ATTACK DETECTION IN SMART GRID USING MACHINE LEARNING METHODS

Mr. A. Saibabu[1], K. Sireesha[2], T. Keerthana[3], K. Meena[4]

Article History: Received: 11.02.2023 **Revised:** 26.03.2023 **Accepted:** 11.05.2023

Abstract

In the context of the smart grid, a number of attack scenarios are framed as statistical learning problems using batch- or real-time data collection. Machine learning techniques are used in this methodology. Assist in determining whether a metric has been manipulated. The proposed method includes an attack detector that could possibly use previous system information to succeed due to the fragmentary nature of the assignment. We combine decision- and showcase-fusion, common batch- and online-learning approaches, and supervised and semi supervised learning to model the attack detection problem. In order to uncover unobservable attacks by using statistical learning techniques.

Keywords: Smart Grid, Attack, Detection, Training, Power.

[1]Assistant Professor, Department of CSE, Sridevi Women's Engineering College, Hyderabad, Telangana, India.
[2,3,4]UG Student, Department of CSE, Sridevi Women's Engineering college, Hyderabad, Telangana, India.

Email: [1]swecsaibabuyadav@gmail.com, [2]sireesha7879@gmail.com, [3]keerthanatamma@gmail.com, [4]klaxmanlaxman616@gmail.com

DOI: 10.31838/ecb/2023.12.s3.292

1. Scope:

Machine learning (ML) techniques for attack detection in the smart grid are extensive and have a wide range of applications. Intrusion detection, fault detection, malware detection, and cybersecurity risk assessment are some potential applications for ML-based attack detection in smart grids.

2. Introduction

Several machine learning-based approaches for power system monitoring and control have been proposed in the micro-grid literature. Provides present an abstract framework for system design that uses machine learning approaches to forecast when specific system components fail. Machine learning algorithms control the energy and environmental performance. Using machine learning methods at the network level, researchers have overcome the challenges of vulnerability screening and the forecasting of hostile behaviour in communication between smart grid equipment. In this study, we concentrate on the challenge of identifying physical-layer attacks on the smart grid, namely those involving the introduction of bogus data. Using the model of the measurements in a system with this hierarchical structure are clustered, which makes them subject to distributed sparse attacks, as described in, which attempt to tamper with the known locations recorded by either local telecom providers or smart phasor measuring units. Additionally, network administrators who employ statistical learning techniques to find attacks are well-versed in metric matrices, cluster-level measurements, and general network architecture. In this study, we concentrate on the challenge of identifying physical-layer attacks on the smart grid, namely those involving the introduction of bogus data. Additionally, the fake data injection attacks, also known as unobservable attacks, cannot be recognized if the injected vectors are located in the Jacobian matrix's column inches and satisfy specific sparsity constraints. The primary contributions of this work are listed below. We also doubt the veracity of several of the statistical learning theory-based fundamental assumptions underlying the smart grid. Then, we provide general attack building methods, such as semi-supervised and online learning techniques. Second, we examine the impact of attacks based on the insertion of false data into the measurement space on the relationship between measurement vector distances. This inspires the creation of algorithms for measuring the euclidean distance, identifying invisible assaults, calculating the attack methods, and anticipating attacks using previous data. Third, we present actual data demonstrating that statistical learning algorithms outperform SVE-based attack detection algorithms in identifying both observable and unobservable attacks. Additionally, at a value of, the efficiency of machine learning models experiences a phase change.

Related Work

"In Machine Learning for NYC's Electric Grid,"

Preventative maintenance in the power industry may be improved with application of techniques for data mining and automated machine learning. In order to predict the possibility of system and component failures, we present a general approach for transforming data from of the electrical grid's history into models. Utilities might use these models to choose which maintenance tasks to focus on first. Feeder failures, cables, joints, cyborgs, and transformer, feeder MTBF estimates, and manhole event vulnerabilities are all graded using this process, which is iterated several times to obtain the most accurate results. Overall, the approach makes use of state-of-the-art machine learning methods for prioritization, as well as for assessing results via cross-validation as well as blind test, and it is flexible enough to deal with inputs that are diverse, noisy, historical (static), quasi-real-time, or in-the-moment. Our overall modeling approach relies on the fact that deep learning features are meaningful to subject matter experts, that processing of data is transparent, and that forecasting results are reliable enough just to endorse sound choices, allowing for business management integrations beyond ranked lists and MTBF estimates to incorporate the predictive performance directly into corporation planning and decision support. We discuss the challenges of utilizing historical data from the electricity network to make forecasts. Nonetheless, the predictive methods that can be derived from this procedure are sufficiently precise to aid in the maintenance of Nyc City's electrical system, which stands in stark contrast to the "rawness" of these data. Adaptive random control (ASC) for the Smart Grid, powered by approximate dynamic programming (ADP), has the potential to provide the electric grid the autonomous intelligence needed to improve its efficiency and self-healing capacities to those of the internet. In order to maximize Smart Grid management of dispersed generation and storage, we show how to regulate both loads and sources.

