

## Research Article

# The XinJiang Astronomical Observatory NSRT Pulsar Data Archive

Hailong Zhang ,<sup>1,2</sup> Markus Demleitner,<sup>3</sup> Jie Wang,<sup>1</sup> Na Wang,<sup>1,2</sup> Jun Nie,<sup>1,2</sup> and Jianping Yuan<sup>1,2</sup>

<sup>1</sup>Xinjiang Astronomical Observatory, Chinese Academy of Sciences, Urumqi 830011, China

<sup>2</sup>Key Laboratory of Radio Astronomy, Chinese Academy of Sciences, West Beijing Road, Nanjing, Jiangsu 210008, China

<sup>3</sup>Heidelberg University, Zentrum für Astronomie, Mönchhof St. 12-14, 69120 Heidelberg, Germany

Correspondence should be addressed to Hailong Zhang; [zhanghailong@xao.ac.cn](mailto:zhanghailong@xao.ac.cn)

Received 31 October 2018; Accepted 30 December 2018; Published 23 January 2019

Guest Editor: Junhui Fan

Copyright © 2019 Hailong Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The sensitivity of our pulsar observation system with typical integration time of 16 minutes has reached 0.4 mJy after we have upgraded the receiver cooling system of NanShan 25-m radio telescope (NSRT) at Xinjiang Astronomical Observatory (XAO) in 2002. About 280 pulsars were observed after the upgrade. The NanShan station pulsar data is transmitted to the data center at XAO headquarters via a 300Mb/s dedicated data link for long-term storage, and the remote backup is realized by synchronous data transmission between NanShan station and XAO headquarters. Metadata extraction, archiving, and releasing are completed by the data release servers which are located in XAO headquarters. At present, more than 84000 pulsar observation data files in the PSRFITS format have been released with a maximum file size of 1GB. XAO pulsar data online service provides cone search and multiobject constraint query. We have published our data based on Virtual Observatory (VO) standards. The online pulsar data service is registered in the International Virtual Observatory Alliance's Registry. Users can access the data through common VO tools.

## 1. Introduction

A Pulsar [1] is a highly magnetized rotating neutron star that emits a beam of electromagnetic radiation. It is the product of massive star evolution and has the characteristics of small volume, large density, high-speed rotation, and strong magnetic field. Pulsar observation research is helpful to reveal the mystery of the birth and evolution of the universe and has very important scientific significance to promote the development of astrophysics, particle physics, and spacecraft navigation. It has always been a hot spot in the field of astronomy and astrophysics. Pulsar observation and theoretical research have made a rapid progress after Jocelyn Bell discovered the first radio pulsar in 1967 [2]. By October 2018, 2658 pulsars have been confirmed (<http://www.atnf.csiro.au/research/pulsar/psrcat/>).

The NanShan 25-m Radio Telescope [3] (latitude  $43^{\circ}28'47''$  N, longitude  $87^{\circ}10'42''$  E) has a major upgrade in 2014, after that the antenna diameter is now 26 meters. NSRT

is an important member of VLBI network in China. It has carried out regular pulsar timing observation since Jan 2010. NSRT has a Declination range from  $-40^{\circ}$  to  $90^{\circ}$  and became an important base for monitoring pulsar timing behavior. How to store, archive, and release the ever-increasing pulsar data is a vital issue in operating the system and explore the data scientific value.

Pulsar data taken by the Parkes Radio Telescope (<https://www.parkes.atnf.csiro.au/>) (Parkes) as far back as the early 1990s is archived for long-term storage in the CSIRO's Data Access Portal (DAP) (<https://data.csiro.au/dap/>), in CSIRO's data center, Canberra, Australia. The CSIRO DAP allows querying and download of pulsar data by any user; however some data are subject to an embargo period. In this case, members of a science team can log in with their account to access their data during the embargo period. Parkes pulsar data archive [4] can be visualized and manipulated with Virtual Observatory (VO) Services. The NSRT pulsar data retrieval service provides similar functionality to Parkes.

