A Survey on RL based Routing in Cognitive Network

Shweta Goltar[1], Mr.Trupesh Patel[2]
[1]*Computer Engineering Department, Silver Oak College of Engineering and Technology Ahmedabad*
[2]*Assistant Professor, Computer Engineering Department, Silver Oak College of Engineering and Technology Ahmedabad*

Abstract- **Designing an efficient routing protocol for cognitive radio networks is critical due to the dynamic behavior of the primary users. Based on empirical studies, the primary users activity on the licensed channels has periodicity comprised of several stages, and that the model of primary users activity may change during different stages. This paper has identified two main challenges facing designers: how to transmit packets via a stable route, and how to ensure imposing of minimal interference on the primary users. To address these, they propose a routing protocol which is based on a generalized version of Q-learning and which exploits the said model of primary users behavior. Degradation of QoS of secondary users stem from lack of attention to the multi-stage periodic behavior of primary users.**

Index Terms- Reinforcement learning, Cognitive radio network, Collaborative learning

I. INTRODUCTION

The radio spectrum has become one of the most precious and rarest natural resources, especially in recent decades due to the uncontrolled growth of wireless applications. According to FCC [1], the current use of the radio spectrum is inefficient and undesirable. Given that the licensed radio spectrum is rarely used, the dynamic allocation of radio spectrum based on cognitive radio technology assuages this problem. Following J. Mitolas initial definition [2, 3], cognitive radio is an approach with the ability of sensing the radio spectrum and of learning how to use the best available licensed channel.

A cognitive radio network consists of users equipped with a cognitive radio. They are called secondary users (SUs) because of not having been authorized to use the licensed radio spectrum. On the other hand, the users with the permission to use the licensed radio spectrum are primary users (PUs). In fact, SUs should access the licensed radio spectrum opportunistically so that the imposed interference on PUs is negligible and can be interpreted as noise. The cognitive radio network mentioned here is ad hoc which we call it cognitive radio ad hoc network (CRAHN), according to Akyildiz et al. [4].

Our approach aims to improve existing proposals because of considering the influence of multi-stage periodic behavior of PUs. The PUs activity may change over time. Based on Wellens et al. [8], there is daily periodicity in the behavior of PUs on using their own licensed channels. Moreover, they are multi-stage during a period based on Wellens et al. [8].

The PUs activity may change during a day (period); i.e. high during some intervals and low during other intervals. So, there are some stages in PUs activity during a day. To the best of our knowledge, our proposal on reinforcement learning-based routing scheme in CRAHN is the first that addresses PUs having multi-stage periodicity. This is a routing scheme based on RL (Q-learning) that learns the most stable path during each stage via storing the learning information of that stage.

Applying Q-learning based on irrelevant learning information of other stages will result in making incorrect routing decision, i.e. finding path that are affected by the PUs activity, and imposing interference on them. All previous studies didn't consider this model of PUs activity, so they could not learn the real PUs activity on the channels and didn't find the most stable path. Our proposed scheme is also based on a generalized version of Q-learning method whose components change with respect to the network condition. This will improve the convergence speed of the learning process according to Macone et al. [9].

II. RELATED WORK

Let's begin with Mahsa Soheil Shamaee, Mohammad Ebrahim Shiri and Masoud Sabaei [1]'s proposal, suppose that CRAHN consists of N secondary users {S1, S2... SN} and that it coexist with some primary networks. Each primary network purchases a licensed channel, where {C1, C2,...,CM} denote the licensed channels. Moreover, each primary network consists of a primary transmitter and a number of primary receivers. In this paper, they focus on the primary transmitters which they denote by {P1, P2,...,PM}.

The PU Pi has the license to use the channel Ci for sending data where $1 \leq i \leq M$. Let us call Ci the licensed channel of the PU Pi, and Pi the PU of the licensed channel Ci where $1 \leq I \leq M$. In CRAHN, two SUs are physical neighbors if and only if they are in the interference range of each other, and are neighbors if and only if they not only are physical neighbors but also have at least one common free licensed channel. The set of physical neighbors of Si is denoted by Ni. The underlying network is the network of SUs which are communicated with each other through the common control channel which we call CCC. Each SU is equipped with two radios, a cognitive radio and an ordinary radio to communicate control packets through CCC [21].

There is also Al-Rawi et al. [13] with their study on the application of the RL-based algorithm to routing in CRAHNs. Firstly, the traditional RL approach was applied to a routing scheme. Next, an RL feature, namely reward function, was investigated. Wang et al. [14] also proposed a distributed routing mechanism for CRAHNs over multiple licensed channels. They showed that the routing MDP can be decomposed into the layered MDPs, in which the interactions between neighbor SUs with their local next-hop selection were modeled as the local stochastic games.

