https://doi.org/10.31449/inf.v49i7.7037 Informatica **49** (2025) 157–164 **157**

Research on Operation and Anomaly Detection of Smart Power Grid Based on Information Technology Using CNN+Bidirectional LSTM

Yong Zhang*, Yueda Gao, Zhe Zhao
E-mail: zhangy_y@hotmail.com
Electric Energy Measurement Branch of Inner Mongolia Power (Group) Co., Ltd., Inner Mongolia 010020, China
* Corresponding author

Keywords: smart grid, abnormal power consumption, intelligent algorithm, integrated learning

Received: August 30, 2024

The accurate detection of abnormal users in the grid is conducive to maintaining the stability of the smart grid. This paper briefly introduces the smart power grid and the intelligent algorithm used to detect users with abnormal power consumption in the power grid. The intelligent algorithm combined the bidirectional long short-term memory (LSTM) and a convolutional neural network (CNN) to extract the features from the power consumption data of the users and then used the adaptive boosting (AdaBoost) model to classify the users. The field operation test was carried out in a small substation. The proposed method was compared with the single bidirectional LSTM and CNN methods. The findings showed that the proposed method had the best performance in the simulation experiment, with a precision of 98.7%, a recall rate of 97.9%, and a false drop rate of 3.6%, and its receiver operator characteristic (ROC) curve deviated the most from the diagonal line and had the largest area enclosed. In the field operation test, the proposed method obtained a lower and more stable false detection rate (approximately 3.6%).

Povzetek: Raziskava je omogočila izdelavo CNN+Bidirectional LSTM za zaznavanje anomalij v pametnem omrežju, kar omogoča stabilno in učinkovito identifikacijo nepravilne porabe električne energije.

1 Introduction

With the rapid advancement of information technology and the optimal adjustment of energy structure, the smart grid has emerged as a crucial development direction of future power systems, gathering increasing attention [1]. Characterized by information, automation, and intelligence, the smart grid can realize real-time monitoring, intelligent scheduling, and efficient management of the power network [2]. In the operation process of the smart grid, the power transported by the power grid will produce technical and non-technical losses. The former refers to the inherent loss of all equipment in the power grid and is an unavoidable loss, while the latter is caused by abnormal power consumption, equipment failure, and network system failure [3]. Abnormal power consumption is the main cause of non-technical loss of the smart grid and will affect the stability and security of the smart grid. Therefore, rapid and accurate detection of abnormal power consumption in the smart grid is conducive to maintaining its stability [4]. Artificial intelligence has been making significant breakthroughs in various new fields, such as the perception of real-world signals. In these areas, artificial intelligence has occasionally surpassed human capabilities and will continue to do so even more in the future. Artificial intelligence can also be applied to the management of smart grids. The relevant research, as shown in Table 1, is all related to improving the performance of smart grids. Some studies focus on demand distribution calculation in smart grids, some on the interaction efficiency between users and intelligent systems in smart grids, and some on intelligent controllers in smart grids. This paper focuses on identifying abnormal electricity users in smart grids by using a convolutional neural network (CNN) to extract electricity consumption features from users and then using bidirectional long short-term memory (LSTM) for identification. By utilizing a deep learning algorithm to improve the efficiency and accuracy of identifying abnormal users in smart grids, it provides an effective reference for enhancing the stable operation of smart grids.

Table 1: Related research

Author	Research content	Research results
Deng et al. [5]	They proposed a dual decomposition-based distributed approach to improve demand response in smart grid.	The test results verified the effectiveness of the proposed algorithm.
Jo et al. [6]	They proposed a lightweight privacy-protecting metering protocol for bidirectional communication between smart grid users and power	The test results showed that the protocol can further improve the speed of message authentication.

	systems.	
Grilo et al. [7]	They developed a solution for extracting relevant data from specific locations in low-voltage power grids and efficiently transmitting the data to intelligent controllers via wireless sensor networks.	The experimental results showed the effectiveness of the proposed method.

