Math-Net.Ru

Общероссийский математический портал

М. Д. Миссаров, Р. Г. Степанов, О задачах комбинаторной оптимизации в ультраметричных пространствах, *ТМФ*, 2003, том 136, номер 1, 164–176

DOI: <https://doi.org/10.4213/tmf209>

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением

<http://www.mathnet.ru/rus/agreement>

Параметры загрузки:

IP: 54.197.130.99

26 апреля 2023 г., 11:51:12

О ЗАДАЧАХ КОМБИНАТОРНОЙ ОПТИМИЗАЦИИ В УЛЬТРАМЕТРИЧНЫХ ПРОСТРАНСТВАХ

Исследованы решения некоторых известных задач комбинаторной оптимизации в d -мерных p -адических пространствах, включая задачу о минимальном паросочетании, задачу о минимальном остовном дереве, а также задачу коммивояжера. Оказалось, что в ультраметричном пространстве “жадные” алгоритмы дают оптимальные решения этих задач, что позволяет получить явные выражения для оценки их средних значений. Исследована асимптотика этих значений при бесконечном увеличении числа точек, обнаружены некоторые сходные черты с евклидовым случаем, а также новые неожиданные свойства.

Ключевые слова: задача коммивояжера, минимальное паросочетание, ультраметричность, жадные алгоритмы, минимальное остовное дерево, p -адические пространства, свойство самоусреднения.

1. ВВЕДЕНИЕ

В литературе по статистической физике и теории вероятностей отмечают сходство задач комбинаторной оптимизации с некоторыми задачами теории неупорядоченных систем [1].

В качестве примера мы рассмотрим три хорошо известные задачи комбинаторной оптимизации – задачу коммивояжера (ЗК), задачу о минимальном паросочетании (МП) и задачу о минимальном остовном дереве (МОД). В ЗК требуется найти кратчайший замкнутый путь, проходящий через заданные точки некоторого метрического пространства. В задаче МП требуется разбить множество точек на пары и соединить точки каждой полученной пары ребром. При этом, если общее число точек нечетно, то одна точка останется не спаренной. Требуется, чтобы сумма длин полученных ребер была наименьшей. В задаче МОД требуется найти дерево, вершины которого совпадают с заданными точками, а сумма длин ребер является наименьшей.

Как было отмечено в работах Мезарда, Паризи, Киркпатрика и др., стохастические задачи комбинаторной оптимизации, в которых точки разбрасываются случайным образом по некоторой области евклидова пространства, можно в определенном смысле

* Казанский государственный университет, Казань, Россия.
E-mail: Moukadas.Missarov@ksu.ru; rstepanov@mail.ru

интерпретировать как модели теории неупорядоченных систем. Действительно, допустимые решения задачи комбинаторной оптимизации можно рассматривать как состояния физической системы, значение целевой функции для допустимого решения – как энергию состояния, а оптимальное решение – как основное состояние системы.

Пусть $f_{n,d}(\omega)$ обозначает оптимальное решение той или иной задачи комбинаторной оптимизации, где ω – конфигурация n точек, равномерно распределенных в единичном шаре d -мерного евклидова пространства (под конфигурацией точек здесь понимается набор их координат, т.е. элементарный исход в модели случайного распределения точек по шару). В работах [2] было показано, что для вышеперечисленных задач комбинаторной оптимизации справедливо свойство “самоусреднения”: для почти всех (в вероятностном смысле) ω существует предел

$$\lim_{n \rightarrow \infty} \frac{f_{n,d}(\omega)}{n^{1-1/d}} = C(d), \quad (1)$$

где константа зависит от вида задачи и размерности пространства, но не зависит от ω .

Мезарди Паризи рассмотрели задачу МП в предположении, что “расстояния” между точками являются независимыми случайными величинами с плотностью распределения $p(x)$ такой, что $\lim_{x \rightarrow 0} p(x) = 1$. Используя метод реплик, они нестрого показали, что в задаче МП среднее (по ω) $\overline{f_{n,1}} \xrightarrow{n \rightarrow \infty} \pi^2/12$.

В настоящей статье мы рассматриваем задачи комбинаторной оптимизации для точек, лежащих в ультраметричном пространстве. Естественными примерами таких пространств являются дискретная решетка с иерархическим расстоянием и ее континуальный аналог – p -адическое пространство. Как известно, свойство ультраметричности приводит к существенному упрощению некоторых моделей математической физики. В частности, преобразование ренормгруппы Вильсона в иерархических моделях локально, что позволяет дать строгое описание метода ренормгруппы в иерархической модели Дайсона [3] и даже приводит к точному решению фермионной иерархической модели [4]. p -Адические модели в теории струн и в квантовой механике рассматривались в работах [5]. Отметим также, что свойство ультраметричности p -адического расстояния позволяет получить явные формулы для p -адических фейнмановских амплитуд во всех порядках теории возмущений [6].

