

基于多样信息素的蚁群算法

高 尚^{1,2} 孙玲芳³ 侯志远¹ 杨静宇⁴

(江苏科技大学电子信息学院 镇江 212003)¹ (苏州大学江苏省计算机信息处理技术重点实验室)²

(江苏科技大学经济管理学院)³ (南京理工大学计算机系 南京 210094)⁴

摘 要 根据蚁群算法信息素更新的特性,提出了求解旅行商问题的多样信息素的蚁群算法。把蚁群的三种不同的信息素更新方式混合在一起,既利用了局部信息,又考虑了整体信息,将局部搜索和全局搜索相结合,使收敛性得到提高。针对旅行商问题的仿真实验结果,表明了该混合算法的有效性。

关键词 蚁群算法,信息素,旅行商问题

Ant Colony Optimization Algorithm Based on Multiplicate Pheromone

GAO Shang^{1,2} SUN Ling-Fang³ HOU Zhi-Yuan¹ YANG Jing-Yu⁴

(School of Electronics and Information, Jiangsu University of Science and Technology, Zhenjiang 212003)¹ (Provincial Key Laboratory of Computer Information Processing Technology, Suzhou 215006)² (School of Economic and Management, Jiangsu University of Science and

Technology, Zhenjiang 212003)³ (Department of Computer, Nanjing University Science and Technology, Nanjing 210094)⁴

Abstract By use of the properties of pheromone of ant colony algorithm, an ant colony algorithm based on multiplicate pheromone is proposed to solve the traveling salesman problems(TSP). Three modes of updating the pheromone are hybridized. The method uses not only local information but also global information and combines the local search with the global search to improve its convergence. The simulation results for TSP show the validity of this algorithm.

Keywords Ant colony algorithm, Pheromone, Traveling salesman problem

1 引言

上世纪 50 年代中期创立了仿生学。人们从生物进化的机理中受到启发,提出了许多用以解决复杂优化问题的新方法,如遗传算法、进化规划、进化策略等。目前研究的热点是模拟生物群落生活的特点而提出的群智能算法。作为一种新兴演化计算技术,它与人工生命,特别是进化策略和遗传算法有着极为特殊的联系,是一种能够有效解决大多数全局优化问题的新方法,其潜在并行性和分布式特点为处理大量的以数据库形式存在的数据提供了技术保证。蚁群算法是对蚁群落食物采集过程的模拟,由意大利学者 Colormi A 和 M. Dorigo 等人于 1992 年首先提出来^[1]。目前国内外研究者用蚁群算法研究了旅行商问题、指派问题、调度问题等,取得了一系列较好的实验结果^[2,3]。蚁群算法在电信路由优化^[4]、数据聚类分析^[5]、数据分类规则提取^[6]等方面效果也很好,已显示出蚁群算法在求解复杂优化问题(特别是离散优化问题)方面的一些优越性,证明它是一种很有发展前景的方法。

2 基本蚁群算法

蚂蚁个体之间通过一种称之为外激素的物质进行信息传递,从而能相互协作,完成复杂的任务。蚂蚁在运动过程中,能够在它所经过的路径上留下该种物质,而且蚂蚁在运动过程中能够感知这种物质的存在及其强度,并以此指导自己的运动方向。蚂蚁倾向于朝着该物质强度高的方向移动。因此,由大量蚂蚁组成的蚁群的集体行为便表现出一种信息正反馈现象:某一路径上走过的蚂蚁越多,则后来者选择该路径

的概率就越大。蚂蚁个体之间就是通过这种信息的交流达到搜索食物的目的。蚁群算法首先成功应用于旅行商问题,下面简单介绍其基本算法。

设有 m 个蚂蚁,每个简单蚂蚁有以下特征:它根据以城市距离和连接边上外激素的数量为变量的概率函数选择下一个城(设 $\tau_{ij}(t)$ 为 t 时刻边 $e(i, j)$ 上外激素的强度)。规定蚂蚁走合法路线,除非周游完成,不允许转到已访问城市,有禁忌表控制(设 $tabu_k$ 表示第 k 个蚂蚁的禁忌表, $tabu_k(s)$ 表示禁忌表中第 s 个元素)。蚂蚁完成周游后,在它每一条访问的边上留下外激素。

