Intelligent Detection of Overlapping Fiber Anomalies in Optical Networks Using Machine Learning

Intelligent Detection of Overlapping Fiber Anomalies in Optical Networks Using Machine Learning

Gulmina Malik
Politecnico di Torino, Italy
gulmina.malik@polito.it

Imran Chowdhury Dipto
Politecnico di Torino, Italy
imran.dipto@polito.it

Muhammad Umar Masood
Politecnico di Torino, Italy
muhammad.masood@polito.it

Mashboob Cheruvakkadu Mohamed
Politecnico di Torino, Italy
mashboob.cheruvakkadu@polito.it

Stefano Straullu
Links Foundation, Italy
stefano.straullu@linksfoundation.com

Sai Kishore Bhyri
Nokia
sai.bhyri@nokia.com

Gabriele Maria Galimberti
Nokia
gabriele.galimberti@nokia.com

João Pedro
Nokia
joao.pedro@nokia.com

Antonio Napoli
Nokia
antonio.napoli@nokia.com

Walid Wakim
Nokia
walid.wakim@nokia.com

Vittorio Curri
Politecnico di Torino, Italy
vittorio.curri@polito.it

***Abstract*—We propose a machine learning approach leveraging state-of-polarization dynamics to detect overlapping fiber anomalies. Simulated disturbances and XGBoost classification achieve near-perfect accuracy under noise, enabling precise identification of concurrent events and enhancing both fault detection and physical layer security in optical communication networks.**

***Index Terms*—Machine learning, XGBoost, state of polarization, optical fiber, fiber anomalies.**

I. Introduction

Fiber optic networks form the backbone of global telecommunications, enabling high-speed data transmission with minimal latency. However, as these networks evolve, they become increasingly susceptible to overlapping anomalies such as bends, breaks, or splice losses. These overlapping disturbances complicate identification, necessitating more sophisticated detection techniques [1]. Traditional techniques, such as Optical Time-Domain Reflectometry (OTDR) and threshold-based monitoring, rely on predefined heuristics and struggle to differentiate the overlapping anomalies, particularly when multiple fault signatures interact [1, 2]. Recent advancements in machine learning (ML) have enabled more sophisticated anomaly detection in optical networks. For instance, in [3], a vision transformer-based model was introduced to identify and locate simultaneous anomaly occurrences; however, it requires a large amount of processing power for training and inference. These challenges highlight the need for intelligent real-time anomaly detection systems that can take preventive measures without compromising the network integrity and economic crisis.

Among ML-based approaches, state-of-polarization (SOP) monitoring has emerged as a promising technique for detecting mechanical disturbances in optical fibers. SOP represents the orientation of the optical signal's electric field as it propagates through the fiber. Changes in SOP can indicate external disturbances, making it a valuable metric to detect fiber anomalies. The SOP angular speed (SOPAS) further defines the rate of change in polarization states (Stoke parameters) on the Poincar´e sphere at a certain angle. This allows for a more granular capture of the disturbances on the fiber, such as bending, shaking, and tapping [1]. However, current SOP-based methods are not very granular in their classification of overlap anomalies, and they frequently do not differentiate between hostile intrusions (such as eavesdropping) and harmless environmental vibrations (such as wind). By modeling complicated concurrent events using polarization signatures, ML provides a revolutionary solution [4].

This study investigates ML-driven SOP analysis for detecting overlapping anomalies in optical networks. Leveraging XGBoost [5], our proposed model achieves high accuracy in classifying overlapped fiber anomalies. The results validate the effectiveness of ML for proactive network maintenance and enhanced security of fiber infrastructure, surpassing traditional fault detection methods in precision and reliability.

II. Polarization-based Anomaly Detection Setup

Every disrupting event has a distinct polarization signature. We have generated several signatures by using Stokes parameters to monitor the polarization state change on the Poincarè sphere. To emulate real-world perturbations, we reproduced three mechanical events: shaking, bending, and fiber hitting, using an Arduino-controlled robotic arm, programmed to generate precise and repeatable perturbations as described in [6]. The experimental setup consists of two G.652 standard single-mode fibers (SSMF) spanning 13 kilometers. A continuous wave of light pulses of 1530 nm is injected into the sensing fiber, and at the receiving end, a Novoptel PM1000 polarimeter

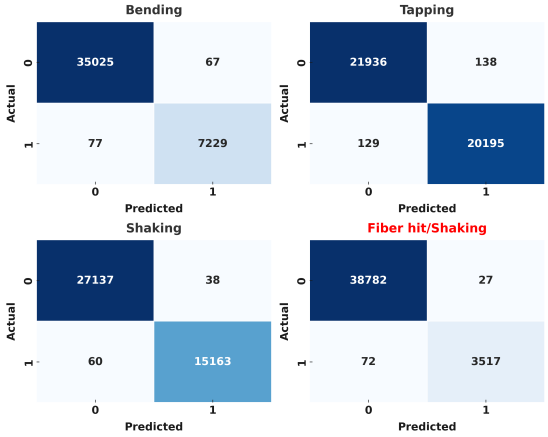

Fig. 1. Confusion metrics of XGBoost for different events.

