Routing and Spectrum Assignment Based on Reinforcement Learning in Multi-Band Optical Networks

Abdennour Ben Terki
Scuola Superiore Sant'Anna
Pisa , Italy
abdennour.benterki@sssup.it

João Pedro
Infinera Unipessoal Lda
Instituto de Telecomunicações, IST
Lisboa, Portugal
jpedro@infinera.com

António Eira
Infinera Unipessoal Lda
Carnaxide, Portugal
aeira@infinera.com

Antonio Napoli
Infinera
Munich, Germany
anapoli@infinera.com

Nicola Sambo
Scuola Superiore Sant'Anna
Pisa , Italy
nicola.sambo@sssup.it

***Abstract*—A routing and spectrum assignment strategy based on Reinforcement Learning (RL-RSA) is proposed for multi-band optical networks. RL-RSA accounts for Stimulated Raman Scattering using the Generalized Gaussian Noise model. Simulation results show that RL-RSA increases the throughput by 20%.**

***Index Terms*—Routing and spectrum Assignment, Reinforcement Learning, Multi-band, Optical networks**

I. Introduction

Multi-band transmission and networking are gaining attention, as exploitation of unused portions of the spectrum (e.g., S- and E-band) could accommodate traffic growth and postpone the deployment of new fibers [1]–[3]. Recently, several advances have been made in enabling technologies (e.g., amplification [4] and switching [5]). In parallel, machine learning (ML) techniques have been explored in optical networks [6] for network optimization and automation, including Quality of Transmission (QoT) estimation [7], failure detection and identification [8]. Reinforcement Learning (RL) [9] is a promising ML approach based on trial and error interactions, which presents the advantage of not requiring a training dataset. In the context of optical networks, RL has been applied to routing and wavelength assignment [10], routing [11], and restoration strategies [12]. Moreover, Deep Reinforcement Learning (DRL) has been investigated in the literature for resource allocation in optical networks [13]–[15]. However, contrary to RL, DRL requires a training data set, whose acquisition may be complex and time consuming. In general, differently from DRL, the use of RL in optical networks is either unexplored or much less explored, especially regarding multi-band optical networks. In this paper, we propose an RSA strategy based on RL (RL-RSA) specifically designed for multi-band optical networks. L, C, S, and E bands are considered with wideband transmission impairments (including Stimulated Raman Scattering – SRS) through the Generalized Gaussian Noise (GGN) model. Simulation results show that RL-RSA increases the network throughput by up to 20%.

This work has received funding from EU Horizon 2020 MENTOR program under the Marie Sklodowska Curie grant agreement 956713 and by the Horizon Europe SEASON (GA 101096120).

II. RL-RSA

An RL model is composed of the following main components: (i) agent, (ii) environment, (iii) state, (iv) action, and (v) reward. The RL model acts as an *agent* that learns through trial and error while interacting with the environment (i.e., the multi-band network in our case). The agent decides which path and portion of the spectrum to assign to a connection request between a source and a destination. The *environment* consists of the network, including links, paths, supported bands, and the Quality of Transmission (QoT) as figure of merit. In this paper, because of the adoption of the GGN model, the QoT figure of merit is the generalized signal-to-noise ratio (GSNR) [3]. The *state* represents the requested connection and the active connections, while the *action* is the path and spectrum attempted by the agent for the request. The *reward* is the feedback score assigned based on the action's success or failure (i.e., that spectrum satisfies continuity constraint over the path or not). Scores are stored in a Q-Table. The proposed RL-RSA works as follows considering a network supporting dual polarization 16 quadrature amplitude modulation (DP-16QAM) and DP- quadrature phase shift keying (DP-QPSK). It is assumed that the connection rate is fixed and the requested connection rate is fulfilled with a single channel if DP-16QAM is supported or with two channels in the case of DP-QPSK.

