

文章编号: 1005-8451 (2016) 04-0006-04

基于开源数据引擎的铁路基础设施地理空间 数据管理方法的研究

程智博¹, 魏小娟²

(1. 中国铁道科学研究院, 北京 100081;

2. 中国铁道科学研究院 电子计算技术研究所, 北京 100081)

摘要: 通过对目前铁路基础设施管理定位现状的研究, 针对大部分采用商业型数据引擎增加成本的问题, 提出一种基于PostGIS的线性参考模型。对铁路一维线路空间数据进行组织和存储, 实现铁路线性资产空间位置信息的定位管理, 为验证该方法应用的可行性给出运行实例, 利用铁路地理信息平台上应用服务, 达到正确描述铁路线性资产的地图位置和降低成本的目的。

关键词: PostGIS; 线性参考; 定位管理

中图分类号: U2: TP39 **文献标识码:** A

Methods of railway infrastructure geo-spatial data management based on Open Source Database Engine

CHENG Zhibo¹, WEI Xiaojuan²

(1. China Academy of Railway Sciences, Beijing 100081, China;

2. Institute of Computing Technologies, China Academy of Railway Sciences, Beijing 100081, China)

Abstract: This article researched on the location situation of railway infrastructure management, proposed a kind of line reference model based on PostGIS to solve the problem of increased cost caused by using commercial database engine. The one-dimensional railway line data were managed and stored to implement the location management of railway linear assets spatial position information. The operation instance was given to verify the feasibility of the model application. The railway geographic information platform was used to upload application service, achieve the purpose of describing the map position of railway linear assets correctly and reducing the cost.

Key words: PostGIS; linear referencing; location management

铁路基础设施是指构成铁路路网的固定设施, 主要包括轨道、线路、路基、桥涵、隧道等。这些基础设施的空间数据多以线状表达, 具有海量、类型复杂等地理特征, 需要使用空间数据库将固定设施的空间位置进行管理。现有的铁路基础设施台账管理方法中, 采用线路-里程的方式进行位置定位, 借助于一维线性度量单位表达铁路沿线设备的位置。通过WebGIS 二维空间坐标展示铁路线路信息时, 并没有考虑到铁路设施数据的一维线性特征, 如何正确地融合两者成为管理和维护铁路基础设施地理信息数据的关键手段。因此, 在铁路空间数据库中, 组织和存储铁路一维的线路空间数据, 并且能够完善地管理、表

达设施设备的空间位置信息, 成为亟待解决的问题^[1]。

随着多种GIS 软件支持对PostgreSQL 空间数据库的访问, 利用PostgreSQL 和PostGIS 数据引擎管理铁路线性空间数据, 能够增强空间数据库的集成部署的灵活性、降低铁路GIS 项目的开发成本, 相对于传统的商业型数据引擎, 具有强大的优势。

1 线性参考和PostGIS

1.1 线性参考

线性参考 (Linear Referencing) 即利用已有线性要素的相对位置存储地理数据的一种方法, 无需显示 (X, Y) 坐标而唯一地标识线路位置的过程^[2]。主要特点是, 仅用一个表达里程的参数来定位属性和事件, 动态地参考和创建线性结构的各个部分^[3]。

收稿日期: 2015-10-08

作者简介: 程智博, 在读博士研究生; 魏小娟, 助理研究员。

线性参考主要使用具有唯一标识 ID 的路径度量值 (M) 描述与沿线自定义起点的相对位置。 M 值存储在事件表中, 将线路从多段线 (*Polygon*) 类型变成带有里程值的多段线 (*Polygon M*) 类型。例如通过带有里程值的线路进行设备资产定位, 描述列车的位置“距京沪高速起点 60 km 处”比描述 (116.72,39.49) 的 GPS 坐标更有意义。图 1 所示不同维度表达坐标点和里程值的关系。

