The X-Layer Optimization in CRN Using Deep Q-Network for Secure High Speed Communication

Chowdhury Sajadul Islam
Computer Science and Engineering (CSE),
Uttara University, Dhaka Bangladesh
sajadul.cse@uttarauniversity.edu.bd

Md. Sarwar Hossain Mollah
Computing and Information System (CIS),
Daffodil International University, Dhaka Bangladesh
headcis@daffodilvarsity.edu.bd

***Abstract*— Cognitive Radio Networks (CRNs) offers an effective solution for radio spectrum scarcity problem and cyber-security to the growth of 5G technologies and the Internet of Things (IoT) devices. In this paper, the CRNs do cross-layer (X-layer) optimization by finding the parameters from the physical layer (Layer-1) of the OSI model and the network layer (Layer-3) of the OSI model so as to progress the end-to-end secure transmission of cognition for 5G and IoT traffic. The proposed model follows the invention target by applying a Deep Q-Network (DQN) to select after that hop for sending based on the waiting duration found on every router when keeping SINR lower than threshold determines by primary channel. A fully linked feed-forward Multilayer Perceptron (MLP) model is applied by secondary users (SUs) to estimate the activity value function. The activity value contains SINR to the primary user (PU) at the layer-1 and following hop to the routers for every packet at the layer-3. The advantage to the neural network (NN) is Mean Opinion Score (MOS) for secure encrypted high-resolution video traffic over 5G network which depends upon the packet drop rate and the bit error rate applied for transmission. As evaluated to the execution of DQN learning at the physical layer, this system gives for 37% gain in the video quality for routers with short queues and besides reaches a balanced load upon a network with routers with different service rates.**

Keywords—deep neural network; CRN; DQN; IoT; 5G

I. Introduction

The key concern of this paper is to research an *X*-layer (layer-1 and layer-3) system for resource allotment in underlay Dynamic Spectrum Access (DSA) to progress the end-to-end quality of experience (QoE) for secure 5G video traffic. Here, we build up a system to apply the cognition got from the network layer variables with physical layer variables to do resource allocation in the transmission of video traffic above the secondary network. Multi-agent DQN framework applies the artificial neural network (ANN) in SUs to do reinforcement taking. The DQN based learning framework is wont to enforce reinforcement learning in SUs to think routing waiting duration at the network layer to find next hop of packets with maintaining SINR got to the primary network under the threshold level. Our simulation outcomes illustrate that grew system exceeds the former system that applies DQN to detect the optimum SINR by taking just physical layer parameters.

II. Background

NNs have been applied towards many applications where they require finding optimum results from greater activity spaces. In [1], reinforcement learning is applied to get solution throughout big distinct action spaces. The reinforcement learning is also used in the production of semiconductor scheduling algorithm [2]. In [3], the SUs discover the most beneficial routing path by with the dynamic spectrum allotment. This constitutes an overlay DSA scheme; the SUs determine the unused spectrum sub-bands. The SUs send information on these frequencies as deciding the best routing path. In [4], suggests an *X*-layer opportunist spectrum entrance and dynamic routing algorithm for CRNs. This algorithm mutually considers routing, spectrum assignment, power allocation, security and (potentially) congestion control in a distributed way. This research proposed a joint flow control, routing, scheduling, and power control scheme based on a Layer optimization framework to increase network quality and scheduling accuracies. In [5], decentralized and localized algorithms used for joint dynamic routing, relay appointment and spectrum allotment under a spread and dynamic situation are considered for co-operative CRNs. In [6], CRNs are considered to build the analytic model to do congestion control throughout the transport layer (layer-4). This model is as well developed for an overlay advance where the cognitive process on cognitive radios (*CR*s) will be delayed while the *PU* begins utilizing the channel for its communicating purpose. All the former research discussed what is done for CRNs applying the overlay proficiency of DSA for spectrum sharing. In this paper, we research and propose an *X*-layer spectrum distribution towards the underlay DSA improvement. The novel idea of this research is to build an *X*-layer resource allotment strategy utilizing a multi-agent DQN discovering a model for many *SUs* in CRNs.

