Asymmetric Cooperative Communications Based Spectrum Leasing via Auctions in Cognitive Radio Networks

Sudharman K. Jayaweera, *Senior Member, IEEE*, Mario Bkassiny, *Student Member, IEEE*, and Keith A. Avery, *Member, IEEE*

***Abstract*—Dynamic spectrum leasing (DSL) was proposed recently as a new paradigm for dynamic spectrum sharing (DSS) in cognitive radio networks (CRN's). In this paper, we propose a new way to encourage primary users to lease their spectrum: The secondary users (SU's) place bids indicating how much power they are willing to spend for relaying the primary signals to their destinations. In this formulation, the primary users achieve power savings due to asymmetric cooperation. We propose and analyze both a centralized and a distributed decision-making architecture for the secondary CRN. In the centralized architecture, a Secondary System Decision Center (SSDC) selects a bid for each primary channel based on optimal channel assignment for SU's. In the decentralized cognitive network architecture, we formulate an auction game-based protocol in which each SU independently places bids for each primary channel and receivers of each primary link pick the bid that will lead to the most power savings. A simple and robust distributed reinforcement learning mechanism is developed to allow the users to revise their bids and to increase their rewards. The performance results show the significant impact of reinforcement learning in both improving spectrum utilization and meeting individual SU performance requirements.**

***Index Terms*—Cognitive radios, cooperative communications, distributed dynamic spectrum leasing, dynamic spectrum access, dynamic spectrum sharing, auction game, game theory.**

I. Introduction

IN [1], [2] the authors introduced the concept of *dynamic spectrum leasing* (DSL) as a new paradigm for dynamic spectrum sharing (DSS) in cognitive radio networks (CRN's). They were motivated by the observation that the passive participation, or rather the non-participation, of primary users as assumed in the previously proposed dynamic spectrum access (DSA) schemes is inefficient in terms of fully utilizing the spectrum. This is because, in DSA, the secondary users (SU's) are responsible for managing the spectrum sharing process while not compromising the primary Quality-of-Service (QoS). The primary users do not have a stake in the process, and thus act completely oblivious to the existence of the SU's as well as the ongoing dynamic spectrum coexistence. On the other hand, in the Dynamic Spectrum Leasing framework, as originally proposed in [1], [2], the primary users are allowed to proactively manage the amount of secondary activity in their licensed spectrum band. Earlier, the idea of spectrum leasing was proposed as a static or offline spectrum sharing technique [3]. However, a similar concept to [1], [2] was proposed in [4], but the latter case relied on cooperative communications between primary and secondary users and does not consider an underlay cognitive architecture as in [1], [2]. The Dynamic Spectrum Leasing presumes that there is a reward for primary users for accepting secondary activity whenever it is affordable without compromising their own QoS. Cognitive Radios (CR's) as envisioned in [5] as radio devices capable of learning and adapting to their RF environment, make an ideal platform for both DSS in general.

As mentioned earlier, the DSA architecture does not consider any participation from the primary system in determining the spectrum sharing process. It was shown in [1], [2], [6]–[9] that both primary and secondary systems could benefit if the primary users were to play an active role, however small, in managing the spectrum sharing process. The Dynamic Spectrum Leasing is shown to be implementable in a game theoretic framework in which both primary and SU's are considered as the players, in contrast with DSA in which only SU's are assumed to be the players. Previous Dynamic Spectrum Leasing proposals focused only on spectrum underlay architectures. Thus, the utility of primary users in previously considered Dynamic Spectrum Leasing games was expressed as a monetary reward proportional to the tolerated interference from the SU's. The utility of SU's were allowed to be of many forms, as in previous DSA proposals. For example, these could be secondary throughput or energy efficiency [1], [10].

In this paper, we propose a completely new way to encourage primary users to lease their spectrum, whenever affordable: Rather than a monetary reward, in the proposed framework the primary reward is accrued in terms of savings on their communication resources, namely the power. This is achieved by proposing an asymmetric cooperative communications architecture in the combined network consisting of both primary and secondary systems. The proposed asymmetric cooperative communications can be realized with very little inter-system information exchange. Note that, user cooperation has previously been considered for data transmission in cognitive networks albeit without the assumption of Dynamic

Manuscript received November 16, 2010; revised March 10, 2011 and May 16, 2011; accepted May 25, 2011. The associate editor coordinating the review of this paper and approving it for publication was M.-O. Pun.

S. K. Jayaweera and M. Bkassiny are with the Department of Electrical and Computer Engineering, University of New Mexico, Albuquerque, NM, USA (e-mail: {jayaweera, bkassiny}@ece.unm.edu).

K. A. Avery is with the Space Vehicles Directorate, Air Force Research Laboratory (AFRL), Kirtland AFB, Albuquerque, NM, USA.

This research was supported in part by the Space Vehicles Directorate of the Air Force Research Laboratory (AFRL), Kirtland AFB, Albuquerque, NM.

Digital Object Identifier 10.1109/TWC.2011.11.102044

Spectrum Leasing [11]–[13]. Indeed, they only considered spectrum underlay models in which secondary nodes relay the primary signal to its destination, in order to mitigate the effect of additional interference to the primary caused by the secondary signals. In the proposed framework, the SU's spend a portion of their transmit powers to asymmetrically relay the primary signals to their destinations. This asymmetry results from having the SU to relay the primary signal while the primary user transmits only its own signal. In return, the primary users lease a certain portion of their spectrum resource to the SU's. This could be interpreted as having the SU's to use their power as currency to buy the bandwidth, thus establishing an exchange rate between power and bandwidth. In this formulation, the primary reward is the power saving they achieve due to cooperative relaying. Compared to previous Dynamic Spectrum Leasing proposals [1], [2], [6], [7], we also relax the assumption of a centralized primary system: In our proposal, each *primary link* (i.e. transmitter and a receiver pair) is allowed to act autonomously in making decisions on spectrum leasing. It will entertain bids from the secondary system specifying how much power would be spent for relaying primary signals. The Dynamic Spectrum Leasing game, thus, leads to an auction in which primary users act as the auctioneers. On the other hand, we propose cooperative communications-based Dynamic Spectrum Leasing frameworks suitable for both centralized and decentralized CRN's. The centralized cognitive network model assumes that there is Secondary System Decision Center (SSDC) [14] that is responsible for making spectrum leasing decisions for the whole secondary system. The SSDC decides which SU should cooperate with which primary user/link. Such decision-making by the secondary system allows it to better negotiate with the primary users. However, each primary user may accept an offer of cooperation with a SU picked by the SSDC only if this would result in at least a certain minimum power saving. If the offer is too low, the corresponding primary user may simply decline the offer and the access to its channel would be denied until the next bidding interval. Each primary user may keep its threshold power level as private information so as to encourage bids as high as possible from the secondary system.

While it gives the secondary system to have more control over its relaying power bids, the feasibility of centralized decision-making in CRN's that operate as secondary systems may be questionable. It requires dedicated (control) channels with enough bandwidth to support reporting of all spectrum sensing (in this case primary leasing offers) measurements at distributed CR's to the SSDC, as well as channel and power allocation decisions from the SSDC back to the distributed radios. While such centralized models are widely assumed in existing literature [7], [15], [16], it is not clear how realistic they might be in practice. On the other hand, we believe that true CR's may very well be the one's that can operate autonomously, yet efficiently. Thus, next we consider a secondary CRN in which users make their own spectrum access decisions without any centralized control. In such a decentralized secondary network, SU's may compete with each other in order to gain access to available primary channels. This leads to a new *distributed dynamic spectrum leasing* (D-DSL) architecture that may be suitable for future heterogeneous wireless network scenarios. The proposed auctioning-based D-DSL framework is applicable to both spectrum interweave and underlay architectures.

