Research Article

Opportunistic Spectrum Access in Self-Similar Primary Traffic

Xiangyang Xiao, Keqin Liu, and Qing Zhao

Department of Electrical and Computer Engineering, University of California, Davis, CA 95616, USA

Correspondence should be addressed to Qing Zhao, qzhao@ece.ucdavis.edu

We take a stochastic optimization approach to opportunity tracking and access in self-similar primary traffic. Based on a multiple time-scale hierarchical Markovian model, we formulate opportunity tracking and access in self-similar primary traffic as a Partially Observable Markov Decision Process. We show that for independent and stochastically identical channels under certain conditions, the myopic sensing policy has a simple round-robin structure that obviates the need to know the channel parameters; thus it is robust to channel model mismatch and variations. Furthermore, the myopic policy achieves comparable performance as the optimal policy that requires exponential complexity and assumes full knowledge of the channel model.

1. Introduction

1.1. Opportunistic Spectrum Access. The "spectrum paradox" is by now widely recognized. On the one hand, the projected spectrum need for wireless devices and services continues to grow, and virtually all usable radio frequencies have already been allocated. Such an imbalance in supply and demand threatens one of the most explosive economic and technological growths in the past decades. On the other hand, extensive measurements conducted in the recent years reveal that much of the prized spectrum lies unused at any given time and location [1]. For example, in a recent measurement study of wireless LAN traffic [2], a typical active FTP session has about 75% idle time, and voice-over-IP applications such as Skype have up to 90% idle time.

These measurements of actual spectrum usage highlight the drawbacks of the current static spectrum allotment policy that is at the root of this spectrum paradox. They also form the key rationale for *Opportunistic Spectrum Access (OSA)* envisioned by the DARPA XG program and currently being considered by the FCC [3]. The idea of OSA is to exploit instantaneous spectrum opportunities by opening the licensed spectrum to secondary users. This would allow secondary users to identify available spectrum resources and communicate nonintrusively by limiting interference to primary users. Even for unlicensed bands, OSA may be of considerable value for spectrum efficiency by adopting a hierarchical pricing structure to support both subscribers and opportunistic users.

1.2. Opportunistic Spectrum Access in Self-Similar Primary Traffic. Since the seminal work of Leland et al. [4], extensive studies have shown that self-similarity manifests in communications traffic in diverse contexts, from local area networks to wide area networks, from wired to wireless applications [5–8]. In this paper, we consider opportunistic spectrum access in self-similar primary traffic processes with long-range dependency. We adopt a multiple time-scale hierarchical Markovian model for self-similar traffic processes proposed in [9, 10]. A decision theoretic framework is developed based on the theory of Partially Observable Markov Decision Processes (POMDPs).

Unfortunately, solving a general POMDP is often intractable due to the exponential complexity. A simple approach is to implement the myopic policy, which only focuses on maximizing the immediate reward and ignores the impact of current action on the future reward. We show in this paper that the myopic policy has a simple and robust structure under certain conditions. This simple structure obviates the need to know the transition probabilities of the underlying multiple time-scale Markovian model and allows automatic tracking of variations in the primary traffic model. Compared to Markovian channel models, the model at hand

is more general but requires more parameters, it is thus more important to have policies that are robust to model mismatch and parameter variations. The strong performance of the myopic policy with such a simple and robust structure is demonstrated through simulation examples.

1.3. Related Work. This paper is perhaps the first that addresses OSA in self-similar primary traffic. It builds upon our prior work on a POMDP framework for the joint design of opportunistic spectrum access that adopts a first-order Markovian model for the primary traffic. Specifically, in [11–13], a decision-theoretic framework for tracking and exploiting spectrum opportunities is developed using a first-order Markovian model for the primary traffic. A fundamental result on the principle of separation for OSA [14, 15] and structural opportunity tracking polices [16, 17] have been established, leading to simple, robust, and optimal solutions.

The first-order Markovian model of the primary traffic, however, has its limitations. It cannot capture the long-range dependency exhibited in a wide range of communications traffic. In this paper, we extend the decision-theoretic framework developed in [11, 12, 14, 15] to incorporate self-similar primary traffic with long-range dependency. We show that the structure and optimality of the myopic sensing policy established in [16, 17] under a first-order Markovian model are preserved under certain conditions in self-similar primary traffic modeled by a multiple time-scale hierarchical Markovian process.

2. A Multiple Time-Scale Hierarchical Markovian Model for Self-Similar Traffic

A fundamental property of a self-similar process is the "scale-invariant behavior." The process is stochastically unchanged when it is zoomed out by stretching the time domain [18]. Specifically, $\{X(t) : t \in \mathcal{R}\}$ is a self-similar process if for any $k \geq 1, t_1, \ldots, t_k \in \mathcal{R}$, and $a, H \in \mathcal{R}^+$,

$$(X(at_1), \ldots, X(at_k)) \overset{d}{=} \left(a^H X(t_1), \ldots, a^H X(t_k)\right), \quad (1)$$

where $\overset{d}{=}$ represents equivalence in distribution. It has been shown that for $1/2 < H < 1$, the autocorrelation of a self-similarly process decreases to zero polynomially, leading to a long-range dependency behavior.

Based on traffic traces from physical networks, several models for self-similar traffic have been developed, among which is a multiple time-scale hierarchical Markovian model proposed in [9, 10]. Under this model, traffic is an aggregation of hierarchical Markovian on-off processes with disparate time-scales. Illustrated in Figure 1 is a two-level hierarchical on-off process. The higher level process has a much slower transition rate than the lower one. The resulting traffic process is "on" (busy) when both Markovian processes are in state 0 and "off" (idle) otherwise. This hierarchical model with two to three levels has been shown to approximate a self-similar process and fit well with measured traffic traces. It is motivated by the physical process of traffic
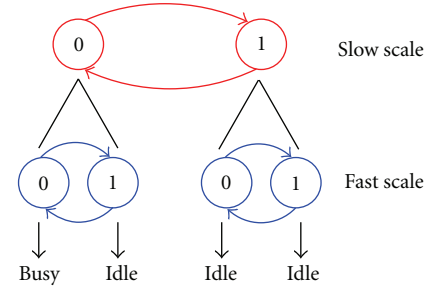

FIGURE 1: A multiple time-scale hierarchical Markovian model for self-similar primary traffic.

generation [9, 10]. Specifically, for a packet to appear in the physical channel, several events at different time scales have to occur, including, for example, establishing a session, releasing a message to the network by a transport protocol like TCP, then releasing a packet to the channel by the MAC and physical layers [9, 10].

