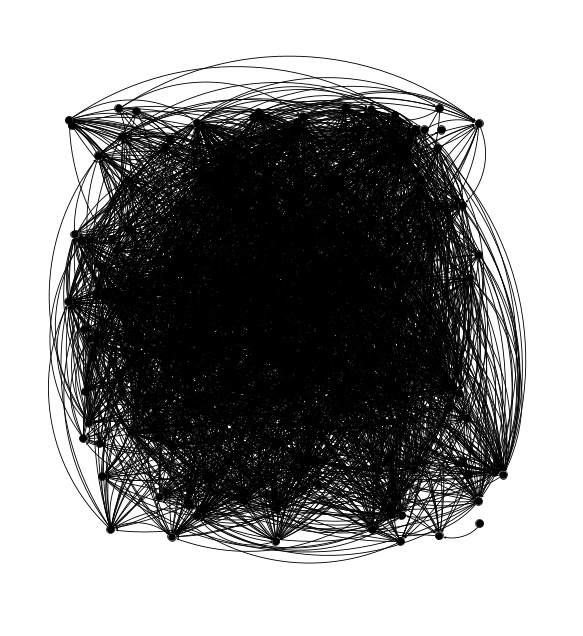
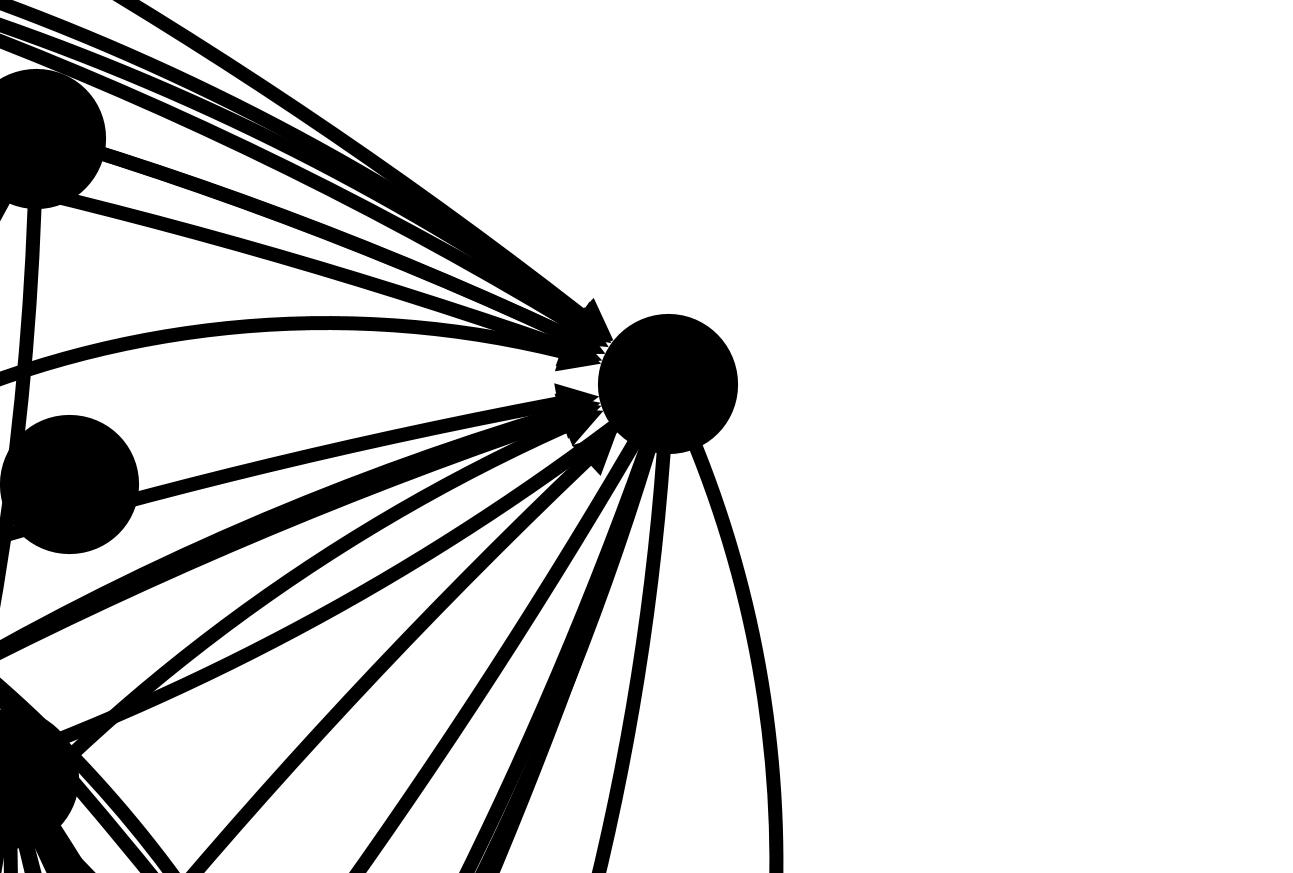
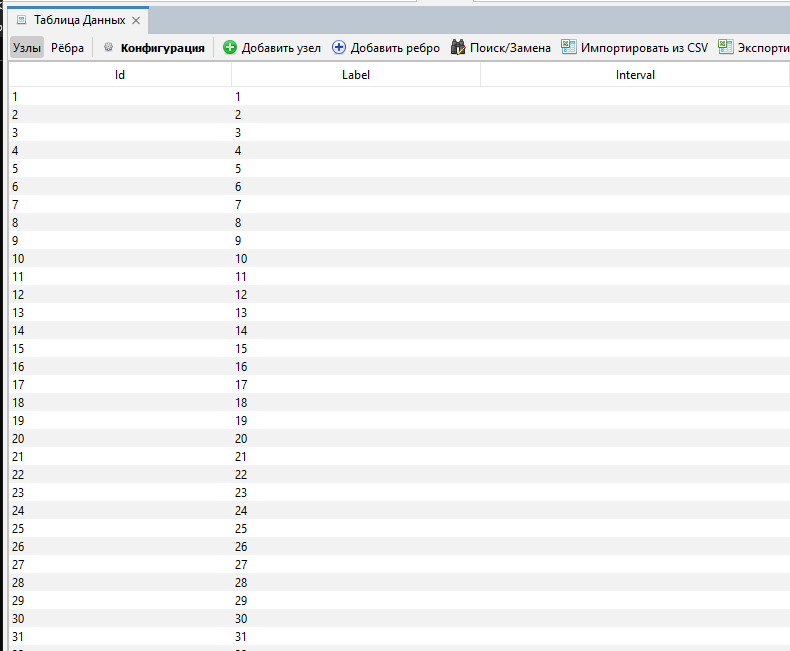
Графы и Гефи Gephi