The GSNR is computed per channel. The algorithm 1 shows the initialization of the Q-Table, where the initial values for the scores are calculated as a function of: (i) the length of the path (π), (ii) the GSNR value of the channel on the path (γ). For instance, the shorter the path, the highest the score (given that α is a negative number). The score also depends on the coefficient β, whose value depends on the GSNR (as an example for the same route, different bands – or channels – experience different transmission performance). In particular, GSNR is compared with the GSNR thresholds of the modu-

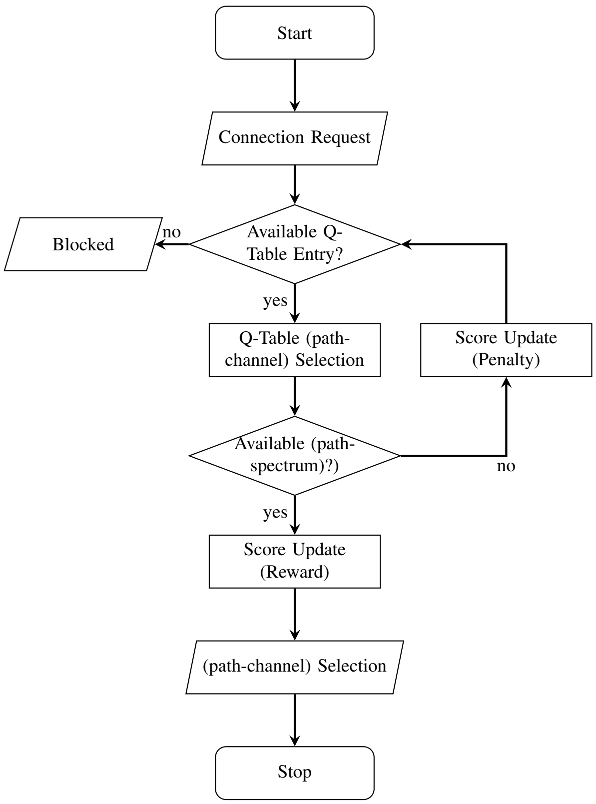

Fig. 1. RL RSA flow chart.

lation formats considered ($\mathrm{TH}_{\mathrm{DP-16QAM}}$ and $\mathrm{TH}_{\mathrm{DP-QPSK}}$). The value of β is assigned in order to give preference to the selection of the more spectral efficient DP-16QAM ($\beta = 5/12$) rather than DP-QPSK ($\beta = 1/16$). Then, the final score of the (path-channel) pair is given by $\alpha \times \pi + \beta \times \gamma$. In case DP-QPSK is selected, two spectrally adjacent channels should be set up. Then, upon connection request, the agent selects the action (path,channel) with the highest score in the Q-table. The score in the Q-Table related to that action is updated positively or negatively depending on the success or failure of the chosen (path-channel) pair as shown in Algorithm 2. The agent's learning is iterative, allowing for continuous improvement of the Q-Table. The reward system, defined in Algorithm 2, encourages the use of higher modulation formats and discourages choosing the same path-channel if a request cannot be established previously (e.g., because the channel does not satisfy spectrum continuity constraint). Fig. 1 shows the flowchart of the proposed RL-RSA. Upon connection request, the RL agent checks the availability of entries in the Q-Table for that connection request. After that, a (path-channel) pair is selected from the Q-Table with the highest score. In the case where the selected pair is available, its score will be updated positively – Algorithm 2– and (path-channel) will be assigned to the request and the connection will be established. If the selected (path-channel) pair is not available – for example if the channel does not satisfy the spectrum continuity constraint over the path –, the selected pair's score will be updated negatively, as described in Algorithm 2 and another available entry will be attempted based on the highest score. If all entries in the Q-Table have been attempted (there is no available path-channel pair), the request will be blocked.

III. Simulation Results

A custom-built simulator is used to evaluate the performance of the proposed RL-RSA for multi-band networks in comparison with a benchmark strategy based on k-shortest paths (K-SP) and load balancing for path computation and First-Fit (FF) for spectrum assignment [16] and with another strategy named *RL-based routing and Lowest GSNR* described in [11]. With the latter, the RL agent selects the path with the highest score (e.g., scores are initialized preferring shorter routes); regarding spectrum assignment, the available channel with the lowest GSNR supporting the highest-order modulation format is considered. A Japanese network topology [17] with 14 nodes and 44 links in is adopted. Traffic follows a Poisson distribution with rate λ. Connection holding time is exponentially distributed with an average of $1/\mu$ = 1 hour. Traffic load (λ/μ) is varied with λ. 400-Gb/s requests and 64-Gbaud symbol rate are assumed. Thus, 400 Gb/s requests can be served via a single DP-16QAM channel in 75 GHz or via 2×200Gb/s DP-QPSK channels over 150 GHz. GSNR is computed per channel with GNPy [18] accounting for SRS.

