Математические
структуры и моделирование
2002, вып. 9, с. 1–4

УДК

ОЦЕНКА ПРИБЛИЖЕННОГО РЕШЕНИЯ АЛГОРИТМОМ ХВАТАЛА ЗАДАЧИ О ПОКРЫТИИ

А.В. Пролубников

In the paper we consider the Set Cover Problem. This problem is NP-hard[2]. So it becomes very important to find approximate solution of this problem in polinimial time. Algorithm developed by V. Chvatal is such algorithm and it is most effective algorithm developed for set cover problem. Followed statement true for this algorithm[1]:

$$\frac{c(Chv)}{c(Opt)} \leq \ln m + 1.$$

We have found another proof for this statement. Our proof is more simple and vivid by our opinion.

В задаче о покрытии даны: конечное множество $U, |U| = m$ и конечный набор множеств $P = \{P_1, \ldots, P_n\}$ таких, что

$$\forall j \ P_j \subseteq U, \ \bigcup_{j=1}^{n} P_j = U.$$

Будет рассматриваться также набор индексов $I = \{1, 2, \ldots, n\}$, соответствующий P.

Набор множеств $P' = \{P_{j_1}, \ldots, P_{j_p}\}$ $(P_{j_i} \in P)$ называется *покрытием* множества U, если $\bigcup\limits_{i=1}^{p} P_{j_i} = U$. Набору множеств P' соответствует набор индексов входящих в него множеств $I' = \{j_1, \ldots, j_p\}$.

Задан также набор положительных чисел $\{c_1, \ldots, c_n\}$ *стоимостей* множеств из P. Иначе говоря, задана аддитивная неотрицательная функция $c : P \to R_+$. Далее будем обозначать $c_j = c(P_j)$. *Стоимость* набора множеств $P' = \{P_{j_1}, \ldots, P_{j_p}\}$ равна

$$c(P') = \sum_{i=1}^{p} c(P_{j_i}) = \sum_{i=1}^{p} c_{j_i}.$$

Целью является отыскание покрытия минимальной стоимости.

Задача о покрытии является NP-трудной [2]. Следовательно, важным является поиск приближенных решений задачи за полиномиальное время. Широко

E-mail: prolubnikov@math.omsu.omskreg.ru
Омский государственный университет

известным алгоритмом отыскания приближенного решения задачи о покрытии является алгоритм Хватала, для которого верна [1] оценка

$$\frac{c(Chv)}{c(Opt)} \leq H(|A|) \leq H(m) \leq \ln m + 1, \tag{1}$$

где Chv – решение, которое дает алгоритм Хватала, Opt – оптимальное решение задачи, m – мощность покрываемого множества U, $|A|$ – максимальная мощность множеств из оптимального покрытия, $H(p) = \sum\limits_{k=1}^{p} \dfrac{1}{k}$.

Алгоритм В. Хватала является полиномиальным алгоритмом, который находит приближенное решение задачи о покрытии. В схеме алгоритма Хватала множества в покрытие (набор множеств Chv) выбираются соответственно минимуму отношения $\dfrac{c(P_j)}{|P_j|}$ – *удельной стоимости* множества следующим образом:

Алгоритм Хватала

Шаг 0. $Chv := \emptyset$;
Шаг 1. Если Chv – покрытие U, то работу завершить. Иначе – **Шаг 2.**
Шаг 2. Выбираем k так, что
$\dfrac{c_k}{|P_k|} = \min_j \Big(\dfrac{c_j}{|P_j|} : P_j \in P,\, P_j \not\subset \cup(P_l : \ l \in Chv) \Big)$;
$Chv := Chv \cup \{P_k\},\ \forall P_j : \ P_j := P_j \setminus P_k$. Перейти на **Шаг 1.**

Для решений, находимых *алгоритмом Хватала*, верна оценка:

$$\frac{c(Chv)}{c(Opt)} \leq H(|A|) \leq H(m), \tag{2}$$

где $|A|$ – максимальная мощность множеств из оптимального покрытия. Существует несколько вариантов доказательства этой оценки, например доказательство самого В. Хватала [1]. В доказательстве В. Хватала задача о покрытии рассматривается как задача целочисленного линейного программирования. Нами получено доказательство, где задача о покрытии рассматривается как комбинаторная задача. И хотя можно сказать, что получена более слабая оценка для решений, отыскиваемых алгоритмом Хватала, а именно оценка:

$$\frac{c(Chv)}{c(Opt)} \leq H(m),$$

полученное доказательство более наглядно демонстрирует мажорирование отношения $\dfrac{c(Chv)}{c(Opt)}$ гармоникой $H(m)$. Доказательство основано на ключевом соотношении (**лемма** 2):

$$\frac{c_i}{|P_i|} \leq \frac{c(Opt_i)}{m_i},$$

где P_i – множество, которое выбирается алгоритмом на i-й итерации; Opt_i – оптимальное покрытие части множества U, непокрытой к i-й итерации работы алгоритма, m_i – мощность этой части множества U.

Лемма 1. *Пусть* $a,\ b,\ c,\ d > 0$. *Тогда* $\frac{a}{b} \leq \frac{c}{d} \Rightarrow \frac{a}{b} \leq \frac{a+c}{b+d} \leq \frac{c}{d}$.

