CISUC
CENTRO DE INFORMÁTICA E SISTEMAS DA UNIVERSIDADE DE COIMBRA

SMART 5GRID

Towards An Intrusion Detection System For Smart Grids

A Federated Approach

Author: Dylan Gonçalves Perdigão
Advisors: Pedro Henriques Abreu and Tiago J. Cruz
University of Coimbra, CISUC, DEI, Portugal

Introduction

This project aims at researching and developing an **Intrusion Detection System (IDS)** for critical infrastructure, such as Smart Grids, which impacts energy distribution and production. **Smart Grids (SGs)** [1] are modernised electricity networks that use advanced digital technology to improve the power grid system's efficiency, reliability, and sustainability. This can include the integration of renewable energy sources, the use of advanced sensors and control systems, and the ability to support two-way communication and power flow [2]. The SGs are composed of **substations** disseminated among the whole infrastructure controlling an **autonomous geographical area**. Inside the substation, their servers have a **hierarchical structure** for processing data from sensors (IoT devices) in the grid.

Proposed Approach

Due to the heterogenicity of the data, applying common Machine Learning (ML) algorithms to detect intrusions can lead to poor classification results. **Federated Learning (FL)** [3] can solve this issue since it enables a **distributed training** of the data among different nodes. The FL process is shown in Figure 1. First, a central server initialises the client's local models. Secondly, the clients update their local models with their data. Finally, the central server aggregates the client's local models.

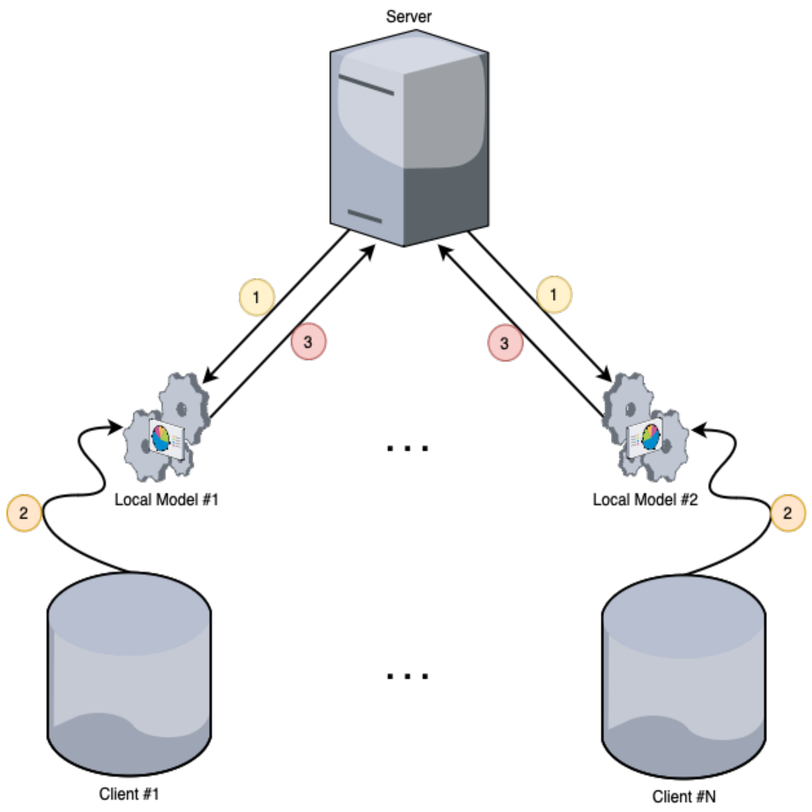

Figure 1: Federated Learning architecture

However, **anomaly detection** problems, such as detecting attacks in a system, are **affected by imbalanced data**. This means that the samples used to train ML models have a kind o class more represented than others (e. g. more examples of normal scenarios than attacks). This leads to poor performance of the models. A solution is to **classify the minority class points** (Fig. 2) into four categories: **Safe, Borderline, Rare, and Outliers** [4]. This typology is computed with the **5-Nearest Neighbors (5NN)** belonging to the same class to the point. This classification makes it possible to rebalance the data with more informative examples using **SMOTE** algorithm [5].

