國 立 交 通 大 學

電信工程學系

碩 士 論 文

利用限制隨機賽局在認知網路下動態管理功率

Dynamic Power Management in Cognitive Radio Networks based on Constrained Stochastic Games

研 究 生：王佳偉

指導教授：方凱田 教授

中 華 民 國 九 十 八 年 六 月

利用限制隨機賽局在認知網路下動態管理功率

Dynamic Power Management in Cognitive Radio Networks based on
Constrained Stochastic Games

研 究 生 ：王佳偉　　Student：Chia-Wei Wang

指導教授 ：方凱田　　Advisor：Kai-Ten Feng

國 立 交 通 大 學
電 信 工 程 學 系
碩 士 論 文

A Thesis
Submitted to Department of Communications Engineering
College of Electrical and Computer Engineering
National Chiao Tung University
in partial Fulfillment of the Requirements
for the Degree of
Master of Science
in

Communications Engineering

June 2009

Hsinchu, Taiwan, Republic of China

中華民國九十八年六月

利用限制隨機賽局在認知網路下動態管理功率

學生：王佳偉 指導教授：方凱田教授

國立交通大學電信工程學系碩士班

摘 要

近年研究指出,已分配頻帶的頻寬使用效率低，而為了促使頻寬使用效率的提升，認知網路(Cognitive Radio, CR)則被提出來動態的使用這些已分配卻效率不高的頻帶。其中在認知網路中使用者之間的訊號干擾與功率分配則被提出許多的相關研究。因此，在這篇論文中,使用賽局理論(Game Theory)的限制隨機賽局(Constrianed Stochastic Game)在動態的通道環境與存在頻寬的使用者下，求出此問題的最佳決策。內容的研究中，分別求出有限時間和無限時間下，包含了已分配與未分配頻帶的最佳決策。而在求解的過程中，均對信號的干擾做了限制，因此CR的使用者對頻寬的擁有者不會造成嚴重的干擾。根據賽局理論模型的表示，可以證明存在賽局的奈許平衡解(Nash equilibrium)，而此奈許平衡解可以使每個CR使用者在彼此競爭的情況下得到個人的最佳化。在模擬的部分，驗證了確實可達到奈許平衡解，也顯示可優於貪婪式的演算法(Greedy mechanism)，並且對有通道感測的誤差下仍可達到可預期的結果。

Dynamic Power Management in Cognitive Radio Networks based on Constrained Stochastic Games

student：Chia-Wei Wang Advisors：Dr. Kai-Ten Feng

Department of Communication Engineering
National Chiao Tung University

ABSTRACT

Recent studies have been conducted to indicate the ineffective usage of licensed bands due to the static spectrum allocation. In order to improve the spectrum utilization, the cognitive radio is therefore suggested to dynamically exploit the opportunistic primary frequency spectrums. The interference from the secondary users to the primary user consequently draws the attention to the spectrum and power management for the cognitive radio networks. In this paper, the constrained stochastic games are utilized to exploit the optimal policies for power management by considering the variations from both the channel gain and the primary traffic. Both the underlay and overlay waveforms are considered within the network scenarios for the proposed power management scheme. Constraints for allowable interferences will be applied in order to preserve the communication quality among the primary and the secondary users. With the assumption of the Markovian property of dynamic environment, finite and infinite time horizon scenarios are both considered in target function. According to the formulation of the constrained stochastic games, the existence of the constrained Nash equilibrium will be validated with rigorous proofs. Simulation results further validate the correctness of the theoretically-derived policies, compare with the greedy mechanism and examine the effect of channel sensing error for dynamic power management.

誌　　謝

回頭看兩年的時間過去就是覺得特別快，但在交通大學的研究生活卻是我難忘的深刻體驗。畢業典禮的到來讓我開始回想起在這兩年研究的經歷，從一開始的對研究的無所適從，一步步對研究的探討，到最後論文的完成，都是在 Mint Lab 裡教授、學長和同學的互相學習上完成的。感謝這個實驗室對我的幫助，也懷念在研究的過程中帶來的快樂。

這篇論文的完成，要感謝指導教授方凱田老師的教導，在研究的討論與寫作的問題中，老師總是會很細心的幫助我在研究中所遇到的問題，也能夠給我很大的空間在研究上發揮，讓我順利完成這篇論文。另外也要感謝研究助理許裕彬學長在研究上的帶領與學習，在我對研究方向和繁雜的數學推導中，都能給我很大的幫助，也在我遇到困難的情況中，給我信心去面對問題。祝福你在美國的博士生涯能三年內順利完成。而口試時，蘇育德老師和黃經堯老師也提供了研究中實現上的問題點與其他寶貴意見，能讓我這篇論文有更完整的改進，使我明白能藉由不同的觀點去思考研究問題。

在實驗室中，博士班的仲賢學長、建華學長、文俊學長和伯軒學長都曾在研究上給過我幫助，而更重要的是他們在研究上所表現的態度與專業更讓我明白研究生的模範表現是如何。而學弟其懋、萬邦、俊宇、承澤和惟能則是研究室的新生力軍，除了為實驗室帶來更歡樂的氣氛，也從他們身上看到他們對自己的研究所投注的努力與熱情。

同屆的同學瑞廷、俊傑和佳仕，兩年的時間中一起上課、討論、運動，帶來的快樂和互相幫助，是我研究生涯中愉快的時光。瑞廷的研究和我一樣是 CR，因此在彼此的討論中就有比較多的互動，外加在實驗室中的座位也在我旁邊，生活上也能更認識到他一些正面的特質，如流利的英語口說和清楚的語言表達，讓我有深刻的印象。俊傑的優秀表現，讓我體會到厲害的研究生也可以在研究與生活中取得平衡，而他對同學間的關懷也讓我感到很溫馨。而佳仕同學和他在修課方面有很多的接觸，讓我感受到他專業與邏輯上的特長，每次與他討論，都能學到清楚的觀念與知識。相信他在未來的博士生涯，也會很順利的畢業。心中很感謝也覺得很幸運有這一幫好同學，從你們身上學到太多太多東西讓我成長。

最後要感謝我的家人，在我離家學習的兩年在背後默默的給我支持，也給了我很大的空間讓我在學習上能夠更自由，期待未來的日子，盡我所能回報給最愛的家人。

王佳偉謹誌　于新竹國立交通大學

Contents

List of Figures

Chapter 1

Introduction

Due to rapid development of wireless systems, the demand for wireless spectrums has resulted in spectrum scarcity based on the conventional fixed allocation schemes. Even with the intensive usage of frequency spectrums, it has been studied by extensive measurements [1] that 62% of spectrum still remains unoccupied by the licensed primary user (PU). Cognitive radio (CR) is an intelligent wireless communication system that is perceptible to its surroundings. It is advanced as an emerging technology to effectively exploit the under-utilized spectrum in order to overcome the overcrowded spectrum problem.

There are two types of spectrum sharing that are defined for the CR networks (CRNs), including the underlay and the overlay waveforms. The underlay waveform represents that the unlicensed secondary users (SUs) are allowed to simultaneously share the primary frequency spectrum with the

PUs. The transmission power of the SUs are in general limited in order not to cause excessive interferences to the PUs. On the other hand, an overlay waveform allows the SUs to perform packet transmission under the existence of a spectrum hole. The spectrum hole is defined as a frequency band authorized to PUs, however, it is vacant at a particular time and geographic location. With the overlay waveform, the SUs can sense and identify the existence of spectrum hole for data communications. Therefore, spectrum utilization can be enhanced with these frequency-agile features. The research work in the CRNs has been investigated from various aspects. The work proposed in [2; 3] presents the techniques for spectrum sensing and detection; while [4; 5] investigate the spectrum allocation problem for the CR. There are also research [6; 7] focusing on the medium access control design for the CRNs.

Game theory [8] has been considered a feasible mathematical tool for solving the resource allocation problems in distributed CRNs. The fundamental concept of game theory is to resolve the conflict and cooperation between intelligent rational decision-makers (DMs). Instead of reaching a globally optimized solution based on identical objective, the DMs within the gaming formulation are seeking for solutions selfishly without the knowledge of other DMs' decisions. The primary reason is due to the inherent conflicts between the objectives that are assigned among the DMs, which can be adopted to model the behaviors of both PUs and SUs within the CRNs. After reaching the optimized solution (i.e. Nash equilibrium (NE) [8]) based on the game theory, each individual DM will not benefit from any action to deviate from the NE. In other words, by considering the conflicted interests between the DMs, the solutions obtained at the NE will provide every DM to possess the optimal resource allocation.

In general, two different types of games are categorized for the game theory, i.e. the strategic games and the extensive games. With the objective of reaching the NE, all DMs simultaneously select their strategies only for one-time by adopting the strategic games [8], which have been exploited to resolve the power control problem for the CRNs in recent research work [9; 10]. The work in [9] proposed an algorithm for distributed multi-channel power allocation based on the strategic gaming model;while the pricing-based games are utilized in [10] to achieve a higher signal-to-noise ratio with the guarantee of reliable data transmissions. However, computation of NEs in strategic game will introduce some computation time at each time.

On the other hand, the extensive games [8; 11; 12] represent a class of gaming models where the DMs repeatedly conduct decision-making numerous times for resource allocation. Unlike the strategic games that each DM considers his strategy only at the beginning of the game, the extensive games is implemented whenever a decision has to be made in order to increases the spectrum efficiency by the multi-stage gaming model. The scheme proposed in [12] utilized the repeated game to solve NE point under underlay waveform. But it can't character the variation of CRN environment. In addition, constrained stochastic games [13; 14] are formulated by extending the extensive games for dynamically-changing environments with the consideration of certain constraints for optimization. It can be considered as an extension of the Markov decision process from a single DM to multiple DMs. The power allocation algorithm proposed in [15] imposes both the power and the buffer length constraints under the environments with varying channel states. It is noticed that only independent states between the DMs are considered in [15], i.e. the states of power and buffer length for each DM is independent to those from other DMs. So, constrained stochastic games can be applied to the resource management problems for CRNs.

In this paper, the constrained stochastic games are adopted and extended to study the dynamic power management problem in CRNs. The dynamic environments occurred from the channel variations and the uncertain spectrum holes will be modeled as the ergodic Markov decision process. It is noticed that the spectrum holes are considered the dependent states for each SU since the SUs are sharing to utilize the spectrum holes while the original licensed PU is temporarily releasing the frequency band. Moreover, each SU can perceive its own current state but is unaware of the states and strategies from the other SUs. As the licensed spectrum is occupied by the PUs, the underlay waveform is executed by the SUs with the introduction of reasonable interferences to the PUs. On the other hand, the SUs will share the spectrum hole with the overlay waveform as the primary traffic is absent. Constraints for allowable interferences will also be imposed to preserve the communication quality among the SUs under the existence of spectrum holes. With the satisfaction of the defined constraints, the constrained NE suggests an optimal solution to the dynamic power assignment according to the SUs' current state within the CRNs. In finite and infinite time horizon, i.e.time non-converge and converge to stable point respectively , existence of constrained NE will be proved.

Therefore, considering all of the issues mentioned above, two stochastic game are proposed in this paper to describe the finite time and infinite time horizon respectively in the CRNs. Similar Dynamic programming method will prove the existence of constrained NE in finite time horizon. Using the stable property of CRNs the existence of constrained NE will be demonstrated in infinite time horizon.

The rest of this paper is organized as follows. chapter 2 presents the system models of finite and infinite time horizon of CRNs. The corresponding proofs for the existence of constrained Nash equilibrium are provided in chapter 3 and chapter 4 respectively. Numerical evaluation is performed in chapter 5; while chapter 6 draws the conclusions.

Chapter 2

System Model for Dynamic Power Management with Constrained Stochastic Games

The schematic diagram of the CRN is illustrated in Fig. 2.1, where a synchronous slotted time structure is considered. A PU is communicating with its primary base station; while there exists $N = 2$ SU pairs where SU(Tx) is intending to transmit its data packets to the respective SU(Rx) within the same frequency spectrum as the PU. The overlay waveform is shown at the time slot 2 where a spectrum hole happens for the SUs to share the licensed band without the existence of the PU. At both time slots 1 and 3, with tolerable interferences to the PU, the SUs coexist with the PU to conducts their transmissions under the execution of the underlay waveform.

