AI/ML-Based State-of-Polarization Monitoring in Optical Networks: Concepts and Challenges

Citation for the original published paper (version of record):
Sadighi, L., Natalino Da Silva, C., Karlsson, S. et al (2025). AI/ML-Based State-of-Polarization Monitoring in Optical Networks: Concepts and Challenges. Optical Fiber Communication Conference (OFC) 2025

(article starts on next page)

AI/ML-Based State-of-Polarization Monitoring in Optical Networks: Concepts and Challenges

Leyla Sadighi[1,*], Carlos Natalino[1], Stefan Karlsson[2], Lena Wosinska[1], Marco Ruffini[3], and Marija Furdek[1]

[1] *Department of Electrical Engineering, Chalmers University of Technology, 412 96 Gothenburg, Sweden.*
[2] *Swedish Defense Material Administration, 586 63 Linköping, Sweden.*
[3] *School of Computer Science and Statistics, CONNECT Centre, Trinity College Dublin, Ireland.*

[*] *sadighi@chalmers.se*

Abstract: Optical networks are vulnerable to various disturbances that can jeopardize service availability or privacy. We discuss AI/ML-based analysis of the incurred state-of-polarization changes for cognitive management of complex disturbances.

1. Introduction

Optical networks form critical communication infrastructure that supports the majority of digital services in today's society. Their critical role requires very high reliability, security and resilience. These are jeopardized by diverse human activities such as construction work leading to accidental fiber cuts, deliberate sabotage or eavesdropping attempts. Early detection of and reaction to various external disturbances is important to ensuring the high performance and survivability of network services. The recent proliferation of artificial intelligence & machine learning (AI/ML) techniques capable of observing intricate changes and trends in optical performance monitoring data has unleashed a tremendous potential for novel diagnostic capabilities. Combined with the advancements in network monitoring techniques, AI/ML drives new use cases related to sensing various environmental changes and disturbances to the optical network.

Traditionally, optical time-domain reflectometer (OTDR) has been used in optical networks as a means to detect and localize anomalies in fiber links. This approach relies on expensive and specialized hardware to detect changes in the optical signal power. Coherent OTDR represents a significant advancement over traditional OTDR by employing coherent detection to capture both the amplitude and phase of backscattered light. A more advanced approach is distributed acoustic sensing (DAS) which leverages Rayleigh backscattering in optical fibers to achieve high spatial resolution for detecting mechanical vibrations along the fiber length. While this technology is highly sensitive and precise, its implementation is costly and presents significant challenges, particularly in integration with legacy optical networks [1]. In recent years, substantial efforts have been made to utilize pre-installed optical fibers for sensing environmental changes [2]. The state of polarization (SOP) represents the orientation of a light wave's electric field as it propagates through an optical fiber and comprises values such as Stokes parameters (S_0, S_1, S_2, S_3), which define the light's polarization. Its high sensitivity to external perturbations, often attributed to stress-induced birefringence [3], makes the SOP a key property for environmental sensing. As a result, it has emerged as a promising technique for monitoring optical fiber networks, garnering significant attention from academia and industry. This solution is particularly interesting due to the possibility of using lightly-modified transceivers to perform the SOP data collection, with promising results reported upon applying AI/ML for the analysis of such data [4, 5].

Despite the high demonstrated potential, widespread adoption of AI/ML-based SOP monitoring in optical networks faces several challenges. This paper overviews the basics of SOP data collection, analysis, and AI/ML processing and discusses several challenges of carrier-grade network-wide SOP monitoring deployment.

2. AI/ML-Based State-of-Polarization Monitoring in Optical Networks

Fig. 1 presents a network architecture where traditional monitoring data, e.g., coming from reconfigurable optical add-drop multiplexers (ROADMs) and transceivers, is collected in conjunction with SOP data. The data can be collected using standard protocols discussed in the literature and available in commercial products. The figure depicts two channels for which SOP data is collected with the aid of an SOP analyzer co-located with a ROADM at the destination node. The gathered SOP data are sent to the network management system (NMS) and stored in a telemetry database, usually implemented as a time-series database to allow queries of network parameter evolution over time. Since the telemetry database stores information about all the SOP channels in the network, it is possible to, in addition to analyzing them individually, correlate the SOP effects collected from multiple channels.

