

一种基于退火策略的混沌神经网络优化算法*

王 凌 郑大钟
(清华大学自动化系·北京, 100084)

0224

摘要: Hopfield 网络(HNN)中引入混沌机制,首先在混沌动态下粗搜索,并利用退火策略控制混沌动态退出和逆分岔出现,进而 HNN 梯度优化搜索,提出了一种具有随机性和确定性并存的优化算法.对经典旅行商(TSP)的研究,表明算法具有很强的克服陷入局部极小能力,较大程度提高了优化、时间和对初值的鲁棒性能,同时给出了模型参数对性能影响的一些结论.

关键词: 退火策略;混沌神经网络;优化 TSP;

文献标识码: A

混沌算法

旅行商问题

A Kind of Chaotic Neural Network Optimization Algorithm Based on Annealing Strategy

WANG Ling and ZHENG Dazhong

(Department of Automation, Tsinghua University · Beijing, 100084, P. R. China)

Abstract: This paper presents a self-organization optimization algorithm, which combines stochastic with deterministic property to introduce chaos mechanism into Hopfield neural network(HNN) to coarsely search the optimum under chaotic dynamics and control the chaotic dynamics by annealing strategy to perform inverse bifurcation and disappear. After that, the gradient property of HNN is used to reach stable point. Simulation results about two typical TSP problems show that such an algorithm, which is robust with initial states, can avoid getting stuck in local minima and has better convergence property as well as time property. Moreover, some conclusions about the effect of parameters on the model are summed up.

Key words: annealing strategy; chaotic neural network; optimization; TSP

1 引言(Introduction)

传统优化算法或易陷入局部极小,或由于计算量和存储量巨大而难于应用.模拟退火算法(SA)^[1]作为随机搜索方法,其突跳特性可避免局部极小,但收敛缓慢且难于确定“最佳”参数.HNN^[2]是一种重要优化方法且易于硬件实现,但难以克服陷入局部极小,对旅行商问题(TSP)的求解表明其计算量大,鲁棒性差,优化性能和可靠性差^[3].由于混沌神经网络在远离平衡点处有丰富的时空动态,利用动态演变导致吸引子的转移可对优化问题构成性能优越的算法,把混沌和神经网络相融合可形成一类有学习能力的系统^[4].本文提出一种基于退火策略的混沌神经网络(ACNN)模型,HNN 中引入混沌机制形成混沌动态,利用其遍历性在解空间中“随机”粗搜索,再由退火策略控制混沌动态逐渐消失并产生逆分岔,一旦逆分岔过程结束便转入 HNN 极度优化.进

而,以经典 TSP 为例进行仿真研究,结果表明本文算法具有很强的克服陷入局部极小能力,较大程度提高了收敛性能和初值鲁棒性.

2 ACNN 模型及其优化机制(ACNN model and its optimization mechanism)

熟知,HNN 求解优化问题的连续模型^[2]和其基于 Euler 法的离散模型分别为:

$$\frac{du_{xi}(t)}{dt} = -u_{xi}(t)/\tau + \beta \cdot \left[\sum_j w_{xi,j} v_{yj}(t) + I_{xi} \right], \quad (1)$$

$$u_{xi}(k+1) = (1 - \Delta t/\tau) \cdot u_{xi}(k) + \Delta t \cdot \beta \cdot \left[\sum_j w_{xi,j} v_{yj}(k) + I_{xi} \right] = \alpha \cdot u_{xi}(k) + \beta \cdot \left[\sum_j w_{xi,j} v_{yj}(k) + I_{xi} \right], \quad (2)$$

进而将混沌机制引入反馈网络,并结合退火策略控

* 国家自然科学基金(69684001)和国家攀登计划资助项目.
收稿日期:1997-10-31;收修改稿日期:1998-11-16.

混沌动态,构造如下的具有自组织特性和克服陷入局部极小能力的基于退火策略的混沌神经网络(ACNN)模型:

$$\begin{cases} u_{xi}(k) = \alpha \cdot u_{xi}(k-1) + \\ \beta \cdot \left[\sum_j w_{xi,yj} v_{yj}(k) + I_{xi} - s_1 \right] - \\ z(k-1) \cdot [v_{xi}(k-1) - s_2], & (3) \\ v_{xi}(k) = 1/[1 + \exp(-\mu \cdot u_{xi}(k))], & (4) \\ z(k) = z(k-1) \cdot \\ \ln[e + \lambda \cdot (1 - z(k-1))], & (5) \end{cases}$$