As per Bing Xia, Muhammad Husni Wahab, Yang Yang , Zhong Fan, and Mahesh Sooriya bandara's study on the Reinforcement Learning Based Spectrum-aware Routing in Multi-hop Cognitive Radio Networks, two adaptive reinforcement learning based spectrum-aware routing protocols are introduced. Q-Learning and Dual Reinforcement Learning are applied respectively for them. The cognitive nodes store a table of Q values that estimate the numbers of available channels on the routes and update them while routing. So they can adaptively learn good routes which have more available channels from just local information. Compared to the previous spectrum aware routing protocols in multi-hop cognitive radio networks, they are simpler and easier to implement, more cost-effective, and can avoid drawbacks in on-demand protocols but still keep adaptive and dynamic routing. Both of our protocols perform better than the spectrum-aware shortest path protocol when network load is not too low. In the meantime, spectrum-aware DRQ-routing learns the optimal routing policy 1.5 times as fast as the spectrum-aware Q-routing at low and medium network load. It also learns a routing policy which is more than seven times as good as that of spectrum-aware Q-routing at high network load.

In another work, Kun Zheng, Husheng Li, Robert C. Qiu and Shuping Gong's article Multi-objective Reinforcement Learning based Routing in Cognitive Radio Networks: Walking in a Random Maze. In this, the routing procedure in cognitive radio networks with dynamic spectrum activities is studied. The spectrum statistics are assumed to be unknown. Moreover, the performance is measured using multiple metrics like average delay and packet loss rate. To address the challenges of randomness, uncertainty and multiple metrics, the multi-objective reinforcement learning algorithm is applied for the routing in cognitive radio networks. The effectiveness of the learning procedure is demonstrated by numerical simulations.[12]

III. VARIOUS METHODS

CRQ-ROUTING [13]

CRQ-routing is based on Q-routing [3], which is a RL-based routing model inspired by the popular Q-learning model [12], and it has been shown to improve routing performance in various types of wired and wireless networks [5-7]. Nevertheless, the application of Q-learning to hybrid-system based routing scheme is novel. The main objective of CRQ routing is to improve the co-existence of Pus and Sus, as well as to improve the Sus' network performance. Generally speaking, this is achieved by learning routes that reduce Sus' interference to Pus, and selecting a route with lower link-layer delay, which is the time required to successfully deliver a SU's packet to the SU next-hop node including any retransmission time due to PU-SU packet collisions. The main objectives are to reduce collisions with

Pus' activities along the route, and to enhance network performance. The rest of this section presents RL model for CRQ-routing, knowledge update procedure and action selection procedure. [2]

DYNAMIC SPECTRUM ACCESS

We consider a multi-hop CRN that operates in the overlay mode to access K non-overlapping PU channels. Each channel can be occupied by a single SU at one time instant when it is temporarily free. According to the empirical study of the PU operation pattern [7], we assume that the Pus employ an unslotted transmission scheme, and the dynamics of a single PU channel can be modeled by a binary-state Markov process with two operational states, namely, "idle" (s=0) and "busy" (s =1).[3]

SPECTRUM-AWARE Q-ROUTING

In our implementation of spectrum-aware Q-routing, the routing decision maker at each node x makes use of a table of values Qx(y,d), where each value is an estimate, for a neighbor y and destination d , of the sum of numbers of channels available of each link on the path to destination d , if sent via neighbor y. When the node has to make a routing decision it simply chooses the neighbor y for which Qx(y,d) is maximum. Learning takes place by updating the Q values.

MORL-BASED ROUTING DESIG

In this, we will focus on two metrics, namely transmission delay and packet loss rate. When we consider the algorithm design as an optimization problem, there could be two choices: 1) Minimize transmission delay under desired constraint of packet loss rate, which would be suitable for the scenario of a best effort application. 2) Minimize packet loss rate under desired constraint of transmission delay, which would be proper for the scenario of a real time application. In our proposed MORL-based routing, we will adopt the first one. In fact, it is also easy to implement the latter one under our proposed framework. Meanwhile, in order to simplify the MORL, we propose to use two Q-tables in our algorithm, one for the transmission delay and the other one for the packet loss rate. As mentioned above, the Q-table stores the accumulated reward regarding to the state and corresponding action. Thus the first Q-table uses the one hop transmission delay as the immediate reward while the second one uses the accumulated packet loss rate as the immediate reward. The optimal action is to find the forwarding node for the next hop in order to minimize the Q-values. Below is the procedure of our proposed algorithm. Besides, the routing loop and link layer retransmission are also considered.

IV. CONCLUSION

In this article, we studied various routing methods. Our main aim is to provide both the primary and secondary users efficient channel availability. The problem of channel availability is used using Q-learning algorithm. Both the users were able to send the packets comfortably without each other's disturbance. Here, based on the availability of destination node secondary users are also given a priority to send the packets.