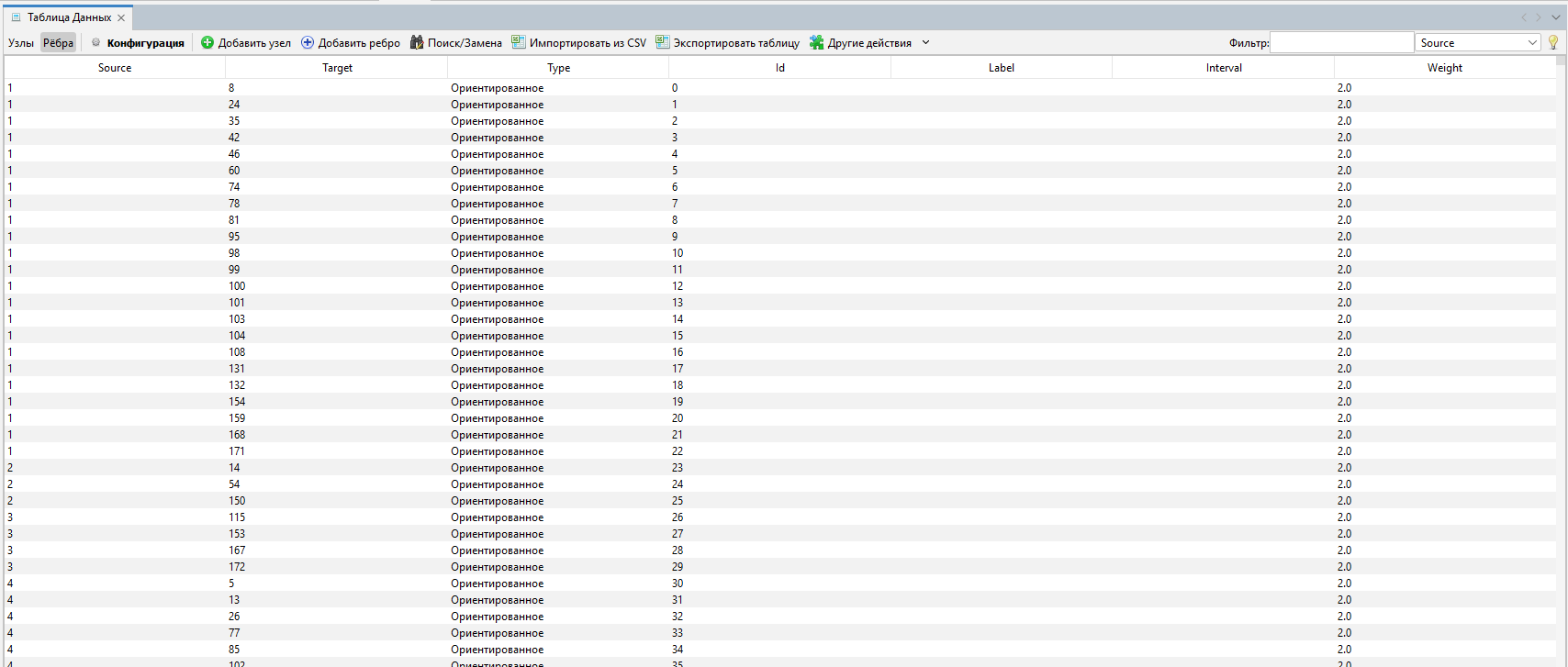
Для начала запустим какой-нибудь датасет с набором данных в виде графа: например, Jazz musicians network (//TODO ссылка на GitHub (<https://gephi.org/datasets/cpan-distributions.gexf.zip> ||| <https://github.com/gephi/gephi/wiki/Datasets> )), сеть с музыкантами по джазу. При открытии датасета в Gephi мы видим вот такое пятно из каракуль:



