

DOI: 10.20040/j.cnki.1000-7709.2023.20221390

基于数据融合并行特征提取的调峰电源设备状态评估方法

郝峰¹, 方冰², 祁炜雯², 郭勤慧², 朱传古³, 潘伟峰³

(1. 国网水新部, 北京 100031; 2. 国网浙江省电力有限公司绍兴供电公司, 浙江 绍兴 310000;
3. 国网电力科学研究院有限公司, 江苏 南京 211106)

摘要: 调峰电源设备状态评估有利于电力系统安全稳定运行。系统监测数据本质为一系列指标量的时间序列, 评价系统状态时需综合考虑不同指标量间的耦合关系, 且实时监测系统对模型的处理速度也有所要求。对此, 提出了一种基于数据融合并行特征提取的调峰电源设备状态评估方法。并行提取时间序列统计特征, 基于多重假设检验选择重要特征, 再利用所设计的分层图卷积网络实现信息整合。实验表明, 对比现有模型, 该方法的识别准确率大幅提高, 运行时间缩短, 可迁移性提升。

关键词: 调峰电源设备状态评估; 并行时间序列处理; 图网络; 特征提取; 数据融合

中图分类号: TM712 **文献标志码:** A **文章编号:** 1000-7709(2023)05-0203-04

1 引言

深入分析调峰电源设备状态、开展相关理论研究和实践检验, 是保障电力系统供电可靠性的关键。水电机组设备众多且结构复杂, 其故障模式机理复杂、多种因素耦合, 基于电路或机械模型的常规故障分析难度较大。近年来, 随着各种在线监测装置的安装, 相关案例中的关键数据被逐步收集, 通过大数据挖掘和人工智能算法实现调峰电源设备状态的自动评价成为可能。现有评估调峰电源设备状态的方法主要分为传统算法和机器学习方法两类。前者需要大量人为修正和调整以保证模型性能, 且难以实现复杂系统有效建模; 另一方面, 直接将机器学习算法应用到调峰电源状态评价的研究较少, 而在其他电力设备在线监测领域有不少应用^[1]。但将这些方法直接应用于调峰电源状态评价中存在以下问题: ①调峰电源监测数据为一系列指标量的时间序列, 系统对模型的处理速度有要求; ②不同指标量间存在耦合关系, 评价系统状态时需综合考虑。对此, 本文提出了一种全新的数据融合并行特征提取算法框架, 可并行提取不同指标量序列的统计特征, 利用多重假设检验选择这些特征, 再利用多层图网络有效

整合和修正数据, 该网络分别对设备层和评价指标层进行建模, 在保证运算效率的同时, 充分挖掘不同指标间的关联性; 最终结合常用分类器即可评价调峰电源设备进行状态。试验表明, 对比现有模型, 该方法大幅提升识别调峰电源设备状态评估准确率, 同时算法的可迁移性和运算速度均得到了保障。

2 研究方法

针对调峰电源设备状态评价任务, 设计了基于数据融合并行特征提取的算法框架, 该框架包含评价体系建立, 并行时间序列特征提取及多源特征分层融合三个阶段。

2.1 调峰电源设备状态评价体系

整个监测系统可分为电磁单元、机械单元、控制单元和综合单元, 其中电磁单元为一次设备对应的电磁量, 包含发电机及变压器的局部放电量、发电机组的气隙、磁场强度、铁芯接地电流; 机械单元为水轮机、发电机的机械部分对应状态量, 包含机架和支持盖振动情况、定子铁芯振动情况、上导和水导摆动情况、不同部位水压情况; 控制单元为控制系统对应监测量, 包含机端电压、有功功率、无功功率、水头、机组转速、导叶开度、定子电

收稿日期: 2022-07-03, 修回日期: 2022-10-12

基金项目: 国家电网有限公司总部科技项目(5700-202140381A-0-0-00)

