

基于 LSTM-UKF 的全电船用锂电池 SOC 容量估计

辛浩东¹, 耿攀¹

(1. 上海海事大学 物流工程学院, 上海 201306)

摘要: 全电船舶(AES) 采用锂电池和燃料电池等清洁能源替代化石燃料, 在减少污染物和温室气体排放发挥着重要的作用。电池荷电状态(SOC)作为电池的一个重要状态量, 能够估计全电船舶的运行时间, 是电池管理系统(BMS)进行电池的监控管理的重要的基础。首先建立了高精度电池模型并使用无迹卡尔曼滤波器(UKF)对锂电池 SOC 进行估计。然后采用了一种长短期记忆网络(LSTM)的 SOC 估算方法, 使用电压、电流和温度变量对模型进行训练, 通过 UKF 滤除噪声, 进一步提升了 LSTM 网络的估算精度, 模型对 SOC 初始值的依赖程度低, UKF 算法得到了明显的优化。结果表明, 采用 LSTM-UKF 的网络模型相比 LSTM 网络模型的估算稳定性和精确度有了较大的提升, SOC 估算误差低于 0.8%, RMSE 仅为 0.22%。

关键词: 锂电池; SOC; UKF; LSTM; 全电船舶;

中图分类号: U661.43 **文献标识码:** A

State-of-charge Estimation Method of Lithium-ion Batteries in All-Electric Ships Based on LSTM-UKF

XIN Hao-dong¹, GENG Pan¹

(Shanghai Maritime University, Shanghai 201306, China)

Abstract: All-electric ships (AES) use lithium batteries and fuel cells to replace fossil fuels and play an important role in reducing pollutants and greenhouse gas emissions. As an important state quantity of the battery, State of charge (SOC) which is an important basis for the battery management system (BMS) to monitor and manage the battery, can estimate the running time of AES. First, a high-precision battery model is established and the Unscented Kalman Filter (UKF) is used to estimate the SOC of the lithium battery. Then a long-short-term memory network (LSTM) SOC estimation method is adopted. The model is trained using voltage, current and temperature variables, and the noise is filtered through UKF, which further improves the estimation accuracy of the LSTM network. At the same time, the model's dependence on the initial value of SOC is low, and the UKF algorithm has been significantly optimized. The results show that compared with the LSTM network, the LSTM-UKF model has a greater improvement in the estimation stability and accuracy. The SOC estimation error is less than 0.8%, and the RMSE is only 0.22%.

Key words: Lithium-ion batteries; SOC; UKF; LSTM; All-electric ships;

0 引言

清洁能源越来越被人们所重视。在陆上, 电动汽车行业正在蓬勃发展, 全电动交通工具不会排放污染物

和温室气体，逐渐被人们所接受，其被视为改善环境，提高空气质量的重要方式。在海上，海洋运输是大宗货物运输最节能的方式，但是其造成的空气污染依然不可以被忽视，解决的方式是在在船舶上引入大规模的储能系统(ES)^[1]。锂电池具有高能量密度、低自放电率和寿命长的优点，其在船舶上的使用率正在持续地增长。目前有些船舶采用了混合动力甚至全电力的方式，当船舶停靠港口时，船舶岸电系统可以为其充电，图 1 展示了一种零排放的船舶示意图，该船采用了锂电池组、超级电容和光伏发电的方式为其供电。

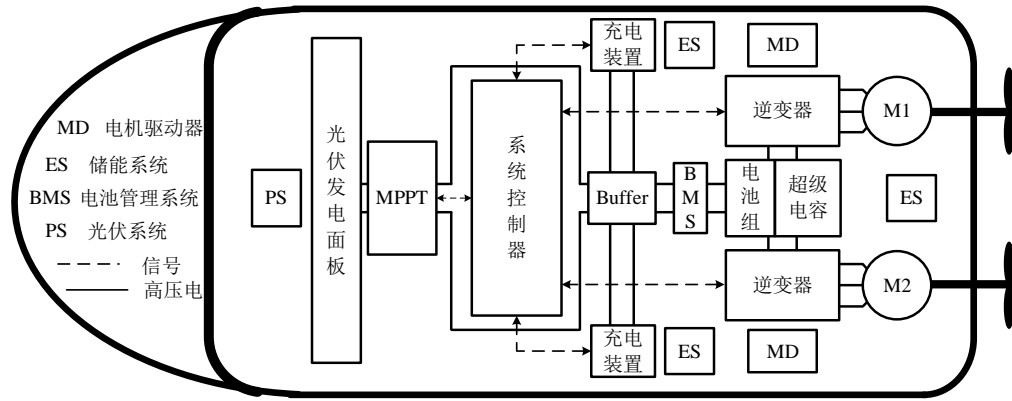


图 1 一种零排放全电动船舶示意图

蓄电池的荷电状态是设计和实现电池管理系统(BMS)的关键参数，是 BMS 研发的核心和难点。准确地估计 SOC 十分的重要，它是 BMS 对电池组进行管理的基础。准确的 SOC 不仅能够反映电池的剩余电量，还能为能量管理系统实现能量均衡分配提供重要依据^[2]。

