QoT Estimation with Margin-Driven Transfer Learning in Time-Varying Optical Networks

QoT Estimation with Margin-Driven Transfer Learning in Time-Varying Optical Networks

Piotr Lechowicz[1,*], Carlos Natalino[1], and Paolo Monti[1]

[1] *Department of Electrical Engineering, Chalmers University of Technology, Gothenburg, Sweden*

[*] *piotr.lechowicz@chalmers.se*

Abstract: Estimating transmission quality in an optical network is critical for resource efficiency but challenging due to the infrastructure time-varying state. We propose a transfer learning solution to adapt a data-driven model to network changes.

1. Introduction

Optical network infrastructure characteristics evolve during their lifetime due to network upgrades, equipment replacements, and aging. As a result, beginning of life (BoL) parameters might quickly become outdated. This behavior makes estimating quality of transmission (QoT) particularly challenging since it requires detailed and up-to-date knowledge of physical layer parameters. Promptly obtaining such parameters from the network is costly and sometimes infeasible, depending on the network scale and device capabilities. Machine Learning (ML) models for QoT estimation can provide good estimations of physical layer parameters also in the presence of uncertainties [1]. With time-varying parameters, a QoT estimation model can quickly become outdated, undermining its purpose and potentially prompting lightpath re-establishment procedures, which operators would prefer to avoid. One potential solution to this issue is to retrain the model whenever its accuracy drops below a certain threshold. However, this approach introduces additional complexity, requiring collecting a larger set of training parameters and additional time to re-train a model. In this context, transfer learning (TL) allows the transfer of knowledge of ML models from a source to a target domain to improve the learning without requiring as many samples compared to training a new model from scratch. The literature on TL currently focuses on training models where physical layer parameters might be uncertain but do not vary with time, i.e., transferring models between two different network topologies using the same hardware [2], or transferring a model from one fiber type to another [3]. In [4], TL has been applied to transfer QoT forecasts over a short period of time between lightpath (LP)s. To the best of our knowledge, the benefits of using TL in the presence of time-varying physical layer parameters have not been studied so far. In particular, the questions related to when to trigger model updates and for how long a model is suitable have not been answered.

This paper proposes a framework for handling the QoT estimation in time-varying optical networks using TL. We evaluate a data-driven trigger for model updates and propose a margin-driven TL that tracks the impact of generalized signal-to-noise ratio (GSNR) inaccuracies on the LP provisioning. We compare the proposed solution to two benchmarks: an analytical model with up-to-date knowledge of network parameters and a model periodically relearned from scratch while the network ages. Results show that TL outperforms the fully retrained model at the beginning of the network lifetime. As the network changes, the model that uses TL becomes less effective, making a complete retraining cycle more suitable.

2. Network Model and Proposed Margin-Driven Transfer Learning

We assume an elastic optical network (EON) scenario. The deployed optical network, presented in Fig. 1(a), comprises fiber links built as sequences of amplified spans. Each span is characterized by a fiber loss and the amplifier noise figure. These two parameters are subject to variations, i.e., degradation in the case of aging or improvements in the case of replacement. Such variations will impact the GSNR of the lightpaths traversing the components. Transceivers at nodes are endpoints to LPs, and can also monitor their QoT. The controller depicted in Fig. 1(b) is responsible for provisioning incoming LP requests and monitoring the network performance. The controller consists of a monitoring module that tracks network parameters based on telemetry data, a degradation estimation module that determines network aging based on the measurements, a QoT estimation module, and an Routing, Modulation and Spectrum Assignment (RMSA) module responsible for LP provisioning. A QoT-aware modulation format (MF) selection is adopted, which takes advantage of the results of the QoT estimator module and can apply various strategies to determine the best MF as part of the RMSA process.

