Dynamic rate adaptation in cognitive radio considering time-dependent channel access models

Álvaro Gonzalo-Ayuso and Jesús Pérez
University of Cantabria
Department of Communication Engineering
Santander-39140, Spain
Email: {alvaro, jperez}@gtas.dicom.unican.es

Abstract—Recent studies showed that, in practical situations, the primary user (PU) access to the channel is better modeled as a time dependent random process. Taking this into account, we address the optimal rate adaptation problem of a cognitive radio (CR) link. A secondary user (SU) link detects an idle channel and starts the transmission with the goal of transmitting a given amount of data packets within a given time. During the transmission, and taking into account frame retransmission, the transmitter dynamically adapts the frames rate, from a finite number of available rates. If the PU accesses the channel, the SU immediately stops the transmission. The problem is formulated as an episodic Markov decision process (MDP). We show that selecting the best stationary policy (using the same rate for the whole transmission) can perform close to optimal.

I. 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 frame retransmission through an ARQ mechanism.

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 noncooperative spectrum sharing where each SU makes its own decision on the spectrum access strategy, based on its own delay constrains and local observation of the spectrum dynamics. We assume that the SU's can adapt the transmission rate according to the current delay constrain, the channel fading conditions 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.

Models for channel occupancy patterns commonly used to date in CR research are usually limited in scope and based on oversimplifications or assumptions. Recently, in [2] a complete study on spectrum occupancy patterns of the PU in CR links was conducted. In order to propose realistic and accurate models of such patterns, several practical networks were measured and modeled. In this work we make use of some of the proposals in [2] to model the PU access to channel as a time dependent random process. This is the main novelty of this work.

Rate adaptation of SU links in CR has been widely addressed in the technical literature, [3], [4], [5]. However,

none of the above works consider frames retransmission. In [6] frames retransmission was taken into account, but assuming a memoryless channel occupancy model and with no delay constrain. To the best of our knowledge, optimal rate adaptation while considering retransmissions of failed frames, with delay constrain and over time-dependent channel occupancy models has not been addressed so far in the context of OSA.

We formulate this rate adaptation problem as an episodic Markov decision process (MDP) [7], [8], [9]. 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 aim of this work is not to propose a practical rate adaptation scheme, we only intend to explore the possible solutions to this problem and to find theoretical performance bounds. The proposed scheme is not feasible as a practical system because it would require accurate and complete information about the channel occupancy patterns as well as the frame error rates.

The remaining of this paper is organized as follows. We present the system model in section II and in section III we introduce the MDP formulation. In IV we provide some insight on MDP problems solving and in section V we present a simple way to compute success and fail probabilities. In sectionVI we present numerical results to evaluate an compare the optimal rate adaptation with stationary policies (the frame rate is fixed for the transmission of the entire file). Finally, section VII presents the conclusions of this work.

II. SYSTEM MODEL

We consider an SU that periodically senses the channel. Once it detects that the channel is idle, it begins the transmission of a fixed size file comprising N_p packets. Each of these fixed size packets is encoded in a single frame. The SU has the capability of adapting its transmission rate, i.e. the duration of each frame.

The objective of the SU is to maximize the probability of transmitting the N_p packets within a given time, T_t, and before a PU reclaims the channel. To achieve this goal the SU has to select a transmission rate for each frame. The SU has K different types of frames available, each one with duration T_k and frame error rate (FER) p_k, where $k \in \{1, \dots K\}$. We assume that, the channel, the transmit power, and the SNR at the receiver remain constant during the sojourn time of the PU's idle state, i.e. the FER remains constant.
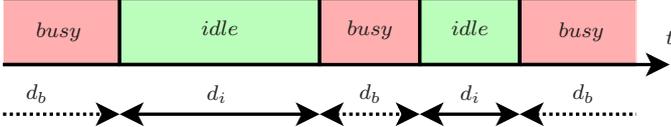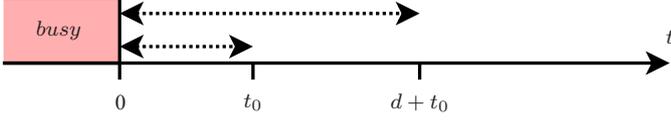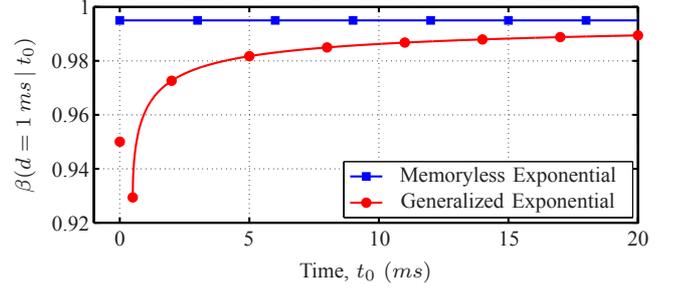
Fig. 1. Channel occupancy process.

