Advanced Journal of Engineering and Scientific Applications (AJESA)

Research Paper Open Access

Available online at www.ajesa.com.ng Vol.2, No.1, 2025.pp.21-27

A Computational Framework for Smart Grid Tampering Detection Using Unsupervised Machine Learning

[1]B.O.Osuesu, [2]G.O. Asoronye, [3]Ilokanuno O., [4]B.J. Robert, [5]A.C. Okafor, [6]R.I. Ogbuikwu and [7]N.J. Akpan

[1]Electrical Electronic Engineering Department, Akanu Ibiam Federal Polytechnic Unwana, Afikpo, Ebonyi State, Nigeria

[2]Computer Engineering Department, ,Akanu Ibiam Federal Polytechnic Unwana, Afikpo, Ebonyi State, Nigeria

[3]Electronic and Computer Engineering Department, Nnamdi Azikiwe University, Awka, Nigeria

[4]Electrical Electronic Engineering Department, Akanu Ibiam Federal Polytechnic Unwana, Afikpo, Ebonyi State, Nigeria

[5,6,7]Electrical Electronic Engineering Department, Federal Polytechnic Ohodo, Enugu State, Nigeria

***Corresponding Author: [2]asorgay@gmail.com**

Abstract

This paper presents a computational framework for detecting tampering in smart grids using unsupervised machine learning algorithms. Leveraging Advanced Metering Infrastructure (AMI) and integrating algorithms such as K-Nearest Neighbourhood with Isolation Forest (KNN + IF), Local Outlier Factor (LOF), and Lightweight Online Detector of Anomaly (LODA), the framework achieves robust tampering detection. Experimental evaluations on a smart grid IoT Coordinating Infrastructure (SG_IoT_CI) showed that the proposed system significantly improved tampering detection accuracy, achieving up to 31.11% accuracy with KNN + IF compared to lower-performing methods. Additionally, the framework demonstrated enhanced service rates, with KNN + IF yielding a 35.09% improvement and LODA achieving the highest throughput at 28.12%. Latency was reduced by 11.98% for KNN + IF and 8.98% for LODA. These results underscore the efficiency of the proposed approach in addressing zero-day attacks and reducing false positives, paving the way for scalable and secure smart grid ecosystems.

Keywords: Smart Grid Security, Tampering Detection, Unsupervised Machine Learning, Advanced Metering Infrastructure (AMI), Anomaly Detection

1. INTRODUCTION

Smart grids are advanced electrical systems that integrate digital communication technologies with traditional power grids to enhance their efficiency, reliability, and sustainability (Teixeira et al., 2020). Key components include the Advanced Metering Infrastructure (AMI) (Guerrero et al., 2023), which enables two-way communication between utility providers and consumers and Distributed Energy Resources (DER) (Liu et al., 2024), which incorporate renewable energy sources like solar and wind into the grid (Bhutta et al., 2024). Another critical aspect is the Internet of Things (IoT) which facilitates real-time monitoring and control of grid operations by connecting devices and sensors across the network (Chen et al., 2019). Despite their numerous advantages, smart grids face persistent challenges, including tampering (Bhattacharjee et al., 2017), which undermines their reliability and security (Dehghani et al., 2023).

Tampering in smart grids refers to unauthorized activities, such as energy theft, manipulation of meter readings, or disruption of grid components (Althobaiti et al., 2021). This not only causes significant financial losses to utility providers but also compromises the stability of power supply. In countries like Nigeria, where energy theft is prevalent, tampering exacerbates existing issues such as inadequate power distribution and frequent outages (Althobaiti et al.,

2021). Cyber-attacks targeting the communication and control systems of smart grids further escalate the risks, making it imperative to address these vulnerabilities (Sahani et al., 2023).

