

文章编号: 1671-7848(2007)04-0387-04

## 基于遗传算法的前向神经网络结构优化

王宏刚, 钱 锋

(华东理工大学 自动化研究所, 上海 200237)



**摘 要:** 对近几年应用遗传算法(Genetic Algorithm, GA)优化设计前向神经网络结构的研究进行了评述。指出了神经网络结构优化设计的重要性和目前各种方法存在的不足。介绍了神经网络结构设计原理和应用 GA 优化设计神经网络应着重考虑的两个问题: 即结构表达策略和适应度函数设计。分别对近来应用 GA 优化设计多层感知器、径向基函数神经网络和径向基概率神经网络结构的研究进行了细致介绍和分析。指出了目前研究工作的不足和未来研究工作的方向。

**关键词:** 遗传算法; 前向神经网络; 结构优化; 多层感知器; 径向基函数; 径向基概率神经网络

中图分类号: TP 183

文献标识码: A

## Structural Optimization of Feedforward Neural Network Based on Genetic Algorithms

WANG Hong-gang, QIAN Feng

(Institute of Automatic Control, East China University of Science and Technology, Shanghai 200237, China)

**Abstract:** A review of the current research state on applying genetic algorithm to structural optimization for feedforward neural networks in recent years is presented. The significance of structural design optimization for ANN and the drawback of the current methods are analyzed. The principle of ANN structural design is introduced and two critical issues on architecture encoding strategy and fitness function design are proposed. The research on applying genetic algorithm to design optimization for multilayer perception, radial basis function and radial basis probabilistic neural networks are introduced and analyzed respectively in detail. According to the problems in current research, the future research directions are pointed out.

**Key words:** genetic algorithms; feedforward neural networks; structural optimization; multilayer perception; radial basis function; radial basis probabilistic neural networks

### 1 引言

人工神经网络(ANN)经历近 60 年的发展后, 其应用已成功渗透到各个领域。虽然 Hornik 等人证明<sup>[1]</sup>只用一个隐含层的前馈网络就能逼近任意复杂度的函数, 但实际应用中神经网络的结构选择仍取决于使用者的经验难以确定, 有学者称其为 NP 问题<sup>[2]</sup>。神经网络的结构选择直接影响到其使用效果。以隐含层节点数为例, 隐含层节点数量太少, 拟合精度很难达到要求; 隐含层节点数量过多, 可能出现过度拟合问题, 不仅降低泛化能力还会增加训练时间。可见, 有必要寻找一种能够代替人进行决策的优化方法。

近年来相继被提出的神经网络结构优化方法主

要有规则化方法、增长方法、修剪方法和进化方法。然而, 规则化方法随着候选模型的增加, 获得最佳模型的概率降低; 而增长方法和修剪方法分别面临“何时终止增加神经元”和“何时停止修剪”的问题, 极可能导致网络过拟合和欠拟合。

目前, 用遗传算法优化神经网络结构是一个研究热点, 本文综述近 6 年应用 GA 优化三种前向神经网络结构的研究工作, 分析讨论未来发展方向。

### 2 GA 优化 ANN 网络结构

1) ANN 结构优化 它包括拓扑结构和节点传递函数的确定。对于多层感知器(MLP)网络<sup>[1]</sup>, 结构优化主要包括隐层数、各隐层节点数及其对应的连接数的确定; 对于径向基函数(RBF)网络<sup>[3]</sup>, 由

收稿日期: 2006-06-13; 收修定稿日期: 2006-07-03

基金项目: 国家 973 计划资助项目(2002CB3122000); 上海市自然科学基金资助项目(05ZR14038); 上海科委科技攻关基金资助项目(04DZ11010\_05DZ11C02); 上海市科委重大基础研究资助项目(05DJ14002)

作者简介: 王宏刚(1983-), 男, 辽宁铁岭人, 研究生, 主要研究方向为智能仿生优化算法与复杂工业过程建模与控制等; 钱 锋(1961-), 男, 江苏扬中人, 教授, 博士生导师。

