

# 短时间尺度用电行为相关性分析网络模型

陈鹏伟, 陶 顺, 肖湘宁, 李 璐, 张 剑

(新能源电力系统国家重点实验室(华北电力大学), 北京市 102206)

**摘要:**用电行为短时间尺度相关关系是用户精细化管理与智能配用电的重要决策基础。建立了针对短时间尺度用电行为相关性的网络化分析方法。首先,由描述短时间尺度用电功率序列间相关性的皮尔逊相关系数构造用电行为相关矩阵,为减小随机用电行为与测量误差等无序噪声信息对相关矩阵的影响,引入随机矩阵理论对相关矩阵进行去噪,并提出利用特征值谱熵的正则化修正方法。然后,基于相关矩阵构建了相关矩阵分阶与同级网络序列,提出了基于社团结构的相关性聚类方法与基于最小树的相关性等级结构挖掘方法;定义了量度用电相关性的拓扑指标,从而构造了用电行为相关性聚类分析与等级分析的网络模型,并就模型的典型应用场景进行了讨论。最后,通过实例数据验证了上述去噪方法、相关性分析网络模型及指标的有效性。

**关键词:**相关性;随机矩阵理论;谱熵;社团结构;最小生成树

## 0 引言

智能配用电是智能电网的重要环节,随着分布式电源、储能、电动汽车、需求响应等多种新型供电形式、各类先进理念和技术的不断出现,社会对供电服务的需求日趋多样化<sup>[1]</sup>,除电力成本更低、用电安全舒适等需求外,个性化、便捷性及互动性强的供电服务和优质电力供应亦成为用户多元需求的重要组成部分<sup>[2]</sup>。同时,随着用电信息采集、配电自动化等系统的不断完善,配电与用电的基础与运行数据近年来出现爆炸性增长,并逐渐呈现出海量、数据项复杂、处理逻辑复杂、存储周期长、计算频度高等大数据特征<sup>[3]</sup>。由于配用电数据隐藏着用户用电行为特性,且用电行为特性之间具有一定的关联性与相似性,并呈现多元化特征,因此如何从配用电数据中挖掘出用户类型与用电行为特性,已成为电网进一步拓展个性化与差异化服务的重要决策基础,由此亟须研究体系化的理论模型、挖掘方法和支撑技术,并结合用电互动机制加以应用,以期提高智能配用电水平<sup>[4-5]</sup>。

针对用户用电行为的分析,已有一些学者开展了相关研究。如文献[6]对终端电力用户进行了聚类研究,基于自组织映射神经网络实现了负荷曲线

的聚类;文献[7]将  $k$ -means 聚类算法与云平台并行计算相结合,实现多维数据下的居民用户用电类型挖掘;文献[8]进一步利用智能电表测量数据对用户按用电模式进行相似性聚类,并将聚类结果应用于系统级日内负荷预测;文献[9]分别针对电力用户用电数据分析技术、配电变压器负荷影响模型构建及短期网供负荷预测等典型场景应用进行了探讨;文献[10]建立了峰谷分时电价下的用户行为模型、响应参数识别及校正模型;文献[11]考虑了电动汽车充电需求特征,提出了一种多级时变负荷行为模型;文献[12]则利用实测数据建立了电动汽车充电站的负荷行为模型。可以发现,现有研究多集中于负荷分类与单一负荷或单一情景下的用电行为分析,对于不同时段多用户间相关关系动态变化过程中的用电行为分析研究鲜有开展。

多用户用电行为短时间尺度(小时级)相关关系估计是用电行为动态变化特性分析与相关预测的研究基础,但由于用户用电行为除受自身不确定需求引导外,还受复杂社会系统各方面因素的交互影响与制约,相关关系随时间变化,且随负荷类型不同而呈现多样性。若以有限的时间序列估计相关关系,则用户随机用电行为的影响(如工业非计划生产、电动汽车随机充电行为、居民用户按临时不确定需求用电等)与测量误差会引入噪声信息,而用较长的时间序列来规避这个问题,则容易忽略短时间尺度非平稳的相关性,如何估计短时间尺度相关关系是一个值得研究的难题。

为建立针对用电行为相关关系动态变化过程的

收稿日期:2016-01-27;修回日期:2016-05-12。

上网日期:2016-07-18。

国家自然科学基金资助项目(51207051);中央高校基本科研业务费专项资金资助项目(2016XS02)。

分析模型,本文首先采用皮尔逊相关系数对用电功率序列进行相关性分析,然后构建了相关矩阵分阶与同级网络序列,提出了基于社团结构的相关性聚类方法与基于最小树的相关性等级结构挖掘方法,以及相关性聚类分析与等级分析的网络模型,并就模型典型应用场景进行了探讨。

## 1 用电行为相关性模型

### 1.1 用电行为相关矩阵

设馈线或用户  $i$  在某一时段的用电功率序列为  $p_i = \{P_i(1), P_i(2), \dots, P_i(k), \dots, P_i(L)\}$ , 其中  $L$  为序列长度,按下式进行归一化处理:

$$r_i(k) = \frac{P_i(k) - \mu_i}{\sigma_i} \quad (1)$$

式中:  $\mu_i$  为序列样本均值;  $\sigma_i$  为样本标准差。则变换后用户  $i$  的用电功率序列符合标准正态分布。继而,由任意用电功率归一化序列  $i$  与  $j$  之间的皮尔逊线性相关系数  $\rho_{i,j}$  可以得到用电行为的相关矩阵  $\mathbf{R}$ , 即维数为  $N \times N$  阶的用电功率序列样本协方差矩阵,  $N$  为馈线数或用户数。

### 1.2 基于随机矩阵极限谱理论的相关矩阵去噪处理

由于用电功率序列中存在随机因素引起的用电行为与测量误差等干扰短时间尺度非平稳相关关系估计的噪声信息,因此需要对相关矩阵  $\mathbf{R}$  进行去噪处理以提高分析准确性。

随机矩阵理论起源于量子物理的发展和研究, Wigner 最早将随机矩阵与量子物理结合起来,并引入到多维数学统计中<sup>[13]</sup>。随机矩阵理论的预测代表了所有可能相互作用的一种平均,通过与随机多维时间序列统计特性的比较,可以体现实际数据中对随机的偏离程度,并揭示实际数据中整体关联的行为特性<sup>[14]</sup>。随后,随机矩阵理论被广泛应用于物理、金融、生物统计、网络科学、大数据等领域<sup>[15-17]</sup>。

