

航空发动机复杂磨损趋势的神经网络多变量预测模型

陈 果 杨虞微

南京航空航天大学,南京,210016

摘要:针对非等间隔的受加油因素影响的光谱油样分析数据的建模预测问题,建立了 BP 神经网络的多变量预测模型,充分考虑了油样分析数据的非等间隔性及受加油因素影响的特点,同时,用遗传算法对网络参数进行了优化。最后,利用两组实际的航空发动机油样光谱分析数据对模型进行了验证。结果表明,所提出的神经网络多变量预测模型能有效解决实际的受多因素影响的油样分析数据建模问题,具有较强的工程实用价值和通用性。

关键词:光谱油样分析;多变量预测;非等间隔;BP 神经网络;遗传算法

中图分类号:O329;F201 **文章编号:**1004—132X(2007)01—0070—05

Artificial Neural Network Multi - Variable Forecasting Model of Aero - Engine Complex Wear Trend

Chen Guo Yang Yuwei

Nanjing University of Aeronautics and Astronautics, Nanjing, 210016

Abstract: Multi - variable predicting model by Back - Propagation Neural Networks (BPNN) was established, and it considered the complex characteristics of the oil analysis data fully such as unequal interval sampling and affected by adding oil factor. In addition, the effect of ANN's parameters on the predicting accuracy was also discussed, and Genetic Algorithm (GA) was used to optimize ANN's parameters. Finally, two time series and aero - engine spectrometric oil analysis data were used to verify this model. The results show that this model can solve effectively the prediction problem of oil analysis data, which is affected by multi - factors such as unequal interval sampling and adding oil factor. The new method has important engineering application value, and it is a common method of forecasting complex time series, which is affected by multi - factors.

Key words: spectrometric oil analysis; multi - variable forecasting; unequal interval sampling; BP neural network; GA

0 引言

油样分析数据具有明显的非等间隔性。机器在正常使用期故障率较低,采样间隔往往比较大,而在损耗阶段故障率较高,这时往往要缩短采样间隔。采样时间间隔的不同就形成了非等间隔时

间序列,另外,由于油样分析中,实际采样时间的安排本身也具有非等间隔性,难以进行完全等间隔采样。由此可见,非等间隔性具有普遍意义,等间隔时间序列只是其特殊情况。

同时,由于机器运行时将消耗润滑油及出现润滑油泄漏的情况,因此,需要不定期地进行加油、补油或换油。显然,由于每次取样后的加油影

收稿日期:2005—11—22

- [5] Davies M A, Dutterer B, Pratt J R, et al. On the Dynamics of High Speed Milling with Long, Slender End - mills[J]. Annals of CIRP, 1998, 47(1): 55-60.
- [6] Lin R S, Koren Y. Efficient Tool - path Planning for Machining Free - form Surfaces [J]. ASME Journal of Engineering for Industry, 1996, 118: 20-28.
- [7] Lo C C. A New Approach to CNC Tool Path Generation[J]. Computer - aided Design, 1998, 30(8): 649-655.

- [8] 周济,周艳红. 数控加工技术[M]. 北京:国防工业出版社,2002.
- [9] Xu H Y, Zhou Y H, Zhang J J. Angular Interpolation of Bi - parameter Curves[J]. Computer - aided Design, 2003, 35(12): 1211-1220. (编辑 何成根)

作者简介:徐立国,男,1963年生。吉林大学机械科学与工程学院博士研究生。主要研究方向为智能精密制造。发表论文 20 篇。赵继,男,1959年生。吉林大学机械科学与工程学院教授、博士研究生导师。徐卫,女,1962年生。空军航空大学计算机教研室副教授。

