

# 基于加权欧氏距离的空间 Co-location 模式挖掘算法研究

周剑云<sup>1</sup> 王丽珍<sup>2</sup> 杨增芳<sup>3</sup>

(普洱学院计算机科学系 普洱 665000)<sup>1</sup> (云南大学信息学院计算机科学与工程系 昆明 650091)<sup>2</sup>  
(玉溪师范学院信息技术工程学院 玉溪 653100)<sup>3</sup>

**摘 要** 空间 Co\_location 模式挖掘关注空间对象实例在一定区域内同时出现的关系。目前大多数研究都是把空间对象的各个实例按同等权重对待,但现实中容易发现同一类对象的不同实例其大小规模、重要程度或是影响力覆盖范围都是不一样的。因此考虑空间对象实例的影响力因素,引入加权欧氏距离阈值参与计算,能发现更具实际价值的 Co\_location 模式。

**关键词** 空间数据挖掘, Co-location 模式挖掘, 加权欧氏距离

中图法分类号 TP392 文献标识码 A

## Algorithm of Mining Spatial Co-location Patterns Based on Weighted Euclidean Distance

ZHOU Jian-yun<sup>1</sup> WANG Li-zhen<sup>2</sup> YANG Zeng-fang<sup>3</sup>

(Department of Computer Science, Puer University, Puer 665000, China)<sup>1</sup>

(Department of Computer Science and Engineering, School of Information and Engineering, Yunnan University, Kunming 650091, China)<sup>2</sup>

(School of Information Technology and Engineering, Yuxi Normal University, Yuxi 653100, China)<sup>3</sup>

**Abstract** The spatial Co-location pattern mining concerns about occurrence relationship of spatial objects' instances at the same time in a certain area. Most of the papers studied based on the spatial objects' instances have the same weight, but in reality it is easy to find that the different instances of the same type of object probably has different size, different importance, or different influence. In this paper, we consider the influence factors of the spatial objects' instances. A concept of the weighted Euclidean distance was introduced and corresponding mining algorithm was designed, in order to find more valuable spatial co-location rules.

**Keywords** Spatial data mining, Co-location pattern mining, Weighted euclidean distance

## 1 引言

空间数据库存储了大量与空间有关的数据,并具有海量、非线性、多尺度、高维和模糊性等复杂性特点。空间数据挖掘(Spatial Data Mining, SDM)是指对空间数据库中非显式存在的知识、空间关系或其他有意义的模式的提取。Ester 把空间数据挖掘技术(SDM)分为 5 个主要类<sup>[1]</sup>:空间聚类、空间分类、趋势分析、空间关联规则挖掘和空间同位模式挖掘(即 Co-location 模式)。其中前 4 种挖掘研究较早,而 Co-location 模式挖掘是目前刚起步的一个研究课题。空间关联规则挖掘与空间同位模式挖掘比较相似但不同,相同的是都表现 if-then 关系,不同的是空间关联规则的挖掘主要是通过对空间谓词的定义来发现不同空间对象和属性对象之间的相关关系,而空间 Co-location 模式挖掘则主要关注一组空间对象实例在一定区域内同时出现的概率关系。国内部分专家把 Co-location 模式也称为同位模式。

## 2 研究现状及问题提出

S. Shekhar 等人<sup>[13]</sup>最早提出基于空间相关的同位模式,它把事务概念泛化,以包括邻域集合,并且将关联规则的概念泛化为同位模式,获取同位模式规则时考虑了空间相关性。在此基础上文献<sup>[14]</sup>提出了不需要支持度剪枝,挖掘可信空间同位模式的方法。文献<sup>[15]</sup>提出了在扩展的空间对象(例如线和多边形)上挖掘空间同位规则的框架。然后 Jin SoungYoo 在文献<sup>[16]</sup>中提出了基于部分连接操作挖掘空间同位模式的算法,且在文献<sup>[17]</sup>中提出基于无连接操作的空间同位模式挖掘方法。国内专家也陆续发表了一系列研究成果,例如文献<sup>[1-12]</sup>,其中文献<sup>[1-5]</sup>主要针对同位模式挖掘算法的改进,文献<sup>[6-10]</sup>主要是针对模糊数据提出的算法,文献<sup>[11,12]</sup>则是针对稀有特征的同位模式挖掘算法。

以上研究都是把空间对象的各个实例按相同规模或是相同重要程度来对待,但是现实生活中却存在这样的情况,如同

本文受云南省教育厅科研基金项目:支持网络课程建设的数据挖掘构架研究(2013Y107),国家自然科学基金项目:非线性环境取能系统随机动力学问题研究(11265012)资助。