由随机矩阵理论中关于相关矩阵的 M-P 律<sup>[18]</sup>,对于随机互相关矩阵  $\mathbf{S}_n$ :

$$\mathbf{S}_n = \frac{1}{L} \mathbf{A} \mathbf{A}^T \quad (2)$$

式中:  $\mathbf{A}$  是由均值为 0、方差为  $\sigma^2$  的独立同分布的随机变量  $x_{ij}$  ( $1 \leq i \leq N, 1 \leq j \leq L$ ) 组成的矩阵。

当  $N \rightarrow \infty, L \rightarrow \infty, W = L/N (> 1)$  固定时,随机互相关矩阵  $\mathbf{S}_n$  的经验谱分布依概率 1 收敛到极限谱  $F_W(\lambda)$ , 其密度函数为:

$$h(\lambda) = \begin{cases} \frac{W \sqrt{(\lambda_{\max} - \lambda)(\lambda - \lambda_{\min})}}{2\pi\sigma^2 \lambda} & \lambda_{\min} \leq \lambda \leq \lambda_{\max} \\ 0 & \text{其他} \end{cases} \quad (3)$$

式中:  $\lambda$  为  $\mathbf{S}_n$  的特征值;  $\lambda_{\min} = \sigma^2(1 - \sqrt{1/W})^2$ ;  $\lambda_{\max} = \sigma^2(1 + \sqrt{1/W})^2$ 。  $\sigma^2 = 1$  时为标准 M-P 律。

相关矩阵  $\mathbf{R}$  为半正定实 Hermite 矩阵,因而谱可分解为:

$$\mathbf{R} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^T \quad (4)$$

式中:  $\mathbf{U} \mathbf{U}^T = \mathbf{I}$ ;  $\mathbf{\Lambda} = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_N\}$ ;  $\mathbf{I}$  为单位矩阵。

比较相关矩阵  $\mathbf{R}$  与随机相关矩阵的特征值统计性质,  $\mathbf{R}$  的特征值中有一部分符合随机矩阵性质,即属于  $[\lambda_{\min}, \lambda_{\max}]$  体现随机用电行为和测量误差的噪声部分。为保留含有真实信息的差异部分,可用零替代相关矩阵噪声特征值<sup>[19]</sup>, 即令

$$\mathbf{\Lambda}_{\text{New}} = (\mathbf{\Lambda} - \mathbf{\Lambda}_r) + \mathbf{\Lambda}_0 \quad (5)$$

式中:  $\mathbf{\Lambda}_r$  为由噪声特征值  $\lambda_i$  ( $\lambda_{\min} \leq \lambda_i \leq \lambda_{\max}, 1 \leq i \leq N$ ) 组成的矩阵;  $\mathbf{\Lambda}_0$  为由零组成的矩阵。

则去噪后相关矩阵为:

$$\mathbf{R}_{\text{New}} = \mathbf{U} \mathbf{\Lambda}_{\text{New}} \mathbf{U}^T \quad (6)$$

为保证  $\mathbf{R}_{\text{New}}$  与原矩阵迹相同,需将  $\mathbf{R}_{\text{New}}$  的对角元素全设置为 1,从而实现极限谱估计下相关矩阵噪声信息的过滤。

### 1.3 考虑特征值谱熵的去噪正则化修正

由于式(3)反映的仅是参数为  $W$  与  $\sigma^2$  时随机相关矩阵的极限谱,并不是相关矩阵  $\mathbf{R}$  中实际噪声信息的准确估计,因而在对不同时段用电行为相关性进行估计时,式(5)与式(6)的过滤方式存在一定的滤除误差。为减小滤除误差,提出一种考虑特征值谱熵的正则化修正方法。

谱熵<sup>[20]</sup>是熵在矩阵统计学中引申定义,用来量度特征值的信息量,具体定义如下:

$$\Psi_{\text{SE}} = -\frac{1}{\log N} \sum_{i=1}^N \Omega(i) \log \Omega(i) \quad (7)$$

式中:  $w(i) = \lambda(i) - \lambda(i-1)$ ,  $\lambda$  为相关矩阵的特征值,且

$$\Omega(i) = \frac{w^2(i)}{\sum_{i=1}^N w^2(i)} \quad (8)$$

由式(7)和式(8)可以看出,谱熵具有以下性质。

1)  $\Psi_{\text{SE}} \geq 0$ , 当且仅当  $\Omega_1, \Omega_2, \dots, \Omega_{N-1}$  中有一个为 1 时,即只存在一个最大特征值且其他特征值为 0 时,  $\Psi_{\text{SE}} = 0$ , 其他情况特征值谱熵均大于 0。

2) 当且仅当  $\Omega_1, \Omega_2, \dots, \Omega_{N-1}$  均相等时,即所有特征值是等间隔的情况下,特征值熵才有最大值,为 1。

对于随机相关矩阵,其特征值最近邻间隔近似服从 Wigner-Dyson 分布,而对于贫关联体系的相关矩阵  $\mathbf{R}$  则近似 Poisson 分布,且随着贫化程度加

深,越趋近于 Poisson 分布<sup>[17]</sup>,对应谱熵的积累速度也逐渐加快。由此,针对式(5)与式(6),由式(7)与式(8)构建正则化特征值判定函数:

$$f(\lambda_i) = \lambda_i - \lambda_{\max} + (\lambda_{\max} - \lambda_{\min}) \left( 1 + \frac{1}{\Psi_{SE} \log N} \sum_{k=1}^{i-1} \Omega(i) \log \Omega(i) \right) \quad (9)$$

当  $f(\lambda_i) \leq 0$  时,则对特征值  $\lambda_i$  进行零替代,否则保留。正则化项的引入使得在滤除噪声特征值时对  $\lambda_{\max}$  附近特征值的归属实现自适应判定与过滤。

## 2 用电行为相关性分析网络模型的建立

### 2.1 相关矩阵分阶与同级网络构建

相关矩阵  $\mathbf{R}_{\text{New}}$  虽然描述了各用电功率序列之间的相关性,但不能直接反映各序列在用电功率序列样本总体中的等级结构及总体序列样本的群体特性。因此,采用网络化处理思路,在相关矩阵  $\mathbf{R}_{\text{New}}$  基础上,定义用电功率序列间的相关性距离:

$$d_{ij} = \sqrt{2(1 - \rho_{i,j})} \quad (10)$$

由式(10)可知,序列间相关性越大,则相关性距离越小,  $\rho_{i,j} = 1$  完全相关时,相关性距离  $d_{ij} = 0$ 。通过式(10)使  $\mathbf{R}_{\text{New}}$  转化为有权邻接矩阵  $\mathbf{D}$ ,从而可建立体现各序列间相关性的无向连通相关性网络  $G(V, E)$ ,  $V$  为体现用电功率序列的节点集合,  $E$  为体现序列间相关性的边集合。

