Neural network architecture to detect system faults / cyberattacks anomalies within a photovoltaic system connected to the grid

Giovanni Battista Gaggero, Mansueto Rossi, Paola Girdinio and Mario Marchese
Department of Electrical, Electronic and Telecommunications Engineering, and Naval Architecture
DITEN, University of Genoa, Italy, Via Opera Pia 11A, 16145, Genoa, Italy

***Abstract*—Anomaly detection is an important issue heavily investigated within different research areas and application domains. Its application in the industrial systems sector may be essential also for the protection of critical infrastructures. Due to the huge amount of involved data and to their complexity the use of machine learning may be the clue. The basic idea is describing an industrial process by a series of key attributes whose measures (the features) compose a state vector including heterogeneous types of measurements. Each feature should be a key attribute which can help discriminate between a normal functioning condition and an anomaly. In this context, the paper presents the use of a deep neural network architecture called autoencoder to detect anomalies due to either system faults or cyberattacks. The chosen application field is a photovoltaic system connected to the grid. The results, even if preliminary, are really promising.**

***Index Terms*—anomaly detection, industrial systems, neural networks**

I. Introduction

Under the pressure of environmental problems, the electrical grid is moving toward a large use of renewable energy sources (RES). Because of the uncertainty of RESs energy production and the high scalability of some solutions (like solar panel), the electrical grid needs to change paradigm from a centralized way to produce energy to a more distributed one. Distributed energy resources (DER) usually refers to a small-scale unit of power generation that is connected to a larger power grid at the distribution level. These may include solar panels, batteries and other storage systems, micro-turbines often cogenerative, and so on. However, the large use of DER brings many problems to the grid. One of the main concerns is related to security [1] [2] [3] [4]: DER are automated systems often connected by a telecommunication network to a control system, for example a Supervisory Control and Data Acquisition (SCADA), in order to coordinate the energy production between the sources. In this context, common communication protocols are Modbus, DNP3 and IEC 61850. All these protocols present severe security issues and fail to ensure the integrity and confidentiality of the communication because of the lack of encryption [5] [6]. DERs can produce a lot of information regarding its working conditions, which has to be monitored by human experts. However, given the complexity of these systems and the high number of plants to monitor, an anomaly detection system that can mimic the human behaviour is a very important issue to manage the grid. The paper is structured as follows. The next section reports a short review of the state of the art concerning anomaly detection applied over industrial systems. Section III describes the methodology and the main idea which this paper is based on: the use of a deep neural network architecture called autoencoder to detect anomalies due to either faults or cyberattacks. Section IV considers the example of a photovoltaic system and applies the designed autoencoder scheme to it. Section V contains a preliminary performance evaluation and Section VI the conclusions.

978-1-7281-3729-2/19/$31.00 2019 IEEE

II. State of the art

Anomaly detection is an important issue that has been investigated within different research areas and application domains [7]. The action aimed at identifying all the behaviors that differ in some way from the normal one is usually called novelty detection or outlier detection [8] [9], for which [10] and [11] propose different autoencoder architectures and [12] suggest a one-class neural network model. Different ML solutions have been investigated to identify faults in power generation systems: in [13] a k-nn supervised algorithm is used to identify faults in the direct-current portion of a solar power plant composed by many arrays. A Support Vector Machine (SVM) classification algorithm is used in [14] to detect faults in Power Generation Systems Based on Solid Oxide Fuel Cells. Artificial Neural Networks (ANN) are applied in [15] in order to identify malicious control of DERs in a grid with high penetration of photovoltaic (PV) generators. An artificial neural network is used in [16] to solve a regression problem in order to predict the power produced by a photovoltaic plant and detect anomalies. [17] introduces an autoencoder architecture to discriminate between fault conditions in an electric motor and make a comparison with a One Class Support Vector Machine classifier.

