

National Research University Higher School of Economics (Higher School of Economics/HSE)  
Faculty of Computer Science  
Bachelor's Programme Data Science and Business Analytics  
01.03.02 Applied Mathematics and Computer Science

## Internship report

Fulfilled by

Krauze Natalia Olegovna  
*(Surname, Given Name, Middle Name if any)*

\_\_\_\_\_  
*(signature)*

Checked by

\_\_\_\_\_  
(job or academic title)

\_\_\_\_\_  
*(surname, initials)*

\_\_\_\_\_  
(signature)

**Moscow, 2019**

## **Contents:**

1. Formulation of the task
2. Getting the data by using VK API
3. Creating and visualisation of the graph of the users of VK community
4. Conclusions
5. Educational Internship Schedule(Plan)
6. Bibliography

## **1. Formulation of the task**

In this internship there was a task to analyse some online communities in one specific social network - VK by using algorithms connected with graph theory. In our case, we choose the group called 'КБЕКТЧН', when we worked with it there were 442 users in most cases from HSE from the faculty of computer science.

In global, we divided our big task into several sub-tasks: getting the data, constructing a graph based on our data and writing codes of algorithms and analysing how fast and correct it works. I did the part connected with getting data and constructing the graph. Also, I decided to visualise our graph to understand the necessities.

