

基于信息压缩的四层前向网络模型 体系结构及其应用^①

谢 涛 张育林

(国防科学技术大学航天技术系, 长沙, 410073)

摘 要: 提出了一种基于信息压缩原理的四层前向网络体系结构, 用以为液体火箭发动机部分参数建立网络模型。利用所选参数的时间序列对这些参数实现平滑、滤波与预测等处理, 提高了基于试车数据驱动的故障监控系统对量测噪声、传感器故障与系统过程噪声的鲁棒性。该四层结构大大缩小了网络权值规模, 网络模型的训练采用演化策略。

主题词: 人工神经网络, 信息压缩⁺, 网络结构, 液体推进剂火箭发动机, 故障检测

分类号: V434.3

FOUR-LAYERED FORWARD ANN HIERARCHY FOR HUGE PERCEPTUAL PROBLEM BASED ON INFORMATION COMPRESSION

Xie Tao Zhang Yulin

(Dept. of Aerospace Technology, National Univ. of Defense Technology, Changsha, 410073)

Abstract: Due to the shortcomings of three-layered forward neural network, a four-layered forward neural network model was constructed based on the information compression principle, and this network hierarchy was used to build a start-up plus main-stage model for parts of the parameters of liquid rocket engine system. The time series of the input parameters were used to smooth and filter as well as predict themselves, which were required in the health monitoring system for rocket engine. Thus the robustness of the health monitoring system to the measurement and process noises as well as the sensor faults were greatly improved. Evolution Strategies were used to train the weight and coefficients of the activation functions. A procedural training strategy was applied in the optimization of the huge 4-layered network model in which the network was initially trained globally in a whole, then broke down into layers for locally training. In addition, as this network hierarchy has less weight and hidden nodes than the conventional 3-layered forward network model, it is a classical network hierarchy for huge forward network model for function approximation.

Subject terms: Artificial neural networks, Information compression⁺, Lattice structure, Liquid propellant rocket engine, Fault detection

1 引 言

鉴于三层 BP 模型存在的固有缺陷^[1], 对样本数据携带的噪声鲁棒性解比较差, 而且随着隐节点数的增大, 网络权重矩阵很快增加。在为某液体火箭发动机部分试车参数建模时, 本

^① 收稿日期: 1998-01-21, 修回日期: 1998-04-01, 本课题为国家自然科学基金资助项目, 批准号:

文提出了一种四层前向感知机的网络体系结构。该网络结构基于信息压缩与解压原理,以远少于第一层输入节点数的第二层隐节点压缩庞大的输入层信息,然后再以与输入层节点数相当的第三层隐节点实现对第二层压缩信息的解压。第四层即输出层,由平滑层、滤波层与预测层所组成。网络模型用来对发动机启动段与主级段的试车参数进行平滑、滤波与预测^[2]。网络输入层利用发动机参数的相邻时间序列,以提高参数的平滑、滤波与预测效果。用来训练的数据样本集是经滤波处理的一次正常试车数据的前 20s 采样数据,采样频率为 50Hz。网络训练采用演化策略^[3],并在训练巨型网络模型中采用分级分别层训练的策略思想。

2 四层前向感知机的体系结构

基于数据压缩原理的四层前向网络模型如图 1 所示。本文选择 17 个发动机参数作为待平滑、滤波与预测的参数,且每个参数的相邻 6 个时间序列值作为网络的输入节点,即输入节点数量为 17×6 。第二层取 10 个隐节点作为数据压缩层,第三层取 100 个隐节点作为数据解压层,第四层为输出层,该层由 3 个并行且具类似结构的平滑层、滤波层与预测层构成,可对输入层的 17 个发动机参数进行平滑、滤波与预测,发动机参数的平滑、滤波与预测值都是发动机故障监控系统所需要的^[2]。

