10

Cognitive MAC Protocols for Dynamic Spectrum Access

Qing Zhao[1], Yunxia Chen[1], and Ananthram Swami[2]

[1] Department of Electrical and Computer Engineering
University of California at Davis†, USA
{qzhao,yxchen}@ece.ucdavis.edu
[2] Army Research Laboratory, USA
aswami@arl.army.mil

10.1 Introduction

The paradox between the overly crowded spectrum and the pervasiveness of idle frequency bands in both time and space indicates that spectrum shortage results from the current static spectrum management policy rather than the physical scarcity of usable radio frequencies [1]. To improve spectrum efficiency, researchers in the engineering, economics, and regulation communities have been actively searching for better spectrum management strategies. Under the general term of dynamic spectrum access, various spectrum reform ideas have been proposed. We provide below a taxonomy to illustrate the relationship among these diverse ideas.

10.1.1 Dynamic Spectrum Access

The term "dynamic spectrum access" has broad connotations that encompass various approaches to spectrum reform, and should be contrasted with the current static spectrum management policy. As illustrated in Fig. 10.1, dynamic spectrum access strategies can be generally categorized under three models.

1. Dynamic exclusive use model: This model maintains the basic structure of the current spectrum regulation policy: spectrum bands are licensed to services for exclusive use. The main idea is to introduce flexibility to improve spectrum efficiency. Two approaches have been proposed under this model: *spectrum property rights* [2, 3] and *dynamic spectrum allocation* [4]. The former approach allows licensees to sell and trade spectrum and to freely choose technology. Economy

† This work was supported in part 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.

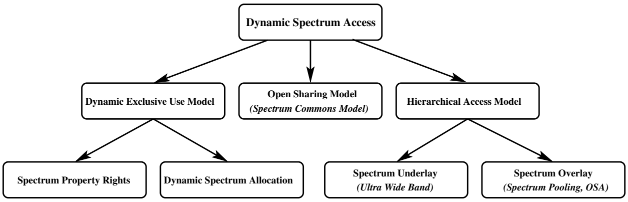

Fig. 10.1. A taxonomy of dynamic spectrum access.

and market will thus play a more important role in driving toward the most profitable use of this limited resource. Note that even though licensees have the right to lease or share the spectrum for profit, such sharing is not mandated by the regulation policy.

The second approach, dynamic spectrum allocation, was brought forth by the European DRiVE project [4]. It aims to improve spectrum efficiency through dynamic spectrum assignment by exploiting the spatial and temporal traffic statistics of different services. Similar to the current static spectrum allotment policy, such strategies allocate, at a given time and region, a portion of the spectrum to a radio access network for its exclusive use. This allocation, however, varies at a much faster scale.

Based on an exclusive-use model, these approaches cannot eliminate white space in spectrum resulting from the bursty nature of wireless traffic.

2. Open sharing model: Also referred to as spectrum commons [5, 6], this model employs open sharing among peer users as the basis for managing a spectral region. Advocates of this model draw support from the phenomenal success of wireless services operating in the unlicensed ISM band (e.g., WiFi). Centralized [7, 8] and distributed [9–11] spectrum sharing strategies have been initially investigated to address technological challenges under this spectrum management model.
3. Hierarchical access model: Built upon a hierarchical access structure with primary and secondary users, this model can be considered as a hybrid of the above two. The basic idea is to open licensed spectrum to secondary users and limit the interference perceived by primary users (licensees). Two approaches to spectrum sharing between primary and secondary users have been considered: *spectrum underlay* and *spectrum overlay*.

The underlay approach imposes severe constraints on the transmission power of secondary users so that they operate below the noise floor of primary users. By spreading transmitted signals over a wide frequency band (UWB), secondary users can potentially achieve short-range high data rate with extremely low transmission power. Based on a worst-case assumption that primary users transmit all the time, this approach does not exploit spectrum white space.

Spectrum overlay was first envisioned by Mitola [12] under the term "spectrum pooling" and then investigated by the DARPA XG program [13] under the term "opportunistic spectrum access (OSA)". Differing from spectrum underlay, this approach does not necessarily impose severe restrictions on the transmission power of secondary users, but rather on when and where they may transmit. It directly targets at spatial and temporal spectrum white space by allowing secondary users to identify and exploit local and instantaneous spectrum availability in a non-intrusive manner.

Compared to the dynamic exclusive use and open sharing models, this hierarchical model is perhaps the most compatible with the current spectrum management policy and legacy wireless systems. Furthermore, the underlay and overlay approaches can be employed simultaneously to further improve spectrum efficiency.

We point out that the hierarchical access model is sometimes categorized under the open sharing model (see, e.g., [6]). Spectrum sharing between primary and secondary users is, however, fundamentally different from spectrum sharing among peer users in both technical and regulatory aspects. We have thus separated the hierarchical access model from the open sharing model in the above taxonomy.

10.1.2 Cognitive Radio

Cognitive radio is often used as a synonym for dynamic spectrum access. We provide below a brief introduction to software-defined radio and cognitive radio.

The terms "software-defined radio" and "cognitive radio" were coined by Mitola in 1991 and 1998, respectively. Software-defined radio, sometimes shortened to software radio, is generally a multi-band radio that supports multiple air interfaces and protocols and is reconfigurable through software run on DSP or general-purpose microprocessors [14]. Cognitive radio, built upon a software radio platform, is a context-aware intelligent radio capable of autonomous reconfiguration by learning and adapting to the communication environment [15]. While dynamic spectrum access is certainly an important application of cognitive radio, cognitive radio represents a much broader paradigm where many aspects of communication systems can be improved via cognition.

10.2 Cognitive MAC for Opportunistic Spectrum Access

In this chapter, we focus on the overlay approach under the hierarchical access model (see Fig. 10.1). The term opportunistic spectrum access (OSA) will be adopted throughout. Our emphasis is on the design of cognitive medium access control (MAC) protocols for secondary users in OSA networks.

10.2.1 Basic Components of Cognitive MAC

Basic design components of cognitive MAC for OSA include (1) a sensing policy for real-time decisions about whether to sense and where in the spectrum to sense and (2) an access policy that determines whether to access based on the sensing outcomes.

The purpose of the sensing policy is twofold: to identify a spectrum opportunity for immediate access and to obtain statistical information on spectrum occupancy for improved future decisions. A balance must be reached between these two often conflicting objectives, and the trade-off should adapt to the bursty traffic and energy constraint of the secondary user. For example, when there are energy costs associated with sensing, a secondary user may decide to skip sensing when its current estimate of spectrum occupancy indicates that no channels are likely to be idle. Clearly, such decisions should balance the reward in energy savings with the cost in lost spectrum information and potentially missed spectrum opportunities.

The objective of the access policy, on the other hand, is to minimize the chance of overlooking an opportunity without violating the constraint of being non-intrusive. Whether the secondary user should adopt an aggressive or a conservative access policy depends on the operating characteristics (probability of false alarm vs. probability of miss detection, and permissible level of interference) of the spectrum sensor. A joint design of MAC protocols and spectrum sensors at the physical layer is thus necessary to achieve optimality. Energy constraints will further complicate the design of access policies. For energy-constrained OSA in fading environments, the secondary user may avoid transmission when the sensed channel is in a deep fade. Even the residual energy level will play an important role in decision-making. When the battery is depleting, should the user wait for increasingly better channel conditions for transmission or should it lower the requirement on channel conditions given that sensing also costs energy? How is such a decision affected by the accuracy and energy consumption characteristics of the spectrum sensor? And how sensitive are such policies to incomplete models and inaccurate model parameter estimates?

The above discussion highlights some of the complexities in the design of a cognitive MAC for OSA in a dynamic network environment with fading, sensing errors, and energy constraints. It demonstrates that the optimal design of cognitive MAC for OSA calls for a cross-layer approach that integrates signal processing with networking.

In this chapter, we aim to illuminate the interactions between the physical and the MAC layers in OSA networks. We focus, in particular, on the impact of sensing errors and channel fading conditions at the physical layer on the optimal sensing and access policies at the MAC layer. In particular, we present a decision-theoretic framework first developed in [16–19]. Based on the theory of partially observable markov decision process (POMDP), this framework integrates the basic components of OSA, leading to an optimal joint design of signal processing algorithms for opportunity identification and MAC protocols for opportunity exploitation.

10.2.2 Related Work

A majority of the existing work focuses on spatial spectrum opportunities that are static or slowly varying in time. Example applications include the reuse of certain TV-bands that are not used for TV broadcast in a particular region. Due to the slow temporal variation of spectrum occupancy, real-time opportunity identification is not as critical a component in this class of applications, and the prevailing approach to OSA tackles network design in two separate steps: (1) opportunity identification assuming continuous full-spectrum sensing; (2) opportunity allocation among secondary users assuming full knowledge of spectrum opportunities. Opportunity identification in the presence of fading and noise uncertainty has been studied in [20–24]. Spatial opportunity allocation among secondary users can be found in [25–28] and references therein. Differing from these works, we focus on the exploitation of temporal spectrum opportunities resulting from the bursty traffic of primary users, For an overview of challenges and recent development in OSA, readers are referred to [29, 30].

