Задача Штейнера

Самсонов Владислав
395 группа ФИВТ ПМИ
e-mail: vvladxx@gmail.com

29 ноября 2015

1 Введение

В данной работе рассматривается задача построения дерева Штейнера. Кроме того, будет рассмотрен метрический вариант задачи в пространстве (X, ρ) с произвольной метрикой и решена задача построения дерева Штейнера в евклидовом пространстве $\mathbb{E}^2$. Алгоритм, представленный в работе, имеет оптимальное время работы $O(n \log n)$ и показатель аппроксимации $a = \frac{2\sqrt{3}}{3} \approx 1.15...$.

2 Теоретические сведения

2.1 Базовые определения

$\mathbf{MST}(G)$	Минимальное остовное дерево для графа G.
$\mathbf{ST}(G, V_0)$	Дерево Штейнера графа G.
$\mathbf{EST}(P)$	Евклидово дерево Штейнера для набора точек P.
$\mathbf{DT}(P)$	Триангуляция Делоне для набора точек P.

Деревом Штейнера для графа $G = (V, E)$ и набора вершин $V_0 \subseteq V$ называется дерево $T \subseteq G$ минимального веса, проходящее по всем вершинам V_0.

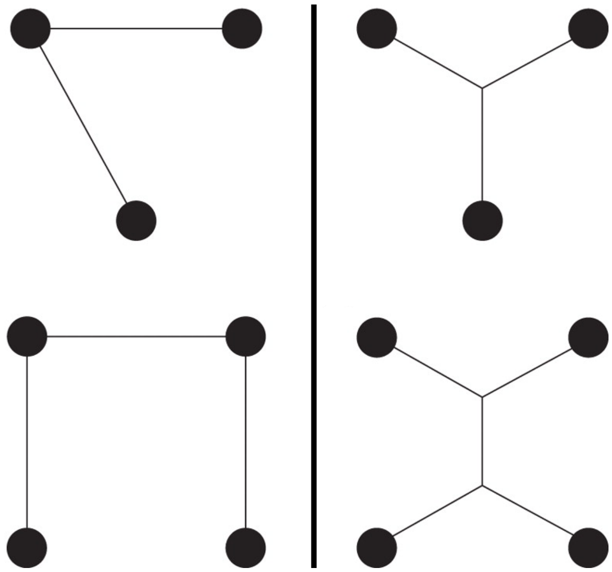

Рис. 1: MST (слева) и дерево Штейнера (справа) в $\mathbb{E}^2$

Метрическим деревом Штейнера в пространстве (X, ρ) для конечного набора точек $P \subset X$ называется $\mathbf{ST}(G, P)$, где $G = (V, E)$ – полный граф, $V = X$, а вес ребра $(v, u) \in X \times X$ определяется метрикой $\rho(v, u)$.

Евклидовым деревом Штейнера для конечного набора точек $P \subset \mathbb{E}$ называется метрическое дерево Штейнера в евклидовом пространстве с евклидовой метрикой.

Минимальным остовным деревом для графа $G = (V, E)$ называется дерево $T \subseteq G$ минимального веса, проходящее по всем вершинам V.

Триангуляцией Делоне для набора точек P в $\mathbb{E}^2$ называется разбиение плоскости на непересекающиеся треугольники с вершинами из P таким образом, что все точки P принадлежат хотя бы одному треугольнику и для любого треугольника t триангуляции выполняется критерий Делоне: никакая из точек $p \in P$ не лежит строго внутри окружности, описанной вокруг t.