Оказывается, что сложные задачи комбинаторной оптимизации для точек, лежащих в ультраметричном пространстве, также упрощаются. Ультраметричность приводит к тому, что “жадные” алгоритмы дают оптимальные решения вышеперечисленных задач. Для стохастических версий этих задач, когда точки разбрасываются случайно и равномерно (относительно меры Хаара) по единичному шару d -мерного p -адического пространства, мы получаем рекуррентные соотношения для средних значений оптимальных решений, используя соображения, близкие к идеям ренормализационной группы. Эти соотношения имеют решения в виде рядов.

Если обозначить среднее значение оптимального решения одной из вышеназванных

задач через $\overline{f_{n,d,p}}$, то при $d > 1$ справедливы неравенства

$$c_1 < \liminf_{n \rightarrow \infty} \frac{\overline{f_{n,d,p}}}{n^{1-1/d}} < \limsup_{n \rightarrow \infty} \frac{\overline{f_{n,d,p}}}{n^{1-1/d}} < c_2,$$

где $0 < c_1 < c_2 < +\infty$.

В отличие от евклидова случая (1), предел

$$\lim_{n \rightarrow \infty} \frac{\overline{f_{n,d}}}{n^{1-1/d}}$$

не существует и, следовательно, явление самоусреднения в этих моделях отсутствует.

2. ОБОСНОВАНИЕ АЛГОРИТМОВ

Напомним, что ультраметричным пространством называется метрическое пространство S , в котором метрика ρ удовлетворяет *неравенству ультраметричности*: для любых $x, y, z \in S$

$$\rho(x, y) \leq \max\{\rho(x, z), \rho(y, z)\}. \quad (2)$$

Из этого неравенства следует, что в ультраметричном пространстве все треугольники либо равнобедренные, либо равносторонние. Кроме того, можно доказать следующее свойство ультраметричного пространства.

ЛЕММА. Пусть в ультраметричном пространстве даны четыре точки x_1, x_2, x_3, x_4 , и при этом $\rho(x_1, x_2) = \min_{1 \leq i < j \leq 4} \rho(x_i, x_j)$. Тогда выполняется неравенство

$$\rho(x_1, x_2) + \rho(x_3, x_4) \leq \rho(x_1, x_3) + \rho(x_2, x_4). \quad (3)$$

Рассмотрим алгоритмы, которые в случае ультраметричного пространства позволяют найти решения вышеперечисленных задач. Оказывается, что такими алгоритмами являются “жадные” алгоритмы.

2.1. Алгоритм решения задачи МП. Пусть имеется множество X , состоящее из n точек. Выберем из этого множества такие точки x_1, x_2 , что

$$\rho(x_1, x_2) = \min_{x, y \in X} \rho(x, y), \quad (4)$$

и объединим эти точки в пару. Среди оставшихся $n - 2$ точек снова найдем две ближайшие точки и объединим их в пару и т.д. Когда в множестве X не останется свободных точек для объединения в пару, получим решение исходной задачи.

Докажем, что приведенный алгоритм позволяет найти оптимальное решение задачи МП в ультраметричном пространстве. Для этого достаточно показать, что существует оптимальное решение G , содержащее пару (x_1, x_2) , обладающую свойством (4). Зафиксируем произвольное оптимальное решение G_1 и предположим, что оно не содержит данную пару.

Возможны два случая: либо в разбиении G_1 вершинам x_1 и x_2 соответствуют парные вершины y_1 и y_2 , либо одна из этих точек (предположим, точка x_1) вовсе не имеет пары, что возможно в случае нечетного числа точек.

Если вершина x_1 в решении G_1 не имеет пары, а точке x_2 соответствует парная вершина y , то от G_1 переходим к G заменой пары (x_2, y) на пару (x_1, x_2) . Очевидно, полученное решение G также является МП и содержит пару (x_1, x_2) .

Если же вершины x_1 и x_2 в разбиении G_1 спарены соответственно с вершинами y_1 и y_2 , то переходим к графу G заменой ребер (x_1, y_1) , (x_2, y_2) на ребра (x_1, x_2) , (y_1, y_2) .

В силу леммы данная замена не приводит к увеличению суммы длин ребер.

2.2. Алгоритм решения задачи МОД. Пусть имеется множество X , состоящее из n точек. Выберем из него произвольную точку x_1 , найдем для нее точку $x_2 \in X$, для которой $\rho(x_1, x_2) = \min_{y \in X, y \neq x_1} \rho(x_1, y)$ и зафиксируем ребро (x_1, x_2) в решении. Затем применим данную процедуру к $n-1$ точкам множества $X \setminus x_1$ и т.д. В конце концов полученное множество зафиксированных ребер будет образовывать граф, который, как легко видеть, является остовным деревом.

Пользуясь методом индукции, докажем, что данное остовное дерево является оптимальным решением. Предположим, что для $|X| = n-1$ это верно (что в случае двух точек очевидно), и докажем данное утверждение для $|X| = n$.

