Virtual Observatory Interface to a Stellar Spectral Library

Report of Summer Internship work at The Inter-University Centre for Astronomy and Astrophysics, Pune

Under the guidance of

Professor A. K. Kembhavi

submitted by
Parikh Madhura Madhukar*

(ENGS3645)

*Sardar Vallabhbhai National Institute of Technology, Surat.

Virtual Observatory Interface to a Stellar Spectral Library

Abstract

The *Virtual Observatory* (VO) is a revolutionary concept that aims at making astronomical and astrophysical data, globally accessible. It is also a platform that provides data analysis tools that can help scientists to effectively and efficiently analyse this data. In this project we develop one such VO interface for a stellar spectral library. The first section describes the interface that was thus developed and the functionalities that it provides. In the next section we will explore briefly the technologies that were used to make such an interface as well as some programming aspects of the project. In the third section we examine the *Observational Table Access Protocol (ObsTAP)* - a data model created by the *International Virtual Observatory Alliance (I-VOA)* to make global data discovery easy.

1 Introduction

The Virtual Observatory(VO)[1] is a highly innovative concept that provides a platform for integrating astronomical data archives from around the world as well as computational tools for interpreting this data. It is a truly unprecedented move, that makes it possible for researchers to avail top-notch data, which they would otherwise have no means of accessing. Virtual Observatory-India(VO-I), developed at IUCAA, is India's contribution to the VO project. In this summer project, we intend to develop a VO tool that interfaces with the Near-Infrared stellar spectral library[2], based on the observations made with the Gurushikhar Infrared Telescope(GIRT), Mt. Abu, India. The library is an important database for solar type stars, with spectra collected at the J, H and K bands. The complete library may be accessed online[1]. Though the library hosts valuable data, there are no query tools currently available, that may aid the user to extract and analyze only portions of the data that he(she) requires. The user is instead required to download the entire datafile for the object of interest. We aim to remove this constraint by developing a web platform through which the user can query over a set of parameters, and can quickly and efficiently extract the desired data. A detailed description of this web interface is given in the next two sections. We also made a study of the ObsTAP protocol, one emerging standard in the VO technology.The ObsTAP is a I-VOA recommended standard that is intended to help a remote user to query several databases from different services simultaneously with a single query. The protocol lays down the model that an ObsTAP service should follow to make this possible. An overview of this protocol is given in the third section.

[1] `http://vo.iucaa.ernet.in/~voi/NIR_Header.html`

2 The Web-based VO Interface

2.1 A brief introduction to some VO tools and techniques

The VOTable: The VOTable is the standard for data exchange in the Virtual Observatory platform. It is constructed with XML(eXtensible Markup Language). It is a powerful and efficient data model that meets the flexibility and interoperability goals for VO data[2].

SIMBAD: SIMBAD stands for Set Identification, Measurement and Bibliography for Astronomical Data.As is evident from the name it is an extensive database for objects outside the Solar System and has been developed at CDS,Strasbourg[3].

NVO registry: It is maintained by the National Virtual Observatory, USA. As they explain the registry is a sort of yellow pages or high level directory of astronomical data archives and computational tools, originating from a wide variety of instruments, across the globe. It is therefore a preferred resource when astronomers need to look up some data[4].

VizieR: VizieR, is an exhaustive collection of online astronomical catalogues and is maintained by CDS[5].

Aladin: Aladin is a widely used VO tool developed again at CDS. It is a sky atlas software that allows the user to visualize astronomical images and datasets[6].

VOStat: VOStat is a VO tool that allows the user to perform statistical computations on the dataset. It uses the open source software package 'R' to perform the computations[7].

