

基于结构自适应径向基神经网络的 油样光谱数据建模*

杨虞微 陈 果

(南京航空航天大学民航学院 南京 210016)

摘要 基于光谱分析数据的机械磨损状态预测有利于发现机械系统的早期磨损故障。由于神经网络对于非线性模型的辨识和非平稳信号的预测,与传统预测模型相比具有明显的优势,将神经网络预测方法运用于光谱分析,提出了基于神经网络预测的光谱分析监测技术。在预测模型中采用了在函数逼近、分类能力和学习速度均优于BP网络的径向基函数(RBF)神经网络模型,针对RBF网络的结构对于信号预测或模型辨识的精度具有影响很大的问题,提出了结构自适应RBF网络预测模型。利用遗传算法,对神经网络输入节点数、径向基函数分布系数及网络训练误差进行了优化,得到了最优的RBF网络预测模型。最后,对某航空发动机实际的光谱分析数据进行了预测和分析,并与ARMA模型进行了比较,结果充分表明了文中方法的有效性和优越性。

关键词 光谱油样分析 预测 RBF神经网络 遗传算法 状态监测

中图分类号 TH165.3 O329 **文献标识码** A **国家标准学科分类代码** 460.1540

Modeling of Spectrometric Oil Analysis Data Based on Structure Self-adaptive RBF Neural Network

Yang Yuwei Chen Guo

(Civil Aviation College, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China)

Abstract Forecasting machine wear state through spectrometric oil analysis(SOA) has an advantage to find out machine wear fault early. Because artificial neural network(ANN) possesses obvious advantages to traditional forecasting models for identifying non-linear system and forecasting non-even signal, the monitoring technique by SOA based on ANN forecasting was brought forward. In forecasting model, radial basis function(RBF) network is adopted because its performance is better than BP network at aspects of function approximating, classing, and learning speed. Aiming at the problem that the structure of RBF network has great effect on forecasting precision, a structure self-adaptive RBF neural network model is put forward. The genetic algorithm(GA) is used to optimize the node number of input layer, the spread coefficient of radial basis function, and training error, in order to obtain the optimum forecasting model of RBF network. Finally, the practical SOA data of an aero-engine are analyzed and forecasted by RBF network, and the results fully show superiority and effectivity of the new method.

Key words Spectrometric oil analysis(SOA) Forecasting RBF neural network Genetic algorithm(GA)
Condition monitoring

* 本文于2004年6月收到。

1 引言

光谱分析操作简单, 精度高, 分析速度快, 而且在有色金属元素的检测方面具有独特的优势, 因此广泛应用于机械设备的故障诊断中, 其应用领域涉及到机车柴油机、船舶柴油机、航空发动机及大型柴油机发电机组等机械设备的有色金属磨粒的检测和识别及早期故障监测。

由于机械磨损状态是一个逐步发展的过程, 因此可以利用光谱分析数据来进行机械系统磨损状态的预测, 以期尽早对故障的发展趋势进行预测和评估, 从而避免重大事故的发生和及时安排维修工作。

传统的 ARMA 预测模型要求时间序列是平稳的, 且服从正态分布, 文献 [1] 运用 Box-Jenkins 法建立了 ARMA 模型, 并对光谱分析数据进行了预测分析。但是机械系统本身具有很多非线性因素, 特别是机械系统出现故障时, 光谱分析得到的元素含量将出现严重的非平稳特征, 并表现出明显的随时间增长的磨损趋势, 在这种情况下, 传统的 ARMA 模型将不能对机械系统磨损状态进行准确的预测和评估。因此, 有必要寻找其他预测方法。

相关研究 [2,3] 表明: 神经网络能拟合任意的非线性函数并且具有很强泛化能力, 进行非线性预测具有很大的潜力。在神经网络模型中, 径向基函数 (radial basis function) 神经网络在函数逼近、分类能力和学习速度方面具有明显优势 [4], 因此广泛应用于非线性预测。

但是, RBF 神经网络应用到预测还有许多问题需要解决 [5], 其中最为突出的问题就是神经网络结构对预测精度有很大的影响, 难于确定最合适的网络结构。文中将 RBF 神经网络方法引入光谱分析, 提出结构自适应 RBF 神经网络预测模型, 利用遗传算法对网络结构进行优化, 从而实现对机械磨损故障的精确预测。