III. Routing of *X*-Layer in CRN

In this paper, we developed an X-layer routing strategy applying DQN-based model for reinforcement learning. This method, the SUs does dynamic spectrum allotment throughout some the layer-1 besides the layer-3 [7]. To analysis, layer-3 activity identifies for each SU using reinforcement learning. The expanded activity space constitutes of SINR on the transmission connection and following hop for packets based on routing delay detected at

978-1-7281-4055-1/19/$31.00 ©2019 IEEE

the intermediate routers. The neural network (NN) applies DQN-based model to execute optimization throughout the action space, $S_A = [\alpha_1, \alpha_2 ...... \alpha_n, g_1, g_2 ...... g_n]$ where $\alpha_1, \alpha_2 ... \alpha_n$ is SINR and $g_1, g_2 ...... g_n$ present adjacent hop for each *SU*.

A. Routing of X-layer in CRN

The DQN framework for *X*-layer routing has the state space as $S_s = (I_s, F_s)$ which reflects intervention got by the *SUs* under the constraints of the threshold set by *PU* and validity of allotment. The *X*-layer routing takes optimum throughout in the layer-3 towards IPTV video transmissions across the 5G network for IoT devices [8]. As such, the activity space is elaborated to take the layer-3 parameter, following hop for the packet routing. Routing delays detected by the *SUs* are applied as an input to the neural network. This routing wait time is the total delay of service time for the packet going out the queue (ψ) and the time needed to service packets already shown in the queue. Following hop is guessed by *NN* for each packet gave based on this routing delay. The time passed by a packet in the routing system (R_s) is given by Eq. 1

$$R_s(t) = (1 + qlen(t)) * ipktlen * \psi \tag{1}$$

where $qlen(t)$ is length of queue at a time t, $ipktlen$ is the length of IPTV packets and ψ is the service rate.

The following algorithm depicts the phases for *X*-layer routing. *SUs* apply the ε-greedy advance for choosing the first activity parameter, sending rate. The video transmissions throughout the network are executed applying this preferred transmit rate. For every video packet that is being transmitted throughout the network, the next activity parameter (next hop) is preferred based upon delay checked at each router into the system. Activity is updated with the following hop to the router with minimal delay.

There are 5 videos, each of 10-second length, are transmitted throughout the network. As a packet in current video transmission reaches the router on an entire queue, then the packet arrives dropped having a packet loss effect. After all the videos are sent across the system, mean PLF is computed for every video. PLF is the amount of packet dropped burst cases complete a 10 seconds length. Advantage, MOS, is computed for these identified parameters.

***Algorithm**- X-Agent DQN-Based Discovering Framework:*

for all SU_i, $i = 1, ..., N$ **do**
  Init. replay memory
  Init. of the *NN* for action-value fun. O_i with arbitrary weights φ_i
  Init. of the *NN* for mark action-value function O'_i on $\varphi'_i = \varphi_i$
end for
for Monte-Carlo simulation = 1:K **do**
  for $t < T$ **do**
  Init. the arguments of routing scheme to 0
  Choose a random activity with probability ε
  Otherwise select the action
  $A_t^{(i)} = \arg\max \ O_i\left(s_t^{(i)}; A_t^{(i)}; \varphi_i\right) A_t^{(i)}$
  Modify λ with the activity chose.
for $tt <$ *Video Duration* **do**
  if Probability of packet arrival $< \lambda * \delta$ **then**
    for all SU_i, $i = 1, ..., N$ **do**
      Comparison router waits and update following hop towards the packet
    end for
  end if
  if line up fully at router **then**
    Packet is lost. Increase the packet drop counter
  else
    Position packet in line and compute receive time
  end if
end for
Modify the state $s_{t+1}^{(i)}$ and the rewards $R_t^{(i)}$
Accumulate $e_i(t) = \left(A_i(t), s_i(t), r_i(t), s_i(t+1)\right)$ in experience replay memory of SU_i, V_i
Modify arguments (φ) of activity-value function $O\left(s_t^{(i)}, A_t^{(i)}; \varphi_i\right)$ by sampling arbitrary mini-batch of conversions from V_i
Compute mean PLF above the amount of videos sent.
Each phase update arguments of mark activity-value function $\varphi'_i = \varphi_i$
  end for
end for

