DECENTRALIZED COGNITIVE MAC FOR DYNAMIC SPECTRUM ACCESS

Qing Zhao, Lang Tong, and Ananthram Swami

ABSTRACT

We consider the problem of opportunistic dynamic spectrum access (DSA) in an ad hoc network in which unlicensed secondary users communicate through channels not used by the primary users. Decentralized cognitive medium access control protocols are presented that allow secondary users to recognize spectrum opportunity and transmit based on a partial observation of the instantaneous spectrum availability. Under a framework of Partially Observable Markov Decision Process (POMDP), we derive optimal and suboptimal decentralized strategies for the secondary users to decide which channel(s) to sense and access for the maximization of the overall network throughput.

1. INTRODUCTION

Static spectrum allocation strategies effectively bypass the problems of spectrum coordination and, with adequate guard bands, avoid interference. Such fixed allocations, however, can be wasteful when the primary user has no data to transmit or may result in unacceptable drops and delays if the user's demand is too high. Dynamic Spectrum Access (DSA) [1] represents a new paradigm of spectrum management, a shift from static allocation to dynamic access, allowing opportunistic communications based on user demands and channel availability. DSA will become increasingly important as overlay schemes and the use of unlicensed spectra increase. DSA is also critical in coping with traffic load variations over time and space and for heterogeneous networks to coexist without explicit coordination.

1.1. Dynamic Spectrum Access

Two approaches to DSA have been envisioned: dynamic spectrum allocation and opportunistic spectrum access. While sharing certain common features, these two approaches differ in their rationale, technological challenges, and domains of applications. Dynamic spectrum allocation brought forth by the European DRiVE project [2] mainly focuses on long-term commercial applications such as UMTS and DVB-T. By exploiting temporal and spatial traffic statistics, it aims to improve spectrum efficiency through time- and space-dependent spectrum sharing among coexisting radio services. For example, the amount of spectrum allocated to UMTS and DVB-T varies over region and the-time-of-day. Similar to the current static spectrum allotment policy, such DSA strategies allocate, at a given time and region, a portion of the spectrum to a radio access network for its exclusive use. As such, white space in spectrum due to bursty traffic cannot be eliminated.

Different from dynamic spectrum allocation which uses the statistics of spectrum occupancy, opportunistic spectrum access envisioned by the DARPA XG program [3] aims to exploit the instantaneous spectrum availability by opening licensed spectrum to secondary users. The idea is to allow secondary users to identify available spectrum resources and communicate opportunistically in a manner that limits the level of interference perceived by primary users. It thus has the potential of eliminating white space in the spectrum. Such DSA strategies are more relevant to applications requiring rapid but short-term deployment and applications denied of cooperation from existing radio access networks. Examples include military units penetrating deep in unknown and/or hostile territories, wireless networks established for particular social events, or sensor networks deployed for specific tasks. Requiring little cooperation from the spectrum licensees, opportunistic spectrum access can be overlayed with the current static allotment policy as well as the envisioned dynamic spectrum allocation. Besides software defined radio, the technological underpinning of opportunistic spectrum access includes efficient spectrum sensing for opportunity identification and adaptive medium access and networking protocols for opportunity utilization.

1.2. Decentralized Cognitive MAC

In this paper, we focus on DSA of the second kind in ad hoc networks: opportunistic spectrum access based on instantaneous network state. One of the most crucial and difficult challenges in such networks is the design of cognitive medium access control (MAC) that recognizes and utilizes spectrum opportunities for optimal network performance. For ad hoc networks without a central authority, it is desirable to have a Decentralized Cognitive MAC (DC-MAC) where each node decides individually how to sense the spectrum and how to gain access. Because users cannot exchange local information on channel availability before agreeing on a

Q. Zhao is with the Department of Electrical and Computer Engineering, University of California, Davis, CA 95616, qzhao@ece.ucdavis.edu

L. Tong is with the School of Electrical and Computer Engineering, Cornell University, Ithaca, NY 15853, ltong@ece.cornell.edu

A. Swami is with the Army Research Laboratory, Adelphi, MD 20783, aswami@arl.army.mil

communication channel, such a protocol should not rely on cooperation among secondary users. Thus the design of DC-MAC for DSA networks is more challenging than that for standard ad hoc networks.

We focus on cognitive MAC where each node must sense the channel intelligently by exploiting statistical traffic behavior. We do not assume that each secondary node has perfect knowledge of the availability of all channels; such knowledge implies a full spectrum sensing synchronous among users. We assume instead that each node can choose to sense a subset of the possible channels and must decide if transmission is possible based on the sensing outcome. As such, each node observes only a partial, not the full, state of the network. Furthermore, we allow sensing errors: the overlook of an available channel and the mistake of identifying an unavailable channel as an opportunity.

The key step in the design of DSA is to capture the dynamic behavior of the spectral vacancy which fundamentally determines the network performance. The modelling of DSA dynamics needs to incorporate channel availability, channel bandwidth, and traffic pattern. The last plays a crucial role in protocol design. For example, if we know that a primary user favors a particular channel and tends to occupy it for a long period of time, that channel would be less likely available for a secondary user, and sensing it would likely be a waste of time and energy.

There is also a need to consider sensing and access jointly, which leads to a cross-layer design for DSA. For example, if a node has no packet to send, does it make sense to perform channel sensing? Sensing costs energy but gains information about the network state. Sensing proactively better prepares the node for transmission but at the cost of energy consumption. What would then be the tradeoff between energy consumption and spectrum utilization?