图 1 的四层网络中除输入层外,各节点动作函数均取 Sigmoidal 函数 $z(x) = (h_1 - h_2)/(1 + e^{-ax}) + h_2$, $h_1 > h_2$,其中 h_1 与 h_2 分别为动作函数的极大值与极小值, a 为动作函数的敏感系数,三者均为待训练的函数参数。

因此,可以计算出该网络模型中待训练的权值与动作函数参数总个数: $N = 7603$ 。而同等信息容量规模的三层网络中,待训练的权值与动作函数参数总个数为 $N > 21573$,几乎是四层压缩网络的 3 倍。

对于这样一个庞大的网络模型,通过训练要找到一个最优的权值与参数集,一些改进的 BP 算法在仿真试验中被证明是无效的,表现为迭代过程中出现不收敛的振荡现象或发散。模拟退火算法^[4]效率太低,我们选用进化计算技术中的演化策略作为网络训练算法。

虽然理论上已经证明演化策略可以找到任意函数的全局最优解,但在有限的计算机储量与计算时间限制下,这种巨型网络模型是极难找到其它局最优权值与参数集的。鉴于要解决的问题并不是求解模型的最优解,而是要搜索出一组网络的权值与参数集,使得该网络模型相对于实际训练数据样本集的平滑、滤波与预测总误差满足给定的要求。因此,我们采用网络全局优化与局部优化相结合的方法,利用演化策略有步骤地分层优化该巨型网络模型。

3 巨型网络模型训练算法

演化策略是 60 年代末基于模拟生物进化提出的一种求解数值优化问题的仿生类算

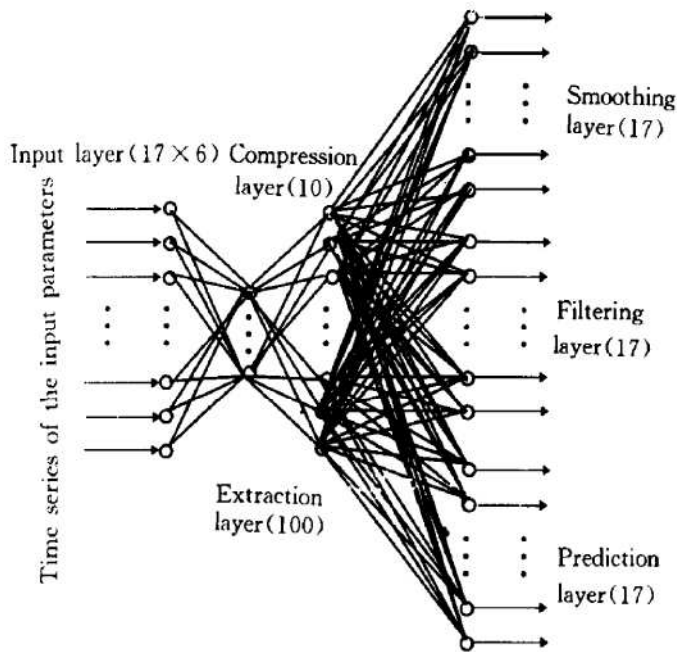


Fig. 1 The ANN model of 17 critical parameters of liquid rocket engine for parameter smoothing and filtering as well as prediction