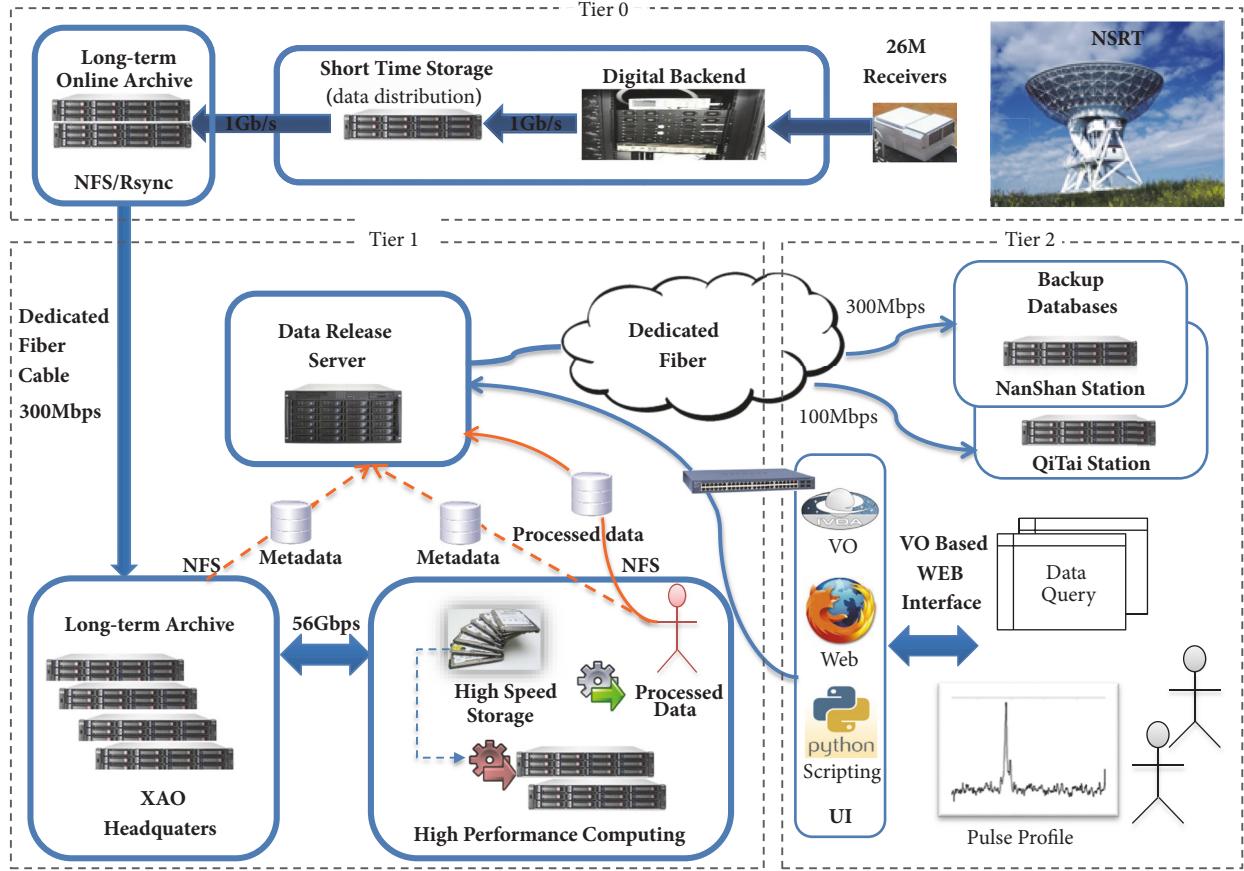


FIGURE 1: Overall design of NSRT pulsar data management system.

## 2. Pulsar Observing System

The pulsar de-dispersion observation system was set up at the end of 1999, which consists of a two-channel room-temperature receiver at L band, a down converter, a multichannel filterbank, and a digitizer. The receiver has a center frequency of 1540 MHz with bandwidth of 320 MHz, and the de-dispersion is implemented by a 2x128x2.5 MHz multichannel analog filterbank. Analogue signals from the filterbank are 1-bit digitized. After sampling, the raw data of each channel are folded at the apparent pulsar period to form subintegrations of 60 s which are stored in the “Timer” format [5]. The observation is started by 1-s pulse and the start time is checked using the 5-s pulse derived from the Observatory H-maser. A GPS time-transfer system is employed to align the Observatory clock with UTC. The pulsar timing program started since January 2000 and 74 strong pulsars [6] with flux density greater than 4mJy were detected.

