Dynamic Spectrum Access for Delay-sensitive Multimedia Applications in Cognitive Radio

Cuong T. Do, Choong Seon Hong
Email: {dtcuong, cshong}@.khu.ac.kr
Department of Computer Engineering, Kyung Hee University

Abstract

This paper investigates the channel selection strategy for secondary users' delay-sensitive multimedia applications over cognitive radio networks with heterogeneous channel capacity. We claim that in order to avoid the costly collision due to the miss-detection and false alarm of primary users, a secondary user may desire an optimal policy which maximizes its reward. In case of explicit model of the channel statistics, Markov decision process (MDP) is applied to derive the optimal channel selection strategy.

1. Introduction

Spectrum decision is an important process in CR system which helps the SU select the best channel to transmit data from candidate channels [1-2]. The author in [1] and [2] designed a channel selection scheme without negotiation for multiuser and multi-channel cognitive radio systems by using multi-agent Q-learning. The purpose of this paper is to apply Markov decision process (MDP) into spectrum access policy for delay-sensitive multimedia applications of secondary users. In our spectrum access policy, the reward is assigned to the used resource based on the result of the reward function. Cognitive radio users select spectrum resources to use based on the expected reward assigned to the spectral resources - resources with higher expected reward are chosen. Furthermore, we investigate the effect of false alarm and missed detection probability to the spectrum access policy. We will provide results and more details in the following sections.

2. System model

We assume a Cognitive Radio System where a Secondary User (i.e., the CR base station who serves the secondary devices) and multiple Primary Users are being operated in the same frequency bands.

We consider a scenario where the SU can utilize N PU channels simultaneously. We assume that the packets of delay-sensitive multimedia applications of SUs have the highest priority over SUs' packets and are named high priority packets. As a high priority packet (HPP) of the SU arrive at time slot t, the SU need to decide which channel for the HPP using as Figure 1. We assume the SU have N queue for N channel as in Fig.1. After choosing the channel for the HPP, the HPP will arrive at the head of the queue and be ready for transmitting in the next time slot due to its high priority. In the next time slot, if the selected channel is idle, the HPP is transmitted successfully. Otherwise, if the selected channel is busy, the HPP must wait since it has lower priority than PUs' packet.

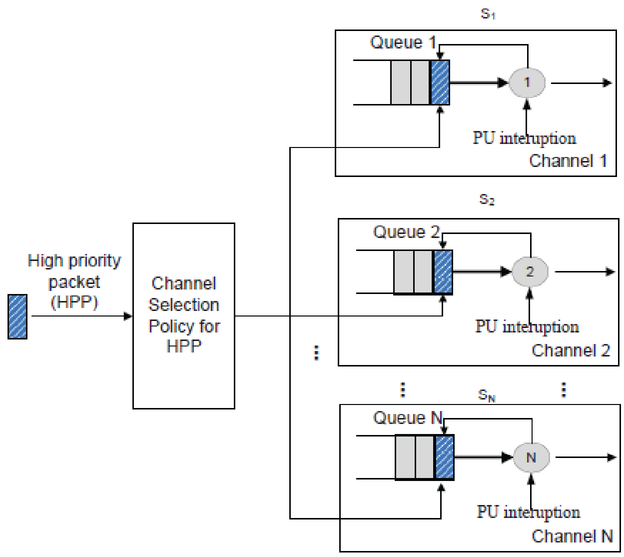

Fig.1.The Channel Selection Policy for HPP.

The Channel Selection Policy for HPP (CSPHPP) works as follows: At the beginning of each slot, when a HPP arrive at CSPHPP, CSPHPP will perform channel selection in order to maximize the HPP's long term reward. We assume that high priority traffics are light. This is reasonable assumption because the fact that the data traffics are often heavier than the real-time or high priority traffic in the networks. Therefore, we consider at each time slot, there is at most one HPP arriving at CSPHPP. Then, CSPHPP deliveries the HPP to the head of the queue of the selected channel. In the next time slot, the HPP is ready to transmit in the selected channel if it is idle. Then, at the end of the slot, the receiver acknowledges a successful data transmission. Our goal is to determine optimal strategy for the Secondary User which maximizes the long term benefit of HPPs by exploiting the decision history and the channels occupancy statistics.

