How Uncertainty on the Fiber Span Lengths Influences QoT Estimation Using Machine Learning in WDM Networks

J. Pesic(1), M. Lonardi(1), N. Rossi(2), T. Zami(2), E. Seve(1) and Y. Pointurier(1)
(1) Nokia Bell Labs France, (2) Nokia France
jelena.pesic@nokia-bell-labs.com

Abstract: We investigate how a machine learning-based QoT estimator performs depending on different features selections, on homogeneity of the learned light paths and on uncertainty of their span lengths using artificial database for the France43 network. © 2020 Nokia Bell Labs

1. Introduction

Although still in its infancy, machine learning (ML) for optical networking has attracted a lot of attention in the scientific community. Different types of ML have been successfully reported when applied to environments without underlying models [1]. Today, the quality of transmission (QoT) estimation for establishing light paths (LPs) in WDM networks is done using a QoT tool usually based on the well-studied analytical Gaussian noise (GN) model [2]. When ML was introduced for QoT estimation, two main research directions appeared: the first one to increase the accuracy of the input parameters of the QoT tool, aimed at reducing margins to optimize network capacity and cut capital expenditure [3]. The second direction was for completely replacing the existing QoT tool [4-9]. The relevance of this replacement is still in question and this paper underlines the potential challenges, pitfalls and roadmap of using ML in the future, instead of the existing QoT tool, particularly when the ML training and testing data sets exhibit some uncertainties stemming from inaccurate knowledge of the network fiber span lengths.

2. Data Set

This work tackles the challenge of using ML for the estimation of the generalized optical signal-to-noise ratio (G-OSNR), including fiber nonlinear and filtering OSNR penalties considering 32 GBaud channel modulation formats, relying upon an artificial data base. We generate a data set, based on the 5 shortest light paths connecting each pair of nodes in the France43 backbone network topology, shown in Fig. 1(a). Span lengths range between 30 km and 80 km, with a mean of 62.5 km. For a total of 9030 unidirectional LPs, we calculate the receiver G-OSNR with the GN model for the following 9 polarization multiplexing channel modulation formats including 4 time-domain hybrid (TDH) modulations: 100 Gb/s QPSK, 125 Gb/s TDH QPSK/8QAM, 150 Gb/s 8QAM, 175 Gb/s TDH 8QAM/16QAM, 200 Gb/s 16QAM, 225 Gb/s TDH 16QAM/32QAM, 250 Gb/s 32QAM, 275 Gb/s TDH 32QAM/64QAM and 300 Gb/s 64QAM. The reference model and the considered physical parameters for filtering penalties can be found in [10]. The data set consists of the list of the 9030 light paths (as input) and their respective G-OSNR [dB] for the highest supported modulation format (as output), when assuming a fully loaded 50 GHz-spaced WDM comb of 80 channels. A feed-forward shallow neural network (NN) with 20 fully connected sigmoid-activated neurons for the hidden layer and an identity-activated neuron at the output layer is trained with the Levenberg-Marquardt backpropagation algorithm with 70%/15%/15% division for training/validation/testing.

We test 3 different features (input parameters) selections at the NN input. In the first one, each LP is represented by the total number of spans it consists of. In the second one, each LP is represented by the total number of spans and the total length of the LP, while in the third one each LP is represented by the list of span lengths. In this study, we consider homogeneous amplifier noise figure (NF = 4.5 dB) and fiber attenuation (0.22 dB/km) along the transmission. Each transparent node traversal is represented in a data set by a span of 80 km.

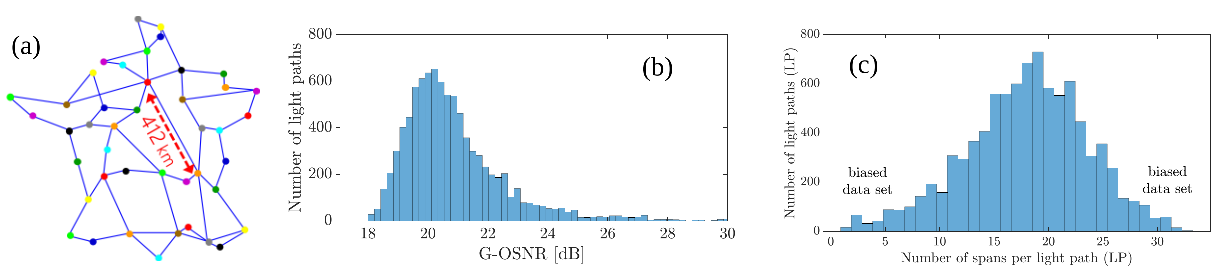


Fig. 1. For France43 network (a) topology, (b) G-OSNR distribution, (c) Distribution of number of spans per light path, for 5 shortest light paths connecting each pair of nodes

