

遗传算法与蚂蚁算法融合的 马尔可夫收敛性分析¹⁾

丁建立 陈增强 袁著祉

(南开大学信息技术科学学院 天津 300071)

(E-mail: jianliding@yahoo.com.cn)

摘 要 遗传算法具有快速随机的全局搜索能力,但不能很好地利用系统的反馈信息. 蚂蚁系统是一种并行的分布式正反馈系统,但初始求解速度慢. 遗传算法与蚂蚁算法的融合,优势互补. 基于上述思想,提出遗传算法与蚂蚁算法融合模型与方法,对该方法的收敛性进行了马尔可夫理论分析,并证明其优化解满意值序列是单调不增的和收敛的. 且对 NP-hard 问题中的 30 城市 TSP 和中国 CHN144 城市 TSP 两个实例进行了实验分析,仿真数据表明该方法不仅是一个逐步收敛的过程,而且求解速度和求解效果都非常好.

关键词 遗传算法, 蚂蚁算法, 融合, 马尔可夫过程, 收敛性

中图分类号 TP301

On the Markov Convergence Analysis for the Combination of Genetic Algorithm and Ant Algorithm

DING Jian-Li CHEN Zeng-Qiang YUAN Zhu-Zhi

(College of Information Technology and Science, Nankai University, Tianjin 300071)

(E-mail: jianliding@yahoo.com.cn)

Abstract Genetic algorithm has the ability of quickly and stochastically global searching, however, it can not make good use of enough output information for systems. Ant system is a parallel-process and distributive-forward system with a relatively slow velocity for providing the solution. Combining genetic and ant algorithms can increase the merits each other. Based on the idea above, the model and method from the combination of genetic and ant algorithms are proposed, and the convergence of the method based on the Markov theory is analysed. Moreover, the conclusion can be drawn that the solution sequence is monotonically decreasing and convergent. The experiment and analysis are carried out for the cases of TSP30 and CHN144 on an NP-hard problem. The results of simulation show that not only the mixed algorithm is a step-by-step convergent process,

1) 国家自然科学基金(60174021,60374037)和河南科技攻关项目(0124140141)资助

Supported by National Natural Science Foundation of P. R. China (60174021,60374037), Key Technology Project of Henan (0124140141)

收稿日期 2002-12-27 收修改稿日期 2003-06-23

Received December 27, 2002; in revised form June 23, 2003

but also its velocity and effect of solving are quite satisfactory.

Key words Genetic algorithm, ant algorithm, combination, Markov process, convergence

1 遗传算法与蚂蚁算法的融合思想与过程

1.1 遗传算法与蚂蚁算法的融合思想

遗传算法(genetic algorithm, GA)具有大范围快速全局搜索能力,但当求解到一定范围时往往作大量无为的冗余迭代,对于系统中的反馈信息利用不够,求精确解效率低. 蚂蚁算法^[1](ant algorithm, AA)主要是通过蚂蚁群体之间的信息素的传递和更新达到最终收敛于最短路径上,其原理是一种正反馈机制,但初期信息素匮乏,求解速度慢. 遗传算法与蚂蚁算法的融合(genetic algorithm-ant algorithm, GAAA)正是基于遗传算法的快速全局搜索能力和蚂蚁算法的正反馈收敛机制,初期采用遗传算法过程生成信息素分布,后期利用蚂蚁算法正反馈求精确解,优势互补.

1.2 遗传算法与蚂蚁算法的融合过程

1.2.1 GAAA 中遗传算法的定义与设置

GAAA 中的遗传算法是基于优胜选择遗传算法的原理与定义. 其鲜明特色是在选择、杂交、变异的过程中最后一个个体取原种群中适应值最小的个体,因而保留了种群序列满意适应值的单调不增性等优良特性,使遗传过程不断进化. 具体是,采用十进制实数编码并确定相应目标的适应值函数;根据适应值函数选择准备进行交配的一对染色体父串;采用 Davis 顺序交叉方法;采用逆转变异方法^[2],即经逆转后适应值有提高保留,否则逆转无效.

