

文章编号：1005-8451（2011）01-0043-03

基于通信线路资源管理系统异构数据迁移的实现

王源，刘云

（北京交通大学 通信与信息系统北京市重点实验室，北京 100044）

摘要：本程序基于通信线路资源管理系统，需要把老系统数据导入新系统中。由于新老系统数据结构不同，且基于 WebGIS 地理信息系统，数据特殊，数据量大，手工导入数据费时费力，用数据导入软件又不能很好满足本系统导入数据需要。所以需要设计一套可移植性和可操作性较好的程序来进行数据的迁移，能够对老系统的地理信息进行准确处理，从而将有用数据迁移到新系统的数据库中。

关键词：数据迁移；地理信息；坐标转换

中图分类号：U285 **文献标识码：**A

Implementation of heterogeneous data migration based on lines of Communication Resource Management System

WANG Yuan, LIU Yun

(Key Laboratory of Communication & Information Systems, Beijing Municipal Commission of Education, Beijing Jiaotong University, Beijing 100044, China)

Abstract: This Program Resource Management System based on communication lines, needed to import the Old System data to the New System. The System was WebGIS-based Geographic Information System. As the Old and the New System data structures were different, it was time-consuming manual import data, and there was not a software to meet the needs of the System. So it was necessary to design a better portability and operability of the procedures for data migration. The Old System could be accurate geographical information processing, useful data could be migrated to the new System's database.

Key words: data migration; geographic information; coordinate transformation

铁通通信网络资源管理系统采用 GIS 技术，80%以上数据具有地理属性，构成通信网络资源的线路资源、设备资源和信息资源等所涉及到的数据和地理位置、图形信息的关系十分密切。GIS 技术与电信管道资源分布情况结合起来，在 GIS 中以图形化方式直观体现。老系统不能很好满足新系统导入数据需要，新老系统的结构层次不同，系统属性字段不同，存在很多一对多或者多对一的对应关系，用存储过程实现起来比较麻烦。因此需要一种方法既能进行批量导入老系统的数据，又能使异构的老系统数据匹配新系统的数据库。

1 数据结构

系统主干数据结构分4个类型：机房内的局站设备，机房外的管道杆路，光缆设备和电缆设备。

收稿日期：2010-05-24

基金项目：国家“863”基金项目（2009AA01Z423），北京交通大学

基本科研业务费项目（2009JBZ005）。

作者简介：王源，在读硕士研究生；刘云，教授。

每种类型又细分出多种层次的数据类型。如图 1。

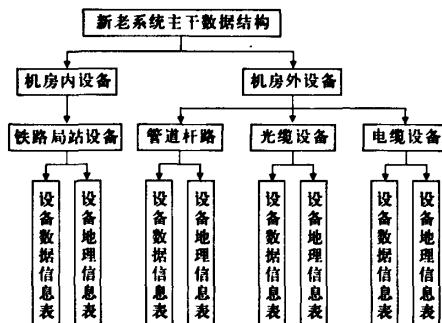


图 1 新老系统主干数据结构

2 Oracle Spatial 结构与坐标转换

2.1 Oracle Spatial 的数据存储结构

在 ArcGIS 中通过 SDE 存储空间数据到 Oracle 存储方式有：二进制 Long Raw、ESRI 的 ST_Geometry 以及基于 Oracle Spatial 的 SDO_Geometry 等。这里主要用到了 Oracle Spatial 的存储方式，通过这种存储方式几何列 Shape 的字段类型为

mdsys.sde_geometry 类型。Oracle Spatial 定义的 SDO_GEOMETRY 类型为：

```
CREATE TYPE sdo_geometry AS OBJECT(
    SDO_GTYPE NUMBER,//前面字符串为
    字段名；后面字符串为字段类型
    SDO_SRID NUMBER,
    SDO_POINT SDO_POINT_TYPE,
    SDO_ELEM_INFO
    SDO_ELEM_INFO_ARRAY,
    SDO_ORDINATES
    SDO_ORDINATE_ARRAY);
```

其中 SDO_geometry AS OBJECT，标识该类型为对象类型。至于该类型中的 SDO_POINT_TYPE、SDO_ELEM_INFO_ARRAY、SDO_ORDINATE_ARRAY 也是 Oracle Spatial 自定义的类型和 SDO_geometry 是一样的。

