Distributed Sensing and Access in Cognitive Radio Networks

Keqin Liu, Qing Zhao, Yunxia Chen
Department of Electrical and Computer Engineering
University of California, Davis, CA 94536
{kqliu,qzhao,yxchen}@ucdavis.edu

***Abstract*—We consider an ad hoc network of secondary users searching for idle frequency bands in a spectrum consisting of multiple channels. In each slot, a secondary user chooses one channel to sense and decide whether to access based on the sensing outcome. A sensing strategy for intelligent channel selection is crucial to track the rapidly varying spectrum opportunities. In an ad hoc network without a central controller or common control channels, a secondary user can only resort to its local observations in the decision making. The tradeoff is between choosing the channel most likely to be idle and avoiding other competing secondary users. We show that the problem can be formulated as a decentralized Partially Observable Markov Decision Process (POMDP). A suboptimal randomized sensing policy is then proposed. This policy effectively addresses this design tradeoff and offers significant improvement in network throughput over the optimal single-user design.**

***Index Terms*—Cognitive radio, spectrum opportunity tracking, multi-user diversity, decentralized POMDP, randomized policy.**

I. INTRODUCTION

Opportunistic spectrum access (OSA), also referred to as spectrum overlay, is one of several approaches proposed to resolve the paradox between the overly crowded radio spectrum and the pervasiveness of spatial and temporal idle frequencies revealed by the actual spectrum usage measurements [1]. The basic idea is to open the spectrum to secondary users. Equipped with cognitive radios that are capable of autonomous reconfiguration by learning from and adapting to the communication environment, secondary users are able to identify and exploit local and instantaneous spectrum opportunities without causing unacceptable interference to primary users.

A basic component of OSA is a sensing strategy at the MAC layer for spectrum opportunity tracking. Due to hardware limitations and energy constraints, a secondary user may not be able to sense all channels in the spectrum simultaneously. In this case a sensing strategy for intelligent channel selection is necessary to track the rapidly varying spectrum opportunities.

By modeling primary users' channel occupancy as a Markov process, the design of sensing strategies is formulated as a partially observable Markov decision process (POMDP) in [2], [3], where the objective is to maximize the throughput of an individual selfish secondary user. The interaction among secondary users is not taken into account in the design of the sensing strategy.

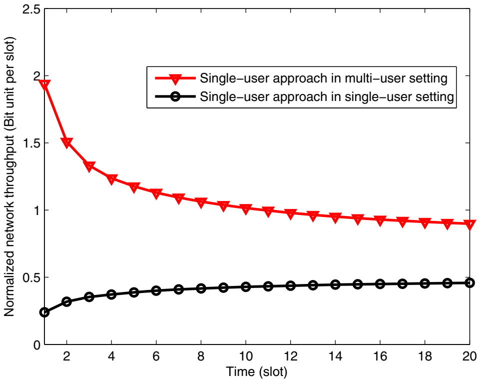

Fig. 1. Network throughput of the optimal single-user sensing strategy.

The optimal sensing strategy designed for individual users is, however, suboptimal in terms of network throughput. Intuitively, to maximize the network throughput, secondary users with data to transmit should seek spectrum opportunities in different channels. Synchronized channel selections where multiple secondary users choose the same channel, leaving opportunities in other channels unexploited, are undesirable. Unfortunately, this is likely to be the case when every secondary user employs the same sensing strategy designed to maximize its own throughput. Shown in Figure 1 is the network throughput performance of the optimal single-user sensing policy applied to a multiuser setting, where we consider 16 secondary users with Poisson packet arrivals competing for opportunities in 4 channels. In the first slot, each secondary user randomly selects a channel to sense, and there is a multiuser diversity gain when secondary users exploit spectrum opportunities asynchronously, resulting in realizing opportunities in different channels simultaneously. As secondary users learn the environment through opportunity sensing and tracking, they obtain similar observations and converge to the same opportunity assessment, thus losing the multiuser diversity gain. In a sharp contrast to the increasing throughput through cognition as we observe in a single-user

---

[0] This work was supported by the Army Research Laboratory CTA on Communication and Networks under Grant DAAD19-01-2-0011 and by the National Science Foundation under Grants CNS-0627090 and ECS-0622200.

setting, the network throughput decreases precipitately with time.