$x, i, y, j = 1, 2, \dots, n; k = 1, 2, \dots, m.$

其中,式(3)为网络动态方程,式(4)为神经元激励函数,式(5)为退温函数。 $u_{xi}(k)$ 和 $v_{xi}(k)$ 为神经元 (x, i) 的输入输出, $w_{xi,yj}$ 为连接权, τ 为时间常数, $z < 1$ 为控制参数, $\alpha(0 \leq \alpha \leq 1)$ 为阻尼因子, μ 激励函数零输入时的斜率, I_{xi} 为外部偏置, s_1, s_2 为设定阈值, $\beta(\beta > 0)$ 和 λ 分别为尺度和退温参数。

ACNN模型综合了随机性和确定性算法的优点,优化过程分为基于混沌的“粗搜索”和基于HNN的“细(梯度)搜索”两个阶段。粗搜索中的混沌搜索同时具有随机性和轨道遍历性,具有克服陷入局部极小的能力。随机性保证大范围搜索能力,轨道遍历性使算法能按系统自身的行为不重复地遍历所有可能状态,有利于克服一般随机算法中以分布遍历性为机制搜索带来的局限性。由此,我们的算法机制上保证了同时具有克服陷入局部极小的能力和良好的搜索效率。具体两阶段描述如下:1)粗搜索阶段:选定控制参数使系统一开始就进入大范围意义下的较强混沌动态,利用混沌搜索的随机性和轨道遍历性在大范围内按其自身规律进行遍历性搜索。搜索过程按混沌轨道进行遍历,不受目标函数限制,从而有很强避免陷入局部极小的能力。伴随混沌搜索的进行,控制参数开始退温,退温到一定程度并由系统的自组织能力,混沌动态将退出而进入逆分岔阶段。此后,搜索过程将限制在一些周期解,且能随控制参数的进一步减小而缩小搜索区域,当逆分岔现象随控制参数的进一步退温而达到周期1形态时,算法得到一个易于得到全局意义下最优解的某一解,从而粗搜索过程结束,转入HNN梯度搜索。2)细搜索阶段:随着粗搜索结束,系统方程中控制参数决定项的作用很弱,优化过程基本按能量函数的梯度下降

方向进行。寻优过程将由粗搜索得到的解小范围内按HNN梯度下降机制进一步搜索,较快收敛到全局意义下较满意的解。

退温函数式(5)使得 $z(k+1)/z(k)$ 随 k 增大而加快来控制 $z(k)$ 的下降速率,如此设计目的是使混沌动态有足够长的进程以提高粗搜索性能,并利用 z 较小时退温速率的加快以加快梯度搜索来改善时间性能。仿真研究表明,这种退温策略能够产生较好的优化效果。

3 基于ACNN模型的仿真研究(Simulation study based on ACNN model)

以典型TSP为例,在Pentium586/133(16M RAM)计算机上,以Borland C++为工具进行仿真研究,能量函数 E 参照HNN^[5]构造,网络权值和偏置阈值可通过求导相应得到。

$$E = \frac{A}{2} \left\{ \sum_{x=1}^n \left(\sum_{i=1}^n v_{xi} - 1 \right)^2 + \sum_{i=1}^n \left(\sum_{x=1}^n v_{xi} - 1 \right)^2 \right\} + \frac{B}{2} \sum_{x=1}^n \sum_{j=1}^n \sum_{\substack{i=1 \\ i \neq x}}^n (v_{x,i+1} + v_{x,i-1}) v_{xi} d_{xi}. \quad (6)$$

1) 10城市TSP^[2]:Hopfield^[2]20次实验中只得16次合法解,10次满意解;Wilson等^[13]在100次实验中只有15次在1000步迭代内收敛到合法解,且仍为次优解;文献[6]在100次实验中只得到47次合法解,12次满意解。本文模型参数取为: $A = B = 1, s_1 = 1, s_2 = 0.65, \mu = 250$,神经元初始输入为 $(-1.0, 1.0)$ 的随机数,其它参数见表1。能量函数的变化连续10次小于0.01作为收敛条件。各参数下50次随机仿真的统计结果见表1,其中波动率定义为平均回路长度相对最优回路长度的偏差。由表可见本文算法:在第1,5,9组参数下每次实验均得到最优解;仅第3组参数下出现3次非法解(约占6%);次优解出现几率远好于文献结果;波动率性能和时间性能均很好。虽然不适当的参数导致出现个别非法解,但出现几率较HNN大大减少。非法解出现原因分析为:①混沌动态过分或不充分导致陷入性能差的不动点;②网络还未完全稳定,即收敛条件中阈值设置不够充分。

