

类号\_\_\_\_\_

密级\_\_\_\_\_

UDC\_\_\_\_\_

编号\_\_\_\_\_

# 中国科学院研究生院 硕士学位论文

构建虚拟太阳天文台的中国数据结点

-----  
刘健  
-----

指导教师： 林钢华 高级工程师

中国科学院国家天文台

申请学位级别： 硕士      学科专业名称： 天文技术与方法

论文提交日期： 2008年5月      论文答辩日期： 2008年5月

培养单位： 中国科学院国家天文台

学位授予单位： 中国科学院研究生院

答辩委员会主席\_\_\_\_\_

**National Astronomical Observatories  
Chinese Academy of Sciences**

**Design and Implementation of Chinese  
Data Node of Virtual Solar Observatory**

**Liu Jian**

**Directed by: Ganghua Lin**

**National Astronomical Observatories,  
Chinese Academy of Sciences**

**May 2008**

## 摘要

传统天文学经过几十年的观测积累，加上近些年新型天文观测卫星、大型地面望远镜不断投入使用，观测数据已呈指数增长，观测波段也已从可见光扩展到X射线和射电波段。这些多波段数据为天文学家提供良好的研究平台、促进交叉学科发展的同时，其劳动密集型的数据查找方式也严重影响和妨碍着天文学家进行多波段研究。

2002年底，国际上几个著名的太阳数据中心和研究机构发起了构建虚拟太阳天文台（VS0）的倡议，这一倡议很快引起了国际太阳物理领域的强烈反映并得到了大力支持。不久之后，由GONG（太阳全球振荡监测网）的首席科学家 Frank Hill 为首起草了《虚拟太阳天文台白皮书》，它明确了VS0初期的技术目标——在异地异构海量多波段数据之上构建一个应用平台，通过此平台用户能够方便地查询、获取和处理分布在各个天文台的多波段数据。经过一系列的论证准备工作VS0项目正式启动，VS0还得到了美国NSF（National Science Foundation）、NASA（National Aeronautics and Space Administration）和欧洲联盟的支持。

将怀柔的数据资源与VS0进行整合不仅能够补充VS0的数据类型，而且能够提高怀柔数据的利用率，以怀柔数据为先例，促进中国太阳数据在积累、规范和共享方面真正地与国际接轨。“分布式系统及分布式处理在太阳物理研究中的应用”得到了国家自然科学基金项目（基金项目：60673158）的资助，“构建虚拟太阳天文台的中国数据结点”是其中的一部分工作。

2006年在科技部“地球系统科学数据共享网”项目的支持下，我们构建了怀柔太阳数据查询系统。此系统采用浏览器+后台数据库的形式，使用户查询下载怀柔数据更加方便快捷，目前此系统已经投入使用。

论文首先对VS0的概念、数据组成、体系结构等做了简要介绍，接着详细阐述了构建VS0数据结点所涉及的相关技术，并重点介绍了怀柔数据结点的系统设计方案、实现方法等。最后论文从产生背景和系统方案设计等方面对怀柔太阳数据查询系统做了详细介绍。

关键字：虚拟太阳天文台，WebService，数据查询系统

## Abstract

With the multi-channel observation of the solar and “data avalanche” faced in the astronomy and the application of information technologies such as Web Services, the Virtual Solar Observatory (VSO) came into being. The initial goal of the VSO was making VSO to be an effective tool to resolve the query and download of complicated and abundant solar data. The VSO was also supported by NSF (National Science Foundation), NASA (National Aeronautics and Space Administration) and The European union.

Contacting the data sets of Huairou Solar Observing Station (HSOS) with the VSO not only can complement the data types of the VSO but also promote more and more people use the HSOS data. “Design and Implementation of Chinese Data Node of Virtual Solar Observatory” was a part of job of “The Application of Distributed System and Distributed processing in the Solar Physics Research” which was funded by “National Natural Science Foundation of China” .

In this thesis, we introduced the conception and framework of the VSO, and then we introduced the technology, framework of the VSO Data Node.

Besides that we also introduced the “Huairou Query System” in terms of the background and the frame.

Keywords: VSO, Web Services, Data Query System

---

# 目录

第一章 国际虚拟太阳天文台概述.....	- 1 -
1.1 概念的提出.....	- 1 -
1.2 数据组成.....	- 2 -
1.3 怀柔太阳观测站简介.....	- 5 -
第二章 构建虚拟太阳天文台结点的相关技术概述.....	- 9 -
2.1 Web Services 技术.....	- 9 -
2.1.1 ML.....	- 10 -
2.1.2 SOAP.....	- 11 -
2.1.3 WSDL.....	- 13 -
2.2 数据模型.....	- 14 -
2.3 虚拟太阳天文台的体系结构.....	- 16 -
第三章 怀柔太阳天文台数据结点的设计与实现.....	- 21 -
3.1 模块介绍.....	- 21 -
3.1.1 查询处理模块.....	- 21 -
3.1.2 数据下载模块.....	- 22 -
3.1.3 辅助方法.....	- 24 -
3.2 数据层的构建.....	- 24 -
3.2.1 数据命名方式.....	- 25 -
3.2.2 数据目录.....	- 26 -
3.2.3 数据库设计.....	- 29 -
3.3 资源注册构造.....	- 29 -
3.4 VSO 的使用.....	- 31 -
第四章 怀柔太阳数据查询系统的设计与实现.....	- 37 -
4.1 背景介绍.....	- 37 -
4.2 系统方案设计.....	- 37 -
4.2.1 体系结构.....	- 37 -
4.2.2 用户工作流程.....	- 39 -
4.3 系统运行实例.....	- 40 -
第五章 总结与展望.....	- 45 -
附录.....	- 49 -
数据库表结构及说明:.....	- 49 -
核心代码: 查询模块.....	- 53 -
核心代码: 下载模块.....	- 59 -
致谢.....	- 67 -



## 第一章 国际虚拟太阳天文台概述

虚拟太阳天文台 (Virtual Solar Observatory, VSO) 是一套分布式太阳档案和分析软件的联合系统,通过它能够实现对分散在世界各地天文台的太阳数据进行查询、获取和分析。由于这个系统具有提供太阳天文学信息的“天文台”功能,但其以大量太阳信息的形式存在于 Internet 上,并不是一个物理实体,所以称其为虚拟太阳天文台<sup>[1]</sup>。

### 1.1 概念的提出

目前全世界有60多个太阳天文台,有的拥有综合性仪器可获取一系列标准波长范围的每日全日面图像,有的专门获取物理参数(如速度场和磁场),有的则以较高的空间和时间分辨率去研究太阳上的某些特定现象,如活动区或日珥。然而,无论这些数据档案在局域范围里多么完整,都不能与其他类似档案直接相连。为了寻找匹配专门标准的数据,研究人员不得不访问很多单个档案站点,联系很多数据负责人。Roudier和Malherbe<sup>[2]</sup>指出,1995年仅在欧洲就有几十个可用的数据库。显然,访问这些数据库搜索有关的数据是非常耗时的任务,甚至几乎不可能完成<sup>[3]</sup>。另一方面,研究人员搜索少数几个著名天文台的观测数据时,通常得到的也只是极其简化的数据,不能满足科学研究的需要。

运行在哥达得的太阳数据分析中心是试图把太阳数据集中在一个单一位置的例子,但它的数据只局限于主要天文台和卫星。1997年Sanchez—Duarte等人<sup>[4]</sup>首先提出将太阳档案综合成一个统一系统的思想。1998年Dimitoglou等人<sup>[3]</sup>进一步讨论了这一思想。同年,意大利Capodimonte天文台的Kevin Reardon提出<sup>[5]</sup>,太阳活动周期变化的详细研究需要分析横跨多年的观测记录,这需要根据标准协议将多个单独维护的数据库无缝隙地连接起来,实现跨多个档案的复杂检索。意大利8个天文台和研究所的天文工作者也提出“全部太阳数据目录”(Whole Sun Catalog, WSC)项目<sup>[6]</sup>,即有效连接遍布世界各地的可用的已提供的单个太阳档案,简化执行检索任务和扩展研究人员的查询范围。

1998年GONG(太阳全球振荡监测网)的首席科学家Frank Hill<sup>[9]</sup>详细描述了VSO的概念。2001年7月NASA空间科学日地结合部举行了两年一次的运作使命

和数据中心回顾会议<sup>[8]</sup>，其中太阳数据分析中心(SDAC)也被包括在此次会议内。会上对 VSO 提出了研究、原型、开发方面的经费预算。在此基础上，2002 底国际上几个著名的太阳数据中心和研究机构发起了构建虚拟太阳天文台 (VSO) 的倡议，这一倡议很快引以了国际太阳物理领域的强烈反映并得到了大力支持。不久之后，由 Frank Hill 为首起草了《虚拟太阳天文台白皮书》<sup>[12]</sup>，它明确了 VSO 初期的技术目标——在异地异构海量多波段数据之上构建一个应用平台，通过此平台用户能够方便地查询、获取和处理分布在各个天文台的多波段数据。经过一系列的论证准备工作 VSO 项目正式启动，VSO 还得到了美国 NSF (National Science Foundation)、NASA (National Aeronautics and Space Administration) 和欧洲联盟的支持。由前期发展而演化为 VSO 的项目还有 JOSO (Joint Organization for Solar Observation)<sup>[10]</sup>、SOLAR (SOHO Long-term Archive)<sup>[7]</sup>、ARTHEMIS (Archive Thermal Emission Imaging System)<sup>[7]</sup>、法国基于地基太阳观测的长期太阳档案项目 BASS2000<sup>[10]</sup>等。

与此同时，数据档案及其挖掘对于太阳研究的重要性，得到了美国国家研究委员会空间研究理事委员会的认可<sup>[11]</sup>。该委员会推荐开发“一个由 NSF (National Science Foundation) 和 NASA (National Aeronautics and Space Administration) 联合支持、通过 www 访问的分布数据档案”。嗣后，美国国家研究委员会天文学天体物理调查委员会郑重签署了有关协议，把国家虚拟天文台 (NVO) 作为顶级优先启动的小型项目，NVO 是一个与 VSO 相似的项目，它显示了天文界对交叉相关可行性研究日益增长的关注。

## 1.2 数据组成

VSO得到了著名太阳数据中心的积极响应，最初就有10个单位计划提供各自拥有的数据作为VSO的一部分，这些数据源于地面和空间观测，主要来自美国(包括NSF和NASA)，当前数据总量为104.3 TB。原先主要由NSO(National Solar Observatory)、SDAC(Solar Data Archive Center)、斯坦福大学和蒙大纳州立大学提供数据服务。NSO提供的数据包括GONG(Global Oscillation Network Group)、GONG+、SOLIS(Synoptic Optical Long-term Investigations of the Sun)等；SDAC提供各种空间和补充数据；斯坦福大学提供Wilcox太阳天文台、SOHO MDI(Michelson Doppler Imager)及其他日震观测数据源；蒙大纳州立大学则提



供扩展的Yohkoh数据库。随后又有很多单位加入了提供数据服务的行列。表1.1列出了VSO 的数据组成情况:

表1.1 VSO的数据组成

数据档案	归属
KPVT <sup>1)</sup> , GONG SOLIS	THE NSO Digital Library
OSO-7 <sup>2)</sup> , SMM <sup>3)</sup> , Yohkoh CGRO BATSE <sup>4)</sup> SOHO <sup>5)</sup> , TRACE <sup>6)</sup>	NASA/GSFC SDAC <sup>7)</sup>
SOI/MDI <sup>8)</sup> , TON, Wilcox Solar Observatory	The Stanford University SSSC <sup>9)</sup>
MDI, TRACE, Yohkoh/SXT <sup>10)</sup> , LaPalma archive	Lockheed
TRACE, Yohkoh/SXT, the Mees Solar Observatory	Montana State University
Helioseismology Data Set of Dopplergrams	USC/Mt. Wilson 60-ft Fower <sup>11)</sup>
Full-disk Halpha, Ca K and White Light Images	BBDO/NJIT <sup>12)</sup>
ARTHEMIS	Italy
BASS2000	France
Coroanl, PSPT <sup>13)</sup> , ECHO <sup>14)</sup> ASP <sup>15)</sup> CHIP <sup>16)</sup>	HAO <sup>17)</sup>

- 1) KPVT 即Kitt Peak Vacuum Telescope;
- 2) OSO-7 即The Seventh Orbiting Solar Observatory;
- 3) SMM 即The Solar Maximum Mission;
- 4) CGRO BATSE 即Compton Gamma-Ray Observatory Burst and Transient Source Experiment;
- 5) SOHO 即 Solar and Heliospheric Observatory;
- 6) TRACE 即Transition Region and Coronal Explorer;
- 7) GSFC / SDAC 即Goddard Space Flight Center / Solar Data and Analysis Center;

- 8) SOI 即Saturn Orbit Insertion;
- 9) SSSC 即the Stanford SOI Science Support Center;
- 10)SXT 即Soft X-ray Telescope;
- 11)USC 即University of Southern California;
- 12)BBSO / NJIT 即Big Bear

VSO 的数据按类型<sup>[12]</sup>可以分为: intensity、equivalent\_width、polarization\_vector、LOS\_magnetic\_field、vector\_magnetic\_field、LOS\_velocity、vector\_velocity、wave\_power、number\_density、particle\_flux、particle\_velocity、thermal\_velocity、composition、wave\_phase、oscillation\_mode\_parameters 等。

表1.2 VSO的数据类型

数据类型	含义
intensity	强度
equivalent_width	等值宽度
polarization_vector	偏振矢量
LOS_magnetic_field	视向磁场
vector_magnetic_field	矢量磁场
LOS_velocity	视向速度场
vector_velocity	矢量速度场
wave_power	波强度
number_density	数密度
particle_flux	粒子流
particle_velocity	粒子速度
thermal_velocity	热速度
composition	合成
wave_phase	波相位
oscillation_mode_parameters	振动模型参数

### 1.3 怀柔太阳观测站简介

中国科学院国家天文台怀柔太阳观测基地以观测太阳磁场、速度场和太阳爆发活动等太阳物理领域中的前沿研究课题为任务，其良好的矢量磁场资料受到国际同行的高度重视。

目前怀柔太阳观测站拥有的望远镜有35cm太阳磁场望远镜、14cm H $\alpha$ 色球望远镜、10cm全日面矢量磁场望远镜、20cm全日面H $\alpha$ 望远镜等。这些望远镜与10.7cm波长2.5m口径的射电望远镜、云南全日面色球望远镜，南京大学的太阳塔等以及美国、欧洲等其它国家地理位置的太阳望远镜形成互补，形成全天24小时全时段、全年365天不间断观测。

