Evol-TL: Evolutionary Transfer Learning for QoT Estimation in Multi-Domain Networks

Che-Yu Liu, Xiaoliang Chen, Roberto Proietti, S. J. Ben Yoo
University of California, Davis, Davis, CA 95616, USA
{*cyliu,xlichen,rproietti,sbyoo*}*@ucdavis.edu*

Abstract: We propose an evolutionary transfer learning approach for QoT estimation in multi-domain optical networks. The results demonstrate that our approach can reduce the amounts of required training data by 10× while achieving accuracies of > 90%. © 2020 The Author(s)

OCIS codes: (060.1155) All-optical networks; (060.2330) Fiber optics communications

1. Introduction

The rapid growth of emerging network applications and their stringent quality-of-service requirements are driving today's multi-autonomous system (multi-AS) backbone networks to evolve toward multi-domain elastic optical networks (MD-EONs), which can support high-capacity and user-customized end-to-end services across ASes [1]. To realize resource-efficient service provisioning in EONs, accurate quality-of-transmission (QoT) modeling techniques are indispensable. In this context, recent studies have reported a number of machine learning (ML)-aided cognitive QoT estimation designs that can model complex network dynamics (e.g., dynamic traffic profiles) and uncertainties (e.g., uncertain device conditions) using big data analytics [2, 3]. However, these existing works focus on single-domain scenarios and cannot be applied to MD-EONs, where only a very limited amount of domain information is available due to domain privacy concerns. Therefore, we lately proposed a hierarchical learning approach for QoT estimation in MD-EONs [4], where domain managers (DMs) work cooperatively with a broker plane by learning domain-level and inter-domain QoT estimators, respectively. Nonetheless, a major challenge remaining unmet is that the approach entails a significant amount of performance monitoring data for every inter-domain lightpath, which can be very costly (i.e., collecting data for each inter-domain lightpath requires nontrivial efforts from the broker plane and DMs) and unscalable. Fortunately, the invention of transfer learning (TL) has enabled to significantly reduce the amount of efforts required for training a ML task by reusing knowledge learned from relevant tasks [5]. The application of TL for QoT estimation was first studied in [6]. However, the work in [6] only adopts a very simple and straightforward TL scheme, and more importantly, does not address the challenges in MD-EONs.

In this paper, we propose Evol-TL, an evolutionary transfer leaning approach, for enabling scalable QoT estimation in MD-EONs. Evol-TL exploits a broker-based MD-EON architecture, where a broker plane performs end-to-end QoT estimation by collecting encoded features from DMs. A generic algorithm (GA) is designed to enable Evol-TL to determine the proper neural network architectures and the right sets of parameters for transferring through iterative optimizations. Evaluations with experimental data show that Evol-TL can significantly reduce the amount of required training data for new tasks without sacrificing the estimation accuracies.

2. Framework

Fig. 1(a) depicts the schematic of Evol-TL in an MD-EON with broker orchestration. A broker plane works with DMs to provide inter-domain services (e.g., QoT-aware lightpath provisioning), following mutual service level agreements. Each DM reports an abstraction of its domain to the broker plane for preventing disclosure of confidential domain information. Particularly, to assist QoT estimation for inter-domain lightpaths, each DM employs an encoder (e.g., a neural network) to map the performance monitoring data collected along the corresponding intra-domain path segments to a new feature space and reports the encoded features to the broker plane. The broker plane combines the received features, and hereby, builds and trains a deep neural network (DNN)-based QoT estimator for each of the inter-domain paths. Differently from the work in [4], where QoT estimators are trained independently, we apply Evol-TL to pursue a more scalable and effective realization of end-to-end QoT estimation. Specifically, as the nature of different QoT estimation tasks (for different paths) are similar, we exploit and transfer the knowledge learned by a pre-trained QoT estimation model to the training of new models so that only small amounts of additional data will be needed to fit the models. Such knowledge transfers can be achieved by initializing the new models with weights copied from the pre-trained model. Fig. 1(b) illustrates the principle of