This hierarchical on-off process can be described by a Markov process with augmented state. For example, the above two-level hierarchical on-off process can be treated as a Markov process with 4 states. The resulting traffic model is thus a hidden Markov model: the state $(0, 0)$ is directly observable and mapped to "on," and the remaining 3 states are mapped to a single state "off." This hidden Markovian interpretation is the key to our POMDP formulation of opportunity tracking and exploitation in self-similar primary traffic as shown in the next section.

3. A POMDP Framework

In this section, we show that under the multiple time-scale hierarchical Markovian model, opportunity tracking can still be formulated as a POMDP similar to that developed in [11–15] under a first-order Markovian model.

3.1. Network Model. Consider a spectrum consisting of N channels, each with transmission rate $B_n(n = 1, \ldots, N)$. These N channels are allocated to a primary network with slotted transmissions. The primary traffic in each channel is a self-similar process following the hierarchical Markovian model with L levels. Each channel can thus be represented by an augmented Markov chain with 2^L states (see Figure 1 above where $L = 2$). The availability (idle or busy) of a channel, that is, the primary traffic trace, is determined by the state of the corresponding augmented Markov chain.

Let $\{p_{ij}^{(n,k)}\}_{i,j=0,1}$ denote the transition probabilities of the kth $(1 \leq k \leq L)$ level Markov process for channel n. We thus have $p_{ii}^{(n,k)} \gg p_{ii}^{(n,k+1)}$ and $p_{ij}^{(n,k)} \ll p_{ij}^{(n,k+1)}$ for $i \neq j$, where $1 \leq k \leq L - 1$. In other words, the kth level Markov process varies much slower than the mth level Markov process for $m > k$. It can be shown that the kth level Markov process for channel n is *positively correlated* when $p_{11}^{(n,k)} > p_{01}^{(n,k)}$, and *negatively correlated* when $p_{11}^{(n,k)} < p_{01}^{(n,k)}$. We notice that the Markov processes at higher levels (i.e., with smaller level

indexes) can be considered as positively correlated due to their slow transition rates.

Consider next a pair of secondary transmitter and receiver seeking spectrum opportunities in these N channels. In each slot, they choose a channel to sense. If the channel is idle, the transmitter sends packages to the receiver through this channel, and a reward $R(t)$ is accrued in this slot (i.e., the number of bits delivered). It is straight forward to generalize the POMDP framework and the results in Section 4 to multichannel sensing scenarios. We assume here that the secondary user has reliable detection of the channel availability.

Our goal is to develop the optimal sensing policy to maximize the throughput of the secondary user during a desired period of T slots.

3.2. POMDP Formulation. The sequential decision-making process described above can be modeled as a POMDP.

Specifically, the underlying system state is given by the state of the augmented Markov chain at the beginning of each slot. Let $S_n(t) = (S_n^{(1)}(t), S_n^{(2)}(t), \ldots, S_n^{(L)}(t))$ denote the state of channel n in slot t, where $S_n^{(k)}(t) \in \{0, 1\}$ represents the state of the kth level Markov process for channel n in slot t. The transition probabilities of this augmented Markov chain can be easily obtained from $\{p_{ij}^{(n,k)}\}_{i,j=0,1} (1 \leq k \leq L)$. Let $O_n(t) \in \{0, 1\}$ denote the availability of channel n in slot t, that is, $O_n(t) = 0$ (busy) when $S_n^{(k)}(t) = 0$ for all $1 \leq k \leq L$ and $O_n(t) = 1$ (idle/opportunity) otherwise.

The reward in each slot is the number of bits that can be delivered by the secondary user. Given sensing action $a(t)$, the immediate reward $R_{a(t)}(t)$ is given by

$$R_{a(t)}(t) = O_{a(t)}(t) B_{a(t)}. \tag{2}$$

Due to the hidden Markovian model of channel availability and partial sensing, the state $S_n(t)$ of the augmented Markov chain representing each channel cannot be fully observed. The statistical information on $S_n(t)$ provided by the entire decision and observation history can be encapsulated in a belief vector $\Lambda_n(t) = \{\lambda_{(n,s)}(t) : s = \{s_k\}_{k=1}^L \in \{0, 1\}^L\}$, where

$$\lambda_{(n,s)}(t) = \Pr\left(S_n(t) = s \mid \{a(i), O_{a(i)}(i)\}_{i=1}^{t-1}\right) \tag{3}$$

represents the conditional probability (given the decision and observation history) that the state of channel n is s in slot t. The whole system state is given by the concatenation of each channel's belief vector:

$$\Lambda(t) = [\Lambda_1(t), \Lambda_2(t), \ldots, \Lambda_N(t)]. \tag{4}$$

This system belief vector $\Lambda(t)$ is a sufficient statistic for making the optimal action in each slot t. Furthermore, $\Lambda(t+1)$ for slot $t+1$ can be obtained from $\Lambda(t)$, $a(t)$, and $O_{(a(t))}(t)$ via Bayes rule as shown in what follows. Let $q_{s's}^{(n)} (s', s \in \{0, 1\}^L)$ denote the transition probability from state s' to s

for channel n, it is easy to see that $q_{s's}^{(n)} = \prod_{k=1}^L p_{s_k' s_k}^{(n,k)}$. Define $s_0 \triangleq \{s_k : s_k = 0\}_{k=1}^L$ and

$$\lambda'_{(n,s)}(t) \triangleq \begin{cases} \dfrac{\lambda_{(n,s)}(t)}{1 - \lambda_{(n,s_0)}(t)}, & \text{if } s \neq s_0, \\ 0, & \text{if } s = s_0. \end{cases} \tag{5}$$