$G$  为  $(N-1)$ -正则的非平面图,总边数为  $N(N-1)/2$ ,其网络复杂度随着  $N$  的增大而平方倍增长。为便于研究用电行为动态演化过程与拓扑特性的可视化及针对性比较,需要对  $G$  分阶处理,具体做法如下:①根据相关性分析的对象规模与精度需求构建分析区间  $\Theta = \{\theta_n | \Delta\theta = \theta_n - \theta_{n-1}, \theta_n \in (0, 1]\}$ ,  $\Delta\theta$  为间隔常数;②利用每个区间刻度  $\theta_n$  对有权邻接矩阵  $\mathbf{D}$  进行重置。构造网络样本  $g_{\theta_n}$ ,映射规则如式(11)所示,从而构成分阶网络序列  $g_{\Theta} = \{g_{\theta_n}\}$  与同级网络序列  $G_{\Theta} = \{G_{\theta_n} | G_{\theta_n} = g_{\theta_n} - g_{\theta_{n-1}}\}$ 。

$$\begin{cases} d_{ij} \rightarrow d_{ij} & d_{ij} \leq \theta_n \\ d_{ij} \rightarrow \infty & d_{ij} > \theta_n \end{cases} \quad (11)$$

分阶处理后,对于  $g_{\Theta}$  中网络样本  $g_{\theta_n}$ ,在相关性距离水平  $[0, \theta_n]$  (即相关性水平  $[1 - \theta_n^2/2, 1]$ ) 下,定义序列占有率(sequence share, SS)来量度总体序列的相关性分布水平:

$$\Phi_{SS} = \frac{1}{N} \sum_{k=1}^K N_k = \frac{N_T}{N} \quad (12)$$

式中:  $K$  为区间划分作用下  $g_{\theta_n}$  中出现的连通子图数;  $N_k$  为连通子图  $k$  包含的节点数。

基于连通图任意节点间平均最短距离,定义相关性距离水平  $[0, \theta_n]$  内所含序列间特征平均相关度(mean characteristic relevance, MCR),量度含多个连通子图情况下网络结构的密集程度(正相关性强弱):

$$\zeta_{\text{MCR}} = \sum_{k=1}^K \left[ \frac{1}{N_T(N_k - 1)} \sum_{i,j,i \neq j} l_{ij} \right] \quad (13)$$

式中:  $l_{ij}$  为节点  $i$  与  $j$  间最短路径。

### 2.2 基于社团结构的用电相关性聚类

由于区间的划分作用,  $G_{\Theta}$  各样本连通子图会出现结构内连接稠密而结构间连接稀疏的网络子图,即社团结构,如图1所示。社团结构作为研究网络拓扑结构和网络组成的重要特征,包含模块、类、群、组等各种含义,目前还没有广泛认可的唯一定义,较为常用的是基于相对连接频数的定义<sup>[21]</sup>。该结构体现了结构内所含序列的群体相关程度。以图1所示网络样本为例,在区间  $[\theta_{n-1}, \theta_n]$  内,社团结构1与社团结构2各序列间会呈现自相关性且小于  $\theta_{n-1}$ ,而社团外序列  $G, F$  到序列  $A, B, C, D$  的距离大于  $\theta_{n-1}$ ,反之亦然。通过对同级网络序列  $G_{\Theta}$  连通子图社团结构的挖掘,从而实现不同相关水平范围内用电行为相关性聚类。

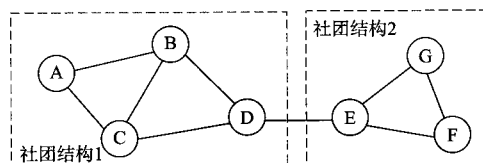


图1 基于社团结构的相关性聚类  
Fig.1 Correlation clustering based on community structure

社团结构划分的方法可分为基于网络拓扑结构、基于网络动力学与基于模块度优化等三类算法,考虑  $G_{\Theta}$  中可能存在的稀疏网络元素,本文选择模块度优化代表算法——Fast Newman 算法<sup>[22]</sup>进行社团结构挖掘。

Girvan 和 Newman 提出了网络划分质量的模块度函数用于衡量网络划分的质量标准<sup>[23]</sup>:

$$Q = \text{Tr}(\mathbf{e}) - \|\mathbf{e}^2\|_1 = \sum_{i=1}^n (e_{ii} - a_i)^2 \in [0, 1] \quad (14)$$

式中:  $\mathbf{e}$  为社团划分结果的  $n \times n$  阶矩阵;  $\text{Tr}(\mathbf{e})$  为  $\mathbf{e}$  的迹;  $\|\mathbf{e}^2\|_1$  为矩阵范数,  $\mathbf{e}$  中元素  $e_{ij}$  表示连接社团  $i$  与社团  $j$  节点的边占有所有边的比例,  $e_{ii}$  表示社团  $i$  内部节点的边占有所有边的比例,  $a_i$  为与社团  $i$

内部节点相连的边占有所有边的比例,即

$$a_i = \sum_{j=1}^n e_{ij} \quad (15)$$

针对不同划分方式,  $Q$  会有相应的值,一般认为其峰值对应的划分结果为最优划分,其值称为最优模块度。Fast Newman 算法基于贪婪思想,按一个节点即一个社团进行初始化,并沿  $Q$  增大最多或减小最少的方向逐次合并,对应最大  $Q$  值的合并结果即为最优划分;结合网络样本连通子图的划分,可最终实现同级网络序列  $G_\theta$ 。按社团结构的用电行为相关性聚类。

### 2.3 基于最小树的相关性等级结构挖掘

为进一步挖掘各用电功率序列在相关性距离水平  $[0, \theta_n]$  约束时的等级结构,采用最小生成树(minimum spanning tree, MST)对分阶网络序列  $g_\theta$  中各网络样本元素进行重构。MST 使用原网络的  $N-1$  条边连接  $N$  个节点而形成树结构,树中各边的权值综合最小且不产生回路。通过 MST 对  $g_\theta$  的重构,可选择出正相关性最强(即距离最小)的  $N-1$  条边构成紧密联系且具有规则拓扑结构的简化网络,对  $g_\theta$  最小生成树的分析就可对用电功率序列间相关性等级结构进行定量量测与典型用电功率序列挖掘。