1.2.2 GAAA 中对蚂蚁算法模型选择与改进

GAAA 中对蚂蚁算法选择基于蚂蚁圈模型和 MMAS(MAX-MIN ant system)算法^[3],在吸取其各自优点的基础上并进行改进.

信息素更新模型:采用蚂蚁圈模型进行信息素更新,即一圈中只有最短路径的蚂蚁才进行信息素修改增加,其轨迹强度的更新方程和蚂蚁的转移概率可定义^[1]为

$$\tau_{ij}(t+1) = \rho \cdot \tau_{ij}(t) + \sum \Delta \tau_{ij}^k(t) \quad (1)$$

$$p_{ij}^k(t) = \begin{cases} \frac{[\tau_{ij}(t)]^\alpha [\eta_{ij}]^\beta}{\sum_{l \in U} [\tau_{il}(t)]^\alpha [\eta_{il}]^\beta}, & j \in U \\ 0, & j \notin U \end{cases} \quad (2)$$

信息素的初值设置:为了更加充分地进行寻优,MMAS 把各路径信息素初值设为最大值 τ_{\max} ,为了避免算法过早收敛非全局最优解,MMAS 将各路经的信息素浓度限制在 $[\tau_{\min}, \tau_{\max}]$ 之间,实验表明在防止算法过早停滞及有效性方面取得了很好的效果^[4]. 这里我们通过遗传算法得到了一定的路径信息素,所以把信息素的初值设置为 $\tau_s = \tau_c + \tau_g$,其中 τ_c 是一个根据具体求解问题规模给定的一个信息素常数,相当于 MMAS 算法中的 τ_{\min} , τ_g 是遗传算法求解结果转换的信息素值.

2 遗传算法与蚂蚁算法融合的马尔可夫收敛性分析

设 $S^N = \{x = (x_1, x_2, \dots, x_N), x_i \in S (i \leq N)\}$ 为种群(优化解)空间, $S^2 = \{(x_1, x_2), x_1, x_2 \in S\}$ 称为母体空间, S 称为个体空间, x_i 称为个体, 其长度 l 称为链长, N 称为种群(优化解)规模. n 为优胜选择遗传算法的迭代次数, m 为蚂蚁圈算法的移动周期, P 为 S^N 上的概率分布. 在这里, 我们可以把遗传算法与蚂蚁算法融合的 GAAA 算法看成是优胜选择遗传算法的扩展, 它由四个过程构成: 选择算子 T_s ; 杂交算子 T_c ; 变异算子 T_m ; 信息素算子 T_g . 其基本概念定义如下.

定义 1^[5]. 适应值函数 $f: S \rightarrow R^+$, 即个体空间到正实数空间的映射, 则全局最优解集为 $M = \{x; \forall y \in S, f(x) \leq f(y)\}$; 满意解集为 $B = \{x \in B; \forall y \notin B, f(x) \leq f(y)\}$.

定义 2^[5]. 选择算子 $T_s: S^N \rightarrow S$, 即在一个种群中选择一个个体的随机映射, 则母体选择算子的概率为

$$P\{T_s(x) = (x_i, x_j)\} = \frac{f(x_i)}{\sum_{k=1}^N f(x_k)} \cdot \frac{f(x_j)}{\sum_{l=1}^N f(x_l)} \quad (3)$$

定义 3^[5]. 杂交算子 $T_c: S^2 \rightarrow S$, 即母体空间到个体空间的映射, k 为可以生成 y 的基因位置的个数, 则单点杂交和单点随机杂交概率分别为

$$p_c = P\{T_c(x_1, x_2) = y\} = \frac{k}{l} \quad (4)$$

$$P\{T_c(x_1, x_2) = y\} = \begin{cases} \frac{kp_c}{l}, & y \neq x_1 \\ (1 - p_c) + \frac{kp_c}{l}, & y = x_1 \end{cases} \quad (5)$$

定义 4^[5]. 变异算子 $T_m: S \rightarrow S$, 即个体空间到个体空间的随机映射.

$$P\{T_m(x) = y\} = p_m^{d(x,y)} (1 - p_m)^{l-d(x,y)} \quad (6)$$

$d(x, y) = \sum_{i=1}^l |a_i - b_i|$ 表示 X 与 Y 之间的 Hamming 距离, p_m 表示变异的概率.

