Root Cause Analysis for Autonomous Optical Networks: A Physical Layer Security Use Case

Root Cause Analysis for Autonomous Optical Networks: A Physical Layer Security Use Case

Carlos Natalino[(1,*)], Andrea Di Giglio[(2)], Marco Schiano[(2)], Marija Furdek[(1)]

[(1)] Department of Electrical Engineering, Chalmers University of Technology, Sweden carlos.natalino@chalmers.se
[(2)] Telecom Italia, Turin, Italy.
[(*)] The source code of this work is available at `https://github.com/carlosnatalino/2020-ECOC-RCA`

Abstract

To support secure and reliable operation of optical networks, we propose a framework for autonomous anomaly detection, root cause analysis and visualization of the anomaly impact on optical signal parameters. Verification on experimental physical layer security data reveals important properties of different attack profiles.

Introduction

Optical networks are characterized by costly and highly human-dependable operation[1]. Optical Performance Monitoring (OPM) and the detection of anomalies (e.g., caused by physical layer intrusions or device degradation) are key tasks during optical network operation. The detection of anomalies is particularly challenging to be performed by humans or analytical models, and Machine Learning (ML) models for anomaly detection have shown promising performance[2],[3]. However, ML-based anomaly detection (enabled mainly by semi- or unsupervised learning) typically only detects the anomaly, without determining its causes. To achieve trustworthy and agile optical network operation, investigating the cause of a detected anomaly is pivotal for triggering appropriate and effective countermeasures.

Anomaly cause investigation is a complex task that may require significant effort and result in an excessive response delay[4]. Previous works have investigated ways to facilitate the identification of anomaly causes, e.g. via Root Cause Analysis (RCA) using a distance metric[4] or analysing shifts in OPM parameters correlation[5]. However, they either work with supervised learning (requiring prior knowledge of the anomalies to be detected) or with a few key OPM parameters.

In this work, we propose a framework to assist RCA of anomalies. The RCA module uses the information provided by an anomaly detection module based on unsupervised ML to compute the changes in the OPM parameters incurred by the anomaly. The framework is validated on a physical layer security use case, using experimental data obtained by inserting harmful jamming signals and fiber squeezing for external polarization modulation[3]. While these attacks can cause severe service disruption, no exact theoretical models for their effects are known to date, which hinders their detection and counteraction. The results of applying our framework indicate that typically disregarded OPM parameters can be insightful for RCA of physical-layer attacks.

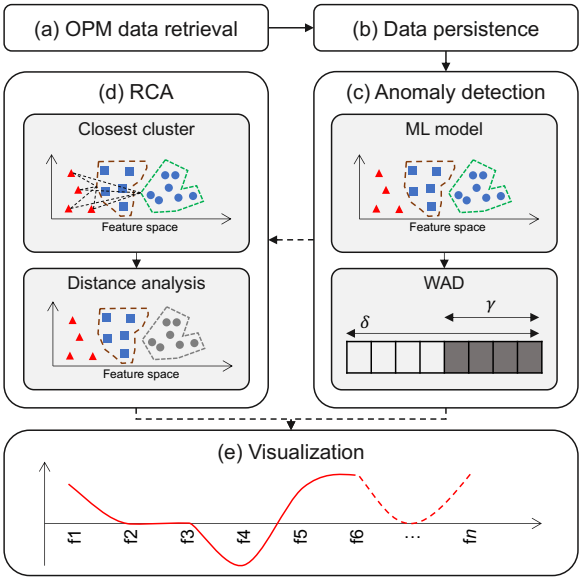

Fig. 1: Framework for Root Cause Analysis (RCA) composed of Optical Performance Monitoring (OPM), data persistence, anomaly detection using Window-based Anomaly Detection (WAD), and visualization modules.

The Root Cause Analysis Framework

RCA requires a framework that provides insightful information about the anomalies potentially occurring in the network and can accurately detect emerging anomalies. Such analysis is especially needed when using semi-supervised and unsupervised learning ML models that flag OPM samples as anomalies without providing any further

information on their nature. At this point, network management staff usually has the difficult task of analyzing the OPM parameters and identifying the cause of the anomaly.

To aid automation of this complex task, we propose an RCA framework comprising 5 modules, as shown in Fig. 1. The first two modules ((a) and (b)) retrieve and store the OPM data from the optical devices. Module (c) performs anomaly detection by pre-processing the data, executing the ML anomaly detection model (based on semi-supervised or unsupervised learning), and executing Window-based Anomaly Detection (WAD). WAD is responsible for reducing the negative impact of sparse false positives or false negatives common in anomaly detection models, especially in cases where false negatives have a detrimental effect, which is the case in security applications[3].

