Automating Optical Network Fault Management with Machine Learning

Automating Optical Network Fault Management with Machine Learning

Journal:	*IEEE Communications Magazine*
Manuscript ID	COMMAG-22-00110.R1
Manuscript Type:	Series Optical Communications and Networks (S1)
Date Submitted by the Author:	01-Jul-2022
Complete List of Authors:	Chen, Xiaoliang; Sun Yat-Sen University, School of Electronics and Information Technology Liu, Che-Yu; UC Davis, Dept. of Computer Science Proietti, Roberto; UC Davis, ECE Li, Zhaohui; Sun Yat-Sen University, School of Electronics and Information Technology Yoo, Ben; University of California, Davis, Electrical and Computer Engineering
Keywords:	fault management, soft failure, machine learning, cognitive optical networking

Automating Optical Network Fault Management with Machine Learning

Xiaoliang Chen, Che-Yu Liu, Roberto Proietti, Zhaohui Li, S. J. Ben Yoo

***Abstract*—Effective fault management is essential for quality-of-service assurance in optical networks. Conventional fault management designs for optical networks mainly rely on threshold-based rules which can hardly characterize the complex fault patterns therein. This article discusses the application of machine learning (ML) in actuating an automated optical network fault management architecture. The architecture is built upon advanced optical performance monitoring (OPM) techniques and software-defined-networking-enabled programmable network control and management. With the viability of abundant OPM data, the architecture employs various ML models to automatically learn fault patterns and thereby to realize data-driven cognitive fault management. We review the state of the art on fault detection, identification and localization based on such an architecture, focusing specifically on soft failures. To overcome the applicability and scalability issues encountered by existing soft failure detection designs, we present a hybrid learning solution that combines the merits of supervised, unsupervised and cooperative ML. In particular, we present a self-taught mechanism that self-learns fault patterns with unsupervised learning and then trains supervised learning classifiers with the learned patterns for online detection. We further introduce a broker-plane-aided federated learning framework to enable collaborative training of classifiers from multiple network domains while complying with domain privacy constraints. Performance evaluations show that the hybrid learning design can achieve high fault detection rates with negligible false alarms using less than ten abnormal data samples ($\sim 0.1\%$ of the scale of normal samples) for training.**

I. Introduction

THE proliferation of emerging applications and networking paradigms (e.g., 5G, edge computing) is posing greater challenges to the underlying optical infrastructures, demanding not only larger capacities but also more stringent quality-of-service guarantees. In this context, it becomes imperative to advance optical network fault management for assuring consistent service performance while sustaining the latest yet more sophisticated data plane technologies.

One of the most critical fault management tasks is tackling soft failures. Unlike hard failures (e.g, fiber cuts) that will disrupt connections immediately, soft failures refer to faults that can gradually deteriorate service performance with mild to moderate intensities, for instance, equipment ageing/malfunctioning [1], misconfiguration, and physical-layer attacks [2]. Agile detection and restoration of soft failures can prevent the occurrence of the consequent hard failures and promote higher resource efficiency by allowing provisioning with reduced margins reserved for the potential faults. Nevertheless, because soft failures often exhibit heterogeneous and complex patterns, their effective management is a non-trivial task.

Traditional fault management designs targeting soft failures mostly make use of threshold-based rules for fault detection and manual inspections by network experts for fault identification and localization. While it can be hard to characterize soft failures with simple threshold-based rules, especially when network conditions keep evolving, the heavy involvement of human efforts leads to poor operational efficiency and scalability issues. Recently, machine learning (ML) has been receiving significant attention as a key enabler for building next-generation intelligent optical networks. ML models can potentially learn complex network rules (e.g., quality-of-transmission models, traffic trends) or operation policies (e.g., routing and wavelength assignment policies) from high-dimensional network state data/traces without explicit programming. With the late advances in optical performance monitoring (OPM) techniques and the maturity of programmable network control and management enabled by software-defined networking (SDN) [3], it is possible to deploy ML models to automatically extract fault patterns from real-time OPM data and thereby to realize data-driven cognitive fault management [1], [4]–[15].

In this article, we present an ML-aided automated fault management architecture for optical networks. We first briefly describe the related system layout, including several key functionalities and their workflows for enabling automated fault management. Then, we review the state-of-the-art proposals for soft failure detection (SFD), identification (SFI) and localization (SFL) based on such an architecture. To overcome the applicability and scalability issues from the existing soft failure detection designs, we present a hybrid learning solution that combines the merits of supervised, unsupervised and cooperative ML. Performance evaluations show that the hybrid learning design can achieve high fault detection rates with negligible false alarms when only less than ten abnormal samples are used for training. Finally, we provide concluding remarks and discuss remaining challenges toward automating optical network fault management.

X. Chen and Z. Li are with the Guangdong Provincial Key Laboratory of Optoelectronic Information Processing Chips and Systems, Sun Yat-sen University, Guangzhou 510275, China. Z. Li is also with the Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), 519000, China (Email: xlichen@ieee.org, lzhh88@mail.sysu.edu.cn).

