

# 控制与决策

Control and Decision

基于HI-DD-AdaBoost. RT的锂离子动力电池SOH预测

田慧欣, 秦鹏亮, 李坤, 王红一

引用本文:

田慧欣, 秦鹏亮, 李坤, 等. 基于HI-DD-AdaBoost.RT的锂离子动力电池SOH预测[J]. 控制与决策, 2021, 36(3): 686–692.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2019.0764>

---

## 您可能感兴趣的其他文章

Articles you may be interested in

[基于矩阵的双论域模糊概率粗糙集增量更新算法](#)

Incremental updating of fuzzy probability rough sets over two universes based on matrix method

控制与决策. 2021, 36(3): 553–564 <https://doi.org/10.13195/j.kzyjc.2019.0692>

[一种基于深度学习的时间序列预测方法](#)

A time series prediction method based on deep learning

控制与决策. 2021, 36(3): 645–652 <https://doi.org/10.13195/j.kzyjc.2019.0809>

[基于多维泰勒网的超前d步预测模型](#)

d-step-ahead predictive model based on multi-dimensional Taylor network

控制与决策. 2021, 36(2): 345–354 <https://doi.org/10.13195/j.kzyjc.2019.0722>

[参数未知的离散系统Q-学习优化状态估计与控制](#)

Q-learning optimal state estimation and control for discrete systems with unknown parameters

控制与决策. 2020, 35(12): 2889–2897 <https://doi.org/10.13195/j.kzyjc.2019.0180>

[基于强化学习的小型无人直升机有限时间收敛控制设计](#)

Finite time control based on reinforcement learning for a small-size unmanned helicopter

控制与决策. 2020, 35(11): 2646–2652 <https://doi.org/10.13195/j.kzyjc.2019.0328>

# 基于 HI-DD-AdaBoost.RT 的锂离子动力电池 SOH 预测

田慧欣<sup>1,2†</sup>, 秦鹏亮<sup>1,2</sup>, 李坤<sup>3</sup>, 王红一<sup>1,2</sup>

(1. 天津工业大学 电气工程与自动化学院, 天津 300387; 2. 天津工业大学 电工电能新技术天津重点实验室, 天津 300387; 3. 天津工业大学 经济与管理学院, 天津 300387)

**摘要:** 锂离子电池是一个复杂的电化学动态系统, 实时准确的健康状态 (SOH) 估计对电动汽车动力锂电池的维护至关重要, 传统建模方法难以实现 SOH 的在线估算. 基于此, 从实时评估电池的 SOH 出发, 在增量学习的基础上, 选取与电池健康状态相关的指标建立 SOH 预测模型. 考虑到增量学习中的耗时性问题, 提出融合滑动窗口技术的 HI-DD 算法, 该算法可以检测概念漂移是否发生, 从而指导和确定模型更新位置; 设计出 HI-DD 与 AdaBoost.RT 结合的模型更新策略, 进而提高模型的在线学习性能和预测精度, 最后使用 CALCE 提供的电池老化实验数据对所提出的方法进行验证. 结果表明, 基于增量学习的 HI-DD-AdaBoost.RT 预测算法具有较强的在线更新能力和较高的预测精度, 能够满足 SOH 在线预测的实际需求.

**关键词:** 锂离子动力电池; SOH; 增量学习; HI-DD; 概念漂移; AdaBoost.RT

中图分类号: TM714

文献标志码: A

DOI: 10.13195/j.kzyjc.2019.0764

开放科学(资源服务)标识码(OSID):



**引用格式:** 田慧欣, 秦鹏亮, 李坤, 等. 基于 HI-DD-AdaBoost.RT 的锂离子动力电池 SOH 预测 [J]. 控制与决策, 2021, 36(3): 686-692.

## Prediction of Li-ion battery SOH based on HI-DD-AdaBoost.RT

TIAN Hui-xin<sup>1,2†</sup>, QIN Peng-liang<sup>1,2</sup>, LI Kun<sup>3</sup>, WANG Hong-yi<sup>1,2</sup>

(1. School of Electrical Engineering and Automatic, Tianjin Polytechnic University, Tianjin 300387, China; 2. Key Laboratory of Advanced Electrical Engineering and Energy Technology, Tianjin Polytechnic University, Tianjin 300387, China; 3. School of Economics and Management, Tianjin Polytechnic University, Tianjin 300387, China)

**Abstract:** Real-time and accurate estimation of the state of health (SOH) is of great importance for the maintenance of Li-ion battery in electric vehicles. Li-ion batteries is a complex electrochemical dynamic system, so traditional modeling methods are difficult to achieve online estimation of SOH. Therefore, starting from real-time assessment of SOH of battery, this paper chooses indicators related to the state of battery health and establishes a SOH prediction model based on incremental learning. A HI-DD algorithm based on sliding window technology is proposed for considering the time-consuming problem of online learning in incremental learning, which can detect whether concept drift occurs, thus guiding and determining the updating position of the model. Furthermore, the model updating strategy based on HI-DD and AdaBoost.RT is designed to improve the online learning performance and prediction accuracy of the model. Finally, the proposed method is verified by using the experimental data of battery aging provided by CALCE. The results show that the HI-DD-AdaBoost.RT prediction algorithm based on incremental learning has strong updating ability of online and high accuracy of prediction, which can meet the actual needs of SOH online prediction.