In July 2002, an L band dual-polarization cryogenic receiver was used for pulsar observations. The noise temperature of the receiver was less than 10 K, the noise temperature of the system reached 20 K and 22 K, respectively, and the observation sensitivity reached 0.4 mJy [7] with integration time of 16 minutes. The sensitivity enables NSRT to observe approximately 280 pulsars.

In January 2010, the Digital FilterBank (DFB) system [8], which was developed by Australia Telescope National Facility

(ATNF)(<https://www.atnf.csiro.au/>), was commissioned and put into operation. It has higher timing resolution which makes it detect 10 millisecond pulsars. The observation system has a calibration probe in the feed and receives orthogonal linear polarizations. A pulsed noise source is used to inject a linearly polarized calibration signal which is employed to calibrate the flux and polarization of pulsar. Absolute timing is provided by the Peripheral Component Interconnect (PCI) based Australia Telescope Distributed Clock (ATDC). The ATDC accepts a 5MHz maser signal which provides the base frequency reference for all of the DFB timing. The IPPS signal, which is generated from the 5MHz input, is used to determine the absolute time. For each observation, the ephemeris from PSRCAT (<http://www.atnf.csiro.au/people/pulsar/psrcat/>) is used online for folding the signal with integration time of 30 s, where the total integration time is 4 to 16 minutes with a cadence of about 10 days. The DFB data format is PSRFITS (a file format customized for pulsar data storage, based on FITS format), and PSRCHIVE [9] software is used to analyze offline data, including the RFI mitigation, de-dispersion, fold, and ephemeris update.

## 3. Overall Design of Pulsar Data Management System

The overall design of NSRT pulsar data management system is shown in Figure 1. The design consists of three tiers, where

tier 0 is responsible for pulsar observation data acquisition and online archiving, tier 1 manages remote backup of original observation data and the metadata extraction, and tier 2 handles the publication and query facilities of the pulsar data.

*Tier 0.* NSRT data acquisition and online archiving were completed in NanShan Station. Tier 0 consists of telescope, receiver, digital backend, data temporary storage, and archiving. Pulsar data recording and preprocessing happen entirely within the telescope receiving system. The data capture program will perform the data preprocessing, de-dispersion, pulse period calculation, period folding, data storage, observation outline query, graphic output, antenna attitude control, and so on in real time. The typical sampling rate in the observation is 1 ms, and the time of each observation is determined by pulsar's flow intensity in the observation frequency band; generally an observation lasts from 4 to 16 minutes. All data will be preprocessed and written into a temporary storage server and long-term archived after validity check.

*Tier 1.* Pulsar observation data is synchronized between the storage system of NanShan station and the XAO headquarters through a dedicated 300Mbps line. Currently, rsync (<https://rsync.samba.org/>) is used to perform incremental backups from NanShan station to XAO headquarters. All the metadata will be imported into the corresponding database to prepare for data release after extracting the metadata information from pulsar data. The data storage server, the Taurus high performance computing system, and the data release server use a common NFS server to implement the data transmission. For database data, there are backups in XAO headquarters and NanShan station. Users are allowed to log in to Taurus analysis system to download and process the data. A 56 Gbps InfiniBand switch connects the Taurus HPC and the long-term archive in XAO headquarters. All the data processed by users can be archived and released as needed.

*Tier 2.* XAO uses GAVO's DaCHS [10] server as the main data release framework to implement metadata operations and the data release based on VO standards. Users are allowed to access the online data retrieval platform through web browsers, standard VO tools, various scripts, etc. At present, the data release system supports cone search and multiconstraint target retrieval based on source name, observation date, observation ID, etc. Users can customize the limited amount of data output, output fields, sorting, and other contents. Currently the pulsar data retrieval service supports HTML (data is returned in your web browser; you can select additional columns in your output from an input field that pops down when you mouse over it), VOTable (<https://wiki.ivoa.net/twiki/bin/view/IVOA/IvoaVOTable>) (data is returned in IVOA's standard data format, the VOTable), FITS Table (the data is in the first extension; this contains much less meta information than a VOTable of the same data and thus should only be used if your backend tools do not understand VOTables), CSV (Comma Separated