于只有一个隐层,故结构优化只包括隐层节点数和其对应的基中心确定;对于近年来迅速发展的径向基概率神经网络(RBPNN)<sup>[4]</sup>,其结构优化内容比 RBF 网络多一项:核函数控制参数。

ANN 结构优化问题可形式化为结构空间的优化问题<sup>[5]</sup>:该空间中每一点代表一种结构,给定一种性能指标(如最低网络训练误差或网络复杂性),各网络结构在结构空间中形成各自结构性能曲面,曲面最高点对应最优网络结构。性能曲面所特有的离散性、多峰性、不可微性、复杂有噪声、欺骗性等特征<sup>[6]</sup>使得 GA 成为最合适的优化工具。

2) GA 优化 ANN 结构原理 应用 GA 优化 ANN 结构的过程为:首先随机产生不同结构的神经网络集(种群),把每个网络结构编码成个体,即每个个体代表一种网络结构;再用不同的初始权值对种群中的网络结构训练,并计算适应度大小;选择若干适应度值较大的个体直接继承到下一代,对当前代群体进行遗传操作,产生下一代群体,重复执行上述操作,直到某个个体达到要求为止。

应用 GA 优化 ANN 结构的首要问题是网络结构的染色体表达策略,根据个体的编码包含多少网络结构信息,可分为直接编码策略和间接编码策略。

直接编码策略:网络结构的每个连接权值都编码成二进制串,网络的拓扑结构由一连接矩阵表示,矩阵元素的值(1 或 0)代表对应网络节点是否有连接,直接编码就是将连接矩阵的行或列元素串接形成染色体。该编码策略简单直观,染色体中包含的结构信息较完备,结构的进化更加精确细微,易于进化出优质紧凑的网络结构,适用于小结构神经网络的进化。缺点是随着网络规模的增加,编码长度变长,导致计算量成指数级增加;另外该策略不能保证所有网络结构都是合理的。

间接编码策略:每个基因位与 ANN 的每个连接权不再有一一对应关系,只对网络结构的重要特征进行编码,其他结构细节则根据先验知识预定或通过结构“生成”规则确定,特点是从网络结构到其基因表达至少需要一层解释机制。已经提出的间接编码策略主要有两种:图生成系统方法<sup>[7]</sup>和参数表达方法<sup>[8]</sup>。该编码策略由于染色体长度显著缩短,可提高进化效率,并且保证不会出现不合理结构,但会导致进化规则更加复杂,而且染色体中一位的变化会导致完全不同的网络结构。

应用 GA 优化 ANN 结构的另一重要问题是适应度函数设计。ANN 结构优化的目标为针对领域问题获得良好的网络结构,并使之尽可能满足训练速度快、逼近精度高和泛化能力强等性能指标,故 GA 的适应度函数设计考虑的因素主要为这三个性能指标或三者的折衷。

### 3 基于 GA 的前向 ANN 结构优化综述

Miller 等人<sup>[6]</sup>于 1989 年最早提出应用 GA 对 ANN 结构优化设计,随后便有大量研究和文献综述,主要集中对 MLP 网络<sup>[9]</sup>和 RBF 网络<sup>[10]</sup>的结构优化。本节主要针对近年来国内外应用 GA 对 MLP 网络、RBF 网络以及 RBP 网络结构优化设计的工作进行综述。

1) GA 优化 MLP 网络结构 Boozarjomehry 和 Svrcek<sup>[11]</sup>应用 L 系统<sup>[12]</sup>将演化规则集编码成二进制染色体,采用间接编码机制设计网络结构。为避免出现计算量随神经元增加成指数级增加,令规则集从单方向演化;为提高搜索速度,限制搜索空间为前向网络;适应度函数采用训练误差、校验误差和网络连接数目的线性组合,采用改进的 GENITOR 算法<sup>[13]</sup>。根据 pH 中和过程和 CSTR 反应的动态性能预测,所得网络在训练速度和预测精度上均优于常规方法。

Arifovic 和 Gencay<sup>[14]</sup>采用直接编码机制的遗传算法演化 MLP 网络结构。每个二进制染色体分三段分别表示隐层节点数、输入节点和初始权值范围;随机生成 500 个权值用共轭梯度法进行监督式训练。算法新颖之处是采用联赛选择和竞选操作算子<sup>[15]</sup>结合进行选择,自适应改变变异率加快算法收敛速度。外汇回报预测结果表明:与传统模型选择方法 SIC<sup>[16]</sup>相比,预测精度提高,但隐层节点数较多。

