Volume 12, Issue 4, Page 163-170, 2023 https://doi.org/10.46810/tdfd.1383065 Research Article

BİNGÖL UNIVERSITY

Türk Doğa ve Fen Dergisi
Turkish Journal of Nature and Science
www.dergipark.gov.tr/tdfd

An effective DNN-based Approach for Detecting Energy Theft in Smart Grids through User Consumption Patterns

Muhammed Zekeriya GÜNDÜZ[1*], Resul DAŞ[2]

[1] Bingöl University, Vocational School of Technical Sciences, Department of Computer Science and Technology, Bingöl, Türkiye
[2] Fırat University, Technology Faculty, Department of Software Engineering, Elazığ, Türkiye
Muhammed Zekeriya GÜNDÜZ ORCID No: 0000-0003-4278-7123
Resul DAŞ ORCID No: 0000-0002-6113-4649

**Corresponding author: mzgunduz@bingol.edu.tr*

(Received: 30.10.2023, Accepted: 15.12.2023, Online Publication: 28.12.2023)

Keywords
AMI,
DNN,
Cyber security,
Energy theft,
Smart grid security

Abstract: The advancement of the Internet has been progressively easing human life. The development of mobile communication technologies has led to the widespread adoption of Internet of Things (IoT) applications. Thus, most systems and devices have connected to the Internet more efficiently. The integration of communication systems into critical infrastructures, such as electricity grids, has given rise to the concept of IoT-based smart grids. In smart grid systems, data communication is facilitated through the Advanced Metering Infrastructure (AMI). Due to the inherent characteristics of communication systems, AMI may be vulnerable to cyber-attacks. Some vulnerabilities have resulted in the emergence of cyber-attack vectors against energy consumption data obtained from smart meters. In this study, an effective energy theft intrusion detection system (IDS) based on users' consumption patterns is proposed. A Deep Neural Network (DNN) based classification model was employed to assess the predictability of both honest and malicious consumption patterns. The proposed model exhibits high and adjustable performance. Extensive experiments have been carried out on a real consumption dataset of approximately 2000 customers. Manipulated data from real readings with two different attack vectors were injected into the dataset. K-fold cross-validation technique was used. The proposed model performed a high performance reaching up to 97.4% accuracy.

Kullanıcı Tüketim Kalıpları Aracılığıyla Akıllı Şebekelerdeki Enerji Hırsızlığını Tespit Etmek İçin Etkili Bir DNN Tabanlı Yaklaşım

Anahtar Kelimeler
AMI,
DNN,
Siber güvenlik,
Enerji hırsızlığı,
Akıllı şebeke güvenliği

Öz: İnternetin ilerlemesi insan hayatını giderek kolaylaştırmaktadır. Mobil iletişim teknolojilerinin gelişmesi, Nesnelerin İnterneti (Internet of Things-IoT) uygulamalarının yaygın olarak benimsenmesine yol açmıştır. Böylece, çoğu sistem ve cihaz internete daha verimli bir şekilde bağlanmıştır. İletişim sistemlerinin elektrik şebekeleri gibi kritik altyapılara entegre edilmesi, IoT tabanlı akıllı şebekeler kavramını ortaya çıkarmıştır. Akıllı şebeke sistemlerinde veri iletişimi, Gelişmiş Ölçüm Altyapısı (Advanced Metering Infrastructure - AMI) aracılığıyla sağlanmaktadır. İletişim sistemlerinin doğal özellikleri nedeniyle, AMI siber saldırılara karşı savunmasız olabilir. Bazı güvenlik açıkları, akıllı sayaçlardan elde edilen enerji tüketim verilerine karşı siber saldırı vektörlerinin ortaya çıkmasına neden olmuştur. Bu çalışmada, kullanıcıların tüketim modellerine dayalı etkili bir enerji hırsızlığı saldırı tespit sistemi önerilmektedir. Hem dürüst hem de kötü niyetli tüketim kalıplarının tahmin edilebilirliğini değerlendirmek için Derin Sinir Ağı (Deep Neural Network - DNN) tabanlı bir sınıflandırma modeli kullanılmıştır. Önerilen model yüksek ve ayarlanabilir performans sergilemektedir. Yaklaşık 2000 müşteriden oluşan gerçek bir tüketim veri kümesi üzerinde kapsamlı deneyler gerçekleştirilmiştir. Veri kümesine iki farklı saldırı vektörü ile gerçek okumalardan elde edilen manipüle edilmiş veriler enjekte edilmiştir. K-katlı çapraz-doğrulama tekniği kullanılmıştır. Önerilen model %97,4 doğruluğa ulaşarak yüksek bir performans göstermiştir.