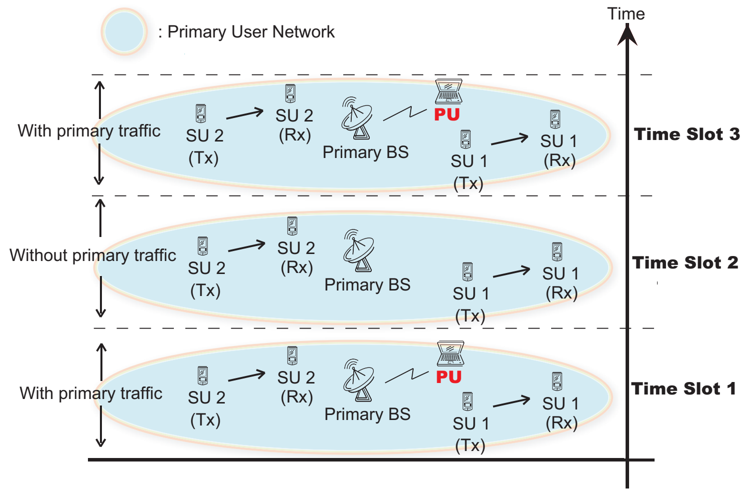

Figure 2.1: The schematic diagram of the cognitive radio network for dynamic power management. (Tx : transmitter , Rx : receiver)

At each time slot t, each SU(Tx) i forwards its data packets with a specific power level $p_i^t \in \boldsymbol{p}_i \triangleq \{p_{i,0}, p_{i,1}, \cdots, p_{i,\max}\}$, which is referred as the action set in the game theory. The global set of the power level for the entire CRN is denoted as $\mathbf{P} = \prod_{i=1}^{N} \boldsymbol{p}_i$. The dynamic environment in CRN is modeled as an ergodic Markov chain [16], where feedback information is considered available for each SU pair, i.e. from SU(Rx) to SU(Tx). In other words, each SU(Tx) will possess the information about all the current states that are detected by its corresponding SU(Rx). The compound state s_i^t of each SU i at the time slot t is constructed by two elements ϕ_i^t and g_i^t, i.e. $s_i^t = (\phi_i^t, g_i^t)$. The parameter $\phi_i^t \in \boldsymbol{\phi}_i \triangleq \{0, 1\}$ is utilized to denoted the status of the PU, where $\phi_i^t = 0$ indicates the absence of the primary traffic, and $\phi_i^t = 1$ represents the existence of the PU within the CRN. It is noted that, at each time slot t, the indication of the primary traffic ϕ_i^t is considered equal for all the SUs i that share the licensed spectrum. Therefore, the global space can be obtained as $\mathbf{\Phi} = \prod_i^N \boldsymbol{\phi}_i = \{\alpha, \cdots, \alpha\}$, where $\mathbf{\Phi}$ has N elements with $\alpha \in \{0, 1\}$. Moreover, the state of the channel gain for each SU i at time slot t is denoted by the index $g_i^t \in \boldsymbol{g}_i \triangleq \{0, \cdots, L_i - 1\}$. The compound state s_i^t will therefore belong to the set $\boldsymbol{s}_i = \boldsymbol{\phi}_i \times \boldsymbol{g}_i$ with the length of state vector equal to $2L_i$. The global state space of s_i^t considering all the N SUs can also be represented as $\mathbf{S} = \prod_{i=1}^N \boldsymbol{s}_i$. The immediate utility of SU i is defined as r_i which is a function of $(\boldsymbol{s}^t, \boldsymbol{p}^t)$. Furthermore, $P_{xy}^i = \mathcal{M}(s_i^{t+1} = y | s_i^t = x)$ is utilized to express the state transition probability, where $\mathcal{M}(\varepsilon)$ is the probability measure over an event ε.

A history at time epoch t of SU i is a time sequence of its current state as well as its previous states and actions, which is denoted as $\boldsymbol{h}_i^t = (s_i^0, p_i^0, s_i^1, p_i^1, \cdots, s_i^{t-1}, p_i^{t-1}, s_i^t)$ with $s_i^k \in \boldsymbol{s}_i$ and $p_i^k \in \boldsymbol{p}_i$. Let $\mathbf{H}_i^t$ be the collection of all possible histories of length t for SU i. A policy employed by SU i can be denoted as a sequence $\boldsymbol{u}_i = (u_i^0, u_i^1, \cdots, u_i^t)$, where $u_i^t : \mathbf{H}_i^t \to \mathcal{M}(\boldsymbol{p}_i)$ is a function mapping from the histories to the probability measure over the action sets of SU i. The elements within the policy u_i^t indicate the occurring probabilities for their corresponding power level $p_{i,j}$ for $j = 0$ to max. It is noted that the decision of the policy u_i^t for each SU is independent to that for the other SUs. The set of all reasonable policies for SU i is in the policy space $\mathbf{U}_i$, i.e. $\boldsymbol{u}_i \in \mathbf{U}_i$. Therefore, with the consideration of all the N SUs, the global policy space $\mathbf{U} = \prod_{i=1}^N \mathbf{U}_i$ is called the class of multi-policies. In addition, the multi-policy except SU i is defined as $\boldsymbol{u}_{-i} = (\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_{i-1}, \boldsymbol{u}_{i+1}, \cdots, \boldsymbol{u}_N) \in \boldsymbol{U}_{-i}$. Moreover, the stationary policies are characterized as the policy that is independent of the histories, i.e. $u_i^t : \boldsymbol{s}_i \to \mathcal{M}(\boldsymbol{p}_i)$ as a function mapping only from the current state $\boldsymbol{s}_i$. The union of all possible stationary policies is denoted as $\mathbf{U}_i^S \in \mathbf{U}_i$, and $\mathbf{U}^S = \prod_{i=1}^N \mathbf{U}_i^S \in \mathbf{U}$ represents the class of stationary multi-policies.

2.1 Finite Time Horizon

The expected utility of SU i with the policy $\boldsymbol{u} = (\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_N) \in \mathbf{U}$ and the initial state $\boldsymbol{s}^0 = (s_1^0, s_2^0, \cdots, s_N^0) \in \mathbf{S}$ can be obtained as

$$R_i^T(\boldsymbol{s}^0, \boldsymbol{u}) = \frac{1}{T} \sum_{t=0}^{T-1} E_{\boldsymbol{s}^0}^{\boldsymbol{u}} \Big[r_i(\boldsymbol{s}^t, \boldsymbol{p}^t) \Big] \tag{2.1}$$

where $E_{\boldsymbol{s}^0}^{\boldsymbol{u}}$ is the operator for the computation of expectation value. Furthermore, the allowable interferences between the SUs and the PU are considered in order to guarantee the quality of service (QoS) of the CRN. The supreme expected allowable interference at the SU i(Rx) is obtained as

$$I_{i,m}^T(\boldsymbol{s}^0, \boldsymbol{u}) = \frac{1}{\sum_{t=0}^{T-1} E_{s^0} \Big[\delta_0(\phi_i^t) \Big]} \sum_{t=0}^{T-1} \sum_{\substack{k=1 \\ k \neq m}}^{N} E_{\boldsymbol{s}^0}^{\boldsymbol{u}} \Big[p_k^t \cdot \nu_{km}(s_k^t) \cdot \delta_0(\phi_k^t) \Big] \quad \forall m \neq i \tag{2.2}$$

and

$$I_p^T(\boldsymbol{s}^0, \boldsymbol{u}) \;=\; \frac{1}{\sum_{t=0}^{T-1} E_{s^0}\Big[\delta_1(\phi_i^t)\Big]} \sum_{t=0}^{T-1}\sum_{k=1}^{N} E_{\boldsymbol{s}^0}^{\boldsymbol{u}}\Big[p_k^t \cdot \nu_{kp}(s_k^t) \cdot \delta_1(\phi_k^t)\Big] \forall p \in \{1, \cdots, M\} \tag{2.3}$$

where δ is the Kronecker delta function. The function $\nu_{km}(s_k^t)$ and $\nu_{kp}(s_k^t)$ represent the corresponding channel gains from SU j(Tx) to SU i(Rx) and SU j(Tx) to PU in state s_k^t respectively. In (2.7), $I_{i,m}^T(\boldsymbol{s}^0, \boldsymbol{u})$ indicates the case with the absence of primary traffic, i.e. $\delta_0(\phi_i^t = 0) = 1$; while $I_p^T(\boldsymbol{s}^0, \boldsymbol{u})$ denotes the case with primary traffic, i.e. $\delta_1(\phi_i^t = 1) = 1$. Under the usage of licensed band from PU, the influence occurred from the SUs is confined by $I_p^T(\boldsymbol{s}^0, \boldsymbol{u}) \leq C_1$ to assure the QoS of the PU, where C_1 denotes the the PU's tolerable interference. Considering the case without the primary traffic, the allowable interference between the SUs are constrained by $I_{i,m}^T(\boldsymbol{s}^0, \boldsymbol{u}) \leq C_0$, where C_0 indicates the QoS constraint among the SUs that share the common spectrum band. Therefore, the set of feasible policies can be defined as $\boldsymbol{u} \in \mathbf{U}$ in order to satisfy the condition $I_{i,m}^T(\boldsymbol{s}^0, \boldsymbol{u}) \leq C_0 \;\forall m \neq i$ and $I_p^T(\boldsymbol{s}^0, \boldsymbol{u}) \leq C_1 \;\forall p$.

Definition 1. A multi-policy $\boldsymbol{u}^* = (\boldsymbol{u}_1^*, \boldsymbol{u}_2^*, \cdots, \boldsymbol{u}_N^*) \in \mathbf{U}$ is a constrained Nash equilibrium (CNE) if it is a feasible policy such that for all SUs i

$$R_i^T(\boldsymbol{s}^0, \boldsymbol{u}^*) \geq R_i^T(\boldsymbol{s}^0, [\boldsymbol{u}_{-i}^*|\boldsymbol{v}_i]) \tag{2.4}$$

for any feasible policies $[\boldsymbol{u}^*_{-i}|\boldsymbol{v}_i]$, where the policy $[\boldsymbol{u}^*_{-i}|\boldsymbol{v}_i]$ means that SU i uses the policy $\boldsymbol{v}_i$ while other SUs $k \neq i$ takes the policy $\boldsymbol{u}^*_k$.

2.2 Infinite Time Horizon

The expected utility of SU i is

$$R_i(\boldsymbol{s}^0, \boldsymbol{u}) = \lim_{T\to\infty} sup \frac{1}{T} \sum_{t=0}^{T-1} E^{\boldsymbol{u}}_{\boldsymbol{s}^0}\left[r_i(\boldsymbol{s}^t, \boldsymbol{p}^t)\right] \tag{2.5}$$

The expected allowable interference at the SU i(Rx) are

$$I_{i,m}(\boldsymbol{s}^0, \boldsymbol{u}) = \lim_{T\to\infty} sup \frac{1}{\sum_{t=0}^{T-1} E_{s^0}\left[\delta_0(\phi_i^t)\right]} \cdot \\ \sum_{t=0}^{T-1} \sum_{\substack{k=1 \\ k\neq m}}^{N} E^{\boldsymbol{u}}_{\boldsymbol{s}^0}\left[p_k^t \cdot \nu_{km}(s_k^t) \cdot \delta_0(\phi_k^t)\right] \quad \forall m \neq i \tag{2.6}$$

and

$$I_p(\boldsymbol{s}^0, \boldsymbol{u}) = \lim_{T\to\infty} sup \frac{1}{\sum_{t=0}^{T-1} E_{s^0}\left[\delta_1(\phi_i^t)\right]} \cdot \\ \sum_{t=0}^{T-1} \sum_{k=1}^{N} E^{\boldsymbol{u}}_{\boldsymbol{s}^0}\left[p_k^t \cdot \nu_{kp}(s_k^t) \cdot \delta_1(\phi_k^t)\right] \quad \forall p \in \{1, \cdots, M\} \tag{2.7}$$

A multi-policy $\boldsymbol{u}^* = (\boldsymbol{u}^*_1, \boldsymbol{u}^*_2, \cdots, \boldsymbol{u}^*_N) \in \mathbf{U}$ is a CNE in infinite time horizon if it is a feasible policy such that for all SUs i

$$R_i(\boldsymbol{s}^0, \boldsymbol{u}^*) \geq R_i(\boldsymbol{s}^0, [\boldsymbol{u}^*_{-i}|\boldsymbol{v}_i]) \tag{2.8}$$

It is considered that the SUs are rational [8] such that all SUs are intending to maximize their corresponding utilities in (2.5). Furthermore, the decision for each SU i to transmit packets with the power level p_i^t at the beginning of time slot t is determined without additional knowledge about the states and actions from the other SUs. As a result, the constrained Nash equilibrium (CNE) [14] will be utilized to facilitate the power management problem from the perspective of game theory, which is defined as follows.

The purpose of this paper is to provide the mechanism for dynamic power management based on the optimal polices that are derived from the CNE. The existence of CNE for the finite and infinite time horizon problems will be acquired in chapter III and IV respectively.

