

文章编号: 1001 - 9014 (2010) 02 - 0136 - 04

用入侵的自适应遗传算法训练人工神经网络

王改良, 武妍

(同济大学 计算机科学与技术系, 上海 201804)

摘要: 给出了一种能和网络结构一一对应的、合适的染色体编码方法. 用物种入侵的遗传算法训练人工神经网络, 在入侵过程中, 遗传算法自适应地调整交叉算子和变异算子. 提出了一种根据平均适应度值确定入侵物种规模的方法, 并详细描述了算法步骤, 最后通过实验证明了本文算法的有效性和优越性.

关键词: 入侵; 遗传算法; 自适应; 人工神经网络; 编码

中图分类号: TP183 **文献标识码:** A

TRAINING ARTIFICIAL NEURAL NETWORK BY INVADING ADAPTIVE GENETIC ALGORITHM

WANG GaiLiang, WU Yan

(1. Department of Computer Science and Technology, Tongji University, Shanghai 201804, China)

Abstract: A suitable chromosome encoding method, which could correspond with the network one by one, was proposed. The species invasion genetic algorithm was used to train artificial neural networks. In the invading process, the genetic algorithm adjusts adaptively crossing operation and mutation operation. A method based on the average fitness values was proposed to determine the scale of invasion species, and a detailed description of the algorithm steps was given. Finally, the validity and superiority of the algorithm are proved by the experimental results.

Key words: invading; genetic algorithm; adaption; artificial neural network; coding

引言

用遗传算法来设计和训练人工神经网络已被证实能有效地解决很多问题^[1,4,6]. 针对染色体编码与网络对应问题, 文章 [2, 3] 都提出了自己的方法. 文章 [3] 先用 GA 训练人工神经网络的权值和阈值, 然后用 Levenberg Marquardt (LM) 算法处理, 得到较好的效果; 也有用改进交叉变异算子的遗传算法优化神经网络^[5,7]. 文章 [9] 直接用遗传算法优化神经网络参数; 还有用遗传算法和 BP (Back Propagation) 算法同时训练网络的^[8]; 文章 [11] 中提出混沌搜索自适应入侵遗传算法训练网络参数, 但没有对网络结构优化.

生物中存在种群入侵, 入侵的种群带来新的物种——新的网络拓扑结构, 促进个体进化; 不同物种之间的杂交能产生更适应环境的个体, 加速个体进化速度. 根据这一特点, 文章给出一种染色体编码方法, 使染色体很好地与网络对应并适应相同物种和

不同物种间个体的交叉, 并给出了交叉和变异算子. 本文提出的算法能优化网络权值、阈值以及网络拓扑结构.

1 基于入侵自适应遗传算法的网络优化

采用三层前向神经网络, 对其结构 (指网络隐层节点数) 以及权值、阈值进行优化, 使网络节点数最佳, 权值、阈值落在全局最优区域.

1.1 染色体编码

就具体问题而言, 网络输入层节点数及输出层节点数都已确定, 其隐层节点数不确定. 图 1 是图 2 网络对应的染色体编码. 两隐层节点阈值之间的等位基因表示前一隐层节点与输入层之间连接权值, 权值为 0 表示无连接. 输出层节点也是如此.

1.2 物种入侵过程以及入侵物种的生成

1.2.1 入侵过程

在网络输入层和输出层节点数确定的情况下, 将隐层节点数相同而权值和阈值不同的网络看作同

收稿日期: 2009 - 03 - 26, 修回日期: 2009 - 10 - 12

Received date: 2009 - 03 - 26, revised date: 2009 - 10 - 12

作者简介: 王改良 (1983-), 男, 湖南永州人, 硕士研究生, 主要从事智能系统与模式识别研究.



