RATE ADAPTATION IN COGNITIVE RADIO LINKS WITH TIME-VARYING CHANNELS

Álvaro Gonzalo-Ayuso and Jesús Pérez

University of Cantabria
Department of Communication Engineering
Santander-39140, Spain

ABSTRACT

In this work we address the optimal rate adaptation problem of a cognitive radio (CR) link in time-variant channels. A secondary user (SU) link detects an idle channel and starts the transmission with the goal of transmitting a given amount of data packets. During the transmission the transmitter dynamically adapts the frames rate, from a finite number of available rates, according to the channel state. If a frame is decoded with error, the corresponding data must be retransmitted in further frames. If a primary user (PU) access the channel during the process, the CR link immediately stops the transmission. The rate adaptation problem is formulated as an infinite-horizon Markov decision process (MDP). We split the problem in a finite number of much simpler MDP problems that can be efficiently solved by conventional MDP solving algorithms. So, the optimal policy and the corresponding maximum probability of successful transmission can be easily obtained.

Index Terms— Cognitive radio (CR), Markov decision processes (MDP), rate control

1. INTRODUCTION

Recently, IEEE 802.22 working group has released the first cognitive radio (CR) standard for wireless regional area networks [1]. This standard supports rate adaptation using adaptive modulation and coding. It also allows secondary users (SU) to support frames retransmission through an automatic repeat request (ARQ) mechanism so that SU's can setup ARQ enabled connections.

This work focuses on opportunistic spectrum access (OSA) in hierarchical CR networks where the SU's only use the licensed spectrum when primary users (PU) are not transmitting. We consider a single SU link which makes its own decision on the spectrum access strategy, based on local observation of the spectrum dynamics. We assume that the SU's can adapt the transmission rate according to the channel fading dynamics and the PU's channel access statistics. We also assume that the SU's support ARQ protocol, so when a frame is decoded with error, its data is retransmitted in a further frame.

Rate adaptation of SU links in CR has been widely addressed in the technical literature, [2], [3], [4]. However, none of the above works consider frames retransmission. In [5] frames retransmission was taken into account, but assuming time-invariant channel. To the best of our knowledge, optimal rate adaptation while considering retransmissions of failed frames over time-varying channels has not been addressed so far in the context of OSA.

We formulate this rate adaptation problem as an infinite-horizon finite Markov decision process (MDP) [6], [7], [8]. From the Bellman's equation, we decompose the problem into a finite number of much simpler MDP problems so the optimal rate adaptation policy can be easily obtained by any conventional dynamic programming (DP) algorithm, as well as the probability of successful transmission.

The remaining of this paper is organized as follows. First, we present the system model in section 2. Section 3 briefly review the time-invariant channel case and in section 4 we present the its generalization to the time-variant case. Section 5 presents numerical result to support the proposed model and finally section 6 presents the conclusions of this work.

2. SYSTEM MODEL

Let us consider a SU link that periodically senses the spectrum band. Once it detects an idle channel, it starts the transmission with the goal to transmit a fixed-size file, comprising N packets. During the transmission, the SU adapts the transmission rate aiming to maximize the probability of transferring the entire file before a PU reclaims the channel.

We assume that all the packets have the same length and that each one is encapsulated in single frame. We consider K different types of frames characterized by its rate and its duration $t_C(k)$. We consider a conventional ARQ mechanism to overcome transmission errors. Once the receiver receives a frame, it sends an ACK (acknowledgement) packet back to the transmitter through an instantaneous error-free feedback

This work was supported by the Spanish Government, Ministerio de Ciencia e Innovación (MICINN), under projects COSIMA (TEC2010-19545-C04-03) and COMONSENS (CSD2008-00010, CONSOLIDER-INGENIO 2010).

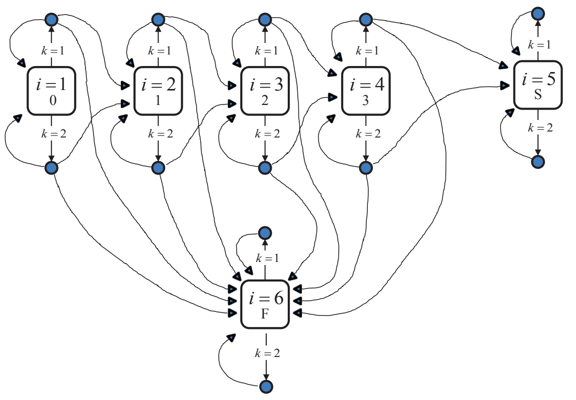

Fig. 1. Transition graph in the case of $K = 2$ available rates

channel to inform whether the frame has been correctly decoded or not.

