

Multi-band Spectrum Selection Framework based on Partial Observations

J. Pérez-Romero, A. Raschellà, O. Sallent, A. Umbert

Dept. of Signal Theory and Communications

Universitat Politècnica de Catalunya (UPC)

Barcelona, Spain

e-mail: [jorperrez, alessandror, sallent, annau]@tsc.upc.edu

Abstract— The demand for higher data rates, capacity and better Quality of Service (QoS) is constantly growing; therefore, there is a pressing need for efficient use of wireless network resources. In this context, the application of Cognitive Radio (CR) principle has emerged as a solution to the problem of spectrum scarcity. In this respect, this paper presents an algorithm that enables multi-band spectrum selection for wireless applications under interference variations in the available spectrum blocks. The algorithm is formulated as a Partially Observable Markov Decision Process (POMDP) targeting the maximization of a reward function that captures how suitable each allocated spectrum block is for the data transmission in each radio link depending on the application bit rate requirements.

Keywords-spectrum selection; Partially Observable Markov Decision Process (POMDP)

I. INTRODUCTION

Regulatory bodies at international, European and national levels are actively working towards efficient and flexible spectrum regulation by fostering technology and service neutral spectrum management, spectrum trading and promotion of collective use of spectrum as well as shared use of spectrum. In such framework, pursuing an increase in spectrum usage efficiency through Dynamic Spectrum Access (DSA), Cognitive Radio (CR) paradigm appears as a key enabling technology [1][2]. In this context, spectrum selection refers to choosing the most appropriate portion of spectrum to be used in DSA/CR communication systems. Several works have addressed the spectrum selection [3]-[5] highlighting the importance of efficient decision-making criteria.

In order to perform an efficient spectrum selection, the cognitive cycle paradigm -including observation, analysis, decision and action- is exploited in this paper. The observation of the radio environment and the analysis of such observations will lead to acquire knowledge about the state of the potential spectrum blocks that can be selected as well as their dynamic behavior. Observations typically involve making measurements at the terminal side and reporting back to the infrastructure side, then resulting costly in terms of signaling overhead, battery consumption, etc. Consequently, decision-making strategies able to efficiently operate with the minimum amount of measurements would be of high interest. In this respect, Partially Observable Markov Decision Processes (POMDPs) [6] become a powerful decision making tool since they can achieve an optimized performance relying on partial observations of the system in which just a limited number of

spectrum blocks are measured while the others are characterized statistically.

This paper proposes an algorithm for efficiently deciding the spectrum block assignment to a number of radio links intended to support applications with specific bit rate requirements in the presence of external interference variations in the candidate spectrum blocks. The problem is formulated as a POMDP in which the agents making the decisions do not have a full knowledge about the state of all the available spectrum blocks, but they only rely on observations made at some instants, thus resulting in a partial observation of the system. As a first difference from previous works that have considered POMDP for spectrum selection such as [7], this paper does not rely only on binary (i.e., idle/occupied) measurements but it is able to capture different degrees of interference. Moreover, the proposed framework captures the heterogeneity in user requirements.

Section II describes the system model and formulates the considered spectrum selection problem as a POMDP, presenting the corresponding spectrum selection policy. Section III presents some initial performance evaluation results and finally, Section IV points out some concluding remarks.

II. SYSTEM MODEL AND PROBLEM FORMULATION

The system is characterized by a set of radio links $j=1, \dots, N$ each one intended to support data transmission with bit rate requirement $R_{req,j}$ between a pair of terminals and/or infrastructure nodes. The potential spectrum to be assigned to the links is organized in $i=1, \dots, M$ spectrum blocks with different central frequencies and bandwidths. They can belong to different bands subject to different interference conditions. The available bit rate for the j -th link in the i -th spectrum block $R_{j,i}$ will depend on both the propagation conditions between the j -th link transmitter and receiver as well as on the interference experienced by the receiver in the i -th block. Then, the spectrum selection problem considered here consists in performing an efficient allocation of the spectrum blocks to the radio links by properly matching the bit rate requirements with the achievable bit rate in each spectrum block.

