How to Integrate Machine-Learning Probabilistic Output in Integer Linear Programming: a case for RSA

Matteo Salani[1], Cristina Rottondi[2], Massimo Tornatore[3]*

[1]*Dalle Molle Institute for Artificial Intelligence, Lugano, Switzerland*
[2]*Dept. of Electronics and Telecommunications, Politecnico di Torino, Turin, Italy*
[3]*Dept. of Electronics, Information and Bioengineering, Politecnico di Milano, Milan, Italy*
**massimo.tornatore@polimi.it*

Keywords: Machine Learning; Optical fiber networks; Quality of service; Routing;

Abstract

We integrate machine-learning-based QoT estimation in reach constraints of an integer linear program (ILP) for routing and spectrum assignment (RSA), and develop an iterative solution for QoT-aware RSA. Results show above 30% spectrum savings compared to solving RSA with ILP using traditional margined reach computation.

1 Introduction

Routing and Spectrum Assignment (RSA) in Elastic Optical Networks (EONs) requires pre-deployment estimation of the Quality of Transmission (QoT) of each lightpath, to assess which modulation formats can be safely used while considering the impact of physical-layer impairments.

Traditional approaches for QoT estimation rely either on analytical models (e.g., the Gaussian noise model [1]) or on backpropagation methodologies that require solving an inverse nonlinear Schroedinger equation through the fiber to estimate the transmitted signal (e.g., the Split Step Fourier Transform [2]). The former approach requires full knowledge of the transmission parameters (e.g., fiber type and length, amplifier noise figure etc.) to ensure accurate estimation. Since knowledge of such parameters is often affected by uncertainties in real network deployments [3], conservative approximations on their values might lead to higher and unnecessary OSNR margins and hence significant underutilization of network resources. The latter approach is computationally too expensive to be invoked in real-time decision processes. Conversely, novel Machine-Learning (ML) based QoT estimation techniques leverage the knowledge extracted from field measurements (e.g., the Bit Error Rate - BER – of existing lightpaths) acquired by means of optical performance monitors deployed at receiving nodes to build prediction models for the QoT of unestablished lightpaths (see [4] for a thorough overview). However, how to integrate ML-based QoT estimation in an RSA framework leading to low-margin EON design is still an open problem: while some heuristic RSA approaches incorporating ML estimations to take resource allocation decisions have recently been proposed (see e.g. [5] and [6]), the design of ML-aided models for optimal RSA has not been addressed yet.

This paper discusses how to integrate the probabilistic output of a ML classifier for QoT estimation [7] in an Integer Linear Program (ILP) formulation for RSA [8] aimed at minimizing the overall spectrum occupation, considering multiple modulation formats. The classifier predicts the probability that the BER of a candidate lightpath will exceed a given threshold, based on five features (pertaining to the characteristics of the lightpath to be deployed): the lightpath total length and maximum link length, the number of traversed links, the amount of traffic to be transmitted and the modulation format to be adopted. Optionally, other features (representing network state when deploying a new lightpath) can be included, such as guardband size, traffic volume and modulation format of the spectrally nearest left and right neighbor channels (co-propagating with the considered lightpath along at least one of its links). Given a traffic request (i.e., a new request for lightpath establishment) containing the above-mentioned features, the classifier outputs a probability p_{th} that the ligthpath configuration will satisfy a given threshold *th* on the BER measured at the receiver. Then, to obtain a yes/no answer, the probability output is binarized according to the following rule: given a reference value p^*_{th}, the ligthpath is considered to be feasible if $p^*_{th} \geq p^*_{th}$, infeasible otherwise.

Once an initial ILP solution is found, the classifier can be queried iteratively to further refine it by including features about the neighbour channels of each lightpath, to account for inter-channel interference. This allows to insert additional constraints in the ILP formulation to prevent that the deployment of neighboring channels causes excessive QoT degradation at the receiver nodes. We show that, in our case study, the proposed framework saves up to 36% in spectrum occupation with respect to a traditional RSA solution with margined reach computations.

2. Integration of ILP and ML for low-margin RSA

As depicted in Fig. 1, the ILP model for RSA and the ML-based QoT classifier can be integrated according to three variants, depending on how reach constraints are incorporated in the ILP formulation: *margined*, *ML-based* and *ML-based iterative*. The margined approach is used as benchmark to assess the performance of the other two approaches. In the following, we discuss each approach in detail.

