IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 11, Issue 5, May 2023

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.379

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 5, May 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1105051 |

A Survey on Statistical Framework for Detecting Electricity Theft Activities in Smart Grid Distribution Networks

Chaitanya Karnakar, Krishna Patil, Aditya Karkhele, Yash Mulik, Dr.Geeta Navale

UG Student, Dept. of Computer Science, Sinhgad Institute of Technology and Science, Narhe, Pune, India

Dept. of Computer Science, Sinhgad Institute of Technology and Science, Narhe, Pune, India

ABSTRACT: Electricity distribution networks have undergone rapid change with the introduction of smart meter technology, that have advanced sensing and communications capabilities, re- sulting in improved measurement and control functions. However, the same capabilities have enabled various cyber-attacks. A par- ticular attack focuses on electricity theft, where the attacker alters (increases) the electricity consumption measurements recorded by the smart meter of other users, while reducing her own measurement. Thus, such attacks, since they maintain the total amount of power consumed at the distribution transformer are hard to detect by techniques that monitor mean levels of consumption patterns. To address this data integrity problem, we develop statistical techniques that utilize information on higher order statistics of electricity consumption and thus are capable of detecting such attacks and also identify the users (attacker and victims) involved. The models work both for independent and correlated electricity consumption streams. The results are illustrated on synthetic data, as well as emulated attacks leveraging real consumption data.

KEYWORDS: Electricity theft detection , smart grid , robustness , CNN , Smart meter

I. INTRODUCTION

Electricity theft has been a major concern worldwide and costs utility companies significant revenue losses [1], [2]. It takes various forms, ranging from physical interventions through illegal connections and meter tampering, to billing irregularities and unpaid bills by customers. The introduction of advanced metering infrastructure has the potential to reduce the risk of electricity theft through its increasing frequency monitoring capabilities. In addition, smart meter technology can lead to effective and accurate load forecasting and on-time troubleshooting for outage remediation and network controlla- bility (see, e.g., [3], [4], [5]). At the same time, it offers new opportunities for tampering with operations of the power grid through cyber-attacks both locally and remotely, that take the form of false data injections. The consequences range from compromising demand response schemes for selected targeted areas, to endangering the power grid's state estimation process or even inducing power outages [6].

There is a growing literature on false data injection (FDI) attack activities (a brief summary is given in [7]). A lotof attention has been paid to the impact of FDI on the grid's state estimation problem [8] and how coordinate attacks can occur [9],[10]. [11] proposes an adaptive procedure to test whether there is a data attack activity combined with a multivariate hypothesis testing method in order to avoid the wrong grid-state estimate. Addressing the problem from a different angle, [12] attempts to prevent the state estimation from being compromised, by approaching the problem from a graph theoretic method aiming to design an optimal set of meter measurements. [13] considers a setting where multiple simultaneous nefarious data attacks are launched and proposes a game theoretic framework to build a defense system

Another thrust has focused on the electricity theft problem and there are two general streams in the literature. One of them focuses on using machine learning and data mining techniques to detect anomalies in the consumption patterns of a household or business, based on smart meters' historical data -see e.g. [14], [15], [16], [17], [18], [19], [20]- potentially augmented with information about the consumer type [20]. These methods can be further subdivided to supervised ones that leverage labels (known FDI vs non-FDI) samples in the training data, and unsupervised ones that try to identify abrupt changes from normal consumption patterns. Supervised methods can be powerful, but availability of labeled FDI samples remains a big challenge. Unsupervised methods are susceptible to the impact of non-malicious factors that alter consumption patterns; e.g. seasonality, change of appliances, change of occupants, and so forth [21].