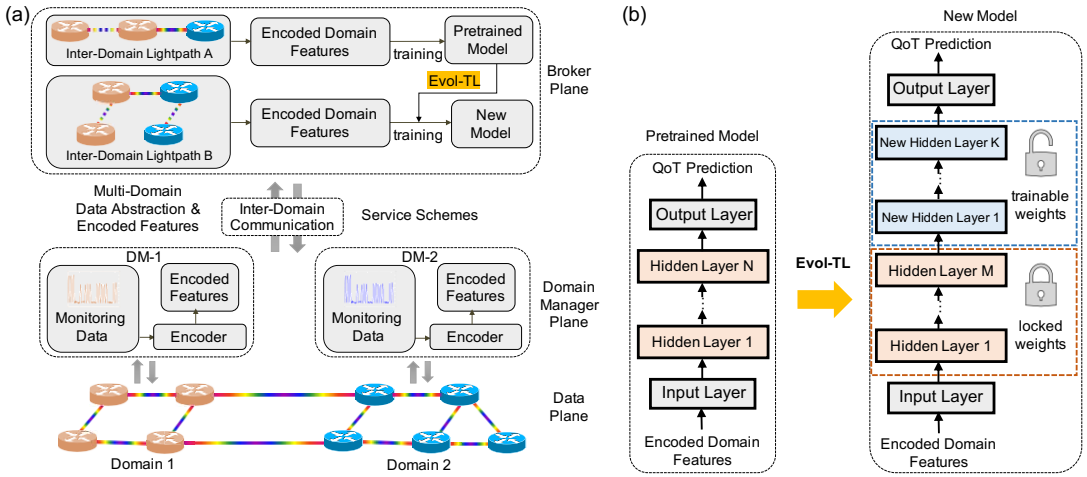


Fig. 1. (a) Schematic of Evol-TL in an MD-EON with broker orchestration, and (b) principle of knowledge transfer in Evol-TL.

knowledge transfer in Evol-TL. We first copy a few hidden layers (typically, the lower layers, as they are less over-fitted) from the pre-trained model to the new model and set the corresponding weights locked. A few randomly initialized and trainable hidden layers are then added on the top to mitigate overfitting. Eventually, we fine tune the obtained model with data for the new task.

3. GA-based Optimization

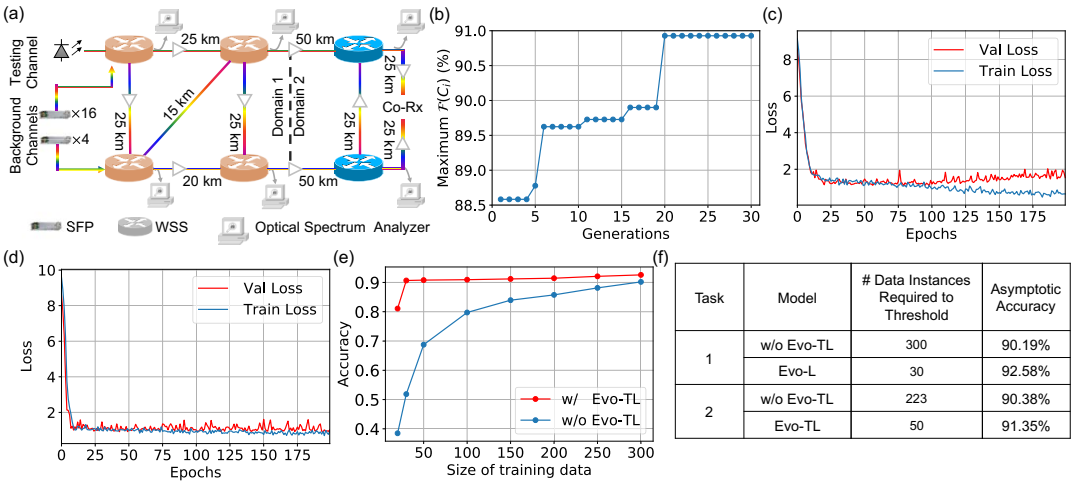
Determining proper DNN architectures to use and the most effective sets of knowledge to transfer have always been a challenging task, whereas previous works mostly rely on human experiences or brute-force searches. In this work, we propose to optimize the DNNs with GA-assist evolutionary learning. Let $P = \{C_i, i \in [1, I]\}$ denote a population of I individuals. Each individual C_i is encoded as $[L_t, L_n, \mathbf{K}, \mathscr{H}, \mathscr{G}]$ to convey the information of: 1) L_t, number of hidden layers transferred from the pre-trained model to the new model; 2) L_n, number of trainable hidden layers added; 3) $\mathbf{K}$, numbers of neurons in each trainable layer; 4) $\mathscr{H}$, activation function of the DNN, and 5) $\mathscr{G}$, optimizer used in training. We measure the performance of each C_i by defining a fitness function $\mathscr{F}(C_i)$ that calculates the average QoT estimation accuracy with the model given by C_i. Table 1 summarizes the procedures of the proposed GA-based optimization. Firstly, we randomly initialize a population P (**Step 1**) and evaluate the fitness function $\mathscr{F}(C_i)$ for each individual $C_i \in P$ by performing a training process using the model encoded by C_i (**Step 2**). In **Steps 3-4**, to facilitate a positive evolution, we sort the individuals in the descending order of $\mathscr{F}(C_i)$ and select the top 80% of the individuals as parents, which are then used to produce offsprings with crossover (exploiting advantageous knowledge) and mutation (exploring) operations. We maintain P of a fixed length by padding with new randomly generated individuals. Finally, we repeat the above procedures for M iterations.

