Reconfiguring Multicast Sessions in EONs Adaptively with Deep Reinforcement Learning

(*Invited Paper*)

Xiaojian Tian, Baojia Li, and Zuqing Zhu
University of Science and Technology of China, Hefei, Anhui 230027, China, Email: zqzhu@ieee.org

***Abstract*—We proposed a deep reinforcement learning (DRL) based approach to reconfigure the multicast sessions in an elastic optical network (EON) adaptively. Simulation results demonstrate that our proposal maintains the optimality of light-trees with less reconfigurations and reduces blocking probability.**

***Index Terms*—Elastic optical networks (EONs), Multicast, Deep reinforcement learning (DRL), Reconfiguration.**

I. Introduction

Nowadays, the volume of multicast services has been increasing dramatically in the Internet, due to the rise of cloud services, video-related applications and distributed computing [1, 2]. Since 2020, the increasing trend becomes even more remarkable due to the surge in the demands for video conferencing and online classroom services during the epidemic. This has put great pressure on the infrastructure and service provisioning of the Internet, and thus promoted the research on network virtualization [3, 4] and network reconfiguration [5]. Meanwhile, it is known that optical networks transmit high-throughput traffic over long distances in the most cost-effective way. More promisingly, recent advances on the flexible-grid elastic optical networks (EONs) make the optical layer more spectrum-efficient, adaptive, and application-aware [6, 7].

Note that, realizing multicast directly in the optical layer has a few advantages, such as less bandwidth overheads and much larger multicast capacity [8], while the agility of EONs further promotes these advantages. Hence, since the inception of EONs, the problem of multicast provisioning has attracted intensive interests, and numerous algorithms were proposed [9–11]. However, due to the $\mathcal{NP}$-hardness of the problem, the proposed algorithms either are not time-efficient or cannot guarantee the performance gap to optimal solutions. Moreover, the dynamic nature of multicast services determines that each multicast session needs to be updated consistently, *i.e.*, the members in its multicast group can join or leave dynamically during its life-time [12]. This makes multicast in EONs even more complex, as the provisioning scheme of each multicast session should be readjusted adaptively to maintain optimality.

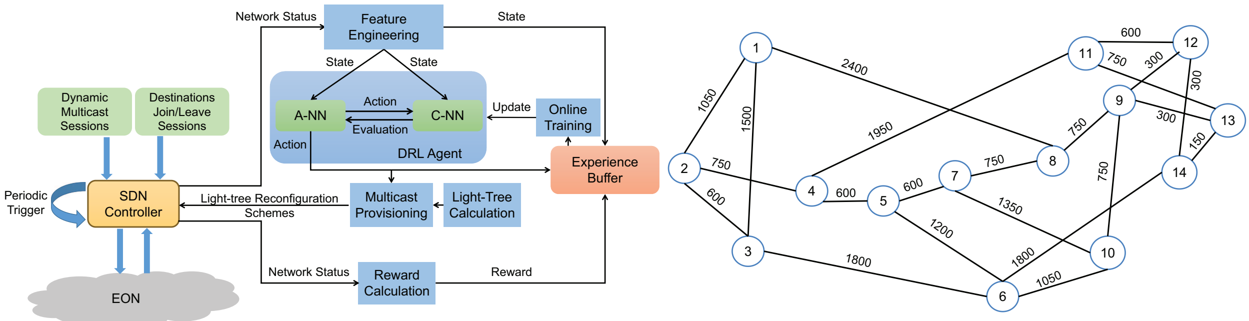
Previously, in [12], people studied how to formulate and reconfigure multicast sessions dynamically in EONs. Specifically, they divided the light-tree reconfiguration into two sub-problems (*i.e.*, tree selection and tree rearrangement), and designed algorithms to tackle them. For the tree selection, they considered two strategies, which were based on D-value and Q-value, respectively. They also developed algorithms to solve the second sub-problem with full and partial rearrangements. The performance of light-tree reconfiguration can be evaluated in two asperts, *i.e.*, the number of reconfigurations and the overall blocking probability. Specifically, to maximize the efficiency, we need to invoke the least number of reconfigurations to get the lowest blocking probability. Nevertheless, to the best of our knowledge, how to optimize this tradeoff has not been fully explored yet. This is because in a dynamic EON, it is difficult to use a deterministic algorithm to optimize the selection ratio of light-trees for the D-value-based tree selection (DTS) and Q-value-based tree selection (QTS) schemes [12].