Traditional tampering detection methods rely heavily on rule-based systems and supervised machine learning models, which are often limited in their ability to adapt to dynamic and evolving threats. Rule-based systems depend on predefined patterns, making them ineffective against novel or complex tampering scenarios (Liu et al., 2021). Supervised machine learning models, on the other hand, require extensive labeled datasets for training, which are not always available or reliable (Khalid et al., 2024). Additionally, these models often struggle with high false-positive rates, leading to unnecessary interventions and reduced trust in the system (Husnoo et al., 2024). Another limitation is the lack of scalability in existing solutions (Zhang et al., 2023). As smart grids expand to accommodate more consumers and integrate additional renewable energy sources, the volume and complexity of data increase exponentially. Many traditional systems are unable to process this data efficiently, resulting in delayed or inaccurate tampering detection (Radoglou-Grammatikis et al., 2020). Furthermore, these systems often fail to account for privacy concerns, as they rely on detailed consumer energy usage data for analysis (Samanthula & Patel, 2023). This can deter consumers from adopting smart grid technologies, thereby hindering their widespread implementation. In this context, unsupervised machine learning algorithms offer a promising alternative (Almalki, 2024). Unlike supervised models, unsupervised algorithms do not require labeled data, making them suitable for identifying anomalies in real-time. By analyzing patterns and deviations in energy consumption data, these algorithms can detect tampering activities without prior knowledge of their specific characteristics (Zhang et al., 2023). Additionally, integrating unsupervised learning with IoT and cloud computing infrastructure enhances the scalability and efficiency of detection systems, enabling them to address the challenges posed by large-scale smart grid deployments (Mohammadi et al., 2023). This paper introduces a computational framework that leverages unsupervised machine learning to detect tampering in smart grids. By addressing the limitations of existing solutions, the proposed framework aims to enhance the security and reliability of smart grid ecosystems, paving the way for their sustainable growth and adoption.

1.1 Related Works

Machine learning techniques have increasingly become pivotal in addressing cybersecurity challenges within smart grid ecosystems (Bhutta et al., 2024), particularly in detecting electricity theft and anomalous behaviors. Recent research has explored diverse approaches to tackle the complex problem of smart grid security detection.

1.2 Machine Learning Approaches for Anomaly Detection

Contemporary studies have demonstrated significant advancements in electricity theft detection methodologies. Li et al.,(2019) proposed a Deep Neural Network (DNN) model that effectively resolved dimensionality challenges and selected critical features for fraud detection in smart grid systems. Similarly, Iftikhar et al.,(2024) developed a hybrid deep learning model combining Multi-Layer Perceptron (MLP) and Gated Recurrent Unit (GRU) for comprehensive electricity theft detection.

1.3 Unsupervised Learning Techniques

Researchers have increasingly focused on unsupervised learning methods to address key challenges in areas such as detecting zero-day attacks in complex cyber-physical systems (Miller et al., 2024), identifying unknown anomalies in real-time data streams (Huang et al., 2024), and handling class imbalance in electricity consumption datasets (Yan & Wen, 2022). Innovative approaches have emerged, such as the ensemble technique proposed by Zhu et al.,(2024), which leverages Advanced Metering Infrastructure (AMI) to assess customer metering data and energy usage patterns. This approach distinguishes between classification-based and clustering-based schemes, providing a meticulous approach to electricity theft detection.

1.4 Challenges in Smart Grid Security

The existing literature highlights critical challenges in tampering detection such as High dimensionality of smart grid data (Kaur et al., 2018), Complex class imbalance in electricity consumption patterns (Syed et al., 2020), Detection of Non-Technical Losses (NTLs) (Esmael et al., 2021), and in the handling of diverse theft patterns (Pamir et al., 2023).

1.5 Performance Evaluation

Recent studies have emphasized multiple performance metrics. For example, accuracy is a common metric used in many studies (Zhang et al., 2019), False positive rates are also considered, particularly in applications where minimizing false alarms is crucial (Wahed et al., 2025). Detection efficiency is another important metric, often expressed as true positive rate or other related measures (Althobaiti et al., 2021). Finally, computational complexity is a key consideration, especially for real-time applications or resource-constrained environments (Bhattacharjee et al,. 2017). The research papers highlight the importance of considering these metrics individually and in combination to provide a comprehensive evaluation of model performance.

1.6 Contribution of Current Research

The proposed framework builds upon these foundations by integrating multiple unsupervised machine learning algorithms, addressing limitations in existing detection approaches, and providing a comprehensive computational framework for smart grid tampering detection. While existing research has made significant strides, a critical need remains for robust, scalable, and efficient tampering detection mechanisms that can adapt to evolving cyber threats in smart grid ecosystems.

