维普资讯 http://www.cqvip.com

集合多覆盖问题的乘性权重更新分析

崔 鹏[1] 钱丽艳[2]
（中国人民大学信息资源管理学院 北京 100872）[1] （北京大学信息学院 北京 100871）[2]

摘 要 集合多覆盖问题的简单贪心算法的近似比是 $\ln n+1$。本文提出简单贪心算法的一个变形，宽度优先贪心算法，并且证明其有近似比 $(\ln n)/r+\ln\ln n+O(1)$，其中 r 是覆盖要求。这个结果比由随机取整方法得到的近似比 $O((\ln n)/r+\sqrt{(\ln n)/r})+1$ 为优。宽度优先贪心算法的设计可以归入 Arora 等最近提出的乘性权重更新方法的框架。

关键词 集合多覆盖，宽度优先贪心算法，乘性权重更新方法

Multiplicative Weights Update Analysis of Set Multicover

CUI Peng[1] QIAN Li-Yan[2]
(School of Information Resource Management, Renmin University of China, Beijing 100872)[1]
(School of EECS, Peking University, Beijing 100871)[2]

Abstract The simple greedy algorithm for set multicover has approximation ratio . This paper presents a variation of the simple greedy algorithm, Breadth First Greedy Algorithm, and proves its approximation ratio can be $(\ln n)/r+\ln\ln n+O(1)$, where r is the cover requirement. This result is better than the approximation ratio $O((\ln n)/r+\sqrt{(\ln n)/r})+1$ obtained by randomized rounding method. The design of this algorithm fits in the framework of multiplicative weights update method, proposed by Arora etc. recently.

Keywords Set multicover, Breadth first greedy algorithm, Multiplicative weights update method

1 引言

集合覆盖问题是一个得到大量关注的 NP 难问题，其输入是集合 U 和一个 U 的子集的集合 C，$|U|=n$。如果 C 的子集 C' 满足，U 的每个元素属于 C' 的至少一个成员，则称 C' 是 U 的一个集合覆盖。问题的目标是找最小的集合覆盖。

集合多覆盖问题（加权情况）是集合覆盖问题的推广情况。输入包含 U，C，覆盖要求 $r\in Z^+$，和权函数 $w: C\rightarrow Z^+$。一个多重集 $C'\subseteq C\cdot r$ 是 U 的一个 r-集合覆盖，如果对每个 $e\in U$，$\in(e,C')=r$，这里 $\in(e,C')$ 是 C' 中包含 e 的子集的个数。对每个多重集 $\bar{C}\subseteq C\cdot r$，令 $\bar{C}$ 的总权为

$w(\bar{C})=\sum_{c\cdot r_c\in\bar{C}}(r_c w(c))$，

这里 r_c 是 c 在 $\bar{C}$ 内的重数。问题的目标是找总权最小的 r-集合覆盖。

定义 $\bar{C}$ 的覆盖度量为

$\#(\bar{C})=\sum_{e\in U}\max(r-\in(e,\bar{C}),0)$，

定义 C 相对于 $\bar{C}$ 的覆盖度量为

$\#(c,\bar{C})=\#(\bar{C})-\#(\bar{C}\cup\{c\})$。

集合多覆盖问题的简单贪心算法（Simple Greedy Algorithm，简称 SimG）可以描述为

SimG
1. $\bar{C}\leftarrow\emptyset$；
2. 在 $C-\bar{C}$ 中选择 c 使得 $\#^M(c,\bar{C})/w(c)$ 最大；
3. $\bar{C}\leftarrow\bar{C}\cup\{c\}$；
4. 如 $\bar{C}$ 是 r-集合覆盖，返回 $\bar{C}$，否则转 2。

其中，称 $\bar{C}$ 为部分 r-集合覆盖。如果一个元素被 $\bar{C}$ 中少于 r 个子集覆盖，则称这个元素是活的。SimG 的近似比是 $\ln n+1$[1]。

2 MWU 方法

不同领域的一些算法都有这样的特点：在一个特定集合上维护一个分布，在每一步元素上的概率被乘或除这个元素的某种“回报”因子。算法运行时间的分析依赖于一个位势函数。

作为一个例子，Young 给出了集合覆盖问题的“健忘取整”技术[2]。他观察到未被覆盖的元素个数恰好是一个位势函数，近似算法在每步只需要缩小这个位势函数。这样他得到众所周知的近似比 $\ln n+1$ 的另一个证明。

