**引言**

设备的失效或故障会带来极大的安全隐患，对设备剩余使用寿命的预测和提前维护是非常有必要的。目前RUL（剩余使用寿命）预测模型主要有三类，基于经验知识的方法、基于数据驱动的方法、基于机理模型的方法。基于知识的方法有专家系统法、模糊系统法，因这些方法太过依赖于专家的先验知识导致泛化能力太差。基于机理模型的方法主要有基于马尔科夫模型的方法、基于卡尔曼滤波的方法、基于粒子滤波的方法，由于目前设备越来越趋于复杂化，导致很难建立精确的物理模型。基于数据驱动的方法是目前的研究热点，其中有统计分析方法和人工智能方法，统计学习方法主要有趋势外推法、AR方法等，人工智能方法主要有ANN、SVM以及近些年非常流行的深度学习方法。统计学习的方法要经过多步递推后再预测剩余寿命，所以无法保证实时性，而浅层人工智能方法存在特征抽取不精确的问题，所以目前基于深度学习的方法是RUL预测中最流行的方法之一。

传统的深度学习方法有堆叠自动编码器方法，该方法能够提取到更加精确的特征，但是由于RUL预测很明显的一个特点就是时间相关性，而AE方法无法提取到时序数据的时间相关性特征，因此后来人们提出循环神经网络模型（RNN）,RNN能够建立数据之间的时间相关性关系，但是随着时间的推移，后面节点对前面节点的感知能力下降，无法记住较长时间的信息，即存在梯度消失的问题。鉴于上述问题，长短期记忆神经网络（LSTM）在保持RNN网络结构的基础上重新设计了计算节点，使其能够记住较长时间之前的信息，使得其在语音识别等领域效果显著。然而LSTM在设备RUL预测上的应用是极少的，尤其是在重要领域发挥着非常关键作用的设备上RUL预测的研究和应用就更少了。比如说航天器上的关键部件，由于很难采样到全生命周期的数据，所以如何利用这些珍贵的非全生命周期数据，以及如何利用LSTM的优势对这类数据建立有效的RUL预测模型是非常重要的问题。

本文以网络参数优化为目标，提出了一种基于代价最小化的参数实时更新LSTM预测模型。能够在不具有设备的全生命周期数据情况下建立RUL预测模型，并随着设备服役时间的增加，利用采集到的数据对已建立的模型参数进行逐步优化更新。通过与多种典型时间序列预测模型的实验对比，验证了所提出的实时更新 LSTM 模型在小样本的时间序列预测中具有很强的适用性和更高的准确性。

The remaining part of this paper is organized as follows: Section 2 introduces LSTM theories and techniques; Section 3 introduces the proposed LSTM prediction model and its parameters real-time update method; Section 4 presents the experimental results and analysis of their results; Sections 5 conclude and look forward to future work.

**LSTM**

一般NN模型无法解决时间相关性问题，而RNN是一种改进的多层神经网络，由输入层、隐藏层、输出层组成。RNN可以被看作是同一个神经网络的多次复制，每个神经网络模块把消息传递给下一个，我们将其按时间展开后如图1所示，可以看出RNN是一种链式结构，而链式的特征本质上是与序列和列表相关的。当前隐藏层的输入不仅有当前时刻的输入还有上一时刻的隐藏层的输出，这就使得之前的信息可以向后传递，这也是RNN能够解决时间相关性的原因。RNN的参数训练是通过随时间反向传播（BPNN）算法实现的，但RNN的一个最大的问题就是训练时梯度需要随时间进行反向传播，当输入序列比较长时，会发生梯度消失和梯度爆炸问题。



Fig.1 Unfolding of single RNN hidden layer unit

LSTM通过引入一组记忆单元，允许网络学习何时遗忘历史信息，何时用新信息更新记忆单元，可以很好的解决上述RNN中存在的问题。LSTM与RNN具有同样的链式结构，不同的是，在标准的RNN中，如图1所示的时刻的隐藏层，实际上只有一个简单的tanh层，而在LSTM中却是四个交互的层，结构如图2所示。



Fig.2 Structure of LSTM unit

首先LSTM的最为关键的就是细胞状态, 细胞状态类似于传送带，信息可以在整个链上流传。其次由于LSTM有一些精心设计的门结构，所以具有去除或者增加信息到细胞状态的能力。门是一种让信息选择性通过的方法。门结构的核心是一个sigmoid神经网络层和一个pointwise乘法操作，sigmoid层的输出在0到1之间，0表示不允许任何量通过，1表示允许任意量通过。LSTM有三个门，分别是遗忘门、输入门、输出门，用这三个门来保护和控制细胞状态。

LSTM的网络训练过程描述如下：

步骤一、进行训练数据预处理（归一化等操作）。

步骤二、初始化网络的权重偏置参数以及网络的超参。

步骤三、计算公式5所示的LSTM的前向运算过程，得到LSTM的预测值。

步骤四、求出预测值与实际值之间误差（常见有均方误差等）。

步骤五、判断是否达到误差阈值或最大迭代次数，是则退出训练。否则利用BPTT算法进行网络参数的更新，然后转向步骤三循环进行，直到满足条件，退出循环。LSTM工作流程图如图3所示：



Fig.3 The training process of LSTM

**参数实时更新的LSTM预测模型**

本文从实际问题出发，给出在只有少量的非全生命周期样本下如何建立有效模型的方法。首先要合理利用已知历史数据建立LSTM预测模型，其次当实际在线数据来时能够利用已经建立好的预测模型得出相应的预测值，下一时刻新的数据来时即可作为该预测值的真实值，将求出的预测值与真实值的误差加入样本整体误差中，然后利用误差最小化方法对模型参数进行迭代更新，以这样循环更新参数的方法，使得模型随着时间的进行能够越来越准确。这种建模思想更有利于实际系统的在线监测，以及故障预警。当预测值达到故障点时即发出故障预警信号。