周剑云(1974-),女,硕士,副教授,主要研究方向为数据挖掘、远程教育, E-mail: smszzy@126.com。王丽珍 女,教授,博士生导师,主要研究方向为数据库和数据挖掘技术。

样是居民小区却有住户多少之分,同样是超市却有规模大小之分等类似情况。也就是说在考虑一定区域内(用户设定的距离值)挖掘同位模式时,应该根据各个对象实例的影响力范围的不同,设定对应的与此实例有同位关系的距离域值。本文对此做出尝试,提出了加权空间距离的 Co-location 模式挖掘算法。

### 3 相关概念

**定义 1(空间对象实例)** 设空间集合  $S$  中包含  $m$  个空间对象类集合  $S = \{C_1, C_2, \dots, C_m\}$ , 如一个城市中的 {居民小区, 医院, 超市, ...}, 其中每个对象类又有许多具体的实体, 这些实体称为该空间对象实例, 例如居民小区具体有 {长城小区, 阳光小区, ...}。

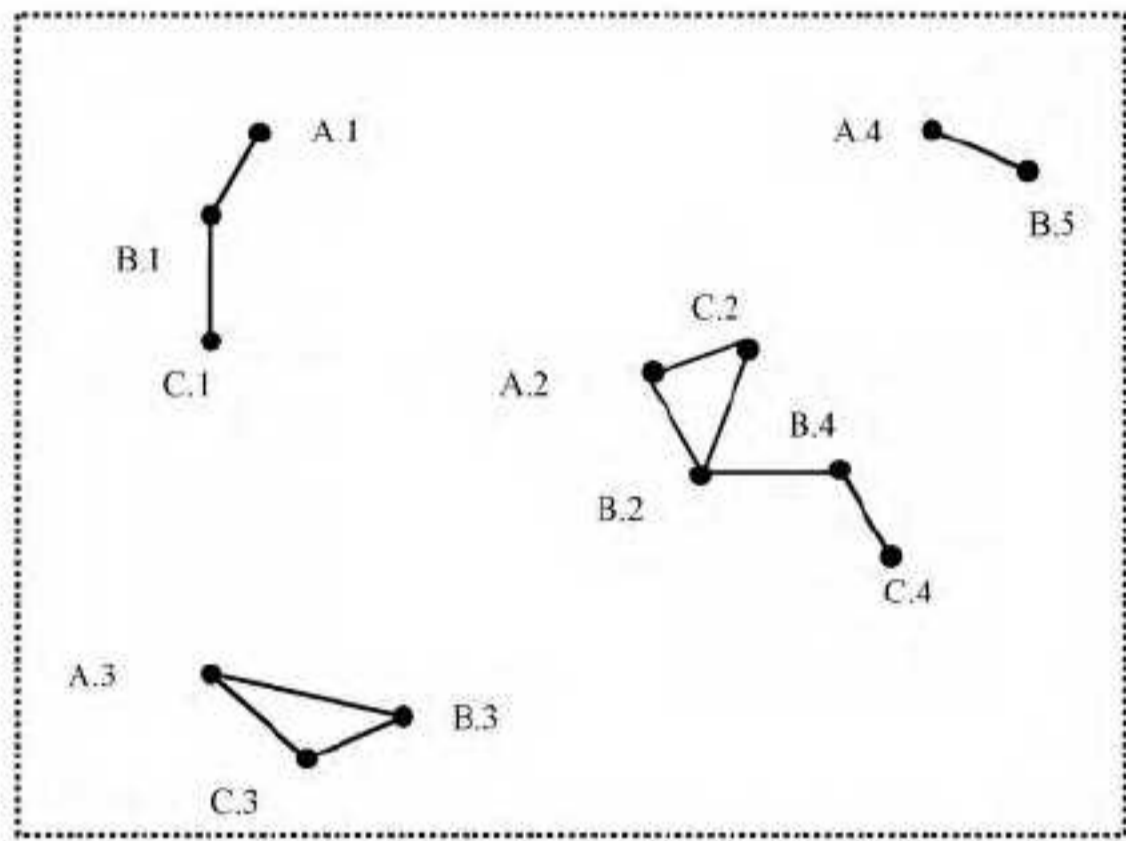


图 1 空间对象实例图

**定义 2(空间互邻)** 若两个空间对象实例之间满足某种空间关系  $R$ , 则称这两个空间对象互邻。空间关系可分为拓扑关系、度量关系以及方位关系。互邻对象之间可用实线相连。目前一般用欧几里德距离  $d$  作为空间度量关系的值要求<sup>[1]</sup>。即:  $R(A.1, B.1) \Leftrightarrow (distance(A.1, B.1) \leq d)$ 。

