Short Paper: On Optimal Sensing and Transmission Strategies for Dynamic Spectrum Access

Senhua Huang, Xin Liu, and Zhi Ding
University of California Davis
Davis, CA 95616, USA
Email: senhua@ece.ucdavis.edu

***Abstract*—The listen-Before-Talk (LBT) strategy has been prevalent in cognitive radio networks where secondary users opportunistically access under-utilized primary band. To minimize the amount of disruption from secondary users to primary signals, secondary users generally are required to detect the presence of the primary user reliably, and access the spectrum intelligently. The sensing time has to be long enough to achieve desirable detection performance. Weaker primary signals require longer sensing time, thereby reduce the secondary transmission opportunities. In this paper, we generalize the packet-level LBT strategy by allowing the secondary user to potentially transmit multiple packets after one sensing, and study the optimal control policy to determine the conditions under which the secondary user should sense the channel. We show that the optimal spectrum access control policy has a simple threshold-based structure, where the secondary user transmits consecutive packets until the estimated probability of the primary user being idle falls below a threshold, and senses the channel otherwise. The result applies to systems with both perfect and imperfect packet collision detection with the primary users.**

I. INTRODUCTION

We consider opportunistic spectrum access by secondary users to the spectrum allocated to a legacy user (primary user). The primary user (PU) has higher priority over the spectrum. Hence, the transmission from the secondary user (SU) should cause little interruption to the PU. Note that the SU and the PU may belong to the same spectrum owner. However, the PU may be unable or unwilling to update its hardware and/or software to facilitate the opportunistic spectrum access from the SU. It is the responsibility of the SU to discover the idle spectrum bands, and access the spectrum resources intelligently such that the interference to the PU is minimized. Therefore, the spectrum access of the SU often adopts the Listen-Before-Talk (LBT) principle, according to which the SU senses the channel before transmission. While this principle is well justified, many existing works on opportunistic spectrum access assume a typical packet-level sensing model (as illustrated in Fig. 1) implicitly or explicitly, i.e., before transmitting every packet, the SU senses the PU channel for a fixed amount of time [1], [2], [3]. The sensing cost in terms of time overhead is often neglected in the problem formulation for packet-level sensing.

While packet-level sensing simplifies the design of PHY/MAC layer protocol, it is not necessarily optimal, especially when sensing may consume too much transmission time. For example, weak PU signal requires a large number of data samples to achieve a good detection performance that is crucial to the protection of the QoS of the PU. Alternatively, when cooperative sensing schemes are used by a group of SUs, the time spent on exchanging sensing information could be in the order of packet length. In these scenarios, the overhead introduced by fixed sensing time per transmission is no longer negligible.

Additionally, perfect collision detection is often assumed in previous works, where acknowledgment message from the secondary receiver can accurately reflect whether the SU packet collides with the PU. However, this is not true in practical systems, as it is difficult to detect the collision with the PU without error. The difficulties result from the captured effects at both primary receiver and secondary receiver due to random fluctuation of the desired signal power and interference power of other users.

Therefore, we study the following two questions:

- What is the optimal sensing-transmitting strategy? Is it optimal to perform LBT on packet level, especially when the sensing time is large?
- What is the impact of imperfect collision detection, which may not faithfully indicate the presence/absence of the PU during a secondary packet transmission?

Using the optimal stopping theory [4], we show that one optimal sensing and transmission policy has a simple threshold-based structure, where the posterior probability of the PU being idle is compared to the threshold. This result coincides with the heuristic that the SU should continue transmitting packets until the estimated idle probability falls below the threshold. For a special case, we derive the closed-form solution for the optimal control on the SU's spectrum access.

II. RELATED WORKS

Encouraged by the potential spectrum policy reform in FCC [5] and the progress made in DARPA XG program [6], the research field of cognitive radio has thrived in recent years.

Both centralized spectrum leasing/auction schemes (e.g. [7], [8]), channel probing and selection (e.g., [9]), and distributed channel allocation and scheduling algorithms (e.g. [10], [11], [12]) have been proposed to enable the spectrum sharing among users with cognitive radio capabilities in different network scenarios. For example, in [9], authors study how to select the best channel to probe and/or transmit among

---

This work is supported in part by the National Science Foundation Grant CNS-0520126 and Grant CNS-0448613.