"An intrusion detection system for smart grid communications"

The machine-to-machine industry is now most promising in the area of smart grid. Smart grid relies on M2M technologies, which have improved to the point that meters and sensors are no longer anticipated to need human input to characterize power use and distribution. Information from these multiple sensors, such as power usage and monitoring signals, may be sent back. Yet, due to the fact that it is part of an energy management and distribution system, SG necessitates prompt action in the face of hostile events like Smart meter DDoS attacks. The article's Gaussian process is used to model potentially harmful and/or atypical occurrences that might put smart grid users' security and privacy at risk. The model is the basis for a revolutionary early warning system designed to foresee harmful activities in the SG network. The SG command center will be able to foresee these malevolent actions with the help of the warning system, allowing them to take preventative measures. Using computational modeling, we confirm the practicality of the suggested early warning system.

3. Methodology

Several stages are required to identifying attack detection in smart grids using machine learning:

Data collection: First, information is gathered from the smart grid's different sensors and devices, including phasor measuring units (PMUs), SCADA systems, and intelligent electronic devices (IEDs).

Feature Extraction: The next step is to extract features from the Collected data so they can be fed into the machine learning models. Voltage, frequency, power, and energy usage are frequent features utilized in smart grid attack detection.

Data Preprocessing: Preprocessing of the data involves removing noise, handling missing values, and normalising the data once it has been gathered.

Ml Model Selection: An appropriate machine learning model is chosen. Which can include Decision trees, support vector machines (SVMs), logistic regression, and KNN are typical models used in smart grid attack detection.

Model evaluation: Afterward, different metrics, including accuracy, precision, recall, and F1 score, are used to assess the trained model's performance.

Attack Detection: Using the data gathered from the smart grid and the trained model, attacks are finally detected in real time. If an attack is found, an alert is generated so that the attack can be mitigated as necessary.

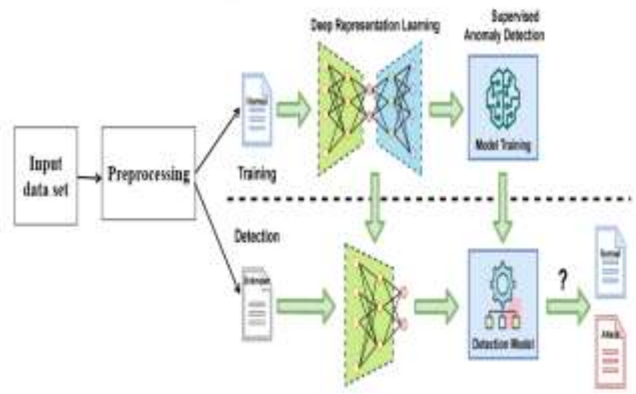
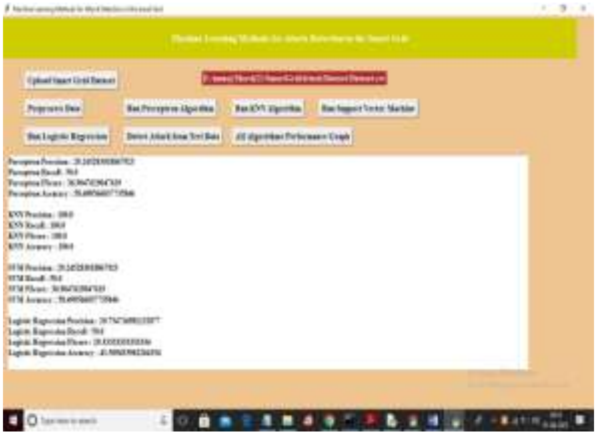


Figure 1 System Architecture

4. Results

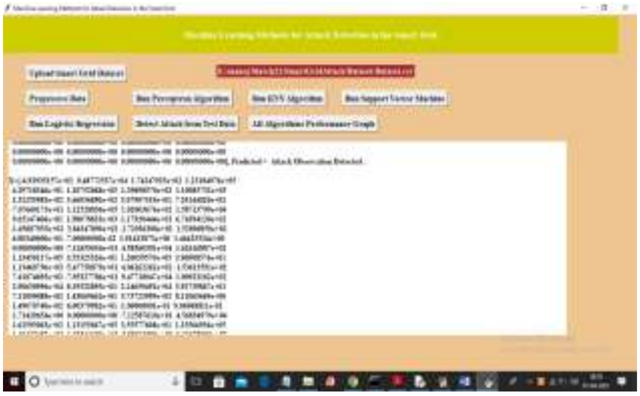
Smart grids, which are energy networks that allow for bidirectional movement of power and data utilizing digital communications technology, are distinguished by their capacity to detect, react to, and proactively handle variations in demand and a range of problems. Intelligent power grids can fix themselves and offer consumers a say in how the grid is run. This smart grid controls the flow of power, the opening and shutting of doors, and other such functions in some "smart cities." A malicious person may launch an attack on the smart grid system by spreading or injecting false information, leading to the execution of the smart grid based on the false information provided.