When an anomaly is detected, the framework invokes the RCA module (d), responsible for extracting information that can be insightful for the network management staff. For instance, in Fig. 1 we illustrate a clusterization technique used as the ML model for anomaly detection. The identified clusters can provide insight into reasons for considering certain samples as anomalous. First, the RCA module identifies clusters closest to the anomalous samples. Then, a feature-wise analysis can be performed between the normal and anomalous samples. Depending on the use case, different metrics can be used for the feature-wise analysis, such as distance[4] and correlation[5]. The family of diagnostic tools supporting RCA can be enriched with the Anomaly Vector (AV) whose elements contain the average difference between the OPM features in the baseline and anomaly condition. Our framework can be used with any anomaly detection technique. If a used technique does not perform clusterization, feature-wise distance calculation can be done between all the normal and all the anomalous samples. Finally, the results of anomaly detection and the insights gathered by the RCA module can be graphically represented by the visualization module (e).

DBSCAN-based Root Cause Analysis (RCA)

This section details the implementation of the framework adopted in this work. Our anomaly detection module uses a combination of Density-Based Spatial Clustering of Applications with Noise (DBSCAN) and WAD. DBSCAN is an unsupervised learning algorithm that clusters the samples by analyzing their pair-wise distances[6].

Algorithm 1 DBSCAN-based RCA

Data: Set of features F, pre-processed dataset X, DBSCAN parameters M and ϵ, WAD parameters δ and τ

Result: Anomaly detection flag $\{true, false\}$, anomaly vector AV (optional)

1 $Y \leftarrow$ DBSCAN(M, ϵ).fit_predict(X)
2 $w \leftarrow$ WAD(δ, τ, Y)
3 **if** w **then**
4 $\quad N \leftarrow \bigcup_{i=0}^{S}\{X_i\} : Y_i \geq 0$
5 $\quad A \leftarrow \bigcup_{i=0}^{S}\{X_i\} : Y_i = -1$
6 $\quad P \leftarrow$ cluster in N closest to A
7 $\quad AV_i \leftarrow \frac{\sum_{j=0}^{|A|} A_{i,j}}{|A|} - \frac{\sum_{k=0}^{|P|} P_{i,k}}{|P|}, i{=}0..F$
8 $\quad$ **return** $true, AV$
9 **else return** $false$;

The algorithm has two parameters: ϵ, defining the neighborhood radius around each sample, and M, defining the minimum number of neighbors a sample should have to be considered normal; samples that do not have M neighbors are considered anomalies. DBSCAN processes a dataset X with S latest samples collected from the optical devices, each containing F features (e.g., OPM parameters). It results in the set Y ($|Y|{=}S$) of cluster indices for each sample, whose values are non-negative for normal and -1 for anomalous samples. WAD works by considering a window with δ samples out of Y, returning $true$ if τ samples within the window have been flagged as anomalies, and $false$ otherwise[3].

Alg. 1 presents the proposed RCA method. The algorithm always returns a flag, informing whether an anomaly is detected ($true$) or not ($false$). If an anomaly is detected, the algorithm also returns the Anomaly Vector (AV), which contains the RCA results. First, DBSCAN is executed over the dataset X (line 1). The resulting set Y is used by WAD (line 2) to determine whether an alarm should be triggered or not. In the positive case (lines 3–8), the RCA module is executed (lines 4–7). It first extracts the normal samples N (line 4) and the anomalous samples A from X (line 5). Then, it identifies the cluster from N that is the closest to the anomaly samples in A (line 6). The AV is formed by computing the difference between the average value of each feature in the anomaly samples and the average value in the closest cluster P (line 7).

Use Case and Numerical Results

The proposed framework is validated on a physical layer security use case, where the anoma-

lies are characterized by physical layer attacks. An experimental optical network testbed with coherent transceivers is used to obtain OPM parameter samples. The OPM parameters of interest are: pre-FEC Bit Error Rate (BER-FEC), post-FEC Bit Error Rate (BER-PF), Loss of Signal (LOS), Optical Power Received (OPR), Chromatic Dispersion (CD), Differential Group Delay (DGD), Optical Signal-to-Noise Ratio (OSNR), Polarization Dependent Loss (PDL) and Q-factor. Two optical channels are monitored characterizing a baseline (no attack) scenario. Then, three attack strategies are launched in the network, namely, In-Band (IB) and Out-of-Band (OOB) jamming, and Polarization Modulation (PM). Each attack strategy is launched with two intensities (light and strong). We refer to our previous work for more information[3].

DBSCAN is configured with $M{=}15$ and $\epsilon{=}1.5$, which results in a 0.034 false positive rate and 0.189 false negative rate. WAD is configured with $\delta{=}20$ and $\tau{=}9$, obtaining a false detection rate of 7.24e-9 and 0.9999 true detection rate, which illustrates the benefits of using WAD over considering the results from DBSCAN directly.

Fig. 2 shows the AVs of the proposed RCA. The results are averaged over 50 random continuous sample windows containing 10:1.5 normal to attack sample ratio. Negative values mean that the attack incurs a lower value of a feature than in the baseline, while positive values mean that the attack incurs a higher value. Comparing the three attack strategies, IB jamming causes the most significant change in the feature values (up to -3), PM causes a medium change (up to -1.5), while OOB jamming changes the values least significantly (up to -0.75). As expected, light intensity attacks impact the values less significantly than their stronger counterparts.