法^[3~5]。定义父本集合 U , 子代集合 Θ , 第 i 代种群总体集合 Π_i , 并且设 $|U| = \mu$, $|\Theta| = \lambda$, 其中 $|U|$ 和 $|\Theta|$ 分别表示集合 U 与集合 Θ 中的元素的个数。 $I = (X, \sigma)$ 表示个体, 其中 X 为 n 维向量解(解的数据串表示), 本文中即指由待训练的四层网络的所有权值与函数参数分层分类地排列成的一条实数串, $X = \{x_1, x_2, x_3, \dots, x_{7603}\}$; σ 为相应于 X 的 7603 维变异方差(因子)向量。杂交算子 $R(\cdot)$: 设 $I' = (X', \sigma')$, $I'' = (X'', \sigma'')$, $I = (X, \sigma)$, 则 $R(I', I'') \rightarrow I$; 采用随机中和算子, 以概率 P_e 使得 $X_i = \frac{1}{2}(X'_i + X''_i)$, $\sigma_i = \frac{1}{2}(\sigma'_i + \sigma''_i)$, $\forall i$ 。变异算子 $M(I') \rightarrow I$: 设 $I = (X, \sigma)$, $I' = (X', \sigma')$; 以概率 P_m 使得 $\sigma_j = \sigma'_j \exp(N_o(0, \Delta\sigma))$, $x_j = x'_j + N_o(0, \sigma_j)$, $\forall i$ 。其中, $N_o(0, \sigma)$ 是一个方差为 σ 的高斯随机变量产生函数, $\Delta\sigma = \sigma_0 - \delta \cdot t + |N(0, 0.3)|$, t 为种群更新代数, $\delta < \sigma_0/gen$, gen 为种群总代数。

$f(\cdot)$: $X \rightarrow R$ 为一解串的适应值评估函数, 本文即指网络对整个训练样本数据集的平滑、滤波与预测误差的总和, 其中网络对每一数据样本的三种误差均取其偏差的绝对值, 即

$$f(X) = \sum_{i=0}^T \sum_{j=1}^{17} |y_{ij,s}(X) - \bar{y}_{ij,s}| + \sum_{i=0}^T \sum_{j=1}^{17} |y_{ij,f}(X) - \bar{y}_{ij,f}| + \sum_{i=0}^T \sum_{j=1}^{17} |y_{ij,p}(X) - \bar{y}_{ij,p}|$$

其中 $T = 50 \times (\text{启动段时间} + \text{部分主级段时间}) = \text{样本数据个数}$; $y_{ij,s}$ 与 $\bar{y}_{ij,s}$ 分别为在第 i 秒时刻第 j 个参数的平滑值与期望值; $y_{ij,f}$ 与 $\bar{y}_{ij,f}$ 分别为在第 i 秒时刻第 j 个参数的滤波值与期望值; $y_{ij,p}$ 与 $\bar{y}_{ij,p}$ 分别为在第 i 秒时刻第 j 个参数的预测值与期望值。 $S(\Omega) \rightarrow U$ 为父本更新机制; $\Omega = U \cup \Theta$ 为根据当前种群总体中个体的评估值从集合 Ω 中选择 μ 个适应值最高的个体作为下一代种群的父本集。

Mate-Select ($f(I_x)$, U) 为随机杂交父本选择算子, 以与父本集合 U 中的个体 I_x 的性能指标评估值成比例的概率 P_x 随机选取父本集合 U 中的个体 I_x ; 其中 $P_x = f(I_x) / \sum_{i=1}^{\mu} f(I_i)$ 。

演化策略的基本步骤为:

- (1) 随机产生 μ 个父本个体, 计算其适应值并按序排列;
- (2) 基于父本的被选概率 P_x , 随机选取 λ ($\lambda > \mu$) 对父本进行杂交与变异, 计算产生 λ 个子代个体的适应值并按序排列;
- (3) 按 $(\mu + \lambda) - ES$ 更新机制选取下一代父本;
- (4) 判断最优父本个体是否满足问题要求: 满足则终止种群更新循环, 输出其结果; 否则转步骤 (2), 继续种群更新循环。

为了防止遗传算法的过早收敛, 本文采取动态改变每代中父本的适应值并引进共享函数, 以限制父本中的所有个体趋于一致。

巨型网络优化策略: 先找模型的全局次优解, 缩小最优解的搜索范围, 再在该全局次优解的附近进行循环微搜索, 进一步逼近问题的最优解。模型的全局次优解的求法即通常的网络训练算法: 全部待搜索的网络权值与动作函数的参数值都设为可变量, 其值通过演化策略、模拟退火技术或 BP 算法来训练。