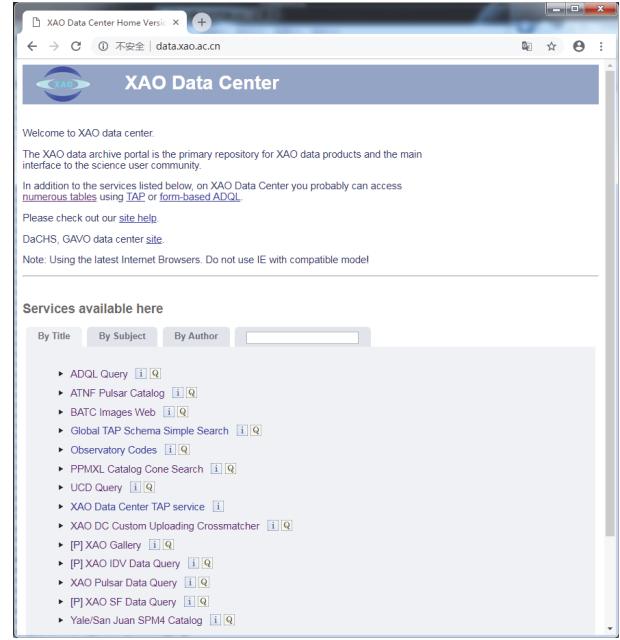


FIGURE 2: The main page of XAO data archive portal.

Values: this format carries almost no metadata as well, but it is understood by many database programs, spreadsheets, etc.), tar (users can download all matching items in a tar file), and other data output formats.

## 4. XAO Data Center and Pulsar Data Retrieval

The XAO data archive portal is the primary repository for XAO data products and the main interface to the science user community. In XAO Data Center (URL: <http://data.xao.ac.cn>), the interface is shown in Figure 2. Currently, several services such as ADQL (<https://wiki.ivoa.net/internal/IVOA/ADQL/WD-ADQL-2.1-20171129.pdf>) query, PPMXL [11] Catalog Cone Search, XAO DC Custom Uploading Cross-matcher, and XAO Pulsar Data Query, etc. have been released in the data center. The released data resources include Pulsar data, Active Galactic Nuclei data, and Molecular Line data.

*4.1. Web Retrieval.* The link of XAO pulsar data query page is <http://data.xao.ac.cn/pul/pulsar/q/form>. Users can also open the page through the “XAO Pulsar Data Query” link in the data center. A screenshot of the retrieval page is shown in Figure 3. At present, 84078 records of pulsar data have been archived with a maximum file size of 1GB. The data format is mainly PSRFITS, and subsequent data processing can be completed through PSRCHIVE.

### The Query Interface:

(1) “Position/Name”: “Position/Name”: it is mainly used to input coordinates, the Right Ascension and Declination information of pulsars, or position information of SIMBAD-resolvable object, for example, RA ~ 278.4, Dec ~ -3.6, M31, etc.

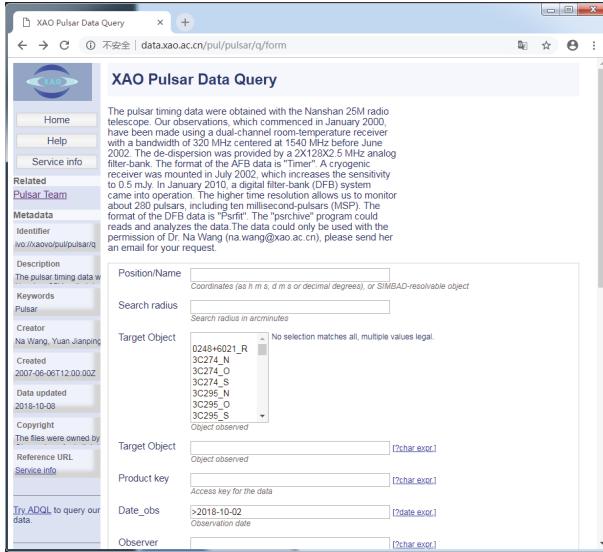


FIGURE 3: XAO pulsar data retrieval interface.

(2) “Search radius”: Search radius is according to the information in “Position/Name” and used to define the scope of data retrieval, for example, SR ~ 6; the unit of Search radius is arcminutes.

