ENERGY-AWARE SPECTRUM-AGILE MEDIUM ACCESS IN FADING

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

Ananthram Swami
Army Research Laboratory
Adelphi, MD 20783
aswami@arl.army.mil

ABSTRACT

This paper addresses the design of distributed cognitive medium access control (MAC) protocols for opportunistic spectrum access (OSA) under an energy constraint on the secondary users. The objective is to maximize the expected number of information bits that can be delivered by a secondary user during its battery lifetime without causing interference to primary users. By absorbing the residual energy level of the secondary user into the state space, we formulate the energy-constrained OSA problem as an unconstrained partially observable Markov decision process (POMDP) and obtain the optimal spectrum sensing and access policy. We analyze and reduce the computational complexity of the optimal policy. We also propose a suboptimal solution to energy-constrained OSA, whose computational complexity can be systematically traded off with its performance. Numerical examples are provided to study the impact of spectrum occupancy dynamics, channel fading statistics, and energy consumption characteristics of the secondary user on the optimal sensing and access decisions.

1. INTRODUCTION

Network centric warfare depends on assured global wireless networks that can be established rapidly, anywhere and anytime. One of the most difficult challenges is the increasingly crowded spectrum under the existing static allotment policy on the one hand, and the increasing demand for bandwidth to support more services. The number of potentially netted devices is large (and increasing) and their needs are so dynamic that the overhead in distributing and updating topology information can be prohibitive. Indeed, the spectrum allocation problem is extremely difficult for multi-nation coalition operations, as the composition of the operating units can change rapidly. A further impediment to rapid deployment is that spectrum allotment policy varies from country to country (sometimes from service to service), often with mandated a priori assignment of spectrum to services before deployment. This should be contrasted with the commercial wired world, where infrastructure provides fiber with virtually unlimited bandwidth, and efficient utilization is largely dictated by technology and cost (routers, transmitters, switches).

[0]This work was supported by Army Research Laboratory CTA on Communication and Networks under Grant DAAD19-01-2-0011.

Opportunistic spectrum access (OSA), envisioned by the DARPA XG program (DARPA, 2005), aims to exploit the instantaneous spectrum availability using sophisticated signal processing and networking techniques. The idea is to identify available channel resources (frequency, space-time, and codes) and communicate opportunistically in a manner that limits the level of interference perceived by primary users. Such strategies are particularly relevant to small units penetrating deep in an unknown territory.

Related Work The design of medium access control (MAC) for OSA has been received great attention recently (DySPAN, 2005; CrownCom, 2006). The majority of the literature (Mangold et al., 2004; Larcher et al., 2004, Papadimitratos et al., 2005) focus on a network of geographically distributed secondary users, each affected by a different set of primary users whose spectrum access activities are static or slowly varying in time. The design objective is to allocate these spatially varying spectrum opportunities among secondary users so that the network-level spectrum efficiency is maximized subject to some regulatory constraint on interference to primary users.

In (Zhao et al., 2005a; Zhao et al. 2006), the exploitation of temporal spectrum opportunities resulting from the bursty traffic of primary users has been studied. Within the framework of partially observable Markov decision process (POMDP), the optimal distributed cognitive MAC protocol that allows secondary users to independently search for and exploit instantaneous spectrum opportunities has been developed in (Zhao et al., 2005a). This protocol consists of a sensing strategy that determines which channels in the spectrum to sense based on spectrum occupancy dynamics and an access strategy that determines whether to transmit over the sensed channels based on sensing outcomes. The energy constraint of secondary users is, however, not taken into account in (Zhao et al., 2005a; Zhao et al. 2006).

Contribution The ability to dynamically and efficiently share the spectrum depends upon how much information

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **01 NOV 2006**	2. REPORT TYPE **N/A**	3. DATES COVERED **-**
4. TITLE AND SUBTITLE **Energy-Aware Spectrum-Agilemedium Access In Fading**		5a. CONTRACT NUMBER
		5b. GRANT NUMBER
		5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)		5d. PROJECT NUMBER
		5e. TASK NUMBER
		5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Department of Electrical and Computer Engineering University of California, Davis, CA 94536**		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSOR/MONITOR'S ACRONYM(S)
		11. SPONSOR/MONITOR'S REPORT NUMBER(S)
12. DISTRIBUTION/AVAILABILITY STATEMENT **Approved for public release, distribution unlimited**		
13. SUPPLEMENTARY NOTES **See also ADM002075.**		
14. ABSTRACT		
15. SUBJECT TERMS		

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**UU**	**8**	

is available about spectrum occupancy, how accurate the information is, and how quickly and fully this information can be exchanged among secondary users. Given that sensing and estimation outcomes cannot be error free, we do not insist on binary (black and white) estimates of spectrum occupancy, rather we estimate occupancy likelihoods. Given that spectrum usage will be dynamic in the scenarios of interest, we do not insist on complete channel occupancy information. But spectrum usage maps will be local (*e.g.*, local interference; the hidden node problem, etc.), the cost of sensing channels and of exchanging information can be high, and the variation in the timescales involved (coherence time for channel occupancy, vs. time required to transmit, receive and exploit this information) can be large. Hence, in this paper we focus on pairs of secondary users arriving at consensus. It is worth stressing that we do not assume the existence of a pre-determined coordination channel, since this is at the heart of the problem that DSA must solve. The incorporation of energy constraint can significantly complicate the cognitive MAC design. Under an energy constraint, sensing decisions should be made based on not only spectrum occupancy dynamics but also channel fading statistics, and access decisions should take into account not only the availability but also the fading condition of the sensed channels. This makes the optimal sensing and access strategies opportunistic in both spectrum and time. Even the residual energy of the secondary user will play an important role in decision-making. For example, when the battery is depleting, should the user wait for increasingly better channel realizations for transmission or should it lower the requirement on channel given that sensing also costs energy? Clearly, such decisions depend on the energy consumption characteristics of secondary users.