Chapter 3

Existence of CNE for Finite Time Horizon Stochastic Game

In this chapter, the constrained optimization problem with finite time horizon considering a single SU will be introduced in Problem 1. The Markov strategy which will be defined in Definition 2 is also a CNE. The similar dynamic programming method will prove existence of CNE from time slot $T-1$ to 0 sequentially.

3.1 Expected Utility and Markov Strategy

The expected utility of SU i when deciding in time slot t is

$$\frac{1}{T-t} \left. \quad r_i(\boldsymbol{s}^t, \boldsymbol{p}^t) + \sum_{s^{t+1}} V_{i(t+1)}(s^{t+1}) P_{s^t s^{t+1}} \right) \tag{3.1}$$

where s^t is the state that occurs in time slot t and $P_{s^t s^{t+1}}$ is the state transition probability from s^t to s^{t+1}. $V_{i(t+1)}(s^{t+1})$ is the utility that SU i expects to receive in the future starting from time t+1.

Problem 1 (Constrained Optimization Problem (COP) With Finite Time Horizon)**.** Given a fixed set of policies $\boldsymbol{u}_{-i} \in \mathbf{U}_{-i}$, find an optimal policy $\boldsymbol{v}_i^*$ for SU i in order to maximize the expected utility

$$R_i^T(\boldsymbol{s}^0, [\boldsymbol{u}_{-i}|\boldsymbol{v}_i]) \tag{3.2}$$

subject to

$$I_{i,m}^T(\boldsymbol{s}^0, [\boldsymbol{u}_{-i}|\boldsymbol{v}_i]) \leq C_0 \quad \forall m \neq i \tag{3.3}$$

and

$$I_p^T(\boldsymbol{s}^0, [\boldsymbol{u}_{-i}|\boldsymbol{v}_i]) \leq C_1 \quad \forall p \in \{1, \cdots, M\} \tag{3.4}$$

For a COP with finite time horizon with terminal time T expected utility of SU i from the strategy combination u is given by

where the first term on the right-hand side of the equation is the expected utility using the strategy u in time slot 0, the second term is the expected utility from using the strategy u in time slot 0 and 1 and so on till the last term which is the expected utility in time slot T-1 when using the strategy

u throughout the game. Next, defined a special strategies, namely Markov strategies.

Definition 2. A Markov strategy for SU i denoted by $u_{i,Mar}$ is a sequence $\{u^t_{i,Mar}\}^T_{t=0}$ such that $u^t_{i,Mar} : s^t_i \rightarrow \mathcal{M}(\boldsymbol{p}_i)$ is measurable for every t. A Markov strategy combination u_{Mar} is a combination of Markov strategies.

Since Markov strategies restrict SUs to make their decisions conditional only on the current self state, this can be a fairly severe restriction on the kind of strategies SUs can use. However, with the assumptions of Markovian nature of transition probabilities, a SU can do just as well by using a Markov strategy. This is so because the current and future utility of a SU is given by (3.1). If every SUs uses a Markov strategy then the optimal p_i for SU i given the current state s_i is optimal no matter what the past history. That is, if every SUs uses a Markov strategy, then an optimal Markov strategy of SU i in time slot t is an optimal strategy. This thus means that if an equilibrium in Markov strategies is found then we have obtained an equilibrium.

Definition 3. A Markov strategy for SU i denoted by $u^*_{i,Mar}$ is an equilibrium if for any s_i in any time slot and for any SU i

$$V_{it}(s_i|u^*_{i,Mar}) \geq V_{it}(s_i|u_i) \tag{3.5}$$

3.2 Existence of CNE

Based on an backward recursion argument, we show the proof that can be used to construct equilibria in COP with finite time horizon.

Theorem 1. *There exists a Markov strategy $\boldsymbol{u}_{Mar} \in \mathbf{U}$ as the CNE for dynamic power management problem of the considered CRN in finite time horizon.*

Proof. At time slot T-1, given the state s_i^{T-1}, the expected utility of SU i from time T-1 to T-1 is denoted as follows

$$\begin{aligned}
& E_{s_i^{T-1}}[r_i(s^{T-1}, p^{T-1})] \\
= & \sum_{p_i^{T-1}} \sum_{s_{-i}^{T-1}} \sum_{p_{-i}^{T-1}} r_i(s,p) u_i^{T-1}(p_i^{T-1} = p_i | s_i^{T-1} = s_i) \cdot \prod_{j \neq i} \frac{u_j^{T-1}(p_j^{T-1}|s_j^{T-1})\pi_{s_j}}{\sum_{\substack{s_k \\ \phi_k=\phi_i}} u_j^{T-1}(p_k|s_k)\pi_{s_k}} \\
= & \sum_{p_i^{T-1}} \left(\sum_{s_{-i}^{T-1}} \sum_{p_{-i}^{T-1}} r_i(s,p) \cdot \prod_{j \neq i} \frac{u_j^{T-1}(p_j^{T-1}|s_j^{T-1})\pi_{s_j}}{\sum_{\substack{s_k \\ \phi_k=\phi_i}} u_j^{T-1}(p_k|s_k)\pi_{s_k}} \right) \cdot u_i^{T-1}(p_i^{T-1} = p_i | s_i^{T-1} = s_i)
\end{aligned} \tag{3.6}$$

which is a strategic game. Besides, the expected interference without PU traffic can be described as

$$E_{s_i^{T-1}}[\sum_{\substack{k=1 \\ k \neq m}}^{N} p_k^{T-1} v_{km}(s_k^{T-1})\delta_0(\phi_k^{T-1})] = \sum_{s_{-i}^{T-1}} \sum_{p_{-i}^{T-1}} p_k^{T-1} v_{km}(s_k^{T-1})\delta_0(\phi_k^{T-1}) \cdot$$
$$\prod_{j \neq i} \frac{u_j^{T-1}(p_j^{T-1}|s_j^{T-1})\pi_{s_j}}{\sum_{\substack{s_k \\ \phi_k=\phi_i}} u_j^{T-1}(p_k|s_k)\pi_{s_k}} + \sum_{p_i^{T-1}} p_i v_{im}(s_i) u_i^{T-1}(p_i^{T-1} = p_i | s_i^{T-1} = s_i) \leq C_0 \tag{3.7}$$

By the same procedure, the expected interference with PU traffic can be depicted as

$$E_{s_i^{T-1}}[\sum_{k=1}^{N} p_k^{T-1} v_{kp}(s_k^{T-1})\delta_1(\phi_k^{T-1})] = \sum_{s_{-i}^{T-1}} \sum_{p_{-i}^{T-1}} p_k^{T-1} v_{kp}(s_k^{T-1})\delta_1(\phi_k^{T-1}) \cdot$$
$$\prod_{j\neq i} \frac{u_j^{T-1}(p_j^{T-1}|s_j^{T-1})\pi_{s_j}}{\sum_{\phi_k=\phi_i}^{s_k} u_j^{T-1}(p_k|s_k)\pi_{s_k}} + \sum_{p_i^{T-1}} p_i v_{ip}(s_i) u_i^{T-1}(p_i^{T-1} = p_i|s_i^{T-1} = s_i) \leq C_1 \quad (3.8)$$

According to equation (3.7) and (3.8), the policy set of SU i is nonempty, compact and convex set at time slot T-1. Because of equation (3.6), the expected utility function is both continuous and quasi-concave in its policy. So, there exits a CNE at time slot T-1.

At time slot $T-2$, given the state s_i^{T-2} and ${u^*}^{T-1}$, the expected utility of SU i from time $T-2$ to $T-1$ is denoted as follows

$$\begin{aligned}
&E_{s_i^{T-2}}[r_i(s^{T-2}, p^{T-2}) + r_i(s^{T-1}, p^{T-1})] \\
= &\sum_{p_i^{T-2}} \sum_{s_{-i}^{T-2}} \sum_{p_{-i}^{T-2}} r_i(s,p) u_i^{T-2}(p_i^{T-2} = p_i|s_i^{T-2} = s_i) \cdot \prod_{j\neq i} \frac{u_j^{T-2}(p_j^{T-2}|s_j^{T-2})\pi_{s_j}}{\sum_{\phi_k=\phi_i}^{s_k} u_j^{T-2}(p_k|s_k)\pi_{s_k}} \\
+ &\sum_{s_{-i}^{T-1}} \left(\sum_{p_i^{T-1}} \sum_{s_{-i}^{T-1}} \sum_{p_{-i}^{T-1}} r_i(s,p) {u_i^*}^{T-1}(p_i^{T-1} = p_i|s_i^{T-1} = s_i) \cdot \right. \\
&\left. \prod_{j\neq i} \frac{u_j^{T-1^*}(p_j^{T-1}|s_j^{T-1})\pi_{s_j}}{\sum_{\phi_k=\phi_i}^{s_k} u_j^{T-1^*}(p_k|s_k)\pi_{s_k}} \right) P_{s_i^{T-2} s_i^{T-1}} \qquad (3.9)
\end{aligned}$$

where the last term is a constant. It's also a strategic game. The same procedure as equation (3.7), (3.8) and (3.9), we can obtain that there exists a CNE at time slot T-2. So, by this recursion procedure we prove that there exist a CNE in finite time horizon stochastic game.

□

Chapter 4

Existence of CNE for Infinite Time Horizon Stochastic Game

In this chapter, the constrained optimization problem for dynamic power management considering a single SU will first be introduced in Problem 2. The linear programming methodology as formulated in Problem 3 will be associated with Problem 2 based on the proofs in Lemmas 1 to 3. Consequently, the dynamic power management problem as defined in Definition 1 will be proved in Theorem 2 for the entire N SUs in the CRN. Consider fixed policies for the other SUs, a constrained optimization problem for a single SU can be formulated to obtain the best response [8] as follows.

Problem 2 (Constrained Optimization Problem (COP))**.** Given a fixed set of policies $\boldsymbol{u}_{-i} \in \mathbf{U}_{-i}$, find an optimal policy $\boldsymbol{v}_i^*$ for SU i in order to maximize the expected utility

$$R_i(\boldsymbol{s}^0, [\boldsymbol{u}_{-i}|\boldsymbol{v}_i]) \tag{4.1}$$

subject to

$$I_{i,m}(\boldsymbol{s}^0, \boldsymbol{u}) \leq C_0 \quad \forall m \neq i \tag{4.2}$$

$$I_p(\boldsymbol{s}^0, \boldsymbol{u}) \leq C_1 \quad \forall p \in \{1, \cdots, M\} \tag{4.3}$$

Therefore, a CNE multi-policy $\boldsymbol{u}^* \in \mathbf{U}$ in Definition 1 can be verified while $\boldsymbol{u}_i^*$ represents the optimal policy in Problem 1 for all SU i providing other SUs take the policies $\boldsymbol{u}_{-i}^*$. In order to resolve Problem 2, the defined COP can be correlated with a linear programming problem by extending from the previous studies [14; 17; 18]. A linear programming problem is defined as follows.

