

**Министерство науки и высшего образования Российской Федерации  
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ  
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ  
«Национальный исследовательский университет ИТМО»  
(Университет ИТМО)**

**Факультет Инфокоммуникационных технологий (ИКТ)**

**Образовательная программа Мобильные и сетевые технологии**

**О Т Ч Е Т**

**по Лабораторной работе**

Дисциплина: Методы сетевого анализа.

Специальность: 09.03.03 Прикладная информатика.

Выполнил:

Балцат К. И.,

студент группы К33401

Санкт-Петербург  
2023

## ЗАДАНИЕ

### 1. \*\*Выберите 2 тематических сообщества.\*\*

Желательно, чтобы сообщества относились к одной тематике. Например можно исследовать 2 разные группы, посвящённые Гарри Поттеру, или предположить, что фанатам Гарри Поттера могут нравиться ещё и Голодные Игры или Сумерки и сравнить их сообщества между собой.

### 2. \*\*Сохраните списки участников выбранных сообществ. Есть ли пользователи, относящиеся к обоим сообществам?\*

### 3. \*\*Сохраните последние 2000 постов каждого из сообществ (ничего страшного, если их вдруг чуть меньше, в таком случае берём все, что есть). Используются ли в постах хэштеги? Если используются, то составьте топ хэштегов по встречаемости для каждой группы, визуализируйте полученные результаты. Сравните списки на предмет пересечений.\*\*

Иными словами, найдите непрерывные наборы символов, начинающиеся со знака #. Сохраните такие наборы по отдельности для обоих сообществ, посчитайте количество повторений, проверьте на пересечения. Для визуализации подойдут matplotlib, plotly express и другие библиотеки с аналогичным функционалом.

### 4. \*\*Посчитайте количество постов за каждый час суток для обоих сообществ, визуализируйте результаты. В какое время суток наиболее активны участники социальных групп? Совпадают ли часы с наибольшей активностью для обоих сообществ?\*

### 5. \*\*Постройте социальные графы для обоих сообществ.\*\*

Если во 2 пункте пересечений не было обнаружено, постройте 2 отдельных графа. Если пересечения по пользователям всё же есть, смотрите на то, насколько большие сообщества вы взяли. Графы можно объединить в 1 или исследовать по отдельности, если общий граф тяжело воспринимать визуально.

6. \*\*Сколько узлов получилось в каждом из графов? А сколько рёбер?\*\*

7. \*\*При помощи кластеризации выделите \*сообщества\* внутри обоих графов. Сколько \*сообществ\* получилось выделить в исследуемых графах?  
\*\*

Внутри крупной социальной группы, как правило, можно выделить небольшие \*сообщества\*. Например, в учебной группе часто есть несколько небольших компаний друзей. В получившихся после выполнения предыдущего пункта социальных графах выделите такие \*сообщества\*, применив кластеризацию.

