Advanced Formulation of QoT-Estimation for Un-established Lightpaths Using Cross-train Machine Learning Methods

(Article begins on next page)

Advanced Formulation of QoT-Estimation for Un-established Lighpaths Using Cross-train Machine Learning Methods

Ihtesham Khan[1], **Muhammad Bilal**[1], **Vittorio Curri**[1]
[1]*Politecnico di Torino, Corso Duca degli Abruzzi, 24, 10129, Torino, Italy*
e-mail: ihtesham.khan@polito.it

ABSTRACT Planning tools with excellent accuracy along with precise and advance estimation of the quality of transmission (QoT) of lightpaths (LPs) have techno-economic importance for a network operator. The QoT metric of LPs is defined by the generalized signal-to-noise ratio (GSNR) which includes the effect of both amplified spontaneous emission (ASE) noise and non-linear interference (NLI) accumulation. Typically, a considerable number of analytical models are available for the estimation of QoT but all of them require the *exact* description of system parameters. Thus, the analytical models are impractical in case of *un-used* network scenarios. In this study, we exploit an alternative approach based on three machine learning (ML) techniques for QoT estimation (QoT-E). The proposed ML based techniques are cross-trained on the characteristic features extracted from the telemetry data of the already *in-service* network. This new approach provides a reliable QoT-E and consequently assists the network operator in network planning and also enables the reliable low-margin LP deployment.

Keywords: Cross-train Machine Learning; Quality of Transmission Estimation; Generalized OSNR.

1. INTRODUCTION

During the past two decades the dramatic increase in the global IP traffic, urged by the introduction of 5G technology along with the expansion of bandwidth hungry applications; such as high definition (HD) video and virtual and augmented reality (VR and AR) contents, is forecast for the coming years [1]. These recent trend of growing traffic generated much research interest towards the flexible and dynamic optical network architecture. In the prosecution of this, technologies such as elastic optical network (EON) and software-defined networking (SDN) are introduced in the past few years. The classical feature of EON and SDN paradigm in the optical networks is dynamic and adaptive provisioning of network resources both in control and data plane. In the data plane, the EON paradigm [2] has explored a completely unique optical network architecture able to provision LPs based on the actual traffic demands. This flexibility makes the LP provisioning problems much more challenging as compared to traditional fixed-grid wave-length division multiplexing (WDM) networks. Apart from this in the control plane, SDN provides on-demand configuration and virtualization of a network. Furthermore, the dynamic configuration of the optical network offers a profusion of design parameters such as single/multi carrier transmission, baud rate, adaptive modulation format, forward error correction (FEC) and adaptive channels spacing provides good degree of versatility to the network operator during the planning phases. In this reference, the preliminary QoT-E of the LP is significantly important for the planning phases of optical networks. Generally, a considerable number of analytical models are available for QoT-E, but all of them require the *exact* description of system parameters. Thus, the analytical models are nonfunctional in the case of *un-used* network scenarios, i.e., the networks where the operator doesn't have an exact knowledge of the working point of network elements (gain and noise figure ripples in amplifiers, insertion losses, etc...)

In this context, an alternative approach has been recently investigated based on the ML techniques which has the potential to estimate the QoT of the LP of *un-used* network using the characteristics features extracted from telemetry data of already *in-service* network. For a state of the art WDM optical transport based on coherent optical technologies, QoT is well assessed by the generalized SNR (GSNR), which includes the effect of ASE noise and NLI accumulation [3]. Generally, the QoT parameter is the the most challenging parameter during planning phase owing to the uncertainties on the knowledge of the exact working point of network elements imply uncertainty in QoT-E. To counteract this, we use three data-driven ML methods for an accurate low margin QoT-E of LP prior to its actual deployment. A ML paradigm has already been expertly used in the optical networks; consider [4], [5], [6], [7] for network performance monitoring, [8], [9] for QoT prediction using ML approach. In [10] the authors used ML technique for controlling optical line system (OLS). An overall survey of ML applied applications in optical networks are discussed in [11].

