Предметная область: анализ графа друзей в социальных сетях.

**Методика нахождения необходимых значений параметров**

Задача аналитика - найти связность двух или нескольких узлов на возможные общие интересы/места проживания/места работы и т.д. В данном случае качество связей игнорируется для упрощения задачи.

Для анализа будущих экспериментов введём следующую функцию:

F = f(p\_1, ..., p\_N), где

F – функция времени решения; решения чего? (\*решения аналитической задачи на графе в зависимости от визуальных параметров графа)

p\_i — набор визуальных параметров графа, наиболее существенно влияющих на восприятие графа и удобство работы с ним. В частности, выбраны следующие параметры:

p\_1 - ширина рамки узла;

p\_2 – коэффициент прозрачности рамки узла;

p\_3 – коэффициент в функции расчёта диаметра узла;

p\_4 – среднее расстояние между связанными узлами;

p\_5 – ширина линий связей;

p\_6 – коэффициент прозрачности линий связей;

Изменяя параметры p\_i, необходимо найти локальныйминимум функции F.

Для нахождения глобального минимума необходимо провести перебор всевозможных значений параметров p\_i и вычислить все соответствующие значения функции F. Для поиска локального минимума необходимо сузить диапазон значений параметров p\_i. Исходя из специфики задачи, примем, что параметры p\_i являются дискретными. В таком случае, для каждого параметра p\_i мы можем указать нижнее и верхнее возможные значения, а также несколько значений внутри этого отрезка. Допустим, что каждый из 6-ти параметров может принять одно из 5-ти возможных значений. В таком случае, в результате полного перебора всевозможных их комбинаций мы получаем 56=15625. Очевидно, что перебрать 15625 комбинаций можно, но крайне затруднительно. Поэтому далее предлагается упрощённая методика минимизации функции, при которой мы допускаем возможность “упустить” наименьший минимум.

Предлагается схема, при которой последовательно для всех параметров p\_i ищутся оптимальные значения (при которых достигается минимум F).

Это возможно сделать с помощью следующего алгоритма:

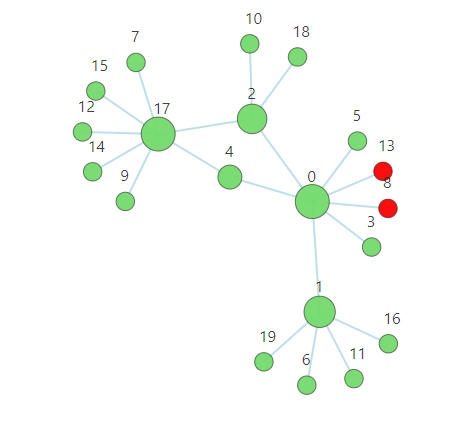
1. Установить случайные значения параметров p\_1, p\_2…p\_6
2. В цикле от i = 0 до i = 6 (то есть в цикле для каждого параметра):
   1. Для параметра p\_i выбрать минимальное значение
   2. В цикле:
      1. С определенным шагом увеличить значение параметра p\_i
      2. Провести эксперимент и записать результаты
      3. Выйти из цикла если параметр p\_i принял максимально заданное значение
   3. Подвести итог, при каком значении параметра p\_i функция F принимает минимальное значение
   4. Установить параметр p\_i, равный полученному ранее значению, для всех следующих экспериментов
   5. Увеличить i на 1 (Перейти к следующему параметру)
3. Найден минимум функции F (по выбранной схеме минимизации) при найденных ранее параметрах p\_1, p\_2,…p\_6

Данный подход достаточно быстро поможет найти совокупность значений параметров p\_1, p\_2,..p\_6, которые дадут удовлетворительное значение функции F.

**Методика проведения экспериментов**

Ещё раз: Задача аналитика найти связность двух или нескольких узлов на возможные общие интересы/места проживания/места работы и т.д. В данном случае качество связей игнорируется для упрощения задачи.

Для описания какие именно эксперименты лучше проводить, следует разбить процесс решения задачи на подпроцессы:

1. При анализе графа аналитик первым делом смотрит насколько далеко интересующие его узлы находятся друг от друга (под расстоянием имеется ввиду минимальное количество узлов между субъектами).
2. Если расстояние не очень велико, то необходимо найти узлы на таком же минимальном расстоянии
   1. Т*о есть если у субъектов есть как минимум 1 общий узел, то нужно искать остальные общие узлы*
   2. *А если, например, у субъектов нет общих узлов, но есть общие на расстоянии двух узлов (друг моего друга), то ищутся все узлы, которые находятся на таком же расстоянии (друзья друзей) и т.д.*
3. Из отобранных общих узлов (если такие есть), для аналитика важно выделить узлы с большим количеством связей (вероятно местный инфлюенсер)

На этом работа аналитика с графом заканчивается, оценивать степень общих интересов и занятий следует уже не из графа.

Анализируя действия аналитика можно выделить несколько подпроцессов и использовать их в экспериментах:

1. 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) – связано с первым пунктом работы аналитика.*
2. Предложить 1 случайный узел и попросить экспериментатора определить все связанные с ним узлы *(Проверяется все ли узлы были определены и за какое время) – связано со вторым пунктом работы аналитика.*
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) – связано со вторым пунктом работы аналитика.*
4. 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) – связано с третьим пунктом работы аналитика.*

Эксперименты подобраны не случайно, каждый из них это более упрощенный этап работы аналитика при выполнении задачи и поиска нужны узлов.