中图分类号:　　　　　　　　　　　　　　　　　　学校代码: 10055

UDC:　　　　　　　　　　　　　　　　　　　　　　密级: 公开

南開大學
博 士 学 位 论 文

多层与随机设施选址问题的近似算法

Approximation algorithms for multi-level and stochastic facility location problems

论文作者 吴晨晨　　指导教师 杨庆之教授

申请学位 理学博士　　培养单位 数学科学学院

学科专业 计算数学　　研究方向 组合优化

答辩委员会主席 韩继业　　评 阅 人 匿名评审

南开大学研究生院

二〇一四年四月

南开大学学位论文使用授权书

根据《南开大学关于研究生学位论文收藏和利用管理办法》，我校的博士、硕士学位获得者均须向南开大学提交本人的学位论文纸质本及相应电子版。

本人完全了解南开大学有关研究生学位论文收藏和利用的管理规定。南开大学拥有在《著作权法》规定范围内的学位论文使用权，即：(1)学位获得者必须按规定提交学位论文(包括纸质印刷本及电子版)，学校可以采用影印、缩印或其他复制手段保存研究生学位论文，并编入《南开大学博硕士学位论文全文数据库》；(2)为教学和科研目的，学校可以将公开的学位论文作为资料在图书馆等场所提供校内师生阅读，在校园网上提供论文目录检索、文摘以及论文全文浏览、下载等免费信息服务；(3)根据教育部有关规定，南开大学向教育部指定单位提交公开的学位论文；(4)学位论文作者授权学校向中国科技信息研究所及其万方数据电子出版社和中国学术期刊(光盘)电子出版社提交规定范围的学位论文及其电子版并收入相应学位论文数据库，通过其相关网站对外进行信息服务。同时本人保留在其他媒体发表论文的权利。

非公开学位论文，保密期限内不向外提交和提供服务，解密后提交和服务同公开论文。

论文电子版提交至校图书馆网站：http://202.113.20.163:8001/paper/index.jsp。

本人承诺：本人的学位论文是在南开大学学习期间创作完成的作品，并已通过论文答辩；提交的学位论文电子版与纸质本论文的内容一致，如因不同造成不良后果由本人自负。

本人同意遵守上述规定。本授权书签署一式两份，由研究生院和图书馆留存。

作者暨授权人签字：吴晨晨

2014 年 5月 20日

南开大学研究生学位论文作者信息

论文题目	多层与随机设施选址问题的近似算法				
姓　　名	吴晨晨	学号	1120110030	答辩日期	2014年5月29日
论文类别	博士■　学历硕士□　硕士专业学位□　高校教师□　同等学力硕士□				
院/系/所	数学科学学院	专　业	计算数学		
联系电话	13426252066	**Email**	chenchen86711@gmail.com		
通讯地址(邮编)：天津市南开区南开大学数学科学学院					
备注：			是否批准为非公开论文	否	

注：本授权书适用我校授予的所有博士、硕士的学位论文。由作者填写(一式两份)签字后交校图书馆，非公开学位论文须附《南开大学研究生申请非公开学位论文审批表》。

Y2755412

南开大学学位论文原创性声明

本人郑重声明：所呈交的学位论文，是本人在导师指导下进行研究工作所取得的研究成果。除文中已经注明引用的内容外，本学位论文的研究成果不包含任何他人创作的、已公开发表或者没有公开发表的作品的内容。对本论文所涉及的研究工作做出贡献的其他个人和集体，均已在文中以明确方式标明。本学位论文原创性声明的法律责任由本人承担。

学位论文作者签名： 吴晨晨 2014 年 5 月 20 日

非公开学位论文标注说明

(本页表中填写内容须打印)

根据南开大学有关规定，非公开学位论文须经指导教师同意、作者本人申请和相关部门批准方能标注。未经批准的均为公开学位论文，公开学位论文本说明为空白。

论文题目			
申请密级	□限制(≤2年) □秘密(≤10 年) □机密(≤20年)		
保密期限	20 年 月 日至20 年 月 日		
审批表编号		**批准日期**	20 年 月 日

南开大学学位办公室盖章(有效)

注：限制★2年(可少于2年);秘密★10年(可少于10年);机密★20年(可少于20年)

中文摘要

设施选址问题是组合优化中经典的NP-困难问题之一。除非 $P = NP$，否则这类问题不可能在多项式时间内求得最优解。很多NP-困难问题可以利用近似算法有效求解。近似算法在多项式时间内得到可估计的可行解。在本论文中，我们考虑两阶段随机设施选址问题，带线性惩罚的两阶段随机设施选址问题，2-层设施选址问题和带软容量限制的设施选址问题。

在第二章，我们考虑两阶段随机设施选址问题。在该问题中，被服务的顾客集合依赖于实现的场景及其相应的概率，设施的开设费用依赖于第一/第二阶段。问题的目标是极小化设施开设费用和连接费用之和的期望。与常用的近似比不同，我们考虑估计实现场景解质量的单场景界。利用线性规划舍入技巧，我们得到2.3613单场景界。

在第三章，我们考虑带线性惩罚的两阶段随机设施选址问题。在两阶段随机设施选址问题的基础上，有些顾客可以不被服务。因而需要支付一定的惩罚费用，最终极小化设施开设费用，连接费用以及不被服务的顾客惩罚费用的总期望。我们给出带线性惩罚的随机设施选址问题基于线性规划的3.0294-近似算法。

在第四章，我们考虑2-层设施选址问题。在该问题中，顾客需要连接到具有先后关系的两层设施上。对任意的正常数 $\varepsilon > 0$，我们给出一个原始-对偶 $3(1+\varepsilon)$-近似算法。区别于标准的原始-对偶算法，需要利用近似分离神谕得到对偶可行解。进一步，利用贪婪调整的技巧，可将近似比改进到 $2.172(1+\varepsilon)^2$。

在第五章，我们考虑带软容量限制的 k-层设施选址问题。在该问题中，每个设施都有软容量限制。通过支付多倍的设施开设费用，每个设施的容量可以增加相同倍数。通过 k-层设施选址问题的双因子，利用所谓的双因子归约，我们可以得到5.5053-近似算法。

关键词： 组合优化; 设施选址; 近似算法。

Abstract

Facility location problem is one of the classical NP-hard problems in combinatorial optimization which can not be solved in polynomial time, unless P=NP. Many NP-hard problems can be handled efficiently by approximation algorithms which output feasible solutions with estimated value in polynomial time. In this thesis, we consider the two-stage stochastic facility location problem, the two-stage stochastic facility location with linear penalties, the 2-level facility location problem, and the k-level facility location problem with soft capacities.

In Chapter 2, we consider the two-stage stochastic facility location problem in which the client set depends on the realized scenario associated with the corresponding probability and the facility opening cost depends on the first/second stage. Our goal is to minimize the expected total cost including the facility opening cost and the connection cost. Instead of the usual approximation ratio, we consider the per-scenario bound which estimates quality for each realized scenario. Using linear programming rounding technique, we obtain the per-scenario bound 2.3613.

In Chapter 3, we consider the stochastic facility location problem with linear penalties. In this problem, each client may be un-served which incurrs a penalty cost associated with this specified client. Our goal is to minimize the expected total cost including the facility opening cost, the connection cost, and the penalized cost. We propose a linear programming rounding 3.0294-approximation algorithm.

In Chapter 4, we consider the 2-level facility location problem in which each client must be served by two-level facilities. We present a primal-dual $3(1+\varepsilon)$-approximation algorithm, where ε is an arbitrary positive constant. Comparing with the standard primal-dual scheme, we use an approximated separation oracle technique to obtain a dual feasible solution. We further improve the approximation ratio to $2.172(1+\varepsilon)^2$ by adopting a greedy adjustment technique.

In Chapter 5, we consider the k-level facility location problem with soft capacities. In this problem, the capacity of each facility can be increased multiply by paying the same multiple facility opening cost. Using the so-called bifactor reduction which estimates the desired approximation ratio via a bifactor of the k-level facility location problem, we obtain a 5.5053-approximation algorithm.

Key Words: Combinatorial optimization; Facility location problem; Approximation algorithm.

目录

第一章 绪论

选址问题在社会生产，生活中有极其重要的作用。所谓设施，是指生产运作过程得以进行的硬件手段，通常是由工厂、办公楼、车间、设备、仓库等物质实体所构成。而在选址问题中，设施选址(facility location)是一个重要分支。所谓设施选址，是指如何运用科学的方法决定设施的地理位置，使之与企业的整体经营运作系统有机结合，以便有效、经济地达到企业的经营目的。简单的说，设施选址问题是指为了满足顾客的需求，在备选的地址上选择一些置放设施最终使得成本最小。实际上，这个问题也可以理解为在每个地址上都有一个待开设的设施，需要决策选择哪些设施开设最终使得满足所有顾客需求的成本最小。设施选址问题是运筹学以及计算几何中的重要问题，受到了极大的关注。

最简单的设施选址问题是由德国著名的经济学家和社会学家 Alfred Weber提出的“韦伯问题”(见图1.1)。在韦伯问题中，顾客到开设设施的成本利用顾客到设施的之间距离衡量，需要开设一个设施，使得所有的顾客到这一被选的设施之间的距离和最小。事实上，这个问题可以通过试验所有备选的地址，在其中找到成本最小的方法来求解。

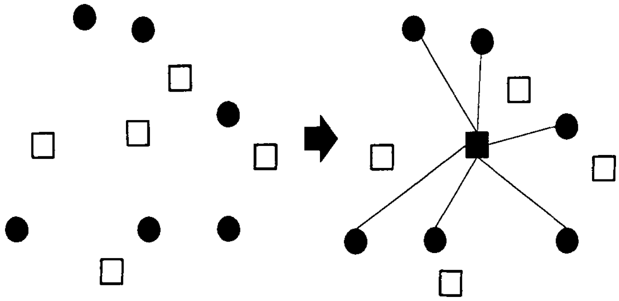


图 1.1 韦伯问题

现在所讨论的选址问题一般是指，选择多个地址置放设施使得置放的费用和顾客到设施的距离费用之和最小。这个问题可视为集合覆盖问题的特殊情况。集合覆盖问题是指，给定一个基集以及基集的一个子集簇，每个子集都有相应的费用，需要在子集簇中找到一些子集使得所有子集的并集即为基集并且所选

子集的费用和最小。

简单的说，在计算复杂性理论中，P-问题是指在关于输入的多项式时间内可以找到问题的最优解。NP-完全问题是指在多项式时间可以判定解的正确性，并且可以由其他 NP-完全问题都在多项式时间内归约。NP-困难问题是指任何 NP-完全问题都可在多项式时间内归约到这个问题。NP-困难问题不一定在多项式时间内判定解的正确性。1971年，Cook和 Leonid Levin相对独立的提出 P是否等于 NP这一难题。到现在为止，多数计算机科学家相信 P≠NP。现在已知的结果有在 Unique game的假设下，P≠NP。当 P≠NP时，P-问题，NP-完全问题和 NP-困难问题的关系可见图1.2。具体的计算复杂性的理论请参见文献 [1]。

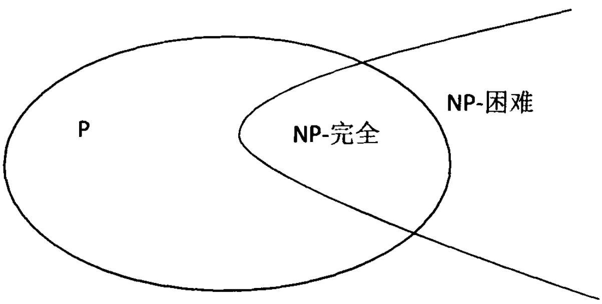


图 1.2　P，NP-完全和 NP-困难问题的关系

集合覆盖是 Karp[2] 提出的21个 NP-困难问题之一。设施选址可以归约成集合覆盖问题，因而也是 NP-困难的，也就是，在 P≠ NP 的条件下，关于输入规模的多项式时间内找不到问题的最优解。

处理 NP-困难问题的方法有很多种。有谚语说,“Fast. Cheap. Reliable. Choose two.” (快速，经济，可靠，选择其中的两个。) Williamson和 Shmoys[3]在“The Design of Approximation Algorithm”一书中指出“If P≠NP，we can’t simultaneously have algorithms that (1) find optimal solutions (2) in polynomial time (3) for any instance.” (如果 P≠NP，我们不能得到一个算法同时满足(1)找到最优解; (2)多项式时间; (3)对任意的实例。) 如果选择(1)和(2)，就是考虑 NP-困难问题的可在关于输入规模多项式时间内求解的特殊情况。如果选择(1)和(3)，就是不考虑算法的运行时间而找到问题的最优解，例如整数规划的分支定界方法等。如果选择(2)和(3)，就是不要求找到问题的最优解，而是需要在有效的时间内找到问题的近似最优解，例如，近似算法和启发式算法等。启发式算法是在工程应用中

常见的方法。与近似算法不同，在最坏情况下启发式算法所得解的质量无法保证。同时，已有许多专家学者致力于设施选址问题的近似算法研究(见[4, 3, 5])。本论文中，我们关注近似算法。徐大川和张家伟[6]总结了各类设施选址问题的近似算法。

定义 1.1 对于极小化问题，一个算法被称为 ρ-近似算法，如果该算法输出的解对应的目标值 $\mathrm{Cost}(Alg)$ 与最优值 Opt 的比值不超过 ρ，即，

$$\frac{\mathrm{Cost}(Alg)}{\mathrm{Opt}} \leq \rho,$$

其中 ρ 被称为近似比。

显然，极小化问题的近似比大于 1。

注释 1.2 如果对于极大化问题，ρ-近似算法的定义为

$$\frac{\mathrm{Cost}(Alg)}{\mathrm{Opt}} \geq \rho.$$

显然，极大化问题的近似比小于 1。

由于在选址问题中包括两部分费用，设施开设费用以及连接费用。因而除了近似比，还有一种更紧的刻画算法所得解质量的标准—双因子。(γ_f, γ_c) 被称为算法的双因子，如果算法所得解的费用不超过 $\gamma_f \mathrm{Cost}_f(\mathrm{Opt}) + \gamma_c \mathrm{Cost}_c(\mathrm{Opt})$，其中 $\mathrm{Cost}_f(\mathrm{Opt})$ 和 $\mathrm{Cost}_c(\mathrm{Opt})$ 分别表示最优解中设施费用和连接费用。显然，$\max\{\gamma_f, \gamma_c\}$ 即为算法的近似比。

根据选址问题中的距离所满足的性质，设施选址问题可以分为度量型的设施选址以及非度量型的设施选址问题。在度量型的设施选址问题中，距离需满足以下条件。

- 非负性: 所有的距离都是非负的。
- 对称性: 同一组设施和顾客，设施到顾客的距离与顾客到设施的距离相同。
- 满足三角不等式: 设施和顾客构成的图中，任意三点可构成三角形。在该三角形中，两边之和大于第三边，两边之差小于第三边。

若距离不满足以上三个性质，则被称为非度量型的设施选址问题。在本论文中，若不特别提出，我们都是考虑度量型的设施选址问题。

一般的说，算法可以分为组合的和非组合的。所谓组合算法是指在算法过程中，只涉及加，减，乘，除以及比较运算。否则的话，则称为非组合算法。非组合算法一般需要求解松弛规划，基于规划的解通过一定的技巧再舍入成整

数解。这类算法由于在一个松弛解的基础上进行舍入，一般比组合算法得到的近似比好。但是，我们仍考虑组合算法的原因可以分为以下几点.