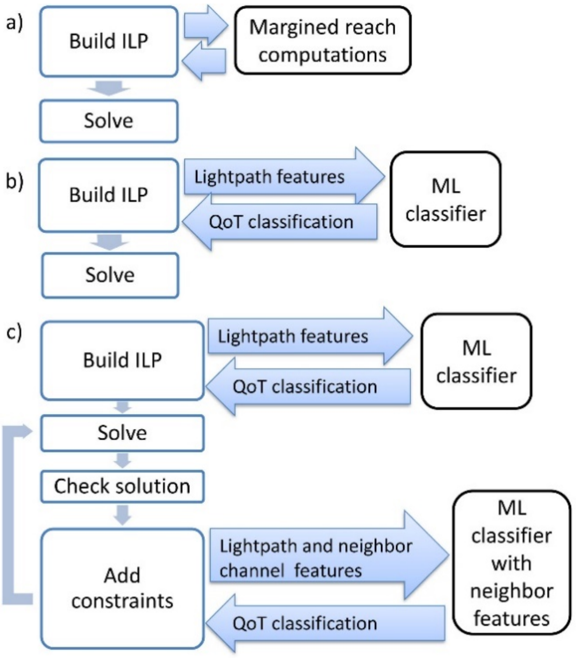

Fig. 1 The three possible strategies to integrate RSA and QoT estimation: margined (a), ML-based (b) and ML-based iterative approach (c).

2.1 Margined Approach

Traditionally, reach constraints in ILP models for RSA are set based on pre-computed reach values obtained via margined analytical formulas (see Fig. 1a). Such formulas identify, for each modulation format, the maximum distance that can be covered by the optical transmission while still ensuring an acceptable QoT. More in detail, let M be the set of modulation formats available for transmission, let K be the set of candidate ligthpaths for a given traffic request and let x_{mk} be a binary variable set to 1 if modulation format $m \in M$ is used to transmit a traffic request over lightpath $k \in K$, to 0 otherwise. Reach constraints can be defined as:

$$x_{mk} \leq \lambda_{mk} \quad \forall\, m \in M \tag{1}$$

where λ_{mk} is a binary parameter set to 1 if the reach of modulation format m computed via margined formulas exceeds the length of lightpath k, to 0 otherwise. Constraints are included in the RSA formulation using Eqn. (1). Due to limited space, we do not report the entire ILP formulation, which can be found in [8].

2.2 ML-Based Approach

In this second approach, the margined reach computation seen above is replaced by the output a ML-based QoT classifier (see Fig. 1b). Our proposed ML classifier (see [5]) takes as an input a possible combination of traffic volume, route and modulation format, and it returns the probability p_{th} that the considered combination is feasible. Based on such output, the reach constraints in Eqn. (1) can now be redefined by setting the value of λ_{mk} as follows:

$$\begin{cases} \boldsymbol{\lambda_{mk}} = \mathbf{1} \text{ if } \boldsymbol{p_{th}} \leq \boldsymbol{p^*_{th}} \\ \boldsymbol{\lambda_{mk}} = \mathbf{0} \text{ otherwise} \end{cases}$$

2.3 ML-Based Iterative Approach

A limitation of ML-based QoT classifiers (as the one in [5] used here) is that they are prone to errors (i.e., a lightpath could be erroneously classified as feasible, despite its actual received BER exceeds the threshold, or viceversa). Such errors could be reduced if information on the spectrally adjacent channels co-propagating with the considered lightpath are provided as additional input to the classifier. reach constraints that consider the crosstalk impairments introduced by the spectrally nearest left and right neighbor channels would be computationally impractical, as their number would be in the order of 10^6 or above in realistic scenarios [8]. In addition, note that, with this approach, it would be necessary to run an a-priori ML-based estimation of all possible configurations of the candidate lightpaths together with their neighboring channels, to calculate the value of the parameters to be inserted in such constraints. To avoid such scalability issues, an initial RSA solution can be found using the ML-based classifier without including neighbor-related features, then the classifier can be queried again to refine its estimation by including features about the neighboring channels of each lightpath (hence being able to account for inter-channel interference) considering the deployment obtained in the initial RSA solution (see Fig. 1c). If the classifier returns a negative outcome (i.e., the ligthpath QoT is not acceptable), an additional constraint is added to the ILP formulation to exclude from the RSA solution the unacceptable combination of a lightpath together with its neighbors. Note that, by using this iterative approach, we do not need to issue a query to the ML classifier for all candidate lightpaths and every possible choice of neighboring channels, but only for those included in the initial RSA solution. More in detail, given an unfeasible triplet of lightpaths (k; k_l; k_r) returned by the QoT classifier (where k_l is the spectrally nearest left neighbour channel and k_r is the spectrally nearest right neighbor channel of lightpath k) and their respective modulation formats (m; m_l; m_r), the constraints to be added to the ILP can be expressed as:

$$\boldsymbol{x_{mk} + x_{m_r k_r} \leq 1} \tag{2}$$

$$\boldsymbol{x_{mk} + x_{m_l k_l} \leq 1} \tag{3}$$

After Eqns. 2 and 3 have been added, the updated ILP is solved again and the procedure to add constraint, if needed, is repeated (see Fig. 1c). The iterative approach stops when the probability p_{th} that the estimated BER is below the target threshold is higher than the reference value p^*_{th} for all the lightpaths or when a stopping criterion, either on computational time or number of iterations, is reached.

