
This version is available at https://strathprints.strath.ac.uk/62466/

Strathprints is designed to allow users to access the research output of the University of Strathclyde. Unless otherwise explicitly stated on the manuscript, Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Please check the manuscript for details of any other licences that may have been applied. You may not engage in further distribution of the material for any profitmaking activities or any commercial gain. You may freely distribute both the url (https://strathprints.strath.ac.uk/) and the content of this paper for research or private study, educational, or not-for-profit purposes without prior permission or charge.

Any correspondence concerning this service should be sent to the Strathprints administrator: strathprints@strath.ac.uk

The Strathprints institutional repository (https://strathprints.strath.ac.uk) is a digital archive of University of Strathclyde research outputs. It has been developed to disseminate open access research outputs, expose data about those outputs, and enable the management and persistent access to Strathclyde's intellectual output.
Feasibility Study of Applicability of Recurrence Quantification Analysis for Clustering of Power System Dynamic Responses

Panagiotis N. Papadopoulos, Member, IEEE
Jovica V. Milanović, Fellow, IEEE
School of Electrical and Electronic Engineering
The University of Manchester
Manchester, UK
panagiotis.papadopoulos@manchester.ac.uk
jovica.milanovic@manchester.ac.uk

Pratyasa Bhui
Nilanjan Senroy, Member, IEEE
Department of Electrical Engineering
Indian Institute of Technology Delhi
New Delhi, India
nsenroy@ee.iitd.ac.in

Abstract—A methodology based on Recurrence Quantification Analysis (RQA) for the clustering of generator dynamic behavior is presented. RQA is a nonlinear data analysis method, which is used in this paper to extract features from measured generator rotor angle responses that can be used to cluster generators in groups with similar oscillatory behavior. The possibility of extracting features relevant to damping and frequency of oscillations present in power systems is studied. The k-Means clustering algorithm is further used to cluster the generator responses in groups exhibiting well or poorly damped oscillations, based on the extracted features from RQA. The effectiveness of RQA is investigated using simulated responses from a modified version of the IEEE 68 bus network, including renewable energy resources.

Index Terms—clustering, k-means, recurrence quantification analysis, renewable energy resources.

I. INTRODUCTION

Due to various social, environmental economic and technical reasons, power systems may be driven to operate closer to their stability limit. New techniques and algorithms need to be implemented, that will facilitate close to real time identification of power system dynamic behavior that will in turn enable corrective control measures. Identifying generators with similar dynamic behavior is a key issue that can guide the implementation of special protection schemes for the prevention of impeding instabilities [1].

There are several clustering algorithms available in the literature that have been applied to identify coherent groups of generators, such as fuzzy C-means, principal and independent component analysis and support vector clustering [2][5]. Different features of time domain responses can be used to define similarity between generator responses before applying the clustering algorithms. Hierarchical clustering has been also applied in [1] to define groups of generators that exhibit instability together. The obtained clusters from this approach are directly related to transient stability rather than to slow coherency. However, this method focuses on unstable cases and all stable generator responses are clustered in one group. Therefore, a further clustering approach for clustering the stable generators which are exhibiting oscillatory behavior is needed, that could be subsequently used to determine the appropriate stabilizing control actions for groups of generators exhibiting similar poorly damped oscillatory behavior.

In this paper, the application of Recurrence Quantification Analysis (RQA) is investigated as a tool to extract relevant features from measured responses that will help in clustering generators exhibiting oscillatory behavior. Most of the above mentioned methods have been applied in power systems with conventional synchronous generators. RQA has the advantage that it can be applied to noisy, nonlinear, non-stationary data, such as those usually obtained from power system dynamic responses after large disturbances. This becomes more important with the increasing penetration of Renewable Energy Resources (RES), which are mainly interfaced using power electronics and are expected to change the system dynamic behavior.

RQA has been used in power system dynamic studies in the past in order to define, among the other, a transient stability index based on voltage measurements [6]. In this study, RQA is applied on generator rotor angles to test the efficiency of the method in extracting useful features that are afterwards used to cluster the generator responses. The extracted features reflect changes in the damping and frequency of oscillation of the generators. This way, groups of generators with poorly or even negatively damped oscillatory behavior can be identified.

