

# 利用 FME 进行 GIS 数据的无损转换

陈 影, 程耀东, 闫浩文

(兰州交通大学数理与软件工程学院, 兰州 730070)

**【摘 要】**在研究 FME 语义映射文件的组成格式、定义方法、执行过程的基础上, 利用修改 FME 语义映射文件的方法, 对 DWG 到 SHAPE 格式的数据无损转换进行了研究。结合科研项目, 作者实现了数据从 AutoCAD 的 DWG 格式到 ArcGIS 的 SHAPE 格式的无损、快速、高效地转换, 有效地解决了“如何实现数据在不同平台之间的转换, 实现数据共享”这个制约 GIS 项目进度的瓶颈问题。

**【关键词】**地理信息系统; 数据格式转换; 空间数据操作引擎; 语义映射文件

**【中图分类号】**P208

**【文献标识码】**A

**【文章编号】**1009-2307(2007)02-0075-03

## 1 引言

FME (Feature Manipulation Engine 空间数据操作引擎) 是 safe 公司推出的新产品, 一个第三方专门用于数据转换的工具, 它有非常强大的功能, 支持近百种格式的数据相互间的转换, 如: AutoCAD DWG/DXF, ESRI Arc/Info Generate, ESRI Export (E00 Uncompressed), ESRI SDE, ESRI Shape, Intergraph MGE, Intergraph/MicroStation Design File, MapGuide SDL, MapInfo MIF/MID, MapInfo TAB, Oracle Spatial Cartridge (Relational), Oracle SQL Loader, Smallworld, SDE2.1 等<sup>[1]</sup>, 并能够同时进行几何数据及其属性的无损转换, 也可以实现自主扩展需要转换的数据格式, 让用户添加自己定义的数据格式, 与其他数据格式进行转换。可以对数据进行批处理, 这样大大提高了工作效率。

## 2 FME 的组成部分

FME 的版本包括: FME Desktop Edition、FME Professional Edition、FME Enterprise Edition 等。最新版本包括以下四部分: ①FME Registration 是映射文件的登记系统, 用修改语义映射文件的方法转换数据, 如果修改后不登录登记系统, 它的执行命令就不能被 FME 所识别, 所以必须建立与系统的链接关系, 使它成为 FME 不可分割的一部分, 只有这样才能起到修改文件的目的, 所以语义映射文件只有在此登记后才能在数据转换中随时被调用; ②FME Universal Translator 是转换系统, 在这个系统中可以根据需要选择源数据和目标数据的格式以及路径, 并进行自动转换; ③在 FME Universal Viewer 中可以随意打开源数据和目标数据, 以便进行检查、对照; ④FME Workbench 是面向对象的转换器, 在此系统中可以根据需要手工设计转换器, 以实现数据的无损转换<sup>[1]</sup>。新版本在各方面都进行了完善。

## 3 FME 语义映射文件

### 3.1 语义映射文件的组成

FME 语义映射文件应用的是一种宏语言, 有自成体系的变量和函数。很多映射文件要引用输入数据集、输出数据集, 或者其他映射文件 (例如 mapinfoMacros.fmi), 此时可能用到 FME 预先定义的一些宏。另外它利用 TCL (Tool Command Language) 作为内置语言。

语义映射文件的内容一般包括: 读写器定义及其配置、要素表、函数、要素动态生成器四部分, 其中最灵活的是要素表。用户应根据目标数据的要求选择不同的分类器, 软件将在调用映射文件时按照要素表选用的函数顺序执行转换。

在项目设计中通常要按照国标或其他标准对地图要素进行分类, 在对数据转换前应保证图面上的任一图元均属于特定的要素类, 属于同一要素类的图元具有某些相同的特征, 如图层、线型、符号名相同等, 可通过 Feature 表来描述要素类的特征, 凡是符合这些特征的图元均属于这个要素类<sup>[2]</sup>。

Feature 表是用来描述要素类特征和控制要素类行为的, 是按照设计要求或作业标准即目标数据的格式设计的控制文件。Feature 表包括两部分内容: 一部分用于描述要素类特征, 具备这些特征的都属于同一要素类; 另一部分是附加部分, 例如在进行数据转换时, 需要调入地理信息系统的目标数据所需的某些属性在源数据中没有体现, 那么就要在转换前在语义映射文件中添加。

