

基于 LightGBM 的 蓄电池容量预测方法研究

Research on Battery Capacity Prediction Method Based on LightGBM


杨泽昆¹, 竹梦圆¹, 周明千² (1. 中讯邮电咨询设计院有限公司, 北京 100048; 2. 中讯邮电咨询设计院有限公司郑州分公司, 河南 郑州, 450007)

Yang Zekun¹, Zhu Mengyuan¹, Zhou Mingqian² (1. China Information Technology Designing & Consulting Institute Co., Ltd., Beijing 100048, China; 2. China Information Technology Designing & Consulting Institute Co., Ltd. Zhengzhou Branch, Zhengzhou 450007, China)

摘要:

通过对通信机房中浮充状态的电池的放电电压曲线进行分析, 确定了放电曲线中对容量影响显著的陡降复升段的特征。利用 LightGBM 算法进行训练, 建立了预测容量的数学模型。通过实验对模型进行了验证, 实验结果表明所提出的预测方法能对 10 h 放电条件下的容量做出准确预测, 其平均绝对百分比误差最小可达 8%。同传统的检测方法相比, 该方法所需的测试时间短, 具备广泛的应用前景。

关键词:

铅酸蓄电池; 容量预测; 特征分析; LightGBM
doi: 10.12045/j.issn.1007-3043.2023.07.013
文章编号: 1007-3043(2023)07-0070-05
中图分类号: TN86
文献标识码: A
开放科学(资源服务)标识码(OSID): 

Abstract:

By analyzing the discharge voltage curve of the battery in the floating charge state in the communication room, the characteristics of the "coup de fouet" that have a significant impact on capacity in the discharge curve have been determined. The LightGBM algorithm is used for training, and a mathematical model for predicting capacity is established. The method is validated through experiments, and the experimental results show that the proposed prediction method can accurately predict the capacity under 10 hour discharge conditions, with an average absolute percentage error of at least 8%. Compared with the traditional detection methods, this method requires less testing time and has broad application prospects.

Keywords:

Lead-acid battery; Capacity prediction; Characteristic analysis; LightGBM

引用格式: 杨泽昆, 竹梦圆, 周明千. 基于 LightGBM 的蓄电池容量预测方法研究[J]. 邮电设计技术, 2023(7): 70-74.

0 引言

阀控式铅酸蓄电池被广泛应用在移动通信设施、船舶以及不间断电源中, 其优点是放电时电动势较稳定、工作电压平稳、使用温度及使用电流范围宽、造价较低^[1]。在实际使用过程中, 充放电方式不正确、充电时间不稳定、放电深度过深等, 都会导致电池自身的正常的化学物理变化发生变化, 导致电池容量永久退化^[2]。根据 IEEE 标准 1188-1996 中的规定, 当铅酸蓄电池可用容量逐渐减少到初始容量的 70%~80% 时,

电池失效, 应更换电池^[3]。正确估计蓄电池的容量, 可以避免过度使用造成蓄电池安全事故发生, 同时对用电设备的可靠运行具有重大意义。

在通信行业, 数据中心、机房、基站等通信设备的主要电源系统是 UPS 供电系统、48 V 供电系统。每套供电系统都有若干蓄电池组构成, 蓄电池组的最小组成单位是单体电池。蓄电池保证着市电停电到油机发电之间断电间隔的通信供电, 是保障通信系统正常运行的重要基础设施。不同于电动汽车行业的锂电池处于充放电循环工作模式, 通信系统所用的备用电池长期处于在线浮充工作模式, 其容量损失主要是长时间浮充造成的^[4]。在长时间浮充后, 蓄电池的现存

收稿日期: 2023-05-25

容量决定着断电后可支持的放电时长。目前运营商常采取定时放电和离线放电的方式来判断蓄电池的容量。运营商的运维人员通常根据蓄电池的运行年限和运行状况,每隔1年或者2年进行一次核对性放电试验,以验证电池的优劣^[5]。核对性放电风险大,效率低,既浪费电能又费时费力。核对性放电只能测试整组电池的容量,以容量最低的一节作为整组容量,部分电池由于放电深度不够,其劣化程度还不能完全暴露,因而缺乏对每一节单体电池容量的评估。而完全的离线放电需要拆卸电池,在实验室环境下进行放电测试,成本较大。考虑到目前在实际蓄电池维护中还存在一种维护性放电,即对在线运行的蓄电池进行短时放电,其时间控制在10~20 min,本文详细分析了放电过程中的电压数据,通过机器学习中的回归方法预测剩余容量,从而在线诊断电池的健康状态。