1.3. Scope, Contribution, and Related Work

Scope   The scope of this paper is limited to the technical (and more analytical) aspects of the MAC design for DSA ad hoc networks. Specifically, we focus on the theoretical formulation and characterization of decentralized medium access and the development of optimal DC-MAC protocols.

While we are not concerned about specific implementation details (*e.g.,* packet format and bit definitions), we are concerned about implementation complexity, both in computation and storage. We take an analytical approach, imposing certain idealized modelling assumptions. It is our hope, however, that the results presented here can provide insights into the design of more complicated DSA networks under more realistic assumptions.

Contribution   The contribution of this paper is threefold. We provide (i) an analytical framework of DC-MAC in DSA ad hoc networks; (ii) a characterization of the optimal protocol; (iii) the development of a low complexity suboptimal

greedy algorithm.

Analytical Framework   We present an analytical framework for the design of DC-MAC for DSA ad hoc networks. This framework includes three components. Jointly, they define protocols that integrate channel sensing and access.

The first component is a channel occupancy model that captures the dynamics of channel availability. By using a Markov chain formulation, we incorporate traffic characteristics of the primary and secondary users. For example, given that a channel is currently occupied by a primary user, the probability that the primary user will need it for the next slot is modelled by the state transition probability.

The second is the performance metric, the objective function that defines the optimal strategy. In this paper, we focus on maximizing spectrum utilization by designing a DC-MAC that maximizes the average throughput. The formulation, however, can be easily tailored to incorporate energy consumptions by imposing constraints or penalty.

The third component is a decision theoretic approach to selecting which channel to sense and access given the node's past sensing history, channel occupancy statistics, and the reward-cost of sensing and transmission.

Optimal DC-MAC   We present next an optimization framework based on the theory of Partially Observable Markov Decision Process (POMDP). The network state is partially observable due to a more practical and more general sensing model that allows a user to sense, not all, but a subset of channels. The structure of the optimal DC-MAC is obtained following the classical work of Smallwood and Sondik [4]. Our formulation is also amenable to other POMDP techniques.

Suboptimal Greedy DC-MAC   The optimal DC-MAC requires the update and storage of an information vector with a dimension exponentially growing with the number of channels. We show that the required sufficient statistic has a dimension linear in the number of channels. This leads to a much simplified suboptimal algorithm based on a greedy approach.

Related Work   A majority of existing work on DSA focuses on the approach of dynamic spectrum allocation [2, 5–16]. The European DRiVE project [2] focuses on dynamic spectrum allocation in heterogeneous networks by assuming a (logical) common coordination channel. The efficiency of DSA will depend upon the ability to predict traffic load (thus spectrum occupancy). A simulation study of the impact of load prediction based on load history and simple regression schemes is reported in [15]. Regulatory aspects and issues in DSA across multiple networks are discussed in [5]. Two centralized DSA protocols that rely on a super-base station are described in [8] and their performance evaluated via simulations.

There have been several attempts on developing cognitive MAC for opportunistic spectrum access [17–21]. These

techniques tackle the problem in two separate steps: (i) the development of an opportunity identification module using classical detection and estimation techniques assuming continuous full-spectrum sensing; (ii) the design of an opportunity allocation module by, for example, graph coloring techniques assuming full knowledge of spectrum opportunities. Missing in this line of approaches are two ingredients: the energy-efficient sensing and the ability to handle bursty traffic. First, the assumption of continuous full-spectrum sensing, while simplifying the design of cognitive MAC, is undesirable and impractical due to the energy consumption and the hardware implication. Second, traffic characteristics, especially the bursty nature, should play a crucial role in any efficient cognitive MAC scheme. Why pay for the knowledge of every opportunity in the whole spectrum all the time when a user only has sporadic needs for spectrum access? Here lies a clear disadvantage of existing approaches that decouple opportunity identification and opportunity allocation.

2. PROBLEM STATEMENT

Network and Channel Model Consider a spectrum containing N channels[1], each with bandwidth B_i ($i = 1, \cdots, N$). These N channels are shared among primary users and a large number of secondary users seeking spectrum opportunities. The traffic statistics of the primary users are such that these N channels are synchronous and slotted. We also assume that the spectrum usage statistics remain unchanged for T slots. The energy and hardware constraints restrict the secondary users from monitoring more than one channel within one slot. Extensions to a more general sensing model are discussed in Section 5.
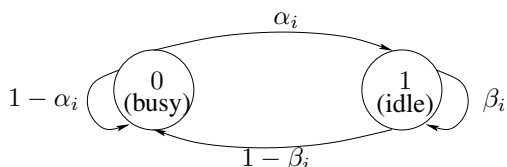

Fig. 1: The Markov channel model

We focus on a decentralized DSA ad hoc network where a large number of secondary users join/exit the network and sense/access the spectrum independently without exchanging local information. We assume that when the network reaches a steady-state, each channel independently presents itself as an opportunity to a secondary user according to a Markov process. As illustrated in Figure 1, channel states are represented by 0 (busy) and 1 (idle thus available to the secondary user). State transitions occur at the beginning of each slot with transition probabilities given by $\{\alpha_i, \beta_i\}$ ($i = $

[1] Here we use the term channel broadly. A channel can be a frequency band with certain bandwidth. It can also be a collection of spreading codes in a CDMA network or a set of tones in an OFDM system.

$1, \cdots, N$). Since the unavailability of a channel may also be caused by channel fading, the Markov chain model can also include fading statistics.

