Field Trial to Assess Bayesian Optimization for Improving QoT Estimation

Reda Ayassi[(1),(2)], Ahmed Triki[(1)], Bertrand Le Guyader[(1)], Florian Frank[(1)], Thierry Marcot[(1)], Esther Le Rouzic[(1)], Maxime Laye[(1)], Roberto Minerva [(2)], Noel Crespi[(2)]

[(1)] Orange Innovation, 2 Avenue Pierre Marzin, 22300 Lannion ahmed.triki@orange.com [(2)] Samovar, Telecom SudParis, Institut Polytechnique de Paris, 91120 Palaiseau, France

Abstract *A Bayesian optimization-based approach was investigated with data extracted from a live network as part of a field trial to retrieve hard-to-measure equipment parameters. These parameters were then utilized to improve the accuracy of the Quality of Transmission estimation.*

Introduction

The Quality of Transmission (QoT) includes all the metrics that assess the quality of the received optical signal, particularly the Signal-to-Noise Ratio (SNR), the Bit Error Rate (BER) and the Q-factor. These metrics are crucial during network design and for assessing the feasibility of a lightpath when establishing a new service between two nodes in an optical network. The QoT estimation is traditionally performed using analytical models[1]. These models vary in their accuracy (i.e. the difference between the measured and the computed QoT), execution time and working hypothesis, which leaves room for improvement.

Several studies have suggested using Artificial Intelligence (AI)-based solutions to estimate or improve the QoT. In[2], we distinguish three categories of solutions according to their scope: i) checking lightpath feasibility; ii) estimating a lightpath's QoT; and iii) enhancing analytical models. While the first category focuses on providing a binary result concerning the feasibility of a path in terms of a QoT metric, the second category aims to retrieve the exact value of the QoT. The third category of solutions instead uses AI to assist the analytical QoT estimation tools to improve the accuracy of the QoT computation.

Our solution falls into the third category, as we use an optimization algorithm to refine the values of a set of parameters considered as uncertain. Value uncertainty stems from several factors, including equipment aging and the difficulty of taking measurements in the field.

In this paper, we improve our Bayesian optimization-based model proposed in[3] by using both SNR and power measurements in the objective functions. Moreover, we validate this new model by applying it in a live operational network as part of a field trial.

Problem Formulation

We denote Q_{l_i} as the QoT indicator of a candidate lightpath l_i, to be deployed in a network $G = (N, L)$, where N and L are the set of nodes and links respectively. $\mathcal{P}_{l_i}$ represents the set of parameters related to nodes N_i and links L_i composing the l_i route as well as the equipment configuration K_i (i.e., $\mathcal{P}_{l_i} = (N_i, L_i, K_i)$). To compute the QoT of several lightpaths, the analytical model $\mathcal{M}$ relies on a set of input parameters $\mathcal{P} = \bigcup_{l_i} \mathcal{P}_{l_i}$. We split the parameters $\mathcal{P}$ into certain $\mathcal{P}_C$, which represents parameters whose values are trusted, and uncertain parameters $\mathcal{P}_U$, which represents parameters whose values might be missing or have changed during the life-cycle of the network, (i.e., $\mathcal{P} = \mathcal{P}_C \cup \mathcal{P}_U$).

We assume that we have the measured values of the QoT of M lightpaths in the network $Q = \{Q_{l_i} : i \in [1, M]\}$. We compute the estimated QoT of those lightpaths using an analytical model $\mathcal{M}$ to obtain $\tilde{Q} = \{\tilde{Q}_{l_i} : i \in [1, M]\}$. Let f_{V_U} be the parametric function used by $\mathcal{M}$ to compute the QoT of a candidate lightpath l_i for some given values of uncertain parameters $V_U = \{v_p : p \in \mathcal{P}_U\}$ such that $\tilde{Q}_{l_i} = f_{V_U}(l_i)$. As f_{V_U} is supposed explicitly unknown, we formulate our problem as a black box multi-objective general optimization problem[4] with the goal of determining $V_{U^*} = argmin_{V_U}\{E_{l_i} = \left|Q_{l_i} - f_{V_U}(l_i)\right| : i \in [1, M]\}$. By finding the uncertain parameters' values V_{U^*} that optimize the QoT of a set of lightpaths M, we aim to improve the QoT estimation of the soon-to-be-deployed lightpaths.

