# Demonstration of Distributed Collaborative Learning with End-to-End QoT Estimation in Multi-Domain Elastic Optical Networks

### XIAOLIANG CHEN,[1] BAOJIA LI,[2] ROBERTO PROIETTI,[1] CHE-YU LIU,[1] ZUQING ZHU,[2] AND S. J. BEN YOO[1, *]

[1]*Department of Electrical and Computer Engineering, University of California, Davis, Davis, CA 95616, USA*
[2]*School of Information Science and Technology, University of Science and Technology of China, Hefei, Anhui 230027, P. R. China*
[*]*sbyoo@ucdavis.edu*

**Abstract:** This paper proposes a distributed collaborative learning approach for cognitive and autonomous multi-domain elastic optical networking (EON). The proposed approach exploits a knowledge-defined networking framework which leverages a broker plane to coordinate the operations of multiple EON domains and applies machine learning (ML) to support autonomous and cognitive inter-domain service provisioning. By employing multiple distributed ML blocks learning domain-level features and working with broker plane aggregation ML blocks (through the chain rule-based training), the proposed approach enables to develop cognitive networking applications that can fully exploit the multi-domain EON states while obviating the need for the raw and confidential intra-domain data. In particular, we investigate end-to-end quality-of-transmission estimation application using the distributed learning approach and propose three estimator designs incorporating the concepts of multi-task learning (MTL) and transfer learning (TL). Evaluations with experimental data demonstrate that the proposed designs can achieve estimation accuracies very close to (with differences less than 0.5%) or even higher than (with MTL/TL) those of the baseline models assuming full domain visibility.

## 1. Introduction

Multi-domain elastic optical networking (EON) has become one of the most appealing solutions for the next-generation Internet infrastructure [1]. The flexible spectrum allocation mechanism of EON [2, 3] can potentially support high-capacity and quality-of-service-aware networking across multiple autonomous systems (ASes) [4]. However, coordinating the operations of multiple autonomous elastic optical networks (EONs) to realize effective multi-domain networking while complying with stringent domain administrative constraints is a non-trivial task. Previous researches have reported a number of flat or hierarchical network control and management (NC&M) paradigms for multi-domain EON, where inter-domain services are provisioned by interactions of peer domains in a distributed manner [1, 5] or by coordination of a higher NC&M hierarchy (often referred to as multi-domain orchestrator) in a semi-centralized manner [6, 7]. These works typically employ fixed and rule-based service provisioning designs (e.g., shortest-path routing) relying on the limited information advertised by ASes. Due to the lack of knowledge about the essential (and hidden) rules of multi-domain EONs, such designs can hardly fully exploit the benefits of EON in the context of multi-domain [8]. Meanwhile, as network conditions (e.g., topologies, traffic profiles) may evolve, the performance of these designs can downgrade over time, entailing significant efforts from networking experts for periodical policy updates, thus leading to poor network scalability.

Recently, machine learning (ML) has attracted intensive research interests and has been

successfully applied to different applications (e.g., image recognition, online gaming) owing to its capability of learning complex functions from big data [9]. Particularly, this capability enables knowledge-based optical networking [10–12] equipped with cognitive impairment modeling (by learning quality-of-transmission (QoT) models [13,14]), fault management (by learning abnormal network behaviors [15,16]), and resource provisioning (by learning routing, modulation and spectrum assignment (RMSA) policies [17,18]) functions. However, most of the existing works only considered single-domain scenarios, which assume that network states are fully observable. Whereas, in multi-domain EONs, only very limited yet abstracted intra-domain information is available, making these existing approaches difficult to be applied. In [19], based on the conjecture that the QoT of inter-domain lightpaths can be derived from the QoT estimations of a sequence of intra-domain path segments, we demonstrated a hierarchical learning approach for QoT estimation in multi-domain EONs. Nevertheless, the hierarchical approach still requires ASes to report optical performance monitoring data from domain border nodes, which may conflict with the autonomy rule of the current Internet. Moreover, because the performance of this approach is highly dependent on the correctness of the artificial task partitioning and featuring scheme used (i.e., the conjecture about the nexus between end-to-end and domain-level QoT), it can hardly be generalized for other more complicated multi-domain networking applications.

