CN 43-1258/TP
ISSN 1007-130X

计算机工程与科学
Computer Engineering & Science

第 35 卷第 4 期 2013 年 4 月
Vol. 35, No. 4, Apr. 2013

文章编号:1007-130X(2013)04-0111-04

限制设施选址问题的近似算法*

刘玉堂,方奇志
(中国海洋大学数学科学学院,山东 青岛 266100)

摘　要:提出了设施选址问题的一个新变体—限制设施选址问题,给出了一个基于随机线性规划舍入的近似算法,并分析了算法的近似度。

关键词:设施选址问题;近似算法;随机线性规划舍入
中图分类号:O224　　**文献标志码**:A
doi:10.3969/j.issn.1007-130X.2013.04.020

Approximation algorithm for limited facility location problem

LIU Yu-tang, FANG Qi-zhi
(School of Mathematical Sciences, Ocean University of China, Qingdao 266100, China)

Abstract: The paper presented an approximation algorithm for the Limited Facility Location problem (LFL), which is a new variant of the classical Uncapacitated Facility Location problem(UFL). The algorithm is based on randomized LP rounding, and its approximation ratio was analyzed.

Key words: facility location problem; approximation algorithm; randomized LP rounding

1　引言

目前,设施选址问题是运筹学中研究最多的问题之一,无容量限制的设施选址 UFL(Unlimited Facility Location)问题是最基础的模型。在一般的 UFL 问题中,给定城市集合 D 和设施集合 F;每个设施 $i \in F$ 对应一个非负的开放费用 f_i;每个城市 $j \in D$ 与每个设施 $i \in F$ 之间存在连接费用 c_{ij}。设施选址问题的目标为:选择哪些设施进行开放,且如何将每个城市连接到一个开放的设施,才能使得设施开放费用和城市与设施的连接费用之和达到最小。

UFL 问题是经典的 NP-hard 问题[1],也就是说,除非 P=NP,否则该问题不可能存在多项式时间算法。给定一个最小化 NP-hard 问题 $\prod$,如果对于 $\prod$ 的任何实例,算法 A 能够在多项式时间内得到目标函数值均不超过最优值 ρ 倍的可行解,则称该算法 A 是一个 ρ-近似算法,并称 ρ 为该算法 A 的近似比。有关 UFL 问题近似算法的研究有很多,例如,线性舍入、原始对偶方法、局部搜索以及各种方法的组合及变种。迄今为止,Li[2]给出的算法得到了最好近似比(1.488);Guha 和 Kuller[3]则证明了 UFL 问题不可能存在近似比小于或等于 1.463 的多项式时间近似算法,除非 NP$\subseteq$DTIME[$n^{O(\log\log n)}$]。设施选址问题的很多变形也得到了广泛的研究[4,5]。

现实中常常出现这种情况:若干个可选设施所起的作用或类型是相同的,因而只需要选取一个进行建设。针对这种情况,本文引入限制设施选址 LFL(Limited Facility Location)问题。在 UFL 模型基础上,限制设施选址问题对设施的开放增加了一个限制:设施集合 F 被划分为有限个互不相交的子集 $F_k(k=1,2,\cdots,R)$,每个子集 F_k 内至多开放一个设施。显然,UFL 问题可以看作是一类特殊的 LFL 问题,其中设施集合 F 被划分为单元集。由此可知,LFL 问题是比 UFL 问题更为广泛的一类问题。

本文组织如下:第 2 节引入 LFL 问题的整数

* 收稿日期:2012-05-16;修回日期:2012-08-20
通讯地址:266100 山东省青岛市松岭路 238 号中国海洋大学数学科学学院 方奇志
Address: School of Mathematical Sciences, Ocean University of China, 238 Songling Rd, Qingdao 266100, Shandong, P. R. China

规划模型,并给出基于随机线性规划舍入技术的近似算法;第3节证明算法的近似比为 $1+(A-1)/e$(其中 A 为与问题实例有关的常数);第4节给出算法的数值模拟结果。

