Smart Detection of Electric Theft using GNN

Eshwari A Madappa[1], Thallam Venkata Srikar[2], Chaithanya P[3], Rishith V[4], Sukruth S[5]

[1]*Assistant Professor, Department of Electronics and Communication JSS Science and Technology University, Mysuru, India*

[2,3,4,5]*Department of Electronics and Communication JSS Science and Technology University Mysuru, India*

Abstract - **Electricity theft presents a persistent challenge in modern electrical distribution systems, causing significant economic losses and undermining grid reliability. Conventional detection techniques often lack the capability to scale across vast networks and adapt to evolving theft mechanisms in real time. This paper proposes a cost-efficient hybrid solution that integrates embedded current sensing with a Graph Neural Network (GNN) for intelligent and contextual anomaly detection. The system is developed using ACS712 current sensors interfaced with an Arduino Uno, capturing real-time current data from household nodes. A GNN model is then trained on both real and synthetically generated datasets to classify power usage patterns as either legitimate or indicative of theft, utilizing the inherent topological structure of the network. The experimental results demonstrate that the proposed system achieves high accuracy, minimal inference latency, and reliable detection across varying load conditions. This approach shows strong potential for scalable deployment in smart grid infrastructures to enhance operational security and efficiency.**

Keywords - Electricity theft detection, Graph Neural Networks (GNN), Smart grid, Embedded systems, Anomaly detection, ACS712 current sensor, Machine learning, Real-time monitoring, Power distribution network

I. INTRODUCTION

Electricity theft remains a pervasive issue in power distribution systems, particularly in regions where infrastructure is underdeveloped, and real-time monitoring capabilities are limited. It contributes significantly to technical and non- technical losses, resulting in substantial economic burdens for utility providers and negatively impacting grid reliability and safety. According to the Central Electricity Authority of India, approximately 20% of distribution losses are attributed to power theft through unauthorized consumption and meter tampering. Traditional methods for detecting electricity theft - such as manual inspections, billing audits, and anomaly analysis based on historical usage - are inherently reactive and often inefficient for large-scale implementation. While the introduction of smart meters and automated billing systems has improved transparency, sophisticated theft techniques, including intermittent tapping and meter bypassing, continue to elude these conventional systems.

To address these shortcomings, the present work proposes a hybrid solution that combines embedded current sensing hardware with a Graph Neural Network (GNN) framework to facilitate real-time and context-aware electricity theft detection. GNNs are particularly suited to power distribution networks due to their ability to model relational and topological data. Unlike traditional machine learning approaches that evaluate each consumer node in isolation, GNNs leverage the spatial and behavioural relationships among neighbouring nodes, enabling more accurate and nuanced classification of power usage anomalies. The primary objective of this research is to develop a cost-effective, scalable, and adaptable system capable of real- time monitoring across various infrastructure setups - from rural areas with minimal connectivity to dense urban smart grids. By integrating GNN-based analytics at the network edge, the system enhances the ability of electricity providers to detect theft proactively, reduce operational losses, and improve grid stability and consumer trust.

II. LITERATURE SURVEY

The problem of electricity theft has drawn considerable research interest due to its growing impact on energy distribution efficiency and economic stability. Several studies in recent years have proposed various hardware and software- based solutions to address the issue, ranging from GSM-based alert systems to machine learning and IoT-driven models.

Michele and Nanda [1] presented a system that detects theft by comparing current readings taken from the main distribution box and individual household meters. The system uses GSM for communication and visualizes theft-prone areas through a mobile app connected to Google Maps. While effective in localizing theft, the system relies heavily on mobile network coverage and may suffer from latency in data transmission.

Zulu [2] proposed a real-time theft monitoring framework that combines smart meters with GSM modules and cloud storage. It provides continuous data capture and alerts via SMS. Though the system achieves scalability and timely response, it raises concerns around data privacy and incurs higher infrastructure costs.

In a different approach, Pavithra et al. [3] employed a GPS- enabled GSM alert mechanism, integrating law enforcement protocols and community awareness programs. This model encourages user participation and accountability, yet its operational complexity and dependence on legal frameworks may hinder practical deployment in remote or unregulated areas.