This paper briefly introduces the smart power grid and the intelligent algorithm used to detect users with abnormal power consumption in the power grid. The intelligent algorithm integrated the bidirectional LSTM with a CNN to extract the features from the power consumption data of users and then used the adaptive boosting (AdaBoost) model to classify these users. Moreover, simulation experiments were carried out in a laboratory. A field operation test was carried out in a small substation.

2 The detection method of abnormal operation in the smart power grid

2.1 Smart grid

With the expansion of the power grid scale, traditional power management approaches have been unable to meet the needs of power grid operation. However, with the progress of information technology and the Internet, it has become easier to collect and analyze operational data of power grids [9], giving rise to smart grids. Compared with the conventional power grid, the smart grid can use the sensors installed on the equipment to automatically collect and analyze the operation data produced by the power grid, then intelligently schedule the power grid, and detect the anomalies in the power grid [10].

Figure 1 presents the basic architecture of smart grid operation monitoring. The corresponding sensing facilities are installed in power grid equipment such as power transmission lines, transformer equipment, distribution station equipment, and smart meters. The sensing facilities collect real-time operation data from the corresponding power grid equipment and then transmit the data to the smart grid control center via the Internet. The control center analyzes and processes the collected data [11].

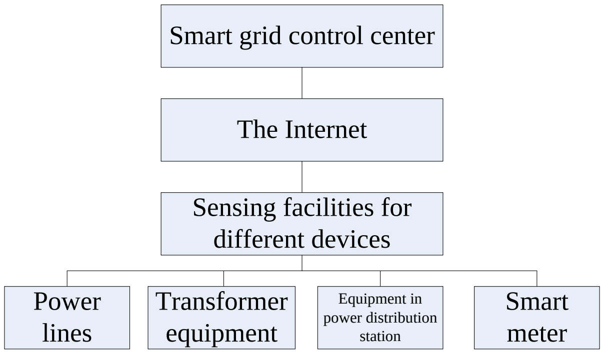

Figure 1: Basic architecture of smart grid operation monitoring.

2.2 Detection of abnormal power consumption

The smart grid will produce technical loss and non-technical loss during operation. The former is the unavoidable loss of power grid equipment during operation, while the latter is the power loss caused by other reasons, including abnormal power consumption and equipment failure, among which abnormal power consumption is the main reason [12]. The causes of abnormal power consumption include illegally stealing electricity, privately pulling wires, private change of electricity meters, and other informal means. The main goal of these informal means of electricity consumption is to pay for less electricity or for temporary convenience. Therefore, during the smart grid operation, it is necessary to detect abnormal users from the grid in time [13].

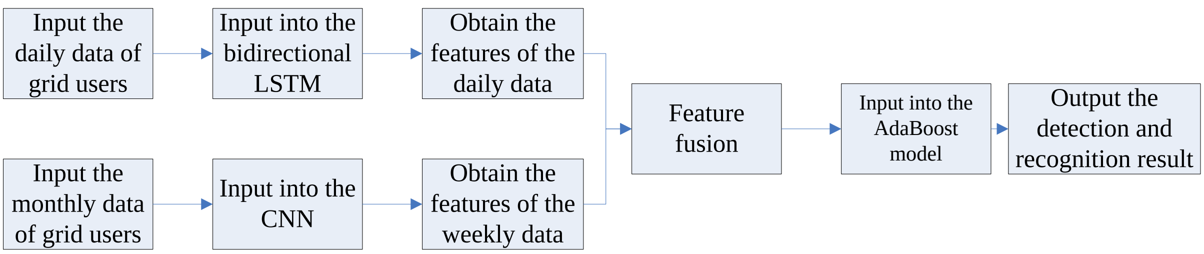

Figure 2: Abnormal power consumption detection process based on bidirectional LSTM+CNN.