1 蓄电池放电电压曲线研究

当前移动基站等基础设施的蓄电池在线监控系统可以对基站内的后备电源蓄电池进行监测管理,记录单体电池的电压、极柱温度和内阻。在蓄电池浮充状态下,浮充电压、浮充电流和内阻等都处于缓慢变化的状态。在恒流放电过程中,由于电量计量特性,蓄电池的端电压会随着时间逐渐降低。先前的研究表明,铅酸蓄电池在满电状态下搁置一段时间后进行放电时会出现一段短时间的电压跌落,而后又出现电压回升至平台电压,然后进入正常放电阶段^[6]。其中电压陡降复升段发生在转入放电状态的3~5 min之内^[7]。短时间的放电对电源系统的影响较小,而且有利于保持常处于浮充状态的电池极板物质的活性,这使得利用短时放电的特征数据进行容量预测成为可能^[4]。

对现网服役的UPS系统中的12 V/65 Ah电池进行放电测试,放电方式主要为3 h和10 h恒流放电,单体终止电压为9 V,设置蓄电池组的终止电压为288 V。采用600 V/100 A放电机,能够实时记录电池两端的电压电流大小,其中的3 h放电记录间隔达10 s,而10 h放电记录间隔可达1 min,测量结果精确,能够满足实验需求。

图1为在3 h和10 h放电率下,蓄电池组总电压随放电时间的变化。由图1可知,在相同截止电压条件下,蓄电池的实际放电电流越小,电池的电压能维持稳定的时间越长。放电电流越大,电压下降得越快。

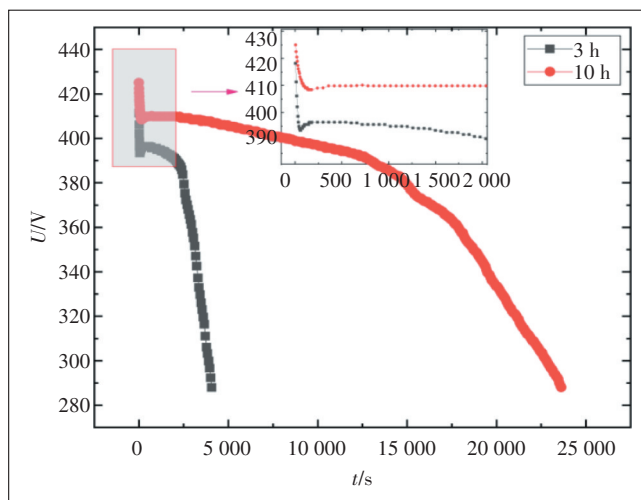


图1 3 h和10 h放电率下的电压曲线

以3 h放电为例,在放电后期蓄电池组端电压几乎直线下降。从化学角度理解,电流大小影响着电化学反应的速率,从而改变着电池的性能参数,因而不同放电率下的容量会存在一定差异。

图2为3 h放电率下组内部分单体电池端电压随时间的变化,在蓄电池组下降到截止电压时,代表性地选取了2类电池,一类电池的单体电压高于9 V(如图2中a),而另一类电池的单体电压低于9 V(如图2中b)。由图2可知,2类电池放电电压曲线的主要区别在后半段,单体电压低于9 V的电池在后半段的短时间内出现了电压的急速下降,导致了过早地到达了截止电压,而且不同单体电池开始出现急速下降的时间和急速下降段维持的时间并不一样。而单体电压高于9 V的电池在后半段电压下降速率增大,但还处于刚开始或者尚未开始的阶段,因而在放电末期还维持着较高的电压位。在这2类电池的放电前段,电压趋势基本一致。

