DOI: 10.3901/JME.2019.06.083

基于双隐层量子线路循环单元神经网络的状态退化趋势预测^{*}

李 锋^{1,2,3} 向 往¹ 陈 勇¹ 汤宝平² 王家序³

(1. 四川大学制造科学与工程学院 成都 610065;

2. 重庆大学机械传动国家重点实验室 重庆 400044;

3. 四川大学空天科学与工程学院 成都 610065)

摘要: 针对现有人工智能预测方法在旋转机械状态退化趋势预测中存在预测精度较差、计算效率较低等缺点, 提出基于双隐层量子线路循环单元神经网络(Double hidden layer quantum circuit recurrent unit neural network, DHL-QCRUNN)的旋转机械状态退化趋势预测方法。首先采用量纲一化排列熵误差构建状态退化特征集, 然后将该特征集输入 DHL-QCRUNN 以完成旋转机械状态退化趋势预测。在所提出的 DHL-QCRUNN 中, 设计双隐层结构以提高网络的非线性映射能力; 并引入量子相移门和多受控非门以实现信息的传递; 通过双隐层的量子反馈机制获得输入序列的整体记忆; 最后采用输出层激发态的概率幅表示输出, 通过以上方法改善了网络的非线性逼近能力和泛化性能, 使所提出的旋转机械状态退化趋势预测方法具有较高的预测精度。此外, 通过量子 Levenberg-Marquadt(LM)算法更新 DHL-QCRUNN 的网络参数, 提高该网络的收敛速度, 使所提出的状态退化趋势预测方法具有较高计算效率。滚动轴承状态退化趋势预测实例验证了该方法的有效性。提出了基于 DHL-QCRUNN 的旋转机械状态退化趋势预测新方法, 该方法具有较高的预测精度和较高的计算效率。

关键词: 双隐层量子线路循环单元神经网络; 量子计算; 排列熵误差; 趋势预测; 旋转机械

中图分类号: TP393; TH17

State Degradation Trend Prediction Based on Double Hidden Layer Quantum Circuit Recurrent Unit Neural Network

LI Feng^{1,2,3} XIANG Wang¹ CHEN Yong¹ TANG Baoping² WANG Jiaxu³

(1. School of Manufacturing Science and Engineering, Sichuan University, Chengdu 610065;

2. The State Key Laboratory of Mechanical Transmissions, Chongqing University, Chongqing 400044;

3. School of Aeronautics and Astronautics, Sichuan University, Chengdu 610065)

Abstract: In view of the shortcomings such as poor prediction accuracy and low computational efficiency of the existing prediction methods based on artificial intelligence in state degradation trend prediction of rotating machinery, a novel state degradation trend prediction method is proposed based on double hidden layer quantum circuit recurrent unit neural network (DHL-QCRUNN). In this method, a normalized permutation entropy error feature set is firstly constructed, and then this feature set is input into DHL-QCRUNN to accomplish state degradation trend prediction of rotating machinery. In the proposed DHL-QCRUNN, a double hidden layer structure is designed to raise the network nonlinear mapping ability; quantum phase-shift gates and multi-qubits controlled NOT gates are introduced to this network for information transfer; the overall memory of input sequences can be obtained by the quantum feedback mechanism in double hidden layers; moreover, the final output is described by the probability amplitudes of excited states in output layer. Therefore, the nonlinear approximation capability and generalization property of network are improved, and then the higher prediction accuracy of the proposed state degradation trend prediction method based on DHL-QCRUNN is obtained for rotating machinery. Besides, the network parameters can be renewed by the quantum Levenberg-Marquardt (LM) algorithm to improve convergence speed of DHL-QCRUNN, accordingly, higher computational efficiency can be obtained for the

^{*} 机械传动国家重点实验室开放基金(SKMT-KFKT-201718)、中国博士后科学基金第60批面上(2016M602685)和四川大学泸州市人民政府战略合作(2018CDLZ-30)资助项目。20180318 收到初稿, 20181212 收到修改稿

proposed trend prediction method. The example of state degradation trend prediction for rolling bearing demonstrates the effectiveness of the proposed method. A novel state degradation trend prediction method is proposed based on DHL-QCRUNN for rotating machinery, which owns higher prediction accuracy and higher computational efficiency.

Key words: double hidden layer quantum circuit recurrent unit neural network; quantum computation; permutation entropy error; trend prediction; rotating machinery

0 前言

机械系统状态预测是设备故障诊断中必不可少的一个环节,准确预测机械系统的状态退化趋势,可以为设备维护管理提供有价值的状态信息和足够的响应时间,避免因设备损坏造成的额外损失^[1]。经典的人工智能预测方法有多变量支持向量机(Multivariable support vector machine, MSVM)、贝叶斯预测模型(Bayesian prediction model, BPM)、极限学习机(Extreme learning machine, ELM)、人工神经网络(Artificial neural network, ANN)等。申中杰等^[2]提出一种用于滚动轴承剩余寿命预测的 MSVM 框架;HERP 等^[3]利用 BPM 进行风电机组轴承失效预测;MAO 等^[4]将 ELM 用于轴承非平稳故障的在线顺序预测;XU 等^[5]将反向传播神经网络(Back-propagation neural network, BPNN)应用于高速钢钻头磨损预测。此外,循环神经网络(Recurrent neural network, RNN)^[6-7]在这类时间序列预测中也有广泛应用。然而,上述预测方法仍存在诸多不足。对于 MSVM,核函数和参数都是人工设定的,在很多情况下预测结果都不确定;由于 BPM 往往不能及时找出最需要推理的因素,无法使该因素及时得到预测值;ELM 的输入权值和隐含层偏差随机产生,建模结果具有较大的随机性;经典神经网络如 BPNN 和 RNN 往往收敛效果差,学习速度慢,网络学习记忆稳定性差,不能获得理想的输出。以上不足使现有的方法在旋转机械状态退化趋势预测中存在预测精度较差、计算效率较低等缺点。

