

一种启发式 A* 算法和网格划分的空间可达性计算方法

马林兵, 曹小曙*

(中山大学地理科学与规划学院, 广州 510275)

摘要: 本文提出了一个适用于研究城市内部的个体或商业区位的微观可达性计算方法, 该方法的核心是将研究区域进行等距的网格划分, 通过计算每个网格的可达性指标, 来研究整个区域的可达性的空间分布特点。在可达性计算中, 利用网格内的道路密度和土地利用状态这两个因素来模拟计算每个网格的交通成本, 引入了启发式 A* 空间搜索算法来计算网格间路径的交通成本, 并且加入适当的启发信息, 提高了搜索效率, 使搜索结果更符合实际需求。最后, 基于本文提出的方法, 利用 GIS 二次开发工具 ArcEngine 开发了计算程序, 收集了多源数据, 以广州市商务区的可达性作为计算对象, 进行了商务区的可达性和易达性案例计算。

关键词: 可达性; 启发式 A* 算法; 城市交通

文章编号: 1000-0585(2008)01-0093-07

前言

空间可达性 (Accessibility) 是衡量一个城市的空间结构、交通状况、土地利用合理性的重要指标, 在规划领域得到广泛的应用。可达性研究得到了城市规划、交通地理以及从事区域研究的众多学者的深入进行^[1]。可达性的大小受到三个方面的因素影响^[2]: 土地使用因素、交通因素和个体因素。其中土地使用因素包括土地的使用性质 (如居住、工作、购物、文化娱乐等) 和土地使用的空间分布 (如位置、规模、强度等); 交通因素包括交通需求、交通供给 (含交通工具因素)、出行距离及时间等; 个体因素包括个人需要、个人能力及个体其他属性等。如今, 可达性所应用的研究范围非常广泛, 关于可达性的定义仍没有一个比较精确的定义^[3], 一般来说, 可达性是指利用一种特定的交通系统从某一给定区位到达活动地点的便利程度。

可达性的计算方法有很多种, Karst T. Geurs^[4] 在其综述中, 将可达性评价计算方法分为四类: 1) 基于基本交通设施的方法 (Infrastructure-based measures); 2) 基于区域位置的方法 (Location-based measures); 3) 基于个人行为的方法 (Person-based measures); 4) 基于收益效用的方法 (Utility-based measures)。这些方法应用于不同的研究目的和空间规模。在可达性的评价、测算和建模过程中, 通常要依赖于空间的简化, 绝大多数研究

收稿日期: 2007-08-09; 修订日期: 2007-11-24

基金项目: 国家自然科学基金项目 (40571052); “985 工程” GIS 与遥感的地学应用科技创新平台资助项目 (105203200400006); 教育部地理信息系统重点实验室<武汉大学>开放基金 (wd200604)

作者简介: 马林兵 (1968-), 博士, 讲师。现主要从事移动对象空间数据库、地理信息系统相关的理论及其在城市规划、交通中的应用研究。E-mail: malb@mail.sysu.edu.cn

* 通讯作者: 曹小曙 (1970-), 男, 甘肃人, 博士, 副教授, 博士生导师。主要从事交通地理与城市地理研究。E-mail: Caoxsh@mail.sysu.edu.cn

将不同尺度的研究区域抽象为点状区域,并将它们作为起点和终点来评价各节点的可达性^[1,5]。这种把某一小区的影响可达性的综合因素人为归结到某一节点(通常为交通节点)进行计算的方法,主要有两个缺点:其一,是计算单元(小区)的划分无标准方法,采用不同的方案划分小区,小区的尺寸和形状不一样,计算所得的可达性值也会不一样^[2];其二,以计算所得到的节点的可达性值代替整个小区的可达性值,空间的精度不够。

为了避免空间简化抽象产生的问题,本文提出一种基于网格划分的可达性计算方法,其核心思想是将研究区域划分成等距的网格来进行可达性计算,并引入了A*启发式搜索算法的思想,作为本方法的核心算法。本文提出的可达性计算方法具有一定的通用性,作为城市地理学定量研究的手段,特别适用于研究城市内部的个体或商业区位的微观可达性评价。