To overcome the above challenges and to realize knowledge-defined multi-domain EON effectively and practically, this paper proposes a distributed collaborative learning approach. The proposed approach employs multiple distributed learning blocks to learn domain-level features and to work with aggregation learning blocks through the chain rule-based training [9]. Consequently, the proposed approach enables ML applications that can fully exploit the multi-domain EON states while obviating the need for the raw and confidential intra-domain information. In Section 2, we first present a knowledge-defined multi-domain networking framework, which leverages a broker plane to coordinate the operations of multiple autonomous EONs and applies ML to realize cognitive service provisioning. Based on the framework, we elaborate on the proposed distributed learning approach in Section 3. In Section 4, we study end-to-end QoT estimation as a use case of the proposed approach and propose three estimator designs by incorporating the concepts of multi-task learning (MTL) and transfer learning (TL). Section 5 show the performance evaluations. Finally, we summarize the paper in Section 6.

## 2. Knowledge-defined networking framework

The proposed distributed learning approach relies on a multi-broker-based NC&M architecture depicted in Fig. 1(a) [20]. Specifically, each domain manager (DM) operates its domain (an EON) autonomously, while working with third-party broker agents for inter-domain service provisioning. DMs disclose different degrees of intra-domain information to brokers (by signing different mutual service level agreements according to their actual service requirements and privacy considerations), while brokers calculate and distribute inter-domain service schemes. Thus, brokers can be seen as constituting a virtual broker plane lying above DMs for coordinating the operations of ASes for inter-domain service deployment. The activities of DMs and brokers are incentive-driven. DMs can subscribe to multiple brokers for more diversified and higher-quality services. Meanwhile, brokers can cooperate or compete for pursuing the desired market sharing (i.e., partitions of inter-domain service requests). Such an incentive-driven mechanism motivates more efficient resource utilization across domains while complying with the autonomy rule of the current Internet by avoiding the presence of a single multi-domain administrative entity.

We introduce ML to the service provisioning designs of DMs and brokers to enable a knowledge-defined cognitive multi-domain networking framework. Figure 1(b) shows the system layout of a DM. Each DM operates its network using a software-defined networking (SDN) paradigm, where a centralized SDN controller communicates with the data plane equipment to receive service requests, collect network state data (e.g., topology, resource utilization, optical
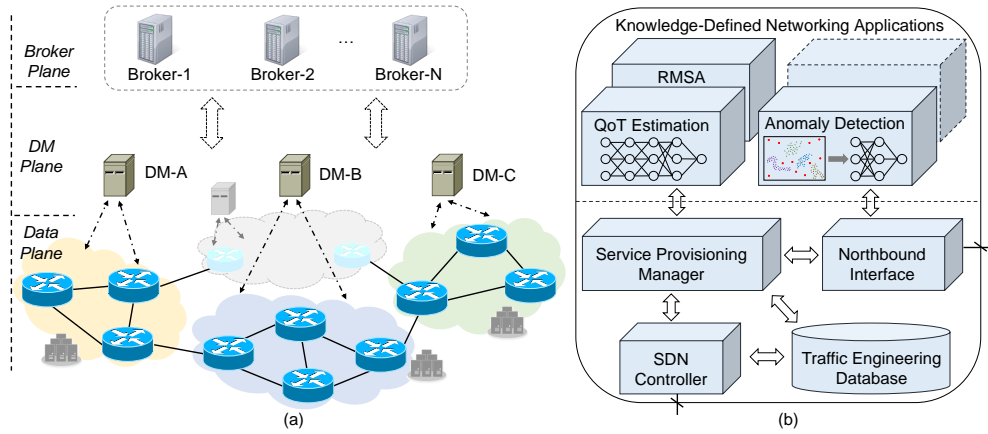
Fig. 1. Knowledge-defined multi-domain networking framework: (a) multi-broker-based multi-domain NC&M architecture, (b) domain manager (DM) system layout.

performance monitoring), and distribute configuration commands. A set of knowledge-defined networking applications embedded with ML-based data analytic capabilities are deployed as key enablers for cognitive networking. Each application handles a specific task, for instance, QoT estimation, RMSA [21], or network slicing [22, 23]. The system modular design of brokers resembles that of DMs but is simpler as brokers do not need to manage data plane operations. The knowledge-defined applications of brokers and DMs work cooperatively to provision inter-domain services, following an observe-analyze-act cycle [24]. In particular, upon receiving an inter-domain service request, a broker first inquires DMs for the related domain-level knowledge (learned by domain-level ML models from abundant and comprehensive network state data and traces) and necessary domain abstractions (*observe*). Here, a domain abstraction can be a virtual topology consisting of domain border nodes and virtual links connecting them [10] or simply a virtual node. The broker then aggregates the collected information and invokes its corresponding ML model to extract knowledge about the behaviors and state of the entire multi-domain EON (*analyze*). With the acquired knowledge, the broker eventually determines the most appropriate inter-domain service scheme (*act*).