响,油样分析所得到的磨损颗粒浓度将产生变化。这样就导致了实际油样分析数据将表现出受加油因素影响的特征。

目前,传统方法均不能有效地解决此类问题。传统 ARMA 法^[1]是针对平稳的、零均值和正态分布的随机序列的线性模型,因此不适合进行非线性预测。统计回归法进行非线性预测时需要知道序列的非线性模型,且回归阶次难于确定。灰色系统模型相当于指数回归模型,但在实际应用中,一个非负序列的累加生成序列未必有指数规律,累加生成和累减还原的方法建模必将加大模型的误差,因此在实际应用中,要慎用灰色系统法^[2],通常,当序列的趋势呈指数函数递增或递减时,灰色模型 GM(1,1)具有相当高的精度,当系统发生转折或周期性变化时,其精度就变得相当差,灰色预测-校正模型^[3]尽管能弥补该缺陷,但需要对原始序列进行平滑处理,同时需要人工选取分段点,建模复杂,且通用性较差。神经网络方法就是由于人工神经网络具有逼近任意非线性函数和一定的泛化能力,在预测领域显示出了强大的生命力和巨大的发展潜力。相关研究^[4]均表明了神经网络较其他方法具有优越性,尤其是在多变量预测技术中,神经网络技术更加显示了其突出的优势。据上述分析可知,要精确预测航空发动机复杂磨损趋势,必须充分考虑多种因素(加油、补油、换油和非等间隔等),实施多变量预测方案。

本文旨在建立神经网络多变量预测模型,综合解决非等间隔的受加油因素影响的油样分析数据的建模和预测问题,从而为航空发动机的视情维修提供更为可靠的决策依据。本文采用三层 BP 网络来建立神经网络预测模型^[5]。

1 BP 网络的多变量时间序列预测模型

1.1 神经网络多变量时间序列预测方法

对于多变量时间序列 $(x_{1,i}, x_{2,i}, \dots, x_{p,i}), i = 1, 2, \dots, n$, 它有 p 个时间变量,认为时间序列的未来值与其前面的 m 个值之间有某种函数关系:

$$(x_{1,n+k}, x_{2,n+k}, \dots, x_{p,n+k}) = F((x_{1,n}, x_{2,n}, \dots, x_{p,n}), (x_{1,n-1}, x_{2,n-1}, \dots, x_{p,n-1}), \dots, (x_{1,n-m+1}, x_{2,n-m+1}, \dots, x_{p,n-m+1})) \quad (1)$$

利用神经网络拟合函数 F , 然后即可用它进行预测。进行多变量时间预测的神经网络结构分为两种,即单步预测网络和多步预测网络。单步预测网络输出的个数是 p , 即为一个多变量时间序

列的变量个数,一次计算得到所有变量一步的预测值;多步预测网络的输出个数为 $k \times p$ 个,一次计算可得到所有 p 个变量的 k 步预测结果。多变量时间序列的多步预测实现方法是进行迭代多步预测。

利用传统的预测方法进行多变量时间序列预测的建模与预测非常复杂。而利用神经网络进行多变量时间序列预测的方法,就如同进行单变量时间序列预测一样简单,这是传统预测方法无法比拟的。

1.2 油样分析数据的神经网络多变量预测模型

航空发动机光谱油样数据的非等间隔性是由于不定期取样和数据丢失等原因造成的,因此,非等间隔的光谱油样数据是很常见的,通常的做法是先将其近似作为等间隔处理,或通过插值的方法得到等间隔数据,然后再运用通常的等间隔数据建模方法(如 ARMA 模型)对数据进行动态建模处理。但这种处理方法会造成较大的误差,与实际情况不尽符合。有效的方法是先对非等间隔数据进行建模,然后,运用模型对未来短期内的任意时刻进行预测。对时间序列进行预测,实际上是通过非线性系统产生的时间序列进行建模,在把握非线性系统的运动规律和运动状态的基础上,从系统产生的时间序列中抽取动力系统,重构相空间,最常用的方法是时延法。

设所研究的时间序列为 $\{x(t)\}, t = 1, 2, \dots, N$, 则当前状态的信息可以表示为

$$x(t+k) = f(x(t), x(t-), \dots, x(t-(m-1))) \quad (2)$$

式中, m 为嵌入维数; t 为时间延迟; k 为预测步数,通常取采样间隔值。

由相空间重构理论的嵌入定理可知^[6]:假设动力系统的维数为 d , 如果 $m \geq 2d + 1$, 则这种映射产生的伪相空间和系统的状态空间微分同胚及拓扑等价,它们的动力学特性定性意义上是完全相同的。

由此可见,对时间序列的预测,关键在于根据已知时间序列数据进行相空间的重构,找出从 m 维空间映射到一维空间的映射函数。而神经网络正好可用来拟合这一非线性函数。

对应于式(2)的状态信息可表示为

$$x(t+k) = f(x(t), x(t-), \dots, x(t-(m-1))); (t+k-t), t-(t-), \dots, [t-(m-2)] - [t-(m-1)]) \quad (3)$$

