Министерство образования и науки РФ

ФГБОУ ВО «Тверской государственный университет»

Факультет прикладной математики и кибернетики

Направление «Прикладная математика и информатика»

Кафедра информатики

«Минимизация пересечений»

|  |
| --- |
| Выполнил: |
| *Наймушин А. В., студент группы 1сп* |

|  |
| --- |
| Научный руководитель: |
| *Карлов Б.Н., к.ф.-м.н.* |
|  |

Тверь 2017

**Введение**

В период с 5 мая 2017 года до 1 января 2016 года, была пройдена практика в федеральном государственном бюджетном образовательном учреждении высшего образования «Тверской государственный университет» на кафедре информатики.

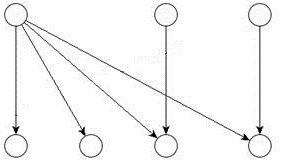
В рамках прохождения данной практики, были поставлены следующие задачи:

* изучение теоретического материала;
* реализация метода физической укладки;
* разработка эвристик, улучшающих качество укладки графа на плоскость.

1. **Минимизация пересечений**

Минимизация пересечений ребер у графа является одним из часто используемых на практике эстетических критериев укладки графа на плоскость. Используя данный критерий можно достичь высокого уровня читабельности графа, так как большую часть его связей будет легче проследить. Кроме этого, задача минимизации пересечений напрямую связана с эффективностью той или иной укладки, хотя в большинстве случаев полностью избавиться от пересечений невозможно.

Проблема минимизации количества пересечений является комбинаторной проблемой, которая состоит в том, чтобы выбрать упорядочение вершин, с наименьшим количеством пересечений. Стоит отметить, что данная задача является NP-полной даже в случае Сигуяма-подобной укладки графа с двумя уровнями. Более того, задача по-прежнему остается NP-полной, если известно оптимальное расположение узлов для одного из уровней.



*(Рис. 1 Пример двухслойного графа)*

Подходы к решению задачи минимизации пересечений могут сильно различаться в зависимости используемой укладки графа. Наиболее часто применяемым на практике методом является метод барицентров, предложенный Сигуямой в 1981 году. Смысл данного метода заключается в том, что положение вершины на плоскости определяется как среднее арифметическое координат её соседей. Таким образом, координаты вершины *u* можно вычислить по следующей формуле:

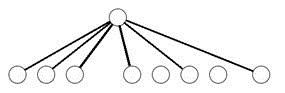
где – степень вершины*,* – множество вершин, для которых есть ребро из *u*, а *x(v)* – координата вершины *v*. Если же после определения координаты двух вершин совпадают, то к координате одной из них прибавляют небольшое значение, чтобы отдалить вершины друг от друга.

После этого, вершины следует упорядочить в соответствии с их барицентрами. Для этого можно воспользоваться любым методом сортировки, например методом быстрой сортировки.

Стоит отметить огромную эффективность данного метода. Время вычисления барицентра одной вершины пропорционально её степени, поэтому барицентры всех вершин можно вычислить за линейное время, а сложность вычислений быстрой сортировки – .

Для метода барицентров существует очень полезная теорема о том, что если для графа *G* возможно размещение вершин без пересечений каких-либо ребер, то метод его обязательно найдет.

Из недостатков данной эвристики можно назвать тот факт, что она предназначена только для иерархической укладки, поэтому использовать данный подход в любых других укладках невозможно.



*(Рис. 2 Иллюстрация метода барицентров)*

Кроме обычного метода барицентров, также известны несколько его модификаций. Наиболее популярным из них является метод медиан. В отличие от метода барицентров, каждой вершины с четной степенью присваивается среднее арифметическое значение позиций вершин соседнего уровня, а для вершин с нечетными степенями – среднее значение координат соседей.

Изучая проблему минимизации пересечений укладки графа, нельзя не затронуть такое понятие «число пересечений».