The Primary Users system has priority and must not be interrupted by the Secondary User. Therefore, the Secondary User can performs the spectrum sensing procedure before it transmits data to protect the Primary Users' transmission. However, due to sensing error, the detection process is characterized by false alarm probability P_f and missed detection probability P_m. The false alarm probability is the probability that a free slot is decided to be occupied. The miss detection probability is the probability that an occupied slot is detected incorrectly as a free slot. Spectrum sensing is used to inform the SU whether the channel is busy or idle. If the channel is idle, the Secondary User can transmit a HPP successfully. Then, it gets a reward. In contrast, when channel is busy, the Secondary User cannot transmit the HPP. The spectrum handoff procedure is performed to return the channel to PUs. This kind of listen-before-talk channel access schemes has been adopted in the quite period technique of the IEEE 802.22 standard [3]. The interruptions by PUs will incur a penalty for the delay in the system. Furthermore, because of missed detection error, the Secondary User may transmit the HPP while the selected channel is occupying by Primary Users. It makes collision and the Secondary User must be punished. Figure 3 shows an example of transmission process for HPPs under affect from PUs.

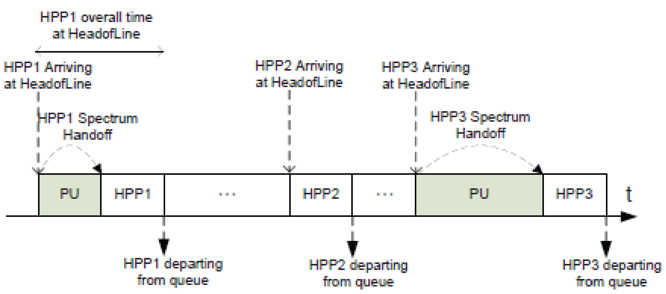

Fig.2. An example of transmission process for HPPs.

We consider a time slot system where we can define the following two states of the channels:

0: the channel is not occupied by Primary Users or the channel is idle.

1: the channel is occupied by Primary Users or the channel is busy.

Similarly to [4], we assume the channel state process is Markovian, i.e., the current state probability depends on only the previous state. We consider a Cognitive Radio System with a set of N Primary channels. We assume N channels have heterogeneous capacity $C_1, C_2, \cdots, C_N$. The traffic statistic of Primary Users are such that the occupancy of these N channels follows a discrete-time Markov process with $M = 2^N$ states. Specifically, the network state in slot t is given by $[S_1(t), \cdots, S_N(t)]$ where $S_n(t) \subset \{ 1(\text{occupied}), 0(\text{idle})\}$. The state transition probability is denoted by P^n_{s1s2} when the state transits from state s1 to state s2 of n-th channel. For example, P^n_{10} denotes the transition probability from the busy state to the idle state of n-th channel. We denote by P^n_I and P^n_B the stationary probabilities of states 0 (idle) and 1 (busy) of n-th channel. It is easy to verify that these stationary probabilities are given by $P^n_I = P^n_{10}/(P^n_{10} + P^n_{01})$ and $P^n_B = P^n_{01}/(P^n_{10} + P^n_{01})$.

3. A Spectrum-Decision-Policy Based On MDP

In this section, we develop a Spectrum-Decision-Strategy approach in Cognitive Radio networks.

We now describe how to formulate the spectrum decision problem as an MDP. An MDP model consists of the following fie elements: 1) states; 2) transition probabilities; 3) rewards; 4) actions; 5) decision epochs.

1) states: At the beginning of each slot, the Secondary User performs channels sensing process. Given that the current state of the underlying Markov process is i, the Secondary User observes state $\Theta_i \in \{0,1\}^{|N|}$ which indicates the availability of each sensed channel. S denotes the set of states consisting 2^N states.

2) transition probabilities: We assume in this section that the state transition probabilities $\{p_{i,j}\}$are known. We example state Θ_i is denoted by $[S_1^i, \ldots, S_N^i]$ and state Θ_j is denoted by $[S_1^j, \ldots, S_N^j]$. Then the transition probability $p_{i,j}$ can be defined as

$$p_{i,j} = P^1_{S_1^i S_1^j} P^2_{S_2^i S_2^j} \ldots P^N_{S_N^i S_N^j} \tag{1}$$

3) actions: Based on the observation, the Secondary User chooses action $a_t = n$ which means that the SU uses the channel n to transmit the HPP. Then, the action taken in slot t consists of the index $a_t \in \{1, \ldots, N\}$ of the channel. The objective is to select channel sequentially

in each slot so that the total expected reward accumulated over long term (wherein the channel occupancy statistics remain unchanged) is maximized. The reward gained by the Secondary User in each slot can be defined in many ways depending on the design objective.

