

Санкт–Петербургский Государственный Университет

**Поляков Иван Михайлович**

**Отчёт по домашнему заданию №1**

*«Анализ сети»*

*«Метрики узлов сети»*

*«Роли вершин в сети»*

Направление 01.04.02: «Прикладная математика и информатика»  
Образовательная программа ВМ.5505.2021: «Математическое и информационное  
обеспечение экономической деятельности»

Руководитель:  
кандидат физ.-мат. наук,  
доцент Воронкова Ева Боруховна

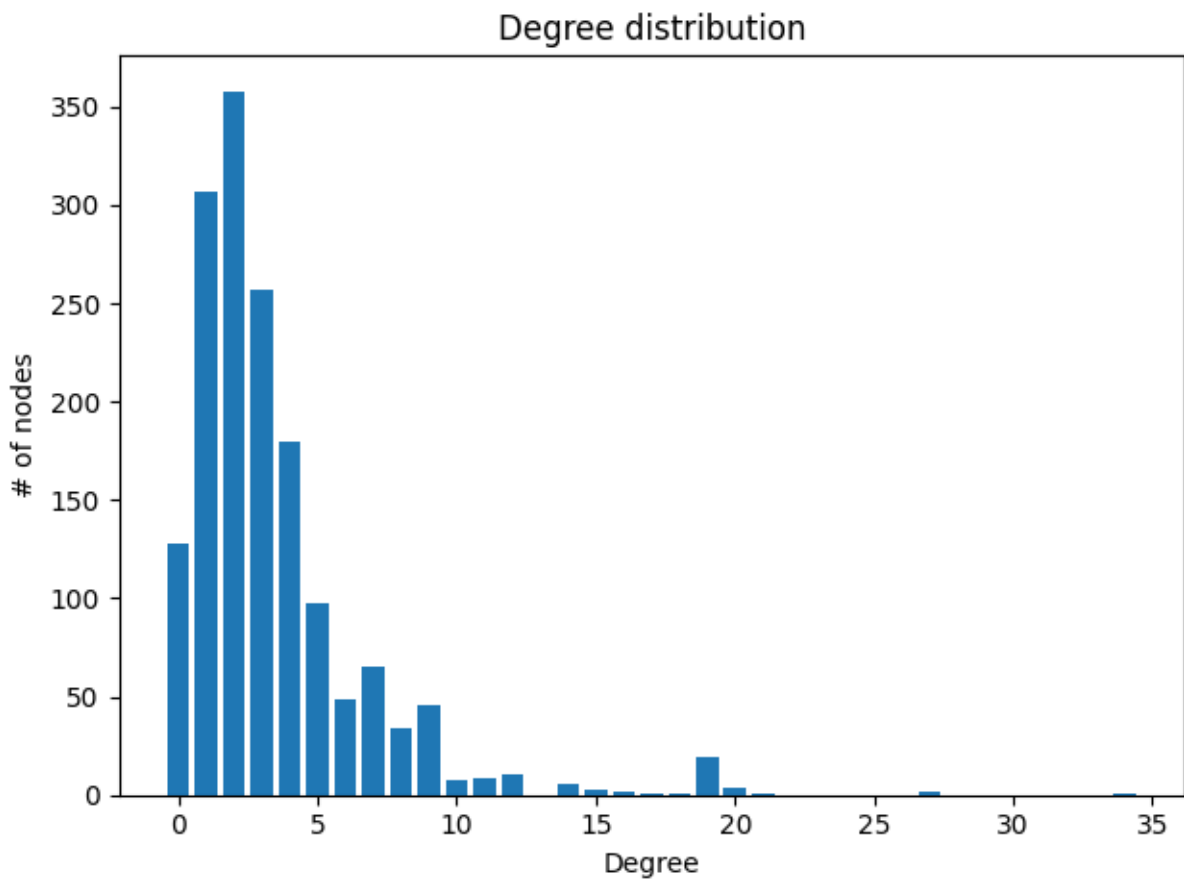
Санкт-Петербург  
2022 г.

## Анализ сети

В данной части работы использовалась сеть научных цитирований *netscience.gml*. В соответствии с заданием, были найдены следующие основные характеристики сети:

- Число узлов: 1589;
- Число рёбер: 2742;
- Средняя степень вершин: 3.541;
- Число вершин степени, большей средней: 539;
- Число компонент связности: 396;
- Число вершин, принадлежащих наибольшей компоненте связности: 379 (23.9%);
- Число рёбер, принадлежащих наибольшей компоненте связности: 914 (33.3%);

Также была построена гистограмма распределения степеней вершин:



Для наибольшей компоненты связности графа дополнительно были вычислены:

- Радиус: 9;
- Диаметр: 17;
- 90-перцентиль расстояния между вершинами: 9.0;

# Метрики узлов сети

В данной части работы использовалась сеть друзей ВКонтакте.