(3) “Target Object”: “Target Object”: it can let the user directly select or enter a pulsar name in the input form, for example, Target Object ~ J2157+4017, or ~ \*j2157\*, and other related information; the ~ sign means contains and \* sign stands for all characters. The \_R in target objects represents calibration file.

(4) “Date\_obs”: observation time. Enter “> year-month-day” which means query the data after a certain day of a certain year. For example, “<2018-09-20” means select values earlier than September 20th, 2018; “year-month-day +/- num” means that select values num days around year-month-day. For example, “2018-09-20 +/- 3” means select 6 days from September 17th, 2018, to September 23rd, 2018.

(5) “Observation Frequency”: observation frequency is an important factor affecting observation of pulsars, which has certain influence on pulsar’s dispersion effect, scattering effect and spectrum, and can be filled according to actual needs; the value type is numeric.

(6) “Output format”: “Output format” includes HTML (default), Text (fixed columns), tar, Text (with Tabs), JSON, VOTable, FITS table, and CSV.

Taking J0534+2200 as an example, 569 items of relevant data were obtained with the conditions: observation frequency greater than 1000MHz, observation bandwidth greater than 500 MHz, and the Date-obs after August 13<sup>th</sup>, 2017; the file name includes “\*.rf”. For example, the pulsar data is on August 14<sup>th</sup>, 2017, the recording time is 2017-08-14T07:11:59Z, and the data is named “x170814\_074136.rf”. The search results are shown in Figure 4. Click “[Preview (pav-DFTp)]” to view pulsar pulse profile (Figure 5).

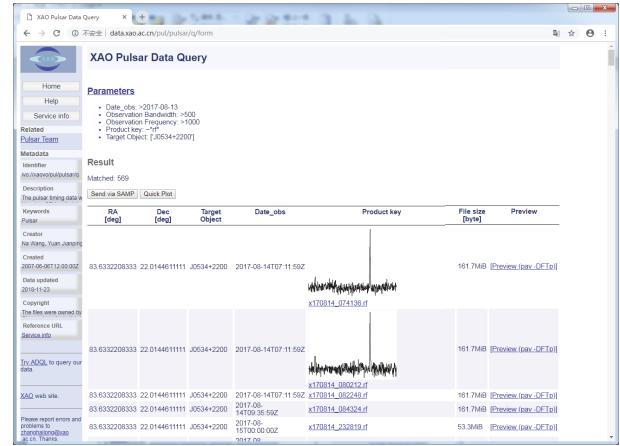


FIGURE 4: Observational data retrieval results of J0534+2200.

Freq: 1556.250 MHz Bw: 512 Length: 1199.959 S/N: 15.349

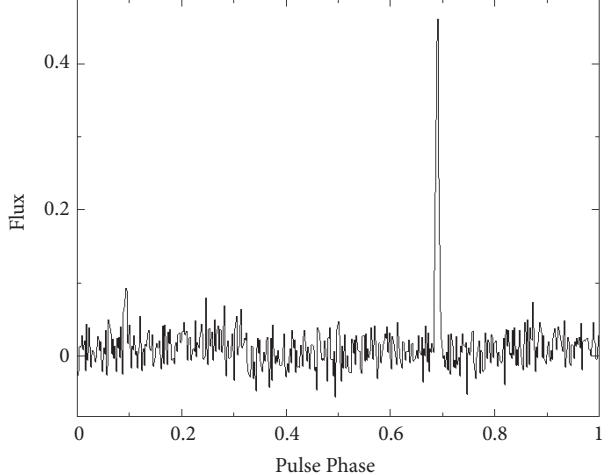


FIGURE 5: Pulse profile of PSR J0534+2200 (file name: x170814\_074136.rf, the flux is not calibrated.)

**4.2. VO Tools Retrieval.** The VO Registry provides a mechanism through which VO applications can discover and select resources related to specific scientific issues. The XAO has completed the registration of pulsar data services.

(1) **TOPCAT** (<http://www.star.bris.ac.uk/~mbt/topcat/>). TOPCAT is an interactive graphic editor for tabular data, which is convenient for astronomers to analyze astronomical data. Thanks to the use of VO standards, it can work smoothly with other tools, service systems and data sets inside and outside VO. It has been widely used in Starlink, AstroGrid, VOTech, AIDA, GAVO, GENIUS, DPAC, and other projects.