Fig. 1(c) shows the GSNR evolution of three LPs (denoted as LP_1, LP_2, and LP_3), established in different moments. For simplicity, it is assumed that LPs are provisioned using the same MF with required GSNR denoted as MF_{th}. To account for aging, a margin is needed, denoted as $MF_{th}+m$. The estimated GSNR at the LP BoL is denoted as g'_1, g'_2, and g'_3. As the network ages, the experienced GSNR decreases, denoted as $g1, g_2$ and g_3. To compensate for GSNR estimation inaccuracy, the controller can add a penalty to the margin $MF + pm$. In the illustrative example of Fig. 1(c), the estimated g'_2 has enough GSNR margin to compensate for aging. However, LP_3

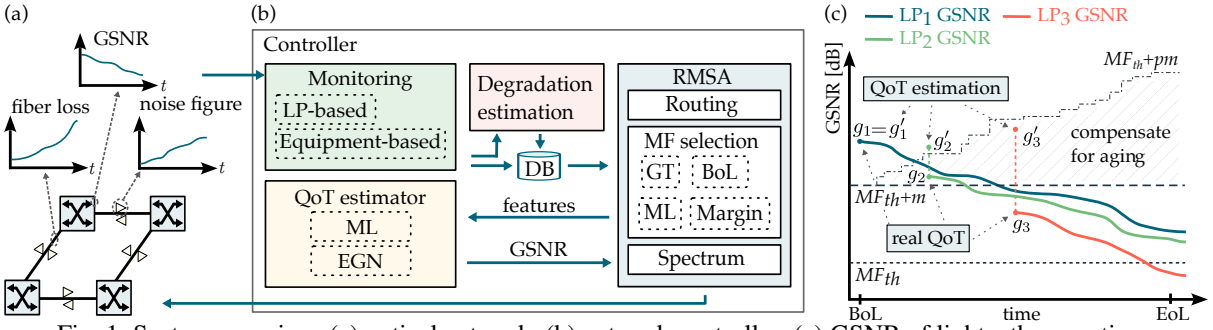

Fig. 1: System overview; (a) optical network; (b) network controller; (c) GSNR of lightpaths over time.

is in risk, as the estimated GSNR g'_3 is above $MF_{th}+m$ at provisioning time, but cannot assure sufficient QoT by the LP end of life (EoL). After detecting sufficient accuracy degradation of the QoT estimator, the controller can trigger a mitigation. The accuracy degradation can be data-driven, e.g., by tracking the mean absolute error (MAE) between the estimated and experienced GSNR, and triggering the mitigation when it crosses a threshold. Alternatively, we propose in this paper to use a margin-driven approach, i.e., to track the impact of GSNR inaccuracies on the LP provisioning. If an analytical model is adopted, increasing the penalty can be a mitigation. If a ML-based model is adopted, the selected mitigation method may refer to a full model retraining, or to TL.

3. Numerical Results

We adopt a dynamic RMSA scenario over an EON to evaluate the impact of network evolution on the overall network performance. We consider an EON over the European network topology with 26 nodes and 42 links. The aging is characterized by an increase in the fiber loss (e.g., due to splicing of broken fiber), and and increased amplifier noise figure due to aging of components. We consider five modulation formats: QPSK, 8-, 16-, 32-, and 64-QAM, with their GSNR threshold set to 6.72, 10.84, 13.24, 16.16, and 19.01 dB, respectively. Launch power is fixed at 0 dBm. The GSNR ground truth (GT) is computed using the enhanced Gaussian noise (EGN) model [5] assuming up-to-date knowledge of all network parameters over its lifetime. The RMSA works as follows. The routing selects the shortest-path among a set of five shortest-paths pre-computed between each node pair in the topology. A QoT-aware MF selection is considered, where the MF with highest spectral efficiency possible is selected provided that the GSNR requirements are met.