**例 1** 图 1 中 A.1 与 B.1 互邻, B.1 与 C.1 互邻, 等。

**定义 3(团)** 设有空间实例集合  $I = \{i_1, i_2, \dots, i_k\}$ , 若各实例之间有互邻关系, 即  $\{R(i_m, i_n), 1 \leq m \leq k, 1 \leq n \leq k\}$ , 则称  $I$  是一个团。

**例 2** 图 1 中  $\{A.1, B.1, C.1\}, \{A.2, B.2, C.2, B.4, C.4\}$  等是团。

**定义 4(空间 Co-location 模式)** 是一组空间对象的子集  $c = \{C_1, C_2, \dots, C_n\}, c \subseteq S$  ( $S$  见定义 1)。

**例 3** 图 1 中  $c = \{A, B, C\}$  是一个空间 Co-location 模式。

**定义 5(行实例、表实例)** 如果团  $I$  包含了 Co-location 模式  $c$  中的所有对象类, 并且  $I$  没有任何一个子集可以包含  $c$  中的所有对象, 那么  $I$  是 Co-location 模式  $c$  的一个行实例。Co-location 模式  $c$  的所有行实例的集合称为表实例。

**例 4** 图 1 中, Co-location 模式  $c = \{A, B, C\}$  的表实例为  $\{\{A.2, B.2, C.2\}, \{A.3, B.3, C.3\}\}$ 。

**定义 6(参与率)** 对象  $O$  的参与率 (participation ratio) 表示为  $PR(c, O)$ , 是对象  $O$  的实例在 Co-location 模式  $c$  的表实例中不重复出现的个数与  $O$  的总实例个数的比率。公式如下:

$$PR(c, O) = \prod_O (table-instance(c)) / |table-instance(O)|$$

**例 5** 图 1 中  $A$  有 4 个实例即  $\{A.1, A.2, A.3, A.4\}$ ,  $B$  有 5 个实例,  $C$  有 4 个。Co-location 模式  $c = \{A, B, C\}$  的表实

例为  $\{\{A.1, B.1, C.1\}, \{A.2, B.2, C.2\}, \{A.3, B.3, C.3\}\}$ , 因为  $A$  的 4 个实例中有 3 个出现在  $c$  的表实例中, 所以  $PR(c, A) = 3/4$ 。同理  $PR(c, B) = 3/5, PR(c, C) = 3/4$ 。

**定义 7(参与度)** 在 Co-location 模式中衡量模式支持度的标准称为参与度  $PI$  (participation index), 它是模式  $c$  中所有空间对象参与率的最小值。公式为:  $PI(c) = \min\{PR(c, O_i)\}, O_i \in c$ 。

**例 6** 若  $c = \{A, B, C\}$ , 则模式  $c$  的参与度  $PI(c) = \min\{PR(c, A), PR(c, B), PR(c, C)\} = \min\{3/4, 3/5, 3/4\} = 3/5 = 0.6$ 。

**定义 8(频繁 Co-location 模式)** 是一组在一定空间中对象实例频繁关联的空间对象子集。设  $c = \{A, B\}$ , 若  $A \Leftrightarrow B$  ( $PI$ ), 其中  $A \subset S, B \subset S, A \cap B = \Phi, PI$  为参与度 ( $S$  见定义 1),  $PI(c) \geq \min-prev$  ( $\min-prev$  是用户给定最小参与度阈值), 则称 Co-location 模式  $c = \{A, B\}$  是频繁的。

**例 7** 图 1 中若  $\min-prev = 0.6$ , 因  $PI(c) = 0.6$  (见定义 7 中计算), 即  $PI(c) \geq \min-prev$ , 所以  $c = \{A, B, C\}$  是一个频繁 Co-location 模式。

**定义 9(欧氏距离)** 设有  $n$  个样品, 每个样品测量它的  $p$  个指标, 以  $x_{ij}$  表示第  $i$  个样品的第  $j$  个指标, 则样品集表示为  $\{x_{ij}, 1 \leq i \leq n, 1 \leq j \leq p\}$ 。若  $x_i = (x_{i1}, \dots, x_{ip})$  和  $x_j = (x_{j1}, \dots, x_{jp})$  分别表示两个具有  $p$  维属性的样品, 则它们的欧氏距离为:  $d(x_i, x_j) = \sqrt{(x_{i1} - x_{j1})^2 + \dots + (x_{ip} - x_{jp})^2}$ 。

