Detection and Identification of Generator Disconnection Using Multi-layer Perceptron Neural Network Considering Low Inertia Scenario

Abstract—This research paper presents a method that uses measurements of voltages angles, as provided by phasor measurement units (PMU), to accurately detect the sudden disconnection of a generation unit from a power grid. Results in this research paper have demonstrated, in a practical fashion, that a multi-layer perceptron (MLP) neural network (NN) can be appropriately trained to detect and identify the sudden disconnection of a generation unit in a multi-synchronous generation unit power system. Synthetic time-series bus voltage angles considering low inertia scenarios in the IEEE 39 bus system were used to train the MLP NN. The training process is speeded up by using four GPUs hardware. The simulations results have confirmed the successful detection and identification of the generator outage.

Keywords—Artificial neural network, deep learning, machine learning, outage detection and identification, power system dynamics.

I. INTRODUCTION

In the last years, there has been substantial renewable energy sources (RES) penetration in the modern power system; it is primarily due to the massive integration of photovoltaic (PV) and wind power generation units [1], [2]. This phenomenon brings new challenges, essentially because now most of the power system generator units are not synchronous machines (SM) but power electronics converters-based generation [3]. The main difference between them is the lack of inertia of the last ones [4], [5]. For this reason, energy storage systems are also being integrated to achieve the capability to provide inertia, among other ancillary services [6]. Rotational inertia is essential for the power system since it allows the SM an instantaneous increase of the active power injection. Furthermore, when a frequency change occurs, the rotational inertia helps to damp the frequency variation [7]. Appropriate levels of rotational inertia are required to ensure the reliable and secure operation of the power system [8]. Several blackouts related to frequency stability have demonstrated the importance of appropriate frequency control and inertia levels, e.g., Italy 2003 [9] and Australia, 2016 [10] and Uruguay-Argentina 16th June 2019 [11]. It is clear that modern power systems are facing changes and updates to accommodate the integration of RES and new PEC-based generation technologies, electrical markets, regulations, measuring and communication technologies, among others [12]. This transition brings new challenges to network operation, stability, and security [13], [14].

The system operators (SOs) worldwide rely on the system wide-area monitoring, control, and protection (WAMPAC) to operate the power system securely and economically. As the rotational inertia is declining in modern power systems, and the transients are becoming faster, and extremes [15], [16], the power system monitoring of the power system status is becoming a crucial essential task and a topic in constant research [17]. Keeping the balance between generation and demand in a power system is crucial. Consequently, the online detection of the sudden disconnection of one or more generation units becomes critical in defining emergency operation plans [18]. Detecting the disconnection of a generator is very important for the system operation; as a consequence, the traditional method is based on the status of the circuit breaker of the generation unit. However, monitoring the circuit status using traditional SCADA results in prohibitive time scale, another option is using signal tripping coming from the protective devices and the time response of this signal depends on the technology used at the circuit breaker and station protection control, modern system can take advantages of fast communication using IEC 61850 but that is not the case in all part of a system or all systems.

P. Palensky
Department of Electrical Sustainable Energy
Delft University of Technology
Delft, The Netherlands
P.Palensky@tudelft.nl

J.L.RuedaTorres@tudelft.nl

Paula Páramo Balsa
Department of Electrical Engineering
Universidad de Sevilla
Seville, Spain
pparamo@us.es

Martha Nohemi Acosta Montalvo
Department of Electrical Engineering, Information Technology and Cybernetics
University of South-Eastern Norway
Porsgrunn, Norway
Martha.Acosta@usn.no

Francisco Gonzalez-Longatt
Department of Electrical Engineering, Information Technology and Cybernetics
University of South-Eastern Norway
Porsgrunn, Norway
fglongatt@fglongatt.org

Jose Luis Rueda Torres
Department of Electrical Sustainable Energy
Delft University of Technology
Delft, The Netherlands
J.L.RuedaTorres@tudelft.nl

Monterrey, México
manuel.andrade@uanl.edu.mx

Manuel A. Andrade
School of Mechanical and Electrical Engineering
Universidad Autónoma de Nuevo León
Monterrey, México
manuel.andrade@uanl.edu.mx

Alejandro Verduzco
School of Mechanical and Electrical Engineering
Universidad Autónoma de Nuevo León
Monterrey, México
alejvcm@gmail.com

Peter Palensky
Department of Electrical Sustainable Energy
Delft University of Technology
Delft, The Netherlands
P.Palensky@tudelft.nl

Monterrey, México
Martha.Nohemi@uanl.edu.mx

Monterrey, México
Martha.Acosta@usn.no

Monterrey, México
Francisco.Gonzalez-Longatt@tudelft.nl

Monterrey, México
Jose.Luis.RuedaTorres@tudelft.nl

Monterrey, México
Alejandro.Verduzco@tudelft.nl
Recently, numerous machine learning (ML) techniques have been widely used in modelling and monitoring complex applications, including security [19], frequency control, etc [20]. It includes the use of artificial neural networks (ANN), decision tree (DT), principal component analysis (PCA), support vector machines (SVM), convolutional neural networks (CNN), among others.

