

文章编号: 1001-0920(2006)01-0093-04

基于均匀设计的蚁群算法参数设定

黄永青¹, 梁昌勇¹, 张祥德²

(1. 合肥工业大学 计算机网络系统研究所, 合肥 230009; 2 东北大学 数学系, 沈阳 110006)

摘要: 蚁群算法的参数设置一直是依靠经验和试验来确定, 造成试验工作量大且难以得到最优的参数组合, 影响了算法的使用。通过将蚁群算法基本模型的参数设定问题描述成均匀设计中多因素多水平的试验设计, 从而能够用较少的试验很快设定算法参数的取值。TSP问题的仿真试验表明, 利用均匀设计得到的参数组合可使蚁群算法获得较优的运行性能, 说明了该方法的可行性和有效性。

关键词: 蚁群算法; 均匀设计; 旅行商问题

中图分类号: TP18 **文献标识码:** A

Parameter Establishment of an Ant System Based on Uniform Design

HUANG Yong-qing¹, LIANG Chang-yong¹, Zhang Xiang-de²

(1. Institute of Computer Network Systems, Hefei University of Technology, Hefei 230009, China; 2 Department of Mathematics, Northeastern University, Shenyang 110006, China Correspondent: HUANG Yong-qing, E-mail: yongqinghuangbs@sina.com)

Abstract: The enactment of the parameters of an ant system is determined by experience and experiment. This leads to heavy workload and makes the optimal combination of the parameters difficult to obtain. Uniform design method is used to convert the problem of parameter establishment into the experimental design of multi-factor and multi-level and reduces the workload of experiment greatly of simulation. The method applied to the traveling salesman problem shows good performance feasibility and effectiveness.

Key words: Ant system; Uniform design; Traveling salesman problem

1 引言

基本蚁群算法是意大利学者Dorigo首先提出的^[1], 并用于求解TSP, ATSP, QAP和JSP等组合优化问题。因蚁群算法具有正反馈、分布式计算等优点, 吸引了众多的研究者的注意, 其研究及应用领域越来越广泛和深入。如: 文献[2]对基本蚁群算法进行了拓展, 可用于连续参数优化; 文献[3]则将蚁群算法运用于数据挖掘。蚁群最优化(ACO)已成为相关领域的热门研究方向, 新模型、新方法、新应用层出不穷^[4]。

用蚁群算法求解问题, 都存在算法有关参数的

设定问题, 较好的参数组合将会使得算法具有全局搜索能力和较快的收敛速度; 相反, 不好的参数组合则会使得算法收敛速度过快或过慢, 并且容易陷入局部最优解。由于缺乏理论指导, 算法参数一般采用试探法来确定, 造成工作量大, 并且很难得到最优的参数组合, 从而影响了算法的使用性能。文献[1]对基本蚁群算法的4个参数作了对比试验, 即设定蚂蚁数目 m 等于城市的数目; 信息素残留系数 ρ {0.3, 0.5, 0.7, 0.9, 0.999}; 参数 α {0, 0.5, 1, 2, 5}和 β {0, 1, 2, 5}, 分别表示信息素水平和城市间的距离对转移概率的贡献程度。为了分析这些参数组合对算法性能的影响, 文献[1]固定前两个参

收稿日期: 2004-12-10; 修回日期: 2005-03-10

基金项目: 国家自然科学基金项目(70171033, 70471046); 安徽省基金项目(01041176, AHSK03-04D44)

作者简介: 黄永青(1974—), 男, 安徽东至人, 讲师, 博士生, 从事进化计算、决策分析等研究; 梁昌勇(1965—), 男, 安徽肥西人, 教授, 博士生导师, 从事人工智能、决策分析等研究