In this paper, we aim to address the deficiency of a direct application of the single-user sensing strategy to a network setting. We consider an ad hoc secondary network with distributed users who join/exit the network and sense/access the spectrum independently without exchanging information on their observations and actions. There is no central controller or common control channel to coordinate the actions of users. We also consider random traffic arrivals at each secondary users, which renders a fixed channel partition among users inefficient. We show that the problem of distributed spectrum sharing among secondary users can be formulated as a decentralized POMDP [4]. Built upon the insight obtained in designing the optimal single-user sensing strategy, we propose a randomized policy to avoid user synchronization. This randomized sensing policy effectively addresses the tradeoff between choosing the channel that is most likely to offer a spectrum opportunity and avoiding other competing secondary users.

Related Work An overview of challenges and recent developments in OSA can be found in [1]. The majority of existing work on OSA focuses on the spatial domain where spectrum opportunities are considered static or slowly varying in time [5]–[7]. As a consequence, real-time opportunity identification and tracking is not as critical a component in this class of applications, and the prevailing approach often assumes perfect knowledge of spectrum opportunities at any time and location over the entire spectrum of interest.

Exploiting temporal spectrum opportunities requires a joint design of spectrum detectors at the physical layer and spectrum sensing and access strategies at the MAC layer [3]. A separation principle for the optimal joint design has been established in [3]; it shows that the design of the sensing policy for opportunity tracking can be decoupled from that of the spectrum detector and the access strategy, leading to simple structural solutions. The structure and the optimality of the myopic approach to opportunity tracking has been addressed in [8], [9] in a single-using setting.

A simple yet sufficiently accurate statistical model of spectrum occupancy is crucial to the efficiency of spectrum opportunity tracking. Measurements obtained from spectrum monitoring test-beds [10] demonstrate the Markovian transition between busy and idle channel states in 802.11b, which is the model used in this paper.

II. NETWORK MODEL

Consider a spectrum consisting of N channels, each with bandwidth B_i $(i = 1, \cdots, N)$. These N channels are licensed to a primary network whose users communicate according to a synchronous slot structure. We model the spectrum occupancy as a discrete-time homogenous Markov process with 2^N states. A state $\mathbf{S}(t) \overset{\Delta}{=} [S_1(t), \ldots, S_N(t)]$ denotes the channel occupancy, where $S_n(t) \in \{0 \text{ (busy)}, 1 \text{ (idle)}\}$ is the state of channel n in slot t.

We consider a secondary ad hoc network with M users independently searching for and accessing spectrum opportunities in these N channels. In each slot, a secondary user with packets to transmit chooses one of the N channels to sense. If the channel is sensed to be busy, the user refrains from transmission and chooses a potentially different channel in the next slot. If the channel is sensed to be idle, the user transmits using carrier sensing. Specifically, it generates a random backoff time and transmits when the timer expires and no other secondary users have claimed the channel. The framework presented in this paper applies when secondary users transmit directly over an idle channel without carrier sensing. Note that collisions among secondary users may occur. The receivers will acknowledge successful transmissions at the end of a slot.

We model the random packet arrivals at each secondary user as a Poisson process with rate λ. Packets are buffered until they are successfully transmitted.

The objective is to design the sensing strategy that governs the channel selection for each user in each slot to maximize the network throughput, *i.e.,* the expected number of bits delivered across the network over a finite horizon of T slots.

III. A POMDP FRAMEWORK

In this section, we present the POMDP formulation of designing the optimal spectrum sensing strategy for a single secondary user. For simplicity, we assume that this user is fully backlogged. Insights obtained from the single-user formulation will help us tackle the multi-user problem.

