|  |  |
| --- | --- |
| Gerb-BMSTU_01 | **Министерство науки и высшего образования Российской Федерации**  **Федеральное государственное бюджетное образовательное учреждение**  **высшего образования**  **«Московский государственный технический университет**  **имени Н.Э. Баумана**  **(национальный исследовательский университет)»**  **(МГТУ им. Н.Э. Баумана)** |

ФАКУЛЬТЕТ «Информатика и системы управления»\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

КАФЕДРА «Программное обеспечение ЭВМ и информационные технологии»\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Лабораторная работа №6**

**по курсу «Проектирование рекомендательных систем»**

|  |  |
| --- | --- |
| **Тема: «Сравнение алгоритмов выделения сообществ»**  **Студент Горячев В. Г.**  **Группа ИУ7-33М**  **Преподаватель Быстрицкая А. Ю.** |  |

Москва

2023 г

## Описание алгоритмов

***Лувенский метод***

Алгоритм иерархической кластеризации, который обнаруживает сообщества в больших сетях. Метод оценивает, насколько плотно связаны элементы в сообществе, и сравнивает результат с тем, насколько связаны элементы в случайной сети.

***Fast-Greedy Modularity Optimization***

Жадный алгоритм, который, считая изначально все узлы уникальными, по метрике объединяют их в пары, пока это возможно.

***Label Propagation Algorithm***

Узлы получают метку, обозначающую сообщество, к которому они принадлежат. Членство в сообществе меняется в зависимости от меток, которыми обладают соседние узлы; метки распространяются до тех пор, пока это возможно, в результате чего выделяются сообщества.

***Библиотеки***

В качестве языка программирования был выбран язык Python вместе с интерактивной средой Jupyter Notebook, поскольку они предоставляют удобный инструментарий для исследования, в частности, для выполнения лабораторных работ. Это определило выбор библиотек — нужно было найти совместимые с языком программирования.

В качестве источника алгоритмов использовалась библиотека по работе с графами, содержащая инструменты для работы с графовыми алгоритмами – networkx.

***Сравнение***

Поскольку в этой работе используются инструменты визуализации, результаты работы алгоритмов можно наглядно представить в виде изображений (на двух разных графах). Также было измерено время работы (7 запусков, 10 повторов), представленное в таблице 1.

Таблица 1 — Время работы алгоритмов

|  |  |  |  |
| --- | --- | --- | --- |
|  | Label Propagation, мс | Лувен, мс | Fast Greedy, мс |
| Малый граф | 0,499 ± 0,035.9 | 1,77 ± 104 | 4,68 ± 0,045 |
| Большой граф | 0,423 ± 0,187 | 170 ± 2,62 | 1770 ± 3,82 |