作为一种极富前景的非线性模型,量子计算被认为是改进神经计算的有效途径之一^[8]。量子过程神经网络^[9-11]利用了量子计算的优势特别是其并行计算特性,具有比经典神经网络更强的并行计算能力和更大的数据处理能力^[12];另外,量子过程神经网络能够直接处理过程数据,具有时变信息处理能力和时变系统输入与输出之间的非线性映射能力,适于解决状态退化趋势预测这类与时变过程相关的问题。本文提出了一种新型量子过程神经网络——双隐层量子线路循环单元神经网络(Double hidden layer quantum circuit recurrent unit neural network, DHL-QCRUNN)。在 DHL-QCRUNN 中,

设计双隐层结构以提高复杂输入输出关系条件下网络的非线性映射能力;在网络中引入了量子相移门和多位受控非门以实现信息从输入层传递到输出层;DHL-QCRUNN 还利用双隐层的量子反馈机制来获得输入序列的整体记忆。另一方面,通过量子 Levenberg-Marquardt(LM)算法更新量子相移门相位以提高 DHL-QCRUNN 网络参数的更新速度。因此,与其他人工智能方法相比, DHL-QCRUNN 具有更好的非线性逼近能力、泛化性能和更快的收敛速度。

利用上述 DHL-QCRUNN 的优点,本文提出基于 DHL-QCRUNN 的旋转机械状态退化趋势预测方法。将该方法用于旋转机械状态退化趋势预测,可达到较高的预测精度和计算效率。

1 双隐层量子线路循环单元神经网络

1.1 量子相移门

量子相移门记为 $R(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$, 其线路表示如图 1 所示。假设量子比特的初始状态是 $|\phi_0\rangle = \begin{pmatrix} \cos\theta_0 \\ \sin\theta_0 \end{pmatrix}$, 则 $|\phi_0\rangle$ 的状态可通过 $R(\theta)$ 变换如下

$$|\phi'_0\rangle = R(\theta)|\phi_0\rangle = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \cos\theta_0 \\ \sin\theta_0 \end{pmatrix} = \begin{pmatrix} \cos(\theta_0 + \theta) \\ \sin(\theta_0 + \theta) \end{pmatrix} \quad (1)$$

式中, $|\phi'_0\rangle$ 表示经过相移变换后的最新状态。

图1 相移门线路

1.2 多位受控非门

在真实的量子系统中,单个量子比特的状态往往受多个量子比特的联合控制^[13]。这种动态行为可以用多位受控非门 $C^n(X)$ (图 2) 来描述。其中, $|x_1\rangle, |x_2\rangle, \dots, |x_n\rangle$ 是控制比特, $|\bar{\phi}\rangle$ 是输入目标比特, $|\phi\rangle$ 是输出目标比特。 $C^n(X)$ 定义如下

$$C^n(X)|x_1x_2\cdots x_n\rangle|\bar{\phi}\rangle = |x_1x_2\cdots x_n\rangle X^{x_1x_2\cdots x_n}|\bar{\phi}\rangle \quad (2)$$

式中, $X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ 是单比特量子非门; X 的指数表

示全为 $1(|x_i\rangle=|1\rangle)$, 则多位受控非门 $C^n(X)$ 作用于 $|\bar{\phi}\rangle$ 使其翻转得到 $|\phi\rangle$ 。

图2 多位受控非门

一般控制量子比特可表示为 $|x_i\rangle = \alpha_i|0\rangle + \beta_i|1\rangle$, 目标量子比特输入态可表示为 $|\bar{\phi}\rangle = \bar{\alpha}|0\rangle + \bar{\beta}|1\rangle$, 则目标量子比特输入态可表示为 $|\bar{\phi}\rangle = \bar{\alpha}|0\rangle + \bar{\beta}|1\rangle$, 则 $C^n(X)$ 的输出推导如下

$$\begin{aligned} C^n(X)|x_1x_2\cdots x_n\rangle|\bar{\phi}\rangle &= |x_1x_2\cdots x_n\rangle X^{x_1x_2\cdots x_n}|\bar{\phi}\rangle = \\ &|x_1\rangle\otimes|x_2\rangle\otimes\cdots\otimes|x_n\rangle\otimes|\bar{\phi}\rangle + \\ &\beta_1\beta_2\cdots\beta_n\bar{\beta}\left|\overbrace{11\cdots10}^n\right\rangle + \beta_1\beta_2\cdots\beta_n\bar{\alpha}\left|\overbrace{11\cdots11}^n\right\rangle - \\ &\beta_1\beta_2\cdots\beta_n\bar{\alpha}\left|\overbrace{11\cdots10}^n\right\rangle - \beta_1\beta_2\cdots\beta_n\bar{\beta}\left|\overbrace{11\cdots11}^n\right\rangle \end{aligned} \quad (3)$$

式中, $\alpha_i^2 + \beta_i^2 = 1$; $\bar{\alpha}^2 + \bar{\beta}^2 = 1$; $\otimes$ 表示张量积。

由式(3)可知, $C^n(X)$ 的输出处于 $n+1$ 个量子比特的纠缠态中, $|\phi\rangle$ 处于状态 $|1\rangle$ 的概率为

$$P = \bar{\beta}^2 + (\beta_1\beta_2\cdots\beta_n)^2(\bar{\alpha}^2 - \bar{\beta}^2) \quad (4)$$

因此在 n 个控制比特的联合控制下, $|\phi\rangle$ 可以表示为

$$|\phi\rangle = \sqrt{1-P}|0\rangle + \sqrt{P}|1\rangle = \cos\varphi|0\rangle + \sin\varphi|1\rangle \quad (5)$$

式中, $\varphi = \arcsin\sqrt{P}$ ($P \in [0,1], \varphi \in [0,0.5\pi]$) 为 $|\phi\rangle$ 的相位。