Fig. 2. Idle period.

Fig. 3. $\beta(d = 1\,ms \mid t_0)$ for the GE distribution ($\lambda = 5$, $\mu = 0.5ms$ and $\alpha = 0.5$) and the memoryless exponential ($\lambda = 5$).

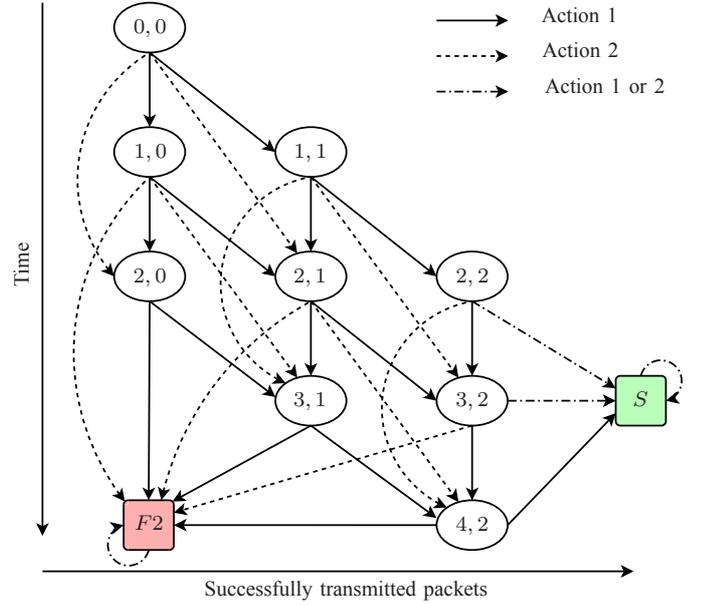Fig. 4. Example with $N_p = 3$, $N_t = 5$, $K = 2$ and $\mathbf{t} = [1, 2]$.

In the subsequent we will consider discrete time steps of duration T seconds, where T represents the g.c.d. (greatest common divisor) of the frame durations. Hence, we define the discrete time duration of each type of frame as $t_k = T_k/T$ and the total number available time slots for the transmission, N_t, as $N_t = \lfloor T_t/T \rfloor$.

We consider a traditional and ideal ARQ mechanism to detect frame transmission errors. When the receiver receives a frame, it sends back an ACK packet to the transmitter through an instantaneous error-free feedback channel to inform whether the frame has been correctly decoded or not. Whenever a frame is decoded with error, the corresponding packet must be retransmitted in a further frame.

Channel access model

Figure 1 depicts the channel occupation process. The channel state changes alternatively between idle and busy periods over time and the duration of the idle/busy periods is given by two random variables denoted by d_i and d_b respectively. Let $F_i(d_i)$ and $F_b(d_b)$ be the corresponding cumulative distribution functions (CDF). Several realistic and accurate models have been proposed for these CDFs for different radio networks [2]. Without loss of generality, in this work we modeled d_i as a generalized exponential (GE) random variable with CDF

$$F_i(d_i) = \left[1 - e^{-\lambda(d_i - \mu)}\right]^\alpha, \tag{1}$$

where $d_i \geq \mu > 0$, $\lambda > 0$ and $\alpha > 0$.

Let us consider an idle period starting at time $t_0 = 0$ (see figure 2), and let $\beta(d|t_0)$ denote the conditional probability that the channel remains idle at time $t_0 + d$ given that it was idle at time t_0. Using Bayes' theorem $\beta(d|t_0)$ is given by

$$\beta(d|t_0) = \begin{cases} \dfrac{1 - F_i\left(t_0 + d\right)}{1 - F_i\left(t_0\right)}, & t_0 > 0 \\ 1 - F_i\left(d\right), & t_0 = 0 \end{cases} \tag{2}$$

Generally, $\beta(d|t_0)$ depends on t_0, i.e. the channel access probability has memory. The well known exponential distribution is an exception to this rule. Continuing with the GE distribution example, notice that for $\alpha = 1$ and $\mu = 0$ the GE distribution becomes the exponential distribution with parameter λ. If we derive $\beta(d|t_0)$ for the exponential case

$$\beta(d|t_0) = \frac{1 - \left(1 - e^{-\lambda(t_0 + d)}\right)}{1 - \left(1 - e^{-\lambda(t_0)}\right)} = e^{-\lambda d}, \tag{3}$$

we prove that it is not dependent on t_0, i.e. it is a memoryless channel. The memoryless exponential model has been extensively used in the literature, however in many practical cases it is not a realistic model.