图1 染色体编码
Fig.1 Chromosome code

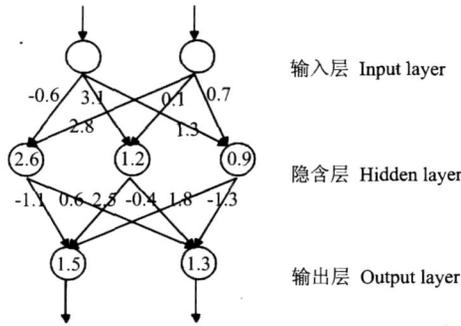


图2 网络结构
Fig.2 Network architecture

一物种, 仅将隐层节点数不相等的网络看作不同物种. 在网络训练中, 先形成一种或几种物种, 每一代都产生新物种入侵, 与已有物种竞争, 优胜劣汰.

1.2.2 入侵物种生成

根据经验公式 $N_h = \sqrt{N_i + N_o} + l, l \in [1, 10]^{[5]}$ 随机生成正整数 k 使得 $k \in [\sqrt{N_i + N_o} + 1, \sqrt{N_i + N_o} + 10]$ 接着随机生成一个实数序列长度为 N , N 根据式 (1) 计算, 其中 k 是隐含层节点数, N_i 是输入层节点数, N_o 是输出层节点数.

$$N = (N_i + N_o + 1) \times k + N_o \quad (1)$$

假设根据式 (1) 随机生成的实数序列为: $A = [a_1, a_2, \dots, a_n] \{-1 < a_i < 1 | i = 1, 2, \dots, n\}$, 则入侵物种个体染色体编码为 $x = [k, a_1, a_2, \dots, a_n]$.

1.3 适应度函数

式 (2) 为网络性能评价函数. X_i 是单个个体, s_{sum} 是样本总数, e_i 是单个样本网络输出误差平方, c 是常数 ($0 < c < 1$).

$$f(X_i) = \sum_{i=1}^{s_{sum}} e_i + c \quad (2)$$

式 (3) 为适应度函数. 由式 (2) 和式 (3) 可知当个体适应度值 $Fit(X_i) = \frac{1}{c}$ 时, 个体适应度值达到最大, 此个体最优.

$$Fit(X_i) = \frac{1}{f(X_i)} \quad (3)$$

1.4 交叉

同一物种间交叉算子: 设交叉的两个父代个体

染色体编码分别是 x_A, x_B . 两个个体按式 (5) 和式 (6) 进行交叉运算产生下一代个体染色体编码分别是 x'_A, x'_B

$$x'_A = x_B + (1 - \lambda) x_A \quad (5)$$

$$x'_B = x_A + (1 - \lambda) x_B \quad (6)$$

本文采用式 (7) 自适应调整参数 λ . f_{av} 个体平均适应度值, f_{max} 是两个个体中适应度函数值较大的. 随着个体适应度值增大而减小. 适应度值大的个体在较小的范围内进行搜索, 避免搜索的波动性, 保护适应度值大的个体, 加快算法的搜索速度.

$$\lambda = \frac{1}{f_{av}} \times \left(\frac{1}{c} - f_{max} \right) \quad (7)$$

不同物种间的交叉算子: 设交叉的两个个体染色体分别是 Y_A, X_A , 长度分别是 L_Y, L_X , 设 $L_Y < L_X$, 交叉后产生新的后代 Y_A, X_A . Y_A, X_A 长度、第一个等位基因分别与父代 Y_A, X_A 一致. 具体交叉如下.

从个体 X_A 染色体中随机选取 $L_Y - 1$ 个等位基因 (不包括第一个等位基因) 组成序列 x_A ; 个体 Y_A 去掉第一个等位基因得到序列 y_A , 序列 x_A, y_A 按式 (5) 和式 (6) 交叉产生新序列 x'_A, y'_A . 把个体 Y_A 染色体第一个等位基因加到序列 y'_A 的第一位, 产生新个体 Y'_A ; 用序列 x'_A 对应替换掉 X_A 中的 x_A , 得到新的个体 X'_A .

