Transfer learning Aided QoT Computation in Network Operating with the 400ZR Standard

(Article begins on next page)

Transfer learning Aided QoT Computation in Network Operating with the 400ZR Standard

Fehmida Usmani
National University of Sciences & Technology (NUST), Pakistan
fusmani.dphd18seecs@seecs.edu.pk

Ihtesham Khan
DET, Politecnico di Torino, Italy
ihtesham.khan@polito.it

Muhammad Umar Masood
DET, Politecnico di Torino, Italy
muhammad.masood@polito.it

Arsalan Ahmad
National University of Sciences & Technology (NUST), Pakistan
arsalan.ahmad@seecs.edu.pk

Muhammad Shahzad
National University of Sciences & Technology (NUST), Pakistan
mshehzaad@hotmail.com

Vittorio Curri
DET, Politecnico di Torino, Italy
curri@polito.it

***Abstract*—The current increase in bandwidth-hungry applications and the progressively evolving concept of connected "smart" devices through the internet have increased internet traffic exponentially. To hold this expansion of internet traffic, the network operators insist on the full capacity utilization of already deployed hardware infrastructure. In this context, accurate and earlier calculation of the quality of transmission (QoT) of the lightpaths (LPs) is critical for minimizing the required margins that arise due to the uncertainty in the operating point of network elements. This article proposes a novel framework in which a transfer learning assisted QoT-Estimation (QoT-E) is made. The transfer learning agent acquired the knowledge from a traditional fully operational network operating on C-band and utilized this knowledge to assist the operator in estimating the LP QoT on a state-of-the-art newly functioning network on an extended C-band operating with 400ZR standards. The measurement parameter considered to estimate the QoT of LP is the generalized signal-to-noise ratio (GSNR). The dataset used in this analysis is generated synthetically by utilizing well tested GNPy platform. Promising results are achieved in terms of reducing the overall required margin and better utilization of the residual network capacity.**

***Index Terms*—Transfer learning, Quality of transmission, Generalized signal-to-noise ratio , Wide-band networking**

I. Introduction

In the current telecommunication scenario, the rapid increase in global IP traffic, driven by the introduction of revolutionary technologies such as 5G, the Internet of Things (IoT), and cloud services, has increased the pressure on the core optical networks for capacity enhancement [1]. Generally, two possible solutions can be implemented to address this issue; installing new infrastructure or increasing the capacity utilization of already implemented optical networks. The first solution requires a significant CAPEX and is not suitable from an operator's point of view. However, the other solution is more feasible as it can increase the returns on the already deployed network infrastructure.

Most of the traditional optical transport system has been implemented uses Wavelength Division Multiplexing (WDM) in the C-band around a spectral window of $\approx 4\,\mathrm{THz}$. To optimize the existing WDM systems, it is necessary to enhance the capacity utilization along with network disaggregation. To exploit this, the capacity can be increased by the deployment of new technologies, such as Band Division Multiplexing (BDM), which enables the use of multi-band systems (extended C-band) on the existing WDM transport systems over the optical fibers (ITU G.652.D). This extended C-band increases the transmission capacity of the state-of-the-art optical network up to $\approx 4.8\,\mathrm{THz}$. This capacity enhancement and the latest technologies such as coherent transmission systems, Elastic Optical Networks (EONs), and Software-Defined Networking (SDN) paradigm empower the current state-of-the-art optical network with high data rate, good degree of flexibility, and better network control.

The EONs give the network controller the flexibility to upscale or downscale the resources based on traffic demands to efficiently use the available spectrum. At the same time, the SDN implementation allows each Network Element (NE) to be managed within a virtualized environment. SDN can facilitate the automation of complex network operations and give users greater control over the network, resulting in flexibility in implementing new services, optimized network resource utilization, lower power consumption, and more return on the CAPEX. Both attributes allow for a disaggregated approach of an optical network that allows the virtual network to be opened and sliced. The primary step towards the disaggregated and flexible optical network is to abstract the WDM optical transport as the weighted graph of LPs traversed through each NE in terms of GSNR obtained mainly from Optical Line Systems (OLS), including amplifiers and fibers [2]. The OLS controller, which runs in the control plane, determines the amplifier operating