II. METHODOLOGY

RQA is a method of nonlinear data analysis which quantifies the number and duration of recurrences of a dynamical system presented by its state space trajectory [7].
Several measures of recurrence have been presented in [7], based on recurrent plots [8]. In this paper, the basic properties of Recurrence Rate (RR) and Determinism (DET) are used to investigate whether these measures are suitable for the extraction of relevant features to cluster the dynamic behavior of power systems.

RR is defined by (1) and DET by (2), where $R$ is the recurrence matrix, $N$ is the number of samples and $P(l)$ is the number of times a diagonal of length $l$ occurs. The recurrence matrix $R$ is a matrix which stores the data points $R_{ij}$ of the measured response that are recurrent, i.e., the distance between the points is less than a specified threshold value. $r$ is the density of recurrent points and DET quantifies the percentage of recurrent points that form a diagonal of minimum length $l_{min}$. In this study a default value of $l_{min}$ is set to 2. More information on RQA can be found in [6]-[8].

\[
RR = \frac{1}{N^2} \sum_{i,j=1}^{N} R_{ij} 
\]

\[
DET = \frac{\sum_{l=l_{min}}^{N} \sum_{i,j=1}^{N} R_{ij}}{\sum_{i,j=1}^{N} R_{ij}} \cdot P(l)
\]

A. Use of RQA on Typical Power System Responses

The aim of clustering generator responses is mainly to distinguish between three attributes: amplitude, damping and frequency of oscillations. RQA is therefore tested to investigate the effectiveness of the method in extracting the relevant features that can reflect changes in these three attributes. Most commonly, oscillations observed in power system responses have the form of damped sinusoids, defined by (3). Four signals, $y_1$-$y_4$, are used as an example to illustrate the ability of RQA to distinguish between signals with different amplitude, frequency and damping, as shown in Table I. RR and determinism are calculated for these four signals and the results are show in Fig. 1. RR varies mainly with the change in damping and very slightly with the change in frequency but it does not change when the amplitude changes. On the other hand DET mainly varies with the change in frequency. Therefore, the two RQA measures could be used as features that will distinguish between damping and frequency. Lower values of RR and DET are expected as the damping of oscillations becomes smaller and as the frequency increases, respectively.

\[
y = A \cdot e^{-\sigma t} \cos(2\pi f t)
\]

### Table I. Test Signals Parameters

<table>
<thead>
<tr>
<th>Test signal</th>
<th>$A$</th>
<th>$\sigma$(1/s)</th>
<th>$f$(Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$y_1$</td>
<td>1</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>$y_2$</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>$y_3$</td>
<td>1</td>
<td>0.2</td>
<td>0.5</td>
</tr>
<tr>
<td>$y_4$</td>
<td>1</td>
<td>0.5</td>
<td>0.9</td>
</tr>
</tbody>
</table>

B. Clustering Generator Responses

The aim is to cluster generators into groups based on how oscillatory their behavior is after a disturbance. As a first step, the “first-swing” unstable cases can be identified using the method proposed in [11]. For the cases that do not exhibit “first-swing” instability, RR and DET are calculated for the rotor angle responses of each one of the generators within the system. Clustering is then applied using RR and DET as features in order to group the generators according to how well or poorly damped oscillatory behavior they exhibit.

K-Means is used a simple partitioning method to cluster generators of a system into $k$ groups based on the features extracted from RQA (i.e. RR and DET). K-Means is an iterative centroid based partitioning technique that uses the centroid $c_i$ of a cluster to represent that cluster. The centroid $c_i$ is defined as the mean value of the points within the cluster. Initially, $k$ points are randomly selected as centroids and the rest of the points are assigned to the respective clusters for which they have the smallest Euclidean distance. The iterative partitioning method afterwards minimizes the sum, over all clusters, of the within-cluster sums of point to cluster-centroid distances [9]. The number of clusters can be defined by observing how RR and DET vary for a large number of simulated contingencies for a given system. A simple approach is to use two clusters (i.e. $k=2$) to distinguish between the generators exhibiting poorly damped and well damped oscillatory behavior. In case the number of groups is not predefined, different clustering techniques such as hierarchical clustering can be also used with RR and DET as features. However, a threshold value that will distinguish between well and poorly damped oscillations needs to be defined in this case. K-Means is chosen in this paper since the focus is to mainly distinguish between well and poorly damped oscillatory behavior and therefore the number of generator groups is predetermined.

