

Web 环境下地理空间数据的开放式表达体系研究

周文生¹ 毛 锋¹ 胡 鹏²

(1 清华大学建筑学院, 北京市海淀区清华园, 100084)

(2 武汉大学资源与环境科学学院, 武汉市珞喻路 129号, 430079)

摘 要: 对Web环境下的适合分布式异构地理空间数据表达的地理空间数据模型进行了研究, 在此基础上提出了虚拟地图、实体地图以及可视化地图的概念, 并分别基于XML、GML以及SVG创建了开放式地理空间数据的三层表达体系, 该体系的建立为开放式WebGIS的数据互操作奠定了基础。

关键词: XML; GML; SVG; 地理空间数据表达

中图分类号: P208

经过近几年的发展, WebGIS的理论、技术、产品以及应用都取得了很大的进步。然而, 目前的WebGIS产品是在不同的环境中独立开发的, 有着自己的文化背景、领域背景和技术背景, 形成了自己的数据模型和功能结构。虽然这些产品在功能和问题描述能力方面大同小异, 但实际操作时差别却很大, 加之内部空间数据组织互相保密, 形成了不同系统间的壁垒。这种封闭、独立的系统对空间数据的处理保持着一种完全封闭的状态, 导致系统间无法共享数据和处理方法。这种状况严重阻碍了WebGIS的进一步发展, 同时也对国家空间数据基础设施(NSDI)、区域空间数据基础设施(RSDI)、全球空间数据基础设施(GSDID)以及数字地球(DE)的构想和实施十分不利。

针对上述问题, 笔者认为, 其根本出路在于地理空间数据在Web环境下的开放式表达, 而扩展标记语言(extensible markup language, XML)作为一种能够描述复杂数据结构的语言, 无疑会对地理空间数据的开放式表达产生重要的影响。

1 Web 环境下的地理空间数据模型研究

地理空间数据模型是用数据对现实世界地理现象的抽象表达, 是地理空间数据库中用于信息表示和提供操作手段的形式构架, 它是研究地理

空间数据表达和管理的基础。目前, 人们对地理空间数据模型的研究已比较深入^[1~3], 然而对在Web环境下适于地理空间数据表达的地理空间数据模型的研究还很欠缺。为了能对Web环境下的地理空间数据进行有效的表达, 需要对此进行深入的研究。笔者从地理空间实体的分类和地理空间数据的组织两个方面来对Web环境下的地理空间数据模型进行研究。

1.1 地理空间实体的分类

地理空间实体是一种在现实世界中不能再划分为同类现象的现象, 它是组成地理空间数据的基础, 是地理空间数据模型乃至GIS的基础。由于地理空间实体具有复杂性和多样性, 对地理实体的分类通常需要同时兼顾其空间要素和属性要素两个方面。考虑Web环境下地理空间数据表达的需要, 并参照文献[1~3], 笔者将空间实体根据其几何形态划分为5种类型, 即点状实体、线状实体、面状实体、注记体以及栅格体。

1.2 地理空间数据及其组织

考虑到Web环境的特殊性, 地理空间数据的组织应采用简单的数据结构, 同时考虑到大多数现有的地理空间数据采用的多是分层模式, 因此, 笔者采用了分区(title)、分层(layer)的数据组织、管理模式。区域是一定地理范围内所有地理空间数据的集合, 图层则是一定地理范围内具有相同性质的空间实体的集合。通常可将地理区域划分为若干

工作区,各工作区包含若干图层,每个图层又包含若干地理空间实体。这样,地理区域、工作区、图层以及地理空间实体就构成了一种层次模型。

2 Web 环境下地理空间数据的三层表达体系

Web 环境下地理空间数据的三层表达体系就是从 WebGIS 对地理空间数据处理的角度,来研究上述地理空间数据模型在地理空间数据的不同处理阶段的表达方法。这 3 个表达层面分别是基于 XML 的虚拟地图表达、基于 GML 的实体地图表达以及基于 SVG 的可视化地图表达,并分别对应于 WebGIS 中的地图数据的请求、地图数据的集成以及地图数据的显示。图 1 给出了这三者之间的关系。