C. Liu is with the Department of Computer Science, University of California, Davis, Davis, CA 95616, USA (Email: cyliu@ucdavis.edu).

R. Proietti is with the Department of Electronics and Telecommunications, Politecnico di Torino, 10129, Torino, Italy (Email: roberto.proietti@polito.it).

S. J. B. Yoo is with the Department of Electrical and Computer Engineering, University of California, Davis, Davis, CA 95616, USA (Email: sbyoo@ucdavis.edu).

Manuscript received March 23, 2022

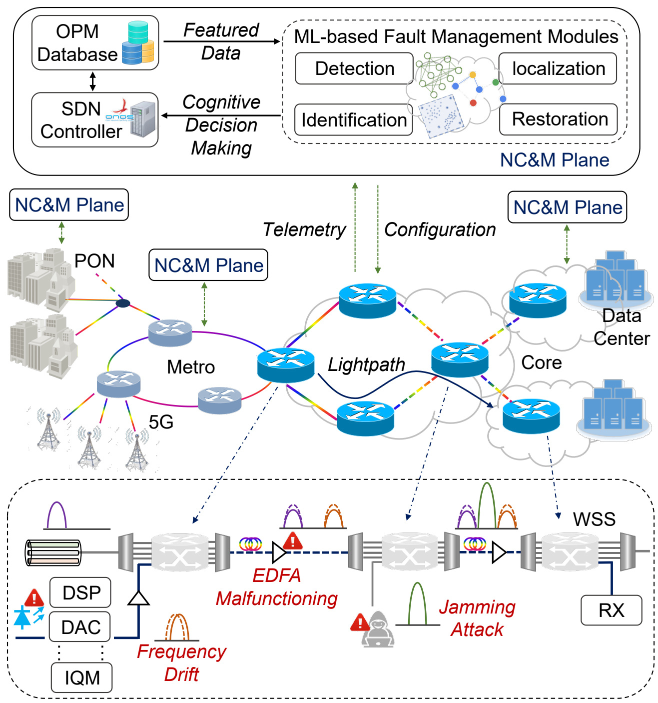

Fig. 1. Illustrative optical network architecture with automated fault management. NC&M: network control and management; OPM: optical performance monitoring; SDN: software-defined networking; PON: passive optical network; DSP: digital signal processing; DAC: digital-to-analog converter; IQM: in-phase/quadrature modulator; EDFA: erbium-doped fiber amplifier; WSS: wavelength selective switch; RX: receiver.

II. Architecture

Soft failures in optical networks can originate from devices at arbitrary network locations with diverse intensities and root causes. Fig. 1 depicts a representative optical network architecture spanning access, metro, and core segments, with multiple autonomous network domains. We sketch out at the bottom of the figure the system structure with respect to an end-to-end lightpath with three instances of soft failures that can degrade the overall system performance. Specifically, a faulty source laser can emit signals with central frequency drifting from the initial configuration, resulting in misalignment of signal and filter bandwidth and consequently increased bit error rate (BER) at the receiver side. The intensity of the signal power is barely affected in this case [1]. Whereas the malfunctioning of an amplifier placed in the middle of the lightpath can directly cause the deterioration of the optical signal-to-noise ratio (OSNR). Aside from factors stemming from the network itself, a malicious user can invade into a node and launch a physical-layer attack by jamming a high-power signal at an adjacent channel and causing interference through the nonlinear effects introduced [2].

The heterogeneity of soft failures makes the traditional threshold-based policies hard to be effective and scalable, and therefore, motivates the development of more powerful fault management architectures with self-learning and cognitive decision-making capabilities. Fig. 1 illustrates an abstracted layout of such an architecture, where each network manager (NM) employs SDN-based network control and management owing to its high programmability. The architecture realizes automated fault management through an observe-analyze-act-based operation cycle. In particular, by leveraging the communication protocols and interfaces offered by SDN (e.g., P4, NETCONF/YANG [3]), NMs are able to implement telemetry services for collecting OPM data related to a customized set of parameters (e.g., signal power, chromatic dispersion) on demand (**observe**). The viability of rich OPM data then enables the deployment of various ML models aiming at different fault management tasks, such as fault detection, identification, localization, and restoration. These models can learn sophisticated (spatial and temporal) fault patterns or rules from OPM traces (**analyze**) and in turn assist in cognitive decision making by the SDN controllers (**act**). In the case of inter-domain networking, NMs can transfer learned knowledge mutually or even perform cooperative learning under the coordination of a broker plane (a trusted third-party entity engaged in multi-domain service provisioning) for pursuing enhanced performance of the ML models.

III. State of the Art

Based on the architecture discussed above, recent studies have demonstrated a number of novel ML-based fault management designs, mostly targeting SFD, SFI, and SFL, as briefly summarized in Table I.

