Discrete & Continuous Models
& Applied Computational Science
ISSN 2658-7149 (online), 2658-4670 (print)

2022, 30 (3) 258–268
http://journals.rudn.ru/miph

Research article

UDC 519.6
PACS 07.05.Tp
DOI: 10.22363/2658-4670-2022-30-3-258-268

Detection of cyber-attacks on the power smart grids using semi-supervised deep learning models

Eugeny Yu. Shchetinin[1], **Tatyana R. Velieva**[2]

[1] *Financial University under the Government of Russian Federation,*
49, Leningradsky Prospect, Moscow, 125993, Russian Federation
[2] *Peoples' Friendship University of Russia (RUDN University),*
6, Miklukho-Maklaya St., Moscow, 117198, Russian Federation

(received: July 15, 2022; revised: July 21, 2022; accepted: August 8, 2022)

Abstract. Modern smart energy grids combine advanced information and communication technologies into traditional energy systems for a more efficient and sustainable supply of electricity, which creates vulnerabilities in their security systems that can be used by attackers to conduct cyber-attacks that cause serious consequences, such as massive power outages and infrastructure damage. Existing machine learning methods for detecting cyber-attacks in intelligent energy networks mainly use classical classification algorithms, which require data markup, which is sometimes difficult, if not impossible. This article presents a new method for detecting cyber-attacks in intelligent energy networks based on weak machine learning methods for detecting anomalies. Semi-supervised anomaly detection uses only instances of normal events to train detection models, which makes it suitable for searching for unknown attack events. A number of popular methods for detecting anomalies with semi-supervised algorithms were investigated in study using publicly available data sets on cyber-attacks on power systems to determine the most effective ones. A performance comparison with popular controlled algorithms shows that semi-controlled algorithms are more capable of detecting attack events than controlled algorithms. Our results also show that the performance of semi-supervised anomaly detection algorithms can be further improved by enhancing deep autoencoder model.

Key words and phrases: smart energy grids, cyber-attacks, semi-supervised anomaly detection, deep learning, autoencoder

1. Introduction

There are many problems in traditional power grids, such as the lack of automated analysis and situational awareness, poor visibility and slow response time, which makes them unable to meet the significantly increased demand and consumption of electricity in the 21st century [1]. With the help of modern information and communication technologies, intelligent networks provide a bidirectional flow of electricity and information, which ensures a more efficient and stable supply of electricity and better demand management [2, 3]. The intelligent energy network consists of four main components: generation, transmission, distribution and consumption, which are connected through a three-level hierarchical structured communication network [4] (see figure 1). The first level of the communication network is the home network, which is responsible for communication at the consumption stage to connect smart devices in consumers' homes to the smart grid with smart meters for more efficient energy management and demand response. The second level of the communication network, the district network, is responsible for communication at the distribution stage, which collects data from smart meters and sends back control commands for advanced accounting applications.