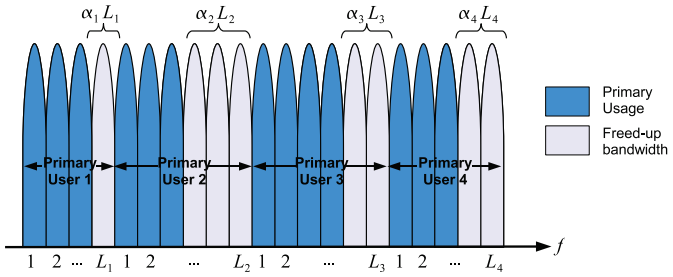

Fig. 1. Distributed dynamic spectrum leasing (D-DSL) in an OFDMA-based wireless network. Each user/link dynamically decides to lease an α_i fraction of its allocated sub-carriers.

We believe that sophisticated autonomous learning to be the defining feature of future CR's. In a decentralized CRN, the SU's who do not win a favorable channel at the beginning of the dynamic spectrum auction process will employ cognitive learning to win a bid for a channel in subsequent bidding times. Each winning secondary node (one per each primary channel) may also use learning to revise its bid in subsequent bidding times to improve its own power savings. In this paper, we develop a simple, yet robust, reinforcement learning mechanism to achieve distributed and autonomous learning from the past experience without any supervision. We show that without any centralized control both primary and secondary radios can learn to arrive at an equilibrium in a completely distributed Dynamic Spectrum Leasing framework. Note that, recently there has been a growing interest in applying Reinforcement Learning techniques to CR's [17], [18]. It permits the cognitive users to learn by interacting with their environment. Other learning method can be found in the literature, such as the Markov model and neural networks [19], [20]. However, these methods are of little interest when there is no full knowledge about the system or in the absence of supervision. That's why we propose a reinforcement learning technique and we show, through simulations, how effective the proposed auction-based D-DSL framework in utilizing the spectrum resources as well as the significant impact of reinforcement learning in both improving spectrum utilization and meeting individual SU performance requirements.

The rest of this paper is organized as follows: Section II defines the system model, Sections III and IV describe the proposed Dynamic Spectrum Leasing model with both the centralized and decentralized CR architectures, respectively. In Section V we show the simulation results, and finally, Section VI concludes the paper by summarizing our results.

II. System Model

The centralized Dynamic Spectrum Leasing (C-DSL) architecture of [7] assumes that all primary and SU's coexist in the whole spectrum band of interest. However, in almost all wireless systems the total spectrum is usually divided into a multiple number of (primary) channels. A channel allocation scheme either dynamically or statically allocates

these channels to primary user, as needed. In the following we assume that there are L primary users/links on L distinct primary channels. Thus, we will use the terms primary user, primary channel and primary link interchangeably. To be general, let us assume that the allocated bandwidth of channel $i \in \mathcal{C} = \{1, \cdots, L\}$ is W_i. At this point, it is perhaps worth mentioning that these channels do not have to be necessarily frequency channels. For example, they could be TDMA channels, in which case the channel resource would be the time slot length T_i of channel/user i. Also the proposed D-DSL architecture can be adapted for an OFDMA-based primary system, in which the i-th primary user can be assumed to be allocated an L_i number of OFDMA sub-carriers as in Fig. 1. In the following, to save space, we will always discuss things in the context of primary channels being distinct FDMA channels. For simplicity, in this paper, we will also assume that each SU has the capability to transmit only over one channel at a time, and that each primary TX-RX pair (link) is allowed to be leased to only one SU at a time.

The time horizon is assumed to be split into time frames of duration T_f each by the primary system, and each time frame is divided into a number of equal-length time slots. We assume that the channel fading varies slowly within a time frame, and thus fading can be considered constant in this time duration. The fading model that we consider can represent the slow-fading channels that result from large-scale changes in the user's location. A possible scenario could occur when a CR moves for a long duration in a certain region, which would change the average power that is received by the CR at each location [21]. Suppose that the maximum transmit power of i-th primary user is $\bar{P}_i$. As required QoS, the RF interference and the observed channel fading (state) conditions change from one time frame to another, the i-th primary user may be able to achieve its required QoS by using only $(1-\alpha_i)$ fraction of its allocated bandwidth W_i, for $\alpha_i \in [0, 1]$. This is the origin of the so-called spectrum holes that leads to the spectrum under-utilization. In existing proposals for DSS based on DSA, the primary users do not pay any attention to this phenomenon, and the SU's are expected to sense the spectrum and detect these opportunities: Whichever the SU that successfully detects these seemingly random spectrum holes will get to access them, perhaps on a contention-basis. Certainly, according to existing DSA proposals, there is no reason for the primary users to pay any attention to who accesses these white spaces, because they do not have anything to gain. By default, in DSA proposals, the focus is on *just utilizing* the spectrum holes rather than *efficient utilization* of spectrum holes.

In contrast, according to our proposed D-DSL framework, if at the beginning of a time frame the i-th primary user determines that it can achieve its required QoS by using only $(1 - \alpha_i)$ fraction, for $0 \leq \alpha_i \leq 1$, of its bandwidth W_i (or sub-carriers L_i, or time slot T_i), then it *consciously* decides to free-up up to an α_i fraction of its bandwidth W_i for SU's to lease. Note that, if there is frequency selective channel fading across the bandwidth W_i, then the i-th primary user will have the freedom to decide which parts of its allocated bandwidth to be freed-up. Although this may be an important aspect in practice, to avoid notational complexity, in this paper we will assume that each primary channel is frequency flat. Thus, to be concrete, we may assume that always the last $\alpha_i W_i$ portion of each channel will be freed-up.

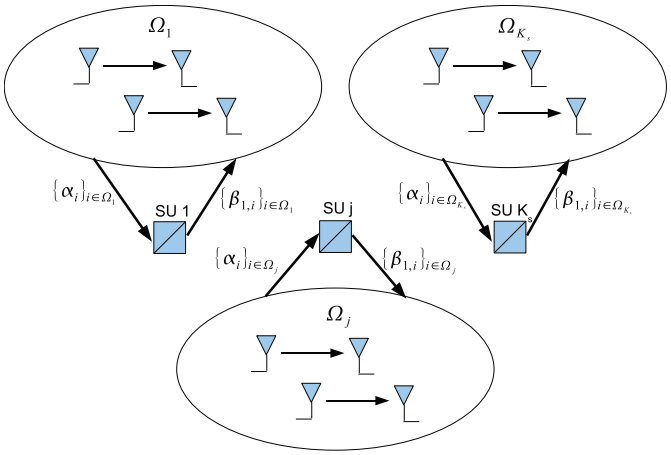

Fig. 2. Distributed dynamic spectrum leasing (D-DSL) based on auction game.

