ДИСКРЕТНЫЙ АНАЛИЗ И ИССЛЕДОВАНИЕ ОПЕРАЦИЙ
Апрель—июнь 2007. Серия 1. Том 14, № 2. 95–101

УДК 519.176

АЛГОРИТМ ПРИБЛИЖЁННОГО РЕШЕНИЯ ОДНОМЕРНОЙ ЗАДАЧИ О ПОСЛЕДОВАТЕЛЬНОСТИ МЕДИАН*)

В. В. Шенмайер

Задача о последовательности медиан (online median) заключается в отыскании последовательности вложенных медиан возрастающей мощности. Рассматривается частный случай задачи, когда клиенты и предприятия расположены в точках вещественной прямой. Лучший известный алгоритм для одномерного случая имеет относительную оценку точности решений 8. Предлагается полиномиальный алгоритм с оценкой точности 5,83.

Введение

Задача о последовательности медиан отличается от обычной задачи о k-медиане тем, что число k открываемых предприятий в ней не задано — вместо этого требуется последовательно открывать предприятия таким образом, чтобы на каждом шаге открытые к этому времени предприятия образовывали решение (точное либо приближённое) задачи о медиане соответствующей мощности.

Пусть заданы: конечное множество клиентов (потребителей) C; конечное множество предприятий F; расстояние $d(u, f) \geqslant 0$, определённое для каждого клиента u и предприятия f; а также вес $w(u) \geqslant 0$, определённый для каждого клиента u. Стоимость $\operatorname{cost}(X)$ произвольного множества предприятий X равна взвешенной сумме $\sum\limits_{u\in C} w(u)\, d(u, X)$, где расстояние $d(u, X)$ от клиента u до множества X есть расстояние от u до ближайшего предприятия из множества X, т. е. $d(u, X) = \min\limits_{f\in X} d(u, f)$.

В задаче о k-медиане задано также положительное целое число k. Требуется найти множество предприятий X мощности k минимальной стоимости среди всех множеств мощности k.

Задача о последовательности медиан состоит в отыскании последовательности таких вложенных множеств предприятий $F_1 \subset F_2 \subset \cdots \subset F_n$,

*)Исследование выполнено при финансовой поддержке Российского фонда фундаментальных исследований (проект 05–01–00395).

где n — мощность множества F, что при каждом $k = 1, 2, \ldots, n$ мощность множества F_k равна k, а стоимость минимальна среди всех множеств мощности k.

Поскольку в большинстве случаев такой последовательности вложенных точных медиан, по-видимому, не существует, имеет смысл говорить о приближённых решениях задачи. Говорим, что решение $F_1, F_2, \ldots, F_n$ задачи о последовательности медиан имеет точность δ, если при любом $k = 1, 2, \ldots, n$ для любого множества предприятий X мощности k имеет место неравенство $\mathrm{cost}(F_k) \leqslant \delta \, \mathrm{cost}(X)$. Другими словами, каждое множество F_k отличается по стоимости от точного решения задачи о k-медиане не более чем в δ раз.

Первый константный алгоритм для метрической задачи о последовательности медиан был получен в работе [4]. Наилучший известный алгоритм имеет оценку точности решений $8\,c$, где c — точность решения задачи о k-медиане [2]. Таким образом, в классе полиномиальных алгоритмов задача может быть решена с точностью $24+\varepsilon$, поскольку лучший известный полиномиальный алгоритм для решения задачи о k-медиане имеет точность $3 + \varepsilon$ [1].

Из отрицательных результатов отметим предел неприближаемости в классе полиномиальных алгоритмов для задачи о k-медиане, равный $1+2/e$ ($\approx 1{,}74$) [3], и предел неприближаемости в классе всех детерминированных алгоритмов для задачи о последовательности медиан, равный $2 - 1/(n - 1)$ [4].