对于非等间隔数据,令

$$x(t+k) = x(t_{n+1}), x(t) = x(t_n), x(t-) = x(t_{n-1}), x(t-(m-1)) = x(t_{n-m+1})$$

$$(t+k) - t = t_n, t - (t -) = t_{n-1}, \dots, [t - (m - 2)] - [t - (m - 1)] = t_{n-m+1}$$

则式(3)可写为

$$x(t_{n+1}) = f(x(t_n), x(t_{n-1}), \dots, x(t_{n-m+1}); t_n, t_{n-1}, \dots, t_{n-m+1}) \quad (4)$$

由此可见,等间隔序列预测是非等间隔序列预测的特殊情况。下一时刻的数值不仅与前 m 个数据值有关,而且还与序列间的时间间隔有关,因此在建立非等间隔预测时,需要在神经网络的输入向量中考虑序列的时间间隔因素。

对于航空发动机润滑油光谱数据组成的非等间隔时间序列而言,设 $C(t)$ 为 t 时刻元素的质量浓度 (ng/mL),则由式(4)可得其预测公式:

$$C(t_{n+1}) = f(C(t_n), C(t_{n-1}), \dots, C(t_{n-m+1}); t_n, t_{n-1}, \dots, t_{n-m+1})$$

另外,由于航空发动机的润滑油存在泄漏及受热蒸发等特点,所以在每次取样之间均将消耗一部分润滑油,为了保证机器充分润滑,每次取样后均要对润滑油进行补充。而润滑油的补充经常将导致油样质量浓度数据的变化。下面对两者的关系进行分析。

设容积为 V 的润滑油中 Fe 元素的累计磨损量 W_{Fe} (ng) 随时间 t (h) 的变化关系为 $W_{Fe}(t)$,设加油量 O 与时间 t 的关系为 $O(t)$,则 i 时刻润滑油中 Fe 元素的质量浓度与时间的关系为

$$C_{Fei} = (W_{Fei} - \sum_{k=1}^{i-1} dW_{Fek} - C_{Fei}O_i) / V \quad (5)$$

$$C_{Fei} = (W_{Fei} - \sum_{k=1}^{i-1} dW_{Fek}) / V$$

$$dW_{Fek} = C_{Fek}O_k$$

显然, i 时刻的 Fe 元素的质量浓度不仅与当前时刻的 Fe 元素的累计磨损量 W_{Fei} 有关,还与前 $i - 1$ 时刻的加油量有关。因此式(4)可写为

$$C(t_{n+1}) = f_1(C(t_n), C(t_{n-1}), \dots, C(t_{n-m+1}); t_1, t_2, \dots, t_{n-m+1}; O(t_n), O(t_{n-1}), \dots, O(t_{n-m+1})) \quad (6)$$

由此可见,下一时刻的 Fe 元素的质量浓度与其前一时刻的质量浓度和加油量均有关系。通常,在工程实际中,往往只能得到润滑油消耗率,加油量通过两次之间的取样时间和润滑油消耗率相乘而得到,即 $O_{n-1} = Q_{n-1} t$,其中, t 为第 n 次与第 $n - 1$ 次取样时间间隔 (h), Q_{n-1} 为第 $n - 1$ 次取样时的润滑油消耗率 (L/h),该值也有变化。因此,加油量为时间间隔和润滑油消耗率的函数,故式(6)又可表示为

$$C(t_{n+1}) = f_2(C(t_n), C(t_{n-1}), \dots, C(t_{n-m+1}); t_1, t_2, \dots, t_{n-m+1}; Q(t_n), Q(t_{n-1}), \dots, Q(t_{n-m+1})) \quad (7)$$