We set up a learning process to optimize network parameters after each service deployment as follows: after the deployment of the k_{th} service s_k, the network parameters are optimized using the performance metrics of the services already deployed $S_D = \{s_j : j \in [1, k]\}$ as objective func-

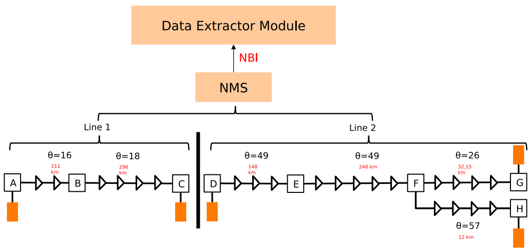

Fig. 1: The subnetwork transmission lines used for the optimization process

tions. The expected performance of the services not yet deployed $S_{ND} = \{s_j : j \in [k+1, N]\}$ are then estimated using the new uncertain parameters' values. We assess the performance of our solution by computing the QoT estimation accuracy on S_D and S_{ND}. In the remainder of this paper, we refer to the error in the estimated SNR of deployed services S_D as the training error and the error in the estimated SNR of non deployed services S_{ND} as the validation error.

The learning process described above is applied to the Bayesian optimization in the same way as[3]. We consider four uncertain parameters: the noise figure of amplifiers (NF), connector loss (CL), fused loss in the fiber (FL), and the power equalization indicator (PEI), which controls the output power of the Reconfigurable Optical Add Drop Multiplexer (ROADM). These parameters are not the only source of uncertainty but they are the most meaningful contributors to noise in SNR computation. For instance, the NF impacts on the linear noise, while the CL, FL and PEI impact on signal power, and thereby on non linear noise. In addition to the SNR, we extend the objective functions by considering the total output power from the amplifiers. This constrains the problems by reducing the degrees of freedom in the parameter space. The multiple objective functions related to SNRs and the output powers are aggregated into one using the root mean squared error (RMSE) operator.

Field Trial Description

Tab. 1: List of monitored services in the two transmission lines

Line	Route	List of Service
1	A-B	S_{11}
	B-C	S_2, S_3, S_4, S_5, S_6, S_7, S_8, S_9
	A-C	S_1, S_{10}, S_{12}
2	D-F	S_1, S_7, S_8
	D-G	S_2, S_3
	D-H	S_4, S_5, S_6

The field trial was performed in two optical transmission lines that belong to a real operational transport network. The network is managed by a Network Management System (NMS) and is composed of 8 ROADMs and extended over 1000 km. These two lines are independent of each other and have different equipment configurations. Each Optical Multiplex Section (OMS) is composed of fiber spans of different lengths as depicted in Fig. 1.

We developed a data extractor module to collect data related to the two transmission lines. The extractor uses the North Bound Interface (NBI) of the NMS to provide: i) topology-related data that mainly concern nodes, network elements cards and links; ii) service-related data that provide for instance the modulation format, bit rate and routing of each service; and iii) performance-related data that provide the output power of each amplifier and the BER of each service. The BER is transformed to SNR using transponder B2B characterization as in[5]. Data extracted from the NMS are not exhaustive, which leads us to complete the models with external information such as equipment characterizations and specifications. Several services (between 16 and 49) are transmitted through the setup with different source-destination nodes leading to a different spectrum load distribution in each OMS. As depicted in Tab. 1, we only monitored the performance of a selected list of services, 12 services in Line 1 and 8 services in Line 2.

Since the monitored services already exist in the field, we emulate a progressive deployment of services to test the potential effect of the learning process on QoT estimation. In this scenario, we assume, at the beginning of each step, that a new service is deployed (i.e., one service is removed from S_{ND} and added to S_D). Then, we perform the optimization based on performance collected from S_D. Finally, we compute the training error and the validation error corresponding to the error between the measured and the estimated SNR in S_D and S_{ND}, respectively.