Algorithm 1 Q-Table Initialization per (path, channel)

Require: $(Path, Channel) \leftarrow Score$
  $\alpha \leftarrow \frac{-1}{1000}$
  $\pi \leftarrow$ Path length in km
  $\gamma \leftarrow$ Channel GSNR in dB
  if GSNR(Path, Channel) $\geq \mathrm{TH}_{\mathrm{DP-16QAM}}$ **then**
    $\beta \leftarrow \frac{5}{12}$
    Q-Table[state][action] = $\alpha \times \pi + \beta \times \gamma$
  else if GSNR(Path, Channel) $\geq \mathrm{TH}_{\mathrm{DP-QPSK}}$ **then**
    $\beta \leftarrow \frac{1}{16}$
    Q-Table[state][action] = $\alpha \times \pi + \beta \times \gamma$
  else if GSNR(Path, Channel) $< \mathrm{TH}_{\mathrm{DP-QPSK}}$ **then**
    Q-Table[state][action] $\leftarrow$ *Removed*
  end if

Algorithm 2 Score Update per (path,channel)

Require: $(path, channel) \leftarrow score\ update$
  if the connection can be established **then**
    if one channel with DP-16 QAM modulation format is assigned **then**
      $Reward \leftarrow Q\text{–}Table[state][action] \times 0.75$
    else if Two channels with DP-QPSK modulation format are assigned **then**
      $Reward \leftarrow Q\text{–}Table[state][action] \times 0.25$
    end if
  else if a connection cannot be established on that path and channel **then**
    $Penalty \leftarrow -(Q\text{–}Table[state][action] \times 0.25)$
  end if

Fig. 2. Blocking probability versus traffic load.

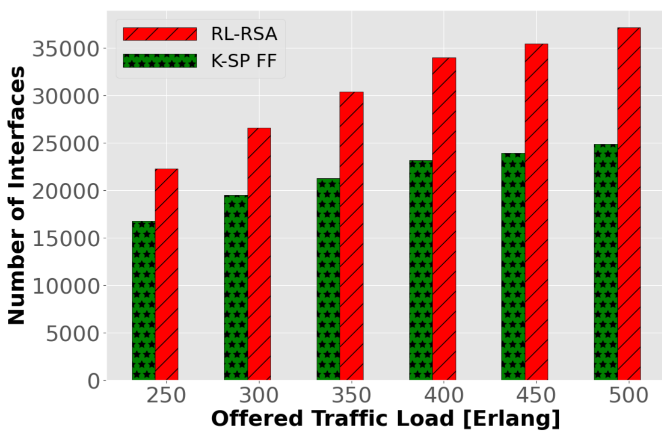

Fig. 3. Interfaces usage.

A L-C-S-E multi-band system is assumed with the supported spectrum as in [1]. The assumed GSNR thresholds are 24 dB for DP-16QAM and 16 dB for DP-QPSK. Fig. 2 shows the blocking probability at varying traffic loads. RL-RSA achieves the lowest blocking probability. K-SP FF, which performs load balancing, achieves better performance than RL-Routing Lowest GSNR, whose routing algorithm is mainly based on the distance. RL-RSA significantly decreases blocking, e.g. by half compared to K-SP for a load of 250 Erlang. This is because RL-RSA dynamically updates its network view by altering scores after choosing an action, such as a path-channel pair, with rewards assigned according to Algorithm 2. With RL-RSA, the blocking reduction reflects in a throughput increase: e.g., for a blocking of 2×10^{-2}, the load rises by 20% with respect to K-SP FF (from 250 to 300 Erlang). Fig. 3 illustrates the number of used transmitter/receiver interfaces for RL-RSA and K-SP FF algorithms. For traffic loads between 250 and 500 Erlang, K-SP FF has a lower usage of interfaces compared to RL-RSA ranging from 26% to 34%. The higher interface usage with the RL-RSA algorithm is due to two reasons: i) RL-RSA reduces blocking compared to K-SP FF, as shown in Fig. 2, meaning that RL-RSA successfully routes more traffic for the same offered load, thus requiring more interfaces; ii) RL-RSA achieves more effective load balancing, which necessitates longer connections and a higher use of the lower-order modulation format (i.e., DP-QPSK), increasing the interface count.