SFD: accurate detection of soft failures is the premise of effective fault management. In [1], the authors developed a finite-state-machine-based algorithm for detecting significant BER degradation due to signal overlap, filter bandwidth tightening, and shift of laser or filter central frequency. A set of states related to the evolution of BER are defined in reference to several adaptive boundaries. In [4], the authors proposed to detect abnormal variations of signal power with a two-stage approach, which makes use of extreme studentized deviate test to perform preliminary detection and an artificial neural network (ANN) classifier for ultimate diagnoses. The former enables a quick screening that can largely reduce the computation-intensive processing by the ANN. Experimental results show a higher accuracy from the cognitive approach compared with threshold-based methods. Later, the work in [5] compared different ML techniques from the angles of accuracy, complexity and sensitivity when applied to detection of abnormal BER variations. The authors concluded that, for the specific scenario under test, random forest achieves a good trade-off between the different performance metrics. Similar to the idea presented in [4], Shu *et al.* proposed a two-stage design for SFD [6]. The first stage utilizes a low-complexity Gaussian distribution model to detect faults from monitoring of BER and power. To reduce false alarms due to the limited accuracy of the Gaussian distribution model, the second stage calls for extract digital spectrum features (e.g., spectrum area) and applies one-class support vector machine (SVM) to further scrutinize the alarms raised by the first stage. More recently, Lun *et al.* leveraged the generative adversarial network technique to learn a mapping from electrical spectra to a latent space where the distributions of normal and abnormal samples are evidently distinguishable [7]. The advantage of the approach is that only normal samples are needed for

TABLE I

A BRIEF SUMMARY OF RECENT PROGRESSES IN ML-BASED FAULT MANAGEMENT DESIGN. N.S.: NOT SPECIFIED; ANN: ARTIFICIAL NEURAL NETWORK; SVM: SUPPORT VECTOR MACHINE; FEC: FORWARD ERROR CORRECTION; TL: TRANSFER LEARNING.

Literature	Type of Fault	Task			OPM/State Data	Approach
		SFD	SFI	SFL		
Vela *et al.* [1]	signal overlap; filter shift; filter tightening	✓	✓		received power; BER	finite state machine; analytical model
Rafique *et al.* [4]	N.S.	✓			received power	extreme studentized deviate test; ANN
Shahkarami *et al.* [5]	EDFA malfunctioning; filter tightening	✓	✓		BER	SVM; ANN; random forest
Shu *et al.* [6]	EDFA ageing; filter shift filter tightening; laser drift	✓	✓		received power; digital spectrum features; BER	SVM; Gaussian distribution
Lun *et al.* [7]	filter tightening; filter shift	✓	✓		electrical spectra	generative adversarial network
Furdek *et al.* [8]	in-band/out-of-band jamming; polarization modulation attack	✓			BER; block errors; chromatic dispersion etc.	density-based clustering; SVM
Abdelli *et al.* [9]	fiber reflective faults	✓		✓	optical time-domain reflectometry power	long short-term memory network
Du *et al.* [10]	loss of signal	✓			number of FECs; power; signal quality; delay etc.	random forest; XGBoost; TL etc.
Lun *et al.* [11]	EDFA malfunctioning; filter shift; filter tightening; Kerr nonlinear effect		✓		power spectrum density	convolutional neural network
Panayiotou *et al.* [12]	N.S.			✓	fault alarms; routing paths	Gaussian process classifier; heuristic
Li *et al.* [13]	N.S.			✓	fault alarms; correlation between alarms	graph neural network
Mayer *et al.* [14]	transponder/EDFA/fiber malfunctioning			✓	transponder/EDFA power, OSNR	ANN

training (also true for one-class SVM), but there is a lack of established theory supporting the distinguishability assumption in more generic settings. Unlike the aforementioned studies that mainly address common faults of lasers, amplifiers and filters, the works in [8], [9] focused on detecting optical-layer attacks or fiber reflective faults. In particular, Furdek *et al.* analyzed the performance of clustering and SVM in detecting in-band/out-of-band jamming and polarization modulation attacks, assuming the availability of diverse OPM data [8]. In [9], the authors devised a long short-term memory-based neural network model that can fulfill fiber reflective fault detection, quantification and localization simultaneously by exploiting power traces measured by optical time-domain reflectometry. They showed that the multi-task learning design outperforms the independent ones. Lately, Du *et al.* demonstrated the potential of forecasting loss of signal events using time-series and heterogeneous performance metrics reported by layer-1/2 modules [10]. Different classification algorithms, i.e., random forest, XGBoost and BRITS, were investigated. The authors also presented a transfer learning approach to enable knowledge sharing between networks.