We assume that there are K_s number of SU's, each with maximum transmit power $\bar{P}_j$ for $j \in \mathcal{K}_s$ where $\mathcal{K}_s = \{1, \cdots, K_s\}$. At the beginning of each time frame, each SU $j \in \mathcal{K}_s$ receives all α_i's from $i \in \Omega_j$ where $\Omega_j \subseteq \{1, \cdots, L\}$ denotes the set of neighboring primary channels (i.e. the primary channels that can be sensed) of the j-th SU, as shown in Fig. 2. Note that the Ω_j sets are not necessarily disjoint. The j-th SU uses the available Channel State Information (CSI) to compute the portion $\beta_{j,i}$ of its power $\bar{P}_j$, where $\beta_{j,i} \in [0, 1]$, that can be allocated to relay the primary signals of the i-th channel for each $i \in \Omega_j$. Each SU computes this set of $\{\beta_{j,i}\}_{i\in\Omega_j}$ such that if it spends $\beta_{j,i}\bar{P}_j$ amount of its power to relay the i-th primary user's signal, it can still achieve a minimum probability of error ϵ over a transmission bandwidth of $\alpha_i W_i$.

In the following, we first consider the Dynamic Spectrum Leasing auctions for centralized cognitive secondary networks and derive the optimal decision-making policies for both primary and SU's. Next, we consider the Dynamic Spectrum Leasing auction based on asymmetric cooperative communications for decentralized cognitive secondary networks in which CR's are equipped with learning capabilities, and derive the equilibrium point.

III. Asymmetric Cooperative Communications-based DSL for Centralized Cognitive Secondary Networks

Suppose that, at the beginning of a given time frame, the i-th primary TX determines it can free-up an α_i fraction of its bandwidth W_i. The objective of the primary user is to gain power savings in return via possible asymmetric cooperative communications facilitated by the SU's.

The assumed asymmetric cooperative system is depicted in Fig. 3: The SU $j \in \mathcal{K}_s$ *coherently* relays the signal of the primary user $i \in \mathcal{C}$ over a link with a fading coefficient $h_{j,i}$. For the sake of illustrating the method, we assume a genie-aided cooperation so that the secondary relay knows the

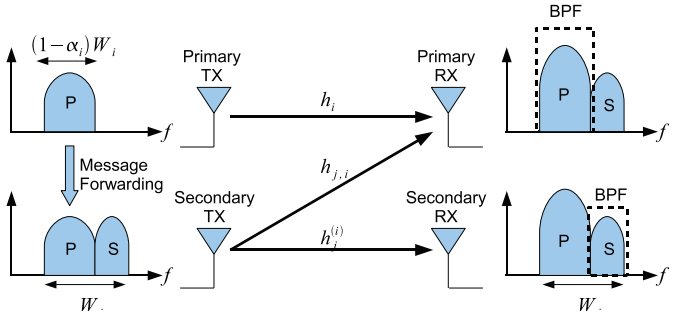

Fig. 3. Asymmetric cooperative communication achieved by primary users with the help of secondary relays.

primary message to be relayed instantaneously. In practice, this assumption can be implemented by assuming that the primary and secondary transmitters are close to each others [22], compared to the other nodes, then, their channel will have a relatively high gain which allows the primary to transmit its message to the secondary TX in a short duration. Afterwards, the primary and secondary TX's transmit, simultaneously, the primary message to its destination. Hence, the relayed signal is transmitted over the bandwidth $(1-\alpha_i)W_i$ that the primary user uses for its own transmission. Note that, this assumption can be easily dropped by adapting a practical cooperative protocol at the expense of more elaborate notation [4], [23]. The SU transmits its own signal over the freed-up bandwidth $\alpha_i W_i$. We denote by h_i the fading coefficient between the primary TX i and the corresponding RX, and $h_j^{(i)}$ the fading coefficient between the secondary TX j and its corresponding secondary RX, when transmitting over channel i.

A. Primary and Secondary Actions

Suppose that the i-th primary user needs a minimum data rate of $R_i^{(\min)}$ on its link. While transmitting at its nominal power level of $\bar{P}_i$, the rate that a primary user can achieve by freeing-up α_i fraction of its bandwidth for leasing is

$$R_i(\alpha_i) \quad = \quad (1-\alpha_i)W_i \log\left(1+\Gamma_i(\alpha_i)\right) \qquad (1)$$

where $\Gamma_i(\alpha_i) = \frac{|h_i|^2 \bar{P}_i}{(1-\alpha_i)W_i N_i}$ is the resulting signal-to-noise ratio (SNR). Suppose that with the current realization of CSI h_i on the primary link, it can achieve a rate of $R_i^{(\min)}$ with only using a minimum of $(1-\alpha_i^{(\max)})$ fraction of its bandwidth (if it transmits at its nominal transmit power $\bar{P}_i$). Then, the primary i may free-up an $\alpha_i \in [0, \alpha_i^{(\max)}]$ fraction of its spectrum resource without degrading its QoS.

Each SU $j \in \mathcal{K}_s$ receives all α_i's from primary users $i \in \Omega_j$, as shown in Fig. 4, and computes the power fractions $\beta_{j,i}$'s for all $i \in \Omega_j$ so that $\beta_{j,i} \in [0, \beta_{j,i}^{(\max)}]$ where $\beta_{j,i}^{(\max)}$ is the maximum power fraction it can allocate to relaying the i-th primary user signal while maintaining a minimum Bit Error Rate (BER) of ϵ with respect to its own receiver over the channel i (the portion $\alpha_i W_i$):

$$\beta_{j,i}^{(\max)} = \arg \max_{\beta_{j,i}\in[0,1]} \beta_{j,i} \qquad \text{s. t.} \quad P_e^{(j,i)}\left(\beta_{j,i}\right) < \epsilon \ , \qquad (2)$$

where $P_e^{(j,i)}(.)$ is the BER of the j-th secondary link if transmitting on primary channel i. If SU j gets to to transmit over channel i, then it will receive a utility of:

$$u_j(\beta_{j,i}, \alpha_i) = \alpha_i W_i \log\left(1+\gamma_{j,i}\right) q\left(\epsilon - P_e(\beta_{j,i})\right) \ , \qquad (3)$$

where $\gamma_{j,i} = \frac{\bar{P}_j(1-\beta_{j,i})|h_j^{(i)}|^2}{\alpha_i W_i N_i}$ is the SNR of the SU j on the leased channel of primary i, and $q(.)$ is the unit-step function. When employing BPSK transmission, $P_{e,j,i}(\beta_{j,i}) = Q\left(\sqrt{\gamma_{j,i}}\right)$ for any $\alpha_i \neq 0$, so that $\beta_{j,i}^{(\max)}$ can be obtained by numerically solving (2). If $\alpha_i = 0$, we let $\beta_{j,i}^{(\max)} = 0$ for all $j \in \mathcal{K}_s$, meaning that SU's will not relay primary user's signal who is not willing to lease any portion of its available bandwidth.

Suppose that the primary users decide to free-up the spectrum segments $\{\alpha_i W_i\}_{i=1}^L$. If the objective of the cognitive users is to maximize the secondary network sum-rate, the optimal choice would be to let $\beta_{j,i} = 0$ for $\forall j$ and $\forall i$. This will enable the SU's to use all their power resources exclusively for the secondary transmission. Of course, the primary users then will not have an incentive to lease spectrum, and thus would rather keep transmitting over the whole spectrum without freeing any portion. Thus, in our proposed C-DSL model, we assume that each primary user expects to be able to reduce its transmit power below a certain threshold $P_i^{th} \leq \bar{P}_i$ due to the cooperative communication advantage with the SU's. In general, this threshold P_i^{th} of user $i \in \mathcal{C}$ is unknown to the SU's. Hence, in our C-DSL model, if a primary user does not receive an offer $\beta_{j,i}$ from the secondary system that will enable it to meet the target power reduction it expects, it may not accept the offer and not lease the spectrum portion. For that reason, the SU's will attempt to choose their $\beta_{j,i}$ values closer to $\beta_{j,i}^{(\max)}$.

