ТРУДЫ ИНСТИТУТА МАТЕМАТИКИ И МЕХАНИКИ УрО РАН

Том 21 № 3 2015

УДК 519.16 + 519.85

PTAS ДЛЯ ЗАДАЧИ Min-k-SCCP В ЕВКЛИДОВОМ ПРОСТРАНСТВЕ ПРОИЗВОЛЬНОЙ ФИКСИРОВАННОЙ РАЗМЕРНОСТИ[1]

Е. Д. Незнахина

Изучается задача Min-k-SCCP о разбиении полного взвешенного графа на k вершинно-непересекающихся циклов минимального суммарного веса. Данная задача является обобщением известной задачи коммивояжера (TSP) и частным случаем классической задачи маршрутизации транспорта (VRP). Известно, что задача Min-k-SCCP в общем случае NP-трудна в сильном смысле и сохраняет свойство труднорешаемости даже в геометрической постановке. Для евклидовой задачи Min-k-SCCP в $\mathbb{R}^d$ при $k = O(\log n)$ построена полиномиальная приближенная схема, обобщающая подход, предложенный для решения задачи Min-2-SCCP на плоскости. Для произвольного $c > 1$ она находит $(1 + 1/c)$-приближенное решение задачи за время $O(n^{O(d)}(\log n)^{(O(\sqrt{d}c))^{d-1}})$.

Ключевые слова: цикловое покрытие размера k, задача коммивояжера (TSP), NP-трудная задача, полиномиальная приближенная схема (PTAS).

E. D. Neznakhina. A PTAS for the Min-k-SCCP in a Euclidean space of arbitrary fixed dimension.

We study the Min-k-SCCP on a partition of a complete weighted digraph into k vertex-disjoint cycles of minimum total weight. This problem is a generalization of the known traveling salesman problem (TSP) and a special case of the classical vehicle routing problem (VRP). It is known that the problem Min-k-SCCP is strongly NP-hard and preserves its intractability even in the geometric statement. For the Euclidean Min-k-SCCP in $\mathbb{R}^d$ with $k = O(\log n)$, we construct a polynomial-time approximation scheme, which generalizes the approach proposed earlier for the planar Min-2-SCCP. For any fixed $c > 1$ the scheme finds a $(1 + 1/c)$-approximate solution in $O(n^{O(d)}(\log n)^{(O(\sqrt{d}c))^{d-1}})$ time.

Keywords: cycle covering of size k, traveling salesman problem (TSP), NP-hard problem, polynomial-time approximation scheme (PTAS).

Введение

В классической постановке задачи коммивояжера (Traveling Salesman Problem, TSP) в качестве входных данных задан полный реберно взвешенный граф, требуется найти минимальный по весу гамильтонов цикл. Задача коммивояжера занимает особое место в комбинаторной оптимизации и исследовании операций и принадлежит к числу типовых NP-трудных задач [1].

В 70-е годы был получен ряд результатов, связанных с исследованием вычислительной сложности задач, возникающих в таких областях, как теория графов, математическое программирование, комбинаторная оптимизация, математическая логика и т. д. Р. Карп обосновал NP-полноту серии задач, в том числе задачи распознавания в графе гамильтонова цикла, путем сведения к ним проблемы выполнимости булевых форм [2]. Неаппроксимируемость задачи коммивояжера в классической постановке была доказана в работе [3].

Исследование частных случаев задачи коммивояжера представляется важным с точки зрения практических приложений. Отдельный интерес представляют два подкласса задачи TSP: Metric TSP и Euclidean TSP. Эти подклассы отличаются входными данными: в первом случае задан неориентированный граф, веса ребер которого удовлетворяют неравенству треугольника; во втором случае предполагается, что вершины графа являются точками в $\mathbb{R}^d$, а веса ребер определяются попарными расстояниями между ними.

[1]Исследования поддержаны Российским научным фондом, грант № 14-11-00109.

Задача коммивояжера является NP-трудной даже в евклидовой постановке [4], следовательно, точные решения задач Metric TSP и Euclidean TSP не могут быть найдены за полиномиальное время в предположении $P \neq NP$. Для указанных подклассов разработаны эффективные приближенные алгоритмы. Для задачи Metric TSP известен 2-приближенный алгоритм [5] и 3/2-приближенный алгоритм Н. Кристофидеса [6]. Тем не менее Metric TSP не обладает полиномиальной аппроксимационной схемой, если $P \neq NP$ [7], а оценка порога эффективной аппроксимируемости данной задачи остается одним из важных открытых вопросов в исследовании задачи коммивояжера. Для Euclidean TSP на плоскости Дж. Митчелл разработал полиномиальную приближенную схему [8]. Задача коммивояжера в евклидовой постановке в пространстве произвольной фиксированной размерности обладает асимптотически точным алгоритмом [9] и полиномиальной приближенной схемой, предложенной С. Аророй [10]. Заметим, что алгоритмы Дж. Митчелла и С. Ароры были разработаны практически одновременно и независимо, однако подход Митчелла не удалось обобщить на случай $\mathbb{R}^d$.

Предметом исследования настоящей работы является задача о минимальном по весу цикловом покрытии графа (Minimum-weight k-Size Cycle Cover Problem, Min-k-SCCP) [11; 12], являющаяся естественным обобщением задачи коммивояжера.

