Operational Data Based Intrusion Detection System for Smart Grid

Georgios Efstathopoulos*, Panagiotis Radoglou Grammatikis†, Panagiotis Sarigiannidis†, Vasilis Argyriou, Antonios Sarigiannidis‡, Konstantinos Stamatakis¶, Michail K. Angelopoulos§ and Solon K. Athanasopoulos¶

***Abstract*—With the rapid progression of Information and Communication Technology (ICT) and especially of Internet of Things (IoT), the conventional electrical grid is transformed into a new intelligent paradigm, known as Smart Grid (SG). SG provides significant benefits both for utility companies and energy consumers such as the two-way communication (both electricity and information), distributed generation, remote monitoring, self-healing and pervasive control. However, at the same time, this dependence introduces new security challenges, since SG inherits the vulnerabilities of multiple heterogeneous, co-existing legacy and smart technologies, such as IoT and Industrial Control Systems (ICS). An effective countermeasure against the various cyberthreats in SG is the Intrusion Detection System (IDS), informing the operator timely about the possible cyberattacks and anomalies. In this paper, we provide an anomaly-based IDS especially designed for SG utilising operational data from a real power plant. In particular, many machine learning and deep learning models were deployed, introducing novel parameters and feature representations in a comparative study. The evaluation analysis demonstrated the efficacy of the proposed IDS and the improvement due to the suggested complex data representation.**

***Index Terms*—Anomaly detection, Cybersecurity, Intrusion Detection System, Machine learning, Operational Data, Smart Grid**

I. Introduction

The next generation of the electrical grid known as Smart Grid (SG) will address multiple challenges of the existing conventional one, such as centralised generation, one-way communication (only electricity transmission), limited control and manual restoration. In particular, an integral part of SG is the convergence of the Information and Communication Technology (ICT) services and especially the Internet of Things (IoT) with the electrical engineering, thus providing two-way communication (both electricity and information), increased reliability, self-healing, remote monitoring, as well as better utilization of the renewable resources [1]–[3]. According to [1], SG is going to be the largest paradigm of the IoT technology called Enernet, combining multiple sensors and communication protocols. Even though SG offers numerous benefits, at the same time, it introduces numerous and significant challenges, especially concerning the security domain due to its heterogeneous nature.

Concerning the cybersecurity aspect, SG inherits the vulnerabilities of the various technologies it integrates. More specifically, the operation of the existing electrical grid is largely based on the Supervisory Control and Data Acquisition (SCADA) systems that monitor and control various processes during the electricity transmission and distribution. However, these systems utilise legacy industrial protocols such as Modbus, IEC-60870-5-104, Profinet and Distributed Network Protocol (DNP3) that are characterised by severe cybersecurity flaws, since they do not involve authentication and authorisation mechanisms [4], [5]. On the other hand, the advent of IoT generates crucial security concerns since it is based on the Internet, which is insecure by its nature [6]. Moreover, it combines novel technologies such as Wireless Sensor Networks (WSNs) that bring the corresponding cybersecurity issues, such as sinkhole, sybil and wormhole cyberattacks. Finally, the ability of various objects to interact with each other and their physical environment without any human intervention and control increases the security and privacy concerns.

In general, the cyberattacks against SG target the *Confidentiality*, *Integrity* and *Availability* (CIA) of the involved systems and communications. In particular, Man-in-The-Middle (MiTM) and False Data Injection (FDI) attacks violate the confidentiality and integrity, respectively. On the other hand, the various Denial of Service (DoS) attacks threaten the availability principle. Furthermore, a more dangerous category of cyberthreats against SG is the Advanced Persistent Threat (APT) which is organised by security specialists for a long period pursuing a specific goal. A characteristic APT against SG was the cyberattack against an Unkranian electric substation resulting in the power blackout for more than 225,000 people [4]. Moreover, another characteristic example was the Stuxnet worm against the Iranian nuclear programme, which exploited four zero-days vulnerabilities [4]. Also, in 2009, various reconnaissance cyberattacks were performed against the US electrical grid by Chinese and Russian cyberattackers [4]. Similarly, in 2014 a campaign of various cyberattacks called Dragonfly [4], attempted to violate the electrical grid infrastructures of many countries such as the US, France, Germany, Italy, Poland, Spain and Turkey.

