基于进化规划的多层前向网络结构优化 [1]

李 映 白本督* 焦李成

(西安电子科技大学雷达信号处理重点实验室 西安 710071)
*(西安电子科技大学 西安 710071)

摘 要 基于进化规划 (EP) 方法，该文提出了设计多层前向网络拓扑结构和权值分布的一种新算法——EPANN 算法. EPANN 算法能同时进化网络的结构和连接权值 (包括阈值)，在进化过程中，强调父代和子代之间的行为联结，结构变异既有结点删除，又有结点增加，不同于单纯的删除算法或构造算法，且结点删除总是先于结点增加，保证了网络规模尽可能小而泛化能力尽可能强.

关键词 进化规划，网络设计，EPANN 算法
中图号 TN-052

1 引 言

在多层前向网络中，采用 BP 算法以及其它一些基于梯度的算法时总是假定网络结构 (规模) 固定，一般来说，只有正确选择了网络结构它们才是有用的，网络规模太小则不能很好地学习问题，而规模太大又会导致过学习，使性能变差. 因此网络结构的优劣对网络的处理能力有很大的影响，一个好的结构应能圆满解决问题，同时可尽可能减少冗余结点和冗余连接. 遗憾的是，神经网络结构的设计目前基本上还依赖于人的经验，尚没有一个系统的方法来设计一个适当的网络结构. 传统上人们在自动设计网络结构时，多采用探测性的删除算法 [1] 或构造算法 [2] . 但是这样的探测方法既不能避免 BP 算法的固有不足 (特定问题的的网络结构难以确定及网络训练慢且经常陷于局部收敛)，也必将使训练出的结果网络的优劣依赖于具体的探测过程，并且一旦当前探测过程失败，目前得到的中间结果也不能继续沿用.

(准) 最优网络结构的设计可以表示为结构空间的搜索问题，其中每个点代表一种结构. 由于这样的结构空间具有规模大，不可微，复杂度高以及多模态的特点 [3]，使得传统的搜索方法很难找到 (准) 最优的结构. 而进化算法 (包括遗传算法、进化规划和进化策略等) 是一类鲁棒性很强的、并行的、全局搜索的进化方法，比删除算法或构造算法更适于网络的结构设计问题，近年来受到了广泛的重视和研究 [3−8] .

本文在 Fogel 的进化规划基础上提出一种新的算法—— EPANN 算法 (Evolutionary Programming Artificial Neural Networks) 来进化前向网络. 之所以采用进化规划方法是基于两点考虑：一是由于进化规划采用实数编码方案，这就省略了二进制数编码方案中复杂的网络编码表示与解码计算，更适合表示和操作网络；二是进化规划只需用变异算子，避免了遗传算法中的交叉算子在操作时很容易破坏已进化出的网络结构 [4] .

2 EPANN 算法的实现

2.1 基本 EPANN 算法叙述

在 EPANN 算法的实现中，对可进化的神经网络类型的约束条件很少，网络不必严格分层或全连接，并且隐结点可以具有不同的变换函数.

EPANN 算法框图如图 1 所示，其具体步骤如下：

[1] 1999-11-30 收到，2000-06-30 定稿
国家自然科学基金和国家“863”计划资助项目

(1) 随机产生一个初始群体，并对其中的每个个体 (对应一种网络结构) 进行权值训练；

(2) 按照误差值从小到大对个体进行排序；

(3) 如果最好的个体满足要求或进化时限已到，则停止；否则继续；

(4) 基于排序选择一个父代，对父代进行权值训练生成一个子代。若子代优于父代，则取代父代并返回 (2)；否则抛弃这个子代继续；

(5) 从父代的隐节点中均匀地随机选取一个进行删除，并对删除后的网络进行权值训练得到一个子代。若子代优于群体中的最差个体，则取代之并返回 (2)；否则抛弃这个子代继续；