- 组合算法的运行时间相对较低。
- 组合算法一般较依赖于问题的结构，因而更能揭示问题的结构。
- 组合算法较容易编程实现。

第一节 无容量限制的设施选址问题

在最简单，最经典的设施选址问题中，设施可以提供的服务没有限制，即，设施可以为任何多个顾客提供服务，这类问题被称为无容量限制的设施选址问题(uncapacitated facility location problem，简称 UFLP)。具体的说，在 UFLP中，给定设施集合 F 和顾客集合 D，每个设施 i 的开设费用为 f_i，设施 i 和顾客 j 之间满足度量性的单位连接费用为 c_{ij}(即前面提到的距离，也称为单位服务费用)，每个顾客的需求量为 d_j。我们的目标是开设一些设施 $F' \subseteq F$，并将顾客连接到一个开设设施，使得开设费用和连接费用之和最小，即，

$$\min_{F' \subseteq F} \left\{ \sum_{i \in F'} f_i + \sum_{j \in D} \min_{i \in F'} d_j c_{ij} \right\}.$$

一般情况，在本论文余下的内容中，若不做特殊说明，d_j 均假设为 1。在UFLP中，需要做两类决策，一为开设哪些设施，为此对每个设施 i 引入 $0-1$ 变量 y_i，$y_i = 1$ 表示设施 i 开设，否则表示设施 i 不开设。二为顾客连接到哪个设施上，为此引入 $0-1$ 变量 x_{ij}，$x_{ij} = 1$ 表示顾客 j 连接到设施 i 上，否则表示不连接。需要将顾客连接到一个设施上，即，$\sum_{i \in F} x_{ij} = 1$。并且当顾客连接到一个设施上，该设施一定需要开设，即，强制约束 $x_{ij} \le y_i$。这样，就可以给出UFLP的整数线性规划

$$\begin{aligned} \min \quad & \sum_{i \in F} f_i y_i + \sum_{i \in F} \sum_{j \in D} c_{ij} x_{ij} \\ \text{s. t.} \quad & \sum_{i \in F} x_{ij} = 1, \qquad \forall j \in D, \\ & x_{ij} \le y_i, \qquad \forall i \in F, j \in D, \\ & x_{ij}, y_i \in \{0,1\}, \qquad \forall i \in F, j \in D. \end{aligned} \tag{1.1}$$

事实上，规划(1.1)中第一组等式约束等价于不等式约束

$$\sum_{i \in F} x_{ij} \ge 1, \qquad \forall j \in D.$$

这是因为如果 $\sum_{i\in F}x_{ij}$ 严格大于 1 时，说明有多个设施使得对应的 x_{ij} 为 1，只取其中之一，其余的都重新赋值为 0，仍然得到可行解，并且目标值不增加。因此，若将第一组等式约束修改成不等式约束，最优解和最优值不发生改变。因而得到常用的 UFLP整数线性规划

$$\begin{aligned}
\min \quad & \sum_{i\in F} f_i y_i + \sum_{i\in F}\sum_{j\in D} c_{ij}x_{ij} \\
\text{s. t.} \quad & \sum_{i\in F} x_{ij} \geq 1, \qquad \forall j \in D, \\
& x_{ij} \leq y_i, \qquad \forall i \in F, j \in D, \\
& x_{ij}, y_i \in \{0,1\}, \qquad \forall i \in F, j \in D.
\end{aligned} \tag{1.2}$$

为了得到规划(1.2)的线性松弛，松弛规划(1.2)中的 $0-1$ 变量约束，将其松弛为 $x_{ij}, y_i \in [0,1]$。但事实上，这组松弛的约束等价于 $x_{ij}, y_i \geq 0$。类似的，这是因为如果 $x_{ij}, y_i > 1$，可以将其重新定义成 1。这样，新的解不违背约束仍是可行解，同时，目标值不增加。因而，规划(1.2)的线性规划松弛为

$$\begin{aligned}
\min \quad & \sum_{i\in F} f_i y_i + \sum_{i\in F}\sum_{j\in D} c_{ij}x_{ij} \\
\text{s. t.} \quad & \sum_{i\in F} x_{ij} \geq 1, \qquad \forall j \in D, \\
& x_{ij} \leq y_i, \qquad \forall i \in F, j \in D, \\
& x_{ij}, y_i \geq 0, \qquad \forall i \in F, j \in D.
\end{aligned} \tag{1.3}$$

给定规划(1.3)的一个解 (x,y)，称每个设施 i 是分数开设的，每个顾客 j 分数连接到集合 $\{i : x_{ij} > 0\}$ 上。

同时，利用对偶理论，可以得到规划(1.3)的对偶规划

$$\begin{aligned}
\max \quad & \sum_{j\in D} \alpha_j \\
\text{s. t.} \quad & \alpha_j \leq \beta_{ij} + c_{ij}, \qquad \forall i \in F, j \in D, \\
& \sum_{j\in D} \beta_{ij}, \qquad \forall i \in F, \\
& \alpha_j, \beta_{ij} \geq 0, \qquad \forall i \in F, j \in D.
\end{aligned} \tag{1.4}$$

规划(1.4)中，对偶变量 α_j 可以看做顾客 j 的预算，β_{ij} 可以看做顾客 j 对设施 i 的开设所做的贡献。对偶变量的含义对设计原始-对偶算法具有重要的意义。

除了以上提出的整数线性规划松弛，UFLP还可以刻画为集合覆盖形式的模型。在集合覆盖问题中，基集 I 的子集簇中的子集 S 对应的费用为 c_S，为此对每个集合 S 引入 $0-1$ 变量 x_S。为了描述集合覆盖约束，即，对于每个顶点 v 必然存在一个子集使得 v 在其中，因此，引入约束 $\sum_{S:v\in S} x_S \geq 1$。这样可以得到集合覆形式的线性整数规划

$$\begin{aligned} \min \quad & \sum_{S} c_S x_S \\ \text{s.t.} \quad & \sum_{S:v\in S} x_S \geq 1, \qquad \forall v \in I, \\ & x_S \in \{0,1\}, \qquad S \subseteq I. \end{aligned} \tag{1.5}$$

由于UFLP的最优解是由一系列的设施 i 以及连接到该设施的顾客 J 构成的星 (i,J)(见图1.3)，最优值是所有的星对应的费用总和。因此可以将UFLP看做基集为顾客集合，以及基集的子集簇为设施和顾客子集构成的星，即 $(i,J)(\forall i \in F, J \subseteq D)$。若顾客 $j \in J$，则称 j 被星 (i,J) 覆盖。同时，星 (i,J) 对应的费用为设施 i 的开设费用以及 J 中的顾客连接到设施 i 上的连接费用之和，即，

$$C_{iJ} = f_i + \sum_{j\in J} c_{ij}.$$

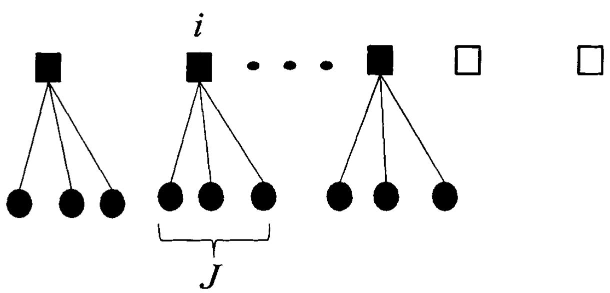


图 1.3 UFLP最优解的结构

对任意的 (i,J) 引入 $0-1$ 变量 X_{iJ}，表示 (i,J) 是否被选到最优解中，即，J 中的顾客是否连接到设施 i 上，$X_{iJ}=1$ 表示星 (i,J) 选到最优解中。因此，得到UFLP集合覆盖形式的整数规划

$$\min \quad \sum_{i\in F}\sum_{J\subseteq D} C_{iJ} X_{iJ}$$

$$\text{s. t.} \quad \sum_{i \in F} \sum_{J: j \in J} X_{iJ} \geq 1, \qquad \forall j \in D, \tag{1.6}$$
$$X_{iJ} \in \{0,1\}, \qquad \forall i \in F, J \subseteq D.$$

对应的线性规划松弛及其对偶规划分别为

$$\begin{aligned} \min \quad & \sum_{i \in F} \sum_{J \subseteq D} C_{iJ} X_{iJ} \\ \text{s. t.} \quad & \sum_{i \in F} \sum_{J: j \in J} X_{iJ} \geq 1, \qquad \forall j \in D, \\ & X_{iJ} \geq 0, \qquad \forall i \in F, J \subseteq D. \end{aligned} \tag{1.7}$$

以及

$$\begin{aligned} \max \quad & \sum_{j \in D} \alpha_j \\ \text{s. t.} \quad & \sum_{j \in J} \alpha_j \leq C_{iJ}, \qquad \forall i \in F, J \subseteq D, \\ & \alpha_j \geq 0, \qquad \forall j \in D. \end{aligned} \tag{1.8}$$

UFLP的集合覆盖形式的整数线性规划对于利用对偶拟合技巧设计近似算法具有重要意义。

考虑 UFLP的近似算法，Hochbaum[7] 利用集合覆盖贪婪算法的思想，给出UFLP的第一个近似算法，得到近似比为 $O(\log n)$，其中 n 是顾客的个数，即，$n = |D|$。在文献[7]中，连接费用不要求满足三角不等式。因而该算法也是非度量型设施选址问题的近似算法。之后，1997年，Shmoys等 [8] 利用线性规划舍入的技巧得到了 UFLP的 4-近似算法，通过随机选取参数将近似比降低到 3.16。线性规划舍入的思想为从规划(1.3)的(分数)最优解出发，将分数解舍入成整数解，其间损失的解质量即为分析近似比所需估计的值。Chudak和 Shmoys[9] 在文献[8]的基础上，随机的开设设施，得到解的费用期望为 1.736 倍最优值，即得到了 1.736-近似算法。Sviridenko[10] 构造连接费用凸函数的上界，利用管道舍入(pipage rounding)的技巧可以得到 1.58-近似算法。

2001年，Jain和 Vazirani[11] 基于线性规划的对偶，构造对偶可行解，再将对偶可行解对应的整数解进行删减得到原问题的整数可行解，分析出近似比为 3。这种技巧被称为原始-对偶。Charikar和 Guha[12] 利用原始-对偶算法 [11] 所得解做为初始解，贪婪的增加开设设施，减少总和费用以改进现行解直到解的质量不能改进，最终得到 1.8526-近似算法。

2003年，Jain等[13]从线性规划对偶的角度利用选址问题的特殊性重新描述了Hochbaum[7]的贪婪算法。首先得到对偶不可行解，同时这个对偶不可行解对应了原问题的一个整数可行解。之后，再将对偶不可行解进行按比例放缩得到对偶可行解。因此，按比例放缩的因子即为近似比。Jain等[13]利用求解线性规划分析出近似比为1.61，同时得到UFLP的双因子$(1,2)$。这种技巧被称为对偶拟合。Mahdian等[14]利用按比例放缩设施费用得到新的实例，对新实例运行Jain等[13]的算法，可以将近似比改进到1.52，同时得到$(1.11,1.78)$的双因子。

此外，还有利用局部搜索的技巧研究UFLP。该技巧的思想为从一个可行解出发，利用定义解的运算，对现行解进行改进，直到找到某个局部最优解，即，所有定义的运算都不能改进算法，得到算法的局部最优解。局部搜索对于任何的组合优化问题都可奏效，但是这种算法是启发式的，一般来说分析不出近似比。但是这类算法对于设施选址问题可以分析出近似比。2004年，对任意的$\varepsilon > 0$，Arya等[15]利用局部搜索的技巧得到了UFLP的$(3+\varepsilon)$-近似算法。若同比例提高设施的费用，运行局部搜索算法，可以将近似比降低到$1+\sqrt{2}+\varepsilon$。

单纯的利用某种技巧已经不能改进设施选址问题的近似比，现在更多的算法是综合考虑以上四种技巧。Bryka和Aardal[16]给出新的线性规划舍入方法，再结合Mahdian等[14]$(1.11,1.78)$的双因子可以得到1.50-近似算法。到现在为止，最好的近似比为Li[17]基于取服从脉冲分布以及均匀分布的组合分布的随机参数设计新的线性规划舍入的近似算法，结合Mahdian等[14]$(1.11,1.78)$得到的1.488-近似算法。

考虑近似比下界，Guha和Khuller[18]指出除非$NP \subseteq DTIME(n^{O(\ln\ln n)})$，否则UFLP的下界为1.463。Sviridenko[19]将这个结论的条件放松到$NP \subseteq P$。

第二节 选址问题的变形

UFLP是选址问题中结构最简单也是最基础的一类模型。实际生产中的模型更为复杂，在UFLP基础上加不同的约束即得到不同的变形。在本节，将介绍几类UFLP的重要变形。

1.2.1 k-层设施选址问题

在UFLP中，顾客的需求只需要一道工序(一层选址)就可以得到满足。但在实际中，由于顾客服务需求的复杂性，需要多道工序才能得以满足，因而需

要多层的选址。这就是多层设施选址问题，也称为 k-层设施选址问题(k-level facility location problem，简称为 k-LFLP)，其中 k 为选址的层数。具体的说，k-LFLP是给定有先后层次关系的设施集合 $F_1,F_2,\cdots,F_k$，以及顾客集合 D。每个设施 $i\in F_1\cup F_2\cup\cdots\cup F_k$ 的开设费用为 f_i。第一层设施 $i_1\in F_1$ 与顾客 $j\in D$ 之间的连接费用为 $c_{i_1 j}$，第 $i_{l-1}\in F_{l-1}$ 层设施与第 $i_l\in F_l$ 层设施之间的连接费用为 $c_{i_{l-1}i_l}$。这里的连接费用均满足度量要求。需要在每层设施中开设一些 $F_1',F_2',\cdots,F_k'$，这些开设的设施组成开设路径 $P'=F_1'\times F_2'\times F_k'$，将每个顾客连接到一条开设路径上使得开设费用和连接费用之和最小，即，

$$\min_{F_l'\subseteq F_l,l=1,2,\cdots,k}\left\{\sum_{l=1}^{k}\sum_{i\in F_l'}f_i+\sum_j\min_{p=(i_1,i_2,\cdots,i_k)\in P'}\left(c_{i_1 j}+\sum_{l=2}^{k}c_{i_{l-1}i_l}\right)\right\}.$$

可以看到，当 $k=1$ 时，k-LFLP即退化为 UFLP。为了描述方便，引入路径集合的记号 $P:=F_1\times F_2\times\cdots\times F_k=\{p=(i_1,i_2,\cdots,i_k):i_1\in F_1,i_2\in F_2,\cdots,i_k\in F_k\}$，设施的集合 $F:=F_1\cup F_2\cup\cdots\cup F_k$，以及顾客 j 到路径 $p=(i_1,i_2,\cdots,i_k)$ 的连接费用 $c_{pj}=c_{i_1 j}+\sum_{l=2}^{k}c_{i_{l-1}i_l}$。记号 $i\in p$ 表示路径 p 上的设施 i。这样，引入 $0-1$ 变量 y_i 表示设施 $i\in F$ 是否开设，$y_i=1$ 表示设施 i 开设，否则设施 i 不开设。x_{pj} 表示顾客 j 是否连接到路径 p 上，$x_{pj}=1$ 表示顾客 j 连接到路径 p 上，否则表示不连接。k-LFLP可以描述为以下整数线性规划

$$\begin{aligned}
\min\quad & \sum_{i\in F}f_iy_i+\sum_{p\in P}\sum_{j\in D}c_{pj}x_{pj} \\
\text{s. t.}\quad & \sum_{p\in P}x_{pj}\geq 1, \qquad \forall j\in D, \\
& \sum_{p:i\in p}x_{pj}\leq y_i, \qquad \forall p\in P,j\in D, \\
& x_{pj},y_i\in\{0,1\}, \qquad \forall i\in F,p\in P,j\in D.
\end{aligned}\tag{1.9}$$

在上述规划中，第一组约束表示每个顾客都必须连接到一条路径上，第二组约束中 $i\in p$ 表示路径 p 上的设施 i，则该约束说明如果顾客连接到某条路径上，则这条路径上的每个设施都必须开设。