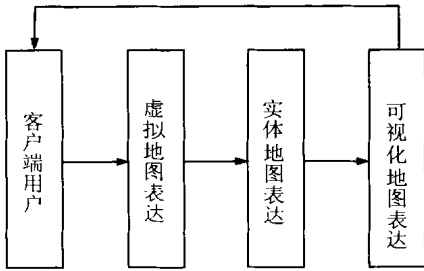


图 1 Web 环境下地理空间数据的三层表达体系
Fig. 1 Three-tier Frame of Representation for Geospatial Data in Web

2.1 基于 XML 的虚拟地图表达

虚拟地图是指网络环境下在地理空间元数据的支持下,对分布式异构地理空间数据的一种高度抽象化的描述,它是对分布式异构地理空间数据集成的一种手段。在虚拟地图中并不包含实际的地图数据,它只是用来描述组成地图的各图层数据的来源、图层数据显示样式等地图图层的控制信息。

虚拟地图的表达需要在地理空间数据资源元数据库的支持下才能完成。地理空间数据资源元数据库是用来对分布式异构地理空间数据的元数据进行统一管理的动态信息库,分布在网络各节点处的地理空间数据的拥有者,可将需要进行共享的数据资源注册到该地理元数据库中。

在地理空间数据资源元数据库的支持下,虚拟地图的描述内容用 XML 的 DTD 表达如下:

```
<! ELEMENT Map (BoundingBox, MapLayers)>
<! ATTLIST Map
  name CDATA #REQUIRED
  format (SVG | GML) #REQUIRED
```

```
>
<! ELEMENT BoundingBox (Left, Right, Bottom, Top)>
<! ELEMENT MapLayers (MapLayer+)>
<! ELEMENT MapLayer (LayerID, LayerStyle)>
<! ATTLIST MapLayer
  Attribute (true | false) #IMPLIED
>
<! ELEMENT LayerID (#PCDATA)>
<! ELEMENT LayerStyle EMPTY>
<! ATTLIST LayerStyle
  Color CDATA #IMPLIED
  SymbolID CDATA #IMPLIED
>
```

2.2 基于 GML 的实体地图表达

实体地图是指在网络环境下,按照虚拟地图的描述,通过虚拟地图处理程序而抽取的分布式异构地理空间数据集中的地理特征数据。它是虚拟地图的具体表现形式,也是分布式异构地理空间数据处理的基础。尽管分布式异构地理空间数据各自所采用的数据模型或存储方法不同,但在实体地图中通过统一的数据表达方法有效地消除了这种差异,使分布式异构地理空间数据通过实体地图进行有效的集成,同时也使分布式异构地理空间数据可以采用相同的方式进行处理。

与虚拟地图一样,实体地图中地理特征数据的表达也可以采用多种方式,但 GML 是进行实体地图表达的最理想方法。基于 GML 进行实体地图的表达,有两个问题需要解决:①在 GML 中只是对地理特征的表达,而没有图层的概念,而在实体地图中采用的是一种以图层为单位的数据组织模式;②由于 GML 在设计时只关注数据内容,并不关心它如何进行显示,因而没有地理特征的显示样式信息,所有这些都是要在实体地图中需要表达的内容。为此,笔者在 GML 的基础上增加了图层元素,以便对分布式异构地理空间数据采用前述的地理空间数据模型来重新组织。此外,考虑到地图数据的显示要求,在图层中还增加了图层显示样式元素来存放图层显示样式的数据。同时,为进行分布式异构地理空间数据的同步编辑处理,在图层元素中还增加了数据版本信息,以便处理程序能根据该版本信息来保证数据处理的一致性。GMLMap 文档 DTD 定义的片断如下:

```
<! ENTITY % gmlgeometry. dtd SYSTEM
" http://www. opengis. org/techno/specs/00-029/gmlgeometry.
dtd" >
%gmlgeometry. dtd;
<! ELEMENT GMLMap ( gml: name, gml: boundedBy,
MapLayers)>
```

```

<! ELEMENT MapLayers (MapLayer?)
<! ELEMENT MapLayer (Style? Features)
<! ATTLIST MapLayer
  Type (Annotation | Line | Point | Polygon | Raster) #
REQUIRED
  Name CDATA #REQUIRED
  Version CDATA #REQUIRED
>