Доказательство. $\frac{a}{b} \leq \frac{c}{d} \Leftrightarrow ad \leq bc$;

$$\frac{a}{b} = \frac{a}{b} \cdot \frac{a+c}{a+c} \cdot \frac{b+d}{b+d} = \frac{a+c}{b+d} \cdot \frac{a(b+d)}{b(a+c)} = \frac{a+c}{b+d} \cdot \frac{ab+ad}{ab+bc} \leq \frac{a+c}{b+d} \cdot \frac{ab+bc}{ab+bc} = \frac{a+c}{b+d};$$

$$\frac{c}{d} = \frac{c}{d} \cdot \frac{a+c}{a+c} \cdot \frac{b+d}{b+d} = \frac{a+c}{b+d} \cdot \frac{c(b+d)}{d(a+c)} = \frac{a+c}{b+d} \cdot \frac{bc+cd}{ad+cd} \geq \frac{a+c}{b+d} \cdot \frac{ad+cd}{ad+cd} = \frac{a+c}{b+d}.$$

■

Следствие 1. $\frac{a_1}{b_1} \leq \frac{a_2}{b_2} \leq \ldots \leq \frac{a_k}{b_k} \Rightarrow \frac{a_1}{b_1} \leq \frac{\sum\limits_{i=1}^{k} a_i}{\sum\limits_{i=1}^{k} b_i}$. ■

Пусть $(P_1, P_2, \ldots,\ P_l)$ – покрытие, которое дает алгоритм Хватала, и алгоритм Хватала выбирает множество P_i на i-й итерации, $(A_1, A_2, \ldots,\ A_k)$ – оптимальное покрытие, a_i – стоимость множества A_i.

Лемма 2. $\frac{c_1}{|P_1|} \leq \frac{c(Opt)}{m} = \frac{\sum\limits_{i=1}^{k} c(A_i)}{m}$.

Доказательство. Положим $A'_1 = A_1,\ A'_2 = A_2 \setminus A_1,\ \ldots,\ A'_k = A_k \setminus \left(\bigcup\limits_{i=1}^{k-1} A_i\right) \Rightarrow A'_i$ – попарно непересекающиеся. Упорядочим множества A_i и введем новую индексацию в соответствии с неубыванием отношения $\frac{a_i}{|A'_i|}$:

$$\frac{a_1}{|A'_1|} \leq \frac{a_2}{|A'_2|} \leq \ldots \leq \frac{a_k}{|A'_k|} \Rightarrow \frac{a_1}{|A'_1|} \leq \frac{\sum\limits_{i=1}^{k} a_i}{\sum\limits_{i=1}^{k} |A'_i|} = \frac{c(Opt)}{m}$$

(по **лемме 1**). По алгоритму $\frac{c_1}{|P_1|} \leq \frac{a_1}{|A_1|} = \frac{a_1}{|A'_1|} \leq \frac{c(Opt)}{m}$. ■

Теорема 1.

$$\frac{c(Chv)}{c(Opt)} \leq H(m).$$

Доказательство. Пусть на i-й итерации алгоритм Хватала выбирает множество P_i^i, где P_i^i – множество P_i к i-й итерации: $P_i^i = P_i \setminus (\bigcup\limits_{k=1}^{i-1} P_k)$; обозначим $p_i = |P_i^i| \Rightarrow \sum\limits_{i=1}^{l} p_i = m$. Обозначим Opt_i – оптимальное покрытие для i-й задачи (т.е. задачи на i-й итерации). По **лемме 2**

$$\frac{c_1}{p_1} \leq \frac{c(Opt)}{m},\ \frac{c_2}{p_2} \leq \frac{c(Opt_2)}{m-p_1},\ \ldots, \frac{c_l}{p_l} \leq \frac{c(Opt_l)}{m-p_1-p_2-\ldots-p_{l-1}}.$$

Так как $c(Opt) \geq c(Opt_2) \geq \ldots \geq c(Opt_l)$, то

$$c(Chv) = \sum_{i=1}^{l} c_i \leq \left(\frac{p_1}{m} + \frac{p_2}{m-p_1} + \ldots + \frac{p_l}{m-p_1-p_2-\ldots-p_{l-1}}\right) \cdot c(Opt) \leq$$
$$\leq \left[\left(\frac{1}{m} + \frac{1}{m-1} + \ldots + \frac{1}{m-p_1+1}\right) + \left(\frac{1}{m-p_1} + \frac{1}{m-p_1-1} + \ldots\right.\right.$$
$$\left.\ldots + \frac{1}{m-p_1-p_2+1}\right) + \ldots + \left(\frac{1}{m-p_1-\ldots-p_{l-1}} + \frac{1}{m-p_1-\ldots-p_{l-1}-1} +\right.$$
$$\left.\left.+ \ldots + \frac{1}{m-p_1-p_2-\ldots-p_l+1}\right)\right] \cdot c(Opt) = \left[\frac{1}{m} + \frac{1}{m-1} + \ldots + \frac{1}{2} +\right.$$
$$\left. +1\right] \cdot c(Opt) = H(m) \cdot c(Opt). \qquad \blacksquare$$

Литература

1. Chvatal V. *A greedy heuristic for the set-covering problem* // Mathematics of operation research. 1979. V.4, N 3. P.233–235.
2. Гэри М. *Джонсон Д. Вычислительные машины и труднорешаемые задачи.* М.: Мир, 1982.