Estimating Quality of Transmission in a Live Production Network using Machine Learning

Jasper Müller[(1)], Tobias Fehenberger[(1)], Sai Kireet Patri[(1)], Kaida Kaeval[(1)], Helmut Griesser[(1)], Marko Tikas[(2)], and Jörg-Peter Elbers[(1)]

[(1)]*ADVA, Fraunhoferstr. 9a, 82152 Martinsried/Munich, Germany* [(2)]*Tele2 Estonia, Tallinn, Estonia*
jmueller@adva.com

Abstract: We demonstrate QoT estimation in a live network utilizing neural networks trained on synthetic data spanning a large parameter space. The ML-model predicts the measured lightpath performance with $<$ 0.5dB SNR error over a wide configuration range.

1. Introduction

Due to fast growing demand in transmission capacity of optical networks the efficient usage of the available equipment and therefore physical-layer aware optimization of the throughput in a network becomes increasingly important. An accurate method for estimating the quality of transmission (QoT) of unestablished lightpaths can be utilized to lower margins. For the optimization of channel configurations a fast method for QoT estimation is required. Determining the QoT of a lightpath entails computing the linear ASE noise and the nonlinear interference (NLI), for which detailed and accurate knowledge on the system and network components is required. While ASE noise calculation is relatively straightforward, the NLI computation can be a challenging task. Several models are available for this, such as numerical split-step simulations and Gaussian noise (GN) models [1]. They all follow the general trade-off between complexity and accuracy.

In recent years, machine learning (ML) models for QoT estimation of unestablished lightpaths have been explored, using mainly synthetic databases to leverage ML models for fast and accurate QoT estimation. In [2], a comparison of multiple models for predicting signal to noise ratio (SNR) margins is carried out, not taking the spectral assignment into account by considering fully loaded links only. Multiple classification models have been shown to achieve high accuracy for the bit-error rate threshold using a simplified parameter space, such as a neural networks and support vector machines (SVM), assuming homogeneous fiber spans in [3] or random-forest algorithms in [4], only considering the closest neighboring channels of a lightpath. In [5], the Q-value of multiple channels is predicted simultaneously with an artificial neural network (ANN) on a single testbed link collecting data from the link for training and verification. Such an ANN is a well suited ML technique for this task because it is able to learn highly nonlinear relationships between the input parameters. Previous works [2] and [3], as well as our own test, show that an ANN is able to outperform other ML algorithms when used for QoT estimation.

In this paper we show the first ANN-based QoT estimation in a live network with production channels. We generate a synthetic database, suitable to train an ML model that can accurately estimate the NLI of unestablished lightpaths for a large variety of link parameters and grid configurations. We use a significantly higher dimensional input parameter space than what is shown in previous work, reflecting the added complexity by also investigating more complex ML model architectures. We manage to train an ML model on the synthetic data base that shows a high generalization ability translating its performance onto data pulled from a live network. A maximum SNR deviation of less than 0.5 dB and an average SNR difference of less than 0.2 dB is achieved while computing the QoT of a single lightpath in microseconds, which is orders of magnitude faster than full-form GN models.

2. Versatile ML Model based on Synthesized Simulation Data

2.1. Simulation-Based Data Generation

An ML model that works for a large variety of system and link configurations usually requires for its training a big data set covering a large parameter space. Gathering this data in a live network is generally not possible because the variety of parameters available in a network is limited by the number of its links and traffic cannot be disrupted. We therefore resort to using numerical simulations for generating the necessary data on which the ML model is trained and then apply the learned model to the network under consideration. Due to the large required size of the data set and the large considered bandwidth, the EGN model [1] was used, which is capable of accurately computing NLI including Raman amplification and allowed creating a large database of $130,000$ data points in a reasonable amount of time. We used the nonlinear coefficient η as output of the model, which is related to SNR as $\mathrm{SNR} = P_{\mathrm{tx}}/(\sigma^2 + \eta P_{\mathrm{tx}}^3)$, with the linear noise σ^2 computed offline with the well-known closed-form expressions.

