MACHINE-LEARNING AIDED OSNR PREDICTION IN OPTICAL LINE SYSTEMS

Andrea D'Amico[(1)], Stefano Straullu[(2)], Antonino Nespola[(2)], Ithesham Khan[(1)], Sherif Abdelfattah[(1)], Emanuele Virgillito[(1)], Stefano Piciaccia[(3)], Aberto Tanzi[(3)], Gabriele Galimberti[(3)], Silvio Abrate[(2)], and Vittorio Curri[(1)]

[1]*Politecnico di Torino, Corso Duca degli Abruzzi, 24, 10129, Torino, Italy*
[2]*LINKS Foundation, Via Pier Carlo Boggio 61, 10138 Torino, Italy*
[3]*Cisco Photonics, Via Santa Maria Molgora 48C, 20871 Vimercate, Italy*
**E-mail: andrea.damico@polito.it*

Keywords: OPTICAL NETWORK AUTOMATION, QOT-E, EDFA, MACHINE-LEARNING,

Abstract

We suppose an OLS agnostic operation by relying only on the OCM telemetry. Analyzing an experimental data-set from an 11 EDFA OLS, we show the need for 3.6 dB margin without knowledge on EDFAs, reduced to 0.4 dB by DNN machine learning.

1 Introduction

Data traffic demand will experience a dramatic increase over the next years, driven by the implementation of 5G access and by the expansion of bandwidth hungry applications such as high definition video, virtual and augmented-reality contents [1]. Backbone optical networks will sustain such a growth, and a key operators' request is to exploit the existing infrastructure to maximize their returns over investments [2]. Such a need is directly related to the capability of orchestrating all the network layers to *squeeze out* all the available capacity from the data transport [3]. In optical networks, the enabler for the optimal exploitation of data transport is the capability of controllers for the optical line systems (OLS) to predict the lighpath (LP) quality of transmission (QoT) before its actual deployment. The QoT is given by generalized SNR (GSNR) considering both ASE noise and NLI [4, 5]. $GSNR = P_{Rx}/(P_{ASE} + P_{NLI}) = 1/(1/OSNR + 1/SNR_{NL})$, where $OSNR = P_{Rx}/P_{ASE}$, $SNR_{NL} = P_{Rx}/P_{NLI}$, P_{Rx} is the power of the channel under test (CUT), P_{ASE} is the power of the ASE noise and P_{NLI} is the intensity of the NLI disturbance. Among the two disturbances, the most challenging to estimate is the ASE noise accumulation setting the optical signal-to-noise ratio (OSNR), because it depends on the working point of erbium-doped fiber amplifiers (EDFA)[6], that is spectral-load dependent [7]. The NLI can be accurately predicted when the ASE noise accumulation is well characterized [8–12].

Fig. 1 schematically shows the possible operations of the QoT estimator (QoT-E) calculating the GSNR, wavelength per wavelength. Some data are available from static characterization of devices – the (1) option in Fig. 1 – and are highly significant for single-vendor systems. An agnostic operation required in open OLS mainly relies on data from telemetry, coming from the optical channel monitor (OCM) and from EDFA telemetry. We can suppose that these data are available only from the network status, without any stored training data-set: the case (2) in Fig. 1. From the statistical characterization of the data-set, we can evaluate the uncertainty on the OSNR and the consequent needed margin. Option (3) in Fig. 1 is the availability of a data-set that has been generated before the in-service operation of the OLS. Data can train a machine-learning (ML) algorithm that will be used to predict the QoT in actual LP deployment and in controlling the OLS power levels.

In this work, we analyze the (2) and (3) options, being the ones characterizing an agnostic operation of the OLS. We suppose a completely agnostic scenario, by relying only on data coming from the OCM available at the end of the line system.

To emulate an OLS, we experimentally setup an 11-EDFA line and we changed spectral loading, collecting data from an OSA mimicking the OCM, as shown in Fig. 2. The Cisco© commercial EDFAs are used as black-boxes just setting the spectrally-averaged gain to the nominal level. The channel combs spectrally loading the OLS have been obtained by shaping ASE noise. Such an approach is not limiting the generality of results because of the large time constant characterizing physical effects in EDFAs. The output of the ASE noise source is shaped by means of a programmable optical wave-shaper filter (Finisar 1000 S), to generate a 100 GHz-spaced 35-channel WDM comb centered at 1550 nm and amplified by a booster amplifier ($EDFA_0$ in Fig. 2). The choice of the 100 GHz spacing was forced by the hardware availability as well as the limitation to 3.5 THz. These restrictions are not limiting the generality of results because of the bandwidth of the phenomena that are properly captured. The optical line is composed by 11 spans, each made of a VOA, setting the optical span attenuation to 10 dB, followed by an EDFA, operating at a constant output power of -1 dBm per channel. We generated experimentally a data-set made of 2291 cases: almost 70 for each spectral load with 2 to 34 on channels; 35 cases with only one channel *on* and the all channel *on* scenario.