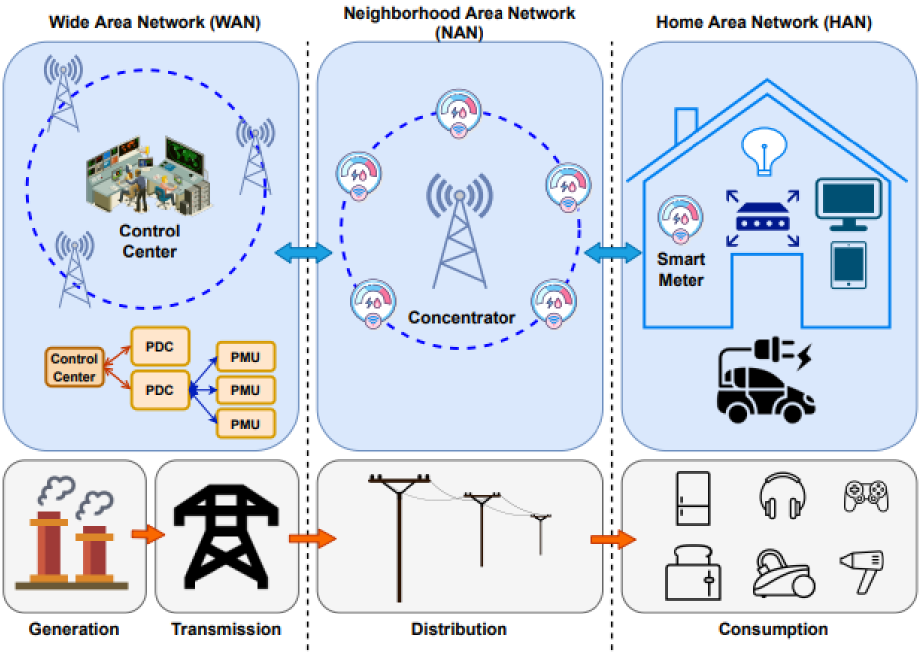

Figure 1. Diagram of smart grid energy consumption system

At the last level, the global network connects with utility management centers, forming the basis of an intelligent network for the communication needs of the stages of electricity production and transmission. Although the integration of advanced ICTs into traditional power grids brings significant benefits for the delivery and management of electricity, it also creates new vulnerabilities in security systems [5]. Cyber-attacks can target any of the four components of a smart grid — from smart home gateways in HAN to control rooms in the global network [6].

In this work, we used measurements from the Power Measurement Units (PMU) to detect cyber-attacks. PMU is a sensor device deployed at the global network level of the smart grid network, which provides real-time measurements of the state of the power system for a wide range of monitoring, protection and control. In the Global Monitoring System (GMS), several

PMUs are connected to a phasor data hub. The GMS central authority then collects information from the PDCS. PMU measurements combine both physical and cyber domains, making them a suitable choice for detecting cyber-attacks targeting the physical domain of an intelligent network, such as False Data Entry (FDE) attacks and malicious shutdown attacks.

In most widely used models have been built to detect cyber-attacks in intelligent networks using controlled learning algorithms. To train supervised algorithms, both normal and attacking data are required. However, collecting representative instances of various attack events is usually a difficult task, if not impossible, which can lead to poor model performance when detecting certain attacks, especially types of attacks not represented in the training data.

In this article, we proposed a method for detecting cyber-attacks in power smart grids with semi-supervised anomaly detection. Unlike supervised models, semi-supervised anomaly detection algorithms use only data from normal events to train a detection model that is capable of detecting unknown types of attacks. We have investigated a number of anomaly detection algorithms and identified the most effective ones for detecting cyber-attacks in smart energy grids. The performance of semi-supervised algorithms was compared with the characteristics of popular supervised algorithms to show their superiority in finding attack events. We have also supplemented semi-supervised anomaly detection with deep learning to extract features to further improve attack detection performance.

2. Related work

Traditional approaches use PMU measurements to assess the state of the power system and compare the difference between the observed and estimated measurements with a threshold for detecting cyber-attacks. A lightweight scheme was proposed in the paper [4], which explores the spatial-temporal correlations between network state estimates and applies confidence voting to detect abnormal state estimates in intelligent networks caused by real-time FDI attacks.

Recently, machine learning has been widely used to detect cyber-attacks in smart grids, where most of the proposed approaches are based on supervised learning algorithms. In the paper [7] a number of supervised learning algorithms were investigated for recognizing violations in the power system and cyber-attacks. A One-Class SVM (OC-SVM) was used in [8] to create an intrusion detection module for detecting malicious attacks in a dispatch control system and data collection system using network traces. The paper [9] applied several popular supervised algorithms, including perceptron, k-nearest neighbor (KNN), support vector machines (SVM) and logistic regression (LR) with ensemble training and combining functions at the object level to predict FDI attacks. Their experimental results demonstrate that machine learning algorithms are superior to state-based algorithms. Singh [10] compared SVM, KNN models for detecting both direct and covert attacks in intelligent networks. Compared to number of methods based on supervised learning algorithms, only a limited number of studies have been conducted on the use of unsupervised anomaly detection algorithms to detect attacks in intelligent networks.

