Optical Layer and Quantum Digital Twins for Enhancing Secure Autonomous Network Operation

Marc Ruiz* **and Luis Velasco**
Optical Communications Group (GCO), Universitat Politècnica de Catalunya (UPC), Barcelona, Spain
e-mail: marc.ruiz-ramirez@upc.edu

ABSTRACT
This paper reviews the recent research advances and open issues related with the use of digital twins for secure autonomous network operation. In particular, we focus on two different solutions that co-exist as part of the overall intelligent network control and management system. On the one hand, the OCATA optical layer network digital twin is presented as a tool that can detect perturbation and variations in optical constellation samples in order to detect and identify attacks on machine learning pipelines compromising secure autonomous network operation. On the other hand, the DARIUS quantum digital twin is presented as a tool that allows improving the behavior of real quantum key distribution (QKD) systems by increasing their key exchange rate while keeping the outstanding ability to detect eavesdropping attacks.

Keywords: ***Digital Twins, Network Security, Optical Networks, Quantum Networks***

I. INTRODUCTION

Digital Twins (DT), in combination with software-defined networking (SDN) control, have demonstrated significant potential across various optical network operation use cases, such as estimating Quality of Transmission (QoT) for optical connection provisioning and managing failures [1]. These solutions provide end-to-end network control by means on fundamental pillars such as the collection of pervasive telemetry across the different network devices and the use of ML models that provide intelligence [2]. Thus, ML pipelines are necessary to keep the proper data ingestion and model processing in support of autonomous network operation. However, the deployment of ML pipelines and DT entities increase the attack surface, which poses additional challenges to solve to increase robustness and security of network operation against multiple threats.

In co-existence with classical optical networks, Quantum Key Distribution (QKD) systems are attracting great attention as they generate secure keys that can be used for different use cases [3], which largely increases overall network security. The main drawback of these kind of systems is the yet low secret key generation rate, which is even worsen in real networks subject to many sources of loss and environmental events [4]. Under this situation, the performance of QKD systems can be potentially improve by using dedicated DTs that can mitigate some of these negative effects, while keeping its ultra-high security functionalities.

In this paper, we present two different use cases related to DTs and security. On the one hand, we provide an overview of the OCATA optical time domain DT [5], that supports several optical transmission technologies including multiband transmission [6]. Then, we show how its ability to detect Man-In-The-Middle (MitM) attacks can be cancelled if the adversary also attacks the ML pipeline in charge of supervising and validating secure network operation. On the other hand, we present the high-level architecture and main functionalities of the DARIUS Quantum DT for polarization-encoded QKD systems [4], which allows understanding fiber stressing events and triggering proper configurations for maximizing performance of QKD systems.

II. THE OCATA OPTICAL LAYER NETWORK DIGITAL TWIN

Fig. 1 illustrates the reference architecture of the OCATA DT, which models optical signal propagation in the time domain. It uses Deep Neural Networks (DNNs) to simulate how in-phase (I) and quadrature (Q) constellations are affected by linear (LI) and nonlinear (NLI) noise across components like Reconfigurable Optical Add-Drop Multiplexers (ROADMs) and optical links. Since DTs rely on telemetry, IQ constellations are collected from transponders (TP).

