

模糊时空 XML 数据 Schema 的研究

柏禄一¹, 严丽², 马宗民¹, 徐长明¹

(1. 东北大学 信息科学与工程学院, 辽宁 沈阳 110819; 2. 东北大学 软件学院, 辽宁 沈阳 110819)

摘 要: 为了管理模糊时空 XML 数据, 将若干个模糊构造子引入到 XML 文档中, 并对 XML 文档的 Schema 进行模糊时空扩展, 同时, 通过修改 XML 文档的 Schema 以容纳这些新引入的模糊构造子. 由此形式化提出模糊时空 XML 数据模型, 之后分别对模糊时空 XML 数据模型中 OID 部分、ATTR 部分、FP 部分、FM 部分以及 FT 部分的 Schema 进行扩展, 并举出具体实例加以说明. 通过对 Schema 的扩展, 使其兼容模糊时空信息, 方便管理模糊时空 XML 数据, 从而可以进一步表示、推理和查询模糊时空 XML 数据.

关 键 词: 数据库; 模糊时空数据; XML 模型; Schema; 模糊构造子

中图分类号: TP 311 **文献标志码:** A **文章编号:** 1005-3026(2015)05-0622-04

Research on Schema of Fuzzy Spatiotemporal XML Data

BAI Lu-yi¹, YAN Li², MA Zong-min¹, XU Chang-ming¹

(1. School of Information Science & Engineering, Northeastern University, Shenyang 110819, China; 2. School of Software, Northeastern University, Shenyang 110819, China. Corresponding author: BAI Lu-yi, E-mail: baily@neuq.edu.cn)

Abstract: In order to manage fuzzy spatiotemporal XML data, several fuzzy constructs are integrated in XML documents. Consequently, Schema of XML documents need to be fuzzy spatiotemporal extended. In the meantime, the integrated constructs should be compatible by modifying Schema of XML documents. Fuzzy spatiotemporal XML data model was formally proposed firstly. Subsequently, Schema of the OID part, the ATTR part, the FP part, the FM part, and the FT part in fuzzy spatiotemporal XML data model were extended, respectively. In succession, an example was introduced to give illustrations. After extending Schema, the XML documents were compatible with fuzzy spatiotemporal information, which was easy to manage fuzzy spatiotemporal XML data. What's more, it could be further used for representing, reasoning, and querying fuzzy spatiotemporal XML data.

Key words: database; fuzzy spatiotemporal data; XML model; Schema; fuzzy construct

数据库技术是当今世界计算机应用中最重要
的基础领域之一. 作为现代数据库的重要分支, 时
空数据库的管理问题是时空数据应用领域中的热
点及难点研究问题^[1-2]. 在时空数据应用领域, 由
于时空数据的爆炸性增长以及时空数据在现实应
用中多以模糊的形式存在, 再加上非模糊时空数
据相关研究的快速发展, 因此对模糊时空数据的
管理受到了越来越高的重视^[3-5]. 模糊时空数据
管理能力在很大程度上依赖于数据模型的设

计^[6-7], 模糊时空数据模型是实现模糊时空数据
管理的基础. 为了实现模糊时空 XML 数据的建
模, 若干个模糊构造子被引入到 XML 文档中^[8].
因此, 需要对 XML 文档的 Schema 进行模糊时空
扩展, 同时, 需要通过修改 XML 文档的 Schema
以容纳这些新引入的模糊构造子^[9-10].

本文首先提出模糊时空 XML 数据模型, 之
后分别对模糊时空 XML 数据模型中 5 个部分的
Schema 进行扩展, 并举出具体实例加以说明.

收稿日期: 2014-04-12

基金项目: 国家自然科学基金资助项目(61402087, 61370075, 61073139); 河北省自然科学基金资助项目(F2015501049); 中央高
校基本科研业务费专项资金资助项目(N130323006); 河北省教育厅资助科研项目(QN2014339); 东北大学秦皇岛分校
博士基金资助项目(XNB201428).

作者简介: 柏禄一(1984-), 男, 辽宁丹东人, 东北大学秦皇岛分校讲师, 博士; 马宗民(1965-), 男, 黑龙江齐齐哈尔人, 东北大学
教授, 博士生导师.

1 模糊时空 XML 数据模型

XML 数据模型最基本的结构是数据树. 因此, 本文将每一个模糊时空数据看成一棵模糊时空 XML 文档树. 根据模糊时空数据的特点, 首先给出其基于 XML 的数据模型.

定义 1 模糊时空 XML 数据可以表示为一个 5 元组: $FSP = (OID, ATTR, FP, FM, FT)$. 其中: OID 表示模糊时空数据的变化历史; $ATTR$ 表示模糊属性; FP 表示模糊位置; FM 表示模糊运动; FT 表示模糊时间.

