

以事件为核心的面向对象时空数据模型

林广发,冯学智,王雷,都金康

(南京大学城市与资源学系,江苏南京 210093)

An Event-centric Object Oriented Spatio-temporal Data Model

LIN Guang-fa, FENG Xue-zhi, WANG Lei, DU Jin-kang

(Department of Urban & Resources Science, Nanjing University, Nanjing 210093, China)

Abstract Rather than a mathematic approach, the intrinsic changing law of spatio-temporal phenomena of the real world should be concerned to attack the temporal problem in GIS. The event is a fundamental factor that results in data's change. But each event only affects the data within a certain range, which is intrinsically corresponds with the cohesion between object properties. And the cohesion determines the object granularity to be marked with time. To find a balance between data redundancy and system complexity, the law must be complied with in designing a spatio-temporal data model. According to this point of view, the paper analyzed several existing typical spatio-temporal data models and discussed a few correlative questions briefly. And then proposed an event-centric object-oriented spatio-temporal data model. In an application example of real estate information system, the model is proved to be good for establishing a foundation for further study of model-abstract, prediction and decision-making.

Key words spatio-temporal data model; event; cohesion; object granularity; real estate information system

摘 要 GIS 中时态问题的解决不仅要在数学上寻求方法的支持,更应该在现实世界的时空现象中去探求其内在的变化规律。事件是驱动时空数据发生变化的根本原因,但每一次事件对数据对象的影响总是局限于一定的范围,这内在地表现为对象的属性之间的时空内聚性并决定了时态数据管理的对象粒度。时空数据模型的设计必须参照这一规律才能在冗余与效率之间取得平衡。在简要分析现有的几种典型时空数据模型的基础上,根据上述观点设计了一种以事件为核心的面向对象时空数据模型。房产信息系统的设计与实现证明,该模型对于时空数据管理效果较好,为进一步抽取变化的模式和预测决策奠定了基础。

关键词 时空数据模型;事件;内聚性;对象粒度;房产信息系统

1 引 言

GIS 描述的现实世界永远处于不断变化之中,但“传统 GIS 却只能以静态的方式表现它^[1]”。随着 GIS 应用的日益广泛和深入,它在

时态方面的功能缺陷与现实应用对过程模拟、趋势预测的迫切需求之间的矛盾也日益突出。虽然自 20 世纪 80 年代中期以来,GIS 界和计算机科学界都分别从不同方向对此相关的问题作了大量研究并取得了许多重要成果,但成熟的时态地

理信息系统(Temporal GIS, 简称 TGIS)模型与商品化的时空分析模块仍未出现^[1]。有些学者认为, TGIS 问题的最终解决要待到“可与拓扑论相类比的”全新数学思想的出现^[2]。

但笔者认为, 时态问题的解决不仅要在数学上寻求方法的支持, 更应该在现实世界的时空现象中去探求其内在的变化规律。从近年来 TGIS 的研究情况看, 在经历了时空快照、基态修正、时空复合等模型的讨论之后, 国外已开始更侧重于时空变化的内部规律的研究与表达^[1, 3~5], 而国内在具体实现方法如非第一范式^[6, 7]、面向对象^[8, 9]等方面不断深入研究之时亦已注意到时空现象自身的规律性^[10~12]。文献[1]提出: “关注具有共同的时空行为的一组特定的空间现象是一种富有前景的方法”, 本文在分析、借鉴前人的时空数据管理方法的基础上, 进一步研究这种共同的时空行为产生的根本原因——对象的时空内聚性及其与事件的影响范围之间的关系, 并由此针对房产信息系统中的时空数据管理问题, 给出了一个初步的应用模型。

2 已有的时空数据模型

2.1 时空数据模型必须表达的基本内容

任何时空数据表示方法都必须与具体的应用目的相联系。图 1 所示的必须以时态为基础的一系列重要功能中, 时空分析是 GIS 应用从初级阶段跨越到高级阶段的关键环节。而实体对象之间的因果关系是最重要但也是最勾杂的时态关系, 建立因果关系通常是绝大多数时空分析的终极目标^[4], 这也是作进一步的预测决策的前提。但对对象的状态变化过程仅仅反映了状态之间的时间先后关系, 并不能说明前后状态之间存在因果关系, 在状态序列分析与因果推导之间还需要一个关键环节——驱动状态变化的事件。所以, 时空数据模型既要表达空间实体及实体间的空间关系^[13], 也要表达实体间的时间序列关系, 最后还要记录直接导致状态变化的事件。