In this work, we solve this problem by proposing a novel approach based on the deep reinforcement learning (DRL) [13], which can obtain the statistically optimal ratio of light-trees to reconfigure based on network status. Our results show that the proposed DRL-assisted approach outperforms the existing algorithms, *i.e.*, it uses less number of reconfigurations to achieve lower overall blocking probability.

II. Network Model and Algorithm Design

We mode the topology of an EON as $G(V,E)$, where V and E are the sets of nodes and fiber links, respectively. Here, considering the complicated structure and expensive cost of multicast-capable (MC) switches [11], we assume that each node $v \in V$ contains a common bandwidth-variable optical switch that is multicast-incapable (MI), similar as the assumption used in [12]. Hence, in the EON, each multicast session is established by combining several lightpaths to build a logic light-tree. Each fiber link $e \in E$ can accommodate F frequency slots (FS'), each of which has a bandwidth of 12.5 GHz, and provides a capacity of 12.5 Gbps if BPSK is used. When setting up a lightpath, we select the modulation format according to the length of its routing path, and assume that the feasible modulation formats are BPSK, QPSK, 8QAM and 16QAM. Each multicast session is defined as $MR(s,D,b,t)$, where s denotes the source, D represents the set of destinations, b is the bandwidth demand in Gbps, and t is its life-time.

When a multicast session first comes in, we use the spectrum-flexible member-only relay (SFMOR) algorithm [11] to set up the lightpaths for provisioning it. Meanwhile, during the life-time of each multicast session, D is actually time-variant, which means that destination nodes can join and leave



(a) (b)

Fig. 1. (a) System architecture and operation principle, and (b) NSFNET topology

the session on-the-fly. Therefore, to adapt to the dynamic environment, the provisioning schemes of the multicast sessions need to be reconfigured timely. We develop a novel DRL model to effectively address the aforementioned problem of multicast session reconfiguration. Specifically, the four basic elements of the DRL model are designed as follows.

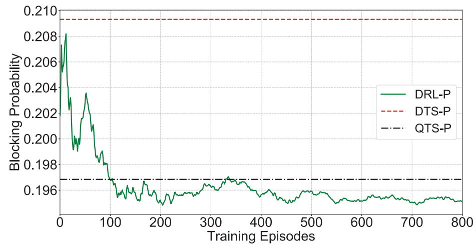
Agent: The agent uses the asynchronous advantage actor-critic (A3C) framework [14]. As shown in Fig. 1(a), A3C achieves intelligent decision-making in a dynamic environment with two neural networks (NNs). The actor NN (A-NN) chooses the best action to address the current state, while the critic NN (C-NN) evaluates the selected action to improve the performance of decision-making. The actions, states and rewards are stored in the experience buffer, with which the A-NN and C-NN are trained in the online manner.

State: We define the environment state based on the spectrum assignment and remaining life-time of each multicast session. For the i-th session $MR_i(s, D, b, t)$, we denote its current light-tree as $\mathcal{T}_i$, and meanwhile, we use the SFMOR algorithm to get a new light-tree $\mathcal{T}_i^*$ for MR_i by treating it as a new session. Then, we define three functions: 1) $slots(\mathcal{T}_i)$ returns the total number of FS' allocated to light-tree $\mathcal{T}_i$, 2) $cuts(\mathcal{T}_i)$ provides the total number of spectrum cuts [15] caused by $\mathcal{T}_i$, which is the spectrum fragmentation induced by provisioning $\mathcal{T}_i$, and 3) $time(\mathcal{T}_i)$ returns the remaining life-time of the multicast session that uses $\mathcal{T}_i$. Hence, the state of multicast session M_i is defined as

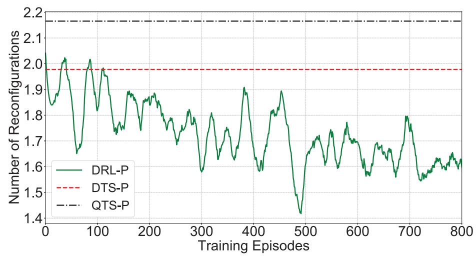
$$C_i = \frac{slots(\mathcal{T}_i^*)}{slots(\mathcal{T}_i)} \cdot \{1 + k_1 \cdot [cuts(\mathcal{T}_i^*) - cuts(\mathcal{T}_i)]\} \cdot [1 - k_2 \cdot time(\mathcal{T}_i)], \quad (1)$$

where k_1 and k_2 are the coefficients for normalization. Apparently, the smaller the first two terms on the right side of Eq. (1) are, the better the new light-tree $\mathcal{T}_i^*$ is. The third term considers the remaining life-time of MR_i, and if the value is larger, the session is closer to be expired and thus the worth of its reconfiguration will be less. Therefore, a smaller value of state C_i suggests that multicast session MR_i should be selected with a larger probability for reconfiguration. To jointly consider all the in-service sessions, we use the mean and variance of their states to denote the EON's state.