**Keywords:** Li-ion battery; SOH; incremental learning; HI-DD; concept drift; AdaBoost.RT

## 0 引言

锂离子电池作为电动汽车的重要组成部分, 由于其具有能量密度高、成本价格低、环保无污染等优点而受到人们的广泛青睐<sup>[1]</sup>. 然而在实际的使用过程中, 电池的健康状态会因外部环境、过充电、自放电

等因素严重影响电动汽车行车安全与续航里程<sup>[2]</sup>. 因此, 电池的健康状态预测对于保障电动汽车运行的动力性、经济性和安全性具有重要的意义.

目前, SOH (state of health) 的预测可分为基于模型的预测和基于数据驱动的预测. 基于模型的预测

收稿日期: 2019-05-31; 修回日期: 2019-08-30.

基金项目: 国家自然科学基金项目(71602143, 51806150, 62073067); 天津科技特派员项目(18JCTPJC62600); 天津市自然科学基金项目(18JCYBJC22000, 18JCQNJC04400); 天津市高等学校创新团队培养计划项目(TD13-5038).

†通讯作者. E-mail: icedewl@163.com.

首先需要建立一个电池模型,常见的电池模型有等效电路模型<sup>[3]</sup>、电化学反应模型<sup>[4]</sup>、经验模型<sup>[5]</sup>等,虽然基于电池模型的方法可以实现SOH的预测,但存在参数辨识复杂、计算量大等问题.数据驱动算法如神经网络、支持向量机、高斯回归等,因无需充分了解电池内部复杂的反应机理被国内外学者广泛应用于SOH的预测中,但是这些数据驱动算法预测精度难以提高且模型泛化能力差,而集成学习算法<sup>[6]</sup>通过构建并结合多个数据驱动模块提高了预测精度并可以防止过拟合问题的出现,因此可以很好地解决传统数据驱动算法中存在的问题.

除此之外,SOH的预测算法大都在离线方式下进行,而实时的SOH预测对于电动汽车安全行驶具有重要的意义.增量学习策略为SOH的实时预测提供了有效途径,但是传统的增量学习方法会增加学习的时间且学习效率低,即传统的增量学习方法无法解决数据流中发生的概念漂移问题,如果使用单一模型进行学习并不断更新将会耗费大量的时间,因此在线学习中,概念漂移检测<sup>[7]</sup>成为一个重要的任务.

上述研究在SOH的预测中取得了一定成果,但SOH预测方法仍然存在着精度难以提高、学习时间长、模型更新效率低等瓶颈问题.基于此,本文提出HI-DD(drift detection based on hoeffding inequality)算法实时检测增量学习过程中发生的概念漂移,从而确定与指导增量学习中模型更新的位置,减少模型学习的时间.同时,为提高SOH的预测精度,采用AdaBoost.RT集成学习算法并加入时间遗忘因子,建立当前时刻SOH预测的总模型,从而实现动力电池SOH预测模型的实时更新功能.最后,通过电池老化实验数据验证本文方法的有效性.

## 1 SOH定义与预测模型

SOH是用来表征电池健康状态的,所以SOH可以从不同方面去定义.目前比较常见的是从电池的最大可用容量、电池内阻等角度定义SOH,本文使用最大可用容量来表征SOH,其表征公式如下所示:

$$SOH = C_m / C_n. \quad (1)$$

其中: $C_m$ 是当前测量得到的电池容量; $C_n$ 是电池标称容量,也可以理解为出厂电池容量.

随着电池的循环使用,电池的相关参数会随着电池健康状态的变化而发生改变.本文以CS2-36型号电池为研究对象,电池的具体参数如表1所示.根据马里兰大学CALCE(center advanced life cycle engineering)提供的实验数据<sup>[8]</sup>对电池的健康状况进行分析,选取电池的最大可用放电容量、最大充电量、

最大放电量、内阻、每次充放电静置后的电池端电压、静置后的电池电流、恒流充电时间、恒流放电时间这8个关键指标,同时考虑到电池老化过程中SEI膜的生成和内阻会不断上升的问题,又构造出容量再生的电压上升幅度,总共9个指标来评估电池的SOH.

表1 CS2-36型号电池参数

参数	
额定容量/mAh	1 100
重量/g	21.1
尺寸	5.4 × 33.6 × 50.6
负极材料	LiCoO <sub>2</sub>
正极材料	石墨

指标选定后,在增量学习的基础上,结合数据驱动技术建立SOH预测模型,将选取的9个指标作为模型的输入,以最大可用容量定义的SOH作为模型的输出,从而实现动力锂离子电池SOH的预测.

