Выбор предметной области: Граф друзей в социальных сетях. Случай использования – аналитические системы, системы анализа данных из открытых источников.

Для отображения графа в указанной предметной области могут существовать жесткие ограничения со стороны «заказчика» (клиента). Соответственно в графическом представлении графа выделяются всевозможные параметры (критерии). Часть из которых, как было сказано ранее, жестко заданы. Вариации же остальных параметров позволяют нам выбрать более-менее эргономичное представление графа.

Допустим введены следующие жесткие ограничения:

* ~~Изображение графа статическое. То есть представляет собой неподвижное изображение.~~
* ~~Возможность изменения масштаба графа.~~
* ~~Перемещение видимой области графа.~~
* Узлы графа имеют круглую форму.
* Рёбра должны быть прямые. Отсутствуют изломы и т.д.

Удобство работы с графом заключается в следующем:

1. Высокая скорость нахождения для любого узла всех связанных с ним узлов.
2. Высокая скорость нахождения узлов с большим количеством связей (ключевых узлов).
3. Сортировка ключевых узлов по уменьшению/возрастанию связей.

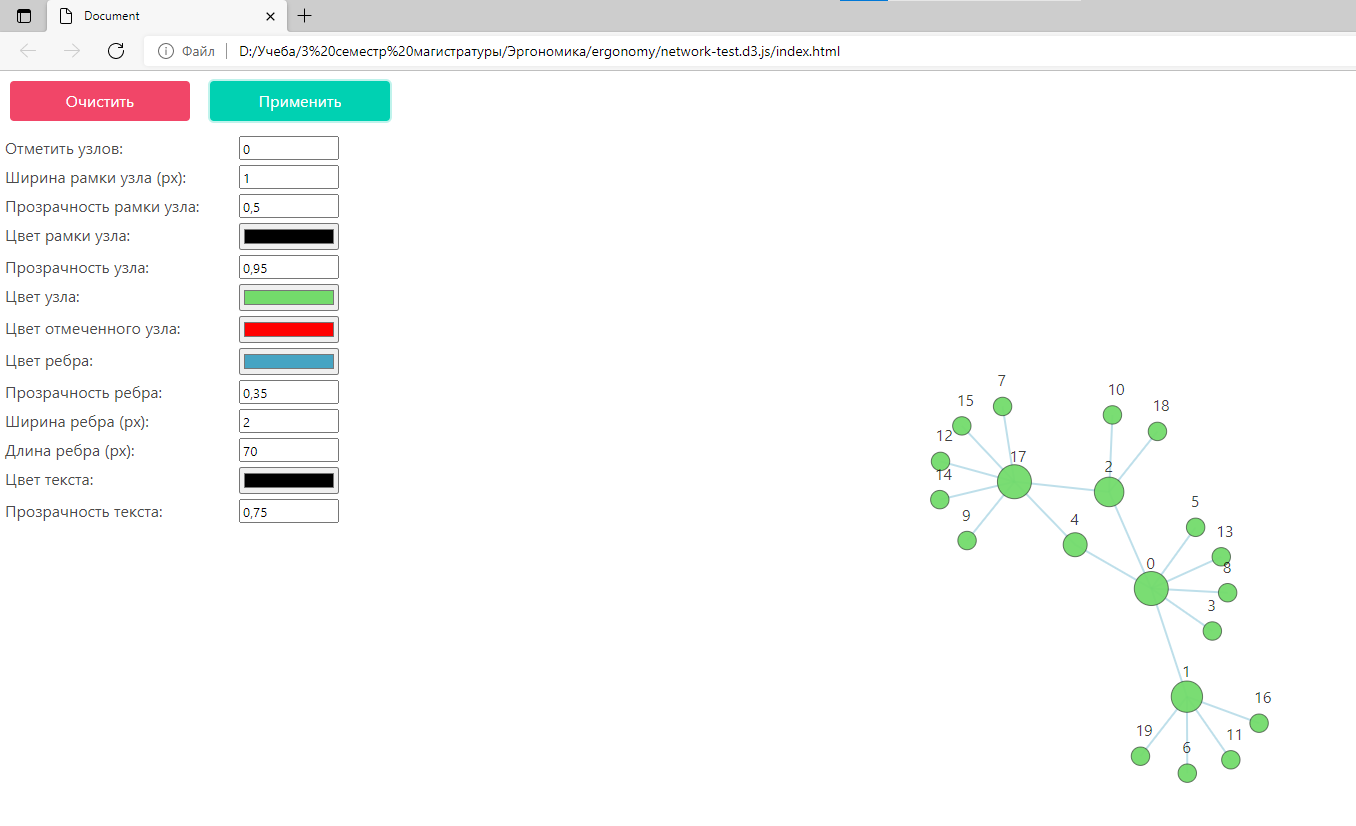
Изменяемые параметры:

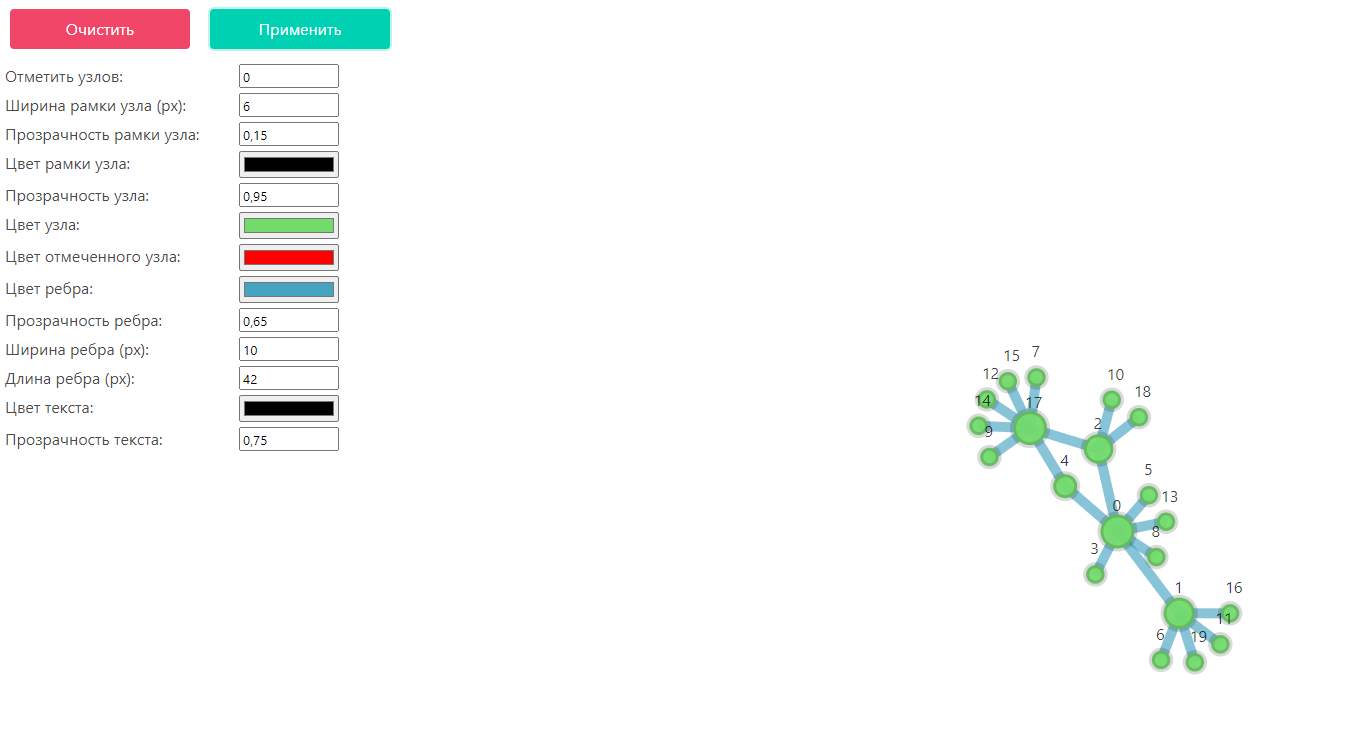
1. Изменять диаметр узла в зависимости от количества связей
   1. Изменять функцию
   2. Изменять коэффициенты в функции
2. Ширину рамки узла
3. Прозрачность рамки узла
4. Прозрачность узла
5. Расстояние между связанными узлами
6. Прозрачность связей
7. Ширину линий связей
8. ~~Цвета~~

