Deep-reinforce-learning-assisted network orchestration for VNF-SC provisioning in inter-DC elastic optical networks

Qi Chen[1], Min Zhu[1,2], Tianyu Shen[3], Jiahua Gu[1,2], Chunping Yan[4], Pingping Gu[4]
[1] National Mobile Communications Research Laboratory, Southeast University, Nanjing, 210096, China, [2] Purple Mountain Laboratories, Nanjing 211111, China, [3] School of Electronic Science and Engineering, Southeast University, Nanjing, 210096, China
*[4] TAICANG T&W Electronics Co. Ltd., Taicang, 215400, Jiangsu. *minzhu@seu.edu.cn*

Abstract: We propose a deep-reinforce-learning (DRL)-based virtual network function (VNF) provisioning algorithm which guarantees efficient VNF reusing while consuming lower spectrum resource. The simulation results show the proposed algorithm can achieve better performance than heuristics.

1. Introduction

With the explosive growth of new networking applications (e.g., cloud computing), network function virtualization (NFV) becomes more and more attractive owing to its agile resource allocation and fast deployment of new service. Decoupling network functions from specific network infrastructure, VNFs could be realized by generic network resources (e.g., bandwidth, CPU cycles and memory space), which greatly reduces CAPEX and OPEX for network operators. Aiming to attain requirement of these network functions, each VNF instance will be deployed in high-performance servers in DCs. Then, operators need to select the optimal VNF location and route data traffic through a series of VNFs on DCs (i.e., service chaining (SC)) to achieve maximum utilization of resources. Especially, SC is always required to map VNFs to DCs sequentially. With high transmission capacity and spectral efficiency, elastic optical networking (EON) has emerged as one of most promising networking technologies for the next-generation backbone networks. Through virtualized optical transponder, EON can achieve bandwidth-variable super channels by grooming series of finer-granularity (e.g., 6.25 GHz) subcarriers and adapting the modulation formats.

On the one hand, optical SC has many advantages like high bandwidth capacity and low power consumption, which benefits inter-DC networks exceedingly. On the other hand, issues that mapping VNF-based SCs (VNF-SCs) to inter-DC EONs make corresponding service provisioning designs more complicated, which has attracted massive researches in recent years. The authors in [1] investigated how to orchestrate multicast-oriented NFV trees in both offline and online scenario. The literature [2] considered load balancing of IT and spectrum resources and proposed joint balancing algorithm, while ignored the fact that VNF resources can be reusable. The NFV placement problem has been solved in [3], but in fact it was not optimized jointly with the spectrum allocation on fibers. The authors in [4] firstly formulated an integer linear programming model to solve the problem exactly, and then a longest common subsequence (LCS)-based algorithm (LBA) was proposed to jointly optimize spectrum and IT resources. Although the heuristic method is efficient, in practice case, we often find that LCS is not unique and even using the same LCS, provision schemes show the variety. Therefore, how to choose a best LCS to achieve optimum allocation both in spectrum and IT resources becomes an interesting task.

Recently, reinforce learning (RL) has demonstrated fabulous level in large-scale control tasks and successfully been applied in the field of resource allocation problems such as multi-resource management in computer systems [5]. And there are also some similar attempts in network orchestration. The literature [6] proposed a DRL-based scheme to manage heterogeneous VNF nodes and IoT network devices, but it doesn't involve optical layer. A DRL-based framework is proposed for routing, modulation and spectrum assignment (RMSA) in EON, but it only considered spectrum resources [7]. Considering the difficulty of choosing the best LCS and provision scheme, this paper proposes DRL-based algorithm to achieve further optimization both in spectrum and IT resources.