Action: As the DRL agent needs to determine the ratio of light-trees to reconfigure based on the current state, the action is just the selection ratio. Then, we get the number of light-trees to reconfigure with the ratio, sort the in-service light-trees in ascending order of their states (*i.e.*, C_i in (1)), and select the top-ranked light-trees accordingly.



(a) Overall blocking probability



(b) Number of reconfigurations

Fig. 2. Training performance of DRL (at traffic load of 45 Erlangs).

Reward: The objective of the DRL model is to find the optimal ratio of light-trees to reconfigure, such that we can invoke the least number of reconfigurations to achieve the lowest overall blocking probability. Hence, we design the reward to contain two factors. The first factor is the instant blocking probability (BP) in the period from after this reconfiguration to before the next one, while the second one is the number of lightpaths (N) to form the new light-trees after reconfiguration. Then, the instant reward is defined as $r = k_3 \cdot \log(BP) + k_4 \cdot N$, where k_3 and k_4 are the negative coefficient for normalization.

Fig. 1(a) explains the system architecture and operation procedure of our proposal. The EON is operated based on the software-defined networking (SDN) architecture [16]. The controller reconfigures the provisioning schemes of multicast sessions periodically with a fixed interval. When a reconfiguration is about to happen, the controller forwards the current network status to the feature engineering module,

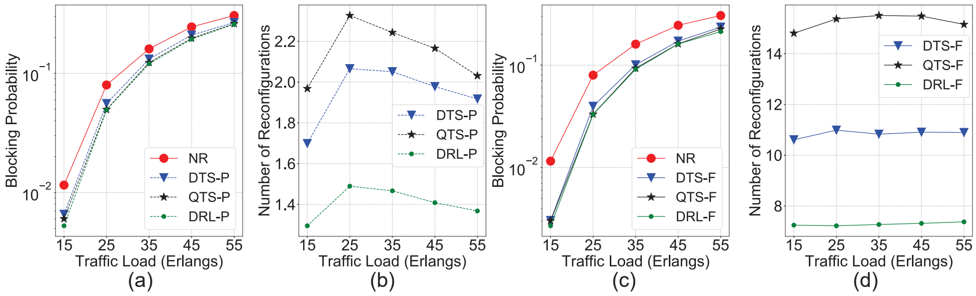


Fig. 3. Performance comparisons for partial rearrangement ((a) and (b)), and full rearrangement ((c) and (d)).

which extracts the state to send to the DRL agent. Then, the agent leverages its A-NN to choose the selection ratio, based on which we select the light-trees to reconfigure.

Next, new light-trees are obtained with the SFMOR algorithm, and the controller implements the light-tree reconfiguration schemes in the EON. Before the next reconfiguration, the controller collects the network status again, and sends it to the reward calculation module, which then can get the instant reward of the last action from the agent. We organize the state, action and reward as a sample to we insert in the experience buffer. Finally, the EON proceeds for the next interval. When enough samples have been accumulated in the experience buffer, an on-line training is triggered to update the DRL.

III. Performance Evaluations

The simulations consider an EON with the NSFNET topology in Fig. 1(b), where there are $F = 100$ FS' on each fiber link. The multicast sessions are generated dynamically according to the Poisson traffic model. For each session $MR_i(s, D, b, t)$, the source s and destinations D are randomly selected, where the initial size of D is within $[2, 5]$, the bandwidth demand b uniformly distributes within $[50, 400]$ Gbps, and the life-time t follows the exponential distribution with an average of 500 time-units. Meanwhile, during the life-time of an MR_i, destinations can join or leave it dynamically. The interval between two adjacent reconfigurations is 100 time-units. We use the algorithms developed in [12] (*i.e.*, DTS and QTS) as benchmarks, and consider both the partial and full rearrangement schemes in [12] for DTS, QTS and DRL.