本文采用 Kruskal 算法进行 MST 重构,并提出以下在相关性距离水平  $[0, \theta_n]$  下序列间相关性等级结构的拓扑指标,用于多个连续短时间尺度的用户间相关性网络动态演化特点的测度。

1) 根据网络样本  $g_{\theta_n}$  各连通区域的重构最小树,定义 MST 的归一化相关度(normalized tree relevance, NTR)用来量度 MST 的联系强弱:

$$\alpha_{NTR} = \sum_{k=1}^K \left[ \frac{1}{N_T(N_k - 1)} \sum_{d_{ij} \in \Phi_k} d_{ij} \right] \quad (16)$$

式中:  $\Phi_k$  为  $g_{\theta_n}$  连通子图  $k$  最小树所有边的集合。

2) 定义非叶子节点数占有率(non-leaf node share, NNS)用来量度 MST 的松散程度:

$$\beta_{NNS} = \sum_{k=1}^K \frac{N_k^t}{N_T} \quad (17)$$

式中:  $N_k^t$  为  $g_{\theta_n}$  连通子图  $k$  最小树的非叶子节点数,即树中度大于 1 的节点。

需要指出,  $\Phi_{SS}$ ,  $\zeta_{MCR}$ ,  $\alpha_{NTR}$  和  $\beta_{NNS}$  均是通过对相关矩阵分阶网络  $g_\theta$  边和节点的量度来体现群体间整体相关性程度,特别地,  $\alpha_{NTR}$  和  $\beta_{NNS}$  着重针对的是  $g_\theta$  的 MST 网络。上述指标除  $\Phi_{SS}$  表示不同相关性水平下序列的分布特点外,  $\zeta_{MCR}$ ,  $\alpha_{NTR}$  和  $\beta_{NNS}$  的值越小,则分别从不同角度表征网络联系越紧密,即相关性越强。

## 3 模型典型应用场景讨论

通过上述短时间尺度相关性分析网络模型对负荷数据多时段的分析,可获得某区域各负荷间用电相关性在各时段的演变过程和不同时段不同相关性水平下的负荷覆盖范围及组成,从而可服务于智能配用电及其相关领域。以下就模型宜应用的典型场景进行分析和探讨。

### 1) 极端事件下用电行为响应特性分析

短时间尺度用电序列受极端降雨、高温、低温、台风等天气事件和假日节庆、团体活动等大型社会活动直接影响,用户用电行为对这些强激励因素的响应特性存在一定程度的一致性。现有关于极端事件对用电影响的研究多聚焦于两者之间的关联性,而本文所提模型则关注事件对用户用电行为的共性影响及其影响范围,属于另一个研究角度,可以作为分析工具用于极端事件的挖掘、影响范围分析和预测,为事件应对提供决策基础。

### 2) 智能配电网(微电网)特性分析

所提模型不仅适用于负荷,且同样适用于分布式电源间、电源和负荷间的相关性分析。对于智能配电网(微电网)概率潮流等不确定分析方法而言,分布式电源及负荷间的相关性需有所考虑,文献[24]即验证了风力发电系统出力相关性对计算结果的影响。在短时间尺度下对系统不确定性分析时,利用本文模型获得的分布式电源、负荷及两者间短时间尺度相关性估计和聚类结果可简化相关性条件,提高分析精度,从而可进一步指导分布式电源和分布式储能的扩展规划,以及含储能系统的运行优化。

### 3) 负荷分群分时管理

不同时段相关性聚类结果是负荷分群分时管理的划分依据之一。在结合用户类型及其基础电价后,聚类结果可为分时或实时电价制定提供一定参考,即电价引导使得同类用户用电正相关性增强而不同类别间用户用电相关性减弱,从而实现电价调节范围内削峰填谷等目标。

## 4 实例验证及结果分析

### 4.1 数据来源

本文的实验数据来源于某市 220 kV 变电站供电区域内 10 kV 用户馈线计量数据,收集时间为 2014-01-01T00:00:00 至 2014-01-02T24:00:00,采样间隔 15 min。选择该配电区域内 100 条馈线作为分析对象,部分馈线监测数据参见附录 A 图 A1。分析窗长设为 6 h,为满足短时间尺度相关性分析

的需要,对时窗内数据进行 3 min 间隔三次样条插值以满足基于随机矩阵理论去噪方法的序列长度限值,分析时  $\Delta\theta$  取 0.05 或 0.1。

#### 4.2 相关矩阵去噪前后对比

##### 1) 关键节点与 MST 对比

取 00:00—05:57 监测数据构造原始相关矩阵,设区间刻度 0.8,对应相关性区间  $[0.68, 1]$ ,则  $g_{\theta_n}=0.8$  原始网络与去噪网络关键节点及节点度值如表 1 所示。

节点的度值反映了节点的重要性。由表 1 可知,关键节点的度变化较小,去噪网络保留了原始相关性网络的关键信息,基于随机理论的相关矩阵去噪处理仅过滤了相关性网络中的非关键信息。为进一步可视化这一特性与控制网络样本节点规模,取较小区间刻度 0.6,并对原始网络与去噪网络进行 MST 重构,以用作网络平面展开,原始网络与去噪

网络主连通子图的 MST 分别如图 2(a)和图 2(b)所示。

表 1 原始网络与去噪网络关键节点对比  
Table 1 Comparison of key nodes between original network and denoising network

原始网络		去噪网络	
节点编号	节点度值	节点编号	节点度值
8	44	96	43
95	42	93	43
93	42	95	41
96	41	80	41
80	41	8	41
7	41	7	41
66	40	98	39
90	39	37	39
37	39	66	38
98	38	90	38