Решение G , построенное по предложенному алгоритму, является объединением ребра (x_1, x_2) и полученного по тому же алгоритму МОД G' для $n-1$ точек множества $X \setminus \{x_1\}$. Следовательно, для доказательства утверждения достаточно проверить, что существует МОД G_1 , являющееся объединением ребра (x_1, x_2) и некоторого МОД G'' для $n-1$ точек множества $X \setminus \{x_1\}$. Действительно, если бы такое МОД G_1 существовало, то G также было бы МОД.

Сначала докажем, что существует оптимальное решение G_2 , содержащее ребро (x_1, x_2) , для которого $\rho(x_1, x_2) = \min_{y \in X, y \neq x_1} \rho(x_1, y)$. Пусть G_3 — произвольное МОД, и пусть в G_3 точки x_1, x_2 не соединены ребром. Так как G_3 — дерево, то существует *единственный* путь, соединяющий точки x_1, x_2 . Соединим их в G_3 ребром, тем самым получив единственный цикл, который содержит ребро (x_1, x_2) . Удалим из данного цикла ребро (x_1, y) , отличное от ребра (x_1, x_2) . Легко видеть, что полученный граф G_2 является также остовным деревом, сумма длин ребер которого не превышает суммы длин ребер G_3 и, следовательно, является МОД.

Итак, существует оптимальное решение G_2 , содержащее ребро (x_1, x_2) . Теперь необходимо доказать, что существует оптимальное решение G_1 , в котором содержится ребро (x_1, x_2) , и не содержится других ребер вида (x_1, y) .

Для этого покажем, что если в решении G_2 , содержащем ребро (x_1, x_2) , содержится ребро (x_1, y) , $y \neq x_2$, то ребро (x_1, y) может быть заменено на ребро (x_2, y) , и полученный граф G_4 также будет являться МОД.

Легко видеть, что граф G_4 является допустимым остовным деревом и сумма длин ребер дерева G_4 не превышает суммы длин ребер G_2 . Последнее следует из того, что в ультраметричном пространстве из неравенства $\rho(x_1, x_2) \leq \rho(x_1, y)$ следует равенство $\rho(x_1, y) = \rho(x_2, y)$.

Заменяя в G_2 все ребра вида (x_1, y) , $y \neq x_2$, на ребра (x_2, y) , мы получаем оптимальное решение G_1 , содержащее ребро (x_1, x_2) и не содержащее других ребер вида (x_1, y) .

Очевидно, что полученное дерево G_1 является объединением ребра (x_1, x_2) и некоторого МОД G'' для $n - 1$ точек множества $X \setminus \{x_1\}$. Следовательно, для $|X| = n$ утверждение доказано.

2.3. Алгоритм решения ЗК. (Заметим, что данный алгоритм известен как “метод ближайшего города”.)

Пусть имеется множество X , состоящее из n точек, и произвольная начальная точка маршрута $x_1 \in X$. Найдем для x_1 точку x_2 такую, что

$$\rho(x_1, x_2) = \min_{y \in X, y \neq x_1} \rho(x_1, y).$$

Зафиксируем ребро (x_1, x_2) в решении. Затем применим данную процедуру к точкам множества $X \setminus x_1$, выбрав начальной точку x_2 , и т.д. В итоге получим путь $x_1 - x_2 - \dots - x_n$. Добавлением ребра (x_n, x_1) получим маршрут $x_1 - x_2 - \dots - x_n - x_1$.

Докажем, что указанный алгоритм позволяет найти оптимальное решение ЗК в ультраметричном пространстве. Пусть G — произвольное оптимальное решение ЗК. Предположим, что оно отличается от решения $x_1 - x_2 - \dots - x_n - x_1$, найденного по предложенному алгоритму. Это означает, что существуют такие значения $i < n$, что G не содержит всю цепочку $x_1 - x_2 - \dots - x_i - x_{i+1}$. Обозначим через k минимальное значение i . Легко видеть, что $1 \leq k < n - 1$.

В соответствии со сделанными предположениями оптимальный маршрут G имеет вид $x_1 - x_2 - \dots - x_k - y_1 - \dots - y_j - x_{k+1} - z_1 - \dots - z_s - x_1$, где $\{y_1, \dots, y_j, z_1, \dots, z_s\}$ — некоторая перестановка вершин $x_i, i > k + 1$, причем $j \geq 1$.

Покажем, что существует решение G_1 , сумма длин ребер которого не превышает суммы длин ребер G , и в котором содержатся все ребра цепочки $x_1 - x_2 - \dots - x_{k+1}$. Действительно, если рассмотреть маршрут G_1 вида $x_1 - x_2 - \dots - x_{k+1} - y_j - \dots - y_1 - z_1 - \dots - z_s - x_1$, то длина маршрута G_1 не больше длины G , поскольку $L(G_1) - L(G) = \rho(x_k, x_{k+1}) + \rho(y_1, z_1) - \rho(x_k, y_1) - \rho(x_{k+1}, z_1) \leq 0$ в силу леммы.