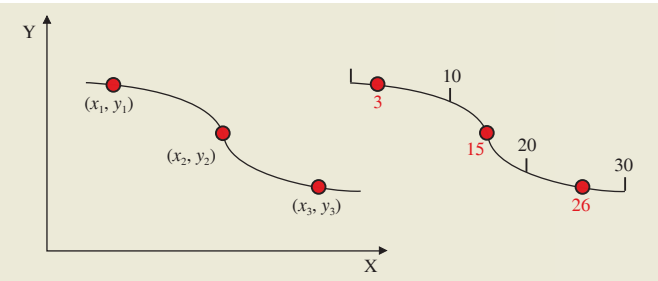


图1 一维和二维的不同线性表达方式

1.2 PostGIS

空间数据的存储与管理是 GIS 的基本功能, 如何对一维线性的铁路资产在空间数据库中进行管理也是研究热点。目前, 主流的商业型空间数据库引擎有 Oracle Spatial、ArcSDE、SQL Server Spatial 等, 然而高昂的购置成本有时让人望而却步。随着开源数据库的发展, 并且遵循开放地理空间联盟 (OGC) 制定的规范, PostgreSQL 和 MySQL 等免费的空间数据库越来越受到用户的欢迎。作为较主流的两款开源数据库, 其空间插件 PostGIS 和 MySQL spatial extensions 管理地理空间数据具有自身的特色, 相比较而言, PostGIS 能够提供更多的空间索引类别、常用的投影坐标系、复杂的数据操作和分析功能, 将属性信息和空间信息存储在一张表里对空间对象进行完整性表达^[4], 这些优势增加了空间数据的易用性和互操作性, 本文选择 PostGIS 数据引擎管理铁路线性数据。

2 基于PostGIS的线性参考模型

采用开源数据库平台 PostgreSQL 9.4 和 PostGIS 2.1, 基于线性参考方法进行建模, 以验证一维线性铁路设备资产在开源数据库中的定位可行性。通过 PostGIS 提供的扩展程序封装线性参考的算法, 能够

解决一维线性数据和二维地理坐标数据的里程转换。图 2 是线性参考模型示意图。

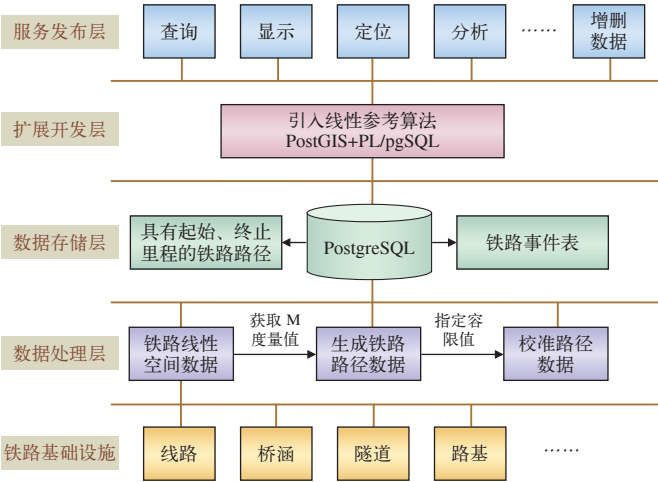


图2 线性参考模型流程图

2.1 数据处理层

数据处理层是基础层, 根据铁路不同业务部门的设施设备台账和已有地图数据建立线路的空间实体, 尽量保证线路的连续性和完整性。可以根据自定义或者读取其他字段的方式获取路径的度量值, 相应地设置其他参数, 得到具有 M 字段的路径。通过与线路对应的点文件进行校准, 进行线路纠正, 得到与之匹配的线路要素类, 此时线上的每一个点都具有 M 值。满足式 (1):

属性值 ∈ {objectID, Shape, lineID} (1)

其中: ObjectID 是属性表每行具有唯一标识的 ID; Shape 类型是 (*Polygon M*), lineID 是标识线路的 ID。

2.2 数据存储层

数据的存储是数据入库的过程, PostGIS 是通过自身内建模板数据库进行新建空间数据库, 其空间支持功能加载到 postgis_template 模板空间数据库中, 这样通过 SQL 语句即可创建数据库。选择 ogr2ogr 命令导入 shp、KML、XML 等格式的文件, 不同的文件格式采取相应的数据导入的方式, 入库的要素类支持路线空间数据库和事件属性数据库的关联。存储的数据由 PostGIS 支持的文本表达方法 (WKT) 和二进制表达方法 (WKB) 描述几何类型。

2.3 扩展开发层

编写用户定义函数 (UDF) 是 PostgreSQL 的强

大功能之一。PL/pgSQL 是一个功能强大的 SQL 脚本语言, 存在于绝大多数的 PostgreSQL 安装文件中, 基本不需要对它进行重复设置。该层的主要功能是实现二维空间和一维线性位置参照系的转换, 其转换方式有 2 种: (1) 从一维线性的数据到二维坐标系的转换; (2) 从二维地图坐标到一维线性数据的转换。

2.3.1 第1种转换方式

试图求出在线路长度范围内某里程值 (M) 所对应的点的二维空间坐标 (X, Y)。这种转换需要插值和拟合即可。为了保证线路的光滑度, 采用 3 次样条插值方法, 数学模型建立如下:

设 $y=f(x_i)$ 是区间 $[m_1, m_n]$ 上的连续函数, f 在 $[m_1, m_n]$ 的各区间函数 $s(x_i)$, 满足:

$s(x)$ 在每个区间内都是三次多项式;

$s(x)$ 在每个节点内具有二阶连续导数;

$s(x_i)=y_i$

得到每个小区间 $[m_{i-1}, m_i]$ 的表达式:

$$s(x_i)=a_{i0}+a_{i1}x+a_{i2}x^2+a_{i3}x^3, (i=0,1,2,\cdots,n) \quad (2)$$

该表达式所绘制的曲线能够较好地逼近原始线路, 通过拟合方式解析函数, 将整个线路的图上距离 (Shape Length) 进行归一化, 通过相对于整条线路的浮点数 $A \in [0,1]$ 处的里程值得到相应的二维点坐标。同时, 也可以通过自定义起始里程和终止里程, 得到在此区间内的子线路。

使用 PL/pgSQL 语言编写算法、封装函数, 通过传入 Float 型的里程值和 Integer 型的线路标识, 对指定线路具有 (MulitiLineString) M 类型的多段线进行赋值、插值等操作, 最终返回可被解析为文本类型的点坐标 PointM 或 LineString。

2.3.2 第2种转换方式

根据已知的二维平面点坐标 (X, Y), 得到该点 (在线路上) 或者距离线路最近的点在线路上的里程值 (M)。在铁路线性要素数据的表达过程中, 随着线性要素属性的变化, 将线路在需求的位置处打断以反映数据的变化。因此采用分段曲线拟合的方式对两种位置的关系进行转换^[5]。

假设铁路线路上有 $(n+1)$ 个采样点, 分为 L 段, 则有:

$$n+1=n_1+n_2+\cdots+n_L \quad (3)$$

可以建立该点 (X, Y) 到线路的距离 ($1 \geq 0$) 的 n 次多项式:

$$R_k(m)=\begin{cases} a_{k,0}+a_{k,1}m+a_{k,2}m^2+\cdots+a_{k,t}m^{t_k} \\ b_{k,0}+b_{k,1}m+b_{k,2}m^2+\cdots+b_{k,t}m^{t_k} \end{cases} \quad (4)$$

其中, $k=1, 2, \cdots, L$ 。

根据最小二乘法, 进行分段插值, 得到最小值的 $a_{k,0}, a_{k,1}, a_{k,2}, \cdots, a_{k,t}$, 以及 $b_{k,0}, b_{k,1}, b_{k,2}, \cdots, b_{k,t}$ 。根据求

出的每一段的 $\begin{cases} X_k(m) \\ Y_k(m) \end{cases} (k=1,2,\cdots,L)$, 得到 $R(m)$ 的表达式:

$$R(m)=\begin{cases} R_1(m), m \in (m_0, m_{n_1}] \\ R_2(m), m \in (m_{n_1}, m_{n_1+n_2}] \\ \vdots \\ R_L(m), m \in (m_{n_1+n_2+\cdots+n_{L-1}}, m_n] \end{cases} \quad (5)$$

通过分段式的最小二乘法进行拟合插值, 能够保证相邻分段之间的平滑性和连续性。通过传入 POINT 型的 (X, Y), 判断该点是在线路上还是线路外, 若在线路上, 直接返回距离线路起始点的里程值; 若在线路外, 根据其最短距离映射到线路上得到里程值。使用该方法可避免使用 PostGIS 内部函数在弧段、折角等处使用直线距离取代曲线距离而产生的误差。

2.4 服务发布层

针对 PostgreSQL 是一个自由的关系型数据库, 可以采用 Web Services 技术, 达到实现异构系统平台之间的互操作性的目的^[6], 将系统的数据显示和需求作为一种服务, 通过服务的请求调用获得数据库的数据共享和应用。基于 Web Service 的关系型数据动态发布服务系统主要包括数据库管理模块、构建服务模块、解析信息模块、服务发布模块, 最终生成服务数据库, 可供用户调用服务进行查询参数和结果, 发布服务流程图如图 3 所示。

基础数据库采用 PostgreSQL 数据库, 存储处理后的基础线性数据, 并可以对数据进行添加、删除、修改和查看等基本操作。通过对数据库添加所需的配置表和必须的字段信息进行构建服务模块。在解析信息模块中, 通过调用 PostGIS 中写入的函数, 将约束条件映射到信息模块中, 使得数据库能更好