作者简介: 郝峰(1983-), 男, 高级工程师, 研究方向为抽水蓄能智能化, E-mail: feng-hao@sgcc.com.cn

通讯作者: 方冰(1986-), 女, 高级工程师, 研究方向为电力系统及其自动化, E-mail: dianxueke80639576@126.com

流、励磁电流、系统相关控制参数与给定、开关量信息；综合单元则对应上导轴瓦温、定子温度等机组相关温度量，及辅助设备油位、压力等信息。

参考相关评价导则，同时结合历史故障和缺陷记录、巡检记录，综合评估调峰电源设备历史状态，即对不同单元状态进行计分并综合。评分过程如下：①根据历史记录，判断各状态量的状态程度，进行相应扣分；②根据各状态量的扣分值和相应权重系数，计算不同设备单元的得分情况；③根据各单元得分情况，计算设备总体得分并判断设备整体所处状态。权重系数通过层次分析法中获得判断矩阵，即两两比较法分析不同设备的重要性。通过特征值分解、归一化和一致性检验最终得到权重向量。不同分值区间对应不同状态等级，区间 $[0, 75]$ 、 $(75, 85]$ 、 $(85, 95]$ 、 $(95, 100]$ ，四种状态等级分别为正常、注意、异常、严重。

2.2 并行时间序列特征提取

调峰电源设备多层异源状态评价体系的本质是并行时间序列。对于系统中的第 i 台设备，其对应的第 j 种状态量为时间序列 $s_{i,j}(t_1)$, $s_{i,j}(t_2), \dots, s_{i,j}(t_{n_i^{(j)}})$ ，这里 t 表示对应的采样时刻， $n_i^{(j)}$ 表示该状态量的采样长度。时间序列可用矩阵 $S_{i,j}$ 表示，即：

$$S_{i,j} = [s_{i,j}(t_1), s_{i,j}(t_2), \dots, s_{i,j}(t_{n_i^{(j)}})]^T = (s_{i,j,1}, s_{i,j,2}, \dots, s_{i,j,n_i^{(j)}})^T \quad (1)$$

式中， $S_{i,j} \in \mathbf{R}^{n_i^{(j)} \times c}$ ， c 为通道数，即对于标量状态量 $c=1$ ，对于向量状态量 c 等于数据维度。

水电站设备的状态量符合统计规律，即运维人员通过分析统计特征获取设备运行工况信息，判断是否需要更换或维修设备。常用统计特征可视为特定的映射函数，本文所使用的统计量包含最大值、最小值、均值、标准差、偏度、峰度等。由于指标数量较多，特征工程有利于去除干扰，提高模型数据使用效率。对各指标统计量，计算特征相关性、测试重要程度并进行选择。对于分类问题，特征相关性可表述为不同类别下该特征的分布是否存在明显差异，可通过假设检验方法判断。FDR 多重检验策略^[2]为对 p 值序列进行排序，根据错误发现率设定值计算阈值 r_φ ，保留 p 值低于阈值的统计量，其中阈值 r_φ 计算公式为：

$$r_\varphi = \varphi q / \left(n_\varphi \sum_{\mu=1}^q \frac{1}{\mu} \right) \quad (2)$$