Objectives We first seek answers to a fundamental question: what is the optimal DSA strategy that maximizes the average number of bits transmitted by the secondary user in T slots. Specifically, we seek the optimal DSA protocol for the secondary user to determine in each slot which channel to monitor and subsequently access so that the average number of bits transmitted in T slots is maximized. We then exploit the specific structure of the problem in search of simpler but near optimal solutions.

Protocol Specifics We present here a CSMA-based implementation. We first discuss the basic structure of the protocol and then comment on several implementation issues.

Protocol Structure We assume that channels are slotted, and the slot timing is broadcast[2]. The beginning of each slot is dedicated for channel sensing. The primary users have the highest priority, and they sense the channel first based on certain priority related back-off scheme. For example, we can impose a minimum value on the backoff time of secondary users. The primary user can claim the slot before secondary users start sensing. The choice of the minimum backoff time for secondary users depends on the propagation delay among neighboring nodes and how much the network can tolerate interference from secondary users.

A secondary user with data to transmit will have to decide which channel to sense. Such decisions are based on its past sensing history and channel statistics using the optimal or suboptimal protocols presented in Section 3 and Section 4. If it decides to sense a particular channel, it will generate a random backoff, possibly a function of its energy level or its channel state [22], and it will transmit when the backoff timer expires and no one claims the channel.

Transmitter-Receiver Synchronization The transmitter and the receiver need to tune to the same channel in order to communicate, and they need to hop synchronously. 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. Here we borrow the idea of receiver-oriented code assignment in CDMA ad hoc networks [23]. 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 DC-MAC protocol which

[2] The slot information can be broadcast by the primary users.

governs channel selection in each slot. In this paper, we focus on the design of DC-MAC protocols assuming that the initial handshake has been established.

Collision Resolution and Avoidance   In a network with a large number of secondary users seeking spectrum opportunities independently, there needs to be a mechanism to deal with collision. The proposed DC-MAC schemes described in Section 3 and Section 4 make the access decision based on the sufficient statistic captured by the information vector, which in a way randomizes the choices of secondary users and reduces the probability of collision.

Collisions can be further minimized by the use of classical random access techniques such as CSMA or ALOHA. Specifically, a secondary user who has identified an opportunity senses the carrier using a random backoff time before accessing the channel. While such techniques do not solve the problem of hidden/exposed terminals, our protocol can be tailored to incorporate busy-tone based techniques.

We first study the design of DC-MAC assuming perfect collision avoidance. We then extend in Section 5 the proposed protocols to incorporate collision which can be modelled as misidentification of spectrum opportunity.

3. OPTIMAL STRATEGY

In this section, we formulate the DC-MAC problem as a POMDP. Under this decision-theoretic framework, we derive sufficient statistics and establish the optimal CD-MAC protocol.

3.1. The POMDP Formulation

The system of N channels given in Section 2 can be modelled by a discrete-time Markov chain with $M = 2^N$ states where the state is defined as the availability of each channel. The transition probability $p_{i,j}$ can be readily obtained from $\{\alpha_i, \beta_i\}_{i=1}^N$. The state diagram for $N = 2$ is illustrated in Figure 2 where $\bar{\alpha}_i = 1 - \alpha_i$ and state $(0, 1)$ indicates the first channel is busy whereas the second channel is available.
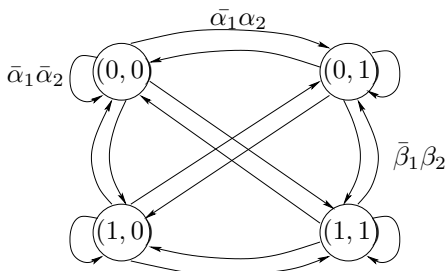
Fig. 2: The underlying Markov process for $N = 2$.

Since in each slot, the user can only select one channel to monitor, the state of the system is only partially observable.

The problem of designing a DSA protocol that maximizes the transmission rate in T slots can then be formulated as a POMDP over a finite horizon. Specifically, this POMDP consists of

- Decision intervals $\{1, \cdots, T\}$: slots;
- States $S \in \{1, \cdots, M\}$: availability of each channel;
- Transition probabilities $p_{i,j}$: functions of $\{\alpha_i, \beta_i\}_{i=1}^N$;
- Actions $a \in \{1, \cdots, N\}$: sense channel a and access if available;
- Observation $\Theta_{j,a} \in \{0, 1\}$: availability of the chosen channel a at state j;
- Reward $w_{a,\theta} = \theta B_a$: number of transmitted bits when the observation is θ under action a.

Following the illustration given in [4], we show in Figure 3 the sequence of operations in a decision interval. Specifically, at the beginning of a decision interval, the system state transits according to $p_{i,j}$. According to a chosen action a which specifies the channel to be sensed in this decision interval (slot), the user senses the channel and transmits if the chosen channel is available. The result of channel sensing is given by $\Theta_{j,a}$ which indicates the availability of the chosen channel. A reward $w_{a,\theta}$, determined by the observation and the action, is obtained at the end of this slot.
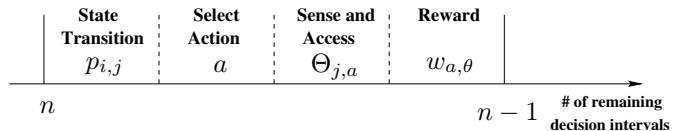
Fig. 3: The sequence of operations in a slot.