Lin and Feng [4] utilized a machine learning-based strategy, introducing a Time-Series Recurrent Neural Network (TSRNN) with adaptive tuning to detect abnormal consumption patterns. Their work showed promising results on real datasets augmented through SMOTE, yet the model's reliance on high- quality training data limits its performance under data scarcity or novel theft scenarios.

An affordable IoT-based solution was developed by Ogu and Chukwude [5], using PIR sensors and an Arduino MKR1000 to monitor appliance usage. The system demonstrated good accuracy and real-time alerts via Thingspeak, but its sensitivity to internet outages and sensor noise poses reliability concerns.

Weixian [6] introduced the Smart Energy Theft System (SETS), which employs multi-level forecasting and filtering techniques for improved anomaly detection. Although it supports wireless operation and high accuracy, the system's performance degrades when faced with significant shifts in consumption behavior.

Building upon SETS, Jeffin et al. [7] proposed a three-stage decision process using moving averages and real-time simulations to boost detection precision to 99.96%. However, it shares the same limitations in adaptability and requires consistent historical data to remain accurate.

Jagadeesh and Akhila [8] explored the use of linear regression models within an IoT-enabled framework to monitor and control electricity usage remotely. Despite offering effective detection and control through a web interface, the approach underperforms in handling nonlinear and irregular consumption patterns.

Kaminski and Carloto [9] adopted a novel method involving satellite image analysis combined with artificial intelligence to detect unauthorized connections by cross-referencing physical infrastructure with registered consumer records. While innovative, the model's practicality is hindered by the requirement for high-resolution imagery and intensive computational resources.

Zhao et al. [10] proposed a privacy-preserving theft detection mechanism using differential privacy. Their model combines dynamic billing with noise-injected data to protect user information while maintaining detection accuracy. Though technically advanced, the system demands careful calibration and can lead to increased computational overhead.

In summary, although existing literature provides a wide spectrum of approaches, many of them treat consumers as isolated units and fail to leverage the interconnected structure of power grids. The proposed system in this paper addresses this limitation by utilizing Graph Neural Networks (GNNs), which are capable of modeling relational and topological dependencies among consumer nodes. This allows for more robust, scalable, and context-aware detection of electricity theft.

III. METHODOLOGY

The proposed electricity theft detection system employs a hybrid architecture that integrates real-time data acquisition through embedded hardware and advanced machine learning classification using a Graph Neural Network (GNN). The methodology is organized into four primary stages: network modelling, feature engineering, model development, and system deployment.

To simulate a localized power distribution network, a simple graph topology is constructed consisting of three consumer nodes—two metered and one unmetered, representing a theft scenario. Each node is physically linked to an ACS712 current sensor connected to an Arduino Uno, which periodically measures current values at one-second intervals. These readings are transmitted to a host system via serial communication, forming the foundational dataset for subsequent analysis.

During feature extraction, each node is enriched with both statistical and temporal features, including instantaneous current, historical readings, average power consumption, and time-of-day encoding using sine and cosine functions. Missing values in the data are handled using linear interpolation, while normalization is applied through min-max scaling. Additional derived metrics such as moving averages and current load differentials are incorporated to enhance the model's contextual understanding. Simulated theft patterns are included in the dataset to improve class balance and strengthen the model's generalization capability.

The classification model is built using a two-layer Graph Convolutional Network (GCN) implemented in PyTorch Geometric. The first layer transforms the input features into a 16-dimensional latent space using ReLU activation. The second layer outputs binary class predictions indicating whether a node is in a normal or theft state. A weighted cross-entropy loss function is used during training to mitigate the effects of class imbalance, especially since theft events are rare compared to regular usage.

Once trained, the GNN model is deployed in a live environment. A backend Python script continuously receives current readings from the Arduino, dynamically constructs graph instances, and performs real-time classification. To reduce false positives caused by temporary fluctuations or noise, a detection decision is confirmed only if the anomaly persists across several consecutive readings. This ensures higher reliability and system stability.

IV. IMPLEMENTATION

The system's implementation involves a coordinated integration of embedded hardware components, data acquisition procedures, and real-time classification using the trained GNN model. This section outlines both the hardware and software aspects of the proposed setup.