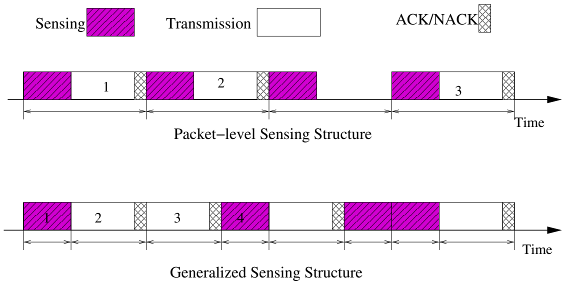

Fig. 1. Two Sensing Structures

a number of channels in order to maximize the reward of cognitive radios, and show that the optimal joint channel probing and transmission strategy has a threshold-based structure. However, they assume that the state of each channel remains unchanged during probing and transmission, which is not true for unslotted PU activities, and they do not consider the penalty on the collision with the PU. In [12], authors consider the cross-layer optimization on flow routing, scheduling, channel allocation, and power control of cognitive radios to improve the spectrum efficiency of multihop software defined networks. However, there is a lack of consideration on the interplay between PU behavior and SU access, which is important especially for the opportunistic spectrum access of the SUs within the spectrum interweave framework, where Listen-Before-Talk principle is emphasized.

The closest related works within the framework of spectrum interweaving include the joint PHY/MAC designs of channel selection, channel sensing/probing, operating points of spectrum sensor for cognitive radios (e.g. [1], [2], [3], [13], [14]). For slotted PU activities, where the PU can only change its state (idle/busy) on the boundary of time slots, partially observable MDP theory (POMDP) is used in [1], [13] to derive the structure for the optimal dynamic spectrum access policy (including channel selection, operating point of the spectrum sensor, and the access decision) with constraint on the collision probability observed by the PU. It was shown that myopic policy is often optimal for many cases by exploiting the inner structure of the spectrum access problem. Furthermore, authors in [14] show that the results in [13] can be extended to unslotted PU activities. However, they assume a packet-level LBT structure, and perfect indication of collision with the PU via acknowledgment mechanism between secondary transmitter and secondary receiver. In [3], a simple periodic sensing scheme with packet-level LBT structure for choosing primary channels with high idle probability is proposed for the SU to exploit the spectrum opportunities in multiple primary channels.

Our work differs on that we use a more flexible sensing/transmission structure based on LBT principle, and we do not assume that the collision detection is perfect with the acknowledgment mechanism. Instead, we consider the impact of imperfect collision detection on the control policy of the opportunistic spectrum access in our problem formulation, which applies to more realistic systems.

III. SYSTEM MODEL

We consider a system consisting of a primary link, and a secondary link that opportunistically accesses the PU channel. The PU activities follow an alternative ON-OFF pattern, i.e., its state is either busy or idle. The duration of the idle (denoted by V_p) and busy (denoted by L_p) states follows exponential distributions with means denoted by v_p and l_p, respectively. With this model, the state transition of the PU can be illustrated as in Fig. 2. The PU transmits its traffic at will. In other words, it does not perform any sensing functionality.

The SU uses a spectrum sensor to detect the status of the PU. The sensing time is denoted by T_s. We assume that sensing time is long enough such that the sensing is accurate. Note that the spectrum sensor at the SU is required to achieve reasonably good performance (especially on missed detection probability) to protect the PU in worst cases; hence the perfect sensing assumption here is justified. However, sensing outcome on current PU state cannot predict whether the PU will remain idle during the next time slot.

The access of the SU follows a slotted-structure, as shown in Fig. 3. We assume that the SU transmits a packet with a fixed length Δ, and that the duration of SU packet is much shorter than the PU busy/idle cycles, i.e., $\Delta \ll v_p$, $\Delta \ll l_p$. For example, the WLAN packets are about several ms, while the busy time of a voice call session is in seconds. Here, we also assume that the SU knows the PU idle/busy time distributions through measurement and estimation as in [15].

