Demonstration of Machine-Learning-Assisted Security Monitoring in Optical Networks

DEMONSTRATION OF MACHINE-LEARNING-ASSISTED SECURITY MONITORING IN OPTICAL NETWORKS

Marija Furdek[1] , Carlos Natalino[1], Fabian Lipp[2], David Hock[2], Nieke Aerts[2], Marco Schiano[3], Andrea Di Giglio[3], Lena Wosinska[1]*

[1]*Department of EE, Chalmers University of Technology, Chalmersplatsen 4, 412 96 Gothenburg, Sweden*
[2]*Infosim GmbH & Co. KG, Landsteinerstraße 4, 97074 Würzburg, Germany.*
[3]*Telecom Italia, Turin, Italy.*
**E-mail: furdek@chalmers.se*

Keywords: PHYSICAL-LAYER SECURITY, ATTACK DETECTION AND LOCALISATION, OPTICAL NETWORK TELEMETRY, SUPERVISED LEARNING, UNSUPERVISED LEARNING

Abstract

We report on the first demonstration of machine-learning-assisted detection, identification and localisation of optical-layer attacks integrated into network management system and verified on real-life experimental attack traces from a network operator testbed.

1 Introduction and Relevance

The introduction of machine learning (ML) algorithms represents a key enabler for autonomous monitoring, management and control of optical networks [1], allowing to fully exploit the telemetry capabilities of coherent devices [2]. The use of ML can reduce the monitoring overhead and enhance the accuracy of estimation methods in optical networks [3]. Various approaches [1, 3, 4] and demonstrators [5–9] showcased the advantages and the integration of ML algorithms in optical network monitoring, control and management. However, apart from few preliminary studies [10, 11], issues related to physical-layer security remain largely unaddressed.

Security assurance in optical networks, as critical communication infrastructure, is becoming increasingly relevant in the face of growing threats aimed at service disruption [12]. Encompassing and cognitive security management requires development and integration of a set of tailored techniques. Firstly, continuous optical performance monitoring (OPM) is needed. State-of-the-art coherent transceivers enable the collection of a broad range of optical parameters without the need for costly specialized equipment. The OPM data collected for different channels in the network then needs to be analyzed in order to perform attack detection and identification (ADI), i.e., to interpret whether individual channels are affected by an attack, and what type of attack it is. Finally, the location of the breach source should be identified as a prerequisite for its neutralisation and service recovery.

This demo showcases the integration of an ML-based ADI and an attack localisation module into an automated network management system capable of continuously monitoring the network, and experimentally demonstrates the system effectiveness for different real-life physical-layer attack techniques in an operator's multi-vendor testbed. The demo enables the audience to interact with the various modules in the holistic monitoring loop, selecting among several possible configurations. By integrating the modules via standard communication interfaces, the capabilities, limitations and challenges of implementing such functionalities are demonstrated.

2 Demo Architecture and Presentation

The monitoring platform showcased in the demo is composed of three main modules, as illustrated in Fig. 1. The monitoring and visualisation server (MVS) plays a central role, being responsible for receiving the OPM reports from the devices located at the network nodes (i.e., coherent receivers), storing the OPM data in a database, communicating with the machine-learning-based attack detection and identification module (ML-ADIM) and the attack localisation module (ALM), triggering alarms, and visualizing the network status.

Fig. 2 describes the communication between the modules for one monitoring cycle, i.e., from the querying of OPM data to

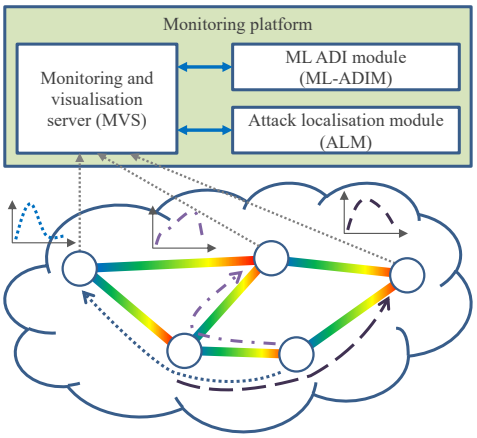

Fig. 1: Architecture of the monitoring platform.

