This is a postprint version of the following published document:

Gotti, Davide; Amaris, Hortensia; Ledesma Larrea, Pablo (2021). A Deep Neural Network Approach for Online Topology Identification in State Estimation. *IEEE Transactions on Power* Systems, 36(6), pp.: 5824 - 5833.

DOI: https://doi.org/10.1109/TPWRS.2021.3076671

© 2021 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.
See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Power Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TPWRS.2021.3076671

A Deep Neural Network Approach for Online Topology Identification in State Estimation

Davide Gotti, *Student Member, IEEE*, Hortensia Amaris, *Senior Member, IEEE*, and Pablo Ledesma

***Abstract*—This paper introduces a network topology identification (TI) method based on deep neural networks (DNNs) for online applications. The proposed TI DNN utilizes the set of measurements used for state estimation to predict the actual network topology and offers low computational times along with high accuracy under a wide variety of testing scenarios. The training process of the TI DNN is duly discussed, and several deep learning heuristics that may be useful for similar implementations are provided. Simulations on the IEEE 14-bus and IEEE 39-bus test systems are reported to demonstrate the effectiveness and the small computational cost of the proposed methodology.**

***Index Terms*—Topology identification, deep neural network, state estimation, bad data detection and identification.**

I. Introduction

POWER system dynamics are becoming increasingly complex due to the integration of nonsynchronous generation, energy storage devices and demand response technologies. As a result, new techniques of dynamic state estimation (DSE) are being developed to monitor the dynamics of electrical variables and improve the control and protection of power systems [1] and [2].

Deployment of phasor measurement units have enabled the development of fast DSE methods that require equally fast complementary tools such as topology identification (TI) algorithms. Significant efforts to properly identify topology changes in power systems have been made over the last several decades. Power system state estimation relies on the perfect a priori knowledge of the network topology provided by the network topology processor (NTP), which analyzes the status of the switching devices and provides the corresponding bus/branch model to the state estimator. However, in many cases, the NTP can be affected by the errors related to the status of the switching devices, which can result in the use of a wrong admittance matrix by the state estimator. Topology errors tend to have a stronger influence than network parameter errors and can cause the state estimation process to be strongly biased.

The residual analysis method reported in [3] assumes that the first state estimation iteration converges successfully so that the residual vector can be computed and used to identify the branch outage. This assumption is not always true because the state estimation convergence is not guaranteed due to the dramatic impact that the topology errors tend to have on the measurement residuals.

The state vector augmentation method, which is also described in [3], includes the branch status in the state vector and requires several state estimation iterations to adjust the state vector, to add the model constraints and, eventually, to augment the state vector. The iterations required to detect the changes in the topology can be a limiting factor for DSE applications because the computational times of the algorithm increase.

In [4], a fuzzy c-means clustering method is proposed for TI and bad data processing. Using this method, the fuzzy pattern vector expands with the size of the network, and the number of possible topologies to be considered increases, causing a loss of resolution that affects the accuracy of the TI and the state estimation.

An event-triggered topology identification is proposed in [5], where a recursive Bayesian approach is used when a network topology change is suspected. This approach is precise and reliable if the topology configurations to be estimated are limited, and it is sensitive to high measurement noise levels. However, the reported computational times are too long for DSE applications.

A generalized state estimation algorithm for topology estimation that includes the switching devices status in the state vector is reported in [6]. This approach requires the incorporation of three additional state variables for each switching device, which will significantly increase the computational burden of the algorithm for large electric networks.

The authors of [7] present a method for topology identification in the generalized state estimation framework. In the first stage, a bad data analysis is used to identify the region of the electric network affected by the error. Second, a network reduction is conducted to restrict the analysis to the suspected area. Finally, the Lagrange multiplier of the state estimation constraints are used to select the type of the anomaly. The reported results indicate that the methodology is accurate, but the CPU times are well above one second, making it unsuitable for DSE applications.

In [8], a measurement-based approach to directly estimate the electric network bus admittance matrix is proposed. This methodology requires the monitoring of each bus voltage phasor through a phasor measurement unit (PMU) or a modern relay capable of providing the voltage phasor measurement. This condition is currently not fulfilled in many power systems.

For the usage of neural networks in TI applications, pioneering work has been conducted in [9] and [10]. In [9] a counter propagation network (CPN) for topology processing is proposed. The results show a certain feasibility for online

The authors are with the Department of Electrical Engineering, Universidad Carlos III de Madrid, Madrid, Spain (e-mail: dgotti@ing.uc3m.es; hortensia.amaris@uc3m.es; pablole@ing.uc3m.es)

© 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Power Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TPWRS.2021.3076671

applications, as reported outcomes exhibit computational times lower than 100 ms. However, the CPN needs accurate information of the switching devices status along with power flow and injection measurements. In this paper, the proposed TI algorithm relies only on the measurement set used for the state estimation and does not need the information of the circuit breakers status, making the algorithm more reliable in case of wrong information from the NTP.

In [10], an artificial neural network (ANN) approach is used for TI. The algorithm can distinguish between topological and gross measurement errors by means of normalized innovations, which are obtained adding a previous state forecasting step to the state estimation algorithm. This methodology implies the implementation of four ANNs for each electrical network branch, which in the case of large-scale systems entails the implementation of a massive number of ANNs, leading to a cumbersome training process. With the approach proposed in this work, the topology determination relies only on one deep neural network (DNN) and does not need any pre-estimation step that implies an increase in the computational cost.

As previously mentioned, the TI must be reliable and fast to enable the DSE to operate in a short time frame. In this regard, the abovementioned algorithms are not suitable for DSE applications due to their high CPU times. The methodology proposed in this work aims to bridge this gap.

The contribution of this work is a TI algorithm that overcomes the limitations previously mentioned. Deployment of DNNs and modern deep learning techniques, that are duly discussed in the next sections, accurately extract the relations between the set of measurements and the topology information without the necessity to receive the switching device status from the NTP or to implement several ANNs for each branch of the electrical network.

The methodology proposed in this work introduces a significant reduction of the computational time required to perform the topology identification. The proposed algorithm is suitable for DSE applications and provides a fast identification of the network topology using the measurements collected for state estimation. The proposed TI algorithm is based on a deep neural network and relies on a set of measurements free of gross errors. To demonstrate the TI algorithm, a bad data detection and identification algorithm, and a state estimator are coupled with the proposed method. For this purpose, a modified version of the algorithm introduced in [11] is used as a bad data detection and identification algorithm even though others methods can be used, e.g., [12] and [13]. Once the measurement set has been checked and possibly modified, the topology configuration is estimated, and subsequently, a network state estimation is performed using an unscented Kalman filter (UKF). However, it is important to note that other state estimation algorithms are compatible with the proposed TI DNN.

The main advantages of the proposed method are:

- It uses only the measurements required for the state estimation and it requires a single DNN.
- It is suitable for DSE applications as it significantly reduces the computational time required to perform the topology identification in comparison with current methods.

The rest of the paper is organized as follows. Section II presents a brief explanation of both the DNNs and the UKF state estimator used in this work. In Section III the proposed methodology is described. Section IV provides the case study description, with a particular focus on the DNN training and testing process. Section V shows how the proposed TI DNN integrates with the state estimator and relative results on the IEEE 14-bus and the 39-bus test networks are reported. Finally, Section VI concludes the paper.

II. Description of the algorithms

A. Deep Neural Network

DNNs are defined as ANNs with multiple hidden layers between the input and output layers. The presence of multiple hidden layers allows the network to learn complex tasks by extracting significant features that enable input-output value mapping.

During the training of a feed-forward neural network, the input signal is propagated forward, and then, the obtained output is compared with the desired output values, and an error signal is computed. This error is then backpropagated using the so-called backpropagation algorithm, which allows the network to adjust its parameters to minimize the cost function determined by the error signal. During the training phase, the process is repeated until a desired accuracy on the training set is achieved. As described in [14], the training can be executed in sequential mode (one training example at a time is introduced) or in batch mode (all the training examples are introduced simultaneously).

Feed-forward propagation: During the forward computation, the input patterns propagate forward through the network and appear at the output end as an output signal. The input signal propagates from one neuron to the next layer of neurons passing through the synaptic weights. The signal is processed as follows:

$$\boldsymbol{z}_i = \boldsymbol{W}_{i,i-1}\boldsymbol{a}_{i-1} + \boldsymbol{b}_i, \tag{1}$$

where $\boldsymbol{a}_{i-1}$ is a vector representing the previous layer outputs, $\boldsymbol{W}_{i,i-1}$ represents the synaptic weights matrix of (n, m) dimensions (with n and m corresponding to the number of neurons in the i_{th} and $i-1_{th}$ layer, respectively), $\boldsymbol{b}_i$ represents the vector of the neuron bias values and $\boldsymbol{z}_i$ represents the i_{th} layer input vector.

Subsequently, each neuron performs the computation of the activation function, usually expressed as a nonlinear function of the input signal:

$$\boldsymbol{a}_i = \boldsymbol{f}_i(\boldsymbol{z}_i), \tag{2}$$

where $\boldsymbol{f}_i$ is the neuron activation function and $\boldsymbol{a}_i$ is the i_{th} layer output vector.

The choice of the activation function can radically modify how changes in the weights and bias during the training phase affect the variations in the output values. Once the input vector is propagated through the network, an error signal is computed:

$$\boldsymbol{e}_j = \boldsymbol{\zeta}(\boldsymbol{y}_j, \boldsymbol{o}_j), \tag{3}$$

© 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Power Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TPWRS.2021.3076671

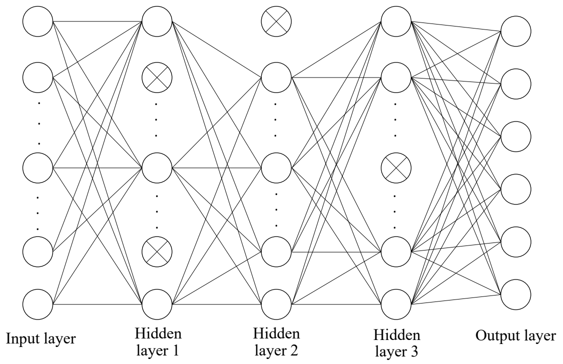

Fig. 1. A deep neural network with 3 hidden layers applying dropout (crossed units represent dropped neurons)

where $\boldsymbol{e}_j$ represents the error vector of the j_{th} output neuron, and $\boldsymbol{\zeta}$ is a generic function that relates the output error with the obtained output vector $\boldsymbol{y}$ and the desired output vector $\boldsymbol{o}$ known from the training set.

A variety of activation functions and error signals formulations and their characteristics are thoroughly described in [15].

Backpropagation: The backpropagation algorithm applies a correction to the synaptic weights and neuron bias that is proportional to the partial derivative of the error signal with respect to the synaptic weights and bias. These partial derivatives represent a sensitivity factor, which allows the determination of the change in the weights and bias hyperspace that minimize the cost function, i.e., the error signal. In the case of the output layer, the partial derivative can be directly computed because there exists a direct relationship between the error signal and the neurons belonging to this layer:

$$\frac{\partial \boldsymbol{e}}{\partial \boldsymbol{W}_{k,k-1}} = \frac{\partial \boldsymbol{e}}{\partial \boldsymbol{y}_k}\frac{\partial \boldsymbol{y}_k}{\partial \boldsymbol{z}_k}\frac{\partial \boldsymbol{z}_k}{\partial \boldsymbol{W}_{k,k-1}}, \tag{4}$$

where k represents the output layer index. In the case of the hidden layers neurons, the chain rule is adopted to express this gradient:

$$\frac{\partial \boldsymbol{\xi}_i}{\partial \boldsymbol{W}_{i,i-1}} = \frac{\partial \boldsymbol{\xi}_i}{\partial \boldsymbol{e}_i}\frac{\partial \boldsymbol{e}_i}{\partial \boldsymbol{a}_i}\frac{\partial \boldsymbol{a}_i}{\partial \boldsymbol{z}_i}\frac{\partial \boldsymbol{z}_i}{\partial \boldsymbol{W}_{i,i-1}}, \tag{5}$$

where ξ represents the error signal propagated through the neural network up to the considered i_{th} hidden layer. Once the partial derivative for a layer is computed, the weights are updated:

$$\boldsymbol{W}_{i,i-1} = \boldsymbol{W}_{i,i-1} - l_r\frac{\partial \boldsymbol{\xi}_i}{\partial \boldsymbol{W}_{i,i-1}}, \tag{6}$$

where l_r is the learning rate that permits the adjustment of the trajectory rate of change in the weight and bias space.

The same procedure is executed for bias actualization.

Dropout: When the training data are limited, DNNs may experience overfitting problems; i.e., a good performance is achieved during the training process, but poor results are obtained for the test set. This issue is caused by the fact that the derivative received by each unit gives an indication of how the synaptic weights and neurons bias should change to reduce the cost function considering how all other components are acting. Thus, units may vary to fix the errors generated by other units, leading to complex co-adaptation phenomena.

The dropout technique introduced in [16] gives a solution to this problem. This approach consists of stochastically dropping neurons and their synaptic weights at each training iteration. Thus, co-adaptations are prevented since the compensating effect of other units is not certain. In fact, under these random changes, every unit must perform well under a wide variety of structural configurations. Fig. 1 shows the structure of a DNN after applying the dropout technique.

B. Unscented Kalman Filter

The UKF is a nonlinear version of the well-established Kalman filter, described in [17]. It belongs to the wider class of sigma-point Kalman filters that make use of the statistical linearization technique known as unscented transformation [18].

The following nonlinear system is considered with a set of states to be estimated and an observation model with additional noises:

$$\boldsymbol{x}_k = \boldsymbol{f}(\boldsymbol{x}_{k-1}) + \boldsymbol{w}_{k-1} \tag{7}$$

$$\boldsymbol{z}_k = \boldsymbol{h}(\boldsymbol{x}_k) + \boldsymbol{v}_k, \tag{8}$$

where $\boldsymbol{x}$ represents the state vector, $\boldsymbol{f}(x)$ is the state estimation function, and $\boldsymbol{h}(x)$ is the observation function that relates the measurement set z_k with the state vector x_k at instant k. Variables $\boldsymbol{w}$ and $\boldsymbol{v}$ are the process and measurement noises, respectively. Similar to the other classic Kalman filter versions, the UKF algorithm executes two main steps: the estimation step and the update step.

Estimation step: Once the sigma points and their weights are calculated as indicated in [18], they are propagated through the nonlinear estimation function, and the predicted states are computed as follows:

$$\widehat{\boldsymbol{x}}_{k,k-1} = \sum_{i=0}^{2n} W_i \boldsymbol{\chi}_{i,k,k-1}, \tag{9}$$

where $\boldsymbol{\chi}_i$ and W_i represent the i_{th} sigma point value and weight, respectively. Then, the estimate covariance matrix is calculated as:

$$\boldsymbol{Pxx}_{k,k-1} = \sum_{i=0}^{2n} W_i[\boldsymbol{\chi}_{i,k,k-1} - \widehat{\boldsymbol{x}}_{k,k-1}][\boldsymbol{\chi}_{i,k,k-1} - \widehat{\boldsymbol{x}}_{k,k-1}]^T \tag{10}$$

Propagating the sigma points through the observation function, the same procedure is applied to compute the observation mean $\widehat{\boldsymbol{z}}_{k,k-1}$ and the covariance matrix $\boldsymbol{Pzz}_{k,k-1}$. Then, following a similar approach, the cross-covariance matrix $\boldsymbol{Pxz}_{k,k-1}$ is determined.

Update step: Once a new measurement set becomes available, the update step is applied:

$$\boldsymbol{K}_k = \boldsymbol{Pxz}_{k,k-1}\boldsymbol{Pzz}_{k,k-1}^{-1} \tag{11}$$

$$\boldsymbol{Pxx}_{k,k} = \boldsymbol{Pxx}_{k,k-1} - \boldsymbol{K}_k\boldsymbol{Pzz}_{k,k-1}\boldsymbol{K}_k^{-1} \tag{12}$$

$$\boldsymbol{x}_{k,k} = \boldsymbol{x}_{k,k-1} + \boldsymbol{K}_k[\boldsymbol{z}_k - \widehat{\boldsymbol{z}}_{k,k-1}] \tag{13}$$

A full description of UKF development can be found in [19]. This state estimator has been chosen for this work because, as reported in [20], it shows better performance in terms of

© 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Power Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TPWRS.2021.3076671

accuracy than other state estimation algorithms such as the weighted least square or the extended Kalman filter. However, as pointed out previously, other state estimation algorithms are compatible with the proposed TI methodology.

III. Method description

The main contribution of this work is to provide a fast and reliable TI algorithm. The methodology is based on DNNs and the TI is performed using the same set of measurements used by the state estimator. Fig. 2 illustrates how the proposed TI DNN method is integrated into a general state estimation problem.

The proposed TI method uses a feed-forward DNN as described in Section II and is formulated to solve a classification problem. The DNN input is the measurement set, consisting of a number of active power, reactive power, voltage magnitude and voltage angle measurements. Before the inputs are fed to the DNN, they are normalized with the following procedure:

$$z_{i_{normalized}} = \frac{z_i - z_{min}}{z_{max} - z_{min}}, \tag{14}$$

where z_i is the i_{th} measurement to be normalized, and z_{max} and z_{min} are the maximum and the minimum values, respectively, in the normalizing range. Thus, the normalized z_i assumes a value between 0 and 1. The normalization is applied separately for power flow/injection measurements, voltage magnitude measurements and voltage angle measurements. This procedure has proved to be very effective for avoiding the generalization problems and accelerating the training process [15]. The output neurons of the DNN have a binary formulation, in which each possible set of output values is associated with a topology configuration. The topology configurations considered for both test systems of Section IV represent the loss of every single branch, according to the N-1 outage criterion. They are reported in [21].

The TI DNN contains several layers, whose number and size depends on the specific power system. For both the test systems presented in Section IV, hidden and output layer neurons use the following sigmoid activation function:

$$a_i = \frac{1}{1 + e^{-z_i}} \tag{15}$$

The error signal used during the training phase is expressed as follows:

$$e_i = (y_i - o_i)^2 \tag{16}$$

The proposed TI DNN allows an accurate mapping of the relations between the measurements and the topology configuration. Once properly trained, the TI DNN is able to accurately predict the network topology even under network operation scenarios for which the TI DNN has not been previously trained.

In order to demonstrate the proposed methodology, the TI DNN is combined with a bad data detection and identification algorithm that ensures that the set of measurements is free of gross errors. Once the TI has been carried out, the corresponding admittance matrix is loaded into the state estimator observation function and the power system states are calculated.

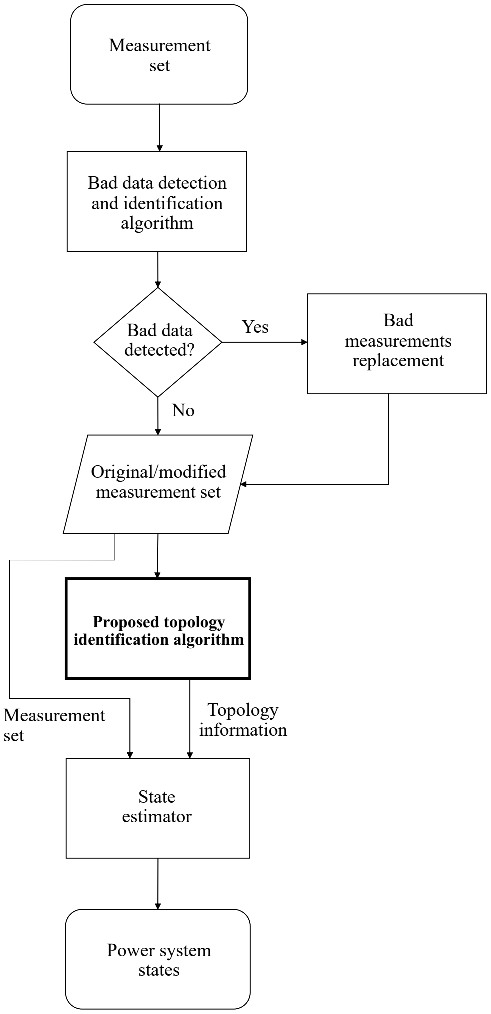
Measurement set
Bad data detection and identification algorithm
Bad data detected?
Yes
No
Bad measurements replacement
Original/modified measurement set
Proposed topology identification algorithm
Measurement set
Topology information
State estimator
Power system states

Fig. 2. Flowchart of the proposed methodology

IV. Case study

The algorithms described in the previous sections are tested on the IEEE 14-bus and IEEE 39-bus test systems. In this section, the structure of the DNN used for the TI, the bad data detection and identification and the calibration of the UKF state estimator for both test networks are reported. Particular attention is directed at the election of the structure of the TI DNN and the creation of the training set, which are extremely important to properly estimate the electric network topology.

The set of 50 and 127 measurements used for the IEEE 14-bus system and the IEEE 39-bus system are listed in [21] and are also shown in Figs. 3 a) and 3 b), respectively. In the case of the IEEE 14-bus system, to prove how the proposed TI DNN performs with different set of measurements, two distinct set of 40 measurements are also used to compare the testing accuracy and relevant considerations are made in this section. All the measurements used for TI, bad data detection and identification, and state estimation are generated using the software PowerFactory. The standard PowerFactory synchronous machine model 2.2 is employed, and voltage- and frequency-dependent loads are considered. A normally distributed random error with zero mean and a 3σ standard

© 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Power Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TPWRS.2021.3076671

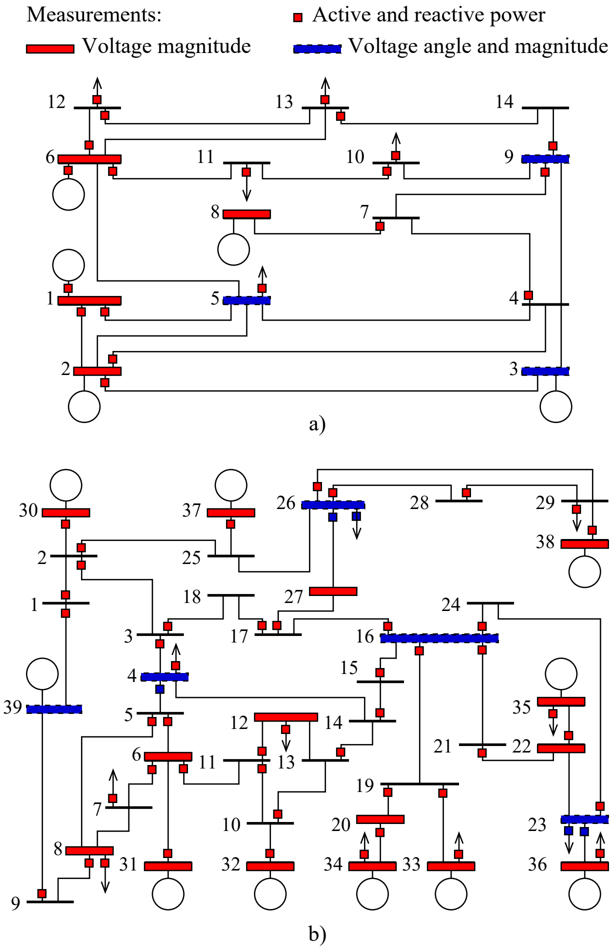

Fig. 3. Single-line diagram and measurement set of a) the IEEE 14-bus test system, and b) the IEEE 39-bus test system.

deviation, corresponding to the measurement accuracy, is added to the measurements. The accuracy values used in this work are described in [22].

All the algorithms presented in this work have been coded in MATLAB. The computer used in this work is an Intel(R) Core(TM) i7-3770 CPU@3.40 GHz 3.40 GHz processor, with 6 GB RAM memory.

A. DNN for Topology Identification

Different DNN structures for both test systems have been attempted, and the final DNN structure for the two test systems is reported below.

TABLE I
ACCURACY OF DIFFERENT DNN STRUCTURES - IEEE 39-BUS SYSTEM

	DNN 90-60-30 no dropout	DNN 75-25-10 no dropout	DNN 75-25-10 dropout 1-0.9-0.95-1	DNN 90-30-10 dropout 1-0.9-0.95-1	DNN 90-30-10 dropout 1-0.7-0.9-1
Training accuracy	98.3%	99.9%	99.3%	98.7%	99.7%
Testing accuracy	84.58%	87.95%	89.57%	95.36%	99.92%

- IEEE 14-bus System
  The TI DNN in the case of the IEEE 14-bus system is structured with three hidden layers, with 40, 20 and 10 neurons, separately. There are 50 input neurons corresponding to the number of measurements used for the state estimation, and there are 5 output neurons because the number of the considered topology configurations is 21 and thus can be expressed as a 5-digit binary number. The learning rate used during the training phase is $lr = 10^{-6}$. No dropout procedure was necessary to correctly identify the topology changes during the testing phase. The training time of this DNN is approximately 12.5 hours, which is consistent with other computational times indicated in [23].
- IEEE 39-bus system
  In the case of the IEEE 39-bus system, the DNN has also three hidden layers with 90, 30 and 10 neurons, separately. There are 127 input neurons and 6 output neurons, as the number of possible topology configurations is 47. The learning rate adopted in this case is $lr = 2 \cdot 10^{-6}$. In the present case, an inverted dropout algorithm is applied since the DNN trained without dropout shows co-adaptation phenomena that compromise the test performance. The keeping probabilities used for the input layer, the first, the second and the third hidden layers are 1, 0.7, 0.9 and 1 respectively. Lower keeping probabilities lead to no-convergence problems during the training phase, and the adoption of higher keeping probabilities is not effective against co-adaptation, as shown in Table I. Useful heuristics for effective dropout application can be found in [24]. The training phase of this DNN was carried out for approximately 94 hours. This value is consistent with other training times reported in [23].

In both test systems, the sigmoid activation function is used, as shown in (15), and the error signal has the same formulation of (16). Other activation functions such as the rectified linear unit and the hyperbolic tangent have also been tested, but the performance obtained during the testing phase with the sigmoid function was significantly superior, and the training time was shorter.

Some of the attempts realized to calibrate the TI DNN for the IEEE 39-bus test system are reported in Table I, where the first three numbers represent the number of neurons of the hidden layers and the last numbers represent the keeping probability used during the dropout procedure. Although with all the represented structures, a good estimation on the training set is achieved, the test accuracy is strongly influenced by the DNN structure and the use of the dropout procedure, as can be easily observed from Table I. In fact, it is found that the DNN with the chosen structure offers better accuracy during the testing. In particular, the proposed DNN was able to correctly detect all of the considered topology configurations reported in [21] in almost every topology estimation. DNNs with a lower number of hidden layers have been tested, but their structure proved to be insufficient for mapping the relations between the upcoming measurements and the actual network topology.

Extensive simulations on both test systems show that in order to correctly extract the relationship between the mea-

© 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Power Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TPWRS.2021.3076671

surements values and the actual topology configuration, a wide variation of the input values during the DNN training process is essential. Hence, for the 14-bus system and the 39-bus system, different generation and load profiles variations were simulated so that during the training process, the bus voltage magnitude fluctuates from approximately 0.75 p.u. to 1.25 p.u. and the current flows from 0 to 130% of the line rated capacity. These variations have proved to be of pivotal importance to effectively extract the input-output mapping. In this manner, the DNN is trained so that its estimations remain reliable even outside the training set. This training process is generated by modeling the previously mentioned generation and loading events for each topology configuration, i.e., for each branch placed out of service. For every topology configuration, approximately 350 training examples are used on both test systems to train the DNN using batch mode [14]. Overall, the DNN for the 14-bus test system employs 7470 training samples, and the DNN for the 39-bus system is trained with 16590 samples. In both test systems, the DNN synaptic weights and bias are initialized using independent normalized Gaussian random variables with zero mean and unit variance.

To observe how the proposed method performs with lower global measurement redundancy and to analyse its dependence from the type of measurements, two TI DNNs are implemented for the IEEE 14-bus system, both with a measurement set composed of 40 elements. In both cases, 10 measurements are withdrawn from the set of 50 measurements reported in [21]. In the first set of measurements, namely type 1 in Table II, 10 active and reactive power injection measurements (relative to bus 5, 6, 10, 12 and 13) are removed from the set of 50 element. In the second set, namely type 2 in Table II, 10 active and reactive power flow measurements are withdrawn from the set of 50 measurements (relative to branch 1-2, branch 2-4, branch 6-11, branch 10-11 and branch 12-13). In this manner, it is possible to observe how the accuracy achieved during the testing phase depends on the measurement type. Results reported in Table II indicate that the accuracy obtained during the training phase with the two reduced sets of measurements is comparable with the training accuracy achieved with the set of 50 elements. Nevertheless, the testing accuracy varies significantly. As can be observed in Table II, in the case of the type 1 set of measurements the testing accuracy decreases only to 95.26%, whereas in the case of the type 2 the accuracy strongly decreases to 86.69%. This result indicates that power flow measurements tend to carry more relevant information than power injection measurements for the topology identification process. In fact, this is quite intuitive, as the former type of measurements are more related to the topology configuration of the network. Nevertheless, the TI DNN trained with 50 measurements is the one which shows better accuracy, because power injection measurements also provide information for the topology identification process. In conclusion, these results suggest that both the number of measurements and their type have an impact on the effectiveness of the proposed method.

For both the IEEE 14-bus and 39-bus systems, the testing phase is conducted considering the sudden outage of a branch, and the simulation is repeated for all the network branches. The simulation lasts 5 s, a short-circuit is simulated at t=1 s and the affected branch is placed out of service after 100 ms. The simulation time step used is 33 ms, and thus, the TI DNN is called to estimate the network topology 150 times for each simulation. The procedure is repeated for all of the considered topology configurations.

TABLE II
ACCURACY WITH DIFFERENT SET OF MEASUREMENTS - IEEE 14-BUS SYSTEM

	Set of 50 measurements	Set of 40 measurements - type 1	Set of 40 measurements - type 2
Training accuracy	99.32%	99.24%	99.17%
Testing accuracy	99.59%	95.26%	86.69%

B. Bad data detection, identification and replacement

To correctly estimate the actual network topology and successively execute the state estimation, gross errors in the measurement set must be detected, identified and replaced. Thus, a bad data detection and identification algorithm is needed. In this work, the methodology proposed in [11] is modified and implemented, although other algorithms can be used. This algorithm is based on an ANN and has been implemented for the IEEE 39-bus system to prove its compatibility with the proposed TI identification algorithm. However, it must be noted that instead of using 1 hidden layer, as detailed in [11], the neural network used in this work is structured with 2 hidden layers because it shows better performance during testing. Since the measurement set for the IEEE 39-bus system has 127 elements, the neural network is provided with 127 input neurons and 127 output neurons, while both hidden layers are equipped with 80 neurons. With this structure, the neural network shows good performance during the test phase. Similar to the DNN for TI, the input set is normalized using (14) and the training set includes wide variations in the input measurements. The activation function used for the hidden layers and output layer neurons is:

$$a_i = \frac{1-e^{-z_i}}{1+e^{-z_i}} \tag{17}$$

Similar to the TI algorithm, the synaptic weights and neurons bias are initialized with Gaussian random values with zero mean and unit variance. No dropout procedure is implemented. The identification rule used in this work has the same formulation proposed in [11] and is expressed as follows:

$$(z_i - \theta_i)^2 > {r_i}^2, \tag{18}$$

where z_i and θ_i are the i_{th} measurement and its estimated value, respectively, and r_i is the desired threshold that flags the bad measurement input. The identification threshold applied in this work is $r_i = 10 \cdot \sigma_i$, where σ_i is the i_{th} measurement standard deviation.

C. State estimation algorithm based on UKF

The diagonal initial estimate covariance matrix $\boldsymbol{P}_{0,0}$ of the UKF is composed by the initial state variances of 0.05^2 and 0.01^2, corresponding to the voltage magnitudes and voltage

© 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Power Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TPWRS.2021.3076671

angles, respectively. The diagonal process noise covariance matrix $\boldsymbol{Q}$ is formed by the terms of 0.005^2 and 0.05^2, corresponding to the voltage magnitude and angle process variances, respectively. In accordance with the measurements variances described above, the diagonal measurement uncertainty matrix $\boldsymbol{R}$ is determined. The initial state vector $\boldsymbol{x}_{0,0}$ is derived from a flat start hypothesis; i.e., the voltage magnitude and voltage angle states are initialized with ones and zeros, respectively. All of the previous assumptions are made for the state estimators implemented in both the IEEE 14-bus and 39-bus test systems.

The index used to evaluate the accuracy of the state estimator is the root mean square (RMS) of the residuals Ψ, which is calculated separately for the voltage magnitudes and angles as shown in [20]. The observation function $\boldsymbol{h}(\boldsymbol{x})$ is determined using the same procedure described in [20].

V. Test Results

As described in the previous section, to test the effectiveness of the proposed TI algorithm, all of the topology configurations reported in [21] have been simulated and successfully detected during the testing phase. In this section, some study cases are reported to show how the integration of the proposed TI DNN can benefit the state estimation process and to show its DSE applicability due to its small computational cost. All of the cases reported in this section are implemented with the set of measurements described in [21]. The global measurement redundancy used for the IEEE 14-bus system and the IEEE 39-bus system is 1.79 and 1.63, respectively. In all of the simulations cases, the time step is 20 ms and a random Gaussian noise as described in Section IV is applied to the measurement set.

A. Case 1: Topology change in the IEEE 14-bus system

In this case, a topology change in the IEEE 14-bus test system is studied. Specifically, a three-phase short-circuit current with zero fault impedance is simulated at the middle of line 2-4 at t=0.5 s, and the fault is cleared 100 ms later by opening the line. The simulation is carried out for 10 s. Once the faulted branch opens, a topology variation occurs. Consequently, some elements in the network admittance matrix that are used in the observation function equations vary. If the topology change is not detected and the admittance matrix values are not updated accordingly, a biased estimation occurs. Usually, the most affected states from this error are the voltage magnitudes and angles of the bus connected to the affected branch. For this reason, the voltage angles at the buses 2 and 4 are reported in Fig. 4 and Fig. 5. An examination of the figures shows that the TI DNN can detect the correct topology after the disconnection of line 2-4 and avoid a biased estimation. It is observed that the algorithm improves not only the steady-state estimation but also the estimation of the transient behavior.

Table III shows the RMS residuals of the state estimation. It is observed that the estimation of the voltage angles has improved significantly with the correct topology information, while the improvement in the voltage magnitude residuals is marginal. This result can be explained based on observation of the measurement set reported in [21]. In fact, the bus 2 voltage magnitude is measured and does not need to be estimated through a load-flow formulation in the observation function that relies on the topology information. While the bus 4 voltage magnitude is not measured, the adjacent bus 2, 3, and 5 voltage magnitudes are monitored, leading to a lower dependence from the topology configuration information. These results are consistent with the considerations made in [8], in which the advantages of using a measurement-based approach in state estimators are highlighted.

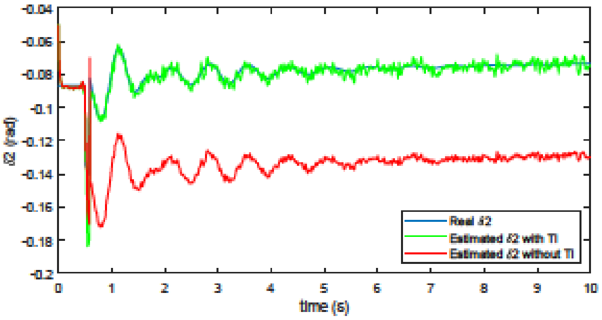

Fig. 4. Case 1-$\delta 2$ estimation with and without TI

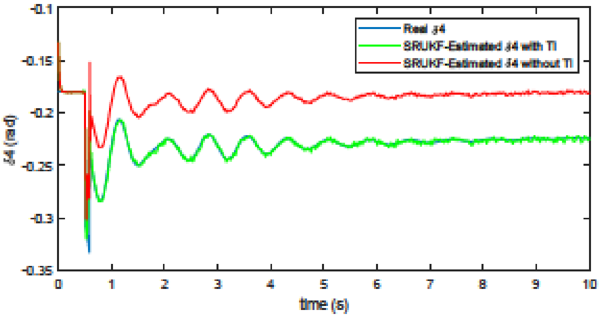

Fig. 5. Case 1-$\delta 4$ estimation with and without TI

TABLE III
Case study 1

Comparison index	Estimation with topology identification	Estimation without topology identification
Ψ_δ (rad)	$5.929 \cdot 10^{\quad 3}$	$2.640 \cdot 10^{\quad 2}$
Ψ_V (-)	$1.120 \cdot 10^{\quad 2}$	$1.325 \cdot 10^{\quad 2}$

In addition to excellent precision over a wide range of testing simulations, the algorithm proposed in this work showed a computational time of approximately $1.3 \cdot 10^{-4}$ s for each topology estimation on this 14-bus test system, thus demonstrating its feasibility for DSE applications, in contrast to the other algorithms mentioned above in Section I.

For this case study, a comparison in terms of computational time with the residual analysis [3] is reported. The residual analysis method exhibits computational times of approximately 0.98 s, whereas the proposed approach processes the topology identification in $1.3 \cdot 10^{-4}$ s. This clearly indicates that the proposed method is significantly faster and suitable for DSE applications. It is worth noticing that, due to the reasons mentioned in Section I, also other TI algorithms perform

© 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

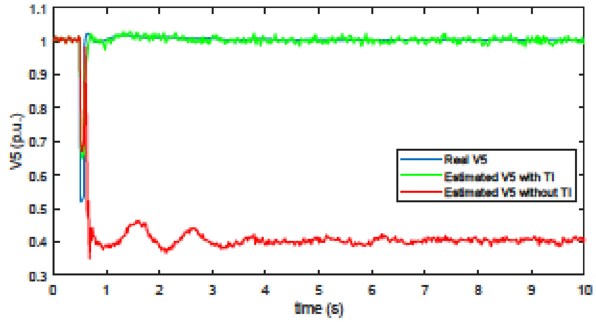

Fig. 6. Case 2-V_5 Estimation with and without TI

TABLE IV
CASE STUDY 2

Comparison index	Estimation with topology identification	Estimation without topology identification
Ψ_δ (rad)	$6.351 \cdot 10^{3}$	$4.011 \cdot 10^{2}$
Ψ_V (-)	$1.173 \cdot 10^{2}$	$1.232 \cdot 10^{1}$

with worse CPU times compared with the proposed TI DNN method, as their computational burden is sensibly higher.

B. Case 2: Topology change in the IEEE 39-bus system

In this case, a three-phase short-circuit current at the middle of line 5-8 is considered. The fault occurs at t=0.5 s, and the line circuit breaker opens 100 ms later. The simulation is carried out for 10 s. In this case as well, the most affected states from the topology change are those adjacent to the faulted branch, i.e., bus 5 and 8 states. While the presence of the voltage measurement device at bus 8 permits the UKF to estimate with a good accuracy even without a correct TI, the absence of a measurement device placed at the bus 5 leads the state estimator to lose track on this state, as can be readily observed from Fig. 6. Similar to the previous study case, with the proposed TI DNN, the states are estimated correctly, and a biased estimation is avoided. Table IV shows the RMS residuals of this case. It is observed that the presence of the TI DNN algorithm reduces the RMS residuals by approximately one order of magnitude since it prevents the UKF state estimator from losing track of several states, specifically those that are not measured.

In this case, the TI DNN requires a computational time of approximately $1.4{\cdot}10^{-4}$ s. It is observed that although a denser DNN structure with respect to the previous study case is used, the computational time for the topology estimation remains almost equal and is perfectly suitable for DSE applications. The UKF state estimator requires approximately $16.9 \cdot 10^{-3}$ s to perform an estimation on this test network; thus, the integration with the TI algorithm does not significantly increase its computational burden, and its online applicability is not affected. The CPU time comparison of both the state estimators and the TI DNNs implemented for the two test systems is presented in Table V.

TABLE V
CPU TIME COMPARISON OF THE STUDY CASES

	TI DNN estimation time (s)	State estimation time (s)
IEEE 14-bus test system	$1.3 \cdot 10^{4}$	$3.3 \cdot 10^{3}$
IEEE 39-bus test system	$1.4 \cdot 10^{4}$	$16.9 \cdot 10^{3}$

TABLE VI
CASE STUDY 3-GROSS MEASUREMENT REPLACEMENTS

Measurement	Wrong Value (p.u.)	Estimated Value (p.u.)	True value (p.u.)
$P_{9\ 39}$	-0.40	-0.18	-0.20
$Q_{9\ 39}$	-0.82	-0.62	-0.69
$P_{23\ 24}$	3.13	3.53	3.53
$Q_{23\ 24}$	0.20	0.38	0.50
$P_{16\ 19}$	-5.82	-5.12	-5.02
$Q_{16\ 19}$	-0.88	-0.46	-0.48

C. Case 3: Gross errors in the measurement set on the IEEE-39 bus system

In this case, the importance of coupling a bad data detection and identification algorithm with the DNN TI algorithm is demonstrated. This simulation is carried out for 1 s and from t=0, 1 s until the end of the test, 6 of the 127 measurements are supposed to be affected by a gross error, as indicated in Table VI. These gross errors are correctly detected, identified and replaced by the bad data detection and identification algorithm. An examination of the data presented in Table VI shows that the replaced measurements are not perfect but are still relatively close to their true values. Using these values, the TI DNN can correctly continue processing the topology configuration. As pointed out in [11], under some circumstances, these measurement estimations may not be very accurate. A possible solution to solve this problem is to replace the identified bad data with pseudo-measurements, using the methodology proposed in [25].

If the bad data detection algorithm is not considered, the bad data measurements fed to the DNN can provoke a false topology configuration detection. In this particular case, the TI DNN, instead of classifying the topology with the binary output of "000000" (i.e., all the branches are connected [21]), detects a "011000" topology, wrongly assuming that line 23-24 opened and consequently leading to a biased estimation of the voltage angle at bus 24, as observed in Fig. 7. In this case, the estimation error is relatively small, but as shown in the previous study case, a wrong topology processing can lead to a strong biased estimation of some states.

This case illustrates that a fast bad data detection and identification algorithm that replaces wrong measurements is a complementary element to the proposed TI DNN algorithm. The CPU cost of the bad data detection and identification algorithm is approximately $3{\cdot}10^{-4}$ s, showing execution time compatibility with the proposed TI DNN.

VI. CONCLUSION

This paper proposes a fast TI algorithm suitable for DSE applications based on DNNs. For the purpose of demonstration, the algorithm is applied to the IEEE 14-bus and 39-bus

© 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Power Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TPWRS.2021.3076671

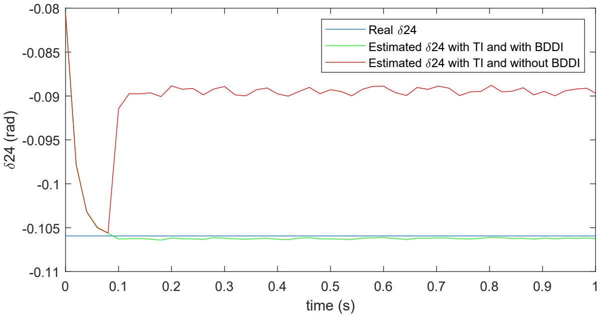

Fig. 7. Case 3-δ24 Estimation with TI and with and without the bad data detection and identification algorithm (BDDI)

test systems while considering the disconnection of any single branch in the system; however, the same formulation can be applied to areas of larger power systems or to other changes in the topology.

The algorithm shows good performance both in terms of reliability and computational time. Regarding the reliability, the results on the IEEE 39-bus system show an accuracy higher than 99% during the testing phase for a testing set of 7050 samples. With respect to the speed, the computational times of the proposed methodology are quite low (approximately 10^{-4} s), making the TI DNN suitable for power systems DSE applications in the area of control and protection. To the best of our knowledge, no other published work has demonstrated a topology processing method with such a small computational cost. In fact, all of the algorithms briefly introduced in Section I present CPU times from 10^{-1} s up to few seconds, limiting their DSE applications.

It is worth noticing that in case a new branch is added in the system, the TI DNN should be re-trained accordingly to take into account new possible topology configurations. However, transmission power systems experience network extensions only occasionally, and these changes are usually planned in advance. Thus, the re-training process does not represent a limiting factor for the application of the proposed methodology.

Finally, the importance of using a set of measurements free of gross errors has been demonstrated. Hence, the proposed algorithm is coupled with a fast bad data detection and identification algorithm that substitutes wrong measurements with estimated values.

REFERENCES

[1] J. Zhao, A. Gómez-Expósito, M. Netto, L. Mili, A. Abur, V. Terzija, I. Kamwa, B. Pal, A. K. Singh, J. Qi, Z. Huang, and A. P. Meliopoulos, "Power System Dynamic State Estimation: Motivations, Definitions, Methodologies, and Future Work," *IEEE Transactions on Power Systems*, vol. 34, no. 4, pp. 3188–3198, jul 2019.

[2] Y. Cui, R. G. Kavasseri, and S. M. Brahma, "Dynamic State Estimation Assisted Out-of-Step Detection for Generators Using Angular Difference," *IEEE Transactions on Power Delivery*, vol. 32, no. 3, pp. 1441–1449, jun 2017.

[3] A. Abur and A. Gómez-Expósito, *Power system state estimation : theory and implementation*. Marcel Dekker, 2004.

[4] D. Singh, J. P. Pandey, and D. S. Chauhan, "Topology identification, bad data processing, and state estimation using fuzzy pattern matching," *IEEE Transactions on Power Systems*, vol. 20, no. 3, pp. 1570–1579, aug 2005.

[5] B. Hayes, A. Escalera, and M. Prodanovic, "Event-triggered topology identification for state estimation in active distribution networks," in *IEEE PES Innovative Smart Grid Technologies Conference Europe*. IEEE Computer Society, jul 2016.

[6] G. N. Korres and N. M. Manousakis, "A state estimation algorithm for monitoring topology changes in distribution systems," in *IEEE Power and Energy Society General Meeting*, 2012.

[7] E. M. Lourenço, E. P. Coelho, and B. C. Pal, "Topology Error and Bad Data Processing in Generalized State Estimation," *IEEE Transactions on Power Systems*, vol. 30, no. 6, pp. 3190–3200, nov 2015.

[8] O. Lateef, R. G. Harley, and T. G. Habetler, "Bus Admittance Matrix Estimation Using Phasor Measurements," in *2019 IEEE Power and Energy Society Innovative Smart Grid Technologies Conference, ISGT 2019*. Institute of Electrical and Electronics Engineers Inc., feb 2019.

[9] D. M. Vinod Kumar, S. C. Srivastava, S. Shah, and S. Mathur, "Topology processing and static state estimation using artificial neural networks," *IEE Proceedings: Generation, Transmission and Distribution*, vol. 143, no. 1, pp. 99–105, 1996.

[10] J. C. Souza, A. M. Da Silva, and A. P. Da Silva, "Online topology determination and bad data suppression in power system operation using artificial Neural Networks," *IEEE Power Engineering Review*, vol. 17, no. 12, p. 57, 1997.

[11] H. Salehfar and R. Zhao, "A neural network preestimation filter for bad-data detection and identification in power system state estimation," *Electric Power Systems Research*, vol. 34, no. 2, pp. 127–134, aug 1995.

[12] T. Zhang, Y. Wang, J. Ning, and M. Zhai, "Detection and Identification of Bad Data Based on Neural Network and K-means Clustering," in *2019 IEEE PES Innovative Smart Grid Technologies Asia, ISGT 2019*. Institute of Electrical and Electronics Engineers Inc., may 2019, pp. 3634–3639.

[13] M. S. Uddin, A. Kuh, Y. Weng, and M. Ilić, "Online bad data detection using kernel density estimation," in *IEEE Power and Energy Society General Meeting*, vol. 2015-September. IEEE Computer Society, sep 2015.

[14] S. Haykin, *Neural Networks: A Comprehensive Foundation*. Prentice Hall PTR, 1998.

[15] M. Nielsen, *Neural Networks and Deep Learning*. [Online]. Available: http://neuralnetworksanddeeplearning.com

[16] G. E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, and R. R. Salakhutdinov, "Improving neural networks by preventing co-adaptation of feature detectors," jul 2012. [Online]. Available: http://arxiv.org/abs/1207.0580

[17] D. Simon, *Optimal state estimation : Kalman, H∞ and nonlinear approaches*. Wiley-Interscience, 2006.

[18] S. J. Julier and J. K. Uhlmann, "Unscented filtering and nonlinear estimation," in *Proceedings of the IEEE*, vol. 92, no. 3, mar 2004, pp. 401–422.

[19] E. A. Wan and R. Van Der Merwe, "The unscented Kalman filter for nonlinear estimation," in *IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium, AS-SPCC 2000*. Institute of Electrical and Electronics Engineers Inc., 2000, pp. 153–158.

[20] D. Gotti, P. Ledesma, and H. Amaris, "Comparative Analysis between State Estimation Algorithms under Static and Dynamic Scenarios," in *Proceedings - 2020 IEEE International Conference on Environment and Electrical Engineering and 2020 IEEE Industrial and Commercial Power Systems Europe, EEEIC / I and CPS Europe 2020*. Institute of Electrical and Electronics Engineers Inc., jun 2020.

[21] "Measurement set and output binary classification used for topology identification," 2020. [Online]. Available: http://dx.doi.org/10.13140/RG.2.2.13660.82569

[22] M. Pau, F. Ponci, A. Monti, S. Sulis, C. Muscas, and P. A. Pegoraro, "An Efficient and Accurate Solution for Distribution System State Estimation with Multiarea Architecture," *IEEE Transactions on Instrumentation and Measurement*, vol. 66, no. 5, pp. 910–919, may 2017.

[23] F. Iandola, "Exploring the design space of deep convolutional neural networks at large scale," Ph.D. dissertation, University of California, Berkeley, 2016.

[24] N. Srivastava, G. Hinton, A. Krizhevsky, and R. Salakhutdinov, "Dropout: A Simple Way to Prevent Neural Networks from Overfitting," Tech. Rep. 56, 2014. [Online]. Available: http://jmlr.org/papers/v15/srivastava14a.html

[25] E. Manitsas, R. Singh, B. C. Pal, and G. Strbac, "Distribution system state estimation using an artificial neural network approach for pseudo measurement modeling," *IEEE Transactions on Power Systems*, vol. 27, no. 4, pp. 1888–1896, 2012.

© 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Power Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TPWRS.2021.3076671

Davide Gotti (S'21) received the B.Sc. and M.Sc. degrees in electrical engineering from Politecnico di Milano, Italy, in 2011 and 2014, respectively, and the M.Sc. degree (best academic record) in renewable energy in electrical systems from Universidad Carlos III de Madrid, Spain, in 2018. Since November 2019, he has been working toward the Ph.D. degree at the Universidad Carlos III de Madrid, Madrid, Spain. His research interests include state estimation, power system modeling and machine learning algorithms.

Hortensia Amaris (M'00-SM'19) received both the Electrical Engineer degree and the Ph.D. degree from the Technical University of Madrid (UPM), Madrid, Spain, in 1990 and 1995, respectively. She is currently a Full Professor of electrical engineering with University Carlos III of Madrid, Spain, where she has been working mainly on operation and control of power networks, power quality, power-electronic converters, renewable energy sources and smart grids.

Pablo Ledesma Larrea received his Ph.D. in 2001 at University Carlos III of Madrid, where he is currently an Associate Professor. He has worked with the Spanish TSO on several projects on large-scale integration of renewable energy. He has been an Academic Visitor at Chalmers University, Sweden and Strathclyde University, UK. His areas of research are transient stability and dynamic modeling of power systems.

© 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.