The scale of the smart grid is huge; hence, the amount of data collected by sensors is very large. It is difficult to carry out statistical analysis of big data by manual alone, and it is necessary to use computers to assist in processing [14]. This paper employs a deep learning algorithm to detect users with abnormal power consumption in the smart grid. The power consumption data generated by users in the process of using a smart grid has different time scales, such as daily data, weekly data, and monthly data. The data at different time scales have different hidden characteristics. Although the data at a single time scale can also reflect the users with abnormal power consumption, it is not comprehensive enough. Therefore, this paper uses different deep learning algorithms to extract the features from the electricity consumption data at different time scales, and then fuses the features for the detection and identification of abnormal users. As shown in Figure 2, this paper selects the bidirectional LSTM to extract the daily data features from power grid users and the CNN to extract the monthly data features. After the fusion, the AdaBoost model [15] is employed to detect and identify abnormal users among power grid users. The specific process is described as follows.

① The daily data of power grid users collected by the device sensor is input into the bidirectional LSTM. The computational formula of the one-directional LSTM is:

$$\begin{cases} f_t = f(W_f \cdot [h_{t-1}, x_t] + \theta_f) \\ i_t = f(W_i[h_{t-1}, x_t] + \theta_i) \\ \tilde{C}_t = \tanh(W_C[h_{t-1}, x_t] + \theta_C) \\ C_t = f_t \cdot C_{t-1} + i \cdot \tilde{C}_t \\ o_t = f(W_o[h_{t-1}, x_t] + \theta_o) \\ h_t = o_t \cdot \tanh(C_t) \end{cases}, (1)$$

where $\tilde{C}_t$ and C_t are the temporary state and updated state of the current memory unit, h_t is the hidden data state input at the current time, x_t is the current input, f_t, i_t, and o_t are the output of three gated units, forget, input, and output, at the current time, W_f, W_i, and W_o are the weight in corresponding gated units, and θ_f, θ_i, and θ_o are the bias in corresponding gated units.

The formula of the bidirectional LSTM is:

$$\begin{cases} h_i^+ = \overrightarrow{LSTM}(W_1^+ x_i, W_2^+ h_{i-1}^+) \\ h_i^- = \overleftarrow{LSTM}(W^- x_i, W_2^- h_{i+1}^-) \\ h_i = h_i^+ \oplus h_i^- \end{cases}, (2)$$

where $\overrightarrow{LSTM}()$ is the calculation function of the hidden layer of the forward LSTM, i.e., the application of equation (1), $\overleftarrow{LSTM}()$ is the calculation function of the hidden layer of the backward LSTM, which is the application of equation (1) (its only difference is reversely inputting x_t), h_i^+, h_i^-, h_i are the current forward LSTM hidden layer output, the current backward LSTM hidden layer output, and the current final output, x_i is the current input, $W_1^+, W_2^+, W_1^-, W_2^-$ are weights of the forward and backward LSTM calculations [16].

② The weekly data of power grid users collected by the device sensor is input into a CNN in the form of two-dimensional arrangement for convolutional feature extraction. The convolutional formula of the CNN is:

$$x_j^l = f\left(\sum_{j \in M} x_i^{l-1} \cdot W_{ij}^l + b_j^l\right), (3)$$

where x_j^l is the feature map of the convolver output, x_i^{l-1} is the feature output of the i-th convolutional kernel in the last convolutional layer after pooling, W_{ij}^l is the weight parameter between the i-th convolutional kernel and the j-th convolutional kernel, b_j^l is the bias of j convolutional kernels of l layers, M is the number of convolutional kernels, and $f(\bullet)$ is an activation function.

③ In the first two steps, the bidirectional LSTM finally outputs hidden state h_i, and the CNN finally outputs the convolution feature. They are processed into features of the same dimension, and then the two features are fused.