**定义 10(加权欧氏距离)** 在欧氏距离的基础上, 对每个属性根据其重要程度赋予一个权重值即为加权欧氏距离, 表示为:  $d_z(x_i, x_j) = \sqrt{w_1(x_{i1} - x_{j1})^2 + \dots + w_p(x_{ip} - x_{jp})^2}$ 。

这里考虑的权重  $w_i$  可根据各属性的覆盖范围或是规模大小等指标, 由专家系统计算得出。

### 4 基于加权欧氏距离的空间 Co-location 模式挖掘算法

基于欧氏距离的思想, 我们提出加权欧氏距离的空间 Co-location 模式挖掘算法 (Algorithm of mining spatial Co-location patterns based on weighted Euclidean distance, 以下简称 WED Co-location 算法)。本小节主要对 WED Co-location 的算法思想以及算法描述加以说明。

#### 4.1 WED Co-location 算法思想

目前大多数算法中, 不考虑数据样本中的每个对象实例的影响力、覆盖程度, 而是将它们一样对待, 用统一的欧氏距离来考查实例之间的邻近关系, 这样不能准确地反映各实例在数据集中的不同作用, 不能挖掘出更具有实际应用价值的 Co-location 模式。因为 Co-location 模式挖掘应该不仅仅只是体现考查对象之间的相近程度, 还应体现它们的内在关系, 也就是每个对象实例的重要性有所不同, 则用于度量邻近程度的距离域值也应该有所不同。如图 2 所示, 虚线圆表示的是用固定距离域值度量的邻近关系范围, 而实线圆表示的是加权计算后的邻近关系示意范围。可见对于  $A$  对象的实例 A.1, A.2, A.3 由于规模或重要程度不同, 因此它们的覆盖范围应该也不同。加权之后, 则 Co-location 模式  $c = \{A, B, C\}$  的表实例为  $\{\{A.1, B.1, C.1\}, \{A.2, B.2, C.2\}, \{A.2, B.4, C.2\}, \{A.3, B.3, C.3\}\}$ , 参与度  $PI(c) = \min\{PR(c, A), PR(c, B), PR(c, C)\} = \min\{3/4, 4/5, 3/4\} = 3/4 = 0.75$ 。若  $\min-prev = 0.7$ , 则 Co-location 模式  $c$  是频繁的。否则, 若依然采

用未加权的距离域值, 则  $c = \{A, B, C\}$  的表实例为  $\{\{A.1, B.1, C.1\}, \{A.2, B.2, C.2\}\}$ , 则  $PI(c) = \{2/4, 2/5, 2/4\} = 0.4 < \min\text{-prev}$ , 则此时模式  $c$  为非频繁的 Co-location 模式。

也就是说, 如果不考虑加权处理, 那  $\{A, B, C\}$  就不会被发现是频繁 Co-location 模式, 而通过加权算法, 以相同的参与度域值是能够发现其为频繁 Co-location 模式的。

