4

Integrate Big Data for Better Operation, Control, and Protection of Power Systems

Guang Lin

CONTENTS

4.1 Introduction ........................................................................ 47
   4.1.1 Types of Big Data in Power Systems and Acquisition ....... 48
   4.1.2 Big Data Integration Quality Challenges in Power Systems ... 48

4.2 Integrating Big Data for Model Reduction and Real-Time Decision Making ............................................. 48
   4.2.1 Data Analysis Algorithms for Model Reduction ............... 49
      4.2.1.1 \( k \)-Means Clustering-Based Model Reduction .......... 49
      4.2.1.2 Dynamic-Feature Extraction, Attribution, and Reconstruction Algorithm for Model Reduction .......... 49
      4.2.1.3 Case Study .................................................... 52

4.3 Summary ........................................................................... 59

4.4 Glossary ............................................................................ 59

References ............................................................................. 60

As deploying the smart grid infrastructure to power systems, new challenges rise in handling extremely large datasets for better decision making. It is critical to utilize such large datasets in a power system for better operation, control, and protection. In particular, we will focus on discussing data types and acquisition, data management and how to integrate big data for feature extraction, systematic integration, and contingency analysis of power systems.

4.1 Introduction

The term big data usually refers to large datasets that are difficult to store, process, and analyze with traditional database and analysis tools. As we are building the smart grid infrastructure, we are facing new challenges in analyzing big data for more accurate and robust operation, control, and protection. There are three characteristics of big data in power systems as indicated in [9]: large volume, high velocity, and increasing variety. While building smart grid infrastructure, many metering devices have been introduced to the power systems, which dramatically increase the data volume. In addition, many phasor measurement units (PMU) or synchrophasors have been employed into power systems, which significantly increase the data velocity. The PMU device measures the electrical waves on an electricity grid, using a common time source for synchronization. Time synchronization...
allows synchronized real-time measurements of multiple remote measurement points on the grid. The PMU device can measure 50/60 Hz AC waveforms (voltages and currents) typically at a rate of 48 samples per cycle (2880 samples per second). The collection of many different types of meter data dramatically increases the variety of big data and places additional difficulty on analyzing such big datasets.

### 4.1.1 Types of Big Data in Power Systems and Acquisition

Various types of data are tagged with metadata and annotations. Metadata consist of additional information that would help data use, security, and validation. For example, the interpretation, the ownership, the meter device, and the collection location and the date are part of the metadata. Annotations to data can greatly enhance the data search capability and enable better collaboration and workflow development.

The types of acquired data in power systems are equipment parameters, network topology, and connectivity information among different power system components, and anomalous network conditions.

The data in power systems could be either static or real time. Static information is often used for planning for which steady-state data are sufficient. On the other hand, real-time data contain the information that is operational such as current, voltage, and power flow. Such real-time data are often used for real-time operation and control, and in close collaboration with independent system operators and regional transmission organizations, and so on. If big datasets are employed for improving reliability and emergency response, real-time data are required. In contrast, data that require minimal updating can be used for research and planning efforts.

### 4.1.2 Big Data Integration Quality Challenges in Power Systems

Anomalies in grid data and inconsistencies among data sources must be resolved before applying the data. Inconsistencies may exist in data from different sources. The dataset must pass through rigorous data validity and quality assurance procedures. The structure and requirements of this process affect model acceptance, validity, and credibility.

