

DOI: 10.19783/j.cnki.pspc.191302

## 基于多指标面板数据特征提取的同调机组分群方法

张智<sup>1</sup>, 陈艳波<sup>1</sup>, 宋新甫<sup>2</sup>, 刘建琴<sup>3</sup>, 李高望<sup>3</sup>, 曾鉴<sup>4</sup>, 陈浩<sup>1</sup>

(1. 华北电力大学电气与电子工程学院, 新能源电力系统国家重点实验室, 北京 102206;  
2. 国网新疆电力有限公司经济技术研究院, 新疆 乌鲁木齐 830000; 3. 国网经济技术研究院有限公司,  
北京 102200; 4. 国网四川省电力公司经济技术研究院, 四川 成都 610000)

**摘要:** 同调机群识别对电力系统的动态等值、主动解列控制具有重要的意义。针对已有同调机群识别方法存在的指标选取单一(仅依据发电机功角曲线特征)且特征提取不充分的问题, 提出一种基于多指标面板数据特征提取的同调机组分群方法。首先, 将反映机组同调性的功角、机端电压及转子角速度3个指标的量测数据形成面板数据。其次, 基于面板数据的指标维度和时间维度特征, 提取以上3个指标在时间序列上的“绝对值”特征、“波动”特征、“偏度”特征、“峰度”特征以及“趋势”特征。接着借鉴层次分析模型, 利用AHP-熵权法求取每个指标及其各特征量权重, 进而计算机组间的加权距离矩阵, 然后利用系统聚类法实现同调机组分群。最后以EPRI-36节点系统和华北电网为例, 验证了所提方法的有效性。

**关键词:** 同调机群; 面板数据; 多指标; 特征提取; 层次分析; 系统聚类

### Coherent clustering method based on feature extraction for multi-index panel data

ZHANG Zhi<sup>1</sup>, CHEN Yanbo<sup>1</sup>, SONG Xinfu<sup>2</sup>, LIU Jianqin<sup>3</sup>, LI Gaowang<sup>3</sup>, ZENG Jian<sup>4</sup>, CHEN Hao<sup>1</sup>

(1. State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources, School of Electrical and Electronic Engineering, North China Electric Power University, Beijing 102206, China; 2. Power Economic Research Institute, State Grid Xinjiang Electric Power Co., Ltd., Urumqi 830000, China; 3. State Grid Power Economic Research Institute Co., Ltd., Beijing 102200, China; 4. Power Economic Research Institute, State Grid Sichuan Electric Power Corporation, Chengdu 610000, China)

**Abstract:** Coherent cluster identification plays an important role in the dynamic equivalent and active splitting control of power systems. At present, coherent clustering mainly depends on the generator's power angle curve features. There are problems with single selection indicators and insufficient feature extraction. For these problems, this paper proposes a coherent clustering method based on multi-index panel data feature extraction. First, it selects three indicators that can reflect the coherence of power angle increment, terminal voltage and rotor angular velocity of the generator to form panel data. Secondly, based on the characteristics of the panel data with both the cross-sectional and the time-series dimensions, it extracts the "absolute value", "fluctuation", "skew degree", "kurtosis" and "trend" features of the three indices in the time series. Then it builds a hierarchical analysis cluster model. It uses the AHP-entropy weight method to calculate the weight of each index and each feature, and calculates the weighted distance matrix between generator groups, to achieve the final coherent group by system clustering. Finally, the EPRI-36 node system and North China Power Grid are taken as examples to verify the effectiveness of the proposed method.

This work is supported by Science and Technology Project of the Headquarter of State Grid Corporation of China (No. SGXJJY00GHJS1800003): "Study on Safety Quality Evaluation and Grid Strengthening Scheme Optimization of Weak Delivery Network with UHV/EHVDC".

**Key words:** coherency identification; panel data; multiple indicator; feature extraction; analytic hierarchy; system clustering

## 0 引言

随着“西电东送, 南北互联”的实施, 全国电

网逐渐向大区域互联电网方向发展, 导致电力系统的安全稳定分析愈发困难<sup>[1-2]</sup>。为提高计算效率, 一般需要在电磁暂态分析中减小系统的规模, 动态等值成为解决此问题的有效途径<sup>[3]</sup>。同调机组分群是实现电力系统动态等值的先决条件。此外, 同调机群的准确识别对于快速找出解列断面亦具有重要作用<sup>[4-6]</sup>。

基金项目: 国家电网有限公司总部科技项目资助(SGXJJY00GHJS1800003)“含特/超高压直流接入的弱送端电网安全品质评价及网架加强方案优化研究”

为有效识别同调机群,文献[7]利用系统中各机组间的电气和机械距离,提出了一种综合的机电距离分群方法。文献[8]提出一种基于状态空间的同调机组识别模型。文献[9]通过相关因子进行同调机组分群。文献[10]利用慢同调法识别同调机群。以上方法原理相对简单,但需对系统进行线性化,并且计算效率与计算规模之间存在一定矛盾。

研究表明,发生扰动后系统中的发电机组有同调现象,即某些机组的功角轨迹具有一致性或相似性,故基于扰动轨迹的相似性,可对发电机组进行同调分群<sup>[11-14]</sup>。文献[15]基于最小二乘支持向量机方法进行在线同调机组分群。文献[16]基于广域量测信号的主成分,采用聚类分析方法进行同调分群。文献[17]通过对广域量测信号进行Koopman模式分解进行同调机群辨识。以上研究工作表明基于广域量测信号可以实现同调机群辨识;但在辨识中需要处理大量数据,其辨识效率依赖于对海量数据中关键信息的快速提取。为此,一些研究中通过降维方法对功角信号进行处理获得特征量,如拉普拉斯变换法<sup>[18]</sup>、小波分析法<sup>[19]</sup>、希尔伯特-黄变换法<sup>[20]</sup>及非负矩阵分解法<sup>[21]</sup>等。

以上研究做了很有意义的工作,但在同调分群过程中仅考虑功角特征,存在指标选取单一且特征提取不充分的问题,由此可能造成分群不准确的现象。考虑到基于面板数据的聚类方法可综合考虑多个指标<sup>[22]</sup>,若将其用于同调机群的识别将有望取得良好的聚类效果。为此,本文提出一种基于多指标面板数据特征提取的同调机组分群方法,其主要步骤如下:(1)基于功角曲线、机端电压以及转子角速度3个指标的量测数据形成面板数据;(2)在时间序列上提取3个指标的“绝对值”特征、“波动”特征、“偏度”特征和“峰度”特征以及“趋势”特征;(3)建立层次分析模型,根据AHP-熵权法计算每个指标及其各特征量的权重,进而计算机组间的加权距离矩阵,利用系统聚类实现同调机组分群。最后以EPRI-36节点系统和华北电网为例,测试所提出方法的分群识别效果。

## 1 同调判别理论及指标选取

### 1.1 同调判别

发电机同调是指系统受扰后,在研究时间之内,若其中两台发电机的母线电压复向量之比为常数,则认为这两台机组是同调的<sup>[13]</sup>,即

$$\frac{\dot{V}_i(t)}{\dot{V}_j(t)} = \frac{V_i(t)}{V_j(t)} e^{j[\delta_i(t) - \delta_j(t)]} = \text{const} \quad (1)$$