1.3 量子线路循环单元

量子线路循环单元由量子相移门和多位受控非门组成, 如图3所示。其中 $|x_t^i\rangle$ 是 t 时刻的控制比特, $|\bar{\phi}_t\rangle$ 是 t 时刻的输入目标比特, $(\varphi_x)_i$ 是第 i 个量子相移门的相位, $|\phi_t\rangle$ 是 t 时刻的输出目标比特(也为下一时刻输入目标比特, 即 $|\bar{\phi}_{t+1}\rangle = |\phi_t\rangle$)。令 $t=1,2,\cdots,T$ 表示 T 个采样时间点, $|x_t^i\rangle$ 可表示为

$$|x_t^i\rangle = \cos(\theta_x)_i^i|0\rangle + \sin(\theta_x)_i^i|1\rangle \quad (6)$$

图3 量子线路循环单元

输出目标比特 $|\phi_t\rangle$ 可表示为

$$|\phi_t\rangle = \cos\varphi_t|0\rangle + \sin\varphi_t|1\rangle \quad (7)$$

则根据式(1)~(5)可推导出 $|\phi_t\rangle$ 的相位如下

$$\begin{aligned} \varphi_t &= \arcsin\sqrt{P} = \arcsin\left\{\sin^2\varphi_{t-1} + \right. \\ &\left.\left[\prod_{i=1}^m \sin^2\left((\theta_x)_i^i + (\varphi_x)_i\right)\right]\left(\cos^2\varphi_{t-1} - \sin^2\varphi_{t-1}\right)\right\}^{\frac{1}{2}} = \\ &\arcsin\left(\sqrt{\sin^2\varphi_{t-1} + S_r \cos 2\varphi_{t-1}}\right) \end{aligned} \quad (8)$$

式中, $S_r = \prod_{i=1}^m \sin^2\left((\theta_x)_i^i + (\varphi_x)_i\right)$, 当 $t=1$ 时, $\varphi(t-1)=0$ 。

1.4 双隐层量子线路循环单元神经网络(DHL-QCRUNN)模型

在 DHL-QCRUNN 中, 设计双隐层结构以获得输入序列的整体记忆, 并采用输出层激发态的概率幅来表示最终输出。其线路表达如图4所示, 其中 $\bar{x}_t^1, \bar{x}_t^2, \cdots, \bar{x}_t^m$ 表示量纲一化输入序列; $|x_t^1\rangle, |x_t^2\rangle, \cdots, |x_t^m\rangle$ 表示输入比特; $|\bar{h}_1\rangle_t^1, |\bar{h}_1\rangle_t^2, \cdots, |\bar{h}_1\rangle_t^p$ 表示第一隐层输入; $|(h_1)_t\rangle^1, |(h_1)_t\rangle^2, \cdots, |(h_1)_t\rangle^p$ 表示第一隐层输出; $|\bar{h}_2\rangle_t^1, |\bar{h}_2\rangle_t^2, \cdots, |\bar{h}_2\rangle_t^q$ 表示第二隐层输入; $|(h_2)_t\rangle^1, |(h_2)_t\rangle^2, \cdots, |(h_2)_t\rangle^q$ 表示第二隐层输出; $|y_t^1\rangle, |y_t^2\rangle, \cdots, |y_t^n\rangle$ 表示输出比特; $\bar{y}_t^1, \bar{y}_t^2, \cdots, \bar{y}_t^n$ 表示最终输出

设 $|x_t^i\rangle \in (|x_t^1\rangle, |x_t^2\rangle, \cdots, |x_t^m\rangle)^T$ 为

$$|x_t^i\rangle = \cos(\theta_x)_i^i|0\rangle + \sin(\theta_x)_i^i|1\rangle \quad (9)$$

式中, $(\theta_x)_i^i = 0.5\pi\bar{x}_t^i$ 为 $|x_t^i\rangle$ 的相位, $\bar{x}_t^i$ 为量纲一化输入样本。

根据量子线路循环单元的输入输出关系, 第一隐层和第二隐层的输入分别表示为

$$|\bar{h}_1\rangle_t^j = \begin{cases} |0\rangle, & (t=1) \\ |\bar{h}_1\rangle_{t-1}^j, & (t \geq 2) \end{cases}$$

$$|\bar{h}_2\rangle_t^k = \begin{cases} |0\rangle, & (t=1) \\ |\bar{h}_2\rangle_{t-1}^k, & (t \geq 2) \end{cases}$$

$$j=1,2,\cdots,p, \quad k=1,2,\cdots,q。$$

记第一隐层输出为

$$|(h_1)_t\rangle^j = \cos(\theta_{h1})_t^j|0\rangle + \sin(\theta_{h1})_t^j|1\rangle \quad (10)$$

根据式(6)~(8), 可推导出 $|(h_1)_t\rangle^j$ 的相位如下

图4 双隐层量子线路循环单元神经网络

$$(\theta_{h1})_t^j = \arcsin \sqrt{P} =$$

$$\arcsin \sqrt{\sin^2(\theta_{h1})_{t-1}^j + (S_q)_t^j \cos(2(\theta_{h1})_{t-1}^j)} \quad (11)$$

式中, $(\theta_{h1})_t^j$ 为 $|(\theta_{h1})_t^j\rangle$ 的相位, 当 $t=1$ 时 $(\theta_{h1})_{t-1}^j = 0$;

$(S_q)_t^j = \prod_{i=1}^m \sin^2((\theta_x)_t^i + (\varphi_x)_{ij})$, 其中 $(\varphi_x)_{ij}$ 表示输入层量子相移门的相位。

同理, 第二隐层输出为

$$|h_2^k(t)\rangle = \cos(\theta_{h2})_t^k |0\rangle + \sin(\theta_{h2})_t^k |1\rangle \quad (12)$$

$$(\theta_{h2})_t^k = \arcsin \sqrt{\sin^2(\theta_{h2})_{t-1}^k + (S_r)_t^k \cos(2(\theta_{h2})_{t-1}^k)} \quad (13)$$

式中, $(\theta_{h2})_t^k$ 为 $|(\theta_{h2})_t^k\rangle$ 的相位; 当 $t=1$ 时, $(\theta_{h2})_{t-1}^k = 0$;