### 4.2 Integrating Big Data for Model Reduction and Real-Time Decision Making

Past lessons learned from large blackouts that affect the grid in recent years have shown the critical need for better situation awareness about network disturbances such as faults, sudden changes from renewable generation, and dynamic events. Hence, integrating big data for real-time decision making will greatly enhance the efficiency and security of grid operation. However, the dynamic state estimation model for a large power network is computational expensive, which cannot meet the need for real-time decision making. To achieve this objective, in interconnected power systems, dynamic model reduction can be applied to generators outside the area of interest (i.e., study area) to reduce the computational cost associated with transient stability studies using the real-time big data. Our approach is to utilize real-time big data to reduce a large system of power grid governing equations to a smaller systems for which the solution is an approximate solution to the original system. Our goal is to employ big data that show how generators are behaving to (1) cluster
the generators based on similar behavior, and then (2) choose a representative element from each cluster. In doing this we have reduced the system model from the full number of generators to a smaller set of generators. When the full system is too computationally intensive to simulate for forecasting, islanding, contingency analysis, or other simulation activities to support real-time decision making, we can use the reduced system that contains all of the same fundamental behaviors as the full grid. Using this reduced system allows us to save computational time, which is a necessary reduction for real-time decision making.

One of the big data in interconnected power systems is PMU datasets, which collect data synchronously across the power grid. These units are deployed to many systems throughout the grid and are time synchronized so that measurements taken at different locations can be correlated together. The data that we care about, collected from the PMUs, are rotor angle data. These data are collected at the millisecond resolution and give a picture of the oscillation of the rotor angle at each generator. For each generator a time series of rotor angles is collected, \( \vec{\delta}_i = \langle \delta_{i}^{(1)}, \delta_{i}^{(2)}, \ldots, \delta_{i}^{(m)} \rangle \), where \( \delta_{i}^{(j)} \) is the rotor angle of generator \( i \) at the \( j \)th time.

### 4.2.1 Data Analysis Algorithms for Model Reduction

In this section, we describe several algorithms we use for model reduction of power systems and characteristic generator identification using real-time big data, PMU data.

#### 4.2.1.1 \( k \)-Means Clustering-Based Model Reduction

\( k \)-means clustering is a standard clustering technique [14] and is widely used in many applications [8,10]. The general idea is a recursive algorithm that, in each step, computes centroids of each of the clusters and then reassigns points to the cluster whose centroid it is closest to. This needs an initialization of cluster centroids, which is often done randomly. The algorithm then runs as follows: assign each point to the cluster whose randomly chosen centroid is closest, recompute centroids, reassign points to clusters, recompute centroids, and so on. This is repeated for some predetermined number of steps, or until clusters do not change and the algorithm has converged. One problem with \( k \)-means clustering is that there is no guarantee that it will terminate in a globally optimal clustering; it is possible to get stuck in a local minimum. Because the initialization is done randomly, there can be multiple clusterings from the same input data.

#### 4.2.1.2 Dynamic-Feature Extraction, Attribution, and Reconstruction Algorithm for Model Reduction

The dynamic-feature extraction, attribution, and reconstruction (DEAR) algorithm [15] is based on the singular value decomposition (SVD) or randomized algorithm [13] to extract the first few principal components for model reduction. The DEAR model reduction method can be summarized as follows:

- **Step 1:** Dynamic-feature extraction—analyze the dynamic response vectors of the original system for a disturbance, and use the SVD or randomized algorithm to extract the first few principal components as the optimal orthogonal bases of these responses.

- **Step 2:** Feature attribution—determine characteristic generators by identifying generators with responses that are highly similar to the optimal orthogonal bases.

Analyze the similarity between \( \delta_i \) and \( x_i \); those \( \delta_i \) with the highest similarity to \( x \) are selected to form a subset of \( \delta \), which is called a set of characteristic generators.
• Step 3: Feature reconstruction—use characteristic generators to describe the full model.

Responses of the noncharacteristic generator are approximated by linear combinations of characteristic generators so that only dynamic equations of characteristic generators are kept and a reduced model is obtained.

4.2.1.2.1 Dynamic-Feature Extraction: Finding the Optimal Orthogonal Bases

Here, we discuss the concept of optimal orthogonal bases of a system’s dynamic responses and how to identify them using dynamic PMU big dataset.

