

蚁群智能及其在大区域基础设施选址中的应用

何晋强, 黎 夏, 刘小平, 陶 嘉

中山大学 地理科学与规划学院, 广东 广州 510275

摘 要: 提出了基于蚁群智能的空间选址模型, 通过蚁群智能和 GIS 的结合来解决复杂的空间优化配置问题。这种启发式的智能搜索方法大大提高了空间搜索能力。为符合选址问题的求解, 从信息素更新方式和禁忌表调整策略两方面对基本蚁群算法进行改进。同时, 为了使得模型能实用于大区域的基础设施选址, 提出了“分步逼近”的策略, 取得了较好的效果。将所提出的模型应用于广州市公共设施的空间优化选址。实验结果表明, 该方法比简单搜索方法和遗传算法更有优势。

关键词: 蚁群算法, GIS, 选址, 简单搜索算法, 遗传算法

中图分类号: TP751.1

文献标识码: A

1 引 言

在资源环境的管理、规划和利用中, 经常需要面对如何在空间上配置这些资源, 以产生最大利用效益值的问题(Feng *et al.*, 1999)。譬如设施选址、土地利用规划、水资源的优化配置等。在 GIS 数据没有使用以前, 传统的优化搜索方法一般只用来解决数据量不大的问题(Church, 1999)。但当把 GIS 空间数据引入到空间优化决策模型中时, 所涉及的数据量非常大, 可能存在的配置方案将是一个十分庞大的数字。我们要想一一比较这些方案的优劣是不可能的, 也是不现实的。因此, 寻求一种有效的空间优化搜索方法具有较为重要的意义。

当需要配置的资源预先就是确定的, 并且数目较小, 多准则分析(multi criteria analysis, MCA)不失为一种有效的空间优化搜索方法(Carver, 1991; Nijkamp *et al.*, 1990; Janssen, 1991)。但 MCA 不适合大数据量的搜索, 并且具有较强的主观性。线性整数规划(linear integer programming, LIP)是目前较为常用的空间优化方法(Bettinger *et al.*, 1999; Crowe *et al.*, 2003), 该方法属于穷尽搜索(brute-force search), 速度较慢, 处理的数据量也有限, 并且是线性方法, 不能很好的处理地理复杂问题。最

近, 一些具有智能的启发式方法开始运用在选址模型中, 包括遗传算法(Brookes, 2001; Li, 2004)(genetic algorithm, GA)、模拟退火算法(Lockwood *et al.*, 1993; Aerts *et al.*, 2002)(simulated annealing, SA)、禁忌算法(Bettinger *et al.*, 1997; Brumelle *et al.*, 1998)(Tabu search)等。与穷尽搜索方法相比较, 启发式方法并不是求取最优解, 而是在启发式函数的引导下, 得到近似最优解, 不需要对所有的方案进行搜索, 因此其速度较快, 比较适合于大数据量的空间优化问题。

本文提出了基于蚁群智能的空间选址模型。蚁群智能算法(Ant colony optimization, ACO)是一种基于群体智能的仿生学优化算法, 由意大利学者 Dorigo 等于 1991 年提出(Coloni *et al.*, 1991), 其本质是一个复杂的多智能体系统, 由大量的简单智能体——蚂蚁(ants)所组成的团体, 通过相互合作能够有效地完成复杂任务, 例如寻找食物的最优路径。每个蚂蚁智能体根据路径上的信息作随机选择, 系统无中心控制, 但最终整个蚁群能够得到优化。这样的系统更具有鲁棒性(Robust), 不会由于一个或者某几个智能个体的故障而影响整个问题的求解, 该算法是群集智能的典型实现。本文尝试利用蚁群智能来解决空间决策中的优化配置问题。在空间选址中, 往往涉及大量的空间数据, 常规的穷尽搜索(brute-force search)方

收稿日期: 2007-11-07; 修订日期: 2008-04-14

基金项目: 国家杰出青年基金资助项目(编号: 40525002)、国家自然科学基金重点资助项目(编号: 40830532)、国家 863 计划资助项目(编号: 2006AA12Z206)。

第一作者简介: 何晋强(1982—), 男, 中山大学地理规划学院研究生, 主要从事地理信息模型研究, 已发表论文数篇。E-mail: hjq099@163.com。

法无法找到最优的解决方案,搜索的时间长,达不到实用的程度。蚁群智能作为一种启发式的智能方法,不仅能够智能搜索、全局优化,而且具有强鲁棒性、正反馈机制等特点(Dorigo *et al.*, 1996, 1997; Duan *et al.*, 2004)。对基本蚁群算法进行了改进,以符合选址问题的求解;同时,通过分步逼近(multi-scale)的优化策略,克服了蚁群算法在大范围区域内搜索时间过长的不足。改进后的分步逼近优化选址模型能大大缩短空间搜索时间,更加具有实用价值,为解决资源环境的空间决策问题提供了良好的手段。

2 基本蚁群算法

蚁群智能算法是受到对真实的蚁群行为的研究启发提出的。昆虫学家发现,虽然单个蚂蚁的行为极其简单,但它们所组成的蚁群群体却表现出极其复杂的行为。蚂蚁在寻找食物源时,能在其走过的路径上释放一种蚂蚁特有的分泌物——信息素,并以此指导自己的运动方向,使蚂蚁倾向

于朝着信息素浓度高的方向移动。因此,由大量蚂蚁形成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。蚂蚁这种选择路径的过程被称之为蚂蚁的自催化行为。由于其原理是一种正反馈机制,因此也可将蚂蚁王国理解成增强型学习系统。

下面以旅行商问题(TSP)为例来阐述其数学模型。 C 表示所有城市集合, n 为TSP规模也就是城市数量, m 为蚂蚁数目, $\tau_{ij}(k)$ 表示第 k 次迭代时路径 (i,j) 上的信息量,初始时刻各路径上的信息量相等,设 $\tau_{ij}(1) = \text{const}$,蚂蚁 $r(r = 1, 2, 3, \dots, m)$ 在运动过程中,根据各条路径上的信息量决定其转移方向。这里用禁忌表 tabu_r 来记录蚂蚁 r 当前所走过的城市,集合随着 tabu_r 的进化过程动态调整。在搜索过程中,蚂蚁根据各条路径上的信息量及路径上的启发信息来计算状态转移概率。 $p_{ij}^r(k)$ 表示在第 k 次迭代过程中蚂蚁 r 由城市 i 转移到城市 j 的状态转移概率

$$p_{ij}^r(k) = \begin{cases} \frac{\left[\tau_{ij}(k) \right]^\alpha \times \left[\eta_{ij}(k) \right]^\beta}{\sum_{s \in \text{allowed}_r} \left[\tau_{is}(k) \right]^\alpha \times \left[\eta_{is}(k) \right]^\beta}, & \text{若 } j \in \text{allowed}_r \\ 0 & \text{否则} \end{cases} \quad (1)$$

