POLITECNICO DI TORINO
Repository ISTITUZIONALE

Assessment of Domain Adaptation Approaches for QoT Estimation in Optical Networks

Original
Assessment of Domain Adaptation Approaches for QoT Estimation in Optical Networks / Di Marino, Riccardo; Rottondi, Cristina; Giusti, Alessandro; Bianco, Andrea. - STAMPA. - (2020). (Intervento presentato al convegno OFC 2020 - The Optical Networking and Communication Conference & Exhibition tenutosi a San Diego, USA nel March 2020) [10.1364/OFC.2020.Th3D.2].

Availability:
This version is available at: 11583/2826023 since: 2020-05-17T15:59:08Z

Publisher:
OSA Publishing

Published
DOI:10.1364/OFC.2020.Th3D.2

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright
Optica Publishing Group (formely OSA) postprint/Author's Accepted Manuscript

"© 2020 Optica Publishing Group. One print or electronic copy may be made for personal use only. Systematic reproduction and distribution, duplication of any material in this paper for a fee or for commercial purposes, or modifications of the content of this paper are prohibited."

(Article begins on next page)

12 October 2024

Assessment of Domain Adaptation Approaches for QoT Estimation in Optical Networks

Riccardo Di Marino[1], **Cristina Rottondi**[1], **Alessandro Giusti**[2] **and Andrea Bianco**[1]

[1]*Dept. of Electronics and Telecommunications, Politecnico di Torino, Italy*
[2]*Dalle Molle Institute for Artificial Intelligence, Lugano, Switzerland*

{*cristina.rottondi,andrea.bianco*}*@polito.it, alessandrog@idsia.ch*

Abstract: We evaluate the performance of two domain adaptation approaches for machine learning assisted quality of transmission estimation of an optical lightpath, for a fixed/variable number of available training samples from the source/target domain. © 2020 The Author(s)

OCIS codes: 060.4250 Fiber optics and optical communications, Networks; 060.4510 Fiber optics and optical communications, Optical communications

1. Introduction

Predicting the Quality of Transmission (QoT) of a candidate lightpath prior to its establishment plays a pivotal role for an effective design and management of optical networks. In the last few years, Machine Learning (ML) techniques for QoT estimation [1] have been advocated as promising alternatives to *i*) approximated mathematical models or *ii*) simulation frameworks that model the propagation of the optical signal along the fiber core, the former often introducing high margins to conservatively compensate for simplifying assumptions and/or for uncertainties in input parameters values, the latter typically requiring prohibitive computational effort when applied to real-scale scenarios. Supervised ML-based approaches learn a mapping from a set of *features*, e.g., characteristics of a lightpath such as length, amount of served traffic and adopted modulation format to a *target variable*, e.g., an indicator of the expected QoT in the lightpath, such as the Bit Error Rate (BER). To make the learning phase effective, a large amount of samples (training set) must be provided to the learning algorithm: Each sample consists of the features of an already established lightpath associated with the actual value of the target variable (*ground truth*), which can be measured at the receiver [2].

On one hand, supervised ML methods assume that the training set is large and *representative* of the samples that will be processed when the model is exploited to predict the QoT. On the other hand, the collection of training samples is often hindered by practical issues (e.g., lack of dedicated telemetry equipment in every network node) or is too costly to permit the acquisition of large datasets. This is especially true for networks in the early stage of deployment, where the number of already installed (and thus, monitorable) lightpaths is very limited. However, it is sometimes possible to rely on large training datasets from a different network than the one on which the ML model operates. Therefore, we assume that a large amount of training data **S** is given for a *source domain* (e.g., a backbone network monitored for a long operational period), and is used to train a model that predicts the QoT of lightpaths to be established in a different *target domain* (e.g., a newly deployed network), for which a small labeled training dataset **T** is available [3–5]. In such a scenario, we wish to exploit at best the data from the source domain to tailor a good model to the target domain. This approach is known in ML research as *Domain Adaptation* (DA) and leverages the intuition that samples from one domain provide useful information concerning the solution of the QoT estimation problem in the other domain.

To evaluate the applicability of existing DA approaches for ML-based QoT estimation of candidate lightpaths, we focus on two networks characterized by different topologies, but adopting the same fiber type and transmission equipment. We assess the performance of two DA techniques depending on the number of available training instances from the target domain. We describe the applied methods in Sec. 2 and discuss their performance in Sec .3, drawing some conclusions.