常见的 SOC 估算方法可以分为三类，一类是基于电池的电压、电流和内阻参数的计算方法，常见的有安时积分法、开路电压法和内阻法。另一类是基于电池等效电路模型的估算方法，如卡尔曼滤波(KF)、龙贝格观测器等。最后一类是机器学习类算法，包括支持向量机(SVM)、长短期记忆网络(LSTM)等。

常见的 SOC 算法都有其各自的局限性：安时积分法需要获取精准的 SOC 初始值，且误差会被逐渐积累，需要对其不断校正；开路电压法利用电池 SOC 和开路电压关系相对稳定的特性，测得 OCV-SOC 曲线的数据，并由此计算 SOC，缺点是需要将电池长时间静止来获取稳定的 OCV 值；卡尔曼滤波法需要建立复杂的电池模型并获取其精确的参数数据，在实验室条件下可以取得很好的效果，然而电池的实际运行情况却十分复杂。目前 SOC 算法研究的主流是结合各种算法的优缺点，去寻找一种符合需求的 SOC 估算方法，本文采用了一种 LSTM-UKF 的 SOC 估算方法，利用 UKF 算法滤除噪声，使用已经训练的 LSTM 网络快速准确地更新 SOC 值，结合两种算法各自的优点，以提高 SOC 算法的鲁棒性和实际应用的能力。

1 基于 LSTM 的锂电池建模

1.1 LSTM 神经网络

LSTM 是一种循环神经网络(RNN)，RNN 的优势在于对时间序列的处理，与传统网络的区别在于，RNN 会将前次的结果代入下一次的隐藏层中，形成一种循环结构。在训练过程中 RNN 时常发生梯度爆炸或消失的现象^[5]，RNN 有记忆期短，无法处理长序列数据，训练成本大等缺点。RNN 模型展开结构如图 2 所示。

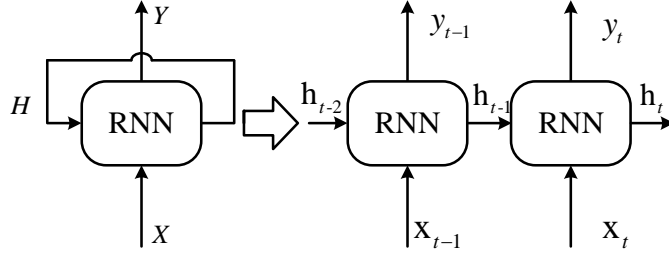


图2 RNN 模型展开

长短期记忆模型(LSTM)是一种强大的时间序列模型。在语音识别, 自然语言处理(NLP), 以及时间序列相关的领域等都有广泛的应用。LSTM 模块(或单元)有 5 个基本组件, 允许对长期和短期数据建模。

LSTM 针对 RNN 存在的问题做出了改进。引入了 LSTM 细胞结构(图 3)取代 RNN 的隐藏层神经元, 构建了三个门结构, 使网络能够记住长期的信息, 而且是只保留了“重要”的信息, “不重要”的信息被选择遗忘。LSTM 的关键就是细胞状态, 门结构可以实现细胞状态信息的增加和去除。门实际是一种选择信息通过的方法, 包括一个 Sigmoid 函数和矩阵点乘操作。

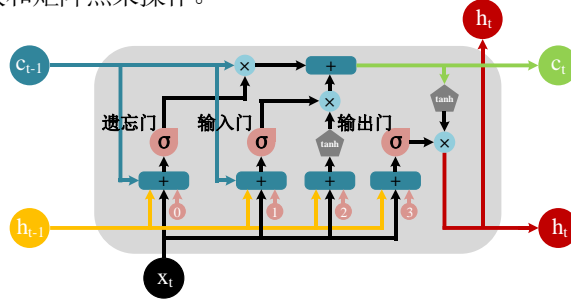


图3 LSTM cell 结构

LSTM 的三个门均采用 sigmoid 函数, 激励函数采用 tanh。以下公式中: i_t 、 f_t 、 o_t 和 c_t 分别表示输入门、遗忘门、输出门和细胞状态; W_i 、 W_f 、 W_o 和 W_c 表示对应的权重; b_i 、 b_f 、 b_o 和 b_c 表示状态的偏置量。

输入门部分:

$$i_t = \sigma(W_i h_{t-1} + W_i x_t + b_i) \quad (1)$$

$$c_t = f_t c_{t-1} + i_t \tanh(W_c h_{t-1} + W_c x_t + b_c) \quad (2)$$

遗忘门部分:

$$f_t = \sigma(W_f h_{t-1} + W_f x_t + b_f) \quad (3)$$

输出门部分:

$$o_t = \sigma(W_o h_{t-1} + W_o x_t + b_o) \quad (4)$$

$$h_t = \tanh(c_t) o_t \quad (5)$$

1.2 LSTM 神经网络的 SOC 估计

按时积分法定义的 SOC 如下。可以看到, 电池当前 SOC 与过去 SOC 之间存在着联系, LSTM 在时间序列的处理方面具有优势, 可以通过构建 LSTM 网络可以用来学习此种相关性。

$$\text{SOC} = \text{SOC}_0 - \frac{\int idt}{C_n} \quad (6)$$

锂电池在工作时存在着各种不确定因素，构建 LSTM 网络时应考虑实际运行情况，输入数据应与电池 SOC 存在相关性，同时输入数据应该容易获取，如开路电压，内阻等难以直接获取电池参数不能用于模型的训练。本文采用了电池电压、电流和温度作为输入数据，这些数据可以由传感器获取，模型训练完成后，BMS 可以直接获取输入数据，调用已训练 LSTM 网络模型完成 SOC 的估算。