表 1 各参数下的 ACNN 对 10 城市 TSP 的仿真统计结果

Table 1 Simulation statistics result of 10 cities TSP under different parameters

序号	β	α	λ	$z(0)$	最优解 数目	次优解 数目	非法解 数目	波动率 %	平均迭代 步数
1	0.02	0.88	0.01	0.05	50	0	0	0	216
2	0.018	0.88	0.01	0.05	49	1	0	0.22	239
3	0.015	0.88	0.01	0.05	42	5	3	1.26	293
4	0.022	0.88	0.01	0.05	43	7	0	1.71	194
5	0.02	0.90	0.01	0.05	50	0	0	0	200
6	0.02	0.86	0.01	0.05	47	3	0	0.93	225
7	0.02	0.88	0.008	0.05	48	2	0	0.63	259
8	0.02	0.88	0.012	0.05	47	3	0	0.74	187
9	0.02	0.88	0.01	0.06	50	0	0	0	267
10	0.02	0.88	0.01	0.04	39	11	0	1.93	147

2) 30 城市 TSP^[7]: 本文算法仿真中将各坐标归一化到 $[0,1]$ 区间, 结果输出时再折合回, 模型参数选择为: $A = 1, B = 0.5, s_1, s_2, \mu$ 和神经元初始输入取值范围同前, $\alpha = 0.9, \beta = 0.015, \lambda = 0.007, z(0) = 0.08$, 能量函数变化连续 10 次小于 0.00001 作为收敛条件; 标准 SA^[1] 参数选择为: 初温 5, 终温 0.01, 退温速率 0.8, 抽样步数 1000, 两点互换为状态转移函数. 50 次随机仿真实验中, 本文算法得到 15 次最优解, 15 次相对误差小于 3% 的次优解, 无非法解出现, 波动率为 6.27%, 平均迭代步数为 653; SA 算法得到仅 3 次最优解, 10 次相对误差小于 3% 的次优解, 波动率为 10.47%, 平均迭代步数为 23000 步.

由此可见, 本文算法具有很强的克服陷入局部极小能力, 各方面性能均优于 HNN 和 SA 算法, 反映了算法中“确定性和随机性搜索机制相混合”的优点. 图 1, 2 给出了本文算法优化过程中神经元输出动态的演变过程和能量函数的变化过程.

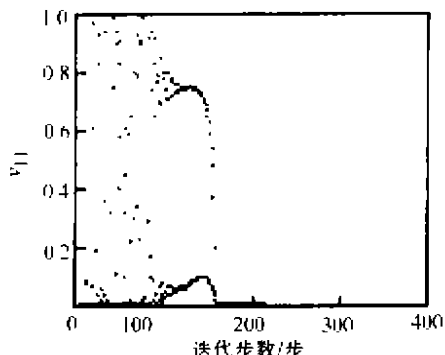


图 1 10 城市 TSP 在第 1 组参数下神经元输出 v_{11} 的演变过程

Fig. 1 Evolution of neuron output v_{11} in 10 cities TSP optimization under parameters group 1

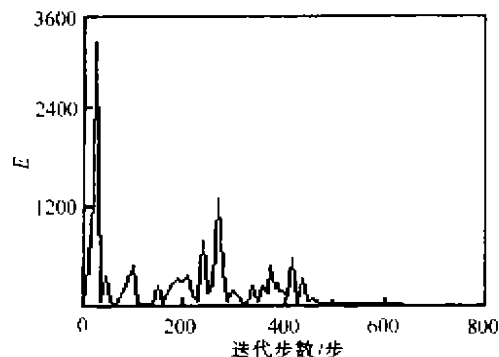


图 2 30 城市 TSP 优化中能量函数的变化曲线
Fig. 2 Varying curve of energy function in 30 cities TSP optimization

4 参数选择的几点结论 (Some conclusions for selecting parameters)

结论 1 尺度参数 β 的影响: β 减小导致迭代步数增大, β 过大或过小会导致次优解增多, β 过小也可能导致非法解出现. 分析网络动态方程可知, β 太小使得能量函数的变化不能充分影响系统动态的演变, 从而导致非法解或性能差的局部极小解; β 过大, 能量函数的影响过大, 不能产生充分的混沌动态, 影响搜索性能.

结论 2 阻尼因子 α 的影响: α 减小导致迭代步数增大和次优解数目的增多. 由式 (2) 直接看出, α 减小相当于增大仿真步长 Δt , 从而影响仿真精度和收敛性. 由于 β 未变, Δt 增大相当于 β' 减小, 式 (2) 中能量函数变化项因而变弱, 导致迭代步数增大. 相反, α 过大虽能使迭代步数减少, 但会出现类似 β 过大的情况.