模型的局部优化即分层优化: 从网络结构中选择一层节点, 暂时固定其连接权值与动作函数的参数值, 剩下的网络层的权值与函数参数作为一个新的优化参数集, 其值同样可以通过演化策略、模拟退火技术或 BP 算法来训练。然后依次对该剩下的网络层进行同样的分级局

部优化，直至分解到只剩下一层网络模型优化。然后再对以上各步中分离出来的单层网络的权值与函数参数进行优化，而固定其余各网络层权值与动作参数值。

由于整个网络优化策略的各步都采用演化策略,每步所得的结果以实数解串形式表示,因此各步骤之间的相互转移可以通过其实数解串来实现。各步骤中的演化操作算子仅作用于相应网络节点的有关权值与动作函数的参数上，而对其余各节点的相关权值与动作函数的参数则保持固定。

4 网络训练结果

网络训练的结果以该网络模型对各输入参数的平滑、滤波与预测误差进行评价，误差分为最大误差与平均误差。表 1 为各参数自发动机启动后 6s 内之 5 步平滑、滤波和预测的误差分布情况。虽然网络在有些采样点上误差过大，但这只是初步训练结果，网络还可继续优化、训练。

Table 1 The 5-step smoothing, prediction and filtering error for 17 parameters

Parameter	Smoothing error (%)		Filtering error (%)		Prediction error (%)	
	Max	Average	Max	Average	Max	Average
PER	34.6732	4.4549	244.928	9.24204	75.1577	4.6035
POVR	77.7406	7.33381	25.6779	2.69652	207.657	8.68055
PXY	84.9672	5.29289	133.885	4.25472	74.5026	3.88169
POY	0.415868	1.114113	1.70247	0.28182	0.719296	0.155837
PEY	1.03879	0.235807	1.2204	0.201508	1.77497	0.243924
POVY	80.226	3.04823	127.78	5.69796	126.91	9.11007
PK	3.16683	0.590232	6.22657	1.0314	7.36874	1.16001
POL	10.3355	2.09517	13.5636	3.20157	16.9525	2.00488
PY	59.0659	3.69611	65.3754	3.59064	115.951	3.8484
PEVY	43.2334	2.48889	23.9074	2.5909	231.768	6.21624
PFR	3.58466	0.252161	0.916476	0.206456	0.697566	0.160885
POVFR	8.22911	0.997792	16.4779	1.5018	18.6975	1.33731
PEO	39.2502	3.22317	26.3233	1.83691	82.5326	6.82424
PEWR	11.9401	1.68452	22.7334	2.31001	24.8448	2.67092
PWEY	36.3031	6.29161	50.1917	4.47132	117.864	6.03425
PWZY	106.143	4.96069	69.4572	4.98997	44.234	3.66319
PGY	4.39296	0.901334	3.82399	0.379612	3.85488	0.364936

图 2~5 反映部分参数的平滑、滤波与预测结果，可见，该网络体系结构是比较合理的。演化策略的训练效率以网络模型总误差的收敛性来评价。图 6 及图 7 分别反映部分全局训练及各输出层训练的演化效率。数值仿真中还发现，如果稍微增加压缩层的节点数，则网络的

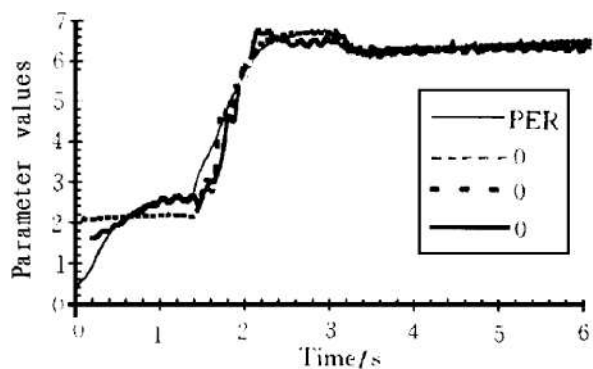


Fig. 2 The output of sensor 0, its smoothed value and filtered value as well as predicted value