IV. Setup of Monte-Carlo Simulation

Monte-Carlo simulations were applied to analyze the selection of activity value. For *PU*, threshold SINR is fixing at 1dB, transmission-power (T_p) is 10mW and powers of AWGN are 1nW. The *PU* and every *SUs* are distributed in 300m radius about the primary base station and secondary base station severally. Length within primary and secondary base stations (*BS*s) is 2km. The path loss model for carrier reaches is log-distance model with loss exponent of 2.81 transmission rate for an individual *SU* might be selected within a range of 0.1 to 0.5 Mbps.

In the IPTV communicating model, each simple stream is changed in an interleaved stream of time-stamped Packetized Elementary Stream (*PES*) packets. The transport stream (*TS*) is made by interrupting the *PES* packets into set-sized *TS* packets of 190 bytes that are referenced to independent time bases [9]. As usually, 1 IP packet consists of 7 *TS* packets [10]. Thus the amount of bits in 1 packet = 7*190*8 =10640 bits. 10 videos of 10 seconds each are carried over the routing network. The routing scheme contains 2 routers on *M/M/1/K* queues [11]. The transmitting rate is set by the *SUs* in order that set the SINR. This is potential owing to the AMC strategy applied in *SU*s. The reaching rate of packets is fixed by this transfer rate range towards the *SUs* which is 0.1 to 0.5 Mbps. This value is the outcome of action chosen by the *NN* in *SUs*.

DQN has been learning rate set to $\alpha = 0.01$ and discounting factor set to $\gamma = 0.9$ for ε-greedy approach, ε is set to 0.8 initially, after convergence, it reduces to 0. Two separate feed-forward MLP networks are used to approximate action-value and target action-value functions. Each of these neural networks has 3 fully connected hidden layers with 4, 3 and 2 neurons respectively. The input layer consists of four nodes representing the state (I_s, F_s) and chose action to be selected (SINR and following hop for the packets). The output layer is 1 neuron representing to the chosen activity SINR. The *NNs* were as well tried out on diverse configurations of the number of hidden layers and number of neurons in every hidden layer. The alternating form tested was of 2 hidden layers with 10 neurons every one. The convergence of DQN algorithm is a bit quicker with this setup of the neural network and the outcomes for MOS are coherent with those with former setup. The capability of replay memory and mini-batch is 100 and 10 severally. The total of Monte-Carlo simulation iterations is set to 500.

V. Results of Monte-Carlo Simulation

A. *Algorithm Benchmarking System*

The DQN model provides optimal bit-rate as each *SU* found the environmental situations. The mean transferring rate is illustrated in Table I. These transmission rates were chosen from a set of 0.1 to 0.5 Mbps rates by the DQN algorithm which is narrated above.