![Figure 1. a) RR and determinism for test signals b) Test signals.](image)

### III. Application Results

A. Test System

The test network used, is a modified version of the IEEE 68 bus/16 machine reduced order equivalent model of the New England Test System and the New York Power System (NETS – NYPS). The system consists of five areas, NETS, NYPS and three external areas represented by G14, G15 and G16 respectively. The conventional part of the test network is adopted from [10] and ten RES units are connected at the buses shown in Fig. 2. For each RES unit, two types of RESs are connected at each bus: Type 3 Doubly Fed Induction
Generators (DFIGs), representing wind generators and Type 4 Full Converter Connected (FCC) units, representing both wind generators and Photo-Voltaic (PV) units. The amount of connected RES (installed capacity) is considered to be equal to 20% of the installed conventional generation of each area. The instantaneous penetration of RES is varying according to the modeled uncertainties as described later. The models used are available in DigSILENT/PowerFactory[11] software and are suitable generic RMS models for large scale stability studies following a similar approach to [12]. More information on the modified test network and RES modelling can be found in [13]. All RES units are considered to provide Fault Ride Through (FRT) capability.

Figure 2. Modified IEEE 68 bus system.

B. Monte Carlo Simulations

A large number of contingencies are simulated using Monte Carlo (MC) to generate representative dynamic responses for the system under study. 6000 contingencies are simulated in this study considering self-clearing three phase faults. Since the purpose of this paper is to investigate the dynamic behavior of the system for cases that do not exhibit non-oscillatory instability, these cases are identified and removed from the dataset. The effect of the intermittent behavior of wind generators and Photo-Voltaic (PV) units is modeled using appropriate probability distributions as explained below. The uncertainties related to the conventional part of the system such as system loading, fault location and duration are also included in this study.

For system loading and PV generation, typical daily curves are initially used, obtained from national grid data[14] and the literature[15] respectively. The system load is considered to vary from 0.6 to 1 pu. The hour of the day is sampled randomly and for each hour, an appropriate PDF is used to model the uncertainty within the hour. For system loading a normal distribution with mean value 1 pu and standard deviation 3.33% is used[11] while for PV generation, a beta distribution with a and b parameters 13.7 and 1.3[16] For wind generation, a Weibull distribution is used with $\varphi=11.1$ and $k=2.2$[17] assuming constant wind speed for every hour within the day. For each load within the system and each RES unit, the PDFs are sampled separately, to consider independent behavior of each unit. For example, assuming the total load is 0.8 pu for a given hour of the day, one load might have a value of 0.75pu and another 0.85pu. The fault location is also sampled randomly following a uniform distribution along each line within the system. Moreover, the uncertainty of the fault duration is modeled using a normal distribution with mean value 13 cycles and standard deviation 6.67%[11].

After considering the uncertainties, Optimal Power Flow (OPF) is solved to determine the conventional generators dispatch $P_{ig}$, where $g=1…16$ is the generator number for the specific test network and $i=1…N$, is the number of simulated contingencies. The nominal capacity of each generator $S_{ng}$ is then adjusted by adding 15% spare capacity according to $S_{ng} = P_{ig} / 0.85$ for each simulated case. In case the resulting $S_{ng}$ is larger than the nominal apparent power of the generators it is set to the nominal value. This means that there is no room for conventional generation disconnection in this case. The disconnection of conventional generation due to both load variations and RES penetration is considered in the following way. Since the generators are considered equivalent generators, reducing the nominal power, is equivalent to a reduction in the moment of inertia of the power plant and an increase in the generator reactance. It should be noted that introducing RES and disconnecting conventional generators is expected to cause changes in the dynamic behavior of the generators.