1. INTRODUCTION

IoT is a revolutionary technology that connects daily objects and devices to the Internet [1]. Thanks to the IoT, daily devices and systems can communicate with each other, share information, and make autonomous decisions more efficiently. It leads to development of new applications that can facilitate processes and improve quality of life [2]. Smart grids, which are critical infrastructures, have increased their usability with the development of IoT technologies. An IoT-based smart grid application uses IoT technologies to optimize energy generation, transmission, distribution, and consumption processes. AMI and smart meter are main elements of a smart grid system [3]. A smart meter is an electronic device designed to continuously monitor and log electricity consumption at regular time intervals [4]. Smart meters facilitate real-time monitoring of electricity consumption and offer valuable analytics for consumers and utilities [5]. By enabling bidirectional communication between service providers and consumers, smart meters improve the accuracy of billing, promote demand response, and empower customers to make conscious decisions regarding their energy consumption [6]. Cyber security issues are critical for the development of smart grid applications [7]. AMI causes many different cyber-attacks that occur in the smart grid applications due to its nature [7, 8]. Most of the attacks exploit security vulnerabilities in smart meters. While consumption data in a smart meter, which is an important component of AMI, passes through different stages, it can be sensitive against local and remote data tampering [9]. Storing and transmitting the data are the stages. Illegal manipulation of users' energy consumption data represents an important and major cyber security problem for smart grid systems [10]. Therefore, it is important to develop an effective IDS that can detect data manipulation attacks such as False Data Injection (FDI) in AMI with high accuracy [11].

Non-technical losses (NTLs) resulting from FDI attacks present a significant global concern [12]. While complete elimination of fraud may be unattainable for smart grid suppliers but implementing measures to detect, prevent, and reduce fraud is a viable approach [13]. Classification-based methods utilize detailed electricity consumption data obtained from smart meters. Customer consumption follows a specific statistical pattern in normal conditions [14]. Usage pattern irregularities may indicate malicious activities [15]. Deep learning (DL) models can be used to train a classifier based on real samples and synthetically generated samples. In our work honest samples are obtained from the real dataset. Malicious samples are obtained synthetically. We present a consumption pattern-based energy theft detector that utilize DNN.

DL approaches have gained popularity due to their superior performance compared to traditional machine learning (ML) methods in recent times [16]. The DNN-based IDS was trained using historical data of the honest consumers and synthetic attack data. The attack datasets were generated from honest samples. The classifier was then used to determine whether a new sample is honest or malicious. We have presented a novel IDS for detecting energy theft in AMI. It detects anomalies in the consumption pattern of customers, offering a cost-efficient and high-performing solution for identifying energy theft and detecting data manipulation. In place of current classification-based techniques, the introduced IDS is resilient to attacks and benign alterations in consumption patterns. It achieves higher accuracy, sensitivity, and lower FPR. We obtained promising results with the model that we proposed. The proposed model was trained separately on balanced and imbalanced datasets using cross-validation technique. As a result of the training and testing, confusion matrices were created to measure the performance of the model. Performance metrics were obtained from these matrices. We tested the performance of the IDS with real data from nearly 2000 customers. The dataset serves as a valuable reference point for assessing and comparing various energy theft detection methods. The outcomes affirm the efficiency of our method.

FDI cyber-attacks, which can result in NTL, involve transmitting consumption data to the center with a value that is less than the actual one [17]. Many academic studies have been conducted to detect and prevent these cyber-attacks. Viegas et al. [14] used statistical and ML methods in the study focusing on electricity demand profile and achieved a forecast success of up to 76%. Jokar et al. [18] proposed an IDS using statistical methods based on consumption pattern in AMI and achieved accuracy in the range of 83.25% and 98.75%. Nagi et al. [12] proposed an approach with the SVM method based on load profile for NTL detection and achieved 60% accuracy. Otuoze et al. [19] proposed a framework for energy theft detection insight of smart city planning, but did not mention the performance rate of the proposed framework. Baskaran et al. [8] proposed a framework for detecting FDI data falsification attacks that may occur in AMI, but did not mention the performance rate of the proposed framework. Na et al. [10] created an FDI detector model using CNN and weighted random forest together and achieved an accuracy of up to 95.7% from the model. Kocaman and Tümen [17] created a LSTM-based model achieved an accuracy of up to 93.60%. DNN-based IDS system that we recommended has achieved 97.46% accuracy.

The rest of the paper is organized as follows: materials and methods are presented in section 2, and the experimental results are presented in section 3. Consequently, the conclusion remarks and future work are given in section 4.

2. MATERIAL AND METHOD

Real smart meter consumption data has recently been made available to researchers for academic and commercial studies. The Irish Social Science Data Archive (ISSDA) shares smart meter data that is difficult for individuals to obtain through real-life applications [20]. However, it is almost impossible to obtain data that

has been subjected to real cyber-attacks [21]. Therefore, hacked synthetic datasets are widely used in cyber security works. This, facilitates the development of artificial intelligence applications. Current DL models generally provide better performance than ML methods [22, 23]. Therefore, an up-to-date and high-performance DNN model was presented in this study.

2.1. Dataset and Attack Vectors

The ISSDA made available a dataset in January 2012, obtained through a collaboration with the Irish Commission for Energy Regulation (CER) [20]. This dataset comprises half-hourly electricity usage data for over 5,000 Irish households and businesses over a 536-day period in 2009 and 2010. The participants had smart meters and willingly took part in the project. All the data came from honest users. The size of dataset, diversity of participants, extended data collection period, and public accessibility make it a valuable resource for research in the analysis of smart meter data. In our tests, consumption data of 1948 customers for ten weekdays are used in half-hour periods by creating a 1948 rows and 480 colums dataset. Honest samples are labeled as 1, malicious samples are labeled as 0.

