

资源环境数据格网的构建及其应用

诸云强¹, 朱少霞², 李爽¹

(1 中国科学院地理科学与资源研究所, 北京 100101; 2 首都师范大学资源环境与旅游学院, 北京 100037)

摘要:资源环境数据共享的紧迫性需要一个全新的资源环境研究平台,在该平台上不仅需要实现资源环境数据资源的共享,而且还要实现专业应用系统、计算模型、工具软件的共享。根据上述应用需求,利用数据格网技术,提出了资源环境数据格网的概念。遵循开放式格网服务体系结构构建了三层的资源环境数据格网,探讨了资源环境数据格网构建的关键技术。以1991~2001年环北京地区土地沙化分析为例,阐述了利用资源环境数据格网的应用。

关键词:资源环境;数据共享;格网技术;数据格网

中图分类号:P208

1 引言

数据格网技术的出现改变了人们进行资源环境研究的方式和手段,人们在网络上可以更自由、方便地使用计算资源、数据资源^[1]。利用数据格网技术,构建资源环境数据格网,实现资源环境数据、计算模型、工具软件的共享,为资源环境研究提供“一站式”平台。

资源环境的共享希望能在一个网络平台上实现各种资源环境问题的研究,而不用考虑数据资源来自何方,采用何种数据库进行存储;也不用理解计算分析的具体模型算法等^[2]。资源环境数据格网是以数据格网技术为支撑,将异地分布式的资源环境数据、工具软件、计算模型联接一起,提供一个集成的资源环境研究平台。

2 资源环境数据格网体系结构及其关键技术

2.1 资源环境数据格网体系结构

资源环境数据格网以通用格网中间件为核心,遵循OGSA(open grid services architecture)体系架构,结合资源环境研究特点,分三层进行构建。如图1所示。

资源环境数据格网从逻辑上分为三层:格网门户层、格网业务层和格网资源层。格网门户层是指格网门户网站,通过门户网站,用户可向平台提交各种应用要求,搜索数据、空间数据处理、资源环境现状分析等,平台处理完成后将结果反馈在门户层上。格网业务逻辑层以通用格网组件为核心,结合资源环境数据资源及业务特征新增或改造适应资源环境研究的资源环境数据格网中间件,如资源环

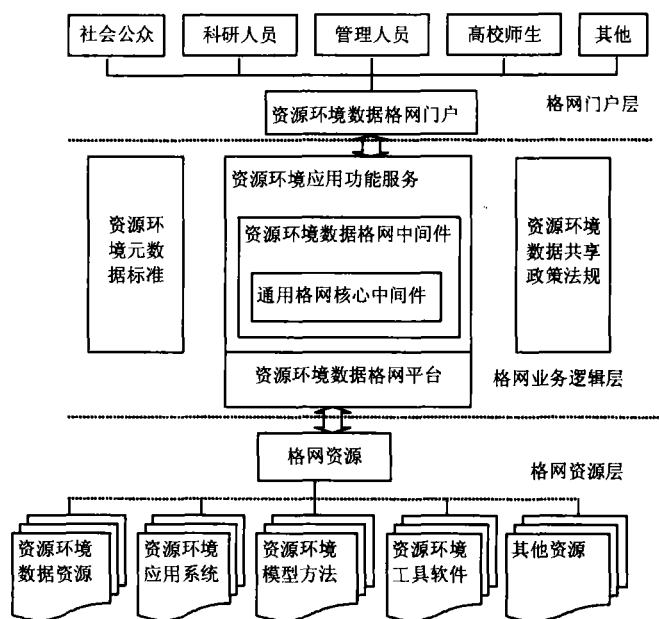


图1 资源环境数据格网体系框架

Fig.1 Structure of resources and environment data grid

收稿日期:2004-12-15; 修回日期:2005-06-12.

资助项目:国家科学数据共享工程试点项目(2002DEA30030)

作者简介:诸云强(1977-),男,博士研究生,研究方向为地学数据共享.

