第 21 卷 第 7 期　　控 制 与 决 策　　2006 年 7 月
Vol. 21 No. 7　　Control and Decision　　July 2006

文章编号: 1001-0920(2006)07-0763-04

蚁群优化算法的收敛性分析

朱 庆 保
(南京师范大学 计算机科学系, 南京 210097)

摘　要: 有关蚁群优化算法收敛性分析的研究还很少, 不利于进一步改进其算法. 为此, 较详细地分析了用蚁群优化算法求解 TSP 问题的收敛性, 证明了当 $0< q_0< 1$ 时, 算法能够收敛到最优解. 分析了封闭路径性质、启发函数、信息素和 q_0 对收敛性的影响, 据此给出了提高算法收敛速度的几点结论.

关键词: 蚁群优化算法; 收敛性分析; 启发函数; TSP 问题
中图分类号: TP301. 6　　文献标识码: A

收敛性

Analysis of Convergence of Ant Colony Optimization Algorithms

ZHU Qing-bao
(Department of Computer Science, Nanjing Normal University, Nanjing 210097, China. E-mail: zhuqingbao@njnu.edu.cn)

Abstract: The convergence problems of ant colony optimization algorithm is studied. To improve the algorithm, the convergence of this algorithm applied to TSP is analyzed in detail. The algorithm will be certain to converge to the optimal solution under the condition $0< q_0< 1$. In addition, the influence on its convergence caused by the properties of the closed path, heuristic functions, the pheromone and q_0 is analyzed. Based on it, some conclusions about the improvement of the speed of convergence are obtained.

Key words: Ant colony optimization algorithm; Convergence analysis; Heuristic function; Traveling salesman problem

1 引　言

目前, 蚁群优化算法[1]已成为研究的热点之一, 并得到广泛的应用. 蚁群算法是一种新的仿生算法, 其理论研究较少, 大多数研究只给出了算法或应用方法, 没有给出收敛性分析. 直到近几年, 有些学者才给出了某些算法的收敛性证明, 但有关研究和文献却非常少. 理论上的严重匮乏成为制约该算法改进和发展的瓶颈.

迄今检索到的最主要的收敛性证明有文献[2~5]. 这些文献对不同算法进行收敛性分析, 但都不同程度地存在一定的局限性. 文献[5]构建一种符合随机过程的分支蚁群算法, 根据蚂蚁数量、路径数量及其出生率、死亡率等, 用随机过程的思想方法证明了算法的收敛性, 其理论和结果显然仅适合于这类分支算法. 文献[3, 4]给出了基于图的蚁群算法的收敛性证明, 其理论基础是把最优路径上的信息素强度描述为离散时间的非齐次马尔可夫随机过程, 据此进行证明. 其局限性在于: 为了符合这一描述, 对算法作了较多的条件约束, 例如算法中的优化路径仅有一条, 仅在这条最优路径上进行全局信息素更新, 不支持局部更新, 所有蚂蚁都从唯一的起点出发. 文献[2]仅依据信息素进行分析, 分析条件与实际情况也有一些差距. 这些文献的共同特点是主要考虑了信息素的影响, 且只考虑了局部或全局信息素更新中的一种情况, 对于算法的相关参数(如启发信息和 q_0 等)对算法的影响没有进行分析, 实际指导意义不强.

本文以求解 TSP 问题为例, 充分考虑启发函

收稿日期: 2005-06-01; 修回日期: 2005-08-08.
基金项目: 江苏省自然科学基金项目(01KJB520007).
作者简介: 朱庆保(1955—), 男, 山东沂源人, 教授, 从事人工智能、智能控制等研究.

数、信息素、q_0 及路径的性质等对算法收敛的影响，对求解 TSP 问题的蚁群优化算法进行全新的分析，并据此提出了算法改进的建议.