Now, considering the importance of precise and appropriate detection of generation disconnection in low inertia power systems and the recent development in the context of machine learning, this research paper is trying to answer the research question: Is it possible to detect the sudden disconnection of a generator in a power system using a multi-layer perceptron (MLP) neural network? The following hypothesis is formulated to answer the research question: (H0) If a multi-layer perceptron neural network is appropriately trained, the sudden disconnection of a generation unit in a multi-machine power system can be detected and identified?. This paper is part of the “Flexible solution (software/hardware) to create artificial intelligence in power systems: FlexAlps” inside the Digital Energy Systems Laboratory (DigEnSys-Lab) [21]. This paper only concerns about detecting and identifying the disconnection of generators, however, the methodology proposed in this paper can be extended to the over frequency events and detect sudden loss disconnection of loads (an event with less probability), in such case the detection mechanism will be able to detect the total power imbalance.

II. BACKGROUND

A. Artificial neural network (ANN)

An ANN learns complex relations between input and output to make accurate predictions. This type of network consists of several simple nodes, also called neurons, connected to form either a single layer or multiple layers. There are different types of ANN structures like multi-layer perceptron (MLP), radial basis function (RBF) network, self-organising maps (SOM) [22]. Nowadays, MLP (also known as the universal approximator) is the most extended and most frequently used type of neural network [23]. An MLP feedforward NN is chosen in this research paper since it can form arbitrarily complex decision boundaries and represent any Boolean function. The MLP has three (or more) layers created from nonlinearly activating nodes: one input layer (IL), one (or more) hidden layer (HL) and one output layer (OL), as shown in Fig. 1.

The learning process in the MLP occurs in the perceptron; a training mechanism is used for changing connection weights after each piece of data is processed. In this paper, a supervised learning process is selected for the MLP. During the training process, the MLP will receive a group of vectors (organised in a matrix) of inputs, also called features (X) and a group of vectors of desired outputs, also called targets (D). The information only circulates from the neurons of the input layer towards the output layer. The action of a neuron depends on its activation functions, which is described as:

$$y_i = f\left(\sum_{j=1}^{n} w_{ij}x_j + b_i\right)$$

where $$y_i$$ is the output of the $$i$$-th neuron, $$f(\cdot)$$ is the activation function, $$w_{ij}$$ is the weight from the $$j$$-th input to the $$i$$-th neuron, $$x_j$$ is the $$j$$-th input of the $$i$$-th neuron, and $$b_i$$ is the bias of the $$i$$-th neuron. The activation function is a non-linear function that describes the reaction of $$i$$-th neuron with inputs $$x_i(t), j = 1, \ldots, n$$. The input signal is connected to a group of hidden layers which use a tanh function. These are likewise connected to an output layer that uses softmax function. The supervised training of the ANN modifies the weights of the connections with the objective of minimising the error of the ANN output concerning the desired output. Usually, MLPs are trained in a supervised manner with a popular algorithm known as backpropagation [24]. When the ANN is trained, a new input can be introduced to estimate the output.

Nevertheless, during the training process, an undesired phenomenon called overfitting may happen. This phenomenon causes the ANN to memorise the training data. It would result in the loss of generalisation of the ANN and, consequently, loss of the capacity to work as desired when new inputs are presented. The input data is divided into Training set (70%), Validation set (15%), and Test set (15%) to avoid overfitting. The Training set is used to train the ANN and vary the weights concerning the output error. The Validation set adjusts some training related parameters. The Test set does not change the learning of the ANN; it is used to test the performance of the ANN when given data not seen before.

III. MATERIAL AND METHODS

In this paper, the authors confirm the hypothesis (H0) that an MLP neural network can be appropriately trained. The sudden disconnection of a generation unit in a multi-machine power system can be detected and identified.

A. Proposed MLP to detect and identify generator outage

This section proposes using an MLP neural network to detect and identify the disconnection of a generator from the multi-machine power system. The representation of the MLP based detection and identification method of generator outage is presented in Fig. 2.

Corollary 1: A multi-machine power system is fully observable if the voltage phasors at all system buses can be uniquely estimated using the available measurements.

The approach used in this paper assumes that by using a wide-area monitoring platform (WAM) based on multiple phasor measurement units (PMUs) the power system is fully observable. Therefore, the proposed detection and identification method used an MLP neural network, where the input signals are the time series of the bus voltage angles ($$\delta$$) of the power system, where:

$$\delta = [\delta_1 \delta_2 \ldots \delta_n]^T$$

Fig. 1. Diagram showing a multi-layer perceptron (MLP), it depicts four inputs in the IL, one hidden layer (fully connected network) and an output at the OL.
where \( \delta_i \) is the time series representing the voltage angle of the \( i \)-th bus \((n = \text{number of buses})\), \( \{\delta_i \in T\} \) or simply \( \{\delta_i\} \).

\[ X_r = \begin{bmatrix} \delta_1^1 & \delta_1^2 & \cdots & \delta_1^n \\ \delta_2^1 & \delta_2^2 & \cdots & \delta_2^n \\ \vdots & \vdots & \ddots & \vdots \\ \delta_n^1 & \delta_n^2 & \cdots & \delta_n^n \end{bmatrix} \] \hspace{1cm} (3)

\[ D_r = \begin{bmatrix} D_1^1 & D_1^2 & \cdots & D_1^n \\ D_2^1 & D_2^2 & \cdots & D_2^n \\ \vdots & \vdots & \ddots & \vdots \\ D_n^1 & D_n^2 & \cdots & D_n^n \end{bmatrix} \] \hspace{1cm} (4)

where \( r \) is the number of the simulation the information is extracted from, \( n \) is the number of buses of the power system, \( s \) is the number of samples per simulation selected for the training, and \( n_g \) is the number of generators in the system. As shown before, both matrices share the same number of samples \( s \) (columns) due to every feature vector corresponding to the desired output vector (targets vector). In this paper, the authors selected 989 samples of features and targets of every simulation. It represents 9.86 s of each simulation, where the behaviour before, during and after disturbance can be observed. The selected samples amount was the result of a search for desired performance and fast training. Consequently, 1,210 features and targets matrices were created. All feature matrices were concatenated as well as all targets matrices. The result is a features matrix \( X \) with dimensions of \( 39 \times 1,196,690 \) and a targets matrix \( D \) with dimensions of \( 10 \times 1,196,690 \), both matrices having information of all the RMS simulations.

C. Creation of features and targets matrices

The data created in the previous section produced time series of several electro-mechanical variables; however, the bus voltage angles were selected to form the features matrix \( (X_r = \delta) \). The rationale behind this decision is that the electro-mechanical dynamics that the generator disconnection causes in a multi-machine power system are reflected in the bus voltage angles faster than in frequency inductive networks, which is the case of high voltage grids. Also, from the practical point of view, this research project investigates the real implementation of the proposed approach, and the availability of data streams from PMU offers the information related to the voltage angles, but also the power system state estimator might help with data in the case of the availability of the data stream of one PMU.

The per-unit current of each generator, coming from the synthetic data, is used to identify if the machines are connected or not immediately after the disturbance. This information is used to label the data set during the training process in the targets matrix \( (D_r) \). Each generator is assigned a binary status variable to indicate the connection status, \( D_{ri} = 1 \) if the \( i \)-th generator is disconnected to the power system and \( D_{ri} = 0 \) if it is connected. For instance, if at time \( t \), the generator G7 is disconnected from the test system and the rest of the machines are connected, the targets vector in that instant would be \( D_r = [0, 0, 0, 0, 0, 1, 0, 0, 0]^T \).

Organising all the features and targets vectors in matrices, \( X_r \) and \( D_r \), results in:

\[ X_r = \begin{bmatrix} \delta_1^1 & \delta_1^2 & \cdots & \delta_1^n \\ \delta_2^1 & \delta_2^2 & \cdots & \delta_2^n \\ \vdots & \vdots & \ddots & \vdots \\ \delta_n^1 & \delta_n^2 & \cdots & \delta_n^n \end{bmatrix} \] \hspace{1cm} (3)

\[ D_r = \begin{bmatrix} D_1^1 & D_1^2 & \cdots & D_1^n \\ D_2^1 & D_2^2 & \cdots & D_2^n \\ \vdots & \vdots & \ddots & \vdots \\ D_n^1 & D_n^2 & \cdots & D_n^n \end{bmatrix} \] \hspace{1cm} (4)

where \( r \) is the number of the simulation the information is extracted from, \( n \) is the number of buses of the power system, \( s \) is the number of samples per simulation selected for the training, and \( n_g \) is the number of generators in the system. As shown before, both matrices share the same number of samples \( s \) (columns) due to every feature vector corresponding to the desired output vector (targets vector). In this paper, the authors selected 989 samples of features and targets of every simulation. It represents 9.86 s of each simulation, where the behaviour before, during and after disturbance can be observed. The selected samples amount was the result of a search for desired performance and fast training. Consequently, 1,210 features and targets matrices were created. All feature matrices were concatenated as well as all targets matrices. The result is a features matrix \( X \) with dimensions of \( 39 \times 1,196,690 \) and a targets matrix \( D \) with dimensions of \( 10 \times 1,196,690 \), both matrices having information of all the RMS simulations.

D. Creation of the ANN

An ANN is defined by the number of layers, the number of neurons in each layer, the connections links, and the weights. The data enters the ANN via the IL, go through the HL, and the output comes out of the OL. In this research paper,
an MLP feedforward neural network is employed. None of the neuron’s outputs in feedforward networks serves as an input to neurons of the same layer or previous layers. The input signal will only propagate forward, from input neurons to output neurons. Since using MLPs, two hidden layers are necessary for full generality; two hidden layers were chosen for this work [20].

The number of hidden neurons was determined by using the \textit{geometric pyramid rule} suggested in [26], which has shown good behaviour. The \textit{geometric pyramid rule} is a rough approximation of the ideal number of hidden neurons and for two hidden layers shows that:

\begin{align*}
NHN_1 &= Mr^2 \quad (5) \\
NHN_2 &= Mr \quad (6) \\
r &= \frac{N}{M} \quad (7)
\end{align*}

where \(N\) is the number of inputs, \(M\) is the number of outputs, \(NHN_1\) is the number of neurons in hidden layer 1, and \(NHN_2\) is the number of neurons in hidden layer 2. It results in \(NHN_1 = 24\) and \(NHN_2 = 15\), given 39 inputs and ten outputs are being considered.

\section*{E. Training process of MLP}

MLPs usually are trained using the error \textit{backpropagation} (BP) algorithm. BP is a gradient descent-based method to minimise the error output computed by the ANN. BP can be viewed as two stages process: (i) \textit{forward stage} and (ii) \textit{backpropagation}. In the first stage, an input vector is applied to the ANN input nodes, and its effect propagates through the network layer by layer. The first stage produces a set of outputs as the actual response of the network. During the forward stage, the weights of the connections are all fixed. In the second stage, the weights are all adjusted following an error-correction rule: the actual response of the ANN is subtracted from the desired response (target) to produce an error signal. This error signal is propagated backwards through the network. This process is repeated layer by layer until completing the artificial neural network. The weights of each neuron are readjusted to make the actual response of the ANN move closer to the desired response, i.e., minimising the error.

The training algorithm chosen is BP based on the optimisation technique \textit{Broyden-Fletcher-Goldfarb-Shanno} (BFGS) \textit{quasi-Newton}. This method has successfully solved nonlinear equations systems [27]. The authors of the BFGS quasi-Newton method suggests the method “substantially outperforms known conjugate gradient methods on a wide class of problems” [28]. A detailed discussion of the BFGS quasi-Newton method can be found in [29].

Cross-entropy (CE) loss, also known as log loss, was selected as the error function to minimise. CE loss measures the performance of a classification model whose output is a probability between 0 and 1. CE loss increases as the predicted probability moves away from the actual classification. Minimising cross-entropy leads to good classifiers. The CE loss for every pair of output-target elements is calculated as:

\[ CE = -d \log(x) \quad (8) \]

where \(d\) is the desired output and \(x\) is the actual output of the ANN.

\section*{F. Training the MLP feedforward neural network}

In this research paper, the synthetic data is divided randomly into three parts: training set (70%), validation set (15%) and test set (15%), avoiding overfitting of the ANN. The MLP consists of two hidden layers, 24 and 15 hidden neurons in each hidden layer (due to results of (5), (6) and (7)). The training process starts by giving the \(X\) and \(D\) matrices. Table I shows the parameters used for the training process.

\begin{table}[h!]
\centering
\caption{LIMITS OF TRAINING CRITERIA}
\begin{tabular}{|l|c|}
\hline
\textbf{Name} & \textbf{Value} \\
\hline
epochs & 1000 \\
\hline
\(CE\) & 0 \\
\hline
\(\text{min grad}\) & \(1 \times 10^{-6}\) \\
\hline
val checks & 6 \\
\hline
\end{tabular}
\end{table}

Epoch is a term used in machine learning to define the number of times all the data in the training set have passed through the learning algorithm. At each epoch, all the training data goes through the model so that the model learns from it.

\section*{IV. RESULTS}

The proposed neural network was implemented in MATLAB 2020a using the \textit{Deep Learning Toolbox}. The data pre-processing, training and testing is performed using Intel Xeon W-3235 Processor, 12-Core, 24-Thread, 3.3/4.4GHz, 64GB RAM, Windows 10 professional and four NVIDIA® GeForce® RTX 2080Ti graphics card. The use of four GPU and parallel processing allowed a substantial speed-up in the training process (approx. 312.87 s).

Fig. 4 shows the evolution of the ANN training and how the error decreases through the epochs until the optimal point is found. The minimum CE loss reached was 2.4\times10^{-4}; this is the mean of the individual CE values (as shown in (8)). The overall accuracy of the trained ANN is 89.8%, considering all three sets. This means that of the total observations presented during training, the ANN classified correctly 89.8% of the observations and misclassified 10.2%. For instance, a misclassification would be the ANN output showing generator G7 was disconnected when actually generator G4 is the one offline or showing generator G3 is online when it is currently offline.

Fig. 5 shows the graphical results of an illustrative example of testing the trained ANN. The ANN received as input 4,000 samples (up to 39.98 seconds of the RMS simulation) of the 39 buses' voltage angles. For this specific example, the simulation case considers the initial rotational inertia of the synchronous generators in the system. Then, the outage of the generator G1 (at \(t = 1.00\) s) is presented.
Fig. 5 depicts the change of the status of generator G1 from $D_{G1} = 0$ (online status) to $D_{G1} = 1$ (offline status), showing the ANN maintains high confidence the generator G1 is disconnected during the rest of the simulation. The remaining generators maintain a very close to $D_{Gi} = 0$ ($i = 2, 3, \ldots, 10$) status, confirming the successful detection and identification of the generator outage.

Fig. 5. Example of output of trained ANN when detecting disconnection of generator 1.

Fig. 6 demonstrates the performance of the trained ANN in a low inertia situation. The ANN received 4,000 samples of the 39 buses voltage angles in the input layer. The input data considering a 10% reduction in the inertia of all machines. From Fig. 6, it can be seen that at the moment of the generator disconnection ($t = 1.00$ s), the MLP detects this event instantaneously and correctly identifies which machine is now offline (in this test, G4) by modifying the status of G4 from zero to one. Meanwhile, the rest of the generators maintain their status at zero value, meaning the ANN identifies those generators are still connected to the power grid.

The tests shown prove that, with proper training, an MLP ANN can successfully identify the disconnection of a generator considering different inertia level situations. Moreover, it can help in the decision process of SOs in a very fast and precise manner, in the case, the right data is available. This is especially important since the total inertia in modern power grids changes throughout the day depending on the power generation. Using synthetic data of the voltage angles of the IEEE 39-bus test system, the ANN presents preliminary results of a more extended project, but it has been able to demonstrate the suitability of the approach by using synthetic data of the voltage angles of the IEEE 39-bus test system.

Fig. 6. Example of output of trained ANN when detecting disconnection of generator 4, in a case of reduced inertia.

V. CONCLUSIONS

The reduced rotational inertia of modern power system is pressing the system operator to make decisions in an increasingly reduced time frame. Traditional mechanism of detecting the sudden disconnection of a generator has different time response, using modern IEC 61850 framework provided the faster and more reliable way and traditional signalling using SCADA system is the slowest and less reliable; however, it is clear that a real power system has not possibility of having a homogeneous IEC61850 framework. As a consequence, this research proposed an alternative for the detection and identification of generator disconnection. The proposed approach uses voltage angle measurements time series coming from phasor measurement units a technology that is growing in penetration (nowadays, many protection relays are already software enabled with the option of PMU). This research paper has demonstrated the hypothesis by appropriately training a multi-layer perceptron neural network; it can successfully detect and identify the sudden disconnection of a generation unit in a multi-machine power system observing only the bus voltage angles. This paper presents preliminary results of a more extended project, but it has been able to demonstrate the suitability of the approach by using synthetic data of the voltage angles of the IEEE 39-bus test system.

ACKNOWLEDGMENT

Martha N. Acosta and Alejandro Verduzco want to acknowledge the financial support given by CONACYT (México) and the support of Universidad Autónoma de Nuevo León. Martha N. Acosta would like to recognise the support of the University of South-Eastern Norway. Prof F. Gonzalez-Longatt would like to express his gratitude to DiGSIlicant for supporting his research and strong collaboration.

REFERENCES