2. Domain Adaptation Approaches for QoT Estimation

We adopt the ML framework for QoT classification proposed in [6], which uses the following features to characterize a lightpath: total length, number of traversed links, maximum link length, amount of traffic to be transmitted and modulation format to be adopted for transmission. Given a candidate lightpath, the classifier provides as output a binary variable whose value depends on whether the predicted probability that the lightpath configuration satisfies a given threshold γ on the BER measured at the receiver is above or below 50%. In this study we opt

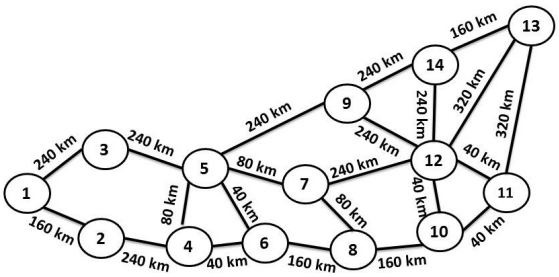


Fig. 1: Japan network topology

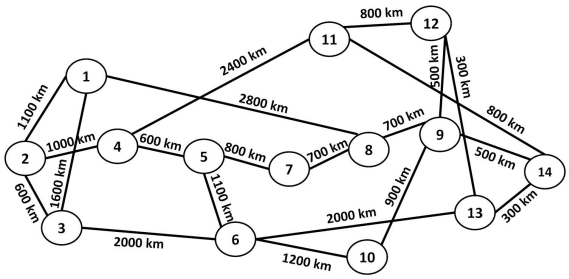


Fig. 2: NSF network topology

for a Support Vector Machine (SVM) learning model, which proved to provide better performance than Random Forests (adopted in [6]) when applying DA techniques in our context. In particular, we compare three baselines and two DA approaches.

Only Source Domain Baseline (SDB) trains the model only on $\mathbf{S}$.

Only Target Domain Baseline (TDB) trains the model only on $\mathbf{T}$.

Dataset Mixing Baseline (DMB) consists in training the classifier on $\mathbf{S} \cup \mathbf{T}$, in a standard supervised fashion.

Feature Augmentation (FA) implements a simple approach [7] which encodes the domain of a sample by augmenting its feature vector. In particular, we triple the length of the original feature vector $\mathbf{x}$ with a rule that depends on the domain: If the sample belongs to $\mathbf{S}$, the resulting feature vector is computed as $\mathbf{x}' = \langle \mathbf{x}, \mathbf{x}, \mathbf{0} \rangle$; otherwise, for a sample in $\mathbf{T}$, $\mathbf{x}' = \langle \mathbf{x}, \mathbf{0}, \mathbf{x} \rangle$. This augmentation transformation is applied to all samples, both in the training and test phases. It is expected that such transformation allows a classifier to learn and exploit both the commonalities between the two domains and the unique characteristics of each domain [7].

Correlation Alignment (CORAL) [8] is an unsupervised DA technique that minimizes domain shift by aligning the second-order statistics of source and target datasets. This is done by applying a transformation ϕ that re-colors whitened source features of the source domain with the covariance of the distribution of the dataset gathered from the target domain. In this case, we assume that a large amount $\mathbf{T}_{\text{unlabeled}}$ of unlabeled data (i.e., samples for which the associated BER is not known) from the target domain is available in addition to $\mathbf{T}$, and we implement the following steps: 1) estimate the transformation ϕ from the source to the target feature spaces using $\mathbf{S}$ and $\mathbf{T}_{\text{unlabeled}}$; 2) train a classifier on $\phi(\mathbf{S}) \cup \mathbf{T}$.