Spectrum selection decision making is executed in a centralized entity in the infrastructure node that controls the existing links in the network. The overall process follows the steps of the classical cognitive cycle, in which the spectrum selection decisions are supported by the information stored in a Knowledge Database (KD) that includes the knowledge resulting from the analysis of the measurements (observations) made on the different spectrum blocks. Decisions made are

This work has been performed in the context of the OneFIT project (Opportunistic networks and Cognitive Management Systems for Efficient Application Provision in the Future Internet, www.ict-onefit.eu) which is supported by the European Community's Seventh Framework Program (FP7). This paper reflects only the authors' views and the Community is not liable for any use that may be made of the information contained therein. The work is also supported by the Spanish Research Council and FEDER funds under ARCO grant (ref. TEC2010-15198).

translated into actions to configure the existing links with the assigned spectrum blocks.

The considered interference model denotes as $I_{j,i}(t) = I_{\max,j,i} \sigma_i(t)$ the interference spectral density measured at time t by the j -th link receiver in the i -th spectrum block due to other external transmitters (i.e. outside the control of the decision making entity). In order to capture the time variations of the interference, $\sigma_i(t)$ is a spectrum block-specific term between 0 and 1 (i.e. $\sigma_i(t)=0$ when no interference exists and $\sigma_i(t)=1$ when the interference reaches its maximum $I_{\max,j,i}$). For modeling purposes, the values of $\sigma_i(t)$ are converted to a discrete set of interference states $S_i(t) \in \{0, 1, \dots, K\}$ where state $S_i(t)=k$ corresponds to $\sigma_{k-1} < \sigma_i(t) \leq \sigma_k$ for $k > 0$ and to $\sigma_i(t) = \sigma_0 = 0$ for $k=0$. Note also that $\sigma_K=1$. Moreover, the dynamics of the interference state $S_i(t)$ for the i -th block is modeled as a discrete-time Markov process that evolves in discrete time intervals of duration Δt with state transition probability from state k to k' given by $p_{k,k'}^i$. The state of each spectrum block evolves independently from the other blocks.

The execution of the spectrum selection decision-making algorithm results into *actions*. Let denote as $a_j(t) \in \{1, \dots, M\}$ the action made for link j at time t corresponding to the selected spectrum block among those currently available (i.e. not allocated to other links). An action is taken for a given link whenever a data transmission session is initiated on this link.

As a consequence of the different actions and resulting spectrum block assignments, each radio link with a data session in course (i.e., an active link) will obtain a reward that measures the obtained performance. Then, let denote $r_{j,i,S_i(t)} \in [0, 1]$ the reward that the j -th link gets at time t when using its allocated spectrum block i and the interference state is $S_i(t)$. It captures how suitable the i -th spectrum block is for the j -th radio link/application, depending on the achievable bit rate $R_{j,i,k}(t)$ relative to the required bit rate $R_{req,j}$. The total system reward $T_R(t)$ is given by the sum of rewards of all the active links at time t .

As a general target, the spectrum selection decision making should follow the optimal policy that maximizes the performance in terms of the expected long-term total system reward $T_R(t)$ accumulated over a certain time horizon tending to infinity. For this purpose, the decision-making entity would ideally need to know the actual interference state of all the spectrum blocks at time t . However, this would impact in terms of increasing signaling overheads and battery consumption to perform all the required observations (i.e., measurements) and report them to the decision-making entity.

To overcome this issue, this paper proposes to make the decisions based on a statistical characterization of the interference state of the different spectrum blocks rather than on actual exhaustive observations. Observations about the interference state of the spectrum blocks are carried out only at specific time instants defined according to a certain observation strategy. Then, due to the partial knowledge that the decision making process has about the actual interference state of the spectrum blocks, the spectrum selection process can be modeled as a POMDP and the statistical characterization of the spectrum blocks at time t is given in terms of the so-called belief vector $\mathbf{Y}(t) = [b_{i,k}(t)]$ where

component $b_{i,k}(t)$ is the probability that the i -th block will be in state $S_i(t)=k$ at time t .

In a POMDP the complexity associated to finding the optimal policy that maximizes the expected long-term system reward is usually prohibitive, mainly because the number of states $(K+1)^M$ grows exponentially with the number of spectrum blocks. Consequently, this paper proposes to use instead the so-called *Myopic Policy* that maximizes the immediate system reward $T_R(t)$. Specifically, considering that the spectrum block selection is made in time t for just one link j and among the set of available blocks so the selection will not impact on the immediate reward of any other link, the myopic spectrum selection policy becomes:

$$a_j(t) = \arg \max_{i \in \{1, \dots, M\}} \sum_{k=0}^K b_{i,k}(t) \sum_{k'=0}^K p_{k,k'}^i r_{j,i,k'}(t + \Delta t) \quad (1)$$

Based on all the above, the implementation of the spectrum selection decision making following (1) requires that the KD stores the state transition probabilities for the different spectrum blocks $p_{k,k'}^i$, the values of the reward $r_{j,i,k}$ that the different radio links can obtain in each spectrum block for each interference state, and the belief vector values $b_{i,k}(t)$.