Рассмотрим частный случай задачи о последовательности медиан, когда клиенты и предприятия расположены в точках вещественной прямой. Несмотря на то, что данный случай является наиболее простой моделью для задач размещения, рассматриваемая задача остаётся на нём приблизительно такой же труднорешаемой как и в общем случае. Лучший известный алгоритм для одномерной задачи имеет оценку точности решений 8 [2] (в одномерном случае $c = 1$). Предлагается полиномиальный алгоритм, решающий одномерную задачу о последовательности медиан с точностью $(1 + \sqrt{2})^2$ (приблизительно 5,83).

1. Описание алгоритма

Основная идея алгоритма совпадает с идеей алгоритма из [2]. Алгоритм начинает свою работу с нахождения решений $F_1^*, F_2^*, \ldots, F_n^*$ обычной задачи о k-медиане при $k = 1, 2, \ldots, n$. В данном случае речь идёт о точных решениях, поскольку одномерная задача о k-медиане полиномиально разрешима.

На следующем шаге алгоритм находит множество K, состоящее из непродолжаемой последовательности номеров $i_1, i_2, \ldots, i_t$, в которой $i_1 = 1$ и номер i_{k+1}, $k = 1, 2, \ldots, t-1$, является минимальным номером таким, что $\mathrm{cost}(F^*_{i_{k+1}}) \leqslant \mathrm{cost}(F^*_{i_k})/\alpha$, где $\alpha = 1 + \sqrt{2}$. Отметим, что в случае использования алгоритма из [2] аналогичная константа равна 2.

Далее строится частичное решение задачи о последовательности медиан, соответствующее номерам из множества K. А именно, находится последовательность вложенных множеств $F_{i_1} \subset F_{i_2} \subset \cdots \subset F_{i_t}$, которые строятся рекурсивно в обратном порядке: сначала определяется множество F_{i_t}, затем множество $F_{i_{t-1}}$ и т. д. Множество F_{i_t} полагается равным $F^*_{i_t}$ и при $k = t-1, \ldots, 1$ множество F_{i_k} определяется следующим образом. Для каждого предприятия f из $F^*_{i_k}$ рассматривается множество $C(f)$ клиентов, для которых предприятие f является ближайшим среди предприятий из $F^*_{i_k}$ (такое множество называется *кластером* или *зоной обслуживания* предприятия f). Находятся ближайшие к f слева и справа предприятия из множества $F_{i_{k+1}}$ (по крайней мере одно из них существует). Из них выбирается лучшее с точки зрения минимизации взвешенной суммы расстояний до клиентов из кластера $C(f)$. Выбранные таким образом предприятия (для всех $f \in F^*_{i_k}$) образуют множество F_{i_k}. В дальнейшем описанную выше процедуру построения множества F_{i_k} будем обозначать через Γ, $F_{i_k} = \Gamma(F_{i_{k+1}}, F^*_{i_k})$.

Заметим, что число точек во множестве F_{i_k} не превосходит мощности множества $F^*_{i_k}$, равной i_k. Если мощность множества F_{i_k} меньше i_k, к F_{i_k} добавляется недостающее число произвольно выбранных точек из множества $F_{i_{k+1}} \setminus F_{i_k}$. Таким образом, частичное решение $F_{i_1}, F_{i_2}, \ldots, F_{i_t}$ задачи о последовательности медиан построено.

Полученное решение произвольным образом достраивается до полного путём добавления к множеству F_{i_k}, $k = 1, 2, \ldots, t$, либо точек из множества $F_{i_{k+1}} \setminus F_{i_k}$, если $k < t$, либо точек из множества $F \setminus F_{i_t}$, если $k = t$.

Теорема. *Описанный выше алгоритм имеет оценку точности решений* $\delta = (1 + \sqrt{2})^2$.

Отметим, что при построении множества F_{i_k} можно взять заведомо лучшее множество: решение задачи об i_k-медиане на множестве предприятий $F_{i_{k+1}}$. Данный вариант не только лучше с точки зрения целевой функции, но и более универсален, поскольку определён для произвольной метрической задачи, когда нет понятия «ближайший слева/справа». Заметим также, что в алгоритме из [2] при построении множества F_{i_k} замена предприятия f из $F^*_{i_k}$ определялась как ближайшее к нему пред-

приятие из множества $F_{i_{k+1}}$.

