Clustering based Online Identification of Secondary Dynamic Parameters for Measurement based Composite Load Modeling

Shuqing Zhang, Member, IEEE, Xiaofei Liu, Junwei Cao, Senior Member, IEEE, Guoyu Tu, Yuxin Wan, Chao Lu, Senior Member, IEEE, and Yingduo Han, Senior Member, IEEE

Abstract—accurate load modeling and parameter identification of power systems is always a difficult problem, remaining unsolved but critical for stability analysis, prediction and decision-making of power systems. The development of wide area measurement system (WAMS) provides possible ways to further address the challenge. In this work, based on an existing load modeling method for online identification of dominant parameters, we put forward an improvement with the clustering method, to get the reactance of the composite load model as a secondary dynamic parameter. Corresponding theoretical analysis, design principles and system implementation are presented. The reactive power dumping time constant during disturbance is chosen as the clustering feature. Simulation results show effectiveness of our improvement with satisfactory accuracy.

Index Terms—parameter identification, clustering, secondary dynamic parameters, WAMS, composite load modeling

I. INTRODUCTION

POWER system modeling and simulation serves as useful tools in the analysis, design, construction and operation of power systems[1]. Mathematical models that can effectively represent characteristics and performance of actual power systems are essential to obtain accurate simulation results. Nowadays, the rapid growing of power consumption and the expanding scale of inter-connected power grid has resulted in higher requirements for power system modeling.

As we know, power systems consist of three parts: generation, transmission and distribution, and loads. The models and parameters of the first two parts are getting mature. However, load modeling is still a challenging issue due to the complexity, time variation and stochastic characteristics of loads. The development of phasor measurement units (PMU) and wide area measurement systems (WAMS) paves the way to address the load modeling challenge [2][3]. WAMS can provide synchronized real-time data at a regional and national scale[4], which are likely to lead to measurement based composite load modeling with higher accuracy and real-time characteristics[5].

In previous work, online real-time dynamic measurements based identification is proposed in [6][7]. Due to high requirement for real-time performance, this method can only identifies 4 dominant parameters of the composite load mode, ignoring some secondary dynamic parameters. In this work, challenges for online identification of secondary dynamic parameters are further addressed.

Studies in [6] and simulation results in this paper show that secondary dynamic parameters, such as rotor reactance and stator reactance, can greatly influence identification results. Unfortunately, the reactance parameters can hardly be identified directly. According to theoretical analysis, direct response analysis gets more of qualitative results due to the system complexity. Instead of getting identifying values, we employ clustering based methods to roughly get secondary parameters to meet practical requirements. Performance evaluation shows that satisfactory classification accuracy can be achieved.

Measurement-based methods have been proposed to obtain mechanism model of a certain load cluster [8][9]. Theoretical and practical issues relevant to load modeling and identification are hot research topics [10]; and various approaches are applied and developed for composite load modeling. A hybrid learning algorithm combines genetic algorithm (GA) and nonlinear Levenberg-Marquardt (L-M) algorithm [11], which takes advantages of the global search ability of GA and the local search ability of L-M algorithm. However, heuristic algorithms are not feasible for fast online identification due to the massive search space. Reducing identified parameters has been proposed in [12], with similar motivation of our work. Instead of full parameter identification, an adequate parameter set can be chosen to reduce computing overhead while still maintaining the model’s capability on describing load dynamics[13][14]. While support vector based clustering methods are also adopted in [14][15] for load modeling, we apply clustering for online identification of secondary dynamic load parameters. Although clustering method is used in component-based modeling [15], its usage in measurement based load modeling is rare.

The rest of the paper is organized as follows. The issue of dynamic parameter identification is described in Section II, with special focus on the necessity of secondary parameter identification; our clustering approach is presented in Section III with both theoretical analysis and simulation study. Detailed information on design principles, system implementation and performance evaluation is included in Section IV; Section V concludes the paper with future research directions.

This work was supported in part by National Natural Science Foundation of China (grants No. 61233016 and No. 51037002). This work was also funded by Academic Exchange Foundation of Tsinghua National Laboratory for Information Science and Technology (TN List).

