

多层前向神经网络的自适应禁忌搜索训练^{*}

贺 一^{1,3} 刘光远² 雷开友¹ 贺 三⁴ 邱玉辉¹

(西南师范大学计算机与信息科学学院 重庆400715)¹ (西南师范大学电子信息工程学院 重庆400715)²
(重庆师范大学现代信息管理系 重庆400047)³ (西南石油学院 成都610500)⁴

摘要 针对BP算法属于局部优化算法的不足,提出了一种新的全局优化算法——自适应禁忌搜索作为前向神经网络的训练算法。该算法通过邻域和候选集的相互配合,动态地调整候选集中分别用于集中性搜索与多样性搜索的元素个数,提高了算法运行的质量和效率。以经典的异或问题(XOR)为例,进行了对比研究。实验结果表明,该算法与BP算法相比明显提高了网络的收敛概率和收敛精度。

关键词 禁忌搜索,前向神经网络,全局优化,集中性与多样性

Applying An Adaptive Strategy-Based Tabu Search to Neural Network

HE Yi^{1,3} LIU Guang-Yuan² LEI Kai-You¹ HE San⁴ QIU Yu-Hui¹

(Faculty of Computer & Information Science, Southwest-China Normal University, Chongqing 400715)¹

(School of Electronic & Information Engineering, Southwest-China Normal University, Chongqing 400715)²

(Dept. of modern Information Management, Chongqing Normal University, Chongqing 400047)³

(Southwest-China Petroleum Institute, Chengdu 610500)⁴

Abstract Aiming at BP algorithm's drawbacks that it is essentially a local optimization algorithm, a novel and global optimization algorithm, Adaptive Tabu Search, is proposed to train feed-forward neural networks. This algorithm improves the quality and efficiency of training neural network by adjusting dynamically the numbers of intensification elements and diversification elements in candidate list and by the cooperating of neighborhood and candidate list. Taking the classical XOR problem as an example, a compare investigation is implemented. It shows that this algorithm has obviously superior convergence rate and precision compared to the BP algorithm.

Keywords Tabu search, Feed-forward neural network, Optimizing globally, Intensification and diversification

1 引言

人工神经网络已被广泛应用于模式识别、智能控制、系统优化、信号及信息处理等诸多领域。其中,多层前馈神经网络以其优良的非线性逼近性能,成为研究和应用最为广泛的一种神经网络,其训练算法一般采用BP算法,因此又叫BP网络。但BP算法采用误差导数指导学习过程,本质上属于局部优化算法,存在着易陷入局部最优、收敛速度慢、动态特性不够理想、学习速度慢等不足。针对这些缺点,国内外研究人员提出了许多改进的BP算法及其变种,详见文[1~6]。这些方法虽然在克服BP网络的不足方面取得了一定进步,但从本质上看,仍未脱离梯度法的局部寻优。随着遗传算法(GA)的出现,大量的研究人员开始把GA应用于神经网络的训练和学习,进行了比较深入的研究,并取得了很好的效果^[7~10],但GA存在着易早熟的缺点。近年来,随着一种新的全局优化算法——禁忌搜索(Tabu Search, TS)的出现^[11~12],人们开始尝试将TS应用于神经网络的训练,并取得了一些研究成果^[13~14],但研究得还不够。

TS算法是一种亚启发式(meta-heuristic)搜索技术,由

Glover在1986年首次提出^[11,15]。TS以其独有的记忆功能、爬山能力强、全局迭代寻优等一系列优点受到广泛关注,现已成功应用于工业调度、组合优化、机器学习、电路设计、函数优化、模式识别、神经网络等诸多领域^[16~22]。在构成TS算法的诸多要素中,集中性搜索(Intensification)与多样性搜索(Diversification)成为算法设计的焦点,也是一个难点。针对这一难点,我们提出了一种集中性与多样性的自适应搜索策略,并成功应用于求解中大规模TSP问题,详见文[16]。由于该策略的框架灵活,且无需问题特殊信息,因此,完全可推广应用于其它领域(如神经网络)。鉴于此,本文以应用最为广泛的前向神经网络作为优化对象,以求解经典的异或问题(XOR)为例,将基于该策略的自适应禁忌搜索作为神经网络的训练算法,进行了深入的研究。仿真实验表明,与BP算法相比,基于该策略的TS算法明显提高了神经网络的收敛概率和收敛精度,也证实了上述想法的可行性和有效性。