总体上,对比图1和图2,无论是蓄电池组端电压还是单个电池端电压,在放电前期均存在电压陡降复升段。由图1中的局部放大图可知,电流越大,初始电压与陡降复升段电压最小值的差值越大,电压复升又下降的趋势更加明显。铅酸蓄电池的内阻主要包含欧姆内阻和极化内阻,欧姆内阻与电池的尺寸、结构、装配等有关,而极化电阻是电流通过电极时,电极电势偏离平衡电极电势引起的,且极化内阻随电流密度增加而增大。根据铅酸蓄电池的简单等效电路模型 $U=E-IR$,当蓄电池处于浮充电状态时,在线放电电流的增大会导致电压的陡降,理论与实际情况较为相

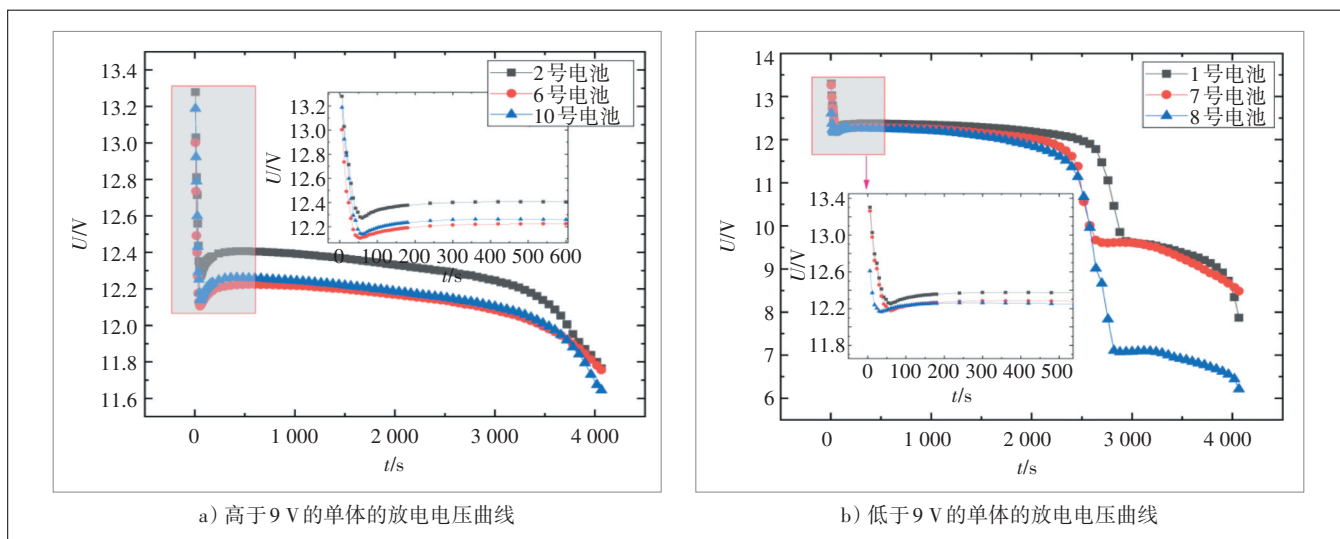


图2 3 h放电率情况下单体电压曲线变化

符。

2 浮充容量在线预测方法

2.1 问题分析

在通信行业,作为后备电源,阀控式铅酸蓄电池的年放电次数相对较少,甚至出现部分电池自安装使用后从未进行过完全放电。除此之外,现网的蓄电池组相邻2次放电过程之间的时间间隔很长,中间甚至会对电池组进行均衡充电,以维持蓄电池组中各单体均处于完全充满的状态。从浮充电压上看,各单体的端电压通常通过电池管理系统进行设置,其浮充电压一般是恒定的,且相差不大。在市电断电后,蓄电池放电时却能观察到部分单体容量下降较快。因而,尽管浮充电压、浮充时间、前次放电深度和温度等因素都会影响到电池容量,但最终的直接表现是放电曲线的变化。相关研究也表明,随着放电电流的减小以及环境温度的升高,陡降复升段的谷底电压和峰值电压将呈近似线性地增大^[8]。放电初期蓄电池电压的陡降复升是比较复杂的,不同的放电方式和环境因素都会对其产生影响,目前尚没有研究表明除谷底电压和峰值电压外的特征对蓄电池容量有影响。因而,可以认为蓄电池的实际可用容量是关于陡降复升段特征的函数,进而可以简化蓄电池容量预测模型。