In our proposed model, each SU may pick the $\beta_{j,i}$ fraction using a particular distribution, or weighting, over $[0, \beta_{j,i}^{(\max)}]$ (not necessarily uniform). For example, if it has a large battery life remaining it can pick a value closer to $\beta_{j,i}^{(\max)}$ and vice versa. Thus, a possible method for picking up $\beta_{j,i}$ could be as $\beta_{j,i}^{(\max)}\left(1-e^{-aT_j^{(res)}}\right)$, where $T_j^{(res)}$ is the residual battery life of the j-th SU.

B. Optimal Channel Assignment at the SSDC

In a centralized cognitive secondary network, we may assume that an SSDC is responsible for making the secondary bid decisions, and then broadcasting these decisions to the SU's through a control channel [14]. At the beginning of each time frame, each SU computes the fraction $\beta_{j,i}$ of its power that it is willing to allocate for relaying the primary signal, and informs these $\beta_{j,i}$ values to the SSDC through a control channel. The SSDC uses the $\beta_{j,i}$ values and the knowledge of channel fading coefficients to determine the channel assignment for each SU so as to maximize the secondary system's sum-rate, as shown in Fig. 4, where $j(i)$ denotes the index of the SU that is assigned to transmit over the channel $i \in \mathcal{C}$. If none of the SU's is assigned to channel i, we let $j(i) = 0$ (0 denoting a dummy SU).

We define the mapping $\phi : \mathcal{K}_s \rightarrow \mathcal{C} \bigcup 0$ as the scheduling function used by the SSDC that assigns each SU to a primary channel. We let $\phi(j) = 0$ to denote that SU j is not assigned

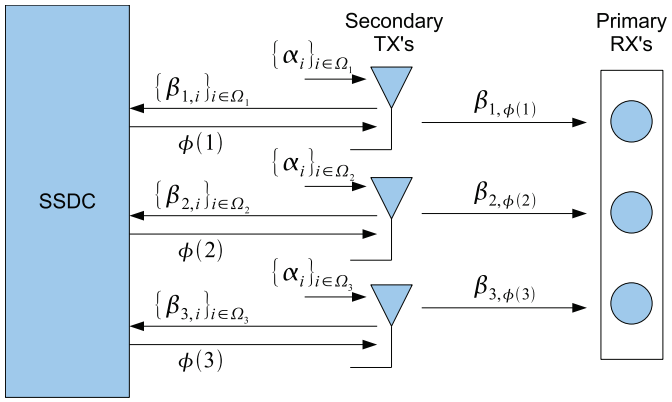

Fig. 4. Secondary System Decision Center (SSDC) operation.

to any primary channel. The optimal channel assignment ϕ^* of SU's is given by the optimization:

$$\phi^* = \arg\max_{\phi} \sum_{j=1}^{K_s} u_j(\beta_{j,\phi(j)}, \alpha_{\phi(j)}), \tag{4}$$

where $u_j(\beta_{j,i}, \alpha_i)$ is as defined in (3). We let $u_j(\beta_{j,\phi(j)}, \alpha_{\phi(j)})|_{\phi(j)=0} = 0$, meaning that the utility of a SU that is not assigned to any available channel is 0. The solution of (4) can be obtained via the Hungarian algorithm since it can be identified as a bipartite matching problem that consists of the bipartite sets $\mathcal{C}$ and $\mathcal{K}_s$. The Hungarian algorithm [24] finds the optimal matching between the elements of the bipartite sets such that it maximizes the sum of the edge weights. If the edge weight between primary $i \in \mathcal{C}$ and secondary $j \in \mathcal{K}_s$ is defined to be the utility $u_j(\beta_{j,i}, \alpha_i)$ then this solution leads to the optimal channel assignment that maximizes the secondary sum-rate. The advantage of this algorithm is that it can find the optimal channel assignment at a cubic complexity. A description of this algorithm can be found, for example, in [11], [25].

At the beginning of each time frame, the SSDC informs the optimal channel assignment $\phi^*(j)$ to each SU $j \in \mathcal{K}_s$. Afterwards, each SU $j \in \mathcal{K}_s$ sends, at its maximum power $\bar{P}_j$, the value of $\beta_{j,\phi^*(j)}$ to its assigned primary user. The primary user decides whether to accept or reject the offer of cooperation, depending on how much power saving it can achieve through cooperative communications. The primary users who accept the offers will start the cooperative communications. Otherwise, the primary user will reject the offer and will keep transmitting over its licensed frequency band during the corresponding time frame.

The primary user i makes its cooperation decision as follows: It receives the bid from the j-th secondary at a received power level of $P^R_{i,j} = |h_{j,i}|^2 \bar{P}_j$. Then it may compute the received SNR it will get if the secondary j transmits at the bid power level of $\beta_{j,i}\bar{P}_j$ to be

$$\Gamma_{j,i} = \frac{\beta_{j,i} P^R_{i,j}}{(1-\alpha_i) W_i N_i}. \tag{5}$$

The i-th primary RX then uses either the Maximum-Ratio Combining (MRC), Maximum-SNR Selection or Coherent Relay detection to compute the resulting overall SNR, if it combines the received signals from both paths: direct path from the i-th primary TX itself and the relayed path from the secondary node j. To be specific, in the followings we will consider coherent relay detection. Denote by $R^f_i(P_i)$ the resulting final primary rate if the primary user i transmits at a power P_i:

$$R^f_i(P_i) = (1-\alpha_i) W_i \log\left(1 + \frac{P_i|h_i|^2 + \bar{P}_j \beta_{j,i} |h_{j,i}|^2}{(1-\alpha_i) N_i W_i}\right). \tag{6}$$

Let $P^{(\min)}_i(\beta_{j,i})$ be the minimum transmit power the i-th transmitter needs to transmit at to achieve $R^f_i(P_i) \geq R^{(\min)}_i$ if it accepts the j-th SU's bid for relaying:

$$P^{(\min)}_i(\beta_{j,i}) = \bar{P}_i \wedge \left[\frac{(1-\alpha_i) N_i W_i \bar{\gamma}_i(\alpha_i) - \bar{P}_j \beta_{j,i} |h_{j,i}|^2}{|h_i|^2}\right]^+, \tag{7}$$

where $\bar{\gamma}_i(\alpha_i) = 2^{\frac{R^{(\min)}_i}{(1-\alpha_i)W_i}} - 1$, $x \wedge y \triangleq \min\{x, y\}$ and $[x]^+ = \max\{0, x\}$. Then, the primary user i decides to cooperate with SU j if $P^{(\min)}_i(\beta_{j,i}) \leq P^{th}_i \leq \bar{P}_i$. Note that, in the above computations, the primary RX assumes that the channel from primary TX to the secondary relay is error-free. This can be a reasonable assumption under many scenarios [22], and it is also possible to modify the above method to take into account such error at the expense of additional system complexity. In general, the channel between the primary and secondary transmitters is not ideal, therefore, (6) can be considered as an upper bound on the primary transmission rate, which is still a valid criterion for making the DSL channel allocation in the general case.