2 RBF 网络预测模型及预测精度的影响因素分析

设 1 个单变量时间序列 $\{x_1, x_2, \dots\}$, 对它进行预测的前提是认为其未来值与其前面的 m 个值之间有着某种函数关系, 可描述为:

$$x_{n+k} = F(x_n, x_{n-1}, \dots, x_{n-m+1}) \quad (1)$$

神经网络预测就是利用神经网络来拟合这种函数

关系 F , 并用它来推导未来的值。径向基函数 (radial basis function, RBF) 神经网络由 3 层组成, 其预测模型结构如图 1 所示。输入层只传递输入信号到隐层, 隐层节点由像高斯核函数那样的辐射状作用函数构成, 而输出层节点通常是简单的线性函数。

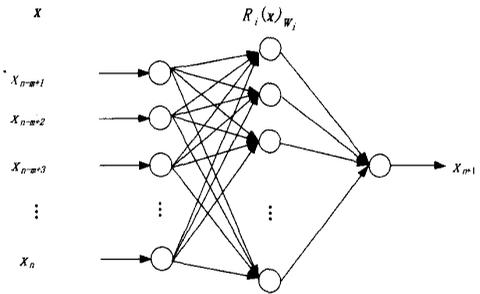


图 1 RBF 网络预测模型结构

隐层节点中的作用函数 (基函数) 对输入信号将在局部产生响应, 当输入信号靠近基函数的中央范围时, 隐层节点将产生较大的输出, 由此可见径向基函数网络具有局部逼近能力。

径向基函数有多种形式, 通常选择高斯函数:

$$R_i(x) = \exp\left[-\frac{\|x - c_i\|^2}{2\sigma_i^2}\right] \quad i = 1, 2, 3, \dots, k \quad (2)$$

式中: x 为 m 维输入向量, $x = \{x_{n-m+1}, x_{n-m+2}, \dots, x_n\}^T$; c_i 为第 i 个基函数的中心, 与 x 具有相同维数的向量; σ_i 为第 i 个感知的变量 (可以自由选择参数), 它决定了该基函数围绕中心点的宽度; k 为感知单元的个数。

由式 (2) 可见, $\|x - c_i\|$ 是向量 $x - c_i$ 的范数, 它通常表示 x 和 c_i 之间的距离, $R_i(x)$ 在 c_i 处有唯一的最大值, 随着 $\|x - c_i\|$ 的增大, $R_i(x)$ 迅速衰减到零。对于给定输入 $x \in R^m$, 只有一小部分靠近 x 的中心被激活。

输入层实现从 $x \rightarrow R_i(x)$ 的非线性映射, 输出层实现从 $R_i(x) \rightarrow x_{n+1}$ 的线性映射。即:

$$x_{n+1} = \sum_{i=1}^k w_i R_i(x) \quad (3)$$

在 MATLAB 神经网络工具箱中, 径向基函数神经网络的设计函数为:

$$[net, tr] = newrb(P, T, GOAL, SPREAD) \quad (4)$$

式中: P 为输入向量; T 为目标输出向量; $GOAL$ 为网络均方误差目标值; $SPREAD$ 为径向基函数的分布系数, 网络越大, 网络输出就越平滑, 网络泛化能力也越强。该函数通过自动增加径向基神经元的方法, 来不断地减少网络输出的均方误差, 直到该误差达到参数 $GOAL$ 的要求时, 网络的训练结束。

由此可见, RBF 网络进行时间序列的自回归预测时, 需要确定的 RBF 网络结构参数为:

- (1) 输入层节点数 m ;
- (2) 径向基函数的分布系数 $SPREAD$;
- (3) 网络均方误差目标值 $GOAL$ 。

$$ARV = \frac{\sum_{i=1}^N [x(i) - \hat{x}(i)]^2}{\sum_{i=1}^N [x(i) - \bar{x}]^2} \quad (5)$$

3 RBF 网络预测模型优化的遗传算法构造

通过上述分析发现, RBF 网络预测能力的影响参数为输入层节点数 m 、径向基函数的分布系数 $SPREAD$ 及网络均方误差目标值 $GOAL$, 为了获取最佳预测性能的 RBF 网络, 需要得到最佳的 m 、 $SPREAD$ 和 $GOAL$ 。显然这是一个优化问题, 如果采取穷举的方式搜索最优值, 计算量将十分巨大以至无法实现。由于遗传算法^[6]具有隐含的并行性和强大全局搜索能力, 可以在很短的时间内搜索到全局最优点。因此利用遗传算法来进行 RBF 网络的预测模型的结构优化。

构造 RBF 网络预测模型优化的遗传算法需要以下步骤:

(1) 设定种群数目 n , 种群数目太小, 遗传算法的性能将变得很差或根本找不出问题的解, 太大, 则会增加计算量, 使收敛时间增长, 种群数目一般取为 30~50 个。

(2) 对输入层节点数 m 、径向基函数的分布系数 $SPREAD$ 及网络均方误差目标值 $GOAL$ 进行二进制编码, 并随机产生初始种群。根据优化参数的取值范围, 将其量化值(用二进制串表示) 编码成基因串 $a = \{\alpha_0, \dots, \alpha_{b-1}; \alpha_b, \dots, \alpha_{b-1}; \alpha_{2b}, \dots, \alpha_{3b}\}$, 其中 a 中的前 B 个量化值代表输入层节点数 m , 中间 B 个量化值代表径向基函数的分布系数 $SPREAD$, 最后 B 个量化值代表网络均方误差目标值 $GOAL$ 。所以每个基因串长度为 $3B$ 个比特位, 此时的搜索空间有 2^{3B} 个点。

(3) 适应度函数的确定。

在实际应用中, 通常用预测值与实际值的差别来交叉验证模型的有效性。对于实际测得的时间序列 $\{x_1, x_2, \dots\}$, 可以利用一部分数据(通常为前 75%) 来建模, 而用另一部分数据(后 25%) 来对所建模型进行验证, 文中采用平均相对变动值 (average relative variance, ARV)^[7], 其定义为:

式中: N 为比较数据个数, $x(i)$ 为实测数据值, \bar{x} 为实测数据平均值, $\hat{x}(i)$ 为预测值。显然, 平均相对变动值 ARV 越小, 也表明预测效果越好, $ARV = 0$ 表示达到了理想预测效果, 当 $ARV = 1$ 时, 表明模型仅达到平均值的预测效果。因此, 定义适应度函数为 $f = 1/ARV$ 。

(4) 对种群解码, 计算每条基因串的适应度。

(5) 将适应度最大的个体, 即种群中最好的个体无条件地复制到下一代新种群中, 然后对父代种群进行选择、交叉和变异等遗传算子运算, 从而繁殖出下一代新种群其它 $n-1$ 个基因串。通常采用转轮法作为选取方法, 适应度大的基因串选择的机会大, 从而被遗传到下一代的机会大, 相反, 适应度小的基因串选择的机会小, 从而被淘汰的机率大。交叉和变异是产生新个体的遗传算子, 交叉率太大, 将使高适应度的基因串结构很快被破坏掉, 太小则使搜索停止不前, 一般取为 0.5~0.9。变异率太大, 将使遗传算法变为随机搜索, 太小则不会产生新个体, 一般取为 0.01~0.1。

(6) 如果达到设定的繁衍代数, 返回最好的基因串, 并将其作为输入层节点数和隐层节点数, 算法结束。否则, 回到(4)继续下一代的繁衍。

4 算 例

为了验证文中方法的有效性, 选取某航空发动机的光谱分析数据作为算例, 分析仪器为美国 Bird 公司的原子发射光谱分析仪。文中遗传算法参数为: 种群数 $n = 30$, 基因串(染色体)采用二进制编码, 交叉率和变异率分别为 0.750 和 0.05, 进化代数为 100 代。

m 、 $SPREAD$ 及 $GOAL$ 均用 5 位二进值编码, 为了避免节点数为 0, 规定解码后, 对于输入层节点数加上 1 得到 m , 而 $SPREAD$ 加上 1 以得到 $SPREAD$, 对 $GOAL$, 通过计算 $GOAL = 10^{-(GOAL \times 2/32)}$ 得到网络训练允许误差目标值。适应度函数取为 $f = 1/ARV$ 。

表 1 遗传算法优化得到的 RBF 神经网络模型(前 8 个)

No	1	2	3	4	5	6	7	8
m	22	21	22	23	22	21	22	22
$SPREAD$	22	21	20	22	26	21	18	20
$GOAL$	0.0237	0.0750	0.0133	0.0237	0.0237	0.0316	0.0237	0.0237
$1/ARV$	10.3964	8.0235	6.2037	5.3816	4.6447	4.3556	3.1277	2.7889

