Deep Reinforced Energy Efficient Traffic Grooming in Fog-Cloud Elastic Optical Networks

Ruijie Zhu[1,*], Shihua Li[1], Peisen Wang[1], Lulu Li[1], Aretor Samuel[1], Yongli Zhao[2]
Zhengzhou University, No.100 Kexue Road, Zhengzhou.
Beijing University of Posts and Telecommunications, No. 8 Xitucheng Road, Beijing.
**zhuruijie@zzu.edu.cn*

Abstract: We propose a novel energy efficient traffic grooming algorithm based on deep reinforcement learning in fog-cloud elastic optical networks. Simulation results show that it can achieve much lower energy consumption than the state-of-art algorithm.
OCIS codes: (060.2330) Fiber optics communications; (060.4250) Networks.

1. Introduction

Cloud computing is good at providing various services [1], however, with the explosive growth of IoT devices worldwide, the huge amounts of data produced by these devices are not suitable to be processed through cloud computing, as the latency is too high. To meet the demand of IoT solutions and to curve the weakness of traditional cloud computing, fog computing emerges as a great candidate [2], whose advantages are low latency, high security, improved user experience, and power-efficiency. Elastic optical networks (EONs) is a promising network infrastructure for communication between fog nodes and cloud datacenters, which can provide flexible and high-efficient service. To take full advantage of the flexibility and achieve energy-efficient EONs, traffic grooming is a promising technique, which can aggregate multiple fine-grained IP traffic flows into the optical layer by existing light-paths flexibly. Especially with the development of substrate devices, such as sliceable optical transponder, sliceable optical regenerator, traffic grooming can achieve higher energy efficiency.

Previous works have intensively studied the traffic grooming problem, and they focused on how to save energy [3], how to save the expense cost [4], and how to utilize the bandwidth efficiently [5]. However, existing studies only apply fixed traffic grooming policies regardless of the time-varying EONs states or rely on simple empirical polices based on manually extracted features, i.e., lack of comprehensive perceptions of holistic EONs states, and therefore they are unable to achieve real adaptive traffic grooming policy in EONs. In the meantime, deep reinforcement learning (DRL) has been demonstrated its effectiveness in solving large scale tasks [6]. By adopting agents, deep reinforcement learning can interact with the environment, and it can score its action to optimize its status. The target of DRL is to maximize the accumulative reward. Researchers have attempted to apply DRL in EONs in several aspects [7], such as on resource assignment, on the routing issue, and on the QoE maintenance. Nevertheless, the application of DRL for solving the traffic grooming problem in EONs has not been explored.

In this paper, we first design an energy model of traffic grooming, based on which a deep reinforced energy efficient traffic grooming algorithm (DREG) is proposed. Simulation results show that it can achieve much lower energy consumption than the state-of-art algorithm.

2. Energy Model of Traffic Grooming

Traffic grooming grooms small bandwidth requests onto the same wavelength to save resource and energy. The total energy consumption mainly consists three parts, IP porter, transponder, and regenerator. We first model the energy consumption of these three parts separately.

IP Porter: We consider 400Gbps, whose basic energy consumption is 560W, and total consumption is $E_{IPT}(W)$.

Optical Transponder: Its energy consumption depends on the required line rate, for each line rate unit, the power