式中: $\dot{V}_i(t)$ 、 $\dot{V}_j(t)$ 分别为发电机*i*、*j*母线上的复电压相量; $V_i(t)$ 、 $V_j(t)$ 为母线电压幅值; $\delta_i(t)$ 、 $\delta_j(t)$ 为电压相角; $t$ 为时间序列上的时间点。

### 1.2 指标选取

现阶段的同调分群方法,一般仅根据发电机功角曲线的特征进行聚类,忽略了电压幅值等因素,指标选取单一,由此可导致分群结果不准确。为提高分群精度,可选取多个反映发电机同调性的指标进行分群。研究表明,反映同调性的指标主要包括3类:静态工况数据、系统模型数据以及瞬间动态数据,且这3类数据包含的同调信息量依次增加。本文基于瞬间动态数据选取3个指标量。

#### (1) 发电机功角增量

发电机功角摇摆曲线直接反映系统的同调特性,可作为发电机组同调判别的依据。定义第*i*台机组的功角增量数据为

$$x_i^{(1)}(t) = \Delta\delta_i(t) \quad (2)$$

式中: $i=1,2,\dots,N$ , $N$ 表示系统中机组总数; $t=0,1,\dots,T$ , $T$ 表示时间序列长度。

#### (2) 发电机的机端电压

电压变化是系统故障后的直接时域响应,传统的同调分群方法认为发电机的机端电压幅值是恒定的,故将其忽略。文献[23]利用扰动后发电机端节点电压幅值实现了暂态功角稳定性的在线评估。在本文中,将机端电压作为同调分群的指标,定义第*i*台发电机的机端电压数据为

$$x_i^{(2)}(t) = V_i(t) \quad (3)$$

式中, $V_i(t)$ 为机端电压幅值的标么值。

#### (3) 发电机转子角速度

考虑转子角速度的方法已在紧急扩展等面积准则法中得到运用。系统受到扰动后,发电机电磁功率快速发生变化,原动机机械功率与电磁功率不平衡,发电机的功角因转速变化而产生相对变化。当电力系统发生严重故障时,发电机的电磁功率变化明显,发电机转子角速度变化更为明显。因此利用发电机转子角速度分群将进一步提高分群精度。为此,定义第*i*台发电机转子角速度数据为

$$x_i^{(3)}(t) = \omega_i(t) \quad (4)$$

式中, $\omega_i$ 为发电机*i*的转子角速度标么值。

## 2 基于多指标面板数据特征提取的同调机组分群方法

### 2.1 面板数据形成

反映发电机同调性的3个指标(功角曲线、机端电压以及转子角速度)具有面板数据的截面维度和

时间维度特征,采用 min-max 标准化方法将量测数据按指标分别做归一化处理,其每个数据可用 3 个子标量表示,如  $x_i^{(k)}(t)$ ,其中  $i$  代表发电机编号, $k$  代表指标, $t$  代表数据的时间点。 $N$  台机组的三维样本总体  $\mathbf{X}$  可表示为  $[\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N]$ 。其中第  $i$  台机组的二维面板数据  $\mathbf{X}_i$  可表示为

$$\mathbf{X}_i = \begin{bmatrix} x_i^{(1)}(1) & x_i^{(1)}(2) & \dots & x_i^{(1)}(T) \\ x_i^{(2)}(1) & x_i^{(2)}(2) & \dots & x_i^{(2)}(T) \\ x_i^{(3)}(1) & x_i^{(3)}(2) & \dots & x_i^{(3)}(T) \end{bmatrix} = \begin{bmatrix} \mathbf{X}_i^{(1)} \\ \mathbf{X}_i^{(2)} \\ \mathbf{X}_i^{(3)} \end{bmatrix} \quad (5)$$

式中:  $\mathbf{X}_i^{(k)} = [x_i^{(k)}(t)]$ ,  $t=1,2,3,\dots,T$ ;  $k=1,2,3$  代表机组  $i$  的第  $k$  个指标组成的时间序列。

## 2.2 基于面板数据的结构差异度聚类指标定义

当电力系统受到扰动后,发电机的多个指标量在时间序列上呈现出动态变化。多指标面板数据反映多个指标在一段时期内的变化趋势、波动程度等动态演化规律。面板数据在截面维度上,体现指标值的绝对水平;在时间维度上,体现指标值在时间序列上的波动性。为了较全面、综合地考虑系统故障后发电机各指标曲线之间的特征,本节基于发电机面板数据定义了绝对水平、趋势、波动以及偏度和峰度 5 个特征量<sup>[22]</sup>。

定义 1: 发电机组  $i$  的第  $k$  个指标在时间序列上的“均值”特征(Average Value Feature),记为  $\text{AVF}(\mathbf{X}_i^{(k)})$ 。

$$\text{AVF}(\mathbf{X}_i^{(k)}) = \bar{x}_i^{(k)} = \frac{1}{T} \sum_{t=1}^T x_i^{(k)}(t) \quad (6)$$

式中,  $\text{AVF}(\mathbf{X}_i^{(k)})$  是机组  $i$  的第  $k$  个指标在时间序列上的均值,均值水平是发电机特征轨迹的相对距离的度量,反映两条轨迹在空间上的接近程度。

定义 2: 发电机组  $i$  的第  $k$  个指标在时间序列上的“波动”特征(Variance Feature),记为  $\text{VF}(\mathbf{X}_i^{(k)})$ 。

$$\text{VF}(\mathbf{X}_i^{(k)}) = \sqrt{\frac{\sum_{t=1}^T (x_i^{(k)}(t) - \bar{x}_i^{(k)})^2}{T-1}} \quad (7)$$

式中,  $\text{VF}(\mathbf{X}_i^{(k)})$  表示机组  $i$  的第  $k$  个指标在时间序列上的波动程度,其反映了发电机在受到扰动后特征量的波动情况,若两台机组是同调的,其特征量在时间序列上的“波动特征”应该是接近的。

定义 3: 发电机组  $i$  的第  $k$  个指标在时间序列上的“偏度”特征(Skewness Coefficient Feature),记为  $\text{SCF}(\mathbf{X}_i^{(k)})$ 。

$$\text{SCF}(\mathbf{X}_i^{(k)}) = \frac{\sum_{t=1}^T (x_i^{(k)}(t) - \bar{x}_i^{(k)})^3}{T(\sigma_i^{(k)})^3} \quad (8)$$

式中:  $\sigma_i^{(k)} = \sqrt{\frac{\sum_{t=1}^T (x_i^{(k)}(t) - \bar{x}_i^{(k)})^2}{T-1}}$  表示机组  $i$  的第  $k$

个指标在时间序列的标准差;  $\text{SCF}(\mathbf{X}_i^{(k)})$  衡量机组  $i$  的第  $k$  个指标在时间序列上是否对称,反映了样本的分布特征。当系统发生故障后,具有同调性的机组的各指标曲线的振荡是相似的,故各指标的数据分布也是接近的。所以当两机组同调时,其偏度指标值是接近的。

定义 4: 发电机组  $i$  的第  $k$  个指标在时间序列上的“峰度”特征(Kurtosis Coefficient Feature),记为  $\text{KCF}(\mathbf{X}_i^{(k)})$ 。

$$\text{KCF}(\mathbf{X}_i^{(k)}) = \frac{\sum_{t=1}^T (x_i^{(k)}(t) - \bar{x}_i^{(k)})^4}{T(\sigma_i^{(k)})^4} - 3 \quad (9)$$