2 问题描述与定义

为了叙述简便，首先给出如下定义:

二维平面上的凸多边形有限区域记为 AS，其内部分布着 n 个城市. 令 n 个城市的集合为 $C = \{c_1, c_2, \cdots, c_n\}$，城市的下标集或序号集为 $R = \{1, 2, \cdots, n\}$. 在 AS 中建立直角坐标系 Σ_0，则 $c_i \in C, i \in R$. Σ_0 有确定的坐标 (x_i, y_i)，记作 $c_i(x_i, y_i)$. 城市 $i \in R$ 称为节点，任意两城市间的连线称为边，记作 e_{ij}，$i, j \in R$.

定义 1 任意两城市 c_i 和 c_j 之间的距离或边长记为 $d(c_i, c_j)$，$i, j \in R$，简记作 d_{ij}，且满足 $d_{ij} = d_{ji}$. 距离 d_{ij} 可记作边长 d_l，d_{ij} 可计算如下:

$$d(c_i, c_j) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}. \quad (1)$$

定义 2 m 只蚂蚁的集合表示为 $\mathrm{ant} = \{1, 2, \cdots k, \cdots, m\}$，某只蚂蚁表示为 $k \in \mathrm{ant}$，蚂蚁于 t 时刻在 $e_{ij}(i, j \in R)$ 上残留的信息量表示为 $\tau_{ij}(t)$.

定义 3 蚂蚁 k 于任意时刻在 AS 中所处的位置为 P，所有 P 在 Σ_0 都有确定的坐标 (x, y). k 在 t_i 时刻处于某节点的位置记为 $P(x_i(t_i), y_i(t_i))$，简记作 P_i 或 $P(t_i)$. 若它的坐标与 $c_j(x_j, y_j) \in$ AS 的坐标相同，则 P_i 与 c_j 等价，记作 $P_i \sim c_j$. k 已走位置的集合表示为 tabu_k，它随着蚂蚁的行走动态调整. 这些已走过的位置不允许再走，因此 tabu_k 称为禁忌表. 可行点集记为 $V = C - \mathrm{tabu}_k$.

定义 4 记 $T = \{t_0, t_1, \cdots, t_i, \cdots, t_{n+1}\}$，$t_0 < t_1 < \cdots < t_i < \cdots < t_{n+1}$，$t_{n+1}$ 为有限时刻. $\forall P \in$ AS，$\forall P(t_i) \in$ FS，$i \in R$. 连续映射 $f: T \to$ AS 使得 $f(t_0) = P_0, f(t_1) = P_1, \cdots, f(t_i) = P_i, \cdots, f(t_n) = P_0$，$i, j \in R$. 映射 f 称为 AS 中从 P_0 到 P_0 的一条封闭巡回路径，像集 $f(T)$ 称为从 P_0 到 P_e 的一条封闭通道. 显然，该封闭通道构成了 AS 中连接点 P_0 到 P_e 的连续曲线，称为路程. 路程长度记作 $L(P_0, P_e)$，简记为 L. 其中 d_l 由式(1) 计算，L 计算如下:

$$L = \sum_{l=1}^{e+1} d_l, d_l = d(c_i, c_j), c_i, c_j \in C, i, j \in R. \quad (2)$$

定义 5 c_i 的相邻节点集记为 $\mathrm{BR}_i(c_i(x_i, y_i)) = \{c \mid c \in C, d(c, c_i) \leqslant d_{\min}\}$，其中 $d_{\min}$ 是根据具体问题和所需相邻集大小设定的距离阈值. 用 z 表示相邻域内的可行点集，则有 $z = (C - \mathrm{tabu}_k) \cap \mathrm{BR}_i$.

定义 6 $\eta_j = 1/d(c_i, c_j)$，称为蚂蚁从节点 i 选择节点 j 的启发函数.

3 蚁群优化算法

为了证明方便起见，根据第 2 节所作的问题描述和定义，结合 ACS 算法，对蚁群算法重新描述如下:

Step 1: 初始化: 将 m 只蚂蚁随机地分配到 n 个城市，并将出发点城市置入禁忌表 tabu_k. 设置代数计数器 $N_c = \max$，设定 β, α 和 ρ 的值，$\forall \tau_{ij} = \tau_0$，$i, j \in R$.