IV. Conclusions

In this work, we proposed a Routing and Spectrum Assignment (RSA) strategy fully assisted by Reinforcement Learning (RL). The Generalized Signal-to-Noise Ratio (GSNR) is used as a physical layer metric, accounting for linear and nonlinear effects, as well as for Stimulated Raman Scattering. Our findings reveal that RL assistance in RSA can significantly reduce blocking compared to the k-Shortest Path routing and First Fit spectrum assignment strategy. This may lead to an increase in supported network load of up to 20%.

References

[1] N. Sambo et al., ”Provisioning in Multi-Band Optical Networks,” in Journal of Lightwave Technology, vol. 38, no. 9, 2020.
[2] N. Sambo et al., ”Network upgrade exploiting multi band: S- or E-band?,” in Journal of Optical Communications and Networking, 2022
[3] A. Ferrari et al., ”Assessment on the Achievable Throughput of Multi-Band ITU-T G.652.D Fiber Transmission Systems,” in JLT 2020.
[4] Z. Chen et al., ”On-Chip Waveguide Amplifiers for Multi-Band Optical Communications: A Review and Challenge,” in JLT, 2022.
[5] J. Renaudier et al., ”Devices and Fibers for Ultrawideband Optical Communications,” in Proceedings of the IEEE, vol. 110, no. 11, 2022.
[6] F. Musumeci et al., ”An Overview on Application of Machine Learning Techniques in Optical Networks,” in IEEE Communications Surveys and Tutorials, vol. 21, no. 2, pp. 1383-1408, Second quarter 2019.
[7] F. Musumeci et al., ”Domain adaptation and transfer learning for failure detection and failure-cause identification in optical networks across different lightpaths [Invited],” in Journal of Optical Communications and Networking, vol. 14, no. 2, 2022.
[8] F. Musumeci et al., ”Domain adaptation and transfer learning for failure detection and failure-cause identification in optical networks across different lightpaths [Invited],” in Journal of Optical Communications and Networking, vol. 14, no. 2, 2022.
[9] M. H. Krishna and M. M. Latha, ”Complexity and Performance Evaluation of Segmented and Recursive Reinforcement Learning,” 2021 IEEE 4th International Conference on Computing, Power and Communication Technologies (GUCON), 2021.
[10] N. Di Cicco et al., ”On Deep Reinforcement Learning for Static Routing and Wavelength Assignment,” in IEEE Journal of Selected Topics in Quantum Electronics, vol. 28, no. 4, 2022.
[11] A. B. Terki, J. Pedro, A. Eira, A. Napoli and N. Sambo, ”Routing and Spectrum Assignment Assisted by Reinforcement Learning in Multi-band Optical Networks,” Proc. of ECOC, 2022.
[12] Z. Zhao et al., ”Service Restoration in Multi-Modal Optical Transport Networks with Reinforcement Learning,” 2020 International Conference on Computing, Networking and Communications (ICNC), 2020.
[13] J. Pinto-Ríos, et al., “Resource allocation in multicore elastic optical networks: A deep reinforcement learning approach,” Complexity, 2023.
[14] A. Beghelli and P. Morales, “Approaches to dynamic provisioning in multiband elastic optical networks,”, Proc. of ONDM 2023.
[15] N. E. D. E. Sheikh, et al., “Multi-band provisioning in dynamic elastic optical networks: a comparative study of a heuristic and a deep reinforcement learning approach,” Proc. of ONDM 2021.
[16] N. Sambo, et al., ”Lightpath provisioning in wavelength switched optical networks with flexible grid,” Proc. of ECOC, 2011.
[17] M. Salani, C. Rottondi and M. Tornatore, ”Routing and Spectrum Assignment Integrating Machine-Learning-Based QoT Estimation in Elastic Optical Networks,” IEEE INFOCOM 2019.
[18] GNPy github repository, https : / / github . com / Telecominfraproject / oopt - gnpy, Accessed: 2022-11-21.