最近，Arora 等提出了一个简洁的元算法来统一这些表面上不同的算法，把它们变成元算法的实例[3]。这个元算法借鉴了 Young 的思想，是机器学习理论里的加权多数算法的随机化和一般化[4]。他们称这个元算法为乘性权重更新方法（Multiplicative Weights Update Method，简称 MWU），并建议视其为与分治方法、动态规划和随机采样等类似的基本工具。

MWU 可以用“专家-事件”语言来描述。开始时每个专家有一个权重。算法每一步，按照专家间权重的比例决定各个专家的概率。专家们按照这个概率做预测，即决定发生那个事件。事件发生后，如专家的约束被满足，他就受到“惩罚”，即权重降低一个因子。为了尽快减低权重，算法寻找极大的不利事件，使得惩罚最大。

本文作者为集合多覆盖问题设计了带乘性因子 $M>1$ 的宽度优先贪心算法（Breadth First Greedy Algorithm，简称 BFG^M），并给出了改进的近似比。BFG^M 可以归入 MWU 的框架，而且它比集合覆盖更适合于说明乘性更新规则（权重的乘性因子恰好是 M）。

崔 鹏 博士，研究方向为计算机图形学。钱丽艳 工程师，研究方向为计算机应用。

维普资讯 http://www.cqvip.com

定义 $\bar{C}$ 的带乘性因子 M 的覆盖度量为

$$\#^M(\bar{C})=\sum_{e\in U}\sum_{p=0}^{r-1-\in(e,\bar{C})}M^p$$

定义 c 相对于 $\bar{C}$ 的带乘性因子 M 的覆盖度量为

$$\#^M(c,\bar{C})=\#^M(\bar{C})-\#^M(\bar{C}\cup\{c\})$$

BFG^M 可以描述为

BFG^M
1. $\bar{C}\leftarrow\emptyset$;
2. 在 $C-\bar{C}$ 中选择 c 使得 $\#^M(c,\bar{C})/w(c)$ 最大;
3. $\bar{C}\leftarrow\bar{C}\cup\{c\}$;
4. 如 $\bar{C}$ 是 r-集合覆盖,返回 $\bar{C}$,否则转 2。

其中,覆盖 e 对覆盖度量 $\#^M(c,\bar{C})$ 的贡献是 $M^{r-1-\in(e,\bar{C})}$。e 被覆盖的次数越少,覆盖 e 的贡献越大;e 被覆盖的次数越多,覆盖 e 的贡献越小。就是说,算法倾向于选择覆盖子集以覆盖尽可能多的被覆盖次数少的元素。

3 主要结果

在本节中,我们证明 BFG^M 的近似比可以为

$$(1+1/(M-1))((\ln n)/r+\ln M)+1。$$

令 $M=\ln n+1$,这个近似比即

$$(\ln n)/r+\ln\ln n+O(1)。$$

我们注意到这个近似比是综述[5]没有列出的新的结果,且比由随机取整方法得到的近似比 $O((\ln n)/r+\sqrt{(\ln n)/r})+1$ 为优[6]。

首先我们不加证明地给出两个事实。

事实 1 对任意 $0<x<1$,有 $(1-x)^{1/x}<1/e$。

事实 2 对任意 $0<x<1$ 和 $q\geqslant1$,有 $1-qx\leqslant(1-x)^q$。

然后我们给出一个引理。

引理 设 C^* 是一个最小集合覆盖,在部分 r 集合覆盖是 $\bar{C}$ 时算法选择 $\bar{c}$,则

(a) $\#(\bar{c},\bar{C})/w(\bar{c})\geqslant\sum_{c\in C^*}\#(c,\bar{C})/w(C^*)$。

(b) $w(\bar{c})\leqslant w(C^*)$。

证明:

(a)由算法的定义,对任意 $c\in C^*$,有

$$\#(\bar{c},\bar{C})/w(\bar{c})\geqslant\#(c,\bar{C})/w(c)$$

所以

$$\#(\bar{c},\bar{C})/w(\bar{c})\geqslant\sum_{c\in C^*}\#(c,\bar{C})/w(C^*)$$

(b) 显然,$\#(\bar{c},\bar{C})\leqslant\#(\bar{C})$,且 $\#(\bar{C})\leqslant\sum_{c\in C}\#(c,\bar{C})$。再由(a)得

$$\begin{aligned}w(\bar{c})&\leqslant\#(\bar{c},\bar{C})w(C^*)/\sum_{c\in C^*}\#(c,\bar{C})\\&\leqslant\#(\bar{C})w(C^*)/\sum_{c\in C^*}\#(c,\bar{C})\\&\leqslant w(C^*)\end{aligned}$$