2. Problem Description

We consider the inter-DC EON as a directed graph $G(N\&V, E)$, where N and V represent the set of DC nodes and optical nodes respectively and E denote the set of fiber links. We assume that N can route light path and be deployed with VNFs while V are only able to route. We denote the i-th VNF-SC request as R_i (s_i, d_i, T_i, B_i), where s_i and d_i mean source and destination node, respectively. $T_i = (t_{i,1}, t_{i,2}, \ldots, t_{i,J_i})$ represents i-th VNF-SC, in which J_i VNFs can be depoyed. Similarly, $B_i = (b_{i,0}, b_{i,1}, b_{i,2}, \ldots, b_{i,J_i})$ indicates bandwidth (BW) requirements in terms of frequency slot (FS), where $b_{i,0}$ is initial BW requirement and $b_{i,j}$ is BW requirement after steering through VNF $t_{i,j}$. In addition, we set boolean variable $z_{e,f}$ and $h_{v,t}$, in which $z_{e,f}=1$ if the f-th FS of fiber $e \in E$ is occupied and $h_{v,t}=1$ if DC $v \in V$ has deployed VNF $t \in T$. Note that, because we only study scheduling tasks in static network, all demands are known in advance. As for optimization objectives, there are two factors we need to consider. On the one hand, the prerequisite is that our VNF instances are a kind of reusable resources, hence we should attempt to reuse them as much as possible rather than instantiating a new one. On the other hand, if we reuse VNF excessively, VNF-SC requests may have to route a longer lightpath to a remote node deployed with needed VNF types, leading to numerous wastage of spectrum resource. Generally speaking, reuse means the detour risk, so it is a trade-off issue obviously. Therefore, the objective is to minimize the spectrum utilization ratio and deployed VNFs in the network jointly, as

$$\text{Minimize}\left(\alpha\cdot\frac{\sum_e\sum_f z_{e,f}}{|E||F|}+\beta\cdot\frac{\sum_v\sum_t h_{v,t}}{|T||V|}\right) \qquad \text{* MERGEFORMAT (1)}$$

where $|E|$, $|F|$, $|T|$, $|V|$ represent number of fibers in network, number of FS in one fiber, number of all possible VNF types and number of DCs, respectively. The α and β are factors introduced to adjust the importance of these two terms. The first term reflects the spectrum utilization ratio and the second term denotes the normalized number of deployed VNFs.

3. DRL-based Algorithm for VNF-SC Provisioning

We use Deep Q-learning (DQL) to implement our VNF provisioning algorithm, which is presented as a deep neural network (DNN). In the DQL, experience reply and target network are used to improve the convergence performance of the proposed scheme in this work. Fig. 1 shows the mutual reactions of state, action, and environment for an agent in our DRL model. The DRL agent executes a sequence of actions and observes states and rewards, with the major components of value function, policy, and model.

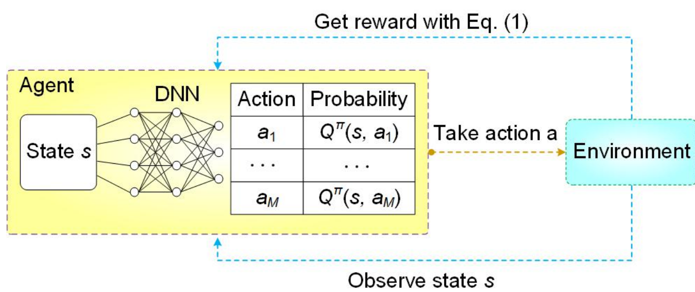

Fig. 1. A reinforcement learning model with a deep neural network

1) State: The state s_t is represented by an $1\times(M(T+1)+2J+3)$ array containing the information of existed VNF types in DCs, network topology, VNF-SC request R_i (s_i, d_i, T_i, B_i). The array is defined as

$$s_t=\{\{N_m,\ \{h_{n_m,VNF_2},\ h_{n_m,VNF_2},\ \cdots,\ h_{n_m,VNF_T}\}\}|_{m\in[1,M]}\},\ \{s,\ d,\ \{t_1,t_1,\cdots,t_J\},\ \{b_0,b_1,\cdots,b_J\}\},\ SRN\} \quad \text{* MERGEFORMAT (2)}$$