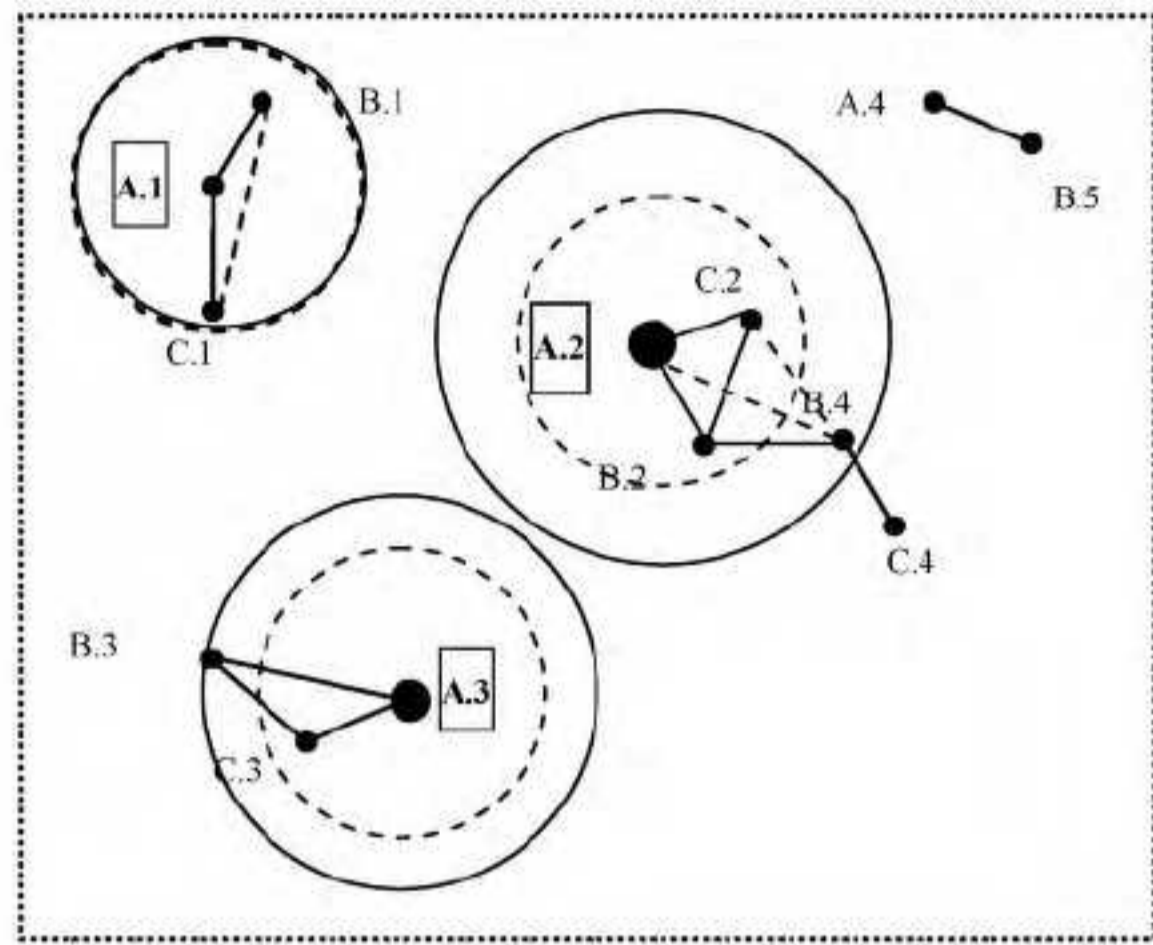


图 2 空间对象加权实例图

#### 4.2 WED Co-location 算法描述

##### 算法 1 WED Co-location 基本算法

输入:  $S = \{C_1, C_2, \dots, C_m\}$  空间对象集合;  
 $I_n = \{i_1, i_2, \dots, i_k\}$ : 空间对象  $C_n (1 \leq n \leq m)$  的实例集合;  
 $D$ : 邻近关系距离阈值;  
 $\text{Min-prev}$ : 最小参与度阈值。  
 输出: Co-location 模式集  $FP$ 。

变量:

$K$ : Co-location 的阶;  
 $D'$ : 加权后邻近关系距离阈值集合;  
 $C_K$ :  $K$  阶 Co-location 候选集;  
 $P_K$ :  $K$  阶 Co-location 频繁集;  
 $Q_K$ :  $K$  阶 Co-location 模式。

步骤:

- 扫描数据库得到考察对象类集  $S$ ;
- 根据考察对象类的权重计算各实例的加权邻近关系距离阈值集合  $D'$ ;
- 对于每个  $i_m \in I_n$ , 以加权距离阈值为半径范围扫描得到 2 阶频繁 Co-location 模式集  $P_2$ ;
- 据  $K+1$  阶频繁模式由  $K$  阶频繁产生的原则挖掘所有频繁同位模式集, 伪代码如下:
  - $FP = \emptyset; K = 2$ ;
  - while  $P_K \neq \emptyset$ 
    - $FP \cup P_K$ ;
    - $K++$ ;
    - 产生  $K$  阶候选模式  $C_K$ ;
    - 计算满足最小参与度的表实例;
    - 产生  $K$  阶频繁模式  $P_K$ ;
- 返回  $FP$ 。

例 8 以图 2 实例为例, 若考察对象类为  $A$ , 则:

第 1 步 扫描数据库得  $s = \{A.1, A.2, A.3, A.4\}$ ;

第 2 步 根据实例权重计算各欧氏加权距离, 图中虚线圆是未加权距离的覆盖范围, 实线圆为加权后的覆盖范围;