怀柔太阳矢量磁场数据具有分辨率高、观测历史长（自1986年起就陆续获取太阳观测数据）的特点，而目前空间太阳望远镜最多只有10年寿命。研究者通过使用怀柔数据产生了大量世界先进水平的研究成果，发现了在太阳观测史上许多未曾观测到的现象，如通过光球磁场观测研究发现：太阳耀斑通常出现在色球速度反变线红移区一侧，并在此基础上提出了太阳活动区磁场和速度场的等效原理等等。目前国际上矢量磁场观测最好的是Huairou(China)、Hawaii(USA)、Huntsville(USA)，其次是Big Bear(USA)、Mitaka(Japan)。

以国家天文台怀柔太阳观测基地为主的中国太阳物理界，与美国、日本、俄罗斯、德国等国家签署了正式的中美、中日、中法、中俄太阳物理合作研究计划；多年来怀柔基地成功地与国际上著名观测站SOHO、TRACE、YOHKOH等进行联测；与日本进行地基观测资料与空间资料的交换；与美国大熊湖天文台开展了“日不落”联合观测；形成怀柔，昆明，大熊湖，奥地利的H $\alpha$ 全日面像的全球观测网，每天观测资料从怀柔和大熊湖网站上发布；形成了以太阳磁场观测为主体的良好国际联系和氛围。

VSO试图解决异地异构海量多波段太阳数据的查询、获取和处理问题，不仅Hinode卫星而且今后的地面空间望远镜数据都将通过VSO进行发布。将怀柔数据的资源与VSO进行整合不仅能够补充VSO的数据类型，而且能够提高怀柔数据的利用率，以怀柔数据为先例，促进中国太阳数据在积累、规范和共享方面真正地与国际接轨。“分布式系统及分布式处理在太阳物理研究中的应用”项目得到了国家自然科学基金项目（基金项目：60673158）的支持，“构建虚拟太阳天文台的

中国数据结点”是其中的一部分工作，它将为今后中国其他数据站点与VSO进行整合提供实例。

参考文献:

- [1] 林钢华, 虚拟太阳天文台及其发展: 《天文学进展》第24卷 第2期
- [2] Rouder Th, Malherbe J M. Obs. de Paris Publ., 1997, 45:6
- [3] Leka K D. Sol. Phys., 1994, 155:301
- [4] Sanchez-Duarte L, Fleck B, Bentley R. ASP Conf. Ser., 1997, 118:382
- [5] <http://arthemis.na.astro.it/wsc/proposal/cnaa.proposal.html>. 2002
- [6] Dimitoglou G, Mendibource C, Reardon K et al. ASP Conf. Ser., 1998, 155:297
- [7] Cora A, Volpicelli C A, Antonucci E et al. MelTl. S. A. It. , 2003, 74: 462
- [8] <http://vso.nso.edu//vsowp.pdf>, 2000
- [9] Gurman J B, In: Wilson A ed. Proc. SOHO 11 Symp., SP-508, Beitrag: ESA, 2002:66
- [10] Mendiboure C. ASP Conf. Ser., 1998, 155: 302
- [11] Hill F. ASP Conf. Ser., 1998. 155: 297
- [11] <http://vsol.nascom.nasa.gov/docs/wiki/PhysicalObservable>
- [12] Richard. S. Bogart, Frank Hill, Piet Martens. The Virtual Solar Observatory Design Proposal



## 第二章 构建虚拟太阳天文台结点的相关技术概述

### 2.1 Web Services 技术

Web Services<sup>[1]</sup> (Web 服务) 定义了应用程序如何在 Internet 上实现互操作, 它极大地拓展了应用程序的功能, 实现了软件的动态提供。在 VSO 系统中, 数据提供者主要利用 Web Services 技术实现与虚拟太阳天文台的整合; 同时其他的应用者也可以利用 VSO 暴露出的 API 接口构造实现某种功能的应用平台。

Web Services 是自包含、模块化的应用程序, 它可以在网络中被描述、发布、定位以及调用。简单地说, 它就是一个向外界暴露出的能够通过 Internet 进行调用的 API 或者应用程序。我们能够用一般的编程方法通过 Internet 来调用这些 Web Services 应用程序, 调用这些 Web Services 的应用程序被称作为服务消费者。

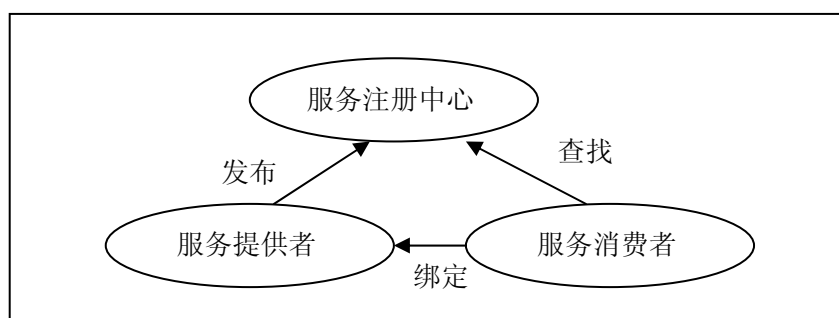


图 2.1 Web Services 的体系结构

Web Services 的体系结构一面向服务的体系结构 SOA (Service Oriented Architecture) 有三种角色, 这三种角色之间使用三种操作, 如图 2.1 所示: 服务提供者把它的服务发布到服务注册中心里, 服务消费者通过搜索服务注册中心找到满足需要的服务, 获得其相关信息之后服务消费者与服务提供者之间进行绑定使用该服务。服务消费者可以是一个人或一个程序。以下是我们在使用 Web Services 时所涉及到的技术:

- 1) XML (可扩展的标记语言): Web Services 平台中表述数据的基本格式, 它不仅解决了数据的表示问题, 而且还定义了一套标准的数据类型。
- 2) SOAP (简单对象访问协议): Web Services 进行信息交换的轻量级协议。它

提供了标准的 RPC 方法来调用 Web Services。SOAP 规范中定义了 SOAP 消息的格式，以及怎样通过 HTTP 协议来使用 SOAP，SOAP 是基于 XML 语言和 XSD 标准的，其中 XML 是 SOAP 的数据编码方式。

- 3) WSDL (WS 描述语言): W3C 提出的基于 XML 的描述 Web Services 接口的语言, WSDL 既是机器可阅读的，又是人可阅读的。

### 2.1.1 XML

XML 是 Extensible Markup Language (可扩展标记语言) 的缩写，由 W3C 发布的一种标准，是 Web 上表示结构化信息的一种标准文本格式。使用 XML 是因为它提供了与平台无关的数据描述并能够很自然地描述数据的层次关系。它以一种开放的自我描述方式定义数据结构，在描述数据内容的同时能突出对结构的描述，从而体现出数据之间的关系，这样组织出来的数据对于应用程序和用户都很友好。XML 的优势之一就是它允许各个组织、个人建立适合自己需要的标记集合，并且这些标记可以迅速地投入使用。XML 的特性如下所示：

- 1、可扩展性 XML 允许使用者创建和使用他们自己的标记。甚至特定的行业也可以一起来定义该领域的特殊标记语言，作为该领域信息共享和数据交换的基础。
- 2、灵活性 XML 提供了一种结构化的数据表示方式，使得用户界面分离于结构化数据。所以，Web 用户所追求的许多先进功能在 XML 环境下更容易实现。
- 3、自描述性 XML 文档通常包含一个文档类型声明，因而 XML 文档是自描述的。不仅人能读懂 XML 文档，计算机也能处理。XML 表示数据的方式真正做到了独立于应用系统，并且数据能够重用。XML 文档被看作是文档的数据库化和数据的文档化。

在 VSO 的服务注册中心里存储了关于数据提供者的仪器、数据类型、数据波长等信息，其中 hsos.xml 描述了怀柔太阳观测基地的数据信息，描述 20cm 全日面 H $\alpha$  数据的片断如下：

```
<xml>
.....
<dataset>
```



```

<source>HSOS</source>           //观测源
<instrument>SMAT</instrument>    //仪器
<detector>20FHalpa</detector> //望远镜
<physobs>intensity</physobs>     //观测类型
<wave>
    <wavemin>6561</wavemin>      //波长最大值
    <wavemax>6565</wavemax>     //波长最小值
    <waveunit>Angstrom</waveunit> //波长单位
    <wavetype>LINE</wavetype>   //波类型
</wave>
<time>                             //时间
    <start>19921216</start>
</time>
<datatype>image</datatype>       //数据类型
<extent>
    <type>FULLDISK</type>
</extent>
</dataset>
.....
</xml>

```

### 2.1.2 SOAP

简单对象访问协议 (SOAP) 是一种基于 XML 的消息规范, 它描述了数据类型的消息格式以及一整套串行化规则, 包括结构化类型和数组。另外, 它还描述了如何使用 HTTP 来传输消息。SOAP 作为一种有效的服务请求被发送到一些网络节点, 节点采用远程过程调用 (RPC)、组件对象模型 (COM)、等方法在任意平台上执行。

SOAP 提供了应用程序之间的交互能力, 这些应用程序可以在异构的平台上运行, 而且可以使用不同的程序语言和不同的技术来实现。

在 SOAP 机制中, 客户和服务器间的消息以 XML 编写、并用 HTTP 进行传输,

它提供了一种在无中心分布式的环境中使用 XML 交换结构化数据的通信机制,它是在松散的联合和动态的集成之间建立的无缝跨平台互操作的协议,它和 XML 是 Web Services 中最为适合的消息交换协议,提供数据访问与操作系统平台的无关性。

SOAP 主要包括四个部分:

- 1、SOAP Envelope 定义一个描述消息中的内容、发送者、接受者、处理者以及如何处理的框架。
- 2、SOAP 编码规则(encoding rule) 用于表示使用数据类型的实例。
- 3、SOAP RPC (RPC representation) 表示远程过程调用和应答的协议。

SOAP既可以用于服务器端也可以用在客户端。在VSO中Perl在客户端的接口如下:

```
use SOAP::Lite;
$soap=SOAP::Lite
->url('http://vso.stanford.edu/MDI')
->proxy('http://vso.stanford.edu/mdi.cgi');
$result=$soap->Query();
```

Perl在服务器端的接口如 :

```
Use SOAP: : Transport::HTTP;
SOAP::Transport::HTTP::CGI
->dispatch_to('MDI')
->handle;
package MDI;
sub Query{. . . }
```

SOAP 是网络胶合剂,它使得 VSO 和数据提供者通话,还能使 VSO 和其他虚拟天文台项目 EGS0、CoSEC 结合(分布服务器通过 SOAP 协议和每一个 VSO 实例通信)。其中 SOAP::Lite 模块是 API 的关键所在,它提供数据访问与操作系统平台的无关性。

### 2.1.3 WSDL

WSDL 是网络描述语言 (Web Services Description Language) 的英文缩写, 它是一个描述 Web Services 功能的协议。WSDL 描述了 Web Services 所采用的协议和格式。WSDL 文档将 Web Services 定义为服务访问点或端口的集合。Web Services 的 WSDL 文档把服务访问点和消息的抽象定义与具体的服务部署和数据格式的绑定分离开来, 因此可以对抽象定义进行重用。WSDL 文档中的消息是指对数据的抽象描述, 而端口类型是指操作的抽象集合, 端口类型使用的具体协议和数据格式规范构成了一个绑定, 将 Web 访问地址与可再次使用的绑定相关联来定义一个端口, 而端口的集合则定义为服务。

在一个 WSDL 中, Web Services 一般使用下列元素:

- 1、Type (类型): 数据类型定义器
- 2、Message (消息): 消息数据结构的抽象类型化定义
- 3、Operation (操作): 对服务中所支持的操作的抽象描述
- 4、PortType (端口类型): 某个访问入口点类型所支持的操作的抽象集合
- 5、Binding (绑定): 特定端口类型的具体协议和数据格式规范的绑定。
- 6、Port (端口): 协议/数据格式绑定与具体 Web 访问地址组合的单个服务访问点。
- 7、Service (服务): 相关服务访问点的集合, 它集成了一组相关的端口。

VSO 的 WSDL 如下所示:

```
<?xml version="1.0" ?>
+ <definitions name="VSO" xmlns="http://schemas.xmlsoap.org/wsdl/"
xmlns:soapenc="http://schemas.xmlsoap.org/soap/encoding/"
xmlns:vso="http://virtuelsolar.org/VSO/VSOi"
xmlns:wsdl="http://schemas.xmlsoap.org/wsdl/"
xmlns:soap="http://schemas.xmlsoap.org/wsdl/soap/"
targetNamespace="http://virtuelsolar.org/VSO/VSOi">
+ <types>
+ <schema xmlns="http://www.w3.org/2001/XMLSchema"
targetNamespace="http://virtuelsolar.org/VSO/VSOi">
+ </types>
+ <message name="QueryInput">
+ <message name="QueryOutput">
+ <message name="GetDataInput">
+ <message name="GetDataOutput">
+ <portType name="VSOiPort">
```

```

+ <operation name="Query">
+ <operation name="GetData">
</portType>
+ <binding name="VSOiBinding" type="vso:VSOiPort">
  <soap:binding style="rpc"
transport="http://schemas.xmlsoap.org/soap/http" />
+ <operation name="Query">
+ <operation name="GetData">
  </binding>
+ <service name="VSOiService">
  </definitions>

```

从这份 WSDL 文件中，我们得知 VSO 向外界提供了 Query 接口和 GetData 接口，以及接口输入输出参数的格式。应用者利用这些信息就可以与 VSO 进行整合，并且还可以利用 VSO 的 API 搭建自己的应用平台。

## 2.2 数据模型

VSO 的数据模型(DM)定义了一个统一的全世界范围的视图，从而解除了数据类型各异的问题。它尽可能全面、完整地定义数据查询时可能用到的查询项。它为每一个数据提供者提供描述站点数据的数据集，使它们在服务资源注册中心里保持一致的数据结构。用户查询及结果返回都以 DM 数据表示，在 VSO 内部的编程实现中，也均以 DM 作为内部数据结构。

数据提供者在数据存储时不需要遵循 DM 格式，但在接收到查询参数之后，数据提供者需要将 DM 格式的查询参数转换为符合当地数据规范的格式；同时，在将查询结果返回 VSO 之前，也需要将查询结果按照 DM 格式进行转换。DM 中的数据标识更多使用了惯例和惯例的统一，这使得数据描述和接口对于用户来说更为直观。采用 DM 不仅有利于整合现有资源，使其易于维护和升级换代，而且它还可以提供动态服务，从而适应网络和资源提供处的变化。

在 VSO 的 DM 中，统一规定用以下数据标识描述数据资源。今后 DM 会涉及更多的数据标识，从而能够更加详尽地描述数据资源。下面列出的是 VSO 计划使用或者作为查询参数已经投入使用的数据标识，其中已经投入使用的标识被标记上了星号<sup>[2]</sup>：