To sum up, existing cognitive solutions for SFD mostly apply supervised [4], [5], [9], [10], semi-supervised [5]–[8], or unsupervised [8] ML approaches. While the supervised ML approaches (e.g., ANN classifier) entail laborious labeling of OPM data and are only applicable for detecting fault patterns already in repository, the unsupervised approaches (e.g., clustering) can potentially detect unseen incidents as they directly mine the patterns of data. The defects of the latter relate to their limited scalability or accuracy performance. The semi-supervised ML approaches (e.g., one-class SVM) seek trade-offs between applicability and scalability by learning the boundaries of normal OPM data distributions. Such approaches eliminate the need for prior knowledge about fault patterns and meanwhile do not bring scalability concerns, but effective modeling of boundaries can be intractable in some cases.

SFI: upon detection of soft failures, SFI is usually invoked to facilitate subsequent localization and restoration of equipment where faults yield. Apart from the work in [1] which uses

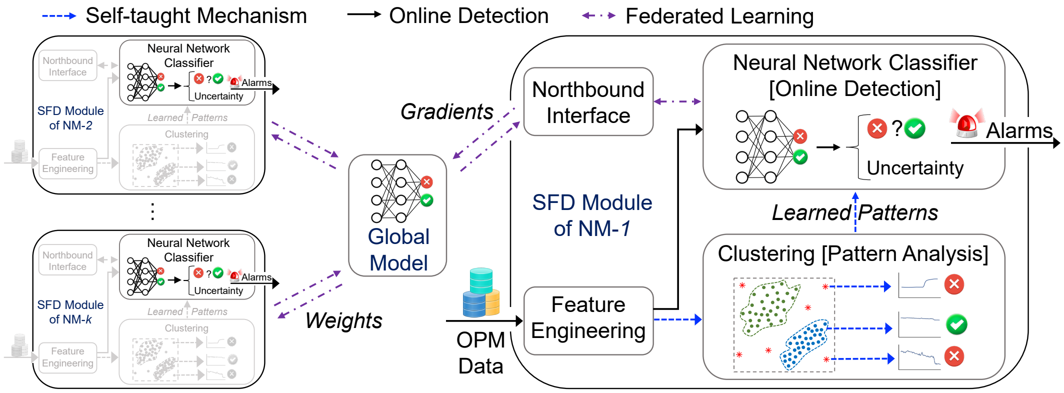

Fig. 2. Schematic of hybrid-learning-based SFD design.

an analytical model, the references listed in Table I for SFI all employ supervised ML models, i.e., SVM or neural network-based classifiers. For instance, the authors of [11] trained one-dimensional convolutional neural networks for classifying four types of faults taking as input signal power spectrum densities.

SFL: conventional solutions for fault localization typically adopt mathematical models to correlate alarms from different network locations or (probing) lightpaths. With the goal of reducing probes in localizing single-link failures, Panayiotou *et al.* proposed a joint heuristic and ML approach, where a graph-based correlation algorithm first finds a set of potential failed links and a Gaussian process classifier built on historical failure statistics further isolates the failures [12]. The proposed approach makes use of only the state (normal or abnormal) and routing information of each lightpath. In [13], the authors modeled the dependency between alarms as knowledge graphs and performed SFL with a graph neural network that explores correlations in the graphs. Besides, recent literature has reported neural-network-based designs that try to learn mappings from OPM data to fault locations without explicitly investigating the inherent correlations [9], [14].

IV. Soft Failure Detection with Hybrid Learning

In this section, we present a hybrid learning design that combines the merits of supervised, unsupervised, and cooperative ML for SFD. Fig. 2 shows the schematic of the design. Given a large amount of OPM data, the feature engineering block first extracts and formats relevant features readily to be processed by the ML blocks. Then, the unsupervised clustering block performs pattern analysis on the data by identifying a set of clusters and outliers. Each of the clusters is formed by samples sharing high similarities (quantified by certain distance metrics, e.g., Euclidean distance). We apply a density-based clustering algorithm called DBSCAN, for its advantage of detecting clusters of arbitrary shapes and sizes. DBSCAN initiates clusters with high-density samples (samples that have large numbers of neighboring samples) and repeatedly expands the clusters by adding the neighboring samples of those selected. Here, neighboring samples refer to samples that are within a predefined distance to a given sample, for instance, random fluctuations related to a normal state. Samples that cannot form clusters are left as outliers. Since soft failures lead to network states deviating from the normal ones (beyond their neighborhoods) and occur infrequently (i.e., correspond to just a few samples out of a large data set), it is natural that we label the outliers as indications of probable faults. Note that, to account for normal but rare or emerging network conditions, such as the deployment of a novel modulation format, a traditional SFD scheme can be employed for further inspection of the outliers. This way, the clustering block actualizes data-driven and generalized SFD, not relying on prior knowledge characterizing particular fault patterns. Nevertheless, the clustering block may suffer from scalability issues when applied to online detection because it needs to revisit the whole data set every time SFD is called, whose runtime complexity scales up with the size of the data set. Inspired by self-supervised learning methods that boost training of ML models utilizing knowledge learned for pretext tasks by unsupervised learning, we adopt a *self-taught* mechanism [15] to facilitate computationally efficient online SFD. Specifically, we train a neural-network-based binary classifier using the learned patterns (labeled data) and make it predict whether each newly collected sample is abnormal or not. In other words, we transfer knowledge from the clustering block to the neural network classifier. After being trained, the complexity of the classifier is only determined by the scale of the neural network (number of weight coefficients), often in the magnitude ranging from hundreds to a few thousand considering the moderate difficulty of SFD tasks.

