

一种数据采集和地图快速更新的方案设计

A Scheme Design for Data Collection and Quick Map Update

于池

Chi Yu

河北省制图院 中国·河北 石家庄 050031

Hebei Provincial Institute of Cartography, Shijiazhuang, Hebei, 050031, China

摘要: 近年来地理信息技术发展日新月异,各种新应用、新技术不断涌现。在数据收集、存储、处理、输出和应用方面有很多优秀的开源数据库和软件可供选择。适当利用这些开源软件,不但能够很好地解决问题,而且还可以节约成本。论文主要描述了利用 Python、PostgreSQL 等技术软件采集不动产单元属性信息采集,同时根据所采集的信息动态更新地图专题图的方案,满足了工作需求,为后续工作的开展奠定了良好的基础。

Abstract: In recent years, geographic information technology has undergone rapid development, with various new applications and technologies constantly emerging. There are many excellent open source databases and software options available for data collection, storage, processing, output, and application. Proper use of these Open-source software can not only solve problems well, but also save costs. The paper mainly describes the scheme of using technical software such as Python and PostgreSQL to collect attribute information of real estate units, and dynamically updating thematic maps based on the collected information, which meets the work requirements and lays a good foundation for subsequent work.

关键词: 地图更新; Arcpy; PostgreSQL; PostGIS

Keywords: map updating; Arcpy; PostgreSQL; PostGIS

DOI: 10.12346/se.v5i1.8121

1 引言

近年来,笔者单位一直承担中国河北省集体土地所有权数据、宅基地和集体建设用地使用权等数据的质量检查任务,具有丰富的基础数据。根据工作需要,要实现不动产单元的登记状况信息汇总,并在地图专题图上实现自动更新。为此我们设计了一种数据集采集和地图快速更新的方案,主要思路是基于现有数据,按行政区划生成调查任务数据,通过 FTP 服务器实现任务数据的下发和调查成果的上交,通过 Python 自动化监测、检查上交的调查成果,符合要求的成果自动更新入库,不符合要求的调查成果,反馈检查结果,直至成果符合要求,整体实现登记信息地图专题图的自动更新。

2 关键技术介绍

2.1 Python 与 Arcpy

Python 是一种高级的跨平台解释性语言,所谓的解释语

言就是指在程序运行之前,不需要提前将代码编译成机器语言,而是边解释边运行。Python 语言从诞生发展到现在已经得到广泛应用,拥有非常丰富的第三方代码库可供选择,可以用来编写脚本,执行流程化操作,利用 Pandas、Numpy 代码库进行数据分析,利用 Tkinter 等代码库进行 GUI 编程,利用 Django、Tornado 等框架进行网站建设,利用 Scrapy 等框架进行编写爬虫程序。可以说 Python 是一个功能强大的多面手,同时由于其跨平台以及语法简洁明了的特性^[1],成为很多项目理想的语言。

ArcPy 是由 Esri 公司开发可以跟 ArcMap 桌面软件实现高度融合的 Python 代码库^[2]或者叫软件包,能够独立或者在 ArcMap 等软件内执行地理空间数据处理、分析、管理功能。利用 Arcpy 软件包以及其 Python 第三方代码库创建工作流,可以实现对文件夹、文件的流程化操作。目前,Arcpy 软件包凭借高可用性、便捷性、敏捷开性能,已广泛

【作者简介】于池(1989-),男,中国河北邢台人,硕士,工程师,从事地理信息系统研究。

应用到空间数据处理中^[3]。

2.2 FTP 服务器

FileZilla 是一个免费开源的 FTP 服务器软件，在功能和性能上不亚于付费服务器。FileZilla 可以通过设置用户、用户组，如图 1 所示，为不同的用户和用户组设置相应的共享文件目录、用户名和密码，可以方便地实现用户隔离，也可以为每个用户设置读取、修改、删除等权限，如图 2 所示。只需简单配置即可架设高性能的 FTP 服务器^[4]。