Under this formulation, the optimal DC-MAC protocol is given by the optimal policy (in terms of maximizing the expected reward in T slots) of this POMDP over a finite horizon[3].

3.2. Sufficient Statistic

Let n $(n = 1, \cdots, T)$ denote the number of remaining decision intervals. For a finite horizon POMDP over T slots, the problem is to select in slot $T - n + 1$ an action a that will optimize the system performance in the remaining n decision intervals.

Since the underlying Markov process is only partially observable, the internal state of the system is unknown. Our

---

[3]The optimal DC-MAC can also be formulated as a POMDP over an infinite horizon. Since the statistics of channel occupancy vary with time due to changes in traffic load, it is more appropriate to consider the finite horizon formulation. It is, however, straightforward to extend our formulation to an infinite horizon setup in which stationary policies can be obtained.

knowledge of the internal state of the system based on all the past decisions and observations can be encoded as an information vector $\pi = [\pi_1, \cdots, \pi_M]$ where π_i is the conditional probability (given all the past sensing history) that the state of the system is i at the beginning of the current decision interval prior to the state transition.

It has been shown in [4] that at any time the information vector π is a sufficient statistic for the optimal policy. It is easy to see that the dynamic behavior of the information vector π is itself a discrete-time continuous-state Markov process. Given the prior information π on the state of the system, our current knowledge π' of the system after observing θ under action a can be easily obtained via the Bayes' rule.

$$\pi' \triangleq [\pi'_1, \cdots, \pi'_M] \triangleq \mathcal{T}(\pi|a, \theta), \tag{1}$$

$$\pi'_j = \frac{\sum_{i=1}^{M} \pi_i p_{i,j} \Pr[\Theta_{j,a} = \theta]}{\sum_{i=1}^{M} \sum_{j=1}^{M} \pi_i p_{i,j} \Pr[\Theta_{j,a} = \theta]}. \tag{2}$$

where $\mathcal{T}(\pi|a, \theta)$ is the updated information vector from π based on observation θ and action a.

We have assumed that the state transition probabilities $\{\alpha_i, \beta_i\}$ are known. In practice, this may not be available. The problem then becomes one of POMDP with unknown transition probabilities. Such formulations and algorithms exist in the literature [24] and they are applicable to our problem.

3.3. Optimal Strategy

The optimal policy determines the action in each decision interval so that the expected total reward is maximized. Let $V_n(\pi)$ denote the maximum expected reward that can be accrued in the remaining n decision intervals when the current information vector is π. We can obtain a recursive equation for $V_n(\pi)$ as

$$V_n(\pi) = \max_{a=1,\cdots,N} \{\sum_{i=1}^{M} \pi_i \sum_{j=1}^{M} p_{i,j} \sum_{\theta=0}^{1} \Pr[\Theta_{j,a} = \theta]$$
$$(w_{a,\theta} + V_{n-1}(\mathcal{T}(\pi|a, \theta)))\}, \tag{3}$$

where $\mathcal{T}(\pi|a, \theta)$ is the a posteriori information vector given by (1).

It is shown in [4] that $V_n(\pi)$ is piecewise linear and convex and can thus be written as

$$V_n(\pi) = \max_k \pi\gamma_k(n) \tag{4}$$

for some finite set of M-dimensional column vectors $\{\gamma_k(n)\}$. In other words, the space of information vectors can be divided into a finite number of convex regions separated by hyperplanes. Within each region, $V_n(\pi) = \pi\gamma_k(n)$ for some k. Following the example given in [4], we illustrate the structure of $V_n(\pi)$ in Figure 4. We consider a three-state system. An information vector π is represented by a point in

an equilateral triangle. As shown in Figure 4, entries of π are given by the distances to the sides of the triangle. After we observe θ in a decision interval, the information vector is transformed into a point in the space of information vectors for the succeeding decision interval (see (1)). For the example given in Figure 4, the space of information vectors when there are $n-1$ decision intervals remaining is partitioned into four regions with the corresponding γ-vectors given by $\{\gamma_1(n-1), \cdots, \gamma_4(n-1)\}$.
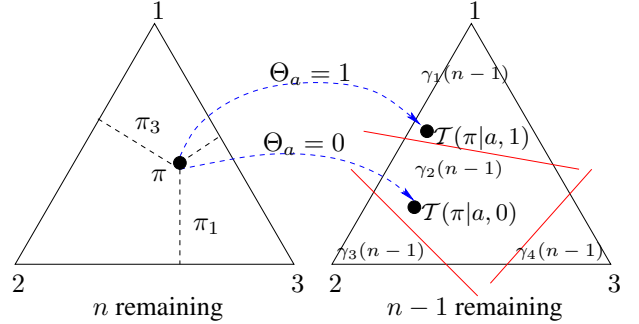

Fig. 4: The structure of $V_n(\pi)$.