10.3 The Network and Protocol Model

10.3.1 The Network Model

Consider a spectrum consisting of N channels,[1] each with bandwidth B_n ($n = 1, \cdots, N$). These N channels are licensed to a primary network whose users communicate according to a synchronous slot structure. The traffic statistics of the primary network are such that the occupancy of these N channels follows a discrete-time Markov process with 2^N states. Specifically, the network state in slot t is given by $S(t) \overset{\Delta}{=} [S_1(t), \cdots, S_N(t)]$ where $S_n(t) \in \{0 \text{ (occupied)}, 1 \text{ (idle)}\}$ is the occupancy state of channel n. The state diagram for $N = 3$ and a sample path of the state evolution are illustrated in Figs. 10.2 and 10.3, respectively. We assume that the spectrum usage statistics of the primary network remain unchanged for T slots. We further assume that the state transition probabilities of the underlying Markov model are known: $P_{s,s'} \overset{\Delta}{=} \Pr\{S(t+1) = s' \,|\, S(t) = s\}$, for every $s, s' \in \{0,1\}^N$. In Section 10.4.3, we discuss OSA with unknown or mismatched Markov model.

We consider a secondary network that seeks spectrum opportunities in these N channels (see Fig. 10.3). We focus on an ad hoc network where secondary users join/exit the network and sense/access the spectrum independently without exchanging local information. In each slot, a secondary user chooses a set of channels to sense and a set of channels to access. Limited by its hardware constraints and energy supply, a secondary user can sense no more than L_1 ($L_1 \leq N$) and access no more than L_2 ($L_2 \leq L_1$) channels in each slot.[2] For the ease of presentation, we assume $L_1 = L_2 = 1$. Results presented in this chapter can be extended to general cases as discussed in [17, 18, 31].

Our goal is to develop cognitive MAC protocols for the secondary network. For an ad hoc OSA network without a central coordinator or a dedicated communication channel, it is desirable to have a decentralized MAC protocol where each secondary user independently searches for spectrum opportunities, aiming at optimizing

[1] Here we use the term channel broadly. A channel can be a frequency band with specified bandwidth, a collection of spreading codes in DS-CDMA network, a set of hopping codes in FH-SS, or a set of subcarriers in an OFDM system.

[2] In principle, we can let $L_2 = N$, i.e., access decisions need not be confined to the currently sensed set of channels.

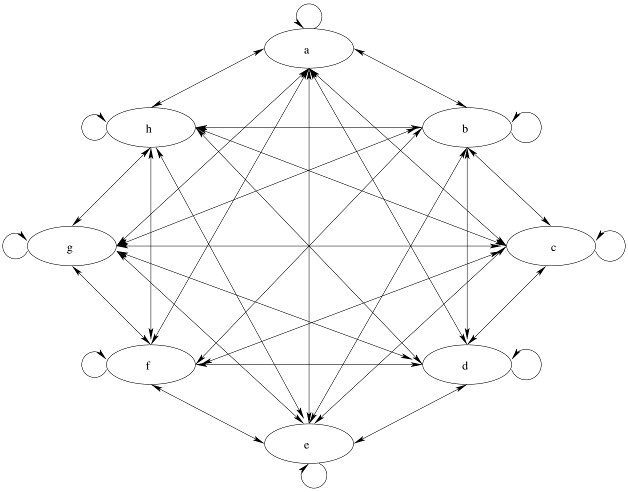

Fig. 10.2. The underlying Markov process for $N = 3$.

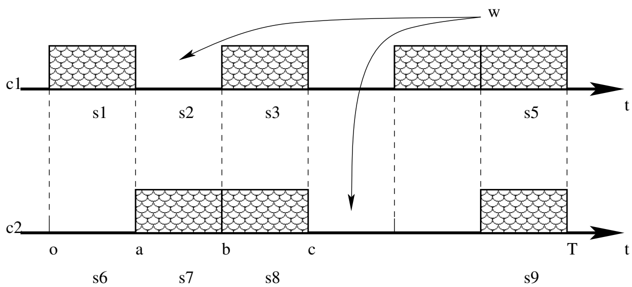

Fig. 10.3. A sample path of spectrum occupancy.

its own performance. Such decentralized protocols do not rely on cooperation among secondary users.

10.3.2 The Basic Protocol Structure

Without delving into protocol details (which are given in Sect. 10.6), we present here the basic protocol structure. At the beginning of each slot,[3] a secondary user with data to transmit chooses a channel to sense and decides whether to access based on the sensing outcome. When the secondary user decides to transmit, it generates a random backoff time, and transmits when this timer expires and no other secondary user has already accessed that channel during the backoff time. At the end of the slot, the receiver acknowledges a successful data transmission. The basic slot structure is illustrated in Fig. 10.4.

[3] Secondary users can synchronize to a slot structure broadcasted by the primary network.

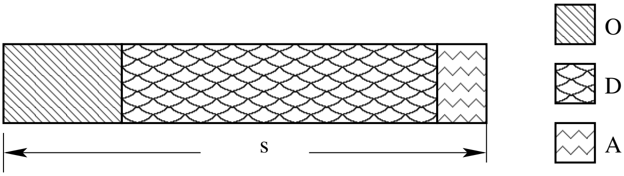

Fig. 10.4. The slot structure.

10.4 The Impact of Sensing Errors on Non-intrusive Cognitive MAC

We study the impact of sensing errors at the physical layer on the design of cognitive MAC protocols. We formulate the joint PHY-MAC design of OSA networks as a constrained partially observable Markov decision process (POMDP). Involved in the design are three basic components: a spectrum sensor at the physical layer; a sensing policy, and an access policy, both at the MAC layer.

10.4.1 Problem Formulation

10.4.1.1 Spectrum Sensor

The spectrum sensor of a secondary user detects, at the beginning of each slot, the availability of the chosen channel. It essentially performs a binary hypotheses test: $\mathcal{H}_0$ (null hypothesis indicating that the sensed channel is idle) vs. $\mathcal{H}_1$ (alternative indicating a busy channel). Let Θ_a be the sensing outcome (the result of the hypotheses test): $\Theta_a = 1$ (idle) and $\Theta_a = 0$ (busy).

If the sensor mistakes $\mathcal{H}_0$ for $\mathcal{H}_1$, a false alarm occurs, and a spectrum opportunity is overlooked by the sensor. On the other hand, when the sensor mistakes $\mathcal{H}_1$ for $\mathcal{H}_0$, we have a miss detection. Let $\epsilon \stackrel{\Delta}{=} \Pr\{\Theta_a = 0 \,|\, S_a = 1\}$ and $\delta \stackrel{\Delta}{=} \Pr\{\Theta_a = 1 \,|\, S_a = 0\}$ denote, respectively, the probabilities of false alarm and miss detection. The performance of a sensor is specified by the receiver operating characteristic (ROC) curve which gives the probability of detection $1 - \delta$ as a function of ϵ (see Fig. 10.5). We point out that analyzing the ROC curve of the spectrum sensor in a wireless network environment can be complex. We assume here that the ROC curve of the spectrum sensor has already been obtained, and we focus on the tradeoff between false alarm and miss detection. Specifically, we seek to answer the following question: which point δ on the given ROC curve should the spectrum sensor operate at?

If the secondary user completely trusts the sensing outcome in decision-making, false alarms result in wasted spectrum opportunities whereas miss detections lead to collisions with primary users. To optimize the performance of the secondary user while limiting its interference to the primary network, we should carefully choose the sensor operating point. Meanwhile, the spectrum access decisions should be made by taking into account the sensor operating characteristics. A joint design of the

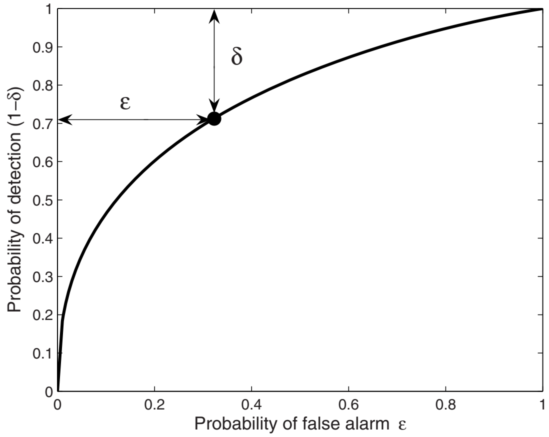

Fig. 10.5. The ROC curve of a spectrum sensor.

spectrum sensor at the physical layer and the access policy at the MAC layer is thus necessary to achieve optimality.

10.4.1.2 Sensing and Access Policies

The sensing policy specifies, in each slot, which channel to sense, and the access policy determines whether to transmit based on the sensing outcome. At the beginning of a slot, a secondary user with data to transmit chooses a channel $a \in \{1, \ldots, N\}$ to sense. Based on the sensing outcome Θ_a, the secondary user decides whether to transmit over the sensed channel: $\Phi_a \in \{0\text{ (no access)}, 1\text{ (access)}\}$. At the end of the slot, the receiver acknowledges a successful data transmission: $K_a \in \{0\text{ (unsuccessful)}, 1\text{ (successful)}\}$. Note that an acknowledgement $K_a = 1$ is obtained if and only if the secondary user chooses to access $\Phi_a = 1$ and the channel is idle $S_a = 1$, i.e.,

$$K_{\mathrm{a}} = 1_{[S_a=1,\Phi_a=1]}. \tag{10.1}$$

A reward $R_{K_a}^{(a,\Phi_a)}$ is accrued depending on K_{a}. Assuming that the number of information bits that can be transmitted is proportional to the channel bandwidth, we define the reward $R_{K_a}^{(a,\Phi_a)}$ obtained by choosing sensing and access action (a, Φ_a) as

$$R_{K_a}^{(a,\Phi_a)} = K_a B_a. \tag{10.2}$$

Due to partial spectrum monitoring and sensing errors, the secondary user and the receiver cannot directly observe the current state of the spectrum occupancy. We thus have a POMDP.