As a starting point to a detailed study of energy-constrained OSA, this paper establishes the fundamental limit on the expected number of information bits that can be delivered by a secondary user during its battery lifetime. By absorbing the residual energy level of the secondary user into the state space, we show that the energy-constrained OSA problem can be formulated as an unconstrained POMDP. Based on the theory of POMDP, we obtain the optimal sensing and access policy which not only provides a performance benchmark but also enables us to study the impact of spectrum occupancy dynamics, channel fading statistics, and energy consumption characteristics of the secondary user on the optimal sensing and access decisions. However, our complexity analysis indicates that the optimal policy is computationally expensive. We therefore exploit the underlying structure of the problem to reduce the computational complexity of the optimal policy. We also provide a suboptimal solution whose computational complexity can be systematically traded off with its performance. Referred to as the greedy-w strategy, this approach maximizes the throughput of the secondary user in a fixed time window of w slots. Simulation result shows that as the window size w increases, the performance of the greedy-w strategy quickly approaches the optimal performance.

Army Relevance The Future Combat System (FCS) (`http://www.army.mil/fcs/index.html`) relies on technologies that enable a fully mobile, fully communicating, agile, situation-aware, and survivable lightweight Army force with internetted C4ISR systems. The FCS Brigade Combat Team is a network of soldiers, interoperating with other networks: unattended ground sensors, small unmanned vehicle systems, non-line-of-sight weapons systems, UAVs, possibly armed robots, etc. The number of radio units in the *soldier system* itself is expected to increase. An enabling technology is jointly designed physical (PHY) and MAC protocols that can support sensing, detection, and exploitation of spectrum opportunities, with strict constraints on the effect on primary users. To this end, an energy-efficient, spectrum-adaptive medium access approach will have a significant impact on achieving agile and situation-aware communications for lightweight army forces. Frequency allocation has been a manual task for well over a century now, and thus cannot adapt to battlefield tempo. Thus the approaches outlined here provide a framework for the automatic management of spectrum, and will be critical in dynamic coalition operations, as joint forces come together and separate. They could also be useful in peacekeeping scenarios where military, other government, and NGO networks must co-exist. Multi-channel systems such as JTRS could also benefit from the adaptive multi-channel cognitive MAC reported in this paper.

2. PROBLEM STATEMENT

We consider a spectrum consisting of N channels (*e.g.*, frequency bands), each with bandwidth B_n ($n = 1, \cdots, N$). The spectrum is licensed to a primary network whose users communicate according to a synchronous slot structure. Let $S_n \in \{0 \text{ (occupied)}, 1 \text{ (idle)}\}$ denote the availability of channel n in a slot. We assume that the spectrum occupancy $\mathbf{S} \stackrel{\Delta}{=} [S_1, \ldots, S_N] \in \{0,1\}^N$ follows a discrete Markov process with 2^N states.

Consider an ad hoc secondary network whose users independently seek instantaneous spectrum opportunities in these N channels. Each secondary user is powered by a battery with initial energy $\mathcal{E}_0$. At the beginning of each slot, a secondary user with data to transmit chooses at most M ($1 \leq M \leq N$) channels to sense and then decides whether to access these channels according to the sensing outcomes. Our goal is to determine the sensing and access decisions sequentially in each slot so as to maximize the total expected number of information bits that can be delivered by a secondary user during its battery lifetime. For ease of presentation, we assume $M = 1$. Our results can be generalized to $M > 1$.

2.1 Protocol Structure

Since a channel only presents an opportunity to a pair of secondary users if it is available at both the transmitter and the receiver, spectrum opportunities need to be identified jointly by the transmitter and the receiver (Zhao et al., 2005b). Below, we briefly describe the protocol structure.

Suppose that the transmitter and the receiver have tuned to the same channel after the initial handshake as described in (Zhao et al., 2005b). At the beginning of a slot, the transmitter and the receiver hop to the same channel[1]. If the channel is sensed to be available, the transmitter generates a random backoff time. If the channel remains idle when its backoff time expires, it transmits a short request-to-send (RTS) message to the receiver, indicating that the channel is available at the transmitter. Upon receiving the RTS, the receiver estimates the channel fading condition using the RTS, and then replies with a clear-to-send (CTS) message if the channel is also available at the receiver. The receiver also informs the transmitter of the current fading condition by piggybacking the estimated channel state to the CTS. After a successful exchange of RTS-CTS, the transmitter and the receiver can communicate over this channel. At the end of this slot, the receiver acknowledges every successful data transmission. Note that at the beginning of each slot, the transmitter and the receiver can also choose not to hop to any channel and turn to sleep mode until the beginning of the next slot.

2.2 Energy Model

We assume that channel between the secondary user and its destination follows a block fading model, *i.e.*, the channel gain in a slot is a random variable (RV), identically and independently distributed (i.i.d.) across slots but not necessarily i.i.d. across channels.

Let $E_{\text{tx}}(n)$ denote the energy consumed in transmitting over channel n in a slot. Note that $E_{\text{tx}}(n)$ is a RV depending on the current fading condition of channel n. In general, the better the channel condition, the lower the required transmission energy. Let L be the number of power levels at which the secondary user can transmit and ε_k the energy consumed in transmitting at the k-th power level in a slot. The transmission energy consumption $E_{\text{tx}}(n)$ thus has realizations restricted to a finite set $\mathcal{E}_{\text{tx}}$ given by

$$E_{\text{tx}}(n) \in \mathcal{E}_{\text{tx}} \stackrel{\Delta}{=} \{\varepsilon_k\}_{k=0}^{L}, \tag{1}$$

[1]Note that the protocols developed in this paper can ensure transceiver synchronization in a distributed fashion. See details in Sec **3.3**

where $0 < \varepsilon_1 < \ldots < \varepsilon_L < \infty$ and $\varepsilon_0 = 0$ indicates that the secondary user does not transmit. We also consider the energy e_s consumed by the secondary user in sensing a channel and the energy e_p consumed in its sleeping mode.