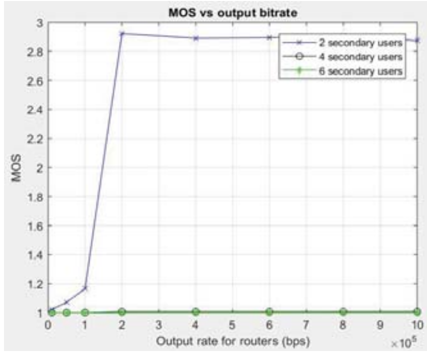
TABLE I. AVERAGE BIT RATES FOR *SUs*

Number of *SUs*	2	4	6
Average bitrate (Mbps)	0.1646	0.1737	0.2439

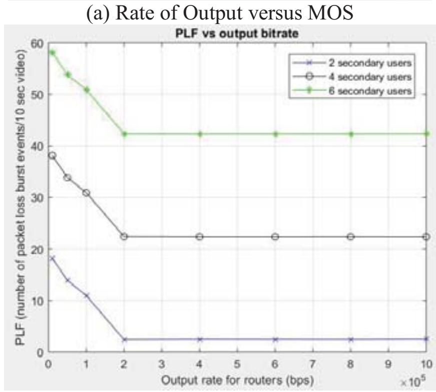
The network constitutes of 2 routers with fixed queues and similar output bit-rate. As the output for both the router is the same, the chance of every router being chosen is 0.5. When IPTV packets are transmitted throughout this network, the mean MOS of 2.5 was observed which is a low quality of video of the last user perception. This can be viewed in Fig. 1a. These outcomes were found by raising the output values towards the routers in the network. The packet reaching rate for IPTV packets transmitted throughout the network is (Average bit-rate)/(IPTV packet duration). Still, as the output rate of the router is really high (1Mbps) as contrasted to the range of packet incoming rates, it is seen that packet drop is even found for this configuration as depicted in Fig. 1b. The poor MOS results flow from the static option of the routers at layer-1 for transmission. To defeat these drawbacks, we proposed a model to apply the routing system environment for a part of the DQN learning algorithm. It is also observed from these outcomes that as the number of *SUs* raises, the packet incoming rate multiplies by the amount of users within the system. This outcome in high packet drops and the MOS losses to around 0. If a greater network with many routers is employed, then this assumption of a large number of *SUs* can be maintained.

B. *Routing Algorithm for X-Layer*

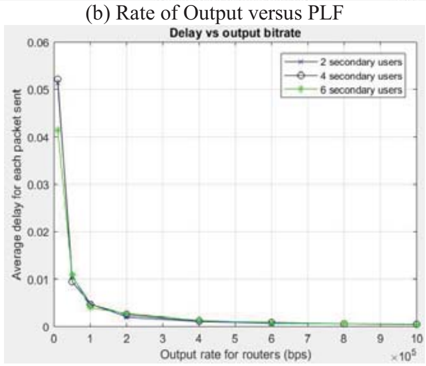
The following algorithm was applied for resource allotment in the *X*-layer routing. Fig. 2(a) depicts changes in MOS for changing the amount of *SUs* in the secondary net. The network arguments fix for these simulations are: output rate of routers = 0.31 Mbps and queue distance = 5 packets. Since it can be viewed in Fig. 2, the increase of more than 2 *SU*s outcome in a considerable decrease in the excellence of video transferred. This is due to the similar cause clarified earlier; 2 routers in the scheme cannot deal the heavy traffic load of many video sources of *SUs*. While the number of *SUs* increases, the rate of



(a) Rate of Output versus MOS

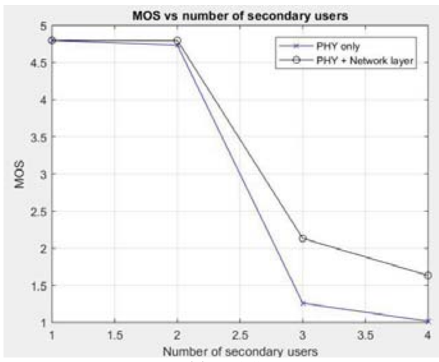


(b) Rate of Output versus PLF

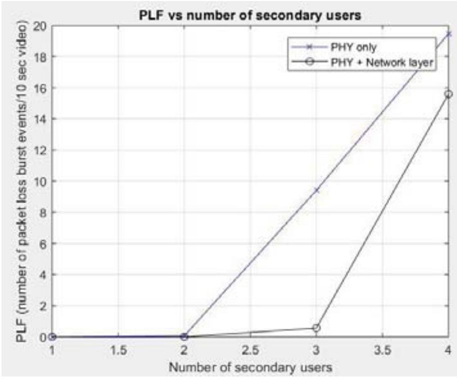


(c) Rate of Output versus average delay

Fig. 1. DQN implementation for only PHY layer



(a) MOS



(b) PLF

Fig. 2. Transferring amount of *SUs* effects.

