

基于历史拓扑和描述子的时空数据模型

金培权, 岳丽华, 龚育昌

(中国科学技术大学 计算机科学技术系, 安徽 合肥 230027)

A Spatiotemporal Data Model Based on History Topology and Descriptors

JIN Pei-quan, YUE Li-hua, GONG Yu-chang

(Dept of Computer Sci & Tech, University of Science and Technology of China, Hefei 230027, China)

Abstract: The problems of representing spatiotemporal information in spatiotemporal data models are analyzed, and a new spatiotemporal data model named STORM is presented. The model which is based on history topology and descriptors. Spatiotemporal information is represented using a hybrid approach, which combines explicit representation and implicit representation. History topology explicitly expresses historical relationships among spatiotemporal objects, while descriptors implicitly represent the internal changes of single spatiotemporal object and spatial topological relationships between spatiotemporal objects. The STORM model can represent both static semantics and dynamic semantics of spatiotemporal information. It can also represent and query spatiotemporal information using extended relational structures and extended relational algebra.

Key words: spatiotemporal data model; history topology; descriptor; static semantics; dynamic semantics

摘 要: 针对已有时空数据模型在表达时空信息上存在的问题, 提出基于一个新的历史拓扑和描述子的时空数据模型 STORM。STORM 采用结合显式表达和隐式表达的混合型方法来表示时空信息, 其中历史拓扑显式表达时空对象间的历史关联, 而描述子隐式表达单个时空对象的内部属性变化以及时空对象间的空间拓扑关系。该模型可以同时表达时空信息的静态语义和动态语义, 并可通过扩展的关系结构和关系代数操作进行表示和查询。

关键词: 时空数据模型; 历史拓扑; 描述子; 静态语义; 动态语义

1 引 言

时态地理信息系统(Temporal GIS)的组织核心是时空数据库。时空数据库的基础是时空数据模型。典型的时空应用如地籍管理、交通管理等都要求数据库能有效地存储和查询对象在某个时间区间里的变化情况。为了回答针对时空变化的查询, 首先必须建立一个有效的时空数据模型。

为了表达时空信息, 人们将时间戳加到空间对象的空间属性或主题属性上, 研究了一系列的时空建模方法, 如时空快照模型^[1]、基态修正模型^[1]、时空立方体模型^[1]、时空复合模型^[1]以及时空对象模型^[2]等。由于这些方法采用版本来表达时空对象的状态, 并通过版本间的比较来实现时空变化的查询, 因此对时空变化的支持相对较弱, 时空变化查询的效率低, 而且不能查询涉及

多个时空对象的时空变化(如分裂、合并等)。G. Langran^[1]、C. Claramunt^[3]、D. Peuquet^[4]、M. Egenhofer^[5,6]、A. Renolen^[7] 以及国内的陈军^[8] 等先后提出了在时空数据模型中显式存储时空变化的方法。显式存储时空变化的观点认为每个时空变化都是和某个特定的事件相关联的,因此时空变化可以通过对事件的存储和表示来实现。但将对象的所有变化都显式存储也带来了额外的开销,即不仅要显式存储对象与对象之间的变化关联,还要显式存储一个对象的各个版本之间的变化关联。在现实世界中,对象与对象之间的变化需要显式的存储,因为应用需要了解一个对象是如何从其他对象演变而来的(是分裂而来,还是合并产生?),而一个对象内部的变化并不都需要显式的存储,因为我们只关心对象的值的改变。

本文提出了一个新的基于历史拓扑和描述子的时空数据模型 STORM。该时空数据模型结合隐式方法和显式方法来表达时空信息,可以更完备地表示和查询时空变化。历史拓扑以显式的方式表达了时空对象间的历史关联,而描述子以隐式的方式表达了时空对象的空间信息和主题属性的变化以及时空对象的空间拓扑关系。