Figure 3 depicts the probability that the channel remains idle during the transmission of frame of duration $1ms$ at time t_0, $\beta(d = 1ms \mid t_0)$, for the GE distribution (with $\lambda = 5$, $\mu = 0.5ms$ and $\alpha = 0.5$), and for the memoryless exponential distribution (with $\lambda = 5$).

III. MDP FORMULATION

In this section we formulate the rate adaptation problem as an MDP.

A. Stages

Each stage of the problem corresponds to a frame transmission. The duration of each stage depends on the type of frame selected, t_k. The process can finish in three different ways: 1) The SU has successfully transmitted the N_p packets, 2) A PU has reclaimed the channel so the transmission has been interrupted, and 3) The allocated time, N_t, is over and the transmission has been interrupted.

TABLE I.   TRANSIENT STATES CHILDREN GENERATION.

```
% Compute child states of (i, j)

t_min = min(t)   ∀ k ∈ {1, 2, ..., K}

for k = 1, ..., K

    if i + t_k ≤ N_t - t_min · (N_p - j)
      if (i + t_k, j) ∉ S   →   S = S ∪ (i + t_k, j)
    end

    if i + t_k ≤ N_t - t_min · (N_p - (j + 1))
      if j + 1 < N_p
        if (i + t_k, j + 1) ∉ S →   S = S ∪ (i + t_k, j + 1)
      end
    end

end
```

B. Actions or controls

The available actions are the K transmission rates. Each action $k \in \{1, 2, \ldots, K\}$ is associated with a frame error probability, p_k, and a frame duration t_k.

C. States

We will relay on a simple example for an easier statement of the problem. Let us consider that we intend to transmit 3 data packets, $N_p = 3$, in not more than 5 time slots, $N_t = 5$ and to do so we have available 2 different types of frames, $K = 2$, with durations of 1 and 2 time slots, $\mathbf{t} = [1, 2]$, and FERs, p_1 and p_2, $\mathbf{p} = [p_1, p_2]$. Figure 4 shows the state transitions diagram. The system states can be classified in two types, transient states which are represented by labeled ellipses, and final states which we represent with rounded labeled squares. The arrows represent the possible transitions between states. For clarity, transitions to terminal state $F1$ are not shown in figure 4 but are possible from any transitional state regardless of the action being taken.

Transient States: Each transient state, (i, j), is described by two parameters, the elapsed time since the PU liberated the channel, $0 \leq i \leq N_t$, and the number of packets successfully transmitted, $0 \leq j \leq N_p$. Regarding the example in figure 4, we can look at the transient states as if they were in a $N_p \times N_t$ grid. States in the same row share the same elapsed time, i, while states in the same column share the same number of successfully transmitted packets, j.

Notice that, given the time and number of packets restrictions as well as the available frames durations, not all the pairs (i, j) are possible states. This leads to a finite state space $\mathcal{S}$. For instance, in our example, state $(3, 0)$ is not a valid transient state because it is not possible to transmit 3 packets when there is only 2 time steps left.

Figure 5 shows all possible transitions from a given state (i, j), the algorithm in Table I describes how the child states of state (i, j) can be obtained. Beginning with state $(0, 0)$, the complete set of transient states can be generated recursively in this manner.