3 Numerical Assessment

3.1 Simulations Settings

We consider the Japan network topology [7] and assume the usage of a flexi-grid with slices of F = 12.5 GHz and elastic transceivers with bandwidth B = 37.5 GHz. The guardband size is set to G = 12.5 GHz. The available spectrum over each link is 4 THz. The set M of modulation formats includes dual polarization (DP)-BPSK, DP-QPSK and DP-n-QAM, with n = {8; 16; 32; 64}, resulting in capacities of from 50 to 300 Gbps. Margined reach values of each modulation format are taken from [8]. We consider 5 all-to-all static traffic matrices with traffic requests uniformly distributed in the range [50; 400] Gbps, with 50 Gbps granularity. We consider either d ={1;2} candidate shortest paths per request thus obtaining the 10 input problem instances reported in Table 1.

For the ML-based approaches, the two classifiers reported in Fig. 2b)-c) have been trained with 20000 "historical" instances. Due to the unavailability of lightpaths' BER measurements gathered from real networks, these instances are synthetically generated by means of a simulator [5] for a linear optical communication system affected by chromatic dispersion and additive white Gaussian noise. Transparent links of dispersion uncompensated standard single-mode fibers are assumed, with optical amplifiers characterized by 20 dB gain and 5 dB noise figure, equally spaced over the links every 100 km. On top of the link budget, a negative-exponential random additional penalty is added to model uncertainty in transmission parameters knowledge. The threshold th is set to 4×10^{-3}. We explore different values of p^*_{th} in the set {0.5; 0.75; 0.90; 0.99} .

3.2 A Note on Feasibility

The margined approach does not admit any solution for instances 3,4,5 and 8,9,10. Since, by design, all traffic demands admit at least one feasible modulation format to be used for transmission, the infeasibility is due to insufficient link capacity for instances with total traffic larger or equal to 38.45 Tbps. Conversely, both ML-based approaches are capable of finding a spectrum assignment for all traffic demands, at the cost of having some lightpaths with BER above threshold. But the number of such lightpaths can be remarkably reduced, e.g., for $p^*_{th} \in$ {0.9; 0.99}, the iterative approach provides a solution with BER below threshold for all but one instance (instance 8 with BER 4.55×10^{-3}).

3.3 Results

For the sake of brevity, we report results obtained only on instances 1,2,6 and 7, where the margined approach is able to provide a solution. Additional details on other instances are reported in [8].

In Figs. 2 and 3 we indicate the margined approach with letter A, the ML-Based approach with letter B and the ML-Based iterative approach with letter C. The thresholds p^*_{th} used in the ML-Based approaches are reported in the legends. Fig. 2 illustrates the average spectrum occupation in THz for instances 1 and 6 (traffic matrix T1) and instances 2 and 7 (traffic matrix T2), while Fig. 3 illustrates the distribution of the BER of all deployed lightpaths (note that the logarithmic y axis scale is truncated at 10^{-15}). We observe that the ML-Based approach substantially reduces the spectrum necessary to route the traffic, with savings up to 36%. ML-Based iterative has slightly higher resource occupation compared to ML-Based (yet, allowing averages savings of 18.2% w.r.t. the margined approach), but this additional resource occupation is paid off by the fact that ML-Based iterative largely reduces unfeasible paths. While ML-based approach produces a total of 18 lightpaths with BER above threshold, the iterative approach produces just 6. In particular, for $p^*_{th} \in$ {0.9; 0.99}, only one above threshold lightpath is produced for instance 1.

Table 1 The considered problem instances.