境元数据描述、空间数据访问、空间数据传输等,在此基础上开发资源环境应用服务功能,包括:资源环境数据处理、资源环境现状分析、综合评价、专题制图等。资源环境数据格网资源不仅包括大量的资源环境数据资源,同时包括功能服务资源;各种应用系统、软件工具、模型算法等。

2.2 构建资源环境数据格网的关键技术

(1) 资源环境数据格网中间件

基于通用数据格网组件,结合资源环境业务特点和数据资源特征,研究资源环境数据格网中间件是资源环境数据格网的关键。

随着 Globus 引入 Web Services 概念,它已经成为了格网的事实标准^[3]。由于资源环境数据格网中涉及到大量的空间数据,空间数据不能够用服务来完全描述,如遥感数据、GIS 数据等;同时空间数据的传输也不同于其他信息的传输,通用的格网数据传输和复制协议可能无法满足空间数据的复杂性和海量分布的特点;空间数据的显示、网络操作同样较其他数据复杂^[4]。因此,为了满足资源环境数据格网的需求,必须根据资源环境研究领域的这些特点,基于 Globus 格网组件的基础,对通用的格网组件或第三方组件进行改造,构建资源环境数据格网中间件。资源环境数据格网中间件主要包括资源环境元数据服务、目录服务、空间数据传输中间件、空间数据挖掘中间件、空间数据显示中间件等。

(2) 资源环境数据资源格网化

数据资源是支撑资源环境研究的基础,是资源环境数据格网最基础的部分。目前,资源环境数据分布在各个研究单位中,或以数据库的形式进行存储或以文件形式进行存储。由于资源环境数据类型、数据格式不一,真正在网络上实现共享相当困难。异构数据资源的统一访问技术是实现数据资源共享和调度的首要前提。通过统一访问技术,将存储、检索数据等高层用户的请求映射为数据资源的底层操作,实现广域范围内对数据有效的统一访问和管理。因此,为了实现异构分布式的资源环境数据的共享,必须对资源环境数据进行格网化,即通过对异构分布式的资源环境元数据的转换,屏蔽它们之间的差异,使得用户能在格网平台上共享资源环境数据资源。资源环境数据资源格网化的步骤:① 异构资源环境元数据标准、异构数据元数据转换;② 资源环境数据格网化封装;③ 资源环境数据

操作(注册、搜索、显示等)。根据资源环境数据特征进行资源环境元数据标准研制(元数据标准采用 XML 配合 Schema 表示),根据元数据标准建立资源环境数据格网模型对多源异构资源环境数据元数据进行转换,实现资源环境数据的格网化封装,完成数据格网化后,才能进行数据注册、数据搜索、定位等功能。用户根据资源环境元数据标准提交查询条件,经过中间件处理后,系统返回元数据查询结果,查询结果中包括相应存储路径和数据集元数据信息。查询检索流程如图 2 所示。