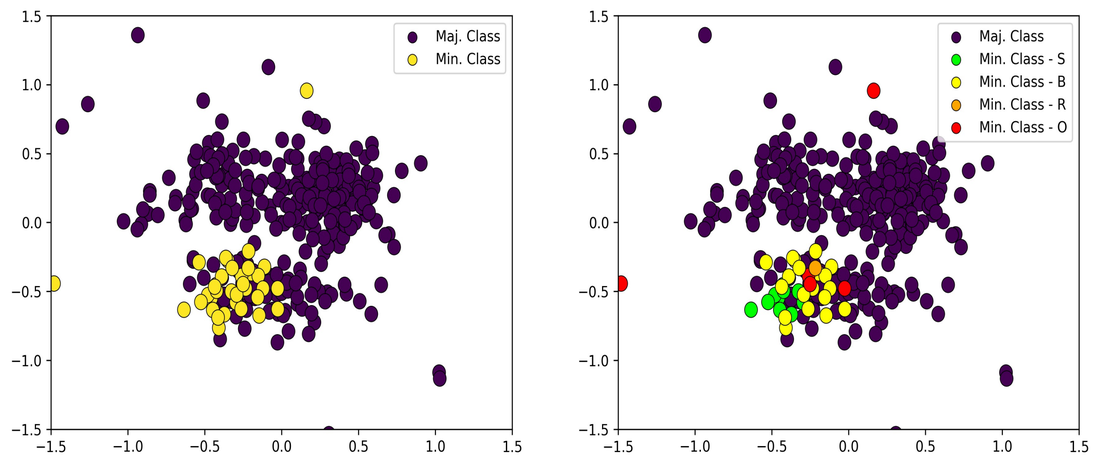

Figure 2: MDS visualisation of the *Ecoli* dataset before and after classifying the minority class

With our framework, a three-level architecture can be built corresponding to the hierarchy of the servers in the substations. The system **monitors in real-time the evolution of the typology of the points in the minority class**.

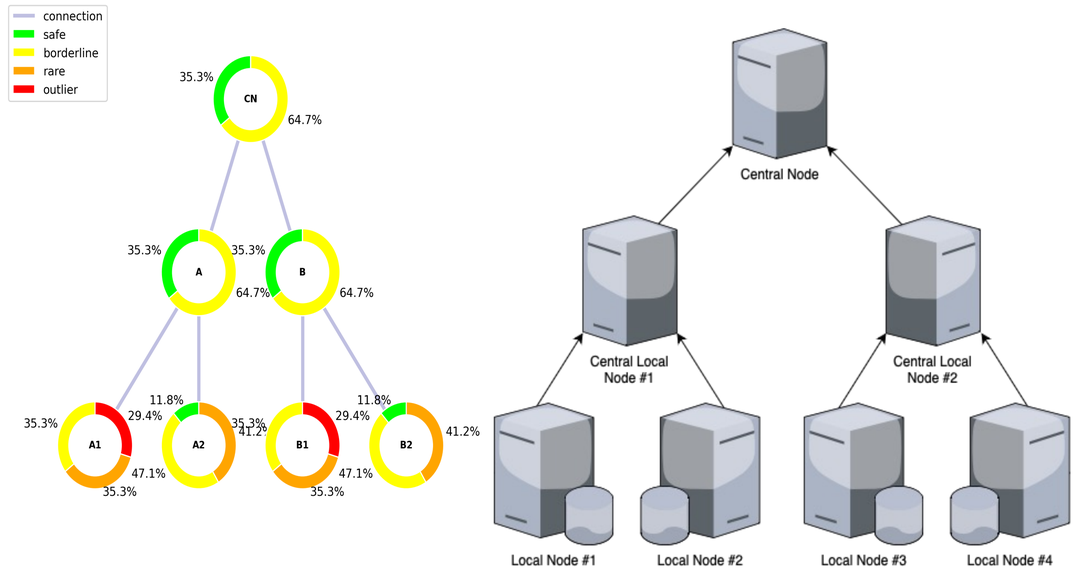

Figure 3: Real-time visualization of the three-level FL system

In Figure 3, at the bottom level, the **Local Nodes (LNs)** are fed with data sent via MQTT protocol by the sensors. The data is aggregated in their **Central Local Nodes (CLNs)** at the second level. Finally, at the top level, the **Central Node (CN)** aggregates the data from the CLNs. The results of the typology of the points are sent to a GUI that shows the percentage for each node.

Conclusion

Considering the current evolution of the typology, the system can decide to **train the classifier with its new data** to achieve a better performance in detecting attacks. Since the data generated from the IoT devices are network logs with a temporality component, using **Time-Series Classifiers (TSC)** is more adequate to solve this problem.

References:

1) Hassan Farhangi. The path of the smart grid. *IEEE Power and Energy Magazine*, 8(1):18–28, 1 2010. ISSN 15407977. doi: 10.1109/MPE.2009.934876.
2) Xi Fang, Satyajayant Misra, Guoliang Xue, and Dejun Yang. Smart grid - The new and improved power grid: A survey, 2012. ISSN 1553877X.
3) H. Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, and Agüera Arcas. Communication-Efficient Learning of Deep Networks from Decentralized Data. In *Artificial intelligence and statistics*, pages 1273–1282, 2 2016.
4) Krystyna Napierala and Jerzy Stefanowski. Types of minority class examples and their influence on learning classifiers from imbalanced data. *Journal of Intelligent Information Systems, 46(3):563–597, 6 2016. ISSN 15737675. doi: 10.1007/s10844-015-0368-1.*
5) *Chawla, N. v., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. Journal of Artificial Intelligence Research, 16, 321–357. https://doi.org/10.1613/JAIR.953*

This work was supported in part by the Fundo Europeu de Desenvolvimento Regional (FEDER)-Competitiveness and Internationalization Operational Program (COMPETE 2020), Portugal 2020 Framework, in the scope of the Smart5Grid Project, under Grant POCI-01-0247- FEDER-047226;