В задаче Min-k-SCCP задан полный реберно взвешенный орграф $G = (V, E, w)$ на n вершинах, весовая функция $w\colon E \to \mathbb{R}$ определяется матрицей $W = (w_{ij})$ $(1 \leq i, j \leq n)$; требуется найти k вершинно-непересекающихся контуров $\{C_1, \ldots, C_k\}$ наименьшего суммарного веса, посещающих в совокупности все вершины графа. Задача Min-k-SCCP NP-трудна в сильном смысле, метрическая и евклидова постановки данной задачи имеют аналогичный статус вычислительной сложности [11]. В работе [13] для Metric Min-k-SCCP предложен эффективный 2-приближенный алгоритм, оценка приближения является асимптотически достижимой. Для Euclidean Min-k-SCCP на плоскости при фиксированном значении параметра $k = 2$ построена полиномиальная приближенная схема [11], развивающая подход С. Ароры [10].

В данной статье рассматривается евклидова постановка задачи Min-k-SCCP и предполагается, что граф G является неориентированным, а веса ребер совпадают с расстояниями между соответствующими вершинами, являющимися точками в $\mathbb{R}^d$. В работе приведено обобщение полиномиальной приближенной схемы для Euclidean Min-2-SCCP на плоскости на случай произвольных фиксированных значений параметра k и размерности пространства.

1. Описание алгоритма полиномиальной приближенной схемы для задачи Euclidean Min-k-SCCP в $\mathbb{R}^d$

О п р е д е л е н и е 1. *Полиномиальной приближенной схемой* (Polynomial-Time Approximation Scheme, PTAS) для задачи комбинаторной оптимизации называется семейство алгоритмов, содержащее для каждого фиксированного $c > 1$ приближенный алгоритм, решающий данную задачу с гарантированной точностью $(1 + 1/c)$ за время, ограниченное сверху некоторым полиномом от длины записи ее исходных данных (при этом порядок и коэффициенты полинома могут, вообще говоря, зависеть от c).

Общая с х е м а а л г о р и т м а развивает подход, предложенный в [11] для решения задачи Min-2-SCCP на плоскости, и состоит из пяти основных частей.

1. Декомпозиция задачи на m $(m \leq k)$ независимых подзадач о цикловом покрытии графа и получение оценки сверху для длин сторон гиперкубов, объемлющих вершины графов, определяющих данные подзадачи, с помощью функции, которая выражает линейную зависимость от веса OPT оптимального циклового покрытия размера k.

2. Доказательство утверждения о том, что произвольной постановке задачи Euclidean Min-k-SCCP и каждому значению параметра c за полиномиальное время может быть сопоставлена округленная постановка так, что произвольный $(1 + 1/c)$-приближенный алгоритм для округленной постановки задачи индуцирует $(1+c_1/c)$-приближенный алгоритм для исходной задачи (для некоторого независимого значения $c_1 > 1$).

3. Построение рекурсивного разбиения объемлющего гиперкуба, содержащего вершины графа, задающего округленную постановку задачи Euclidean Min-k-SCCP.

4. Доказательство теоремы, утверждающей, что с вероятностью, не меньшей $1/2$, в рамках выбранной вероятностной модели существует $(1 + 1/c)$-оптимальный набор маршрутов специального вида, именуемый цикловым (m, r, k)-покрытием графа G. Понятие циклового (m, r, k)-покрытия графа введено в [11].

5. Построение $(1 + 1/c)$-оптимального циклового (m, r, k)-покрытия графа с помощью метода динамического программирования и стандартной схемы дерандомизации.

Пункты 1, 4 и 5 являются ключевыми, утверждение п. 2 — это простое следствие леммы из [11], метод реализации п. 3 совпадает с соответствующим подходом, предложенным для решения задачи Euclidean TSP в $\mathbb{R}^d$ [10].

2. PTAS для задачи Euclidean Min-k-SCCP в $\mathbb{R}^d$

2.1. Декомпозиция евклидовой задачи о цикловом покрытии графа

Как известно, диаметр D множества и радиус R описанного вокруг него шара в евклидовом пространстве размерности d связаны между собой неравенством Юнга [14]

$$\frac{1}{2}D \leqslant R \leqslant \left(\frac{d}{2d+2}\right)^{1/2} D. \tag{1}$$

Построим минимальный остовный лес из k деревьев (k-MSF) $T_1, T_2, \ldots, T_k$, воспользовавшись простой модификацией алгоритма Борувки — Краскала [5], предложенной в [11].

Введем ряд обозначений:

D_i — диаметр множества вершин, входящих в дерево T_i, $i \in \mathbb{N}_k = \{1, \ldots, k\}$;

D — максимальное значение D_i ($i \in \mathbb{N}_k$);

R_i — радиус описанной вокруг T_i сферы;

R — максимальное значение R_i ($i \in \mathbb{N}_k$).

Рассмотрим вспомогательный полный граф $G_T = (V_T, E_T)$, в котором деревья k-MSF образуют множество вершин V_T, а веса ребер определяются функцией $\rho\colon V_T \times V_T \to \mathbb{R}$. Значение $\rho(T_i, T_j)$ равно расстоянию между центрами сфер, описанных вокруг деревьев T_i и T_j, где $\{i, j\} \subset \mathbb{N}_k$.