The hacked consumption data was obtained synthetically. Two different attack vectors were used in the study to create synthetic attack data. The attacks are named f_1 and f_2. While the f_1 attack aims the multiplication of the real data by a random number between $0.1 - 0.8$, the f_2 attack aims the reduction of the values at a specific time of the day to zero. The equations of attack vectors f_1 and f_2 are as in (1) and (2).

$$f_1 = random(0.1\text{-}0.8) * all\ honest\ data \quad (1)$$

$$f_2 = 0 * honest\ data\ (specific\ time\ range,\ t=37\text{-}41) \quad (2)$$

There are 48 consumption data per day for a customer. While the first reading is done at 00.30, the last reading is done at 00.00. In this context, the 37th, 38th, 39th, 40th, 41th readings between 18.30 and 20.30, where the consumption is high, were selected for the f_2 attack vector. It is important to visualize how an FDI attack vector that aims to reduce energy consumption data changes the actual data. In this context, the honest consumer pattern of a customer and the f_1, f_2 attack forms are shown in figure 1.

2.2. Deep Neural Network

A DNN model employs techniques to automatically adjust its weights during training on large datasets, allowing it to capture intricate patterns [16]. A DNN model comprises various layers, such as input, dense (fully connected), convolutional, recurrent, activation, and output layers, depending on the architecture. Each of these output neurons serves as input for subsequent layers. Dense layers, prevalent in DNNs, are often combined with other layers to construct models capable of extracting hierarchical features and representations from the input data.

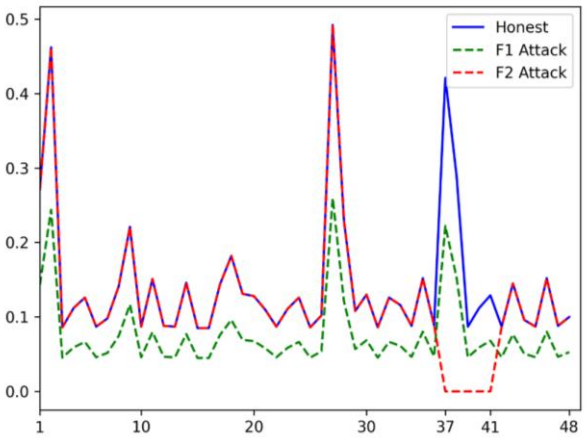

Figure 1. Honest consumer pattern of a customer with $f1, f2$ attacks

The arrangement and quantity of dense layers in a neural network structure are determined by the specific problem at hand and the desired level of model complexity [23]. DNN encompasses diverse areas such as computer vision, natural language processing, and pattern recognition. The training process involves fine-tuning the weights through backpropagation, enabling the DNN to learn and generalize from the provided datasets. DNNs have played a crucial role in advancing the capability of deep learning across a wide range of applications, showcasing their adaptability and effectiveness in capturing complex patterns within data.

2.3. Proposed Approach

Users with no missing data were selected as samples. Hence 10-days consumption data of approximately 2000 customers was designed in 30-minutes periods. The dataset was converted into $1x480$ input data for each customer. Synthetic attack datasets were obtained by applying f_1 and f_2 attacks on all honest-data (h). These datasets were normalized by combining them as $h + f_1, h + f_2, h + f_1 + f_2$. Honest data labeled as 1 and malicious ones as 0. While $h + f_1$ and $h + f_2$ are balance datasets, $h + f_1 + f_2$ is an imbalance dataset. In the training phase, confusion matrices are obtained for each fold using the k-fold cross validation technique. The value of k is five. Model performance results were got from the performance metrics that are obtained from these confusion matrices.

The dataset obtained from ISSDA is consumption data belongs completely honest customers. This honest dataset was exposed to $f_1, f_2, f_1\&f_2$ attacks. Thus, different datasets that were exposed to three different cyber-attacks were obtained. f_1 attack dataset, f_2 attack dataset, and $f_1\&f_2$ attack dataset were generated separately. The dataset created with $f_1\&f_2$ is imbalanced while the others are balanced. After these datasets were normalized, they were given to the DNN algorithm as input for the training and validation processes. Our model consists of seven layers, including the input layer, five fully connected layers and the output layer. The architecture of the energy theft detector is shown in figure 2.

