行政院國家科學委員會專題研究計畫成果報告

協力式感知無線網路之電量管理及高吞吐量協定設計
期中報告

計畫編號：NSC 98-2221-E-009 -065 -MY2
執行期限：98 年 8 月 1 日 至 100 年 7 月 31 日
主持人：方凱田 教授 國立交通大學電信工程學系

以賽局理論為基礎之感知無線網路動態電量管理

由於現存的無線網路系統採用固定頻寬分配的方式，而使得頻譜使用效率不佳，根據研究指出，有 62%的已分配頻帶在任意的時間、地點、頻帶是沒有被使用的。利用感知網路(Cognitive Radio)動態偵測和使用有認證頻帶(licensed band)的技術，可增加頻譜的使用效率，然而不可避免地，感知網路使用者(secondary user)會對主要使用者(primary user)產生干擾，所以干擾和功率的控制是感知網路須解決的問題。

在這篇報告中，考慮通道增益(channel gain)和主要使用者資訊流(primary traffic)下，利用賽局理論(Game Theory)的限制隨機賽局(constrained stochastic game)求得最佳的策略。在管理功率方面，底層波型(underlay waveform)和重疊波型(overlay waveform)的網路環境皆有考慮。另外，為了確保感知網路使用者的干擾不會影響主要使用者，利用可允須干擾(allowable interference)當作行為限制。根據限制隨機賽局的表示法，可以證明奈許平衡解(Nash equilibrium)存在於此功率管理的問題。而模擬的驗證也可確認和推導的結果一致。

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

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. According to the formulation of the constrained stochastic games, the existence of the constrained Nash equilibrium will be validated with rigorous proofs, which will be acquired as the optimal policies for the power management problem. Simulation results further validate the correctness of the theoretically derived policies for dynamic power management.

Publication:

- Chia-Wei Wang, Yu-Pin Hsu, and Kai-Ten Feng, "Dynamic Power Management in Cognitive Radio Networks based on Constrained Stochastic Games," in Proceedings of *IEEE Globe Telecommunications Conference (Globecom 2009)*, Honolulu, HI, Nov. 2009.

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

Chia-Wei Wang, Yu-Pin Hsu, and Kai-Ten Feng
Department of Electrical Engineering
National Chiao Tung University, Hsinchu, Taiwan
{casanova.cm96g, charles880353.ep90}@nctu.edu.tw and ktfeng@mail.nctu.edu.tw

***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. According to the formulation of the constrained stochastic games, the existence of the constrained Nash equilibrium will be validated with rigorous proofs, which will be acquired as the optimal policies for the power management problem. Simulation results further validate the correctness of the theoretically-derived policies for dynamic power management.**

Keywords: Cognitive radio, dynamic power management, constrained stochastic games, constrained Nash equilibrium.

I. 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 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.

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,

[1]This work was in part funded by the Aiming for the Top University and Elite Research Center Development Plan, NSC 96-2221-E-009-016, the MediaTek research center at National Chiao Tung University, and the Universal Scientific Industrial (USI) Co., Taiwan.

978-1-4244-4148-8/09/$25.00 ©2009
This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE "GLOBECOM" 2009 proceedings.

where the scheme proposed in [12] utilized the repeated game to solve NE point under underlay waveform. 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. Furthermore, 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. However, constrained stochastic games can also 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 [14] suggests an optimal solution to the dynamic power assignment according to the SUs' current state within the CRNs.

The rest of this paper is organized as follows. Section II presents the system model. The corresponding proofs for the existence of constrained Nash equilibrium are provided in Section III. Numerical evaluation is performed in Section IV; while Section V draws the conclusions.

II. System Model for Dynamic Power Management with Constrained Stochastic Games

The schematic diagram of the CRN is illustrated in Fig. 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.