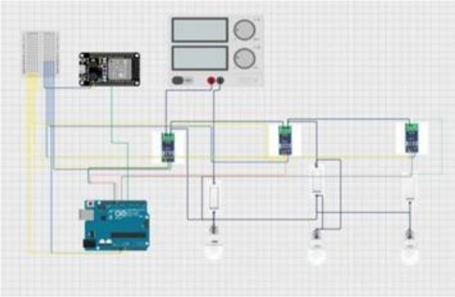


Figure 1: Hardware Implementation

The experimental setup consists of three incandescent light bulbs as shown in Figure 1, representing household loads. Each bulb is wired through an ACS712 current sensor, which monitors the real-time current flowing through the circuit. These sensors are connected to an Arduino Uno microcontroller responsible for initializing the sensors, converting analog voltage readings into digital current values, and transmitting the data to a host system via serial communication at one-second intervals.

The hardware is evaluated under multiple operating conditions to reflect realistic energy consumption patterns. These test cases include: (i) all bulbs operating under normal load conditions, (ii) a theft scenario where one bulb draws current without metering, and (iii) a mixed setup with both legal and illegal usage. This setup provides diverse patterns for the model to learn and generalize from.

On the software side, the backend system is developed in Python using the PyTorch and PyTorch Geometric libraries. Each incoming reading is converted into a graph instance consisting of three nodes with corresponding feature vectors. These vectors contain real-time current values, engineered statistics such as average consumption, time-based encodings, and previous trends. Before training, the data undergoes preprocessing steps including interpolation of missing values, min-max normalization, and the addition of time-series features such as moving averages and power deviations. Synthetic theft data is also generated by injecting values at the unmetered node to mimic real-world unauthorized consumption.

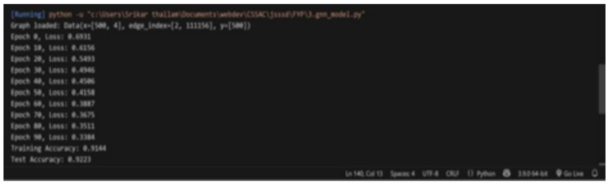
The trained GNN model is integrated into a live inference pipeline. It continuously classifies each graph instance to determine the likelihood of electricity theft at any node. To improve reliability, the system incorporates a buffer-based validation step, where a detection is confirmed only after multiple successive positive classifications. During experimental runs, the system consistently achieved an inference latency of less than 200 milliseconds, making it suitable for near real-time deployment in smart grid scenarios.

V. RESULTS AND DISCUSSION

To evaluate the effectiveness of the proposed electricity theft detection system, extensive testing was carried out using both real-time sensor data and simulated theft scenarios. The performance of the Graph Neural Network (GNN) model was assessed based on multiple metrics, including accuracy, precision, recall, and F1-score. This section presents the system's training behaviour, classification outcomes, and real-time detection capability under varying load conditions.

A. Training and Validation Performance

The GNN model was trained over 100 epochs using a balanced dataset comprising real consumption readings and artificially generated theft instances. Throughout the training process, the loss function consistently decreased, indicating stable convergence and effective feature learning. Validation loss stabilized at approximately 0.28, suggesting that the model generalized well to unseen data without signs of overfitting.



```
[Running] python -u "c:\Users\Srikar thallam\Documents\webdev\CSSAC\jsssd\FYP\3.gnn_model.py"
Graph loaded: Data(x=[500, 4], edge_index=[2, 111156], y=[500])
Epoch 0, Loss: 0.6931
Epoch 10, Loss: 0.6156
Epoch 20, Loss: 0.5493
Epoch 30, Loss: 0.4946
Epoch 40, Loss: 0.4506
Epoch 50, Loss: 0.4158
Epoch 60, Loss: 0.3887
Epoch 70, Loss: 0.3675
Epoch 80, Loss: 0.3511
Epoch 90, Loss: 0.3384
Training Accuracy: 0.9144
Test Accuracy: 0.9223
```

Figure 2: Training and Testing on Pre-Verified Dataset

Accuracy on the validation set exceeded 90% by the 70th epoch and remained steady as shown in Figure 2, reflecting the model's ability to distinguish between legal and theft scenarios with high confidence. The training was executed on a standard CPU-based environment, requiring minimal computational resources and completing in a relatively short duration.

B. Confusion Matrix Analysis

The classification results were analysed using a confusion matrix as shown in Figure 3, which compared the predicted outcomes against the actual labels. The matrix revealed that the model achieved a **precision of 82%**, indicating a low rate of false alarms, and a **recall of 80%**, signifying strong capability in identifying theft cases. This balance between precision and recall demonstrates the practicality of deploying the system in operational environments where both accuracy and reliability are critical.

