

信息科学与技术学院

042 系

17

南京航空航天大学

信息科学与技术学院

(计算机科学与技术系卷 I)

信息科学与技术学院2006年发表论文目录

序号	姓名	职称	单位	论文题目	刊物、会议名称	年、卷、期	类别
1	包磊 秦小麟	博士生 教授	042-1 042-1	随时间演化的不确定区域间拓扑关系判定	中国图象图形学报	2006. 11. 06	
2	包磊 秦小麟	博士生 教授	042-1 042-1	时空聚集计算研究进展	计算机科学	2006. 33. 01	
3	包磊 秦小麟	博士生 教授	042-1 042-1	不确定性时空查询的一种符号表示法	南京航空航天大学学报	2006. 38. 01	
4	包磊 秦小麟 张骏	博士生 教授 博士生	042-1 042-1 042-1	Reasoning the Spatiotemporal Relations Between Time Evolving Indeterminate Regions	LECTURE NOTES IN ARTIFICIAL INTELLIGENCE	2006. 393	
5	张骏 秦小麟 包磊	博士生 教授 博士生	042-1 042-1 042-1	一种支持空间拓扑分析的三维数据模型	中国图象图形学报	2006. 11. 07	
6	张骏 秦小麟 包磊	博士生 教授 博士生	042-1 042-1 042-1	一种三维空间9-交模型的简化方法	南京航空航天大学学报	2006. 38. 11	
7	张骏 秦小麟	博士生 教授	042-1 042-1	利用简化9交模型进行三维拓扑分析	计算机辅助设计与图形学学报	2006. 18. 12	
8	钟勇 秦小麟 包磊	博士生 教授 博士生	042-1 042-1 042-1	一种基于多维集的关联模式挖掘算法	计算机研究与发展	2006. 43. 12	
9	钟勇 秦小麟 郑吉平	博士生 教授 博士生	042-1 042-1 042-1	一种灵活的使用控制授权语言框架研究	计算机学报	2006. 29. 08	
10	郑吉平 秦小麟 崔新春	博士生 教授 博士生	042-1 042-1 042-1	基于数字水印的数据库角色访问控制模型	电子学报	2006. 34. 10	
11	郑吉平 秦小麟 钟勇 孙瑾	博士生 教授 博士生 博士生	042-1 042-1 042-1 042-1	基于SPN模型的可生存性D BMS中恶意事务修复算法的研究	计算机学报	2006. 29. 08	
12	黄添强 秦小麟 王钦敏	博士生 教授 研究员	042-1 042-1 外	空间离群点的模型与跳跃取样查找算法	中国图象图形学报	2006. 11. 09	
13	黄添强 秦小麟 叶飞跃	博士生 教授 博士生	042-1 042-1 042-2	基于方形邻域的离群点查找新方法	控制与决策	2006. 21. 05	
14	管致锦 秦小麟	博士生 教授	042-1 042-1	A Method for Logic Optimization of Large Number of I/O Variables	Journal of Computational Information Systems	2006. 02. 03	
15	管致锦 秦小麟	博士生 教授	042-1 042-1	Reversible Synthesis with Minimum Logic function	IEEE	2006. 02. 03	
16	叶水生 秦小麟	博士生 教授	042-1 042-1	Research on Fuzzy Kohonen Neural Network for Fuzzy Clustering	LECTURE NOTES IN COMPUTER SCIENCE	2006. 4101	

信息科学与技术学院2006年发表论文目录

序号	姓名	职称	单位	论文题目	刊物、会议名称	年、卷、期	类别
17	崔新春 秦小麟	博士生 教授	042-1 042-1	A Robust Algorithm for Watermark Numeric Relational Databases*	Lecture Notes of CIS	2006. 344	
18	朱朝晖	教授	042-1	Similarity between preferential models	Theoretical Computer Science	2006. 353. 01	
19	朱朝晖	教授	042-1	A characterization theorem for injective model classes axiomatized by general rules	Theoretical Computer Science	2006. 360. 01	
20	陈灯川 马维华	硕士生 教授	042-1 042-1	PC看门狗的原理与实现	石家庄铁道学院学报	2006. 19. 04	
21	陈灯川 马维华	硕士生 教授	042-1 042-1	基于FPGA的IDE硬盘数据AES加解密研究与 实现	电子技术应用	2006. 32. 12	
22	陈灯川 马维华	硕士生 教授	042-1 042-1	基于SG830电子语音书的设计与实现	兰州大学学报(自然科 学版)	2006. 42 (增刊)	
23	顾芸 马维华	硕士生 教授	042-1 042-1	对程控交换机中桥接时隙释放机制的设计	湖北电力	2006. 30. 05	
24	史海峰 徐涛	硕士生 教授	042-1 042-1	基于安全审计的监控系统模型的设计	计算机技术与发展	2006. 16. 06	
25	史海峰 徐涛	硕士生 教授	042-1 042-1	一种动态字符录像数据的关键字搜集算法	计算机应用	2006. 26. 03	
26	郭玉堂 徐涛	硕士生 教授	042-1 042-1	Peer-to-Peer对等网络可信性的分析和比较	计算机工程与设计	2006. 27. 24	
27	周惠 徐涛 武小川	硕士生 教授 硕士生	042-1 042-1 042-1	一种抵抗共谋攻击的数字视频水印算法	计算机应用	2006. 26. 04	
28	武小川 徐涛	硕士生 教授	042-1 042-1	基于ASM实现视频中物体的定位	航空计算技术	2006. 36. 02	
29	陈松灿 李道红	教授 硕士生	042-1 042-1	Image binarization focusing on objects	Neurocomputing	2006. 69. 16-18	
30	杨绪兵 陈松灿	博士生 教授	042-1 042-1	基于原型超平面的多类最接近支持向量机	计算机研究与发展	2006. 43. 10	
31	孙廷凯 冯爱民 陈松灿	博士生 讲师 教授	042-1 042-1 042-1	基于相关性度量的伪主成分分析	南京航空航天大学学报	2006. 38. 06	
32	陈蕾 陈松灿 张道强	硕士生 教授 副研	042-1 042-1 042-1	小世界体系的多对多核联想记忆模型及其 应用	软件学报	2006. 17. 02	
33	陈蕾 陈松灿	硕士生 教授	042-1 042-1	Distance-Based Sparse Associative Memory Neural Network Algorithm for Pattern Recognition	NEURAL PROCESS LETTERS	2006. 24. 01	
34	贾新华 王喆 陈松灿	博士生 博士生 教授	042-1 042-1 042-1	Fast screening out true negative regions for Microcalcification detectron in digital Mammograms	南京航空航天大学学报 (英文版)	2006. 23. 01	

信息科学与技术学院2006年发表论文目录

序号	姓名	职称	单位	论文题目	刊物、会议名称	年、卷、期	类别
35	刘俊 陈松灿	博士生 教授	042-1 042-1	Discriminant Common Vecotors Versus Neighbourhood Components Analysis and Laplacianfaces:A comparative study in small sample size problem	Image and Vision Computing	2006.24.03	
36	刘俊 陈松灿	博士生 教授	042-1 042-1	Non-iterative generalized low rank approximation of matrices	Pattern Recognirion Letters	2006	
37	丁军娣 陈松灿 马儒宁	博士生 教授 讲师	042-1 042-1 外	A Fast Directed Tree Based Neighborhood Clustering for Image Segmentation	“International Conference on Neural Information Peocessing” 会议交流(2006)		
38	王立松 李杰 王海峰 丁秋林	副教授 教授	042-1 042-1	Secure Data Synchronization Exchange Service Appl Ication Program Interface Based on SyncML Protocol	南京航空航天大学学报 (英文版)	2006.23.01	

随时间演化的不确定区域间拓扑关系判定

包磊 秦小麟

(南京航空航天大学信息科学与技术学院, 南京 210016)

摘要 不确定性处理是时空数据库技术研究的新领域, 现有研究成果集中在时空不确定性的表示模型方面, 缺乏不确定性对象间时空关系特别是拓扑关系的分析。提出了利用 3 维拓扑分析模型来分析 2 维运动对象的时空关系, 针对随时间演化的 2 维不确定性区域, 将 Egg/Yolk 模型中的 RCC (region connection calculus) 区域扩充至 3 维, 得到 46 种联合完备且互不相交的基本拓扑关系, 根据各基本拓扑关系的时空特性, 将 46 种基本关系归类为 21 类不确定性时空关系。

关键词 时空数据库 拓扑关系 不确定性

中图分类号: TP311.131 文献标识码: A 文章编号: 1006-8961(2006)06-0883-07

The Topological Relations Analysis for Indeterminate Regions Evolving with Time

BAO Lei, Q N Xiao-lin

(College of Information Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 210016)

Abstract The topic of uncertainty management is a new issue in spatiotemporal database research. Some efforts have already been made on the indeterminacy representation models, while the area of topological relations models for indeterminate spatiotemporal objects remain untouched. The goals of these models are to provide a set of Joint Exclusive and Pairwise Disjoint topological relations for uncertain spatiotemporal objects. This paper presents a topological relations model for indeterminate evolving 2D regions. It checks the correspondence between 3d topological relations and spatiotemporal relations, provides restrictions to make the 3d relation model fitful for spatiotemporal objects. Then it extends Egg/Yolk model to the third dimension that can describe the approximate topological relations for indeterminate evolving regions. The result is a collection of relations clusters which have different spatiotemporal nature.