### 3.2 语义映射文件的生成

FME 可以用多种方法生成映射文件, 最简单的方法是用 FME 自动映射文件产生器生成一个映射文件作为模板, 这个模板是根据系统提示自动生成的, 然而它一般不能满足用户的需要, 所以应根据自己的需要进行修改, 登记后即可在日后的转换工程中进行调用。另一种比较简单的方法是利用 Workbench 生成, 在 Workbench 中定义好转换器后, 可以用映射文件形式显示转换过程。转换过程就是每层数据、每个要素以及要素属性的对应过程, 映射文件就是记录这个过程的载体。另外还可以用 INCLUDE 的扩展语法动态生成映射文件。

### 3.3 语义映射文件的定义

语义映射文件的语句以关键字开始, 并用关键字控制后续语句的解释。关键字以 MACRO 开头 (必须是大写), 然后是宏的名字, 在该行后面是需要替换的内容。引用宏时利用 \$( ) 将宏的名字加上括号, 如下是一个宏的生成过程:

```
MACRO blue 8 IGDS 32 igds_ color $(blue) igds_
```

**作者简介:**陈影 (1981-), 男, 河北献县人, 硕士研究生, 专业为地图学与地理信息系统, 研究方向: GIS 应用技术及 GIS 系统建库。

E-mail: chenqy\_2005@163.com

**收稿日期:**2006-03-07

**基金项目:**兰州交通大学“青蓝人才工程资助计划”基金资助项目



style 3

会被表示为:

IGDS 32 igds\_color 8 igds\_style 3

语句太长时, 可以用“\”分行。宏变量用“\$( )”引用。单行注释以“\”开始, 并忽略所有的空格和空行, 多行注释可以用“/\* \*/”方式来完成<sup>[2]</sup>。

当 INCLUDE 目标以 ( ! ) 符号开头时, FME 认为它是程序<sup>[2]</sup>, 其标准输出被包含进来, 而当以 ( [ ] ) 符号开头时, FME 认为它是 TCL 脚本, 在调用文件时被包含进来。

### 3.4 语义映射文件的执行过程

源数据中的要素 (如 MicroStation 中的线要素) 和目标数据中的要素 (如 AutoCAD 中的线) 一般并不能直接建立转换对应关系, 源数据中的要素首先要表达为 FME 的格式 (FME 通过 fme\_type, fme\_geometry 等属性表示源要素), 这是一种中间格式或称为伪格式, 伪格式是 FME 的特定格式。目标数据 (SHAPE 格式) 是依据地理信息系统的入库标准制定的, 可通过编写代码对源要素进行处理, 具体的处理受读写器 Reader 和 Writer 配置的影响, 伪格式中的信息与目标数据的信息是一一对应的, 包括各类要素及要素的属性, 处理的信息就包含在伪格式中。最后将伪格式转换为目标格式, 从而实现数据转换。如果只处理几种要素类型, 可以直接建立源要素和目标要素的对应关系, 这样更简便。

## 4 FME 在项目中的应用

### 4.1 技术方案

江门市规划局在建立数据库过程中所选用的源数据格式是 DWG 格式, 目标数据格式是 ArcGIS 的 SHAPE 格式。

在外业采集数据后, 可利用成图软件进行数据处理。ArcGIS 被人们广泛用于开发地理信息系统, 但它不能直接读入 DWG 格式的数据<sup>[3]</sup>, 所以选用 FME 软件, 用于把原始数据格式转换为 ArcGIS 能够识别的 SHAPE 格式, 以便数据的顺利入库<sup>[4]</sup>。转换过程如图 1 所示:

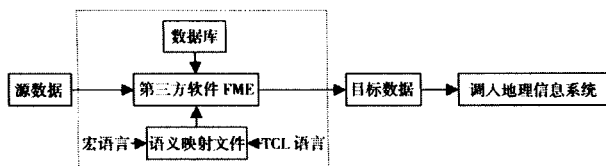


图 1 源数据到目标格式数据的转换过程

### 4.2 前期准备

若要做到数据的快速、高效转换, 除了能对 FME 进行熟练的操作外, 还要对源数据和目标数据格式有具体的了解<sup>[5]</sup>, 这样才能更好地设计要素表。这次项目的源数据是广州开思的 DWG 格式, 目标数据格式是 ArcGIS 的 SHAPE 格式。由于此系统主要用于国土规划, 所以定制了独立的数据库标准, 例如层名、要素名、所需要的特定的要素属性等。

除了用于数据转换, FME 还有纠错功能。在数据转换前可以用 FME 检查源数据, 例如检测源数据分层是否正确, 错误的要进行手工处理, 以便项目的顺利完成。所以要做以下几方面工作:

①掌握源数据及目标数据的数据格式及对应关系<sup>[6]</sup>; ②熟悉地理信息系统的入库标准; ③根据入库要求对源数据做必要修改; ④依据源数据和目标数据作一个数据表。