C. Application of RQA on Generator Rotor Angle Responses

Initially, three representative Test Cases (TCs) chosen from a large number of performed MC simulations are shown as in Fig. 2. TC1 is well damped, TC2 is exhibiting oscillations with growing amplitude and TC3 has sustained oscillations indicating poor damping. RR and DET are calculated for all 16 generators in the system and presented in Fig. 3. In the specific studied system under the given operating conditions considered, the three TCs presented are representative rotor angle responses after a disturbance, which are “first-swing” stable.

In Fig. 4, the values of RR and DET are presented for the three TCs. In cases TC2 and TC3 when growing or sustained oscillations are observed, low values for both RR and DET are obtained for the generators exhibiting this behavior. However, higher values are obtained for the well damped oscillations of TC1. More specifically G1, G8 and G9 in TC2 which exhibit growing oscillations, have the lowest RR and DET values. In a similar manner G2-G7, G9 and G10 in TC3, exhibit poorly damped oscillations and therefore have low values of RR and DET. On the other hand, G10 and G12 in TC2 are examples that exhibit well damped oscillations and therefore have higher RR and DET values. Furthermore, in TC1 G11 and G14-G16 have relatively low RR values (compared to the rest of the generators) but high DET values. The relatively low RR values indicate lower damping in the responses of these generators compared to the rest (G1-G10, G12) for this TC. However, the responses are still well damped and RR has larger value (more than 0.1) compared to oscillatory unstable cases in TC2 and TC3. It can be concluded that the information from the values of RR and DET could be useful to detect poorly damped and oscillatory behavior in the system.

In general, we can summarize three distinct behaviors of the specific test system under the operating conditions studied: 1) Generators that exhibit both low values of RR (lower than approximately 0.1) and low DET values at the same time. These generators either exhibit poorly, or negatively damped oscillations. 2) Generators that exhibit relatively low values of RR (but higher than 0.1) and high values of DET. This usually suggests damped oscillations but with higher settling time.
generators have better damped responses.

In general, values of RR lower than approximately 0.07 and DET lower than approximately 0.85 suggest the existence of either poorly or negatively damped oscillations. It should be mentioned that more than two groups of generators could have been defined in some cases as it can be seen in Figs 5 and 6. However, the number of clusters defined by k-Means is chosen to be 2 since the main aim is to identify the group of generators possibly requiring additional stabilizing control actions.

Apart from the obtained groups through clustering, it is also essential to observe the position of the cluster centroids in order to draw conclusions about the behavior of the group. For example, some of the groups might not require additional stabilizing control actions and vice versa. The centroids of the two clusters for the three representative TCs are presented in Fig. 6b. For TC1 where all oscillations are well damped, both cluster centroids have high DET values and RR value is larger than 0.1. For TC2 one cluster has very low values of both RR and DET, indicating that this is the group with negatively damped oscillations. TC3 is similar to TC2 with the two clusters being closer. For the group exhibiting sustained oscillations (G2-G7, G9, G10) the cluster centroid has RR value lower than 0.07 and DET lower than 0.8. Therefore, by observing the cluster centroids for each case, information on the oscillatory behavior of the generator groups can be extracted.

A. Clustering Generator Responses

In order to identify groups of generators that exhibit similar oscillatory behavior for a given disturbance, the k-Means clustering algorithm is applied using RR and DET of the rotor angles of the 16 generators as features. The number of clusters is chosen as 3, in order to identify two groups of generators for each contingency; one with poorly or negatively damped oscillations and one with well damped oscillations. It should be noted that clustering the generators in two groups does not necessarily mean that one group is well damped and the other poorly damped. There are also cases when all generators exhibit well damped responses. In this case, one group will be relatively more oscillatory than the other, as for example in TC1. For this reason it is also important to observe the cluster centroids as presented later in this section.

