第 29 卷第 6 期　温州大学学报·自然科学版　2008 年 12 月
Vol 29, No 6　Journal of Wenzhou University · Natural Sciences　Dec, 2008

一种求解带权集合覆盖问题的近似算法

张馨东，罗　亮

（兰州交通大学，甘肃兰州　730070）

摘　要：以优化形式描述的集合覆盖问题是一个 NP 难问题，设计快速有效的近似算法，具有重要的理论与现实意义．基于贪心算法思想，提出了一种求解带权集合覆盖问题的近似算法，并讨论了该算法的相对近似比．

关键词：带权集合覆盖；贪心算法；相对近似比

中图分类号：O242.2　　**文献标识码**：A　　**文章编号**：1006-0375(2008)06-0046-03

工业领域里的许多实际问题都可建模为集合覆盖问题，如资源选择问题、设备选址问题等．带权问题的集合覆盖问题可描述为：给定一个包含 n 个元素的集合 X ，F 包含 X 的 m 个子集，即 $F=\{S_1,\cdots,S_i,\cdots,S_m\}$ ，$S_i\,(1\le i\le m)$ 为 X 的子集．对每个子集 S_i ，定义一个权值 $w(S_i>0)$ ，带权的集合覆盖问题的目标是找出一个 F 的子集 C ，使得 C 覆盖整个 X ，并使其权值最小．即

$$X=\bigcup_{S_i\in C}S_i$$

$$\min\sum_{S_i\in C}w(S_i)$$

集合覆盖问题是一个NP完全问题[1]，还没有多项式时间精确算法，实际应用中往往不要求一定得到精确解，因此可以考虑设计出近似多项式时间算法，使该解的目标函数值在某种性能比的定义下，保证是近似最优的[2]．对于很多问题，利用贪心算法的思想[3]可以得到性能较好的近似算法．文献[4]提出了一种随机近似算法．本文借鉴了按密度（价值/重量）贪心的策略，提出一种近似算法．

1 算　法

对于这个问题考虑如下算法，每次选取集合 $S\in F$ ，其权值为 w ，使得 $\dfrac{w}{|S\cap X|}$ 最小，然后将 $S\cap X$ 中的元素从 X 中去掉，并选取下一个集合，直到 X 被完全覆盖．算法设计为：

GreedyPowSetCover(X, F)

（1）$U\leftarrow X$

（2）$C\leftarrow\Phi$

（3）while $U\neq\Phi$ do

收稿日期：2008-05-06

作者简介：张馨东（1984- ），男，甘肃陇西人，硕士研究生，研究方向：算法设计与复杂性

（4）select an $S \in F$ that minimizes $\dfrac{w}{|S \cap U|}$

（5）$U \leftarrow U - S$

（6）$C \leftarrow C \cup \{S\}$

（7）return C

2 相对近似比的分析

对于以上设计的算法有下面的定理成立.

定理 解带权集合覆盖问题的 GreedyPowSetCover（X, F）算法是多项式时间的$(\ln|X|+1)$-近似算法.

证明：假设C是由算法 GreedyPowSetCover(X, F)得到的集合，C^*是最优解集合．设S_i表示第i个被算法选中的集合，权值为w_i，将之分配给该集合中第一次被选中的元素．以w_x表示$x \in X$分配到的权值，在x第一次被覆盖时，权值为：

$$w_x = \frac{w_i}{|S_i - (S_1 \cup S_2 \cup \cdots \cup S_{i-1})|} \quad ,$$

则$W_C = \sum_{x \in X} w_x$．其中W_C表示集合C覆盖X后得到的权值，亦即近似最优解$A(I)$．对最优解集C^*中的集合，依照算法 GreedyPowSetCover(X, F)的取法依次取入，于是有

$$\sum_{S \in C^*} \sum_{x \in S} w_x \geq \sum_{x \in X} \sum_{x \in X} w_x = W_C \quad \text{，即}$$

$$W_C \leq \sum_{S \in C^*} \sum_{x \in S} w_x \quad . \tag{1}$$

下面估计$\sum_{x \in S} w_x$的大小.

