Disaggregated Confidentiality-Preserving Scheme for Fault Detection in Optical Networks

R. F. Sales,[1,*] A. Ribeiro,[1] M. F. Silva,[2] F. R. Lobato,[1] A. Sgambelluri,[3] L. Valcarenghi[3] and J. C. W. Costa[1]

[1] *Institute of Technology, Federal University of Pará, Pará, Brazil*
[2] *Los Alamos National Laboratory, Los Alamos, New Mexico 87544, USA*
[3] *Scuola Superiore Sant'Anna, Pisa, Italy*

[*] *rafael.sales@itec.ufpa.br*

Abstract: We propose a confidentiality-preserving approach based on distributed principal component analysis (PCA) and telemetry data scrambling to detect hard-failures in optical networks. Experiments in a real optical testbed show the suitability of the proposed disaggregated solution.

1. Introduction

An accurate and rapid fault detection is of paramount importance to maximize the availability of transparent optical networks (TON). Although an effective failure management in optical networks reduces service disruptions, the security of the data transmitted on these networks are often unaccounted for. In that regard, the ultimate goal is to detect network failures effectively while preserving the confidentiality of the data.

As the complexity of networks increase rapidly, machine learning (ML) techniques overcomes the conventional threshold-based methods for failure management in TONs. These techniques have become crucial as they enable cognitive and fully-automated networks, coping with large system parameters and applications [1]. However, most ML algorithms used for failure detection in optical networks requires a large volume of data for proper training. It leads to a vast collection of telemetry data, raising concerns about the data confidentiality.

Principal component analysis (PCA) is one of the widely known algorithms for anomaly detection, which turns it very suitable for optical failure detection. For instance, [2] uses PCA for soft failure detection in TON and introduces data scrambling to ensure data confidentiality in a homomorphic computation scheme (i.e., computation is executed directly on the encrypted data). This is possible due to PCA rotation invariant property, which was exploited by shuffling telemetry data to preserve the information contained in the data. In the same sense, [3] shows different dimensionality reduction methods combined with a privacy-preserving approach. Only PCA and singular value decomposition (SVD) maintain the exact performance with and without the data transformation.

Although efficient, all the aforementioned works are based on third-party centralized PCA, which has less security regarding data breach since all data is concentrated in one location. Hence, inspired by the federated learning scheme, we propose a disaggregated PCA scheme that splits the data through several computation nodes for partially solving the global PCA problem. Later, from the partial solutions, we compute a global one without direct access to the entire data at a single location. To improve data confidentiality, each portion of data shared among the several computation nodes are randomly scrambled following the approach proposed in [2,3] to prevent unauthorized access while transmitting the data to each node. This process ensures a distributed processing of a large volume of telemetry data, alleviating the computational loads in training while ensuring data confidentiality.

2. Disaggregated Confidentiality-Preserving Scheme

PCA is a dimensionality reduction method that replaces a set X original variables by a set Q latent variables. The set Q is called principal components (PCs) and are obtained by multiplying the original data matrix by the eigenvectors of its covariance matrix. The dimensionality of Q is defined by the number of eigenvectors used, whether each PC retains a part of the variance of the original data. The quality of the Q set approximation can be measured by the sum of the variances associated with the retained PCs [4]. The eigenvectors of the covariance matrix of X can be calculated using SVD. Applying SVD to X leads $X = UDV^\top$, from which the covariance matrix can be written as $X^\top X = VD^2V^\top$, where V are the eigenvectors and D^2 the eigenvalues associated with the eigenvectors. The original data matrix can be reconstructed by applying $\hat{X} = QV^\top + E$, where E is the error matrix. Training the PCA using only normal data results in a model whose reconstruction errors for data from faulty conditions leads to increasing errors, depending on the level of failure.