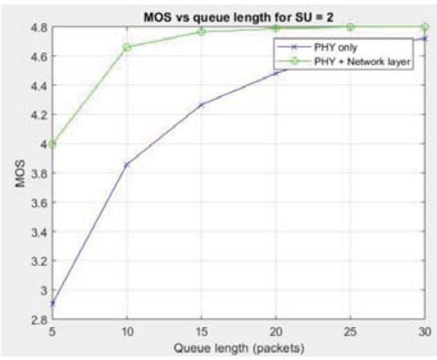
transmitting multiplies by the number of users actively transferring information in the system. This outcome in ρ>1 for a large amount of users, having an unsteady system and over-crowding. This effect enhances in queue durations having large packet drop which can be viewed in Fig. 2(b). Nevertheless, yet with a decrement in MOS owing to large influx of video traffic, the *X*-layer system accomplishes a improve quality than that of single-layer CRNs. For that reason, the *X*-layer network constantly attempts to route to the node with minimal delay. Hence, yet with queue distances rising exponentially, the *X*-layer network attempts to balance the load on both the routers in a try to enhance MOS.

C. Queue Size Increasing Rate

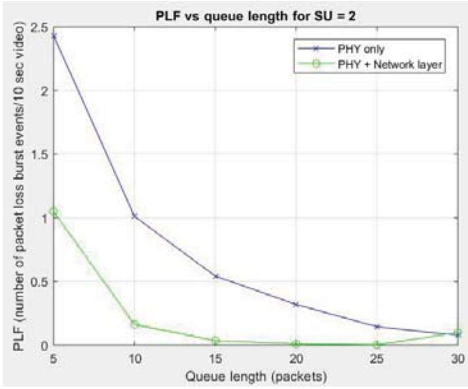
The packets transfer throughout the network was evaluated by changing the queue lengths at the routers. The range of queue length is computed based on packet incoming rates. The arrival rates of packets are calculated by:

$$A_p = \frac{T_{BR}}{L_P} \tag{2}$$

where T_{BR} is the transfer bit rate and L_P is the IPTV packet length = 10640bits. Minimal transfer rate of the *SUs* is 0.1 Mbps. The packet incoming rate for a *SU* transferring with this bit rate is 9.69packets/sec. Mean transmitting rate of *SUs* is 0.31 Mbps. The packet rate for this bit rate is 29.59packets/second.



(a) MOS



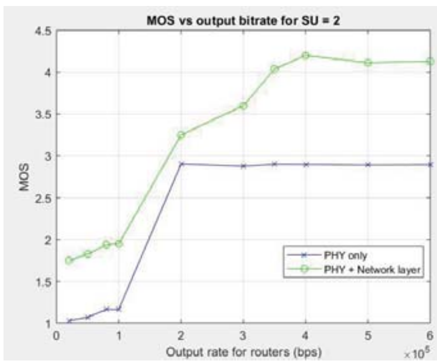
(b) PLF

Fig. 3. Queues lengths altering effects

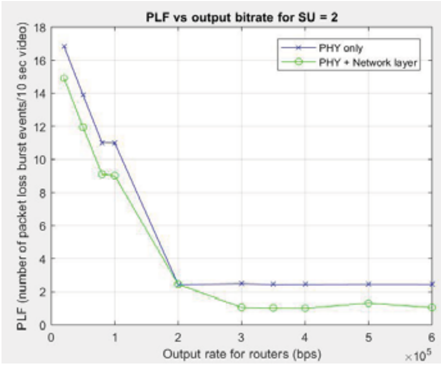
Thus the range of queue up lengths, 5 to 30 packets is wont to find the result on system parameters. Fig. 3(a) depicts 30% more MOS in the *X*-layer routing system for queue size of 5 packets. For upper values of queue size, single-layer CRNs and *X*-layer CRNs have comparable video quality. This can be found by a decrement in packet drop depicted in Fig. 3(b). PLF decreases as more memory is free in routers to store arrival packets because of larger queue lengths which successively outcomes to improve quality.