The data sets studied in our study were created based on the structure of the power system, consisting of intelligent electronic devices, dispatch control systems and network monitoring devices. The power system framework can simulate several operational scenarios to generate data corresponding to three types of events: absence of events, natural events, and attack events. The six types of events are described as follows:

— No event: Normal readings.
— Short circuit: There has been a single ground fault, which can be determined by reading the percentage range in the data.
— Line maintenance: Operators switch one or more IEDs to perform maintenance on certain parts of the power system and its components.
— Remote Shutdown Command Implementation Attack: Attackers can send commands that switch Improvised Explosive devices to switch switches when they can get into the system.
— Attack with changing relay settings: Attackers change the settings, for example, disable the main functions of the settings, as a result of which the IEDs do not switch the switches whenever an acceptable error or command occurs.
— Data Intrusion attack: Attackers modify PMU measurements such as voltage, current, and sequence components to simulate a real malfunction resulting in the disconnection of switches.

The system has four PMUs integrated with relays, where each PMI measures 29 features. A total of 116 functions were obtained from four PMUs. Depending on how to group the scenarios, three groups of datasets were created based on the generated data — binary class data and multi-class data from the framework. Since the purpose of our study was to distinguish attack events from other types of events, we adopted a binary group of datasets in which no events and normal events are treated as ordinary events.

3. Methodology

The proposed method contains two main components: deep representation learning and semi-supervised anomaly detection [11–15]. The first step of the proposed method is to prepare a training dataset that contains only examples of normal events. The dimension of the object space is then reduced by deep representation learning, when a low-dimensional hidden representation is extracted from the input data using a deep autoencoder. Finally, a semi-supervised anomaly detection algorithm using the representation of the studied features is used to train the detection model. At the detection stage, a hidden representation is first created from an unknown input instance by the deep autoencoder, which is then fed into the trained detection model to classify the instance as a normal event or an attack event.

In our study, deep autoencoder is used to extract features, which used to learn robust low-dimensional representations from multidimensional input data. PCA, a popular feature extraction method was used as a method for comparison. After training the autoencoder with the training dataset, the encoder and code layer are retained for feature extraction, while the decoder has been removed from the network. The hidden code-level representation will be used as input for a semi-supervised anomaly detection algorithm.

4. Performance evaluation and results

The binary group of data sets on attacks on power systems adopted in our study contains 15 data sets covering 37 scenarios [16–18]. We used min–max normalization to normalize the data. The parameters used by the algorithms investigated in our study are listed in the table 1. The characteristics of semi-supervised anomaly detection algorithms using all 116 PMU functions were investigated first. Among the usual instances of the dataset, 50% were randomly selected to train the detection algorithm. The remaining 50% of normal instances and all attack instances were then used for testing. Examples of ROC curves obtained using eight algorithms on datasets 1 and 11 at one stage of the experiment are shown in the figure 2. The average AUC values of the algorithms calculated from the results of 10 runs for each of the 15 data sets are shown in the figure 2.

Table 1

Parameters used by the semi-supervised and supervised algorithms investigated in our study

Model	Parameters
OCSVM	RBF kernel, degree=3
LOF	K=25
KNN	K=10
IForest	iTrees=100
SVM	RBF kernel, degree=3, C=10
Deep Autoencoder	Batch_size=8

Figure 2 shows that three most effective algorithms in terms of average AUC are OCSVM, KNN and IForest. These algorithms show significantly better performance than the other four algorithms. Then we used the distance to the angle d to determine the detection threshold of the algorithm to obtain accuracy. The three best algorithms in terms of F1 average score are OCSVM, KNN and iForest.

In our study, two metrics were used to evaluate the effectiveness: the area under the ROC curve (AUC) and the F1 score. The ROC curve shows the relationship between the true positive rate (TPR) and the false positive rate (FPR) by changing the detection threshold. Equations (1) and (2) define TPR and FPR, where TP, TN, FP, and FN are true positive, true negative, false positive and false negative, respectively:

$$R = \frac{\mathrm{TP}}{(\mathrm{TP} + \mathrm{FN})}, \tag{1}$$

$$R = \frac{\mathrm{FP}}{(\mathrm{FP} + \mathrm{TN})}. \tag{2}$$

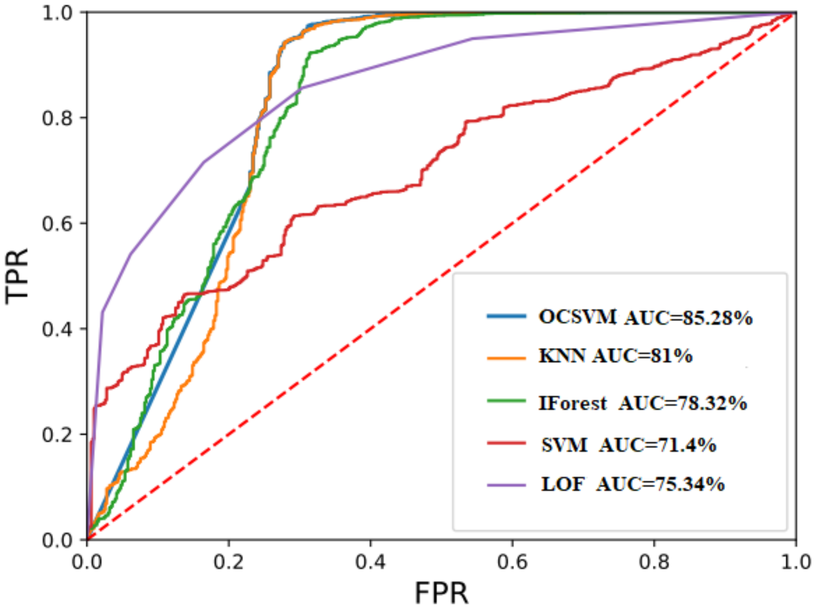

Figure 2. Examples of ROC curves obtained by semi-supervised and supervised algorithms

AUC measures the area under the ROC curve to indicate the performance of the model on distinguishing normal and attack events. A higher AUC value means that model has a better capability to distinguish normal and attack events. F1 score is defined as the harmonic mean of the precision and recall:

$$\text{Precision} = \frac{\text{TP}}{(\text{TP} + \text{FP})}, \tag{3}$$

$$\text{Recall} = \frac{\text{TP}}{(\text{TP} + \text{FN})}, \tag{4}$$

$$\text{F1} - \text{score} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{(\text{Precision} + \text{Recall})}. \tag{5}$$

We adopted the distance to corner (D) as the criterion, which determines the optimal threshold as the point on the ROC curve closest to the corner point $(0, 1)$:

$$D = \sqrt{(1 - \text{TPR})^2 + \text{FPR}^2}. \tag{6}$$

Then we compared three most effective semi-supervised algorithms in terms of AUC (OCSVM, LOF, IForest) with two popular controlled algorithms — KNN and SVM, which were used to detect cyber-attacks in power smart grids. For each of the 15 datasets, a training dataset for semi-supervised algorithms was generated by randomly selecting 50% of the regular instances. These regular instances were combined with the same number of randomly selected attack instances to form a training dataset for controlled algorithms. The remaining 50% of normal instances and attack instances were used to form a test dataset for both semi-supervised and supervised algorithms.

Examples of ROC curves are shown in figures 2, which were obtained using five unsupervised and supervised algorithms. Examples of F1-scores are shown in figures 3, which were obtained using five unsupervised and supervised algorithms. Among all the algorithms, SVM had the worst performance, while the KNN algorithm has significantly better average AUC than other algorithms. The good performance of the controlled KNN algorithm in terms of AUC is due to its significantly better TPR compared to the three semi-supervised algorithms when the FPR is low. On the other hand, one of figure 2 also shows that as the FPR increases, the three semi-controlled algorithms can achieve a high TPR much faster than the controlled KNN algorithm.