- Observing Time
  - Observation\_Time\*
  - Duration

- Time\_Step
- Target Location
  - Observation\_Center\_West
  - Observation\_Center\_North
  - Bounding\_Radius
- Observer Location
- Spectral Range
  - Wave\_Type
  - Wave\_Bands (*may be deleted in future versions*)
  - Wave\_Minimum\*
  - Wave\_Maximum\*
  - Wave\_Step (*may be deleted in future versions*)
- Physical Observable\*
  - intensity
  - equivalent\_width
  - polarization\_vector
  - LOS\_magnetic\_field
  - vector\_magnetic\_field
  - LOS\_velocity
  - vector\_velocity
  - wave\_power
  - wave\_phase
  - oscillation\_mode\_parameters
  - number\_density
  - particle\_flux
  - composition
  - particle\_velocity
  - thermal\_velocity
- Data Organization
  - image
  - map
  - time\_series
  - movie
  - spectrum
  - mode\_spectrum
  - spectral\_temporal
  - spatial\_spectral
- Wave Mode Sampling
  - Degree\_Minimum
  - Degree\_Maximum
  - Degree\_Step (*may be deleted in future versions*)

- Data Source

- Observatory\*
- Instrument\*
- Provider\*

下面是采用 DM 格式对欧文山谷射电天文台 OVRO 数据资源的描述：

```
<xml>
<version>0.6</version>
<provider>OVRO</provider>
<contact>Dale Gary</contact>
<uri>http://l5-m5.stanford.edu/VSO/OVROi</uri>
<proxy>http://l5-m5.stanford.edu/cgi-bin/ovro/ovroi.cgi
</proxy>
<available>1</available>
<dataset>
<source>OVRO</source>
<instrument>OVSA</instrument>
<physobs>intensity</physobs>
<wave>
<wavemin>1.2</wavemin>
<wavemax>18</wavemax>
<waveunit>GHz</waveunit>
</wave>
<time>
<start>20000316162315</start>
</time>
</dataset>
</xml>
```

### 2.3 虚拟太阳天文台的体系结构

虚拟太阳天文台有不同的体系结构，如由美国多个天文机构牵头的VSO、由

欧洲联盟提出的欧洲网格太阳天文台 (European Grid of Solar Observations, EGSO)、协作式日地连接器 (Collaborative Sun Earth Connector, COSEC)。VSO 最先提出, 而且它包括国际上大的太阳数据中心也多。其中VSO的简化体系结构<sup>(3-5)</sup> 如图2.2所示:

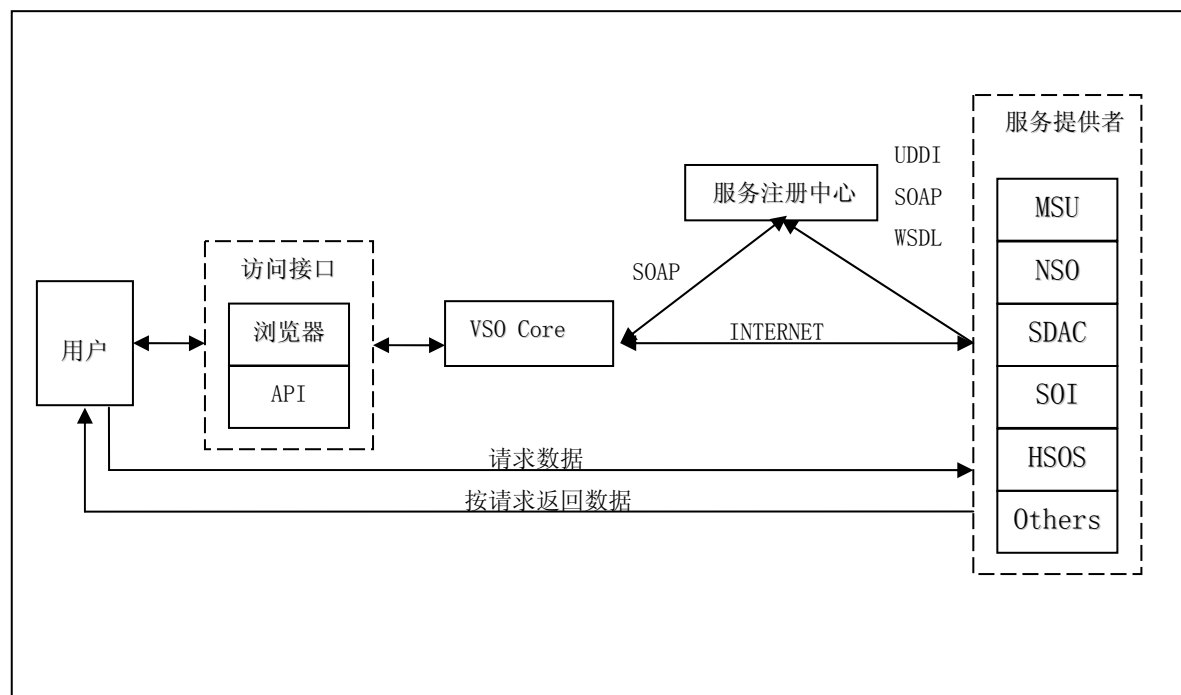


图 2.2 VSO体系结构

它主要由图中5个部分组成:

- (1) 用户: 指访问VSO的人, 如太阳物理学家, 空间物理学家等;
- (2) 访问接口: 可以是浏览器也可以是用户应用程序接口;
- (3) VSO Core: 是VSO核心, 主要用来处理用户查询, 通过查询服务注册中心的资源确定到哪一个站点去查询数据, 同时接收各个数据查询者的返回查询结果信息, 进行统一整理之后, 通过访问方法返回给用户;
- (4) 服务注册中心: 主要用来纪录数据提供者的数据资源及调用方式等信息;
- (5) 数据提供者: 它是实际执行数据查询所在地;

服务部署、发布和调用的过程为:

首先服务提供者构建Web Services (用WSDL描述) 并向服务注册中心进行服

务注册和发布；当用户需要获得某些特定的太阳数据时，需要通过访问接口将查询条件传给VS0 Core<sup>(6)</sup>中的查询引擎（Query Engine）(如图2.3所示)，查询引擎向服务注册中心进行服务查询，服务注册中心将返回符合条件的服务提供者的数据情况和相关的服务描述信息（WSDL），查询引擎根据这些信息基于SOAP直接

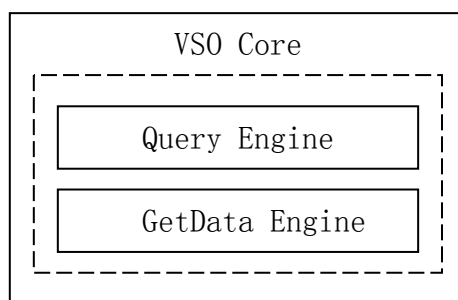


图 2.3 VS0 Core 组成

定位到相关的服务提供者结点上调用查询服务。数据提供者进行查询响应后经VS0的访问接口将查询结果返回给用户。之后用户将需要下载的数据信息传递给VS0 Engine中的数据获取引擎（GetData Engine）（如图2.3所示），数据获取引擎的作用是将用户要下载数据的相关参数传递给相应的数据提供者，数据提供者进行数据下载响应后，会将数据下载的相关信息按规定的格式返回数据获取引擎，经其做相关处理之后将数据下载信息返回给用户，由用户决定下载与否。

因为只有查询条件和查询结果通过VS0进行传递，真正的数据传输发生在VS0之外，因此大大减轻了VS0中的网络数据流通量。

VS0体系结构的成功开发增加了基础系统体系结构在处理分布式系统的扩展性和容错性方面的能力。这种扩展性体现在数据提供者方面：

- (1) 由原来的几个天文台扩展到全美国、乃至全世界，可以精细化数据模型；
- (2) 可以增加其他服务，如连接更为典型的分布式处理项目，如协作式日地连接器（CoSEC）、欧洲网格太阳天文台（EGSO）；
- (3) 现在用户方面，将单一类型用户和界面扩展到多种类型用户和界面；
- (4) 容错性能则体现在这个系统的全部体系结构受到那些不同数据仓库的影响，当增加新数据提供者或新搜索能力时，在原型开发中，开发者小心地识别并孤立出那些慢的或者无反应的数据仓库，使它们不影响VS0搜索的整体质量。对这种额外复杂性的处理能力使搜索者集中精力于搜索目标而不是使用的工具。



参考文献:

- [1] 顾宁、刘家贸、柴晓路等. 《Web Services原理与研发实践》. 第一版. 机械工业出版社, 2006. 10~13
- [2] Implementation of the Virtual Solar Observatory.  
[http://umbra.nascom.nasa.gov/vso/docs/AAS200406/vso\\_implementation.pdf](http://umbra.nascom.nasa.gov/vso/docs/AAS200406/vso_implementation.pdf) 2004
- [3] Richard S. Bogart. The Virtual Solar Observatory.  
<http://umbra.nascom.nasa.gov/vso/docs/AAS200406/vso.pdf> 2004
- [4] Joseph Gurman, George Dimitoglou, Richard Bogart. VSO 0.5: A Working Demonstration with Distributed Queries.  
[http://virtualsolar.org/docs/AGU200312/posters/query\\_construction.pdf](http://virtualsolar.org/docs/AGU200312/posters/query_construction.pdf) 2003
- [5] Gurman J, Dimitoglou G Bogart R, et . Query\_construction.  
[http://virtualsolar.org/docs/AGU200312/posters/data\\_integration.pdf](http://virtualsolar.org/docs/AGU200312/posters/data_integration.pdf) 2003
- [6] Frank Hill, Richard. S. Bogart, et. The Virtual Solar Observatory Design Proposal.  
[http://umbra.nascom.nasa.gov/vso/docs/VSO\\_strawman\\_20021125.pdf](http://umbra.nascom.nasa.gov/vso/docs/VSO_strawman_20021125.pdf) 2002
- [7] J. Hourcle. Data Transfer Negotiation Within the Virtual Solar Observatory.  
[http://lwsde.gsfc.nasa.gov/Hourcle\\_VSO\\_27Oct04.pdf](http://lwsde.gsfc.nasa.gov/Hourcle_VSO_27Oct04.pdf)



## 第三章 怀柔太阳天文台数据结点的设计与实现

各数据站点在 VSO 的体系结构（见图 2.2）中扮演数据提供者的角色，它们通过 SOAP 与 VSO 通信。数据提供者的数据资源和相关处理程序保存和运行在本地服务器中，这使得 VSO 成为没有执行瓶颈的分布式系统。数据提供者的相关程序用 Perl 语言实现。选择 Perl 的原因是 VSO 前期的已开发网络、数据库接口及词汇匹配包都用它编写<sup>(1)</sup>。

### 3.1 模块介绍

怀柔数据结点主要提供了两大模块用以响应 VSO 的调用：查询处理模块和数据下载模块；

#### 3.1.1 查询处理模块

查询处理模块的主要功能是接收从 VSO 传递下来的 QueryRequest 格式的查询参数，进行初步验证后，根据本地数据资源构造相应的 SQL 语句，执行查询后，将查询结果按照 VSOQueryResponse 格式返回 VSO。主要用到以下几个方法：

```
sub _Query {};  
sub _Search {};  
sub _PackageResults {};
```

\_Query 方法的主要功能是接收查询参数，并对参数做基本验证。检测时间格式，观测波长、数据提供者等等，确定本地数据资源能够满足查询条件。

\_Search 方法的主要功能是完成将查询参数从 QueryRequest 格式到本地数据规范的转化，并根据本地数据资源构建相应的 SQL 语句，执行查询后，将结果传递给 \_PackageResults 方法。

\_PackageResults 方法的主要功能是将 \_Search 接收的查询结果按照 VSOQueryResults 的格式附加必要的信息后将查询结果返回给 VSO。

QueryRequest 格式规定了查询参数的格式，如下所示<sup>(2)</sup>：

```
<complexType name="QueryRequest">  
<all>  
<element name="provider" type="string" minOccurs="0" />
```

```

<element name="source" type="string" minOccurs="0" />
<element name="instrument" type="string" minOccurs="0" />
<element name="physobs" type="string" minOccurs="0" />
<element name="time" type="vso:Time" minOccurs="1" />
<element name="wave" type="vso:Wave" minOccurs="0" />
<element name="extent" type="vso:Extent" minOccurs="0" />
<element name="field" type="vso:StringArray" minOccurs="0" />
</all>
</complexType>

```

VSOQueryResponse 格式规定了查询结果的返回格式，其格式如下所示<sup>(2)</sup>：

```

<complexType name="VSOQueryResponse">
<all>
<element name="version" type="float" />
<element name="no_of_records_found" type="int" minOccurs="0" />
<element name="no_of_records_returned" type="int" minOccurs="0" />
<element name="error" type="string" minOccurs="0" />
<element name="debug" type="string" minOccurs="0" />
<element name="status" type="string" minOccurs="0" />
<element name="provider" type="string" />
<element name="source" type="string" minOccurs="0" />
<element name="instrument" type="string" minOccurs="0" />
<element name="physobs" type="string" minOccurs="0" />
<element name="time" type="vso:Time" />
<element name="wave" type="vso:Wave" minOccurs="0" />
<element name="extent" type="vso:Extent" minOccurs="0" />
<element name="fileid" type="string" minOccurs="0" />
<element name="size" type="float" minOccurs="0" />
<element name="extra" type="vso:Extra" minOccurs="0" />
<element name="info" type="string" minOccurs="0" />
</all>
</complexType>

```

### 3.1.2 数据下载模块

数据下载模块的主要功能是接收从 VSO 传递下来的 VSOGetDataRequest 格式的查询参数，经过初步检测之后，构造相应的 SQL 语句，执行本地查询之后，将查询结果按照 VSOGetDataResponse 格式返回 VSO。主要用到以下方法：

```

sub _GetData { };
sub _ProcessDataRequest { };
sub _PackageData { };

```

\_GetData 方法的主要功能是接收来自 VS0 的数据下载请求，对符合 VS0GetDataRequest 格式的参数做简单的检测：用户是否请求了一个可用的下载方法，以及本地数据资源能否满足用户的数据下载请求。

\_ProcessDataRequest 方法的主要功能是处理来自 VS0 的数据下载请求，经过分析构造相应的 SQL 语句，执行 SQL 查询语句后，将结果输入到 \_PackageData 方法中。

\_PackageData 方法的主要功能是将结果按照 GetDataResponse 的格式返回 VS0。

VS0GetDataRequest 格式规定了 VS0 调用数据下载功能时所必需提供给数据提供者的信息，格式如下所示<sup>(2)</sup>：

```
<complexType name="VS0GetDataRequest">
<all>
  <element name="version" type="float" minOccurs="0" />
  <element name="method" type="vso:StringArray" />
  <element name="email" type="string" />
  <element name="host" type="string" minOccurs="0" />
  <element name="user" type="string" minOccurs="0" />
  <element name="directory" type="string" minOccurs="0" />
  <element name="password" type="string" minOccurs="0" />
  <element name="address" type="string" minOccurs="0" />
  <element name="required" type="string" minOccurs="0" />
  <element name="provider" type="string" />
  <element name="fileid" type="vso:StringArray" />
</all>
</complexType>
```

