

文章编号: 2095-2163(2020)03-0040-05

中图分类号: TP202+.1

文献标志码: A

# 改进布谷鸟算法在锂电池剩余寿命预测中的应用

王海洋, 宋万清

(上海工程技术大学 电子电气工程学院, 上海 201620)

**摘要:** 本文提出一种将改进布谷鸟算法(ICS)用于优化相关向量机(RVM)的锂电池剩余寿命预测(RUL)方法。首先, 本文详细介绍了RVM模型原理以及权值系数 $\omega$ 的概率分布式, 同时给出RVM预测模型的主要流程。接着, 通过双参数动态调整策略改进CS算法, 并利用ICS算法对RVM模型中的权值系数 $\omega$ 进行了优化。然后通过美国航空航天局开源数据库中锂电池数据阐述了锂电池RUL预测主要步骤, 并最终通过实际实验与一系列误差指标验证了本文提出的ICS+RVM方法有着更高的锂电池RUL预测准确度。

**关键词:** 锂电池; 改进布谷鸟算法; 相关向量机; 剩余使用寿命; 概率分布函数; 双参数动态调整

## Application of Improved Cuckoo Search Algorithm in lithium-ion batteries remaining useful life prediction

WANG Haiyang, SONG Wanqing

(School of Electronic and Electrical Engineering, Shanghai University of Engineering Science, Shanghai 201620, China)

**[Abstract]** In this paper, an Improved Cuckoo Search Algorithm (ICS) is proposed to optimize the Relevance Vector Machine (RVM) for the prediction of the remaining useful life (RUL) of lithium-ion batteries. Firstly, the principle of RVM model and the probability distribution of weight coefficient  $\omega$  are introduced in detail, and the main process of RVM prediction model is given. Then, CS algorithm is improved by two parameters dynamic adjustment strategy, and the weight coefficient  $\omega$  of RVM model is optimized by ICS algorithm. After that, the main steps of lithium-ion batteries RUL prediction are described through the data of lithium-ion batteries in the open source database of NASA. Finally, the actual experiment and a series of error indices verify that the proposed ICS + RVM method in this paper has a higher accuracy of lithium-ion batteries RUL prediction.

**[Key words]** lithium-ion batteries; Improved Cuckoo Search Algorithm; relevance vector machine; remaining useful life; probability distribution function; two parameters dynamic adjustment

### 0 引言

随着智能手机等一系列电子设备的广泛应用, 锂电池作为这些电子设备的能量来源受到了普遍的关注, 对于其剩余寿命(Remaining Useful Life)的预测研究也已成为目前学界的热点课题<sup>[1]</sup>。

经过多年的发展, 锂电池剩余寿命预测方法陆续涌现, 较为典型的主要有: 模型法<sup>[2]</sup>、数据驱动法<sup>[3]</sup>以及混合法<sup>[4]</sup>。当前研究人员主要利用数据驱动法进行锂电池的剩余寿命预测。朱晓栋等人<sup>[5]</sup>考虑到差异性问题的提出了基于维纳过程的锂电池剩余寿命预测方法。王飞<sup>[6]</sup>则通过改进隐马尔可夫模型从而优化锂电池剩余寿命的预测结果。何成等人<sup>[7]</sup>提出了非线性回归神经网络法来训练锂电池容量数据的方法, 取得不错的效果。陈雄姿等人<sup>[8]</sup>利用改进的最小二乘支持向量回归法对锂

电池剩余寿命进行概率性预测。王腾蛟等人<sup>[9]</sup>提出了加入粒子群优化的粒子滤波算法预测锂电池寿命, 预测精度得到提高。上述数据驱动的锂电池RUL预测都有一定的缺点与局限性, 因此本文提出了一种新颖的锂电池RUL预测方法。

研究中, 通过支持向量机模型并结合概率学习的贝叶斯理论从而提出了相关向量机(Relevance Vector Machine, RVM)模型。由于RVM模型结合了朴素贝叶斯原理, 因此可通过自相关来判断先验与极大似然估计法来计算后验分布。RVM模型相对于SVM模型而言, 具有更好的计算准确度以及更低的计算复杂程度。但由于RVM处理数据过分的稀疏以及容量数据存在动态波动的特征, 从而使得直接利用RVM模型进行各种预测实验时, 预测实际结果的稳定性较差。因此, 本文利用改进的布谷