3. Nominal Errors on G-OSNR Estimate

The results of the ML-based QoT tool are assessed in the form of its absolute estimation error on the G-OSNR, compared to the one given by the analytical GN model (ground truth). Fig. 1(b) illustrates the breakdown of these ground truth G-OSNR over the 9030 5-shortest LPs. Mean absolute error (MAE), maximum absolute error (MaxAE), and root mean square error (RMSE) are used as statistical metrics to measure NN model performance. Over 1000 runs, training and testing are done on a new NN realization from which we obtain in total 1355000 testing estimations (1000 runs x 1355 (15% of 9030 LPs)). The relative importance of these errors can be gauged in comparison with the G-OSNR distribution depicted in Fig. 1(b). As expected, Table 1 shows that the most insightful set of input parameters, which is “list of span lengths”, entails the smallest errors in terms of RMSE and MAE. It also shows very high MaxAE whatever the feature selection. These poor predictions occur when the tested LPs consist of a relatively small or large number of spans. We call these cases “outliers” when the absolute error exceeds 1 dB. They are seldom as depicted in Fig. 1(c) and thus they cannot be accurately learned via NN-based ML. We define the probability of outliers as the number of outliers divided by the number of total testing estimations. In complement to Table 1, Fig. 2 illustrates the distribution of G-OSNR absolute error when focusing on the 0 to 1 dB range. It also clearly quantifies the better accuracy of NN estimate with respect to the list of span lengths since the related distribution is the least spread out. Despite the better nominal performance of this “list of span lengths” feature selection and the smaller outliers’ probability over the whole set of 1355000 testing estimations, Table 1 indicates that it also induces the worst MaxAE. Its more significant number of weights to be managed in NN layers makes it more sensitive to outliers as compared to the 2 other feature selections. Although training with uniform data set is a well-known condition for a efficient ML work, in this study we use on purpose the non uniform one, in terms of number of spans (see Fig. 1(c)), to quantify the pitfall of unsufficiently uniform data set. To overcome this hurdle when setting up data set for ML, in addition to considering more LPs to better characterize the network, more attention should be paid to build a uniform distribution of LPs in terms of number of spans. If this is not possible, we should resort to artificial data generation guaranteeing a fair LP length distribution.

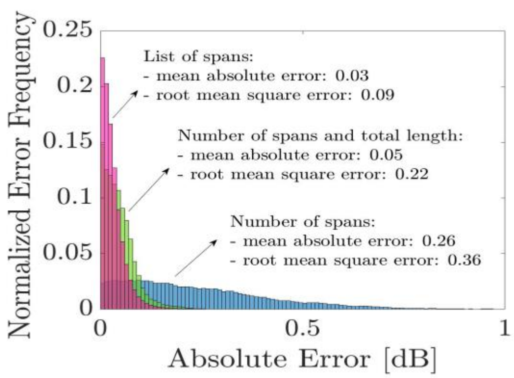


Fig. 2. Absolute error in G-OSNR estimation for 3 feature selections

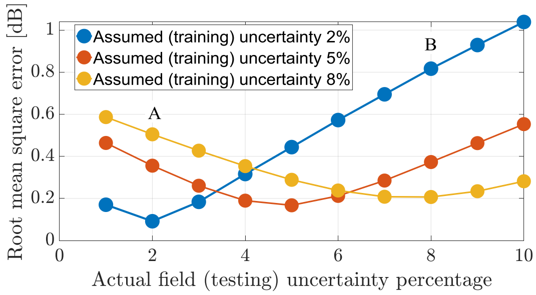
Table 1. Evaluation of the G-OSNR estimation

Feature selections	RMSE [dB]	MAE [dB]	Probability of outliers	MaxAE [dB]
Number of spans	0.36	0.26	0.01	7.5
Number of spans & Total length	0.22	0.05	9.2x10^{-4}	8.5
List of span lengths	0.09	0.03	3.9 x10^{-4}	39

4. Adding Uncertainty on the Span Lengths

The authors in [11] showed that 29% of the fibers in the North American backbone network exhibit more than 5 km length discrepancy between the values assumed for the network design and the actual values in the field. Consequently, we assess the penalties when using ML for the QoT estimate of the non-established LPs, in case of mismatch between the values in the design tool (in our case, training data set) and the values in the field (testing data set). As we consider GN model as a ground truth, we are using the span length values from the design tool with added uncertainty to build a training and testing data sets. Uncertainties are modeled either with a Gaussian or a Uniform process, both additive and zero-mean. It is noteworthy that the uncertainties applied on the spans of each LP from the set of 5-shortest LPs were uncorrelated, even for the LPs partly going through the same spans. This way, we avoid the final resulting distribution of uncertainties being affected by the occurrence rate of each individual span in this set. Fig. 3(a) shows the RMSE on the G-OSNR estimation when training NN networks with assumed uniform uncertainty of 2%, 5% and 8% and testing with the actual field uniform uncertainty up to 10%. Minimum errors on the estimations are when assumed and actual uncertainties are the same and these errors grow along with the difference between those two values. Comparing points A (training with 8%, testing with 2%) and B (training with 2%, testing with 8%) on Fig. 3(a), it appears that despite the same absolute difference of 6% in uncertainty between the training and testing phases, their respective RMSEs notably differ. This means that in order to minimize the final overall resulting G-OSNR uncertainty, in a context where the feature selection of a given

(a)



(b)