FME 有连接数据库的功能, 可以根据字段索引逐一读入记录。是否要建立外部数据库要看数据的复杂程度, 如果需要, 可以用 Microsoft Access 依据源数据和目标数据建

立一个数据对应表。

建立数据对应表时, 应严格遵守地理信息系统空间数据库设计标准的要求, 并参照源数据的要素编码表, 建立源数据与目标数据的对应关系。建立数据对应表的依据是“江门市基础空间数据库设计方案 (大比例尺)”, 这个方案是专门为江门市基础空间数据库提供的建库标准。如表 1 是控制点类的入库标准。

表 1 测量控制点类

图层名称	主要内容	实体类型	层名
测量控制点	测量控制点	P	CTLPNT
控制点注记	测量控制点的注记	P	CTLANN
控制点辅助线	控制点名或点号 与该点高程之间的横线	L	CTLASO

建立的数据对应表见图 2。

ID	主参数	次参数	要素名称	shaper	att	att	shapetype	MAP
2	ZBTZ	Contour 0	防火带	VEGLIN			SL	962003 775
9610000	2	ZBTZ	I4 0	地类界	VEGLIN		SL	961003 774
9540000	3	ZBTZ	(ZBTZ) GC163 半荒漠植物地	VEGLIN			SP	954023 773
9530000	3	ZBTZ	(ZBTZ) GC124 花圃	VEGLIN			SP	953023 772
9520000	3	ZBTZ	(ZBTZ) GC164 植物稀少地	VEGLIN			SP	952023 771
9510000	3	ZBTZ	(ZBTZ) GC161 芦苇地	VEGLIN			SP	951023 770
9440000	3	ZBTZ	(ZBTZ) GC162 灌草地	VEGLIN			SP	944023 769
9430000	3	ZBTZ	(ZBTZ) GC176 人工草地	VEGLIN			SP	943023 768
9420000	3	ZBTZ	(ZBTZ) GC176 改良草地	VEGLIN			SP	942023 767
9410000	3	ZBTZ	(ZBTZ) GC121 天然草地	VEGLIN			SP	941023 766

图 2 源数据与目标数据对应表

数据对应表建好后可以和 FME 连接, 以便在系统运行时被调用。可以在语义映射文件中建立它们的连接关系, 部分语句如下所示:

对于线要素读数据库的表: gtc2000\_compare, 先打开数据库, 再读取数据

```

Relate TABLE_ LOCATION gtc2000 _ compare "
$(FME_MF_DIR) \
gtc2shape.mdb"
Relate TABLE_ DEF gtc2000 _ compare DATABASE
\

```

```

DATABASE_SERVER_ TYPE MDB \
DATABASE_USER_NAME "" \
DATABASE_PASSWORD "" \
DATABASE_NAME "" 以上代码建立了 FME
与数据库的连接关系

```

与数据库的连接关系

```

图层 column_type_unneeded \
主参数 column_type_unneeded \
编码 column_type_unneeded \
shaperlayer column_type_unneeded \
shapercode column_type_unneeded \
Relate RELATION_ DEF " JOINER_line" 1: M \
TABLE gtc2000_compare \
UNIQUE(图层) \
JOIN 图层 TO autocad_layer \
TRANSFER shaperlayer TO shaperlayer
TRANSFER shapercode TO shapercode

```

以上代码读取建立了对应关系的各个字段的数据

### 4.3 语义映射文件的修改

由于数据比较复杂, 用 FME 自动映射文件产生器自动生成的映射文件不能实现对数据的无损转换, 所以要对文件内容做必要的修改。

修改时要参照目标数据的入库标准, 分层处理, 各层再分别对不同要素进行处理。下面以转换居民地要素为例

进行说明。

首先, 设置源数据及目标数据格式, 设计如下所示:

READER\_ TYPE DWG

表示读入的数据格式是 DWG 格式

WRITER\_ TYPE SHAPE

目标格式是 SHAPE 格式

系统的执行语句(读入数据后)如下:

SHAPE\_ DEF RESRGN \

SHAPE\_ GEOMETRY shape\_ polygon \

表示目标数据的多边形要素, 以下依次是房屋属性

设置

CODE char(6) \

表示字长

ADRE char(64) \

表示字长

STRUCT char(80) \

FLOOR number(8, 0) \

SCODE char(12)

以上代码是数据的 SHAPE 定义, 数据入库后 JMD (居民地)要素放于 RESRGN 层, 这个定义要依据预先定制的 SHAPE 数据的入库标准, 要素需要什么属性先要在此定义。

