Federated Privacy-Preserving Strategy for Generalizing Soft-Failure Localization in Multi-Carrier Optical Networks

Forough Shirin Abkenar[1], Ramanuja Kalkunte[1], Venkata Virajit Garbhapu[5], Sifat Ferdousi[1],
Yusuke Hirota[2], Sugang Xu[2], Masaki Shiraiwa[2], Aryanaz Attarpour[3],
Massimo Tornatore[3], Yoshinari Awaji[2], Biswanath Mukherjee[1,4]
[1]University of California, Davis; [2]NICT, Japan; [3]Politecnico di Milano, Italy;
[4]Soochow University, China; [5]Huawei Paris Research Center, France.
Email: fshirina@ucdavis.edu

***Abstract*—We propose a privacy-preserving strategy based on federated learning to localize soft failures in multi-carrier optical networks using a self-supervised approach on unlabeled data. Evaluations conducted on data from a testbed demonstrate the effectiveness of the proposed strategy.**

I. Introduction

Fault management plays a crucial role in efficiently operating optical networks as failures can result in service interruptions and loss of critical data, leading to revenue loss and customer dissatisfaction [1]. In particular, soft failures (i.e., failures that do not entail a complete interruption of the communication but only a degradation of signal quality, in contrast to hard failures) lead to complex non-linear patterns that cannot be easily extracted by manual inspection of performance logs of the network equipment. Thanks to the emergence of machine learning (ML), novel techniques can efficiently analyze such complex patterns.

Relying on ML-based anomaly detection techniques [2], Ref. [3] proposes a hybrid learning approach to localize soft failures in multi-carrier optical networks, where a self-supervised model extracts normal and abnormal patterns from data and a federated learning (FL) model localizes failures at the node/domain levels. Similarly, Ref. [4] proposes to localize soft failures, derived from laser drifts and filter misalignments by training different supervised classifiers specific to each carrier to detect anomalies. The aforementioned approaches require sending telemetry data to a central entity for training, raising concerns about *the privacy of data* belonging to the carriers; in this regard, FL is a promising approach, as it allows model training without requiring raw data to leave the local carriers. For example, Ref. [5] develops a flavor of FL, namely vertical FL model, which incorporates diverse features from different operators using data instances of the same type, to localize soft failures in partially-disaggregated optical networks, in which the transceivers are from a different vendor than the one providing the open line system (OLS). Similarly, Ref. [6, 7] leverages the principal component analysis (PCA) technique to transmit telemetry data to an unreliable third party for further failure localization analysis, thereby ensuring the privacy of data among each carrier.

Existing methods in the literature, such as [4], are heavily based on supervised failure data, which is expensive and time consuming to collect. Furthermore, these methods lack a generalized approach for localizing soft failures in varying magnitudes of the same failure type. For instance, each carrier in [6] shuffles its own data and sends them to the third party. The third party then performs a clustering algorithm on the received data and localizes failures. Although this preserves the privacy of the data, when applied to unsupervised datasets from domains different from the training data, new models often require retraining to achieve high localization accuracy. However, this retraining process is resource intensive and time consuming.

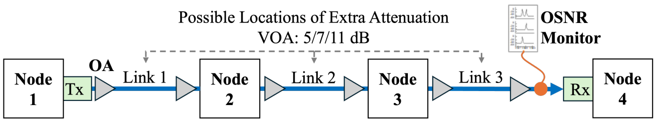


Fig. 1: Testbed topology for failure scenarios.

To address these limitations, this paper introduces a novel framework designed for multi-carrier systems, where differences in the topological characteristics of carriers lead to variation in data distribution. Hence, each carrier operates within a unique domain represented by unlabeled datasets with diverse distributions. Consequently, incorporating contributions from multiple carriers with distinct domains is essential to enhance model generalization and improve localization accuracy. The key contributions of this work are threefold: (*i*) generalize a model that can effectively localize failures of varying magnitudes; (*ii*) deploy a self-supervised learning phase for making localization; and (*iii*) preserve privacy of carriers' data during training phase. To this end, we propose the Privacy-Preserving Strategy (PPS) that integrates domain adaptation and knowledge distillation techniques to generalize the model in a federated manner, preserving data privacy. It also exploits high-performance unsupervised clustering models to support the self-supervised learning phase.

