

基于云平台监控数据的锂离子电池一致性评估

文/张鹏飞¹ 周喜超² 乔雄³

1. 国网定西供电公司; 2. 国网综合能源服务集团有限公司; 3. 国网信通产业集团安徽继远软件有限公司

基于电化学储能监测平台的实时监测数据, 针对储能电站的电池组一致性进行深入分析。首先, 通过电化学储能监测平台对电池的运行数据进行收集和整理, 明确锂离子电池电压、温度和SOC等关键参量, 形成一致性关键参量的表征。其次, 提炼反映电池组一致性的评估特征数据。最后, 研究运用基于DBSCAN (Density-Based Spatial Clustering of Applications with Noise) 聚类算法对异常单体电池进行筛选。实验证明, 该算法能高效获取储能电站运行数据中的关键电气特征量, 准确评估电池组一致性状况, 同时能够迅速定位潜在故障的单体电池。此研究对于提升储能电站运行的稳定性, 实现电池一致性的及时判断和异常单体电池的精准筛选具有显著的实际意义。

在国家大力发展清洁能源, 构建新型电力系统的大背景下, 储能作为平衡电力供需、促进新能源消纳的重要手段, 已经发挥了巨大的作用。锂离子电池作为电池及电池管理的发展核心, 其储能情况得到广泛关注, 亦是目前其大范围推广的掣肘所在。作为一种新型绿色环保电池, 锂离子电池具有比能量高、循环寿命长、输出功率大和可承受较宽充放电倍率等特点^[8]。锂离子电池工作时的热行为对其性能和安全问题有很大的影响^[2], 因此对锂离子电池充放电行为进行监测是十分必要的。

Tian等人^[3]根据电池特征量, 即开路电压、欧姆内阻和极化内阻, 提出了一种基于多特征加权的锂离子电池组在线一致性评价方法。魏刚^[4]基于时域测量原理的电池阻抗谱快速准确测量、阻抗谱建模解析和基于分数阶模型的电池状态评估算法, 提出了一种锂离子电池快速状态评估系统。靳文涛等人^[5]通过分析不同工况下的储能电站电池运行情况, 提取单体电池平均电压、极差电压和电压标准差等参数作为电池组一致性评估特征量, 采用层次

分析法提出储能电池一致性评估方法。王帅等人^[6]通过构建电池模型, 研究电池组放电电压曲线簇, 将影响电池组的一致性特征量分为两类: 初始状态量和状态变化率。其中初始状态量包括初始内阻、初始容量等; 状态变化率包括内阻变化率、容量衰减率等, 并对以上特征量对一致性的影响程度进行了验证。程功^[7]通过电池组循环实验, 分析了15个一致性表征参数的变化特征, 建立表征电池一致性的参数统计特性演化模型, 该模型对预测电池组一致性表征参数具有较高准确性和适应性。

本文选取电化学储能电站的锂离子电池作为研究对象, 记录其充放电行为, 对充放电过程中电压不一致性、温度不一致性、SOC不一致性、电池组和电池簇SOC跳变、电压差、温差以及充放电后电压差进行对比分析; 并在每次充电或放电结束后计算分析充电或放电过程中的数据, 若有异常, 则标记当次的充电或放电过程有异常, 并将异常的结果存入数据表中; 若无异常, 则标记当次的充放电过程无异常。

储能电站数据结构与预处理算法

1. 储能电站电池组数据结构

储能电站的储能系统由多个电池堆组成，每个电池堆包含多个电池簇，且每个电池簇由数百个磷酸铁锂电池串联而成。储能电站配备了电压监测、温度监测和SOC数据监测装置，以实时监测电池堆、电池簇和单体电池的性能参数。

2. 数据预处理

储能电站以15 min为时间间隔存储数据，对所有电池均采集12个电压数据、温度数据和SOC数据，每3 h看作一个数据集，采集电池组内电池的电压、温差和SOC差值。所设计的电池数据预处理算法流程主要包括三个关键环节：文件转码、数据提取和数据填充。首先，对电池数据文件进行转码，将文件均转为utf-8格式，以便后期对数据文件的提取与分析；然后，根据后续分析的需要，提取进行电池组一致性分析时需要的单体电池电压、温度和SOC值。

