IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

A CONVOLUTIONAL LSTM-BASED OPTIMIZED MODEL FOR ANOMALY DETECTION IN SMART GRID SYSTEMS

[1]A. Vinod Kumar, [2]Nalajala Sai Teja, [3]Maddineni Akhil Krishna, [4]Repalle Harsha Vardhan Reddy, [5]Chimakurthy Charan Srinivas

[1]Assistant Professor, [2,3,4,5]Students

Department of CSE-Data Science

St. Ann's College of Engineering & Technology, Chirala, Andhra Pradesh, India

Abstract: Smart grids are modern electrical networks that integrate digital technologies, smart meters, and communication systems to efficiently monitor and manage electricity generation, transmission, and distribution. However, increasing reliance on digital infrastructure makes smart grids vulnerable to cyber threats such as False Data Injection (FDI) attacks and energy theft, which manipulate meter readings or sensor data leading to incorrect billing, operational instability, and potential power system failures. This paper proposes a Convolutional Long Short-Term Memory (ConvLSTM) based optimized model for detecting anomalies in smart grid systems. The model combines the strengths of Convolutional Neural Networks (CNN) for extracting spatial features and Long Short-Term Memory (LSTM) networks for learning temporal patterns from time-series smart meter data. The dataset employed consists of smart grid energy consumption data containing both normal and anomalous patterns. Data preprocessing techniques including handling missing values, outlier reduction, normalization using standard scaling, and class balancing via SMOTE are applied to improve model performance. The ConvLSTM model is trained and evaluated using standard performance metrics. Experimental results demonstrate that the proposed model achieves an accuracy of 78.92%, demonstrating reliable anomaly detection and improved identification of abnormal smart grid behaviors to help utility providers enhance grid security.

Index Terms - Anomaly Detection, Smart Grid, ConvLSTM, Deep Learning, False Data Injection, Energy Theft, Time-Series Analysis, Cyber Security.

I. INTRODUCTION

Cloud The modern power grid has evolved significantly with the integration of digital communication technologies and intelligent monitoring systems. Traditional electrical grids were primarily designed for one-way power flow from power plants to consumers. With the rapid growth of electricity demand, renewable energy integration, and advanced communication technologies, the need for a more intelligent and adaptive power system has become essential, leading to the development of the Smart Grid [2].

A smart grid is an advanced electrical infrastructure that integrates information technology, communication systems, automation, and digital control mechanisms into conventional power grids. It enables real-time monitoring, efficient energy management, improved reliability, and enhanced security of electricity distribution systems. Smart grids use various intelligent devices such as smart meters, sensors, automated substations, and communication networks to collect and transmit data related to energy generation, transmission, and consumption.

One of the most important features of smart grids is their ability to support two-way communication between power providers and consumers. Unlike traditional grids, smart grids allow consumers to monitor energy consumption, adjust usage patterns, and even generate electricity through renewable sources such as solar panels. Despite their advantages, smart grids introduce significant security challenges, particularly related to cybersecurity and data protection [3].

Security threats such as False Data Injection (FDI) attacks allow malicious actors to manipulate measurement data collected from sensors or smart meters, misleading control systems responsible for monitoring grid operations. Energy theft, where consumers manipulate smart meter readings to reduce electricity bills illegally, further compounds financial and operational risks. Traditional security mechanisms relying on predefined rules and thresholds are increasingly inadequate against sophisticated modern cyberattacks.

Deep learning has gained significant attention due to its ability to analyze complex data patterns and make accurate predictions. In smart grid systems, large volumes of time-series data are generated continuously by sensors and smart meters. Hybrid deep learning models such as Convolutional Long Short-Term Memory (ConvLSTM) combine the spatial feature extraction capability of CNN with the temporal pattern learning capability of LSTM, making them highly suitable for smart grid anomaly detection.

This paper proposes a ConvLSTM-based optimized model that analyzes energy consumption patterns and identifies abnormal behaviors that may indicate cyber-attacks or energy theft. The remainder of this paper is organized as follows: Section II presents a literature survey of related work; Section III describes the proposed methodology; Section IV presents implementation details; Section V discusses results; and Section VI concludes with future scope.