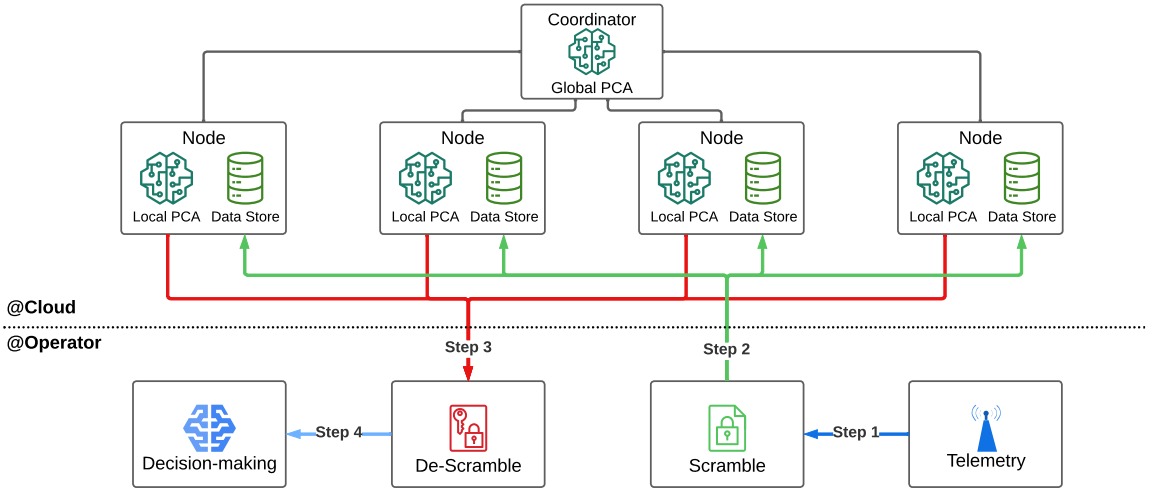

Fig. 1: Disaggregated Confidentiality-Preserving Scheme

All the concepts mentioned before are on the basis of centralized approach. However, in this work, based on [5], we use a disaggregated PCA approach for failure detection. The main difference between centralized and disaggregated PCAs is related to eigenvectors and eigenvalues acquisition. Distributed PCA algorithm calculates the PCs over partitioned data. The entire dataset is firstly partitioned into several subsets (local data). Then, local PCs are calculated for each susbset of data and communicated to a coordinator, which estimates the global covariance matrix. Correspondingly, this matrix is used to obtain the global PCs into which all local data is projected.

Furthermore, leveraging the advantages of distributed learning and homomorphic encryption, this work proposes a scheme that combines the disaggregated PCA with the data scrambling technique proposed in [2]. In most cases, the telemetry data are composed of several time-series with time-dependent structure. To make the data less intelligible, we can change the time order of the time series, leveraging the rotation invariant property of PCA to shuffle telemetry data to preserve information contained in original data. To further increase confidentiality, we use a distributed learning scheme that shares the data with n nodes, partitioning the time-series. This approach prevents unauthorized access to the data by a potentially malicious third-party actor as it does not have access to the entire data. Moreover, the scrambling approach changes the temporal and structural order of the data making it unintelligible in the event of a data breach attack (e.g. man in the middle [6]), adding an extra layer of protection to this disaggregated learning scheme. Since PCA is rotationally invariant, which means that it does not depend on the actual order of the data. We can train the PCA model over the scrambled data, in a homomorphic fashion, without any loss of performance and generality. As has been previously demonstrated in [2, 3].

The proposed approach is showed in Fig. 1. This scheme is composed of operator and cloud components. The operator side is composed of the homomorphic encryption and decision-making components. The cloud side are the machine learning components. In step 1, network telemetry is sent to the scramble block, where the data is scrambled and zero-centered. In step 2, the scrambled data is splitted and shared with n cloud local PCA nodes remotely located, where training is carried out without sharing the telemetry data between nodes. In step 3, reconstruction errors (failure indicators) are sent to the de-scramble block, which reorders the time series into its original form for comparison and failure analysis. Finally, in step 4, the operator receives the reconstruction errors in the original order from which the decision-making process can be derived.

3. Results

To evaluate the proposed approach, we used the telemetry dataset described in [7]. The considered testbed, includes two Ericcson SPO 1400 devices, one Wavelength Selective Switch (WSS) and four EDFA amplifiers. At the end of the WSS a 10 dB attenuator is installed to simulate failures. The optical link between the SPO1 and SPO2 consists of 3 spans of 80 km each. The data is collected with sampling frequency of 3.5 seconds and consists of 4 features, corresponding to the 4 EDFAs input powers. An interpolation technique was employed due to missing values in the original data set, generated at the end 13,948 samples. Among the samples, the first 80% of the data are used for training and the following 20% for testing. It is important to note that only the data corresponding to the test phase presented samples under failure conditions. The configuration of the scheme used for evaluation is