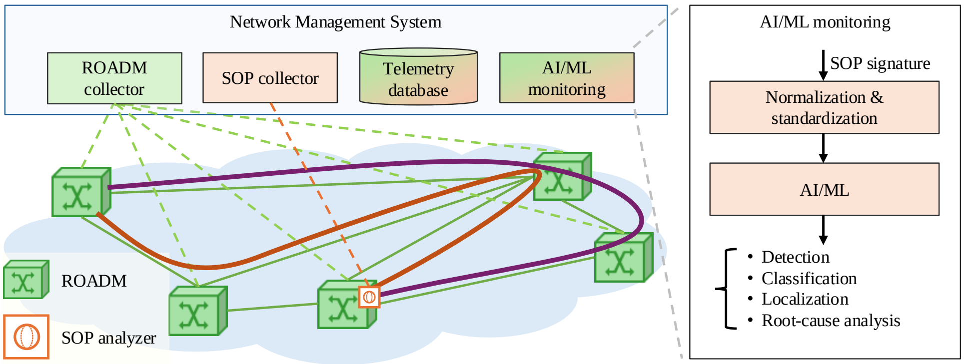


Fig. 1: Network architecture with AI/ML monitoring based on traditional reconfigurable optical add-drop multiplexer (ROADM) and state of polarization (SOP) data collection.

The stored SOP data is analyzed by the AI/ML monitoring module that can, depending on the exact applied technique, perform various functions such as, e.g., detection of anomalies, their classification, localization, or root cause analysis. As detailed in the right-hand side of the figure, pre-processing procedures such as normalization and standardization can be applied to adjust the data to the most suitable format for AI/ML processing.

If AI/ML techniques for anomaly detection are applied, the output of the model is a binary variable that indicates whether the provided sample(s) represent an anomaly. This is commonly implemented by using an unsupervised learning algorithm, which does not require prior training or knowledge about the normal or abnormal operating conditions. Deeper insights into the nature of detected anomalies can be obtained by, e.g., root cause analysis [6]. Although the detection of abnormal events already enables several use cases for using SOP data, a more detailed output of the AI/ML model is often desired. This can be achieved by using supervised learning, e.g., a multi-class classifier, which is able to determine the precise category of an anomalous event based on the analyzed SOP data. However, this is only possible if the model is provided with a dataset containing sufficient samples from each category of events that need to be classified, which may be costly or infeasibile.

Finally, another aspect that can be useful is the localization of the event, which can be characterized in several ways. Firstly, *topological* localization refers to the ability of determining which fiber link in the topology is the source of the detected event. Since SOP data is collected only at the receiver, topological localization depends on the physical routes of lightpaths used to collect the SOP data. Secondly, *spatial* localization defines the exact point of an event along the fiber link (e.g., how many kilometers from its origin). SOP lacks spatial resolution in forward propagation alone, as events can occur anywhere along the fiber. However, spatial localization is possible by propagating a reverse signal and correlating the forward and reverse signals in time, requiring tightly synchronized receivers. The precision of spatial resolution depends on the synchronization accuracy and other factors. The final localization aspect refers to the *temporal* dimension, i.e., determining the start (and possibly the duration) of the detected event, which depends on the monitoring frequency.

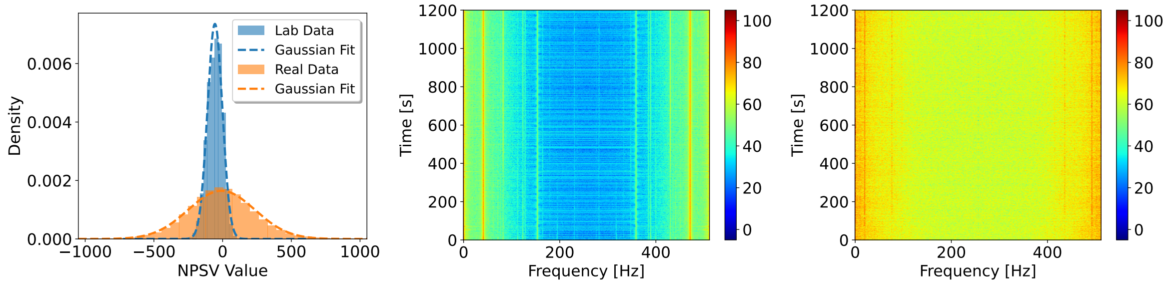
3. Open Challenges

The use of SOP data as an enabler of using the optical network as a sensor for environmental disturbances. However, several remaining challenges should be addressed in order to make the technology suitable for widespread adoption. In the following, we mention a few of the critical challenges.

Data processing: The processing of the SOP data for their use in AI/ML models comprises several steps that involve the detection of analog optical signal, its conversion to the digital domain, the computation of numerical polarization state variation (NPSV), its mapping to the Poincaré sphere, etc. As shown in Fig. 1, part of this processing needs to be done at the receiving node, but other parts can be performed at the SOP collector. One of the challenges in this context is to define at which node (local or central) to carry out certain parts of the processing, which can reveal interesting trade-offs in terms of cost, telemetry efficiency, etc. Another challenge to be considered is related to the selection of the data format to be used as an input to the AI/ML model, where different formats may incur a different number of SOP processing steps.