□

定理 BFG^M 有近似比

$$(1+1/(M-1))((\ln n)/r+\ln M)+1。$$

证明:设 C^* 是一个最优集合覆盖,

$$\#_0=n\sum_{i=0}^{r-1}M^i,$$

及

$$k=w(C^*)\frac{M}{M-1}\frac{1}{r}\ln(n\sum_{i=0}^{r-1}M^i)$$

给定 $\bar{C}$,定义位势函数

$$f(\bar{C})=\#(\bar{C})(1-\frac{M-1}{M}\frac{r}{w(C^*)})^{k-w(\bar{C})}$$

由事实 1,

$$f(\emptyset)\leqslant\#_0(1-\frac{M-1}{M}\frac{r}{w(C^*)})^k<1$$

设接下来算法选择 $\bar{c}$,则由引理(a),

$$\begin{aligned}\#(\bar{c},\bar{C})&\geqslant\frac{w(\bar{c})}{w(C^*)}\sum_{c\in C^*}\#(c,\bar{C})\\&=\frac{w(\bar{c})}{w(C^*)}\sum_{\text{alive }e\in U}\sum_{c:c\in C,e\in c}M^{r-1-\in(e,\bar{C})}\\&=\frac{w(\bar{c})}{w(C^*)}(\sum_{e:e\in U,\in(e,\bar{C})<r-1}(M^{r-1-\in(e,\bar{C})}\in(e,C^*))\\&\quad+\sum_{e:e\in U,\in(e,\bar{C})=r-1}\in(e,C^*))\end{aligned}$$

注意到

$$\begin{aligned}\#(\bar{C})&=\sum_{\text{alive }e:e\in U}\sum_{i=0}^{r-1-\in(e,\bar{C})}M^i\\&=\sum_{e:e\in U,\in(e,\bar{C})<r-1}\sum_{i=0}^{r-1-\in(e,\bar{C})}M^i+\sum_{e:e\in U,\in(e,\bar{C})=r-1}1\end{aligned}$$

我们有

$$\begin{aligned}\#(\bar{C}\cup\{\bar{c}\})&=\#(\bar{C})-\#(\bar{c},\bar{C})\\&\leqslant\sum_{e:e\in U,\in(e,\bar{C})<r-1}(\sum_{i=0}^{r-1-\in(e,\bar{C})}M^i-\frac{w(\bar{c})}{w(C_0^*)}M^{r-1-\in(e,\bar{C})}\in(e,C^*))+\sum_{e:e\in U,\in(e,\bar{C})=r-1}(1-\frac{w(\bar{c})}{w(C^*)}\in(e,C^*))\\&\leqslant\sum_{e:e\in U,\in(e,\bar{C})<r-1}\sum_{i=0}^{r-1-\in(e,\bar{C})}M^i(1-\frac{M-1}{M}\frac{w(\bar{c})}{w(C^*)}\in(e,C^*))+\sum_{e:e\in U,\in(e,\bar{C})=r-1}(1-\frac{M-1}{M}\frac{w(\bar{c})}{w(C^*)}\in(e,C^*))\end{aligned}$$

由 C^* 的定义,对任意 $e\in U,\in(e,C^*)\geqslant r$。再由事实 2,

$$\begin{aligned}\#(\bar{C}\cup\{\bar{c}\})&\leqslant\#(\bar{C})(1-\frac{M-1}{M}\frac{r\cdot w(\bar{c})}{w(C^*)})\\&\leqslant\#(\bar{C})(1-\frac{M-1}{M}\frac{r}{w(C^*)})^{w(\bar{c})}\end{aligned}$$

所以

$$\begin{aligned}f(\bar{C}\cup\{\bar{c}\})&=\#(\bar{C}\cup\{\bar{c}\})(1-\frac{M-1}{M}\frac{r}{w(C^*)})^{k-w(\bar{C}\cup\{\bar{c}\})}\\&\leqslant\#(\bar{C})(1-\frac{M-1}{M}\frac{r}{w(C^*)})^{k-w(\bar{C}\cup\{\bar{c}\})+w(\bar{c})}\\&=\#(\bar{C})(1-\frac{M-1}{M}\frac{r}{w(C^*)})^{k-w(\bar{C})}\\&=f(\bar{C})\end{aligned}$$

设算法返回的 r-集合覆盖是 C',c' 是算法最后一步选择的子集,则 $C'-\{c'\}$ 是一个部分 r-集合覆盖,且 $f(C'-\{c'\})\leqslant f(\emptyset)<1$。