For convenience, the classical generator model is assumed, and rotor angles, δ, are the state variables. The magnitude of the generator internal voltage, $E'$, is assumed to be constant. Suppose $\delta_1, \delta_2, \ldots, \delta_i, \ldots, \delta_m$ are the $m$ rotor angles of the system to be reduced. The term $\delta_i$ is a normalized $n$-dimensional row vector representing the dynamic of rotor angle $\delta_i$ following a disturbance. Its elements are time series: $\delta_i(t_1), \delta_i(t_2), \ldots, \delta_i(t_n)$. Define

$$\delta = [\delta_1; \delta_2; \ldots; \delta_m]$$  \hspace{1cm} (4.1)

Here, $\delta$ is an $m \times n$ matrix. Suppose $x = [x_1; x_2; \ldots; x_i; \ldots; x_r]$ is the first $r$ principal components from SVD. Here, $x_i$ is an $n$-dimensional row vector, and $r < m$. Optimal means that for any $r < m$, $\delta_i$ can be approximated by a linear combination of $x$, and the errors between the approximation, $\hat{s}$, and the actual responses are minimized. In other words, given any $r < m$, we need to find an optimal $x$, such that $\hat{s} = \kappa x$, where $\kappa$ is an $m \times r$ matrix, and

$$||\delta - \hat{s}||_2 = \sum_{i=1}^{m} (\delta_i - \hat{s}_i)^T (\delta_i - \hat{s}_i)$$  \hspace{1cm} (4.2)

is minimized. We adopt the SVD algorithm for a moderate size of dataset or employ a randomized algorithm [13] for big data to solve the above problem. In particular, through SVD we get

$$\delta = U D W^T$$  \hspace{1cm} (4.3)

where $U$ is an $m \times m$ unitary matrix, $D$ is an $m \times n$ rectangular diagonal matrix with nonnegative real numbers on the diagonal, and $W^T$ is an $n \times n$ unitary matrix. The first $r^2$ rows of $W^T$ constitute the optimal orthogonal bases, which can be used to approximate $\delta$. The diagonal elements of $D$ (i.e., singular values of $\delta$) in descending order are scaling factors indicating the strength or energy of the corresponding row vectors of $W^T$. Define

$$x = W^T (1: r, :)$$

$$\kappa = T(:, 1: r)$$  \hspace{1cm} (4.4)

where $T = UD$; $\kappa$ is the first $r$ columns of $T$. As a result, $\delta$ can be approximated by $\hat{s} = \kappa x$. For any $r < m$, the SVD algorithm can guarantee that Equation 4.2 is minimized.

Remark 4.1 Although there may exist thousands of modes in the mathematic model of a large-scale power system, usually only a small fraction of them are noticeably excited by a disturbance. We refer to these modes as dominant modes, and the rest as dormant modes. Usually, dominant modes have strong oscillation energy (shown as the features in system dynamics), while dormant modes have weak energy and are hard to observe. Here, SVD is used to extract the features composed of those dominant modes, which correspond to the first $r$ diagonal elements of matrix $D$ in Equation 4.3, unlike how it is used in traditional linear system model reduction methods (e.g., balanced truncation [11] or Hankel norm reduction [7]). More details about the SVD algorithm can be found in [4] and [2].
4.2.1.2.2 Feature Attribution: Determine Characteristic Generators

The \( r \) optimal orthogonal basis vectors found in the feature extraction step will result in minimal errors when used to approximate \( \delta \). If each of these basis vectors exactly matches the dynamic angle response of one of the generators, then the angle dynamics of the other generators must have very minimal energy impact (because their corresponding singular values are smaller). This means that we can just keep these \( r \) generators in the model and ignore the other generators. Although this will not happen in a real system because generators usually participate in multiple oscillation modes, we still will try to match an optimal basis with the generator whose dynamic response has the highest similarity to this oscillation pattern, and will call the generator a characteristic generator. We will then use the set of characteristic generators as suboptimal bases to represent the entire system. To determine the characteristic generators, we need to find a subset of \( \delta = [\delta_1; \delta_2; \ldots; \delta_m] \) such that this subset has the highest similarity to \( x \), the optimal orthogonal set. In other words, we need to find