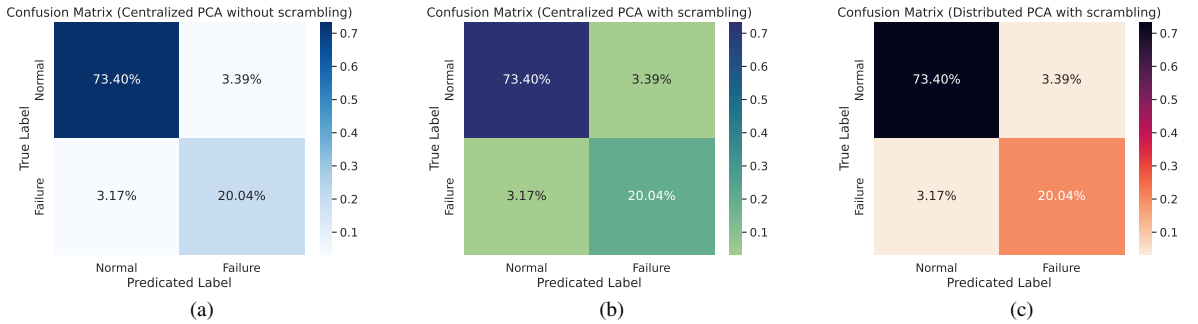

Fig. 2: Confusion Matrix: Centralized PCA (a), Centralized PCA with Scramble (b), Distributed Confidentiality-Preserving PCA (c)

shown in Fig. 1. The scheme has 4 local PCA nodes and 1 global node (coordinator node), where each local PCA trains with 20% of all data (approximately 2,789 samples). The rest of the data is used to the test phase, 5% for each node (approximately 697 samples).

The results are condensed in Fig. 2, where the Fig. 2a is the baseline, the most common application of PCA for anomaly detection. Fig. 2b shows the results of common PCA combined with the data scrambling technique proposed in [2]. As shown in the work [3], this approach does not change the result compared to the baseline, due to PCA rotation invariant property. Fig. 2c shows the results of disaggregated confidentiality-preserving scheme proposed by this work. Although the dataset is distributed among local PCA nodes, the configuration of 4 local nodes and 1 global node combined with data scramble, as seen in Fig. 1, did not change the results compared to the centralized PCA with and without the scramble, obtaining the same result as the other two approaches, as we can see in our results. Achieving an accuracy of 93.44%, 3.39% type I error and 3.17% type II error.

4. Conclusion

In this work, a disaggregated confidentiality-preserving scheme for detecting failures in a TON was proposed. The scheme takes advantage of the distribution of the dataset in combination with a scrambling technique to preserve the data confidentiality in the advent of unauthorized access. Additionaly, it presents the benefits of distributed systems, such as fault tolerance and parallelism. This scheme demonstrates promising results as it achieves the same accuracy as the conventional implementation of unsupervised ML models, delivering data confidentiality and scalability.

Acknowledgement

This work was partly supported by the project CLEVER (Project ID 101097560), that is supported by the Key Digital Technologies Joint Undertaking and its members (including top-up funding by Italian Ministry of University and Research — MUR)

References

1. F. Musumeci, “Machine Learning for Failure Management in Optical Networks,” 2021 Optical Fiber Communications Conference and Exhibition (OFC), San Francisco, CA, USA, 2021, pp. 1-28.
2. M. F. Silva et al., “Confidentiality-preserving Machine Learning Scheme to Detect Soft-failures in Optical Communication Networks,” 2022 European Conference on Optical Communication (ECOC), Basel, Switzerland, 2022.
3. M. F. Silva et al., “Confidentiality-preserving machine learning algorithms for soft-failure detection in optical communication networks,” in Journal of Optical Communications and Networking, vol. 15, no. 8, pp. C212-C222, August 2023, doi: 10.1364/JOCN.481690.
4. Jolliffe IT, Cadima J. 2016 Principal component analysis: a review and recent developments.Phil. Trans. R. Soc. A 374: 20150202.
5. Liang, Yingyu, Maria-Florina Balcan and Vandana Kanchanapally. “Distributed PCA and k-Means Clustering.” (2013).
6. Bhadouria, Aashi. (2022). Study of: Impact of Malicious Attacks and Data Breach on the Growth and Performance of the Company and Few of the World’s Biggest Data Breaches. 10.29322/IJSRP.X.2022.p091095.
7. InRete Lab, 2021, “Optical Failure Dataset,” Scuola Superiore Sant’Anna. [Online]. https://github.com/Network-And-Services/optical-failure-dataset