1 A* 算法介绍

A*算法是人工智能中的一种启发式搜索算法,它在电脑游戏的寻径计算以及最佳路径计算、最佳航线选择^[6]中得到广泛的应用。人工智能计算中,在寻找从初始状态到目标状态过程中会产生许多分枝,使得求解的路径很多,就构成了一个图,称为状态空间图。A*算法能够尽快地求出状态空间搜索的最佳路径。作为一种启发式搜索算法,其核心就是利用一个估价函数评估每次决策的价值,决定先尝试哪一种方案。A*算法在某一节点的估价函数可以表示为:

$$F(n) = G(n) + H(n) \quad (1)$$

其中, $F(n)$ 是节点 n 的评价函数, $G(n)$ 是从起始节点到节点 n 的实际代价, $H(n)$ 是从节点 n 到目标节点的最佳路径的估计函数,它包含了搜索的启发信息。启发式搜索的基本思想是:对于搜索过程中遇到的每个新节点,按估价函数计算出它的最佳代价估计值,然后选出当时估计值最小的节点,从该结点开始继续搜索,它实际上是以节点的代价估计值为标准的最佳优先搜索。关于A*算法的详细描述可参阅文献[5]~[8],本文将利用A*算法的基本思想,来进行可达性的计算。

2 可达性计算方法

2.1 网格的可达性与易达性

城市中某一区域可达性应从两个角度来分析,一方面,分析从城市其他各点、各区到该点、该区的方便性,另一方面,分析从这一点、这一区到城市其他地方的方便性。这两方面的意义并不等同。作为一个店主、商人或者公共娱乐场所的经理,对前一方面的可达性比较感兴趣。而作为一般居民,则对后一方面的可达性更感兴趣^[9]。

因此,将研究区域分成等距的网格,网格分两种类型,一种是服务需求网格,一种是服务提供网格,不同类型的网格的可达性内涵是不一致的。

每一服务需求网格获取某项服务的可达性定义为:所有服务提供网格的服务吸引力与从该需求网格开始克服中间网格的阻力到达提供服务网格的比值的和。其计算公式见式2,它与最通用的可达性计算模型——潜能模型^[2]的公式是一致的:

$$A_i = \sum_{j=1}^n \frac{M_j}{C_{ij}} \quad (2)$$

其中, A_i 为服务需求网格 i 的可达性值; M_j 是服务提供网格 j 的质量(服务能力);

C_{ij} 是服务需求网格 i 到服务提供网格 j 的交通成本； ∂ 为 i 和 j 之间的距离摩擦系数。

所有服务需求网格到某一服务提供网格的方便程度，定义为该服务提供网格的易达性，易达性计算公式见式 (3)。

$$B_j = \sum_{i=1}^n P_i \times C_{ij} \tag{3}$$

其中， B_j 为服务提供网格 j 的易达性值； P_i 是服务需求网格 i 的人口密度， C_{ij} 是服务需求网格 i 到服务提供网格 j 的交通成本。 B_j 值越高，该服务提供网格的易达程度越低。

2 2 网络交通成本

有两种方法来模拟通过网格的交通成本，一种是道路密度法^[10]，一种是土地利用类型相对阻力法^[11, 12]。

一般说来，道路密度越高，交通越方便，网格的道路密度与交通成本反相关，网格的交通成本由式 (4) 计算：

$$C_i = \left[\frac{\sum_{j=1}^m f_j \times L_j}{area(G_i)} \right]^{-\lambda} \tag{4}$$

城市中的道路的类型有多种，不同类型的道路对交通成本的影响是不一样的， m 表示道路类型，这里我们将道路分成四类，即 1—主干道、2—次要道路、3—立交高架路、4—地铁。 L_j 表示落在网格内 j 类道路的长度， f_j 是第 j 类道路的影响因子，地铁最高，次要道路最低。 C_i 表示第 i 个网格的交通成本， $area(G_i)$ 表示第 i 个网格的面积， λ 表示相关系数。

城市土地利用与交通相互联系、相互影响，交通发展与土地利用相互促进。不同的土地利用形态，决定了交通发生量和交通吸引量，决定了交通分布形态，在一定程度上决定了交通结构^[13~17]。不同的土地利用类型对交通产生的阻力是不同的，因此可以根据土地利用类型来模拟交通成本。俞孔坚、周廷刚等^[11, 12] 根据城市中不同土地利用类型，提出了一个土地利用类型空间阻力相对值表，如表 1：

