Privacy Preserving Digital Twin Knowledge Sharing for Multi-domain Networks

Marc Ruiz and Luis Velasco*

Optical Communications Group (GCO), Universitat Politècnica de Catalunya (UPC), Barcelona, Spain
e-mail: luis.velasco@upc.edu

Abstract: Knowledge sharing techniques among OCATA optical layer digital twin instances are proposed for multi-domain scenarios. Intra-domain model transformations are performed to guarantee privacy of intra-domain topology. Remarkable accuracy to estimate multi-domain lightpaths QoT is shown. © 2024 The Authors

1. Introduction

Optical layer digital twins have been proposed to facilitate network automation, as they can support applications from Quality of Transmission (QoT) estimation during optical connection (*lightpath*) provisioning to failure management after the lightpath has been setup [1]. Such digital twin applications generally focus on the operation of single domain networks, where the digital twin has fully network visibility. However, future 6G networks are envisioned to support a large number of services with stringent performance spanning multiple domains [2] and therefore, meeting such end-to-end (e2e) requirements will require from tight coordination among domains. To create e2e models, the different domains supporting an e2e lightpath can share models trained for the intra-domain network, as in [3]. However, distributing such intra-domain models is not secure, as they can include details of the intra-domain network, e.g., the number of hops, the distance of the optical links, or the configuration of optical devices. Note that such information could be of interest to craft specific attacks in case of eavesdropping [4]. Therefore, privacy of the internals of each domain must be enforced when intra-domain models leave the security perimeter of the domain when are shared.

In this paper, we assume the OCATA optical time domain digital twin [5]. OCATA relies on deep neural networks (DNN) to model the expected effects of optical devices (optical filters and amplifiers) and fibers on in-phase and quadrature (IQ) optical constellations. By concatenating DNNs for the elements in the intra-domain route of a lightpath, expected QoT, such as the pre-forward error correction (pre-FEC) bit error rate (BER), as well as other metrics, can be computed [1]. As intra-domain models are concatenations of DNNs, transformations are proposed to secure sharing intra-domain models used for the modelling of e2e multi-domain lightpaths.

2. E2e OCATA Modeling in Multi-Domain Scenarios

The proposed solution is to create *exportable* intra-domain DNN models that preserve privacy at the required level. Such models are created on-demand, e.g., every time a new inter-domain lightpath is provisioned. Exportable models are built from already trained ones and shared among domains supporting an e2e lightpath in order to build the e2e model. In this work, domain boundaries are defined by the network elements under the control of a single Software-Defined Networking (SDN) domain controller, e.g., an operator domain or vendor island. We assume that every domain includes an instance of OCATA that is fed with DNN models of the different transponders (TP), reconfigurable optical add-drop multiplexers (ROADM), and fiber links (referred to as *components*) in the route of a lightpath in the domain (intra-domain route or segment) [5].

Fig. 1 illustrates an e2e multi-domain network scenario with three domains (labeled *D1*, *D2*, and *D3*). Without loss of generality, we assume that the route between sites *A* and *Z* (in orange color in Fig. 1) represents either: *i*) a new multi-domain lighpath which e2e QoT needs to be evaluated during provisioning time; or *ii*) an already established multi-domain lightpath which e2e model needs to be built for computing expected QoT metrics. The main workflow for the e2e model generation is also included in Fig. 1. OCATA first obtains the intra-domain route segment *p* of the lightpath from the local SDN controller (1 in Fig. 1) and builds a *disaggregated model* η that characterizes the segment by concatenating trained component models (2). Such models propagate the set of features *F* that characterize IQ constellation points as bi-variate Gaussian distributions with its mean position (μ_I and μ_Q), its variance (σ_I and σ_Q), and covariance (σ_{IQ}), i.e., five features per constellation point (see [5] for details).

To compose the e2e model of the lightpath, let us assume that source and intermediate domains of the lightpath share intra-domain models with the destination domain through the network orchestrator. Therefore, *intra-domain model synthesis* is carried out in the domains sharing their models to generate exportable segment models φ from the disaggregated ones (3). Models φ hiding the internals of the domain for privacy preserving, are sent to the domain SDN controller (4) and shared with the destination SDN controller through the orchestrator (5). Finally, OCATA in the destination domain generates the e2e model from the received φ models (6-7). Note that the e2e model is a DNN-based model that concatenates the intra-domain models of the domains in the path (φ_{D1} and φ_{D2}),

The research leading to these results has received funding from the Smart Networks and Services Joint Undertaking under the European Union's Horizon Europe research and innovation programme under G.A. No. 101096120 (SEASON), the MICINN IBON (PID2020-114135RB-I00) project and from the ICREA Institution.