II. Topology And Dataset Structure

The primary network topology is a multi-carrier system comprising a central provider-neutral entity (PNE) [8] and multiple carriers. For each carrier, the study uses real telemetry data collected from a testbed at NICT in Sendai, Japan. The testbed topology, shown in Fig. 1, consists of four nodes and three links, each node equipped with pre- and post-optical amplifiers (OAs). Link-level soft failures are emulated as additional attenuation imposed by variable optical attenuators (VOAs), effectively simulating amplifier malfunction. Attenuation levels of +5, +7, and +11 dB were tested, with optical signal-to-noise ratio (OSNR) measured at the receiver node. The datasets are defined as follows: (*i*) 888: 80 km links, 100 Gbps transponders, and 192.4 THz frequency; (*ii*) 555: Similar to 888, but with 50 km links; (*iii*) 888_10G: Similar to 888, but with 10 Gbps

978-3-903176-67-6 © 2025 IFIP

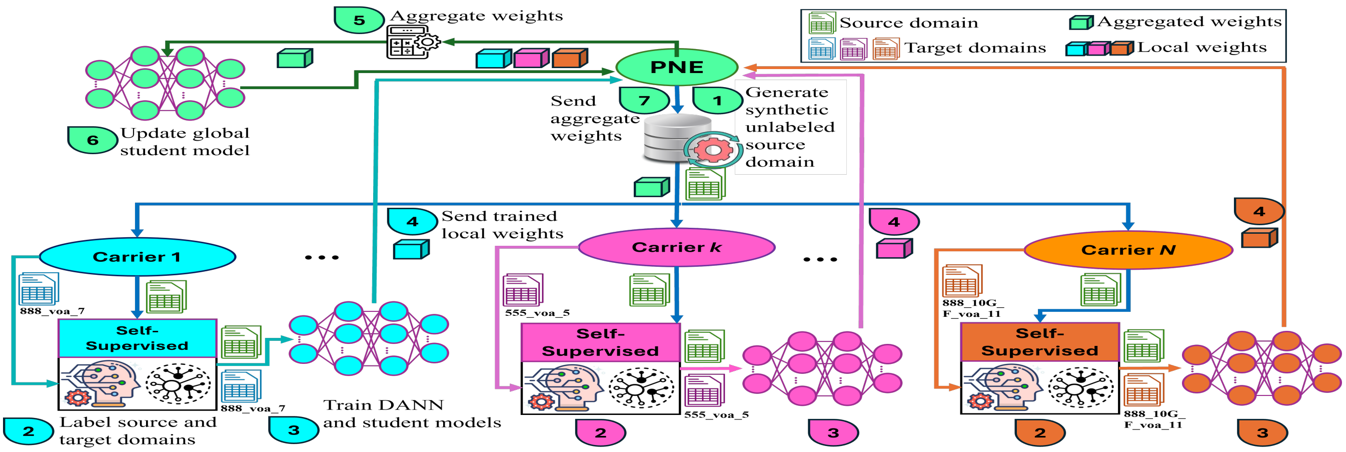


Fig. 2: Flow diagram of PPS.

transponders; (*iv*) 888_F: Similar to 888, but with 194.1 THz frequency; and (*v*) 888_10G_F: 80 km links, 10 Gbps transponders, and 194.8 THz frequency. Datasets are labeled as $\{dataset\_title\}\_\{voa\}\_\{magnitude\}$, e.g., "888_voa_5" for 888 with 5 dB attenuation. Data is captured under four conditions: Normal (no failure), Failure_1 (VOA at link 1), Failure_2 (VOA at link 2), and Failure_3 (VOA at link 3).

III. Proposed Privacy-Preserving Strategy

PPS is designed to generalize a model for failure localization across datasets with diverse domains, while preserving data privacy. To achieve this, PPS leverages three key paradigms: (*i*) FL that ensures data privacy by enabling decentralized model training; (*ii*) domain adaptation that adapts the model to varying domain-specific characteristics; and (*iii*) knowledge distillation that facilitates the localization and generalization of soft failures across different failure magnitudes. Figure 2 illustrates the overall workflow of PPS, comprising seven distinct steps. These steps (detailed below) are distributed between the PNE and the carriers.

Self-Supervised Learning: Since PPS relies on domain adaptation, it is essential to define a source domain for training. To achieve this, we first normalize the OSNR values across different datasets using *MaxMin* normalization which scales the values between zero and one. Our analysis of sample datasets–888_voa_5, 888_voa_7, 888_10G_F_voa_11, and 888_10G_voa_7–presented in Fig. 3 reveals that "Normal", "Failure_3", "Failure_2", and "Failure_1" consistently exhibit the highest to the lowest OSNR values. Notably, dataset 888_10G_voa_7 stands out, as its categories are clearly distinguishable with no overlap between them. However, overlap between failure classes is observed in other datasets which suggests that adapting the categories of all datasets to align with those in 888_10G_voa_7 could enable the model to generalize effectively for failure localization. Therefore, we designate 888_10G_voa_7 as the source domain and all other datasets as target domains.