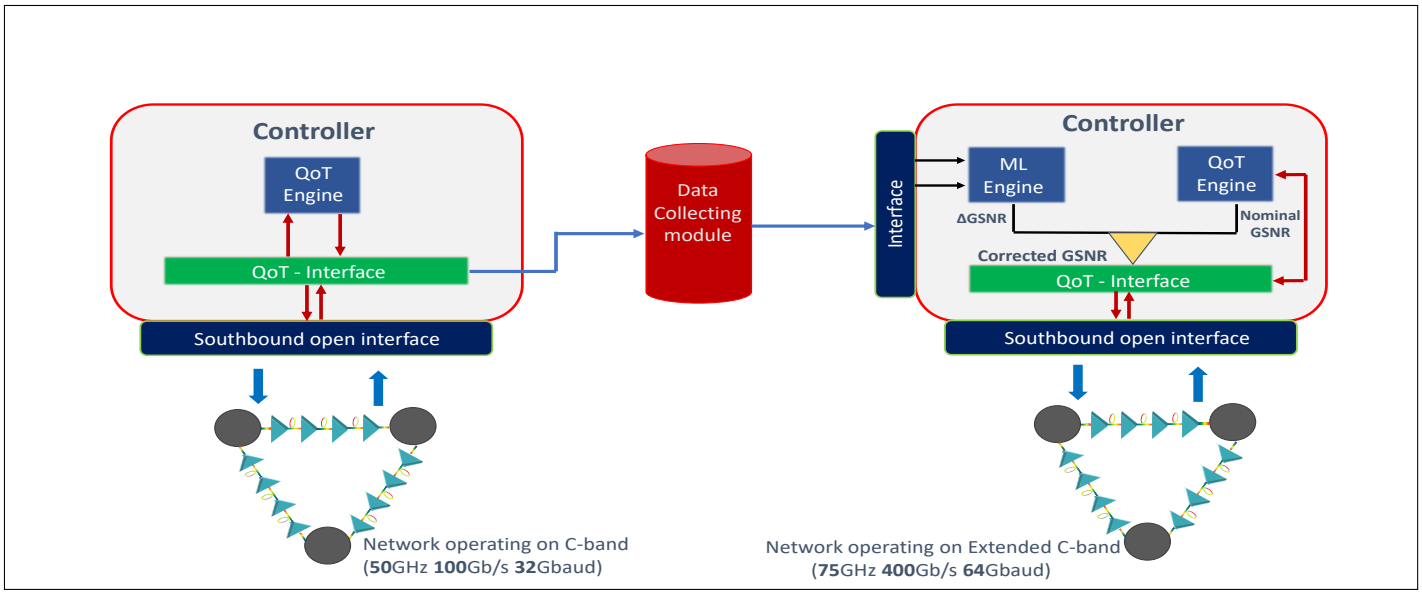


Fig. 1: Transfer learning schematic utilizing traditional C-band knowledge to assist QoT-E engine of extended C-band network.

point and eventually determines the GSNR degradation. In order to better rely on the GSNR degradation, the nominal operating point must be accurately defined. The more accurate these parameters are, the less margin is set in LP deployment, which allows better exploit the installed equipment's capacity. In addition, reliable QoT or GSNR estimation also helps with automatic recovery from network faults by reducing downtime.

The NEs are mainly suffer with some uncertainty in the operating point with respect to the nominal values due to various factors (aging effect, variation in the spectral load and effects of in-field operation). These offset values create a difference between the actual and the nominal GSNR values computed by the QoT estimator engine [3]. These offset values from the nominal point of the GSNR are mainly due to the amplifier's gain ripple and noise figure. In order to cater these uncertainties, a system margin is deployed to avoid network Out-of-service (OOS). In order to minimize the deployed system margin to better utilize the network resources, an accurate GSNR estimation is required.