(6) 对父代进行节点增加操作，训练这个网络得到一个子代。用这个子代取代群体中的最差个体并返回 (2)。

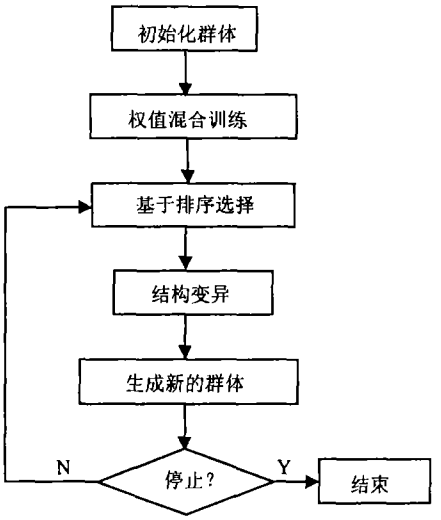

图 1　EPANN 算法框图

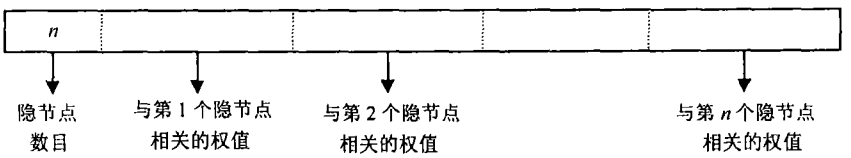

图 2　个体编码形式

2.2 算法实现

2.2.1 *编码模式*　在 EPANN 中，每个个体对应一种网络结构，我们将隐节点数目及网络权值进行混合实数编码，每一个体码串形如图 2 所示。

个体码长随隐节点数目的不同而不同。用这样的表示模式可以很容易地进行结构变异。

2.2.2 *适应度评估和选择机制*　在 EPANN 中，每个个体的适应度用 $f = 1/E$ 来表示，E 为网络的目标函数：

$$E = \sum_{t=1}^{T} \sum_{i=1}^{n} (Y_i(t) - Z_i(t))^2 \tag{1}$$

其中 $Y_i(t)$ 和 $Z_i(t)$ 是模式 t 的第 i 个结点的实际输出和期望输出，n 和 T 分别是输出结点数目和输入模式数目。

在 EPANN 中采用排序选择机制 [5]，群体中的个体就目标函数 E 的值从小到大排序，记为 $0,1,\cdots,M-1$(设群体规模等于 M)，那么第 $(M-j)$ 个个体以下列概率被选择：

$$P(M-j)=j\Big/\sum_{k=1}^{M}k \tag{2}$$

这种选择机制的优点是基于目标函数的排序而不是目标函数本身的大小，因而避免了适应度值差别太大而导致样本多样性损失太多，因此可以较好地解决不成熟收敛问题。

2.2.3 权值的混合训练 因为标准的 BP 算法及其派生的基于准确误差梯度的计算方法收敛速度慢且存在局部收敛的问题，在 EPANN 算法中权值训练采用改进的 Aploex 算法 [9]。这个算法建立在模拟退火思想基础上，搜索过程把随机性和确定性结合起来，既能避免局部最小，又能加快收敛速度。算法描述如下：

在第 k 次迭代：$\Delta E(k)=E(k-1)-E(k-2)$,

若 $\Delta E(k)<0$，权值按原相同方向前进一个 δ 长度，若结果不理想，再回退半个 δ；

若 $\Delta E(k)>0$, 则 $W_{ij}(k)=w_{ij}(k-1)+\delta_{ij}(k)$。其中

$$\delta_{ij}(k)=\begin{cases}-\delta, & \text{以概率}P_{ij}(k)\\ +\delta, & \text{以概率}1-P_{ij}(k)\end{cases}$$

概率 $P_{ij}(k)$ 为玻尔兹曼分布：$P_{ij}(k)=1/\left(1+e^{-(C_{ij}(k)/(T(k))}\right)$，其中 $C_{i,j}(k)$ 表示权值改变与误差改变的相关量，$C_{ij}(k)=\Delta w_{ij}(k)\cdot\Delta E(k)$；$T(k)$ 为温度，体现了概率 $P_{ij}(k)$ 对 $\Delta E(k)$ 的敏感性，应保证无论 $\Delta E(k)$ 大小如何，$P_{ij}(k)$ 都应明显鼓励使 $\Delta E(k)$ 显著减小的权值变化。故自适应地调整 $T(k)$，当 $k<N1$ 时，取一大数，$k>N1$ 时，$T(k)=\frac{\delta}{N}\sum_{k1=k-N1+1}^{k}|\Delta E(k1)|$。