VS0GetDataResponse 格式规定了数据提供者对 VS0 的数据下载请求响应后，将数据下载的相关信息返回时所遵循的格式，如下所示<sup>(2)</sup>：

```
<complexType name="ProviderGetDataResponse">
<all>
  <element name="version" type="float" />
  <element name="info" type="vso:StringArray" minOccurs="0" />
  <element name="provider" type="string" minOccurs="0" />
  <element name="fileid" type="vso:StringArray" minOccurs="0" />
  <element name="url" type="string" minOccurs="0" />
  <element name="provider" type="string" minOccurs="0" />
  <element name="status" type="string" minOccurs="0" />
  <element name="details" type="string" minOccurs="0" />
</all>
```

```

<element name="content" type="string" minOccurs="0" />
<element name="debug" type="string" minOccurs="0" />
<element name="method" type="vso:StringArray" minOccurs="0" />
</all>
</complexType>

```

### 3.1.3 辅助方法

在构建 VSO 数据结点的过程中，还有以下几个必要的辅助方法：

#### ○ Ping

功能：VSO 用来检测某个数据结点网络是否畅通。

#### ○ \_GetDatabaseInfo

功能：返回本地数据库的信息：数据库名，用户，密码等。

#### ○ \_ValidData

功能：简单检测输入的时间参数是否符合规定的格式：14 位数=年[4] +月[2]+日[2]+时[2]+分[2]+秒[2]。

#### ○ \_WaveConversion

功能：对于一个给定的 wave 格式，将其数据值均转换为以埃 (Å) 为单位的值。这将减少 VSO 内部做相同的转换。

#### ○ \_ProviderID

功能：返回数据提供者在 VSO 中的 ID。

#### ○ \_ThrowError

功能：将错误信息按固定的格式返回。

## 3.2 数据层的构建

望远镜获取原始数据后自动将数据传送到采集计算机上，然后采集计算机通过 ftp 将数据和 tab 表文件上传到应用服务器中，之后应用服务器中的后台处理程序会定时自动地将数据信息归档到本地数据库中。数据归档物理结构如图 3.1 所示。

目前怀柔太阳观测基地上传的数据类型有：由 35cm 望远镜观测到的局部矢量磁场和多普勒速度场数据、由 20cm 望远镜观测到的全日面 H $\alpha$  数据和由 10cm 望远镜观测到的全日面矢量磁场数据。上传到应用服务器的资源包括数据和 tab 表文件，其中 tab 表文件是 35cm 望远镜的观测日志，用以纪录数据

的活动区、文件名、时间、视宁度和质量等信息。

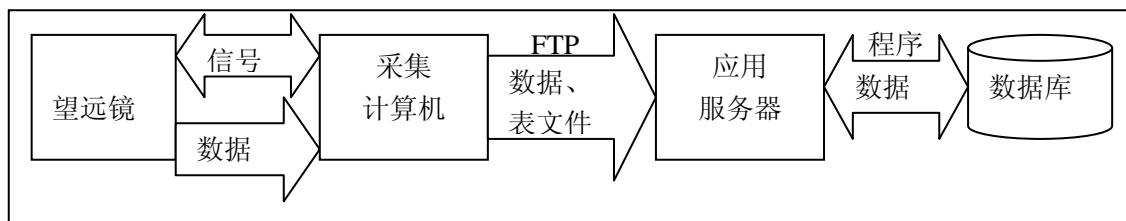


图 3.1 数据归档物理结构

### 3.2.1 数据命名方式

怀柔太阳观测基地的数据文件名都依一定的意义给定<sup>(3)</sup>：

- 局部矢量磁场和速度场数据文件的命名格式均为：**X1X2X3X4X5X6X7.fit**。

各字符位代表含义如下所示：

X1=代表观测类型[1位]；

s：单色像

l：Stokes 的 L 分量

q：Stokes 的 L 分量

u：Stokes 的 L 分量

v：速度场

X2=观测位置[1位]；

5 代表：在观测谱线 FeI 5324.19Å 处

4 代表：在观测谱线 H 4861.34Å 处

X3=活动区[5位]；X4=日[2位]；X5=时[2位]；X6=分[2位]；X7=秒[2位]；

例如：L50703402023322.fit 说明这是一个属于 07034 活动区、02 日 02:33:22（世界时）在谱线 FeI 5324.19Å 处观测到的 Stokes L 分量的数据。

- 全日面 H $\alpha$  的数据文件命名格式为 **X1X2X3X4X5X6X7.fit**。各字符位代表含义如下所示：

X1=代表观测类型[4位]；

X2=年[2位]；X3=月[2位]；X4=日[2位]；X5=时[2位]；X7=秒

[2 位];

X8=待定[4 位];

例如: HrHa080417013804P00A. fit 说明这是在 2008 年 4 月 17 日 01:38:04 (世界时) 观测到的 20cm 全日面 Ha1pha 数据。

- 全日面矢量磁场的文件命名格式为 X1X2X3X4X5X6X7.fit。各字符位代表含义如下所示:

X1=代表观测类型[2 位];

X2=年[2 位]; X3=月[2 位]; X4=日[2 位]; X5=时[2 位]; X7=秒[2 位];

X8=待定[4 位];

x9=Stokes 参数[1 位]

X10=观测谱线位置[1 位]

例如: Fm070501020526P00AU5. fit 说明这是 2007 年 05 月 01 日 02:05:26 (世界时) 在谱线 FeI 5324.19Å 处观测到的全日面矢量磁场的 Stokes U 分量的数据。

### 3.2.2 数据目录

怀柔数据资源和 tab 表文件是按照一定的目录格式存储在应用服务器中, 其存储目录分别如图 3.2 和图 3.3 所示:



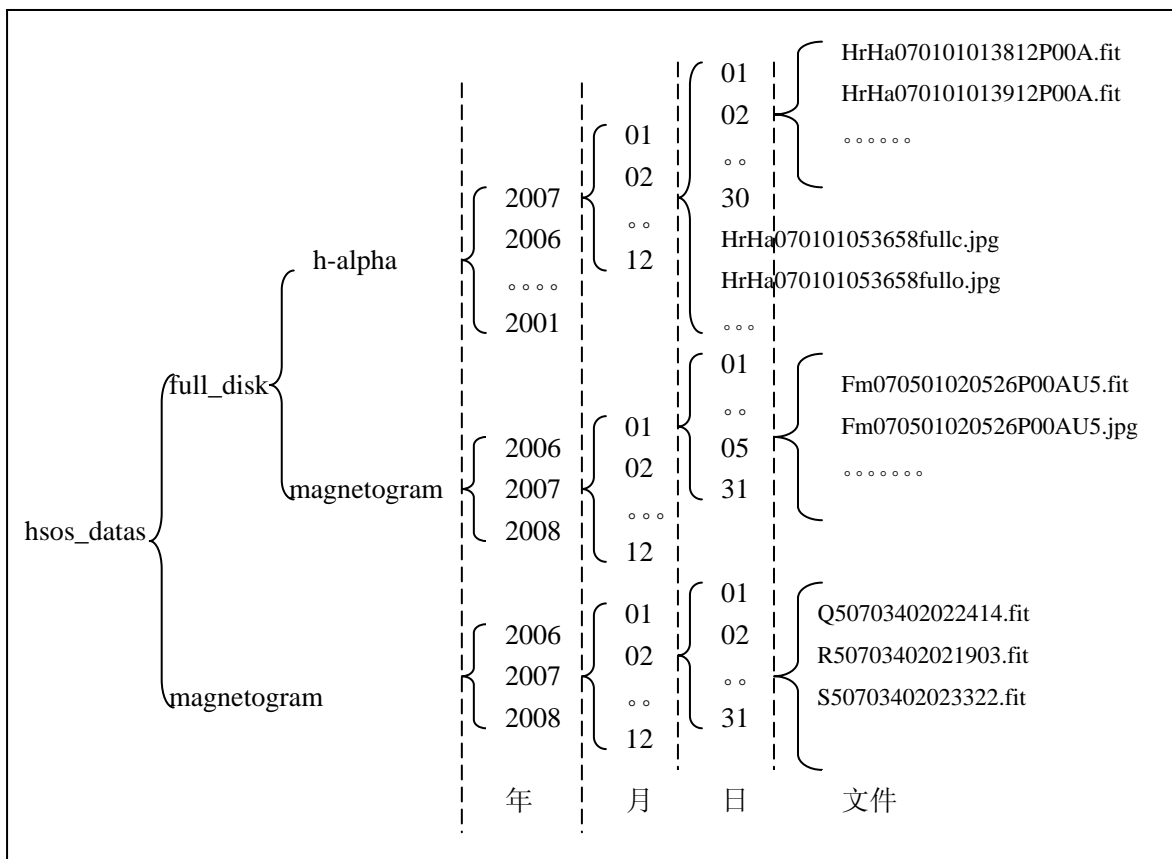


图 3.2 怀柔数据的目录结构

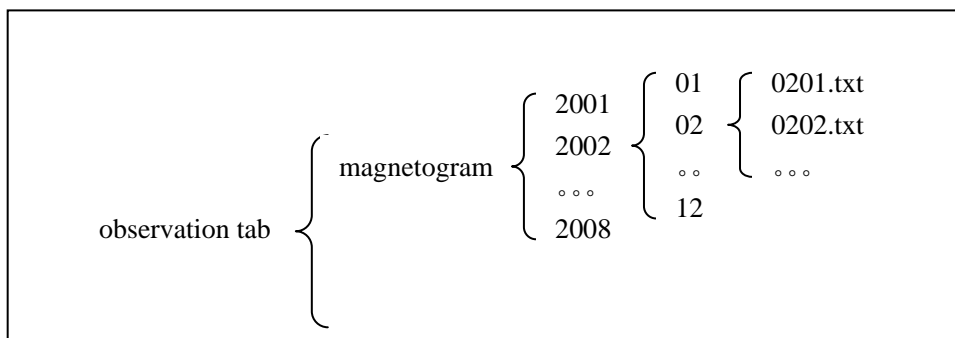


图 3.3 怀柔 tab 表文件的目录结构

在如何得到相关的数据信息这一问题上，我们采用两种算法：

因为数据的文件名和存储位置代表了一定的信息：路径、时间、数据类型等，所以对于提取这些信息，我们采用如下程序伪算法——A 算法：

- 1、扫描 hsos\_datas 目录，找到一个文件。
- 2、根据文件的命名和路径将文件进行归类，确定要更新的表名。

- 3、在文件名中提取信息，结合文件的存储路径构造相应的 SQL 语句。执行 SQL 语句更新相应的数据表
- 4、继续寻找目录中的下一个文件。

对于全日面矢量磁场和 20cm 全日面 H $\alpha$  数据而言，采用 A 算法可以得到足够的描述数据的信息。但是对于局部矢量磁场和多普勒速度场来说，因为其自身的文件名不足以提供足够的信息（如文件发生的年、月），所以我们还必须提取 tab 表文件中日志信息。tab 表文件是 35cm 望远镜的观测日志，纪录数据的活动区、文件名、时间、视宁度和质量等信息。其内容如图 3.4 所示：

```

-----
----- HUAIROU SOLAR OBSERVING STATION -----
-----
OBJECT HSOS.NO: 02086
  TIME POSITION(record time UT): (UT=Local Time of the Station minus 8 Hours)
      Start: 2002-04-08 01:44:09   End: 2002-04-08 04:37:19
  (Lon, Lat): Longitude=55.10, Latitude=1.70
  (B , L) : B=1.74,L=295.50
  SEEING : D
  CLIMATE : overcast
  TEMPERATURE OF THE FILTER: 42.240
  OBSERVER: Bao Kingming
-----

```

CONTENT	TIME POSITION	QUALITY	CALIB	FRAME	GAIN	LEVEL	FILENAME
D4 (H-Beta filtergram. Wavelength=4861.34)	2002-4-8 10:39:30	D	199780.0	256	7.0	41	D40208608023911.fit
D5 (Photospheric filtergram. Wavelength=5324.19)	2002-4-8 10:31:30	D	245340.0	256	4.0	25	D50208608023108.fit
L4 (H-Beta longitudinal magnetogram. Wavelength=4861.34-0.24)	2002-4-8 10:37:28	D	1909770.0	256	6.0	49	L40208608023650.fit
L5 (Photospheric longitudinal magnetogram. Wavelength=5324.19-0.075)	2002-4-8 10:24:13	D	214810.0	256	5.0	15	L50208608022353.fit
	2002-4-8 11:05:15	D	362990.0	256	3.0	27	L50208608030450.fit
	2002-4-8 11:28:22	D	496970.0	256	3.5	45	L50208608032807.fit
	2002-4-8 12:35:15	D	537570.0	256	3.5	47	L50208608043459.fit
Q5 (Photospheric Stokes Q. Wavelength=5324.19)	2002-4-8 10:22:20	D	1136820.0	256	5.0	45	Q50208608022206.fit

图 3.4 tab 文件信息

用于提取 tab 表文件中的日志信息的程序伪算法——B 算法：

- 1、扫描 tab 目录，找到一个文件。
- 2、打开\*.txt 文件，得到如活动区，视宁度、质量、文件名、文件时间等信息。
- 3、构造相应的 SQL 语句，执行 SQL 语句，更新相应的数据表。
- 4、关闭文件，读取下一个文件。

### 3.2.3 数据库设计

怀柔太阳观测基地的数据库中一共有 17 个表,关于数据库表的说明如下所示。具体每个表的结构说明见附录之数据库表结构及说明。

表 3.1 数据库中表的信息

表名	说明
file_in_table	纪录tab表文件中的信息
hsos_no_in_table	活动区表
full_mag	全日面矢量磁场表
full_mag_picture	全日面矢量磁场图表
ha_20	20cm全日面Halpha表
ha_20_c	20cm全日面Halpha的表文件表
mag_file_path	局部矢量磁场和速度场表
mag_picture	局部矢量磁场和速度场图像表
ip_address	ip转化表
user	用户表
user_log	下载数据纪录表
view_user_log	浏览数据纪录表