结论 3 退温参数 λ 的影响: λ 的大小反映参数 z 下降速率的快慢, λ 增大导致迭代步数减小. 但是, λ 过大使得混沌动态消失过快, 从而收敛到性能差的局部极小解或非法状态; 而 λ 过小使得混沌动

态持续太久,严重影响搜索时间性能。

结论 4 控制参数初值 $z(0)$ 的影响: $z(0)$ 减小使迭代步数减小,但增多了次优解数目。分析系统动态方程和退温策略函数可知, $z(0)$ 减小使 $z(k+1)/z(k)$ 减小,从而 z 下降较快,混沌行为过小且不充分影响优化性能。相反, $z(0)$ 过大会使控制参数下降过慢,则混沌对系统动态演化过程的过大影响导致时间性能变差。

此外,对大规模 TSP 问题中参数 B 应选择得比 A 小些以均衡能量函数各项的作用。

参考文献 (References)

- [1] Kirkpatrick S, Gelatt C D and Vecchi C M. Optimization by simulated annealing [J]. Science, 1983, 220(4598): 671 - 680
- [2] Hopfield J J, Tank D W. "Neural" computation of decisions in optimization problems [J]. Biological Cybernetics, 1985, 52(8): 141 -

152

- [3] Wilson G V, Pawley G S. On the stability of the traveling salesman problem algorithm of Hopfield and Tank [J]. Biological Cybernetics, 1988, 58(1): 63 - 70
- [4] Aihara K, Takabe T and Toyoda M. Chaotic neural networks [J]. Physics Letters A, 1990, 144(6, 7): 333 - 340
- [5] Xu X, Tsai W T. Effective neural algorithms for the traveling salesman problem [J]. Neural Networks, 1991, 4(2): 193 - 205
- [6] 庄镇泉, 王煦法, 王东生. 神经网络与神经计算机 [M]. 北京: 科学出版社, 1992
- [7] Fogel D B. Applying evolutionary programming to selected traveling salesman problems [J]. Cybernetics and System, 1993, 24(1): 27 - 36

本文作者简介

王 凌 1972 年生 1995 年毕业于清华大学自动化系 现为该系博士生, 主要研究优化算法, 神经网络等。

郑大钟 1935 年生 现为清华大学自动化系教授, 博士生导师 研究领域包括线性系统理论, 最优控制, 大系统分散控制, DEDS, 混合动态系统等。

中国 2000 年机器人学大会 (CCR'2000)

——中国第六届机器人学术大会暨中国第四届智能机器人学术研讨会

第二次征文通知

主办单位: 中国自动化学会、中国机械工程学会、中国电子学会、中国汽车工程学会、中国人工智能学会、中国宇航学会、国家 863 计划智能机器人专家组、国家 863 计划空间机器人专家组、中国机器人工程协会

承办单位: 中国人工智能学会智能机器人学会

协办单位: 中南工业大学

时 间: 2000 年 10 月 22 日 ~ 25 日

地 点: 湖南长沙中南工业大学

一、征文范围:

· 机器人技术发展趋势及社会经济论题; · 中国加入 WTO 对机器人学的影响和对策; · 面向 21 世纪的智能机器人体系结构及系统技术; · 机器人新型机构及运动学、动力学; · 机器人控制技术及智能控制; · 人工智能在机器人中的应用; · 机器人规划与导航; · 机电一体化系统与机器人应用工程; · 先进制造技术与机器人; · 机器人视觉及非视觉传感技术; · 多媒体技术和多传感集成融合技术; · 神经网络、多智能体、虚拟现实等技术在机器人系统中的应用; · 人——机器人——机器交互技术及接口; · 机器人装配; · 机器人语言编程及仿真; · 工业机器人的新设计方法和新应用; · 遗传算法、进化计算及软计算在机器人中的应用; · 机器人及自动化; · 机器人学的教育与培训; · 智能控制(含递阶控制、专家控制、模糊控制、学习控制、神经控制等); · 智能自动化(含 CIMS、CIPS、CAD、CAM、MIS 等); · 智能系统; · 规划、调度、决策与管理系统; · 故障诊断与容错系统; · 现代制造系统、离散事件驱动系统和 Petri 网; · 信息处理系统; · 办公自动化、通讯自动化; · 电力系统自动化和自动装置; · 现代控制理论与应用; · 系统辨识、优化、建模与估计; · 模式识别与图像处理; · 计算智能(含遗传编程和进化计算学); · 人工生命理论及应用; · 系统鲁棒性与稳定性; · 计算机系统; · 计算机网络; · 计算机控制; · 自动化仪表

二、征文要求:

1. 未在其他会议或刊物上发表过的论文。
2. 反映机器人学、自动化、计算机理论和技术及应用研究成果。
3. 文稿必须正式打印, 录用论文将要求寄文稿软盘。
4. 每篇论文篇幅不超过 6000 字(含英文摘要和关键词), 并附 50 字以内的作者简介。

三、关键日期:

1. 收文截止: 2000 年 5 月 1 日。
2. 录用通知: 2000 年 6 月 10 日。
3. 清样截止: 2000 年 7 月 10 日。

四、来稿请寄:

410083 长沙市中南工业大学中国 2000 年机器人学大会秘书处 刘明

五、论文录用:

录用论文将在全国优秀科技期刊《中南工业大学学报》上以专辑正式出版。

中国人工智能学会智能机器人学会