Построим разбиение множества вершин G_T на m ($m \leq k$) кластеров методом ближайшего соседа [15;16] с пороговым значением расстояния между кластерами $(2k+1)R$, вычислительная сложность процедуры составит $O(k^3)$. Для каждого построенного кластера объединим множества вершин всех деревьев, вошедших в данный кластер, и обозначим эти множества S_i; таким образом получим кластеризацию $S_1, S_2, \ldots, S_m$ вершин исходного графа G. Под диаметром кластера будем понимать максимальное расстояние между входящими в него вершинами.

Утверждение 1. *Любой маршрут, входящий в оптимальное цикловое покрытие размера k для графа G, проходит через вершины, соответствующие только одному из кластеров $S_1, S_2, \ldots, S_m$. Кроме того, диаметры кластеров ограничены сверху и выполняется неравенство*

$$\max_{i \in \mathbb{N}_m} D_{S_i} \leq \left(\frac{d}{2d+2}\right)^{1/2} (2k^2 - k + 1) OPT. \tag{2}$$

Д о к а з а т е л ь с т в о. Предположим от противного, что один из k маршрутов циклового покрытия наименьшего веса содержит вершины как из S_i, так и из S_j, где $\{i, j\} \subset \mathbb{N}_m$.

Обозначим этот маршрут через P. По предположению P содержит хотя бы два ребра, соединяющих вершины из разных кластеров, суммарная длина которых по условию больше $2(2k-1)R$.

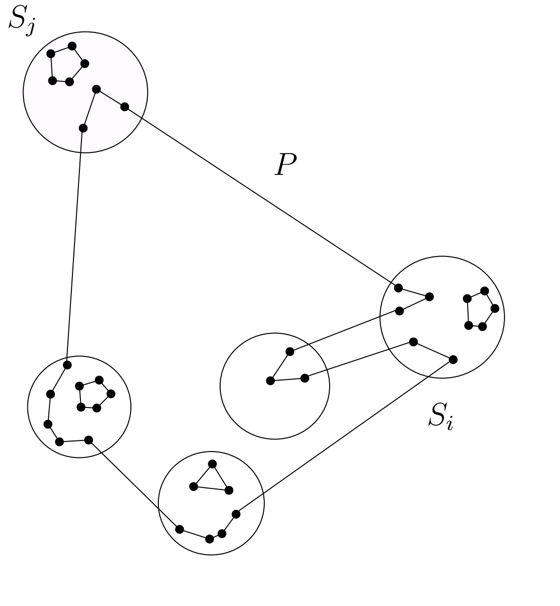

Рис. 1. В P чередуются вершины из разных сфер, $k = 5$.

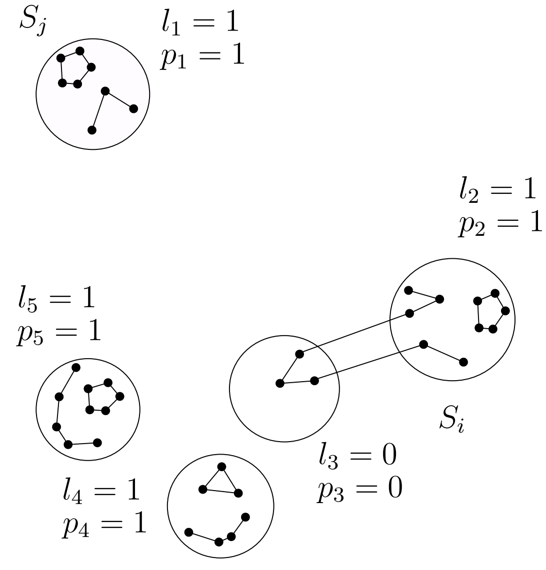

Рис. 2. Разбиение P на фрагменты.

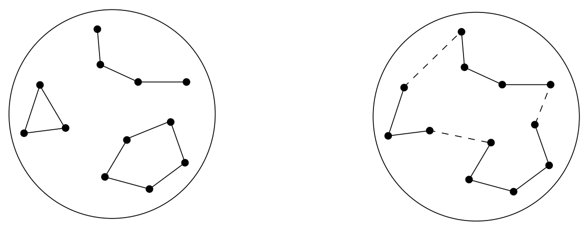

Рис. 3. Замыкание фрагмента P и 2-х циклов в один маршрут ($l_i = 2,\ p_i = 1$).

В маршруте P могут чередоваться вершины из разных сфер (рис. 1). Пусть $\{u, v\}$ — произвольное ребро, входящее в P, такое, что $u \in S_i$, $v \in S_j$. Зафиксировав порядок обхода вершин маршрута P: $u \to u_{i_1} \to \ldots \to v$, построим разбиение P на фрагменты таким образом, чтобы каждой сфере соответствовал фрагмент P между первой и последней посещенными внутри данной сферы вершинами, и удалим ребра маршрута P, соединяющие вершины из различных фрагментов (рис. 2).