Step 2: 对于所有的 k，以当前城市 i 为中心，选择下一个节点 $j \in R$. 根据定义 3，首先从 $n - 1$ 个城市中找出 $|V|$ 个未走过的城市 $\{j_1, j_2, \cdots, j_p, \cdots, j_v\}$，即 $j_p \notin \mathrm{tabu}_k$.

Step 3: 按式(3) 或(4) 在 $|V|$ 个城市中选择节点

$$j = \begin{cases} \arg\max\limits_{j \notin \mathrm{tabu}_k} \{[\tau_{ij}(t)][\eta_j(t)]^\beta, q \leqslant q_0; \\ S, \text{otherwise}. \end{cases} \quad (3)$$

其中: $0 < q_0 \leqslant 1$ 是初始设定的参数，$q \in (0, 1)$ 是随机数，随机变量 S 根据下式决定:

$$p_{ij}^k(t) = \begin{cases} \dfrac{[\tau_{ij}(t)][\eta_j]^\beta}{\sum\limits_{h \notin \mathrm{tabu}_k} [\tau_{ih}(t)][\eta_h]^\beta}, j \notin \mathrm{tabu}_k; \\ 0, j \in \mathrm{tabu}_k. \end{cases} \quad (4)$$

其中: $p_{ij}^k(t)$ 表示 t 时刻蚂蚁 k 由节点 i 转移到节点 j 的概率: β 为启发信息的重要程度. 当 $q > q_0$ 时，计算 $|V|$ 个城市的转移概率 p_{ij}^k，根据赌轮盘规则选择城市 j，并将 j 加入禁忌表 tabu_k.

Step 4: 局部信息素更新: 用参数 $1 - \rho$ 表示信息素消逝的程度，每只蚂蚁走完一条边后，按下式进行局部信息素更新:

$$\tau_{ij}(t+1) = (1 - \rho)\tau_{ij}(t) + \rho\Delta\tau_{ij}. \quad (5)$$

其中

$$\Delta\tau_{ij} = \sum_{k=1}^{m} \Delta\tau_{ij}^k,$$

$$\Delta\tau_{ij}^k = \begin{cases} Q/l_{jb}, \text{蚂蚁 } k \text{ 走过边 } e_{ij}; \\ 0, \text{otherwise}. \end{cases}$$

Q 为常数，l_{jb} 是 k 从开始城市到当前城市走过的路程长度，$\Delta\tau_{ij}^k$ 表示 k 在本次循环中留在边 e_{ij} 上的信息量，$\Delta\tau_{ij}$ 表示本次循环中边 e_{ij} 上信息量的增量，τ_m 是限定的最小信息素强度，为一小的常数. 若 $\tau_{ij}(t+1) < \tau_{\min}$，则令 $\tau_{ij}(t+1) = \tau_{\min}$.

Step 5: m 只蚂蚁选择完节点 j 后，令 $i_{\text{new}} = j$，返回 Step 2 选择下一个节点.

Step 6: m 只蚂蚁游完一周后，用式(2) 计算路程长度 L_k，找出并保留 $L_{k\min}$，$L_{k\min} = \min L_k$.

Step 7: 信息素全局更新: 所有蚂蚁走完全部城

市后,按下式对信息素进行全局更新:

$$\tau_{ij}^{new} = (1-\alpha)\tau_{ij}^{old} + \alpha\Delta\tau_{ij}. \tag{6}$$

其中

$$\Delta\tau_{ij} = \begin{cases} 1/l_k, e_{ij} \in \text{global-beat-tour}; \\ 0, \text{otherwise.} \end{cases}$$

α为全局信息素挥发系数, l_k 为本次最佳路径的长度, $e_{ij} \in$ global-best-tour 表示 k 所走的边 e_{ij} 属于最佳路径. 若 $\tau_{ij}^{new} < \tau_{min}$, 则令 $\tau_{ij}^{new} = \tau_{min}$.