In this work, we use the ML framework, which assists the QoT-E by reducing the margin required due to the uncertainty that arises by amplifier's gain ripple and noise figure. The ML paradigm has already been well analyzed for QoT-E such as; a Cognitive-case-based-reasoning (CBR) method is demonstrated in [4]. In [3], the ML-based approach is used to control OLS in an open environment. Random forest (RF) is utilized to exploit the already accumulated database in [5] to decrease uncertainty in design margins and network parameters. Numerous ML-based approaches are proposed in [6]–[8] for QoT-E of LP. In [9], a binary classifier based on RF is presented to estimate the bit-error-ratio (BER) of LPs before their establishment. In [10], the authors evaluated the performance of two Domain adaption (DA) approaches for ML-assisted QoT-E of an optical LP for a fixed/variable number of available training samples from the source/target domain. In [11], a Convolutional-neural-network (CNN)-based QoT estimator is proposed in the context of DA scenarios. Finally, the authors in [12] analyzed the QoT-E accuracy delivered by a few Active Learning (AL) and DA methods on two different network topologies.

The significant distinction of this work is that we proposed a novel framework in which we exploit the dataset originating from the in-service C-band network and use it to train a transfer learning agent to work in conjunction with the QoT-E engine of the network controller of another sister network which is operating on the extended C-band network with 400ZR standards (see Fig. 1). The purpose of the transfer learning agent is to correct the GSNR estimation of LPs of an extended C-band network and consequently reduce the margin to utilize the network resources better.

II. Simulation Model and Dataset Generation

In this work, a software-defined open optical network is considered, in which OLS is modeled as the edges, and the nodes are characterized as Reconfigurable-optical-add-drop-multiplexers (ROADMs) [13]. The OLS considered are supposed to work on the optimum operating point, and only the ripple gain of the amplifier accounts for the perturbed behavior of the physical layer. These gain ripples fluctuate with the variation of the spectral load. Therefore, OLS controllers can guarantee that they are operating at the nominal operating point with some degree of uncertainty in the operating point. On the lower layer, the LPs are transparently deployed on the WDM flexible grid system, connecting the transceivers and supporting

TABLE I

Topology Details [15]		
Parameters	**EU: Training**	**USA : Testing**
Number of Nodes	28	100
Number of Links	41	171
Average path distance (km)	2014.06	2541.75
Maximum path distance (km)	3051.10	5481.07
Minimum path distance (km)	669.30	568.33
Average number of spans per Link	19.75	27.49

dual-polarization multilevel modulation formats [14]. The LPs, during its propagation, suffer from several kinds of impairments, mainly Amplified spontaneous noise (ASE) and Non-linear impairments (NLI). The ASE noise introduced by each In-line amplifier (ILA) is statistically independent and adds up in the propagation. In contrast, the NLI of every span is statistically correlated with each other [2]. The overall $GSNR$ of each LP crossing through the OLS is given by:

$$\frac{1}{GSNR} = \sum_{n} \frac{1}{GSNR_n} \tag{1}$$

where *n* represents the number of OLSs traversed by the LP during certain path. The GSNR metric accounts for both the ASE and NLI over the given path.

The simulation framework considers two sisters network; traditional C-band network (EU network) and extended C-band (USA network). The two networks under consideration have different topologies but the same hardware; fiber types and amplifiers. The details related to network topologies for both the considered networks are described in Table I. The traditional C-band has a total bandwidth of $\approx 4\,\mathrm{THz}$, which allows carrying 80 channels over a standard 50 GHz grid and extended C-band having a total bandwidth of $\approx 4.8\,\mathrm{THz}$ which allows carrying of 64 channels over 75 GHz grid. The transceivers of traditional C-band and extended C-band operate at 32 Gbaud and 64 Gbaud, respectively, shaped with a Raised-root-cosine filter. The Erbium-doped-fiber-amplifiers (EDFAs) considered for both networks are configured to operate in a constant output power mode with 0 dBm/channel. The connections of both networks are assumed to work with Standard single-mode fiber (SMF) with a span of 80 km. The ILAs in both networks are considered to have a randomly selected noise figure for each amplifier in the 3.5 to 4.5 dB range, along with a random gain ripple with a 1 dB variation. The details of network simulation parameters are reported in Table II.