模糊时空数据的变化历史 OID 的变化不仅可以表示模糊时空数据的变化类型, 还可以表示该模糊时空数据由什么数据变化而来, 变化成什么数据. 本文用 $ATTR$ 表示模糊时空数据的属性, 属性个数可以是一个或多个, 可以是模糊或非模糊; FP 表示模糊时空数据的位置信息, 可以是模糊的; FM 包括模糊时空数据的运动方向和运动值; FT 包括模糊时间点和模糊时间区间.

2 模糊时空 XML 数据的 Schema 扩展

由模糊时空 XML 数据模型可知, 模糊时空数据分为 OID , $ATTR$, FP , FM 以及 FT 5 部分. 因此, 本节从以上 5 部分分别对模糊时空 XML 数据的 Schema 进行扩展.

2.1 OID 部分

对 OID 部分时空 XML 数据的 Schema 扩展定义了存在的元素节点和属性节点, 以及这些节点间的结构关系. 本文将变化类型、前趋以及后继看作非模糊的, 因此这 3 个元素节点没有可能性分布. 值得一提的是, “original - definition” 是通用的节点名, 可以根据实际需要由模糊时空 XML 数据文档中具体的节点名所替代. 另外, “fttype” 表示模糊时间类型, 具体将在 FT 部分时空 XML 数据的 Schema 扩展中讨论. 关于 OID 部分时空 XML 数据的 Schema 扩展如图 1 所示.

2.2 $ATTR$ 部分

对 $ATTR$ 部分时空 XML 数据的 Schema 扩展, 属性节点存在析取和合取两种类型可能性分布, 因此, 需要定义这两种类型的可能性分布, 并定义 $Dist$ 区分这两种类型可能性分布. 此外, 取值为 $[0, 1]$ 的可能性属性 “Poss” 被定义, 用来表

```
<xs:element name = "OID" type = "xs:string" >
  <xs:complexType >
    <xs:sequence >
      <xs:element name = "type" type = "xs:string" />
      <xs:attribute name = "original - definition" type =
"fttype" />
      <xs:element name = "predecessor" type = "xs:string" />
      <xs:attribute name = "original - definition" type = "fttype" />
      <xs:element name = "successor" type = "xs:string" />
      <xs:attribute name = "original - definition" type =
"fttype" />
    </xs:sequence >
    <xs:attribute name = "original - definition" type = "fttype" />
  </xs:complexType >
</xs:element >
```

图 1 OID 部分的 Schema 扩展

Fig. 1 Schema extensions of OID part

示节点的隶属度或可能性分布值. 关于 $ATTR$ 部分时空 XML 数据的 Schema 扩展如图 2 所示.

```
<xs:element name = "ATTR" type = "xs:string" >
  <xs:element name = "original - definition" type = "xs:string"
    minOccurs = "0" maxOccurs = "unbounded" >
    <xs:attribute name = "original - definition" type = "fttype" />
    <xs:element name = "Val" type = "valtype" minOccurs = "1"
      maxOccurs = "unbounded" >
      <xs:complexType name = "valtype" >
        <xs:sequence >
          <xs:element name = "Dist" type = "distype" />
          <xs:complexType name = "distype" >
            <xs:element name = "original - definition" minOccurs =
              "0" maxOccurs = "unbounded" >
              <xs:attribute name = "Poss" type = "xs:fuzzyval"
                minOccurs = "0" maxOccurs = "1" >
              </xs:attribute >
            </xs:element >
            <xs:attribute value = "disjunctive | conjunctive" >
            </xs:complexType >
          </xs:sequence >
        </xs:complexType >
        <xs:attribute name = "Ts" type = "xs:fuzzytime" />
        <xs:attribute name = "Te" type = "xs:fuzzytime" />
      </xs:element >
    </xs:element >
  </xs:element >
```

图 2 $ATTR$ 部分的 Schema 扩展

Fig. 2 Schema extensions of $ATTR$ part

2.3 FP 部分

由于本文主要研究二维模糊时空数据, 对 FP 部分时空 XML 数据的 Schema 扩展仅定义了模糊时空数据的 MBR 部分. 对于一维模糊时空数据, 可将 $sequence$ 部分替换为 “ x_l ” 和 “ x_r ”. 关于 FP 部分时空 XML 数据的 Schema 扩展如图 3 所示.

```

<xs:element name="position" type="xs:string">
  <xs:attribute name="original-definition" type="fttype"/>
  <xs:element name="Val" type="valtype" minOccurs="1"
    maxOccurs="unbounded">
    <xs:complexType>
      <xs:sequence>
        <xs:element name="xmin" type="xs:float"/>
        <xs:element name="xmax" type="xs:float"/>
        <xs:element name="ymin" type="xs:float"/>
        <xs:element name="ymax" type="xs:float"/>
      </xs:sequence>
    </xs:complexType>
  </xs:element>
</xs:attribute>
</xs:element>

