $\mathbf{u}$ as $\{f_i(x_0, \mathbf{u}), (a_i, x_i) \in \mathcal{X}_i \times \mathcal{A}_i, \forall i | f_i(x_0, \mathbf{u}) = \pi_i^{u_i}(x_i) u_i(x_i, a_i)\}$. Note that for each player i and stationary policy u_i, the state of that player is an irreducible Markov chain with one ergodic class. Thus, a unique steady-state probability exists. Therefore, we can omit the initial state distribution x_0.

Proposition 5.3. *For any stationary multi-policy OSA for SUs, we have the following properties:*

1. *If $z^*_{i,\mathbf{u}}$ is an optimal solution of $\mathbf{LP(i,u)}$, then any element $v \in g_i(z^*_{i,\mathbf{u}})$ is an optimal stationary response for SU i against the stationary policy $\mathbf{u_{-i}}$ of other SUs. Moreover, the multi-policy $\mathbf{w} = [v|\mathbf{u_{-i}}]$ satisfies $f_i(\mathbf{w}) = z^*_{i,\mathbf{u}}$.*
2. *The optimal sets $\Gamma_i(\mathbf{u}), \forall i$ are convex, compact, and upper semi-continuous in $\mathbf{u_{-i}}$, where $\mathbf{u}$ is identified with points in $\prod_{i=1}^{N} \prod_{x_i} M_1(A_i(x_i))$.*
3. *For all i, $g_i(z_i)$ is upper semi-continuous in z over the set of solutions for* $\mathbf{LP(i,u)}$.

Proof. The proof of (1) follows from Theorem 2.6 of [109]. The first part of (2) is a direct result of the LP. However, the second part follows by applying the theory of sensitivity analysis of LP by Dantzig et al. [110] in the Theorem 3.6 of [111] to $\mathbf{LP(i,u)}$. The last property follows from the definition of $g_i(z_i)$. □

5.3.2 The Nash equilibrium

We model the interaction between SUs as a non-cooperative game. Let us define the concept of NE between SUs in our model.

Definition 5.4. The NE is defined as a set of strategies (one for each player) $\mathbf{u}^* = (u_1^*, u_2^*, ..., u_N^*)$, such that:

$$\forall i \in \{1, \ldots, N\}, \quad u_i^* = \arg\max_{u_i} R_i(u_i, \mathbf{u}^*_{-\mathbf{i}}). \tag{5.9}$$

A successful transmission for a SU over the licensed channel depends not only on the PUs' activity but also on the competition with other SUs. When a SU senses the channel as idle, it transmits successfully its packet if and only if the action of all other SUs is not to transmit on the licensed channel during the current slot. Indeed, a SU that chooses an action $a \in \{0, 1\}$ does not impact the instantaneous reward of other SUs. Given this remark, we have the following theorem, which states the existence of a NE multi-policy for our OSA problem between SUs.

Theorem 5.5. *There exists a stationary multi-policy* $\mathbf{u}^*$ *that is a Nash equilibrium.*

Proof. Consider a fixed value of the stationary probability that the channel is idle, $\Pi(0)$. Note that for each SU i and any stationary policy u_i, the state process of that SU is an irreducible Markov chain with one ergodic class. Moreover, the strategies chosen by any SU does not depend on the cost realization. Otherwise, a SU could use the costs to estimate the state and actions of other SUs. Thus, from the Theorem of fixed point of Kakutani, a fixed point $u_i \in BR(\mathbf{u}_{-\mathbf{i}})$ exists. Proposition 5.3 implies that the stationary multi-policy $g = \{g_i(z_i)\forall i\}$ is a NE. □

After proving the existence of a NE of our game, the second problem we address now is to determine a particular type of equilibrium: the Symmetric Nash Equilibrium. A symmetric multi-policy $\mathbf{u}^* = (u^*, u^*, \cdots, u^*)$ is an SNE if and only if:

$$R_i(u^*, \mathbf{u}^*_{-\mathbf{i}}) \geq R_i(u_i, \mathbf{u}^*_{-\mathbf{i}}), \ \forall i \text{ and } \forall u_i \neq u^*. \tag{5.10}$$