## ВЫПОЛНЕНИЕ

Два сообщества: БФ Система и Фонд Будущие Лидеры.

```
Add code cell
⌘/Ctrl+M B

def get_users(group_id):
    """Получение всех участников группы"""
    user_list = []
    for offset in range(0, get_offset(group_id) + 1):
        params = {'access_token': vk_config["token"], 'v': 5.131, 'group_id': group_id, 'offset': offset*1000}
        users = requests.get('https://api.vk.com/method/groups.getMembers', params=params).json()['response']
        for user in users["items"]:
            user_list.append(user)
    return user_list

id_1 = 38829395
id_2 = 97561908
clubname_1 = "bfsistema"
clubname_2 = "fbl.russia"
rosmol = get_users(id_1)
tvoyhod = get_users(id_2)

Количество подписчиков: 6200
Количество подписчиков: 11234
```

```
↑ ↓

▼ Поиск пересечений

✓ 0s [5] common = set(tvoyhod).intersection(rosmol) # Ищем пересечения среди подписчиков обоих сообществ
common = list(common) # Переделываем получившийся результат в список
len(common) # Выводим длину списка, которая и будет равна количеству человек

130

✓ 3m [7] def get_posts(group_id, domain, amount):
    """Получаем нужное количество постов из группы и сохраняем текст записей"""
    text_list = []
    for offset in range(0, get_offset_wall(group_id, domain, amount)):
        params = {'access_token': vk_config["token"], 'owner_id': -1*group_id, 'domain': domain, 'offset': offset}
        r = requests.get('https://api.vk.com/method/wall.get', params=params)
        data = r.json()
        data = data['response']['items']
        for i in range(len(data)):
            if len(text_list) < amount:
                text_list.append(data[i]['text'])
            else:
                break
    return text_list

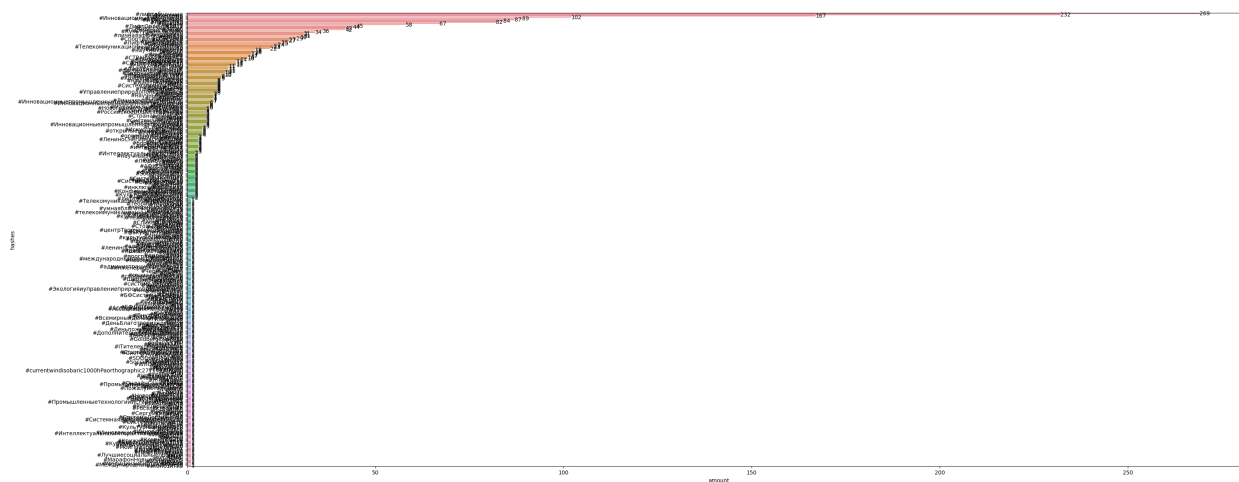
rosmol_posts = get_posts(id_1, clubname_1, 2000)
tvoyhod_posts = get_posts(id_2, clubname_2, 2000)

Количество постов: 2416
Количество постов: 2254

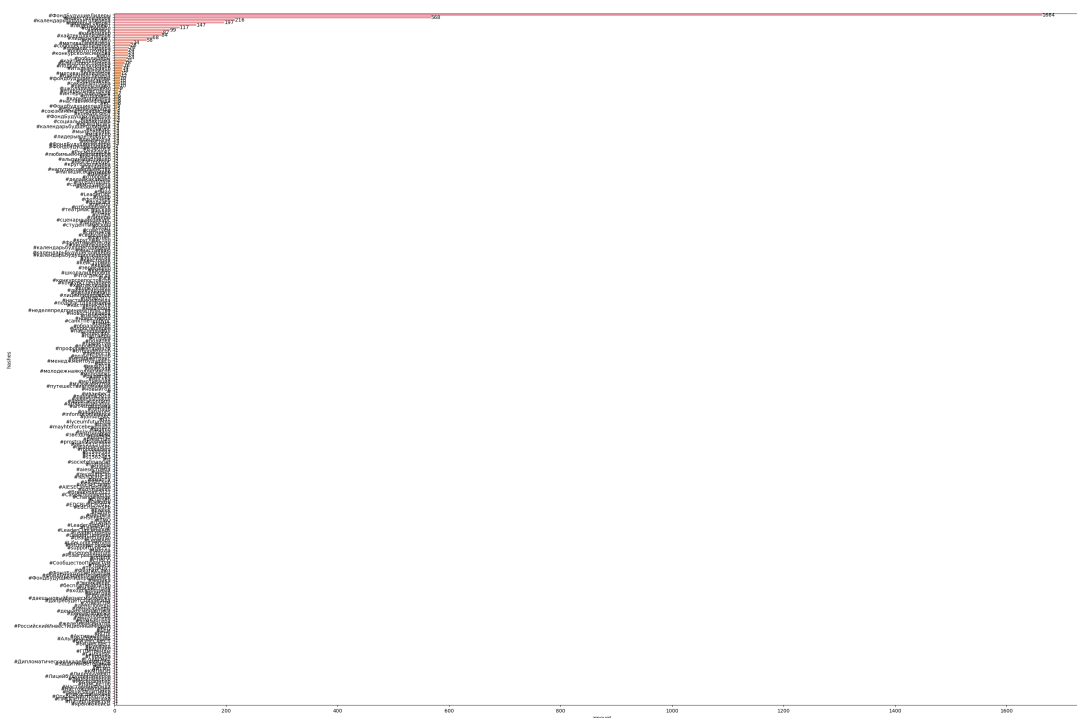
✓ 0s [8] print(len(rosmol_posts)) # Проверка количества текстов записей для сообщества "Росмолодежь"
print(len(tvoyhod_posts)) # Проверка количества текстов записей для сообщества "Твой Ход"

2000
2000
```