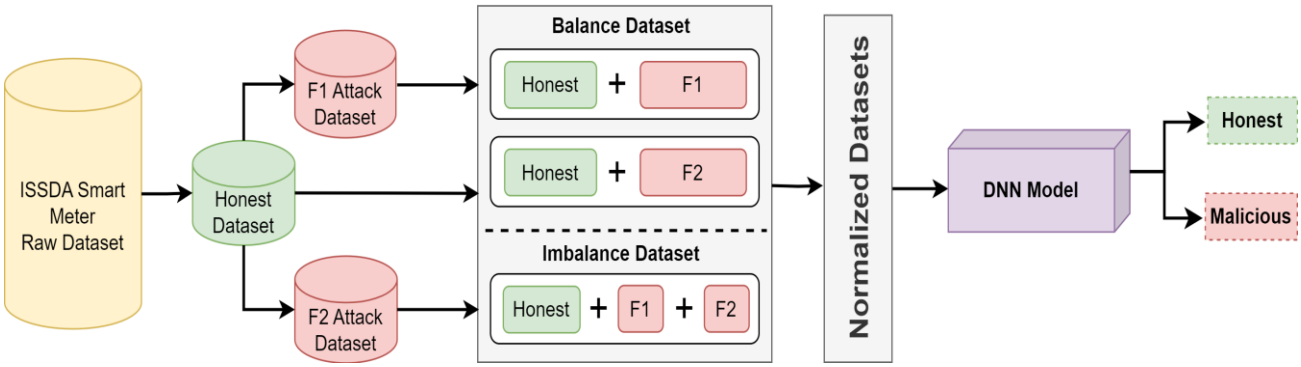

Figure 2. The architecture of the DNN-based energy theft detector

The model accepts 1-by-480 consumption pattern into the input layer. Glorot uniform initializer, ReLU activation function, dropout, early stopping and L2 regularization were preferred to prevent over-fitting with batch normalization in every fully connected layer.

The model flow is as follows:
The input data is represented as a vector. Mathematical output of fully connected (dense) layers is stated as

$$Z^{[i]} = \sum_{i=1}^{5} W^{[i]} A^{[i-1]} + b^{[i]}$$

$$A^{[i]} = \mathrm{Re}\,LU(Z^{[i]})$$

Where:

$A^{[0]}$ represents the input data vector.

$W^{[i]}$ is the weight matrix layer *i*.

$b^{[i]}$ is the bias vector for layer *i*.

$Z^{[i]}$ is the linear operation result at layer *i*.

$A^{[i]}$ is the activation at layer *i*, computed using the *ReLU* activation function. Output layer is stated as

$$Z^{[6]} = W^{[6]} A^{[5]} + b^{[6]}$$

$$\hat{Y} = Sigmoid(Z^{[6]})$$

Where:

$\hat{Y}$ is the output prediction, typically using sigmoid activation function for binary classification.

2.4. Evaluation of the Proposed Model

The confusion matrix, also called as the error matrix, is a common method for evaluating a model's effectiveness. It is generated when the model compares its predicted outcomes with the actual samples, resulting in four key indicators. True Positives (TP - correctly predicted positives), True Negatives (TN - correctly predicted negatives), False Positives (FP - incorrectly predicted positives), False Negatives (FN - incorrectly predicted negatives) [16]. The number of TP and TN represent the number of correctly recognized malicious and honest samples respectively. The number of FP and FN represent the number of misclassified malicious and honest samples respectively. The structure of the confusion matrix used in the study is shown in figure 3.

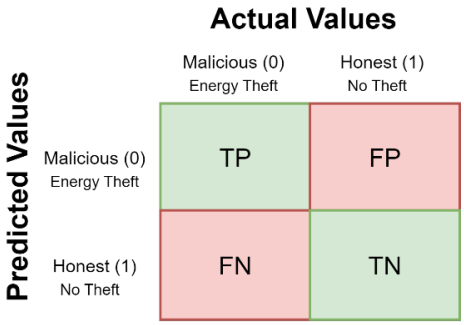

Figure 3. Confusion matrix for 2-classes

Evaluation metrics derived from a confusion matrix include several key indicators used to assess the performance of classification models. These metrics are calculated based on the values present in the confusion matrix. Table 1 shows prominent performance metrics such as sensitivity, specificity, precision, F1-score, accuracy, and False Positive Rate (FPR) obtained from the confusion matrix. Meanwhile recall, True Positive Rate (TPR), hit rate, Detection Rate (DR) and sensitivity represent the same metric. These metrics help in assessing different aspects of a classification model's performance, considering both the correct and incorrect classifications made by the model in comparison to the actual values.

In our experiments, we employed k-fold cross-validation technique to ensure reliable and generalized outcomes. This approach involves partitioning the dataset into training and validation sets to evaluate and compare model performance. During the k-fold cross-validation process, every data point is utilized for both training and validation across each fold of the dataset. K-fold cross-validation serves two fundamental objectives in model assessment. Firstly, it rigorously evaluates a model's performance using the provided dataset, developing its ability to yield accurate predictions or classifications on entirely novel, unseen data. This process ensures a comprehensive understanding of the model's generalization capabilities beyond the training set.

Table 1. Performance metrics obtained from the confusion matrix

Metrics	Equation	Short Description
Sensitivity	$\frac{TP}{TP+FN}$	The performance of the model in detecting malicious samples.
Specificity	$\frac{TN}{TN+FP}$	The performance of the model in detecting honest samples.
Precision	$\frac{TP}{TP+FP}$	The ratio of malicious samples predicted as malicious to all malicious samples.
F1-score	$\frac{2TP}{2TP+FP+FN}$	It is expressed as the harmonic mean between precision and sensitivity.
Accuracy	$\frac{TP+TN}{TP+FP+FN+TN}$	Measures the overall correctness of predictions by the model.
FPR	$\left(\frac{FP}{FP+TN}\right)$	Measures the proportion of actual honest that were incorrectly classified as malicious samples.

Secondly, by systematically partitioning the data into subsets for training and validation, k-fold cross-validation facilitates a comparative analysis of the model performances. This methodological approach aids in determining the optimal model among various algorithms based on their performance metrics within the dataset. As a result, it empowers the informed selection of the algorithm that exhibits superior predictive ability and robustness, crucial considerations for addressing specific problem domains effectively.