Step 8: 本次周游得到的 L_{kmin} 与最优长度 l_d 进行比较, 若 $l_{kmin} < l_d$, 则用 l_{kmin} 替换 l_d, 并替换最优路径表.

Step 9: 若设置的计数值 $N_C - 1 \neq 0$, 则清空并初始化禁忌表, 重复上述过程, 直到 $N_C - 1 = 0$ 为止.

4　收敛性分析

定义 7　设封闭巡回路径为 f_i, 若 $\forall t_1, t_2 \in [t_0, t_n]$, $t_2 = t_1 + \Delta t$, 且有

$$d(P(t_1), P(t_2)) = \min\{d(P(t_1), P), P \in V\},$$

满足 $L(P(t_2), P_e) < L(P(t_1), P_e)$, 则 $f_i(t_2) = P(t_2)$ 是向 P_e 严格趋近的节点位置, f_i 是从 P_0 到 P_e 严格趋近的封闭路径. 所有严格趋近的封闭路径的集合记作 F_p, 非严格趋近的封闭路径记作 f_n, 所有 f_n 的集合记作 F_N. Δt 是 k 从当前节点走到下一节点所需时间.

引理 1　对于所有的 e_{ij}, 均有 $\tau_{min} \leqslant \tau_{ij} \leqslant \tau_{max}$, 其中

$$\tau_{max} = \max\left(\frac{\tau_m}{\rho}, \frac{\tau_m}{\rho}\left[\frac{(1-\alpha)\rho+\rho}{(1-\alpha)\rho+\alpha}\right]\right),$$

$$\tau_m = \max(\rho\Delta\tau_{ij}, \alpha\Delta\tau_{ij}).$$

证明　设某代蚂蚁觅食得到最优路径为 f_0, 根据算法步骤和式(3)～(6), 可分以下 3 种情况:

1) 蚂蚁 k 从未走过边 e_{ij}, 该边上的信息素只消散, 不增加. 根据式(5) 有

$$\begin{cases} \tau_{ij}^{min}(t) = (1-\rho)^t\tau_0, t\to\infty, \tau_{ij}\to 0; \\ \tau_{ij} = \tau_{min}\text{(被强制到}\tau_{min}\text{)}. \end{cases} \tag{7}$$

2) 蚂蚁 k 走过边 e_{ij}, 但 $e_{ij} \notin f_0$, 此时该边上的信息素只进行局部更新. 根据式(5) 有

$$\tau(1) = (1-\rho)\tau(0) + \tau_m, \tau_m = \rho\Delta\tau_{ij},$$
$$\tau(2) = (1-\rho)\tau(1) + \tau_m = (1-\rho)^2\tau(0) + (1-\rho)\tau_m + \tau_m,$$
$$\vdots$$
$$\tau(t) = (1-\rho)^t\tau(0) + (1-\rho)^{t-1}\tau_m + (1-\rho)^{t-2}\tau_m + \cdots + \tau_m = (1-\rho)^t\tau(0) + \sum_{i=1}^{t}(1-\rho)^{t-i}\tau_m.$$

其中: $\tau(0), \tau(1), \cdots, \tau(t)$ 分别表示初始状态, 第 1 代, …, 第 t 代搜索后的信息素强度; $\sum_{i=1}^{t}(1-\rho)^{t-i}$ 为一等比数例, 公比为 $1-\rho$, 其前 t 项之和为 $\frac{1-(1-\rho)^t}{1-(1-\rho)}$. 因此有

$$\begin{cases} \tau_{ij}^{max}(t) = (1-\rho)^t\tau_0 + \sum_{i=1}^{t}(1-\rho)^{t-i}\tau_m; \\ \tau_m = \rho\Delta\tau_{ij}, t\to\infty, \tau_{ij}^{max}(t)\to\frac{1}{\rho}\tau_m. \end{cases} \tag{8}$$