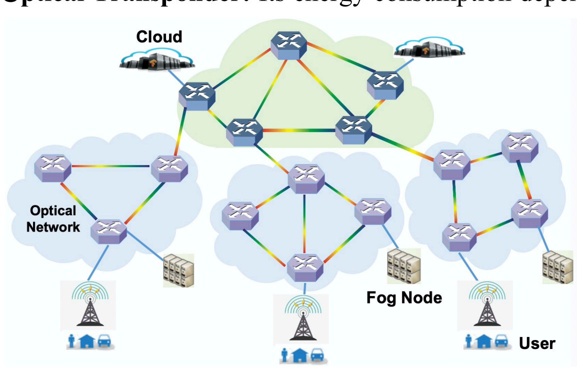

Fig.1. The structure of Fog-Cloud elastic optical networks

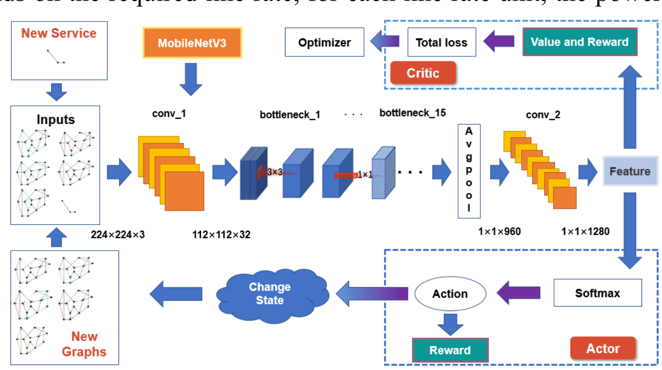

Fig.2. DREG algorithm

consumption is $1.683W$ ($\eta = 1.683W/Gbps$). The calculation equation is as follows,

$$E_{OPT}^{i}(W) = \eta \times TR(Gb/s) + 91.333(W), \ (1) \qquad E_{OPT}(W) = \sum_{i=1}^{N_{OPT}} \mathrm{E}_{OPT}^{i}, \ (2)$$

where TR is the traffic rate, and N_{OPT} is the number of optical transponders. In this paper, we mainly consider two types of line rate 40Gbps and 100Gbps.

Optical Regenerator: The basic energy consumption of each optical regenerator is $\mu = 100W$. Its extra energy consumption depends on the required line rate. For 40Gbps and 100Gbps line rates, the corresponding extra energy consumption is 25W, and 50W, separately. The calculation equation of its energy consumption is as follows,

$$E_{OPR}^{i}(W) = \mu \times N_{OPR} + \theta, \ (3) \qquad E_{OPR}(W) = \sum_{i=1}^{N_{OPR}} \mathrm{E}_{OPR}^{i}, \ (4)$$

where θ is the extra energy consumption, and N_{OPR} is the number of optical regenerators.

Thus, the total energy consumption can be calculated as $E_{TG}(W) = E_{IPT}(W) + E_{OPT}(W) + E_{OPR}(W)$.

3. Deep Reinforced Energy-efficient Traffic Grooming (DREG) Algorithm

To cope with the network state efficiently, we first convert the network states into images, which can represent different kinds of network states, and the feature can be easily extracted by deep learning algorithms. Then we design a deep reinforcement learning architecture to learn the traffic grooming policies.

3.1 Graph Conversion

The structural information contained in the network is discrete, thus it is not easy to extract features by existing deep learning algorithms directly. But the pixels of a single object in the picture are continuous, and the features of the objects in the picture can be conveniently extracted. More importantly, the network topology consists network node devices and communication media, which can describe the system more intuitively. Therefore, we propose a form of progr the information of the network and service requests into a graph as an input to the deep learning network.

The spectrum resources are divided into five parts, and each part contains 10 spectrum slots. Then five colorful images are generated based on the resource states of the corresponding parts. For each image, the locations of all nodes are first drawn according to the node location of the topology. The node connectivities of each image are based on the resource states of the connected links. We use 11 different colors to represent the resource usage of the link. When a request arrives, the route of this request will be first calculated by Dijkstra Shortest Path algorithm. The source node, the destination node, and the path will be plotted into an image. In this way, the route of a request is converted into an image and it can be processed with the spectrum images at the same time.

3.2 DREG Algorithm

In this paper, we proposed a Deep Reinforced Energy-efficient Traffic Grooming (DREG) Algorithm. We use actor-critic (AC) reinforcement learning algorithm, in which the actor selects the part that can be used for traffic grooming to reduce energy consumption, and the critic is used to judge the status of the network and analyze the action's quality.