3. EXPERIMENTAL RESULTS

We used *Python 3.10* and *Google Colab Pro* (A100 GPU and V100 GPU) platforms for all processes. The honest and malicious datasets were trained and tested using the five-fold cross-validation method. Hyper parameters are given in table 2.

Table 2. Hyper parameters used in the DNN-based model

Hyper parameters	Values
Epoch	100
Batch Size	64
Optimizer	SGD
Learning Rate	0.01
L2 Regularization	0.0001
Kernel Initializer	Glorot_uniform
Seed	12

Balance datasets and imbalance dataset were trained with the model using the statified five-fold cross-validation technique. Confusion matrices were obtained for each fold of each dataset. The confusion matrices obtained for each fold of each dataset is given in table 3.

The performances of the models were measured with the accuracy, precision, specificity, sensitivity, F1-score, and FPR performance metrics. The standard deviations were also calculated. Performance metrics were obtained with the related confusion matrices presented in table 3. The results of these metrics are given in table 4. According to the results, 84.68% accuracy was obtained with the $Honest + f_1$ attack dataset, while 97.46% accuracy was obtained with the *Honest*+f_2 attack dataset. These results were obtained from balanced datasets. An accuracy of 85.28% was achieved with the *Honest*+ f_1 + f_2 attack dataset, which is the imbalance dataset. According to these results, the best performance result was obtained with $Honest + f_2$ attack dataset, which includes $f_2$attack vector. Accuracy, specificity, sensitivity, and FPR performance metrics were used sequentially. Therefore, 84.68%, 89.16%, 81.32%, 10.84% were achieved in the *Honest*+ f_1 attack respectively. 97.46%, 99.90%, 95.38%, 0.10% were achieved in the *Honest*+ $f_2$attack respectively. 85.28%, 81.08%, 87.30%, 18.92% were achieved in the *Honest*+f_1+f_2 attack respectively.

Figure 4 shows the ROC curves of the models trained with different balanced datasets. The AUC values of the model trained with $Honest + f_1$ attack data at each fold are 89.97%, 90.60%, 90.67%, 90.89%, and 90.59%, respectively. The AUC values of the model trained with $Honest + f_2$ attack data at each fold are 99.94%, 99.16%, 99.84%, 99.42%, and 99.71%, respectively. As can be seen from the curves and AUC values in figure 4, the $Honest + f_2$ attack dataset showed a higher accuracy rate than other datasets.

The ROC curve demonstrates the model's capacity to reasonably differentiate malicious samples, even within the context of the dataset's inherent imbalance. ROC analysis serves as a critical indicator of the stability of the trained deep learning model, particularly in challenging scenarios involving imbalanced dataset, where achieving high accuracy presents difficulties.

Figure 5 exhibits ROC curves of the folds corresponding model trained with imbalanced ($Honest + f_1 + f_2$) dataset. The AUC values of the model trained with $Honest + f_1 + f_2$ attack data on each fold were calculated as 92.50%, 91.90%, 91.02%, 90.59%, and 92.61%, respectively. Additionally, figure 5 shows loss-accuracy graphs of the folds corresponding to the imbalanced dataset. When examining the epochs, it becomes apparent that the accuracy generally increases while the loss curve decreases at a certain rate, ultimately reaching a reasonable level. Hence, the model has undergone the learning process in a determined manner, effectively avoiding overfitting.

The information regarding the balancing or imbalancing of datasets in articles is often unclear, which adds complexity to examining the distribution nature of the datasets utilized in various studies. Table 5 shows the comparison of our study with some existing studies. All studies work on the same dataset with different sample rates. In [14], statistical and ML methodologies were explored to forecast consumer demand profiles. Among these models, the support vector machine (SVM) algorithm achieved an accuracy rate as high as 75.8% for this particular application.

Table 3. Confusion matrices for each fold of the proposed model

Dataset	Fold 1		Fold 2		Fold 3		Fold 4		Fold 5	
Honest + f_1 *Attack*	362[a]	28[b]	355[a]	34[b]	354[a]	35[b]	339[a]	51[b]	349[a]	41[b]
	97[c]	293[d]	104[c]	286[d]	80[c]	310[d]	61[c]	328[d]	66[c]	323[d]
Honest + f_2 *Attack*	390[a]	0[b]	389[a]	0[b]	388[a]	1[b]	390[a]	0[b]	389[a]	1[b]
	9[c]	381[d]	20[c]	370[d]	12[c]	378[d]	49[c]	340[d]	7[c]	382[d]
Honest + f_1&f_2 *Attack*	695[a]	84[b]	720[a]	59[b]	686[a]	93[b]	717[a]	63[b]	738[a]	41[b]
	79[c]	311[d]	116[c]	274[d]	92[c]	298[d]	113[c]	276[d]	120[c]	269[d]

TP ([a]), FP ([b]), FN ([c]), TN ([d])

Table 4. Performance results and standard deviation of the proposed model according to balance and imbalance datasets