In order to find an SNE, we assume that $N-1$ SUs use a strategy u_0, and a tagged SU (without loss of generality, the user N) uses the strategy u_N. Therefore, a multi-policy $\mathbf{u} = (u_0, \cdots, u_0, u_N) := (\mathbf{u}_{-\mathbf{N}}, u_N)$ is an SNE if and only if:

$$u_N = u_0 \in BR(\mathbf{u}_{-\mathbf{N}}). \tag{5.11}$$

5.3.3 Properties of the Nash equilibrium

Let us define by $P_{tr}(u_i)$ the attempt rate for a SU i. $P_{tr}(u_i)$ is expressed as follows:

$$P_{tr}(u_i) = \sum_{x'_i \in \mathcal{X}} \pi_i^{u_i}(x'_i) u_i(x'_i, 2), \tag{5.12}$$

where $\pi_i^{u_i}(x_i)$ is the stationary probability that the state of the SU i is x_i, and u_i is the mixed strategy of the SU i. The following proposition states that for each SU i, its attempt rate is always the same at different SNE of the game.

Proposition 5.6. *Consider two SNE $\mathbf{u_1^*} \neq \mathbf{u_2^*}$, such that $\mathbf{u_1^*} = (u_1^*, \cdots, u_1^*)$ and $\mathbf{u_2^*} = (u_2^*, \cdots, u_2^*)$. Therefore, the attempt rates for any SU i at the SNE are unique and equal:*

$$\forall i \in \{1, \cdots N\}, \quad P_{tr}(u_1^*) = P_{tr}(u_2^*) := P^*.$$

Proof. Consider z_0^* the solution of the LP that maximizes $\bar{r}_i(x_i, a_i, \mathbf{u_{-i}})$, and z_ϵ^* the solution of the LP that maximizes $\bar{r}_i(x_i, a_i, \mathbf{u_{-i}}) + \epsilon \mathbb{1}_{a_i=2}$. Note that, in the second problem, the reward for the action 2 is increased, compared to the first one. Assume that $\sum_{x_i} z_0^*(2, x_i) > \sum_{x_i} z_\epsilon^*(2, x_i)$, then we obtain:

$$\begin{aligned} & \sum_{x_i} \sum_{a_i} z_\epsilon^*(a_i, x_i) \bar{r}_i(x_i, a_i, \mathbf{u_{-i}}) + \epsilon \sum_{x_i} z_\epsilon^*(2, x_i), \\ \leq & \sum_{x_i} \sum_{a_i} z_0^*(a_i, x_i) \bar{r}_i(x_i, a_i, \mathbf{u_{-i}}) + \epsilon \sum_{x_i} z_\epsilon^*(2, x_i), \\ < & \sum_{x_i} \sum_{a_i} z_0^*(a_i, x_i) \bar{r}_i(x_i, a_i, \mathbf{u_{-i}}) + \epsilon \sum_{x_i} z_0^*(2, x_i). \end{aligned} \tag{5.13}$$

Therefore, z_0^* is the optimal solution that maximizes $\bar{r}_i(x_i, a_i, \mathbf{u_{-i}}) + \epsilon \mathbb{1}_{a_i=2}$, which leads to a contradiction as z_ϵ^* is assumed to be the optimal solution of $\bar{r}_i(x_i, a_i, \mathbf{u_{-i}}) + \epsilon \mathbb{1}_{a_i=2}$. The first inequality is because z_0^* maximizes $\bar{r}_i(x_i, a_i, \mathbf{u_{-i}})$, and the second one is due to the assumption. Then, we obtain that $\sum_{x_i} z_0^*(x_i, 2) \leq \sum_{x_i} z_\epsilon^*(x_i, 2)$.

Note that the attempt rate of the SU i is expressed by $P_{tr}(u_i)$, and the attempt rate of other SUs is expressed by $\bar{P}_{tr}(\mathbf{u_{-i}})$. Therefore, if the attempt rate of other SUs decreases, the reward $\bar{r}_i(x_i, 2)$ increases and then the attempt rate $P_{tr}(u_i)$ increases. In fact, a SU decreases its attempt rate if all the other SUs increase their attempt rates. Finally, the BR function of SU i decreases with the attempt rate of other users $\bar{P}_{tr}(\mathbf{u_{-i}})$.