3. 缺失值处理

在数据采集过程中，由于系统硬件故障和数据信号传输问题，会出现数据点缺失和样本行缺失的现象。因此，需要对数据缺失值进行处理，保证数据片段尽量完整。由于数据缺失方式较多（单个缺失、单行缺失和多行缺失等），需要采取多种方法对应填充。具体的填充策略为：当第*i*行数据、第*j*个字段出现缺失值，首先判断缺失数据行时间是否连续，然后分析缺失时间段（多行缺失）。如果时间数据保持连续性，并且缺失数据时间跨度小于30 s（3行数据），不管出现多少个缺失值（包括整行缺失），都采用插值法填补。具体步骤为：选择缺失值附近10个点（默认使用第*i*-5行至第*i*+5行数据，如果缺失值在数据开始或结束附近），采用牛顿插值法填充；但如果时间不连续，或者缺失时间大于30 s，则采用删除法，直接将整行数据删除。

数据编码

电化学储能电站的数据量巨大、设备繁杂，需要设计高效稳定的处理系统，以便在储能电站端采集和处理海量数据。通过有限的电力带宽向集控平台稳定高效传送重要数据信息，平均1 MW储能单

元约有1万点的数据量，100 MW的储能电站就会有100万点的数据量，储能数据都采用秒级传输，尤其是电芯数据秒级并发一般几十万点规模；储能电站设备种类众多，同一类型的设备，不同的厂家在功能、数据格式和协议之间存在较大差异。

因此，面向动辄百万点且设备繁杂的储能电站，需要高效、易处理和易扩展的建模与传输方式，并且设计高效的数据组织方式，以便在电力网络有限的带宽下以可接受的时延向云平台或其他平台传送数据。方案确定采用消息队列和json/protobuf的通信模式，该模式能够应对秒级大数据传输，并且具备高压缩比、高安全性的数据编码，经过多个百兆瓦储能站点的运行验证，是一种安全高效的模式。

1. 数据传输网络拓扑

储能电站运行数据传输时，首先对在安全I区的储能电站二次系统PCS和BMS进行采集，其中I区的监控系统将采集数据传送至II区数据汇聚服务器，穿过正向隔离装置传送III区的数据计算服务器，数据服务器将汇集的电站所有实时监测数据通过电力专线，穿过防火墙上上传至云平台。拓扑图如图1所示。

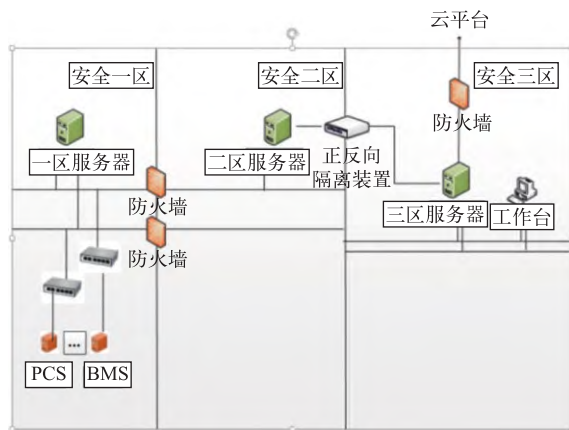
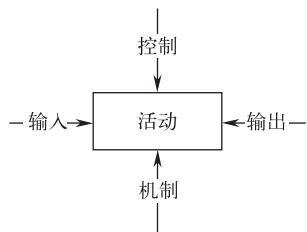


图1 数据传输拓扑图

2. 数据分类

在储能电站数据分类设计中，应用IDEF0建模方法。首先，数据分类采用系统工程建模工具IDEF0，自上而下，面向业务属性，对储能电站进行层次结构识别和划分，最终识别出最小业务单元。IDEF0建模的主要元素是由方盒和箭头组成，如图2所示。方盒代表完成某种功能的活动，

