DYNAMIC GENERATIVE RESIDUAL GRAPH CONVOLUTIONAL NEURAL NETWORKS FOR ELECTRICITY THEFT DETECTION

SITA SOWJANYA PRAKHYA[1], sowjipavan14@gmail.com
CH. SHYAMALA RANI[2], rani.shyamala@gmail.com
MATHANGI.VASUNDHARA[3], mathangivasundhara112@gmail.com
Assistant Professor[1,2&3], Department Of It, Mvsr Engineering College

ABSTRACT

Illegal electricity users pose a significant threat to the economic and security aspects of the power system by illicitly accessing or manipulating electrical resources. With the widespread adoption of Advanced Metering Infrastructure (AMI), researchers have turned to leveraging smart meter data for electricity theft detection. However, existing models rely on methods that model a single electricity load curve and cannot capture the temporal dependencies, periodicity, and underlying features between electricity consumption cycles. This study introduces a novel electricity theft detection method based on dynamic residual graph networks. Innovatively, it proposes a dynamic topological graph construction technique that allows for the real-time updating of adjacency matrices during the training process, thereby effectively capturing the complex relationships in electricity usage patterns. Utilizing the MixHop graph convolutional network, it delves into the temporal sequence dependencies, periodicity, and hidden characteristics within user electricity consumption data. Additionally, to address the issue of model instability caused by scarce theft data, we employ the SMOTE (Synthetic Minority Over-sampling Technique) oversampling technique and enhance overall classification performance by modifying class weights in the loss function. We trained this network architecture on the real SGCC (State Grid Corporation of China) dataset, and the results demonstrate its superiority over other mainstream existing models.

INTRODUCTION

In the realm of electricity transmission and distribution, the impact of electrical energy losses looms large, exerting a significant influence on the safety and economic efficiency of the power system. These losses are primarily categorized into technical and non-technical losses [1]. As the name suggests, technical losses represent inevitable energy dissipation during electricity transmission, while non-technical losses stem primarily from illicit activities such as electricity theft by end-users. Statistical data reveal that electricity theft results in staggering annual losses of up to $96 billion globally [2]. These nefarious activities engender substantial economic repercussions for electricity distribution companies and nations and raise grave safety concerns. These risks encompass accidents during grid construction, localized power outages, abnormal damage to electrical equipment and infrastructure, and electrical fires [3].

Consequently, research on the detection of electricity theft has paramount practical significance. Electricity theft detection aims to use effective technological means to promptly and accurately identify and prevent unauthorized electricity consumption. This safeguarding is crucial to ensure the power system's regular operation and stable supply. Traditional methods for detecting electricity theft rely heavily on manual inspections and billing analysis, necessitating specialized personnel for on-site investigations. These inspections entail scrutinizing the connections of electrical devices and thestatus of electric meters, with

anomalies identified through comparisons between historical electricity consumption and current usage patterns [4], [5]. Nevertheless, these conventional methods often grapple with inefficiency, rendering them inadequate for overseeing the demands of a large user base. Furthermore, these methods often lack accuracy and real-time capabilities, making it difficult to promptly detect and prevent electricity theft. With the advancement of smart grid technology and the development of Advanced Metering Infrastructure (AMI), smart meters have gained widespread use as a crucial component of AMI [6]. This progress has ushered in a new era in research concerning electricity theft detection. By harnessing the data collected from smart meters, researchers can now leverage advanced machine learning techniques to achieve intelligent monitoring of users' electricity consumption behavior and detect instances of electricity theft. These research approaches can be categorized into three main types: methods based on Support Vector Machines (SVM), methods based on statistical inference, and methods based on artificial neural networks.

LITERATURE REVIEW

Rule-based classification of energy theft and anomalies in consumers load demand profile

- Sonal Jain, Kushan Choksi, N. Pindoriya
- Published in IET Smart Grid 31 July 2019

The invent of advanced metering infrastructure (AMI) opens the door for a comprehensive analysis of consumers consumption patterns including energy theft studies, which were not possible beforehand. This study proposes a fraud detection methodology using data mining techniques such as hierarchical clustering and decision tree classification to identify abnormalities in consumer consumption patterns and further classify the abnormality type into the anomaly, fraud, high or low power consumption based on rule-based learning. The proposed algorithm uses real-time dataset of Nana Kajaliyala village, Gujarat, India. The focus has been on generalizing the algorithm for varied practical cases to make it adaptive towards non-malicious changes in consumer profile. Simultaneously, this study proposes a novel validation technique used for validation, which utilizes predicted profiles to ensure accurate bifurcation between anomaly and theft targets. The result exhibits high detection ratio and low false-positive ratio due to the application of appropriate validation block. The proposed methodology is also investigated from point of view of privacy preservation and is found to be relatively secure owing to low-sampling rates, minimal usage of metadata and communication layer. The proposed algorithm has an edge over state-of-the-art theft detection algorithms in detection accuracy and robustness towards outliers.