Anticipating the cybersecurity issues of SG, both academia

* G. Efstathopoulos is with the 0INF, Imperial Offices, London, UK, E6 2JG - E-Mail: george@0inf.com

† P. Radoglou Grammatikis and P. Sarigiannidis are with the Department of Electrical and Computer Engineering, University of Western Macedonia, Kozani 50100, Greece - E-Mail: {pradoglou, psarigiannidis}@uowm.gr

‡ A. Sarigiannidis is with Sidroco Holdings Ltd, 3113, Limassol, Cyprus - E-Mail: asarigia@sidroco.com

§ M. Angelopoulos is both with the Department of Economics, University of Piraeus and the Testing Research & Standards Center / Public Power Corporation SA, Leontariou 9, Kantza, 15351, Athens, Attica - E-Mail: {m.angelopoulos}@dei.com.gr

¶ K. Stamatakis, and S. Athanasopoulos are with the Testing Research & Standards Center / Public Power Corporation SA, Leontariou 9, Kantza, 15351, Athens, Attica - E-Mail: {K.Stamatakis, m.angelopoulos, S.Athanasopoulos}@dei.com.gr

and industry have provided proper authentication and access control mechanisms to enhance the existing protection solutions. For instance, the IEC 62351 standard provides solutions, thereby addressing the security gaps of many industrial protocols. However, many manufacturers and vendors cannot implement or integrate these solutions. Also, such solutions may not prevent various threats such as FDI and DoS cyberattacks. An effective countermeasure against cyberthreats is to apply an Intrusion Detection System (IDS) which audits and analyses security information to detect timely possible malicious security violations. A significant benefit of these systems is that they can detect zero-day cyberattacks or unknown anomalies by adopting Artificial Intelligence (AI) mechanisms.

In this paper, a novel IDS is proposed capable of detecting possible anomalies in a power plant, by utilising real-life operational data in terms of temperature values. In particular, the proposed IDS adopts many supervised learning techniques based on a dataset coming from a large scale power plant in the area of Lavrio, Greece. The power plant belongs to the Public Power Corporation (PPC) premises, who is the main power generator in Greece. The outline architecture of this plant is depicted in Fig. 1. The introduced IDS extracts security events, thus informing the system operator or/and the security administrator about possible anomalies coming from known or unknown cyberattacks or system's disturbances.

The rest of the paper is organised as follows: Various relevant works are discussed on Section II. In Section III, we provide a background about the IDS systems. Section IV presents and analyses the architecture of the proposed IDS. Section V is devoted to the evaluation process, which demonstrates the efficacy of our IDS. Finally, Section VI provides the concluding remarks of this paper, by summarising its main contribution and providing directions for future work.

II. Related Work

Relevant works providing IDS systems for SG are discusses on this section. Each paragraph is devoted to a separate IDS by summarising its architecture and performance.

In [7], A. Patel et al. proposed an anomaly-based IDS relying on a Support Vector Machine (SVM), an Ontology Knowledge Base (OKB) and a fuzzy analyser. In particular, this IDS can monitor the entire SG ecosystem and consists of numerous Host-Based IDS (HIDS) and Network-Based IDS (NIDS) agents that each of them applies an SVM model which was trained by combining records from the KDD CUP 1999 dataset and experiments carried out by the authors. Moreover, in order to reduce the false positives generated by the previous SVM model, a fuzzy logic technique was adopted capable of determining a risk value between 0 and 1 for each entity of SG. Finally, an OKB was used to identify the target of the possible attacks. Based on the evaluation process, the Area Under Curve (AUC) reaches 0.994.

In [8], the authors presented an IDS for the Advanced Metering Infrastructure (AMI) consisting of three units that monitor the network traffic generated by smart meters, data collectors and the AMI headend respectively. Concerning the detection process, the algorithm evaluates seven machine learning algorithms by using both KDD CUP 1993 and NSL-KDD datasets. The algorithms evaluated are: 1) Single Classifier Drift, 2) Bagging using Adaptive-Size Hoeffding Tree, 3) Bagging using ADWIN, 4) Limited Attribute Classifier, 5) Leveraging Bagging, 6) Active Classifier, 7) Accuracy Updated Ensemble. Based on the experimental results, the Single Classifier Drift and the Active Classifier are suggested for the smart meters, the Leveraging Baggin for the data collectors while the Active Classifier for the AMI headends. Also it is worth mentioning the relevant work in [9]–[13].