D. Output Changing Rates of Routers

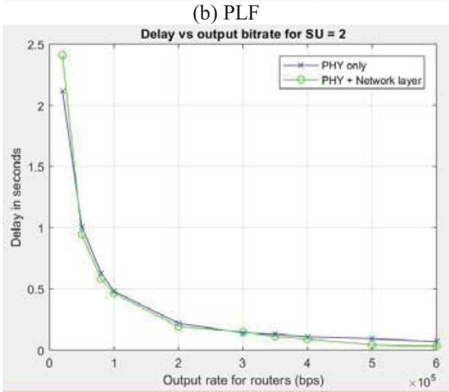
The rate of transmission which can be chosen by *SU* scopes within 0.1 to 0.6 Mbps. Thus, to analyze the result of service rates evaluates for output bit rate for the routers are chosen between 0.03 to 0.61 Mbps. The queue lengths are adjusted to minimal of 5 packets. Fig. 4(a) depicts that the mean video quality for DQN learning throughout just the physical layer satisfies to MOS = 2.91 when the output rate of the queue lengths is 0.21 Mbps. Yet while the output rate of the queue lengths is raised, these factors produce to more or less packet drop as depicted in Fig. 4(b). Nevertheless, for *X*-layer CRNs regarding the network layer (layer-3) argument queue length delay), MOS > 4. This is for the most beneficial path is chose for every packet based on the delay time monitored in the network which reduces the probability of packet drop. The effect is 37% improved in MOS. Afterward, the production rate for routers in both the system is the similar, the waiting period found are the same as shown in Fig. 4(c). The mean waiting period for packets in the network system is different in the two events evaluated if the network has distinct routes to end and more amount of routers.



(a) MOS



(b) PLF



(c) Delay

Fig. 4. Altering effects of output rate for queue lengths with minimal queues

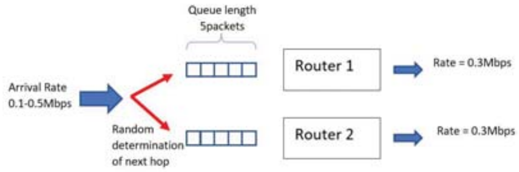
E. Load to Router

Use of a router in this system is specified by the divide of generated packets sent to the router. The use for j^{th} router in the network can be calculated by:

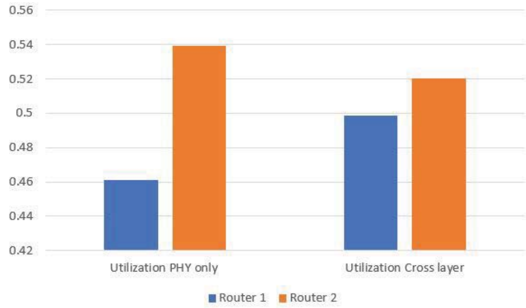
$$U_j = \frac{Ps_j}{\sum_{j=1}^{R_n} Ps_j} \tag{3}$$

where Ps_j is the amount of packets transmitted to the j^{th} router and R_n is the total number of routers. Hence, utilization 50% of router-1 shows that 50% of the created packets were forwarded to this router. For computing the usage of the queue, queue distance is fixing to minimum 5 packets.

Equal Output Rates - The packet rates at both the routers is fixed to 0.31 Mbps to compute use of a queue. Whenever both the routers have the same rate of service or output, the routing cost on both the directions to the end node is the same. Thus, the probabilistic routing advance is applied for single-layer CRNs. Therein advance; the following hop for the packet takes arbitrarily as the cost to the destination is equal through both routes with a probability of 0.5 for every router. Nevertheless, the *X*-layer routing system applies DQN to decide next hop for queue based on respected packet delays. As required, for arbitrary router choice and DQN based router choice, the router use is around 50%. This is for both the strategies apply the routers equally to transfer the data. In Fig. 5 depicts