III. Methodology

A. Machine learning overview

Lets consider a portion of an automated industrial process, for which we have different type of measured key features. The aim is to describe the functioning of the process by this group of key features. The features measurements

$x_1(t), x_2(t)...x_n(t)$ are collected so to build the vector shown in (1) at each sampling time t. (1) is called state vector and each element of the vector is a feature.

$$X(t) = \{x_1(t), x_2(t)...x_n(t)\} \tag{1}$$

The vector can include heterogeneous type of measurements: physical measurements from the process, environmental measurements, parameters of the telecommunication network, and so on. Each feature should be a key attribute which can help discriminate between a normal functioning condition and an anomaly. An anomaly can be defined into different ways: it could be the fault of a component or of a sensor, or be caused by a malicious manumission. Regardless of the cause, the anomaly is defined here as an undesired working condition of the physical process. A cyber attack that targets a system like an industrial process "translates" into a bad working condition: the purpose of the proposed algorithm is to automatically detect it by observing the vector (1) that describes the physical behaviour.

Many scientific works focus on the dynamic state estimation to detect anomalies by implementing the equations that describe the system and by choosing a threshold for the error between estimated and measured parameters. The application of such type of approach may have some drawbacks:

- It requires to know the exact behaviour of the system, which means to know the exact parameters of the equations.
- It could be very hard to write a closed-form equation which takes into account heterogeneous types of measurements.
- It requires a customized design.

Machine learning approaches could be useful to face up such type of problems. Traditionally, anomaly detection is considered as a ML classification problem: intrusion detection is a solid field of research in which large datasets containing normal and abnormal examples of behavior are collected and used to train a classification algorithm. Usual algorithms are linear algorithms (like logistic regression and support vector machine), decision trees, instance based or lazy algorithms (like k-nearest neighbour), or deep neural networks. The main problem of these approaches in the field of physical behaviour-related anomaly detection is to find appropriate datasets. It is really difficult to dispose of labeled datasets of faults, even more difficult to have labeled datasets of bad behaviors if the anomaly is caused by a cyber attack. A solution may be simulating different types of faults, but it is hard to forecast all of the possible bad operating conditions. Another type of solution is represented by the regression problem: first we choose some target variables to monitor, then we define a threshold for the error on the forecasted values. If the error exceeds the threshold, then we report the anomaly. The main drawback of this solution is that we have to choose a limited number of features as output.

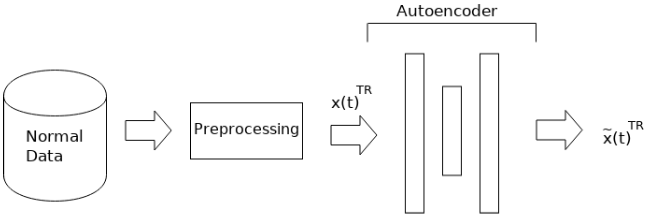

Fig. 1. Training phase

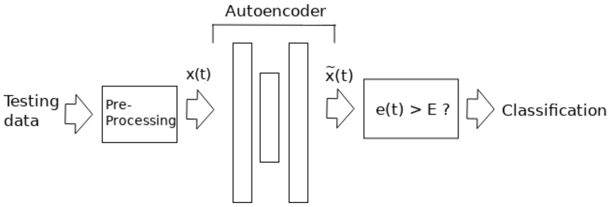

Fig. 2. Test phase

B. Autoencoders

The proposed solution uses a deep neural network architecture called autoencoder. Autoencoders are a type of unsupervised learning algorithms in which the neural network learns to reconstruct its input: trained the NN with the normal behavior, we expect that the network reconstructs abnormal data with an higher error and normal data with a lower error; in this way, we can use the magnitude of the error to classify new data. So, the anomaly detection algorithm is composed of two phases: in the training phase, shown in Fig. 1, the network is trained with only normal behavior data; after the training phase, we set the threshold for the error between the input and the output of the autoencoder. Then, in the test phase, in Fig. 2, we provide the neural network with unlabeled data and the error is calculated: data are classified as abnormal if the error exceeds the threshold. In detail, we call $x^{TR}(t)$ the input vector of time t of the train dataset and $\tilde{x}^{TR}(t)$ the reconstructed vector; we define the error as in (2)

$$e^{TR}(t) = \left\| x^{TR}(t) - \tilde{x}^{TR}(t) \right\| \tag{2}$$

Evaluating the mean $e^{TR}(t)$ of the training dataset, we empirically set the threshold E. Then we evaluate the error on the test dataset as in (3)

$$e(t) = \| x(t) - \tilde{x}(t) \| \tag{3}$$

if $e(t)$ exceeds E, the sample is classified as anomalous, otherwise as normal.