Upon receiving the packet from the secondary transmitter (SU-Tx), the secondary receiver (SU-Rx) may feedback an acknowledgment message. The acknowledgment here serves two purposes. First, it validates the packet transmission of the SU in the MAC layer. More importantly, it provides some information to the SU-Tx on whether a collision with the PU has occurred. In an ideal scenario in the sense of collision detection, a NACK (ACK) from the SU-Rx accurately signifies that a collision with the PU has (not) happened during last time slot.

However, under what is known as the captured effect, the SU-Rx may be able to decode the SU-Tx's packet even when the PT is transmitting in practical systems. Moreover, the lack of ACK could result from collision with the PU, deep SU channel fade, or interference from other users. The ACK itself is also possibly weak. Therefore, we view the acknowledgment as inaccurate (in the sense of collision detection), which is different from most of existing works. We define the following two probabilities:

$$\begin{aligned}\gamma_1 &= Pr[\text{NACK}|\text{Collision with PU}]\\ \gamma_0 &= Pr[\text{NACK}|\text{No Collision with PU}].\end{aligned}$$

Since the interference from the busy primary transmitter can only worsen the packet error rate of SU-Rx, we have $\gamma_1 > \gamma_0$. When the ACK/NACK of the SU faithfully reflects the collision result with the PU, we have $\gamma_0 = 0$ and $\gamma_1 = 1$.

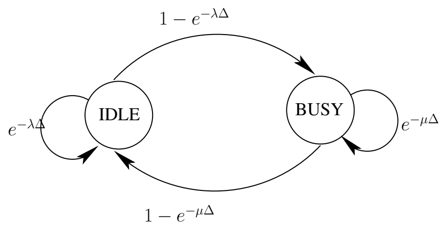

Fig. 2. Primary user's state transition

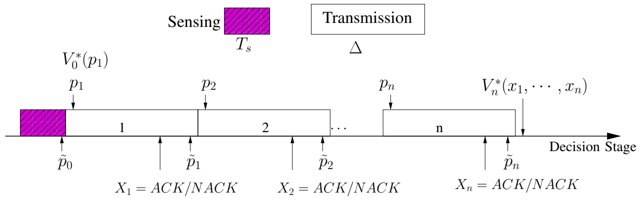

Fig. 3. Sequences of Spectrum Access

IV. PROBLEM FORMULATION

In order to exploit the spectrum opportunities, the SU-Tx decides dynamically whether to transmit a packet or sense the channel at each slot. Since sensing is one of two decision outcomes, the SU does not have complete information on the PU state (idle/busy).

Though it is natural to formulate this spectrum access problem as a partially observable Markov decision process (POMDP) problem, we find it easier to define the system state information as the posterior probability that the PU is idle and the problem becomes a MDP problem with fully observable state and uncountable state space. We use p_t to denote the state at time slot t, $p_t \in [0, 1]$. Since the SU's information about the PU state at current slot t depends on the information in slot $t-1$, the action a_{t-1}, and the observation result at slot $t-1$, p_t is a sufficient statistic to determine the optimal policy for the originally partially observable Markov decision process [16]. The action space of the SU is denoted by $\mathbb{A} = \{1(\text{Transmit}), 0(\text{Sense})\}$.

For each successful transmission, the SU receives a unit reward. Furthermore, because the PU has a higher priority on the spectrum resource, the SU will be charged a cost C for each packet collision with the PU. Obviously, without the collision penalty, the SU will always transmit if no other constraints is imposed. Thus, the collision penalty is important to control the aggressiveness of the SU's access activities. Then, the immediate reward $r_t(p_t, a_t)$ at time slot t with state p_t and action a_t is expressed as follows:

$$\begin{aligned} r_t(p_t, 1) &= p_t(1-\gamma_0) + (1-p_t)(1-\gamma_1 - C) \\ r_t(p_t, 0) &= 0. \end{aligned}$$

To prevent the extreme case that the SU never stops transmitting, we require that $\bar{p} - (1-\bar{p})C < 0$, where $\bar{p}$ is the limiting distribution of channel being idle. Since $\bar{p} = \frac{v_p}{v_p + l_p}$, we have $C > v_p/l_p$. On the other hand, C should be be too large to prevent any secondary transmission. Hence, we have:

$$(1 - e^{-\Delta/v_p}) \cdot 1 - (1 - e^{-\Delta/v_p})C > 0$$

which is equal to have $C < \frac{e^{-\Delta/v_p}}{1-e^{-\Delta/v_p}}$.