Retraction Retraction: An Efficient Power Theft Detection Using Mean-Shift Clustering and Deep Learning in Smart Grid (IOP Conf.

- G. Johncy, A. AnishaFelise
- Published 2020

Energy theft constitutes a major concern for the utility operators in these modern smart homes. The task of detecting and reducing the energy losses has been highly challenging due to insufficient inspection techniques. Energy distribution is comprised of Technical and Non- Technical Losses (NTL). Energy theft generates a major share of Non-Technical Losses which also prompts budgetary misfortunes for the service organizations. The data in the modern smart meters are transmitted in wireless mode. Therefore the smart homes are vulnerable to energy theft. Many new technologies have been adopted to resolve the issue of energy theft in Advance Metering Infrastructure (AMI) in Smart Grids. Consumption pattern must be derived to identify illegal energy consumers. Computational method has been derived to

analyze and identify energy consumption patterns based on data mining techniques. Machine learning technique improves the got client energy utilization readings and guides them on contrasting irregularities in use. Deep Learning method as Convolution Neural Network is implemented on the activity of order methods on client energy utilization and illegal use of electricity and the amount of consumption by energy theft.

Electricity Theft Pinpointing Through Correlation Analysis of Master and Individual Meter Readings

- P. Biswas, Hongyun Cai, +3 authors Vincent W. Zheng
- Published in IEEE Transactions on Smart… 1 July 2020

Electricity theft costs utility companies billions of dollars worldwide annually. The electricity consumption data recorded by consumers' smart meters, coupled with the aggregate energy supply data recorded by master meters provide a new opportunity to pinpoint the source of electricity theft. Existing works on electricity theft pinpointing either assume linear attack modes which often limit their capability in identifying nonlinear electricity theft behaviours, or incur extra cost for model training or sensor installation. Our insight hinges upon the fact that the value of electricity theft loss (ETL) should be more correlated to the meter readings of energy thieves than to those of honest consumers. Guided by this insight, we formulate the problem of electricity theft pinpointing as a time-series correlation analysis problem which does not require linearity assumption of attack modes or any cost of training. Two coefficients are defined to evaluate the suspicion level of a consumer's reported energy consumption pattern. A comprehensive set of experiments has been conducted on a real-world energy usage dataset with several types of attacks, and the results show that our proposed technique significantly improves the pinpointing accuracy when compared with other state-of-the-art methods

EXISTING SYSTEM

The present study utilizes authentic electricity consumption data spanning 1035 days from 42372 users, as made available by SGCC1 for the years 2014 to 2016. There are 38757 legitimate users, while only 3615 instances of electricity theft are identified. An evident imbalance exists between the numbers of normal users and electricity theft cases. This disparity will be addressed in subsequent sections of this paper, where corresponding solutions will be provided. The original dataset exhibits substantial variations in daily electricity consumption among users due to differing energy demands. We performed a standardization operation on the data to facilitate analysis, mapping it to the 0-1 range. Figure 1 shows the electricity load curves of theft and normal users over 56 days.

As depicted in Figure 1, electricity theft users and legitimate users exhibit significant fluctuations in electricity consumption within specific ranges. Therefore, more than one-dimensional load curves are required to analyze the electricity consumption patterns of different user types. In previous research, converting electricity load data into two-dimensional representations has been proven effective [24], [25]. In our study, we partitioned the electricity consumption load of users based on cycles, transforming the one-dimensional power curve into two-dimensional data. Figure 2 illustrates the electricity load curves for both user categories over eight weeks. Figure 2 depicts that the electricity load curves for different user categories exhibit distinct characteristics. Load curves for legitimate users demonstrate a certain level of periodicity, while those for electricity theft users appear erratic and irregular. For instance, the load curve of legitimate users maintains fluctuations within a specific range for five weeks, displaying

