A Graph Neural Network Multi-Task Learning-Based Approach for Detection and Localization of Cyberattacks in Smart Grids

ICASSP 2023 – Rhodes Island, Greece

Abdulrahman Takiddin, Rachad Atat, Muhammad Ismail, Katherine Davis, Erchin Serpedin

Introduction

- Smart grids rely on measurement data to ensure proper supply and demand management and system stability.
- The cyber-physical nature of smart power grids makes them vulnerable to false data injection attacks (FDIAs) where malicious entities manipulate power system measurement data.
- Such attacks can bypass traditional bad data detectors, which lead to making wrong decisions and may result in system instability.
- Existing FDIA detectors perform one of two tasks, detection or localization and offer limited detection performance.

Contributions

We propose a multi-task learning graph neural network (GNN) detector that offers the following:

- It performs two tasks: graph classification to determine the system status (under attack/normal operation) and node classification to localize the attack (the attacked node). This is performed efficiently using a three-stage GNN with joint, task-specific, and fusing layers.
- It captures the complex patterns of measurement data and spatial aspects of power grids using convolutional Chebyshev graph layers. It is examined against FDIAs on IEEE 14, 39, and 118-bus systems.
- It makes decisions based on features learned from both tasks and offers enhanced system status identification and attack localization with detection rates (DRs) of 98.5-100% and 99-100%, respectively.

System Modeling

Power System Modeling

- Modeled using an undirected graph where buses and power lines are represented by nodes and edges, respectively.
- **Spatial aspects:** we adopt IEEE 14, 39, and 118-bus systems.
- **Temporal aspects:** electric power injections and flows.
- Power flow analysis using Newton's method is carried out to determine the real and reactive power flows in the system.

Benign Samples

- Measurement data during normal operation.
- Include 96 daily power dynamics timestamps over six months.

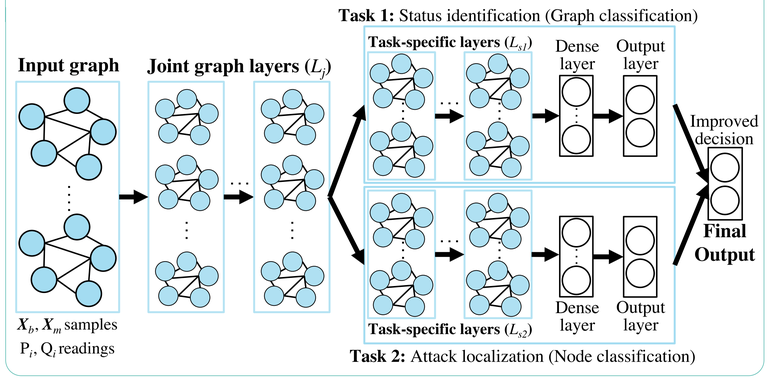
Attack Data

- We adopt three FDIA functions that are used to mimic the power system's operation under attack.
- **Direct attack:** randomly applies bounded perturbations into samples.
- **Replay attack:** uses a false repetition of a reading from a previous timestamp to replace the reading of a current timestamp.
- **General attack: g**enerates malicious samples using a range of true measurement values.
- We use an equal number of benign and malicious samples.

Model Architecture

We propose a multi-task learning-based approach that performs system status identification and attack localization simultaneously using a GNN with stacked convolutional Chebyshev graph layers. The model is divided into three stages:

- The first stage presents the joint graph layers that are used to extract preliminary features from the data that are needed for both tasks.
- The second stage consists of task-specific graph layers that are designated to capture relevant features for a specific task.
- The third stage is the final decision of the two tasks based on previously learned features, which boosts the performance.
- We adopt such a structure to calculate the initial weights and parameters (shared parameters) once, which are then transmitted to the next stage for further processing.



Experimental Results

- We examine task-specific benchmark detectors that perform tasks separately including auto regressive integrated moving average (ARIMA), support vector machine (SVM), multi-layer perceptron (MLP), recurrent neural network (RNN), convolutional neural network (CNN), autoencoder (AE), convolutional graph neural network (CGNN), and graph AE (GAE).

Task 1 (system status identification)

- The proposed multi-task GNN detector offers superior DR by 23.2 - 30.1%, 15 - 21.2%, and 9.4 - 11.5% compared to shallow, deep, and graph-based benchmarks, respectively.

Task 2 (attack localization)

- The proposed multi-task GNN detector offers superior DR by 20.2 - 26.6%, 9.6 - 17.3%, and 4.1 - 5.8% compared to shallow, deep, and graph-based benchmarks, respectively.

Conclusions

- Smart power grids are subject to cyber false data injection attacks (FDIAs) where detection and localization of attacks are critical.
- The proposed detector performs both tasks with improved detection performance compared to task-specific detectors.
- The proposed there-stage structure helped in boosting the detection performance since the final decision is based on the outputs of the two tasks (graph and node classification).

Acknowledgment

This work is supported by NSF EPCN Awards 2220346 and 2220347.