Если присмотреться (увеличить) изображение, то мы видим узлы и дуги: //TODO Что такое ориентированнось в графах (пояснение)   
  


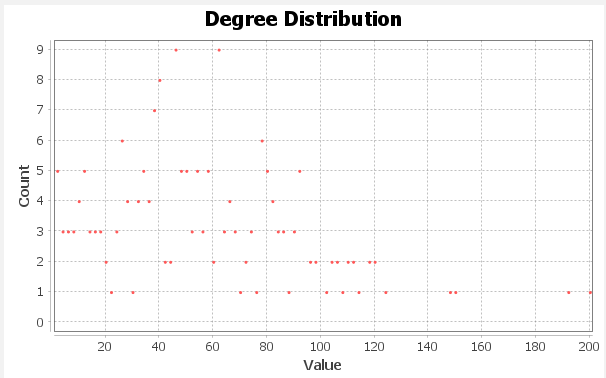
Граф — это структура, состоящая из набора объектов, где некоторые пары объектов в некотором смысле «связаны». Объекты представлены абстракциями, называемыми вершинами (также называемыми узлами или точками), а каждая из связанных пар вершин называется ребром (также называемым связующим звеном или линией).

Вершина (узел) — это элемент (точка) графа, обозначающий объект любой природы, входящий в множество объектов, описываемое графом. Вот, кстати, как таблично выглядят узлы:  
  
Дуга — это направленный путь, который идёт только в одну сторону. Обычно говорят, что дуга идёт из одной вершины в другую.

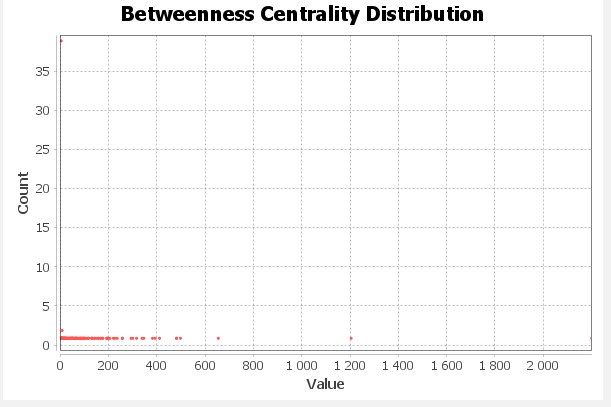
Ребро — это ненаправленный путь, который идёт в обе стороны. Вот, кстати, таблица с ребрами графа:  