We first show that a completely agnostic use of the OLS may require the deployment of up to 3.6 dB of system margin. Then, using the TensorFlow© platform [13], we show that using deep neural network (DNN) algorithms, with some optimization,

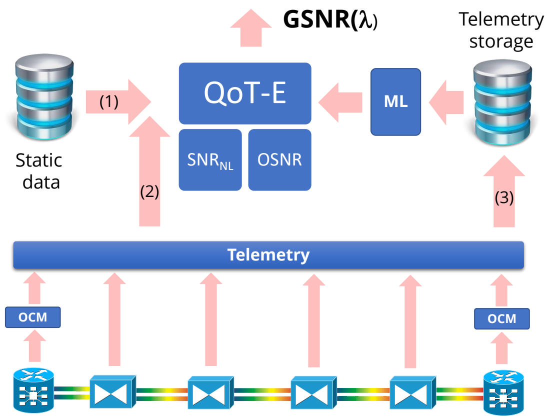

Fig. 1 QoT-E module predicting the GSNR(λ) using static data from factory characterizations of network elements (1), data form telemetry (2) or stored data from telemetry feeding a ML module (3).

and properly exploiting the data-set, we can reduce the uncertainty on the predicted OSNR down to 0.2 dB, minimizing the needed margin to 0.4 dB.

2 Results

We analyze the experimental data-set in order to give an assessment on its statistics and consequently deriving the needed system margin. Then, we suppose the data-set has been acquired prior the OLS activation, and show the ML benefits.

2.1 Data-set statistics

The main purpose of this analysis is to characterize the OSNR fluctuations of this process due to the specific spectral load features. The data-set consists of pairs *combination-outpower* for each channel identified by a frequency occupation slice with 100 GHz width. These quantities have been measured over 2291 *on-off* configurations, a subset of the 2^{35} possibilities given 35 channels. The set of configurations considered is uniform over all the channels and the number of total on channels. Through the whole analysis, any uncertainty in the measurements is negligible with respect to the variances due to the lack of information, therefore, we have not take them into. Since

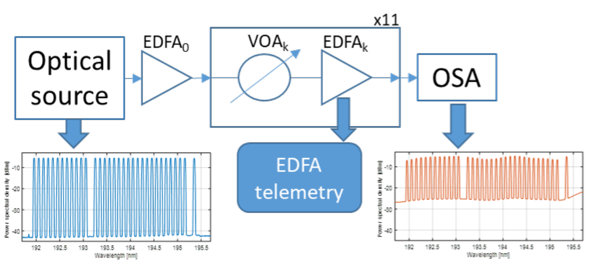

Fig. 2. Experimental setup generating the data-set.

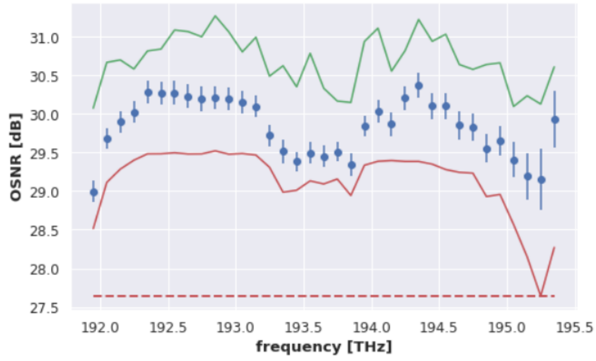

Fig. 3 Blue dots: mean value of the OSNR measured on each channel; errorbars: standard deviations. In green and red, respectively, the maximum and the minimum for each channel . The dashed line indicates the overall minimum OSNR.

the sample has a relative limited dimension, in what follows, we address to the standard deviation calculated with the weight $N - 1$, in order to obtain a less biased estimator of any quantity variance. By means of this first approach, we obtain the statistical description shown in Fig. 3. The average values of the OSNR, dots in the figure, sketch a characteristic figure of the EDFAs amplification process and take place between 29 and 30.5 dB with standard deviations from 0.13 to 0.4 dB. Omitting the description of the shape outlined by the OSNRs, in this work, we would like to arrange this behavior in a practical case. Not far from a realistic scenario, let us suppose that the line user experiences a condition of total lack of information of the EDFAs cascade features. Unaware of the frequency resolution of the amplification process, he/she has to set a constant nominal value of the OSNR over all the spectrum, which would be indeed beneath the minimum value measured, the dashed red line in the Fig. 3. As shown in the same figure, this solution would be suboptimal with a required margin up to 3.6 dB.