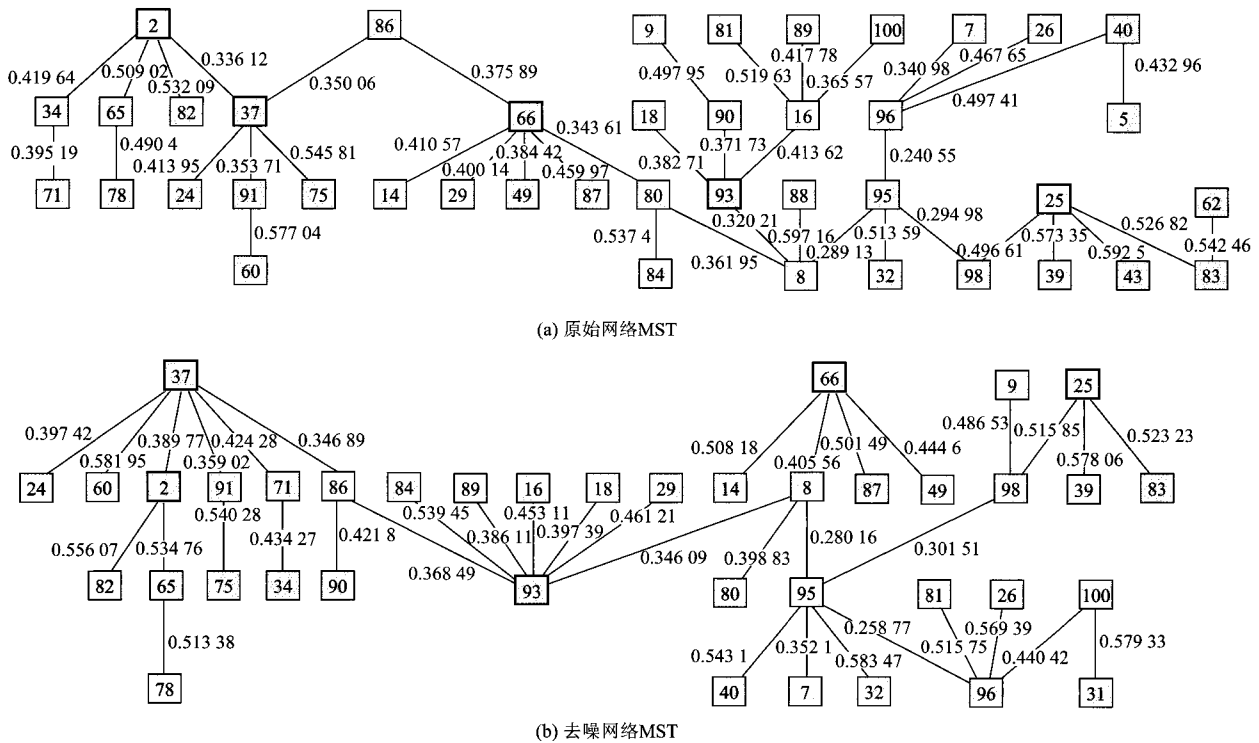


图 2 原始网络与去噪网络的 MST 对比 ( $\theta_n=0.6$ )

Fig.2 Comparison of MST between original network and denoising network ( $\theta_n=0.6$ )

由图 2(a)原始网络 MST 与图 2(b)去噪网络 MST 的对比可见,去噪处理对关键节点的影响较小,如节点 2,25,37,66,93 等。需要指出,MST 就节点信息而言存在局部性与表征不足,但适用于分析网络等级结构,能够平面化展示与突出网络中心节点。

选择图 2 中节点 37,24,91,75,进一步分析所

提去噪方法的效果。这 4 个节点的原始监测数据参见附录 A 图 A2,原始序列间的相关性  $\rho_{37,24}$ ,  $\rho_{37,91}$ ,  $\rho_{37,75}$  分别为 0.914 3,0.937 4 和 0.851 0,去噪后  $\rho_{37,24}=0.921$  0,  $\rho_{37,91}=0.935$  6,  $\rho_{37,75}<0.68$ 。可以发现,基于随机矩阵理论的去噪方法可以较好地滤除随机因素对相关性的影响,如馈线 24 与 37 间的相关性,使得相关系数增大,而类似馈线 75 序列上

体现的非随机用电行为(多个连续大幅值波动),则予以保留与突出,即馈线 75 与 37 间去噪后的相关系数较之实际序列计算结果有所减小。

2)模块度与特征平均相关度对比

对 1)中所述原始相关矩阵分阶处理,区间刻度依次取 0.4 至 0.9,  $\Delta\theta=0.05$ ,原始网络与去噪网络的最优模块度与特征平均相关度如表 2 所示。

表 2 原始网络与去噪网络模块度与特征平均相关度对比  
Table 2 Comparison of  $Q$  and  $\zeta_{MCR}$  between original network and denoising network

$\theta_n$	原始网络		去噪网络	
	最优模块度 $Q$	特征平均相关度 $\zeta_{MCR}$	最优模块度 $Q$	特征平均相关度 $\zeta_{MCR}$
0.90	0.100 2	1.535	0.092 9	1.349
0.85	0.105 6	1.272	0.093 9	1.215
0.80	0.117 3	1.153	0.105 3	1.148
0.75	0.150 1	1.155	0.135 8	1.186
0.70	0.160 8	1.225	0.143 6	1.223
0.65	0.179 9	1.112	0.174 6	1.110
0.60	0.194 3	1.053	0.188 5	1.027
0.55	0.222 1	1.067	0.198 9	1.002
0.50	0.284 1	0.904	0.257 0	0.804
0.45	0.345 7	0.858	0.287 0	0.880
0.40	0.415 0	0.992	0.343 5	0.942

如前面所述,模块度反映的是节点分布与节点间联系的均匀程度,而特征平均相关度则体现分布的密集程度,即正相关性(节点间距离)的强弱。由表 2 可知,通过去噪处理,用电功率序列相关性网络的特征平均相关度均有所减小,表征节点间联系变强,且最优模块度有所减小,表征节点分布更均匀。

表 3 基于动态时间弯曲距离的序列相似性测度  
Table 3 Sequence similarity measure based on dynamic time warping

序列	A1	A2	A3	B1	B2	B3	C1	C2	C3
A1	0	0.068 7	0.031 5	0.123 9	0.204 7	0.102 6	0.052 7	0.116 0	0.116 0
A2	0.068 7	0	0.073 1	0.380 8	0.359 6	0.338 9	0.191 2	0.119 9	0.137 7
A3	0.031 5	0.073 1	0	0.095 6	0.127 9	0.096 4	0.078 8	0.033 2	0.059 3
B1	0.123 9	0.380 8	0.095 6	0	0.030 9	0.067 8	0.135 6	0.065 2	0.149 9
B2	0.125 7	0.359 6	0.127 9	0.030 9	0	0.067 8	0.149	0.094 5	0.163 3
B3	0.204 7	0.338 9	0.096 4	0.067 8	0.067 8	0	0.156 2	0.086 2	0.101 3
C1	0.102 6	0.191 2	0.078 8	0.135 6	0.149 9	0.156 2	0	0.061 4	0.079 8
C2	0.052 7	0.119 9	0.033 2	0.065 2	0.094 5	0.086 2	0.061 4	0	0.046 5
C3	0.116 0	0.137 7	0.059 3	0.149 9	0.166 3	0.101 3	0.079 8	0.046 5	0