A. Reward and Design Objective

Let π_s denote a sensing policy that decides, sequentially, which channel to sense in each slot. The design of π_s can be formulated as a POMDP [2]. Specifically, the underlying state space is $\mathbf{S} = [S_1, \cdots, S_N] \in \{0, 1\}^N$, and the action space is $a \in \{1, \cdots, N\}$. A natural choice of reward $R_a(t)$ for choosing channel a in slot t is

$$R_a(t) = S_a(t)B_a,$$

which represents the number of bits delivered. We can then define the objective function as the total number of bits transmitted in T slots:

$$J \overset{\Delta}{=} \mathbb{E}_{\pi_s}[\sum_{t=1}^{T} R_a(t)], \tag{1}$$

where $\mathbb{E}_{\pi_s}$ represents the conditional expectation given that a sensing policy π_s is employed. Note that the reward $R_a(t)$ obtained in slot t depends on the sensing action (which channel a to sense) and the state of the underlying Markov process (channel availability) in slot t.

B. Sufficient Statistics

Due to partial spectrum sensing, the system state $\mathbf{S}$ cannot be directly observed. We can, however, infer it from the sensing outcomes. It has been show in [11] that the *a posterior* distribution of the system state that exploits the entire sensing and decision history characterizes our knowledge about the system state and is a sufficient statistic for optimal decision making. Specifically, at the beginning of slot t, our knowledge of the system state based on all past decisions and observations can be summarized by a belief vector

$$\Lambda(t) = [\lambda_1(t), \cdots, \lambda_{2^N}(t)]$$

where $\lambda_j(t)$ is the conditional probability (given the decision and observation history) that the system state is j ($j = 1, \cdots, 2^N$) in slot t.

For independently evolving channels, it has been shown in [2] that the marginal conditional distribution is a sufficient statistic, *i.e.,* we can consider the following belief vector

$$\Omega(t) = [\omega_1(t), \cdots, \omega_N(t)]$$

where $\omega_i(t)$ denotes the conditional probability that channel i is available in slot t. Note that the dimension of the belief vector is reduced from 2^N to N when channels are independent. For simplicity, we focus on independent channels. The extension to the general case with correlated channels is straightforward.

C. Optimal Single-User Sensing Policy

With the concept of belief vector, a deterministic sensing policy π_s essentially defines the mapping from $\Omega(t)$ to the index a of the channel to be sensed for each slot t:

$$\begin{aligned} \pi_s &= [\mu_1, \cdots, \mu_T], \\ \text{where } \mu_t &: \Omega(t) \in [0,1]^N \to a \in \{1, \cdots, N\}. \end{aligned}$$

For a randomized sensing policy π_s, it defines the mapping from $\Omega(t)$ to the probability mass function of the channel index $a(t)$, which now is a random variable taking values in $\{1, 2, \cdots, N\}$. In other words, the action to be chosen in each slot is a probability distribution. For an unconstrained POMDP as we have here, optimality can always be achieved by a deterministic policy. Nevertheless, for a POMDP over a finite horizon, the optimal policy is generally non-stationary; the mapping from the belief vector to the optimal action varies over time.

IV. ACHIEVING MULTI-USER DIVERSITY

A direct application of the optimal single-user sensing strategy to a multiuser setting results in the undesirable user synchronization and decreasing network-level throughput as demonstrated in Figure 1. Specifically, when multiple secondary users happen to select the same channel in a particular slot, they obtain the same observation, which leads to similar updated belief vectors. In turn, similar belief vectors are more likely to result in the same channel selection in the future slots. This positive feedback loop eventually leads to user synchronization.

In this section, we consider multi-user sensing policies by explicitly taking into account the competition among secondary users. We first formulate the problem as a decentralized POMDP. A suboptimal randomized policy is then proposed and its performance studied.

A. A Decentralized POMDP Formulation

Spectrum sharing among M distributed secondary users can be formulated as a decentralized POMDP. The underlying state space is again $\mathbf{S} = [S_1, \cdots, S_N] \in \{0,1\}^N$, indicating the availability of each channel[1]. The joint action $A(t) = \{a_1(t), \cdots, a_M(t)\}$ indicates the channel selection of each user in slot t.