对于 k-LFLP，第一个近似算法是 Meyerson等 [20] 提出的 $O(\ln n)$-近似算法，其中 n 是顾客的个数。该算法是纯组合的，不借助线性规划。Ageev等 [21] 利用路径归约，将 k-LFLP归约到单层设施选址问题，即，UFLP，分析得到 3.27-近似算法。Aardal等 [22] 给出 k-LFLP的整数规划(有指数多个变量)，基于

该规划给出线性规划舍入的 3-近似算法。随后，Gabor和 van Ommeren[23] 给出了 k-LFLP的一个新的整数规划，该规划只有多项式个变量，基于线性规划舍入得到近似比仍为 3。上述的整数规划都是基于 k-LFLP的路径性质，但实际上，k-LFLP的最优解是一些树组成的森林(见图1.4)。Byrka和 Rybicki[24] 基于最优解森林结构的特性，提出一个新的整数规划，得到近似比是关于 k 是单调递增的近似算法，并且当 $k\to\infty$，近似比趋向于 3。由于 k-LFLP解的森林结构，Bumb和 Kern[25] 利用原始-对偶技巧得到 6-近似算法。Zhang[26] 利用拟贪婪的算法技巧得到 2-层设施选址问题(2-level Facility location problem，简称 2-LFLP)的 1.77-近似算法。

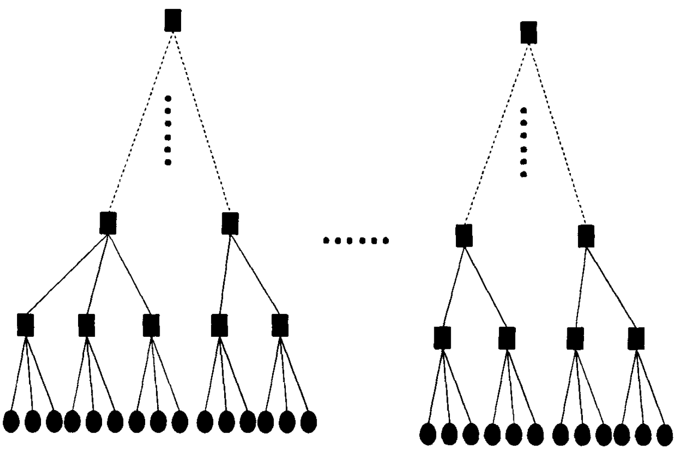


图 1.4 k-LFLP最优解的结构

关于近似比下界， Krishnaswamy和 Sviridenko[27] 证明出除非 $NP \subseteq DTIME$ ($n^{O(\ln\ln n)}$)，2-LFLP($k=2$)和一般的 k-LFLP的近似比的下界分别为 1.539 和 1.61。由此可以看出，单层设施选址问题(UFLP)，2-层设施选址问题以及一般的多层设施选址问题在算法设计的难度上是不同的。

1.2.2 带容量限制的设施选址问题

在 UFLP问题中，设施可以提供的服务量是无限的。但实际中，由于资源有限，这种假设是理想的。实际中更多的是每个设施都有能够提供服务的上界，加上这种约束，即为带容量限制的设施选址问题。在该问题中，每个设施 i 都有容量限制 u_i。具体的，带容量限制的设施选址问题是指给定设施的集合 F 以及顾客的集合 D，每个设施有一个开设费用 f_i 以及容量限制 u_i，支付 f_i, 设施 i 即可提供 u_i 个单位的服务量。设施 i 和顾客 j 之间的单位连接费用为 c_{ij}。需要开设一些设施，将顾客连接到设施上并且不违背开设设施的容量使得开设费用和连接费用之和最小。

如果每个设施可以通过多次开设获得多倍的容量，这类带容量的设施选址问题称为带软容量限制设施选址问题(facility location problem with soft capacities，简称 FLPWSC)。否则，若每个设施至多只能开设一次，则称为带硬容量限制的设施选址问题(facility location problem with hard capacities，简称 FLPWHC)。FLPWSC和 FLPWHC的区别为是否可以复制设施，即，多次开设设施，获得多个设施的容量。因而对设施 i 的开设引入变量 y_i，则在 FLPWSC中 y_i 在自然数域 $\mathbb{N}$ 中取值表示设施 i 的开设次数。但在 FLPWHC中 y_i 为 $0-1$ 变量，$y_i=1$ 表示设施 i 开设，否则表示设施 i 不开设。对顾客 j 是否连接到设施 i 上引入变量 x_{ij}，$x_{ij}=1$ 表示顾客 j 连接到设施 i 上，否则表示不连接。在 FLPWSC中 x_{ij} 为 $0-1$ 变量。而根据顾客的需求是否可以连接到多个设施上以获得满足，FLPWHC可以分为可分的带硬容量限制设施选址问题(splittable facility locaiton with hard capacities)以及不可分的带硬容量设施选址问题(un-splittable facility location problem with hard capacities)。在可分的带硬容量限制设施选址问题中，x_{ij} 可以为介于 0 和 1 之间的分数，而在不可分的带硬容量限制的设施选址问题中，x_{ij} 只能为 $0-1$ 变量。

这样，FLPWSC可以刻画为以下规划。

$$
\begin{aligned}
\min\ & \sum_{i\in F} f_i y_i + \sum_{i\in F}\sum_{j\in D} c_{ij}x_{ij} \\
\text{s. t.}\ & \sum_{i\in F} x_{ij} \geq 1, \qquad \forall i\in F, \\
& x_{ij} \leq y_i, \qquad \forall i\in F, j\in D, \\
& \sum_{j\in D} x_{ij} \leq u_i y_i, \qquad \forall i\in F,
\end{aligned} \tag{1.10}
$$

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

$$y_i \in \mathbb{N}, \qquad \forall i \in F.$$

在上述规划中，第一组约束表示每个顾客都需要连接到一个设施上，第二组约束表示如果顾客 j 连接到设施 i 上，设施 i 必须开设，第三组约束表示连接到设施 i 上的顾客个数不能超过设施 i 的容量限制。

对于软容量设施选址问题，第一个近似比是 Shmoys等 [8] 利用线性规划舍入技巧得到的 9-近似算法，通过随机选取参数可以将近似比改进到 7.63。Chudak和 Shmoys[28] 改变在舍入过程中的聚类方法并且随机选取开设设施，对所有的设施容量都相同的特殊情况，得到 3-近似算法。Jain和 Vazirani[11] 利用对偶，将带容量限制的设施选址问题转化为无容量限制设施选址问题(UFLP)。通过 UFLP的原始-对偶算法可以得到 FLPWSC的 4-近似算法。 Arya等 [15] 利用局部搜索的思想，得到 3.72-近似算法。Jain等 [13] 利用对偶拟合的技巧得到了 3-近似算法。 Mahdian等 [29] 利用归约的方法，从 UFLP的双因子算法可以得到带软容量设施选址问题的近似比，从而得到 2.89-近似算法。Mahdian等 [30] 同样利用归约的方法，从 UFLP的双因子算法可以得到带软容量设施选址问题的双因子，从而得到 2-近似算法。

同时，规划(1.10)的整数间隙为2，即，整数最优值和分数最优值的比值不超过 2。事实上，考虑以下实例: 只有一个设施 i 以及 k 个顾客，设施 i 的容量为 $k-1$，开设费用为 $f_i = 1$。所有的连接费用都是 0(见图1.5)。

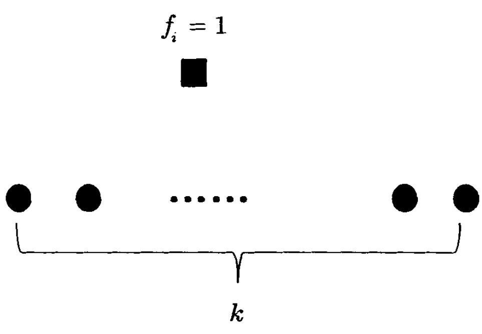


图 1.5 规划(1.10)整数间隙的实例

由于设施容量为 $k-1$，共有 k 个顾客，因此设施 i 需要开设 2 次，因此，规划(1.10)的整数最优值为 2。同时，规划(1.10)线性规划松弛的最优解为将设施

i 开设 $1+\frac{1}{k-1}$ 次，这样最优值为 $1+\frac{1}{k-1}$。这样，可以得到规划(1.10)的整数间隙为

$$\frac{2}{1+\frac{1}{k-1}}=\frac{2(k-1)}{k}.$$

当 k 趋向于无穷，整数间隙趋向于 2。在分析近似比时，一般不知道整数最优值，而是利用松弛问题的分数最优值作为下界。所以，任意基于规划(1.10)的线性规划松弛的算法，得到的近似比不可能低于 2。所以，文献[30]中提出的算法是基于线性规划规划(1.10)的松弛最紧的近似算法，近似比得不到改进。

可分的带硬容量限制的设施选址问题可以刻画为以下规划。

$$\begin{aligned}
\min \quad & \sum_{i\in F} f_i y_i+\sum_{i\in F}\sum_{j\in D} c_{ij}x_{ij} \\
\text{s. t.} \quad & \sum_{i\in F} x_{ij}\geq 1, \qquad \forall i\in F, \\
& x_{ij}\leq y_i, \qquad \forall i\in F, j\in F, \\
& \sum_{j\in D} x_{ij}\leq u_i y_i, \qquad \forall i\in F, \\
& x_{ij}\geq 0, \qquad \forall i\in F, j\in D, \\
& y_i\in\{0,1\}, \qquad \forall i\in F.
\end{aligned} \tag{1.11}$$

不可分的带硬容量的限制的设施选址问题可以刻画为以下规划。

$$\begin{aligned}
\min \quad & \sum_{i\in F} f_i y_i+\sum_{i\in F}\sum_{j\in D} c_{ij}x_{ij} \\
\text{s. t.} \quad & \sum_{i\in F} x_{ij}\geq 1, \qquad \forall i\in F, \\
& x_{ij}\leq y_i, \qquad \forall i\in F, j\in F, \\
& \sum_{j\in D} x_{ij}\leq u_i y_i, \qquad \forall i\in F, \\
& x_{ij}, y_i\in\{0,1\}, \qquad \forall i\in F.
\end{aligned} \tag{1.12}$$

由于不可分的带硬容量设施选址问题难度大，现在为止，考虑较多的为可分的带硬容量设施选址问题的近似算法。如果考虑可分的带硬容量限制的设施选址问题，其线性规划的整数间隙为无穷，即，整数最优值和分数最优值的比值趋向于无穷。考虑下面的实例(见图1.6): 给定 2 个设施，设施的容量均为 u，设施的费用分别为 $f_1=0$ 和 $f_2=1$。顾客的个数为 $u+1$，设施和顾客之间的连接费用均为 0。

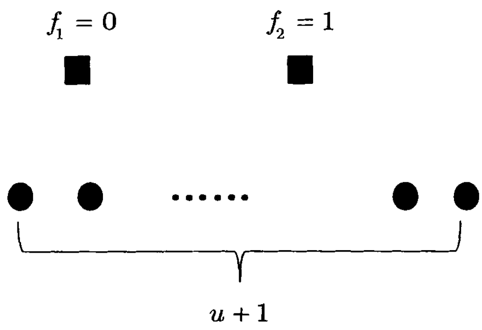


图 1.6 规划(1.11)整数间隙的实例。实心方形表示开设的设施。圆形表示顾客。

由于设施的容量为 u，顾客的个数为 $u+1$，因而两个设施都必须开设才能满足需求。所以，上述实例整数规划最优值为 2。现考察上述实例的分数最优解，即，

$$y_1 := 1, \qquad y_2 := \frac{1}{u}, \qquad x_{1j} := \frac{u}{u+1}, \qquad x_{2j} := \frac{1}{u+1}.$$

因此，分数最优值为 $\frac{1}{u}$。这样，可以得到规划(1.11)线性规划松弛的整数间隙为

$$\frac{1}{1/u} = u.$$

随着 u 趋向于无穷，整数间隙趋向于无穷。因此，若基于规划(1.11)的线性规划松弛设计近似算法，得到的近似比不可能是常数。因而，对硬容量设施选址问题一般是利用局部搜索的技巧设计近似算法。

Korupolu等 [31] 利用局部搜索这种启发式算法的思想，提出在当前解中增加一个设施和减少一个设施两种运算，分析出当所有设施的容量都相同时，即 $u_i \equiv u (\forall i \in F)$，给出 FLPWHC问题的第一个常数近似比的近似算法。对任意给定的 $\varepsilon > 0$，近似比为 $8+\varepsilon$。算法的运行时间是关于问题输入以及 $1/\varepsilon$ 的多项式。虽然文献[31]给出的近似比较大，但是给出了一个启发式算法的近似比，具有很大的理论意义。在文献[31]这一开创性的结果之后，有大量的文章利用局部搜索的技巧研究 FLPWHC。 Chudak和 Williamson[32] 简化了文献[31]中算法的分析过程，同时将近似比降低到 $6+\varepsilon$。对于设施容量不相同的情形， Pal等 [33] 对当前解引入增加一个设施，关闭一个设施开设一个设施集合，以及开设一个设施关闭一个设施集合的运算，分析得到 $(9+\varepsilon)$-近似算法。Zhang等 [34] 增加了开设一个设施集合关闭一个设施集合的运算，并且证明了这种运算是可以在多项

式时间内实现的，最终分析得到对任意 $\varepsilon>0$，FLPWHC存在一个 $(5.83+\varepsilon)$-近似算法。

1.2.3 随机设施选址问题

在 UFLP中，顾客的需求是固定的。但实际中所有的问题都可以看做是不确定的。随机设施选址问题(two-stage stochastic facility location problem，简称SFLP)给定每个顾客随机需求的分布，并且可以将这种概率分布转化为由可能需要被服务的顾客集合及其相应出现的概率。具体的说，在 SFLP中，选址过程分为两个阶段，第一阶段给出设施的集合 F 和潜在的顾客集合 D，设施 i 若在该阶段被开设则需支付 f_i^0 的开设费用。第二阶段，顾客集合 $D_s\subseteq D(s=1,2,\cdots,S)$ 以概率 q_s 出现。若顾客集合 $D_s\subseteq D$ 出现需要被服务，称场景 s 出现。在这一阶段，作为补充设施 i 可以以 f_i^s 的开设费用开设。在这一阶段开设的设施只能服务场景 s 中出现的顾客(即，D_S 中的顾客)。需要决策在每个阶段以及每个场景中开设哪些设施，使得开设费用和连接费用的总期望最小，即，

$$\min_{F_0,F_s\subseteq F,s=1,2,\cdots,S}\left\{\sum_{i\in F_0}f_i^0+\sum_{s=1}^{S}q_s\left[\sum_{i\in F_s}f_i^s+\sum_{j\in D_s}\min_{i\in F_0\cup F_s}c_{ij}\right]\right\}.$$

SFLP可以看做动态设施选址问题(dynamic facility location problem，简称DFLP)的特殊情况。DFLP是指给定设施的集合 F 和顾客集合 D，选址过程分为 T 个阶段，在每个时间段，顾客的需求和设施的开设费用都不相同。设施 i 在时间段 s 的开设费用为 f_i^s。顾客 $j\in D$ 在时间段 t 的需求量 d_j^t，并且这一时间段的顾客需求只能由这一时间段之前开设的设施服务。顾客和设施之间的单位连接费用为 c_{ij}。需要在每个时间段开设一些设施，将每个时间段的顾客连接到在该时间段之前的开设设施上，使得开设费用和连接费用之和最小，即，

$$\min_{F_t\subseteq F,t=1,2,\cdots,T}\sum_{t=1}^{T}\sum_{i\in F_t}f_i+\sum_{t=1}^{T}\sum_{j\in D}\min_{i\in\cup_{l=1}^{t}F_l}d_j^tc_{ij}.$$