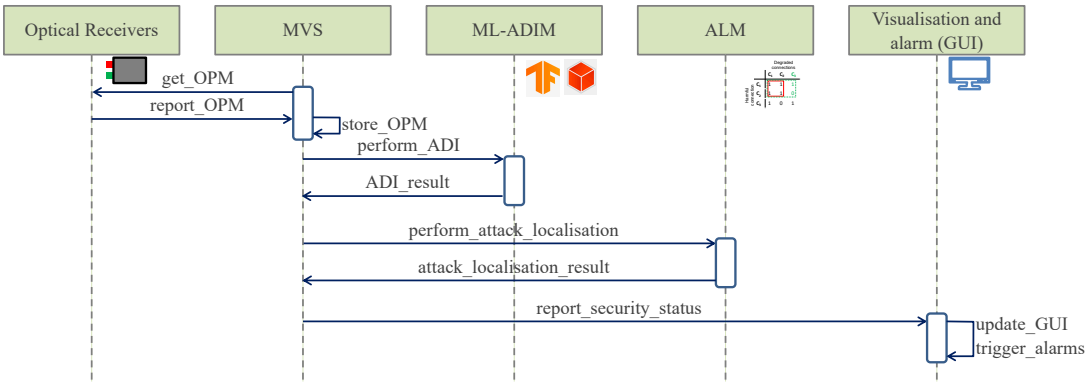

Fig. 2: Communication diagram between the modules composing the demo.

the security status interpretation and visualisation. The MVS periodically polls the optical receivers using simple network management protocol (SNMP), which report back the OPM data for storing in a database and further processing. When OPM data is received from the transponders, the MVS calls the ML-ADIM, which performs detection and identification of attacks affecting individual optical channels. To prioritize between accuracy and complexity performance, the ML-ADIM can be executed in three modes: *(i)* supervised learning (SL) only; *(ii)* unsupervised learning (UL) only; and *(iii)* supervised and unsupervised learning.

SL algorithms, on the one hand, provide efficient and precise ADI [13], but are not suitable for previously unseen attacks [14]. A pre-trained model is able to identify the type and intensity of an attack by evaluating only the current OPM data obtained from the optical devices, without requiring any historical data, which contributes to a low complexity of the algorithm. The identification of the attack type can assist the ALM, which can leverage the identified attack to enhance localisation. However, when the network experiences a new type of attack, this technique is not useful since the model was not trained for that type, possibly leading to harmful false results. UL models, on the other hand, are able to detect attacks not seen previously, but cannot identify the particular type of attack [14]. By leveraging on anomaly detection algorithms, it is possible to identify attacks based on the principle that their presence, in the beginning, is extremely rare, i.e., most OPM samples are attack-free. As a result, UL models enable early detection of new types of attacks without prior knowledge of the attack signature. However, many UL algorithms traverse the entire dataset (or a big portion of it), which increases complexity [15, 16].

The MVS combines the diagnostic information of individual channels received from the ML-ADIM with general information such as the network topology graph and resource allocation, and forwards it to the ALM for network-wise attack localisation to identify the source of the attack, i.e., the breached link and/or the harmful connection. The ALM deduces the possible source of the attack depending on the information provided, as well as the specific characteristics of the attack, if available. Finally, the monitoring platform's graphical user interface (GUI) provides a visual representation of the network security status, representing the channels affected by an attack as well as its likely source. Both ML-ADIM and ALM expose their functionalities through representational state transfer (REST), which enables their use by any network monitoring, control and management software. They are developed in the form of containers, facilitating their deployment in the form of virtual network functions (VNFs), which can be used for scalability. The modules are implemented as stateless, i.e., do not store the data, but rely on the MVS to supply it.