Поставим в соответствие сфере, описанной вокруг дерева T_i, числа l_i $(0 \leq l_i \leq k-1)$ и p_i $(0 \leq p_i \leq 1)$, равные числу циклов и количеству фрагментов маршрута P, отнесенных данной сфере (рис. 2). Если цикл соединяет вершины из нескольких деревьев и пересекает соответствующие описанные сферы, то сопоставим его одной из этих сфер, выбранной произвольным образом. По построению верны следующие соотношения

$$\sum_{i=1}^{k} l_i = k - 1 \quad \text{и} \quad 2 \leq \sum_{i=1}^{k} p_i \leq k. \tag{3}$$

Пусть q — это количество сфер, для которых $l_i + p_i = 0$. Рассмотрим два случая: $q = 0$ и $q \geq 1$. В первом случае построим цикловое покрытие размера k следующим образом. Для i-й сферы объединим в один цикл l_i соответствующих замкнутых маршрутов и фрагмент P с помощью добавления не более, чем $(l_i + p_i)$ новых ребер внутри данной сферы, вес каждого из которых не превосходит $2R$ (рис. 3). Выполним указанное преобразование для каждой из k сфер, при этом стоимость циклового покрытия увеличится не более, чем на

$$2R\sum_{i=1}^{k}(l_i+p_i)\le 2(2k-1)R$$

в силу соотношений (3), в то время как суммарный вес удаленных ребер превышает эту величину. Таким образом, в предположении $q=0$ обосновано существование циклового покрытия размера k для графа G, вес которого строго меньше веса первоначального решения.

Во втором случае $q\ge 1$, и следовательно $(k-q)$ сферам, для каждой из которых $l_i+p_i>0$, соответствуют $(k-1)$ цикл и не менее двух фрагментов P. Заметим, что одной сфере не может быть отнесено более одного фрагмента P. Тогда не менее $(q+1)$ цикла распределено по сферам, в которых уже содержится цикл либо фрагмент P. Исключим q циклов из рассмотрения таким образом, чтобы количество сфер, для которых $l_i+p_i=0$, не изменилось. Получим, что $(k-q)$ сферам приписаны $(k-q-1)$ цикл и фрагменты P. Справедливы следующие соотношения:

$$l_i+p_i>0,\quad i=1,\dots,k-q,$$
$$\sum_{i=1}^{k-q}l_i=k-q-1\quad\text{и}\quad 2\le\sum_{i=1}^{k-q}p_i\le k-q. \tag{4}$$

Для указанных $(k-q)$ сфер выполним преобразование, аналогичное проведенному в первом случае. Для каждой сферы объединим в один цикл сопоставленные ей замкнутые маршруты и фрагмент P с помощью добавления не более, чем (l_i+p_i) новых ребер внутри данной сферы (рис. 3). Таким образом, мы построим цикловое покрытие размера $(k-q)$ с помощью модификации, увеличивающей вес циклового покрытия не более, чем на

$$2R\sum_{i=1}^{k-q}(l_i+p_i)\le 2(2k-2q-1)R,$$

в силу соотношений (4). Добавим к полученному цикловому покрытию q циклов, ранее исключенных из преобразования, и получим цикловое покрытие размера k, вес которого также строго меньше веса исходного покрытия.

Тем самым обосновано существование k-циклового покрытия, вес которого строго меньше веса исходного решения, что противоречит предположению об оптимальности последнего.

Оценим диаметр полученных кластеров. В пространстве $\mathbb{R}^d$ диаметр кластера будет наибольшим, если центры сфер, описанных вокруг деревьев k-MSF и образующих данный кластер, лежат на одной прямой. Любой кластер из разбиения множества вершин V исходного графа G содержит не более k сфер, поэтому выполняется неравенство

$$\max_{i\in\mathbb{N}_m}D_{S_i}\le(k-1)(2k+1)R+2R.$$

Применяя неравенство (1) и учитывая очевидную двустороннюю оценку $D\leqslant MSF\leqslant OPT$ для веса минимального остовного леса $\{T_1,T_2,\dots,T_k\}$, получаем

$$R\leqslant\Big(\frac{d}{2d+2}\Big)^{1/2}D\leqslant\Big(\frac{d}{2d+2}\Big)^{1/2}OPT.$$

Отсюда верна оценка

$$\max_{i\in\mathbb{N}_m}D_{S_i}\le\Big(\frac{d}{2d+2}\Big)^{1/2}(2k^2-k+1)OPT.$$

Утверждение доказано.

Таким образом, если $m=k$, то задача Euclidean Min-k-SCCP разбивается на k независимых подзадач TSP. В частности, PTAS для задачи Euclidean Min-k-SCCP может быть построена комбинацией PTAS для этих подзадач.

Однако, если количество построенных кластеров строго меньше k, то необходимо перебрать все возможные случаи распределения k циклов по m кластерам, для каждого случая решить m независимых подзадач о цикловом покрытии и выбрать оптимальное решение. Количество таких случаев совпадает с числом композиций k длины m и равно биномиальному коэффициенту из $(k-1)$ по $(m-1)$ [17].

Остановимся на рассмотрении частного случая задачи Euclidean Min-k-SCCP, для которого $m = 1$, так как он является наихудшим с точки зрения вычислительной сложности.

2.2. Округленная постановка задачи

Для полноты изложения приведем определение округленной задачи Euclidean Min-k-SCCP, принятое в [10; 11].

Определение 2. *Округленной* называется постановка задачи Euclidean Min-k-SCCP, для которой выполняются следующие ограничения: все вершины графа имеют целочисленные координаты и вес любого ребра e_{ij} больше или равен 4.

Чтобы получить округленную постановку задачи, необходимо и достаточно провести следующие преобразования.

1. Положим

$$L = \left(\frac{d}{2d+2}\right)^{1/2}(2k^2 - k + 1)MSF \leq \left(\frac{d}{2d+2}\right)^{1/2}(2k^2 - k + 1)OPT$$

и построим объемлющий вершины графа G гиперкуб $\mathcal{S}$ со стороной L. Исходная постановка задачи Euclidean Min-k-SCCP удовлетворяет неравенству (2), следовательно, такое построение возможно.

