

О кросс-идентификации объектов для планирования наблюдений на наземных и космических обсерваториях



Ладейщиков Дмитрий Антонович (УрФУ)

Постановка задачи:

необходим инструмент для кросс-идентификации большого количества каталогов с разной плотностью источников, различным объёмом (от 500 Кб до ~50 Гб) и в различных спектральных диапазонах.



Решение:

- 1) Объединить все искомые каталоги в одну таблицу, которая содержит лишь координаты источников
- 2) С помощью алгоритма DBSCAN произвести поиск групп
- 3) Каждой группе присвоить свой идентификатор (напр.: G12.5323+67.8732) в соответствии со средними координатами всех источников в группе
- 4) Все источники из различных каталогов, входящие в одну группу, приобретают один внутренний идентификатор

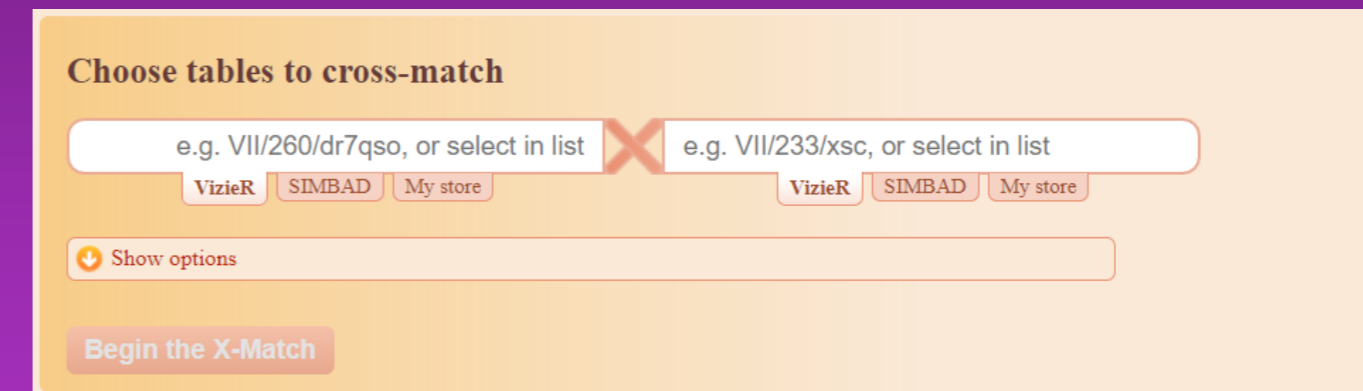


Результат:

Мы имеем возможность связать источники из различных каталогов и изучать их параметры взаимосвязанно (корреляция, двумерные диаграммы и т.д.)

Существующее решение:

Сервис **CDS x-match** позволяет эффективно кросс-идентифицировать два различных каталога



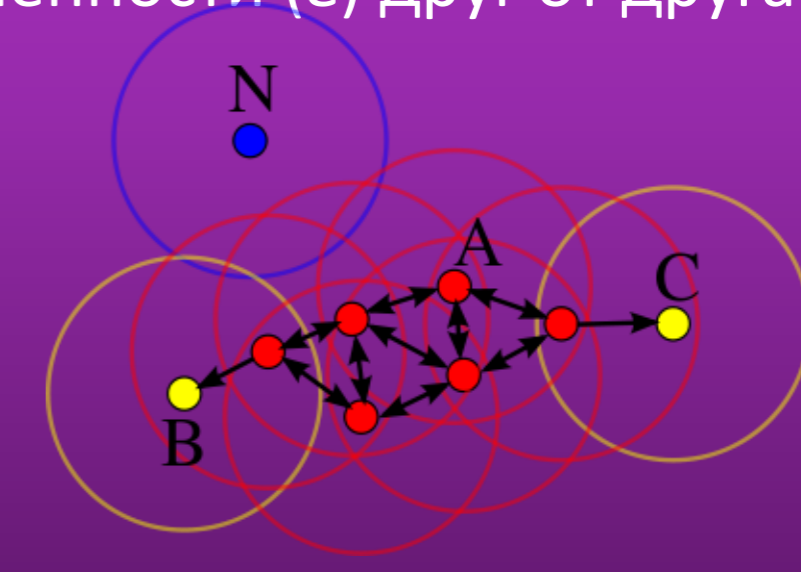
Но данная система не позволяет связывать более двух различных таблиц:

* You can not make query on more than two tables.



Алгоритм DBSCAN (Ester et al., 1996)

Позволяет эффективно искать группы источников на основании их взаимной удалённости (ϵ) друг от друга.



Оптимальное значение порога (ϵ) выбирается в зависимости от решаемой задачи и спектрального диапазона. Пример:

Для УФ диапазона: $\epsilon \sim 5$ уг. сек.