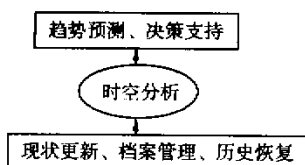


图 1 时空数据管理的主要功能模块

Fig. 1 Main functions of spatio-temporal GIS

2.2 事件的定义

“事件”是一个仍然在探讨中的概念。蒋捷等学者在总结以往的事件概念后, 认为“人们对事件与空间目标之间的因果关联表达等问题的认识尚不清楚, 事件不但是时空目标状态开始或终结的标志, 而且还表示状态变化的原因”, 并由此将土地审批、划拨过程看做是由一系列“原子事件”构成的几个“复合事件”, 提出了基于事件的土地划拨时空数据库的设计思想^[12], 将事件的研究向前推进了一步。但笔者认为该文中的“事件”仅仅是政府土地管理职能中固定的日常业务程序, 而不是导致土地现状变化的一个事件(event), 对应的状态相当于办公过程中的一个事务(transaction)未完成时数据的各个中间版本管理环节, 它更适用于数据的完整性管理。而土地利用类型发生变化的真正原因往往隐含在申请、审批的文本中, 需要从中作进一步的语义抽取。

每个对象的产生、发展、灭亡都有其直接原因, 本文把这种驱动对象发生变化的直接原因狭义地定义为该对象所具有的事件, 即原子事件, 而具有因果关联的多个原子事件就组合成为一个复合事件, 它对应于一个更宏观的对象的变化。事件的这种等级嵌套划分即事件的尺度是相对的, 它决定于具体研究目标的对象尺度及对变化过程的研究详细程度。

2.3 时态数据管理的对象粒度

对于时态信息的存储, 除了将时间看做另一维的如 4 维空间栅格十六叉树编码^[14]等方法外, 其他方法都将时间看做一个属性, 涉及如何标记时间值(或数据版本)的问题。这又包括空间数据和属性数据 2 个方面。

对空间数据, 最早的快照方法将时态标记在区域快照上。为了减小冗余, 基态修正^[3]、Spatio-temporal Object^[15]等模型都将时态标记在空间数据的点、线、面等要素的几何描述上, 国内龚健雅等在面向对象的概念下进一步发展了这种方法^[8, 9]。对此, 文献[9]认为: “在拓扑数据中考虑时态数据是极为复杂和效率低下的”。对属性数据的时态标记, 数据库领域的学者具有更长的研究历史, 他们大多集中在关系模型中第一范式(1NF)与非第一范式(N1NF)的讨论上。一般认为, N1NF 是未来的发展方向^[16, 17]。但 N1NF 的最大缺点是把同时发生的变化存储在树形结构不同层次的结点上, 给时态查询分析带来了很大的

困难。陈军等按变化发生的同时性将属性归并分组,得到了经时间规范化处理后的 NINF 关系表^[6]。这样以较小的冗余代价提高了系统效率,下文将对这种规范化的内涵作进一步的探讨。

这些方法实际上都包含了一个时态数据管理的“对象粒度^[18]”问题,即对象的结构会随着粒度的细化而趋于复杂,而数据冗余却随之减小。因此“增大对象的粒度是面向对象方法中用来降低软件复杂性的基本技术^[19]”,但如果粒度太大,单元内的局部变化将导致整个单元重复存储,这是数据冗余的主要来源。因此,时态数据管理的对象粒度与数据冗余、系统复杂性 3 者之间存在如图 2 所示的相关关系。

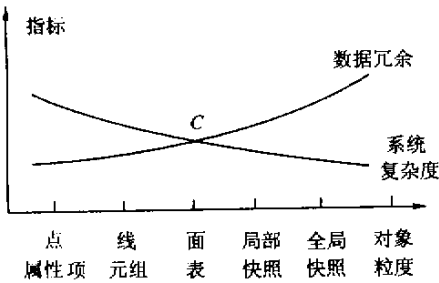


图 2 对象粒度与数据冗余、系统复杂性之关系
Fig. 2 The relation between object granularity, data redundancy and system complexity