The major difference between the previous literature and the present work is that we exploit a more realistic approach by using telemetry data of GSNRs responses to specific traffic configurations of LPs of the already *in-service* network in an open environment. The telemetry data are generated *synthetically* by perturbing the nominal working point of network elements of *in-service* network. Exploiting this telemetry data of already well developed used network, we cross-train the proposed ML techniques to estimate the QoT of an *un-used* network. The cross-trained ML techniques empowers the network operator to obtain a reliable and accurate QoT-E that can be used for planning, control and reliably deploy the LP with a minimum margin.

2. System Model & Data Generation

For mimicking the telemetry data of *in-service* network, a reliable and well-tested open source GNPy library is used for physical layer abstraction [12]. This library outlines an end to end simulation environment which develop the network models for physical layer, using this capability we configure GNPy to mimic the telemetry data. A typical SDN empowered backbone optical network is considered in which edges are modeled by OLSs that comprise of fibers and amplifiers while nodes are defined as re-configurable optical add-drop multiplexer (ROADM) sites. The given OLSs are supposed to operate at the nonlinear-propagation optimal working point and the random behavior of physical layer is considered through amplifier gain ripple. The merit of QoT i.e., the $GSNR$ of any candidate LP routed through a particular OLSs is given by $1/GSNR = \sum_n 1/GSNR_n$, where n is the number of OLSs contributing in the routing of particular LP. The $GSNR$ metric of candidate LP is given by Equation. 1, where P_{Rx} is the power of the channel at the receiver, P_{ASE} is the power of the ASE noise and P_{NLI} is the power of the NLI.

$$\mathrm{GSNR} = \frac{P_{\mathrm{Rx}}}{P_{\mathrm{ASE}} + P_{\mathrm{NLI}}}\,, \qquad (1) \qquad\qquad \mathrm{MSE} = \frac{\sum_{i=0}^{n} (\mathrm{GSNR}_i^{\mathrm{p}} - \mathrm{GSNR}_i^{\mathrm{a}})^2}{n}\,, \qquad (2)$$

In addition to this, fiber impairments such as fiber attenuation (α = 0.3 dB/km), dispersion (D = 16 ps/nm/km) and insertion losses are also considered. In order to make simulation more realistic, the statistics of insertion losses are determined by an exponential distribution with $\lambda = 4$ as described in study [13]. The considered OLSs carry only 76 channels over the standard 50 GHz grid on the C-band, having total bandwidth close to 4THz due to the limitation of computational resources. We do not expect substantial difference in results when considering standard 96 channels on the entire C-band. We supposed to rely on transceivers at 32 GBaud, shaped with a root-raised-cosine filter. The amplifiers in the OLSs are configured to work at a constant output power mode of 0 dBm per channel. All channels are operated at their optimum launch power ranges (0-0.006) dBm and all network links are supposed to operate on standard single-mode fiber with average span length of 80 km. In-line amplifiers are supposed with noise figure uniformly distributed for each amplifier in a range of (6 to 11) dB along with uniform random gain ripple with 1 dB variation.

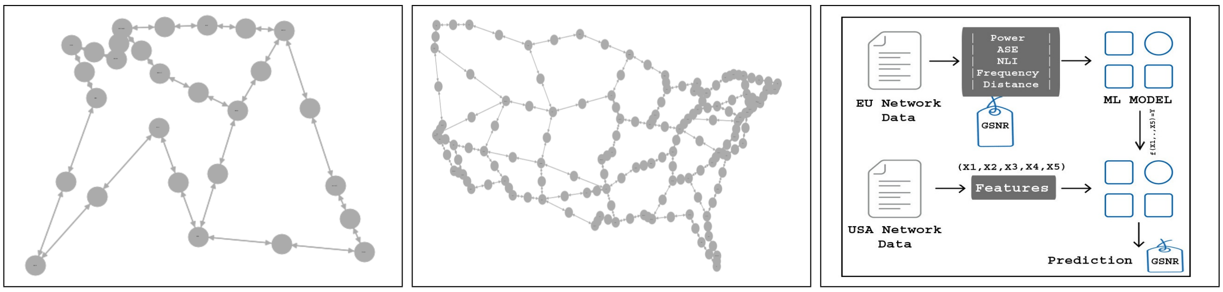

Figure 1: (a) European Network Topology; (b) USA Network Topology; (c) Machine Learning Black-Box