When provisioning an upcoming VNF-SC request, we find K shortest paths according to its source and destination nodes and pick one path containing LCS with VNF sequence of the request. $\{N_m, \{h_{Nm,VNF1}, h_{Nm,VNF2}, \ldots, h_{Nm,VNF\text{-}T}\}\}|_{m\in[1,M]}$ indicates M selected DCs from the path and deployed VNF types in each DC. For the sake of simplicity, we divide each VNF-SC request R (s, d, T, B) into J sub requests as $\{\{b_0, t_1\}, \{b_1, t_2\}, \ldots, \{b_{J-1}, t_J, b_J\}\}$ and handle them in turn. For accurate representation, the sub request number (SRN) of the sub request we are operating is also added into the *State*, as shown in Eq. (2).

2) Action: As mentioned above, the agent has to decide which DCs are assigned to the sub request that is being handled. There are M DCs are selected, and the DRL agent execute one action at each step, *action* ϵ $\{N_1, N_2, \ldots N_M\}$. In each step, action denotes the DCs where we map the VNF request and spectrum resource will be allocated according to mapped VNF.

3) Reward: In our VNF-SC provision scheme based on DRL in EON, Eq. (1) is used to generate reward metric. In each step, Eq. (1) is calculated twice, before and after taking an action, as O_t and O_{t+1} respectively. The reward is defined as $r_t=O_t-O_{t+1}$.

4) Training: The DNN is trained in an episodic manner. At the beginning, the agent initiates an empty experience buffer Λ and sets ε as 0.2. Each request is provisioned in several steps in turn and in each step, our agent firstly observes current state s_t as input of DNN and takes an action a_t according to output Q value. Then, agent observes state as s_{t+1}, derives reward r_t and terminals this steps with storing $\{s_t, a_t, r_t, s_{t+1}\}$ into Λ. After all requests solved, we train DNN through experiences in Λ, update its parameters and increase ε a little until 1. In particular, the capacity of Λ is limited, which means that it has to remove some old experiences when reaching a specific amount of storage.

4. Simulation Results and Discussions

We evaluate the performance of the proposed algorithm with the 28-node US Backbone topology. In the simulations, the number of nodes that have local DCs is set to be 15 and these DC nodes are selected from topology randomly. We assume that there are $|T|$=8 types of VNFs and 250 FSs in each link. The number of VNFs requested is randomly distributed between 1 and 3, the initial bandwidth requests are 6, 7 or 8 FSs and when data go through different kinds of VNFs, the bandwidth request will increase or decrease according to VNF types. Eq. (1) generally uses $\alpha \leqslant \beta$, which is because that the spectrum resources are more difficult to optimize than IT resource. The proposed DRL-based algorithm is compared with some benchmark algorithms. The first one is the shortest-path and batch VNF deployment algorithm (SRA), which selects the shortest path and maps VNFs randomly [4]. The second and third are decentralized and centralized LBA, which choose the most decentralized and centralized LCSs to reuse VNFs as many as possible, respectively. Specially, centralization means more VNFs are deployed in same DCs.