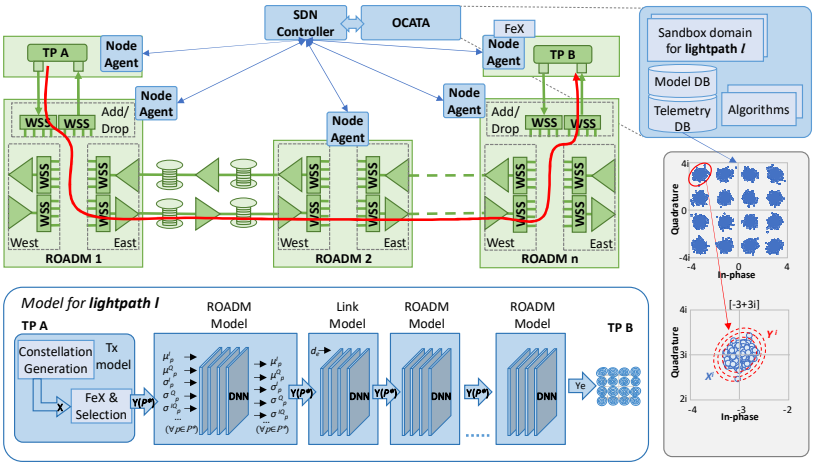

Fig. 1: OCATA reference architecture

979-8-3315-9777-1/25/$31.00 ©2025 IEEE

In OCATA, lightpaths are modeled by chaining DNN-based components. Signals using m-QAM modulation (optical constellations) are processed, where each constellation point (CP) i is represented by a feature vector Y_i capturing the mean and covariance matrices that characterize the CP as bi-variate Gaussian distribution. For scalability, only a subset of CPs—two interior and two exterior—are modeled directly; others are derived from these. The transmitter (Tx) model generates initial constellation X and its feature set Y, using a pseudo-random bit sequence. A feature extraction (FeX) module applies Gaussian Mixture Models (GMM) to compute Y.

OCATA can estimate the expected constellation X_e and features Y_e, enabling comparison with actual received data X_r. A discrepancy function or model can then identify deviations, which may indicate anomalies and/or failures [1], or even attacks [7].

III. MACHINE LEARNING PIPELINE SECURITY

Assuming the previous reference network and OCATA architecture and functionalities, let us analyze the impact of a MitM attack [7]. In addition, let us assume that OCATA has an algorithm that allows computing the expected lightpath length L as a function of expected propagated features Y_e, as presented in [5]. Fig. 2a considers a lightpath under a MitM attack. Without loss of generality, we consider that Mallory has gained physical access to the intermediate site, so the attacker can receive the original optical signal from Alice, alter data and forward the new signal to Bob. In addition, Mallory could need remote access to other sites along the route of the lightpath, e.g., if physical access to an intermediate site in the route of the lightpath is not possible and some sort of configuration of switching in the optical nodes is needed. Finally, let us assume that the attack can be conducted with negligible optical transmission disruption, and therefore no loss-of-light alarm is raised.

OCATA can detect differences w.r.t. the expected lightpath length L, by inferring that the optical signal received by Bob has been generated by a transmitter located at $L' \cong L_2$. Therefore, a warning notification will be sent to the SDN controller, which eventually can trigger countermeasures. In view of that, Mallory needs to perform additional processing to bypass security systems based on OCATA. In particular, the HOCUS module presented in [7] can be used to attack the ML pipeline that collects and process optical constellation samples and compares with expected features. In particular, HOCUS reshapes optical symbols in collected samples in a way that OCATA would infer a value of the distance travelled by the received signal as desired by Mallory. Fig. 2b illustrates the bi-variate Gaussian distributions of constellation point (-3+3i) before and after the application of HOCUS. HOCUS receives source symbols (x_a) with length L_a and produces modified symbols (x_b) with a dispersion that resembles L_b.

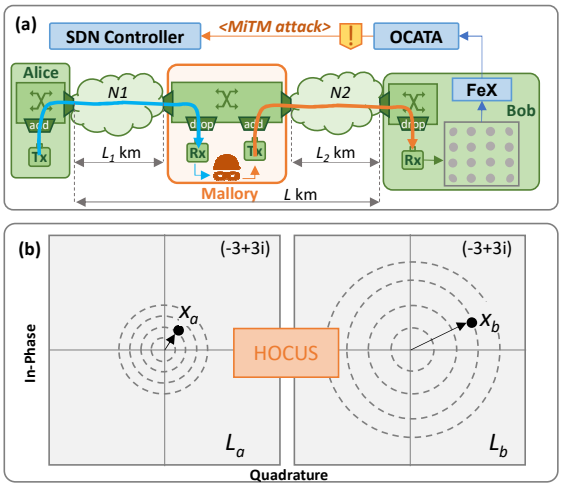

Fig. 2: MitM attack (a) and ML Pipeline Attack (b)

To evaluate HOCUS for ML pipeline attack, a simulation environment based on the data and models components used in [5] was setup. Fig. 3a shows the performance of HOCUS for a wide range of L_1 distances. The figure shows the normalized prediction of the lightpath estimation algorithm and the confidence interval (in green) when no MitM is performed. As can be observed, the attacker can effectively bypass any tampering attack detection by performing at the same time a ML pipeline attack with HOCUS, since lightpath length estimations fall in the confidence region area. Fig. 3b shows the detail of CP (-3+3i) before (original) and after the MitM attack, without and with HOCUS, for the case <L = 800 km, L_1 = 720 km>. The figure allows confirming that HOCUS introduces additional dispersion in the symbols that results in CPs similar to the original ones.