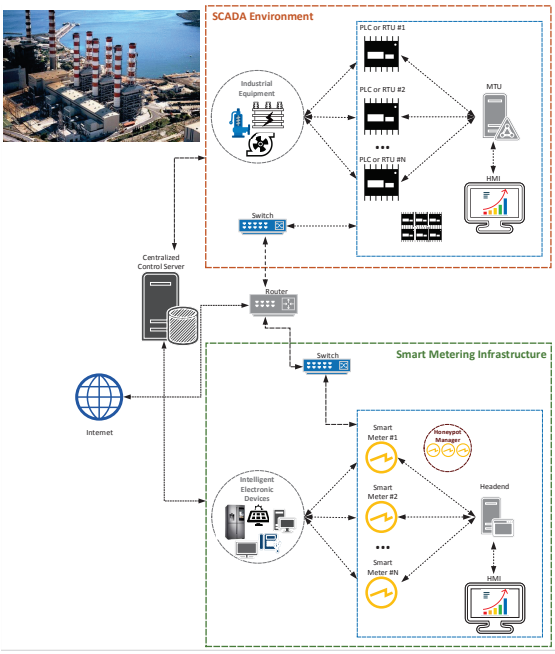

Fig. 1: PPC Power Plant - Lavrio, Greece Unit 5.

In [14], the authors developed an anomaly-based IDS for AMI, which monitors and controls the bidirectional Transmission Control Protocol/Internet Protocol (TCP/IP) network flows, which are aggregated periodically in the data collector component. The proposed IDS consists of four modules, namely 1) the Network Monitoring Module, 2) the Network Flow Extraction Module, 3) the Analysis Engine Module and 4) the Response Module. Regarding the detection process implemented by the Analysis Engine Module, a Classification And Regression Tree (CART) decision tree was deployed by utilising the CICIDS2017 dataset. Based on the evaluation analysis, the accuracy and the True Positive Rate (TPR) of the proposed IDS reach 0.996 and 0.993 respectively.

In [15], B. Kang et al. implemented a signature-based IDS for IEC 61850 substations by using the Suricata IDS. More detailed, a stateful analysis plugin was implemented into Suricata, whose architecture is divided into three units, namely 1) Manufacturing Message Specification (MMS) decoder, 2) rule match engine and 3) state manager. The first unit decodes the MMS packets by extracting their attributes. The second

unit applies the signature rules, while the role of the last unit is to update the state of the protected devices. Concerning the evaluation process, two cyberattacks were performed and detected successfully.

In [16], Y. Yang et al. implemented a specification-based IDS devoted to protecting synchrophasor systems utilising IEEE C37.118. In particular, their IDS is composed of 1) access control rules, 2) protocol rules and 3) behaviour rules. The access control rules determine the legitimate Medium Access Control (MAC) and the Internet Protocol (IP) addresses as well as the corresponding transport layer ports permitted to transmit and receive network packets. The protocol rules define that only IEEE C37.118 network packets can be transmitted by the various entities. Finally, the last category adopts a deep packet inspection process, thus defining behaviour rules based on the attributes of IEEE C37.118. Concerning the evaluation process, the False Positive Rate (FPR) is calculated approximately at 0%.

A survey of different anomaly detection techniques in various application domains, including energy, was presented in [17]. Also, a more specific study related to IDS systems for SG was presented in [4]. Methods based on data analytics and statistics are commonly used approaches to identify abnormal behaviours using network and operational data. Statistical techniques aim to evaluate if the testing datasets fit to previously modelled distributions. Jakkula and Cook use statistics and clustering techniques to identify outliers in datasets collected from smart environments [18], but they have not considered the impact of the exogenous variables, e.g., weather temperature on the electricity consumption. Adnan et al. combine linear regression with clustering techniques for getting better results [19]. Zhang et al. [20] further use piecewise linear regression to t the relation between energy consumption and weather temperature. Brown et al. use K-Nearest Neighborhood (KNN) in fast kernel regression to predict electricity consumption [21], which requires large datasets. Nadai et al. combine Autoregressive Integrated Moving Average (ARIMA) and adaptive Articial Neural Network (ANN) to detect anomaly consumption [22] using a relatively small data set that is from a few buildings.