定义 5. 信息素算子 $T_g: S^N \rightarrow S$, 即在一个蚂蚁圈序列中选择一个最优个体信息素更新的映射. 信息素算子概率为

$$P\{T_g(x) = x_i\} = \frac{f(x_i)}{\sum_{k=1}^N f(x_k)} \quad (7)$$

引理 1^[5]. 优胜选择遗传算法种群序列 $\{x(n); n \geq 0\}$ 是有限齐次马尔可夫链.

引理 2^[6]. 蚂蚁系统序列 $\{\tau(m), x(m), f^*(m)\} (m = 1, 2, \dots)$ 是有限齐次马尔可夫链.

定理 1. GAAA 算法的优化解序列 $\{x(n); n \geq 0\}$ 是有限齐次马尔可夫链.

证明. 由引理 1 知优胜选择遗传算法 $x(n+1)$ 仅与 $x(n)$ 有关, 与迭代次数 n 无关. 由引理 2 知, 蚂蚁系统 $\{\tau(m+1), x(m+1), f^*(m+1)\}$ 仅与 $\{\tau(m), x(m), f^*(m)\}$ 有关, 而与循环周期 m 无关. 同时, 由蚂蚁圈模型知, 每一次信息素仅更新最优蚂蚁圈, 它具有优胜选择遗传算法所有的特征. 因而, 我们取

$$x(n+1) = (x_i(n+1) = T_g^i \circ T_m^i \circ T_c^i \circ T_s^i(x(n)); (i \leq N-1); x_N(n+1) = x_{i_0}(n)) \quad (8)$$

这里 $i_0 = \arg \min_j \{f(x_j(n))\}$ 表示使 $f(x_j(n))$ 取最小值的个体为 $x_j(n)$, 且转移概率矩阵为

$$P\{x, y\} = P\{x(n+1) = y/x(n) = x\} = \begin{cases} \prod_{k=1}^N P\{T(x(n))_k = y_k\}, & \exists i_0 \in M(x), \text{使 } y_N = x_{i_0} \\ 0, & \text{else} \end{cases} \quad (9)$$

其中 $M(x) = \{i; f(x_i) = \min\{f(x_j)\}\}$, 则

$$T(x(n)) = T_g^i(T_m^i(T_c^i(T_s^i(x(n)))) = T_g^i \circ T_m^i \circ T_c^i \circ T_s^i(x(n)) \quad (10)$$

上式表明, $x(n+1)$ 仅与 $x(n)$ 有关, 而与 n 无关. GAAA 算法的优化解序列 $\{x(n); n \geq 0\}$ 是有限齐次马尔可夫链. 证毕.

定理 2. GAAA 算法的马尔可夫链序列的优化解满意值序列是单调不增的, 即对于任意的 $n \geq 0$, 有 $F(x(n)) \geq F(x(n+1))$.

证明. 首先, 由于 GAAA 中采用的是优胜选择遗传算法, $i_0 = \arg \min_j \{f(x_j(n))\}$, 则有

$$x_N(n+1) = x_{i_0}(n), F(x(n)) \geq F(x_N(n+1)) \geq F(x(n+1)) \quad (11)$$

其次, GAAA 中采用了蚂蚁圈算法, $i_0 = \arg \min_j \{f(x_j(m))\}$, 同样有

$$x_N(m+1) = x_{i_0}(m), F(x(m)) \geq F(x_N(m+1)) \geq F(x(m+1)) \quad (12)$$

而蚂蚁圈算法是以优胜选择遗传算法结果为初始分布, 具有优值继承性, 因而对于任意的 $n \geq 0$, 有 $F(x(n)) \geq F(x(n+1))$, 即 GAAA 算法的马尔可夫链序列的优化解满意值序列是单调不增的. 证毕.