The SU's who did not get the chance to access a channel might send new offers $\beta_{j,i}$ to the SSDC in the following *time slots* and the SSDC computes the optimal channel assignment based on the new $\beta_{j,i}$ values. Thus, the SU's in the centralized CRN may learn to increase their action variables $\beta_{j,i}$ within a time frame hoping that the primary users would accept the new offers. Conversely, a SU $j \in \mathcal{K}_s$ that has accessed a channel $i \in \mathcal{C}$ in a given time slot might decrease its $\beta_{j,i}$ value in order to reduce the power spent on relaying the primary signal. We refer to this model as the *centralized CRN with learning*, in contrast with the above described *static centralized CRN* in which the SU's fix $\beta_{j,i}$ within a time frame.

In a centralized CRN with learning, at the beginning of each time frame, the SU's $j \in \mathcal{K}_s$ determine their bids $\beta_{j,i} \in [0, \beta^{(\max)}_{j,i}]$ for $\forall i \in \mathcal{C}$. In the subsequent time slots, the SU's update their $\beta_{j,i}$ values and the SSDC computes the optimal assignment of SU's to available primary channels (from (4)) based on the new $\beta_{j,i}$ values, during each of the time slots. The new bids are sent to the primary users who then will decide whether to accept or reject those offers. The accepted SU's start asymmetric cooperation based transmission on their assigned channels. At each time slot, the SU's apply a simple reinforcement learning algorithm to update their $\beta_{j,i}$ values as follows:

$$\begin{aligned} \text{Winning Node: } \beta^{(new)}_{j,i} &= \beta_{j,i} - \mathcal{I}_{E(j,i)} \Delta\beta, \text{ for } \beta^{(new)}_{j,i} > 0 \\ \text{Losing Node: } \beta^{(new)}_{j,i} &= \beta_{j,i} + \Delta\beta, \text{ for } \beta^{(new)}_{j,i} \leq \beta^{(\max)}_{j,i}, \end{aligned} \tag{8}$$

where $\Delta\beta > 0$ is some step size and $\mathcal{I}_{E(j,i)}$ is the indicator function of the event $E(j,i) = \{$SU j has never lost a bid on channel i in the current time frame$\}$. This reinforcement learning algorithm converges to fixed $\beta_{j,i}$ values after a sufficient number of time slots.

IV. An Auction-Based DSL Protocol for Autonomous SU's

The channel access in a decentralized CRN is based on the competition among the SU's. This competition can be formulated as an auction game in which each SU $j \in \mathcal{K}_s$ places a bid $\beta_{j,i}$ for each primary channel $i \in \Omega_j$. After computing its set of bids $\{\beta_{j,i}\}_{i\in\Omega_j}$, each SU j sends these values to corresponding primary receivers (or could be the same receiver) at its maximum power level $\bar{P}_j$.

Receivers of each primary link will be responsible for determining the winning bid for that channel. If there are any ties among SU's for winning a particular channel, then the corresponding primary user will randomly pick one of them. Each primary user then informs its chosen winning secondary node of its bid being successful. In some cases, a particular SU's bids may be selected by more than one primary channel as winning bids. Then this SU decides to accept the invitation to cooperate with the primary channel that permits it to achieve the largest secondary rate. The remaining channels do not lease their spectrum to any user, thus encouraging the losing SU's to increase their bids on those particular channels in the next time slot. Once the bid selection is done, then the primary and winning SU's start to transmit based on the asymmetric cooperative communications, as shown in Fig. 3.

The primary users will only recompute their α_i values only at the beginning of a frame, since channel conditions are assumed to be almost constant over the duration of a frame. However, at every time slot, each primary user may adapt its freed-up channel portion α_i so that it receives bids to cooperate from more SU's. Similarly, the SU's are free to place new bids at the beginning of each time slot within a given frame. This allows each SU to revise its bids in order to maximize its chance of getting access to the most favorable channel, while minimizing the relay power $\beta_{j,i}\bar{P}_j$ it needs to spend. Thus, during each frame the primary-secondary interaction can be modeled as a repeated auction game as follows:

1) *Players*: L primary TX-RX pairs on L channels and K_s SU's.
2) *Actions*: Primary TX-RX pair i's action is to choose $\alpha_i \in [0, \alpha_i^{(\max)}]$ such that it satisfies the primary transmission rate requirements. Each SU j's action is to choose a set of power division ratios $\beta_{j,i}$'s, for each $i \in \Omega_j$.

Each SU will aim to transmit at the lowest $\beta_{j,i}$ possible. However, this might reduce its chances in gaining channel access because a primary user would prefer a SU that is willing to spend as much power as possible for relaying its signal so that it minimizes its transmit power P_i.

A. Selection of Winning Bids for Cooperative Communications

The objective of primary users in the proposed D-DSL framework is to minimize their own power expenditure by exploiting cooperative communications facilitated by SU's. This objective is achieved by maximizing the primary utility function:

$$u_i\left(\alpha_i, \beta_{j(i),i}\right) = \frac{\bar{P}_i - P_i(\beta_{j(i),i})}{\bar{P}_i} q\left(R_i(\alpha_i) - R_i^{(\min)}\right), \tag{9}$$

where $P_i(\beta_{j(i),i})$ is the primary i's transmit power with $P_i(\beta_{0,i}) = \bar{P}_i$ indicating that if primary i does not reach an agreement with any SU then it will be transmitting at its maximum power. The i-th primary receiver then chooses the SU j that will lead to the smallest $P_i^{(\min)}$ (given in (7)) such that $R_i^f(P_i) \geq R_i^{(\min)}$ as the winning bid for asymmetric cooperation on its channel, such that $j(i) \triangleq j^* = \arg\min_{j\in\mathcal{K}_s} P_i^{(\min)}(\beta_{j,i})$. The winning bid selection simplifies to:

$$j(i) \triangleq j^* = \arg\max_{j\in\mathcal{K}_s} \beta_{j,i}\bar{P}_j|h_{j,i}|^2. \tag{10}$$

B. Repeated Auction Game Model for D-DSL with Reinforcement Learning

In the subsequent plays of the repeated game, if the channel conditions stay fixed, the SU's can learn the others strategies and try to win the auction for spectrum leasing. At the beginning of each time slot, primary users take new bids $\beta_{j,i}$. The SU's update their bids again using the simple reinforcement learning strategy given in (8). However, note that in this case each individual SU updates its own bid $\beta_{j,i}$ independently of other secondary users.

On the other hand, the primary users also learn and adapt their actions α_i at every time step. Primary users take distinct actions depending on whether a SU was selected for cooperation or not: A primary user who did not get a SU to cooperate with will try to increase its α_i values so that more SU's becomes interested in cooperating with it, and vice versa. The primary learning algorithm is as follows:

$$\begin{aligned} \text{Coop.:}\ \alpha_i^{(new)} &= \alpha_i - \Delta\alpha \ \text{ for } \ \alpha_i^{(new)} > 0 \\ \text{No Coop.:}\ \alpha_i^{(new)} &= \alpha_i + \Delta\alpha \ \text{ for } \ \alpha_i^{(new)} \leq \alpha_i^{(\max)}, \end{aligned}$$

where $\Delta\alpha > 0$ is some step size. However, when the primary users are adapting their α_i according to the secondary actions, the values of α_i might decrease and thus, the sum rate of SU's might decrease as well. For that reason, we assume that primary users learn with a probability $\zeta \in [0,1]$, meaning that they adapt their actions in each time slot (within a *learning period*) with a probability ζ. The *learning period* consists of K_l time slots at the beginning of a time frame. The α_i values are not supposed to change outside a *learning period*.