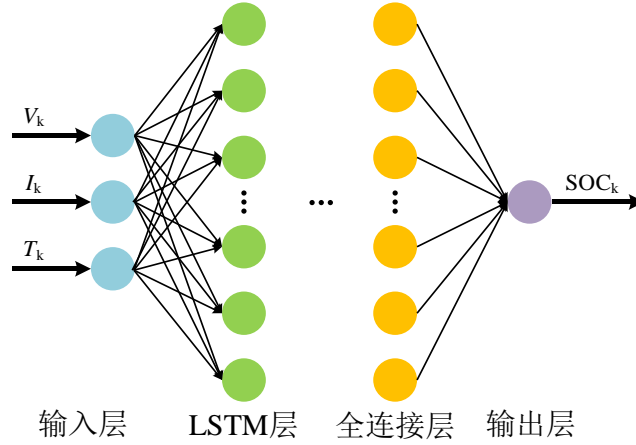


图4 LSTM 网络 SOC 估算模型

LSTM 网络的结构如图 4 所示，可以分为三部分，输入层主要负责数据的归一化，将其传递给 LSTM 层；LSTM 层主要负责学习输入向量之间，和他们在时间维度上的联系；输出层完成数据反归一化，输出 SOC 预测值。

2 基于 UKF 的锂电池建模

2.1 无迹卡尔曼滤波器

卡尔曼滤波器(KF)属于一种最优观测器，通过反馈调整所用模型的状态变量，使估计输出与实测输出之间的实时误差最小。建立系统的数学模型后，可以间接得到无法测量的系统状态。通过校正卡尔曼增益 K 来改善模型的动态特性，每次迭代时的 K 由状态与测量的误差和噪声（不确定性）来计算。KF 只适用于线性系统，不能直接用来估算电池 SOC。对于非线性系统人们提出了扩展卡尔曼滤波(EKF)和无迹卡尔曼滤波(UKF)，相比 UKF，EKF 忽略了系统高阶项，估算精度还有待改善^[3]。

UKF 是一种非线性的滤波的方法。UKF 利用 UT 变换(Unscented Transform)，在估计点的附近确定一些采样点，用这些点来表示的高斯密度近似状态的概率密度函数来做进一步的状态估计^[2]。UT 变换是指，在原状态分布中按某一规则选取一些采样点，使他们的均值和协方差与原状态分布的均值和协方差相同。将这些点代入非线性函数中，得到非线性函数值点集，然后通过得到的点集来求取变换后的均值与协方差。UT 变换后的均值与协方差的精度至少具有两阶精度的泰勒展开。UT 变换和 UKF 算法过程如下：

计算 $2n+1$ 个 Sigma 点和采样点对应的权值， n 是状态维数。

$$\begin{cases} \mathbf{X}^{(0)} = \bar{\mathbf{X}}, i = 0 \\ \mathbf{X}^{(i)} = \bar{\mathbf{X}} + (\sqrt{(n+\lambda)\mathbf{P}})_i, i = 1 \sim n \\ \mathbf{X}^{(i)} = \bar{\mathbf{X}} - (\sqrt{(n+\lambda)\mathbf{P}})_i, i = n+1 \sim 2n \end{cases} \quad (7)$$

$$\begin{cases} \omega_m^{(0)} = \frac{\lambda}{n + \lambda} \\ \omega_c^{(0)} = \frac{\lambda}{n + \lambda} + (1 - a^2 + \beta) \\ \omega_m^{(i)} = \omega_c^{(i)} = \frac{\lambda}{2(n + \lambda)}, i = 1 \sim 2n \end{cases} \quad (8)$$

式中：下标 m、c 分别表示均值和协方差，上标为采样点； $\alpha \in (0,1]$ 控制采样点的分布状态（采样点距离均值）；待选参数 $\beta \geq 0$ 为非负的权，可以合并方程高阶项的动差，增加高阶项的影响；参数 $\lambda = \alpha^2(n + \kappa) - n$ 表示缩放比例参数，用以减少总的预测误差， κ 是待选参数，通常应确保矩阵 $(n + \lambda)P$ 是半正定的。

UKF 计算及更新过程见表 1：

表 1 UKF 计算及更新过程

● 初始化状态 $X_0^a = E[x_0]$ ，协方差矩阵 $P_0 = E[(x_0 - x_0^a)(x_0 - x_0^a)^T]$

● 获得一组采样点（称 Sigma 点集）和对应的权值：

$$\hat{X}_k^{(i)} = [\hat{X}_k \quad \hat{X}_k + \sqrt{(n + \lambda)P_k} \quad \hat{X}_k - \sqrt{(n + \lambda)P_k}]$$