It has been shown in [32] that the knowledge of the current spectrum occupancy state based on all past decisions (i.e., sensing and access actions) and observations (i.e., acknowledgements) can be summarized by a belief state

$\boldsymbol{\lambda}(t) \stackrel{\Delta}{=} \{\lambda_s(t)\}_{s\in\{0,1\}^N}$, where $\sum_s \lambda_s(t) = 1$. Each element $\lambda_s(t)$ of the belief state $\boldsymbol{\lambda}(t)$ is the conditional probability (given the decision and observation history) that the current spectrum occupancy state is given by $s \in \{0,1\}^N$ prior to the state transition in slot t. Hence, a sensing policy π_{s} is given by a sequence of functions: $\pi_{\mathrm{s}} = [\mu_1, \ldots, \mu_{\mathrm{T}}]$ where $\mu_t : [0,1]^{2^N} \to \{1,\ldots,N\}$ maps the belief state $\boldsymbol{\lambda}(t) \in [0,1]^{2^N}$ at the beginning of slot t to a channel $a \in \{1,\ldots,N\}$ to be sensed. An access policy π_c is given by a sequence of functions: $\pi_c = [\nu_1, \ldots, \nu_{\mathrm{T}}]$ where $\nu_t : [0,1]^{2^N} \times \{0,1\} \to \{0,1\}$ maps the belief state $\boldsymbol{\lambda}(t) \in [0,1]^{2^N}$ and the sensing outcome $\Theta_a \in \{0,1\}$ of the chosen channel a to an access action $\Phi_a \in \{0,1\}$.

10.4.1.3 Design Objective

We want to determine the optimal sensor operating point δ and the optimal sensing and access policies $\{\pi_{\mathrm{s}}, \pi_{\mathrm{c}}\}$. The objective is to maximize the total expected reward (equivalently the throughput of the secondary user) in T slots under the collision constraint:

$$\begin{aligned}
&\{\delta^*, \pi_{\mathrm{s}}^*, \pi_{\mathrm{c}}^*\} = \arg \max_{\delta,\pi_{\mathrm{s}},\pi_{\mathrm{c}}} \mathbb{E}_{\{\delta,\pi_{\mathrm{s}},\pi_{\mathrm{c}}\}} \left[\sum_{t=1}^{T} R_{K_a}^{(a,\Phi_a)}(t) \,\middle|\, \boldsymbol{\lambda}(1) \right] \\
&\text{s.t. } P_a(t) = \Pr\{\Phi_a(t) = 1 \,|\, S_a(t) = 0, \boldsymbol{\lambda}(t)\} \leq \zeta \text{ holds} \\
&\quad \text{for any } a \text{ and } t \text{ such that } \Pr\{S_a(t) = 0 \,|\, \boldsymbol{\lambda}(t)\} > 0
\end{aligned} \tag{10.3}$$

where $\mathbb{E}_{\{\delta,\pi_{\mathrm{s}},\pi_{\mathrm{c}}\}}$ is the expectation given that sensing and access policies $\{\pi_{\mathrm{s}}, \pi_{\mathrm{c}}\}$ are employed and sensor operates at point δ, $\boldsymbol{\lambda}(1)$ is the initial belief state which is usually given by the stationary distribution of the spectrum occupancy states. Note that when $\Pr\{S_a(t) = 0 \,|\, \boldsymbol{\lambda}(t)\} = 0$, i.e., channel a is available with probability 1 in slot t, the constraint in (10.3) becomes irrelevant and the secondary user's access decision is simply $\Phi_a(t) = 1$. In the rest of this section, we consider the non-trivial case where $\Pr\{S_a(t) = 0 \,|\, \boldsymbol{\lambda}(t)\} > 0$ in any channel a and slot t.

10.4.2 Separation Principle for Optimal Joint Design

The design objective given in (10.3) is a constrained POMDP, which usually requires randomized policies to achieve optimality. In this case, a sensing policy determines the mapping from the current belief state to the probability of choosing each channel and an access policy the mapping from the current belief state to the transmission probabilities under different sensing outcomes. Since there exist uncountably many probability distributions, randomized policies are computationally prohibitive. In this section, we establish a separation principle for the optimal joint design. This separation principle reveals the existence of deterministic optimal sensing and access policies, leading to significant complexity reduction. It also enables us to obtain, in closed-form, the optimal sensor operating point and the optimal access policy.

10.4.2.1 The Impact of Sensor Operating Point on Access Policy

Let $f_a^\theta(\boldsymbol{\lambda}(t), t)$ be the probability of transmitting over chosen channel a given sensing outcome $\Theta_a = \theta$ and belief state $\boldsymbol{\lambda}(t)$ at the beginning of slot t. In Theorem 10.1, we provide closed-form optimal transmission probabilities $(f_a^1(\boldsymbol{\lambda}(t), t), f_a^0(\boldsymbol{\lambda}(t), t))$ for different sensor operating points δ.

Theorem 10.1. The optimal access policy is time-invariant and belief-independent. Specifically, the optimal transmission probabilities are solely determined by the sensor operating point δ and the maximum allowed probability of collision ζ, i.e., for any chosen channel a, belief state $\boldsymbol{\lambda}(t)$, and slot t, we have

$$(f_a^1(\boldsymbol{\lambda}(t), t), f_a^0(\boldsymbol{\lambda}(t), t)) = \begin{cases} (1, \frac{\zeta-\delta}{1-\delta}), & \delta < \zeta \\ (1, 0), & \delta = \zeta \\ (\frac{\zeta}{\delta}, 0), & \delta > \zeta. \end{cases} \qquad (10.4)$$

Proof. See [33] for details.

Theorem 10.1 enables us to study the impact of sensor operating characteristics on the optimal access policy. As illustrated in Fig. 10.6, the ROC curve can be partitioned into two regions: the “conservative” region ($\delta > \zeta$) and the “aggressive”

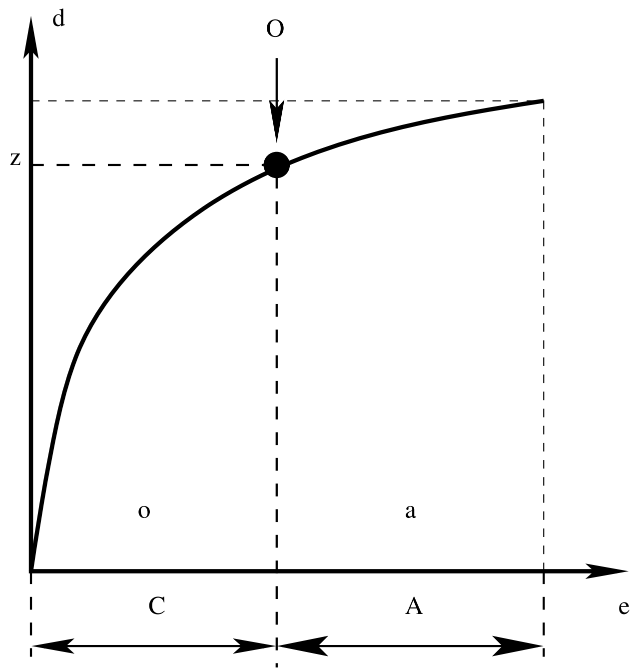

Fig. 10.6. The partition of an ROC curve.

region ($\delta < \zeta$). When $\delta > \zeta$, the spectrum sensor is more likely to misidentify an opportunity (i.e., a busy channel is sensed to be idle). Hence, the access policy should be conservative to ensure that the probability of collision is bounded below ζ. Specifically, even when the sensing outcome $\Theta_a = 1$ indicates that the channel is available, the user should only transmit with probability $\frac{\zeta}{\delta} < 1$. When the channel is sensed to be busy: $\Theta_a = 0$, the user should trust the sensing outcome and refrain from transmission. On the other hand, when $\delta < \zeta$, the spectrum sensor is more likely to overlook an opportunity (i.e., an idle channel is sensed to be busy). Hence, the user should adopt an aggressive access policy: always transmit when the channel is sensed to be available and transmit with probability $\frac{\zeta-\delta}{1-\delta} > 0$ even when the channel is sensed to be busy. When $\delta = \zeta$, the optimal access policy is deterministic: always trust the sensing outcome.

10.4.2.2 The Separation Principle

Given belief state $\boldsymbol{\lambda}(t)$ at the beginning of slot t, we rewrite the design constraint in (10.3) as

$$\begin{aligned}P_a(t) &= \sum_{\theta=0}^{1} \Pr\{\Phi_a = 1 \mid \Theta_a = \theta\} \Pr\{\Theta_a = \theta \mid S_a(t) = 0\} \\ &= \delta f_a^1(\boldsymbol{\lambda}(t), t) + (1-\delta) f_a^0(\boldsymbol{\lambda}(t), t). \qquad (10.5)\end{aligned}$$

Careful inspection of (10.4) and (10.5) reveals that the constraint given in (10.3) is satisfied regardless of the chosen channel. We thus have a separation principle (Theorem 10.2) for the optimal joint OSA design, which decouples the design of spectrum sensor and access policy from that of sensing policy. Following this separation principle, we obtain closed-form optimal sensor operating point δ^* and access policy π_c^* in Theorem 10.3.

Theorem 10.2. Separation Principle The joint design of OSA formulated in (10.3) can be obtained in two steps without losing optimality. First, choose sensor operating point δ and access policy π_c according to (10.4) to maximize the expected immediate reward. Second, choose sensing policy π_s to maximize the expected total reward.