1.5 变异

采用动态单点变异, 设个体 $A = [a_1, a_2, \dots, a_k, \dots, a_n]$, a_k 发生变异, 变异后基因 a_k 按式 (8)、(9) 计算, t 是当前迭代次数, T 是总迭代数, $Fit(A)$ 个体 A 适应度值. 由式 (9) 可以看出, 随着迭代次数增加, 以及个体适应度值升高, 越趋近 1, $1 - \lambda$ 趋近 0, 基因变异范围变小; 特别当 $\lambda = 1$ 时, 式 (8) 变为 $a_k = a_k$, 说明个体不发生变异完全适应环境. 式 (9) 是算法自适应的调节变异算子.

$$a_k = a_k \pm (1 - \lambda) a_k, k = 1, \dots, n \quad (8)$$

$$\lambda = \frac{t}{T} \times Fit(A) \times c \quad (9)$$

1.6 入侵物种数量确定

$$N_v = N_c \times \left[\frac{1}{2} - \frac{1}{1 + \exp[f_{av} - 1/c]} \right] \quad (10)$$

N_v 是入侵物种个体总数, N_c 是当前种群中个体总数, f_{av} 是当前种群个体平均适应度值. 式 (10) 能很好地调节入侵物种的规模, 模拟自然界的入侵规律.

1.7 网络训练算法步骤

步骤一: 设置总迭代次数 T , 交叉概率 λ , 变异概率 μ , 种群大小 M ; 置当前迭代次数 $t = 0$; 随机生

成 M 个个体, 每个个体的第一个等位基因 k 应满足 $k \in [\sqrt{N_i + N_o} + 1, \sqrt{N_i + N_o} + 10]$, N_i 是输入层节点数, N_o 是输出层节点数; 选择激活函数, 设置常数 c , 计算所有个体适应度值。

步骤二: 用轮盘赌的方法选择个体配对交叉, 变异产生下一代。

步骤三: 随机生成入侵种群, 入侵种群规模由式 (10) 确定, 每个个体的第一个等位基因 k 满足 $k \in [\sqrt{N_i + N_o} + 1, \sqrt{N_i + N_o} + 10]$, 入侵种群和遗传变异种群合并成一个种群, 计算该种群中各个体的适应度值。最高适应度值达到要求, 或者达到最大迭代次数, 算法停止, 否则选择 M 个函数适应度值最高的个体, 形成新一代种群, 加 1 转向步骤二。

步骤四: 输出适应度值最高的个体, 如果有多个适应度值相同个体, 则选择染色体长度最短的个体作为输出。

2 实验及结果分析

2.1 异或问题

为了证明本文算法能快速得到更简单的网络结构, 将该算法运用于解决异或问题并与文献 [2] 提出的 NNDT 算法和文献 [11] 提出的 NEVT 算法比较, 借此进一步说明能对复杂网络结构进行优化。算法中常数 $c = 0.01$, 初始种群大小 100, 经过 200 次实验的平均数据以及与其它算法的比较如表 1 所示 (其它算法的数据来自文献 [2])。

文献 [2] 提出 NNDT 使用结构化网络编码方法, 用启发式规则约束网络拓扑结构的变异概率以及变异操作的选择趋势, 并使用小生境机制进行拓扑变异保护; 文献 [10] 提出 NEVT 用自适应调整变异概率遗传算法训练人工神经网络。从表 1 可以看出本文提出的算法在计算单元数比其它两种算法的要少, 计算单元数越少算法速度越快。本文最终得到的平均网络结构比 NNDT 算法以及 NEVT 算法得到的网络结构简单。因此本文算法能快速训练前馈网络并能得到更简单的网络结构。这是由于物种入侵能带来更多新网络, 促使算法从更加广阔的解空间

进行搜索, 找到更优的解。

2.2 Iris 分类

为证明本文算法快速收敛, 泛化能力强, 进一步证明本文算法的高效性和优越性, 接着将其用于 Iris 三种花型分类, 并用简单遗传算法训练人工神经网络 (简称 GA-ANN) 和 BP 算法来训练神经网络以及文章 [3] 算法与本文算法作比较。文献 [3] 提出人工神经网络与遗传算法相接合的混合方法基于多种群遗传算法优化网络结构和权值, 然后用 LM 算法对网络进一步训练。