2 时空信息的表达

2.1 时空信息的静态语义和动态语义

时空信息是随时间而变化的空间信息。时空信息表达了时空对象的静态语义和动态语义。静态语义是时空对象在某个特定时刻的语义,它只表示了当前时刻或某个确定时刻时空对象的状态。动态语义是时空对象在某个时间区间里的语义,它表达了时空对象随时间而发生变化的历史信息。时空信息的表达必须要能够同时表达时空信息的静态语义和动态语义。

2.2 历史拓扑和描述子

本文采用历史拓扑和描述子来表达时空信息。其中历史拓扑表达了时空对象之间的演变过程,而描述子则表达了单个时空对象内部的变化信息。历史拓扑和描述子表达了时空信息的动态语义,也表达了时空信息的静态语义。

定义 1 历史拓扑: 一个历史拓扑是一个时空对象与其他时空对象在时间轴上的关联。

定义 2 描述子: 时空对象的描述子是定义在时间域上的一个函数 $F(\text{Time}, \text{SubObject})$, 其中 Time 是时间域, SubObject 是时空对象的一个

属性子集。如果 SubObject 是时空对象的空间属性,则称为空间描述子;如果 SubObject 是时空对象的主题属性子集,则称为属性描述子。

一个时间区间里的描述子状态可以是一个连续状态,也可以是离散状态。连续状态表示在一段时间里,时空对象的属性随时间连续变化,而离散状态表示时空对象的属性在给定时间区间里保持不变。

本文通过相应的图形来表示历史拓扑和描述子。图 1 给出了时空信息表达中使用的符号。

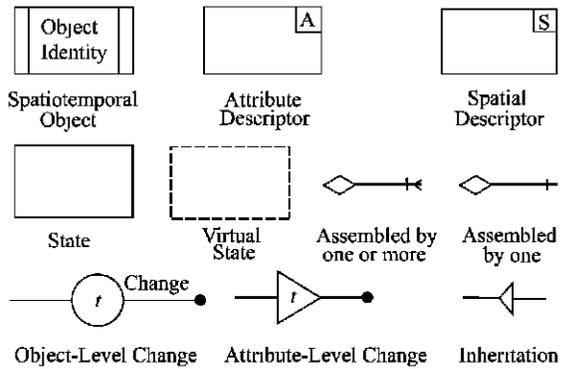


图 1 时空信息表达所使用的符号

Fig. 1 Notations to represent spatiotemporal information

2.3 时空信息表达实例

下面以地块变化的例子来说明通过历史拓扑和描述子表达时空信息的思想。图 2 显示了地块的时空变化。

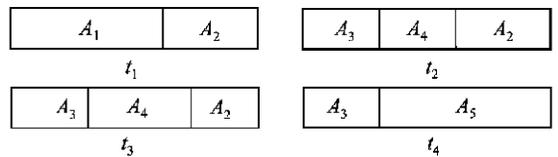


图 2 地块的时空变化信息

Fig. 2 Spatiotemporal changes of fields

我们假设每个地块有一个空间位置和一个所有者,并且假设有下面的时空变化:

1. 初始时只有 A_1 和 A_2 两个地块,所有者分别为“Amey”和“Batt”;
2. t_2 时刻地块 A_1 分裂为两个地块 A_3 和 A_4 ,“Amey”成为 A_3 和 A_4 的所有者;
3. t_3 时 A_4 的空间区域增加同时 A_2 的空间区域减小,同时 A_2 的所有者由“Batt”变为“Catty”;
4. t_4 时刻 A_4 和 A_2 合并为地块 A_5 , A_5 的所有者为“Doggy”。