The clusters obtained match with the described behavior for all three TCs and are presented in Figs. 5 and 6. As mentioned before, in TC1 the responses of all generators are well damped. Both groups have high values of DET (greater than 0.95) and RR values greater than 0.1. The first group, consisting of G1-G10 and G12, has better damped responses while the second group consisting of G11 and G14-G16 has relatively lower damping. In TC2, the second group consisting of G1, G8 and G9 exhibit responses with negative damping while the responses of the rest of the generators are positively damped. The second group has low values of both DET (lower than 0.65) and RR (lower than 0.07). G14-G16 also exhibit relatively low damping with RR values higher than 0.07. However, DET has values higher than 0.9. For TC3, the second group, consisting of G2-G7, G9 and G10 exhibit sustained oscillations with very low damping while the rest of the generators have better damped responses.

Figure 3. Generator rotor angles for a) well damped response – TC1, b) growing oscillations – TC2 and c) sustained oscillations – TC3.

Figure 4. a) RR and b) determinism for TC1-TC3.

IV. IDENTIFYING GENERATOR GROUPS BASED ON SIMILARITY OF THEIR OSCILLATORY BEHAVIOR

A. Clustering Generator Responses

In order to identify groups of generators that exhibit similar oscillatory behavior for a given disturbance, the k-Means clustering algorithm is applied using RR and DET of the rotor angles of the 16 generators as features. The number of clusters is chosen as 2, in order to identify two groups of generators for each contingency; one with poorly or negatively damped oscillations and one with well damped oscillations. It should be noted that clustering the generators in two groups does not necessarily mean that one group is well damped and the other poorly damped. There are also cases when all generators exhibit well damped responses. In this case, one group will be relatively more oscillatory than the other, as for example in TC1. For this reason it is also important to observe the cluster centroids as presented later in this section.

The clusters obtained match with the described behavior for all three TCs and are presented in Figs. 5 and 6. As mentioned before, in TC1 the responses of all generators are well damped. Both groups have high values of DET (greater than 0.95) and RR values greater than 0.1. The first group, consisting of G1-G10 and G12, has better damped responses while the second group consisting of G11 and G14-G16 has relatively lower damping. In TC2, the second group consisting of G1, G8 and G9 exhibit responses with negative damping while the responses of the rest of the generators are positively damped. The second group has low values of both DET (lower than 0.65) and RR (lower than 0.07). G14-G16 also exhibit relatively low damping with RR values higher than 0.07. However, DET has values higher than 0.9. For TC3, the second group, consisting of G2-G7, G9 and G10 exhibit sustained oscillations with very low damping while the rest of the generators have better damped responses.

In general, values of RR lower than approximately 0.07 and DET lower than approximately 0.85 suggest the existence of either poorly or negatively damped oscillations. It should be mentioned that more than two groups of generators could have been defined in some cases as it can be seen in Figs 5 and 6. However, the number of clusters defined by k-Means is chosen to be 2 since the main aim is to identify the group of generators possibly requiring additional stabilizing control actions.

Apart from the obtained groups through clustering, it is also essential to observe the position of the cluster centroids in order to draw conclusions about the behavior of the group. For example, some of the groups might not require additional stabilizing control actions and vice versa. The centroids of the two clusters for the three representative TCs are presented in Fig. 6b. For TC1 where all oscillations are well damped, both cluster centroids have high DET values and RR value is larger than 0.1. For TC2 one cluster has very low values of both RR and DET, indicating that this is the group with negatively damped oscillations. TC3 is similar to TC2 with the two clusters being closer. For the group exhibiting sustained oscillations (G2-G7, G9, G10) the cluster centroid has RR value lower than 0.07 and DET lower than 0.8. Therefore, by observing the cluster centroids for each case, information on the oscillatory behavior of the generator groups can be extracted.
The proposed clustering procedure can be used in an online manner after a disturbance occurs. An online identification method of power system transient stability status as proposed in [11] can be initially used to distinguish between the stable and unstable cases. If a “first-swing” stable case is detected the clustering methodology presented in this section can be applied to identify groups of generators with poorly or well damped oscillations.

### B. Application of Clustering to a Large Number of MC Simulations

The k-Means clustering algorithm (with k=2) is also applied to all the MC simulated contingencies. The overall number of distinct patterns identified is 323 due to the large number of possible combinations between the 16 generators of the system. The most common patterns (observed more than 50 times) and accounting for approximately 70% of all simulated cases are presented in Table II.