Final States: We consider the following final states:

- *Busy Channel* ($F1$): A PU has accessed the channel, the SU stops the transmission so it does not interfere with the PU. All the previously transmitted packets are considered to be lost.
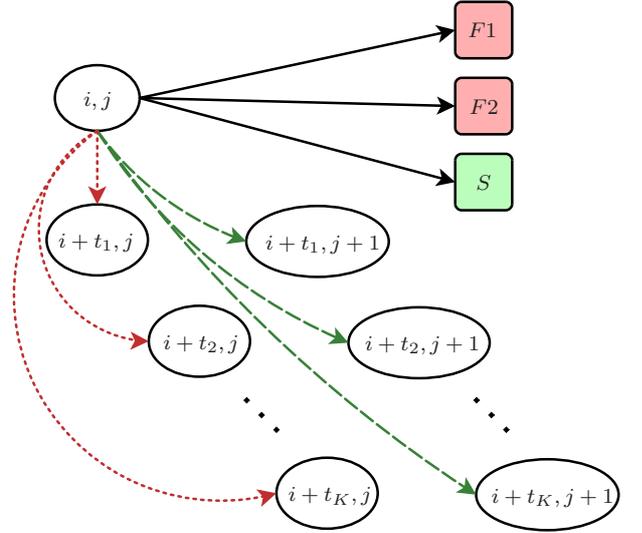

Fig. 5.   Possible transitions from state (i, j).

- *Time-out* ($F2$): Not all the N_p packets have been (or can be) transmitted before the time runs out. State $F1$ groups all those (i, j) pairs in which time-out is inevitable.

- *Success* (S): The N_p packets have been successfully transmitted within the allocated time.

The final states are absorbing, so once the system falls in a final state it remains in it indefinitely. Notices that the set of available actions may not be the same for all the states. In our example, action 2 is not allowed in state $(4, 2)$ because there is not enough time available to take action 2. For a given transient state only actions than can lead to another transient state or to the successful state are allowed.

D. Transition Probabilities $\mathcal{T}$

We denote the probability of transitioning from state s to state s' when action k is perform by $\tau(s, s', k) = P_r(s'|s, k)$.

Transient States: From state (i, j) the following transitions have to be considered (any other transition has probability zero):

- If the channel remains idle, the transmission of a frame of type k fails, and $(i + t_k, j) \in \mathcal{S}$, then:

$$P_r((i + t_k, j)|(i, j), k) = \beta(t_k \cdot T|i \cdot T)p_k. \quad (4)$$

On the other hand, if $(i + t_k, j) \notin \mathcal{S}$, then it means a time-out is inevitable and therefore:

$$P_r(F2|(i, j), k) = \beta(t_k \cdot T|i \cdot T)p_k. \quad (5)$$

- If a frame of type k is transmitted successfully and the destination state $(i + t_k, j + 1) \in \mathcal{S}$, then

$$P_r((i + t_k, j + 1)|(i, j), k) = \beta(t_k \cdot T|i \cdot T)(1 - p_k). \quad (6)$$

However, if $j + 1 = N_p$, the destination state is S, and therefore:

$$P_r(S|(i, j), k) = \beta(t_k \cdot T|i \cdot T)(1 - p_k). \quad (7)$$

- If the PU accesses the channel:

$$P_r(F1|(i,j),k) = 1 - \beta(t_k \cdot T|i \cdot T) \quad (8)$$

Final States: As its name implies, the only possible transition is to stay in the same state with probability 1 regardless of the chosen action:

$$P_r(s|s,k) = 1, \quad (9)$$

for $s \in \{F1, F2, S\}$ and $k \in \{1, \ldots, K\}$.

E. Rewards $\mathcal{R}$

Every time an action k is perform, and, as a result a transition between states s and s' occurs, a reward $R_t(s, s', k)$ is received. The reward is 0 for all transitions except those that reach the success state from the last column of the state diagram, $(i, N_p - 1)$, in such case, the assigned reward is 1. In our example, only transitions from states $(2,2)$, $(3,2)$ and $(4,2)$ to S are rewarded with 1.

IV. Solving the Markov Decision Process

Solving an MDP implies finding a policy for the decision maker, in this case the SU link, a function $\pi(s)$ that specifies the action to be taken in each state s. The goal is to find the optimal policy, $\pi(s)^*$, that maximizes the cumulative reward over a potentially infinite horizon of time steps. Associated to a policy $\pi(s)$, there is always a value function, $V(s)^{\pi(s)}$, which is defined as the cumulative reward expected to be earn when starting in state s and following policy π thereafter.

Solving an MDP is finding an optimal policy $\pi(s)^*$, which maximizes $V(s)$ so $V^{\pi(s)^*} \geq V^\pi$ for every state and for any other possible π. The most common approach for solving MDPs are the Dynamic Programing algorithms (DP). Among this family of algorithms we chose value iteration, [10], for our simulations.