第 3 步 扫描距离加权后的 2 阶频繁模式, 我们认为在同一个实线圆内的对象与考察对象是满足邻近关系的。因此

可以得到对象实例的近邻关系, 如图 3 所示。

$c$	A B	A C	B C
加权后的邻近对象实例 $O_i$	A.1 B.1	A.1 C.1	B.1 C.1
	A.2 B.2	A.2 C.2	B.2 C.2
	A.2 B.4	A.3 C.3	B.4 C.2
	A.3 B.3		B.4 C.4
	A.4 B.5		B.3 C.3
参与率 $PR(c, O_i)$	4/4 5/5	3/4 3/4	4/5 4/4
参与度 $PI(c)$	1	3/4	4/5

图 3 加权 2 阶 Co-location 模式

若  $\min\text{-prev} = 0.7$ , 则 2 阶 Co-location 模式  $\{A, B\}, \{A, C\}, \{B, C\}$  都是频繁的。

但若是没有进行欧氏距离加权运算, 扫描得到的对象实例的近邻关系如图 4 所示。

$c$	A B	A C	B C
加权后的邻近对象实例 $O_i$	A.1 B.1	A.1 C.1	B.1 C.1
	A.2 B.2	A.2 C.2	B.2 C.2
	A.3 B.3	A.3 C.3	B.3 C.3
	A.4 B.5		B.4 C.4
参与率 $PR(c, O_i)$	4/4 4/5	3/4 3/4	4/5 4/4
参与度 $PI(c)$	4/5	3/4	4/5

图 4 传统 2 阶 Co-location 模式

第 4 步 由  $K$  阶频繁生成  $K+1$  阶频繁 Co-location 模式, 如图 5 所示。

	A B	A C	B C
2 阶频繁连接生成 3 阶模式的过程示例	A.1 B.1	A.1 C.1	B.1 C.1
	A.2 B.2	A.2 C.2	B.2 C.2
	A.2 B.4	A.3 C.3	B.4 C.2
	A.3 B.3		B.4 C.4
	A.4 B.5		B.3 C.3
3 阶 Co-location 模式	A	B	C
	A.1	B.1	C.1
	A.2	B.2	C.2
	A.2	B.4	C.2
	A.3	B.3	C.3
	3/4	4/5	3/4

图 5 从加权 2 阶频繁到 3 阶频繁 Co-location 模式

从图中可见, 因为  $A.1, B.1, C.1$  互有近邻关系, 所以  $\{A.1, B.1, C.1\}$  可以添加到 3 阶 Co-location 模式集中; 同理  $\{A.2, B.2, C.2\}, \{A.2, B.4, C.2\}, \{A.3, B.3, C.3\}$  都可添加到 3 阶 Co-location 模式集中。

由于  $PI(c = \{A, B, C\}) = \min(3/4, 4/5, 3/4) = 0.75$ , 大于  $\min\text{-prev} = 0.7$ , 因此 Co-location 模式  $\{A, B, C\}$  是频繁的。

但若是由传统计算方法也即未加权的算法进行运算, 得到的数据如图 6 所示。

	A B	A C	B C
2 阶频繁连接生成 3 阶模式的过程示例	A.1 B.1	A.1 C.1	B.1 C.1
	A.2 B.2	A.2 C.2	B.2 C.2
	A.4 B.5	A.3 C.3	
3 阶 Co-location 模式	A	B	C
	A.1	B.1	C.1
	A.2	B.2	C.2
		2/4	2/5 2/4

图 6 从传统 2 阶频繁到 3 阶频繁 Co-location 模式

由于  $PI(c = \{A, B, C\}) = \min(2/4, 2/5, 2/4) = 0.4$ , 小于  $\min\text{-prev} = 0.7$ , 因此此时 Co-location 模式  $\{A, B, C\}$  是非频繁的。

## 5 算法复杂度及实验分析

### 5.1 算法复杂度

通常空间数据集的空间对象不会很多,多的是空间对象实例,所以算法的时间复杂度与实例数紧密相关。WED Co-location 算法的时间复杂度主要集中在步骤 3 和步骤 4 即扫描空间对象集合获得 2 阶频繁模式以及为获得高阶频繁模式的连接运算上。设有  $n$  个对象,  $k_i (i=1, 2, \dots, n)$  是各对象实例数, 则算法的时间复杂度为  $O(\sum_{i=2}^n (k_{i-1} * k_i))$ 。