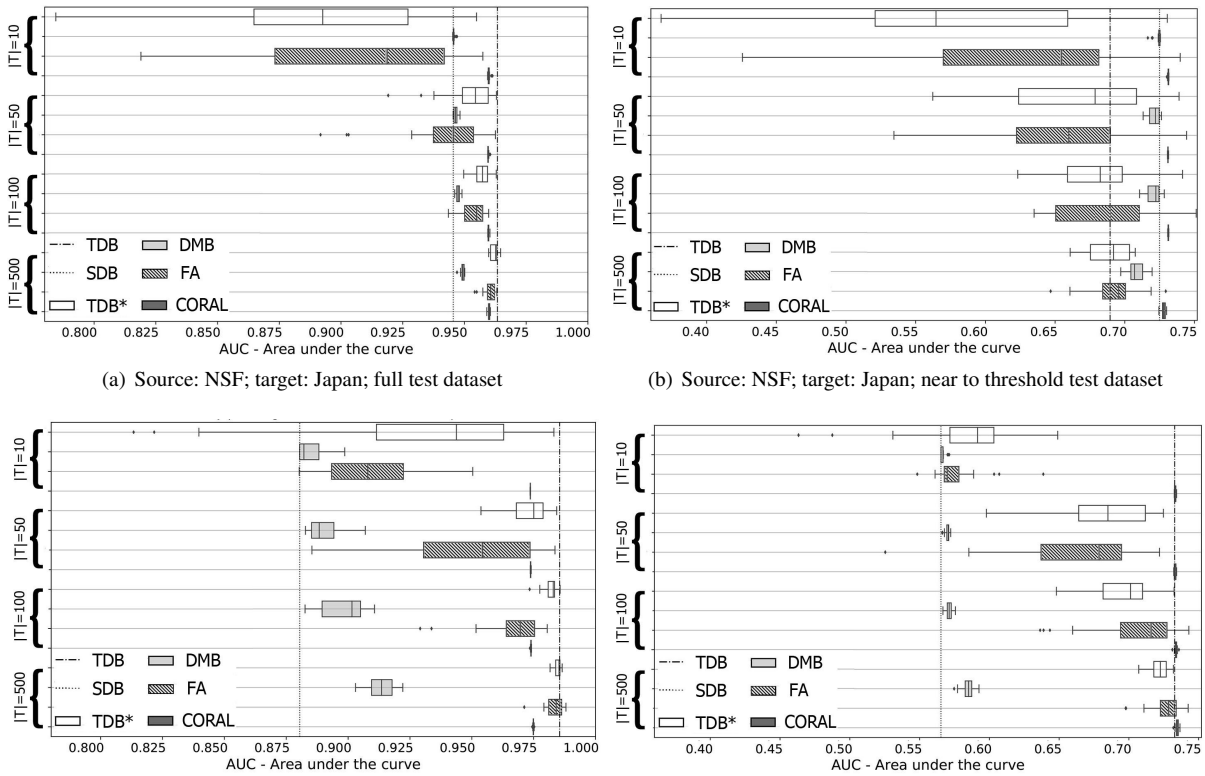
3. Results

We consider the Japan and NSF networks depicted in Figs. 1 and 2 and evaluate the performance of the DA approaches presented in Section 2 in terms of Area Under the ROC Curve (AUC).

Dataset generation. To generate synthetic BER measurements we use the QTool presented in [6]. Given a candidate lightpath, traffic volume and modulation format, the QTool calculates the BER as a function of the signal-to-noise ratio measured at the channel decoder via the approximated AWGN model of dispersion uncompensated transmission over single mode fibers. The Qtool also adds randomly-distributed penalties to account for model uncertainties. We assume a flexi-grid scenario with 12.5 GHz slice width and transceivers operating at 28 Gbaud with 37.5 GHz optical bandwidth, using a modulation format chosen among dual polarization (DP)-BPSK, QPSK and n-QAM, with $n = 8, 16, 32, 64$. Traffic demands exceeding the capacity of a transceiver are served by superchannels containing multiple adjacent transceivers. To construct the training dataset $\mathscr{R}_{\text{source}}$ for each topology, we produce 10000 instances by randomly choosing a source-destination node pair, a modulation format and a traffic demand uniformly selected in the range $[50-500]$ Gbps with 50 Gbps granularity and calculating the BER with the Qtool. We set the BER threshold to $\gamma = 4 \cdot 10^{-3}$. The test dataset is constructed for both topologies by generating a separate set $\mathscr{E}_{\text{target}}$ of 90000 instances, with the same procedure used to produce the training dataset.