IV. APPLICATION

A. Case study overview

Lets consider a typical scheme of a photovoltaic systems shown in Fig. 3, composed of: solar panels, DC/DC converter (usually a booster converter) that is controlled by the Maximum Power Point Tracking (MPPT) algorithm, DC link, and finally Current-source inverter connected to the grid. We categorize all the collectible information in five groups:

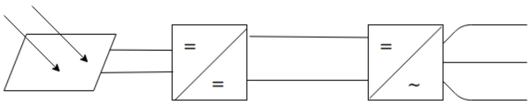

Fig. 3. Photovoltaic System

- Alternated Current (AC) side electrical information: active and reactive power, voltages (Root Mean Square, RMS), currents (RMS), frequencies, total harmonic distortion (THD)
- Direct Current (DC) side electrical information: voltages and currents
- PV information: voltage, current, temperature of the cells
- Environmental information: irradiance, temperature of the air
- Electronic information: maximum power point, dc/dc converter dutycycle

We expect that the anomaly detection algorithm learn the correlation between the physical parameters; for example, we expect a correlation between the measured irradiance and the AC-side currents, even if it would be difficult to write an equation in a closed form. We collect all information at a certain time so to create the state vector (1). The state vector is then sent to the anomaly detection algorithm.

x_1	Irradiance: the solar irradiance hitting the panel
x_2	T_air: the temperature of the environment
x_3	T_pv: the temperature of the PV's cells
x_4	V_pv: the voltage measured at the terminals of the panel
x_5	I_pv: the current emitted by the panel
x_6	V_dc: the voltage measured at the DC link
x_7	I_c: the average current in the DC capacitor
x_8	δ: the dutycycle of the DC/DC converter
x_9	V_a: the voltage of phase a (AC side)
x_{10}	V_b: the voltage of phase b (AC side)
x_{11}	V_c: the voltage of phase c (AC side)
x_{12}	I_a: the current of phase a
x_{13}	I_b: the current of phase b
x_{14}	I_c: the current of phase c
x_{15}	f_a: the frequency of phase a
x_{16}	f_b: the frequency of phase b
x_{17}	f_c: the frequency of phase c
x_{18}	THD_a: the total harmonic distortion of the voltage on phase a
x_{19}	THD_b: the total harmonic distortion of the voltage on phase b
x_{20}	THD_c: the total harmonic distortion of the voltage on phase c
x_{21}	Q: the reactive power emitted by the inverter
x_{22}	P: the active power emitted by the inverter

TABLE I
STATE VECTOR

B. Anomaly model

Photovoltaic systems are prone to different types of faults [18] [19]. While electrical faults like short circuits must be considered by electrical protections, there are different problems that produce a degradation of the performances, like partial shading faults, open-circuits faults, soiling and so on. Moreover, aging of the cells translate into a loss of efficiency. There are different types of techniques to detect such type of faults; our work proposes a different approach based on the automatic analysis of data generated by DERs.

DERs are prone to different cyber attacks as well. As mentioned, the lack of security of many industrial protocols leads to the risk of man-in-the-middle attacks following a breach in the perimeter defense. However, the attacks can come not only from the network: components can run a malicious firmware caused by insiders, receive supply chain attacks or even be victim of human errors. All these threats translate into abnormal physical working conditions.

V. PERFORMANCE EVALUATION

In order to preliminary validate the proposed scheme, we set up a simulation environment with of a photovoltaic system connected to the grid by using MATLAB/Simulink, as shown in Fig. 4. The electrical model is electromagnetic and includes the simulation of the heat exchange with the environment and the control system of the electrical converters. We run the simulations for different working conditions and extract the features as shown in Table I. Each samples contains 22 features, which comprehend electrical data, thermal data and the DC/DC converter's dutycycle, as detailed above.