Keywords spatiotemporal database, topological relation, uncertainty

1 引言

当前有关时空数据库的研究成果主要针对确定性时空对象, 即要求时空对象具有精确的空间位置、边界以及确定的时空关系。建立在此前提下, 已有研究成果包括各类时空对象的表示模型^[1]和对象间的时空关系分析模型^[2]。然而, 对于现实世界中大量随时间演化的运动对象来说, 无法获得其精确的时间和空间信息, 如高速运动的敌方目标、湖泊

的水位线以及动物的繁衍区域等等。20 世纪 90 年代末, 一些学者提出了不确定性处理在时空数据库领域的重要性^[3], 指出必须对时空对象的位置不确定性、几何属性不确定性以及运动关系的不确定性的表示和处理作深入研究。

目前时空数据库领域内不确定性处理的研究主要集中在表示模型研究上。Tossebro 扩充了模糊空间数据类型, 给出模糊时空数据类型, 包括模糊时空点和模糊时空区域的抽象数据类型和其离散实现模型^[4]。Wolfson 等人以 3 维柱状结构近似表示位置

基金项目: 国家自然科学基金项目 (69973032); 江苏省自然科学基金项目 (BK2001045)

收稿日期: 2005-03-14; 改回日期: 2005-08-25

第一作者简介: 包磊 (1977~), 男, 副教授。现在海军工程大学指挥自动化教研室工作, 为南京航空航天大学博士研究生。研究方向为时空数据库、作战仿真系统。E-mail: blnj2000@nuaa.edu.cn

不确定的运动车辆的运动轨迹^[5],在可扩充数据库管理系统 Informix DS 2000 基础上实现了 DOMINO 原型系统。基于灰集的时空数据抽象模型和离散实现模型^[6,7],侧重于时空运动趋势的预测分析。

在不确定性时空关系分析模型方面,由于不确定性同时存在于时间维和空间维上,加之时空不确定性互相依赖、互为存在条件以及时空的连续性、无限性等特点,到目前为止,可查的资料中还没有出现完整的同时考虑到时空特性和不确定性的关系分析模型。可供借鉴的相关研究领域有,不确定性空间对象间拓扑分析模型研究和确定性时空关系分析模型研究。有关不确定性空间对象间拓扑分析方法的研究有:引入模糊逻辑或者概率函数来表示不确定的空间拓扑关系^[8,9];将对象空间不确定性定义为其宽边界,在此基础上扩充 Egenhofer 的 9-intersection 模型来分析宽边界不确定区域间拓扑关系,获得了 44 种联合完备且互不相交 (J. E. P. D.) 的基本拓扑关系^[10]; Egg/Yolk 模型^[11]将不确定性空间区域表示为一对确定性区域,获得了 46 种基本拓扑关系,同样也是 J. E. P. D. 的。

有关确定性时空关系的研究有: Muller 在文献 [12] 中给出的定性分析模型,定义了 leave, reach, collide, inner, outer, intersect 共 6 种基本时空关系; Erwig 引入时空谓词作为时空关系的形式化表示,这些谓词实际上是定义在时空抽象数据类型基础上的函数; GTDM 模型^[13]将时态关系扩展到高维用以描述对象间的拓扑和方向关系,是一种离散化的定量模型。Andrea 提出的基于对象移动时段和空间关系时段的模型,分析定量的方向和距离关系^[14]。

本文对不确定性时空对象间时空拓扑关系的分析判定模型进行了研究,提出了一种随时间不断演化的不确定性区域间时空关系的形式化分析模型。其思路是将 2 维 Egg/Yolk 模型扩充至 3 维,用不确定性 3 维拓扑关系来分析不确定性时空关系。

2 随时间演化不确定区域的表示模型

不确定性时空对象的形式化定义是其抽象数据类型。时空类型是空间基本类型的时态提升。设 A 是空间类型, $Evolving(A)$ 是其对应时空类型,定义为一个从时间到 A 的函数:

$$Evolving(A) : time \rightarrow A \quad (1)$$

对于随时间演化的 2 维不确定区域类型

$Evolving(Region)$, 可表示为一组近似表示对象时空关系的函数,如图 1 所示。

$$Evolving(Region) : time \xrightarrow{\text{indeterminate}} Region \quad (2)$$

$$\forall t \in T, Evolving(Region)(t) \in URegion$$

其中, T 是区域的定义时间周期, $URegion$ 是 2 维不确定性区域类型,其含义为随时间演化的 2 维不确定区域在任意一个时刻表现为一个 2 维不确定区域。

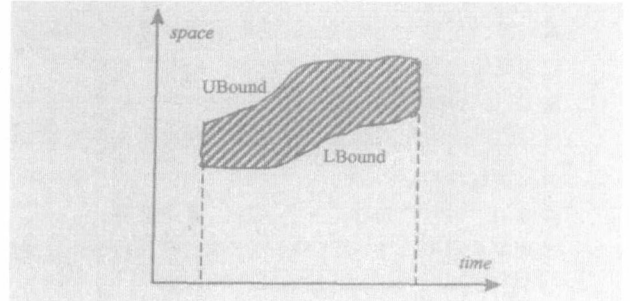


图 1 时空类型的抽象表示

Fig 1 Abstraction of indeterminate spatiotemporal types

$Evolving(Region)$ 的离散实现一般通过对时间段进行切片分段来实现。如图 2 所示,随时间演化的不确定性 2 维区域按时间分割为多个单元 unit,以 A. V. L. 树组织^[4]。在每个单元内部,区域的演化近似表示为一对上下界函数,其形式为线性插值函数、固定值或者灰差分方程^[7]。当时态不确定性与空间不确定性相互影响时,演化函数的计算还需要经过一系列合并算法和几何变形算法^[4]得到。

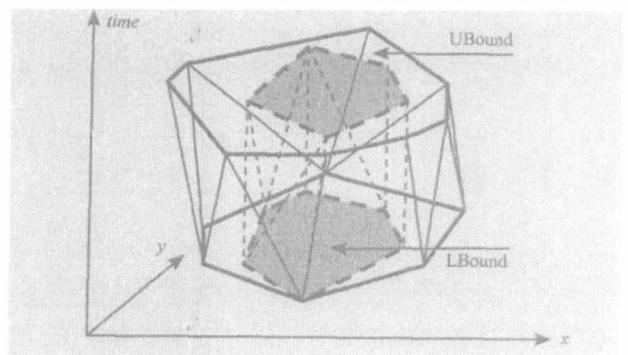


图 2 时间演化不确定性区域的离散表示

Fig 2 Discrete model of evolving indeterminate region

3 2 维 Egg/Yolk 模型及其 3 维扩展

Cohn 等人基于区域对模型给出了一种不确定简单区域的拓扑关系模型,称为 Egg/Yolk 模型^[11]。

Yolk对应下分明, Egg对应上分明, 分别相当于区域的确 定部分和整体。该模型中的区域满足 RCC (region connection calculus)公理, 该模型排除了分明区域, 并要求 Yolk 是 Egg 的一个真部分。Egg/Yolk模型利用两个区域的 Egg 和 Yolk 之间的 RCC-5 关系, 对不确定性区域间的关系进行分类, 建立了相应的 J. E. P. D. 关系集合, 这些关系根据一个 4 元素矩阵来确定, 如图 3 所示。