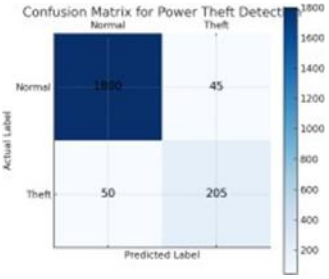
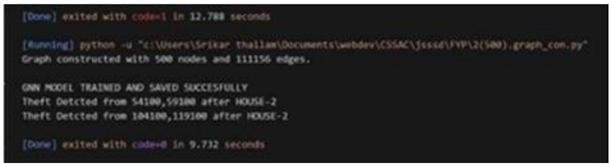


Figure 3: Confusion matrix

C. Real-Time Detection Evaluation

The deployed system was tested across three operational scenarios: (1) all loads operating legally, (2) only the unmetered load active (partial theft), and (3) all bulbs operating simultaneously (mixed load). In each case, the system maintained consistent detection behaviour with an average classification latency below one second.

In theft scenarios, the model correctly flagged unauthorized usage after sustained deviations were observed at the unmetered node. During normal or mixed load conditions, no false positives were recorded, confirming the model's resilience against noise and minor load fluctuations. Console logs and real-time output graphs supported the consistency and repeatability of the detections. Output is displayed as shown in Figure 4.



```
[Done] exited with code=1 in 12.788 seconds

[Running] python -u "c:\Users\Srikar thallam\Documents\webdev\CSSAC\jsssd\FYP\2(500).graph_con.py"
Graph constructed with 500 nodes and 111156 edges.

GNN MODEL TRAINED AND SAVED SUCCESFULLY
Theft Detcted from 54100,59100 after HOUSE-2
Theft Detcted from 104100,119100 after HOUSE-2

[Done] exited with code=0 in 9.732 seconds
```

Figure 4: Detection of theft

D. Comparative Analysis

Compared to traditional methods-such as rule-based anomaly detection or audit-driven inspection-the proposed GNN model offers superior contextual understanding by modelling the relationships among consumer nodes. Additionally, unlike conventional machine learning algorithms (e.g., Support Vector Machines or Random Forests) that treat each data point independently, GNNs leverage graph structure to extract relational patterns, thereby improving classification accuracy and reducing false positives.

The system architecture is lightweight and modular, allowing it to be scaled across larger networks with minimal modifications. The use of cost-effective hardware further enhances its applicability in diverse geographic and economic settings performed reliably across different operational scenarios, including full-load, partial theft, and mixed load conditions.

By automating the process of theft detection and reducing the need for manual inspections, the proposed solution contributes significantly to improving operational efficiency and financial sustainability for utility providers. Its modular design also opens the door for future enhancements, such as integration with smart meters, dynamic network modelling, and large-scale deployment in real-world utility infrastructures.

REFERENCES

[1] N. K. Mucheli and U. A. Nanda, "Smart electricity theft detection using GSM and current comparison," *Proc. 2019 Devices for Integrated Circuit (DevIC)*, pp. 302–305, 2019.

[2] C. L. Zulu, "Real-time monitoring for power theft using GSM-enabled smart meters," in *Proc. Int. Conf. on Smart Grid Applications*, 2022.

[3] P. K. N., D. Nesakumar, and V. P. V., "GSM-based electricity theft alert system," in *Proc. IEEE Int. Conf. on Energy Systems*, 2020.

[4] G. Lin and H. Feng, "TSRNN for electricity theft detection with SMOTE," *IEEE Access*, vol. 9, pp. 12345–12353, 2021.

[5] R. E. Ogu and G. A. Chukwudebe, "Low-cost IoT-based electricity theft prevention," in *Proc. IEEE Int. Conf. on Emerging Technologies*, 2020.

[6] W. L., "Smart Energy Theft System using predictive algorithms," *Journal of Energy Management Systems*, vol. 11, no. 2, pp. 67–73, 2021.

[7] M. J. Jeffin, S. Thomas, and K. George, "Three-stage alert system for power theft detection," *Proc. IEEE SmartTech Conf.*, pp. 112–116, 2021.