2. INTRODUCTION

2.1 System Architecture

As depicted in Figure 1, the SG_IoT_CI system integrates AMI with cloud infrastructure to enable real-time monitoring and tampering detection. The architecture consists of three layers: edge devices (smart meters), fog computing for preliminary data processing, and cloud computing for advanced analytics and storage.

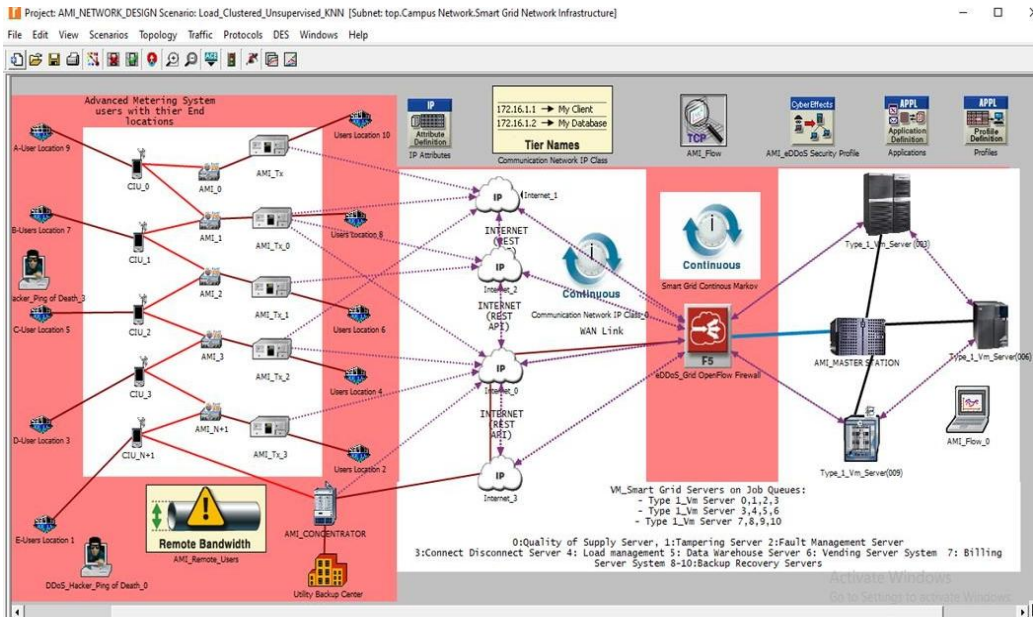

Figure 1: Architecture of the SG_IoT_CI system.

2.2 Computational Framework

The proposed computational framework integrates multiple unsupervised machine learning algorithms to detect tampering within the smart grid ecosystem. The methodology involves using K-Nearest Neighbourhood with Isolation Forest (KNN + IF), which clusters data points and isolates anomalies through an ensemble of decision trees. Additionally, the framework employs Local Outlier Factor (LOF) to identify outliers by analyzing local density deviations, ensuring improved anomaly detection. Furthermore, the Lightweight Online Detector of Anomalies (LODA) is incorporated as a lightweight model capable of identifying deviations in high-dimensional data streams with minimal computational overhead. By combining these algorithms, the detection system ensures robustness and adaptability to real-time variations in energy consumption patterns, providing a scalable and effective approach for tampering detection in smart grid environments.

2.3 Data Acquisition and Preprocessing

To evaluate the proposed framework, energy consumption data was collected from AMI systems in a controlled testbed environment. The preprocessing steps involved standardizing energy usage patterns through data normalization to eliminate inconsistencies. Feature extraction was then performed to identify key metrics such as voltage fluctuations, power consumption rates, and frequency patterns. Lastly, noise filtering was applied to remove irrelevant or erroneous data points, thereby improving the accuracy of tampering detection.

2.4 Model Training and Implementation

The unsupervised learning models were trained using real-time AMI data through a structured implementation process. Initially, energy consumption records were segmented into manageable time-series datasets, ensuring an organized approach to data handling. Feature engineering was then conducted to extract meaningful attributes necessary for anomaly detection. The selected algorithms were subsequently executed on the test dataset, with detection parameters

being refined to enhance accuracy. Finally, the models were evaluated based on performance metrics such as precision, recall, and response time. By leveraging cloud computing resources, the framework ensures scalability and real-time processing capabilities for large-scale smart grid deployments.