The piecewise linearity and convexity of $V_n(\pi)$ lead to a linear programming procedure for calculating the optimal policy and the corresponding expected reward. Specifically, from (3) and (4) we obtain

$$V_n(\pi) = \max_{a=1,\cdots,N} \{\sum_{i=1}^{M} \pi_i \sum_{j=1}^{M} p_{i,j} \sum_{\theta=0}^{1} \Pr[\Theta_{j,a} = \theta]$$
$$(w_{a,\theta} + \mathcal{T}(\pi|a, \theta)\gamma_{l(\pi,a,\theta)}(n-1))\}, \tag{5}$$

where $l(\pi, a, \theta)$ denotes the corresponding γ-vector index for the region containing the transformed information vector $\mathcal{T}(\pi|a, \theta)$. Thus, if the set of γ-vectors for $V_{n-1}(\cdot)$ has been calculated, we can obtain from (5) the optimal action and the corresponding γ-vector for any specified information vector for the n-horizon case. A linear programming algorithm is provided in [4] for computing the γ-vectors and the corresponding mapping of these vectors onto the set of actions. Thus, the optimal policy is given by the partition of the space of the information vectors into convex regions, the γ-vectors associated with each convex region, and the mapping between the γ-vectors and the optimal actions. Note that the computation of the partition, the γ-vectors, and the mapping can be done off-line and the result stored in a table.

4. SUFFICIENT STATISTIC WITH REDUCED DIMENSION

We exploit the structure of the underlying Markov process to reduce the dimension of the sufficient statistic from exponential to linear (in N). Based on this sufficient statistic with reduced dimension, we derive a suboptimal greedy algorithm that maximizes per-slot throughput.

4.1. Reduced-State POMDP

As stated in Section 3, the 2^N-dimensional information vector π is a sufficient statistic. We show in this section that by exploiting the specific structure of the underlying Markov process, we can obtain a sufficient statistic with dimension reduced from *exponential* to *linear* with respect to N, *i.e.,* from $M = 2^N$ to N.

Proposition 1 *Let $\Lambda = [\lambda_1, \cdots, \lambda_N]$ where λ_i is the probability that channel i is available to the secondary user. Then at any time, Λ is a sufficient statistic for the above specified DSA system.*

Proof: To prove Λ is a sufficient statistic, we need to show that (i) Λ summarizes all the information on the availability of each channel obtained from the history of observations; (ii) the maximum expected reward $V_n(\cdot)$ is completely determined by Λ.

To show (i), we define $\mathcal{I}(t)$ as the total available information about the process at the end of decision interval t. Please note that the time variable t increases with time while the time variable n used in the rest of the paper denotes the number of remaining decision intervals, thus decreasing with time. Since the only information obtained during decision interval t is our observation $\Theta_a(t)$ under action $a(t)$, we have

$$\mathcal{I}(t) = \{a(t), \Theta_a(t), \mathcal{I}(t-1)\}. \tag{6}$$

Let $S_i(t)$ denote the state of channel i in slot t. We then have

$\Pr[\text{channel } i \text{ is available in slot } t | \mathcal{I}(t)]$

$= \Pr[S_i(t) = 1 | \mathcal{I}(t)]$

$= \Pr[S_i(t) = 1 | a(t), \Theta_a(t), \mathcal{I}(t-1)]$

$$= \begin{cases} 1 & \text{if } a(t) = i, \Theta_a(t) = 1 \\ 0 & \text{if } a(t) = i, \Theta_a(t) = 0 \\ \lambda_i(t-1)\beta_i + \bar{\lambda}_i(t-1)\alpha_i & \text{if } a(t) \neq i \end{cases} \tag{7}$$

where $\bar{\lambda}_i(t-1) = 1 - \lambda_i(t-1)$, and (7) follows from the independent Markov model on the dynamics of the channels. From (7) we see that the calculation of $\Lambda(t)$ based on the whole history of observations requires only $\Lambda(t-1)$ and the newly obtained information in decision interval t. Thus, $\Lambda(t-1)$ summarizes all the information on channel availability gained prior to decision interval t and represents a sufficient statistic for the past sequence of observations $\mathcal{I}(t-1)$.

We now prove (ii) by induction. For $n = 1$, we have

$$V_1(\Lambda) = \max_{a=1,\cdots,N}(\lambda_a\beta_a + (1-\lambda_a)\alpha_a)B_a. \tag{8}$$

Clearly, $V_1(\cdot)$ is completely determined by Λ. Assume $V_{n-1}(\cdot)$ is determined by Λ. It then follows that

$$\begin{aligned} V_n(\Lambda) &= \max_{a=1,\cdots,N}\{(\lambda_a\beta_a + \bar{\lambda}_a\alpha_a)B_a \\ &\quad + \sum_{\theta=0}^{1}\Pr[\Theta_a = \theta|\Lambda, a]V_{n-1}(\mathcal{T}(\Lambda|a,\theta))\} \\ &= \max_{a=1,\cdots,N}\{(\lambda_a\beta_a + \bar{\lambda}_a\alpha_a)B_a \\ &\quad + (\lambda_a\bar{\beta}_a + \bar{\lambda}_a\bar{\alpha}_a)V_{n-1}(\mathcal{T}(\Lambda|a,0)) \\ &\quad + (\lambda_a\beta_a + \bar{\lambda}_a\alpha_a)V_{n-1}(\mathcal{T}(\Lambda|a,1))\} \end{aligned} \tag{9}$$

where $\mathcal{T}(\Lambda|a,\theta)$ denotes the updated information on channel availability given the observation θ under action a. From (7) we know that $\mathcal{T}(\Lambda|a,\theta)$ is completely determined by Λ for given θ and a. We then conclude from (9) that Λ presents a sufficient statistic for calculating $V_n(\cdot)$.

□□□

Proposition 1 shows that by exploiting the statistical independence among channels, we can reduce the dimension of the sufficient statistic from 2^N to N. The optimal policy can thus be obtained from the space of information vectors Λ whose dimension increases linearly instead of exponentially with the number of channels (see (9)) for the recursive equation on $V_n(\Lambda)$). This result has the potential of significantly reducing the computational complexity and memory requirement as demonstrated in the greedy approach presented next.