引理 3^[5]. 优胜选择遗传算法种群马尔可夫链序列 $\{x(n); n \geq 0\}$ 以概率 1 收敛到满意种群集 M^* 的子集 M_0^* , $M_0^* = \{y = (y_1, \dots, y_N); y_N \in M\}$, 即

$$\lim_{n \rightarrow \infty} P\{x(n) \in M_0^* / x(0) = x_0\} = 1 \quad (13)$$

定理 3. GAAA 算法的优化解马尔可夫序列以概率 1 收敛到满意解集 B 的子集 B_0^* , $B_0^* = \{y = (y_1, \dots, y_N); y_N \in B\}$, 即

$$\lim_{n \rightarrow \infty} P\{x(n) \in B_0^* / x(0) = x_0\} = 1 \quad (14)$$

证明. 设 x' 是 $f(x)$ 的唯一最小值解, 由定理 1 和引理 3 知 $P\{x, y\}$ 有性质:

1) 当 $x, y \in B_0^*$ 时, $P\{x, y\} > 0, P\{y, x\} > 0$, 即 $x \leftrightarrow y$; 2) 当 $x \in B_0^*, y \notin B_0^*$ 时, $P\{x, y\} = 0$, 即 $x \rightarrow y$.

因而, B_0^* 为正常返的非周期的不可约闭集, $S^N \setminus B_0^*$ 为非常返的状态集.

$$\lim_{n \rightarrow \infty} P\{x(n) = y/x(0) = x_0\} = \begin{cases} \pi(y), & y \in B_0^* \\ 0, & y \notin B_0^* \end{cases} \quad (15)$$

$$\text{于是} \quad \lim_{n \rightarrow \infty} P\{x(n) \in B_0^* / x(0) = x_0\} = 1 \quad (16)$$

证毕.

推论 1. 对 GAAA 算法, 如果 $\{x(n); n \geq 0\}$ 对于任意满意解集是收敛的, 则必收敛到全局最优解集.

因为对于 $\epsilon > 0$ 有满意解集 $B(\epsilon) = \{y; f(y) \leq f(x^*) + \epsilon\}$, 其中 x^* 表示最优解, 即 $x^* \in M$, 亦即全局最优解集 M 即是所有满意解集之交集, M 为最小满意解集. 实际问题中只要求收敛到满意解集 $B(\epsilon)$ 即可.

3 仿真实验分析

我们采用 GAAA 算法分别对典型的 NP-hard 问题 30 城市 TSP 问题和中国 CHN144 城市问题进行了实验, GAAA 中遗传算法迭代次数分别选定为 30 代和 144 代, 蚂蚁算法中各路径信息素初值 τ_c 分别设为 60 和 600, 遗传算法求解结果转换的信息素值是经过路径加 2, 轨迹更新分别取 $\rho = 0.8, Q = 1000$ 和 $\rho = 0.9, Q = 100000$. 表 1、表 2、图 1 和图 2 是仿真实验结果.

表 1 GAAA 算法优化解数据逼近过程
Table 1 An approach process of GAAA optimization-solution

GAAA 过程	TSP30 优化解值			CHN144 优化解值		
	最大值	最小值	平均值	最大值	最小值	平均值
初始随机生成优化解值	1500	1209	1318.8	88763	65739	76316
选择算子作用优化解值	1198	1101	1139.1	65564	61278	63211
杂交算子作用优化解值	1095	1005	1050.3	56405	54329	55323
变异算子作用优化解值	987	817	912.3	51994	51344	51508
蚂蚁算法作用优化解值	452	424	433.7	31153	30354	30612

表 2 GAAA 算法随机求解的 30 个优化解值分布
Table 2 A distribution of GAAA 30 optimization-solution value

TSP30 优化解值分布 ($\alpha=1, \beta=2, \rho=0.8, Q=1000$)								CHN144 优化解值分布 ($\alpha=2, \beta=2, \rho=0.9, Q=100000$)							
436	430	431	439	426	437	433	429	30355	30585	30643	30730	30807	30928	31153	30810
434	439	426	438	424	426	425	446	30748	30617	30507	30358	30379	30357	30595	30551
449	426	424	443	434	427	452	436	30676	30760	30736	30788	30563	30354	30360	30431
426	425	448	431	440	430			30566	30739	30792	30743	30380	30356		

