

A Learning-based Multiuser Opportunistic Spectrum Access Approach in Unslotted Primary Networks

Sachin Shetty
ECE Department
Rowan University
Glassboro, NJ 08028
Email: shetty@rowan.edu

Min Song
ECE Department
Old Dominion University
Norfolk, VA 23517
Email: msong@odu.edu

Chunsheng Xin
CS Department
Norfolk State University
Norfolk, VA 23504
Email: cxin@nsu.edu

E. K. Park
CSEE Department
University of Missouri at Kansas City
Kansas City, MO 64110
Email: ekpark@umkc.edu

Abstract—Opportunistic spectrum access presents a new approach to wireless spectrum utilization and management. In this paper, we propose a non-cooperative based OSA approach: *learning-based approach* to allow multiple secondary users to achieve maximal throughput in an unslotted opportunistic spectrum access (OSA) network. In this approach, collisions among secondary users are taken into consideration while making channel sensing decisions. Spectrum maps for secondary users are estimated based on occurrence of collisions. Our approach allows secondary users to achieve maximal throughput by seeking independent spectrum opportunities without exchanging any control information among secondary users. Numerical results show that the learning-based approach obtains near-optimal performance in most of the scenarios.

I. INTRODUCTION

The proliferation of a wide range of wireless devices has resulted in an overly crowded radio spectrum. In contrast to this scarcity in spectrum availability is the pervasive existence of spectrum opportunities. Real measurements show that, at any given time and location, a large portion of licensed spectrum lies unused [1]. To exploit the abundant spectrum opportunities, cognitive radio networks have been proposed as a novel approach to improve spectrum utilization. Opportunistic spectrum access (OSA) is one of the approaches envisioned for dynamic spectrum management in cognitive radio networks [2]. The basic idea of OSA is to allow secondary users to identify and exploit spectrum opportunities under the constraint that they do not cause harmful interference to primary users.

Most of the existing works on OSA strategies assume the presence of a slotted primary network [2]–[9]. These OSA strategies maximize the throughput of an individual secondary user in a multi-channel slotted primary network. In presence of multiple channels, a key decision for every secondary user is to determine which channel to sense. With multiple secondary users contending for spectrum opportunities, the sensing decision must take into account the possibility that the good channels may be desired by other users. At the beginning of every slot, every secondary user senses an idle primary channel to potentially transmit over. Based on the sensing decision in the beginning of the slot, secondary users decide if the channel will be idle for the remaining of the slot. However, if the primary users adopt an unslotted transmission scheme,

the channel can switch between idle and busy state at any time. This leads to potential collisions with the primary users or missed opportunities in spite of perfect sensing during the sensing period. The secondary users need to make a decision to transmit based on its sensing decision. The secondary users should minimize missed opportunities due to inaccurate detection outcomes and limit collisions with the primary users.

In this paper, we design a learning-based OSA approach to enable secondary users maximize their spectrum opportunities in unslotted primary networks. Our learning-based approach explores the interaction between secondary users to make the channel sensing decision. Specifically, each secondary user takes into consideration a partial spectrum map of colliding secondary users. The partial spectrum map is obtained by estimating colliding secondary user's channel belief vector. The belief vectors are implicitly learned and updated based on the occurrence of collisions. This is an important property as secondary users cannot completely ignore the presence of other secondary users. The channel sensing decision is more effective in every subsequent time slots as each secondary user learns from the collision information. Collisions provide an approximate estimate of the level of activity for secondary users in the channels. Another important property of our learning-based approach is that there is no explicit exchange of spectrum maps between secondary users.

The rest of the paper is organized as follows. Section 2 briefly reviews the related work. We formulate the problem and provide the network model in Section 3. In Section 4, we introduce the learning-based approach and conduct the performance analysis. The simulation results are presented in Section 5. Finally, Section 6 provides the concluding remarks.

II. RELATED WORK

Research on OSA strategies have primarily been conducted in the following three scenarios: single-user setting, multiuser setting, and unslotted primary networks. In [3], [5], the strategy for a single secondary user to maximize total throughput in a slotted primary network is proposed. However, in real OSA networks, multiple secondary users opportunistically seek spectrum access from primary users. In [4], [6], multiple secondary users reserve channels by sending control messages on coordination channels. The presence of a common control

channel is only advantageous in networks where there are high availability of unused channels. Finally, there have been very few research efforts on OSA strategies for unslotted primary networks [7], [8]. But both [7], [8] do not model contention among multiple secondary users. In this paper, we propose a learning based approach which does not require cooperation or coordination and uses feedback information from collisions to assign channels to multiple secondary users in an unslotted primary network.