```

上述 DTD 定义中, 图层类别通过 MapLayer 的 type 属性来表示, 其中点图层、线图层、面图层、笔记层以及栅格层分别与 point、line、polygon、annotation 以及 raster 对应。图层的样式数据通过 <style> 元素来表示, 地理特征的属性数据通过 <properties> 元素和 <property> 元素来表示。

2.3 基于 SVG 的可视化地图表达

可视化地图就是实体地图在客户端的可视化表示, 这种可视化表达既可采用栅格的形式(如 JPG、GIF 格式), 也可采用矢量的形式(如 SVG、VML)。

考虑到 SVG 所具有的特点, 笔者采用 SVG 作为可视化地图的描述工具。由于 W³C 设计 SVG 的主要目标是用于一般矢量数据的显示, 而地理空间数据是一种很复杂的数据, 它包括图形几何数据、属性数据、栅格数据以及拓扑数据等多种类型的数据。目前的 SVG 标准并不支持拓扑关系、时态、地理坐标系以及三维数据, 特别是不直接支持地理空间对象属性信息。然而, 由于 SVG 是基于 XML 的, 因此, 可根据地理空间数据的特点对 SVG 进行必要的扩展, 以便通过 SVG 能够有效地描述地理空间数据。

通过对 SVG 的分析, 笔者认为, 直接利用 SVG 进行地理空间数据的表达涉及三方面的内容: ①在 SVG 文档中, 对地理空间数据进行合理、有效的组织, 以便在客户端能够有效地控制地图数据。考虑到 SVG 对图形的描述能力(不支持拓扑关系), 笔者采用前述的空间数据模型来进行地理空间数据的组织。②对 SVG 进行必要的扩充, 增加新的与地理空间数据表达有关的元素(这里称为地理元素), 以支持对地理空间数据的表达。③对根据上述两点要求所组织的地图提供相应的地图操作接口, 地图操作的内容包括地图的显示控制, 如地图的缩放、漫游、地图的图层显示控制、地图的属性数据查询等。对于地图操作接口的实现可参见文献[1]。

2.3.1 地理元素的设计

通过分析, 笔者认为, 将 SVG 用于地理空间数据的描述需增加以下必要的地理元素。

1) 地图描述元素 GeoMap

地图信息包含描述地图的基本信息, 是地图的元数据, 它在空间数据的应用中有很重要的作用。然而在 SVG 中, 只在根元素 SVG 中设定了 width、height 等属性, 远不能满足描述地图基本信息的需要。为此, 引入 GeoMap 元素。其定义为:

```

<! Element GeoMap CDATA # IMPLIED
<! ATTLIST GeoMap
  MapName # REQUIRED
  Scale CDATA # REQUIRED //地图的比例尺, 用比例尺分母表示
  CoordType CDATA # REQUIRED //地图文档采用的坐标系统
  Projection CDATA # REQUIRED //地图文档采用的投影方式
  Xmin # REQUIRED //以下 4 个属性值表示地图的范围
  ...
  XGeomMin # REQUIRED //与(Xmin, Ymin)点对应的大地坐标
  YGeomMin # REQUIRED
>