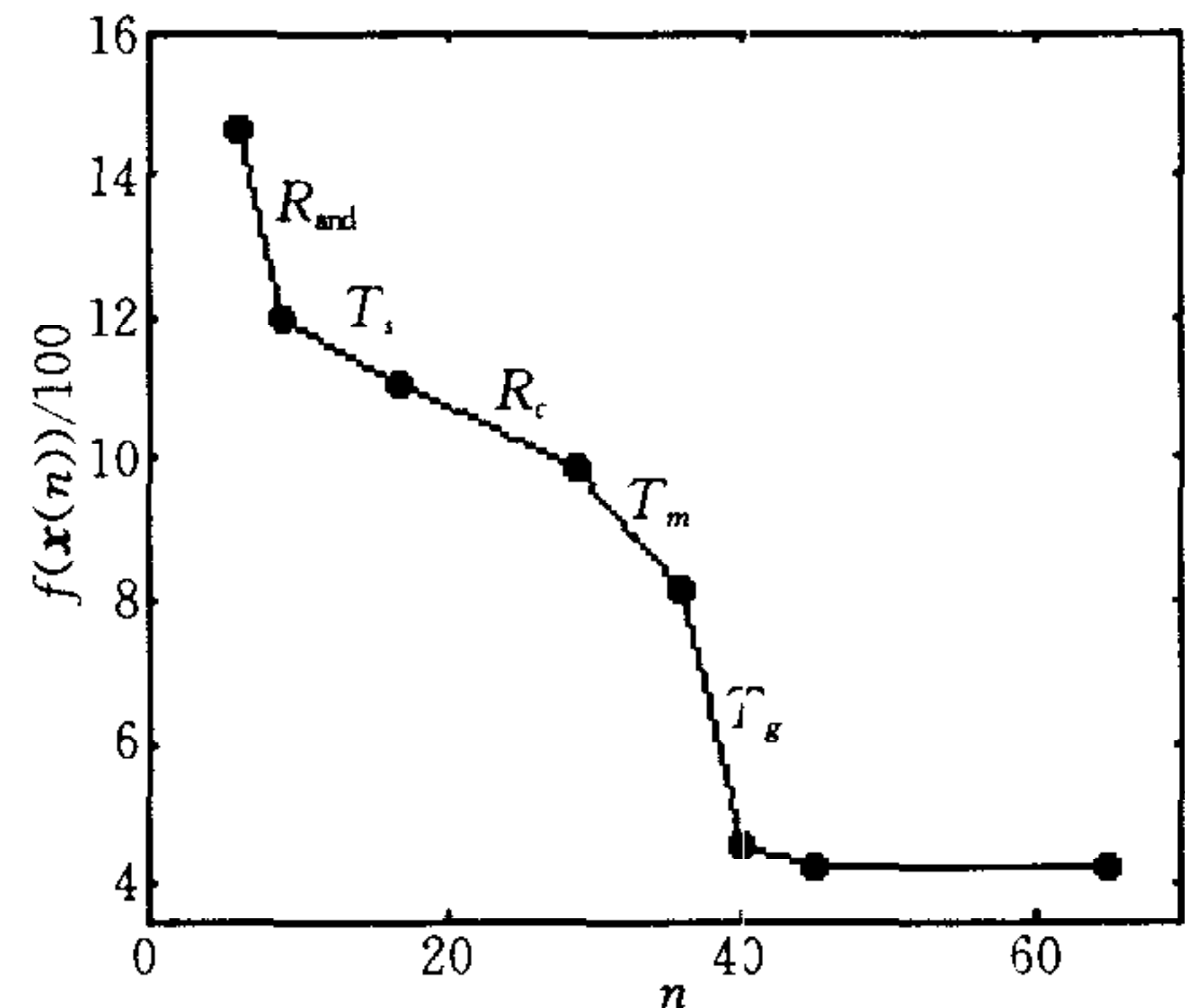


图 1 GAAA 一次随机求解过程(TSP30)