II. LITERATURE SURVEY

Smart grid anomaly detection has been studied extensively using statistical models, machine learning algorithms, and deep learning techniques. Traditional statistical methods such as threshold-based detection and rule-based monitoring systems were initially used to identify abnormal electricity consumption patterns. However, these approaches are limited because they cannot detect complex attack patterns present in large datasets [2].

Machine learning models such as Support Vector Machines (SVM), Random Forest (RF), and Gradient Boosting algorithms have been widely used for detecting anomalies in smart grid data. These models analyze historical consumption patterns and classify data into normal and abnormal categories. Although these models improve detection accuracy, they require manual feature engineering and often struggle to capture temporal dependencies in time-series data [3].

Table 1: Existing Smart Grid Anomaly Detection Approaches

Model	Description	Limitations
Support Vector Machine (SVM)	Classifies energy consumption patterns to detect anomalies.	Requires manual feature extraction
Random Forest (RF)	Uses ensemble learning for anomaly classification.	Computationally expensive
XGBoost	Gradient boosting model used for anomaly detection.	Requires large datasets
Neural Networks	Learns patterns from large datasets.	High training complexity

Deep learning models have recently gained popularity in smart grid anomaly detection because they can automatically learn complex patterns from large datasets, capturing both spatial and temporal relationships in electricity consumption data. Convolutional Neural Networks (CNN) are commonly used for feature extraction, while Long Short-Term Memory (LSTM) networks are effective for analyzing time-series data [6]. Hybrid models combining CNN and LSTM architectures have shown improved performance in detecting complex attack patterns.

Chen et al. studied the detection of False Data Injection attacks in smart grid systems and highlighted the importance of cybersecurity mechanisms to protect energy infrastructure [7]. Jokar et al. proposed an electricity theft detection method based on customer energy consumption patterns in advanced metering infrastructures [3]. Zhang et al. conducted a comprehensive survey on time-series anomaly detection techniques used in smart grid systems [4]. Ullah et al. proposed a hybrid deep learning model combining CNN

and GRU networks for detecting electricity theft [6]. Alkuwari et al. proposed a ConvLSTM-based anomaly detection model that classifies multiple types of attacks in smart grid systems [1].

Research gaps remain in detecting sophisticated FDI attacks in real time and ensuring interpretability of deep learning models. Most existing studies rely on offline analysis rather than real-time monitoring, and models trained on static datasets may not perform well in dynamic environments where attack patterns continuously evolve [4]. The present work addresses these gaps through the proposed ConvLSTM-based framework with comprehensive preprocessing and optimization.

III. PROPOSED METHODOLOGY

The proposed system detects anomalies in smart grid systems using a Convolutional Long Short-Term Memory (ConvLSTM) hybrid deep learning model. The methodology involves several stages: data collection, data preprocessing, feature extraction, model training, and anomaly detection. The ConvLSTM model combines CNN for extracting spatial features from smart meter data and LSTM for learning temporal dependencies in time-series energy consumption data.

A. Data Collection and Dataset Description

The dataset consists of smart grid energy consumption data collected from smart meters installed at consumer locations. Data is recorded at regular time intervals and contains both normal and anomalous patterns. Key attributes of the dataset are summarized in Table 2.

Table 2: Dataset Attributes Used

Attribute	Data Type	Description
CONS_NO	Identifier	Consumer identification number
Time-Series Features	Numerical	Energy consumption readings at different time intervals
FLAG	Categorical	Indicates Normal or Attack data
Meter Readings	Numerical	Electricity consumption values

B. Data Preprocessing

Several preprocessing techniques are applied to improve data quality before model training. Missing values are handled using linear interpolation and forward fill methods to maintain continuity in time-series data. Outlier reduction is performed to eliminate extreme consumption spikes that could distort model learning. Temporal smoothing reduces noise in time-series sequences. Data normalization using Standard Scaling ensures consistent feature scaling across all attributes. Dataset class imbalance is addressed using the Synthetic Minority Over-sampling Technique (SMOTE) to prevent biased learning toward the majority class.