式中， φ 为按照 p 值排序后的假设序号； q 为错误发现率设定值； n_φ 为假设总数。

整个计算过程的示意图见图 1。

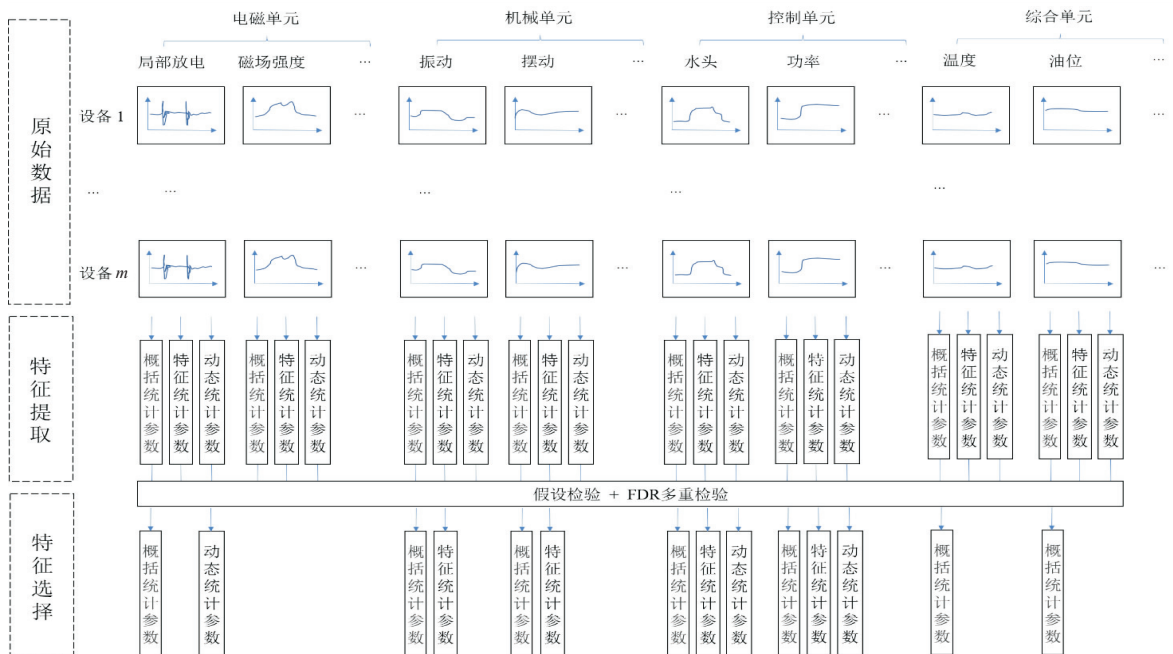


图 1 并行时间序列特征提取过程示意图

Fig. 1 Illustration of parallel time series feature extraction process

2.3 基于图卷积的特征融合

通过构造分层图卷积网络并行时间序列进一步融合特征以挖掘不同设备间、不同参量间的相关性。该网络的第一层为设备层，用于建模不同设备间的联系，第二层为评价指标层，用于建模同

一设备下不同指标间的联系。对于同一设备同一指标下的不同统计量，首先利用多层感知机进行初步融合：

$$f_{mlp1}(s_{i,j}^\varphi) = \mathbf{v}_{i,j} \quad (3)$$

式中， f_{mlp1} 为多层感知机对应的映射函数； $s_{i,j}^\varphi \in$

$\mathbf{R}^{N_{i,j}^e \times c}$ 为状态量序列 $\mathbf{S}_{i,j}$ 对应的统计量序列; $N_{i,j}^e$ 为通过检验的统计量个数; c 为通道数; $\mathbf{v}_{i,j} \in \mathbf{R}^{C_i}$ 为映射后的特征向量; C_i 为映射后的通道数,该参数在同一设备的不同状态量间共享。

此外,对于单一状态量,若单位时间内存在多条记录,单独统计每条记录,再通过池化操作获取最终的统计量序列。对同一设备的不同状态量,利用图卷积网络进行融合,即评价指标层的融合在不同设备间独立进行,可表示为:

$$\mathbf{V}_i^{(L+1)} = f_\sigma(\tilde{\mathbf{L}}^{(i)} \mathbf{V}_i^{(L)} \mathbf{W}_i^{(L)}) \quad (4)$$

式中, $\mathbf{V}_i^{(L)}$ 为第 i 个设备的第 L 层图特征矩阵; $\mathbf{V}_i^{(0)}$ 即经过多层感知机后的输入特征矩阵; f_σ 为激活函数; $\tilde{\mathbf{L}}^{(i)}$ 为该设备的状态量对应图网络的拉普拉斯矩阵; $\mathbf{W}_i^{(L)}$ 为该设备的第 L 层卷积核权重。

对于融合后的设备 i 对应的图特征矩阵 $\mathbf{V}_i^{(n_{L1})}$,首先输入多层感知机映射成向量 \mathbf{u}_i ,并作为第二层设备层图网络节点 i 对应的输入。此时,融合不同设备间特征,并得到整个系统的图特征矩阵 $\mathbf{U}^{(n_{L2})}$ 。同理,利用多层感知机可将整个系统矩阵映射成向量 \mathbf{y} 。此过程可表示为:

$$f_{\text{mlp2}}(\mathbf{V}_i^{(n_{L1})}) = \mathbf{u}_i \quad (5)$$

$$f_{\text{gen}}([\mathbf{u}_1, \dots, \mathbf{u}_i, \dots, \mathbf{u}_m]) = \mathbf{U}^{(n_{L2})} \quad (6)$$

$$f_{\text{mlp3}}(\mathbf{U}^{(n_{L2})}) = \mathbf{y} \quad (7)$$

式中, f_{mlp2} 、 f_{mlp3} 分别为不同多层感知机对应的映射函数; n_{L1} 、 n_{L2} 分别为不同网络层数; f_{gen} 为设备层网络映射函数。