2 禁忌搜索应用于前向神经网络

(1)禁忌对象 本文选取神经网络的训练误差,即均方差mse (mean squared error)作为禁忌对象。由于mse是一个实

^{*} 本文受到教育部科学技术重点项目(No. 104262)和重庆市科委基金项目(2003-7881)共同资助。贺 一 副教授,博士研究生,研究方向为自然计算、并行计算及神经网络等。雷开友 副教授,博士研究生,研究方向为神经网络与计算智能。刘光远 教授,硕士生导师,主要研究方向为神经网络、模糊控制及现代优化方法。贺 三 博士研究生,目前研究方向为天然气管道优化运行的控制理论与技术。邱玉辉 教授,博士生导师,长期从事非单调推理、近似推理、神经网络、机器学习和分布式人工智能的教学和研究工作。

数,而在 TS 中直接禁忌一个实数是很难奏效的,因此,要对某一个 mse 实行禁忌,就以这个 mse 为基准点,左右偏移一个极小的范围,形成一个禁忌区间。这个偏移范围称为禁忌标准 TC(tabu criterion)。

(2)邻域结构 邻域定义为:神经网络的每一个权值和阈值在其上一步迭代中所取值的基础之上,加上一个修正量。这个修正量取某一区间内的一个随机实数,这个区间,我们称为邻域中元素值(包括权值和阈值)的偏移范围,简称权重偏移范围 WOR(weights offset region)。

(3)集中性(Intensification)与多样性(Diversification)的自适应搜索策略 集中性与多样性的搜索策略是 TS 算法设计中的焦点,本文采用文[16]中所提出的自适应搜索策略。在邻域中,构造集中性元素时,其权值偏移范围 WOR_1 的取值很小,本文取为 $\pm 0.1 \sim \pm 1.0$,以进行集中性搜索;构造多样性元素时,其权值偏移范围 WOR_2 的取值较大,本文取为 $\pm 1.0 \sim \pm 10.0$,以便进行多样性搜索。除此之外,自适应搜索策略的控制措施与文[16]所述完全一样(见图1)。由于该策略的自适应性,我们将基于该策略的禁忌搜索称为自适应禁忌搜索。

(4) 算法基本流程

Step1 设定算法参数。

Step2 用随机法产生初始权值和阈值,其取值为 $[-1, +1]$ 内的随机实数。

Step3 按上述规则和控制策略,生成邻域。

Step4 按上所述规则和控制策略,从邻域中选取元素组成候选集。

Step5 根据评价函数,从候选集中选出未被禁忌表禁忌的当前局部最优解。

Step6 更新初始解。

Step7 更新禁忌表。

Step8 根据上述规则和控制策略,调整候选集中的分界点长度 DL。

Step9 判断是否满足终止条件,若满足,则输出最优解,终止算法;若不满足,则转 Step3。

(算法中的终止条件为限定最大迭代步数)

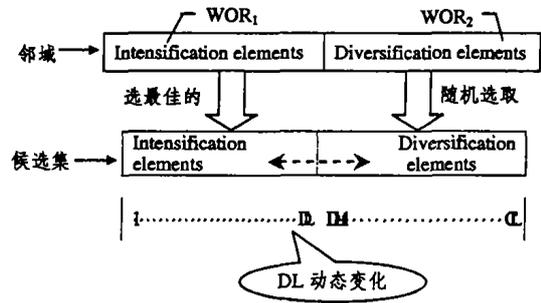


图1 邻域、候选集的构成

3 计算机仿真实验

(1) 几点说明

·测试算例:以非线性分类中的典型代表——异或问题(XOR)为例来检测本文算法的有效性。目标函数为:均方差 mse (mean squared error)。误差目标精度为: 1×10^{-6} 。