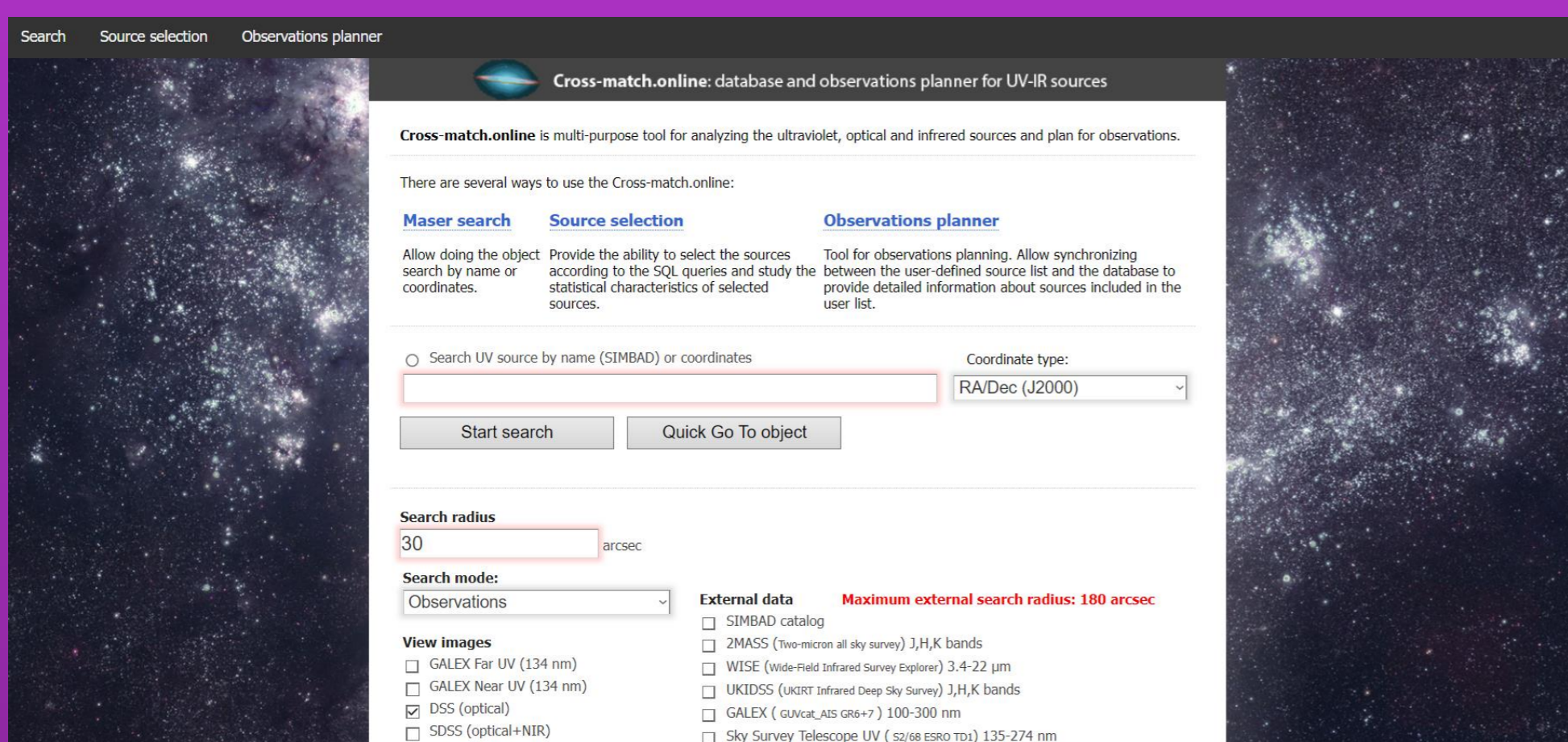
Для ММ диапазона: $\epsilon \sim 15-30$ уг. сек.

Для радио диапазона: $\epsilon \sim 30-60$ уг. сек.

С использованием данного метода автором в настоящее время разрабатываются несколько инструментов для подготовки и планирования наблюдений, в том числе на наземных и космических обсерваториях