Сложность алгоритма определяется сложностью поиска начальной последовательности $F_1^*, F_2^*, \dots, F_n^*$, что в свою очередь совпадает со сложностью поиска одной k-медианы, поскольку метод динамического программирования, с помощью которого решается одномерная задача о k-медиане, вкладывает её в серию задач о k-медианах для разных k.

2. Доказательство верхней оценки

Определим следующее понятие. *Конфигурацией* назовём набор, состоящий из произвольного взвешенного множества клиентов C и двух множеств предприятий S, S^*, из которых второе является точным решением задачи о k-медиане, в которой $k = |S^*|$ и множеством клиентов является C. Обоснование верхней оценки точности решений, получаемых алгоритмом, основано на следующей лемме.

Лемма. *Для произвольной конфигурации C, S, S^* справедливо неравенство*

$$\mathrm{cost}(\Gamma(S, S^*)) \leqslant \beta\,(\mathrm{cost}(S) + \mathrm{cost}(S^*)), \text{ где } \beta = (1+\sqrt{2})/2. \tag{1}$$

Доказательство леммы будет приведено ниже.

Доказательство теоремы. Согласно лемме стоимость множества F_{i_k}, $k = 1, 2, \dots, t$, не превосходит суммы ряда $\beta\,\mathrm{cost}(F_{i_k}^*) + \beta^2\,\mathrm{cost}(F_{i_{k+1}}^*) + \dots$. Учитывая выбор номеров $i_1, i_2, \dots, i_t$, получаем

$$\mathrm{cost}(F_{i_k}) \leqslant \beta\,\mathrm{cost}(F_{i_k}^*)\,(1 + \beta/\alpha + (\beta/\alpha)^2 + \dots) = \mathrm{cost}(F_{i_k}^*)\,\beta\alpha/(\alpha - \beta).$$

Таким образом, получена верхняя оценка точности частичного решения $F_{i_1}, F_{i_2}, \dots, F_{i_t}$. Оценим стоимость произвольного множества F_s, $s = 1, 2, \dots, n$, $s \notin K$. Пусть i_k — максимальный номер из множества K такой, что $i_k < s$. Тогда $\mathrm{cost}(F_s^*) > \mathrm{cost}(F_{i_k}^*)/\alpha$ по построению множества K. Поскольку целевая функция не возрастает с ростом числа открытых предприятий, то отсюда следует, что

$$\mathrm{cost}(F_s) \leqslant \mathrm{cost}(F_{i_k}) \leqslant \mathrm{cost}(F_{i_k}^*)\,\beta\alpha/(\alpha-\beta) < \mathrm{cost}(F_s)\,\beta\alpha^2/(\alpha-\beta).$$

Заметим, что $\alpha = 2\beta$ (это значение является оптимальным для константы α). Следовательно, точность алгоритма не превосходит величины $4\beta^2$, равной $(1+\sqrt{2})^2$. Теорема доказана.

Доказательство леммы. Предположим, для некоторой конфигурации C, S, S^* выполнено обратное, т. е.

$$\mathrm{cost}(\Gamma(S, S^*)) > \beta\,(\mathrm{cost}(S) + \mathrm{cost}(S^*)). \tag{2}$$

Вывод противоречия будет состоять из двух этапов: построение существенно более простой конфигурации, в которой также выполнено неравенство (2), и вывод противоречия с использованием простоты полученной конфигурации. Первый этап состоит из шести описанных ниже упрощений конфигурации.

Упрощение 1. В силу неравенства (2) найдётся хотя бы одно предприятие $c \in S^*$ такое, что вклад клиентов кластера $C(c)$ в левую часть неравенства превосходит вклад в правую. Следовательно, если $S^* = \{c\}$, а множество клиентов $C = C(c)$, то неравенство (2) по-прежнему будет выполнено. Отметим, что получившийся набор $C, S, \{c\}$ является конфигурацией, поскольку точка c является 1-медианой своего кластера.