2.2 陡降复升段特征分析与提取

典型的陡降复升段如图3所示,放电起始时所对应时间点为 t_0 ,在电压陡降复升区内最低点电压为谷底电压,其对应的时间为 t_1 ,复升区间的最大电压为

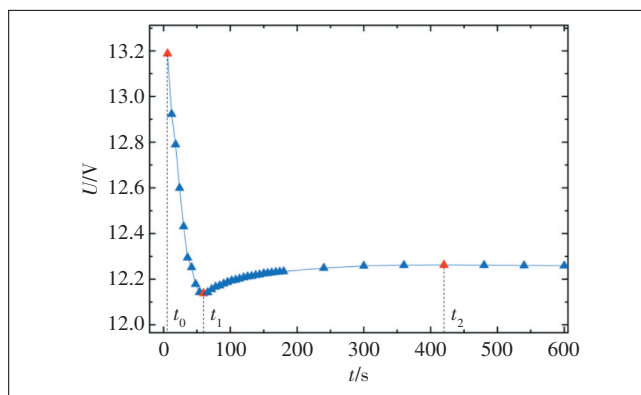


图3 陡降复升段电压变化曲线

峰值电压,其对应的时间为 t_2 。从端电压维度和时间维度进行6个特征的预提取,在端电压维度上包含放电起始电压与谷底电压的差值(ΔU_1)、峰值电压与谷底电压的差值(ΔU_2)以及谷底电压(V_b)3种特征,在时间维度上包含放电起始电压与谷底电压所处时间点之差(ΔT_1)、峰值电压与谷底电压所处的时间点之差(ΔT_2)以及电压陡降段持续的时间占整个陡降复升段持续的时间的比例(ratio),其公式为
$$\text{ratio} = \frac{\Delta T_1}{\Delta T_1 + \Delta T_2}$$
。除此之外,定义电压相对于时间变化量的特征,以陡降段为例,单位时间内端电压的下降为 $\Delta U_1/\Delta T_1$,同理可得,单位时间内电压的上升量可表示为 $\Delta U_2/\Delta T_2$ 。以3 h放电率的数据为例,计算预选特征与不同放电率下放电容量之间的相关性(见表1)。

可见,复升段单位时间内的电压变化量特征、峰

表1 不同放电率下特征与容量的相关性

放电率	ΔU_1	ΔU_2	ΔT_1	ΔT_2	$\Delta U_1/\Delta T_1$	$\Delta U_2/\Delta T_2$	V_b	ratio
3 h	-0.07	0.34	-0.06	0.50	-0.07	-0.24	-0.08	-0.37
10 h	-0.10	-0.17	0.04	0.23	-0.17	-0.23	0.07	-0.17

值电压与谷底电压的电压差值和时间差值、电压陡降段持续的时间占整个陡降复升段持续的的时间的比例与容量具有较好的相关性。特别地,在10 h放电率下,陡降段单位时间内电压变化量特征与放电容量间的相关性也较大。由表1可知,最优特征虽然并不完全一样,但相似性较大,可以取2种放电率下特征的并集作为最终的预选特征。这些特征的获得均来自于陡降复升段,相比较于核容实验更易获取,具备作为表征电池健康状态的参数的潜力。

2.3 LightGBM算法模型

2.3.1 梯度提升决策树

梯度提升决策树(Gradient Boost Decision Tree, GBDT)是一种用于机器学习的回归算法^[9],由GB(gradient boosting)算法和CART二叉回归树组合而成。梯度提升可将弱学习器提升为强学习器,属于重要的集成学习技术。梯度提升方法在迭代的每一步构建一个能够沿着梯度最陡的方向降低损失的学习器来弥补已有模型的不足。理论上,GB算法可以选择各种不同的学习算法作为基学习器,其中用得最多的基学习器是决策树,通常是CART回归树。决策树算法需要更少的特征工程,便于处理特征间的交互关系并且是非参数化的。单独使用决策树算法时,容易出现过拟合问题。但通过限制树的深度等方法来抑制决策树的复杂性,降低单棵决策树的拟合能力^[10],再通过梯度提升方法集成多个决策树,能够缓解过拟合问题。由此可见,梯度提升方法和决策树算法可以取长补短,进而提高精度。

