

Optimal Dynamic Spectrum Access via Periodic Channel Sensing

Qianchuan Zhao,¹ Stefan Geirhofer,² Lang Tong,² and Brian M. Sadler³

¹Department of Automation, Tsinghua University, Beijing, 100084, China

²School of Electrical and Computer Engineering, Cornell University, Ithaca, NY 14853, USA

³Army Research Laboratory, Adelphi, MD 20783, USA

Abstract—The problem of dynamically accessing a set of parallel channels occupied by primary users is considered. The secondary user is allowed to sense and to transmit in a single channel. By exploiting idle periods between bursty transmissions of primary users, and by using a periodic sensing strategy, optimal dynamic access is achieved by maximizing the throughput of the secondary user while constraining collision probability with the primary user. The optimal dynamic spectrum access problem can then be formulated within the framework of Constrained Markov Decision Processes (CMDPs). The optimal control policy is identified via a linear program, and its performance is analyzed numerically and through Monte Carlo simulations. Finally, we compare the optimal scheme to an ideal benchmark case when simultaneous sensing of all channels is assumed.

I. INTRODUCTION

The wireless communications spectrum has densely been allocated by regulators as of today making it unlikely that the bandwidth needs of emerging technologies can be satisfied. At the same time, however, recent measurement campaigns [1] illustrate that it is the regulator's static frequency allocation rather than the spectrum's heavy utilization that leads to this scarcity. The emerging area of dynamic spectrum access aims at resolving this contrast by allowing so-called secondary users to transmit in assigned bands, provided that no (significant) amount of interference is generated.

In order to meet this constraint secondary users can exploit various degrees of freedom to achieve orthogonality with the primary system. This paper focuses on dynamic spectrum access in the time domain [2], exploiting the fact that many of the wireless communications systems in use today show a bursty transmission behavior. Measurement results have shown that the spectrum is generally only lightly used making it worthwhile to exploit the idle periods between consecutive packet transmissions.

When it comes to deriving access strategies for the secondary system, a model for the primary user's medium access is indispensable [3]. Moreover, a suitable balance between

prediction accuracy and computational complexity needs to be struck. In the scope of this paper we shall assume that the primary user can be modeled by a two-state Markov process consisting of a 'busy' and 'idle' state. The process evolves continuously in time and, by the Markov property, has exponentially distributed holding times in either state.

A. Main contribution

We assume that multiple primary user channels evolve independently in time (for instance, consider multiple non-overlapping WLAN channels). While all of these channels can be used for dynamic spectrum access, hardware limitations make it possible to sense and transmit in one channel at a time (although transmitting and sensing can be performed in different channels). This partial observability makes the problem analytically challenging. We counter this intractability by introducing a periodic sensing scheme. In this way we can recast the problem as a constrained Markov decision process (CMDP), which has been well-studied [4]. We show how to use linear programming techniques to find the optimal stationary randomized policy for this problem and verify the correctness of our result by Monte-Carlo simulations. Finally, we numerically compare the performance to heuristic schemes in order to assess the significance of the periodic sensing postulation.

B. Related Work

One class of DSA schemes assumes a hierarchy of primary and secondary users where the secondary users are allowed to access the spectrum when they do not interfere with or have limited interference on the transmissions of primary users [5]. This can be accomplished in the time domain by a sense-then-transmit strategy as first proposed in [2]. Under the assumption that the primary users follow a slotted transmission protocol, and their traffic pattern can be modeled as a discrete-time Markov chain, Zhao *et al.* derive the optimal access protocol based on the Partially Observable Markov Decision Process (POMDP) [2], [6].

In addition to the complexity of POMDP, the assumption of slotted transmissions, however, may not be valid for some practical systems. Indeed, wireless LAN transmissions do not have a slot structure. Experimental results show that WLAN traffic is modeled more closely by a continuous-time semi-Markov process [7], [8]. This paper removes the assumption of

¹This paper was prepared through collaborative participation in the Communications and Networks Consortium sponsored by the U.S. Army Research Laboratory under the Collaborative Technology Alliance Program, Cooperative Agreement DAAD19-01-2-0011. The work was done when Qianchuan Zhao was visiting Cornell University as a visiting Professor. Part of the research of Qianchuan Zhao was supported by a faculty exchange program between Tsinghua University and Cornell University and NSFC grant no. 60574067. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation thereon.

slotted transmission by the primary user and assumes instead a continuous-time Markov process for the primary user's behavior.