Shuqing Zhang is Assistant Professor of Department of Electrical Engineering, Tsinghua University, Beijing 100084, P. R. China (e-mail: zsq@tsinghua.edu.cn).
II. DYNAMIC PARAMETER IDENTIFICATION

A. Dominant Parameter Identification

The algorithm we tried to improve belongs to the measurement-based modeling approach [16], aiming at identifying dominant dynamic parameters of a mechanistic model of power loads. Fig. 1 shows the structure of the model.

![Model structure](image)

Fig. 1 Model structure

The model is the parallel connection of a constant resistance and an induction motor, shown in Fig. 2.

![Induction motor model structure](image)

Fig. 2 Induction motor model structure

According to the mechanism analysis of the algorithm [17], there are four dominant parameters in the model, the portion of dynamic load $P_{ct}$, initial slip $s_0$, rotor resistance $R_r$, and inertia time constant $H$. To search in the four-dimension parameter, the algorithm is divided into two steps. The first step is based on analysis of the “0+” response of the composite load model to identify the percentage of the dynamic load and the angel of transient electrical force of rotor. After that, based on Volterra Model and pattern recognition [18], the algorithm is able to derive four parameters. Other parameters in the model are called secondary parameters and set as typical values. However, sensitivity analysis [19] and simulation experiments show that so called secondary parameters such as rotor reactance $X_r$ and stator reactance $X_s$ are also of great sensitivity to model accuracy.

B. Secondary Parameter Identification

In actual power systems, secondary parameter values have great dispersion and cannot be represented by typical values. Only identifying four dominant parameters is not enough in practice. A case study is given below where two loads with same load parameters except for $X_s$ are linked to the same bus (to ensure the synchronization of voltage). The IEEE-14 Bus System [20] is selected as the simulated system, as illustrated in Fig. 3.

![IEEE-14 Bus System structure](image)

Fig. 3 IEEE-14 Bus System structure

In the simulation experiment, the two loads are linked to Bus 14. Load 2 is set as typical values while Load 1 is not.

Their parameter settings are listed below:

<table>
<thead>
<tr>
<th></th>
<th>$H$</th>
<th>$P_{ct}$</th>
<th>$R_s$</th>
<th>$X_s$</th>
<th>$X_m$</th>
<th>$R_r$</th>
<th>$X_r$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Load 1</td>
<td>10</td>
<td>.40</td>
<td>.0116</td>
<td>.095</td>
<td>3.50</td>
<td>.020</td>
<td>.12</td>
</tr>
<tr>
<td>Load 2</td>
<td>10</td>
<td>.40</td>
<td>.0116</td>
<td>.295</td>
<td>3.50</td>
<td>.020</td>
<td>.12</td>
</tr>
</tbody>
</table>

Then parameters are identified using the algorithm [6][7] via perturbation analysis. The result is shown below:

<table>
<thead>
<tr>
<th></th>
<th>$P_{ct}$</th>
<th>$R_s$</th>
<th>$s_0$</th>
<th>$H$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Load 1</td>
<td>3.52</td>
<td>0.0378</td>
<td>0.0052</td>
<td>1.5323</td>
</tr>
<tr>
<td>Load 2</td>
<td>3.52</td>
<td>0.0182</td>
<td>0.0096</td>
<td>2.3134</td>
</tr>
</tbody>
</table>

As shown in above results, secondary parameters can greatly influence identification results. Hence it is necessary to gain more information about secondary parameters. Unfortunately, the rotor reactance $X_r$ and stator reactance $X_s$ can hardly be identified directly. According to theoretic analysis in Section III, direct response analysis is more of qualitative results due to system complexity. Instead of identifying exact values, clustering is adopted to roughly identify secondary parameters for fast online usage.

III. THE CLUSTERING APPROACH

The basis of clustering is feature selection and extraction. The reactive power response curves of two loads are shown in Fig. 4.

![Reactive power response curves of two loads](image)

Fig. 4 Reactive power response curves of two loads

There are three major differences between the responses of the two loads. The first one is the dumping time constant during the disturbance. The second one is the peak amplitude after the disturbance and after the resection of the disturbance. The third one is the oscillation mode after the resection of the disturbance. The third one requires more data samples, while the second one is greatly influenced by many other factors such as the seriousness of the disturbance. Hence, we finally selected the first difference as the clustering feature.