(2) **TOPCAT Retrieval.** Start TOPCAT and searching for the keyword “xaovo” through Cone Search in VO menu can find 4 XAO registered services including XAO Pulsar Data Query, PPMXL Catalog Cone Search, etc.

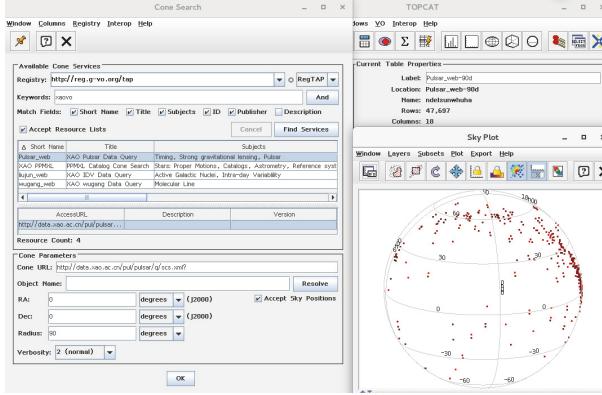


FIGURE 6: VO tool retrieval results.

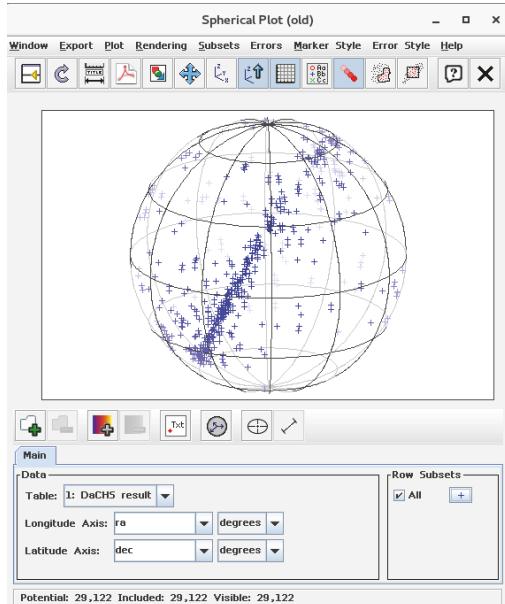


FIGURE 7: TOPCAT visualization results.

As shown in Figure 6, this example selects XAO Pulsar Data Query service to automatically generate Cone URL “<http://data.xao.ac.cn/pul/pulsar/q/scs.xml?>” to retrieve RA~0, Dec~0, Radius~90, Verbosity~3(maximum). 47697 items of relevant data can be retrieved (the left panel on Figure 6). Among them, “J2000” with the year 2000 epochs (named after the letter “B” with year 1950 epochs).

**4.3. Pulsar Data Visualization.** Through TOPCAT, the retrieved pulsar data can be visualized: turn on TOPCAT and select “Send via SAMP” in pulsar retrieval result interface to transfer the data to TOPCAT for visualization. In this example, TOPCAT version 4.6-1 was used to visualize 29122 results with Date-obs: < 2014-02-01 and Product key: “`*.rf`” (PULSAR data format generated by DFB3 system after January 2010); as shown in Figure 7, the data is marked in 3D galactic coordinates by Spherical Plot.

**4.4. Pulsar Data Access.** Since the launch of the pulsar data release server in 2016, the data access statistical information is shown in Figure 8. The left part of Figure 8 shows the total number of accesses with an average of about 10,000 accesses per month, and the right part of Figure 8 shows the number of hosts accessed with an average of nearly 500 hosts accessing the data server per month.

## 5. Conclusion

The XAO data archiving and release system is available to worldwide researchers. Data release is based on VO protocols. Users can retrieve and access data through browsers, various scripts, and standard VO tools. The pulsar data service has been registered to the IVOA.