初始时刻,各条路径上的信息量相等,设 $\tau_{ij}(0) = C$ (C 为常数)。蚂蚁 $k(k=1, 2, \dots, m)$ 在运动过程中,根据各条路径上信息量决定转移方向, $p_{ij}^k(t)$ 表示在 t 时刻蚂蚁 k 由位置 i 转移到位置 j 的概率:

$$p_{ij}^k = \begin{cases} \frac{\tau_{ij}^{\alpha}(t) \cdot \eta_{ij}^{\beta}(t)}{\sum_{s \in allowed_k} \tau_{is}^{\alpha}(t) \cdot \eta_{is}^{\beta}(t)} & \text{if } j \in allowed_k \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

其中, $allowed_k = \{0, 1, \dots, n-1\} - tabu_k$ 表示蚂蚁 k 下一步允许选择的的城市。与实际蚁群不同,人工蚁群系统具有记忆功能, $tabu_k(k=1, 2, \dots, m)$ 用以记录蚂蚁 k 当前所走过的城市,集合 $tabu_k$ 随着进化过程做动态调整。 η_{ij} 表示边弧 (i, j) 的能见度,用某种启发式算法算出,一般取 $\eta_{ij} = 1/d_{ij}$, d_{ij} 表示城市 i 与城市 j 之间的距离。 α 表示轨迹的相对重要性, β 表示能见度的相对重要性, ρ 表示轨迹的持久性。 $1 - \rho$ 理解为轨迹衰减度随着时间的推移,以前留下的信息逐渐消失,用参

高 尚 副教授,主要从事系统理论等方面的研究。

数 $1-\rho$ 表示信息消逝程度,经过 n 个时刻,蚂蚁完成一次循环,各路径上信息量要根据下式做调整:

$$\tau_{ij}(t+n) = \rho \cdot \tau_{ij}(t) + \Delta\tau_{ij} \quad (2)$$

$$\Delta\tau_{ij} = \sum_{k=1}^m \Delta\tau_{ij}^k \quad (3)$$

$\Delta\tau_{ij}^k$ 表示第 k 只蚂蚁在本次循环中留在路径 ij 上的信息量, $\Delta\tau_{ij}$ 表示本次循环中路径 ij 上的信息量增量, L_k 表示第 k 只蚂蚁环游一周的路径长度, Q 是常数。

$$\Delta\tau_{ij}^k = \begin{cases} \frac{Q}{L_k} & \text{若第 } k \text{ 只蚂蚁在本次循环经过 } ij \\ 0 & \text{否则} \end{cases} \quad (4)$$

$\tau_{ij}(t), \Delta\tau_{ij}(t), p_{ij}^k(t)$ 的表达形式可以不同,要根据具体问题而定。M. Dorigo 曾给出 3 种不同模型^[7], 分别称为 ant-cycle system, ant-quantity system, ant-density system, 它们的差别在于表达式(4)的不同。(4)式称为 ant-cycle system。在 ant-quantity system 模型中,

$$\Delta\tau_{ij}^k = \begin{cases} \frac{Q}{d_{ij}} & \text{若第 } k \text{ 只蚂蚁在本次循环经过 } ij \\ 0 & \text{否则} \end{cases} \quad (5)$$

在 ant-density system 模型中,

$$\Delta\tau_{ij}^k = \begin{cases} Q & \text{若第 } k \text{ 只蚂蚁在本次循环经过 } ij \\ 0 & \text{否则} \end{cases} \quad (6)$$