```

图 3 FP 部分的 Schema 扩展

Fig. 3 Schema extensions of FP part

2.4 FM 部分

对 FM 部分时空 XML 数据的 Schema 扩展分别定义了运动变化方向和运动变化值的 Schema. 由于存在 3 种运动变化方向, 因此在时空 XML 数据的 Schema 扩展中将这 3 种运动变化方向进行枚举. 关于 FM 部分时空 XML 数据的 Schema 扩展如图 4 所示.

```

<xs:element name="motion" type="xs:string">
  <xs:attribute name="original-definition" type="fttype"/>
  <xs:element name="Val" type="avaltype" minOccurs="1"
    maxOccurs="unbounded">
    <xs:element name="xaxis" type="axistype"/>
    <xs:restriction base="xs:string">
      <xs:enumeration value=" ">
      <xs:enumeration value=" ">
      <xs:enumeration value=" ">
    </xs:restriction>
    <xs:element name="yaxis" type="axistype"/>
    <xs:restriction base="xs:string">
      <xs:enumeration value=" ">
      <xs:enumeration value=" ">
      <xs:enumeration value="↑">
    </xs:restriction>
    <xs:element name="xval" type="float"/>
    <xs:element name="yval" type="float"/>
  </xs:element>
</xs:attribute>
</xs:element>

```

图 4 FM 部分的 Schema 扩展

Fig. 4 Schema extensions of FM part

2.5 FT 部分

对 FT 部分时空 XML 数据的 Schema 扩展, 实际上是对模糊时空数据的其他 4 部分元素节点的时间属性进行模式扩展. 本文用“ T_s ”和“ T_e ”分别表示时间属性的起始时间点和结束时间点, 其类型为“xs:fuzzyinteger”, 表示起始时间点和结束时间点是模糊的. 关于 FT 部分时空 XML 数据

的 Schema 扩展如图 5 所示.

```

<xs:complexType name="fttype">
  <xs:sequence>
    <xs:element name="Ts" type="xs:fuzzyinteger">
    <xs:element name="Te" type="xs:fuzzyinteger">
  </xs:sequence>
</xs:complexType>

```

图 5 FT 部分的 Schema 扩展

Fig. 5 Schema extensions of FT part

3 实例说明

在对模糊时空 XML 数据进行 Schema 扩展之后, 下面举出一个具体实例对其加以说明. 代码 1~36 行是带有模糊时空信息的 XML 文档片断, 下面分别从模糊时空 XML 数据的 OID, ATTR, FP, FM 以及 FT 部分进行说明.

代码片断:

```

1. <clouds>
2.   <Val Poss = 0.9>
3.     <cloud OID = "Wiz Khalifa">
4.       <type Ts = "t0" type="create" />
5.       <covered cities Ts = "t1" Te = "t2">
6.         <Dist type = "conjunctive">
7.           <Val Poss = 0.8>Shenyang</Val>
8.           <Val Poss = 0.7>Dalian</Val>
9.         </Dist>
10.      </covered cities>
11.      <cloud density Ts = "t3" Te = "t4">
12.        <Dist type = "disjunctive">
13.          <Val Poss = 0.85>thick</Val>
14.          <Val Poss = 0.75>thin</Val>
15.        </Dist>
16.      </cloud density>
17.      <position Ts = "t5" Te = "t6">
18.        <Val Poss = 0.8>
19.          <xmin>3</xmin>
20.          <ymin>5</ymin>
21.          <xmax>6</xmax>
22.          <ymax>8</ymax>
23.        </Val>
24.      </position>
25.      <motion Ts = "t7" Te = "t8">
26.        <Val Poss = 0.7>
27.          <xaxis>"("</xaxis>
28.          <yaxis>"("</yaxis>
29.        </Val>
30.        <Val Poss = 0.8>
31.          <xval>4</xval>

```

```

32.      <yval>6</yval>
33.      </Val>
34.    </motion>
35.  </Val>
36. <clouds>

```

第3行,表示 cloud 名为“Wiz Khalifa”的可能性分布为 0.9. 当元素节点的可能性分布等于 1 时,标签 <Val Poss = 1.0> 和标签 </Val> 通常被省略,例如第4行的变化类型是非模糊的,因此标签 <Val Poss = 1.0> 和标签 </Val> 被省略.