**基金项目:** 上海自然科学基金项目(14ZR1418500); 上海工程技术大学项目(18XJC002)。

**作者简介:** 王海洋(1994-), 男, 硕士研究生, 主要研究方向: 短期电力负荷预测、锂电池剩余寿命预测; 宋万清(1962-), 男, 教授, 主要研究方向: 状态监测与故障诊断、大数据处理与分析。

**通讯作者:** 王海洋 Email: whysues@163.com

**收稿日期:** 2019-11-22

鸟算法对其进行了系数优化,从而提高最终预测准确性。

## 1 相关向量机模型

### 1.1 RVM 模型原理

相关向量机方法是 Tipping 根据朴素贝叶斯理论的框架提出的一种将低维非线性的问题转化至高维线性的方法。RVM 通过计算最大化后验概率求解函数里相关向量的权值系数。对于给定的训练样本集  $\{x_n, t_n\}$  ( $n = 1, 2, \dots, N$ ), 样本的输出  $t_n$  表达式如下:

$$t_n = y(x_n; \omega) + \varepsilon_n, \quad (1)$$

其中,  $x_n$  表示第  $n$  个样本的输入,  $\varepsilon_n$  为噪声项。根据式(1), 推得 RVM 模型的函数表达式如下:

$$y(x, \omega) = \sum_{i=1}^N \omega_i k(x, x_i) + \varepsilon_n, \quad (2)$$

其中,  $\omega_i$  ( $i = 0, 1, \dots, N$ ) 为权值系数的向量,  $k(x, x_i)$  表示核函数。

假设输出  $t$  是独立的, 并且噪声为白噪声(即  $\varepsilon_n$  服从均值为 0 且方差为  $\sigma^2$  的正态分布), 那给定训练样本集中的  $t$  的概率分布函数为:

$$p(t | \omega, \sigma^2) = \prod_{i=1}^N N(t_i | y(x_i; \omega), \sigma^2) = (2\pi\sigma^2)^{-\frac{N}{2}} \exp\left\{-\frac{1}{2\sigma^2} \|t - \Phi\omega\|^2\right\}, \quad (3)$$

其中,  $t = (t_1, t_2, \dots, t_N)^T$ ,  $\omega = (\omega_1, \omega_2, \dots, \omega_N)^T$ ,  $\Phi = [\varphi(x_1), \varphi(x_2), \dots, \varphi(x_N)]^T$  且  $\varphi(x_i) = [k(x_1, x_N), k(x_2, x_N), \dots, k(x_N, x_N)]^T$ 。

为了防止过拟合的发生, 将权值系数  $\omega$  的先验概率分布定义如下:

$$p(\omega | \alpha) = \prod_{i=1}^N N(\omega_i | 0, \alpha_i^{-1}), \quad (4)$$

其中,  $\alpha$  为  $N + 1$  维的超参数向量。

根据朴素贝叶斯原理,  $\omega$  的后验分布表达式如下:

$$p(\omega | t, \alpha, \sigma^2) = \frac{p(t | \omega, \sigma^2) p(\omega | \alpha)}{p(t | \alpha, \sigma^2)} = (2\pi)^{-\frac{N+1}{2}} \times |\Sigma|^{-\frac{1}{2}} \times \exp\left\{-\frac{1}{2} (\omega - \mu)^T \Sigma^{-1} (\omega - \mu)\right\}, \quad (5)$$

其中,  $\Sigma = (\sigma^{-2} \Phi^T \Phi + A)^{-1}$  为协方差矩阵;  $A$  为对角矩阵  $\text{diag}(\alpha_1, \alpha_2, \dots, \alpha_N)$ ;  $\mu = \sigma^{-2} \Sigma \Phi^T t$  为均值。而  $\alpha_i$  和  $\sigma^2$  都是需要估计得出的参数。其迭代表达式如下:

$$\alpha_i^{new} = \frac{\gamma_i}{\mu_i}, \quad (6)$$

$$(\sigma^2)^{new} = \frac{\|t - \Phi\mu\|}{N - \sum_i \gamma_i}. \quad (7)$$

其中,  $\mu_i$  为第  $i$  个  $\mu$  的均值;  $\gamma_i = 1 - \alpha_i \Sigma_{ii}$ ;  $\Sigma_{ii}$  为  $\Sigma$  的第  $i$  个对角元素。

