Advanced Machine Learning Techniques for Cyberattack Detection in Smart Grids: A Comprehensive Review.

Md Rakibul Ahasan
Miami University

***Abstract*—This review paper presents recent advancements in machine learning-based defense strategies for securing smart power grids against cyberattacks, such as False Data Injection Attacks (FDIAs) and electricity theft. These threats challenge the stability and security of power systems, and the need for improved detection methods. Key contributions include the use of deep autoencoders with attention and Long Short-Term Memory (LSTM) structures for enhanced electricity theft detection in Advanced Metering Infrastructures (AMIs), ensemble learning-based anomaly detectors for robustness against evasion attacks, and Graph Neural Networks (GNNs) for better generalization across power system topologies and unseen attack types. The paper highlights the progress in developing more adaptable and robust ML-based defense mechanisms for smart power grid security. Additionally, it outlines future research directions focusing on the security of Connected Autonomous Vehicles (CAVs) as a potential Cyber-Physical System (CPS).**

***Index Terms*—Electricity theft, Deep learning, Evasion Attacks, Smart Grids, Robust Detection, Autoencoders, Hypermeter Optimization, False Data Injection Attacks, Graph Autoencoder, Graph Neural Network, Connected Autonomous Vehicle.**

I. Introduction

Electricity theft and FDIAs pose significant challenges to the integrity and reliability of smart power grids, leading to financial losses and compromised grid performance. AMIs with smart meters are deployed to monitor energy consumption and detect anomalies. However, these systems are prone to cyber-attacks that manipulate consumption data, deviating from the grid's stability and efficiency. The cyber-physical nature [14] of smart grids requires robust security measures to safeguard against FDIAs, which can lead to incorrect operational decisions and system instability. Developing smart defense strategies to detect and counter these theft attacks is crucial for ensuring the security and resilience of smart power grids. To mitigate this issue the authors performed various research and found some state-of-the-art anomaly detectors that are better than existing benchmark detectors. The key contributions are:

- A study investigates autoencoders for detecting electricity cyberattacks in smart grids, comparing fully connected feed-forward, LSTM-based seq2seq structures, and various autoencoder models. The research highlights the effectiveness of LSTM-based RNN autoencoders and AEA models, with the AEA anomaly detector achieving a 94% detection rate (DR) and a 5% false alarm rate (FA), outperforming other detectors.
- Another research examines the impact of evasion attacks on electricity theft detectors in smart grids. It proposes two strong evasion attacks (NNP and NND) and tests them under different settings. A robust electricity theft anomaly detector is designed, combining an AAE, convolutional-recurrent, and feed-forward model using sequential ensemble learning, showing stable performance against high levels of evasion attacks.
- The proposed GAE-based detector for FDIAs in power grids is topology-aware and provides unsupervised anomaly detection. It achieves superior detection rates of 87-91% and 94-99% in topology-specific and generalized settings, respectively. The detector is scalable and performs graph classification to determine system status and node classification to localize attacks.

The rest of the article is organized as follows, Section II explains the data set used by the authors, and the details of multiple cyber attacks in smart grids. Section III talks about the data partitioning of different data sets and Section IV discusses the structure of the deep learning anomaly detector used and proposed by the authors. Section V outlines the results obtained from the research and Section VI concludes the article with a potential future work in security strategies in connected autonomous vehicles.

II. Dataset and Relevant Cyber Attack Types

The author uses energy consumption profiles from AMI to train and test the anomaly detectors. Regarding the smart grid, they used the bus topology from the IEEE bus systems and modeled a power system. The AMI dataset is collected from the State Grid Corporation of China (SGCC) [1] and the Irish Smart Energy Trial (ISET) [2].

A. AMI Datasets

The SGCC dataset contains the daily electricity consumption of 40,000 customers for over three years. Let's consider $E_c(d,t)$ as the electricity consumption of customer c at day d. The $R_c(d)$ is the electricity consumption from the AMI, then in benign sample $R_c(d) = E_c(d)$. The malicious sample in the SGCC data is replacing the actual energy consumption value with the zero value.