SFLP问题是由 Ravi和 Sinha[35] 提出的，他们首先刻画出 SFLP的整数规划，再利用线性规划舍入的技巧得到 8-近似算法。接下来有一系列的改进的算法。Shmoys和 Swamy[36] 利用 SFLP的线性规划松弛将顾客分成 $S+1$ 个部分，分别指定其在第一阶段以及第二阶段各个场景中连接，构造出 $S+1$ 个 UFLP。对每个 UFLP运行 Mahdian等 [14] 的算法，可以得到 3.378-近似算法。Srinivasan[37] 基于文献[36]中的算法思想，对每个 UFLP运行对偶拟合的双因子算法 [13]，可以

得到 2.369-近似算法。 Bryka等 [38] 对 UFLP运行线性规划设施算法 [9]，将近似比改进到 2.2957。 Ye和 Zhang[39] 利用原始-对偶算法给出了 DFLP的 3-近似算法。运用贪婪增加设施的技巧，可以将近似比改进到 1.8526。这是到目前为止，这是 SFLP最好近似比的近似算法。

然而，常用的近似比是估计算法得到的费用期望和最优费用期望之间的关系。然而，这种估计是在平均意义下的，不能给出真实出现的场景解的质量。为此， Swamy[40] 引入单场景界。具体的说，单场景界是指由算法得到的解对应每个场景的费用和最优解中每个场景的费用。很显然，由单场景界可以得到近似比。Sirinivasan [37] 中在得到 2.369-近似算法基础上，利用原始问题的分数解做为连接费用上界得到 3.25-单场景界。Byrka等 [38] 利用文献[16]中的线性规划舍入技巧得到了 SFLP 单场景界为 2.4957 的近似算法。

1.2.4 带惩罚的设施选址问题

在 UFLP中，每个顾客都需要连接到开设的设施上。但实际上，有时顾客距离设施较远，服务该顾客需要花费的成本较大，这时设施可能选择不服务其。但是不服务就需要支付一定的惩罚费用，这就是带惩罚的设施选址问题。根据惩罚费用的形式，可以将带惩罚的设施选址问题分为带线性惩罚的设施选址问题和带次模惩罚的设施选址问题。具体的，给定设施的集合 F 和顾客的集合 D，设施 i 的开设费用 f_i，以及设施和顾客之间满足度量性的单位连接费用 c_{ij}。此外，还给定关于不被服务的顾客集合的惩罚费用函数 $h(\cdot)$。当 $h(\cdot)$ 是关于不被服务的顾客集合的线性函数时，则被称为带线性惩罚的设施选址问题(facility location problem with linear penalty，简称 FLPLP)。当 $h(\cdot)$ 是关于不被服务的顾客集合的单调不减次模函数时，则被称为带线性惩罚的设施选址问题(facility location problem with submodular penalty，简称 FLPSP)。一个集合函数 $h(\cdot)$ 是单调不减函数，如果给定集合 $T_1 \subseteq T_2$，有 $h(T_1) \leq h(T_2)$. 一个集合函数 $h(\cdot)$ 被称为次模函数，如果对于任意给定两个集合 T_1, T_2，有

$$h(T_1) + h(T_2) \geq h(T_1 \cap T_2) + h(T_1 \cup T_2).$$

此外，次模函数还有一个等价的定义，若 $T_1 \subseteq T_2$，以及任意 $i \notin T_2$，都有

$$h(T_1 \cup \{i\}) - h(T_1) \geq h(T_2 \cup \{i\}) - h(T_2).$$

事实上，次模函数表示的是边际费用随着集合的增大而减小。这样的性质常见

于经济，管理等领域的实际问题。显然，线性函数(上述不等式取恒等式)为特殊的次模函数。

带惩罚的设施选址问题可以描述为

$$\min_{F'\subseteq F, P'\subseteq D}\left\{\sum_{i\in F'} f_i+\sum_{j\in D\setminus P'}\min_{i\in F'} c_{ij}+h(P')\right\}.$$

在带惩罚的设施选址问题中，需要决策需要开设哪些设施，对每个设施 i 引入 $0-1$ 变量 y_i，$y_i=1$ 表示设施 i 开设，否则表示设施 i 不开设。同时，引入 $0-1$ 变量 x_{ij} 表示顾客 j 是否连接设施 i 上，$x_{ij}=1$ 表示顾客 j 连接到顾客 i 上，否则表示不连接. 对顾客的任一子集 $J\subseteq D$ 引入 $0-1$ 变量 z_J 表示顾客集合 J 中所有的顾客是否被惩罚，$z_J=1$ 表示顾客集合 J 被惩罚，否则表示顾客集合 J 不被惩罚。这样，可以刻画出带惩罚的设施选址问题的线性整数规划。

$$\begin{aligned}
\min\quad & \sum_{i\in F} f_i y_i+\sum_{i\in F}\sum_{j\in D} c_{ij}x_{ij}+\sum_{J\subseteq D} h(J)z_J \\
\text{s. t.}\quad & \sum_{i\in F} x_{ij}+\sum_{J:j\in J} z_j\geq 1, \qquad \forall i\in F, \\
& x_{ij}\leq y_i, \qquad \forall i\in F, j\in D, \\
& x_{ij}, z_J\in\{0,1\}, \qquad \forall i\in F, J\subseteq D.
\end{aligned}\tag{1.13}$$

在上述规划中，第一组约束表示对于如果顾客 j 没有连接到一个设施上，那么它一定包含在某个被惩罚的顾客集合中。第二组约束表示如果顾客连接到某个设施上，那么这个设施一定要开设。

对于带线性惩罚的设施选址问题，Charikar等 [41] 第一次引入这种选址变形，并给出基于原始-对偶技巧的 3-近似算法。Xu和 Xu[42] 给出了基于线性规划随机舍入技巧的 $2+2/e\cong 2.736$。Geunes等 [43] 指出对于经典问题基于线性规划的 α-近似算法，其相应的线性惩罚问题的近似比为 $\frac{1}{1-e^{-1/\alpha}}$。利用 Li[17] 给出的 UFLP的 1.488-近似算法，可以得到 FLPLP的 2.0435-近似算法。Xu和 Xu[44] 利用贪婪改进原始-对偶算法将近似比降低到 1.8526。利用 Li[17] 的思想，Li等 [45] 得到迄今为止 FLPLP最好近似比的 1.5148-近似算法。

Hayrapetyan等 [46] 首次提出将惩罚费用函数推广为次模函数，并提出若存在传统问题基于线性规划的 α-近似算法，则其相应次模惩罚问题存在一个 $(1+\alpha)$-近似算法。利用当前 UFLP最好的基于线性规划的 1.488-近似算法 [17]，则可得到 FLPSP存在一个 2.488-近似算法。该算法基于问题的凸松弛，需要调

用椭球算法求解，运行时间较高。Du等[47]利用原始-对偶的技巧得到了运行时间较低的3-近似算法。Li等[48]在Du等[47]算法的基础上，进行局部贪婪调整，可以将近似比改进到2.375。同时，Li等[45]将Hayrapetyan等[46]关于线性惩罚的结论，推广到次模惩罚。

Xu等[49]将次模惩罚引入到随机设施选址问题，利用原始-对偶技巧得到3-近似算法。

1.2.5 极大化的设施选址问题

设施选址问题除了极小化的形式，还有极大化的形式。具体的说，极大化的设施选址问题是给定设施的集合 F 和顾客的集合 D，每个设施 i 的开设费用为 f_i，顾客 j 连接到设施 i 即得到 d_{ij} 利润，这里的利润不一定满足三角不等式。需要开设一些设施使得顾客的服务得以满足，最终使得总利润减去设施开设费用的净利润最大。

$$\max_{F'\subseteq F}\left\{\sum_{j\in D}\max_{i\in F'}d_{ij}-\sum_{i\in F'}f_i\right\}.$$

对每个设施 i 是否开设引入 $0-1$ 变量 y_i，$y_i=1$ 表示设施 i 开设，否则设施 i 不开设。同时，对每个顾客 j 是否连接到设施 i 引入 $0-1$ 变量 x_{ij}，$x_{ij}=1$ 表示顾客 j 连接到设施 i 上获得利润，否则表示不连接。这样，可以得到极大化的设施选址问题的整数线性规划

$$\begin{aligned}\max\quad & \sum_{i\in F}\sum_{j\in D}d_{ij}x_{ij}-\sum_{i\in F}f_iy_i\\ \text{s. t.}\quad & \sum_{i\in F}x_{ij}\le 1, \qquad \forall j\in D,\\ & x_{ij}\le y_i, \qquad \forall i\in F, j\in D,\\ & x_{ij},y_i\in\{0,1\}, \qquad \forall i\in F, j\in D.\end{aligned} \tag{1.14}$$

在上述规划中，第一组约束表示每个顾客至多连接到一个设施上，第二组约束表示当顾客连接到一个设施，那么这个设施一定需要开设。由于极大化的设施选址问题的可行解目标值可能为负值，因此，在讨论可行解目标值为负数的优化问题时，需要将近似比 ρ 的定义修改为

$$\text{Cost}(Alg)-C^L\ge\rho(\text{Opt}-C^L),$$

其中 Cost(Alg) 表示算法所得解的费用。Opt 是极大化的设施选址问题的最优值。C^L 为极大化的设施选址问题的一个下界，即，将所有设施都开设，顾客连接到最近的开设设施，$C^L := \sum_{j\in D}\min_{i\in F} d_{ij} - \sum_{i\in F} f_i$。显然，$C^L$ 为极大化的设施选址问题的下界。

Cornuejols等 [50] 利用简单的贪婪算法可以得到 $(1-1/e)$-近似算法，其中 $1-1/e \cong 0.632$。 Aardal和 Sviridenko[51] 利用线性规划舍入的方法得到了 $(2\sqrt{2}-2)$-近似算法，其中 $2\sqrt{2}-2 \cong 0.828$。

对于 k-层设施选址问题也有极大化的形式。具体的说，极大化的设施选址问题是给定 k-层设施的集合 $F := F_1 \cup F_2 \cup \cdots \cup F_k$ 和顾客的集合 D，每个设施 $i \in F$ 的开设费用为 f_i，顾客 $j \in D$ 连接到设施路径集合中 $P = F_1 \times F_2 \times \cdots \times F_k$ 中的 $p = (i_1, i_2, \cdots, i_k)$ 即得到 d_{pj} 利润，需要开设一些设施使得顾客的服务得以满足，最终使得总利润减去设施开设费用的净利润最大，即，

$$\max_{P'\subseteq P}\left\{\sum_{j\in D}\max_{p\in P'} d_{pj} - \sum_{i\in \cup_{p\in P'} p} f_i\right\}.$$

对每个设施 i 是否开设引入 $0-1$ 变量 y_i，$y_i = 1$ 表示设施 i 开设，否则设施 i 不开设。同时，对每个顾客 j 是否连接到设施路径 p 引入 $0-1$ 变量 x_{pj}，$x_{pj} = 1$ 表示顾客 j 连接到路径 p 上，否则表示不连接。这样，可以得到极大化的设施选址问题的整数线性规划

$$\begin{aligned}
\max\quad & \sum_{p\in P}\sum_{j\in D} d_{pj}x_{pj} - \sum_{i\in F} f_i y_i \\
\text{s. t.}\quad & \sum_{i\in F} x_{pj} \le 1, \qquad \forall j \in D, \\
& \sum_{p:i\in p} x_{pj} \le y_i, \qquad \forall i \in F, j \in D, \\
& x_{pj}, y_i \in \{0,1\}, \qquad \forall i \in F, j \in D, p \in P.
\end{aligned} \tag{1.15}$$

Bumb[52] 将文献[51]中的方法推广到极大化的 2-层设施选址问题，得到 0.47-近似算法。这里，提到的近似比与极大化的单层设施选址问题类似，极大化的 k-层设施选址问题的近似比为 ρ，如果存在一个算法使得

$$\text{Cost}(Alg) - C^L \ge \rho(\text{Opt} - C^L),$$

其中 Cost(Alg) 是算法所得解的费用。Opt 是极大化的 k-层设施选址问题的最优值。C^L 是极大化的 k-层设施选址问题的一个下界，即，将所有设施都开设，顾

客连接到最近的开设设施路径，$C^L := \sum_{j\in D}\min_{p\in P} d_{pj} - \sum_{i\in F} f_i$。Zhang和 Ye[53] 在开设所有设施以及每层设施中费用最低的设施两个解中选择一个，最终可以得到 $\frac{1}{2}$-近似算法。

第三节 研究的问题及技巧

在本论文中，我们考察选址问题的几种重要的变形，包括两阶段随机设施选址问题，带线性惩罚的两阶段设施选址问题，2-层设施选址问题，以及带软容量限制的 k-层设施选址问题。我们分别利用设施选址问题常用的线性规划舍入，原始-对偶，以及组合的归约技巧设计近似算法。

1. 我们考虑两阶段随机设施选址问题的单场景界。在随机设施选址问题中，需要决策在第一阶段和第二阶段的每个场景开设设施，以及，将顾客连接到第一阶段和第二阶段相应场景中开设设施上。首先利用随机设施选址问题的线性规划松弛，对每个顾客-场景对在第一阶段和其对应场景中选择一类作为自己的邻域。再通过选择中心以及邻域的相交性对所有顾客-场景对进行聚类，再在每个聚类中随机开设设施。最终，分析得到算法的可行解对于每个场景的费用不超过 2.3613 倍分数最优值中该场景的费用。
2. 我们考虑随机设施选址问题的一个重要变形带线性惩罚的两阶段随机设施选址问题。和第二章一样，考虑带线性惩罚的随机设施选址问题的单场景界。首先利用线性规划松弛的解，设定一个阈值确定被惩罚的顾客。对剩余的顾客，首先利用带线性惩罚随机设施选址问题的线性规划松弛，将所有潜在的顾客分成 $S+1$ 个集合，分别指定其被第一阶段和第二阶段 S 个场景中开设的设施服务。每个集合和相应的设施集合构成一个无容量限制设施选址问题的实例。通过无容量限制的设施选址问题的线性规划舍入算法，得到带线性惩罚的随机设施选址问题的可行解。最终，分析得到算法的可行解对于每个场景的费用不超过 3.0294 倍分数最优值中该场景的费用。
3. 我们考虑 2-层设施选址问题。首先得到 2-层设施选址问题基于集合覆盖的线性整数规划。基于该规划，对任意的常数 $\varepsilon>0$ 利用原始-对偶技巧得到 $3(1+\varepsilon)$-近似算法。一般的原始-对偶算法需要利用对偶问题的多项式时间可解的分离神谕(separation oracle)。但是 2-层设施选址问

题的分离神谕可以看做极大化的设施选址问题，是 NP-困难的。因而，利用近似求解极大化的设施选址问题，近似的求解分离神谕。为保证可行性，离散化时间使得对偶增长的过程只在离散的点上取得。接着，利用贪婪增加一个树费用是否可以降低的技巧，可以将近似比降低为 $2.172(1+\varepsilon)^2$。

4. 我们考虑带软容量限制的 k-层设施选址问题。首先得到 k-层设施选址问题和带软容量限制的 k-层设施选址问题之间双因子的递推关系。之后，利用线性规划舍入技巧得到 k 层设施选址问题的双因子算法，最终得到 5.5053-近似算法。

第四节 论文的结构安排

第二章给出了两阶段随机设施选址问题的线性规划舍入 2.3613 单场景界。第三章给出了带线性惩罚的两阶段随机设施选址问题的基于线性规划舍入的 3.0294 单场景界。第四章对 2-层设施选址问题给出了基于近似分离神谕的原始-对偶 3-近似算法。第五章基于双因子归约技巧给出了带软容量限制的设施选址问题的 5.5053-近似算法。