## 3. Distributed collaborative learning approach

This section elaborates on the principle of the distributed collaborative learning approach. As shown in Fig. 2, each ML model (e.g., a deep neural network, DNN [9]) for inter-domain networking consists of multiple distributed learning blocks (one per domain) and a broker plane aggregation learning block. The distributed learning blocks have full access to the network state data of the corresponding domains, from which they can extract proper features representing the state of every domain. Note that, as the structures and scales of domain state data can be different (for instance, due to different topologies, routing schemes, or state modeling schemes), each DM needs to decide an appropriate distributed block design separately to enable adequate data representation and knowledge extraction capabilities while mitigating overfitting issues. The features learned by the distributed blocks are then reported to the broker plane. The aggregation learning block takes as input the collected domain-level features, as well as certain amounts of additional data featuring the observations by the broker plane depending on a specific application considered (for instance, performing RMSA requires information of inter-connectivity among domains), to generate the learning targets for inter-domain networking. The distributed learning blocks essentially work as black boxes and it is difficult to derive the inputs of each block only from the reported features without knowing the structure of the learning block (e.g., the number
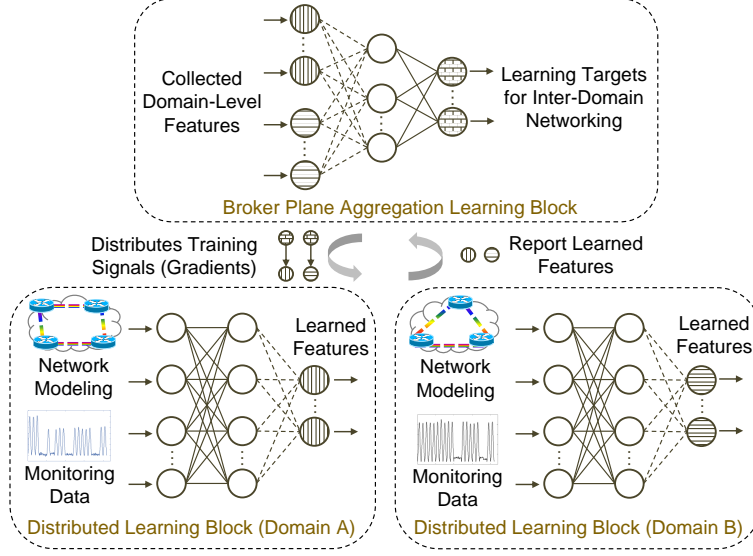
Fig. 2. Principle of distributed collaborative learning.

of layers and the number of neurons at each layer of the neural network adopted), the scale and dimension of the inputs, and how input data are preprocessed. Thus, the proposed approach enables strict domain privacy while still allowing the ML models to exploit the original and complete network state data. Note that, the proposed approach is different from the federated learning approach in [25], which attempts to reduce the communication overhead due to data collection by employing multiple learning blocks (with exactly the same structure) that learn in parallel from distributed data. Therefore, federated learning is still a perfect-information model and cannot be applied to multi-domain networking.

We adopt the chain rule to train the distributed and the aggregation learning blocks as ensembles. In other words, the learning blocks are trained jointly under the supervision of common training signals for achieving the same goals, thus, obviating the need for artificial domain-level featuring. We define a distributed collaborative learning ML model with $M$ distributed blocks as,

$$
\begin{aligned}
y &= F_{\theta_0}(x_1, \cdots, x_M, s_0), \\
x_i &= f_{\theta_i}(s_i),
\end{aligned}
\tag{1}
$$

where $f_{\theta_i}(\cdot)$ represents the learning block belonging to domain $i$ ($1 \le i \le M$) with parameter set $\theta_i$ and input $s_i$, $F_{\theta_0}(\cdot)$ represents the aggregation learning model with parameter set $\theta_0$, $s_0$ is the additional broker plane input, and $y$ is the eventual learning target. Given a dataset $\mathcal{D}$ of $N$ instances, the training aims to minimize a loss function $\mathcal{L}(\theta)$ ($\theta = \{\theta_1, \cdots, \theta_M, \theta_0\}$). For instance, when considering a supervised regression task, where each data instance $n$ is associated with a label $\hat{y}_n$, $\mathcal{L}(\theta)$ is often calculated as the mean square error (MSE) over $\mathcal{D}$ for minimizing the overall prediction error from the model, i.e.,

$$
\mathcal{L}(\theta) = \frac{1}{N} \sum_{1 \le n \le N} (y_n - \hat{y}_n)^2.
\tag{2}
$$