After finalizing the network configuration, the most delicate part is the spectral load of simulated links which in the proposed work is the subset of 2^{76}, overall possible combinations of spectral load given by 76 channels. Among this subset each source-to-destination (s –> d) pair has 1024 realizations of random traffic ranging from 34% to 100% of total bandwidth utilization. The data set is mimicked against the European Union (EU) Network Topology shown in Fig. 1a, which is used as already *in-service* network and for an *un-used* network case the required data set is generated against the USA Network Topology shown in Fig. 1b. The configuration used for an *un-used* network scenario is the same as for the former *in-service* network case with the exception of different amplifier gain ripple and noise figure, which is normally the most random part in the OLS. In order to mimic this behaviour, GNPy is configured to generate 4096 data realizations for 4 (s –> d) pairs of *in-service* EU Network and 35840 data realizations against 35 (s –> d) pairs of an *un-used* USA Network.

3. Visualize Machine Learning Black-Box

The proposed models cognate the features and labels of the candidate LP of the *in-service* network using ML. The manipulated parameters used to define the features for ML models include received signal power, NLI, ASE, channel frequency and distance between source to destination node, while the exploit label is manipulated by $GSNR$ parameter of the candidate LP depicted in Fig .1c. The total number of input features for proposed ML models consists of *380* entries, as we have 76 entries against each manipulated parameters *(76 X 5 = 380)*. In this study, the three ML based models are build-up using high level python application program interface (API) of open source ML library called *scikit-learn (SKL)*. *SKL* provides variety of algorithms for ML tasks ranging from classification, regression, dimensionality reduction, and clustering. Moreover, it has also some modules for features extraction, pre-processing data and evaluation of the models [14]. The proposed *SKL* based ML models are evaluated by *mean square error (MSE)* as a loss function expressed in Equation. 2, where $\mathrm{GSNR}_i^{\mathrm{a}}$ and $\mathrm{GSNR}_i^{\mathrm{p}}$ are the actual and predicted values of the any candidate channel for the ith spectral load respectively and n is the total number of realizations in the test data set.

4. Machine Learning Models

In the present work, the three ML based models are developed using python API of *SKL* library. The models are listed as decision tree regressor (DTR), random forest regressor (RFR) and multi-layer perceptron regressor (MLPR) [14]. The models are configured for training, validation and testing by the conventional rule 70/15/15 having training-steps of 1000, in order to give the models sufficient intelligence. Each proposed model is cross-trained by 4 different paths of EU-Network; *Amsterdam-Berlin, Brussels-Bucharest, Frankfurt-Istanbul, Vienna-Warsaw* having 4096 realizations and is tested explicitly on 35 different paths of *un-used* USA-Network having 35840 realization depicted in Fig. 3.

4.1 Decision Tree Regressor

The proposed DTR constructs a tree based on several decisions inferred from the data features. The present DTR consists of three basic modules; pre-processing, training and testing. Before passing the data to train module of the DTR, the data set is standardized using pre-processing module [15]. After standardizing the data set, it is moved to the training module where DTR is cross-trained on 4 different paths of *in-service* EU-Network. The training module has two main basic tuning parameters; *min_samples_leaf* and *max_depth*. In order to get the optimum values of these parameters in the present scenario, considering the computational time and model over fitting, *min_samples_leaf* = 3 and *max_depth* = 100 are selected. Soon after cross-training, the test module start testing 35 different paths of *un-used* USA-Network.