式中, $\text{allowed}_r = \{C - \text{tabu}_r\}$ 表示蚂蚁 r 下一次允许选择的城市; α 为信息素启发式因子。反映蚂蚁在运动过程中所积累的信息在蚂蚁运动时所起的作用,其值越大,则蚂蚁倾向于选择其他蚂蚁经过的路径; β 为期望启发式因子,反映蚂蚁在运动过程中启发信息在蚂蚁选择路径中的受重视程度; $\eta_{ij}(k)$ 为启发函数,其表达式如下:

$$\eta_{ij}(k) = \frac{1}{d_{ij}} \quad (2)$$

式中, d_{ij} 表示相邻两个城市之间的距离,对蚂蚁 r 而言, d_{ij} 越小,则越大, $p_{ij}^r(k)$ 就越大,蚂蚁倾向于选择与之距离较短的城市。

为了避免残留信息过多淹没启发信息,在每次迭代完成后,对残留信息进行更新处理。由此第 $k+1$ 次迭代时在路径 $\tau_{ij}(k)$ 的信息量按如下规则进行调整。

$$\tau_{ij}(k+1) = (1 - \rho) \times \tau_{ij}(k) + \Delta\tau_{ij}(k) \quad (3)$$

$$\Delta\tau_{ij}(k) = \sum_{r=1}^m \Delta\tau_{ij}^r(k) \quad (4)$$

式中, ρ 表示信息挥发系数,其取值范围为 $[0, 1]$; $\Delta\tau_{ij}(k)$ 表示本次迭代中路径 (i,j) 上的信息素增量, $\Delta\tau_{ij}^r(k)$ 表示第 r 只蚂蚁第 k 次迭代在路径 (i,j) 上留下的信息量

其中 $\Delta\tau_{ij}^r(k)$ 按下式进行计算:

$$\Delta\tau_{ij}^r(k) = \begin{cases} \frac{Q}{L_r}, & \text{若第 } r \text{ 只蚂蚁在本次循环中经过 } (i,j) \\ 0, & \text{否则} \end{cases} \quad (5)$$

式中, Q 表示信息素强度, L_r 表示第 r 只蚂蚁在本次循环中所走路径的总长度。可通过指定迭代次数或

者当迭代所得到的解不再发生变化作为算法停止条件。

3 基于蚁群智能的空间优化选址

选址研究中的典型问题,如 p -median 问题、覆盖问题、 p -center 问题是现在选址研究的热点 (Hansen *et al.*, 1997; Kariv *et al.*, 1979; Toregas *et al.*, 1971), 大多数研究集中在优化算法的设计和改进行上。本文也是属于这类的研究,选择 p -median 问题作为基于蚁群智能的空间优化选址研究案例。 p -median 问题是选定 p 个公共设施的位置,使多数或平均性能最优的问题。通常是使成本最小,如使总的运输距离最小、总的运输时间最少,或者使总运输费用最小等,故又称为最小和问题。在本文中是指在 $N \times M$ 的栅格空间中确定 p 个目标栅格的位置 (如医院、图书馆等公共设施),使得其余的栅格到距其最近的目标栅格的距离之和最小。

在 TSP 中,基本蚁群算法的目的是确定城市的连接顺序,即各城市如何连接,使得连接所有城市的路径之和最短,且所有的城市是最终结果的组成部分;而在 p -median 选址问题中,其目标是从 $N \times M$ 个栅格中确定 p 个目标栅格,使得费用之和最小,所选出的目标栅格只是候选栅格中极少的一部分。因此,根据 p -median 选址问题的上述特征,对基本蚁群算法进行改进,使其符合问题的求解方式是十分必要的。

3.1 定义目标函数

目标函数用于判定算法得到的结果的好坏,在 TSP 中,目标函数就是连接所有城市的路径之和;而 p -median 选址问题中,以人们出行时总的交通时间或费用为准则。本文以所选取的公共设施与服务对象间的费用和最小为目的,目标函数如下:

$$F_1 = \sum_{i=1}^M \sum_{j=1}^N d_{\min}(i, j) \times p_{\text{den}}(i, j) \times A \quad (6)$$

式中, M 表示实验区域的宽度,即栅格的行数; N 表示实验区域的长度,即栅格的列数; $d_{\min}(i, j)$ 表示栅格 (i, j) 到与其最近的目标栅格的欧几里德距离; $p_{\text{den}}(i, j)$ 表示栅格 (i, j) 的人口密度; A 表示每个栅格在实际中对应的面积。在真实的选址应用中,通过获取实验区域内任意两栅格的真实路径长度,用其替换两栅格间的欧几里德距离即可。也可以定义其他的目标函数,只要对公式 (6) 作相应调整即可,这种调整和改变在蚁群算法中实施起来相当方便。

进行空间优化决策时,针对不同的应用目的,可以定义不同的目标函数。在上面的目标函数中,以总的出行交通时间或费用为标准,用欧几里德距离代替真实距离,而实际情况中,总费用与交通条件密切相关,对上面目标函数进行修改,加入一个道路状态因子。修改后公式如下:

$$F_2 = w_1 \times F_1 + w_2 \times X_{\text{road}} = w_1 \times \sum_{i=1}^M \sum_{j=1}^N d_{\min}(i, j) \times p_{\text{den}}(i, j) \times A + w_2 \times \sum_{p=1}^K e^{-\rho \times D_{\text{road}}^p} \quad (7)$$

式中, w_1 和 w_2 表示影响因子的权重, K 为要选取的目标栅格个数, ρ 为道路系数, D_{road}^p 为第 p 个目标栅格的交通状况,即该目标栅格到主干道的距离。 $1/F_2$ 能够被当作选址结果的适应度值,后面的对比实验中,取目标函数的倒数 $1/F_2$ 适应度值来进行比较。

3.2 定义启发函数

信息素是表征过去信息的载体,而启发函数是表征未来信息的载体,它们直接影响到算法的全局收敛性和求解效率。实验结果表明,启发函数对保证蚁群算法在合理的时间内搜索到全局最优解很重要。按照常理,目标函数也表明,选择那些人口密度大的栅格作为目标栅格趋向于使总的费用或时间最小。

$$\eta_{ij}(k) = p_{\text{den}}(i, j) \quad (8)$$

3.3 信息素更新策略

与 TSP 不同的是, p -median 选址问题是从大量候选栅格中选出为数较少的目标栅格,如按 TSP 中的方式更新信息素,只有极少数被蚂蚁选中栅格的信息素在挥发后会增加,这样蚂蚁在路径选择过程中的正反馈机制被大大的削弱甚至被淹没,不利于蚂蚁搜索最优解;另外由于相邻栅格具有较强的依赖关系,被选中的目标栅格附近的栅格有可能比目标栅格本身更优。