实验中从三类数据中各随机选 40 组数据用于训练, 余下的数据用于测试, 测试数据总共 30 组。网络输入层节点数为 4, 输出层节点数为 1。本文算法: 种群大小 100, 交叉概率 0.9, 变异概率 0.05, 迭代次数 100 次, $c = 0.01$, 经训练得隐层输入节点 9; GA-ANN 算法: 在相同条件下设置隐层节点数为 9; BP 算法: 迭代次数 100 次, 隐含层节点数也设置为 9, 学习率为 0.75; 文章 [3] 算法: 种群大小 100, 交叉概率 0.9, 变异概率 0.05, 隐层节点数集合 6~10, 迭代次数 100 次; 表 2 是四种算法对 Iris 数据分类的结果。表 3 是四种算法对训练数据完全正确分类的情况下比较测试数据分类正确率。

从表 2 可以看出, 本文算法对训练数据分类正确率和 GA-ANN 一样, 但比 BP 算法分类正确率高, 因此可以得知本文算法对训练数据分类能力与 GA-ANN 相当, 但比 BP 算法和文章 [3] 提出的算法分类能力强; 而测试数据分类正确率比 BP 算法、GA-ANN、文章 [3] 算法正确率高, 说明本文算法训练后的网络的泛化能力比 BP 算法、GA-ANN、文章 [3] 算法训练的网络泛化能力强。表 3 进一步说明本文算法的效率比 BP 算法、GA-ANN 和文章 [3] 算法高, 同时表明本文算法的泛化能力比其它三种算法强。这是由于物种入侵形成竞争, 促进物种进化使算法能快速找到更优解。

表 2 Iris 分类结果比较

Table 2 Comparison result of Iris classification

算法	训练数据分类正确率 (%)	测试数据分类正确率 (%)
本文算法	99.56	98.21
BP	98.61	94.87
GA-ANN	99.56	95.65
文章 [3] 算法	99.32	97.39

表 1 异或问题结果比较

Table 1 Comparison results for XOR problem

算法	平均计算单元数	隐含层平均节点数	平均连接数	种群大小
本文算法	2986.75	2.23	6.95	100
NNDT	3308.62	2.33	7.91	100
NEVT	5904.64	4.62	15.3	13

表3 Iris分类测试数据结果比较

Table 3 Comparison result of test data of Iris classification

算法	迭代次数	测试数据分类正确率(%)
本文算法	150	98.57
BP	200	96.94
GA-ANN	210	97.85
文章[3]算法	195	98.19

3 结论

提出了一种基于物种入侵自适应遗传算法的设计和训练人工神经网络的算法.给出了一种与网络对应的染色体编码方法.该编码方法能应用于相同物种和不同物种个体之间的交差操作,并使染色体与网络一一对应;给出了相应的交差算子和具体的网络训练算法.物种入侵扩大算法在解空间中的搜索范围,使算法能得到更优解——即网络结构更简单,隐层节点数更少,连接数更少,泛化能力更强的人工神经网络;自适应调整,加速算法的收敛过程,提高了算法的效率.最后的仿真实验有力地证明了本文算法的高效性和优越性.