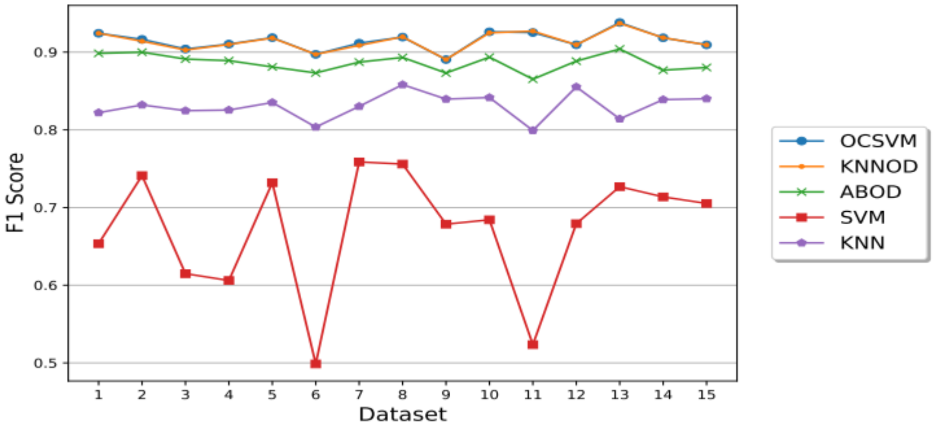

Figure 3. Performance comparison of the semi-supervised algorithms with supervised algorithms in terms of average precision, recall, and F1 score

5. Performance improvement using a deep autoencoder

Finally, the impact of DAE-based deep representation training on the performance of semi-supervised anomaly detection algorithms for detecting cyber-attacks in power smart grids was investigated. The PCA method has been adopted as a reference method for comparison. We set the number of extracted objects to 30 for both DAE and PCA. The input and hidden encoder layers in the DAE have 116 and 60 nodes, respectively. The three most efficient semi-supervised algorithms OCSVM, KNN and IForest in terms of AUC, F1-score accuracy metrics were included in this study. The results in terms of the average AUC are shown in figure 4, which were obtained by averaging the results of all runs of all 15 datasets. Figures 4, 5 show that DAE can further improve the performance of three semi-controlled algorithms in terms of both performance indicators. Statistical tests (paired t-test with two samples, $\alpha = 0.05$) show that the AUC obtained using three semi-controlled algorithms with deep learning representation is significantly higher than when using all functions. The F1 scores obtained by OCSVM and KNN using deep representation learning are also significantly higher than when using all functions.

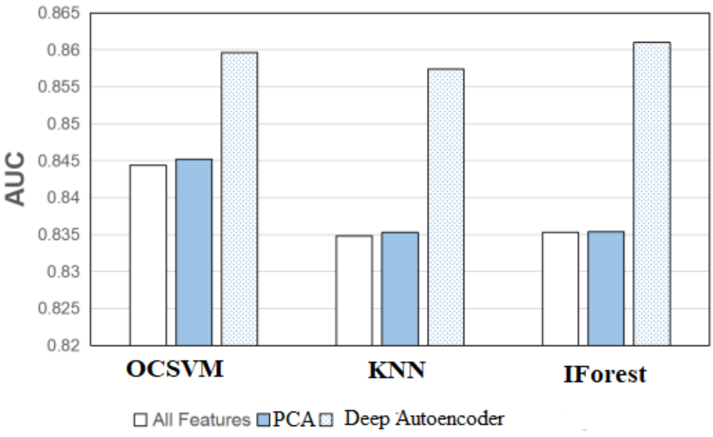

Figure 4. Performance of semi-supervised algorithms with and without feature extraction in terms of average AUC

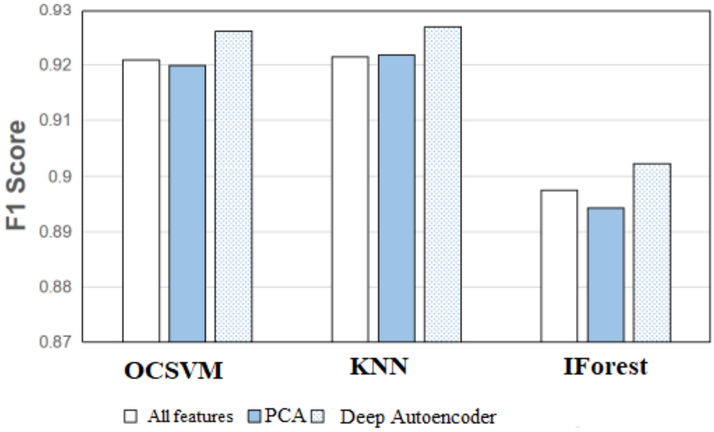

Figure 5. Performance of semi-supervised algorithms with and without feature extraction in terms of average F1 score

6. Conclusion

Smart grids ensure efficient supply of electricity to various facilities and its management through the introduction of advanced digital technologies into traditional power grids. On the other hand, the vulnerabilities that have appeared in their security can be used to carry out cyber-attacks that lead to devastating damage. Using PMU measurements that connect the physical and cybernetic domains, the article develops a method based on semi-controlled anomaly detection and deep learning to detect cyber-attacks in smart energy grids. Unlike supervised algorithms, semi-supervised anomaly detection algorithms use only instances of normal events to train detection models, which makes them capable of detecting events of unknown types of attacks.