2 模型和算法

限制设施选址问题可正式描述为:给定城市集合 D 和设施集合 F;每个设施 $i\in F$ 对应一个非负的开放费用 f_i;每个城市 $j\in D$ 与每个设施 $i\in F$ 之间存在连接费用 c_{ij};设施集合 F 被划分为有限个互不相交的子集 $F_k(k=1,2,\cdots,R)$,每个子集 F_k 内被限制至多开放一个设施。LFL 的问题是:选择哪些设施进行开放,并且如何将每个城市连接到一个开放设施,才能使得设施开放费用和城市与设施的连接费用之和达到最小。

令 $x_{ij}(i\in F,j\in D)$、$y_i(i\in F)$ 均为 0-1 变量:若城市 $j\in D$ 连接到设施 $i\in F$,则 $x_{ij}=1$,否则 $x_{ij}=0$;若设施 $i\in F$ 开放,则 $y_i=1$,否则 $y_i=0$。那么,LFL 问题的 0-1 规划模型如下:

$$\text{Min } w=\sum_{i\in F}f_iy_i+\sum_{i\in F}\sum_{j\in D}c_{ij}x_{ij}$$

$$\text{s.t.}\quad \sum_{i\in F}x_{ij}=1,\forall j\in D \tag{1}$$

$$\text{(ILP)}\quad \sum_{i\in F_k}y_i\leqslant 1,\forall k=1,2,\cdots,R \tag{2}$$

$$x_{ij}\leqslant y_i,\forall i\in F,j\in D \tag{3}$$

$$x_{ij}\in\{0,1\},\forall i\in F,j\in D \tag{4}$$

$$y_i\in\{0,1\},\forall i\in F \tag{5}$$

其中,约束(1)确保每个城市都被连接到一个设施;约束(2)保证在每个设施子集 F_k 中至多开放一个设施;约束(3)确保每个城市连接到的设施必定是开放的。目标函数为设施的总开放费用和城市与开放设施之间的总连接费用之和。

由于无容量限制的设施选址问题是 NP-hard,且 UFL 问题可以看作是一类特殊的问题,因此有如下的复杂性结论:

定理1 限制设施选址 LFL 问题是 NP-hard。

下面利用随机线性规划舍入技术,给出 LFL 的一个随机近似算法。

将上述 0-1 规划(ILP)中的 0-1 变量松弛为区间[0,1]上的变量,则得到相应的线性规划松弛 LPR(Linear Programming Relaxation):

$$\text{Min}\quad w=\sum_{i\in F}f_iy_i+\sum_{i\in F}\sum_{j\in D}c_{ij}x_{ij}$$

$$\text{s.t.}\quad \sum_{i\in F}x_{ij}=1,\forall j\in D$$

$$\text{(LPR)}\quad \sum_{i\in F_k}y_i\leqslant 1,\forall k=1,2,\cdots,R$$

$$x_{ij}\leqslant y_i,\forall i\in F,j\in D$$

$$x_{ij}\geqslant 0,\forall i\in F,j\in D$$

$$y_i\geqslant 0,\forall i\in F$$

由于约束条件(1)、(2)和目标函数(求最小)的限制,容易验证约束 $x_{ij}\geqslant 0(i\in F,j\in D)$ 和 $y_i\geqslant 0(i\in F)$ 等价于 $0\leqslant x_{ij}\leqslant 1(i\in F,j\in D)$ 和 $0\leqslant y_i\leqslant 1(i\in F)$。

借助于线性规划松弛,给出如下随机近似算法。

算法 随机线性规划舍入算法(Randomized LP Rounding)

步骤1 求解线性规划松弛 LPR,得最优分数解 (x^*,y^*) 及最优值 opt_{LPR}。

步骤2 对每个独立的设施子集 $F_k(k=1,2,\cdots,R)$,根据以下规则确定设施是否开放:

假设子集 F_k 内有 s 个设施,设为 $i_1,i_2,\cdots,i_s$,则 $y_{i_1}+y_{i_2}+\cdots+y_{i_s}\leqslant 1$。随机产生一个不超过1的正数 λ,若 $\lambda\leqslant y_{i_1}$,则只开放设施 i_1;若 $y_{i_1}<\lambda\leqslant y_{i_1}+y_{i_2}$,则只开放设施 i_2;以此类推,若 $\lambda>y_{i_1}+y_{i_2}+\cdots+y_{i_s}$(此时 $y_{i_1}+y_{i_2}+\cdots+y_{i_s}<1$),则 F_k 中所有设施都不开放。