A major challenge of training a successful classifier for SFD is the scarcity of abnormal samples, which can result in model overfitting to normal classes and thereby poor detection accuracy. To meet this challenge, we introduce evaluations of *model uncertainty* by applying the Bayesian neural network (BNN) technique. Different from regular neural networks that generate deterministic predictions, BNNs are probabilistic models aiming at learning the distributions of targets characterized by, for instance, means and standard deviations. Thus, BNNs allow us not only to make predictions but also to evaluate the uncertainties about predictions. When the BNN classifier is trained with mostly normal samples, it is less confident about its predictions (higher uncertainty) for data falling out of normal distributions (i.e., abnormal data). Based on this principle, we enhance the classifier by a threshold-

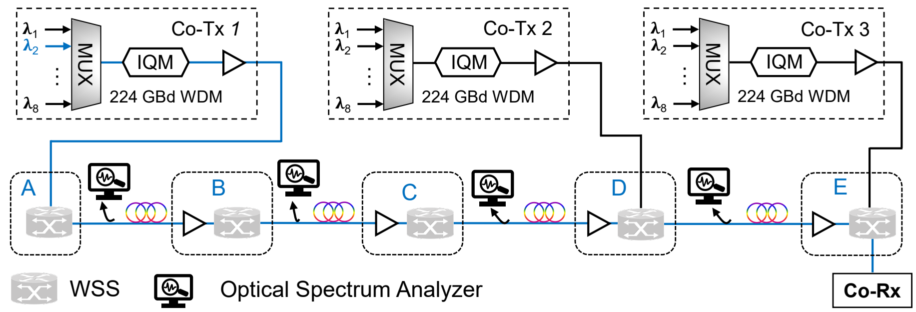

Fig. 3. System setup for data set generation. WDM: wavelength division multiplexing; Co-Tx: coherent transceiver; Co-Rx: coherent receiver; MUX: multiplexer.

based discriminator which claims a sample as abnormal if the prediction uncertainty is higher than a threshold, regardless of the actual prediction by the classifier. We automate the choice of the threshold for a given data set by counting the frequency of occurrence of prediction uncertainty and finding a point beyond which the uncertainty distribution becomes evidently sparse. Such a point can represent an approximate boundary of the uncertainty distribution of normal samples. In addition, since the choice of the threshold involves in a trade-off between fault detection and false alarm rates, that is, a lower threshold can improve fault detection rate but leads to larger numbers of false alarms, the maximum tolerable false alarm rate can serve as an auxiliary reference for determining the threshold. For instance, we can restrict the threshold to be higher than the 99^{th} percentile of the prediction uncertainty to bound false alarm rate by 1%.

To further mitigate the problem incurred by shortage of abnormal samples, the hybrid learning design encapsulates a *federated learning* framework by which NMs can train classifiers cooperatively while securing the confidentiality of each network domain. As illustrated by Fig. 2, the federated learning framework involves a hierarchical architecture where a global classifier model maintained by the broker plane lying in the higher control and management hierarchy coordinates the training of distributed models (owned by NMs) in multiple rounds. All the models employ an identical neural network structure. Note that, the framework does not necessarily require every NM to apply the self-taught mechanism, i.e., an NM can purely rely on a classifier for SFD. In each round of training after initialization, the distributed models are first synchronized to the global model (i.e., by downloading the model weights) and then execute independent training of multiple epochs using local data sets and standard training algorithms like *Adam*. Here, a training epoch refers to that a model is updated by traversing the whole data set once. Afterward, the distributed models upload the derived gradients to the global model, which concludes the round by aggregating the received gradients (e.g., averaging) and updating the model accordingly. Federated learning essentially facilitates knowledge sharing among NMs while protecting the data integrity of domains because only model weights and gradients are exchanged. This can be especially beneficial in the case where NMs possess data that are nonidentically distributed (i.e., complementary to each other).