2. Построим внутри $\mathcal{S}$ ортогональную решетку с расстоянием между гиперплоскостями $L/8cn\sqrt{d}$ и переместим вершины заданного графа в ближайшие к ним узлы решетки. Расстояние между любыми двумя вершинами увеличится при этом не более чем на $L/4cn$, соответственно вес циклового покрытия размера k изменится не более чем на $L/4c$.

3. Изменим расстояние между гиперплоскостями решетки, умножив все координаты на $32cn\sqrt{d}/L$. Тогда минимальное расстояние между вершинами станет равно 4, а длина стороны объемлющего гиперкуба — $O(cn\sqrt{d})$.

4. Поместим начало координат в одном из углов $\mathcal{S}$ и расположим координатные оси таким образом, чтобы ребра гиперкуба принадлежали осям и все вершины графа имели неотрицательные координаты.

Достаточность построения PTAS для округленной Euclidean Min-k-SCCP непосредственно следует из леммы, доказанной в [11], одна из эквивалентных формулировок которой приведена ниже.

Лемма. *Пусть округленная задача* Euclidean Min-k-SCCP *получена из исходной постановки задачи* Euclidean Min-k-SCCP *с помощью преобразований, увеличивающих вес произвольного циклового покрытия размера k не более чем на $O(OPT/c)$. Тогда* PTAS *для округленной* Euclidean Min-k-SCCP *индуцирует* PTAS *для общей задачи* Euclidean Min-k-SCCP.

Далее опишем построение PTAS для округленной задачи Euclidean Min-k-SCCP.

2.3. Рекурсивное разбиение объемлющего гиперкуба $\mathcal{S}$

Построим геометрическое разбиение задачи с помощью структуры данных 2^d-дерево, аналогичной 4-дереву, использованному в случае плоскости [10; 11].

Длину стороны L объемлющего гиперкуба $\mathcal{S}$ примем равной наименьшей подходящей степени двойки. Организуем процедуру построения 2^d-дерева следующим образом: $\mathcal{S}$ — корень дерева, каждый гиперкуб, включая корневой, делим на 2^d равных дочерних гиперкуба. Повторяем данную процедуру рекурсивно до тех пор, пока не получим гиперкубы, содержащие не более одной вершины исходной задачи. Договоримся считать, что $\mathcal{S}$ принадлежит уровню 0, 2^d его дочерних гиперкуба — уровню 1 и так далее. Заметим, что построенное дерево содержит $O(2^d n)$ листьев, $O(\log L) = O(\log(cn\sqrt{d}))$ уровней, и соответственно количество узлов 2^d-дерева можно оценить сверху как $O(2^d n \log(cn\sqrt{d}))$.

Зададимся значением параметра $m \in \mathbb{N}$ и поставим в соответствие каждой $(d-1)$-мерной грани узла дерева ортогональную сетку из $m+2^{d-1}$ порталов. Заметим, что в отличие от случая плоскости не каждое натуральное число является допустимым значением параметра m.

Утверждение 2. *Расстояние между ближайшими порталами на грани узла i-го уровня 2^d-дерева равно $O(L/2^i m^{1/(d-1)})$.*

Доказательство. Для построения ортогональной сетки на грани d-мерного куба необходимо $(m_1+2)^{d-1} - 2^{d-1}$ порталов, где $m_1 \in \mathbb{N}$ — это количество порталов на ребре гиперкуба (рис. 4).

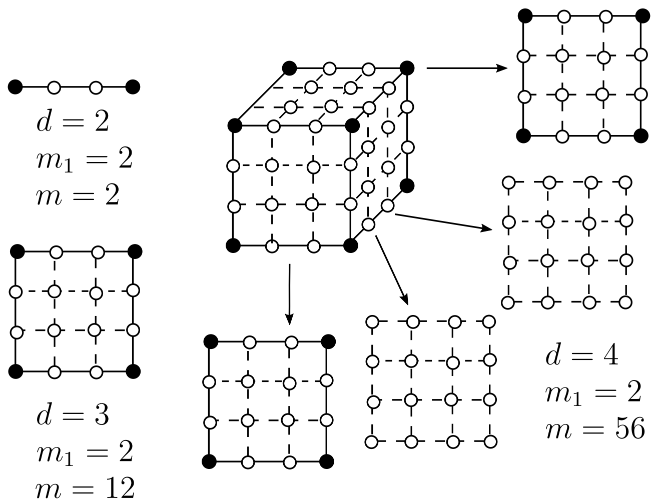

Рис. 4. Пример построения ортогональной сетки порталов при $d = 2, 3, 4$.

Длина стороны гиперкуба i-го уровня равна $L/2^i$ по построению, отсюда расстояние между соседними порталами на грани данного узла вычисляется как $L/2^i(m_1+1)$.