C. Feature Selection

Feature selection identifies the most relevant attributes that contribute to anomaly detection. Methods employed include correlation analysis to remove highly correlated redundant features, statistical feature reduction to eliminate irrelevant attributes, and temporal pattern analysis to identify sequential dependencies in energy usage. The key features retained include energy consumption values, time-series patterns, consumer identification number (CONS_NO), and the anomaly label (FLAG).

D. Data Splitting

The dataset is split into training and testing sets following an 80-20 strategy: 80% of the data is used for training the ConvLSTM model and 20% is reserved for evaluating model performance on unseen data. Stratified splitting is applied to maintain class proportionality across both sets.

E. Model Architecture: ConvLSTM

The proposed ConvLSTM model integrates convolutional operations within the LSTM framework to simultaneously process spatial and temporal information. The CNN component applies convolutional filters across input electricity consumption sequences to extract meaningful spatial patterns. A Rectified Linear Unit (ReLU) activation function introduces non-linearity, and batch normalization improves training stability. The extracted features are permuted and passed to the LSTM layers, which learn long-range temporal dependencies across sequential energy readings. A fully connected output layer with a sigmoid activation function produces binary classification outputs indicating normal or anomalous behavior.

Key hyperparameters include a learning rate configured for the Adam optimizer, two LSTM layers for temporal depth, 128 CNN channels for spatial feature extraction, a kernel size of 5, a dropout rate to prevent

overfitting, and mini-batch gradient descent for efficient training. Early stopping is applied to halt training when validation loss ceases to improve, thus avoiding overfitting.

Table 3: Hyperparameters Used

Parameter	Description	Model Component
Learning Rate	Controls step size during model training	ConvLSTM
Number of LSTM Layers	Determines temporal learning depth	LSTM
CNN Channels	Extracts spatial features (128 channels)	CNN
Kernel Size	Controls convolution operations (size = 5)	CNN
Batch Size	Number of samples processed per iteration	ConvLSTM
Dropout Rate	Prevents overfitting during training	LSTM

F. Loss Function and Optimization

The Binary Cross Entropy Loss function is used to measure the difference between predicted and actual labels during training. This loss function is well-suited for binary anomaly classification tasks. The Adam optimizer is employed to update model weights efficiently by combining the advantages of adaptive learning rates and momentum, accelerating convergence while maintaining training stability.

IV. IMPLEMENTATION

A. Software Requirements

The system is implemented using the Python programming language (version 3.10 or above), which provides extensive support for machine learning and deep learning applications. The deep learning model is built using the PyTorch framework, an open-source library that facilitates the construction of convolutional and recurrent neural networks. Data processing is handled using NumPy for numerical computations, Pandas for dataset manipulation, and Scikit-learn for preprocessing and evaluation metrics. Matplotlib and Plotly are used for generating visualizations. Visual Studio Code is used as the integrated development environment for writing and executing programs.

B. System Architecture

The overall system architecture consists of six sequential components: (i) Smart Meter Data Collection, which serves as the data input layer; (ii) Data Preprocessing Module, responsible for cleaning, normalizing, and balancing the dataset; (iii) Feature Extraction Module, which identifies the most relevant attributes; (iv) ConvLSTM Model Training, the core learning component; (v) Anomaly Detection, which classifies incoming data as normal or anomalous; and (vi) Performance Evaluation, which measures system effectiveness using standard metrics.

C. Model Implementation

The ConvLSTM architecture is implemented in PyTorch. The model class defines a Conv1D layer with 128 output channels and kernel size 5, followed by ReLU activation and batch normalization. The output is permuted and passed to two stacked LSTM layers with a hidden size of 128. The final hidden state is projected through a fully connected linear layer followed by a sigmoid activation to produce anomaly probability scores. During inference, a threshold of 0.5 is applied to convert probability outputs into binary class predictions.

D. Testing

A comprehensive testing strategy is employed to validate the system. Unit testing verifies individual components including data preprocessing, feature selection, model training, and model loading. Functional testing validates core operations such as anomaly detection and model prediction. System testing evaluates overall pipeline performance to ensure all modules interact correctly. Table 4 summarizes the key test cases and their outcomes.