In [23], the authors move away from the concept of single events identified as an anomaly to the concept of the collective anomaly, that is itemsets of events that may be anomalous based on their patterns of appearance. The work in [24] investigates the use of re-sampling techniques for intrusion detection inside of a hierarchical, three-layer SG communication system and the authors in [25] proposed a generative model for anomaly detection that takes into account the hierarchical structure of the network and the data collected from smart meters. A method that applies the K-means algorithm for clustering of traffic data and outlier detection was introduced in [26] for the data transmitted between the utility centre and the smart homes. In [27] an approach based on stacked sparse autoencoder was introduced to extract the high-level representation from massive monitoring data acquired automatically from actual smart meter network. Then softmax is used for classification to detect anomalies and send alarm messages using web technologies.

Undoubtedly, each of the previous works provides significant and useful methodologies for detecting cyberattacks and anomalies against SG. Each methodology includes the corresponding advantages and disadvantages. In particular, the signature-based IDS systems are characterised by high detection performance, but they cannot detect unknown attacks. On the other hand, the anomaly-based techniques usually adopting a machine learning method are able to detect unknown attacks, but they exhibit many false positives. Finally, the specification-based techniques combine the advantages of the previous methods, but they should be updated continuously due to the rapidly evolving nature of SG.

In this paper, we present an anomaly-based IDS utilising a plethora of machine learning and deep learning algorithms. In particular, compared to the previous works, the rationale behind our implementation is coming from a real-life paradigm, where temperature values are utilised to identify possible anomalies in a real power plant. Furthermore, our work adopts a novel complex data representation, evaluating the performance of many algorithms.

III. IDS Background

The goal of an IDS system is to detect possible attacks and anomalies either by informing timely the system operator or the security administrator or performing some countermeasures. The typical architecture of IDS consists of three entities, namely 1) Agents, 2) Analysis Engine and 3) Response Module. Agents are responsible for monitoring and collecting information about the target systems. Two kinds of agents are defined: 1) Host-Based and 2) Network-Based. The first one monitors only one specific host, thus collecting information such as the network traffic generated and received by this host as well as its log files. On the other hand, the second one is able to monitor the network traffic generated by many devices. Based on this discrimination, the IDS systems can be classified into two categories: 1) HIDS and 2) NIDS. Accordingly, the Analysis Engine receives the information collected by the agents and tries to detect possible cyberattack or anomaly patterns. The detection mechanisms applied by the analysis engine can be classified into three categories: 1) signature-based, 2) anomaly-based and 3) specification-based. The first one matches the information collected by the agents with specific attack signatures. The second category attempts to identify possible anomalies by adopting statistical analysis and AI techniques. The last category matches the information collected by the agents with a set determining the legitimate behaviours. Finally, the Response Module informs the responsible administrator about the possible cyberattacks and anomalies.

IV. Architecture of the Proposed IDS

The proposed method for anomaly detection using operational data is based on an supervised learning framework that incorporates a training and a testing stage, as illustrated

in Fig. 2. This section analyses the suggested methodology and provides details for the data acquisition process. Also, all the pre-processing steps are analysed and the proposed novel complex representation of the descriptor is presented.

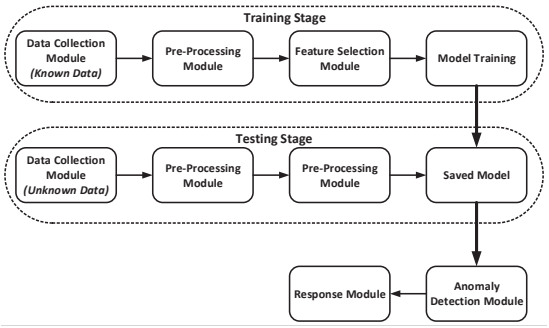

Fig. 2: Architecture of the Porposed IDS.