式中,  $\text{KCF}(\mathbf{X}_i^{(k)})$  衡量机组  $i$  的第  $k$  个指标在时间序列上的集中程度。当系统发生故障,受故障影响越大的机组其振荡幅度越高,其峰度值越大。当两台机组同调时,其峰度值相似。

定义 5: 发电机  $i$  的第  $k$  个指标在时间序列上的“趋势”特征(Tread Feature),记为  $\text{TF}(\mathbf{X}_i^{(k)})$ 。

$$\text{TF}(\mathbf{X}_i^{(k)}) = \frac{\sum_{t=1}^T (x_i^{(k)}(t) - \bar{x}_i^{(k)})(t - T/2)}{\sum_{t=1}^T (t - T/2)^2} \quad (10)$$

式中,  $\text{TF}(\mathbf{X}_i^{(k)})$  衡量机组  $i$  的第  $k$  个指标在时间序列上的变化情况,若两台机组的第  $k$  个指标随时间呈同向变化,变化越协调,两者越相似。

将提取的特征量分别归一化,公式为

$$\text{AVF}(M_i^{(k)}) = \frac{\text{AVF}(\mathbf{X}_i^{(k)})}{\max(\text{AVF}(\mathbf{X}^{(k)}))} \quad (11)$$

式中,  $\text{AVF}(M_i^{(k)})$  为机组  $i$  的第  $k$  个指标的归一化均值特征量。

同法可得其他 4 个指标的归一化特征量  $\text{VF}(M_i^{(k)})$ ,  $\text{SCF}(M_i^{(k)})$ ,  $\text{KCF}(M_i^{(k)})$  及  $\text{TF}(M_i^{(k)})$ 。进而可得发电机组  $i$  的特征矩阵,为一个  $3 \times 5$  的矩阵,记为  $\mathbf{M}_i$ 。

$$\mathbf{M}_i = [\mathbf{M}_i^{(1)}, \mathbf{M}_i^{(2)}, \mathbf{M}_i^{(3)}] = \begin{bmatrix} \text{AVF}(M_i^{(1)}) & \text{VF}(M_i^{(1)}) & \text{SCF}(M_i^{(1)}) & \text{KCF}(M_i^{(1)}) & \text{TF}(M_i^{(1)}) \\ \text{AVF}(M_i^{(2)}) & \text{VF}(M_i^{(2)}) & \text{SCF}(M_i^{(2)}) & \text{KCF}(M_i^{(2)}) & \text{TF}(M_i^{(2)}) \\ \text{AVF}(M_i^{(3)}) & \text{VF}(M_i^{(3)}) & \text{SCF}(M_i^{(3)}) & \text{KCF}(M_i^{(3)}) & \text{TF}(M_i^{(3)}) \end{bmatrix} \quad (12)$$

## 2.3 基于层次分析法的同调机组分群

### 2.3.1 基于 AHP-熵权法的权重确定

层次分析法(Analytic Hierarchy Process, AHP)<sup>[24]</sup>是一种主观权重确定方法,其优点是可发挥专家经

验,缺点是过于主观;熵权法是一种客观的权重确定方法,其优点是可有效提取数据中包含的信息量,缺点是受离散极值影响较大。为充分发挥两种方法的优点,可构建主观与客观相结合的权重确定方法<sup>[25]</sup>。建立每台机组的层次结构,如图1所示。将反映同调性的3个指标定为A层,将每个指标在时间序列上提取的特征定为B层。

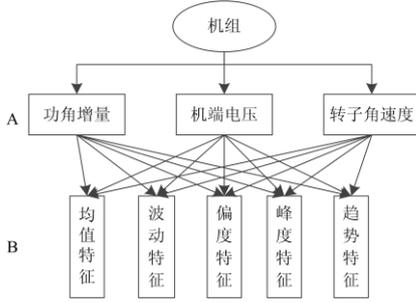


图1 层次结构模型

Fig. 1 Hierarchy model

### 1) B层特征值权重计算

#### (1) AHP法确定权重

基于AHP法确定B层特征值权重的步骤如下。

步骤1:根据专家打分法,邀请多位专家,对B层两两特征进行比较,建立相对重要性矩阵。求该矩阵的最大特征值及特征向量,通过一致性判断,得出B层特征权重。设有 $n$ 位专家,计算出各位专家的特征权重向量 $c_i$ , $i=1,L,n$ 。

步骤2:基于步骤1,将各位专家的特征权重向量 $c_i$ 组成权重矩阵 $C$ ,如式(13)所示。

$$C = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & \dots & c_{15} \\ c_{21} & c_{22} & \dots & c_{25} \\ \vdots & \vdots & \dots & \vdots \\ c_{n1} & c_{n2} & \dots & c_{n5} \end{bmatrix} \quad (13)$$

式中, $c_{ij}$ 代表第 $i$ 个专家对第 $j$ 个特征所赋的权重。

步骤3:基于权重矩阵 $C$ ,根据式(14)计算相关系数矩阵。

$$v_{ij} = 1 - \sqrt{\frac{1}{5} \sum_{l=1}^5 (c_{il} - c_{jl})^2}, \quad i, j = 1, 2, L, n \quad (14)$$

由 $v_{ij}$ 组成指标权重向量的相关系数矩阵 $V$ 。

步骤4:根据式(15),按比例将偏离程度较大的专家权重剔除。

$$v_i = \sum_{j=1}^n v_{ij} \quad (15)$$

式中, $v_i$ 为专家 $i$ 权重意见与其余专家权重意见相似程度之和, $v_i$ 越小表示专家 $i$ 评定的权重偏离程度越大。

步骤5:求筛选后权重矩阵列向量的平均值 $\gamma_A = [\gamma_A(1) \ \gamma_A(2) \ \gamma_A(3) \ \gamma_A(4) \ \gamma_A(5)]$ ,即为B层各特征值的权重,且 $\sum_{l=1}^5 \gamma_A(l) = 1$ 。

#### (2) 熵权法确定权重

由于3个指标的数据在时间序列上各异,故利用熵权法需分别确定每个指标的特征权重。

步骤1:计算第 $i$ 台机组的第 $k$ 个指标对应的第 $l$ 个特征量相对于其他机组的比重。

$$q_i^{(k)}(l) = \frac{M_i^{(k)}(l)}{\sum_{i=1}^n M_i^{(k)}(l)}, \quad l = 1, 2, L, 5 \quad (16)$$

步骤2:计算指标 $k$ 的特征值 $l$ 的熵值。

$$e^{(k)}(l) = -a \sum_{i=1}^N q_i^{(k)}(l) \log(q_i^{(k)}(l)), \quad l = 1, 2, L, 5 \quad (17)$$

式中, $e^{(k)}(l)$ 为特征熵值; $a > 0, e^{(k)}(l) > 0$ 。一般取 $a = 1/\log n$ 。

步骤3:确定特征权重。

由于 $e^{(k)}(l)$ 是逆向指标,对熵值进行正向化和归一化处理,得到特征值 $l$ 的权重系数为

$$\gamma_E^{(k)}(l) = \frac{1 - e^{(k)}(l)}{\sum_{l=1}^5 (1 - e^{(k)}(l))}, \quad l = 1, 2, L, 5 \quad (18)$$