张栋和蔡开元<sup>[17]</sup>提出基于 GA 的前向神经网络两阶段学习方法。作者首先提出 GA 不合同时演化神经网络结构和参数,然后提出先用实数编码的 GA 进行结构选择和初始参数优化,再用 LM 算法<sup>[18]</sup>对网络参数进行精细优化,适应度函数采用 SRC 准则<sup>[19]</sup>,即综合考虑网络复杂度,训练精度和泛化能力。利用该方法设计神经网络对一个两输入单输出的非线性函数进行建模,结果令人满意,但未给出与其他方法的对比。

李智勇和童调生<sup>[20]</sup>提出基于多物种进化遗传算法(SEGA)的神经网络结构进化设计方法。基本思想是将遗传基因分为物种基因(物种属性)和个体基因(个体特征);在结构学习时,ANN 的结构信息编码为物种基因,连接权值编码为其个体基因,这种层次编码方法可克服简单遗传算法存在的适应度评价噪音的问题;物种适应度函数侧重于泛化能力,个体适应度函数侧重于训练精度;该方法能使结构学习与权值学习同时进行,电力负荷建模仿真证明了方法的可行性。

Barrios 等人<sup>[21]</sup>设计出一种能用 GA 自动设计并训练单隐层前向神经网络的遗传神经网络系统

(GANN)。系统由结构设计模块和权值训练模块组成,结构设计模块采用基于“元结构”的二进制编码和海明交叉算子的 GA,所得网络结构解码后再用实数编码的 GA(权值训练模块,采用形态交叉算子)优化权值;适应度函数设计包括网络复杂度和训练精度。“元结构”编码和海明交叉操作的结合提高了 GA 的局部搜索能力和收敛速度,和形态交叉操作结合保证能够得到最小网络;方法优点是适应值对个体基因位的敏感度降低,即一位的变化将产生非常相似的网络结构,一定程度上解决直接编码机制存在的敏感性问题<sup>[22]</sup>。乳腺癌细胞检测识别结果与文献<sup>[23]</sup>的 GANNET 相比,收敛速度和网络复杂度有明显提高。

Sexton 等人<sup>[24]</sup>开发一种同时优化前向神经网络结构和权值的系统(NNSOA)。在算法<sup>[25]</sup>基础上进行改进,新算法的适应度函数包括网络复杂度和预测精度;每次变异操作后,随机对某个个体中的某个连接权置零,原个体和新个体的取舍由适应值大小决定;网络训练采用增长法,记录并自动更新最优个体,若连续三次增加节点未更新则算法停止。对 Building 数据集<sup>[26]</sup>建模结果表明,NNSOA 系统能较好地除去网络不必要连接,提高泛化能力和速度得到网络性能明显优于修剪法和增长法。

Sexton 等人<sup>[27]</sup>在文献<sup>[24]</sup>基础上对 NNSOA 系统的信息挖掘能力研究,将 NNSOA 系统应用于“贷款批准决策”数据集分类测试,结果表明该系统既有较好的分类能力,又能提取重要的输入因素,为工作者提供领域问题的额外信息,一定程度上克服了 ANN 黑箱的缺点。

Dam 和 Saraf<sup>[28]</sup>受文献<sup>[11]</sup>的启发利用 GA 优化设计的前向神经网络用于原油分馏产品的物性预测。算法采用直接编码机制,交叉操作局限于具有相同层数的网络结构,遗传操作后对个体进行合理性检验,适应度函数只考虑训练精度,不考虑网络复杂度;采用精英保留策略的基本 GA 算法,对原油分馏产品的比重预测,预测结果在实测结果的误差范围内,但网络的泛化能力较差,作者将其归因于训练样本的有限性;笔者认为由于算法仅包括训练精度未包括网络复杂度,而训练精度高的网络结构未必有很好的泛化能力<sup>[29]</sup>,所以原因在于适应度函数设计存在问题。