● 进一步预测状态量及协方差矩阵 $i = 1, 2, 3, \dots, 2n + 1$ ：

$$\hat{X}_k^{(i)-} = f(\hat{X}_k^{(i)}, u_k) \quad \hat{X}_k^- = \sum_{i=0}^{2n} \omega_m^{(i)} \hat{X}_k^{(i)-}$$

$$\hat{P}_k^- = \sum_{i=0}^{2n} \omega^{(i)} [\hat{X}_k^- - \hat{X}_k^{(i)-}] [\hat{X}_k^- - \hat{X}_k^{(i)-}]^T + Q$$

● 计算新的 Sigma 点集：

$$\hat{X}_{k+1|k}^{(i)} = [\hat{X}_k^- \quad \hat{X}_k^- + \sqrt{(n + \lambda)P_k^-} \quad \hat{X}_k^- - \sqrt{(n + \lambda)P_k^-}]$$

● 计算预测观测值： $Y_k^{(i)} = g(X_{k+1|k}^{(i)})$

● 计算系统的预测均值及协方差：

$$\hat{Y}_k = \sum_{i=0}^{2n} \omega_m^{(i)} Y_k^{(i)} \quad P_{y_i, y_i} = \sum_{i=0}^{2n} \omega_c^{(i)} [Y_k^{(i)} - \hat{Y}_k] [Y_k^{(i)} - \hat{Y}_k]^T$$

$$P_{x_i, y_i} = \sum_{i=0}^{2n} \omega_c^{(i)} [\hat{X}_{k+1|k}^{(i)} - \hat{X}_k^-] [Y_k^{(i)} - \hat{Y}_k]^T$$

● 计算 Kalman 增益矩阵： $K_k = P_{x_i, y_i} P_{y_i, y_i}^{-1}$

● 系统状态和协方差更新：

$$\hat{X}_{k+1} = \hat{X}_k^- + K_k (Y_k - \hat{Y}_k) \quad P_{k+1} = P_k^- - K_k P_{y_i, y_i} K_k^T$$

2.2 电池建模与参数辨识

卡尔曼滤波算法依赖于研究对象的数学模型，为了将 UKF 算法应用在电池 SOC 估算上，我们首先需要选择合适的电池等效电路模型，在建立电池的状态空间方程的基础上再应用卡尔曼滤波算法。选择电池模型时，应从实际需求出发，兼顾复杂度和精度这两个模型指标。本文采用如图 6 所示的三阶 RC 等效电路模型，RC 单元参数反映的是电池在运行过程对电池参数(温度、SOC 和充放电倍率)的依赖性。

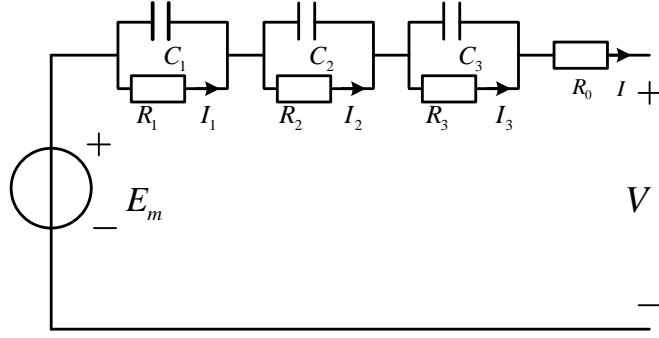


图 6 三阶 Thevenin 等效电路模型

由 SOC 的定义和基尔霍夫电压电流定律, 对图 6 中的电路列写方程, 经离散化处理后得到三阶 Thevenin 等效电路的数学模型:

$$\begin{bmatrix} SOC_{k+1} \\ U1_{k+1} \\ U2_{k+1} \\ U3_{k+1} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & e^{-\frac{T_s}{\tau_1}} & 0 & 0 \\ 0 & 0 & e^{-\frac{T_s}{\tau_2}} & 0 \\ 0 & 0 & 0 & e^{-\frac{T_s}{\tau_3}} \end{bmatrix} \begin{bmatrix} SOC_k \\ U1_k \\ U2_k \\ U3_k \end{bmatrix} + \begin{bmatrix} -\frac{T_s}{3600 \cdot Q_n} \\ R1(1 - e^{-\frac{T_s}{\tau_1}}) \\ R2(1 - e^{-\frac{T_s}{\tau_2}}) \\ R3(1 - e^{-\frac{T_s}{\tau_3}}) \end{bmatrix} [i_k] \quad (9)$$

$$V_k = E_m(SOC_k) - U1_k - U2_k - U3_k - R_0 \cdot i_k \quad (10)$$

式中: T_s 为采样时间, Q_n 为电池额定容量, $\tau_k = R_k C_k$ ($k = 1, 2, 3$) 为时间常数。

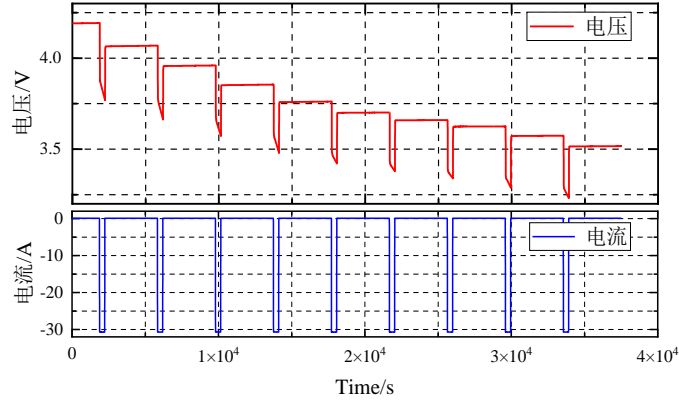


图 7 锂电池恒流放电

由于锂电池是强非线性系统, 电池模型的参数并不是固定不变的, 电池各参数会随着时间的改变而不断地变化。一般认为, 电池模型中电容、电阻和电动势是温度和 SOC 的函数, 即:

$$x = f(SOC(t), T(t)) \quad (11)$$

可以通过图 7 中所示的恒流放电实验确定参数与 SOC 和温度之间的函数关系。首先在不同温度的环境下重复恒流放电试验, 然后通过参数辨识算法确定某一确定温度下的函数关系。

锂电池参数辨识是通过算法不断地改变和优化电池模型中参数的过程。某一电池参数的电池模型在相同的恒流放电试验的电压曲线与真实实验曲线比较，参数辨识算法调整电池参数逐渐使两曲线拟合，最后得到最优的电池模型参数。表 2 为在 20°C 的环境下对锂电池参数的辨识结果。

表 2 20°C 下的参数辨识结果

SOC	Em/V	R0/Ω	R1/Ω	R2/Ω	R3/Ω	tau1/s	tau2/s	tau3/s
0.1	3.4937	0.0099	0.0003	0.0010	0.0027	9.7154	61.4454	650.0073
0.2	3.5526	0.0088	0.0009	0.0010	0.0011	12.1200	50.4659	649.8549
0.3	3.6141	0.0083	0.0013	0.0010	0.0016	12.0371	84.5851	938.7253
0.4	3.6490	0.0087	0.0003	0.0011	0.0013	10.9055	54.8648	349.4190
0.5	3.6886	0.0080	0.0012	0.0010	0.0018	9.2513	55.6604	983.4443
0.6	3.7440	0.0086	0.0005	0.0009	0.0004	13.4646	66.7955	958.5700
0.7	3.8394	0.0084	0.0007	0.0009	0.0009	10.2523	92.0037	962.0886
0.8	3.9485	0.0085	0.0008	0.0010	0.0009	10.0422	81.0404	402.2209
0.9	4.0609	0.0085	0.0008	0.0009	0.0007	9.6508	90.8335	405.7126
1	4.1903	0.0085	0.0008	0.0009	0.0003	10.0849	104.5829	568.2599

2.3 UKF 的 SOC 估计

考虑到系统在 UKF 估计时，存在过程噪声和测量噪声，通常将系统的噪声设定为固定值^[4]。电池额模型系统可以改写为：

$$x_k = A_k x_{k-1} + B_k u_k + \omega_k \quad (12)$$

$$y_k = E_m(SOC_k) - U1_k - U2_k - U3_k - R_0 i_k + v_k \quad (13)$$

式中： y_k 、 x_k 、 ω_k 、 v_k 分别为观测量、状态量、过程噪声和测量噪声。

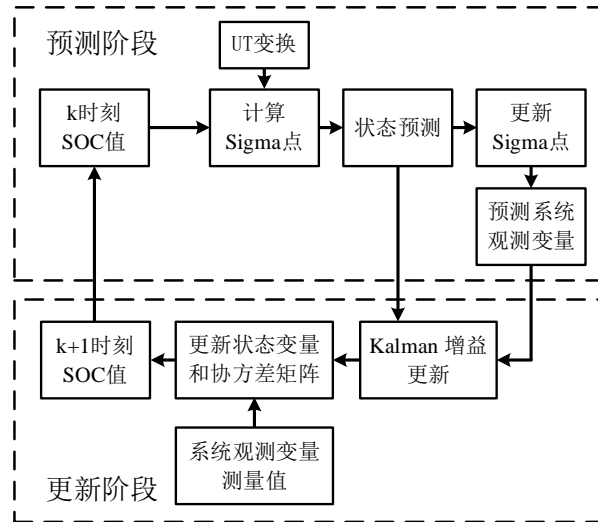


图 8 UKF 估算 SOC 过程

UKF 估算锂电池过程如图 8 所示。UKF 算法相对传统的安时积分法，其对 SOC 初始值的依赖程度小，误差也不会随时间积累，但是在 SOC 初始值设定误差较大的情况下，UKF 需要经过多次迭代更新后才可以很好地跟踪 SOC 真实值。如图 9 与图 10 所示，在本文建立的三阶 Thevenin 电池模型下，UKF 算法在 SOC 初始值确定的时可以快速实现 1% 以内的平均误差，然而 UKF 算法也有其缺点，在电池模型精度不够，和测量干扰较大时 SOC 估算效果往往不是很好。