We then have

$$\lambda_{(n,s)}(t+1) = \begin{cases} \displaystyle\sum_{s' \in \{0,1\}^L} \lambda'_{(n,s')}(t) q_{(s's)}^{(n)}, & a(t) = n, O_n(t) = 1, \\ q_{(s_0 s)}^{(n)}, & a(t) = n, O_n(t) = 0, \\ \displaystyle\sum_{s' \in \{0,1\}^L} \lambda_{(n,s')}(t) q_{(s's)}^{(n)}, & a(t) \neq n. \end{cases} \tag{6}$$

Let $\pi = \{\pi_t\}_{t=1}^T$ be a series of mappings from $\Lambda(t)$ to $a(t)$ for each $1 \leq t \leq T$, which denotes the sensing policy for channel selection. We then arrive at the following stochastic control problem:

$$\pi^* = \arg \max_{\pi} \mathbb{E}_{\pi}\left[\sum_{t=1}^T R_{\pi_t(\Lambda(t))}(t) \mid \Lambda(1)\right], \tag{7}$$

where $\mathbb{E}_{\pi}$ represents the expectation given that the sensing policy π is employed, $\pi_t(\Lambda(t))$ is the sensing action in slot t under policy π, and $\Lambda(1)$ is the initial belief vector. When no information is available to the secondary user at the beginning of the first slot, $\Lambda(1)$ is given by the stationary distributions of the on-off Markov processes at all levels of these N channels. Specifically, $\lambda_{(n,s)}(1)$ is given by

$$\lambda_{(n,s)}(1) = \prod_{k=1}^L \left(s_k \frac{p_{01}^{(n,k)}}{p_{01}^{(n,k)} + p_{10}^{(n,k)}} + (1 - s_k) \frac{p_{10}^{(n,k)}}{p_{01}^{(n,k)} + p_{10}^{(n,k)}}\right). \tag{8}$$

4. The Myopic Policy and Its Semiuniversal Structure

The myopic policy ignores the impact of the current action on the future reward, focusing solely on maximizing the expected immediate reward $\mathbb{E}[R_{a(t)}(t)]$. The myopic action $\hat{a}(t)$ in slot t given current belief vector $\Lambda(t)$ is thus given by

$$\begin{aligned} \hat{a}(t) &= \arg \max_{a=1,\ldots,N} \Pr[O_a(t) = 1 \mid \Lambda(t)] B_a \\ &= \arg \max_{a=1,\ldots,N} (1 - \lambda_{(a,s_0)}(t)) B_a. \end{aligned} \tag{9}$$

In general, obtaining the myopic action in each slot requires the recursive update of the belief vector $\Lambda(t)$ as given in (6). Next we show that under certain conditions, the myopic policy for stochastically identical channels has a semiuniversal structure that does not need the update of the belief vector or the knowledge of the transition probabilities.

Consider stochastically identical channels. Let $\{p_{ij}^{(k)}\}_{i,j=0,1}$ denote the transition probabilities of the kth level Markov process for all channels, B the transmission rate of all channels.

We establish a simple and robust structure of the myopic policy under certain conditions as shown in Theorem 1. Let $\omega_n^{(k)}(t)$ denote the marginal probability that the state of the kth level Markov process for channel n is 1 in slot t, it is easy to see that $\omega_n^{(k)}(t) = \sum_{s:s_k=1} \lambda_{(n,s)}(t)$. We assume that the initial states are independent across all levels, that is, for all $s \in \{0,1\}^L$

$$\lambda_{(n,s)}(1) = \prod_k \left(s_k \omega_n^{(k)}(1) + (1 - s_k)\left(1 - \omega_n^{(k)}(1)\right)\right). \quad (10)$$

Theorem 1. *Suppose that channels are independent and stochastically identical, the Markov processes at all levels are positively correlated, and the initial states of the Markov processes are independent across all levels for each channel. Furthermore, the initial system belief vector $\Lambda(1)$ satisfies the following condition. There exists a channel ordering $(n_1, n_2, \ldots, n_N)$ such that $\omega_{n_1}^{(k)}(1) \geq \omega_{n_2}^{(k)}(1) \geq \cdots \geq \omega_{n_N}^{(k)}(1)$ for all $1 \leq k \leq L$, that is, the channel ordering by the initial states at all levels is the same. The myopic policy has a round robin structure based on the circular channel ordering $(n_1, n_2, \ldots, n_N)$. Starting from sensing channel n_1 in slot 1, the myopic action is to stay in the same channel when it is idle and switch to the next channel in the circular ordering when it is busy.*

Proof. Without loss of generality, we assume $B = 1$. We first prove the following three lemmas.

Lemma 1. *If the states of the Markov processes are independent (conditioned on the past observations) across all levels for each channel in slot t, and there exists a channel ordering $(n_1, n_2, \ldots, n_N)$ such that $\omega_{n_1}^{(k)}(t) \geq \omega_{n_2}^{(k)}(t) \geq \cdots \geq \omega_{n_N}^{(k)}(t)$ for all $1 \leq k \leq L$, then for all $t' > t$, $\omega_{n_1}^{(k)}(t') \geq \omega_{n_2}^{(k)}(t') \geq \cdots \geq \omega_{n_N}^{(k)}(t')$ for all $1 \leq k \leq L$ when no observation is made from t up to t'.*

Proof. Starting from slot t, the independence of the states of the Markov processes (conditioned on the past observations) across all levels for each channel will hold as long as no observation is made. For any channel n, we have, for all $1 \leq k \leq L$,

$$\omega_n^{(k)}(t') = \left(p_{11}^{(k,t'-t)} - p_{01}^{(k,t'-t)}\right)\omega_n^{(k)}(t) + p_{01}^{(k,t'-t)}, \quad (11)$$

where $p_{ij}^{(k,m)}$ is the m-step transition probability from state i to j at the kth level. In particular, we have

$$p_{01}^{(k,m)} = \frac{p_{01}^{(k)} - p_{01}^{(k)}\left(p_{11}^{(k)} - p_{01}^{(k)}\right)^m}{p_{01}^{(k)} + p_{10}^{(k)}},$$
$$p_{11}^{(k,m)} = \frac{p_{01}^{(k)} + p_{10}^{(k)}\left(p_{11}^{(k)} - p_{01}^{(k)}\right)^m}{p_{01}^{(k)} + p_{10}^{(k)}}. \quad (12)$$