另一方面,

$$\begin{aligned}f(C'-\{c'\})&=\#(C'-\{c'\})(1-\frac{M-1}{M}\frac{r}{w(C^*)})^{k-w(C'-\{c'\})}\\&\geqslant(1-\frac{M-1}{M}\frac{r}{w(C^*)})^{k-w(C'-\{c'\})}\end{aligned}$$

所以

$$(1-\frac{M-1}{M}\frac{r}{w(C^*)})^{k-w(C'-\{c'\})}<1,$$

这样 $w(C'-\{c'\})<k$。

由引理(b),

$$\begin{aligned}w(C')&=w(C'-\{c'\})+w(c')\\&<w(C^*)\frac{M}{M-1}\frac{1}{r}\ln(n\sum_{i=0}^{r-1}M^i)+w(C^*)\\&\leqslant w(C^*)((1+1/(M-1))((\ln n)/r+\ln M)+1)\end{aligned}$$

□

结束语 本文讨论的集合多覆盖问题可以进一步推广到

(下转第 237 页)

维普资讯 http://www.cqvip.com

ings of the SIGCHI Conference on Human Factors in Computing Systems, 1997. 359～366
7 http://www.aic.nrl.navy.mil/～severett/vr.html
8 Badler N I, Webber B, Kalita J, et al. Animation from instructions[J]. In: Badler N I, Barsky B A, Zeltzer D, eds. Make Them Move, SanMateo, CA: Morgan-Kaufmann, 1991. 51～93
9 Vilhjalmsson H, Cassell J. BodyChat: autonomous communicative behaviors in avatars[C]. In: Proc. of the 2nd Annual ACM International Conference on Autonomous Agents, Minneapolis, USA, 1998
10 Cassell J, Bickmore T, Campbell L, et al. Requirements for an architecture for embodied conversational characters[J]. In: Computer Animation and Simulation'99, Vienna, Austria: Springer Verlag, 1999
11 Cassell J. A framework for gesture generation and interpretation [A]. In: Cipolla R, Pentland A, eds. Computer Vision in Human Machine Interaction, Cambridge University Press, 2000
12 Binsted K, Luke S. Character design for soccer commentary[A]. In: Asada M, Kitano H, eds. RoboCup-98: Robot Soccer World Cup II, vol 1604/1999. 22
13 Reynolds C W. Computer animation with scripts and actors[C]. In: Proceedings of the 9th Annual Conference on Computer Graphics and Interactive Techniques, 1982. 289～296
14 Perlin K, Goldberg A. Improv: A system for scripting interactive actors in virtual worlds[J]. Computer Graphics, 1996, 29(3): 205～216
15 Sgouros N M. Dynamic generation, management and resolution of interactive plots[J]. Artificial Intelligence, 1999, 107:29～62
16 Mohamed A S, et al. Toward a computational theory for motion understanding: the expert animators model[C]. In: Proc. of Fourth Conference of Artificial Intelligence for Space Applications, 1988. 289～302
17 Rilpkema H, et al. Computer animation of knowledge-based Human GraAPIng[J]. ACM Computer Graphics, 1991, 25(4): 12～33
18 Haskins J. The role of expert systems in creating interactive video animation[C]. In: Proc. of moving Towards Experts Systems Globally in the 21st Century, 1994. 1087～1088
19 Mah S, et al. NSAIL PLAN: an experience with constraint based reasoning in planning and animation[C]. In: Proc. of Computer Animation'94, 1994. 83～92
20 劳志强,潘云鹤. 基于影像动画设计的知识表达模型[J]. 计算辅助设计与图形学学报,1998,10(4):367～376
21 Funge J D. 人工智能在计算机游戏和动画中的应用:认知建模方法[M]. 班晓娟,艾迪明译. 清华大学出版社,2004
22 Funge J, Tu Xiaoyuan, Terzopoulos D. Cognitive modeling: knowledge, reasoning and planning for intelligent characters[C]. In: SIGGRAPH 1999, Los Angeles, CA, USA, 1999
23 Dinerstein J, Egbert P K, Garis H, et al. Fast and learnable behavioral and cognitive modeling for virtual character animation[J]. Computer Animation and Virtual Worlds, 2004, 15(2): 95～108
24 Badler N I. Artificial intelligence, natural language and simulation for human animation[C]. In: State-of-the-Art in Computer Animation. Proceedings of Computer Animation '89, 1989. 19～31
25 Lglesias A, Luengo F. A new based-on- artificial-intelligence framework for behavioral animation of virtual actors[C]. In: Proceedings of the International Conference on Computer Graphics, Imaging and Visualization (CGIV'04), 2004. 45～250
26 Wilhelms J, Skinner R. Motion control[J]. IEEE Computer Graphics and Applications, 1990, 10(3): 14～22
27 Flocks R C. Herds and schools: a distributed behavioral model [J]. ACM Computer Graphics, SIGGRAPH, 1987, 21(4): 25～34
28 Anderson N, McDaniel E, Chenney S. Constrained animation of flocks[C]. In: Proceedings of the 2003 ACM SIGGRAPH/Eurographics Symposium on Computer Animation, 2003
29 Sims K. Evolving virtual creatures[A]. In: Proceedings of SIGGRAPH-1994, Computer Graphics, Annual Conference Series, New York: ACM Press, 1994. 15～22
30 Sommerer C, Mignonneau L. Art as a living system: interactive computer artworks[J]. Leonardo, 1999, 32(3): 165～173
31 Sommerer C, Mignonneau L. GENMA-genetic manipulator[C]. In: Ars Electronica' 96: Memesis The Future of Evolution. Wien/New York: Springer Verlag, 1996. 294～295
32 Maes P. Artificial life meets entertainment: interacting with lifelike autonomous agents[C]. Special Issue on New Horizons of Commercial and Industrial AI, 1995, 38(11): 108～114
33 Blumberg B, Galyean A. Multi-level direction of autonomous creatures for real-time virtual environments[C]. In: Proceedings of SIGGRAPH95, 1995. 47～54
34 Terzopoulos D, Tu Xiaoyuan, Grzeszczuk R. Artificial fishes: autonomous locomotion, perception, behavior and learning in a simulated physical world[J]. Artificial Life, 1994, 1(4): 327～351
35 Tu X, Terzopoulos D. Artificial fishes: physics, locomotion, perception and behavior[C]. In: Proceedings of SIGGRAPH 94, 1994. 24～29
36 陈泓娟,班晓娟,涂序彦,等. 计算机动画的人工生命方法研究——人工鱼的自繁衍模型[J]. 自动化学报,2003,29(6):986～989
37 Ban Xiaojuan, Ai Dongmei, Zeng Guangping, et al. Computer animation based on artificial life and artificial intelligence: the research of artificial fish[C]. In: 2005 IEEE International Conference on Networking, Sensing and Control, 2005. 321～325
38 涂序彦. 广义人工生命及其应用[J]. 见:中国人工智能学会第一届"人工生命及应用"专题学术会议论文集《人工生命及应用》. 北京邮电大学出版社,2004. 1～13
39 Bedau M A. Philosophical aspects of artificial life[A]. Varela, Cambridge, MA: Radford Books/MIT Press, 1992
40 David K, David R, Pavel N, et al. Intelligent virtual agents. In: Lecture notes in computer science, 2003
41 Dinerstein J. Improving and extending behavioral animation through machine learning[D]. In: Brigham Young University, 2005

