道路网约束下的城市事件空间点模式分析*

佘 冰, 朱欣焰, 禹 维†, 杨云涛, 徐 晓
(武汉大学 测绘遥感信息工程国家重点实验室, 武汉 430079)

摘 要: 为对城市管理的资源分配提供有效的决策支持,基于道路网约束下的核密度估计和 K 函数方法,对城市事件的一阶和二阶空间分布特性进行分析,并与平面空间点模式分析结果进行对比。实验结果表明,相比平面空间点模式分析方法,道路网约束下的核密度估计提取的"热点"路段可直观地体现事件沿道路网的分布情况,而道路网约束下的 K 函数能真实地反映出事件在不同空间尺度下的聚集程度。

关键词: 城市事件; 道路网约束; 空间点模式分析; K 函数; 核密度估计
中图分类号: TP391;P208 **文献标志码**: A **文章编号**: 1001-3695(2013)08-2327-03
doi:10.3969/j.issn.1001-3695.2013.08.021

Research on spatial pattern of city events under road network constraints

SHE Bing, ZHU Xin-yan, GUO Wei†, YANG Yun-tao, XU Xiao
(*State Key Laboratory of Information Engineering in Surveying, Mapping & Remote Sensing, Wuhan University, Wuhan* 430079, *China*)

Abstract: To provide effective decision support for resource planning in city management, this article analyzed the first-order and second-order properties of the city event distribution, using the kernel density estimation and the *K*-function methods under road network constraints. The results were compared with that produced by the planar point pattern methods. The results show that the "hotspot" road sections extracted by the kernel density estimation under road network constraints can more directly reveals the distribution of events along road network, and the *K*-function under road network constraints more faithfully reflects spatial clustering pattern in different spatial scales.

Key words: city events; road network constraints; point pattern analysis; *K*-function; kernel density estimation

0 引言

城市现代化建设进程不断加快,发展的同时也给城市的管理提出了更高的要求。城市网格化管理,是以城市公共区域的事件和部件为管理对象,实现对管理对象和管理责任的精细化管理,提高城市管理的综合化和信息化水平,同时实现专业处置部门和网格监督员联动的管理模式和信息资源共享系统[1]。2005 年 8 月,武汉市在江汉区启动数字化城市管理系统建设试点[2]。经过多年持续运行,城市网格管理系统已经累积大量城市事件历史数据。为了充分发挥数字化城市管理系统的优势,针对这些历史事件数据,利用时空数据分析技术,深入挖掘城市管理事件时空分布规律,对城市管理人力资源分配方案的制定及城市管理决策支持有重要的指导意义。

城市事件的发生大多数情况下在空间和时间上并不是随机的,很多情况下,它们在几何平面上形成聚集。基于城市事件的空间分布特性,本文采用空间点模式分析方法对其进行分析。空间点模式分析主要用于探测点事件的分布特性,在生态学、环境管理、公共卫生等领域已被广泛应用[3~5],而专门针对城市事件的空间模式进行分析的相关研究仍然较少。此外,传统的平面空间点模式方法在应用到具有明显沿道路网发生趋势的事件时会生成错误的聚集模式[6]。因此,研究人员开始将网络下的空间点模式方法应用到各自领域的数据分析中[7,8]。本文利用道路网约束条件下的空间点模式分析方法,针对武汉市江汉区城市事件数据,探究城市事件中占道经营事件的空间模式,以期为数字化城市管理提供科学参考。

1 研究区域和数据来源

武汉市江汉区位于长江北岸,2005 年作为试点进行数字化城市管理系统的建设。武汉市政府根据江汉区自然地理布局和行政区域划分现状,在全区 13 条街道、114 个社区的基础上进一步通过网格划分,把全区划分为 1 003 个网格。本文使用的占道经营数据来自武汉城市网格化管理与服务系统,该系统登记了江汉区每天发生的所有城市事件,涉及 4 大类 28 小类共 134 种类型的事件。其中,占道经营事件是城市管理事件中的一类主体事件。本文主要针对 2011 年 3 月份的占道经营事件进行分析。