### 5.2 实验分析

本小节通过实验验证所提出的 WED Co-location 算法的有效性, 并与传统经典 join-based 算法挖掘结果进行比较。算法采用 VB 语言编写, 在 intel core(TM) i5-3210M CPU, 2.5GHz, 4G 内存的计算机上运行。

采用模拟数据, 设定数据空间为  $500 \times 500$ ; 对象类为 5 个; 对象实例总数为 3000; 以 A 对象类为考察类, 计算 A 对象类与其它类的加权距离公式中的  $\{w_1, w_2, \dots, w_n\}$  由随机函数产生, 在实际应用中应该是由领域专家评定给出。

#### 5.2.1 WED Co-location 算法与经典算法的对比

为了验证算法的有效性, 在改变数据实例规模的情况下, WED Co-location 算法与经典的 join-based 算法挖掘出的 Co-location 模式数量呈比较一致的变化趋势, 但又不完全相同, 如图 7 所示。这首先说明 WED Co-location 算法是有效的, 挖掘出的 Co-location 模式符合数据规模对算法的影响的规律, 也就是说得到的 Co-location 模式数差异不大, 因为距离加权时, 相比均值权重可能更大也可能更小, 所以与不加权的挖掘结果是相当的, 同时加权后算法挖掘出的 Co-location 模式在某些情况下更贴近实际情况, 反映出的折线图也就有所变化。

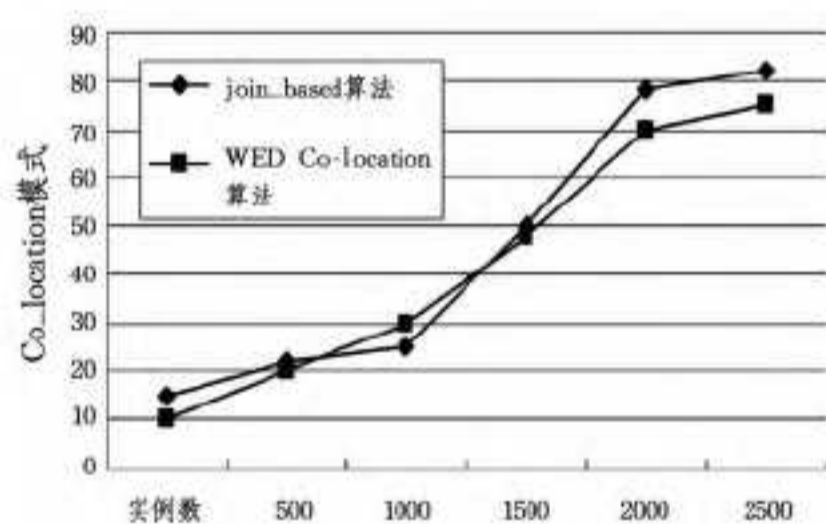


图 7 WED Co-location 算法与经典算法的对比

#### 5.2.2 min\_prev 参数阈值对算法的影响

随着参与度阈值逐渐增大, join-based 算法与 WED Co-location 算法得到的 Co-location 模式数呈逐渐减少趋势, 如图 8 所示。也就是参与度阈值要求越高, 则能满足条件的 Co-location 模式就越少, 这个规律 WED Co-location 算法也能正确表现。