\[
\xi = [\delta_p; \delta_q; \ldots; \delta_z] \quad (r = z - p + 1)
\]  

such that \( \xi \) is highly similar to \( x \). According to the last subsection, any \( \delta_q \) can be approximated by a linear combination of the optimal orthogonal bases:

\[
\delta_p \approx \kappa_{p1} x_1 + \kappa_{p2} x_2 + \cdots + \kappa_{pi} x_i + \cdots + \kappa_{pr} x_r
\]

\[
\delta_q \approx \kappa_{q1} x_1 + \kappa_{q2} x_2 + \cdots + \kappa_{qi} x_i + \cdots + \kappa_{qr} x_r
\]

Here, \( x \) is the optimal orthogonal basis and \( \delta \) is normalized [1]. A larger \(|\kappa_{qi}|\) indicates a higher degree of collinearity between the two vectors (\( \delta_q \) and \( x_i \)). For example, if \(|\kappa_{qi}| > |\kappa_{pi}|\), it indicates that the similarity between \( x_i \) and \( \delta_q \) is higher than that between \( x_i \) and \( \delta_p \). \( \delta_q \) will have the highest similarity to \( x_i \), if the inequality in Equation 4.7 holds.

\[
|\kappa_{qi}| > |\kappa_{pi}| \quad \text{for } \forall p \in \{1, 2, \ldots, m\} \quad p \neq q
\]  

By doing so, a rotor angle response of the highest similarity can be identified for each optimal orthogonal basis. As a result, \( \xi \) in Equation 4.5 is determined.

Remark 4.2 The same characteristic generator can appear two or more times in Equation 4.5 when using the criteria in Equation 4.7. For example, if \( \delta_q \) has the highest similarity to both \( x_i \) and \( x_j \), then we will have two \( \delta_q \) entries in Equation 4.5. In that case, delete one of the entries, and thus, \( r = r - 1 \).

Remark 4.3 From an engineering perspective, some generators may be of particular interest, and detailed information about them is preferred. Dynamic equations for these generators can be kept in the reduced model without being approximated if they are not identified as characteristic generators.

4.2.1.2.3 Feature Reconstruction: Model Reduction Using the Linear Combination of Characteristic Generators

According to Equations 4.3–4.7, \( \delta \) can now be arranged as Equation 4.8:

\[
\delta = \begin{bmatrix} \xi \\ \bar{\xi} \end{bmatrix} \approx \begin{bmatrix} \kappa_{\xi} \\ \kappa_{\bar{\xi}} \end{bmatrix} x
\]

where \( \delta \) is an \( m \times n \) matrix defined in Equation 4.1, \( \xi \) is an \( r \times n \) matrix defined in Equation 4.5, representing rotor angle dynamics of characteristic generators, and \( \bar{\xi} \) is an
\((m - r) \times n\) matrix representing the dynamics of noncharacteristic generators; \(x\) is an \(r \times n\) matrix and can be calculated from Equation 4.5; and \(\kappa_{\xi}\) is an \(r \times r\) square matrix; \(\kappa_{\xi}\) is an \((m - r) \times r\) matrix. Normally, \(\kappa_{\xi}\) is invertible. We have two different approaches to finding the approximate linear relations between \(\bar{\xi}\) and \(\xi\). The first approach is to solve the following overdetermined equation:

\[
\bar{\xi} = C\xi
\]

where \(C\) is an \((m - r) \times r\) matrix and can be determined by the least-squares method, namely, \(C = \bar{\xi}[((\xi\xi^T)^{-1}\xi]^T\). Another approach is to use the approximate linear relations in Equation 4.8. According to Equation 4.8, we have

\[
\xi \approx \kappa_{\xi}x
\]

and

\[
\bar{\xi} \approx \kappa_{\bar{\xi}}x
\]

Premultiplying \(\kappa_{\xi}^{-1}\) on both sides of Equation 4.9 yields

\[
x \approx \kappa_{\xi}^{-1}\xi
\]

Substituting Equation 4.11 into 4.10 yields

\[
\bar{\xi} \approx \kappa_{\xi}\kappa_{\xi}^{-1}\xi
\]

Equation 4.8 or 4.12 establishes the approximate linear relations between the rotor angle dynamics of characteristic generators and that of noncharacteristic generators. The dynamics of all generators in the original system then can be reconstructed by using only the dynamic responses from characteristic generators.