Proof. See [33] for details.

Theorem 10.3. The optimal sensor operating point is $\delta^* = \zeta$. The optimal access policy π_c^* is given by $\Phi_a^* = \Theta_a$.

Proof. See [33] for details.

Theorem 10.3 reveals the existence of deterministic optimal access policy for the constrained POMDP given in (10.3). Specifically, the optimal access policy π_c^* is to simply trust the sensing outcome: $\Phi_a^* = \Theta_a$, i.e., access if and only if the channel is detected to be available.

10.4.2.3 The Optimal Sensing Policy

In Theorem 10.3, we have obtained the optimal sensor operating point δ^* and the optimal access policy π_{c}^*. Since δ^* and π_{c}^* have been chosen to ensure the constraint regardless of the chosen channel, we are free to search for the optimal sensing policy π_{s}^* over the whole design space. The design of the sensing policy thus becomes an unconstrained POMDP, where optimality can be achieved by deterministic policies.

Let $V_{\mathrm{t}}(\boldsymbol{\lambda}(t))$ denote the maximum total expected reward obtained from slot t, $1 \leq t \leq T$, given the belief state $\boldsymbol{\lambda}(t)$ at the beginning of slot t. Given sensor operating point δ^* and access policy π_{c}^*, we obtain $V_{\mathrm{t}}(\boldsymbol{\lambda}(t))$ recursively by

$$
\begin{aligned}
V_{\mathrm{t}}(\boldsymbol{\lambda}(t)) &= \max_a \sum_{s\in\{0,1\}^N} \sum_{s'\in\{0,1\}^N} \lambda_{s'}(t) P_{s',s} \sum_{k_a=0}^{1} Q_s(k_a) \\
&\quad \times [k_a B_a + V_{t+1}(\mathcal{T}(\boldsymbol{\lambda}(t)\,|\,a,k_a))], \quad 1 \leq t < T \\
V_{\mathrm{T}}(\boldsymbol{\lambda}(T)) &= \max_a \sum_{s\in\{0,1\}^N} \sum_{s'\in\{0,1\}^N} \lambda_{s'}(t) P_{s',s} Q_s(1) B_a
\end{aligned}
\tag{10.6}
$$

where $Q_s(0) = 1 - Q_s(1)$, $Q_s(1) \stackrel{\Delta}{=} \Pr\{K_a = 1\,|\,S(t) = s\} = 1_{[s_a=1]}(1-\epsilon^*)$ is the probability of successful transmission when the current spectrum occupancy $S(t)$ is in state $s = [s_1, \ldots, s_N]$. Note that $1_{[s_{\mathrm{a}}=1]}$ indicates whether channel a is idle given $S(t) = s$ and ϵ^* is the probability of false alarm that can be achieved when the spectrum sensor operates at δ^*. The updated belief state $\boldsymbol{\lambda}(t+1) = \mathcal{T}(\boldsymbol{\lambda}(t)\,|\,a,k_a)$ can be obtained via Bayes rule as

$$
\lambda_s(t+1) = \frac{\sum_{s'\in\{0,1\}^N} \lambda_{s'}(t) P_{s',s} Q_s(k_a)}{\sum_{s\in\{0,1\}^N} \sum_{s'\in\{0,1\}^N} \lambda_{s'}(t) P_{s',s} Q_s(k_{\mathrm{a}})}. \tag{10.7}
$$

The optimal sensing policy π_{s}^* can be obtained by solving the optimality equation given in (10.6). It is shown in [32] that $V_{\mathrm{t}}(\boldsymbol{\lambda}(t))$ is piecewise linear and convex, leading to a linear programming procedure for calculating π_{s}^*.

Suboptimal sensing policies with reduced complexity are developed in [16, 17, 34].

10.4.3 Simulation Examples

In this section, we provide simulation examples to study the cognitive nature of the MAC protocols developed within the POMDP framework and the impacts of sensor operating point δ and mismatched Markov model on the performance of the optimal OSA.

10.4.3.1 Simulation Setup

We consider $N = 3$ independently evolving channels with the same bandwidth $B_n = 1$. As illustrated in Fig. 10.7, the state transition of spectrum occupancy can be

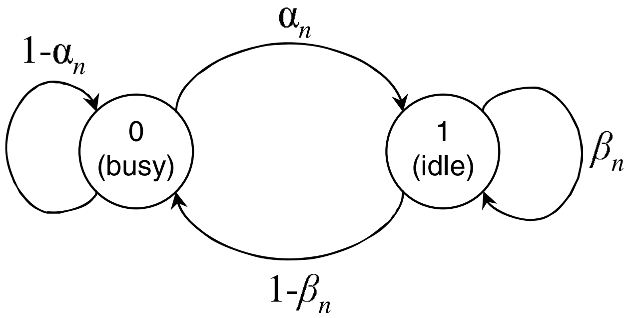

Fig. 10.7. The Markov model of independently evolving channels.

characterized by $\boldsymbol{\alpha} \stackrel{\Delta}{=} [\alpha_1, \alpha_2, \alpha_3]$ and $\boldsymbol{\beta} \stackrel{\Delta}{=} [\beta_1, \beta_2, \beta_3]$, where α_n denotes the probability that channel n transits from state 0 (busy) to state 1 (idle) and β_n denotes the probability that it stays in state 1. We assume that the spectrum occupancy dynamics remain unchanged over $T = 10$ slots. The throughput of the secondary user is measured by the expected total reward per slot, i.e., $V_1(\boldsymbol{\lambda}(1))/T$, where $\boldsymbol{\lambda}(1)$ is given by the stationary distribution of the underlying Markov process.

At the beginning of each slot, the spectrum sensor takes M measurements $\{Y_i\}_{i=1}^M$ of the chosen channel. We assume that both the channel noise and the signal of primary users can be modeled as white Gaussian processes $\mathcal{N}$. Then, the spectrum sensor performs the following hypotheses test:

$$\begin{cases} \mathcal{H}_0 \text{ (idle channel)}: & Y_i \sim \mathcal{N}(0, \sigma_0^2),\ i = 1, \cdots, M \\ \mathcal{H}_1 \text{ (busy channel)}: & Y_i \sim \mathcal{N}(0, \sigma_1^2),\ i = 1, \cdots, M \end{cases}$$

where σ_0^2 is the noise power and σ_1^2 is the primary signal power. The energy detector is optimal under Neyman–Pearson (NP) criterion [35, sect. 2.6.2]:

$$\sum_{i=1}^{M} Y_i^2 \gtrless_{\mathcal{H}_0}^{\mathcal{H}_1} \eta \tag{10.8}$$

where the threshold η determines the false alarm and miss detection rates of the detector. The ROC curve of the energy detector is given by [35, Sect. 2.6.2]

$$1 - \delta = 1 - \gamma\left(\frac{M}{2}, \eta\frac{\sigma_0^2}{\sigma_1^2}\right), \quad \epsilon = 1 - \gamma\left(\frac{M}{2}, \eta\right) \tag{10.9}$$

where $(\sigma_1^2 - \sigma_0^2)/\sigma_0^2$ is the SNR and $\gamma(n, a) = \frac{1}{\Gamma(n)} \int_0^a t^{n-1} \mathrm{e}^{-t}\, \mathrm{d}t$ is the incomplete gamma function. In all the figures, we assume $M = 10$ and SNR = 5 dB.

10.4.3.2 The Cognitive Nature of POMDP Modeling

As discussed in Sect. 10.2.1, a fundamental tradeoff in the design of sensing policies is between obtaining immediate spectrum access and gaining spectrum statistical information for future use. To illustrate this, we consider a simple static sensing strategy that chooses the channel most likely to be available (weighted by its bandwidth) based on the stationary distribution of the underlying Markov process. In this case, the secondary user simply waits on a particular channel predetermined by the spectrum occupancy statistics and the channel bandwidths. Such an approach ignores information, about the underlying state of the Markov process, that can be obtained from the sensing outcomes. Missing in this approach is that every sensing outcome provides information on the state of the underlying Markov process. Channel selection should be based on the *a posterior* distribution of channel availability that exploits the whole history of sensing outcomes, i.e., the belief state. As demonstrated in this section, the optimal sensing strategy is one of sequential decision making that achieves the best trade-off between gaining immediate access in the current slot and gaining system state information for future use. We illustrate in Fig. 10.8 the potential gain of optimally using the observation history assuming perfect sensing. Plotted in Fig. 10.8 is the throughput of the secondary user as a function of time. We see from this figure that the performance of the optimal approach improves over time, which results from the increasingly accurate information on the system state obtained by accumulating observations. Approximately 40% improvement is achieved over the static approach.

10.4.3.3 Impact of Sensor Operating Point on MAC Performance

Figure 10.9 illustrates the impact of sensor operating point δ on the throughput and the optimal access policy of the secondary user. The upper figure plots the maximum throughput of the secondary user for each given sensor operating point δ. The optimal access policy is specified by the transmission probabilities (f_a^0, f_a^1), which are shown in the middle and the lower figures, respectively. We can see that the maximum throughput is achieved at $\delta^* = \zeta = 0.05$ and the transmission probabilities change with δ as given by Theorem 10.1. Interestingly, the throughput curve

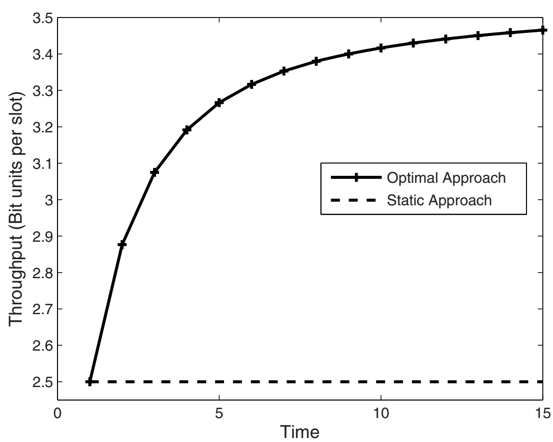

Fig. 10.8. The cognitive nature of POMDP modeling.