We use $\beta(k)$ to denote the probability that the PU's do not access the channel during the transmission of a frame of length $t_C(k)$. Notice that we are considering a memoryless channel in the sense that $\beta(k)$ solely depends on the frame duration and not on the lapsed time since the last PU access.

3. TIME-INVARIANT CHANNEL

Assuming a SU non-fading channel, the frame error rate (FER) will solely depend on the type of frame. It will be denoted by $p_e(k),\ k = 1, ..., K$.

3.1. Problem formulation as a MDP:

- **Controls:** The controls are the available rates: $k \in \{1, 2, ..., K\}$.

- **States:** There are $N_S = N + 2$ possible states, that are indexed and classified as follows

  - Transient states: $1 \leq i \leq N$.
  - Success state (S): $i = N + 1$
  - Fail state (F): $i = N + 2$

  Each transient state is defined by the number of packets successfully transmitted during the process, so the system is in state i when $i - 1$ packets have been already transmitted. The success state (S) is an absorbing state and corresponds to the situation where all packets have been transmitted, whereas the system falls in the absorbing fail state (F) when a PU has reclaimed the frequency band before all packets have been transmitted. To illustrate this, figure 1 shows the transition graph for $K = 2$ and $N = 4$.

- **Transition probabilities:** There are three types of transitions: 1) transitions from a transient state to itself when the transmitted frame has been decoded with error, 2) transition from a transient state to another transient state or to the success state when the frame has been successfully transmitted, 3) transitions from a transient state to the fail state when a PU reclaims the channel. Therefore, the transition probability from state i to j when control k is applied is

$$p_{i,j}^k = \begin{cases} 1, & i = j > N \\ \beta(k)p_e(k), & i = j \leq N \\ 1 - \beta(k), & i \leq N \ \wedge \ j = N + 2 \\ \beta(k)(1 - p_e(k)), & i \leq N \ \wedge \ j = i + 1 \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

- **Rewards:** We define the transition reward from state i to j when control k is applied as follows

$$r_{i,j}^k = \begin{cases} 1, & i = N \ \wedge \ j = N + 1 \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

  In words, there is not reward until all blocks have been successfully decoded. The expected immediate reward, when the system is in state i and control k is applied, will be

$$q_i^k = \sum_{j=1}^N p_{i,j}^k r_{i,j}^k = \begin{cases} \beta(k)(1 - p_e(k)), & i = N \\ 0, & i \neq N. \end{cases}$$

- **Policies:** A policy is defined by a column vector $\mathbf{d} = [d_1, d_2, ..., d_N]^T$, where entry d_i denotes the control to be used when the system is in state i and the superscript T denotes transpose.

3.2. Optimal policy for maximum probability of successful transmission

In [5], it is shown that the policy that maximizes the probability of transmit the entire file (optimal policy $\mathbf{d}^*$) is stationary and given by

$$d_i^* = \arg\max_k a(k), \forall i \tag{3}$$

where

$$a(k) = \frac{\beta(k)(1 - p_e(k))}{1 - \beta(k)p_e(k)}, \quad k = 1, 2, \ldots, K. \tag{4}$$

4. TIME-VARYING CHANNEL

The following subsection presents the model for the time-variant SU link channel.

4.1. Channel model

We consider a point-to-point frequency-flat block-fading channel, where the channel remains constant during the transmission of a frame, and can change for consecutive frames. Let γ denote the channel power gain process. In general, γ at different frames are time correlated, and the amount of correlation depends on the frames duration.

To model the dynamics of γ we consider K discrete-time first-order Markov chains; each one with time discretized to $t_C(k), k = 1, 2, \ldots, K$. The fading range $0 \leq \gamma < \infty$ is discretized into M regions so that the m-th region is defined as $R_m = \{\gamma : A_m \leq \gamma < A_{m+1}\}$, where $A_1 = 0$ and $A_{M+1} = \infty$.

Let us consider the discrete random process $\{\gamma_d\}$, defined as the channel state during the frames transmission. For a given frame, the channel is in state $\gamma_d = m$ if $\gamma \in R_m$ during the frame transmission. The transition probability from channel state m to channel state n during the transmission of a frame of length $t_C(k)$ is denoted by $t^k_{m,n}$.