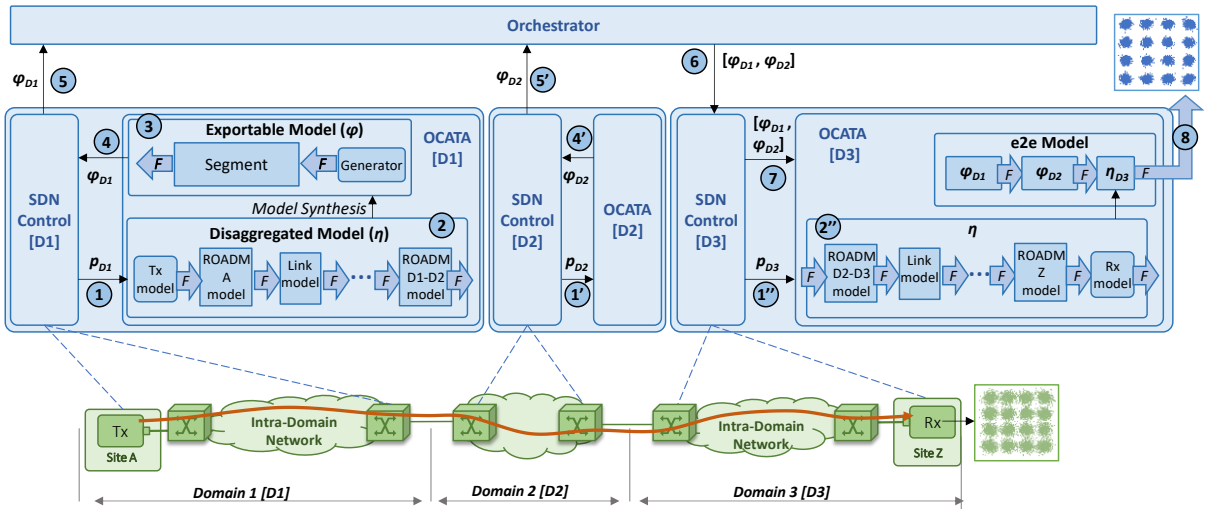

Fig. 1: Multi-domain scenario and proposed workflow

together with the disaggregated model for the local domain (η_{D3}). By propagating features F from source to destination, expected IQ constellations can be obtained (8) and used for the selected application.

3. Exportable Model Synthesis

This section details the procedure carried out by source and intermediate domains of an e2e lightpath to generate exportable intra-domain models φ from disaggregated models η. To this aim, two different sets of component models are available in OCATA. On the one hand, the *symmetric non-linear biased* (SNLB) set contains highly accurate DNNs with: *i*) equal number of neurons per layer; and *ii*) non-linear activation functions (e.g., hyperbolic tangent, tanh) and non-zero bias for every hidden and output neuron. This set is used to build η, which is expected to provide the highest fidelity to model the intra-domain lightpath segment. On the other hand, the *asymmetric linear un-biased* (ALUB) set contains component models where: *i*) each layer has a different number of neurons; and *ii*) the activation function of the first hidden layer and output layer is linear and the bias of the first hidden layer and output layer is zero. This set is used to generate an alternative disaggregated model η'. Although η is more accurate than η', the latter exhibits good properties for security that allow easily hiding the concatenation of consecutive components by merging layers, thus obtaining a layered intra-domain model whose components cannot be isolated.