Fig. 2 shows the training performance of our DRL agent. Fig. 2(a) indicates that with partial rearrangement, the blocking performance of DTS-P and QTS-P is worse than that of our DRL-assisted approach. As each light-tree is a logic one and built with several lightpaths, the "number of reconfigurations" in Fig. 2(b) is defined as the average number of lightpath reconfigurations per session. The results in Fig. 2 suggest that compared with the benchmarks, our proposal leverages less lightpath reconfigurations to achieve better blocking performance. Then, we consider more traffic loads, and plot the results on blocking probability and number of reconfigurations in Figs. 3(a) and 3(b), respectively. Similar trends can still be observed. We also compare the algorithms' performance in Figs. 3(c) and 3(d), for when they are using full rearrangement, *i.e.*, the advantages of our proposal become even more obvious.

IV. Summary

We proposed a DRL-based approach to reconfigure the provisioning schemes of multicast sessions in an EON dynamically. Our results verified that the proposal has the intelligence to maintain the optimality of light-trees adaptively.

References

[1] Z. Pan *et al.*, "Advanced optical-label routing system supporting multicast, optical TTL, and multimedia applications," *J. Lightw. Technol.*, vol. 23, pp. 3270–3281, Oct. 2005.

[2] P. Lu *et al.*, "Highly efficient data migration and backup for Big Data applications in elastic optical inter-data-center networks," *IEEE Netw.*, vol. 29, pp. 36–42, Sept./Oct. 2015.

[3] L. Gong and Z. Zhu, "Virtual optical network embedding (VONE) over elastic optical networks," *J. Lightw. Technol.*, vol. 32, pp. 450–460, Feb. 2014.

[4] L. Gong, H. Jiang, Y. Wang, and Z. Zhu, "Novel location-constrained virtual network embedding (LC-VNE) algorithms towards integrated node and link mapping," *IEEE/ACM Trans. Netw.*, vol. 24, pp. 3648–3661, Dec. 2016.

[5] J. Liu *et al.*, "On dynamic service function chain deployment and readjustment," *IEEE Trans. Netw. Serv. Manag.*, vol. 14, pp. 543–553, Sept. 2017.

[6] Z. Zhu, W. Lu, L. Zhang, and N. Ansari, "Dynamic service provisioning in elastic optical networks with hybrid single-/multi-path routing," *J. Lightw. Technol.*, vol. 31, pp. 15–22, Jan. 2013.

[7] Y. Wang, P. Lu, W. Lu, and Z. Zhu, "Cost-efficient virtual network function graph (vNFG) provisioning in multidomain elastic optical networks," *J. Lightw. Technol.*, vol. 35, pp. 2712–2723, Jul. 2017.

[8] L. Sahasrabuddhe and B. Mukherjee, "Light trees: optical multicasting for improved performance in wavelength routed networks," *IEEE Commun. Mag.*, vol. 37, pp. 67–73, Feb. 1999.

[9] Q. Wang and L. Chen, "Performance analysis of multicast traffic over spectrum elastic optical networks," in *Proc. of OFC 2012*, pp. 1–3, Mar. 2012.

[10] L. Gong *et al.*, "Efficient resource allocation for all-optical multicasting over spectrum-sliced elastic optical networks," *J. Opt. Commun. Netw.*, vol. 5, pp. 836–847, Aug. 2013.

[11] X. Liu, L. Gong, and Z. Zhu, "On the spectrum-efficient overlay multicast in elastic optical networks built with multicast-incapable switches," *IEEE Commun. Lett.*, vol. 17, pp. 1860–1863, Sept. 2013.

[12] M. Zeng *et al.*, "Control plane innovations to realize dynamic formulation of multicast sessions in inter-DC software-defined elastic optical networks," *Opt. Switch. Netw.*, vol. 23, pp. 259–269, Jan. 2017.

[13] X. Chen *et al.*, "Deep-RMSA: A deep-reinforcement-learning routing, modulation and spectrum assignment agent for elastic optical networks," in *Proc. of OFC 2018*, pp. 1–3, Mar. 2018.

[14] V. Mnih *et al.*, "Asynchronous methods for deep reinforcement learning," in *Proc. of ICML 2016*, pp. 1928–1937, Jun. 2016.

[15] Y. Yin *et al.*, "Spectral and spatial 2D fragmentation-aware routing and spectrum assignment algorithms in elastic optical networks," *J. Opt. Commun. Netw.*, vol. 5, pp. A100–A106, Oct. 2013.

[16] Z. Zhu *et al.*, "Demonstration of cooperative resource allocation in an OpenFlow-controlled multidomain and multinational SD-EON testbed," *J. Lightw. Technol.*, vol. 33, pp. 1508–1514, Apr. 2015.