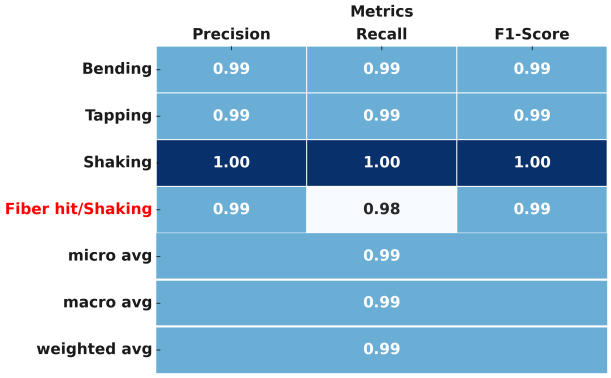

Fig. 2. Classification report of XGBoost for different events.

is connected, which tracks the polarization signatures on the Poincarè sphere. The robotic arm movements are pre-programmed, ensuring accurate emulation of real-world disturbances.

Fiber bending is characterized as eavesdropping by manipulating the exposed fiber, using a commercial optical fiber identifier (OFI) with a bend diameter of 0.25 mm. The handgrip controls the clamping; we clamp for 5 seconds before releasing. An eavesdropping event is generated when the device detects light leakage. The same test is repeated 10 times to ensure consistency. For shaking, the robotic arm moves the fiber up and down with a frequency of 3 Hz and an angle of deviation 90°. For fiber hit, we manually hit the fiber with 1 tap/s. Each of these events is recorded as Stokes parameters, reflecting their unique fingerprint. However, for the generation of overlapping events, the coordinated action of fiber shaking, by the robotic arm, and fiber hit is synchronized.

In order to ensure improved safety protocols and proactive maintenance, we used ML algorithms to precisely categorize and forecast the true nature of overlapping occurrences, even amidst noisy disturbances.

III. ML Model Performance in Anomaly Detection

The experimental results provide a demonstration of the effectiveness of the proposed ML model in accurately identifying anomalous events that could potentially compromise the optical fiber infrastructure. To simulate a real-world situation, feature engineering techniques such as adding differences of Stokes parameters $\{S_1, S_2, S_3\}$ and rolling mean of the parameters are used. To ensure the models are resilient, we included some environmental noise in the data. This helps the models to acquire features instead of memorizing the training data.

80% data is used for training and the remainder 20% is reserved for testing. We tested various classifiers like random forest, XGBoost, decision tree, and logistic regression. The findings indicate that XGBoost outperformed other classifiers in terms of accuracy scores since it can predict non-linearities and requires less processing time; thus, we selected it for further analysis. Further breakdown of XGBoost model performance is shown in Figure 1, where it is clear that the model has generalized data and made accurate predictions with only 2.06% of overlap events misclassified.

The metric scores achieved by XGBoost are illustrated in Figure 2 where the micro average aggregates all true and false positives, and false negatives across all the classes before computing metrics. Macro average computes the metric for each of the classes separately and then averages them equally, whereas the weighted average computes the metric of each class separately then takes a weighted sum based on the number of true instances per class. This Figure 2 shows that the model was able to predict "Bending" and "Tapping" with 99% Precision, Recall and F1-Scores and it was able to predict "Shaking" with a metric score of 100%. The model's ability to forecast the multi-event "Fiber hit/shaking" with a precision of 99%, further demonstrates its predictive capability.

Real-time response to multi-threat scenarios is made possible by the model's ability to reliably distinguish between benign shaking and hostile intrusions (like tapping) through training on overlapping tampering and noise-augmented data. This provides a strong foundation to protect networks from dynamic, interconnected faults. Future studies will investigate the combination of adaptive ML models that dynamically enhance performance as they come across fiber anomalies with real-time data streams.

Acknowledgment

This work has been supported from the project PNRR-NGEU (MUR–DM117/2023), and from the EU's Horizon Europe research and innovation program under GA No. 101092766 (ALLEGRO Project) and DN NESTOR GA No. 101119983

References

[1] K. Abdelli *et al.*, "Machine-learning-based anomaly detection in optical fiber monitoring," *Journal of optical communications and networking*, vol. 14, no. 5, pp. 365–375, 2022.

[2] S. Pellegrini *et al.*, "Overview on the state of polarization sensing: application scenarios and anomaly detection algorithms," *Journal of Optical Communications and Networking*, vol. 17, no. 2, pp. A196–A209, 2025.

[3] K. Abdelli *et al.*, "Anomaly detection and localization in optical networks using vision transformer and sop monitoring," in *Optical Fiber Communication Conference*, pp. Tu2J–4, Optica Publishing Group, 2024.

[4] L. Sadighi *et al.*, "Machine learning-based polarization signature analysis for detection and categorization of eavesdropping and harmful events," in *2024 Optical Fiber Communications Conference and Exhibition (OFC)*, pp. 1–3, IEEE, 2024.

[5] C. Zhang *et al.*, "Cause-aware failure detection using an interpretable xgboost for optical networks," *Optics Express*, vol. 29, no. 20, pp. 31974–31992, 2021.

[6] G. Malik *et al.*, "Machine learning for predictive multi-event detection in fiber optic systems," in *International Conference on Machine Learning and Communication Networks (ICMLCN)*, 2025.