Inst.	Tot.[Tbps]	Aver. [Gbps]	Max. [Gbps]	d
1	26.70	146.70	200	1
2	33.75	185.44	300	1
3	38.45	211.26	400	1
4	42.90	235.71	300	1
5	48.05	264.81	400	1
6	26.70	146.70	200	2
7	33.75	185.44	300	2
8	38.45	211.26	400	2
9	42.90	235.71	300	2
10	48.05	264.26	400	2

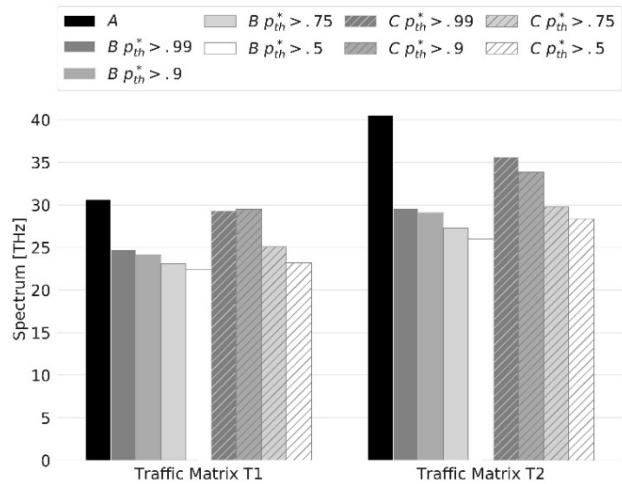

Fig. 2 Spectrum occupation.

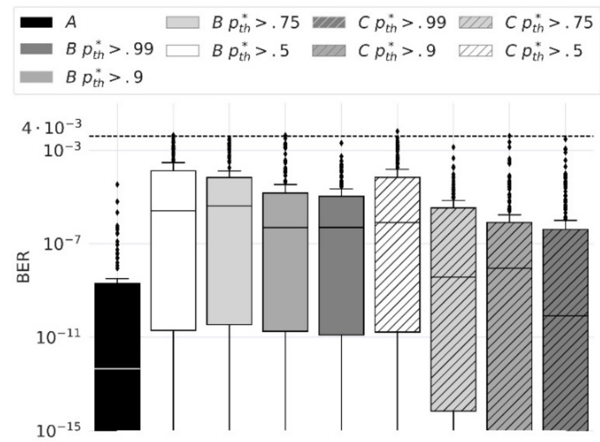

Fig. 3 BER of the deployed lightpaths.

4 Conclusion

We propose an iterative process to integrate the probabilistic outputs of a ML-based QoT estimator in an ILP for the solution of the RSA problem in EONs, considering multiple modulation formats. Results show potential for significant reduction of OSNR margins with respect to traditional approaches based on margined reaches. On the other hand, compared to non-iterative ML-based approaches, the iterative approach requires a minor increase in resource occupation, but guarantees much less BER threshold violations.

5 References

[1] Bosco, G., Curri, V., Carena, A., Poggiolini, P., and Forghieri, F.: ‘On the performance of Nyquist-WDM terabit superchannels based on PM-BPSK, PM-QPSK, PM-8QAM or PM-16QAM subcarriers’. IEEE/OSA Journal of Lightwave Technology, 2011, 29(1), pp. 53-61.

[2] Ip, E., and Kahn, J. M.: ‘Compensation of dispersion and nonlinear impairments using digital backpropagation’. IEEE/OSA Journal of Lightwave Technology, 2008, 26(20), pp. 3416-3425.

[3] Proietti, R., Chen, X., Castro, A., et al.: ‘Experimental demonstration of cognitive provisioning and alien wavelength monitoring in multi-domain EON’. Proc. of Optical Fiber Communication Conference, San Diego (CA), USA, 2018, March, pp. W4F-7

[4] Musumeci, F., Rottondi, C., Nag, A., Macaluso, I., Zibar, D., Ruffini, M., & Tornatore, M.: ‘An overview on application of machine learning techniques in optical networks’. IEEE Communications Surveys & Tutorials, 2018, in press. DOI:10.1109/COMST.2018.2880039

[5] Seve, E., Pesic, J., Delezoide, C., Bigo, S., and Pointurier, Y.: ‘Learning process for reducing uncertainties on network parameters and design margins’. IEEE/OSA Journal of Optical Communications and Networking, 2018, 10(2), pp. A298-A306.

[6] Panayiotou, T., Chatzis, and Georgios Ellinas. "Performance analysis of a data-driven quality-of-transmission decision approach on a dynamic multicast-capable metro optical network." IEEE/OSA Journal of Optical Communications and Networking 9(1), pp 98-108, 2017.

[7] Rottondi, C., Barletta, L., Giusti, A., and Tornatore, M.: ‘Machine-learning method for quality of transmission prediction of unestablished lightpaths’. IEEE/OSA Journal of Optical Communications and Networking, 2018, 10(2), pp. A286-A297.

[8] Salani, M., Rottondi, C., Tornatore, M.: ‘Routing and Spectrum Assignment Integrating Machine-Learning-Based QoT Estimation in Elastic Optical Networks’, Proc. Of IEEE International Conference on Computer Communications, Paris, France, 2019, May.