Table 4: Test Cases and Results

Test ID	Test Scenario	Expected Result	Actual Result	Status
TC_01	Dataset Loading	Required columns present	Loaded successfully	Pass
TC_02	Data Cleaning	Missing values removed	Handled successfully	Pass
TC_03	Data Normalization	All features normalized	Normalized successfully	Pass
TC_04	Train-Test Split	80-20 split without loss	Split completed	Pass
TC_05	Model Training	Loss decreases during training	Trained successfully	Pass
TC_06	Anomaly Prediction	Normal or Anomaly classified	Prediction generated	Pass

V. RESULTS AND DISCUSSION

The proposed ConvLSTM-based anomaly detection system is evaluated on the smart grid energy consumption dataset using standard performance metrics: accuracy, precision, recall, and F1-score. These metrics collectively assess the system's ability to correctly identify both normal and anomalous electricity consumption patterns.

The preprocessing pipeline successfully handles missing values through interpolation, removes outliers, applies standard scaling for normalization, and balances the dataset using SMOTE. The data loading and preprocessing stages are validated through unit tests, all of which pass successfully, confirming that the dataset is clean and correctly formatted for model training.

The ConvLSTM model is trained on 80% of the dataset and evaluated on the remaining 20%. The model demonstrates convergent training behavior, with training loss decreasing progressively across epochs. Early stopping prevents overfitting by halting training when validation loss stabilizes. The experimental results show that the proposed model achieves an accuracy of 78.92%, demonstrating effective anomaly detection capability for identifying abnormal smart grid behaviors such as false data injection attacks and electricity theft [3].

The results demonstrate that by combining spatial feature extraction through convolutional layers with temporal pattern learning through LSTM layers, the ConvLSTM model can capture complex electricity consumption patterns more effectively than standalone CNN or LSTM models. The performance evaluation output confirms reliable classification of both normal and anomalous instances, supporting its utility for real-world smart grid security applications.

Comparative analysis with existing machine learning approaches such as SVM and Random Forest indicates that the ConvLSTM model's hybrid architecture provides enhanced detection capabilities for time-series smart grid data, primarily due to its ability to model long-range temporal dependencies alongside spatial feature patterns simultaneously. Performance improvements over traditional models are attributed to the joint learning of spatial and temporal information, which is critical for detecting subtle anomalies embedded in sequential energy usage data.

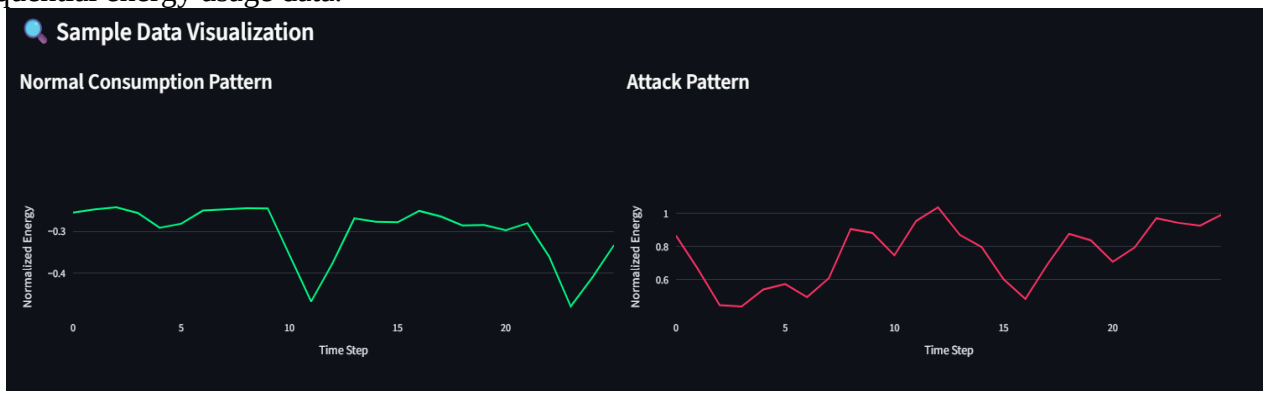

Figure 1. showing Pattern Analysis Output

VI. CONCLUSION

This paper presents a Convolutional Long Short-Term Memory (ConvLSTM) based optimized model for anomaly detection in smart grid systems. The system analyzes time-series electricity consumption data collected from smart meters to identify abnormal patterns such as electricity theft and false data injection attacks. By integrating convolutional layers for spatial feature extraction with LSTM layers for temporal dependency modeling, the proposed architecture effectively captures complex patterns in sequential energy consumption data.