Let E denote the residual energy level of a secondary user at the beginning of a slot. Note that E is a RV determined by the channel conditions and the sensing and access decisions in all previous slots. Thus, E belongs to finite set $\mathcal{E}_r$ given by

$$\begin{aligned} E \in \mathcal{E}_r \stackrel{\Delta}{=} \{e : e = \mathcal{E}_0 - \sum_{k=0}^{L} c_k(e_s + \varepsilon_k) - ce_p, \\ e \geq 0, c, c_k \geq 0, c, c_k \in \mathbb{Z}\} \cup \{0\}, \end{aligned} \tag{2}$$

where c_k is the number of slots when the secondary user chooses to sense a channel and then transmit over it at the k-th power level and c is the number of slots when the secondary user turns to sleeping mode.

3. OPTIMAL ENERGY-CONSTRAINED OSA

The energy-constrained OSA problem can be formulated as a constrained POMDP, which is usually more difficult to solve than an unconstrained one. By absorbing the residual energy level of the secondary user into the state space, we reduce a constrained POMDP to an unconstrained one. Based on the theory of POMDP, we obtain the optimal spectrum sensing and access policy.

3.1 An Unconstrained POMDP Formulation

State Space In each slot, the network state is characterized by the current spectrum occupancy state $\mathbf{S} \in \{0, 1\}^N$ and the residual energy $E \in \mathcal{E}_r$ of the secondary user at the beginning of this slot. The state space $\mathcal{S}$ of the POMDP can be defined as

$$(\mathbf{S}, E) \in \mathcal{S} \stackrel{\Delta}{=} \{(\mathbf{s}, e) : \mathbf{s} \in \{0, 1\}^N, e \in \mathcal{E}_r\}. \tag{3}$$

Action Space After the state transition of spectrum occupancy at the beginning of each slot, the secondary user can either choose a channel $a \in \{1, \ldots, N\}$ to sense or turn to sleep ($a = 0$). If the secondary user chooses channel a to sense, then it will obtain a sensing outcome $\Theta_a \in \{0, 1, \ldots, L\}$ which reflects the occupancy state and the fading condition of the chosen channel: $\Theta_a = 0$ indicates that channel a is busy (*i.e.*, $S_a = 0$) and $\Theta_a = k$ ($k = 1, \ldots, L$) indicates that channel a is idle (*i.e.*, $S_a = 1$) and the fading condition requires the secondary user to transmit at the k-th power level (*i.e.*, $E_{\text{tx}}(a) = \varepsilon_k$). Given sensing outcome Θ_a, the secondary user decides whether to transmit over the chosen channel. Let $\Phi_a(k) \in \{0 \text{ no access}, 1 \text{ access}\}$ ($k = 0, \ldots, L$) denote the access decision under sensing outcome $\Theta_a = k$. Since we have assumed perfect spectrum sensing, the access decision under $\Theta_a = 0$ (busy) is simple: $\Phi_a(0) = 0$ (no access). In this case, secondary users will not collide with primary users.

The action space $\mathcal{A}$ consists of all sensing decisions a and access decisions $\mathbf{\Phi}_a \triangleq [\Phi_a(1), \ldots, \Phi_a(L)]$:

$$(a, \mathbf{\Phi}_a) \in \mathcal{A} \triangleq \{(0, [0, \ldots, 0])\} \cup \{(a, \boldsymbol{\phi}) : a \in \{1, \ldots, N\}, \\ \boldsymbol{\phi} \triangleq [\phi(1), \ldots, \phi(L)] \in \{0,1\}^L\}. \quad (4)$$

Note that the access decision $\mathbf{\Phi}_0$ associated with sensing action $a = 0$ (sleeping mode) is determined by $\Phi_0(k) = 0$ for all $1 \leq k \leq L$.

Network State Transition At the beginning of each slot, the spectrum occupancy state $\mathbf{S}$ transits independently of the residual energy E according to transition probabilities $\{p_{\mathbf{s},\mathbf{s}'}\}$, where $p_{\mathbf{s},\mathbf{s}'}$ denotes the probability that the spectrum occupancy state transits from $\mathbf{s} \in \{0,1\}^N$ to $\mathbf{s}' \in \{0,1\}^N$. In this paper, we assume that the spectrum occupancy dynamics $\{p_{\mathbf{s},\mathbf{s}'}\}$ are known and remain unchanged during the battery lifetime of the secondary user.

If channel $a \in \{1, \ldots, N\}$ is chosen in this slot, the secondary user will consume e_s in sensing and $\Phi_a(\Theta_a)\varepsilon_{\Theta_a}$ in transmitting. Thus, at the end of this slot, the residual energy of the secondary user reduces to $E' = \mathcal{T}_E(E \mid a, \Theta_a, \Phi_a(\Theta_a))$:

$$\mathcal{T}_E(E \mid a, \Theta_a, \Phi_a(\Theta_a)) \\ = \begin{cases} E - e_p, & a = 0, \\ \max\{E - e_s - \Phi_a(\Theta_a)\varepsilon_{\Theta_a}, 0\}, & a \neq 0, \end{cases} \quad (5)$$

where e_p is energy consumed in the sleeping mode.

Observations Due to partial spectrum sensing, the secondary user does not have full knowledge of the spectrum occupancy state in each slot. It, however, can obtain the occupancy state of the chosen channel $a \in \{1, \ldots, N\}$ from sensing outcome (*i.e.*, observation) $\Theta_a \in \{0, 1, \ldots, L\}$. Let $q_{\mathbf{s}}^{(a)}(k)$ be the probability that the secondary user observes $\Theta_a = k$ in the chosen channel a given current spectrum occupancy state $\mathbf{S} = \mathbf{s}$. Under perfect spectrum sensing, we have that

$$q_{\mathbf{s}}^{(a)}(k) = \Pr\{\Theta_a = k \,|\, \mathbf{S} = \mathbf{s}\} \\ = \begin{cases} 1_{[k \neq 0]} p_a(k), & \text{if } a \neq 0, s_a = 1, \\ 1_{[k=0]}, & \text{if } a \neq 0, s_a = 0, \end{cases} \quad (6)$$

where $p_a(k) \triangleq \Pr\{E_{\text{tx}}(a) = \varepsilon_k\}$ is the probability that the fading condition of channel n requires the secondary user to transmit at the k-th power level, and $1_{[x]}$ is the indicator function: $1_{[x]} = 1$ if x is true and 0 otherwise. Note that $\{p_a(k)\}_{k=1}^L$ are determined by the fading statistics of channel a and are independent of the spectrum occupancy state. From (6), we can see that $\sum_{k=0}^L q_{\mathbf{s}}^{(a)}(k) = 1$ for any spectrum occupancy state $\mathbf{s} \in \mathcal{S}$ and any chosen channel $a \in \{1, \ldots, N\}$. For notation convenience, we define $q_{\mathbf{s}}^{(0)}(k) = 1_{[k=0]}$.

