

# 粗神经网络的禁止搜索训练算法研究

梅晓丹, 孙圣和

(哈尔滨工业大学自动化测试与控制系, 黑龙江哈尔滨 150001)

**摘 要:** 粗神经网络不仅可以处理明确定量的输入信息, 而且可以处理传统神经网络不能处理的语义形式的输入信息. 本文给出了一种新的训练方法——基于禁止搜索算法的粗神经网络训练方法. 仿真结果表明本文提出的训练算法在收敛速度和网络性能上都有显著提高.

**关键词:** 粗神经网络; 禁止搜索算法; 遗传算法

**中图分类号:** TP183      **文献标识码:** A      **文章编号:** 0372-2112 (2001) 12A-1908-04

## Rough Neural Network Using Tabu Search Algorithm

MEI Xiao-dan, SUN Sheng he

(Dept. of Automatic Test and Control, Harbin Institute of Technology, Harbin, Heilongjiang 150001, China)

**Abstract:** Rough neural network can be used to process not only definite input but also semantic input which conventional neural network can not process. This paper proposes a new training method based on tabu search algorithm. Experimental results illustrate that the new training algorithm improves in convergence speed and network performance.

**Key words:** rough neural network; tabu search algorithm; genetic algorithm

### 1 引言

传统的神经网络, 如前馈神经网络、反馈神经网络以及自组织神经网络中, 每个神经元的输入、输出在训练以及应用中都是一个明确的值. 即使在模糊神经网络中, 神经元也是用一个定量值来代替一个固定的隶属度, 不能处理语义形式的输入信息. Lingras<sup>[1,2]</sup>应用一个粗神经元来代替传统神经网络中的一个神经元, 以求增加神经网络的信息理解能力. 但实质上这仍然是一个传统的神经网络. 而由文献[3]提出的粗神经网络则完全由粗神经元构成, 无论在结构和组成上都是一种全新的神经网络, 其输入具有语义性, 提高了神经网络多信息的综合能力. 但由于粗神经网络结构的特殊性, 很难应用传统的神经网络学习算法进行训练, 文献[3]中采用遗传算法<sup>[4-6]</sup>对网络进行训练, 但收敛慢, 导致训练时间过长. 本文提出采用禁止搜索算法对粗神经网络进行训练.

禁止搜索算法是一种简单灵活的全局优化技术. 它通过一系列随机移动来搜索可行解, 并禁止某些搜索方向以避免死循环和跳离局部极小. 它的随机移动可以部分地或完全地记录在禁止表中, 从而避免以后搜索过程中的重复搜索. 它使用多个候选解以引导搜索向最优目标函数值的方向进行. 同时, 算法作次优的移动允许从先前的局部最优点出发继续搜索. 禁止表主要记忆已搜索过的解的关键特征, 那么后面的搜索将根据禁止表中所记忆的特征来阻止搜索过程回到以前已经搜索过的解, 从而使得搜索过程经济有效地进行. 禁止搜索

已经在组合优化、智能控制、机器学习等领域中得到广泛应用<sup>[7]</sup>.

为了减少网络训练时间和提高网络性能, 本文提出用禁止搜索算法训练粗神经网络, 以求尽可能找到全局优化的网络权值, 并同遗传训练算法进行了比较.

文献[3]中提出的粗神经网络模型是用于一维和二维高斯数据的分类识别问题. 因此本文也采用该分类问题进行仿真.

本文的第二部分介绍了粗神经网络的结构; 第三部分阐述了禁止搜索算法的原理; 第四部分介绍了基于禁止搜索算法的粗神经网络学习; 第五部分给出了仿真结果及分析; 第六部分得出了结论.

### 2 粗神经网络

粗集理论是 Z. Pawlak 教授等提出的一个用于数据分析的数学分支<sup>[8-10]</sup>. 它可以对不完整、不精确和不确定数据进行分析 and 推理, 提取出有用信息, 是智能信息处理的有力数学工具. 文献[3]将粗集理论和人工神经网络技术结合在一起, 根据粗集理论中等价关系  $R$  上、下近似的概念, 把一个神经元分成上、下两层分别代表粗集的上下近似, 并用这样的粗神经元组成整个神经网络, 形成了全新意义的粗神经网络.

粗神经网络与传统神经网络的不同主要表现在神经元的构成和连接方式, 如图 1 所示. 由图可以看出, 每个粗神经元

由上下两层组成. 两个粗神经元的基本连接如图 1(a) 所示, 它们有 15 种连接组合, 常见组合如图 1(b) 所示.