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 5, May 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1105051 |

A different stream in the literature utilizes information about the architecture of a neighborhood area network in the smart grid [22], [23], [24], [25], [26]. Specifically, it assumes that the electricity provider builds a distribution station within every neighborhood that acts as an "electricity router" to distribute power from the substation to all consumers, A master smart meter (known as the *collector*) measures aggregate power supply from the power provider to all consumers within a certain time interval. Further, smart meters installed at each consumer (households or businesses) record their correspond- ing energy consumption for the same time interval. [24] proposed a method that utilizes such measurements, together with information about the resistances of lines connecting the consumption points to the distribution transformers to estimate technical losses due to low voltage power lines, as well as intrinsic inefficiencies in the transformers. [25], [26] em- ployed such measurements and a linear regression framework to identify electricity theft, wherein the dependent variable corresponds to the aggregate measurement by the collector, and the predictor variables to the household/business smart meter measurements. However, for this approach to work, it is assumed that the predictor variables are uncorrelated, an assumption that is automatically violated when theft occurs, as technically demonstrated in Section II below. Note that this regression framework would work to identify faulty individual smart meters, since their measurements will most likely be random and hence uncorrelated.

II. RELATED WORK

There exists a variety of data mining model with respect to distinguish areas and their electricity theft detection behavior mechanism to perform better than a traditional detection process without using fixed connotation guidelines. This section describes some of the standard benchmark methods that has been used for the said problem. Among the non-technical losses of an electrical supplier, power theft has the most severe and harmful impacts. Fraudulent power usage diminishes supply quality, increases generating load, forces legal customers to pay exorbitant electricity bills. Authors proposed a system for detecting power theft based on a convolutional neural network (CNN) and a long short-term memory (LSTM) approach (Hassan et al., 2022). CNN is a frequently used approach for automating extraction of features and classification. Due to the imbalanced dataset in the area of power theft, however, many AI-based algorithms are susceptible to under-fitting. To circumvent this issue and identify as many forms of theft assaults as feasible, authors proposed an approach based on local outlier factor (LOF) and clustering (Peng et al., 2021). Initially, the load profiles are analyzed via k-means. The LOF is then applied to determine the anomalous degrees of candidates for outliers. Then, a corresponding framework for application in practice is created. Finally, numerical tests based on genuine datasets demonstrate the method's effectiveness

The electricity losses are mostly categorized in to the two types such as technical losses (TLs) and non-technical losses (NTLs) (Schonheit ¨ et al., 2021). The TLs are related to the hardware parts of the whole network i.e. any fault in hardware of the system can cause brown-outs or black-outs which results in electricity losses. The TLs can be recovered by reinstallation of the transformers or the electric meters or power supplies. However, the TLs have some limitations such as, reinstallation costs are usually very high, hardware devices are vulnerable to the weather conditions and the process can be difficult and consuming. The NTLs are difficult to detect because these losses usually occur because of any malicious action performed by the consumer. Therefore, the data driven approaches are used to analyze electricity data consumption and apply various classification techniques on them to detect any normal or malicious behavior in the electricity consumption. The existing electricity-theft detection methods mostly use support vector machine (SVM) classifier to classify the normal and fraudulent consumer. However, this classification technique provides very low accuracy results. Anomalies are the unexpected behaviors or patterns that exist in the data. These anomalies are also called outliers. Anomaly detection is the basic process used to detect any fraudulent behavior. Many techniques are proposed to detect the frauds, thefts that usually occur in the large industrial systems. The anomaly detection is helpful in providing reliability and operational control and theft detection in smart grids. As smart grids have smart sensors which make them vulnerable to the malicious attacks.

III. SYSTEM ARCHITECUTURE

THE PROPOSED ELECTRICITY THEFT DETECTION SYSTEM

This section focuses on the design aspects of the proposed electricity theft detection system. Using statistical examination of the consumption of electrical energy data of both thieves of energy and usual consumers, one can find that the electricity consumption data of energy thieves are typically less non-frequent or frequent, associated with that of usual consumers. This monitoring can facilitate classifying the irregular using of electricity and the periodicity of the electricity consumption.Nevertheless, it is challenging to examine the periodicity of the electricity-consuming data because of many reasons as: 1) it is problematic to study the periodicity of the electricity-consuming data because it is

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 5, May 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1105051 |