表 1 土地利用类型空间阻力相对值表

Tab 1 Relative spatial resistance for difference land use types

土地利用类型	相对阻力	土地利用类型	相对阻力
道路、广场	1	仓储用地	100
居住用地	3	工业用地	100
公共绿地	12	公共设施用地	100
市政设施用地	5	河流	999

根据周廷刚 (2004)

通过每一个网格的交通相对阻力为：

$$R_i = \sum_{k=1}^n R_k^* \times P_k \tag{5}$$

其中， R_k^* 表示 k 类土地的空间阻力相对值， P_k 表示 k 类土地在网格内所占面积比重。

综合式 (4) 和 (5)，每个网格的交通成本可以用式 (6) 表示。由于网格是对城市区域的硬性划分，在实际计算中，发现有些相邻网格间的交通成本跳跃较大，因为交通成本

仅是一个相对值，这里通过取对数，可以使计算结果平稳一些。

$$C_i = \log \left[\left(\frac{\sum_{j=1}^m f_j \times L_j}{\text{area}(G_i)} \right)^{-\lambda} \times \sum_{k=1}^n (R_k^* \times P_k) \right] \tag{6}$$

2 3 网格路径交通成本计算

从一个网格到另一个网格的路径有多条，例如，图 1 中列出从网格 G(2, 2) 到 G(4, 7) 中的三条路径。两个相邻网格之间的交通成本为两个网格中所有点之间交通成本的平均值，如图 2，网格 1 与网格 2 相邻，则网格 1 到网格 2 的交通成本为：(C₁ + C₂)/2，网格 1 到网格 3 的交通成本为：(C₁ + C₃)/√2/2。

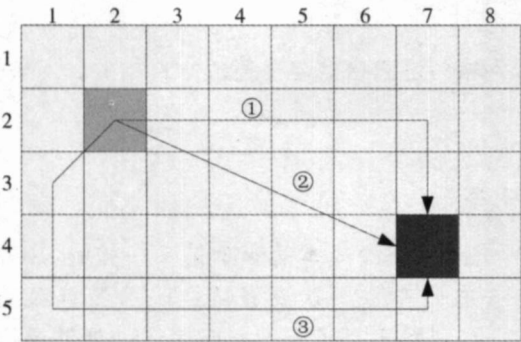


图 1 网格路径示例
Fig. 1 Examples for grid path

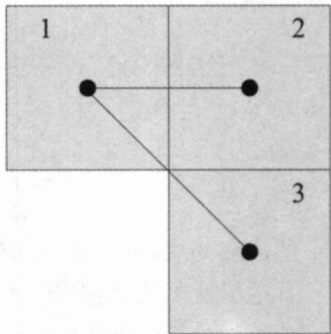


图 2 网格间平均交通成本
Fig. 2 Average traffic cost between grids

如何求取网格间路径的交通成本是本文研究的关键。Dijkstra 算法能保证求取网格间的最小交通成本，但它是较多的搜索空间和时间作为代价的。A* 算法不但能提高搜索效率，更重要的是还能加入一些启发式搜索信息，使搜索的结果更符合实际需求。

利用 A* 算法的基本思想，搜索网格间的最佳路径的算法过程如下：

- (1) 建立一个 OpenTable，放入起始网格编号 S；
- (2) 重复如下步骤：
 - ①寻找 OpenTable 中 F 值最低的网格。称它为当前格 I；
 - ②把 I 放入到 CloseTable 中；
 - ③找到 I 的八邻域网格，对每一个网格做如下处理：
 - ◆如果它已在 CloseTable 中，则略过；
 - ◆如果它不在 OpenTable 中，把它添加进 OpenTable，并以 I 作为它的 Father，计算这一网格的 F 值、G 值、H 值。
 - ◆如果它已在 OpenTable 中，以 G 值为参考，检查到从 I 网格到达它的路径是否最好，如果是，就把这一格的 Father 指针改为当前网格 I，并重新计算这一网格的 G 值和 F 值；
 - ④停止。当目的网格 O 添加进入 OpenTable 中，则路径被找到，或者没有找到目的网格 O，但 OpenTable 空了，则路径不存在。
- (3) 从目的网格 O 开始，沿着每一网格的 Father 指针回到 S，则得到基于启发函数 H(n) 的最佳路径。