Finally, from a practical viewpoint, a sensing-based real-time implementation for sharing the spectrum with WLAN channels has been described in [9]. The testbed focuses on heuristic algorithms to share the spectrum with the WLAN.

II. SYSTEM MODEL

Assume that we can access one of the N primary user channels that evolve independently in time. Furthermore, let each of these channels $\{X_i(t), 0 \leq i \leq N-1\}$ be represented by a time-homogeneous continuous-time Markov process with an idle ($X_i(t) = 0$) and a busy ($X_i(t) = 1$) state, respectively. As a consequence of the Markovian assumption the holding times in either state are exponentially distributed with parameters λ_i (idle) and μ_i (busy), respectively. We stress that the primary system is *not* slotted; primary users can access the channel at any given time.

In contrast to the primary user, the secondary user employs a slotted communication protocol as depicted in Fig. 1 (consider Bluetooth as a practical example). In each slot the secondary user (i) senses the channel at the beginning of the slot, (ii) uses this sensing result to decide if and in which channel to transmit, and (iii) possibly receives an acknowledgement by the secondary receiver.

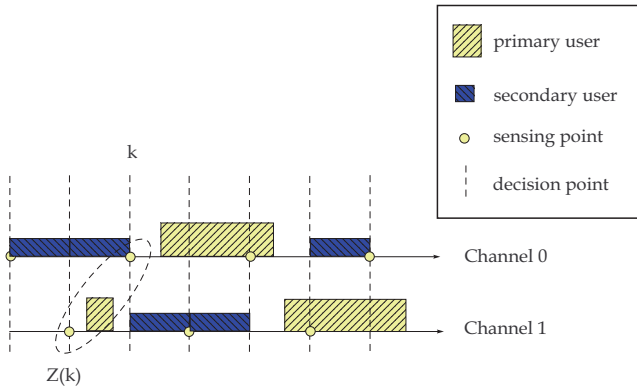


Fig. 1. Slot structure.

A block diagram of the system is shown in Fig. 2. The signal captured by the antenna is passed through an analog frontend and sampled within the sensing block. A decision is made on whether the primary user is present and this sensing result is passed on to a controller that decides whether it is safe to transmit (and if yes, in which channel). If a transmission occurs the secondary user's data is fed to the transmit modem, which in turn interfaces the analog frontend.

There are several approaches to ensure synchronization among secondary transceivers. Provided that all secondary users share the same observation of the channel, using the same random seed within the controller ensures that the secondary system will tune to the same channel. Additionally, ac-

knowledgements and signaling information can be multiplexed with data to ensure synchronization in periodic intervals.

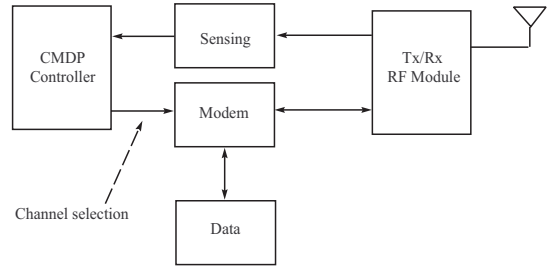


Fig. 2. System block diagram.

III. PERIODIC SENSING APPROACH

Hardware limitations restrict the secondary user to sense only one of the N channels in every slot. This limitation makes the problem of finding an optimal access strategy challenging, since the state of the system at any time is only partially observed. In this paper, we render the problem tractable by postulating a periodic sensing approach, thus separating the sensing and the access part of the problem. While imposing a periodic sensing strategy is in general suboptimal, it leads to a fully observable Markov Decision Process, thus simplifying the problem considerably.

A. Induced Markov chain

We propose to sense the channel periodically and in increasing order, i.e., at the beginning of the k -th slot $I_k = [kT_s, (k+1)T_s)$, channel $q = k \bmod N$ is sensed, where T_s denotes the slot size, and 'mod' denotes the modulus operation. Let the vector $\mathbf{Z}(k) = [Z_1(k), \dots, Z_N(k)]$ represent the last N sensing results from all channels with $Z_i(k) = 0$ and $Z_i(k) = 1$ indicating an idle and busy channel, respectively. The state space of $\mathbf{Z}(k)$ is thus $\mathcal{S} = \{0, 1\}^N$.