④ The fused features are input into the AdaBoost model for recognition and detection. When training the AdaBoost model, the weight of the samples with classification errors is increased, and then the new weak classifier is trained by random sampling according to the weights. The above steps of "sampling according to weights - training weak classifiers - increasing the weight of samples with classification error" are repeated, and multiple weak classifiers are obtained. The weight of weak classifiers and the updated formula of the weight of training samples are:

$$\begin{cases} \alpha_m = \frac{1}{2}\log\frac{1-e_m}{e_m} \\ w_{m+1,i} = \frac{w_{m,i}\exp(-\alpha_m y_i G_m(x_i))}{Z_m} \\ Z_m = \sum_{i=1}^{N} w_{m,i}\exp(-\alpha_m y_i G_m(x_i)) \end{cases}, (4)$$

where α_m is the weight of the m-th classifier, e_m is the classification error of the m-th classifier, $w_{m,i}$ is the weight of sample i when training the m-th classifier, y_i is the result label of sample i, $G_m(x_i)$ is the computation function of the m-th classifier, x_i is the input of sample i, and Z_m is the normalization factor.

3 Simulation experiment

3.1 Experimental environment

The simulation experiment was carried out in laboratory servers, which were configured as Windows 11 operating system, Core i7 processor, and 32G memory.

3.2 Experimental data

The dataset used was from the open dataset provided by the State Grid Corporation of China (http://www.sgcc.com.cn/). The dataset selected in this paper covered the period from May 1, 2015, to June 1, 2018. The dataset contained 35,689 users, of whom 93.5% were normal users. Abnormal users accounted for 6.5%. The public dataset from the State Grid Corporation of China was collected from real grid users, so the proportion of normal and abnormal users was realistic, with only a small number of abnormal users. The dataset with the unbalanced proportion of normal and abnormal users was not suitable for training intelligent algorithms, and it was easy to cause the intelligent algorithms to be biased. Therefore, this paper used SMOTE to balance the sample number, and after that, the ratio of normal users to abnormal users was 1:1. The basic process of the SMOTE algorithm is as follows: ① The Euclidean distance between each sample in the default bond set and other samples in that set was calculated, and a neighbor sample set was obtained for each sample based on the Euclidean distance. ② For each sample, several samples were randomly selected from the neighbor sample set. ③ New samples were generated according to the following formula:

$$x_{new} = x + rand(0,1) \cdot (x - x_m) , (5)$$

where x refers to the original sample in the default bond set, x_m is the random sample in the neighbor sample set of x, $rand(0,1)$ is a random number between 0 and 1, and x_{new} is the newly generated sample.

3.3 Experimental setup

The relevant parameters of the abnormal electricity detection and recognition algorithm proposed in this paper are shown in Table 2. Moreover, in order to verify the effectiveness of the proposed algorithm, it was compared with a single bidirectional LSTM and a single CNN. The relevant parameters of the two algorithms were consistent with those of the corresponding part of the proposed algorithm.

Table 2: Relevant parameters of the proposed algorithm.

	Structure	**Parameter**	**Structure**	**Parameter**
Bidirectional LSTM	Input layer	An input with 1,125 dimensions	Forward hidden layer	One layer, 256 nodes, the sigmoid activation function
	Backward hidden layer	One layer, 256 nodes, the sigmoid activation function	Output layer	An output with 256 dimensions
CNN	Input layer	An input with 7 × 161 dimensions	Convolution layer 1	Eight 3× 1 convolution kernels, a step length of 2
	Pooling layer 1	A pooling box with a specification of 3×1, a step length of 2	Convolution layer 2	Sixteen 3 × 1 convolution kernels, a step length of 1
	Pooling layer 2	A pooling box with a specification of 3×1, a step length of 1	Output layer	An output with 256 dimensions
AdaBoost	Weak classification learner	Linear predictor	Number of classifiers	50
	Error threshold	0.1		

In addition to training and testing the algorithm using the open dataset, this study also conducted a one-month field test in a local small-scale substation and compared the predicted results of the algorithm with the measured results of the substation. The above two single algorithms for comparison were also tested for a one-month field test.