Problem 3 (Linear Programming (LP) problem)**.** Consider a set of state-action pairs for SU i characterized by $\mathbf{K}_i = \{(s_i, p_i) : s_i \in \mathbf{S}_i, p_i \in \mathbf{P}_i\}$ as well as $\mathbf{K} = \prod_i^N \mathbf{K}_i$ and $\mathbf{K}_{-i} = \prod_{j \neq i}^N \mathbf{K}_i$. Given a set of stationary policies

$\boldsymbol{u}_{-i} \in \mathbf{U}_{-i}^{S}$, find $\boldsymbol{z}_{i,\boldsymbol{u}_{-i}}^{*} = \{z_{i,\boldsymbol{u}_{-i}}^{*}(s_i, p_i) : (s_i, p_i) \in \mathbf{K}_i\}$ which maximizes

$$\mathscr{R}_i(\boldsymbol{z}_{i,\boldsymbol{u}_{-i}}) = \sum_{(s_i,p_i)\in\mathbf{K}_i} \mathcal{R}_{i,\boldsymbol{u}_{-i}}(s_i, p_i) \cdot z_{i,\boldsymbol{u}_{-i}}(s_i, p_i) \tag{4.4}$$

subject to

$$\mathscr{I}_{i,m}(\boldsymbol{z}_{i,\boldsymbol{u}_{-i}}) = \sum_{\substack{(s_i,p_i)\in\mathbf{K}_i \\ \phi_i = j}} \mathcal{I}_{i,\boldsymbol{u}_{-i}}(s_i, p_i) \frac{z_{i,\boldsymbol{u}_{-i}}(s_i, p_i)}{\boldsymbol{Z}_{i,j}} \le C_0 \ \forall m \neq i \tag{4.5}$$

$$\mathscr{I}_{p}(\boldsymbol{z}_{i,\boldsymbol{u}_{-i}}) = \sum_{\substack{(s_i,p_i)\in\mathbf{K}_i \\ \phi_i = j}} \mathcal{I}_{i,\boldsymbol{u}_{-i}}(s_i, p_i) \frac{z_{i,\boldsymbol{u}_{-i}}(s_i, p_i)}{\boldsymbol{Z}_{i,j}} \le C_1 \ \forall p \in \{1, \dots, M\} \tag{4.6}$$

$$\sum_{(s_i,p_i)\in\mathbf{K}_i} z_{i,\boldsymbol{u}_{-i}}(s_i, p_i) \left[\delta_{r_i}(s_i) - P_{s_i r_i}^{i}\right] = 0 \quad \forall r_i \in \mathbf{S}_i \tag{4.7}$$

$$\sum_{(s_i,p_i)\in\mathbf{K}_i} z_{i,\boldsymbol{u}_{-i}}(s_i, p_i) = 1 \tag{4.8}$$

$$z_{i,\boldsymbol{u}_{-i}}(s_i, p_i) \ge 0 \quad \forall (s_i, p_i) \in \mathbf{K}_i \tag{4.9}$$

where $P_{s_i r_i}^{i}$ in (4.7) is the transition probability from state s_i to r_i for SU i. The value of $\delta_{r_i}(s_i)$ in (4.7) is equal to 1 as the state $s_i = r_i$, otherwise $\delta_{r_i}(s_i) = 0$. The denominator $\boldsymbol{Z}_{i,j}$ in (4.5) is utilized for normalization purpose as

$$\boldsymbol{Z}_{i,j} = \sum_{\substack{(s_k,p_k)\in\mathbf{K}_i\\ \phi_k=j}} z_{i,\boldsymbol{u}_{-i}}(s_k,p_k) \tag{4.10}$$

The functions $\mathcal{R}_{i,\boldsymbol{u}_{-i}}(s_i,p_i)$ in (4.4) and $\mathcal{I}_{i,\boldsymbol{u}_{-i}}(s_i,p_i)$ in (4.5) are the expected immediate utility and the allowable interference while SU i executes the power level p_i at the state s_i under the case that the other SUs are adopting the policy $\boldsymbol{u}_{-i}$. Both functions can be expressed as

$$\mathcal{R}_{i,\boldsymbol{u}_{-i}}(s_i,p_i) = \sum_{\substack{(s,p)_{-i}\in\mathbf{K}_{-i},\\ \phi_k=\phi_i,\forall k\neq i}} \prod_{m\neq i} \Omega_{i,m}\cdot r_i(\boldsymbol{s},\boldsymbol{p}) \tag{4.11}$$

$$\mathcal{I}_{i,\boldsymbol{u}_{-i}}(s_i,p_i) = \left. \sum_{\substack{(s,p)_{-i}\in\mathbf{K}_{-i},\\ \phi_k=\phi_i,\forall k\neq i}} \prod_{m\neq i} \Omega_{i,m} \sum_{k=1}^{N} p_k\, \nu_{ki}(s_k)\right) \tag{4.12}$$

where $\Omega_{i,m}$ corresponds to the probability of the state-action pair (s_m,p_m) for SU m. Let the stationary distribution of the state s_m for SU m be $\pi_m(s_m)$, $\Omega_{i,m}$ can be computed as

$$\Omega_{i,m} = \frac{u_m(p_m|s_m)\pi_m(s_m)}{\sum_{\substack{(s_k,p_k)\in\mathbf{K}_m,\\ \phi_k=\phi_i}} u_m(p_k|s_k)\pi_m(s_k)} \tag{4.13}$$

where $u_m(p_m|s_m)$ denotes the probability measure for SU m to conduct action p_m based on the state s_m. The normalized term in the denominator of (4.13) is utilized to indicate that common spectrum among all the SUs will result in the correlation among the states of each SU, i.e. $\phi_m = \phi_i$ for all $m \neq i$.

A set of nonnegative real numbers is defined as $\boldsymbol{\omega}_i = \{\omega_i(s_i, p_i) : (s_i, p_i) \in \mathbf{K}_i\}$. The probability $\boldsymbol{\gamma}_i(\boldsymbol{\omega}_i) = \{\gamma_{s_i}^{p_i}(\boldsymbol{\omega}_i) : (s_i, p_i) \in \mathbf{K}_i\}$ can be define as $\gamma_{s_i}^{p_i}(\boldsymbol{\omega}_i) = \omega_i(s_i, p_i)/\sum_{p_k} \omega_k(s_k, p_k)$ in the case that $\sum_{p_k} \omega_k(s_k, p_k) \neq 0$. Otherwise, an arbitrary value is assigned to $\gamma_{s_i}^{p_i}(\boldsymbol{\omega}_i)$ such that $\sum_{p_k} \gamma_{s_i}^{p_i}(\boldsymbol{\omega}_i) = 1$. The parameter $\boldsymbol{\lambda}_i(\boldsymbol{\omega}_i)$ represents a set of stationary policies for SU i that selects its power level p_i at the state s_i with the probability $\gamma_{s_i}^{p_i}(\boldsymbol{\omega}_i)$. Furthermore, $f_i(s_i^0, \boldsymbol{u}_i; s_i, p_i)$ is denoted as the limiting point of the time sequence $\{f_i^t(s_i^0, \boldsymbol{u}_i; s_i, p_i)\}_t$. The expected state-action frequency $f_i^t(s_i^0, \boldsymbol{u}_i; s_i, p_i)$ [18] for SU i at time t can be obtained as

$$f_i^t(s_i^0, \boldsymbol{u}_i; s_i, p_i) = \frac{1}{t} \sum_{k=0}^{t-1} P_{s_i^0}^{\boldsymbol{u}_i}(s_i^k = s_i, p_i^k = p_i) \tag{4.14}$$

where $P_{s_i^0}^{\boldsymbol{u}_i}(\varepsilon)$ is the the probability measure over the event ε with the policy $\boldsymbol{u}_i$ and the initial state s_i^0. Based on the definition of the state-action frequency, the relationship between the COP and the LP problem can be constructed as follows.

Lemma 1. *Given a set of stationary policies $\boldsymbol{u}_{-i} \in \mathbf{U}_{-i}^S$, for any $\boldsymbol{z}_{i,\boldsymbol{u}_{-i}}$ that satisfies (4.7) to (4.9) will result in $\mathscr{R}_{i,j}(\boldsymbol{z}_{i,\boldsymbol{u}_{-i}}) = R_i(\boldsymbol{s}^0, [\boldsymbol{u}_{-i}|\boldsymbol{\lambda}_i(\boldsymbol{z}_{i,\boldsymbol{u}_{-i}})])$ for SU i.*

Proof. Based on the definition of $R_i(\boldsymbol{s}^0, \boldsymbol{u})$ in (2.5), the following equation can be obtained:

$$R_i(\boldsymbol{s}^0, \boldsymbol{u}) = \lim_{T\to\infty} sup \frac{1}{T} \sum_{t=0}^{T-1} E_{\boldsymbol{s}^0}^{\boldsymbol{u}} \left[r_i(\boldsymbol{s}^t, \boldsymbol{p}^t) \right] \tag{4.15}$$

$$= \lim_{T\to\infty} sup \frac{1}{T} \sum_{t=0}^{T-1} \sum_{(s_i,p_i)\in \mathbf{K}_i} \sum_{\substack{(s,p)_{-i}\in \mathbf{K}_{-i} \\ \phi_l=\phi_i, \forall l\neq i}} r_i(\boldsymbol{s}, \boldsymbol{p}) \cdot$$

$$P_{s_i^0}^{\boldsymbol{u}_i}(s_i^t = s_i, p_i^t = p_i) \prod_{j\neq i} \frac{P_{s_j^0}^{\boldsymbol{u}_j}(s_j^t = s_j, p_j^t = p_j)}{\sum_{\substack{(s_k,p_k)\in \mathbf{K}_j \\ \phi_k=\phi_i}} P_{s_j^0}^{\boldsymbol{u}_j}(s_j^t = s_k, p_j^t = p_k)} \tag{4.16}$$

$$= \sum_{(s_i,p_i)\in \mathbf{K}_i} f_i(s_i^0, \boldsymbol{u}_i; s_i, p_i) \cdot$$

$$\left[\sum_{\substack{(s,p)_{-i}\in \mathbf{K}_{-i} \\ \phi_l=\phi_i, \forall l\neq i}} r_i(\boldsymbol{s}, \boldsymbol{p}) \prod_{j\neq i} \frac{f_j(s_j^0, \boldsymbol{u}_j; s_j, p_j)}{\sum_{\substack{(s_k,p_k)\in \mathbf{K}_j \\ \phi_k=\phi_i}} f_j(s_j^0, \boldsymbol{u}_j; s_k, p_k)} \right] \tag{4.17}$$

$$= \sum_{(s_i,p_i)\in \mathbf{K}_i} f_i(s_i^0, \boldsymbol{u}_i; s_i, p_i) \cdot \mathcal{R}_{i,\boldsymbol{u}_{-i}}(s_i, p_i) \tag{4.18}$$

It is noted that the equality from (4.16) to (4.17) is mainly due to the assumption of stationary multi-policy. By substituting $\boldsymbol{u}_i$ in (4.18) with $\boldsymbol{\lambda}_i(\boldsymbol{z}_{i,\boldsymbol{u}_{-i}})$, it can be obtained that $f_i(s_i^0, \boldsymbol{\lambda}_i(\boldsymbol{z}_{i,\boldsymbol{u}_{-i}}); s_i, p_i) = z_{i,\boldsymbol{u}_{-i}}(s_i, p_i)$. The relationship between (4.1) and (4.4) can therefore be established, which completes the proof. □

Lemma 2. *Given a set of stationary policies* $\boldsymbol{u}_{-i} \in \mathbf{U}_{-i}^S$. *By choosing* $\boldsymbol{z}_{i,\boldsymbol{u}_{-i}}$ *based on (4.7) to (4.9), the following relationship can be obtained:* $\mathscr{I}_{i,m}(\boldsymbol{z}_{i,\boldsymbol{u}_{-i}}) = I_{i,m}(\boldsymbol{s}^0, [\boldsymbol{u}_{-i}|\boldsymbol{\lambda}_i(\boldsymbol{z}_{i,\boldsymbol{u}_{-i}})])$ *and* $\mathscr{I}_p(\boldsymbol{z}_{i,\boldsymbol{u}_{-i}}) = I_p(\boldsymbol{s}^0, [\boldsymbol{u}_{-i}|\boldsymbol{\lambda}_i(\boldsymbol{z}_{i,\boldsymbol{u}_{-i}})])$. *Moreover,* $\boldsymbol{\lambda}_i(\boldsymbol{z}_{i,\boldsymbol{u}_{-i}})$ *is considered a feasible policy for the COP if* $\boldsymbol{z}_{i,\boldsymbol{u}_{-i}}$ *ad-*

ditionally satisfies (4.5).

