Low-overhead Heuristic Algorithms for Spectrum Sensing in Cognitive Radio Networks

V. Miliotis\textsuperscript{1}, I. Koutsopoulos\textsuperscript{2,1}, L. Alonso\textsuperscript{1} and C. Verikoukis\textsuperscript{*}

\textsuperscript{1}Technical University of Catalonia, Department of Signal Theory and Communications, Barcelona, Spain
\textsuperscript{2}Department of Computer Science, Athens University of Economics and Business (AUEB), Athens, Greece
\textsuperscript{*}Centre for Research and Technology Hellas (CERTH), Greece

Email: \{vasileios.miliotis, luigsl\}@tsc.upc.edu, jordan@aueb.gr, cveri@cttc.es

Abstract—Cognitive Radio (CR) has matured as a research field and has the goal to improve the utilization of spectrum. This technology enables a CR user to function in unlicensed or licensed bands without causing interference to primary users. In order to achieve this functionality a CR user needs to sense the spectrum and exploit efficiently the transmission opportunities. In this paper we propose two heuristic sensing algorithms for ad-hoc CR networks with low computational requirements, where each user tries to exploit “spectrum holes” based only on its own observations. The first algorithm refers to an immediate reward approach, according to which, each time a “spectrum hole” is detected, it is used instantaneously. The second refers to an immediate-versus-future-reward approach, where the CR may wait to sense the next channel instead of transmitting in the current one, if the expected reward when using the next channel is greater than the instant. In our proposed algorithms we examine both known and unknown statistics of channels. For the case of unknown channel parameters we introduce an initial sensing time and we investigate through simulations the impact of different sensing periods on the system’s throughput.

I. INTRODUCTION

In order to enable a Cognitive Radio (CR) to exploit the spectrum opportunistically, it is necessary that it is able to (i) examine the spectrum (spectrum sensing), in which it will try to detect “spectrum holes”, (ii) analyse the sensing outcome and (iii) decide which frequency band will be used, as well as the transmission rate depending on its requirements and on the estimated availability of the chosen band. As soon as the spectrum decision is made, the CR can start transmitting. Nevertheless, since the electromagnetic environment changes over time and geographic location, the CR needs to monitor these changes, in order to leave the current channel when it becomes unavailable and switch to another one. In this way, seamless communication can be achieved. Any change during transmission time such as the appearance of a primary user, the movement of the user or a traffic variation must cause the CR to adjust its parameters on the new spectral conditions.

Spectrum sensing and accessing in a CR network depends mainly on the implemented network architecture. Thus, in a hierarchically structured network, each user could sense the spectrum separately in a distributed manner and the sensing outcomes could be sent to a network coordinator who would finally decide how the allocation of the free channels should be done. In the case of an ad-hoc CR network, every user operates autonomously, i.e. based on its own observation results decides which is the best channel to access. In this paper we consider this case. Most of the existing works on spectrum exploration and exploitation propose sensing algorithms which require computationally demanding operations. The optimality of some of the proposed algorithms is proved, but their real-time implementation on a CR would lead to consumption of high energy and computational resources consumption. In fact, there has been a lot of progress in theoretical analysis of CR sensing but little has been put into practice.

In this paper we propose two heuristic sensing algorithms with low computational requirements for the exploration and exploitation of transmission opportunities by a CR user that operates autonomously in an ad-hoc CR network. These algorithms operate in a look-ahead fashion, however their advantage is that they are easy to implement. The first algorithm refers to an immediate reward approach, where every time a transmission opportunity is detected, it is exploited instantly. The second refers to an immediate-versus-future-reward approach, where a CR user may wait for a better transmission opportunity instead of using the current one. The basic steps that we followed for the design of these algorithms comprise the modeling of the wireless medium, the assumption on the a priori knowledge of the CR about the statistics of the examined channels and the decision criteria for access to a specific channel.

II. RELATED WORK

In the past few years, the augmenting congestion in the usage of the electromagnetic spectrum has led the research activity in the investigation of systems that provide for opportunistic spectrum access. A basic feature in such systems, that is required for CR users, is the ability of sensing the channels, aiming to detect transmission opportunities. This operation requires particular sensitivity of the reception antennas.