**Давайте разберемся с графом по детальнее, а именно, центральность степени, центральность посредничества, центральность близости, центральность собственного вектора.**

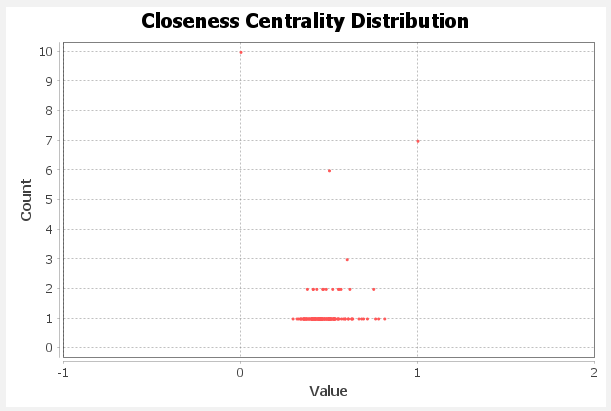
Центральность по степени (количество связей)– это то, насколько важна конкретная вершина графа с точки зрения количества связей с другими вершинами в сети



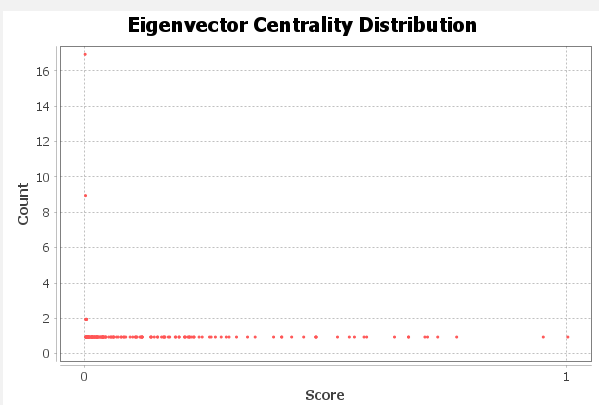
Степень посредничества (это количество раз, когда узел является промежуточным между двумя другими узлами (мост/посредник))— это мера центральности в графе, основанная на кратчайших путях. Для любой пары вершин в связном графе существует по меньшей мере один путь между вершинами, для которого минимально либо число рёбер, по которым путь проходит, либо сумма весов этих рёбер. Степень посредничества для каждой вершины равна числу этих кратчайших путей через вершину.