III. PROBLEM DEFINITION

A. Network Model

We consider a unslotted primary network with $m = 1, 2, \dots, M$ secondary users. Each secondary user can sense $n = 1, 2, \dots, N$ primary channels, each with bandwidth B_j ($j = 1, 2, \dots, N$). The occupancy of the N channels by primary users are modeled as independent continuous-time Markov processes [7]. The availability of channel j for secondary user i is modeled as a two-state continuous-time Markov chain with a state of $S_j^i(t)$, where $S_j^i(t) = 1$ indicates that there is an opportunity for secondary user i in channel j , and $S_j^i(t) = 0$ otherwise. The idle and busy periods for channel j are exponentially distributed with parameters λ_j and μ_j , respectively. This leads to an unslotted primary network, where the primary users can access the channel at any time. Secondary users, however, adopt a slotted transmission structure with a slot length of τ . At each slot, a secondary user chooses one of the N channels to sense and decide whether to transmit over the chosen channel based on the sensing outcome. The operations in a secondary slot are shown in Fig. 1.

Consider the transmission slot of length τ starts at time t . The beginning of the slot is used for sensing one of the N channels which takes τ_s seconds. We assume that a secondary user can distinguish primary user traffic from other secondary users. Based on the current sensing decision and past sensing results, the secondary user can choose to either transmit on one of the N channels or not transmit at all. The channel access takes place during the access period $[t + \tau_s, t + \tau_s + \tau_t]$ of the slot. If the channel remains idle for the entire duration of $[t + \tau_s, t + \tau_s + \tau_t]$, the transmission is successful; otherwise, a

collision occurs. At the end of each time slot, a secondary user predicts the states of the channels in the next time slot with the latest sensing outcome. Notice that even if a secondary user has sensed the channel to be idle during the sensing period τ_s , the primary user may become busy during the time duration $[t + \tau_s, t + \tau_s + \tau_t]$.

B. Problem Formulation

We formulate the opportunistic channel access problem as a constrained POMDP represented by the tuple (S, A, R, W) given below.

- S represents the state of the underlying system at the beginning of each slot for every secondary user. For secondary user i , the state in channel j is given by $S_j^i(k) \triangleq S_j^i(t)|_{t=(k-1)\tau}$, where $k = 1, \dots, T$ is the slot index. The system state in slot k for all secondary users is thus $S(k) = [S_1^1(k), S_2^1(k), \dots, S_N^1(k), S_1^2(k), \dots, S_N^M(k)] \in \{0, 1\}^{M, N}$. Recall that $S_j^i(k)$ is a discrete-time Markov chain for secondary user i and channel j . The state transition probabilities of the Markov chain are as follows,

$$\begin{aligned} p_{00}^{ij}(k) &= 1 - \frac{\mu_j}{\lambda_j + \mu_j} (1 - \exp(-(\lambda_j + \mu_j)k)) \\ p_{01}^{ij}(k) &= \frac{\mu_j}{\lambda_j + \mu_j} (1 - \exp(-(\lambda_j + \mu_j)k)) \\ p_{10}^{ij}(k) &= \frac{\lambda_j}{\lambda_j + \mu_j} (1 - \exp(-(\lambda_j + \mu_j)k)) \\ p_{11}^{ij}(k) &= 1 - \frac{\lambda_j}{\lambda_j + \mu_j} (1 - \exp(-(\lambda_j + \mu_j)k)) \quad (1) \end{aligned}$$

- A is the combined action space for all secondary users. Specifically, after every sensing operation, every secondary user can either choose to transmit in one of the N channels or, alternatively, not transmit at all. We use $a^i(k)$ to denote the action taken by user i in time slot k . An acknowledgement (ACK) is piggybacked to indicate whether the transmission by secondary user i on channel j is successful or not; $ACK_j^i(k) \in \{0(\text{no success}), 1(\text{success})\}$.
- R represents the reward accrued by a successful transmission. The reward is defined as the number of bits delivered when a secondary user senses and transmits on the channel chosen by action $a^i(k)$ in the current time slot. We use $R_j^i(k)$ to denote the award accrued by secondary user i on channel j during time slot k .
- W represents the belief vector. Each secondary user cannot directly observe the entire system state due to limited sensing. However, a secondary user can infer the system state from its decision and observation history. The statistical information on the system state provided by the entire decision and observation history can be encapsulated in a belief vector. A belief vector for user i at time slot k , namely $W^i(k)$, is a N -dimensional vector $(w_1^i(k), w_2^i(k), \dots, w_N^i(k))$, where $w_j^i(k)$ denotes the conditional probability for secondary user i to access channel j in time slot k given $S_j^i(k) = 1$.