In the technical literature (see [9] and references therein), there are several models to analytically obtain these transition probabilities as function of the number of channel states M, the intervals limits $\{A_m\}$ and the channel normalized Doppler frequency f_D. The later determines the rate of variation of the channel with respect to the frames duration. Although the physical wireless channel is inherently non-Markovian, it has been shown that stationary first-order Markov chains can capture the essence of the channel dynamics when the number of regions/states is low and the channel fades slow enough [9]. Now, the FER's will depend not only on the frame rate but also on the channel state. Accordingly, hereafter $p_e(k,m)$ will denote the FER when a frame of rate k is transmitted and the channel state is m.

4.2. Formulation as a MDP

- **Controls:** The controls are the available rates.
- **States:** We consider a two-dimensional state space (as depicted in figure 2) where the states components (m,i) denote the channel state and the number of packets already transmitted, respectively. The total number of states is $N_S = M(N+2)$.
- **Transition probabilities:** The probability of transitioning from state (m,i) to state (n,j), when a frame of type k is transmitted, is

$$p^k_{(m,i),(n,j)} = \qquad (5)$$
$$t^k_{m,n} \cdot \begin{cases} 1, & i = j > N \\ \beta(k) p_e(k,m), & i = j \leq N \\ 1 - \beta(k), & i \leq N \ \wedge \ j = N+2 \\ \beta(k)(1 - p_e(k,m)), & i \leq N \ \wedge \ j = i+1 \\ 0, & \text{otherwise.} \end{cases}$$

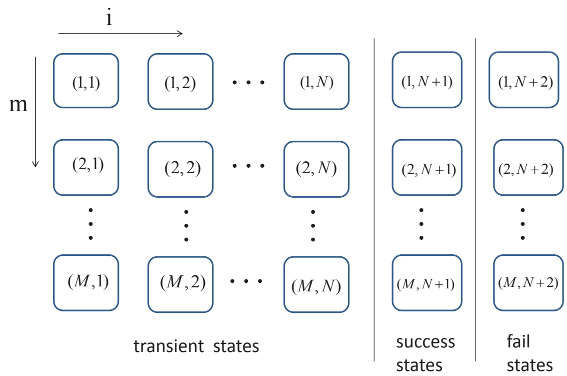

Fig. 2. Two-dimensional states arrangement.

Note that, unlike the time-invariant case, the FER, $p_e(k,m)$ depends on both, the type of frame end the channels state.

- **Rewards:** We consider the following transition rewards

$$r^k_{(m,i),(n,j)} = \begin{cases} 1, & i = N \ \wedge \ j = N+1 \\ 0, & \text{otherwise.} \end{cases} \qquad (6)$$

where $r^k_{(m,i),(n,j)}$ denotes the transition reward from state (m,i) to state (n,j) when a frame of type k is transmitted. So there is no reward until the entire file has been successfully transmitted.

Considering (5) and (6) the expected immediate reward when the system is in state (m,i) and a frame of type k is transmitted is

$$q^k_{m,i} = \begin{cases} \beta(k)(1 - p_e(k,m)), & i = N \\ 0, & \text{otherwise.} \end{cases} \qquad (7)$$

- **Policies:** We write $d_{m,i} = k$ to denote that the policy assigns control k to the state (m,i). Note that policy values $d_{m,N+1}$ and $d_{m,N+2}$ are not relevant because, when the system is in that states, the transmission has already finish or a PU has reclaimed the channel.

4.3. Probability of success for a given policy

In the following we derive a simplified algorithm to compute the probability of successful transmission for a fixed policy, that is, the value vector of the policy.

We define the policy vector

$$\mathbf{d} = \left[(\mathbf{d}_1)^T, (\mathbf{d}_2)^T, \ldots, (\mathbf{d}_{N+2})^T \right]^T, \qquad (8)$$

where $\mathbf{d}_i = [d_{1,i}, d_{2,i}, \ldots, d_{M,i}]^T$, $i = 1, \ldots, N$. Policies for states with the same i (same number of packets transmitted) are grouped in vectors $\mathbf{d}_i$.