C. Equilibrium of the Reinforcement Learning in the D-DSL auction game

Given our proposed D-DSL model, we observe that the auctions are independent among all the primary channels.

Thus, we can analyze the equilibrium on each channel $i \in \mathcal{C}$ separately. Obviously, the $\beta_{j,i}$ values may converge only if $\{\alpha_i\}_{i\in\mathcal{C}}$ are fixed during a certain period.

By applying the reinforcement learning algorithm in (8), and after sufficiently many time steps, the winning SU on channel $i \in \mathcal{C}$ will be (at equilibrium):

$$\hat{j}(i) \triangleq \hat{j} = \arg\max_{j\in\mathcal{K}_s} \beta_{j,i}^{(\max)} \bar{P}_j |h_{j,i}|^2 \ . \tag{11}$$

In this case, the equilibrium point is obtained as:

$$\hat{\beta}_{\hat{j},i} = \max_{k\neq\hat{j},k\in\mathcal{K}_s} \beta_{k,i}^{(\max)} \frac{\overline{P}_k |h_{k,i}|^2}{\overline{P}_{\hat{j}} |h_{\hat{j},i}|^2} + \Delta\beta - \delta, \tag{12}$$

for some $\delta \in [0, \Delta\beta]$. Also, at equilibrium, $\hat{\beta}_{k,i} = \beta_{k,i}^{(\max)}$ for all $k \neq \hat{j}$.

Moreover, if $\Delta\beta \to 0$, then $\hat{\beta}_{\hat{j},i} \to \max_{k\neq\hat{j},k\in\mathcal{K}_s} \beta_{k,i}^{(\max)} \frac{\overline{P}_k |h_{k,i}|^2}{\overline{P}_{\hat{j}} |h_{\hat{j},i}|^2}$. In this case, it can be easily verified that the equilibrium point of the reinforcement learning algorithm converges to the Nash equilibrium of a second-price auction [26].

V. Performance Results

To verify the convergence of the proposed asymmetric cooperative communications based D-DSL framework implemented as an auction game, we consider primary system with $L = 3$ and a secondary system having $K_s = 5$ users. We also assume Rayleigh fading channels with $\mathbb{E}\{|h|^2\} = 1$. The maximum transmit power of the primary and SU's are $\bar{P}_i = 30mW$ and $\bar{P}_j = 30mW$, respectively. We assume that all channels have a bandwidth ($W_i = 10kHz$), and the noise level at the receivers is $N_i = 0.1\mu W/Hz$. The minimum transmission rate requirement of a primary user is set to $R_i^{(\min)} = 10kbps$ and we assume that SU's require a BER smaller than $\epsilon = 0.05$. The primary system is assumed to be static by having $\zeta = 0$, and we set the step size of the secondary learning algorithm to $\Delta\beta = 0.02$. First, we show in Fig. 5 the convergence of the secondary action variables $\beta_{j,i}$ as a function of time, over 3 time frames with 50 slots each.

In Fig. 6, we let $L = 3$ and $K_s = 3$ and plot the secondary sum-rate and the average per-user primary power as a function of $\bar{P}_j$, for both centralized and decentralized CRN's, implemented based on either a static or a learning framework. Note that the static scenario refers to setting α_i to $\alpha_i^{(\max)}$ and $\beta_{j,i}$ to $\beta_{j,i}^{(\max)} \left(1 - e^{-aT_j^{(res)}}\right)$ during the whole frame duration, where $a = \frac{1}{30*\text{number of slots/frame}}$ and $T_j^{(res)}$ is the total remaining number of slots at the beginning of a frame. In each of the centralized or decentralized CRN's, the learning permits the secondary network to achieve a higher sum-rate, compared to the static scenario. Moreover, in either case, we observe that the centralized CRN outperforms the decentralized CRN only if $\bar{P}_j \geq 50mW$. This is because when the secondary system is centralized, the i-th primary user will agree to cooperate with a SU only if cooperation leads to it being able to transmit at a power level less than a certain threshold level of P_i^{th}. Since P_i^{th} is unknown to the SU's, this forces them to allocate at least a minimum amount of power to relay the primary message, if they are to have any chance

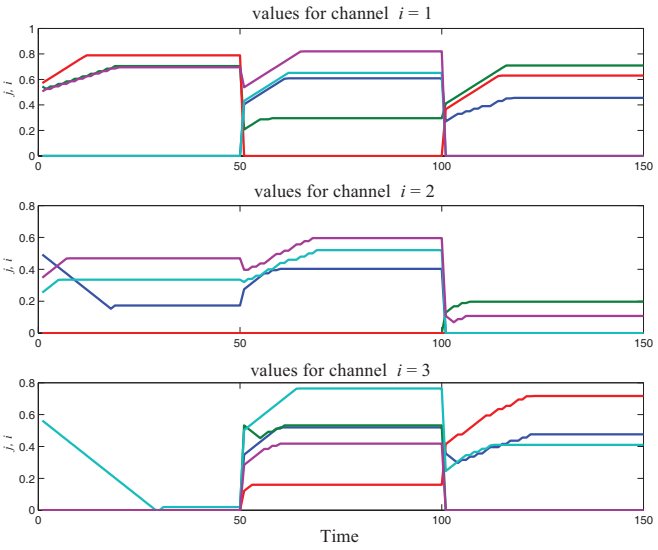

Fig. 5. $\beta_{j,i}$ values for $L = 3$, $K_s = 5$ and 3 time frames.

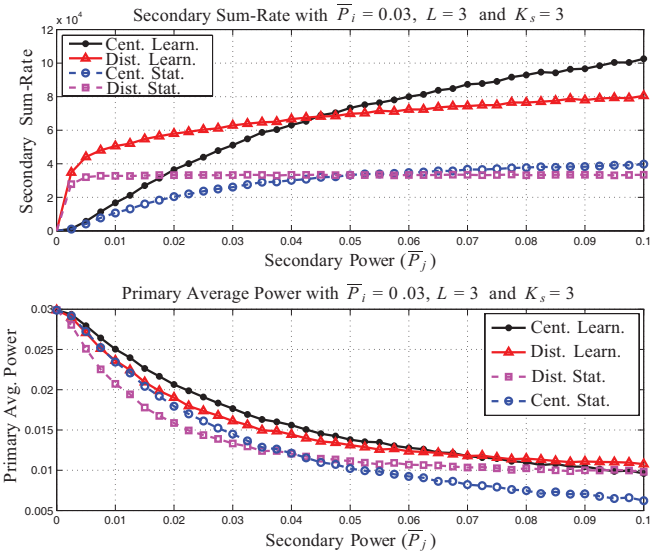

Fig. 6. Primary and secondary performance vs. $\bar{P}_j$.

of winning access to the i-th primary channel, for $i \in \mathcal{C}$. Therefore, the SU's will have an upper bound on the amount of power that they can use to transmit their own signals. However, if the secondary network is decentralized, the minimum power allocated to relay the primary message is only based on the competition among SU's. In this case, the primary users do not have the expectation to reduce their own transmit power below a hard threshold (such as P_i^{th} for $i \in \mathcal{C}$). Instead, they accept the best bid from the competing SU's, irrespective of how much small this bid is. This difference makes both the primary power savings as well as the secondary sum-rate to be lower in the centralized case compared with the decentralized case when the secondary power $\bar{P}_j$ is very small.