Покажем, что верно неравенство

$$\frac{L}{2^i(m_1+1)} \leq \frac{3}{2}\,\frac{L}{2^i m^{1/(d-1)}};$$

это эквивалентно

$$m_1 + 1 \geq \frac{2}{3}((m_1+2)^{d-1} - 2^{d-1})^{1/(d-1)}.$$

При стремлении d к бесконечности значение выражения $((m_1+2)^{d-1} - 2^{d-1})^{1/(d-1)}$ асимптотически стремится снизу к величине $m_1 + 2$, поэтому

$$\frac{m_1+1}{((m_1+2)^{d-1} - 2^{d-1})^{1/(d-1)}} \geq \frac{m_1+1}{m_1+2} \geq \frac{2}{3}.$$

Таким образом, если гиперкуб в 2^d-дереве имеет уровень i, то расстояние между ближайшими порталами можно оценить как $O(L/2^i m^{1/(d-1)})$.

Утверждение доказано.

Понятия центральной точки и дерева с циклическим сдвигом, определенные для решения задачи Euclidean Min-2-SCCP на плоскости, легко переносятся на случай произвольной фиксированной размерности пространства.

Назовем *центральной* d-мерную точку, все координаты которой равны $L/2$.

Определение 3. Пусть $a_1, a_2, \ldots, a_d \in \mathbb{N}_L$; *$2^d$-деревом* $T(a_1, a_2, \ldots, a_d)$ назовем 2^d-дерево, для которого точка $((L/2 + a_1) \mathrm{mod}\ L, (L/2 + a_2) \mathrm{mod}\ L, \ldots, (L/2 + a_d) \mathrm{mod}\ L)$ является центральной.

Гиперкубы, принадлежащие произвольному уровню $i \geqslant 1$ дерева $T(a_1, a_2, \ldots, a_d)$, как и его центральная точка, подвергаются циклическому сдвигу по каждой из d координатных осей, при этом положение $\mathcal{S}$ и координаты вершин графа G остаются неизменными.

По аналогии со случаем $d = 2$ будем называть *m-регулярным множеством порталов* и обозначать $P(a_1, a_2, \ldots, a_d; m)$ объединение разбиений порталами $(d-1)$-мерных граней всех узлов 2^d-дерева $T(a_1, a_2, \ldots, a_d)$ (за исключением $\mathcal{S}$).

2.4. Теорема существования

Обоснуем в этом разделе существование циклового покрытия размера k для заданного графа G, обладающего рядом свойств.

Обозначим через $V(C)$ множество вершин произвольного цикла C в графе G ($V(C) \subseteq V$). Сопоставим C (m,r)-аппроксимацию, т. е. замкнутую ломаную $l(C)$, удовлетворяющую следующим условиям:

(i) множество вершин $l(C)$ является подмножеством $V(C) \cup P(a_1, a_2, \ldots, a_d; m)$;

(ii) ломаная $l(C)$ обходит вершины $V(C)$ в порядке, задаваемом маршрутом C;

(iii) $l(C)$ пересекает каждую грань произвольного узла дерева $T(a_1, a_2, \ldots, a_d)$ не более r раз ($r \in \mathbb{N}$), причем исключительно в точках множества $P(a_1, a_2, \ldots, a_d; m)$.

Приведем одну из эквивалентных формулировок теоремы 5 из [10].

Теорема 1 (Structure Theorem for Euclidean TSP in $\mathbb{R}^d$). *Пусть постановка округленной задачи* TSP *в* $\mathbb{R}^d$ *задается полным евклидовым графом G, длина стороны объемлющего гиперкуба $\mathcal{S}$ которого равна L, и задана константа $c > 0$. Пусть дискретные случайные величины $a_1, a_2, \ldots, a_d$ распределены равномерно и независимо на множестве $\mathbb{N}_L$.*

Тогда для любого $\eta \in (0,1)$ найдутся D_1, $D_2 > 0$ такие, что при $r = \left\lceil (D_1\sqrt{d}c)^{d-1} \right\rceil$, $m = \left\lceil (D_2 dc \log L)^{d-1} \right\rceil$ для произвольного простого цикла C в графе G веса $W(C)$ с вероятностью, не меньшей $1 - \eta$, существует (m,r)-аппроксимация $l(C)$, вес которой не превосходит $(1 + 1/c)W(C)$.

Определение 4 [11]. Пусть $\mathcal{C} = \{C_1, \ldots, C_k\}$ — произвольное цикловое покрытие размера k графа G, $l(C_i)$ — (m,r)-аппроксимация цикла C_i. Множество $\mathcal{L}(\mathcal{C}) = \{l(C_1), \ldots, l(C_k)\}$ назовем *цикловым (m,r,k)-покрытием* графа G.

Обобщим результат теоремы 3 из [11] на случай задачи Euclidean Min-k-SCCP в пространстве произвольной фиксированной размерности и докажем следующую теорему.

Теорема 2. *Пусть $c > 0$ — произвольная постоянная, L — размер объемлющего гиперкуба $\mathcal{S}$ для округленной постановки задачи* Euclidean Min-k-SCCP *в* $\mathbb{R}^d$. *Дискретные случайные величины $a_1, a_2, \ldots, a_d$ распределены равномерно и независимо на множестве $\mathbb{N}_L$. Тогда при $m = (O(dc \log L))^{d-1}$ и $r = (O(\sqrt{d}c))^{d-1}$ с вероятностью, не меньшей $1/2$, существует цикловое (m,r,k)-покрытие стоимости, не превышающей $(1 + 1/c)\,OPT$.*

Доказательство. Рассмотрим $\mathcal{C}^* = \{C_1^*, C_2^*, \ldots, C_k^*\}$ — цикловое покрытие размера k наименьшего веса как решение рассматриваемой округленной постановки задачи Euclidean Min-k-SCCP. Как обычно, обозначим вес $\mathcal{C}^*$ через OPT, т. е. $\sum_{i=1}^k W(C_i^*) = OPT$.