Введём некоторые обозначения. Пусть a, b — ближайшие соседи точки c из множества S слева и справа соответственно. Можно считать, что оба соседа существуют (в противном случае в качестве отсутствующего соседа можно взять бесконечно удалённую точку). Не нарушая общности, будем считать, что точка c находится не правее точки $d = (a+b)/2$ (случай, когда $c > d$, полностью симметричен и анализируется аналогично). Пусть f, f^c, f^a и f^b — стоимости множеств $S, \{c\}, \{a\}$ и $\{b\}$ соответственно. Тогда неравенство (2) принимает вид

$$\min(f^a, f^b) > \beta(f^c + f). \tag{3}$$

Упрощение 2. Рассмотрим интервал вещественной прямой слева от точки a. Сместим клиентов, расположенных в этом интервале, в точку a, а предприятия исключим из множества S. Заметим, что все величины, содержащиеся в неравенстве (3), не возрастут, причём величины f^c, f^a и f^b уменьшатся на одинаковую константу. Следовательно, неравенство (3) по-прежнему будет выполнено. Аналогично поступим с клиентами и предприятиями, расположенными правее точки b. Заметим, что после таких упрощений во множестве S останутся только предприятия a и b.

Упрощение 3. Заметим, что взвешенная сумма расстояний от предприятия до клиентов, расположенных от него по одну сторону, равна расстоянию от предприятия до центра масс этих клиентов, умноженному на их суммарный вес. Отсюда следует, что величины f^a, f^b, f^c и f не изменятся, если множество клиентов, расположенных на любом из интервалов $(a, c), (c, d)$ и (d, b), заменить на любое другое, имеющее прежний суммарный вес и центр масс. Пользуясь данным наблюдением, множество клиентов каждого из этих интервалов заменим на множество, состоящее из двух *клиентов, лежащих на краях интервала* и имеющих прежний суммарный вес и центр масс. В результате останется только четыре клиента a, c, d и b.

Упрощение 4. Сместим клиента d в точку c. Заметим, что при этом величина f^b не уменьшится, а величины f^a, f^c и f уменьшатся на одинаковую константу. Следовательно, неравенство (3) по-прежнему будет выполнено. Таким образом, число клиентов уменьшилось еще на одного.

Упрощение 5. Уменьшим веса клиентов a и b на величину $\Delta w = \min(w(a), w(b))$. Величина f при этом останется неизменной (поскольку в точках a и b находятся предприятия множества S), а величины f^a, f^b и f^c уменьшатся на одинаковую константу $\Delta w\,(b-a)$. Таким образом, неравенство (3) по-прежнему будет выполнено.

Заметим, что получившийся набор $C = \{a, c, b\}$, $S = \{a, b\}$, $\{c\}$ всё ещё является конфигурацией. Действительно, согласно известному свойству 1-медианы произвольная точка прямой является медианой тогда и только тогда, когда суммарный вес клиентов, расположенных по одну сторону от неё, не превосходит суммарного веса клиентов, расположенных с противоположной стороны и в самой точке [5]. Поскольку упрощения 2–5 не нарушают этого свойства, точка c по-прежнему является 1-медианой на множестве клиентов C.

Упрощение 6. Сделаем нормализацию весов клиентов и координат всех точек получившейся конфигурации. Поскольку после упрощения 5 один из клиентов a, b имеет нулевой вес, согласно указанному выше свойству 1-медианы вес клиента c должен быть максимальным. Следовательно, после нормализации получим конфигурацию, состоящую из трёх клиентов: $a = 0$, $c = x \leqslant 1/2$ и $b = 1$ с весами соответственно y, 1 и z, где $y, z \leqslant 1$ и один из весов y, z равен 0.