Figure 6 also shows the average power spent by each primary user in each of the above mentioned scenarios. The average power spent by a primary user will be the highest, if the CRN is equipped with learning capabilities, in either the centralized or the decentralized case. In a CRN with learning,

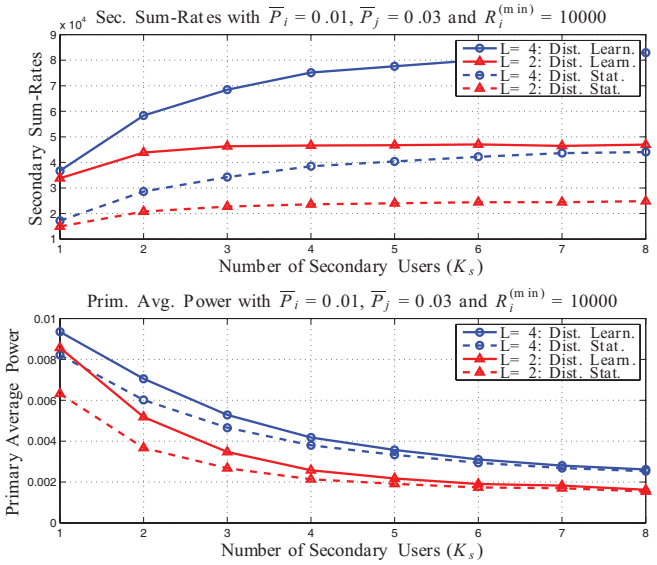

Fig. 7. Primary and secondary performance vs. K_s.

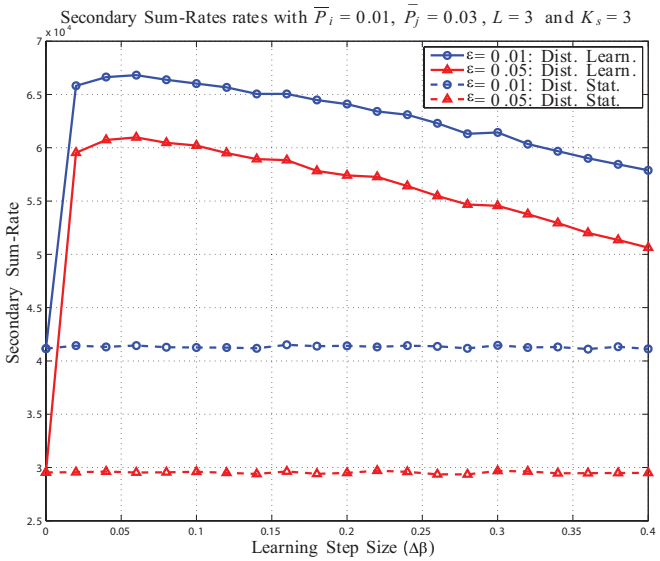

Fig. 8. Secondary sum rates vs. $\Delta\beta$.

the SU's learn to allocate just the minimum necessary power to relay the primary signals, while still gaining access to the primary channels. However, if the CRN is static, the SU's try to place high $\beta_{j,i}$ values at the beginning of the time frame because they will not have another chance to adapt their action variables within the same time frame. As a result, the primary users will take advantage of the static behavior of the CRN and achieve higher power savings. On the other hand, Fig. 7 shows that the sum-rate of SU's increases with the number of SU's (K_s) because of the increased diversity. It shows also that a significant gain can be achieved in the secondary sum-rate when the SU's employ reinforcement learning.

In Fig. 8 we plot the the secondary sum-rate versus learning step-size $\Delta\beta$. We observe that the secondary sum-rate reaches a maximum near $\Delta\beta = 0.06$. Note that, a small step size could slow down the convergence to the optimal point and a large step size makes $\beta_{j,i}$ to deviate from the equilibrium of the second-price auction. Therefore, $\Delta\beta$ should be carefully

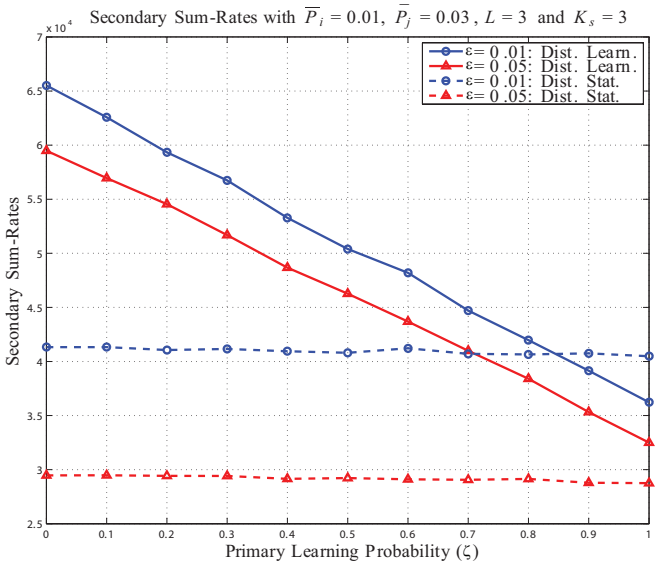

Fig. 9. Secondary sum rates vs. primary learning probability ζ.

adjusted in order to take advantage of the learning procedure.

Finally, in Fig. 9, we allow the primary users to adapt their actions during the first 25% of each time frame, and we show the effect of the primary learning on the secondary throughput. We see that the secondary performance degrades as the primary users try to learn more frequently. In fact, the primary learning procedure allows the primary users to decrease α_i, which reduces the available bandwidth for secondary transmission. Of course, it is more advantageous for SU's to cooperate with a static non-adaptive primary system, which will facilitate the adaptation of SU's to their environment and prevent them from being exploited by the primary users.

VI. Conclusion

In this paper, we have proposed both centralized and distributed Dynamic Spectrum Leasing architectures that allow primary users to reduce their power expenditure by using the SU's as relay nodes. In return for this asymmetric cooperative communication gains, primary users free-up a portion of their spectrum resources to SU's. In the centralized case, we derived the optimal channel assignment of SU's by using the Hungarian algorithm. Also, we developed a repeated auction game for D-DSL, in which the autonomous SU's learn by interacting with their environment so that they distributively reach an equilibrium. We proposed a reinforcement learning algorithm for both primary and SU's to learn and revise their actions. Our simulation results showed that the proposed reinforcement learning permits to enhance the performance of both centralized and decentralized CRN's.

References

[1] S. K. Jayaweera and T. Li, "Dynamic spectrum leasing in cognitive radio networks via primary-secondary user power control games," *IEEE Trans. Wireless Commun.*, vol. 8, no. 6, pp. 3300–3310, July 2009.

[2] S. K. Jayaweera, G. Vazquez-Vilar, and C. Mosquera, "Dynamic spectrum leasing: a new paradigm for spectrum sharing in cognitive radio networks," *IEEE Trans. Veh. Technol.*, vol. 59, no. 5, pp. 2328–2339, May 2010.

[3] J. Bae, E. Beigman, R. Berry, M. L. Honig, H. Shen, R. Vohra, and H. Zhou, "Spectrum markets for wireless services," in *Proc. 3rd IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks*, Oct. 2008, pp. 1–10.