V. Performance Assessment

We assessed the performance of the hybrid learning design with data sets collected using the VPItransmissionMaker™ Optical Systems simulator. Fig. 3 shows the five-node system setup. Each adjacent node pair is connected by a 100-kilometer standard single-mode fiber, with the fiber loss compensated by a gain-controlled amplifier that has a noise figure of 4 dB. The three coherent transceivers attached to nodes A, D and E were used to inject signals on eight wavelength channels at 224 Gbauds. We set up lightpaths from node A to E on λ_2 (the signals of interest) with various modulation formats (4QAM, QPSK, or 8PSK), symbol rates ($[9, 28]$ Gbauds), and launch power ($[0.22, 3.6]$ milliwatt). By varying the number of signals injected by each transceiver and adding perturbations (0.3 or 0.6 dB) to the gain of the amplifier in node C, we emulated the dynamic network condition each lightpath undergoes. We introduced one of the following soft failures at certain simulation time points: *i)* 10 dB reduced gain by the amplifier in node D (amplifier malfunctioning); *ii)* injection of a 3 milliwatt signal on λ_3 by transceiver 2 (jamming attack); *iii)* injection of a signal on λ_2 by transceiver 2 (signal overlapping due to misconfiguration); *iv)* drift of laser central frequency by half of channel spacing; and *v)* narrowing of lowpass filter bandwidth at the receiver side by 20% (tight filtering). The optical spectrum analyzers and the coherent receiver constantly monitor the signal power and BER values, respectively. We recorded the BER evolution (five consecutive BER values) and the power variation along the routing path with respect to every lightpath to represent both the temporal and spatial characteristics of the system. In total, $15,654$ normal samples and 30 abnormal samples were collected.

We implemented the DBSCAN algorithm with the related parameters tuned according to the method discussed in [15]. The BNN classifier was realized by the Monte-Carlo dropout method based on a fully-connected neural network architecture of two hidden layers (each consisting of ten neurons). The Monte-Carlo dropout method applies a dropout rate (0.1 used in the evaluations) at which a neuron is deactivated in each training/inference operation, thereby, produces a probabilistic classifier model. We evaluated the model uncertainties by computing the mutual information of 100 feed-forward calculations during each inference. By referring to the uncertainty distributions of testing samples, we decided a rough interval of the uncertainty threshold and tested different configurations

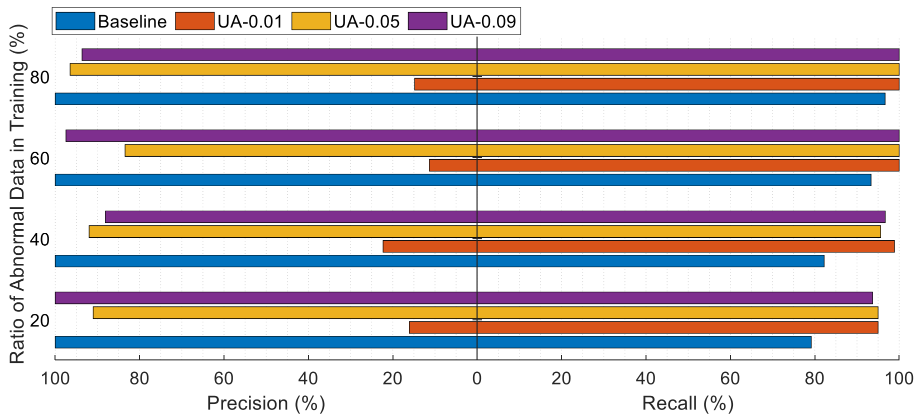

Fig. 4. Performance comparison between the uncertainty-aided (UA) design with different threshold setups and a baseline without uncertainty awareness.

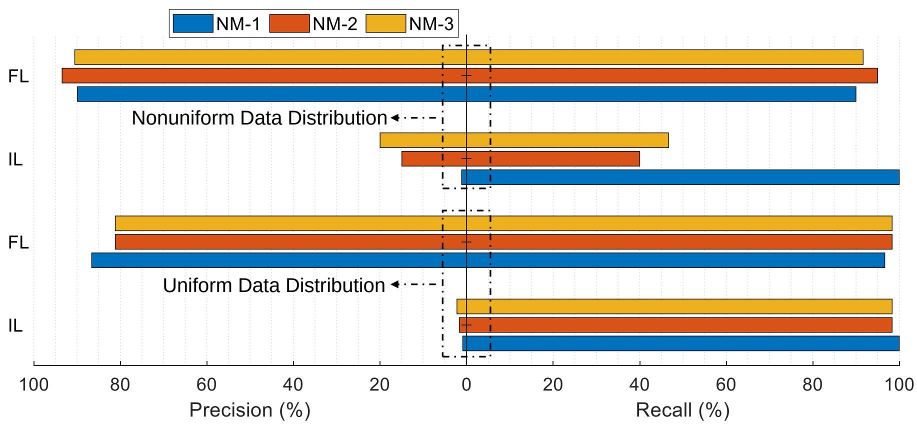

Fig. 5. Performance comparison between federated learning (FL) and independent learning (IL) for SFD.

