

Improving data transmission rate for SU in CR Network with multi-radio by using Q-learning method

Le Anh Tuan, Choong Seong Hong
Department of Computer Engineering, Kyung Hee University
E-mail: {latuan, cshong}@khu.ac.kr

Abstract

This paper concerns about handover problem of Secondary User (SU) in Cognitive Radio Network (CRN). A new scheme with optimal SU's handover is proposed to improve the data transmission rate in CRN with a multi-radio. Specifically, our problem is formulated as reinforcement learning approach (Q-learning) to maximize expected reward (data transmission rate) in a long term. Simulation results show the effectiveness of our proposed scheme.

I. Introduction

In CRN network, the spectrum access problem of SU is important to improve QoS of CR system. It is optimized to improve data transmission rate in the Joint Relay Section and Channel Access [1] and select network based on the best performance in terms of resulting PU's collision probability, SU's throughput, and SU's energy consumption [2]. However, most previous works hasn't considered the SU's data transmission issue via interface multi-radio. Thus, we would like to propose a solution to solve the preceding issue in our work. By combination among spectrum sensing, Adaptive Modulation and Coding scheme in physical layer, and fading channel gain, SU performs the optimum of spectrum access decision to maximize the data transmission rate of SUs. Specifically, reinforcement learning approach (Q-learning) would be applied to maximize expected reward (data transmission rate) in a long term.

The remaining of this paper is organized as follows: Section II provides system model and problem of the data transmission rate with multi-radio. Next, reinforcement learning (Q-learning) method applied for spectrum access is addressed in Section III. Then Section IV performs simulation results. And finally, conclusions and future works are in section V.

II. System model

Assume that, having N secondary nodes in CR system, each node may use a spectrum pool of licensed spectrum (with M equal-bandwidth radio channels). Moreover, a Mobile SU has L radio interfaces and the SU' moving is very slow in the coverage of N nodes. In primary system, the primary users transmitters (PU_T) is transmitting the data to the primary user receivers (PU_R). In secondary system mobile secondary users (mobile SU) connect to

secondary nodes by using licensed spectrum.

In order to transmit a data unit by parallel multi-radio, each mobile SU should obtain bandwidth from multiple Nodes. After allocating the bandwidth, each mobile SU experiences different channel gains on each bandwidth in each node. For mobile SU with L interfaces connect to L nodes, from Shannon capacity formula for Gaussian channel, the achievable data rate r of mobile SU can be defined as follows[3]:

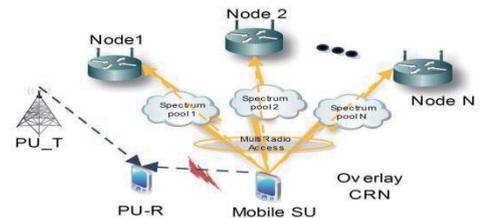


Fig. 1 System model

$$r_{su} = \sum_{l=1}^L \beta_l b_l \log \left(1 + \frac{|H_l|^2 \cdot p_l}{N_l \cdot b_l} \right) \quad (1)$$

where b_l is notation of allocated bandwidth to the mobile SU from Node l, p_l is transmission power of mobile SU to Node l, and β_l , ($0 \leq \beta_l \leq 1$) represents the efficiency which can be guaranteed by mobile user l to Node l, H_l is channel gain function and N_l is noise power spectral density

III. Q-learning applied for joint access

Decision making of SU is always effected by either current situation or next situations and thus, reinforcement learning is highly appreciated as an appropriate method for SU to make decisions precisely. This technique consists of a policy, a reward function and a value function. Suppose that at time k, mobile SU's state is s_k and emits an action a_k then at next time, a consequence of its action, the SU mobile receives a reward r_{k+1} .

and receives a new state s_{k+1} . Reinforcement learning (Q-learning) issue is address more detail in [4,5] and state s_k , action a_k , reward $r_s^a(k)$ expected reward $E(r_k)$, policy φ is describe as follow:

$$(s, a, s') \quad P(s, a, s') = \Pr[s_{k+1} = s' | s_k = s, a_k = a] \quad (2)$$

$$(s, a, s') \rightarrow R(s, a, s') = E[r_k | s_{k+1} = s', s_k = s, a_k = a] \quad (3)$$

$$s \rightarrow \varphi(s, a) = \Pr(a_k = a | s_k = s) \quad (4)$$

The value of taking action a in take s under a policy φ is denoted by $Q^\varphi(s, a)$:

$$Q^\varphi(s, a) = E_\varphi\{R_k | s = s_k, a = a_k\} \quad (5)$$

State s_t includes the state probability of fading channel and spectrum sensing.

$$s(k) = \{p_{ij}^{PU}(k), p_{ij}(k)\} \quad (6)$$

Where $p_{i'j'}(k)$ is the fading channel transition probability at next state j' given that current channel state is i' . $p_{ij}^{PU}(k)$ is the spectrum state transition probability (busy-idle) at next state i given that current channel state is j .

