

# 基于 Voronoi 图的有障碍物空间聚类<sup>\*</sup>

薛丽霞<sup>1,2</sup> 汪林林<sup>2</sup> 王佐成<sup>1,2</sup> 李永树<sup>1</sup>

(西南交通大学土木工程学院 成都 610031)<sup>1</sup> (重庆邮电学院 重庆 400065)<sup>2</sup>

**摘要** Voronoi 图具有侧向临近特性,可以方便表达空间临近关系。根据这一性质,解决在空间数据聚类分析中的有实体障碍物所面临的问题。通过生成考虑实体障碍物的 Voronoi 图,建立 Voronoi 图的距离度量方法,提出采用循障碍物求距离法,可以方便地完成有障碍物的空间聚类分析。该方法使得有实体障碍物情况下空间聚类更加准确,并通过实验进行验证。

**关键词** Voronoi 图,空间聚类,Voronoi 距离

## Spatial Clustering in the Presence of Obstacles Based on Voronoi Diagram

XUE Li-Xia<sup>1,2</sup> WANG Zuo-Cheng<sup>1,2</sup> LI Yong-Shu<sup>1</sup> Wang Lin-Lin<sup>2</sup>

(Civil Engineering College, Southwest Jiaotong University, Chengdu 610031)<sup>1</sup>

(Chongqing University of Post and Telecommunication, Chongqing 400065)<sup>2</sup>

**Abstract** Voronoi diagram has lateral spatial adjacency character, based on it, we can express the spatial lateral adjacency relation conveniently and solve the problem derived from spatial clustering in the presence of obstacles. The paper explains the method of Voronoi diagram building, proposes the distance measure of Voronoi diagram, and the distance measure algorithm of following the obstacle. In virtue of the method, the spatial clustering in the presence of obstacles can be accomplished conveniently, and more precisely. Experiments with the method show that it is efficient.

**Keywords** Voronoi diagram, Spatial clustering, Voronoi distance

## 1 前言

空间数据库技术的发展为空间数据挖掘提供了许多有效的工具,然而却不能解决现实世界中许多存在约束条件的空间数据挖掘问题,例如具有实体障碍物空间聚类挖掘。在现实空间中,两个相邻的实体并不一定相连,例如道路两旁的两栋房屋与道路相邻这种侧向近邻关系,两栋房屋虽然临近,但是如果道路是具有隔离带的高速公路,则其距离关系就必须重新考虑。当前国际上流行的各种基于平面图空间数据模型的空间数据库是靠建立拓扑数据结构表达点、线、面空间目标之间的邻接关系。拓扑数据结构难以直接表达这种十分重要的侧向空间相邻关系<sup>[1]</sup>。Voronoi 图能够明晰地表达侧向临近关系,将 Voronoi 图方法应用于空间分析,能够解决有实体障碍物的空间数据挖掘面临的问题。

## 2 有实体障碍物的空间数据挖掘

在空间距离关联规则挖掘和空间聚类中,通常采用欧氏距离,两点之间的距离定义为直接的欧氏距离。但是,在两个对象之间有实体障碍物的实际问题的解决过程中,欧氏距离成为“理想距离”,影响数据挖掘的知识准确性,甚至导致错误的模式出现。例如为解决银行 ATM 机的分布问题而进行的空间聚类分析,如果分析区域有河流障碍存在,河流成为用户聚类分析中不得不考虑的重要因素。

实际上,这类问题归结为如何处理有实体障碍物的空间对象之间的距离问题。这里首先对实体障碍物进行分析,以上示例中的实体障碍物需要满足一个基本条件才能成为“必须考虑”意义上的障碍物。假设  $O_1$  和  $O_2$  为两个研究的空

对象, $d$  代表  $O_1$  与  $O_2$  的欧氏距离, $L$  代表障碍物, $D$  为从  $O_1$  通过实际路径到达  $O_2$  经过的距离, $L$  成为必须考虑的障碍物的条件是: $D \gg d$  ( $D$  远远大于  $d$ )。在实际的工作中,主要是线状地物或者山体等成为我们必须考虑的障碍物。

## 3 Voronoi 图的生成

Voronoi 图(或称 Dirichlet 铺盖、泰森多边形 Thiessen polygon)与一些自然结构十分相象,具有许多有趣而惊人的数学特性,早在 17 世纪就有科学家进行研究。上世纪 80 年代,有关 Voronoi 图的理论方法和应用成为研究的热点内容<sup>[2]</sup>。Voronoi 图为研究地学、计算机科学和数学等领域中的一些问题提供了有力的工具。在空间数据处理领域,Voronoi 图被用于实现空间临近操作、Buffer 分析、空间内插,数字化过程中的断点捕捉和多边形构造等,空间目标的 Voronoi 图成为又一个研究热点。