由图 A3 与表 3 可见,A1,A2 与 A3 间的动态时间弯曲距离明显小于与其他两类间的距离,表明相似度高,其他两类亦呈现明显的组内相似性与类间差异性,从而形成表 3 对角成块的特点,见红色标示部分。需要指出,在表 3 中可以发现,C1 与 A1 相似度高,但与 A2,A3 相似度低,同时与 C2,C3 比

由图 4 直接可见,去噪处理增加了网络的稳固程度,从而有利于进一步的序列间相关性拓扑分析。

4.3 不同相关性水平下社团结构聚类

取 06:00—11:57 馈线监测数据构造原始相关矩阵,设 $[\theta_{n-1},\theta_n]=[0.5,0.6]$ ,对应相关性区间为 $[0.820,0.875]$ ,对  $G_{\theta_n=0.6}$  主连通子图进行社团挖掘,划分树状图如图 3 所示,其中最大  $Q$  值为 0.300 2。

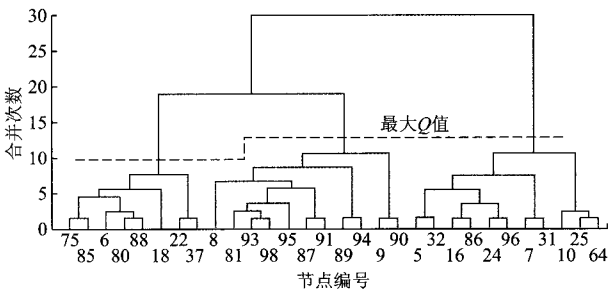


图 3 区间 $[0.5,0.6]$ 下社团结构划分树状图  
Fig.3 Decomposition tree of community structure with  $[0.5, 0.6]$

由图 3 可见,在相关水平 $[0.820,0.875]$ 内,用电功率序列经过最优社团挖掘可分为 3 类:A 类包含 6,18,22,37,75 等序列;B 类包含 8,9,81,87,89 等序列;C 类包含 5,7,16,24,31 等序列。3 类序列部分监测数据(每类取 3 个序列)可参见附录 A 图 A3。

采用动态时间弯曲距离<sup>[25]</sup>对附录 A 图 A3 所示序列间的相似性进行量化,测度结果如表 3 所示。

较,C1 属于C 类较之 A 类更合适,从而进一步验证了按社团结构进行不同相关性水平下序列间相关性聚类的有效性 with 准确性。

4.4 基于 MST 的相关性网络演化测度

为验证基于 MST 的相关性等级结构挖掘方法对相关性网络动态演化过程的测度能力,将 48 h 采

集数据按 6 h 的时间尺度划分为 8 个阶段,各分阶段序列不同阶段序列占有率  $\Phi_{ss}$  如图 4 所示。与附录 A 图 A1 所示用电功率序列及附录 A 图 A4 所示 100 条馈线总功率序列对比可知,由于元旦假期强社会因素的影响,2014-01-01T00:00:00—2014-01-02T23:45:00 时间段内存在 3 个相关性高峰,即阶段 1 假期开始、阶段 6 假期后复工与阶段 8,对应附录 A 图 A4 则体现为趋势的保持与曲线的光滑程度,其中尤以对阶段 6 假期后复工这一强社会激励因素的影响表征得最为突出。

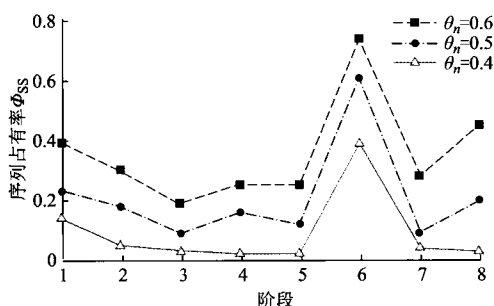


图 4 各分阶段序列不同阶段序列占有率  $\Phi_{ss}$   
Fig.4  $\Phi_{ss}$  at all stages with different  $\theta_n$

取  $\theta_n=0.4$ , 针对阶段 6 进行 MST 重构, 如图 5 所示。在阶段 6 的 MST 中, 由于 MST 重构对强联系的选择, 可筛选出中心节点 56, 9, 62 等典型负荷序列。

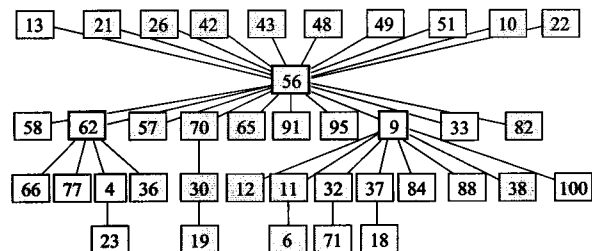


图 5 阶段 6 重构 MST ( $\theta_n = 0.4$ )  
Fig.5 MST Reconstruction of stage 6 with  $\theta_n = 0.4$

进一步验证 MST 网络对相关距离水平  $\theta_n$  约束时序列间相关性等级结构的表征能力, 取区间刻度 0.6, 不同阶段各指标量度结果如图 6 所示。

由图 6 与图 4 对比可知, 8 个阶段各指标演化趋势与序列占有率  $\Phi_{ss}$  的变化过程存在明显的负相关关系, MSR, NTR 与 NNS 均能对序列相关性网络的相关程度有所体现, 如在阶段 6 时  $\Phi_{ss}$  出现最高峰而其他 3 个指标出现最低谷。基于网络的特征平均相关度  $\zeta_{MCR}$  考虑所有节点间的相关度, 但由于区间重置作用, 其敏感度较低, 而基于 MST 的归一化相关度  $\alpha_{NTR}$  与非叶子节点数占有率  $\beta_{NNS}$  由于 MST 的筛选作用, 保留了节点间的强联系, 敏感性

更高, 验证了基于 MST 的相关性等级结构挖掘方法及其对用电行为相关性网络演化测度的有效性与优越性。

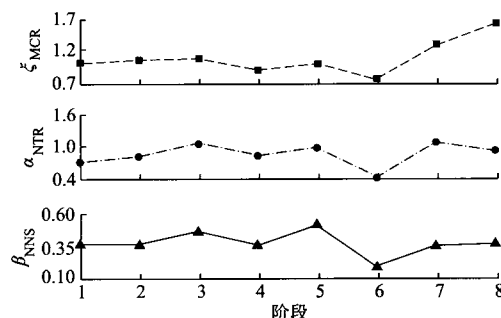


图 6 不同阶段各指标量度结果  
Fig.6 Measure results of indexes at different stages

本文的实例验证虽然采用了 10 kV 馈线数据作为分析对象, 但部分馈线属于大用户直供电, 因而所提网络模型及验证结果同样适用于大规模终端用户间用电行为短时间尺度相关性分析。