它们的区别在于:后两种模型中,利用的是局部信息,而前者利用的是整体信息,在求解 TSP 问题时,性能较好,因而通常采用它,称为基本模型。解旅行商问题的蚁群算法的基本步骤是:

步骤 1 $nc \leftarrow 0$; (nc 为迭代步数或搜索次数)各 τ_{ij} 和 $\Delta\tau_{ij}$ 的初始化;将 m 个蚂蚁置于 n 个顶点上;

步骤 2 将各蚂蚁的初始出发点置于当前解集中;对每个蚂蚁 $k(k=1, 2, \dots, m)$, 按概率 p_{ij}^k 移至下一顶点 j ; 将顶点 j 置于当前解集;

步骤 3 计算各蚂蚁的路径长度 $L_k(k=1, 2, \dots, m)$; 记录当前的最好解;

步骤 4 按更新方程修改轨迹强度;

步骤 5 对各边弧 (i, j) , 置 $\Delta\tau_{ij} \leftarrow 0, nc \leftarrow nc + 1$;

步骤 6 若 $nc <$ 预定的迭代次数且无退化行为(即找到的都是相同解), 则转步骤 2;

步骤 7 输出目前最好解。

3 多样信息素更新规则

多数文献采用 ant-cycle system 和 ant-quantity system 这两个系统。文献^[7]还对这 3 个模型做了比较,认为 ant-cycle system 模型较好。实际上这 3 个模型在某种意义上有其合理的一面。真正蚁群社会中的蚁群是有组织的、有分工的,不同的蚁群有不同的信息素调控机制。采用单样信息素更新规则,并不全面。这里提出把蚁群分成 3 种,分别采用这 3 种模型,混合更新信息素。这种混合更新规则既利用了局部信息,又考虑了整体信息。将局部搜索和全局搜索相结合,可以使收敛性得到提高。为了避免搜索停滞,路径上的信息素浓度被限制在 $[\tau_{\min}, \tau_{\max}]$ 范围内。另外,信息素的初始值被设为其取值上限,这样有助于增加算法初始阶段的搜索能力^[8]。该方法称为 MAX-MIN Ant System。这里也采用 MAX-MIN Ant System 技术。

多数文献^[7]对 α, β, ρ 选取做了研究,对初始信息数 C 和

Q 的选取未考虑。实际上, C 与 $\Delta\tau_{ij}$ 是有联系的, $\Delta\tau_{ij}$ 不能比 C 小很多,否则影响力很小; $\Delta\tau_{ij}$ 比 C 大很多也不好,否则前几只蚂蚁的影响力过大,容易陷入局部最优。因此, $\Delta\tau_{ij}$ 要取合适一点。经过实验发现 $\Delta\tau_{ij}$ 取 $\frac{1}{5}C$ 左右时,效果较好。因此公式(4)、(5)和(6)中的 Q 是不一样的,要根据实际问题进行估算。

4 算法测试

选用 Oliver30(最好解为 423.7406)和 att48(TSPLIB 提供的最好解为 33522)作为实验例子,来研究 3 种蚁群的蚂蚁数的比例 $m_1 : m_2 : m_3$ 的选取。算法参数如下: $\alpha=1, \beta=5, m=60, C=10, \rho=0.99, Q$ 的设置如表 1 所示;并且与模拟退火算法、遗传算法进行比较。模拟退火算法采用文[9]的算法,起始温度 $T=100000$,终止温度 $T_0=1$,退火速度 $\alpha=0.99$ 。遗传算法程序采用 MATLAB 的遗传算法工具箱^[10],参数如下:染色体个数 $N=30$,交叉概率 $P_c=0.2$,变异概率 $P_m=0.5$,迭代次数 100。对各种算法测试 20 次,结果如表 2 所示。从表 2 可以看出,采用多样信息素更新规则的混合效率较高,特别蚂蚁数的比例 $m_2 : m_1 : m_3$ 取 $2 : 1 : 1$ 的效果更好。图 1 是解 Oliver30 混合算法最好的解,总路程为 423.7406。图 2 是解 att48 混合算法最好的解,总路程为 33522。