按照上述方法计算各指标的特征权重,即得到基于熵权法的特征权重矩阵为

$$\gamma_E^{(k)} = \begin{bmatrix} \gamma_E^{(1)}(1) & \gamma_E^{(1)}(2) & \gamma_E^{(1)}(3) & \gamma_E^{(1)}(4) & \gamma_E^{(1)}(5) \\ \gamma_E^{(2)}(1) & \gamma_E^{(2)}(2) & \gamma_E^{(2)}(3) & \gamma_E^{(2)}(4) & \gamma_E^{(2)}(5) \\ \gamma_E^{(3)}(1) & \gamma_E^{(3)}(2) & \gamma_E^{(3)}(3) & \gamma_E^{(3)}(4) & \gamma_E^{(3)}(5) \end{bmatrix} \quad (19)$$

#### (3) 熵权法对AHP法的修正

分别计算出AHP法和熵权法的特征权重值后,引入熵值变量,将AHP法和熵权法结合,最终权重如式(20)所示。

$$\gamma^{(k)}(l) = \gamma_A(l) e^{(k)}(l) + \gamma_E^{(k)}(l) (1 - e^{(k)}(l)) \quad (20)$$

式中: $l = 1, L, 5$ ;  $\sum_{l=1}^5 \gamma^{(k)}(l) = 1$ ;  $\gamma_A(l)$ 和 $\gamma_E^{(k)}(l)$ 分别为AHP法和熵权法所确定的权重。

基于式(20),可得基于AHP-熵权法的特征权重矩阵为

$$\gamma^{(k)} = \begin{bmatrix} \gamma^{(1)}(1) & \gamma^{(1)}(2) & \gamma^{(1)}(3) & \gamma^{(1)}(4) & \gamma^{(1)}(5) \\ \gamma^{(2)}(1) & \gamma^{(2)}(2) & \gamma^{(2)}(3) & \gamma^{(2)}(4) & \gamma^{(2)}(5) \\ \gamma^{(3)}(1) & \gamma^{(3)}(2) & \gamma^{(3)}(3) & \gamma^{(3)}(4) & \gamma^{(3)}(5) \end{bmatrix} \quad (21)$$

#### 2) A层指标权重计算

将各机组的功角增量、机端电压以及转子角速度的特征值加权求和,得到各机组3个指标的加权数据;然后利用AHP-熵权法计算3个指标的权重,设最终得到的权重向量为 $w = [w_1, w_2, w_3]$ 。

### 2.3.2 定义加权距离矩阵

基于欧氏距离, 定义系统中任两台发电机组  $i$  和  $j$  第  $k$  个指标的改进距离函数为

$$D_{ij}^{(k)} = \sqrt{\sum_{l=1}^5 \gamma^{(k)}(l) (M_i^{(k)}(l) - M_j^{(k)}(l))^2} \quad (22)$$

式中,  $M_i^{(k)}(l)$  表示机组  $i$  的第  $k$  个指标对应的第  $l$  个特征量 ( $i=1,2,L,N; k=1,2,L,3; l=1,2,L,5$ )。

根据式(22), 计算  $N$  台机组各指标距离矩阵  $\mathbf{D}^{(k)}$ , 矩阵维数为  $N \times N$ 。进一步将功角增量距离矩阵  $\mathbf{D}^{(1)}$ 、机端电压距离矩阵  $\mathbf{D}^{(2)}$  以及转子角速度距离矩阵  $\mathbf{D}^{(3)}$  加权求和得到加权距离矩阵:

$$\mathbf{D}_{N \times N} = w_1 \times \mathbf{D}^{(1)} + w_2 \times \mathbf{D}^{(2)} + w_3 \times \mathbf{D}^{(3)} \quad (23)$$

Ward 法把方差分析的思想用于系统聚类, 其优点是可有效衡量类内及类与类间的距离, 具有较好的聚类效果。基于得到的加权距离矩阵, 通过 Ward 法实现对同调机组的分群。

### 2.4 最优聚类簇数确定

为了确定最佳聚类簇数, 引入 DB 指标, DB 指标通过描述样本的类内散度与各聚类中心的间距, DB 越小表示类与类之间的相似度越低, 从而对应越佳的聚类结果。其定义为

$$DB(K) = \frac{1}{K} \sum_{i=1}^K \max_{j=1-K, j \neq i} \left( \frac{W_i + W_j}{L_{ij}} \right) \quad (24)$$

式中:  $K$  为分群数;  $W_i$  表示类  $L_i$  中的机组面板数据到其聚类中心的平均加权距离, 聚类中心取类  $L_i$  中所有机组面板数据平均值;  $W_j$  表示类  $L_j$  中的所有机组到类  $L_j$  中心的平均加权距离;  $L_{ij}$  表示类  $L_i$  和  $L_j$  中心之间的距离。

### 2.5 基于多指标面板数据特征提取的同调机组分群流程

基于广域量测系统提供的功角、机端电压及转子角速度数据, 以上提出的基于多指标面板数据特征提取的同调机组分群方法可概括为图 2。

## 3 算例分析

为了验证本文提出的基于多指标面板数据特征提取的同调机组分群方法的有效性, 利用 EPRI-36 节点标准系统和华北电网进行测试。仿真采用 PSASP, 算法采用 Matlab 编程实现。

### 3.1 EPRI-36 节点系统算例

EPRI-36 节点系统包含 8 台发电机。测试中, 0 时刻在母线 bus-9 处设置三相短路故障, 0.1 s 后切除故障; 设置积分步长为 0.01 s, 量测时间从故障

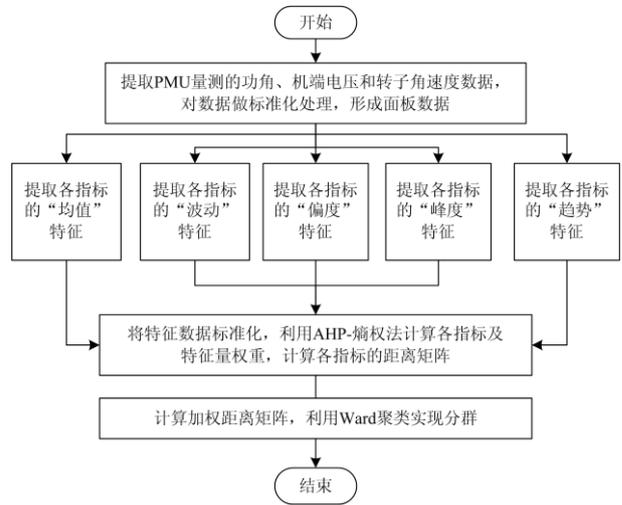


图 2 同调分群流程图

Fig. 2 Flow chart of coherence clustering

发生时刻起到 1 s 结束, 记录 8 台机组的功角、机端电压以及转子角速度数据, 量测数据曲线如图 3 所示。对发电机功角、机端电压以及转子角速度数据进行采样, 总时间为 1 s, 采样间隔为 0.01 s, 对数据进行标准化处理, 形成面板数据。

