УДК 519.6

ОБ АППРОКСИМАЦИОННОЙ СЛОЖНОСТИ ЗАДАЧИ О МИНИМАЛЬНОМ КОМИТЕТЕ

Хачай М.Ю.

Институт математики и механики Уральского отделения Ран
ул. С.Ковалевской, 16, г.Екатеринбург, Россия, 620219

Abstract

The minimal Committee combinatorial problem (MC) is studied. It's know that this problem is NP-hard (see e.g. [1]). In this paper we show that unless $NP \not\subset TIME(m^{O(\log\log m)})$ there are no approximation algorithms for this problem with factor $(1-\varepsilon)\ln m$ for every $\varepsilon > 0$ even when all sets used in it's formulation are finite. To prove this estimate we reduce to it a SET COVER problem for which the same result is know [2].

Введение

В статье рассматривается одна комбинаторная задача — задача о Минимальном комитете (MC). Она тесно связана с тремя разделами исследования операций: теорией голосования, оптимизацией и распознаванием образов. В теории голосования исследуются коллективные процедуры принятия решений, базирующиеся на различных логиках подсчета голосов (демократиях). Комитет является математической моделью голосования на основе правила простого большинства.

В теории распознавания образов важное место занимают т.н. *персептронные алгоритмы*, восходящие ещё к работам Ф. Розенблатта. Легко убедиться, что 2-слойный персептрон с дополнительным условием неотрицательности весов нейрона второго слоя математически эквивалентен комитету.

Наконец, в теории оптимизации часто приходится иметь дело с т.н. *несобственными* задачами [3]. Как известно, существует несколько условий, при которых оптимизационная задача является несобственной. В терминах линейного программирования, например, задача является несобственной при условии несовместимости прямой, двойственной (или обеих) систем ограничений. Разработаны несколько методов коррекции несобственных задач, например, метод Чебышевской аппроксимации. Метод комитетных решений также активно используется для этих целей (см. обзор в [1]).

Во всех описанных выше случаях всегда желательно искать наиболее простой комитет, приводящий к задаче «Минимального Комитета».

Задача «Минимальный Комитет» (MC)

Пусть заданы множество X $D_1, D_2, \ldots, D_m$ его непустых подмножеств. Рассмотрим абстрактную систему включений

$$x \in D_j, \ (j \in N_m = \{1, 2, \ldots, m\}) \tag{1}$$

Система (1) не обязательно совместна, т.е. допустимо выполнение соотношения $ID_j = \oslash$. Как обычно, назовём *комитетным решением с q элементами системы* (1) *(или просто комитетом) конечную последовательность* $Q = (x^1, x^2, \ldots, x^q)$ *удовлетворяющую условию* $|\{i : x^j \in D_j\}| > \frac{q}{2}$ *при каждом* $j \in N_m$.

Задача «МИНИМАЛЬНЫЙ КОМИТЕТ»(MC). *Заданы множество X и набор его непустых подмножеств $D_1, D_2, \ldots, D_m$. Требуется найти комитетное решение системы* (1) *с наименьшим возможным q (или установить, что система не имеет комитетных решений).*

Ограничимся в этой работе рассмотрением частного случая задачи MC — задачей MCF, в рамках которой все множества $D_1, D_2, \ldots, D_m$ и X конечны. Эта задача имеет комбинаторную природу и трудно решаема (существование полиноминального алгоритма для неё влечёт равенство $P = NP$).