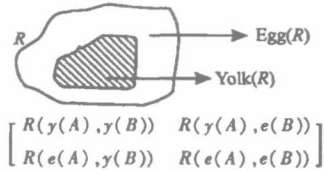


图 3 Egg/Yolk模型和其关系判定矩阵
Fig 3 Egg/Yolk model and its relationship matrix

矩阵中每个元素是一个 RCC-5 关系, 包括 DR (分离), PO (相交), PP (真部分), EQ (相等) 和 PPI (反真部分)。如图 4 所示。

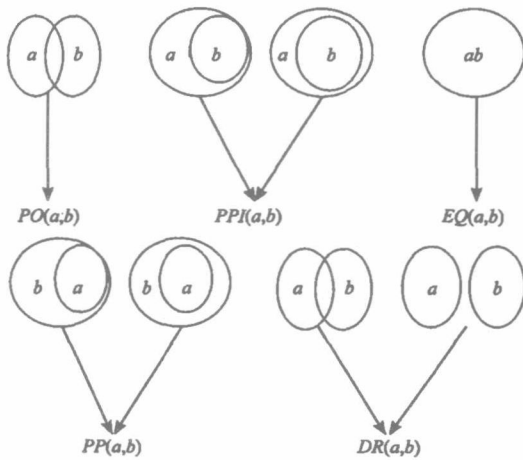


图 4 简单区域对象间的 RCC-5 关系
Fig 4 RCC-5 relations between simple regions

RCC-5 关系已经被证明是 J. E. P. D. 的。因此根据图 3 中关系矩阵, 不确定性区域间的拓扑关系有 5^4 种, 构成一个 J. E. P. D. 的关系集合。由于 Egg 和 Yolk 之间存在固有约束关系, 比如 Yolk 是 Egg 的真部分、Yolk 不能为空等等, 在 625 种基本关系中有大量不可能出现的情况, 例如 $DR(Egg(a), Egg(b)) \rightarrow DR(Yolk(a), Yolk(b))$, 所以 $\begin{bmatrix} \neg DR & * \\ * & DR \end{bmatrix}$ 就是不可能出现的关系矩阵。考虑到各类约束, 最终在 625

个关系中仅有 46 种可能发生的空间关系, 构成一个 J. E. P. D. 的关系集合, 用于描述 2 维不确定性区域间的拓扑关系。如图 5 所示。

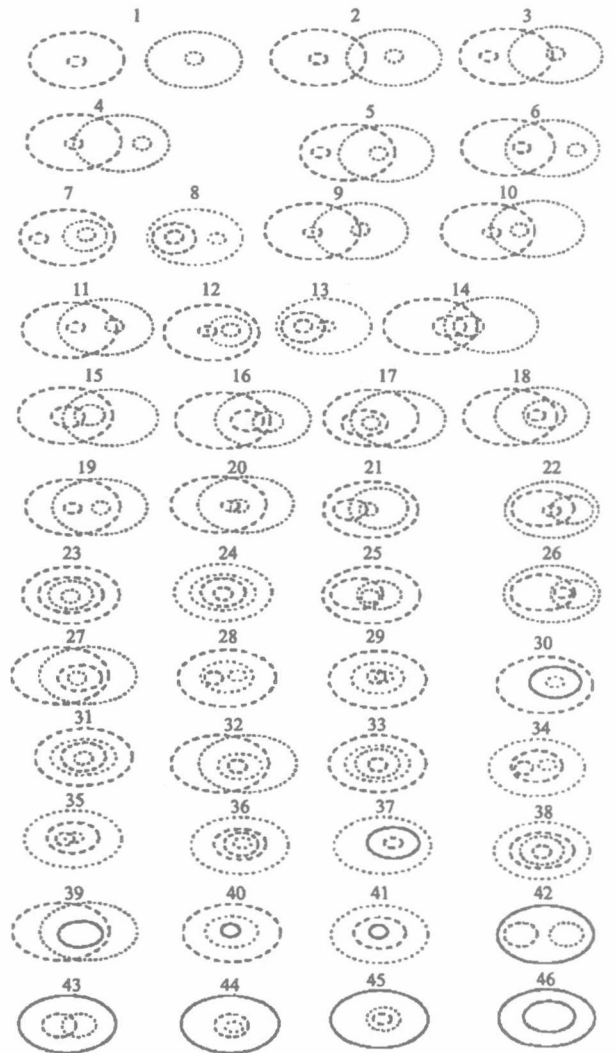


图 5 2 维简单不确定性区域对象间的关系简图
Fig 5 46 relations between 2D indeterminate regions

在第 2 节介绍的表示模型中, 2 维不确定性随时间演化区域表示为一对确定性演化区域, 可以进一步看成是以时间维作第 3 个空间维所构成 3 维空间中的纯几何对象, 也就是一对确定的 3 维空间体对象。Egg/Yolk 模型规定所有区域满足 RCC 公理, 该公理被证明可适用于任意维空间内的同维区域间的关系描述, 所以当扩充到 3 维空间后, Egg/Yolk 模型可以获得相同的 46 个 J. E. P. D. 基本关系。利用这 46 个基本关系可以描述随时间演化不确定区域间的拓扑关系。

如图 6 所示,图示的 3 维体对象 A, B 间的拓扑关系 $overlap(A, B)$ 蕴含不确定演化区域 A, B 有时候可能相交的信息。

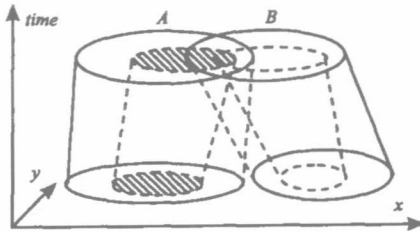


图 6 overlap ←有时可能相交

Fig 6 overlap ←sometimes possibly overlap

从直观上得出结论:某些 3 维几何体间的拓扑关系与 2 维随时间演化区域间的时空关系是相同的。引用有关拓扑空间和拓扑同构的定义,这种关系的形式化的简要定义和证明如下:

定义 1 设集合 $X, P(X)$ 是 X 的幂集,设有 $\Gamma \subseteq P(X)$, 则 (X, Γ) 为一个拓扑空间,当且仅当:

- (1) $X \in \Gamma, \emptyset \in \Gamma$
- (2) $U, V \in \Gamma \Rightarrow U \cup V \in \Gamma$
- (3) $S \in \Gamma \Rightarrow \bigcup_{A \in S} A \in \Gamma$. Γ 中的元素称为开集,其在 X 中的对应元称为闭集,简要起见,将 Γ 省略,以 X 表示一个拓扑空间。

定义 2 设 X, Y 是拓扑空间,称函数 $f: X \rightarrow Y$ 连续,当且仅当对于 Y 的每个开子集 $U, f^{-1}(U) = \{x \in X | f(x) \in U\}$ 是 X 中的闭集。称 X, Y 拓扑同构,当且仅当存在函数 $f: X \rightarrow Y$ 是连续的且有 $f^{-1}: Y \rightarrow X$ 也是连续的。

设有 3 维欧氏空间 \mathbb{R}^3 内的几何体集合:

$volum es = \{V \subseteq \mathbb{R}^3 | V \text{ 与 } \mathbb{R}^3 \text{ 内的闭球拓扑同构},$

$$\forall (\min(z | (x, y, z) \in V) \leq z \leq (\max(z | (x, y, z) \in V))) : \{(x, y) | (x, y, z) \in V\} \in region\}$$

则 $volum es$ 与随时间演化的区域是拓扑同构的,因为存在连续且逆函数也连续的函数:

$$\alpha(V) = \{(z, r) | z \in time, (\min(z | (x, y, z) \in V) \leq z \leq (\max(z | (x, y, z) \in V))), r = \{(x, y) | (x, y, z) \in V\}\}$$

从几何特性上说, $volum es$ 包含了欧氏空间上那些封闭的、连通的、无孔的 3 维体,且其在任何一个 z 坐标上与 xy 平面平行的平面相交都是一个 2 维 RCC 区域。它通过一个时间映射函数: $\alpha: volum e \rightarrow$