```

2) 图层描述元素 GeoLayer

图层信息是反映一个图层的结构化数据, 尽管在 SVG 中提供了起注释作用的 <desc> 元素和 <title> 元素, 但它们并不适合描述图层信息。为此, 引入 GeoLayer 元素。其定义为:

```

<! Element GeoLayer
<! ATTLIST GeoLayer
  LayerName CDATA # REQUIRED //图层名称
  LayerType (Point | Line | Polygon | Annotation | Raster) # REQUIRED //图层类别
  ScaleMax CDATA # IMPLIED //图形显示时的最大比例尺, 用比例尺分母表示
  ScaleMin CDATA # IMPLIED //图形显示时的最小比例尺, 用比例尺分母表示
>

```

3) 地理属性数据元素 GeoAttribute

属性数据是地理空间数据中重要的组成部分, 由于 SVG 中并没有提供表示这类数据的专用元素, 因此, 在这里引入地理属性数据元素 GeoAttribute, 以便将属性数据直接与图形数据组织在一起, 并可为以后的数据处理提供统一的操作接口, 有利于系统功能的实现。地理属性数据元素 GeoAttribute 的定义为:

```

<! Element GeoAttribute (*Field) //属性组元素
<! ATTLIST FieldNum CDATA # REQUIRED //属性域个数
<! Element Field CDATA # REQUIRED //属性域元素

```

```
<! ATTLIST Name CDATA # REQUIRED //属性域名称
<! ATTLIST Type (Integer | Float | Sting | Date |
Boolean) //属性域类型
```

2.3.2 地理空间数据的编码

从前述的地理空间数据模型的描述可以看出,可视化地图需要描述的对象有三大类,即地图对象、图层对象以及地物对象。为了能够对所描述的对象进行有效的操作,在这里对所有的描述对象均采用 SVG 所提供的分组元素 *g* 作为主元素,同时为了能够区别这三类对象,在这里规定,对于地图对象,分组元素的 *id* 属性代表地图编号,以“Map”单词开始;对于图层对象,分组元素的 *id* 属性代表图层编号,以“Layer”单词开始。

1) 地图对象的编码

地图对象包含两部分内容,即地图信息以及所包含的各图层对象,其中地图信息通过地图描述元素 *GeoMap* 来表示。

2) 图层对象的编码

图层对象通过一分组元素来表示,该元素的 *id* 属性表示图层的标识符,对于图层的显示样式,不同类型的图层有不同的规定。①点图层。由于点状地物是通过点状符号类进行表达的,因此,点状图层除了需要定义点状符号的颜色外,还要对点状符号进行定义。②线图层。在 SVG 中,对于传统的线状符号支持不够,目前只能通过线宽、线型(通过 *stroke-dasharray* 可任意定义)两个参数来进行描述。③面图层。在 SVG 中,面状符号可通过 *Pattern* 元素来实现,由于面状符号是通过点状符号按一定规律进行的排列,因此,面状符号是在点状符号的基础上进行定义的。④注记层。相对来说,注记层的样式要简单得多,只需指定 *style* 属性中的 *stroke* 即可。⑤栅格层。栅格层不需要指定任何显示样式。

3) 地物对象的属性数据编码

对于地物对象的属性数据可采用内嵌法和外联法两种编码方法。内嵌法是指将所关联的属性数据与几何数据放在同一个地物分组元素中,由于 SVG 并未提供对属性数据的描述标记,因此,在这里使用 *GeoAttribute* 元素来标记属性数据。外联法是指属性数据通过地物标识号存储在外部数据库中。为此,地物标识号需要进行进一步的约定,即地物标识号除了以“F*”开始外,它由地物类别号和目标标识号两部分组成,其中通过地物类别号可确定与该类地物所对应的属性关系表,而通过目标标识号来定位属性表中的记录号。此外,也可直接通过超链接将属性数据页面(通过 ASP 技术动态生