3. RESULTS AND DISCUSSION

3.1 Performance Metrics

To evaluate the effectiveness of the proposed tampering detection framework, several key performance metrics were analyzed, including Full Scale Query Response Time (FSQRT), latency, service rate, and throughput. The results demonstrate that the use of unsupervised learning models significantly enhances tampering detection capabilities compared to traditional methods. The Full Scale Query Response Time (FSQRT) was measured under an open-flow security control model. Results showed that SG_IoT_CI AMI Overflow performed at 47.82%, while Non-Open Flow achieved 52.17%, indicating that real-time processing in open-flow networks is more efficient. Latency, which refers to the delay between data collection and anomaly detection, varied across different models. The Secure SG_IoT_CI AMI latency values were 20.96% for LPBSVR, 11.98% for KNN + IF, 19.31% for SVM, 19.76% for LPBNN, 19.01% for LOF, and 8.98% for LODA. These results indicate that LODA achieved the lowest latency, making it the most efficient algorithm in real-time tampering detection. Service rate measures the efficiency of anomaly detection in handling smart grid transactions. The service rates recorded were 14.03% for LPBSVR, 35.09% for KNN + IF, 5.26% for SVM, 8.77% for LPBNN, 15.79% for LOF, and 21.05% for LODA. The highest service rate was recorded for KNN + IF, demonstrating its efficiency in handling smart grid event processing. Throughput, which reflects the overall capacity of the system to process anomaly detection tasks, varied across the models. The throughput results showed that LPBSVR achieved 19.13%, KNN + IF recorded 25.04%, SVM had 19.27%, LPBNN registered 15.47%, LOF reached 18.28%, and LODA attained the highest throughput at 28.12%. These findings highlight that LODA had the highest throughput, making it the most scalable model for large-scale smart grid deployments.

3.2 Graphical Analysis

To visualize the performance of the proposed framework, key performance indicators such as latency, service rate, and throughput are presented in graphical form.

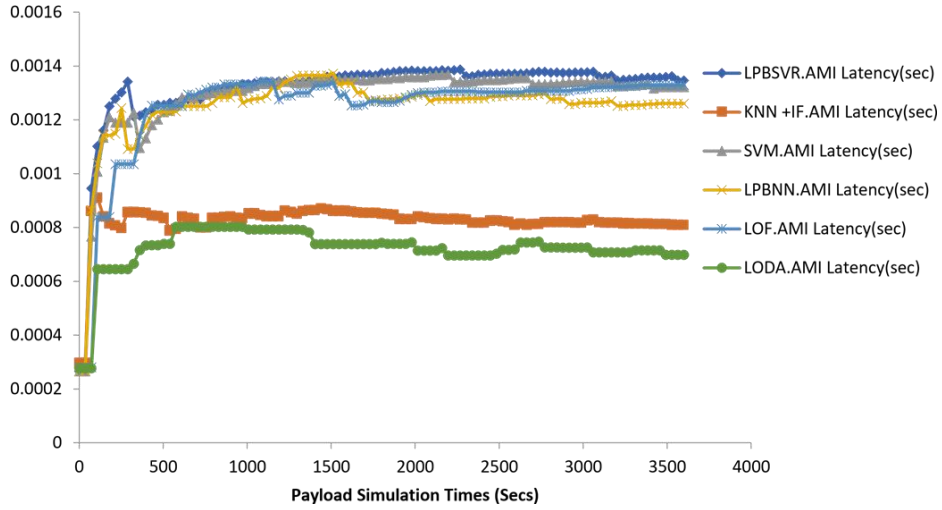

Figure 2: Latency performance

The latency performance comparison as depicted in Figure2 illustrates the response times of different machine learning models, highlighting the efficiency of LODA, which recorded the lowest latency. This graph provides insight into the real-time detection capabilities of each model, showcasing how quickly anomalies can be identified and flagged.