Proof. The allowable interference in (2.7) can be expressed via the state-action frequency as

$$I_{i,m}(\boldsymbol{s}^0, \boldsymbol{u}) = \sum_{\substack{(\boldsymbol{s},\boldsymbol{p})\in\mathbf{K} \\ \phi_m=j,\forall m}} \sum_{k=1}^{N} p_k\, \nu_{ki}(s_k)\Bigg) \cdot \prod_{l=1}^{N} \frac{f_l(s_l^0, \boldsymbol{u}_l; s_l, p_l)}{\sum_{\substack{(s_k,p_k)\in\mathbf{K}_l \\ \phi_k=j}} f_l(s_l^0, \boldsymbol{u}_l; s_k, p_k)} \tag{4.19}$$

By adopting similar procedures as that from the proof of Lemma 1, the relationship that $\mathscr{I}_{i,m}(\boldsymbol{z}_{i,\boldsymbol{u}_{-i}}) = I_{i,m}(\boldsymbol{s}^0, [\boldsymbol{u}_{-i}|\boldsymbol{\lambda}_i(\boldsymbol{z}_{i,\boldsymbol{u}_{-i}})])$ and $\mathscr{I}_p(\boldsymbol{z}_{i,\boldsymbol{u}_{-i}}) = I_p(\boldsymbol{s}^0, [\boldsymbol{u}_{-i}|\boldsymbol{\lambda}_i(\boldsymbol{z}_{i,\boldsymbol{u}_{-i}})])$can be easily acquired. Furthermore, since $I_{i,m}(\boldsymbol{s}^0, [\boldsymbol{u}_{-i}|\boldsymbol{\lambda}_i(\boldsymbol{z}_{i,\boldsymbol{u}_{-i}})])$ $\mathscr{I}_{i,m}(\boldsymbol{z}_{i,\boldsymbol{u}_{-i}}) \leq C_0$ and $I_p(\boldsymbol{s}^0, [\boldsymbol{u}_{-i}|\boldsymbol{\lambda}_i(\boldsymbol{z}_{i,\boldsymbol{u}_{-i}})]) = \mathscr{I}_p(\boldsymbol{z}_{i,\boldsymbol{u}_{-i}}) \leq C_1$, it can be found that $\boldsymbol{\lambda}_i(\boldsymbol{z}_{i,\boldsymbol{u}_{-i}})$ will be a feasible policy for the COP. This completes the proof. □

Lemma 3. *Given the set of policies $\boldsymbol{u}_{-i} \in \mathbf{U}_{-i}^S$ and $\boldsymbol{z}_{i,\boldsymbol{u}_{-i}}^*$ as an optimal solution for the LP problem. It is discovered that $\boldsymbol{\lambda}_i(\boldsymbol{z}_{i,\boldsymbol{u}_{-i}}^*)$ will be the best response for the COP.*

Proof. Based on Lemmas 1 and 2 associated with Theorem 3.6 in [17], the proof of this lemma can be achieved. □

In order to extend the results to N SUs, the following parameters are defined. Given the set $\boldsymbol{z} = (\boldsymbol{z}_1, \boldsymbol{z}_2, \cdots, \boldsymbol{z}_N)$ such that $\boldsymbol{z}_i = \{z_i(s,p) : (s,p) \in \mathbf{K}_i\}$ will satisfy (4.5) to (4.9), where $\boldsymbol{u} = (\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_N)$ with $\boldsymbol{u}_i = \boldsymbol{\lambda}_i(\boldsymbol{z}_i)$.

The set $\mathbf{Z}_i$ is composed by the elements $\boldsymbol{z}_i$ as stated above, and the global space $\mathbf{Z} = \prod_{i=1}^{N} \mathbf{Z}_i$. By considering the mapping function $\Psi_i(\boldsymbol{z}) : \mathbf{Z} \to \mathbf{Z}_i$, the set of optimal solutions for the LP problem in Problem 3 for each SU i can be denoted as $\Psi_i(\boldsymbol{z}) = \{z^*_{i,\boldsymbol{u}_{-i}}(s,p) : (s,p) \in \mathbf{K}_i\}$. Moreover, its product space can also be defined as $\Psi(\boldsymbol{z}) : \mathbf{Z} \to \mathbf{Z}$ where

$$\Psi(\boldsymbol{z}) = \prod_{i=1}^{N} \Psi_i(\boldsymbol{z}) \tag{4.20}$$

Theorem 2. *There exists a stationary multi-policy $\boldsymbol{u} \in \mathbf{U}^S$ as the CNE for dynamic power management problem of the considered CRN.*

Proof. According to the association of both the COP and the LP problem as described in Lemma 3, it remains to show if there exists a fixed point (i.e. $\boldsymbol{z} \in \Psi(\boldsymbol{z})$) to the vector-valued function as in (4.20). The domain of $\Psi_i(\boldsymbol{z})$ (i.e. $\mathbf{Z}_i$) is considered a compact and convex set by investigating (4.5) to (4.9), and so is its product space $\mathbf{Z}$. It is noted that $\Psi_i(\boldsymbol{z})$ is defined as

$$\Psi_i(\boldsymbol{z}) = \underset{\boldsymbol{z}_{i,\boldsymbol{u}_{-i}} \in \mathbf{Z}_i}{\arg\max} \mathscr{R}_i(\boldsymbol{z}_{i,\boldsymbol{u}_{-i}}) \tag{4.21}$$

where $\mathscr{R}_i(\boldsymbol{z}_{i,\boldsymbol{u}_{-i}})$ is observed to be a continuous function in terms of $\boldsymbol{z}_{i,\boldsymbol{u}_{-i}}$. Therefore, both $\Psi_i(\boldsymbol{z})$ and its product space $\Psi(\boldsymbol{z})$ are considered non-empty based on the extreme value theorem [19]. Furthermore, $\Psi(\boldsymbol{z})$ is a convex set for all $\boldsymbol{z} \in \mathbf{Z}$ due to the linearity of $\mathscr{R}_i(\boldsymbol{z}_{i,\boldsymbol{u}_{-i}})$. The continuity of $\mathscr{R}_i(\boldsymbol{z}_{i,\boldsymbol{u}_{-i}})$ results in the closed graph of $\Psi(\boldsymbol{z})$. The proof can consequently be completed by adopting the Kahutain's fixed point theorem [8]. $\square$

Remark 1. *Given $\boldsymbol{z}^* \in \Psi(\boldsymbol{z}^*)$, the set of stationary multi-policies $\{\boldsymbol{\lambda}_1(\boldsymbol{z}_1^*), \boldsymbol{\lambda}_2(\boldsymbol{z}_2^*) \cdots, \boldsymbol{\lambda}_N(\boldsymbol{z}_N^*)\}$ is a CNE to the dynamic power management problem for the considered CRN.*

Chapter 5

Numerical Evaluation

In this chapter, there are three issues conducted to verify the results attained from the derivation of the optimal policy. Additionally, the computation of CNE can be obtained by [8; 20]. First, we want to validate the correctness of theoretic result and examine whether to satisfy the interference constraint. According to different C_0 and C_1, we look into the simulation results. Secondly, we compare the proposed scheme with greedy approach which each SUs maximize power level to get more utility. We observe the outcomes in different interference constraints C_0 and C_1. Finally, we detect the effect of channel sensing error in proposed scheme. Substitute different amount of error to see the difference between non-error policy and error one. The error percent is defined as follow.

$$erroramount = errorpercent \times \max \mathcal{M}(\varepsilon) \quad \forall \varepsilon \in S_i, \varepsilon \in \Phi_i \tag{5.1}$$

Moreover, it is noted that the immediate utility function r_i are defined in two types :

$$r_{i,data}(\boldsymbol{s}^t, \boldsymbol{p}^t) = B \cdot \log_2 \left(1 + \frac{p_i^t\, \nu_{ii}(s_i^t)}{\sum_{j \neq i} p_j^t\, \nu_{ji}(s_j^t) + \sigma_i^2 + \varepsilon_i \phi_i^t} \right) \tag{5.2}$$

and

$$r_{i,pricing}(\boldsymbol{s}^t, \boldsymbol{p}^t) = B \cdot \log_2 \left(1 + \frac{p_i^t\, \nu_{ii}(s_i^t)}{\sum_{j \neq i} p_j^t\, \nu_{ji}(s_j^t) + \sigma_i^2 + \varepsilon_i \phi_i^t} \right) - c \times p_i^t \tag{5.3}$$

where $\boldsymbol{s}^t = (s_1^t, s_2^t, \cdots, s_N^t) \in \mathbf{S}$ and $\boldsymbol{p}^t = (p_1^t, p_2^t, \cdots, p_N^t) \in \mathbf{P}$. Equation (5.3) represents the utility function which want to achieve the fairness, i.e. the more power spread the more cost. In addition, Table I illustrates the relevant parameters that are utilized in the analysis and simulations.

Table I : System Parameters

Number of PU (M)	1
Number of SU (N)	2
Bandwidth (B)	1M (Hz)
Power level (P_i)	{0,10mW}
Channel gain ($v_{ii}; v_{ji}$)	{0.05,0.1} ; {0.025,0.05}
PU interference (ε_i)	5mW
AWGN (σ_0^2)	0.5mW
Pricing factor (c)	5M
Interference constraints ($C_0; C_1$)	

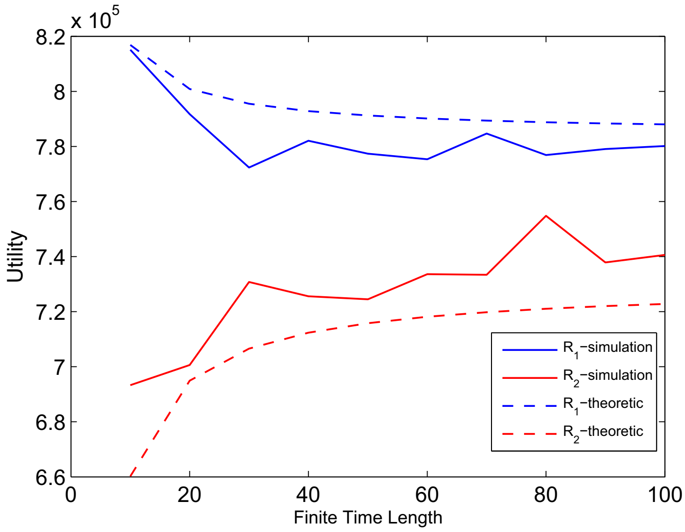

Figure 5.1: Finite Time Horizon : Time length versus expected utility under $C_0 = 0.5mW$, $C_1 = 0.5mW$ and $r_{i,data}$

5.1 Finite Time Horizon

5.1.1 Validate

Fig.(5.1 - 5.4) and Fig () show the validations of theoretic and simulation results by different utility function, 5.2 and 5.2 respectively. Because the status of expected utility doesn't reach stable, results may have a little variation. In addition, C_0 which represents the constraint with the absence of PU mainly affects the amount of expected utility, i.e. maximal value of expected utility happened when $C_0 = 0.5mW$.

Fig.(5.5 - 5.8) and Fig () present the validations of theoretic and simulation interference by different utility function, 5.2 and 5.2 respectively. The results show that all satisfy the interference constraint under the proposed scheme.

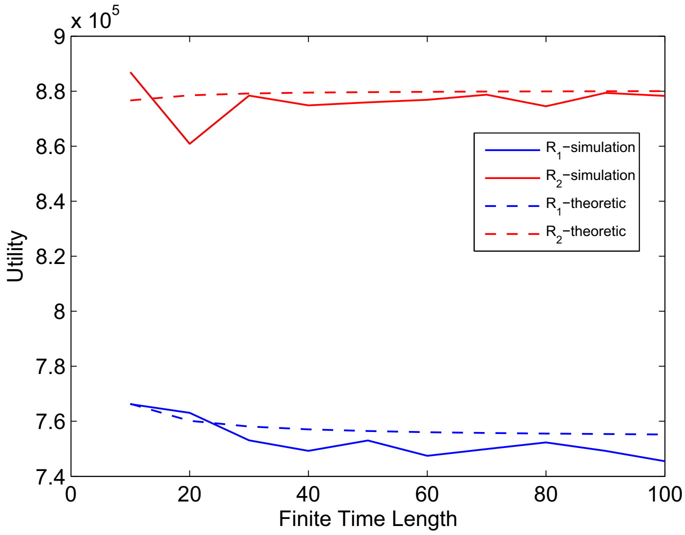

Figure 5.2: Finite Time Horizon : Time length versus expected utility under $C_0 = 0.5mW$, $C_1 = 1mW$ and $r_{i,data}$

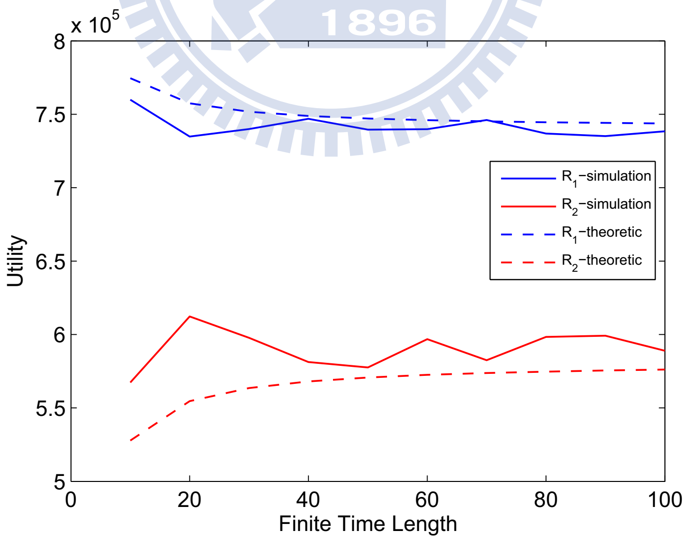

Figure 5.3: Finite Time Horizon : Time length versus expected utility under $C_0 = 0.25mW$, $C_1 = 0.5mW$ and $r_{i,data}$