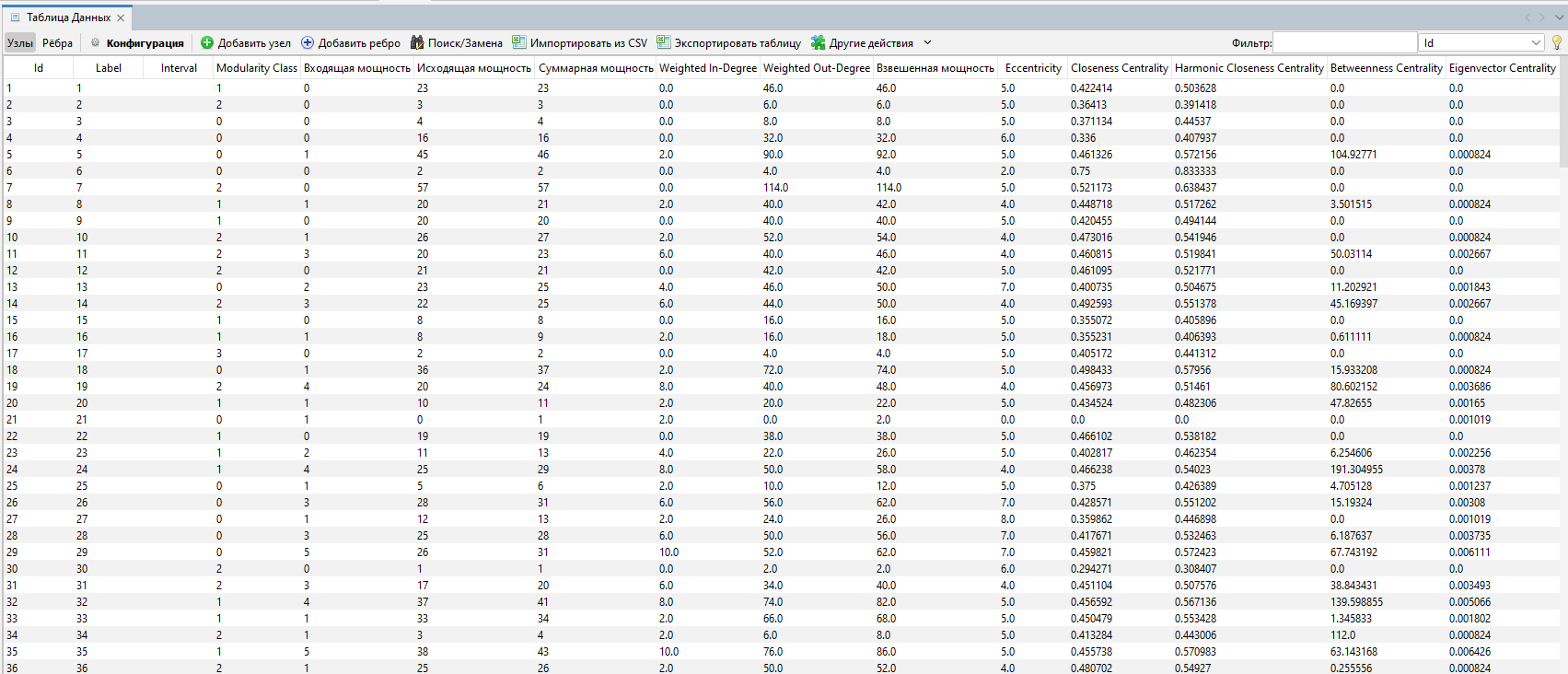


Центральность близости у графа (насколько узел близок между остальными узлами графа) — это насколько узел близок к другим узлам графа. Она определяется как обратная сумма расстояний от данного узла до всех остальных узлов. Чем меньше расстояние от узла до остальных, тем выше его центральность близости.

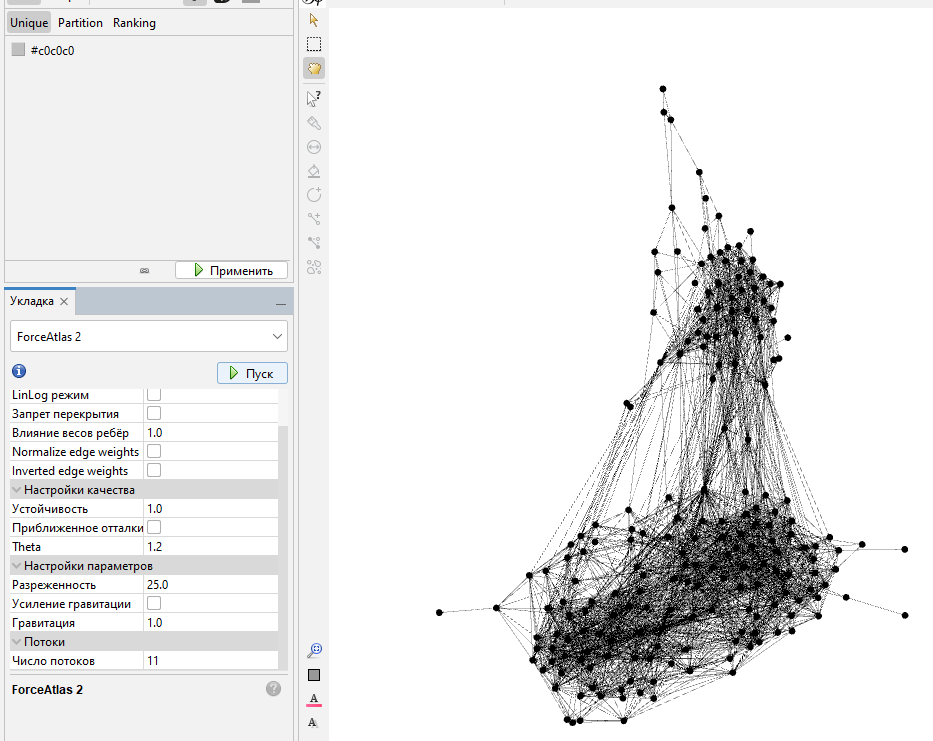


Центральность собственного вектора (значимость – количество связей и их важность) — измеряет влияние узла с учетом его соседей. Узел считается важным, если он соединен с другими важными узлами. Центральность основана на собственных векторах матрицы смежности графа.