$(S_r)_t^k = \prod_{j=1}^p \sin^2((\theta_{h1})_t^j + (\varphi_{h1})_{jk})$, 其中 $(\varphi_{h1})_{jk}$ 表示第一隐层量子相移门的相位。

网络最终输出 $|y_t^l\rangle \in (|y_t^1\rangle, |y_t^2\rangle, \dots, |y_t^n\rangle)^T$ 为

$$|y_t^l\rangle = \cos(\theta_y)_t^l |0\rangle + \sin(\theta_y)_t^l |1\rangle \quad (14)$$

式中, $(\theta_y)_t^l$ 为 $|y_t^l\rangle$ 的相位, $l=1, 2, \dots, n$ 。

根据式(1)~(5), $(\theta_y)_t^l$ 可表示为

$$(\theta_y)_t^l = \arcsin \sqrt{P} =$$

$$\arcsin \left(\prod_{k=1}^q \sin((\theta_{h2})_t^k + (\varphi_{h2})_{kl}) \right) \quad (15)$$

式中, $(\varphi_{h2})_{kl}$ 表示第二隐层量子相移门的相位。

为简化计算过程, 用 $|y_t^l\rangle$ 处于状态 $|1\rangle$ 的概率幅来表示最终输出, 即

$$\bar{y}_t^l = \sin(\theta_y)_t^l = \prod_{k=1}^q \sin((\theta_{h2})_t^k + (\varphi_{h2})_{kl}) \quad (16)$$

1.5 DHL-QCRUNN 的学习算法

在 DHL-QCRUNN 的训练过程中, 每一步训练

都要更新相位 $(\varphi_x)_{ij}$ 、 $(\varphi_{h1})_{jk}$ 、 $(\varphi_{h2})_{kl}$ 。记第 s 步训练中输入层和隐层量子相移门的相位为 $(\varphi_x)_{ij}(s)$ 、 $(\varphi_{h1})_{jk}(s)$ 、 $(\varphi_{h2})_{kl}(s)$, 相应的增量为 $\Delta(\varphi_x)_{ij}(s)$ 、 $\Delta(\varphi_{h1})_{jk}(s)$ 、 $\Delta(\varphi_{h2})_{kl}(s)$, 则更新后的相位为

$$(\varphi_x)_{ij}(s+1) = (\varphi_x)_{ij}(s) + \Delta(\varphi_x)_{ij}(s) \quad (17)$$

$$(\varphi_{h1})_{jk}(s+1) = (\varphi_{h1})_{jk}(s) + \Delta(\varphi_{h1})_{jk}(s) \quad (18)$$

$$(\varphi_{h2})_{kl}(s+1) = (\varphi_{h2})_{kl}(s) + \Delta(\varphi_{h2})_{kl}(s) \quad (19)$$

本文采用量子 Levenberg-Marquardt(LM)算法更新 $(\varphi_x)_{ij}$ 、 $(\varphi_{h1})_{jk}$ 、 $(\varphi_{h2})_{kl}$, 设 DHL-QCRUNN 的逼近误差函数为

$$E_t = \frac{1}{2} \sum_{l=1}^n (e_t^l)^2 = \frac{1}{2} \sum_{l=1}^n (y_t^l - \bar{y}_t^l)^2 \quad (20)$$

式中, y_t^l 和 $\bar{y}_t^l$ 分别表示实际输出和期望输出, 显然, E_t 是关于 $(\varphi_x)_{ij}$ 、 $(\varphi_{h1})_{jk}$ 、 $(\varphi_{h2})_{kl}$ 的一致连续函数。

根据梯度下降法, $(\varphi_x)_{ij}$ 的梯度推导如下

$$\frac{\partial e_t^l}{\partial (\varphi_x)_{ij}} =$$

$$\frac{\partial e_t^l}{\partial (\theta_{h2})_t^k} \left(\frac{\partial (\theta_{h2})_t^k}{\partial (\theta_{h1})_t^j} \left(\frac{\partial (\theta_{h1})_t^j}{\partial (\varphi_x)_{ij}} + \frac{\partial (\theta_{h1})_t^j}{\partial (\theta_{h1})_{t-1}^j} \frac{\partial (\theta_{h1})_{t-1}^j}{\partial (\varphi_x)_{ij}} \right) + \frac{\partial (\theta_{h2})_t^k}{\partial (\theta_{h2})_{t-1}^k} \frac{\partial (\theta_{h2})_{t-1}^k}{\partial (\theta_{h1})_{t-1}^j} \frac{\partial (\theta_{h1})_{t-1}^j}{\partial (\varphi_x)_{ij}} \right) \quad (21)$$

$$\frac{\partial e_t^l}{\partial (\theta_{h2})_t^k} = -y_t^l \cot((\theta_{h2})_t^k + (\varphi_{h2})_{kl}) \quad (22)$$

$$\frac{\partial (\theta_{h2})_t^k}{\partial (\theta_{h1})_t^j} =$$

$$2(S_r)_t^k \frac{\cos(2(\theta_{h2})_{t-1}^k) \cot((\theta_{h1})_t^j + (\varphi_{h1})_{jk})}{\sin(2(\theta_{h2})_t^k)} \quad (23)$$

$$\frac{\partial(\theta_{h1})_t^j}{\partial(\varphi_x)_{ij}} = 2(S_q)_t^j \frac{\cos(2(\theta_{h1})_{t-1}^j) \cot((\theta_x)_t^i + (\varphi_x)_{ij})}{\sin(2(\theta_{h1})_t^j)} \quad (24)$$

$$\frac{\partial(\theta_{h2})_t^k}{\partial(\theta_{h2})_{t-1}^k} = (1 - 2(S_r)_t^k) \frac{\sin(2(\theta_{h2})_{t-1}^k)}{\sin(2(\theta_{h2})_t^k)} \quad (25)$$

$$\frac{\partial(\theta_{h1})_t^j}{\partial(\theta_{h1})_{t-1}^j} = (1 - 2(S_q)_t^j) \frac{\sin(2(\theta_{h1})_{t-1}^j)}{\sin(2(\theta_{h1})_t^j)} \quad (26)$$