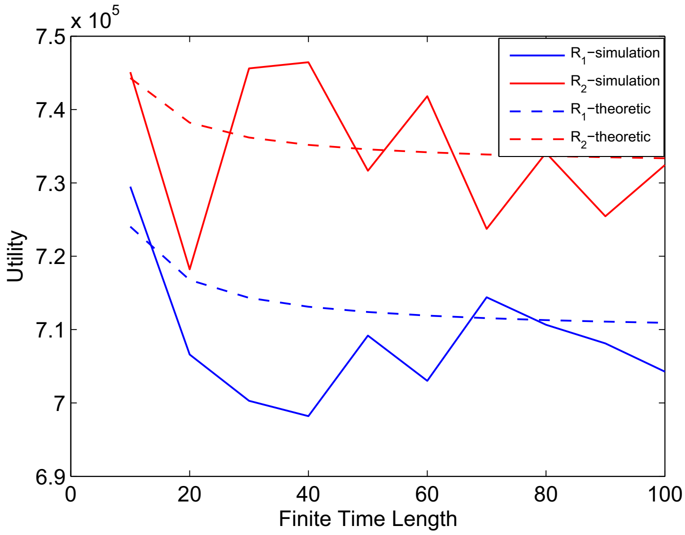

Figure 5.4: Finite Time Horizon : Time length versus expected utility under $C_0 = 0.25mW$, $C_1 = 1mW$ and $r_{i,data}$

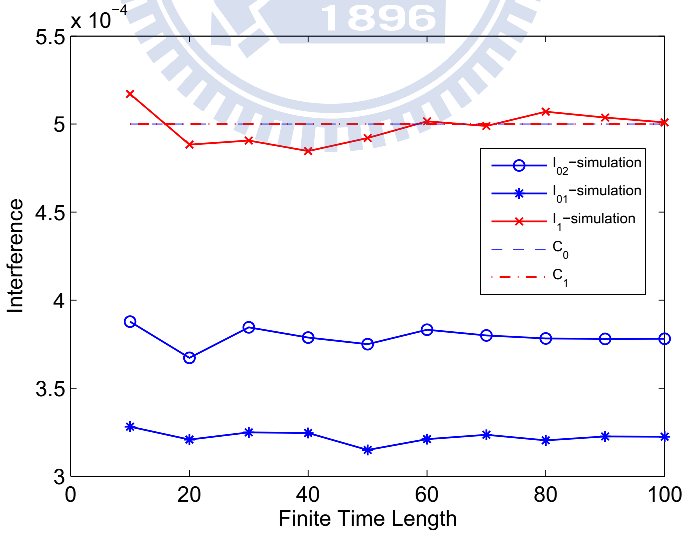

Figure 5.5: Finite Time Horizon : Time length versus expected interference under $C_0 = 0.5mW$, $C_1 = 0.5mW$ and $r_{i,data}$

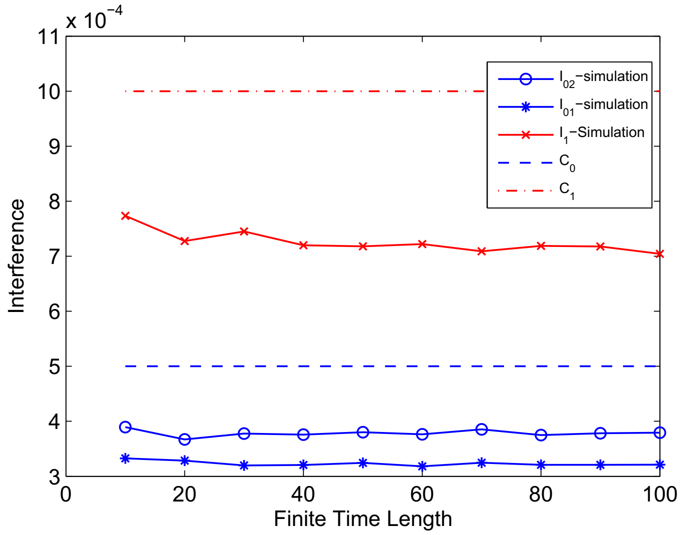

Figure 5.6: Finite Time Horizon : Time length versus expected interference under $C_0 = 0.5mW$, $C_1 = 1mW$ and $r_{i,data}$

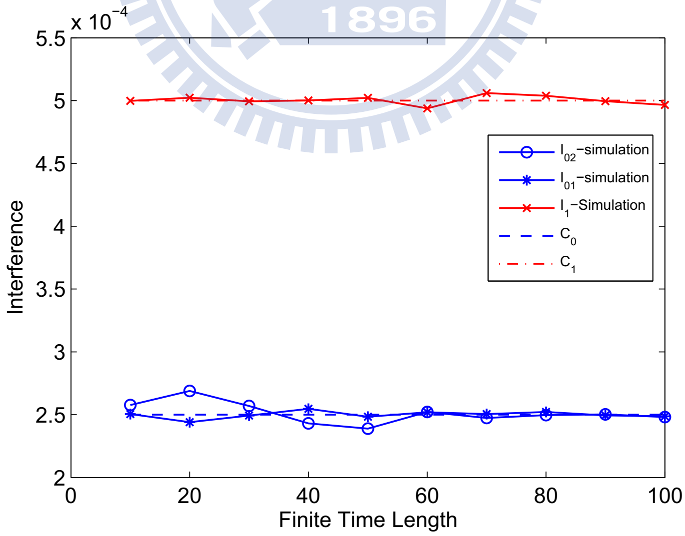

Figure 5.7: Finite Time Horizon : Time length versus expected interference under $C_0 = 0.25mW$, $C_1 = 0.5mW$ and $r_{i,data}$

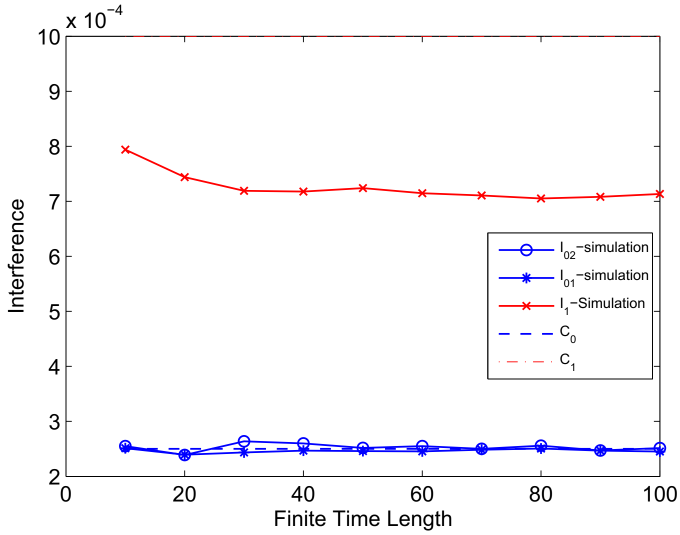

Figure 5.8: Finite Time Horizon : Time length versus expected interference under $C_0 = 0.25mW$, $C_1 = 1mW$ and $r_{i,data}$

5.1.2 Compare with greedy mechanism

Fig.(5.9 - 5.10) display the comparison of proposed and greedy mechanisms in equation (5.2). These outcomes don't show the advantage of proposed scheme due to the design of utility function. However, Fig.(5.11 - 5.12) show that proposed scheme have better performance than greedy one. Because of the curve of the equation (5.3), game theory has the ability to adjust the action to the maximal value. On the other hand, the greedy scheme always choose the maximum power which not the optimal decision. In addition, we can observe the existence of optimal action when $C_0 \geq 0.25mW$ and $C_1 \geq 0.6mW$ in Fig.(5.11) and Fig.(5.12) respectively.

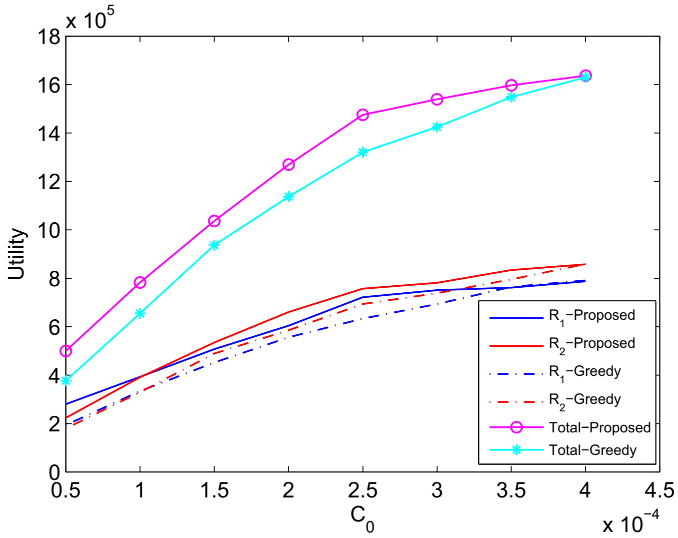

Figure 5.9: Finite Time Horizon : C_0 versus expected interference under $r_{i,data}$

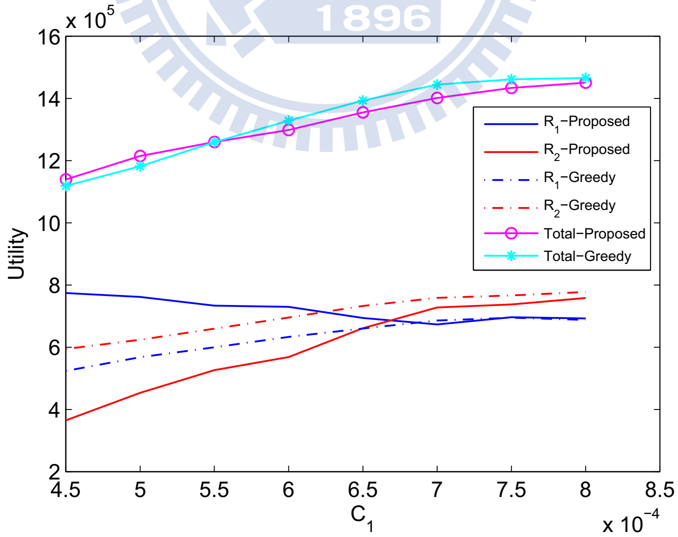

Figure 5.10: Finite Time Horizon : C_1 versus expected interference under $r_{i,data}$

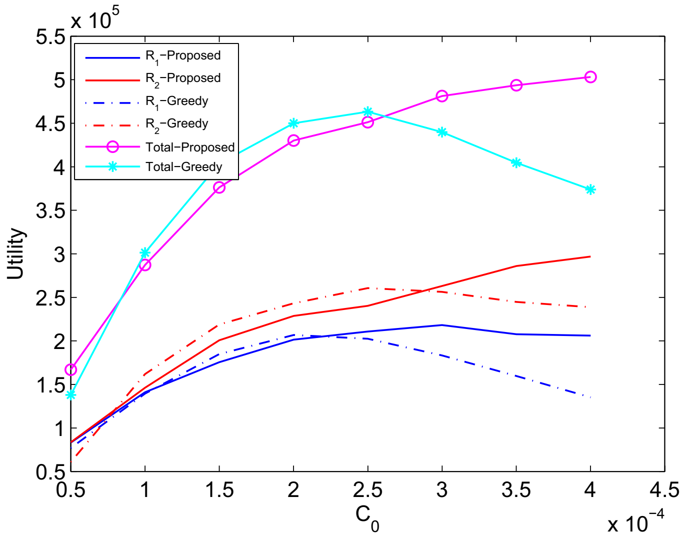

Figure 5.11: Finite Time Horizon : C_0 versus expected interference under $r_{i,pricing}$

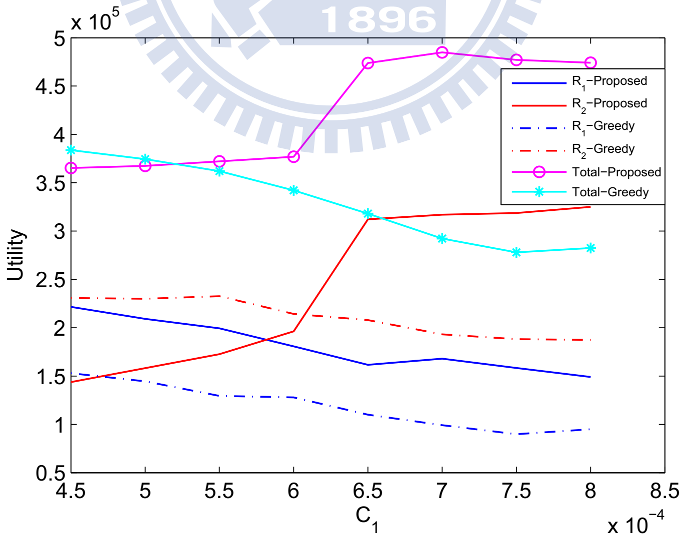

Figure 5.12: Finite Time Horizon : C_1 versus expected interference under $r_{i,pricing}$