We define the immediate reward vector analogously

$$\mathbf{q}^{\mathbf{d}} = \left[\mathbf{0}^T, \ldots, \mathbf{0}^T, \left(\mathbf{q}_N^{\mathbf{d}_N}\right)^T, \mathbf{0}^T, \mathbf{0}^T\right]^T, \qquad (9)$$

where $\mathbf{0}$ denotes the all-zeros $M \times 1$ vector and

$$\mathbf{q}_N^{\mathbf{d}_N} = \left[q_{1,N}^{d_{1,N}}, q_{2,N}^{d_{2,N}}, \ldots, q_{M,N}^{d_{M,N}}\right]^T \qquad (10)$$

The values of each policy are also grouped in a vector as

$$\mathbf{v}^{\mathbf{d}} = \left[\left(\mathbf{v}_1^{\mathbf{d}}\right)^T, \left(\mathbf{v}_2^{\mathbf{d}}\right)^T, \ldots, \left(\mathbf{v}_N^{\mathbf{d}}\right)^T, \mathbf{0}^T, \mathbf{0}^T\right]^T, \qquad (11)$$

where $\mathbf{v}_i^{\mathbf{d}} = \left[v_{1,i}^{\mathbf{d}}, v_{2,i}^{\mathbf{d}}, \ldots, v_{M,i}^{\mathbf{d}}\right]^T$, $i = 1, \ldots, N$. Notice that, the values of states with identical i (same number of packets transmitted) are grouped in vector $\mathbf{v}_i^{\mathbf{d}}$ and that the values for the absorbing states are always zero, for all policies.

From (5), the transition matrix for any policy $\mathbf{d}$, $\mathbf{P}^{\mathbf{d}}$, is a block sparse upper triangular matrix with the following structure

$$\begin{bmatrix}
\mathbf{P}_0^{\mathbf{d}_1} & \mathbf{P}_1^{\mathbf{d}_1} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{P}_2^{\mathbf{d}_1} \\
\mathbf{0} & \mathbf{P}_0^{\mathbf{d}_2} & \mathbf{P}_1^{\mathbf{d}_2} & \cdots & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{P}_2^{\mathbf{d}_2} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{P}_0^{\mathbf{d}_N} & \mathbf{P}_1^{\mathbf{d}_N} & \mathbf{0} & \mathbf{P}_2^{\mathbf{d}_N} \\
\mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{T}^{\mathbf{d}_{N+1}} & \mathbf{0} \\
\mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{T}^{\mathbf{d}_{N+2}}
\end{bmatrix},$$

where the blocks are $M \times M$ matrices given by

$$\begin{aligned}
\mathbf{P}_0^{\mathbf{d}_i} &= \mathbf{B}^{\mathbf{d}_i}\mathbf{P}_e^{\mathbf{d}_i}\mathbf{T}^{\mathbf{d}_i}, \\
\mathbf{P}_1^{\mathbf{d}_i} &= \mathbf{B}^{\mathbf{d}_i}(\mathbf{I} - \mathbf{P}_e^{\mathbf{d}_i})\mathbf{T}^{\mathbf{d}_i}, \\
\mathbf{P}_2^{\mathbf{d}_i} &= (\mathbf{I} - \mathbf{B}^{\mathbf{d}_i})\mathbf{T}^{\mathbf{d}_i}, \\
\mathbf{T}^{\mathbf{d}_i} &= [t_{m,n}^{d_{i,m}}]_{M \times M}, \\
\mathbf{B}^{\mathbf{d}_i} &= \text{diag}([\beta(d_{1,i}), \ldots, \beta(d_{M,i})]), \\
\mathbf{P}_e^{\mathbf{d}_i} &= \text{diag}([p_e(d_{1,i}, 1), \ldots, p_e(d_{M,i}, M)]).
\end{aligned}$$

The Bellman equation for a policy $\mathbf{d}$ will be

$$\mathbf{v}^{\mathbf{d}} = \mathbf{q}^{\mathbf{d}} + \mathbf{P}^{\mathbf{d}}\mathbf{v}^{\mathbf{d}}. \qquad (12)$$

Considering (12) and the particular structures of $\mathbf{v}^{\mathbf{d}}$, $\mathbf{q}^{\mathbf{d}}$ and $\mathbf{P}^{\mathbf{d}}$ in our problem, the computation of the policy value vector can be split in N iterations as follows

$$\begin{aligned}
\mathbf{v}_{N+1}^{\mathbf{d}} &= \mathbf{v}_{N+2}^{\mathbf{d}} = \mathbf{0} \\
\mathbf{v}_N^{\mathbf{d}} &= \mathbf{P}_0^{\mathbf{d}_N}\mathbf{v}_N^{\mathbf{d}} + \mathbf{q}_N^{\mathbf{d}_N}, \\
\mathbf{v}_i^{\mathbf{d}} &= \mathbf{P}_0^{\mathbf{d}_i}\mathbf{v}_i^{\mathbf{d}} + \mathbf{P}_1^{\mathbf{d}_i}\mathbf{v}_{i+1}^{\mathbf{d}}, \;\; i < N.
\end{aligned} \qquad (13)$$