同理, $(\varphi_{h1})_{jk}$ 的梯度为

$$\frac{\partial e_t^l}{\partial(\varphi_{h1})_{jk}} = \frac{\partial e_t^l}{\partial(\theta_{h2})_t^k} \left(\frac{\partial(\theta_{h2})_t^k}{\partial(\varphi_{h1})_{jk}} + \frac{\partial(\theta_{h2})_t^k}{\partial(\theta_{h2})_{t-1}^k} \frac{\partial(\theta_{h2})_{t-1}^k}{\partial(\varphi_{h1})_{jk}} \right) \quad (27)$$

$$\frac{\partial(\theta_{h2})_t^k}{\partial(\varphi_{h1})_{jk}} = 2(S_r)_t^k \frac{\cos(2(\theta_{h2})_{t-1}^k) \cot((\theta_{h1})_t^j + (\varphi_{h1})_{jk})}{\sin(2(\theta_{h2})_t^k)} \quad (28)$$

同理, $(\varphi_{h2})_{kl}$ 的梯度为

$$\frac{\partial e_t^l}{\partial(\varphi_{h2})_{kl}} = -y_t^l \cot((\theta_{h2})_t^k + (\varphi_{h2})_{kl}) \quad (29)$$

令 $\mathbf{p}$ 表示由输入层和双隐层量子相移门的相位组成的矢量, $\mathbf{v}_t$ 表示误差矢量, $\mathbf{J}_t$ 表示导数矩阵, 其定义分别如下

$$\mathbf{p} = ((\varphi_x)_{11}, \dots, (\varphi_x)_{mp}, (\varphi_{h1})_{11}, \dots, (\varphi_{h1})_{pq}, \dots, (\varphi_{h2})_{11}, \dots, (\varphi_{h2})_{qn})^T \quad (30)$$

$$\mathbf{v}_t = (e_t^1, e_t^2, \dots, e_t^n)^T \quad (31)$$

式中, $e_t^l = y_t^l - \bar{y}_t^l$, $l = 1, 2, \dots, n$,

$$\mathbf{J}_t = \begin{pmatrix} \frac{\partial e_t^1}{\partial(\varphi_x)_{11}}, \dots, \frac{\partial e_t^1}{\partial(\varphi_x)_{mp}}, \dots, \frac{\partial e_t^1}{\partial(\varphi_{h2})_{11}}, \dots, \frac{\partial e_t^1}{\partial(\varphi_{h2})_{qn}} \\ \frac{\partial e_t^2}{\partial(\varphi_x)_{11}}, \dots, \frac{\partial e_t^2}{\partial(\varphi_x)_{mp}}, \dots, \frac{\partial e_t^2}{\partial(\varphi_{h2})_{11}}, \dots, \frac{\partial e_t^2}{\partial(\varphi_{h2})_{qn}} \\ \vdots \\ \frac{\partial e_t^n}{\partial(\varphi_x)_{11}}, \dots, \frac{\partial e_t^n}{\partial(\varphi_x)_{mp}}, \dots, \frac{\partial e_t^n}{\partial(\varphi_{h2})_{11}}, \dots, \frac{\partial e_t^n}{\partial(\varphi_{h2})_{qn}} \end{pmatrix} \quad (32)$$

通过上述准备, 利用量子 LM 算法更新 $\mathbf{p}$

$$\mathbf{p}(s+1) = \mathbf{p}(s) + \Delta \mathbf{p}(s) =$$

$$\mathbf{p}(s) - \sum_{t=1}^T \frac{(\mathbf{J}_t)^T \mathbf{v}_t}{(\mathbf{J}_t)^T \mathbf{J}_t + \mu \mathbf{I}} \quad (33)$$

式中 $\Delta \mathbf{p} = (\Delta(\varphi_x)_{11}, \dots, \Delta(\varphi_x)_{mp}, \dots, \Delta(\varphi_{h2})_{11}, \dots, \Delta(\varphi_{h2})_{qn})$ 为 $\mathbf{p}$ 的增量, $\mathbf{I}$ 为单位矩阵, μ 是确保矩阵 $(\mathbf{J}_t)^T \mathbf{J}_t + \mu \mathbf{I}$ 可逆的一个微小正数。

至此, DHL-QCRUNN 中所有相位的增量均已求得, 将 $\Delta \mathbf{p}$ 里的增量分别代入式(17)~(19), 即可实现各个相位的更新, 每更新一次相位就完成一次对 DHL-QCRUNN 的训练。

2 基于 DHL-QCRUNN 的旋转机械状态退化趋势预测方法

所提出的基于 DHL-QCRUNN 的状态退化趋势预测方法实现流程如图 5 所示。该流程说明如下:

- ① 为所有采集的旋转机械振动样本数据构造排列熵^[14];
- ② 将排列熵输入 DHL-QCRUNN 中进行训练, 完成排列熵预测;
- ③ 计算各时间点实际的排列熵值与该点的排列熵预测值的误差, 并将该误差值作为状态退化特征;
- ④ 将排列熵误差输入 DHL-QCRUNN 中对网络进行训练, 并完成旋转机械状态退化趋势预测。

图 5 基于 DHL-QCRUNN 的状态退化预测方法实现流程

在训练 DHL-QCRUNN 之前, 通过相空间重构将样本数据 $\mathbf{s} = (x_a, x_{a+1}, \dots, x_b)$ 分为训练样本集 $\mathbf{X}_{\text{train}}$ 和相应的期望输出集 $\mathbf{y}_{\text{train}}$ (即训练目标集)两部分, 其表达式分别如下

$$\mathbf{X}_{\text{train}} = \begin{pmatrix} x_a & x_{a+1} & \cdots & x_{b-m} \\ x_{a+1} & x_{a+2} & \cdots & x_{b-m+1} \\ \vdots & \vdots & & \vdots \\ x_{a+m-1} & x_{a+m} & \cdots & x_{b-1} \end{pmatrix} \quad (34)$$

$$\mathbf{y}_{\text{train}} = (x_{a+m} \quad x_{a+m+1} \quad \cdots \quad x_b) \quad (35)$$