4.2.1.2.4 Generalization to High-Order Models

In classical models, it is assumed that the magnitude of the generator internal voltage \(E'\) is constant, and only its rotor angle \(\delta\) changes after a disturbance. In reality, with the generator excitation system, \(E'\) will also respond dynamically to the disturbance. The dynamics of \(E'\) can be treated in the same way as the rotor angle \(\delta\) in the above-mentioned model reduction method to improve the reduced model, except that the set of characteristic generators needs to be determined from \(\delta\). This way, both \(\delta\) and \(E'\) of noncharacteristic generators will be represented in the reduced model using those of the characteristic generators.

4.2.1.2.5 Online Application of the DEAR Method

For offline studies, the DEAR process can be performed at different conditions and operating points of the target system (external area) to obtain the corresponding reduced models. For online applications, however, computational cost may be very high if SVD has to be calculated every time the system configuration changes. A compromise can be made by maintaining a fixed set of characteristic generators, which is determined by doing SVDs for multiple scenarios offline and taking the super set of the characteristic generators from each scenario. During real-time operation of the system, the approximation matrix \(C\) from Equation 4.8 used for feature reconstruction, is updated (e.g., using the recursive least-squares method) based on a few seconds data right after a disturbance. This way, SVD is not needed every time after a different disturbance occurs.

4.2.1.3 Case Study

In this section, the IEEE 145-bus, 50-machine system [5] in Figure 4.1 is investigated. There are 16 and 34 machines in the internal and external areas, respectively. Generator 37 at
Bus 130 in the internal area is chosen as the reference machine. All generators are modeled using classical models. A three-phase, short-circuit fault (F1) is configured on Lines 116–136 at Bus no. 116 at $t = 1$ s. The fault lasts for 60 ms, and then the line is tripped to clear the fault. Postfault rotor angle dynamics in the time interval of $1.2 \leq t \leq 5$ s are analyzed to perform model reduction, using inertial aggregation [6] (one of the coherency-based reduction methods) and the DEAR method, so that their performance can be compared. Many methods are available for coherency identification. In this case study, the principal component analysis method presented by Anaparthi et al. [1] and Moore [12] is chosen to identify coherency groups, and the MATLAB® clustering toolbox is used to aid the analysis. Clustering results according to the rotor angle dynamics in the external area are shown in Figure 4.2. In Figure 4.2, the horizontal axis represents generator numbers, and the vertical axis scales distances between generator groups. Here the distance is defined in the three-dimensional Euclidean space expanded by the first three columns of the matrix $T$ in Equation 4.5. Depending on the distance selected between clusters, different number of coherency groups can be obtained. For example, at a distance larger than 9, two groups are formed (level-2 clustering). Generators 23, 30, and 31 comprise one group, and the other generators comprise another group. Similarly, there are 10 generator groups at level 10, which are shown in the following: Group 1 (generators 30, 31); Group 2 (generator 23); Group 3 (generators 9 and 10); Group 4 (generator 16); Groups 5 (generators 7, 13, and 15); Group 6 (generator 3); Group 7 (generators 32 and 36); Group 8 (generators 8, 18, 25, 33, 34, and 35); Group 9 (generators 2 and 6); Group 10 (generators 1, 4, 5, 11, 12, 14, 17, 19–22, 24, 26,
Fewer groups result in a simpler system. The normalized (i.e., subtracted by the mean value of the data and divided by its standard deviation) angle dynamics of the 10 groups at level 10 are shown in Figure 4.3, where coherency can be observed between generators in the same group. These coherent machines are then aggregated using the inertial aggregation method reported by Chow et al. [6]. Finally, we obtain a reduced system with 10 aggregated generators for the external system. Following the procedure described above, the optimal orthogonal bases are first obtained by Equations 4.3 and 4.5 and by setting $r = 10$. These 10 basis vectors are shown as the blue solid lines in Figure 4.4. Then, the corresponding 10 characteristic generators are identified using Equation 4.7. The rotor angle dynamics of these characteristic generators are shown as dashed red lines in Figure 4.4. An approximate linear relation between the characteristic generators and the noncharacteristic generators is then established to get the reduced model. Notice that, in this case, $\delta_2$ has the highest similarity to orthogonal bases $x_7$ and $x_8$. Therefore, the set of characteristic generators contains only nine elements, which is $\xi = [\delta_{27}, \delta_{15}, \delta_{30}, \delta_{36}, \delta_{23}, \delta_2, \delta_{18}, \delta_4]^T$. With the reduced models developed using both coherency aggregation and the DEAR method, the performance of these two methods can be compared. Under the foregoing disturbance, the dynamic responses of generator G42 (connected to the faulted line) from these two reduced models and from the original model are shown in Figure 4.5. The blue solid line represent the original model. The red dashed-dotted line represents the reduced model by coherency aggregation, and the black dotted line by the DEAR method. The reduced model by the DEAR method appears to have smaller differences from the original model, and outperforms the coherency aggregation method. Another important metric for evaluating the performance of model reduction is the reduction ratio, which is defined as

