Задача для собеседования на кафедру

Игорь Вожга Б05-872

Май 2020

Задача TSP

Формулировка: Дана карта, на которой отмечено N поселений. Поселения находятся на различном расстоянии друг от друга. Торговец отправляется из поселения N_0. Предложите алгоритм, который позволит найти оптимальный маршрут для обхода всех N поселений и вернуться в точку старта.

Ответ должен содержать один или несколько алгоритмов (можно псевдокод) и пояснение о эффективности данных решений.

Решение:

Для решения задачи будем использовать жадный алгоритм greedy и алгоритм $2 - opt$, реализующий 2-оптимальную эвристику.

Для начала примитивынй алгоритм greedy построит цикл, который будет близок к тому, к чему мы стремимся. После этого полученный greedy цикл я улучшу спомощью $2opt$.

2-оптимальная эвристика основана на построении окрестности для данного тура τ. То есть для τ строится множество туров τ' полученных из τ спомощью удаления двух рёбер (a, b) и (c, d) и добавления рёбер (a, c) и (b, d).

А алгоритм $2 - opt$ являеятся реализацией данной эвристики для задачи TSP. Построим псевдокод:

(Так как мы в итоге строим цикл, то там по факту не важно с какой позиции мы стартуем, поэтому примем N_0 за единицу)

$In : N$ - колво поселений, множество поселений $P = \{p_1, \dots, p_N\}$; Координаты поселений (x_i, y_i);

Первоначальный тур: $\tau = (p_{0_1}, \dots, p_{0_n})$

Out : Тур $T = (p'_1, \dots, p'_n)$

$begin$: Обозн $Pairs_0 = \{(i, j) | i, j \in 1, \dots, n; i \neq j\}$ - множество пар поселений

$Pairs = Pairs_0$

$repeat$

$Pairs - (a, b)$

$t' = \{p_1, \dots, p_{a-1}, p_b, p_{b-1} \dots, p_{a+1}, p_a, p_{b+1} \dots\}$

$if\ length(\tau') < length(\tau)\ then$ - где $lentgh$ - длина тура

$begin$:

$\tau = \tau'$

$Pairs = Pairs_0$

end;

$until\ |Pairs| = 0$ - пока не переберём все пары для текущего тура

$return\ \tau'$

end;

Оптимальность алгоритма:

Для начала докажем, что если для входа x алгоритм $2 - opt$ нашёл тур $T = (p_1, \dots, p_n)$, то количество q пар (a, b) поселений для которых

$\rho(a, b) > \frac{2l^*(x)}{\sqrt{i}}\ \forall i \in 1, \dots, n \to q < i$ (1)

(где $l^*(x)$ - оптимальная длина цикла являющегося решением для данного входа x)

Допустим, что это не так, тогда $q \geq i$. Рассмотрим пары поселений (a, b) удовлетворяющие условию (1).

Покажем, что число конечных точек b ограниченно.

Рассмотрим окружность радиуса $\frac{l^*(x)}{\sqrt{i}}$ и предположим, что точки $b_1, \dots b_s$ лежат в данной окружности. Причём $s \geq \sqrt{i}$. А точки $a_1, \dots a_s$ - это соответствующие им начальные.

Тогда по предположению расстояние между $b_1, \dots b_s$ не больше $\frac{2l^*(x)}{\sqrt{i}}$. А из этого будет следовать, что между начальными $a_1, \dots a_s$ расстояние не меньше $\frac{2l^*(x)}{\sqrt{i}}$, потому что иначе удалив рёбра (a_1, b_1) и (a_1, b_1) и вставив два других: (a_1, a_2) и (b_1, b_2) В T получили бы более короткий тур $'$, но это противоречит факту, что T - локально оптимальный тур полученный алгоритмом $2 - opt$. Поэтому мы доказали, что существует $s > \sqrt{i}$ точек на расстоянии не меньше $\frac{2l^*(x)}{\sqrt{i}}$. А значит оптимальный цикл на поселениях $a_1, \ldots a_s$ имеет длину не меньше чем $2l^*(x)$. Однако из неравенства треугольника мы знаем, что если $a_1, \ldots a_s$ является подмножеством входного множества поселений P, то тк добавление точек не может уменьшить длину тура, то длина тура на точках $a_1, \ldots a_s$ должна быть меньше $l^*(x)$. А значит мы пришли к противоречию. То есть внутри такой окружности может лежать $s \leq \sqrt{i}$ конечных точек $b_1, \ldots b_s$.