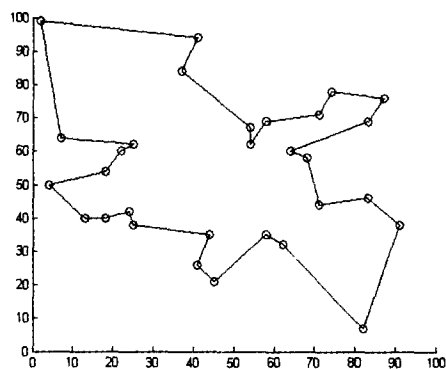


图 1 用混合算法解 Oliver 的最好的解

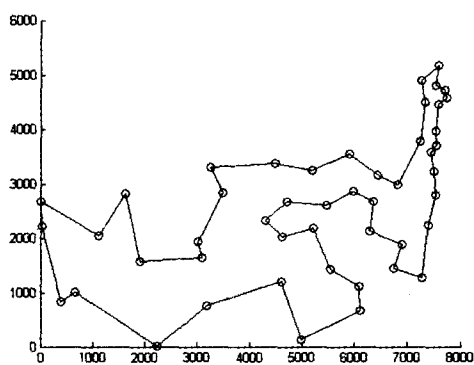


图 2 用混合算法解 att48 的最好的解

表 1 Q 的取值

Q	Oliver30	att48
公式(4)	1000	100000
公式(5)	100	10000
公式(6)	2	2

表2 几种算法测试结果

算法	Oliver30			att48			
	最好解	平均值	最差解	最好解	平均值	最差解	
模拟退火算法	438.5223	424.6918	479.8312	34958	35176	40536	
基本遗传算法	483.4572	467.6844	502.5742	38541	38732	42458	
ant-density system	425.6490	429.1071	433.1215	33786	35688	36559	
ant-quantity system	424.6727	429.7738	435.2420	33902	35714	36111	
ant-cycle system	423.7406	429.7032	432.4568	33780	35595	36534	
$m_1 : m_2 : m_3$	1:1:1	423.7406	429.2475	432.4168	33896	35792	36357
	2:1:1	423.7406	427.9214	431.4568	33522	35561	36100
	3:2:1	423.7406	429.3649	432.7944	33657	35599	36297

结束语 利用蚁群算法信息更新特性,提出的多样信息素的蚁群算法可以显著提高计算效率,具有较大的实用价值。尽管国内外研究蚁群算法的比较,但还有许多问题值得研究,如算法的参数选择只能通过仿真实验,无法给出理论指导。但从当前的应用效果来看,这种模仿自然生物的新型系统寻优思想无疑具有十分光明的前景,更多深入细致的工作还有待于进一步展开。

参考文献

- Colomi A, Dorigo M, Maniezzo V. An investigation of some properties of an ant algorithm [A]. In: Proc. of the Parallel Problem Solving from Nature Conference (PPSN'92)[C]. Brussels, Belgium; Elsevier Publishing, 1992. 509~520
- 吴庆洪,张纪会,徐心和. 具有变异特征的蚁群算法[J]. 计算机研究与发展, 1999, 36(10): 1240~1245
- 马良,项培军. 蚂蚁算法在组合优化中的应用[J]. 管理科学学报, 2001, 4(2): 32~33
- Gunes M, Sorges U, Bouazizi I. ARA the ant colony based routing

algorithm for MANETs [A]. In: Proceedings International Conference on Parallel Processing Workshops [C]. Unconver, B C, Canada, 2002. 79~85