**2. Число пересечений**

Отправной точкой к изучению задачи минимизации значения пересечений стала задача о кирпичной фабрике, поставленная венгерским математиком Палом Тураном еще в 1945 году. Во время Второй мировой войны, математику приходилось работать на кирпичной фабрике, толкая телегу груженную кирпичами. Всякий раз телегу было толкать тяжелее на месте пересечения колей. Это и привило Турана к постановке задачи о нахождении минимального числа пересечений в рисунке полного графа.

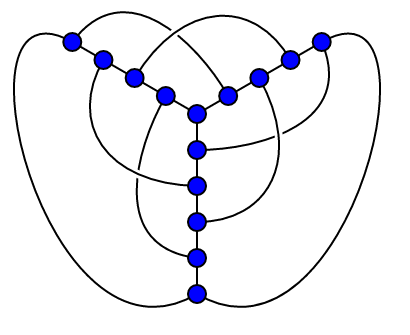
Определим понятие числа пересечений графа. Числом пересечений графа G (с англ. *crossing number*) называют минимально возможное число пересечений ребер графа любой укладки. Обозначается это число как *cr(G)*.

Стоит отметить, что понятие числа пересечений является одной из характеристик степени непланарности (от англ. *degree of non-planarity*) графа.

Как было показано Гери и Джонсоном в 1983 году, проблема нахождения числа пересечений графа является NP-полной проблемой. Исследователи смогли доказать этот факт, сведя данную задачу к задаче о линейном размещении графа на плоскость, которая, как ими ранее была доказана, является NP-полной.

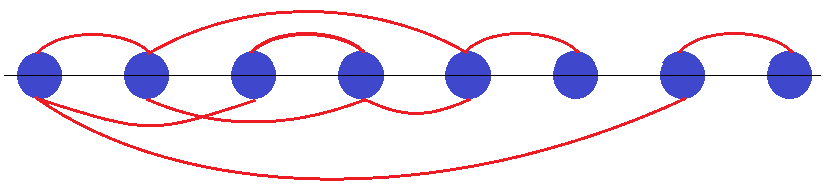
Помимо обобщенного понятия числа пересечений, в литературе очень часто встречаются и вариации определений для различных укладок. Одно из наиболее часто встречающихся определений является прямолинейное число пересечений. Данное определение полностью идентично определению обычного числа пересечений, однако на него накладывается дополнительное требование в виде использования только укладок с прямолинейными ребрами. Прямолинейное значение пересечений обозначается как .

Можно определить несколько полезных теорем, которые позволяют установить связь между обобщенным числом и прямолинейным. Одной из таких теорем является теорема Бьенстока и Дина. Она гласит, что если , то . Также ими была доказана и другая теорема, которая гласит, что для любого числа существует такой граф *G* для которого = 4, однако .



*(Рис. 3 Граф Хивуда, для которого )*

Другим распространенным определением числа пересечений является линейное число пересечений. Пусть дан некий граф *G = (V, E).* Линейным числом пересечений называют наименьшее число пересечений, среди всевозможных линейных укладок графа *G* вдоль прямой линии. Обычно в литературе такое число обозначается как .



*(Рис. 4 Пример линейной укладки графа на горизонтальную линию)*

Для линейного числа пересечений также можно установить взаимосвязь с обобщенным числом пересечений. Эта связь была установлена Николсоном в своей теореме. Согласно данной теореме, для любого плоского графа можно построить изоморфный ему граф *G’* (т.е. между двумя данными графами можно построить взаимно-однозначное соответствие) такой, что все его вершины будут расположены на одной прямой, а ребра будут представлены в виде дуг.

Как и в предыдущих случаях с другими вариациями определений, проблема проверки того, что не превосходит некоторое число *k* также является NP-полной проблемой.

На практике, для того чтобы найти число пересечений какого-нибудь графа, очень часто прибегают к различным эвристикам. И хотя они не позволяют найти точное значение, однако они позволяют, за разумное время, получить неплохое его приближение. Методика одной из эвристик заключается в следующем. Пусть дан некий граф *G=(V,E).* Из данного графа удаляют некоторое подмножество ребер так, чтобы новый граф , являлся планарным. Затем из множества *F* выбирают ребро *f* и пытаются найти укладку графа с наименьшим числом пересечений (обычно сводя данную задачу, к задаче о нахождении кратчайшего пути). Далее эта процедура повторяется до тех пор, пока множество .