此页不缺内容

第二章 两阶段随机设施选址问题的单场景界

第一节 问题介绍

在 UFLP中，所有的参数都是确定性的。但是在实际情况中，在选址之前顾客的信息是未知的。因此，本章研究 UFLP一类重要的变形，两阶段随机设施选址问题(two-stage stochastic facility location problem，简记 SFLP)。SFLP过程分为两个阶段，在第一阶段给定设施的集合 $F(m=|F|)$，潜在的顾客的集合 $D(n=|D|)$，以及设施和顾客之间满足三角不等式的单位连接费用 $c: F\times D\to\mathbb{R}_+$。在这一阶段还给定可能出现顾客的分布函数，即给定顾客集合 $D_s\subseteq D(s=1,2,\cdots,S)$为需要服务的顾客集合，及其出现的概率 q_s。可以在第一阶段以 f_i^0 的费用开设设施 i，并且这一阶段开设的设施可以服务之后出现的任何顾客。在第二阶段，场景 $s(s=1,2,\cdots,S)$实现，为了降低连接费用，作为补充设施 i 可以以费用 f_i^s 开设。一般来说，$f_i^s\geq f_i^0$。在这一阶段开设的设施只能为这一阶段实现的顾客提供服务。最终，需要分别在第一阶段和第二阶段分别开设一些设施使得最终开设费用和连接费用的总和最小，即，

$$\min_{F_0,F_s\subseteq F}\left\{\sum_{i\in F_0}f_i^0+\sum_{s=1}^{S}q_s\left(\sum_{i\in F_s}f_i^s+\sum_{j\in D_s}\min_{i\in F_0\cup F_s}c_{ij}\right)\right\}.$$

为了方便起见，记出现在 $D_t(t=1,2,\cdots,S)$中的顾客 j 为 (j,t)，称为顾客-场景对，将第一阶段的设施以及第二阶段各个场景中的设施 i 都称为设施-场景对，分别记为 $(i,0)$ 以及 $(i,s)(s=1,2,\cdots,S)$. 设施-场景对的集合记为 $\mathscr{F}:=\{(i,s):i\in F,s=0,1,2,\cdots,S\}$ 和顾客-场景对的集合记为 $\mathscr{D}:=\{(j,t):j\in D_t,t=1,2,\cdots,S\}$。此外，重新定义连接费用

$$c_{ij}^{st}:=\begin{cases}c_{ij}, & \text{如果 } s=0 \text{ 或者 } s=t,\\ +\infty, & \text{否则}.\end{cases}$$

令 $q_0:=1$，这样，SFLP可以看做开设一些设施-场景对，以服务顾客-场景对，最终使得开设费用和连接费用总和最小(见图2.1)。设施-场景对 (i,s) 和顾客-场景对 (j,t) 之间的单位连接费用为 c_{ij}^{st}，每个设施-场景对 (i,s) 的开设费用为

$q_s f_i^s$，每个顾客-场景对 (j,t) 的需求为 q_t。因此，SFLP可以刻画为以下整数规划。

$$
\begin{aligned}
\min \quad & \sum_{(i,s)\in\mathscr{F}} q_s f_i^s y_i^s + \sum_{(i,s)\in\mathscr{F}} \sum_{(j,t)\in\mathscr{D}} q_t c_{ij}^{st} x_{ij}^{st} \\
\text{s. t.} \quad & \sum_{(i,s)\in\mathscr{F}} x_{ij}^{st} \geq 1, \qquad \forall (j,t)\in\mathscr{D}, \\
& x_{ij}^{st} \leq y_i^s, \qquad \forall (i,s)\in\mathscr{F},(j,t)\in\mathscr{D}, \\
& x_{ij}^{st}, y_i^s \in \{0,1\}, \qquad \forall (i,s)\in\mathscr{F},(j,t)\in\mathscr{D}.
\end{aligned} \tag{2.1}
$$

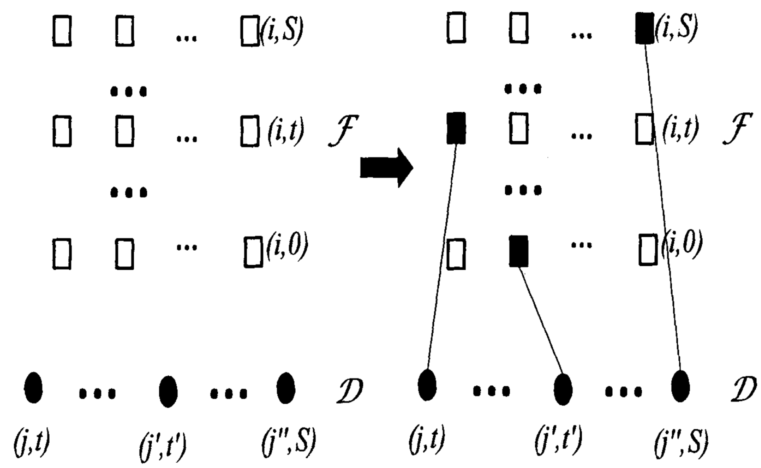


图 2.1 SFLP

在上述规划中，变量 y_i^s 表示设施-场景对 (i,s) 是否开设，$y_i^s=1$ 表示设施-场景对 (i,s) 开设，否则表示不开设。变量 x_{ij}^{st} 表示顾客-场景对 (j,t) 是否连接到设施-场景对 (i,s) 上，$x_{ij}^{st}=1$ 表示顾客-场景对 (j,t) 连接到设施-场景对 (i,s) 上，否则表示不连接。第一组约束表示每个顾客-场景对必须被某个设施-场景服务。第二组约束表示如果顾客-场景 (j,t) 连接到设施-场景对 (i,s)，那么设施-场景对 (i,s) 一定要开设。规划(2.1)的线性规划松弛是

$$
\begin{aligned}
\min \quad & \sum_{(i,s)\in\mathscr{F}} q_s f_i^s y_i^s + \sum_{(i,s)\in\mathscr{F}} \sum_{(j,t)\in\mathscr{D}} q_t c_{ij}^{st} x_{ij}^{st} \\
\text{s. t.} \quad & \sum_{(i,s)\in\mathscr{F}} x_{ij}^{st} \geq 1, \qquad \forall (j,t)\in\mathscr{D},
\end{aligned} \tag{2.2}
$$

$$x_{ij}^{st} \le y_i^s, \qquad \forall (i,s) \in \mathscr{F}, (j,t) \in \mathscr{D},$$
$$x_{ij}^{st}, y_i^s \ge 0, \qquad \forall (i,s) \in \mathscr{F}, (j,t) \in \mathscr{D}.$$

假设规划(2.2)的分数最优解为 $(\bar{x},\bar{y})$。在第一章中，我们介绍近似比是研究算法得到的解质量估计。但是在 SFLP中算法得到的费用是实现场景的费用期望，这种是场景平均意义下的费用。因此，近似比不能估计出算法得到实际实现场景的费用。这样，考虑到单场景界。对于 SFLP的算法 ALG，令 ALG_t 表示由算法 ALG 得到的场景 t 对应的费用。如果

$$\frac{ALG_t}{\sum\limits_{i\in F}\left[(f_i^0\bar{y}_i^0 + f_i^t\bar{y}_i^t) + \sum\limits_{s=0,t}\sum\limits_{j\in C_t} c_{ij}^{st}\bar{x}_{ij}^{st}\right]} \le \gamma,$$

则称算法 ALG 的单场景界为 γ。

在本节中，对于规划(2.2)的解 (x,y)，顾客-场景对 (j,t) 以及设施-场景对集合 A，将 $\sum_{(i,s)\in A} c_{ij}^{st}x_{ij}^{st}/\sum_{(i,s)\in A} x_{ij}^{st}$ 称为在解 (x,y) 下 (j,t) 到 A 的平均距离，也称为 (j,t) 分数连接到 A 的距离。同时，将在解 (x,y) 下 (j,t) 到 $\mathscr{F}$ 的平均距离简记为在解 (x,y) 下 (j,t) 的平均距离。

在分数最优解 $(\bar{x},\bar{y})$ 的基础上，定义顾客-场景对 (j,t) 的分数连接费用 $\tilde{C}_{j,t} := \sum_{(i,s)\in\mathscr{F}} c_{ij}^{st}\bar{x}_{ij}^{st}$。不失一般性，假设 $(\bar{x},\bar{y})$ 是一个完全解，即，如果 $\bar{x}_{ij}^{st} > 0$，那么 $\bar{x}_{ij}^{st} = \bar{y}_i^s$。如果 $(\bar{x},\bar{y})$ 不是一个完全解，则可以通过分割的技巧构造一个完全解。具体的说，如果存在一个设施-场景对 (i,s) 以及不满足完全解的顾客-场景对集合 V，即，对任意 $(j,t)\in V$ 都满足 $\bar{x}_{ij}^{st} < \bar{y}_i^s$。可以对 V 中的顾客-场景对按 $\bar{x}_{ij}^{st}$ 排序，不妨设，$\bar{x}_{i1}^{st_1} \le \bar{x}_{i2}^{st_2} \le \cdots \le \bar{x}_{iv}^{ss_v}$。将设施-场景对 (i,s) 分割成 (i_1,s) 和 (i_2,s) 使得 $\bar{y}_{i_1}^s = \bar{x}_{i1}^{st_1}$ 和 $\bar{y}_{i_2}^s = \bar{y}_i^s - \bar{y}_{i_1}^s$。同时，对任意 $(j,t)\in V$ 做同样的分割，即，令 $\bar{x}_{i_1j}^{s\,t} = y_{i_1}^s$ 以及 $\bar{x}_{i_2j}^{s\,t} = \bar{x}_{ij}^{st} - \bar{x}_{i_1j}^{s\,t}$。更新 $V := V\setminus\{(j_1,t_1)\}$。重复以上过程直到所有的设施-场景对以及顾客-场景对都满足完全解的定义。将这个过程称为分割技巧。

第二节 线性规划舍入算法

在本节，我将介绍基于线性规划舍入的随机设施选址问题的算法。总体来说，这类算法的基本思想是首先得到线性规划的分数最优解，通常需要对这个分数解进行一定的调整，之后再利用调整后的分数解舍入成整数可行解。因此，

先将分数最优完全解 $(\tilde{x},\tilde{y})$ 按比例提高 $\alpha > 2$ 倍，得到新的解 $(\bar{x},\bar{y}) := (\alpha\tilde{x},\alpha\tilde{y})$，即，$\bar{x}_{ij}^{st} = \alpha\tilde{x}_{ij}^{st}$ 和 $\bar{y}_i^s = \alpha\tilde{y}_i^s$。可以假设对每个 $(i,s) \in \mathscr{F}$ 和 $(j,t) \in \mathscr{D}$，$\bar{x}_{ij}^{st} \leq 1$ 和 $\bar{y}_i^s \leq 1$。如果不满足以上条件，类似的，利用分割的技巧得到满足条件的解。

为了叙述方便，首先引入一些记号。对每个顾客-场景对 (j,t)，

- $F_{j,t}$: (j,t) 分数连接的设施-场景对集合，即，$F_{j,t} := \left\{(i,s) : \bar{x}_{ij}^{st} > 0\right\}$;
- $\bar{C}_{jt}$: (j,t) 在解 $(\bar{x},\bar{y})$ 下的平均距离，即，$\bar{C}_{jt} := \frac{1}{\alpha}\sum_{(i,s)} c_{ij}^{st}\bar{x}_{ij}^{st}$;
- $\tilde{C}_{jt}$: (j,t) 在解 $(\tilde{x},\tilde{y})$ 下的平均距离，即，$\tilde{C}_{jt} := \sum_{(i,s)} c_{ij}^{st}\tilde{x}_{ij}^{st}$;
- 按照到 (j,t) 的连接费用递增顺序，分别重新标记所有第一阶段和第二阶段场景 t 的设施-场景对。假设这个顺序是 $c_{1j}^{0t} \leq \cdots \leq c_{ij}^{0t} \leq \cdots \leq c_{nj}^{0t}$，同样有 $c_{1j}^{tt} \leq \cdots \leq c_{ij}^{tt} \leq \cdots \leq c_{nj}^{tt}$。

注意到 $\sum_{(i,s)} \bar{x}_{ij}^{st} = \alpha \geq 2$，这就意味着 $\sum_{(i,0)} \bar{x}_{ij}^{0t} \geq 1$ 和 $\sum_{(i,t)} \bar{x}_{ij}^{tt} \geq 1$ 至少有一个不等式成立。

- 如果 $\sum_{i=1}^{n} \bar{x}_{ij}^{0t} \geq 1$，令 $i_0^* := \min\left\{i' : \sum_{i=1}^{i'} \bar{x}_{ij}^{0t} = 1\right\}$ (如果有需要，当等式不成立时应用分割的技巧)，令 $d_{j,t}^0 := c_{i_0^*j}^{0\,t}$，以及 $N_{j,t}^0 := \{(i,0) : 1 \leq i \leq i_0^*\}$。否则，令 $d_{j,t}^0 := +\infty$。
- 如果 $\sum_{i=1}^{n} \bar{x}_{ij}^{tt} \geq 1$，令 $i_t^* := \min\left\{i' : \sum_{i=1}^{i'} \bar{x}_{ij}^{tt} = 1\right\}$ (如果有需要，当等式不等式时应用分割的技巧)，令 $d_{j,t}^t := c_{i_t^*j}^{t\,t}$ 和 $N_{j,t}^t := \{(i,t) : 1 \leq i \leq i_t^*\}$。否则，令 $d_{j,t}^t := +\infty$。

因此，为了得到每个顾客-场景对 (j,t) 的邻域集 $N(j,t)$，计算 $d_{j,t}^0$ 和 $d_{j,t}^t$。

- 如果 $d_{j,t}^0 \leq d_{j,t}^t$，定义 $d_{j,t} := d_{j,t}^0$ 和 $N(j,t) := N_{j,t}^0$。
- 如果 $d_{j,t}^0 > d_{j,t}^t$，定义 $d_{j,t} := d_{j,t}^t$ 和 $N(j,t) := N_{j,t}^t$。

最终，定义 (j,t) 到 $N(j,t)$ 的平均距离 $\Delta_{j,t} := \sum_{(i,s)\in N(j,t)} c_{ij}^{st}\bar{x}_{ij}^{st}$。

在我们的算法中，首先迭代通过选取中心对顾客-场景对聚类，在每个聚类中随机开设设施，再将每个顾客连接到开设的最近设施。

算法 2.1

步 0. 令 U 是未连接的顾客-场景对的集合，$\mathscr{D}'$ 是中心的集合，和 $\mathscr{F}$ 是开设的设施-场景对的集合。

步 1. 初始时，所有的顾客-场景对都是未连接，均不是中心。并且所有的设施-场景对都没有开设。也就是说，$U := \mathscr{D}$，$\mathscr{D}' := \emptyset$，和 $\mathscr{F} := \emptyset$。

步 1.1. 在 U 中选取 $d_{j,t} + \Delta_{j,t}$ 最小的顾客-场景对做为中心，即，找出 (j',t')

满足

$$(j',t') := \arg\min_{(j,t)\in U}\{d_{j,t}+\Delta_{j,t}\}.$$

将 (j',t') 放入中心集合中, 即, 更新 $\mathscr{D}' := \mathscr{D}' \cup \{(j',t')\}$。

步 *1.2.* 根据 $N(j',t')$, 在 U 中找出与 $N(j',t')$ 分数相连的顾客-场景对集合, 记为 $B(j',t')$。也就是, 令

$$B(j',t') := \{(j,t) \in U : N(j,t) \cap N(j',t') \neq \emptyset\}.$$

同时, 称 $B(j',t')$ 中的顾客-场景对已连接, 即, 更新 $U := U \setminus B(j',t')$。将 $\{(j',t'), N(j',t'), B(j',t')\}$ 记为以 (j',t') 为中心的聚类(见图2.2)。