3) 蚂蚁 k 走过边 e_{ij} 且 $e_{ij} \in f_0$, 此时该边上的信息素进行局部和全局更新. 根据式(5) 和(6) 有

$$\begin{cases} \tau_{ij}^{max}(t) = \\ (1-\rho)^t(1-\alpha)^t\tau_0 + \sum_{i=1}^{t}(1- \\ \rho)^{t-i}(1-\alpha)^{t-i}[(1-\alpha)\tau_m + \tau_m], \\ t\to\infty, \tau_{ij}^{max}\to\frac{\tau_m}{\rho}\left[\frac{(1-\alpha)\rho+\rho}{(1-\alpha)\rho+\alpha}\right]. \end{cases} \tag{9}$$

其中 $\tau_m = \max(\rho\Delta\tau_{ij}, \alpha\Delta\tau_{ij}^k)$.

比较以上 3 种情况可知, 对于所有的 τ_{ij}, 都有 $\tau_{min} \leqslant \tau_{ij} \leqslant \tau_{max}$. □

引理 2　当 $q_0 = 1$ 时, $\forall e_{ij} \in f_i, f_i \in F_p$, $\forall e_{kl} \in f_n, f_n \in F_N$, 均有 $\tau_{kl} \leqslant \tau_{ij}$.

证明　当 $q_0 = 1$ 时, k 恒用式(3) 选择下一个节点. 根据算法步骤和式(3), (5), (6), 分为以下 3 种情况:

1) $\forall e_{kl}(k, l \in R) \notin f_i, e_{kl} \in f_n$, 在可行边集中, $\exists e_{ij}(i, j \in R) \in f_i$, 使得 $\eta_{ij} < \eta_{kl}$. 初始化时 $\forall \tau_{ij}(i, j \in R) = \tau_0$, 因而 k 用式(3) 选择并行走边 e_{kl} 的概率为 0, 从而 $\tau_{kl} = \tau_{min}$.

2) $\forall e_{ij} \in f_i, e_{kl} \notin f_0$, τ_{ij}^{max} 由式(8) 计算.

3) $\forall e_{ij} \in f_i, e_{kl} \in f_0$, τ_{ij}^{max} 由式(9) 计算.

比较以上 3 种情况可知, 均有 $\tau_{ij} \geqslant \tau_{kl} = \tau_{min}$. □

定理 1　设客观存在的全局最优解为 f_i 类型, 当 $q_0 = 1$ 时, 算法一定能收敛于最优解.

证明　$\forall t_i \in [t_0, t_n]$, 设 k 的当前位置为 $P(t_i)$, $P(t_i) \sim c_i$, k 从 c_i 选择并走到下一个节点 c_j, $c_j \sim P(t_{i+1})$. 由于全局最优解为 f_i 类型, 根据定义 4 和定义 7, 必然存在一个严格趋近的节点位置 $P(t_{i+1})$, 即 $\forall P(t_{i+1}) \in V$, $\exists P(t_{i+1})$, 使得

$$d(P(t_i), P(t_{i+1})) = \min\{d(P(t_i), P), P \in V\},$$

满足 $L(P(t_{i+1}), P_e)$ 为最小, 记为 $P(t_j)$. 设 k 从 $P(t_i)$ 选择 $P(t_j)$ 的概率为 p, 由于 $q_0 = 1$, 算法恒用式(3) 选择 c_j. 由式(3) 和引理 2 可得 $p \equiv 1$, 因而 k 在 t 代觅食过程中, 选择最优解的概率 $p_t = 1 - (1 - p^n)^t \equiv 1$. □

从定理 1 的证明可知, 若全局最优解为 f_i 类型,

则算法不仅能收敛于最优解,而且能较快收敛.

推论 1 设客观存在的全局最优解 $f_0 \in f_i$ 类型,则 $q_0 < 1$ 时的收敛时间一定比 $q_0 = 1$ 时长.