Considered Scenarios The two DA methods (i.e., FA and CORAL) are benchmarked against the SDB, TDB and DMB baselines, where SDB assumes $|\mathbf{S}| = 10000$ and TDB assumes $|\mathbf{T}| = 10000$. In the DMB baseline and in the two DA approaches, the training phase assumes $|\mathbf{S}| = 10000$ (where $\mathbf{S} = \mathscr{R}_{\text{source}}$) and $|\mathbf{T}| = 10, 50, 100, 500$ (where $\mathbf{T} \subset \mathscr{R}_{\text{target}}$), whereas, for CORAL only, we assume $\mathbf{T}_{\text{unlabeled}} = 10000$ (where $\mathbf{T}_{\text{unlabeled}}$ contains the feature vectors of the elements in $\mathscr{R}_{\text{target}}$). Note that, for the QoT estimation task, collecting unlabeled samples of the set $\mathbf{T}_{\text{unlabeled}}$ from the target domain is trivial, as we simply need to select route, traffic volume and modulation format of a perspective ligthpath to derive its feature vector, but we do not need to measure its BER. Moreover, we report an additional baseline TDB* that trains the SVM classifier using only the samples of the target domain leveraged by the DA approaches (i.e., $|\mathbf{T}| = 10, 50, 100, 500$, where $\mathbf{T} \subset \mathscr{R}_{\text{target}}$). The trained classifiers are tested using the whole $\mathscr{E}_{\text{target}}$ set. Experiments are repeated 20 times, with random extraction of the elements of $\mathbf{T}$ from set $\mathscr{R}_{\text{target}}$.

Numerical assessment. Fig. 3(a) plots the AUC obtained when the NSF topology is used as source domain and the Japan topology as target domain. Results show that, when $|\mathbf{T}|$ is low, CORAL significantly improves the



(a) Source: NSF; target: Japan; full test dataset

(b) Source: NSF; target: Japan; near to threshold test dataset

(c) Source: Japan; target: NSF; full test dataset

(d) Source: Japan; target: NSF; near to threshold test dataset

Fig. 3: Performance comparison of the considered DA approaches.

AUC w.r.t. the SDB, TDB* and DMB baselines, while FA achieves on average lower AUC than DMB and shows comparable performance to TDB*. Conversely, when $|\mathbf{T}| = 500$, TDB*, FA and CORAL provide similar results, which closely approach TDB: Indeed, when the number of available samples from the target domain is quite large, DA techniques are expected to be less useful, as those samples are already sufficiently representative of the whole feature space.

In Fig. 3(b) we consider the same scenarios of Fig. 3(a), but we restrict the test set to a subset of $\mathscr{E}_{\text{Japan}}$ of 6000 elements exibiting a BER within the range $[10^{-5}, 10^{-2}]$, i.e., close to the system threshold γ. Such samples are expected to be more difficult to classify. Though the AUC is generally lower, the considered DA approaches show AUC trends similar to those reported in Fig. 3(a), with CORAL even exceeding the TDB baseline.

In Figs. 3(c)-(d), we report results obtained using the Japan topology as source domain and the NSF topology as target domain. In this case, FA and CORAL always outperform SDB and FMB. In Fig. 3(c), for large $|\mathbf{T}|$, FA shows comparable performance to TDB* and more closely approaches TDB than CORAL. Note that, in Figs. 3(c)-(d), the gap between SDB and TDB is larger than in Figs. 3(a)-(b). Indeed, the average lightpath length in the Japan network is considerably lower than in the NSF network. Thus, little knowledge about the BER of very long lightpaths can be obtained through the samples of the Japan network, i.e., the source domain used to compute SDB.

Based on the above reported results, we conclude that DA approaches show considerable improvement in the AUC w.r.t. standard ML techniques, especially when the number of available samples from the target domain is very limited. Further research will be devoted to investigating how the degree of dissimilarity between source and target domains impacts the performance of DA.

References

1. F. Musumeci et al., "An overview on application of machine learning techniques in optical networks," *IEEE Comm. Surveys and Tutorials*, vol. 21, no. 2, pp. 1383–1408, 2018.
2. K. Christodoulopoulos et al., "ORCHESTRA-Optical performance monitoring enabling flexible networking," Proc. ICTON 2015.
3. W. Mo et al. "ANN-based transfer learning for QoT prediction in real-time mixed line-rate systems," Proc. OFC 2018.
4. Q. Yao, et al. "Spectrum optimization for resource reservation based on transductive transfer learning in space division multiplexing elastic optical networks" Proc. ECOC 2018.
5. L. Xia et al. "Transfer learning assisted deep neural network for OSNR estimation," *Optics Express* 27(14), pp. 19398–19406, 2019.
6. C. Rottondi et al., "Machine-learning method for quality of transmission prediction of unestablished lightpaths," *IEEE/OSA J. of Optical Comm. and Netw.*, vol. 10, no. 2, pp. A286–A297, Feb 2018.
7. H. Daumé III, "Frustratingly easy domain adaptation," arXiv preprint arXiv:0907.1815, 2009.
8. B. Sun et al., "Return of frustratingly easy domain adaptation,", Proc. AAAI Conf. on Artificial Intelligence 2016.