图4为GBDT的训练和测试流程。在训练集上进行多轮迭代,每轮迭代产生一个基模型,每个基模型在上一轮基模型的残差基础上进行训练。基模型满足足够简单、低方差和高偏差的要求,如分类回归树。训练的过程是通过降低偏差来不断提高最终模型的精度。最终的回归模型由每轮训练得到基模型加权求和得到^[11],其公式如下:

$$F_m(x) = \sum_{m=1}^M T(x; \theta_m) \quad (1)$$

其中, M 为基模型的个数, T 为基模型, θ_m 为每个

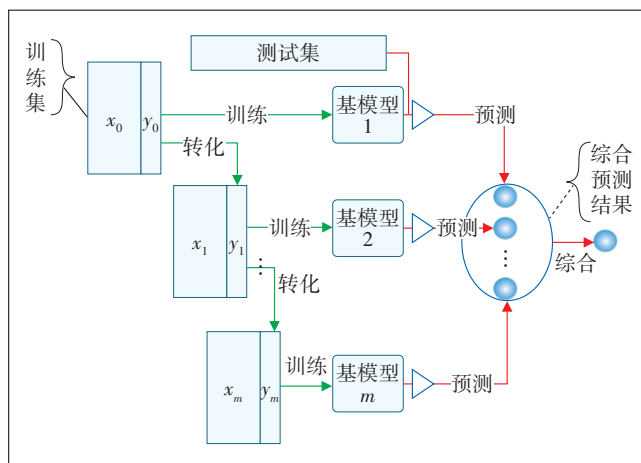


图4 GBDT的训练和测试流程

基模型的可学习参数。

2.3.2 LightGBM

LightGBM是一种对GBDT的优化模型。虽然GBDT算法表现出良好的机器学习效果,但其需要构建一定数量的决策树,划分最优分割点,并对特征值进行排序,随着数据量的几何式增长,GBDT面临着易过拟合、训练速度慢等问题^[12]。LightGBM对特征进行分桶,将连续的浮点特征值离散化成整数,根据特征所在的区间进行梯度累加和个数统计,遍历寻找最优的分割点。特征的离散使得存储方便、运算更快、鲁棒性强、模型更加稳定。使用互斥特征捆绑方法将许多互斥的特征绑定为一个特征,达到降维效果并且降低计算复杂性。采用单边梯度采样算法,在保留大梯度样本的同时,减少大量只具有小梯度的样本,仅用保留的样本计算信息增益,在减少数据和保证精度之间达到平衡^[13]。除此之外,LightGBM直接支持类别特征,特征、数据、投票并行,多项的工程优化大大提高了算法的性能。

3 基于LightGBM的容量预测结果与分析

从基站、机房等收集了现网近300多组蓄电池组的核对性放电数据,基本情况如下:蓄电池组的运行时间为3~10年,单体蓄电池额定容量为200~1 000 Ah,额定电压为2 V和12 V,放电方式为3 h、10 h恒流放电。统计分析在不同放电倍率下陡降复升段持续的时间,其中3 h放电情况下的短时放电时间为7~10 min,而10 h放电情况下的短时放电时间为23~30 min。从短时放电过程中提取陡降复升段,形成8个预选特征,并建立LightGBM容量预测模型。特征归一化能够

使机器学习模型快速收敛,同时可以避免输入数据量纲不一致的问题。预选特征归一化计算如下:

$$y = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (2)$$

容量归一化计算如下:

$$C_{\text{nor}} = \frac{C_o}{C_s} \quad (3)$$

其中, C_o 为原始计算出来的放电容量的数据, C_s 为相应放电倍率下电池的额定容量。针对回归问题,常用的评价指标为平均绝对误差(MAE)和平均绝对百分比误差(MAPE)。MAE可以避免计算误差时相互抵消的问题,能够准确反映实际预测误差的大小;MAPE常用于衡量预测的准确性,越接近0则预测误差越小^[14]。

本文采用LightGBM分别对额定电压为2 V和12 V的蓄电池容量进行回归预测,对同一额定电压和放电倍率的所有单体电池计算MAE和MAPE,其实验结果如表2所示。在相同的10 h放电率的条件下,额定电压为2 V和12 V的蓄电池容量预测的MAPE分别为8%和11%。在3 h放电率的情况下,MAE在47 Ah左右,MAPE会高一些,在额定电压为2 V时,MAPE较10 h放电增加了31%,但考虑到只使用了放电开始后7~10 min的电压数据已经能够达到40%的MAE,说明了所提出的方法的潜力。在额定电压为12 V的情况下,3 h容量预测的MAPE小于2 V蓄电池大约13个百分点,达到了不错的精度。综合上述结果来看,所提出的预测方法可作为判断蓄电池剩余容量的参考。