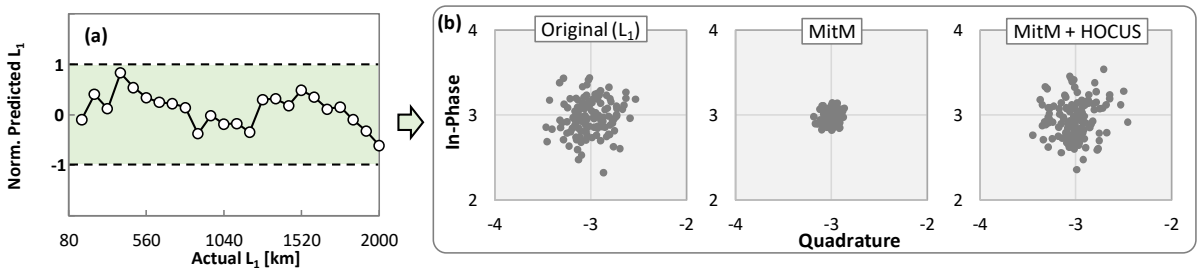

Fig. 3: HOCUS performance for ML pipeline attacks

In view of that, it is clear that additional algorithms for not only MitM attack detection but also ML pipeline attack detection are necessary to be adopted in OCATA to prevent advanced threats that can seriously compromise

secure autonomous operation of optical networks.

IV. THE DARIUS QUANTUM DIGITAL TWIN

Fig. 4 shows the architecture of DARIUS quantum DT (Fig. 4a), as well as details of the QKD system it models (Fig. 4b). DARIUS concatenates models of the optical components that form the QTx and QRx, as well as the fiber connecting them, to create a digital replica of the quantum channel. DARIUS includes methods to discern eavesdropping from fiber stressing events. Moreover, DARIUS help a DNN-powered compensation method running in the receiver to take counter actions against events on the fiber, in order to increase effective KER.

The eavesdropping detection takes advantage of Monitoring (Mo) intervals to monitor discrepancies in SOP, quantum Bit Error Rate (qBER), and Key Exchange Ratio (KER) between Mo and key exchange (Ke) intervals (Fig. 4c). Specifically, Mo intervals are assumed to track SOP fluctuations in case of fiber stressing events; keys exchange is interrupted and the QTx sends polarized photons during T_O period, which the QRx can measure and tune its electronic polarization controller (EPCs) during T_R. Mo intervals reduce KER and therefore, true Stokes measurements cannot be performed very frequently, which makes QKD systems especially vulnerable against episodes of large SOP variation which leads to large qBER and eavesdropping become indistinguishable.

DARIUS can distinguish between eavesdropping and excessive qBER, which will allow to continue with the key exchange in case of the latter. The SOP evolution is traceable when events caused by human operator works or environmental conditions affect the optical fiber. In contrast, eavesdropping results into unrecognizable SOP changes. SOP trajectory measurements during Mo and Ke intervals help DARIUS to detect eavesdropping [4].

