A Three-stage Training Framework for Customizing Link Models for Optical Networks

Xiaomin Liu, Huazhi Lun, Mengfan Fu, Yunyun Fan, Lilin Yi, Weisheng Hu, and Qunbi Zhuge*
State Key Laboratory of Advanced Optical Communication Systems and Networks, Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai, 200240, China
**Corresponding author: qunbi.zhuge@sjtu.edu.cn*

Abstract: We propose a link model customization framework to increase modeling accuracy for each specific link in an optical network. In addition, an active acquisition method is employed in this framework to improve tolerance to link parameter uncertainties.
OCIS codes: (060.2330) Fiber optics communications, (060.1155) All optical networks.

1. Introduction

Quality of transmission (QoT) estimation [1] and monitoring [2] are the key functions to build low-margin and intelligent optical networks for higher capacity and efficiency. Analytical approaches in modeling link performance often suffer from insufficient accuracy and/or high computation complexity in the context of diverse optical link conditions [3, 4]. More recently, machine learning (ML) has received increasing attentions since it provides new opportunities to build *autonomous optical networks* [5,6]. Although ML models can be trained with large datasets obtained by simulations, the training dataset might not cover all real deployment situations, which can result in large deviations in some cases. In [7], authors propose a learning strategy to augment training datasets using the data measured from established or probe lightpaths. However, a single model is employed for all links of a network, which compromises the accuracy of each link in order to achieve an overall optimal performance. Also, to retrain a ML model for a specific channel impairment based on real-time measurement, extracting the real value of the target impairment can be difficult. In addition, the uncertainty of link parameters will make the measurements in real scenes less accurate, which may worsen the retraining performance.

In this paper, we propose a three-stage training framework to build a customized model for each link to further improve modeling accuracy in a self-learning manner. In addition, an active acquisition method is demonstrated to acquire real values of linear or nonlinear impairments. In our framework, an initial model is first trained by a large dataset generated by simulations. Next, this model is retrained with a small dataset from simulations with the configurations of a target link. Finally, the proposed active acquisition method is applied to acquire real values using a probe channel in order to further retrain the model. We evaluate the framework in the modeling of nonlinear impairment, which is regarded as one of the most challenging link models [3, 4]. Based on extensive simulations, we demonstrate that the accuracy improves by 2dB for each target link and the root mean square error (RMSE) reduces by more than 50% compared with the previous modeling method with ML [8].

2. Operation principle

The concept of the model customization proposed in this work aims to maximize the accuracy of each specific link by avoiding the compromise of performance across links. In our framework, a uniform ML algorithm such as an artificial neural network with a fixed structure is employed for all links and the set of ML model parameters are individually trained for each link. In this case, the model customization only needs to add additional parameters, which will not induce a significant increase in controller complexity. In this section, we will describe the three-stage training framework and the embedded active acquisition method. Fiber nonlinearity modeling is used as an example in demonstrating the efficiency of this framework.

2.1 Three-stage training framework for model customization

The proposed framework for model customization is illustrated in Fig. 1, and the three stages are described as follows.

First stage: model initialization. The model is first trained with a large dataset, of which samples consist of heterogeneous link/channel configurations. This stage will provide an initialized model for the following customization procedures. Since the initial model is the same for all links, this stage only needs to be conducted once. The obtained model can provide coarse results. However, the results in some links especially for those not covered in the training dataset may have large deviations. If link parameter uncertainties exist, the performance of the initial model will be further degraded.

Second stage: offline customization. At this stage, the initial model is retrained with a much smaller dataset from simulations, where only diverse channel configurations are considered for a specific link configuration. In this process, the model will be retrained to learn more about the particularities of this link to significantly improve the predicting precision. Since the model is retrained based on a pre-trained one, the training time will be short.

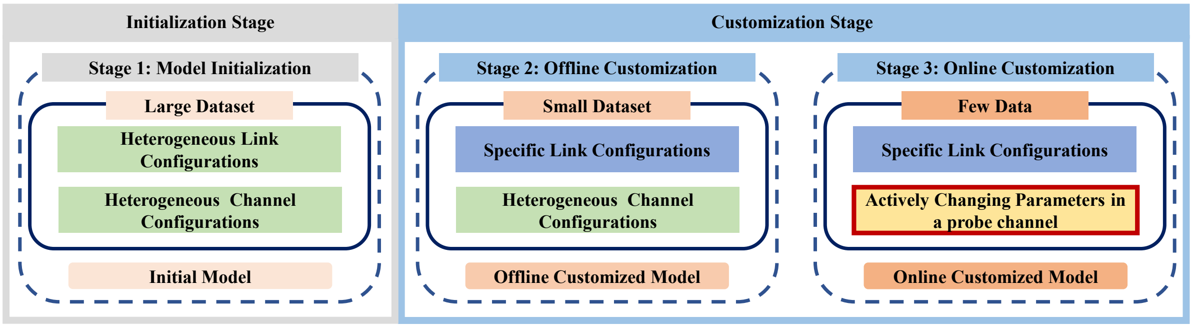

Fig. 1. The framework of the three-stage customization procedure for a specific link.