A. Data Collection Module

Initially, operational data and more specifically temperature values coming from the incoming cooling water and the generator winding was collected from the Lavrio Unit 5 Power Plant overtime sampled every minute. As depicted in Fig. 1 the specific unit consists of multiple logic controllers such as Programmable Logic Controllers (PLCs) and Remote Terminal Units (RTUs) that monitor and control the operations of the industrial equipment existing such as power generators and transformers. Moreover, it includes many smart meters that record the energy consumption of the various Intelligent Electronic Devices (IEDs) utilised such as monitors and computers. It should be noted that all information generated by the power plant devices is recorded in a centralised control server. The ground truth was provided by the power plant engineers indicating the anomalies and the events that triggered them. The input is a time-series and the training is performed only with normal data. Fig. (3) shows a sample of the dataset utilised.

For the proposed machine learning approach the dataset was split to training and testing subsets and simple k-fold Cross Validation (CV) was also used for analysing the state of the system and accessing the reliability of the proposed method.

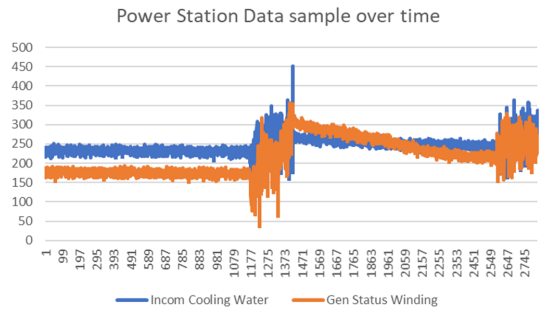

Fig. 3: Sample of the operational data from the power plant.

B. Pre-processing and Feature selection Modules

Several pre-processing approaches were implemented and tested for the training and evaluation of the anomaly detection. Feature standardisation was considered making the values of each feature in the data have zero-mean and unit variance. Let $f(\mathbf{m}), \mathbf{m} = [x, y]^T E \in R$ denote the feature vector with x and y to represent the water and generator temperature respectively. Initially, the distribution mean and standard deviation for each feature x and y is determined, and subsequent the mean from each feature is subtracted and the values of each one are divided by its standard deviation.

$$f' = \frac{f - \bar{f}}{\sigma} \tag{1}$$

Where f is the original feature vector, $\bar{f}$ is the mean of that feature vector, and σ is its standard deviation.

During the data capturing and feature generation stages, it is common that complex data is obtained but decomposed to independent values without considering if there is a correlation between them and they are computed independently by a classier. To overcome this issue, vectorial features can be represented more precisely using a complex representation [28], [29]. Since, in our case, vectorial features $\mathbf{m} = [x, y]$ are the primary source of information, a complex representation of these features allows better correlation between them [28]–[30].

Considering a complex vector z representation for the pre-processed features we have $f(z), z = x + iy \in C$ that can be also denoted using the Euler representation $z = re^{i\phi}$ where $r = |z| = \sqrt{x^2 + y^2}$ is the magnitude of z and $\phi = argz = atan2(y, x)$. As a result a solid representation of the selected complex features is obtained, while the computational complexity does not increase significantly. A vector f can be decomposed in components that are linearly independent, and therefore, they can reconstruct the original data by linearly combining them. However, a correlation may exist between the components x and y from the statistical point of view (i.e., water and generator temperatures are not uncorrelated). The proposed complex descriptor does not affect the overall performance if the components are independent, but this complex representation considers and takes advantage of that improving the performance, if there is a correlation. The proposed method capturing the dependencies within the two temperature sensors exploits the complex representation.

C. Anomaly Detection Module

For the anomaly detection, several machine learning methods were considered including One Class-SVM, Isolation Forests, Angle-Base Outlier Detection (ABOD), Stochastic Outlier Selection (SOS), Principal Component Analysis (PCA) and deep fully connected autoencoders. The proposed complex feature vectors over a sliding time window were used as input for all these approaches.

During the training stage, all the methods were trained using normal data (without anomalies) and regarded the above methods, linear kernels were used for the PCA and One Class

TABLE I: The evaluation results for time window 20.