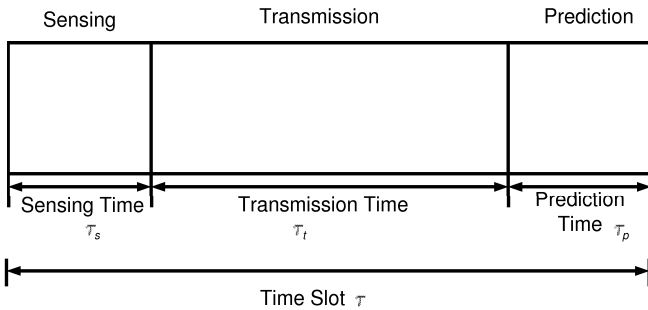


Fig. 1. Operations in a secondary slot.

We propose a learning-based OSA approach. The goal of our OSA approach is to achieve maximal throughput for secondary users. Here, we measure the throughput as the total number of bits that can be delivered by secondary users in T slots, which can be computed by summing the expected reward for all secondary users. Thus, the problem can be formulated as follows,

$$\begin{aligned} & \max \left(\sum_{t=1}^T \sum_{i=1}^M R_j^i(k) \right) \\ & \text{s.t. } P_j^{pu}(k) = Pr\{\Theta_j(k) = 1 | \Psi_j(k) = 0\} = 0, \forall j, k \end{aligned} \quad (2)$$

where $\Theta_j(k) \in 0(\text{noaccess}), 1(\text{access})$ denotes the decision taken by the secondary users to access channel j during the access time period. The collision constraint defined in the above equation indicates that the probability of collision $P_j^{pu}(k)$ perceived by the primary users in any channel j and slot k is equal to zero. $\Psi_j(k)$ is defined as the availability of an idle channel j :

$$\Psi_j(k) = \begin{cases} 1, & S_j(t) = 1 \forall t \in [(k-1)\tau + \tau_s, k\tau] \\ 0, & \text{otherwise} \end{cases}$$

Our learning-based OSA approach involves different degrees of learning among the secondary users. In each time slot, the learning-based approach allows each secondary user to implement the three main operations from Fig. 1 in following four phases: i) channel sensing phase, ii) channel access phase, iii) reward phase, and iv) prediction phase. Next, we will present the details of the approach.

IV. LEARNING-BASED OSA APPROACH

The basic idea of the learning-based approach is that secondary users should learn from collision events in order to maximize the network throughput. The strategy for each secondary user is to decide with what probability to sense a channel so that the collisions with other secondary users are minimized. Though collision reduces the network throughput, it provides information to estimate the belief vector of secondary users. In our approach, each secondary user not only maintains an estimated belief vector of itself, but also maintains an estimated belief vector for other secondary users by learning from collision events. We assume that a secondary user can correctly identify the signal of primary users and the other secondary users. This can be implemented by adding an identification preamble in front of every frame transmitted by secondary users. Next, we explain the implementation of the four phases and conduct the performance analysis.

A. Channel Sensing Phase

Let $p_{i,j}^i(k)$ be channel j 's idle probability in time slot k from the perspective of user i . At the beginning of slot k , user i first retrieves its reward and the channel idle probabilities information which are calculated in phases of reward and prediction (explained later). It then decides with what probability to sense each channel so that the collisions with other users are minimized.