At the beginning of an iteration, the state of the network is initialized. At this time, no resource allocation has been taken and all resources are available. When a request arrives, observing the current state of the network and conducting grooming to get a reward. If the service is successfully groomed, the network state changes. When five requests are assigned, the actor and the critic networks are updated with state, action and reward values. Since the AC shares a neural network, the total loss is calculated according to Equations (5), (6) and (7). l_v is the mean square error of the total reward value and value function, and l_a is the cross entropy of the difference between the strategy function, and the total reward value and the value function. Finally, the network parameters are updated by the gradient descent method. The entropy which is represented by e is introduced to evaluate the possibility of an action. When e converges to a certain value, the system learns a better strategy and can groom all services energy-efficiently

Condition	Action		Reward
NO Available Resource	No Action		-3
NO Available Resource	Other Action		-4
Available Resource	No Action		-5
Available Resource	Action is not available, using First-Fit		-7
Available Resource	Selected Available Parts	No New Regenerator and No New Transceiver and Grooming	3.5
Available Resource		No New Regenerator and No New transceiver and Not Grooming	3
Available Resource		New Regenerator and No New Transceiver and Grooming	2.5
Available Resource		No New Regenerator and New Transceiver and Grooming	2
Available Resource		New Regenerator and No New Transceiver and Not Grooming	1.5
Available Resource		No New Regenerator and New Transceiver and Not Grooming	1
Available Resource		New Regenerator and New Transceiver and Grooming	0.5
Available Resource		New Regenerator and New Transceiver and Not Grooming	0

Table I. Rewards and punishments table

$$l_v = \frac{1}{n}\sum_{i=1}^{n}(R_i - V(s,\theta))^2 \qquad (5)$$

$$l_a = -\frac{1}{n}\sum_{n=1}^{n}\big(R_i - V(s,\theta)\big) \cdot \log\left(\pi(a|s),\theta\right) \qquad (6)$$

$$l_t = l_v \cdot c_v + l_a \cdot c_i + e \cdot c_e \qquad (7)$$

$$H - Swish[x] = x\,\frac{ReLU6(x+3)}{6} \qquad (8)$$

The actor and the critic share a Convolutional Neural Network-MobilenetV3. It is used to extract the features of the input grayscale graphs as shown in Figure 2. MobilenetV3 is a lightweight convolutional neural network that decomposes the standard convolutional layer into deep convolution and point convolution. It can improve the speed of the operation greatly, and use the H-Swish function as shown in Equation (8) as the activation. The function can improve the accuracy of the network compared to the ReLU function. It should be noted that there is no activation function in last layer, because the extracted features will be destroyed after dimensionality reduction.

When a service is groomed, a reward value or a penalty value is calculated to judge the quality of this grooming. When all the business distribution is completed, the overall energy consumption of the network is less. Therefore, in thousands of iterations, the way of processing the services is different, and the impact on the network is different. The impact is reflected by the reward and punishment value, as shown in Table I. When a request arrives, if the network resources are not available, it should be blocked. If there are available network resources, the service will be groomed, and different rewards and punishments will be set for different situations of configurations of transceivers and regenerators. It is worth noting that when the request selects the part with unavailable resources, we use First-Fit for distribution, which ensures that the request can be successfully allocated while reducing energy consumption.