地块的信息通过一个空间描述子和一个属性

描述子来描述,其中属性描述子AD描述地块的所有者变化历史,SD表示地块的空间信息变化历史。

一个地块与其他地块之间的历史关联通过历史拓扑来表示。图3显示了图2例子对应的时空变化。

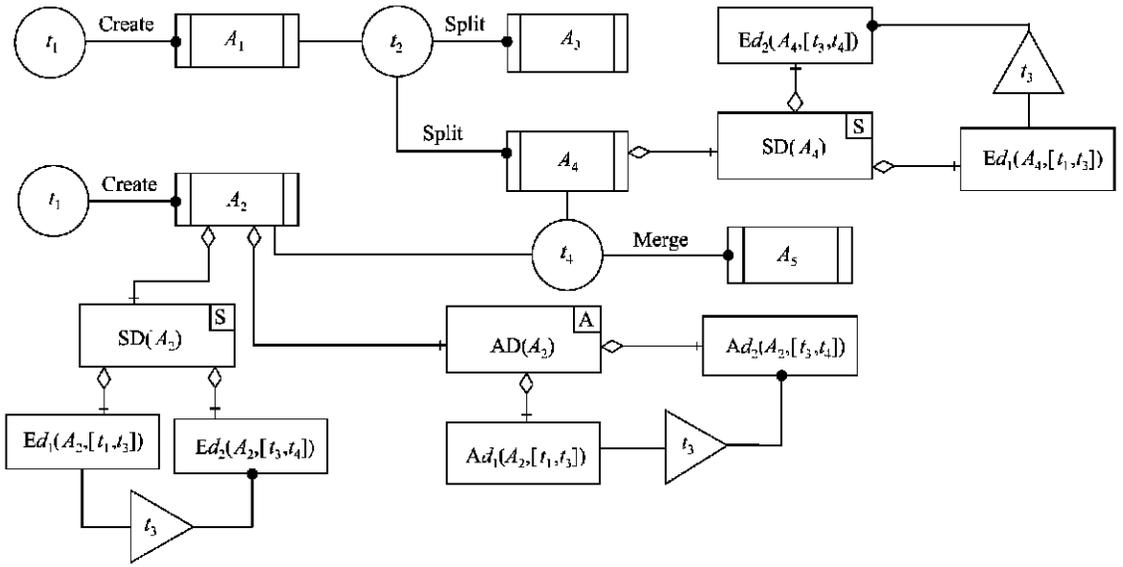


图3 地块的时空变化表示

Fig. 3 Representing spatiotemporal changes of field

图3中,每个圆表示一次历史拓扑(即对象级变化,object-level change)。t₂时的历史拓扑表示了A₁在t₂时刻分裂为A₃和A₄这一时空变化,而t₄时的历史拓扑表示了A₂和A₄在t₄时刻合并为A₅这一时空变化。例子中共发生了两次空间变化,分别是A₂在t₃时刻的空间变化和A₄在t₃时刻的空间变化。这两个空间变化通过A₂的空间描述子SD(A₂)和A₄的空间描述子SD(A₄)表示,在图3中以右上角带“S”标记的矩形框表示。其中A₂的空间描述子由两个空间状态Ed₁(A₂, [t₁, t₃])和Ed₂(A₂, [t₃, t₄])构成,这两个状态之间的变化是一次属性级变化(attribute level change)(空间属性发生变化),在图3中以三角形表示。A₄的空间变化表示与A₂类似。例子中共发生了一次主题属性变化,即t₃时刻A₂的所有者发生了改变。这一变化以A₂的属性描述子AD(A₂)表示,在图3中以右上角带“A”标记的矩形框表示。AD(A₂)包含了两个主题属性状态Ad₁(A₂, [t₁, t₃])和Ad₂(A₂, [t₃, t₄]),两个状态之间的变化也是一次属性级变化(主题属性发生变化)。

图3中限于篇幅省略了A₁、A₃和A₅的空间描述子和属性描述子,以及A₄的属性描述子,因为A₁、A₃、A₅的属性(包括空间属性和主题属性)以及

A₄的主题属性在其存在时间里都没有发生变化。

3 时空数据模型 STORM 的数据结构

STORM 中一个时空对象表示为一个元组,我们将其定义为时空元组。时空元组集合构成了时空关系。

定义3 给定一组域 D₁, D₂, ..., D_n, 其中 D_n 是历史拓扑类型, 并且至少存在一个域 D_i (1 ≤ i ≤ n) 是时空数据类型, 笛卡儿积 D₁ × D₂ × ... × D_n 的每一个元素称为一个时空元组(Spatiotemporal Tuple)。