Algorithm 1: Exportable model synthesis procedure

IN: η, ALUB, *thr*	**OUT**: φ
1:	$\eta' \leftarrow \emptyset$
2:	**for** *i* in 1..|*p*| **do**
3:	$\eta' \leftarrow \eta' \cup$ getMostDiverse(ALUB, $\eta[i]$.type, η')
4:	$\varphi \leftarrow$ mergeComponents(η') (eq. (1) and Fig. 2)
5:	$\varphi \leftarrow$ shuffleHiddenNeurons(φ)
6:	**return** truncateCoeffs(φ, *thr*)

The overall procedure is detailed in Algorithm 1, which receives the disaggregated intra-domain model η, models set ALUB, and a threshold *thr* used for truncating non-significant model coefficients. Without loss of generality, we assume that every single component model in η has k alternative models in ALUB, each of them with different number and configuration of hidden layers. After initializing model η' (line 1 in Algorithm 1), a model for each component in η is chosen from the ALUB set and added to η' (lines 2-3). At every iteration, the selected model is the one that maximizes the diversity (computed by means of the Euclidean distance in the multi-dimensional space of model parameters and coefficients) with respect to the component models already added to η'. Then, a merging procedure between consecutive component models is applied (line 4). This procedure is sketched in Fig. 2 for models u and v (each with n layers), where $u(i)$ denotes the i-th layer of model u. Thus, weights α (for connections between the last hidden layer and the output of model u) and β (for connections between the input layer and the first hidden layer of model v) can be linearly merged to obtain new weights γ using eq (1).

$$\gamma_{ij} = \sum_{k \in F} \alpha_{ik} \cdot \beta_{kj} \quad \forall i \in u(n-1), j \in v(2) \qquad (1)$$

Although the exportable model φ is ready at this point, layer *obfuscation* makes more difficult getting lightpath details by model inspection. Thus, neurons are randomly shuffled within their layer and weights and biases are truncated to 0 if they are lower than *thr* (lines 5-6).

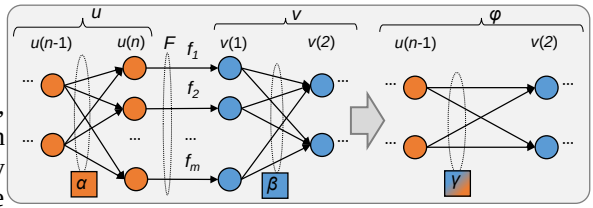

Fig. 2: DNN layer merging

4. Illustrative Results

In order to evaluate the performance of the proposed methodology, we implemented a Python-based simulator emulating a transparent multi-domain network. The OCATA instance of each domain was loaded with component

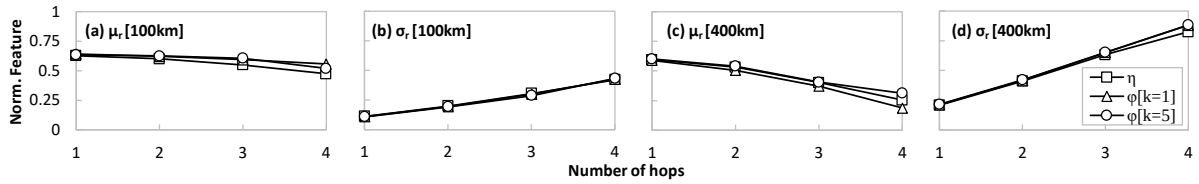

Fig. 3: Feature propagation performance of constellation point [-3+3i]

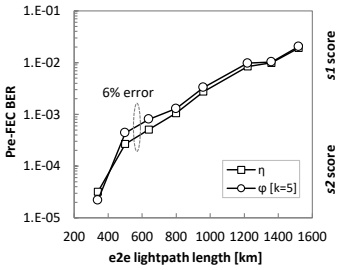

Fig. 4: e2e QoT estimation

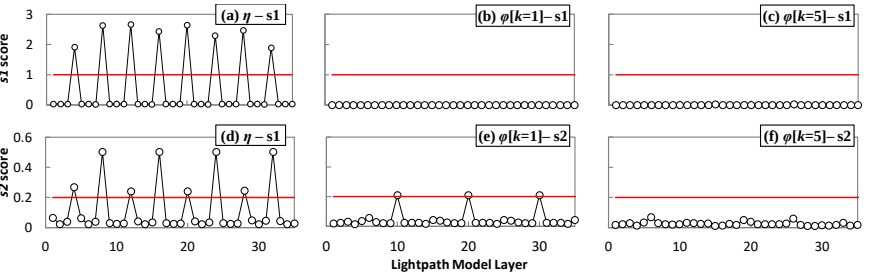

Fig. 5: Privacy evaluation of lightpath example (4 hops of 400-km links)