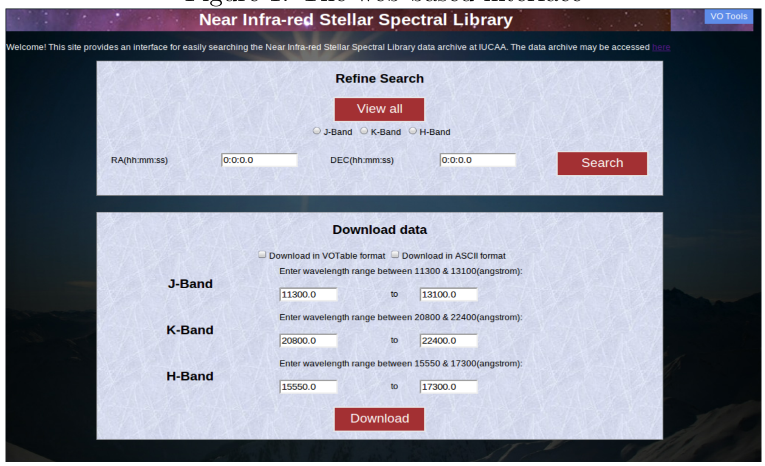
2.2 Functionalities provided by the interface

The IUCAA data archive has files where each file contains the observed data for one particular object. These observations were made in 3 different wavelength bands : J , K and H. Thus corresponding to an object say $HR1948$ there would be 3 data files - 1948*j.dat*, 1948*k.dat* and 1948*h.dat* - for observations made in each of the fore-mentioned bands. (However for some objects observations may not have been taken in all the three bands, in that case only data files corresponding to the relevant bands for the objects would be present). These data files are ASCII files and consist of two space-separated columns of values. The first column represents the wavelength in angstroms (with $5\mathring{A}$ bins) and the second column is the relative flux observed at those wavelengths. The earlier website for the spectral library provided links to individual files. The user had two options : One could download all files of interest, one by one, by clicking on the individual links or one could download the entire archive containing all the files. One required use case is that the user may want to extract data corresponding to only a particular wavelength range, rather than downloading the complete file. We now look at some of these features that have been incorporated in the VO interface we developed.

- The user can click on the $View\ all$ button to see all the files that are present in the database. The user can then tick the checkbox against those files that he wishes to download. Alternatively the user can use the options: $J-Band$, $K-Band$ or $H-Band$, to view only the files related to a particular band.

[2] `http://www.ivoa.net/cgi-bin/twiki/bin/view/IVOA/IvoaVOTable`
[3] `http://simbad.u-strasbg.fr/simbad/`
[4] `http://www.us-vo.org/`
[5] `http://vizier.u-strasbg.fr/viz-bin/VizieR`
[6] `http://aladin.u-strasbg.fr/`
[7] `http://vo.iucaa.ernet.in/~voi/VOStat.html`

Figure 1: The web-based interface



- Since the position of an object is frequently used to query an astronomical database, we have provided a search interface wherein the user can enter the *RA* and *DEC* co-ordinates of an object and search for its data files. The *RA* and *DEC* values were not present in the original data. However a program that can query the SIMBAD astronomical database for an object and fetch its *RA* and *DEC* values, had been written at VO-I. We used this program to populate our database with the *RA* and *DEC* values of all the objects in the archive.
- The interface also allows the user to input the starting and ending wavelengths corresponding to each of the bands in text-boxes. Only the data contained in these ranges will be outputted in the user-selected files when they are downloaded.
- An option is available to the user to download all the files in VOTable format rather than as ASCII files. Since VOTable is the standard format for data exchange in the Virtual Observatory, this will help the user to easily use other available VO tools for analysing the data.
- Along with each file listing, we have also provided the SIMBAD, NVO registry and VizieR links for that object. Clicking on these, the user can access all the information for that object across datasites globally.
- As the data may be downloaded in VOTable format, the user can now easily avail of several VO tools. On the top right corner of our website we provide a drop down menu that links to some such VO tools such as Aladin and VOStat, that may be launched online as an applet.
- Another interesting option that we provide the user is to plot the spectrum of the object in VOPlot. The user can view the spectrum by clicking on the 'Plot spectrum in VOPlot' link. The VOI team, IUCAA made modifications in the existing VOPLot, to allow for this functionality in the website.

Figure 2: A part of the file listing with links to SIMBAD and NVO. The search results with the queried RA and DEC are returned highlighted.