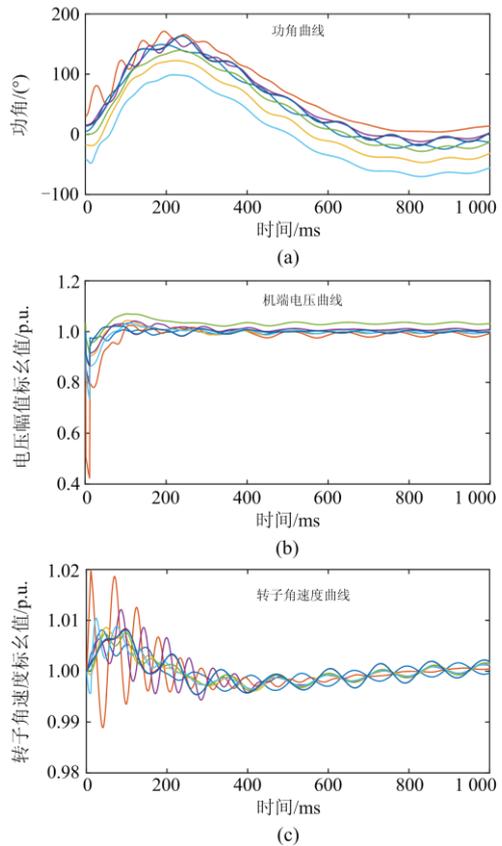


图 3 EPRI-36 节点系统的量测数据曲线

Fig. 3 Measurement data curve of EPRI-36 bus system

本节对基于多指标面板数据的同调机组分群方法进行分析。根据图 2，首先提取 8 台机组各指标的特征量；然后邀请 5 位专家进行打分，然后根据 AHP-熵权法得到特征权重矩阵。

基于本文所提方法进行同调机组分群，分群结果如图 4 的聚类树所示，其中，横坐标表示发电机编号，纵坐标表示各同调机群间的 SSD 距离。

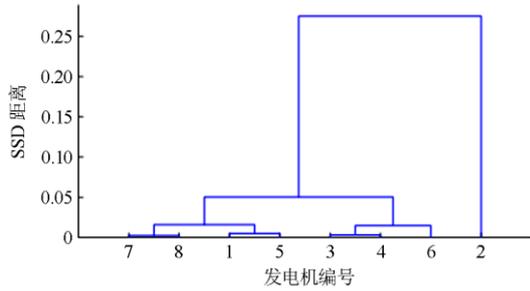


图 4 EPRI-36 节点系统的聚类树型

Fig. 4 Diagram of clustering-tree of EPRI-36 bus system

为了说明本文所提同调机组分群方法的有效性，将本文方法与基于功角轨迹的 Max-Min、K-Means 方法，以及基于小波变换(WT)-模糊聚类(FCM)、主成分分析(PCA)的同调机组分群方法进行对比。以上方法都是基于功角数据进行分群。因此，先将上述基于功角曲线的分群方法进行对比，根据式(11)计算上述分群方法的 DB 指标，判断各种方法的分群效果，如表 1 所示。

表 1 不同分群方法的 DB 指标

Table 1 DB indicator value of different clustering methods

| 分群数 $k$ | 2    | 3    | 4    | 5    | 6    | 7    |
|---------|------|------|------|------|------|------|
| Max-Min | 1.12 | 1.03 | 1.12 | 1.07 | 1.04 | 1.08 |
| K-Means | 1.14 | 0.92 | 1.12 | 1.07 | 1.06 | 1.08 |
| WT-FCM  | 1.12 | 0.78 | 1.11 | 1.08 | 1.08 | 1.08 |
| PCA     | 1.12 | 0.78 | 1.11 | 1.05 | 1.08 | 1.08 |

表 1 通过 DB 指标比较 4 种分群方法的效果，当  $K=3$  时，各种分群方法的分群结果为最优。当  $K=3$  时，由 DB 指标可看出，Max-Min 分群方法的效果最差。相较于基于轨迹的同调分群方法，基于 XW-FCM、PCA 特征提取的同调分群结果更加准确。当  $K=3$  时对各种分群方法的结果做进一步分析，分群结果如表 2 所示。

表 2  $K=3$  时各分群方法的分群结果

Table 2 Clustering results for each grouping method at  $K=3$

| 分群方法    | 分群结果            |
|---------|-----------------|
| Max-Min | 1 4/2/3 5 6 7 8 |
| K-Means | 1/2/3 4 5 6 7 8 |
| WT-FCM  | 2/3 4/1 5 6 7 8 |
| PCA     | 2/3 4/1 5 6 7 8 |

为了将本文提出的基于多指标面板数据的同调识别方法与基于单指标的分群方法进行对比，首先通过 DB 指标得到基于本文方法得到的最佳聚类簇数，如表 3 所示。当  $K=3$  时，DB 指标最小，因此 3 为最佳分群数。

表 3 不同分群数的 DB 指标

Table 3 DB indicator value of different number clusters

| $K$   | 2    | 3    | 4    | 5    | 6    | 7    |
|-------|------|------|------|------|------|------|
| DB 指标 | 1.25 | 0.81 | 0.89 | 0.94 | 1.07 | 1.08 |

当  $K=3$  时，对比本文方法和基于 PCA 方法的分群结果。基于 PCA 的分群方法和本文分群方法的不同主要在于 6 号机组。这两种方法分别将 6 号机组和 1、5、7、8 号机组及 3、4 号机组分为一群。

图 5 给出了 PCA 方法和本文方法的功角曲线对比。从图 5(a)和图 5(d)可以看出，相较于 1、5 号机组，6 号机组与 3、4 号机组的功角曲线相似度更高。但从总体上来看，系统故障后，各台发电机的功角曲线的总体变化趋势是一致的。

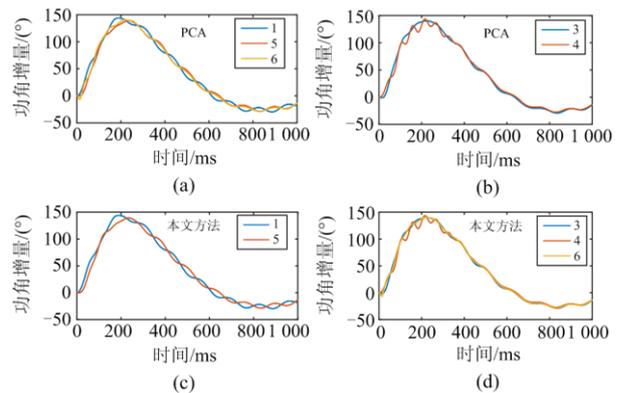


图 5 功角曲线对比

Fig. 5 Comparison of power angle curve

由于各机组的功角曲线的总体趋势相似，因此仅通过功角曲线的特征难以直观判断各机组间的同调性。图 6 给出了机端电压曲线对比。从图 6(a)、图 6(d)可看出，相对于 1 号和 5 号机组，6 号与 3、4 号机组的机端电压曲线在时间序列上有一致的变化，与本文方法分群结果一致。通过分析可以得出，本文所提方法不仅保证了同调机群功角曲线变化的一致性，同时保证了机端电压曲线变化的一致性。