根据上述分析,本文建立了综合考虑数据间隔信息、加油量信息的光谱油样数据多变量神经

网络预测模型。为了简化模型,加油量的时间延迟与油样数据和采样时间均取值 m ,如图 1 所示。

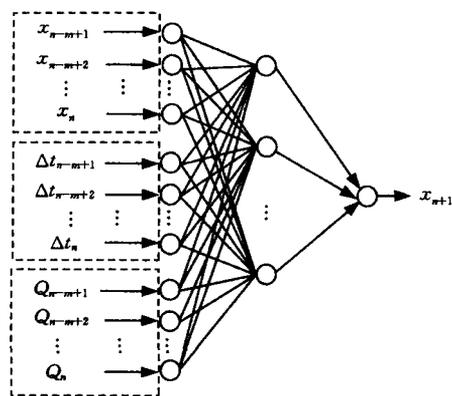


图 1 油样分析数据的神经网络多变量预测模型

(1) 将数据间的采样时间间隔 t 作为一个输入信息,从而形成一个既包含数据信息又包含时间间隔信息的训练样本和预测样本,这样使时间序列的间隔信息反映到了模型中。

(2) 将每次取样时的润滑油消耗率 Q 作为一个变量,从而就形成了一个既包含数据信息又包含加油信息的训练样本和预测样本,这样使加油信息反映到了模型中。

不失一般性,本文仅考虑单步预测情况,为统一解决非等间隔和受加油因素影响油样分析数据而建立的多变量预测模型是对单变量模型的推广,即需要拟合如式(7)所示的函数。

2 用遗传算法优化神经网络预测模型参数^[7,8]

神经网络预测能力的主要影响参数为输入层节点数 N 、中间隐层节点数 M 及网络允许训练误差,为了获取最佳预测性能的 BP 网络,需要得到最佳的 N 、 M 和网络允许训练误差目标值。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。因此,本文利用文献[7,8]中的神经网络预测模型结构优化方法得到最优的 BP 网络预测模型。

该方法的基本思路是,首先对输入层节点数 N 、中间隐层节点数 M 和训练允许误差进行二进制编码,并随机产生初始种群。用平均相对变动值 (average relative variance, ARV) 的倒数构造适应度函数,在适当的交叉率和变异率下,通过逐代遗传,最后搜索出具有最佳预测能力的网络模型。其中衡量预测效果的平均相对变动值定义为

$$ARV = \frac{\sum_{i=1}^{N_c} [x(i) - \bar{x}(i)]^2}{\sum_{i=1}^{N_c} [x(i) - \bar{x}(i)]^2} \quad (8)$$

式中, N_c 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,表明预测效果越好, $ARV = 0$ 表示达到了理想预测效果, $ARV = 1$,表明模型仅达到平均值的预测效果。

3 应用实例

为了对本文方法进行充分验证,分别用两组某型军用飞机发动机的非等间隔的受加油因素影响的润滑油光谱分析的 Fe 元素含量数据对所建模型进行验证。光谱数据采用美国 BIRD 公司的原子发射光谱仪获得。在建模分析中,均采用本文提出的神经网络多变量预测模型。

3.1 实例一

某型号某台航空发动机在从 18h 到 275h 的运行过程中,得到了 66 个光谱油样分析数据,将数据分成两个部分,用其中的 90% (前 59 个数据) 建模,剩余的 10% (后 7 个数据) 用于对所建模型进行外推验证。图 2 所示为样本时间间隔与发动机运行时间的关系,其中,润滑油消耗速率均为 0.1L/h,因此,加油量的变化规律与时间间隔完全一致。从图 2 中可以观察到时间间隔变化较大,其中最小间隔为 1h,最大间隔为 10h。用遗传算法对神经网络进行优化得到:在训练样本比例为 0.9、允许训练步数为 1000、隐层节点数为 6、输入节点数为 2 和训练精度为 0.001 的情况下,达到了最佳预测效果,其一步预测值与实测值的比较结果如图 3 所示。从图 3 中可以看出模型达到了相当高的预测精度,所有样本的平均相对变动值为 0.0348。

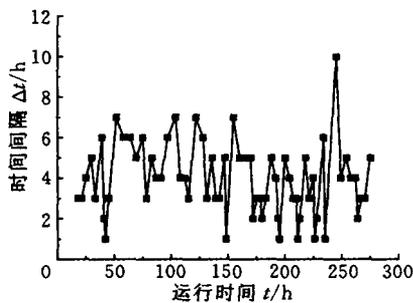
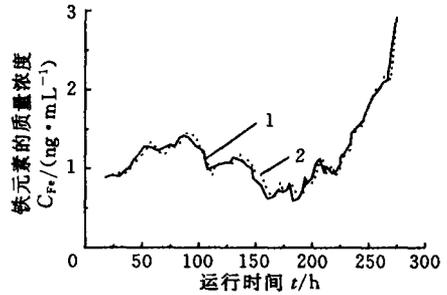