3.4 Evaluation indicator

This paper used the precision, recall rate, false detection rate, and receiver operator characteristic (ROC) curve as the evaluation indicators of the detection performance of the algorithm. The formula of precision, recall rate, and false drop rate is:

$$\begin{cases} P = \dfrac{TP}{TP+FN} \\ R = \dfrac{TP}{TP+FP} \\ W = \dfrac{FP+FN}{TP+FP+TN+FN} \end{cases} , (6)$$

where P refers to precision, R refers to the recall rate, and W is the false drop rate.

The formula for calculating the true and false positive case rates of the ROC curve is:

$$\begin{cases} TPR = \dfrac{TP}{TP+FN} \\ FPR = \dfrac{FP}{TN+FP} \end{cases} , (7)$$

where TPR is the true positive rate and FPR is the false positive rate.

3.5 Test results

Figure 3 shows the performance of the three abnormal power user detection algorithms. It can be seen that there was little difference in precision, recall rate, and false detection rate between a single bidirectional LSTM and a single CNN. The AdaBoost recognition algorithm based on bidirectional LSTM+CNN had significantly better detection and recognition performance than the other two algorithms.

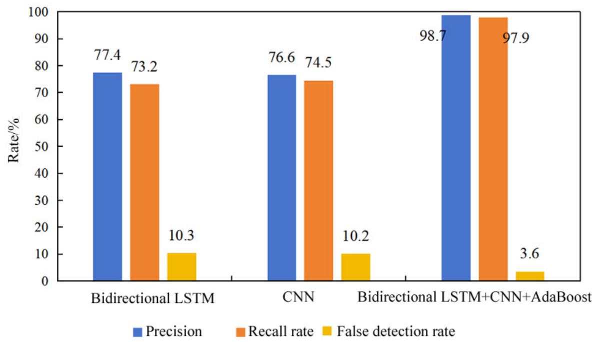

Figure 3: Performance of three abnormal user detection algorithms.

In addition to the above three evaluation indicators, the ROC curve was also used to evaluate the detection algorithm, and the final result is shown in Figure 4. The ROC curve reflects the positive and false detection rates of a prediction model under different recognition thresholds. When the prediction model cannot make predictions and can only give random results, its ROC curve is a diagonal line. When the prediction model is ideal, its ROC curve is a broken line coincident with axes $x=0$ and $y=1$. In other words, for a prediction model, the closer its ROC curve is to the diagonal line, the worse its prediction performance will be; the closer it is to the broken line, the better its prediction performance will be. It can be seen that the performance of a single bidirectional LSTM was close to that of a single CNN, while the bidirectional LSTM+CNN+AdaBoost detection algorithm had better prediction performance.

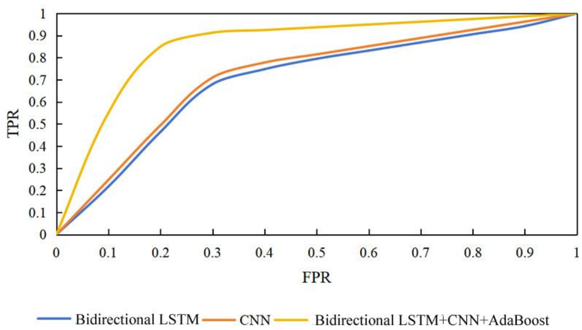

Figure 4: The ROC curves of three detection algorithms.

Finally, the three detection algorithms were deployed in a small substation for one month, and the changes in the false detection rates of the three detection algorithms during the field operation period are shown in Figure 5. With the passage of field operation time, the false detection rates of a single bidirectional LSTM and a single CNN gradually increased, and the CNN increased more. The proposed method had a stable false detection rate, which was lower than that of the other two detection algorithms.

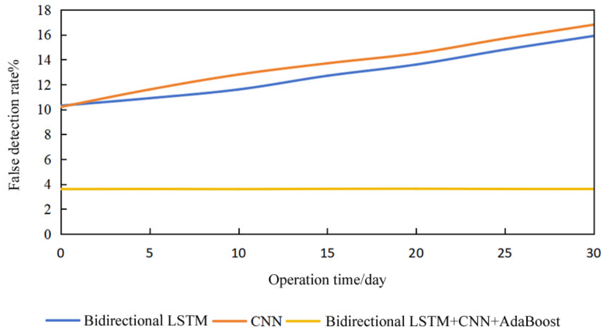

Figure 5: False detection rates of three detection algorithms in the one-month substation field operation.