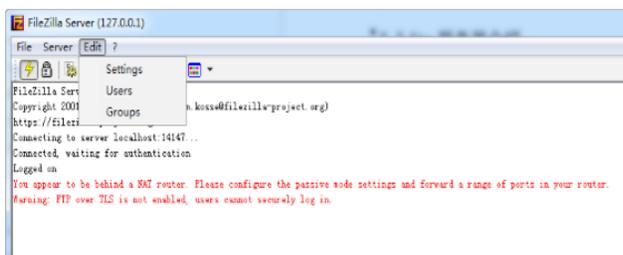


图 1 设置用户

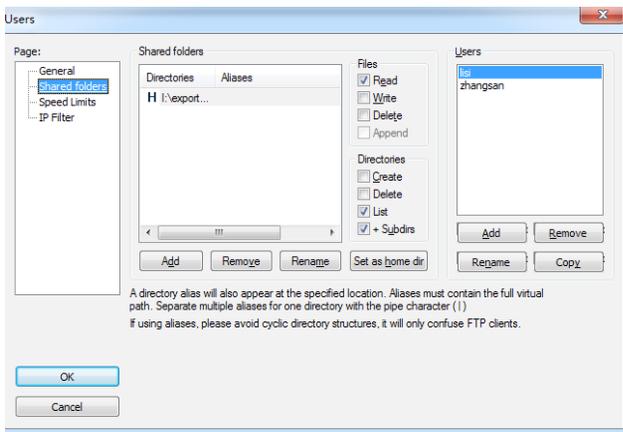


图 2 设置目录权限

2.3 PostgreSQL 数据库

PostgreSQL 是一种开源关系型数据库。基本上能够支持大部分标准 SQL 语法，可以减少技术人员的学习成本。该数据库具有通用关系型数据库查询、Key、触发器、事务处理、版本控制等功能^[5]。同时还提供了 PostGIS 扩展插件，支持对地理空间数据的存储管理，能够与常用的地理信息平台软件直连调用。能够方便地作为数据源进行相关的图形发布操作。PostgreSQL 凭借其强大的地理数据存储处理能力，可以很好地应用于地理空间数据处理^[6]。目前在 GIS 领域得到了广泛应用。本研究采用 PostgreSQL 和 PostGIS，满足了任务需求，取得了良好的应用效果。

3 方案设计

3.1 总体业务流程

工作的总体业务流程为：

①在省级服务器中，以县（区）行政区划为单位，建立文件存放目录（主文件夹），每个主文件夹下再建立 3 个子

目录，用于存放调查任务数据、调查成果数据、问题清单数据。通过 ftp 服务器对每个主文件夹设置用户名、密码以实现用户隔离。

②通过插件的导出功能，按照行政区划，将全省宅基地和集体建设用地图斑中符合要求的属性数据导出，形成调查任务数据（Excel），存放至省级服务器指定的文件存放目录（调查任务数据目录）。

③各县级自然资源部门，通过自然资源专网，下载本地区的调查任务数据，按照要求填报本地区宅基地和集体建设用地登记情况，形成调查成果数据。

④各地填报完成后，通过自然资源专网，将调查成果数据上传至省级服务器指定文件存放目录（调查成果数据）。

⑤插件自动监测调查成果数据目录，发现有新成果上传后，自动读取成果数据，进行数据完整性、有效性检查，通过检查的可以直接导入；对有问题的调查成果数据，生成问题清单并推送到指定目录（问题清单数据），待相关地区修改上传后，重新检查。