式中, $\mathbf{X}_{\text{train}}$ 的每一列表示一个 $m(m < b-a)$ 维训练样本, 其训练目标是 $\mathbf{y}_{\text{train}}$ 对应列中的元素。显然, 训练样本总数为 $b-a-m+1$ 。

DHL-QCRUNN 训练过程如下: ① 在第一步训练中初始化相位, 把 $[0, 2\pi]$ 内的值随机赋给 $(\varphi_x)_{ij}$ 、 $(\varphi_{h1})_{jk}$ 、 $(\varphi_{h2})_{kl}$; ② 将训练样本集输入 DHL-QCRUNN 计算输出 $\bar{\mathbf{y}}_{\text{train}} = [\bar{x}_{a+m} \quad \bar{x}_{a+m+1} \quad \cdots \quad \bar{x}_b]$; ③ 利用量子 LM 算法更新 $(\varphi_x)_{ij}$ 、 $(\varphi_{h1})_{jk}$ 、 $(\varphi_{h2})_{kl}$ 以得到下一个训练周期的初始相位; ④ 重复步骤(2)和步骤(3)直到 $\mathbf{y}_{\text{train}}$ 和 $\bar{\mathbf{y}}_{\text{train}}$ 的均方误差小于设定的阈值, 或达到最大训练步数。

DHL-QCRUNN 训练完成后, 采用多步向前预测法进行状态退化趋势预测, 预测过程如下: 首先, 将训练集最后 m 个样本 $(x_{b-m+1}, x_{b-m+2}, \cdots, x_b)$ 作为 DHL-QCRUNN 的初始输入, 得到 $b+1$ 时刻的输出 $\bar{x}_{b+1}$; 然后, 用 $(x_{b-m+2}, x_{b-m+3}, \cdots, \bar{x}_{b+1})$ 作为 DHL-QCRUNN 的输入, 得到 $b+2$ 时刻的输出 $\bar{x}_{b+2}$; 以此类推, 向 DHL-QCRUNN 输入 $(\bar{x}_{b-m+N}, \bar{x}_{b-m+1+N}, \cdots, \bar{x}_{b-1+N})$, 得到 $b+N$ 时刻的输出 $\bar{x}_{b+N}$, 完成第 N 步预测, 至此得到共 N 个预测值。

随着旋转机械状态的不断恶化, 排列熵往往突变越来越明显, 以致实际的排列熵与其排列熵预测值之间的误差越来越大, 即排列熵误差随着旋转机械状态的不断恶化呈现整体逐渐递增的稳定趋势, 因此排列熵误差的可预测性较好。鉴于此, 为了能平稳、准确预测旋转机械全寿命状态退化的动态趋势, 本文采用排列熵误差作为状态退化特征。在 $b+N$ ($b=120, 121, \cdots, 932$) 时刻的排列熵误差为

$$(E_{\text{pe}})_{b+N} = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_{b+i} - \bar{x}_{b+i})^2} \quad (36)$$

式中, N 表示预测样本个数; x_{b+i} 和 $\bar{x}_{b+i}$ 分别表示 $b+i$ 时刻的排列熵实际值和排列熵预测值。

3 实例分析

采用 Cincinnati 大学实测的滚动轴承状态退化数据^[15]对本文所提方法进行验证, 试验装置如图 6 所示。轴承试验台的转轴上安装四个航空专用轴承, 这些航空轴承为 Rexnord 公司制造的 ZA-2115 双列滚子轴承。交流电动机通过带传动以 2 000 r/min 的恒定转速带动转轴旋转, 试验过程中轴承被施加 6 000 lbs 的径向载荷。采样频率为 20 kHz, 采样长度为 20 480 个点, 每隔 10 min 采集一次轴承的振动

数据, 轴承持续运行直到出现故障。试验台运行到第 7 天时 1 号轴承出现外圈故障而失效。本文采用采集到的该 1 号轴承状态退化数据验证本文所提方法。取 1 号轴承这七天的状态退化的振动数据(即 984 个样本)来构造排列熵, 排列熵的嵌入维数设为 $d=5$, 时间延迟设为 $\tau=3$ 。

图 6 滚动轴承状态退化试验装置

构造的排列熵如图 7 所示, 从第 548 点开始出现明显下降曲线, 表明轴承处于初始退化阶段(早期故障阶段)。从第 548 到第 697 点开始出现明显不规则曲线, 表明轴承出现了较大外圈缺陷, 在 697 点以后不规则曲线密集程度明显增多, 表明该轴承已出现严重故障, 濒临失效。

图 7 1 号轴承运转 7 天的排列熵曲线

这里, 首先通过 DHL-QCRUNN 预测排列熵, DHL-QCRUNN 的参数设定如下: 输入层维数 $n=8$; 第一隐层维数 $p=6$; 第二隐层维数 $q=10$; 输出层维数 $n=1$; 最大训练步数 $N_{\text{max}}=5\,000$; 均方误差阈值 $E_{\text{mse}}=0.0001$ 。预测样本数取 $N=50$ 。分别令

$$\begin{cases} a=1 \\ b=120 \end{cases}, \begin{cases} a=2 \\ b=121 \end{cases}, \cdots, \begin{cases} a=813 \\ b=932 \end{cases}$$

根据第 2 节的多步向前预测法使用 DHL-QCRUNN 来预测 $N=50$ 个排列熵值, 再根据式(36)计算出相应的排列熵误差 $((E_{\text{pe}})_{170}, \cdots, (E_{\text{pe}})_{982})$, 并经过量纲一化处理得到最终量纲一化的排列熵误差 $((E_{\text{pen}})_{170}, \cdots, (E_{\text{pen}})_{982})$, 结果如图 8 所示。