Since $p_{11}^{(k)} \geq p_{01}^{(k)}$, we have $p_{11}^{(k,m)} \geq p_{01}^{(k,m)}$ for any $m \in \mathcal{Z}^+$. Consider two channels a and b with $\omega_a^{(k)}(t) \geq \omega_b^{(k)}(t)$. From (11), it is easy to see that $\omega_a^{(k)}(t') \geq \omega_b^{(k)}(t')$ for any $t' > t$. Lemma 1 thus holds. $\square$

Lemma 2. *If the states of the Markov processes are independent (conditioned on the past observations) across all levels for each channel in slot t and the chosen channel n is observed in state 0, then channel n will have the smallest probability $\omega_n^{(k)}(t+1)$ of being in state 1 at level k for all $1 \leq k \leq L$ in slot $t+1$.*

Proof. Given observation $O_n(t) = 0$ in slot t, we have $\omega_n^{(k)}(t+1) = p_{01}^{(k)}$ for all $1 \leq k \leq L$. From (11), p_{01}^k is the smallest belief value at level k among all channels in slot $t+1$. $\square$

Lemma 3. *Consider channel n with current belief $\Lambda_n(t)$ in slot t. For $t' > t$, let $\lambda_{n,s_0}^{(k)}(t')$ $(0 \leq k \leq t' - t)$ denote the belief value in slot t' if k1s are successively observed from slot t to $t + k - 1$. One has*

$$\lambda_{n,s_0}^{(t'-t)}(t') \leq \lambda_{n,s_0}^{(0)}(t'). \quad (13)$$

Proof. From (12), we have

$$p_{11}^{(k,j)} \geq p_{01}^{(k,j)},$$
$$p_{00}^{(k,j)} \geq p_{10}^{(k,j)}. \quad (14)$$

Let $q_{(s,s')}^{(j)}$ denote the j-step transition probability from channel state s to s'. From (14), it is easy to see that $q_{(s,s)}^{(j)} \geq q_{(s',s)}^{(j)}$ for all $s, s' \in \{0,1\}^L$. We thus have

$$
\begin{aligned}
\lambda_{n,s_0}^{(k)}(t') &= \sum_{s \in \{0,1\}^L, s \neq s_0} \frac{\lambda_{(n,s)}^{(k-1)}(t+k-1)}{1 - \lambda_{(n,s_0)}^{(k-1)}(t+k-1)} q_{(s,s_0)}^{(t'-t-k+1)} \\
&\leq \sum_{s \in \{0,1\}^L, s \neq s_0} \lambda_{(n,s)}^{(k-1)}(t+k-1) q_{(s,s_0)}^{(t'-t-k+1)} \\
&\quad + q_{(s_0,s_0)}^{(t'-t-k+1)} \lambda_{(n,s_0)}^{(k-1)}(t+k-1) \\
&\quad \times \sum_{s \in \{0,1\}^L, s \neq s_0} \frac{\lambda_{(n,s)}^{(k-1)}(t+k-1)}{1 - \lambda_{(n,s_0)}^{(k-1)}(t+k-1)} \\
&= \sum_{s \in \{0,1\}^L, s \neq s_0} \lambda_{(n,s)}^{(k-1)}(t+k-1) q_{(s,s_0)}^{(t'-t-k+1)} \\
&\quad + \lambda_{(n,s_0)}^{(k-1)}(t+k-1) q_{(s_0,s_0)}^{(t'-t-k+1)} \\
&= \sum_{s \in \{0,1\}^L} \lambda_{(n,s)}^{(k-1)}(t+k-1) q_{(s,s_0)}^{(t'-t-k+1)} \\
&= \lambda_{n,s_0}^{(k-1)}(t').
\end{aligned}
\quad (15)
$$

We thus have $\lambda_{n,s_0}^{(t'-t)}(t') \leq \lambda_{n,s_0}^{(t'-t-1)}(t') \leq \cdots \leq \lambda_{n,s_0}^{(0)}(t')$. $\square$

We are now ready to prove the theorem. Assume we observed 0 on channel n_1 in slot 1. Based on the independence of the initial states of the Markov processes across all levels for each channel, channel n_1 will have the smallest probability of being idle in the next slot according
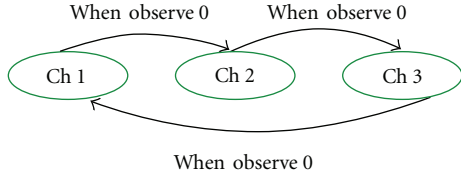
FIGURE 2: The round robin structure of the myopic policy ($N = 3$).

to Lemma 2. Furthermore, the order of the probabilities of being in state 1 at each level for all unobserved channels will be the same according to Lemma 1, and channel n_1 will have the smallest probability of being in state 1 at each level. The independence of the states (conditioned on the past observations) and ordering conditions on the initial system belief vector still hold in the next slot with channel ordering $(n_2, \ldots, n_N, n_1)$. On the other hand, if we observe 1 on channel n_1 in slot 1, channel n_1 will have the largest probabilities of being idle as long as the observations are 1 on this channel according to Lemmas 3 and 1. When a 0 is observed after the 1s, channel n_1 will again have the smallest probability of being in state 1 at each level. The independence of the states (conditioned on the past observations) and ordering conditions on the initial system belief vector still hold in the next slot with channel ordering $(n_2, \ldots, n_N, n_1)$. By induction, it is easy to see that the myopic policy has a round robin structure with circular ordering $(n_1, n_2, \ldots, n_N)$. □

Figure 2 shows an example of the round robin structure of the myopic policy when $N = 3$ with a circular channel order of $(1, 2, 3)$. The myopic action is to sense the three channels in turn with random switching times (when the current channel is busy).