定义4 一个时空关系(Spatiotemporal Relation)是一个时空元组的集合。

图4显示了STORM中的基本数据结构。其中Thematic Attributes是时空对象的主题属性,Spatiotemporal Attributes是扩展的时空数据类型,表示时空对象的时空变化。History Topology Attribute是扩展的历史拓扑类型,表示了时空对象在不同时间里与其他时空对象间的历史关联。

Primary Key	Thematic Attributes	Spatiotemporal Attributes	History Topology Attributes
-------------	---------------------	---------------------------	-----------------------------

图4 时空对象的数据结构

Fig. 4 Data structure of a spatiotemporal object

4 时空关系的表达

现实世界中时空实体之间的关系主要有两种:¹ 静态时空关系, 即某个特定时刻时空对象之间的空间拓扑关系; ④ 动态时空关系, 即时空对象与时空对象沿时间轴方向发生的关联, 如一个地理对象 A 与另一个地理对象 B 在 t 时刻合并为对象 C , 然后对象 C 又在时刻 t' ($t' > t$) 分裂为对象 D 和对象 E 。这两类时空关系在 STORM 中通过不同的方法来表示。

4.1 静态时空关系的隐式表示

静态时空关系表达了时空信息的静态语义, 它通过将时空对象投影到时间维上得到时空对象在某个时刻的状态以及与其他时空对象的空间拓扑关系。时空对象间的静态时空关系通过时空数据类型在特定时刻的空间拓扑关系操作来表达。空间拓扑关系操作包括以下 6 种静态时空关系: disjoint, meet, contain, intersects, overlap 和 equal。

设 μ_1 和 μ_2 是时空数据类型, 则 μ_1 和 μ_2 的任意实例在特定时刻 (instant 类型) 的静态时空关系可以通过以下的空间拓扑操作表示 (其中 when 操作返回给定时刻的时空对象状态):

1. 相离 disjoint: $\text{when}(\mu_1, \text{instant}) \times \text{when}(\mu_2, \text{instant}) \rightarrow \text{bool};$
2. 相邻 meet: $\text{when}(\mu_1, \text{instant}) \times \text{when}(\mu_2, \text{instant}) \rightarrow \text{bool};$
3. 部分覆盖 overlap: $\text{when}(\mu_1, \text{instant}) \times \text{when}(\mu_2, \text{instant}) \rightarrow \text{bool};$
4. 包含 contain: $\text{when}(\mu_1, \text{instant}) \times \text{when}(\mu_2, \text{instant}) \rightarrow \text{bool};$
5. 相交 intersects: $\text{when}(\mu_1, \text{instant}) \times \text{when}(\mu_2, \text{instant}) \rightarrow \text{bool};$
6. 完全覆盖 equal: $\text{when}(\mu_1, \text{instant}) \times \text{when}(\mu_2, \text{instant}) \rightarrow \text{bool};$

静态时空关系在 STORM 中采用了隐式表示的方法。在时空数据库应用中, 任何两个空间对象间都存在着联系, 并且这种联系还会随时间而变化。以汽车与桥梁为例, 无论什么时刻, 一个汽车与一个桥梁之间总是存在一种空间拓扑关系。因此, 任何两个时空对象间在不同时刻存在着不同的静态时空关系, 不能采用传统的 ER 方法来预先指定时空对象间的静态时空关系, 而应当通过每个时刻时空对象间的空间拓扑关系来确定。

4.2 动态时空关系的显式表示

动态时空关系是一个时空对象沿时间轴与其他时空对象间发生的联系。动态时空关系表达了时空信息的动态语义, 它表达了时空对象之间的演变过程。描述子也表达了动态语义, 但它表达的是单个时空对象的演变过程。