## 2 基于增量学习的HI-DD-AdaBoost.RT预测算法

为满足SOH的在线预测需求,在增量学习的基础上进行SOH的在线预测,由于SOH在线预测相关数据的分布与模式时时刻刻发生着变化,即SOH数据中会有概念漂移的发生,如果某时刻数据与前一时刻数据之间未发生概念漂移,即两数据窗口间数据隐藏的分布模式和所选参数与SOH值之间关联程度并没有发生改变,则无论是否将当前时刻窗口的数据加入模型进行更新,预测效果并不会发生较大变化,即加入概念漂移的检测后算法的预测稳定性较高.因此,在增量学习的基础上加入概念漂移的检测,提出Hoeffding不等式和滑动窗口结合的HI-DD算法,通过其检测结果指导和确定模型更新的位置;同时,考虑到算法的预测精度问题,结合AdaBoost.RT集成算法,提出一种具有历史数据遗忘能力的HI-DD-AdaBoost.RT在线学习算法.

### 2.1 基于Hoeffding不等式与滑动窗口结合的概念漂移检测算法HI-DD

考虑到增量学习中存在的耗时性问题和SOH数据中有概念漂移的发生,本文提出一种滑动窗口与Hoeffding不等式结合的HI-DD算法进行概念漂移的检测,通过检测结果简化增量学习中在线学习与模型更新的过程.

#### 2.1.1 Hoeffding不等式

Hoeffding不等式刻画的是某个事件的真实概率与不同的伯努利试验中观察到的频率之间的差异,在使用机器学习算法时,通过建立抽象模型来模拟真实数据中所隐含的规律<sup>[9]</sup>,其描述如下所示.

设  $X_1, X_2, X_3$  是取值范围在  $[a_i, b_i]$  上的  $n$  个独立的随机变量, 可得到一般的 Hoeffding 不等式

$$P(\bar{x} - E[\bar{x}] \geq t) \leq \exp\left(-\frac{2n^2t^2}{\sum_{i=1}^n (b_i - a_i)^2}\right). \quad (2)$$

特殊地, 若  $X_1, X_2, \dots, X_n$  为  $[0, 1]$  的独立随机变量, 即  $0 \leq x_i \leq 1$ , 则有

$$P(\bar{x} - E[\bar{x}] \leq t) \leq \exp(-2nt^2). \quad (3)$$

其中:  $\bar{x} = \frac{1}{n}(X_1 + \dots + X_n)$ ,  $\bar{x}$  为经验均值,  $E(\bar{x})$  为真实的均值.

### 2.1.2 加入滑动窗口机制的 HI-DD 算法

滑动窗口技术在在线学习中发挥了很重要的作用, 滑动窗口模型随着数据样本的到达, 新的数据不断插入到滑动窗口中, 如图 1 所示, 有效地提高了模型的实时性和准确性.

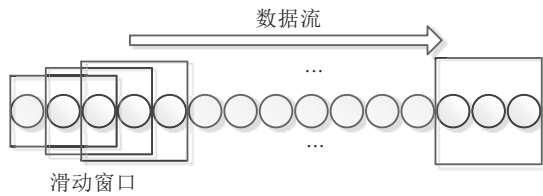


图 1 滑动窗口示意

在 Hoeffding 不等式及其理论的基础上, 结合滑动窗口进行概念漂移的检测. 设当前窗口数据为  $S_2$ , 前一时刻的窗口数据为  $S_1$ , 当前时刻与前一时刻的窗口数据集合为  $S$ , 接下来对两个窗口发生概念漂移时的阈值进行推导.

设窗口  $S_1$  内的数据为  $X_1$ , 样本均值为  $\lambda_1$ ,  $S_2$  内的数据为  $X_2$ , 样本均值为  $\lambda_2$ , 总窗口  $S$  内的数据为  $X$ , 样本均值为  $\lambda$ , 作零假设  $H_0$ , 即

$$E(\lambda_1) = E(\lambda_2) = E(\lambda), \quad (4)$$

$$D(\lambda_1) = D(\lambda_2) = D(\lambda). \quad (5)$$

根据上述所作的零假设, 两个窗口的数据分布相同且没有发生概念漂移, 为判断  $S_1$  窗口与  $S_2$  窗口是否有显著性差异, 构造检验统计量  $\lambda_1 - \lambda_2$ , 设置信度为  $\delta$ , 则有

$$P[(\lambda_1 - \lambda_2) - E(\lambda_1 - \lambda_2) \geq t] \leq \delta. \quad (6)$$

为消除两窗口数据量纲不同引起的误差, 对式 (6) 进行中心标准化, 即

$$P\left[\frac{(\lambda_1 - \lambda_2) - E(\lambda_1 - \lambda_2)}{\sqrt{D(\lambda_1 - \lambda_2)}} \geq t\right] \leq \delta. \quad (7)$$