However, sharing 888_10G_voa_7 with all entities in the network will violate the privacy concerns. To avoid such a violation, PPS exploits Metropolis Hasting (MH) algorithm [9], which generates synthetic data very similar to the ground-truth samples, using the statistical measurements of a predefined distribution. Accordingly, in step **1**, the PNE employs the MH algorithm and generates synthetic unlabeled source domain data based on 888_10G_voa_7. Each carrier includes its own dataset, referred to as target domain, which is unlabeled. In step **2**, the corresponding carrier employs K-Means clustering model on target domain and source domain, separately, to form initial clusters for training purpose. Considering all four categories of failures, the number of clusters in K-Means is set to four.

For training purpose, we combine and separately normalize all failure classes within each dataset. However, the classes of failure are not known in test dataset. To address this limitation, we develop a random forest (RF) model trained on unnormalized training data alongside its corresponding normalized data. The model learns to normalize unseen data effectively in our framework.

Domain Adaptation: Step **3** implements the domain-adaptation phase. Each carrier in PPS involves a domain adversarial neural network (DANN) model [10] consisting of three main components: a feature extractor, a label predictor, and a domain classifier. The feature extractor is a deep neural network (DNN) with an input layer, two hidden layers, and an output layer. The first and second hidden layers are composed of 100 and 50 neurons, respectively. The label predictor is a neural network with a hidden layer of 100 neurons and an output layer of size 4, corresponding to the four failure categories. The domain classifier includes two hidden layers with 100 neurons each and an output layer with 2 neurons to differentiate between the source and target domains. Each carrier trains its own DANN model using the pseudo-labeled source and target domains obtained from the self-supervised learning phase in step **2**. The trained models are referred to as teacher models. Notably, all hyperparameter values are determined through extensive trials.

Knowledge Distillation: To generalize a failure localization model without retraining for changes in attenuation magnitude, PPS leverages an offline knowledge distillation mechanism [11]. The PNE defines a smaller DANN-based student model that learns failure localization from teacher models trained by the carriers. In step **4**, carriers normalize a portion of their unlabeled test data using the developed RF model, train the student model, and send the updated weights back to the PNE. PNE then averages the weights, updates the global student model, and sends the weights back to the carriers during steps **5**-**7**, enabling PPS to generalize failure localization without sharing sensitive carrier details.

IV. Numerical Evaluation

We compare the performance of PPS with two counterpart DNN models, namely DNN Supervised (DNN-S) and DNN

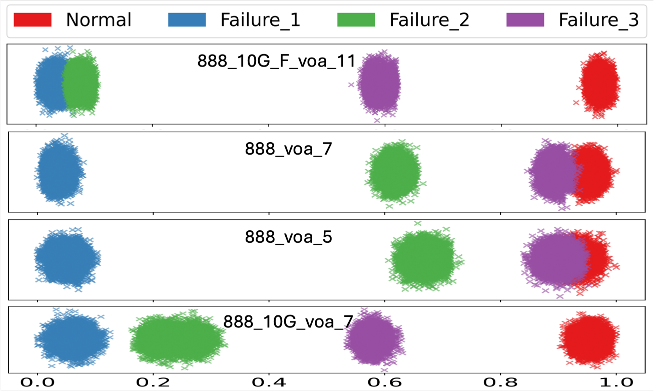


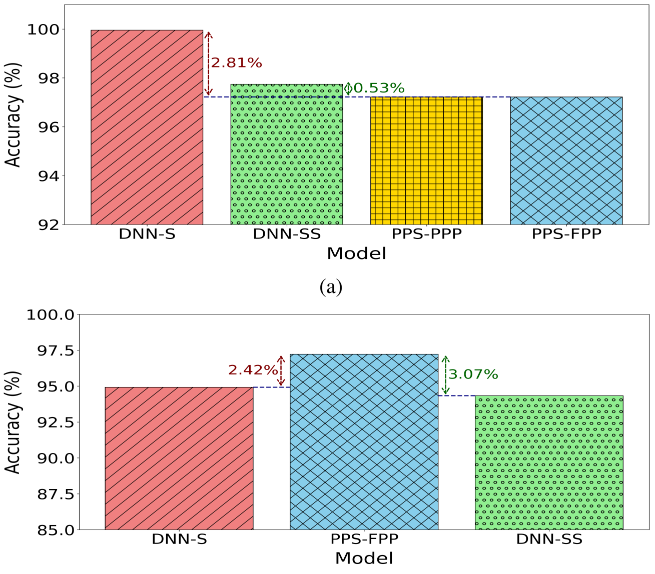
Fig. 3: Normalized OSNR.

Self-Supervised (DNN-SS). The former is a model developed on fully-labeled datasets, wherein the information of all carriers is shared as a centralized dataset with DNN-S. The latter assumes that data are unlabeled and leverages K-Means clustering to label data. Similar to DNN-S, all data belonging to carriers are shared with DNN-SS for training purposes.