The discussed scenario is simulated using an open-source GNPy library to create synthetic datasets that abstract the physical layer. The GNPy library created an end-to-end simulation environment that generates network models for the physical layer. The datasets are generated for the C-band network (EU network) and extended C-band (USA network). The generated dataset for the traditional C-band network is the subset of 2^{80}, with 80 channels as the total possible realization of the spectral load, while for an extended C-band network, it is 2^{64}, with

TABLE II

Simulation Parameters	
Launch Power/ Channel	0 dBm
Dispersion (D)	16.0 $ps/nm/km$
Attenuation coefficient (α)	0.2 dB/km
Channel Spacing (C-Band)	50 GHz
Channel Spacing (Extended C-Band)	75 GHz
Span Length	80 km
WDM Comb (C-Band)	80
WDM Comb (Extended C-Band)	64
Baud Rate (C-Band)	32 $Gbaud$
Baud Rate (Extended C-Band)	64 $Gbaud$
Amplifier Noise Figure	[3.5 - 4.5] dB [16]
Nominal Amplifier Noise Figure	4 dB
Amplifier Gain Ripple	Variation of 1 dB
Nominal Amplifier Gain Ripple	Flat
Fiber Type	Standard SMF

overall 64 operating channels. The variation in traffic load of total bandwidth utilization for both networks ranges from 34% to 100%.

The proposed architecture exploits the knowledge of EU network to train the transfer learning agent application program interface (API) integrated along with the core QoT estimator engine in the USA network controller. The core QoT estimator engine estimates the LPs GSNR using the nominal parameters. Typically, the working point of NE changes during its operational phase causes an uncertainty in the GSNR estimation calculated by the central QoT estimator engine using these nominal parameters provided by the vendors. The presence of this uncertainty in GSNR estimation arises a demand for putting some margin, which subsequently reduces the deployable traffic rate and causes underutilization of network resources. In the proposed scenario, the transfer learning agent trained on the dataset of already operating network (EU network) is used to assists the core QoT estimator engine of other agnostic newly deploying (USA network in this case). The main focus of this work is to target the uncertainty that arises due to amplifier ripple gain and noise figure.

III. Transfer learning agent

This work introduces a transfer learning module capable of supporting the core QoT estimator engine to correct the estimated GSNR of a particular LP of the newly deployed extended C-band network using the acquired knowledge from the already operating C-band network. The transfer learning agent proposed in the SDN controller is based on a homogeneous regression model of an Artificial-neural-network (ANN). The input feature space is comprised of power, ASE noise, NLI, span numbers, and total distance. Conventionally, optimizing the ANN model parameters by minimizing the Mean-square-error (MSE) is not difficult but the case is different for transfer learning since only a few datasets from the sister network are available; the structure of the previously trained model may affect the performance of the transfer learning agent. Therefore, we perform extensive simulations to determine the appropriate settings for the ANN model.

TABLE III: Statistics of GSNR margin

Paths	Without Transfer learning				Transfer learning
	Nominal GSNR mean (dB)	**+ve Error** $\Delta\mathrm{GSNR} > 0$ **(dB)**	**-ve Error** $\Delta\mathrm{GSNR} < 0$ **(dB)**	**WCS Margin** $GSNR_{max} - GSNR_{min}$ **(dB)**	**Prediction Error MSE (dB)**
Birmingham → Bismarck	11.17	0.58	-0.6	1.20	0.079
Bismarck → Boston	10.53	0.58	-0.5	1.07	0.079
Boston → Buffalo	15.84	0.38	-0.53	1.0	0.085
Charlotte → Chicago	13.49	0.54	-0.72	1.26	0.098
Cleveland → Columbus	22.14	0.51	-0.94	1.45	0.083
Dallas → Denver	14.19	0.6	-0.9	1.49	0.083
Detroit → ElPaso	10.48	0.6	-0.54	1.13	0.083
ElPaso → Fresno	14.27	0.63	-0.96	1.58	0.065
Greensboro → Hartford	14.45	0.42	-0.56	1.0	0.094

The proposed ANN model consists of an input layer, two hidden layers, both with the same number of neurons (96), a dropout layer, and an output layer. For all neurons, a ReLU-based activation function is used to overlook vanishing gradient problems. Adaptive Moment Estimation (Adam) and MSE (Eq. 2) are used as an optimizer and loss function, respectively.

$$MSE = \frac{1}{n}\sum_{i}^{n}\left(\Delta\mathrm{GSNR}_i^{\mathrm{predicted}} - \Delta\mathrm{GSNR}_i^{\mathrm{actual}}\right)^2 \quad (2)$$