models (ROADM and fiber links ranging from 100 km to 400km) trained with data for 16- quadrature amplitude modulation (QAM) optical constellations available in [6]. As introduced in Section 2, each constellation point is characterized by 5 features that are propagated through the DNN models. To reduce DNN models' complexity, features for only four selected constellation points ([*-3+3i*], [*1+1i*], [*-1-1i*], and [*1-3i*]) are propagated. Then, the SNLB set contains one single DNN model per component type, with 20 input and output neurons, and 3 hidden layers with 20 tanh neurons. Regarding the ALUB set, *k* models per component were trained with a variable number of hidden layers (from 3 and 6), number of neurons per layer (from 10 to 30), and activation function, while ensuring the characteristics of first hidden and output layer of component models described in Section 3.

We firstly focus on evaluating the procedure in Algorithm 1 to obtain φ models for a wide range of segments with 1 to 4 hops, which leads to total intra-domain distances between 100 km and 1,600km for the lightpath segments. Three approaches are compared: *i*) using the disaggregated model η (for benchmarking purposes), *ii*) φ with k=1 (i.e., without model diversity), and *iii*) φ with k=5 (with model diversity). In both φ models, *thr* was set to 0.01. Fig. 3 shows the expected evolution of μ_l and σ_l (normalized to training values) of constellation point [-3+3i] for segments with links of 100km (Fig. 3a-b) and 400km (Fig. 3c-d). We observe in the results that all models provide similar performance in estimating propagated features, which validates the accuracy of φ models.

Next, we reproduce the scenario and workflow in Fig. 1 considering different distances and link lengths for the domains. The approach in [1] for estimating the pre-FEC BER from propagated features is used as e2e QoT evaluation metric in the OCATA instance of domain D3. Fig. 4 shows the estimated pre-FEC BER as a function of e2e lightpath length for the considered approaches. We observe that accurate e2e QoT estimation is provided by exportable model φ (~6% maximum error in the log scale with respect to benchmarking using η model).

Finally, we evaluate privacy for a 4-hop lightpath with 400-km links. Fig. 5 shows two score-based methods that aim at finding those intermediate model layers that are boundary (beginning or end) of components, as a way to infer some characteristics of the topology of the intra-domain network topology. In the first method (*s1*), intermediate layer outputs are retrieved and compared with model output in terms of mean square error (MSE) (Fig. 5a-c), while the second method (*s2*) tries to find patterns of model coefficients that repeat along the model (Fig. 5d-f). In both cases, the score is inversely proportional to MSE and large values indicate a layer that is suitable to be a component boundary (red lines show significant levels). As expected, η model provides detailed internal information; one peak per component (ROADMs except last and links) is clearly observed. On the contrary, components cannot be isolated by *s1* when analyzing φ models. Moreover, adding diversity (k=5) prevents detecting coefficient patterns by *s2*, which validates the proposed exportable model synthesis procedure.

5. Conclusions

The composition of e2e models for multi-domain lightpaths has been proposed and validated to increase security during model sharing. OCATA builds exportable models for intra-domain lightpaths segments that are shared. Results showed high accuracy of the e2e model together with valuable intra-domain privacy preserving properties.

References

[1] L. Velasco *et al*., "Applications of Digital Twin for Autonomous Zero-Touch Optical Networking [Invited]," in Proc. ONDM, 2023.
[2] E. Bertin *et al*., *Shaping Future 6G Networks: Needs, Impacts, and Technologies*, Wiley & Sons, 2021.
[3] F. Tabatabaeimehr *et al*., "Cooperative Learning for Disaggregated Delay Modeling in MultiDomain Networks," IEEE TNSM, 2021.
[4] P. Porambage *et al*., "The Roadmap to 6G Security and Privacy," IEEE Open Journal of the Communications Society, 2021.
[5] D. Sequeira *et al*., "OCATA: A Deep Learning-based Digital Twin for the Optical Time Domain," IEEE/OPTICA JOCN, 2023.
[6] M. Ruiz *et al*., "Optical Constellation Analysis (OCATA)," https://doi.org/10.34810/data146, CORA, V2, 2022.