4.2. A Greedy Approach

Searching for the optimal policy can be computationally intense, especially when T and N are large. In this section, we propose a suboptimal protocol based on a greedy approach that maximizes per-slot throughput. Since Λ is a sufficient statistic, the optimal action a_* that maximizes per-slot throughput is completely determined by Λ. Specifically,

$$a_* = \arg\max_{a=1,\cdots,N}(\lambda_a\beta_a + \bar{\lambda}_a\alpha_a)B_a. \tag{10}$$

Let $W_n(\Lambda)$ denote the expected throughput in the remaining n slots achieved by the greedy approach. We obtain a recursive equation for $W_n(\Lambda)$ as

$$\begin{aligned} W_n(\Lambda) &= \max_{a=1,\cdots,N}\{(\lambda_a\beta_a + \bar{\lambda}_a\alpha_a)B_a\} \\ &\quad + \sum_{\theta=0}^{1}\Pr[\Theta_{a_*} = \theta|\Lambda, a_*]V_{n-1}(\mathcal{T}(\Lambda|a_*,\theta)) \\ &= \max_{a=1,\cdots,N}\{(\lambda_a\beta_a + \bar{\lambda}_a\alpha_a)B_a\} \\ &\quad + (\lambda_{a_*}\bar{\beta}_{a_*} + \bar{\lambda}_{a_*}\bar{\alpha}_{a_*})V_{n-1}(\mathcal{T}(\Lambda|a_*,0)) \\ &\quad + (\lambda_{a_*}\beta_{a_*} + \bar{\lambda}_{a_*}\alpha_{a_*})V_{n-1}(\mathcal{T}(\Lambda|a_*,1)) \end{aligned} \tag{11}$$

where a_* is given by (10), and $\mathcal{T}(\Lambda|a,\theta)$ denotes the updated information on channel availability given the observation θ under action a. The calculation of $\mathcal{T}(\Lambda|a,\theta)$ is given in (7).

5. VARIATIONS AND EXTENSIONS

In this section, we discuss several variations and extensions of the proposed analytical framework and optimal/suboptimal DC-MAC protocols. We focus on the variations of the greedy approach; extensions to the optimal strategy follows directly.

5.1. Overlook and Misidentification of Opportunities

The Markov model for the channel occupancy allows easy incorporation of overlook and misidentification of spectrum opportunity. Consider first the overlook of opportunity. Let ϵ_i ($i = 1, \cdots, N$) denote the probability that an idle channel i is mis-sensed as busy. The channel a_* selected by the greedy approach is thus given by

$$a_* = \arg \max_{a=1,\cdots,N} (\lambda_a \beta_a + \bar{\lambda}_a \alpha_a)(1 - \epsilon_a) B_a. \qquad (12)$$

A modification of (7) leads to the update of the information vector.

$$\mathcal{T}_i(\Lambda|a,\theta) = \begin{cases} 1 & \text{if } a = i, \theta = 1 \\ \frac{(\lambda_i\beta_i + \bar{\lambda}_i\alpha_i)\epsilon_i}{(\lambda_i\beta_i + \bar{\lambda}_i\alpha_i)\epsilon_i + \lambda_i\bar{\beta}_i + \bar{\lambda}_i\bar{\alpha}_i} & \text{if } a = i, \theta = 0 \\ \lambda_i\beta_i + \bar{\lambda}_i\alpha_i & \text{if } a \neq i \end{cases},$$
$$(13)$$

where $\mathcal{T}_i(\Lambda|a,\theta)$ denotes the ith entry of the transformed information vector. A recursive equation for $W_n(\Lambda)$ can be readily obtained by modifying $\Pr[\Theta_{a_*} = \theta]$ in (11).

Consider next the scenario where with probability δ_i, the secondary user mistakes a busy channel i as an idle one. When a misidentification occurs, the user will transmit, leading to a collision. We assume instantaneous and error-free acknowledgement from the receiver to the secondary user after a successful transmission. The secondary user can thus identify the event of collision and use this information to ensure correct update of the information vector. Specifically, let $K \in \{0, 1\}$ denote whether the secondary user receives the acknowledgement ($K = 1$ indicates the reception of acknowledgement), we have

$$\mathcal{T}_i(\Lambda|a,\theta,K) = \begin{cases} 1 & \text{if } a = i, \theta = 1, K = 1 \\ 0 & \text{if } a = i, \theta = 1, K = 0 \\ 0 & \text{if } a = i, \theta = 0 \\ \lambda_i\beta_i + \bar{\lambda}_i\alpha_i & \text{if } a \neq i \end{cases}.$$

The channel a_* selected by the greedy approach in this case is given by (10), and a recursive equation for $W_n(\Lambda)$ can be obtained by modifying (11). The above two scenarios can be easily combined to model both overlook and misidentification of spectrum opportunity.

Besides sensing error, hidden and exposed terminals can also lead to overlook and misidentification of spectrum opportunities. Specifically, a hidden terminal (a transmitting node within the range of the receiver but not the transmitter) results in a misidentification while an exposed terminal

(a transmitting node within the range of the transmitter but not the receiver) causes an opportunity overlook. The problem of hidden and exposed terminals can be handled in the same manner as sensing errors discussed above. Note that while opportunity overlook and misidentification degrade the throughput performance, they do not affect the synchronization between the transmitter and the receiver. By checking whether a packet has been successfully received, the receiver is aware of the sensing results at the transmitter and can thus maintain the same update of the information vector as the transmitter. Similarly, the acknowledgement scheme discussed above ensures that both the transmitter and the receiver can incorporate misidentification into the update of the information vector.