表2 容量预测结果表

放电率	2 V		12 V	
	MAE/Ah	MAPE	MAE/Ah	MAPE
3 h	47.50	0.39	48.60	0.26
10 h	44.40	0.08	61.30	0.11

4 结语

本文提出了基于LightGBM的蓄电池容量预测方法,该方法可以在线检测蓄电池健康状态。相比较于其他的电池容量检测方法,本文提出的方法基于蓄电池的浅放电过程,大大降低了对电源系统可靠性的影响。除此之外,详细分析了放电过程中陡降复升段的电压特征,通过相关性分析,考虑不同特征对放电容量的影响。采用基于LightGBM的机器学习模型预测

不同放电情况下的放电容量,并在现网收集的蓄电池数据上进行验证性实验,在10 h放电的情况下,MAPE可控制在10%左右。本文提出的方法容易实现容量在线监测,且不会对电池造成不可逆的损害,对确保蓄电池的安全稳定运行具有重大意义。

参考文献:

- [1] HURLEY W G, WONG Y S, WOLFLE W H. Self-equalization of cell voltages to prolong the life of VRLA batteries in standby applications [J]. IEEE Transactions on Industrial Electronics, 2009, 56 (6): 2115-2120.
- [2] SAWAI K, TSUBOI Y, SHIOTA M, et al. Corrosion of Pb-Ca-Sn alloy during potential step cycles [J]. Journal of Power Sources, 2008, 175(1):604-612.
- [3] 葛礼嘉, 宋政湘, 张国钢. 变电站用阀控式铅酸蓄电池浮充寿命研究 [J]. 电力电容器与无功补偿, 2020, 41(6):191-195, 201.
- [4] HE Q, ZHA Y B, SUN Q, et al. Capacity fast prediction and residual useful life estimation of valve regulated lead acid battery [J]. Mathematical Problems in Engineering, 2017, 2017: 7835049.
- [5] 汪昌元, 罗群, 徐晓峰. 一种铅酸蓄电池剩余容量的预测方法 [J]. 广西电力, 2020, 43(4):60-62.
- [6] 袁世魁. 阀控式铅酸蓄电池性能在线诊断方法的研究 [D]. 南京: 东南大学, 2018.
- [7] 冯真得. 变电站用VRLA蓄电池SOC及SOH估计算法研究 [D]. 合肥: 合肥工业大学, 2016.
- [8] 袁世魁, 程力. 基于Coup de fouet现象的蓄电池SOH估测 [J]. 蓄电池, 2018, 55(2):65-68.
- [9] FRIEDMAN J H. Greedy function approximation: a gradient boosting machine [J]. The Annals of Statistics, 2001, 29(5): 1189-1232.
- [10] 翟冠强. 基于决策树理论的大规模风电场输出功率超短期预测研究 [D]. 长春: 东北电力大学, 2018.
- [11] 肖迁, 焦志鹏, 穆云飞, 等. 基于LightGBM的电动汽车行驶工况下电池剩余使用寿命预测 [J]. 电工技术学报, 2021, 36(24): 5176-5185.
- [12] 崔逊航. 基于改进LightGBM的风力发电机组叶片故障预测算法研究 [D]. 重庆: 重庆邮电大学, 2021.
- [13] KE G L, MENG Q, FINLEY T, et al. LightGBM: a highly efficient gradient boosting decision tree [C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach, California, USA: Curran Associates Inc., 2017: 3149-3157.
- [14] 许隽杰, 陈永保, 章立新, 等. 基于LightGBM算法的短期电力负荷预测研究 [J]. 建模与仿真, 2022, 11(4): 1071-1082.

作者简介:

杨泽昆, 助理工程师, 硕士, 主要从事计算机视觉相关算法工作; 竹梦圆, 工程师, 硕士, 主要从事计算机视觉、视频AI领域的算法研发工作; 周明千, 高级工程师, 硕士, 主要从事通信电源技术研究和双碳能源支撑工作。