Воспользуемся теоремой 1: при $\eta = 1/2k$ для каждого цикла C_i^* с вероятностью, не меньшей $1 - 1/2k$, существует (m,r)-аппроксимация $l(C_i^*)$ веса

$$W(l(C_i^*)) \leqslant (1 + 1/c)W(C_i^*). \tag{5}$$

Значения параметров m и r удовлетворяют оценкам $(O(dc\log L))^{d-1}$ и $(O(\sqrt{d}c))^{d-1}$ соответственно.

В силу равномерности распределения случайных величин $a_1, a_2, \dots, a_d$ вероятность объединения обратных событий, состоящих в отсутствии для цикла C_i^* (m,r)-аппроксимации $l(C_i^*)$, вес которой удовлетворяет (5), ограничена сверху 1/2. Таким образом, с вероятностью, не меньшей 1/2, существует такое цикловое (m,r,k)-покрытие $\{l(C_1^*), l(C_2^*), \dots, l(C_k^*)\}$, что

$$\sum_{i=1}^{k} W(l(C_i^*)) \leqslant \Big(1+\frac{1}{c}\Big)\sum_{i=1}^{k} W(C_i^*) = \Big(1+\frac{1}{c}\Big)\,OPT.$$

Теорема доказана.

2.5. Замечания к процедуре динамического программирования

Процедура поиска циклового (m,r,k)-покрытия $\{l_1, l_2, \dots, l_k\}$ наименьшего веса в задаче Euclidean Min-k-SCCP основана на методе динамического программирования и развивает подход, предложенный в [10; 11]. Время работы алгоритма составит $O(n(\log n)^{(O(\sqrt{d}c))^{d-1}})$.

Внутренняя подзадача для узла 2^d-дерева S состоит в поиске части циклового (m,r,k)-покрытия, целиком находящейся внутри S и посещающей все расположенные в нем вершины исходного графа, минимальной стоимости.

В начале процедуры динамического программирования рассматриваются листья 2^d-дерева $T(a_1, a_2, \dots, a_d)$. Пусть S — произвольный лист дерева. По построению S содержит не более одной вершины исходного графа G, и соответствующая подзадача может быть решена прямым перебором за $O(2dr)$ операций.

Рассмотрим случай, когда S не является листом дерева $T(a_1, a_2, \dots, a_d)$. Обозначим через $S^1, \dots, S^{2^d}$ его дочерние гиперкубы, для которых предполагаем решенными внутренние подзадачи. Построим решение для S реккурентно, считая выполненным условие, что ответ состоит из отрезков (m,r)-аппроксимаций $\{l_1, l_2, \dots, l_k\}$.

Для оценки трудоемкости процедуры динамического программирования сохраним обозначение, введенное в [11]: $\mathfrak{P}$ — семейство всевозможных мультимножеств P, состоящих не более чем из $2dr$ порталов, расположенных на внутренних относительно S гранях дочерних гиперкубов $S^1, \dots, S^{2^d}$. По построению на каждой такой грани расположено $m + 2^{d-1}$ портала, и в силу ограничений, накладываемых на цикловое покрытие, грань может быть пересечена не более r раз. Получим, что $|\mathfrak{P}| = O((m+2^{d-1})^{2dr})$.

Для произвольного мультимножества $P \in \mathfrak{P}$ существует $O((dr)^{2dr})$ способов назначить каждому порталу ломаную, соответствующую одной из k (m,r)-аппроксимаций, и $O((2dr)!)$ вариантов разбиения мультимножества на упорядоченные пары.

Трудоемкость подзадачи, соответствующей S, может быть оценена сверху величиной

$$O\big((m+2^{d-1})^{2dr}(dr)^{2dr}(2dr)!\big).$$

Задача поиска циклового (m,r,k)-покрытия минимального веса эквивалентна решению подзадачи для объемлющего гиперкуба $\mathcal{S}$.

Оценим сверху число подзадач для вывода общей трудоемкости процедуры динамического программирования. Заметим, что каждому узлу 2^d-дерева S соответствует $O((m+2^{d-1})^{2dr})$ способов выбора мультимножества P порталов на $(d-1)$-мерных гранях S.

Каждому такому мультимножеству соответствует $O((2dr)!)$ способов разбиения его на упорядоченные пары, и каждому такому разбиению — $O(k^{dr})$ способов распределения этих пар между маршрутами $\{l_1, l_2, \dots, l_k\}$. Принимая во внимание, что общее число узлов дерева $T(a_1, a_2, \dots, a_d)$ составляет $O(2^d n\log(cn\sqrt{d}))$, получаем искомую оценку трудоемкости процедуры для произвольной фиксированной центральной точки $(a_1, a_2, \dots, a_d)$:

$$O\big(2^d n\log(cn\sqrt{d}) \times (m+2^{d-1})^{4dr}((2dr)!)^2(dr)^{2dr} \times k^{dr}\big). \tag{6}$$

Переменные d и c не являются частью входа задачи Euclidean Min-k-SCCP. При значениях параметров $m = (O(dc\log(cn\sqrt{d})))^{d-1}$ и $r = (O(\sqrt{d}c))^{d-1}$, а также, с учетом того, что количество способов распределения k циклов по m кластерам равно $O(2^k)$, оценка (6) эквивалентна

$$O\big(n(k\log n)^{(O(\sqrt{d}c))^{d-1}}2^k\big).$$

З а м е ч а н и е. Вычислительная сложность стандартной схемы дерандомизации, связанной с полным перебором сдвигов 2^d-дерева $T(a_1, a_2, \dots, a_d)$, составляет $O(n^d)$.