2 研究方法

道路网约束下的空间点模式分析方法可以被分为两类:一阶特性和二阶特性。其中一阶特性是检测空间过程的一阶影响,探究点事件的均值随着空间的变化而变化的过程,其代表为核密度估计方法;而二阶特性检测空间过程的二阶影响,主要考察点事件之间的空间依赖性,其代表为 K 函数分析方法。

收稿日期: 2012-12-30; **修回日期**: 2012-02-04 **基金项目**: 国家"863"计划资助项目(2011AA010502);国家科技支撑计划资助项目(2012BAH35B03); 测绘遥感信息工程国家重点实验室开放基金资助项目; 中央高校基本科研业务费专项资金资助项目

作者简介: 佘冰(1987-),男,湖北黄石人,博士研究生,主要研究方向为空间数据挖掘;朱欣焰(1963-),男,教授,博导,主要研究方向为空间数据库和地理信息服务;禹维(1981-),男(通信作者),湖北公安人,讲师,博士,主要研究方向为地理信息服务(guowei98032@gmail.com);杨云涛(1986-),女,硕士研究生,主要研究方向为空间分析;徐晓(1990-),男,硕士研究生,主要研究方向为空间统计.

2.1 道路网约束下的核密度估计方法

核密度估计基本思想是假定研究区域内任意一个位置都有一个潜在的事件密度。平面核密度估计方法是以空间上某点为圆心,统计出半径为 r 的圆内事件发生的数量,除以圆的面积。所以在事件点 s 处的密度可以表示如下[9]:

$$\lambda(s)=\sum_{i=1}^{n}\frac{1}{\pi r^2}k\left(\frac{d_{is}}{r}\right) \tag{1}$$

其中:r 为带宽;d_{is} 为点 i 到事件点 s 的距离;k 为核函数,常为对称密度函数,可以理解为点 i 的权重。核估计可以理解为每个点突起值的平均值,核函数 k 决定突起的形状,而带宽 r 决定突起的宽度。

道路网约束下的核密度估计法,核心思想在于将道路网空间分解为基本等长的线性单元,记为 lixel(即 linear pixel),然后用最短网络距离替代欧氏距离作为距离度量方式[9]。算法基本步骤如下:

a)将原有的道路网转换为基于路段的线性参考系统,路段即为相邻两个道路交叉点之间的线段,如图 1 中的 AB 路段。

b)把每个路段划分为等长的线性单元,记为 lixel。

c)对每个 lixel 保存一个计数器。遍历点对象集合,针对每一个事件点,求得距离该点最近的 lixel,将该 lixel 的计数器值加 1。

d)计算每个 lixel 的中心到邻近 lixel 的中心的最短网络距离 d,选择在带宽 r 之内的 lixel 作为其邻居。针对每个 lixel,利用式(2)计算该 lixel 处的密度值:

$$\lambda(s)=\sum_{i=1}^{n}c_i\frac{1}{r}k\left(\frac{d_{is}}{r}\right) \tag{2}$$

其中:d_{is} 为两个 lixel 在网络上最短路径距离;r 为带宽;k 为核函数;c_i 为 lixel 的事件计数器值;n 为 lixel s 的邻居 lixel 数目。

需要指出的是,道路网约束下的核密度估计分析与平面核密度估计分析方法不同之处主要体现在对于带宽的距离计算是基于道路网得到的最短网络距离,并且密度值的计算是针对每个线性单元,而非二维圆形面积。Okabe 等人[10,11]给出了网络下核密度估计的计算和优化算法以及无偏的核函数表达式。与平面核密度分析的参数性质类似,核函数的选取对道路网约束下的核密度估计分析结果影响不大,而带宽和 lixel 长度则会对结果造成很大影响。