图 2 中平衡点 C 所对应的粒度并不是固定的,可因具体应用问题的不同而变化,而且当应用系统有特别要求时,在数据冗余与系统效率之间可有一定的取舍。问题是,C 点的确存在规律性吗?存在。在 TGIS 中,事件是驱动数据发生变化的根本原因,但每一事件对实体对象的影响总是局限于一定的范围,如房产中房屋的扩建、拆除与产权变更等事件对房屋的影响范围都不相同而且相对固定。这影响范围内在地表现为时空数据的“空间/属性内聚性和时态内聚性^[8]”,只有依照这种内聚性划分得到的对象粒度才是平衡点 C 所对应的粒度。事实上,对象粒度在时态中的重要性还远不止于此,不恰当的粒度划分将弱化对象的结构完整性,无法表达变化的真正对象。这在时空分析上具有重要意义。

2.4 变化前后状态的关联追溯

时间在人的意识中表现为状态的序列。因此, TGIS 中时态问题的关键不在于如何将时间值标记在状态中,而在于如何表达状态的变化过程。这方面常用的方法是父子对象之间的指针链

接^[11]。常征等则提出了一种新的方法,即通过分析地块变更过程中的空间相交性和时间相接性来构造查询地块链的约束条件^[10]。但这种方法对于房屋这样的空间实体却不适用,因为房地产行业很常见的一种情况是旧房拆除后在原地又马上建新房,这前后 2 幢房子是不存在父子关系的独立的 2 个实体。这样就存在一个问题,即如何判断变化前后的对象是同一个对象呢?文献^[20]认为,对每一个实体必须定义一个能标识它的本质要素,这就是下文中讨论的对象的“内核”。

2.5 面向对象的优点

“面向对象”是一种认识客观世界的世界观。它将客观世界看成是由许多不同种类的对象构成的,每个对象都有自己的内部状态和运动规律,对象之间的相互联系和相互作用构成了完整的客观世界^[18]。“面向对象”还是一种新的数据组织方法。它是支撑空间复杂对象建模的最有效手段^[8],张巍等提出的复杂地物管理方法就是一个很好的例子,该方法现已成功地应用到 GeoStar 软件的开发中^[21]。在时态表达方面,面向对象数据模型有一个最基本的优点就是打破了关系模型范式的限制,直接支持对象的嵌套和变长记录^[9]。

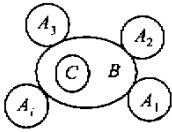
3 以事件为核心的面向对象时空数据模型

基于上述讨论,本文以房产信息系统为例,提出了一种以事件为核心的面向对象时空数据模型。其要点是:①采用面向对象的方法组织时空数据,以面向对象数据库作为功能实现的平台;②定义每一类对象的“内核”来表示该类对象的“本质”,对象类的每一个实例的对象标识(Object Identity,简称 OID)与该实例的“核”一一对应;③归纳提取驱动各类对象发生变化的原子事件的集合,定义这些事件将导致哪些对象的哪些属性同时发生变化,如何变化;④将对象的属性(含空间图形单元)按变化的同时性(即“时态内聚性”)分组存储,每组为一个数据管理基本单元;⑤将事件归类编码,时态信息用“时间+事件编码”字符串标记在各基本单元上;⑥定义事件之间的连锁关系,以事件为核心建立数据完整性控制表,事件驱动的数据更新过程以此为据。

3.1 对象的结构定义

根据第 2 部分的讨论,本文设计了如图 3 所

示的类似于原子结构的对象概念结构。



C ——对象的内核, B ——内核的紧密层
 $A_1 \sim A_i$: 对象的 i 个松散层颗粒

图 3 时态对象概念结构示意图

Fig. 3 The concept structure of spatio-temporal objects

表 1 面向对象数据库中房屋对象的结构

Tab. 1 The object structure of house in object oriented database

OID	Shape	UserID	Tbegin	Tend	Structure	Areal	Adjunct	Bound
-----	-------	--------	--------	------	-----------	-------	-------	---------	-------