The ISET dataset has readings from 3,000 residential units, where the data granularity is 30 minutes and the total data collection period is one and a half years. That sums to around 25,000 data readings. It is established that the dataset is malicious when the $R_c(d) \neq E_c(d)$. The authors adopted a false data injection approach with assumptions that a customer's action is responsible for his own meter tempering but not affecting the other customer. The types of cyber attacks are:

Partial reduction attacks: This attack function $f(E_c(d,t))$ reduces the energy consumption by a fixed value α, and the value of α is randomly selected between 0.1 to 0.8. Moreover, another function $f(E_c(d,t))$ deviates the energy consumption by a dynamic random function β where $\beta(d,t) = rand(0.1, 0.8)$.

Selective by-pass attacks: This attack function reports zero energy consumption for a time interval $[t_i(d), t_f(d)]$ and the original energy consumption in the rest of the time interval. A low-level attack has a duration of 4 Hr of zero energy reporting, whereas a high-level attack has a duration of 24 Hr.

Price-based load control attacks: There are cases where the electricity pricing varies throughout the day. In this attack, the attacker moves the electricity consumption where the electricity price is low as $f(E_c(d,t)) = E_c(d, T-t+1)$.

Evasion attacks: The evasion attack represents a malicious dataset that fools the anomaly detector by pretending to be a benign sample. The authors formulate two evasion attacks [5] namely NNP which depend on average perturbation value and NND attacks which use the average Euclidean distance of its own and surround reading from the customer.

The NNP attack generates adversarial samples by adding a perturbation value ϵ to the target electricity reading $E_c(d,t)$. The perturbation value ϵ is calculated using the gradient of the model's loss function concerning $E_c(d,t)$ and the surrounding readings of the customer on the same day. The goal is to generate a malicious reading R_c^{adv} that maximizes the model's loss, thereby increasing the probability that the theft will go undetected. This process is iterative, and after each time step, a clipping function is applied to ensure that the generated and original readings have similar patterns.

The NND attack fools the detector by minimizing the Euclidean distance between the target electricity reading $Ec(d,t)$ and the perturbation value ϵ. For each generated adversarial sample R_c^{adv}, the perturbation value ϵ varies based on the average of $Ec(d,t)$ and its k neighboring readings.

B. Power System Modelling from IEEE Bus System

Due to the non-disclosure agreements and national security reasons [3] power system data is not readily available. Hence the authors built a power system from an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{W})$ where buses are nodes $\mathcal{V}$, power lines are edges $\mathcal{E}$. Line admittance is represented by a weighted adjacency matrix $\mathcal{W} \in \mathbb{R}^{n \times n}$. If buses i and j are connected, a weight $\mathcal{W}_{ij}$ is associated with edge $e = (i,j)$ based on line admittance. The active power P_i, reactive power Q_i and the power flow analysis are conducted using Newton's method in MATLAB MATPOWER toolbox to determine active and reactive power flows. The benign data samples $x_b(t,i)$ denotes the normal data types at bus i and timestamp t. That includes 96 power dynamics for 17,000 timestamps over six months. The types of attacks are:

Direct attack: The direct attack constructs x_m by adding random perturbations to a benign sample x_b, with the perturbations limited by a scaling factor α such that $|\alpha| \leq 0.05$.

Replay attack: The replay attack creates x_m by repeating a value from a past time step $t-1$ to substitute for the value at the current time step t.

General attack: The general attack formulate x_m from the authentic measurement values. The construction is as follows:

$$x_m(t,i) = x_b(t,i) + (-1)^{\delta} \alpha . \gamma . \text{Range}(x_b(t,i)),$$

where δ is a binary random variable and γ is a uniformly distributed random variable within the interval [0, 1].

III. Data partitioning

In the AMI datasets, the authors used sampling. Let B and M represent the normalized benign and malicious datasets. The benign dataset B is divided into two parts, B_1 and B_2, at a 2:1 ratio. The training set, denoted X_{TR}, consists of B_1. The test set is created by merging M with B_2, labeling benign samples as 0 and malicious samples as 1. To maintain balance and prevent bias, the ADASYN sampling [4] technique is used to equalize the prevalence of both classes in the test set, resulting in X_{TST} with labels Y_{TST}. Similarly, for supervised detectors, after balancing the combined dataset using ADASYN, it is split into training and test sets in the same 2:1 ratio.