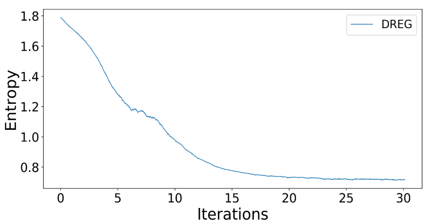

Fig.3. Entropy of DREG

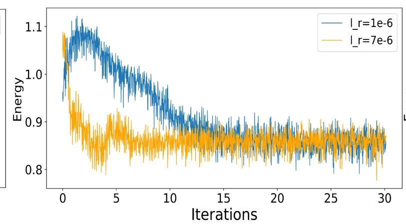

Fig.4. Energy consumption with different learning rates

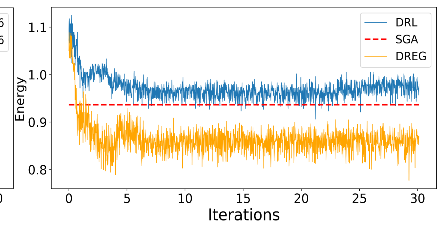

Fig.5. Energy consumption

4. Simulation Results

In this section, we evaluate the proposed algorithm with the NSFNET (14 nodes, 21 links) topology network. Supposing the network can be divided into 5 parts, each part has 10 spectrum slots on each link. There are 100 requests, and the traffic requirement of each request is equally distributed between 40Gbps and 100Gbps.

To illustrate the performance of DREG, we compare it with Deep Reinforcement Learning Algorithm (DRL) and Stateful-aware Grooming Algorithm (SGA) [1]. To represent the results more clearly, The number of iterations in each result graph is reduced to 1/1000 of the true value, and the energy consumption value is reduced to 1/100000 of the true value. Figure 3 shows the performance of entropy value with the increasing iterations, when the basic learning rate is $1e-6$. If the entropy is larger, the choice of action will be more random. We find that the entropy converges after 20000 iterations. This shows that as the network continues to learn, a stable grooming method has been discovered. To speed up the convergence, we increase the basic learning rate to be $7e-6$. Figure 4 depicts the changes of the energy consumption with two different basic learning rates (l_r). The blue line and the red line represent the basic learning rate which is $1e-6$ and $7e-6$, respectively. It can be seen that when the basic learning rate is smaller, the convergence is slower and more stable. When the basic learning rate is larger, the convergence is faster and there is some fluctuation. But in both cases, the energy consumption finally converges to the same level. This is because both learning rates are small enough that $7e-6$ can only speed up the convergence without affecting the results, therefore we take $7e-6$ as the basic learning rate.

Figure 5 shows the total energy consumption using different algorithms. Blue line indicates DRL, red line indicates SGA, and orange line indicates DREG. It can be seen that when the basic learning rate is the same, the convergence speed is almost the same. The energy consumption of DREG algorithm decreases about 13% than DRL, and the total energy consumption of DREG algorithm decreases about 8% than that of SGA. The reason is that DRL only adopts reinforcement learning to allocate the requests. SGA can only groom services with the same source and destination nodes. DREG can groom two services on the same link according to certain rules. Through continuous learning, the requests can be allocated to the suitable part making the network consumes less energy.

5. References

[1] J. Doyle, et al., “Dithen: A Computation-as-a-Service Cloud Platform for Large-Scale Multimedia Processing,” TOCC, 7.2, 2019.
[2]Z. Ning, et al., “Green and Sustainable Cloud of Things: Enabling Collaborative Edge Computing,” CM, 57.1, 2019.
[3] Z. Zhong, et al., “Energy efficiency and blocking reduction for tidal traffic via stateful grooming in optical networks,” JOCN, 8.3, 2016.
[4] B. Gangopadhyay, et al., “5G-Ready Multi-failure Resilient and Cost-effective Transport Networks,” JLT, 1-10, 2019.
[5] A. Celik, et al., “Design and Provision of Traffic Grooming for Optical Wireless Data Center Networks,” TOC, 67. 3, 2245-2259, 2019.
[6] F. Wang, et al., “Intelligent Edge-Assisted Crowdcast with Deep Reinforcement Learning for Personalized QoE,” IEEE INFOCOM, 2019.
[7] N. Cong, et al., “Applications of Deep Reinforcement Learning in Communications and Networking: A Survey,” Com. Sur. & Tutor., 2019.