2) GA 优化 RBF 网络结构 Sumanthi 等人<sup>[30]</sup>采用 GA 优化 RBF 网络的隐中心矢量。网络参数用二进制编码,采用随机保留选择机制和矩阵交叉操作,第二层网络权值由奇异值分解算法得出。方法应用于图像分割和心脏病数据分类,结果表明 RBF 网络分类效果有明显提高。

Leung 等人<sup>[31]</sup>应用 GA 搜索 RBF 网络的最优隐

中心、方差和隐层节点数。用实数编码网络参数,并采用近似的 Akaike 信息准则(AIC)控制网络复杂度。对复杂环境中物体的探测结果表明,该 GA/RBF 系统能够抽取对象的非线性特性,与文献<sup>[32]</sup>所得结果相比,探测效果有显著提高。

受 Liu<sup>[33]</sup>的启发,Jiang 等人<sup>[34]</sup>构造一种结构模块化神经网络(SMNN)。该网络将 BP 网络和 RBF 网络结合,期望得到泛化能力强和计算速度快的网络,应用 GA 优化网络的隐节点和 RBF 网络的高斯函数基中心,染色体有两部分,固定长度的二进制串代表 BP 节点数目;可变长度的实数编码串代表 RBF 隐节点位置和数目。染色体解码后利用 LM 算法训练得到的均方误差(MSE)作为其适应值。水泥强度建模实验结果显示,SMNN 在预测准确性和网络复杂性方面优于单独的 BP 网络和 RBF 网络,不足之处为 GA 算法运行时间太长。

3) GA 优化 RBP 网络结构 Guo 等<sup>[35]</sup>首次使用 GA 优化径向基概率神经网络(RBPNN)的结构。优化对象为隐中心矢量的数目和位置,以及高斯核函数的形状参数。采用二进制与实数混合编码:二进制基因位“1”表示样本被选择,“0”表示不被选择;染色体的最后一位用实数编码,代表相应的核函数控制参数。双螺旋问题分类的结果表明,用 GA 优化后的网络不仅结构缩小,泛化能力也明显提高。

赵温波等<sup>[36]</sup>在文献<sup>[37]</sup>的基础上提出用微遗传算法(uGA)<sup>[38]</sup>结合最大绝对误差算法(MAEA)对径向基概率神经网络结构优化。MAEA 能克服 GA 与 ROLSA 的运算速度慢和易陷入局部最小的缺点,并继承了增长方法的优点,用于优选隐中心矢量;uGA 是基于小种群和再初始化操作的特殊遗传算法,具有较好的收敛速度和精度,用于优化核函数控制参数。uGA 的交叉算子使用两点随机交叉,不使用变异操作。双螺旋问题和 IRIS 问题分类结果表明,与 ROLSA 算法和 MKM 算法相比,该算法效率更高,所得网络泛化能力更强。

## 4 结 语

本文综述近 6 年国内外应用 GA 对 MLP 网络、RBF 网络和 RBP 网络结构优化设计的研究。与回归、聚类、交叉校验法、增长法和修剪法相比,应用 GA 优化前向 ANN 能得到“更满意”的网络结构,体现出广阔的应用前景;然而,应用 GA 优化前向 ANN 网络结构存在一些问题。

首先是搜索速度有待提高。目前实验性研究主要限于小规模网络和离线使用,而实际工业过程应用常要求大规模网络和在线使用,优化过程应快速及时。然后是结构表达策略需要改进。直接编码策

略导致的基因型结构与显形网络的一对多映射使适应度评价存在噪声；间接编码策略易导致敏感性问题，即染色体中一位的变化会导致完全不同的网络结构。另外是多种不确定因素给操作者带来不便。GA/ANN系统的调试结果受多种因素影响：GA编码方式、种群初始化方法、遗传算子设计、适应度设计、ANN训练方法等。