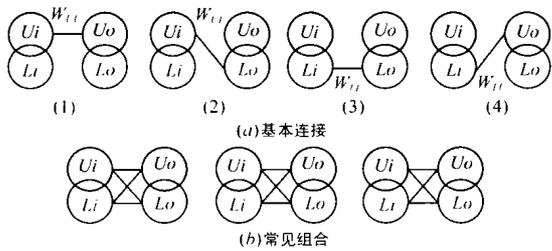


图 1 粗神经元的连接方式

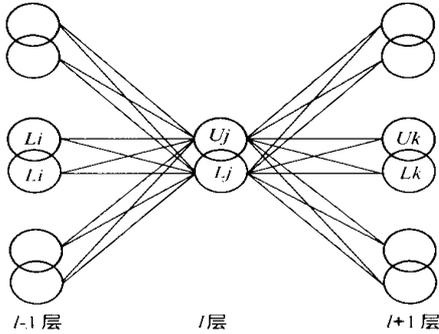


图 2 粗神经元连接

考虑如图 2 所示的粗神经元连接, 其中  $i$  是和  $j$  相连的  $l-1$  层中的粗神经元. 对于  $l$  层的粗神经元  $j$ , 其输入状态计算如下:

$$\text{Input}_{Uj} = \sum_{\text{connection from } i \text{ to } j} (w_{Uij} \times \text{Output}_{Li} + w_{LUj} \times \text{Output}_{Li}) \quad (1)$$

$$\text{Input}_{Lj} = \sum_{\text{connection from } i \text{ to } j} (w_{Lij} \times \text{Output}_{Li} + w_{LUj} \times \text{Output}_{Li}) \quad (2)$$

粗神经元  $j$  的输出可用下式计算:

$$\text{Output}_{Uj} = \max\{f_U(\text{Input}_{Uj} - \theta_{Uj}), f_L(\text{Input}_{Lj} - \theta_{Lj})\} \quad (3)$$

$$\text{Output}_{Lj} = \min\{f_U(\text{Input}_{Uj} - \theta_{Uj}), f_L(\text{Input}_{Lj} - \theta_{Lj})\} \quad (4)$$

其中  $\theta_{Uj}$  和  $\theta_{Lj}$  为粗神经元  $j$  上下两层的阈值函数,  $f_U$  和  $f_L$  为上下两层的传递函数. 它们可以相同也可以不同. 在本文中,  $f_U$  和  $f_L$  均取为正切 Sigmoid 函数, 有:

$$f_U(x) = f_L(x) = \frac{2}{1 + e^{-x}} - 1 \quad (5)$$

当传递函数确定后, 我们可以调节权值改变粗神经元上下两层的抑制和激励关系. 在全互连的情况下, 假定  $w_{LU} =$

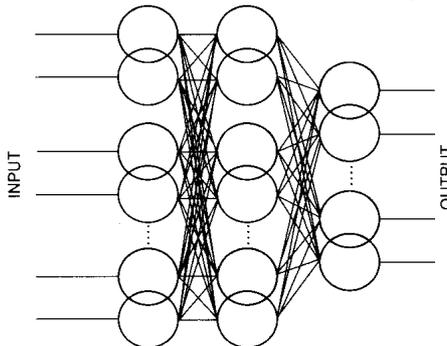


图 3 粗神经网络

$w_{LU} = w_{LU} = w_{LL}$ , 则粗神经元的上下两层的输出相同. 此时一个粗神经元相当于有两个相同输出的传统神经元. 通过分别设定权值和连接情况, 我们得到多种不同的神经网络结构. 图 3 所示为一个粗神经网络的一般连接. 由于每个神经元的输出有两值且有大小之分, 很难直接求其导数. 因此传统的 BP 算法较难于粗神经网络的训练.