The three most common patterns are number 3, 6 and 5 which correspond to well damped responses with one group exhibiting better damped responses than the other. The group with better damped responses varies slightly between these patterns and consists mainly of generators of NETS and NYPS area (pattern 6), with some generators from NYPS area being removed from this group according to the specific pattern. For example, G11 is removed from the group exhibiting better damping in pattern 3 and both G11 and G12 in pattern 5. Patterns 1, 4, 9 and 12 are examples of patterns when poorly damped oscillations are observed. Patterns 1 and 12 are again slight variations with G10 and G12 being added to the group exhibiting poorly damped oscillations in pattern 12.

In Fig. 7 the cluster centroids for all the cases that pattern 1 occurs, are presented as an example. Group 1 and 2 are defined in the order shown in Table II. The two groups are clearly defined for all the cases. However, this might not be the case for all the patterns. Moreover, there can be variations in the location of the centroids from case to case. For the pattern shown in Fig. 7, the maximum variation of the RR value of the centroid is approximately 25%. It is therefore important to observe the centroids of the obtained clusters for each specific case in order to draw conclusions about the dynamic behavior of the generators belonging to each group.

From the MC simulations performed, the dynamic behavior of each individual generator is also investigated. K-Means is used again but this time to cluster the rotor angle responses of each generator separately into poorly or well damped for all the MC simulated cases. RR and DET are again used as features but the observations this time are the N\(_0\) simulations for each generator separately. This way, two cluster centroids for each generator are obtained that describe its dynamic behavior, considering the uncertainties described earlier. The centroids for the 16 generators are shown in Fig. 8.

By observing the locations of the centroids of the two clusters for each generator we can conclude that G1-G10 exhibit poorly damped oscillations in some cases, since one of their cluster centroids is exhibiting low values of both RR and DET (as in TC2, TC3). On the other hand, for G11-G16 both clusters have relatively higher values of RR and especially DET which points to better damped oscillations (as in TC1). A representative example for G9 and G16 is also shown in Fig. 9 for all the simulated responses. It is concluded, that the cluster
centroids are a rough indication of the generator dynamic behavior. Individual cases with different behavior extending relatively far from the cluster centroid might exist, as shown in Fig. 9.

![Figure 8. Cluster centroids for the 16 generators of the system for all MC simulations.](chart8)

![Figure 9. Clustering results for all MC simulated cases for a) G9 and b) G16.](chart9)

V. CONCLUSIONS

The paper presented a Recurrence Quantification Analysis based methodology to cluster generators based on similarity of their post disturbance oscillatory behavior. The methodology focuses on cases that do not exhibit aperiodic first-swing instability. It is shown that two RQA metrics (Recurrence Rate and Determinism) can be used to identify features relevant to frequency and damping but not the amplitude of the generator oscillatory responses. RR and DET of generator rotor angle responses can be used as parameters for the k-Means clustering algorithm to cluster generators in groups exhibiting well or poorly damped oscillations. The centroids of the clusters need to be observed in order to draw conclusions about the dynamic behavior of the generators belonging to each group. In general low values of RR and DET (lower than 0.07 and 0.85 respectively) suggest poorly or negatively damped oscillations. The methodology presented can be used online (following the disturbance) to trigger possible stabilizing control actions in cases that do not lead to aperiodic instability. The dynamic behavior of individual generators is also investigated by applying k-Means clustering on a large number of rotor angle responses with RR and DET used as input parameters. The obtained centroids can provide a rough estimate of the oscillatory behavior of each generator, considering the uncertainties present in power systems.

In general, it is shown that RQA metrics can provide sufficient information that can be used by a clustering algorithm to distinguish between well damped and poorly damped oscillations. RQA can be directly applied to nonlinear non-stationary data and thus its application is straightforward. However, since RR and DET do not directly have a physical meaning related to damping, frequency and especially amplitude of oscillations, observation of a large number of disturbances for a given system is required to draw useful conclusions.

REFERENCES