基于以上两点,本文采取信息素递减扩散的策略对信息素进行更新。具体实现如下:如果某栅格被选为目标栅格,则在信息素更新时,以目标栅格为中心,以指定的长度为边长 (如 5×5) 的正方形小区域内,按信息素增量从中心往四周递减的策略更新信息素,正方形小区域的边长根据实验区域的大小和选择的目标栅格的数目做相应的调整,如区域越大,边长可相应调大,选择的目标栅格越少,边长可相应调大。公式如下:

$$\Delta\tau_{ij}^r(k) = \begin{cases} \frac{Q}{(d_{\text{centre}}^p(i,j) + 1) \times F_2}, & \text{若栅格}(i,j) \text{ 在目标栅格 } p \text{ 所在的小区域内} \\ 0, & \text{否则} \end{cases} \quad (9)$$

式中, Q 为信息素强度, $d_{\text{centre}}^p(i,j)$ 表示在目标栅格对应的小区域内栅格与中心栅格 p (目标栅格) 的欧几里德距离, F_2 为目标函数的值。

蚁群算法在求解初期具有全局搜索和快速收敛能力, 可当得到的解接近最优解时, 这种随机搜索算法求解速度变慢, 收敛困难, 且容易出现停滞现象 (Colorni *et al.*, 1991; Dorigo *et al.*, 1996)。为加快收敛速度, 找到更优的解, 在算法的求解后期或者当算法多次迭代后所得的解仍未改进时, 我们采取以下策略加快求解和收敛速度: 除按上面提到的信息素更新策略更新信息素外, 为充分利用已找到的最优解, 额外增加已找到的最优解所在栅格的信息素, 加强已找到的最好解对后续求解过程的正反馈作用。

3.4 禁忌表调整策略

在 TSP 中, 禁忌表用来记录蚂蚁已走过的城市, 相应的在 p -median 选址问题中, 禁忌表用来记录已被蚂蚁选中的目标栅格。在实际应用中, 可通过禁忌表来表达选址过程中的限制因素, 比如某些栅格由于特定因素而不能被选作目标栅格, 只要在初始化时, 将这些栅格置于禁忌表中即可。

根据信息素更新策略, 蚂蚁所选择的目标栅格附近的栅格都赋予了较高的信息素。为了使信息素更新策略发挥促进作用, 即防止后面的蚂蚁在搜索时, 选择前面的蚂蚁所选择的目标栅格对应的小区域内的两个或两个以上栅格作为目标栅格, 而相距很近的栅格同时作为目标栅格与最优解互为矛盾, 本文采取与信息素更新策略相类似的禁忌表调整策略, 即蚂蚁在搜索过程中, 选中一个目标栅格后, 以目标栅格为中心的小的正方形小区域内的栅格与目标栅格一起并入到禁忌表中, 从而防止蚂蚁选择相距很近的栅格同时作为目标栅格, 正方形小区域的长度调整也与上面信息素更新策略一致。

3.5 算法描述

以下为执行整个蚁群选址过程的算法描述:

Pheromone = Constant /* 所有栅格上的信息素初始化的值 */

Do while ($p < 500$ and $M < 5$) /* p 为迭代次数, M 为一次迭代过程中找到的解连续相同的次数 */

For $K = 1$ to num_of_ants /* 蚂蚁数 */

Tabu = NULL /* 初始化禁忌表 */
For $N = 1$ to num_of_sites /* 目标栅格数 */
Select a site /* 按概率选择一个目标栅格 */
Tabu update /* 更新禁忌表 */
Next N
If (The result excels the best result already found)
Replace the best result with the existing result
End if
If (The result is the same as the previous one) /* 检查收敛情况 */
then
 $M = M + 1$
Else
 $M = 0$
End if
Next K
Pheromone update /* 根据选址的结果更新信息素 */
 $p = p + 1$
Loop

3.6 基于“分步逼近”策略的蚁群智能的优化选址

为了使得模型能应用于大区域的基础设施选址, 必须在蚁群智能中采用“分步逼近”的策略, 以缩短空间优化搜索所需的时间。具体实现如下: 以 50×50 的实验区域为例, 假如在其中选择 6 个目标栅格, 每个栅格有其对应的人口密度, 从实验区域的左上角开始, 按 5×5 的窗口合并原有栅格, 把原有的 50×50 的区域缩减为 10×10 的缩减空间, 新的栅格的人口密度为原有 5×5 窗口区域的平均人口密度值, 即进行尺度转换操作。利用蚁群智能方法在转换后的空间上进行优化搜索, 得到 6 个目标栅格。再把缩减空间还原为原来尺度 (50×50) 的空间。初步搜索到的目标栅格还原到对应原尺度空间的中心栅格, 即得到原尺度空间的 6 个目标栅格。一般而言, 还原后的目标栅格在原尺度空间内并不是最优解, 可能只是近似最优解。此时, 我们再运行一次蚁群算法, 这次的搜索与前一次有所不同, 前面的初步选址是在整个区域内进行搜索, 而这次是在已得到一个近似最优解的前提下运行算法。也就是说每个目标栅格的大体位置

已定,还原到原始栅格数据层运行算法时,只是在还原后每个目标栅格的附近进行搜索,即在以目标栅格中心的 5×5 的栅格区域内进行搜索,为了防止上面提到的不一致情况出现,可适当扩大目标栅格中心的

搜索区域,如调整为 7×7 ;另外与第一次运行蚁群算法不同的是,此次运行蚁群算法不需要调整禁忌表,这是由于在每个搜索空间只选择一个目标栅格。蚁群分步优化选址的过程如图 1。