There are three major differences between the responses of the two loads. The first one is the dumping time constant during the disturbance. The second one is the peak amplitude after the disturbance and after the resection of the disturbance. The third one is the oscillation mode after the resection of the disturbance. The third one requires more data samples, while the second one is greatly influenced by many other factors such as the seriousness of the disturbance. Hence, we finally selected the
first difference as the clustering feature.

A. Theoretical Analysis

To theoretically prove that the dumping time constant during the disturbance is the main feature for clustering, eigenvalue analysis is used. Theoretical analysis is also compared with simulation results in the following section.

A widely used electric motor mathematic model is shown below:

\[
\begin{align*}
\frac{dE_d}{dt} &= -\frac{1}{T_E} E_d + (X - X^t) I_d - 2\pi f_0 (\omega_t - 1) E_d \\
\frac{dE_q}{dt} &= -\frac{1}{T_E} E_q - (X - X^t) I_q + 2\pi f_0 (\omega_t - 1) E_q \\
\frac{d\omega}{dt} &= \frac{1}{2H} (T_E - T_m)
\end{align*}
\]

(1)

where:

\[ f_0 \text{ is 50Hz, the power-frequency; } I_d \text{ is the current in Axis d; } I_q \text{ is the current in Axis q; } T_E \text{ is electromagnetic torque; } T_m \text{ is mechanical torque. And it has:} \]

\[ \begin{align*}
\omega_0 &= 1 - \varepsilon; X = X + X^t; X^t = X^t \times X_{\omega} + X_{\varepsilon}; \\
T_E &= E_d I_d + E_q I_q T_m = (A\omega_0 + B_0 + C) T_0; \\
A\omega_0^2 + B_0\omega_0 + C = \frac{2E_{max}}{MVABSE}; \\
I_d &= \frac{1}{R^2 + \omega^2} [R(U_d - E_d) + X(U_q - E_q)]; \\
I_q &= \frac{1}{R^2 + \omega^2} [R(U_q - E_q) + X(U_d - E_d)];
\end{align*} \]

The states are the voltage in Axis d \( E_d \), the voltage in Axis q \( E_q \), and the per-unit rotational speed of the motor \( \omega \).

Since the mathematic model is nonlinear, linearization is necessary before eigenvalue analysis. We use Taylor expansion for linearization. The first step is to get initial values.

Given the initial voltage \( U_{0d} \) and set \( B=0 \),

\[ X_{m} = X + X_m^t; X_m = X_m^t; X_\omega = X; X^t = X^t_m^t; \]

\[ A^t = R_m^2 + X^t_m^t; B^t = 2R_m^2; C^t = X^t_m^t + R_m^2; \]

\[ A = \frac{A^t P}{U_{C^t}} - R_m B = \frac{B^t P}{U_{C^t}} - X^t_m^t C^t = \frac{C^t P}{U_{C^t}} - R_m X_m^t; \]

\[ s_0 = \frac{R}{R_m^2} - \frac{B + \sqrt{B^2 - 4AC}}{2A}; \]

(3)

The initial value of \( \omega_0 \) is:

\[ \omega_0 = 1 \times s_0 \]

(4)

And given \( U_q = U_{0q} \), \( U_{0d} = 0 \),

According to equations,

\[ \frac{dE_{d}}{dt} = -\frac{1}{T_E} E_{d} + (X - X^t) I_{d} - 2\pi f_0 (\omega_t - 1) E_{d} \]

\[ \frac{dE_{q}}{dt} = -\frac{1}{T_E} E_{q} - (X - X^t) I_{q} + 2\pi f_0 (\omega_t - 1) E_{q} \]

\[ I_{d} = \frac{1}{R^2 + \omega^2} [R(U_d - E_d) + X(U_q - E_q)] \]

\[ I_{q} = \frac{1}{R^2 + \omega^2} [R(U_q - E_q) - X(U_d - E_d)] \]

The initial value of the three states is easy to figure out. Abbreviate the mathematic model as,

\[ x = f(x) + \hat{b} \]

where

\[ x = \begin{bmatrix} E_d & E_q & \omega \end{bmatrix}^T \]