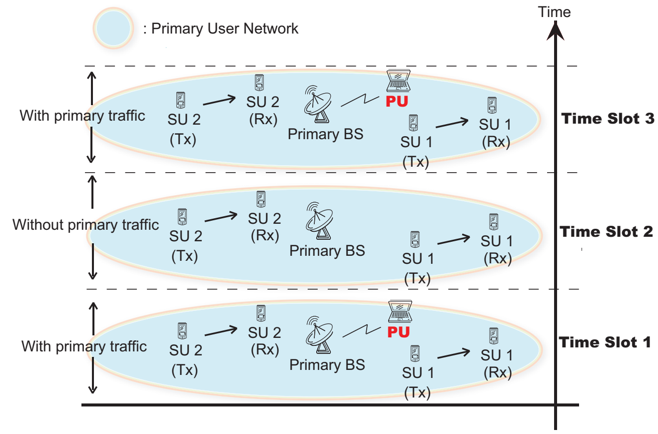

Fig. 1. The schematic diagram of the cognitive radio network for dynamic power management. (Tx : transmitter , Rx : receiver)

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 $\boldsymbol{\Phi} = \prod_i^N \boldsymbol{\phi}_i = \{\alpha, \cdots, \alpha\}$, where $\boldsymbol{\Phi}$ 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 $\mathbf{S} = \prod_{i=1}^N \boldsymbol{s}_i$. 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 $u_i^t : \mathbf{H}_i^t \rightarrow \mathcal{M}(\boldsymbol{p}_i)$ is a function mapping from the

978-1-4244-4148-8/09/$25.00 ©2009
This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE "GLOBECOM" 2009 proceedings.

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. $u_i^t : \boldsymbol{s}_i \rightarrow \mathcal{M}(\boldsymbol{p}_i)$ 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.

In this paper, the immediate utility of SU i is defined as r_i which is a function of $(\boldsymbol{s}^t, \boldsymbol{p}^t)$. For example, the achievable transmission data rate of SU i can be applied to the immediate utility as follows

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

where $\boldsymbol{s}^t = (s_1^t, s_2^t, \cdots, s_N^t) \in \mathbf{S}$ and $\boldsymbol{p}^t = (p_1^t, p_2^t, \cdots, p_N^t) \in \mathbf{P}$. The parameter B denotes the bandwidth of the licensed spectrum with unit in Hz. The function $\nu_{ji}(s_j^t)$ represents the corresponding channel gain from SU j(Tx) to SU i(Rx) in state s_j^t, and σ_i^2 is the power of the noise under the assumption of additive, white, and Gaussian distribution. ε_i denotes the interference from the primary traffic that imposes on SU i. 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

$$R_i(\boldsymbol{s}^0, \boldsymbol{u}) = \limsup_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T-1} E_{\boldsymbol{s}^0}^{\boldsymbol{u}} \left[r_i(\boldsymbol{s}^t, \boldsymbol{p}^t) \right] \quad (2)$$

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

$$I_{i,j}(\boldsymbol{s}^0, \boldsymbol{u}) = \limsup_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T-1} \sum_{k=1}^{N} E_{\boldsymbol{s}^0}^{\boldsymbol{u}} \left[p_k^t \cdot \nu_{ki}(s_k^t) \cdot \delta_j(\phi_k^t) \right] \quad (3)$$

where δ is the Kronecker delta function. In (3), $I_{i,0}(\boldsymbol{s}^0, \boldsymbol{u})$ indicates the case with the absence of primary traffic, i.e. $\delta_0(\phi_i^t = 0) = 1$; while $I_{i,1}(\boldsymbol{s}^0, \boldsymbol{u})$ 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 $I_{i,1}(\boldsymbol{s}^0, \boldsymbol{u}) \leq C_1$ 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 $I_{i,0}(\boldsymbol{s}^0, \boldsymbol{u}) \leq C_0$, 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 $I_{i,j}(\boldsymbol{s}^0, \boldsymbol{u}) \leq C_j$ for $j \in \{0, 1\}, \forall i$.

It is considered that the SUs are rational [8] such that all SUs are intending to maximize their corresponding utilities in (2). 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.

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

$$R_i(\boldsymbol{s}^0, \boldsymbol{u}^*) \geq R_i(\boldsymbol{s}^0, [\boldsymbol{u}_{-i}^* | \boldsymbol{v}_i]) \quad (4)$$

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 $\boldsymbol{u}_k^*$.