As shown in Fig. 3, at the end of decision slot t, the SU will update its estimation on the PU idle probability $\tilde{p}_t$ based on the observation it receives after its chosen action. The observation can be either the sensing outcome or the ACK/NACK. Specifically, we have the following observation model:

$$\begin{aligned} \tilde{p}_t^{(a_t=0)}(p_t) &= \begin{cases} 1, & \text{sensing “idle”} \\ 0, & \text{sensing “busy”} \end{cases} \\ \tilde{p}_t^{(a_t=1)}(p_t) &= \begin{cases} \tilde{p}_t(\text{ACK}), & \text{ACK} \\ \tilde{p}_t(\text{NACK}), & \text{NACK}, \end{cases} \end{aligned} \tag{1}$$

where

$$\begin{aligned} \tilde{p}_t(\text{ACK}) &= \frac{p_t(1-\gamma_0)}{p_t(1-\gamma_0) + (1-p_t)(1-\gamma_1)}, \\ \tilde{p}_t(\text{NACK}) &= \frac{p_t\gamma_0}{p_t\gamma_0 + (1-p_t)\gamma_1}. \end{aligned}$$

The system changes to state p_{t+1} based on the following rules:

$$p_{t+1} = \tilde{p}_t e^{-\frac{\Delta}{v_p}} + (1-\tilde{p}_t)(1 - e^{-\frac{\Delta}{l_p}}). \tag{2}$$

The access policy of the SU is to decide at each instant t whether to transmit or sense on the channel. The optimal sensing and transmission strategy maximizes the average reward over the whole access period, which consists of L repeated trials, i.e.,

$$\max \qquad \lim_{L\to\infty} \frac{(\sum_{l=1}^{L}\sum_{t=1}^{N_l} r_t)/L}{(\sum_{l=1}^{L}(T_s + N_l\Delta))/L}, \tag{3}$$

where N_l is the number of packets transmitted in the lth trial of the decision process. When it causes no confusion, we use N to denote the stopping rule that decides whether to stop transmission based on the current system state p_t. For stopping rule N, define

$$Y_N = \sum_{t=1}^{N} r_t, \qquad T_N = (T_s + N\Delta), \tag{4}$$

i.e., Y_n is the accumulated reward until stage n, T_n is the total time spent to reach stage n.

V. OPTIMAL SENSING AND TRANSMISSION STRATEGY

Since the exponential distribution is memoryless, the PU state transition probability does not depend on time (as illustrated in Fig. 2). Therefore, the SU always "restarts" the access from sensing the channel being idle, i.e., $\tilde{p}_0 = 1$. For given policy π, the expected sum of reward is identically and independently distributed (i.i.d.), as well as the number of packets transmitted in each run. Therefore, maximizing the average reward per unit of time is equal to maximizing the rate of return $E(Y_N)/E(T_N)$ [4], i.e.,

$$\lim_{L\to\infty} \frac{(\sum_{l=1}^{L} Y_{N_l})/L}{(\sum_{l=1}^{L} T_{N_l})/L} = \frac{E(Y_N)}{E(T_N)} \quad \text{almost surely.} \tag{5}$$

