JOURNAL OF SOUTHERN COMMUNICATION ENGINEERING
ISLAMIC AZAD UNIVERSITY BUSHEHR BRANCH

E-ISSN: 2980-9231
https://sanad.iau.ir/journal/jce

https://doi.org/...

Vol. x/ No. x/xxx

Research Article

Detection of False Data Injection Intrusions in Smart Power Grids: A Machine Learning Fusion Approach

Mohammadreza Pourshirazi[1] | Mohsen Simab[2*] | Alireza Mirzaee[3] | Bahador Fani[4]

[1]Department of Electrical Engineering, Faculty of Engineering, University of Marvdasht, Iran.
Mohammadreza.pourshirazi@iau.ac.ir

[2] Department of Electrical Engineering, University of Larestan, Lar, Iran.
msimab@lar.ac.ir

[3]Department of Electrical Engineering, Faculty of Engineering, Islamic Azad University, Dariun Branch, Shiraz, Iran.
alireza.mirzaee@iau.ac.ir

[4]Department of Electrical Engineering, Faculty of Engineering, Isfahan (Khorasgan) Branch, Isfahan, Iran.
bfani@khuisf.ac.ir

Correspondence
Mohsen Simab, Assistant Professor of Electrical Engineering, Faculty of Engineering, University of Larestan, Lar, Iran.
msimab@lar.ac.ir

Main Subjects:
Smart Power Grids

Paper History:
Received: 3 March 2024
Revised: 19 June 2024
Accepted: 23 June 2024

Abstract

In this paper, a novel approach is proposed for False Data Injection (FDI) intrusion detection in smart power grids, which utilizes the fusion of four machine learning (ML) based methods including Random Forest (RF), Gradient Boosting Machine (GBM), Linear Discriminant Analysis (LDA), and Logistic Regression (LR). The results of these methods are integrated using an intelligent algorithm based on the Adaptive Neuro-Fuzzy Inference System (ANFIS) to achieve higher accuracy and efficiency. The base classifiers are trained and tested using optimal state vectors obtained from Phasor Measurement Units (PMUs) using the Random Forest technique Then, after the hyperparameter optimization procedure, by intelligently integrating the predicted outputs of the base models using an ANFIS-based algorithm, it is possible to distinguish normal samples from anomalies with high accuracy and low sensitivity to noise and uncertainty in the PMU data. The IEEE 14-bus power system was utilized to test the proposed method, and the results show that it performs superior to the baseline techniques by achieving a 25%-30% higher detection rate (DR) compared to the base ML methods. By dramatically reducing the false positive rate (FPR) and increasing the detection rate (DR), this method shows that it can be used to improve smart grid security against FDI anomalies.

Keywords: Smart grids, Power systems, Intelligent algorithms, False data injection, Data fusion, Intrusion detection, Machine learning.

Highlights

- A novel ONF-ML scheme combining multiple machine learning models with ANFIS for FDIA.
- Scalability and resilience against varying attack intensities and noisy datasets.
- High detection accuracy (over 95%) and low false positive rates (below 2%) across diverse attack scenarios.
- Experimental validation on IEEE 14-bus system with real-world load data.

Citation: … [in Persian].

1. Introduction

With the rapid advancement of new technologies and the connection of smart devices to communication networks, smart power grids have emerged as one of the vital pillars of modern energy systems. These grids, in the electricity industry, with the ability to collect and process real-time data from measurement units and various sensors, are capable of improving efficiency and reducing costs in energy management systems. However, cyber threats, especially False Data Injection (FDI) intrusions, pose security and performance challenges to the growth of these grids [1]. FDI attacks, in which the data collected from sensors and measurement units are deliberately altered, can disrupt the estimation of power system states by manipulating this data. As a result, there may be poor decisions, system failures, and eventually less grid stability. Attackers can cause anomalies and drastically lower the efficiency of measurement and control systems by altering network data [2, 3].

This paper proposes an innovative approach to detect FDI attacks in energy systems using an integration of multiple machine learning algorithms. This approach integrates machine learning (ML) algorithms including Logistic Regression (LR), Linear Discriminant Analysis (LDA), Random Forest (RF), and Gradient Boosting Machine (GBM). To enhance detection rate accuracy and other performance metrics in identifying FDI anomalies, the results of these algorithms are then combined exploiting an Adaptive Neuro-Fuzzy Inference System (ANFIS). The main goal of this method is to improve detection accuracy and reduce the false positive rate compared to traditional methods that often fail to detect complex FDI attacks.