similar rising and falling trends. However, significant fluctuations occurred in the second, fourth, and fifth weeks, which can be understood as outliers generated during the data collection process or due to changes in user consumption patterns.For users engaged in electricity theft, there appears to be a noticeable volatility in their consumption patterns. It is crucial to highlight that missing data points were observed on the fifth day of the eighth week, a phenomenon commonly encountered in real-world datasets, often attributed to electricity theft or meter malfunctions.In general, both electricity theft users and legitimate users display irregular electricity consumption fluctuations over specific periods. Relying solely on manual data analysis methods to identify electricity usage patterns and promptly detect electricity theft presents significant challenges. Moreover, during the data acquisition process, factors such as power interruptions, construction activities, smart meter malfunctions, and transmission line damages can lead to missing values and outliers in the dataset. This further complicates the task of accurately identifying electricity theft using conventional models. Hence, there is a need for more advanced models capable of extracting latent features from electricity load curves, enabling a more precise identification of electricity theft behavior.To address the issue of imbalanced data, researchers such as Zanetti et al. [12], Rodriguez et al. [13], and Martino et al. [14] have proposed using one-class SupportVector Machines (SVM) for electricity theft detection. Unlike the classical SVM model that requires two classes of training samples (normal and abnormal), a one-class SVM only requires a specific course of training samples (in our case, using standard consumption patterns). A one-class SVM learns a decision function to classify new data, determining whether it is similar to the training data [15]. Krishna et al. [17] introduced a detector based on the Kullback-Leibler divergence. The corresponding customer is flagged suspicious if consumption readings significantly deviate from the historical distribution over a week. This detector can identify sophisticated electricity theft attacks that bypass checks conducted at the internal nodes of a radial network topology. In [18], a fuzzy C-means method is employed for clustering analysis of user electricity data sets, targeting different types of users, including industrial, commercial, residential, and governmental units. It calculates the Euclidean distance between electricity consumption and the typical cluster vector to determine the user's cluster affiliation. However, the values of the characteristic vectors in this method rely heavily on manual design, rendering it less effective in identifying complex electricity usage patterns.

Disadvantages

• The complexity of data: Most of the existing machine learning models must be able to accurately interpret large and complex datasets to detect Electricity Theft Detection.

• Data availability: Most machine learning models require large amounts of data to create accurate predictions. If data is unavailable in sufficient quantities, then model accuracy may suffer.

• Incorrect labeling: The existing machine learning models are only as accurate as the data trained using the input dataset. If the data has been incorrectly labeled, the model cannot make accurate predictions..

PROPOSED SYSTEM

we propose a model based on a residual graph convolutional network and design a unique end-to-end network architecture. We introduce a dynamic topological graph generation technique that divides the data into 7-day intervals, dynamically generating an adjacency matrix continuously updated during training to capture more accurate node relationships. Subsequently, we employ graph convolutional modules to extract the temporal and spatial characteristics among power load nodes. Additionally, we utilize the SMOTE (Synthetic Minority Over-sampling Technique) method to expand the dataset

before training and apply a cost-sensitive loss function to increase the penalty weight for the minority class.

Advantages :

• We proposed a dynamic topological graph generation method that allows updating connections between nodes during training. This approach offers fewer parameters while delivering more accurate node relationships than static construction methods.

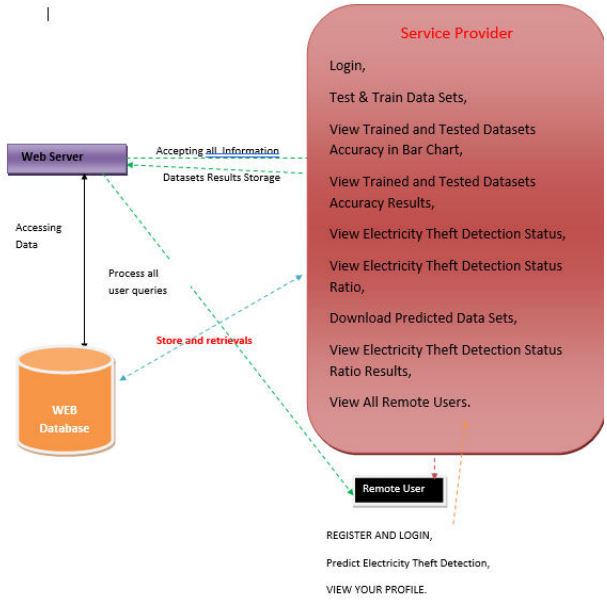
• We employed the Mix-hop graph convolutional network to facilitate the extraction of deep-seated temporal dependencies, periodic patterns, and latent features within user electricity consumption data. By incorporating residual connections, we increased the depth of the network, effectively mitigating the problem of gradient explosions during training and thereby enhancing the model's robustness.

