CROSS-IDENTIFICATION OF RADIO SOURCES FROM A LARGE NUMBER OF HETEROGENEOUS CATALOGUES USING VO TOOLS

B. Vollmer, S. Derrière, T. Boch, B. Gassmann, P. Dubois, F. Genova, F. Ochsenbein\(^1\) and T.P. Krichbaum\(^2\)

\(^1\)CDS, Observatoire astronomique de Strasbourg, UMR 7550, 11 rue de l’université, F-67000 Strasbourg
\(^2\)Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, D-53121 Bonn

ABSTRACT

The SPECFIND catalogue of radio cross-identifications and spectra, which is available in VizieR (Ochsenbein et al. 2000), has been used to search for Gigahertz peaked source candidates. These sources were observed quasi-simultaneously with the Effelsberg 100-m radio telescope at 6 cm (4.85 GHz), 2.8 cm (10.45 GHz), and 9 mm (32 GHz). It turned out that this is an efficient procedure to discover new Gigahertz peaked sources, which are believed to be AGNs at the beginning of their radio evolution. A new SPECFIND V2.0 catalogue is presented. It contains the cross-identification of 87000 radio objects from 105 catalogues or 3.76 million radio sources. With an increase of 8% of available sources we increase the number of radio objects with associated radio spectra by 25%. Going from 20 catalogues in the previous SPECFIND release to 105 catalogues in the new release was only possible due to the development of three Virtual Observatory (VO) tools within the European VO-TECH project. These new tools (i) identify pertinent radio catalogues in the VO registry using Unified Content Descriptors (UCDs), (ii) extract relevant data, and (iii) normalize these for the determination of radio spectra.

Key words: E\(\LaTeX\); Virtual Observatory.

1. INTRODUCTION

The cross-identification of radio sources observed in the centimeter to meter wavelength domain with different instruments is a rather difficult task, because of huge differences in sensitivity and/or spatial resolution. On the other hand, most sources show a power-law spectral energy distribution in these wavelengths due to synchrotron or thermal emission. Synchrotron emission produces a power law spectrum with a possible cut-off or reversal of the spectral index at low frequencies due to self-absorption or comptonization. The spectrum of thermal electrons is flat in the optically thin domain. The SPECFIND cross-identification tool takes advantage of the power-law shape of the spectra. In addition, it takes into account the angular resolution of the observations, the source size, and the flux densities at a given frequency. The SPECFIND tool also insures that a radio source cannot be assigned to more than one physical object. In a first release (Vollmer et al. 2005a,b) we cross-identified the sources of the 20 largest radio catalogues in VizieR (Ochsenbein et al. 2000), representing 3.5 million sources. Our work lead to more than 700,000 independent radio cross-identifications and \(\sim 67,000\) independent radio spectra with more than two independent frequencies. In this article we present a scientific project using the SPECFIND catalogue (Sect. 2). For a significant increase of independent radio cross-identifications an input of sources from more than a hundred radio catalogues is needed. This goal could only be attained by taking advantage of Virtual Observatory (VO) capabilities, which are described in Sect. 3. The new release of the SPECFIND catalogue is presented in Sect. 4.

2. IDENTIFICATION OF GIGAHertz PEAKED SOURCES FROM THE SPECFIND CATALOGUE

A good illustration of the scientific use of the SPECFIND database in VizieR we have searched for sources with peculiar spectra, which are flat or inverted. These sources represent Gigahertz Peaked Sources (GPS) candidates or future mm-VLBI targets. Gigahertz Peaked Sources are supposed to be young radio-loud AGN, and are not well studied (for a review see O’Dea 1998). Existing GPS samples are small (\(N < 200\), see, e.g. Labiano et al. 2007) and an extension of this sample is needed to investigate whether these sources are the precursor of the local FRI and FRII radio galaxies.

We extracted 220 objects with flat or inverted spectra from the SPECFIND catalogue and observed them quasi-simultaneously at 4.85 GHz (6 cm), 10.45 GHz (2.8 cm), and 32.0 GHz (9 mm) with the Effelsberg 100m telescope (Vollmer et al. 2008). Since the SPECFIND spectra are made from multi-epoch data, we can assess in this way
the percentage of variable sources in our sample. These variable sources represent potential candidates for intra-day variable (IDV) sources. Using the VLBA calibrator survey (VCS) we have investigated the parsec-scale morphology of the sources. About 45% of the sources in our sample are classified by us as Gigahertz peaked source (GPS) or High Frequency Peaker (HFP) candidates. We add 65 new GPS/HFP candidates to existing samples. We confirm the expected tendency that HFP are more compact on a milliarcsecond scales than the classical GPS sources, which peak at lower frequencies.