动态时空关系的表示不能采取静态时空关系的隐式表示方法。这是因为动态时空关系不仅要指明一个时空对象从哪些时空对象演变而来, 最终演变成哪些时空对象, 而且还要指出这种演变是以什么方式进行的。例如在城市规划中, 行政区 A 是从以前的行政区 B 中演变而来的, 对于应用而言, 不仅要知道 A 是从 B 演变而来的, 而且还要知道演变的方式 (例如是从 B 中划分出来的) 以及演变的时间。这些信息无法直接从时空对象的时空数据中计算得到。

在 STORM 中, 动态时空关系采用了显式的表示方法, 即通过历史拓扑类型存储的时空对象的历史拓扑值来表示动态时空关系。每一个时空对象 (时空元组) 都有一个历史拓扑类型的属性。每当时空对象发生标识变化, 就更新相关时空对象的历史拓扑值, 即在当前的历史拓扑值中附加一个历史拓扑状态。例如, 要记录 t 时刻 A 从 B 中分裂出来的历史拓扑信息, 在数据库中插入 A 时, A 的历史拓扑值为“ $\langle \{B\}, \text{NULL}, 2, t \rangle$ ”; 同时, 由于 B 也发生了标识变化 (假设分裂为 A 和 C), 则在 B 的历史拓扑值后附加一个新的历史拓扑状态值“ $\langle \{\text{NULL}, \{A, C\}\}, 2, t \rangle$ ”。

5 时空查询操作

基于 STORM 的时空数据库中的数据被组织成时空关系。时空关系上的查询操作与关系代数操作类似。下面给出 STORM 上的 5 个基本关系代数操作的定义。由于 5 个基本关系代数操作构成了关系代数操作的完备集合, 因此 STORM 上的其他查询操作都可以通过这 5 个基本操作实现。

1. 时空并: 设 R 和 S 是两个时空关系, R 和 S 的并是由属于 R 或属于 S 的时空元组构成的集合。时空关系的并用 $R \cup^s S$ 形式定义如下:

$$S \cup^s R = \{t \mid t \in R \vee t \in S, t \text{ 是时空元组变量}\}$$

2. 时空差: R 和 S 的差是由属于 R 但不属于 S 的时空元组构成的集合。时空关系的差用 $R -^s S$ 表示, 形式定义如下:

$R - {}^s S \underline{\Delta} \{t \mid t \in R \wedge t \in S, t \text{ 是时空元组变量}\}$

3. 时空笛卡儿积: 设 R 和 S 的属性个数分别为 r 和 s 。定义 R 和 S 的笛卡儿积是一个具有 $(r + s)$ 个属性的时空元组集合, 其中每个时空元组的前 r 个属性值来自 R 的一个时空元组, 后 s 个属性值来自 S 的一个时空元组。时空关系的笛卡儿积用 $R \times {}^s S$ 表示, 形式定义如下:

$R \times {}^s S \underline{\Delta} \{t \mid t = \langle t', t'' \rangle \wedge t' \in R \wedge t'' \in S, t \text{ 是时空元组变量}\}$

t 是时空元组变量

4. 时空选择: 选择操作根据某些条件对时空关系作水平分割, 即选取满足条件的时空元组。条件用命题公式 F 表示。 F 有两种成分。一种是“运算对象”, 另一种是“运算符”。“运算对象”可以是常数、属性名或函数, 其中函数可以是空间/时态几何操作、空间/时态属性操作或空间/时态拓扑操作。“运算符”可以是算术比较运算符 ($>, \geq, <, \leq, \neq, =$) 或逻辑运算符 (\neg, \wedge, \vee)。时空关系 R 关于公式 F 的选择操作作用 $\sigma_F^s(R)$ 表示, 形式定义如下:

$\sigma_F^s(R) \underline{\Delta} \{t \mid t \in R \wedge F(t) = \text{true}, t \text{ 是时空元组变量}\}$

t 是时空元组变量

5. 时空投影: 投影操作对一个时空关系进行垂直分割。设 R 是一时空关系, R 的属性记为 (A_1, A_2, \dots, A_r) , 一个时空元组 t 在其属性 A_i 上的值记为 t_{A_i} , R 在属性 $A_{i_1}, A_{i_2}, \dots, A_{i_m}$ ($m \leq r, i_1, \dots, i_m$ 为 1 到 r 的整数) 的投影用 $\pi_{A_{i_1}, A_{i_2}, \dots, A_{i_m}}^s(R)$ 表示, 它是一个具有 m 个属性的时空元组集合, 形式定义如下:

$\pi_{A_{i_1}, A_{i_2}, \dots, A_{i_m}}^s(R) \underline{\Delta} \{t \mid t = \langle t_{A_{i_1}}, t_{A_{i_2}}, \dots, t_{A_{i_m}} \rangle \wedge \langle t_{A_1}, t_{A_2}, \dots, t_{A_r} \rangle \in R\}$

6 实验分析

为了检验 STORM 模型的实用性, 我们对 STORM 进行了实现并通过一个实验例子来说明 STORM 对时空应用的支持。整个实现结构如图 5 所示。

STORM 的实现以 Informix Dynamic Server 为底层的 DBMS, 采用 Informix DBDK 工具集进行开发。其中 C 函数库是实现时空数据类型和时空操作的代码库, 可以在运行时被链接到 Informix DBMS 中。时空 DataBlade 提供了对对象

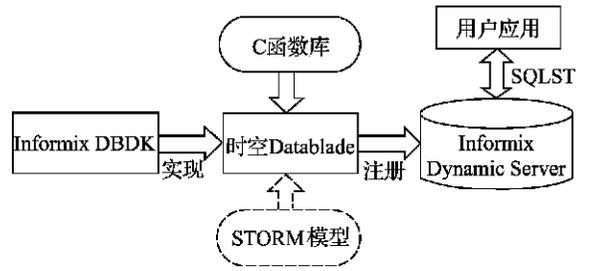


图 5 STORM 的实现框架

Fig. 5 Framework to implementing STORM

关系数据库管理系统的时空支持, 用 C 语言编写。时空 DataBlade 被注册后, 用户应用就可以通过 SQL 方便地获得时空数据管理的支持, 不需要任何额外的工作。

由于 STORM 的实现相当于数据库的一种插件, 其本身并没有显式的界面, 因此我们通过一个例子数据库 storm-demo 来说明基于 STORM 模型的实现过程以及对时空应用的支持。在下面的文字中, 倾斜体文字都是时空 DataBlade 中的时空数据类型或时空操作。

首先创建地块表 land (land_id int, boundary stdregion, history mHT), 并插入若干条记录。其中类型 stdregion 对应地块的空间描述子, mHT 对应地块的历史拓扑。插入的数据对应图 6 中的例子。

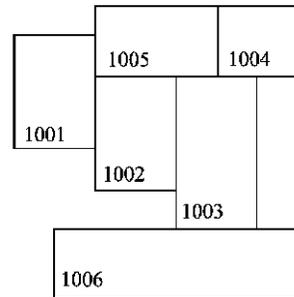


图 6 插入的地块数据

Fig. 6 Fileds data inserted

例 1 查询例子(空间拓扑——静态时空关系查询): 查询“1999-12-31”时与地块 1002 相邻的所有地块的地块号和边界。

```
Select b.land_id, b.boundary
From land a, land b Where
meet(when( a.boundary, '1999-12-31
00:00:00'), when( b.boundary, '1999-12-31
00:00:00')) and a.land_id= 1002;
```

该查询的结果见图 7。



图 7 例 1 的查询结果

Fig. 7 The querying results of example 1

例 2 时空变化表示(空间变化——单个地块的时空变化): 在“2000-12-31”, 地块 1006 的边界扩大为“(2, 1), (9, 1), (9, 3), (2, 3)”