综上所述，本文对基于GA的前向神经网络结构优化的研究方向归纳为：

- ①对GA的改进算法有待继续研究，如继续发展并行遗传算法(PGA)，以提高GA的搜索速度。
- ②开展ANN结构表达策略的研究，针对现有表达策略的适应度噪声和敏感性问题特点，应大胆将图论、集合、形态学等工具引入ANN结构表达策略中。
- ③加强对GA和ANN的理论研究，将数学引入到GA和ANN的定量分析，并争取得出确定性的结论作为操作准则。
- ④将现有的不同结构优化方法思想融合，集成各方法的优点，扬长避短，最大限度发挥现有方法的优势。

#### 参考文献(References)：

- [1] Hornik K M, Stinchcombe M. Multilayer feedforward networks are universal approximators[J]. *Neural Networks*, 1989, 2(2): 359-366.
- [2] Judd J S. Learning in networks is hard[C]. San Diego, CA: Proc the 1th IEEE International Conference on Neural Networks, 1987.
- [3] Park J, Sandberg I W. Universal approximation using radial basis function networks[J]. *Neur Comput*, 1991, 3(2): 246-257.
- [4] Huang D S. Radial basis probabilistic neural networks model and application[J]. *International Journal of Pattern Recognition and Artificial Intelligence*, 1999, 13(7): 1083-1101.
- [5] 何永勇, 褚福磊, 钟秉林. 基于进化计算的神经网络设计与实现[J]. *控制与决策*, 2001, 16(3): 257-262. (He Yongyong, Chu Fulei, Zhong Binglin. Artificial neural networks design and implementation based on evolutionary computation[J]. *Control and Decision*, 2001, 16(3): 257-262.)
- [6] Miller G F, Todd P M, Hegde S U. Designing neural networks using genetic algorithms[C]. Morgan Kaufmann: Proceedings of International Conference on Genetic Algorithms and Their Applications, 1989.
- [7] Kitano H. Designing neural networks using genetic algorithms with graph generation system[J]. *Complex Systems*, 1990, 4(4): 461-476.
- [8] Vonk E, Jain L C. Using genetic algorithm grammar encoding to generate neural networks[J]. *Proc IEEE Int Conf Neural Networks*, 1995, 4(6): 1928-1931.
- [9] Yao X. Evolving artificial neural networks[J]. *Proc IEEE*, 1999, 87(9): 1423-1447.
- [10] Harpham C, Dawson C W, Brown M R. A review of genetic algorithms applied to training radial basis function networks[J]. *Neural Comput & Applic*, 2004, 13(3): 193-201.
- [11] Boozarjomehry R B, Svrcak W Y. Automatic design of neural network structure[J]. *Computers and Chemical Engineering*, 2001, 25(7): 1075-1088.
- [12] Lindenmayer A. Mathematical models for cellular interactions in development inputs[J]. *Journal of Theoretical Biology*, 1968, 18(3): 280-289.
- [13] Boozarjomehry R. Application of artificial intelligence in feedback linearization[D]. Calgary: The University Calgary, 1997.
- [14] Arifovic J, Gencay R. Using genetic algorithms to select architecture of a feedforward artificial neural network[J]. *Physica A*, 2001, 289(3): 574-594.
- [15] Arifovic J. Genetic algorithm learning and the Cobweb mode[J]. *J Econ Dyn Control*, 1994, 18(1): 3-28.
- [16] Schwarz G. Estimating the dimension of a mode[J]. *The Annals of Statistics*, 1978, 6(2): 461-464.
- [17] 张栋, 蔡开元. 基于遗传算法的神经网络两阶段学习方案[J]. *系统仿真学报*, 2003, 15(8): 1088-1090. (Zhang Dong, Cai Kaiyuan. A genetic-algorithm-based two-stage learning scheme for neural networks[J]. *Journal of System Simulation*, 2003, 15(8): 1088-1090.)
- [18] Hagan M T, Menhaj M B. Training feedforward networks with the Marquardt algorithm[J]. *IEEE Trans Neural Networks*, 1994, 5(6): 989-993.
- [19] Wang L, John Y. Extracting fuzzy rules for system modeling using a hybrid of genetic algorithm and Kalman filter[J]. *Fuzzy Sets and System*, 1999, 101(3): 353-362.
- [20] 李智勇, 童调生. 基于多物种进化遗传算法的神经网络结构学习方法[J]. *计算机工程与应用*, 2003, 22(8): 87-90. (Li Zhiyong, Tong Tiaosheng. The ANN architecture learning method based on SEGA[J]. *Computer Engineering and Applications*, 2003, 22(8): 87-90.)