对上式进行化简, 得到

$$P\left[(\lambda_1 - \lambda_2) \geq t\sqrt{\frac{D(X_1)}{n_1} + \frac{D(X_2)}{n_2}}\right] \leq \delta. \quad (8)$$

由 Hoeffding 的一般不等式, 将式 (2) 代入 (8), 得

$$P\left[(\lambda_1 - \lambda_2) \geq t\sqrt{\frac{D(X_1)}{n_1} + \frac{D(X_2)}{n_2}}\right] \leq 2 \exp\left(-\frac{2n^2t^2\left(\frac{D(X_1)}{n_1} + \frac{D(X_2)}{n_2}\right)}{\sum_{i=1}^n (b_i - a_i)^2}\right), \quad (9)$$

即有

$$2 \exp\left(-\frac{2n^2t^2\left(\frac{D(X_1)}{n_1} + \frac{D(X_2)}{n_2}\right)}{\sum_{i=1}^n (b_i - a_i)^2}\right) = \delta, \quad (10)$$

则所求阈值为

$$t = \sqrt{\frac{\ln\left(\frac{2}{\delta}\right)\left(\sum_{i=1}^n (b_i - a_i)^2\right)}{2n^2\left(\frac{D(X_1)}{n_1} + \frac{D(X_2)}{n_2}\right)}}. \quad (11)$$

其中:  $n_1, n_2$  分别为  $S_1$  和  $S_2$  数据窗口中的数据个数,  $t$  为两窗口发生概念漂移的阈值,  $D(x)$  为数据窗口  $S$  的方差.

另外, 与上述求导阈值进行比较的检验统计量  $er$  为

$$er = \frac{|\lambda_1 - \lambda_2|}{\sqrt{\frac{D(x_1)}{n_1} + \frac{D(x_2)}{n_2}}}. \quad (12)$$

因此, 根据  $er$  与  $t$  的比较结果, 若

$$er \geq t, \quad (13)$$

则发生了概念漂移, 否则未发生概念漂移.

根据以上描述, 所提出的 HI-DD 算法的基本原理如图 2 所示.

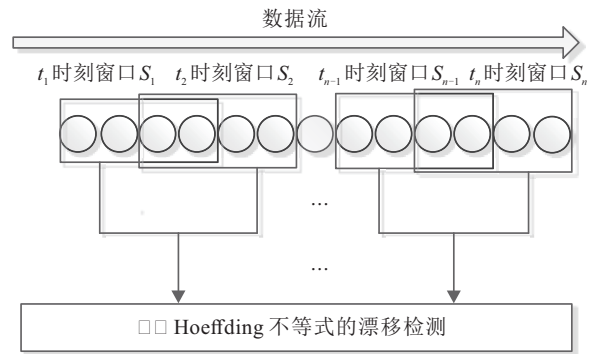


图 2 HI-DD 算法原理

根据上述原理, HI-DD 算法的具体步骤如下.

#### 算法 1 HI-DD 概念漂移检测算法.

输入: 数据集  $num = (a_1, a_2, \dots, a_n)$ , 滑动窗口集合  $S = []$ , 长度  $size_k$ , 滑动窗口每次向前滑动的长度  $u$ ;

输出: 发生概念漂移的滑动窗口数据集合  $B$ .

1. for  $i$  in  $n$  do
  - $S = [S; \text{num}(i)];$
2. if  $\text{length}(S) > \text{size}_k$  and  $(\text{length}(S) - \text{size}_k)/u = 1$  then
3. //维持窗口长度不变
4. for  $q$  in  $u$  do
  - $S(1) = [];$
5. end for
6. //设当前窗口数据为  $S_2$ , 前一窗口数据为  $S_1$ , 两窗口的集合为
  - $SS = [S_1; S_2].$
7. //找到当前窗口  $SS$  中的最小值  $a_i$  和最大值  $b_i$ , 并计算当前窗口构造的检验统计量  $er$  和发生概念漂移的阈值  $t$ .
8. if  $er \geq t$  then
9. //发生了概念漂移
  - $B = SS;$
10. else
  - $B = [];$
11. end if
12. end if
13. end for

所提 HI-DD 算法巧妙地结合了 Hoeffding 一般不等式, 适用于任何数据的概念漂移检测, 同时在 Hoeffding 一般不等式的基础上融入了滑动窗口技

术, 滑动窗口应用在在线学习中, 可以降低在线学习的复杂度, 两者融合可以有效地检测到概念漂移的发生, 从而可以依据检测结果进行模型的更新, 并且可以在预测精度稳定的情况下减少学习的时间, 满足在线学习的需要。