2.2.4 结构变异 在 EPANN 算法中，结构变异算子为结点删除和结点增加。输入结点和输出结点数目由给定问题而定，实际对隐结点进行删、减操作，并且结点删除总是先于结点增加，只有当删除结点不能产生一个好的子代时，才增加结点。这样顺序变异的目的是为了保证网络规模尽可能的小，避免了在适应度函数中增加一个网络复杂性调整项，例如 $f=f_{\text{error}}+\alpha\cdot f_{\text{complexity}}$，这需要反复试验才能确定合适的系数 α 值。

另外，在一般的算法中，结点增加操作都是随机增加一个具有全连接的隐结点 (或隐层)，连接权的初始值随机选择，这样做很容易破坏父代已经学习好的行为，使父代和子代之间的行为联结变差。而在 EPANN 算法中，结点增加是通过分裂一个现有的隐结点来实现的，Odri 等人 [10] 称之为“细胞分裂”，“细胞分裂”比随机增加结点能更好地维持父代和子代之间的行为联结。在所有隐结点中均匀地随机选取一个结点进行分裂，分裂后的结点具有下列连接权值：

$$\left.\begin{aligned}&w_{ij}^1=w_{ij}^2=w_{ij}, && i\geq j\\ &w_{ki}^1=(1+\alpha)w_{ki},\ w_{ki}^2=-\alpha\cdot w_{ki}, && i<k\end{aligned}\right\} \tag{3}$$

其中 w 是原来结点 i 的权矢量，w^1 和 w^2 是分裂后的两个结点的权矢量。

3 实 验

我们采用 EPANN 算法来设计一个三层前向网络的雷达目标分类器。

表 1　雷达目标识别的实验参数

群体规模	20	权值训练步长	0.005
最大生成代数	500	权值训练迭代次数	100
初始隐结点数	2~5	Aploex 中的温度设置	100
		温度调整参数	10

在微波暗室中，利用雷达转台对三类飞机目标 F1, F2, F3 的缩比模型进行成像实验，每类飞机各成像 750 个，对应观测角为 $0^\circ \sim 100^\circ$，经频域变换和归一化处理后，得到目标的 64 维的距离像模式样本集。图 3 画出了三类飞机在某一方位角的一维距离像。均等地抽取每类目标的 200 个像，并分成四组，每组样本包括各类飞机的 50 个 64 维的模式向量。其中三组用来训练网络，另外一组用来对网络进行测试。

根据样本维数和类别数确定输入层结点数为 64，输出层结点数为 3，以输出 $(1,0,0)$, $(0,1,0)$, $(0,0,1)$ 分别代表目标 F1， F2， F3 的理想识别结果。实验参数如表 1 所示。网络经过训练后，得到分类器的最佳结构为 64-3-3 。把识别结果按正确识别、错误识别和无法识别三种情况对各类飞机的识别率统计如表 2 所示。

从实验结果中可以看到，由 EPANN 算法可以设计出非常紧凑的网络结构 (64-3-3)，这对于分类器来说非常重要，因为分类器的结构越精简，则其泛化能力越好。这一点从分类器的识别结果中得到了验证：分类器对检测样本的总正确识别率为 90.4%，表明该分类器具有较好的泛化能力。

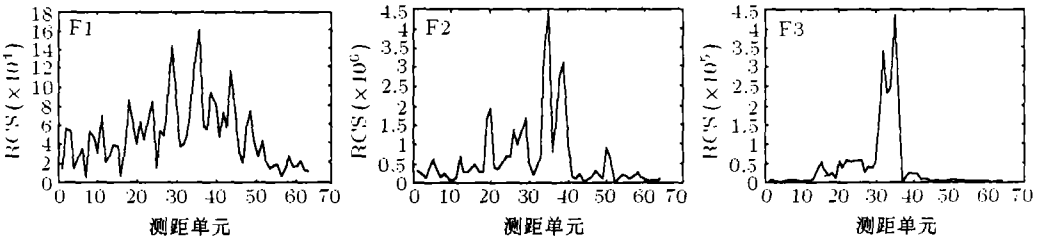

图 3　三类飞机在某一方位角的一维距离像

表 2　分类器的识别结果

识别情况	正确识别 (%)			错误识别 (%)			无法识别 (%)		
飞机类型	F1	F2	F3	F1	F2	F3	F1	F2	F3
训练样本	100	100	0	0	0	0	0	0	0
检测样本	89.2	90.3	91.8	3.2	3.4	2.5	7.6	6.3	5.7

