Power Grid Anomaly Detection via Hybrid LSTM-GIN Model (Student Abstract)

Amelia Jobe[*1], **Richard Ky**[*2], **Sandra Luo**[3],
Akshay Dhamsania[4], **Sumit Purohit**[5], **Edoardo Serra**[1]

[1]Boise State University,
[2]San José State University,
[3]University of Texas at Dallas,
[4]Texas A&M University,
[5]Pacific Northwest National Laboratory,
ameliajobe@u.boisestate.edu, richard.ky@sjsu.edu, sandra.luo@utdallas.edu,
akshay.dhamsania2002@gmail.com, sumit.purohit@pnnl.gov, edoardoserra@boisestate.edu

Abstract

Cyberattacks on power grids pose significant risks to national security. Power grid attacks typically lead to abnormal readings in power output, frequency, current, and voltage. Due to the interconnected structure of power grids, abnormalities can spread throughout the system and cause widespread power outages if not detected and dealt with promptly. Our research proposes a novel anomaly detection system for power grids that prevents overfitting. We created a network graph to represent the structure of the power grid, where nodes represent power grid components like generators and edges represent connections between nodes such as overhead power lines. We combine the capabilities of Long Short-Term Memory (LSTM) models with a Graph Isomorphism Network (GIN) in a hybrid model to pinpoint anomalies in the grid. We train our model on each category of nodes that serves a similar structural purpose to prevent overfitting of the model. We then assign each node in the graph a unique signature using a GIN. Our model achieved a 99.92% accuracy rate, which is higher than a version of our model without structural encoding, which had an accuracy level of 97.30%. Our model allows us to capture structural and temporal components of power grids and develop an attack detection system with high accuracy without overfitting.

Introduction

As power grid systems become more complex, they become more vulnerable to cyberattacks. Attacks typically result in abnormal readings in power output, frequency, current, and voltage. If not detected quickly, abnormalities can propagate throughout the system, leading to cascading effects such as power outages if not promptly identified and mitigated.

Many related works have studied the problem of cyberattacks on power grids through supervised machine learning or graph neural networks. These models focus on classifying power grid patterns as normal or anomalous. Some research has developed anomaly detection models, but these models do not capture critical structural and temporal information of the power grid. To the best of our knowledge, no research has developed an anomaly detection system for power grids that captures key structural and temporal elements of the graph while also preventing overfitting.

Our research utilizes the interconnected structure of power grids to construct a network graph. We employ a hybrid model that merges the capabilities of Long Short-Term Memory (LSTM) models with a Graph Isomorphism Network (GIN) to pinpoint anomalies in the grid with precision. Our model achieved a 99.92% accuracy rate, which is higher than a version of our model without structural encoding, which had an accuracy level of 97.30%. Our work contributes to the larger goal of power grid protection.

Related Work

To protect power grid systems, (Ringsquandl et al. 2021) explores the application of GNNs to electrical grids and found that power grid models with GNNs are more accurate and robust as compared to baseline models.

Several papers highlight the importance of encoding structural information about nodes and edges to improve anomaly detection models. (Cai et al. 2021), (Haghshenas, Hasnat, and Naeini 2023), and (Boyaci et al. 2021) found that modeling the power grid structure in their models significantly improved model accuracy.

In (Presekal et al. 2023), researchers combined a Long Short-Term Memory Model with a Graph Convolution Network to create a hybrid model. This model performed with high accuracy in classification-based anomaly detection in power grid systems and was capable of generating an attack graph map in near real time. (Haghshenas, Hasnat, and Naeini 2023) developed a temporal graph neural network to detect false data injection attacks. While a significant amount of research has been done on anomaly classification of power grids, to our knowledge there is no existing research utilizing a joint LSTM-GIN model to capture the structural and temporal elements of a power grid network as well as perform detailed anomaly detection.

Methodology

The power grid we used contains 123 nodes and was generated using NATI[P]G (Bel et al. 2023). Our work involves

[*]These authors contributed equally.

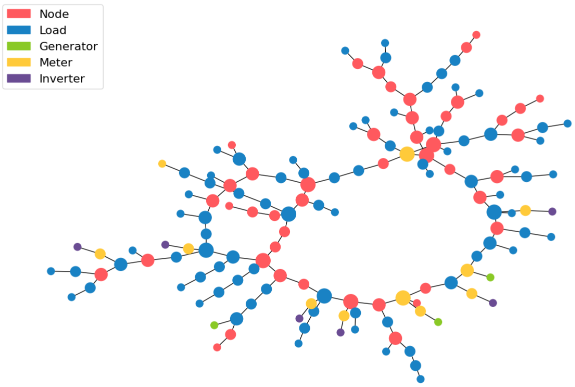

Figure 1: Power Grid Network Graph

two data sets: a baseline data set, which consists of uniform data, and an anomaly data set, which contains discernible attacks at 13, 23, and 26 seconds.