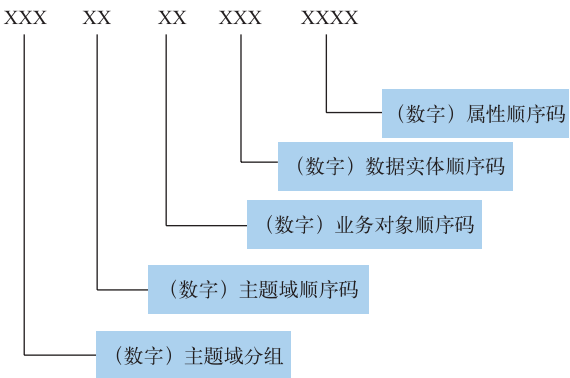
箭头表示活动所需或活动产生的真实信息或对象。IDEF0基本功能活动由活动、控制、机制、输入和输出5个元素组成，控制是该活动执行的机制和条件。



IDEF0方法采用自顶向下逐层分解的方式来构建系统的功能模型。根据数据分类理论依据，面向业务属性，对储能电站业务域进行层次结构识别和划分，将数据确定为包含主题域分组、主题域、业务对象、数据实体和属性5个层次的层次架构，从而确定其类目的隶属与关联关系。

3. 电化学储能电站数据编码的方法

电化学储能电站数据编码是在储能电站数据分类体系的基础上，对储能电站各层级数据进行系统性编码的过程。电化学储能电站数据编码应遵循唯一性、匹配性、可扩充性和简洁性等基本原则，编码方法宜采用层次编码方法。电化学储能电站编码格式如图3所示。



电池组的一致性评价算法

电池组中的单体电池一致性情况影响整个系统的性能和效率、安全风险、电池的寿命以及储能电站的稳定运行。通过定期评估电池的一致性状态，可以及早发现潜在问题，采取相应措施来均衡电池

的使用，提高系统的安全性，延长电池组的寿命，最终降低维护成本，确保储能系统的可靠性和经济性。

1. 电池组数据的特征量选取

首先，储能监测平台每15 min为一个采样点，对所有电池均采集12个电压数据、温度数据和SOC数据，每3 h看作一个数据集合。采集电池组内电池的温差、电压和SOC差值计算公式为

$T_a = (T_1 + T_2 + T_3 + \cdots + T_{12}) / 12$ (1)

$T_{diff}^n = (T_n - T_a) \quad (n = 1, 2, 3, \dots, 12)$ (2)

$V_a = (V_1 + V_2 + V_3 + \cdots + V_{12}) / 12$ (3)

$V_{diff}^n = (V_n - V_a) \quad (n = 1, 2, 3, \dots, 12)$ (4)

$S_a = (S_1 + S_2 + S_3 + \cdots + S_{12}) / 12$ (5)

$S_{diff}^n = (S_n - S_a) \quad (n = 1, 2, 3, \dots, 12)$ (6)

式中， T_n 为电池组在第 n 个采集点时，电池单体的温度； V_n 为电池组在第 n 个采集点时，电池单体的电压； S_n 为电池组在第 n 个采集点时，电池单体的SOC数值； T_{diff}^n 是每个采集节点的单体电池温度和单体电池平均温度的差值； V_{diff}^n 是每个采集节点的单体电池电压和单体电池平均电压的差值； S_{diff}^n 是每个采集节点的单体电池电压和单体电池平均SOC的差值。

2. 一致性评价算法

在计算出每个单体电池的相关数值后，将差值与设定的电压阈值（0.1 V）、温度阈值和SOC阈值进行对比，超过阈值视为差异，最后记录单体对应的数据，并将其保存起来。对于电池组、电池簇SOC跳变判断规则，监测平台会计算电池组或电池簇每分钟的SOC差值，若SOC差值超过阈值5%，则认为发生跳变，将跳变的电池组或电池簇编号存入数据表中，记录异常类型“SOC跳变”（非空的判断）。同时，监测平台还会进行电压差计算和温差计算。即同一个时间点，同一电池簇下单体最高电压和最高温度与单体最低电压和最低温度的差值，若压差超过阈值（0.1 V）或温差超过阈值（10 ℃），则视为电压差超限或温差超限，即会告警。

当通过得到某个电池组的告警时，就需要一种针对电池组内异常电池的筛选算法，对电池组内的单体电池进行逐个分析，找到电池组内表现差的单体电池后，对表现较差的单体电池进行维修或替换，即可有效提高电池组的效率以及安全性。