(6)

f(x) =

\[
\begin{align*}
-\frac{1}{T_E}E_d + (X - X^t)I_d - 2\pi f_0 (\omega_t - 1)E_d \\
-\frac{1}{T_E}E_q - (X - X^t)I_q + 2\pi f_0 (\omega_t - 1)E_q \\
\frac{1}{2H}(T_E - T_m)
\end{align*}
\]

(8)

and abbreviate f(x) as:

\[ f(x) = \begin{bmatrix} A_{E_d} + A_{\omega_0}E_d + A_{\omega_0}\omega_0 E_d \\
A_{E_q} + A_{\omega_0}E_q + A_{\omega_0}\omega_0 E_q \\
A_{E_d}^2 + A_{\omega_0}E_d^2 + A_{\omega_0}\omega_0 E_d + A_{\omega_0}\omega_0 E_d \end{bmatrix} \]

(9)

According to the initial values that have been figure out,

\[ f(x) = \begin{bmatrix} A_{E_d} + A_{\omega_0}E_d + A_{\omega_0}\omega_0 E_d + A_{\omega_0}E_d \omega_0 \\
A_{E_q} + A_{\omega_0}E_q + A_{\omega_0}\omega_0 E_q + A_{\omega_0}E_q \omega_0 \\
A_{E_d}^2 + A_{\omega_0}E_d^2 + A_{\omega_0}\omega_0 E_d^2 + A_{\omega_0}\omega_0 E_d \omega_0 \\
+ (2A_{\omega_0}E_d + A_{\omega_0}\omega_0 E_d) \omega_0 \end{bmatrix} \]

(10)

Abbreviate above expression as:

\[ A + Cx \]

(11)

Equation 11 lays out the linear analytical description of the composite dynamic load model. We can get three eigenvalues of matrix \( C \). Two of the eigenvalues represent a second-order damped oscillation. Another eigenvalue represent first-order attenuation. The damping of the second-order damped oscillation is too large that the influence of this mode can be ignored. Hence by analyzing the first-order attenuation mode, we can get the dumping time constant under different value of \( X_t \), because the inverse of the only real eigenvalue is the dumping time constant.

B. Simulation Results

Theoretical analysis has proved the feasibility of clustering. Instead of identification of exact values, we can divide parameter values into intervals, and classify possible parameter intervals into several groups. Every group is defined as a class in the clustering algorithm. The whole method is divided into two steps. The first step is clustering samples with unknown secondary parameters and creating the training set, in which all of samples belong to a class. The second step is classifying new samples with unknown parameters using the training set. Hence we can roughly know the secondary parameters of the new samples.

The algorithm used in the first step is the nl-means algorithm and the k-nn algorithm [19] for the second step. Another important issue is feature extraction for clustering and classifying. After several simulation experiments, we selected system identification method, for its better accuracy. According to theoretical analysis, the first-order attenuation has more impact. During the disturbance, the load can be considered as the following system.

\[ \text{Voltage} \rightarrow \frac{K}{T + \tau} \rightarrow \text{Reactive Power} \]

Fig. 5 Approximate model during disturbance

By using the System Identification Toolbox of Matlab, voltage and reactive power are chosen as input and output. Using
this tool, the dumping time constant can be easily figured out from the PSD-BPA [21] simulation result. Here we link loads to Bus 12, select four groups of load parameters with different \( X_s \) and acquired their average eigenvalues. Simulation results are listed in Table I.

**TABLE I**

<table>
<thead>
<tr>
<th>Disturbance Location</th>
<th>( X_s )</th>
<th>( 0.095 )</th>
<th>( 0.145 )</th>
<th>( 0.195 )</th>
<th>( 0.295 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bus9 vs Bus7</td>
<td>23.3182</td>
<td>18.9883</td>
<td>15.9678</td>
<td>9.020386</td>
<td></td>
</tr>
<tr>
<td>Bus 12 vs Bus 13</td>
<td>25.2876</td>
<td>20.5300</td>
<td>17.1656</td>
<td>12.61925</td>
<td></td>
</tr>
<tr>
<td>Bus 9 vs Bus 14</td>
<td>24.09</td>
<td>19.3165</td>
<td>16.3102</td>
<td>11.52459</td>
<td></td>
</tr>
<tr>
<td>Bus 13 vs Bus 14</td>
<td>24.09</td>
<td>19.3165</td>
<td>16.3102</td>
<td>11.52459</td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>23.8849</td>
<td>19.4804</td>
<td>15.9111</td>
<td>10.95429</td>
<td></td>
</tr>
</tbody>
</table>

Compare simulation results with theoretical analysis, we can get the following results, as illustrated in Fig. 6. According to the curve, the eigenvalue reduces nonlinearly along with the reduction of \( X_s \). This result shows the relationship between reactive power response curve and inner model parameters, which is the basis of the clustering approach.