表 1 是房产数据库中房屋对象的结构实例。表 1 中从左到右依次为房屋的 OID、平面图对象、用户标识码、状态开始时态、状态结束时态、建筑结构、占地面积、房屋附属物对象、房屋境界线对象。表 1 中省略号之后的均为关联的松散层。从房屋的结构可见它是一个具有 NINF 特性的内嵌了对象的复杂对象,但面向对象方法封装了复杂的内部结构,提供了更灵活的操作机制。

3.2 对象之间的主要关联方法

1. OID: 在面向对象系统中,对象之间的关系是根据对这些对象的 OID 的引用来表示的。因此,OID 就是通过一个复杂对象网络的、一个方便的导航手段^[22]。

2. UserID: 如房产中按城区、丘、幢、层、户与权属的层次隶属关系逐级编码,低层标识码包含了高层标识码。权属的 UserID 为“区号+丘号+幢号+层号+户号+权属号”,当层号+户号=“00000”时,表示该房屋为单产权,否则为多产权。由截取出的子串就可以直接查找其上层对象。UserID 据具体的情况设计,与 OID 结合解决房产中常见的多对多关联问题。

3. 纵向的时态关联:OID 在系统中是惟一的,它的时态标记值是线性递增的,这相当于由 OID 与时间、事件构成了一个 3 维格网,OID 维上每条线段都与系统中某个对象的某个状态惟一对

应。内核在概念上是对象的根本性主体(如房屋的结构主体),内核紧密层的变化标志着该对象的一些质变事件。在分析阶段,内核的定义起到辅助划分对象类的作用,对这种划分,文献[6]提出了一个属性分层聚类操作方法;在实现阶段,以 OID 代表该对象封装后的内核。各层都可以是一内嵌的表或对另一个对象的引用,如房屋空间图形就是一个内嵌的多边形对象。将驱动变化的事件的代码与对应的时间合并成一个字符串类型的时态属性项,标记在相应层的状态中,如表 1 所示。

应。线段的两端点指示了该状态开始/结束的时间和原因,线段序列即状态序列。因此,这种时空数据结构为时空分析提供了坚实的基础。有关的时空查询分析,限于篇幅另文论述。

3.3 事件驱动的对象更新机制

1. 原子事件的语义抽取。在系统分析阶段归纳提取现实世界中事件的集合,据此定义在数据对象上可能执行的动作的集合,将这些动作编码。如房屋对象的事件集合为 $E_{house} = \{\text{新建, 拆除, 扩建, 部分拆除, 加层, 丘变更}\}$,对应编码为“301~306”;产权对象的事件集合为 $E_{right} = \{\text{新增, 购买, 转卖, 转让, 分割, 继承, 合并, 删除}\}$,对应编码为“401~408”。

2. 定义每一个事件所改变的数据对象、改变方式及激发的其他事件,在系统设计时予以实现。

3. 找出事件之间的连锁关系。文献[5]对于不同实体在变化过程中的相互影响有深入的论述。本文将其中的思想应用于研究事件之间的关联。如图 4 中,一条道路的新建引起 A、B 2 个丘的形状发生变化:房屋 4 从 A 转移到 B,房屋 3 被拆掉一部分,房屋 6 被完全拆除,相关的产权失效,住户可能转移到别处并购买新的产权等等,这一系列事件都是相互关联的。事件关系的定义一方面可以保证数据的完整性和数据之间的逻辑一致性,另一方面有利于抽取变化的模式。

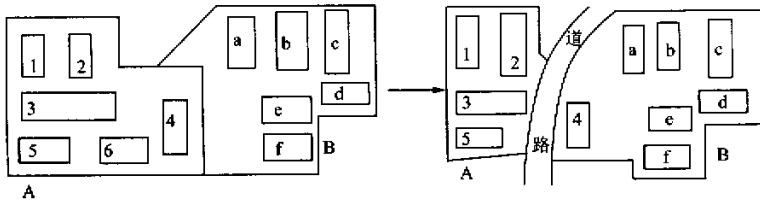


图 4 事件之间的连锁关系 道路的开拓导致房屋的变更 ,再引起产权的变化

Fig.4 The chain among events of road building , houses breaking and rights changing