成)与几何数据联系起来。

4) 地理实体的几何数据编码

由于 SVG 已经提供了一些基本的图形元素,如圆、多边形、路径等,因此,对于几何数据的描述可直接采用这些元素。①点状实体。由于在点状实体所在的图层中已经通过 *symbol* 元素定义了点状符号,因此,在这里只要对所定义的符号通过 *use* 元素引用即可。②线状实体、面状实体。由于在 GML 中,地理特征的几何属性仅限于简单几何体,因此,这里的线状实体与面状实体都通过 *path* 元素来表示,至于图形的具体显示样式通过对图层显示样式的定义来完成。③注记体通过 *text* 元素直接表示。④栅格体直接通过 *image* 元素进行表示。

3 结 语

基于 XML 的地理空间数据的三层表达体系,由于采用的是一种完全开放的技术,因此,地理空间数据在 Web 环境下是以“透明”的方式来进行处理和传输的,这就使得基于这种技术的 WebGIS 更容易将分布式异构地理空间数据进行集成处理,使系统的开放性大大加强。笔者基于上述技术提出的基于 XML 的开放式 WebGIS 整体解决方案^[4],其主要特征就是依据该方案构建的 WebGIS 系统完全符合了 OGC 对 WebGIS 所提出的 4 个基本要求,即互操作性、可扩展性、数据、服务分布性以及交互性,从而有效地解决了目前 WebGIS 所存在的瓶颈问题。

参 考 文 献

- 1 龚健雅. 整体 GIS 的数据组织与处理方法. 武汉: 武汉测绘科技大学出版社, 1993
- 2 郭仁忠. 空间物体分类及空间物体构造. 武汉测绘科技大学学报, 1994, 19(1): 22~27
- 3 崔伟宏. 空间数据结构研究. 北京: 中国科学技术出版社, 1995
- 4 周文生. 基于 XML 的开放式万维网地理信息系统的理论与实践. [博士论文]. 武汉: 武汉大学, 2002
- 5 孙一中. XML 理论和应用基础. 北京: 北京邮电大学出版社, 2000
- 6 Geography Markup Language (GML). <http://www.opengis.org/techno/specs/00-029/GML.html>, 2003

第一作者简介: 周文生, 博士后, 副教授。现主要从事网络地理信息系统、摄影测量与遥感以及城市规划与设计方面的研究。代表成果: 基于 SVG 的 WebGIS 研究; Web 环境下基于 SVG 的地图浏览器的实现; 基于 Web 的厂区管网地理信息系统的开发等。

E-mail: zwswh@sohu.com

Study of Three-Tier Frame for Open Representation of Geospatial Data in Web

ZHOU Wensheng¹ MAO Feng¹ HU Peng²

(1 College of Architecture, Tsinghua University, Tsinghuayuan, Haidian District, Beijing 100084, China)

(2 School of Resource and Environment Science, Wuhan University, 129 Luoyu Road, Wuhan 430079, China)

Abstract: This paper studies the geospatial data model suitable for representation for distributed heterogeneous geospatial data in Web, then proposes the new concepts of virtual map, entity map and visual map. A three-tier frame of open representation for the geospatial data based on XML, GML and SVG is presented, namely the virtual map based on XML, the entity map based on GML and the visual map based on SVG.

Key words: XML; GML; SVG; representation of geospatial data

About the first author: ZHOU Wensheng, post-doctoral fellow, associate professor, engaged in the research on WebGIS, photogrammetry and remote sensing, city planning and design. His achievements include: research of WebGIS based on SVG, design and implement of the SVG map browser in Web; development of pipeline networks management information system based on Web; etc.

E-mail: zwswh@sohu.com

(责任编辑: 晓平)

《武汉大学学报·信息科学版》编辑委员会

名誉主任: 宁津生

主任: 李德仁

委员: 毋河海 王新洲 刘 甬 刘经南 刘耀林 朱元泓 朱灼文

仲思东 张正禄 张祖勋 苏光奎 杜清运 杜道生 李建成

李清泉 郑肇葆 柳建乔 晁定波 龚健雅 舒 宁 詹庆明

主 编: 李德仁(兼)

副 主 编: 柳建乔(常务)