Win20	Accuracy Norm	Accuracy Comp	F1s Norm	F1s Comp	AUC Norm	AUC Comp
PCA	0.554	**0.975**	0.608	**0.975**	0.710	**0.967**
OneClass	0.595	**0.893**	0.648	**0.869**	**0.735**	0.657
Iforest	0.575	**0.875**	0.628	**0.884**	0.729	**0.851**
ABOD	0.521	**0.688**	0.584	**0.718**	**0.586**	0.547
SOS	0.951	**0.975**	0.947	**0.975**	0.842	**0.967**
Auto	0.560	**0.619**	0.614	**0.669**	0.718	**0.763**

TABLE II: The evaluation results for time window 30.

Win30	Accuracy Norm	Accuracy Comp	F1s Norm	F1s Comp	AUC Norm	AUC Comp
PCA	0.527	**0.597**	0.585	**0.652**	0.687	**0.736**
OneClass	0.553	**0.905**	0.611	**0.885**	**0.702**	0.681
Iforest	0.548	**0.791**	0.606	**0.819**	0.697	**0.864**
ABOD	0.530	**0.799**	0.591	**0.791**	**0.642**	0.568
SOS	0.976	**0.989**	0.976	**0.990**	0.921	**0.994**
Auto	0.538	**0.603**	0.596	**0.658**	0.693	**0.740**

TABLE III: The evaluation results for time window 50.

Win50	Accuracy Norm	Accuracy Comp	F1s Norm	F1s Comp	AUC Norm	AUC Comp
PCA	0.459	**0.550**	0.523	**0.614**	0.627	**0.684**
OneClass	0.516	**0.913**	0.581	**0.909**	0.659	**0.773**
Iforest	0.480	**0.617**	0.545	**0.675**	0.637	**0.757**
ABOD	0.574	**0.739**	0.640	**0.765**	**0.603**	0.598
SOS	0.989	**0.995**	0.989	**0.995**	0.960	**0.997**
Auto	0.466	**0.553**	0.530	**0.617**	0.628	**0.684**

SVM, the Euclidean distance was used to obtain the dissimilarity matrix and T-distributed Stochastic Neighbor Embedding (tSNE) to calculate the affinity matrix for SOS. The angle-based outlier factor used for ABOD is defined as the variance over the angles between the feature vectors weighted by their distance and for the Isolation forest approach the algorithm is using full decision trees and we measure the average path length between the root and each leaf (feature point) with the abnormal data points to be the ones with relatively short average path. About the designed deep architecture, the following autoencoder was designed with six fully connected layers as it is shown in Fig. (4).

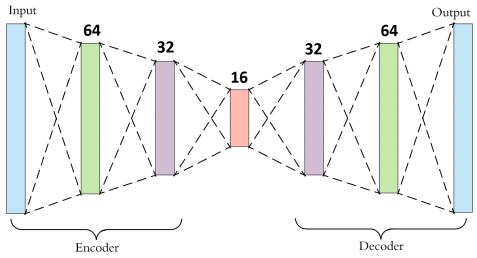

Fig. 4: An overview of the architecture of the proposed autoencoder that was used for anomaly detection.

The objective of the testing stage is the detection of anomalies in new operational data. Therefore, new incoming operational data is pre-processed, transformed into a complex feature representation, following the same pre-processing steps as in the training stage. Once the features are extracted, all the models created during the training are used to detect anomalies at the new input data and determine the presence of any attacks.

D. Response Module

This module receives the output of the *Anomaly Detection Module* and undertakes to inform the security operator or the security administrator about the possible cyberattacks by extracting the appropriate security events. A web-based platform was developed for this purpose, providing also related statistics.

V. Evaluation Analysis

This section shows and analyses the anomaly detection results obtained using the normal real uncorrelated features and the proposed complex representation for all the machine learning and deep learning approaches. The results are represented by the Accuracy, the F1 score, and AUC.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{2}$$

$$F1 = \frac{2 \cdot precision \cdot recall}{precision + recall} \tag{3}$$

$$AUC = \int_a^b TPR(FPR^{-1}(x))dx = P(X_1 > X_0) \tag{4}$$

Where X_1 is the score for a positive instance and X_0 is the score for a negative instance.