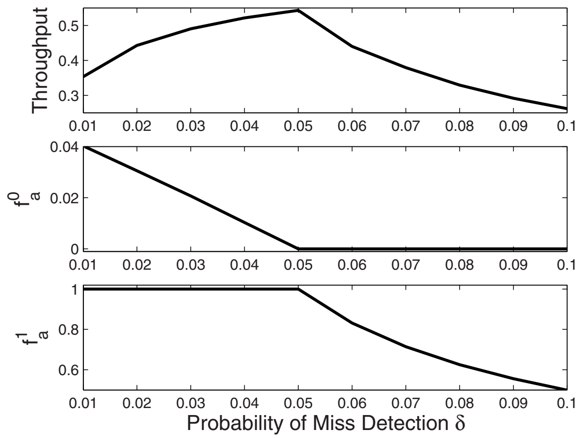

Fig. 10.9. The impact of sensor operating point on the throughput in normalized units. $\boldsymbol{\alpha} = [0.2, 0.4, 0.6]$, $\boldsymbol{\beta} = [0.8, 0.6, 0.4]$, $\zeta = 0.05$.

is concave with respect to δ in the "aggressive" region ($\delta < \zeta$) and convex in the "conservative" region ($\delta > \zeta$). The performance thus degrades at a faster rate when the sensor operating point drifts toward the "conservative" region. This suggests that miss detections (which lead to collisions) are more harmful to the performance of OSA than false alarms (which represent missed opportunities).

10.4.3.4 OSA with Unknown or Mismatched Model

If the transition probabilities of the Markov model are unknown, formulations and algorithms for POMDP with an unknown model exist in the literature [36] and can be applied to the problem of OSA design. Here we study the impact of mismatched Markov model on the performance of the optimal OSA.

We assume that the spectrum occupancy evolves according to the transition probabilities given by $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ while the secondary user employs the optimal OSA policy based on inaccurate transition probabilities $\boldsymbol{\alpha}'$ and $\boldsymbol{\beta}'$. In the upper plot of Fig. 10.10, we plot the relative throughput loss of the secondary user as a function of the relative error ψ in transition probabilities which is given by $\psi = \frac{\alpha'_n - \alpha_n}{\alpha_n} \times 100\% = \frac{\beta'_n - \beta_n}{\beta_n} \times 100\%$. Clearly, the maximum throughput is achieved when the relative error is zero (i.e., the secondary user has accurate information on transition probabilities). Inaccurate transition probabilities can cause performance loss. We find that the relative performance loss is below 4% even when the absolute relative error is up to 20%. In the lower figure, we examine the probability of collision perceived by

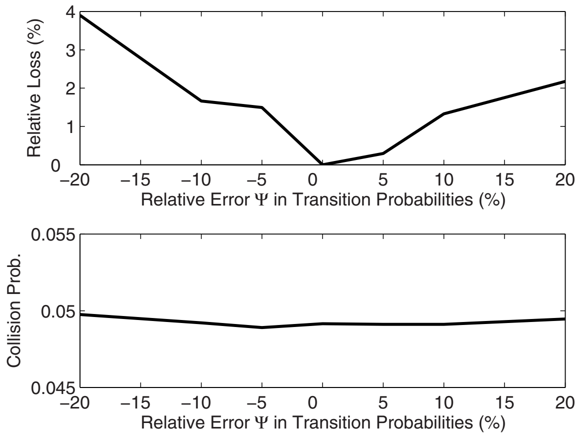

Fig. 10.10. The impact of inaccurate transition probabilities on the throughput of the secondary user. $\boldsymbol{\alpha} = [0.2, 0.4, 0.6]$, $\boldsymbol{\beta} = [0.8, 0.6, 0.4]$, $\zeta = 0.05$.

the primary network. We find that the probability of collision is not affected by mismatched transition probabilities. The reason behind this observation is the separation principle: the optimal sensor operating point and the optimal access policy, which determine the probability of collision, are independent of the spectrum occupancy dynamics.

10.5 The Impact of Fading on Energy-Constrained Cognitive MAC

In this section, we study the impact of channel fading conditions at the physical layer on the design of cognitive MAC protocols under energy constraints. We show that the problem can again be formulated within the framework of POMDP. Optimal and suboptimal sensing and access policies with reduced complexity are obtained for energy-constrained OSA networks in fading environments. To isolate the effect of energy constraint on the design of cognitive MAC, we assume that sensing errors are negligible.

10.5.1 Energy and Fading Model

The network model is the same as that given in Sect. 10.3. We present below the energy and channel fading model.

We assume that channels between the secondary user and its destination follow a block fading model. That is, 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 $E_s(n)$ and $E_{tx}(n)$ denote, respectively, the energy consumed in sensing and accessing channel n in a slot. For simplicity, we assume that sensing energy consumption $E_s(n)$ is identical for all channels: $E_s(n) = e_s$ for every n. Note that the transmission energy consumption $E_{tx}(n)$ 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 kth power level in a slot. The transmission energy consumption $E_{tx}(n)$ thus has realizations restricted to a finite set $\mathcal{E}_{tx}$ given by

$$E_{tx}(n) \in \mathcal{E}_{tx} \stackrel{\Delta}{=} \{\varepsilon_k\}_{k=0}^{L} \tag{10.10}$$

where $0 < \varepsilon_1 < \cdots < \varepsilon_L < \infty$ and $\varepsilon_0 = 0$ indicates that the secondary user does not transmit. We also consider the energy e_p consumed in the sleeping mode of the secondary user.

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

$$E \in \mathcal{E}_r \stackrel{\Delta}{=} \{e : e = \mathcal{E}_0 - \sum_{k=0}^{L} c_k(e_s + \varepsilon_k) - ce_p, e \geq 0, c, c_k \geq 0, c, c_k \in \mathbb{Z}\} \cup \{0\} \tag{10.11}$$

where c_k is the number of slots when the secondary user chooses to sense a channel and then transmit over it at the kth power level and c is the number of slots when the secondary user turns to sleeping mode. Section 10.6 discuses how the secondary user can obtain knowledge of the required power level.

10.5.2 Optimal Energy-Constrained OSA

The energy-constrained OSA 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 sensing and access policies.

10.5.2.1 An Unconstrained POMDP Formulation

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

$$(S, E) \in \mathcal{S} \stackrel{\Delta}{=} \{(s, e) : s \in \{0, 1\}^N, e \in \mathcal{E}_r\}. \tag{10.12}$$

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 go 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 $\bar{\Phi}_a \stackrel{\Delta}{=} [\Phi_a(1), \ldots, \Phi_a(L)]$:

$$(a, \bar{\Phi}_a) \in \mathcal{A} \stackrel{\Delta}{=} \{(0, [0, \ldots, 0])\} \cup \{(a, \boldsymbol{\phi}) : a \in \{1, \ldots, N\}$$
$$\boldsymbol{\phi} \stackrel{\Delta}{=} [\phi(1), \ldots, \phi(L)] \in \{0, 1\}^L\}. \tag{10.13}$$

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

Network State Transition. Recall that the network state consists of two parts: the spectrum occupancy S and the residual energy E of the secondary user. At the beginning of each slot, the spectrum occupancy S transits independently of the residual energy E according to transition probabilities $\{P_{s,s'}\}$. As stated in Section 10.3, we assume that the spectrum occupancy dynamics $\{P_{s,s'}\}$ are known and remain unchanged during the battery lifetime of the secondary user.

If the secondary user decides to sense channel $a \in \{1, \ldots, N\}$ in this slot, then it 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 $E' = \mathcal{T}_E(E \mid a, \Theta_a, \Phi_a(\Theta_a))$:

$$\mathcal{T}_E(E \mid a, \Theta_a, \Phi_a(\Theta_a)) = \begin{cases} E - e_p, & a = 0 \\ \max\{E - e_{\text{s}} - \Phi_a(\Theta_a)\varepsilon_{\Theta_a}, 0\}, & a \neq 0 \end{cases} \tag{10.14}$$

where e_{p} is the 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 $q_s^{(a)}(k)$ be the probability that the secondary user observes $\Theta_a = k$ in the chosen channel a given current spectrum occupancy state $S = s$. Under perfect spectrum sensing, we have that

$$q_s^{(a)}(k) = \Pr\{\Theta_a = k \,|\, S = s\} = \begin{cases} 1_{[k \neq 0]} p_a(k), & \text{if } a \neq 0, s_a = 1 \\ 1_{[k=0]}, & \text{if } a \neq 0, s_a = 0 \end{cases} \tag{10.15}$$

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 (10.15), we can see that $\sum_{k=0}^{L} q_s^{(a)}(k) = 1$ for any spectrum occupancy state $s \in \mathcal{S}$ and any chosen channel $a \in \{1, \ldots, N\}$.