Concerning $p_{k,k'}^i$ and $r_{j,i,k}$, they could in practice be obtained based on some specific initial acquisition mechanisms including measurements of the different links and spectrum blocks. As for the belief vector values $b_{i,k}(t)$, they should be dynamically updated with time resolution Δt based on the discrete-time Markov process that models the interference state in each spectrum block. To perform this update, the knowledge about the real interference of the spectrum blocks obtained through observations (i.e. measurements) performed at certain time instants can be exploited to obtain a more accurate estimation of the probability that the i -th block will be in state k at a later time. More precisely, let define as $o_i(t)$ the observation made at time t in the spectrum block i . This observation provides the actual interference state of the spectrum block, that is $o_i(t)=k$. Using the available observations $o_i(t)$, the values of $b_{i,k}(t)$ are updated for all the spectrum blocks every Δt as follows:

$$b_{i,k'}(t + \Delta t) = \begin{cases} p_{k,k'}^i & \text{if } (o_i(t) = k) \\ \sum_{n=0}^K p_{n,k'}^i b_{i,n}(t) & \text{otherwise} \end{cases} \quad (2)$$

According to the above, an observation strategy is required to determine the time instants in which the observations of the different spectrum blocks are carried out. In the context of this paper, it is assumed that observations are executed every Δt only for the spectrum blocks that are allocated to some link.

It is worth mentioning that this paper assumes that the network operates in a stationary environment, so that the values of the state transition probabilities and the rewards for the different links/spectrum blocks do not change. Operation in non-stationary environments is left for future work.

III. INITIAL PERFORMANCE EVALUATION RESULTS

This section presents some initial results to obtain a first insight into the potentials of the proposed methodology. $L = 2$ links are considered, each one generating sessions with duration exponentially distributed with average 120s (link #1) and 60s (link #2). The time between the end of a session and the beginning of the next one is exponentially distributed with average 200s (link #1) and 100s (link #2). Required bit rates are 200 Mb/s (link #1) and 100 Mb/s (link #2). $M = 3$ spectrum blocks are considered, each one with 3 interference states. State 0 is the case without interference and state 2 is the case with maximum interference. Bit rates and reward values are presented in Table I. Average durations (minutes) in states $\{0,1,2\}$ are $\{3, 1, 1\}$ for block B1, $\{10,5,5\}$ for block B2, and $\{8,6,6\}$ for block B3. $\Delta t=1$ s has been considered.

TABLE I. BIT RATES (MB/S) AND REWARD VALUES OF THE LINKS IN THE DIFFERENT SPECTRUM BLOCKS

Link	Spectrum Block	State $S_i=0$		State $S_i=1$		State $S_i=2$	
		$R_{ii,0}$	$r_{ii,0}$	$R_{ii,1}$	$r_{ii,1}$	$R_{ii,2}$	$r_{ii,2}$
1	B1	264	0.92	150	0.85	87	0.21
	B2	297	0.86	246	0.95	87	0.21
	B3	365	0.74	308	0.84	73	0.11
2	B1	145	0.87	40	0.16	8	0.00
	B2	204	0.68	151	0.85	12	0.00
	B3	263	0.55	184	0.68	6	0.00

To illustrate first the behavior of the POMDP-based algorithm, Fig. 1 plots an example of time evolutions of the belief vector values for the interference states of spectrum block B1. In the upper part, the time instants when B1 is assigned to link #2 are presented for two consecutive sessions of this link. Link #1 is inactive during this example. Although not shown here, in this example the real interference state of B1 has remained at 0 during all the time of link #2 activity. Therefore, during the two periods in which link #2 is active and thus observations are performed, the belief $b_{1,0}(t)$ of state 0 remains at a high value $p_{0,0}^1 = 0.994$ because, knowing after an observation that the real interference state is 0, it is very likely that it will continue at the same state. Similarly, the belief values $b_{1,1}(t)$ and $b_{1,2}(t)$ of states 1 and 2 are kept at very low values associated to the low transition probability of moving from state 0 to states 1 and 2. The belief values remain constant until approximately 110s, when link #2 is released and thus no more observations are made. At this stage, $b_{1,0}(t)$ starts to decrease, reflecting that, as the elapsed time since the last observation increases, it progressively becomes less likely that B1 still remains in state 0. Similarly, $b_{1,1}(t)$ and $b_{1,2}(t)$ progressively increase. At around 180s link #2 becomes again active. At this point, the belief value $b_{1,0}(t)$ is still at a high value of around 0.7, meaning that assigning B1 is still a good option since it is likely that it will keep the state of the last observation. Correspondingly, B1 is assigned and the belief value $b_{1,0}(t)$ increases again to 0.994 and follows a similar behavior than at the beginning of the execution.