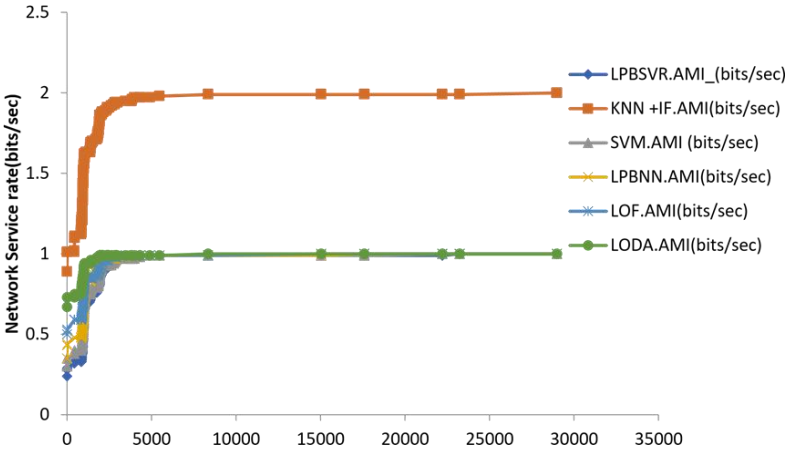

Figure 3: Service rate

The service rate comparison as shown in Figure 3 depicts the efficiency of each algorithm in processing smart grid events. The graph shows that KNN + IF achieved the highest service rate, indicating its capability to handle a large number of tampering detection requests with minimal delay.

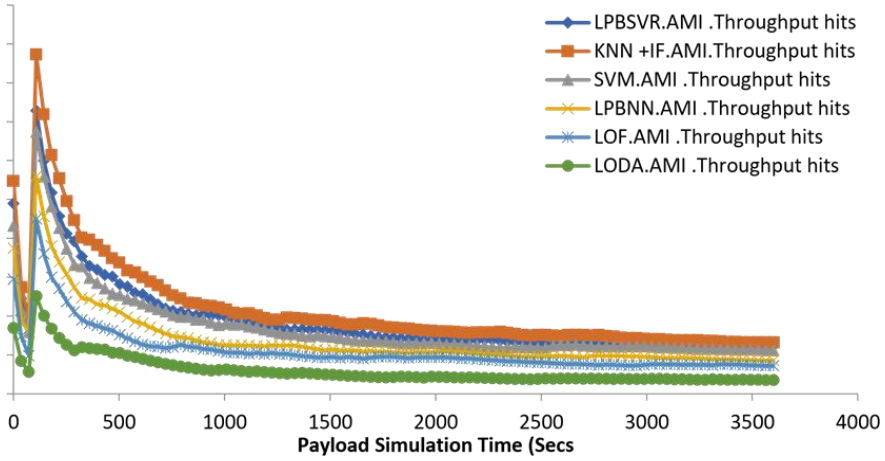

Figure 4: Throughput Efficiency

Figure 4 depicts the graph for the throughput efficiency compares the data processing capacities of the different models. LODA exhibited the highest throughput, confirming its ability to manage extensive datasets effectively and ensuring scalability in large-scale deployments.

3.3 Discussion

The results validate the effectiveness of unsupervised machine learning algorithms in detecting tampering within smart grid systems. The combination of KNN + IF, LOF, and LODA allowed for accurate anomaly detection while maintaining efficient processing speeds. Compared to traditional rule-based or supervised models, the unsupervised approach eliminated the dependency on labeled datasets and enhanced adaptability to new tampering methods. KNN + IF demonstrated the highest detection accuracy, making it the most reliable model for identifying tampering events, while LODA exhibited the best throughput, ensuring the framework's scalability. Additionally, the reduced latency observed across all models ensures rapid response to potential security threats, a critical requirement for real-time smart grid monitoring. Despite these advantages, some limitations remain. The computational overhead of training and deploying multiple unsupervised learning models may be a challenge, especially in resource-constrained environments. Additionally, while false positives were minimized, further refinement in threshold tuning and feature selection could improve overall reliability. LODA's lower detection accuracy compared to KNN + IF suggests that hybrid models combining unsupervised learning with reinforcement learning or deep learning techniques could further optimize tampering detection. Overall, this study presents a scalable and adaptable solution for improving smart grid security. By leveraging multiple unsupervised algorithms, the proposed framework enhances tampering detection, offering a robust alternative to conventional approaches while addressing existing limitations.