### 3.1 Voronoi 图的基本思想

Voronoi 图的定义是:假设  $P = \{p_1, p_2, \dots, p_n\}$  ( $3 \leq n < \infty$ ) 是欧几里得平面上的一个点集,并且这些点互不相同,即  $x_i \neq x_j, i \neq j, i, j \in I_n(1, 2, \dots, n)$ ,用  $d(p_i, p_j)$  表示点  $p_i$  与  $p_j$  的欧氏距离。设  $x$  为平面上的点,则区域  $V(i) = \{x \in E^2 | d(x, p_i) \leq d(x, p_j), j = 1, 2, \dots, n, j \neq i\}$  称为 Voronoi 多边形 ( $V$ -多边形,记为  $V(p_i)$ ),各点的  $V$  多边形共同组成 Voronoi 图<sup>[1]</sup>。

$V$ -图可以理解为对空间的一种分割方式(一个  $V$ -多边形内的任意一点到本  $V$ -多边形中心点的距离都小于到其它  $V$ -多边形中心点的距离)<sup>[3]</sup>,也可以理解为对空间的一种内插方式(空间中的任意一个未知点的值都可以由距离它最近的已

<sup>\*</sup> 基金项目:重庆市自然科学基金资助项目(CSTC 2005BB2065)。薛丽霞 讲师,博士研究生,主要研究方向为空间数据模型,空间数据库,数据挖掘,3S 技术及其应用。

知点,即采样点的值来代替),图1是离散生长点的 Voronoi 图。

Voronoi 图具有侧向临近的特性,两个侧向相邻的空间实体并不一定相连,如河流边的房屋与河流相邻,但在几何上并不相连。由于河流与房屋之间没有其他实体,因此它们的 Voronoi 多边形必有一条公共边,所以,只要判断两目标对象是否具有公共边,就可以判断两者之间是否侧向相邻。图2表示了房屋与河流的侧向近邻关系。

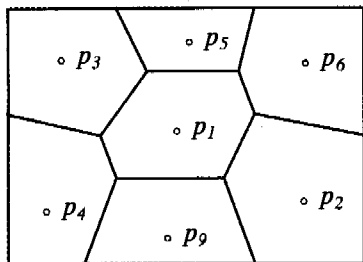


图1 离散点的 Voronoi 图

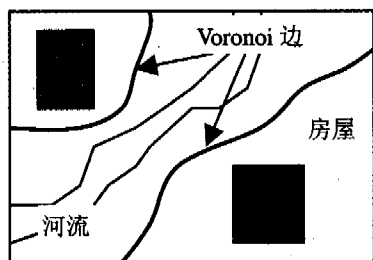


图2 侧向临近关系表达

### 3.2 顾及障碍物的 Voronoi 图的生成

Voronoi 图的生成是应用 Voronoi 图空间数据模型的前提,早期主要研究点状目标的 Voronoi 图生成算法,近年来逐步发展到研究线、曲线、面与混合目标的 Voronoi 图生成算法<sup>[4]</sup>,以及高维 Voronoi 图,多阶 Voronoi 图、球面 Voronoi 图、基于其他度量的 Voronoi 图等<sup>[5]</sup>。就生长算法的特点而言,主要分为矢量空间的生成算法和栅格空间中的生成算法,简称矢量法(vector-based)与栅格法(raster-based)。就本文的研究主题,我们仅对顾及障碍物的点与线混合 Voronoi 图的生成进行讨论。

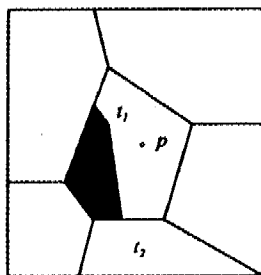


图3  $t_1 t_2$  分段剖分区

从通达性考虑障碍物,如河流、高速公路等都有架设的通行桥隧,在研究区域它们将障碍物分段。在普通 Voronoi 图的生成基础上,通过障碍物扫描算法<sup>[1]</sup>,可以确定障碍物所穿过的诸 Voronoi 区域,同时把一个障碍物分成了若干段,每一段都独自位于一个 Voronoi 多边形内,即每段都和相应的生长目标相联系。通过交换障碍物段的起、终点的方法规范障碍物段和相应生长目标之间的空间位置关系,然后生成障碍

物段的剖分区。图3中的阴影区表示  $t_1 t_2$  的分段剖分区。

在分段剖分区的基础上,对分段剖分区进行合并,形成若干整体剖分区。在有若干个障碍物的情况下,将整体剖分区合并成若干个统一剖分区进行统一处理。将剖分区进行合并后,相应的障碍物段也合并为一段。图4中的阴影区表示由两个分段剖分区合并而成的一个整体剖分区。通过处理共享边,采用分簇算法和最适搜索算法<sup>[1]</sup>处理公共边,得到可视最短路径 Voronoi 图,如图5。