这一推论是显然的,用定理 2 的证明方法可证.

推论 2 设客观存在的全局最优解 $f_0 \notin f_i$ 类型,当 $q_0 = 1$ 时,算法不能收敛到最优解.

证明 由于客观存在的全局最优解 $f_0 \notin f_i$ 类型,$\exists e_{kl} \in f_n$,由引理 2 的证明可知,$q_0 = 1$ 时恒用式(3) 选择非趋近节点的概率为 0. 设 f_0 包括 x 个趋近节点和 y 个非趋近节点,选择一个趋近节点的概率为 R,选择一个非趋近节点的概率为 r, 则 k 在任意一代觅食过程选择 f_0 的概率 $p = R^x r^y = 0$. □

定理 2 设客观存在的全局最优解 $f_0 \notin f_i$ 类型,当 $0 < q_0 < 1$ 时, 经过充分长的时间 t_m,算法能收敛到最优解.

证明 设任意一个一般解 F_{p_i} 包括 x 个趋近的节点和 y 个非趋近的节点. 选择一个任意趋近节点的概率为 R, 选择一个任意非趋近节点的概率为 r, 估算 R 和 r.

根据算法,$q \leqslant q_0$ 时 k 根据式(3) 选择节点,否则根据式(4) 选择节点. 这些节点即可以是趋近节点,也可以是非趋近节点. 设 p_3^x 和 p_3^y 分别为 $q \leqslant q_0$ 时选择一个趋近和非趋近节点的概率,p_4^x 和 p_4^y 分别为 $q > q_0$ 时选择一个趋近和非趋近节点的概率. 根据算法有

$$R = q_0 p_3^x + (1 - q_0) p_4^x,$$
$$r = q_0 p_3^y + (1 - q_0) p_4^y.$$

由引理 1 和引理 2,定义 6,式(3) 可知,$p_3^y = 0, p_3^x = 1, \forall \tau > \tau_{\min} > 0, \eta_{\min} > 0$. 因此,由式(4) 选择 F_p 上任意一个节点的极限最小概率为

$$p_{\min} \geqslant \frac{\tau_{\min} \eta_{\min}^{\beta}}{\sum\limits_{h \in (|V| - 1)} \tau_{\max} \eta_{\max}^{\beta} + \tau_{\min} \eta_{\min}^{\beta}} > 0.$$

当 $q > q_0$ 时,用式(4) 计算出的概率 $p_{ij} > 0$, 因而 $p_4^x > 0, p_4^y > 0, 0 < q_0 < 1$. 由此推知 $R > 0, r > 0$. k 经过 t 代觅食至少有一次找到最优解的概率为

$$p(t) \geqslant 1 - (1 - R^x r^y)^t, t \to \infty, P(t) \to 1. \tag{10}$$

即当 $t = t_m$ 是一个较大的值时,有 $p(t) = 1 - \epsilon \approx 1$. 定理得证. □

定理 2 说明,当 $0 < q_0 < 1$ 时,不管信息素如何变化,经过充分长的时间总能收敛到最优解. 通过证明可知,当最优解的多个节点不属于 f_i 且不考虑信息素作用时,其收敛到最优解的时间会很长.

下面分析 $q_0 < 1$ 时信息素变化对收敛速度的影响. 令 g 为最大与最小信息素之比,根据引理 1 有

$$g = \tau_{\max} / \tau_{\min} = \max\Big(\frac{\tau_m}{\rho}, \frac{\tau_m[(1 - \alpha)\rho + \rho]}{\rho[(1 - \alpha)\rho + \alpha]}\Big) \Big/ \tau_{\min}.$$

设 k 从 $P(t_i)$ 选择 $P(t_{i+1}) \in f_0$ 的最大概率分别为 $p_{\max}$ 和 $p_{\max}^{\tau}$. 前者是不考虑信息素时的最大概率,则有

$$p_{\max} \leqslant \frac{\eta_{\max}^{\beta}}{\sum\limits_{h \in (|V| - 1)} \bar{\eta}^{\beta} + \eta_{\max}^{\beta}} = \frac{b\bar{\eta}^{\beta}}{\sum\limits_{h \in (|V| - 1)} \bar{\eta}^{\beta} + b\bar{\eta}^{\beta}} = \frac{b}{|V| - 1 + b}. \tag{11}$$

其中: $\bar{\eta}$ 为启发函数平均值,$b\bar{\eta}^{\beta} = \eta_{\max}^{\beta}$.