·神经网络结构:网络结构选为:2-2-1,前两层的传递函数为标准 S 型函数,输出层的传递函数为纯线性函数。

·算法参数:由于 TS 的性能与关键参数有关^[11,15],因此,本文对一些关键参数的取值进行了不同程度的组合,具体如下。

表1 算法参数组合

Group	TL	TC	NL
1	50	0.01%	100
2	100	0.01%	200
3	200	0.03%	500

其中:Group:组数;TL:禁忌长度(Tabu List),即禁忌表的大小;TC:禁忌标准(Tabu Criterion);NL:邻域长度(Neighborhood List),即邻域的大小,指邻域中的元素个数。

对于每组参数,最大迭代步数均取为2000,候选集长度均取为10。

·编程工具:Matlab6.5。

(2)仿真实验结果 对于每组参数情况,程序均运行100次,具体结果见表2,典型的收敛过程见图2。

表2 收敛概率、收敛精度、收敛速度统计

Group	收敛概率	收敛精度	算法终止时的 最小误差精度	收敛速度(Steps)		
				mean	max	min
1	94%	1×10^{-6}	1.2785×10^{-9}	502.99	801	115
2	95%	1×10^{-6}	3.0759×10^{-9}	499.60	969	174
3	91%	1×10^{-6}	8.3038×10^{-10}	329.33	637	91
总体平均	93.3%	1×10^{-6}		443.97		

4 讨论

(1)收敛过程分析 从图2可以看出,TS 在训练 FNN 过程中,能够根据解的情况自动进行集中性搜索或多样性搜索。当搜索未陷入局部极小时,则进行集中性搜索(见图中未加虚框的部分);当搜索入局部极小时,则进行多样性搜索(见图中的虚框部分)。在整个搜索过程中,mse 呈平稳下降趋势,其间虽有小的波动,但这不是振荡,也不是循环,恰恰是体现了 TS 算法所独具的优势,即 TS 可以接受劣解。

(2)与其它算法的比较 文[3]提出了一种超线性 BP 算

法,同样用于求解 XOR 问题,并与传统 BP 算法作了对比,现引用其结果,与本文算法的3组测试结果的平均值进行比较,详见表3。从表3不难看出,本文 TS 算法与 BP 算法相比较而言,具有更高的收敛精度和更高的收敛概率,且学习步数少。

表3 TS 算法与其它算法的学习结果比较

学习算法	最大学习步数	收敛精度	收敛概率	收敛学习步数
简单 BP 算法	8000	1×10^{-3}	89%	2206
最速下降算法	8000	1×10^{-3}	86%	1221
超线性 BP 算法	8000	1×10^{-3}	82%	548
本文 TS 算法	2000	1×10^{-6}	93.3%	43.97

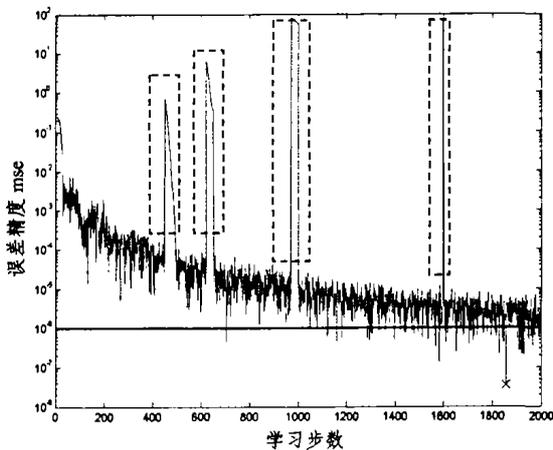


图2 典型收敛过程示意图

结论与展望 综上所述,我们可得出结论:

•将 TS 作为 ANN 的学习算法时,与其它算法相比较而言,它在本质上是一种全局寻优算法,因而具有更高的收敛精度、收敛概率。

•从表2可以看出,在3组不同参数的情况下,最终测试结果相差并不大。这主要是因为,集中性与多样性的自适应搜索策略改善了 TS 的不足,使得:禁忌长度和邻域长度对搜索结果的影响相对较小。

•TS 在训练 ANN 时,不需要计算传递函数的导数。也就是说,除了 sigmoid 函数外,我们可以选择其它非连续可微的函数作为传递函数,来满足一些特殊要求,这正是笔者下一步要进行的工作。

参考文献

- 1 刘光远,邱玉辉.基于稳健误差估计器的快速 BP 算法.计算机科学,1997,24(2):66~68
- 2 裴浩东,苏宏业,褚健.多层前向神经网络的权值平衡算法.电子学报,2002,30(1):139~141
- 3 梁久桢,何新贵,黄德双.前馈网络的一种超线性收敛 BP 学习算法.软件学报,2000,11(8):1094~1096

(上接第81页)

世界中的实体和过程,以及如何让一个主体在一定的时间内根据这些信息进行推理并作出决策。在系统中,我们采用慎思主体结构来构建软件 Agent。

(3)各种类型的 Agent 之间通信的实现。将软件 Agent 技术引入到基于 Web 的数据仓库体系结构中的各个部分,需要构建多种类型的 Agent,要求这些 Agent 相互配合,共同完成任务。那么,各种类型的 Agent 之间如何通信是实现我们的这个系统的关键问题之一。目前国际上著名的主体通信语言 (Agent Communication Language, 简称 ACL) 是美国 ARPA 的知识共享计划中提出的两个相关语言:一是 KQML^[12] (Knowledge Query and Manipulation Language), 另一个是 KIF (Knowledge Interchange Format)。KQML 定义了一种主体间传递信息的标准语法以及一些“动作表达式”,且这些动作从言语行为理论演化而来。KIF 则给信息的内容提供一种语法,它基本上是用类似 LISP 的语法书写一阶谓词演算。在系统中,我们采用 KQML 来表示各个 Agent 间的通信原语,即用 KQML 定义各个 Agent 的通信原语格式。

结束语 随着因特网的日益普及,Web 技术在各个领域得到了广泛的应用。论文深入分析了基于 Web 的数据仓库体系结构存在的不足,尤其是针对如何把有用的 Web 数据集成并入到数据仓库中这一目前的研究热点问题,提出了结合 Agent 技术,将 HTML 页面转化为 XML 数据源的解决方案,从而可实现将 Web 数据转换为数据仓库的数据源,充分发挥