现在, 通过 DHL-QCRUNN 预测量纲一化排列熵误差。DHL-QCRUNN 的参数设定如下: 输入层

图 8 1 号轴承运转 7 天的量纲一化排列熵误差

数维数 $m=8$ ；第一隐层维数 $p=6$ ；第二隐层维数 $q=10$ ；输出层维数 $n=1$ ；最大训练步数 $N_{\max}=5000$ ；均方误差阈值 $E_{\text{mse}}=0.0001$ 。预测样本数取 $N=50$ 。图 8 中第 600 到 679 个采样点处于初始退化阶段，是状态退化趋势观测中最重要的时间序列数据。因此，取这 80 个样本的实际量纲一化排列熵误差进行预测验证，其中前 30 个样本点作为训练样本(即令 $a=600$ ， $b=629$)，后 50 个样本点作为测试样本，使用 DHL-QCRUNN 并根据第 2 节的多步向前预测方法来预测后 50 个测试样本的量纲一化排列熵误差值，最终预测曲线如图 9 所示。

图 9 DHL-QCRUNN 预测的量纲一化排列熵误差曲线

图 9 结果表明，被预测的在 $[630, 679]$ 区间内的

(a) BPNN 预测的量纲一化排列熵误差曲线

50 个量纲一化排列熵误差与图 8 对应的实际量纲一化排列熵误差非常接近，即由预测的量纲一化排列熵误差拟合的曲线与实际量纲一化排列熵误差的动态趋势(即实际状态退化趋势)一致。说明基于 DHL-QCRUNN 的趋势预测方法是有效的，可以将该方法用于双列滚子轴承的状态退化趋势预测。

为了验证所提出的基于 DHL-QCRUNN 的状态退化趋势预测方法的优势，首先将其状态退化趋势预测精度与 BPNN、RNN、长短时记忆神经网络(Longshort-term memory neural network, LSTMNN)、LS-SVM 和 ELM 进行比较。后五种预测方法也采用多步向前预测法的预测方式。BPNN、RNN、LSTMNN 的学习率均设为 $\alpha=0.5$ ，隐层节点数为 6，且这三种神经网络的其他网络参数，即最大训练步数和均方误差阈值的设置分别都与 DHL-QCRUNN 相同；为 LS-SVM 选择 RBF 核函数，通过留一交叉验证参数优化算法自动优化其核参数和正则化参数；ELM 的隐层节点数与上述三种神经网络一致，设为 6，激活函数为 Sigmoid 函数。采用这些方法得到的双列滚子轴承的状态退化趋势预测曲线如图 10a~10e 所示。

为进一步定量评价预测精度，采用方均根误差(RMSE)作为评价指标，其表达式为

$$e = \sqrt{\frac{1}{N} \sum_{i=1}^N \left((E_{\text{pen}})_{b+i} - (\bar{E}_{\text{pen}})_{b+i} \right)^2} \quad (37)$$

式中， N 表示预测样本个数； $(E_{\text{pen}})_{b+i}$ 和 $(\bar{E}_{\text{pen}})_{b+i}$ 分别表示第 $b+i$ 个样本点的实际的量纲一化排列熵误差和预测的量纲一化排列熵误差。

在 DHL-QCRUNN、BPNN、RNN、LSTMNN、LS-SVM 和 ELM 的参数设置保持不变的条件下，用这五种预测方法重复进行 100 次预测，取最小方均根误差 $e_{\min}$ 、最大方均根误差 $e_{\max}$ 、平均方均根

(b) RNN 预测的量纲一化排列熵误差曲线

(c) LSTMNN 预测的量纲一化排列熵误差曲线

(d) LS-SVM 预测的量纲一化排列熵误差曲线

(e) ELM 预测的量纲一化排列熵误差曲线

图 10 五种方法预测的量纲一化排列熵误差曲线

差 $\bar{e}$ 作为 100 次预测的预测误差评价指标, 结果见表 1。

表 1 6 种状态退化趋势预测方法的预测误差比较

预测方法	最小预测误差 $e_{\min}/\times 10^{-2}$	最大预测误差 $e_{\max}/\times 10^{-2}$	平均预测误差 $\bar{e}/\times 10^{-2}$
BPNN	1.80	81.87	11.93
RNN	4.32	14.13	10.42
LSTMNN	2.18	13.88	9.97
LS-SVM	2.81	24.69	10.56
ELM	2.33	46.19	13.81
DHL-QCRUNN	1.76	13.02	7.09

图 9、10 和表 1 结果表明, DHL-QCRUNN 的平均预测误差相对最小, 且预测误差的波动范围也相对最小, 说明 DHL-QCRUNN 具有良好的非线性逼近能力和泛化性能, 将其用于典型旋转机械——双列滚子轴承的性态退化趋势预测, 相比于 BPNN、RNN、LSTMNN、LS-SVM、ELM 可以取得更高的预测精度。

然后, 本文再用 BPNN、RNN 和 LSTMNN 与 DHL-QCRUNN 进行收敛速度对比。各神经网络参

数设置保持不变, 这四种神经网络的逼近误差如图 11 所示。

图 11 四种神经网络的逼近误差

由图 11 可知, 当训练步数增加时, 四种神经网络的逼近误差均不断下降, 但本文所提出的 DHL-QCRUNN 的逼近误差始终比 BPNN、RNN、LSTMNN 要小。

最后, 将 BPNN、RNN、LSTMNN、LS-SVM、ELM 进行状态退化趋势预测的计算时间(即训练时间与预测时间之和)与 DHL-QCRUNN 的计算时间

进行对比, 结果如图 12 所示。

图 12 六种状态退化趋势预测方法的计算时间比较

由图 12 可知, DHL-QCRUNN 的计算时间仅为 0.870 189 s, BPNN 的计算时间为 1.215 257 s, RNN 的计算时间为 1.252 406 s, LSTMNN 的计算时间为 4.856 815 s, LS-SVM 的计算时间为 0.494 650 s, ELM 的计算时间为 0.294 022 s。显然, DHL-QCRUNN 的计算时间比 BPNN、RNN 和 LSTMNN 明显短得多, 仅比 LS-SVM 和 ELM 略长(两者仍处于同一数量级)。