Таким образом, от решения, содержащего k ребер цепочки $x_1 - \dots - x_n$, можно перейти к решению, содержащему $k + 1$ ребер. Повторяя данную процедуру, мы в конце концов получим решение, совпадающее с маршрутом $x_1 - \dots - x_n - x_1$.

3. СРЕДНИЕ ЗНАЧЕНИЯ ОПТИМАЛЬНЫХ РЕШЕНИЙ ЗАДАЧ КОМБИНАТОРНОЙ ОПТИМИЗАЦИИ В p -АДИЧЕСКИХ ПРОСТРАНСТВАХ

Примером континуального ультраметричного пространства является поле p -адических чисел Q_p . Напомним, что поле p -адических чисел Q_p определяется как пополнение поля рациональных чисел Q по норме, задаваемой следующим правилом: если $m/n \in Q$ представляется в виде $m/n = p^\alpha m'/n'$, где m' и n' взаимно просты с p , p — фиксированное простое число, то нормой является $|m/n|_p = p^{-\alpha}$.

Пусть Q_p^d — d -мерное p -адическое пространство:

$$Q_p^d = \{x = (x_1, \dots, x_d): x_i \in Q_p, i = 1, \dots, d\}.$$

Зададим норму на Q_p^d как $|x|_p = \max_i |x_i|_p$. Заметим, что пространство Q_p^d также ультратрично относительно расстояния $|x - y|_p$. Элементы из Q_p^d можно складывать и умножать на p -адические числа.

Обозначим через $Z_p^d = \{x \in Q_p^d: |x|_p \leq 1\}$ единичный шар (“целых p -адических векторов”). Легко видеть, что он разбивается на p^d шаров диаметра p^{-1} :

$$Z_p^d = \bigcup_j V_j,$$

где $V_j = \{j + pZ_p^d\}$, j пробегает множество векторов $J = \{(j_1, \dots, j_d): j_i \in \{0, \dots, p-1\}, i = 1, \dots, d\}$. Нормированная трансляционно-инвариантная мера μ на Q_p^d определяется следующим простым соотношением: $\mu(Z_p^d) = 1$, $\mu(V_j) = p^{-d}$, $j \in J$.

В свою очередь, шары V_j также разбиваются на p^d шаров диаметра p^{-2} и объема p^{-2d} , и т.д. В дальнейшем шар диаметра p^{-k} , содержащийся в единичном шаре Z_p^d , будем называть шаром k -го уровня.

В этом разделе мы рассмотрим стохастические версии задач комбинаторной оптимизации, в которых точки разбрасываются случайным образом по шару Z_p^d равномерно по мере Хаара. Наша задача состоит в том, чтобы вычислить средние (относительно меры Хаара) значения оптимальных решений в задачах МП, МОД и ЗК. Обозначим эти средние как $L_n^{\text{МП}}$, $L_n^{\text{МОД}}$, $L_n^{\text{ЗК}}$, соответственно, где n — число точек.

ТЕОРЕМА 1. Пусть $L_n^{\text{МП}}$ — среднее значение оптимальных решений задачи МП для n точек, равномерно распределенных по мере Хаара в шаре Z_p^d . Тогда верна формула

$$L_n^{\text{МП}} = -\frac{1 - (-1)^n}{4} + \frac{p-1}{4} \sum_{k=1}^{\infty} p^{k(d-1)} (1 - (1 - 2p^{-kd})^n). \quad (5)$$

ДОКАЗАТЕЛЬСТВО. Разобьем исходный шар радиуса 1 на p^d шаров радиуса p^{-1} . Расстояние между двумя точками, одновременно расположенными в одном из этих шаров, будет не больше p^{-1} . Расстояние же между двумя точками из разных шаров равно 1.

Будем рассматривать решение задачи МП, полученное по описанному ранее алгоритму. Так как на каждом этапе алгоритма фиксируется ребро наименьшей длины и расстояния между точками одного шара 1-го уровня всегда меньше расстояний между точками из разных шаров 1-го уровня, то алгоритм будет “предпочитать” пары точек из одного шара 1-го уровня парам точек из разных шаров 1-го уровня.

Рассмотрим этап алгоритма, на котором было зафиксировано последнее ребро, соединяющее точки, лежащие в одном шаре 1-го уровня. (Если таких ребер нет, подсчет суммы длин ребер МП не вызывает затруднений, так как это означает, что в каждом шаре содержится не более одной точки, и все ребра МП имеют длину 1.)

Рассмотрев объединение всех ребер, зафиксированных до этого момента включительно, легко видеть, что точки каждого шара 1-го уровня, рассмотренные отдельно от остальных точек, разбились по парам оптимальным образом, и на каждый шар, в котором выпало нечетное число точек (“нечетный шар”), приходится одна точка, не имеющая пары (“свободная” точка). На последующих этапах разбиваются по парам эти свободные точки, а составляющие данное разбиение ребра имеют длину 1.

Таким образом, длина минимального паросочетания для n точек единичного p -адического шара складывается из суммы длин минимальных паросочетаний для каждого из p^d шаров 1-го уровня и суммы длин ребер, соединяющих свободные точки нечетных шаров.