$Evolving(\text{region})$ 与随时间演化区域拓扑同构。

4 随时间演化不确定区域间时空关系

尽管随时间演化的不确定 2 维区域可表示为一对确定性演化区域,而确定性 2 维演化区域又拓扑同构于 3 维欧氏空间中的一些封闭连通几何体,3 维不确定性空间拓扑关系分析结果还是不能直接用于不确定性时空关系的分析。

图 7 给出了不确定性变化区域 A, B 的两种关系,从拓扑关系上看,他们具有相同的的关系矩阵,但是他们的时空关系明显不同。

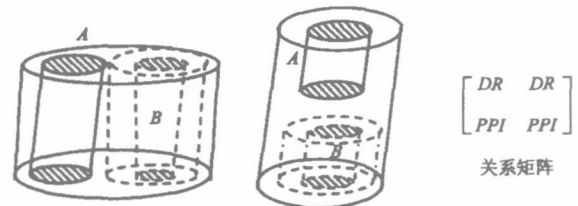


图 7 相同拓扑关系不同时空关系的情况

Fig 7 Different spatiotemporal relation with the same topological relation

为了避免这种由于对象定义时间不同而造成的特殊情况,做出以下假设:

假设 在讨论随时间演化的不确定性区域 A, B 的时空关系时,认为 A, B 具有相同的定义时间,且在其定义时间内任何一个时刻, A, B 的 $Yolk$ 不为空,即

$$t \in tim\ span(A) \cap tim\ span(B),$$

$$Yolk(at(A, t)) \neq \emptyset, Yolk(at(B, t)) \neq \emptyset$$

$tim\ span(A)$ 返回 A 的定义时间, $at(A, t)$ 返回 A 在 t 时刻的空间状态。

该假设是可满足的,因为对不满足条件的对象 A, B ,可以采用时间分段的方法将其分割为满足条件的部分来讨论,如图 8 所示。

图 8 中,经过时间分段, t_1, t_2, t_3 之间的部分满足假设条件。通过考察各个分段间的关系,可以获得对象的整体关系。

在使用 3 维不确定性空间拓扑关系分析结果分析不确定性时空关系时,还有一个问题,即时空关系不仅包含了空间拓扑关系,还具有时态语义,如图 9 所示。

图 9 中的 2 种情况具有不同的时空关系,但是其空间拓扑关系是相同的。

为此,引入关键字 $sometimes$ 和 $always$, 用以区

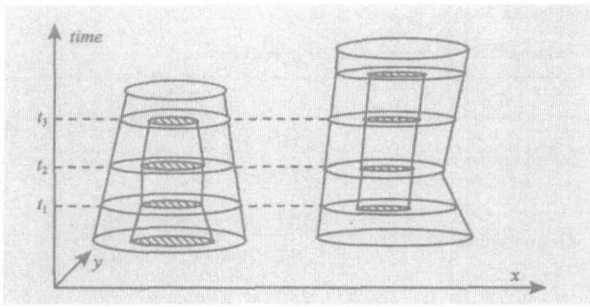
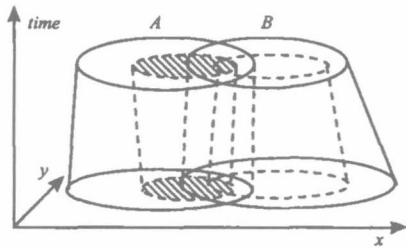
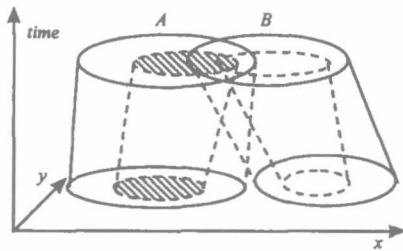


图 8 通过时间分段来满足假设条件

Fig 8 Time slicing



(a) 一直相交



(b) 有时相交

图 9 不同的时态特性

Fig 9 Different temporal relations

分这两种不同的时空关系。并建立了 3 维 RCC-5 关系和这 2 种时空关系的对应:

$$\text{Sometimes } R(A, B) \Leftrightarrow t \in \text{time span}(A) \cap \text{time span}(B), \\ R(\text{at}(A, t), \text{at}(B, t))$$

$$\text{Always } R(A, B) \Leftrightarrow \forall t \in \text{time span}(A) \cap \text{time span}(B), \\ R(\text{at}(A, t), \text{at}(B, t))$$

规则 1 $DR^{3d} \leftarrow$ always disjoint

规则 2 $PP^{3d} \leftarrow$ always inside

规则 3 $EQ^{3d} \leftarrow$ always equal

规则 4 $PPI^{3d} \leftarrow$ always Contain

规则 5 $PO^{3d} \leftarrow$ sometimes overlap

选取规则 1 证明之:

规则 1 若 A, B 满足 RCC 关系 DR^{3d} (分离), 则有时空关系 *always disjoint* (一直分离)

证明: “ \rightarrow ”,

设 $Z = \{z \in \mathbb{R} \mid (\min(z \mid (x, y, z) \in A) \leq z \leq (\max(z \mid (x, y, z) \in A)))\}$,

$$DR^{3d}(A, B) \Rightarrow \forall z \in Z, \text{disjoint}(\text{Intersectat}(A, z), \\ \text{Intersectat}(B, z)),$$

$\text{Intersectat}(A, z)$ 获得 A 在 z 位置与 xy 平面平行切平面相交得到的区域对象。使用映射函数 $\alpha: \text{volume} \rightarrow \text{Evolving}(\text{region})$, 可将此关系映射到时空对象上, 有:

$\forall t \in \text{time span}(A), \text{disjoint}(\text{at}(A, t), \text{at}(B, t))$, 得证。反之亦然, 容易证明 “ \leftarrow ” 的成立。

根据这些规则, 可以获得时变区域对之间的时空关系, 从而可对 3 维空间拓扑关系的 46 种基本关系进行分析和重新归类, 最终获得随时间演化区域对象的 21 种不确定性时空关系, 可描述 2 维不确定时变区域之间的所有可能时空拓扑关系, 其关系名和对应关系矩阵如表 1 所示。

在判定任意 2 个随时间演化的不确定 2 维区域 A, B 间的时空拓扑关系时, 只需要计算 2 对确定性 3 维空间体对象 $\text{Egg}(A), \text{Egg}(B), \text{Yolk}(A), \text{Yolk}(B)$ 间的 RCC-5 关系, 生成相应关系矩阵, 然后根据关系矩阵查表即可获得其时空拓扑关系。简单算法如下:

algorithm Judge_Relation(a, b)

let $Rcc-5_b = \{DR, PO, PP, PPI, EQ\}$; //RCC-5 关系枚举类型

let $Relationship_Clusters = \{\text{Always disjoint, Sometimes overlap, } \dots, \text{Always equal}\}$; //21 个关系组

input

a, b is the 2 indeterminate evolving regions need the relationship judgement

output

r : Relationship_Clusters

method:

Relation_Matrix: $2 * 2$ matrix of RCC-5;

Relation_Matrix[1][1] = Calculate_RCC_5(Yolk(a), Yolk(b));

Relation_Matrix[1][2] = Calculate_RCC_5(Yolk(a), Egg(b));

Relation_Matrix[2][1] = Calculate_RCC_5(Egg(a), Yolk(b));

Relation_Matrix[2][2] = Calculate_RCC_5(Egg(a), Egg(b));

$r = \text{LookUp_From_Table}(\text{Relation_matrix})$;

return r ;

end Judge_Relation

表 1 随时间演化不确定性区域间的 21 种不确定性时空关系
Tab 1 Clusters of topological relations between indeterminate evolving regions