In the previous section, we set all the rewards in our model to 0 except those associated with transitions to final state S from nonterminal states, which were given 1. Under this particular choice of rewards [10], the value function $V^\pi(s)$ equals the probability of reaching state S from state s following policy π.

Going back to the example in figure 4, and, regarding that the process always starts in state $(0,0)$, it is easy to see that $V((0,0))$ is the probability of reaching S. When we obtain the optimal policy, what we are really doing is maximizing the value function and hence, the probability of reaching final state S.

V. Probability of the final states

Our interest is on computing the probabilities $Pr(F1)$, $Pr(F2)$ and $Pr(S)$ of reaching states $F1$, $F1$ and S, after the N_t time steps for a general policy, and particularly for the optimal one. To describe how this probabilities are computed we will again relay on the example in figure 4. Traveling column wise down and then right we associate an index to each one of the transient states. For instance, state $(0,0)$ is number 1 and state $(4,2)$ is state 9. We also index final states $F1$, $F1$ and S as 10, 11 and 12 respectively.

Let P^π denote the 12×12 state transition probabilities matrix whose entries are $\tau_{s \to s'}^\pi = Pr(s'|s, \pi(s))$ for a given policy π:

$$\mathbf{P}^\pi = \begin{bmatrix} 0 & \tau_{1\to2}^\pi & \tau_{1\to3}^\pi & \cdots & \tau_{1\to F1}^\pi & 0 & 0 \\ 0 & 0 & \tau_{2\to3}^\pi & \cdots & \tau_{2\to F1}^\pi & \tau_{2\to F2}^\pi & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & 1 \end{bmatrix} \quad (10)$$

Then, it is possible to compute the probability of being in each one of the states after the N_t time steps as:

$$[\mathbf{P}^\pi]^{N_t} = \begin{bmatrix} 0 & 0 & \cdots & Pr(F1) & Pr(F2) & Pr(S) \\ 0 & 0 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 \end{bmatrix} \quad (11)$$

The probabilities or reaching each one of the states after the N_t time steps is given by the corresponding elements of the first row of (11). Since we have a time limit of N_t time steps, after that time the process can only be in one of the 3 final states.

VI. Numerical Results

In this section we present numerical results to illustrate the performance of the optimal rate adaptation scheme. We choose the GE channel access model in figure 3 ($\lambda = 5$, $\mu = 0.5ms$ and $\alpha = 0.5$) and we consider $K = 3$ available rates. The frame durations are 4, 3 and $2ms$ (hence $\mathbf{t} = [4, 3, 2]$ and T=$1ms$) and the corresponding frame error rates are $\mathbf{p} = [0.001, 0.01, 0.3]$.

Figure 6 shows the probability of successful transmission as a function of the number of packets that conform the file, N_p, assuming that the available time is $T_t = 120ms$. The probability is evaluated for the $K = 3$ possible stationary policies, (always transmit the same type of frame) as well as for the optimal policy π^*. Clearly, the time restriction is very loose, meaning that the probability of time-out (reaching state $F2$) is, in practice 0, therefore, it behaves as if there were no time restriction at all. For this reason it is enough to represent only $Pr(S)$ since $Pr(F2) = 1 - Pr(S)$ in this case. We see how the optimal policy behaves like the stationary policy 3, this suggest that the optimal policy is stationary or at least very close to stationary.

Next, we set a much more restrictive time limit, $T_t = 60ms$. In this case the restriction is tighter, specially for frame types 1 and 2. Figure 7 shows the probabilities $Pr(F1)$, $Pr(F2)$ and $Pr(S)$ versus the number of packets. We can see how the time restriction makes reaching S impossible with stationary policies with frame types 1 or 2 for N_p greater than 16 and 21 respectively. We can also see that now the optimal policy is not stationary but it is still quite close to it.

To illustrate this fact we show a state-action diagram in figure 8 for $N_p = 20$ and $T_t = 60ms$. Each square in the diagram represents a state of the MDP and it is colored depending on the optimal type of frame according to $\pi(s)^*$,
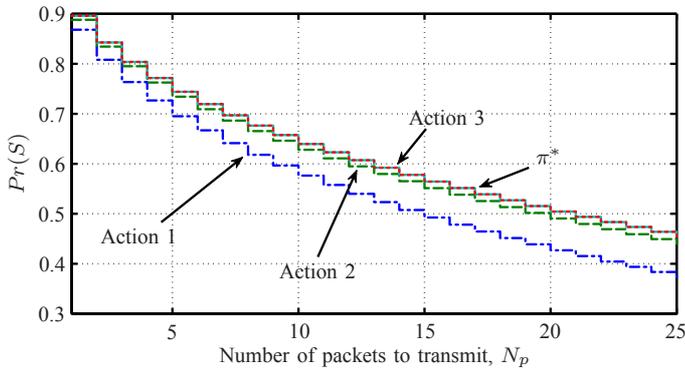
Fig. 6. Probability of Success versus number of packets to transmit, N_p, for different policies when $T_t = 120ms$.

the empty squares represent states that are not included in the model. Figure 8 shows that the optimal policy is not stationary. In [6] it was proven that memoryless channels occupancy models, such us the exponential, lead to optimal policies that are always stationary. In this work we generalize to more complex models and we also include a time restriction, this two changes lead to nonstationary optimal policies. However, figure 7 shows that the stationary policy $k = 3$ achieves close to optimal performance.

Finally, we study the effect of T_t value for a fixed $N_p = 20$. Figure 9 shows the probabilities $Pr(F1)$, $Pr(F2)$ and $Pr(S)$ as a function of the available time T_t between 30 and $90ms$. Again, the optimal policy behaves only slightly better than the best stationary policy. We have observed this kind of behavior for a range of different channel occupancy models with realistic parameters.

VII. CONCLUSIONS

In this work we have study the rate adaptation problem of SU links in hierarchical CR networks from a cross layer perspective. We consider that the SU opportunistically accesses the channel with the goal to transmit a given number of packets in a given time during the sojourn time of the PU's idle state. Unlike other related works we have adopted a more general an realistic channel access model. We consider that the probability that the channel remains idle during the transmission of a frame depends not only on the frame duration but also on the elapsed time since the PU liberated the channel.

Numerical results showed that even thought the channel memory induces not stationary policies, in practice, the actual performance degradation is negligible. Therefore, we can conclude that in practice the problem is reduce to the selection of the best stationary policy.

Stationary policies have an additional advantage, it is not necessary to solve the MDP problem and it is possible to obtain a close form expression to compute $Pr(S)$, $Pr(F2)$ and $Pr(F2)$.

We have also show how the time restriction can influence the system behavior, non the less, the optimal policy performance remains close to the best stationary policy.
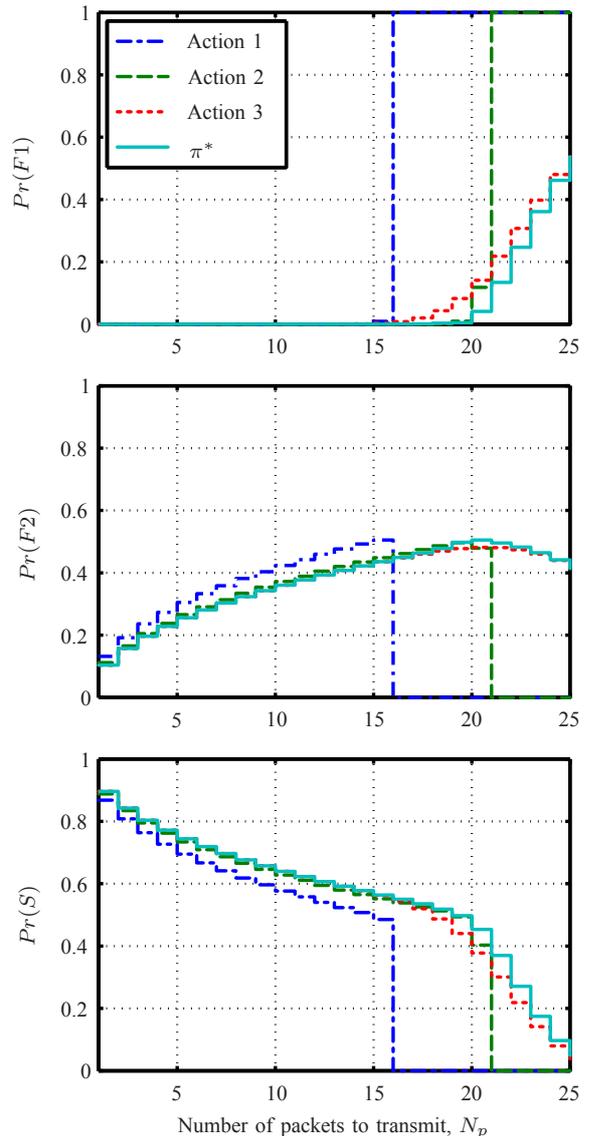

Fig. 7. Probability of success, busy channel and time-out versus number of packets to transmit for different policies when $T_t = 60ms$.