[4] I. Stanojev, O. Simeone, Y. Bar-Ness, and T. Yu, "Spectrum leasing via distributed cooperation in cognitive radio," in *Proc. IEEE International Conference on Communications*, May 2008, pp. 3427–3431.

[5] J. Mitola, "Cognitive radio: an integrated agent architecture for software defined radio," Ph.D. dissertation, Royal Institute of Technology (KTH), Stockholm, Sweden, 2000.

[6] S. K. Jayaweera and C. Mosquera, "A dynamic spectrum leasing (DSL) framework for spectrum sharing in cognitive radio networks," in *43rd Annual Asilomar Conf. on Signals, Systems and Computers*, Nov. 2009.

[7] K. Hakim, S. K. Jayaweera, G. El-Howayek, and C. Mosquera, "Efficient dynamic spectrum sharing in cognitive radio networks: centralized dynamic spectrum leasing (C-DSL)," *IEEE Trans. Wireless Commun.*, vol. 9, no. 9, pp. 2956–2967, Sep. 2010.

[8] S. Toroujeni, S.-S. Sadough, and S. Ghorashi, "Time-frequency spectrum leasing for OFDM-based dynamic spectrum sharing systems," in *Proc. 6th Conference on Wireless Advanced*, June 2010, pp. 1–5.

[9] O. Simeone, I. Stanojev, S. Savazzi, Y. Bar-Ness, U. Spagnolini, and R. Pickholtz, "Spectrum leasing to cooperating secondary ad hoc networks," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 1, pp. 203–213, Jan. 2008.

[10] P. Lin, J. Jia, Q. Zhang, and M. Hamdi, "Dynamic spectrum sharing with multiple primary and secondary users," in *Proc. IEEE International Conference on Communications*, May 2010, pp. 1–5.

[11] M. Bkassiny and S. K. Jayaweera, "Optimal channel and power allocation for secondary users in cooperative cognitive radio networks," in *Special Session on Advanced Radio Access Techniques for Energy-Efficient Communications in 2nd International Conference on Mobile Lightweight Wireless Systems*, May 2010.

[12] O. Simeone, J. Gambini, Y. Bar-Ness, and U. Spagnolini, "Cooperation and cognitive radio," in *Proc. IEEE International Conference on Communications*, June 2007, pp. 6511–6515.

[13] W. Su, J. Matyjas, and S. Batalama, "Active cooperation between primary users and cognitive radio users in cognitive ad-hoc networks," in *Proc. IEEE International Conference on Acoustics Speech and Signal Processing*, Mar. 2010, pp. 3174–3177.

[14] Y. Li, S. K. Jayaweera, M. Bkassiny, and K. A. Avery, "Optimal myopic sensing and dynamic spectrum access in centralized secondary cognitive radio networks with low-complexity implementations," in *IEEE 73rd Vehicular Technology Conference (VTC Spring 2011)*, May 2011.

[15] H. Sun, D. Laurenson, J. Thompson, and C.-X. Wang, "A novel centralized network for sensing spectrum in cognitive radio," in *Proc. IEEE International Conference on Communications*, May 2008, pp. 4186–4190.

[16] C. Zhao, M. Zou, B. Shen, B. Kim, and K. Kwak, "Cooperative spectrum allocation in centralized cognitive networks using bipartite matching," in *Proc. IEEE Global Telecommunications Conference*, Dec. 2008, pp. 1–6.

[17] K.-L. Yau, P. Komisarczuk, and P. Teal, "Applications of reinforcement learning to cognitive radio networks," in *Proc. 2010 IEEE International Conference on Communications Workshops*, May 2010, pp. 1–6.

[18] H. Li, "Multi-agent Q-learning of channel selection in multi-user cognitive radio systems: a two by two case," in *Proc. IEEE International Conference on Systems, Man and Cybernetics*, Oct. 2009, pp. 1893–1898.

[19] N. Baldo and M. Zorzi, "Learning and adaptation in cognitive radios using neural networks," in *Proc. 5th IEEE Consumer Communications and Networking Conference*, Jan. 2008, pp. 998–1003.

[20] Q. Zhao, L. Tong, A. Swami, and Y. Chen, "Decentralized cognitive MAC for opportunistic spectrum access in ad hoc networks: a POMDP framework," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 3, pp. 589–600, Apr. 2007.

[21] T. Su, H. Ling, and W. Vogel, "Markov modeling of slow fading in wireless mobile channels at 1.9 ghz," *IEEE Trans. Antennas Propag.*, vol. 46, no. 6, pp. 947–948, June 1998.

[22] N. Devroye, M. Vu, and V. Tarokh, "Cognitive radio networks," *IEEE Signal Process. Mag.*, vol. 25, no. 6, pp. 12–23, Nov. 2008.

[23] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity—part I: system description," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1927–1938, Nov. 2003.

[24] H. W. Kuhn, "The Hungarian method for the assignment problem," *Naval Research Logistics Quarterly*, vol. 2, pp. 83–97, Mar. 1955.

[25] Y.-J. Choi, J. Kim, and S. Bahk, "Downlink scheduling with fairness and optimal antenna assignment for MIMO cellular systems," in *Proc. IEEE Global Telecommunications Conference*, vol. 5, Nov. 2004, pp. 3165–3169.

[26] V. Krishna, *Auction Theory*, 2nd edition. Academic Press, 2010.

Sudharman K. Jayaweera (S'00, M'04, SM'09) received the B.E. degree in Electrical and Electronic Engineering with First Class Honors from the University of Melbourne, Australia, in 1997 and M.A. and Ph.D. degrees in Electrical Engineering from Princeton University in 2001 and 2003, respectively. He is currently an assistant Professor in Electrical Engineering at the Department of Electrical and Computer Engineering at University of New Mexico, Albuquerque, NM. From 2003-2006 he was an assistant Professor in Electrical Engineering at the Department of Electrical and Computer Engineering at Wichita State University. Dr. Jayaweera is currently an associate editor of *EURASIP Journal on Advances in Signal Processing*. His current research interests include cooperative and cognitive communications, information theory of networked-control systems, statistical signal processing and wireless sensor networks.

Mario Bkassiny (S'06) received the B.E. degree in Electrical Engineering and the M.S. degree in Computer Engineering both from the Lebanese American University, Lebanon, in 2008 and 2009, respectively. He is currently working towards the Ph.D. degree in Electrical Engineering at the Communication and Information Sciences Laboratory (CISL), Department of Electrical and Computer Engineering at the University of New Mexico, Albuquerque, NM, USA. His current research interests are in cognitive radios, distributed learning and reasoning, cooperative communications, dynamic spectrum leasing and game theory.

Keith A. Avery is the Program Lead for the Integrated Microsystems program at the Air Force Research Laboratory focusing on advanced packaging and optoelectronics for space. He received his BS degree from DeVry Institute of Technology in 1983. For the first 12 years of his career he worked in the commercial sector designing digital and analog circuits for commercial, industrial, and telephony applications. Prior to joining AFRL he worked as a government contractor performing design activities for space experiments, advanced packaging techniques, and radiation effects on micro-electronics. During his career he has increased his level of responsibility for design activities and program management. He has authored or co-authored numerous papers on designs for space and radiation effects. Mr. Avery is a member of IEEE/NPSS and AIAA.