### 1.2 RVM 预测步骤

上面一部分介绍了 RVM 模型的基本原理及参数推导过程, 这里拟对利用 RVM 模型进行预测, 并做阐释分述如下。

**步骤 1** 选择适当的核函数, 将特征向量映射到高维空间。RVM 模型可以选择的核函数有许多类型, 研究人员通常都是利用其中几种常用的核函数, 例如线性核函数、多项式核函数以及高斯核函数(即 RBF 核函数)等来应用于预测与分类。尤其以 RBF 核函数被使用的次数最多。选择 RBF 核函数最重要一点是选择合适的带宽, 因为带宽参数的过大与过小都会影响预测的最终效果, 使得预测的结果不太准确。

**步骤 2** 根据朴素贝叶斯理论, 计算最大后验概率分布, 得到权值系数  $\omega$  分布。

**步骤 3** 初始化  $\alpha_i$  和  $\sigma^2$ 。在 RVM 模型中  $\alpha_i$  和  $\sigma^2$  是通过迭代估计计算出的, 即需进行初始化处理。同时, 考虑到初始化的随机性, 因此初始化初值的选择不影响最终的预测的结果。

**步骤 4** 将上面计算与迭代估计的参数代入 RVM 模型, 进行最终的预测实验。

## 2 ICS 算法原理

本文采用一种优化算法对 RVM 模型的核函数的系数的权值系数  $\omega$  进行优化。本文选择布谷鸟算法(Cuckoo Search Algorithm)对 RVM 模型的权值系数  $\omega$  进行优化。布谷鸟搜索(CS)算法是自启发式算法, 其通过模拟布谷鸟的寄生繁衍过程来解决全局优化问题。CS 算法采用了 Levy 飞行搜索机制, 遵循 3 条准则, 具体如下:

**准则 1** 布谷鸟一次只产一个蛋, 并且会把所产蛋丢在随机选择的某个鸟巢里。

**准则 2** 只有能最佳孵化出的杜鹃蛋的鸟巢才会被保存下来, 以供下一代使用。

**准则 3** 可供选择的寄主鸟巢数量是一定的。杜鹃鸟产下的蛋以一定的概率被寄主鸟发现, 此概率定为  $p_a \in [0, 1]$ 。

根据上述三个准则, 基于 Levy 飞行的 CS 算法计算公式为:

$$X_{t+1} = X_t + \alpha \oplus \text{levy}(\beta), \quad (8)$$

其中,  $t = 1, 2, 3, \dots$ ,  $\alpha$  为步长控制因子;  $\oplus$  为点对点乘法;  $\text{levy}(\beta)$  为 Levy 飞行并满足:

$$\text{Levy}(\beta) \sim u = t^{-\lambda}, 1 < \lambda \leq 3, \quad (9)$$

其中,  $\lambda$  为 Levy 飞行参数。

CS 算法里有 4 个参数, 分别是: 鸟巢数量  $N$ 、发现概率  $p_a$ 、步长  $\alpha$  以及 Levy 飞行中的  $\lambda$ 。由于鸟巢数量与 Levy 飞行参数在初始化后基本就确定了, 于是对于改进的 CS 算法主要便是对  $p_a$  和  $\alpha$  的改进。本文提出了双参数动态调整的策略, 首先便是对于步长  $\alpha$  的改进。原本的步长也是固定不变的, 这使得调整迭代次数时会降低预测精度, 本文提出步长与迭代次数同步指数递减的方法来动态调整  $\alpha$ , 具体公式如下:

$$\alpha'(t) = \alpha(t) \cdot \max\_it \cdot \exp\left\{-\frac{i}{\max\_it}\right\}, \quad (10)$$

其中,  $i$  为当前迭代次数,  $\max\_it$  为最大迭代次数。

接着, 就是对于发现概率  $p_a$  的改进。  $p_a$  作为概率参数对于寻优十分关键, 将其概率改为两段的变化概率。也就是在最大迭代次数  $\max\_it$  前半部分为负正弦自适应递减, 而在最大迭代次数  $\max\_it$  后半部分为了搜寻最佳的鸟巢位置采取负余弦自适应递减, 公式如下:

$$p_a = \begin{cases} p_{\alpha\max} - \sin\left(\frac{\pi}{2}\right) \cdot \frac{i-1}{\max\_it-1} \cdot \frac{\alpha}{\alpha} (p_{\alpha\max} - p_{\alpha\min}), & i \leq \frac{\max\_it}{2}; \\ p_{\alpha\max} - \cos\left(\frac{\pi}{2}\right) \cdot \frac{i-1}{\max\_it-1} \cdot \frac{\alpha}{\alpha} (p_{\alpha\max} - p_{\alpha\min}), & i \geq \frac{\max\_it}{2}. \end{cases} \quad (11)$$