记得到的开放设施集合为 F^*。若 F^* 为空集,则重复步骤2直至 F^* 至少包含一个开放设施。

步骤3 对于每个城市 $j\in D$,计算它到 F^* 中设施的最小连接费用,并将它连接到这一最小连接费用所对应的设施。计算当前得到的可行解的目标函数值。

步骤4 多次重复步骤2和步骤3,输出最小目标函数值所对应的可行解作为原 LFL 问题的近似最优解。

由于线性规划 LPR 可在多项式时间内求解,上述算法其他步骤也均可在多项式时间内完成,故该随机线性规划舍入算法是一个多项式时间算法,算法复杂度由线性规划的规模所决定。

3 算法近似比分析

本节重点分析上一节给出的随机线性规划舍入算法的近似比。首先定义 $Max(j)=\max_{i\in F}\{c_{ij}\}$,$Min(j)=\min_{i\in F}\{c_{ij}\}$。

定理2 给定限制设施选址(LFL)问题的实例,若存在正常数 A 使得任意城市 $j\in D$ 都有

$Max(j) \leqslant A \cdot Min(j)$,则 LFL 问题的随机线性规划舍入算法所得到的可行解的目标函数值的期望 $E[w]$ 满足:

$$E[w] \leqslant (1+\frac{A-1}{e}) \cdot opt_{\mathrm{LPR}}$$

证明 设 LFL 问题的随机线性规划舍入算法输出结果所对应的 0-1 解为 $(\hat{x},\hat{y})$。由算法步骤 2 可知,对于 $i \in F$,$\hat{y}_i$ 以概率 y_i^* 取值 1,以概率 $1-y_i^*$ 取值 0。因此,设施开放费用的期望值为:

$$E[\sum_{i\in F} f_i \hat{y}_i] = \sum_{i\in F} f_i E[\hat{y}_i] = \sum_{i\in F} f_i y_i^*$$

下面考虑连接费用的期望 $E[\sum_{j\in D}\sum_{i\in F} c_{ij}\hat{x}_{ij}]$。

对于每个城市 $j\in D$,不妨设 $c_{1j} \geqslant c_{2j} \geqslant \cdots \geqslant c_{|F|j}$。定义随机变量 X_i:若 $i,i+1,\cdots,|F|$ 中有设施开放,则 $X_i=1$;否则,$X_i=0$。因此,城市 j 的连接费用为:

$$\sum_{i\in F} c_{ij}\hat{x}_{ij} = \sum_{i=1}^{|F|} X_i(c_{ij}-c_{(i-1)j}) + X_1 c_{0j}$$

令 $c_{0j}=c_{1j}$。因为 $1,2,\cdots,|F|$ 中一定有开放设施,故 $X_1=1$。城市 j 的连接费用的期望为:

$$E[\sum_{i\in F} c_{ij}\hat{x}_{ij}] = E[\sum_{i=1}^{|F|} X_i(c_{ij}-c_{(i-1)j})] + c_{1j} =$$
$$\sum_{i=1}^{|F|} E[X_i](c_{ij}-c_{(i-1)j}) + c_{1j}$$

由不等式 $1-\sum_{i=1}^{n} a_i \leqslant \prod_{i=1}^{n}(1-a_i)(n=1,2,3,\cdots)$ 和 $1-x \leqslant e^{-x}$ 可得:

$$\Pr[X_i=0] = \prod_{k=1}^{R}(1-\sum_{\substack{l=i\\ l\in F_k}}^{|F|} y_l^*) \leqslant \prod_{l=i}^{|F|}(1-y_l^*) \leqslant$$
$$\prod_{l=i}^{|F|} e^{-y_l^*} = e^{-\sum_{l=i}^{F} y_l^*}$$

由 $x_{lj}^* \leqslant y_l^*(l=1,2,\cdots,|F|)$ 可得:

$$e^{-\sum_{l=i}^{F} y_l^*} \leqslant e^{-\sum_{l=i}^{F} x_{lj}^*}$$

再由不等式 $1-e^{-x} \geqslant (1-\frac{1}{e})x(0\leqslant x\leqslant 1)$ 可得:

$$E[X_i] = \Pr[X_i=1] \geqslant$$
$$1-e^{-\sum_{l=i}^{F} x_{lj}^*} \geqslant (1-\frac{1}{e})\sum_{l=i}^{|F|} x_{lj}^*$$