Therefore, the optimal spectrum access problem can be expressed as:

$$\max_{N\in\mathcal{C}} \qquad \frac{E(Y_N)}{E(T_N)}, \tag{6}$$

where

$$\mathcal{C} = \{N : N \geq 1, E(T_N) < \infty\} \tag{7}$$

is the set of stopping rules for which $E(T_N) < \infty$. Since $C > v_p/l_p$, the strategy of never stopping transmission is not optimal, and thus the optimal stopping rule always resides in $N \in \mathcal{C}$. The optimal average reward per time unit is then expressed as:

$$\alpha^* = \max_{N\in\mathcal{C}} \frac{E(Y_N)}{E(T_N)}. \tag{8}$$

Optimal stopping theory [4] is used to characterize the structure of the optimal stopping rule for the secondary spectrum access. Define

$$S_n(p) = Y_n - \alpha T_n, \tag{9}$$

where p is the initial state, and α can be regarded as a cost per time unit to reach stage n. According to *Theorem 6.1* in [4] , if for some α, $\sup_{N\in\mathcal{C}} E(S_N) = 0$, then $\sup_{N\in\mathcal{C}} E(Y_N)/E(T_N) = \alpha$. In addition, the policy which attains $\sup_{N\in\mathcal{C}} E(S_N) = 0$ achieves the maximum rate of return, i.e., α^*. Then, we translate the problem of maximizing rate of return to an ordinary stopping time problem with reward at stage n denoted by S_n. Define $V_0^*(p) = \sup_{N\in\mathcal{C}} E(S_N(p))$ as the maximum expected return given we start from state p. First, we show the existence of the optimal rule for the ordinary stopping time problem:

$$\max_{N\in\mathcal{C}} \qquad E(S_N). \tag{10}$$

Specifically, we have:

Proposition 1. [1] *There exists an optimal stopping rule N^* for problem (10).*

Relying on the optimality equation, the optimal rule has the following form:

$$N^* = \min\{n \geq 0 : S_n \geq V_n^*(x_1, \cdots, x_n)\}, \tag{11}$$

where $V_n^*(x_1, \cdots, x_n) = \sup_{N\geq n} S_N(x_1, \cdots, x_n)$ is the maximum expected reward given the observation $X_1 = x_1, \cdots, X_n = x_n$. Notice that, the optimal value obtained at decision stage n is determined by state p_{n+1} (as illustrated in Fig. 3), which abstracts all the observed information from $(x_1, x_2, \cdots, x_n)$. This indicates a time invariance property of the optimal value function, i.e., the expected payoffs at stage n after observing $X_1, \cdots, X_n$ is the same as it was at stage 0, except for an additional cost (or reward) to reach state p_{n+1}. Similar arguments can be found in [4]. Specifically, we have the following results:

$$\begin{aligned} V_n^*(X_1, \cdots, X_n) &= V_n^*(p_{n+1}) \\ &= V_0^*(p_{n+1}) - \alpha n\Delta + \sum_{t=1}^{n} r_i, \end{aligned} \tag{12}$$

where $-\alpha n\Delta + \sum_{t=1}^{n} r_i$ denotes the total "reward" accumulated upto stage n. Then, the rule given by the principle of optimality is reduced to the following form:

$$\begin{aligned} N^* &= \min\{n \geq 0 : S_n = V_n^*(p_{n+1})\} \\ &= \min\{n \geq 0 : S_n = V_0^*(p_{n+1}) - \alpha\Delta n - \alpha T_s\} \\ &= \min\{n \geq 0 : V_0^*(p_{n+1}) + \alpha T_s = 0\}. \end{aligned} \tag{13}$$

For general stopping time problem with uncountable state space, it is very difficult to find the structure of the optimal stopping rule using (13). In this section, we show that for the spectrum access model considered here, the optimal policy has a threshold-based structure as follows.

Theorem 1. The optimal stopping rule to maximize the rate of return is:

$$\pi^* : a_t = \begin{cases} 1(\text{Transmit}), & \text{if } p_t \geq p^* \\ 0(\text{Sense}), & \text{o.w.} \end{cases}, \tag{14}$$

where $p^* = \max\{p : V_0^*(p) + \alpha T_s = 0\}$.

The result is intuitive. When the channel is more likely to be idle ($> p^*$), the SU should continue packet transmission rather wasting spectrum opportunity on sensing. On the other hand, it worth mentioning that the typical challenge of finding optimal sequential decision is to balance between the immediate reward and all possible future payoffs. Without any structure in the optimal policy, it requires an exhaustive search over the set of all possible policies (which is practically impossible) to obtain the maximum throughput per time unit for the SU. However, with the shown well-defined structure for the optimal policy here, we can find the optimal value of p^* by searching over $p \in [0, 1]$.