基于标签 <Val Poss> 和标签 </Val>, 可以表示元素节点的隶属度和属性节点的可能性分布. 对 ATTR 部分 XML 模式的模糊时空扩展, 本文引入 Dist 区分属性节点存在唯一值或多个值. 第6行到第9行是合取可能性分布, 这是因为名为“Wiz Khalifa”的云层可能在同一时间覆盖多个城市. 第12行到第15行是析取可能性分布, 这是因为云层的密度在同一时间是唯一值.

第17行到第24行表示云层的位置信息, 其中该云层模糊位置的 MBR 表示点为 (3, 5) 和 (6, 8). 第18行和第23行表示该云层处于此位置的隶属度等于 0.8.

第25行到第34行表示云层的运动信息. 其中, 第27行和第28行表示云层运动方向是右上方向(东北方向). 第26行和第29行表示云层的运动方向是右上方向(东北方向)的隶属度等于 0.7. 第31行和第32行表示云层的运动变化值在水平方向扩大 4 倍、在垂直方向扩大 6 倍. 第30行和第33行表示云层的运动变化值为第31行和第32行所示的隶属度等于 0.8.

第4行到第5行、第11行、第17行、第25行都包括时间信息, 并且标识为“ T_s ”和“ T_e ”的时间信息可能是模糊的. 第4行的时间信息表示名为“Wiz Khalifa”的云层在时间 t_0 时被创建. 由于模糊时空数据的变化类型是非模糊的, 因此 t_0 是非模糊的; 第5行的时间信息表示云层在 t_1 和 t_2 的时间区间上所覆盖的城市是第7行的 Shenyang 和第8行的 Dalian, 其中 t_1 和 t_2 可以是模糊时间; 第11行的时间信息表示云层的密度在 t_3 和 t_4 的时间区间上是第13行可能性分布为 0.85 的浓和第14行可能性分布为 0.75 的稀, 其中 t_3 和 t_4 可以是模糊时间; 第17行的时间信息表示云层在 t_5 和 t_6 的时间区间上位置在第19行到第22行上,

其中 t_5 和 t_6 可以是模糊时间; 第25行的时间信息表示云层在 t_7 和 t_8 的时间区间上运动变化方向是第27行和第28行的右上方向, 运动变化值是第31行的水平方向扩大 4 倍和第28行的垂直方向扩大 6 倍, 其中 t_7 和 t_8 可以是模糊时间.

4 结 语

本文形式化提出了模糊时空 XML 数据模型, 并分别对模糊时空 XML 数据模型中 5 个部分的 Schema 进行扩展, 通过具体实例说明其兼容模糊时空信息, 以达到管理模糊时空 XML 数据的目的.

参考文献:

- [1] Wang X Y, Zhou X F. Spatiotemporal data modeling and management: a survey [C]//Proceedings of the 36th International Conference on Technology of Object-Oriented Languages and Systems. Xi'an, 2000: 202-211.
- [2] Abraham T, Roddick J F. Survey of spatiotemporal databases [J]. *GeoInformatica*, 1999, 3(1): 61-99.
- [3] Stefanakis E. A unified framework for fuzzy spatiotemporal representation and reasoning [C]//Proceedings of the 20th International Cartographic Conference. Beijing, 2001: 2678-2687.
- [4] Sözer A, Yazici A, Oğuztüzün H, et al. Modeling and querying fuzzy spatiotemporal databases [J]. *Information Sciences*, 2008, 178(19): 3665-3682.
- [5] Ribaric S, Hrkac T. A model of fuzzy spatiotemporal knowledge representation and reasoning based on high-level Petri nets [J]. *Information Systems*, 2012, 37(3): 238-256.
- [6] Pelekis N, Theodoulidis B, Kopanakis I, et al. Literature review of spatiotemporal database models [J]. *The Knowledge Engineering Review*, 2004, 19(3): 235-274.
- [7] Nandal R. Spatiotemporal database and its models: a review [J]. *Journal of Computer Engineering*, 2013, 11(2): 91-100.
- [8] Ma Z M, Yan L. Fuzzy XML data modeling with the UML and relational data models [J]. *Data Knowledge Engineering*, 2007, 63(3): 972-996.
- [9] Yan L, Ma Z M, Liu J. Fuzzy data modeling based on XML schema [C]//Proceedings of the 2009 ACM Symposium on Applied Computing. Hawaii, 2009: 1563-1567.
- [10] Franceschet M, Montanari A, Gubiani D. Modeling and validating spatiotemporal conceptual schemas in XML Schema [C]//Proceedings of the 18th International Conference on Database and Expert Systems Applications. Regensburg, 2007: 25-29.