### 2.2 AdaBoost.RT与HI-DD结合的SOH预测算法

基于上述所提出的 HI-DD 算法, 使用滑动窗口策略进行数据的实时分块, 并对所分相邻数据块之间进行概念漂移的实时检测, 如果检测到发生概念漂移, 则说明该 SOH 的数据模式发生了变化, 需要对检测到的数据块进行建模并更新总模型, 如果没有检测到概念漂移的发生, 则保留原模型无需进行更新。另外, 针对单一的数据驱动算法预测精度低且容易出现过拟合的问题, 本文采用集成学习算法对数据块进行建模预测。集成学习具有强大的学习能力并且有很高的精确度, 其通过集成策略将多个子学习机融合, 达到提高学习性能及预测精度的目的。AdaBoost 是集成学习中的主要分支之一, Solomatine 等<sup>[10]</sup> 提出的 AdaBoost.RT 算法在回归问题的应用具有较好的性能, 具有强大的学习能力, 所以选取 AdaBoost.RT 算法对检测到的数据块进行建模并预测 SOH。在模型更新过程中考虑到历史数据和实时数据对现有模型的重要程度不同, 引入时间遗忘因子并结合模型的误差率共同指导模型的更新, 因此根据上述描述, 该算法的基本原理如图3所示。

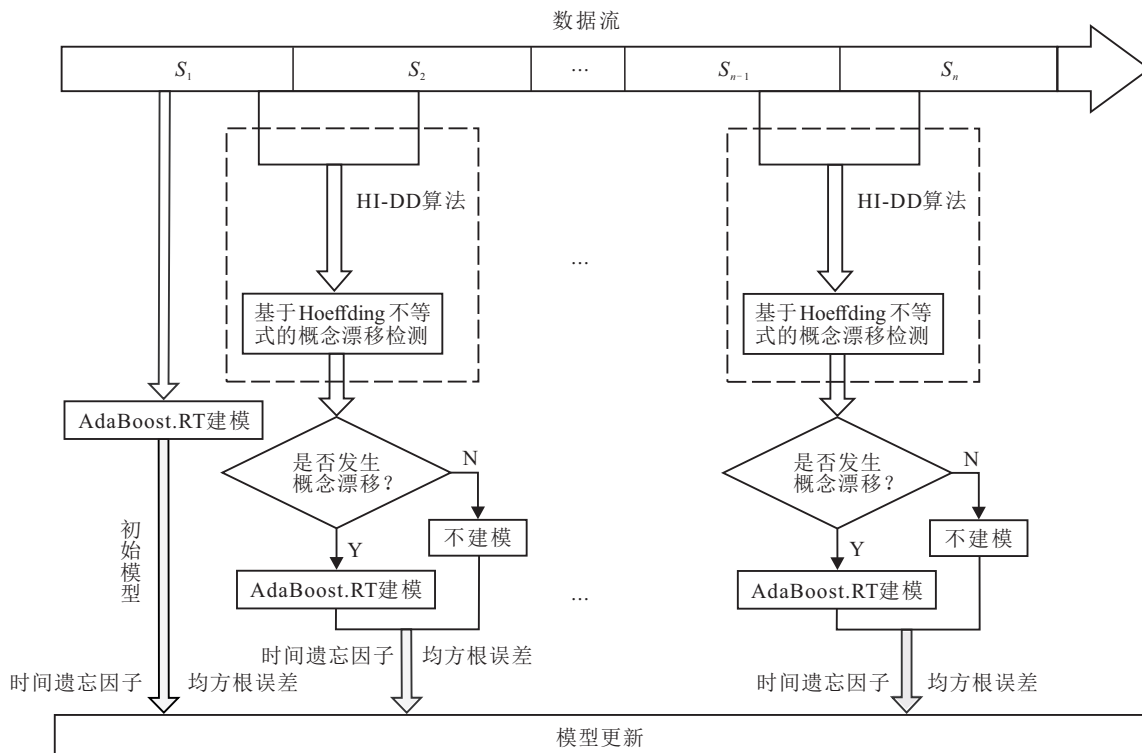


图3 HI-DD-AdaBoost.RT算法原理

根据上述原理,该算法的具体步骤如下.

**算法2** HI-DD-AdaBoost.RT在线预测算法.

输入: 数据流  $\text{num} = [(x_1, y_1), \dots, (x_n, y_n)]$ , 时间遗忘因子  $\lambda$ , 滑动窗口集合  $S = []$ , 长度  $\text{size}_k$ , 滑动窗口每次向前滑动的长度  $u$ ;

$$\text{输出: } F(x) = \frac{\sum_j \left( \frac{1}{\text{RMSE}_j} \right) (1-\lambda)^{J-j} F_j(x)}{\sum_j \left( \frac{1}{\text{RMSE}_j} \right) (1-\lambda)^{J-j}}.$$