На этом первый (наиболее сложный) этап вывода противоречия закончен. Переходим ко второму этапу, на котором, используя простоту полученной конфигурации, покажем невыполнимость неравенства (3).

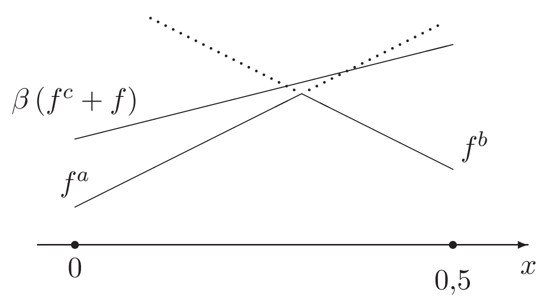

Рис. 1

Выпишем в явном виде, чему равны составляющие данного неравенства: $f^a = x+z$, $f^b = 1-x+y$, $f = x$, $f^c = xy+(1-x)z$. Поскольку правая

часть неравенства является линейной функцией от x, а левая — кусочно-линейной, неравенство должно быть выполнено либо на краях интервала $(0, 1/2)$, либо в точке излома левой части (рис. 1). При $x = 0$ неравенство не выполнено, поскольку $\min(f^a, f^b) = z$, $f = 0$ и $f^c = z$. Аналогично, при $x = 1/2$ имеем $f^a = 1/2 + z$, $f^b = 1/2 + y$, $f = 1/2$ и $f^c = y/2 + z/2$. Следовательно, $\min(f^a, f^b) \leqslant (f^a + f^b)/2 = 1/2 + z/2 + y/2 = f + f^c$, т. е. неравенство (3) также не выполнено.

Точка излома в левой части неравенства (3) определяется соотношением $f^a = f^b$, что означает, что $x + z = 1 - x + y$ и $x = (1 + y - z)/2$. Так как $x \leqslant 1/2$, то отсюда получаем, что $y \leqslant z$. Поскольку одна из переменных y, z равна нулю, то $y = 0$. Следовательно, $z = 1 - 2x$ и $f^c = (1-x)z = 1-3x+2x^2$. В результате неравенство (3) приобретает вид $1-x > \beta\,(x + 1-3x+2x^2)$ или $2\beta x^2 + x\,(1-2\beta) + \beta - 1 < 0$. Дискриминант этой квадратичной формы равен $1 - 4\beta + 4\beta^2 - 8\beta^2 + 8\beta = -4\beta^2 + 4\beta + 1$. Но константа β выбрана так, что последнее выражение равно нулю. Следовательно, парабола $2\beta x^2 + x\,(1 - 2\beta) + \beta - 1$ касается оси x, не заходя в нижнюю полуплоскость. Таким образом, неравенство (3) не выполнено и в точке излома левой части. Полученное противоречие завершает доказательство леммы.

ЛИТЕРАТУРА

1. **Arya V., Garg N., Khandekar R., Munagala K., Pandit V.** Local search heuristic for k-median and facility location problems // Proc. of the 33rd symposium on theory of computing (STOC). New York: ACM Press, 2001. P. 21—29.
2. **Chrobak M., Kenyon C., Noga J., Young N.** Online medians via online bribery // arXiv: cs.DS/0504103 (http://arxiv.org/abs/cs.DS/0504103), 2005.
3. **Jain K., Mahdian M., Saberi A.** A new greedy approach for facility location problems // Proc. of the 34th symposium on theory of computing (STOC). New York: ACM Press, 2002. P. 731-–740.
4. **Mettu R., Plaxton G.** The online median problem // SIAM J. Comput. 2003. V. 32. N 3. P. 816-–832.
5. **Mirchandani P., Francis R.** (eds.). Discrete location theory. New York: Wiley, 1990.

Адрес автора:

Институт математики
им. С. Л. Соболева СО РАН,
пр. Академика Коптюга, 4,
630090 Новосибирск,
Россия.
E-mail: shenmaier@mail.ru

Статья поступила
11 сентября 2006 г.

Переработанный вариант —
9 января 2007 г.