又因 $c_{ij}-c_{(i-1)j} \leqslant 0$,故:

$$\sum_{i=1}^{|F|} E[X_i](c_{ij}-c_{(i-1)j}) + c_{1j} \leqslant$$
$$\sum_{i=1}^{|F|}(1-\frac{1}{e})\sum_{l=i}^{|F|} x_{lj}^*(c_{ij}-c_{(i-1)j}) + c_{1j} =$$
$$(1-\frac{1}{e})\sum_{i=1}^{|F|} c_{ij}(\sum_{l=i}^{|F|} x_{lj}^* - \sum_{l=i+1}^{|F|} x_{lj}^*) -$$
$$(1-\frac{1}{e})\sum_{l=1}^{|F|} x_{lj}^* c_{1j} + c_{1j} =$$
$$(1-\frac{1}{e})\sum_{i=1}^{|F|} c_{ij}x_{ij}^* + \frac{1}{e}c_{1j}$$

由此可得:

$$E[\sum_{j\in D}\sum_{i\in F} c_{ij}\hat{x}_{ij}] \leqslant (1-\frac{1}{e})\sum_{j\in D}\sum_{i\in F} c_{ij}x_{ij}^* + \frac{1}{e}\sum_{j\in D} c_{1j}$$

因此

$$E[\sum_{j\in D}\sum_{i\in F} c_{ij}\hat{x}_{ij}] + E[\sum_{i\in F} f_i\hat{y}_i] \leqslant$$
$$(1-\frac{1}{e})\sum_{j\in D}\sum_{i\in F} c_{ij}x_{ij}^* + \frac{1}{e}\sum_{j\in D} c_{1j} + \sum_{i\in F} f_i y_i^* =$$
$$(1-\frac{1}{e})\sum_{j\in D}\sum_{i\in F} c_{ij}x_{ij}^* + \frac{1}{e}\sum_{j\in D} Max(j) + \sum_{i\in F} f_i y_i^* \leqslant$$
$$(1-\frac{1}{e})\sum_{j\in D}\sum_{i\in F} c_{ij}x_{ij}^* + \frac{A}{e}\sum_{j\in D} Min(j) + \sum_{i\in F} f_i y_i^* \leqslant$$
$$(1-\frac{1}{e})\sum_{j\in D}\sum_{i\in F} c_{ij}x_{ij}^* + \frac{A}{e}\sum_{j\in D}\sum_{i\in F} c_{ij}x_{ij}^* + \sum_{i\in F} f_i y_i^* \leqslant$$
$$(1+\frac{A-1}{e})(\sum_{j\in D}\sum_{i\in F} c_{ij}x_{ij}^* + \sum_{i\in F} f_i y_i^*) =$$
$$(1+\frac{A-1}{e}) opt_{\mathrm{LPR}}$$

□

推论 1 限制设施选址 LFL 问题的随机线性规划舍入算法的近似比为 $(1+\frac{A-1}{e}) \cdot opt_{\mathrm{LPR}}$。

推论 2 对于限制设施选址 LFL 问题,存在近似比为 $(1+\frac{A-1}{e}) \cdot opt_{\mathrm{LPR}}$ 的确定性多项式时间近似算法。

证明 根据定理 2,由条件期望的标准技术[6],将算法去随机化可得所得到的可行解的目标函数值 w 满足 $w \leqslant (1+\frac{A-1}{e}) \cdot opt_{\mathrm{LPR}}$。

推论 1 和 2 中 A 的定义同定理 2。 □

4 数值模拟

本节将利用随机生成的实例来检验算法的精度和效率。

随机生成的实例及通过随机线性规划舍入算法所得结果如表 1 所示。

表 1 中,m 表示城市的数目;n 表示设施的个数;各实例中设施集合平均分为六个子集;w 为随机线性规划舍入算法的输出结果;opt_{LPR} 为线性规划松弛 LPR 的最优目标函数值。