According to (7) and (10), $\mathbf{q}_N^{\mathbf{d}_N} = \mathbf{1} \cdot \mathbf{P}_1^{\mathbf{d}_N}$, where $\mathbf{1}$ is the all-ones vector of size $M \times 1$. Therefore, the values of a policy can be efficiently computed in a backward recursively way as follows

$$\mathbf{v}_i^{\mathbf{d}} = \mathbf{A}(\mathbf{d}_i) \cdot \begin{cases} \mathbf{v}_{i+1}^{\mathbf{d}}, & i < N \\ \mathbf{1}, & i = N \end{cases}, \;\; i = N, \ldots, 1. \qquad (14)$$

where $\mathbf{A}(\mathbf{d}_i)$ is a $M \times M$ matrix given by

$$\mathbf{A}(\mathbf{d}_i) = \left(\mathbf{I} - \mathbf{P}_0^{\mathbf{d}_i}\right)^{-1}\mathbf{P}_1^{\mathbf{d}_i}, \;\; i = 1, \ldots, N. \qquad (15)$$

The entry (m, n) of $\mathbf{A}(\mathbf{d}_i)$ is just the transition probability from the transient state (m, i) to the state $(n, i+1)$ under the policy vector $\mathbf{d}_i$, taking into account the frame retransmissions.

Note that equations (15) can be viewed as the generalizations of (4) to the case of time-varying channel.

4.4. Optimal policy

From (14), the optimal values and policy can be computed backward iteratively as follows

$$\mathbf{v}_i^* = \max_{\mathbf{d}} \; \mathbf{A}(\mathbf{d}) \cdot \begin{cases} \mathbf{v}_{i+1}^*, & i < N \\ \mathbf{1}, & i = N \end{cases}, \qquad (16)$$

$$\mathbf{d}_i^* = \arg\max_{\mathbf{d}} \; \mathbf{A}(\mathbf{d}) \cdot \begin{cases} \mathbf{v}_{i+1}^*, & i < N \\ \mathbf{1}, & i = N \end{cases}$$

Each iteration of (16) can be viewed as a simple M-states MDP problem

$$\tilde{\mathbf{v}}^* = \max_{\tilde{\mathbf{d}}} \; \tilde{\mathbf{P}}^{\tilde{\mathbf{d}}}\tilde{\mathbf{v}}^* + \tilde{\mathbf{q}}^{\tilde{\mathbf{d}}}, \qquad (17)$$

where $\tilde{\mathbf{d}}$, $\tilde{\mathbf{q}}^{\tilde{\mathbf{d}}}$ and $\tilde{\mathbf{v}}^{\tilde{\mathbf{d}}}$ are $M \times 1$ vectors and $\tilde{\mathbf{P}}^{\tilde{\mathbf{d}}}$ are $M \times M$ transition matrices. Moreover, the transition matrices are the same for all iterations; only the vectors of immediate rewards $\tilde{\mathbf{q}}^{\tilde{\mathbf{d}}}$ changes from one iteration to another as follows

$$\tilde{\mathbf{q}}_i^{\tilde{\mathbf{d}}} = \begin{cases} \mathbf{P}_1^{\tilde{\mathbf{d}}}\mathbf{v}_{i+1}^{\tilde{\mathbf{d}}}, & i < N \\ \mathbf{q}_i^{\tilde{\mathbf{d}}}, & i = N \end{cases} \qquad (18)$$

Therefore, in each iteration, the optimal policy and optimal values can be efficiently computed with any conventional DP algorithms thanks to the reduced number of channel states.

5. NUMERICAL RESULTS

This section shows some numerical simulations to illustrate the benefits of the rate adaptation scheme. We consider $K = 3$ available rates and $M = 4$ channel states. The corresponding FER's, $p_e(k, m)$, are shown in table 1. We model the PU's

Action	$\mathbf{m = 1}$	$\mathbf{m = 2}$	$\mathbf{m = 3}$	$\mathbf{m = 4}$
1	0.0537	0.0151	0.0033	0.0006
2	0.1987	0.0917	0.0316	0.0083
3	0.2858	0.1785	0.0927	0.0398

Table 1. Frame error probabilities for different rates and channel states (m).

access to the channel as a Poisson process with access rate λ, so the probability that the PU's do not access the channel during the transmission of a frame of type k is $\beta(k) = \exp^{-\lambda t(k)}$.