The aggregation learning block can be trained with existing gradient-based optimization methods (e.g., Adam [26]) by iteratively computing and applying the gradients of $\mathcal{L}(\theta)$ regarding $\theta_0$ (denoted as $\nabla_{\theta_0} \mathcal{L}$). In the meantime, the aggregation block distributes the gradients of $\mathcal{L}(\theta)$ regarding $x_i$ to the distributed blocks, which in turn can obtain the gradients of $\mathcal{L}(\theta)$ regarding $\theta_i$

applying the chain rule. The chain rule can be formulated as,

$$\nabla_{\theta_i}\mathcal{L} = \mathbb{E}\left[\nabla_{x_i}\mathcal{L}|_{s_0=s_{0,n}, x_i=f_{\theta_i}(s_{i,n})} \cdot \nabla_{\theta_i}f_{\theta_i}(s_i)|_{s_i=s_{i,n}}\right],\tag{3}$$

where $\{s_{1,n}, \cdots, s_{M,n}, s_{0,n}\}$ is the input of a specific data instance (indexed by $n$) in $\mathcal{D}$. Such a chain rule-based training mechanism allows each distributed block to be trained independently, and thus, enables customized learning block designs by DMs.

The distributed collaborative learning approach serves as a generic framework and can be applied to inter-domain networking applications using different ML techniques (e.g., supervised learning, reinforcement learning). For instance, we could extend the deep reinforcement learning scheme for RMSA proposed in [18] to a multi-domain scenario by implementing the policy and value DNNs with distributed learning. Another application, which will be discussed and demonstrated in the following section, can be end-to-end QoT estimation for inter-domain lightpaths. A practical concern from the network operation point of view is the communication overhead introduced by the interactions between the learning blocks. Let $Z_i$ denote the number of output features of domain $i$ and consider a dataset $\mathcal{D}$ of size $N$ with which we train a distributed learning model, the amount of numerical values sent by the distributed blocks to the aggregation block during each training epoch is $N \cdot \sum Z_i$, and vice-versa. Further, consider a multi-domain EON with ten domains, $Z_i = 100$, and a training process of a thousand epochs, the overall communication overhead could range from several to a few hundred gigabytes, depending on the value of $N$. However, since training or retraining is usually performed offline (i.e., do not impose very short and strict deadlines and can be scheduled at nighttime), the resultant overhead can be deemed as tolerable. Once an ML model is trained and deployed for online operations, the overhead becomes negligible as the model only needs to process data instances for a few service requests at a time.

## 4. Use case: end-to-end QoT estimation

We investigate the design of end-to-end QoT estimators as a use case of the distributed collaborative learning approach. Figure 3 gives an illustrative example of the proposed designs. Specifically, the learning blocks are built upon fully-connected neural networks. Similar to our previous work in [19], the distributed learning blocks take as input the optical performance monitoring data (e.g., link load, channel power) collected at each intermediate monitoring point along inter-domain routing paths $(P_1, \cdots, P_k)$ and output domain-level features. As different intra-domain path segments can be composed of different numbers of links (i.e., consisting of different numbers of monitoring points), leading to different scales of inputs, each DM employs multiple learning blocks (with different structures) for different segments. The domain-level features are reported to the broker plane, which in turn can employ the following designs for aggregation blocks.

*1) Independent learning.* A straightforward design is to employ multiple aggregation blocks, each for predicting the QoT of a different inter-domain path (see the illustration in Fig. 3(a)). Thus, we can obtain multiple independent QoT estimators and train them by directly applying the chain rule-based mechanism discussed in Section 3 for minimizing the estimation errors. We denote this design as DisLearn.