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 considered problem will be acquired in the next section.

III. Existence of CNE for Dynamic Power Management

In this section, the constrained optimization problem for dynamic power management considering a single SU will first be introduced in Problem 1. The linear programming methodology as formulated in Problem 2 will be associated with Problem 1 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 1 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 1 (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

$$R_i(\boldsymbol{s}^0, [\boldsymbol{u}_{-i} | \boldsymbol{v}_i]) \quad (5)$$

subject to

$$I_{i,j}(\boldsymbol{s}^0, [\boldsymbol{u}_{-i} | \boldsymbol{v}_i]) \leq C_j \quad \forall j \in \{0, 1\} \quad (6)$$

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 1, 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.

978-1-4244-4148-8/09/$25.00 ©2009

Problem 2 (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

$$\mathscr{R}_i(\boldsymbol{z}_{i,\boldsymbol{u}_{-i}}) = \sum_{(s_i,p_i)\in\mathbf{K}_i} \mathcal{R}_{i,\boldsymbol{u}_{-i}}(s_i,p_i) \cdot z_{i,\boldsymbol{u}_{-i}}(s_i,p_i) \tag{7}$$

subject to

$$\mathscr{I}_{i,j}(\boldsymbol{z}_{i,\boldsymbol{u}_{-i}}) = \sum_{\substack{(s_i,p_i)\in\mathbf{K}_i \\ \phi_i=j}} \mathcal{I}_{i,\boldsymbol{u}_{-i}}(s_i,p_i) \frac{z_{i,\boldsymbol{u}_{-i}}(s_i,p_i)}{\boldsymbol{Z}_{i,j}} \leq C_j \ \forall j \in \{0,1\} \tag{8}$$

$$\sum_{(s_i,p_i)\in\mathbf{K}_i} z_{i,\boldsymbol{u}_{-i}}(s_i,p_i)\left[\delta_{r_i}(s_i) - P_{s_i r_i}^i\right] = 0 \quad \forall r_i \in \mathbf{S}_i \tag{9}$$

$$\sum_{(s_i,p_i)\in\mathbf{K}_i} z_{i,\boldsymbol{u}_{-i}}(s_i,p_i) = 1 \tag{10}$$

$$z_{i,\boldsymbol{u}_{-i}}(s_i,p_i) \geq 0 \quad \forall (s_i,p_i) \in \mathbf{K}_i \tag{11}$$

where $P_{s_i r_i}^i$ in (9) is the transition probability from state s_i to r_i for SU i. The value of $\delta_{r_i}(s_i)$ in (9) 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 (8) is utilized for normalization purpose as

$$\boldsymbol{Z}_{i,j} = \sum_{\substack{(s_k,p_k)\in\mathbf{K}_i \\ \phi_k=j}} z_{i,\boldsymbol{u}_{-i}}(s_k,p_k) \tag{12}$$

The functions $\mathcal{R}_{i,\boldsymbol{u}_{-i}}(s_i,p_i)$ in (7) and $\mathcal{I}_{i,\boldsymbol{u}_{-i}}(s_i,p_i)$ in (8) 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

$$\mathcal{R}_{i,\boldsymbol{u}_{-i}}(s_i,p_i) = \sum_{\substack{(s,p)_{-i}\in\mathbf{K}_{-i}, \\ \phi_k=\phi_i,\forall k\neq i}} \prod_{m\neq i} \Omega_{i,m} \cdot r_i(\boldsymbol{s},\boldsymbol{p}) \tag{13}$$

$$\mathcal{I}_{i,\boldsymbol{u}_{-i}}(s_i,p_i) = \sum_{\substack{(s,p)_{-i}\in\mathbf{K}_{-i}, \\ \phi_k=\phi_i,\forall k\neq i}} \prod_{m\neq i} \Omega_{i,m} \left(\sum_{k=1}^N p_k\, \nu_{ki}(s_k)\right) \tag{14}$$