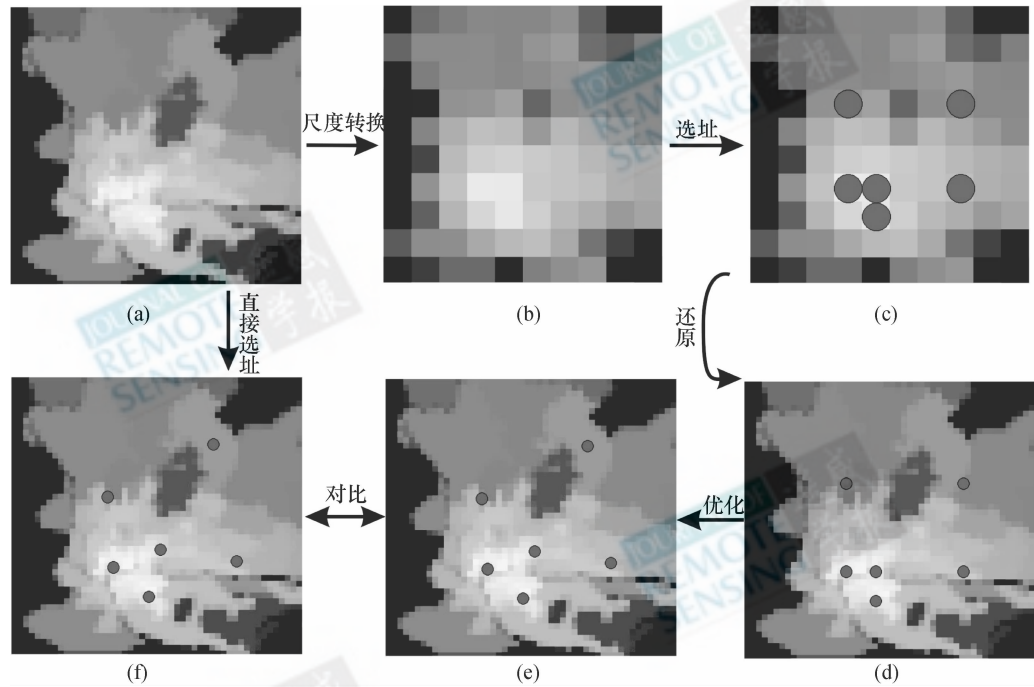


图 1 基于“分步逼近”的蚁群算法的优化选址示意图

(a)原始数据层;(b)尺度转换后数据层;(c)尺度转换后数据层上选址;
(d)原始数据层上对应的选址;(e)进一步优化后的选址;(f)原始数据层上直接选址

Fig. 1 The procedure of multi-scale site selection using ACO

(a) original resolution layer; (b) coarser resolution layer;
(c) results on coarser resolution layer; (d) initial results on finer resolution layer;
(e) optimization result on finer resolution layer; (f) direct site selection results on original layer

为了验证上面所说分步选址的有效性,即通过分步选址能否找到最优解。我们从广州市截取了 50×50 的区域进行对比实验,当目标栅格数目从 1 到 16 个时,用本文提出的分步蚁群算法和未分步蚁群算法所找到的最优解进行对比,也就是对图 1 中相应的 (e) 图和 (f) 图所得到的选址结果的目标函数值进行比较,对比结果见表 1。

从表 1 中可知,当目标栅格数个数较少时,分步算法与未分步算法所找到的选址结果相同;当目标栅格数目较多时,两种方式所找到的选址结果的目标函数值很相近,差距非常小,表 1 中的最大比值差距在千分之二内。此对比实验说明我们上面的分步方式是可行和有效的,通过分步方式,蚁群算法能找到最优解或近似最优解。

表 1 “分步逼近”蚁群算法选址和直接蚁群算法选址适应度值 ($1/F_2$) 对比

Table 1 Comparison of fitness values between multi-scale site selection and single-scale site selection using ACO

目标数量	分步选址/ $\times 10^{-6}$	直接选址/ $\times 10^{-6}$	分步选址/ 直接选址
1	0.609	0.609	1.0000
2	0.786	0.786	1.0000
3	1.028	1.028	1.0000
4	1.176	1.176	1.0000
5	1.310	1.310	1.0000
6	1.452	1.452	1.0003
7	1.567	1.567	0.9998
8	1.666	1.666	1.0000
9	1.772	1.771	1.0006
10	1.858	1.860	0.9989
11	1.967	1.967	1.0001
12	2.064	2.061	1.0013
13	2.149	2.148	1.0007
14	2.230	2.230	0.9997
15	2.322	2.323	0.9996
16	2.400	2.400	0.9998

4 实验结果及分析

4.1 研究区域

本文的实验区域为广州市城区、包括珠海区、越秀区、荔湾区的全部和天河区及白云区的人口密集区。以公共设施的空间选址为例,检验蚁群智能算法在基础设施的空间优化布局方面的有效性。人口数据来源为2003年广州人口普查数据,以街区为统计单位;交通数据为2003年该区域内交通主干道分布数据。

4.2 利用 GIS 获取空间数据

GIS 提供了空间优化模型所需的基本空间数据,根据人口普查数据和各街区边界的矢量空间数据,可以很方便地计算出每个街区内的平均人口密度。在模型运行前,需要将矢量数据转为栅格数据,栅格的分辨率为 100m × 100m,共有 250 × 250 个栅格(图 2),交通状况数据获取方式如下:先得到广州市交通主干道分布数据,然后计算研究区域到主干道的欧几里德距离,得到交通状况栅格图(图 3),图像分辨率及大小与人口数据一致。由此可见,处理所涉及的数据量相当大,简单的穷尽搜索方法很难在较短时间内获得最优解。

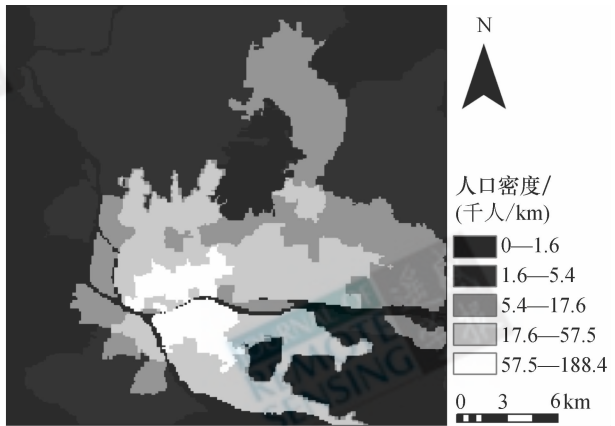


图 2 广州市人口密度

Fig. 2 The population density in district of Guangzhou

4.3 实验结果的对比

本文程序用 VB6.0 实现,在 P4 2.0G 处理器上运行。我们用简单搜索算法、遗传算法与本文的算法进行对比,按公式(7)中定义目标函数做对比实验,同时考虑人口的分布和交通状况两个因素对选址的影响。