系统受到大扰动后，发电机电磁功率快速发生变化，原动机机械功率与电磁功率不平衡，发电机的功角因转速变化而产生相对变化。电力系统发生严重故障时，发电机转子角速度变化更为明显。图 7 给出了转子角速度曲线的对比。

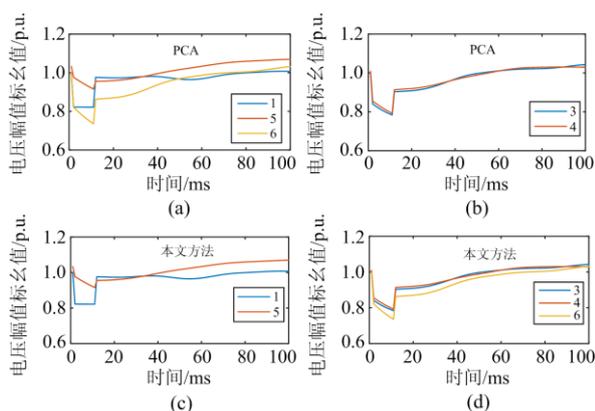


图 6 机端电压曲线对比

Fig. 6 Comparison of terminal voltage curve

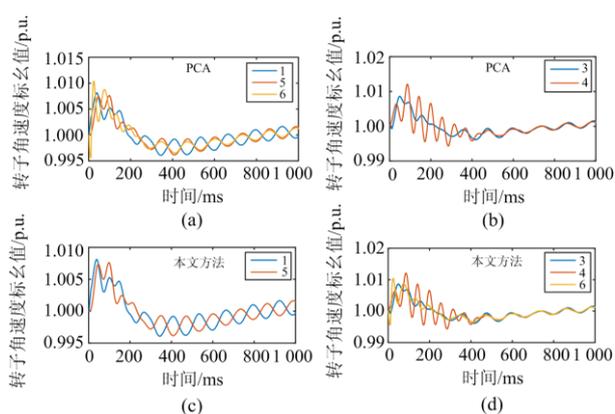


图 7 转子角速度曲线对比

Fig. 7 Comparison of rotor angular speed

从图 7(a)、图 7(c)可看出, 6 号机组与 1 号和 5 号机组的角速度变化曲线的相似度较低; 1 号机组和 5 号机组在故障发生后短时间内的变化一致。从图 7(b)、图 7(d)中可以看出, 3、4、6 号机组的转子角速度曲线在故障切除后变化逐渐趋于一致。综合以上对功角曲线、机端电压曲线以及转子角速度曲线的综合对比分析可判断出, 6 号机组与 3、4 号机组有较高的同调性, 与本文同调识别方法的分群结果一致。通过对多种基于单指标同调识别方法和本文提出的基于多指标面板数据特征提取的同调分群方法在 EPRI-36 节点系统中分群结果的对比分析, 可看出本文提出的方法具有更好的同调分群效果。

### 3.2 华北电网算例

进一步选取华北电网验证本文所提方法的有效性。华北电网主要包括北京(简称为 B)、天津(简称为 T)、河北(简称为 H)、山西(简称为 S)和山东(简称为 L)。由于华北电网规模较大, 机组较多, 故挑选 70 台主要发电机组作为研究对象。表 4 给出了主要的研究对象。

表 4 研究对象

Table 4 Study objects

| 区域 | 机组                                                                                                 |
|----|----------------------------------------------------------------------------------------------------|
| 北京 | BCR, BGE, BLD, BSH, BTR, BXG, BXS, BYD, BZC                                                        |
| 天津 | TBJ, TDG, TGJ, TJR, TJS, TJC, TLR, TNJ, TPS, TWR                                                   |
| 河北 | HBX, HCD, HCZ, HDC, HDJ, HHD, HHS, HLS, HMT, HQR, HAS, HXB, HXT                                    |
| 山西 | SAP, SCY, SDN, SDR, SFZ, SFL, SGY, SGY, SHG, SHG, SHZ, SJY, SLC, SSQ, STM, SWP, SWX, SXN, SYG, SZQ |
| 山东 | LBB, LDH, LHZ, LHD, LLW, LJN, LLZ, LLY, LLB, LQD, LRZ, LTA, LWH, LWF, LYD, LLC, LLR                |

故障设置在锡盟-山东 1 000 kV 特高压交流输电通道的津津济 K1-鲁济京 K1 线路上, 设置三相短路故障, 0.2 s 后切除故障。从故障发生时刻起到 0.5 s 结束, 记录 70 台机组的功角、电压以及角速度数据, 如图 8 所示。

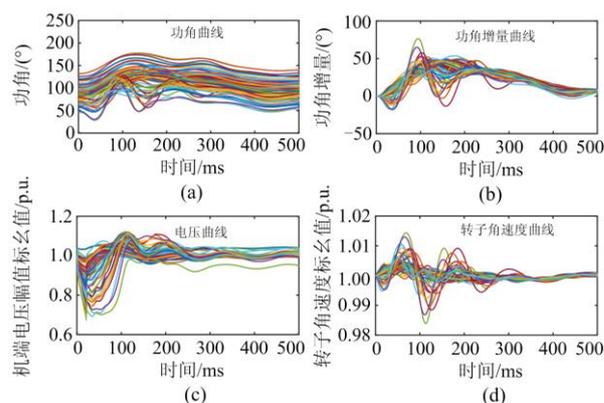


图 8 华北电网机组故障后的量测数据

Fig. 8 Measurement data of North China Power Grid generators after fault

基于故障后量测的功角、电压幅值以及转子角速度数据, 然后利用本文所提方法进行同调机组分群, 得到聚类树如图 9 所示。

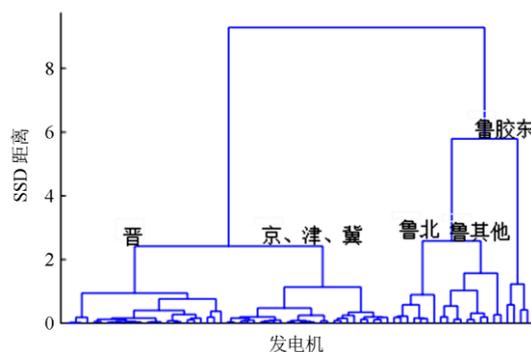


图 9 华北电网的同调机群聚类树

Fig. 9 Cluster tree of coherence generators in North China Power Grid

从图 9 可看出, 1 000 kV 特高压交流线路津济 K1-鲁济京 K1 短时间三相短路后, 华北电网的发电机根据地域被划分为不同的同调机群。山西电网的机组具有一致的同调性; 京、津、冀电网的大部分机组具有较高的同调性; 山东电网的机群被划分为 3 群, 从聚类树图可看出鲁胶东地区与山东其他地区机组呈现出较高的非同调性。当  $K=5$  时, DB 指标为 0.73, 其值最小, 故最佳分群数为 5, 分群结果如表 5 所示。