The action is determined as follows:

$$a_t = \{a_{Node}^l(k), a_{cs}(k), a_{ca}(k), a_{AMC}(k)\} \quad (7)$$

where $a_{Node}^l(k)$ is action for selecting node l in set of candidate L nodes at time k , $a_{ca}(k)$ is action for selecting a channel in a set of channels $\{1, 2, \dots, M\}$ of node l at time k , the channel sensing decision is denoted by $(a_{cs}(k) \in \{0(\text{no sense}), 1(\text{sense})\})$ and $a_{AMC}(k)$ is notation of the AMC schemes decision.

$$\text{Reward is defined: } r_s^a(k) = r_{s(k)}^{a(k)}(\beta_l, b_l, H_l, p_{l, N_l}) \quad (8)$$

The value of $Q(s_k, a_k)$ is learned as follows: When a mobile SU performs its communication, it calculates the immediate reward r_k equal to instant reverse the date transmission rate and updates the matrix $Q(s_k, a_k)$ corresponding to the class i of the mobile SU:

$$Q(s_k, a_k) = Q(s_k, a_k) + \alpha(r + (\gamma \cdot Q_{max}(s_{k+1}, \text{all actions}) - Q(s_k, a_k))) \quad (9)$$

We can check convergence as follows: If $\Delta Q(s, a) < \epsilon, \forall s \in S, a \in A(s)$, convergence occurs. Otherwise, back to the beginning to continue learning the value of $Q(s_k, a_k)$ until converge. Finally, mobile SU computes optimal policy to maximize the expected reward r in a long-term.

IV. Numerical Results

Assume that, system model includes three Nodes, and each node has two channels. The mobile SU has only two interfaces that can connect to two Nodes at once. Bandwidth of each Node i ($i = 1, 2, 3$) is 5, 5

and 10Mhz, respectively. Each network has the same 2 AMC (QPSK, 64-QAM), 2 respective states SNR (Bad, Good) are 1dB, 10.2488dB.

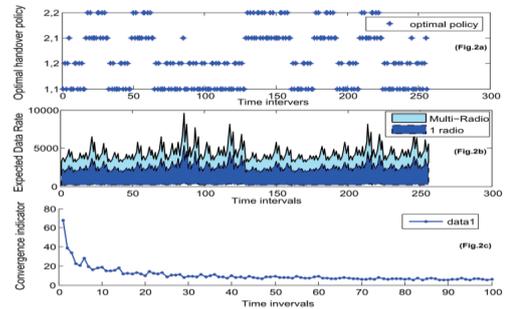


Fig.2 Simulation results

After performing Q-learning, we achieved optimal policy (Fig.2a) by using 2 channels for data transmission, expected reward (data rate) in Fig.2b) and convergence indicator (fig.2c) for learning of mobile SU. Compared to access by using a radio for data transmission, our scheme achieves higher data transmission rate (Fig.2b).

V. Conclusions

In this paper, we analyzed one scenario of data transmission in the CR System based on optimizing spectrum access of SU. Through combining among the spectrum sensing, AMC, fading channel, data rate AMC and using Q-learning method, we obtained the optimum of the data transmission rate with multi-radio interface. And the miss detection and false alarm will be considered in our coming works.

Acknowledgement

This research was supported by the IT R&D program of KCA(10913-05004: Study on Architecture of Future Internet to Support Mobile Environments and Network Diversity). *Dr. CS Hong is corresponding author.

Reference

- [1] Changqing Luo, F.Richard Yu "Centralized Scheme for Joint Relay Selection and Channel Access in Partially Sensed Cognitive Radio Cooperative Networks" IEEE Globecom 2011.
- [2] Chonggang Wang, Kazem Sohrawy, "Network Selection for Secondary Users in Cognitive Radio System", IEEE INFOCOM 2011.
- [3] Yonghoon Choi, Hoon Kim, Sang-wook Han, senior Member, IEEE, "Joint Resource Allocation for Paraleel Multi-Radio Access in Heterogeneous Wireless Networks." IEEE Transactions on Wireless Communications, 2010
- [4] L.Saker, S.Ben Jemaa, "Q-learning for joint access decision in heterogenous networks" WCNC2009
- [5] Richard S.Sutton and Andrew G.Barto "Reinforcement Learning"