Занумеруем шары 1-го уровня индексом i , $1 \leq i \leq p^d$. Пусть n_i – число точек, попавших в шар с номером i . Так как радиус этого шара равен p^{-1} , то среднее значение длины МП для точек, попавших в i -й шар, равно $L_{n_i}^{\text{МП}}/p$.

Очевидно, что число нечетных шаров равно

$$\sum_{i=1}^{p^d} \frac{1 - (-1)^{n_i}}{2}.$$

Сумма длин ребер, соединяющих в пары свободные точки всех нечетных шаров, равна половине данной величины, точнее целой части от половины данной величины, т.е.

$$\sum_{i=1}^{p^d} \frac{1 - (-1)^{n_i}}{4} - \frac{1 - (-1)^n}{4}.$$

Итак, при условии, что распределение точек по шарам задается набором $(n_1, n_2, \dots, n_{p^d})$, средняя величина оптимального паросочетания равна

$$\sum_{i=1}^{p^d} \left(\frac{L_{n_i}^{\text{МП}}}{p} + \frac{1 - (-1)^{n_i}}{4} \right) - \frac{1 - (-1)^n}{4}.$$

Так как распределение целочисленного вектора $(n_1, n_2, \dots, n_{p^d})$ подчиняется мультиномиальному закону, то величина $L_n^{\text{МП}}$ удовлетворяет следующему рекуррентному уравнению (по формуле полной вероятности):

$$L_n^{\text{МП}} = \frac{1}{p^{dn}} \sum_{n_1+n_2+\dots+n_{p^d}=n} \frac{n!}{n_1! \dots n_{p^d}!} \left[\sum_{i=1}^{p^d} \left(\frac{L_{n_i}^{\text{МП}}}{p} + \frac{1 - (-1)^{n_i}}{4} \right) - \frac{1 - (-1)^n}{4} \right]. \quad (6)$$

Обозначив

$$H_n = \frac{4L_n^{\text{МП}} + 1 - (-1)^n}{p^d(1 - p^{-1})}, \quad (7)$$

нетрудно из (6) получить уравнение

$$H_n = \frac{p^{d-1}}{p^{nd}} \sum_{k=0}^n C_n^k H_k (p^d - 1)^{n-k} + 1 - \left(\frac{p^d - 2}{p^d} \right)^n. \quad (8)$$

Вычисления показывают, что решение уравнения (8) имеет вид

$$H_n = p^{1-d} \sum_{k=1}^{\infty} p^{k(d-1)} (1 - (1 - 2p^{-kd})^n). \quad (9)$$

Из (7) и (9) следует утверждение теоремы 1.

ТЕОРЕМА 2. Пусть $L_n^{\text{МОД}}$ – среднее оптимальное решение задачи МОД для n точек, равномерно распределенных по мере Хаара в шаре Z_p^d . Тогда верна формула

$$L_n^{\text{МОД}} = (p-1) \sum_{k=0}^{\infty} p^{k(d-1)} (1 - (1 - p^{-kd})^n) - p\delta_n, \quad (10)$$

где

$$\delta_n = \begin{cases} 0, & n = 0, \\ 1, & n \geq 1. \end{cases}$$

ДОКАЗАТЕЛЬСТВО. Будем рассматривать вариант описанного ранее алгоритма для задачи МОД, согласно которому на каждом этапе мы выбираем не любую точку x_1 , а такую, для которой существуют точки, находящиеся с x_1 в одном шаре. И только если таких точек x_1 нет, то x_1 – любая точка.

Аналогично рассуждениям, использованным при доказательстве теоремы 1, заметим, что длина МОД складывается из суммы длин МОД для каждого из шаров 1-го уровня и из суммы длин ребер длины 1, соединяющих деревья из разных шаров в одно дерево. Легко видеть, что число ребер длины 1 в этом дереве равно $\sum_{i=1}^{p^d} \delta_{n_i} - \delta_n$.

Отсюда получаем следующее рекуррентное соотношение для $L_n^{\text{МОД}}$:

$$L_n^{\text{МОД}} = \frac{1}{p^{dn}} \sum_{n_1+n_2+\dots+n_{p^d}=n} \frac{n!}{n_1! \dots n_{p^d}!} \left[\sum_{i=1}^{p^d} \left(\frac{L_{n_i}^{\text{МОД}}}{p} + \delta_{n_i} \right) - \delta_n \right], \quad (11)$$

где $L_0^{\text{МОД}} = 0$.

Пусть

$$V_n = \frac{L_n^{\text{МОД}} + p\delta_n}{p-1}.$$

Тогда из соотношения (11) нетрудно получить, что

$$V_n = \frac{p^{d-1}}{p^{nd}} \sum_{k=0}^n C_n^k V_k (p^d - 1)^{n-k} + \delta_n.$$

Данному соотношению удовлетворяют только значения

$$V_n = \sum_{k=0}^{\infty} p^{k(d-1)} (1 - (1 - p^{-kd})^n),$$

откуда и получается результат для $L_n^{\text{МОД}}$.