From the point of view of performance, Table II presents the results in terms of the average reward achieved by each link, the total system reward and the satisfaction probability (i.e. the probability that the achieved bit rate is above the requirement). The POMDP approach is compared to a random

selection among the available blocks. It can be observed that the POMDP approach achieves a significant improvement of 25% with respect to the simpler random spectrum selection in terms of rewards and satisfaction probabilities.

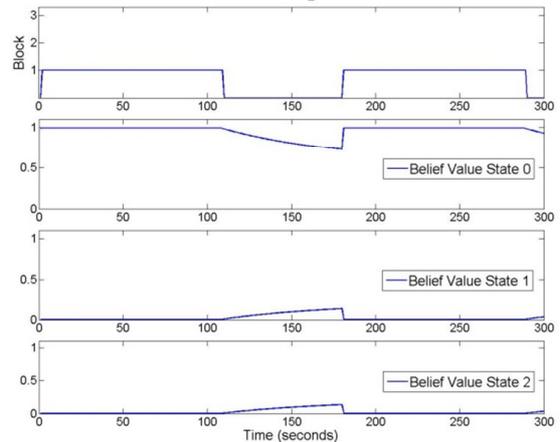


Figure 1. Example of time evolution of the belief vector values for the states of spectrum block B1 and instants when B1 is assigned to link #2.

TABLE II. COMPARISON BETWEEN POMDP AND RANDOM SELECTION

		Reward	Total Reward	Satisfaction (%)
POMDP	Link 1	0.84	1.55	90
	Link 2	0.71		86
Random	Link 1	0.69	1.24	71
	Link 2	0.55		70

IV. CONCLUSIONS

This paper has presented a novel POMDP-based framework for multi-band spectrum selection in wireless CR networks. It operates in a centralized way using the statistical characterization of the interference variations in the available spectrum blocks. The use of POMDPs allows an efficient operation in spite of performing only partial observations of some spectrum blocks. The proposed approach inherently considers heterogeneity in the QoS requirements of the different applications to be established by maximizing a reward function that considers the different suitability of each spectrum block to each radio link/application. Initial results reflect gains of 25% with respect to a random selection.

REFERENCES

- [1] J. Mitola III, "Cognitive radio: an integrated agent architecture for software defined radio," Ph.D. dissertation, KTH Royal Institute of Technology, 2000.
- [2] I.F. Akyildiz, W.-Y. Lee, M.C. Vuran, S. Mohanty, "Next generation/dynamic spectrum access/cognitive radio wireless networks: a survey", Comput. Networks (Elsevier) 50 (13) (2006) 2127–2159
- [3] J. Vartiainen, M. Hoyhtya, J. Lehtomaki, and T. Braysy, "Priority channel selection based on detection history database," CROWNCOM 2010, June 2010.
- [4] Y. Li, Y. Dong, H. Zhang, H. Zhao, H. Shi, and X. Zhao, "QoS provisioning spectrum decision algorithm based on predictions in cognitive radio networks," WiCOM 2010, Sept. 2010.
- [5] W.-Y. Lee and I. Akyildiz, "A spectrum decision framework for cognitive radio networks," IEEE Transactions on Mobile Computing, vol. 10, 2011, pp. 161–174.
- [6] K.P. Murphy, "A Survey of POMDP Solution Techniques", available at <http://http.cs.berkeley.edu/~murphyk/Papers/pomdp.ps.gz>, 2000.
- [7] Q. Zhao, L. Tong, A. Swami, Y. Chen, "Decentralized Cognitive MAC for Opportunistic Spectrum Access in Ad Hoc Networks: A POMDP Framework", IEEE Journal on Selected Areas in Communications, Vol. 25, No. 3, Apr. 2007.