对于 $m=20,n=30$ 的实例,每个设施子集内有五个设施,故每个子集至多有六种设施开放选择(包括不开放任何设施的情形),因此最多会有

Table 1 Result of numerical computation

表 1 数值模拟结果

m	n	重复次数	w	opt_{LPR}	w/opt_{LPR}
20	30	6	6 538	6 125	1.067 4
20	60	6	6 117	5 646.7	1.083 3
40	60	6	9 929	8 556.7	1.160 4
40	60	15	9 929	8 556.7	1.160 4
40	60	30	9 891	8 556.7	1.155 9
40	60	50	9 803	8 556.7	1.145 7
40	60	100	9 389	8 556.7	1.097 3
40	84	100	10 148	9 206.6	1.102 2
50	120	100	11 233	9 374.2	1.198 3

$6^6-1=46655$ 种情形，而算法仅用了六次迭代就得到了比值为 1.067 4 的算法结果，证明该算法效果好且效率高。

由表 1 还可以得出，相同的实例规模（m 与 n 相同）下，随机线性规划舍入算法重复次数（步骤 4）越多，输出的结果越好；迭代次数相同时，实例规模越小，算法输出的结果越好。

为了检验常数 A 对算法近似比的影响，我们设计了不同 A 值下相同规模的若干实例。算法的数值模拟结果如表 2 所示。

Table 2 Results in different A

表 2 不同 A 值下相同规模实例的模拟结果（m=40，n=60，重复次数=20）

A	w	opt_{LPR}	w/opt_{LPR}
1	41 002	41 002	1
2	26 409	25 993	1.016 0
5	16 099	15 754	1.021 9
10	13 192	12 031	1.096 5
50	9 941	9 180.6	1.082 8
100	10 580	9 162.8	1.154 7
1 000	10 054	8 551.4	1.175 7

从表 2 可以得出，在实例的规模相同、且随机线性规划舍入算法重复（步骤 4）次数也相同的情况下，A 的值越小，得到的结果越好。当 A 的值较大时，可通过增加算法重复（步骤 4）次数来提高算法输出结果的精度。

5 结束语

本文从实际问题出发讨论了一类限制的设施选址 LFL 问题：在无容量设施选址 UFL 模型基础上，对设施的开放增加限制，即将设施的集合划分成若干不相交子集、且每个子集内至多开放一个设施。该问题是一般无容量设施选址问题的推广，具有良好的理论和应用价值。

本文设计了一个近似比为 $1+(A-1)/e$（A 为与问题实例有关的常数）的基于随机线性规划舍入的随机近似算法，并对算法进行了数值模拟。模拟实验结果表明该随机近似算法不仅精度高，而且运算效率也令人满意。

参考文献：

[1] Vazirani V V. Approximation algorithms[M]. Berlin Heidelberg Springer-Verlag,2003.

[2] Li S. A 1.488 approximation algorithm for the uncapacitated facility location problem[C]// Proc of the 38th International Colloquium on Automata, Languages and Programming, 2011:77-88.

[3] Guha S, Kuller S. Greedy strikes back: Improved facility location algorithms[J]. Journal of Algorithms, 1999,31(1): 228-248.

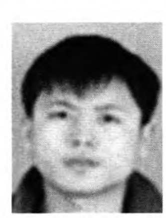
[4] Zhang J. Approximating the two-level facility location problem via a quasi-greedy approach[J]. Math Program,2006,108 (1):159-176.

[5] Aardal K,Chudak F A,Shmoys D B. A 3-approximation algorithm for the k-level uncapacitated facility location problem [J]. Information Processing Letters,1999,72:161-167.

[6] Erdös P, Selfridge J L. On a combinatorial game[J]. Journal of Combinatorial Theory,1973,14(3):298-301.

[7] Chudak F A, Shmoys D B. Improved approximation algorithms for the uncapacitated facility location problem[J]. SIAM Journal on Computing,2003,33:1-25.

作者简介：



刘玉堂（1988 -），男，山东莒县人，硕士生，研究方向为组合最优化和近似算法。**E-mail**：wwwwww1988723@163.com

LIU Yu-tang, born in 1988, MS candidate, his research interests include combinatorial optimization, and approximation algorithms.

方奇志（1966 -），女，安徽安庆人，博士，教授，CCF 会员（E200009124S），研究方向为组合最优化和近似算法。**E-mail**：qfang@ouc.edu.cn

FANG Qi-zhi, born in 1966, PhD, professor, CCF member(E200009124S), her research interests include combinatorial optimization, and approximation algorithms.