其中,  $p_{\alpha\max}$  和  $p_{\alpha\min}$  分别表示发现概率的最大值与最小值。于是通过上述两个参数的调整来探究 ICS 算法寻优能力。

利用 ICS 算法来优化权重系数  $\omega$  的主要步骤如下:

**步骤 1** 目标函数的初始化。

**步骤 2** 设置 CS 算法参数, 步长  $\alpha$ , 发现概率  $p_a$  及其最大值与最小值  $p_{\alpha\max}$  和  $p_{\alpha\min}$ , 鸟巢数量  $N$ , Levy 飞行参数  $\lambda$ , 初始的位置  $nest_0 = [d_1^0, d_2^0, \dots, d_N^0]$ , 设将各参数代入目标函数计算得出的均方根误差 (RMSE) 为适应值  $fbest_0 =$

$[y_1^0, y_2^0, \dots, y_N^0]$ , 最后设最大迭代次数  $\max\_it$ 。

**步骤 3** 选择得到的最佳的第一代的鸟巢位置  $d_h^i$ , 根据式(10)计算步长, 再代入到 Levy 飞行来搜寻鸟巢位置, 记新的鸟巢为  $S$ , 将  $S$  代入目标函数计算出新的适应值  $fnew$ 。将  $fnew$  与上一代的最佳适应值  $y_h^i$  相比, 如果  $fnew < y_h^i$ , 则令基于 Levy 飞行所得鸟巢位置设为新的最佳巢位, 反之亦然。

**步骤 4** 根据公式(11)更新  $p_a$ , 设置随机数  $rand(0, 1)$ 。如果  $rand > p_a$  则随机选择鸟巢位置, 并将其替换为最差巢位, 反之亦然。

**步骤 5** 如果满足最大迭代次数, 则停止搜索; 否则返回步骤 3。

**步骤 6** 输出整个过程中适应值最小的鸟巢位置, 并将此位置所计算出的  $\omega$  作为最优解。

### 3 锂电池 RUL 预测过程

本文的锂电池预测数据采用了美国航空航天局开源数据库里面 2014 年发布的随机游走 (Random Walk) 退化的锂电池数据集作为实验对象<sup>[10]</sup>。本实验中, 4 组锂电池 (RW9, RW10, RW11, RW12) 容量退化过程如图 1 所示, 本实验 4 组电池初始容量都为 2.0 Ahr。

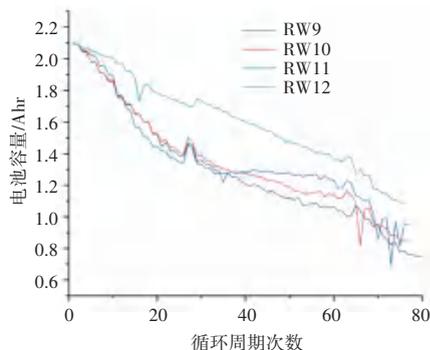


图 1 4 组锂电池退化数据

Fig.1 4 sets of lithium-ion battery degradation data

为了介绍锂电池 RUL 预测的具体流程, 本文利用 4 组电池数据中的 RW9 数据为例, 给出锂电池 RUL 预测步骤详述如下。

**步骤 1** 首先, 便是锂电池退化状态的识别, 也就是利用上述锂电池数据, 代入 RVM 模型进行训练并利用 ICS 算法进行权值系数  $\omega$  的优化, 由此建立起优化后的 RVM 模型进行 RUL 预测。

**步骤 2** 然后, 是各种指标的设置。根据国际标准认为锂电池容量退化至初始容量的 70%~80% 则认为电池失效, 此处的电池容量值规定为失效阈值 (Failure Threshold, FT)。而当锂电池不断进行充放电循环, 其循环周期到达所设失效阈值的那个时

刻,即称为寿命终结点(End of Life, EOL)。

**步骤3** 接着是锂电池 RUL 预测过程。选择不同区间的训练样本使得初始预测时间点不同,预测到达失效阈值的退化曲线。

**步骤4** 最后,通过计算概率密度函数得出最终的预测的 RUL。当所预测退化曲线达到所设定的失效阈值时,为了更加准确地得出最终的 RUL,再计算出达到点处的概率密度(Probability Density Function, PDF)并描绘出如图2所示的概率密度函数分布图。其计算公式可表示为:

$$P(L_k | Q_{FT}) = \sum \omega_i \delta(L_k - \hat{L}_k). \quad (12)$$

其中,  $Q_{FT}$  表示达到退化阈值的电池容量;  $L_k$  表示  $k$  起始预测时刻的实际锂电池 RUL;  $\hat{L}_k$  表示  $k$  起始预测时刻的预测出的锂电池 RUL;  $\delta$  为狄拉克函数。