4 Discussion

With the rapid development of technology, the smart grid, as an important component of modern power systems, is increasingly being emphasized for its stability and security. By integrating computation, communication, and physical environment, smart grids enable efficient management and monitoring of power networks. However, the complexity of smart grids also brings many challenges, especially in anomaly detection. Traditional detection methods often suffer from low accuracy and slow response, making it difficult to meet the requirements of modern power systems. The emergence of deep learning algorithms with data mining capabilities provides a new approach for processing and analyzing operational data in smart grids. As a deep learning model, the CNN has shown excellent performance in areas such as image processing and video analysis. In smart grids, the CNN can be applied to feature extraction and classification of power data. Key features can be extracted from massive amounts of power data through training the CNN model. The convolutional layer structure of the CNN possesses the ability to extract local features, enabling it to capture subtle variations in power data and improve the recognition capability for abnormal power data. The LSTM is a special type of recurrent neural network that can handle long sequence data, while the bidirectional LSTM further considers reverse sequences based on the LSTM. In the smart grid, a bidirectional LSTM can be used for processing and analyzing time series data. By capturing the temporal dependencies in power data, the bidirectional LSTM enables the detection of abnormal user behavior in the grid.

This paper combined a CNN and a bidirectional LSTM to analyze electricity consumption data from users in the smart grid and detect users with abnormal electricity usage.

Afterward, simulation experiments were conducted in the laboratory using power grid data collected from the State Grid Corporation of China. Subsequently, a one-month actual test was carried out in a local small-scale substation. This paper demonstrated that combining a CNN and a bidirectional LSTM yielded higher performance in identifying abnormal power grid users than using only a bidirectional LSTM or a CNN alone. During the actual operation process of the small-scale substation, our algorithm maintains lower and more stable false detection rates. The reason is that the CNN can extract local features of power data, while bidirectional LSTM can extract time series features of power data. Compared with a single bidirectional LSTM or a single CNN, it can obtain more comprehensive power data features. In addition, the AdaBoost model combines multiple weak classifiers into a strong classifier to further improve the recognition performance of the algorithm.

5 Conclusions

This paper briefly introduces the smart grid and the intelligent algorithm used to detect users with abnormal power consumption. This algorithm combined the bidirectional LSTM with the CNN to extract the features from the power consumption data of the users. Then, the AdaBoost model was employed to classify the users. A field operation test was carried out in a small substation. The proposed method was compared with a single bidirectional LSTM and a CNN. The precision, recall rate, and false detection rate were not much different between the single bidirectional LSTM and single CNN, and the performance of the proposed was obviously better than the other two algorithms. The ROC curves showed that the performance of the single bidirectional LSTM was close to that of the single CNN, while the predictive performance of the proposed detection algorithm was better. In the field operation, the false detection rate of the single bidirectional LSTM and single CNN gradually increased, and the CNN increased more; the proposed method had a more stable and smaller detection rate than the other two algorithms.

References

[1] Khan A R, Mahmood A, Safdar A, Khan ZA, Khan NA (2016). Load forecasting, dynamic pricing and DSM in smart grid: A review. *Renewable & Sustainable Energy Reviews*, 54(FEB.), pp. 1311-1322. https://doi.org/10.1016/j.rser.2015.10.117

[2] Collier SE (2017). The Emerging Enernet: Convergence of the Smart Grid with the Internet of Things. *IEEE Industry Applications Magazine*, 23(2), pp. 12-16. https://doi.org/10.1109/MIAS.2016.2600737