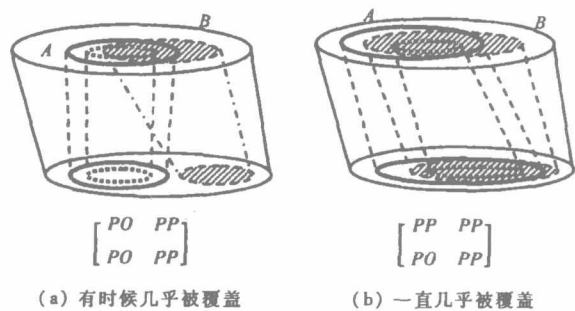
关系名	矩阵内容	关系名	矩阵内容
A always disjoint	$\begin{bmatrix} DR & DR \\ DR & DB \end{bmatrix}$	Sometimes overlap	$\begin{bmatrix} PO & PO \\ PO & PO \end{bmatrix}$
A always inside	$\begin{bmatrix} PP & PP \\ PP & PB \end{bmatrix} \begin{bmatrix} PP & PP \\ EQ & PB \end{bmatrix}$	A always contain	$\begin{bmatrix} PPI & PP \\ PPI & PP \end{bmatrix} \begin{bmatrix} PPI & EQ \\ PPI & PP \end{bmatrix}$
Sometimes possibly meet	$\begin{bmatrix} DR & DR \\ DR & PC \end{bmatrix} \begin{bmatrix} DR & DR \\ PO & PC \end{bmatrix} \begin{bmatrix} DR & PC \\ DR & PC \end{bmatrix}$	Sometimes possibly overlap	$\begin{bmatrix} DR & PC \\ PO & PC \end{bmatrix} \begin{bmatrix} DR & PP \\ DR & PC \end{bmatrix} \begin{bmatrix} DR & PP \\ PO & PC \end{bmatrix}$
A always Covered by boundary	$\begin{bmatrix} DR & PP \\ DR & PB \end{bmatrix}$	A always possibly covered by boundary	$\begin{bmatrix} DR & PP \\ PO & PB \end{bmatrix}$
A always covers with boundary	$\begin{bmatrix} DR & DR \\ PPI & PPI \end{bmatrix}$	A always possibly covers with boundary	$\begin{bmatrix} DR & PO \\ PPI & PPI \end{bmatrix}$
Sometimes boundary overlap	$\begin{bmatrix} DR & PP \\ PP & PC \end{bmatrix}$	A always boundary overlap	$\begin{bmatrix} DR & PP \\ PPI & PPI \end{bmatrix} \begin{bmatrix} DR & PP \\ PP & PPI \end{bmatrix} \begin{bmatrix} DR & PP \\ PPI & EQ \end{bmatrix}$
Sometimes strong overlap	$\begin{bmatrix} PO & PP \\ PPI & PC \end{bmatrix} \begin{bmatrix} PO & PP \\ PPI & PB \end{bmatrix} \begin{bmatrix} PO & PP \\ PPI & PP \end{bmatrix}$ $\begin{bmatrix} PO & PP \\ PPI & EQ \end{bmatrix} \begin{bmatrix} PPI & PP \\ PPI & PC \end{bmatrix} \begin{bmatrix} PP & PP \\ PPI & PC \end{bmatrix}$	A always strong overlap	$\begin{bmatrix} EQ & PP \\ PPI & PC \end{bmatrix}$
Sometimes nearly covered by	$\begin{bmatrix} PO & PP \\ PO & PC \end{bmatrix} \begin{bmatrix} PO & PP \\ PO & PB \end{bmatrix}$	A always nearly covered by	$\begin{bmatrix} PP & PP \\ PO & PB \end{bmatrix} \begin{bmatrix} PP & PP \\ PO & PC \end{bmatrix}$
Sometimes nearly covers	$\begin{bmatrix} PO & PC \\ PPI & PC \end{bmatrix} \begin{bmatrix} PO & PC \\ PPI & PP \end{bmatrix}$	A always nearly covers	$\begin{bmatrix} PPI & PO \\ PPI & PP \end{bmatrix} \begin{bmatrix} PPI & PC \\ PPI & PC \end{bmatrix}$
A always nearly fill	$\begin{bmatrix} PP & PP \\ PPI & PB \end{bmatrix} \begin{bmatrix} PP & PP \\ PPI & PP \end{bmatrix} \begin{bmatrix} PP & PP \\ PPI & EQ \end{bmatrix}$ $\begin{bmatrix} EQ & PP \\ PPI & PP \end{bmatrix}$	A always nearly filled by	$\begin{bmatrix} EQ & PP \\ PPI & PP \end{bmatrix} \begin{bmatrix} PPI & PP \\ PPI & PP \end{bmatrix} \begin{bmatrix} PPI & PP \\ PPI & PB \end{bmatrix}$ $\begin{bmatrix} PPI & PP \\ PPI & EQ \end{bmatrix}$
A always equal	$\begin{bmatrix} EQ & PP \\ PPI & EQ \end{bmatrix}$		

算法 Judge_Relation 中, Calculate_RCC_5 函数计算 2 个 3 维 RCC 区域间的 RCC-5 关系, 这部分工作已经基于 *Reals^[5]* 进行了实现, 函数 *LookUp_From Table* 根据输入的参数查表 1 获得相对应的关系类型。

5 与相关工作的比较

直接相关的工作是不确定性空间区域的拓扑分析模型^[10, 11]。我们的工作建立在文献 [11] 中提供的 2 维不确定性区域间拓扑分析模型——Egg/Yolk 模型基础上。将其扩充到 3 维空间, 并考虑了利用 3 维拓扑关系来表示时空关系的一些问题。着重分

析了时空关系的时态特性, 区分了 *always* 和 *sometimes* 的不同情况。如图 10 所示。



(a) 有时候几乎被覆盖 (b) 一直几乎被覆盖

图 10 3 维 Egg/Yolk 不能区分的情况

Fig 10 Relations can't be distinguished by 3D Egg/Yolk

在 Egg/Yolk 2 维模型中,图 10 中的两种关系同归于 *nearlycoveredby* 关系类,而在的分析结果中,它们分别属于两个不同的关系类 *sometimes nearlycoveredby*, *always nearlycoveredby*。这样就区分了两种关系,在时态约束上的不同。注意到 *Clan entini* 的宽边界 9 交模型^[10]同样也是可扩充至 3 维空间的,之所以选择 Egg/Yolk 模型作为研究基础是因为二者所得到的基本拓扑关系集合是相同的,且 RCC-5 关系更适合建立与 *sometimes* 和 *always* 关键字的对应。

移动目标数据库 (Moving objects database) 领域内也有一些可借鉴成果。文献 [15] 中规定了一些表示时空不确定性的查询关键字并进行了讨论,但其专门针对位置不确定的移动点对象,无法给出分析操作的通用计算方法。

6 结 论

分析了使用不确定性 3 维拓扑关系来表示不确定性时空关系的一些问题,将 2 维 Egg/Yolk 模型扩充到 3 维空间,分析随时间演化的不确定性区域间的时空关系,结果为 21 种基本时空关系。

还需要进一步研究方面包括:

(1) 复杂不确定性时空对象间关系的分析方法,如随时间变化的不确定性复杂区域间关系,不确定性运动点和运动区域之间的关系;

(2) 关系矩阵各元素的快速计算算法。

参考文献 (References)