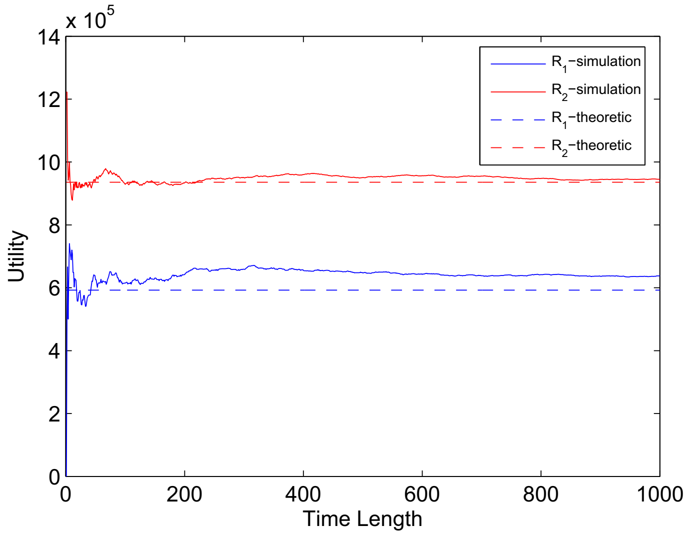

Figure 5.13: Infinite Time Horizon : Time length versus expected utility under $C_0 = 0.5mW$, $C_1 = 0.5mW$ and $r_{i,data}$

5.2 Infinite Time Horizon

5.2.1 Validate

Fig.(5.13 - 5.16) and Fig () show the validations of theoretic and simulation results by different utility function, equation (5.2) and (5.3) respectively. These results show that the proposed scheme can predict the expected utility when time length large enough. It noted that in Fig.(5.15) and Fig.(5.16) the expected utility have a few variation in former time slot. Due to the strict interference constraint of C_0, SUs have lower probability to transmit data when absence of PU. So, it may need more time to converge the theoretic value of expected utility.

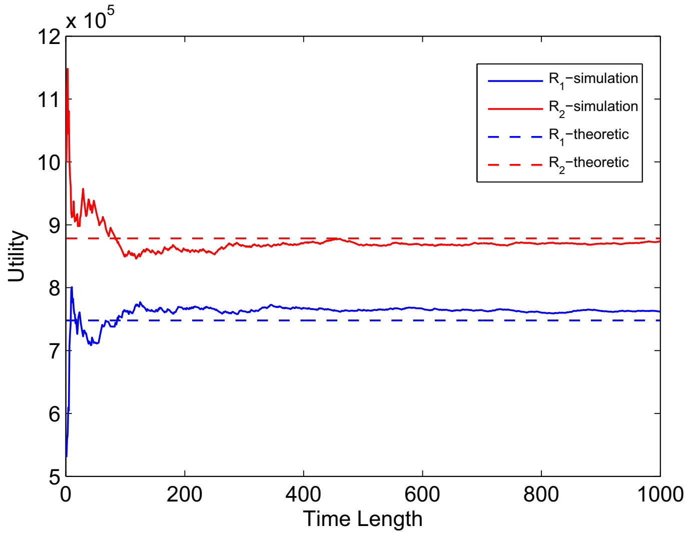

Figure 5.14: Infinite Time Horizon : Time length versus expected utility under $C_0 = 0.5mW$, $C_1 = 1mW$ and $r_{i,data}$

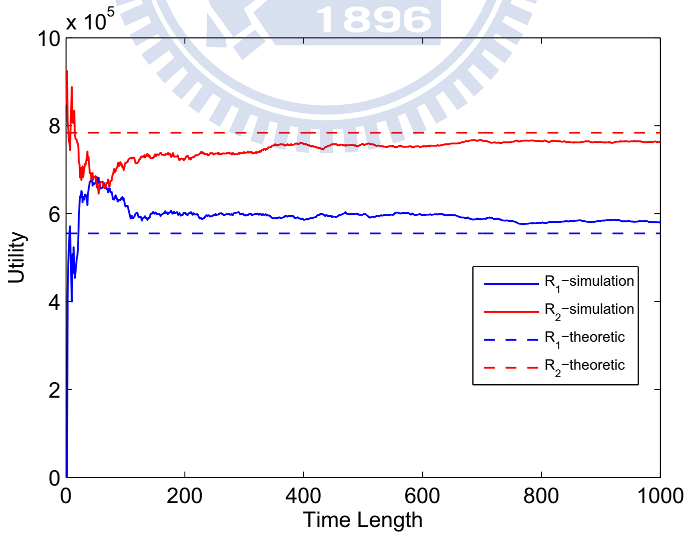

Figure 5.15: Infinite Time Horizon : Time length versus expected utility under $C_0 = 0.25mW$, $C_1 = 0.5mW$ and $r_{i,data}$

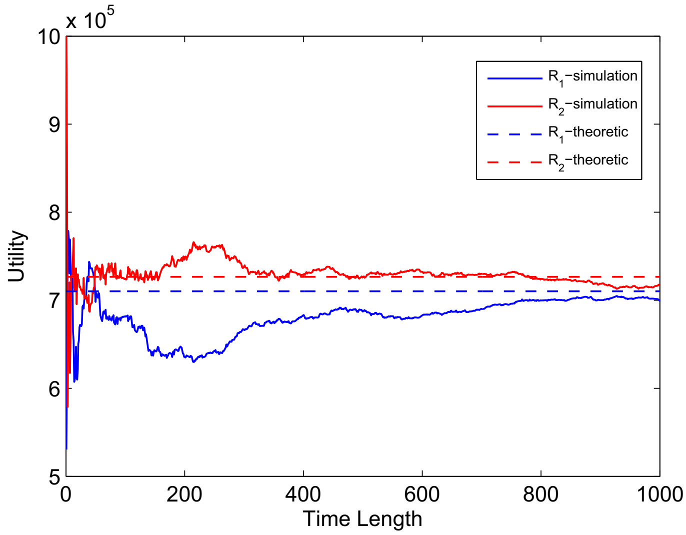

Figure 5.16: Infinite Time Horizon : Time length versus expected utility under $C_0 = 0.25mW$, $C_1 = 1mW$ and $r_{i,data}$

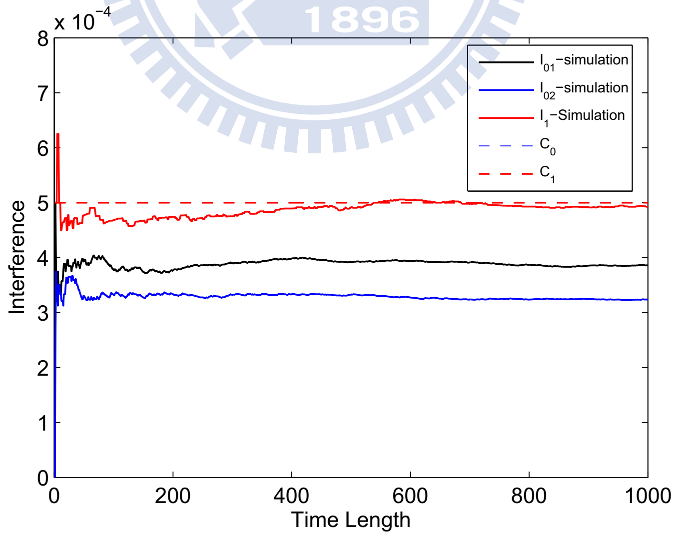

Figure 5.17: Infinite Time Horizon : Time length versus expected interference under $C_0 = 0.5mW$, $C_1 = 0.5mW$ and $r_{i,pricing}$

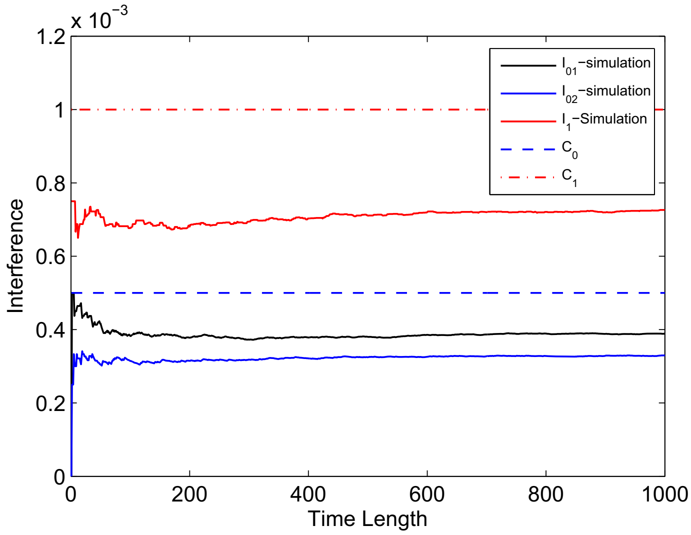

Figure 5.18: Infinite Time Horizon : Time length versus expected interference under $C_0 = 0.5mW$, $C_1 = 1mW$ and $r_{i,pricing}$

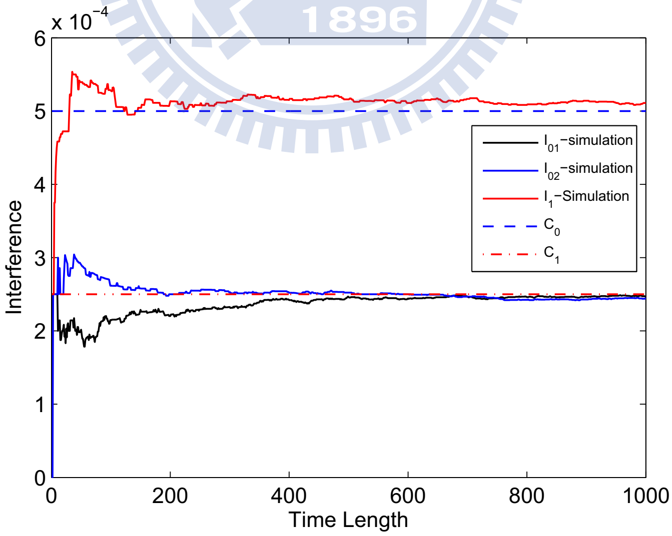

Figure 5.19: Infinite Time Horizon : Time length versus expected interference under $C_0 = 0.25mW$, $C_1 = 0.5mW$ and $r_{i,pricing}$

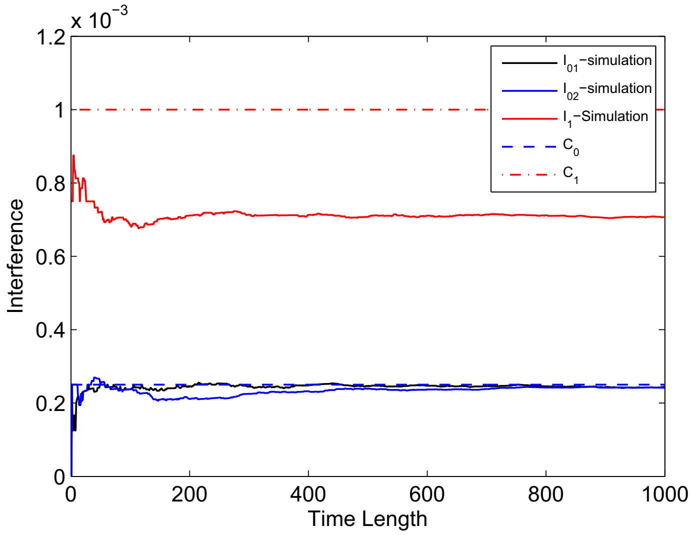

Figure 5.20: Infinite Time Horizon : Time length versus expected interference under $C_0 = 0.25mW$, $C_1 = 1mW$ and $r_{i,pricing}$

5.2.2 Compare with greedy mechanism

Fig.(5.21 - 5.24) show the comparing of proposed and greedy mechanisms in equation (5.2) and (5.3) respectively. These outcomes show that proposed scheme always better than the greedy scheme.

5.2.3 Effect of channel sensing error

Fig.(5.25 - 5.26) illustrate the effect of channel sensing error in equation (5.2) and (5.3) respectively. When error percent in equation (5.1) lower than 0.2, the expected interference doesn't exceed the constraint. However, it will cause higher interference when error percent overstep 0.2. According to this situation, we can set the strictly (e.g. $C_0 = 0.45mW$) to make up the effect of sensing error.