During the demo live presentation, the attendees can interact with the platform by selecting the type and the location of an attack to be launched, the type of ML algorithm used by ML-ADIM (UL/SL), and the working mode of the ALM (link/connection localisation). The SL mode of the ML-ADIM uses a pre-trained artificial neural network (ANN) hosted by TensorFlow Serving. The UL mode uses the Scikit-Learn implementation of the density-based spatial clustering of applications with noise (DBSCAN). The experimental data from the coherent transceivers in the testbed for the described attack techniques of different intensities are collected beforehand and replayed in real time. For a selected attack scenario and the ML-ADIM+ALM modes, the configuration, the effects, and the security assessment result are visualised using the GUI based on StableNet [17].

3 Use case

The multi-vendor testbed used to collect the data for the demo is based upon a commercial optical transport network with 6 ROADM nodes and 11 links. The optical signals under test are two 200 Gbit/s polarisation multiplexed 16QAM signals generated by coherent transponders and reported through their monitoring interface. Three physical-layer attack techniques are implemented: *(i)* in-band jamming (IBJ), where the intrusion signal is a continuous wave (CW) low power signal with frequency within the bandwidth of the signal under test; *(ii)* out-of-band jamming (OBJ), where the intrusion signal is a CW signal with a frequency outside the bandwidth of the signal under test, and *(iii)* polarisation scrambling (PS), where a polarisation state modulator is activated to cause transmission errors when the induced polarisation variation is faster than the receiver's polarisation recovery algorithm.

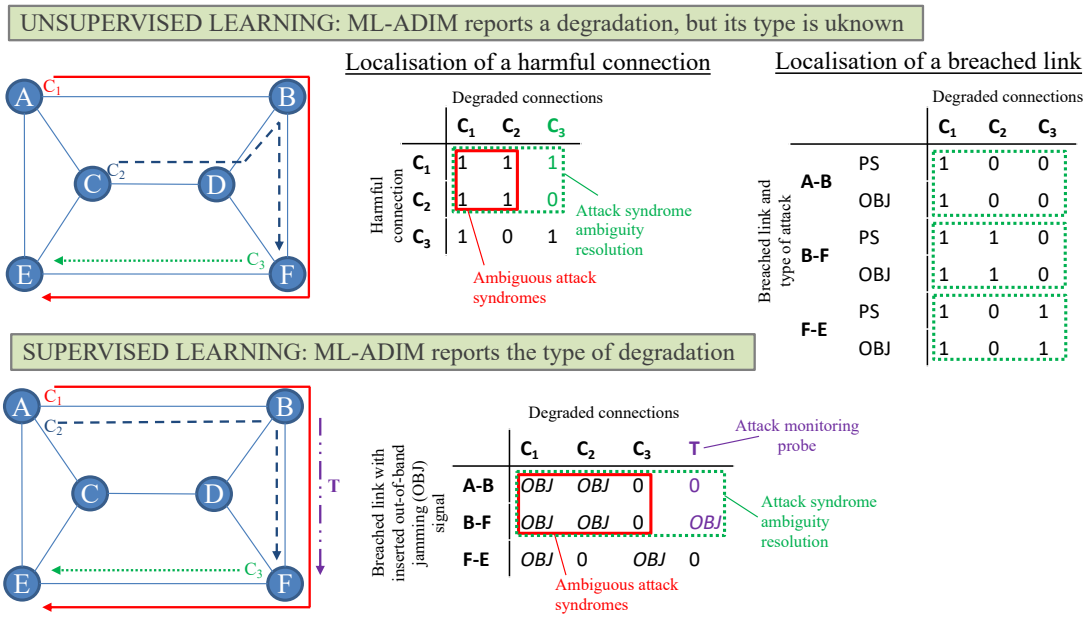

Fig. 3: Monitoring out-of-band jamming (OBJ) and polarisation scrambling (PS) attacks in unsupervised (top) and supervised learning mode (bottom).