The Gaussian Noise Like Interference model is considered as the the analytical model $\mathcal{M}$. Thus, we choose the implementation provided by the GNPy library[6] to calculate the estimated value of the SNR of each service (i.e., $\tilde{Q}$), as well as the power at the output of each amplifier. The goal of the experiment is to test if our optimization model is able to minimize the error between the SNR measured in the field and those computed

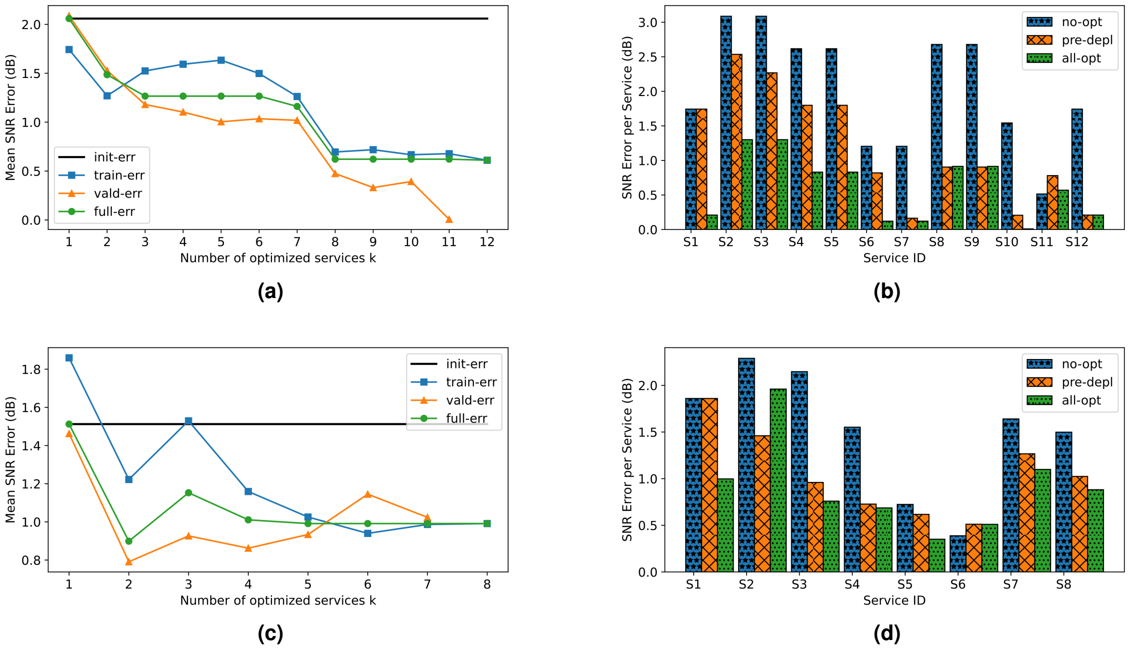

Fig. 2: Optimization results of the progressive deployment scenario. (a) Mean SNR error progression for Line 1. (b) SNR error per service for Line 1. (c) Mean SNR error progression for Line 2. (d) SNR error per service for Line 2

by GNPy.

Learning Process and Results Discussion

The graphs in Fig. 2a and Fig. 2c depict the progression of the mean SNR error after each successive optimization for Line 1 and Line 2 respectively. The initial error (init-err) shows the expected performance without the optimization process and serves as a benchmark. The training error (train-err) and validation error (vald-err) are calculated at each step k by averaging the SNR error for the S_D and S_{ND} services respectively. We note that the initial error is around 2 dB for Line 1 and 1.5 dB for Line 2. Those errors are mainly due to the fact that data extracted from the live network and used to model the two lines in GNPy are not reliable (e.g., missing or not up-to-date values). The high initial error stresses the importance of refining network parameters. After performing the optimization, the mean SNR error (i.e., full-err) is reduced by 1.44 dB for Line 1 and 0.52 dB for Line 2. The ability of the algorithm to minimize the error depends on the number of parameters to optimize and the number of services already deployed. For instance, the algorithm needs more iterations to converge to the correct uncertain parameters' values in Line 2 than in Line 1 as the topology of Line 2 is more complex. The training and validation error slightly fluctuate depending on which service is optimized first, but the overall trend shows that the error decreases proportionally to the number of services taken into consideration during the optimization.