In our experiments, the most effective semi-supervised algorithms were identified using publicly available datasets on attacks on intelligent energy systems. A comparison of performance with popular controlled algorithms has shown that semi-supervised algorithms have a better ability to detect cyber-attacks. In addition, our results showed that the detection performance of semi-supervised algorithms can be further enhanced by deep representation training based on DAE.

References

[1] G. Dileep, "A survey on Smart Grid technologies and applications," *Renewable Energy*, vol. 146, pp. 2589–2625, 2020. DOI: 10.1016/j.renene.2019.08.092.

[2] V. C. Gungor, D. Sahin, T. Kocak, S. Ergut, C. Buccella, C. Cecati, and G. P. Hancke, "Smart Grid technologies: communication technologies and standards," *IEEE Transactions on Industrial Informatics*, vol. 7, no. 4, pp. 529–539, 2011. DOI: 10.1109/TII.2011.2166794.

[3] T. Flick and J. Morehouse, *Securing the Smart Grid: Next Generation Power Grid Security.* Syngress, 2010.

[4] S. Aftergood, "Cybersecurity: the cold war online," *Nature*, vol. 547, no. 7661, pp. 30–31, Jul. 2017. DOI: 10.1038/547030a.

[5] C. Chio and D. Freeman, *Machine learning and security: protecting systems with data and algorithms.* O'Reilly Media, 2018.

[6] D. S. Berman, A. L. Buczak, J. S. Chavis, and C. L. Corbett, "A survey of deep learning methods for cyber security," *Information*, vol. 10, no. 4, 2019. DOI: 10.3390/info10040122.

[7] D. Wang, X. Wang, Y. Zhang, and L. Jin, "Detection of power grid disturbances and cyber-attacks based on machine learning," *Journal of Information Security and Applications*, vol. 46, pp. 42–52, 2019. DOI: 10.1016/j.jisa.2019.02.008.

[8] S. Ahmed, Y.-D. Lee, S.-H. Hyun, and I. Koo, "Unsupervised machine learning-based detection of covert data integrity assault in Smart Grid networks utilizing isolation forest," *IEEE Transactions on Information Forensics and Security*, vol. 14, pp. 2765–2777, 2019.

[9] M. Ozay *et al.*, "Machine learning methods for attack detection in the Smart Grid," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 27, pp. 1773–1786, 2016.

[10] V. K. Singh and M. Govindarasu, "Decision tree based anomaly detection for remedial action scheme in Smart Grid using PMU data," in *IEEE Power & Energy Society General Meeting PESGM*, 2018, pp. 1–5. DOI: 10.1109/PESGM.2018.8586159.

[11] G. Pang, C. Shen, L. Cao, and A. V. D. Hengel, "Deep learning for anomaly detection: a review," *ACM Comput. Surv.*, vol. 54, no. 2, 2021. DOI: 10.1145/3439950.

[12] Z. E. Huma, S. Latif, J. Ahmad, Z. Idrees, A. Ibrar, Z. Zou, F. Alqahtani, and F. A. Baothman, "A hybrid deep random neural network for cyberattack detection in the Industrial Internet of Things," *IEEE Access*, vol. 9, pp. 55 595–55 605, 2021. DOI: 10.1109/ACCESS.2021.3071766.

[13] M. S. Minhas and J. Zelek, "Semi-supervised anomaly detection using autoencoders," *Journal of Computational Vision and Imaging Systems*, vol. 5, no. 1, p. 3, 2019.

[14] M. Wieler. "Weakly supervised learning for industrial optical inspection." (2007), [Online]. Available: https://hci.iwr.uni-heidelberg.de/node/3616.