### 3.3 资源注册构造

在 VSO 的服务注册中心里,包含了所有数据提供者的数据资源信息及其他相关信息。通过 2.3 节中关于 VSO 体系结构的介绍,可以知道 VSO 系统通过查询服务注册中心的信息从而筛选出满足查询条件的数据提供者。每个数据提供者利用 VSO 的 DM 描述各自的数据资源<sup>(4)</sup>。

hsos.xml 是利用 DM 对怀柔太阳数据资源进行描述的文件,将 hsos.xml 及接口信息提交给 VSO,系统将相关信息存储到服务注册中心中完成资源注册。这样当用户通过 VSO 发出查询请求时候,系统就可以据此判断怀柔数据站点能否提供相应的数据。服务注册中心里怀柔太阳数据服务的信息如下所示:

```
<xml>
<version>1.5</version>
<provider>NAOC</provider>
```

```
<contact></contact>
<uri>http://virtualsolar.org/VSO/NAOC</uri>
<proxy>http://...</proxy>
<available>1</available>
<dataset>
  <source>HSOS</source>
  <instrument>SMFT</instrument>
  <detector>35LVM</detector>
  <physobs>intensity</physobs>
  <physobs>LOS_magnetic_field</physobs>
  <physobs>vector_magnetic_field</physobs>
  <physobs>LOS_velocity</physobs>
  <wave>
    <wavemin>5323</wavemin>
    <wavemax>5325</wavemax>
    <waveunit>Angstrom</waveunit>
    <wavetype>LINE</wavetype>
  </wave>
  <time>
    <start>19860612</start>
  </time>
  <datatype>image</datatype>
  <extent>
    <type>PARTIAL_SUN</type>
  </extent>
</dataset>

<dataset>
  <source>HSOS</source>
  <instrument>SMAT</instrument>
  <detector>10FVM</detector>
  <physobs>vector_magnetic_field</physobs>
  <wave>
    <wavemin>5323</wavemin>
    <wavemax>5325</wavemax>
    <waveunit>Angstrom</waveunit>
    <wavetype>LINE</wavetype>
  </wave>
  <time>
    <start>19860612</start>
  </time>
  <datatype>image</datatype>
  <extent>
```

```
        <type>FULLDISK</type>
    </extent>
</dataset>
<dataset>
    <source>HSOS</source>
    <instrument>SMAT</instrument>
    <detector>20FHalpHa</detector>
    <physobs>intensity</physobs>
    <wave>
        <wavemin>6561</wavemin>
        <wavemax>6565</wavemax>
        <waveunit>Angstrom</waveunit>
        <wavetype>LINE</wavetype>
    </wave>
    <time>
        <start>19921216</start>
    </time>
    <datatype>image</datatype>
    <extent>
        <type>FULLDISK</type>
    </extent>
</dataset>
</xml>
```

在hsos.xml文件中，每个dataset元素都从观测类型、观测时间、观测波段等方面对数据资源进行了描述，同时文件中也给出了调用接口信息。VSO根据hsos.xml判断怀柔的数据资源是否符合查询要求，如果符合，那么将通过hsos.xml获取怀柔的接口信息，然后将接口信息传递给VSO Core中的查询引擎，使其能够直接调用怀柔数据站点的API。

### 3.4 VSO 的使用

图 3.5 是 VSO 为用户提供查询的起始界面，通过这个界面用户可以对查询参数进行自由组合制订个性化的查询界面，这些查询参数是：时间、观测类型、仪器/观测站/数据提供者，波长范围，别名等，其中时间参数是必选项。提交之后，系统就会返回相应的查询界面。

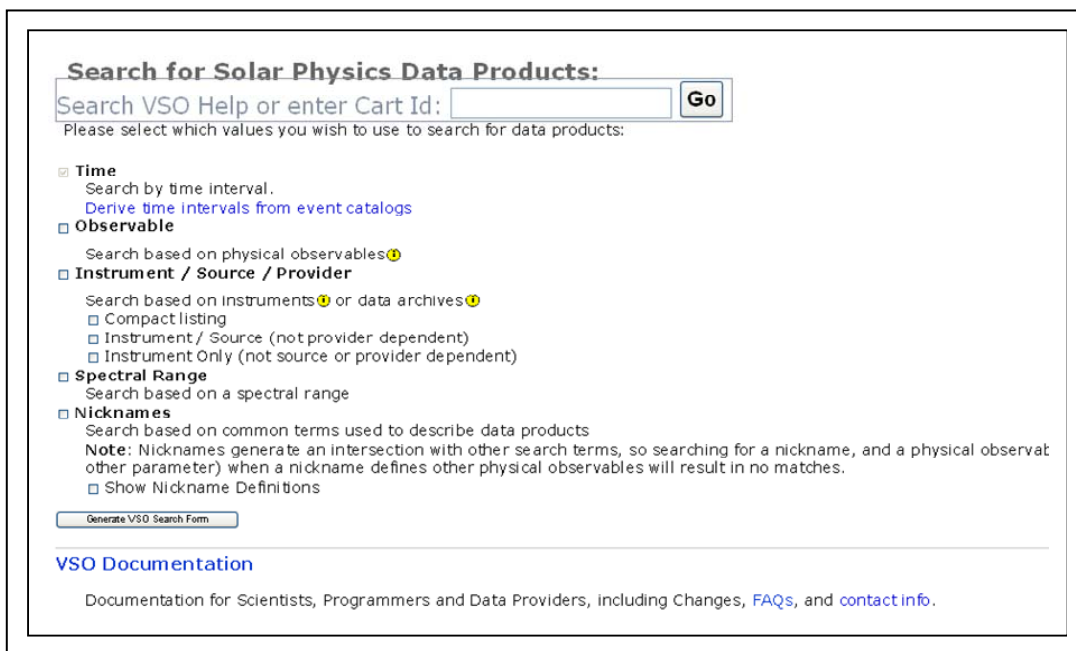


图 3.5 VSO 的起始查询界面

例如，我们以时间 time 和仪器 / 观测站 / 数据提供者 Instrument/Source/Provider 这两个参数制订查询界面，提交之后 VSO 会返给我们如图 3.6 所示的查询界面：

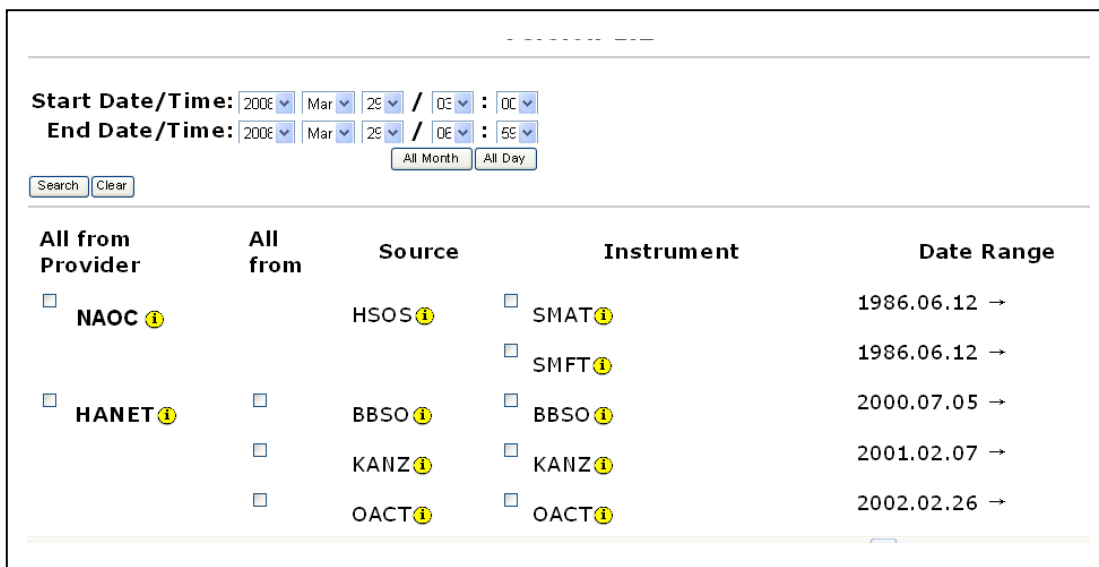


图 3.6 VSO 的查询界面

在图 3.6 中输入查询条件，以查询 2007-03-27 00:00 至 2007-03-27 23:00 来自怀柔 (HSOS) 的 SMFT 望远镜和 HANCET 的 KANZ 望远镜的数据为例，经过 VSO 的后台执行——查询引擎通过查询服务注册中心得到适合的数据提供者的

信息资源、调用其相应的应用程序，以及各数据提供者进行执行响应并将结果返回——之后，可以得到图 3.7，图中展示了查询结果。

<input type="checkbox"/> Thumbnail	<input type="checkbox"/> Time Start	<input type="checkbox"/> Time End	<input type="checkbox"/> Min Spectral Range	<input type="checkbox"/> Max Spectral Range	<input type="checkbox"/> Wave Type	<input type="checkbox"/> Observable	<input type="checkbox"/> Source	<input type="checkbox"/> Instrument	<input type="checkbox"/> Extent
<input type="checkbox"/>	2007.03.27 07:01:36	2007.03.27 07:01:36	5323 Å	5325 Å	LINE	N/A	HSOS	SMFT	PARTIAL_SUN
<input type="checkbox"/>	2007.03.27 07:01:36	2007.03.27 07:01:36	5323 Å	5325 Å	LINE	N/A	HSOS	SMFT	PARTIAL_SUN
<input type="checkbox"/>	2007.03.27 07:02:58	2007.03.27 07:02:58	5323 Å	5325 Å	LINE	N/A	HSOS	SMFT	PARTIAL_SUN
<input type="checkbox"/>	2007.03.27 07:02:58	2007.03.27 07:02:58	5323 Å	5325 Å	LINE	N/A	HSOS	SMFT	PARTIAL_SUN
<input type="checkbox"/>	2007.03.27 07:03:30	2007.03.27 07:03:30	5323 Å	5325 Å	LINE	N/A	HSOS	SMFT	PARTIAL_SUN
<input type="checkbox"/>	2007.03.27 07:03:30	2007.03.27 07:03:30	5323 Å	5325 Å	LINE	N/A	HSOS	SMFT	PARTIAL_SUN
<input type="checkbox"/> <a href="#">IMG Link</a>	2007.03.27 07:09:09	2007.03.27 07:09:09	6558 Å	6568 Å	line	intensity	KANZ	KANZ	N/A
<input type="checkbox"/>	2007.03.27 07:14:56	2007.03.27 07:14:56	5323 Å	5325 Å	LINE	N/A	HSOS	SMFT	PARTIAL_SUN
<input type="checkbox"/>	2007.03.27 07:14:56	2007.03.27 07:14:56	5323 Å	5325 Å	LINE	N/A	HSOS	SMFT	PARTIAL_SUN
<input type="checkbox"/>	2007.03.27 07:15:51	2007.03.27 07:15:51	5323 Å	5325 Å	LINE	N/A	HSOS	SMFT	PARTIAL_SUN
<input type="checkbox"/>	2007.03.27 07:15:51	2007.03.27 07:15:51	5323 Å	5325 Å	LINE	N/A	HSOS	SMFT	PARTIAL_SUN

图 3.7 VSO 查询结果显示界面

在 Thumbnail 中，我们可以通过点击 IMG Link 来预览数据生成的图像。用户在图 3.7 中选择需要下载的数据，进行提交之后。在图 3.8 中 VSO 将显示数据提供者所提供的下载方法供用户进行选择。

Sessions : 25-Apr-2008 07:03:04 UTC

Provider	Select Transfer Method
NAOC <sup>?</sup>	<input checked="" type="radio"/> URL-FILE <sup>?</sup>
HANET <sup>?</sup>	<input checked="" type="radio"/> URL-FILE <sup>?</sup>

图 3.8 VSO 数据下载方法选择

提交之后，VSO 将调用相应数据提供者的\_getData 方法，等 VSO 接收到数据提供者的响应结果之后将下载信息返回给用户，如图 3.9 所示。根据这些信息用户就可以直接到数据提供者那里下载数据了。

<b>CART ID: VSO-SDAC_TEST-080425-071</b>			
<b>Request Status</b>			
Session 1 : 25-Apr-2008 07:03:04 UTC			
Provider	Time	State	Comments
	25-Apr-2008		
CAS	07:44:18 UTC	DONE	<a href="http://sun.bao.ac.cn/hsos_datas/full_disk/magnetogram/2006/05/15/Fm060515140134P00A">http://sun.bao.ac.cn/hsos_datas/full_disk/magnetogram/2006/05/15/Fm060515140134P00A</a>
	25-Apr-2008		
HANET	07:44:19 UTC	DONE	<a href="ftp://ftp.bbso.njit.edu/pub/archive/2007/03/28/kanz_half_fr_20070328_074246.fts.gz">ftp://ftp.bbso.njit.edu/pub/archive/2007/03/28/kanz_half_fr_20070328_074246.fts.gz</a> <a href="ftp://ftp.bbso.njit.edu/pub/archive/2007/03/29/kanz_half_fr_20070329_071111.fts.gz">ftp://ftp.bbso.njit.edu/pub/archive/2007/03/29/kanz_half_fr_20070329_071111.fts.gz</a>

图 3.9 VSO 的数据下载界面

由于只有查询条件和查询结果在 VSO 中传递，真正的数据传输发生在 VSO 之外，因此大大减轻了 VSO 的网络负担。



参考文献:

- [1] 林钢华. 虚拟天文台技术与展望. 计算机工程, 2006 年 4 月第 8 期:6~8
- [2] [http://vso.nascom.nasa.gov/API/sdacVSOi\\_strict.wsd1](http://vso.nascom.nasa.gov/API/sdacVSOi_strict.wsd1)
- [3] <http://sun.bao.ac.cn/observation/readme>
- [4] Richard. S. Bogart, Frank Hill, Piet Martens.  
[http://virtualsolar.org/docs/VSO\\_strawman\\_20021125.pdf](http://virtualsolar.org/docs/VSO_strawman_20021125.pdf) 2002



## 第四章 怀柔太阳数据查询系统的设计与实现

### 4.1 背景介绍

自 1986 年起, 国家天文台怀柔太阳观测基地就陆续获取太阳观测数据, 观测类型分为局部矢量磁场、多普勒速度场、全日面 Halpha 和全日面矢量磁场等。就数据的时间积累和观测类型而言, 怀柔太阳观测基地在国际太阳领域中占有重要位置。但是在数据管理方面, 我们仍沿用通过文件目录进行管理的形式, 当用户查找某些特定的数据时需要花费很大的精力甚至需要在基地相关观测人员的帮助下才能够得到需要的数据, 随着怀柔基地观测的积累及新仪器的投入使用, 观测的数据量逐步增大、数据类型也不断增多, 从而使通过采用文件目录进行管理所带来的局限性显得更为突出。

2006 年在科技部“地球系统科学数据共享网”项目的支持下, 我们对怀柔的数据管理系统进行了改造, 构建了怀柔太阳数据查询系统。此系统采用浏览器+后台数据库的方式, 用户通过查询页面将查询条件提交给后台服务器, 后台服务器立即对用户的请求进行响应, 同时系统还将数据查询参数扩展为: 时间、数据类型和活动区, 从而极大地方便了用户查找和获得特定的数据。

### 4.2 系统方案设计

怀柔数据查询系统是出于方便用户检索下载相关数据而设计的。它是建立在 Linux 操作系统、Apache 服务器、MySQL 数据库管理系统的平台之上、利用 PHP 开发的一个基于浏览器/服务器模式的原型系统。它通过改进工作流程, 在后台自动实现了一系列的数据归档工作, 而无须人工干预。同时, 它还为用户提供了一个可以通过网络访问系统的图形接口, 用户可以通过此接口在线检索和下载数据, 以及查看数据的使用情况、删除不合格数据等(仅限管理员)。下面将详细介绍此系统的方案设计。