To better understand the optimization results, we plot the improvement in SNR estimation for each service in graphs Fig. 2b and Fig. 2d. The error without optimization (no-opt) is the initial error per service. The error before service deployment (pre-depl) shows the expected improvement in SNR estimation if the optimization process is performed just before deployment. The error after all optimizations (all-opt) is the expected performance after all optimizations have been performed. We observe that the initial SNR error for each service varies between 0.5 dB and 3 dB. After successive optimizations, the performance of services having an initial error over 0.5 dB is significantly reduced and the improvement in SNR error can reach up to 1.78 dB. Furthermore, the all-optimization error shows that continuous optimizations serve to further increase the accuracy of the QoT estimation and do not penalize the already deployed services. By analyzing the pre-depl error, we can expect up to 1.6 dB of improvement in QoT estimation.

Conclusion

We present the results of a field trial where a Bayesian optimization-based solution was used to improve QoT estimation provided by an analytical model. By refining data directly extracted from the live network, the accuracy of QoT estimation was improved by up to 1.78 dB per service.

References

[1] P. Poggiolini and Y. Jiang, “Recent advances in the modeling of the impact of nonlinear fiber propagation effects on uncompensated coherent transmission systems”, *J. Lightwave Technol.*, vol. 35, no. 3, pp. 458–480, Feb. 2017. [Online]. Available: https://opg.optica.org/jlt/abstract.cfm?URI=jlt-35-3-458.

[2] R. Ayassi, A. Triki, N. Crespi, R. Minerva, and M. Laye, “Survey on the use of machine learning for quality of transmission estimation in optical transport networks”, *J. Lightwave Technol.*, vol. 40, no. 17, pp. 5803–5815, Sep. 2022. [Online]. Available: https://opg.optica.org/jlt/abstract.cfm?URI=jlt-40-17-5803.

[3] R. Ayassi, A. Triki, M. Laye, E. L. Rouzic, N. Crespi, and R. Minerva, “Bayesian optimization-based algorithm to improve the quality of transmission estimation”, in *OSA Advanced Photonics Congress 2021*, Optica Publishing Group, 2021, NeF2B.3. DOI: 10.1364/NETWORKS.2021.NeF2B.3. [Online]. Available: https://opg.optica.org/abstract.cfm?URI=Networks-2021-NeF2B.3.

[4] M. Pelikan, D. E. Goldberg, and E. Cantú-Paz, “Boa: The bayesian optimization algorithm”, in *Proceedings of the 1st Annual Conference on Genetic and Evolutionary Computation - Volume 1*, ser. GECCO’99, Orlando, Florida: Morgan Kaufmann Publishers Inc., 1999, pp. 525–532, ISBN: 1558606114.

[5] G. Grammel, V. Curri, and J.-L. Auge, “Physical simulation environment of the telecommunications infrastructure project (tip)”, in *Optical Fiber Communication Conference*, Optica Publishing Group, 2018, p. M1D.3. DOI: 10.1364/OFC.2018.M1D.3. [Online]. Available: https://opg.optica.org/abstract.cfm?URI=OFC-2018-M1D.3.

[6] A. Ferrari, M. Filer, K. Balasubramanian, *et al.*, “Gnpy: An open source application for physical layer aware open optical networks”, *J. Opt. Commun. Netw.*, vol. 12, no. 6, pp. C31–C40, Jun. 2020. DOI: 10.1364/JOCN.382906. [Online]. Available: https://opg.optica.org/jocn/abstract.cfm?URI=jocn-12-6-C31.