For a special scenario, where no ACK/NACK exists to facilitate the SU-Tx to detect the collision with the PU, we can obtain a closed-form expression on the optimal sensing/transmission strategy for the dynamic spectrum access as in [17].

VI. DISCUSSION

We have shown that the optimal sensing/transmission strategy for the exponential idle time distribution is a single threshold-based policy. The criterion we adopted here is the maximum average reward (minus the collision penalty) per time unit. Since it is an unconstrained dynamic programming problem, the optimal strategy is a strategy without randomization. However, it is nontrivial to extend the results here to general distributions of PU idle time.

In some practical systems, the PU may impose strict requirement on the interruption from the SU. One such requirement is to limit the packet collision probability observed by the PU

---

[1] The proof of this proposition and the following theorem is removed due to space limit. Interested readers can find the proof in [17]

(denoted by p_p^c). Then, the design of dynamic spectrum access of the SU becomes a constrained optimization problem. In this case, we can show that the transmission strategy maximizing the successful transmission time of the SU subject to the PU packet collision probability constraint is also a threshold-based policy ([18]). To state the optimal policy, we first define the decision metric for the SU as:

$$g(t) = \frac{1 - F_{V_p}(t)}{f_{V_p}(t)}, \tag{15}$$

where $f_{V_p}(t)$ and $F_{V_p}(t)$ are the probability density function and cumulative density function of the PU idle time, respectively, and t is the time duration during which the PU has remained idle. In other words, $g(t)$ is the likelihood of successful transmission (without colliding with the PU) given that the PU has been idle for t. Then, we have the following result:

Theorem 2. For a given distribution $f_{V_p}(t)$ for the PU's idle time, the following listen-before-talk spectrum access policy is optimal under the collision probability constraint $p_p^c \leq \eta$:

$$a^*(t) = \begin{cases} 1, & \text{if} \quad g(t) > \gamma^* \\ 1 \quad \text{with probability } p^*, & \text{if} \quad g(t) = \gamma^* \\ 0, & \text{otherwise,} \end{cases}$$

where n_p is the average number of PU packets in a busy period, and the values of γ^* and p^* are determined by

$$\int_{\tau: g(\tau) > \gamma^*} f_{V_p}(\tau) d\tau + p^* \int_{\tau: g(\tau) = \gamma^*} f_{V_p}(\tau) d\tau = n_p \eta.$$

Randomization is required when $g(t) = \gamma^*$.

For exponential idle time distribution, it is optimal for SU to transmit with probability $p^* = n_p \eta$ when it detects the PU being idle, which is different from the policy we derive in this paper. Note that the result in Theorem 2 applies to cases where the SU can detect the collision with the PU after transmitting a packet. It is our current work to study the optimal spectrum access (constrained or unconstrained) for systems in which the SU cannot detect the collision with the PU perfectly, and the PU idle time follows a general distribution.

VII. CONCLUSIONS AND FUTURE WORKS

We generalize the Listen-Before-Talk sensing/transmission structure at the packet level. Given a reward mechanism for successful SU transmission and collision penalty, we develop an adaptive control policy to decide whether to sense/transmit in each decision stage. We also consider the impact of inaccurate collision detection with PU traffic on spectrum access policy. We found the optimal spectrum access policy to have a simple threshold-based structure. The optimal access decision is to continue transmitting a packet when the posterior idle probability of the PU is higher than the threshold, and to sense the channel otherwise. With this structure, the optimal policy can be found by simply searching for the optimal threshold, with which the computation complexity is greatly reduced. One possible direction for future work is to study the problem for general idle/busy time distributions. In addition, it is also of interest to develop an adaptive MAC protocol in response to the dynamic changing PU behavior.

REFERENCES

[1] Q. Zhao, L. Tong, A. Swami, and Y. Chen, "Decentralized cognitive MAC for opportunistic spectrum access in ad hoc networks: A POMDP framework," *IEEE Journal on Selected Areas in Communications (JSAC): Special Issue on Adaptive, Spectrum Agile and Cognitive Wireles Networks*, vol. 25, no. 3, pp. 589–600, 2007.