Detection and localisation of attacks using the developed platform is illustrated with a simplified example in Fig. 3 for different connection routing, attack types (OBJ and PS) and working modes of the ML-ADIM (UL/SL) and ALM modules (link/connection localisation). The top part shows the case where UL is applied, i.e., the ML-ADIM module reports only the degradation of a subset of connections. As the cause of the degradation is unknown to the module, no attack type information is provided. The information from ML-ADIM is represented in ALM using binary words called attack syndromes [18], where the status of each degraded connection is denoted by 1, while 0 denotes regular, unaffected status at the receiver. Attack syndromes are pre-constructed for different sources of attacks and, if unique, can unambiguously identify the breach source depending on the subset of affected connections. The attack syndromes are constructed during initialisation according to the two ALM working modes, stored in the MVS, and used for lookup of the reported syndrome. The top left table of Fig. 3 shows localisation of the harmful connection. If we assume that connection C_1 carries the harmful signal (e.g., of excessive power, which could be caused by tampering with the power levels), it will degrade C_2 in their shared fibre link (B-F). If connection C_3 were not present in the network, the attack syndromes of C_1 and C_2 as potential attack sources would be equal (denoted with a red frame), and unambiguous attack localisation would not be possible. Thus, C_3 serves as the tie-breaker enabling correct attribution of the attack source to connection C_1 (denoted with a green frame). The top right table illustrates localisation of the breached link for representative links and the two attack techniques. The link-wise unique attack syndromes allow for the breached link to be deduced from the subset of degraded connections.

The bottom part of Fig. 3 shows the SL mode where ML-ADIM reports more detailed information about the attack method that caused degradation of each connection. This diagnostic tool provides a better insight into the performance of an individual channel, and is able to identify a harmful connection based on pertinent OPM parameters (e.g., excessive power), which makes its localisation a less challenging task than in the UL mode. For this reason, we only focus on the localisation of the breached link. The corresponding table with attack syndromes contains the type of the attack identified by ML-ADIM for each connection. In this example, we assume that connections C_1, C_2 and C_3 are established in the network, and an OBJ signal is inserted on link A-B, B-F, or F-E. While the syndrome of link F-E as the insertion point is unique, the syndromes of A-B and B-F are identical because the same subset of connections is affected. Therefore, an attack monitoring probe T is needed to resolve the ambiguity and enable precise breach localisation.

4 Final Remarks

The demo showcases a holistic security monitoring framework capable of detecting physical-layer attacks, identifying the type of known attacks, and localizing their source at the network level. By demonstrating the effectiveness of the system's implementation based on standard, open interfaces in an operator's testbed, we outlined the main challenges and capabilities of integrated data analytics for physical-layer network security diagnostics.

5 Acknowledgements

We gratefully acknowledge Infinera for providing the Groove G30 transponder, as well as Roberto Morro for providing the data acquisition application. This article is based upon work from Celtic-Next projects SENDATE-EXTEND and SENDATE-PLANETS, and COST Action 15127 RECODIS.

6 References

[1] Rafique, D., Szyrkowiec, T., Grießer, H., et al.: 'Cognitive assurance architecture for optical network fault management', Journal of Lightwave Technology, 2018, **36**, (7), pp. 1443–1450. DOI: 10.1109/JLT.2017.2781540.

[2] Paolucci, F., Sgambelluri, A., Cugini, F., et al.: 'Network telemetry streaming services in SDN-based disaggregated optical networks', Journal of Lightwave Technology, 2018, **36**, (15), pp. 3142–3149. DOI: 10.1109/JLT.2018.2795345.

[3] Meng, F., Mavromatis, A., Bi, Y., et al.: 'Self-learning monitoring on-demand strategy for optical networks',Journal of Optical Communications and Networking, 2019, **11**, (2), pp. A144–A154. DOI: 10.1364/JOCN.11.00A144.

[4] Vela, A.P., Shariati, B., Ruiz, M., et al.: 'Soft failure localization during commissioning testing and lightpath operation', Journal of Optical Communications and Networking, 2018, **10**, (1), pp. A27–A36. DOI: 10.1364/JOCN.10.000A27.