图 11、12 结果表明: 将 DHL-QCRUNN 用于双列滚子轴承的状态退化趋势预测比 BPNN、RNN 和 LSTMNN 具有更高的收敛速度和计算效率。

4 结论

(1) 在所提出的 DHL-QCRUNN 中, 通过设计双隐层结构, 构造量子相移门和多位控制非门并循环利用量子线路循环单元中的历史信息, 提高了网络输入量子态与输出量子态的匹配程度, 因此 DHL-QCRUNN 具有良好的非线性逼近能力和较理想的泛化性能。

(2) 在 DHL-QCRUNN 的学习训练过程中, 采用量子 LM 算法对输入层和双隐层中的量子相移门相位进行更新。由于结合了量子计算在并行处理和 LM 算法在求解优化问题上的优势, DHL-QCRUNN 具有较快的收敛速度和较低的计算成本。

(3) 由于 DHL-QCRUNN 在非线性逼近能力、泛化性能和收敛速度方面的优势, 基于 DHL-QCRUNN 的旋转机械状态退化趋势预测方法具有较高的预测精度和计算效率。

(4) 双列滚子轴承状态退化趋势预测实例验证了该方法的有效性。

参 考 文 献

- [1] LEI Y G, LIN P, GONTARZ S, et al. A model-based method for remaining useful life prediction of machinery[J]. IEEE Transactions on Reliability, 2016, 65(3): 1314-1326.
- [2] 申中杰, 陈雪峰, 何正嘉, 等. 基于相对特征和多变量支持向量机的滚动轴承剩余寿命预测[J]. 机械工程学报, 2013, 49(2): 183-189.
SHEN Zhongjie, CHEN Xuefeng, HE Zhengjia, et al. Remaining life predictions of rolling bearing based on relative features and multivariable support vector machine[J]. Journal of Mechanical Engineering, 2013, 49(2): 183-189.
- [3] HERP J, RAMEZANI, MPHAMMAD H, et al. Bayesian state prediction of wind turbine bearing failure[J]. Renewable Energy, 2018, 116: 164-172.
- [4] MAO W T, HE L, YAN Y J, et al. Online sequential prediction of bearings imbalanced fault diagnosis by extreme learning machine[J]. Mechanical Systems and Signal Processing, 2017, 83(15): 450-473.
- [5] XU J, YAMADA K, SEIKIYA K, et al. Effect of different features to drill-wear prediction with back propagation neural network[J]. Precision Engineering, 2014, 38(4): 791-798.
- [6] GUO L, LI N, JIA F, et al. A recurrent neural network based health indicator for remaining useful life prediction of bearings[J]. Neurocomputing, 2017, 240: 98-109.
- [7] XIANG W, LI F, WANG J X, et al. Quantum weighted gated recurrent unit neural network and its application in performance degradation trend prediction of rotating machinery[J]. Neurocomputing, 2018, 313: 85-95.
- [8] ESKANDARI E, AHMADI A, GOMAR S. Effect of spike-timing-dependent plasticity on neural assembly computing[J]. Neurocomputing, 2016, 191: 107-116.
- [9] CAO H, CAO F, WANG D. Quantum artificial neural networks with applications[J]. Information Sciences, 2015, 290: 1-6.
- [10] DA S, ADENILTON J, DE O, et al. Comments on "quantum artificial neural networks with applications"[J]. Information Sciences, 2016, 370-371: 120-122.
- [11] 李鹏华, 柴毅, 熊庆宇. 量子门 Elman 神经网络及其梯度扩展的量子反向传播学习算法[J]. 自动化学报, 2013, 39(9): 1511-1522.
LI Penghua, CHAI Yi, XIONG Qingyu. Quantum gate Elman neural network and its quantized extended gradient back-propagation training algorithm[J]. Acta Automatica Sinica, 2013, 39(9): 1511-1522.
- [12] LI P, XIAO H, SHANG F, et al. A hybrid quantum-inspired neural networks with sequence inputs[J]. Neurocomputing, 2013, 117: 81-90.
- [13] MAGALHAES S, LEAL V, HORTA I. Modelling the relationship between heating energy use and indoor

- temperatures in residential buildings through artificial neural networks considering occupant behavior[J]. *Energy and Buildings*, 2017, 151: 332-343.
- [14] YAN R Q, LIU Y B, GAO R X. Permutation entropy: A nonlinear statistical measure for status characterization of rotary machines[J]. *Mechanical Systems and Signal Processing*, 2012, 29: 474-484.
- [15] LEE J, QIU H, YU G. "Bearing data set", IMS, University of Cincinnati, NASA Ames prognostics data repository[DB/OL]. <http://ti.arc.nasa.gov/project/prognostics-data-repository>, NASA Ames, Moffett Field, CA, 2007.

作者简介: 李锋, 男, 1982 年出生, 博士, 副教授, 硕士研究生导师。主要研究方向为设备状态监测与故障诊断。

E-mail: lifeng19820501@163.com

向往, 男, 1993 年出生。主要研究方向为设备状态监测与故障诊断。

E-mail: 719970338@qq.com

陈勇, 男, 1993 年出生。主要研究方向为设备状态监测与故障诊断。

E-mail: 1919264568@qq.com

汤宝平(通信作者), 男, 1971 年出生, 博士, 教授, 博士研究生导师。

主要研究方向为设备状态监测与故障诊断、虚拟仪器、无线传感器网络。

E-mail: bptang@cqu.edu.cn

王家序, 男, 1954 年出生, 博士, 教授, 博士研究生导师。主要研究方向为机电传动与智能控制、空间飞行器机构。

E-mail: 313844507@qq.com