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

$$\Omega_{i,m} = \frac{u_m(p_m|s_m)\pi_m(s_m)}{\sum_{\substack{(s_k,p_k)\in\mathbf{K}_m, \\ \phi_k=\phi_i}} u_m(p_k|s_k)\pi_m(s_k)} \tag{15}$$

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 (15) 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

$$f_i^t(s_i^0, \boldsymbol{u}_i; s_i, p_i) = \frac{1}{t}\sum_{k=0}^{t-1} P_{s_i^0}^{\boldsymbol{u}_i}(s_i^k = s_i, p_i^k = p_i) \tag{16}$$

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 (9) to (11) 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), the following equation can be obtained:

$$R_i(\boldsymbol{s}^0,\boldsymbol{u}) = \limsup_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1} E_{\boldsymbol{s}^0}^{\boldsymbol{u}}\left[r_i(\boldsymbol{s}^t,\boldsymbol{p}^t)\right] \tag{17}$$

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

$$P_{s_i^0}^{\boldsymbol{u}_i}(s_i^t = s_i, p_i^t = p_i)\prod_{j\neq i}\frac{P_{s_j^0}^{\boldsymbol{u}_j}(s_j^t = s_j, p_j^t = p_j)}{\sum_{\substack{(s_k,p_k)\in\mathbf{K}_j \\ \phi_k=\phi_i}} P_{s_j^0}^{\boldsymbol{u}_j}(s_j^t = s_k, p_j^t = p_k)} \tag{18}$$

$$= \sum_{(s_i,p_i)\in\mathbf{K}_i} f_i(s_i^0,\boldsymbol{u}_i;s_i,p_i)\cdot \left[\sum_{\substack{(s,p)_{-i}\in\mathbf{K}_{-i} \\ \phi_l=\phi_i,\forall l\neq i}} r_i(\boldsymbol{s},\boldsymbol{p})\prod_{j\neq i}\frac{f_j(s_j^0,\boldsymbol{u}_j;s_j,p_j)}{\sum_{\substack{(s_k,p_k)\in\mathbf{K}_j \\ \phi_k=\phi_i}} f_j(s_j^0,\boldsymbol{u}_j;s_k,p_k)}\right] \tag{19}$$

$$= \sum_{(s_i,p_i)\in\mathbf{K}_i} f_i(s_i^0,\boldsymbol{u}_i;s_i,p_i)\cdot\mathcal{R}_{i,\boldsymbol{u}_{-i}}(s_i,p_i) \tag{20}$$

It is noted that the equality from (18) to (19) is mainly due to the assumption of stationary multi-policy. By substituting $\boldsymbol{u}_i$ in (20) 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 (5) and (7) 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 (9) to (11), the*

978-1-4244-4148-8/09/$25.00 ©2009

following relationship can be obtained: $\mathscr{I}_{i,j}(\boldsymbol{z}_{i,\boldsymbol{u}_{-i}}) = I_{i,j}(\boldsymbol{s}^0, [\boldsymbol{u}_{-i}|\boldsymbol{\lambda}_i(\boldsymbol{z}_{i,\boldsymbol{u}_{-i}})])$*. Moreover,* $\boldsymbol{\lambda}_i(\boldsymbol{z}_{i,\boldsymbol{u}_{-i}})$ *is considered a feasible policy for the COP if* $\boldsymbol{z}_{i,\boldsymbol{u}_{-i}}$ *additionally satisfies (8).*

Proof: The allowable interference in (3) can be expressed via the state-action frequency as

$$I_{i,j}(\boldsymbol{s}^0, \boldsymbol{u}) = \sum_{\substack{(\boldsymbol{s},\boldsymbol{p})\in\mathbf{K} \\ \phi_m = j, \forall m}} \left(\sum_{k=1}^{N} p_k \, \nu_{ki}(s_k) \right) \cdot \prod_{l=1}^{N} \frac{f_l(s_l^0, \boldsymbol{u}_l; s_l, p_l)}{\sum_{\substack{(s_k,p_k)\in\mathbf{K}_l \\ \phi_k = j}} f_l(s_l^0, \boldsymbol{u}_l; s_k, p_k)} \quad (21)$$