图 2 时间间隔与发动机运行时间的关系

为了衡量模型的预测时间及精度,在一步预测的基础上,进行多步预测,其思路是将上一步预测的结果代入下一步预测,通过迭代实现多步预测。表 1 为多步预测结果,从表 1 中可以看出,三步以内其预测精度很好,基本上在 5% 以内,但在四步以后其预测精度则在 10% 以上。表 2 为对同样数据点进行一步预测的结果,很明显,其预测精度基本上在 5% 以内。



1. 实测值 2. 预测值

图 3 Fe 元素的质量浓度一步预测值与实测值的比较

表 1 多步预测精度分析

步数	时间 t (h)	时间间隔 t (h)	加油速率 (mL/h)	真实值 y _s (μg/L)	预测值 y _y (μg/L)	绝对误差 y _y - y _s	相对误差 y _y - y _s / y _s (%)
1	254	5	0.5	1.924	1.9269	0.0029	0.15073
2	258	4	0.4	1.984	2.0568	0.0728	3.669 35
3	262	4	0.4	2.076	2.1895	0.1135	5.467 24
4	264	2	0.2	2.108	2.3675	0.2595	12.31025
5	267	3	0.3	2.124	2.5850	0.461	21.704 33
6	270	3	0.3	2.506	2.9500	0.444	17.717 48
7	275	5	0.5	2.922	3.6207	0.6987	23.9117

表 2 一步预测精度分析

步数	时间 t (h)	时间间隔 t (h)	加油速率 (mL/h)	真实值 y _s (μg/L)	预测值 y _y (μg/L)	绝对误差 y _y - y _s	相对误差 y _y - y _s / y _s (%)
1	254	5	0.5	1.924	1.9269	0.0029	0.150 73
1	258	4	0.4	1.984	2.0500	0.066	3.326 61
1	262	4	0.4	2.076	2.0303	0.0457	2.201 35
1	264	2	0.2	2.108	2.1765	0.0685	3.249 53
1	267	3	0.3	2.124	2.1556	0.0316	1.487 76
1	270	3	0.3	2.506	2.1366	0.3694	14.740 62
1	275	5	0.5	2.922	2.9712	0.0492	1.683 78

3.2 实例二

某型号某台航空发动机在从 3h 到 345h 的运行过程中,得到了 78 个光谱油样分析数据,将采样数据分成两个部分,用其中的 90% (前 70 个数据) 进行建模,剩余的 10% (后 8 个数据) 用于对所建模型进行外推验证。图 4 所示为样本时间间隔与发动机运行时间的关系,该数据中润滑油消耗率均为 0.1L/h,因此,加油量的变化规律与时间间隔完全一致。从图 4 中可以观察到时间间隔变化较大,其中最小间隔为 1h,最大间隔为 10h。用遗传算法对神经网络进行优化得到:在训练样本比例为 0.9、允许训练步数为 1000、隐层节点数

为 6、输入节点数为 4 和训练精度为 0.001 的情况下,达到了最佳预测效果,其一步预测值与实测值的比较结果如图 5 所示。从图 5 中可以看出模型达到了相当高的预测精度,所有样本的平均相对变动值 ARV 为 0.0134。

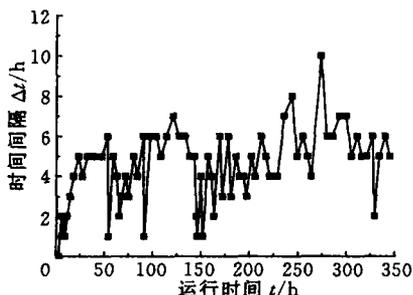
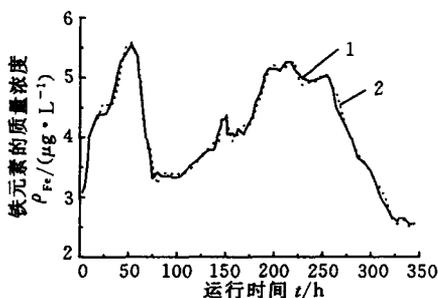