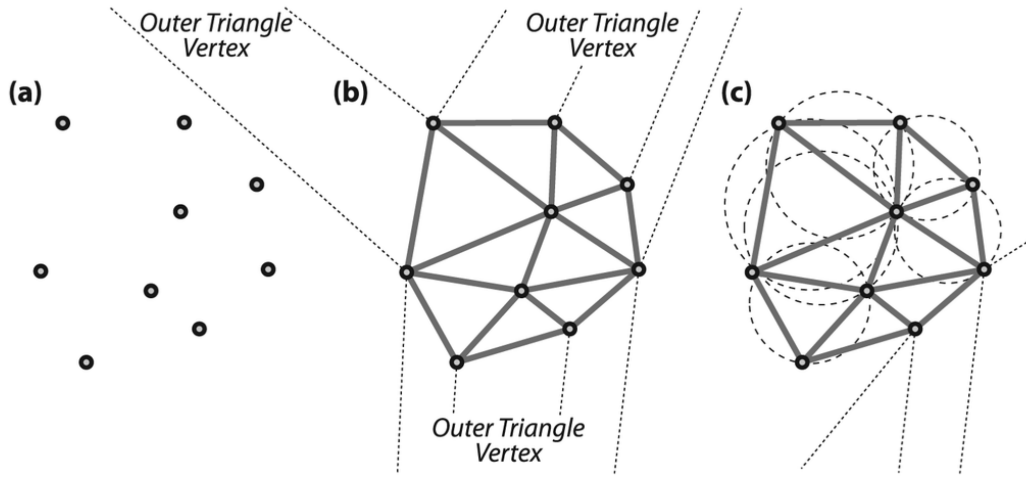

Рис. 2: Триангуляция Делоне

2.2 Задача Штейнера

Далее будем рассматривать соответствующую задачу распознавания: существует ли дерево Штейнера весом не более k.

Теорема 1. *Все постановки задачи Штейнера (общая/метрическая/евклидова) лежат в* **NP**.

Доказательство. В качестве сертификата можно предъявить искомый набор рёбер. Все проверки на корректность дерева и сумму рёбер выполняются очевидным образом и за полиномиальное время. □

Теорема 2. $\mathbf{ST}(G, V_0)$ *является* **NP***-полной.*

Доказательство. Из теоремы 1 и следующей теоремы очевидным образом следует утверждение этой теоремы, поскольку метрический вариант задачи является частным случаем общей постановки и сводится к общей. Будем сразу доказывать **NP**-полноту метрической задачи Штейнера. □

Теорема 3. *Метрическая задача Штейнера является* **NP***-полной.*

Доказательство. Покажем, что задача о вершинном покрытии ($\mathbf{VC}(G)$) полиномиально сводится к метрической задаче Штейнера.
Пусть дан граф $G = (V, E)$ из задачи о вершинном покрытии. Построим

полный граф G' следующим образом: в нём $|V|+|E|$ вершин, v' соответствуют $v \in V$ и v'_{uv} соответствуют $(u,v) \in E$.
Определим расстояния между вершинами G':

- $\forall (u,v) \in E : \rho(u', v'_{uv}) = \rho(v', v'_{uv}) = \rho(v'_{uv}, u') = \rho(v'_{uv}, v') = 1$
- $\forall u, v \in V : \rho(u', v') = \rho(v', u') = 1$
- $\forall u \in V \ \forall (v,z) \in E$, т.ч. $u \neq v \neq z$: $\rho(u', v'_{vz}) = \rho(v'_{vz}, u') = 2$
- $\forall e, e' \in E$: если рёбра e и e' имеют общую вершину, то $\rho(v'_e, v'_{e'}) = \rho(v'_{e'}, v'_e) = 2$, иначе $\rho(v'_e, v'_{e'}) = \rho(v'_{e'}, v'_e) = 3$

Расстояния, определённые таким образом, удовлетворяют неравенству треугольника, а значит G' – корректная постановка метрической задачи Штейнера.

Лемма 1. *Если в G есть вершинное покрытие размера k, то в G' есть дерево Штейнера весом $m+k-1$.*

Доказательство. Пусть $S \subset V$ – вершинное покрытие. Рассмотрим точки $\{v' \in V' : v \in V\}$ и точки $\{v'_{uv} \in V' : (u,v) \in E\}$ и посмотрим на все рёбра длины 1 между ними. В результате такого описания получился индуцированный связный подграф G', потому что любая вершина u' соединена ребром длины 1 с любой другой v', а каждая ω'_{uv} соединена с u'. Любое остовное дерево этого связного подграфа в G' имеет $m+k-1$ рёбер длины 1, а значит условие леммы выполнено. $\square$

Лемма 2. *Если в G' есть дерево Штейнера весом $\leq m+k-1$, то в G есть вершинное покрытие размера k.*

Доказательство. Пусть T – дерево Штейнера весом $\leq m+k-1$. Сперва мы изменим дерево таким образом, что у него не будет рёбер длины 2 или 3. Будем повторять следующие действия:

- Если есть ребро длины 2 между w' и v'_{uv}, мы удаляем ребро (ставим ∞) из дерева и добавляем 2 новых ребра длины 1: (w', u') и (u', v'_{uv}).
- Если есть ребро длины 2 между v'_{uv} и v'_{vw}, мы удаляем ребро (ставим ∞) из дерева и добавляем 2 новых ребра длины 1: (v'_{uv}, v') и (v', v'_{vw}).