Reward Structure At the end of each slot, the secondary user obtains a non-negative reward $R_{E,\Theta_a}^{(a,\Phi_a(\Theta_a))}$ depending on its residual energy E at the beginning of this slot, the sensing outcome Θ_a, and the sensing and access decisions $(a, \Phi_a(\Theta_a))$. Assuming that the number of information bits that can be transmitted over a channel in one slot is proportional to the channel bandwidth, we define immediate reward $R_{E,\Theta_a}^{(a,\Phi_a(\Theta_a))}$ as

$$R_{E,\Theta_a}^{(a,\Phi_a(\Theta_a))} \triangleq \begin{cases} 0, & a = 0, \\ \Phi_a(\Theta_a) B_a 1_{[E - e_s - \varepsilon_{\Theta_a} \geq 0]}, & a \neq 0. \end{cases} \quad (7)$$

Note that no reward will be accumulated once the battery energy level drops below $e_s + \varepsilon_1$, where ε_1 is the least required transmission energy. Hence, the total expected accumulated reward represents the total expected number of information bits that can be delivered by the secondary user during its battery lifetime.

Belief State At the beginning of a slot, the secondary user has the information of its own residual energy E but not the current spectrum occupancy state $\mathbf{S}$. Its knowledge of $\mathbf{S}$ based on all past decisions and observations can be summarized by a belief state $\boldsymbol{\lambda} = \{\lambda_{\mathbf{s}}\}_{\mathbf{s} \in \{0,1\}^N}$ (Smallwood and Sondik, 1971), where $\lambda_{\mathbf{s}}$ is the conditional probability (given the decision and observation history) that the network state is $\mathbf{S} = \mathbf{s}$ at the beginning of this slot prior to the state transition of spectrum occupancy.

At the end of a slot, the secondary user can update the belief state $\boldsymbol{\lambda}$ for future use based on sensing action a and sensing outcome Θ_a in this slot. Specifically, let $\boldsymbol{\lambda}' \triangleq \mathcal{T}_\lambda(\boldsymbol{\lambda} \,|\, a, k)$ denote the updated belief state whose element $\lambda'_{\mathbf{s}}$ represents the probability that the current spectrum occupancy state is $\mathbf{S} = \mathbf{s}$ given belief state $\boldsymbol{\lambda}$ at the beginning of this slot and the observation $\Theta_a = k$ of chosen channel a in the current slot. Applying Bayes rule, we obtain $\lambda'_{\mathbf{s}}$ as

$$\lambda'_{\mathbf{s}} = \Pr\{\mathbf{S} = \mathbf{s} \,|\, \boldsymbol{\lambda}, a, k\} \\ = \begin{cases} \sum_{\mathbf{s}'} \lambda_{\mathbf{s}'} p_{\mathbf{s}',\mathbf{s}}, & a = 0, \\ \dfrac{\sum_{\mathbf{s}'} \lambda_{\mathbf{s}'} p_{\mathbf{s}',\mathbf{s}} 1_{[s_a = 1_{[k \neq 0]}]}}{\sum_{\mathbf{s}''} \sum_{\mathbf{s}'} \lambda_{\mathbf{s}'} p_{\mathbf{s}',\mathbf{s}''} 1_{[s''_a = 1_{[k \neq 0]}]}}, & a \neq 0, \end{cases} \quad (8)$$

where the summations are taken over the space $\{0,1\}^N$ of spectrum occupancy state $\mathbf{S}$. Note that when the secondary user turns to sleeping mode ($a = 0$), no observation is made and the belief state is updated according to the spectrum occupancy dynamics $\{p_{\mathbf{s},\mathbf{s}'}\}$.

Unconstrained POMDP Formulation We have formulated the energy-constrained OSA as a POMDP problem. A policy π of this POMDP is defined as a sequence of functions:

$$\pi \triangleq [\mu_1, \mu_2, \ldots], \quad \mu_t : [0,1]^{2^N} \times \mathcal{E}_r \to \mathcal{A},$$

where $\{a, \mathbf{\Phi}_a\} = \mu_t(\boldsymbol{\lambda}, E)$ maps every information state $(\boldsymbol{\lambda}, E)$, which consists of belief state $\boldsymbol{\lambda} \in [0,1]^{2^N}$ and residual energy $E \in \mathcal{E}_r$, at the beginning of slot t to a sensing decision $a \in \{0, 1, \ldots, N\}$ and a set of access decisions $\mathbf{\Phi}_a = [\Phi_a(1), \ldots, \Phi_a(L)] \in \{0,1\}^L$.

The design objective is to find the optimal policy π^* that maximizes the total expected reward:

$$\pi^* = \arg\max_{\pi} \mathbb{E}_\pi \left[\sum_{t=1}^{\infty} R^{(a, \Phi_a(\Theta_a))}_{(\mathbf{S}, E), \Theta_a}(t) \;\middle|\; \boldsymbol{\lambda}_0 \right], \qquad (9)$$

where $\boldsymbol{\lambda}_0$ is the initial belief state given by the stationary distribution of spectrum occupancy. We thus have an unconstrained POMDP.