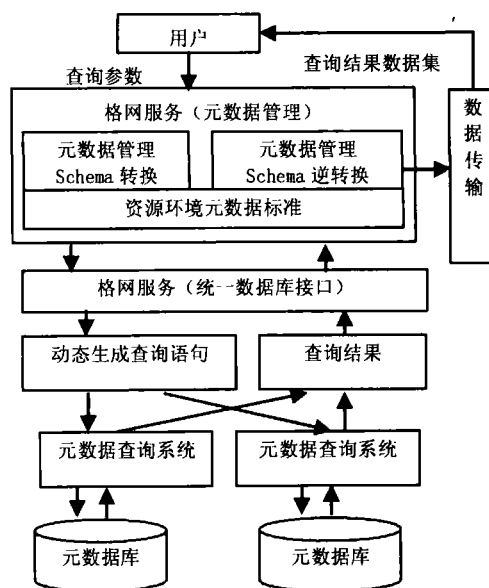


图 2 资源环境数据查询流程

Fig.2 The flow of querying resources and environment data

(3) 资源环境数据格网应用功能

包括:空间数据管理功能和资源环境专业应用功能。在格网平台上,利用 Web Service 功能模块进行封装。Web Service 能够用标准的服务描述语言(WSDL)对服务进行描述,并以 XML 文件形式保存。通过对该 XML 文件的访问,使用服务的用户可以获取该服务的接口特征,如接口名称,接口参数个数,接口参数类型等。WSDL 被赋予网络地址,可以通过 Internet 访问。用户可以直接通过该 WSDL 文件的 URL 访问服务描述文件,也可以通过注册服务器访问该文件^[5]。OGSA 环境中提供注册服务,该服务作为 OGSA 标准核心服务部署在 OGSA 环境中。注册服务是一种特殊的格网服务,它可以管理其他格网服务的 GSH(Grid Service Handle)和 Service Data。GSH 是格网服务的唯一 ID 号,通过

该 GSH 可以访问格网服务;Service Data 是格网服务的一系列属性值,属性值由设计格网服务的人员制订。Service Data 反映格网服务的服务内容。OGSA 体系利用 WSDL 和 Service Data 帮助用户理解和发现格网服务^[6]。

(1)空间数据管理功能,主要包括:空间数据的显示浏览、动态更新、编辑处理;空间数据和属性信息之间的互查;缓冲区分析、叠加分析、邻近分析等多种空间分析功能,其结果依据要求可制作成不同类型的专题图件等。

(2)资源环境专业应用功能,主要包括:资源环境动态信息提取与监测分析,如土地利用构成分析、植被覆盖度分析、水土流失分析、资源环境时空变化分析等;利用数据模型对不同时空尺度的资源环境进行综合评价;资源环境趋势预测,在变化分析的基础上,总结归纳生态环境的区域差异、变化过程特点,发现资源环境变化原因,模拟预测资源环境变化趋势。

3 资源环境数据格网的应用实例分析

资源环境数据格网的应用:

(1)资源环境数据的共享。数据生产者可以注册发布自己的数据,数据使用者可以发现调用自己所要的数据,完全与数据格式、数据资源所处的位置无关。

(2)资源环境计算共享。资源环境数据格网不仅可以实现数据资源的共享,而且可以实现资源环境计算模型、功能操作、应用系统、工具软件的共享。

(3)资源环境集成应用。在资源环境数据格网中,利用 workflow 技术,根据资源环境应用流程提供应用服务链,将数据服务和功能服务有机地集成在一起,通过服务的自动查找、动态选择,自动执行用户的应用请求。当然用户也可以自己定制应用流程,借助资源环境数据格网,完成应用研究。

土地沙化是土地退化重要的组成部分,按土地利用类型可进一步分为沙化耕地、沙化林地、沙化牧草地。以环北京地区(与北京毗邻的河北、内蒙、山西三省区的 51 个县级行政区域)1991~2001 年土地沙化分析为例,说明如何利用资源环境数据格网进行资源环境应用研究。

资源环境数据格网的土地沙化分析基本流程(如图 3):

(1)用户通过资源环境数据格网查询需要的数据资源、空间叠加分析服务和专题图制作服务。在资源环境数据格网中空间数据和空间操作功能函数遵循 OGC 的规范,采用 Web Service 技术被封装成数据服务和功能服务注册到 UDDI(Universal Description,Discovery and Integration,统一描述、发现与集成)注册中心中,平台根据服务分类和服务标识查询得到相应的数据服务和功能服务的 URL。