	Accuracy (%)	Precision (%)	Specificity (%)	Sensitivity (%)	F1-score (%)	FPR (%)
Honest + f_1 *Attack*	84.68±1.41	90.30±1.99	89.16±1.55	81.32±2.88	85.51±1.0	10.84±1.55
Honest + f_2 *Attack*	97.46±1.95	99.90±0.13	99.90±0.13	95.38±3.44	97.56±1.81	0.10±0.13
Honest + f_1&f_2 *Attack*	85.28±0.76	91.27±2.38	81.08±3.55	87.30±1.47	89.20±0.67	18.92±3.55

Mean ± standard deviation

Table 5. A comparision of exist studies on the same dataset

Reference	Year	Simulation Platform	Poroposed Model	Dataset Resource	Accuracy(%)
[14]	2015	N/A	SVM	ISSDA	75.8
[24]	2017	N/A	Density-based clustering	ISSDA	93.2
[25]	2016	N/A	SVM-based	ISSDA	94.0
[26]	2020	N/A	MP-ANN	ISSDA	93.4
[13]	2018	Python 3.x	DNN-based	ISSDA	93.0
Our study	**2023**	**Python 3.10**	**DNN-based**	**ISSDA**	**97.4**

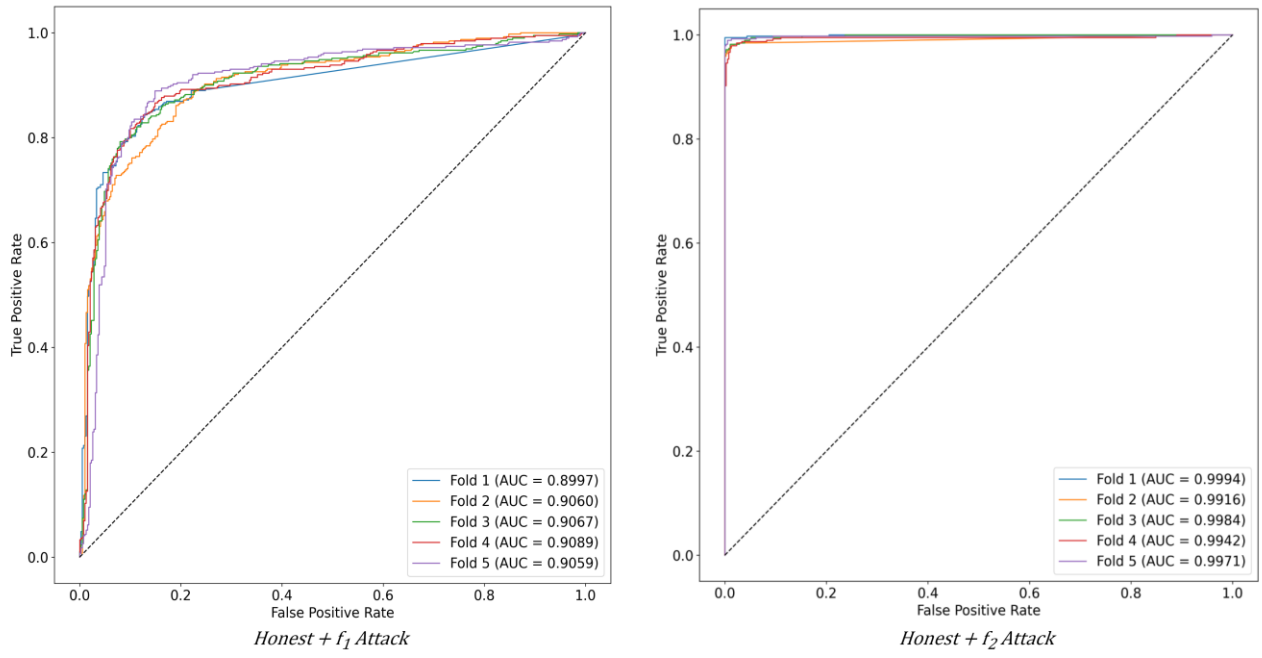

Figure 4. The ROC curves of the DNN models for balanced datasets

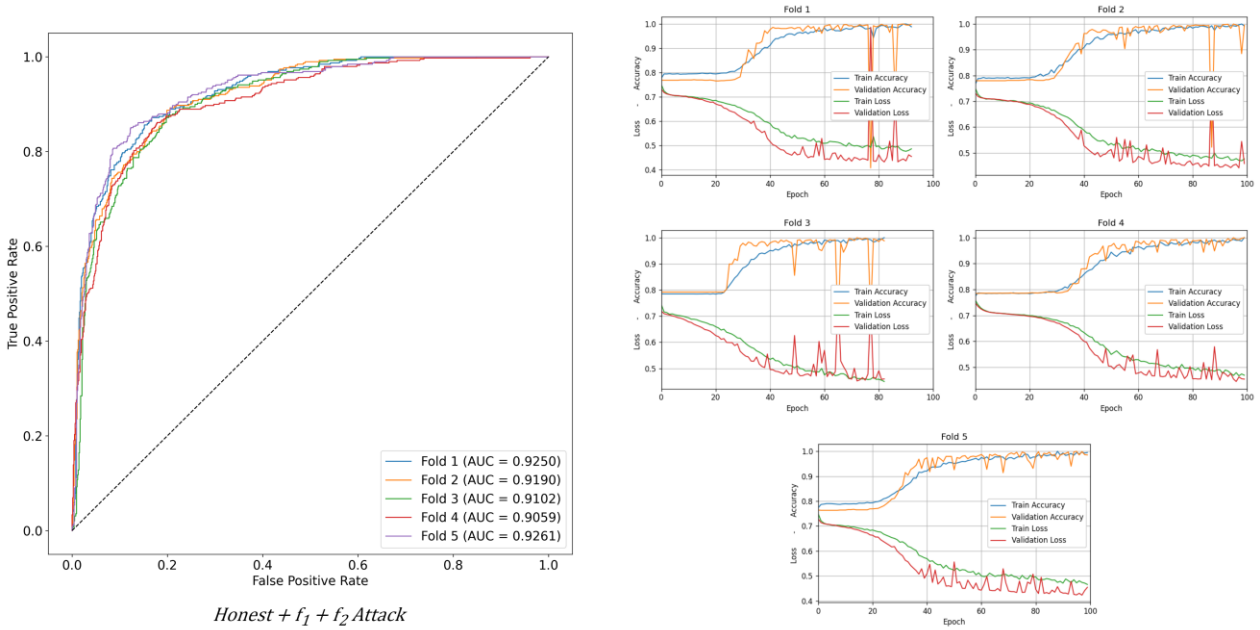

Figure 5. The ROC curve and loss-accuracy graphs along with epochs for imbalanced dataset

In [24], the method using smart meter data effectively detects energy theft by focusing on abnormal load profiles, outperforming traditional unsupervised techniques like k-means clustering, GMM clustering, and DBSCAN. It has high accuracy. The technique is helpless for attackers that do not generate load profiles with an abnormal shape. Moreover, the proposed SVM-driven detector in [25] exhibits notable efficacy, demonstrating an average detection rate of 94% alongside a false alarm rate of 11%. In [26], a multilayer perceptron was developed the detection of energy theft in distribution systems using the Multilayer Perceptron Artificial Neural Network (MP-ANN) algorithm. They obtained 93.4% accuracy. In [13], a DNN based system was designed for identifying energy theft on a per-customer basis. This specialized DNN-based detector attains a detection rate reaching 93%. Simulation outcomes underscore a considerable advancement in performance when compared with state-of-the-art shallow detection methodologies. In some studies, it has not been clearly stated the real data exposed to which attack vectors and whether the data is imbalanced or balanced. In our proposed model, real data were exposed to f_1 and f_2 attacks separately and together, as explained before. Our energy theft detector provides up to 97.4% accuracy.

4. CONCLUSION