1. for  $i$  in  $n$  do
  1.  $S = [S; \text{num}(i)]$ ;
  2. if  $\text{length}(S) = \text{size}_k$
  3. 使用 AdaBoost.RT 集成算法建立初始模型  $F_1$ ;
  4. end if
  5. if  $\text{length}(S) > \text{size}_k$  and  $(\text{length}(S) > \text{size}_k)/u = 1$
  6. //维持窗口长度不变
  7. for  $q$  in  $u$  do
    1.  $S(1) = []$ ;
  8. end for
  9. //根据 HI-DD 算法检测是否发生概念漂移, 如果发生返回发生概念漂移的滑动窗口数据集  $B$ , 否则不更新模型
  10. if  $\text{length}(B) \neq 0$
  11. //对检测到的第  $j(2, 3, \dots, J)$  个滑动窗口数据集应用 AdaBoost.RT 算法建立模型  $F_j$ . 其中, 数据集  $B$  有  $m$  个数据, 即

$$B_j = [(x_1^j, y_1^j), (x_2^j, y_2^j), \dots, (x_m^j, y_m^j)];$$

12. //模型的更新, 计算数据集  $B_j$  的均方根误差

$$\text{RMSE}_j = \sqrt{\left( \sum_{d=1}^m (F_j(x_d) - y_d^j)^2 \right)};$$

13.  $y_d^j$  表示第  $j$  个数据集中第  $d$  个样本的输出;
14. end if
15. end if
16. end for

算法首先在增量学习基础上进行 SOH 的预测, 其中 HI-DD 算法通过融合滑动窗口技术进行数据的分块并检测数据中发生的概念漂移, 简化了增量学习的建模与更新过程. 同时, AdaBoost.RT 集成学习算法解决了传统数据驱动算法的预测精度低和参数调整复杂的问题, 使 SOH 的预测精度更高. 最后, 加入遗忘因子和均方根误差实时建立总模型, 解决了增量学习中将历史数据与实时数据视为等同权重的问题, 实现了增量学习中的模型更新与预测功能, 可以

很好地实现电动汽车动力锂电池 SOH 的在线预测.

### 3 实验仿真及分析

#### 3.1 电池的充放电实验

本文采用 CALCE 官网公开的 CS2-36 型电池老化数据, 该电池是在恒流恒压模式下进行快速充放电循环实验, 每次循环的充电速率为 0.5 C, 直到电压达到 4.2 V, 再维持 4.2 V 恒压充电直到充电电流降到 0.05 A 以下, 充满电静置后再以 1 C 速率进行放电, 直到电池的放电截止电压降为 2.7 V.

#### 3.2 基于增量学习的 HI-DD-AdaBoost.RT 算法的在线 SOH 预测

在所选取电池实验数据的基础上, 本文所有算法实验均在 Windows 10 系统和 Matlab R2016b 环境下运行, PC 机硬件配置为: Intel(R) Core(TM) i7-6700 CPU, 32 GB 内存, 500 GB 硬盘. 选取第 1 节的 9 个评价指标作为模型的输入, 最大可用容量定义的 SOH 作为模型的输出, 采用基于增量学习的 HI-DD-AdaBoost.RT 算法进行在线预测.

##### 3.2.1 HI-DD 算法的实验验证及结果分析

为检验所提出 HI-DD 算法的有效性, 以公开的 KDDCUP99<sup>[11]</sup> 网络攻击数据集验证概念漂移检测方法 HI-DD 的有效性. 首先对 KDDCUP99 数据集进行预处理, 将提供的字符型特征转换为数值型特征; 然后选取前 50 000 个测试数据对网络的状态进行概念漂移的检测, 检测结果如图 4 所示. 其中, 网络的状态有 28 种类型, “0” 表示当前网络状态正常, “1~28” 表示不同类型的网络攻击.

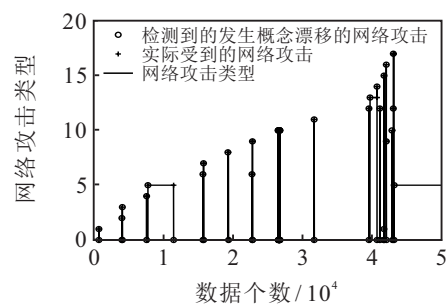


图4 KDD99数据集概念漂移检测结果

由检测结果可知, 基于 HI-DD 算法的概念漂移检测方法可以检测到不同类型的网络攻击, 因此可以很好地根据数据分布的变化找到概念漂移发生的位置, 说明基于 HI-DD 的概念漂移检测方法可行有效.

##### 3.2.2 基于增量学习的 HI-DD-AdaBoost.RT 的 SOH 预测实验及结果分析

使用 HI-DD 算法实时检测数据的概念漂移, 并将检测到的概念漂移数据应用 AdaBoost.RT 算法进

行建模. 首先选取900个样本作为训练样本集, 剩余样本作为测试样本集; 然后设置一个大小为300的检测窗口, 该窗口的长度可随意选定, 其次设定滑动窗口每次滑动的数据长度, 由于电池的SOH值会随着循环使用不断降低, 其数据曲线整体呈抛物状的下降趋势, 而且越到电池的使用后期其SOH下降速率越快, 因此窗口每次滑动的长度将决定概念漂移检测的灵敏度. 针对SOH数据的变化趋势, 从理论上可分析出当窗口滑动的数据个数越多时, 两窗口间差异会越大, 概念漂移的检测越灵敏, 学习过程中模型更新的次数会越多, 因此通过迭代得到概念漂移发生的概率(发生漂移的数据块个数/总数据块个数)与窗口向前移动长度间的关系, 如图5所示.