- Наконец, если есть ребро длины 3 между v'_{uv} и v'_{wz}, мы удаляем ребро (ставим ∞) из дерева и добавляем 3 новых ребра длины 1: (v'_{uv}, v'), (v', w') and (w', v'_{wz}).

Мы повторяем действия, описанные выше, пока в графе есть рёбра длины 2 или 3. Этот процесс не увеличивает стоимость и в результате получается связный граф. Мы можем получить дерево удалением рёбер (и уменьшением стоимости) если нужно.
В полученном дереве есть только рёбра длины 1 и его вес $\leq m + k - 1$, т.е. оно содержит $\leq m + k$ вершин. m вершин v'_{uv} должны быть в дереве, значит в дереве $\leq k$ вершин v'. Пусть S – множество этих вершин. Можно заметить, что это и есть вершинное покрытие для G. В самом деле, для любого ребра $(u, v) \in G$ вершина v'_{uv} соединена с v' в дереве только рёбрами длины 1, что означает, что либо u', либо v' в дереве, и либо u, либо v в S. $\square$

Таким образом, мы полиномиально свели $\mathbf{VC}(G)$ к метрической задаче Штейнера. $\square$

Теорема 4. *Длина MST в метрической задаче Штейнера не более удвоенного веса минимального дерева Штейнера.*

Доказательство. Обозначим вес минимального дерева Штейнера за L_s. Заменим в этом дереве каждое ребро на двойное ребро. Получили дерево с чётным количеством рёбер, а значит в нём можно выделить эйлеров цикл. Теперь склеим двойные рёбра обратно в одно. Эйлеров цикл будет проходить по всем вершинам графа G (возможно по несколько раз) и иметь длину $2L_s$. Но, с другой стороны, граф G полный, а значит в нём существует гамильтонов цикл минимального веса L_g. Тогда $L_g \leq 2L_s$. Обозначим длину MST за L_m. Если удалить одно ребро из минимального гамильтонова цикла, то получится какое-то остовное дерево длины L, при этом $L \leq L_g$ (т.к. удалили ребро) и $L_m \leq L$ (т.к. L_m – длина минимального остовного дерева).
В итоге получаем следующую цепочку: $L_m \leq L \leq L_g \leq 2L_s$. $\square$

2.3 Алгоритм 2-аппроксимации для метрической задачи Штейнера

Пусть в пространстве (X, ρ) требуется решить метрическую задачу Штейнера для графа $G(X, E)$ и конечного набора точек P.

Тогда достаточно построить минимальное остовное дерево для индуцированного подграфа G вершинами P.

Теорема 5. *Данный алгоритм имеет показатель аппроксимации не больше 2.*

Доказательство. Прямое следствие из теоремы 4:

$$a = \sup \frac{p^*}{p} = \sup \frac{L_m}{L_s} \leq 2$$

Тут p^* – найденное решение, p – минимальное решение. □

Теорема 6. *Данный алгоритм имеет сложность $O(|P|^2)$.*

Доказательство. Так как индуцированный подграф полный и содержит $C^2_{|P|} \sim |P|^2$ рёбер, то задачу построения MST можно решать алгоритмом Прима для полных графов за $O(|P|^2)$. □

2.4 Алгоритм для евклидовой задачи Штейнера

Рассмотрим задачу нахождения дерева Штейнера в $\mathbb{E}^2$, то есть на плоскости. В этом случае, разумеется, можно решать задачу алгоритмом для метрической задачи Штейнера. Однако, в этом частном случае алгоритм можно ещё немного улучшить, как в показателе аппроксимации, так и по времени работы.

Теорема 7 (Нижняя оценка для показателя аппроксимации)**.** $a \geq \frac{2\sqrt{3}}{3}$.

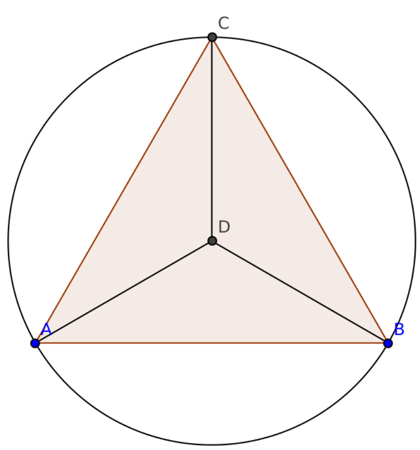

Рис. 3: Пример

Доказательство. Рассмотрим конфигурацию точек $P = \{A, B, C\}$, показанную на рисунке. Точки ABC образуют равносторонний треугольник со стороной 1. Деревом Штейнера будут отрезки AD, BD, CD длиной $\frac{\sqrt{3}}{3}$ каждый, то есть вес дерева Штейнера $\sqrt{3}$. Но вес минимального остова 2, откуда получаем $a \geq \frac{2}{\sqrt{3}} = \frac{2\sqrt{3}}{3}$. □

Теорема 8 (Верхняя оценка для показателя аппроксимации)**.** $a \leq \frac{2\sqrt{3}}{3}$.