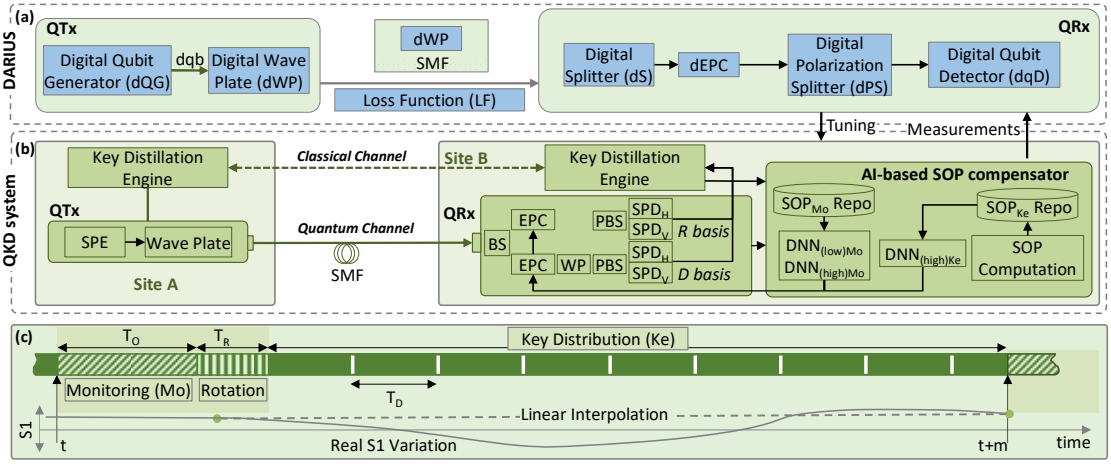

Fig. 4: DARIUS Architecture and SOP interpolation method

V. IMPROVING PERFORMANCE OF QKD SYSTEMS

Environmental events introduce fluctuations on the SOP of transmitted photons with differential velocity. In this case, the DNN model under operation in the SOP compensator ($DNN_{Low(Mo)}$) is not able to foresee the SOP of incoming photons. Therefore, another DNN model trained for higher velocity events ($DNN_{High(Mo)}$) is needed. DARIUS can detect the increased SOP velocity and change the model under operation, which would increase KER by SOP distortion compensation in different environmental conditions. In addition to DNN models, a rotation plan is continuously computed to rotate EPC in steps in order to maximize performance with the minimum number of rotations in each base.

The architecture of DARIUS and the QKD system presented in Fig. 4 have been evaluated on a simulation environment developed in Python (see details in [4]). We assume photon generation and detection rate of 100 Mb/s. In addition, we assume regular BB84 protocol with sifted KER of 45% and privacy amplification rate of 10%. As a result, a nominal KER of 4.5 Mb/s (100*45%*10% Mb/s) can be achieved with these specifications, in the absence of SOP perturbations and eavesdropping. In this work, we assume the most sophisticated scenario for an eavesdropper, i.e., he/she has detected Mo intervals and he/she can insert photons with the right polarization during Mo periods. In this case, discrepancies in qBER during Mo intervals ($qBER_{Mo}$) and during KE intervals ($qBER_{Ke}$) will reveal eavesdropping, i.e., if the difference exceeds a given threshold *thr*.

Fig. 5 evaluates eavesdropping detection in the absence (Fig. 5a) and presence (Fig. 5b) of high velocity events. For the sake of generality, we computed *eavesdropping rate* as ratio of the transmitted photons that are actually tampered by the eavesdropper. In Fig. 5a we can see the difference between expected and real (measured) qBER. Assuming *thr* = 2%, the analysis clearly reveals eavesdropping even with only 10% eavesdropping rate. Note also that KER drops when the eavesdropping rate increases, which could lead to false diagnosis if SOP measurements in Ke intervals are not analyzed. In Fig. 5b, qBER estimation during key distillation can show that $qBER_{Mo}$ and $qBER_{Ke}$ (expected and real *qBER*) are clearly different. Therefore, considering same *thr*=2%, as in the case where

no environmental events, eavesdropping rates as low as 10% can be detected.

A complete example is eventually showcased in Fig. 6 to illustrate how DARIUS can help the QKD system to address the high qBER in case of higher velocity events. Initially, the $DNN_{(low)Mo}$ model is in operation in the AI-based SOP compensator, which uses linear interpolation to compensate for low speed events, and qBER is well under the typical eavesdropping threshold (10%). Then, a fiber stressing event with increasing velocity reaches a critical speed (3rad/s), so DARIUS stops compensation until learning the new conditions. As a result, qBER exceeds the threshold and keys are discarded. DARIUS checks for eavesdropping and both scenarios are checked and confirm no attack. Then, DARIUS checks the velocity before and after the increased qBER and determines that velocity has increased from 2rad/s to 5 rad/s. In consequence, the AI model in operation is changed to DNN (high) Mo, and threshold -based interpolation using DNN (high) Ke model is applied. Now, the AI-based SOP compensation method provide again good predictions and can perfectly compensate for the higher velocity events.