A* 算法的关键是确定合适的启发估价函数 $H(I)$, 它负责计算从当前网格 I 到目的网格 O 的预估交通成本, 它对最终的搜索效率影响很大。在城市中从一个地点到达另一个地点行进时受距离和角度两个因素约束, 因此, 这里采用欧基里德距离结合行进角度来计算预估交通成本。有: $H(I) = \exp(1 + \sin\theta) \sum_{i=1}^m C_i$ (7)

式中, $\sum_{i=1}^m C_i$ 表示从 I 网格到 O 网格采用欧基里德法生成路径所经过网格的交通成本 C_i 的和, 其中, 欧基里德法生成路径所经过的网格可以利用计算机图形学中的 Bresenham^[18] 直线扫描生成算法得到。 θ 是以 S 网格和 O 网格的连线为基线, I 网格与 O 网格连线与基线的夹角。

在实际应用中, 为了使计算结果更符合实际需求, 可引入更多的启发式信息, 如有公交站点的网格优先搜索, 地铁线经过的网格优先搜索等。

3 计算案例

我们选择广州市作为案例研究区域, 以商务区的可达性作为计算对象。研究区域包括广州市几个中心行政区: 天河区、海珠区、越秀区、荔湾区、白云区、芳村区等。道路的数据采用广州市 1:10000 交通电子地图, 分为主干道、次要道路、立交高架路、地铁等道路类型。为了完成计算案例, 我们从有关部门获得广州市各街区的五普人口数和人口密度, 广州市分级商务区空间分布图, 商务区分 6 个级别 (如图版 3 图 3); 广州市规划土地利用数据。

由于这些数据来源不一样, 空间参考不一致, 我们利用 ArcInfo 的空间配准工具对这些数据配准, 基本上能满足计算精度。另外, 由于这些数据所覆盖的空间范围不一致, 在尽量保留重合的空间数据的前提下, 对数据做了一定的取舍裁剪。在 ArcInfo9 的二次开发工具 ArcEngine 环境下, 以 C# 语言开发了比较通用的可达性计算包, 主要功能包括网格生成、研究区域网格划分、网格内交通成本计算、A* 网格搜索、网格可达性计算等。

将研究区域分成 500m 等距网格, 包含商务区的网格为服务提供网格, 将商务区的等级作为商务区的服务质量。其他没有包含商务区的网格为服务需求网格。根据式 (6) 计算出的广州市交通成本空间分布网格图, 如图版 3 图 4。

根据式 (2)、式 (3) 计算出广州市商务区网格可达性指数和网格易达性指数, 其空间分布如图版 3 图 5、6。

根据图 3, 越秀区集中了几个级别比较高、面积较大的商务区, 且交通成本比较低。因此, 从图 5 可以看出, 以越秀区为中心, 整个城市的商务区服务需求网格可达性指数由高到低呈向外递减分布, 图 5 已清晰地给出可达性指数等势圈。根据图 6, 易达性较好的商务区服务提供网格集中在海珠区北部和越秀区一带地区, 这与该地区发达的城市交通网络 (地铁、立交桥) 有一定关系。相对来说, 天河区以北和白云区的商务区服务提供网格的易达性比较差。这些计算结果将为广州市商务区的总体规划提供一些定量的参考。

4 结论

本文提出的可达性计算方法有以下特点:

- (1) 适用于研究城市内部的个体或商业区位的微观可达性评估;

(2) 数据来源获取和数据处理相对比较容易;

(3) 将研究区域以网格方式进行划分, 避免了一般的空间简化方法所带来的不确定性, 并且空间粒度可以随网格间距大小进行调整;

(4) 采用启发式 A* 算法进行网格路径交通成本计算, 提高了计算效率, 它虽然不能保证网格路径交通成本最小, 但可以通过加入不同的启发信息, 使搜索结果符合实际需求。