Update Land Set

```
boundary = Replace ( stdregionOut ( boundary ), '
NOW', "'2000-12-31 00:00:00" )
|| ', "'2000-12-31 00:00:00" NOW 2 1 9 1 9 3 2 3'
Where land- id= 1006;
```

例 3 时空变化表示(历史拓扑——地块间的时空变化): 在“2001-12-31”, 地块 1006 分裂为两个地块

```
Insert Into Land Values( 1007, "'2001-12-31
00:00:00" NOW 2 1 5 1 5 3 3 1 2 1', '3" 2001-12-31
00:00:00" 1006, NULL' );
Insert Into Land Values( 1008, "'2001-12-31
00:00:00" NOW 5 1 9 1 9 3 5 3 5 1', '3" 2001-12-31
00:00:00" 1006, NULL' );
Update Land Set
boundary = Replace ( stdregionOut ( boundary ), '
NOW', "'2001-12-31 00:00:00" ), history = history || ' | 3" 2001-12-31 00:00:00" NULL, 1007
1008' Where land- id= 1006;
```

例 4 时空查询: 查询地块 1006 在 1999-01-01 到 2003-01-01 之间的空间变化

```
Select history( boundary, "'1999-01-01
00:00:00, 2003-01-01 00:00:00" :: period ) as
st- change From land Where id= 1006
```

例 5 时空查询: 查询地块 1006 在 1999-01-01 到 2003-01-01 间的标识变化信息(历史拓扑)

```
Select history( history, "'1999-01-01
00:00:00, 2003-01-01 00:00:00" :: period ) as
st- change From land Where id= 1006
```

通过实验可以看到, STORM 模型可以有效地存储时空数据和时空变化, 并能够有效地表示

时空查询。同时, 由于 STORM 模型可以在对象关系数据库之上实现, 时空应用通过 SQL 语言就可以存取时空数据库, 因此 STORM 模型可以有效地支持时空应用。

7 结束语

本文针对已有时空数据模型存在的问题, 提出了一个基于历史拓扑和描述子的时空数据模型 STORM。STORM 通过历史拓扑显式地表达时空对象间的历史演变过程, 同时通过描述子隐式地表达单个时空对象内部的属性变化和时空对象间的空间拓扑关系, 从而可以同时表达时空信息的静态语义和动态语义。STORM 的数据结构采用简单的时空关系, 并通过扩展的关系代数操作进行查询, 因此数据表示和操作非常简单, 也易于实现。

参考文献:

- [1] LANGRAN G. Time in Geographic Information Systems[M]. London: Taylor & Francis. 1992.
- [2] WORBOYS M F. A Unified Model for Spatial and Temporal Information [J]. The Computer Journal, 1994, 37(1): 26-34.
- [3] CLARAMUNT C, et al. Toward Semantics for Modeling Spatio-temporal Processes within GIS[A]. Proc of 7th International Symposium on Spatial Data Handling[C]. Delft: Taylor & Francis, 1996. 47-63.
- [4] PEUQUET D J, et al. An Event-Based Spatiotemporal Data Model (ESTDM) for Temporal Analysis of Geographical Data[J]. International Journal of Geographical Information Systems, 1995, (1): 7-24.
- [5] HORNSBY K, et al. Identity-Based Change Operations for Composite Objects[A]. Proc of 8th International Symposium on Spatial Data Handling[C]. Vancouver: International Geographical Union, 1998. 202-213.
- [6] HORNSBY K, et al. Identity-Based Change: A Foundation for Spatio-temporal Knowledge Representation [J]. International Journal of Geographical Information Science, 2000, (3): 207-224.
- [7] RENOLEN A. History Graphs: Conceptual Modeling of Spatiotemporal Data[A]. Proc of GIS Frontiers in Business and Science[C]. Brno, Czech Republic: International Cartographic Association, 1997. 31-43.
- [8] CHEN Jun, et al. An Event-Based Approach to Spatio-temporal Data Modeling in Land Subdivision Systems[J]. GeoInformatica, 2000, (4): 387-42.