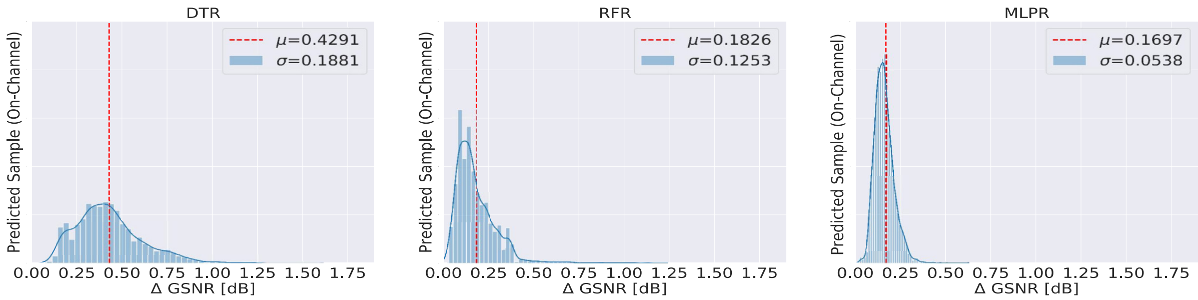

Figure 2: Distribution of $\Delta GSNR$ for *SKL* Based Learning Methods

4.2 Random Forest Regressor

The RFR model combines several decision trees, cross-trained separately on randomly selected subsets of the training data set using Bagging technique [15]. The proposed RFR also consists of same three basic modules exactly like DTR; pre-processing, training and testing. The standardized data set after pre-processing is moved to the training module, where RFR is trained on 4 different paths of *in-service* EU-Network. Similar to DTR parameter tuning, we also do the same for RTR and decide *min_samples_leaf* = 3 and *max_depth* = 100 in the particular simulation scenario. After cross-training, the RFR test module start testing 35 different paths of *un-used* USA-Network.

4.3 Multi-layer Perceptron Regressor

The MLPR has generally two or more layers of perceptrons which form a directed acyclic graph, where each layer of MLPR is fully connected to the subsequent layer. The proposed MLPR also consists of three basic modules; pre-processing, training and testing same. The standardized data after pre-processing is moved to the training module where MLPR is trained on 4 different paths of *in-service* EU-Network. The train module of MLPR is configured with back propagation algorithm along with default *stochastic gradient descent (SGD)* optimizer having *learning rate* = 0.01 and L_2 regularization = 0.001 [16]. In addition to this, MLPR consists of 3 hidden layers along with 20 neurons for each hidden layer, having *relu* as activation function that allows translation of the given input features into the prediction of label of our point of interest with less complexity [17]. After cross-training, the test module start testing 35 different paths of *un-used* USA-Network.

5. Results and Discussion

In this section, we exploit the prediction performance of the proposed three ML models; DTR, RFR and MLPR. The assessment of the proposed ML models are done by calculating the prediction error $\Delta GSNR$, where $\Delta GSNR$ = $GSNR_{Predicted}$ - $GSNR_{Actual}$. In Fig. 2, a prediction error metric is plotted for DTR, RFR, and MLPR against the test samples of on channels realization only. For the given simulation scenario, DTR is unable to find underlying relationship and irregularities. On the other hand, RFR took average of various trees

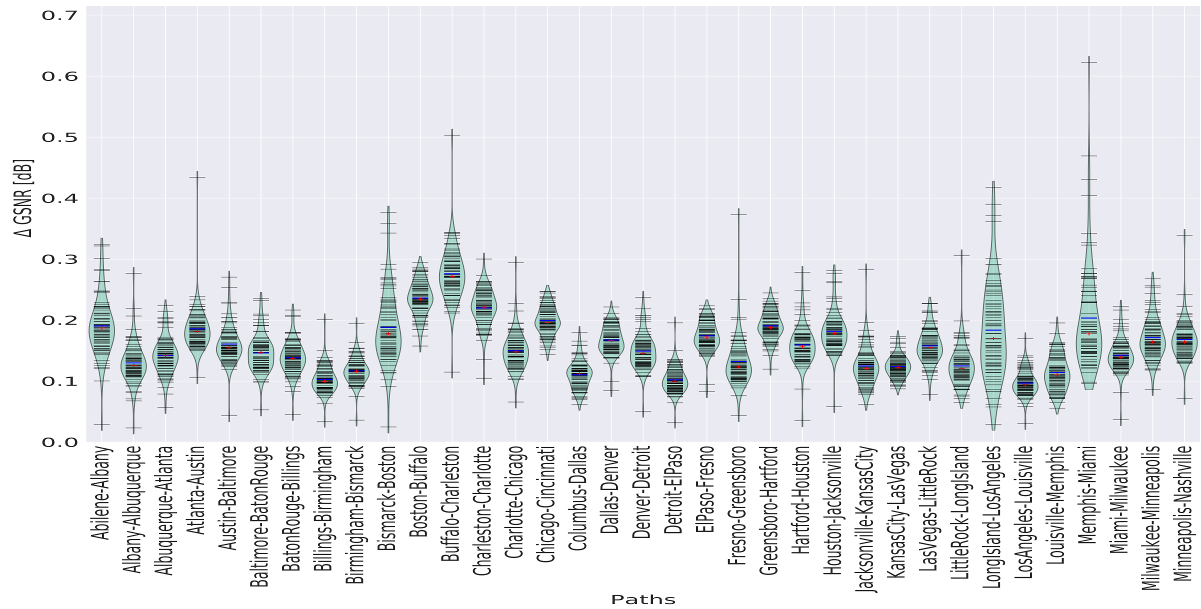

Figure 3: Distribution of $\Delta GSNR$ for the simulated links of USA Network using *MLPR*