对$S \in F$，令$u_i = |S - (S_1 \cup S_2 \cup \cdots \cup S_i)|$表示算法 GreedyPowSetCover($X, F$)选择$S_1, S_2, \cdots, S_i$后集合$S$中未被覆盖的元素个数，令$k$是使$u_k = 0$的$u_k$的第一个下标，使得$S$中每一个元素被$S_1, S_2, \cdots, S_i$其中的一个覆盖，可知$u_{i-1} \geq u_i$，且$S$中元素$u_{i-1} - u_i$是第一次为$S_i$所覆盖．因而

$$\sum_{x \in S} w_x = \sum_{i=1}^{k} (u_{i-1} - u_i) \frac{w_i}{|S_i - (S_1 \cup S_2 \cup \cdots \cup S_{i-1})|}$$

$$\leq \sum_{i=1}^{k} (u_{i-1} - u_i) \frac{w_S}{|S - (S_1 \cup S_2 \cup \cdots \cup S_{i-1})|} \leq \sum_{i=1}^{k} \sum_{j=u_i+1}^{u_i-1} \frac{w_S}{j} = \sum_{i=1}^{k} (\sum_{j=1}^{u_{i-1}} \frac{w_S}{j} - \sum_{j=1}^{u_i} \frac{w_S}{j}) .$$

令$H(d) = \sum_{i=1}^{d} \frac{1}{i}$，$H(0) = 0$，则

$$\sum_{x \in S} w_x \leq w_S \sum_{i=1}^{k} (H(u_{i-1}) - H(u_i)) = w_S H(u_0) = w_S H(|S|)$$

$$\le w_S H(\max\{|S|: S \in F\}) . \qquad (2)$$

由式（1）与式（2）可得

$$A(I) = W_C \le \sum_{S\in C^*}\sum_{x\in S} w_x \le \sum_{S\in C^*} w_S H(\max\{|S|: S\in F\})$$

$$= H(\max\{|S|: S \in F\})\sum_{S\in C^*} w_S = H(\max\{|S|: S\in F\})OPT(I) .$$

又对调和级数 $H(n)$，有

$$\ln(n+1) \le H(n) = \sum_{k=1}^{n}\frac{1}{k} \le \ln n + 1 .$$

对 S，$|S| \le |X|$．故

$$A(I) \le H(|X|)OPT(I) \le (\ln|X|+1)OPT(I) .$$

命题成立.

3. 结　语

在一些应用中，$\max\{|S|: S\in F\}$ 是一个比较小的常数，算法 GreedyPowSetCover(X，F)可返回最多是最优解值 $OPT(I)$ 常数倍的解值．本文通过讨论算法的相对近似比，说明 GreedyPowSetCover(X，F)是 $(\ln|X|+1)$-近似多项式时间算法，并能返回较为满意的近似解.

参考文献

[1] Karp R M. Reducibility among Combinatorial Problems [C] // Karp R M. Complexity of Computer Computations. New York: Plenum press, 1972: 85-103.

[2] 张德富. 算法设计与分析: 高级教程[M]. 北京: 国防工业出版社, 2007: 94-101.

[3] 王晓东. 算法设计与分析[M]. 第 2 版. 北京: 电子工业出版社, 2005: 303-314.

[4] 姚国辉, 朱大铭, 马绍汉, 等. 带权集合覆盖问题的一种随机近似算法[J]. 吉林大学学报: 工学版, 2007, 3(2): 429-432.

Approximation Algorithm for Weighted Set Cover Problem

ZHANG Xindong, LUO Liang

(School of Mathematics, Physics and Software Engineering, Lanzhou Jiaotong University, Lanzhou, China 730070)

Abstract: The set cover problem described with maximal form is an NP complex one. The authors in this paper proposed an approximative algorithm for weighted set cover problem, and then discussed the relative approximative ratio of the algorithm.

Key words: Weighted set cover problem; Greedy algorithm; Relative approximative ratio

（编辑：王一芳）