Для наибольшей компоненты связности были рассчитаны следующие метрики, а также выведены первые 10 наибольших их значений:

- Центральность по степени

- 0.5484
- 0.5161
- 0.4839
- 0.4516
- 0.3548
- 0.3226
- 0.3226
- 0.2903
- 0.2903
- 0.2903

- Центральность по близости

- 0.6327
- 0.6200
- 0.5849
- 0.5741
- 0.5636
- 0.5536
- 0.5345
- 0.5345
- 0.5254
- 0.5167

- Центральность по посредничеству

- 0.2146
- 0.1482
- 0.1393
- 0.1240
- 0.1173
- 0.0668
- 0.0619
- 0.0567
- 0.0517
- 0.0468

- Центральность по собственному вектору

- 0.3564
- 0.3411
- 0.3346
- 0.3207
- 0.2765
- 0.2414
- 0.2395
- 0.2239
- 0.2099
- 0.2093

Ниже приведена визуализация каждой метрики и её распределение значений на данном подграфе.

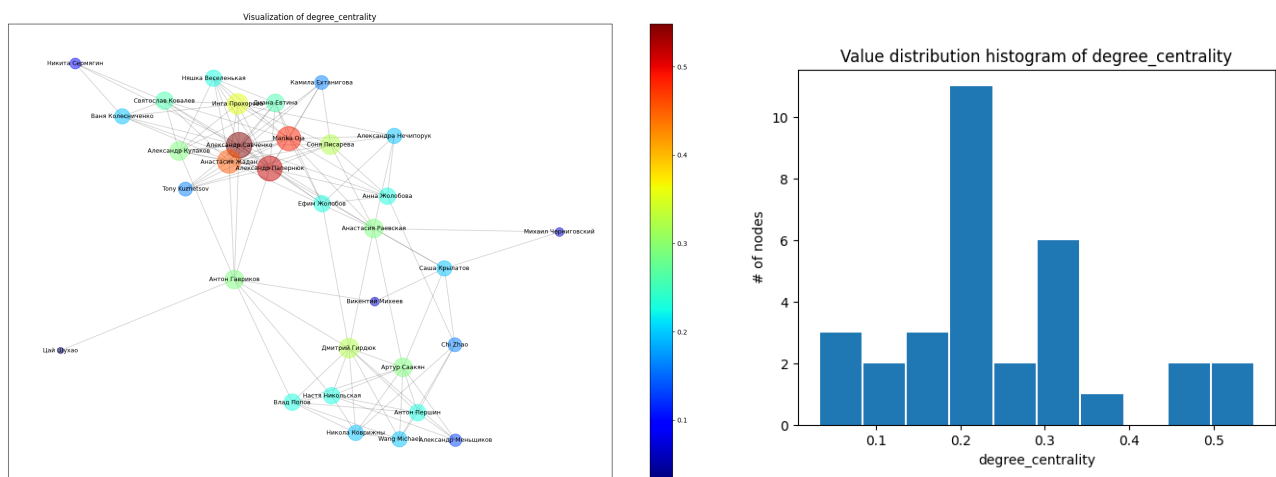


Рис. 1: Центральность по степени

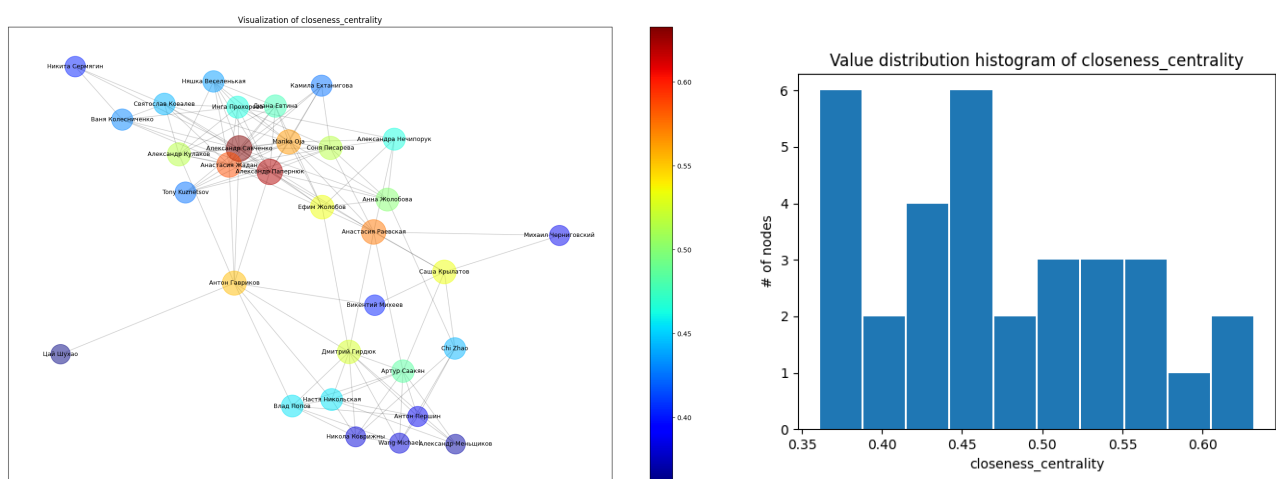


Рис. 2: Центральность по близости

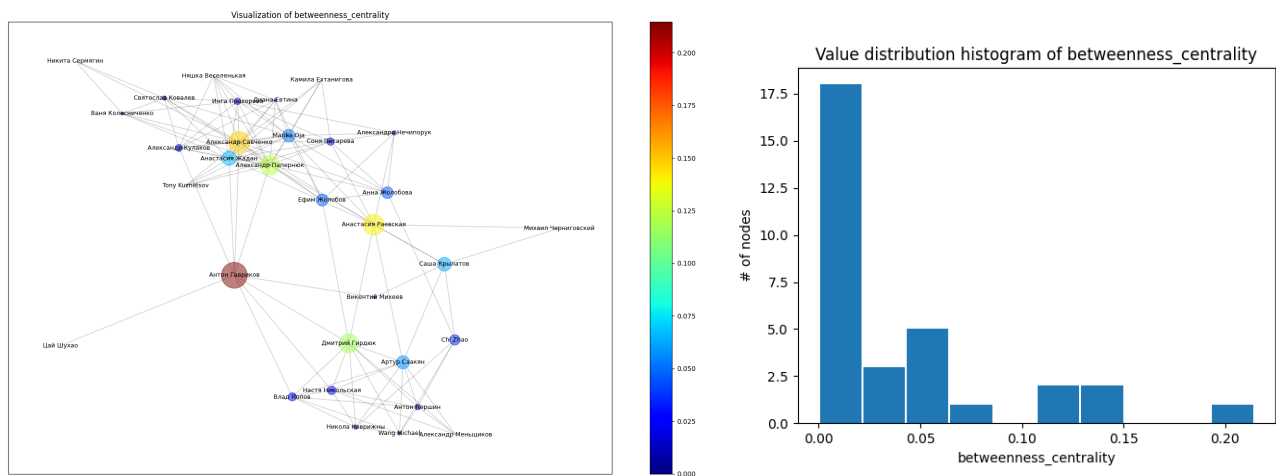


Рис. 3: Центральность по посредничеству

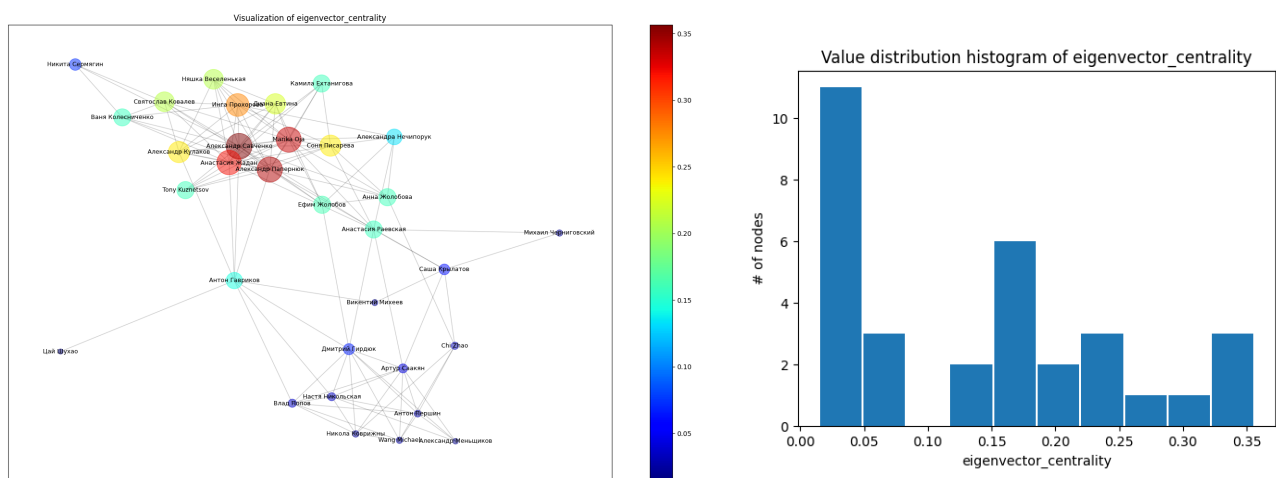