- 4 高雪鹏,丛爽. BP 网络改进算法的性能对比研究[J]. 控制与决策, 2001,16(2):167~171
- 5 Engoziner S, Tomsen E. An accelerated learning algorithm for multilayer perceptron: Optimization layer by layer. IEEE Transactions on Neural Network, 1995,6(1):31~42
- 6 Scalero R S, Tepedelenlioglu N. A fast new algorithm for training feedforward neural networks. IEEE Transactions on Signal Processing, 1992,40(1):202~210
- 7 Friedrich S, Klaus P A. Clinical monitoring with fuzzy automata. Fuzzy Sets and Systems, 1994,61(1):37~42
- 8 Delgado M, Mantas C, Pegalajar M C. A genetic procedure to tune perceptrons. In: Proc. of Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU'96), 1996, 2: 963~969
- 9 Blanco A, Delgado M, Pegalajar M C. A genetic algorithm to obtain the optimal recurrent neural network. International Journal of Approximate Reasoning, 2000,23 (1): 67~83
- 10 Blanco A, Delgado M, Pegalajar M C. A real-coded genetic algorithm for training recurrent neural networks. Neural Networks, 2001,14(1):93~105
- 11 Glover F, Laguna M. Tabu Search. Boston: Kluwer Academic Publishers, 1997
- 12 Glover F, Hanafi S. Tabu search and finite convergence. Discrete Applied Mathematics, 2002,119 (1-2): 3~36
- 13 Sexton R S, Alidaee B, Dorsey R E, et al. Global optimization for artificial neural network: A tabu search application. European Journal Operational Research, 1998,106(2-3):570~584
- 14 于志伟. Tabu 机——一种新的全局优化神经网络. 电子学报, 1999,27(2):117~119
- 15 王凌. 智能优化算法及其应用. 北京:清华大学出版社,2001
- 16 贺一,刘光远,邱玉辉. Tabu Search 中集中性和多样性的自适应搜索策略. 计算机研究与发展, 2004,41(1):162~166
- 17 贺一,刘光远. 基于变异方法的禁忌搜索. 计算机科学, 2002,29(5):115~116
- 18 方永慧,刘光远,贺一,邱玉辉. 一种基于插入法的禁忌搜索. 西南师范大学学报(自然科学版), 2003,28(6):887~891
- 19 Liu Guangyuan, He Yi, Qiu Yuhui, Yu Juebang. Research on Influence of Solving Quality Based on Different Initializing Solution Algorithm in Tabu Search. In: Proc. of Intl. Conf. on Communications Circuits and Systems and West Sino Expositions, Chengdu, China, IEEE Press, 2002. 1141~1145
- 20 Liu Guangyuan, He Yi, Fang Yonghui, Yuhui Qiu. A Novel Adaptive Search Strategy of Intensification and diversification in Tabu Search. In: Proc. of IEEE Intl. Conf. on Neural Network and Signal Processing, Nanjing, IEEE Press, Dec. 2003. 428~431
- 21 Zhang Hongbin, Sun Guangyu. Feature selection using tabu search method. Pattern Recognition, 2002,35 (3):701~711
- 22 Ferland J A, Ichoua S, Lavoie A, Gagné E. Scheduling using tabu search with intensification and diversification. Computer & Operations Research, 2001,28(11):1075~1092

数据仓库技术和 Web 技术相结合的优势。文中还分别研究了在基于 Web 的数据仓库体系结构各个层次中软件 Agent 技术的应用,针对该结构目前存在的一些主要问题提供了有效的解决方案,为实现一个更灵活、可伸缩、高效的 Web 数据仓库系统提供了更为广阔的前景。

参考文献

- 1 吴宏旻,陈奇,俞瑞钊. 关于数据仓库若干问题的讨论. 计算机科学, 1999,26(2):39~43
- 2 毛国君. 数据仓库的质量管理问题和办法. 计算机科学, 2003,30(8):88~91
- 3 Inmon W H. Building the Data Warehouse. 2nd ed., New York: John Wiley & Sons, Inc., 1996
- 4 李秀,廖璋,刘文煌. 基于 Web 的数据仓库系统的研究. 计算机工程, 2001,27(11):44~46
- 5 许亮,李明,王惠琴. 基于 Web 的数据仓库体系研究. 甘肃工业大学学报, 2002,28(1):68~71
- 6 Michael R G, Steven P K. Software Agents [J]. Communications of the ACM, 1998,37(7):48259
- 7 Wooldridge M, Jennings N R. Intelligent agents: Theory and practice [J]. Knowledge Engineering Review, 1994,10(2):1152152
- 8 Pham V A, Karmonch A. Mobile Software Agents: An Overview. IEEE Common Magazine, 1998(7):26~37
- 9 蒋文伟,许华虎,唐毅. 基于 Agent 的数据仓库的研究. 计算机工程, 2001,27(3):29~32
- 10 徐忠健,袁捷,杨倩. 基于 Agent 的三层数据仓库系统体系结构的研究. 计算机工程, 2003,29(3):58~60
- 11 刘振岩,王万森. 基于 XML 的数据挖掘的研究. 计算机科学, 2003,30(5):42~43
- 12 Finin T, Wiederhodl. An overview of KQML: A Knowledge Query and Manipulation Language, Stanford University Computer Science Department, 1991