In practice, the channel ordering assumption on the initial system belief vector in Theorem 1 means that at the beginning if one level of a channel (say channel m) is more likely to be idle than the same level of another channel (say channel n), then every level of channel m is more likely to be idle than the same level of channel n. For example, in the first slot, if it is less likely to have a session established over channel m than over channel n, then a message is less likely to be released by a primary user over channel m than over channel n. If the initial channel ordering assumption is not satisfied, then before all channels have been visited, the myopic policy may not have the round-robin structure. However, after all channels have been visited, the channel ordering assumption will be satisfied and the structure of the myopic policy will hold thereafter.

We notice that the secondary user usually has no initial information about the channel availability. In this case, the initial system belief vector is given by the stationary distributions of the underlying Markov processes as given in (8). The channel ordering assumption on the initial system belief vector in Theorem 1 is thus satisfied since stochastically identical channels have the same stationary distribution at the same level. The circular channel ordering in the round-robin structure can be set arbitrarily.

For two-level hierarchical Markovian channel models ($L = 2$), we can relax the condition on the initial system belief vector in Theorem 1 without affecting the round robin structure of the myopic policy.

Theorem 2 (relaxation of initial condition). *Suppose that channels are independent and stochastically identical, the Markov processes at all levels are positively correlated, and the initial states of the Markov processes are independent across all levels for each channel. The round robin structure of the myopic policy given in Theorem 1 remains unchanged when for any two channels i and j with $\omega_i^{(1)}(1) \geq \omega_j^{(1)}(1)$, the following two equations hold:*

$$
\begin{aligned}
&\prod_{k=1}^{2}\left(1 - \omega_i^{(k)}(1)\right) \leq \prod_{k=1}^{2}\left(1 - \omega_j^{(k)}(1)\right), \\
&\omega_i^{(2)}(1) - \omega_j^{(2)}(1) \\
&\geq -\frac{\left(1 - p_{11}^{(2)}\right)\left(p_{11}^{(1)} - p_{01}^{(1)}\right)\left(\omega_i^{(1)}(1) - \omega_j^{(1)}(1)\right)}{\left(1 - p_{11}^{(1)}\right)\left(p_{11}^{(2)} - p_{01}^{(2)}\right)}.
\end{aligned}
\tag{16}
$$

Proof. The proof follows a similar process to Theorem 1. We first prove the following lemma.

Lemma 4. *If the states of the Markov processes are independent (conditioned on the past observations) across all levels for each channel in slot t, and there exists a channel ordering $(n_1, n_2, \ldots, n_N)$ such that $\omega_{n_1}^{(1)}(t) \geq \omega_{n_2}^{(1)}(t) \geq \cdots \geq \omega_{n_N}^{(1)}(t)$ and for any $1 \leq i < j \leq N$,*

$$
\prod_{k=1}^{2}\left(1 - \omega_{n_i}^{(k)}(t)\right) \leq \prod_{k=1}^{2}\left(1 - \omega_{n_j}^{(k)}(t)\right),
\tag{17}
$$

$$
\begin{aligned}
&\omega_{n_i}^{(2)}(t) - \omega_{n_j}^{(2)}(t) \\
&\geq -\frac{\left(1 - p_{11}^{(2)}\right)\left(p_{11}^{(1)} - p_{01}^{(1)}\right)\left(\omega_{n_i}^{(1)}(t) - \omega_{n_j}^{(1)}(t)\right)}{\left(1 - p_{11}^{(1)}\right)\left(p_{11}^{(2)} - p_{01}^{(2)}\right)}.
\end{aligned}
\tag{18}
$$

One then has for all $t' > t$ under the condition that no observation is made from t to t', (17) and (18) still hold if t is replaced by t', and $\lambda_{n_1, s_0}(t') \leq \lambda_{n_2, s_0}(t') \leq \cdots \leq \lambda_{n_N, s_0}(t')$.

Proof. By induction, we only need to prove the lemma for $t' = t+1$. Since no observation is made from t to $t+1$, we have that the states of the Markov processes for each channel are independent from t to $t + 1$. We thus have that the expected immediate reward $f(\omega_n^{(1)}(k), \omega_n^{(2)}(k))$ for channel n at time $k (t \leq k \leq t + 1)$ is given by

$$
\begin{aligned}
&f\left(\omega_n^{(1)}(k), \omega_n^{(2)}(k)\right) \\
&= 1 - \lambda_{n, s_0}(k) = 1 - \left(1 - \omega_n^{(1)}(k)\right)\left(1 - \omega_n^{(2)}(k)\right).
\end{aligned}
\tag{19}
$$

Assume that for channel i and j, we have $f(\omega_i^{(1)}(t), \omega_j^{(2)}(t)) \geq f(\omega_i^{(1)}(t), \omega_i^{(2)}(t))$. Define $\Delta^{(k)} \triangleq p_{11}^{(k)} - p_{01}^{(k)}$ for $k = 1, 2$. We

then have $\Delta^{(1)} \geq \Delta^{(2)} > 0$, $\omega_j^{(1)}(t) \geq \omega_i^{(1)}(t)$, $\omega_n^{(k)}(t+1) = \Delta^{(k)}\omega_n^{(k)}(t) + p_{01}^{(k)}$, and

$$
\begin{aligned}
&\omega_j^{(2)}(t) - \omega_i^{(2)}(t) \\
&\geq -\frac{\left(1 - p_{11}^{(2)}\right)\Delta^{(1)}\left(\omega_j^{(1)}(t) - \omega_i^{1}(t) - \omega_i^{(1)}(t)\right)}{\left(1 - p_{11}^{(1)}\right)\Delta^{(2)}} \\
&\Longleftrightarrow \left(1 - p_{11}^{(1)}\right)\left(\Delta^{(2)}\right)^2\left(\omega_j^{(2)}(t) - \omega_i^{(2)}(t)\right) \\
&\quad + \left(1 - p_{11}^{(2)}\right)\Delta^{(1)}\left(\Delta^{(2)}\right)\left(\omega_j^{(1)}(t) - \omega_i^{(1)}(t)\right) \geq 0 \quad (20) \\
&\Longrightarrow \left(1 - p_{11}^{(1)}\right)\left(\Delta^{(2)}\right)^2\left(\omega_j^{(2)}(t) - \omega_i^{(2)}(t)\right) \\
&\quad + \left(1 - P_{11}^{(2)}\right)\left(\Delta^{(1)}\right)^2\left(\omega_j^{(1)}(t) - \omega_i^{(1)}(t)\right) \geq 0 \\
&\Longrightarrow \left(1 - p_{11}^{(1)}\right)\Delta^{(2)}\left(\omega_j^{(2)}(t+1) - \omega_i^{(2)}(t+1)\right) \\
&\quad + \left(1 - p_{11}^{(2)}\right)\Delta^{(1)}\left(\omega_j^{(1)}(t+1) - \omega_i^{(1)}(t+1)\right) \geq 0.
\end{aligned}
$$