[5] Vela, A.P., Gifre, L., De Dios, O.G., et al.: 'CASTOR: A monitoring and data analytics framework to help operators understand what is going on in their networks'. European Conference on Optical Communication (ECOC), Rome, Italy, September 2018, pp. TuDS.1. dOI: 10.1109/ECOC.2018.8535500.

[6] Troia, S., Rodriguez, A., Martín, I., et al.: 'Machine-learning-assisted routing in SDN-based optical networks'. European Conference on Optical Communication (ECOC), Rome, Italy, September 2018, pp. TuDS.6. DOI: 10.1109/ECOC.2018.8535437.

[7] Morro, R., Lucrezia, F., Gomes, P., et al.: 'Automated end to end carrier ethernet provisioning over a disaggregated WDM metro network with a hierarchical SDN control and monitoring platform'. European Conference on Optical Communication (ECOC), Rome, Italy, September 2018, pp. TuDS.4. DOI: 10.1109/ECOC.2018.8535422.

[8] Sadasivarao, A., Syed, S., Lu, B., et al.: 'Demonstration of advanced open WDM operations and analytics, based on an application-extensible, declarative, data model abstracted instrumentation platform'. Optical Fiber Communication Conference (OFC), San Diego, CA, USA, March 2019. pp. M3Z.1. DOI: 10.1364/OFC.2019.M3Z.1.

[9] Bouda, M., Oda, S., Akiyama, Y., et al.: 'Demonstration of continuous improvement in open optical network design by QoT prediction using machine learning'. Optical Fiber Communication Conference (OFC), San Diego, CA, USA, March 2019. pp. M3Z.2. DOI: 10.1364/OFC.2019.M3Z.2.

[10] Zhu, Z., Kong, B., Yin, J., et al.: 'Build to tenants' requirements: On-demand application-driven vSD-EON slicing', Journal of Optical Communications and Networking, 2018, **10**, (2), pp. A206–A215. DOI: 10.1364/JOCN.10.00A206.

[11] Li, Y., Hua, N., Yu, Y., et al.: 'Light source and trail recognition via optical spectrum feature analysis for optical network security', Communications Letters, 2018, **22**, (5), pp. 982–985. DOI: 10.1109/LCOMM.2018.2801869.

[12] Skorin.Kapov, N., Furdek, M., Zsigmond, S., et al.: 'Physical-layer security in evolving optical networks', Communications Magazine, 2016, **54**, (8), pp. 110–117. DOI: 10.1109/MCOM.2016.7537185.

[13] Natalino, C., Schiano, M., Giglio, A.D., et al.: 'Field demonstration of machine-learning-aided detection and identification of jamming attacks in optical networks'. European Conference on Optical Communication (ECOC), Rome, Italy, September 2018, pp. We2.58. DOI: 10.1109/ECOC.2018.8535155.

[14] Furdek, M., Natalino, C., Schiano, M., et al.: 'Experiment-based detection of service disruption attacks in optical networks using data analytics and unsupervised learning'. Metro and Data Center Optical Networks and Short-Reach Links II, San Francisco, CA, USA, February 2019. DOI: 10.1117/12.2509613.

[15] Chen, X., Li, B., Proietti, R., et al.: 'Self-taught anomaly detection with hybrid unsupervised/supervised machine learning in optical networks', Journal of Lightwave Technology, 2019, **37**, (7), pp. 1742–1749. DOI: 10.1109/JLT.2019.2902487.

[16] Varughese, S., Lippiatt, D., Richter, T., et al.: 'Identification of soft failures in optical links using low complexity anomaly detection'. Optical Fiber Communication Conference (OFC), San Diego, CA, USA, March 2019. pp. W2A.46. DOI: 10.1364/OFC.2019.W2A.46.

[17] 'StableNet – unified network and services management', https://www.infosim.net/stablenet/, accessed 1 May 2019.

[18] Furdek, M., Chan, V.W.S., Natalino, C., et al.: 'Network-wide localization of optical-layer attacks'. Conf. on Optical Network Design and Modelling (ONDM), Athens, Greece, May 2019.