Рис. 4: Центральность по собственному вектору

Также была рассчитана косинусная мера близости вершин по матрице смежности. Выведены первые 10 пар вершин, которые являются наиболее близкими друг к другу, со следующими значениями:

- Косинусная мера
  - 1.00000
  - 1.00000
  - 0.92582
  - 0.86603
  - 0.84887
  - 0.84266
  - 0.82496
  - 0.82496
  - 0.82496
  - 0.81650

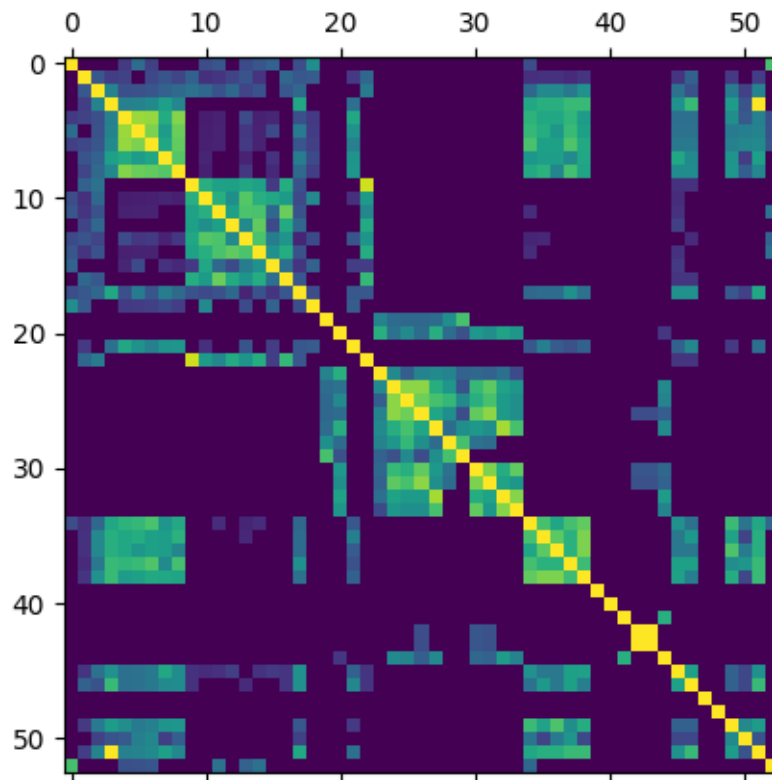


Рис. 5: Косинусная мера близости по матрице смежности

Нельзя не отметить различия в графиках и распределениях значений метрик. Это связано с тем, что граф является малым по числу вершин и неплотным, что и сказывается на отклонениях значений, несмотря на близость вершин. Однако, первые 10 значений косинусной близости могут означать наличие нескольких подсообществ в данном подграфе, а также то, что две пары с максимальными значениями косинусной близости могут представлять собой отдельные компоненты связности.

## Роли вершин сети

В заключительной части данной работы использовался тот же граф друзей ВКонтакте, что и в предыдущем задании.

Каждому узлу  $u$  был сопоставлен вектор  $\tilde{V}_u$ , элементами которого являются:

- Степень вершины;
- Число рёбер в эго-сети данной вершины;
- Число рёбер, соединяющих эго-сеть рассматриваемого узла с остальным графом.

По рассчитанным векторам характеристик были получены косинусные меры близости вершин сети. Выведены наиболее близкие друг к другу вершины со следующими значениями:

- Косинусная мера

- 1.00000
- 1.00000
- 1.00000
- 1.00000
- 1.00000
- 1.00000
- 1.00000
- 1.00000
- 1.00000
- 0.99998

По полученным значениям можно сделать вывод, что данный способ поиска различий в графе не является наилучшим, так как все вершины в данном графе хорошо коррелируют между собой по данным трём характеристикам. Различия косинусных мер, полученных в данном и предыдущем пункте работы, не стоит считать основополагающими: в первом случае расчёты производились на всём графе, когда как в предыдущем пункте – на наибольшей компоненте связности.