By adopting similar procedures as that from the proof of Lemma 1, the relationship that $\mathscr{I}_{i,j}(\boldsymbol{z}_{i,\boldsymbol{u}_{-i}}) = I_{i,j}(\boldsymbol{s}^0, [\boldsymbol{u}_{-i}|\boldsymbol{\lambda}_i(\boldsymbol{z}_{i,\boldsymbol{u}_{-i}})])$ can be easily acquired. Furthermore, since $I_{i,j}(\boldsymbol{s}^0, [\boldsymbol{u}_{-i}|\boldsymbol{\lambda}_i(\boldsymbol{z}_{i,\boldsymbol{u}_{-i}})]) = \mathscr{I}_{i,j}(\boldsymbol{z}_{i,\boldsymbol{u}_{-i}}) \leq C_j$ for all j, 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 (8) to (11), 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 2 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

$$\Psi(\boldsymbol{z}) = \prod_{i=1}^{N} \Psi_i(\boldsymbol{z}) \quad (22)$$

Theorem 1. *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 (22). The domain of $\Psi_i(\boldsymbol{z})$ (i.e. $\mathbf{Z}_i$) is considered a compact and convex set by investigating (8) to (11), and so is its product space $\mathbf{Z}$. It is noted that $\Psi_i(\boldsymbol{z})$ is defined as

$$\Psi_i(\boldsymbol{z}) = \underset{\boldsymbol{z}_{i,\boldsymbol{u}_{-i}} \in \mathbf{Z}_i}{\arg\max} \, \mathscr{R}_i(\boldsymbol{z}_{i,\boldsymbol{u}_{-i}}) \quad (23)$$

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]. □

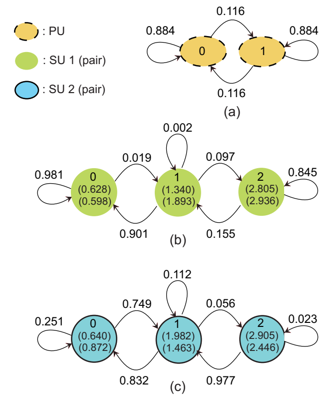

Fig. 2. Simulation parameters: (a) transition probabilities of primary traffic states (0: without primary traffic, 1 : with primary traffic); (b) transition probabilities of channel gain states for SU 1; (c) transition probabilities of channel gain states for SU 2. Thereinto, the value within the parenthesis for the channel gain state indicate the channel gain in that state.

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.*

IV. Numerical Evaluation

In this section, simulations are conducted to verify the results attained from the derivation of the optimal policy. The computation of CNE can be obtained by [8; 20]. It is noted that the immediate utility function r_i as described in (1) is utilized as the objective to be achieved for the power management problem. With the bandwidth of $B = 10$ MHz for the considered licensed spectrum, one PU and two SU pairs are considered in the CRN with the period of $T = 1000$ time slots for power management. The power constraints without and with the primary traffic are defined as $C_0 = 1000$ and $C_1 = 100$ in the unit of watt respectively, i.e. the allowable interferences among the SUs and from the SUs to the PU. The action set $\boldsymbol{p}_i = \{0, 100, 200, 300, 400\}$, interference from primary traffic $\varepsilon_i = 5$, and the noise power $\sigma_i^2 = 1$ are assumed for both SUs.

Fig. 2 illustrates the transition probabilities for both the primary traffic states and channel gain states. As in Fig. 2 (a), the value of the primary traffic state with 0 denotes the absence of primary traffic; while value with 1 stands for the existence of primary traffic. The transition probabilities of channel gain states for SU 1 and 2 are depicted in Fig. 2 (b) and (c) respectively where the the channel gains are also denoted for each state. In each state, the first value is the channel gain of self-traffic and the second value denotes that from other SU. For example, for channel gain state 2 of SU 1, the channel gain from SU 1(Tx) to SU 1(Rx) is 2.805 and that

978-1-4244-4148-8/09/$25.00 ©2009

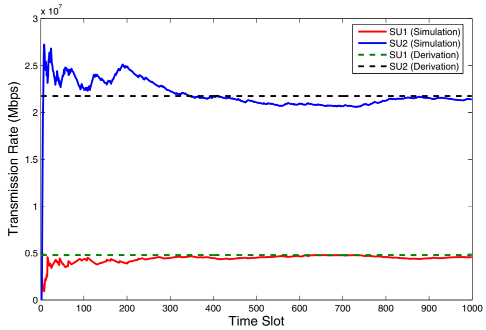

Fig. 3. Performance comparison: the transmission rates by adopting the optimal policies for SU 1 and 2 (solid lines: simulations, dashed lines: theoretical derivation).