In the power system dataset, The evaluated supervised models are trained and tested using both benign x_b and malicious x_m samples, while unsupervised models are trained on x_b and tested on both x_b and x_m. The datasets for training (X_{TR}), validation (X_{VA}), and testing (X_{TS}) are split such that they each contain an equal proportion of benign and malicious samples, with 80% of the samples used for training, and 10% each for validation and testing.

IV. Electricity Theft Detection Details

The authors uses different deep-learning approaches to detect thefts in the smart grid. The approaches use different autoencoders [5], an ensembling method of an autoencoder, and a convolutional neural network (CNN) [6], or a multi-task graph neural network (GNN) [7] [8]. The selection depends on the type of cyber attacks residing in the dataset and based on the types of data set either AMI or Power system dataset. The details of deep learning are below:

Fully Connected Stacked Autoencoder (FC-SAE): It consists of an encoder and a decoder. The encoder includes an input layer with several neurons, multiple dense hidden layers, and a latent layer. The decoder has several dense hidden layers and an output layer. The input layer receives a data vector, which is then compressed by the encoder into a lower-dimensional latent representation. The training process of the SAE involves optimizing these weights and biases to minimize the reconstruction error, typically using gradient descent algorithms.

Sequence-to-Sequence Stacked Autoencoder (Seq2Seq SAE): It is designed for time-series data like energy consumption profiles, which exhibit temporal correlations. Unlike the Fully Connected SAE (FC-SAE), which cannot capture these correlations, the Seq2Seq SAE utilizes a deep Recurrent Neural Network (RNN) based on Long Short-Term Memory (LSTM) units to address the vanishing gradient problem and better learn temporal dependencies over long intervals. The Seq2Seq SAE comprises two deep LSTM-RNNs: an LSTM encoder that compresses the time-series input into a hidden state, and an LSTM decoder that reconstructs the original time-series data from the encoded state. The model aims to minimize the reconstruction mean-square error.

Fully Connected Variational Autoencoder (FC-VAE): It has a similar architecture to the Fully Connected Stacked Autoencoder (FC-SAE) concerning the input layer and hidden dense layers. However, the VAE stands out in generative modeling due to its continuous latent space, which facilitates random sampling for interpolation. This is achieved by generating mean (μ) and variance (σ^2) vectors for the latent variable, which are then utilized by the decoder for sample generation. For theft detection, the reconstruction probability is computed and compared against a threshold. The loss function of the VAE includes both a reconstruction term and a regularization term. The training algorithm for the FC-VAE is similar to that of the FC-SAE, with additional steps for generating the distribution parameters (μ and σ^2) and updating the model parameters using backpropagation.

Sequence-to-Sequence Variational Autoencoder (LSTM-VAE) It incorporates Long Short-Term Memory (LSTM) structures for both the probabilistic encoder and decoder, originally used for language modeling but here employed for anomaly detection. The LSTM-VAE architecture mirrors that of the LSTM-SAE, with the addition of a latent distribution similar to the FC-VAE. The training algorithms for both LSTM-SAE and LSTM-VAE follow a similar logic, with the LSTM-VAE generating mean (μ_x) and variance (σ_x^2) vectors, which are then used by the decoder to generate samples.

Autoencoder with Attention (AEA): In SAE and VAE, the encoder compresses the input sequence into a fixed-length context vector, while the decoder generates the output sequence based on this vector. The AEA model introduces an attention layer between the encoder and decoder, which assigns different weights to each time step, allowing for a dynamic context vector that better captures the entire sequence. This attention mechanism is handy for anomaly detection, as it can highlight time steps with higher contributions to the desired output, resulting in a more accurate reconstruction error. The AEA model follows a sequence-to-sequence algorithm, with the attention layer receiving inputs from both the encoder and decoder's hidden states. The attention weights are computed using an alignment scoring function, and the context vector is a weighted sum of the encoder's hidden states. The model is trained to minimize the mean squared error between the original and reconstructed sequences for SAE and AEA, or the reconstruction probability for VAEs. A sample is labeled as malicious if the cost function exceeds a specific threshold, otherwise, it is considered benign.