Используя понятие числа пересечений для задачи минимизации пересечений в графе можно сформулировать аналогичную задачу о числе пересечений. Пусть на вход подается некий граф *G* и число *k > 0*. Задача заключается в ответе на вопрос: можно ли найти такую укладку графа, при которой ?

Многими исследователями были предприняты многочисленные попытки оценить значения границ числа пересечений графа. Одной из известных теорем в теории графов является теорема польского математика Казимежа Заранкевича. Занимаясь изучением проблемы о кирпичной фабрике, он смог найти выражение, которое позволяло точно определить значение числа для любого двудольного графа. К сожалению, в его доказательстве нижней границы была найдена ошибка, однако он смог точно найти верхнюю границу числа пересечений:

В 1960 году была впервые опубликована задача о нахождении числа пересечений полного графа. Энтони Хилл и Джон Эрнест смогли не только сформулировать саму задачу, но также высказали свою гипотезу о том, что верхняя граница числа пересечений графа равняется:

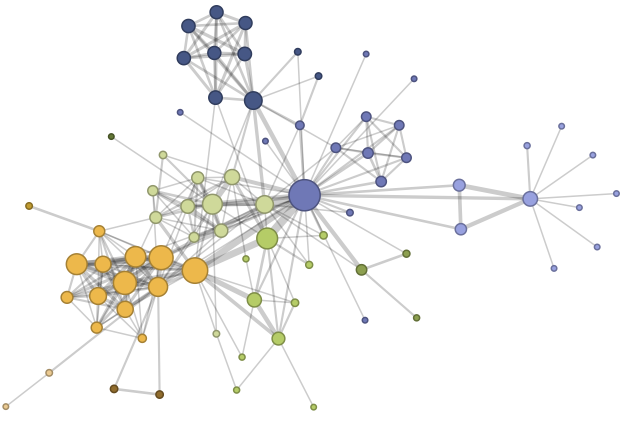
К 2017 году, значения чисел пересечений известны лишь для небольшого числа семейств графов. Были предприняты многочисленные попытки нахождения нижних границ, однако точного результата найдено не было.

**3. Укладка с применением физической модели**

Одними из наиболее применяемых методов укладок графов на плоскость являются методы с применением физических моделей.

Данные методы появились в работах Тутте в 1963 году, который показал, что графы с большим количеством узлов могут быть хорошо уложены на плоскость с применением физических моделей, добавив каждому ребру различные силы и позволив системе самой достигнуть состояния равнения.

Смысл данных методов заключается в том, что для каждого ребра или узла, задаются некоторые силы, обычно согласно закону Гука, т.е. , где ***F*** – сила упругости, ***k*** – жесткость тела, а – деформация тела. Затем данный алгоритм пытается минимизировать энергию полученной системы для получения самой укладки.



*(Рис. 5 Укладка с применением физической модели)*

Сформулируем сам алгоритм укладки с применением физической модели. Пусть на вход подается некий граф .

1. **Инициализация**. Каждому узлу графа задается некоторая случайная точка на координатной плоскости.

**2**. **Итерации алгоритма**. Для каждого узла имеем векторную переменную *displ*, которая характеризирует смещение узла на данной итерации. Перед началом каждой итерации, данная переменная для всех узлов равна 0.

**2.1**. **Вычисление отталкивающих сил**. Для каждой пары узлов, *,* вычисляем значение смещения по следующим формулам:

где – некоторые переменные, ограничивающие площадь, в которой будет размещен граф, *|x|* - оператор нормы векторной величины *x.*

**2.2**. **Вычисление притягивающих сил**. Для каждого ребра , произведем корректировку смещений узлов *u* и *v* по следующим правилам:

где *u.pos –* векторная переменная, определяющая позицию узла *u* на плоскости.