Furthermore, all the methods and features were tested for three different sliding time windows of 20, 30 and 50 minutes. Tables I, II and III show the Accuracy, the F1 scores and the AUC values for all the models and both features (normal

Fig. 5: The overall performance of the machine learning and deep learning methods with and without the proposed complex feature representation and the affect of the size of the sliding time window.

and complex), indicating an improvement for the proposed complex representation. Furthermore, the average values for each method and the overall effect of the size of the time window are shown in Fig. (5). The overall average ccuracy was increased by 29%, the F1 score by 22% and the AUC by 8%. Regarding the sliding time window, it is observed that the performance is reduced for larger sizes, something that was expected since it affects the on-time estimation of sudden outliers.

VI. Conclusions

In this work, a novel approach for cyberattacks detection on SGs has been introduced based on anomaly detection over operational data. Furthermore, a complex representation of the input data was suggested aiming to exploit the correlation in-between the data values improving the overall accuracy of anomaly detection. Several machine learning and deep learning methods were used in a comparative study demonstrating the improved performance of the proposed methodology. Furthermore, real operational data from a power plant was used and different parameters were considered. Overall the use of operational data and anomaly detection methods provide a new mechanism for accurate detection of cyberattacks and anomalies that may affect the performance of the various devices on the grid. Regarding the future plans it is expected to improve the proposed architecture considering more advanced DNNs based on LSTM layers and manifold representations for multi-dimensional operational data.

VII. Acknowledgement

This project has received funding from the European Unions Horizon 2020 research and innovation programme under grant agreement No. 787011 (SPEAR).

References

[1] S. Tan, D. De, W. Song, J. Yang, and S. K. Das, "Survey of security advances in smart grid: A data driven approach," *IEEE Communications Surveys Tutorials*, vol. 19, no. 1, pp. 397–422, Firstquarter 2017.

[2] A. Triantafyllou, P. Sarigiannidis, and T. D. Lagkas, "Network protocols, schemes, and mechanisms for internet of things (iot): Features, open challenges, and trends," *Wireless Communications and Mobile Computing*, vol. 2018, 2018.

[3] C. Dalamagkas, P. Sarigiannidis, D. Ioannidis, E. Iturbe, O. Nikolis, F. Ramos, E. Rios, A. Sarigiannidis, and D. Tzovaras, "A survey on honeypots, honeynets and their applications on smart grid," in *The 1st International Workshop on Cyber-Security Threats, Trust and Privacy Management in Software-defined and Virtualized Infrastructures*, 06 2019, pp. 1–8.

[4] P. I. Radoglou-Grammatikis and P. G. Sarigiannidis, "Securing the smart grid: A comprehensive compilation of intrusion detection and prevention systems," *IEEE Access*, vol. 7, pp. 46 595–46 620, 2019.

[5] S. K. Goudos, P. Sarigiannidis, P. I. Dallas, and S. Kyriazakos, *Communication Protocols for the IoT-Based Smart Grid*. Cham: Springer International Publishing, 2019, pp. 55–83. [Online]. Available: https://doi.org/10.1007/978-3-030-03640-9_4

[6] P. I. R. Grammatikis, P. G. Sarigiannidis, and I. D. Moscholios, "Securing the internet of things: Challenges, threats and solutions," *Internet of Things*, vol. 5, pp. 41 – 70, 2019. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S2542660518301161

[7] A. Patel, H. Alhussian, J. M. Pedersen, B. Bounabat, J. C. Júnior, and S. Katsikas, "A nifty collaborative intrusion detection and prevention architecture for smart grid ecosystems," *Computers & Security*, vol. 64, pp. 92–109, 2017.

[8] M. A. Faisal, Z. Aung, J. R. Williams, and A. Sanchez, "Data-stream-based intrusion detection system for advanced metering infrastructure in smart grid: A feasibility study," *IEEE Systems Journal*, vol. 9, no. 1, pp. 31–44, March 2015.

[9] V. Hadjioannou and et al., "Security in smart grids and smart spaces for smooth iot deployment in 5g," *Internet of Things (IoT) in 5G Mobile Technologies, Springer Cham*, pp. 371–397, 2016.

[10] J.-M. Batalla and et al., "Efficient media streaming with collaborative terminals for the smart city environment," *IEEE Communications Magazine*, vol. 1, no. 55, pp. 98–104, 2017.