4. CONCLUSION AND RECOMMENDATION

4.1 Conclusion

This study presents a robust computational framework for detecting tampering in smart grids using unsupervised machine learning. By leveraging AMI and integrating KNN + IF, LOF, and LODA algorithms, the framework enhances detection accuracy, scalability, and efficiency. The experimental results demonstrate its ability to minimize latency, improve throughput, and detect anomalies effectively. Compared to traditional methods, this approach provides adaptability to evolving tampering techniques without relying on labeled datasets. Despite some computational overhead, the model proves to be a viable solution for real-time tampering detection in large-scale smart grid environments. Future work should focus on optimizing computational efficiency and exploring hybrid learning techniques to further enhance detection capabilities.

4.2 Recommendations

To further improve tampering detection in smart grids, it is recommended to integrate reinforcement learning techniques to enhance adaptability. Additionally, deploying the framework in real-world smart grid environments will provide further validation of its effectiveness and scalability. This study introduces a computational framework for smart grid tampering detection using unsupervised machine learning. By leveraging AMI and integrating advanced algorithms, the framework achieves enhanced accuracy and efficiency, addressing critical challenges in smart grid security. Future work will focus on reducing computational overhead and extending the framework to other smart grid applications.

References

Almalki, A. J. (2024). Unsupervised learning with hybrid models for detecting electricity theft in smart grids. IEEE Access, 12, 187027–187040.

Althobaiti, A., Jindal, A., Marnerides, A., & Roedig, U. (2021). Energy theft in smart grids: A survey on data-driven attack strategies and detection methods. IEEE Access, pp 1–1.

Bhattacharjee, S., Thakur, A., Silvestri, S., & Das, S. K. (2017). Statistical security incident forensics against data falsification in smart grid advanced metering infrastructure. Proceedings of the Seventh ACM on Conference on Data and Application Security and Privacy, 163-174.

Bhutta, M. S., Li, Y., Abubakar, M., Almasoudi, F. M., Alatawi, K. S., Altimania, M. R., & Al-Barashi, M. (2024). Optimizing solar power efficiency in smart grids using hybrid machine learning models for accurate energy generation prediction. Scientific Reports, 14(1), 1-12.

Chen, S., Wen, H., Wu, J., Lei, W., Hou, W., Liu, W., Xu, A., & Jiang, Y. (2019). Internet of Things Based Smart Grids Supported by Intelligent Edge Computing. IEEE Access, 7, 74089–74102.

Dehghani, M., Niknam, T., Ghasemigarpachi, M., Alhelou, H. H., Pourbehzadi, M., Javidi, G., & Sheybani, E. (2023). Public policies for cyber security of sustainable dominated renewable smart grids. IET Generation, Transmission & Distribution.

Esmael, A. A., Silva, H. H., Ji, T., & Silva Torres, R. (2021). Non-Technical Loss Detection in Power Grid Using Information Retrieval Approaches: A Comparative Study. IEEE Access, 9, 40635–40648.

Guerrero, J., Martín, A., Parejo, A., Marín, D. F. L., Molina, F. J., & León, C. (2023). A general-purpose distributed analytic platform based on edge computing and computational intelligence applied on smart grids. Sensors, 23(7), 3087.

Huang, Q., Tang, Z., Weng, X., He, M., Liu, F., Yang, M., & Jin, T. (2024). A Novel Electricity Theft Detection Strategy Based on Dual-Time Feature Fusion and Deep Learning Methods. Energies.

Husnoo, M. A., Anwar, A., Haque, M. E., & Mahmood, A. N. (2024). Decentralized Federated Anomaly Detection in Smart Grids: A P2P Gossip Approach. arXiv.Org, abs/2407.15879.

Iftikhar, H., Khan, N., Raza, M. A., Abbas, G., Khan, M., Aoudia, M., Touti, E., & Emara, A. (2024). Electricity theft detection in smart grid using machine learning. Frontiers in Energy Research.

Kaur, D., Aujla, G., Kumar, N., Zomaya, A. Y., Perera, C., & Ranjan, R. (2018). Tensor-Based Big Data Management Scheme for Dimensionality Reduction Problem in Smart Grid Systems: SDN Perspective. IEEE Transactions on Knowledge and Data Engineering, 30, 1985–1998.