1-D time series data with enormous size, 2) The electricity consumption data is frequently incorrect and loud, 3) Several traditional methods of data investigation, e.g., ANN and support vector machine (SVM) cannot be straight carried out to the consumption of electricity data because of the calculation difficulty and the restricted simplification ability. Thus, to face the above challenges, the CNN approach has been adopted in this work.
A realistic electricity consumption dataset released by State Grid Corporation of China is used to train the models. This work is intended to identify electricity theft from the power consumption pattern of users, utilizing CNN-based deep learning and BM techniques. This classifier model is trained to utilize a dataset consisting of daily power consumption data of both normal and fraudulent users in a supervised manner. First, the data is prepared by a data-preprocessing algorithm to train the model. The preprocessing step also involves synthetic data generation for better performance. At the next stage, the proposed model is hyper-tuned and finally, the optimized model is evaluated via the test data
.

IV. CONCLUSION

In this paper, we have primarily focused on how to address coordinated power theft activities detection problem by consid- ering independent and dependent smart meter data generating mechanism. For each case, two scenarios, pairwise and one attacker-many victims, have been thoroughly investigated.

We have separately developed an easy-to-implement detec- tion algorithm to detect attacks and identify attackers and victim nodes. The implementation of the strategy leverages a regularized covariance estimator, followed by close exami- nation of patterns in the resulting matrix. Extensive numerical results based on both synthetic and real data illustrate the superior performance of the proposed methodology.

Note that there is a plethora of machine learning approaches that addresses the detection problem. However, identifying “attackers” and their corresponding “victims” is a more chal- lenging problem that few of these approaches can address Hence, this constitutes an important feature of the proposed methodology.

There are some open problems that merit additional in- vestigation, including scenarios involving multiple attackers and multiple victims. However, such coordinated attacks are more difficult to launch, since they require a higher level of sophistication from the attacker’s perspective.

REFERENCES

1. T. B. Smith, “Electricity theft: a comparative analysis,” *Energy policy*, vol. 32, no. 18, pp. 2067–2076, 2004.
2. T. Ahmad, H. Chen, J. Wang, and Y. Guo, “Review of various modeling techniques for the detection of electricity theft in smart grid environ- ment,” *Renewable and Sustainable Energy Reviews*, vol. 82, pp. 2916–2933, 2018.
3. A. Ipakchi and F. Albuyeh, “Grid of the future,” *IEEE Power and Energy Magazine*, vol. 7, no. 2, pp. 52–62, March 2009.
4. H. Gharavi and R. Ghafurian, “Smart grid: The electric energy system of the future [scanning the issue],” *Proceedings of the IEEE*, vol. 99, no. 6, pp. 917–921, June 2011.
5. G. B. Giannakis, V. Kekatos, N. Gatsis, S. J. Kim, H. Zhu, and B. F. Wollenberg, “Monitoring and optimization for power grids: A signal processing perspective,” *IEEE Signal Processing Magazine*, vol. 30, no. 5, pp. 107–128, Sept 2013
6. V. Pappu, M. Carvalho, and P. M. Pardalos, *Optimization and security challenges in smart power grids*. Springer, 2013.
7. G. Liang, J. Zhao, F. Luo, S. R. Weller, and Z. Y. Dong, “A review of false data injection attacks against modern power systems,” *IEEE Transactions on Smart Grid*, vol. 8, no. 4, pp. 1630–1638, July 2017.
8. Z. H. Yu and W. L. Chin, “Blind false data injection attack using pca approximation method in smart grid,” *IEEE Transactions on Smart Grid*, vol. 6, no. 3, pp. 1219–1226, May 2015.
9. X. Liu, Z. Bao, D. Lu, and Z. Li, “Modeling of local false data injection attacks with reduced network information,” *IEEE Transactions on Smart Grid*, vol. 6, no. 4, pp. 1686–1696, July 2015.
10. S. Bi and Y. J. Zhang, “Graphical methods for defense against false-data injection attacks on power system state estimation,” *IEEE Transactions on Smart Grid*, vol. 5, no. 3, pp. 1216–1227, May 2014.

INNO SPACE
SJIF Scientific Journal Impact Factor

Impact Factor: 8.379

doi cross ref

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details