Сообщество 1

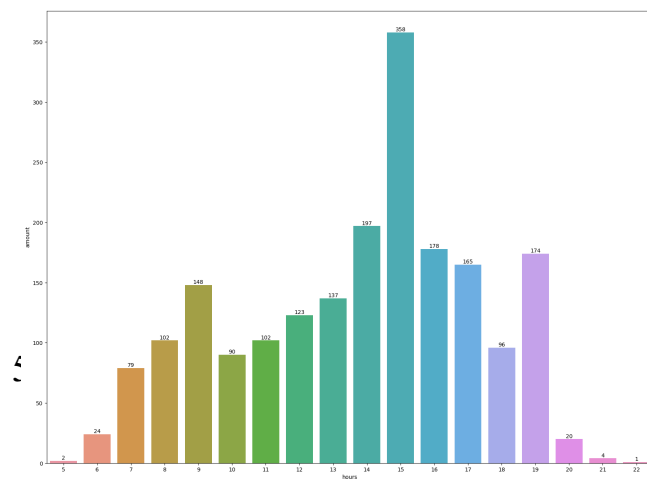
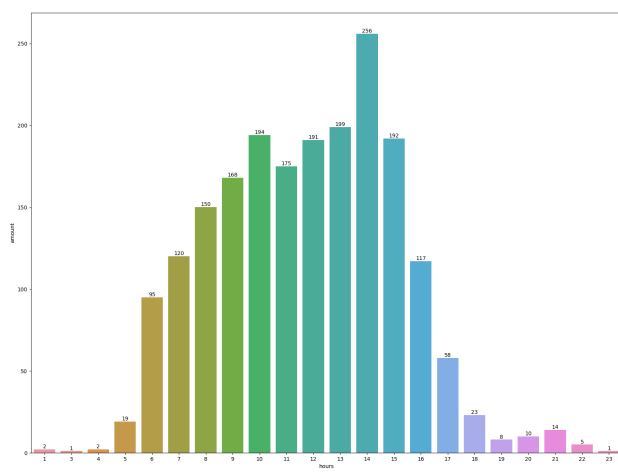