The choice of parameter space is based on representing the specifications and capabilities of real networks and up-to-date equipment while being sufficiently broad such that the resulting ML model is versatile and can be applied to various configurations. Hence, we consider uncompensated links using EDFAs at the end of each span, potentially supported by Raman amplification for high-loss spans. Restricting ourselves to standard single

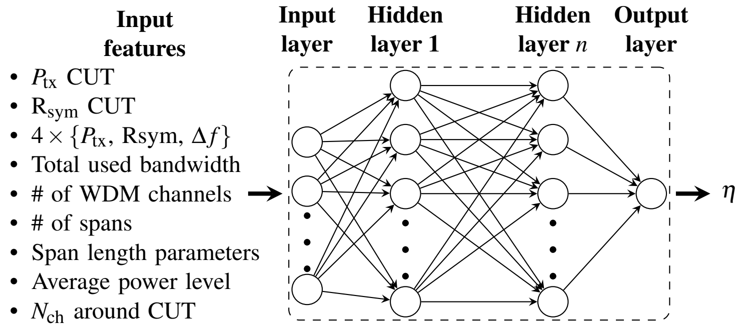

Fig. 1: Schematic of the employed ML models.

Table 1: Data generation parameters

Parameter	Range
# of spans	1 to 60, step=2
L_{span} [km]	10 to 120, step=1
α [dB/km]	0.19 to 0.275
Modulation	QPSK, 16QAM
Rsym	35 GBd, 69 GBd
Data rate	100G, 200G
P_{tx} [dBm]	-6 to 2.5, step=0.5

mode fiber of fixed dispersion (D=16.7ps/nm/km) and nonlinearity (γ=1.3 1/W/km), several fiber and system parameters were varied uniformly over a wide range, as listed in Table 1. Data channels that carry 100G (QPSK at 35GBd) or 200G (QPSK at 69GBd or 16QAM at 35GBd) are generated, with the WDM spectral width adjusted to either 50 or 75 GHz, depending on the symbol rate of the channel. The channel configurations reflect a subset of configurations available with ADVA's TeraFlex transponder [6]. The spectral occupancy over the entire C-band was selected randomly between 75% to 95%. The channels transmit power was varied as shown in Table 1. For further randomization the channel under test (CUT) was uniformly chosen between all generated channels. Evaluating the NLI for this setting constitutes one data point, and we generate a data set of 130,000 data points.

2.2. Specifications of the ML Model

Two different feed-forward neural networks (NNs) were employed for QoT estimation. The baseline is a standard ANN using the leaky rectified linear unit (ReLU) activation function. This ANN is compared to an ANN using the scaled exponential linear unit (SeLU) activation function, also called self-normalizing neural network (SNN), which has been shown to outperform other NNs on a variety of data sets [7]. Preliminary tests suggested the use of a cone-shaped ANN in which the number of neurons in each layer are successively halved from the first hidden layer to the output layer. For the SNN, a constant number of neurons was used. Both models were trained for 50 epochs, minimizing the RMS error with the adamax optimizer. We used a batch size of 64 and started with a learning rate of 0.01, divided by 10 every 10th epoch. A grid search was used to determine the number of hidden layers and the number of neurons for the ANN and the SNN. The NN architecture and the used input parameters are shown in Fig. 1. The choice of input parameters was found to be crucial for the performance of the ML model. While a large input space potentially allows for a finer regression model, the required size of the data set for training becomes prohibitively large. We heuristically determined a mix of scalars and averaged metrics that have a strong influence on NLI. For example, instead of up to 60 span lengths as input, we found that the link architecture is well represented by using the average, min., max., variance and the average of the cumulative sum. For the grid we use specific parameters (symbol rate (R_{sym}), transmit power and distance to CUT) for the four closest neighbors to the CUT. Additionally we use the number of channels N_{ch} in the 10 closest 150 GHz areas around the CUT center frequency and further parameters shown in Fig. 1 representing the full grid. The best performance for the ANN was achieved with 8 hidden layers, starting with 512 neurons on the first hidden layers, and for the SNN with 16 layers with 64 neurons each. For training and evaluation, we followed the usual 70/15/15 split of the data into training, validation and test set, respectively, with hyper-parameters optimized on the validation set.