### 3 禁止搜索算法

禁止搜索算法的基本原理是通过一系列移动来搜寻可行解的空间, 禁止某些搜索方向以避免死循环(避免重复搜索路径)和跳离局部极小(全局优化). 这些移动的特征部分地或完全地记录在禁止表中, 以禁止后面的迭代重复前面进行过的搜索动作. 基本的禁止搜索算法可以大致描述如下:

禁止搜索算法( )

- {
- 随机产生初始解并作为当前解和最优解;
- 如果迭代次数没达到
- {
- 在当前解的邻域内产生候选解;
- 求出各候选解相应的目标函数值;
- 按目标函数值及禁止表来更新和修改当前解、最优解和禁止表以求逐渐向最优解逼近;
- }
- 最终的最优解即为所求;
- }

### 4 基于禁止搜索算法的粗神经网络训练

采用禁止搜索算法训练粗神经网络, 算法中最优解的搜索过程就是寻找网络最优权值和阈值的过程, 我们结合仿真试验采用的粗神经网络来说明禁止搜索的过程. 仿真试验采用的粗神经网络如图 4 所示, 其输入层由 4 个神经元组成, 只起到数据输入的作用, 隐层由 4 个神经元组成, 输出层有 1 个神经元. 根据上下近似的概念, 每个神经元有 2 层权值, 即上层权值和下层权值.

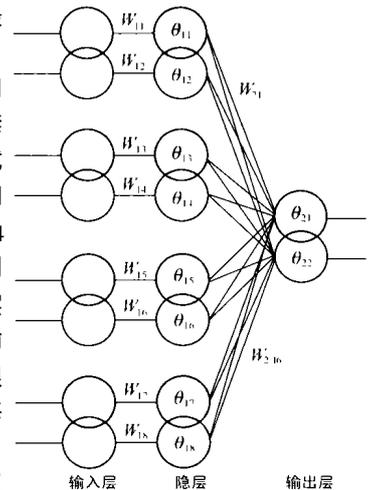


图 4 仿真试验采用的粗神经网络结构

禁止搜索算法的主要步骤如下:

步骤一: 初始化.

首先确定解的组成和结构. 对于图 4 所示的粗神经网络, 解的组成如下图所示:

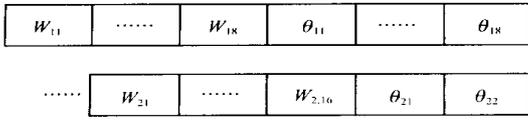


图 5 解的组成

随机地产生一个含有  $ls$  个单元的当前解, 其中  $ls$  ( $ls = 8 + 8 + 16 + 2 = 34$ ) 表示一个解中权值和阈值的个数, 然后随机交换两个单元, 产生  $N_s$  个候选解, 将交换单元的位置标号记录. 同时设置迭代次数  $t = 1$ .

步骤二: 评估.

计算每个候选解的误差函数值, 即网络实际输出与理想输出的差值.

步骤三: 排序.

按照误差由小到大排列这  $N_s$  个候选解. 相应的位置标号纪录也要按此顺序重新排列.

步骤四: 禁止.

排列后的位置标号纪录一一同禁止表比较, 如果不同, 则把该位置对应的候选解作为当前解进行下一次迭代, 并将其位置标号添加到禁止表中; 如果相同, 则判断该位置对应的候选解的误差是否小于最优解的误差, 如果小于, 把该候选解作为当前解, 并更新最优解, 继续迭代, 否则重复这次迭代.

步骤五: 禁止表溢出.

如果禁止表中的纪录个数已达到规定数目, 下一个填充的禁止纪录将从禁止表的起始位置插入, 禁止表的最后一个纪录被淘汰溢出.

步骤六: 终止条件判断.

若迭代次数达到预定值, 结束迭代. 此时, 最优解对应的参数作为粗神经网络的参数.

### 5 仿真试验

仿真试验是基于粗神经网络在数据分类方面的应用.

第一个仿真试验是对一维高斯分布数据进行分类. 对于一维高斯数据, 有:

$$\begin{aligned} \xi &\sim N(0, 1): & H_1 \\ \eta &\sim N(1, 1): & H_0 \end{aligned}$$

其中  $N(m, \sigma^2)$  表示均值为  $m$  方差为  $\sigma^2$  的高斯分布, 如

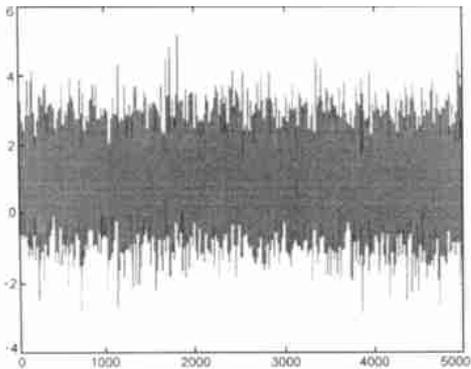


图 6 一维高斯数据分布

图 6 所示.