Define $p_{S,j}^i(k)$ as the probability of user i to sense channel j in time slot k . User i then computes the channel sensing probability vector $P_S^i(k) = (p_{S,1}^i(k), p_{S,2}^i(k), \dots, p_{S,n}^i(k))$. In order to minimize the collisions with other secondary users, an optimization process needs to be carried out. The optimization equation is given by:

$$\begin{aligned} P_S^i(k) = \arg \max_{P_S^i(k)} & \left(\sum_{i=1}^M \sum_{j=1}^N (p_{S,j}^i(k) p_{I,j}^i(k)) \right. \\ & \left. R_j^i(k) \prod_{l \in M, l \neq i} (1 - p_{S,j}^l(k)) \right) \\ & \text{s.t. } \sum_{j=1}^N p_{S,j}^i(k) \leq 1 \end{aligned} \quad (3)$$

In Eq. (3), $p_{S,j}^i(k) \prod_{l \in M, l \neq i} (1 - p_{S,j}^l(k))$ denotes the probability that at least one user is about to sense channel j . $(p_{S,j}^i(k) p_{I,j}^i(k) R_j^i(k) \prod_{l \in M, l \neq i} (1 - p_{S,j}^l(k)))$ is the expected throughput obtained from channel j . After summing over all secondary users and channels, the throughput in the whole spectrum is obtained. The above optimization process is implemented on each secondary user independently. Therefore, each secondary user has a different view of the channel probabilities. At the end of this phase, each secondary user has computed the channel sensing probabilities. In the next phase, one of the channels will be chosen to access.

B. Channel Access Phase

In this phase, the channel sensing probabilities will be mapped to a concrete decision in determining which channel to access. Specifically, user i needs to perform the action $a^i(k) \in \{0, 1, \dots, N\}$ by choosing one specific channel for transmission. Notice that it is possible that $a^i(k) = 0$; this means that user i will have no transmission at all. The value for $a^i(k)$ is chosen such that the expected immediate reward is maximized. Mathematically,

$$a^i(k) = \begin{cases} 0, & P_S^i(k) = 0 \\ \arg \max_{j=1, \dots, N} (P_S^i(k)), & \text{otherwise} \end{cases} \quad (4)$$

As can be seen the first two phases map secondary users to various channels across the whole spectrum in a distributive fashion and thus is able to achieve a maximal throughput for secondary users. The next two phases collect information needed by the next time slot.

C. Reward Phase

Recall that in an unslotted primary network, the occupancy of each channel by primary users follows an unslotted transmission scheme. Collisions may occur in spite of the fact that the channel is idle during the sensing period. In this phase, each secondary user computes its reward based on the actual transmission quality. A successful transmission will accrue a reward, and cost is incurred if a collision occurs.

The reward $R_j^i(k)$ is defined as the total number of bits successfully transmitted when secondary user i accesses channel j for transmission in slot k . Given the acknowledgement $ACK_j^i(k)$, transmission rate r_j , and transmission time τ_t , the reward $R_j^i(k)$ can be calculated as follows,

$$R_j^i(k) = ACK_j^i(k)r_j\tau_t \quad (5)$$

From Eq. (5), we get the reward accumulated by a secondary user in the current time slot. By substituting Eq. (5) in Eq. (2), the reward accumulated by all secondary users in the current time slot is computed. Thus, we have the maximal throughput for all secondary users. As shown in Eq. (2), $R_j^i(k)$ influences the computation of $P_S^i(k)$. A zero value for $R_j^i(k)$ indicates a collision has taken place. In every subsequent time slot, the maximal throughput is attained by reduced collisions which leads to a higher reward.

D. Prediction Phase

In this phase, the belief vectors are first computed and then are used for the calculation of channel idle probabilities. Based on two network scenarios, we present two distinct computations of belief vectors. In the first scenario, the number of channels exceeds the number of secondary users. In this case, each secondary user estimates the belief vector of other secondary users. In the second scenario, the number of channels is less than the number of secondary users. In this case, some secondary users will not be able to transmit for long periods of time as they are not able to sense any channel. Thus, the secondary user independently pick the channel for which it has a higher belief of seeing an opportunity.

In the first scenario, the secondary user maintains a belief vector of itself and estimates the belief vectors of other secondary users by learning from the collision events. The estimation of belief vectors ensures that there is no message exchange between secondary users. Also the collision information provides accurate belief vector estimation for other secondary users. Should a collision occur, every secondary user identifies the colliding users based on the identification provided in the preamble. The initial belief vector is set to the stationary distribution of the underlying Markov process if no information on the initial system state is available. The update of belief vector in every time slot is shown in Eq. (6).

$$w_j^l(k+1) = \begin{cases} p_{11}^{lj}(k)w_j^l(k) + (1 - (w_j^l(k)))p_{01}^{lj}(k), & \text{if } a^i(k) \neq j \\ p_{11}^{lj}(k)w_j^l(k) + (1 - (w_j^l(k)))p_{01}^{lj}(k), & \text{if } a^i(k) = j \text{ and } a^l(k) \neq j \text{ and } S_j^i(k) = 1 \\ & \text{and } collision(i, l, k) = FALSE \\ p_{11}^{lj}(k), & \text{if } a^i(k) = j \text{ and } collision(i, l, k) = TRUE \end{cases} \quad (6)$$

where $collision(i, l, k)$ represents whether a collision occurs between secondary users l and i in time slot k , $a^l(k)$ is the

predicted action of secondary user l by secondary user i using the estimated belief vector $w^l(k)$.