Теорема 1. Задача MCF NP — трудна

Удобно переформулировать задачу в терминах целочисленного линейного программирования. Пусть множества $J_1, J_2, \ldots, J_Q$ суть индексные множества всех максимальных по включению подсистем (м.с.п.) системы (1)*. Рассмотрим две $m \times Q$ матрицы инциденций: A и B, в которых*

$$a_{ji} = 1 \wedge b_{ji} = 1, \text{ \textit{если} }, j \in J_i,$$

$$a_{ji} = 1 \wedge b_{ji} = -1, \text{ \textit{в противном случае,}}$$

и, соответственно, две задачи ЦЛП:

$$\min\{(e,t)|Bt \geqslant f, t \in Z_+^Q\} \tag{2}$$

$$\min\left\{ \begin{array}{l} s : \alpha_s \leqslant 2s - 1, s \in N \\ \alpha_s = min\{(e,t)|At \geqslant sf, t \in Z_+^Q\} \end{array} \right\} \tag{3}$$

Здесь e и f — векторы из единиц, принадлежащие пространствам E_Q и E_m соответственно.

Теорема 2. Задача MCF, (2) и (3) полиномиально эквивалентны

Доказательство эквивалентности задач MCF и (2) *приведено в* [4]*, случай задач* (2) *и* (3) *очевиден.*

Следствие 1. Задачи (2) и (3) NP — трудны

Можно убедится, что описанные выше оптимизационные задачи остаются NP-трудными даже при дополнительном условии взаимной недоминируемости столбцов их матриц.

Аппроксимационный порог

Общим подходом, применяемым при исследовании NP-трудной оптимизационной задачи является разработка т.н. приближенных алгоритмов. Как обычно, приближенным алгоритмом (с точностью аппроксимации r) для задачи комбинаторной минимизации назовём алгоритм позволяющий для каждой её конкретной постановки

$$f^* = min\{f(x)|x \in M\}$$

за полиномиальное время находить допустимое решение x_{app} с условием

$$\frac{f(x_{app})}{f^*} \leqslant r.$$

Наряду с разработкой приближенных алгоритмов, как правило, рассматривается вопрос нахождения *аппроксимационного порога* задачи. Аппроксимационным порогом называется число, являющееся нижней оценкой точности аппроксимации произвольного приближённого алгоритма исследуемой задачи (при некотором здравом предположении, например, $NP \neq P$). В работе [2] найден такой порог для известной задачи о покрытии множества (Set Cover).

Задача «ПОКРЫТИЕ МНОЖЕСТВА»(Set Cover). *Заданы множество $S, |S| = n$ и непустое множество его подмножеств $C = \{c_1, c_2, \ldots, c_m\} \subseteq 2^S$. Требуется найти минимальное по мощности подмножество $C' \subseteq C$, покрывающее S (т.е. $\cup C' = S$).*

Теорема 3. [2] Для произвольноого $\varepsilon > 0$ существование приближенного алгоритма задачи SETCOVER с точностью аппроксимации$(1 - \varepsilon)\ln n$ влечет $NP \subset TIME(n^{O(\log\log n)})$.

Замечательно, что подобное утверждение справедливо и для задачи MCF (а, значит, и для задачи MC в целом).

Теорема 4. Если справедливо условие $NP \notin TIME(n^{O(\log\log m)})$, то для произвольного $\varepsilon > 0$ не существует приближенного алгоритма задачи MSF с точностью $(1 - \varepsilon)\ln m$.

Доказательство. 1. Сведём задачу SetCover к задаче MCF. Пусть зафиксированы множества $S = \{s_1, s_2, \ldots, s_n\}$ и $C = \{c_1, c_2, \ldots, c_m\}$. Нам достаточно за время ограниченное сверху полиномом от m до n указать подходящую формулировку задачи MCF, после чего показать, что для натурального k подмножество $C' \subset C, |C| = k$ является покрытием S тогда и только тогда, когда задача MCF имеет допустимое комитетное решение из $2k - 1$ элементов.