To avoid overfitting, a dropout layer is added at a rate of 0.20, dropping 20% of the random neurons to eliminate co-adaptive learning in each iteration. The model ANN is trained with 500 epochs. To determine the optimal number of epochs, an early stopping approach is used. The features data of 80 channels and the $\Delta\mathrm{GSNR}$ of the target channel are fed into the input layer. The transfer learning agent is trained on the EU network (C-band) data of 12000 samples. For validation purposes, the agent uses 6000 samples in order to estimate the QoT ($\Delta\mathrm{GSNR}$) error in the channel, and for testing purposes, the remaining 6000 samples are used. Transfer learning depends on the similarity of the marginal distribution probability in two different networks and the variety of samples available for retraining.

After obtaining the well-trained and tested model from the C-band (EU network) dataset this model is then used to run the transfer learning system with the small additional dataset obtained from the extended C-band network (US network) to adjust the weights of the hidden layers. To achieve a more reliable transfer learning performance, we randomly selected the re-training samples. As the number of input features is smaller in the extended C-band network (64 channels of 75 GHz) than the C-band network (80 channels of 50 GHz), a new input layer is added. We adjusted the weights of the hidden layers of the C-band network and held the current knowledge from the C-band network to the remaining layers of the model. Once the accuracy of the model predictions is achieved, the trained transfer learning module can be used together with the core QoT estimator engine to improve the accuracy in GSNR estimation of the LP for its deployment in an extended C-band network.

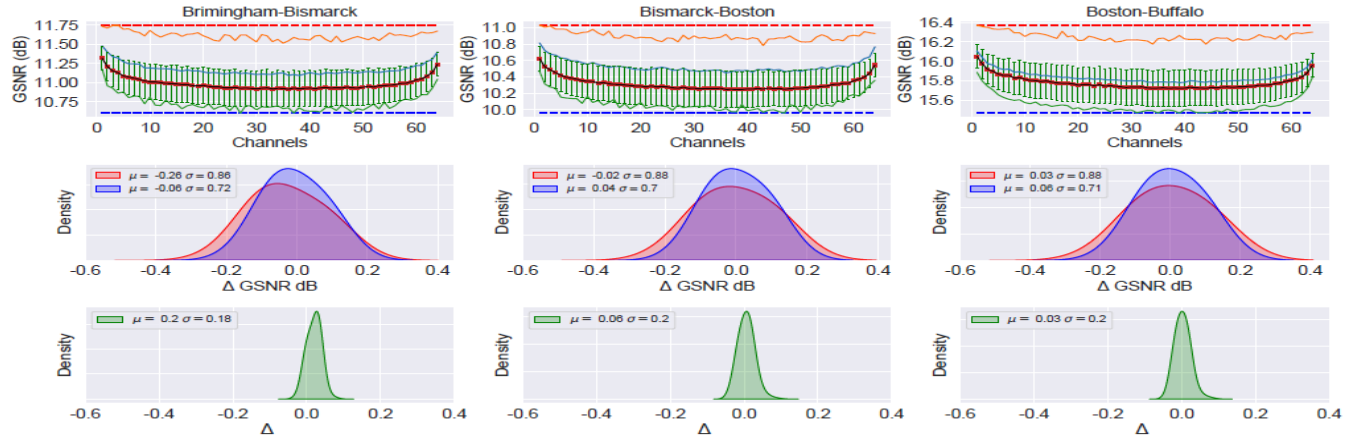
IV. Results and Discussion

During the initial deployment of the (USA network - extended C-band), the network controller relies only on the nominal description of the system parameters (see Table II) to estimate the GSNR. This estimated nominal GSNR is subjected to some degree of uncertainty due to the variation in the operating points of the NEs. This section describes results related to the transfer learning agent to cater to this induced uncertainty. In Fig. 2, the top three plots of each set show the GSNR statistics of the three paths of the newly established (USA network - extended C-band) network. In this figure, the GSNR statistics for all 64 channels and all samples of all USA test paths, i.e., Birmingham → Bismarck, Bismarck → Boston, and Boston → Buffalo from the first set, Charlotte → Chicago, Cleveland → Columbus, and Dallas → Denver from the second set, Detroit → ElPaso, ElPaso → Fresno, and Greensboro → Hartford from the third set are plotted in the frequency domain.