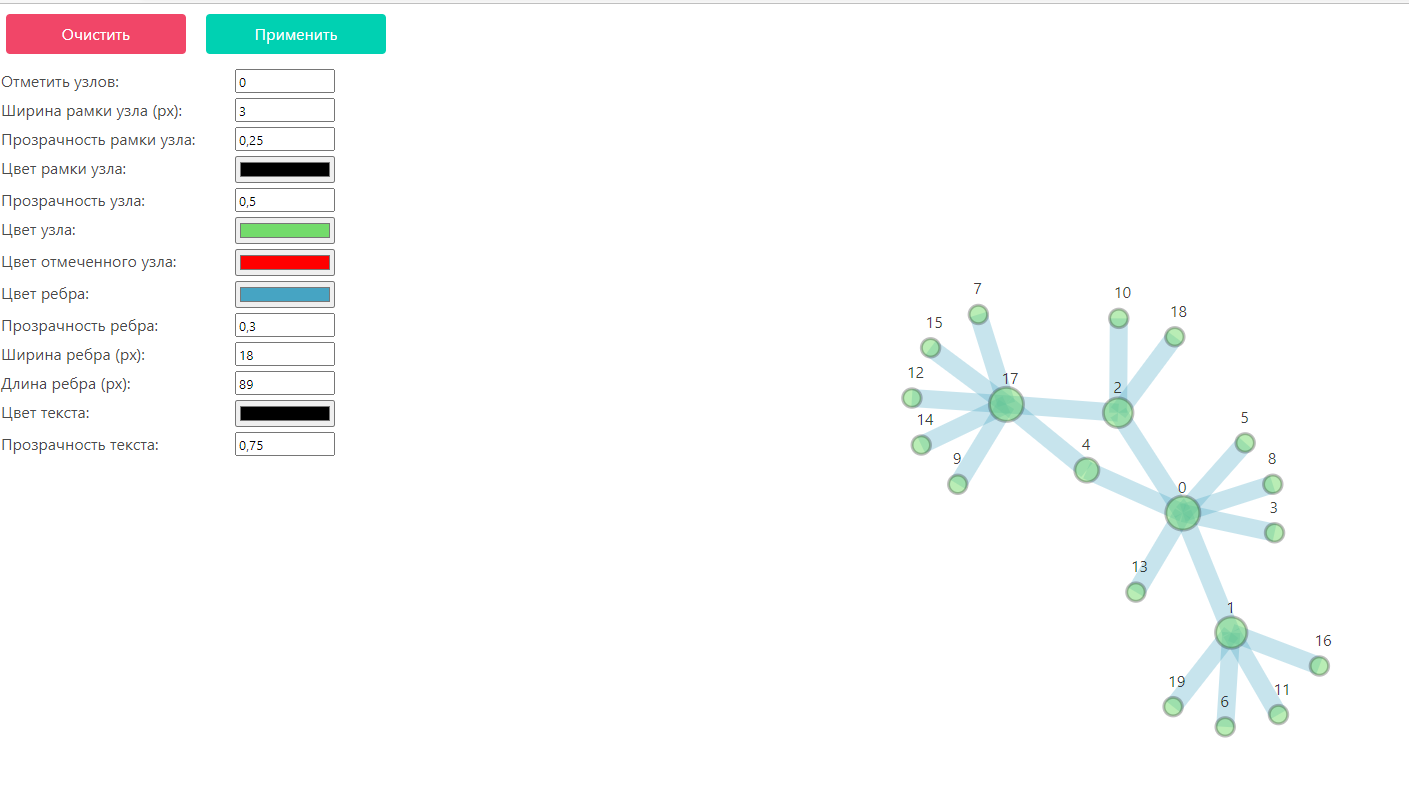
Эксперименты:

1. Предложить 1 случайный узел и попросить экспериментатора определить все связанные с ним узлы *(Проверяется все ли узлы были определены и за какое время)*
2. Предложить 3 случайных узла и попросить экспериментатора визуально определить у какого меньше и больше всего связей *(Проверяется правильно ли экспериментатор отсортировал по количеству связей эти 3 узла и за какое время)*
3. Предложить 2 случайных узла и попросить экспериментатора перечислить количество узлов, которые имеют связь с первыми двумя *(Проверяется все ли узлы были определены и за какое время)*
4. Найти самый короткий маршрут (или хотя бы оптимальный) от точки А до точки Б в графе. *(Проверяется действительно ли экспериментатор нашёл самый короткий путь, считается ошибка + учитывается время нахождения)*
5. Найти в графе самый сильно связанный и слабо связанные узлы (те у которых больше/меньше всего связей). *(Проверяется правильно ли найдены узлы и за какое время)*

Пример изменения параметров графа в приложении (небольшой пример):





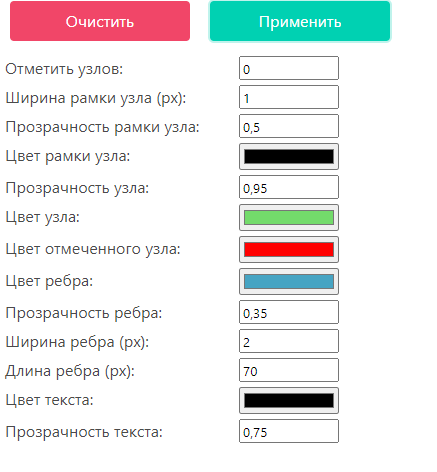


Планируется: провести вышеперечисленные эксперименты засекая время выполнения, а также проверяя правильность выполнения (изменяя каждый параметр графа независимо друг от друга с определенным шагом). После выполнения всех экспериментов для каждого параметра графа получим N\*M таблиц (Где N – количество изменяемых параметров, M – количество видов экспериментов).

Изменяемые параметры:

Пока что не добавил, как должен изменяться диаметр узлов, так как я думаю сделать это логарифмической функцией от количества связей. То есть диаметр возрастает, чем больше связей имеет узел – показывает важность данного узла (пример можно посмотреть выше на картинках).

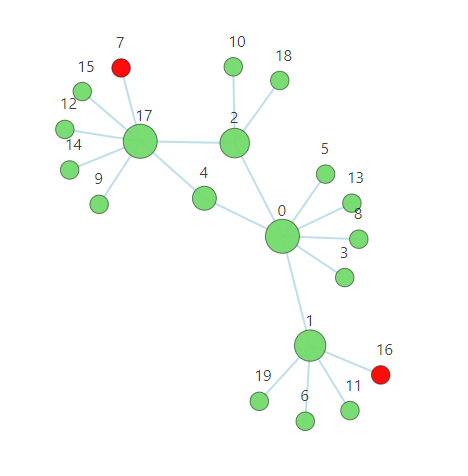
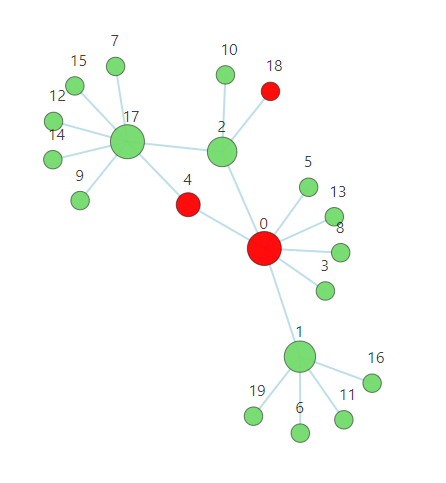
(Цвета изменять не буду, эти параметры я уберу)