数, 而对 α 与 β 的各种组合进行分析, 总共需要作 20 组试验; 但是, 若考虑参数 ρ 的影响, 想做彻底的分析则需要作 100 组试验, 再加上蚂蚁数目 m , 则试验的组数变得难以承受。更重要的一点是, 这些参数实际上相互耦合, 相互联系, 具有复杂的关系, 仅取这些粗糙的值进行组合, 难以找到参数的最优组合。文献[5] 则通过仿真试验, 探讨了基本蚁群算法中有关参数的设定原则, 给出参数的有效取值范围, 但其参数的组合取值仍较粗糙, 难以得到最优的参数组合。随着蚁群最优化的应用越来越广泛, 迫切需要一种有效的理论指导, 能够以较少的试验确定出较优的参数取值, 均匀设计正是适应这种要求的方法。

本文对基本蚁群算法的参数设定进行了研究, 将其中 4 个主要参数的设定问题描述成 4 个因素多水平的均匀设计, 找出这些参数的最优组合。

2 基本蚁群算法

以旅行商问题(TSP) 为例说明基本蚁群算法的步骤。给定 n 个城市, 从城市 i 出发到城市 j 的转移概率定义为

$$p_{ij}^k(t) = \begin{cases} \frac{[\tau_{ij}(t)]^\alpha [\eta_j]^\beta}{\sum_{s \in \text{allow ed}_k} [\tau_{is}(t)]^\alpha [\eta_s]^\beta}, & j \in \text{allow ed}_k; \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

其中: $\tau_{ij}(t)$ 表示 t 时刻城市 i 与城市 j 之间的信息素水平; $\eta_j = 1/d_{ij}$, d_{ij} 是城市 i 与城市 j 之间的距离; allow ed_k 表示第 k ($k = 1, 2, \dots, m$) 只蚂蚁还未访问的城市集合; α 和 β 则分别表示信息素水平和城市间的距离对转移概率的贡献程度。

m 只蚂蚁经过 n 个时刻 (n 个城市) 完成算法的一次循环, 得到 m 个 TSP 问题的解, 然后需要对各个城市之间的信息素水平进行更新, 其计算公式为

$$\tau_{ij}(t+n) = \rho \cdot \tau_{ij}(t) + \sum_{k=1}^m \Delta \tau_{ij}^k \quad (2)$$

其中: $\rho \in [0, 1)$ 表示信息素在本次循环中的残留程度; $\Delta \tau_{ij}^k$ 表示第 k 只蚂蚁在本次循环中在城市 i 与城市 j 之间留下的信息素的量, 其计算公式常采用 (ant cycle system 模型)

$$\Delta \tau_{ij}^k = \begin{cases} Q/L_k, & \text{本次循环蚂蚁 } k \text{ 经城市 } i \text{ 到城市 } j; \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

其中: Q 为常数, L_k 表示蚂蚁 k 在本次循环中搜索到解。经过若干次循环后, 可采用适当的终止条件结束算法, 输出问题的解。

3 均匀设计

从算法的描述可以看出, 基本蚁群算法的参数

包括: 蚂蚁的数目 m , 参数 α 和 β , 信息素的残留系数 ρ 及常数 Q 。而前面 4 个参数的选择, 对算法的性能影响极大 (由于常数 Q 对以上算法的性能没有明显的影响^[5], 故不予考虑)。但由于这些参数相互耦合、联系紧密, 与算法的性能之间有复杂的关系, 只靠经验和试探性的试验很难得到最优的参数组合。

对上面的蚁群算法来说, 参数个数 $s = 4$, 若每个参数取 $n = 5$ 个水平 (即在参数的取值范围内取 5 个代表值), 现在想要找到这些参数的最优组合, 如果每个水平组合做一次试验, 则共需进行 $n^s = 625$ 次试验, 这是一个很大的数目。若参数的水平数取的更大, 则试验的组数变得很大而难以实现。一个自然的想法就是, 可否从这些试验组合中选取有代表性的水平组合进行试验, 以找到算法参数的最优组合? 这就是“部分因子试验法”。均匀设计是一种非常有效的部分因子试验方法, 是一种稳健试验设计。与正交设计相比, 均匀设计为试验者提供更多的选择, 可以用更少的试验, 即可获得所期望的结果; 利用均匀设计来设定算法的参数是非常成功的, 这方面的研究有待深入, 潜力很大^[6]。文献[7] 利用均匀设计对遗传算法的参数进行了成功设定, 取得了很好的效果。均匀设计是想让试验点足够的“均匀”, 使试验点具有更好的代表性, 从而可以通过较少的试验获取较多的信息。