对于融合后的特征采用多层感知机模型分类。图 2 为所提框架流程,其中各阶段的实现思路如前所述,算法整体输入为多源状态量时间序列,输出为设备整体状态。

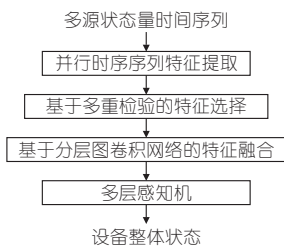


图 2 本文所提算法框架流程图

Fig. 2 Flowchart of the proposed framework

3 实例分析

3.1 数据来源

为验证所提方法的可行性,选取监测系统记录的某地水电站运行数据进行分析。数据收集时

间为 2019 年 7~12 月,设备类型包含发电机组、变压装置、配电装置、辅助系统等。对于不同传感器采集的数据,监测诊断的时间单位统一为 d。监测指标包括:①机组工况。统计不同机组每日有功功率和无功功率;②温度情况。统计不同机组每日定子、冷风、热风温度;③轴承情况。统计不同机组每日上导轴承的最高瓦温、最低瓦温、平均瓦温、最高油温、油槽油位、摆度,推力轴承的最高瓦温、最低瓦温、平均瓦温、最高油温、油槽油位、摆度,水导轴承的最高瓦温、最低瓦温、平均瓦温、最高油温、油槽油位、摆度;④辅助信息。统计不同机组每日供水水压、供水方向和顶盖水位。统计指标主要为最大值、最小值和平均值。根据相关阈值和趋势分析,首先由系统初步打判断各设备状态,再由专家进一步确认及核实。除判断设备状态外,部分异常监测数据的修正工作同样由专家完成。图 3 为一段时间内某发电机组部分数据统计情况示例,包含水导、上导在不同方向上的摆动距离峰峰值及冷风、热风、定子的每日最高温度情况。

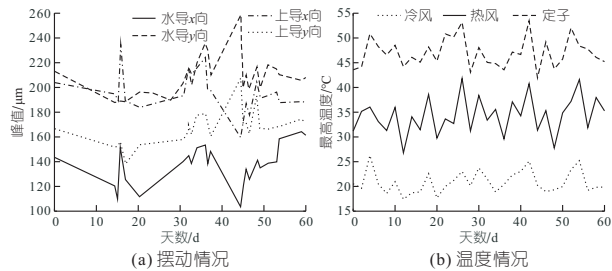


图 3 某发电机组部分数据统计情况示例

Fig. 3 An example of some data statistics results for a generator set

对于设备级状态评价问题,单条案例即为单日某设备对应的不同监测数据记录,经统计该数据集共包含 2 816 条案例,对应时长为 176 d。对于每条案例,根据设备种类及运行情况分为正常、注意、异常、严重四个等级,其中各个类别的数量分别为 1 968、533、240、75 条。不同设备类型中四种评价等级的分布情况见图 4。随机选取 123 d 的数据作为训练集,剩余 53 d 的数据作为测试集。

3.2 试验结果

最佳参数下本文模型的运行结果见表 1,测试时评价指标采用 F1 分数。可以看到,在不同等级中正常状态下模型的准确率最高,这是因为该等级下样本数量最多,模型训练更为充分,后续通过收集相应数据可进一步提升其余等级下的模型性能。此外,在不同设备类型中辅助设备对应的准确率最高,这是因为相较于其他类型,辅助设备评价指标更少,评价标准也更清晰,因此模型识

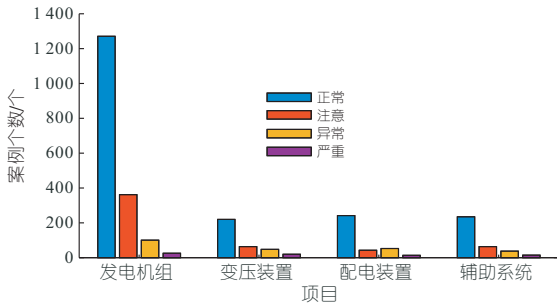


图 4 不同设备类型中四种评价等级的分布情况

Fig. 4 Rating distribution in different equipment types

别难度较低。表 2 为不同模块对性能的提升情况。可以看到,对于识别难度较大的设备类型,并行时间序列特征提取模块、基于 FDR 多重检验的特征选择模块及基于多层图卷积的特征融合模块充分挖掘了对评价指标间关联性进行,大幅提升了识别准确率。

表 1 本文所提方法对应预测准确率

Tab. 1 Predication accuracy of the proposed method

设备类型	F1 分数	状态等级	F1 分数
发电机组	0.860	正常状态	0.950
变压装置	0.872	注意状态	0.929
配电装置	0.885	异常状态	0.912
辅助系统	0.901	严重状态	0.730
整体	0.880	整体	0.880