Historical data extracted from Phasor Measurement Units (PMUs) are used for preprocessing and extracting distinguishing features from normal and infiltrated states [4]. After dimensionality reduction and feature selection using the Random Forest technique and the Gini index, optimal state vectors are prepared for training and testing the models. Subsequently, the base algorithms RF, GBM, LDA, and LR are trained separately, and their outputs are combined via ANFIS for final detection. One of the notable features of this study is the use of the ANFIS algorithm to integrate the outputs of various models, enabling the system to more accurately identify disturbed states. This approach, in addition to its high accuracy, demonstrates lower sensitivity to the magnitude of FDI intrusions, which is particularly important in real-world environments with substantial data fluctuations and dynamic conditions of smart grids.

2. Innovation and contributions

One of the main achievements of this study is the increased detection accuracy of intrusions. The proposed method is capable of identifying complex and new patterns in the data that separate algorithms cannot detect. The use of Adaptive Neuro-Fuzzy Inference Systems (ANFIS) to combine the outputs of various machine learning algorithms significantly improves the detection accuracy of FDI attacks. The use of ANFIS in this study reduces the sensitivity to the extent of FDI interference and improves the performance of the system against more complex and dynamic attacks, resulting in more accurate and efficient detection. This article presents an integrated approach that provides a new and effective solution for detecting FDI intrusions in smart grids to resolve traditional methods' insufficiencies in detecting subtle intrusions in anomalous data [5].

Among the innovations applied in this study, the following can be stated:

First, the proposed method significantly improves the accuracy of false data injection (FDI) attack detection and is capable of identifying complex and novel patterns in the data that individual algorithms cannot detect. Integrating the Adaptive Neuro-Fuzzy Inference System (ANFIS) with various machine learning algorithms improves detection performance by increasing the detection rate (DR) and reducing the false positive rate (FPR). This improvement in performance leads to more accurate detection of false data injections.

Secondly, this method reduces sensitivity to the range of FDI attacks. Traditional methods are usually sensitive to variations in attack data; however, the use of ANFIS in this study reduces the system's sensitivity to different magnitudes of FDI attacks. This feature makes the proposed system more resilient to complex and dynamic attacks and enables it to detect anomalies with higher accuracy and efficiency. In addition, this method has high flexibility, making it easy to adapt to different types of ML algorithms and datasets.

The efficiency and adaptability of the system to different operating conditions can be improved by the ability to adjust the ANFIS parameters, allowing performance optimization in different scenarios. Furthermore, this flexibility makes the system more robust to changing circumstances. Finally, another important contribution is scalability. This approach is appropriate for high-performance distributed systems that are growing because the proposed method is capable of working with large-scale power networks. Scalable and effective detection systems are essential to protect against FDI attacks as smart grids expand. This need is met by the proposed approach, which allows to effectively detect FDI attacks in distributed and scalable systems.

3. Materials and Methods

The data used were obtained from Phasor Measurement Units (PMU). These data include current and voltage measurements from various points in the power grid, which are used in the process of system state estimation. The mentioned data usually include normal values and values contaminated with False Data Injection (FDI) attacks. The initial pre-processing of the data includes normalization of values, extraction of critical features

required for detecting FDI intrusions, and cleaning of missing and anomalous data. Feature extraction methods were applied to find unusual patterns in the data. The differences between FDI and normal states are identified by extracting distinctive features. These features consist of the standard deviation, mean, and other statistical attributes that are derived from the PMU data. The number of features is then decreased, and key features are chosen, using dimensionality reduction techniques like Random Forest. To increase the models' precision and effectiveness, this feature selection is carried out using metrics like the Gini index. The mean, standard deviation, and other statistical traits derived from the measurement data may be among these attributes.

In this paper, four basic ML algorithms have been used for detecting FDI intrusions. Random Forest (RF), Gradient Boosting Machine (GBM), Linear Discriminant Analysis (LDA), and Logistic Regression (LR) are the basic algorithms that are trained separately to classify the data into normal and anomalous categories based on the extracted features. RF has been employed as one of the main choices because, due to its high power in preventing overfitting, it can enhance the accuracy of the detection system and prevent false errors. GBM, by combining multiple weak prediction models, produces a strong predictive model that can be useful for identifying complex patterns in the data. LDA is used to reduce data dimensions and classify them based on distinguishing features, and LR is used to predict the probability of occurrence of a specific event (here, FDI intrusion) based on input data features. In the next step, the outputs of the basic algorithms are combined using an Adaptive Neuro-Fuzzy Inference System (ANFIS). This approach is exploited as an intelligent model for combining the outputs of different ML algorithms. ANFIS automatically adjusts fuzzy parameters and optimizes the incorporated model for detecting FDI intrusions. This method is particularly effective in identifying complex anomalies in the data, which individual algorithms are unable to detect.

4. Results and Discussion

The results of simulations were performed under two scenarios—incremental and decremental false data injections—into the PMU measurement system of the standard IEEE 14-bus power network. The baseline algorithms and the proposed hybrid model were used to evaluate the system's ability to detect FDIs. The proposed method achieved a detection rate of 93.2% in the first scenario, with incorrect data being inserted into PMU measurements on bus number 2 at a rate of -6%. This significantly outperformed base methods such as RF (66.3%) and GBM (64.3%).