Note that if the secondary user turns to sleep, then it will not have any sensing outcome. We can define $\{q_s^{(0)}(k)\}$ as arbitrary values that satisfy $\sum_{k=0}^{L} q_s^{(0)}(k) = 1$. For simplicity, we define $q_s^{(0)}(k) = 1_{[k=0]}$.

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

$$R_{E,\Theta_a}^{(a,\Phi_a(\Theta_a))} \triangleq \begin{cases} 0, & a = 0 \\ \Phi_a(\Theta_a) B_a 1_{[E - e_{\mathrm{s}} - \varepsilon_{\Theta_a} \geq 0]}, & a \neq 0. \end{cases} \tag{10.16}$$

That is, a reward is obtained if and only if the secondary chooses to sense and access (i.e., $a \neq 0$, $\Phi_a(\Theta_a) = 1$) an idle channel (i.e., $\Theta_a \neq 0$) and its residual energy is enough to cope with the channel fade in the selected channel (i.e., $E - e_{\mathrm{s}} - \varepsilon_{\Theta_a} \geq 0$). Note that no reward will be accumulated once the battery energy level drops below $e_{\mathrm{s}} + \varepsilon_1$, 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 S. As stated in Section 10.4, its knowledge of S based on all past decisions and observations can be summarized by a belief state $\boldsymbol{\lambda} = \{\lambda_s\}_{s \in \{0,1\}^N}$ [32], where λ_s is the conditional probability (given the decision and observation history) that the spectrum occupancy is in state s at the beginning of this slot prior to the state transition.

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 λ'_s denotes the probability that the current spectrum occupancy state is $S = s$ given belief state $\boldsymbol{\lambda}$ at the beginning of this slot and the observation $\Theta_a = k$ of the chosen channel a in the current slot. Applying Bayes rule, we obtain λ'_s as

$$\begin{aligned} \lambda'_s &= \Pr\{S = s \,|\, \boldsymbol{\lambda}, a, k\} \\ &= \begin{cases} \sum_{s'} \lambda_{s'} P_{s',s}, & a = 0 \\ \dfrac{\sum_{s'} \lambda_{s'} P_{s',s} 1_{[s_a = 1_{[k \neq 0]}]}}{\sum_{s} \sum_{s'} \lambda_{s'} P_{s',s} 1_{[s_a = 1_{[k \neq 0]}]}}, & a \neq 0 \end{cases} \end{aligned} \tag{10.17}$$

where the summations are taken over the space $\{0,1\}^N$ of spectrum occupancy state 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 $\{P_{s,s'}\}$.

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:

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

where $\{a, \bar{\Phi}_a\} = \mu_t(\boldsymbol{\lambda}, E)$ 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 $\bar{\Phi}_a = [\Phi_a(1), \ldots, \Phi_a(L)] \in \{0,1\}^L$.

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

$$\pi^* = \arg\max_{\pi} \mathbb{E}_\pi \left[\sum_{t=1}^{\infty} R_{E,\Theta_a}^{(a,\Phi_a(\Theta_a))}(t) \,\middle|\, \boldsymbol{\lambda}_0 \right] \tag{10.18}$$

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

10.5.2.2 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 (10.16) 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:

$$V(\boldsymbol{\lambda}, E) = \max_{(a,\boldsymbol{\phi}) \in \mathcal{A}} \sum_{k=0}^{L} u_k^{(a)} [R_{E,k}^{(a,\phi(k))} + V(\mathcal{T}_\lambda(\boldsymbol{\lambda} \,|\, a, k), \mathcal{T}_E(E \,|\, a, k, \phi(k)))] \tag{10.19}$$

where $\mathcal{T}_\lambda(\boldsymbol{\lambda} \,|\, a, k)$ is the updated belief state given in (10.17), $\mathcal{T}_E(E \,|\, a, k, \phi(k))$ is the reduced battery energy given in (10.14), and $u_k^{(a)} \stackrel{\Delta}{=} \Pr\{\Theta_a = k \,|\, \boldsymbol{\lambda}\}$ 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:

$$u_k^{(a)} = \sum_{s' \in \{0,1\}^N} \lambda_{s'} \sum_{s \in \{0,1\}^N} P_{s',s} \; q_s^{(a)}(k). \tag{10.20}$$

In principle, by solving (10.19), we can obtain the optimal sensing and access actions (a^*, Φ_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 $V_{\text{opt}} = V(\boldsymbol{\lambda}_0, \mathcal{E}_0)$, where $\boldsymbol{\lambda}_0$ is the initial belief state.

10.5.3 Optimal Policy with Reduced Complexity

Although the value function given in (10.19) 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.

10.5.3.1 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.

We notice from (10.19) 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 [34], we can calculate the number of all possible information states $(\boldsymbol{\lambda}, E)$ during the secondary user's battery lifetime. Specifically, noting from (10.17) 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 on the order of $\mathrm{O}((N(L+1))^{T-1})$, which is exponential in the battery lifetime T and polynomial in the number N of channels. Moreover, from (10.19) and (10.20), we can see that it requires $\mathrm{O}(3|\mathcal{A}|2^N 2^N(L+1))$ 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 on the order of $\mathrm{O}(3|\mathcal{A}|2^N 2^N(L+1)(N(L+1))^{T-1})$. We can see that the complexity is mainly caused by the following three factors: (1) the number $\mathrm{O}((N(L+1))^{T-1})$ 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 10.5.4. In this section, we focus on the other two factors.

10.5.3.2 Reduction of Action Space Size

Careful inspection of (10.14), (10.16) and (10.19) 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 (10.19) only depends on the k-th entry $\phi(k)$ of the access decision $\bar{\boldsymbol{\phi}}$ and is independent of $\phi(i)$ ($i \neq k$). We can thus simplify (10.19) as

$$V(\boldsymbol{\lambda}, E) = \max_{a \in \{0,1,\ldots,N\}} \{\sum_{k=0}^{L} u_k^{(a)} \max_{\phi(k) \in \{0,1\}} [R_{E,k}^{(a,\phi(k))} + V(\mathcal{T}_\lambda(\boldsymbol{\lambda}\,|\,a,k), \mathcal{T}_E(E\,|\,a,k,\phi(k)))]\}. \tag{10.21}$$

The maximization in (10.21) is taken over a space with size $\mathrm{O}(2NL)$, increasing linearly with the number L of power levels, while that in (10.19) is taken over the action space $\mathcal{A}$ whose size $\mathrm{O}(N2^L)$ increases exponentially with L.

Proposition 10.1 states that the optimal access decision Φ_a^* is a threshold policy.

Proposition 10.1. Given the belief state $\boldsymbol{\lambda}$ and the residual energy level E of the secondary user at the beginning of a slot, there exists a threshold k_a^* associated with sensing action $a \in \{1, \ldots, N\}$ such that the optimal access decision $\Phi_a^* = [\phi_a^*(1), \ldots, \phi_a^*(L)]$ is given by

$$\phi_a^*(k) = \begin{cases} 1, & \text{if } k \leq k_a^* \\ 0, & \text{if } k > k_a^*. \end{cases} \tag{10.22}$$

Proof. See [18].

Proposition 10.1 can help us avoid the search for optimal access decisions in some scenarios, resulting in further complexity reduction. Specifically, for each sensing action $a \neq 0$, we can calculate the optimal access decisions $\phi_a^*(k)$ in a decreasing order of sensing outcome k. Once we have $\phi_a^*(k^*) = 1$ for a certain value of k^*, we can determine the optimal access decisions for all remaining sensing outcomes $k < k^*$ without further computation.

10.5.3.3 Reduction of Belief State Dimension

Assume that the spectrum occupancy evolves independently across channels. It has been shown in [16] that $\boldsymbol{\omega} \stackrel{\Delta}{=} [\omega_1, \ldots, \omega_N]$, 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 $\mathrm{O}(N)$ with the number N of channels while that of $\boldsymbol{\lambda}$ increases exponentially $\mathrm{O}(2^N)$.

Using the belief state $\boldsymbol{\omega}$, we can simplify the value function given in (10.21). 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, (10.21) reduces to

$$\begin{aligned} \hat{V}(\boldsymbol{\omega}, E) = & \max_{a \in \{0,1,\ldots,N\}} \{(1 - \omega_a')\hat{V}(\hat{\mathcal{T}}_\lambda(\boldsymbol{\omega} \,|\, a, 0), \mathcal{T}_E(E \,|\, a, 0, 0)) \\ & + \omega_a' \sum_{k=1}^{L} p_a(k) \max_{\phi(k) \in \{0,1\}} [R_{E,k}^{(a,\phi(k))} + \hat{V}(\hat{\mathcal{T}}_\lambda(\boldsymbol{\omega} \,|\, a, k), \mathcal{T}_E(E \,|\, a, k, \phi(k)))]\} \end{aligned} \tag{10.23}$$

where $\omega_0' \stackrel{\Delta}{=} 0$, $\omega_a' = \omega_a \beta_a + (1 - \omega_a)\alpha_a$ ($a \in \{1, \ldots, L\}$) is the probability that channel a is available in the current slot given $\boldsymbol{\omega}$, $\mathcal{T}_E(E \,|\, a, k, \phi_a(\mathrm{k}))$ is the reduced battery energy given in (10.14), and the updated belief state $\hat{\boldsymbol{\omega}} \stackrel{\Delta}{=} [\omega_1, \ldots, \omega_N] = \hat{\mathcal{T}}_\lambda(\boldsymbol{\omega} \,|\, a, k)$ is given by

$$\hat{\omega}_n = \begin{cases} 0, & \text{if } a \neq 0, n = a, k = 0 \\ 1, & \text{if } a \neq 0, n = a, k \neq 0 \\ \omega_n', & \text{otherwise.} \end{cases} \tag{10.24}$$

10.5.4 Suboptimal Cognitive MAC with Reduced Complexity

We notice from (10.19) 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 very favorably traded off with its performance.