表 2 不同模块对性能的提升

Tab. 2 Performance improvements by using different modules

项目	子项	特征统计	特征选择	数据融合
设备类型	发电机组	0.126	0.139	0.117
	变压装置	0.098	0.112	0.080
	配电装置	0.105	0.128	0.084
	辅助系统	0.100	0.067	0.050
整体		0.108	0.112	0.083
状态等级	正常	0.070	0.078	0.056
	注意	0.105	0.120	0.091
	异常	0.140	0.152	0.085
	严重	0.119	0.098	0.102
整体		0.108	0.112	0.083

3.3 与现有模型对比

为体现本文所提方法(DFPFE)的优越性,与基于最小二乘曲面的异常状态检测模型^[3](LSS)、基于复杂网络系统的状态评价模型^[4](CNM)和故障树分析模型^[5](FTA)进行对比,见表 3。由表 3 可知,本文所提方法远优于其余三种模型。这是因为本文所提方法在特征提取、选择及融合阶段充分考虑了不同设备、不同指标间的关联性,同时特征提取阶段中不同类型的统计量保证了模型对各指标的全面认知,特征融合阶段中多层图卷积网络又进一步提升了模型的表征能力。此外,如图 5 所示,由于采用并行时间序列特征提取,模型的处理效率大幅提升,这对于实时

表 3 不同模型预测准确率对比

Tab. 3 Comparison of performance of different models

项目	子项	模型			
		LSS	CNM	FTA	DFPFE
设备类型	发电机组	0.725	0.825	0.840	0.860
	变压装置	0.716	0.772	0.778	0.872
	配电装置	0.701	0.800	0.790	0.885
	辅助系统	0.750	0.840	0.860	0.901
整体		0.723	0.809	0.817	0.880
状态等级	正常	0.784	0.860	0.889	0.95
	注意	0.705	0.823	0.810	0.929
	异常	0.681	0.802	0.815	0.912
	严重	0.720	0.750	0.754	0.730
整体		0.723	0.809	0.817	0.880

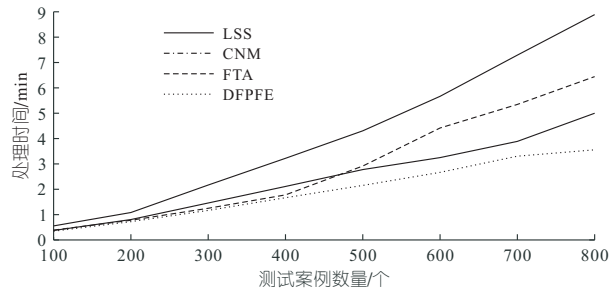


图 5 不同模型处理时间变化曲线

Fig. 5 Processing time varying with number of samples from different models

在线监测意义重大。综上所述,本文所提方法相较于现有模型更适合于多因素耦合、多指标关联等复杂系统下的调峰电源状态评估任务。

4 结论

调峰电源设备状态评估任务中存在数据结构复杂、指标间耦合关系强、算法处理时间长等一系列实际问题,对此提出了一种数据融合并行特征提取算法框架。实验证明,该方法能有效解决复杂系统中的时空性建模问题,大幅提升预测准确率,且借助并行处理架构保证模型运算效率。此外,该框架还可作为数据特征映射模块加入到现有模型中,具有较好的扩展性。

参考文献:

- [1] 胡梓民. 基于机器学习的复杂电力设备故障诊断模型研究[D]. 北京:华北电力大学,2019.
- [2] BENJAMINI Y, DRAI D, ELMER G, et al. Controlling the false discovery rate in behavior genetics research[J]. Behavioural brain research, 2001, 125 (1/2): 279-284.
- [3] 安学利,潘罗平,桂中华,等. 抽水蓄能电站机组异常状态检测模型研究[J]. 水电能源科学,2013,31 (1):157-160.
- [4] 刘海鹏. 大中型水电站设备健康状态评估体系研究[D]. 重庆:重庆大学,2010.
- [5] 杨贤. 面向智能水电站的远程监测与分析系统[D]. 武汉:华中科技大学,2012.

显,故障相电流略有降低,而有功损耗变化显著且易于测量,可将其作为主要状态特征量,为变压器状态监测提供参考。

b. 在轴向方向上,故障股与非故障股均受到向中部的挤压力,内侧股整体受到向下的电动力,外侧股整体受到向上的电动力;在辐向方向上,故障股与非故障股均受到向铁心的压缩力。辐向力与轴向力分布不均,绕组更易失稳变形。应注重低压绕组中部导线换位处的绝缘,并加强绕组端部夹件的紧固力及绕组中部撑条的机械强度,必要情况下,可考虑在两股并联导线之间也增设撑条。

参考文献:

[1] 范慧芳, 咸日常, 咸日明, 等. 改进蜂群算法在大型电力变压器故障诊断中的应用[J]. 水电能源科

学, 2021, 39(9):197-200.