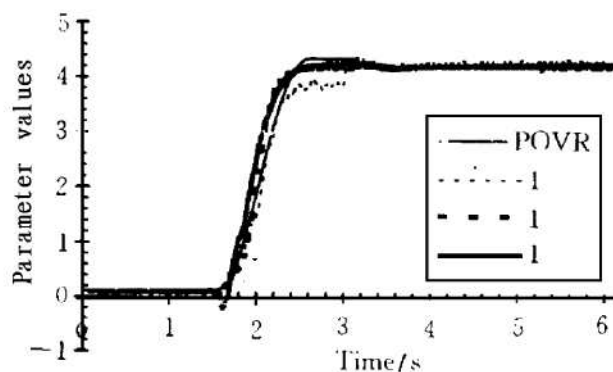


Fig. 3 The output of sensor 1, its smoothed value and filtered value as well as predicted value

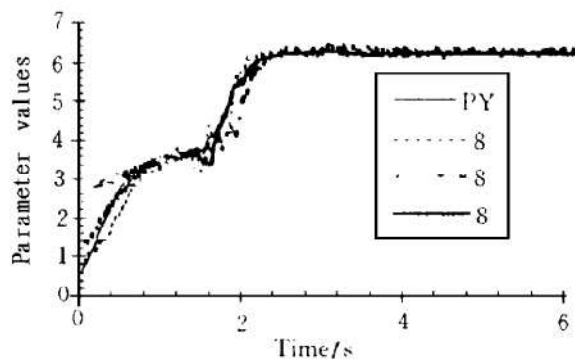


Fig. 4 The output of sensor 8, its smoothed value and filtered value as well as predicted value

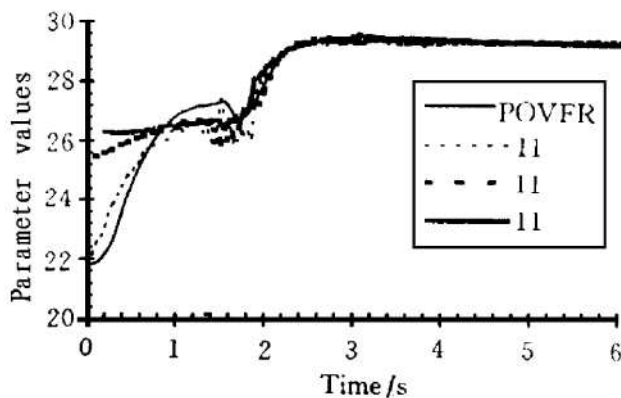


Fig. 5 The output of sensor 11, its smoothed value and filtered value as well as predicted value