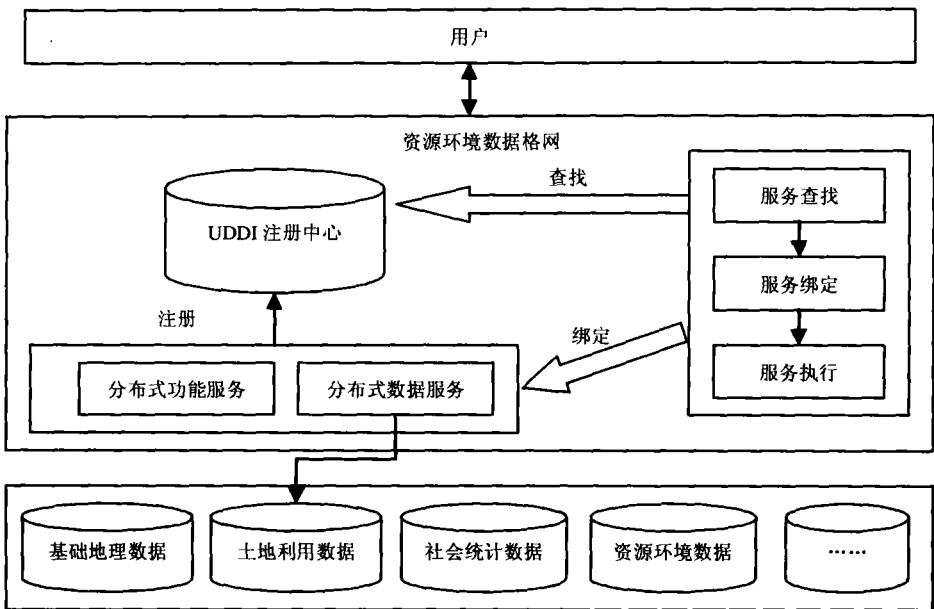


图 3 基于资源环境数据格网的土地沙化分析流程

Fig.3 The analysis flow of sandy desertification based on resources environment data grid

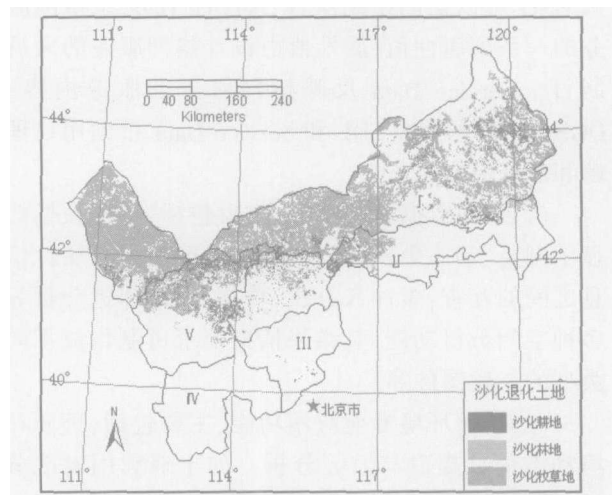
(2)根据空间叠加分析服务的调用规则,执行空间叠加分析服务,完成空间叠加运算。空间叠加分析服务要求输入图层具有相同的坐标参考系,如果数据的坐标参考系不同则需要转换成一致的坐标参考系。空间叠加分析服务输入参数包括叠加处理所需要的二个图层 URL(可以是点图层、线图层和面图层)、空间叠加处理的类型(包括点与多边形叠加、线与多边形叠加、多边形与多边形叠加,土地沙化分析采用后者中的叠合运算)。其操作结果是使用 GML 编码的叠加结果数据 URL。

(3)获得处理结果后,调用专题图制作服务,对土地沙化结果进行分类着色,制作土地沙化分布图。

在环北京地区土地沙化分析中,为了方便用户的操作,将土地沙化分析流程制作成服务链以服务的形式发布到注册中心。用户在土地沙化分析中可直接调用土地沙化分析服务链,同时输入分析的时空参数。格网平台根据服务链设置的抽象服务描述进行服务的查找和动态选择,得到 1991~2001 年环北京地区的土地利用现状数据、叠加分析和专题图制作。根据服务的 URL 将抽象服务绑定到具体的服务,按工作流程自动执行叠加分析运算和专题图制作功能,得到 1991~2001 年环北京地区土地沙化分布图,如图 4 所示。其中,沙化耕地主要分布在阴山北麓及周边的丘陵地区;沙化林地,主要分布在正蓝旗中西部、多伦北部、林西山地、敖汉东北等地;沙化牧草地主要分布在四子王旗、镶黄旗、正镶白旗、正蓝旗、克什克腾西部、翁牛特旗东部、阿鲁克尔沁旗南部等地。在所有沙化退化土地中,沙化牧草地占绝大部分面积。