[15] R. Qi, C. Rasband, J. Zheng, and R. Longoria, "Semi-supervised outlier detection and deep feature extraction for detecting cyber-attacks in Smart Grids using PMU data," *Advances in Intelligent Systems and Computing*, vol. 1134, pp. 509–515, 2020. DOI: 10.1007/978-3-030-43020-7_67.

[16] E. Y. Shchetinin, "On methods of quantitative analysis of the company's financial indicators under conditions of high risk of investments," *Discrete and Continuous Models and Applied Computational Science*, vol. 28, no. 4, pp. 346–360, 2020. DOI: 10.22363/2658-4670-2020-28-4-346-360.

[17] E. Y. Shchetinin, "Modeling the energy consumption of smart buildings using artificial intelligence," in *CEUR Workshop Proceedings*, vol. 2407, 2019, pp. 130–140.

[18] E. Y. Shchetinin, "Development of Energy Saving Technologies for Smart Buildings by Using Computer Algebra," *Programming and Computer Software*, vol. 46, pp. 324–329, 2020. DOI: 10.1134/S0361768820050084.

For citation:

E. Y. Shchetinin, T. R. Velieva, Detection of cyber-attacks on the power smart grids using semi-supervised deep learning models, Discrete and Continuous Models and Applied Computational Science 30 (3) (2022) 258–268. DOI: 10.22363/2658-4670-2022-30-3-258-268.

Information about the authors:

Shchetinin, Eugeny Yu. — Doctor of Physical and Mathematical Sciences, Lecturer of Department of Mathematics, Financial University under the Government of Russian Federation (e-mail: riviera-molto@mail.ru, ORCID: https://orcid.org/0000-0003-3651-7629)

Velieva, Tatyana R. — Candidate of Sciences in Physics and Mathematics, Senior lecturer of Department of Applied Probability and Informatics of Peoples' Friendship University of Russia (RUDN University) (e-mail: velieva-tr@rudn.ru, phone: +7(495)9520250, ORCID: https://orcid.org/0000-0003-4466-8531, ResearcherID: Q-6304-2016, Scopus Author ID: 56695390200)

УДК 519.6
PACS 07.05.Tp
DOI: 10.22363/2658-4670-2022-30-3-258-268

Обнаружение кибератак на интеллектуальные энергосистемы с использованием неконтролируемых моделей глубокого обучения

Е. Ю. Щетинин[1], **Т. Р. Велиева**[2]

[1] *Финансовый университет при Правительстве Российской Федерации,*
Ленинградский проспект, д. 49, Москва, 125993, Россия
[2] *Российский университет дружбы народов,*
ул. Миклухо-Маклая, д. 6, Москва, 117198, Россия

Аннотация. Современные интеллектуальные энергосети объединяют передовые информационные и коммуникационные технологии в традиционные энергосистемы для более эффективного и устойчивого снабжения электроэнергией, что создаёт уязвимости в их системах безопасности, которые могут быть использованы злоумышленниками для проведения кибератак, вызывающих серьезные последствия, такие как массовые перебои в подаче электроэнергии и повреждение инфраструктуры. Существующие методы машинного обучения для обнаружения кибератак в интеллектуальных энергетических сетях в основном используют классические алгоритмы классификации, которые требуют разметки данных, что иногда сложно, а то и невозможно. В данной статье представлен новый метод обнаружения кибератак в интеллектуальных энергетических сетях, основанный на слабых методах машинного обучения для обнаружения аномалий. Полуконтролируемое обнаружение аномалий использует только экземпляры обычных событий для обучения моделей обнаружения, что делает его подходящим для поиска неизвестных событий атак. В ходе исследования был проанализирован ряд популярных методов обнаружения аномалий с полууправляемыми алгоритмами с использованием общедоступных наборов данных о кибератаках на энергосистемы для определения наиболее эффективных из них. Сравнение производительности с популярными управляемыми алгоритмами показывает, что полууправляемые алгоритмы лучше способны обнаруживать события атак, чем управляемые алгоритмы. Наши результаты также показывают, что производительность полуконтролируемых алгоритмов обнаружения аномалий может быть дополнительно улучшена за счёт усовершенствования модели глубокого автоэнкодера.

Ключевые слова: интеллектуальные энергетические сети, кибератаки, частично контролируемое обнаружение аномалий, глубокое обучение, автоэнкодер