- Lumer E, Faieta B. Diversity and adaptation in populations of clustering ants [A]. In: Proc. of the 3 Conf on Simulation of Adaptive Behavior [C]. MIT Press, 1994. 499~508
- Parpinelli R S, Lopes H S, Freitas. Data mining with an Ant Colony optimization algorithm [J]. IEEE Transactions on Evolutionary Computation, 2002, 6(4): 321~332
- Dorigo M, Maniezzo V, Colomi A. Ant system: optimization by a colony of cooperating agents [J]. IEEE Trans on System, Man and Cybernetics, 1996, 26(1): 28~41
- Stutzle T, Hoos H. The MAX-MIN ant system and local search for the traveling salesman problem [A]. In: Proceedings of the IEEE International Conference on Evolutionary Computation (ICEC'97)[C]. Indianapolis, USA, 1997. 309~314
- 康立山,谢云,尤矢勇,等. 模拟退火算法[M]. 北京: 科学出版社, 1994. 130~151
- 高尚. 基于 MATLAB 遗传算法优化工具箱的优化计算[J]. 微型电脑应用, 2002, 18(8): 52~54

(上接第 130 页)

BD_1, BD_2, \dots, BD_m 分别为维表 D_1, D_2, \dots, D_m 的磁盘块数, U 为结果 T 的磁盘块数。因为事实表的记录数远远大于维表,所以 $BF \gg BD_1, BF \gg BD_2, \dots, BF \gg BD_m$, 总的磁盘存取块数近似等于 $BF+U$ 。

对于传统的多表连接处理方法,如果两两表连接的次序为 $F \times D_1 \times D_2 \times \dots \times D_m$, 则因为 $BF \times D_1 \times D_2 \times \dots \times D_i > BF$, 最终连接运算所要求的磁盘块存取数将大于 $(2m-1)BF+U$, 其中读操作为 m 次, 写操作为 $m-1$ 次。显然, 当 $m > 1$ 时, 多表排序连接算法要比普通的多连接算法磁盘存取块数少很多。

总结 在当今信息时代,企业往往被淹没在来源于多个渠道的庞大、丰富的海量数据中。只有及时地将数据有机地组合在一起,及时地将信息转化为知识和智能,才能更好地指导企业进行商业决策和行动。商务智能的作用就在于此,它帮助管理者做出科学的计划、判断、决策,避免主观、片面等因素引起的重大失误。本文主要从技术的角度对商务智能的基本概念、技术架构和涉及到的技术进行了探讨,并对具体实现进行了深入研究。大连商品交易所计算机系统监控项目已经投入到应用中,对整个期货交易系统起到保驾护航的作用,有效地增强了企业的敏捷度和竞争力。通过集成化的商务智能系统的构建,企业高效地利用现有的信息基础设施中的数据,

大大提高了用户对数据查询分析和决策分析的能力。商务智能是一个跨学科的方向,还存在很多没有解决的问题,比如在数据仓库中的可视化的数据挖掘系统等问题,需要进一步进行研究。

参考文献

- Chaudhuri S, Dayal U. An overview of data warehousing and OLAP technology. ACM Sigmod Record, 1997, 26(1): 65~74
- Inmon W H. Building the Data Warehouse [M]. John Wiley & Sons Inc, 1996
- O'Neil P, Quass D. Improved query performance with variant indexes. ACM Sigmod Record, 1997, 26(2): 38~49
- Wang L, Li Y J, Wijesekera D, et al. Precisely answering multidimensional range queries without privacy breaches. Proceedings of the Eighth European Symposium on Research in Computer Security (ESORICS'03), 2003
- Gao Aiqiang, Li Qingzhong. The Data Manipulation and Querying Optimization. In SCDDWS, Computer Engineering and Application, 2002
- 王珊. 数据仓库技术与联机分析处理. 北京: 科学出版社, 1999
- 陈晓云,郭朝珍. 数据析取分类研究与设计. 计算机应用, 2001, 21(8): 1~2
- 王裕明,吴忠. 商务智能中元数据管理模型研究. 计算机应用与软件, 2005(8)
- 王卫平,徐宏发. 基于 Web Services 的商务智能的网络研究. 计算机系统应用, 2005(7)