The reward $R(t)$ in slot t for a given joint action $A(t)$ and system state $\mathbf{S}(t)$ is the total number of bits successfully delivered by all M users. Let

$$m_n(t) = |\{a_i(t) : a_i(t) = n\}|$$

denote the number of users choosing channel n, which is determined by the joint action $A(t)$. We have

$$R(t) = \sum_{n=1}^{N} S_n(t) P_{cs}(m_n(t)) B_n, \qquad (2)$$

where $P_{cs}(m_n)$ denotes the probability that one secondary user successfully claims the channel through carrier sensing when m_n users are contending for the same channel. Assume that a user will succeed in claiming the channel if its backoff time is smaller than other users' by at least δ (account for the propagation and processing delay). The probability of success for $m > 1$ is then given by

$$P_{cs}(m) = \int_{\delta}^{\bar{\tau}} F_1(x - \delta) f_2(x) \mathrm{d}x,$$

where $\bar{\tau}$ is the maximum backoff time, $F_1(x)$ is the CDF of the first order statistic, and $f_2(x)$ the PDF of the second order statistic of m backoff times. For uniformly distributed backoff times, we have, for $m > 1$,

$$\begin{aligned} P_{cs}(m) &= n(n-1)\left[\frac{1}{m-1}(1-\frac{\delta}{\bar{\tau}})^{m-1} - \frac{1}{2m-1}(1-\frac{\delta}{\bar{\tau}})^{2m-1} \right. \\ &\quad \left. -\frac{1}{m}(1-\frac{\delta}{\bar{\tau}})^{m} + \frac{1}{2m}(1-\frac{\delta}{\bar{\tau}})^{2m}\right]. \end{aligned}$$

Note that we can easily update the reward function (2) to incorporate multipacket reception at receivers or direct transmission without carrier sensing among secondary users.

Local observation $o_i(t)$ at user i who chooses channel a_i in slot t consists of two parts: the state $S_{a_i}(t)$ of the channel a_i it chooses and whether it succeeds in carrier sensing. Note that the second observation along with user i's backoff time can produce an estimate of the number m_{a_i} of users who choose the same channel a_i [12]. This estimate may be used in choosing future actions to avoid crowded channels.

For each secondary user i, a local policy π_i maps its local observation history $\{o_i(1), \cdots, o_i(t-1)\}$ to local actions $a_i(t)$ to slot t. The objective is to design the local policy for each user to maximize the expected total reward obtained in T slots.

---

[1] We assume here that all secondary users are fully backlogged. For random traffic arrivals, the queue length of each user may be incorporated into the system state.

B. A Randomized Policy for Multiple Competing Users

Optimal solutions to decentralized POMDPs are often intractable. Leveraging insight obtained from the optimal single-user sensing strategies, we aim to obtain simple suboptimal multi-user policies that are robust to varying user populations and random packet arrivals.

One approach to avoiding synchronous channel selection among users is to consider randomized tracking policies. In this case, based on its belief vector, the user decides the *probability* of selecting each channel, *i.e.,* the probability mass function (PMF) of the channel to be selected. With a randomized sensing policy, even when two users have the same belief, their channel selections can be different, leading to different updated beliefs and future actions. The design of randomized policies are, unfortunately, significantly more complex due to the uncountable action space of probability distributions.

We propose a heuristic multiuser policy where the action (the PMF of channel selection) is simply the normalized belief vector. Specifically, for user i with belief vector $\Omega^{(i)}(t) = [\omega_1^{(i)}(t), \cdots, \omega_N^{(i)}(t)]$, the probability $p_n^{(i)}(t)$ of choosing channel n in slot t is given by

$$p_n^{(i)}(t) = \frac{\omega_n^{(i)}(t)}{\sum_{k=1}^{N} \omega_k^{(i)}(t)}.$$

At the end of slot t, user i updates its belief vector via Bayes' rule by incorporating its local action $a_i(t)$ and local observation $S_{a_i}(t)$.

This heuristic design is based on the intuition that higher selecting probabilities should be given to channels more likely to be opportunities, and the belief provides precisely each channel's likelihood of being an opportunity based on past observations. This simple randomized policy performs surprisingly well as demonstrated in Figure 2. Again, we observe the improved performance over time; user synchrony is avoided, and the benefit of cognition through belief updates restored.