组内异常单体电池筛选算法

聚类是一种无监督学习方法，可以在没有数据标签的情况下将包含多种样本的集合通过某种内在关系（相似度或距离）分为多个类别，每一个类别称为一个“簇”，簇之间通常是不相干的。聚类的目的是根据样本特征的相似度将行为相近的样本聚集在同一簇中，特征不相似的样本分散在不同簇中。

1. DBSCAN聚类算法

DBSCAN (Density-Based Spatial Clustering of Applications with Noise) 是一种密度聚类算法，用于将数据集中的样本点划分为具有相似特征的簇，并能够有效处理噪声数据。该算法基于样本点周围的密度来识别簇的形状，不需要预先设定簇的数量，对于非凸形状的簇和噪声点的处理相对较好。DBSCAN算法的核心思想是通过定义一个样本点周围的半径和最小样本数，将密度足够大的样本点划分为一个簇。具体来说，算法的步骤如下。

初始化参数：设置两个参数，半径（eps）和最小样本数（min_samples）。

1) 选择未被访问的样本点。从数据集中选择一个未被标记的样本点。

2) 计算半径内的邻域。计算选定样本点到数据集中所有其他样本点的距离，确定距离在半径eps内的样本点。

3) 判断核心点。如果半径内的样本点数量大于等于最小样本数min_samples，则将当前样本点标记为核心点。

4) 扩展簇。将当前核心点所在的簇扩展，将其半径内的所有样本点加入该簇。

5) 标记已访问的样本点。将簇中的所有样本点标记为已访问。

6) 重复过程。重复上述过程，选择下一个未被访问的样本点，直到所有样本点都被访问。

7) 判断噪声点。未被访问的样本点如果不是任何簇的核心点的一部分，将其标记为噪声点。

8) 输出结果。算法执行完毕后，得到样本点的簇划分结果，包括核心点、边界点和噪声点。

2. 特征量选取和聚类分析

假设要判断当天电池组中单体电池的一致性，应选取当天储能电站的电池数据。而对电池进行聚

类分析时，需要选取一个或多个表征电池一致性情况的特征量。特征量的选取直接关系到聚类结果的准确性，所以使用Z标准化方法提取电池数据的特征量。

若一个电池组有 n 个电池，且一天时间内有 p 个数据采集点，即每个电池都有 p 个温度数据、电压数据和SOC数据。使用Z标准化方法计算每个采集点温度数据、电压数据和SOC数据的Z-score值 $V_{z,i}^p$ ，其中 i 为电池的编号， p 为采集点序号。 $T_{z,i}^p$ 、 $V_{z,i}^p$ 和 $S_{z,i}^p$ 代表在采集点采集数据时，此单体电池与电池组内其他电池的离散情况。然后，分别计算每个单体电池在所有采集点的电池温度、电压和SOC的Z-score平均值，即可得到单体电池在当天的温度、电压和SOC的平均Z-score值 $\text{avg}V_{z,i}$ 。

$$\text{avg}T_{z,i} = (T_{z,i}^1 + T_{z,i}^2 + T_{z,i}^3 + \cdots + T_{z,i}^n) / n \quad (7)$$

$$\text{avg}V_{z,i} = (V_{z,i}^1 + V_{z,i}^2 + V_{z,i}^3 + \cdots + V_{z,i}^n) / n \quad (8)$$

$$\text{avg}S_{z,i} = (S_{z,i}^1 + S_{z,i}^2 + S_{z,i}^3 + \cdots + S_{z,i}^n) / n \quad (9)$$

然后将电池组内所有单体电池的温度、电压和SOC的平均Z-score值 $\text{avg}V_{z,i}$ 组成两个特征数据集，使用这两个特征数据集作为两个特征量，通过DBSCAN聚类算法对电池进行二维聚类。聚类图中的横坐标代表单体电池的电压离散情况，纵坐标代表温度离散情况，每一个电池为一个对象点。DBSCAN聚类后，若对象点为离群点，则证明此电池数据与组内其他的电池数据距离较远，即电池可能出现故障，最终故障结果将展示在储能监测平台之上（见图1）。