$$R = \frac{(N_F - N_R)}{N_F} \quad (4.13)$$
FIGURE 4.3

where:

\[ N_R \] is the total number of state variables of the reduced model of the external system \\
\[ N_F \] is that of the original model

The mismatch between the black dotted line and the blue solid line in Figure 4.5 is 0.1630, and the reduction ratio defined by Equation 4.13 is \( R = (34 - 9)/34 = 0.7353 \), both of which represent the performance of the DEAR method. The mismatch between the red dashed-dotted line and the blue solid line in Figure 4.5 is 0.4476 and \( R = (34 - 10)/34 = 0.7059 \), both representing the performance of the coherency aggregation method. Therefore, it can be concluded that the DEAR method performs better, even under a slightly higher reduction ratio. We now investigate if the same conclusion can be drawn under different reduction ratios and for generators other than G42 shown in Figure 4.5. Define a comprehensive metric shown in Equation 4.14 for all the internal generators.
FIGURE 4.4

\[ J(i) = \frac{1}{N} \sum_{i \in \phi} J_s(i) \]  

(4.14)

where \( \phi \) is the set of all the generators in the internal system, \( N \) is the total number of these generators, and \( J_s(i) = \sqrt{\frac{1}{(t_2-t_1)} \int_{t_1}^{t_2} (\delta^q_i(t) - \delta^f_i(t))^2 dt} \). A performance comparison of the DEAR method and the traditional coherency aggregation is shown in Figure 4.6, in which the horizontal axis represents the reduction ratio defined in Equation 4.13, and the vertical coordinates represent the error defined in Equation 4.14. It is apparent that the DEAR method consistently performs better than the coherency method. To demonstrate the basic idea of a super set of characteristic generators in Section III.F, three faults (three-phase fault lasting for 60 ms) are configured on Lines 116–136, Lines 116–143, and Lines 115–143,
FIGURE 4.5