**2.3**. **Корректировка позиции узлов**. Для каждого узла произведем коррекцию позиции по следующему правилу:

**2.4.** **Критерий останова**. Если среднеквадратичное значение всех смещений достигло некоторого порогового значения, то выходим из алгоритма, иначе переходим на пункт 2.

Преимущества укладки с использованием физических моделей заключаются в том, что они способны оптимизировать сразу несколько эстетических критериев, таких как максимизация симметрии или равномерное распределение вершин. Также, данные методы привлекательны своей гибкостью. Можно с легкостью применять различные эвристики, улучшающие качество укладок, при этом, практически не изменяя самого алгоритма метода.

К недостаткам данных методов, можно отнести то, что они имеют достаточно высокое время работы: , где *I* – число итераций, *n* – количество вершин в графе. Однако на практике, методы данного класса сходятся достаточно быстро.

Главным недостатком данных методов является проблема локального минимума. Это объясняется тем, что данные методы стремятся уменьшить общую энергию системы и очень часто, найденный минимум не является глобальным минимумом функционала энергии. Это приводит к ухудшению качеству укладки.

**4. Эвристики улучшения качества укладки**

Проблема локального минимума представляет собой одну из важнейших проблем, с которой приходится сталкиваться при использовании укладки графа с помощью физических моделей. Чтобы попытаться улучшить качество укладки графа, в данной работе были рассмотрены несколько возможных эвристик.

Идея первой эвристики была позаимствована из другой математической задачи, в которой также встречается проблема нахождения локального минимума. Этой задачей является обучение нейронной сети методом обратного распространения ошибки.

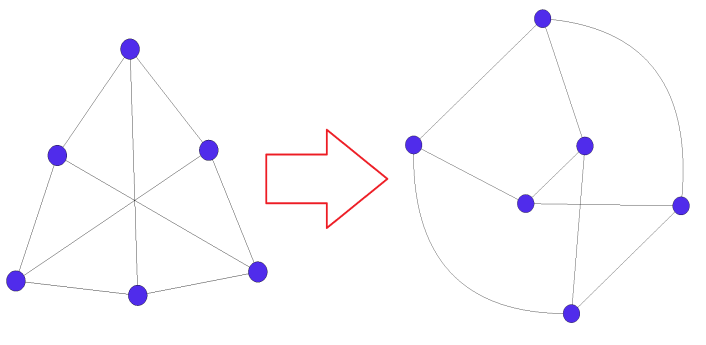
Для корректного обучения нейронных сетей, в конце итерации обучения вычисляется значение среднеквадратичной ошибки, которая была допущена при распознании образов из множества обучающих примеров. Обучение нейронной сети следует остановить в том случае, когда изменение этой самой ошибки между итерациями становится меньше некоторого порогового значения. К сожалению, поверхность функционала ошибки содержит в себе очень много точек экстремума, в частности, различных локальных минимумов. Как только нейронная сеть достигает одну из таких точек, разница в изменении значения ошибки резко падает, что приводит к тому, что обучение сети заканчивается слишком рано и, соответственно, данная сеть не сможет выдавать корректный отклик.

Для решения проблемы стабилизации функционала ошибки в локальных минимумах, одним из простейших способов является «встряска» весов (от англ. jog of weights). Идея данного метода заключается в том, что при каждой стабилизации значения ошибки, следует производить случайные модификации всех весовых коэффициентов в достаточно большой окрестности текущего значения. Это приведет к тому, что метод градиентного спуска начнет поиск минимального значения из других точек и, возможно, сможет обойти «ловушку» локального минимума, в которую он попал в прошлый раз.

Идея данного подхода была рассмотрена и в задаче нахождения укладки графа на поверхность. На каждом этапе корректировки позиции узлов графа, будем также сдвигать узлы на небольшое расстояние, в случайном направлении. Это позволит алгоритму укладки начать поиск оптимального расположения узлов из новых точек и, возможно, избежать «ловушки» локального минимума.

