Machine Learning-Based Polarization Signature Analysis for Detection and Categorization of Eavesdropping and Malicious Attacks

Leyla Sadighi, Stefan Karlsson, Marija Furdeck

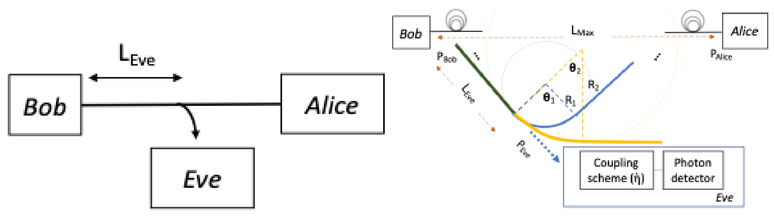
Department of Electrical Engineering, Optical Networks Unit

Abstract

This experimental study enhances optical network security through detecting and categorizing eavesdropping and malicious attacks by analyzing polarization state changes data with supervised machine learning classifiers. Notably, the XGBoost classifier reached an impressive 92.29% accuracy.

Introduction

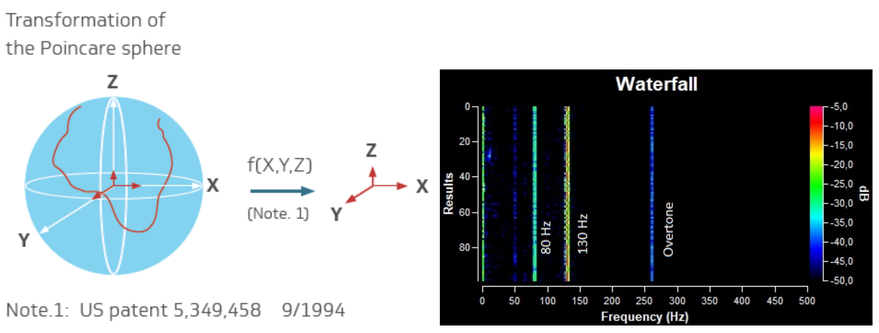
Addressing security threats in fiber optic-based infrastructures is susceptibility crucial due to their increasing adoption and to undetectable eavesdropping and malicious acts. Recent years have marked an increase in sabotage attempts on these systems, alongside the ever-present risk of unauthorized data interception, which is exacerbated by advances in computational and quantum computing. Optical fibers are particularly vulnerable to eavesdropping attacks, wherein unauthorized light coupling techniques can intercept data.



We focused on enhancing the security of fiber optical networks by using machine learning on concluded signatures from the binary data collected from our devices. Our research presents promising results in detecting and classifying eavesdropping and melicious attacks, thereby bolstering optical network security.

Polarization State changes

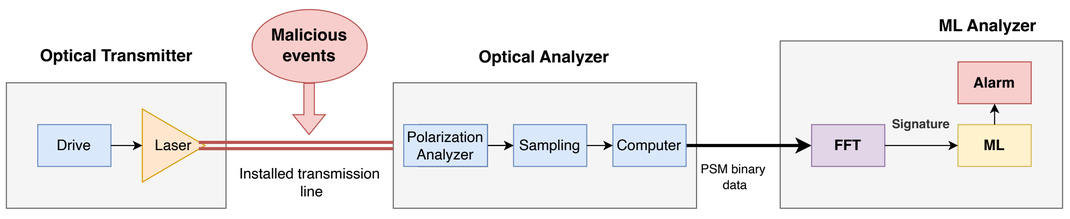
- **Signature:**a signature is defined as a sequence of power changes in a specific time and frequency derived after processing the polarization state change data.
- **Waterfall:** a waterfall is a unique plot assigned to its equivalent specific signature where the horizontal and vertical axes denote frequency and time, respectively.
- This demonstration will show how different security threats can be detected and categorized by observing the changes in polarization state in the transmitted light by using machine learning techniques, achieving more precise identification of unauthorized activities than previously visualization-based and coherent detection approaches[1, 2, 3] while minimizing false alarms.



Methodology

Our research employs Machine Learning (ML) algorithms to analyze the polarization signatures derived from real experiments. Our innovative approach has the capability to sense the difference between the different signatures of eavesdropping and malicious attacks of three different cable types. We successfully segregate signatures indicative of malicious activities from those attributable to benign factors, thereby fortifying system security.