(a) For single - layer CRNs routing algorithm

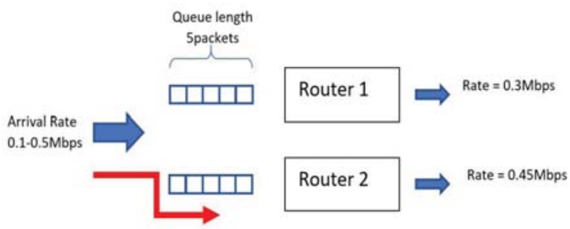


(b) Use of router

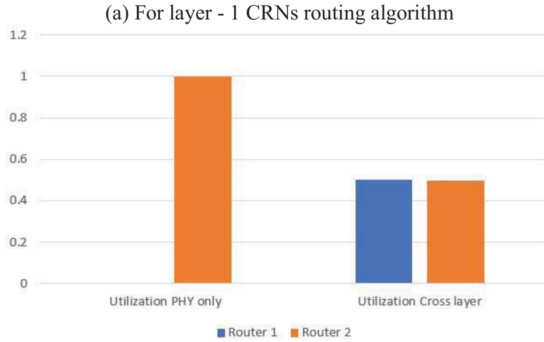
Fig. 5. Same output rates between two routers.

Mean the use of the routers for both schemes. The physical layer optimization system is applied in the different queue because of selecting the next hop for a packet randomly for a video transmission system. Still, the chance of router being picked out as the following hop is approximately 0.46 to 0.56.

Different Output Rates -The rate of output for router-1 is fixed to 0.31 Mbps as the output rate of router-2 is fixed to 1.5 times that at router-1 (0.46 Mbps). Queue lengths of both the routers are 5 packets, router preference at the physical layer is accomplished by routing via the path with maximal output to the end node. In this event, towards the layer-1 only access all the packets to acquire routed to router-2 contributing its usage of 100%. As the layer-1 only DQN optimization does not look at wait or queue lengths, the *SU* routes packets to this router even



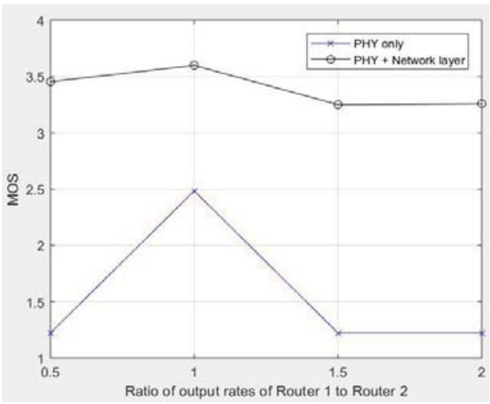
(a) For layer - 1 CRNs routing algorithm



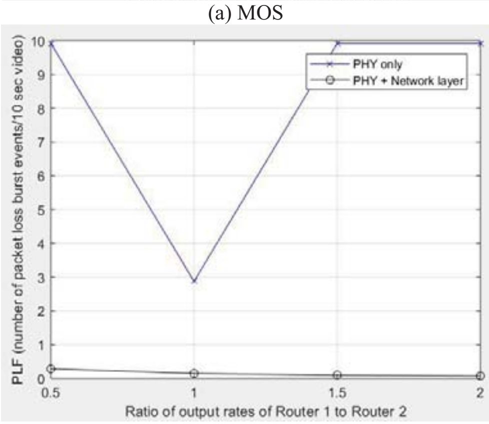
(b) Use of router

Fig. 6. Different output rates between two routers.

if they find lost at the routing destination node. The *X*-layer scheme finds late at the following hop and consequently determines the activity. When queues of the router with more output increases; waiting time at this router also raises. In this scenario, the *X*-layer system routes packets to router-1. This can be checked from Fig. 6 where the usage for the *X*-layer system is same for both the routers, hence, depicting that the packets are evenly routed to both routers to forbid congestion between them. The result of choosing special multiplication component for output bit rate for router 2 is as depicted in Fig. 7. Router-1 preserves throughput rate of 0.31Mbps. The ratio of output rates within the 2 routers is set to [0.5, 1, 1.5, and 2]. As can be depicted, the execution metric for the *X*-layer system is around constant for various values of the output rate of router-2. This ensures the adequate use of routers yet when they have distinct output rates. For perfection on layer-1 only, the system has improved MOS and minimum packet drop only if both routers



(a) MOS



(b) PLF

Fig. 7. Transferring proportion of output rates within routers

have same output values, i.e. the multiplication factor for router-2 is 1. Because of the same throughput on both routers, following hop is arbitrarily taken which outcomes are minimum packet drop than while just one router on more throughput is chosen.