均匀设计是一种简单易行的试验设计方法, 设定蚁群算法参数的步骤如下:

- 1) 对所研究的问题, 确定要设定的 s 个因素的取值范围;
- 2) 对每个参数各自取 n 个水平; 选择相应的等水平的均匀设计表 $U_n(n^s)$, 并按照表的水平组合编制均匀设计试验方案进行试验, 记录下算法的输出值。为获得有效的统计信息, 每组试验方案可以进行多次连续试验;
- 3) 从已做的试验点中挑一个指标最优的, 相应的参数组合即为所要求的最优组合。

以旅行商问题 bays29 为例, 结合文献[5] 建议的参数试验范围, 对基本蚁群算法考虑其中 4 个主要因素 (注意 $Q = 100$), 分别为: 蚂蚁数目 m ($m \in [8, 22]$); 信息素挥发度 ρ ($\rho \in [0.4, 0.932]$); 参数 α ($\alpha \in [0.8, 5]$); 参数 β ($\beta \in [0.8, 5]$)。均匀设计表可以自己构造 (如采用好格子点法^[6]), 也可以从文献[6] 的附录中得到。这里采用均匀设计表 $U_{15}(15^4)$, 并按照表的水平组合编制的试验方案如表 1 所示。表 1 中去掉括号中的数值后, 即为 4 因素 15 个水平的均匀设计表; 对每个参数进行 14 等分后, 得到 15 个从小到大排列的参数值, 以相应均匀

设计表中的数值为序号, 即得到括号中的参数值, 这些参数的水平组合即为均匀设计试验方案。不难看出, 按每个参数组合进行一次试验计算, 进行完全的试验需要作 15^4 次试验, 而采用均匀试验设计后, 只需要选择有很好的均匀性和代表性的 15 次试验即可得到满意的参数组合。由于试验点充分均匀分散, 试验点中最优因素组合值离在试验范围内通过全面试验寻求的最优因素组合值不会很远。这个方法看起来粗糙, 但通过大量的试验证明, 它是十分有效的^[6~8]。

表 1 均匀设计表 $U_{15}(15^4)$ 与均匀设计试验方案

组	m	ρ	α	β
1	1(8)	8(0.666)	6(2.3)	12(4.1)
2	2(9)	2(0.438)	12(4.1)	6(2.3)
3	3(10)	12(0.818)	3(1.4)	2(1.1)
4	4(11)	15(0.932)	9(3.2)	9(3.2)
5	5(12)	4(0.514)	8(2.9)	15(5.0)
6	6(13)	9(0.704)	15(5.0)	4(1.7)
7	7(14)	6(0.590)	1(0.8)	7(2.6)
8	8(15)	13(0.856)	13(4.4)	14(4.7)
9	9(16)	1(0.400)	4(1.7)	10(3.5)
10	10(17)	7(0.628)	10(3.5)	1(0.8)
11	11(18)	14(0.894)	5(2.0)	5(2.0)
12	12(19)	5(0.552)	14(4.7)	11(3.8)
13	13(20)	10(0.742)	2(1.1)	13(4.4)
14	14(21)	3(0.476)	7(2.6)	3(1.4)
15	15(22)	11(0.780)	11(3.8)	8(2.9)