3.2 Optimal Policy

Let $V(\boldsymbol{\lambda}, E)$ be the value function, which denotes the maximum expected remaining reward that can be accrued when the current information state is $(\boldsymbol{\lambda}, E)$. We notice from (7) that the value function is given by $V(\boldsymbol{\lambda}, E) = 0$ for any information state $(\boldsymbol{\lambda}, E)$ with residual energy $E < e_s + \varepsilon_1$. For any other information state, its value function $V(\boldsymbol{\lambda}, E)$ is the unique solution to the following equation:

$$\begin{aligned} V(\boldsymbol{\lambda}, E) = \max_{(a, \boldsymbol{\phi}) \in \mathcal{A}} \sum_{k=0}^{L} u_k^{(a)} [R^{(a, \phi(k))}_{E,k} \\ + V(\mathcal{T}_\lambda(\boldsymbol{\lambda} \,|\, a, k), \mathcal{T}_E(E \,|\, a, k, \phi(k)))], \end{aligned} \qquad (10)$$

where $\mathcal{T}_\lambda(\boldsymbol{\lambda} \,|\, a, k)$ is the updated belief state given in (8), $\mathcal{T}_E(E \,|\, a, k, \phi(k))$ is the reduced battery energy given in (5), and $u_k^{(a)} \triangleq \Pr\{\Theta_a = k \,|\, \boldsymbol{\lambda}\}$ is the probability of observing $\Theta_a = k$ given belief state $\boldsymbol{\lambda}$, which is determined by the spectrum occupancy dynamics and the channel fading statistics:

$$u_k^{(a)} = \sum_{\mathbf{s}' \in \{0,1\}^N} \lambda_{\mathbf{s}'} \sum_{\mathbf{s} \in \{0,1\}^N} p_{\mathbf{s}', \mathbf{s}} \quad q_{\mathbf{s}}^{(a)}(k). \qquad (11)$$

In principle, by solving (10), we can obtain the optimal sensing and access actions $(a^*, \mathbf{\Phi}_a^*)$ that achieve the maximum expected reward $V(\boldsymbol{\lambda}, E)$ for each possible information state $(\boldsymbol{\lambda}, E)$. We can also obtain the maximum expected number of information bits V_{opt} that can be delivered by a secondary user during its battery lifetime as $V_{opt} = V(\boldsymbol{\lambda}_0, \mathcal{E}_0)$, where $\boldsymbol{\lambda}_0$ is the initial belief state.

3.3 Transceiver Synchronization

Transceiver synchronization is a key issue in distributed MAC design for OSA networks (Zhao et al., 2005b). Specifically, a secondary user and its intended receiver need to hop to the same channel at the beginning of each slot in order to communicate (Zhao et al., 2005b). Here we show that the optimal sensing and access policy developed in Section **3.2** ensures transceiver synchronization.

The protocol structure described in Section **2.1** ensures that both the transmitter and the receiver have the same information on the occupancy state and the fading condition of the sensed channel in each slot. Hence, at the end of each slot, the transmitter and the receiver will reach the same updated belief state $\boldsymbol{\lambda}$ using (8) and the same residual energy E of the transmitter using (5). Since the channel selection is determined by the information state $(\boldsymbol{\lambda}, E)$, the transmitter and the receiver will hop to the same channel in the next slot, *i.e.,* transceiver synchronization is maintained.

4. OPTIMAL POLICY WITH REDUCED COMPLEXITY

Although the value function given in (10) can be solved iteratively, it is computationally expensive. In this section, we first identify the sources of high complexity of the optimal policy and then reduce the complexity accordingly.

4.1 Complexity of the Optimal Policy

We measure the computational complexity of a policy as the number of multiplications required to obtain all sensing and access actions during the secondary user's battery lifetime T when initial belief state and battery energy are given.

From (10), we notice that the optimal sensing and access action in the first slot depends on the value functions of all possible information states during the battery lifetime T. Hence, the computational complexity of the optimal policy is determined by the number of multiplications required to calculate the value functions of all possible information states.

Following the complexity analysis in (Djonin et al., 2007), we can calculate the number of all possible information states $(\boldsymbol{\lambda}, E)$ during the secondary user' battery lifetime. Specifically, noting from (8) that the updated belief state is the same under all non-zero sensing outcomes $(k \neq 0)$, we can see that each information state $(\boldsymbol{\lambda}, E)$ can transit to at most $L+1$ different information states under sensing action $a \neq 0$ but only one under sensing action $a = 0$. Hence, for fixed initial information state $(\boldsymbol{\lambda}_0, \mathcal{E}_0)$, the number of all possible information states is of order $\mathcal{O}((N(L+1))^{T-1})$, which is exponential in the battery lifetime T and polynomial in the number N of channels. Moreover, from (10) and (11), we can see that it requires $\mathcal{O}(3|\mathcal{A}|2^N 2^N(L+1))$ multiplications to calculate each value function, where $|\mathcal{A}|$ is the size of the action space, 2^N is the dimension of the belief state, and $L+1$ is the number of possible observations. Therefore, the computational complexity of the optimal policy is of order $\mathcal{O}(3|\mathcal{A}|2^N 2^N(L+1)(N(L+1))^{T-1})$. We can see that the complexity is mainly caused by the following three factors: 1) the number $\mathcal{O}((N(L+1))^{T-1})$ of possible information states; 2) the size $|\mathcal{A}|$ of the action space, and 3) the dimension 2^N of the belief state. We will address the first factor in Section **5**. In this section, we focus on the other two factors.