2.2 道路网约束下的 K 函数

道路网约束下的 K 函数继承于平面 K 函数(又称 Ripley's K 函数)。Ripley's K 函数的基本思想是:假定依次在各个事件中心设置半径为 r 的圆,计算落入圆内其他事件的数量,然后针对所有的圆,计算出圆内事件的平均数量除上密度。对一系列距离值 d 不断重复,可定义为[12]

$$K(d)=\frac{1}{\lambda}\times\frac{\sum_{i=1}^{n}\#[d_{\min}(i)<r]}{n}=\frac{A}{n^2}\sum_{i=1}^{n}\#[d_{\min}(i)<r] \tag{3}$$

其中:A 为研究区域面积;n 为点总数;λ 为点的密度;$\#[d_{\min}(i)<r]$ 表示落在以点 i 为圆心、r 为半径的圆区域内的点的数量(不包括点 i)。

道路网约束下的 K 函数是由 Okabe 和 Yamada 对平面 K 函数方法进行改进将其扩展到网络空间的[13]。假设存在分布在道路网 R 上的事件集 ξ,记道路网的边集合为 L_T,其总长度为 l_T,那么在完全空间随机过程的假设下道路网约束下的 K 函数定义为

$$K(t)=\omega^{-1}\times E(t,\xi,L_T) \tag{4}$$

其中:ω 为 R 上的点密度;$\omega=n/l_T$,n 为事件点总数;E 为 ξ 中以 t 为搜索距离的事件数目在二项点过程假设下的期望值。网络距离 t 定义为 L_T 中两个事件点的最短网络距离。由此道路网约束下 K 的估计值可表达为

$$\hat{K}(t)=\omega^{-2}\times l_T^{-1}\times\sum_{i=1}^{n}n_{P_i}(t) \tag{5}$$

其中:$n_{P_i}(t)$ 为事件点 P_i 在道路网的边集合 L_T 网络距离 t 内的所有邻域点数目。道路网约束下 K 函数采用蒙特卡罗模拟方法检验空间分布模式的显著性,但是由于这里使用的是网络距离作为度量方式,所以计算量会比传统 K 函数的检验大很多。其检验过程为:按照 CSR 过程生成 n 次的分布数据,计算每一次 CSR 过程的 $K(t)$,如果 $K(t)$ 的观测值小于给定的 t 尺度上对应的 CSR 过程中 $K(t)$ 的最小值或大于最大值时,即可判定点模式在这一尺度上显著地异于 CSR。

3 分析结果

本文运用 SANET 4.0 进行道路网约束下的 K 函数计算及核密度估计,结合 CrimeStat 3.0 对平面 K 函数和核密度估计进行对比分析。K 函数的计算结果通过 R 语言进行可视化,核密度分析结果通过 ArcGis 10.0 进行可视化。

3.1 核密度估计分析

对于道路网约束下的核密度估计分析方法,主要参数有核函数、带宽、lixel 长度。研究发现,不同的核函数选择对整体的密度模式影响很小[9]。本文统一采用的是 SANET 软件中的 equal split 核函数。图 2 为选取江汉区局部路段对不同带宽下进行核密度估计分析的结果,路段宽度与密度值大小成正比。可见,带宽越小,聚集的细节变化信息越丰富;带宽越大,小的聚集将被合并成大的"热点",空间上点密度的变化更光滑,细节信息减少,但直观可视化效果更好。Lixel 长度与带宽对结果的影响类似。在实际的数据分析中,带宽的大小反映了事件之间的影响强度,需要根据事件的实际特征来确定,而 lixel 长度反映了计算的最小单元,需要根据城市管理的需要和尺度来确定。

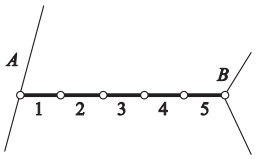
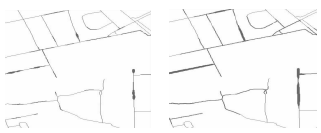