ТЕОРЕМА 3. Пусть $L_n^{\text{ЗК}}$ – среднее оптимальное решение ЗК для n точек, равномерно распределенных по мере Хаара в шаре Z_p^d . Тогда верна формула

$$L_n^{\text{ЗК}} = L_n^{\text{МОД}} + \delta_n \frac{1 - p^{-d(n-1)}}{1 - p^{-d(n-1)-1}}. \quad (12)$$

ДОКАЗАТЕЛЬСТВО. Для $n = 0$ утверждение очевидно. Докажем его для $n > 0$.

Сравнение алгоритмов для ЗК и задачи МОД в ультраметричном пространстве показывает, что первые $n - 1$ ребер маршрута $x_1 - \dots - x_n$, построенного по алгоритму ЗК, дают оптимальное решение задачи МОД. Следовательно, разность между решениями ЗК и задачи МОД равна длине ребра (x_n, x_1) , присоединяемого по алгоритму ЗК в последнюю очередь. Легко видеть, что эта длина равна диаметру наименьшего шара, содержащего все n точек, т.е. равна 1, если существуют хотя бы два непустых шара 1-го уровня, p^{-1} , если есть только один непустой шар 1-го уровня и хотя бы два непустых шара 2-го уровня, и т.д.

Итак, если все точки попали в один шар k -го уровня, но есть хотя бы два непустых шара $(k+1)$ -го уровня, то разность между решениями ЗК и задачи МОД составит p^{-k} . Так как вероятность того, что все точки попадут в один шар k -го уровня, равна $p^{-kd(n-1)}$, то с вероятностью $p^{-dk(n-1)} - p^{-d(k+1)(n-1)}$ диаметр наименьшего шара, содержащего все n точек, равен p^{-k} .

Отсюда

$$L_n^{\text{ЗК}} - L_n^{\text{МОД}} = \sum_{k=0}^{\infty} (p^{-dk(n-1)} - p^{-d(k+1)(n-1)}) p^{-k} = \frac{1 - p^{-d(n-1)}}{1 - p^{-d(n-1)-1}}.$$

Утверждение теоремы 3 доказано.

Обратимся к вопросу асимптотического поведения величин $L_n^{\text{МП}}$, $L_n^{\text{МОД}}$ и $L_n^{\text{ЗК}}$ при $n \rightarrow \infty$. Заметим, что задача отыскания асимптотики этих величин сводится к отысканию асимптотики ряда

$$\sum_{k=0}^{\infty} p^{k(d-1)} (1 - (1 - ap^{-kd})^n),$$

где $0 < a < 1$, поскольку мы можем записать

$$L_n^{\text{МП}} = -\frac{1 - (-1)^n}{4} + \frac{p-1}{4} p^{(d-1)} \sum_{k=0}^{\infty} p^{k(d-1)} \left(1 - \left(1 - \frac{2}{p^d} p^{-kd} \right)^n \right),$$

$$L_n^{\text{МОД}} = (p-1) p^{(d-1)} \sum_{k=0}^{\infty} p^{k(d-1)} \left(1 - \left(1 - \frac{1}{p^d} p^{-kd} \right)^n \right) - \delta_n.$$

ТЕОРЕМА 4. Пусть

$$G_n = \sum_{k=0}^{\infty} p^{k(d-1)} (1 - (1 - ap^{-kd})^n), \quad 0 < a < 1.$$

При $d = 1$ величина G_n представима в виде

$$G_n = \log_p n + O(1), \quad n \rightarrow \infty. \quad (13)$$

В случае $d > 1$ существуют $c_1, c_2 > 0$ такие, что для всех n выполняется неравенство

$$c_1 < \frac{G_n}{n^{1-1/d}} < c_2. \quad (14)$$

Предела величин $G_n/n^{1-1/d}$ при $n \rightarrow \infty$ не существует. Однако при фиксированном вещественном x и натуральном параметре m существует предел

$$\lim_{m \rightarrow \infty} \frac{G_{p^{(m+x)d}}}{(p^{(m+x)d})^{1-1/d}} = U_d(x), \quad (15)$$

где $U_d(x)$ — периодическая функция с периодом 1, равная

$$U_d(x) = \sum_{k=-\infty}^{+\infty} p^{(1-d)(k+x)} (1 - e^{-ap^{d(k+x)}}). \quad (16)$$

ДОКАЗАТЕЛЬСТВО. Пусть $d = 1$. Тогда можно записать

$$\begin{aligned} G_n &= \sum_{k=0}^{\infty} p^{k(d-1)} (1 - (1 - ap^{-kd})^n) = an \sum_{k=0}^{\infty} \int_0^{1/p^k} (1 - ax)^{n-1} dx = \\ &= an \sum_{k=0}^{\infty} \sum_{j=k}^{\infty} \int_{1/p^{j+1}}^{1/p^j} (1 - ax)^{n-1} dx = an \sum_{j=0}^{\infty} (j+1) \int_{1/p^{j+1}}^{1/p^j} (1 - ax)^{n-1} dx. \end{aligned}$$