4 仿真试验

为进行比较, 利用文献[1]中的参数组合设计了相应的试验方案, 即: 设定蚂蚁数目 m 等于城市的数目; 信息素残留系数 $\rho = 0.5$; 参数 $\alpha \in \{0.5, 1, 2, 5\}$ 和 $\beta \in \{1, 2, 5\}$ (α, β 取 0 的组合不予考虑), 共需要进行 12 组试验。对旅行商问题 bays29, 在 JBuilder 9.0 环境下, 每个试验方案运行 20 次, 每次运行 5 000 次循环, 发现只有两种参数组合能够发现最优解 (如表 2 所示), 其他参数组合则不能找到最优解 (未列出相应结果)。其中表 2 最后一列表示的是 20 次试验中发现的最好解及发现的次数。

表 2 bays29 (最优解为 2020) 的试验结果

($m = 29, \rho = 0.5$)

参数组合	解的平均值	标准差 (σ)	循环次数的平均值	最好解 (次数)
$\alpha = 2, \beta = 5$	2 026.90	1.104 3	2 064.8	2 020(1)
$\alpha = 5, \beta = 5$	2 025.10	0.984 6	1 856.5	2 020(2)

仍以旅行商问题 bays29 为例, 对蚂蚁数目 m , 信息素残留系数 ρ , 参数 α 和 β 四个参数, 利用表 1 编制的均匀设计试验方案, 在相同试验条件下得到的运行结果如表 3 所示。

表 3 bays29 (最优解为 2020) 的试验结果 (对应表 1 中的试验方案)

组	解的平均值	标准差 (σ)	循环次数的平均值	最好解 (次数)
1	2 027.25	1.573 0	3 428	2 020(6)
2	2 077.25	12.631 5	1 277	2 022(1)
3	2 153.85	8.961 9	2 858	2 067(1)
4	2 075.00	7.950 2	2 688	2 026(2)
5	2 024.80	1.232 1	2 764	2 020(7)
6	2 054.75	4.563 9	2 507	2 026(1)
7	2 067.60	5.286 5	2 791	2 036(1)
8	2 060.5	11.332 6	2 310	2 020(3)
9	2 032.15	1.712 1	2 755	2 020(1)
10	2 199.10	10.795 6	2 200	2 066(1)
11	2 056.15	2.894 2	2 159	2 034(2)
12	2 022.75	0.817 4	2 119	2 020(9)
13	2 032.50	1.018 6	1 697	2 020(1)
14	2 139.45	10.382 3	2 651	2 028(1)
15	2 029.10	1.877 4	2 336	2 020(1)

从表 2 可以看出, 12 组参数组合中只有 2 组能够使算法发现最优解; 并且在 20 次试验中, 仅有 1 到 2 次能够找到最优解, 说明在这些试验参数的选择下, 发现最优解的能力并不强, 比较容易陷入局部优化。这些参数选择虽然在文献[1]中对 Oliver30 的测试结果较好, 但对 bays29 效果却不理想, 说明参数的选择是依赖于问题的, 某个问题的最佳参数组合对另一个问题可能并非最佳。

从表 3 可以看出, 好的参数组合, 能够保证算法有较好的全局收敛性能, 同时具有较快的收敛速度, 并且算法的稳定性和可靠性也较高。相应的表 1 中第 1, 第 5, 第 12 组参数组合, 均取得了较好的效果, 特别是第 12 组参数组合发现最好解的能力很强, 而且综合性能最好, 相应的参数值作为基本蚁群算法的最优参数选择。从表 3 还可以看出, 表 1 中第 1, 第 5, 第 12 组参数组合之间有较明显的差异, 说明基本蚁群算法中蚂蚁数目 m , 信息素残留系数 ρ , 参数 α 和 β 四个参数之间是相互耦合的, 具有复杂的关系, 所以算法中的 4 个参数应当综合考虑。为了使得试验的组合数目较少而又能够获得算法运行的最佳参数组合, 采用均匀设计编制试验方案非常有效, 能够获得满意的结果。

对 bays29 问题, 以均匀设计表 $U_9(9^4)$ 进行 9 组试验, 发现也能找到满意的最优参数组合. 在实际操作中, 可以针对问题规模选择合适大小的均匀设计表. 对 TSP 问题的测试显示, 均匀设计方法是一种简单易行的有效方法.