4) rewards: For the chosen action a_t = n and current state i, together with any next state j, the expected value of the next reward is

$$R_{ij}^n = E[r_{t+1}|s_t = i, a_t = n, s_{t+1} = j]. \quad (2)$$

For this section, R_{ij}^n is completely specify by follows definitions. The reward $R_{ij}^n = R_{n,idle}$ the Secondary User get when the system state is j; n-th channel is selected and sensing results indicates that n-th channel is idle in the next state.

$$R_{n,idle} = p_I^n(1-P_f)C_n - p_B^n P_m C, \quad (3)$$

where C is the collision cost of Secondary User with Primary User due to the missed detection error. Here, the reward when n-th channel is selected is defined as the receiving expected capacity of n-th channel minus the expected collision cost of n-th channel in one time slot. It means that $R_{n,idle}$ is the expected benefit when n-th channel is opportunistically utilized by the Secondary User. On the other hand, $R_{ij}^n = R_{n,busy}$ is the reward when n-th channel is selected but sensing results indicates that n-th channel is busy in the next state. $R_{n,busy}$ can be defined as $R_{n,busy} = \frac{p_B^n C_n}{T}$, where T is the predefined constant represents the threshold number of retransmit or delay threshold of HPP due to PUs' appearance. Hence, $R_{n,busy}$ represents the delay cost due to waiting time of the Secondary User while selecting channel is busy. We ignore the waiting time in queue because we assume that HPPs are light and we already put the HPP at the head of line of the SU queue.

5) decision epochs: The decision epochs is defined as the period from the moment HPP arrives until it can be transmitted successfully.

4. Policy Iteration Algorithm (PIA)

Deterministic Markovian decision rules are function or policy $\pi: S \to A$. The core problem of MDPs is to find a policy for the decision maker: a function $\pi(s_t)$ that specifies the action π that the decision maker will choose when in current state s_t. The goal is to choose a policy π that will maximize the cumulative function of the random rewards, typically the expected discounted sum over a potentially infinite horizon:

$$E[\pi] = \sum_{t=0}^{\infty} \gamma R_{a_t}(s_t, s_{t+1}), \quad (4)$$

where we choose action $a_t = \pi(s_t)$ and γ is the discount factor that satisfied $0 < \gamma < 1$. MDPs can be solved dynamic programming through policy iteration algorithms [5]. The detail of Policy Iteration Algorithm was in [5], due to the limit of space, we do not show it here.

Numerical Results: we present some numerical results for a scenario with the follows parameters. We set number of channels N = 2 with capacity C1 = 4;C2 = 10 respectively. The probability false alarm P_f = 0:1 and the probability missed detection P_m = 0:1. The collision cost C = 40. The transition probabilities are $P_{10}^1 = 0.75, P_{10}^2 = 0.9, P_{01}^1 = 0.3, P_{01}^2 = 0.7, T = 10, \gamma = 0.9$. After performing the PIA algorithm, we get the optimal policy $\pi^* = \{1,2,1,2\}$ which indicates the action a_t corresponding to four states $(\{0,0\}; \{0,1\}; \{1,0\}; \{1,1\})$.

5. Conclusion

We have modeled a spectrum decision process for hight priority packet in Cognitive Radio Networks. Based on the assumption of Markovian transition process, the optimal spectrum access policy is obtain by using Markov Decision Process with Policy Iteration Algorithm. In the future work, when the channel statistics are unknown, we can use online learning algorithm as Q-learning for learning the optimal policy.

Acknowlegement

This paper was supported by NIA (National Information Society Agency). Dr. CS Hong is the corresponding author.

Reference

[1] H. Li. "Multi-agent q-learning of channel selection in multi-user cognitive radio systems: A two by two case," in Pro. IEEE CSMC, pp. 1893-1898, 2009.

[2] H. Li,"Multiagent q-learning for Aloha-like spectrum access in cognitive radio systems" EURASIP J. Wireless Commun. Netw., vol. 2010.

[3] C. Stevenson, G. Chouinard, Z. Lei, W. Hu, S. Shellhammer, and W. Caldwell, Ieee 802.22: The first cognitive radio wireless regional area network standard, Communications Magazine, IEEE, vol. 47, no. 1, pp. 130 .138, january 2009.

[4] Q. Zhao, L. Tong, A. Swami and 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, pp. 589-600, Apr. 2007.

[5] R. S. Sutton and A. G. Barto, Reinforcement learning : an introduction: The MIT Press, 1998.