3. THE ROLE OF VO TOOLS FOR THE NEW RELEASE OF THE SPECFINDD CATALOGUE

In the first release of the SPECFINDD catalogue sources from the 20 largest radio catalogues were cross-identified. To increase the number of cross-identification in a significant way, the number of input catalogues has to be increased. Since the number of catalogues with a given number of radio sources increases rapidly with decreasing number of radio sources, the number of input catalogues has to be increased by an order of magnitude to attain this goal. Before the actual cross-identification, the pertinent radio catalogues have to be identified and the information contained in these catalogues has to be extracted in a uniform way. Carrying out these tasks by hand is very time consuming. We therefore decided to develop VO tools to make the catalogue identification and preparation more efficient.

Within the framework of the VO-TECH Design Study CDS has developed three VO tools at CDS: (i) a tool to search for useful radio catalogues in the Virtual Observatory, (ii) a tool to extract relevant information from these catalogues and to uniformize the catalogue information, and (iii) a tool to characterize the data, i.e. to include additional metadata necessary for the full usage of the data in the VO, as required in the VO “Characterization” data model. During the development the tools were kept as general as possible. They can thus be used in other astronomical contexts. We also plan to include the extended radio catalogue description gathered by the characterization tool into VizieR. These tools are available at http://eurovotech.org/twiki/bin/view/VOTech.

3.1. Registry query tool

This tool identifies VO resources based on Unified Content Descriptors (UCDs). The UCDs are a controlled vocabulary defined by the VO to describe astronomical quantities (Derrière et al. 2004). The tool is written in Java and uses XMlDB API to get data from the VO registry of resources. The user specifies a required set of UCDs. The tool searches the VO registry for all catalogues whose descriptions contain these UCDs (upper part of Fig. 1). For example, in our project the minimum set of parameters needed for the radio cross-identification are source coordinates and a radio flux.

The result of the query is a list of relevant radio catalogues (lower part of Fig. 1). The catalogues can then be sorted into useful and not useful catalogues by displaying the catalogue descriptions (VizieR Readme) in a web browser. A workspace permits to save and restore all actions performed on the catalogues. At the end, a final list of relevant catalogues is established.

3.2. Data homogenization tool

These relevant catalogues can be directly loaded into the data homogenization tool (upper part of Fig. 2). The tool creates homogenized data from a heterogeneous set of catalogues. It is written in Java and works on XML tables. In a first step the user specifies a set of columns for the output table. In our case we defined the output columns according to the needs of SPECFINDD (lower part of Fig. 2). In a second step the tool generates an interface where a column of the entry radio catalogue is assigned to a user specified output column (Fig. 3). The user is free to change the input column that he wants to assign to an output column. It is also possible to assign an arithmetic combination of different input columns or conditions on input columns to an output column. As a result the tool generates an ASCII or VOTable output table for each input radio catalogue. The ASCII output tables can be directly used by SPECFINDD.
Figure 2. Data homogenization tool. Upper part: list of uploaded relevant catalogues. Lower part: user specified set of output columns.

Figure 3. Homogenization interface. Each row corresponds to one catalogue. Each column is predefined by the user. A column of the input catalogue is assigned to a predefined output column (Fig. 2). The user can then accept or change this assignment.

Figure 4. Characterization tool. The user specifies the metadata necessary for the full usage of the data in the VO, as required in the VO “Characterization” data model.

3.3. Characterization tool - CAMEA

We realized that the description of the radio catalogues does not contain all necessary information for the cross-identification. Basic information such as the identity of the instrument, frequency, resolution, and observation dates are not included in the catalogue metadata. We therefore decided to develop a third VO tool which permits to specify this missing information. More generally, it will permit to create a full description of a VO resource based on the VO characterization data model (Louys et al. 2008; Fig. 4). In the future, this tool will serve as input for the data homogenization tool described above.

4. SPECFIND V2.0

The use of the registry query tool and data homogenization tool enabled us to include 105 radio catalogues from VizieR into the SPECFIND radio cross-identification tool. The number of sources from these catalogues is $3.76 \times 10^6$. Compared to the first release of the SPECFIND catalogue this is an increase of available radio sources by $\sim 8\%$. This relatively small increase is due to the fact that the number of catalogues with a given number of radio sources increases rapidly with decreasing number of radio sources. However, the smaller catalogues often provide the missing third flux density to establish a radio spectrum. For example, in the northern hemisphere there is a multitude of radio objects.
Figure 5. Sky coverage of radio sources. Upper panel: from the first release; lower panel: from the SPECFIND V2.0 catalogue.

with available NVSS (1.4 GHz) and WENSS (325 MHz) flux densities. The surveys at higher frequencies, which are included in SPECFIND, are rather shallow and thus did not detect the majority of the sources. Observations at high frequencies underlying small catalogues are almost always more sensitive than observations which give rise to large catalogues, with the drawback that they are made within small areas on the sky. This is the reason why a modest increase of source available for the cross-identification (∼ 8%) leads to a significant increase of cross-identified radio objects of ∼ 25%. The source coverage of the first and the second release of the SPECFIND catalogue are shown in Fig. 5. The SPECFIND V2.0 catalogue will be soon available via Vizier at CDS.

REFERENCES