第二个仿真实验是对二维高斯分布数据进行分类. 对于二维高斯数据, 有:

$$\begin{aligned} (\xi, \eta) &\sim N(0, 0, 1, 1, 0) & : H_1 \\ (\xi_1, \eta_1) &\sim N(1, 1, 1, 1, 0) & : H_0 \end{aligned}$$

其中  $N(m_1, m_2, \sigma_1^2, \sigma_2^2, \rho)$  为相关系数为  $\rho$ , 均值分别为  $m_1$  和  $m_2$ , 方差分别为  $\sigma_1^2$  和  $\sigma_2^2$  的二元高斯分布, 如图 7 所示.

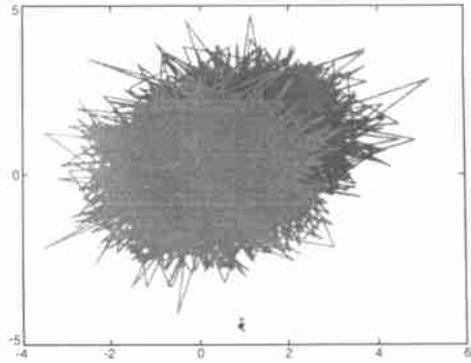


图 7 二维高斯数据分布

该实验所用的粗神经网络如图 4 所示. 其输入层中粗神经元的



图 8 一维高斯数据的输入

数据信息是坐标值. 对于一维高斯数据, 每一个数据从输入层的 4 个神经元输入, 每个神经元的上下层的输入是相同的, 网络的输入数据如图 8 所示, 四个神经元的输入是相同的. 对于二维高斯数据, 取每个数据水平方向的起点  $x_s$  和终点  $x_e$  分别输入一个神经元的上下层; 垂直方向的起点  $y_s$  和终点  $y_e$  分别输入另一个神经元的上下层. 网络的输入数据如图 9 所示, 四个神经元的输入是两两相同的. 粗神经网络最终输出取输出层神经元上下两层输出的均值.

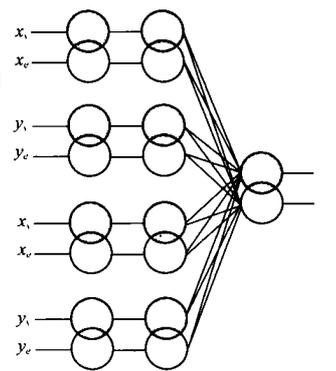


图 9 二维高斯数据的输入

分别用基于遗传算法和禁止搜索算法训练的粗神经网络对这些数据点进行

分类. 第一个实验的训练数据和测试数据为图 6 中两类数据各 500 点. 第二个试验的训练数据和测试数据为图 7 中两类数据各 500 点. 对于  $H_0$  和  $H_1$  两类数据, 神经网络的目标输出分别为 1 和 -1. 测试时阈值分别取 0.9 和 -0.9.

用来训练粗神经网络的遗传算法采用如下的参数, 种群

大小, 交叉概率, 变异概率分别为 20, 0.6 和 0.1. 训练迭代次数从 10, 50 到 100 都做了实验. 禁止搜索算法中解改动概率阈值, 每次迭代产生的候选解个数和禁止表长度分别为 0.06, 20 和 20. 训练迭代次数也从 10, 50 到 100 都做了实验. 表 1 和表 2 给出了两种算法仿真结果的比较. 图 10 和图 11 是迭代 100 次时两种算法误差量级对照.

表 1 GA-RNN 和 TS-RNN 训练一维高斯数据结果比较

| 训练迭代次数 |         | 10                    | 50                    | 100                   |
|--------|---------|-----------------------|-----------------------|-----------------------|
| GA-RNN | 训练时间(s) | 35.98                 | 188.34                | 370.95                |
|        | 误差指标    | $8.39 \times 10^{-2}$ | $6.84 \times 10^{-2}$ | $4.24 \times 10^{-2}$ |
|        | 分类正确率   | 64%                   | 73%                   | 89%                   |
| TS-RNN | 训练时间(s) | 41.3                  | 193.66                | 372.26                |
|        | 误差指标    | $6.60 \times 10^{-2}$ | $4.15 \times 10^{-2}$ | $3.73 \times 10^{-2}$ |
|        | 分类正确率   | 76%                   | 91%                   | 93%                   |