现在估算考虑信息素时选择优化解结点的最大概率 $p_{\max}^{\tau}$. 设 $\forall e_{ij} \in f_0$,有 $\tau_{ij} = \tau_{\max}$, $\forall e_{ij} \notin f_0$,有 $\tau_{ij} = \tau_{\min}$. 根据式(4) 有

$$p_{\max}^{\tau} \leqslant \frac{\tau_{\max}[\eta_{\max}]^{\beta}}{\sum\limits_{h \in (|V| - 1)} \tau_{\min} \bar{\eta}^{\beta} + \tau_{\max}[\eta_{\max}]^{\beta}} = \frac{\tau_{\max} b \bar{\eta}^{\beta}}{\sum\limits_{h \in (|V| - 1)} (\tau_{\max}/g) \bar{\eta}^{\beta} + \tau_{\max} b \bar{\eta}^{\beta}} = \frac{g b}{|V| - 1 + g b}. \tag{12}$$

实际上,非优化解边上的信息素并不全是 $\tau_{\min}$,因而此概率为最大极限值. 两概率之比为

$$\frac{p_{\max}^{\tau}}{p_{\max}} = \frac{|V| - 1 + b}{(|V| - 1)/g + b} > 1. \tag{13}$$

一般 $g \gg 1$,多数情况 $|V| - 1 \gg b$,因而有 $p_{\max}^{\tau}/p_{\max} \gg 1$. 由式(10),(12) 和(13) 可知,在信息素的作用下,会使收敛速度有较大的提高.

推论 3 使用最近邻策略可使收敛速度提高.

证明 k 每次进行节点选择时,仅在相邻域内的可行点集 z 内选择即可. 这时式(11) 和(12) 中的 $|V|$ 由 $|z|$ 代换,由于 $|z| \ll |V|$,由式(10) 和(12) 可知,其收敛速度将有较大提高. 另一方面, 从算法原理看,使用最近邻策略使计算量大幅减少,也有利于收敛速度的提高. □

推论 4 加大最优路径上的信息素,在一定范围内会使收敛速度提高.

证明 根据式(12) 和(13),在一定范围内,g 越大,选择优化解结点的概率越大; 根据式(10),其收敛速度也越快. □

作者在有关改进算法的研究中,取得了与本节分析结论相吻合的结果,证实了本节分析结论的正确性. 有关仿真实验结果参见文献[6].

(下转第 770 页)

提高,而且提供了样本属于所在类中的可信程度. 实验结果表明,与 Platt 的概率建模方法相比,本文提出的基于最大熵估计的概率建模方法具有优良的性能. 在下一步工作中,作者将对支持向量机方法中多类分类问题进行概率建模的研究.

参考文献(**References**)

[1] Wahba G. Support Vector Machines, Reproducing Kernel Hilbert Spaces and the Randomized GACV [A]. *Advances in Kernel Methods Support Vector Learning* [C]. Massachusetts: MIT Press, 1999: 69-88.

[2] Platt J C. Probabilities for Support Vector Machines [A]. *Advances in Large Margin Classifiers* [C]. Massachusetts: MIT Press, 2000: 61-74.

[3] Sollich P. Bayesian Methods for Support Vector Machines: Evidence and Predictive Class Probabilities [J]. Machine Learning, 2002, 46: 21-52.