instead of choosing randomly subsets of the features. So, its overall performance is much better than DTR. On the other hand MLPR performed very well due to its cognitional potentiality provided by internally configured neurons as compared to DTR and RFR. The results are verified by observing the mean (μ) and standard deviation (σ) of $\Delta GSNR$ against each proposed model in Fig. 2. Focusing on the MLPR after observing μ and σ of $\Delta GSNR$, it is quite obvious that MLPR is the best learning methods in the present simulation scenario. We further analyze and elaborate the results related to MLPR by showing the beans-plot in Fig. 3, the $\Delta GSNR$ distribution of all the test paths of *un-used* USA Network. In Fig. 3, the out bound values of $\Delta GSNR$ is also depicted along with the μ value (*red dot in each bean*) of $\Delta GSNR$ for each test path.

6. Conclusion

In summary, we proposed and exploit the ability of three different ML models for QoT-E, considering the scenario of cross-train ML models on telemetry data of *in-service* EU Network and tested on completely *un-used* USA Network. The proposed ML models are developed by using higher level APIs of *scikit-learn* library. The data set used in this particular scenario is generated *synthetically* using a reliable and well-tested GNPy library. Exploiting the ability of cross-trained MLPR, developed by using *scikit-learn* performs better than RFR and DTR due to its cognitive ability of perceptron. It is remarkable that, MLPR proved to be the model achieving the best generalization in terms of average prediction error of 0.1697 dB with standard deviation 0.0538 dB .

References

[1] Cisco, "Cisco Visual Networking Index: Forecast and Trends, 2017–2022," Tech. rep., Cisco (2017).
[2] G. Ori et.al., "Eon: A new dawn for the optical layer?" IEEE CM **50**, s12–s20 (2012).
[3] M. Filer et.al., "Multi-Vendor Experimental Validation of an Open Source..." JLT **36**, 3073–3082 (2018).
[4] M. Freire et.al., "Predicting optical power excursions in erbium..." in *2018 ACP,* (IEEE, 2018), pp. 1–3.
[5] J. Thrane et.al., "Machine learning techniques for optical performance..." JLT **35**, 868–875 (2017).
[6] F. Khan et.al., "Optical performance monitoring in fiber...." in *2018 OFC,* (IEEE, 2018), pp. 1–3.
[7] B. Luca et.al., "Qot estimation for unestablished lighpaths..." in *OFC,* (OSA, 2017), pp. Th1J–1.
[8] S. Ippokratis et.al., "Accurate quality of transmission estimation..." JOCN **11**, 140–150 (2019).
[9] M. Weiyang et.al., "Ann-based transfer learning for qot..." in *2018 OFC,* (IEEE, 2018), pp. 1–3.
[10] A. D'Amico et.al., "Using machine learning in an open optical line system..." JOCN **12**, C1–C11 (2020).
[11] M. Javier et.al., "(ai) methods in optical networks: A comprehensive survey," OSN **28**, 43–57 (2018).
[12] A. Ferrari et.al., "Gnpy: an open source application for physical layer aware..." JOCN **12**, C31–C40 (2020).
[13] Y. Ando, "Statistical analysis of insertion-loss improvement.." IEEE PTL **3**, 939–941 (1991).
[14] G. Hackeling, *Mastering Machine Learning with scikit-learn* (Packt Publishing Ltd, 2017).
[15] "Supervised learning," https://scikit-learn.org/stable/supervised_learning.html.
[16] S. Ruder, "An overview of gradient descent optimization algorithms," arXiv:1609.04747 (2016).
[17] A. Kızrak, "Comparison of activation functions for deep neural networks," (2020).