to evaluate their impact on fault detection and false alarm rates. We first compared the proposed design with a baseline that adopts a regular neural network classifier to evaluate the benefit of incorporating uncertainty analysis in SFD. Fig. 4 shows the results of precision and recall as functions of the proportion of abnormal data used for training. Here, precision and recall are defined as the ratio of true positives (positives being successfully detected) to the number of claimed positives and the ratio of true positives to the total number of positives (i.e., detection rate) [10], respectively. It can be seen that the proposed design can achieve recall of $\sim 95\%$ with just 20% of abnormal data used for training, whereas that from the baseline is only 79% in this case. With 60% of abnormal data used for training, the detection rate from the proposed design reaches 100% under all the three threshold configurations. Meanwhile, we can observe that the choice of a lower threshold slightly improves the detection rate but leads to larger numbers of false alarms, i.e., lower precision. By switching the threshold from 0.01 to 0.05, we can increase precision from less than 25% to $83-96\%$ (corresponding to false positive rates of $0.01-0.04\%$). Overall, the proposed design promotes higher fault detection rates against the baseline, especially when the abnormal data for training are rare, at a reasonable cost in the increase of the number of false alarms.

Next, we assessed the performance of the federated learning framework (denoted by FL) assuming three NMs, each possessing a division of the original data set. To verify the robustness of FL against the distribution of data, we tested both uniform and nonuniform data division schemes. With the former, normal and abnormal data were evenly distributed to the NMs, whereas with the latter, we made NM-1 hold all the abnormal samples assigned for training (60% of the abnormal data set in both cases). We compared FL with an independent learning approach (denoted by IL) where NMs perform independent model training with local data sets. Fig. 5 presents the precision and recall results from the model of each NM under different data distributions. In the case where data are evenly divided, FL achieves recall comparable to that from IL but with over $40\times$ higher precision. In other words, IL sustains high detection rates at the cost of raising excessive false alarms. The performance of the models from different NMs are close in this case. The advantage of FL becomes more notable under nonuniform data distributions. Since only the model of NM-1 got trained with abnormal samples, the models of the other NMs with IL fail to detect soft failures to a large extent (with recall of just around 40%). On the other hand, by enabling knowledge sharing among NMs, the performance of FL remains stable, achieving $> 90\%$ detection rates for all the NMs while preserving high precision.

VI. Conclusion

In this article, we discussed an ML-aided automated fault management architecture for optical networks. We reviewed the state-of-the-art proposals based on this architecture and introduced a hybrid learning approach for soft failure detection to overcome the applicability and scalability issues from existing solutions. Performance assessment verified the effectiveness of the hybrid learning approach.

Open questions include but not limit to: *i)* how to decide the most effective set of optical parameters and the frequency to monitor for realizing desirable trade-off between model accuracies and costs; *ii)* how to further enhance the practicability of the unsupervised clustering block by more advanced algorithm designs and evaluations incorporating more comprehensive network conditions or real traces; and *iii)* how to achieve robust performance in soft failure identification and localization tasks when the number of fault types or potential fault locations is large and the distribution of data over the target classes is skewed.

Acknowledgments

This work was supported in part by the Key-Area Research and Development Program of Guangdong Province under Grant 2020B0101080002, NSFC Project U2001601, NSF award # ICE-T:RC 1836921, and NSFC Project 62035018.

References

[1] A. Vela, M. Ruiz, F. Fresi, N. Sambo, F. Cugini, G. Meloni, L. Potí, L. Velasco, and P. Castoldi, "BER degradation detection and failure identification in elastic optical networks," *J. Lightw. Technol.*, vol. 35, no. 21, pp. 4595–4604, Nov. 2017.

[2] S. Nina, M. Furdek, S. Zsigmond, and L. Wosinska, "Physical-layer security in evolving optical networks," *IEEE Commun. Mag.*, vol. 54, no. 8, pp. 110–117, Aug. 2016.

[3] D. Scano, F. Paolucci, K. Kondepu, A. Sgambelluri, L. Valcarenghi, and F. Cugini, "Extending P4 in-band telemetry to user equipment for latency- and localization-aware autonomous networking with AI forecasting," *J. Opt. Commun. Netw.*, vol. 13, no. 9, pp. D103–D114, 2021.

[4] D. Rafique, T. Szyrkowiec, H. Grießer, A. Autenrieth, and J. Elbers, "Cognitive assurance architecture for optical network fault management," *J. Lightw. Technol.*, vol. 36, no. 7, pp. 1443–1450, Apr. 2018.

[5] S. Shahkarami, F. Musumeci, F. Cugini, and M. Tornatore, "Machine-learning-based soft-failure detection and identification in optical networks," in *Proc. Conf. Opt. Fiber Commun.*, Mar. 2018, paper M3A.5.

[6] L. Shu, Z. Yu, Z. Wan, J. Zhang, S. Hu, and K. Xu, "Dual-stage soft failure detection and identification for low-margin elastic optical network by exploiting digital spectrum information," *J. Lightw. Technol.*, vol. 38, no. 9, pp. 2669–2679, 2020.

[7] H. Lun, X. Liu, M. Cai, Y. Wu, M. Fu, L. Yi, W. Hu, and Q. Zhuge, "GAN based soft failure detection and identification for long-haul coherent transmission systems," in *Proc. Conf. Opt. Fiber Commun.*, 2021, paper Th4J.2.