Figure 3 shows the probability of successful transmission (of transmiting the entire file) as a function of the number of packets N, assuming $\lambda = 5$, $M = 4$ channel states and maximum Doppler frequency $f_D = 20Hz$. The frames duration are $t_C(1) = 4ms$, $t_C(2) = 3ms$ and $t_C(3) = 2ms$. The figure shows the performance of the rate adaptation scheme when the optimal policy is follow (d^*), and the static policies where the same type of frame (d = 1, 2 or 3) is always used. It shows significant improvement when using rate adaptation, the optimal policy $d*$ behaves at least as good as the best static policy.

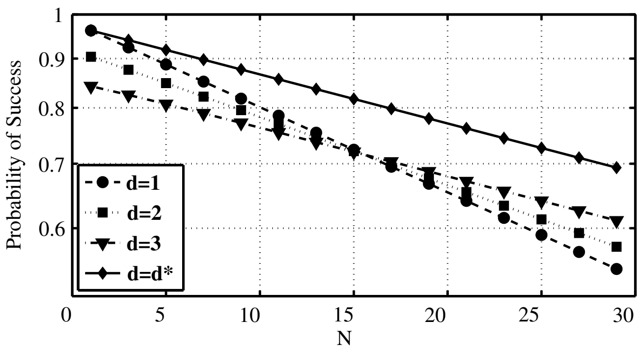

Fig. 3. Probability of success as a function of the number of packets for different policies.

6. CONCLUSIONS

In this work we have studied the rate adaptation problem of a SU link in hierarchical CR networks from a cross layer perspective. Unlike other related works, we have considered that the link channel is time-varying, and we have taken into account the retransmission of erroneous frames. The SU opportunistically access the channel with the goal to transmit a given number of packets (data file) during a sojourn time of the PU's idle state.

We only consider a single SU link operating, taking multiple SU's competing to access the channel into account would require a more complex MDP model. Nonetheless, we suggest that our scheme could be used in simple devices with low requirements in terms of delay, for example some wireless sensors.

The adaptation problem have been formulated as a MDP problem with a two-dimensional state set. Considering the specific structure of the transition probabilities, we have simplified the original (an complex) MDP problem by splitting it in a number of simple MDP problems with much less number of states, that can be easily solved with any conventional DP algorithm.

7. REFERENCES

[1] C. R. Stevenson, G. Chouinard, Z. Lei, W. Hu, S. J. Shellhammer, and W. Caldwell, "Ieee 802.22: The first cognitive radio wireless regional area network standard," *IEEE Communications Magazine*, vol. 47, pp. 130–138, January 2009.

[2] C. C. Chai, "On power and rate adaptation for cognitive radios in an interference channel," in *Proc. 71th IEEE Vehicular Technology Conference, PIMRC*, Taipei, Taiwan, May 2010, vol. 2, pp. 1–5.

[3] L. Gao and S. Cui, "Power and rate control for delay-constrained cognitive radios via dynamic programming," *IEEE Transactions on Vehicular Technology*, vol. 58, pp. 4819–4827, November 2009.

[4] H. H. J. Wang, J. Zhu, and S. Li, "Optimal policy of cross-layer design for channel access and transmission rate adaptation in cognitive radio networks," *EURASIP Journal on Advances in Signal Processing*, vol. 2010, 2010.

[5] J. Pérez and M. Khodaian, "Optimal rate and delay performance in non-cooperative opportunistic spectrum access," in *9th International Symposium on Wireless Communication Systems (ISWCS 2012)*, Paris, France, August 2012.

[6] R. A. Howard, *Dynamic Control and Markov Processes*, M.I.T. Press, 1960.

[7] M. L. Puterman, *Markov Decision Processes. Discrete Stochastic Dynamic Programming*, John Wiley and Sons, 2005.

[8] D. P. Bertsekas, *Dynamic Programming and Optimal Control. Volume II*, Athena Scientific, 2007.

[9] P. Sadeghi, R. A. Kennedy, P. B. Rapajic, and R. Shams, "Finite-state markov modeling of fading channels," *IEEE Signal Processing Magazine*, vol. 57, pp. 57–80, September 2008.