#### 4.2.1 体系结构

本系统以可视化的形式响应用户对服务器请求的每一个操作。它按照功能模块可分为界面模块、数据查询模块、数据下载模块、后台管理模块。系统的

总体结构图如图 4.1 所示。

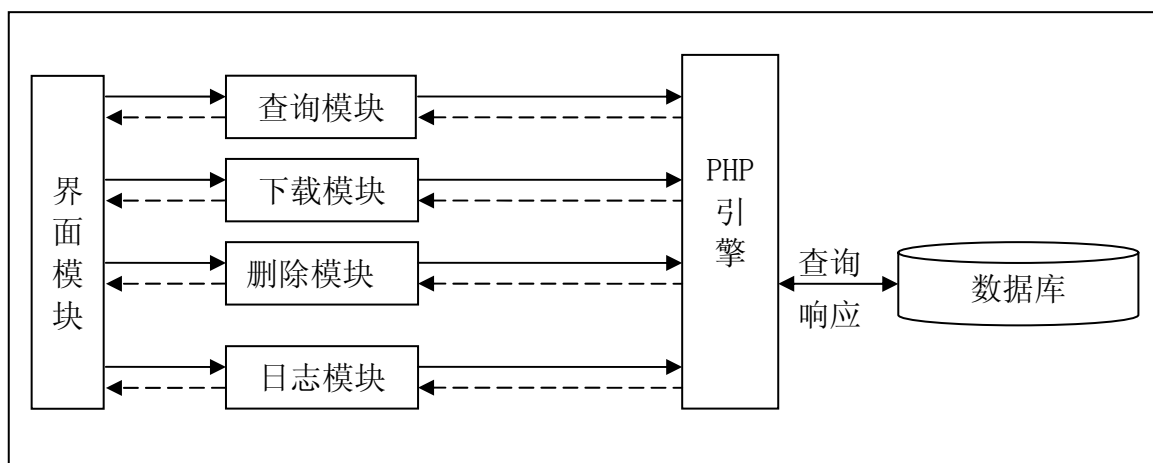


图 4.1 数据查询系统总体结构图

本系统为用户提供以下功能：

- 1、将查询条件扩展为时间、数据类型、活动区等。用户通过在浏览器中输入查询条件向服务器提交查询；
- 2、用户可以浏览、下载查询结果；
- 3、用户（仅限管理员使用）可以查询数据的下载、浏览情况以及删除不合格数据。

以下是各模块的功能介绍：

- 1、界面模块：负责接收用户的调用信息，传递给相应的模块，并将模块响应结果呈现给用户
- 2、查询模块：主要接收传入的查询条件，实现复杂查询语句的构造及执行，将查询结果返回界面模块；
- 3、下载模块：主要接收需要下载的数据的参数，实现复杂查询语句的构造及执行后，将相关数据做打包处理，并将 URL 返回界面模块。；
- 4、删除模块：通过传入的文件参数，进行身份判断和删除确认后，删除相关数据；
- 5、日志模块：主要纪录使用系统的用户 IP 情况、浏览数据量和下载数据量；

查询模块的工作流程如下图 4.2 所示：

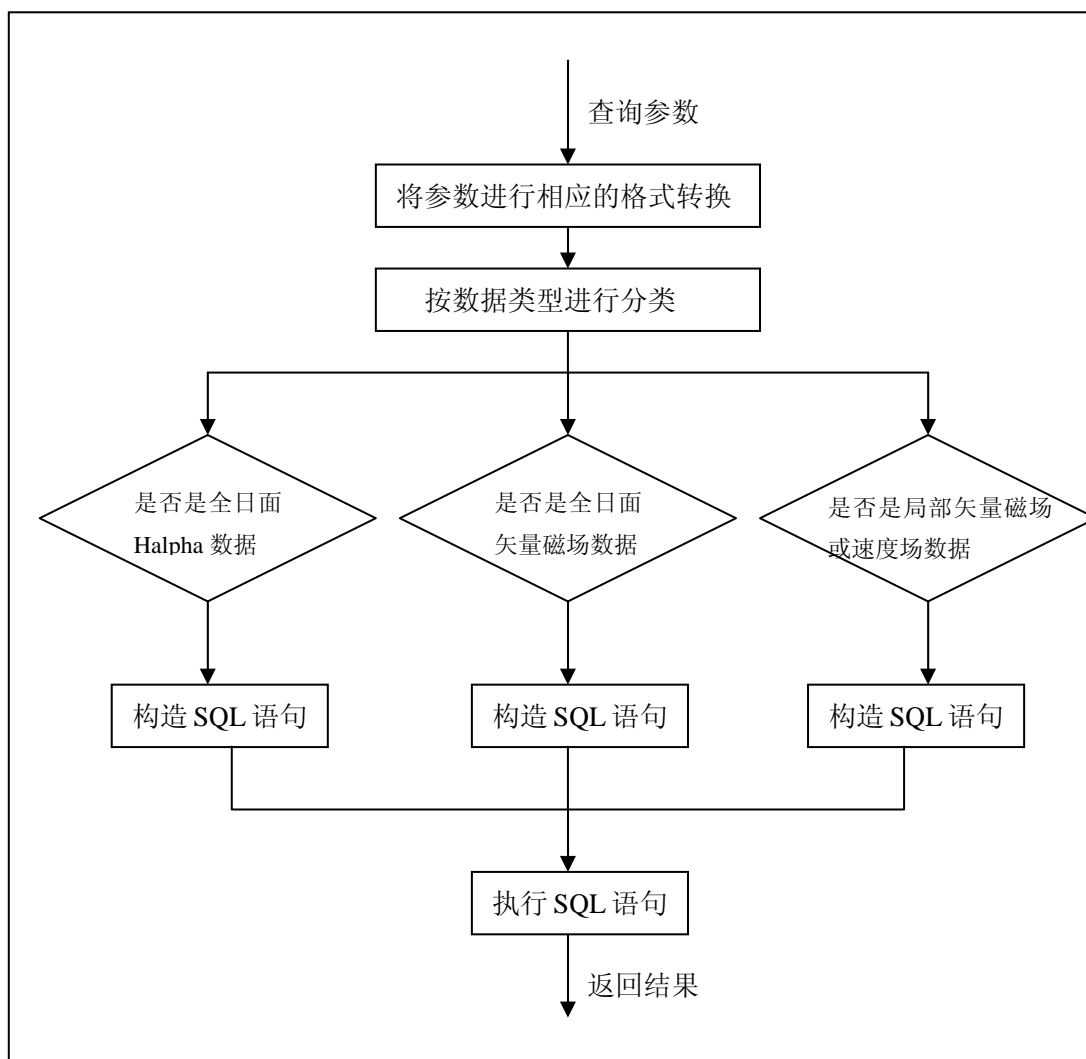


图 4.2 查询模块工作流程图

关于查询模块和下载模块的核心代码见附录。

#### 4.2.2 用户工作流程

用户工作流程如图 4.3 所示。首先用户进行身份验证，主要是为了验证是管理者还是一般的使用者，他们分别拥有不同的权限。如果用户登录成功，则进入相应的主界面，否则重新登录。普通用户的主界面提供了在线查询和下载功能。管理者可以通过相应的界面查看数据的浏览和下载情况，以及修改密码等。

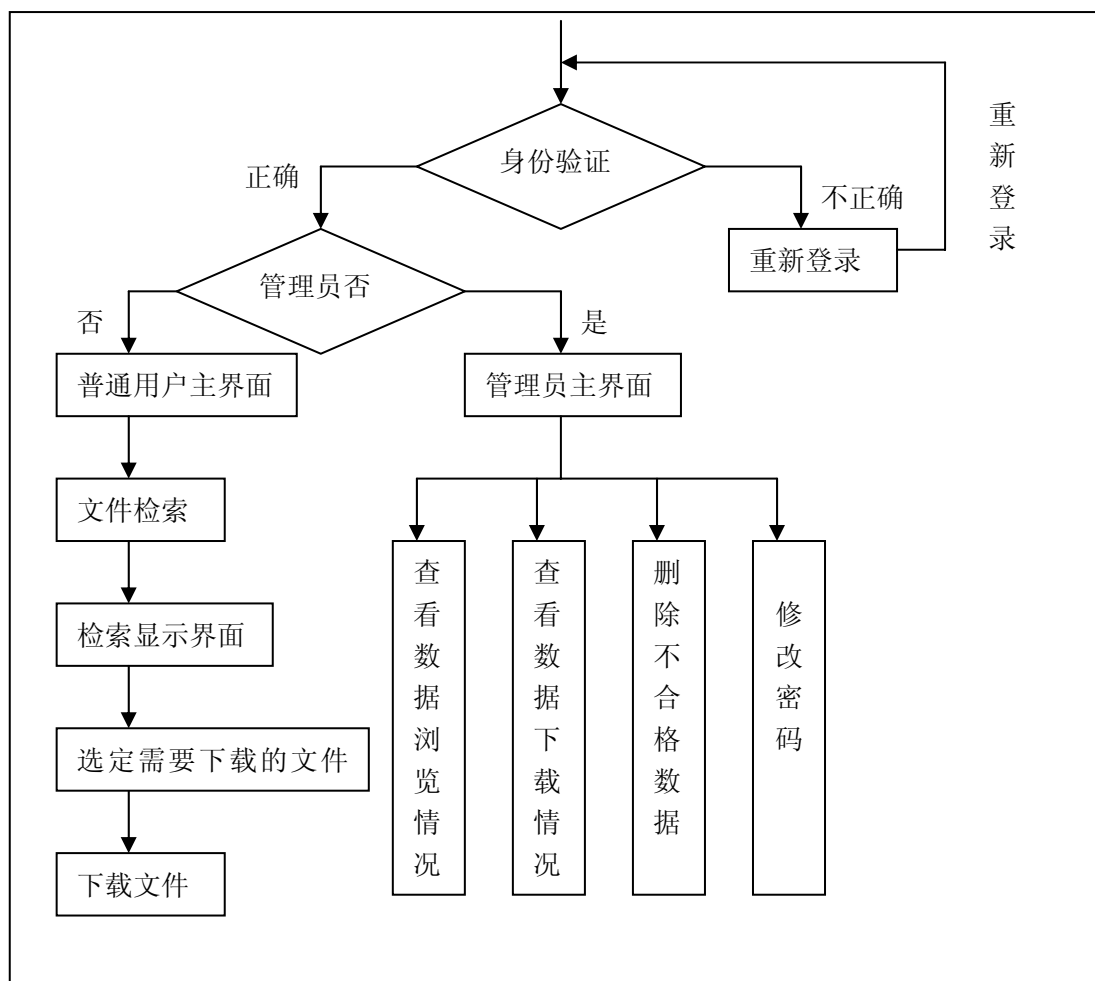
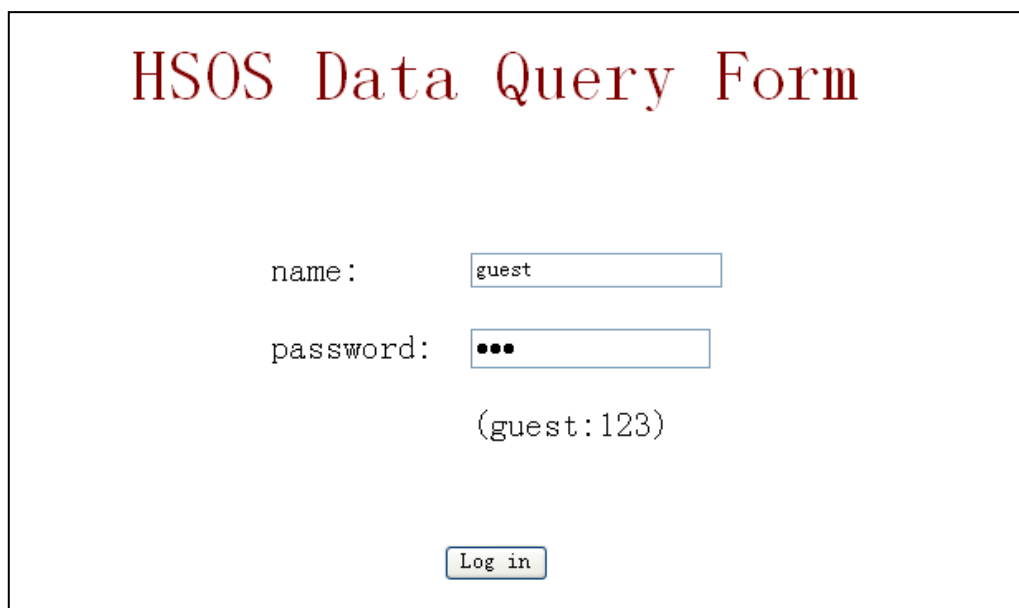


图 4.3 用户工作流程图

### 4.3 系统运行实例

以普通用户为例，展示系统的运行情况：打开浏览器，输入服务器的 IP 地址，首先进入到用户登录界面,见图 4.4:



**HSOS Data Query Form**

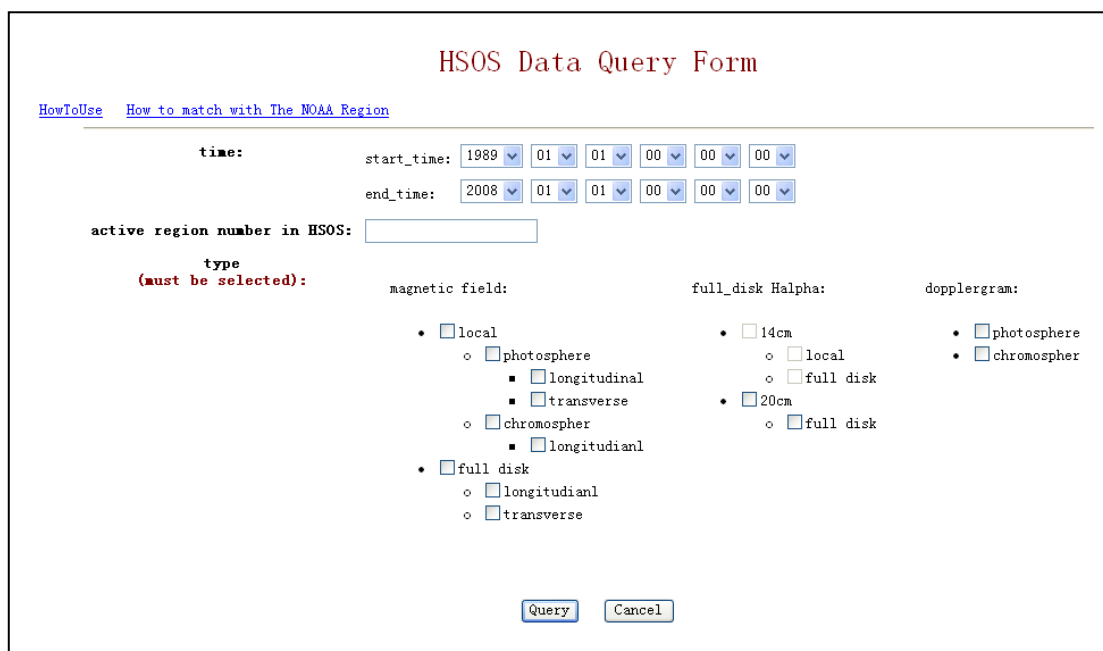
name:

password:

(guest:123)

图 4.4 用户登录界面

正确登录后，进入查询界面，如下图所示：



**HSOS Data Query Form**

[HowToUse](#) [How to match with The NOAA Region](#)

---

time: start\_time:

end\_time:

active region number in HSOS:

**type**  
(must be selected):

<p>magnetic field:</p> <ul style="list-style-type: none"> <li>• <input type="checkbox"/> local           <ul style="list-style-type: none"> <li>◦ <input type="checkbox"/> photosphere               <ul style="list-style-type: none"> <li>▪ <input type="checkbox"/> longitudinal</li> <li>▪ <input type="checkbox"/> transverse</li> </ul> </li> <li>◦ <input type="checkbox"/> chromosphere               <ul style="list-style-type: none"> <li>▪ <input type="checkbox"/> longitudinal</li> </ul> </li> </ul> </li> <li>• <input type="checkbox"/> full disk           <ul style="list-style-type: none"> <li>◦ <input type="checkbox"/> longitudinal</li> <li>◦ <input type="checkbox"/> transverse</li> </ul> </li> </ul>	<p>full_disk Halpha:</p> <ul style="list-style-type: none"> <li>• <input type="checkbox"/> 14cm           <ul style="list-style-type: none"> <li>◦ <input type="checkbox"/> local</li> <li>◦ <input type="checkbox"/> full disk</li> </ul> </li> <li>• <input type="checkbox"/> 20cm           <ul style="list-style-type: none"> <li>◦ <input type="checkbox"/> full disk</li> </ul> </li> </ul>	<p>dopplergram:</p> <ul style="list-style-type: none"> <li>• <input type="checkbox"/> photosphere</li> <li>• <input type="checkbox"/> chromospher</li> </ul>
---	---	--

图 4.5 查询界面

此界面提供了时间、活动区号、数据类型等查询参数，其中时间和数据类型是必选项。用户通过图中 4.5 所示的界面将查询条件提交给服务器，由服务器中的查询模块进行处理。以查询 2007-01-12 00:00:00 ~ 2007-01-12-02:00:00 的全日面矢量磁场数据为例，经服务器处理之后，得到的结果显示如下图 4.6 所示：



图 4.6 查询结果显示界面

在此界面中，用户选择需要下载的数据，进行数据下载提交之后，进入下载方法选择界面，图 4.7 所示：



图 4.7 下载方法选择界面

以选择 email 格式下载为例，后台程序将下载数据打包之后，将文件地址发送到用户的 Email 中。如果选择 URL 格式下载将在界面中直接得到文件的 URL。

如果管理员正确登录的话，则会进入管理员主界面（图 4.8），在此管理员可以进行功能选择：



- 1、查看数据下载情况；
- 2、查看数据浏览情况；
- 3、删除数据；
- 4、修改密码。

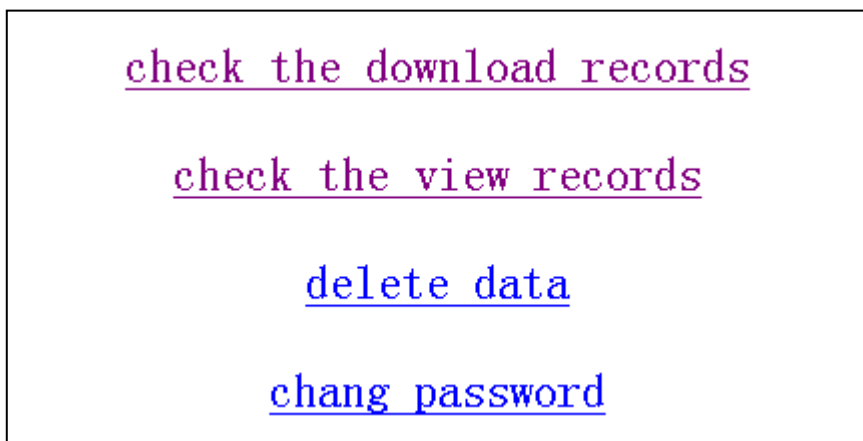


图 4.8 管理员的主界面

其中查看数据的下载情况可以有两种格式：图表格式和矩形图格式。

ip	address	email	view_dt	filetime	hsos_no	type	local_mag	full_mag	l4	ha_1	l4
192.168.254.254	局域网,对方和您在同一内部网	ljnoble@163.com	2007-03-02 10:19:23	2006-01-01 00:00:00-2006-01-02 00:00:00		.u5.q5.t5.r5	3	10	0	0	
150.144.30.169	美国, CZ88.NET	oneiros@grace.nascom.nasa.gov	2007-04-19 00:08:55	2005-01-01 00:00:00-2005-02-01 00:00:00		15.s5.u5.q5.t5.r5.l4.s4	2	8	0	0	
192.168.254.254	局域网,对方和您在同一内部网	gy@bao.ac.cn	2007-04-25 15:12:23	2002-05-22 00:00:00-2002-05-22 23:00:00		15.s5.u5.q5.t5.r5	6	0	0	0	
150.144.30.169	美国, CZ88.NET	oneiros@grace.nascom.nasa.gov	2007-06-06 23:36:38	2007-01-01 00:00:00-2007-01-01 23:00:00		ha_20	0	0	0	0	
81.89.4.185	罗马尼亚, CZ88.NET	crisd@sira.astro.ro	2007-08-25 00:44:24	2002-01-08 00:00:00-2002-01-08 23:00:00		15.s5.u5.q5.t5.r5.l4.s4.v5.d5.v4.d4.ha_20	94	0	0	0	
193.190.231.130	比利时, CZ88.NET	rodriguez@oma.be	2007-09-14 16:23:32	2007-05-19 00:00:00-2007-05-20 00:00:00		ha_20	0	20	0	0	
193.190.231.130	比利时, CZ88.NET	rodriguez@oma.be	2007-09-24 17:51:27	2007-05-18 18:00:00-2007-05-19 15:00:00		ha_20	5	0	0	0	

图 4.9 数据下载情况的表格显示

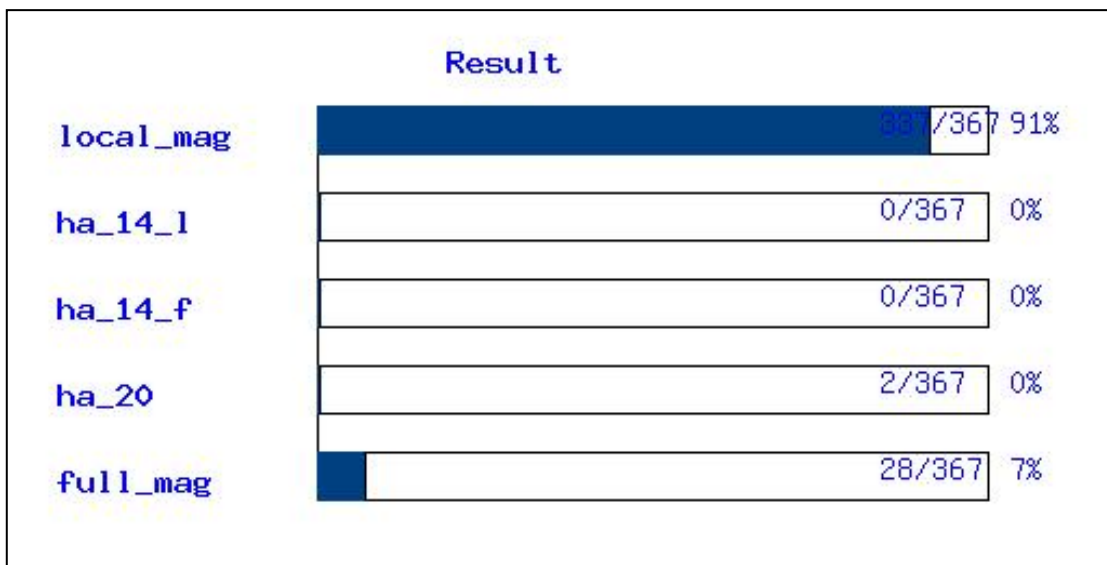


图 4.10 数据下载情况的图表显示

## 第五章 总结与展望

目前，随着探测器和空间技术的发展，天文学业已进入全波段巡天阶段，多波段数据从数量和类型上都急剧增多，VSO 为解决这些异地异构海量多波段数据的查询、获取和处理等问题提供了一个很好的平台，不仅 Hinode 而且今后的地面空间望远镜数据都将通过 VSO 被用户查询到，将怀柔数据资源加入 VSO 为我们今后更好地与国际进行交流打下了良好的基础。

构建怀柔太阳数据查询系统，将查询参数扩展为时间、数据类型和活动区，在数据查询参数方面进行了创新，此系统不仅极大地减少了人力，而且使用户能够方便快捷地查询下载特定的怀柔数据资源。

本论文介绍了在怀柔太阳数据资源的基础上所做的两份工作：构建虚拟太阳天文台的中国数据结点和构建怀柔太阳数据查询系统。论文的主要内容如下：

第一章主要从概念的提出、数据的组成等方面对国际虚拟太阳天文台（VSO）进行了概述，同时对怀柔太阳观测基地进行了简要介绍，指出了怀柔太阳观测基地加入 VSO 的必要性和重要性。

第二章概述了构建虚拟太阳天文台数据结点的相关技术。其中包括 XML、SOAP、WSDL，并着重介绍了 VSO 的数据模型 DM 和 VSO 简要的体系结构及其工作流程。

第三章详细介绍了怀柔太阳天文台数据结点的设计与实现，分别从模块、数据层和资源注册构造等方面进行了详细的阐述。

第四章主要介绍了怀柔太阳数据查询系统。首先介绍了系统的背景，接着从体系结构和用户工作流程等方面对系统方案进行了详细介绍。



## 发表文章目录

刘健, 林钢华. 构建虚拟太阳天文台的中国数据结点, 天文技术与方法, 2008  
(已收录)



## 附录

## 数据库表结构及说明:

file\_in\_table tab: 纪录 tab 表文件中信息

列名	说明	类型
filename	tab 表中纪录的所观测的数据文件名	char(19)
dt	文件的观测时间	datetime
hsos_no	文件所观测的活动区	char(15)
quality	文件的观测质量	char(1)

hsos\_no\_in\_table: 活动区表

列名	说明	类型
hsos_no	活动区编号	char(5)
s_dt	活动区的开始时间	datetime
e_dt	活动区的结束时间	datetime
seeing	视宁度	char(1)
climate	天气指数	char(6)

full\_mag: 全日面矢量磁场表

列名	说明	类型
datetime	时间	datetime
file	文件名	char(25)
path	存储路径	char(87)

full\_mag\_picture: 全日面矢量磁场图表

列名	说明	类型
dt	日期	date
path	存储路径	char (60)

filename	文件名	char (15)
----------	-----	-----------

ha\_20: 20cm 全日面 Halpha 表

列名	说明	类型
datetime	时间	datetime
filename	文件名	char (28)
path	存储路径	char (70)
dt	日期	date

ha\_20\_c: 20cm 全日面 Halpha 的表文件表

列名	说明	类型
dt	日期	date
path	存储路径	char (72)

ha\_20\_picture: 20cm 全日面 Halpha 的图像表

列名	说明	类型
dt	日期	date
path	存储路径	char (73)
filename	文件名	char (29)

mag\_file\_path: 35cm 望远镜观测的数据存储表

列名	说明	类型
filename	文件名	char (22)
path	存储路径	char (60)

mag\_picture: 35cm 望远镜观测数据的图片表

列名	说明	类型
filename	文件名	char (25)
path	存储路径	char (65)

ip\_address: ip 转化表



列名	说明	类型
s_ip	ip 段的首地址	char(15)
e_ip	ip 段的末地址	char(15)
address	ip 的物理地址	char(60)
s	ip 段首地址的数值转化	char(11)
e	ip 段末地址的数值转化	char(11)

user: 用户表

列名	说明	类型
id	ID	int
name	姓名	char (8)
password	密码	char (50)

user\_log: 下载数据纪录表

列名	说明	类型
ip	IP	char(15)
download_dt	下载时间	datetime
file_time	下载文件的时间	char (40)
hsos_no	下载文件的活动编号	char (5)
type	下载文件的类型	char (50)
local_mag	下载的局部矢量磁场的个数	int
14_ha_l	下载的局部 14cmHalpaha 的个数	int
14_ha_f	下载的全日面 14cmHalpaha 的个数	int
20_ha	下载的 20cm 全日面 Halpaha 的个数	int
full_mag	下载的全日面矢量磁场的个数	int
email	下载用户的 email	char (35)
address	下载用户的 IP 对应的物理地址	char (60)

view\_user\_log: 浏览数据纪录表

列名	说明	类型
----	----	----

构建虚拟太阳天文台的数据结点

ip	IP	char(15)
download_dt	下载时间	datetime
file_time	下载文件的时间	char (40)
hsos_no	下载文件的活动编号	char (5)
type	下载文件的类型	char (50)
local_mag	下载的局部矢量磁场的个数	int
14_ha_l	下载的局部 14cmHalpHa 的个数	int
14_ha_f	下载的全日面 14cmHalpHa 的个数	int
20_ha	下载的 20cm 全日面 HalpHa 的个数	int
full_mag	下载的全日面矢量磁场的个数	int
email	下载用户的 email	char (35)
address	下载用户的 IP 对应的物理地址	char (60)