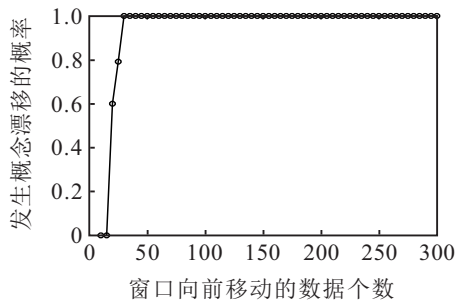


图5 概念漂移的灵敏度与窗口移动大小的关系

由以上结果可知, 窗口向前移动的数据个数越多, 其概念漂移的检测越灵敏, 与之前的理论分析一致. 当窗口每次向前移动的数据个数为20~30时, 概念漂移检测的灵敏度适中, 因此本次实验中将窗口移动的数据长度定为25.

在AdaBoost.RT建模与模型更新的过程中, 子学习机选择BP神经网络, 隐含层节点个数设为16, 使用10个BP子学习机进行集成学习, 滑动窗口数据块的遗忘因子设为0.95.

基于前面的分析与参数设置, 首先应用HI-DD算法对SOH的数据进行概念漂移的检测, 检测结果如图6所示. 由图6可知, 此次实验数据中发生概念漂移点的个数有19个, 加上初始模型的建立需要进行20次模型的更新. 如果不进行概念漂移的检测, 则需要维持窗口的大小加上初始模型的建立进行25次模

型的更新, 由此可见概念漂移的加入会减小模型更新的时间.

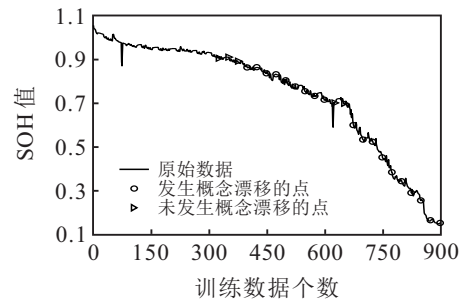


图6 SOH数据的概念漂移检测状态

根据上面检测到的发生概念漂移的点, 分别应用AdaBoost.RT集成算法、基于增量学习的Bagging算法、基于增量学习的HI-DD-Bagging算法、基于增量学习的AdaBoost.RT算法、基于增量学习的HI-DD-AdaBoost.RT算法进行SOH的预测, 其预测的最终结果和预测误差如图7和图8所示.

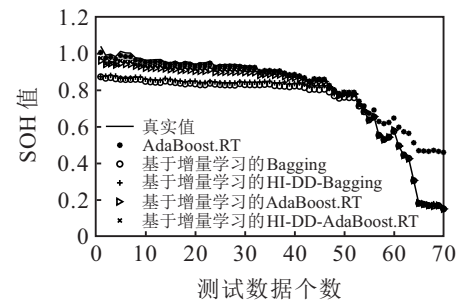


图7 基于不同算法的SOH预测结果

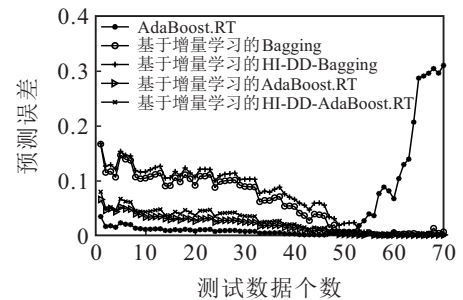


图8 基于不同算法的SOH预测误差

由图7和图8可知, 基于增量学习的在线预测学习算法预测效果较好. 为更加清楚地了解所提算法的优越性, 通过误差与运行时间分别对以上算法进行性能比较, 其比较结果如表2所示.

表2 不同算法性能指标对比

算法	均方误差	平均绝对误差	运行时间/s
AdaBoost.RT	0.009 3	0.046 1	—
基于增量学习的Bagging	0.007 5	0.070 4	43.679 840
基于增量学习的HI-DD-Bagging	0.005 9	0.060 3	32.112 562
基于增量学习的AdaBoost.RT	0.001 0	0.024 6	43.234 740
基于增量学习的HI-DD-AdaBoost.RT	0.000 6	0.018 3	31.953 327

由以上不同算法的预测结果和运行时间比较可以看出,单一的预测算法误差最大,加入基于HI-DD概念漂移检测的增量学习算法,其模型的预测结果和误差与未加基于HI-DD概念漂移检测的增量学习算法相差不大,这也表明所提出的HI-DD-AdaBoost.RT在线学习算法预测稳定性较高,可以在更短的时间内达到同样好的效果,因此本文所提的HI-DD-AdaBoost.RT算法预测效果好且运行时间短,适合动力锂离子电池SOH的在线预测。