Certain initial technical specifications have been published by the IEEE 802.22 Working Group [1], which refers to hierarchical structured secondary networks. The authors in [2] present a decentralized MAC protocol for ad-hoc CR networks, where each user maintains a belief vector about
the states of the channels. The availability of each channel follows a Markov process. As the monitoring capabilities of the spectrum are limited, the user maps the current belief vector to a sensing and accessing decision, in order to maximize its total throughput. The mapping of the belief vector to a sensing and accessing decision is approached as a Partially Observable Markov Decision Process (POMDP). The work in [3] presents a myopic policy for the detection of "spectrum holes", modeling the channels as independent and statistically identical Markov processes. This policy aims at the maximization of immediate reward, overlooking the impact of current decision in the future evolution of the users payoff. The optimality of this policy is proved for \( N = 2 \) channels, while for \( N > 2 \) it was further analysed and proved in [4]. In [5] the equivalence of the sensing and accessing problem in CR networks with the multi-armed bandit problem is presented. The trade-off between exploring and exploiting the channels is studied in order to achieve a balance between seeking for transmission opportunities and exploiting the available channels. The same approach is used in [6], but this time a multi-user CR network is assumed and a form of competition is added in the problem. Consequently, for the avoidance of conflicts, a Carrier Sense Multiple Access/ Collision Avoidance (CSMA-CA) protocol is proposed. An alternative approach from the prevailing proposed techniques is followed in [7], for the case of sensing and accessing of not predetermined portions of the spectrum and for unknown channel statistics, where a sensing policy would lead to waste of time and energy. Thus, the use of a "pilot" channel is proposed, through which relative information on the channels is provided to the users. In [8] and [9] the use of game theory is proposed for a multi-CR system. In this case a user adapts its operation according to his perception for the existence of primary users but also considering the probable adaptation of operation of other secondary users.

III. SYSTEM MODEL

We consider a CR user searching for channels to exploit. The available spectrum consists of \( N \) channels. These \( N \) channels are modeled as independent but not stochastically identical Gilbert-Elliot channels [10], [11]. We assume the existence of several groups of channels. Only channels that belong to the same group are stochastically identical. Every channel can be either in "good" or in "bad" state. For the case of Cognitive Radios we map the "good" state to an unused channel, and the "bad" state to a used channel. The transitions between the above states of a channel form a Markov chain with transition probabilities \( \{p_{bb}, p_{bg}, p_{gb}, p_{gg}\} \), where \( b \) and \( g \) correspond to "bad" and "good" states respectively. The time is assumed in slots and channel states are updated in every slot. The CR senses one channel per slot. If the channel is found unused, the CR transmits and takes a reward equal to 1. If not, it waits for the next slot.

IV. IMMEDIATE-REWARD ALGORITHM

In [2], [5] and [6] the solution for the sensing and accessing problem is based on the calculation of a belief vector, which depicts the probability distribution of the states of the channels. This vector is calculated taking into consideration the decision history and the results of observations. Nevertheless, in [12] it is proved that the additional use of the induced rewards by the adopted decisions decreases the uncertainty in the computation of the belief vector.

In this section we propose two versions of the immediate reward heuristic algorithm for the sensing problem in the context of an ad-hoc CR network. In the first version, named Belief-Based Sensing (BBS), the statistics of the channels are assumed known a priori to the CR user, while in the second, named Score-Based Sensing (SBS), no information is available. For the sake of comparison we describe and implement two more algorithms, one for random and one for sequential sensing.

1) Belief-Based Sensing (BBS): In the two preceding algorithms no knowledge of the transition probability matrices is available to the CR user. For the case of known transition probabilities we develop a Belief-Based Sensing algorithm. A belief vector of dimension \( N \times 1 \), \( \mathbf{v}(t) \), expresses the belief of the CR user that channels are at "good" state at time \( t \). For the first slot choice, the initial values of the elements of the belief vector are equal to the steady-state distribution of the underlying Markov chain of each channel. Every channel \( i \) is described by a two-state Markov chain, with states "bad" and "good", and its transition probability matrix is

\[
\begin{bmatrix}
    p_{bb} & p_{bg} \\
p_{gb} & p_{gg}
\end{bmatrix}
\]

The steady-state distribution is equal to the left-hand eigenvector of the transition probability matrix corresponding to the dominant eigenvalue \( \lambda_1 \) (as the transition probability matrix is stochastic \( \lambda_1 = 1 \)). Specifically, it is

\[
\mathbf{v}_i(0) = \frac{p_{bg}}{p_{bg} + p_{gb}}, \quad i = 1, ..., N
\]