Since we are considering SNE strategies, we have $P_{tr}(u_i) = \bar{P}_{tr}(\mathbf{u_{-i}})$. Suppose that there are two Nash equilibrium strategies, $\mathbf{u^1}$ and $\mathbf{u^2}$ having different attempt rates,

$P_{tr}(u^1) < P_{tr}(u^2)$. As both $\mathbf{u^1}$ and $\mathbf{u^2}$ are SNE, we have the following inequality:

$$P_{tr}(BR_i(\mathbf{u^1_{-i}})) = P_{tr}(u^1) < P_{tr}(u^2) = P_{tr}(BR_i(\mathbf{u^2_{-i}})), \tag{5.14}$$

which lead to a contradiction, as $BR_i(.)$ is a decreasing function with respect to the attempt rate. □

We denote by P^* the attempt rate of a SU when all SUs use a SNE strategy. As usual in non-cooperative games, the utilization of the resource is suboptimal at the NE. In the following section, we look for a network manager's control mechanism in order to optimize an important global metric of the system, the average total throughput.

5.4 Network management

The SNE between SUs has been deeply investigated using a LP technique in the previous section. Note that interactions between SUs induce collisions. Henceforward, we focus on the impact of the PUs' activity on the performance of the global system. Since the resource utilization at the SNE is generally suboptimal, we propose to introduce some control in order to enhance the spectrum utilization. We propose a simple mechanism by introducing some kind of hierarchy in the OSA game. We obtain this hierarchy by introducing a controller, named the network manager. This controller plays as a leader in the Stackelberg game, and the SUs play as followers.

We formulate the problem of maximizing the average total throughput of the system as a Stackelberg game. The objective of the network manager is to maximize the average total throughput of the system at the SNE. Note that the average total throughput of the system is defined as follows:

$$U^* = \frac{1}{N}\sum_{i=1}^{N} P_{tr}(u_i^*)\prod_{j\neq i}(1 - P_{tr}(u_j^*)).$$

From Proposition 5.6 , the attempt rates at the SNE of all SUs are equals. Thus, we obtain:

$$U^* = P^*(1 - P^*)^{N-1}.$$

The following proposition gives us the attempt rate at the SNE that maximizes the average total throughput of the system.

Proposition 5.7. *When the attempt rate at the SNE, P^*, is equal to $\frac{1}{N}$, the average total throughput U^* is maximized.*

Proof. As we have N users transmitting over the same licensed channel, with an average probability of P, we have a successful transmission, if the channel is idle, with probability $P(1-P)^{N-1}$. The probability P^* maximizes $P(1-P)^{N-1}$ if and only if $(1-P^*)^{N-1} - P^*(N-1)(1-P^*)^{N-2} = 0$, then $(1-NP^*)(1-P^*)^{N-2} = 0$. Therefore, when $P^* = \frac{1}{N}$, the utility for SUs is optimal. □

Note that the attempt rate P^* obtained from a multi-policy SNE, given by Theorem 5.5, does not necessarily equal the optimal attempt rate obtained from Proposition 5.7. Then, the network manager makes a decision (an intervention) in order to influence the SNE multi-policy.

The question that we have to answer is how the network manager can impact SUs' policies in order to maximize the average total throughput of the system at the SNE. Before, we state, in the following proposition, some properties of the attempt rate and the channel occupancy. The following proposition shows that increasing the channel occupancy decreases the attempt rate of SUs at the SNE.

Proposition 5.8. *P^* is decreasing when $\Pi(0)$ decreases.*

Proof. Consider two stationary probabilities that the channel is idle $\Pi_1(0)$ and $\Pi_2(0)$, such that $\Pi_1(0) < \Pi_2(0)$. Consider two SNE strategies, $\mathbf{u^1}$ obtained with the stationary probability $\Pi_1(0)$, and $\mathbf{u^2}$ obtained with the stationary probability $\Pi_2(0)$. Note that, for a given value of attempt rate P^*, the immediate reward for the action $a_i = 2$ is higher for the channel having a stationary probability of $\Pi_2(0)$ than for the channel having a stationary probability of $\Pi_1(0)$ (see Equation (5.8)). Let us denote by $P_{tr}(u^1)$ the attempt rate obtained with strategy $\mathbf{u^1}$, and by $P_{tr}(u^2)$ the attempt rate obtained with strategy $\mathbf{u^2}$. We obtain from Proposition 5.6 that $P_{tr}(u^1) < P_{tr}(u^2)$ (decreasing $\Pi(0)$ decreases the instantaneous reward for the action $a_i = 2$).

