

文章编号: 1002-0411(2000)06-0511-05

一种新型的动态递归神经网络及其全自动设计算法

史天运 贾利民 蔡秀生

(铁道部科学研究院通信信号研究所 北京 100081)

摘要: 本文基于非线性自回归滑动平均模型 NARMA 模型和前馈神经网络建模的思想, 提出一种输入层与输出层神经元递归的动态递归神经网络; 基于进化计算中遗传算法和进化策略与自寻优 BP 算法的不同结合方式, 提出两种动态递归神经网络全自动高效设计算法, 实现了网络结构、权重和自反馈增益同时优化学习, 实例应用表明所提网络结构及其设计算法的有效性。

关键词: 动态递归神经网络, 网络结构, 遗传算法, 进化策略, 全自动设计算法

中图分类号: TP13

文献标识码: B

1 引言

动态递归神经网络^[1,2]在工业过程建模、仿真、控制及自学习方面是具有发展潜力的网络, 相比多层前馈神经网络具有如下优势: ① 不需要对系统模型进行事先定阶; ② 离线建模时避免了串并模型建模而并联模型应用的误差; ③ 应用中不必考虑系统输入和输出的历史信息, 网络本身已经考虑这些影响; ④ 避免了大量的输入节点可能对外部噪声产生的敏感。但是目前动态递归神经网络存在学习速度较慢和模型精度较低的缺点, 因而应用较差, 本文基于前馈神经网络高效建模的思路和神经网络快速自动设计的目标, 从网络结构与学习算法入手, 研究动态递归神经网络, 为推动实用化奠定坚实的基础。

2 新型的动态递归神经网络结构

部分动态递归神经网络以其反馈连接较少, 结构简单, 模型辨识效率高, 有较好的应用前景。从最简单的部分动态递归神经网络 E l m a n 网络^[2]入手, 本文引入非线性自回归滑动平均 NARMA 模型和前馈神经网络建模的思想, 从网络拓扑结构变化上, 提出一种新型的部分动态递归神经网络。

为了解决高阶系统的辨识问题, 文献[2]提出了修改的 E l m a n 网络, 如图 1 所示。

在图 1 修改的 E l m a n 网络结构单元中, 有一个固定增益权的反馈连接, 因此, 结构单元在 k 时刻的输出, 将等于隐层在 $k-1$ 时刻的输出加上结构单元在 $k-1$ 时刻输出值的 α 倍, 即:

$$x_{ci}(k) = \alpha x_{ci}(k-1) + x_i(k-1) \\ l = 1, 2, \dots, q$$

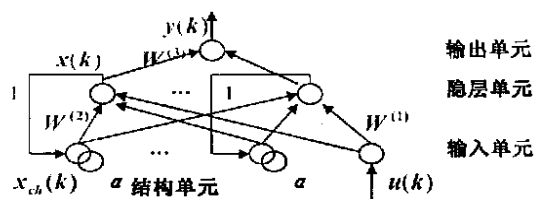


图 1 一种修改的 E l m a n 网络(HDRN 网络)拓扑结构图

其中: $x_{c,l}(k)$ 和 $x_l(k)$ 分别表示 k 时刻第 l 个结构单元和第 l 个隐层单元的输出, α 为自反馈增益. 显然, 当增益 α 为零时, 修改的 $E lman$ 网络就退化为基本的 $E lman$ 网络.

修改 $E lman$ 网络描述的非线性状态空间表达式为:

$$\begin{aligned} x(k) &= f(W^{(2)}x_c(k) + W^{(1)}u(k)) \\ x_{ch}(k) &= x(k-1) + \alpha x_{ch}(k-1) \\ y(k) &= f(W^{(3)}x(k)) \end{aligned}$$

以此结构的动态神经网络模型, 经过实际样本测试表明其学习速度较慢. 对其模型结构剖析, 与 $NARMA$ 模型对照比较, 发现 $NARMA$ 模型建立中直接考虑了系统输入和输出历史信息的影响, 与仅考虑隐层反馈的 $E lman$ 网络非直接对应关系, 而前馈神经网络建模中正是直接考虑了系统输入和输出历史信息对系统当前输出的影响, 并且具有较好的建模效率和模型

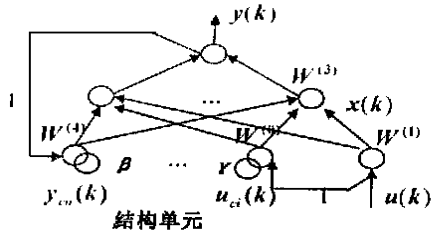


图 2 IO-DRN 网络拓扑结构图

则 IO-DRN 网络描述的非线性状态空间表达式为:

$$\begin{aligned} x(k) &= f(W^{(4)}y_{co}(k) + W^{(0)}u_{ci}(k)) + W^{(1)}u(k) \\ y_{co}(k) &= y(k-1) + \beta y_{co}(k-1) \\ u_{ci}(k) &= u(k-1) + \gamma u_{ci}(k-1) \\ y(k) &= f(W^{(3)}x(k)) \end{aligned}$$

其中: β, γ 分别为输出和输入自反馈增益.