- 1 Erwig M, Guting R H, Schneider M, *et al* abstract and discrete modeling of spatio-temporal data types[J]. ACM-GIS 1998: 131~136.
- 2 Erwig M, Schneider M. Spatio-temporal predicates [J]. IEEE Transactions on Knowledge and Data Engineering, 2002, 14 (4): 881~901.
- 3 Roddick J, Egenhofer M, Hoel E. Spatial, temporal, and spatio-temporal databases—hot issues and directions for Ph. D. research [J]. SIGMOD Record, 2004, 33 (2): 126~131.
- 4 Tossebro E, Nygard M. Uncertainty in spatiotemporal databases[A]. In: Proceedings of 2nd Biennial International Conference on Advances in Information Systems (ADV IS) [C], Trondheim, Norway, 2002: 43~53.
- 5 Bao Lei, Qin Xiao-lin. The overview of grey spatiotemporal data types [A]. In: Proceedings of 2nd Asian Workshop on Foundations of Software (AWFS2003) [C], Nanjing, 2003: 53~57.
- 6 Bao Lei, Qin Xiao-lin, Zhong Yong. A grey spatiotemporal model for the prediction of objects' near future evolution[J]. Journal of Image and Graphics, 2005, 10 (2): 234~240. [包磊,秦小麟,钟勇.一种用于运动趋势预测的灰色时空数据模型[J].中国图象图形学报, 2005, 10 (2): 234~240.]
- 7 Trajcevski G, Wolfson O, Chamberlain S. The geometry of uncertainty in moving objects databases[A]. In: Proceedings of the 8th Conference on Extending Database Technology (EDBT02) [C], Prague, Czech Republic, 2002: 204~218.
- 8 Schneider M. Finite resolution crisp and fuzzy spatial objects[A]. In: Proceedings of the 9th International Symposium on Spatial Data Handling [C], Beijing, 2000: 3~17.
- 9 Winter S. Uncertainty of topological relations in GIS [A]. In: Proceedings of ISPRS Commission III Symposium: Spatial Information from Digital Photogrammetry and Computer Vision [C], Bellingham: SPIE, 1994: 924~930.
- 10 Clementini E, Felice Di. Approximate topological relations [J]. International Journal of Approximate Reasoning, 1997, 16 (2): 173~204.
- 11 Cohn A G, Gotts N M. The 'Egg-Yolk' representation of regions with indeterminate boundaries[A]. In: Proceedings of Geographic Objects with Indeterminate Boundaries [C], London: Taylor & Francis, 1996: 171~187.
- 12 Muller A. A qualitative theory of motion based on spatio-temporal primitives[A]. In: Proceedings of the 6th Int'l Conference on Knowledge Representation and Reasoning [C], Trento: Morgan Kaufmann, 1998: 131~143.
- 13 Niki P, Ivan R, Kia M. Spatio-Temporal modeling in video and multimedia geographic information systems [J]. Geoinformatica, 2001, 5 (4): 375~409.
- 14 Miene A, Ubbo V. Interpretation of spatio-temporal relations in real-time and dynamic environments[A]. In: Proceedings of the Robot Soccer World Cup V [C], Seattle, Springer-Verlag, 2002: 441~446.
- 15 Wolfson O. Querying the uncertain position of moving objects[A]. In: Temporal Databases: Research and Practice [M], Berlin: Springer Verlag, 1998.

时空聚集计算研究进展^{*})

包磊 秦小麟

(南京航空航天大学信息科学与技术学院 南京 210016)

摘要 时空数据库要处理大量的数据。相对于单个时空数据来说,大量数据的聚集计算结果更有信息量。本文综述了时态聚集、空间聚集和时空聚集计算领域的研究现状,着重分析了各类时空聚集算法的研究进展。讨论了目前时空聚集计算存在的问题,并指出了今后的发展方向。

关键词 时空数据库,聚集查询,聚集函数

Research Progress in Spatiotemporal Aggregation Computation

BAO Lei QIN Xiao-Lin

(College of Information Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 210016)

Abstract Spatiotemporal databases need to process vast amounts of data. In such cases, generating aggregate information from the data set is more useful than individually analyzing every entry. In this paper, we study the most relevant techniques for the evaluation of aggregate queries on spatial, temporal, and spatiotemporal data. The research progress of spatio-temporal aggregation is summarized. The problems of current research are discussed and the future directions are pointed out.

Keywords Spatiotemporal database, Aggregate query, Aggregate function

聚集计算是数据库领域的重要研究内容之一。随着近年来时空数据库研究的逐渐深入,时空聚集计算成为该领域内的研究热点。时空数据库每日要处理海量的数据,从查询需求上看,用户一般不会对大量数据中的某个单个信息感兴趣,而是需要获得大量数据的总体变化趋势或者总体统计信息。例如,交通管制系统中,一般需要查询某个路口每小时内的车辆流量而不是某个车辆具体的位置。对于时空数据库来说,提供强大的聚集计算能力非常重要。

与传统关系数据库的聚集计算相比,时空聚集计算有其特殊性。传统关系数据聚集函数针对关系中的某个或者某几个显式属性,各个元组可以抽象为数据集合中的一个独立的点,而时空对象本身在时间和空间上占据一定的范围,时空聚集不仅发生在某些显式属性上,还会在对象所占据的时间或者空间区域上进行。由于时空信息互相依赖、互为存在条件以及时空的连续性、无限性等特点,时空聚集的计算比较困难,需要深入细致的研究,包括对时空聚集的准确描述和时空聚集算法的优化。

当前对时空聚集计算的研究已经展开,包括时态聚集计算、空间聚集计算和时空聚集计算研究,已有成果包括了对各类聚集函数的描述和分类以及聚集算法的研究。有关聚集算法的研究又包括对聚集算法的优化和聚集结果的近似计算等方面,其目的都是力图以较高的性能快速获得所需的计算结果。本文对时空聚集相关领域的研究成果进行介绍和分析,其中第1节介绍聚集函数描述和定义。第2、3、4节分别介绍时态聚集、空间聚集和时空聚集查询的描述、分类以及相关的算法实现。最后对目前时空聚集研究中存在的问题以及未来的发展方向进行了总结。

1 聚集函数及其形式化定义

数据库中的聚集计算一般通过聚集函数来实现。聚集函数考察数据库中的一组元组,概括分析其总体信息,并以一个返回值反映分析结果。SQL92 标准规定了 5 个聚集函数,分别是 Max, Min, AVG, Sum 和 Count。SQL99 在此基础上增加了 Every, Some 和 Any 三个函数。SQL-OLAP 中又规定了 18 个新的聚集函数^[1]。

聚集函数的形式化定义为:

设有关系 R , 其模式为 $\{A_1, A_2, \dots, A_n\}$, A_i 对应的定义域为 D_i , 可数集合 $Agg = \{f_1, f_2, \dots, f_n\}$ 是 R 的聚集函数集合, 对任一 $f_i \in Agg, f_i: D_1 \times D_2 \times \dots \times D_n \rightarrow D_{agg}$, 其中 D_{agg} 是聚集函数的值域。对于某聚集查询, 首先通过谓词 SP 对 R 的元组列表 A 进行划分, 再将相应聚集函数作用于划分元组 $\gamma(A)$, 最终获得结果, 即:

$$Agg_{f_i, SP}(R) = \{s \cdot f_i(\gamma(A)) \mid s \in \pi_A(R)\}$$

在对某组元组应用聚集函数进行计算之前, 一般首先对元组列表进行划分, 然后对获得的划分执行聚集函数。划分方法包括分组 (Group)、分区 (Partition) 和滑动窗口 (Slide Window) 等。

2 时态聚集计算

时态数据只在一段时间上有效, 并且随时间持续变化。对于传统关系数据库, 聚集函数应用于某组元组之上, 针对某个或者某几个显式属性展开计算。对于时态数据库来说, 时间域定义为由基本时间粒 (granules) 组成的全序集合, 而聚集计算一般是时间域上从某个粒度向更粗粒度上的一个概括。

^{*} 基金项目: 本项目由国家自然科学基金资助(69973032), 江苏省自然科学基金(BK2001045)。包磊 讲师, 博士研究生, 研究方向为时空数据库, 作战仿真系统; 秦小麟 教授, 博士生导师, 研究方向为数据库技术、GIS。

例如,“计算今年每个职工的年平均工资”。

2.1 时态聚集算法研究

现有时态聚集算法可分为2类,基于时态索引的聚集计算和无索引聚集计算。基于索引的聚集算法预先对数据作局部聚集计算,并将结果存储在外存中,在需要进行聚集计算时,利用已存储的局部聚集结果快速计算出结果。无索引聚集算法通过对数据库的一次预扫描,在内存中生成一些可用于存储局部聚集结果的数据结构来进行计算。文[2]通过对时态关系的一次扫描,建立聚集树(aggregation tree),利用深度优先搜索法可快速求得聚集结果。聚集树不是平衡树,因此其节点的插入会大大影响算法整体性能。文[3]提出2-3树,利用叶节点存储聚集结果的时间信息,2-3树是平衡树,因此其搜索效率和调整效率较高,但是由于事先需要对各元组按时间排序,因此额外开销很大。文[4]提出的PA树以AVL树组织数据,存储时间戳和局部聚集值,在计算时,中序遍历PA树得到结果。文[5]提出利用meta数组将关系元组分为各个子集,采用分治法对各个子集进行聚集计算,最终结果通过对各子集结果的归并得到。文[6~8]还给出了一些并行时态聚集算法。基于时态索引的聚集计算采用存储于外存的数据结构来提高算法效率。文[9]提出SB树,树中存储了递增的局部聚集结果和其对应的时间段,聚集结果通过对SB树的深度优先搜索得到。在调整时,SB树利于新节点的插入,但是不利于节点的删除,因为对节点的删除会破坏聚集值的递增规律。与SB树不同,MVSB树^[10]为一系列SB树,每个时间戳对应一个SB树,适合时态范围聚集计算的实现。文[11]基于B树和R树,给出了一种聚集结果近似计算算法,能够处理Sum,Count聚集函数的近似计算。