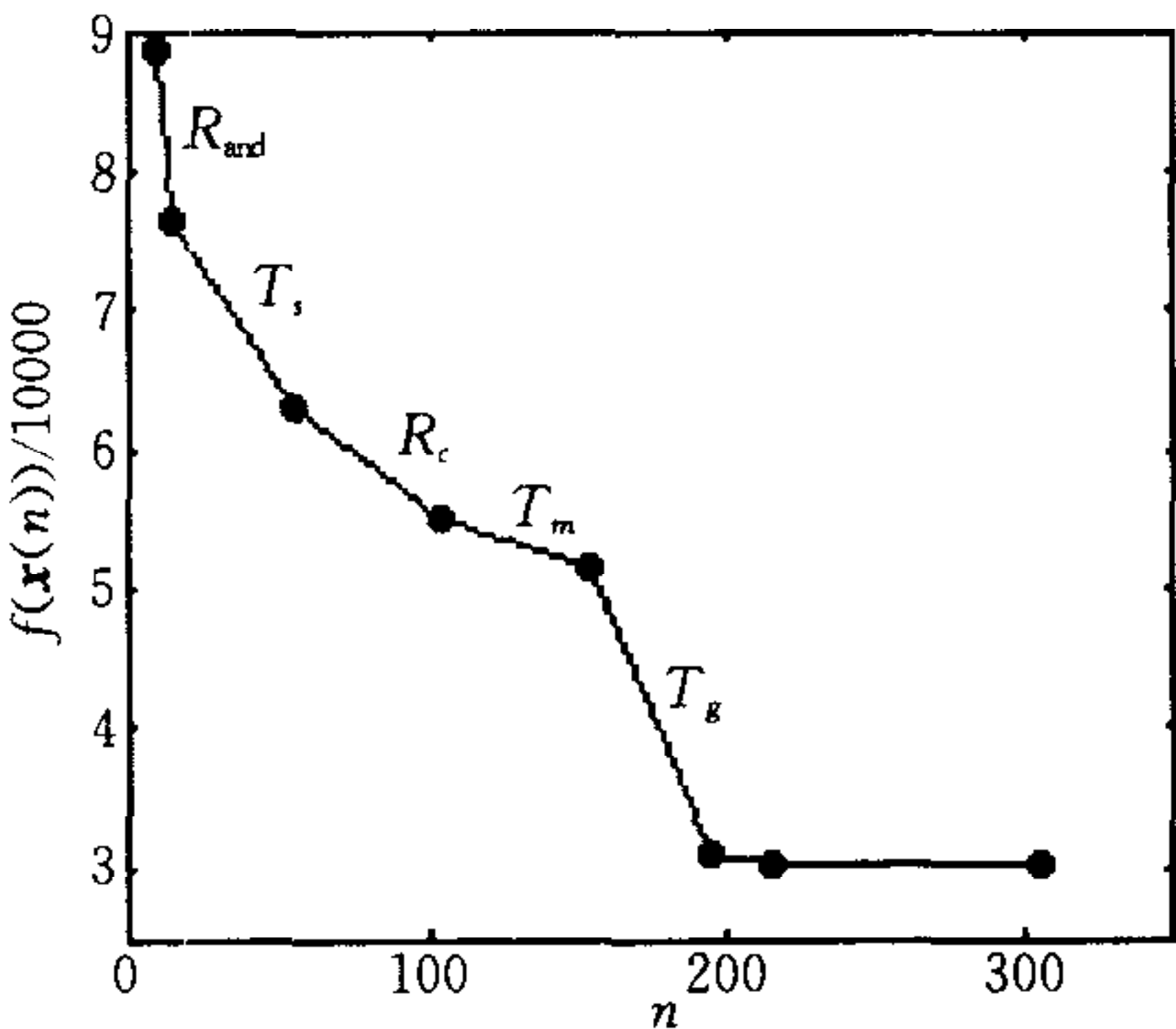


图 2 GAAA 一次随机求解过程(CHN144)

Fig. 1 A stochastic GAAA solving processes(TSP30) Fig. 2 A stochastic GAAA solving processes(CHN144)

表 1 反映 GAAA 算法优化解数据逼近过程, 依次从随机产生初始优化解→选择算子→杂交算子→变异算子→蚂蚁圈更新优化. 无论从最大值、最小值、平均值分布来看, 本文算法是一个逐步收敛的过程, 这和理论上的分析是一致的.