		Assumed uncertainty	
		Gaussian distribution	**Uniform** distribution
Actual uncertainty	**Gaussian** distribution	RMSE = 0.16 dB MAE = 0.13 dB	RMSE = 0.68 dB MAE = 0.65 dB
	Uniform distribution	RMSE = 0.65 dB MAE = 0.62 dB	RMSE = 0.16 dB MAE = 0.11 dB

Fig. 3 (a) RMSE on the G-OSNR estimation when adding uniform uncertainty on the span lengths (b) RMSE and MAE for 5% assumed and 5 % actual uncertainty on the span lengths

network is only known with poor precision, whenever possible one should prefer to train the corresponding NN-based QoT estimator by overestimating the uncertainties versus underestimating them. To evaluate the sensitivity of NN-based QoT estimate to different uncertainty distributions, we examine it while training (assumed uncertainties) and testing (actual uncertainties) using the same or distinct distributions. Hence, Fig. 3(b) reports RMSE and MAE obtained when combining them. Unsurprisingly, the errors are notably smaller when the assumed and actual uncertainty distributions are identical, because NN training learns not only the nominal G-OSNR values but also the corresponding assumed distribution of uncertainties. One of the associated challenges when designing a network is to know this distribution. Therefore, whenever it is impossible to assess the distribution one should report the performance also of the worst (off-diagonal in Fig. 3(b)) case, which impacts the margins the most.

5. Conclusion

We have shown the importance of the feature selection when building training and testing data sets. Even without any added uncertainty on the fiber span lengths, we have shown that looking at single performance metric can be detrimental. As every statistical measure condenses a large amount of data into a single value, it only provides one projection of the model errors emphasizing a certain aspect of the error characteristics of the model performance. For instance, the solely root mean square error does not grasp the harmful impact of biased training data set on some feature selections. We have also highlighted the importance of the uniformity of the training set in terms of the selected features, which plays an essential role in the accuracy of the G-OSNR estimation. If inaccessible with data from the field, uniformity should be achieved by adding artificial data possibly enhanced by actual telemetry.
For the first time we evaluated the impact of the uncertainty on input parameters. We pointed out that both the uncertainty statistical distribution (process) and strength (percentage), must be known in advance to achieve the best estimation performance. Such prior knowledge is one of the most important challenges when designing a network with ML-based QoT tool. If this task is not possible, we presented some rules of thumb for taking into account the worst case, considering different uncertainty distributions for the training and testing sets.

6. References

[1] T. O'Shea and J. Hoydis, "An Introduction to Deep Learning for the Physical Layer," in *IEEE Transactions on Cognitive Communications and Networking*, vol. 3, no. 4, pp. 563-575, Dec. 2017.
[2] P. Poggiolini, "The GN Model of Non-Linear Propagation in Uncompensated Coherent Optical Systems," J. Lightwave Technol. 30, (2012)
[3] E. Seve, J. Pesic, C. Delezoide, and Y. Pointurier, "Learning process for reducing uncertainties on network parameters and design margins," in Optical Fiber Communication Conference, OSA Technical Digest (online) (Optical Society of America, 2017), paper W4F.6
[4] L. Barletta, A. Giusti, C. Rottondi, and M. Tornatore, "QoT Estimation for Unestablished Lighpaths using Machine Learning," in Optical Fiber Communication Conference, OSA Technical Digest (online) (Optical Society of America, 2017), paper Th1J.1.
[5] R. M. Morais and J. Pedro, "Machine Learning Models for Estimating Quality of Transmission in DWDM Networks," J. Opt. Commun. Netw. 10, D84-D99 (2018)
[6] J. Slovak, M. Herrmann, W. Schairer, E. Torrengo, K. Pulverer, A. Napoli, and U. Häbel, "Aware Optical Networks: Leaving the Lab," J. Opt. Commun. Netw. 11, A134-A143 (2019)
[7] I. Sartzetakis, K. K. Christodoulopoulos and E. M. Varvarigos, "Accurate quality of transmission estimation with machine learning," in IEEE/OSA Journal of Optical Communications and Networking, vol. 11, no. 3, pp. 140-150, March 2019. doi: 10.1364/JOCN.11.000140
[8] D. Rafique and L. Velasco, "Machine Learning for Network Automation: Overview, Architecture, and Applications [Invited Tutorial]," J. Opt. Commun. Netw. 10, D126-D143 (2018)
[9] J. Yu, W. Mo, Y. Huang, E. Ip and D. C. Kilper, "Model transfer of QoT prediction in optical networks based on artificial neural networks," in *IEEE/OSA Journal of Optical Communications and Networking*, vol. 11, no. 10, pp. C48-C57, October 2019.
[10] J. Pesic, N. Rossi and T. Zami, "Fiber Type Dependent Benefits When Deploying Up-to-300Gb/s Elastic Transponders Adapting to Ageing of Margins," *2018 Optical Fiber Communications Conference and Exposition (OFC)*, San Diego, CA, 2018, pp. 1-3.
[11] M. S. Zefreh and S. Asselin, "Capacity Enhancement in Optical Networks using Margin Extraction," in *Optical Fiber Communication Conference*, OSA Technical Digest (online) (Optical Society of America, 2018), paper W4A.3.