Sequential ensemble learning-based robust detector: It comprises an input layer followed by four blocks an RNN-based Attentive Autoencoder (AAE), a convolutional-recurrent part, fully connected layers, and an output layer. The sequence aims to differentiate between benign and malicious samples by capturing complex patterns and temporal correlations. The AAE block includes an encoder and decoder with LSTM layers and an attentive layer. The encoder encodes the input time series into a hidden state, while the attentive layer assigns importance to time steps based on their contribution to the desired output. The decoder reconstructs the output with the attentive layer's output. This convolutional-recurrent block applies convolution and max-pooling to the AAE's output for feature extraction and compression. Additional LSTM layers capture more hidden features and temporal correlations. The fully connected layer reshapes the output of the convolutional-recurrent part for the final decision. The output layer has two neurons representing the malicious and benign classes.

Multi-Task GNN Model: The GNN model is structured into three stages: In the first state the joint graph layers are for preliminary feature extraction, then the task-specific graph layers for capturing relevant features, and finally a decision stage that combines learned features to enhance detection performance. The model employs a Graph Neural Network (GNN)-based detector with convolutional Chebyshev graph layers for optimization and feature extraction. Task-specific layers are followed by dense layers that determine the probability of an attack presence using a sigmoid function. The final output layer provides an improved decision based on the information obtained in the previous steps.

V. Result Discussion

The authors compare multiple detectors of different characteristics with their proposed theft detection approach. The detector characteristics are shallow/deep/graph, types of machine learning (supervised/unsupervised), and static/dynamic approach. The evaluation metrics of deep learning models are detection rate (DR) where DR=TP/(TP+FN). The DR symbolizes the detected malicious sample and TP, and FN refer to the true positive and false negative. The false alarm rate (FA) is the false detection of the benign sample as a malicious sample. The FA value is calculated as FA=FP/(FP+TN), where FP is the false positive and TN is the true negative. The accuracy is measured by ACC=(TP+TN)/(TP+TN+FP+FN).

As a starter, the authors researched whether deep learning provides better performance compared to shallow detectors for electricity theft detection. They have used a variable autoencoder setting explained in Section IV. Key changes are in several layers, optimizer, dropout rate, hidden layer activation function, and output activation function. The settings of individual autoencoders are FC-SAE (8, Adam, 0.4, Sigmoid, Softmax), LSTM-VAE (4, Adam, 0.2, Sigmoid, Sigmoid), FC-VAE (8, Adam, 0.4, Relu, Softmax), LSTM-VAE (4, SGD, 0, Tanh, Sigmoid), and AEA (6, SGD, 0, Sigmoid, Sigmoid). This parameter selection is done through an optimization algorithm. In conclusion, they showed that for the datasets SGCC and ISET in Section II, LSTM models capture temporal characteristics of electricity consumption data better than FC models. Variational autoencoders (VAE) optimize performance in the sense of improved detection rate (DR) and false alarm (FA) rate since they better capture the variability. AEA model adds higher value than the two models, SAE and VAE, and presents higher improvement in DR and FA. This is due to the potentiality of the attention layer to support detection capabilities. The LSTM-AEA model records an improvement of up to 21% in DR and 13% in FA when it is compared against other detectors, establishing better capability of detection by the deep learning architecture of electricity theft.

In another research, the authors implemented a robust evasion attack scenario explained in Section II-A. The attacker's knowledge of evasion settings is classified into the white box, gray box, and black box settings. In the white box setting the attacker has full knowledge of the utility's detector and the dataset. In the gray box and black box settings, the attacker has partial knowledge and no knowledge regarding the utility settings respectively. Hence, here they use a substitute dataset Australian Smart-Grid-Smart-City Customer Trial (ASCT) dataset [9]. Moreover, compared to the benchmark evasion attack FGSM [10], BIM [11], AutoAttack [12], and C & W [13] attack the authors introduce robustness in their proposed evasion attack. They worked on small perturbation value ϵ to create the adversarial samples close to the original data pattern. The deterioration of energy consumption is observed for different evasion attack injections from 0% to 100% along with the traditional attack from 100% to 0% has a 25% increment in either attack and sum up to 100% of attacks:

- In white-box settings, where attackers have full access, performance deterioration ranges from $6.3-6.7\%$ to $33.4-35.8\%$ whereas the benchmark attacks result in a lower decrease rate of $4.5-5.3\%$ to $24.7-29.2\%$.
- In gray-box settings, with partial access, the deterioration ranges from $3.9-4.4\%$ to $24-26.9\%$ for proposed attacks and $2.3-3.1\%$ to $15.1-19.8\%$ for benchmark attacks.
- In perfect-match black-box settings, where attackers have no access to the model or dataset, the deterioration ranges from $3-3.4\%$ to $19.2-22.2\%$ for proposed attacks and $1.2-2.1\%$ to $10.4-15.1\%$ for benchmark attacks.

In these circumstances, the authors proposed a detector combining AEA, CNN-RNN, and a fully connected output block. This detector detects anomalies with small perturbation values ranging from 0.2 to 0.4 and k values ranging from 2 to 4. The key parameters for the AAE network and Convolutional networks are as follows:

- **AAE Network:**
  - LSTM cells in encoding layers: (500, 300, 200).
  - LSTM cells in decoding layers: (200, 300, 500).
  - Optimizer: SGD.
  - Dropout rate: 0.2.
  - Weight constraint: 1.
  - Hidden and output activation function: Sigmoid.
- **Convolutional Network:**
  - Optimizer: SGD.
  - Hidden activation: ReLU.
  - Fully connected layer neurons: 500.
  - Dropout rate: 0.2.
  - Weight constraint: 1.
  - Hidden and output activation functions: ReLU.

The result of the proposed detector demonstrates significant improvement over benchmark detectors in various settings, with an average enhancement of $24-32.8\%$ in the white box, $14.2-24.8\%$ in the gray box, and $10-20.4\%$ in black-box settings at 100% evasion percentage. Notably, in the most challenging scenario of white-box setting with NND attack type and 100% evasion level, the detector's performance only declines by 3%. Moreover, combined evasion attacks show that the performance of the proposed detectors only decreases by 0.3-0.8%, enabling robustness against combined attacks. In 50% of evasion injection level with a combination of five types of evasion attacks, the proposed detector outperforms state-of-the-art adversarial defense mechanisms by 12.3-17.8%, 7.5-13.2%, and 5.1-11.1% in white, gray, and black-box settings, respectively.

TABLE I
HYPERPARAMETER FOR SMART GRID THEFT DETECTION AND LOCALIZATION

Model	Hyperparameters
ARIMA	Differencing: 1, Moving Averages: 0
SVM	Kernel: Scale, Gamma: Sigmoid
MLP	Layers: 4, Units: 32, Optimizer: Adamax, Dropout: 0, Activation: ELU
RNN	Layers: 3, Units: 16, Optimizer: Adam, Dropout: 0.2, Activation: ReLU
CNN	Layers: 4, Units: 32, Neighborhood Order: 5, Optimizer: Rmsprop, Activation: ReLU
Graph Models	Layers: 4, Units: 32, Neighborhood Order: 3, Optimizer: Adam, Activation: ReLU

Table I represents the optimal hyperparameter selection adopted by a multi-stage grid search algorithm [5] used to detect the theft and localize the affected node in the smart grid. The types of attacks are mentioned in Section II-B. The theft detection and localization of nodes using the machine learning model are performed in IEEE 14-bus, 39-bus, and 118-bus systems.