More specifically, $\mathbf{Z}(k)$ is a stochastic process indexed by $k \geq N-1$ such that

$$Z_i(k) = \begin{cases} X_i((pN+i)T_s), & i \leq q, \\ X_i(((p-1)N+i)T_s), & i > q \end{cases}$$

where $p = \lfloor \frac{k}{N} \rfloor$ is the integer part of $\frac{k}{N}$.

It is straightforward to show that $\mathbf{Z}(k)$ forms a discrete-time Markov chain. Furthermore, $\mathbf{Z}(k)$ is irreducible and periodic with period N , since every state is recurrent and $\Pr(\mathbf{Z}(k+1)|\mathbf{Z}(k))$ depends only on $q = k \bmod N$ but not on $\lfloor \frac{k}{N} \rfloor$.

The steady-state probabilities of $\mathbf{Z}(k)$ are given by

$$f_q(\mathbf{z}) = \lim_{P \rightarrow \infty} \frac{1}{P} f_q^P(\mathbf{z}), \quad (1)$$

where $f_q^P(\mathbf{z})$ represents the number of times \mathbf{z} appears in the sequence $\{\mathbf{Z}(pN+q), p = 1, \dots, P\}$. The existence of (1) is guaranteed for all q and $\mathbf{z} \in \mathcal{S}$.

Since the primary user processes X_i are independent, the stationary distribution of process $\mathbf{Z}(k)$ has product form,

$$f_q(\mathbf{z}) = \prod_{i=0}^{N-1} (1_{[z_i=0]} f_q(0) + 1_{[z_i=1]} f_q(1)), \quad (2)$$

where $1_{[\cdot]}$ denotes the indicator function and

$$f_q(0) = \frac{\mu_i}{\lambda_i + \mu_i}, \quad f_q(1) = \frac{\lambda_i}{\lambda_i + \mu_i}. \quad (3)$$

B. Actions and rewards

Having characterized the Markov chain, which is induced by the primary system and the adopted slot structure for the secondary user, we need to add a control dimension to our problem. Specifically, after each sensing operation, we can either choose to transmit in one of the N channels, or alternatively not transmit at all. Let the action chosen in slot k under policy π be denoted as $a_k \in \mathcal{A} = \{-1, 0, \dots, N-1\}$. Choosing $a_k \geq 0$ symbolizes to transmit in the a_k -th channel, whereas $a_k = -1$ means that no transmission is taking place.

If we choose to transmit, we **accrue** a reward or **incur** a cost depending on whether the packet transmission is successful or is leading to a collision with the primary user. It is stressed that even if a channel has just been sensed idle, a collision can occur in the subsequent slot since the primary user's medium access is not slotted.

Let us define the **reward** $r(\mathbf{z}(k), a, k)$ accrued by a successful transmission as

$$r(\mathbf{z}(k), a, k) = \begin{cases} \Pr(X_i(t) = 0, \forall t \in I_k | Z_i(k) = z_i), & a \geq 0 \\ 0, & a = -1 \end{cases} \quad (4)$$

which corresponds to the secondary user's transmission being successful. **If the secondary user chooses not to transmit, $a_k = -1$, then no reward is accrued.**

Analogously, we can define the **cost** of choosing action $a_k \geq 0$ as the probability that the transmission leads to a collision with the primary user. Mathematically,

$$c(\mathbf{z}(k), a, k) = \begin{cases} 1 - r(\mathbf{z}(k), a, k), & a \geq 0 \\ 0, & a = -1 \end{cases} \quad (5)$$