Proposed Scheme



Data Collection

- Our collected data for ML analysis encompasses signatures of 13 distinct experimental settings, designed to provide an in-depth analysis of malicious acts across three cable types in different conditions. It contains the signatures of: 1)relaxed case(rlx), which represents normal transmission line signature in the absence of vibrations; 2) bending in FOCS cable (b-fc) and 3) bending in bare fiber(b-bf), designed to help distinguish between FOCS and bare fiber signatures, especially during the manipulation of the outer layer; Additionally, scenarios 4 through 9 address vibration-related variations based on frequency and cable type. Specifically, for FOCS, scenarios 4 and 5 differentiate between normal vibrations at 155 Hz (v-n-fc) and abnormal vibrations at 80 Hz (v-an-fc). For indoor cables, scenarios 6 and 7 characterize normal (v-n-id) and abnormal (v-an-id) vibration signatures. Scenarios 8 and 9 pertain to bare fiber, identifying normal (v-n-br) and abnormal (v-an-br) vibrations, respectively. Additional considered scenarios focus on the interplay of bending and vibrations in indoor cables: 10) Bending + 130 Hz(b-n-v-id), where the cable is flexed over a 10 mm diameter rod with a 130 Hz vibration; 11) Bending + 80 Hz(b-an-v-id), using a similar setup but with an 80 Hz vibration; 12) Bending + dual-frequency vibrations, incorporating both abnormal and normal vibrations at 80 Hz and 130 Hz (b-dl-v-id); and 13) relaxed scenario with dual-frequency vibrations (rlx-dl-v-id), which is the same as scenario 12 but without bending.
- After collecting signatures from these 13 distinct scenarios, we randomly segmented each signature into 70% training (840 points) and 30% testing (360 points) during the data preparation phase. Subsequently, these subsets merged into a unified dataset, which was shuffled to mitigate sequence bias. This led to a training dataset comprising $10,920$ samples and a test set of $4,680$ samples.

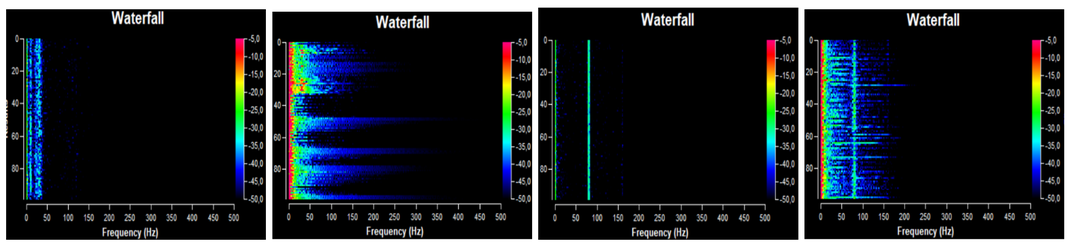
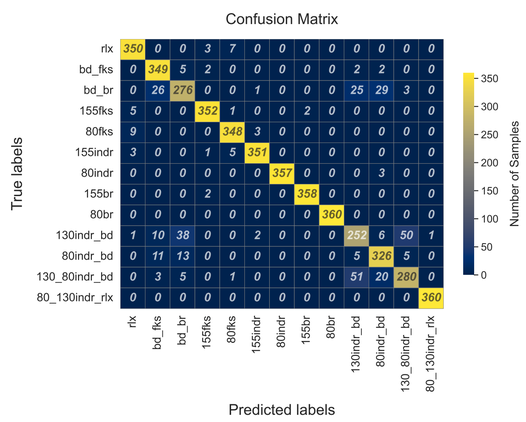


Figure 1: An example of waterfalls of the collected data in 4 different conditions on indoor cable(from left to right) :relaxed case, bending, 80 Hz vibration,bending plus 80 Hz vibration

Results

- We undertook an exhaustive selection process to identify the most suitable classifier for our machine learning task for this specific dataset type, evaluating ten distinct classifiers.
- The classifier adeptly categorized relaxed data and relaxed data plus vibrations on 80 Hz and 130 Hz for indoor cable.

Classifier Name	Accuracy	F1 score
XGBoost	0.923	0.922
Random Forest	0.911	0.909
Gradient Boosted Trees	0.906	0.904
Bagging with Decision Trees	0.878	0.877
Desision Tree	0.826	0.822
SVM	0.823	0.822
Linear Discriminant Analysis	0.821	0.819
Logistic Regression	0.816	0.814
K-NN	0.788	0.779
MLP	0.745	0.744



Conclusions

- By adeptly identifying intricate patterns and subtle malicious indicators in network traffic data, our approach presents a significant advancement in optical network security, offering a promising avenue for safeguarding sensitive data transmitted through fiber optic infrastructures.
- Through an in-depth analysis signatures of polarization state changes data from optical devices, we successfully employed ML techniques, specifically the XGBoost Classifier, to detect and categorize eavesdropping and malicious attacks with an impressive accuracy.

Future Works

- Future works can focus on recording datasets in real-life network installations and detecting real-life eavesdropping and malicious attempts.

References

[1] Y. Aono, E. Ip, and P. Ji, *More than communications: environment monitoring using existing optical fiber network infrastructure*, in *Optical Fiber Communication Conference*, W3G–1, 2020, Optica Publishing Group.

[2] S. Karlsson, R. Lin, L. Wosinska, and P. Monti. *Eavesdropping G. 652 vs. G. 657 fibres: a performance comparison*, in *ONDM 2022*, pp. 1–3, 2022. IEEE.

[3] S. Karlsson, M. Andersson, R. Lin, L. Wosinska, and P. Monti. *Detection of abnormal activities on a SM or MM fiber*, in *OFC 2023* , pp. 1–3, 2023. IEEE.