## Data Availability

All the pulsar data generated by XAO 25m radio telescope have been released in XAO data center, and the way to query the data was described in the article. The address of our data center is <http://data.xao.ac.cn>. Astronomers can repeat our research or query easily.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

The authors gratefully acknowledge the support of the National Natural Science Fund of China (11873082, U1531125, 11803080, and 11503075), National Key Basic Research Program of China, 973 Program 2015CB857100, National Key Basic Research and Development Program 2018YFA0404704, and Youth Innovation Promotion Association CAS. The research is partly supported by the Operation, Maintenance and Upgrading Fund for Astronomical Telescopes and Facility Instruments, budgeted from the Ministry of Finance of China (MOF) and administrated by the Chinese Academy of Sciences (CAS). The algorithm and debugging work has

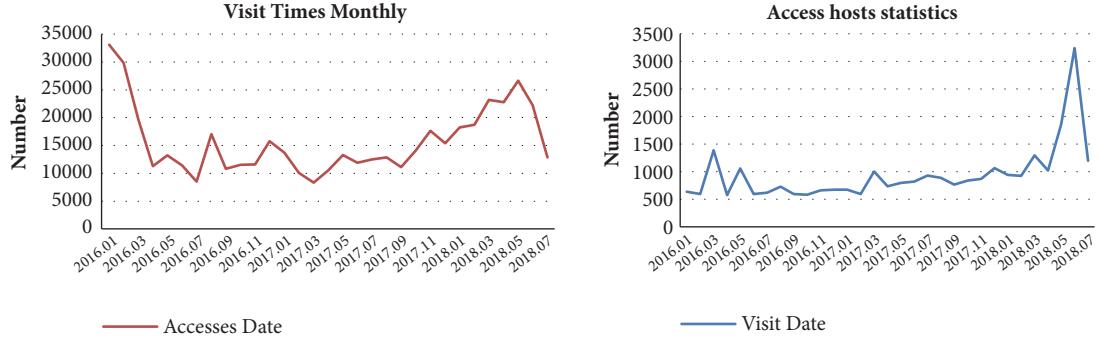


FIGURE 8: NSRT pulsar data service accessing statistics.

applied Taurus High Performance Computing Cluster of Xinjiang Astronomical Observatory, CAS.

## References

- [1] R. D. Lorimer and M. Kramer, *Handbook of Pulsar Astronomy*, Cambridge university press, 2005.
- [2] Z. Weizhen, *Observation and Study of Pulsar Timing and Glitch*, Graduate University of the Chinese Academy of Sciences, 2009.
- [3] Z. Hailong, Z. Yan, N. Jun et al., "Xinjiang astronomical observatory NSRT data storage system," *Astronomical Research & Technology*, vol. 15, no. 2, pp. 181–187, 2018.
- [4] G. Hobbs, D. Miller, R. N. Manchester et al., "The parkes observatory pulsar data archive," *Publications of the Astronomical Society of Australia*, vol. 28, no. 3, pp. 202–214, 2011.
- [5] N. Wang, R. N. Manchester, J. Zhang et al., "Pulsar timing at urumqi astronomical observatory: observing system and results," *Monthly Notices of the Royal Astronomical Society*, vol. 328, no. 3, pp. 855–866, 2001.
- [6] J. Zhang and N. Wang, "Single dish astrophysics observation at urumqi observatory," *Progress of Astronomy*, vol. 18, no. 4, pp. 271–282, 2000.
- [7] N. Wang, *Observation and Study of Pulsar Timing and Interstellar Scintillation*, Peking University, 2001.
- [8] G. Hampson and B. Andrew, *A 1GHz Pulsar Digital Filter Bank and RFI Mitigation System*, 2008, <http://www.jb.man.ac.uk/pulsar/observing/DFB.pdf>.
- [9] A. W. Hotan, W. Van Straten, and R. N. Manchester, "PSRCHIVE and PSRFITS: An open approach to radio pulsar data storage and analysis," *PASA - Publications of the Astronomical Society of Australia*, vol. 21, no. 3, pp. 302–309, 2004.
- [10] M. Demleitner, M. C. Neves, F. Rothmaier, and J. Wambsganss, "Virtual observatory publishing with DaCHS," *Astronomy and Computing*, vol. 7-8, pp. 27–36, 2014.
- [11] S. Roeser, M. Demleitner, and E. Schilbach, "The ppmxl catalog of positions and proper motions on the ICRS. Combining USNO-B1.0 and the two micron all sky survey (2MASS)," *The Astronomical Journal*, vol. 139, no. 6, pp. 2440–2447, 2010.