2.2 Machine-learning aided OSNR prediction

We now describe the improvements which we obtain applying the ML to this system. In our opinion, this scenario is a perfect ground for a ML approach, since it can compensate the lack of information due to the hidden behaviour of the EDFAs cascade. Far from being an exhaustive description of the ML application, the goal of this work is to achieve a better prediction power with respect to the approaches exposed in the previous section, starting from the same data-set. In particular, we take advantage of the well known Tensorflow© platform [13] and we tune, along our requirements, few high level features. As first evidence, we obtain that the deep neural network (DNN) reaches a better precision with respect to the linear regression model. Chosen the former model, we have tried many configurations of hidden layers and number of nodes and, once we reached a suitable reliability, we have opted for the best trade-off between precision and computational time.

Given a channel under test, we select a portion of the data-set, almost one fourth of it, as test subset. We create it choosing randomly various instances from the data-set, with the only requirement to preserve the uniform distribution with respect to the number of on channels in the configurations; this subset, from now on, represents our reference point to questioning the quality of the predictions. As outlined in the previous section, the uncertainty of the system can be split along the variances of the received signal power and the ASE noise. Therefore, we proceed applying the ML training procedure separately. For the prediction of the ASE noise we identify the train set with the whole residual sample. Whereas, for the prediction of the output power when the channel is *on*, we restrict the residual sample to the configurations with the channel under test *on*, in order to feed the machine with the proper information. This process, since the uniform distribution of the sample, reduces significantly the train set in this case. This dimension discrepancy of the train sets, can partially motivate the different accuracy of the final results.

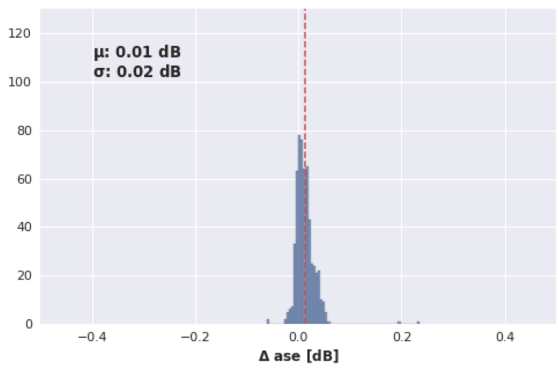

Fig. 4 Distribution of the discrepancy of the ML guesses of the ASE noise for the channel centred in the spectral hole burning.

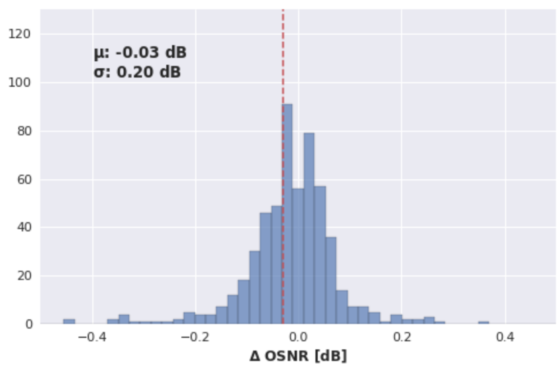

Fig. 5 Distribution of the discrepancy of the ML guesses of the OSNR for the channel centred in the spectral hole burning.

First, we focus our investigation on the channel with the most uncertainty. Unfortunately, the collected data-set does not

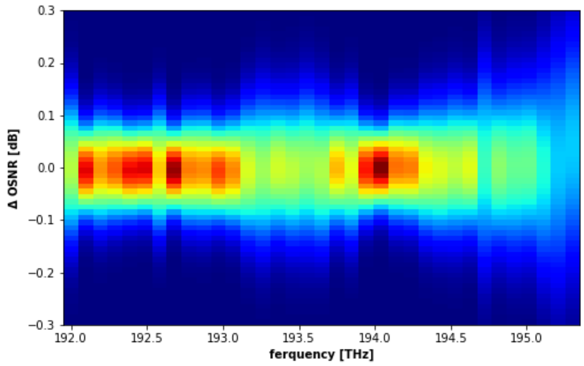

Fig. 6 Each slice of the heat map represents the distribution of the error of the prediction process. Hotter colors stand for higher accuracy.

allow us to train and test the machine directly over those scenarios in which only the channel under test (CUT) changes state. Nevertheless, at least for those configurations with on channels not in the frequency neighborhood of the CUT or for reasonable full spectral loads, the change of the state of one channel does not effectively affects the others. Therefore, we feed the trained ML model with the powers of all the channels but the CUT and we analyze the accuracy of the guess of the CUT OSNR produced by the machine. In Fig. 4-5, we show an overall statistical description of the inaccuracy of the ML, respectively, for the ASE noise and the OSNR guesses. Clearly, as previously anticipated, the ASE prediction is effectively more accurate of the signal power prediction, not pictured here, that leads the final OSNRs inaccuracy of 0.2 dB. We repeat this investigation for all the channels and we picture the results in Fig.6. Each slice of the heat map outlines the distribution of the uncertainty of the ML prediction.