4 结束语

本文在进化规划的基础上提出一种设计前向网络拓扑结构和权值分布的新算法—— EPANN 算法，该算法在进化过程中强调父代和子代之间的行为联结，采用结点删除先于结点增加的结构变异顺序和结点分裂等方法，以保证网络规模尽可能小而泛化能力尽可能强。 EPANN 算法对于可行的网络结构限制条件很少，因而对不同网络的搜索空间很大，并且由于其全局搜索能力，可以避免结构的局部最小。实验结果表明用 EPANN 算法设计出的雷达目标神经网络分类器结构紧凑，具有较好的泛化能力。

参 考 文 献

[1] R. Reed, Pruning algorithms—A survey, IEEE Trans. on Neural Networks, 1993, NN-4(5), 740–747.

[2] T. Y. Kwok, D. Y. Yeung, Constructive algorithms for structure learning in feedforward neural networks for regression problems, IEEE Trans. on Neural Networks, 1997, NN-8(5), 1131–1148.

[3] V. Maniezzo, Genetic evolution of the topology and weight distribution of neural networks, IEEE Trans. on Neural Networks, 1994, NN-5(1), 39–53.

[4] P. J. Angeline, G. M. Saunders, J. B. Pollack, An evolutionary algorithm that constructs recurrent neural networks, IEEE Trans. on Neural Networks, 1994, NN-5(1), 54–64.

[5] X. Yao, Y. Liu, Evolutionary design of artificial neural networks with different nodes, Proc. of ICEC'96, Nagoya, Japan, 1996; 670–675.

[6] X. Yao, Y. Liu, A new evolutionary system for evolving artificial neural networks, IEEE Trans. on Neural Networks, 1997, NN-8(3), 694–713.

[7] G. W. Greenwood, Training partially recurrent neural networks using evolutionary strategies, IEEE Trans. on Speech & Audio Processing, 1997, 5(1), 192–194.

[8] P. G. Harrald, M. Kamstra, Evolving artificial neural networks to combine financial forecasts, IEEE Trans. on Evolutionary Computation, 1997, 1(1), 40–51.

[9] 徐立新，等，神经网络训练的一种随机方法，哈尔滨工业大学报， 1996, 29(3), 85–87.

[10] S. V. Odri, D. P. Petrovacki, G. A. Krstonosic, Evolutional development of a multilevel neural network, Neural Networks, 1993, 6(4), 694–713.

STRUCTURAL OPTIMIZATION OF MULTILAYER FEEDFORWARD NETWORKS BASED ON EVOLUTIONARY PROGRAMMING

Li Ying Bai Bendu* Jiao Licheng

(*Key Lab for Radar Signal Processing, Xidian University, Xi'an* 710071, China)

*(*Xidian University, Xi'an* 710071, China)

Abstract Based on evolutionary programming, a novel algorithm named EPANN for designing the topology and weight distributions of feedforward networks is proposed in this paper. EPANN algorithm evolves network architectures and connection weights (including biases) simultaneously and emphasizes the behavioral links between parents and their offspring in evolution, such as weights training after each architectural mutation and node splitting. Unlike the pure constructive or pruning algorithm, EPANN's architectural mutations include both node deletion and node addition, and prefer node deletion to addition in order to encourage the network architecture as compact as possible and generalization ability as good as possible.

Key words Evolutionary programming, Network design, EPANN algorithm

李 映： 女， 1969 年生，博士生，主要研究方向为模式识别和信号处理，计算智能与混合系统等.

白本督： 男， 1972 年生，硕士生，主要兴趣为模式识别，神经网络等.

焦李成： 男， 1959 年生，教授，博士生导师，主要研究方向包括：计算智能和集成系统，非线性智能信号和图像处理.