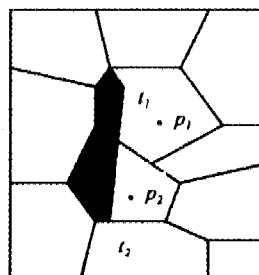


图4  $t_1 t_2$  的整体部分区

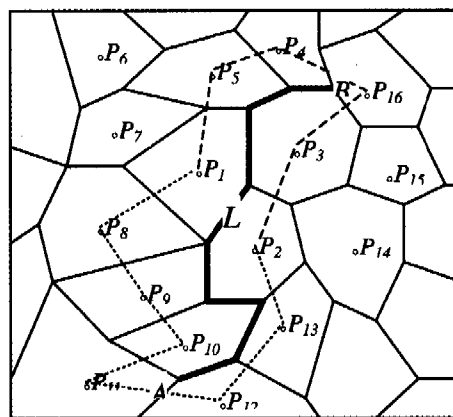


图5 可视最短路径 Voronoi 图

### 4 Voronoi 图的空间 Voronoi 距离

以 Voronoi 图为基础,我们可以发现空间两目标对象 Voronoi 区域之间的 Voronoi 区域数目从一定程度上反映出其间的其他目标的个数,这个数目体现出两空间对象的空间临近关系<sup>[6]</sup>。我们先定义这种 Voronoi 距离:设任意两个空间目标  $P_i, P_j$  之间的 Voronoi 区域的最少数目  $k$  为其间的 Voronoi 距离,记为  $Vd(P_i, P_j)$ ,一般地  $Vd(P_i, P_j) \geq 0$ ,当  $P_i = P_j$  时,  $Vd(P_i, P_j) = 0$ ;特殊情况当  $P_i \subset P_j$  或  $P_j \subset P_i$  时,  $Vd(P_i, P_j) = 0$ 。例如图5中,  $Vd(P_1, P_5) = 1, Vd(P_1, P_7) = 1, Vd(P_1, P_8) = 2$ 。

对有障碍物的 Voronoi 图,其 Voronoi 距离就必须考虑障碍物,图5中的  $L$  为一线状障碍物,在 VORDLL(Voronoi dynamic linking library)的基础上<sup>[7]</sup>,我们提出一种循障碍物求距离法,以图5中的  $P_1$  到  $P_2$  的 Voronoi 距离为例,算法描述如下:

(1)初始化:对障碍物  $L$  进行正规化,确定左右 Voronoi 多边形,起点为  $A$ ,终点为  $B$ 。

(2)查询步:从左 Voronoi 多边形  $P_1$  为起点,调用 VORDLL 的 NBROBJ 方法,得到  $P_1$  所有立即临近 Voronoi 多边形  $\{P_2, P_3, P_5, P_7, P_8\}$ 。

(3)筛选步:以是否具有  $L$  为公共边或者  $L$  上的公共点

进行筛选立即临近 Voronoi 多边形,以  $L$  为公共边的 Voronoi 多边形  $\{P_2, P_3\}$  首先排除,没有  $L$  上的公共点的 Voronoi 多边形  $\{P_7\}$  排除,余下  $\{P_5, P_8\}$ ,作为两个求 Voronoi 距离的方向。

(4)将  $P_5$  和  $P_8$  作为新的起点,重复步骤(2),(3)直到到达  $P_2$  的 Voronoi 多边形。两个方向得到的 Voronoi 多边形分别为  $\{P_1, P_5, P_4, P_{16}, P_3, P_2\}$  和  $\{P_1, P_8, P_9, P_{10}, P_{11}, P_{12}, P_{13}, P_2\}$ 。

(5)计算 Voronoi 距离,按照我们定义的 Voronoi 距离,求得  $P_1$  到  $P_2$  两条路径中的较短的路径为  $\{P_1, P_5, P_4, P_{16}, P_3, P_2\}$ ,Voronoi 距离为 5。

如果起点与障碍物没有公共边,例如求  $P_6$  到  $P_2$  的 Voronoi 距离,则需要将  $P_6$  与  $P_2$  求连线,作为在未与障碍物相遇之前的搜索方向。按照循障碍物求距离法, $P_6$  到  $P_2$  的 Voronoi 距离为 7。

## 5 基于 Voronoi 距离的空间聚类分析实例

我们将基于 Voronoi 距离的聚类分析方法应用于银行业务点的选址分析。为了验证该方法的作用,我们只考虑区域的人口数量和平均收入水平,以楼栋为生长点,单栋楼中的人口和收入水平为权进行加权 Voronoi 图生成。研究区域中共有楼栋 1393 栋,有河流为障碍物,河流被桥梁分段。