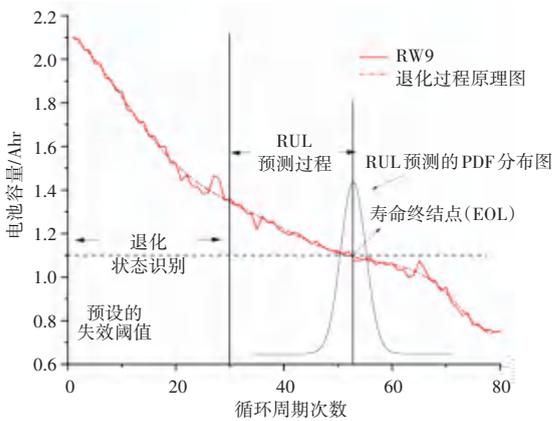


图2 锂电池 RUL 预测流程原理图

Fig. 2 Schematic diagram of lithium-ion battery RUL prediction process

### 4 实验流程与分析

本节将进行实际的锂电池 RUL 预测实验与结果分析。根据上节锂电池失效的国际标准,因此本次实验预设锂电池失效阈值为实际容量的70%。选择 RW12 组锂电池退化数据来进行 RUL 预测实验,因为其对应的退化过程退化至70%的时刻最慢,也最为平缓。实验中,分别利用 RVM 模型与 ICS+RVM 模型对 RW12 组锂电池退化数据进行 RUL 预测。选择12个不同的起始预测点分别进行 RUL 预测,计算出每个预测点最终的 PDF 并绘制出分布图,详见图3和图4。

为了更加直观与全面地展示对于锂电池退化模型的 RUL 预测效果,本文使用了评价整体预测性能的误差定量指标:均方误差(Mean Square Error, MSE)与均方根误差(Root Mean Square Error,

RMSE)两种方式。这里将用到如下数学公式:

$$MSE = \frac{1}{N} \sum_{i=1}^N (RUL_i^f - RUL_i)^2, \quad (13)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (RUL_i^f - RUL_i)^2}. \quad (14)$$

其中,  $N$  表示所预测的锂电池退化模型所选取的初始预测时间点个数;  $RUL_i^f$  表示第  $i$  个起始预测时间点到预测出的 EOL 之间的锂电池 RUL;  $RUL_i$  表示第  $i$  个起始预测时间点到实际的 EOL 之间的锂电池 RUL。误差计算结果见表1。

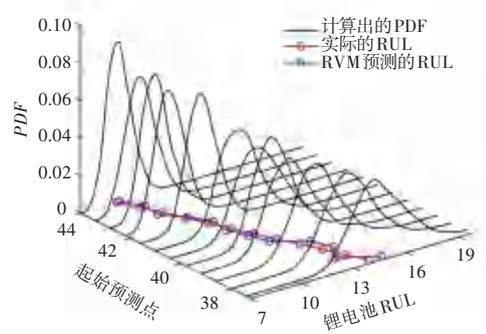


图3 利用 RVM 进行 RUL 预测的结果图

Fig. 3 Results of RUL prediction using RVM

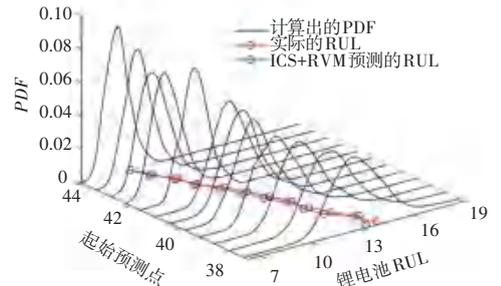


图4 利用 ICS+RVM 进行 RUL 预测的结果图

Fig. 4 Results of RUL prediction using ICS+RVM

表1 预测结果的 MSE 和 RMSE

Tab. 1 Predicted MSE and RMSE

RW12	MSE	RMSE
RVM	8.667	2.943
ICS+RVM	4.333	2.081

### 5 结束语

本文提出了一种利用改进布谷鸟算法优化相关向量机模型的锂电池剩余寿命预测方法。文章首先介绍了基于贝叶斯准则的 RVM 模型原理,并给出 RVM 预测步骤。接着详细介绍了利用 ICS 算法优化权值系数  $\omega$  的过程。通过结合 RW9 组锂电池退

(下转第 50 页)