Object	Band	Type	External Links
HR4033	H	A	Plot spectrum in VOPlot
HR5291	H	A	Plot spectrum in VOPlot
HR3131	H	A	Plot spectrum in VOPlot
HR3975	H	A	Plot spectrum in VOPlot
HR4689	H	A	Plot spectrum in VOPlot
HR4963	H	A	Plot spectrum in VOPlot
HR5531	H	A	Plot spectrum in VOPlot

3 An overview of the technologies used in the interface

In this section we attempt to describe briefly the web technologies and software tools we used in coding up the interface.

The DBMS: We have used the Oracle 10g XE as our DBMS. This is an open source version of the popular Oracle RDBMS. The database schema is fairly simple, comprising of a one relation, that stores all the pertinent information of an object : the object name, its RA and DEC co-ordinates, the band, the spectral type and the pathname of its corresponding data file in the server. Though we have used the Oracle RDBMS, we have also exported the schema to another popular and open source DBMS - PostgreSQL. For this we have used the freely available Ora2Pg software tool.

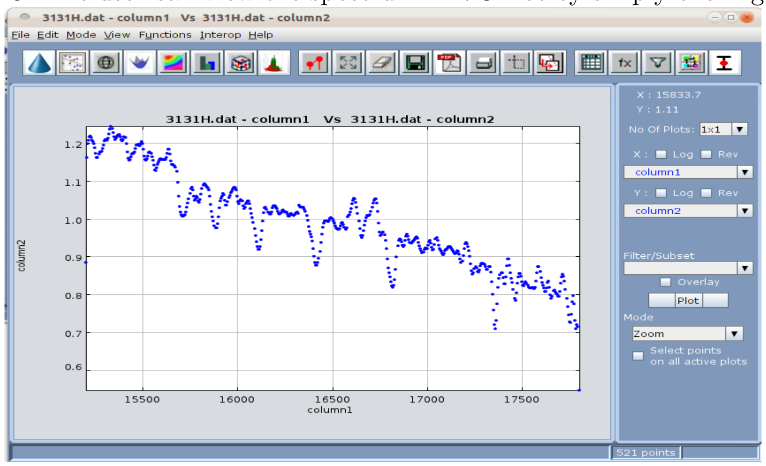
The Server: We use the GlassFish application server for deploying the web application. It is an open source project started by the Sun Microsystems and now supported by the Oracle Corporation.

Java Servlet technology: We use the Java Servlets on the server side for input handling. These receive the HTTP Post and GET requests from the client side and afterwards serve the corresponding HTTP response back to the client. For connecting to the DBMS we use the Java Database Connectivity(JDBC) API. The JDBC allows easy integration of SQL statements into the Java program. A very helpful and hands-on introduction to JDBC and Servlets may be found at the online site of the Stanford InfoLab[8].

AJAX and JavaScript: On the client side we use AJAX and JavaScript. The Asynchronous JavaScript and XML(AJAX) is a revolutionary web technology. Since requests to the server made with AJAX are asynchronous, the browser does not block when the request is made. Further the response from the server does not need to cause a full page refresh. We have used AJAX in sending requests and receiving responses from the server.This

[8] `http://infolab.stanford.edu/~ullman/fcdb/oracle.html`

Figure 3: The user can view the spectrum in VOPlot by simply clicking a link



requires using the XMLHttpRequest object in JavaScript. Also we use the highly flexible and easy to use XML format for sending all the responses, back from the server. We have used JavaScript for performing client side validations.

Shell scripting: One major use case of the web app is to serve out data files while trimming them so that they contain only data only between the user-specified wavelengths. This file processing has been handled via a Unix shell script where we also use the Sed filter.

Converting ASCII files to the VOTable format: The files were available in the IUCAA data archive, originally only in the ASCII format. However significant gains can be made if the data is made available in the VOTable format. To also supply the data to the user in the VOTable format, we use the VOConvert[9] software that was already developed by VO-I,IUCAA. The command line version of this software is interactive. To use it for our purposes, non-interactively, we got around by using the Except Unix tool.

The Eclipse IDE: For developing the entire project we used the Eclipse Integrated Development Environment, upgraded to Java EE (Enterprise Edition) capabilities. The IDE provides a very efficient and user-friendly environment for developing and organizing such a dynamic web project.