下面是设置 DWG 和 SHAPE 格式的对应关系, 在 SHAPE 格式设置中, 依次是系统中的房屋需要的属性, 其中 autocad\_ thickness 是 DWG 格式的数据要素的固有属性, 即要素的编号。这些属性可以在地理信息系统的属性查询列表中显示。实现代码如下:

DWG JMD \

DWG 格式数据的 JMD 层

autocad\_ entity autocad\_ polygon \

AutoCAD 中的多边形要素

autocad\_ thickness% autocad\_ thickness \

%表示取值, 在此表示的是取 AutoCAD 的线的厚度值

shaperlayer" " \

目标数据的层

code\_ 121200 \

给 code\_ 1 赋值为 21200

layer\_ 2RGN \

把 RGN 赋值给 layer\_ 2

SHAPE " \* " \

CODE @ Concatenate( % code\_ 1, 7) \

"@" 表示调用函数, Concatenate 表示连接函数, 把 code\_ 1 的值与 7 连接, code\_ 1 的值为 21200, 连起来为 212007

ADRE " " \

STRUCT " " \

FLOOR 1 \

SCODE % autocad\_ thickness \

SCODE 的值取 AutoCAD 中要素的厚度值

@ FeatureType(@ Concatenate( RES, % layer\_ 2))

#### 4.4 数据转换后的效果图

语义映射文件修改好后, FME 在执行转换操作时将调用该文件。如下两图是单个房屋要素转换前后对比图, 图3为 DWG 格式, 图4为 SHAPE 格式。

对比前后图可以看出, 转换后的图形完全符合入库标准。

如下是点、线、面要素及要素属性入库后的效果图(见图5-7):

其他要素的转换原理与此相似, 通过测试, 转换后数据完全符合入库要求。入库后图形要素及要素属性也没有缺失。

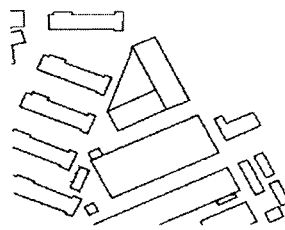


图3 转换前房屋要素图



图4 转换后房屋要素图

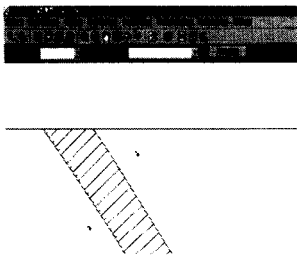


图5 点要素

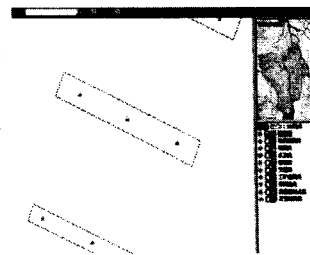


图6 花园

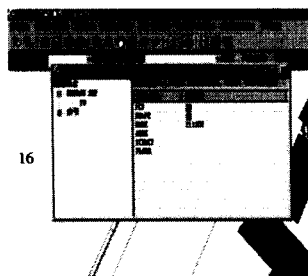


图7 属性对话框

## 5 结束语

经过实践证明, FME 在 GIS 数据跨平台转换方面具有良好的通用性和可操作性, 功能强大。它具有简单的操作界面, 可以用不同的操作方式实现同一种数据在不同平台或不同数据在同一平台的共享, 正逐渐成为各种 GIS 软件共享数据的事实标准<sup>[7]</sup>。作者所研究的数据转换方法对从事地理信息系统开发的技术人员具有一定的参考价值。

### 参考文献

- [1] FME 技术白皮书 [Z]. 北京世纪安图数码科技发展有限公司, 加拿大 Safe Software Inc FME 中国总代理, 2000.
- [2] 李长松, 李辉. FME 在 DLG 数据入库中的应用 [EB/OL]. 黄河规划设计, <http://www.yrcc-design.com.cn/>
- [3] 周小成, 等. 基于 Geodatabase 的 CAD 数据到 GIS 的解决方法 [J]. 现代测绘, 2004, 27(6): 15-17.
- [4] 叶慧芳, 胡小伍. AUTOCAD 与 Arc / Info GIS 文本数据的转换及 GIS 文本数据库的建立 [J]. 地矿测绘, 2004, 20(4): 19-21.
- [5] 刘增量, 刘鹏. GIS 数据向交换格式数据转换的方法与实现 [J]. 北京测绘, 2004, (2): 44-47.
- [6] 钟世彬, 郑贵洲. AutoCAD 和 MAPGIS 间的数据转换 [J]. 测绘科学, 2005, 30(3): 97-98.
- [7] 姚占派, 原瑞红. 简介 FME 在 GIS 数据交换中的应用 [J]. 广东测绘, 2004, (3): 27-29.