Let the Q -matrix for each channel be Q_i , then we have

$$Q_i = \begin{pmatrix} -\lambda_i & \lambda_i \\ \mu_i & -\mu_i \end{pmatrix} \quad (6)$$

and the matrix $e^{Q_i t}$ evaluates to

$$e^{Q_i t} = \begin{bmatrix} 1 - \frac{\lambda_i}{\lambda_i + \mu_i} (1 - e^{-(\lambda_i + \mu_i)t}) & \frac{\lambda_i}{\lambda_i + \mu_i} (1 - e^{-(\lambda_i + \mu_i)t}) \\ \frac{\mu_i}{\lambda_i + \mu_i} (1 - e^{-(\lambda_i + \mu_i)t}) & 1 - \frac{\mu_i}{\lambda_i + \mu_i} (1 - e^{-(\lambda_i + \mu_i)t}) \end{bmatrix}. \quad (7)$$

An analytical expression for the reward is obtained by

$$r(\mathbf{z}(k), i, k) = \left[e^{Q_i \times (q-i)T_s} \right]_{(z_i(k), 0)} e^{-\lambda_i T_s}, \quad (8)$$

where the subscript notation indicates the (i, j) -th element of $e^{Q_i t}$.

If we introduce a table g indexed by \mathbf{z}, q, i ,

$$g(\mathbf{z}, q, i) = \left(\frac{\mu_i}{\lambda_i + \mu_i} + (1_{[z_i=0]} \frac{\lambda_i}{\lambda_i + \mu_i} + 1_{[z_i=1]} \frac{\mu_i}{\lambda_i + \mu_i}) e^{-(\lambda_i + \mu_i)((N+q-i) \bmod N) T_s} \right) e^{-\lambda_i T_s} \quad (9)$$

for $i \geq 0$ and zero otherwise, the immediate reward and cost in k -th slot can be analytically evaluated by

$$r(\mathbf{z}(k), a, k) = \begin{cases} g(\mathbf{z}(k), k \bmod N, a), & a \geq 0 \\ 0, & a = -1 \end{cases} \quad (10)$$

and

$$c(\mathbf{z}(k), a, k) = \begin{cases} 1 - g(\mathbf{z}(k), k \bmod N, a), & a \geq 0 \\ 0, & a = -1 \end{cases} \quad (11)$$

It is worthwhile to note the special case when $q = i$. In this case we have

$$g(\mathbf{z}, q, q) = 1_{[z_q=0]} e^{-\lambda_q T_s}. \quad (12)$$

That is, when $z_q = 0$ and we transmit on Channel q , the immediate reward will be $e^{-\lambda_q T_s}$; when $z_q = 1$ and we transmit on Channel q , no reward will be obtained.

IV. SUBOPTIMAL STATIC ACCESS PROTOCOLS

Under periodic sensing, with the analytical expressions (10) and (11) given in Section III for the immediate reward and collision probability, we introduce two simple heuristic protocols that are easy to implement. They can be used for **comparisons as lower bounds of the achievable throughput under constraints on collision** with primary users.

A. Memoryless Access (MA)

Under periodic sensing, if in k -th slot, the secondary user senses a busy channel $q = k \bmod N$, no transmission is made; otherwise the user transmits in channel q with probability β_q^{MA} . The transmission probability β_q^{MA} is decided such that the probability of collision in k -th slot is below α while maximize the throughput for the secondary user. Denote this heuristic policy as π^{MA} . It is straightforward to show that the transmission probability

$$\beta_q^{\text{MA}} = \min \left(\frac{\alpha}{1 - e^{-\lambda_q T_s}}, 1 \right), \quad (13)$$

and the throughput of this policy is

$$J(\pi^{\text{MA}}) = \frac{1}{N} \sum_{q=0}^{N-1} f_q(0) \min \left(\frac{\alpha}{1 - e^{-\lambda_q T_s}}, 1 \right) e^{-\lambda_q T_s}, \quad (14)$$

where $f_q(0) = \frac{\mu_q}{\lambda_q + \mu_q}$ is the stationary probability for Channel q to be idle.