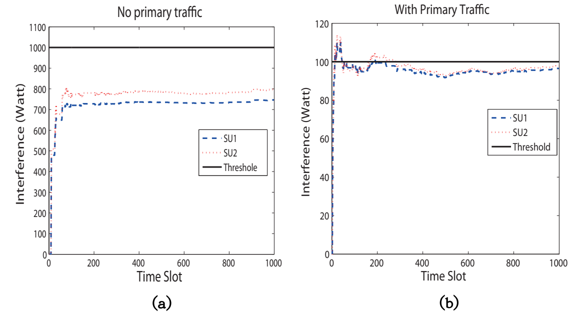

Fig. 4. Performance comparison: the allowable interference by adopting the optimal policies under case (a) without and case (b) with the primary traffic (solid lines: allowable threshold, dashed lines: simulation).

from SU 1(Tx) to SU 2(Rx) is 2.936. Based on the parameters as above, the optimal policies $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ for SU 1 and 2 that are acquired from the derivation in previous sections can be obtained as

$$\boldsymbol{u}_1(g_1,\phi_1) = \begin{bmatrix} \{0,0,1,0,0\} & \{1,0,0,0,0\} \\ \{0,1,0,0,0\} & \{1,0,0,0,0\} \\ \{0.108,0.892,0,0,0\} & \{1,0,0,0,0\} \end{bmatrix}$$

$$\boldsymbol{u}_2(g_2,\phi_2) = \begin{bmatrix} \{0,0,0,0,1\} & \{0,1,0,0,0\} \\ \{0,0,0,0,1\} & \{0,1,0,0,0\} \\ \{0,0,0,0,1\} & \{0.776,0.224,0,0,0\} \end{bmatrix}$$

where the columns indicate the state for the existence of primary traffic, i.e. column 1 without PU and column 2 with PU. The rows represent the channel gain states as denoted in Fig. 2 (b) and (c), i.e. row 1 for channel state 0, row 2 for channel state 1, and row 3 for channel state 2. The elements within the policy matrices $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ represent the probability measures on the action set. For example, $\boldsymbol{u}_1(0,0) = \{0,0,1,0,0\}$ denotes that SU 1 will take action on the power level $p_{1,2} = 200$ with probability of 1; while other power levels will not be executed.

Simulations are conduct to compare the performance with that derived from the formulations in the previous sections. Fig. 3 shows the transmission rates by adopting the optimal policies for SU 1 and 2; while Fig. 4 illustrates the allowable interferences under the cases with and without the primary traffic. It can be observed that the simulations results are consistent with that obtained from the theoretical derivations. For example, the optimal policy adopted by SU 1 results in the transmission rate of around $4.9 \times 10^6(bps)$ after $T = 250$; while the theoretical value is computed as $4.802 \times 10^6(bps)$. The transmission rate for SU 2 is around $2.1 \times 10^7(bps)$ after $T = 450$, which is consistent with the derived value of $2.173 \times 10^7(bps)$. As a result, it can be seen that the optimal policies based on the derivation of the constrained stochastic games is achievable for dynamic power management of the CRN.

V. Conclusion

This paper proposes a dynamic power management scheme for maximizing the transmission rate 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. Associated with the constraints of allowable interferences, the constrained stochastic games are utilized to acquired the optimal policies based on the objective of maximized data transmission rate. 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.

References

[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.

978-1-4244-4148-8/09/$25.00 ©2009
This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE "GLOBECOM" 2009 proceedings.