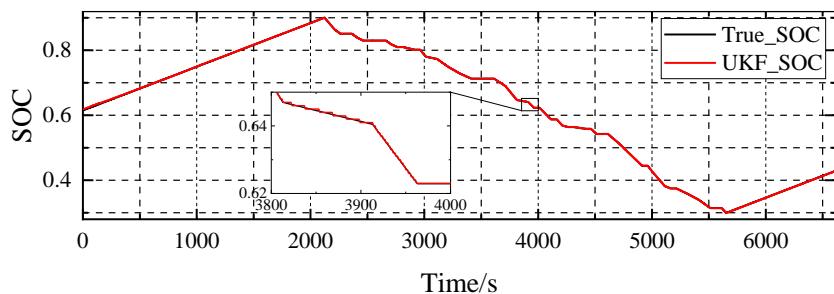


图9 UKF 估计结果

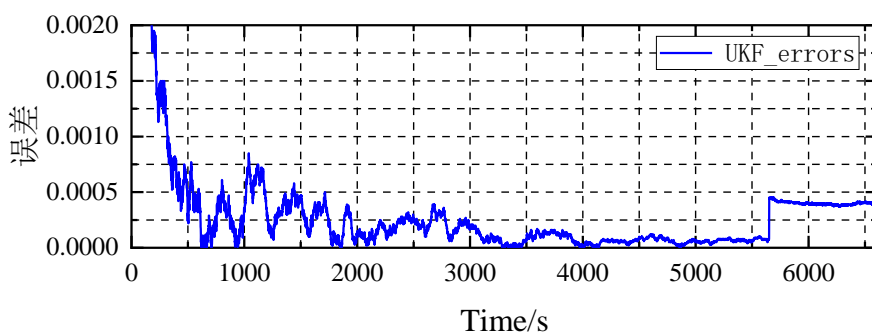


图10 UKF 估计误差

3 实验设计和数据预处理

3.1 LSTM-UKF 的 SOC 估计

LSTM-UKF 的 SOC 估算框图如图 11 所示,利用 LSTM 网络辅助 UKF 不断校正 SOC 的预测值,利用 UKF 算法滤除噪声减少预测误差,提高了 SOC 算法的实际应用能力。

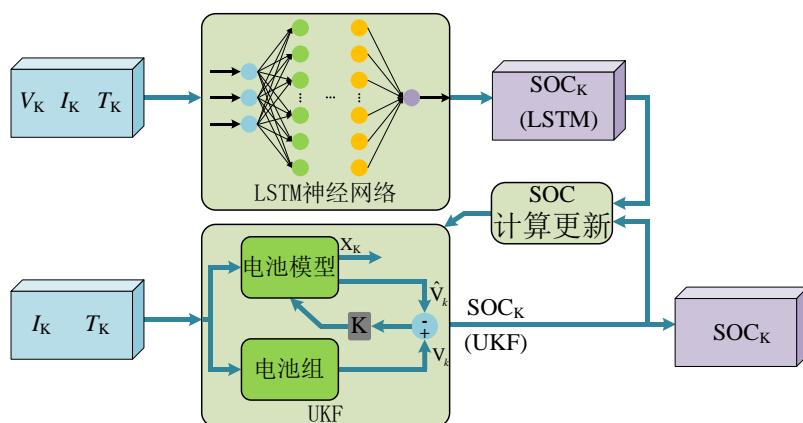


图11 LSTM-UKF 估计 SOC 框图

3.2 数据获取和预处理

实验采用 SANYO 公司 18650 磷酸铁锂电池组, 其额定容量为 31A·h。最大充电电压 4.2V, 放电截止电

压是 2.5V。对其进行充放电测试，采样时间为 1 秒，以获取训练数据，充放电电压电流如图 12 和图 13 所示。

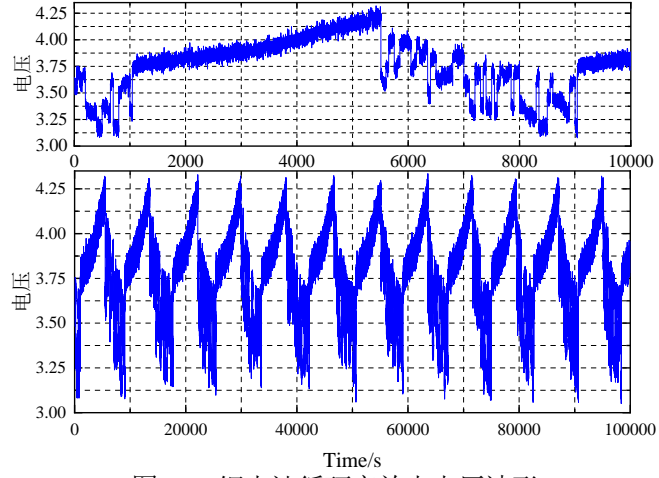


图 12 锂电池循环充放电电压波形

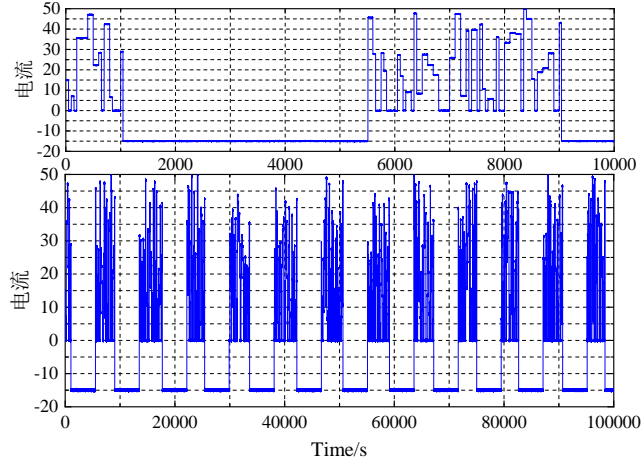


图 13 锂电池循环充放电电流波形

本次实验使用电池电压、电流和温度作为输入数据，其数据范围各有不同。为了使网络训练更加有效可靠，消除量纲加速优化的过程，训练前需要对数据正规化处理。 $SOC \in [0,1]$ ，可以使用以下公式对输入向量进行归一化处理。

$$x = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (14)$$

式中： x_{\min} 和 x_{\max} 分别是神经网络输入、输出数据的最小值和最大值。当网络输出预测结果后，将预测数据反归一化处理，即可得到有实际意义预测数据。