10.5.4.1 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

$$\begin{aligned} &Y_0^{(a)}(\boldsymbol{\lambda}, E) = 0 \\ &Y_w^{(a)}(\boldsymbol{\lambda}, E) = \sum_{k=0}^{L} u_k^{(a)} \max_{\phi(k)\in\{0,1\}} [R_{E,k}^{(a,\phi(k))} \\ &+ \max_{b\in\{0,1,\ldots,N\}} Y_{w-1}^{(b)}(\mathcal{T}_\lambda(\boldsymbol{\lambda} \,|\, a, k), \mathcal{T}_{\mathrm{E}}(E \,|\, a, k, \phi(k)))] \end{aligned} \tag{10.25}$$

where $u_k^{(a)}$, $\mathcal{T}_\lambda(\boldsymbol{\lambda} \,|\, a, k)$, and $\mathcal{T}_{\mathrm{E}}(E \,|\, a, k, \phi(k))$ are given in (10.20), (10.17), and (10.14), respectively. From (10.25), we can see that for any w, $Y_w^{(a)}(\boldsymbol{\lambda}, E) = 0$ if $E < e_{\mathrm{s}} + \varepsilon_1$.

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.,

$$a_w = \arg \max_{a\in\{0,1,\ldots,N\}} Y_w^{(a)}(\boldsymbol{\lambda}, E). \tag{10.26}$$

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

$$\begin{aligned} \phi_{a_w}(k) = \arg \max_{\phi\in\{0,1\}} \{&R_{E,k}^{(a_w,\phi)} \\ &+ \max_{b\in\{1,\ldots,N\}} Y_{w-1}^{(b)}(\mathcal{T}_\lambda(\boldsymbol{\lambda} \,|\, a_w, k), \mathcal{T}_E(E \,|\, a_w, k, \phi))]\}. \end{aligned} \tag{10.27}$$

Next, we consider two extreme cases of the greedy-w strategy.
Case 1: When $w = 1$, the greedy-1 approach focuses solely on maximizing the immediate reward. Specifically, the secondary user employing greedy-1 approach chooses the channel with the maximum expected immediate reward and transmits whenever the channel is sensed to be available:

$$\begin{aligned} a_1 &= \arg \max_{a \in \{1,\ldots,N\}} \sum_{k=1}^{L} u_k^{(a)} R_{E,k}^{(a,1)} \\ \phi_{a_1}(k) &= 1_{[k \neq 0]}. \end{aligned} \tag{10.28}$$

The greedy-1 approach has the lowest computational complexity.
Case 2: Consider the case when window size w exceeds the maximum battery lifetime of the secondary user. In this case, the network reaches a terminating state in less than w slots regardless of the sensing and access strategies. Since no reward is accumulated after the network reaches a terminating state, the greedy-w approach is equivalent to the optimal strategy.

10.5.4.2 Complexity Vs. Performance

We can see from (10.26) and (10.27) 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 on the order of $\mathrm{O}((N(L+1))^{w-1}T)$, which is linear in T. Clearly, the computational complexity of greedy-w approach increases with w.

Next, we compare the performance of the greedy-w approach with the optimal performance $V(\boldsymbol{\lambda}_0, \mathcal{E}_0)$. In Fig. 10.11, we plot the total expected number of information bits that can be delivered by the secondary user during its battery lifetime. We

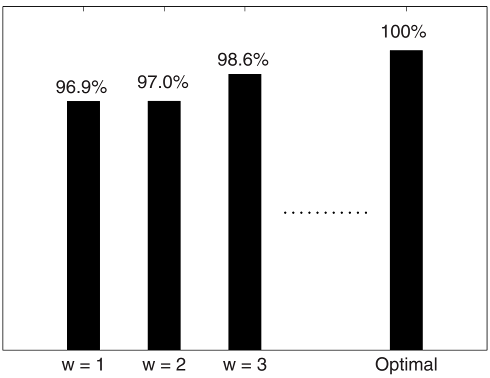

Fig. 10.11. Throughput comparison of the greedy-w and the optimal approaches.

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, usually small, the greedy-w approach can achieve a desired tradeoff between complexity and performance.

10.5.5 Numerical Examples

Equation (10.19) indicates that a sensing and access action $(a, \boldsymbol{\phi}) \in \mathcal{A}$ affects the total expected reward in three ways: (1) it yields 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 spectrum occupancy information up to this slot; (3) it causes a reduction in battery energy from E to $\mathcal{T}_E(E, a, k, \phi(k))$, decreasing the 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.

10.5.5.1 To Sense or Not to Sense?

The secondary user may choose to sense in order to gain immediate reward and spectrum 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 Table 10.5.5.1, we study the optimal sensing decision $1_{[a^* \neq 0]}$ in a particular slot under different spectrum occupancy dynamics and belief states.

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 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

Table 10.1. The impact of spectrum occupancy dynamics α and belief states $\boldsymbol{\omega}$ on the optimal sensing decision $1_{[a^* \neq 0]}$. $N = 2$, $[B_1, B_2] = [1, 1]$, $\mathcal{E}_0 = 4$, $e_s = 0.6$, $e_p = 0.1$, $L = 2$, $\mathcal{E}_{tx} = \{1, 2\}$, $p_n(1) = p_n(2) = 0.5$ for $n = 1, 2$.

α	0.05		0.1		0.4		0.8	
$\boldsymbol{\omega}$	[0.5,0.5]	[0,0]	[0.5,0.5]	[0,0]	[0.5,0.5]	[0,0]	[0.5,0.5]	[0,0]
Sense	X		X		X	X	X	X
Do not sense		X		X				

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 Table 10.5.5.1, 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 states. 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 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 $\hat{V}(\boldsymbol{\omega}_1, E - e_\mathrm{p})$. Since $\hat{V}(\boldsymbol{\omega}, E)$ increases with the battery energy E for every fixed $\boldsymbol{\omega}$, we have $\hat{V}(\boldsymbol{\omega}_1, E) \geq \hat{V}(\boldsymbol{\omega}_1, E - e_\mathrm{p})$ and hence the secondary user should choose to sense whenever it has a stationary belief state.

10.5.5.2 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 the 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 Table 10.2, we study the impact of sensing energy consumptions e_s and channel fading statistics $\{p_n(k)\}_{k=1}^{L}$ on the optimal access decision $\phi^*(k)$ under different observations k. 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. Comparing the optimal access decisions in the two cases of Table 10.2 when sens-

Table 10.2. The impact of sensing energy consumptions e_s and channel fading statistics on the optimal access decision $\phi^*(k)$ under different observations k. $N = 2$, $[B_1, B_2] = [1, 1]$, $\mathcal{E}_0 = 8$, $e_\mathrm{p} = 0.1$, $L = 3$, $\mathcal{E}_{\mathrm{tx}} = \{1, 2, 3\}$. Case 1: $p_n(1) = 0.5, p_n(2) = 0.3, p_n(3) = 0.2$ for $n = 1, 2, 3$. Case 2: $p_n(1) = 0.3, p_n(2) = 0.3, p_n(3) = 0.4$.

Sensing energy e_s		0			0.7			0.8			1.0		
Observation k		1	2	3	1	2	3	1	2	3	1	2	3
Case 1	Access	X			X	X		X	X		X	X	X
	Do not access		X	X			X			X			
Case 2	Access	X			X	X		X	X	X	X	X	X
	Do not access		X	X			X						

ing energy is $e_s = 0.8$. We find that if the probability that the channel experiences deep fading is small (case 1), 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 (case 2), the secondary user should focus on gaining immediate reward because of the long waiting time for better channel realizations.

10.6 Protocol Specifics of Decentralized Cognitive MAC

In this section, we present protocol specifics of the cognitive MAC strategies presented in Sects. 10.4 and 10.5.

10.6.1 Transceiver Synchronization

Without a dedicated communication or control channel, transceiver synchronization is a key issue in distributed cognitive MAC for OSA networks [16, 17, 37]. Specifically, a secondary user and its intended receiver need to hop to the same channel at the beginning of each slot in order to carry out the communication. The synchronization problem can be separated into two phases: the initial handshake between the transmitter and the receiver and the synchronous hopping in the spectrum after the initial establishment of communication.

There are a number of standard implementations to facilitate the initial handshake. As given in [16, 17, 37], we can borrow the idea of receiver-oriented code assignment in CDMA ad hoc networks [38]. Specifically, each secondary user is assigned a set of channels (not necessarily unique) which it monitors regularly to check whether it is an intended receiver. A user with a message for, say, user A will transmit a handshake signal over one of the channels assigned to user A. Once the initial communication is established, the transmitter and the receiver will implement the same spectrum sensing and access strategy which governs channel selection in each slot. As detailed in [17, 18], the sensing and access strategies presented in Sects. 10.4 and 10.5 ensure synchronous hopping between the transmitter and the receiver in the presence of collisions, sensing errors, and fading.

Specifically, the structure of the cognitive MAC protocols developed within the POMDP framework 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 $\boldsymbol{\lambda}$. Since the channel selection is determined by the information state $\boldsymbol{\lambda}$, the transmitter and the receiver will hop to the same channel in the next slot, i.e., transceiver synchronization is maintained.