• In response to the issue of sparse electricity theft data, we employed the SMOTE (Synthetic Minority Over-sampling Technique) method to augment the minority class data. We adjusted the class weights in the loss function. We modified the class weights within the loss function, effectively tackling the class imbalance issue in the real dataset and improving the model's generalization abilities.

• Our experiments on a real dataset from the National Grid Corporation reveal that the proposed model consistently outperforms mainstream models in various experiments

IMPLEMENTATION

SYSTEM ARCHITECTURE



MODULES

SERVICE PROVIDER

In this module, the Service Provider has to login by using valid user name and password. After login successful he can do some operations such as Test & Train Data Sets, View Trained and Tested Datasets Accuracy in Bar Chart, View Trained and Tested Datasets Accuracy Results, View Electricity Theft Detection Status, View Electricity Theft Detection Status Ratio, Download Predicted Data Sets, View Electricity Theft Detection Status Ratio Results, View All Remote Users.

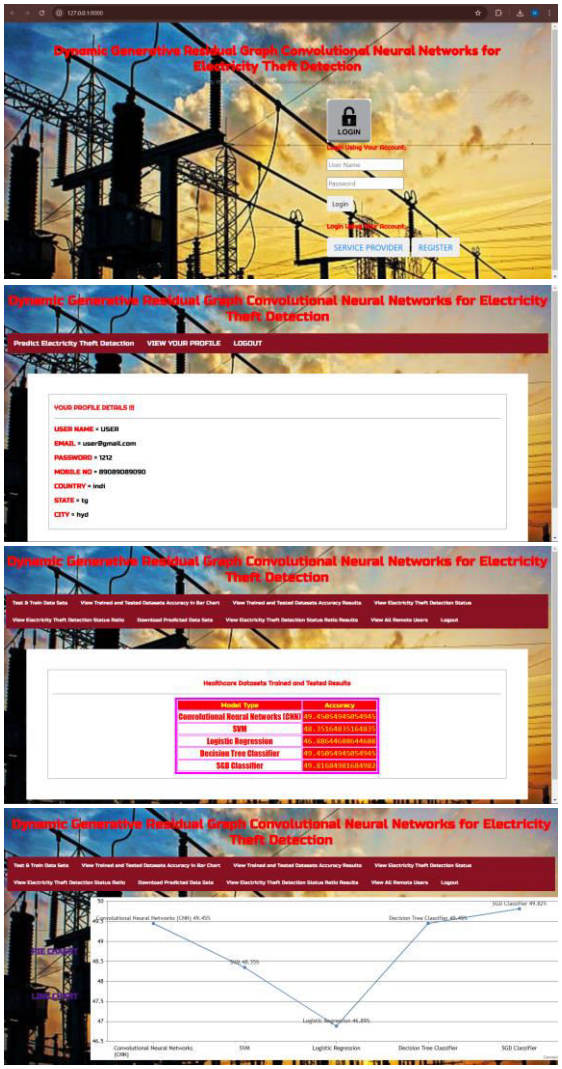
VIEW AND AUTHORIZE USERS

In this module, the admin can view the list of users who all registered. In this, the admin can view the user's details such as, user name, email, address and admin authorizes the users.

REMOTE USER

In this module, there are n numbers of users are present. User should register before doing any operations. Once user registers, their details will be stored to the database. After registration successful, he has to login by using authorized user name and password. Once Login is successful user will do some operations like REGISTER AND LOGIN ,Predict Electricity Theft Detection, VIEW YOUR PROFILE.

RESULT



CONCLUSION

Our research introduces a DGRGNN (Dynamic Generative Residual Graph Neural Networks) model for detecting electricity theft behavior in smart grids. After simulation analysis on a real-world dataset, the following conclusions are obtained:

• We propose a dynamic topological graph generation method to update node connections during training. This approach provides more accurate node relationships with fewer parameters than static construction methods.

• The model utilizes the mix-hop graph convolutional network to better extract deep-seated temporal dependencies and periodic patterns within user electricity consumption data, resulting in improved detection performance. Furthermore, including residual connections increases the network's depth, effectively mitigating the issue of gradient explosions during training and enhancing the model's robustness.

• To address the challenge of sparse electricity theft data, we employ the SMOTE method to augment minority class data and adjust class weights in the loss function. These methods effectively tackle class imbalance in the real dataset, enhancing the model's generalization capabilities.