⑥通过检查的数据成果可以导入数据库，并与原数据库属性表挂接，形成全省的宅基地和集体建设用地情况的数据库，并实现专题图成果的动态更新。

业务流程图如图 3 所示。

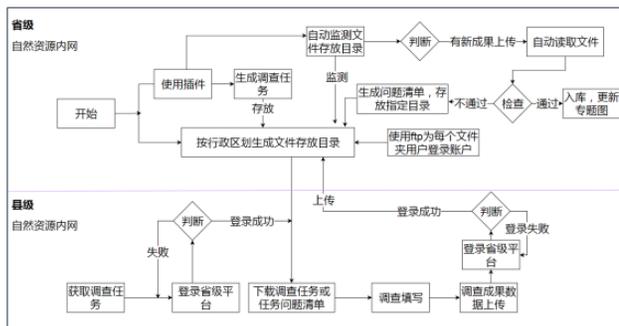


图 3 业务流程图

3.2 功能实现

3.2.1 属性数据导出

利用 Arcpy.da 数据访问模块，按照市、县（区）、乡镇（街道）、村（社区）级行政区读取全省宅基地和集体建设用地图斑中符合要求的属性数据，通过 xlwt、xlrd、xlutils 模块，生成 Excel 类型调查任务数据，并将调查任务数据推送到服务器中指定存放位置。代码实现如下：

```
With arcpy.da.SearchCursor(fc, fields) as cursor:
    totalNum=0
    rowNum = 1
    for row in cursor:
        print(“oper: %s” % row[1].decode(‘utf-8’))
        shiCode = row[2][0:4]
        xianCode = row[2][0:6]
        xiangCode = row[2][0:9]
```

3.2.2 调查数据自动检查

利用 win32file.Read Directory Changes W 模块提供的功能能够监测文件存放目录,当发现有新的调查成果数据(Excel 类型)上传时,自动读取成果数据,进行数据完整性、有效性检查,检查不动产单元号是否变化、登记信息是否按照要求填写,通过检查的可以直接导入;对有问题的调查成果数据,生成问题清单并推送到指定目录,待相关地区修改上传后,重新检查。

3.2.3 调查数据导入

通过检查的数据成果可以导入数据库,并与原数据库属性表挂接,形成全省的宅基地和集体建设用地情况的数据库和专题图成果。代码实现如下:

```
edit=arcpy.da.Editor(workspace)
edit.startEditing(True)
edit.startOperation()
for i in range(1,rowNum):
objid=getExcelStr(worksheet.cell(i,0)) flag=getExcelStr(worksheet.
cell(i,1))
try:
table name=" gis.sde.mapflag"
fields=( 'objectid,' ' flag' )
with arcpy.da.UpdateCursor(table name, fields) as cursor:
for row in cursor:
print "row[0]" +str(row[0])
print "row[1]" +str(row[1])
if str(row[0])==objid:
row[1]=flag
cursor.updateRowrow)
```

```
except arcpy.ExecuteError:
print arcpy.GetMessageO
except Exception as e:
print e.message
edit.stopEditing(True)
```

3.2.4 动态更新专题图

通过 Geo Server 连接 PostgreSQL,将专题图层发布为地图服务,当调查数据导入后,地图服务会根据采集的数据动态更新专题地图,能够保持专题图的现势性。

4 结语

论文从技术角度描述了从数据采集到上传,以及地图自动更新的整体流程。该方案能够提高数据汇交效率,实现专题图数据动态更新,为后续工作提供了较好的技术支撑。

参考文献

- [1] 周崑,徐东炯,沈丽娟.基于Python和ArcPy的专题图批量制作方法[J].测绘与空间地理信息,2020,43(10):17-20.
- [2] 魏智东,李美珍.基于ArcPy地理空间数据库批量裁剪工具设计[J].北京测绘,2021,35(3):385-389.
- [3] 谢正明,王友昆,张君华,等.基于Arcpy和NumPy的DEM的投影形变分析方法[J].测绘地理信息,2022(6):27-31.
- [4] 陈慧民.应用开源软件FileZilla Server架设FTP服务器[J].中国教育网络,2011(7):76-77.
- [5] 钱建国,张天源,吴锋.基于PostGIS的小流域数据库设计与应用研究[J].测绘与空间地理信息,2020,43(2):8-11.
- [6] 梅新,王文荟,钟少波,等.基于PostGIS的线要素空间聚合方法[J].计算机工程,2021(5):308-315.