Сообщество 2



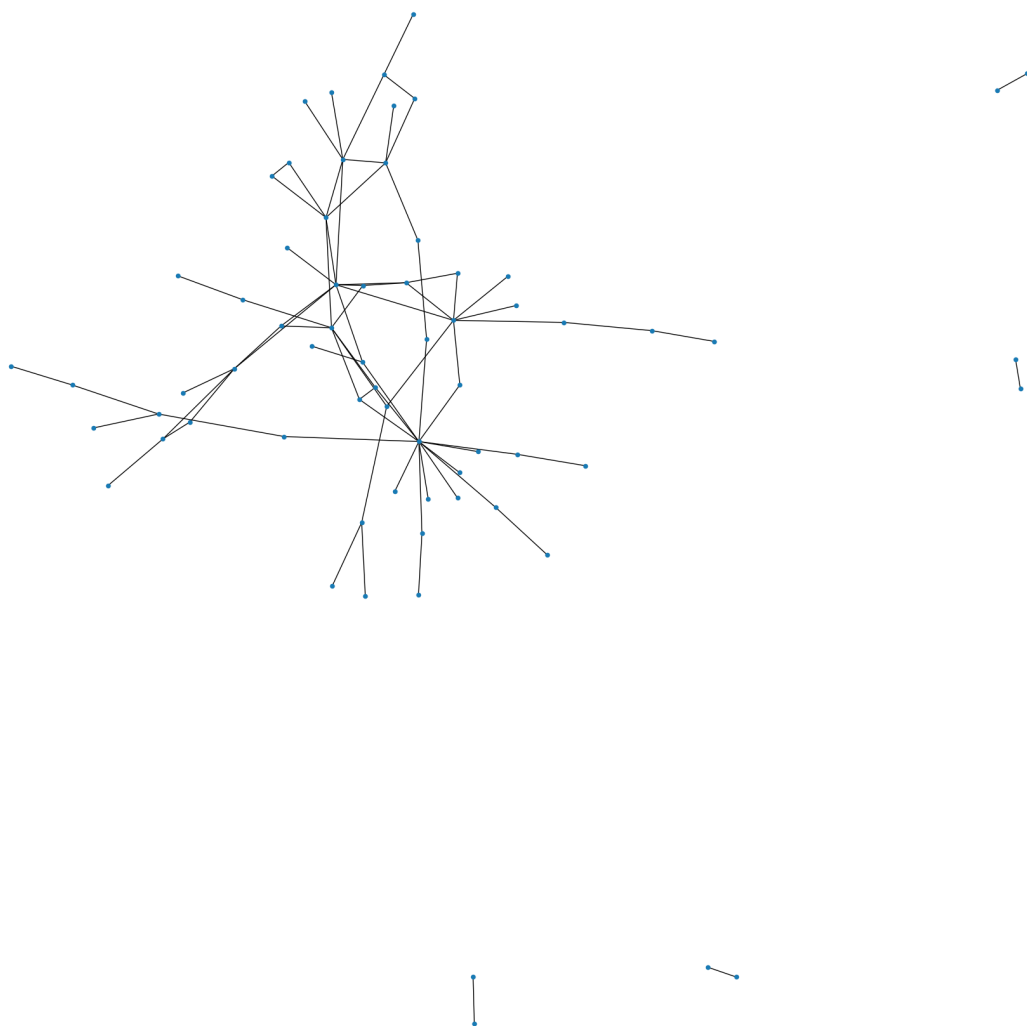
Сообщество 1

Сообщество 2

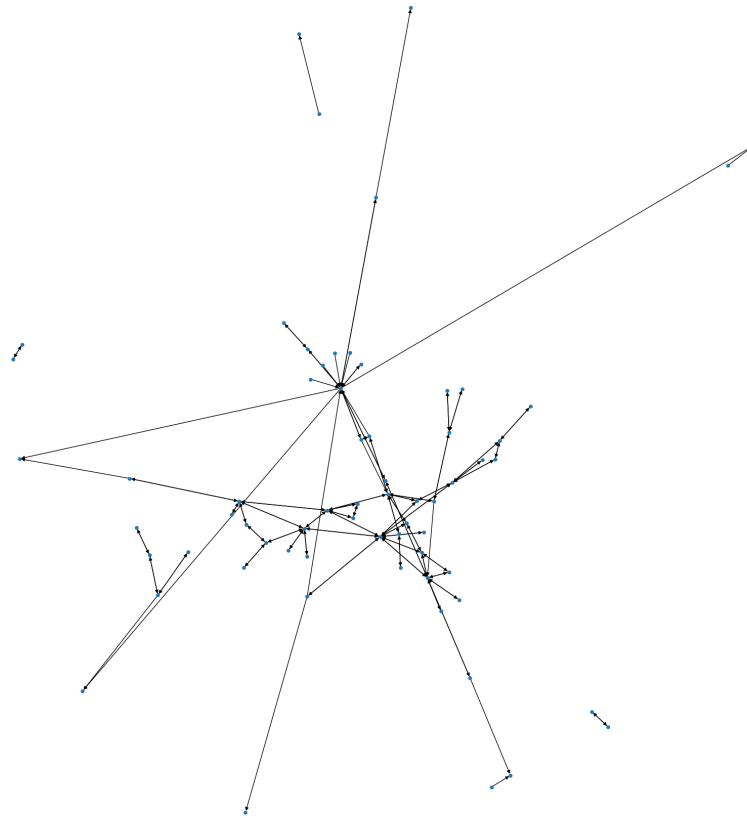


## Социальный граф для пересечений, ориентированный и неориентированный