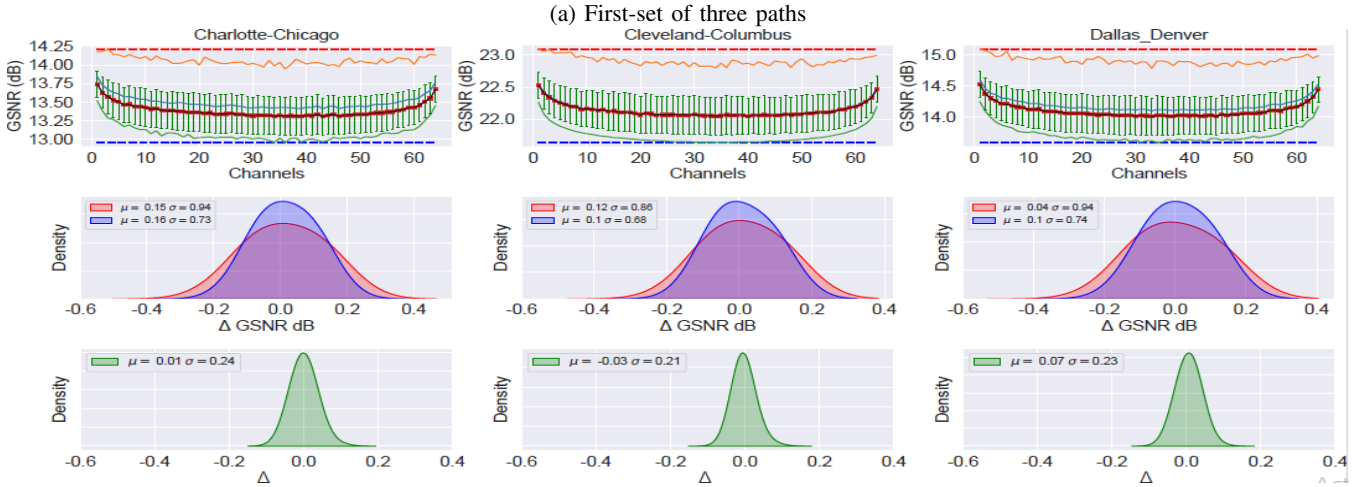
TABLE IV: Traditional C-band (EU network) Paths used for transfer learning agent cognition

Paths	Number of Spans
Amsterdam → Berlin	8
Brussels → Bucharest	30
Frankfurt → Istanbul	34
Vienna → Warsaw	7
London → Madrid	19
Paris → Rome	34

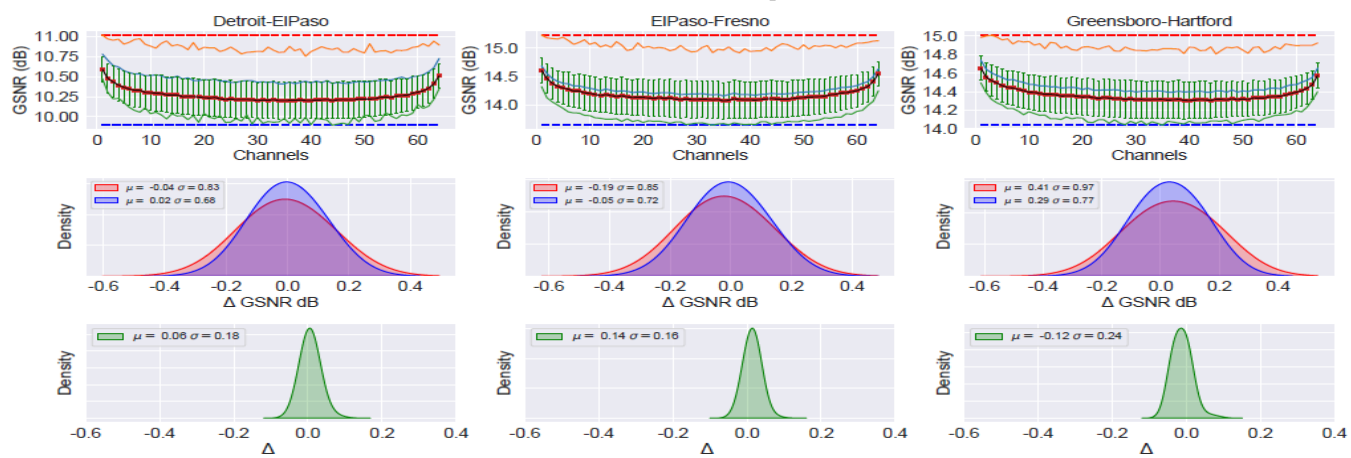
The exact ranges of variation are shown in the top three plots of each set in Fig. 2, where the dashed red (top) and dashed blue (bottom) lines indicate the overall maximum and minimum GSNR ($GSNR_{max}$ and $GSNR_{min}$), respectively. The red dots in the curve shape represent the mean (μ), and the error bars show the standard deviation (σ) of the total realizations for each channel. The black curve approximately equal to the mean values represents the nominal GSNR values for each path. The upper orange line (along with $GSNR_{max}$) and lower green line (along with $GSNR_{min}$) represent the maximum and minimum GSNR of each channel. Following the GSNR variations in Fig. 2, it can be observed that the actual GSNR values vary along with the nominal GSNR values. The uncertainties experienced by the system in terms of GSNR is