在图 4 所示数据的对比实验中,蚁群选址模型的主要参数选择如下:信息素启发因子 α 为 1.5、期望启发因子 β 为 0.3、信息素挥发系数 ρ 为 0.1、蚂蚁数为 5、信息素强度 Q 为 1、信息素和禁忌表边长都为 5。由于蚁群算法涉及的参数较多,且参数关联性较强,如何确定各参数的最优组合目前没有完善的理论根据可寻,一般根据经验而定。在本文的选址模型中,比如蚂蚁数选择多少,与实验区域大小、选址目标个数、信息素挥发系数 ρ 有较大的关联性;实验表明,当实验区域越大,蚂蚁数更多;当目标个数越多,蚂蚁数更少,当信息挥发指数越大,蚂蚁数更多。由于蚁群算法的并行性,较多的蚂蚁数可提高算法的全局搜索能力和算法的稳定性,但过多的蚂蚁使候选点的信息素平均化,正反馈作用减弱,实验表明,蚂蚁过多时,选址模型的求解能力下降;蚂蚁数过少,蚂蚁搜索随机性减弱,收敛加快,容易陷入局部最优,出现早熟现象。关于蚂蚁对实验结果的影响,对以下对比实验,单次迭代过程中,蚂蚁数分别取 1,3,5,7,9,运算 10 次取适应度平均值,如图 4,可以看出,当蚂蚁数在 5—7 之间,选址结果较好。

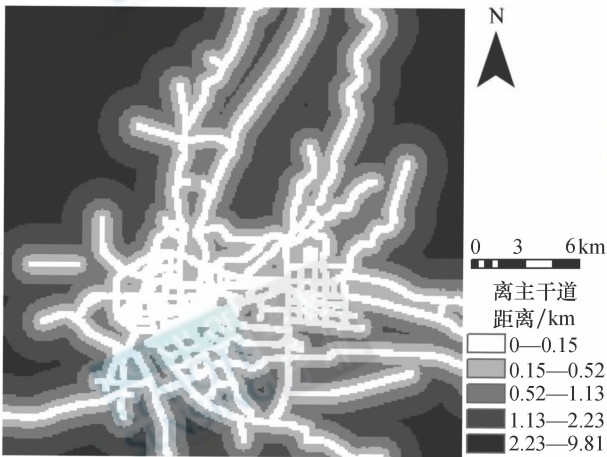


图 3 广州市交通状况

Fig. 3 Proximity to major road in Guangzhou

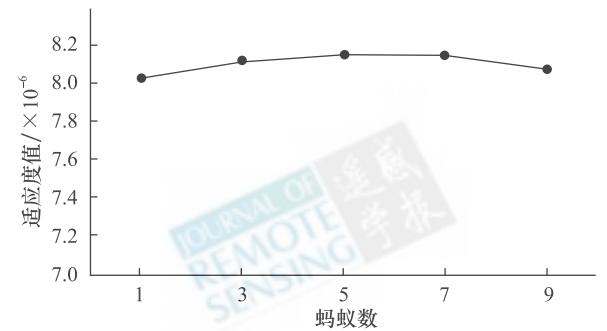


图 4 蚂蚁数与实验结果关系(10 个目标)

Fig. 4 The relationship between the number of ants one iteration and experiment result

根据黎夏等在 2004 年用 GA 做的选址研究, GA 的参数选择如下:初始个体数为 200,变异概率为 0.01,交叉概率为 0.9。简单搜索算法(SSA)的基本思想是,如果在实验的栅格区域($M \times N$)内选择的目标栅格数目为 K 个,该算法先在实验栅格区域内遍历选择使得总费用最小的第一个目标栅格,然后在第一个栅格的基础上选择第二个目标栅格,以此类推,直到选出第 K 个目标栅格。如按在上面的区域运行算法, $M = 250, N = 250$,设 $K = 10$,一次简单搜索算法要计算目标函数 625000 次。经多次实验,运行一次目标函数计算所需的平均时间为 0.48s,这样完成整个搜索过程需要约 83h,相当费时。所以我们将原 250×250 实验数据区域缩小 5 倍得到 50×50 区域进行对比实验。所选的区域见图 5,分别用简单搜索算法、遗传算法和蚁群算法对所选择的区域进行空间优化搜索运算,运算结果见表 2、表 3 及图 6。

表 2 简单搜索算法、遗传算法与蚁群算法适应度值 ($1/F_2$ 对比)

Table 2 Comparison of the improvement of fitness values between the SSA,GA and ACO in site selection					
目标栅格数	2	4	6	8	10
SSA/ $\times 10^{-6}$	3.7	5.0	6.0	6.9	7.7
GA/ $\times 10^{-6}$	3.9	5.2	6.2	7.1	8.0
ACO/ $\times 10^{-6}$	3.9	5.3	6.3	7.2	8.2
(ACO-SSA)/SSA/%	5.3	5.5	2.7	3.9	2.8
(ACO-GA)/GA/%	0.0	0.4	1.3	1.2	2.5

表 3 简单搜索算法、遗传算法与蚁群算法时间对比

Table 3 Comparison of computing time between SSA and GA and ACO

目标栅格数	2	4	6	8	10
SSA/s	32	95	198	333	525
GA/s	45	138	247	419	832
ACO/s	4	28	41	54	73
ACO/SSA/%	12.5	29.5	20.7	16.2	13.9
ACO/GA/%	8.9	20.3	16.6	12.9	8.8

从表 2 看出,在所选择的对比实验区域内,蚁群算法相对于简单搜索算法在不同目标栅格点时有 3.9%—5.6% 的提高,并且找到的结果每次比遗

传算法好,该蚁群算法能够获得相对于简单搜索算法和遗传算法更好的选址结果;从表 3 中可知,蚁群算法在计算时间上具有更大的优势,在所选择的对比实验区域内,蚁群算法运行所花费的时间只有不到简单搜索算法的 1/4,比遗传算法花费的时间更少,并随着目标栅格的数目增加,时间优势更加明显。另外,当搜索区域扩大时,本文提出的“分步逼近”蚁群算法在时间上的优势更加明显,下面的实验将会证实这一点。图 7 进一步比较 ACO 和 GA 的收敛速度,两者具有较快的收敛速度,但 ACO 能找到更好的解。