4 该模型在房产信息系统中的应用

4.1 房产信息的主要特点

房产信息的主要特点为 :① 层次结构较清楚。房产信息基本上可以分为以房屋为中心的自然属性和以权属为中心的社会属性 2 大部分。② 房屋图形较简单 ,颗粒清楚。③ 时间尺度变化大 ,如古建筑的寿命可达千年 ,而商业租房的住户

变更周期则以月计 ,若数据结构设计不当极易造成巨大的数据冗余。④ 变化一般可看作是离散的事件易于定义。

4.2 对象定义

根据上述对象结构 ,本文定义了丘、房屋、层、户、产权等对象。

4.3 事件抽取、定义与编码

部分事件的定义如表 2 所示。

表 2 房产事件定义表举例

Tab.2 The definition of real estate events

事件类型	名称	代码	操作对象	操作方式	激发事件
房屋事件	新建	301	房屋	新增一个房屋对象实例	401
房屋事件	拆除	302	房屋	给房屋对象标记结束状态	408(产权删除)
产权事件	新增	401	产权	新增一个产权对象实例	50(新增户)
产权事件	分割	405	产权	原产权标记删除 ,新增 2 个产权对象实例	408 ,401 ,501
户事件	合并	504	户	原户标记删除 ,新增 2 个户对象实例	
⋮					

4.4 系统实现

系统实现的平台主要用 Arc/Info 8 的地理数据库模型 (geodatabase)^[23]。系统功能结构如图 5 所示。

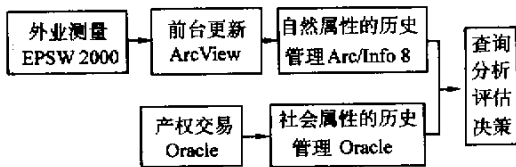


图 5 房产信息系统的功能结构

Fig.5 The structure of real estate information system

5 结束语

事件驱动下时空数据变化规律的研究为数据模型的设计提供了理论依据。本文据此设计了一

种以事件为核心的面向对象时空数据模型。该模型主要有 3 个特点 :① 确定了时态数据管理的最佳对象粒度及对象层次结构 ;② 抽取原子事件 ,将事件编码与事件发生的时间组合成一个时态信息码 ,标记在相应对象中 ,以记录该对象变化的原因 ;③ 根据事件之间的连锁关系控制数据完整性。在该模型的支持下 ,设计的房产信息系统初步实现了城建历史恢复、闲置房变化分析、拆迁评估等具有实践意义的时空分析功能。

由于面向对象方法在实体建模中的灵活性 ,这种数据模型具有较为广泛的适用性。但其前提是对应用领域中可能发生的事件的作用对象、作用方式、影响范围等问题已经研究得比较清楚 ,这样才能合理地构造对象模型、划分对象粒度 ,并对事件作语义抽取与编码。本文进一步的工作 :① 研究如何完善事件的定义及事件语义抽取、编码

方法,以应用于更复杂的时空问题;②研究基于对象、事件、时间3个基本要素的时空分析、因果推断方法。

致谢 株洲市房产局信息中心李跃纲主任在系统研究与实现过程中提出了许多宝贵的意见,谨此表示诚挚的谢意。

参考文献:

- [1] EGENHOFER M J, *et al.* Progress in Computational Methods for Representing Geographical Concepts[J]. I J GIS, 1999, 13(8): 775-796.
- [2] WANG Xiao-dong, CUI Wei-hong. Implementation of Digital Earth's Spatio-temporal Dimensions[J]. Progress in Geography, 1999, 18(2): 147-152. (in Chinese)
- [3] PEUQUET D J, DUAN N. An Event-based Spatio-temporal Data Model for Temporal Analysis of Geographical Data[J]. I J GIS, 1995, 9(1): 7-24.
- [4] PEUQUET D J. Time in GIS and Geographical Databases[A]. Geographical Information Systems[C]. New York: John Wiley & Sons, 1999. 91-103.
- [5] HORNSBY K, EGENHOFER M J. Identity-based Change: A Foundation for Spatio-temporal Knowledge Representation[J]. I J GIS, 2000, 14(3): 207-224.
- [6] CHEN Jun, CHEN Shang-chao, TANG Zhi-feng. Representing Temporal Attributes in GIS Using Non-1NF Approach[J]. Journal of Wuhan Technical University of Surveying and Mapping, 1995, 20(1): 12-17. (in Chinese)
- [7] HUANG Ming-zhi, ZHANG Zu-xun. 1NF Relational Basis of Spatio-temporal Data Model[J]. Acta Geodaetica et Cartographica Sinica, 1997, 26(1): 1-6. (in Chinese)
- [8] SHU Hong, CHEN Jun, DU Dao-sheng, *et al.* An Object-oriented Spatio-temporal Data Model[J]. Journal of Wuhan Technical University of Surveying and Mapping, 1997, 22(3): 229-233. (in Chinese)
- [9] GONG Jian-ya. An Object-oriented Spatio-temporal Data Model in GIS[J]. Acta Geodaetica et Cartographica Sinica, 1997, 26(4): 289-298. (in Chinese)
- [10] CHANG Zheng, CHEN Jun, DU Dao-sheng. Cadastral Spatio-temporal Database Organization and Query Based on Properties of Land Subdivision[J]. Journal of Wuhan Technical University of Surveying and Mapping, 1997, 22(3): 216-221. (in Chinese)
- [11] YUE Yan-fen, CHEN Jun. Organizing Spatio-temporal Data of Land Subdivision with Temporal Land Parcels[J]. Journal of Wuhan Technical University of Surveying and Mapping, 1997, 22(3): 222-228. (in Chinese)
- [12] JIANG Jie, CHEN Jun. Event-based Spatio-temporal Database Design for Land Subdivision System[J]. Acta Geodaetica et Cartographica Sinica, 2000, 29(1): 65-70. (in Chinese)
- [13] CHEN Jun, ZHAO Ren-liang. Spatial Relations in GIS: A Survey on Its Key Issues and Research Progress[J]. Acta Geodaetica et Cartographica Sinica, 1999, 28(2): 95-102. (in Chinese)
- [14] GUO Da-zhi, YANG Wei-ping, HAN Guo-jian. A Spatial and Temporal 4-D Data Model for Mine GIS[J]. Acta Geodaetica et Cartographica Sinica, 1993, 22(1): 33-40. (in Chinese)
- [15] WORBOYS M F. A Unified Model for Spatial and Temporal Information[J]. The Computer Journal, 1994, 37(1): 26-34.
- [16] ZHANG Shi-chao. Temporal Database Review[J]. Computer Science, 1992, 19(3): 37-49. (in Chinese)
- [17] TANG Chang-jie. The Achievements, Deficiency and Future Work in Temporal Databases[J]. Computer Science, 1999, 26(3): 63-65. (in Chinese)
- [18] WANG Zhi-jian, ZHAI Cheng-xiang, XU Jia-fu. Study on the Basic Concepts of Object-oriented[J]. Computer Science, 1992, 19(4): 13-17. (in Chinese)
- [19] LU Ru-qian, JIN Zhi, CHEN Gang. Ontology-oriented Requirements Analysis[J]. Journal of Software, 2000, 11(8): 1 009-1 017. (in Chinese)
- [20] LANGRAN G. A Review of Temporal Database Research and Its Use in GIS Applications[J]. I J GIS, 1989, 3(3): 215-232.
- [21] ZHANG Wei, GONG Jian-ya. Object-oriented Data Model for Complex Feature in GIS[J]. Acta Geodaetica et Cartographica Sinica, 1995, 24(4): 293-300. (in Chinese)
- [22] HE Yan-xiang, ZHENG Zhen-mei, SHI Shu-gang. Object-oriented Database[M]. Wuhan: Wuhan University Press, 1995. 13. (in Chinese)
- [23] ZEILER M. Modeling Our World: The ESRI Guide to Geodatabase Design[M]. California: ESRI Press, 1999. 1-21.