LSTM的第一步就是决定我们从细胞状态中丢弃什么信息，这是通过遗忘门实现的，该门会读取和，然后经过sigmoid后输出一个0到1之间的数值，这个数值再与每一个在细胞状态中的数字进行乘法操作，这样就是实现了从细胞状态中选择性丢弃信息的目的。第二步是确定什么信息可以被保存在细胞状态中，这里包括了两个部分，一是sigmoid输入门层，决定什么值我们将要更新。二是用tanh层创建一个新的候选值相量。该过程是确认更新的信息，并将这两个信息产生对状态的更新。第三步是就是更新旧的细胞状态了，即将更新为，首先将旧状态与遗忘门的输出相乘，丢弃掉不需要的信息。接着加上新的候选值。最后一步就是确定输出什么值，首先是通过一个sigmoid层（实际就是输出门，）确定细胞状态的哪些部分将输出出去，然后，把细胞状态通过tanh进行处理，并将它和输出门的输出相乘，最终确定输出哪些部分。总体来看，在时刻t时LSTM的更新方式如公式1所示。

 公式1

假设实际时间序列为，改进后的LSTM步骤如下：

**Step1**:将扩维成，其中n为时间序列长度，k是样本维度，样本个数为n-k+1，训练数据的标签为。 并对X进行归一化处理，以X的第一列为例：。

**Step2**: 初始化网络参数，例如遗忘门的权重和偏置，同理还有输入门，输出门等。并设置误差阈值Error\_Cost，最大迭代次数Max\_iter等超参。

**Step3**: 计算t时刻需要从t-1时刻的细胞状态遗忘的哪些信息。首先计算遗忘门的输出，然后将遗忘门的输出与t-1时的细胞状态进行乘法操作。

**Step4**:计算t时刻哪些输入信息可以留在t时刻的细胞状态中。这里包括了两个部分，一是输入门的输出，决定哪些值我们将要更新。二是用tanh层创建一个新的候选值相量 然后将输入门的输出与候选值相量进行乘法操做。

**Step5**:计算t时刻的细胞状态。是由遗忘门和输入门对细胞状态的共同作用的结果

**Step6**:计算t时刻的网络输出。首先计算输出门的输出然后将输出门的输出与当前时刻的细胞状态进行乘法操作得到当前时刻的输出 ，即当前时刻的预测值。重复步骤3至6算出所有训练样本的预测值。

**Step7**:求出所有样本预测值与真实值之间的误差error。如果error<Error\_Cost或当前迭代次数iter>Max\_iter，则结束训练，否则利用BPTT算法进行网络参数的更新，迭代次数加一，然后转向步骤三循环进行，直到达到误差阈值或最大迭代次数，退出循环。保存训练好的网络参数

**Step8**:新样本与进行步骤3-6所示LSTM的前向运算得到，利用该网络得出新样本的预测值，当采集到数据时，即可将作为预测值的真实值标签，计算总体误差，再利用BPTT算法将模型参数更新为，其中为学习率，和分别为各层神经元权值和偏置的梯度矩阵和向量。由于加入新样本时参数初始化即为历史样本的最优点附近，故只需简单几步的更新即可达到新样本下更优的全局最优解。

**Step9**:当下一时刻的预测值达到故障点时即发出故障预警，及时做出反应，防止更大的损失甚至安全事故发生。

改进LSTM流程图如图4所示：



Fig.4 The training process of improved LSTM

**实验分析**

本文以NASA埃姆斯中心的锂离子电池部分地面实验数据为基础，采用不同的预测算法，以及改进前后的LSTM算法，预测锂离子电池容量衰退到70%所进行的循环充放电周期，由此实现锂离子电池的寿命预测。

LSTM代价函数值随迭代次数的变化如图4所示，左边的是固定学习率为0.01时代价函数的变化曲线图，中间为时变学习率1/(10+sqrt(iter))的代价函数曲线图，右图为本文实际采用的方法的代价函数曲线图。



Fig.5 Training error of various methods

实验结果如图5所示，实线为实际SOH曲线，红色虚线为AR预测的SOH曲线、蓝色虚线是神经网络预测的SOH曲线、黄色虚线是LSTM预测的SOH曲线、黑色虚线是改进后的LSTM预测的SOH曲线.可以明显看到在具备历史数据的情况下，各种预测方法效果都很好，但是当不具备历史数据的情况下，几种传统方法的预测效果都不好，但是本文提出的方法可以很好解决这种情况，实验结果如图5中黑色虚线所示。



Fig.6 Battery health prognostics based on various methods model



Fig.7 MSE of Various Methods

**结论**

针对传统LSTM模型无法有效利用非全生命周期的数据建立优秀的RUL预测模型以及无法合理利用在线数据的问题,提出以误差最小化为目标的参数实时更新方法，以解决小样本数据的在线建模以及模型修正问题。以NASA锂离子电池地面试验采集的数据为例,验证了所提的LSTM 模型参数实时更新方法在可靠性预测领域中的适用性。同时证明该方法能够有效地用于锂离子电池寿命预测中,在工程应用方面具有较高的实际价值。基于目前的工作，后续我们可以展开进一步研究：比如从LSTM模型参数入手，寻求更有效的参数优化方法。此外，未来也可以将非数值型的数据比如说图片数据进行数据融合，使得抽取出来的特征更加全面，更准确的建立预测模型。