步 *1.3.* 重复步*1.1*以及步*1.2*直到 $U=\emptyset$。

步 *2.* 在每个聚类 $\{(j',t'),N(j',t'),B(j',t')\}$ 中, 以 $\bar{x}_{ij'}^{st'}$ 概率相关的开设 $N(j',t')$ 中一个设施-场景对 (i,s)。以 $\bar{y}_i^s$ 的概率独立的开设所有不在任何一个聚类中的设施-场景对。

步 *3.* 连接顾客-场景对 (j,t) 到距离其最近的第一阶段或者第二阶段场景 t 中开设设施。

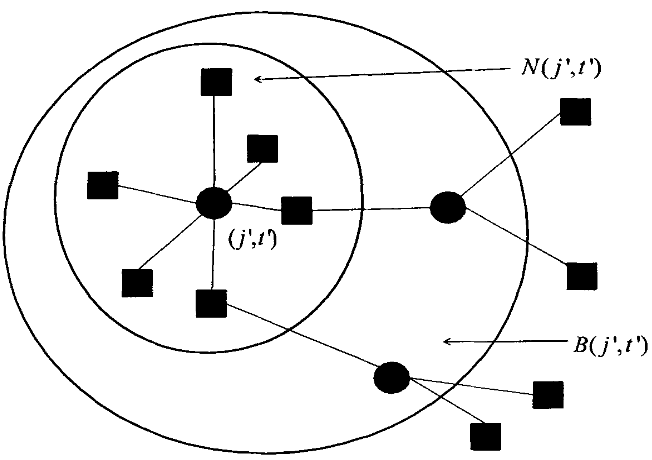


图 2.2 以 (j',t') 为中心的聚类

第三节 分析

在本节，我们将分析算法2.1所得解对每个场景的质量。为了估计连接费用期望，需要分别估计 $d_{j,t}$ 和 $\Delta_{j,t}$。首先，文献[38]已经给出了 $d_{j,t}$ 的上界。为了完整性，本文还是给出如下证明。

引理 2.2 ([38]) 对任意顾客-场景对 (j,t)，

$$d_{j,t} \leq \frac{\alpha}{\alpha-2}\tilde{C}_{j,t}.$$

证明. 我们只证明 $d_{j,t}=d^0_{j,t}$ 的情况，另一种情况可类似得到。注意到如果 $d_{j,t}=d^0_{j,t}$ 意味着 $\sum_{i=1}^{i_0^*}\bar{x}^{0t}_{ij}=1$。考虑以下两种情况。

情况1. $\sum_{i=1}^{n}\bar{x}^{tt}_{ij}\geq 1$。

从 $N^0_{j,t}$ 和 $N^t_{j,t}$ 的定义，可以得到 $\sum_{(i,s)\in\mathscr{F}\setminus(N^0_{j,t}\cup N^t_{j,t})}\bar{x}^{st}_{ij}=\alpha-2$。那么，顾客-场景对 (j,t) 到设施-场景集合 $\mathscr{F}\setminus(N^0_{j,t}\cup N^t_{j,t})$ 的连接费用比 $d_{j,t}$ 高。因此，$d_{j,t}$ 比 (j,t) 到 $\mathscr{F}\setminus(N^0_{j,t}\cup N^t_{j,t})$ 的平均距离小，则有，

$$\begin{aligned}
d_{j,t} &\leq \frac{\sum_{(i,s)\in\mathscr{F}\setminus(N^0_{j,t}\cup N^t_{j,t})}c^{st}_{ij}\bar{x}^{st}_{ij}}{\sum_{(i,s)\in\mathscr{F}\setminus(N^0_{j,t}\cup N^t_{j,t})}\bar{x}^{st}_{ij}}\\
&= \frac{\alpha}{\alpha-2}\sum_{(i,s)\in\mathscr{F}\setminus(N^0_{j,t}\cup N^t_{j,t})}c^{st}_{ij}\tilde{x}^{st}_{ij}\\
&\leq \frac{\alpha}{\alpha-2}\sum_{(i,s)}c^{st}_{ij}\tilde{x}^{st}_{ij}=\frac{\alpha}{\alpha-2}\tilde{C}_{j,t}.
\end{aligned}$$

第二个等式成立由 $\bar{x}$ 是由 $\tilde{x}$ 按比例放缩 α 倍得到。第三个不等式由于求和从部分设施-场景对集合放大为整个设施-场景对集合得到。第四个等式由 $\tilde{C}_{j,t}$ 的定义得到。

情况 2. $\sum_{i=1}^{n}\bar{x}^{tt}_{ij}<1$。

从 $N(j,t)$ 的定义可知，有 $\sum_{(i,0)\in\mathscr{F}\setminus N(j,t)}\bar{x}^{0t}_{ij}\geq\alpha-2$。注意到，顾客-场景 (j,t) 到设施-场景对中 $\mathscr{F}\setminus N(j,t)$ 的连接费用比 $d_{j,t}$ 高。因此，$d_{j,t}$ 比 (j,t) 到设施-场景对集合 $\mathscr{F}\setminus N(j,t)$ 的平均距离小, 则有

$$\begin{aligned}
d_{j,t} &\leq \frac{\sum_{(i,0)\in\mathscr{F}\setminus N(j,t)}c^{0t}_{ij}\bar{x}^{0t}_{ij}}{\sum_{(i,0)\in\mathscr{F}\setminus N(j,t)}\bar{x}^{0t}_{ij}}\\
&\leq \frac{\alpha}{\alpha-2}\sum_{(i,0)\in\mathscr{F}\setminus N(j,t)}c^{0t}_{ij}\tilde{x}^{0t}_{ij}\\
&\leq \frac{\alpha}{\alpha-2}\tilde{C}_{j,t}.
\end{aligned}$$

第二个不等式成立因为 $\tilde{x}$ 是由 $\bar{x}$ 按比例放缩 α 倍得到。
综合以上两种情况，引理即证。 □

其次，估计 $\Delta_{j,t}$ 的费用。

引理 2.3 对每个顾客-场所对 (j,t)，有

$$\Delta_{j,t} \le \alpha\tilde{C}_{j,t}.$$

证明. 注意到 $\tilde{C}_{j,t}$ 表示在解 $(\tilde{x},\tilde{y})$ 下 (j,t) 的平均距离。因此，

$$\begin{aligned}\tilde{C}_{j,t} = \sum_{(i,s)} c_{ij}^{st}\tilde{x}_{ij}^{st} &\ge \sum_{(i,s)\in N(j,t)} c_{ij}^{st}\tilde{x}_{ij}^{st} \\ &= \frac{1}{\alpha}\sum_{(i,s)\in N(j,t)} c_{ij}^{st}\bar{x}_{ij}^{st} \\ &= \frac{1}{\alpha}\Delta_{j,t}.\end{aligned}$$

由于 $\bar{x}$ 的构造可得第三个等式成立。由 $\Delta_{j,t}$ 的定义可得最后一个不等式。
所以，引理成立。 □

给定设施场景对集合 A，令 $\mathscr{A}$ 表示在集合 A 中至少有一个设施-场景对开设的事件。下面给出顾客-场景对连接到 A 中开设的最近设施的连接费用的期望。首先给出一个结论。

事实 2.4 ([54]) 如果序列 $\{a(l)\}_{l=1}^k$，$\{c(l)\}_{l=1}^k$，和 $\{w(l)\}_{l=1}^k$ 中 $\{a(l)\}$ 是单调不增的，$\{c(l)\}$ 是单调不减的，那么，

$$\frac{\sum_{l=1}^k a(l)c(l)w(l)}{\sum_{l=1}^k a(l)w(l)} \le \frac{\sum_{l=1}^k c(l)w(l)}{\sum_{l=1}^k w(l)}.$$

基于事实2.4，可以得到

引理 2.5 ([16,38]) 对每个顾客-场景对 (j,t)，在设施-场景对集合 A 中至少有一个设施-场景对开设的条件下，(j,t) 连接到最近的开设设施-场景对的期望为

$$E\left[\min_{(i,s)\in A:(i,s)\text{ 是开设的}} c_{ij}^{st}|\mathscr{A}\right] \le c\{(j,t),A\},$$

其中 $c\{(j,t),A\}$ 表示 (j,t) 到 A 在解 $(\bar{x},\bar{y})$ 下的平均距离，即，

$$c\{(j,t),A\} := \frac{\sum_{(i,s)\in A} c_{ij}^{st}\bar{x}_{ij}^{st}}{\sum_{(i,s)\in A}\bar{x}_{ij}^{st}}.$$

证明. 不失一般性，可以假设 A 中所有的设施-场景对均为第一阶段或者第二阶段 t 场景中。由聚类的过程，可以将 A 划分为若干个子集 $A_1,A_2,\cdots,A_d$，使得每个子集中的设施-场景对在同一个聚类中。注意到，在不同的子集中的设施-场景对的开设是互相独立的，并且在每个子集中至多有一个设施-场景对开设。因此，可将设施-场景对集合 A_l 转化成一个设施-场景对 (l,t_l)。顾客 j 到这个新的设施场景对的距离为在解 $(\bar{x},\bar{y})$ 下的平均距离，即，

$$c_{lj}^{t_l t} := \frac{\sum_{(i,t_l)\in A_l} c_{ij}^{t_l t}\bar{x}_{ij}^{t_l t}}{\sum_{(i,t_l)\in A_l}\bar{x}_{ij}^{t_l t}} = \frac{\sum_{(i,t_l)\in A_l} c_{ij}^{t_l t}\bar{y}_i^{t_l}}{\sum_{(i,t_l)\in A_l}\bar{y}_i^{t_l}}.$$

并且，(l,t_l) 开设的概率为 $\bar{y}_l^{t_l} := \sum_{(i,t_l)\in A_l}\bar{y}_i^{t_l}$。

现在，可以假设设施-场景对集合 $A := \{(1,t_1),(2,t_2),\cdots,(d,t_d)\}$，且满足 $c_{1j}^{t_1 t} \le c_{2j}^{t_2 t} \le \cdots \le c_{dj}^{t_d t}$。注意到，$t_1,t_2,\cdots,t_d \in \{0,t\}$，也就是，所有 A 中设施-场景对在第一阶段或者第二阶段 t 场景中。因此，由事实2.4可以得到

$$\begin{aligned}
E\left[\min_{(i,t)\in A:(i,t)\text{ 开设}} c_{ij}|\mathscr{A}\right] &= \frac{c_{1j}^{t_1 t}\bar{y}_1^{t_1} + c_{2j}^{t_2 t}\bar{y}_2^{t_2}(1-\bar{y}_1^{t_1}) + \cdots + c_{dj}^{t_d t}\bar{y}_d^{t_d}\Pi_{l=1}^{d-1}(1-\bar{y}_d^{t_d})}{\bar{y}_i^{t_1} + \bar{y}_2^{t_2}(1-\bar{y}_1^{t_1}) + \cdots + \bar{y}_d^{t_d}\Pi_{l=1}^{d-1}(1-\bar{y}_d^{t_d})} \\
&\le \frac{c_{1j}^{t_1 t}\bar{y}_1^{t_1} + c_{2j}^{t_2 t}\bar{y}_2^{t_2} + \cdots + c_{dj}^{t_d t}\bar{y}_d^{t_d}}{\bar{y}_i^{t_1} + \bar{y}_2^{t_2} + \cdots + \bar{y}_d^{t_d}} \\
&= c\{(j,t),A\}.
\end{aligned}$$

引理即证。 □

现在为止，可以估计连接费用的期望了。

引理 2.6 *对每个顾客-场景对 (j,t)，令 $C_{j,t}$ 表示由算法2.1得到的连接费用。有*

- *如果 (j,t) 是中心，那么其连接费用的期望为 $E[C_{j,t}] \le \alpha\tilde{C}_{j,t}$;*
- *如果 (j,t) 不是中心，那么其连接费用的期望为 $E[C_{j,t}] \le \left(1+e^{-\alpha}\ \frac{\alpha^2-\alpha+2}{\alpha-2}\right)\tilde{C}_{j,t}$。*

证明. 由算法过程中，顾客可以分为中心和非中心。因此，分别考虑以下两类情况.

- 如果 (j,t) 是中心，存在一个 $N(j,t)$ 中开设的设施-场景对 (i,s). 那么，由引理2.3有

$$E[C_{j,t}] = \sum_{(i,s)\in N(j,t)} c_{ij}^{st}\bar{x}_{ij}^{st} = \Delta_{j,t} \le \alpha\tilde{C}_{j,t}.$$

- 如果 (j,t) 不是中心，考虑 (j,t) 所在的聚类，令 (j',t') 为该聚类的中心。

情况1. $F_{j,t}$ 中存在一个开设的设施-场景对。

由算法 2.1，$F_{j,t}$ 中每个设施-场景对的概率为 $\bar{y}_i^s$。由于 $(\bar{x},\bar{y})$ 是完全解，则 $\bar{y}_i^s$ 即为 $\bar{x}_{ij}^{st}$。那么，可以得到 $F_{j,t}$ 中存在一个开设的设施-场景对这一事件的概率至少

$$1-\Pi_{(i,s)\in F_{j,t}}\left(1-\bar{x}_{ij}^{st}\right)\geq 1-\Pi_{(i,s)\in F_{j,t}}e^{-\bar{x}_{ij}^{st}}=1-e^{-\sum_{(i,s)\in F_{j,t}}\bar{x}_{ij}^{st}}=1-e^{-\alpha}.$$

由常用不等式 $1-z\leq e^{-z}$ 可得第一个不等式成立。引理2.3和2.5意味着 (j,t) 的连接费用期望至多

$$\frac{\sum_{(i,s)\in N(j,t)}c_{ij}^{st}\bar{x}_{ij}^{st}}{\sum_{(i,s)\in N(j,t)}\bar{x}_{ij}^{st}}\leq\frac{\Delta_{j,t}}{\alpha}\leq\tilde{C}_{j,t}.$$

情况2. $F_{j,t}$ 中的设施-场景对都不开设。

这一事件发生的概率不超过 $e^{-\alpha}$。利用三角不等式(参见图2.3)，接下来分以下两类情况讨论。

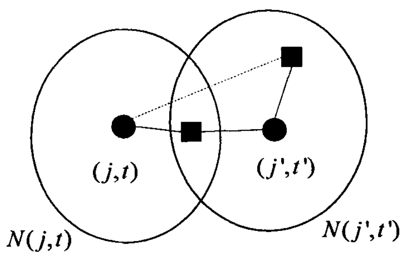


图 2.3 将 (j,t) 连接到 (j',t') 的邻居 $N(j',t')$ 上

情况2.1 $N(j,t)\cap N(j',t')$ 中存在一个设施-场景对 (i,s) 满足 $c_{ij'}^{st'}\leq\Delta_{j't'}$。

假设 $N(j',t')$ 中开设的设施-场景对为 (i',s)。那么，由三角不等式知，$C_{j,t}$ 至多为

$$c_{i'j}^{st}\leq c_{i'j'}^{st'}+c_{ij'}^{st'}+c_{ij}^{st}\leq d_{j't'}+\Delta_{j't'}+d_{j,t}\leq 2d_{j,t}+\Delta_{jt}.$$

由 (j',t') 是中心可得最后一个不等式成立。

情况2.2 $N(j,t)\cap N(j',t')$ 中任何一个设施-场景对 (i,s) 都满足 $c_{ij'}^{st'}>\Delta_{j't'}$。

注意到 (j',t') 到集合 $N(j',t')$ 在解 $(\bar{x},\bar{y})$ 下的平均距离为 $\Delta_{j',t'}$，有

$$\sum_{(i,s)\in N(j',t')\backslash N(j,t)} c_{ij'}^{st'}\bar{x}_{ij'}^{st'} \leq \Delta_{j't'}.$$