TABLE II
DETECTION PERFORMANCE OF THEFT IN SMART GRID

Metric	14-bus	39-bus	118-bus
DR	98.5	99.3	100.0
FAR	0.92	0.54	0.12
ACC	98.1	98.9	99.8

TABLE III
DETECTION OF LOCALIZATION BY THE PROPOSED MULTI-TASK GNN

Metric	14-bus	39-bus	118-bus
DR	99.2	99.7	100.0
FAR	0.72	0.42	0.08
ACC	98.7	99.2	99.7

Table II and III show the proposed multi-task GNN detector performance of detecting theft in a smart grid and localizing the node that is affected by the theft. These results outperform the existing shallow/deep/graph framework because of the proposed GNN model's ability to capture spatial aspects and task-specific features using stacked convolutional Chebyshev graph layers. The improvement in Detection Rate (DR) is maximum from 23.2% to 30.1%, 15% to 21.2%, and 9.4% to 11.5%, respectively for the existing shallow/deep/graph framework. For the attack localization task, the enhancements are 20.2% to 26.6%, 9.6% to 17.3%, and 4.1% to 5.8% compared to the same categories of benchmarks. Moreover, for the higher bus system 118-bus, the proposed GNN model provides the best detection rate compared to the smaller bus 14-bus, and 39-bus systems.

VI. CONCLUSION

In summary, this collection of research has advanced the field of electricity theft detection in smart grids by developing and evaluating various machine learning-based detectors. The studies have identified evasion attacks that significantly deteriorate the performance of existing detectors and proposed a robust detector combining multiple neural network architectures, demonstrating resilience against these attacks. Additionally, deep autoencoder anomaly detectors have been introduced, outperforming shallow and static architectures in detecting electricity theft. These detectors, particularly those based on LSTM-AEA, showed superior detection performance. Furthermore, research on graph neural network (GNN)–based detectors highlighted their effectiveness in detecting FDIAs in smart grids. These topology-aware detectors exhibited superior detection rates compared to classical machine learning-based detectors and demonstrated scalability to larger systems. Overall, these contributions have significantly enhanced the ability to detect and mitigate electricity theft in smart grids, ensuring secure and efficient energy distribution systems.

VII. FUTURE WORK

In future work, I would like to explore the performance of deep learning structure over the same combination of datasets used for shallow architecture. This means that deep learning is trained over the benign sample and malicious sample and tested over the benign sample and malicious sample. Moreover, The deterministic setting of cyber attack injection of evasion and traditional can changed to a random setting. Hence, at a given time the injection percentage varies any value between 0% to 100% as the sum of evasion attack and traditional attack. Then the detector is unaware of the attack volumes and deteriorates from its standard performance.

In application perspective, Connected and autonomous vehicles are set to become the biggest part of future advanced intelligent transportation systems. This shift will be driven by simultaneous progress in machine learning and wireless communication technologies, leading to a more feature-rich and efficient vehicular ecosystem. However, there are significant security concerns associated with using ML in such a critical environment, where a wrong ML decision could be more than just an inconvenience and potentially result in the loss of lives. The types of attacks are sensor spoofing, vehicle-to-vehicle (V2V) communication attacks, and vehicle control system hijacking. In the initial stage, we will formulate the attack function for CAVs and use a suitable dataset for a deep learning-based anomaly detector. It's highly unlikely that the detector used for the smart grid can reused for CAVs hence a literature review of existing work is a prerequisite. Once the attack functions are ready we can use the existing detector or any proposed detector to evaluate the detector's performance in detecting cyber attacks on CAVs.

In line with my thought process, I have reviewed a paper ”Intelligent Sensor Attack Detection and Identification for Automotive Cyber-Physical Systems.” Here the author aims to solve a problem where two out of three sensors associated with the Inertial Measurement Unit (IMU) are attacked and the deep learning detects those attacks without prior knowledge of the attacks. They have used an unmanned ground vehicle robot where the IMU, left, and right wheel sensor is placed and run over a 220 m straight road at constant velocities (0.4, 0.7, 1.0, 1.3, and 1.6 m/s). They have collected 124550 sensor data and used 80% for training and 20% for testing. The input data consists of 8 variables, x/y/z direction acceleration, and velocity from IMU, and left, right velocity from two left and right encoders. The output is classified into 7 Classes where Class 0 refers to normal operation, Class 1, 2, 3 refers to when one of the sensors is attacked, and Class 4, 5, 6 refers to two sensors attacked. During the IMU attack, the car runs at a constant velocity, and during the left and right wheel encoder attack, the vehicle's velocity deviates to the range of -0.3 to +0.3 m/s.