3 动态递归神经网络全自动设计算法

以前, 动态递归神经网络最常用的学习算法是各种改进的 BP 算法, 其训练过程同其训练多层前馈神经网络一样, 仅能对其网络权重进行学习, 而对于网络结构和自反馈增益参数却只能通过人工试凑或其他结构优化方法(比如: 增加法、减少法、遗传进化法和节点删除合并法等)来确定, 并且存在学习速度较慢, 辨识精度较差, 无法实现网络结构与权重同时优化的缺点. 为了充分发挥动态递归神经网络的应用效果, 迫切需要一种高效的全自动设计算法, 本文基于全局优化算法和局部优化算法相结合的思想, 从著名的进化算法中遗传算法和进化策略特点入手, 提出与改进 BP 算法的两种不同结合方式, 不仅大大提高了学习速度, 而且实现了网络结构、权重和自反馈增益的同时优化学习, 形成两种动态递归神经网络全自动设计算法.

3.1 基于遗传算法、自寻优 BP 算法和隐节点删除合并相结合的动态递归神经网络全自动设计算法——GABP

本 GABP 算法权重学习采用遗传算法(Genetic Algorithm, 简称 GA)与改进 BP 算法相结合, 具体利用 GA 全局搜索, 避免局部极小问题, 快速获取权重的最佳初始点, 然后采取改进的 BP 算法快速局部定位. 为克服 GA 二进制编码过长, 采用十进制或十六进制编码方式, 改进的

精度. 作者受此启发, 提出基于输入层与输出层神经元递归的动态递归神经网络——简称 IO-DRN 网络. 为了比较起见, 对于上述修改的 $E lman$ 网络, 由于其仅考虑隐层神经元的递归, 所以为其命名为 HDRN 网络.

IO-DRN 网络采用输入层与输出层神经元递归的方式, 实现策略与 HDRN 网络隐层递归神经网络类似, 其具体结构如图 2 所示.

BP 算法采用具有较高效率的步长自寻优 OSBP 算法^[3], 即: 在每步学习过程中, 自动寻找最优学习率保证网络误差下降最大. 神经网络结构优化采用雷鸣等人提出的隐层节点删除或合并方法^[4], 保证网络权重学习过程中其结构同时自动优化; 自反馈增益参数优化采取一定范围 (0—1) 自动增加选取法, 保证网络递归的合理性. 该综合学习算法实现步骤如下:

- ① 选择单隐层即 3 层网络, 根据经验选择隐层节点数 (r 个输出、 p 个输入, 则隐层节点数 $q = \overline{rp}$; 或, $N_w \geq L / (1 + \log_2^2)$, N_w 为权重数目, L 为最大样本数); 选取自反馈增益参数为 0.05.
- ② 由 GA 全局搜索获取网络权重最佳初始点.
- ③ 调用 OSBP 算法调节步长, 修正权重, 训练网络, 判断迭代步数是否超过规定的步数, 是, 转向④; 否, 继续学习.
- ④ 学习精度是否达到初始要求, 是, 转向⑥; 否, 根据概率选择将隐层节点数增加一个或自反馈增益参数增加 0.01, 转向⑤.
- ⑤ 调用 OSBP 算法继续学习, 这里不再应用 GA 全局搜索获取最佳初始点, 只对新增加的隐节点有关权重选取 $[-1, 1]$ 间的随机数作为初始值, 其它权重值保持不变.
- ⑥ 根据文献^[3]提出隐节点删除或合并规则进行节点删除; 若两条规则都不满足, 则不进行节点删除, 转向步骤⑦, 否则转向③.
- ⑦ 学习精度是否达到最终要求, 否, 继续调用 OSBP 算法学习; 是, 算法终止.

3.2 基于自适应进化策略和自寻优 BP 算法相结合的动态递归神经网络全自动设计算法——ESBP

本 ESBP 算法采用全局优化算法进化策略 (Evolutionary Strategy, 简称 ES) 与改进的 BP 算法相结合自动设计动态递归神经网络.

由于 ES 算法直接用实数进行变异优化, 无需交叉操作, 所以利用 ES 设计动态递归神经网络不需要编码, 且能够实现网络权重、网络结构和自反馈增益的同时进化. 但随着网络结构的变化, 网络权重数目发生变化, 所以当对这些网络参数同时进化时, 需要对其排列一定顺序, 保证进化中随网络结构变化, 相应的权重会正确增加或删除.