Finally, we obtain that the attempt rate P^* decreases when the stationary probability that the licensed channel is idle decreases. □

We have the following relationship between $\Pi(0)$ and β_0.

Lemma 5.9. *$\Pi(0)$ is increasing with β_0.*

Proof. The stationary probability $\Pi(0)$ is defined as follows:

$$\Pi(0) = \frac{\beta_0}{1 - \alpha + \beta_0}.$$

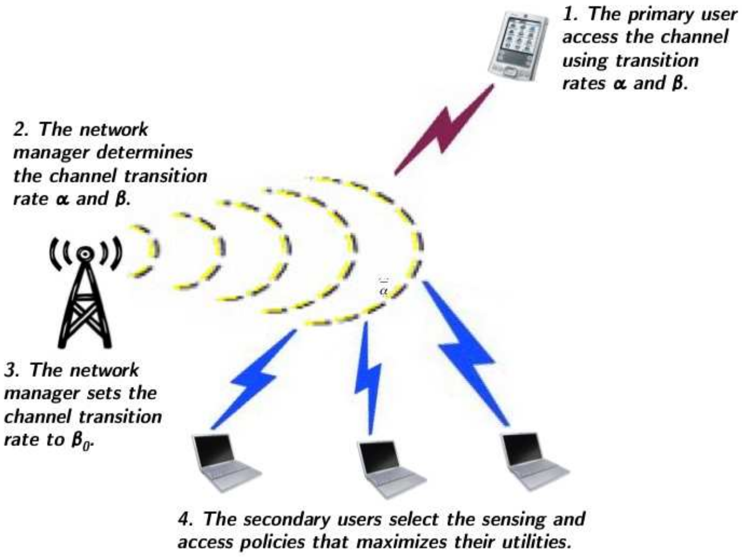

FIGURE 5.3: The Stackelberg game model of the SU throughput maximization.

The derivative of $\Pi(0)$ with respect to β_0 is:

$$\frac{\partial \Pi(0)}{\partial \beta_0} = \frac{1-\alpha}{(1-\alpha+\beta_0)^2}.$$

As $\alpha \in [0,1]$, then the derivative of $\Pi(0)$ with respect to β_0 is always positive. Therefore $\Pi(0)$ is increasing with β_0. □

Given this result, the network manager varies the channel occupancy state in order to maximize the average total throughput of SUs at the SNE. Figure 5.3 depicts the relationships between PUs, the network manager and SUs.

Moreover, the stationary probability that the licensed channel is idle is given by $\Pi(0) = \frac{\beta}{1-\alpha+\beta}$. It is obvious that the stationary probability $\Pi(0)$ is increasing with β. Thus, by reducing β, the network manager can reach a target value of stationary probability $\Pi(0)$ that maximizes the average total throughput of SUs at the SNE. We denote by β_0 the transition rate that maximizes the average total throughput of SUs at the SNE.

Remark 5.10. Note that if $P^* > \frac{1}{N}$, then $\beta_0 < \beta$, and the network manager increases the channel occupancy in order to maximizes the average total throughput of SUs at the SNE. However, if $P^* < \frac{1}{N}$, then the target value β_0 that maximizes the average total throughput at the SNE is above the PUs' transmission rate, i.e. $\beta_0 > \beta$. Therefore, the network manager cannot improve the performance of the system. Indeed, the network manager can only decrease the transition rate from state occupied to idle, by occupying the licensed channel after it was already occupied. Figure 5.4 illustrates the impact of the transition rate β_0 on the attempt rate when using an SNE policy.