5 结 语

基本蚁群算法参数设置的好坏, 直接影响算法的运行性能, 从而影响算法的实用性. 均匀试验设计方法的出现, 为科学有效设定蚁群算法参数提供了手段. 本文将蚁群算法基本模型的参数设定问题描述成均匀设计中多因素多水平的试验设计, 通过选择有代表性的参数组合进行试验, 从而能够用较少的试验次数找到满意的算法参数组合. 仿真试验表明了该方法的可行性和有效性, 为科学设定相关算法参数提供了参考.

参考文献(References)

- [1] Dorigo M, Maniezzo V, Colomi A. The Ant System: Optimization by a Colony of Cooperation Agents[J]. *IEEE Trans on Systems, Man, and Cybernetics-Part B*, 1996, 26(1): 29-41.
- [2] 陈峻, 沈洁, 秦玲. 蚁群算法求解连续空间优化问题的一种方法[J]. *软件学报*, 2002, 13(12): 2317-2323
(Chen L, Shen J, Qing L. A Method for Solving Optimization Problem in Continuous Space by Using Ant Colony Algorithm [J]. *J of Software*, 2002, 13(12): 2317-2323.)
- [3] Rafael S Parpinelli, Heitor S Lopes, Alex A Freitas. Data Mining with an Ant Colony Optimization Algorithm [J]. *IEEE Trans on Evolutionary Computing*, 2002, 6(4): 321-332.
- [4] 刘士新, 宋健海, 唐加福. 蚁群最优化—模型、算法及应用综述[J]. *系统工程学报*, 2004, 19(5): 496-502
(Liu S X, Song J H, Tang J F. Ant Colony Optimization Review: Modeling, Algorithms and Applications [J]. *J of Systems Engineering*, 2004, 19(5): 496-502.)
- [5] 詹士昌, 徐婕, 吴俊. 蚁群算法中有关算法参数的最优选择[J]. *科技通报*, 2003, 19(5): 381-386
(Zhan S C, Xu J, Wu J. The Optimal Selection on the Parameters of the Ant Colony Algorithm [J]. *Bulletin of Science and Technology*, 2003, 19(5): 381-386.)
- [6] 方开泰, 马长兴. *正交与均匀试验设计*[M]. 北京: 科学出版社, 2001.
(Fang K T, Ma C X. *Orthogonal and Uniform Experimental Design* [M]. Beijing: Science Press, 2001.)
- [7] 何大阔, 王福利, 张春梅. 基于均匀设计的遗传算法参数设定[J]. *东北大学学报*, 2003, 24(5): 419-411.
(He D K, Wang F L, Zhang C M. Establishment of Parameters of Genetic Algorithm Based on Uniform Design [J]. *J of Northeastern University*, 2003, 24(5): 419-411.)
- [8] 任露泉. *试验优化设计与分析*[M]. 北京: 高等教育出版社, 2003: 191-204.
(Ren L Q. *Optimal Design and Analysis of Experiments* [M]. Beijing: Higher Education Press, 2003: 191-204.)
- [10] 姚志湘, 钱宇, 李秀喜. 多变量化工过程系统的状态空间及其维数的判定准则[J]. *高校化学工程学报*, 2004, 18(1): 73-78
(Yao Z X, Qian Y, Li X X. The Criterion for Dimensionality Analysis of the State Space of Chemical Processes [J]. *J of Chemical Engineering of Chinese Universities*, 2004, 18(1): 73-78.)
- [11] Qin S Joe, Dunia R. Determining the Number of Principal Components for Best Reconstruction [J]. *J of Process Control*, 2000, 10(2-3): 245-250.
- [12] 李弼程, 罗建书. *小波分析及其应用*[M]. 北京: 电子工业出版社, 2003.
(Li B C, Luo J S. *Wavelet Analysis and Its Application* [M]. Beijing: Publishing House of Electronics Industry, 2003.)

(上接第92页)