We generate an initial training dataset to serve as the basis for training the ML models by generating 100,000 LPs requests. For each LP, a MF is selected based on the GSNR evaluated using EGN analytical formula with a 2 dB margin. Out of 100,000 requests, 88,815 are accepted, for which LP features and resulting GSNR are recorded. We use an artificial neural network (ANN) as the ML-based GSNR estimator, with a similar architecture as in [6]. The input contains 17 features containing node, path and lightpath properties, and hyperbolic tangent activation function. One hidden layer contains 256 neurons and hyperbolic tangent activation function. The output is activated using a linear function. The BoL ML model is trained using generated dataset with 50%-50% train-test split, using mean square error (MSE) loss function, RMSprop optimizer with 10^{-4} learning rate. The achieved MAE of GSNR on test set is equal to 0.145 dB.

To evaluate the impact of aging, we simulate the network operation over 300,000 requests corresponding to 6 years of network operation. Requests arrives following an exponential distribution with mean of 10.5 minutes for connecting a randomly selected node pair. Holding time of requests follow an exponential distribution with mean of 21 hours. Network aging is simulated through an event every 17.5 hours that degrades *(i)* the attenuation of one fiber span randomly selected in the network by $0.05/s_l$ to $0.1/s_l$, where s_l is span length in km; and the noise figure of one randomly selected amplifier by 0.1 to 0.5 dB. We compare 6 GSNR estimation methods during the simulation. GT is an ideal model based on the analytical EGN model and up-to-date knowledge of all the network parameters and their change during network evolution. EGN-BoL is a model based on the EGN considering the BoL state of the network. EGN-Pen adds to the EGN a margin to compensate for the aging, representing a traditional approach to handle aging when using analytical models. EGN-Pen tracks the margin for all the network links individually. While estimating the GSNR for a path, the model accounts for the highest penalty among the links belonging to that path. After establishing a lightpath, a mismatch between the estimated GSNR and monitored one is recorded and link penalties are updated accordingly. ML-MAE represents a data-driven TL approach. ML-MAE performs TL on the ML model, starting with the BoL one, each time the MAE between the estimated and GT GSNR degrades above a given threshold, i.e., 0.25 dB in our case. ML-MRG represents the proposed margin-driven TL approach. ML-MRG performs TL on the ML model, starting with the BoL one, each time the transfer triggering criteria is met. The criteria considered is a 10% threshold of allocation retries due to the wrong MF selection over 1,000 LPs. TL is applied by freezing the parameters of the hidden ANN layer and learning over 20 epochs while minimizing MSE. Finally, ML-R is an ML model retrained every 1 year