Third stage: online customization. Finally, an active acquisition method is employed at the beginning of the link establishment to obtain the real values of the target impairment. Then these data from the real scenes is used to further retrain the customized model from the previous stage. This stage is efficient in reducing the impact of link parameter uncertainties, which can be divided into two types: systematic deviations and random fluctuations. The random fluctuations can be easily addressed by making more measurements. The systematic deviations can be addressed in the retraining process in this stage. The active acquisition is a key for this stage, and it will be detailed as follows.

2.2 Active acquisition method

For the active acquisition, we propose to actively change the configurations of a probe channel before link establishment, in order to obtain real values of certain impairments. This method can be applied to various scenarios. For instance, the real nonlinearity and filtering effects can be measured by changing the launch power and the bandwidth of the probe signal, respectively. Since we focus on the nonlinear modeling in this work, we take the measurement of nonlinear noise-to-signal ratio (NSR) as an example. It is known that in long-haul transmissions, the nonlinear (NLI) noise is mixed with amplifier spontaneous emission (ASE) noise and back-to-back (B2B) transceiver noise. The total NSR can be calculated as $NSR_{total} = NSR_{ASE} + NSR_{NLI} + NSR_{B2B}$. Assuming the B2B noise has been measured for each transceiver during manufacturing, according to the Gaussian Noise (GN) model [3, 4], the total power (P_{total}) of the nonlinear noise (P_{NLI}) and ASE noise (P_{ASE}) can be calculated by

$$P_{total} = P_{NLI} + P_{ASE} = P_{SPM} + P_{XPM} + P_{ASE} = \eta_1 P_{CUT}^3 + \eta_2 P_{INT1} P_{INT2} P_{CUT} + P_{ASE} \tag{1}$$

where η_1 and η_2 are parameters related to link configurations. A series of signals with different input powers (P_{CUT}) are launched in the probe channel for many times to reduce the influence of random fluctuations when the launch power of other established channels remains unchanged. With the launch power of the nth interfering channel (P_{INTn}) and the channel under test (P_{CUT}), the real P_{NLI} can be obtained by fitting the curve of the P_{total} and P_{CUT}. As a result, the SNR of real nonlinear noise is derived for the third-stage customization training.

3. Simulation and discussion

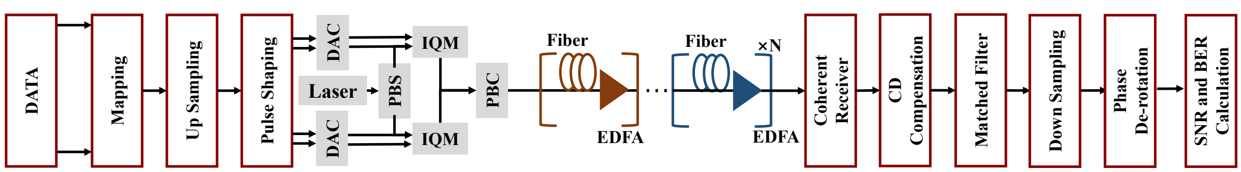

Fig. 2. Simulation setup.

The simulation setup is shown in Fig. 2. A fix-rate system is evaluated with a symbol rate of 35Gbaud and a channel spacing of 50GHz. The symbol length is 2^{16}. At the transmitter (Tx) side, root-raised-cosine (RRC) pulse shaping with a roll-off of 0.02 is applied. For the fiber link, lumped Erbium-doped fiber amplifiers (EDFA) are employed and the noise figure is 5dB. The fiber nonlinearity is simulated by the split-step Fourier method. The receiver (Rx) filters out the center channel for the following signal processing, which includes chromatic dispersion compensation (CDC), matched filtering, down-sampling and phase de-rotation. Since laser phase noise is ignored, only the overall phase rotation caused by fiber nonlinearities is removed. Finally, the SNR is calculated.

Table 1. Summary of link/channel configurations

Span fiber types	Span length	Span number	Launch power	Modulation format	Channel number
SSMF, PSCF, ELEAF	80 km	3:1:17	-3:1:3 dBm	QPSK/16QAM	3:2:17

We choose three links as examples to evaluate our framework. The first and second links have six and twelve spans of standard single mode fiber (SSMF), respectively. The third link has 11 spans, each randomly configured with the three types of fiber in Table 1. For the first-stage training, we simulate 500 samples with heterogeneous link and channel configurations randomly chosen from Table 1 to build the training dataset. The randomness follows a uniform distribution. At the second stage, for each link, we randomly choose 100 channel configurations to build the training set. The center channel is considered and tested. For the active acquisition, the launch power of the center channel is changed from -5dBm to 5dBm to derive the nonlinear SNR. An optical power uncertainty ϵ which follows a normal distribution of $\epsilon \sim (0.5, 0.3)$dB is assumed. The SNR of each input power is measured for ten times to suppress random fluctuations. For each link, 100 samples with different channel configurations and power uncertainties are randomly generated for testing. A two-layer neural network is built where sigmoid and linear functions are applied as the first-layer and second-layer activation function, respectively. The optimizer function is root mean square prop (rmsprop) and the evaluation metric is mean absolute error (MAE). For each stage, 20% samples are used for validation. The training performance is compared with the method proposed in [8] which provides an initial model for all links.