Step 1: Randomly initialize P.
Step 2: Perform training for each model conveyed by C_i and calculate $\mathscr{F}(C_i), \forall C_i \in P$.
Step 3: Sort individuals in the descending order of $\mathscr{F}(C_i)$ and select the first 80% of individuals as parents P_0.
Step 4: Perform crossover and mutation operations with P_0 to obtain offspring P_s and restore the population by adding randomly generated individuals P_r, i.e., $P = P_s \cup P_r$.
Step 5: Repeat **Step 2-4** for M iterations.

Table 1. GA-based optimization procedures adopted in Evol-TL.

4. Results

We evaluated the performance of Evol-TL with experimental data collected from the two-domain EON testbed shown in Fig. 2(a). We set up three inter-domain lightpaths consisting of three, four, and five nodes, respectively. For each of the paths, we generated diversified network conditions by randomly changing the background traffic and attenuation for each wavelength selective switch (WSS), leading to the signal launch power varying between -7dBm and 12dBm. At each run, we measured the actual Q-factor of the testing signal at Co-Rx and record this value as the current data label. Then, we used WSSs to remove the testing signal and recorded the outputs of optical spectrum analyzers (OSAs) as the features. We collected 1095, 1440, 1795 data instances for the three-node, four-node, and five-node paths respectively. We pre-trained a QoT estimator of two fully-connected layers (12 neurons for each) for the three-node path and transferred the knowledge to the training of the QoT estimators for the four-node (task 1) and five-node paths (task 2).



Task	Model	# Data Instances Required to Threshold	Asymptotic Accuracy
1	w/o Evo-TL	300	90.19%
	Evo-L	30	92.58%
2	w/o Evo-TL	223	90.38%
	Evo-TL	50	91.35%

Fig. 2. (a) Two-domain EON testbed implementation; (b) convergence process of Evo-TL; (c-d) loss vs training epochs (c) with Evo-TL and (d) without Evo-TL (number of training data instances being 30); (e) accuracies with different numbers of data instances; (f) results of required training data instances to threshold (accuracies above 90%) and asymptotic accuracy. Task 1: training of the QoT estimator for the four-node path. Task 2: training of the QoT estimator for the five-node path.

Fig. 2(b) shows the evolutionary process of Evol-TL for task 1, where each point represents the fitness of the best individual at certain generation. For each individual, 30 instances of the four-node path data were used for training and another 100 instances were used for evaluation and test, respectively. We can see that the performance of Evol-TL keeps improving steadily with the generation. At generation 20, Evol-TL converges and obtains an individual with $\mathcal{F}(C_i) > 90\%$. We decoded the information in the individual to build a fully-connected DNN with three hidden layers (11 neurons for each), two of which copied from the pre-trained model. Meanwhile, activation function of ELU and gradient-based optimizer of adamax were used for the DNN. Figs. 2(c) and (d) show the comparison between loss (validation and training) versus training epochs without and with Evol-TL for the new model. In Fig. 2(c), we can see that with the increasing of epoch, the validation loss is obviously higher than the training loss, which indicates overfitting due to the small number of training samples. In contrast, with Evol-TL, the validation loss and the training loss are comparable, showing that the model is well fit (Fig. 2(d)). Fig. 2(e) depicts the comparison between accuracy versus different sizes of training dataset with and without Evol-TL in Task 1. Without Evol-TL, we can realize that the accuracy of the new model increases steadily with the increasing of the number of data instances and reach 90% with 300 training data instances. In contrast, with Evol-TL, the accuracy increases drastically when we increase the size of training dataset from 20 to 30 and remain stable afterward. If we set a threshold for achieving a good model as its accuracy being above 90%. The numbers of training data instances required to reach the threshold are 30 and 300 for training with and without Evol-TL, respectively. This means 10× reduction in the amount of required training data. Fig. 2(f) summarizes the results of number of required training data instances to threshold and asymptotic accuracy (accuracy achieved by using full dataset). Overall, the evaluation results demonstrate that Evol-TL enables us to decide proper DNN architectures and the right knowledge to transfer, leading 10× reduction in the amount of required training data.

5. Conclusion

In this paper, we proposed Evol-TL for scalable QoT estimation in MD-EONs. Evaluation results demonstrate 10× reduction in the amount of required training data by Evol-TL.

References

1. X. Chen *et al.*, *IEEE Commun. Mag.*, **vol**. 56, pp. 152–158, 2018.
2. L. Barletta *et al.*, in *Proc. Opt. Fiber Commun. Conf.*, 2017, paper Th1J.1.
3. R. Proietti *et al.*, *J. Opt. Commun. Netw.*, vol. 11, pp. A1–A10, 2019.
4. G. Liu *et al.*, *J. Lightw. Technol.*, vol. 37, pp. 218-225, 2019.
5. S. Pan *et al.*, in *Trans. Knowl. Data Eng.*, vol. 22, pp. 1345-1359, 2010.
6. W. Mo *et al.*, in *Proc. Opt. Fiber Commun. Conf.*, 2018, paper W4F.3.

Acknowledgements: This work was supported in part by DOE DE-SC0016700, and NSF ICE-T:RC 1836921.