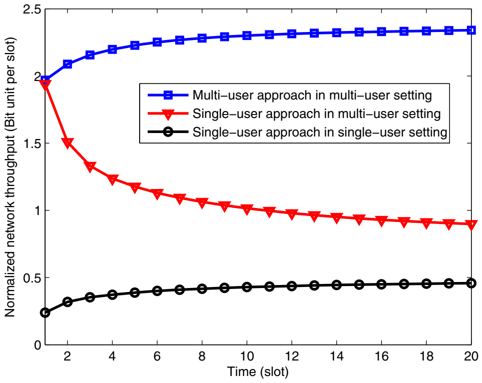

Fig. 2. Network throughput of the randomized sensing policy.

V. CONCLUSION

In this paper, we consider distributed spectrum sharing among competing secondary users. We demonstrate that the optimal sensing strategy designed for individual secondary users leads to suboptimal performance in terms of the network throughput due to synchronized channel selection among secondary users. To effectively address the tradeoff between choosing the channel most likely to offer a transmission opportunity and avoiding other secondary users, we propose a randomized sensing policy. Significant improvement over the optimal single-user design has been observed.

The proposed approach to realizing multiuser diversity has used only the observations of channel occupancy encapsulated in the belief vector. Additional gain may be realized by learning from the collision history how many secondary users are competing for the same opportunity. The key question is how to combine these two pieces of information—the belief indicating channel occupancy states and the estimated number of competing users in the same channel—in distributed decision making. This question will be addressed in the future work.

REFERENCES

[1] Q. Zhao and B. Sadler, "A Survey of Dynamic Spectrum Access," *IEEE Signal Processing magazine: Special Issue on Resource-Constrained Signal Processing, Communications, and Networking*, vol. 24, pp. 79–89, May 2007.

[2] 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: Special Issue on Adaptive, Spectrum Agile and Cognitive Wireles Networks*, vol. 25, pp. 589–600, April 2007. (see also *Proc. of DySPAN*, pp. 224 - 232, Nov. 2005).

[3] Y. Chen, Q. Zhao, and A. Swami, "Joint Design and Separation Principle for Opportunistic Spectrum Access in the Presence of Sensing Errors," *submitted to IEEE Transactions on Information Theory*. available at http://www.ece.ucdavis.edu/∼qzhao/Journal.html (see also *Proc. of IEEE Asilomar Conference on Signals, Systems, and Computers*, Oct. 2006).

[4] D. Bernstein, *Complexity Analysis and Optimal Algorithms for Decentralized Decision Making*. PhD thesis, University of Massachusetts Amherst, September 2005.

[5] H. Zheng and C. Peng, "Collaboration and Fairness in Opportunistic Spectrum Access," in *Proceedings of IEEE International Conference on Communications (ICC)*, 2005.

[6] W. Wang and X. Liu, "List-coloring based channel allocation for open-spectrum wireless networks," in *Proc. of IEEE VTC*, 2005.

[7] S. Sankaranarayanan, P. Papadimitratos, A. Mishra, and S. Hershey, "A Bandwidth Sharing Approach to Improve Licensed Spectrum Utilization," in *Proceedings of the first IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks*, 2005.

[8] Q. Zhao and B. Krishnamachari, "Structure and optimality of myopic sensing for opportunistic spectrum access," in *Proc. of IEEE Workshop on Toward Cognition in Wireless Networks (CogNet)*, June 2007.

[9] Q. Zhao, B. Krishnamachari, and K. Liu, "Low-Complexity Approaches to Spectrum Opportunity Tracking," in *Proc. of the 2nd International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CrownCom)*, August 2007.

[10] S. Geirhofer, L. Tong, and B. Sadler, "Dynamic spectrum access in WLAN channels: empirical model and its stochastic analysis," in *Proc. of the First International Workshop on Technology and Policy for Accessing Spectrum (TAPAS)*, August 2006.

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

[12] M. Maskery, V. Krishnamurthy, and Q. Zhao, "Decentralized Dynamic Spectrum Access for Cognitive Radios: Cooperative Design of a Non-cooperative Game." submitted to *IEEE Transactions on Mobile Comupting.*, March 2007.