(1) SDO_GTYPE：表示要存储的几何类型，如点线面。SDO-GTYPE 参数值，如表 1。

表 1 SDO_GTYPE 参数值举例

SDO_GTYPE 值	几何类型	相关描述
2001	单点 Point 类型	地理对象包含一个普通的点
2002	单线 Polyline 和 Curve 类型	地理对象包含直线或片段 segments
2005	多点 MultiPoint 类型	地理对象包含多个点的集合

(2) SDO_SRID：几何空间参考坐标系，类型为 NUMBER。SDO_SRID 定义空间坐标参考系统。

(3) SDO_POINT：几何类型就是存储点坐标，否则为空。oracle 自定义的 SDO_POINT_TYPE 类型。SDO_POINT 类型的构造方法为：sdo_point_type(x,y,z)，其中 x,y,z 类型为 Double 和 Int 都可。SDO_POINT 字段定义为含有 X、Y、Z 属性的 SDO_POINT_TYPE 类型。

(4) SDO_ELEM_INFO：定义要如何理解 SDO_ORDINATES 中的坐标串。SDO_ELEM_INFO 类型的构造方法为：sdo_elem_info_array(a, b, c)。每个 SDO_ELEM_INFO 属性单元由 SDO_STARTING_OFFSET、SDOETYPE 和 SDO_INTERPRETATION 组成。SDO_STARTING_OFFSET：声明组成当前几何片段的第 1 个坐标在 SDO_ORDINATES 数组中的坐标序号。SDOETYPE。结合 SDO_STARTING_OFFSET 和 SDOETYPE 表来理解。SDOETYPE 值 =1,2,

1 003, 或 2 003，说明几何为简单的几何类型。SDO_INTERP-RETATION：依赖 SDOETYPE 是否是组合元素，如果 SDOETYPE 值 =1,2,1 003, 或 2 003，标识决定了元素坐标队列的翻译顺序。

(5) SDO_ORDINATES：存储实际坐标的，以 X、Y 以及不同点之间都是逗号隔开。

2.2 坐标变换

系统采用 WGS-84 坐标系。该坐标系是目前 GPS 所采用的坐标系统，GPS 发布的星历参数就基于此坐标系统。WGS-84 世界大地坐标系，它是一个地心地固坐标系统。该坐标系的坐标原点位于地球的质心，Z 轴指向 BIH1984.0 定义的协议地球极方向，X 轴指向 BIH1984.0 的启始子午面和赤道的交点，Y 轴与 X 轴和 Z 轴构成右手系。采用椭球参数为：长半轴 $a=6\ 378\ 137\ m$ ，协议地球扁率 $f_{84}^{-1}=298.257\ 223\ 563$ 。

一种形式就是一种坐标系。如空间直角坐标系 (X, Y, Z)、大地坐标系 (B, L) 和平面直角坐标 (X, Y) 等。老系统中存储的是坐标系统下的空间直角坐标，新系统需要的是以经纬度表示的大地坐标系。需要将空间直角坐标系转换成大地坐标系，得到大地坐标 (B, L)。根据公式：

$$\begin{cases} L = \arctan(Y/X) \\ B = \arctan(Z(N+H)/[(X^2+Y^2)^{1/2}(N(1-e^2)+H)]) \\ H = Z \sin B \cdot N(1-e^2) \end{cases}$$

用上式采用迭代法求出大地坐标 (B, L)。L 为大地精度，B 为大地纬度， $N=a/(1-e^2 \sin^2 B)^{1/2}$ ，N 为该点的卯酉圈曲率半径； $e^2=(a^2-b^2)/a^2$ ，a, e 分别为该大地坐标系对应椭球的长半轴和第 1 偏心率。

3 数据迁移程序设计

3.1 数据迁移流程

本程序使用 Java 开发，运行本程序要启动与源数据库和目的数据库的连接。获得源数据表格和字段、目的数据表格和字段及所需条件进行下一步。通过数据处理类，利用刚获得的源数据表格和字段及查询条件，从源表中查询出需要导入的数据放入一个对象中。之后，把该对象内的数据进行格式转化或者特殊处理，根据已获得的目的数据表格和字段，生成可执行的 sql 语句，输出到目的数据库中，并生成一个输出文档。程序流程如图