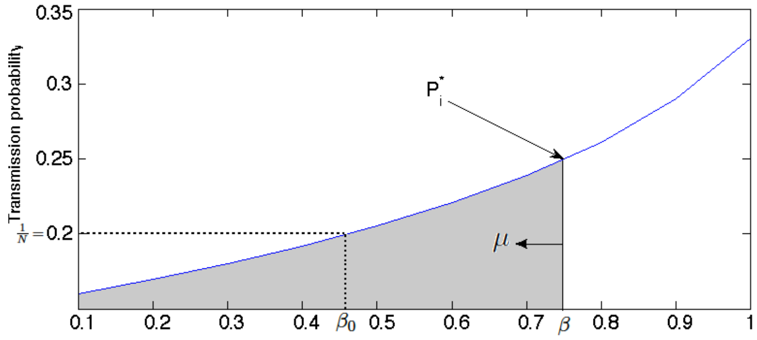

FIGURE 5.4: The attempt rate when using a SNE policy with respect to the transition rate β_0.

Let us define the network manager's (leader) actions by:

- a_1^p: the network manager occupies the licensed channel if this channel was already occupied in the previous slot and becomes idle in the current slot;
- a_2^p: the network manager does not occupy the channel if this channel was occupied in the previous slot and becomes idle in the current slot.

In fact, when the leader chooses the action a_1^p, the licensed channel is not used by PUs but appears occupied for the followers (SUs). Then, the leader's action impacts the SNE of the followers. The set of the leader's actions is denoted $\mathcal{A}_l = \{a_1^p, a_2^p\}$. We define a mixed strategy of the leader by a mapping $\mu : \mathcal{A}_l \rightarrow [0, 1]$, where $\mu(a)$ is the probability that the leader takes the action a. Note that we have $\mu(a_2^p) = 1 - \mu(a_1^p)$. Given a strategy μ of the network manager, the induced transition rate β' is:

$$\beta'(\mu) \quad = \quad (1 - \mu(a_1^p)) \times \beta, \tag{5.15}$$

where β is the transition rate of PUs. Denote by $\mathbf{u}^*(\mu)$ the SNE of the followers when the leader's strategy is μ. In fact, the action of the leader μ changes the transition rate from β to $\beta'(\mu)$, which impacts the SNE of the followers. The objective of the leader (network manager) is therefore to find a strategy μ that maximizes the average throughput of the system:

$$\bar{U}(\mu, \mathbf{u}^*(\mu)) = \frac{1}{N} \sum_{i=1}^{N} Thr_i(\mathbf{u}^*(\mu)) = P^*(\mathbf{u}^*(\mu))(1 - P^*(\mathbf{u}^*(\mu)))^{N-1}. \tag{5.16}$$

The network manager problem can be expressed as follows:

$$\mu^* = \arg\max_{\mu} U(\mu, \mathbf{u}^*(\mu)), \tag{5.17}$$

where $\mathbf{u}^*(\mu)$ is an SNE among the N SUs taking into account the strategy of the leader. The vector of actions $(\mu^*, \mathbf{u}^*(\mu^*))$ is by definition a Stackelberg equilibrium [108], and we have the following theorem, which proves the existence of such equilibrium.

Theorem 5.11. *There exists a Stackelberg equilibrium for our hierarchical game with a network manager and N SUs.*

Proof. We have proved, in Proposition 5.7, that the attempt rate at the SNE P^*, which maximizes the leader's utility should be equal to $P^* = \frac{1}{N}$, where N is the number of SUs. Moreover, we have proved, in Proposition 5.8, that P^* decreases when $\Pi(0)$ decreases, and that $\Pi(0)$ is increasing with β. Thus, the leader computes the value of $\beta' = \min\{\beta_0, \beta\}$, and uses the following strategy:

$$\mu(a_1^p) = 1 - \frac{\beta'}{\beta}, \text{ and } \mu(a_2^p) = \frac{\beta'}{\beta}.$$

Note that SUs converge to an SNE where every SU maximizes its own utility taking into account the new channel transition rates (α, β'). Therefore, there exists a Stackelberg equilibrium between the network manager and SUs. □

5.5 Numerical illustrations

We illustrate, in this section, some Matlab-based simulation results in both saturated ($q_a = 1$) and non-saturated regimes ($q_a < 1$). We consider five SUs ($N = 5$) transmitting opportunistically, and we assume that the deadline delay is 3 slots. The deadline delay is the time by which the packet must be transmitted. We set the transmission cost $c_t = 100$; the sensing cost $c_s = 5$ and the throughput $\Phi = 200$kbit/s. Moreover, we consider a delay penalty function $f(l) = \min\{l, l_{max}\}$, where l_{max} is the deadline delay.