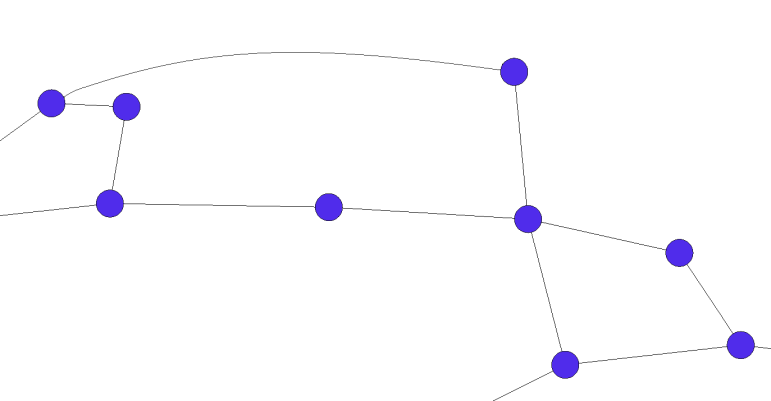
К сожалению, на практике данный подход оказал слабое влияние на качество укладки заданного графа. Лишь в нескольких случаях, эвристика помогла уменьшить количество пересечений ребер конечной укладки. Кроме этого, для сходимости алгоритма потребовалось гораздо больше времени, чем для обычного алгоритма укладки.

Другая рассмотренная эвристика направлена на изменение представления ребер графа не в виде прямых линий, а в виде дуг. Смысл данной эвристики заключается в следующем. Пусть имеется некий граф , а также пара пересекающихся ребер *(u, v)* и *(w, t),* где . Тогда выберем одно из ребер, например *(u, v),* разобьем его на два новых ребра *(u, v’)* и *(v’, v),* добавив в граф новый фиктивный узел *v’*. Затем, запустим процесс нахождения укладки для нового графа . После того, как укладка нового графа будет найдена, заменим ребра графа сплайнами Безье, используя добавленные фиктивные вершины как точки характеристического многоугольника для данного ребра.



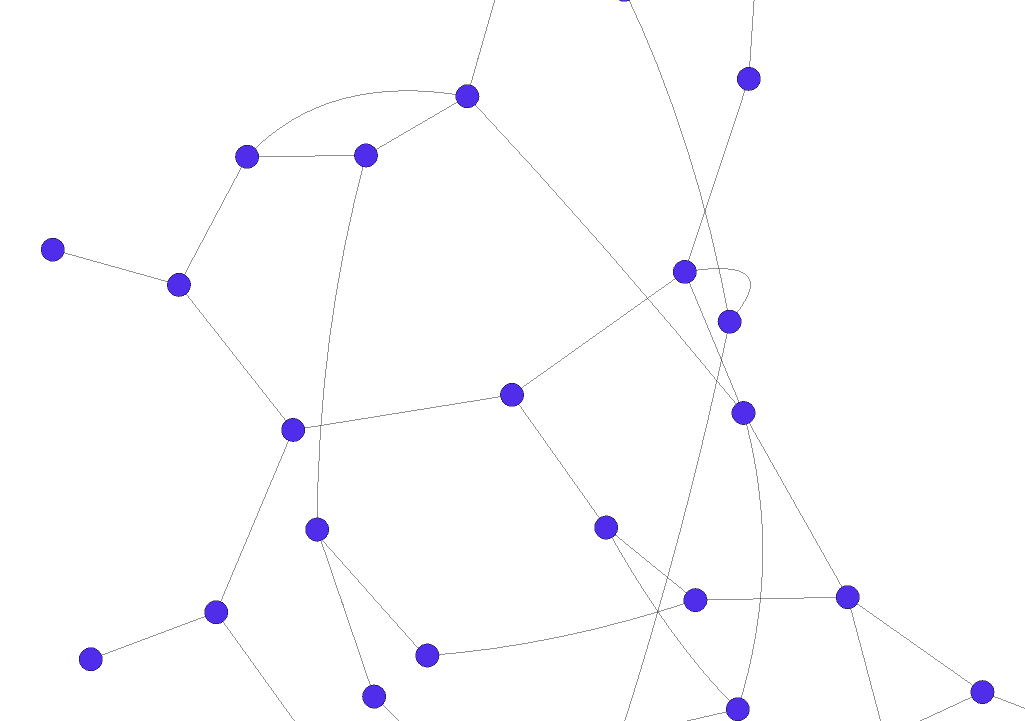
*(Рис. 6 Граф «Домики и колодцы» до и после применения эвристики)*