表 2 是 GAAA 算法求解的 30 个优化解值分布, 从中可以看出求解的精度非常高. 用

GAAA 算法不仅得到目前公认的 30 个城市 TSP 问题的最好解 ($d^* = 423.74$) 及中国 CHN144 城市问题的最好解 ($d = 30354.3$), 而且得到了一大批与最好解非常接近的满意解, 这在实践上非常有意义.

图 1 和图 2 是 GAAA 算法分别对 TSP30 问题和 CHN144 问题的一次随机求解过程, 从中不仅可以看出其单调递减性, 而且它清晰的刻画了每种操作对求解过程的影响. 其中 T_g 初始阶段非常陡峭, 它证实了在经过了遗传选择、杂交、变异操作后, 蚂蚁算法利用其生成信息素迅速逼近最优解. 这就克服了蚂蚁算法因初始信息素匮乏而初始求解速度慢的弊病. 而从迭代的过程来看, TSP30 需经过 40 次左右的迭代运算, 即达到满意解状态, CHN144 需经过 200 次左右的迭代运算, 即达到满意解状态, 显示了极高的求解效率.

4 结论

遗传算法与蚂蚁算法的融合, 优势互补, 取得了非常好的效果. 通过理论分析, GAAA 算法的满意解序列为一个单调递减的马尔可夫过程, 并且一定能够收敛. 仿真分析表明, GAAA 算法在经过优胜选择遗传算法形成信息素分布后, 改进的蚂蚁圈算法以极快的速度逼近最优解. 本文再次揭示了基于智能化方法的融合是未来解决大规模复杂性 NP 问题的重要方法之一.

References

- 1 Dorigo Marco, Gambardella Luca Maria. Ant colony system: A cooperative learning approach to the traveling salesman problem. *IEEE Transactions on Evolutionary Computation*, 1997, 1(1): 53~66
- 2 Wu Qing-Hong, Zhang Ji-Hui, Xu Xin-He. An ant colony algorithm with mutation features. *Journal of Computer Research & Development*, 1999, 36(10): 1240~1245(in Chinese)
- 3 Stutzle, Thomas, Hoos, Holger H. MAX-MIN ant system. *Future Generation Computer System*, 2000, 16(8): 889~914
- 4 Wu Bin, Shi Zhong-Zhi. An ant colony algorithm based partition algorithm for TSP. *Chinese Journal of Computers*, 2001, 24(12): 1328~1333(in Chinese)
- 5 Zhang Wen-Xiu, Liang Yi. Mathematical Foundation of Genetic Algorithms. Xi'an: Xi'an Jiantong University Press, 2000. 80~119(in Chinese)
- 6 Gutjahr W J. A graph-based ant system and its convergence. *Future Generation Computer System*, 2000, 16(8): 873~888

丁建立 博士、教授. 主要研究领域为智能仿生算法、智能决策、网络优化.

(DING Jian-Li Ph. D., professor. His research interests include simulating biology intelligent algorithms and intelligent decision and network optimization.)

陈增强 博士、教授、博士生导师、南开大学自动化系主任. 主要研究领域为智能预测控制、智能网络管理.

(CHEN Zeng-Qiang Ph. D., professor, he is the Director of the Department of Automation Nankai University. His research interests include intelligent forecast control and intelligent network management.)

袁著祉 教授、博士生导师、IEEE 高级会员. 主要研究领域为智能控制、自适应控制和智能决策系统.

(YUAN Zhu-Zhi Professor, IEEE Senior Member. His research interests include intelligent control, and adaptive control and intelligent decision system.)