在本研究中还有许多影响因素没有涉及, 例如, 相同服务质量的服务提供网格之间的竞争效应, 人口密度对网格交通成本的影响, 网格划分的尺度对计算结果的影响, 交通网络密度的对计算结果的非线性影响等, 这将在以后的工作中, 收集更丰富的数据资源, 作进一步的研究。

参考文献:

- [1] 李平华, 陆玉麒. 城市可达性研究的理论与方法评述. 城市问题, 2005, 123(1): 69~ 73
- [2] 杨培军, 宋小东. 基于 GIS 的可达性评价方法比较. 长安大学学报, 2004, 21(4): 27~ 32
- [3] 吴威, 曹有挥, 曹卫东, 梁双波. 开放条件下长江三角洲区域的综合交通可达性空间格局. 地理研究, 2007, 26(2): 391~ 402
- [4] Karst T Geurs, Bert van Wee. Accessibility evaluation of land-use and transport strategies: Review and research directions. Journal of Transport Geography, 2004, 12: 127~ 140
- [5] 王德春, 陈利敏, 张孝芳. 基于 A* 算法的航舰最佳航线选择. 青岛大学学报(自然科学版), 2005, 18(4): 10~ 13
- [6] <http://www-cs.stanford.edu/~amitp/Articles/AStar5.html>
- [7] Ikeda T, Imai H. Enhanced A* algorithms for multiple alignments: Optimal alignments for several sequences and k2 opt approximate alignments for large cases. Theoretical Computer Science, 1999, 210: 341~ 374
- [8] 邹亮, 徐建闽, 朱玲湘. A* 算法在基于电子地图的动态路径诱导中的应用. 武汉理工大学学报(交通科学与工程版), 2006, 30(5): 885~ 888.
- [9] 杨涛, 过秀成. 城市交通可达性新概念及其应用研究. 中国公路学报, 1995, 8(2): 25~ 30
- [10] 马林兵, 曹小曙. 基于 GIS 的城市公共绿地景观可达性评价方法研究. 中山大学学报(自然科学版), 2006, 45(6): 111~ 115
- [11] 俞孔坚, 段铁武, 李迪华, 彭晋福. 景观可达性作为衡量城市绿地系统功能指标的评价方法与案例. 城市规划, 1999, 123(8): 8~ 11.
- [12] 周廷刚, 郭志达. 基于 GIS 的城市绿地景观引力场研究——以宁波市为例. 生态学报, 2004, 24(6): 1158~ 1163
- [13] Harvey J Miller. GIS software for measuring space-time accessibility in transportation planning and analysis. GeoInformatica, 2000, 4(2): 141~ 159
- [14] Davidson K B. Accessibility in transport/land-use modeling and assessment. Environment and Planning, 1997, A9, 521~ 528
- [15] Si-ming Li, Yi-man Shum. Impacts of the national trunk highway system on accessibility in China. Journal of Transport Geography, 2001, 9: 39~ 48
- [16] 毛蒋兴, 闫小培. 城市交通系统对土地利用的影响作用研究. 地理科学, 2002, 25(3): 353~ 360
- [17] 曹小曙, 闫小培. 经济发达地区交通网络演化对通达性空间格局的影响. 地理研究, 2003, 22(3): 305~ 312
- [18] 孙家广. 计算机图形学. 北京: 清华大学出版社, 1998

A computing method of spatial accessibility based on heuristic A^* algorithms and grids partition

MA Lin-bing, CAO Xiao-shu

(School of Geography and Planning, Sun Yat-sen University, Guangzhou 510257, China)

Abstract: An accessibility computing method based on grid partition and heuristic A^* searching algorithms was put forward in the paper, which is used to research accessibility spatial distribution feature in the study area through computing each grid's accessibility.

Traffic cost is an important element in accessibility computing. There are two approaches to stimulate grid traffic cost, one is approach of density of road network, and the other one is approach of relative spatial resistance for different land use status. The density of road network is anti-correlative with traffic cost. Land use has a close relationship with urban traffic, so grid relative traffic cost can be stimulated directly by integrating the two approaches, actual status of road and topological structure of road network was ignored. A computing formula of grid relative traffic cost was defined in the paper.