图 4 时间间隔与发动机运行时间的关系



1. 实测值 2. 预测值

图 5 Fe 元素的质量浓度一步预测值与实测值的比较

表 3 和表 4 所示分别为多步和一步预测精度分析,从表 3 中可以看出,两步以内其预测精度较好,基本上在 10% 以内,但在三步以后其预测精度则在 10% 以上。表 4 为对同样数据点进行一步预测的结果,其预测精度基本上在 5% 以内(除第 1 步)。

表 3 多步预测精度分析

步数	时间 t (h)	时间间隔 t (h)	加油速率 (mL/h)	真实值 y_s ($\mu\text{g/L}$)	预测值 y_y ($\mu\text{g/L}$)	绝对误差 $ y_y - y_s $	相对误差 $ y_y - y_s / y_s (\%)$
1	311	6	0.6	2.9380	3.1257	0.1877	6.388 70
2	316	5	0.5	2.8180	3.0778	0.2598	9.219 30
3	321	5	0.5	2.6700	3.0033	0.3333	12.483 15
4	327	6	0.6	2.5900	2.9425	0.3525	13.610 04
5	329	2	0.2	2.6500	2.9179	0.2679	10.109 43
6	334	5	0.5	2.6440	2.9481	0.3041	11.501 51
7	340	6	0.6	2.5440	2.9087	0.3647	14.335 69
8	345	5	0.5	2.5575	2.9053	0.3478	13.599 22

4 结论

(1) 针对基于油样分析技术的航空发动机磨损状态趋势预测问题,提出了基于神经网络多变

表 4 一步预测精度分析

步数	时间 t (h)	时间间隔 t (h)	加油速率 (mL/h)	真实值 y_s ($\mu\text{g/L}$)	预测值 y_y ($\mu\text{g/L}$)	绝对误差 $ y_y - y_s $	相对误差 $ y_y - y_s / y_s (\%)$
1	311	6	0.6	2.9380	3.1257	0.1877	6.388 70
2	316	5	0.5	2.8180	2.9079	0.0899	3.190 21
3	321	5	0.5	2.6700	2.6488	0.0212	0.794 01
4	327	6	0.6	2.5900	2.5431	0.0469	1.810 81
5	329	2	0.2	2.6500	2.5657	0.0843	3.181 13
6	334	5	0.5	2.6440	2.6129	0.0311	1.176 25
7	340	6	0.6	2.5440	2.4920	0.0520	2.044 03
8	345	5	0.5	2.5575	2.4823	0.0752	2.940 37

量预测模型的非等间隔的受加油因素影响的光谱油样分析数据的建模方法,并用遗传算法对神经网络预测模型参数进行了优化。

(2) 利用实测的两组非等间隔的受加油因素影响的油样光谱数据验证了本文提出的神经网络多变量预测模型的有效性,并进行了预测时间和精度分析。

参考文献:

- [1] 杨叔子,吴雅. 机械故障诊断的时序方法 [M]. 西安:西安交通大学出版社,1989.
- [2] 张玉详,陆士良. 神经网络在工程时间序列预报问题研究中的应用 [J]. 南京理工大学学报,1997, 21(6):522-524.
- [3] Zhang H, Li Z G, Chen Z N. Application of Grey Modeling Method to Fitting and Forecasting Wear Trend of Marine Diesel Engine[J]. Tribology International, 2003, 36: 753-756.
- [4] 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.
- [5] 吴今培,肖建华. 智能故障诊断与专家系统 [M]. 北京:科学出版社, 1997.
- [6] 徐章遂,房立清,王希武,等. 故障信息诊断原理及应用[M]. 北京:国防工业出版社, 2000.
- [7] 陈果. 用遗传算法优化神经网络预测模型结构[J]. 中国机械工程, 2004, 15(增刊): 176-179.
- [8] 陈果. 神经网络模型的预测精度影响因素分析及其优化[J]. 模式识别与人工智能, 2005, 18(5): 528-534.

(编辑 何成根)

作者简介:陈 果,男,1972 年生。南京航空航天大学民航学院副教授。主要研究方向为航空发动机状态监测与故障智能诊断、专家系统、数据融合、神经网络与遗传算法、图像处理及模式识别和机械动力学等。发表论文 30 余篇。杨虞微,男,1972 年生。南京航空航天大学民航学院博士研究生。