В ультрафиолетовом, оптическом и ближнем инфракрасном диапазонах

<http://cross-match.online>

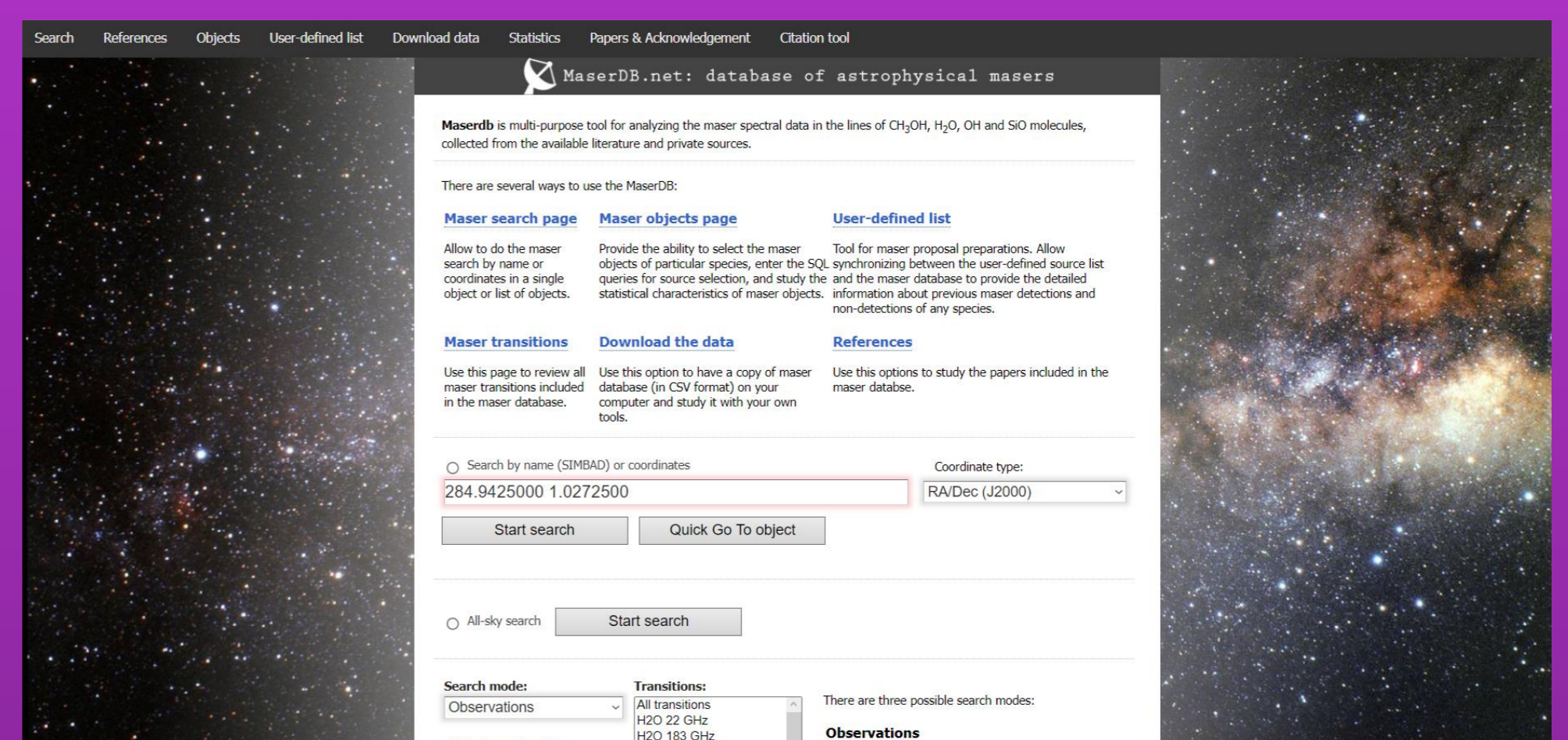


Данная система в своей основе является базой данных ультрафиолетовых источников.

- Позволяет автоматически кросс-идентифицировать следующие популярные УФ каталоги: GALEX, SWIFT/XMM-Newton, MSX
- В перспективе планируется добавление иных каталогов, в том числе: 2MASS, UKIDSS, MSX (IR), WISE и др.

В дальнем инфракрасном, миллиметровом и радио диапазонах

<http://maserdb.net>



Данная система в своей основе является базой данных мазерных источников, но данные в других диапазонах выводятся автоматически.

Публикации:

- Ladeyschikov D.A., Sobolev A.M. et al. The online Multi-wavelength database of water masers in star formation regions // The Astronomical Journal. – 2022 (accepted).
- Ladeyschikov D.A., Urquhart J.S., Sobolev A.M. et al. The Physical Parameters of Clumps Associated with Class I Methanol Masers // The Astronomical Journal. – 2020. – V. 160. – № 5. – P. 213-224.
- Ladeyschikov D.A., Bayandina O.S., Sobolev A.M. Online Database of Class I Methanol Masers // The Astronomical Journal. – 2019. – V. 158. – № 6. – P. 233-247.

Преимущества выбранного подхода:

- Кросс-идентификация каталогов производится заранее на сервере, поэтому каждый пользователь избавлен от необходимости выполнять её заново.
- Возможность использования языка SQL для отбора источников по значениям потока, координат и т.д.
- Возможность оперативного получения

данных (таблицы, изображения) об источниках в различных спектральных диапазонах

- Возможна работа как с отдельными источниками, так и со списками или большими выборками источников.
- Возможность оперативного добавления новых каталогов
- Быстрый поиск данных

Недостатки выбранного подхода:

- Огромный объем некоторых каталогов (к примеру, GAIA) не позволяет включить их полностью в локальную базу данных.
- Высокая сложность разработки исходных кодов для эффективной работы с большими объемами данных (более 500 Мб)