The key point is how to compute traffic cost of grids path. Dijkstra algorithm makes sure to get the lowest traffic cost, but it must spend more searching space and time. Not only can A^* algorithms improve searching efficiency, but also can add some heuristic searching information to make searching result more reasonable. The key of the searching process is the definition of heuristic evaluation function H , which is responsible for estimating traffic cost from intermediate grid to target grid. Usually, marching from one spot to another is constrained by two factors: distance and arrival angle. So Euclid distance and marching arrival angle are adopted to evaluate traffic cost. In application, in order to make computing result accorded with practice, more heuristic information can be imported into computing process.

For verifying our research, a software package was developed with C# language under ArcEngine9 environment. A case study was implemented by using the package. Accessibility of business districts was selected as computing object in Guangzhou city. The computing result by using the method can provide a quantitative reference for urban planning of the city.

Key words: accessibility; heuristic A^* algorithms; urban traffic

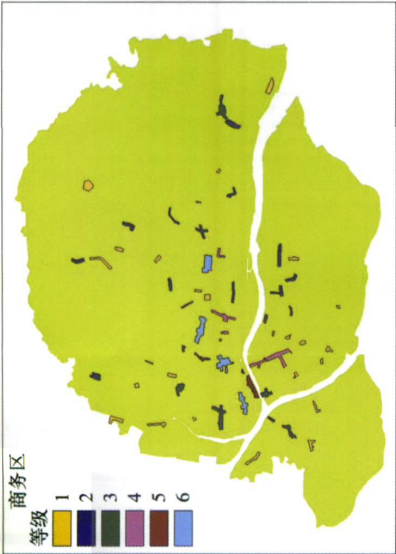


图 3 Spatial distribution of business district in Guangzhou

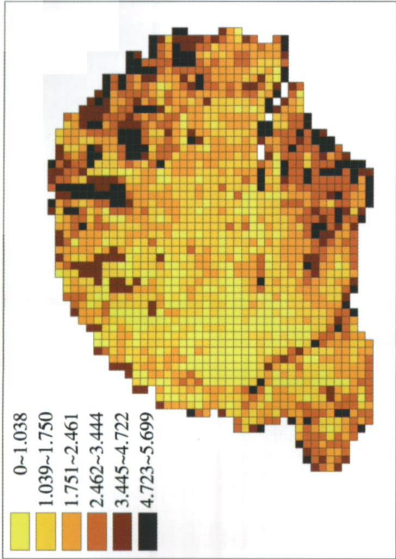


图 4 Spatial distribution of grid traffic cost

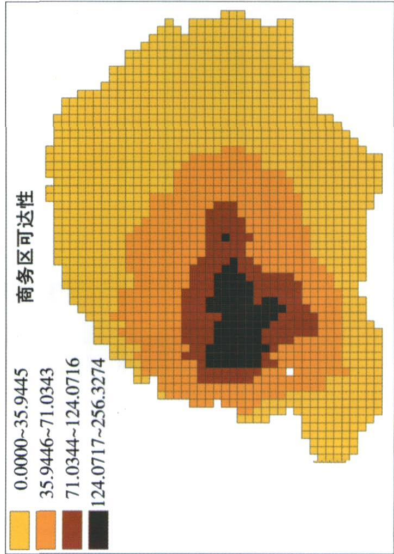


图 5 Spatial distribution of accessibility of needing business service grid

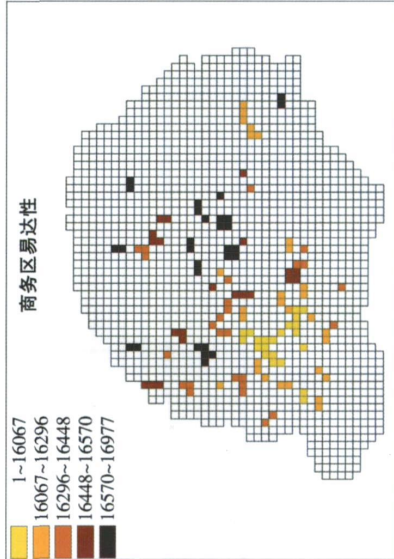


图 6 Spatial distribution of easy-accessing of providing business service grid