图1 Lixel分割示意图

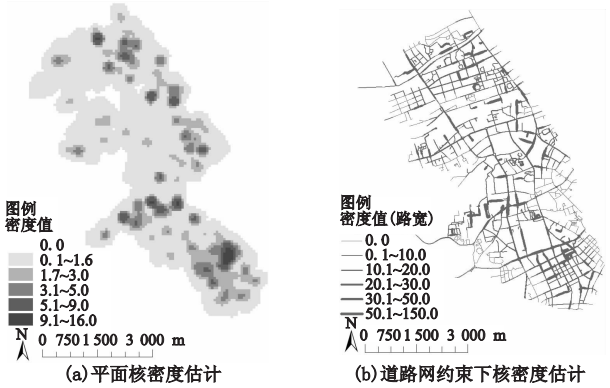


(a)20 m (b)100 m

图2 不同带宽对核密度估计分析结果影响示意(lixel统一选取10 m)

本文采用 250 m 带宽和 10 m 的 lixel 长度进行分析,其显示聚集效果较直观,并且细节变化信息也较明显。图 3 表示了平面核密度估计分析和道路网约束下核密度估计分析的结果,可以看出两者的"热点"区域相似。但平面核密度是对整个地区进行密度分析,结果显示存在五个"热点"地区,热点核心区经纬度为:114.282E,30.582N;114.245E,30.635N;114.252E,30.592N;114.267E,30.595N;114.26E,30.595N。而道路网约束下的核密度分析是针对道路进行密度分析,因此可以找到"热点"路段。由图 3(b)可见,占道经营的"热点"主要集中在三级道路上,在二级道路上发生较少。

可以看出,平面空间点模式分析对于点事件在研究范围内均匀发生的假设,以及单纯地用点事件间的欧氏距离来表示两个事件的距离的思想,显然违背了明显沿道路分布的事件。道路网是网络空间的一种特定实例。

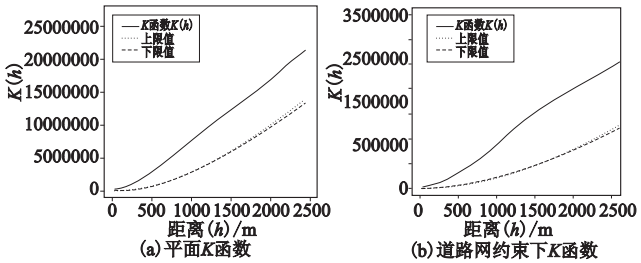


(a)平面核密度估计　　(b)道路网约束下核密度估计

图3　占道经营类事件对比分析

3.2　*K* 函数分析

图 4 显示了占道经营类事件分别通过平面 *K* 函数与道路网约束下 *K* 函数进行分析的结果,图中实线代表了实际观测值的 *K* 函数曲线,而上下虚线分别代表蒙特卡洛模拟的上百分之五分位数和下百分之五分位数 *K* 函数曲线。可以看出,占道经营事件在平面 *K* 函数和道路网约束下 *K* 函数计算下都可以得出其呈聚集分布。但是对比两图中实际观测值与模拟值的比例差别可以看出,占道经营事件在道路网约束下比在平面情况下明显呈现出更强的聚集性。



(a)平面*K*函数　　(b)道路网约束下*K*函数

图4　占道经营类事件平面*K*函数分析与道路网约束下*K*函数分析结果

Lu 等人详细分析了平面 *K* 函数分析沿道路分布事件带来的问题[6]:当区域内道路网密度较小的情况下,由平面 *K* 函数计算出的聚集强度会高于网络 *K* 函数的计算结果,从而产生虚假警报;而当区域内道路网密度较大的情况下,由平面 *K* 函数计算出的聚集强度会低于网络 *K* 函数的计算结果,从而低估了事件的聚集趋势。本文针对占道事件的研究结果符合上述后一种情况,即平面 *K* 函数的结果低估了事件的真实聚集强度。

3.3　城市占道经营事件空间模式特征

城市管理事件主要依靠监督员巡逻发现,而占道经营事件是城市事件的一类主体事件,因此揭示占道经营事件的空间分布特征,并探讨影响其空间格局的相关因素,对于城市管理事件的预防和人力资源配置具有现实指导意义。通过道路网下的 *K* 函数分析可以得出,占道经营事件呈现显著的沿道路聚集的趋势。这一方面揭示出,占道经营事件并不是随机发生在道路网上。由于不同路段所处地区的人口、经济情况都不尽相同,因此导致占道经营事件产生的几率随着所处路段变化而变化。另一方面,由于空间现象往往存在空间自相关性,占道经营事件之间的相互作用也不可忽略。