表 5 华北电网机组分群结果

| 分群数 | 分群结果                                                                                                                        |
|-----|-----------------------------------------------------------------------------------------------------------------------------|
| 1   | SAP, SCY, SDN, SDR, SFZ, SFL, SGY, SHG, SHZ, SJY, SLC, SSQ, STM, SWP, SWX, SXN, SYG, SZQ, BLD, BSH, BZC                     |
| 2   | HBX, HCZ, HHS, HLS, HMT, HQR, HXT, HDC, HAS, HXB, TBJ, TDG, TGJ, TJC, TLR, TNJ, TJR, TJS, TPS, TWR, BCR, BGE, BTR, BXG, BXS |
| 3   | HCD, HDJ, HHD, BYD, LBB, LTA, LLC                                                                                           |
| 4   | LHZ, LLZ, LWH, LLW, LLY, LJN, LLB, LRZ, LLR, LWF                                                                            |
| 5   | LDH, LHD, LQD, LYD                                                                                                          |

对表 5 的分群结果做进一步分析。分群 1 中西北电网所有机组分为一组同调机群, 表现出一致的

同调性; 此外, 北京机群中的 BLD、BSH 和 BZC 被归入到分群 1。津济 K1-鲁济京 K1 特高压输电线路的送端连接京北特、津板桥站, 故障后京、津、冀地区电网受到较大扰动。在表 5 的分群结果中, 京、津、冀 3 个地区的大部分机组被划分为第 2 群。通过对第 1 群、第 2 群机组的功角、电压、转子角速度曲线进行对比, 可得出同调分群结果与时域仿真结果一致。

津济 K1-鲁济京 K1 输电线路的受端连接到鲁泉城 T01 站, 线路发生故障后, 导致山东电网功率不平衡, 出现较大扰动, 通过分群结果看出, 山东地区的机组被划分为 3 个同调机群。除此之外, 北京机群的 BYD、河北机群的 HCD、HDJ、HHD 被归入到第 3 群。图 10 给出了山东电网各同调机群的指标曲线对比。对山东电网中机组的各指标的曲线进行观察, 从图 10 中可看出, 山东电网受到较大扰动, 各地区的机组表现出较大的动态差异。从图 10 中看出, 第 5 组(鲁胶东地区)机组的指标曲线波动较大, 与山东其他地区机组呈现出明显的非同调特性。通过对图 10 中 3、4、5 群中各指标曲线的绝对值距离、变化趋势、波动的综合对比, 认为将山东机群分为 3 群的结果是可靠的。

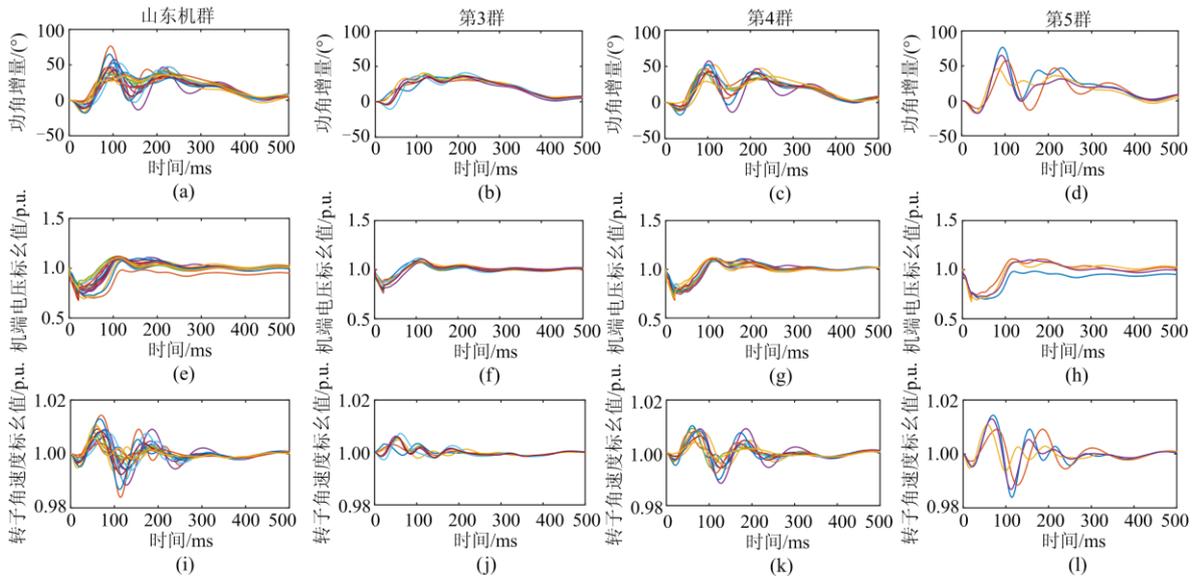


图 10 山东电网中各同调机群的指标曲线对比图

Fig. 10 Comparison of index curves of different groups in Shandong Power Grid

通过对华北电网算例进行分析, 可得出结论: 故障后, 各地区的发电机组总体上有一致的同调性, 但由于故障对各省网的影响不同, 故各区域内的发电机组又有不同的同调分群。本文所提方法给出的分群结果与时域仿真结果吻合, 从而验证了所提方

法的有效性, 同时说明本文方法在实际系统中具有较高的工程应用价值。

### 4 结论

已有的同调识别方法多是基于单一指标进行分

群,且其特征提取不充分,这可能造成分群结果不准确。为解决此问题,本文提出一种基于多指标面板数据特征提取的同调机组分群方法,理论分析和仿真算例验证了所提方法具有如下优点:

1) 客观选取了多个反映机组受扰动后动态特性的3个指标,建立了基于功角、机端电压和转子角速度的同调机组分群方法,提高了分群结果的准确性。

2) 基于面板数据特征提取的方法在时间序列上提取的“均值”特征、“波动”特征、“偏度”特征和“峰度”特征以及“趋势”特征,可充分提取各指标曲线在时间序列上的特征,为高精度分群奠定了基础。

3) 基于AHP-熵权法将主观与客观赋权法相结合,在主观与客观赋权后引入熵值,运用熵值对AHP赋权法进行修正,既考虑了专家的意见,又客观反映了样本数据的定量特性。