B. Greedy Access (GA)

Given $\mathbf{Z}(k) = \mathbf{z}$ and sensing channel $q = k \bmod N$, compute the a posteriori probability

$$g(\mathbf{z}, q, i) = \Pr(X_i(t) = 0, \forall t \in I_k | \mathbf{Z}(k) = \mathbf{z}) \quad (15)$$

in each channel i being idle in slot I_k . Choose the channel $i^* = \arg \max_i g(\mathbf{z}, q, i)$ which is most likely idle. Transmit in Channel i^* with probability $\beta_q^{\text{GA}}(\mathbf{z})$ such that the collision probability $c(\mathbf{z}, q, i^*)$ in slot I_k is below α . Denote this heuristic policy as π^{GA} . It is easy to show that the transmission probability is

$$\beta_q^{\text{GA}}(\mathbf{z}) = \min \left(\frac{\alpha}{1 - \max_i g(\mathbf{z}, q, i)}, 1 \right) \quad (16)$$

and the throughput of this policy is

$$J(\pi^{\text{GA}}) = \frac{1}{N} \sum_{q=0}^{N-1} \sum_{\mathbf{z} \in S} f_q(\mathbf{z}) \beta_q^{\text{GA}}(\mathbf{z}) \max_i g(\mathbf{z}, q, i) \quad (17)$$

where $\beta_q^{\text{GA}}(\mathbf{z})$ is as determined in (16). This strategy is similar to the greedy approach in [2].

V. OPTIMAL POLICY WITH PERIODIC SENSING

We have defined an induced Markov chain modeling the observations of the secondary system and have introduced the control dimension by specifying actions and rewards. We have also introduced two simple heuristics as policies to control the secondary user transmissions. Although they are easy to implement, they may not make the optimal use of the spectrum under periodic sensing. In order to find optimal policies, we need to define an optimization criterion and constraints.

A. Optimization criterion

We ultimately strive to maximize the throughput of the secondary system while abiding by a hard constraint on the interference. Mathematically, we can formulate this goal as maximizing the average number of successful transmissions (of the secondary user),

$$J(\pi) = \lim_{K \rightarrow \infty} \frac{1}{K} \sum_{k=1}^K E_{\pi} r(\mathbf{Z}(k), a_k, k), \quad (18)$$

where the expectation is taken over the probability distribution induced by a policy π . At the same time we have to abide by the constraint on a cost function

$$C(\pi) = \lim_{K \rightarrow \infty} \frac{1}{K} \sum_{k=1}^K E_{\pi} c(\mathbf{Z}(k), a_k, k). \quad (19)$$

The stochastic optimization problem is thus

$$\max_{\pi} J(\pi) \quad (20)$$

subject to

$$C(\pi) \leq \alpha \quad (21)$$

where $0 \leq \alpha \leq 1$ is a given constant. The problem thus falls into the category of constrained Markov decision processes (CMDPs) [4], [10] and can be solved by a linear program.

Note, however, that since the reward and cost are periodic, an extension of the standard CMDP theory is needed. It is well known that the optimal solution to a CMDP is, in general, randomized. The policy π is thus represented by a mapping from the set of observations \mathbf{z} and q to the probability that we choose action i . In the following let the probability that we choose action $i \geq 0$ based on \mathbf{z} and q be denoted by $\beta_{q,i}(\mathbf{z})$. No transmission takes place with probability $\beta_{q,-1}(\mathbf{z}) = 1 - \sum_{i=0}^{N-1} \beta_{q,i}(\mathbf{z})$.

Notice that our problem is a special type of CMDP in the sense that the underlying Markov chain $\mathbf{Z}(k)$ is not affected by the actions chosen by the decision maker. As a CMDP, it is special also because the rewards $r(\mathbf{z}(k), a, k)$ and costs $c(\mathbf{z}(k), a, k)$ at each k are not time independent, instead, they are periodic.

B. Linear programming solution

We introduce a linear program that has the optimal randomized policy $\beta_{q,i}(\mathbf{z})$ as its solution. Consider,

$$\max_{\beta} \frac{1}{N} \sum_{q=0}^{N-1} \sum_{\mathbf{z} \in S} f_q(\mathbf{z}) \sum_{i=0}^{N-1} g(\mathbf{z}, q, i) \beta_{q,i}(\mathbf{z}) \quad (22)$$

subject to

$$\frac{1}{N} \sum_{q=0}^{N-1} \sum_{\mathbf{z} \in S} f_q(\mathbf{z}) \sum_{i=0}^{N-1} (1 - g(\mathbf{z}, q, i)) \beta_{q,i}(\mathbf{z}) \leq \alpha, \quad (23)$$

$$\sum_{i \in A} \beta_{q,i}(\mathbf{z}) = 1, \forall q, \mathbf{z}, \quad \beta_{q,i}(\mathbf{z}) \in [0, 1], \forall q, \mathbf{z}, i. \quad (24)$$