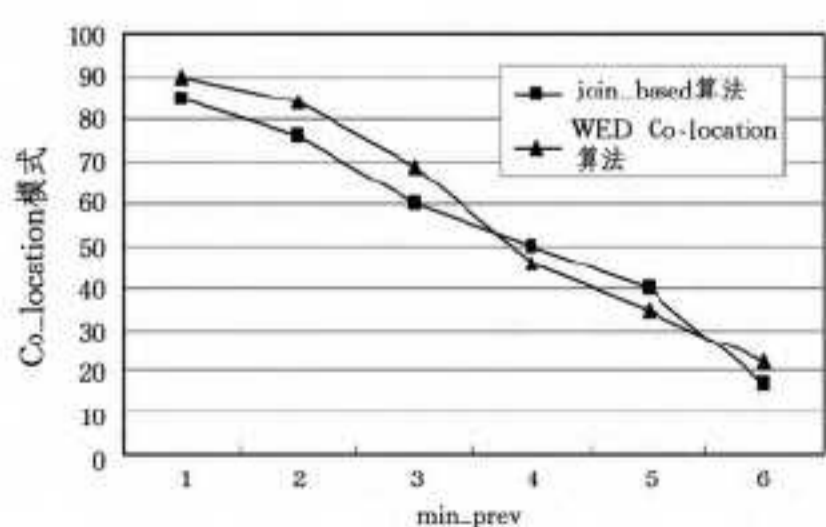


图 8 参与度域值对算法的影响

结束语 本文考虑空间对象实例的影响力因素, 引入加

权欧氏距离阈值参与计算, 能发现更具实际价值的同位模式。传统同位模式发现的是空间数据集内属于不同对象类的空间实例频繁存在的紧密相邻的模式, 其同位关系是双向的。本文的 WED Co-location 算法发现的是对某一类空间对象出现时其他对象类与其存在频繁邻近关系的模式, 这样的挖掘在某种意义上更具实际价值。但是, 算法在确定各实例的权重时会遇到比较大的困难, 可能会因为专家系统在计算该值上的不一致性而造成最终挖掘结果的不完全一致。今后, 将进一步深入研究选择怎样的距离域值才能使挖掘得到的 Co-location 模式更合理、有效、准确。

## 参考文献

- [1] 高世健, 王丽珍, 等. 基于凝聚层次聚类的 co-location 模式挖掘[J]. 广西师范大学学报, 自然科学版, 2011(29):167-173
- [2] 边馥苓, 万幼. k-邻近空间关系下的空间同位模式挖掘算法[J]. 武汉大学学报, 信息科学版, 2009(34):331-338
- [3] 王占全, 王申康, 华成. 空间分类数据同位规则挖掘算法[J]. 计算机辅助设计与图形学学报, 2005(17):2339-2344
- [4] 杨桢, 李勃, 等. 同位模式挖掘中一种改进的极大团事务化算法[J]. 云南民族大学学报, 2010(19):305-308
- [5] 周园, 云正富, 赵国庆. 基于空间同位模式的贝叶斯网构建[C]// Proceedings of 2010 International Conference on Broadcast Technology and Multimedia Communication. Volume 2, 2010
- [6] 陆叶, 王丽珍, 张晓峰. 从不确定数据集中挖掘频繁 Co-location 模式[J]. 计算机科学与探索, 2009, 3(6):656-664
- [7] 高世健, 王丽珍, 肖清. 一种基于 U-AHC 的不确定空间 co-location 模式挖掘算法[J]. 计算机研究与发展, 2011(S3):60-66
- [8] 肖清, 陈红梅, 王丽珍. 基于 DS 理论的不确定空间 co-location 模式挖掘[J]. 云南大学学报, 自然科学版, 2011, 33(S2):182-187
- [9] 欧阳志平, 王丽珍, 周丽华. 实例位置模糊的空间 co-location 模式挖掘研究[J]. 计算机科学与探索, 2012(12):1144-1152
- [10] 欧阳志平, 王丽珍, 陈红梅. 模糊对象的空间 Co-location 模式挖掘研究[J]. 计算机学报, 2011(10):1947-1954
- [11] 冯岭, 王丽珍, 高世健. 一种带稀有特征的空间 co-location 模式挖掘新方法[J]. 南京大学学报, 自然科学版, 2012(1):99-107
- [12] 陆亿红, 王子仁, 黄燕. 适合稀少空间特征的同位模式挖掘算法[J]. 浙江工业大学学报, 2007(4):408-412
- [13] S Sha-shi, Huang Yan. Discovering spatial co-location patterns: A summary of results[C]// Proceedings of the 7th International Symposium on Spatial and Temporal Databases. Redondo Beach, CA, 2001:236-256
- [14] Uang Y, Xiong H, Shekhar S, et al. Mining Confident Co-location Rules without A Support Threshold[C]// Proc. 2003 ACM-Symposium on Applied Computing. New York, NY, USA, 2003:497-501
- [15] Xiong H, Shekhar S, Huang Y, et al. A Framework for discovering co-location patterns in datasets with extended spatial objects [C]// Berry M W, Dayal U, Kamath C, et al., eds. Proceedings of the Fourth SIAM International Conference on Data Mining. Florida, USA, 2004:78-89
- [16] Yoo Jin Soung, Shekhar S. A partial join approach for mining co-location patterns[C]// Foser D P, Cruz I F, Ronthaler M., eds. 12th ACM International Workshop on Geographic Information Systems. Washington, DC, USA, 2004:241-249
- [17] Yoo Jin Soung, Shekhar S, Celik M. A Join-less Approach for Co-location Pattern Mining: A Summary of Results[C]// Proceedings of the IEEE International Conference on Data Mining (ICDM). Houston, USA, 2005