[3] Shan X, Dai Z, Zhang Z, Zhang Y, Bi X, Ge C (2015). Research on and Application of Integrated Smart Alarm Based on Smart Grid Dispatching and Control Systems. *Dianli Xitong Zidonghua/Automation of Electric Power Systems*, 39(1), pp. 65-72. https://doi.org/10.7500/AEPS20140930016

[4] Wang X, Liang Q, Mu J, Wang W, Zhang B (2015). Physical layer security in wireless smart grid. *Security & Communication Networks*, 8(14), pp. 2431-2439. https://doi.org/10.1002/sec.751

[5] Gams M, Kolenik T (2021). Relations between Electronics, Artificial Intelligence and Information Society through Information Society Rules. *Electronics*, 10(4), pp. 1-16. https://doi.org/10.3390/electronics10040514

[6] Deng R, Lu R, Xiao G, Chen J (2015). Fast Distributed Demand Response With Spatially and Temporally Coupled Constraints in Smart Grid. *IEEE Transactions on Industrial Informatics*, 11(6), pp. 1597-1606. https://doi.org/10.1109/TII.2015.2408455

[7] Jo H J, Kim I S, Lee D H (2016). Efficient and Privacy-Preserving Metering Protocols for Smart Grid Systems. *IEEE Transactions on Smart Grid*, 7(3), pp. 1732-1742. https://doi.org/10.1109/TSG.2015.2449278

[8] Grilo A, Casaca A, Nunes M, Bernardo A, Rodrigues P, Almeida JP (2017). A management system for low voltage grids. *2017 IEEE Manchester PowerTech*, IEEE, Manchester, UK. https://doi.org/10.1109/PTC.2017.7980826

[9] Gholian A, Mohsenian-Rad H, Hua Y (2016). Optimal Industrial Load Control in Smart Grid. *IEEE Transactions on Smart Grid*, 7(5), pp. 2305-2316. https://doi.org/10.1109/TSG.2015.2468577

[10] Li Q, Zhu Y, Ding J, Li W, Sun W, Ding L. (2024). Deep Reinforcement Learning Based Resource Allocation for Fault Detection with Cloud Edge Collaboration in Smart Grid. *CSEE Journal of Power and Energy Systems*, 10(3), pp. 1220-1230. https://doi.org/10.17775/CSEEJPES.2021.02390

[11] Good N, Ellis K A, Mancarella P (2017). Review and classification of barriers and enablers of demand response in the smart grid. *Renewable & Sustainable Energy Reviews*, 72, pp. 57-72. https://doi.org/10.1016/j.rser.2017.01.043

[12] Chen K L, Chen Y R, Tsai Y P, Chen N (2017). A Novel Wireless Multifunctional Electronic Current Transformer Based on ZigBee-Based Communication. *IEEE Transactions on Smart Grid*, 8(4), pp. 1888-1897. https://doi.org/10.1109/TSG.2015.2510325

[13] Luo X (2017). Design and Implementation of Wireless Voltage Monitoring System Based on Zigbee. *Chinese Journal of Electron Devices*, 12(3), pp. 83-96. https://doi.org/10.4018/IJITWE.2017070108

[14] Sharma J, Bhattacharya PP, Jha MK (2016). Performance Analysis of IEEE 802.15.4 in Different NLOS Environments of Smart Grids. *International Journal of Sensors Wireless Communications & Control*, 2016, 6(3):216-220. https://doi.org/10.2174/221032790666616092812190 6

[15] Gao Y, Zhang J, Chen Z, Chen X (2024). Anomaly Detection Algorithm of Power System Based on Graph Structure and Anomaly Attention. *Computers, Materials & Continua*, 79(4), pp. 493-507. https://doi.org/10.32604/cmc.2024.048615

[16] Syu J H, Srivastava G, Fojcik M, Cupek R, Lin JCW (2023). Energy grid management system with anomaly detection and Q-learning decision modules. *Computers and Electrical Engineering*, 2023, 107, pp. 108639-1-108639-11. https://doi.org/10.1016/j.compeleceng.2023.108639