Khalid, A., Mustafa, G., Rana, M. R. R., Alshahrani, S. M., & Alymani, M. (2024). RNN-BiLSTM-CRF based amalgamated deep learning model for electricity theft detection to secure smart grids. PeerJ Computer Science, 10.

Li, S., Han, Y., Yao, X., Song, Y., Wang, J., & Zhao, Q. (2019). Electricity Theft Detection in Power Grids with Deep Learning and Random Forests. Journal of Electrical and Computer Engineering, 2019, 4136874:1-4136874:12.

Liu, J., Wang, H., Bao, J., Sun, R., Du, X., & Guizani, M. (2024). RPMDA: Robust and privacy-enhanced multidimensional data aggregation scheme for fog-assisted smart grids. IEEE Internet of Things Journal, 11(11), 16021-16032.

Liu, Q., Hagenmeyer, V., & Keller, H. (2021). A Review of Rule Learning-Based Intrusion Detection Systems and Their Prospects in Smart Grids. IEEE Access, 9, 57542–57564.

Miller, M., Habbak, H., Badr, M. M., Baza, M., Mahmoud, M., & Fouda, M. M. (2024). Electricity Theft Detection Approach Using One-Class Classification for AMI. Consumer Communications and Networking Conference, 260–265.

Mohammadi, M., Shrestha, R., Sinaei, S., Salcines, A., Pampliega, D., Clemente, R., ... & Sanz, A. L. (2023). Anomaly detection using LSTM-autoencoder in smart grid: A federated learning approach. 7th IEEE International Conference on Cloud and Big Data Computing pp. 176–183.

Pamir, S., Javaid, N., Javed, M. U., AH, M., Almasoud, A., & Imran, M. (2023). Electricity theft detection for energy optimization using deep learning models. Energy Science & Engineering, 11, 3575–3596.

Radoglou-Grammatikis, P. I., Sarigiannidis, P. G., Efstathopoulos, G., & Panaousis, E. (2020). ARIES: A Novel Multivariate Intrusion Detection System for Smart Grid. Sensors, 20(17), 4872.

Sahani, N., Zhu, R., Cho, J., & Liu, C. C. (2023). Machine learning-based intrusion detection for smart grid computing: A survey. ACM Transactions on Cyber-Physical Systems, 7(1), 1–31.

Samanthula, B., & Patel, H. (2023). Privacy-Preserving and Outsourced Computation Framework for Power Usage Control in Smart Grids. In Proceedings of the 2023 ACM Workshop on Secure and Trustworthy Cyber-Physical Systems (pp. 1–8).

Syed, D., Abu-Rub, H., Refaat, S., & Xie, L. (2020). Detection of Energy Theft in Smart Grids using Electricity Consumption Patterns. 2020 IEEE International Conference on Big Data (Big Data), 4059–4064.

Teixeira, B., Santos, G., Pinto, T., Vale, Z., & Corchado, J. (2020). Application ontology for multi-agent and web-services' co-simulation in power and energy systems. IEEE Access, 8, 81129-81141.

Wahed, M. A., Alqaraleh, M., Alzboon, M. S., & Al-Batah, M. S. (2025). Evaluating AI and Machine Learning Models in Breast Cancer Detection: A Review of Convolutional Neural Networks (CNN) and Global Research Trends. LatIA.

Yan, Z., & Wen, H. (2022). Performance Analysis of Electricity Theft Detection for the Smart Grid: An Overview. IEEE Transactions on Instrumentation and Measurement, 71, 1–28.

Zhang, Y., Fang, X., Hordiichuk-Bublivska, O., Beshley, H., & Beshley, M. (2023). Modified Masking-Based Federated Singular Value Decomposition Method for Fast Anomaly Detection in Smart Grid Systems. Energies, 16(16), 6228.

Zhang, Y., Fang, X., Hordiichuk-Bublivska, O., Beshley, H., & Beshley, M. (2023). Modified Masking-Based Federated Singular Value Decomposition Method for Fast Anomaly Detection in Smart Grid Systems. Energies, 16(16), 6228.

Zhu, L., Wen, W., Li, J., Zhang, C., Zhou, B., & Shuai, Z. (2024). Deep Active Learning-Enabled Cost-Effective Electricity Theft Detection in Smart Grids. IEEE Transactions on Industrial Informatics, 20, 256–268.