Machine Learning Methods for Attack Detection in the Smart Grid
Upload Smart Grid Dataset
Preprocess Data
Run Perceptron Algorithm
Run KNN Algorithm
Run Support Vector Machine
Run Logistic Regression
Detect Attack from Test Data
All Algorithms Performance Graph

```
Perceptron Precision : [illegible]
Perceptron Recall : [illegible]
Perceptron FScore : [illegible]
Perceptron Accuracy : [illegible]

KNN Precision : 100.0
KNN Recall : 100.0
KNN FScore : 100.0
KNN Accuracy : 100.0

SVM Precision : [illegible]
SVM Recall : [illegible]
SVM FScore : [illegible]
SVM Accuracy : [illegible]

Logistic Regression Precision : [illegible]
Logistic Regression Recall : [illegible]
Logistic Regression FScore : [illegible]
Logistic Regression Accuracy : [illegible]
```

Figure 2 Run algorithm screen

We tested all four algorithms by clicking their respective buttons and comparing their resulting accuracy, precision, recall, and FSCORE values; ultimately, KNN proved to be the most effective of the four. The ML system can now predict whether or not a class label of "normal" or "attack" is appropriate for the uploaded test data. As can be seen in the test data below, ML will be used to predict a class label when none exists. Predicted results are shown below as ATTACK observation values.


Machine Learning Methods for Attack Detection in the Smart Grid
Upload Smart Grid Dataset
Preprocess Data
Run Perceptron Algorithm
Run KNN Algorithm
Run Support Vector Machine
Run Logistic Regression
Detect Attack from Test Data
All Algorithms Performance Graph

```
[illegible] Predicted = Attack Observation Detected
[illegible]
```

Figure 3 Attack Detected Observed

5. Conclusion

The Performance for various supervised, adequate institutional, classifiers and feature field fusing, and online learning methods for attack detection have been examined for various attack scenarios. Both the compromised and protected metrics are classified as two distinct groups in a reflect the amount classification issue. Using an SVE method, we found that the most cutting-edge machine learning algorithms outperformed the most popular attack detection techniques in our studies, which included the detection of both visible and invisible assaults. We found that among these algorithms, perceptron is the least susceptible to system size and k-NN is the most sensitive. The k-NN is hampered by the issue of skewed data, which makes it less effective. As a result, k-NN may do better than competing algorithms in small systems but poorly in big ones. Machine learning techniques aid in the detection of numerous types of attacks, including spoofing, denial-of-service, and injection of bogus data. Therefore, when an unusual action occurs, the system will assist the user in recognizing the attack.

6. References

C. Rudin et al., “Machine learning for the New York City power grid,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 34, no. 2, pp. 328–345, Feb. 2012.

R. N. Anderson, A. Boulanger, W. B. Powell, and W. Scott, “Adaptive stochastic control for the smart grid,” Proc. IEEE, vol. 99, no. 6, pp. 1098–1115, Jun. 2011.

Z. M. Fadlullah, M. M. Fouda, N. Kato, X. Shen, and Y. Nozaki, “An early warning system against malicious activities for smart grid communications,” IEEE Netw., vol. 25, no. 5, pp. 50–55, Sep./Oct. 2011.

Y. Zhang, L. Wang, W. Sun, R. C. Green, and M. Alam, “Distributed intrusion detection system in a multi-layer network architecture of smart grids,” IEEE Trans. Smart Grid, vol. 2, no. 4, pp. 796–808, Dec. 2011.

T. T. Kim and H. V. Poor, "Strategic protection against data injection attacks on power grids," IEEE Trans. Smart Grid, vol. 2, no. 2, June 2011, pp. 326-333.

O. Chapelle, "Training a support vector machine in the primal," Neural Comput., vol. 19, no. 5, 1155-1178, 2007.

Roger S. Pressman, "Software Engineering: A Practitioner's Approach," 6th edition, McGraw-Hill International Edition.

Software Engineering, Sommerville, Pearson Education, 7th edition.

The user manual for the unified modelling language James Rambaugh, Ivar Jacobson, Grady Booch, and Pearson Education.

Mark Lutz's fourth edition of Programming Python: Powerful Object-Oriented Programming

Combining Pattern Classifiers: Methods and Algorithms, L. I. Kuncheva. USA: Wiley, Hoboken, NJ, 2004.

Boosting: Foundations and Algorithms by R. E. Schapire and Y. Freund. 2012. Cambridge, Massachusetts: MIT Press.