因此，由三角不等式可以得到 $C_{j,t}$ 的条件期望不超过 $2d_{j,t}+\Delta_{jt}$。综上，可得到 $C_{j,t}$ 的条件期望的上界为 $2d_{j,t}+\Delta_{jt}$。由引理2.2-2.3以及 $\alpha>2$，有

$$2d_{j,t}+\Delta_{j,t} \leq 2\frac{\alpha}{\alpha-2}\tilde{C}_{j,t}+\alpha\tilde{C}_{j,t}=\frac{\alpha^2}{\alpha-2}\tilde{C}_{j,t}.$$

最终，可以得到 $C_{j,t}$ 的期望为

$$\begin{aligned} E[C_{j,t}] &\leq (1-e^{-\alpha})\tilde{C}_{j,t}+e^{-\alpha}\frac{\alpha^2}{\alpha-2}\tilde{C}_{j,t} \\ &= \left(1+\frac{\alpha^2-\alpha+2}{\alpha-2}e^{-\alpha}\right)\tilde{C}_{j,t}. \end{aligned}$$

定理得证。 □

最后，由引理2.6可以得到我们的单场景界。

定理 2.7 *由算法2.1得到的单场景的费用期望不超过* $\max\left\{\alpha,1+\frac{\alpha^2-\alpha+2}{\alpha-2}e^{-\alpha}\right\}$。

证明. 对每个场景 t，首先考虑场景 t 的设施费用的期望。每个以 (j,t) 为中心的设施-场景对 (i,s)(其中 $s=0$ 或者 t)的开设概率为 $\bar{x}_{ij}^{st}$。由于 $(\bar{x},\bar{y})$ 是完全解，因而，$\bar{x}_{ij}^{st}=\bar{y}_i^s$。更多的，不在任何一个聚类中的设施-场景对的开设概率也为 $\bar{y}_i^s$。因此，场景 t 中的设施费用为

$$\sum_{i\in F}\sum_{s=0,t} f_i^s\bar{y}_i^s=\alpha\sum_{i\in F}\sum_{s=0,t} f_i^s\bar{y}_i^s.$$

对于场景 t 中连接费用的期望，引理2.6意味着其上界

$$\max\left\{\alpha,1+\frac{\alpha^2-\alpha+2}{\alpha-2}e^{-\alpha}\right\}\sum_{j\in D_t}\tilde{C}_{j,t}.$$

对设施和连接费用求和，由算法2.1得到的场景 t 的费用期望不超过

$$\begin{aligned} &\alpha\sum_{i\in F,s=0,t} f_i^s\bar{y}_i^s+\max\left\{\alpha,1+\frac{\alpha^2-\alpha+2}{\alpha-2}e^{-\alpha}\right\}\sum_{j\in D_t}\tilde{C}_{j,t} \\ =\ &\max\left\{\alpha,1+\frac{\alpha^2-\alpha+2}{\alpha-2}e^{-\alpha}\right\}\left(\sum_{i\in F}\left(f_i^0\bar{y}_i^0+f_i^t\bar{y}_i^t\right)+\sum_{i\in F,s=0,t}\sum_{j\in C_t} c_{ij}^{st}\bar{x}_{ij}^{st}\right). \end{aligned}$$

□

通过设置 $\alpha=2.3613$ 使得 $\alpha=1+\frac{\alpha^2-\alpha+2}{\alpha-2}e^{-\alpha}$，可以得到单场景界为 2.3613。

第四节 结论

在本章，我们考虑了设施选址问题中一个重要的变形，两阶段随机设施选址问题。不同于一般问题的近似比，考虑两阶段随机设施选址问题的单场景界。基于线性规划得到了 2.3613 的单场景界，改进了之前最好的 2.4957[38]。未来，一个直接的问题就是能否改进两阶段随机设施选址问题的近似比和单场景界。

此页不缺内容

第三章 带线性惩罚的两阶段随机设施选址的单场景界

第一节 问题介绍

在本章，我们将介绍带线性惩罚的两阶段随机设施选址问题(two-stage stochastic facility location with linear penalty，简记 SFLPLP)。SFLPLP是在 SFLP的基础上，顾客可以不被服务。场景 t 中的顾客 j 不被服务则需支付一定的惩罚费用，不同场景的惩罚费用不一定相同。具体的说，SFLPLP中给定设施集合 F 以及潜在的顾客集合 D。顾客的出现是随机，还给定所有可能被服务的顾客集合 $D_s(s=1,2,\cdots,S)$ 及其出现的概率 q_s。可以在第一阶段以费用 f_i^0 开设设施 i，这一阶段开设的设施可以为之后出现的任何顾客服务。到第二阶段，某个顾客集合 D_s 出现，也称场景 s 实现。在这一阶段可以以 f_i^s 的费用开设设施 i，开设的设施只能为 D_s 中的顾客服务。同时，顾客-场景对 (j,s) 也可以不被服务，若其不被服务需要支付惩罚费用 p_j^s。目标为开设第一阶段和第二阶段各个场景中的一些设施，以及选择一些顾客被惩罚，最终使得开设费用，连接费用，以及惩罚费用总和最小，即，

$$\min_{F_0,F_s\subseteq F,P_s\subseteq D}\left\{\sum_{i\in F_0}f_i^0+\sum_{s=1}^{S}q_s\left(\sum_{i\in F_s}f_i^s+\sum_{j\in D_s\setminus P_s}\min_{i\in F_0\cup F_s}c_{ij}+\sum_{j\in P_s}p_j^s\right)\right\}.$$

与第二章类似地，也引入 $\mathscr{F}$，$\mathscr{D}$ 以及连接费用 c_{ij}^{st}。

- 设施-场景对集合 $\mathscr{F}$: $\mathscr{F}:=\{(i,s):i\in F,s=0,1,\cdots,S\}$。
- 顾客-场景对集合 $\mathscr{D}$: $\mathscr{D}:=\{(j,t):j\in D,s=1,2,\cdots,S\}$。
- 设施-场景对 $(i,s)\in\mathscr{F}$ 和 $(j,t)\in\mathscr{D}$ 之间的连接费用:

$$c_{ij}^{st}:=\begin{cases}c_{ij}, & \text{如果 } s=0 \text{ 或者 } s=t,\\ +\infty, & \text{否则.}\end{cases}$$

同时，定义 $q_0:=1$，因此，SFLPLP可以描述为: 给定设施-场景对集合 $\mathscr{F}$ 以及顾客-场景对集合 $\mathscr{D}$，设施-场景对 (i,s) 和顾客场景对 (j,t) 之间的连接费用 c_{ij}^{st}，每个设施-场景对 $(i,s)\in\mathscr{F}$ 的开设费用为 $q_sf_i^s$，每个顾客-场景对 $(j,t)\in\mathscr{D}$ 的需求为 q_t。顾客-场景对可以不被服务，不被服务的顾客需要支付 $q_tp_j^t$ 的惩罚

费用。最终需要选择一些设施-场景对，为顾客-场景对服务，最终使得开设费用，连接费用和惩罚费用之和最小(见图3.1)。可以看到，如此定义的 SFLPLP与带线性惩罚的设施选址问题类似。但是需要注意的是，这里的连接费用并不满足三角不等式，因此，带线性惩罚的设施选址问题的算法不能直接应用在 SFLPLP。因此，首先给出 SFLPLP的线性整数规划。

$$
\begin{aligned}
\min\quad & \sum_{(i,s)\in\mathscr{F}} q_s f_i^s y_i^s + \sum_{(i,s)\in\mathscr{F}}\sum_{(j,t)\in\mathscr{D}} q_t c_{ij}^{st} x_{ij}^{st} + \sum_{(j,t)\in\mathscr{D}} q_t p_j^t z_j^t \\
\text{s. t.}\quad & \sum_{(i,s)\in\mathscr{F}} x_{ij}^{st} + z_j^t \geq 1, \qquad \forall (j,t)\in\mathscr{D}, \\
& x_{ij}^{st} \leq y_i^s, \qquad \forall (i,s)\in\mathscr{F},(j,t)\in\mathscr{D}, \\
& x_{ij}^{st}, y_i^s, z_j^t \in \{0,1\}, \qquad \forall (i,s)\in\mathscr{F},(j,t)\in\mathscr{D}.
\end{aligned} \tag{3.1}
$$

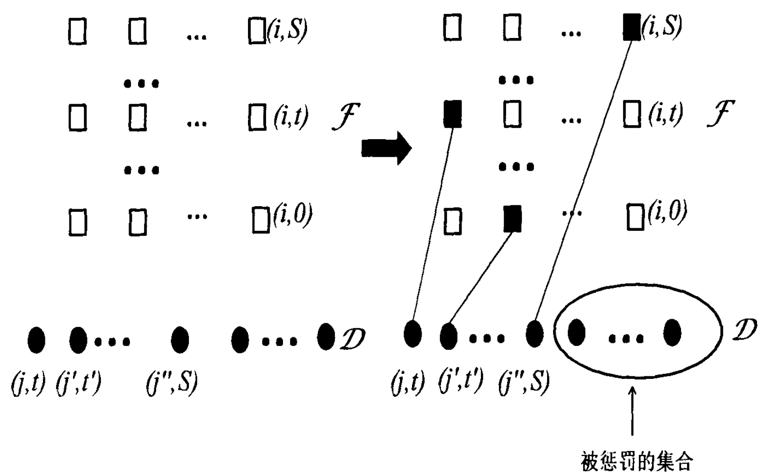


图 3.1 SFLPLP

在以上规划中，x_{ij}^{st} 表示顾客-场景对 (j,t) 是否连接到设施-场景对 (i,s) 上，y_i^s 表示设施-场景对 (i,s) 是否开设，z_j^t 表示顾客-场景对 (j,t) 是否被惩罚。第一组约束表示顾客-场景对 (j,t) 要么连接到某个设施-场景对上，要么被惩罚。第二组约束表示如果顾客-场景对 (j,t) 连接到顾客-场景对 (i,s) 上，那么设施-场景对 (i,s) 必须开设。规划(3.1)的线性规划松弛为

$$
\min\quad \sum_{(i,s)\in\mathscr{F}} q_s f_i^s y_i^s + \sum_{(i,s)\in\mathscr{F}}\sum_{(j,t)\in\mathscr{D}} q_t c_{ij}^{st} x_{ij}^{st} + \sum_{(j,t)\in\mathscr{D}} q_t f_j^t z_j^t
$$

$$\text{s. t.} \quad \sum_{(i,s)\in\mathscr{F}} x_{ij}^{st} + z_j^t \geq 1, \qquad \forall (j,t)\in\mathscr{D}, \tag{3.2}$$
$$x_{ij}^{st} \leq y_i^s, \qquad \forall (i,s)\in\mathscr{F},(j,t)\in\mathscr{D},$$
$$x_{ij}^{st}, y_i^s, z_j^t \geq 0, \qquad \forall (i,s)\in\mathscr{F},(j,t)\in\mathscr{D}.$$

同样的，假设规划(3.2)的分数最优解为完全解(如果不是，可利用分割的技巧得到，参见第二章)。

第二节 算法

在本节，将介绍SFLPLP的线性规划舍入算法。对于一个阈值 $\gamma\in(0,1)$，首先利用规划(3.2)的分数最优解确定被惩罚的顾客-场景对集合。对于剩余的设施-场景对，构造 $S+1$ 个无容量设施选址问题，再对其分别利用线性规划舍入技巧得到整数可行解，最终得到整数可行解。

算法 3.1

步0 给定参数 $\gamma\in(0,1)$ 及 $\alpha\geq 2$。

步1 求解规划(3.2)，得到分数最优解 $(\bar{x},\bar{y},\bar{z})$。不妨设该解是完全解(如果不是，可以利用分割技巧将其转化为完全解)。

步2 令 $\mathscr{P} := \left\{(j,t)\in\mathscr{D} : \bar{z}_j^t \geq \gamma\right\}$ 为被惩罚的顾客-场景对集合。

步3 将 $(\bar{x},\bar{y},\bar{z})$ 中 $\bar{x}$ 和 $\bar{y}$ 部分按比例提高 $\alpha\geq 2$，得到新的解 $(\tilde{x},\tilde{y}) := (\alpha\bar{x},\alpha\bar{y})$，即，$\tilde{x}_{ij}^{st} := \alpha\bar{x}_{ij}^{st}$ 以及 $\tilde{y}_i^s := \alpha\bar{y}_i^s$。同时，可以假设对每个 $(i,s)\in\mathscr{F}$ 和 $(j,t)\in\mathscr{D}$ 都有 $\tilde{x}_{ij}^{st}\leq 1$ 以及 $\tilde{y}_i^s\leq 1$。

步4 对任意的顾客-场景对 $(j,t)\in\mathscr{D}\setminus\mathscr{P}$，按照到 (j,t) 的连接费用递增顺序，重新标记所有第一阶段和第二阶段场景 t 的设施-场景对。假设这个顺序是 $c_{1j}^{0t}\leq\cdots\leq c_{ij}^{0t}\leq\cdots\leq c_{nj}^{0t}$，由于设施地址的相同，同样有 $c_{1j}^{tt}\leq\cdots\leq c_{ij}^{tt}\leq\cdots\leq c_{nj}^{tt}$。

步5 对任意 $s=0,1,2,\cdots,S$，令 $\mathscr{D}_s := \emptyset$。考虑以下情况。

情况5.1 如果 $\sum_{i=1}^n \tilde{x}_{ij}^{0t}\geq 1-\gamma$，定义 $i_0^* := \arg\min\left\{i' : \sum_{i=1}^{i'}\tilde{x}_{ij}^{0t} = 1-\gamma\right\}$(应用分割的技巧保证等式成立)。令 $d_{j,t}^0 := c_{i_0^*j}^{0\,t}$ 以及 $N_{j,t}^0 := \{(i,0) : 1\leq i\leq i_0^*\}$。否则，令 $d_{j,t}^0 = +\infty$。

情况5.2 如果 $\sum_{i=1}^n \tilde{x}_{ij}^{tt}\geq 1-\gamma$，定义 $i_t^* := \arg\min\left\{i' : \sum_{i=1}^{i'}\tilde{x}_{ij}^{tt} = 1-\gamma\right\}$(应用分割的技巧保证等式成立)。令 $d_{j,t}^t := c_{i_t^*j}^{t\,t}$ 以及 $N_{j,t}^t := \{(i,t) : 1\leq i\leq i_t^*\}$。否则，令 $d_{j,t}^0 = +\infty$。

因此，为了得到每个顾客-场景对 (j,t) 的邻域集 $N(j,t)$，计算 $d^0_{j,t}$ 和 $d^t_{j,t}$。

- 如果 $d^0_{j,t} \le d^t_{j,t}$，定义 $d_{j,t} := d^0_{j,t}$ 和 $N(j,t) := N^0_{j,t}$。并令 $\mathscr{D}_0 := \mathscr{D}_0 \cup \{(j,t)\}$。
- 如果 $d^0_{j,t} > d^t_{j,t}$，定义 $d_{j,t} := d^t_{j,t}$ 和 $N(j,t) := N^t_{j,t}$。并令 $\mathscr{D}_t := \mathscr{D}_t \cup \{(j,t)\}$

步6 对 $s = 0,1,\cdots,S$，构造 UFLP的实例 $\mathscr{I}_s$(见图3.2):

- 设施的集合为 F，其中设施 i 的开设费用为 $q_s f^s_i$。
- 顾客的集合为 $\mathscr{D}_s$，并且顾客的需求均为 q_s。
- 每个设施 $i \in F$ 和顾客 (j,t) 之间的连接费用为 c_{ij}。