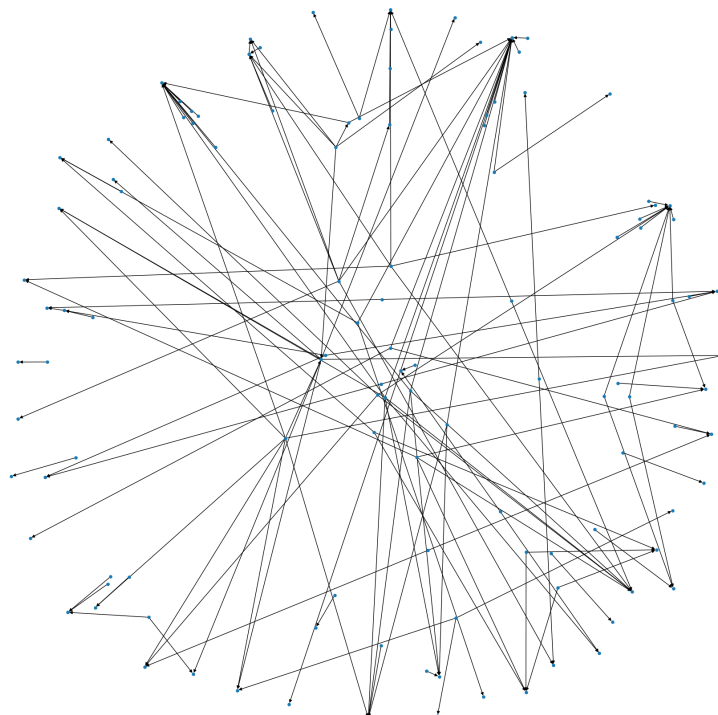
Социальный граф для людей, являющихся членами обоих сообществ



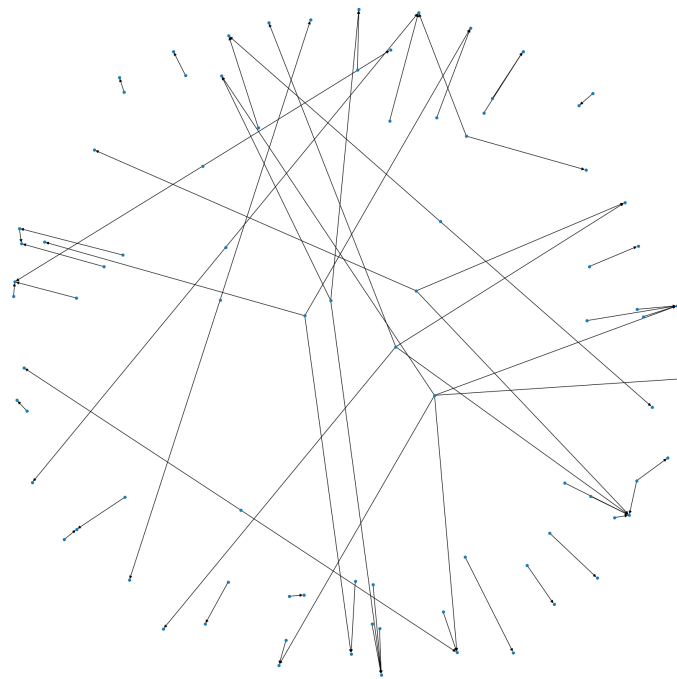
Социальный граф для людей, являющихся членами обоих сообществ