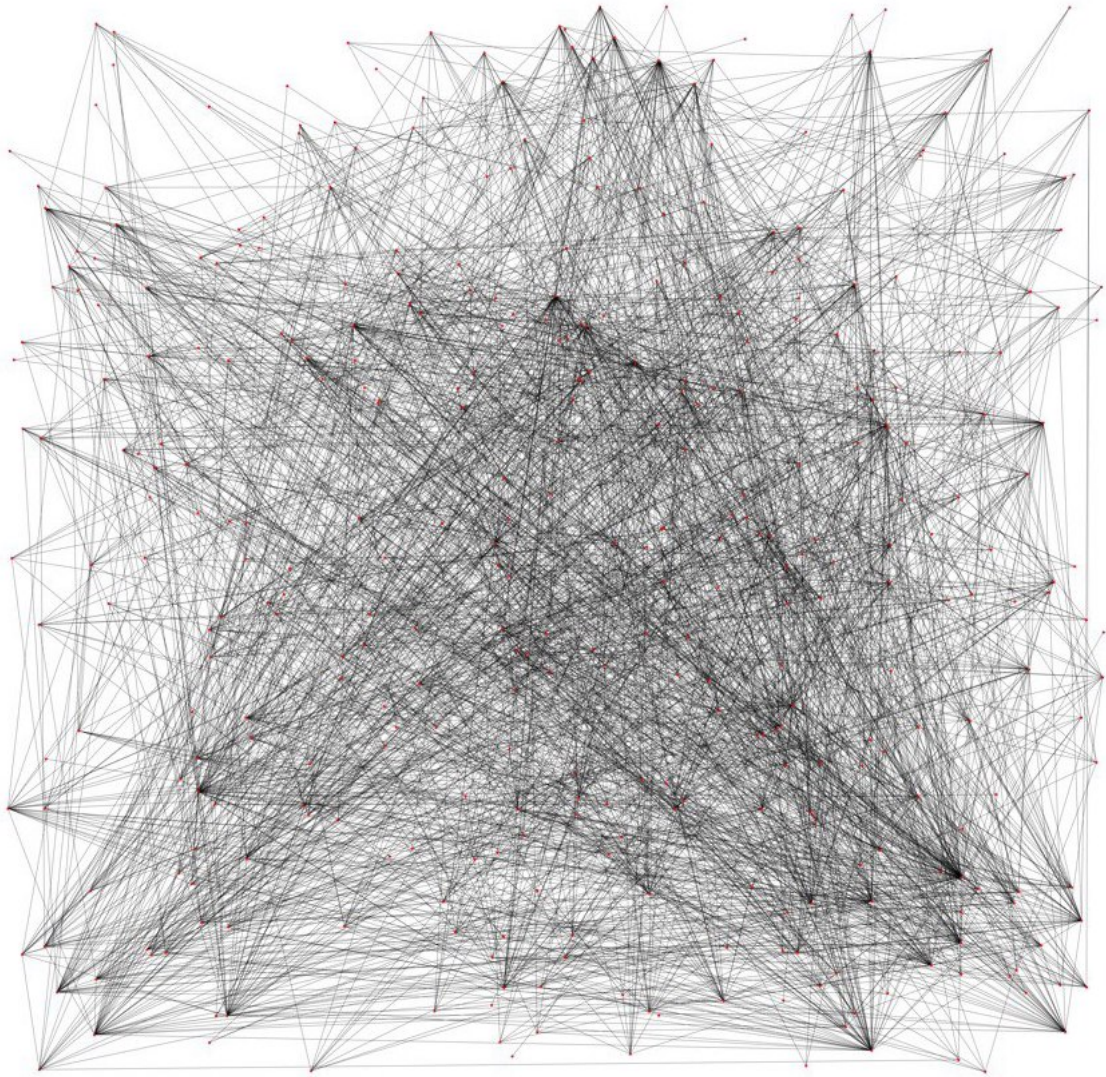
## **2. Getting the data**

To get the data I used VK API. To make requests to VK API I used the special library in Python called 'request'. Firstly, I get access to the list of users of the group by using method called 'get.members'. Then, it was needed to get the list of friends of every user. I did it by the method called 'friends.get'. However, there were some problems. Take closed, private accounts of users of the group and I could not get information about them, some quarter of users profiles were banned or deleted. That is why I was needed to cope with it to escape falls in the programme. Also, VK API allows to do not more than 3 request per second. That is why I used special necessary delays to satisfy the limit of 3 requests per second. Next, I think that request to the server is a constant time as well as other operations. I represented our users as an unordered set - that is why to understand whether the user is a member of a group it is needed to be constant time. I counted the number of users in the group - getting the connections between members of the group (to be friends) took quadratic of the amount of users of the group time.

## **3. Creating and visualisation of the graph of the users of VK community**

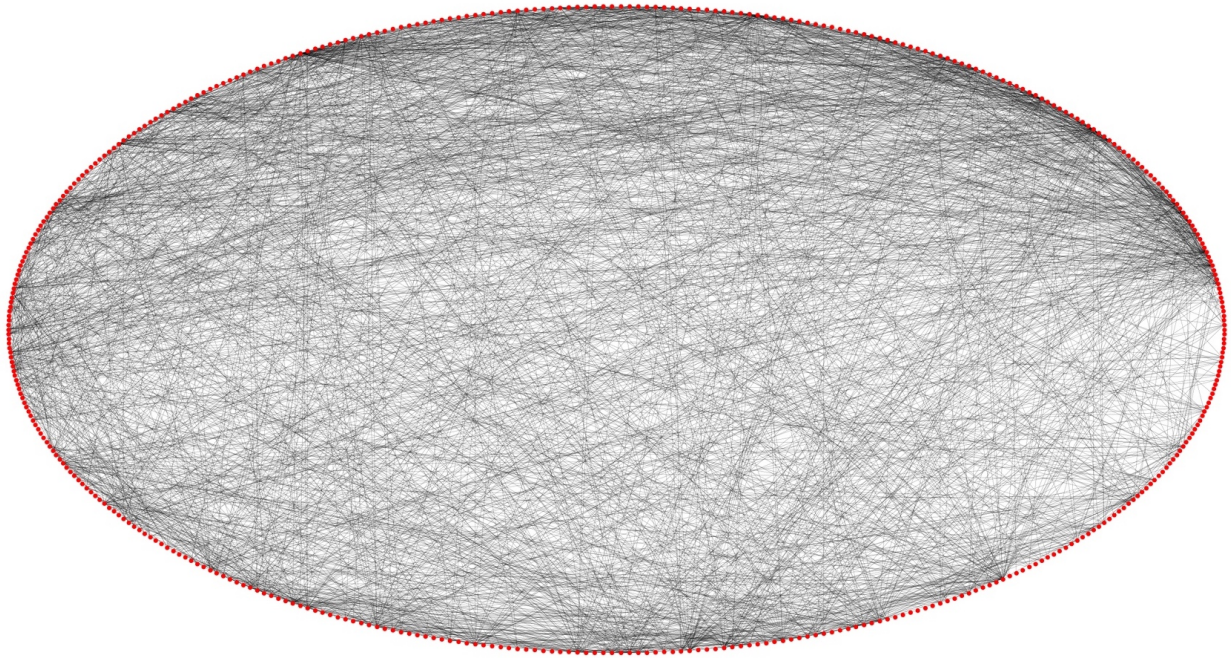
I did some representations of graphs of the users. Graph is constructed in the proper way: vertices - members of the group and edges - constructed between those users who are friends. So, there were several

representations of the graph : the adjacent matrix, the list of adjacent vertices and the list of edges. Got graph consist of 442 vertices and 2337 edges. To get the best representation of the graph I decided to use libraries in Python to visualise it. I did several visualisations - random location of vertices on the plane and the second one with the ordered vertices around the line of an oval. There are my visualisations above:



Visualisation with random location of vertices on the plane





Visualisation with the ordered vertices around the line of an oval

## **4. Conclusions**

During the internship I learned a lot of new things. Firstly, I learned how to work with the library VK API and the library 'request'. Also, I got worthy experience how to work with documentations by myself. Secondly, I knew how from one representation of the graph get new different ones. Moreover, I learned how to implement and visualise graphs with the library 'NetworkX'. To sum up, we were managed to work in team, communicate and solve practical tasks. Also, I learned more about algorithms, analysis and the fields of computer science.

## 5. Educational Internship Schedule

№	Calendar period	Work Plan	Internship Supervisor's signature/
1	01.07.2019	1. Organizational (induction) meeting	
2	01.07.2019	2. Instructing on the requirements of labor protection, safety, fire safety and internal labor regulations	
3	01.07.2019 - 13.07.2019	3. Fulfillment of Individual Assignment	
4	01.07.2019 - 13.07.2019	4. Consultation	
5	14.07.2019	5. Preparation and submission of the Report	

## 6. Bibliography

1. <https://vk.com/dev/manuals> - VK API manual
2. <https://networkx.github.io/documentation/stable/tutorial.html> - NetworkX tutorial