We thus proved that (18) still holds if t is replaced by $t+1$.

From (20) and $f(\omega_j^{(1)}(t), \omega_j^{(2)}(t)) \geq f(\omega_i^{(1)}(t), \omega_i^{(2)}(t))$, we have

$$
\begin{aligned}
&\left(1 - p_{11}^{(1)}\right)\Delta^{(2)}\left(\omega_j^{(2)}(t) - \omega_i^{(2)}(t)\right) \\
&\quad + \left(1 - p_{11}^{(2)}\right)\Delta^{(1)}\left(\omega_j^{(1)}(t) - \omega_i^{(1)}(t)\right) \\
&\quad + \Delta^{(1)}\Delta^{(2)}\left(f\left(\omega_j^{(1)}(t), \omega_j^{(2)}(t)\right) - f\left(\omega_i^{(1)}(t), \omega_i^{(2)}(t)\right)\right) \geq 0 \\
&\Longleftrightarrow \left(1 - p_{11}^{(1)}\right)\Delta^{(2)}\left(\omega_j^{(2)}(t) - \omega_i^{(2)}(t)\right) \\
&\quad + \left(1 - p_{11}^{(2)}\right)\Delta^{(1)}\left(\omega_j^{(1)}(t) - \omega_i^{(1)}(t)\right) \\
&\quad + \Delta^{(1)}\Delta^{(2)}\left(\omega_j^{(1)}(t) + \omega_j^{(2)}(t) - \omega_j^{(1)}(t)\omega_j^{(2)}(t)\right. \\
&\quad \left. - \left(\omega_i^{(1)}(t) + \omega_i^{(2)}(t) - \omega_i^{(1)}(t)\omega_i^{(2)}(t)\right)\right) \geq 0 \\
&\Longleftrightarrow \omega_j^{(1)}(t+1) + \omega_j^{(2)}(t+1) - \omega_j^{(1)}(t+1)\omega_j^{(2)}(t+1) \\
&\geq \left(\omega_i^{(1)}(t+1) + \omega_i^{(2)}(t+1) - \omega_i^{(1)}(t+1)\omega_i^{(2)}(t+1)\right) \\
&\Longleftrightarrow f\left(\omega_i^{(1)}(t+1), \omega_i^{(2)}(t+1)\right) \geq f\left(\omega_j^{(1)}(t+1), \omega_j^{(2)}(t+1)\right).
\end{aligned}
$$
(21)

From (21) and (19), we proved that (17) still holds if t is replaced by $t+1$, and $\lambda_{n_1,s_0}(t') \leq \lambda_{n_2,s_0}(t') \leq \cdots \leq \lambda_{n_N,s_0}(t')$. □

We are now ready to prove the theorem. Assume we observed 0 on channel n_1 in slot 1. Based on the independence of the initial states of the Markov processes across all levels for each channel, then channel n_1 will have the smallest probability of being in state 1 at each level in the next slot according to Lemma 2. From Lemma 4, the order of the probability of being idle for all unobserved channels will not change. Furthermore, we notice that the independence of the states (conditioned on the past observations) and ordering conditions (16) on the initial system belief vector still hold in the next slot with channel ordering $(n_2, \ldots, n_N, n_1)$. On the other hand, if we observe 1 on channel n_1 in slot 1, channel n_1 will have the largest probabilities of being idle as long as the observations are 1 on this channel according to Lemmas 3 and 4. When a 0 is observed after the 1s, channel n_1 will again
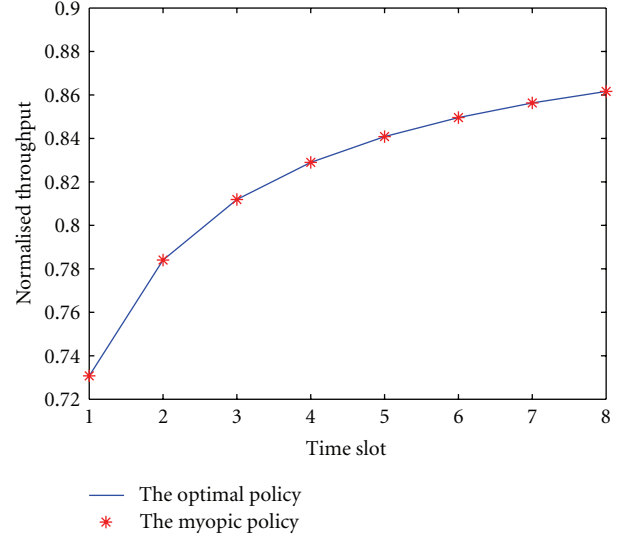

FIGURE 3: The performance of the myopic policy ($N = 3$, $L = 2$, $p_{11}^{(1)} = 0.95$, $p_{01}^{(1)} = 0.05$, $p_{11}^{(2)} = 0.65$, and $p_{01}^{(2)} = 0.3$.).

have the smallest probability of being in state 1 at each level. From Lemma 4, the independence of the states (conditioned on the past observations) and ordering conditions (16) on the initial system belief vector still hold in the next slot with channel ordering $(n_2, \ldots, n_N, n_1)$. By induction, it is easy to see that the myopic policy has a round robin structure with circular ordering $(n_1, n_2, \ldots, n_N)$. □

Theorems 1 and 2 show that the myopic policy is a round-robin scheme (see Figure 2 where $N = 3$) for stochastically identical channels under certain conditions. This semiuniversal structure leads to robustness against model mismatch and variations.