[2] L. Lai, H. E. Gamal, H. Jiang, and H. V. Poor. (2007) Cognitive medium access: Exploration, exploitation and competition. [Online]. Available: http://arxiv.org/abs/0710.1385

[3] Q. Zhao, S. Geirhofer, L. Tong, and B. M. Sadler, "Optimal dynamic spectrum access via periodic channel sensing," in *Proc. Wireless Communications and Networking Conference (WCNC)*, 2007.

[4] T. S. Ferguson. (2006) Optimal stopping and applications. [Online]. Available: http://www.math.ucla.edu/ tom/Stopping/Contents.html

[5] "Facilitating opportunities for flexible, efficient, and reliable spectrum use employing cognitive radio technologies, notice of proposed rule making and order," Federal Communications Commision, Report, Et docket No. 03-322, December 2003.

[6] F. W. Seelig, "A description of the August 2006 XG demonstrations at fort A.P. Hill," *Second IEEE International Symposium on Dynamic Spectrum Access Networks, DySPAN*, pp. 1–12, 2007.

[7] M. Buddhikot, P. Kolodzy, S. Miller, K. Ryan, and J. Evans, "DIMSUMnet: New directions in wireless networking using coordinated dynamic spectrum access," in *IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks (IEEE WoWMoM)*, aromina/Giardini Naxos, Italy, 2005.

[8] S. Gandhi, C. Buragohain, L. Cao, H. Zheng, and S. Suri, "A general framework for wireless spectrum auctions," *Second IEEE International Symposium on Dynamic Spectrum Access Networks, DySPAN*, pp. 22–33, 2007.

[9] N. B. Chang and M. Liu, "Optimal channel probing and transmission scheduling for opportunistic spectrum access," in *Proc. of the 13th annual ACM international conference on Mobile computing and networking*. New York, NY, USA: ACM, 2007, pp. 27–38.

[10] Y. Yuan, P. Bahl, R. Chandra, T. Moscibroda, and Y. Wu, "Allocating dynamic time-spectrum blocks in cognitive radio networks," in *Proc. of the 8th ACM international symposium on Mobile ad hoc networking and computing*. New York, NY, USA: ACM, 2007, pp. 130–139.

[11] J. Zhao, H. Zheng, and G.-H. Yang, "Distributed coordination in dynamic spectrum allocation networks," *First IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks*, pp. 259–268, 8-11 Nov. 2005.

[12] Y. T. Hou, Y. Shi, and H. D. Sherali, "Optimal spectrum sharing for multi-hop software defined radio networks," *in Proc. IEEE INFOCOM, Anchorage, AL*, pp. 1–9, 2007.

[13] Y. Chen, Q. Zhao, and A. Swami, "Joint design and separation principle for opportunistic spectrum access in the presence of sensing errors," *IEEE Transactions on Information Theory*, vol. 54, no. 5, pp. 2053–2071, 2008.

[14] Q. Zhao and K. Liu, "Detecting, tracking, and exploiting spectrum opportunities in unslotted primary systems," *in Proc. of IEEE Radio and Wireless Symposium (RWS)*, 2008.

[15] H. Kim and K. Shin, "Efficient discovery of spectrum opportunities with mac-layer sensing in cognitive radio networks," *IEEE Transactions on Mobile Computing*, vol. 7, no. 5, pp. 533–545, May 2008.

[16] R. Smallwood and E. Sondik, "The optimal control of partially observable markov processes over a finite horizon," *Operations Research*, pp. 1071–1088, 1971.

[17] S. Huang, X. Liu, and Z. Ding, "On optimal sensing and transmission strategies for dynamic spectrum access," UC Davis, Technical Report, August 2008. [Online]. Available: "http://www.ece.ucdavis.edu/˜senhua/TRdyspan08.pdf"

[18] ——, "Optimization of transmission strategies for opportunistic access in cognitive radio networks," UC Davis, Technical Report, April 2008. [Online]. Available: "http://www.ece.ucdavis.edu/˜senhua/TRconstraint.pdf"