VI. Conclusion

The *X*-layer (cross-layer) routing system designed in this research accomplishes resource allocation by discovering the environmental parameters over layer-1 and layer-3 and takes actions to update its own parameters at *X*-layers in order that maximizes QoE for real-time video communication. For smaller queue lengths at the intermediate routers, the *X*-layer scheme depicts to be effective by reaching MOS 1.41 times that of CRNs throughout layer-1. Hence, time-sensitive data such as video or live-streaming could be implemented without damage throughout the secondary network on an *X*-layer DQN resource allotment. As well as this model executes effective usage of the routers by executing load-balancing based on detected packet delays. For a scheme of routers on dissimilar output, this model performs better by preserving a fixed level of live video quality. For this setup, the system surpassed the CRNs for layer-1 only. The gain in MOS for *X*-layer routing is around 63%. This shows that the advance model can be reached a system of many routers with different queue lengths and service rates and the execution will become same. The *X*-layer routing scheme can be applied for any amounts of routers to be more nearer to real routing systems. These schemes may as well deal any number of *SU* transmission system more efficiently by discovering routing routes with the minimization of delays. Graph theory and game theory also has been seen to be applicable in many CRNs implementations.

References

[1] Hessel, Matteo, Joseph Modayil, Hado Van Hasselt, Tom Schaul, Georg Ostrovski, Will Dabney, Dan Horgan, Bilal Piot, Mohammad Azar, and David Silver. "Rainbow: Combining improvements in deep reinforcement learning." In Thirty-Second AAAI Conference on Artificial Intelligence. 2018.

[2] Waschneck, Bernd, et al. "Deep reinforcement learning for semiconductor production scheduling." 2018 29th Annual SEMI Advanced Semiconductor Manufacturing Conference (ASMC). IEEE, 2018.

[3] Du, Yihang, Chun Chen, Pengfei Ma, and Lei Xue. "A Cross-Layer Routing Protocol Based on Quasi-Cooperative Multi-Agent Learning for Multi-Hop Cognitive Radio Networks." Sensors 19, no. 1 (2019): 151.

[4] Fan, Shu. "A Cross-Layer Optimization QoS Scheme in Wireless Multimedia Sensor Networks." Algorithms 12, no. 4 (2019): 68.

[5] Teng, Yinglei, and Mei Song. "Cross-layer optimization and protocol analysis for cognitive ad hoc communications." IEEE Access 5 (2017): 18692-18706.

[6] V. Esmaeelzadeh, E. S. Hosseini, R. Berangi, and O. B. Akan, “Modeling of rate-based congestion control schemes in cognitive radio sensor networks,” Ad Hoc Networks, vol. 36, pp. 177–188, 2016.

[7] Li, Haitao, and Changjun Liu. "Cross□layer optimization for full□duplex cognitive radio network with cooperative spectrum sensing." International Journal of Communication Systems 32, no. 5 (2019): e3895.

[8] Khattak, Hasan Ali, Zoobia Ameer, Ikram Ud Din, and Muhammad Khurram Khan. "Cross-layer Design and Optimization Techniques in Wireless Multimedia Sensor Networks for Smart Cities." Computer Science & Information Systems 16, no. 1 (2019).

[9] Khabbiza, El Hassane, Rachid El Alami, and Hassan Qjidaa. "Peer-Assisted Content Delivery to Reduce the Bandwidth of TSTV Service in IPTV System." International Journal of Digital Multimedia Broadcasting 2019 (2019).

[10] Li, Mingfu, Chien-Lin Yeh, and Shao-Yu Lu. "Real-time QoE monitoring system for video streaming services with adaptive media playout." International Journal of Digital Multimedia Broadcasting 2018 (2018).

[11] Belikaidis, I. P., Georgakopoulos, A., Demestichas, P., Miscopein, B., Filo, M., Vahid, S., ... & Fitch, M. (2017). Multi-RAT dynamic spectrum access for 5G heterogeneous networks: The SPEED-5G approach. IEEE Wireless Communications, 24(5), 14-22.