## 核心代码：查询模块

```
<?php
require('./ip_change.inc');

/*****

连接数据库

*****/

    $mylink=mysql_connect("localhost","name","password") or die ("cannot connect
database!");

    mysql_select_db("hsos_db",$mylink) or die("cannot select the database!");

/*****

接收从浏览器中传入的参数：时间、活动区、数据类型

*****/

$year1=$_GET['year1'];
$month1=$_GET['month1'];
$day1=$_GET['day1'];
$hour1=$_GET['hour1'];
$min1=$_GET['min1'];
$sed1=$_GET['sed1'];

$year2  = $_GET['year2'];
$month2 = $_GET['month2'];
$day2   = $_GET['day2'];
$hour2  = $_GET['hour2'];
$min2   = $_GET['min2'];
$sed2   = $_GET['sed2'];

$hsos_no=trim($_GET['hsos_no']);
```

```
@$ty = $_GET['ty'];
@$ha_p_f = $_GET['ha_p_f']; //ha_14_f
@$ha_p_p = $_GET['ha_p_p']; //ha_14_l
@$ha_n_f = $_GET['ha_n_f']; //ha_20
@$ty_f = $_GET['ty_f'];

$s_dt = "$year1-$month1-$day1 $hour1:$min1:$sed1";
$e_dt = "$year2-$month2-$day2 $hour2:$min2:$sed2";

$str2=array();

$num_row_f = 0;

$name = $_SESSION['name'];
$pass = $_SESSION['pass'];
$_SESSION['s_dt'] = $s_dt;
$_SESSION['e_dt'] = $e_dt;
$_SESSION['hsos_no'] = $hsos_no;
$_SESSION['ha_p_p'] = $ha_p_p;
$_SESSION['ha_p_f'] = $ha_p_f;
$_SESSION['ha_n_f'] = $ha_n_f;
$_SESSION['ty'] = $ty;
$_SESSION['ty_f'] = $ty_f;

$num_rows=0;
$total_num_ha_14_l=0;
$total_num_ha_14_f=0;
$total_num_ha_20=0;
$total_f_m=0;

?>
```

```
<?php
/*****

根据不同的用户显示不同的界面：

admin: 具有删除功能

guest: 有下载功能

*****/

if($name =='admin')           //若是 admin 的话， 显示 delete 按钮

    echo" <form  name = 'formB' action = 'del1.php' method ='get'><h4 align ='right'><input
type ='submit' value ='delete'></h4>";

    else                       // 若不是 admin， 显示 download 按钮

    echo" <form  name ='formB' action = 'download.php' method ='get'><h4 align
='right'><input type ='submit' value ='Request Data'></h4>";

?>

/*****

显示查询结果

*****/

<?php
//显示 35cm 望远镜的查询结果

    if($ty!="")

        $num_rows= showtable_1_mag($name,$s_dt,$e_dt,$hsos_no,$ty,");

//显示 14cm Halpha 局部的查询结果

    if($ha_p_f!="")

        $total_num_ha_14_f= showtable_ha_14_f($s_dt,$e_dt);

//显示 14cm Halpha 全日面的查询结果

    if($ha_p_p!="")

        $total_num_ha_14_l= showtable_ha_14_l($s_dt,$e_dt,$hsos_no);
```

```

//显示 20cm Halpha 全日面的查询结果

if($ha_n_f!="")

    $total_num_ha_20= showtable_ha_20($s_dt,$e_dt);

//显示全日面矢量磁场的查询结果

if($ty_f!="")

    $total_f_m= showfull_mag($s_dt,$e_dt,$ty_f);

/*****

纪录用户的查询情况：IP、查询时间、查询数据的情况——时间段、活动区号、数据
类型、各种类型的返回个数，并将这些信息插入到 view_user_log（数据浏览纪录表）中

*****/

$ip          =          getenv("REMOTE_ADDR");

$address = ip2adr($ip);

$view_dt = date('Y-m-d-H:i:s');

$file_time    = "$s_dt~$e_dt";

// 构建 SQL 语句

$sql = 'insert into view_user_log values';

$sql

=$sql."('$ip','$view_dt','$file_time','$hsos_no','$type','$num_rows','$total_num_ha_14_l','$total_
num_ha_14_f','$total_num_ha_20','$address','$total_f_m')";

mysql_query($sql);          //将纪录插入到 view_user_log 中

?>

<?php

/*****

构建查询全日面矢量磁场的 SQL 语句，执行查询，并以 html 的格式显示查询结果

*****/

function showfull_mag($s_dt,$e_dt,$ty_f)

{

}

```

```

/*****
构建相应的 SQL 语句，用以查询由 35cm 望远镜观测到的数据，执行 SQL 语句，并以
html 格式显示查询结果
*****/

function showtable_1_mag($name,$s_dt,$e_dt,$hsos_no,$ty,$q)
{
}

/*****
构建查询 14cmHalpaha 的局部磁场数据的 SQL 语句，执行查询，并以 html 的格式显示
查询结果
*****/

function showtable_ha_14_l($s_dt,$e_dt,$hsos_no)
{
}

/*****
构建查询 14cmHalpaha 的全日面磁场数据的 SQL 语句，执行查询，并以 html 的格式显
示查询结果
*****/

function showtable_ha_14_f($s_dt,$e_dt)
{
}

/*****
构建查询 20cmHalpaha 的全日面磁场数据的 SQL 语句，执行查询，并以 html 的格式显
示查询结果
*****/

function showtable_ha_20($s_dt,$e_dt)
{
}

?>
```





## 核心代码：下载模块

```
<?php

$email=$_GET['email'];

if(!isset($email))

/*****

未获取用户的 email 信息，则接收用户的下载数据情况，并要求用户提交 Email 信息

*****/

{

//接收用户所下载的数据情况：

@sha_14_f      = $_GET['ha_14_f'];

@filename_ha_p = $_GET['filename_ha_p'];

@filename_1_m  = $_GET['filename_1_m']; //local_mag && v  选中下载的文件

@sha_20_f      = $_GET['ha_20_f'];

@filename_f_m  = $_GET['filename_f_m'];

$_SESSION['ha_14_f']      = sha_14_f;

$_SESSION['filename_ha_p'] = filename_ha_p;

$_SESSION['filename_1_m'] = filename_1_m;

$_SESSION['ha_20_f']      = sha_20_f;

$_SESSION['filename_f_m'] = filename_f_m;

$total_size=0;

//计算所下载数据的大小

if(isset($sha_20_f)) {

while($var=current($sha_20_f)) {

$total_size+=filesize("../$var");

next($sha_20_f);

}

}

if(isset($filename_1_m)) {

while($var=current($filename_1_m)) {
```

```
$total_size+=filesize("../$var");
    next($filename_l_m);
}
}
if(isset($filename_ha_p)) {
    while($var=current($filename_ha_p)) {
        $total_size+=filesize("../$var");
        next($filename_ha_p);
    }
}
if(isset($ha_14_f)) {
    while($var=current($ha_14_f)) {
        $total_size+=filesize("../$var");
        next($ha_14_f);
    }
}
if($filename_f_m != "") {
    while($var=current($filename_f_m)) {
        $total_size +=filesize("../$var");
        next($filename_f_m);
    }
}
$total_size=intval($total_size/(1024*1024));

?>
<html>
.....
<form name="register" method="get" action = download.php >
<p align="left"><br>
<font face ="Alias" color ='#800000">
```

```

An email address is required<p></font><input type="text" name="email" size="20">
<hr>
<input type=radio name=download_type value ="url" <?php echo $k;?>>
<font face ="Alias" color ="#800000">url</font>
<p>
  <input type=radio name=download_type  value ="email" checked>
<font face ="Alias" color ="#800000">email</font>
<p><hr><p>
<font face ="Alias" color ='blue'>you totally want to download <?php //echo
$total_size.'M'; ?><br>
  <?php
    if($total_local_mag>0)          echo "Local Magnetogram:  $total_local_mag
items<br>";
    if($total_mag_f>0)             echo "Full Disk Magnetogram:  $total_mag_f items<br>";
    if($total_ha_20>0)             echo "Full Disk H-alpha(20 cm):  $total_ha_20
items<br>";
  ?>
</font><p><hr><p>
<input type ='hidden' name='size' value='<?php echo $total_size;?>'>
<input type ='hidden' name='size1' value='<?php echo $total_local_mag;?>'>
<input type ='hidden' name='size2' value='<?php echo $total_ha_20;?>'>
<input type ='hidden' name='size3' value='<?php echo $total_mag_f;?>'>
<input type ='hidden' name='size4' value='<?php echo $total_ha_14_l;?>'>
<input type ='hidden' name='size5' value='<?php echo $total_ha_14_f;?>'>

<input type="submit" value="Request Data" name="B1" onclick = "return check();">
<input type="reset" value="cancle" name="B2">
.....
</html>
<?

```

```
}  
else  
//若已经获取用户的 email 信息,则根据用户选择的数据下载方式, 进行相应的处理  
{  
require('./ip_change.inc');  
$email =trim($email);  
$detail_mail = explode("@",$email);  
$d_name = $detail_mail[0];  
//连接数据库  
$mylink=mysql_connect("localhost","name","password") or die ("cannot connect  
database!");  
mysql_select_db("hsos_db",$mylink) or die("cannot select the database!");  
  
//获取所要下载得数据的 URL 信息及其他相关信息  
$ha_14_f      = $_SESSION['ha_14_f'];  
$filename_ha_p = $_SESSION['filename_ha_p'];  
$filename_l_m  = $_SESSION['filename_l_m'];  
$ha_20_f      = $_SESSION['ha_20_f'];  
$filename_f_m  = $_SESSION['filename_f_m'];  
$hsos_no      = $_SESSION['hsos_no'];  
$s_dt         = $_SESSION['s_dt'];  
$e_dt         = $_SESSION['e_dt'];  
$ha_p_p       = $_SESSION['ha_p_p'];  
$ha_p_f       = $_SESSION['ha_p_f'];  
$ha_n_f       = $_SESSION['ha_n_f'];  
$ty           = $_SESSION['ty'];  
@$ty_type     =$_SESSION['ty_type'];  
$download_type = $_GET['download_type'];  
//获取用户的下载信息: IP、下载时间、下载数据的信息,一时间段、数据类型及数据量  
$ip          = getenv("REMOTE_ADDR");
```

```

$address = ip2adr($ip);

$download_time = date('Y-m-d-H:i:s');

$file_time      = "$s_dt~$e_dt";

@$total_local_mag = count($filename_l_m);

@$total_ha_14_l = count($filename_ha_p);

@$total_ha_14_f = count($ha_14_f);

@$total_ha_20 = count($ha_20_f);

@$total_full_mag = count($filename_f_m);

$i=0;

$ex_dt=0;

$dt =date('YmdHi');

$star_name=" ../hsos_datas/tmp/$dt$d_name.tar.gz";

$starname="$dt$d_name.tar.gz";    //构建打包数据数据的名称

if($download_type=="email")

{

    $date2 = mktime("00","00","00",date('m'),date('d')+3,date('Y'));

    $date2 = date('Y-m-d',$date2);

    echo "<font face ='Alias'>your file address will be sent to <font color = 'Blue'>
$email</font><br>and the file will be kept until<font color = '#800000'>
$date2<p></font>Thank you for using the HSOS Data Query Form! Please report any problems
to <a href='mailto:lgh@bao.ac.cn'>lgh@bao.ac.cn</a></font><br>";

}

$star_filename = "";

//打包下载数据

if($ha_14_f!="")          { tar_File($ha_14_f, $starname); }

if($ha_20_f!="")          { tar_File($ha_20_f, $starname); }

if($filename_ha_p!="")    { tar_File($filename_ha_p, $starname); }

if($filename_f_m!="")     { tar_File($filename_f_m, $starname); }

function tar_File($filenames, $starname) { ... } //将$filenames 打包到$starname

```

文件中

```

/*****

获取文件的下载信息，并将其插入到 user_log（数据下载情况表）中

*****/

$type="";

if($ty!="")
{
    $x= str_replace("/", '!', $ty_type);
    $type=$type.$x;
}

if($ha_p_p!="")
    $type= "$type.$ha_p_p";

if($ha_p_f!="")
    $type = "$type.$ha_p_f";

if($ha_n_f!="")
    $type="$type.$ha_n_f";

$sql="insert into user_log values";

$sql=$sql ."('$ip','$download_time','$file_time','$hsos_no','$type','$total_local_mag','$total_
ha_14_l','$total_ha_14_f','$total_ha_20','$email','$address','$total_full_mag)";

mysql_query($sql) or die('user_log wrong');

//若选择 URL 的下载方式，则直接将单个文件的 URL 通过浏览器返给用户

if($download_type == "url")
{
    system("tar czf $tar_name $tar_filename");

    echo "download your file:<br><font color = '#0080ff'>
<a href='/hsos_datas/tmp/$tarname'> $tarname </font></a>";

    echo "<br><br><center><a href='./query_jiemian.php '>return </a></center><br><br>";
}

```

//若选择 Email 下载方式，则将数据进行打包，然后将地址发给用户以供下载

```
if($download_type=="email")
{
    $to = "$email";
    $subject = "Your HSOS Data Request has Completed";
    $msg = "Information on your jobs is available at:

    http://sun.bao.ac.cn/hsos_datas/tmp/$starname

    Attention please: your information will be kept until $date2";
    $headers="From : lgh@bao.ac.cn";
    mail("$to","$subject","$msg","$headers");
    system("tar cvzf $star_name $star_filename >/tmp/null&");
}
.....
?>
```





## 致谢

三年时间转眼即逝，回顾走过的路。在国家天文台的三年是我人生中一笔非常宝贵的财富。在这三年中给我帮助的人太多太多，非常高兴能够有这个机会表达我的感激之情。

首先感谢我的指导老师林钢华，感谢林老师从我入学一直到毕业这三年的时间里给予我一如既往的关怀和孜孜不倦的教诲。三年中，从我的课程选择到科研项目到写文章到写论文，林老师都给予我启迪，拓展我的思路，引导我学术思维，教给我学术方法。虽然林老师日常工作非常繁忙，但是她还是对我的每一个问题都给予细致的指导和帮助，同时还给我提供了良好的软硬件环境和学习环境。导师言传身教是我两年来最大的收获。我将刻苦研究，不断进取，只有这样才能不辜负老师的期望。

感谢怀柔太阳观测基地的首席研究员张洪起老师。虽然张老师工作非常的繁忙但是对我们的个人发展、生活还是非常的关心，给我们以指导。同时张老师为人谦和、大气以及严谨的治学态度和对科学的热爱，为我们树立了学习的榜样。张老师是我非常敬重的一位师长，在此向张老师致以深深的敬意和感谢。

感谢邓元勇老师，虽然邓老师不是我的指导老师，但是对我的工作也给予了同样的关心、提出了很多宝贵的意见和建议，在此向邓老师表示深深的感谢。

感谢杜红荣老师，在三年的学习工作中，杜老师给予了我无私的帮助和关心，杜老师对同学们真诚爱护，为人处事平易近人、严于律己永远是我学习的榜样，在此向杜老师致以崇高的敬意和深深的感谢！

感谢崔辰州老师，我的一部分工作是在崔老师的帮助下完成的。崔老师深厚的技术基础、严谨的治学态度、对事业的执著追求以及平易近人的品质都是我学习的榜样，在此向崔老师致以深深的感谢。

感谢王东光老师，感谢王东光老师在百忙之中给予我的关心和挂念。

感谢张枚老师，张老师严谨的治学态度和对科研的热爱给我留下了深刻的印象，在今后的工作中，我一定以张老师为榜样努力工作。

感谢怀柔基地的其他老师和同学，因为有了大家的关怀，我的学习和生活

才充满了快乐。他们是：李威、王艺、胡科良、包星明、汪国萍、苏江涛、徐海清、郝娟、谢文彬、玄伟佳、王栋、高裕、王传宇、王晓帆、杨尚斌、刘继红、孙英姿、陈洁、申基、肖江等等。

感谢国家天文台的 2005 级的朋友们，特别感谢朱兰、肖丽、宋乔、薛元、刘姝，杨阳，这些可爱的姐妹让我的生活充满了温馨。

感谢来自 LAMOST 的好朋友田海俊，他对我的工作和生活给予了很多的帮助和支持。

感谢我的家人，正是有了你们的关心和爱护，我才有了前进的动力！