• Our experiments on a real dataset released by the National Grid Corporation consistently demonstrate that the proposed model outperforms existing mainstream models in multiple experiments. In fact, the proposed DGRGNN model has a wide range of application areas, such as transformer fault detection in electric power scenarios. Since transformer fault data also belong to time series, the proposed DGRGNN model can capture anomalous data.

REFERENCES

[1] R. Jiang, R. Lu, Y. Wang, J. Luo, C. Shen, and X. Shen, ''Energytheft detection issues for advanced metering infrastructure in smart grid,'' Tsinghua Sci. Technol., vol. 19, no. 2, pp. 105–120, Apr. 2014.

[2] L Source Northeast Group. Accessed: Aug. 6, 2022. [Online]. Available: https://www.prnewswire.com/news-releases/96-billion-is-lost-every-yearto-electricity-theft-300453411.html

[3] X. Xia, Y. Xiao,W. Liang, and J. Cui, ''Detection methods in smart meters for electricity thefts: A survey,'' Proc. IEEE, vol. 110, no. 2, pp. 273–319, Feb. 2022.

[4] M. G. Chuwa and F. Wang, ''A review of non-technical loss attack models and detection methods in the smart grid,'' Electric Power Syst. Res., vol. 199, Oct. 2021, Art. no. 107415. [Online]. Available:https://www.sciencedirect.com/science /article/pii/S0378779621003965

[5] Y. Guo, C.-W. Ten, and P. Jirutitijaroen, ''Online data validation for distribution operations against cybertampering,'' IEEE

Trans. Power Syst., vol. 29, no. 2, pp. 550–560, Mar. 2014.

[6] Nandigama, N. C. (2023). Data-Warehouse-Enhanced Machine Learning Framework for Multi-Perspective Fraud Detection in Multi-Stakeholder E-Commerce Transactions. International Journal on Recent and Innovation Trends in Computing and Communication, 11(5), 592–600. https://doi.org/10.17762/ijritcc.v11i5.11808.

[7] J. Nagi, K. S. Yap, S. K. Tiong, S. K. Ahmed, and M. Mohamad, ‘‘Nontechnical loss detection for metered customers in power utility using support vector machines,’’ IEEE Trans. Power Del., vol. 25, no. 2, pp. 1162–1171, Apr. 2010.

[8] J. Nagi, K. S. Yap, S. K. Tiong, S. K. Ahmed, and F. Nagi, ‘‘Improving SVM-based nontechnical loss detection in power utility using the fuzzy inference system,’’ IEEE Trans. Power Del., vol. 26, no. 2, pp. 1284–1285, Apr. 2011.

[9] J. Nagi, K. S. Yap, S. K. Tiong, S. K. Ahmed, and A. M. Mohammad, ‘‘Detection of abnormalities and electricity theft using genetic support vector machines,’’ in Proc. TENCON IEEE Region 10 Conf., 2008, pp. 1–6.

[10] S. S. S. R. Depuru, L. Wang, V. Devabhaktuni, and R. C. Green, ‘‘High performance computing for detection of electricity theft,’’ Int. J. Electric Power Energy Syst., vol. 47, pp. 21–30, May 2013. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0142061512005947

[11 Nandigama, N. C. (2016). Teradata-Driven Big Data Analytics For Suspicious Activity Detection With Real-Time Tableau Dashboards. International Journal For Innovative Engineering and Management Research, 5(1), 73–78.

[12] M. Zanetti, E. Jamhour, M. Pellenz, and M. Penna, ‘‘A new SVM-based fraud detection model for AMI,’’ in Computer Safety, Reliability, and Security, A. Skavhaug, J. Guiochet, and F. Bitsch, Eds. Cham, Switzerland: Springer, 2016, pp. 226–237.

[13] F. Rodriguez, M. D. Martino, J. P. Kosut, F. Santomauro, F. Lecumberry, and A. Fernández, ‘‘Optimal and linear F-measure classifiers applied to non-technical losses detection,’’ in Prog. Pattern Recognit., Image Anal., Comput. Vis., Appl., A. Pardo and J. Kittler, Eds. Cham, Switzerland: Springer, 2015, pp. 83–91.

[14] M. Di Martino, F. Decia, J. Molinelli, and A. Fernández, ‘‘Improving electric fraud detection using class imbalance strategies,’’ in Proc. Int. Conf. Pattern Recognit. Appl. Methods (IPRAM), 2012.

[15] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, ‘‘Scikit-learn: Machine learning in Python,’’ J. Mach. Learn. Res., vol. 12, pp. 2825–2830, Nov. 2011.