## 5 结语

本文提出了用电行为相关性网络化分析方法。首先, 针对用电功率序列相关矩阵中存在的随机用电行为与测量误差等噪声信息, 引入了随机矩阵理论与矩阵谱熵对相关矩阵进行去噪处理。然后, 针对相关矩阵转化后的网络序列, 提出了基于社团结构的相关性聚类方法与基于最小树的相关性等级结构挖掘方法, 并定义了量度用电相关性的拓扑指标, 从而构成完整的相关性分析网络模型。最后, 以 100 条馈线监测数据作为分析对象, 验证了上述去噪方法、相关性分析网络模型及指标的有效性。

所提出的模型可挖掘出短时间尺度下多用户间用电行为相关关系及其连续变化特性, 后续可进一步面向大规模馈线及终端用户数据, 应用于极端天气事件和大型社会活动等强激励因素影响下的用电行为分析和预测、短时间尺度智能配电网(微电网)运行特性分析和配置运行优化、负荷分群分时管理等方面。

附录见本刊网络版 (<http://www.aeps-info.com/aeps/ch/index.aspx>)。

## 参考文献

- [1] 李扬, 王蓓蓓, 李方兴. 灵活互动的智能用电展望与思考[J]. 电力系统自动化, 2015, 39(17): 2-9. DOI: 10.7500/AEPS20150730004.
- LI Yang, WANG Beibei, LI Fangxing. Outlook and thinking of flexible and interactive utilization of intelligent power [J].

- Automation of Electric Power Systems, 2015, 39(17): 2-9. DOI: 10.7500/AEPS20150730004.
- [2] 杨永标, 颜庆国, 王冬, 等. 居民用户智能用电建模及优化仿真分析[J]. 电力系统自动化, 2016, 40(3): 46-50. DOI: 10.7500/AEPS20150924010.
- YANG Yongbiao, YAN Qingguo, WANG Dong, et al. Intelligent electricity consumption modeling and optimal simulations for residential users[J]. Automation of Electric Power Systems, 2016, 40(3): 46-50. DOI: 10.7500/AEPS20150924010.
- [3] 王继业, 季知祥, 史梦洁, 等. 智能配用电大数据需求分析与应用研究[J]. 中国电机工程学报, 2015, 35(8): 1829-1836.
- WANG Jiye, JI Zhixiang, SHI Mengjie, et al. Scenario analysis and application research on big data in smart power distribution and consumption systems[J]. Proceedings of the CSEE, 2015, 35(8): 1829-1836.
- [4] 孙国强, 李逸驰, 卫志农, 等. 智能用互动体系构架探讨[J]. 电力系统自动化, 2015, 39(17): 68-74. DOI: 10.7500/AEPS20150320001.
- SUN Guoqiang, LI Yichi, WEI Zhinong. Discussion on interactive architecture of smart power utilization [J]. Automation of Electric Power Systems, 2015, 39(17): 68-74. DOI: 10.7500/AEPS20150320001.
- [5] 祝恩国, 窦健. 用电信息采集系统双向互动功能设计及关键技术[J]. 电力系统自动化, 2015, 39(17): 62-67. DOI: 10.7500/AEPS20150414002.
- ZHU Enguo, DOU Jian. Design of two-way interaction function based on electricity information acquisition system and its key technology[J]. Automation of Electric Power Systems, 2015, 39(17): 62-67. DOI: 10.7500/AEPS20150414002.
- [6] 李智勇, 吴晶莹, 吴为麟, 等. 基于自组织映射神经网络的电力用户负荷曲线聚类[J]. 电力系统自动化, 2008, 32(15): 66-70.
- LI Zhiyong, WU Jingying, WU Weilin, et al. Power customers load profile clustering using the SOM neural network [J]. Automation of Electric Power Systems, 2008, 32(15): 66-70.
- [7] 张素香, 刘建明, 赵丙镇, 等. 基于云计算的居民用电行为分析模型研究[J]. 电网技术, 2013, 37(6): 1542-1546.
- ZHANG Suxiang, LIU Jianming, ZHAO Bingzhen, et al. Cloud computing-based analysis on residential electricity consumption behavior[J]. Power System Technology, 2013, 37(6): 1542-1546.
- [8] QUILUMBA F L, LEE W J, HUANG H, et al. Using smart meter data to improve the accuracy of intraday load forecasting considering customer behavior similarities[J]. IEEE Trans on Smart Grid, 2015, 6(2): 911-918.
- [9] 郑海雁, 金农, 季聪, 等. 电力用户用电数据分析技术及典型场景应用[J]. 电网技术, 2015, 39(11): 3147-3152.
- ZHENG Haiyan, JIN Nong, JI Cong, et al. Analysis technology and typical scenario application of electricity big data of power consumers[J]. Power System Technology, 2015, 39(11): 3147-3152.
- [10] 阮文骏, 王蓓蓓, 李扬, 等. 峰谷分时电价下的用户响应行为研究[J]. 电网技术, 2012, 36(7): 86-92.
- RUAN Wenjun, WANG Beibei, LI Yang, et al. Customer response behavior in time-of-use price [J]. Power System Technology, 2012, 36(7): 86-92.
- [11] HAIDAR A M, MUTTAQI K M, HAQUE M H. Multistage time-variant electric vehicle load modelling for capturing accurate electric vehicle behavior and electric vehicle impact on electricity distribution grids[J]. IET Generation Transmission & Distribution, 2015, 9(16): 2705-2716.
- [12] LEOU R C, TENG J H, SU C L. Modelling and verifying the load behavior of electric vehicle charging stations based on field measurements[J]. IET Generation Transmission & Distribution, 2015, 9(11): 1112-1119.
- [13] 姜丹丹. 大维随机矩阵谱理论在多元统计分析中的应用[D]. 长春: 东北师范大学, 2010.
- [14] 徐赞新, 王钺, 司洪波, 等. 基于随机矩阵理论的城市人群移动行为分析[J]. 物理学报, 2011, 60(4): 46-52.
- XU Zanxin, WANG Yue, SI Hongbo, et al. Analysis of urban human mobility behavior based on random matrix theory[J]. Acta Physica Sinica, 2011, 60(4): 46-52.
- [15] 魏大千, 王波, 刘涤尘, 等. 高维随机矩阵描述下的量测大数据建模与异常数据检测方法[J]. 中国电机工程学报, 2015, 35(增刊1): 59-66.
- WEI Daqian, WEI Bo, LIU Dichen, et al. A method for WAMS big data modeling and abnormal data detection with large random matrices[J]. Proceedings of the CSEE, 2015, 35(Supplement 1): 59-66.
- [16] 韩华, 吴翎燕, 宋宁宁. 基于随机矩阵的金融网络模型[J]. 物理学报, 2014, 63(13): 431-440.
- HAN Hua, WU Lingyan, SONG Ningning. Financial networks model based on random matrix [J]. Acta Physica Sinica, 2014, 63(13): 431-440.
- [17] 李蓉, 颜平兰, 陈健, 等. 随机矩阵理论在肺癌基因网络识别中的应用[J]. 物理学报, 2009, 58(10): 6703-6708.
- LI Rong, YAN Pinglan, CHEN Jian, et al. Application of random matrix theory to identification of lung cancer gene networks[J]. Acta Physica. Sinica. 2009, 58(10): 6703-6708.
- [18] SENGUPTA A M, MITRA P P. Distributed of singular values for some random matrices[J]. Physics Review E: Statistical Nonlinear & Soft Matter Physics, 1999, 60(3): 3389-3392.
- [19] VASILIKI P, PARAMESWARAN G, BERND R, et al. A random matrix approach to cross-correlations in financial data [J]. Physical Review E: Statistical Nonlinear & Soft Matter Physics, 2001, 65: 18p.
- [20] KENETT D Y, SHAPIRA Y, BEN-JACOB E. RMT assessments of the market latent information embedded in the stocks' raw, normalized and partial correlations[J]. Journal of Probability and Statistics, 2009: 13p.
- [21] 李晓佳, 张鹏, 狄增如, 等. 复杂网络中的社团结构[J]. 复杂系统与复杂性科学, 2008, 5(3): 19-42.
- LI Xiaojia, ZHANG Peng, DI Zengru, et al. Community structure in complex networks [J]. Complex System and Complexity Science, 2008, 5(3): 19-42.
- [22] NEWMAN M E. Fast algorithm for detecting community structure in networks [J]. Physical Review E: Statistical Nonlinear & Soft Matter Physics, 2004, 69(6): 5p.
- [23] NEWMAN M E, GIRVAN M. Finding and evaluating community structure in networks [J]. Physical Review E: Statistical Nonlinear & Soft Matter Physics, 2004, 69: 16p.
- [24] 艾小猛, 文劲宇, 吴桐, 等. 基于点估计和 Gram-Charlier 展开的含风电电力系统概率潮流实用算法[J]. 中国电机工程学报, 2013, 33(16): 16-22.