*2) Multi-task learning (MTL).* DisLearn requires collecting a sufficient amount of performance monitoring data for each inter-domain path for training accurate QoT estimators, which can be costly and impractical for real network operations. To cope with this issue, we leverage the concept of multi-task learning (MTL) [27] and propose an integrated aggregation block design (denoted as DisLearn-MTL) that performs QoT estimation for multiple inter-domain lightpaths simultaneously. Figure 3(b) shows the principle of DisLearn-MTL. The aggregation block is built of a few shared hidden layers receiving and processing input features from all the distributed learning blocks and separate output heads generating the QoT estimation for
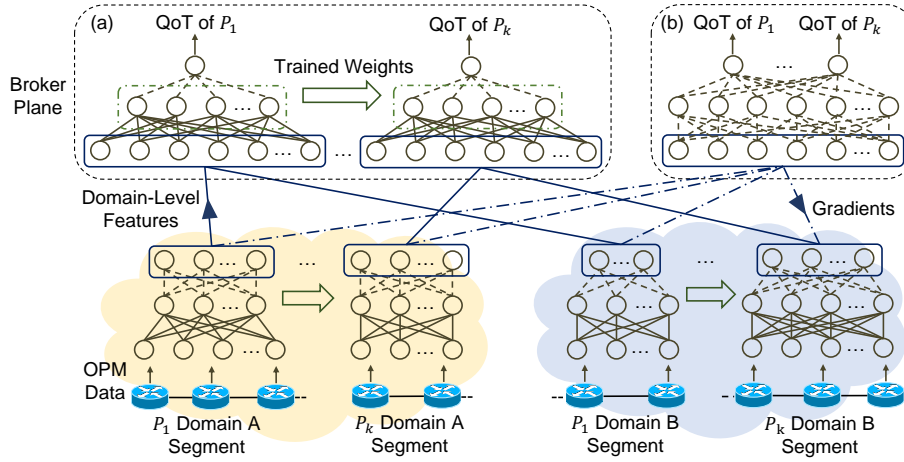
Fig. 3. Distributed learning-based end-to-end QoT estimator designs: (a) separate aggregation learning block design for independent learning or transferring learning, (b) shared aggregation learning block design for multi-task learning. OPM: optical performance monitoring.

each lightpath. Similarly, the aggregation and distributed learning blocks can be trained jointly with the chain rule-based mechanism for minimizing the overall estimation errors taking the data of all the lightpaths. Since MTL enables learning from data of multiple relevant tasks at the same time (enabling knowledge sharing among tasks) and combat overfitting (by learning more generalized feature extraction functions), DisLearn-MTL can potentially achieve improved estimation accuracy with a reduced amount of data for each lightpath.

*3) Transfer learning (TL).* TL is an ML technique that can accelerate and improve the performance of the training of target ML tasks by reusing the knowledge learned for related source tasks [28]. Therefore, TL is another promising technique for enhancing the scalability of the distributed learning-based QoT estimator designs. Figure 3(a) shows the schematic of a third aggregation block design incorporating TL (denoted as DisLearn-TL). The arrangement of the distributed and aggregation blocks in DisLearn-TL is similar to that in DisLearn, where separate aggregation blocks are deployed for different lightpaths. The QoT estimators are trained sequentially according to the chain rule. In particular, each time we first copy a portion of weights from a fine-tuned learning block (typically, weights of lower layers of a neural network as they are less biased or overfitted) to the target learning block and initialize the rest of the weights randomly. Then, a normal training process as that in DisLearn is performed. We envision such a knowledge transferring mechanism to facilitate effective training of QoT estimators while remarkably reducing the amounts of data required for most of the lightpaths.

## 5. Performance evaluation

We evaluated the performance of the proposed QoT estimator designs with experimental data collected from a two-domain EON testbed (with seven nodes) shown in Fig. 4. We set up five inter-domain lightpaths consisting of three, four, three, five, and four nodes, respectively, and measured their end-to-end Q-factors using a coherent receiver. To create adequate numbers of data instances well representing the feature spaces of the QoT estimation tasks, we emulate various link load scenarios and device conditions by injecting different numbers of background wavelength channels and introducing different degrees of wavelength-specific attenuation at the wavelength-selective switches (WSS's) at different locations. Each optical spectrum analyzer (OSA) along the lightpaths monitored the power of all the wavelength channels (in total 21, including one testing and 20 background channels) and the noise floor. Thus, each data instance

corresponds to a network condition where the links on a path have different numbers of active channels with different power and noise levels. We processed the original monitoring data to generate data instances which each has $4 \times H_j$ input features, conveying (i) the power of the testing channel, (ii) the number of active background channels, (iii) the average power of background channels, and (iv) the noise floor observed at $H_j$ OSAs along the corresponding lightpath (indexed by $j$). Figure 5(a) summarily gives a description of the dataset used for evaluations. In total, we obtained 1450, 316, 271, 243, and 278 data instances, respectively, for the five paths. Figure 5(b) shows the frequency of occurrence plot with respect to the measured Q-factors.



Fig. 4. Testbed setup for experimental dataset generation.



| Path ID | # of Nodes in Domain A | # of Nodes in Domain B | Size of Data Set |
|---------|------------------------|------------------------|------------------|
| 1 | 2 | 1 | 1450 |
| 2 | 2 | 2 | 316 |
| 3 | 1 | 2 | 271 |
| 4 | 3 | 2 | 243 |
| 5 | 2 | 2 | 278 |

(a)

Fig. 5. (a) Description of the experimental datasets, (b) distribution of the Q-factors.