5.2. More General Sensing/Access Models

It is straightforward to extend the proposed DC-MAC framework and protocols to accommodate more general hardware models. Specifically, the user can sense up to L_1 channels and access up to L_2 channels simultaneously. The extension of the greedy approach to this general case is straightforward. The optimal strategy also follows directly after modifying the action space to include all $\binom{N}{L_1}$ possibilities of channel selection. It is obvious that without considering cost in sensing and transmission, actions that select less than L_1 channels should not be considered, and the channels to access are those L_2 idle channels with the largest bandwidth.

6. NUMERICAL AND SIMULATION RESULTS

We present in this section numerical and simulation examples on the performance of the proposed DC-MAC protocols. We focus on the performance comparison between the optimal and suboptimal greedy approaches and the effect of opportunity overlook on the performance of DC-MAC.

6.1. Optimal vs. Suboptimal Approaches

In Figure 5 and Figure 6 we compare the transmission rate (in bits/slot) achieved by the optimal and the suboptimal protocols proposed in this paper. As shown in Figure 5, the transmission rate achieved by the greedy approach matches that of the optimal scheme in this particular setup. For the three-channel scenario considered in Figure 6, the performance loss of the greedy approach is within 4%. These examples demonstrate the near-optimal performance achieved by the greedy approach at a much lower complexity. We point out that the transmission rate increases over time. This is due to the improved information about the state of the system drawn from accumulating observations, demonstrating the cognitive nature of the proposed protocols.
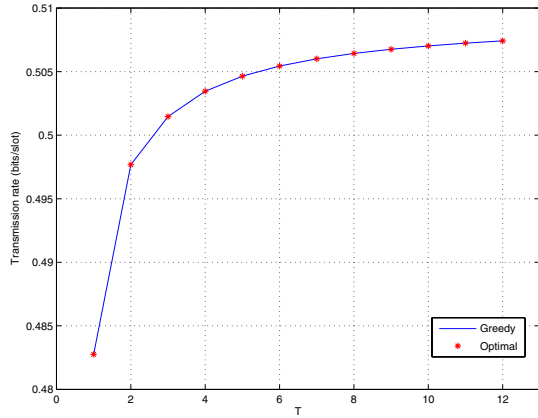
Fig. 5: Transmission rate of the optimal and greedy strategies ($N = 2$, $B_1 = 1$, $B_2 = 2$, $\alpha_1 = 0.44$, $\beta_1 = 0.23$, $\alpha_2 = 0.28$, $\beta_2 = 0.12$, the initial information vector is set to the stationary distribution).
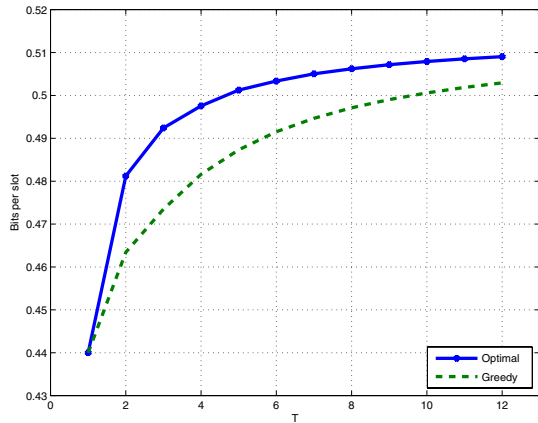

Fig. 6: Transmission rate of the optimal and greedy strategies ($N = 3$, $B = [0.9, 1, 0.8]$, $\alpha = [0.1, 0.5, 0.8]$, $\beta = [0.5, 0.4, 0.3]$, the initial information vector is set to the stationary distribution).

6.2. Robustness to Opportunity Overlook

In this simulation example, we study the performance of the suboptimal greedy approach in the presence of opportunity overlook. As given in (13), incorporating opportunity overlook into the update of the information vector requires the knowledge of the overlook probabilities $\{\epsilon_i\}_{i=1}^N$. We are particularly interested in the performance of the greedy algorithm in the presence of overlook without the knowledge of the overlook probabilities. Shown in Figure 7 is the simulation result where we study the throughput of the greedy approach as a function of the overlook probability $\epsilon_1 = \cdots = \epsilon_N = \epsilon$. We consider here three cases: the greedy approach in the absence of overlook, the greedy approach in the presence of overlook with and without the knowledge of the overlook probability ϵ. Without the knowledge of ϵ, the informa-

tion vector is updated according to (7) as if $\epsilon = 0$. From Figure 7 we can see that overlook results in linear degradation in throughput since only a fraction of $1 - \epsilon$ of spectrum opportunities are utilized. The performance of the greedy approach is, however, robust to the lack of knowledge on the overlook probability. Without the knowledge of ϵ to perform the correct update of the information vector, the performance loss is marginal.
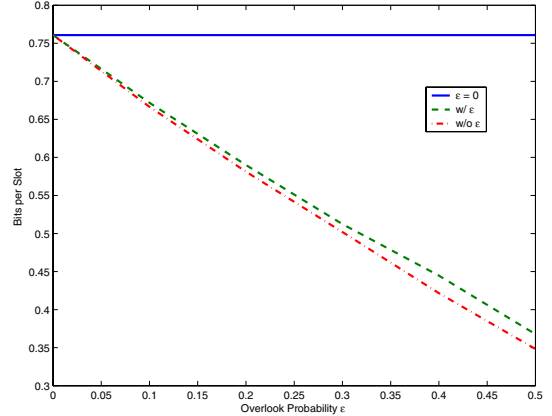

Fig. 7: Transmission rate of the optimal and greedy strategies ($N = 3$, $B = [0.9, 1, 0.8]$, $\alpha = [0.4, 0.6, 0.8]$, $\beta = [0.9, 0.7, 0.5]$, $T = 30$, the initial information vector is set to the stationary distribution).