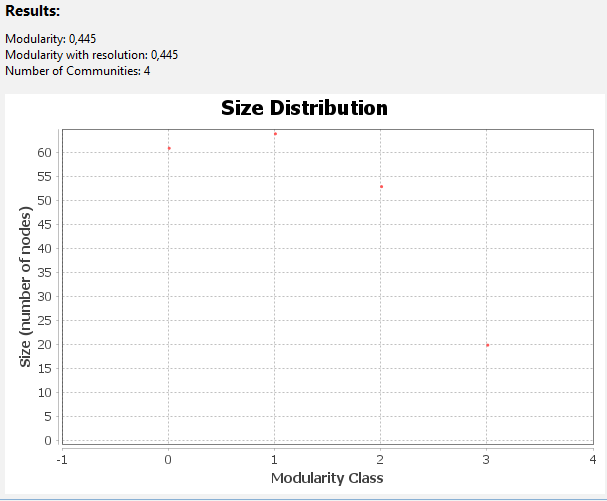


Если мы хотим посмотреть все характеристики каждого узла, надо просто открыть таблицу еще раз: 

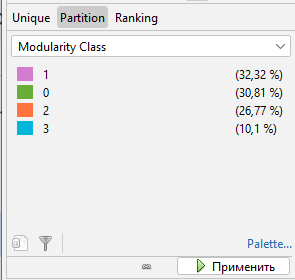
**Давайте уложим граф так, чтобы было более менее понятно, используем укладку ForceAtlas2 с разрежённостью 25, после нее получится такой результат //TODO по какому признаку появилось 4 кластера:**



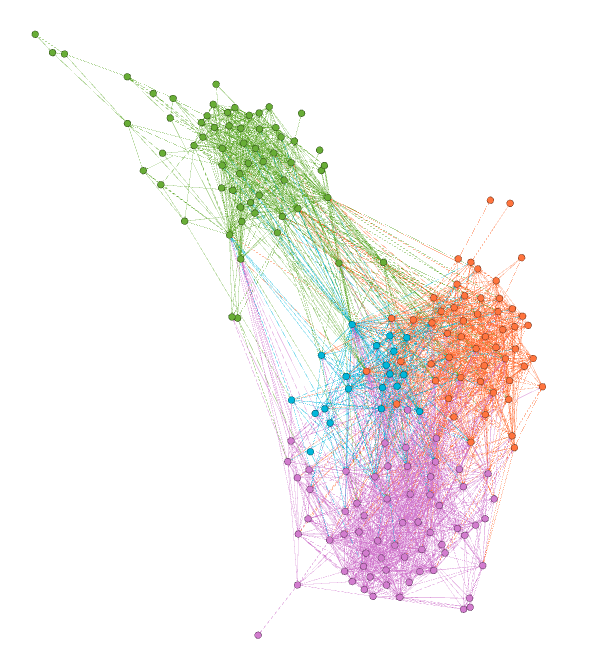
Теперь выделим кластеры… Для этого надо запустить модулярность:  

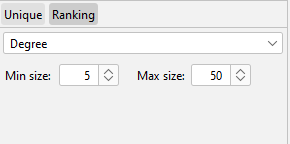

И у нас получилось 4 класса [ОТРЕДАКТИРОВАНО] если узел принадлежит к нулевому классу, то у него 61 присоединенный сосед, если к первому классу, то 64 соседа, если ко второму – 53 соседа, если к третьему – 20 соседей, соседи в данном случае - сообщества???:  


Они отобразились здесь:



Если мы покрасим, то увидим общую картинку такую:



Чтобы в графе не путаться и найти более или менее значимые узлы (те, где примыкают большее или меньшее кол-во узлов), надо просто изменить размер здесь:  


Теперь граф приобрел такой вид \*Чем больше узел, тем он более важный и имеет больше пересечений\*:  