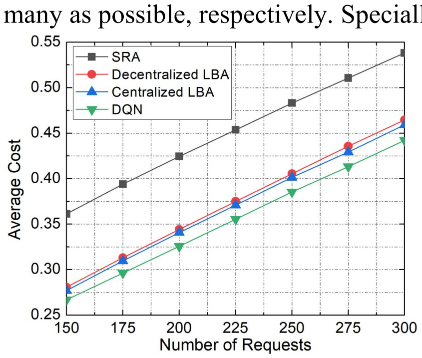

Fig.2 AC for different number of requests

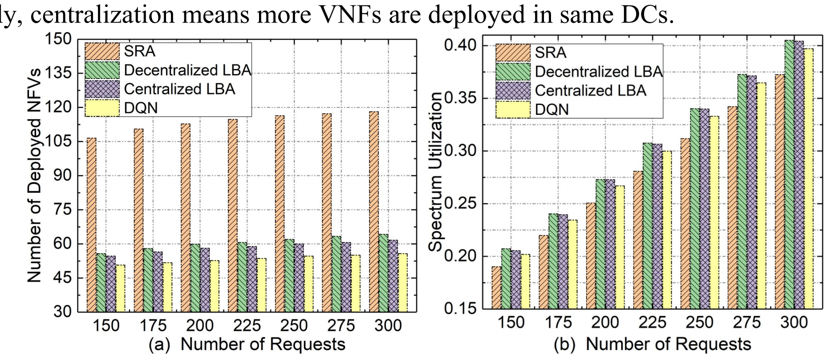

Fig.3 Number of deployed VNFs and spectrum utilization for servicing VNF-SCs.

Fig. 2 shows the results of the average cost (AC) calculated by Eq. (1). As we can see, with increasing number of requests, four algorithms all consume more resources, where SRA is much more than the other three mainly due to lack of ability to reuse already existing VNFs consciously. Decentralized LBA is better than the centralized in terms of AC, while the DQL-based algorithm can achieve best performance. Fig. 3 (a) shows utilization of IT resources. Similarly, the proposed algorithm consumes least. In addition, we find that the number of VNFs changes very slowly with the increasing number of requests. It is because that if a certain number of VNFs have provisioned in network, reusing these VNFs will occur in the allocation process of IT resources. Fig. 3 (b) shows utilization of spectrum resources. When the number of requests increases, the spectrum utilization increases linearly. For the same number of requests, our proposed DQL-based algorithms can attain better performance than LBAs. Because SRA always chooses the shortest path and doesn't consider reusing, it consumes the least spectrum resources.

5. Conclusion

This paper presents the model of reusable VNFs for VNF-SC in EON. It is necessary that the joint allocation of IT and spectrum resources is greatly significant for the VNF-SC provisioning, which can be realized through deep reinforcement learning. The simulation results show that the proposed algorithm can achieve the best performance.

Reference

[1] M. Zeng et al., "Orchestrating multicast-oriented NFV trees in inter-DC elastic optical networks", *Proc. ICC*, 2016.

[2] B. Li, X. Yu, Y. Zhao, Y. Li, Q. Ou, Z. Liu, et al., "Joint IT and spectrum resource load balancing for VNF selection in inter-datacenter elastic optical networks", *Proc. 16th Int. Conf. Opt. Commun. Netw. (ICOCN)*, pp. 1-3, Aug. 2017.

[3] M. Xia, M. Shirazipour, Y. Zhang, H. Green, and A. Takacs, "Network function placement for NFV chaining in packet/optical datacenters," J. Lightw. Technol., vol. 33, no. 8, pp. 1565–1570, Apr. 15, 2015.

[4] W. Fang, M. Zeng, X. Liu, W. Lu and Z. Zhu, "Joint spectrum and it resource allocation for efficient vnf service chaining in inter-datacenter elastic optical networks", IEEE Communications Letters, vol. 20, no. 8, pp. 1539-1542, 2016.

[5] H. Mao, et al.: "Resource Management with Deep Reinforcement Learning," In Proc of HotNets 16. ACM, New York, NY, USA, pp. 50-56, Nov. 2016.

[6] X. Fu, F. R. Yu, J. Wang, Q. Qi and J. Liao, "Dynamic service function chain embedding for NFV-enabled IoT: A deep reinforcement learning approach", IEEE Trans. Wireless Commun., vol. 19, no. 1, pp. 507-519, Jan. 2020.

[7] X. Chen et al., "DeepRMSA: A deep reinforcement learning framework for routing modulation and spectrum assignment in elastic optical networks", *J. Lightw. Technol.*, vol. 37, pp. 4155-4163, Aug. 2019.