4.2 Reduction of Action Space Size

Careful inspection of (5), (7) and (10) reveals that the quantity $R_{E,k}^{(a,\phi(k))}+V(\mathcal{T}_\lambda(\boldsymbol{\lambda}\,|\,a,k),\mathcal{T}_E(E\,|\,a,k,\phi(k)))$ inside the square parenthesis of (10) only depends on the k-th entry $\phi(k)$ of the access decision $\boldsymbol{\phi}$ and is independent of $\phi(i)$ $(i\neq k)$. We can thus simplify (10) as

$$V(\boldsymbol{\lambda},E)=\max_{a\in\{0,1,\ldots,N\}}\{\sum_{k=0}^{L}u_k^{(a)}\max_{\phi(k)\in\{0,1\}}[R_{E,k}^{(a,\phi(k))} + V(\mathcal{T}_\lambda(\boldsymbol{\lambda}\,|\,a,k),\mathcal{T}_E(E\,|\,a,k,\phi(k)))]\}. \tag{12}$$

Note that the maximization in (12) is taken over the space with size $\mathcal{O}(2NL)$ increasing linearly with the number L of power levels, while that in (10) is taken over the action space $\mathcal{A}$ whose size $\mathcal{O}(N2^L)$ increases exponentially with L.

4.3 Reduction of Belief State Dimension

Assume that the spectrum occupancy evolves independently across channels. It has been shown in (Zhao et al., 2005a) that $\boldsymbol{\omega}\triangleq[\omega_1,\ldots,\omega_N]$, where ω_n denotes the probability (conditioned on all previous decisions and observations) that channel n is available at the beginning of a slot prior to the state transition, is a sufficient statistic for belief state $\boldsymbol{\lambda}$. Note that the dimension of $\boldsymbol{\omega}$ increases linearly $\mathcal{O}(N)$ with the number N of channels while that of $\boldsymbol{\lambda}$ increases exponentially $\mathcal{O}(2^N)$.

Applying the belief state $\boldsymbol{\omega}$, we can simplify the value function given in (12). Specifically, let $\alpha_n=\Pr\{S_n'=1\,|\,S_n=0\}$ denote the probability that channel n transits from 0 (busy) to 1 (idle) and $\beta_n=\Pr\{S_n'=1\,|\,S_n=1\}$ the probability that channel n remains idle. Then, (12) reduces to

$$\begin{aligned}\hat{V}(\boldsymbol{\omega},E)=\max_{a\in\{0,1,\ldots,N\}}&\{(1-\omega_a')\\ &\times\hat{V}(\hat{\mathcal{T}}_\lambda(\boldsymbol{\omega}\,|\,a,0),\mathcal{T}_E(E\,|\,a,0,0))\\ &+\omega_a'\sum_{k=1}^{L}p_a(k)\max_{\phi(k)\in\{0,1\}}[R_{E,k}^{(a,\phi(k))}\\ &+\hat{V}(\hat{\mathcal{T}}_\lambda(\boldsymbol{\omega}\,|\,a,k),\mathcal{T}_E(E\,|\,a,k,\phi(k)))]\},\end{aligned} \tag{13}$$

where $\omega_0'\triangleq 0$, $\omega_a'=\omega_a\beta_a+(1-\omega_a)\alpha_a$ $(a\in\{1,\ldots,L\})$ is the probability that channel a is available in the current slot given $\boldsymbol{\omega}$, $\mathcal{T}_E(E\,|\,a,k,\phi_a(k))$ is the reduced battery energy given in (5), and the updated belief state $\hat{\boldsymbol{\omega}}\triangleq[\omega_1,\ldots,\omega_N]=\hat{\mathcal{T}}_\lambda(\boldsymbol{\omega}\,|\,a,k)$ is given by

$$\hat{\omega}_n=\begin{cases}0, & \text{if } a\neq 0, n=a, k=0,\\ 1, & \text{if } a\neq 0, n=a, k\neq 0,\\ \omega_n', & \text{otherwise.}\end{cases} \tag{14}$$

5. SUBOPTIMAL ENERGY-CONSTRAINED OSA

We notice from (10) that the optimal sensing and access decisions in a slot rely on the value functions of all possible information states in the remaining slots, which significantly increases the computational complexity of the optimal policy. In this section, we provide a suboptimal solution to energy-constrained OSA, which reduces the number of value functions used in decision-making. We show that the computational complexity of this suboptimal strategy can be traded off with its performance.

5.1 The Greedy-w Approach

Referred to as greedy-w approach, the proposed strategy maximizes the total expected reward in a time window of w slots. Let $Y_w^{(a)}(\boldsymbol{\lambda},E)$ denote the maximum reward that can be accumulated in a window of w slots given information state $(\boldsymbol{\lambda},E)$ and sensing action a. We can calculate $Y_w^{(a)}(\boldsymbol{\lambda},E)$ recursively by

$$\begin{aligned}&Y_0^{(a)}(\boldsymbol{\lambda},E)=0\\ &Y_w^{(a)}(\boldsymbol{\lambda},E)=\sum_{k=0}^{L}u_k^{(a)}\max_{\phi(k)\in\{0,1\}}[R_{E,k}^{(a,\phi(k))}\\ &\quad+\max_{b\in\{0,1,\ldots,N\}}Y_{w-1}^{(b)}(\mathcal{T}_\lambda(\boldsymbol{\lambda}\,|\,a,k),\mathcal{T}_E(E\,|\,a,k,\phi(k)))],\end{aligned} \tag{15}$$

where $u_k^{(a)}$, $\mathcal{T}_\lambda(\boldsymbol{\lambda}\,|\,a,k)$, and $\mathcal{T}_E(E\,|\,a,k,\phi(k))$ are given in (11), (8), and (5), respectively. From (15), we can see that for any w, $Y_w^{(a)}(\boldsymbol{\lambda},E)=0$ if $E<e_s+\varepsilon_1$.

Given belief state $\boldsymbol{\lambda}$ and residual energy E of the secondary user at the beginning of a slot, the greedy-w approach chooses channel a_w that maximizes the reward obtained in the next w slots to sense, *i.e.,*

$$a_w=\arg\max_{a\in\{0,1,\ldots,N\}}Y_w^{(a)}(\boldsymbol{\lambda},E). \tag{16}$$

Given sensing outcome $k\in\{1,\ldots,L\}$, the access decision $\phi_{a_w}(k)$ of the greedy-w approach is given by

$$\phi_{a_w}(k)=\arg\max_{\phi\in\{0,1\}}\{R_{E,k}^{(a_w,\phi)} + \max_{b\in\{1,\ldots,N\}}Y_{w-1}^{(b)}(\mathcal{T}_\lambda(\boldsymbol{\lambda}\,|\,a_w,k),\mathcal{T}_E(E\,|\,a_w,k,\phi))]\}. \tag{17}$$

Since its channel selection is determined by the information state $(\boldsymbol{\lambda}, E)$, the greedy-$w$ approach ensures transceiver synchronization as shown in Section **3.3**.