(上接第 220 页)

覆盖整数规划问题(Covering Integer Programming,简称CIP),不难看出能够把本文的方法应用到这种一般情况,并得到新的结果。

另外,如果限制C中集合至多被选择一次,就得到约束集合多覆盖问题。这个情况目前最好的近似比是 $\ln n+1$[1]。对简单贪心算法的下界分析表明这个近似比很难改进[7]。如何把本文的设计思路应用到这个情况将是一个带有挑战性的课题。

参 考 文 献

1 Rajagopalan S, Vazirani V V. Primal-Dual RNC approximation algorithms for set cover and covering integer programs. SIAM Journal on Computing, 1999, 28:525～540
2 Young N E. Randomized rounding without solving the linear program. Sixth ACM-SIAM Symposium on Discrete Algorithms (SODA 95), 1995. 170～178
3 Arora S, Hazan E, Kale S. Multiplicative weights method: a meta-algorithm and its applications, Online Survey. http://www.cs.princeton.edu/～arora/pubs/MWsurvey.pdf, 2006
4 Littlestone N, Warmuth M K. The weighted majority algorithm. Information and Computation, 1994, 108:212～261
5 Kolliopoulus S G, Young N E. Approximation algorithms for covering/packing integer programs. Journal of Computer and System Sciences, 2005, 71:495～505
6 Srinivasan A. Improved approximations guarantees for packing and covering integer programs. SIAM Journal on Computing, 1999, 29:648～670
7 Berman P, DasGupta B, Sontag E. Randomized approximation algorithms for set multicover problems with applications to reverse engineering of protein and gene networks. In: 7th International Workshop on Approximation Algorithms for Combinatorial Optimization Problems (APPROX 2004), LNCS 3122, Springer-Verlag, Berlin, 2004. 39～50