4 A look at the I-VOA Observational Table Access Protocol(ObsTAP)

4.1 The I-VOA Table Access Protocol(TAP)

The TAP is a web-service protocol that may be used to access tabular data that is stored in the VO framework. It works atop the HTTP. Queries to the data servers are typically sent with the *A*stronomical Data Query Language(ADQL)-although support may also be available for using query languages like the SQL and Parametrized Query Language(PQL) - and the results are returned in the standard VOTable format, other formats are optional. The queries may be both

[9] http://vo.iucaa.ernet.in/~voi/VOConvert.htm

*s*ynchronous(data is immediately returned in response of the request) or*a*synchronous(data is returned when another request is made after the original request) and these queries are made by passing the query components as parameters of the HTTP GET or POST request made to the URL of the data server. TAP-service tablesets are stored usually in an RDBMS. Any TAP service must mandatorily have in its table sets a table called the TAP_SCHEMA that contains all the metadata pertaining to the data tables. A query posted by the user may either be a :

Data query: This is used to retrieve the astronomical data from the server site.

Metadata queries: These are similar to the previous queries but they are used to query tables that will reveal the information about the schema of data tables stored in the RDBMS at the server site. By using the information obtained from the metadata queries, the user can then know which attributes and relation names to use in the data queries.

Virtual Observatory Support Interface(VOSI): VOSI returns all the necessary metadata of the service in the standard XML format. Its usage is the same across all the TAP services.

A detailed description of various services and parameters that must be implemented by a TAP service provider can be found in the I-VOA TAP documentation[10]. The TAP thus provides a uniform way of accessing tabular data from different data sites and is an important milestone in reaching the I-VOA goal of interoperabilty.

4.2 The Observational Data Model

The Observational data model is devised with the aim of providing a standard common data model that may be used to describe any sort of astronomical observation. Thus objects described in the Observational data model may be for instance an image of the sky or some spectral dataset. Laying down a model that can describe such vastly varied data by a common set of attributes is no small challenge. However this common model for all astronomical observations can dramatically improve interoperability[11].

4.3 ObsTAP - Observational Table Access Protocol

As the name suggests, the ObsTAP combines the Observational data model with the TAP in a standard service that allows the user to simultaneously query multiple sites for a particular data by means of a single query. Any type of data that is stored at any site can be easily discovered in one go. Since the data may be vastly different and complex various other services such as the Simple Image Access(SIA) and the Simple Spectral Access(SSA) may then be used to appropriately handle the data. Simple data files may be directly downloaded as such. Thus a *d*ata-link service is to be provided that communicates with the appropriate VO service (such as SIA, SSA, etc.) for handling the data once that it is discovered. The detailed specifications of how the TAP_SCHEMA in a TAP service should be modified so that it can be upgraded to an ObsTAP service may all be accessed in the I-VOA documentation[12].

5 Acknowledgements

I am very thankful to INSA, NASI and IASc, for this opportunity to undertake fellowship at a wonderful institute like IUCAA. The extraordinarily well-stocked library, great computing

[10] `http://www.ivoa.net/Documents/TAP/`

[11] `www.ivoa.net/Documents/Notes/DMObs/DMObs-20050421.pdf`

[12] `http://www.ivoa.net/Documents/ObsCore/`

facilities and extremely cheerful and helpful staff at IUCAA have made this fellowship a truly memorable experience. I am especially grateful to my mentor Prof. A. K. Kembhavi under whose patient and able guidance, I learnt many new things in the field. I am also very glad for the constant support and help that I received from the VOI-team, IUCAA and specially from Mr. Sharmad Navelkar. I am indebted to my parents for their encouragement and valuable advice, and also to all my teachers at SVNIT. I also would like to thank all my friends who have always helped me out in times of need.

References

[1] Brumfiel G, The heavens at your fingertips, *Nature Vol. 420*, Nov 2002.

[2] Ranade A, et al, A near infra-red stellar spectral library:III : J-Band spectra, *Bull. Astr. Soc. India, 35*, 2007.