After conducting multiple trials, we set the hyperparameters as follows: For both teacher and student models, batch size and learning rate are set to 64 and 0.001, respectively; for the student model, we consider 64 neurons in the first hidden layer of the feature extractor and 32 neurons in the second layer. The number of neurons deployed in hidden layers of the label predictor and domain classifier is set to 64; each teacher model is trained over 1000 epochs, while it is 10 epochs for the student model.

To show the proficiency of PPS in preserving the privacy of source and target domains, we consider two versions of PPS: fully privacy-preserving PPS (PPS-FPP) and partially privacy-preserving PPS (PPS-PPP). The former feeds the teacher model with the synthetic source domain, while the latter directly feeds the teacher model with 888_10G_voa_7. Figure 4a indicates the corresponding results wherein PPS-PPP and PPS-FPP achieve very similar performance in failure localization with accuracy of around 97.2% (their difference is about 0.0004%). As expected, DNN-S achieves the highest accuracy of 99.95% among all four strategies. However, it not only violates the privacy of data for training purposes, but also requires all data to be fully labeled. On the other hand, DNN-SS relaxes the need for labeled data by leveraging the self-supervised learning mechanism; it still violates the privacy of data for training purposes for less than 1% of performance improvement compared to PPS.

Leveraging the nature of domain adaptation and the similarity in data distribution for attenuation magnitudes of 5 and 7 dB, we investigate whether size of the training dataset can be reduced. Our results show that using only data from voa_7 and voa_11 yields promising outcomes. To further validate this, we conducted experiments using an adaptive training set (ATS) where data corresponding to 5 dB was excluded and only data for 7 and 11 dB were utilized. The simulation results obtained in Fig. 4b indicate that PPS-FPP results in high precision of 96.22% and outperforms DNN-S and DNN-SS. This emphasizes the ability of PPS to generalize a model for failure localization, thanks to domain adaptation and knowledge distillation mechanisms; yet, this represents a degradation of 1.03% compared to the non-ATS set which includes data at 5 dB during training. As a result, if the similarity between data distributions is known, we can optimize the training process for teacher models, reducing overall training time.



(a)

(b)

Fig. 4: Performance comparison of proposed PPS and baselines w.r.t. different training sets: (a) Regular, (b) Adaptive.

V. Conclusion

We proposed PPS, a distributed strategy designed to generalize soft-failure localization across varying magnitudes of the same failure in multi-carrier systems. PPS utilizes federated learning to ensure data privacy while employing domain adaptation and knowledge distillation techniques to enhance generalization of the model. Evaluation results demonstrate that PPS achieves high precision in failure localization requiring fewer carriers to participate in the training phase which reduces the overall training burden across the network.

Acknowledgment: This work is supported by US-Japan JUNO3 project: NSF Grant no. 2210384. Massimo Tornatore acknowledges the support by the PRIN PNRR project GASTON, funded by Italian Ministry of University and Research.

References

[1] F. Musumeci *et al.*, “A Tutorial on Machine Learning for Failure Management in Optical Networks,” *JLT*, vol. 37, no. 16, pp. 4125–4139, 2019.

[2] C. Natalino *et al.*, “Spectrum Anomaly Detection for Optical Network Monitoring Using Deep Unsupervised Learning,” *IEEE Communications Letters*, vol. 25, no. 5, pp. 1583–1586, 2021.

[3] X. Chen *et al.*, “On Cooperative Fault Management in Multi-Domain Optical Networks Using Hybrid Learning,” *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 28, no. 4, 2022.

[4] B. Shariati *et al.*, “Learning From the Optical Spectrum: Failure Detection and Identification,” *JLT*, vol. 37, no. 2, pp. 433–440, 2019.

[5] F. Musumeci *et al.*, “Vertical Federated Learning for Failure Localization in Partially Disaggregated Optical Networks,” in *HPSR*, July 2024.

[6] M. F. Silva *et al.*, “Confidentiality-preserving Machine Learning Scheme to Detect Soft-failures in Optical Communication Networks,” in *ECOC*, Sept. 2022.

[7] M. Silva *et al.*, “Confidentiality-preserving machine learning algorithms for soft-failure detection in optical communication networks,” *J. Opt. Commun. Netw.*, vol. 15, no. 8, pp. C212–C222, 2023.

[8] S. Sahoo *et al.*, “Datacenter-Carrier Cooperation over Optical Networks during Disaster Recovery,” in *OFC*, 2022.

[9] M. Kenyeres and J. Kenyeres, “Performance Analysis of Generalized Metropolis-Hastings Algorithm over Mobile Wireless Sensor Networks,” in *Cybernetics & Informatics (K&I)*, 2020.

[10] G. Yaroslav *et al.*, “Domain-Adversarial Training of Neural Networks,” *Journal of Machine Learning Research*, vol. 17, pp. 2096–2030, 2016.

[11] K. Zhang *et al.*, “Student Network Learning via Evolutionary Knowledge Distillation,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 32, no. 4, pp. 2251–2263, 2022.