4 结语

资源生态环境监测等迫切需要加强资源环境数据格网研究。本文就资源环境数据格网的体系结构及其关键技术、资源环境数据格网应用作了初步的探讨,资源环境数据格网的关键技术、应用实践还有待进一步深入地研究。



注: I-内蒙古高原干旱草原生态区(包括浑善达克沙地荒漠草原和昭乌达高原农牧交错区)
II-坝上高原农牧生态区
III-燕山山脉农业生态区
IV-大同盆地及周边地区农业生态区
V-阴山山脉农牧交错生态区

图 4 1991~2001 年环北京地区土地沙化分布

Fig.4 The distribution of sandy desertification in the surrounding areas of Beijing during 1991~2001

参考文献

- [1] 孙九林, 李 爽. 地球科学数据共享与数据网格技术. 地球科学, 中国地质大学学报, 2002, 27(5):539~543.
- [2] 任建武. Grid GIS 关键技术研究. 南京: 南京师范大学, 2004.
- [3] 王意洁, 肖依, 任浩. 数据格网及其关键技术研究. 计算机研究与发展, 2002, 39(8):943~947.
- [4] 李国庆. 空间信息应用格网技术研究. 北京: 中科院地理所, 2004.
- [5] 柴晓路, 梁宇奇. Web Services 技术、架构和应用. 北京: 电子工业出版社, 2002, 12~13.
- [6] 陈 萍, 余华山, 王 彬 等. 网格计算环境 G10bus 介绍. 计算机应用研究, 2003, (8):96~98.

下转 P66

Inter-solution to Coordinate Fit and Creation of Matrix Cells

ZHOU Wei

(Geographical School, Nanjing Normal University, Nanjing 210097, China)

Abstract: In the fields of GIS, CAD, surveying and mapping etc., four-parameter fit and seven-parameter fit are important methods, by which we may get two sets of coordinate conversion parameters for direct and inverse solution. In this paper, the models based on four-parameter and seven-parameter transformation are discussed, inverse models are ascertained by reconstructing the matrixes models, and the solution progress and program for creating corresponding coefficients matrixes are also given.

Keywords: four-parameter; seven-parameter; fit model; direct solution; inverse solution

上接 P62

Construction and Application of Resources and Environment Data Grid

ZHU Yunqiang¹, ZHU Shaoxia², LI Shuang¹

(1 Institute of Geographic Sciences and Natural Resources Research, CAS, Beijing 100101, China;

2 College of Resources Environment and Tourism, Capital Normal University, Beijing 100037, China)

Abstract: Resources environment is the base of continual survival and development of human society. Researching resources environment can know the current status and the changing trends of the resources environment, and can sequentially explore the mechanism of resources environment changes. So there needs an open, distributed research platform to intensively study resources environment. With the platform, not only resources and environment data but also special application systems, models and software tools can be shared. According to the above requirement, this paper puts forward resources and environment data grid. The architecture of the grid mainly has the following types: the five-layered Sandglass Model, the OGSA (open grid services architecture), and the WSRF (Web Service Resource Frame). The OGSA takes services as core which can conveniently constitute uniform standard interface to manage all kinds of entirely different resources, and to accomplish sharing and cooperation. In this paper, the architecture of resources environment data grid is designed into three layers, i.e. grid resources layer, business logic layer and grid port layer, based on the OGSA. Furthermore, three key technologies are analyzed in this paper: resources environment data grid middleware, data resources gridding and the application function of resources environment data grid. Under the environment of J2EE, following the specification of the OGIS and the OWS spatial information, using the Web Services technology, the resources environment data grid can be realized. To illustrate the application of resources and environment data grid platform, a case of the sandy desertification in the surrounding areas of Beijing during 1991~2001 is studied.

Key words: resources and environment; data sharing; grid technology; data grid