5.2 Complexity Vs. Performance

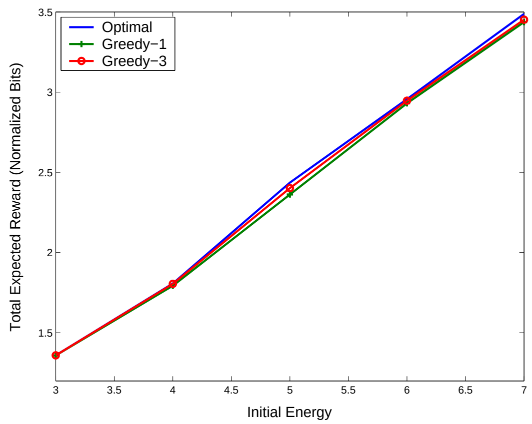

Fig. 1. The number of information bits that can be transmitted by the secondary user during its battery lifetime. $N = 2$, $[B_1, B_2] = [1, 1]$, $[\alpha_1, \alpha_2] = [0.2, 0.6]$, $[\beta_1, \beta_2] = [0.8, 0.8]$, $e_s = 0.5$, $e_p = 0.1$, $L = 2$, $\mathcal{E}_{\text{tx}} = \{1, 2\}$, $p_n(1) = 0.8$, $p_n(2) = 0.2$ for $n = 1, 2$.

We can see from (16) and (17) that the sensing and access decisions made by the greedy-w approach in a slot only depend on the value functions of all possible information states in the next w slots. Hence, the total number of value functions required to determine the sensing and access decisions during battery lifetime T is of order $\mathcal{O}((N(L+1))^{w-1}T)$, which is linear in T. Clearly, the computational complexity of greedy-w approach increases with w.

In Fig. 1, we compare the performance of the greedy-w approach with the optimal performance $V(\boldsymbol{\lambda}_0, \mathcal{E}_0)$ as the initial energy $\mathcal{E}_0$ increases. We consider $N = 2$ independently evolving channels with different occupancy dynamics. As the window size w increases, the performance of the greedy-w approach improves. It quickly approaches the optimal performance as w increases.

The above observations show that the computational complexity of the greedy-w approach increases while its performance loss as compared to the optimal performance decreases as the window size w increases. Hence, by choosing a suitable w, the greedy-w approach can achieve a desired tradeoff between complexity and performance.

6. NUMERICAL EXAMPLES

Careful inspection of (10) reveals that a sensing and access action $(a, \boldsymbol{\phi}) \in \mathcal{A}$ affects the total expected reward in three ways: 1) it acquires an immediate reward $R_{E,k}^{(a,\phi(k))}$ in this slot; 2) it transforms the current belief state $\boldsymbol{\lambda}$ to $\mathcal{T}_\lambda(\boldsymbol{\lambda}, a, k)$ which summarizes the information of spectrum occupancy up to this slot; 3) it causes a reduction in battery energy from E to $\mathcal{T}_E(E, a, k, \phi(k))$, leading to a shorter remaining battery lifetime. Hence, to maximize the total expected reward during battery lifetime, the optimal sensing and access policy should achieve a tradeoff among gaining instantaneous reward, gaining information for future use, and conserving energy. In this section, we study the impact of spectrum occupancy dynamics, channel fading statistics, and energy consumption characteristics on the optimal sensing and access actions.

To sense or not to sense? The secondary user may choose to sense in order to gain immediate reward and channel occupancy information, but not to sense in order to conserve energy. Hence, the optimal decision on whether to sense should strike a balance between gaining reward/information and conserving energy. In Fig. 2, we study the optimal sensing decision $1_{[a^* \neq 0]}$ in a particular slot under different spectrum occupancy dynamics and belief states.

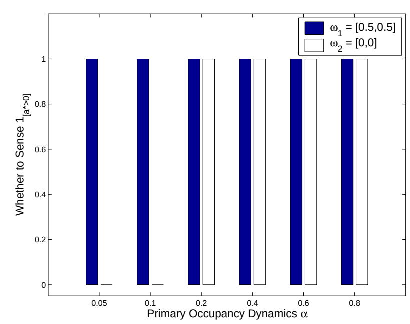

Fig. 2. The optimal decision $1_{[a^* \neq 0]}$ on whether to sense under different spectrum occupancy dynamics and belief states. $N = 2$, $[B_1, B_2] = [1, 1]$, $\mathcal{E}_0 = 4$, $e_s = 0.6$, $e_p = 0.1$, $L = 2$, $\mathcal{E}_{\text{tx}} = \{1, 2\}$, $p_n(1) = p_n(2) = 0.5$ for $n = 1, 2$.