then using a similar argument as in [10] it can be shown that the solution to the above linear program indeed corresponds to the optimal policy of the CMDP (20)-(21). To explain why this is true intuitively, let us consider a fixed optimal policy π^* . If we classify transmissions according to sensing channel q , the objective function can be written in form of

$$J(\pi^*) = \lim_{P \rightarrow \infty} \frac{1}{PN + q} \sum_{p=1}^P \sum_{q=0}^{N-1} E_{\pi^*} r(\mathbf{Z}(pN+q), a_{pN+q}, pN+q). \quad (25)$$

Let us introduce $\beta_{q,i}^{\pi^*,P}(z) \in [0, 1]$ as the frequency of action $i \in \mathcal{A}$ chosen by π^* in slot I_{pN+q} when the observed value of $\mathbf{Z}(pN+q)$ equals to $\mathbf{z} \in S$ in a sample path with $p = 1, 2, \dots, P$. Since both the state space S and the action space \mathcal{A} are finite, the limit $\lim_{P \rightarrow \infty} \beta_{q,i}^{\pi^*,P}(\mathbf{z})$ exists. Let us denote it as $\beta_q^{\pi^*}(\mathbf{z})$. Since the sensing results on primary users are not affected by the transmission policy of the secondary user, the asymptotical transmission rate given a sensing channel q can then be calculated as the expected rate of successful transmission in stationary state, that is

$$\begin{aligned} \lim_{P \rightarrow \infty} \frac{1}{P} \sum_{p=1}^P E_{\pi^*} r(\mathbf{Z}(pN+q), a_{pN+q}, pN+q) \\ = \sum_{\mathbf{z} \in S} f_q(\mathbf{z}) \sum_{a \in \mathcal{A}} g(\mathbf{z}, q, a) \beta_q^{\pi^*}(\mathbf{z}) \end{aligned} \quad (26)$$

Thus we have

$$J(\pi^*) = \sum_{q=0}^{N-1} \frac{1}{N} \sum_{\mathbf{z} \in S} f_q(\mathbf{z}) \sum_{a \in \mathcal{A}} g(\mathbf{z}, q, a) \beta_{q,a}^{\pi^*}(\mathbf{z}). \quad (27)$$

This shows the equivalence between objective functions (20) and (22) since $g(\mathbf{z}, q, -1) = 0$. Similarly, we can show the equivalence between constraints (21) and (23).

Once the solution $\beta = (\beta_{q,i}(\mathbf{z}), i \in \mathcal{A}, q \in \{0, 1, \dots, N-1\}, \mathbf{z} \in S)$ has been obtained, the secondary user stores it as a table in his(her) memory. The secondary user's policy given the observations \mathbf{z} and q is to flip a biased coin such that with probability $\beta_{k \bmod N, i}(\mathbf{Z}(k))$ we transmit in Channel i and with probability $\beta_{k \bmod N, -1}(\mathbf{Z}(k))$ no transmission occurs.

VI. NUMERICAL EVALUATION

In this section we evaluate the performance of the optimal policy numerically. Specifically, we focus on the case of three independent channels, with parameters $\lambda = \lambda_i$ and $\mu = \mu_i, i = 1, 2, 3$.

The choice of λ and μ was motivated from experiments conducted in [7]. In particular, the parameters were chosen based on a "Skype" conference call session with three participating parties. The idle-times, although showing some heavy-tailed behavior, can be approximated by an exponential distribution with parameter $\lambda^{-1} = 4.2$ ms. We assume $\mu^{-1} = 1$ ms for the channel's busy period.

To evaluate the optimal policy under periodic sensing, we compare it with two classes of policies.

(1) the optimal policy under full observation assumption (that is, all channels can be sensed simultaneously). While this assumption is not a reasonable assumption in practice, with its throughput as an upper bound, it allows us to determine how much we loose by restricting ourselves to schemes that employ periodic sensing.

(2) the suboptimal static access protocols MA and GA introduced in Section IV which are also under periodic sensing. They are simpler to implement.