假设网络输入层、隐层和输出层神经元数目依次为 r, q 和 p , 神经元函数为 Sigmoid, 即 $f(\theta, b) = (2 + e^{-\theta})^{-1}$. 令: 输入层与隐层间、结构单元中输入层递归与隐层间、结构单元中输出层递归与隐层间、隐层与输出层间的连接权依次为 $w_{ij}^{(1)}, w_{ij}^{(0)}, w_{oj}^{(4)}, w_{jm}^{(4)}$ 其中: $i = 1, 2, \dots, r; j = 1, 2, \dots, q; o, m = 1, 2, \dots, p$; 则需优化的网络参数应排列如表 1.

表 1 需优化的网络参数排列

自反馈增益		隐层节点数目	有关第 1 个隐节点所有权重				...	有关第 q 个隐节点所有权重			
			输入节点与第一个隐节点间	输出节点与第一个隐节点间	结构单元中输入层递归节点与第一个隐节点间	结构单元中输出层递归节点与第一个隐节点间		输入节点与第 q 个隐节点间	输出节点与第 q 个隐节点间	结构单元中输入层递归节点与第 q 个隐节点间	结构单元中输出层递归节点与第 q 个隐节点间
γ	β	q	$w_{i1}^{(1)}$	$w_{i1}^{(3)}$	$w_{i1}^{(0)}, \dots, w_{i1}^{(0)}$	$w_{i1}^{(4)}, \dots, w_{i1}^{(4)}$...	$w_{iq}^{(1)}$	$w_{iq}^{(3)}$	$w_{iq}^{(0)}, \dots, w_{iq}^{(0)}$	$w_{iq}^{(4)}, \dots, w_{iq}^{(4)}$

采用网络输出均方差 E 作为 ES 算法的适应度, 所有隐节点的权重按照自适应 ES^[5] 进行优化学习. 为了保证神经网络综合性能, 自反馈增益按照其进化概率 p_r 来确定是否进化, 即在

每代根据 P_r 来变异, 具体进化方法就是基本 ES. 为了实现网络结构(即隐层神经元数目)的进化, 附加增加与删除操作: 增加: 进化中按照增加概率 P_a 来决定是否增加一个隐神经元, 同时增加与之关联的连接权重, 新增权重可根据初试范围随机选择. 删除: 进化中按照删除概率 P_e 来决定是否删除一个隐神经元, 若删除, 可根据随机概率确定删除第几个隐神经元, 同时也将与之关联的连接权重删除. 对于自反馈增益进化概率 P_r 、隐层神经元数目增加概率 P_a 和删除概率 P_e , 采用自适应原则: 它们根据解的适应度减小而减小, 保证 $p_a, p_a, p_e \in [0.01, 0.3]$, 这样可保证进化的合理性、提高进化效率

所有网络参数初始值选择原则:

- 所有权重初始值范围为 $[-1, 1]$;
- 自反馈增益参数的初始值为 0.3;
- 隐层神经元数目 q 初始值根据前述的经验公式来确定.

在神经网络学习初期, 作为全局优化算法的进化策略效果较好, 但达到一定程度后还未满足最终精度要求时, 速度极慢, 大规模群体搜索代价太高, 而这时, 网络结构和自反馈增益参数已经固定下来, 为此将与 BP 算法结合, 在确定的网络结构和神经元参数基础上, 将 ES 进化所得的网络权重作为 BP 算法起始点, 采用梯度法进行快速学习. 为了提高学习效率, 采用文献 [3] 提出的步长自寻优改进 BP 算法(即 OSBP)与 $(\mu + \lambda)$ -ES 集成, 则该集成化的动态递归神经网络学习算法如下:

- ① 根据网络参数初始值选择原则随机选择 ES 的初始群体(μ 个个体), 令 $k = 1$;
- ② 对于每个个体计算其网络误差作为 ES 适应度, 并计算第 k 代 μ 个个体平均适应度 \bar{f} 和最佳适应度 f_{\min} ;
- ③ 根据自适应进化策略^[5], 优化确定第 k 代的进化步长 σ_k ;
- ④ 由自适应原则, 确定自适应策略, 计算第 k 代自反馈增益进化概率 P_r , 隐层神经元数目增加概率 P_a 和删除概率 P_e ;
- ⑤ 进化操作: 包括变异、增加、删除和选择;
- ⑥ 连续 6 次最大适应度变化是否小于 ϵ ? 是, 转向⑦; 否, 令 $k = k + 1$, 转向(2);
- ⑦ 固定网络结构和递归参数, 以此时的权重作为 BP 算法初始值, 并调用 OSBP 算法, 对权重进行学习;
- ⑧ 最终网络误差是否满足要求? 是, 停止学习; 否, 转向⑦继续学习.