[8] M. Furdek, C. Natalino, A. Di Giglio, and M. Schiano, "Optical network security management: requirements, architecture, and efficient machine learning models for detection of evolving threats," *J. Opt. Commun. Netw.*, vol. 13, no. 2, pp. A144–A155, 2021.

[9] K. Abdelli, H. Griesser, P. Ehrle, C. Tropschug, and S. Pachnicke, "Reflective fiber fault detection and characterization using long short-term memory," *J. Opt. Commun. Netw.*, vol. 13, no. 10, pp. E32–E41, 2021.

[10] W. Du, D. Cote, C. Barber, and Y. Liu, "Forecasting loss of signal in optical networks with machine learning," *J. Opt. Commun. Netw.*, vol. 13, no. 10, pp. E109–E121, 2021.

[11] H. Lun, M. Fu, X. Liu, Y. Wu, L. Yi, W. Hu, and Q. Zhuge, "Soft failure identification for long-haul optical communication systems based on one-dimensional convolutional neural network," *J. Lightw. Technol.*, vol. 38, no. 11, pp. 2992–2999, 2020.

[12] T. Panayiotou, S. P. Chatzis, and G. Ellinas, "Leveraging statistical machine learning to address failure localization in optical networks," *J. Opt. Commun. Netw.*, vol. 10, no. 3, pp. 162–173, Mar. 2018.

[13] Z. Li, Y. Zhao, Y. Li, S. Rahman, F. Wang, X. Xin, and J. Zhang, "Fault localization based on knowledge graph in software-defined optical networks," *J. Lightw. Techn.*, vol. 39, no. 13, pp. 4236–4246, 2021.

[14] K. S. Mayer, J. A. Soares, R. P. Pinto, C. E. Rothenberg, D. S. Arantes, and D. A. A. Mello, "Machine-learning-based soft-failure localization with partial software-defined networking telemetry," *J. Opt. Commun. Netw.*, vol. 13, no. 10, pp. E122–E131, 2021.

[15] X. Chen, B. Li, R. Proietti, Z. Zhu, and S. J. B. Yoo, "Self-taught anomaly detection with hybrid unsupervised/supervised machine learning in optical networks," *J. Lightw. Technol.*, vol. 37, no. 7, pp. 1742–1749, April 2019.

Xiaoliang Chen received the Ph.D. degree from the University of Science and Technology of China in 2016. He is an Associate Professor at the Sun Yat-Sen University. Prior to joining the Sun Yat-sen University, he was a postdoctoral researcher at the University of California, Davis (UC Davis). His research interests include optical networks, optical interconnection architectures for data centers, software-defined networking, and cognitive networking.

Che-Yu Liu received the B.S. degree in electrical and computer engineering from National Chiao Tung University in 2012, and the M.S. degree in electrical and computer engineering, UC Davis, in 2016. He is currently working toward the Ph.D. degree in computer science, UC Davis. His research interest focuses on optical networks, software-defined networking, and machine learning.

Roberto Proietti received the M.S. degree in telecommunications engineering from the University of Pisa in 2004, and the Ph.D. degree in electrical engineering from Scuola Superiore Sant Anna in 2009. He was a Project Scientist with the Next Generation Networking Systems Laboratory, UC Davis. He is currently an Assistant Professor with Politecnico di Torino. His research interests include optical switching technologies and architectures for supercomputing and data center applications, high-spectral-efficiency coherent transmission systems, and elastic optical networking.

Zhaohui Li received the B.S. degree from the Department of Physics and M.S. degree from the Institute of Modern Optics, Nankai University, in 1999 and 2002, respectively, and the Ph.D. degree from the Nanyang Technological University in 2007. He is working with the School of Electronics and Information Technology, Sun Yat-sen University, as a Professor, since 2009. His research interests are optical communication systems, optical signal processing technology and ultrafine measurement systems.

S. J. Ben Yoo is a Distinguished Professor at UC Davis. His research at UC Davis includes 2D/3D photonic integration for future computing, cognitive networks, communication, imaging, and navigation systems, micro/nano systems integration, and the future Internet. Prior to joining UC Davis in 1999, he was a Senior Research Scientist at Bellcore, leading technical efforts in integrated photonics, optical networking, and systems integration. Prior to joining Bellcore in 1991, he conducted research on nonlinear optical processes in quantum wells, a four-wave-mixing study of relaxation mechanisms in dye molecules, and ultrafast diffusion-driven photodetectors at Stanford University (BS 84', MS 86', PhD 91', Stanford University). Prof. Yoo is Fellow of IEEE, OSA, NIAC and a recipient of the DARPA Award for Sustained Excellence (1997), the Bellcore CEO Award (1998), the Mid-Career Research Faculty Award (2004 UC Davis), and the Senior Research Faculty Award (2011 UC Davis).