REFERENCES

- [1] Zhang Q W, Wang C. Using Genetic Algorithm to Optimize Artificial Neural Network: A Case Study on Earthquake Prediction [C]. Genetic and Evolutionary Computing, 2008: 128—131.
- [2] WU Yan, WAN Wei. Method on designing and training of artificial neural network based on genetic algorithm [J]. *J. Infrared Millim. Waves* (武妍, 万伟. 基于遗传算法设计和训练人工神经网络的方法, 红外与毫米波学报), 2007, 26(1): 65—68.
- [3] Chen Z P, Zhao Y M, Zheng Y, et al. Neural Network Electrical Machine Faults Diagnosis Based on Multi-population GA [C]. IEEE International Joint Conference, 2008, 3795—3798.
- [4] Zhao X M, Li H Y, Zhang J. GA in Optimized Control of Central Air-conditioning System Based on ANN Simulation [C]. Eighth ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing, 2007, 3: 617—622.
- [5] MA Fang-Fang. Optimization of neural networks with a real-coded genetic algorithms [J]. *Microprocessors* (马芳芳. 基于实数编码遗传算法的神经网络优化设计. 微处理机), 2008(3): 107—109.
- [6] Ma J, Wang B S, Ma Y G. ANN-based Real-Time Parameter Optimization via GA for Super Heater Model in Power Plant Simulator [C]. Machine Learning and Cybernetics, 2008 International Conference 2008, 4: 2269—2273.
- [7] Pan S T, Wu C H, Lai C C. The Application of Improved Genetic Algorithm on The Training of Neural Network for Speech Recognition [C]. Information and Control, Second International Conference on Innovative Computing, 2007: 168—171.
- [8] LI Wei-Chao, SONG Da-Meng, CHENG Bin. Artificial neural network based on genetic algorithm [J]. *Computer Engineering and Design* (李伟超, 宋大猛, 陈斌. 基于遗传算法的人工神经网络. 计算机工程与设计), 2006, 27(2): 316—318.
- [9] Grzesiak L M, Meganck V, Sobolewski J, et al. Genetic Algorithm for Parameters Optimization of ANN-based Speed Controller [C]. The International Conference on "Computer as a Tool", 2007: 1700—1705.
- [10] Tsou Y R, Spitsyn V G. Using Genetic Algorithm With Adaptive Mutation Mechanism for Neural Networks Design and Training [C]. The 9th Russian-Korean International Symposium, 2005: 709—714.
- [11] WU Yan, FENG Zhao. Adaptive invading genetic algorithm based on chaos search [J]. *Journal of computer applications* (武妍, 冯钊. 一种基于混沌搜索的自适应入侵遗传算法. 计算机应用), 2008, 28(1): 101—103.
- [12] Adibi A, Buse K, Psaltis D. Two-center holographic recording [J]. *J. Opt. Am. B*, 2001, 18(5): 584—601.
- [13] Adibi A, Buse K, Psaltis D. Theoretical analysis of two-step holographic recording with high-intensity pulses [J]. *Phys. Rev. A*, 2001, 63(2): 023813-1-023813-17.
- [14] ZHANG Tian-Hao, LI Kun, SHAO Wei-Wei, et al. TE and TM mode photorefractive surface waves [J]. *J. Infrared Millim. Waves* (张天浩, 李坤, 邵伟伟, 等. TE和TM模式光折变表面波. 红外与毫米波学报), 2008, 27(1): 60—64.
- [15] SHEN Yan, ZHANG Guo-Quan, FU Bo, et al. Studies on nonvolatile holographic recording for $\text{LNbO}_3:\text{Fe}$ with different stoichiometry [J]. *J. Infrared Millim. Waves* (申岩, 张国权, 付博, 等. 对不同组份 $\text{LNbO}_3:\text{Fe}$ 非挥发全息存储的研究. 红外与毫米波学报), 2005, 24(4): 269—272.
- [16] YAO Jiang-Hong, XUE Liang-Ping, YAN Bo-Xia, et al. Optical parametric oscillator based on periodically poled $\text{MgO}:\text{LNbO}_3$ crystal [J]. *J. Infrared Millim. Waves* (姚江宏, 薛亮平, 颜博霞, 等. 周期极化掺镁铌酸锂晶体的光学参量振荡. 红外与毫米波学报), 2007, 34(2): 209—213.
- [17] SHEN Yan, SUN Xiu-Dong, ZHAO Ye-Quan. Theoretical studies on nonvolatile holographic recording for $\text{LNbO}_3:\text{Fe}:\text{Mn}$ [J]. *J. Infrared Millim. Waves* (申岩, 孙秀冬, 赵业权. $\text{LNbO}_3:\text{Fe}:\text{Mn}$ 晶体非挥发全息存储性能的理论研究. 红外与毫米波学报), 2008, 27(6): 421—424.

(上接 96页)