2.3. Numerical Results Based on Synthetic Simulation Data

We verify the performance of the ML models on a test set of $20,000$ datapoints using the EGN model results as baseline and the SNR deviation $\Delta\text{SNR} = |\text{SNR}_{\text{ML}} - \text{SNR}_{\text{EGN}}|$ as metric. Both ML models show a mean SNR deviation of 0.07 dB on the synthetic data. The SNN has a slightly lower maximum SNR deviation of 0.58 dB compared to 0.64 dB for the ANN. Both models predict the NLI for the entire test set in under 1s ($i7-7500U$ CPU with 12GB RAM), thus having a prediction time in the order of μs and being magnitudes faster than the EGN model, that requires more than 1 minute for a single datapoint on average.

3. Live Network Study

3.1. System and Link Configuration

The ML-based QoT estimation is performed in a pan-European live network operated by Tele2 Estonia spanning several thousand kilometers and has the majority of the C-band filled with high-margin 100G QPSK channels. Within the disruption-free live channels, five ADVA TeraFlex transponders were installed for testing purposes. Our test-channels were inserted to the dedicated ROADM C-port using an 8-port splitter/combiner module and occupied 400 GHz spectrum around 193.95 THz inside the C-band, configured as an optical spectrum-as-a-service. Two different physical loopback locations were evaluated, giving total link lengths of 1792 km and 3751 km.

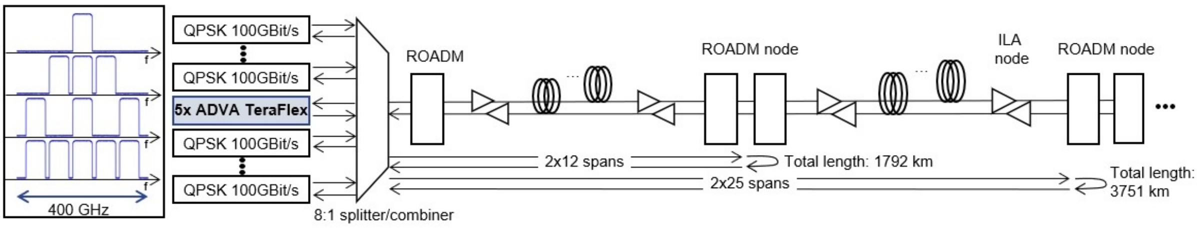

Fig. 2: Live network link with two loopback locations for the evaluated ADVA TeraFlex channels.

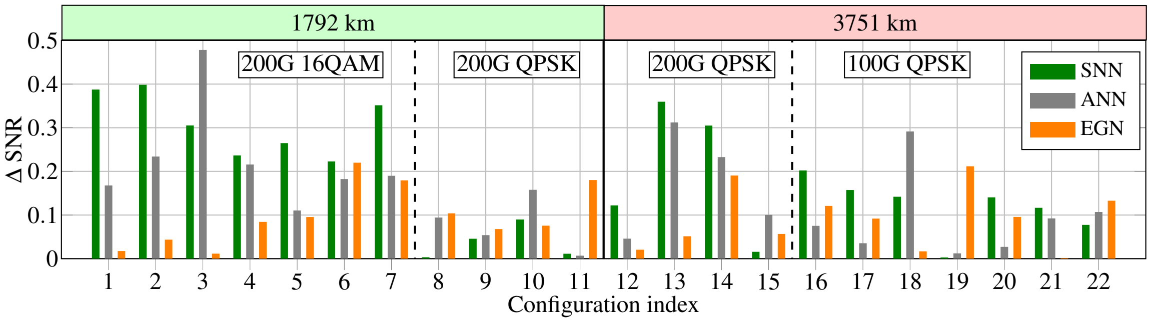

Fig. 3: SNR deviation between live network measurement and the ANN, SNN, and EGN models for 22 different configurations. The boxes give the data rate and modulation format of the CUT.