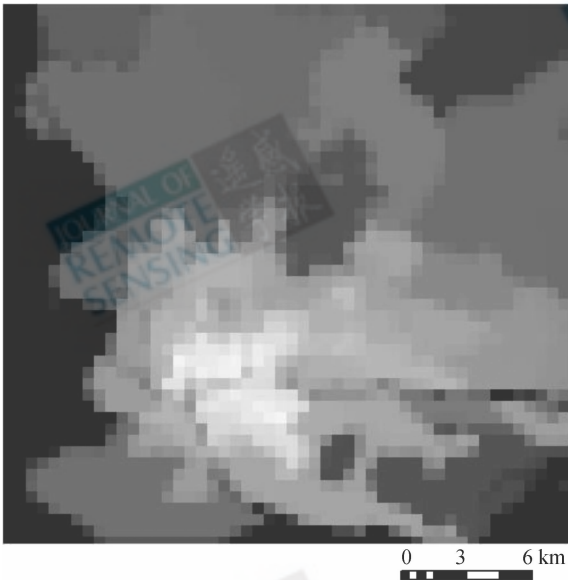


图 5 广州市小尺度人口密度

Fig. 5 The population density of the study area in Guangzhou

从上面的对比实验中可看出蚁群算法相对简单搜索算法和遗传算法能找到更优的选址结果,而且在时间上具有更大的优势。下面我们在原始数据层上,对几种算法所得的运行时间和选址结果进行对比,由于原始数据为 250×250 ,我们加入 3.5 节中的分步逼近策略。正如前面所提到的,在运行简单搜索算法过程中,主要是运行公式(7)所示的目标函数,运行一次目标函数的时间与所实验区域的总栅格数的平方成正比,经过估算,要用运行简单搜索算法在广州市选出 10 个目标栅格将花费近 83 h。将“分步逼近”蚁群算法和蚁群算法、遗传算法、简单搜索算法进行了比较。由于区域比较大,计算费时,我们只对 10 个栅格目标下的选址时间和所得解的优劣进行比较。对比结果如表 4。

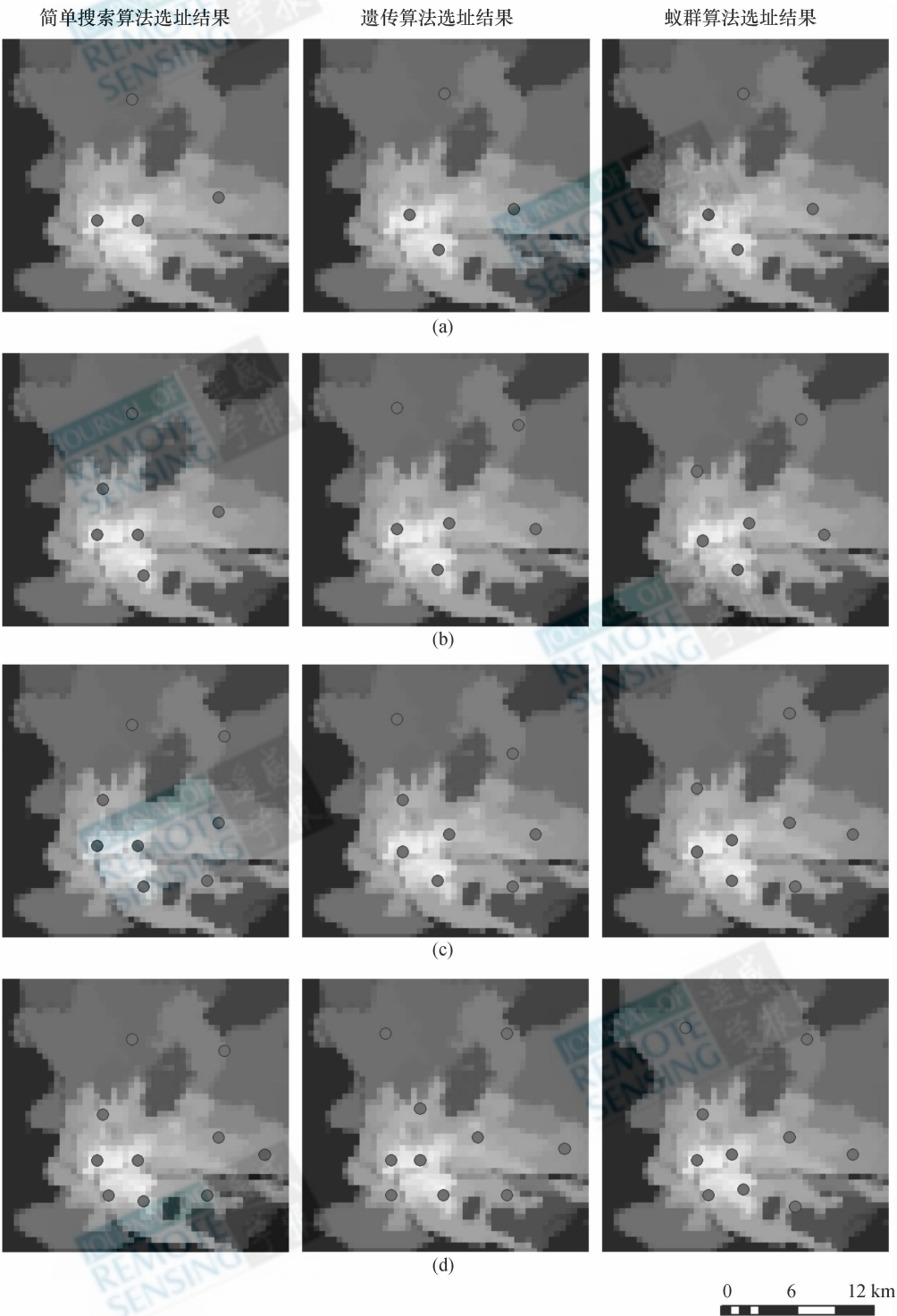


图6 简单搜索算法、遗传算法和蚁群算法选址结果对比图

(a)4 个目标;(b)6 个目标;(c)8 个目标;(d)10 个目标

Fig.6 Site selection for the facility from the SSA, GA and ACO

(a) four targets;(b) six targets;(c) eight targets;(d) ten targets

表 4 目标数量为 10 时简单搜索算法、遗传算法和蚁群算法运行时间及结果对比

Table 4 Comparison of computing time for the SSA, GA, single ACO and multi-scale ACO of ten targets		
算法	计算时间/h	适应度值/ $\times 10^{-6}$
SSA	82.5	7.7
GA	5.2	8.0
single ACO	2.1	8.2
multi-scale ACO	0.5	8.2

从上表看出,“分步逼近”蚁群算法所得到的选址结果优于简单搜索算法 5.5%, 优于遗传算法 2.5%; 所运行的时间明显优于简单搜索算法和遗传算法,进一步体现本文算法在时间上所具有的优势。图 8 是运行各种算法在广州市进行空间优化

选址的结果对比图,图 9 为目标数量为 20 和 30 时,运行分步逼近算法蚁群算法的选址结果图。

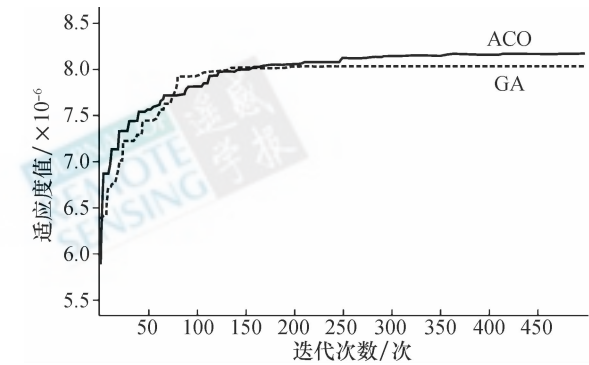


图 7 ACO 与 GA 适应度值收敛曲线图(10 个目标)
Fig.7 Convergence rates of the GA and the ACO of ten targets