Acquired data are then divided into two parts: the first one is used to train the autoencoder while the second one is used to test it. We define two types of anomalies:

- data corruption
- bad physical behavior

In order to simulate a bad data injection or a sensor fault, the value of one data for each sample is modified in order to create a physically impossible/anomalous state vector; for example, keeping the same current and voltage values, injected power is modified. Bad physical behavior is obtained by corrupting the control system of the electrical converters: for example we add a bias in the MPPT algorithm in order to make the PV work at a lower voltage than the maximum power point.

Then we pre-process the data: because of the different scale and variation range of each feature, applying MSE to raw data would produce biased results. So, we firstly normalize the dataset: we calculate the average and the standard deviation for each feature on the test dataset, then we apply (4):

$$\frac{(x_i(t) - \bar{x_i})}{\sigma_i} \tag{4}$$

Where $\bar{x_i}$ is the average value and σ_i is the standard deviation over the training dataset of the i-th feature.

Finally we set up the anomaly detection algorithm. The first part is represented by a shallow autoencoder on MATLAB deep learning toolbox. We set the dimension of the hidden layer at 16. Then we evaluate the mean square error between the real and predicted data on the training dataset and we set the threshold heuristically to classify new data. The overall accuracy is 79.71%. The confusion matrix is reported in Table II, with the threshold set to 0.06.

Changing the threshold in a small range has no evident impact on the overall accuracy but the confusion matrix changes significantly: when changing the threshold False Positives

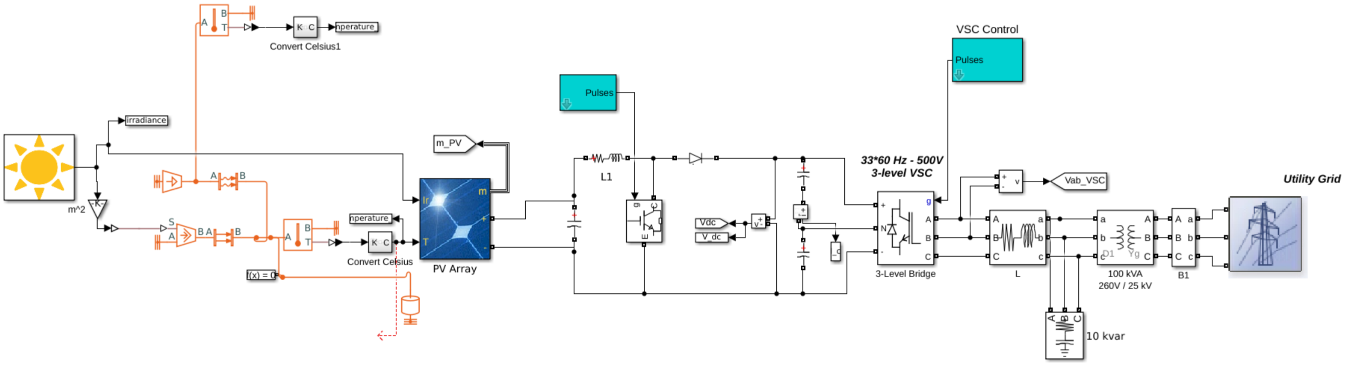

Fig. 4. Simulation environment in Simulink

		Real class	
		Normal	Anomaly
Classified as	Normal	0.245	0.030
	Anomaly	0.172	0.552

TABLE II
CONFUSION MATRIX

increases while False Negative decreases, or vice versa. The size of the hidden layer influences the results: a size between 18 and 10 produces similar results, while out of this range accuracy decreases.

VI. Conclusions

Machine learning algorithms may be fundamental tools to improve the security level of automated systems. In particular, deep autoencoders could be really useful to build physical behaviour-related anomaly detection schemes, thanks to their capability to learn the links between heterogeneous physical parameters. Distributed energy resources can be a good field of application because of the huge amount and complexity of data. It will be worth investigating different types of autoencoding architectures like fully-connected deep architectures, convolutional and recurrent neural network autoencoders.