![Fig. 6 Comparison of simulation results with theoretic analysis](image)

As shown in Fig. 6, simulation result is consistent well with theoretical analysis. That is to say, the dumping time constant can be used to acquire the value of parameter \( X_s \) and can be used as the feature of our clustering algorithms.

IV. SYSTEM IMPLEMENTATION AND PERFORMANCE EVALUATION

A. System Implementation

One of important issues in the implementation of the nl-means algorithm is to choose the number of classes and the initials of clustering centers. The flow chart of our method is shown in Fig. 7.

We get 21 points \( X_s=0.095:0.01:……:0.295 \). Eigenvalues of these points are presented as \( X(i), i=1,2,……,21 \). In the end, we figure out that the number of class is eight and initial clustering centers of each class can be theoretical eigenvalues of load models with \( X_s \) of 0.095, 0.115, 0.135, 0.155, 0.185, 0.215, 0.255, 0.295.

Fig. 7 Flow chart of clustering initialization

We get 21 points \( X_s=0.095:0.01:……:0.295 \). Eigenvalues of these points are presented as \( X(i), i=1,2,……,21 \). In the end, we figure out that the number of class is eight and initial clustering centers of each class can be theoretical eigenvalues of load models with \( X_s \) of 0.095, 0.115, 0.135, 0.155, 0.185, 0.215, 0.255, 0.295.

For performance evaluation of the clustering approach we designed in Section III, we get necessary samples from PSD-BPA disturbance simulation results. After eliminating some sample points that are abnormal due to the limitation of calculation accuracy of PSD-BPA, we get a sample set with 160 sample points.

![Fig. 7 Flow chart of clustering initialization](image)

In particular, samples can only be classified into the right class or adjacent classes. This situation will reduce the impacts of wrong classifications. Such errors may result from many reasons. Firstly, the system is too small that the modes of the disturbance itself will greatly influence the performance of power loads. And as a linked system, the response of each load will obviously influence others. Hence the simulation result will be different from single power load. Such errors can be reduced by using bigger systems with more data. Secondly, the calculation accuracy of PSD-BPA and Matlab is limited.
We group 120 samples as a training set and 40 samples as a testing set randomly. The flow chart of performance evaluation is included in Fig. 8.

### B. Simulation Results

The classification results are included in Table III, which shows that the accuracy of this experiment is 100%. Even though this may not be the case for other situations, this classification result proved that using the proposed clustering approach and choosing the dumping time constant as the clustering feature is a promising way to roughly identify the parameter $X_r$ in the composite load model. Proper training set can alleviate the influence of noises, the mode of disturbance, and calculation errors.

![Flow chart of performance evaluation](image)

**Fig. 8 Flow chart of performance evaluation**

**TABLE III**

<table>
<thead>
<tr>
<th>Class</th>
<th>Training set</th>
<th>Error rate</th>
<th>Classification</th>
<th>Error rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
</tbody>
</table>

### V. Conclusions

The accuracy of load models has important impact on power system simulation. Online identification of dynamic parameters has been proposed for measurement based composite load modeling and existing work is focused on fast identification of dominant parameters.

In this work, it is pointed out some secondary dynamic parameters have also important impact to the model accuracy. We put forward a method to improve a two-step load modeling algorithm. We use clustering algorithms to identify parameter $X_r$ in the composite load model. Based on theoretical analysis and simulation results, we choose the reactive power dumping time constant during disturbance as the clustering feature. Clustering and classification results show the accuracy of our approach.

In the future, experiments should be carried out at a larger scale to further validate the approach. It is necessary to study the influence of the variation of other load parameters on clustering results. More accurate load models can be applied in many scenarios, e.g., energy management systems, power quality management, wide-area monitoring and control.

### REFERENCES