FIGURE 4.6
respectively. Choose the size $r$ of the characteristic generator set for each fault scenario to be 9 here. Define $D_g$ as the number of times each generator appears as a characteristic generator after analyzing the system dynamic response to a fault. For example, generator 2 is a characteristic generator in all the three fault scenarios; it, therefore, has a $D_g$ of 3. $D_g$ of each characteristic generator is shown in Table 4.1. Select $D_g = 2$ as a threshold, that is, generators with $D_g \geq 2$ are chosen to form the super set of characteristic generators. The following generators are selected: generators 2, 15, 18, 23, 30, 3, 4, 22, and 36. With the super set, three different coefficient matrices $C$ in Equation 4.8 can be obtained for the three faults, denoted by $C_1$, $C_2$, and $C_3$. Then a rough estimation of a generalized $C$ can be obtained by, for example, letting $C_g = (C_1 + C_2 + C_3)/3$. When another three-phase fault takes place, for example, F4 on Lines 141–143, measurement data for the first 3 s after fault clearance can be used to refine the coefficient matrix $C_g$. Applying the recursive least-squares method, $C_g$ is replaced by the more accurate coefficients in $C_4$. We thus get a new reduced model, represented by the super set of characteristic generators and the coefficient matrix $C_4$, without performing SVD. The performance of the new reduced model is illustrated in Figure 4.7 using the rotor dynamics of generator 42.

### TABLE 4.1

$D_g$ of characteristic generators.

<table>
<thead>
<tr>
<th>Gen. no.</th>
<th>2</th>
<th>15</th>
<th>18</th>
<th>23</th>
<th>30</th>
<th>3</th>
<th>4</th>
<th>22</th>
<th>36</th>
<th>1</th>
<th>1</th>
<th>1</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D_g$</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

![Graph showing the rotor angle dynamics of generator 42 following fault F4.](image)

**FIGURE 4.7**

4.3 Summary

In this chapter, types of big data in power systems and acquisition are introduced. Several big data analysis algorithms to integrate real-time big dataset for model reduction and real-time decision making are illustrated. Both $k$-means clustering and DEAR methods are presented. The network model is unchanged in the DEAR method, which makes online applications relatively easier and more flexible (e.g., generators of interest can be retained in the reduced model). Tests on the IEEE standard system shows that the DEAR method yields better reduction ratios and smaller response errors under both stable and unstable conditions than the traditional coherency-based aggregation methods. The introduced data analysis tools can effectively integrate real-time big data for better operation, and control and protection of power systems.

4.4 Glossary

Phasor measurement unit: A phasor measurement unit (PMU) or synchrophasor is a device that measures the electrical waves on an electricity grid, using a common time source for synchronization. Time synchronization allows synchronized real-time measurements of multiple remote measurement points on the grid. In power engineering, these are also commonly referred to as synchrophasors and are considered one of the most important measuring devices in the future of power systems.

$k$-Means clustering: $k$-means clustering is a method of vector quantization, originally from signal processing, that is popular for cluster analysis in data mining. $k$-means clustering aims to partition $n$ observations into $k$ clusters in which each observation belongs to the cluster with the nearest mean, serving as a prototype of the cluster.

Singular value decomposition: In linear algebra, the singular value decomposition (SVD) is a factorization of a real or complex matrix, with many useful applications in signal processing and statistics. Formally, the SVD of an $m \times n$ real or complex matrix $M$ is a factorization of the form $M = U\Sigma V^*$, where $U$ is an $m \times m$ real or complex unitary matrix, $\Sigma$ is an $m \times n$ rectangular diagonal matrix with nonnegative real numbers on the diagonal, and $V^*$ (the conjugate transpose of $V$, or simply the transpose of $V$ if $V$ is real) is an $n \times n$ real or complex unitary matrix. The diagonal entries $\Sigma_{i,i}$ of $\Sigma$ are known as the singular values of $M$. The $m$ columns of $U$ and the $n$ columns of $V$ are called the left-singular vectors and right-singular vectors of $M$, respectively.

Randomized algorithm: A randomized algorithm [13] is an algorithm that employs a degree of randomness as part of its logic. The algorithm typically uses uniformly random bits as an auxiliary input to guide its behavior, in the hope of achieving good performance in the average case over all possible choices of random bits.

Model reduction: Model reduction [3] is used to produce a low-dimensional system that has the same response characteristics as the original system with far less storage requirements and much lower evaluation time. The resulting reduced model might be used to replace the original system as a component in a larger simulation or it might be used to develop a low-dimensional controller suitable for real-time applications.
References