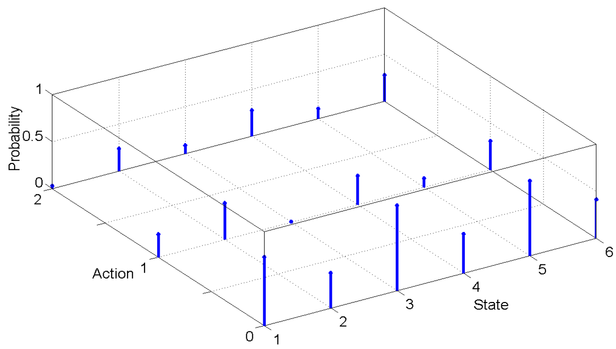

FIGURE 5.5: The equilibrium policy in the saturated case with $\alpha = 0.1$, $\beta = 0.9$ and $c_t = 100$.

5.5.1 Symmetric Nash equilibrium

Consider, first, the saturated regime, where SUs have always packets to transmit. Therefore, we obtain the following set of states:

State index	1	2	3	4	5	6
l	1	1	2	2	2	2
λ	α	β	α	β	$\Omega(\alpha)$	$\Omega(\beta)$

We can observe, in Figure 5.5 obtained with $\alpha = 0.1$ and $\beta = 0.9$, that a SU chooses a mixed strategy composed of the three possible actions: sleeping; sensing; sensing and transmitting. Moreover, when the transmission cost increases $c_t = 500$, we observe, in Figure 5.6, that SUs have less incentive to sense and transmit.

Secondly, we focus on the non-saturated regime with $q_a = 0.85$. When a SU transmits a packet, its local state l becomes 1 if it receives a new packet at the time slot t (with probability q_a), otherwise $l = 0$. Therefore, we obtain the following set of states:

State index	1	2	3	4	5	6	...	18
l	0	0	0	0	0	0	...	2
λ	α	β	$\Omega(\alpha)$	$\Omega(\beta)$	$\Omega^2(\alpha)$	$\Omega^2(\beta)$	...	$\Omega^2(\beta)$

Consider $\alpha = 0.9$ and $\beta = 0.1$, a scenario where the licensed channel stays in the same state during long periods, as it is the case with TV white bands [78]. We plot, in Figure 5.7, the multi-policy SNE obtained after solving the LP. We observe that the probability of sensing when the SU has no packet to transmit, i.e. $a_i = 1$, is increasing with the number of consecutive time slots the SU have not sensed the licensed channel. It means that the SU tries to get information about licensed channels by sensing even if it has no packet to transmit.

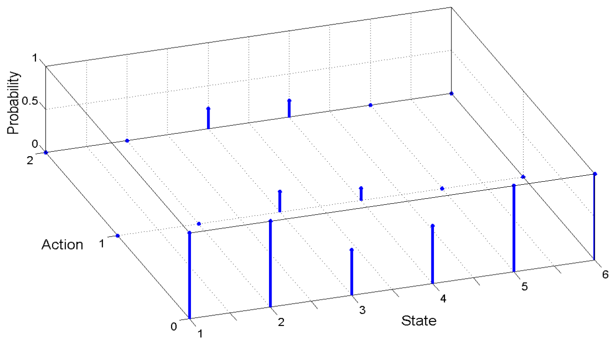

FIGURE 5.6: The equilibrium policy in the saturated case with $\alpha = 0.1$, $\beta = 0.9$ and $c_t = 500$.

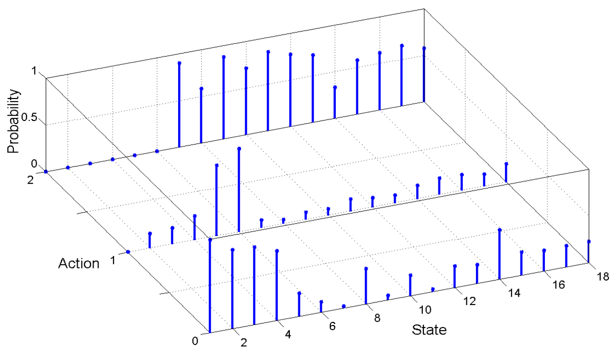

FIGURE 5.7: The equilibrium policy in the non saturated case with $\alpha = 0.9$, $\beta = 0.1$ and $q_a = 0.85$.