Таким образом, доказана следующая теорема.

Теорема 3. *Задача* Euclidean Min-k-SCCP *в* $\mathbb{R}^d$ *обладает полиномиальной приближенной схемой трудоемкости*

$$O\big(n^{d+1}(k\log n)^{(O(\sqrt{d}c))^{d-1}}2^k\big). \tag{7}$$

При $d = 2$ вычислительная сложность PTAS, построенной для задачи Euclidean Min-k-SCCP, совпадает с трудоемкостью PTAS, построенной в [11] для задачи Euclidean Min-2-SCCP, и отличается от нее, как следует из соотношения (7), постоянным множителем $2^k k^{O(c)}$.

Следствие. *Задача* Eclidean Min-k-SCCP *в* $\mathbb{R}^d$ *с условием, что* k *является частью входа, обладает полиномиальной приближенной схемой с трудоемкостью* $O(n^{O(d)}(\log n)^{(O(\sqrt{d}c))^{d-1}})$ *для* $k = O(\log n)$.

Заключение

В работе удалось распространить результат, полученный для задачи Euclidean Min-2-SCCP на плоскости, обосновав полиномиальную приближенную схему для произвольных фиксированных значений параметра k и размерности пространства d. Предложенный алгоритм также является PTAS в случае, если параметр k является частью входа задачи Euclidean Min-k-SCCP в $\mathbb{R}^d$, при $k = O(\log n)$.

СПИСОК ЛИТЕРАТУРЫ

1. **Гэри М., Джонсон Д.** Вычислительные машины и труднорешаемые задачи. М.: Мир, 1982. 416 с.
2. **Karp R.** Reducibility among combinatorial problems // Complexity of Computer Computations: Proc. Sympos. / eds. R. E. Miller and J. W. Thatcher. New York: Plenum Press, 1972. P. 85–103. (The IBM Research Symposia Series).
3. **Sahni S., Gonzales T.** *P*-complete approximation problems // J. Assoc. Comput. Mach. 1976. Vol. 23, no. 3. P. 555–565.
4. **Papadimitriou C.** Euclidean TSP is *NP*-complete // Theoret. Comput. Sci. 1977. Vol. 4, no. 3. P. 237–244.
5. Introduction to Algorithms / T. Cormen, C. Leiserson, R. Rivest, C. Stein. Cambridge; London: MIT Press, 1990. 1292 p.
6. **Christofides N.** Worst-case analysis of a new heuristic for the traveling salesman problem // Proc. of Symposium on New Directions and Recent Results in Algorithms and Complexity. New York: Academic Press, 1976. P. 441.
7. Proof verification and interactability of approximation problems / S. Arora, C. Lund, R. Motwani, M. Sudan, M. Szegedy // Proc. of the 33rd Annual IEEE Symposium on Foundations of Computer Science. 1992. P. 13–22.
8. **Mitchell J.** Guillotine subdivisions approximate polygonal subdivisions: A simple polynomial-time approximation scheme for geometric TSP, *k*-MST, and related problems // SIAM J. Comp. 1999. Vol. 28, no. 4. P. 1298–1309.
9. **Гимади Э.Х., Перепелица В.А.** Асимптотически точный подход к решению задачи коммивояжера // Управляемые системы: сб. науч. тр. / Ин-т математики СО АН СССР. Вып. 12. Новосибирск, 1974. С. 35–45.

10. **Arora S.** Polynomial-time approximation schemes for Euclidean traveling salesman and other geometric problems // J. ACM. 1998. Vol. 45, no. 5. P. 753–782.
11. **Хачай М.Ю., Незнахина Е.Д.** Полиномиальная приближенная схема для евклидовой задачи о цикловом покрытии графа // Тр. Ин-та математики и механики УрО РАН. 2014. Т. 20, № 4. С. 297–311.
12. **Хачай М.Ю., Незнахина Е.Д.** Аппроксимируемость задачи о минимальном по весу цикловом покрытии графа // Докл. АН. 2015. Т. 461, № 6. С. 644–649.
13. **Khachay M., Neznakhina K.** Approximation of Euclidean k-size cycle cover problem // Croat. Oper. Res. Rev. 2014. Vol. 5, no. 2. P. 177–188.
14. **Jung H.** Über die kleinste Kugel, die eine räumliche Figur einschliesst // J. Reine Angew. Math. 1901. Vol. 123. P. 241–257.
15. **Sneath P.** Computers in taxonomy // J. Gen. Microbiol. 1957. Vol. 17. P. 201–226.
16. **Gower J., Ross G.** Minimum spanning trees and single linkage cluster analysis // Appl. Statist. 1969. Vol. 18. P. 54–64.
17. **Эндрюс Г.** Теория разбиений. М.: Наука, 1982. 256 с.

Незнахина Екатерина Дмитриевна
аспирант, математик
Институт математики и механики им. Н. Н. Красовского УрО РАН,
Уральский федеральный университет им. Б. Н. Ельцина
e-mail: eneznakhina@yandex.ru

Поступила 13.05.2015