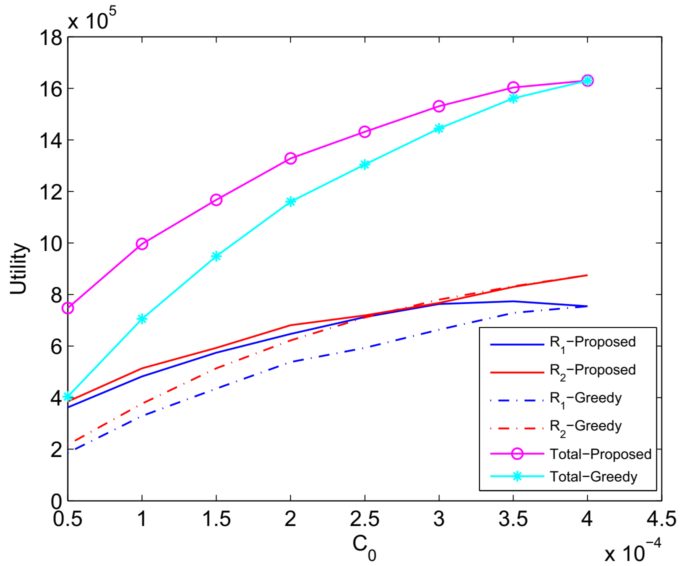

Figure 5.21: Infinite Time Horizon : C_0 versus expected interference under $r_{i,data}$

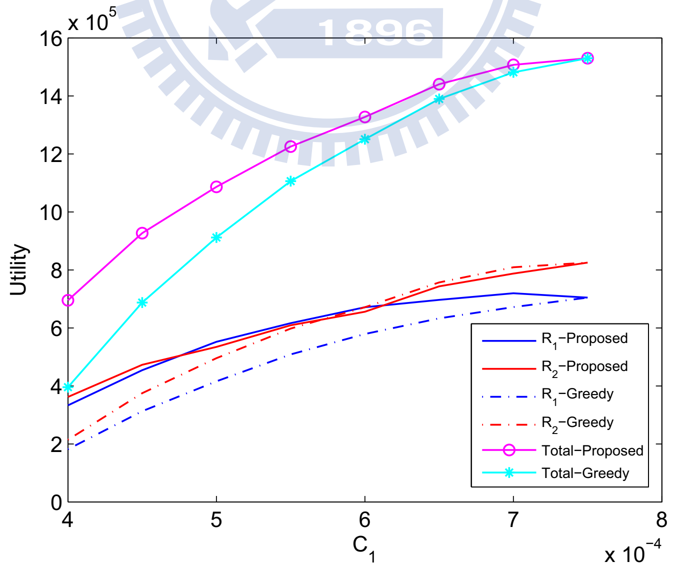

Figure 5.22: Infinite Time Horizon : C_1 versus expected interference under $r_{i,data}$

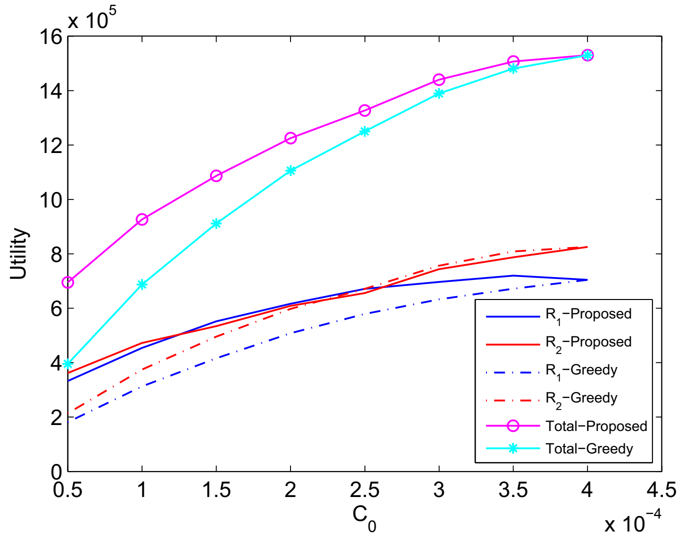

Figure 5.23: Infinite Time Horizon : C_0 versus expected interference under $r_{i,pricing}$

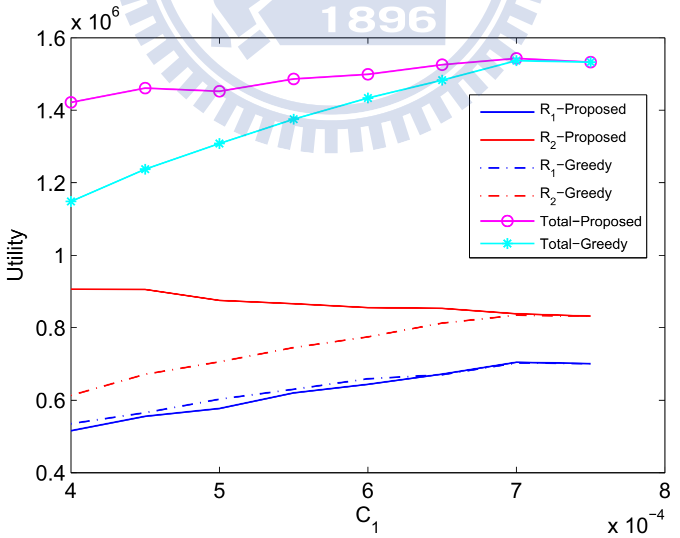

Figure 5.24: Infinite Time Horizon : C_1 versus expected interference under $r_{i,pricing}$

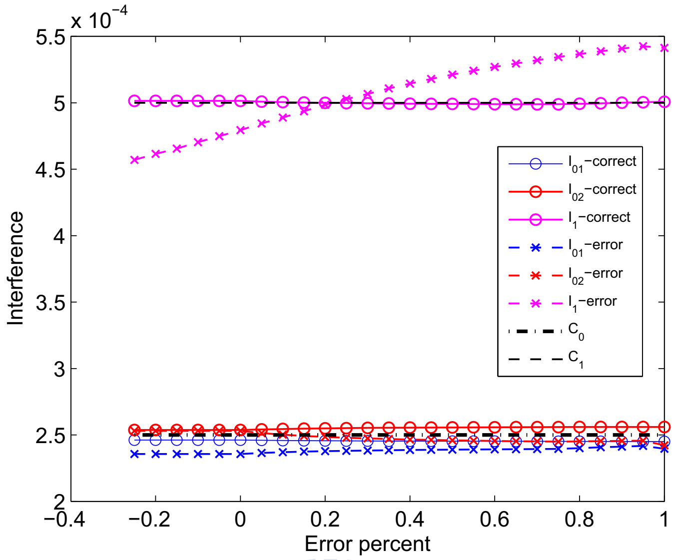

Figure 5.25: Infinite Time Horizon : error percent versus expected interference under $r_{i,data}$

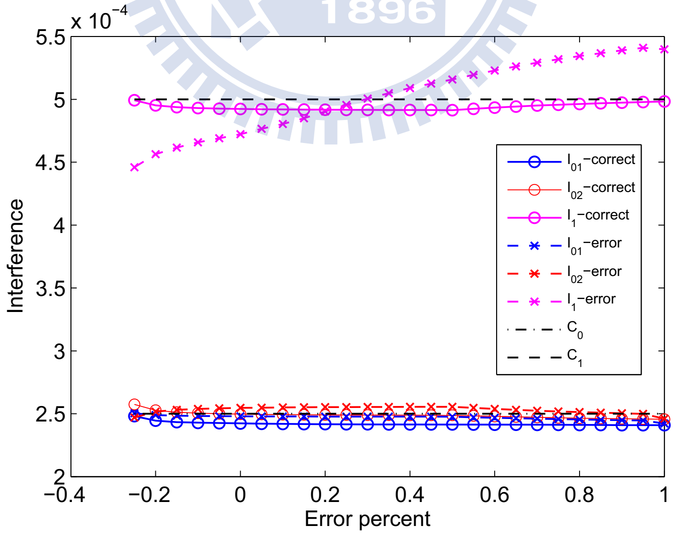

Figure 5.26: Infinite Time Horizon : error percent versus expected interference under $r_{i,pricing}$

Chapter 6

Conclusion

This paper proposes a dynamic power management scheme for maximizing the expected utility function in the cognitive radio networks (CRN). The variations from both the spectrum holes and the channel gains are considered in the network scenarios for the CRN. Based on the Markovian property of dynamic environment, finite and infinite time horizon situations are both investigated. Associated with the constraints of allowable interferences, the constrained stochastic games are utilized to acquired the optimal policies based on the objective of maximized the exptected utility function. The existence of the constrained Nash equilibrium can be proved and is served as the optimal policies for the power management problem. Simulations are performed to validate the correctness of the optimal policies that are proposed for the dynamic power management in CRN. Moreover, the proposed schemes have better performance than greedy mechanism and channel sensing error does not induce severe aberration.

Bibliography

[1] M. McHenry, "Spectrum White Space Measurements," *New America Foundation Broadband Forum*, 2003.

[2] G. Ganesan, Y. Li, and S. Li, "Spatiotemporal Sensing in Cognitive Radio Networks," *IEEE J. Sel. Areas Commun.*, vol. 26, pp. 5–12, 2008.

[3] S. M. Mishra, A. Sahai, and R. W. Brodersen, "Cooperative Sensing among Cognitive Radios," *Proc. IEEE ICC*, vol. 4, pp. 1658–1663, 2006.

[4] F. Wang, M. Krunz, and S. Cui, "Spectrum Sharing in Cognitive Radio Networks," *Proc. IEEE INFOCOM*, pp. 1885–1893, 2008.

[5] J. Jia, Q. Zhang, and X. Shen, "A Hardware-Constrained Cognitive MAC for Efficient Spectrum Management," *IEEE J. Sel. Areas Commun.*, vol. 26, pp. 106 –117, 2008.

[6] H. Su and X. Zhang, "Cross-Layer Based Opportunistic MAC Protocols for QoS Provisionings Over Cognitive Radio Wireless Networks," *IEEE J. Sel. Areas Commun.*, vol. 26, pp. 118–129, 2008.

[7] K. Hamdaoui, B.; Shin, "OS-MAC: An Efficient MAC Protocol for Spectrum-Agile Wireless Networks," *IEEE Trans. Mobile Comput.*, vol. 7, pp. 915 –930, 2008.

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

[9] Y. Wu and D. H. K. Tsang, "Distributed Multichannel Power Allocation Algorithm for Spectrum Sharing Cognitive Radio Networks," *Proc. IEEE WCNC*, pp. 1436–1441, 2008.

[10] J. Li, D. Chen, W. Li, and J. Ma, "Multiuser Power and Channel Allocation Algorithm in Cognitive Radio," *Proc. IEEE ICPP*, pp. 72–72, 2007.

[11] Z. Ji and K. J. R. Liu, "Multi-Stage Pricing Game for Collusion-Resistant Dynamic Spectrum Allocation," *IEEE J. Sel. Areas Commun.*, vol. 26, pp. 182–191, 2008.

[12] R. Etkin, A. Parekh, and D. Tse, "Spectrum Sharing for Unlicensed Bands," *IEEE J. Sel. Areas Commun.*, vol. 25, pp. 517–528, 2007.

[13] L. S. Shapley, "Stochastic Games," *Proc. Nat. Acad. Science*, vol. 39, pp. 1095–1100, 1953.

[14] E. Altman, K. Avrachenkov, N. Bonneau, M. Debbah, R. El-AzouzI, and D. S. Menasche, "Constrained Cost-Coupled Stochastic Games with

Independent State Processes," *Operations Research Letters*, vol. 36, pp. 160–164, 2008.

[15] E. Altaian, K. Avratchenkov, N. Bonneau, M. Debbah, R. El-Azouzi, and D. S. Menasche, "Constrained Stochastic Games in Wireless Networks," *Proc. IEEE GLOBECOM*, pp. 315–320, 2007.

[16] C. Derman, *Finite State Markovian Decision Processes.* Academic Press, Inc. Orlando, FL, USA, 1970.

[17] E. Altman and A. Shwartz, "Sensitivity of Constrained Markov Decision Processes ," *Annals of Operations Research*, vol. 32, pp. 1–22, 1991.

[18] A. Hordijk and L. C. M. Kallenberg, "Constrained Undiscounted Stochastic Dynamic Programming," *Mathematics of Operations Research*, vol. 9, pp. 276–289, 1984.

[19] G. B. Folland, *Advanced Calculus.* Prentice Hall, 2002.

[20] F. D. and L. D. K., *The Theory of Learning in Games.* The MIT Press, 1999.

[21] T. Ferguson, L. Shapley, and R. Weber, "Notes on a Stochastic Game with Information Structure," *International Journal of Game Theory*, pp. 222–228, 2002.

[22]