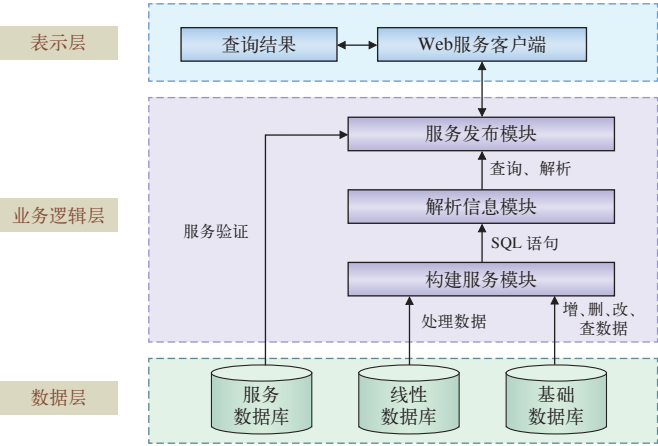


图3 发布服务流程图

地识别 SQL 语句。在服务发布模块中，采用 REST (REpresentational State Transfer) 的结构设计，将数据服务转移向应用服务，利用其提供的服务接口，将查询信息存储到服务数据库中，通过输入查询的关键字等信息，获取查询结果，实现系统间的数据服务应用。

3 应用实例

采用 B/S 架构，应用和数据库部署在同一台服务器上，使用 PostgreSQL 数据库及 PostGIS 插件、Tomcat 7.0 中间件，运行于 Windows Server 2008 R2 操作系统，客户端使用 Windows 7 操作系统，IE 8/9/10/11 浏览器，在铁路地理信息平台中注册服务，在该平台的线性参考模块中查询到上传的服务，界面如图 4 所示。



图4 服务管理界面

根据发布的“坐标转里程”和“里程转坐标”两种 Web 服务，选择线路所在的空间数据库，可以依据不同的线路类型进行查询，如选择铁路正线、站线等，构造完成的查询实例界面如图 5 所示。

查询经过服务引擎转化成对应的 UDF 语句，在

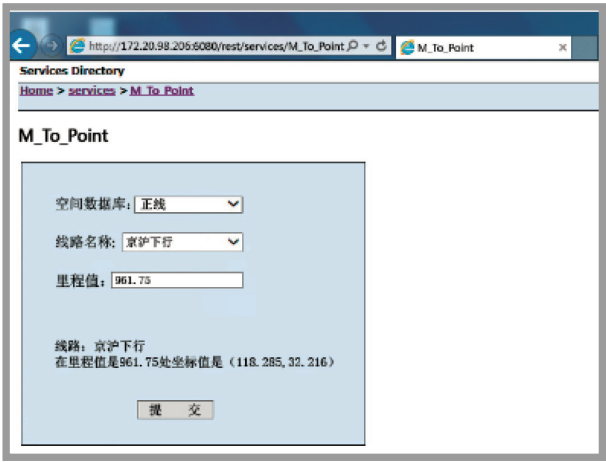


图5 里程转坐标查询界面

调用服务的过程中将输入的数据转化为参数与服务接口中的函数参数相匹配，转到 PostgreSQL 数据库中，将查询结果返回给 Web 界面。

4 结束语

通过对 PostGIS 空间数据引擎的研究，根据铁路线性要素的一维表达方式的实际情况，提出了基于 PostGIS 的线性参考模型，并通过 PostgreSQL 作为后台数据库进行发布 Web 服务。该应用实例证明了 PostGIS 开源数据库引擎的实用性和可操作性，能融合于铁路地理信息平台中，完成铁路一维线性资产的地理定位表达，为今后地理空间数据的引擎选择提供了更广阔的空间。

参考文献：

[1] 史天运,王英杰,李 平.数字铁路框架体系的研究[J].交通运输系统工程与信息,2010,10(6):29-33.

[2] ESRI 环境系统研究所公司(美国).ArcGIS 中的线性参考[Z].美国:ESRI 环境系统研究所公司,2004.

[3] 程 亮,龚健雅,朱明媛,等.线性参考系统技术在道路资产管理系统中的应用研究[J].测绘科学,2008,33(4):206-208.

[4] 侯 伟.开源空间数据库的数据表达,管理与分析[J].测绘通报,2012(S1):602-604.

[5] 蔡先华,齐 征,戚洁平,等.基于线性 LRS 与基于 SRS 的空间数据转换方法[J].公路交通科技,2005,(S2):165-168.

[6] 尹兆冰,王加阳.Web Service 及其关键技术研究综述[J].软件导刊,2010,9(2):121-123.

责任编辑 徐侃春