References

[1] Qi, Junjian, et al. ”Cybersecurity for distributed energy resources and smart inverters.” IET Cyber-Physical Systems: Theory Applications 1.1 (2016): 28-39.

[2] Rawat, Danda B., and Chandra Bajracharya. ”Cyber security for smart grid systems: Status, challenges and perspectives.” SoutheastCon 2015. IEEE, 2015.

[3] Kang, BooJoong, et al. ”Investigating cyber-physical attacks against IEC 61850 photovoltaic inverter installations.” 2015 IEEE 20th Conference on Emerging Technologies Factory Automation (ETFA). IEEE, 2015.

[4] Carter, Cedric, et al. ”Cyber Security Assessment of Distributed Energy Resources.” 2017 IEEE 44th Photovoltaic Specialist Conference (PVSC). IEEE, 2017.

[5] Dzung, Dacfey, et al. ”Security for industrial communication systems.” Proceedings of the IEEE 93.6 (2005): 1152-1177.

[6] Xu, Yikai, et al. ”Review on cyber vulnerabilities of communication protocols in industrial control systems.” 2017 IEEE Conference on Energy Internet and Energy System Integration (EI2). IEEE, 2017.

[7] Chandola, Varun, Arindam Banerjee, and Vipin Kumar. ”Anomaly detection: A survey.” ACM computing surveys (CSUR) 41.3 (2009): 15.

[8] Markou, Markos, and Sameer Singh. ”Novelty detection: a reviewpart 1: statistical approaches.” Signal processing 83.12 (2003): 2481-2497.

[9] Markou, Markos, and Sameer Singh. ”Novelty detection: a reviewpart 2: neural network based approaches.” Signal processing 83.12 (2003): 2499-2521.

[10] Amarbayasgalan, Tsatsral, Bilguun Jargalsaikhan, and Keun Ryu. ”Unsupervised novelty detection using deep autoencoders with density based clustering.” Applied Sciences 8.9 (2018): 1468.

[11] Chen, Jinghui, et al. ”Outlier detection with autoencoder ensembles.” Proceedings of the 2017 SIAM International Conference on Data Mining. Society for Industrial and Applied Mathematics, 2017.

[12] Chalapathy, Raghavendra, Aditya Krishna Menon, and Sanjay Chawla. ”Anomaly detection using one-class neural networks.” arXiv preprint arXiv:1802.06360 (2018).

[13] Harrou, Fouzi, Bilal Taghezouit, and Ying Sun. ”Improved k NN-Based Monitoring Schemes for Detecting Faults in PV Systems.” IEEE Journal of Photovoltaics 9.3 (2019): 811-821.

[14] Costamagna, Paola, et al. ”A classification approach for model-based fault diagnosis in power generation systems based on solid oxide fuel cells.” IEEE Transactions on Energy Conversion 31.2 (2015): 676-687.

[15] Kosek, Anna Magdalena. ”Contextual anomaly detection for cyber-physical security in Smart Grids based on an artificial neural network model.” 2016 Joint Workshop on Cyber-Physical Security and Resilience in Smart Grids (CPSR-SG). IEEE, 2016.

[16] Kosek, Anna Magdalena, and Oliver Gehrke. ”Ensemble regression model-based anomaly detection for cyber-physical intrusion detection in smart grids.” 2016 IEEE Electrical Power and Energy Conference (EPEC). IEEE, 2016.

[17] Principi, Emanuele, et al. ”Unsupervised electric motor fault detection by using deep autoencoders.” IEEE/CAA Journal of Automatica Sinica 6.2 (2019): 441-451.

[18] Pillai, Dhanup S., Frede Blaabjerg, and Natarajan Rajasekar. ”A comparative evaluation of advanced fault detection approaches for PV systems.” IEEE Journal of Photovoltaics 9.2 (2019): 513-527.

[19] AbdulMawjood, Kais, Shady S. Refaat, and Walid G. Morsi. ”Detection and prediction of faults in photovoltaic arrays: A review.” 2018 IEEE 12th International Conference on Compatibility, Power Electronics and Power Engineering (CPE-POWERENG 2018). IEEE, 2018.