Теперь покажем, что когда $q \geq i$ мы можем построить такое множество P', что $|P'| \geq i$ и любые его точки лежат на расстоянии как минимум $\frac{l^*(x)}{\sqrt{i}}$.

Для этого мы возьмём множество всех поселений P и будем добавлять в P' произвольные конечные точки b_i из P и удалять из P все точки находящиеся в окрестности радиусом $\frac{l^*(x)}{\sqrt{i}}$. Таким образом, исходя из предыдущего пункта так как мы могли удалить $s \leq \sqrt{i}$ конечных точек, то в построенном множестве P' не меньше $\sqrt{i}$ точек.

Таким образом, в P' расстояние между точками $> \frac{l^*(x)}{\sqrt{i}}$, точек $> \sqrt{i}$ значит длина цикла на таком множестве P' $l > l^*$, что противоречит неравенству треугольника, тк P' - подмнво P. Таким образом мы доказали утверждение, что количество q пар (a, b) поселений для которых

$\rho(a, b) > \frac{2l^*(x)}{\sqrt{i}}$ $\forall i \in 1, \ldots, n \to q < i$

Теперь, если просуммировать по i для n точек и вместо рёбер уже рассматривать полный цикл, то получим выражение, определяющее оптимальность $2 - opt$:

$$L^{2opt} \leq 4 \cdot OPT\sqrt{n}$$

(где OPT - оптимальный размер цикла, L^{2opt} - размер цикла, который вернул алгоритм $2opt$)

Также следует сказать про алгоритм Кристофидеса, дающий решения в пределах $3/2$ от оптимального.

Пусть G является полным графом на множестве вершин V, а функция w назначает неотрицательные вещественные веса каждому ребру графа G. Согласно неравенству треугольника для любых трёх вершин u, v и x должно выполняться $w(uv) + w(vx) \geqslant w(ux)$

Алгоритм можно описать на псевдокоде следующим образом.

1)Создаём минимальное остовное дерево T графа G.

2)Пусть O будет набором вершин с нечётными степенями в T. Согласно лемме о рукопожатиях, O имеет чётное число вершин.

3)Находим совершенное паросочетание M минимального веса в порождённом подграфе, заданным вершинами из O.

4)Комбинируем рёбра M и T с образованием связного мультиграфа H, в котором каждая вершина имеет чётную степень.

5)Образуем эйлеров цикл в H.

6)Преобразуем цикл, найденный на предыдущем шаге, в гамильтонов цикл путём пропуска повторяющихся вершин (сокращение).

Стоимость решения, полученного алгоритмом, лежит в границах $3/2$ от оптимального. Для доказательства этого факта предположим, что C является оптимальным обходом задачи коммивояжёра. Удаление ребра из C даёт стягивающее дерево, которое должно иметь вес, не меньший веса минимального стягивающего дерева, откуда следует, что $w(T) \leq w(C)$

Далее нумеруем вершины O в циклическом порядке по C и делим C на два множества путей — одно имеет нечётные номера первых вершины в циклическом порядке, а второе имеет чётные номера. Каждый набор путей соответствует совершенному паросочетанию множества O, которое сочетает в пару две конечные точки каждого пути, а вес этого сочетания не превосходит веса путей. Поскольку эти два множества путей разбивают рёбра C, одно из этих двух множеств имеет максимум половину веса C, и благодаря неравенству треугольника их соответствующее паросочетание имеет вес, который также не менее половины веса C. Совершенное паросочетание минимального веса не может иметь больший вес, так что $w(M) \leqslant w(C)/2$. Сложение весов T и M даёт вес эйлерова цикла, который не превосходит $3w(C)/2$. Благодаря неравенству треугольника сокращение не увеличивает вес, так что вес результата также не превосходит $3w(C)/2$.