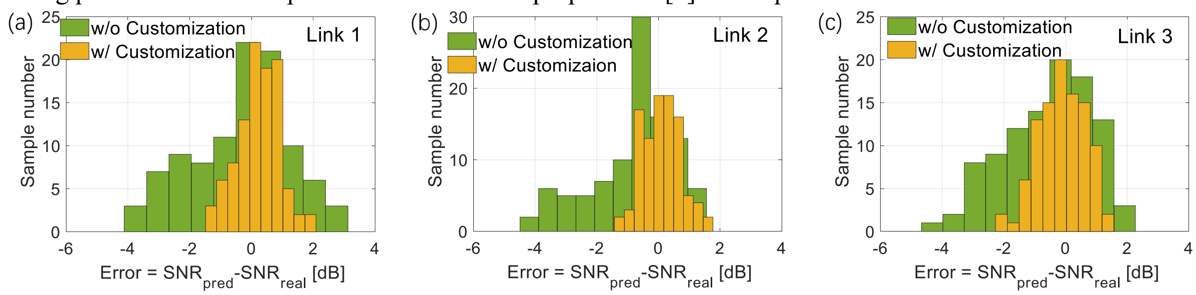

Fig. 3. The comparison of the performance of models with and without the customization stages.

Fig. 3(a), (b) and (c) show the error histograms of the initial model and the customized model for Link 1, 2 and 3, respectively. Results show that the initial model performs differently for each link. After retrained with the customizing framework, the maximum error reduces from 4.3dB to 2dB, illustrating that the three-stage framework effectively mitigates the influence of the power deviation and fluctuation. The training of the model requires ~6000 and ~1000 epochs for the first and second stages to converge, respectively. For the third stage, the model goes through the finetune processing for 10 epochs to avoid overfitting. Therefore, the second and third stages are time-saving for the model customization.

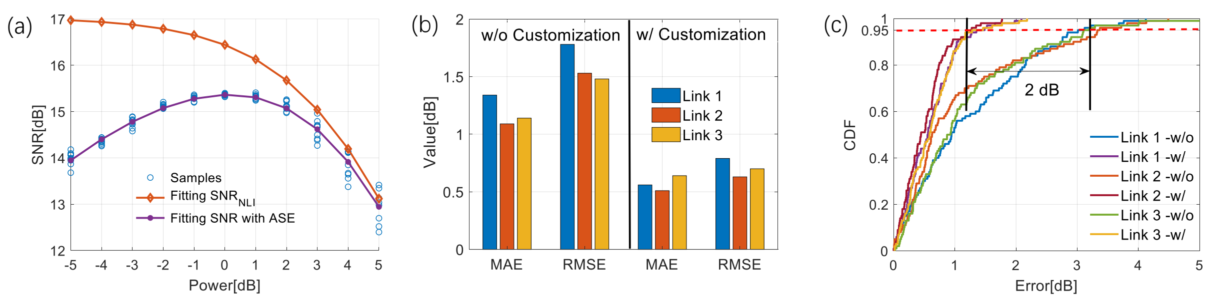

Fig. 4. (a) Fitting results for active acquisition. (b) The comparison of the MAE and root mean square error (RMSE). (c) The cumulative probability of prediction deviation.

To clarify the details of the customization process, Fig. 4(a) shows the fitting procedure of Link 3 in the active acquisition procedure, which is shown to be tolerant to parameter fluctuations. Fig. 4(b) shows the comparison of the MAE and RMSE before and after training with the customization stages, and the RMSE and MAE are significantly reduced. Fig. 4(c) shows that for each link, the accuracy improves by 2dB at 95% cumulative probability. For about 90% of the samples, errors are less than 1dB for all the three links after the customization stages.

4. Conclusion

We propose a framework to customize the model for a specific link in optical networks. The three-stage training framework embedded with an active acquisition method improves modeling accuracy for each link and increases the tolerance to parameter uncertainties, which are verified by extensive simulations.

This work was supported by NSFC (61801291), Shanghai Rising-Star Program (19QA1404600) and National Key R&D Program of China (2018YFB1801203).

[1] I. Sartzetakis *et al*, J. Opt. Commun. Netw. 11, 140-150 (2019)
[2] Z. Dong *et al*, J. Lightw. Technol. 34, 525-543 (2016)
[3] P. Poggiolini *et al*, J. Lightw. Technol. 32, 694-721 (2014)
[4] A. Carena *et al*, Opt. Express 22, 16335-16362 (2014)
[5] D. R *et al*, J. Opt. Commun. Netw, 2018, 10(10)
[6] Gifre Lluís *et al*, J. Opt. Commun. Netw, 10.5(2018)
[7] M. Ruiz *et al*, in OFC (2019), paper M2A.4.
[8] Q. Zhuge *et al*, J. Lightw. Technol. 37, 3055-3063 (2019)