表 2 GA-RNN 和 TS-RNN 训练二维高斯数据结果比较

| 训练迭代次数 |         | 10    | 50                    | 100                   |
|--------|---------|-------|-----------------------|-----------------------|
| GA-RNN | 训练时间(s) | 43.77 | 213.66                | 404.03                |
|        | 误差指标    | 8.93  | $3.86 \times 10^{-1}$ | $4.85 \times 10^{-2}$ |
|        | 分类正确率   | 67%   | 81%                   | 92%                   |
| TS-RNN | 训练时间(s) | 44.93 | 202.29                | 406.29                |
|        | 误差指标    | 7.99  | $2.47 \times 10^{-2}$ | $4.25 \times 10^{-3}$ |
|        | 分类正确率   | 86%   | 94%                   | 99%                   |

由实验结果可以看出, 禁止搜索算法与遗传算法相比, 在训练时间接近时, 网络分类性能显著提高; 若达到同样误差目标, 采用遗传算法训练需 100 次迭代, 而采用禁止搜索仅需要 40-50 次迭代, 由此可见, 禁止搜索不仅使粗神经网络训练的速度大大提高, 且分类能力也显著增强, 对训练样本的数量要求及计算时间上都显著优于遗传算法. 该算法无论在仿真还是实际应用都具有很大价值.

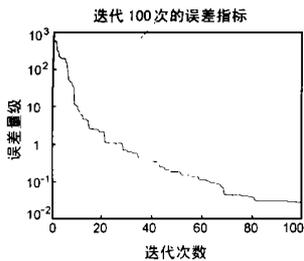


图 10 GA-RNN 算法误差收敛状态

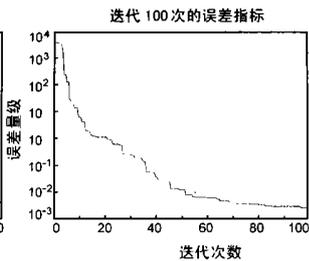


图 11 TS-RNN 算法误差收敛状态

## 6 结论

本文提出了粗神经网络的禁止搜索训练算法, 并给出了同遗传训练算法的比较. 仿真试验的结果表明禁止搜索训练算法的性能明显优于遗传训练算法. 该算法还可以用于声音、图像等信号的处理作进一步的评估.

## 参考文献:

- [ 1 ] LINGRAS P J. Comparison of neofuzzy and rough neural networks [ J ]. Information Sciences, 1998, 110(3- 4): 207- 215.
- [ 2 ] LINGRAS P J. Rough neural networks [ A ]. Proceedings of Sixth International Conference on Information Processing and Management of Uncertainty in Knowledge Based Systems [ C ], Granada, Spain, 1996: 1445- 1450.
- [ 3 ] 张兆礼, 孙圣和. 粗神经网络及其在数据融合中的应用 [ J ]. 控制与决策, 2001, 16(1): 76- 78.
- [ 4 ] GOLDBERGG D E. Genetic Algorithms in Search, Optimisation and Machine Learning [ M ]. Addison Wesley Publishing Company, 1989.
- [ 5 ] QI Xiao feng, Palmieri Francesco. Theoretical analysis of evolutionary algorithms with an infinite population size in continuous space Part I: Basic properties of selection and mutation [ J ]. IEEE Trans. on Neural Networks, 1994, 5( 1 ): 102- 109.
- [ 6 ] QI Xiao feng, Palmieri Francesco. Theoretical analysis of evolutionary algorithms with an infinite population size in continuous space part II: analysis of the diversification role of crossover [ J ]. IEEE Trans. on Neural Networks, 1994, 5( 1 ): 120- 128.
- [ 7 ] F Glover, M Laguna. Tabu Search [ M ]. Kluwer Academic Publishers, London, 1997.
- [ 8 ] PAWLAK Z. Rough Sets Theoretical Aspects of Reasoning About Data [ M ]. Kluwer Academic, Boston, 1991.
- [ 9 ] 曾黄麟. 粗集理论及其应用 [ M ]. 重庆: 重庆大学出版社, 1998.
- [ 10 ] 王珏. Rough Set 约减与数据浓缩 [ J ]. 高技术通讯, 1997, 7( 1 ): 40- 45.

## 作者简介:



梅晓丹 女, 1974 年 9 出生于黑龙江省伊春市. 哈尔滨工业大学自动化测试与控制系博士生, 研究方向为语音信号的识别与压缩.



孙圣和 男, 1937 年出生于山东省. 哈尔滨工业大学自动化测试与控制系教授、博导, 从事计算机测试与控制, 信号处理与系统辨识, 数据压缩的研究.