图 6 K-平均值聚类结果

图 6 表示不考虑障碍物的情况下 K-平均值聚类的实验结果,聚类产生 5 个中心,其中第 2 和 5 簇跨越河流。图 7 表示考虑障碍物的情况下,生成有障碍物的加权 Voronoi 图,按照 Voronoi 距离来进行聚类的实验结果,图 6 中的第 2、5 簇被河流分割,形成第 2、7、5、6 四个簇,第 3 簇由于有桥梁存在,所以变化不大。另外,簇中心发生了移动,簇 1 和簇 4 的

中心向左边移动。通过对楼栋数据进行分析,发现第 1、4 簇的楼栋加权 Voronoi 多边形从左至右有由小变大的趋势,由于我们采用的 Voronoi 距离对 Voronoi 多边形大小不加以考虑而导致中心的移动。从实验可以看出,通过 Voronoi 图能够很好解决有实体障碍物的空间数据挖掘中的空间聚类分析。



图 7 Voronoi 图聚类结果

**总结** 借助 Voronoi 图能够明晰地表达侧向临近关系,将 Voronoi 图应用于空间数据挖掘的聚类分析。通过构建有障碍物的空间 Voronoi 图,使用 Voronoi 距离,采用循障碍物求距离法,能够方便地解决有实体障碍物的空间聚类分析。该方法也可以在空间距离分析中应用。在生成加权 Voronoi 图的方法,以及在 Voronoi 多边形大小差距较大时求 Voronoi 距离仍需要进一步研究。

## 参考文献

- 1 陈军. Voronoi 动态空间数据模型[M]. 北京:测绘出版社,2002
- 2 Goodchild M F, Yang Shiren. A hierarchical data structure for globac information sustems [J]. Computer Vision and Geographic Image Processing, 1992, 54(1): 31~44
- 3 Green P J, Sibson R. Computing dirichlet tessellations in the plane [M]. Computer Journal, 1997, 21: 168~173
- 4 Zhao Xuesheng, Chen Jun. QTM-based algorithm for generating Voronoi-diagram of spherical objects [C]. In: Proceeding of the 3-rd ISPRS Workshop on Dynamic and Muti-dimensional GIS, 2001. 391~195
- 5 Robert J R. Delaunay triangulation and Voronoi diagram on the surface of a sphere [J]. ACM Transaction on Mathematical Software, 1997, 23(3): 416~434
- 6 Zhao Renliang, Chen Jun, Li Zhilin. Define and describe k-order spatial adjacency with Voronoi distance [C]. In: Proceeding of Dynamic nd Muti-Dimensional GIS, 1999. 77~82
- 7 Gold C M, Yang w. Spatial data management tools based on the Voronoi dynamic data model(VORDLL 1.0) User's Manual [M]. 1994
- 8 陈军. Voronoi 动态空间数据模型[M]. 北京:测绘出版社,2002
- 9 段永强, 曹健, 张申生. workflows 系统中的动态任务调度. 中国机械工程, 2002, 13(3): 233~235, 241
- 10 Tramontina G B, Wainer J, Ellis C. Applying scheduling techniques to minimize the number of late jobs in workflow systems. In: Proceedings of 2004 ACM Symposium on Applied Computing, Nicosia, Cyprus, 2004. 3. 1396~1403
- 11 Davulcu H, Kifer M, Ramakrishnan C R, et al. Ramakrishnan. Logic based modeling and analysis of workflows. In: Proceedings of the Seventeenth ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems, 1998. 25~33
- 12 Julia S, de Oliveira F F. A p-time hybrid Petri net model for the scheduling problem of workflow management systems. In: Proceedings of the IEEE International Conference on Systems, Man and Cybernetics, vol. 5. 2004. 4947~4952
- 13 欧阳昱. 基于授权约束的工作流任务指派算法. 计算机工程与应用, 2004, 36: 25~27
- 14 Shen Minxin, Tzeng G-H, Liu D-R. Multi-criteria task assignment in workflow management systems. In: Proceedings of the 36th Hawaii International Conference on System Sciences, 2003. 9
- 15 Lee K M. Adaptive resource scheduling for workflows considering competence and preference. In: Negoita M Gh, et al. eds. KES 2004, LNAI 3214, Springer-Verlag, Berlin, Heidelberg, 2004. 723~730
- 16 van der Aalst W M P, van Hee K. Workflow Management: Models, Methods, and Systems. MIT Press, 2002
- 17 Harper P R, de Senna V, Vieira I T, et al. A genetic algorithm for the project assignment problem. Computers & Operations Research, 2005, 32: 1255~1265

(上接第 160 页)

## 参考文献