By assuming a slot size $T_s = 0.25$ ms, we obtain in Fig. 3 the throughput curves for all four policies (FO, PS, MA, and GA) over the interval $[0, 0.05]$ of the constraint constant α .

We observe that the optimal policy with periodic sensing (PS) achieves the same throughput as the throughput of the optimal policy with full observation (FO) in the region $\alpha \in [0, 0.045]$ and a bit lower in throughput than the full observation in region $\alpha \in (0.045, 0.05]$. The throughput of GA is greater than that of MA, but both GA and MA are not as good as that of PS and FO.

VII. CONCLUSION

In conclusion, we have considered the problem of sharing spectrum in the time domain by exploiting idle periods between bursty transmissions of a primary user. By focusing on a periodic sensing scheme we were able to formulate the problem as a constrained Markov decision process (CMDP), and find the optimal randomized control policy using a linear programming technique. We have evaluated the method's

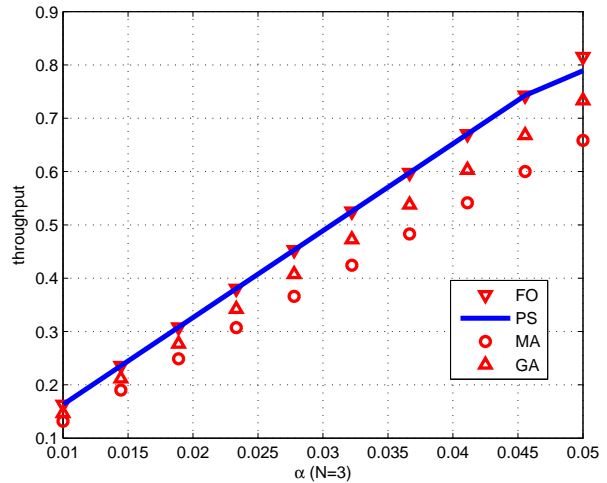


Fig. 3. Comparison among different protocols.

performance for several scenarios. Our results show that the periodic sensing, while limiting the set of admissible policies, is close to optimal in the cases we have considered so far. We also evaluated the robustness of our solution in terms of primary user parameters and the sensing errors. The results are encouraging. Details can be found in our technique report. Future research includes the case of more than two secondary users.

REFERENCES

- [1] R. Tandra and A. Sahai, "Fundamental limits on detection in low SNR under noise uncertainty," in *Proc. International Conference on Wireless Networks, Communications and Mobile Computing*, 2005.
- [2] Q. Zhao, L. Tong, and A. Swami, "Decentralized Cognitive MAC for Dynamic Spectrum Access," in *Proc. First IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks*, Nov. 2005, pp. 224–232.
- [3] S. Geirhofer, L. Tong, and B. M. Sadler, "A Measurement-Based Model for Dynamic Spectrum Access," in *Proc. IEEE Conference on Military Communications (MILCOM)*, 2006.
- [4] E. Altman, *Constrained Markov Decision Processes*. Chapman & Hall/CRC, 1999.
- [5] Q. Zhao and B. M. Sadler, "Dynamic Spectrum Access: Signal Processing, Networking, and Regulatory Policy," *submitted to the IEEE Signal Processing Mag.*, 2006.
- [6] Q. Zhao, L. Tong, A. Swami, and Y. Chen, "Decentralized Cognitive MAC for Opportunistic Spectrum Access in Ad Hoc Networks: A POMDP Framework," *submitted to IEEE J. Select. Areas Commun.*, Feb. 2006.
- [7] S. Geirhofer, L. Tong, and B. M. Sadler, "Dynamic Spectrum Access in WLAN Channels: Empirical Model and Its Stochastic Analysis," in *Proc. First International Workshop on Technology and Policy for Accessing Spectrum*, 2006.
- [8] —, "Dynamic Spectrum Access in the Time Domain: Modeling and Exploiting Whitespace," *submitted to the IEEE Communications Magazine*, 2006.
- [9] S. D. Jones, N. Merheb, and I.-J. Wang, "An experiment for sensing-based opportunistic spectrum access in CSMA/CA networks," in *First IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks*, Nov. 2005, pp. 593–596.
- [10] M. L. Puterman, *Markov Decision Processes. Discrete Stochastic Dynamic Programming*. John Wiley & Sons, Inc., 1994.