- [21] Barrios D, Carrascal A, Manrique D, et al. Cooperative binary-real coded genetic algorithms for generating and adapting artificial neural networks[J]. *Neural Comput & Applic*, 2003, 12(2): 49-60.
- [22] Dorado J. Cooperative strategies to select automatically training patterns and neural architectures with genetic algorithms[D]. Spain: University of La Coruna, 1999.
- [23] Robbins G E, Plumbley M D, Hughes J C, et al. Generation and adaptation of neural networks by evolutionary techniques[GANNET] J]. *Neural Computing and Applications*, 1993, 1(1): 23-31.
- [24] Sexton R S, Dorsey R E, Sikander N A. Simultaneous optimization of neural network function and architecture algorithm[J]. *Decision Support Systems*, 2004, 36(3): 283-296.
- [25] Dorsey R E, Mayer W J. Genetic algorithms for estimation problems with multiple optima, non-differentiability, and other irregular features[J]. *Journal of Business and Economic Statistics*, 1995, 13(1): 53-66.
- [26] Prechelt L. A study of experimental evaluations of neural network learning algorithm: current research practice[R]. Germany: Fakultat fur Informatik, Universitat Karlsruhe, D-76128 Karlsruhe, 1994.
- [27] Sexton R S, McMurtrey S, Cleaver D J. Knowledge discovery using a neural network simultaneous optimization algorithm on a real world classification problem[J]. *European Journal of Operational Research*, 2006, 168(3): 1009-1018.
- [28] Dam M, Saraf D N. Design of neural networks using genetic algorithm for on-line property estimation of crude fractionator products[J]. *Computers and Chemical Engineering*, 2006, 30(4): 722-729.
- [29] Baldi P. Learning in linear neural networks: a survey[J]. *IEEE Trans on Neural Networks*, 1995, 6(4): 837-858.
- [30] Sumathi S, Sivanandam S N, Ravindran R. Design of a soft computing hybrid model classifier for data mining applications[J]. *Engin Intell Sys Electr Engin Comm*, 2001, 9(1): 33-56.
- [31] Leung H, Dubash N, Xie N. Detection of small objects in clutter using a GA-RBF neural network[J]. *IEEE Trans Aero Electr Sys*, 2002, 38(1): 98-118.
- [32] Leung H, Lo T. Chaotic radar signal processing over the sea[J]. *IEEE J Ocean Eng*, 1993, 18(3): 287-295.
- [33] Liu Y, Yao X. Evolutionary design of artificial neural networks with different nodes[A]. *Evolutionary Computation[C]*. Nagoya Japan: Proc IEEE Int Conf, 1996.
- [34] Jiang N, Zhao Z Y, Ren L Q. Design of structural modular neural networks with genetic algorithm[J]. *Adv Eng Soft*, 2003, 34(1): 17-24.
- [35] Guo L, Huang D S, Zhao W B. The optimization of radial basis probabilistic neural networks based on genetic algorithms[C]. Portland: Proceedings of the International Joint Conference on Neural Networks, 2003.
- [36] 赵温波, 王立明, 黄德双. 最大绝对误差结合微遗传算法优化径向基概率神经网络[J]. *计算机研究与发展*, 2005, 42(2): 179-187. (Zhao Wenbo, Wang Liming, Huang Deshuang. Structure optimization of radial basis probabilistic neural networks by the maximum absolute error combined with the micro-genetic algorithm[J]. *Journal of Computer Research and Development*, 2005, 42(2): 179-187.)
- [37] Zhao W B, Huang D S, Guo L. Optimizing radial basis probabilistic neural networks using recursive orthogonal least squares algorithms combined with micro-genetic algorithms[C]. Portland Oregon: The Int'l Joint Conf on Neural Networks, 2003.
- [38] Goldberg D E. Sizing population for serial and parallel genetic algorithms[C]. California: The 3rd Conf on Genetic Algorithms, 1989.