10.6.2 Identification of Spectrum Opportunity and Fading Condition

When every secondary user is affected by the same set of primary users, the state of a channel is the same at both the transmitter and the receiver. Detection of spectrum opportunity can thus be carried out at the transmitter alone. When secondary users are affected by different sets of primary users, however, the state of spectrum occupancy is location dependent; a channel that is idle at a transmitter may not be idle at the corresponding receiver. In this case, spectrum opportunities need to be identified jointly by the transmitter and the receiver[4] [17, 34]

To achieve joint opportunity identification at both transmitter and receiver, a scheme based on RTS-CTS exchange is proposed in [17, 37]. We briefly comment on this scheme using the energy-constrained cognitive MAC given in Sect. 10.5 as an example, where we show that this scheme also facilitates the estimation of channel fading conditions.

At the beginning of a slot, the transmitter and the receiver hop to the same channel. 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 next slot.

We point out that the RTS-CTS exchange has multiple functions. Besides facilitating opportunity identification and channel fading estimation, it also mitigates the hidden and exposed terminal problem as in a conventional communication network [39]. Other collision avoidance schemes such as busy tone and dual busy tone may be incorporated to further reduce the occurrence of collision among secondary users.

Conclusion

In this chapter, we have discussed some of the technical challenges of cognitive MAC for OSA and made an initial attempt to establish a theoretical framework within which these challenges can be systematically and collectively addressed. In particular, the framework of POMDP makes the MAC cognitive; an opportunistic user makes optimal decisions for sensing and access based on the belief state that summarizes the knowledge of the network state based on all past decisions and observations.

[4] In this case, $S_n(t) = 1$ if channel n is available at both the transmitter and the receiver. Otherwise, $S_n(t) = 0$. Strictly speaking, the availability of a channel at the secondary transmitter is determined by primary receivers rather than primary transmitters in its neighborhood [29]. The detection of primary receivers can be transformed to the detection of primary transmitters. A detailed presentation can be found in [29].

This decision-theoretic framework also allows the integration of sensing errors, hardware limitations, and energy constraints into the modeling of cognitive MAC design.

References

1. "FCC spectrum policy task force: Report of the spectrum efficiency working group," Nov. 2002.
2. R. Coase, "The federal communications commission," J. Law Econ., pp. 1–40, 1959.
3. D. Hatfield and P. Weiser, "Property rights in spectrum: Taking the next step," in *Proc. of the first IEEE Symp. New Frontiers Dynamic Spectr. Access Networks*, pp. 43–55, Nov. 2005.
4. L. Xu, R. Tonjes, T. Paila, W. Hansmann, M. Frank, and M. Albrecht, "DRiVE-ing to the Internet: Dynamic radio for IP services in vehicular environments," in *Proc. of 25th Annual IEEE Conference on Local Computer Networks*, pp. 281–289, Nov. 2000.
5. Y. Benkler, "Overcoming agoraphobia: Building the commons of the digitally networked environment," *Harvard J. Law Technol.*, Winter 1997–1998.
6. W. Lehr and J. Crowcroft, "Managing shared access to a spectrum commons," in *Proc. of the first IEEE Symp. New Frontiers Dynamic Spectr. Access Networks*, pp. 420–444, Nov. 2005.
7. C. Raman, R. Yates, and N. Mandayam, "Scheduling variable rate links via a spectrum server," in *Proc. of the first IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks*, pp. 110–118, Nov. 2005.
8. O. Ileri, D. Samardzija, and N. Mandayam, "Demand responsive pricing and competitive spectrum allocation via a spectrum server," in *Proc. of the first IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks*, pp. 194–202, Nov. 2005.
9. S. Chung, S. Kim, J. Lee, and J. Cioffi, "A game-theoretic approach to power allocation in frequency-selective Gaussian interference channels," in *Proc. of IEEE International Symposium on Information Theory*, pp. 316–316, June 2003.
10. R. Etkin, A. Parekh, and D. Tse, "Spectrum sharing for unlicensed bands," in *Proc. of the first IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks*, pp. 251–258, Nov. 2005.
11. J. Huang, R. Berry, and M. Honig, "Spectrum sharing with distributed interference compensation," in *Proc. of the first IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks*, pp. 88–93, Nov. 2005.
12. J. Mitola, "Cognitive radio for flexible mobile multimedia communications," in *Proc. of IEEE International Workshop on Mobile Multimedia Communications*, pp. 3–10, Nov. 1999.
13. "DARPA: The Next Generation (XG) Program." http://www.darpa.mil/ato/programs/xg/index.htm.
14. J. Mitola, *Software Radios: Wireless Architecture for the 21st Century*. Wiley, 2000.
15. J. Mitola, "Cognitive radio." Licentiate proposal, KTH, Stockholm, Sweden.
16. Q. Zhao, L. Tong, and A. Swami, "Decentralized cognitive MAC for dynamic spectrum access," in *Proc. of the first IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks*, pp. 224–232, Nov. 2005.
17. Q. Zhao, L. Tong, A. Swami, and Y. Chen, "Decentralized cognitive MAC for opportunistic spectrum access in ad hoc networks: A POMDP framework," in *IEEE Journal on Selected Areas in Communications: Special Issue on Adaptive, Spectrum Agile and Cognitive Wireless Networks*, Apr., 2007.

18. Y. Chen, Q. Zhao, and A. Swami, "Distributed cognitive MAC for energy-constrained opportunistic spectrum access," in *Proc. of IEEE Military Communication Conference*, Oct. 2006.
19. Y. Chen, Q. Zhao, and A. Swami, "Joint design and separation principle for opportunistic spectrum access," in *Proc. of IEEE Asilomar Conference on Signals, Systems, and Computers*, Oct. 2006.
20. A. Sahai and N. Hoven and R. Tandra, "Some fundamental limits on cognitive radio," in *Proc. of Allerton Conference on Communication, Control, and Computing*, Oct. 2004.
21. D. Cabric, S. M. Mishra, and R. W. Brodersen, "Implementation issues in spectrum sensing for cognitive radios," in *Proc. of IEEE Asilomar Conference on Signals, Systems, and Computers*, pp. 772–776, Oct. 2004.
22. K. Challapali, S. Mangold, and Z. Zhong, "Spectrum agile radio: Detecting spectrum opportunities," in *Proc. of International Symposium on Advanced Radio Technologies*, 2004.
23. B. Wild and K. Ramchandran, "Detecting primary receivers for cognitive radio applications," in *Proc. of the first IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks*, pp. 124–130, Nov. 2005.
24. A. Ghasemi and E. Sousa, "Collaborative spectrum sensing for opportunistic access in fading environments," in *Proc. of the first IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks*, pp. 131–136, Nov. 2005.
25. H. Zheng and C. Peng, "Collaboration and fairness in opportunistic spectrum access," in *Proc. of IEEE International Conference on Communications*, vol. 5, pp. 3132–3136, May 2005.
26. W. Wang and X. Liu, "List-coloring based channel allocation for open-spectrum wireless networks," in *Proc. of IEEE Vehicular Technology Conference*, vol. 1, pp. 690–694, Sept. 2005.
27. S. Sankaranarayanan, P. Papadimitratos, A. Mishra, and S. Hershey, "A bandwidth sharing approach to improve licensed spectrum utilization," in *Proc. of the first IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks*, pp. 279–288, Nov. 2005.
28. M. Steenstrup, "Opportunistic use of radio-frequency spectrum: A network perspective," in *Proc. of the first IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks*, pp. 638–641, Nov. 2005.
29. Q. Zhao and B. Sadler, "A survey of dynamic spectrum access: Signal processing, networking, and regulatory policy," in *IEEE Signal Process. Magazine*, May, 2007.
30. Q. Zhao and A. Swami, "A decision-theoretic framework for dynamic spectrum access", to appear in *IEEE Wireless Communications Magazine*, 2007.
31. Y. Chen, Q. Zhao, and A. Swami, "Joint PHY-MAC design for opportunistic spectrum access in the presence of sensing errors," submitted to *IEEE Trans. Signal Process.*, Jan. 2007.
32. R. Smallwood and E. Sondik, "The optimal control of partially obserable Markov processes over a finite horizon," *Operat. Res.*, vol. 21, pp. 1071–1088, 1973.
33. Y. Chen, Q. Zhao, and A. Swami, "Proof of the separation principle for opportunistic spectrum access," Technical report, University of California at Davis, 2006. http://www.ece.ucdavis.edu/_qzhao/Report.html.
34. D. Djonin, Q. Zhao, and V. Krishnamurthy, "Optimality and complexity of opportunistic spectrum access: A truncated Markov decision process formulation," in *Proc. of IEEE International Conference on Communications (ICC)*, 2007.
35. H. L. V. Trees, Detection, Estimation, and Modulation Theory, Part I. Wiley-Interscience, 2001.

36. D. Aberdeen, “A survey of approximate methods for solving partially observable markov decision processe,” Technical report, National ICT Australia, Dec. 2003. http://users.rsise.anu.edu.au/˜daa/papers.html.
37. Q. Zhao, L. Tong, and A. Swami, “A cross-layer approach to cognitive mac for spectrum agility,” in *Proc. of IEEE Asilomar Conference on Signals, Systems, and Computers*, pp. 200–204, Oct.–Nov. 2005.
38. M. B. Pursley, “The role of spread spectrum in packet radio networks,” *Proc. IEEE*, vol. 75, pp. 116–134, Jan. 1987.
39. A. S. Tanenbaum, Computer Networks, 3rd ed. Upper Saddle River, NJ: Prentice-Hall PTR, 1996.