In this study, the detection of energy theft in smart grids was investigated on synthetic data obtained from a real dataset with a DNN-based approach. The method relies on consumption data patterns. The manipulated values were injected into the dataset with two different attack vectors. Both balanced and imbalanced datasets were investigated with the same DNN network. Stratified five-fold cross-validation technique was used to obtain more generalizable results during the training phase. The results were compared. Widely accepted performance metrics such as accuracy, precision, specificity, sensitivity, FPR, and AUC-ROC were used to evaluate the performance of the model. We observed that the classification done with the *Honest*+f_2 balanced dataset performed better results than the imbalanced dataset. The model achieved 97.46% accuracy, 99.9% precision, 99.9% specificity, 95.38% sensitivity, 97.56% F1-score, 0.1% FPR, and up to 99% AUC-ROC when it was tested on *Honest*+f_2 balanced dataset. Compared to other data-driven methods evaluated on the same dataset, we achieved the best accuracy of 97.46% among existing studies. By proactively identifying irregularities in consumption patterns, the DNN-based approach offers a robust IDS, enhancing the reliability and resilience of energy distribution systems. In future work, a study will be done with more cyber-attack vectors and a CNN-based hybrid model will be developed on a larger balanced dataset.

Declaration of Competing Interest

The authors declare no conflicts of interest.

Acknowledgment

The authors would like to thank ISSDA and CER for providing the real smart meter consumption data.

This manuscript was produced from Muhammed Zekeriya Gündüz's Ph.D. thesis ("Development of New Approaches for the Detection and Solution of Security Vulnerabilities in the Internet of Things Based Smart Grids").

REFERENCES

[1] Gunduz MZ and Das R. Internet of things (IoT): Evolution, components and applications fields. *Pamukkale University Journal of Engineering Sciences.* 2018; 24(2). doi: 10.5505/pajes.2017.89106.

[2] Gunduz MZ and Das R. Analysis of cyber-attacks on smart grid applications. *International*

Conference on Artificial Intelligence and Data Processing (IDAP). 2018; doi: 10.1109/IDAP.2018.8620728.

[3] Gündüz MZ and Daş R. Akıllı Şebekelerde İletişim Altyapısı ve Siber Güvenlik. *Iğdır Üniv. Fen Bil Enst. Der.* 2020;10(2). doi: 10.21597/jist.655990.

[4] Sahoo S, Nikovski D, Muso T, and Tsuru K. Electricity theft detection using smart meter data. *IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT).* 2015; doi: 10.1109/ISGT.2015.7131776.

[5] Emmanuel M and Rayudu R. Communication technologies for smart grid applications: A survey. *Journal of Network and Computer Applications*. 2016;74 doi: 10.1016/j.jnca.2016.08.012.