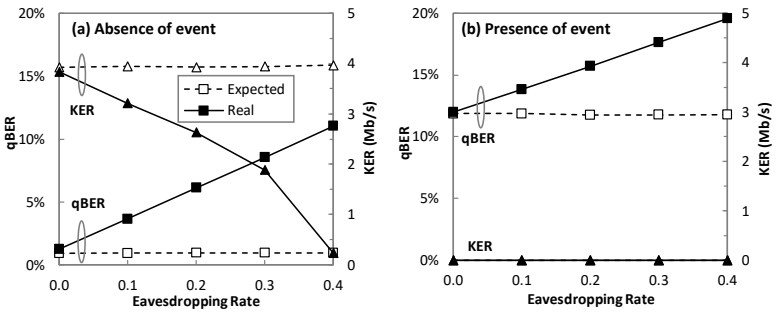

Fig. 5: Eve detection in absence (a) and presence (b) of fiber stressing event

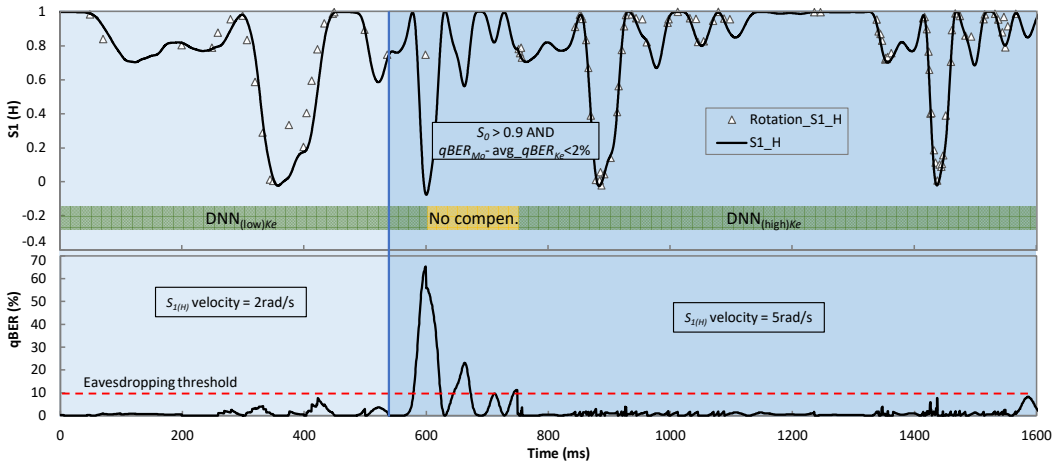

Fig. 6: Illustrative example of DARIUS operation

VI. CONCLUSIONS

DTs are key elements in support of secure autonomous operation of both optical and quantum networks. Although large achievements towards scalable and accurate intelligent operation have been reached so far, DT security requires additional methods to prevent and mitigate ML pipeline attacks, as well as to better discriminate true attacks (eavesdropping) from false attacks (environmental events) for proper network operation.

ACKNOWLEDGEMENTS

This work has received funding from the European Union's Horizon Europe research and innovation program under grant agreement No. 101092766 (ALLEGRO Project) and from the ICREA Institution

REFERENCES

[1] M. Devigili *et al.*, "Applications of the OCATA Time Domain Digital Twin: from QoT Estimation to Failure Management," J. of Opt. Comm. and Netw., 2024.
[2] P. Gonzalez *et al.*, "Near-Real-Time 6G Service Operation Enabled by Distributed Intelligence and In-Band Telemetry," J. of Opt. Comm. and Netw., 2025.
[3] M. Ahmadian *et al.*, "Demonstration of Classical and Hybrid Solutions for Lightpath Security Using Quantum Keys," in Proc. ICTON, 2024.
[4] M. Ahmadian *et al.*, "DARIUS: A Digital Twin to Improve the Performance of Quantum Key Distribution," J. Light. Technol., 2024.
[5] M. Ruiz *et al.*, "Deep Learning -based Real-Time Analysis of Lightpath Optical Constellations," J. of Opt. Comm. and Netw., 2022.
[6] S. Ghasrizadeh *et al.*, "Using the OCATA Digital Twin to Improve QoT of Optical Connections in Multiband Optical Networks," in Proc. ONDM, 2024.
[7] M. Ruiz *et al.*, "Man-in-the-Middle Attacks Through Re-Shaping I-Q Optical Constellations," in Proc. OFC, 2023.