Taking into account the complexity of the QoT estimation tasks and the relatively small amounts of data collected, we implemented the distributed learning-based QoT estimator designs with small-scale neural networks. In particular, for all the five paths, we implemented the distributed learning blocks with two-layer neural networks, each of which has an output layer of four neurons (i.e., outputting four domain-level features) directly connected to the input layer. Three-layer neural network architectures were adopted for the aggregation blocks ([8, 8, 1] for DisLearn and DisLearn-TL, and [8, 8, 3] for DisLearn-MTL). We compared the performance of the proposed designs with that of the perfect-information models, which are based on the assumption of a multi-domain dictator having full intra-domain visibility. The perfect-information models essentially reduce the inter-domain QoT estimation task to an already well-investigated intra-domain one, and thus, provide the upper bounds of performance. For fair comparisons, we implemented the perfect-information models with fully-connected neural networks of two hidden layers ([8, 8]). All the neural networks used *ReLU* as the activation function.

Figure 6 shows the comparisons between DisLearn and the perfect-information model with the data of path 1. We picked 85% of the data to build the training set while the rest were used for validation and testing. The results of training and validation losses (defined as MSE) in Figs. 6(a) and 6(c) show that both of the models converged after training of 200 epochs without the presence of notable overfitting, verifying the appropriateness of the configurations of the learning blocks.
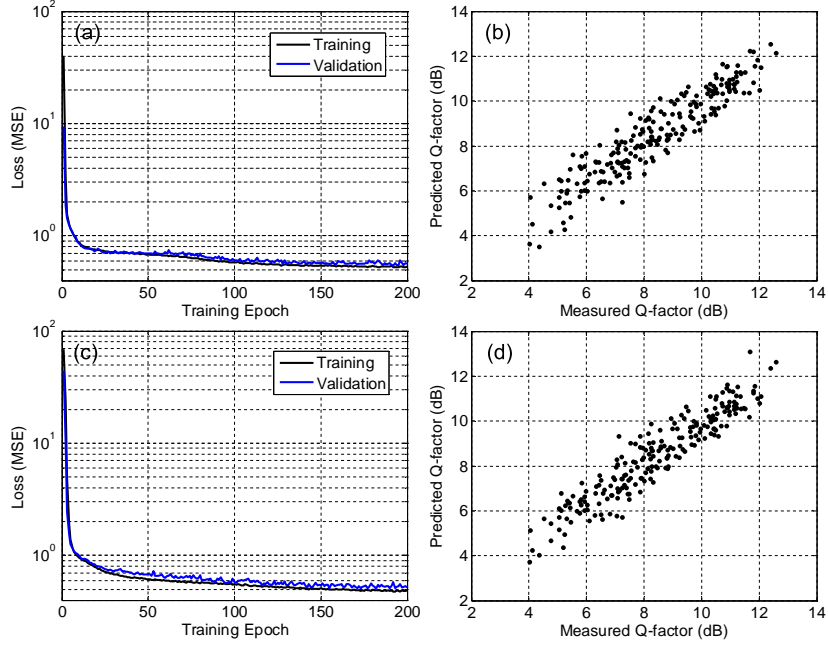
Fig. 6. Comparisons between DisLearn ((a), (b)) and the perfect-information model ((c), (d)) with data of path 1: (a) and (c) losses during training, (b) and (d) comparison between measured and predicted Q-factors.

Table 1. Results of estimation accuracy for paths 1-3.

| Model | Path 1 | | Path 2 | | Path 3 | |
|---|---|---|---|---|---|---|
| | $\eta$ | $\varepsilon$ (dB) | $\eta$ | $\varepsilon$ (dB) | $\eta$ | $\varepsilon$ (dB) |
| DisLearn | 92.69% | 0.59 | 89.71% | 0.90 | 91.21% | 0.67 |
| Perfect-Info | 92.78% | 0.58 | 89.85% | 0.89 | 91.68% | 0.63 |
| DisLearn-MTL | 92.85% | 0.57 | 90.54% | 0.87 | 92.32% | 0.60 |
| DisLearn-TL | - | - | 90.64% | 0.86 | 92.11% | 0.63 |