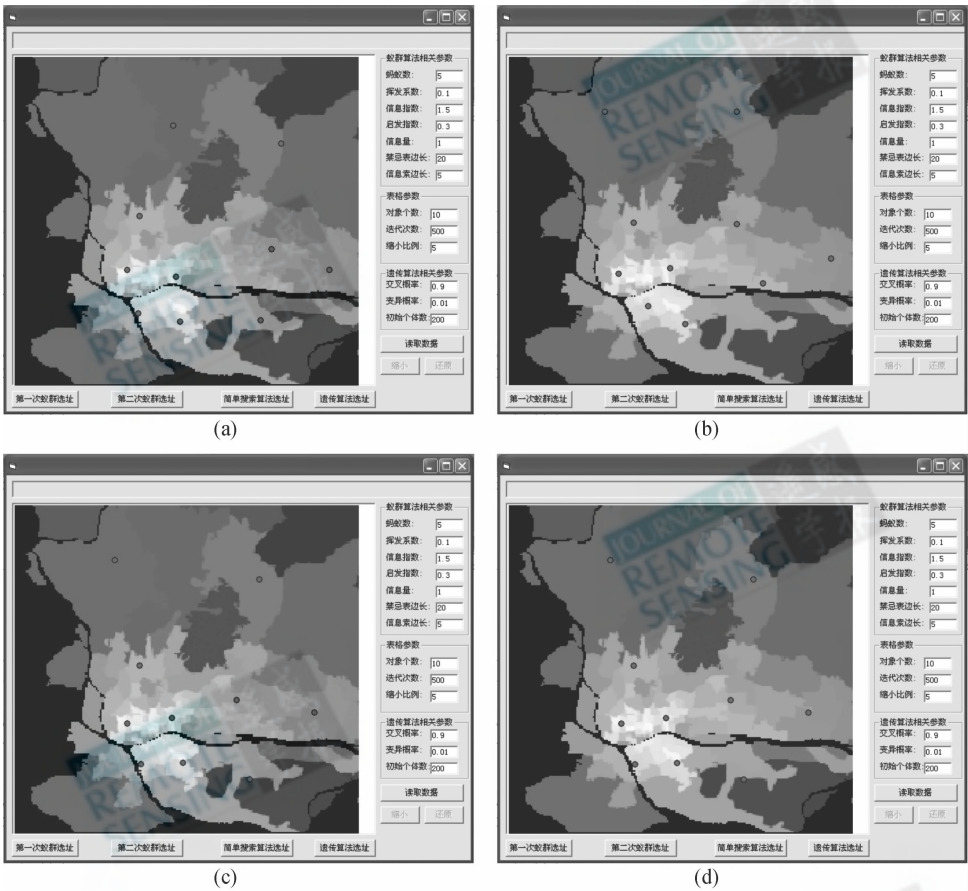


图 8 目标数量为 10 的几种算法的选址结果对比图
(a) 简单搜索算法选址;(b) 遗传算法选址;(c) 蚁群选址;(d) “分步逼近”蚁群选址
Fig.8 Site selection for the facility from the SSA, GA, ACO and multi-scale ACO
(a) results of simple search algorithms;(b) results of genetic algorithms;
(c) results of ant colony optimization;(d) results of multi-scale ant colony optimization

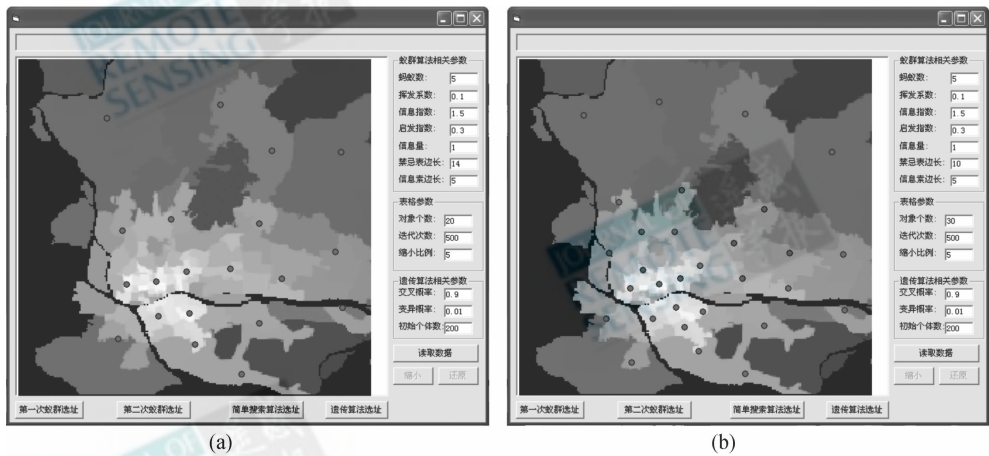


图9 目标数量为20时(a)和30时(b)“分布逼近”蚁群算法选址结果

Fig.9 Site selection for 20(a) and 30(b) targets of the facility from multi-scale ACO

5 结 论

资源和环境的空间决策涉及复杂的非线性关系,以及处理高维和海量的空间数据,常用的穷尽搜索方法在有限的时间内往往难以求取问题的最优解。与穷尽搜索方法相比较,蚁群算法在优化过程中不受函数约束条件的限制,只依赖优化目标的函数值,能够求解非线性、高维的优化问题。并且,蚁群算法是一种启发式的智能算法,不需要对所有的方案进行搜索,因此其速度较快,比较适合于大数据量的空间优化问题。

本文提出了将蚁群智能算法与GIS结合来解决空间优化问题的新方法。从信息素更新方式和禁忌表调整策略两方面对基本蚁群算法进行改进,以符合选址问题的求解。为了克服在大范围内多目标选址运行时间过长的不足,本文采用“分步逼近”的策略进行空间搜索。首先,对搜索区域进行尺度转换,利用蚁群算法在转换后的低分辨率的空间中搜索出初步的最优目标栅格。再把低分辨率的空间还原为原来尺度的空间。随后,在原来尺度空间中利用蚁群算法对初步选出的目标栅格再进行一次更精确的优化搜索,以获取最终的目标栅格。改进后的算法在保证可以找到最优解或近似最优解的情况下大大缩短了运行时间。