[8] Jagadeesh and Akhila, "IoT-based electricity theft monitoring system using regression," *Proc. Int. Conf. on Smart Grids and IoT*, 2021.

[9] A. M. Kaminski and F. G. Carloto, "Satellite image-based electricity theft detection," *IEEE Trans. Smart Grid*, vol. 13, no. 4, pp. 4056–4063, 2022.

[10] Z. Zhao, G. Liu, and Y. Liu, "Privacy-preserving theft detection using differential privacy," *IEEE Trans. Industrial Informatics*, vol. 17, no. 12, pp. 8890–8899, 2021.

VI. CONCLUSION

This study presents a comprehensive framework for detecting electricity theft using an integrated approach that combines embedded current sensing hardware with a Graph Neural Network (GNN) model. The system is designed to operate in real time, using ACS712 sensors and Arduino Uno microcontrollers to capture current consumption data, which is then analysed by a trained GNN to classify energy usage patterns as either normal or indicative of theft. The proposed methodology emphasizes both cost-effectiveness and scalability, making it adaptable for use in a wide range of deployment environments-from rural areas with limited infrastructure to urban smart grids. The inclusion of graph-based learning techniques enables the system to exploit inter-node relationships within the distribution network, offering improved classification performance over traditional machine learning models. Experimental evaluations demonstrated strong results, with classification accuracy exceeding 91%, low inference latency, and robustness against noise and load fluctuations. The system

BIOGRAPHIES OF AUTHOR

Chaithanya P holds a Bachelor of Engineering degree in Electronics and Communication Engineering from Sri Jayachamarajendra College of Engineering (SJCE), under JSS Science and Technology University, Mysuru, Karnataka, India. During their undergraduate studies, they worked extensively on projects related to embedded systems, IoT-based automation, and microcontroller applications. Their passion for electronics has led them to explore research opportunities in intelligent hardware systems and low-power circuit design. They have also participated in several national-level technical symposia and workshops, further enriching their academic experience.

Assistant Professor in the Department of Electronics and Communication Engineering at Sri Jayachamarajendra College of Engineering (SJCE), JSS Science and Technology University, Mysuru, Karnataka. She has been actively engaged in academic teaching, mentoring, and research, with a particular focus on advanced topics in electronics and communication. Her expertise lies in the design and control of power electronic systems, digital signal processing, analog and digital communication, and microcontroller-based embedded system design. With a strong theoretical foundation and practical insight, she specializes in Power Electronics, VLSI Design, and Embedded Systems, aligning with current industry trends and research advancements. Her teaching interests also span across subjects like Control Systems, Communication Networks, and Electronic Circuits, and she has guided several undergraduate projects in these areas. She is passionate about creating industry-relevant learning environments for students and is an advocate of interdisciplinary applications of ECE in fields such as renewable energy and automation.

Thallam Venkata Srikar completed their undergraduate education in Electronics and Communication Engineering from SJCE, a prestigious constituent college of JSS Science and Technology University, located in Mysuru, India. With a keen interest in communication systems, signal processing, and semiconductor technology, they contributed to multiple team-based projects focusing on wireless data transmission and real- time signal filtering. Apart from academic performance, they were also involved in student technical clubs and volunteered in national hackathons and innovation challenges. Their research interests include RF systems, embedded communication, and digital system design.

Rishith V is a graduate in Electronics and Communication Engineering from Sri Jayachamarajendra College of Engineering (SJCE), JSS Science and Technology University, Mysuru. With a strong foundation in both theoretical and applied electronics, their academic journey involved hands-on learning in areas like IoT, AI-integrated electronics, and sensor networks. They have collaborated on academic papers and technical reports during their final-year research on smart grid systems. Their core areas of interest include Internet of Things (IoT), machine learning for embedded platforms, and smart city technologies.

Sukruth S earned a bachelor's degree in Electronics and Communication Engineering from JSS Science and Technology University, formerly SJCE, Mysuru. Throughout their undergraduate years, they displayed a deep commitment to advancing practical knowledge through mini- projects, internships, and participation in technical expos. Their final year project focused on the integration of artificial intelligence into real-time hardware systems. Their interests lie in automation, robotics, digital electronics, and the intersection of AI and embedded hardware. They aspire to contribute to innovative solutions in smart infrastructure and human–machine interfaces.