To visualize the performance of the two models after training, we plotted the comparisons between the estimated and the measured Q-factors in Figs. 6(b) and 6(d). We first assumed that paths 1-3 were established and Table 1 summarizes the results of mean estimation accuracy ($\eta$) and error ($\varepsilon$) from the different models. The accuracy is defined as $\eta = 1 - |Q_{mea} - Q_{est}|/Q_{mea}$, where $Q_{est}$ and $Q_{mea}$ represent the estimated and the measured Q-factors, respectively. Note that, each data point in the table is an average of the results from ten independent experiments with randomly partitioned training and testing datasets. Firstly, we can see that DisLearn can always achieve estimation accuracies very close to those of the perfect-information models (with differences less than 0.5%), which clearly verifies the effectiveness of the proposed distributed collaborative learning approach. The accuracies for paths 2 and 3 are lower than those for path 1 due to the smaller amounts of training data available. Meanwhile, the results indicate that by implicitly exploiting relevant knowledge from the data of multiple paths and directly transferring the knowledge learned, respectively, both DisLearn-MTL and DisLearn-TL achieve higher estimation accuracies than DisLearn. We further evaluated the proposed designs taking paths 4 and 5 as newly added lightpaths in the network. For DisLearn-TL, we reused the learned

parameters from the DisLearn-MTL model for paths 1-3 which performs the best. We did not evaluate DisLearn-MTL taking all the five paths as it entails retraining from scratch. Table 2 presents the results of estimation accuracy, which are consistent with those in Table 1. We can see that the accuracies of DisLearn and the perfect-information models are very close and that TL can facilitate better performance.

Table 2. Results of estimation accuracy for paths 4 and 5.

| Model | Path 4 | | Path 5 | |
|---|---|---|---|---|
| | $\eta$ | $\varepsilon$ (dB) | $\eta$ | $\varepsilon$ (dB) |
| DisLearn | 88.71% | 0.95 | 90.18% | 0.88 |
| Perfect-Info | 88.61% | 0.96 | 90.17% | 0.89 |
| DisLearn-TL | 89.50% | 0.89 | 90.96% | 0.84 |

## 6.  Conclusion

This paper proposed a distributed collaborative learning approach for meeting the challenges of knowledge-defined multi-domain EON and presented a study of end-to-end QoT estimation with the proposed approach. Performance evaluations with experimental data demonstrated the effectiveness of the proposed designs.

As future works, we will consider more comprehensive experimental scenarios with diversified routing schemes, modulation formats, baud rates, channel spacing, fiber and equipment types, and collect larger amounts of data for further verifying the performance of the proposed QoT estimator designs. We will also study the applicability of the distributed learning approach for learning effective resource allocation policies in multi-domain EONs by extending the deep reinforcement learning design in [18]. Meanwhile, we will incorporate securing domain privacy and data confidentiality as a key aspect of our future investigations.

## Disclosures

The authors declare no conflicts of interest.

## References

1. Z. Zhu, X. Chen, C. Chen, S. Ma, M. Zhang, L. Liu, and S. J. B. Yoo, "OpenFlow-assisted online defragmentation in single-/multi-domain software-defined elastic optical networks," J. Opt. Commun. Netw. **7**, A7–A15 (2015).
2. O. Gerstel, M. Jinno, A. Lord, and S. J. B. Yoo, "Elastic optical networking: a new dawn for the optical layer?" IEEE Commun. Mag. **50**, S12–S20 (2012).
3. Z. Zhu, W. Lu, L. Zhang, and N. Ansari, "Dynamic service provisioning in elastic optical networks with hybrid single-/multi-path routing," J. Light. Technol. **31**, 15–22 (2013).
4. P. Lu, L. Zhang, X. Liu, J. Yao, and Z. Zhu, "Highly-efficient data migration and backup for big data applications in elastic optical inter-datacenter networks," IEEE Netw. **29**, 36–42 (2015).

5. Z. Zhu, C. Chen, X. Chen, S. Ma, L. Liu, X. Feng, and S. J. B. Yoo, "Demonstration of cooperative resource allocation in an OpenFlow-controlled multidomain and multinational SD-EON testbed," J. Lightw. Technol. **33**, 1508–1514 (2015).

6. A. Giorgetti, "Proactive H-PCE architecture with BGP-LS update for multidomain elastic optical networks," J. Opt. Commun. Netw. **7**, 1–9 (2015).

7. X. Chen, C. Chen, D. Hu, S. Ma, and Z. Zhu, "Multi-domain fragmentation-aware RSA operations through cooperative hierarchical controllers in SD-EONs," in *Proc. Asia Commun. Photon. Conf.*, (2015). Paper AM2E.3.

8. X. Chen, B. Li, R. Proietti, Z. Zhu, and S. J. B. Yoo, "Multi-agent deep reinforcement learning in cognitive inter-domain networking with multi-broker orchestration," in *Proc. Conf. Opt. Fiber Commun.*, (2019). Paper M2A.2.

9. Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature **521**, 436–444 (2015).