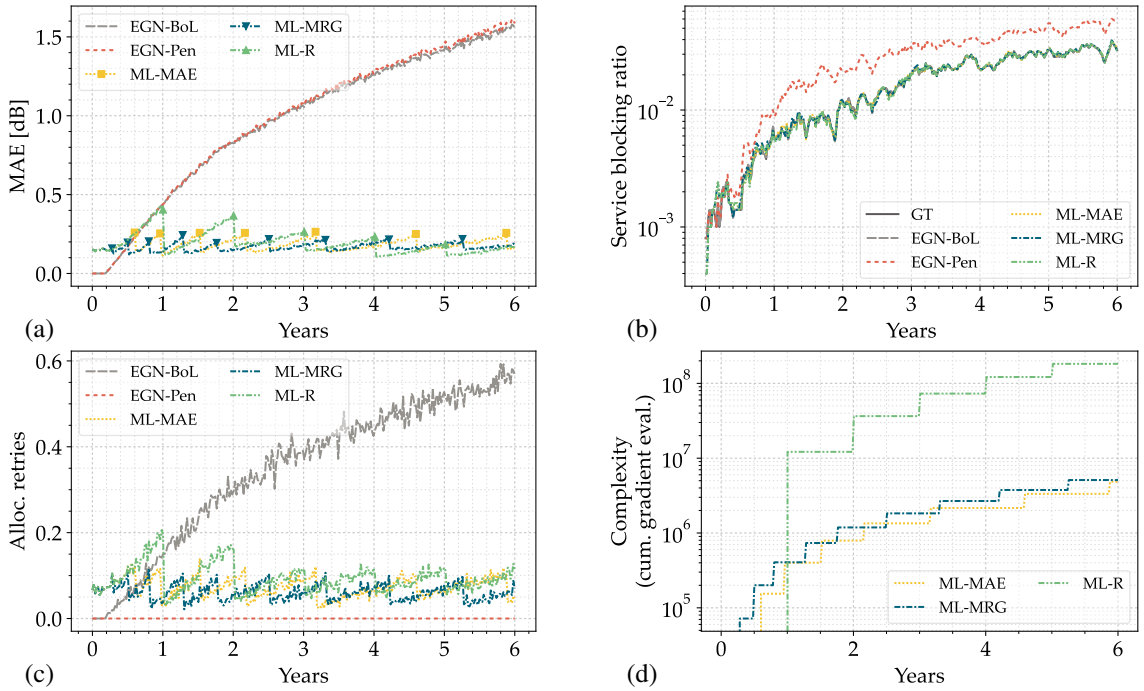

Fig. 2: Performance of proposed QoT estimation strategies over years; (a) mean absolute error (MAE); (b) service blocking ratio for QoT estimation strategies; (c) established lightpaths allocation retries for QoT estimation strategies; (d) complexity as a cumulative sum of gradient evaluations for the ML approaches.

for 200 epochs based on the full collected data over this period. The performance is assessed in terms of MAE between the GSNR of an approach and the GT, the service blocking ratio, the number of LP allocation retries due to insufficient GSNR measured after establishing the LP, and the ML cumulative complexity measured as the total number of gradient evaluations.

Fig. 2a presents MAE for proposed QoT estimation strategies. Progressive accuracy decrease is observed for the BoL model. For ML-MAE, ML-MRG, and ML-R, MAE has a cyclic behavior where the MAE increases, followed by a sharp drop when TL/retraining cycles are triggered. Analyzing the ML-based estimators, TL (ML-MAE and ML-MRG) performs better during the first 3 years, but ML-R shows slightly better accuracy in the later stages of network operation. This behavior indicates that TL can take advantage of the trained knowledge during the early stages of the network. However, as the network ages, a completely new model may be more suitable due to the substantially different state of the underlying infrastructure. Fig. 2b shows service blocking ratio for proposed QoT estimation methods. All the methods, except for EGN-Pen, have similar service blocking. This is expected because as the network ages, the GSNR of selected LP starts to be lower, resulting in the selection of less spectrally efficient MF and, as a consequence, higher blocking. EGN-Pen considers margin and penalty when selecting the MF, resulting in potential underestimation of MF, and consequent higher blocking. The above conclusions are confirmed by Fig. 2c which visualizes normalized number of established lightpath allocation retries. EGN-Pen achieves near-zero retries as its penalty prevents MF overestimation. For BoL methods, the number of retries increases similarly as the MAE, as the methods start to overestimate MF selection. For ML-MRG and ML-R the retries drops according to the transferring/relearning the models. Fig. 2d shows complexity of ML-MAE, ML-MRG, and ML-R models, which intuitively indicate that TL has lower complexity.

4. Conclusion

In this work, we investigate how and to what extent TL is beneficial for QoT estimation of newly established LP in time-varying optical networks. Results show that TL works well when the initial knowledge matches more closely the current one. As the network evolves, full training cycles are required.

References

1. M. Lonardi *et al.*, *OSA Advanced Photonics Congress (AP)*, 2020, 10.1364/NETWORKS.2020.NeM3B.2
2. J. Yu *et al.*, *Journal of Optical Communications and Networking*, 2019, 10.1364/JOCN.11.000C48
3. I. Khan *et al.*, *Journal of Optical Communications and Networking*, 2021, 10.1364/JOCN.409538
4. S. Allogba *et al.*, *Journal of Lightwave Technology*, 2022, 10.1109/JLT.2022.3160379
5. M. R. Zefreh *et al.*, *Journal of Lightwave Technology*, 2020, 10.1109/JLT.2020.2997395
6. G. Bergk *et al.*, *Journal of Optical Communications and Networking*, 2022, 10.1364/JOCN.442733