[11] Y. Nikoloudakis and et al, "A fog-based emergency system for smart enhanced living environments," *IEEE Cloud Computing*, vol. 6, pp. 54–62, 2016.

[12] E. Markakis and et al., "Acceleration at the edge for supporting smes security: The fortika paradigm," *IEEE Communications Magazine*, 2019.

[13] M. Gajewski and et al., "Two-tier anomaly detection based on traffic profiling of the home automation system," *Computer Networks Journal, Elsevier*, vol. 158, no. 15, pp. 46–60, 2019.

[14] P. I. Radoglou-Grammatikis and P. G. Sarigiannidis, "An anomaly-based intrusion detection system for the smart grid based on cart decision tree," in *2018 Global Information Infrastructure and Networking Symposium (GIIS)*, Oct 2018, pp. 1–5.

[15] B. Kang, K. McLaughlin, and S. Sezer, "Towards a stateful analysis framework for smart grid network intrusion detection," in *Proceedings of the 4th International Symposium for ICS & SCADA Cyber Security Research 2016*. BCS Learning & Development Ltd., 2016, pp. 1–8.

[16] Y. Yang, K. McLaughlin, S. Sezer, T. Littler, B. Pranggono, P. Brogan, and H. F. Wang, "Intrusion detection system for network security in synchrophasor systems," in *IET International Conference on Information and Communications Technologies*, April 2013, pp. 246–252.

[17] B. A. K. V. Chandola, V., "Anomaly detection: a survey," in *ACM Comput. Surv.*, vol. 41, no. 3, 2009, p. 15.

[18] C. D. Jakkula, V., "Outlier detection in smart environment structured power datasets," in *6th International Conference on Intelligent Environments*, 2010, p. 2933.

[19] R. Adnan, H. Setan, and M. Mohamad, "Multiple outliers detection procedures in linear regression," in *Matematika*, vol. 19, 2003, p. 2945.

[20] C. W. B. J. Zhang, Y., "Anomaly detection in premise energy consumption data," in *Power and Energy Society General Meeting*, 2011, p. 18.

[21] B.-L. C. B. Z. Brown, M., "Kernel regression for real-time building energy analysis," in *J. Build. Perform. Simul.*, vol. 5, no. 4, 2011, p. 263276.

[22] v. S. M. De Nadai, M., "Short-term anomaly detection in gas consumption through arima and articial neural network forecast," in *IEEE Workshop on Environmental, Energy and Structural Monitoring Systems*, 2015, p. 250255.

[23] B. Rossi, S. Chren, B. Buhnova, and T. Pitner, "Anomaly detection in smart grid data: An experience report," in *2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, Oct 2016, pp. 002 313–002 318.

[24] C. Promper, D. Engel, and R. C. Green, "Anomaly detection in smart grids with imbalanced data methods," in *2017 IEEE Symposium Series on Computational Intelligence (SSCI)*, Nov 2017, pp. 1–8.

[25] R. Moghaddass and J. Wang, "A hierarchical framework for smart grid anomaly detection using large-scale smart meter data," *IEEE Transactions on Smart Grid*, vol. 9, no. 6, pp. 5820–5830, Nov 2018.

[26] D. M. Menon and N. Radhika, "Anomaly detection in smart grid traffic data for home area network," in *International Conference on Circuit, Power and Computing Technologies*, March 2016, pp. 1–4.

[27] Y. Yuan and K. Jia, "A distributed anomaly detection method of operation energy consumption using smart meter data," in *2015 International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP)*, Sep. 2015, pp. 310–313.

[28] T. Adali, P. Schreier, and L. Scharf, "Complex-valued signal processing: The proper way to deal with impropriety," *IEEE Trans. Signal Processing (overview paper)*, vol. 59, no. 11, p. 51015123, 2011.

[29] X.-L. Li, T. Adali, and M. Anderson, "Noncircular principal component analysis and its application to model selection," *IEEE Sig. Proc.*, vol. 59, no. 10, p. 45164528i, 2011.

[30] Z. Chai, K.-K. Ma, and Z. Liu, "Complex wavelet-based face recognition using independent component analysis," *Fifth Intern. Conf. on Intelligent Information Hiding and Multimedia Signal Proc.*, p. 832835, 2009.