AI Xiaomeng, WEN Jinyu, WU Tong, et al. A practical algorithm based on point estimate method and Gram-Charlier expansion for probabilistic load flow calculation of power systems incorporating wind power [J]. Proceedings of the CSEE, 2013, 33(16): 16-22.

- [25] 朱承治,李题印,李先锋,等.基于动态时间弯曲和云模型的电能计量动态误差估计[J].电网技术,2015,39(11):3208-3215.  
ZHU Chengzhi, LI Tiyan, LI Xianfeng, et al. Dynamic error estimation for power energy meters based on dynamic time warping and cloud model [J]. Power System Technology, 2015, 39(11): 3208-3215.

陈鹏伟(1992—),男,通信作者,博士研究生,主要研究方向:电能质量、分布式发电、主动配电网。E-mail: chenpw2014@163.com

陶 顺(1972—),女,博士,副教授,主要研究方向:智能配电网、电能质量。E-mail: taoshun@ncepu.edu.cn

肖湘宁(1953—),男,教授,博士生导师,主要研究方向:电能质量、电力电子。E-mail: xxn@ncepu.edu.cn

(编辑 张焱)

## Network Model for Correlation Analysis of Short-term Electricity Consumption Behavior

CHEN Pengwei, TAO Shun, XIAO Xiangning, LI Lu, ZHANG Jian

(State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources (North China Electric Power University), Beijing 102206, China)

**Abstract:** The short-term correlation of electricity consumption behavior is the important decision basis for user meticulous management and intelligent distribution. In this paper, a network method is first proposed to analyze the correlation matrix of electricity sequences. In order to eliminate the effects of noise information from the random behavior and measurement error, random matrix theory is adopted for correlation matrix denoising, and regularization correction exploiting eigenvalue spectrum entropy is also introduced. Then, according to the tiered and peer network sequences constructed from the filtered correlation matrix, a correlation clustering method based on community structure and a hierarchy mining method based on minimum spanning tree are proposed for detailed analysis, including topological indicators to characterize the dynamic evolution of correlation network, as well as the discussion of the model application. Case studies with instance data verify the validity of the proposed denoising method, network analysis model and indicators.

This work is supported by National Natural Science Foundation of China (No. 51207051) and Fundamental Research Funds for the Central Universities (No. 2016XS02).

**Key words:** correlation; random matrix theory; spectral entropy; community structure; minimum spanning tree

(上接第 38 页 continued from page 38)

## Emergency Control Strategy Against Line Overload Based on Virtual Control Units and Heuristic Search

REN Jianwen, HE Peicheng, HE Chen

(State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources (North China Electric Power University), Baoding 071003, China)

**Abstract:** In order to quickly eliminate overload caused by power flow transfer, an emergency control strategy based on virtual control units and heuristic search is proposed. Firstly, the concept of virtual control unit is introduced, i.e. the control unit that can completely eliminate the overload, whose bus power adjustment equal to maximum overload of lines. Then, the two-level search strategy is presented to find the optimal control unit. At the first level, by setting a rational threshold, the bus pairs with stronger load shedding capacity is selected to create the priority control unit set. At the second level, the heuristic searching procedure based on virtual control units is given to obtain the optimal controllable unit within the priority control unit set, where the load shedding effectiveness and load shedding capacity are integrated. If the control unit can not be found in the above threshold range, the first level search scope should be appropriately enlarged until satisfactory control units are found. Verification of power flow is avoided by taking into account the normal line redundancy constraining the adjustment calculated. The power flow of lines and bus power are amended after a tune-up according to sensitivity relations, while the control speed is improved. The simulation on an IEEE 30-bus system has verified the effectiveness of the proposed algorithm.

This work is supported by National Natural Science Foundation of China (No. 50837002).

**Key words:** virtual control unit; heuristic search; cascading overload; emergency control