2. 需要导入的数据主要分为设备数据信息表和设备地理信息表。

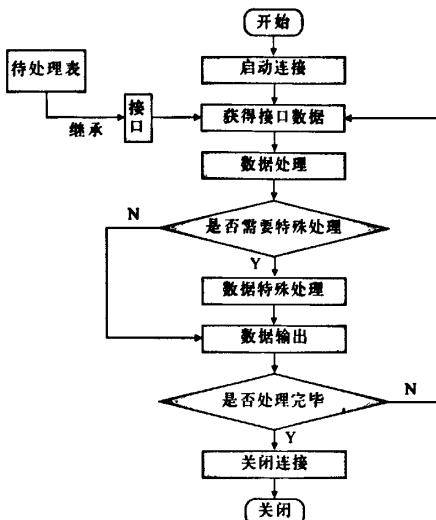


图2 程序流程图

3.2 设备数据信息程序

Java 是一种面向对象的语言。面向待处理的表，设计一个容器的接口程序。程序定义了 6 种类型接口，如表 2。分析待处理的表之间的关系后，就可以新建对应的 Java 类，继承该接口，在新的类中放入这 6 种对象就可以被别的类调用。随着待处理的表格增多，该方式把复杂的程序转化为一种简易的模板而不用反复修改程序，在模板中添加有用信息便可以达到目的，方便快捷。这种通过继承的方法实现数据处理，灵活性、可移植性比较强。另外，经常会用到多级嵌套的 sql 语句。

表2 接口程序设计

接口类型	定义
String[] getSourceTable()	获得源库的表名，源表可能有多个，所以为数组
String[] getSourceFields()	获得源字段名
String getDestTable()	获得目标库表名
String[] getDestFields()	获得目标库字段名
String getCondition()	从源数据片获得数据的条件
void process(String[] destFields, Object[] values, DBA writer)	按照目标库字段的类型处理要导入的值。

3.3 设备地理信息程序设计思想

导入地理设备信息表的难点在于读出的地理信息需要经过坐标转化和 SDO_Geometry 的使用。根据 2.1 节提到的 Oracle Spatial 的数据存储结构，不同的地理信息需要按照相应的格式导入

数据库中。

MDSYS.SDO_Geometry(2001,8307,
MDSYS.SDO_Point_Type(x1,y1,null),null,null)
数据结构表示手孔表的地理信息是单点 Point 类型，WGS-84 坐标系，点坐标是 x1, y1。

MDSYS.SDO_Geometry(2002,8307,null,
MDSYS.SDO_Elem_Info_Array(1,2,1),
MDSYS.SDO_Ordinates_Array(x1,y1,x2,y2))
表示管路表的地理信息的地理对象包含直线，且
第 1 个顶点在 SDO_Ordinates_Array 开始位
置是 x1，元素类型是 2 简单的几何图形，顶点连
接方式是通过直线连接，x1, y1, x2, y2 是该管
路的实际坐标。

数据处理需要对原表中的数据进行处理，原表的 id 比较短，经过 SYS_GUID 处理，即使数据量较大时，SYS_GUID 也会保证它创建的标识符在每个数据库里都是唯一的。还有一些表所属地区有所变更，要在特殊处理时进行。

4 结束语

基于通信线路的 WebGIS 工作原理，设计了导入导出数据程序，实现了把不同结构的数据库匹配导入到新系统数据库中。处理掉老系统中冗余的空值、废值，导入效率达到 90% 以上。该程序面向对象设计，简单方便，可移植性强，可修改性强。利用该程序，提高导入数据的效率，避免软件导入数据可修改性差的弊端，真正匹配系统数据导入的需要，是 Java 面向对象程序设计的范例。

参考文献：

- [1] Kevin Loney, Bob Bryla. Oracle 10g DBA 手册[Z]. 北京：清华大学出版社，2006.
- [2] Gay S.Horstmann, Gray Cornell. java 核心技术卷 II：高级特性[M]. 北京：机械工业出版社，2008.6.
- [3] 郭杰华, 鲍远律, 胡玉锁, 等. 基于 Internet 的地理信息系统 (WebGIS) 的研究和开发[J]. 微机发展, 1999, 9 (1): 61-63.
- [4] 朱志伟, 刘书雷, 李军. 一种基于 Oracle Spatial 的 WebGIS 实现方法[J]. 计算机工程, 2004, 30 (2): 98-99.
- [5] 柳光魁, 王振裸, 等. BJ-54 坐标系与 WGS-84 坐标系转换方法及精度分析[J]. 测绘与空间地理信息, 2007, 6 (30).