[4] Kwok J T Y. Moderating the Outputs of Support Vector Machine Classifiers[J]. *IEEE Trans on Neural Networks*, 1999, 10(5): 1018-1031.

[5] Hastie T, Tibshirani R. Classification by Pairwise Coupling [J]. *The Annals of Statistics*, 1998, 26(1): 451-471.

[6] Vapnik V. *Statistical Learning Theory* [M]. New York: Wiley, 1998.

[7] 张文生, 王珏, 戴国忠. 支持向量机中引入后验概率的理论和方法研究[J]. 计算机研究与发展, 2002, 39(4): 392-397.
(Zhang W S, Wang J, Dai G Z. Study of Theory and Method Introducing Posteriori Probability into Support Vector Machines [J]. *J of Computer Research and Development*, 2002, 39(4): 392-397.)

[8] Lin H T, Lin C J, Weng R C. *A Note on Platt's Probabilistic Outputs for Support Vector Machines*[R]. National Taiwan University, Taibei, 2003.

[9] 张翔. 支持向量机及其在医学图像分割中的应用[D]. 武汉: 华中科技大学, 2004.
(Zhang X. *Support Vector Machine and Its Applications in Medical Image Segmentation*[D]. Wuhan: Huazhong University of Science and Technology, 2004.)

[10] Theodoridis S, Koutroumbas K. *Pattern Recognition* [M]. Amsterdam: Elsevier Press, 2003.

(上接第 766 页)

5 结 论

本文分析了用蚁群算法求解 TSP 问题的收敛性. 得到的结论是: 若全局最优解属于严格趋近的封闭路径, 当设定 $q_0 = 1$ 时, 算法一定能快速收敛于最优解; 当设定 $q_0 < 1$ 时, 会使收敛速度变慢. 若全局最优解不属于严格趋近的封闭路径, 当设定 $q_0 = 1$ 时, 算法不能收敛到最优解; 当设定 $0 < q_0 < 1$ 时, 算法能够收敛到最优解, 但收敛时间将很长. 在一定范围内加大优化路径上的信息素, 会加快收敛. 根据这一结论, 在算法设计时, 应估计或分析优化解的性质, 据此合理地设定 q_0, 并在一定范围内加大全局优化路径上的信息素, 以利于提高收敛速度.

通过本文的分析可以看出, 影响蚁群算法收敛速度的两大因素是信息素和启发函数. 选择合适的启发函数对提高收敛速度至关重要, 启发信息的值过大, 将抑制信息素的作用; 启发信息的值过小, 则会影响收敛速度. 通过分析证明, 采用最近邻选择策略可以提高收敛速度.

参考文献(**References**)

[1] Colorni A, Dorigo M, Maniezzo V. Distributed Optimization by Ant Colonies [A]. *Proc of ECAL'91 European Conf on Artificial Life* [C]. Paris: Elsevier Publishing, 1991: 134-144.

[2] Thomas Stutzle, Marco Dorigo. A Short Convergence Proof for a Class of Ant Colony Optimization Algorithms [J]. *IEEE Trans on Evolutionary Computation*, 2002, 6(4): 358-365.

[3] Amr Badr, Ahmed Fahmy. A Proof of Convergence for Ant Algorithms[J]. *Infomation Sciences*, 2004, 160: 267-279.

[4] Walter J. Gutjahr. A Graph-based Ant System and Its Convergence [J]. *Future Generation Computer Systems*, 2000, 16: 873-888.

[5] Walter J Gutjahr. ACO Algorithms with Guaranteed Convergence to the Optimal Solution [J]. *Infomation Processing Letters*, 2002, 82: 145-153.

[6] 朱庆保, 杨志军. 基于变异和动态信息素更新策略的蚁群算法[J]. 软件学报, 2004, 15(2): 185-192.
(Zhu Q B, Yang Z J. Ant Colony Optimization Algorithm Based on Mutation and Dynamic Pheromone Updating[J]. *J of Software*, 2004, 15(2): 185-192.)