In the second scenario, secondary users do not estimate the belief vectors of other secondary users. The channel access decision is made based on the channel estimation of individual user. Each secondary user then updates its belief vector independently as follows:

$$w_j^i(k+1) = \begin{cases} p_{11}^{ij}, & \text{if } a^i(k) = j \text{ and } S_j^i(k) = 0 \\ p_{01}^{ij}, & \text{if } a^i(k) = j \\ p_{01}^{ij}, & \text{if } a^i(k) \neq j \end{cases} \quad (7)$$

Once the belief vectors are computed, the next step is to calculate the channel idle probabilities which will be used in the next time slot. Given the belief vector $w_j^i(k)$ and state transition probabilities from Eq. (1), the probability that channel j is idle in the next time slot for secondary user i is given by

$$p_{1,j}^i = w_j^i(k)p_{11}^{ij}(k) + (1 - w_j^i(k))p_{01}^{ij}(k) \quad (8)$$

The prediction phase is the final task to be completed in the current time slot. In the subsequent time slot, each secondary user retrieves its reward and the channel idle probabilities from the previous time slot, and then repeats the execution for all four phases starting from the channel sensing phase.

E. Computation Complexity

The computation overhead comes mainly from the computation of P_S^i . From Eq. (3), the optimization equation can be reduced to a maximum weighted matching problem in a bipartite graph $G(V_1, V_2, E)$, where $V_1 = N$ and $V_2 = M$. Edge $e \in E = (i, j)$ means that channel j is available to secondary user i . The maximum weighted matching problem for a bipartite graph can be solved in polynomial time [11]. If $M > N$, P_S^i can be computed in $O(N^2M)$ time slots; otherwise $O(M^2N)$ time slots.

V. NUMERICAL RESULTS

In this section we present comprehensive simulations in MATLAB to evaluate the performance of our learning-based approach. We will compare the performance of the learning-based approach with equal probability strategy and cooperative multi-user approach. In the equal probability strategy, secondary users sense each channel with the same probability. In the cooperative multi-user strategy, each secondary user has a complete view for channel availability for other secondary users by exchanging belief vectors. In our simulations, the values for primary traffic parameters λ and μ are motivated by practical experiments conducted in [9]. The number of primary users in the network is equal to the total number of channels N . The idle-times show heavy-tailed behavior and are approximated by an exponential distribution with parameter $1/\lambda = 4.2ms$. The channel busy period is assumed to be $1/\mu = 1ms$. We assume that the bandwidth $B = 1$ and the length of the time slot $\tau = 1ms$.

TABLE I

IMPACT OF THE VARIATION IN IDLE PROBABILITY DISTRIBUTION ON THE NETWORK THROUGHPUT IN A 6-CHANNEL NETWORK WITH 4 SECONDARY USERS.

Idle Probability	Optimal	Coop	Learn	Equal
[0.1, 0.1, 0.1, 0.2, 0.2, 0.7]	1.15	1.15	1	0.88
[0.8, 0.1, 0.03, 0.03, 0.03, 0.03]	1.24	1.24	1.08	0.96
[0.1, 0.1, 0.1, 0.4, 0.4, 0.5]	0.76	0.75	0.62	0.5
[0.4, 0.15, 0.15, 0.15, 0.15]	1.25	1.25	1.13	1.02
[0.15, 0.15, 0.15, 0.15, 0.15, 0.15]	1.35	1.35	1.22	1.11

A. Performance of learning-based approach in presence of multiple secondary users

In this simulation, we evaluate the network performance under different message arrival rates. In Fig. 2, we present the total throughput experienced by four secondary users in a six channel network. Message arrivals at the secondary users form a Poisson process. The message length is geometrically distributed with an average length of 50 packets. In each slot, secondary users do not participate in the channel sensing and access activities if they do not have packets to transmit. The total number of time slots used in simulations are 1000. Unsurprisingly, the performance of the cooperative approach is closer to the optimal approach with a price of high communication overhead. The throughput obtained in our learning-based approach is lower than the cooperative approach by 10-15%.