V. ACKNOWLEDGEMENT

I, hereby, take an opportunity to convey my gratitude for the generous assistance and cooperation, that I received from the Principal Saurin Shah and the PG Coordinator Prof. Vikas Tulshyan and to all those who helped me directly and indirectly. I am sincerely thankful to my Guide, Asst. Prof. Trupesh Patel for their constant help, stimulating suggestions and encouragement and gave valuable guidance and constructive suggestion and helped me in completing my Literature Review work successfully. I am also thankful to other faculty members who have directly or indirectly helped me whenever it was required by me.

REFERENCES

[1] FCC, E. (2003). Docket no 03-222 notice of proposed rule making and order.

[2] Mitola, J. (2000). Cognitive radio—An integrated agent architecture for software defined radio.

[3] Haykin, S. (2005). Cognitive radio: Brain-empowered wireless communications.IEEE Journal on Selected Areas in Communications, 23(2), 201–220.

[4] Akyildiz, I. F., Lee, W. Y., & Chowdhury, K. R. (2009). Crahns: Cognitive radio ad hoc networks. AD Hoc Networks, 7(5), 810–836.

[5] Al-Rawi, H., & Yau, A. (2013). Routing in distributed cognitive radio networks: A survey.Wireless Personal Communications, 69(4), 1983–2020.

[6] Caleffi, M., & Cacciapuoti, A. S. (2016). Optimal database access for tv white space. IEEE Transactions on Communications, 64(1), 83–93.

[7] Cacciapuoti, A. S., Caleffi, M., & Paura, L. (2016). Optimal strategy design for enabling the coexistence of heterogeneous networks in tv white space. IEEE Transactions on Vehicular Technology, 65(9), 7361–7373.

[8] Wellens, M., Riihija ¨Rvi, J., & Ma ¨Ho ¨Nen, P. (2009). Empirical time and frequency domain models of spectrum use. Physical Communication, 2(1), 10–32.

[9] Macone, D., Oddi, G., & Pietrabissa, A. (2013). Mq-routing: Mobility-, gps-and energy-aware routing protocol in manets for disaster relief scenarios. Ad Hoc Networks, 11(3), 861– 878.

[10] Shiang,H.P.,&VanderSchaar,M.(2009).Distributed resource management in multi hop cognitive radio networks for delay-sensitive transmission. IEEE Transactions on Vehicular Technology, 58(2), 941–953.

[11] Xia, B., Wahab, M. H., Yang, Y., Fan, Z., & Sooriyabandara, M. (2009). Reinforcement learning based spectrum-aware routing in multi-hop cognitive radio networks. In 2009.CROWNCOM'09 4th international conference on cognitive radio oriented wireless networks and communications (pp. 1–5). IEEE.

[12] Zheng, K., Li, H., Qiu, R. C., & Gong, S. (2012). Multi-objective reinforcement learning based routing in cognitive radio networks: Walking in a random maze. In 2012 international conference on computing, networking and communications (ICNC) (pp. 359–363). IEEE.

[13] Al-Rawi, H., Yau, A., Mohamad, H.,Ramli, N., & Hashim W., et al. (2014). reinforcement Learning based routing scheme for cognitive radio ad hoc networks. In 2014 7th IFIP wireless and mobile networking conference (WMNC) (pp. 1–8). IEEE.

[14] Wang, W., Kwasinski, A., & Han, Z. (2014). Routing based on layered stochastic games for multi-hop cognitive radio networks. In 2014 IEEE wireless communications and networking conference (WCNC) (pp. 63–68). IEEE.

[15] Saleem, Y., Yau, A., Mohamad, H., Ramli, N., & Rehmani, M. H. (2015). Smart: A spectrum-aware cluster-based routing scheme for distributed cognitive radio networks. Computer Networks, 91, 196–224.

[16] Geirhofer, S., Tong, L., & Sadler, B. M. (2008). Cognitive medium access: Constraining interference based on experimental models. IEEE Journal on Selected Areas in Communications, 26(1), 95–105.

[17] Sutton, R. S., & Barto, A. G. (1998). Reinforcement learning: An introduction (Vol. 1). Cambridge: MIT Press.

[18] Xing, X., Jing, T., Huo, Y., Li, H., & Cheng, X. (2013). Channel quality prediction based on Bayesian inference in cognitive radio networks. In INFOCOM, 2013 proceedings IEEE (pp. 1465– 1473). IEEE.

[19] Varga, A. (2010). Omnet?? user manual.

[20] Bayhan, S., & Alago ¨z, F. (2012). Distributed channel selection in crahns: A non-selfish scheme for mitigating spectrum fragmentation. Ad Hoc Networks, 10(5), 774–788.

[21] Mahsa Soheil Shamaee, Mohammad Ebrahim Shiri Masoud Sabaei (2018) A Reinforcement Learning Based Routing in Cognitive Radio Networks for Primary Users with Multi-stage Periodicity Springer Science+Business Media, LLC, part of Springer Nature 2018.