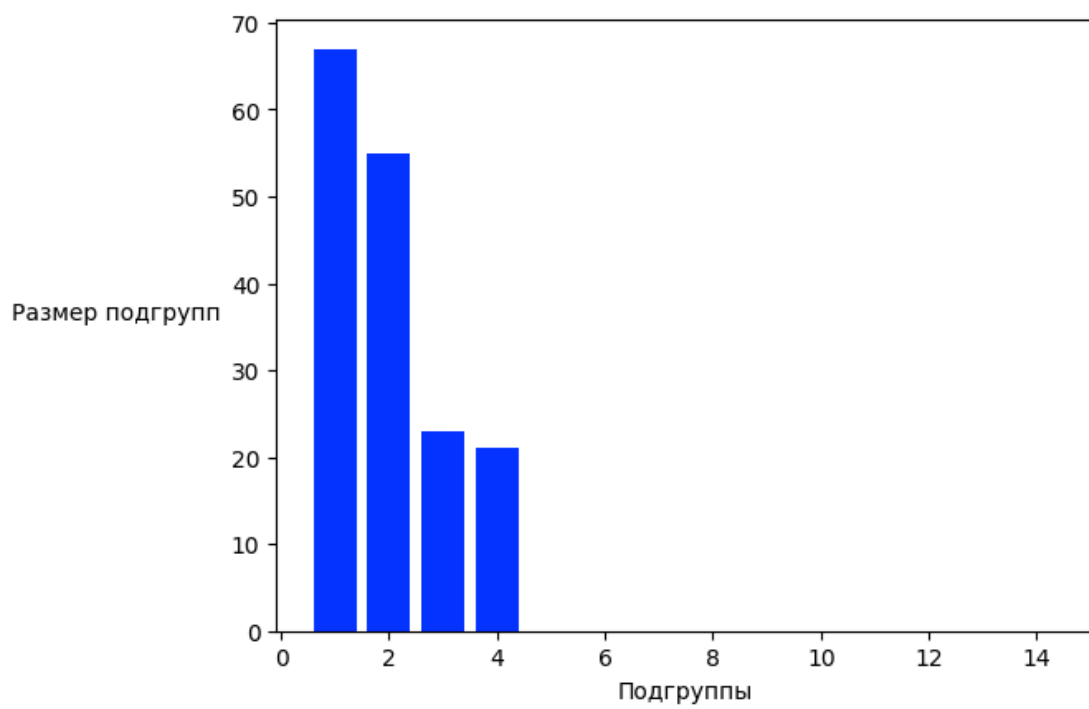
Социальный граф для людей, являющихся членами сообщества bfsistema