The effective SNR, sometimes called generalized SNR (GSNR), of the center channel was recorded. The used configurations were 100G and 200G QPSK on the long link, and 200G QPSK and 16QAM on the short link. For each CUT configuration, the four WDM neighbours, having the same modulation format, were varied according to the on/off modes shown on the left of Fig. 2. Additionally the on/off modes were repeated for 200G QPSK neighbors in the case of the 35 GBd CUT, leading to a total of 22 different configurations. This allows to capture the NLI impact of neighbouring channels on the CUT and thus to efficiently test the model's capability of QoT estimation. The linear SNR is calculated offline, including penalty terms for in-line ROADM filtering of the 69 GBd channels ($\sim$ 2 dB) and for any overlap of the 69 GBd channels on a 75 GHz grid (0.2 dB). Finally, the deviation from the measured values to the EGN and ML models are computed.

3.2. Numerical Results

The performance of the models and the EGN model on this data is shown in Fig. 3 showing the SNR deviations for all 22 configurations. A maximum SNR deviation of 0.22, 0.40, 0.48 dB and a mean deviation of 0.1, 0.18, 0.15 dB is observed for the EGN, SNN and ANN model respectively. At less than 0.5 dB SNR deviation, the SNR of the ML models is computed within microseconds, which is several orders of magnitude faster than the EGN model. The two NNs show similar performance, with the SNN having a slightly lower maximum SNR deviation, but a slightly higher mean deviation than the ANN. Among the excellent fit on all configurations, the highest deviations of both ML models are observed for the 200GBit/s 16QAM channel. This can be attributed to the CUT's modulation format not being included in the models input parameters. As the inclusion into the current data set leads to a disproportional deterioration of the performance on the QPSK channels, this requires a modified and extended data set, which is left for further research.

4. Conclusions

We have introduced an ML model for QoT estimation of unestablished lightpaths in a live production network. The model is trained on a synthetic database populated by the EGN data and uses an input parameter space that models the NLI of a lightpath well, without becoming prohibitively large. In combination with a deep network architecture, this enables to transfer the model performance to a live network, showing a maximum SNR error of less than 0.5 dB. The model computes the NLI of a lightpath within a few microseconds, potentially enabling for real-time network management and operation applications.

The work has been partially funded by the German Ministry of Education and Research in the project OptiCON (**#16KIS0989K**).

References

1. A. Carena *et al.*, "EGN model of nonlinear fiber propagation," Opt. Express, **22**(13), 2014.
2. R. M. Morais *et al.*, "Machine Learning Models for Estimating Quality of Transmission in DWDM Networks," JOCN, **10**(10), 2018.
3. S. Aladin *et al.*, "Quality of Transmission Estimation and Short-Term Performance Forecast of Lightpaths," JLT, **38**(10), 2020.
4. C. Rottondi *et al.*, "Machine-Learning Method for Quality of Transmission Prediction of Unestablished Lightpaths," JLT, **10**(2), 2020.
5. Z. Gao *et al.*, "ANN-Based Multi-Channel QoT-Prediction over a 563.4-KM Field-Trial Testbed,"JLT, **38**(9), 2020.
6. ADVA, "TeraFlex", www.adva.com/en/products/open-optical-transport/fsp-3000-open-terminals/teraflex, Accessed: 2021-01-23.
7. G. Klambauer *et al.*, "Self-Normalizing Neural Networks" arXiv:1706.02515, 2017.