Furthermore, the proposed method achieved the highest values compared to other performance metrics such as precision, accuracy, and recall. Specifically, 134 of 144 intrusion samples were detected and 861 of 864 normal samples were correctly classified. In the second scenario, which involved +6% incremental changes to the PMU measurements, the proposed method again demonstrated superior performance. It identified 131 out of 144 intrusion samples and correctly classified 859 out of 864 normal samples, achieving a detection rate of 90.9%, which was higher than other algorithms. In the third scenario, which analyzed the system’s sensitivity to FDI variations in the range of -20% to +20%, the results showed that the proposed method is significantly resistant to various FDI changes and maintains high performance even under minor FDI injections.

Finally, the comparison results with the method in the paper [6] show that the proposed method has a higher detection rate compared to the previous method under different injection conditions, especially at 5% and 2% injection. These results demonstrate the superiority of the proposed scheme in detecting FDI and enhancing the performance of detection systems, especially under complex and noisy conditions.

5. Conclusion

This paper presents an intelligent hybrid approach for smart power grid False Data Injection (FDI) intrusion detection. The suggested method has greatly increased detection accuracy and decreased the false positive rate by combining machine learning algorithms with Adaptive Neuro-Fuzzy Inference Systems (ANFIS). The results show that this approach outperforms traditional systems in terms of accuracy and efficiency and can detect complex and unknown patterns in data by leveraging adaptive and hybrid features.

The proposed approach is more robust against dynamic and complex attacks due to its high accuracy and lower sensitivity to the various FDI attacks. The flexibility of this method in adapting to various machine learning algorithms and different datasets is one of its main strengths. This feature enables effective use of the proposed method in large networks and massive datasets, making it a suitable option for distributed systems and expanding smart grids.

Simulations performed have evaluated the performance of the proposed method under different scenarios of incremental and decremental FDI injections. The results show that in the first scenario (decremental injection), the proposed method, with the detection of 134 out of 144 intrusion samples (detection rate of 93.2%) and 861 out of 864 normal samples, has performed significantly better than basic methods such as RF (66.3%) and GBM (64.3%). In the second scenario (incremental injection), the proposed method, with the detection of 131 out of 144 intrusion samples (detection rate of 90.9%) and 859 out of 864 normal samples, demonstrated high performance. In the system sensitivity analysis, the performance of the proposed method remained stable even

against minor variations in FDI attacks, and the reduction in detection rate and F1 score was noticeably less than that of basic methods.
These results show that the proposed approach can significantly improve the security of smart grids due to its reliable and widely applicable performance. Last but not least, this study provides a scalable and effective framework that improves the performance of intrusion detection systems and acts as a reliable protection against cyber threats in smart grids. In addition to successfully improving the stability and security of these systems, this approach can serve as a basis for future studies and advances in the field of power grid security.

6. Acknowledgement
This paper is extracted from the doctoral dissertation in the field of Electrical-Power Engineering, approved and defended at the Islamic Azad University, Marvdasht Branch. The authors would like to express their sincere gratitude to the Department of Electrical Engineering and the research authorities of the Islamic Azad University, Marvdasht Branch.

References

[1] M. A. Husnoo, A. Anwar, N. Hosseinzadeh, S. N. Islam, A. N. Mahmood, and R. Doss, "False data injection threats in active distribution systems: A comprehensive survey," *Future Generation Computer Systems,* vol. 140, pp. 344-364, 2023.

[2] Q. Wang, W. Tai, Y. Tang, and M. Ni, "Review of the false data injection attack against the cyber-physical power system," *IET Cyber-Physical Systems: Theory & Applications,* vol. 4, no. 2, pp. 101-107, 2019.

[3] M. Zhang *et al.*, "False data injection attacks against smart gird state estimation: Construction, detection and defense," *Science China Technological Sciences,* vol. 62, no. 12, pp. 2077-2087, 2019.

[4] S. Y. Derakhshandeh and R. Pourbagher, "Robust coupled single-port method based on PMU-based state estimation method for voltage stability assessment," *International Journal of Electrical Power & Energy Systems,* vol. 151, p. 109150, 2023.

[5] M. Ashrafuzzaman, S. Das, Y. Chakhchoukh, S. Shiva, and F. T. Sheldon, "Detecting stealthy false data injection attacks in the smart grid using ensemble-based machine learning," *Computers & Security,* vol. 97, p. 101994, 2020.

[6] M. Mohammadpourfard, A. Sami, and Y. Weng, "Identification of false data injection attacks with considering the impact of wind generation and topology reconfigurations," *IEEE Transactions on Sustainable Energy,* vol. 9, no. 3, pp. 1349-1364, 2017.

Declaration of Competing Interest: Authors do not have conflict of interest. The content of the paper is approved by the authors.

Author Contributions: