Electrical Power and Energy Systems 104 (2019) 215–221

Contents lists available at ScienceDirect

Electrical Power and Energy Systems

journal homepage: www.elsevier.com/locate/ijepes

Signal processing techniques for sensing based generator coherency analysis

Gianluca Bruno, Enrico Maria Carlini, Cosimo Pisani*

TERNA, Italian System Operator, Via Palmiano 101, 00138 Rome, Italy

A R T I C L E I N F O

Keywords:
Coherency
Power system
Clustering

A B S T R A C T

In this paper generator coherency analysis of power system is investigated via some signal processing techniques. Sensor data analysis here designed is based on the fusion of advanced signal processing techniques for sensing-based coherency identification, including *k*-means and fuzzy *k*-means clustering, agglomerative hierarchical cluster tree, and Independent Component Analysis (ICA). Detailed results are presented and discussed in order to prove the effectiveness of the techniques and carry out a comparative assessment.

1. Introduction

The large scale deployment of advanced sensor networks for acquiring and processing synchronized and spatially-distributed measurements is considered one of the most important prerequisite for improving the reliability and the security of large-scale interconnected power systems. This has stimulated the conceptualization of Wide Area Measurement Systems (WAMS), which process phasor data acquired from key buses by Global Positioning System (GPS)-synchronized sensing devices, called Phasor Measurement Units (PMUs) [1]. The large stream of time-synchronized phasor measurements acquired by the PMUs, if properly processed and integrated with other traditional sensors (such as remote terminal units, digital fault recorders, etc.), can be adopted to enhance the power systems observability. This improves the "situational awareness" of Transmission System Operators (TSOs) by enabling advanced proactive functions, such as system integrity protection schemes, adaptive protection, and dynamic on-line security analysis (DSA) [2]. These applications could play a strategic role in modern electrical transmission networks, which are frequently pushed to operate very close to their stability limits.

The identification of the *coherent* group of generators, here referred as Generation Coherence Analysis – GCA, represents strategic and valuable information [3]. In this context, the term *coherent* means that, after the disturbance onset, the generators exhibit similar rotor angle swing curves, which are so close to each other that they can be assumed to oscillate together [4]. One fundamental remark in coherency analysis is that the formation of coherent groups depends on both the nature and location of the disturbance [5].

Several methods have been proposed in literature for GCA, which can be classified in two main categories: *model-based methods* and *measurement based methods*. Model-based methods mainly rely on the availability of a power system dynamic model, which is typically linearized around the current operating point. Although the adoption of these methods has been widely explored in the power system literature, their real-time deployment on large and interconnected power systems is highly challenging due to their huge computation demand. Moreover, they need detailed information on the modeling parameters of each power system component, which is not readily available, or affected by large uncertainties.

To overcome these limitations, the adoption of measurement-based coherency identification methods has been proposed in the literature [6]. These approaches try to extract actionable knowledge from the grid sensors data streaming, such as generator rotor angle and speed, bus voltage magnitude and phase.

The modern literature on measurement-based GCA is vast and [7–11] outlines the main contributions, the open problems, and the research challenges characterizing this emerging research domain.

The analysis of these papers reveals that, although several signal processing based techniques have been proposed for GCA, an experimental assessment of their performances on a real and complex operation scenario is still at its infancy [12].

Armed with such a vision, in this paper generator coherency analysis of power system is investigated via some signal processing techniques. Sensor data analysis here designed is based on the fusion of advanced signal processing techniques for sensing-based coherency identification, including k-means and fuzzy k-means clustering, agglomerative hierarchical cluster tree, and Independent Component Analysis (ICA). Detailed results are presented and discussed in order to prove the effectiveness of the techniques and carry out a comparative assessment.

The remainder of the paper is organized as follows. Section 2 collected the main literature contributions regarding GCA. In Section 3 the

* Corresponding author.
E-mail address: cosimo.pisani@terna.it (C. Pisani).

https://doi.org/10.1016/j.ijepes.2018.06.020
Received 9 February 2018; Received in revised form 7 May 2018; Accepted 5 June 2018
Available online 12 July 2018

theoretical foundations of the adopted sensing-based coherency identification algorithms are analyzed. In Sections 4 the main results are presented and discussed. The conclusive remarks and the future developments are summarized in Section 5.

2. Related works

GCA is one of the most fundamental tool for implementing *dynamic equivalencing* of power systems [13]. Thus, it has attracted large research efforts aimed at defining effective methodologies for identifying coherent areas in interconnected power systems.

In particular, GCA in coherency based DE techniques has been traditionally addressed by deploying linearized power system models. These solution techniques are very straightforward but, as outlined in [6], they could be not suitable for GCA of large-scale power systems in the presence of critical disturbances, due to their inability to describe complex non-linear system dynamics. This has limited their deployment in power system control centers, and stimulated research for alternative solutions based on more advanced modeling techniques. In particular, the solution method proposed in [14] combines the Taylor-series expansion of the generator rotor angles at three different transient phases, with a measure of the electrical coupling among the generators, obtained by defining a distance measure based on the system admittance matrix. According to this approach, the generators' coherency is identified by analyzing the epsilon decompositions of the power flow equations Jacobian matrix. This makes the proposed approach mathematically straightforward, and its algorithmic complexity linear. On the other hand it requires the precise knowledge of detailed parameters of the power systems under analysis, which are very difficult to estimate, and are affected by large uncertainties [15].

To overcome this limitation, more advanced solution methodologies, based on signal processing techniques, have been proposed in the literature. In particular, in [16] a line vulnerability index, which is obtained by processing the post-fault line transient potential energy and the bus voltage magnitudes, is adopted to classify coherent areas in large-scale power transmission systems. The same problem has been solved in other papers by means of spectral analysis techniques, which include: Fast Fourier Transform (FFT) of the generator rotor angles expanded via Taylor series [17], Fourier analysis of generator speed measurements [18], Hilbert Huang Transform of the phase differences among inter-area oscillations and swing curves [19], normalized spectral clustering algorithm and ICA of the generator speed and bus angle signals [20,21], wavelet phase difference analysis of low frequency electromechanical oscillations [22] and multiflock-based analysis of the generator frequencies and phases [23].

Recently, attention has been given to pattern recognition techniques based on neural networks [24] and clustering techniques some of which are presented in Section 3. Although the performance of these methods have been successfully validated on some experimental test-beds, their deployment on realistic transmission systems is still in its infancy, and needs to be researched [25]. In this context, several open problems need to be fixed including the improvement of the computing efficiency, the enhancement of the algorithm scalability, and the complexities in managing large data-sets. These issues are particularly relevant in the context of the ENTSO-E Continental synchronous area, where the complexities deriving by the interconnection of the national power systems, the need for accurate monitoring of the strategic energy corridors.

In trying to address these issues, generator coherency analysis of power system is investigated via some signal processing techniques. The main idea is to extract actionable intelligence from measured data-sets by properly combining multiple signal processing techniques, such as k-means clustering techniques, agglomerative hierarchical cluster tree and ICA. The outcomes of this measurement-based coherency identification paradigm is expected to match with a model-based coherency identification algorithm (e.g. directional cosine [4]), but with a lower computational complexity. This important feature makes the proposed solution particularly suitable for an on-line application.

3. Measurement-based coherency identification algorithms

The measurement based coherency identification algorithms here analyzed are presented in the following with related mathematical framework.

3.1. K-means

K-means is the most simple and popular unsupervised learning algorithm able to deal with clustering problems. The aim of the method is to classify an ensemble of measurements, as the case of the generator rotor measurements (i.e. rotor angle or speed) in a certain number of clusters, k, fixed a priori. Each cluster is defined by a centroid properly placed in the clustering space: different centroid locations imply different results. One approach to initialize the algorithm is to position each centroid as much as possible far away from each other. Then, each point belonging to a given measurements ensemble has to be associated to the nearest centroid. At this point the new k centroids can be updated as barycenters of the clusters resulting from the previous step and a new binding has to be performed among the points of the measurements ensemble and the nearest new centroids. The procedure is iterated until no more change of centroids position is obtained. In mathematical terms, given an ensemble of observations $\mathbf{x} = [x_1, x_2, \ldots, x_n]^T$, i.e. generator rotor measurements, where each of them is a d-dimensional real vector, k-means clustering collects the n observations into k ($\leq$n) sets, $\mathbf{S} = \{S_1, S_2, \ldots, S_k\}$ by minimizing the following objective function:

$$F = \sum_{i=1}^{k} \sum_{x \in S_i} \|\mathbf{x} - \mu_j\|^2 \tag{1}$$

hence:

$$\underset{\mathbf{S}}{\operatorname{argmin}} = \sum_{i=1}^{k} \sum_{x \in S_i} \|\mathbf{x} - \mu_j\|^2 \tag{2}$$

whereas $\|\mathbf{x} - \mu_j\|$ is a properly chosen measure of distance (e.g. Euclidean or L1 distance etc.), among n data measurements $\mathbf{x}$ and the cluster center μ_j. Each individual x_p in $\mathbf{x}$ must be assigned to only one cluster. Standard algorithm for solution of (2) is proposed by Lloyd in [26] and consists in two steps. Starting from an initial guess about k means, the algorithm assigns each observation to the cluster whose mean, m_k, yields the least within-cluster sum of squares:

$$S_i^{(t)} = \{x_p\colon \|x_p - m_i^{(t)}\|^2 \leqslant \|x_p - m_j^{(t)}\|^2 \ \ \forall j, \ 1 \leqslant j \leqslant k\} \tag{3}$$

Each x_p is hence assigned to only one element of $S^{(t)}$ while the new means, to be the centroids of the observations in the new clusters, are updated as follows:

$$m_i^{(t+1)} = \frac{1}{|S_i^{(t)}|} \sum_{x_j \in S_i^{(t)}} x_j \tag{4}$$

The main drawbacks of the method are that (*i*) there is no guarantee of convergence to a global minimum of the (1) and (*ii*) the results are significantly sensitive to the initial randomly selected cluster centers. By running the algorithm multiple times and by recourse to the physical knowledge about the investigated power system both effects can be mitigated.

3.2. Fuzzy k-means

The highlighted mathematical framework suggests how k-means could be a good potential candidate for extension to work with fuzzy theory. In particular, while in k-means algorithm each observation, x_p, cannot be assigned to more than one cluster, in fuzzy k-means this can