Data collection: Data collection is another challenging aspect whenever the use of AI/ML models is considered, with a strong correlation with the training procedures. There are three aspects related to this challenge. Firstly, it is important to collect data that can be used for the appropriate training procedures. For instance, if multiple events need to be categorized, the collected data must follow the same categorization. Otherwise, if the AI/ML will only perform anomaly detection, a simpler categorization of the collected data can be performed. Secondly, the routing of the channels used for SOP monitoring need to be carefully computed. For instance, in Fig. 1, two link-disjoint channels are used for collecting the SOP data, covering six out of the seven fiber links in the topology, and having the SOP analyzer at a single node. For larger topologies, defining paths that cover a representative set of links at a minimal cost while allowing for complete event disambiguation becomes a challenging problem. Finally, while the collection of traditional monitoring data (e.g., Q-factor, optical signal-to-noise ratio) has been addressed by several protocols in the literature and industry, the protocol and the data format to be adopted for transferring SOP data from the SOP analyzer to the NMS remains an open question that deserves investigation.

AI/ML model training: In optical networks, it is expected that each component will incur a distinct effect on the SOP data. Therefore, it is unlikely that the same AI/ML model can be used for different paths/sections of the network. This means that a specific AI/ML model will be needed to analyse SOP of each channel in the network. An example of this behavior is illustrated in Fig. 2, which shows the data related to the same event signature (i.e., 80 Hz vibration) over a lab and a real-world deployment. More details can be found in [7]. Fig. 2(a) shows the density of data values in both collected datasets, which make it evident that the data collected from the real-world testbed has substantially more noise than the data collected from the lab. Figs. 2(b) and (c) show the waterfall diagrams from the two deployments, respectively, which depict the intensity of the SOP changes across frequency and time. We can observe that the data collected from the lab, Fig. 2(b), presents a stronger distinction over frequency than the one from the real-world testbed, Fig. 2(c). Moreover, optical fiber infrastructures are always evolving, and the SOP data needs to be continuously analyzed to detect substantial changes such that AI/ML models can be re-trained if necessary.



(a) Data distribution (b) Waterfall from the lab (c) Waterfall from the real-world testbed

Fig. 2: Illustrative response of 80 Hz vibrations collected at a lab and at a real-world deployment.

Performance trade-offs and targets: A critical challenge when designing a monitoring system is to define how often the monitoring is performed. This aspects defines the overhead of monitoring data transmission in the network, the amount of processing capacity needed to obtain the solution, and the speed of detecting an event. A default assumption that the more frequent monitoring is, the better needs to be revisited when considering AI/ML adoption. AI/ML models can be characterized by metrics such as false positive and false negative rates. These rates are evaluated depending on the number of inferences, i.e., the number of analyzed samples, implying that they scale with the frequency of the data analysis. The more frequently the data is analyzed, the more frequently the system will need to handle incorrect outputs from the AI/ML model. Thus, it is important to consider not only the performance target of AI/ML models, but also the frequency at which these models will be used, in order to clearly evaluate their impact on the overall operational overhead of the network. For instance, a model with a false positive and a false negative rates equal to 0.01 will experience on average two mistakes for every 100 samples evaluated, one false positive and one false negative. If the network is monitored every second, two mistakes will need to be handled every 100 seconds, and this scales linearly with the frequency of monitoring.

4. Conclusion

AI/ML-based SOP monitoring can enable real-time detection of various disturbances affecting optical network infrastructure. Challenges related to the data collection and processing, adaptability to evolving threats, and the computational overhead of real-time AI model deployment need to be overcome through, among other, advanced data gathering, model optimization and integration with NMS to enhance service availability and integrity.

Acknowledgment: This work was supported by the Swedish Research Council (2023–05249), the European Commission's Digital Europe Programme (101127973) through the 5G-TACTIC project, and CONNECT center, OpenIreland, funded by the Science Foundation Ireland grants 13/RC/2077_p2 and 18/RI/5721.

References

1. Z. He and Q. Liu, J. Light. Technol. **39**, 3671–3686 (2021). DOI: 10.1109/JLT.2021.3059771.
2. G. Marra *et al.*, Science **376**, 874–879 (2022). DOI: 10.1126/science.abo1939.
3. S. Pellegrini *et al.*, IEEE Photonics J. **15**, 1–9 (2023). DOI: 10.1109/JPHOT.2023.3281670.
4. A. Tomasov *et al.*, Opt. & Laser Technol. **167**, 109668 (2023). DOI: 10.1016/j.optlastec.2023.109668.
5. A. Tomasov *et al.*, in *Proc. of CLEO,* (2024), p. JTu2A.217. DOI: 10.1364/CLEO_AT.2024.JTu2A.217.
6. C. Natalino *et al.*, IEEE Trans. on Netw. Serv. Manag. **19**, 2702–2713 (2022). DOI: 10.1109/TNSM.2022.3198139.
7. L. Sadighi *et al.*, in *Proc. of ECOC,* (2024), p. Tu4E.5.