VIII. ACKNOWLEDGMENT

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] M. López-Benítez, "Spectrum usage models for the analysis, design and simulation of cognitive radio networks," Ph.D. dissertation, Universitat Politècnica de Catalunya (UPC), Barcelona, May 2011.

[3] C. C. Chai, "On power and rate adaptation for cognitive radios in an interference channel," in *Proc. 71th IEEE Vehicular Technology Conference, PIMRC*, vol. 2, Taipei, Taiwan, May 2010, pp. 1–5.
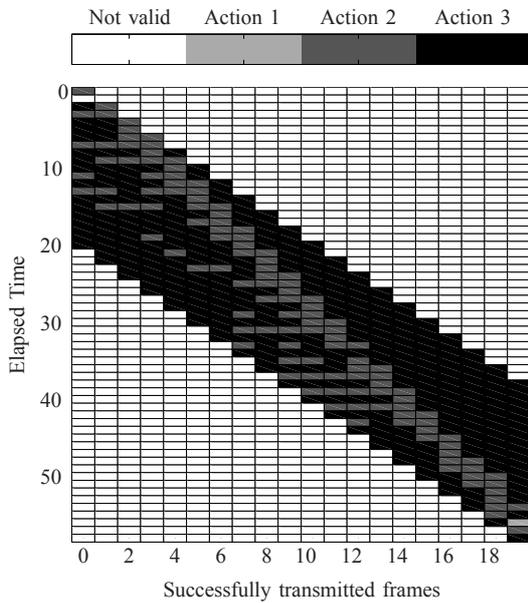
Fig. 8.   State-policy diagram.

[4] 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.

[5] 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.

[6] 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.

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

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

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

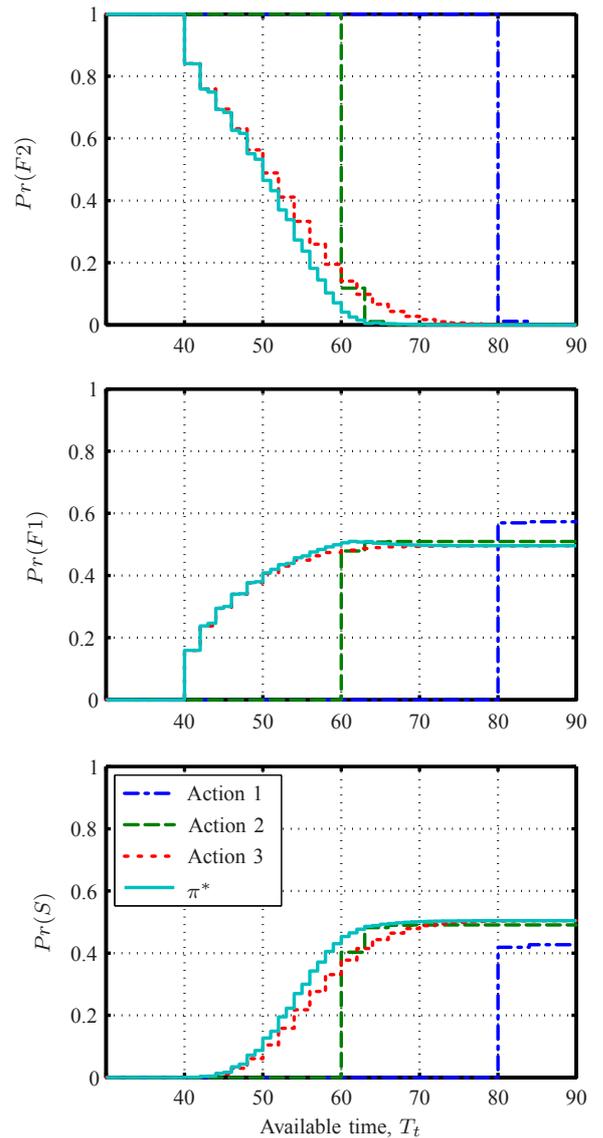[10] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. M.I.T. Press, 1998.

Fig. 9.   Probability of success, busy channel and time-out versus available time, T_t, for different policies when $N_p = 20$.