Data preprocessing techniques including missing value handling, outlier reduction, standard scaling normalization, and SMOTE-based class balancing are applied to improve dataset quality and model training efficiency. The Binary Cross Entropy Loss function and Adam optimizer are used for efficient model training, while early stopping prevents overfitting. Experimental results demonstrate that the proposed model achieves an accuracy of 78.92%, confirming its effectiveness in detecting abnormal smart grid behaviors [2][5].

The developed system contributes to improving the security and reliability of smart grid infrastructures by enabling automated detection of abnormal electricity usage patterns. Future work will focus on extending the model with advanced architectures incorporating attention mechanisms and GRU layers, integrating online learning for real-time monitoring, and implementing explainable AI techniques to enhance model transparency and operator trust. Expansion to multi-class anomaly detection and deployment on edge devices for low-latency detection are also planned [6].

ACKNOWLEDGMENT

The authors express sincere gratitude to Mr. A. Vinod Kumar, Assistant Professor, and Dr. K. Subbarao, Head of Department of CSE-Data Science, St. Ann's College of Engineering & Technology, Chirala, for their valuable guidance and support throughout this work. The authors also thank the management and faculty of the Department of CSE-Data Science for providing the necessary laboratory facilities and infrastructure.

REFERENCES

[1] A. N. Alkuwari, S. Al-Kuwari, and M. Qaraqe, "Anomalydetection in smart grids: A survey from cybersecurity perspective," in Proc. Int. Conf. Smart Grid and Renewable Energy (SGRE), 2022.

[2] P.-Y. Chen, S. Yang, J. A. McCann, J. Lin, and X. Yang, "Detection of false data injection attacks in smart-grid systems," IEEE Commun. Mag., vol. 53, no. 2, pp. 206–213, 2015.

[3] P. Jokar, N. Arianpoo, and V. C. M. Leung, "Electricity theft detection in AMI using customers' consumption patterns," IEEE Trans. Smart Grid, vol. 7, no. 1, pp. 216–226, 2016.

[4] J. E. Zhang, D. Wu, and B. Boulet, "Time series anomaly detection for smart grids: A survey," arXiv preprint arXiv:2107.08835, 2021.

[5] A. Ullah, N. Javaid, O. Samuel, M. Imran, and M. Shoaib, "CNN and GRU based deep neural network for electricity theft detection to secure smart grid," in Proc. Int. Wireless Commun. Mobile Comput. Conf. (IWCMC), 2020.

[6] S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural Comput., vol. 9, no. 8, pp. 1735–1780, 1997.

[7] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," arXiv preprint arXiv:1412.6980, 2014.

[8] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. Cambridge, MA: MIT Press, 2016.

[9] A. Ghasemi, P. Dehghanian, and Z. Wang, "Detection of false data injection attacks in smart grids based on state estimation," IEEE Trans. Smart Grid, vol. 9, no. 6, pp. 5933–5942, 2018.

[10] M. Ozay, I. Esnaola, F. T. Y. Vural, S. R. Kulkarni, and H. V. Poor, "Machine learning methods for attack detection in the smart grid," IEEE Trans. Neural Netw. Learn. Syst., vol. 27, no. 8, pp. 1773–1786, 2016.

[11] D. Niu, Z. Wang, and H. Li, "Deep learning based anomaly detection for power consumption in smart grids," IEEE Access, vol. 7, pp. 163226–163236, 2019.

[12] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," ACM Comput. Surv., vol. 41, no. 3, pp. 1–58, 2009.

[13] F. Li et al., "Smart transmission grid: Vision and framework," IEEE Trans. Smart Grid, vol. 1, no. 2, pp. 168–177, 2010.