结束语

本文以电站行为为研究对象，构建了一种创新的储能电站行为模型。通过智能判断电站动作，以及对储能电站充放电行为的精准分析和态势研判，从而实现对电站运行状态的实时监测和异常判断。首先，通过电站行为模型对电站实时监测数据进行整理和收集，确保了对关键参数的准确把握，为后续分析奠定了坚实的基础。其次，考虑到电站数据采集样本庞大，且在不同工况下设备状态趋势的差异，采用数据编码形式对不同工况进行分类，并单独进行分析，有效减少了单次分析的数据量，降低了数据干扰和误差。同时，本研究运用基于DBSCAN (Density-Based Spatial Clustering of

Applications with Noise) 聚类算法, 对异常单体电池进行筛选。该算法不仅高效获取了储能电站运行数据中的关键特征量, 准确评估了电池组一致性状况, 而且能够迅速定位潜在的故障单体电池, 为及时维护和修复故障电池提供了有力的支持, 进一步提高了电站分析的精准性, 有助于预防潜在安全隐患的发生, 同时提升了整个电站的可靠性。EA

参考文献

[1] 程涛, 王欣媛, 王向阳, 等. 轻质、高稳定石墨纸锂离子电池负极集流体[J]. 中国科学技术大学学报, 2017, 47(12): 971-977.

[2] WANG Qingsong, PING Ping, et al. Thermal runaway caused fire and explosion of lithium ion battery[J]. Journal of Power Sources, 2012(208): 210-224.

[3] TIAN J Q, WANG Y J, LIU C, et al.

Consistency evaluation and cluster analysis for lithium-ion battery pack in electric vehicles[J]. Energy, 2020(194): 116944.

[4] 魏刚. 基于阻抗谱的锂离子电池快速状态评估系统设计[D]. 哈尔滨: 哈尔滨工业大学, 2020.

[5] 靳文涛, 李煜阳, 贾学翠, 等. 储能电站电池一致性综合评估方法研究[J]. 电器与能效管理技术, 2022(5): 23-28.

[6] 王帅, 尹忠东, 郑重, 等. 电池模组一致性影响因素在放电电压曲线簇上的表征[J]. 电工技术学报, 2020, 35(8): 1836-1847.

[7] 程功. 电池组一致性的统计特性与变化规律研究[D]. 北京: 北京交通大学, 2017.

[8] 高欣, 王若谷, 高文菁, 等. 基于运行数据的储能电站电池组一致性评估方法[J]. 储能科学与技术, 2023, 12(9): 2937-2945.

(上接第26页)

CapSA-DBN的工业多元负荷分类研究[J]. 河南师范大学学报(自然科学版), 2023, 51(3): 123-130.

[3] 王艳松, 申晓阳, 李强, 等. 基于PCA-GRD-LWR模型的海上油田中长期最大电力负荷预测[J]. 中国石油大学学报(自然科学版), 2023, 47(2): 129-135.

[4] 刘立巍, 周建新, 刘培栋, 等. 基于KNN-LSTM的区域热负荷短期预测及在机组热电可行域的应用研究[J]. 热能动力工程, 2023, 38(3): 91-97.

[5] 李云松, 张智晟. 基于GRU-TGTransformer的综合能源系统多元负荷短期预测[J]. 电力系统保护与控制, 2023, 51(15): 33-41.

[6] 金立, 张力, 唐杨, 等. 考虑温湿指数与耦合特征的综合能源负荷短期预测[J]. 综合智慧能源, 2023, 45(7): 70-77.

[7] 后理通, 张晨晨, 丛意林, 等. Levy-CPSO-GA算法在空调冷负荷预测模型LSTM中的应用[J]. 青岛大学学报(工程技术版), 2023, 38(1): 16-23.

[8] 胡兆龙, 胡俊建, 彭浩, 等. 基于深度随机森林算法的短期用户负荷预测——以金华地区为例[J]. 电子科技大学学报, 2023, 52(3): 430-437.

[9] 王立则, 谢东, 周立峰, 等. 基于Attention机制和ResNet的CNN-BiLSTM短期电力负荷预测模型研究[J]. 南华大学学报(自然科学版), 2023, 37(1): 33-39, 86.

[10] 欧阳福莲, 王俊, 周杭霞. 基于改进迁移学习和多尺度CNN-BiLSTM-Attention的短期电力负荷预测方法[J]. 电力系统保护与控制, 2023, 51(2): 132-140.