7. CONCLUSION

We have presented an optimization framework for decentralized cognitive MAC for opportunistic spectrum access. The protocol is based on carrier sensing that allows proper assignment of priorities among primary and secondary users. The POMDP formulation allows us to derive optimal and low complexity suboptimal protocols that maximize overall network throughput. Numerical evaluation and simulations indicate that the suboptimal algorithm provides near optimal performance.

The proposed approach can be extended in a number of areas. Existing carrier sensing protocols for ad hoc networks can be easily incorporated into our framework. Opportunistic communication techniques based on channel realizations are also compatible with our framework. Of particular interest are POMDP techniques that do not assume a priori knowledge of transition probabilities.

8. REFERENCES

[1] W. D. Horne, "Adaptive Spectrum Access: Using the Full Spectrum Space," in *Proc. of the International Symposium on Advanced Radio Technologies*, Mar. 2003.

[2] 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.

[3] "DARPA: The Next Generation (XG) Program." http://www.darpa.mil/ato/programs/xg/index.htm.

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

[5] P. Leaves, K. Moessner, R. Tafazolli, D. Grandblaise, D. Bourse, R. Tonjes, and M. Breveglieri, "Dynamic spectrum allocation in composite reconfigurable wireless networks," *IEEE Communications Magazine*, vol. 42, pp. 72 – 81, May 2004.

[6] P. Demestichas, G. Vivier, K. El-Khazen, and M. Theologou, "Evolution in wireless systems management concepts: from composite radio environments to reconfigurability," *IEEE Communications Magazine*, vol. 42, pp. 90 – 98, May 2004.

[7] R. Keller, T. Lohmar, R. Tonjes, and J. Thielecke, "Convergence of cellular and broadcast networks from a multi-radio perspective," *IEEE Personal Communications*, vol. 8, pp. 51 – 56, April 2001.

[8] B. Aazhang, J. Lilleberg, and G. Middleton, "Spectrum sharing in a cellular system," in *Proc. of IEEE Eighth International Symposium on Spread Spectrum Techniques and Applications*, pp. 355 –359, Aug. - Sept. 2004.

[9] Y. Xing, R. Chandramouli, S. Mangold, and Sai. Shankar N, "Dynamic spectrum access to open spectrum wireless networks." Submitted to IEEE Journal on Selected Areas in Communications, 2005.

[10] M. Buddhikot, P. Kolodzy, S. Miller, K. Ryan, and J. Evans, "DIMSUMnet: New Directions in Wireless Networking Using Coordinated Dynamic Spectrum Access," in *in Proc. of IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks*, June 2005.

[11] V. Rodriguez, K. Moessner, and R. Tafazolli, "Market-Driven Dynamic Spectrum Allocation: Optimal End-User Pricing and Admission Control for CDMA," in *to appear in Proc. of 14th IST Mobile & Wireless Communications Summit*, (Dresden, Germany), June 2005.

[12] C. A. Nissen and G. M. Butler, "A Technology Enabled Framework for Dynamic Allocation of the Radio Frequency Spectrum," tech. rep., The MITRE Corporation, Nov. 2004.

[13] P. Leaves, S. Ghaheri-Niri, R. Tafazolli, and J. Huschke, "Dynamic Spectrum Allocation in Hybrid Networks with Imperfect Load Prediction," in *Proc. of IEE Second International Conference on 3G Mobile Communication Technologies*, pp. 444 – 448, May 2002.

[14] S. Mangold, Z. Zhong, K. Challapali, and C. Chou, "Spectrum agile radio: radio resource measurements for opportunistic spectrum usage," in *Proc. of Globecom*, 2004.

[15] K. Challapali, S. Mangold, and Z. Zhong, "Spectrum Agile Radio: Detecting Spectrum Opportunities," in *International Symposium on Advanced Radio Technologies*, 2004.

[16] A. Larcher, H. Sun, M. van der Shaar, and Z. Ding, "Decentralized Transmission Strategies for Delay-sensitive Applications over Spectrum Agile Networks ," in *Proc. of International Packet Video Workshop*, (Irvine, CA), Dec. 2004.

[17] W. Wang, X. Liu, and H. Xiao, "Exploring opportunistic spectrum availability in wireless communication networks," in *submitted*, 2005.

[18] Q. Zhao and L. Tong, "Opportunistic Carrier Sensing for Energy Efficient Information Retrieval in Sensor Networks," *EURASIP Journal on Wireless Communications and Networking*, no. 2, pp. 231–241, 2005.

[19] M. B. Pursley, "The Role of Spread Spectrum in Packet Radio Networks," *Processings of The IEEE*, vol. 75, pp. 116–134, Jan. 1987.

[20] K. Murphy, "A survey of POMDP solution techniques," tech. rep., 2000. Technical Report, UC Berkeley, http://www.cs.ubc.ca/m̃urphyk/papers.html.