At the beginning of each slot the CR decides to sense the channel with the maximum probability of being unused, i.e. the corresponding channel to the maximum element of the belief vector. At the end of slot \( t \), given the sensed channel is \( a \) and the observation \( S_a \), the CR updates the belief vector as follows

\[
\mathbf{v}_i(t+1) = \begin{cases} 
p_{gg}, & \text{if } a=i \text{ and } S_a=\text{"good"} \\
p_{bg}, & \text{if } a=i \text{ and } S_a=\text{"bad"} \\
v_i(t)p_{gg} + (1 - v_i(t))p_{bg}, & \text{if } a \neq i
\end{cases}
\]

2) Score-Based Sensing (SBS): In the score-based sensing algorithm, we enable the CR user to act randomly for a small initial period of time \( t_c \), which is the cognition time. During this period the CR estimates the channel statistics and it does not access any channel to transmit. In order to estimate the steady-state distribution of the underlying Markov chain of each channel we introduce a simple though useful metric, the Hit Ratio, where

\[
\text{Hit Ratio}_i = \frac{\# \text{ slots with } S_i=\text{"good"}}{\# \text{ observations of } i}, \quad i = 1, ..., N
\]
Under these observations at the end of $t_c$, the CR chooses the subset of channels with the higher Hit Ratio and continues to choose randomly, but this time among the subset of channels with the higher Hit Ratio. It is underlined that in this algorithm there is no information on the transition probability matrices of the channels, though its performance is not much subordinate than this of the preceding. Through simulations it is observed that its throughput is similar to that of the Belief Based Sensing algorithm, as shown in Fig. 2.

3) Random Sensing (RS): In the random sensing algorithm, at the start of each slot, the next channel to be examined is chosen randomly with equal probability. As expected, through simulations we can see that poor throughput is achieved.

4) Sequential Sensing (SS): In the sequential sensing (SS) algorithm, the channels are examined sequentially according to their frequencies. Again, transmission opportunities are exploited immediately. The throughput of the sequential sensing does not differ from the random sensing.

V. IMMEDIATE-VERSUS-FUTURE-REWARD ALGORITHM

In this section we propose three versions of the immediate-versus-future-reward heuristic algorithm for sensing and accessing, where the CR user does not exploit immediately the transmission opportunities. We introduce a decision process that enables the CR to compare the immediate reward to the future reward and decide whether it will wait for a possibly higher reward versus adopting the instantaneous one. In the first two algorithms the channel parameters are assumed known a priori, while in the third one, no information is available. In the last algorithm we propose an estimation procedure for the parameters of the Markov chains that describe the system channels. The parameter estimation of the system belongs to the class of estimation problems of a Hidden Markov Model (HMM), where no analytical solution exists for the minimization of the distance between the estimated and the real parameters [13]. The common characteristic of the algorithms of this section is that the channels are sensed sequentially after being ranked according to a quality criterion.

In this section we map the “good” state to a transmission rate equal to $r_1$, and the “bad” state to a transmission rate equal to $r_2$, where $r_1 > r_2$. The transitions between the channel states above, form a Markov chain with transition probabilities $\{p_{11}, p_{12}, p_{21}, p_{22}\}$. The time is assumed in slots and channel states are updated in every slot. The CR senses only one channel per slot. A belief vector $v(t)$ expresses the estimate of the user for the probability that channels are at “good” state in slot $t$. It is updated according to the transition probabilities of the channels as well as the sensing outcome of the last examined channel.

1) Sequential Sensing with Restart (SSR): In this algorithm the transition probabilities of channel states are known (e.g. the local channel statistics are available in a database). The belief of the CR user that the channels are at “good” state is expressed with a belief vector. In the first time slot the initial values of the elements of the belief vector are equal to the steady-state distribution (s.s.d.) of the Markov process that characterizes each channel according to (1). The CR user ranks the channels by decreasing weighted s.s.d. by rates. Thus, the weighted s.s.d. for channel $i$ can be expressed as follows:

$$r_1 \frac{p_{21}^i}{p_{12}^i + p_{11}^i} + r_2 \frac{p_{12}^i}{p_{21}^i + p_{11}^i}$$

which can be written as:

$$1 \frac{r_1 p_{21}^i + r_2 p_{12}^i}{p_{21}^i + p_{12}^i}$$

At the beginning the CR starts sensing the channels according to the above ranking. During the sensing time $t_s$, the belief vector is updated. At the end of each slot a condition $C$ is examined and decision $d(t)$ is taken. If the condition stands $d(t) = 1$, and the CR transmits through the last examined channel for the next $t_d - t_s$ slots. Otherwise $d(t) = 0$ and it continues sensing the next channel. Where $t_d$ is the maximum decision duration. Given the observation $S_i$ in slot $t$, the examined condition $C$ is true if $S_i =$ “good” and the following inequality holds

$$p_{11}^i > v_{i+1}(t+1)p_{21}^i + (1 - v_{i+1}(t+1))p_{21}^i + 1$$

and false otherwise. Consequently

if $C =$ true, $d(t) = 1$ : transmit in channel $i$ in $t + 1$ slot
if $C =$ false, $d(t) = 0$ : scan channel $i + 1$

Every $t_d$ seconds the previous procedure is repeated. During access time the update of the belief vector continues in order to enrich its knowledge of the channels. When the decision duration ends, the CR starts sensing the channels from the beginning according to their initial ranking.

2) Sequential Sensing with Access from Bad Channel also and Restart (SSABCR): This algorithm differs from the previous in that the CR can lock to a channel even if it is found in “bad” state and transmit in lower rate. Now, the condition $C$ is true if $S_i =$ “good” and the following inequality holds

$$p_{11}^i > v_{i+1}(t+1)p_{21}^i + (1 - v_{i+1}(t+1))p_{21}^i$$

or if $S_i =$ “bad” and the following inequality holds

$$p_{21}^i > v_{i+1}(t+1)p_{21}^i + (1 - v_{i+1}(t+1))p_{21}^i$$

and false otherwise. The decision $d(t)$ follows (6).

3) Hit Ratio with Parameter Estimation Algorithm (HRPE): To implement the previous algorithm in case of unknown channel statistics we dedicate an initial time $t_e$ for cognition purposes. In this phase the CR senses the channels and it does not access any channel to transmit. During this time the CR user estimates the s.s.d. of the Markov chains of the channels as well as their transition probabilities. The channels are examined sequentially according to their frequencies, before being ranked. Every channel is sensed for $t_e/N$ continuous slots, where $N$ is the number of the channels. The s.s.d is estimated by the Hit Ratio according to (3). The probability that channel $i$ is in “good” state is expressed as $\hat{p}_i^1$. The transition probabilities are estimated according to (8) and (9).

$$\hat{p}_i^1 = \text{Hit Ratio}_i, \quad i = 1, ..., N$$
The performance is measured under the same channel states and transitions for all algorithms. In Fig. 1 we can see the Hit Ratio of the channels for the SBS algorithm. For the SBS algorithm we set the initial cognition time to \( t_c = 1000 \) slots, as after this period the quality of each channel is clearly estimated according to the Hit Ratio. We observe that after \( t_c \), only the channels with high Hit Ratio are sensed. This policy enables the CR to operate nearly like the BBS algorithm, without knowing the channel statistics at the beginning of the cognition time. This fact becomes clearer in Fig. 2, which depicts the reward over time for all the immediate reward algorithms. Through the simulation we show that the Score Based algorithm is subordinate from the Belief Based for less than 10% regarding the reward obtained over a period of time \( T \), while Random and Sequential Sensing algorithms have lower rewards in the order of more than 50%.

VI. SIMULATION RESULTS

A. Immediate-Reward Algorithms

We simulated the immediate reward algorithms over 12 channels, grouped in 3 different classes according to their quality. The simulation is running for \( T = 12000 \) slots and the performance is measured under the same channel states and transitions for all algorithms. In Fig. 1 we can see the Hit Ratio of the channels for the SBS algorithm. For the SBS algorithm we set the initial cognition time to \( t_c = 1000 \) slots, as after this period the quality of each channel is clearly estimated according to the Hit Ratio. We observe that after \( t_c \), only the channels with high Hit Ratio are sensed. This policy enables the CR to operate nearly like the BBS algorithm, without knowing the channel statistics at the beginning of the cognition time. This fact becomes clearer in Fig. 2, which depicts the reward over time for all the immediate reward algorithms. Through the simulation we show that the Score Based algorithm is subordinate from the Belief Based for less than 10% regarding the reward obtained over a period of time \( T \), while Random and Sequential Sensing algorithms have lower rewards in the order of more than 50%.