3.3 网络设计与训练

LSTM 神经网络的训练过程如图 14 所示，训练数据和测试数据划分比例为 0.85，测试数据窗口长度 n 为 50，即输入为前 49 个的电压、电流和温度数据，输出为当前时刻 SOC 的预测值。

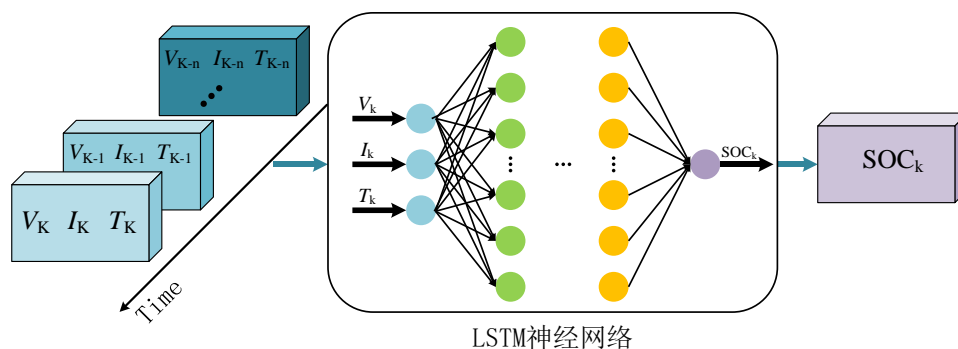


图 14 LSTM 估计 SOC 离线训练框图

4 实验结果与分析

4.1 LSTM的SOC估算结果

LSTM 层的细胞个数对训练有着重要的影响, 分别选取 50-300 的细胞个数对网络进行训练, 训练结果如图 15 所示。考虑 LSTM 网络的复杂度、单步预测消耗时间和预测的准确性, 采用平均绝对误差(MAE)、均方根误差(RMSE)和单步预测时间对四种 LSTM 模型进行了对比, 对比情况见表 3 和图 15。

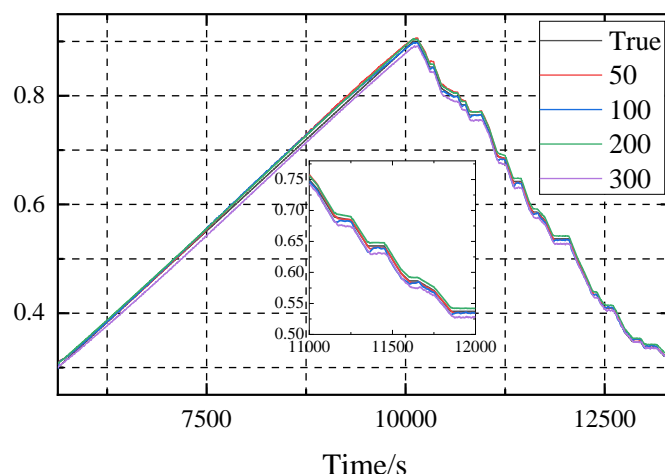


图 15 不同 LSTM 网络 SOC 训练结果

表 3 不同 LSTM 网络下的性能指标

LSTM 细胞数	50	100	200	300
MAE	0.0051	0.0042	0.0056	0.0062
RMSE	0.0065	0.0045	0.0058	0.0068
最大绝对误差	0.0156	0.0097	0.0091	0.0136
单步耗时/s	0.046	0.057	0.068	0.076