For IB jamming (Fig. 2a), OSNR and Q-factor are the most affected features, followed by DGD and OPR. OOB jamming (Fig. 2b) shows a similar profile as the IB, but with a milder effect on the feature values. PM attack (Fig. 2c), on the other hand, exhibits a very different profile. Besides OSNR and Q-factor, this attack also significantly changes the CD and DGD of the received signal. Moreover, an increase in the BER-PF is observed. It is interesting to note that, looking at the physical nature of the attack, CD and DGD features should not be affected by the OOB jamming attack where just OSNR, Q-factor and BER-PF are expected to vary. AV visualization equips

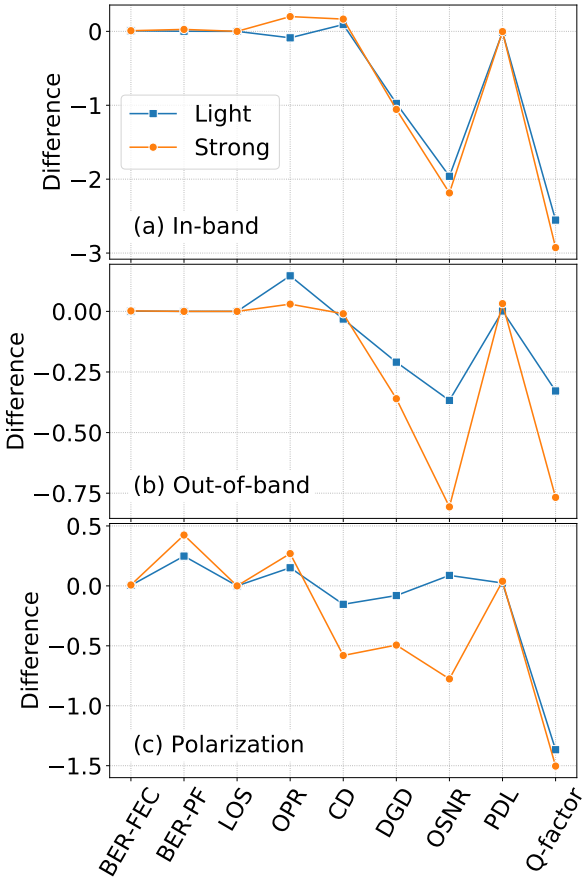

Fig. 2: The anomaly vectors, i.e. the average difference between the detected anomaly samples and the closest cluster for each feature.

the operators with deep insight into the anomaly structure and eases the physical interpretation of the anomaly thus complementing the ML-assisted RCA tool. AV visualization goes beyond the typical historical data plot that is provided by Network Management Systems today, where simple time series of OPM parameters are presented to the operators. This is especially significant when a new, previously undetected anomaly is analyzed.

Conclusions

This study combines RCA with a visualization tool to enable anomaly identification in optical network scenarios. Verified on a physical layer security use case, it shows that some physical layer parameters that are not at the center of usual analyses (e.g., CD) can be a good source of insight. Moreover, some of these parameters are not considered in analytical models because they should not be affected by attacks, but showed significant changes in the real system. The framework and the findings have important implications on the subsequent development of tailored anomaly remediation strategies.

Acknowledgements: VR (2019-05008).

References

[1] D. Rafique and L. Velasco, “Machine learning for network automation: Overview, architecture, and applications”, *IEEE J. Opt. Commun. Netw.*, vol. 10, no. 10, pp. D126–D143, 2018, DOI: 10.1364/JOCN.10.00D126.

[2] X. Chen *et al.*, “Self-taught anomaly detection with hybrid unsupervised/supervised machine learning in optical networks”, *J. Lightwave Techn.*, vol. 37, no. 7, pp. 1742–1749, Apr. 2019, DOI: 10.1109/JLT.2019.2902487.

[3] M. Furdek *et al.*, “Machine learning for optical network security monitoring: A practical perspective”, *J. Lightwave Techn.*, vol. 38, no. 11, pp. 2860–2871, 2020, DOI: 10.1109/JLT.2020.2987032.

[4] D. Rafique *et al.*, “Analytics-driven fault discovery and diagnosis for cognitive root cause analysis”, in *Optical Fiber Communication Conference*, DOI: 10.1364/OFC.2018.W4F.6, 2018, W4F.6.

[5] M. Furdek *et al.*, “Experiment-based detection of service disruption attacks in optical networks using data analytics and unsupervised learning”, in *Metro and Data Center Optical Networks and Short-Reach Links II*, DOI: 10.1117/12.2509613, SPIE, 2019, pp. 73–82.

[6] E. Schubert *et al.*, “DBSCAN revisited, revisited: Why and how you should (still) use DBSCAN”, *ACM Trans. Database Syst.*, vol. 42, no. 3, Jul. 2017, DOI: 10.1145/3068335.