3 Conclusions

We analyzed the prediction of the OSNR component of the GSNR, being the dominant one and the mostly affected by uncertainties induced by the spectral dependence of gain and noise figure of EDFAs on the spectral load, given the nominal gain. We supposed an agnostic use of the optical line systems, by operating the EDFAs as black-boxes setting the nominal gain and by relying only on data from the OCM to predict the spectrally resolved GSNR. We experimentally obtained a data-set from an 11-EDFA OLS, focusing on the OSNR prediction only. We show that, without any specific knowledge, the uncertainty on the GSNR with the possible spectral loads needs the deployment of a margin up to 3.6 dB. Supposing to be able to acquire a training data-set before the actual traffic deployment, we show that DNN ML techniques from the TensorFlow© platform enable a reduction on the uncertainty on the OSNR down to 0.2 dB.

4 Acknowledgements

This work has been supported by Cisco within a SRA project.

5 References

[1] Cisco. Cisco Visual Networking Index: Forecast and Trends, 2017âĂŞ2022. Technical report, Cisco, 2017.

[2] Glenn Wellbrock and Tiejun J. Xia. How will optical transport deal with future network traffic growth? In *2014 The European Conference on Optical Communication (ECOC)*. IEEE, sep 2014.

[3] Vittorio Curri, Mattia Cantono, and Roberto Gaudino. Elastic all-optical networks: A new paradigm enabled by the physical layer. how to optimize network performances? *Journal of Lightwave Technology*, 35(6):1211–1221, mar 2017.

[4] Vittorio Curri, Andrea Carena, Andrea Arduino, Gabriella Bosco, Pierluigi Poggiolini, Antonino Nespola, and Fabrizio Forghieri. Design strategies and merit of system parameters for uniform uncompensated links supporting nyquist-WDM transmission. *Journal of Lightwave Technology*, 33(18):3921–3932, sep 2015.

[5] Mark Filer, Mattia Cantono, Alessio Ferrari, Gert Grammel, Gabriele Galimberti, and Vittorio Curri. Multi-Vendor Experimental Validation of an Open Source QoT Estimator for Optical Networks. *Journal of Lightwave Technology*, 36(15):3073–3082, aug 2018.

[6] Brian Taylor, Gilad Goldfarb, Saumil Bandyopadhyay, Vittorio Curri, and Hans-Juergen Schmidtke. Towards a route planning tool for open optical networks in the telecom infrastructure project. In *Optical Fiber Communication Conference/National Fiber Optic Engineers Conference 2018*, 2018.

[7] Maxim Bolshtyansky. Spectral hole burning in erbium-doped fiber amplifiers. *Journal of lightwave technology*, 21(4):1032–1038, 2003.

[8] Gert Grammel, Vittorio Curri, and Jean Luc Auge. Physical simulation environment of the telecommunications infrastructure project (tip). In *Optical Fiber Communication Conference/National Fiber Optic Engineers Conference 2018*, 2018.

[9] René-Jean Essiambre and Robert W. Tkach. Capacity trends and limits of optical communication networks. *Proceedings of the IEEE*, 100(5):1035–1055, may 2012.

[10] A. Carena, V. Curri, G. Bosco, P. Poggiolini, and F. Forghieri. Modeling of the impact of nonlinear propagation effects in uncompensated optical coherent transmission links. *Journal of Lightwave Technology*, 30(10):1524–1539, may 2012.

[11] M. Cantono, D. Pilori, A. Ferrari, C. Catanese, J. Thouras, J. L. Auge, and V. Curri. On the Interplay of Nonlinear Interference Generation with Stimulated Raman Scattering for QoT Estimation. *Journal of Lightwave Technology*, PP(99):1–1, 2018.

[12] Ronen Dar, Meir Feder, Antonio Mecozzi, and Mark Shtaif. Properties of nonlinear noise in long, dispersion-uncompensated fiber links. *Optics Express*, 21(22):25685, oct 2013.

[13] https://www.tensorflow.org/.