需要指出的是,本文选取的指标仅限于瞬时动态数据,下一步的研究方向将考虑把静态工况数据、系统模型数据纳入到同调分群中。

## 参考文献

- [1] 余贻鑫,李鹏. 大区电网弱互联对互联系统阻尼和动态稳定性的影响[J]. 中国电机工程学报, 2005, 25(11): 6-11.  
YU Yixin, LI Peng. The impact of weak interconnection of bulk power grids to damping and dynamic stability of power systems[J]. Proceedings of the CSEE, 2005, 25(11): 6-11.
- [2] REN Chao, XU Yan, ZHANG Yuchen. Post-disturbance transient stability assessment of power systems towards optimal accuracy-speed tradeoff[J]. Protection and Control of Modern Power Systems, 2018, 3(2): 194-203. DOI: 10.1186/s41601-018-0091-3.
- [3] PODMORE R. Identification of coherent generators for dynamic equivalent[J]. IEEE Transactions on Power Apparatus Systems, 1978, 97(7): 1344-1354.
- [4] 倪敬敏,沈沉,李颖,等. 主动解列控制中电网弱连接的一种在线识别方法[J]. 中国电机工程学报, 2011, 31(4): 24-30.  
NI Jingmin, SHEN Chen, LI Ying, et al. An on-line weak-connection identification method for controlled islanding of power system[J]. Proceedings of the CSEE, 2011, 31(4): 24-30.
- [5] 周虎兵,张焕青,杨增力,等. 二次系统隐性故障的多指标综合风险评估[J]. 电力系统保护与控制, 2019, 47(9): 120-127.  
ZHOU Hubing, ZHANG Huanqing, YANG Zengli, et al. Multi-criteria integrated risk assessment of secondary system hidden failures[J]. Power System Protection and Control, 2019, 47(9): 120-127.
- [6] 陈刚,丁理杰,李旻,等. 异步联网后西南电网安全稳定特性分析[J]. 电力系统保护与控制, 2018, 46(7): 76-82.  
CHEN Gang, DING Lijie, LI Min, et al. Stability characteristics of Southwest China Power Grid after asynchronous interconnection[J]. Power System Protection and Control, 2018, 46(7): 76-82.
- [7] 周海强,鞠平,孔德超,等. 基于机电距离的聚类方法在动态等值中的应用[J]. 电力系统自动化, 2008, 32(3): 14-17.  
ZHOU Haiqiang, JU Ping, KONG Dechao, et al. Application of clustering method based on electromechanical distance between different units in power system dynamic equivalence[J]. Automation of Electric Power Systems, 2008, 32(3): 14-17.
- [8] WU F F, TSAI Y K. Identification of groups of  $\epsilon$ -coherent generators[J]. IEEE Transactions on Circuits and Systems, 1983, 30(4): 234-241.
- [9] KIM H, JANG G, SONG K. Dynamic reduction of the large-scale power systems using relation factor[J]. IEEE Transactions on Power Systems, 2004, 19(3): 1696-1699.
- [10] YUSOF S B, ROGERS G J, ALDEN R T H. Slow coherency based network partitioning including load buses[J]. IEEE Transactions on Power Systems, 1993, 8(3): 1375-1382.
- [11] 戴晨松,薛峰,薛禹胜. 受扰轨迹的分群研究[J]. 电力系统自动化, 2000, 24(1): 13-16.  
DAI Chensong, XUE Feng, XUE Yusheng. Classification of disturbed trajectories[J]. Automation of Electric Power Systems, 2000, 24(1): 13-16.
- [12] 冯康恒,张艳霞,刘志雄,等. 基于广域信息的同调机群在线识别方法[J]. 电网技术, 2014, 38(8): 2082-2086.  
FENG Kangheng, ZHANG Yanxia, LIU Zhixiong, et al. A wide area information based online recognition of coherent generators in power system[J]. Power System Technology, 2014, 38(8): 2082-2086.
- [13] PAI M A, ADGAONKAR R P. Electromechanical distance measure for decomposition of power systems[J]. Electrical Power & Energy Systems, 1984, 6(4): 249-254.
- [14] 王振浩,杜虹锦,李国庆,等. 基于  $t$ -分布邻域嵌入的同调机群无监督识别[J]. 电力系统保护与控制, 2018, 46(22): 64-71.  
WANG Zhenhao, DU Hongjin, LI Guoqing, et al. Unsupervised identification of coherent generators based

- on t-distributed stochastic neighbor embedding[J]. Power System Protection and Control, 2018, 46(22): 64-71.
- [15] 文俊, 刘天琪, 李兴源, 等. 在线识别同调机群的优化支持向量机算法[J]. 中国电机工程学报, 2008, 28(25): 80-85.  
WEN Jun, LIU Tianqi, LI Xingyuan, et al. On-line identification of coherent generator using optimized LS-SVM[J]. Proceedings of the CSEE, 2008, 28(25): 80-85.
- [16] ANAPARTHI K K, CHAUDHRI B, THORNHILL N F. Coherency identification in power systems through principal component analysis[J]. IEEE Transactions on Power Systems, 2005, 20(3): 1658-1660.
- [17] SUSUKI Y, MEZIC I. Nonlinear Koopman modes and coherency identification of coupled swing dynamics[J]. IEEE Transactions on Power Systems, 2011, 26(4): 1894-1904.
- [18] 宋洪磊, 吴俊勇, 郝亮亮, 等. 基于 WAMS 和改进拉普拉斯特征映射的同调机群在线识别[J]. 电网技术, 2013, 37(8): 2157-2164.  
SONG Honglei, WU Junyong, HAO Liangliang, et al. On-line identification of coherent generator based on WAMS and improved Laplacian eigenmap algorithm[J]. Power System Technology, 2013, 37(8): 2157-2164.
- [19] AVDAKOVIC S, BECIROVIC E, NUHANOVIC A, et al. Generator coherency using the wavelet phase difference approach[J]. IEEE Transactions on Power Systems, 2014, 29(1): 271-278.
- [20] 宋洪磊, 吴俊勇, 冀鲁豫. 基于慢同调理论和希尔伯特-黄变换的发电机在线同调识别[J]. 电力自动化设备, 2013, 33(8): 70-76.  
SONG Honglei, WU Junyong, JI Luyu. Online identification of coherent generators based on slow coherency theory and Hilbert-Huang transform[J]. Electric Power Automation Equipment, 2013, 33(8): 70-76.
- [21] 吴兴扬, 卫志农, 孙国强, 等. 基于非负矩阵分解的同调机群识别方法[J]. 电力系统自动化, 2013, 37(14): 59-64.  
WU Xingyang, WEI Zhinong, SUN Guoqiang, et al. A method for identifying coherent generators based on non-negative matrix factorization[J]. Automation of Electric Power Systems, 2013, 37(14): 59-64.
- [22] 党耀国, 侯荻青. 基于特征提取的多指标面板数据聚类方法[J]. 统计与决策, 2016(19): 68-72.  
DANG Yaoguo, HOU Diqing. Multi-index panel data clustering method based on feature extraction[J]. Statistics & Decision, 2016(19): 68-72.
- [23] 魏少攀, 杨明, 韩学山, 等. 一种利用机端电压幅值轨迹 MLE 指标的暂态功角稳定在线辨识方法[J]. 中国电机工程学报, 2017, 37(3): 3775-3786.  
WEI Shaopan, YANG Ming, HAN Xueshan, et al. An online identification method for transient angle stability using MLE index of generator terminal voltage amplitude trajectory[J]. Proceedings of the CSEE, 2017, 37(3): 3775-3786.
- [24] 荆朝霞, 曾丽. 基于改进层次分析法的配电网中压改造项目迫切度分析[J]. 电力系统自动化, 2011, 35(4): 92-95.  
JING Zhaoxia, ZENG Li. Analysis on urgency degrees of distribution network reconstruction scheme based on advanced AHP[J]. Automation of Electric Power Systems, 2011, 35(4): 92-95.
- [25] 董福贵, 张也, 尚美美. 分布式能源系统多指标综合评价研究[J]. 中国电机工程学报, 2016, 36(12): 3214-3222.  
DONG Fugui, ZHANG Ye, SHANG Meimei. Multi-criteria comprehensive evaluation of distributed energy system[J]. Proceedings of the CSEE, 2016, 36(12): 3214-3222.

收稿日期: 2019-10-23; 修回日期: 2019-12-13

作者简介:

张智(1994—), 男, 硕士研究生, 研究方向为电力系统动态等值; E-mail: zhangzhincepu@163.com

陈艳波(1982—), 男, 副教授, 博士生导师, 研究方向为电力系统分析、稳定和控制等。E-mail: yanbochen2008@sina.com

(编辑 魏小丽)