As a result, the authors compared Neural network (NN), Gated Recurrent Unit (GRU), and LSTM performance to detect the security breach and the LSTM outperformed all the deep learning methods and achieved an accuracy of 97.33%. However, this deep learning detection method is not performed over a real vehicle where the vehicle speed is constantly changed, and a real road surface with turning, stop, slope, etc. I am considering this paper as a benchmark and further extend it by modeling an accurate vehicle, road model with a robust attack function. Finally, I will formulate a detector so that it ensures better detection.

References

[1] Z. Zheng, Y. Yang, X. Niu, H.-N. Dai, and Y. Zhou, “Wide and deep convolutional neural networks for electricity-theft detection to secure smart grids,” IEEE Trans. Ind. Informat., vol. 14, no. 4, pp. 1606–1615, Apr. 2018.

[2] “Irish Social Science Data Archive,” UCD Library. [Online]. Available: http://www.ucd.ie/issda/data/commissionforenergyregulationcer/

[3] R. Atat, M. Ismail, M. F. Shaaban,E. Serpedin, andT.Overbye, “Stochastic geometry-based model for dynamic allocation of metering equipment in spatio-temporal expanding power grids,” IEEE Trans. Smart Grid, vol. 11, no. 3, pp. 2080–2091, May 2020.

[4] C. Lu, S. Lin, X. Liu, and H. Shi, “Telecom fraud identification based on ADASYN and random forest,” in Proc. 5th Int. Conf. Comput. Commun. Syst., Shanghai, China, 2020, pp. 447–452.

[5] A. Takiddin, M. Ismail, U. Zafar and E. Serpedin, ”Deep Autoencoder-Based Anomaly Detection of Electricity Theft Cyberattacks in Smart Grids,” in IEEE Systems Journal, vol. 16, no. 3, pp. 4106-4117, Sept. 2022, doi: 10.1109/JSYST.2021.3136683.

[6] A. Takiddin, M. Ismail and E. Serpedin, ”Robust Data-Driven Detection of Electricity Theft Adversarial Evasion Attacks in Smart Grids,” in IEEE Transactions on Smart Grid, vol. 14, no. 1, pp. 663-676, Jan. 2023, doi: 10.1109/TSG.2022.3193989.

[7] A. Takiddin, R. Atat, M. Ismail, K. Davis and E. Serpedin, ”A Graph Neural Network Multi-Task Learning-Based Approach for Detection and Localization of Cyberattacks in Smart Grids,” ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Rhodes Island, Greece, 2023, pp. 1-5, doi: 10.1109/ICASSP49357.2023.10096822.

[8] A. Takiddin, R. Atat, M. Ismail, O. Boyaci, K. R. Davis and E. Serpedin, ”Generalized Graph Neural Network-Based Detection of False Data Injection Attacks in Smart Grids,” in IEEE Transactions on Emerging Topics in Computational Intelligence, vol. 7, no. 3, pp. 618-630, June 2023, doi: 10.1109/TETCI.2022.3232821.

[9] “Smart-Grid Smart-City Customer Trial Data.” [Online]. Available: https://tinyurl.com/9wftuaf2 (Accessed: Mar. 2022).

[10] I. J. Goodfellow, J. Shlens, and C. Szegedy, “Explaining and harnessing adversarial examples,” Mar. 2015, arXiv:1412.6572v3.

[11] A. Kurakin, I. Goodfellow, and S. Bengio, “Adversarial examples in the physical world,” Feb. 2017, arXiv:1607.02533v4.

[12] F. Croce and M. Hein, “Reliable evaluation of adversarial robustness with an ensemble of diverse parameter-free attacks,” in Proc. Int. Conf. Mach. Learn., 2020, pp. 2206–2216.

[13] N. Carlini and D. Wagner, “Towards evaluating the robustness of neural networks,” in Proc. IEEE Symp. Security Privacy, 2017, pp. 39–57.

[14] Yu, Xinghuo, and Yusheng Xue. ”Smart grids: A cyber-physical systems perspective.” Proceedings of the IEEE 104.5 (2016): 1058-1070.