(a) First-set of three paths



(b) Second-set of three paths



(c) Third-set of three paths

Fig. 2: GSNR statistics (I^{st}row of each set), ΔGSNR actual/predicted distribution (2^{nd}row of each set), Prediction error (Δ) (3^{rd}row of each set): USA Network Paths

given by Eq. 3:

$$\Delta\mathrm{GSNR} = \mathrm{GSNR}^{\mathrm{nominal}} - \mathrm{GSNR}^{\mathrm{actual}} \tag{3}$$

Concerning Eq. 3 two possible cases of $\Delta\mathrm{GSNR}$ arise; $\Delta\mathrm{GSNR} > 0$, the actual GSNR value is less than the estimated GSNR value, and the reliance on the QoT-E leads to the undesirable OOS. Another is $\Delta\mathrm{GSNR} < 0$ where the nominal GSNR value is less than the actual GSNR value leading to the underutilization of the available capacity. Both the cases are reported for all nine paths of the deployed (extended C-band) considering the mean of nominal GSNR in Table. III. Along with this, during the Worst Case Scenario (WCS) when the operators do not have any specific knowledge of physical parameters, a total margin required for all nine paths of the deployed (extended C-band) is also reported.

A transfer learning agent is applied to predict GSNR accurately and minimize the overall uncertainties in the GSNR estimation of an extended C-band network. The agent retrieved data of six paths of traditional C-band (EU network) and is trained on four paths with 12000 (3000 samples/path) training samples, validated on two paths with 6000 samples (see Table IV). The agent's performance and scalability are evaluated on nine extended C-band network test paths (3000 samples/path). The transfer learning agent's performance on the nine paths of the USA network is shown as a distribution plot in Fig. 2 (second row of each set), where $\Delta GSNR^{actual}$ is shown as the blue curve and $\Delta GSNR^{predicted}$ is shown as the red curve. By noticing the μ and σ values in Fig. 2 (second row), we observe that the transfer learning agent performs well by reducing the uncertainty in GSNR estimation; the values of MSE are listed in Table. III for each path. The total error in GSNR estimation of the extended C-band network is drastically reduced against each path using a transfer learning agent (see Table. III) .

$$\Delta = \Delta\mathrm{GSNR}^{\mathrm{actual}} - \Delta\mathrm{GSNR}^{\mathrm{predicted}} \tag{4}$$

Furthermore, the error in predicting the error in GSNR by transfer learning; Δ between the $\Delta GSNR^{actual}$ and the $\Delta GSNR^{predicted}$ (see Eq. 4) is also shown in the third row of each set of Fig. 2. The μ and σ value of Δ shows excellent performance of the transfer learning agent in correcting the GSNR values of the extended C-band network using the C-band network dataset.

V. Conclusion

In this work, we proposed the use of a transfer learning agent trained on a dataset of a traditional C-band network (EU network) to correct GSNR estimation in an extended C-band network (USA network). We created synthetic datasets for both networks using the open-source GNPy library for training and testing purposes.