Социальный граф для людей, являющихся членами сообщества@lussia

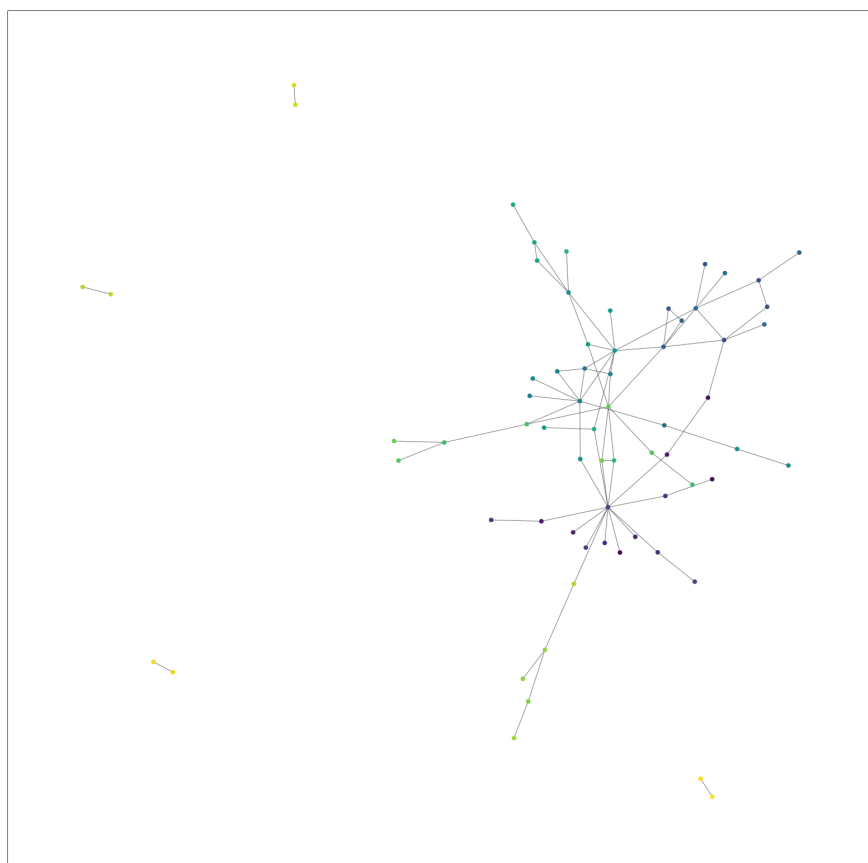
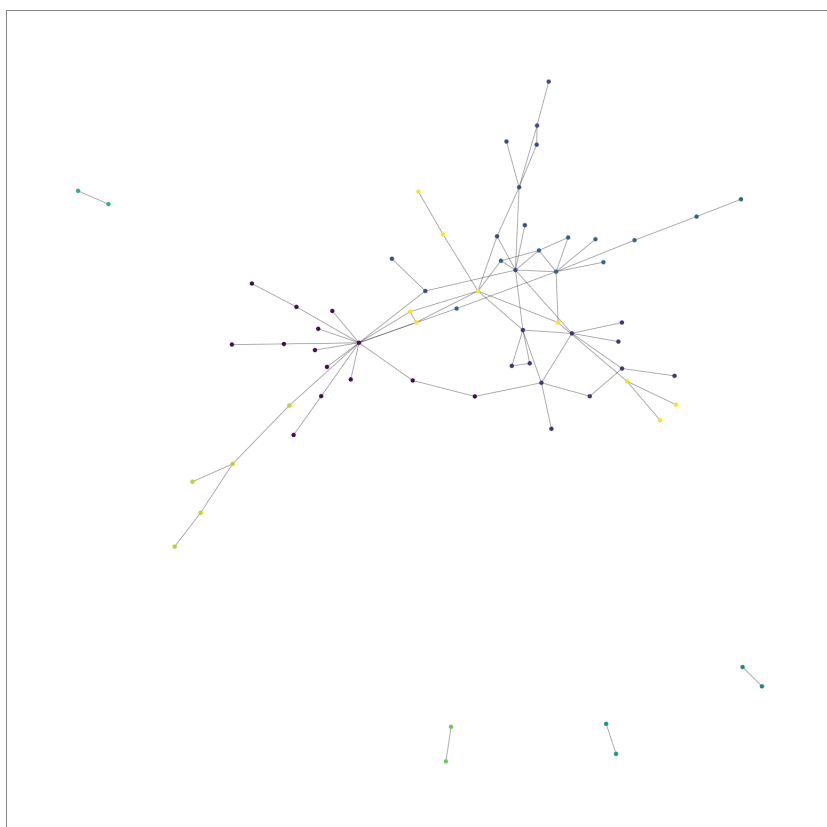


## Сообщества внутри графов





## Кластеризация участников обоих сообществ



## ВЫВОД

Я выполнил лабораторную работу и на практике освоил методы анализа сетей.