На практике, данная эвристика показала довольно неплохие результаты как на графах малой размерности, так для графов с большим количеством вершин.



*(Рис. 7 Пример для произвольного графа)*

К недостаткам данного подхода, можно отнести то, что для сходимости алгоритма укладки потребуется гораздо больше времени, так как граф практически полностью меняет расположение своих узлов.



*(Рис. 8 Пример неудачной укладки)*

Кроме этого, частое использование данной эвристики приводит к резкому возрастанию длины ребер, что также негативно сказывается на качестве конечной укладки.

**Заключение**

В рамках данной практике, был изучен метод укладки графов на плоскость с использованием физических моделей. Была создана программа, которая позволяет пользователю задавать произвольный граф с последующим применением самого метода укладки. Были рассмотрены несколько возможных эвристик, направленных на улучшение качества конечной укладки графов. Указанные эвристики также были реализованы в указанной выше программе.

В дальнейшем планируется провести исследование на тему нахождения наилучшей укладки графа с фиксированным положением узлов.

**Список литературы**

1. Теория графов, Харари Ф., пер. с англ. В. П. Козырева, под ред. Г. П. Гаврилова. Издательство «Мир», Москва, 1973
2. Bounds for rectilinear crossing numbers. D. Bienstock, N. Dean., J. Graph Theory, 17(3):333–348, 1993
3. Crossing number is NP-complete. M. Garey, D. S. Johnson. SIAM Journal on Algebraic and Discrete Methods, 4(3):312-316, 1983.
4. Drawing graphs: methods and models / Michael Kaufmann, Dorothea Wagner (ed.). – Berlin, Heidelberg, New York, Barcelona, Hong Kong, London , Milan , Paris, Singapore , Tokyo : Springer, 2001
5. Efficient Planarity Testing, Hopcroft J., Tarjan R.E., J. ACM, vol. 21, no. 4, pp. 549-568, 1974.
6. Graph Theory with Applications, Bondy J.A., Murty U.S.R., North Holland, Amsterdam, 1976
7. Methods for visual understanding of hierarchical systems, Sugiyama K., Tagawa S., Toda M. // IEEE Trans. Syst., Man,and Cybern. — 1981. — Vol. 11, N. 2. — P. 109–125.
8. NP-Completeness Column: an Ongoing Guide, Johnson D.S., J. of Algorithms, vol. 3, no. 1, pp. 89-99, 1982.
9. NP-Completeness Column: an Ongoing Guide, Johnson D.S., J. of Algorithms, vol. 5, no. 2, pp. 147-160, 1984.
10. On Drawing a Graph with the Minimum Number of Crossings, Ferrari D., Mezzalira L., Technical Report n. 69-11, Istituto di Elettrotecnica ed Elettronica, Politecnico di Milano, 1969.
11. On Straight Lines Representation of Planar Graphs, Fary I., Acta Sci. Math. Szeged, vol. 11, pp. 229-233, 1948.
12. On The Crossing Numbers of Cartesian products with Trees. D. Bokal, Graph Theory 56 (2007), 287–300.
13. Permutation procedure for minimizing the number of crossings in a network. T.A.J. Nicholson. IEE Proceedings, 115:21–26, 1968.
14. Toward a theory of crossing numbers, W. T. Tutte, Combin. Theory 8 (1970), 45–53.