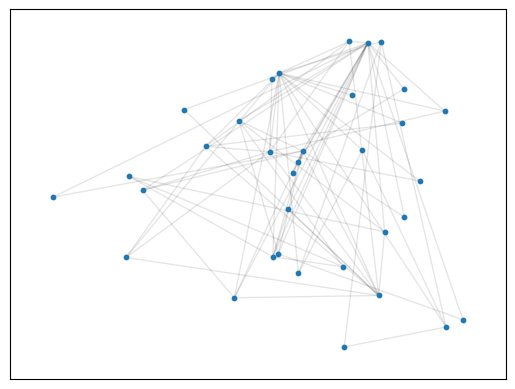


Рисунок 1 — Исходный малый граф

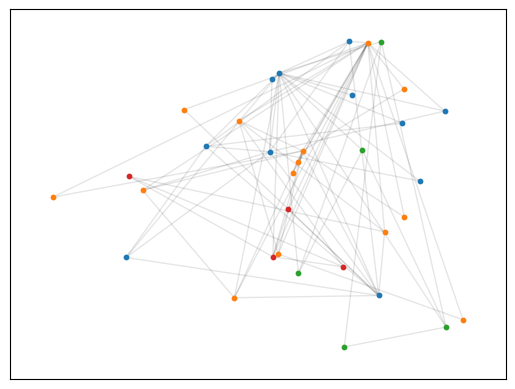


Рисунок 2 — Лувенский алгоритм

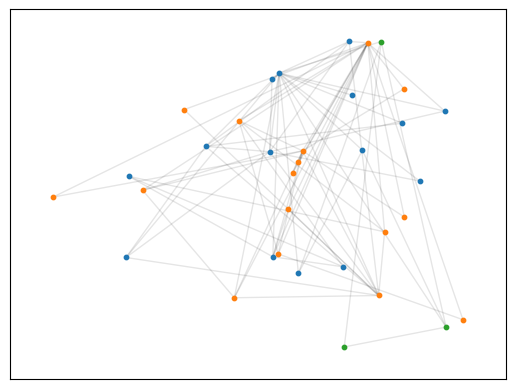


Рисунок 3 — Распространение меток

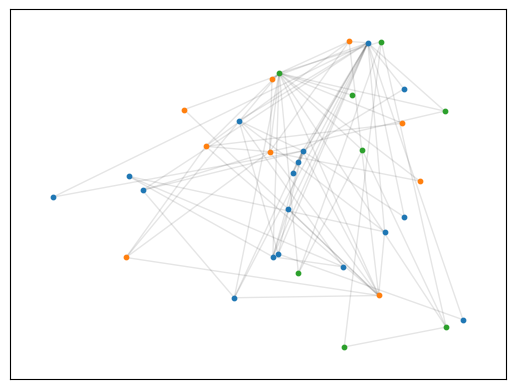


Рисунок 4 — Быстрый жадный…

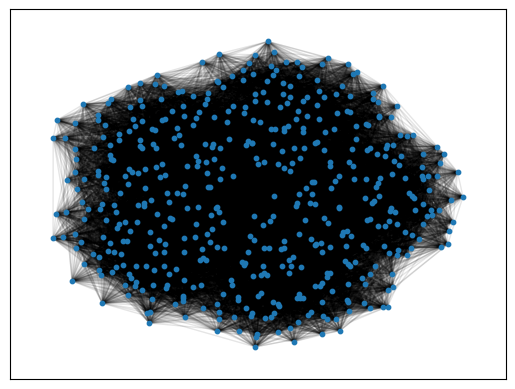


Рисунок 5 — Исходный большой граф

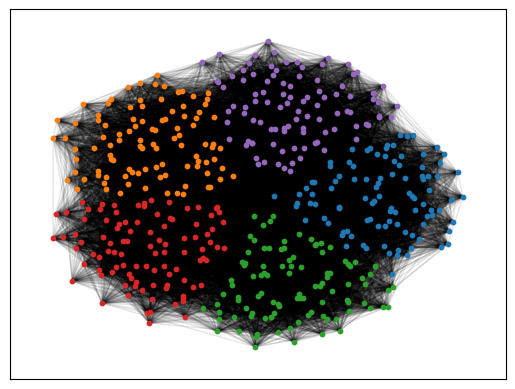


Рисунок 6 — Лувенский алгоритм

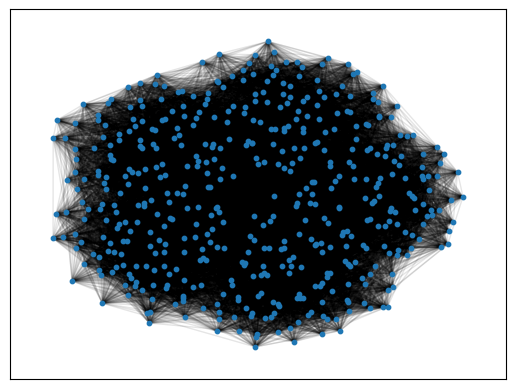


Рисунок 7 — Распространение меток. Не справился с задачей

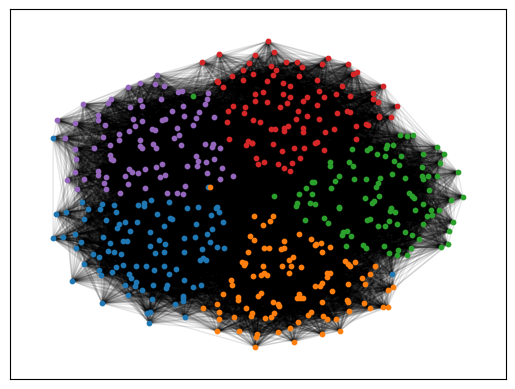


Рисунок 8 — Быстрый жадный…

**Вывод**

В ходе лабораторной работы были сравнены алгоритмы выделения сообществ в графах. Хуже всего себя показал алгоритм распространения меток, поскольку в принципе не справился с выделением сообществ во втором графе (учитывая, что другие алгоритмы справились – возможно, проблема в реализации или в особенностях самого графа), хотя по времени работы был впереди. Быстрый жадный алгоритм работает очень долго на больших графах и, к тому же, метки как-то странно перемешиваются (хоть и в небольших количествах). Поэтому лучшим среди рассмотренных является Лувенский алгоритм – как по результату, так и по времени работы.