[2] ASADI N, KELK H M. Modeling, analysis, and detection of internal winding faults in power transformers[J]. IEEE transactions on power delivery, 2015, 30(6):2419-2426.

[3] XIAO L, JIAN L, WANG Y, et al. Simulative and experimental investigation of transfer function of inter-turn faults in transformer windings[J]. Electric power systems research, 2014, 107:1-8.

[4] 李智, 咸日常, 李思毛, 等. 一起变压器调压绕组短路故障分析[J]. 变压器, 2021, 58(11): 69-72.

[5] 张立静, 盛戈峰, 侯慧娟, 等. 基于电热特性融合分析的油浸式变压器匝间短路故障辨识方法[J]. 电网技术, 2021, 45(7): 2473-2482.

[6] 张俊杰, 刘兰荣, 刘东升, 等. 用场路耦合模拟变压器线圈部分短路故障电流的方法[J]. 电工技术学报, 2015, 30(20): 65-70.

Characteristics Analysis of Single-strand Inter-turn Short Circuit in Low Voltage Winding Wire Transposition of Distribution Transformer

LI Zhi¹, XIAN Ri-chang¹, CUI Chuan², HU Yu-yao¹, GUO Yang¹, HAO Yan¹, LI Jia-qi¹

(1. College of Electrical and Electronic Engineering, Shandong University of Technology, Zibo 255049, China;

2. State Grid Shandong Zibo Power Supply Company, Zibo 255000, China)

Abstract: In order to ensure that the conductor length of multiple parallel coils of power transformer is consistent and the current density is balanced, the structure transposition is usually carried out in the winding process. The actual operation shows that the insulation at the conductor transposition is easy to damage and leads to inter-turn short circuit fault. Aiming at the problems that the fault point is difficult to set and the electrical quantity cannot be directly detected in the actual test, this paper uses the 'field-circuit' coupling principle to take the distribution transformer with the model of S13-M-200/10, the low voltage winding is wound by two flat copper wires and the transposition point is in the middle of the winding as an example. Based on the finite element simulation software, the simulation analysis model with the same size as the physical structure is established. Through the comparison of the actual test value and the simulation analysis value of the performance parameters, on the basis of verifying the correctness of the model, the electrical characteristics and electro-dynamics of the low voltage winding of the power transformer are studied when the single inter-turn short circuit occurs at the conductor transposition. The results show that when a single inter-turn short circuit occurs in the low voltage winding of the transformer, the loop current of dozens of times the rated current flows inside the short circuit ring, the fault phase current decreases slightly, and the active power loss increases sharply. Conductor transposition causes uneven distribution of axial force and radial force of the coil, and the winding is prone to instability and deformation. It is necessary to take mechanical and insulation strengthening measures at the conductor transposition.

Key words: distribution transformer; low-voltage winding; two side by side; conductor transposition; inter-turn short circuit; electrodynamic force

(上接第 206 页)

Condition Assessment of Peaking Power Source Based on Data Fusion and Parallel Feature Extraction

HAO Feng¹, FANG Bing², QI Wei-wen², GUO Qin-hui², ZHU Chuan-gu³, PAN Wei-feng³

(1. Water and Innovation Department, State Grid Corporation of China, Beijing 100031, China;

2. Shaoxing Power Supply Company of State Grid Zhejiang Electric Power Co., Ltd., Shaoxing 310000, China;

3. State Grid Electric Power Research Institute, Nanjing 211106, China)

Abstract: For the task of condition evaluation of peaking power equipment, the monitoring indicators are defined as a set of time series, among which the complex coupling relationships need to be considered. Besides, requirements on processing time are introduced by the real-time systems. A new framework is proposed to address these problems via the technology of parallel feature extraction as well as data fusion. The time series are statistically analyzed in parallel where multiple hypothesis testing is used to select the important features. Using a defined hierarchical graph convolutional network, the related information is integrated for the final condition evaluation task. Compared with the existing models, experiments indicate that the proposed method with stronger transferability and shorter processing time has a much higher predication accuracy.

Key words: condition assessment of peaking power equipment; parallel time series data processing; graph neural network; feature extraction; data fusion