\[
\hat{p}_{12}^i = \frac{\text{# transitions of } i \text{ from "good" to "bad"}}{\text{# observations of } i}, i = 1, \ldots, N \tag{8}
\]

\[
\hat{p}_{21}^i = \frac{\text{# transitions of } i \text{ from "bad" to "good"}}{\text{# observations of } i}, i = 1, \ldots, N \tag{9}
\]

The ranking of the channels here is done by decreasing weighted estimated s.s.d. For every channel we have:

\[
\frac{1}{\hat{p}_{21}^i + \hat{p}_{12}^i} (r_1 \hat{p}_{21}^i + r_2 \hat{p}_{12}^i) \tag{10}
\]

and the belief vector, which initially is equal to the estimated s.s.d, is now updated as follows:

\[
\hat{v}_i(t + 1) = \begin{cases} 
\hat{p}_{11}^i, & \text{if } a = i \text{ and } S_a = "good" \\
\hat{p}_{21}^i, & \text{if } a = i \text{ and } S_a = "bad" \\
\hat{v}_i(t)\hat{p}_{11}^i + (1 - \hat{v}_i(t))\hat{p}_{21}^i, & \text{if } a \neq i 
\end{cases} \tag{11}
\]

B. Immediate-vs-Future-Reward Algorithms

In our simulations we have assumed 12 stochastic and non statistically identical Markov channels. In order to explore the trade-off between cognition time and total system throughput for the HRPE algorithm we repeated the simulation for different \( t_c = \{60, 120, 180, \ldots, 2400\} \) time slots, 100 times for each cognition duration. Every simulation ran for \( T = 12000 \) slots and the maximum decision duration was set equal to \( t_d = 50 \) slots. In Fig. 3 we can see the mean system’s throughput for different cognition durations. The maximum mean throughput for a total time of \( T = 12000 \) slots is achieved for cognition time \( t_c = 300 \) slots. Different cognition durations affect the energy efficiency of the CR. We assess the energy efficiency of the system in Fig. 4 using the power levels provided in [14] for WiFi transceivers for a transmission rate of 54 Mbps. In Fig. 5 we see a throughput comparison for all the immediate-versus-future-reward algorithm versions, for \((r_1, r_2) = (2, 1)\), when the \( t_c = 300 \) slots for the HRPE algorithm. We observe that the SSABCR algorithm is slightly better than the SSR and best in overall performance. The HRPE algorithm has satisfactory performance even if the
channels statistics are initially unknown. In Fig. 6 we present the mean Kendall-tau distance [15] of the channels ranking, based on the estimation outcome for different durations of the cognition time $t_c = \{60, 120, 180, \ldots, 2400\}$. We use Kendall-tau distance as a metric of the pairwise disagreements to compare the channel ranking based on estimation with the channel ranking based on the actual channel statistics. We notice that for $t_c > 500$ the Kendall-tau distance presents less than 5 permutation errors, while the maximum distance is equal to $N(N-1)/2$ permutation errors.

VII. CONCLUSIONS

We considered the problem of spectrum sensing and accessing in the context of Cognitive Radio networks for both immediate reward and immediate-versus-future-reward systems. The proposed algorithms assume known or unknown channel statistics. In the case of unknown channel statistics we introduced heuristic metrics for the estimation of the unknown parameters and we showed that the performance of the corresponding algorithms has not subordinate performance to these with known channel statistics. A prime feature of the proposed algorithms is that they present low computational requirements and are able to provide satisfying performance in real-time systems. They also operate in a look-ahead fashion that can be found in dynamic programming and partially observable Markov decision processes but at the same time they are of low complexity that makes them easy to implement in real CR systems.

ACKNOWLEDGMENTS

This work has been partially funded by the Research Projects GREENET (PITN-GA-2010-264759) and GEOCOM (TEC2011-27723-C02-01). V. Miliotis acknowledges the support of the Greek State Scholarship Foundation (IKY), following a procedure of individualised assessment of the academic year 2011-2012, through the resources of the educational program “Education and Lifelong Learning” of the European Social Fund (ESF) and the NSRF 2007-2013.

I. Koutsopoulos acknowledges the support of ERC08-RECITAL project, co-financed by Greece and the European Union (European Social Fund) through the Operational Program “Education and Lifelong Learning” - NSRF 2007-2013.

REFERENCES