## 4 结论

本文所提的基于增量学习的HI-DD-AdaBoost.RT在线预测算法可以实时地监测电池的健康状态,有利于电池的循环使用.针对SOH在线预测难以实现的问题,选取9个表征电池健康状态的指标,在增量学习的基础上,使用融合滑动窗口技术的HI-DD算法对SOH相关数据进行概念漂移的检测,简化了增量学习中模型建立与模型更新的过程,减小了模型更新的时间;同时考虑到传统算法的预测精度低和增量学习中历史数据的遗忘性问题,使用AdaBoost.RT集成学习算法并结合时间遗忘因子和误差率进行模型的更新,最后在CALCE提供的电池实验数据上进行验证,并与其他方法进行了比较,实验表明基于增量学习的HI-DD-AdaBoost.RT算法具有较高的预测精度,可以满足实际生活中电动汽车锂离子电池SOH的在线预测要求。

## 参考文献(References)

- [1] 韦海燕, 陈孝杰, 吕治强, 等. 灰色神经网络模型在线估算锂离子电池SOH[J]. 电网技术, 2017, 41(12): 4038-4044.  
(Wei H Y, Chen X J, Lv Z Q, et al. Online estimation of lithium ion battery SOH by grey neural network model[J]. Power Grid Technology, 2017, 41(12): 4038-4044.)
- [2] 刘新天, 李涵琪, 魏增福, 等. 基于Drift-Ah积分法的CKF估算锂电池SOC[J]. 控制与决策, 2019, 34(3): 535-541.  
(Liu X T, Li H Q, Wei Z F, et al. Estimation of lithium battery SOC by drift-ah integration based on CKF[J]. Control and Decision, 2019, 34(3): 535-541.)
- [3] 李昌, 罗国阳. 结合支持向量机的卡尔曼预测算法在VRLA蓄电池状态监测中的应用[J]. 电工技术学报, 2011, 26(11): 168-174.  
(Li C, Luo G Y. Application of kalman prediction algorithm combined with support vector machine in VRLA battery state monitoring[J]. Transactions of China Electrotechnical Society, 2011, 26(11): 168-174.)
- [4] Ahmed R, Sayed M E, Arasaratnam I, et al. Reduced order electrochemical model parameters identification and state of charge estimation for healthy and aged li-Ion batteries-Part II: Aged battery model and state of sharge estimation[J]. IEEE Journal of Emerging and Selected Topics in Power Electronics, 2014, 2(3): 678-690.
- [5] Paul S, Diegelmann C, Kabza H. Analysis of ageing inhomogeneities in lithium-ion battery systems[J]. Journal of Power Sources, 2013, 239(10): 642-650.
- [6] Li Y Y, Zhong S M, Zhong Q S, et al. Lithium-ion battery state of health monitoring based on ensemble learning[J]. IEEE Access, 2019, 7: 8754-8762.
- [7] 张杰, 赵峰. 流数据概念漂移的检测算法[J]. 控制与决策, 2013, 28(1): 29-35.  
(Zhang J, Zhao F. Detection algorithm for concept drift of flow data[J]. Control and Decision, 2013, 28(1): 29-35.)
- [8] 焦自权, 范兴明, 张鑫, 等. 基于改进粒子滤波算法的锂离子电池状态跟踪与剩余使用寿命预测方法[J]. 电工技术学报, 2020, 35(18): 3979-3993.  
(Jiao Z Q, Fan X M, Zhang X, et al. State tracking and residual service life prediction method for lithium ion batteries based on improved particle Filter algorithm[J]. Journal of Electrical Technology, 2020, 35(18): 3979-3993.)
- [9] 季一木, 张永潘, 郎贤波, 等. 面向流数据的决策树分类算法并行化[J]. 计算机研究与发展, 2017, 54(9): 1945-1957.  
(Ji Y M, Zhang Y P, Lang X B, et al. Parallelization of decision tree classification algorithm for streaming data[J]. Computer Research and Development, 2017, 54(9): 1945-1957.)
- [10] Solomatine D P, Shretha D L. Experiment with AdaBoost.RT: An improved boosting scheme for regression[J]. Neural Computation, 2006, 18: 1678-1710.
- [11] 李勇, 张波. 一种基于深度CNN的入侵检测算法[J]. 计算机应用与软件, 2020, 37(4): 324-328.  
(Li Y, Zhang B. An intrusion detection algorithm based on deep CNN[J]. Computer Application and Software, 2020, 37(4): 324-328.)

## 作者简介

田慧欣(1978—), 女, 教授, 博士, 从事电动汽车锂离子电池的智能建模、控制与优化等研究, E-mail: icedewl@163.com;

秦鹏亮(1996—), 男, 硕士生, 从事工业大数据技术、电池健康状态预测的研究, E-mail: 1310070913@qq.com;

李坤(1981—), 男, 副教授, 博士, 从事优化算法、人工智能等研究, E-mail: lk\_neu@163.com;

王红一(1984—), 女, 副教授, 博士, 从事人工智能图像处理等研究, E-mail: wanghongyi@tjpu.edu.cn.

(责任编辑: 齐 霖)