4 应用结果

本文提出的新型动态递归神经网络结构及其学习算法应用于实际工程项目玻璃熔窑压力的建模中, 以其实测运行数据作为训练样本. 玻璃熔窑具体描述如图 3.

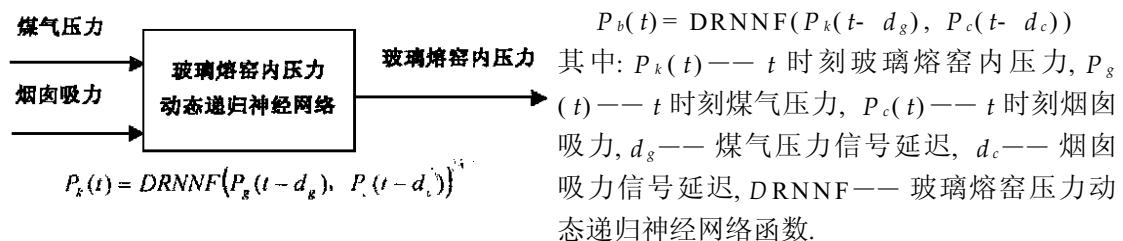


图3 玻璃熔窑描述

实际选择的学习样本和测试样本分别有

250 个和 550 个. GA 算法参数选取为: 群体规模 $n=20$, 十进制编码, 交叉概率 $P_c=0.9$, 变异概率 $P_m=0.09$. ES 算法参数选取为: $\mu=10$, $\lambda=60$, 即(10, 60)-ES; 隐层神经元初选为 5, 6, 7, 8; 网络最终学习误差 E 设定为 0.01.

经过实际训练, 两种算法均确定最优的玻璃熔窑压力递归神经网络模型结构为 $2 \times 6 \times 1$, 其中 6 为隐层节点数, 输出层递归的自反馈增益 $\beta=0.2$, 输入层递归的自反馈增益 $\gamma=0.8$, 两种算法辨识效果和测试精度如表 2.

表 2 两种算法辨识效果和测试精度

两种动态递归神经网络	GABP			ESBP		
	迭代次数(次)	训练时间(s)	测试精度	迭代次数(次)	训练时间(s)	测试精度
HDRN 网络	30+ 281395	12+ 3284	0.021153	40+ 162448	13+ 1875	0.019815
IODRN 网络	40+ 29653	18+ 258	0.019989	60+ 18813	20+ 166	0.019368

5 结论

根据上面的实例训练结果和作者进行的其它实例建模效果均表明: 在复杂系统建模中, 本文所提出的输入层与输出层神经元均递归的动态递归神经网络 IODRN 网络相比, 隐层递归的动态递归神经网络 HDRN 网络不仅具有较小的模型规模, 更且具有很高的学习效率和较好的测试精度, 非常适合于复杂系统的建模. 文中所提的两种动态递归神经网络全自动设计算法均能够实现网络结构、权重和自反馈增益的同时自动优化, 比参数分别学习大大提高了学习效率, 而且在网络设计理论上迈出了一步. 基于自适应进化策略和自寻优 BP 算法相结合设计算法具有算法简单、无需集成结构优化方法、学习效率高的特点, 是具有发展前途的算法.

参 考 文 献

- 1 S Adwankar, R N Banavar. A recurrent network for dynamic system identification. International Journal of System Science, 1997, 28(12): 1239~ 1250
- 2 孙增圻, 张再东. 智能控制理论与技术. 清华大学出版社, 1997
- 3 史天运. FMS 状态监测与故障诊断技术研究. 北京理工大学博士学位论文. 1998
- 4 雷 鸣. 神经网络自适应学习研究. 系统工程与电子技术. 1994, (3): 19~ 27
- 5 史天运, 贾利民. 基于进化策略的前馈神经网络设计研究. 1999 年中国智能自动化学术会议论文集, 清华大学出版社
- 6 B Schenker, M Agarwal. Dynamic modeling using neural network. International Journal of System Science, 1997, 28(12): 1285~ 1298

A NEW DYNAMIC RECURSIVE NEURAL NETWORK AND FULLY AUTOMATIC DESIGN ALGORITHM

SHI Tian-yun JIA Lim in CAI Xiu-sheng

(Communication & Signal Research Institute, China Academy of Railway, Beijing 100081)

Abstract Based on the modeling idea of non-linear auto-regressive moving average model and feedforward neural network, the new dynamic recursive neural network with the input and output neuron recursion is proposed. Based on the different combined ways to the genetic algorithm, evolutionary strategy

(下转第 520 页)