10. X. Chen, R. Proietti, H. Lu, A. Castro, and S. J. B. Yoo, "Knowledge-based autonomous service provisioning in multi-domain elastic optical networks," IEEE Commun. Mag. **56**, 152–158 (2018).

11. F. Musumeci, C. Rottondi, A. Nag, T. Macaluso, D. Zibar, M. Ruffini, and M. Tornatore, "An overview on application of machine learning techniques in optical networks," IEEE Commun. Surv. Tutor. **21**, 1383–1408 (2019).

12. F. Khan, Q. Fan, C. Lu, and A. Lau, "An optical communication's perspective on machine learning and its applications," J. Light. Technol. **37**, 493–516 (2019).

13. C. Rottondi, L. Barletta, A. Giusti, and M. Tornatore, "Machine-learning method for quality of transmission prediction of unestablished lightpaths," J. Opt. Commun. Netw. **10**, A286–A297 (2018).

14. R. Proietti, X. Chen, K. Zhang, G. Liu, M. Shamsabardeh, A. Castro, L. Velasco, Z. Zhu, and S. J. B. Yoo, "Experimental demonstration of machine-learning-aided QoT estimation in multi-domain elastic optical networks with alien wavelengths," J. Opt. Commun. Netw. **11**, A1–A10 (2019).

15. D. Rafique, T. Szyrkowiec, H. Grießer, A. Autenrieth, and J. Elbers, "Cognitive assurance architecture for optical network fault management," J. Light. Technol. **36**, 1443–1450 (2018).

16. X. Chen, B. Li, R. Proietti, Z. Zhu, and S. J. B. Yoo, "Self-taught anomaly detection with hybrid unsupervised/supervised machine learning in optical networks," J. Light. Technol. **37**, 1742–1749 (2019).

17. J. Suárez-Varela, A. Mestres, J. Yu, L. Kuang, H. Feng, P. Barlet-Ros, and A. Cabellos-Aparicio, "Routing based on deep reinforcement learning in optical transport networks," in *Proc. Conf. Opt. Fiber Commun.*, (2019). Paper M2A.6.

18. X. Chen, B. Li, R. Proietti, H. Lu, Z. Zhu, and S. J. B. Yoo, "DeepRMSA: A deep reinforcement learning framework for routing, modulation and spectrum assignment in elastic optical networks," J. Light. Technol. **37**, 4155–4163 (2019).

19. G. Liu, K. Zhang, X. Chen, H. Lu, J. Guo, J. Yin, R. Proietti, Z. Zhu, and S. J. B. Yoo, "The first testbed demonstration of cognitive end-to-end optical service provisioning with hierarchical learning across multiple autonomous systems," in *Proc. Conf. Opt. Fiber Commun.*, (2018). Paper Th4D.7.

20. X. Chen, Z. Zhu, L. Sun, J. Yin, S. Zhu, A. Castro, and S. J. B. Yoo, "Incentive-driven bidding strategy for brokers to compete for service provisioning tasks in multi-domain SD-EONs," J. Light. Technol. **34**, 3867–3876 (2016).

21. Y. Yin, H. Zhang, M. Zhang, M. Xia, Z. Zhu, S. Dahlfort, and S. Yoo, "Spectral and spatial 2D fragmentation-aware routing and spectrum assignment algorithms in elastic optical networks," J. Opt. Commun. Netw. **5**, A100–A106 (2013).

22. L. Gong, X. Zhou, X. Liu, W. Zhao, W. Lu, and Z. Zhu, "Efficient resource allocation for all-optical multicasting over spectrum-sliced elastic optical networks," J. Opt. Commun. Netw. **5**, 836–847 (2013).

23. L. Gong and Z. Zhu, "Virtual optical network embedding (VONE) over elastic optical networks," J. Light. Technol. **32**, 450–460 (2014).

24. S. J. B. Yoo, "Multi-domain cognitive optical software defined networks with market-driven brokers," in *Proc. Eur. Conf. Opt. Commun.*, (2014). Paper WE.2.6.3.

25. H. McMahan, E. Moore, D. Ramage, S. Hampson, and B. Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Proc. Int. Conf. Artif. Intell. Statist.*, vol. 54 (2017), pp. 1273–1282.

26. D. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Int. Conf. Learn. Repr.*, (2015). ArXiv:1412.6980.

27. T. Evgeniou and M. Pontil, "Regularized multi-task learning," in *Proc. Conf. Knowl. Disc. Data Mining,* (2004), pp. 109–117.

28. S. Pan and Q. Yang, "A survey on transfer learning," Trans. Knowl. Data Eng. **22**, 1345–1359 (2010).