5. Simulation Examples

In this section, we illustrate the performance and robustness of the myopic policy for independent and stochastically identical channels. Based on Theorem 1, the myopic policy is implemented in the following steps.

Step 1. Obtain the initial channel ordering $(n_1, n_2, \ldots, n_N)$, that is, , $\omega_{n_1}^{(k)}(1) \geq \omega_{n_2}^{(k)}(1) \geq \cdots \geq \omega_{n_N}^{(k)}(1)$ for all $1 \leq k \leq L$.

Step 2. In the first slot, the myopic policy chooses channel n_1 to sense.

Step 3. For any $t(t \geq 1)$, if the currently sensed channel (say n_i) is idle, then we will sense it again in slot $t+1$. Otherwise we sense the next channel (i.e., channel n_{i+1} if $1 \leq i < N$ or channel n_1 if $i = N$) in the circular ordering $(n_1, n_2, \ldots, n_N)$.

In Figure 3, the system belief vector starts from the stationary distributions of the underlying Markov processes. For this example, the conditions in Theorem 1 are satisfied and the myopic policy obeys a round robin
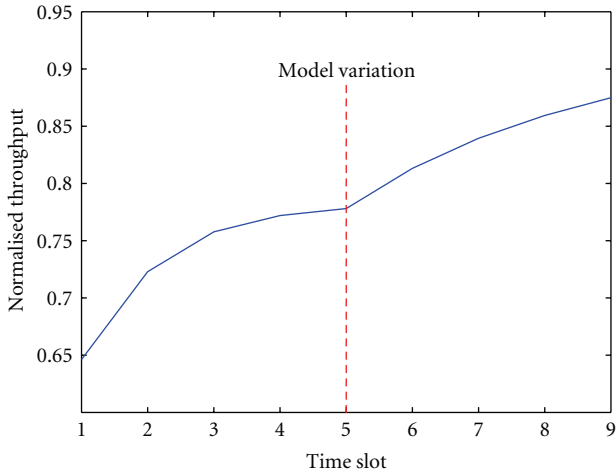
FIGURE 4: The robustness of the myopic policy ($N = 3$, $L = 2$. For $t \leq 4$, $p_{11}^{(1)} = 0.9$, $p_{01}^{(1)} = 0.05$, $p_{11}^{(2)} = 0.2$, and $p_{01}^{(2)} = 0.1$; for $t > 4$, $p_{11}^{(1)} = 0.99$, $p_{01}^{(1)} = 0.69$, $p_{11}^{(2)} = 0.9$, and $p_{01}^{(2)} = 0.8$.).

structure. We observe that the myopic policy achieves identical performance as the optimal policy that requires exponential complexity assumes full knowledge of the transition probabilities at all levels of the hierarchical channel model.

Figure 4 shows an example that the myopic policy can automatically track model variations. The transition probabilities in this example change abruptly at the fifth slot, which corresponds to a drop in the primary traffic load. It can be shown that these variations will not affect the round robin structure of the myopic policy as long as the conditions in Theorem 1 are satisfied. From this figure, we can observe from the change in the throughput increasing rate that the myopic policy effectively tracks the traffic model variations in the primary system.

We point out that when channels are independent but stochastically nonidentical, the myopic policy is not optimal in general. From Figure 5, we observe that the myopic policy has a performance loss compared to the optimal one. However, the myopic policy can still achieve a near optimal performance.

Last, we show an example that the myopic policy is optimal for independent and stochastically identical channels when there are sensor errors. Under the Markovian model (i.e., each channel has only one level), a separation principle that decouples the design of spectrum sensor and access policies from that of the sensing policy has been established in [14, 15]. While the separation principle may not hold under the multiple time-scale hierarchical Markovian model, the separate design still provides a simple and valid solution. Specifically, the spectrum sensor policy is to choose the detection threshold such that the probability of miss detection is equal to the maximum allowable probability of collision to the primary users. The access policy is simply to trust the detection outcome. Using these designs of the spectrum sensor and the access policy, we then design the
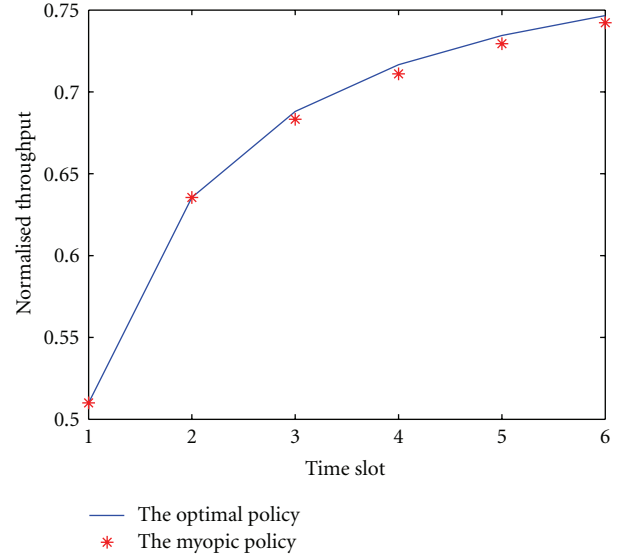