It is evident by observing the obtained promising results that the proposed transfer learning agent can reduce the uncertainties in the estimation of LP QoT in the extended C-band network by exploiting the knowledge of the traditional C-band operational network. The proposed transfer learning agent API works synergistically with the QoT estimator engine of extended C-band to assist it in correcting the GSNR estimation in the software-defined optical networks.

References

[1] Cisco, "Cisco annual internet report (2018–2023) white paper," (2020).

[2] V. Curri, "Software-defined wdm optical transport in disaggregated open optical networks," in *2020 22nd International Conference on Transparent Optical Networks (ICTON),* (IEEE, 2020), pp. 1–4.

[3] A. D'Amico, S. Straullu, A. Nespola, I. Khan, E. London, E. Virgillito, S. Piciaccia, A. Tanzi, G. Galimberti, and V. Curri, "Using machine learning in an open optical line system controller," Journal of Optical Communications and Networking **12**, C1–C11 (2020).

[4] T. Jiménez, J. C. Aguado, I. de Miguel, R. J. Durán, M. Angelou, N. Merayo, P. Fernández, R. M. Lorenzo, I. Tomkos, and E. J. Abril, "A cognitive quality of transmission estimator for core optical networks," Journal of Lightwave Technology **31**, 942–951 (2013).

[5] E. Seve, J. Pesic, C. Delezoide, S. Bigo, and Y. Pointurier, "Learning process for reducing uncertainties on network parameters and design margins," Journal of Optical Communications and Networking **10**, A298–A306 (2018).

[6] I. Khan, M. Bilal, M. Siddiqui, M. Khan, A. Ahmad, M. Shahzad, and V. Curri, "Qot estimation for light-path provisioning in un-seen optical networks using machine learning," in *International Conference on Transparent Optical Networks (ICTON),* (IEEE, 2020).

[7] I. Khan, M. Bilal, and V. Curri, "Advanced formulation of qot-estimation for un-established lightpaths using cross-train machine learning methods," in *International Conference on Transparent Optical Networks (ICTON),* (IEEE, 2020).

[8] I. Khan, M. Bilal, and V. Curri, "Assessment of cross-train machine learning techniques for qot-estimation in agnostic optical networks," OSA Continuum **3**, 2690–2706 (2020).

[9] C. Rottondi, L. Barletta, A. Giusti, and M. Tornatore, "Machine-learning method for quality of transmission prediction of unestablished lightpaths," Journal of Optical Communications and Networking **10**, A286–A297 (2018).

[10] R. Di Marino, C. Rottondi, A. Giusti, and A. Bianco, "Assessment of domain adaptation approaches for qot estimation in optical networks," in *2020 Optical Fiber Communications Conference and Exhibition (OFC),* (IEEE, 2020), pp. 1–3.

[11] F. Usmani, I. Khan, M. U. Masood, A. Ahmad, M. Shahzad, and V. Curri, "Convolutional neural network for quality of transmission prediction of unestablished lightpaths," Microwave and Optical Technology Letters **63**, 2461–2469 (2021).

[12] D. Azzimonti, C. Rottondi, A. Giusti, M. Tornatore, and A. Bianco, "Active vs transfer learning approaches for qot estimation with small training datasets," in *Optical Fiber Communication Conference,* (Optical Society of America, 2020), pp. M4E–1.

[13] A. Ahmad, A. Bianco, E. Bonetto, M. Garrich, and J. R. Oliveira, "Switching node architectures in flexible-grid networks: A performance comparison," in *2014 International Conference on Optical Network Design and Modeling,* (2014), pp. 49–54.

[14] https://www.itu.int/rec/T-REC-G.694.1/en.

[15] I. Khan, M. Bilal, M. Umar Masood, A. D'Amico, and V. Curri, "Light-path qot computation in optical networks assisted by transfer learning," Journal of Optical Communications and Networking **13** (2021).

[16] A. Ferrari, M. Filer, K. Balasubramanian, Y. Yin, E. Le Rouzic, J. Kundrat, G. Grammel, G. Galimberti, and V. Curri, "Gnpy: an open source application for physical layer aware open optical networks," Journal of Optical Communications and Networking **12** (2020).