During data analysis, we assessed the real energy levels for all generators and inverters and the voltage consumption levels of each load node within the data sets. We calculated oscillation levels for each node in the baseline data set and during each attack in the anomalous data set. The anomaly levels in the baseline data were near-zero. The anomalous data set, in contrast, demonstrated high levels of variation within the power grid network during all three attacks. The anomaly levels across impacted nodes were not uniform. Nodes closest to the attack source(s) registered the highest level of anomaly oscillation, while those farthest from the source(s) exhibited minimal to no variation.

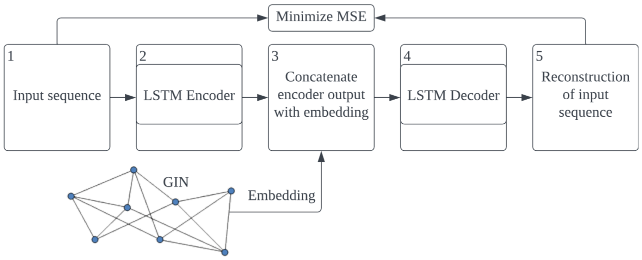
Minimize MSE
1
Input sequence
2
LSTM Encoder
3
Concatenate encoder output with embedding
4
LSTM Decoder
5
Reconstruction of input sequence
GIN
Embedding

Figure 2: Model Framework

The combination of LSTM and GIN models allows us to capture temporal and structural information relevant in the context of power grids. Nodes with a greater number of direct or indirect connections are more likely to experience load fluctuations. As such, it is crucial to incorporate the graph structure of the power grid into our model's design.

The first LSTM in our model is an encoder that uses a sequence of ten timestamps to predict the feature values at the next timestamp. However, since the graph structure has not yet been taken into consideration, the output is not final.

For the next step, we train a GIN on a graph representation of the power grid to produce an embedding for each node.

Model Version	F1 Score Accuracy
Model without GIN	97.30%
Model with GIN	99.92%

Table 1: Anomaly detection accuracy comparison between models with and without the GIN

The embeddings encapsulate the topology of the node's neighborhood and act as unique signatures. We concatenate the output from the first LSTM with the corresponding signatures and use the concatenated result as input for the next LSTM.

The second LSTM is a decoder that uses the node-specific signatures to generate a more informed prediction of the feature values at a specific time given the previous ten timestamps. We employ the mean squared error as the loss function during training and also as a means to determine what percentage of the data has been deemed anomalous by the model. When the error exceeds a chosen threshold, we know the input must differ significantly from the baseline training data and is thus anomalous. To identify the right threshold to classify whether an instance is anomalous or normal, we selected as the threshold the 0.99 percentile computed on the MSE reconstruction errors of all the training samples.

Experimental Findings

The model is trained using data from multiple nodes and subsequently tested on data from a node not previously exposed to the model. Comparing each model prediction to whether a data point is actually anomalous allows us to calculate the F1 score and evaluate the performance of our model. Our results indicate that the model is capable of accurately differentiating between normal and anomalous data. The attack data can be categorized into two types: trivial attacks, which are easily detectable, and pernicious attacks that could elude detection by less sophisticated models. Our model's robust performance in accurately identifying these diverse attacks underpins its utility in enhancing power grid security.

Acknowledgments

This research was made possible by the Research Experience for Undergraduates (REU) program in Data Driven Security at Boise State University sponsored by National Science Foundation (CCF 1950599).

References

Bel, O.; Kim, J.; Hofer, W. J.; Maharjan, M.; Purohit, S.; and Niddodi, S. 2023. Co-Simulation Framework For Network Attack Generation and Monitoring. arXiv:2307.09633.

Boyaci, O.; Umunnakwe, A.; Sahu, A.; Narimani, M. R.; Ismail, M.; Davis, K. R.; and Serpedin, E. 2021. Graph neural networks based detection of stealth false data injection attacks in smart grids. *IEEE Systems Journal*, 16(2): 2946–2957.

Cai, L.; Chen, Z.; Luo, C.; Gui, J.; Ni, J.; Li, D.; and Chen, H. 2021. Structural temporal graph neural networks for anomaly detection in dynamic graphs. In *Proceedings of the 30th ACM international conference on Information & Knowledge Management*, 3747–3756.

Haghshenas, S. H.; Hasnat, M. A.; and Naeini, M. 2023. A Temporal Graph Neural Network for Cyber Attack Detection and Localization in Smart Grids. In *2023 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, 1–5.

Presekal, A.; Ştefanov, A.; Rajkumar, V. S.; and Palensky, P. 2023. Attack Graph Model for Cyber-Physical Power Systems using Hybrid Deep Learning. *IEEE Transactions on Smart Grid*, 1–1.

Ringsquandl, M.; Sellami, H.; Hildebrandt, M.; Beyer, D.; Henselmeyer, S.; Weber, S.; and Joblin, M. 2021. Power to the Relational Inductive Bias: Graph Neural Networks in Electrical Power Grids. In *Proceedings of the 30th ACM International Conference on Information & Knowledge Management*, CIKM ’21, 1538–1547. New York, NY, USA: Association for Computing Machinery. ISBN 9781450384469.