FIGURE 5: The performance of the myopic policy for nonidentical channels ($N = 5$, $L = 2$, $p_{01}^{(1)} = [0.2\ 0.2\ 0.4\ 0.4\ 0.4]$, $p_{01}^{(2)} = [0.4\ 0.4\ 0.45\ 0.45\ 0.45]$, $p_{11}^{(1)} = [0.8\ 0.8\ 0.6\ 0.6\ 0.6]$, and $p_{11}^{(2)} = [0.7\ 0.7\ 0.5\ 0.5\ 0.5]$).
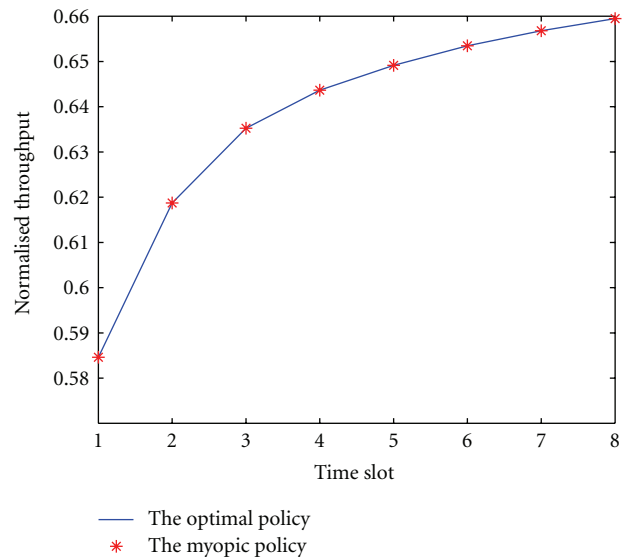

FIGURE 6: The performance of the myopic policy with sensor errors (Prob. of false alarm = 0.2, Prob. of miss detection = 0.3, $N = 3$, $L = 2$, $p_{01}^{(1)} = [0.05\ 0.05\ 0.05]$, $p_{01}^{(2)} = [0.3\ 0.3\ 0.3]$, $p_{11}^{(1)} = [0.95\ 0.95\ 0.95]$, and $p_{11}^{(2)} = [0.65\ 0.65\ 0.65]$).

sensing policy for channel selection, which is reduced to an unconstraint POMDP problem as addressed in this paper. Under this design dictated by the separation principle, we observe from Figure 6 that the myopic sensing policy can still achieve the optimal performance for independent and stochastically identical channels even when there are sensor errors.

6. Conclusion

In this paper, we have considered the multichannel opportunistic access in self-similar primary traffic processes. Under the assumption that the states of the Markov process are positively correlated at each level and initially independent across all levels for each channel, we have shown that for independent and stochastically identical channels when the initial system belief vector satisfies the channel ordering condition as stated in Theorem 1, the myopic policy has a simple and robust structure with strong performance. Future work includes investigating the optimality and throughput limits of the myopic policy for independent and stochastically identical channels, and extending the simple structure of the myopic policy to nonidentical channels.

Acknowledgments

References

[1] FCC Spectrum Policy Task Force, "Report of the spectrum efficiency," working group, November 2002.

[2] S. Geirhofer, L. Tong, and B. M. Sadler, "Dynamic spectrum access in the time domain: modeling and exploiting white space," *IEEE Communications Magazine*, vol. 45, no. 5, pp. 66–72, 2007.

[3] FCC03-322, "Notice of proposed rule making: facilitating opportunities for flexible, efficient, and reliable spectrum use employing cognitive radio technologies and authorization and use of software defined radios," December 2003.

[4] W. Leland, M. Taqqu, W. Willinger, and D. Wilson, "On the self-similar nature of ethernet traffic," in *Proceedings of the ACM International Conference of the Special Interest Group on Data Communication (SIGCOMM '93)*, pp. 183–193, San Francisco, Calif, USA, September 1993.

[5] K. Park and W. Willinger, *Self-Similar Network Traffic and Performance Evaluation*, John Wiley & Sons, New York, NY, USA, 2000.

[6] M. Jiang, M. Nikolic, S. Hardy, and L. Trajkovic, "Impact of self-similarity on wireless data network performance," in *Proceedings of the IEEE International Communications Conference (ICC '01)*, vol. 2, pp. 477–481, June 2001.

[7] D. Radev and I. Lokshina, "Self-similar simulation of IP traffic for wireless networks," *International Journal of Mobile Network Design and Innovation*, vol. 2, pp. 202–208, 2007.

[8] Q. Liang, "Ad hoc wireless network traffic-self-similarity and forecasting," *IEEE Communications Letters*, vol. 6, no. 7, pp. 297–299, 2002.

[9] V. Misra and W.-B. Gong, "A hierarchical model for teletraffic," in *Proceedings of the 37th IEEE Conference on Decision and Control (CDC '98)*, vol. 2, pp. 1674–1679, Tampa, Fla, USA, 1998.

[10] W. Gong, Y. Liu, V. Misra, and D. Towsley, "Self-similarity and long range dependence on the internet: a second look at the evidence, origins and implications," *Computer Networks*, vol. 48, no. 3, pp. 377–399, 2005.

[11] Q. Zhao, L. Tong, and A. Swami, "Decentralized cognitive MAC for dynamic spectrum access," in *Proceedings of the 1st IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN '05)*, pp. 224–232, Baltimore, Md, USA, November 2005.

[12] Q. Zhao, L. Tong, A. Swami, and Y. Chen, "Decentralized cognitive MAC for opportunistic spectrum access in ad hoc networks: a POMDP framework," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 3, pp. 589–599, 2007.

[13] Q. Zhao and A. Swami, "A decision-theoretic framework for opportunistic spectrum access," *IEEE Wireless Communications*, vol. 14, no. 4, pp. 14–20, 2007.

[14] Y. Chen, Q. Zhao, and A. Swami, "Joint design and separation principle for opportunistic spectrum access," in *Proceedings of IEEE Asilomar Conference on Signals, Systems, and Computers (ASLIOMAR '06)*, pp. 696–700, Pacific Grove, Calif, USA, October-November 2006.

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

[16] Q. Zhao, B. Krishnamachari, and K. Liu, "On myopic sensing for multi-channel opportunistic access: structure, optimality, and performance," *IEEE Transactions on Wireless Communications*, vol. 7, no. 12, pp. 5431–5440, 2008.

[17] S. H. Ahmad, M. Liu, T. Javidi, Q. Zhao, and B. Krishnamachari, "Optimality of myopic sensing in multi-channel opportunistic access," to appear in *IEEE Transactions on Information Theory*.

[18] O. Sheluhin, S. Smolskiy,, and A. Osin, *Self-Similar Processes in Telecommunications*, John Wiley & Sons, New York, NY, USA, 2007.