Это утверждение долгое время оставалось открытой гипотезой. Экспериментально гипотеза хорошо подтверждается. Доказательство этой теоремы можно посмотреть в [2], но в нём была найдена ошибка [3]. Поэтому, оставим эту теорему без доказательства.

Теперь будем улучшать время работы.

Лемма 3. *В* $\mathbf{DT}(P)$ *не более* $3|P|-3$ *рёбер и не более* $2|P|-2$ *треугольников.*

Доказательство. Пусть T – количество треугольников, E – количество рёбер. Все рёбра принадлежат ровно двум треугольникам кроме тех K рёбер, которые лежат на выпуклой оболочке $\mathbf{CONV}(P)$ – в этом случае, очевидно, ребро принадлежит только одному треугольнику.

$$2E - K = 3T \implies 2E \geq 3T$$

Далее, пользуемся $|P| - E + (T+1) = 2$ (формула Эйлера для планарных графов), получаем нужные соотношения. □

Лемма 4. $\mathbf{MST}(P) \subseteq \mathbf{DT}(P)$

Доказательство. См. [4] □

Лемма 5. $\mathbf{DT}(P)$ *можно построить за* $O(|P| \log |P|)$.

Доказательство. См. [5] □

Теорема 9. $\mathbf{MST}(P)$ *для* $|P|$ *точек в* $\mathbb{E}^2$ *можно построить за* $O(|P| \log |P|)$.

Доказательство. Прямое следствие из леммы 3-5. Строим $\mathbf{DT}(P)$. Теперь из $\mathbf{DT}(P)$ мы можем выделить $\mathbf{MST}(P)$. Так как по лемме 3 в полученном графе рёбер не слишком много, то достаточно воспользоваться алгоритмом Прима/Краскала и получить требуемый результат. □

Таким образом, полученный алгоритм имеет хороший показатель аппроксимации $a = \frac{2\sqrt{3}}{3} \approx 1.15...$ и оптимальное время работы.
Оптимальность следует из того, что с помощью евклидова дерева Штейнера можно сортировать действительные числа, расположив их на одной прямой. В результате построения дерева штейнера для такой конструкции соседние числа будут соединены ребром, то есть получится двусвязный список отсортированных чисел. Как известно, $N \log N$ – нижняя граница для сортировки N произвольных действительных чисел. Те же самые рассуждения применимы и для минимального остовного дерева. Это значит, что представленный алгоритм является оптимальным.

3 Тестирование

Тестирование проводилось на точках, сгенерированных случайно и равномерно распределённых внутри квадрата $[0, 1] \times [0, 1]$. Программа вычисляла аппроксимацию веса дерева Штейнера. Ошибка вычисляется относительно точного ответа по формуле $1 - \frac{A}{A^{opt}}$. Для большого количества точек точный ответ не известен (слишком долго считается) и ошибка не указана.

Число точек	Время работы (сек.)	Ошибка
10	0.0001	0.12
10^2	0.0002	0.09
10^3	0.0033	0.08
10^4	0.0401	—
10^5	0.1713	—

4 Реализация на языке C++

Исходный код можно найти в *./src*.
main.cpp — основная программа.

Список литературы

[1] U.C. Berkeley — CS172: Automata, Computability and Complexity, Professor Luca Trevisan. 2015.

[2] A.O. Ivanov and A.A. Tuzhilin — Minimal networks: the Steiner problem and its generalizations, CRC press, 1994. Page 167, Theorem 3.1.

[3] A.O. Ivanov and A.A. Tuzhilin — The Steiner Ratio Gilbert-Pollak Conjecture Is Still Open, Algorithmica February 2012, Volume 62, Issue 1, pp 630-632.

[4] Godfried T. Toussaint — The relative neighbourhood graph of a finite planar set, Pattern Recognition Vol. 12, Issue 4, 1980, Pages 261–268.

[5] D.T. Lee, B.J. Schachter — Two algorithms for constructing a Delaunay triangulation, International Journal of Computer & Information Sciences, June 1980, Vol. 9, Issue 3, pp 219-242.