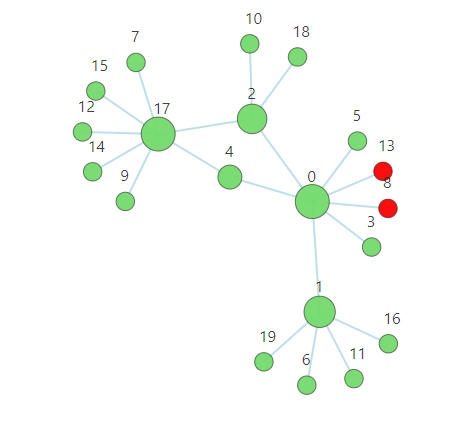
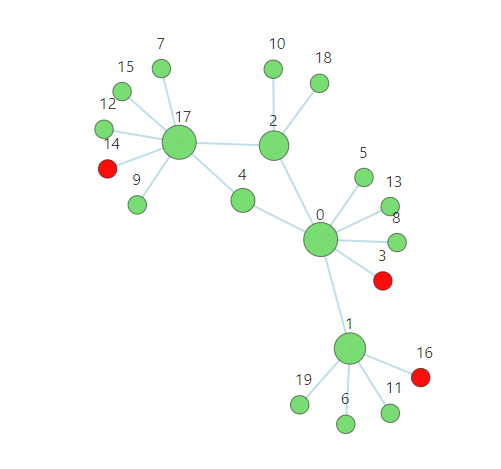


В результате всех экспериментов можно будет решить какие должны быть параметры для лучшего прохождения экспериментов.

Используя генератор случайных чисел, я подсвечиваю необходимое мне количество узлов для выполнения экспериментов. Где нужно предложить экспериментатору необходимое количество узлов.

Примеры:





Общие вопросы-замечания.

1. Мы исследуем граф на предмет ЧЕГО???? Все эксперименты направлены на то, чтобы посчитать количество связных узлов и время. ЧТО НАМ ЭТО ДАСТ? Мы определим СВЯЗНОСТЬ???, если есть такая характеристика. Если это соцсети, мы определим активность??, коммуникабельность??, **ЭТО ГЛАВНЫЙ ВОПРОС**. Если мы докажем (нам будет известно или как то по другому), что 2 приведенных параметра влияют НА ???? ЧТО???
2. По сути экспериментов
3. Есть какой то ВЕКТОР проведения экспериментов? Системность?
4. Сколько экспериментов надо провести?? Сколько групп экспериментов надо провети?
5. После проведения таких спонтанных экспериментов ни на один вопрос утвердительно не ответить: типа достаточно экспериментов? Непонятно. Достаточно изменяемых параметров?? Непонятно. Можно взять другие изменяемые параметры?? Непонятно.
6. То есть берем «от фонаря» чего то измеряем, чего то меняем, чего то получаем, НО НЕПОНЯТНО НИЧЕГО, ЧТО ИЗМЕРЯЕМ, ЗАЧЕМ, ПОЧЕМУ ИМЕННО ЭТО ИЗМЕНЯЕМ И ПОЧЕМУ ИМЕННО ЭТО ИЗМЕРЯЕМ.

**РЕЗЮМЕ: СУЖАЙ И КОНКРЕТИЗИРУЙ ЗАДАЧУ**

Работа над ошибками:

1. Предметная область – визуализация связей людей в социальных сетях.

Предполагается, что аналитику нужно исследовать вопрос:

* общительность конкретного субъекта (количество связей);
* нахождение других субъектов с кем он имеет контакт;
* нахождение субъектов, которые могут иметь общие интересы с несколькими другими субъектами (у исследуемого субъекта есть связи с теми, кто относится к одной группе интересов);
* В случае если граф представляет из себя определенную социальную группу людей:
  + Определить самого главного/весомого субъекта группы (имеет больше всего связей со всеми)
  + Определить насколько один субъект может иметь влияние на второй субъект (дальность в графе)

1. Планирование экспериментов применяется при поиске оптимальных условий.

План: совершить перебор вышеперечисленных экспериментов для локального выявления оптимальных параметров.

Перебор в целом является избыточным для анализа, то есть его данных вполне хватит, так как рассмотрены все варианты параметров.

В конечном итоге результат можно считать успешным, если при проведение финального эксперимента с уже выявленными лучшими параметрами – вероятность правильного ответа будет 0.95

1. Другие измеряемые параметры взять конечно можно, задача специально была сужена для облегчения анализа, иначе задача эргономичного изображения графа стала бы слишком сложной.

Что касается количества экспериментов – тут точно знать не могу, как описать какие виды экспериментов нужны, а какие нет. Я постарался подобрать именно такие эксперименты, которые покрывают задачи аналитика, а дальше сделать полный перебор параметров по этим экспериментам для выявления лучших параметров не пропускаю детали (ведь перебор покрывает все варианты).