5.5.2 Braess paradox

Figure 5.8 illustrates the attempt rate P^* depending on the number of SUs. We observe that the attempt rate at the SNE is decreasing with the number of SUs, which is somehow intuitive, as the collision probability $1 - P^*(1 - P^*)^{N-1}$ increases due to the competition between SUs. In Figure 5.9, a Braess kind of paradox is illustrated. Indeed, there is a degradation of the performance of the system when additional resource is added. Specifically, we have an opposite formulation, saying that reducing system resources induce better performances. When the average spectrum occupancy (stationary probability that the licensed channel is occupied, i.e. $\frac{1-\alpha}{1-\alpha+\beta}$) is less than 0.5, the average throughput of the system increases with the average occupation of the channel.

In order to understand this phenomenon, we study the impact of the average channel occupancy on the average total throughput of the system. The SUs' attempt rate is decreasing when the channel is less available. Surprisingly, the average throughput is not always increasing with the offered channel opportunities. In fact, we observe, in Figure 5.9, that when the channel is available more than 50% of time, the average SUs' throughput is decreasing when the licensed channel is idler. The attempt rate is $P = \frac{1}{5}$

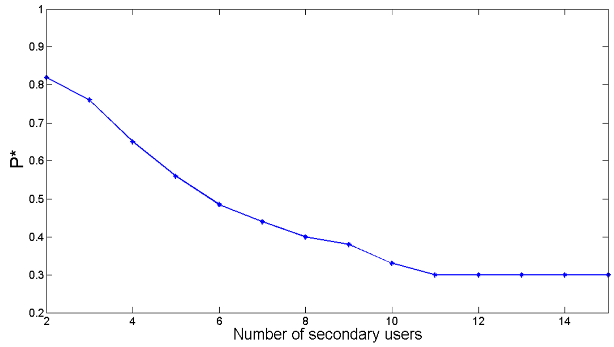

FIGURE 5.8: The attempt rate at the SNE depending on the number of SUs for $\alpha = 0.95$ and $\beta = 0.9$.

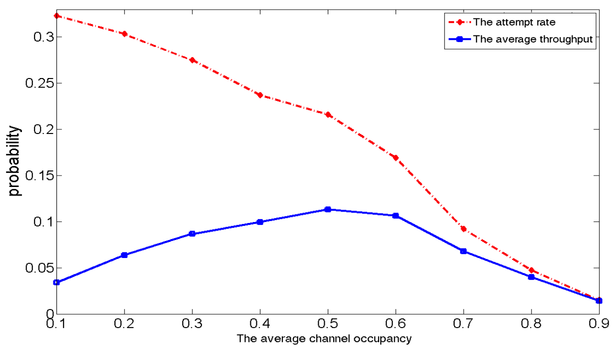

FIGURE 5.9: The attempt rate and the average throughput with the channel occupancy for $c_t = 100$.

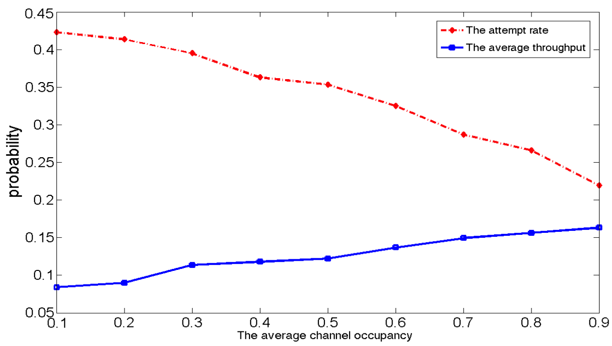

FIGURE 5.10: The attempt rate and the average throughput with the channel occupancy for $c_t = 900$.

when the channel availability is 0.5, and the average throughput is maximal for this channel availability. Note that it has been already proved that the SUs' attempt rate, that maximizes the average total throughout is $\frac{1}{N}$, where N is the number of SUs. In Figure 5.10, there is another example in which the average throughput is always increasing with the average channel occupancy.