We consider $N = 2$ independently evolving channels with identical spectrum occupancy dynamics $\alpha_1 = \alpha_2 = \alpha$ and $\beta_1 = \beta_2 = \beta$. We assume that $\beta = 1 - \alpha$. Hence, the stationary distribution of spectrum occupancy state $\mathbf{S}$ is given by $\boldsymbol{\omega}_1 = [0.5, 0.5]$. Consider another belief state $\boldsymbol{\omega}_2 = [0, 0]$ with which the secondary user has full information on the spectrum occupancy prior to the state transition in this slot. Conditioned on the belief states at the beginning of this slot, the conditional probability that channel n is available can be calculated as $\Pr\{S_n = 1 \,|\, \boldsymbol{\omega}_1\} = 0.5$ and $\Pr\{S_n = 1 \,|\, \boldsymbol{\omega}_2\} = \alpha$ for $n = 1, 2$. From Fig. 2, we find that the secondary user chooses not to sense only when the conditional probability $\Pr\{S_n = 1 \,|\, \boldsymbol{\omega}\}$ that the channel is available is very small. We also find that the secondary user always chooses to sense if the belief state is given by the stationary distribution $\boldsymbol{\omega}_1$ of the spectrum occupancy dynamics. The reason behind this is the monotonicity of the value function $\hat{V}(\boldsymbol{\omega}, E)$ in terms of battery energy E. Specifically, if the secondary user chooses not to sense, then its belief state at the beginning of the next slot will remain $\boldsymbol{\omega}_1$ but its battery energy will be reduced by e_p due to energy consumption in the sleeping mode. The maximum total expected reward that can be obtained is thus given by $\hat{V}(\boldsymbol{\omega}_1, E - e_p)$. Since $\hat{V}(\boldsymbol{\omega}, E)$ increases with the battery energy E for every fixed $\boldsymbol{\omega}$, we have $\hat{V}(\boldsymbol{\omega}_1, E) \geq \hat{V}(\boldsymbol{\omega}_1, E - e_p)$ and hence the secondary user should choose to sense whenever it has a stationary belief state.

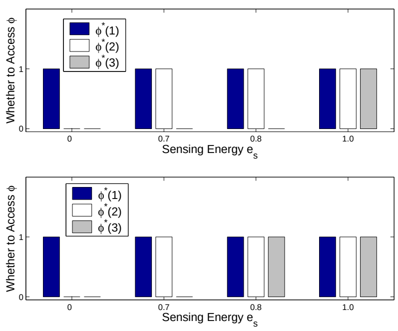

Fig. 3. The optimal access decision under different sensing energy consumptions e_s and channel fading statistics. $N = 2$, $[B_1, B_2] = [1, 1]$, $\mathcal{E}_0 = 8$, $e_p = 0.1$, $L = 3$, $\mathcal{E}_{\text{tx}} = \{1, 2, 3\}$. In the upper plot, $p_n(1) = 0.5, p_n(2) = 0.3, p_n(3) = 0.2$ for $n = 1, 2, 3$. In the lower plot, $p_n(1) = 0.3, p_n(2) = 0.3, p_n(3) = 0.4$.

To access or not to access? Without an energy constraint, the secondary user should always access the channel that is sensed to be available. However, under an energy constraint, the access decision should take into account both the energy consumption characteristics and the channel fading statistics. For example, when the sensed channel is available but has poor fading condition, should the secondary user access this channel to gain immediate reward or wait for better channel realizations to conserve energy? In Fig. 3, we study the optimal access decision ϕ^* under different sensing energy consumptions e_s and channel fading statistics $\{p_n(k)\}_{k=1}^{L}$. We find that when sensing energy consumption e_s is negligible, the secondary user should refrain from transmission under poor channel conditions and wait for the best channel realization. However, when e_s is large, it should always grab the instantaneous opportunity regardless of the fading condition because the sensing energy consumed in waiting for the best channel realization may exceed the extra energy consumed in combating the poor channel fading.

The access decision should also take into account the channel fading statistics. Compare the optimal access decisions in the upper and the lower plots of Fig. 3 when sensing energy is $e_s = 0.8$. We find that if the probability that the channel experiences deep fading is small (see the upper plot), the secondary user should avoid transmitting under poor channel realizations because the waiting time for a better channel realization is short and hence the energy wasted in waiting can still be lower than the extra energy needed to combat the poor channel condition. On the other hand, if the channel tends to have poor fading conditions (see the lower plot), the secondary user should focus on gaining immediate reward because of the long waiting time for better channel realizations.

CONCLUSIONS

In this paper, we obtained the optimal sensing and access policy for energy-constrained OSA by formulating the resulting problem as an unconstrained POMDP. We proposed a suboptimal solution, called greedy-w, whose computational complexity can be systematically traded off with its performance. Numerical results demonstrated that the optimal sensing and access decisions should take into account not only the spectrum occupancy dynamics but also the channel fading statistics and the energy consumption characteristics of the secondary user. The cognitive MAC developed in this paper is energy-efficient and spectrum adaptive, and should find applications in FCS.

REFERENCES

DARPA, 2005: DARPA: The Next Generation Program, http://www.darpa.mil/ato/programs/xg/index.htm.

DySpan, 2005: *Proceedings of the first IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks*, Baltimore, USA.

CrownCom, 2006: *Proceedings of the first International Conference on Cognitive Radio Oriented Wireless Networks and Communications*, Mykonos Island, Greece.

Mangold, S., Z. Zhong, K. Challapali, and C. Chou, 2004: Spectrum agile radio: radio resource measurements for opportunistic spectrum usage. *Proc. of Globecom*, 3467-3471.

Larcher, A., H. Sun, M. van der Shaar, and Z. Ding, 2004: Decentralized transmission strategies for delay-sensitive applications over spectrum agile networks. *Proc. of International Packet Video Workshop*, Irvine, CA.

Papadimitratos, P., S. Sankaranarayanan, and A. Mishra, 2005: A bandwidth sharing approach to improve licensed spectrum utilization. *IEEE Commun. Magazine*, **43**, 10-14.

Zhao, Q., L. Tong, and A. Swami, 2005a: Decentralized cognitive MAC for dynamic spectrum access. *Proc. of the first IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySpan)*.

Zhao, Q., L. Tong, A. Swami, and Y. Chen, 2006: Cross-layer design of opportunistic spectrum access in the presence of sensing error. *Proc. of the 40th Conference on Information Science and Systems (CISS)*.

Zhao, Q., L. Tong, and A. Swami, 2005b: A cross-layer approach to cognitive MAC for spectrum agility. *Proc. of Asilomar Conference on Signals, Systems, and Computers*.

Smallwood, R. and E. Sondik, 1971: The optimal control of partially observable Markov processes over a finite horizon. *Operations Research*, 1071–1088.

Djonin, D., Q. Zhao, and V. Krishnamurthy, 1971: Optimality and complexity of opportunistic spectrum access: a partially observed markov decision process approach. *Submitted to IEEE ICC*.