对广州市公共设施进行了空间选址的实验,实验结果表明,蚁群分步选址算法不但能找到比简单搜索算法较优的解,而且在运行时间上优势明显,特别在大范围,多目标选址情况下,其相对于简单搜索算法和遗传算法在时间上的优势更加突出。

REFERENCES

Aerts C J H and Heuvelink G B M. 2002. Using simulated annealing for resource allocation. *International Journal of Geographical Information Science*, **16**(6): 571—587

Bettinger P, Boston K and Sessions J. 1999. Intensifying a heuristic forest harvest scheduling search procedure with 2-opt decision choices. *Can. J. For. Res.*, **29**: 1784—1792

Bettinger P, Sessions J and Boston K. 1997. Using tabu search to schedule timber harvests subject to spatial wildlife goals for big game. *Ecol. Model.*, **94**: 111—123

Brookes C J. 2001. A genetic algorithm for designing optimal patch configurations in GIS. *International Journal of Geographical Information Science*, **15**(6): 539—559

Brumelle S, Granot D and Halme M. 1998. A tabu search algorithm for finding good forest harvest schedules satisfying green-up constraints. *Eur. J. Oper. Res.*, **106**: 408—424

Carver S J. 1991. Integrating multi-criteria evaluation with geographical information systems. *International Journal of Geographical Information Systems*, **5**(3): 321—339

Church and R L. 1999. Location modeling and GIS. P A Longley, M F Goodchild, D J Maguire, *et al. Geographical Information Systems: Volume 1*. New York: John Wiley & Sons, Inc

Colomi A, Dorigo M and Maniezzo V. 1991. Distributed optimization by ant colonies. *Proceedings of the 1st European Conference on Artificial Life*

Crowe K, Nelson J D and Boyland M. 2003. Solving the area-restricted harvest-scheduling model using the branch and bound algorithm. *Can. J. For. Res.*, **33**: 1804—1814

Dorigo M and Gambardella L M. 1997. Ant colony system: a cooperative learning approach to the traveling salesman problem. *IEEE Transaction on Evolutionary Computation*, **1**(1): 53—56

Dorigo M, Maniezzo V and Colomi A. 1996. Ant system; optimization by a colony of cooperating agents. *IEEE Transactions on Systems*,

- Man, and Cybernetics-Part B, **26**(1): 29—41
- Duan H B, Wang D B and Zhu J Q. 2004. Development on ant colony algorithm theory and its application. *Control and Decision*, **19**(12): 1321—1326
- Feng C M and Lin J J. 1999. Using a genetic algorithm to generate alternative sketch maps for urban planning. *Computers, Environment and Urban Systems*, **23**(2): 91—108
- Hansen P and Mladenovic N. 1997. Variable neighborhood search for the p-median. *Location Science*, **5**: 207—226.
- Janssen R. 1991. Multiobjective Decision Support for Environmental Problems. Free University, Amsterdam
- Kariv O and Hakimi S L. 1979. An algorithmic approach to network location problems, Part 1: The p-centers. *SIAM Journal Applied Mathematics*, **37**(3): 513—538
- Li X and Yeh A G O. 2004. Optimal spatial search using genetic algorithms and GIS. *Acta Geographic Sinica*, **59**(5): 745—753
- Lockwood C and Moore T. 1993. Harvest scheduling with spatial constraints: a simulated annealing approach. *Can. J. For. Res.*, **23**: 468—478
- Nijkamp P, Rietveld P and Voogd H. 1990. Multi Criteria Evaluation in Physical Planning. Amsterdam: North Holland Publishers
- Toregas C, Swain R and ReVelle C. 1971. The location of emergency service facilities. *Operations Research*, **19**: 1363—1373

附中文参考文献

- 段海滨,王道波,朱家强等. 2004. 蚁群算法理论及应用研究进展. 控制与决策, **19**(12): 1321—1326
- 黎夏,叶嘉安. 2004. 遗传算法与 GIS 结合进行空间优化决策. 地理学报. **59**(5): 745—753

Ant colony algorithms for optimal site selection in large regions

HE Jin-qiang, LI Xia, LIU Xiao-ping, TAO Jia

School of geography and planning, Sun Yat-sen University, Guangdong Guangzhou 510275, China

Abstract: Optimal site search for sitting facilities is crucial for the effective use and management of resources and it is also a common task for urban planning. The brute-force method has difficulty in solving complex site search problems especially in large scale areas. In this study, a location model is proposed based on ant colony algorithms (ACO). It combines ant colony intelligence and GIS to solve the problems of complicated spatial optimal allocation. ACO has strong search ability for a huge volume of spatial data.

At first, the algorithm is modified about the strategy of pheromone update and Tabu table adjusting to fit the sites location problem. Spatial allocation problems usually have a large set of spatial data and only a few targets. The pheromone evaporates very fast because the selected cells only amount to a small percentage of all the cells. The positive feedback is too weak to play a role in the optimization. A modification is to incorporate the strategy of neighborhood pheromone diffusion in defining pheromone updating. At the same time, an optimal result for sites selection usually does not include two near candidate cells together, so a restricted Tabu table updating strategy is adopted which resembles the strategy of neighborhood pheromone diffusion. Then another important modification is to adopt a multi-scale approach to alleviate the computational demand in conducting large-scale spatial search. This includes two phases of optimization. First, a coarser resolution is used for the identification of rough locations of targets using ACO. Then the next round of optimization is implemented by just searching the neighborhood around these initial locations in the original resolution. This two-phase procedure of optimization can thus significantly reduce the computation time.

The study area is located in the city of Guangzhou. This optimization problem considers two spatial variables, population distribution and transportation conditions, which are obtained from GIS. The raster layers have a resolution of 100×100 m with a size of 250×250 pixels. A comparison experiment is conducted among the multi-scale ACO, ACO, genetic algorithms and simple search algorithms. Experiments indicate that this multi-scale ACO method can produce similar results but use much lesser computation time, compared with the single ACO method. This method has better performance than the single search method and the genetic algorithm method for solving site search problems. ACO has obtained the utility improvement of 2.7%—5.5%, compared with the single search method. ACO has slight improvement of the total utility value over the GA method, but is able to reduce computation time significantly. Its computation time is only 12.5%—29.5% and 8.8%—20.3% of those of the single search method and the GA method respectively.

In conclusion, integrating ACO with GIS is important for solving real-world site selection problems. The integration allows these two techniques to be mutually benefited from each other. ACO provides an efficient distributed computation algorithm while GIS provides useful spatial information. The comparison experiment indicates that the proposed model has better performance than GA and SSA especially in the computing time for solving site search problems.

Key words: ant colony optimization, GIS, site selection, simple search algorithm, genetic algorithms