经过道路网约束下的核密度估计分析发现,占道经营“热点”或靠近城中区地区,其缺少专业市场,为占道经营的摊贩提供了机会;或处在商业中心和广场附近,其人口流量较大,占道事件也相对较多。这些路段为城市事件高发路段,应为城市管理的治理以及城市资源的合理规划重点考虑的路段区域。值得注意的是,平面核密度分析可以直观地体现出在社区尺度的“热点”,因此对于城市的管理也具有一定实践意义。但相比而言,道路网约束下的核密度估计分析,则从更精细的尺度上研究事件的聚集性,寻找出事件的“热点”路段,可以更直接、准确地为城市管理者提供决策支持。

4　结束语

由于城市事件多沿道路分布,传统的平面空间点模式分析方法会对聚集情况作出错误的估计。本文利用道路网约束下的空间点模式分析方法,发现占道经营事件呈现了明显的沿道路聚集的趋势,且聚集程度较利用传统的平面空间点模式分析结果更为显著。道路网约束下的核密度分析从道路尺度上精确反映了事件的强度信息,相比平面核函数对城市管理更具指导意义。为了深入理解事件的影响因素以及对事件数量进行预测,后续研究还需考虑路网等级因素作进一步的回归建模分析。此外,道路网约束下的空间模式分析方法相比平面下的空间模式分析方法计算量大幅增加,需要研究相关优化和并行算法以提高算法执行效率。

参考文献:

[1] 李德仁,邵振峰,朱欣焰. 论空间信息多级格网及其典型应用[J]. 武汉大学学报:信息科学版,2004,29(11):945-950.

[2] 李德仁,李宗华,彭明军,等. 武汉市城市网格化管理与服务系统建设与应用[J]. 测绘通报,2007(8):1-4,44.

[3] HARRIS K, GENDE S, LOGSDON M, *et al*. Spatial pattern analysis of cruise ship-humpback whale interactions in and near Glacier Bay National Park, Alaska[J]. Environmental Management,2012,49(1):44-54.

[4] VADREVU K, BADARINATH K, ANURADHA E. Spatial patterns in vegetation fires in the Indian region[J]. Environmental Monitoring and Assessment,2008,147(1-3):1-13.

[5] 曹志冬,王劲峰,高一鸽,等. 广州 SARS 流行过程的空间模式与分异特征[J]. 地理研究,2008,27(5):1139-1149.

[6] LU Yong-mei, CHEN Xu-wei. On the false alarm of planar *K*-function when analyzing urban crime distributed along streets[J]. Social Science Research,2007,36(2):611-632.

[7] DAY P L, PEARCE J. Obesity-promoting food environments and the spatial clustering of food outlets around schools[J]. American Journal of Preventive Medicine,2011,40(2):113-121.

[8] DAI D, TAQUECHEL E P, STEWARD J, *et al*. The impact of built environment on pedestrian crashes and the identification of crash clusters on an urban university campus[J]. Western Journal of Emergency Medicine,2010,11(3):294-301.

[9] XIE Zhi-xiao, YAN Jun. Kernel density estimation of traffic accidents in a network space[J]. Computers, Environment and Urban Systems,2008,32(5):396-406.

[10] OKABE A, SATOH T, SUGIHARA K. A kernel density estimation method for networks, its computational method and a GIS-based tool[J]. International Journal of Geographical Information Science,2009,23(1):7-32.

[11] SUGIHARA K, SATOH T, OKABE A. Simple and unbiased kernel function for network analysis[C]//Proc of the 10th International Symposium on Communications and Information Technologies. 2010:827-832.

[12] RIPLEY B D. Spatial statistics[M]. [S. l.]: Wiley-Interscience, 2004.

[13] OKABE A, YAMADA I. The *K*-function method on a network and its computational implementation[J]. Geographical Analysis,2001,33(3):271-290.