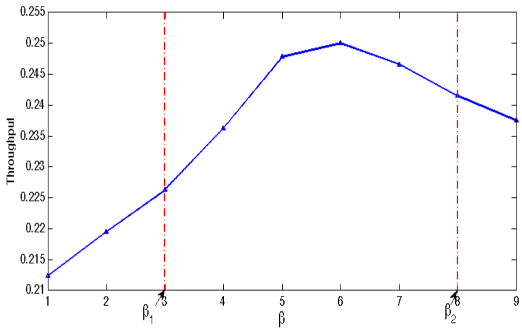

FIGURE 5.11: The average throughput depending on β.

5.5.3 Stackelberg equilibrium

Let us consider a scenario where two SUs are competing in order to access a licensed channel. The PUs' transition rate α is set to 0.1. We consider, first, that $\beta = 0.8$, and we illustrate, in Figure 5.11, the average throughput of the SUs depending on the transition rate β. We observe that the optimal value of β_0, which is also the transition rate at the Stackelberg equilibrium, is equal to 0.6. Therefore, the network manager has to decrease the transition rate from the occupied state to the idle state (i.e. β) from 0.8 to 0.6, which increases the average throughput of SUs from 0.2415 to 0.25.

Secondly, we consider that the PUs' requirement is $\beta = 0.3$. Thus, the network manager has to increase β_0 in order to increase the average throughput of SUs, which require that PUs use less the licensed channel. However, as we have already assumed that the SUs' access is opportunistic, PUs are unaware of the presence of SUs, and the network manager cannot increase β_0. Thus, the optimal action of the network manager is to be inactive ($\beta_0 = \beta$), as it cannot improve the actual SUs' performance.

Finally, Figure 5.12 illustrates the average channel availability ($\Pi(0)$) that maximizes the throughput for SUs at the SNE. We considered that PUs occupy the licensed channel with a probability $\Pi(1) = 0.5$. Then, when the cost is higher than 100, there is no paradox, as we cannot increase the channel availability (the network manager has to increase $\Pi(0)$).

5.6 Conclusion

In this chapter, we have set up a non-cooperative OSA mechanism for CR networks, and we have considered that SUs are in competition in order to access a licensed channel. Both the saturated and the non-saturated regimes have been studied, and we have

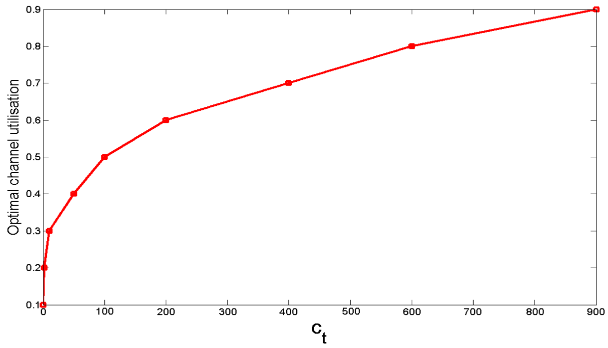

FIGURE 5.12: The optimal channel utilization with the transmission cost.

proved the existence of an SNE multi-policy for the OSA problem, modeled as a non-cooperative game between SUs. Moreover, we have proved that the attempt rate at the SNE is unique. The impact of both the arrival rate and the transmission cost on the system performances has been deeply studied. Simulation results have shown that more opportunities of transmission may decreases the average throughput of the system due to the aggressiveness and the competition between SUs. In fact, we have found Braess paradox where reducing system resources induce better performance. In order to optimize the average throughput of the system, we have proposed a Stackelberg game model for the network manager. We have proved the existence of an optimal strategy for the network manager. This strategy is defined by increasing the average time that the licensed channel is occupied.

In the following part of this thesis, we study self-adaptive congestion control at the transport layer, especially for multimedia applications. We focus, in the next chapter, on the resources management in wireless networks at upper layer of the protocol stack, the transport layer. Specifically, we propose some content-aware congestion control mechanisms for partially observable environments.