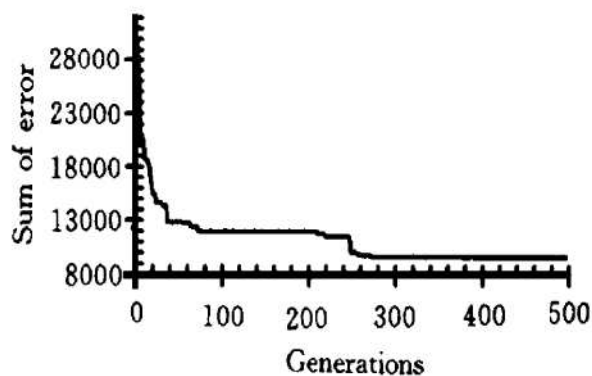


Fig. 6 The convergence characteristic of the globally training

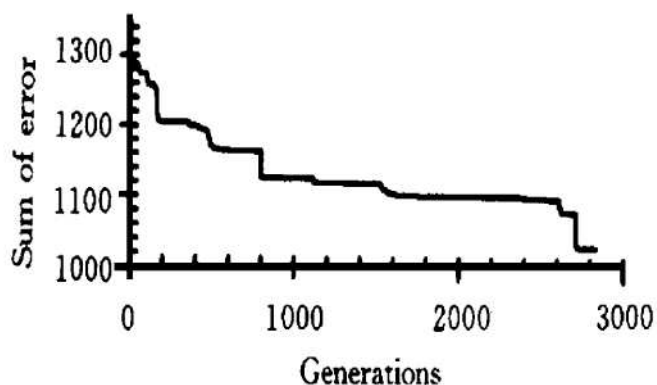


Fig. 7 The convergence characteristic of locally training the output layer

建模精度将大大提高,这说明该网络的信息含量与压缩节点数具有直接的关系。

5 结 论

巨型四层前向网络模型的训练过程与结果表明:

(1) 基于信息压缩原理的四层前向网络体系结构能成功用于高维复杂函数的逼近问题; 该模型虽然很难搜索其全局最优权值与动作函数的参数值, 但可以先搜索其全局次优模型, 然后再分步局部优化各网络层, 而且该分级优化策略对类似函数逼近问题的动态系统建模特别有效;

(2) 对动态系统进行神经网络建模时, 可适当集中样本数据, 各参数的网络误差按其变化特性进行加权处理; 演化策略在网络模型的训练中具有较大的优越性;

(3) 对巨型网络模型的分级优化本质上是对一个超高维函数变量的分部优化求解, 该模型的优化成功为演化策略解决超高维函数的优化提供了经验与策略思想。

(4) 进一步的研究应考虑演化策略与 BP 算法有机结合, 以实现随机搜索与定向搜索相结合的思想; 网络结构与权值的同步优化, 以实现最终结构神经网络建模; 对动态系统进行神经网络建模时, 应对网络的输入参数进行统计主元分析, 并采用算术集群与几何集群技术, 以减少数据点并使样本点尽量分布均匀, 对四层网络模型中的压缩层与解压层的信息压缩比率与解压能力进行定性定量分析。

参 考 文 献

- 1 鲍立威, 何 敏, 沈 平. 关于 BP 模型的缺陷的讨论. 模式识别与人工智能, 1995, 3 (1)
- 2 谢 涛, 张育林. 基于滑动数据窗聚类结果分析的复杂动力系统故障检测与诊断算法框架及其应用. 西安: 第二界全球华人智能大会, 1997.
- 3 谢 涛, 张育林. 非线性混合回归演化算法研究. 推进技术, 1998, 19 (5)
- 4 Rechenberg I. Evolutionstrategie: optimierung technischer systeme nach prinzipien der biologischen evolution. Stuttgart: Frommann-Holzboog Verlag, 1973.
- 5 Schwefel H P. Numerische optimierung of computer models. Chichester, Uk: John Wiley, 1981.

简 讯

通用发动机/威廉姆斯联合研制小型涡轮发动机

通用发动机公司 (GM) 和威廉姆斯国际公司联合研制宇航应用、主要是用于无人驾驶飞机的小型涡轮发动机。

公司相信, 对无人驾驶飞机市场, 齿轮箱可以和推进器合并成一体, 希望原型产品可以维持 18 个月, 在三年中通用发动机公司能加速生产。同时, 威廉姆斯公司正与贝尔直升机公司商议改造贝尔公司“鹰眼”无人驾驶飞机发动机事宜。该飞机目前由艾利逊公司的 C20B 涡轮轴发动机提供动力, 计划改为威廉姆斯公司的 FJX-2 型。

推力为 3kN 的 FJX-2 型发动机目前列入 NASA 的通用航空推进计划的一部分。FJX-2 型将在 2000 年左右获得涡轮和涡轮螺旋桨发动机应用。

先进低成本动力装置的核心是达到一试验标准, 原型产品将于 1999 年初进行试验。威廉姆斯国际公司希望该发动机的价格将不超过 35 万美元。

刘梦威 刘鸣雾 供稿