经过 50 代进化后, 适应度函数达到了最大值, 其优化结果为: 最优的适应度值为 10.3964, ARV 为

0.0962, 输入层节点数 $m = 22$, $SPREAD = 22$, $GOAL = 0.0237$ 。图 2 为该网络对 Fe 元素的光谱数据的预测

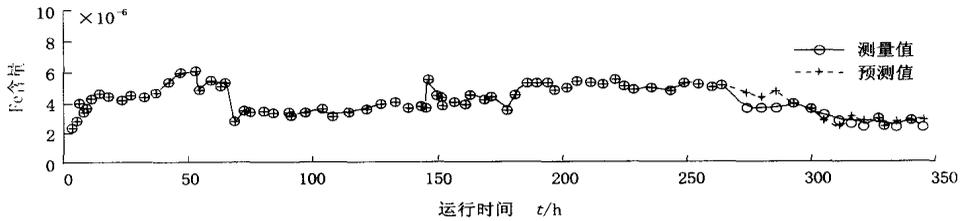


图 2 最优 RBF 网络 ($m = 22, SPREAD = 22, GOAL = 0.0237$) 的预测值与实测值比较

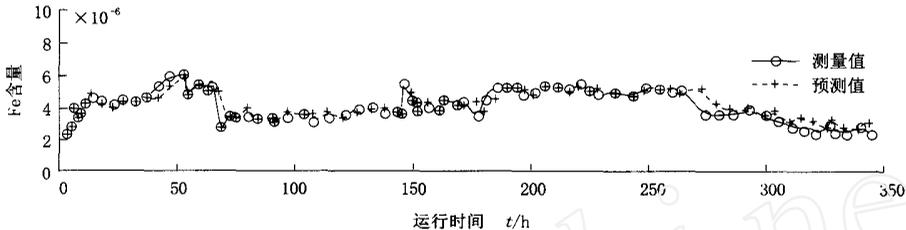


图 3 ARMA (23, 17) 模型的预测值与实测值比较

值与实测值的比较,其中,用前 60 个数据建模,用后 18 个数据来验证。从图中可以看出,测试数据与预测值的拟合程度达到了很高的精度。为了与传统 ARMA 模型进行比较,采用文献[8]的方法得到最佳的 ARMA 预测模型为 ARMA (23, 17)。用相同的数据建模和验证后,得到的实测值与预测值的比较,如图 3 所示,ARV 为 0.2683。对比图 2 和图 3,不难看出,结构自适应神经网络预测模型得到了更佳的预测效果,而且整个建模过程实现了自动化。由此可见,实例充分验证了结构自适应神经网络预测模型的正确合理性。

5 结 论

(1) 提出用神经网络进行光谱数据预测的方法,从而对机械磨损故障实现有效地监测。

(2) 分析了神经网络进行非线性预测、多变量预测的优越性以及神经网络用于预测的缺点所在。

(3) 分析了影响 RBF 网络预测能力的 3 个重要参数——输入层节点数、RBF 分布系数 $SPREAD$ 及网络均方误差目标值 $GOAL$; 提出基于结构自适应 RBF 神经网络的光谱油样分析数据建模技术,用遗传算法构造了同时优化影响神经网络预测精度的参数的算法,得到了最优的 RBF 网络预测模型。

(4) 用某航空发动机的实测光谱数据为例,利用结构自适应 RBF 神经网络建立预测模型,实现了发动机油样光谱数据的精确预测,并与 ARMA 模型进行了比较,充分验证了文中方法的有效性和优越性。

参考文献

- [1] 干敏梁,左洪福,杨忠,等 时序建模方法在滑油光谱分析中的应用 光谱学与光谱分析,2000,20(1): 64-67.
- [2] Lapedes A, Farber Nonlinear signal processing using neural network: prediction and system modeling [R]. Technical Report LA-UR-87-2662, Los Alamos, NM: Los Alamos National Laboratory, 1987.
- [3] Weigend A B, et al Predicting the future: a connectionist approach [J]. International Journal of Neural System, 1990(1): 193-209.
- [4] Xu K, et al Application of neural networks in forecasting engine systems reliability [J]. Applied Soft Computing, 2003(2): 255-268.
- [5] 刘豹,胡代平.神经网络在预测中的一些应用研究 [J].系统工程学报,1999,14(4): 338-344.
- [6] Goldberg D. Genetic algorithms in search, optimization and machine learning [M]. Reading, MA: Addison Wesley, 1989.
- [7] Cholewo T, Zurada J M. Sequential network construction for time series prediction [C]. Proceedings of the IEEE International Joint Conference on Neural Networks, 1997: 2034-2039.
- [8] 陈果.基于遗传算法的 ARMA 模型定阶新技术 机械工程学报,2005,41(1): 41-45.

作者简介

杨虞微 男 1972 年出生 博士研究生 主要研究方向为航空发动机智能诊断与专家系统

E-mail: yangyuwei2003@yahoo.com.cn