总体来说,非索引聚集算法需要预先对关系作一次预扫描,根据查询谓词在主存中生成相应数据结构,现有的算法都需要聚集值满足一定的递增规律,因此只适合某些有分布性的聚集函数。基于索引的时态聚集算法通过事先存储在外存中的索引来提高算法效率,不需要作预扫描,但是不考虑属性上的谓词操作,只能处理作用于单值数据上的聚集函数,如SQL92中所规定的5类聚集函数;聚集函数的最终结果都是通过对外部聚集结果的组织得到,对于无法通过局部结果组合得到的整体聚集函数也无法进行处理,如Median。

2.2 基于数据流的时态聚集算法研究

数据流是数据的有序序列。由于数据流中的每个数据都带有时间戳以记录该数据的产生、接收时间,数据流可看作某种时态数据。数据流在网络控制、通信传输、Web应用以及传感器网络等方面有较广泛的应用。由于数据流包含大量数据,一般来说数据流中的数据只使用一次就被丢弃或者存档,对数据流的聚集计算相对于某个单独数据的查询更加重要^[26]。Datar等人提出了对数据流使用指数直方图来求取Count和Sum聚集函数的近似解的算法^[12],Zhang提出依据粒度层次对数据流进行时态聚集的机制,基本思想是根据数据的已存储时间采用不同粒度作聚集计算,已存储时间越长,粒度越粗^[13]。

3 空间聚集计算

空间聚集根据不同空间粒度对数据元组进行聚集计算,空间数据的聚集查询一般需要先根据对象的空间范围进行分类和划分得到某个元组集合,然后再在此元组集合上执行相应的聚集函数得到最终结果。典型的空间聚集查询的例子为

“计算四川省内各个县的森林覆盖面积”。其中“四川省内”为空间谓词,“各个县”指需要按每个县的空间范围进行划分。

目前已有的空间聚集算法一般都是返回某个方形区域内的所有空间对象的聚集函数值,称为Box Aggregation。其算法都是通过空间索引来进行计算。文[14]指出空间对象的拓扑关系一般不能保证满足聚集函数的分布特性,事先存储的局部聚集值不能直接用于整体聚集结果的计算,必须依照一定的规则对空间对象进行分解,才能得到可用的预聚集结果。文[15]提出aR树在R树节点的最小外接矩形MBR内标注了其所包含的所有对象的聚集值,这样在聚集计算时,可以直接引用该信息,不用对每个对象进行聚集计算。文[16]对aR树进行了优化,给出了专门用于求取Min,Max聚集函数的MR树。对于Sum,Count,Avg等聚集函数,文[18]中使用了递增的聚集索引,包括用于内存的静态ECDF树和用于外存的ECDF-B树^[17],针对ECDF树节点插入成本较高的缺点,提出的BA树可以有效地对节点数目进行调整。文[19]提出的MRA树适合于多维空间中点对象的聚集计算,以节点局部聚集结果对整体结果的不确定程度的影响因子作为该树的遍历顺序,在搜索的过程中可以不断修正结果。文[20]的aP树借用时态聚集算法处理二维空间中的点对象的聚集计算,其最大优点是遍历时间与对象数目无关。

对于空间聚集来说,目前已有的研究成果都是返回某个区域内的所有空间对象的聚集函数值(Box Aggregation),所支持的空间谓词有限,在利用空间谓词对元组进行划分之后,对象的空间信息将丢弃掉,不参与聚集计算,因此仅支持具有分布特性的聚集函数的计算。

4 时空聚集计算

时空对象同时在时间上空间上占据一定的区域,对象的显式属性和空间范围都会随时间不断变化,这种变化可以是离散的,也可能是连续不断的。时空聚集根据不同粒度对数据元组进行进行分类和划分得到某个元组集合,然后再在此元组集合上执行相应的聚集函数得到最终结果。而时空粒度是时间粒度和空间粒度的结合,如“每年每省”,“每小时每平方米”典型的时空聚集查询的例子为“计算10年内中国各省的年平均森林覆盖面积”。在此查询中,“10年内”为时态谓词,“中国境内”为空间谓词,“各省”指需要按省进行空间划分。

针对时空聚集算法的研究目前比较少,相关的研究开始于2002年,Zhang等人在其基于数据流的聚集算法基础上作了扩充,其基本思路是对时空对象进行降维分解,其各个分解部分通过ECDF树或者BA树来完成其聚集计算^[18,21]。Papadias给出了一种基于索引的时空聚集算法^[22],他将空间对象按所在区域进行分组,将各个静态区域以R树组织起来,每个区域对应的时态信息以一棵B树存储,最终获得的数据结构称为aRB树,对aRB树的进一步扩充还有aHRB树和a3DRB树,它们与aRB树的区别在于用于分组的区域可以是动态划分的。Papadias的aRB树无法处理在查询区域内同一时间段内重复出现的对象,这个问题被称为“Distinct counting problem”。为了解决这个问题,Tao对aRB树进行了修改^[23],针对Count和Sum聚集函数,解决了该问题。Sun针对连续运动的对象,提供了聚集函数近似结果的计算方法^[24],它将二维空间划分为 $w \times w$ 个网格单元,每个单元内存入了当前时间内包含的对象数目,通过其提出的数据结

构“适应性多维直方图”(Adaptive Multidimensional Histogram),可快速计算 Count 聚集函数。

时空聚集算法的研究工作集中在对空间聚集算法的扩充方面,因此其算法都是基于相应的时空索引来完成,在利用时空谓词对时空关系进行了划分之后,这些算法舍弃了参与聚集计算对象的时态信息和空间信息,因此才会发生“Distinct count”之类的问题。与对应的空间聚集算法类似,现有的时空聚集算法依赖于预聚集生成的局部聚集结果,只适合几种具有分布特性的聚集函数的计算。

5 存在问题和将来研究方向

本文在简要介绍时空聚集的基础上,对近年来有代表性的时态聚集、空间聚集和时空聚集的研究工作进行了重点介绍。在时空聚集研究领域,单独的时态聚集方面的研究已经比较充分,提出了一些很有价值的算法。但是在空间聚集和时空聚集方面的研究还远远不够,尤其是时空聚集的研究还处于“萌芽”阶段。

现有的时空聚集算法分为2类,基于索引的聚集算法和非索引聚集算法。非索引算法要求在聚集计算之前进行预聚集,获得局部聚集结果,基于索引的聚集算法利用事先存储的索引信息进行聚集计算。通过预聚集,聚集算法可以大大提高,但是这种依赖于预处理所获得的局部聚集值的算法无法处理 Median 等非分布的聚集函数。在文[25]中,Palpanas 在研究传统关系数据库中的聚集查询时注意到了这个问题,提出了适用于所有聚集函数的通用预聚集处理方法,但是其方法只可应用于传统关系数据库中。

现有的时空聚集算法还有一个主要缺点是,在利用时空谓词对关系进行划分之后,算法舍弃了对对象的时空信息进行聚集计算,这样不仅会导致一些错误的发生,而且往往这些时空信息本身才是用户所关心的。例如“返回某暴风云团的最大时速,以及达到最大时速的具体时间和地点”。

通过总结近年来时空聚集的研究进展,我们认为以下几方面是未来研究发展的方向:①支持各类聚集函数的统一聚集算法;②支持各类时态谓词、空间谓词和时空谓词的时空聚集查询处理;③针对某类或者某个聚集查询的高效算法。