При $x \in [1/p^{k+1}, 1/p^k]$ выполняются неравенства $-\log_p x \leq k+1 \leq -\log_p x + 1$. Поэтому можно записать

$$\begin{aligned} an \sum_{j=0}^{\infty} \int_{1/p^{j+1}}^{1/p^j} (-\log_p x) (1 - ax)^{n-1} dx &\leq \\ &\leq G_n \leq an \sum_{j=0}^{\infty} \int_{1/p^{j+1}}^{1/p^j} (-\log_p x + 1) (1 - ax)^{n-1} dx, \end{aligned}$$

откуда

$$0 \leq G_n - W_n \leq 1 - (1 - a)^n, \quad (17)$$

где

$$W_n = \frac{an}{\ln p} \int_0^1 (-\ln x)(1-ax)^{n-1} dx.$$

Сделав замену $y = -\ln x$, получим

$$\begin{aligned} W_n &= \frac{an}{\ln p} \int_0^\infty y(1-ae^{-y})^{n-1} e^{-y} dy = \frac{1}{\ln p} \lim_{A \rightarrow \infty} \int_0^A y d(1-ae^{-y})^n = \\ &= \frac{1}{\ln p} \lim_{A \rightarrow \infty} \left[y(1-ae^{-y})^n \Big|_0^A - \int_0^A (1-ae^{-y})^n dy \right] = \\ &= \frac{1}{\ln p} \lim_{A \rightarrow \infty} \left[A(1-ae^{-A}) - \int_0^A (1-ae^{-y})^n dy \right] = \\ &= \frac{1}{\ln p} \lim_{A \rightarrow \infty} \left[A - Aae^{-A} - A + \sum_{k=1}^n \frac{(1-ae^{-A})^k}{k} - \sum_{k=1}^n \frac{(1-a)^k}{k} \right] = \\ &= \frac{1}{\ln p} \left[\sum_{k=1}^n \frac{1}{k} - \sum_{k=1}^n \frac{(1-a)^k}{k} \right], \end{aligned}$$

откуда

$$\lim_{n \rightarrow \infty} (W_n - \log_p n) = \frac{\gamma + \ln a}{\ln p}.$$

Из этого последнего соотношения и неравенства (17) следует первое утверждение теоремы 4.

Пусть теперь $d > 1$. Оценка (14) может быть получена тем же способом, каким было доказано представление (13). Докажем формулу (15).

Отметим сначала, что в сходимости двухстороннего ряда (16), определяющего функцию $U_d(x)$, легко убедиться, используя тот факт, что $1 - e^{-x} = x + O(x^2)$.

Обозначим

$$\begin{aligned} I_m(x) &= \sum_{k=0}^{\infty} p^{k(d-1)} \left(1 - \left(1 - \frac{a}{p^{kd}} \right)^{p^{(m+x)d}} \right) = G_{p^{(m+x)d}}, \\ J_m(x) &= \sum_{k=0}^{\infty} p^{k(d-1)} \left(1 - \left(1 - \frac{a}{p^{kd}} \right)^{p^{(m+x)d} - ap^{d(m+x-k)}} \right), \\ D_m(x) &= \sum_{k=0}^{\infty} p^{k(d-1)} (1 - e^{-ap^{d(m+x-k)}}). \end{aligned}$$

Нужно доказать, что для любого x выполняется равенство

$$\lim_{m \rightarrow \infty} p^{(1-d)(m+x)} I_m(x) = U_d(x). \quad (18)$$

Убедимся сначала, что $\lim_{m \rightarrow \infty} p^{(1-d)(m+x)} D_m(x) = U_d(x)$. Действительно, при $m \rightarrow \infty$

$$\begin{aligned} p^{(1-d)(m+x)} D_m(x) &= p^{(1-d)(m+x)} \sum_{k=0}^{\infty} p^{k(d-1)} (1 - e^{-ap^{d(m+x-k)}}) = \\ &= \sum_{k=-m}^{\infty} p^{(1-d)(x-k)} (1 - e^{-ap^{d(x-k)}}) = \\ &= \sum_{k=-\infty}^m p^{(1-d)(x+k)} (1 - e^{-ap^{d(x+k)}}) \rightarrow \\ &\rightarrow \sum_{k=-\infty}^{+\infty} p^{(1-d)(x+k)} (1 - e^{-ap^{d(x+k)}}). \end{aligned}$$

Следовательно, для доказательства равенства (18) достаточно показать, что $|I_m - D_m| = o(p^{d(m+x)})$. Докажем, что в действительности $|I_m - D_m| = O(1)$, т.е. существует постоянная величина C , не зависящая от m , такая, что $|I_m - D_m| \leq C$.