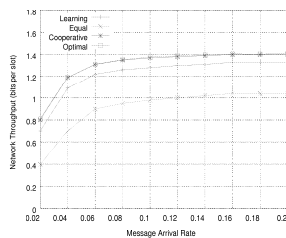
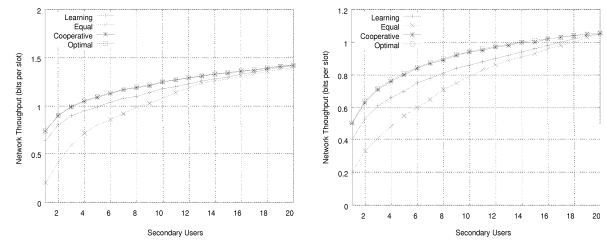


Fig. 2. Throughput for 4 secondary users in a 6-channel network

B. The impact of idle probability distribution on the performance of learning-based approach

In this simulation, we examine the impact of varying idle probabilities from channel to channel on the throughput of secondary users. Fig. 3a shows a scenario where the idle probabilities vary sharply from channel to channel [0.7, 0.2, 0.2, 0.1, 0.1, 0.1]. Fig. 3b shows a scenario where the idle probabilities vary smoothly from channel to channel [0.5, 0.4, 0.4, 0.1, 0.1, 0.1]. In Fig. 3a and Fig. 3b, when the number of secondary users are lesser than the number of channels, the learning-based approach performs better than the equal-probability approach. When the number of secondary users exceeds the number of channels, the performance of both the approaches are closer to each other.

Table II depicts the throughput obtained by four secondary users with different channel idle probability distributions. The performance of optimal approach and cooperative approach



(a) Sharply varying idle probabilities. (b) Smoothly varying idle probabilities.

Fig. 3. Impact of varying idle probabilities.

are indistinguishable. The learning based approach does not perform as well as the cooperative approach, but it performs better than the equal-probability approach.

VI. CONCLUSIONS

In this paper, we proposed a learning-based approach to allow multiple secondary users to achieve maximal throughput by exploiting idle periods in an unslotted primary network. Our approach is computationally less expensive and has no communication overhead as it does not involve exchange of spectrum maps among other secondary users. The performance of the learning-based approach is closer to the cooperative and optimal approaches with no communication overhead.

ACKNOWLEDGMENTS

The research of Min Song is supported by NSF CAREER Award CNS-0644247.

REFERENCES

- [1] M. McHenry, "Spectrum white space measurements," New America Foundation Broadband Forum, 2003.
- [2] Q. Zhao and B.M. Sadler, "A survey of dynamic spectrum access," IEEE Signal. Process. Mag, May 2007.
- [3] Q. Zhao and A. Swami, "Structure and Optimality of Myopic Sensing for Opportunistic Spectrum Access," Proc. of the IEEE International Conference on Communications (ICC), 2007.
- [4] Y. Yuan, P. Bahl, R. Chandra, P. A. Chou, J. I. Ferrell, T. Moscibroda, S. Narlanka, and Y. Wu, "Knows: Cognitive radio networks over white spaces," in Second IEEE International Symposium on Dynamic Spectrum Access Networks, 2007.
- [5] 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, May 2008.
- [6] D. Raychaudhuri and X. Jing, "A spectrum etiquette protocol for efficient coordination of radio devices in unlicensed bands," in Proc. on Personal, Indoor and Mobile Radio Communications, 2003.
- [7] Q. Zhao, S. Geirhofer, L. Tong, and B. M. Sadler, "Opportunistic spectrum access via periodic channel sensing," IEEE Transactions on Signal Processing, Vol. 56, No. 2, Feb. 2008.
- [8] 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.
- [9] S. Geirhofer, L. Tong, B.M. Sadler, "Dynamic spectrum access in WLAN channels: Empirical model and its stochastic analysis," Workshop Technol. Policy Accessing Spectrum (TAPAS), 2006.
- [10] M. Marcus, "Real time spectrum markets and interruptible spectrum: new concepts of spectrum use enabled by cognitive radio," in IEEE International Symposium on Dynamic Spectrum Access Networks, 2005.
- [11] F. Bourgeois and J. Lassalle, "An extension to the munkres algorithm for the assignment problem to rectangular matrices," Communications of the ACM, December 1971.