参考文献

- Melton J. Advanced SQL, 1999. Understanding Object-Relational and Other Advanced Features. The Morgan Kaufman Series in Data Management Systems. Morgan Kaufmann Publishers, San Francisco, CA, 2003
- Kline N, Snodgrass R T. Computing Temporal Aggregates. In: Proc. of the Intl. Conf. on Data Engineering, Taipei, Taiwan, March 1995. 222~231
- Kline R N. Aggregation in Temporal Databases. PhD thesis, University of Arizona, Tucson, Arizona, May 1999
- Kim J S, Kang S T, Kim M-H. On Temporal Aggregate Processing based on Time Points. Information Processing Letters, September, 1999, 71(5-6): 213~220
- Moon B, Vega Lopez I F, Immanuel V. Scalable Algorithms for Large Temporal Aggregation. In: Proc. of the Intl. Conf. on Data Engineering, San Diego, CA, March 2000. 145~156
- Moon B, Vega Lopez I F, Immanuel V. Efficient Algorithms for Large-Scale Temporal Aggregation. IEEE Transactions on Knowledge and Data Engineering, May-June 2003, 15(3): 744~751
- Gendrano J A G, Huang B C, Rodrigue J M, et al. Parallel Algorithms for Computing Temporal Aggregates. In: Proc. of the Intl. Conf. on Data Engineering, Sydney, Australia, March 1999. 418~427
- Ye X, Keane J A. Processing Temporal Aggregates in Parallel. In: IEEE International Conference on Systems, Man, and Cybernetics, Orlando, FL, Oct. 1997. 1373~1378
- Yang J, Widom J. Incremental Computation and Maintenance of Temporal Aggregates. In: Proc. of the Intl. Conf. on Data Engineering, Heidelberg, Germany, April 2001. 51~60
- Zhang D, Markowetz A, Tsotras V J, et al. Efficient Computation of Temporal Aggregates with Range Predicates. In: Proc. of the ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems, Santa Barbara, CA, May 2001. 237~245
- Tao Y, Papadias D, Faloutsos C. Approximate Temporal Aggregation. In: Proc. of the Intl. Conf. on Data Engineering, Boston, USA, March 30-April 2, 2004
- Datar M, Gionis A, Indyk P, Motwani R. Maintaining Streams Statistics over Sliding Windows (Extended Abstract). In: Proc. of the annual ACM-SIAM Symposium on Discrete Algorithms, San Francisco, CA, January 2002. 635~644
- Zhang D, Gunopulos D, Tsotras V J, Seeger B. Temporal Aggregation over Data Streams using Multiple Granularities. In: Proc. of the Conf. on Extending Database Technology, Prague, Czech Republic, March 25-27 2002. 646~663
- Pedersen T B, Tryfona N. Pre-aggregation in Spatial Data Warehouses. In: Proc. of the International Symposium on Advances in Spatial and Temporal Databases, Redondo Beach, CA, July 12-15 2001. 460~480
- Papadias D, Kalnis P, Zhang J, Tao Y. Efficient OLAP Operations in Spatial Data Warehouses. In: Proc. of the Intl. Symposium on Advances in Spatial and Temporal Databases, Redondo Beach, CA, July 12-15 2001. 443~459
- Zhang D, Tsotras V J. Improving Min/Max Aggregation Over Spatial Objects. In: Proc. of the ACM-GIS Conference, Atlanta, GA, November 2001. 88~93
- Bentley J L. Multidimensional Divide-and-Conquer. Communications of the ACM, 1980, 23(4): 214~229
- Zhang D, Tsotras V J, Gunopulos D. Efficient Aggregation Over Objects with Extent. In: Proc. of the ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems, Madison, WI, June 2002. 121~132
- Lazaridis I, Mehrotra S. Progressive Approximate Aggregate Queries with a Multi-Resolution Tree Structure. In: Proc. of the ACM-SIGMOD Conference, Santa Barbara, CA, May 2001. 401~412
- Tao Y, Papadias D, Zhang J. Aggregate Processing of Planar Points. In: Proc. of the Conf. on Extending Database Technology, Prague, Czech Republic, March 25-27 2002. 682~700
- Zhang D. Aggregation Computation over Complex Objects, [PhD thesis]. University of California, Riverside, August 2002
- Papadias D, Tao Y, Kalnis P, Zhang J. Indexing Spatio-Temporal Data Warehouses. In: Proc. of the Intl. Conf. on Data Engineering, San Jose, CA, February 26-March 1 2002. 166~175
- Tao Y, Kollios G, Considine J, Li F, Papadias D. Spatio-Temporal Aggregation Using Sketches. In: Proc. of the Intl. Conf. on Data Engineering, Boston, USA, March 30-April 2, 2004
- Sun J, Papadias D, Tao Y, Liu B. Querying about the Past, the Present, and the Future in Spatio-Temporal Databases. In: Proc. of the Intl. Conf. on Data Engineering, Boston, USA, March 30-April 2, 2004
- Palpanas T, Sidle R, Cochrane R, Pirahesh H. Incremental Maintenance for Non-Distributive Aggregate Functions. In: Proc. of the VLDB Conf., Hong Kong, China, August 2002. 802~813
- Qiao L, Agrawal D, Abbadi A E. RHist: Adaptive Summarization over Continuous Data Streams. In: Proc. of the ACM-ClK Conference, McLean VA, November 2002. 469~476
- Tao Y, Papadias D. The MV3R-Tree: A Spatio-Temporal Access Method for Timestamp and Interval Queries. In: Proc. of the VLDB Conf., Roma, Italy, Sep. 11-14, 2001. 431~440

不确定性时空查询的一种符号表示法

包磊 秦小麟

(南京航空航天大学信息科学与技术学院, 南京, 210016)

摘要: 分析了时态查询、时空查询和不确定性时空查询的各类查询形式及其查询特性, 以 BNF (Backus-normal form) 范式形式给出不确定性时空查询语义的统一符号表示法, 并对典型查询进行举例说明。该表示法建立在时态谓词、时空谓词和不确定性时空谓词基础上, 可消除由不同表现形式带来的语义混淆, 以对各类不确定性时空查询进行统一表示。该方法还可用于不确定性时空数据库原型系统的实现和不确定性时空查询的进一步研究。

关键词: 时空数据库; 查询语言; 不确定性; 符号表示法

中图分类号: TP311.131

文献标识码: A

文章编号: 1005-2615(2006)01-0101-05

Notation for Syntax of Indeterminately Spatiotemporal Queries

Bao Lei, Qin Xiaolin

(College of Information Science and Technology,

Nanjing University of Aeronautics & Astronautics, Nanjing, 210016, China)

Abstract: Different forms and characteristics of temporal, spatiotemporal and indeterminately spatiotemporal queries are discussed. A new notation for indeterminately spatiotemporal queries is proposed. The notation is based on the temporal, spatiotemporal and indeterminate spatiotemporal predicates, and combines indeterminate predicates on the time dimension with predicates on explicit attributes and spatial dimensions. The proposed notation in the BNF form is shown by examples of various queries. Compared with the related work, such as spatiotemporal query language (STQL) and future temporal logic (FTL), it is proved that the notation is simple and extensible, and provides a consistent way for indeterminately temporal or spatiotemporal queries. Results show that the notation can be applied to prototype implementation.

Key words: spatiotemporal database; query language; indeterminateness; notation

引言

时空数据库是处理随时间不断变化的空间对象的数据管理系统, 由时态数据库、空间数据库发展而来, 在地理信息系统相关领域有广泛的应用。从 20 世纪 90 年代初开始, 有关时空数据库的研究工作陆续展开并取得了一定的成果, 其主要成果包括: 时空本体研究, 涉及到对时间、空间本体论及其

相互作用的研究; 时空数据模型研究, 包括表示模型和分析模型研究; 时空数据查询语言的研究。

由于测量误差, 数字化误差以及不完整信息的存在, 不确定性广泛存在于时空数据库应用系统中。某城市卫星导航系统中, 用户需要使用 PDA 查询出通过交通网距离最近的旅馆, 但是由于 GPS 的定位误差, 系统也许无法识别出用户所在道路。军用指挥控制系统中, 通过观通设备和情报

基金项目: 国家自然科学基金(69973032)资助项目; 江苏省自然科学基金(BK2001045)资助项目。

收稿日期: 2005-03-04; **修订日期:** 2005-06-05

作者简介: 包磊, 男, 博士研究生, 副教授, 1977 年 4 月生; 秦小麟(联系人), 男, 教授, 博士生导师, Email: qinxcs@nuaa.edu.cn