[6] Otuoze AO *et al.* Electricity theft detection framework based on universal prediction algorithm. *Indonesian Journal of Electrical Engineering and Computer Science*. *2019;*15(2) doi: 10.11591/ijeecs.v15.i2.pp758-768.

[7] Gunduz MZ and Das R. Cyber-security on smart grid: Threats and potential solutions. *Computer Networks. 2020;*169. doi: 10.1016/j.comnet.2019.107094.

[8] Baskaran H., Al-Ghaili AM, Ibrahim ZA, Rahim FA, Muthaiyah S and Kasim H. Data falsification attacks in advanced metering infrastructure. *Bulletin of Electrical Engineering and Informatics.* 2021;10(1). doi: 10.11591/eei.v10i1.2024.

[9] Das R and Gunduz MZ. Analysis of cyber-attacks in IoT-based critical infrastructures. *International Journal of Information Security Science*. *2019;8(4).*

[10] Na L, Xiaohui X, Xiaoqin M, Xiangfu M, and Peisen Y. Fake Data Injection Attack Detection in AMI System Using a Hybrid Method. *IEEE Sustainable Power and Energy Conference (iSPEC).* 2021. doi: 10.1109/iSPEC53008.2021.9735875.

[11] Bhattacharjee S and Das SK. Detection and Forensics against Stealthy Data Falsification in Smart Metering Infrastructure. *IEEE Transactions on Dependable and Secure Computing. 2021;*18(1) doi:10.1109/TDSC.2018.2889729.

[12] Nagi J, Yap KS, Tiong SK, Ahmed SK, and Mohamad M. Nontechnical Loss Detection for Metered Customers in Power Utility Using Support Vector Machines. *IEEE Transactions on Power Delivery.* 2010;25(2). doi: 10.1109/TPWRD.2009.2030890.

[13] Ismail M, Shahin M, Shaaban MF, Serpedin E and Qaraqe K. Efficient detection of electricity theft cyber attacks in AMI networks. *IEEE Wireless Communications and Networking Conference (WCNC).* 2018. doi: 10.1109/WCNC.2018.8377010.

[14] Viegas JL, Vieira SM, Sousa JMC, Melício R, and Mendes VMF. Electricity demand profile prediction based on household characteristics. *12th International Conference on the European Energy Market (EEM).* 2015. doi: 10.1109/EEM.2015.7216746.

[15] Viegas JL, Esteves PR, Melício R, Mendes VMF, Vieira SM. Solutions for detection of non-technical losses in the electricity grid: A review. *Renewable and Sustainable Energy Reviews*. 2017;80. doi: 10.1016/j.rser.2017.05.193.

[16] Ayaz I, Kutlu F, Cömert Z, DeepMaizeNet: A novel hybrid approach based on CBAM for implementing the doubled haploid technique. *Agronomy Journal*. doi: 10.1002/agj2.21396.

[17] Kocaman B and Tümen V. Detection of electricity theft using data processing and LSTM method in distribution systems. *Sādhanā*. 2020; 45(1) doi: 10.1007/s12046-020-01512-0.

[18] Jokar P, Arianpoo N, Leung VCM. Intrusion detection in advanced metering infrastructure based on consumption pattern.*IEEE International Conference on Communications (ICC).* 2013. doi: 10.1109/ICC.2013.6655271.

[19] Otuoze AO, Mustafa MW, Mohammed OO, Saeed MS, Surajudeen-Bakinde NT, Salisu S. Electricity theft detection by sources of threats for smart city planning. *IET Smart Cities. 2019;* 1(2) doi: 10.1049/iet-smc.2019.0045.

[20] ISSDA, Irish Social Science Data Archive. https://www.ucd.ie/issda/data/commissionforenergyregulationcer/

[21] Şahin K, Hizal S, Zengin A. Design and Implementation of ADevs-Based Cyber-Attack Simulator for Cyber Security. (*IJSIMM). 2022;* 21(1). doi: https://doi.org/10.2507/IJSIMM21-1-587.

[22] Dinçer Y, İnik Ö. Çevresel Seslerin Evrişimsel Sinir Ağları ile Sınıflandırılması. *KONJES.* 2023;11(2). doi: 10.36306/konjes.1201558.

[23] Haq EU, Pei C, Zhang R, Jianjun H, Ahmad F. Electricity-theft detection for smart grid security using smart meter data: A deep-CNN based approach. *Energy Reports.* 2023;9 doi: 10.1016/j.egyr.2022.11.072.

[24] Zheng K, Wang Y, Chen Q, Li Y, Electricity theft detecting based on density-clustering method. *IEEE Innovative Smart Grid Technologies (ISGT-Asia).* 2017. doi: 10.1109/ISGT-Asia.2017.8378347.

[25] Jokar P, Arianpoo N, Leung VCM. Electricity Theft Detection in AMI Using Customers' Consumption Patterns. *IEEE Transactions on Smart Grid.* 2016;7(1) doi: 10.1109/TSG.2015.2425222.

[26] Souza MA, Pereira JLR, Alves GO, Oliveira BC, Melo ID, Garcia PAN. Detection and identification of energy theft in advanced metering infrastructures. *Electric Power Systems Research*. 2020; 182. doi: 10.1016/j.epsr.2020.106258.