4.3 LSTM-UKF的SOC估算结果

在表现最好的 LSTM 网络中引入 UKF 后, LSTM-UKF 算法对 SOC 的预测结果如图 16 至图 19 所示。

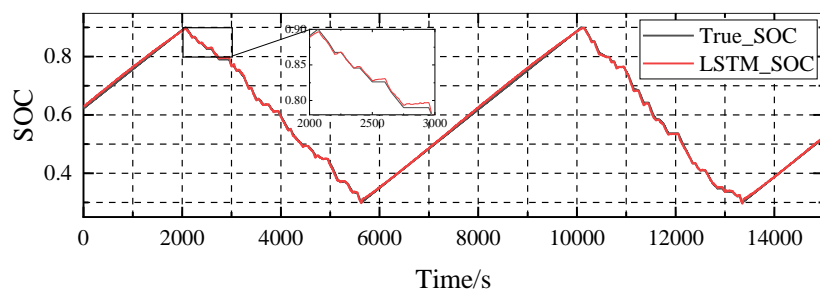


图 16 LSTM 估计结果

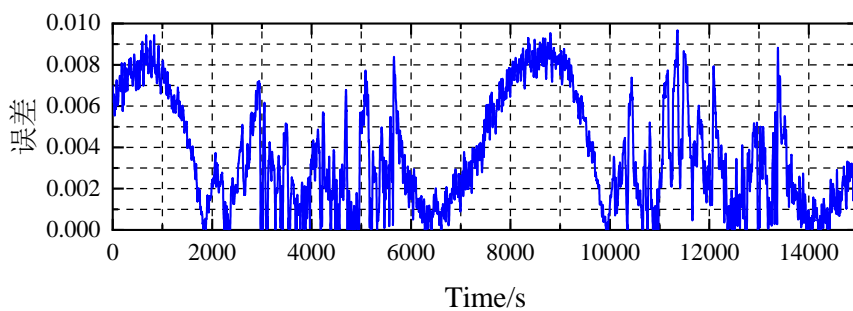


图 17 LSTM 估计误差

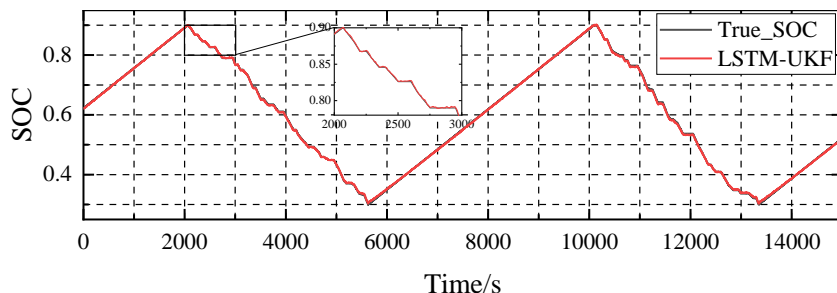


图 18 LSTM-UKF 估计结果

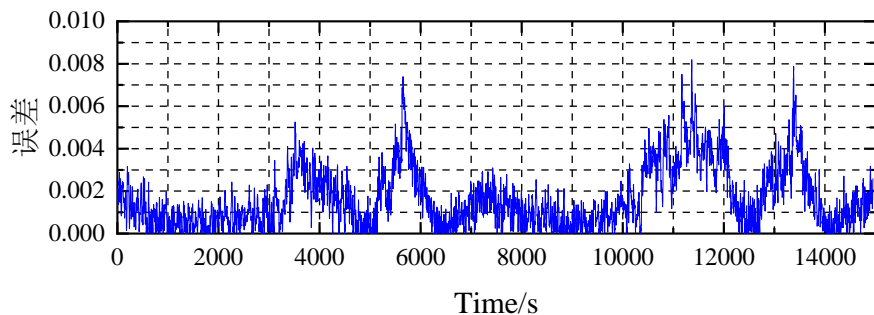


图 19 LSTM-UKF 估计误差

在引入 UKF 算法后 LSTM 网络的误差明显降低，预测的稳定性得到增强。LSTM-UKF 算法的预测平均绝对误差(MAE)从 0.42%下降到 0.17%，最大绝对误差为 0.8195%，误差 RMSE 为 0.22%，单步预测时间为 0.062s。模型在 MAE 减少 59.52%情况下，单步预测时间仅增加了 8.8%，模型的预测结果符合实际需求。

5 总结

本文对全电船舶锂电池 SOC 估算的 UKF 方法和 LSTM 方法分别进行探讨。根据锂电池的电压、电流和温度数据对模型进行训练，选取预测结果最优的 LSTM 网络，在网络中引入了 UKF 算法。UKF 算法滤除了 LSTM 网络输出的噪声，大大减少了预测的误差，使得模型预测的稳定性进一步得到提高。LSTM 网络的引入解决了 UKF 进行 SOC 估算对 SOC 初始值依赖的问题，简化了 UKF 算法因初始状态偏差产生较大估算误差，需要多次迭代更新的过程，从而提升了 UKF 算法的实际应用能力。

参考文献

- [1] Pan Geng and Xiaoyan Xu and Tomasz Tarasiuk. State of Charge Estimation Method for Lithium-Ion Batteries in All-Electric Ships Based on LSTM Neural Network[J]. Polish Maritime Research, 2020, 27(3) : 100-108.
- [2] M.a.hannan,M.s.h.lipu,A.hussain,等. A review of lithium-ion battery state of charge estimation and management system in electric vehicle applications: Challenges and recommendations[J]. Renewable & Sustainable Energy Reviews, 2017, 2017(78): 834-854.
- [3] 黄小平,王岩. 卡尔曼滤波原理及应用: MATLAB仿真[M]. 北京: 电子工业出版社, 2015: 1-179.
- [4] 徐劲力,马国庆. 基于UKF的在线锂离子电池SOC估算研究[J]. 电源技术, 2019, 43(10): 1615-1618, 1644.
- [5] 郑永飞,文怀兴,韩昉,等. 基于 LSTM 神经网络的动力电池 SOC 估算研究[J]. 计算机应用与软件, 2020, 37(3): 78-81, 88.

作者简介：辛浩东（1997），男、硕士研究生。

地址：上海市浦东新区临港新城海港大道 1550 号

邮编：201306

电话：18221527369

E: johnxin2020@163.com