Рассмотрим $n \times m$ матрицу инциденций A множеств S и C. Как и ранее

$$a_{ji} = \begin{cases} 1, & \text{если } s_j \in c_i \\ 0, & \text{в противном случае.} \end{cases}$$

Сопоставим ей $(n+1)\times(m+1)$ матрицу A', полученную из A окаймлением её строкой и столбцом из единиц (см. табл. 1). Элемент полученной матрицы, находящийся в правом нижнем углу положим равным нулю.

		x^1	x^2	$\dots$	x^m	x^{m+1}
		c_1	c_2	$\dots$	c_m	S
D_1	s_1					1
D_2	s_2					1
$\vdots$	$\vdots$			A		$\vdots$
D_n	s_n					1
D_{n+1}	C	1	1	$\dots$	1	0

Таблица 1. Построение задачи MCF

Рассмотрим задачу MSF, соответствующую матрице A'. А именно, определим множество $X = \{x^1, x^2, \dots, x^{m+1}\}$, а подмножества, определяющие систему (1)

$$D_j = \{x^{m+1}\} \cup \{x^i : s_j \in c_i, i \in N_m\}(j \in N_n)$$

$$D_{n+1} = \{x^1, x^2, \dots, x^m\}.$$

Пусть $C' = \{c_i, c_{i_2}, \dots, c_{i_k}\}$ — покрытие множества S, т.е. для каждого $j \in N_n$ найдётся номер

$$i_p = i_p(j) : s_j \in c_{i_p}(j),$$

что влечёт $x^{i_p(j)} \in D_j$, по построению. Следовательно, последовательность

$$Q = (x^{i_1}, x^{i_2}, \dots, x^{i_k}, \underbrace{x^{m+1}, \dots, x^{m+1}}_{k-1})$$

является комитетом системы (1), поскольку каждое множество D_j содержит как минимум k элементов Q.

С другой стороны рассмотрим допустимое комитетное решение Q системы (1) из $2k-1$ элементов:

$$Q = (x^{i_1}, x^{i_2}, \dots, x^{i_{2k-1-p}}, \underbrace{x^{m+1}, \dots, x^{m+1}}_{p}).$$

По выбору множества D_{n+1} имеем $p < k$. Рассмотрим подпоследовательность $(x^{i_1}, x^{i_2}, \dots, x^{i_k})$ последовательности Q. Для каждого $j \in N_n$ существует

$$i_p = i_p(j) : x_{ip}(j) \in D_j,$$

т.к. Q является комитетом и, следовательно, $s_j \in c_{i_p}(j)$ по построению множеств D_j. Поэтому множество $C' = \{x^{i_p} : p \in N_k\}$ — искомое S-покрытие.

2. Допустим, нашелся приближённый алгоритм задачи MCF с точностью аппроксимации r, позволяющий найти допустимое комитетное решение системы (1) из $2k-1$ элементов и

$$1 \leqslant \frac{2k-1}{2s-1} \leqslant r,$$

где минимальный комитет содержит $2s-1$ элементов. Приведённая выше редукция влечёт, что этот алгоритм может быть использован в качестве приближённого алгоритма задачи $SetCover$, причём его аппроксимационная точность может быть оценена как

$$\frac{k}{s} = r\left(1 - \frac{1}{2s}\right) + \frac{1}{2s} \leqslant r\left(1 - \frac{1}{2s}\right)\frac{r}{2s} \leqslant r.$$

Следовательно, для задачи MSF справедливо утверждение теоремы. □

Работа поддержана РФФИ, гранты: НШ-792.2003.1, 01-01-00563 и 04-01-96104, а также Фондом содействия отечественной науке.

Список литературы

1. Mazurov Vl.D., Khachai M.Yu., Rybin A.I.: Commitee Construction for Solving Problems of Selection, Diagnostics and Prediction. Proceedings of the Steklov Institute of mathematics Suppl. 1, (2002), S67-S101.
2. Feige U.A Threshold of ln n for Approximating Set Cover. Journal of the ACM, 45(4).
3. Еремин И.И.. Теория линейной оптимизации. Екатеринбург: УрО Ран, 1998, 248 с.
4. Мазуров Вл.Д., Хачай М.Ю. Комитеты систем линейных неравенств. Автоматика и телемеханика, № 2, (2004), 43-54.