Воспользуемся тем фактом, что для любых α и β , $0 < \alpha < \beta$, выполняется неравенство

$$\left(1 - \frac{\alpha}{\beta}\right)^{\beta-\alpha} \geq e^{-\alpha} \geq \left(1 - \frac{\alpha}{\beta}\right)^{\beta}.$$

Отсюда легко видеть, что

$$0 \leq I_m - D_m \leq I_m - J_m.$$

Покажем, что $I_m - J_m \leq C$ для некоторой постоянной величины C :

$$\begin{aligned} I_m - J_m &= \sum_{k=0}^{\infty} p^{(d-1)k} \left(1 - \frac{a}{p^{dk}}\right)^{p^{d(m+x)}} \left(\left(1 - \frac{a}{p^{dk}}\right)^{-ap^{d(m+x-k)}} - 1 \right) = \\ &= a \sum_{k=0}^{\infty} p^{(d-1)k} \left(1 - \frac{a}{p^{dk}}\right)^{p^{d(m+x)}} p^{d(m+x-k)} \int_0^{ap^{-dk}} (1-y)^{-ap^{d(m+x-k)}-1} dy = \\ &= ap^{d(m+x)} \sum_{k=0}^{\infty} p^{-k} \left(1 - \frac{a}{p^{dk}}\right)^{p^{d(m+x)}} \sum_{j=k}^{\infty} \int_{ap^{-d(j+1)}}^{ap^{-dj}} (1-y)^{-ap^{d(m+x-k)}-1} dy = \\ &= ap^{d(m+x)} \sum_{j=0}^{\infty} \int_{ap^{-d(j+1)}}^{ap^{-dj}} \sum_{k=0}^j p^{-k} \left(1 - \frac{a}{p^{dk}}\right)^{p^{d(m+x)}} (1-y)^{-ap^{d(m+x-k)}-1} dy \leq \\ &\leq ap^{d(m+x)} \int_0^a \sum_{k=0}^{\infty} p^{-k} (1-y)^{p^{d(m+x)} - ap^{d(m+x-k)}-1} dy = \\ &= ap^{d(m+x)} \sum_{k=0}^{\infty} p^{-k} \frac{1 - (1-a)^{p^{d(m+x)} - ap^{d(m+x-k)}}}{p^{d(m+x)}(1 - ap^{-dk})} \leq a \sum_{k=0}^{\infty} \frac{p^{-k}}{1 - ap^{-dk}} = C. \end{aligned}$$

Утверждение доказано.

Следствием данного утверждения является то, что при $d = 1$

$$\begin{aligned} L_n^{\text{МП}} - \frac{p-1}{4} \log_p n &= O(1), \\ L_n^{\text{МОД}} - (p-1) \log_p n &= O(1), \\ L_n^{\text{ЗК}} - (p-1) \log_p n &= O(1). \end{aligned}$$

При $d > 1$ получаем довольно неожиданный результат: оказывается, в этом случае общей асимптотики оптимальных решений рассмотренных задач не существует. Для разных последовательностей $\{n_k\}$, $n_k \rightarrow \infty$, могут существовать разные пределы $L_{n_k}/n_k^{1-1/d}$, хотя множество предельных точек ограничено. Такое поведение отличается от поведения в евклидовом случае, для которого доказана асимптотика $L_n/n^{1-1/d} \rightarrow \infty$.

Список литературы

- [1] *M. Mezard, G. Parisi, M. Virasoro*. Spin-Glass Theory and All Beyond. World Sci. Lect. Notes Phys. V. 9. Singapore: World Scientific, 1987; *S. Kirkpatrick, G. Toulouse*. J. Physique. 1985. V. 46. P. 1277–1292.
- [2] *J. Beardwood, J. H. Halton, J. M. Hammersley*. Proc. Cambridge Philos. Soc. 1959. V. 55. P. 299–327; *J. M. Steele*. Probability Theory and Combinatorial Optimization. Philadelphia: SIAM, 1997; *J. Yukich*. Quasi-additive Euclidean functionals. In: Discrete Probability and Algorithms. Proc. of the Workshops “Probability and Algorithms” and “The Finite Markov Chain Renaissance” (IMA, Univ. of Minnesota, Minneapolis, MN, USA, 1993). IMA Vol. Math. Appl. V. 72. Eds. D. Aldous et al. New York: Springer, 1995. P. 149–158.
- [3] *P. M. Bleher, Ya. G. Sinai*. Commun. Math. Phys. 1973. V. 33. P. 23–42.
- [4] *М. Д. Миссаров*. ТМФ. 1998. Т. 114. № 3. С. 323–336; Т. 117. № 3. С. 471–488.
- [5] *И. В. Волович*. ТМФ. 1987. Т. 71. С. 337–340; *В. С. Владимиров, И. В. Волович, Е. И. Зеленов*. p -Адический анализ и математическая физика. М.: Наука, 1994.
- [6] *E. Yu. Lerner, M. D. Missarov*. Commun. Math. Phys. 1989. V. 32. P. 347–356.

Поступила в редакцию 15.V.2002 г.