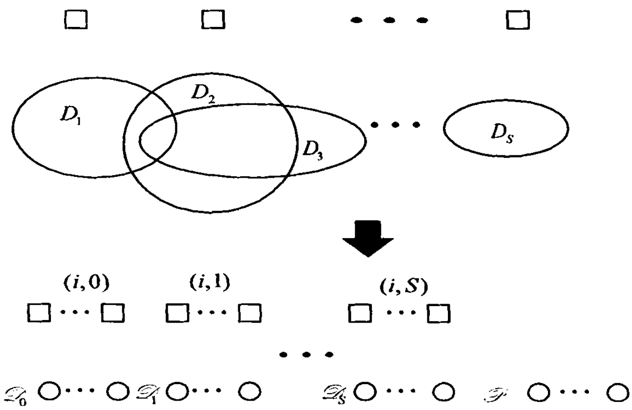


图 3.2 构造 UFLP实例

步7 应用以下算法处理实例 $\mathscr{I}_s$。令 U 是未连接的顾客集合，$\mathscr{D}'$ 是中心顾客的集合。初始时，$U := \mathscr{D}_s$ 和 $\mathscr{D}' := \emptyset$。

步7.1 在 U 中选择 $d_{j,t}$ 最小的顾客-场景对作为中心，即，找出 (j',t') 满足

$$(j',t') := \arg\min_{(j,t)\in U} \{d_{j,t}\}.$$

同时更新 $\mathscr{D}' := \mathscr{D}' \cup \{(j',t')\}$。

步7.2 根据 $N(j',t')$，在 U 中找出与 $N(j',t')$ 分数相连的顾客-场景对集合，记为 $B_{j',t'}$，即，

$$B(j',t) = \{(j,t) \in U : N(j,t) \cap N(j',t') \neq \emptyset\}.$$

同时称 $B(j',t')$ 中的顾客-场景对已连接, 更新 $U := U \setminus B(j',t')$。称 $\{(j',t'),N(j',t'),B(j',t')\}$ 为以 (j',t') 为中心的聚类。

步7.3 重复步7.1和7.2直到 $U = \emptyset$。

步7.4 在每个聚类 $\{(j',t'),B_{j',t'},N_{j',t'}\}$ 中，以概率 $\bar{x}_{ij'}^{st'}$ 相关的开设 $N(j',t')$ 中一个设施-场景对 (i,s)。以 $\bar{y}_i^s$ 独立的概率开设不在任何聚类中的设施-场景对 (i,s)。

步8 将顾客-场景对 $(j,t) \in \mathscr{D} \setminus \mathscr{P}$ 连接到距离其最近的 $\mathscr{I}_0$ 和 $\mathscr{I}_t$ 中开设的设施。

第三节 分析

在本节，将分析算法3.1所得解的质量。由于需要得到单场景界，因此考虑每个场景中的费用。首先考虑每个场景 t 中的惩罚费用。

引理 3.2 在场景 t 中，由算法3.1得到的惩罚费用不超过

$$\frac{1}{\gamma}\sum_{j\in D_t} p_j^t z_j^t.$$

证明. 由算法3.1，知道惩罚的顾客满足 $z_j^t \geq \gamma$，即，$1 \leq \frac{z_j^t}{\gamma}$。因此场景 t 的惩罚费用为

$$\sum_{(j,t)\in\mathscr{P}} p_j^t \leq \sum_{j\in D_t} \frac{z_j^t}{\gamma} p_j^t.$$

引理即证。 □

其次，考虑场景中设施的费用期望。由于 $(\bar{x},\bar{y})$ 是完全解，算法中设施随机开设的概率均为 $\bar{y}_i^s$。

引理 3.3 算法3.1可以得到场景 t 的设施费用期望不超过

$$\alpha \sum_{(i,s)\in\mathscr{F}} q_s f_i^s \bar{y}_i^s.$$

接下来，考虑每个场景中顾客的连接费用期望。在此之前，首先给出 $d_{j,t}$ 的上界。

引理 3.4 对每个 $(j,t) \in \mathscr{D} \setminus \mathscr{P}$，到其邻域集合中最远的设施-场景对的连接费用为

$$d_{j,t} \leq \frac{\alpha}{(\alpha-2)(1-\gamma)}\tilde{C}_{j,t},$$

其中 $\tilde{C}_{j,t} = \sum_{(i,s)\in\mathscr{F}} c_{ij}^{st}\bar{x}_{ij}^{st}$。

证明. 只考虑 $d_{j,t} = d^0_{j,t}$ 的情况，另一种情况 $d_{j,t} = d^t_{j,t}$ 可类似地得到。注意到如果 $d_{i,t} = d^0_{j,t}$ 时，$\sum_{(i,0)} \bar{x}^{0t}_{ij} \geq 1-\gamma$。考虑以下情况。

- $d^t_{j,t} \neq +\infty$，也就是 $\sum_{(i,t)} \bar{x}^{tt}_{ij} \geq 1-\gamma$。由于

$$\sum_{(i,s)\in\mathscr{F}} \bar{x}^{st}_{ij} + z^t_j \geq 1, \qquad \forall (j,t) \in \mathscr{D},$$

以及对任意的 $j \in \mathscr{D} \setminus \mathscr{P}$ 都有 $z^t_j < \gamma$，那么，

$$\sum_{(i,s)\in\mathscr{F}} \bar{x}^{st}_{ij} \geq 1-\gamma, \qquad \forall (j,t) \in \mathscr{F} \setminus \mathscr{P}.$$

再由 $\bar{x}$ 是将 $\tilde{x}$ 同比例扩大 α 倍，因此有

$$\sum_{(i,s)\in\mathscr{F}} \bar{x}^{st}_{ij} \geq \alpha(1-\gamma), \qquad \forall (j,t) \in \mathscr{F} \setminus \mathscr{P}.$$

从算法3.1的步5，有

$$\sum_{(i,0)\in N^0_{j,t}} \bar{x}^{0t}_{ij} = 1-\gamma, \qquad \sum_{(i,t)\in N^t_{j,t}} \bar{x}^{tt}_{ij} = 1-\gamma.$$

那么，

$$\sum_{(i,s)\in\mathscr{F}\setminus\left(N^0_{j,t}\cup N^t_{j,t}\right)} \bar{x}^{st}_{ij} = (\alpha-2)(1-\gamma).$$

此外，对每个 $(i,s) \in \mathscr{F} \setminus \left(N^0_{j,t} \cup N^t_{j,t}\right)$ 都有 $c^{st}_{ij} \geq d_{j,t}$。因此

$$d_{j,t} \leq \frac{\sum_{(i,s)\in\mathscr{F}\setminus\left(N^0_{j,t}\cup N^t_{j,t}\right)} c^{st}_{ij}\bar{x}^{st}_{ij}}{(\alpha-2)(1-\gamma)} \leq \frac{\sum_{(i,s)\in\mathscr{F}} c^{st}_{ij}\bar{x}^{st}_{ij}}{(\alpha-2)(1-\gamma)} = \frac{\alpha}{(\alpha-2)(1-\gamma)}\tilde{C}_{j,t}.$$

- $d^t_{j,t} = +\infty$，也就是 $\sum_{(i,t)} \bar{x}^{tt}_{ij} < 1-\alpha$。那么，

$$\sum_{(i,0)\in\mathscr{F}\setminus N^0_{j,t}} \bar{x}^{0t}_{ij} \geq (\alpha-2)(1-\gamma).$$

此外，对每个设施-场景对 $(i,0) \in \mathscr{F} \setminus N^0_{j,t}$ 都有 $c^{0t}_{ij} \geq d_{j,t}$。因此，有

$$d_{j,t} \leq \frac{\sum_{(i,s)\in\mathscr{F}\setminus N^0_{j,t}} c^{st}_{ij}\bar{x}^{st}_{ij}}{(\alpha-2)(1-\gamma)} \leq \frac{\sum_{(i,s)\in\mathscr{F}} c^{st}_{ij}\bar{x}^{st}_{ij}}{(\alpha-2)(1-\gamma)} = \frac{\alpha}{(\alpha-2)(1-\gamma)}\tilde{C}_{j,t}.$$

引理即证。 □

接下来，可以估计场景 t 的连接费用期望。

引理 3.5 对每个顾客-场景对 $(j,t) \in \mathscr{D} \setminus \mathscr{P}$，场景 t 中的连接费用期望不超过

$$\left(1+2e^{-\alpha(1-\gamma)\frac{\alpha+1}{\alpha-2}}\right)\frac{\tilde{C}_{j,t}}{1-\gamma}.$$

证明. 注意到顾客-场景对 $(j,t) \in \mathscr{D} \setminus \mathscr{P}$ 满足 $\bar{z}_j^t < \gamma$。顾客-场景对 (j,t) 在解 $(\bar{x},\bar{y})$ 下分数连接的设施-场景对都没有开设这一事件发生的概率为

$$\begin{aligned}
& \Pi_{(i,s)\in\{(i,s):\bar{x}_{ij}^{st}\}}\left(1-\bar{y}_i^s\right) \\
\leq\ & \Pi_{(i,s)\in\{(i,s):\bar{x}_{ij}^{st}\}} e^{-\bar{y}_i^s} \\
=\ & e^{-\sum_{(i,s)\in\{(i,s):\bar{x}_{ij}^{st}\}}\bar{y}_i^s} \\
=\ & e^{-\alpha(1-\bar{z}_j^t)} \\
\leq\ & e^{-\alpha(1-\gamma)}.
\end{aligned}$$

在以上不等式中，第一个不等式成立由于不等式当 $y \geq 0$，有 $1-y \leq e^{-y}$。第二个等式由指数函数的运算法则得到。第三个等式成立由于在规划(3.2)中，必然存在一个最优解满足 $\sum\limits_{(i,s)} \bar{x}_{ij}^{st} + \bar{z}_j^t = 1$ 以及 $(\bar{x},\bar{y})$ 是完全解。最后一个不等式成立由于 (j,t) 在集合 $\mathscr{D} \setminus \mathscr{P}$ 中满足 $\bar{z}_j^t < \gamma$。

在这种情况下，假设 (j',t') 是 (j,t) 所在聚类的中心，(i,s) 是 (j',t') 和 (j,t) 邻域交集中的一个设施-场景对，(i',s) 是这个聚类中的开设设施(见图3.3)。利用三角不等式，(j,t) 的连接费用不超过

$$c_{i'j}^{s\,t} \leq c_{ij}^{st} + c_{ij'}^{st'} + c_{i'j}^{s\,t'} \leq d_{j,t} + d_{j',t'} + d_{j',t'} \leq 3d_{j,t}.$$

在以上不等式中，第一个不等式由三角不等式成立，第二个不等式由于 $d_{j,t}$ 是顾客-场景对 (j,t) 分数连接的设施-场景对中距离最远的，第三个不等式由中心的选取准则得到。因此，(j,t) 的连接费用期望不超过 $3\frac{\alpha}{(\alpha-2)(1-\gamma)}\tilde{C}_{j,t}$。

另一方面，如果顾客-场景对 (j,t) 分数连接的设施-场景对中至少有一个设施开设，由引理2.5可以得到 (j,t) 的连接费用不超过 $\frac{\tilde{C}_{j,t}}{1-\alpha}$。

因此，有顾客-场景对 (j,t) 的连接费用期望不超过

$$\begin{aligned}
& \left(1-e^{-\alpha(1-\gamma)}\right)\times\frac{1}{1-\gamma}\tilde{C}_{j,t} + e^{-\alpha(1-\gamma)}\times 3\frac{\alpha}{(\alpha-2)(1-\gamma)}\tilde{C}_{j,t} \\
\leq\ & \left(1+2e^{-\alpha(1-\gamma)}\frac{\alpha+1}{\alpha-2}\right)\frac{\tilde{C}_{j,t}}{1-\gamma}.
\end{aligned}$$

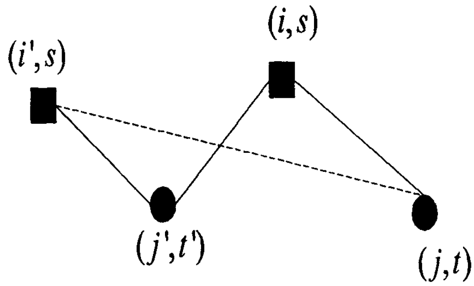


图 3.3 顾客-场景对 (j,t) 连接到 $(i'.s)$

引理即证。 □

到现在为止，可以分析得到单场景界。

定理 3.6 *算法3.1可以得到 SFLPLP的 3.0294 单场景界。*

证明. 由引理3.2，引理3.3以及引理3.5，综合算法3.1可以得到场景 t 的设施费用，连接费用和惩罚费用的总期望不超过

$$\begin{aligned}
&\alpha\sum_{i\in F} f_i^0\bar{y}_i^0+\alpha\sum_{i\in F} q_t f_i^t y_i^t \\
&+\sum_{j\in D_t}\left(1+2e^{-\alpha(1-\gamma)}\frac{\alpha+1}{\alpha-2}\right)\frac{1}{1-\gamma}\sum_{(i,s)\in\mathscr{F}} q_t c_{ij}^{st}\bar{x}_{ij}^{st} \\
&+\frac{1}{\gamma}\sum_{j\in D_t} q_t p_j^t\bar{z}_j^t \\
\leq\ & \max\left\{\alpha,\left(1+2e^{-\alpha(1-\gamma)}\frac{\alpha+1}{\alpha-2}\right)\frac{1}{1-\gamma},\frac{1}{\gamma}\right\}\left(\sum_{i\in F} f_i^0\bar{y}_i^0+\sum_{i\in F} q_t f_i^t y_i^t\right. \\
&\left.+\sum_{j\in D_t}\sum_{(i,s)\in\mathscr{F}} q_t c_{ij}^{st}\bar{x}_{ij}^{st}+\sum_{j\in D_t} q_t p_j^t\bar{z}_j^t\right).
\end{aligned}$$

因此，SFLPLP的单场景界是

$$\max\left\{\alpha,\left(1+2e^{-\alpha(1-\gamma)}\frac{\alpha+1}{\alpha-2}\right)\frac{1}{1-\gamma},\frac{1}{\gamma}\right\}.$$

当取 $\gamma=0.3301$ 和 $\alpha=3.0290$ 时，可以得到单场景界 3.0294。 □

第四节 结论

在本章，我们讨论了带线性惩罚的随机设施选址问题的单场景界。通过分

数最优解，首先确定被惩罚的顾客集合，再将原问题分解成多个经典的设施选址问题。利用线性舍入的技巧得到整数解。最终，可以分析出该算法得到的单场景界为 3.0294。

此页不缺内容

第四章 2-层设施选址问题

第一节 问题介绍

在本章，我们将介绍 2-层设施选址问题(2-level facility location problem，简称 2-LFLP)。一般地，k-层设施选址问题可以描述为: 给定 $(k+1)$-部图 $(F_1 \cup F_2 \cup \cdots \cup F_k \cup D, E)$，其中顶点集为具有层次结构的 k 个分层设施集合构成以及顾客集合 D，E 为任意两个集合中之间的边。每个设施 $i \in F := F_1 \cup F_2 \cup \cdots \cup F_k$ 的开设费用为 f_i。给定边的费用 $c : E \to \mathbb{R}_+$。目标是在每层设施 $F_l(l=1,2,\cdots,k)$ 中开设一些 F_l'，并且将任意顾客 j 连接到由以上分层开设设施 $i_1(j) \in F_1', i_2(j) \in F_2', \cdots, i_k(j) \in F_k'$ 构成的路 $(i_1(j), i_2(j), \cdots, i_k(j))$ 上，使得开设费用和连接费用之和最小，即，

$$\min_{F_l' \subseteq F_l, l=1,2,\cdots,k} \left\{ \sum_{l=1}^{k} \sum_{i \in F_l'} f_i + \sum_{j \in D} c_{ji_1(j)} + c_{i_1(j)i_2(j)} + \cdots + c_{i_{k-1}(j)i_k(j)} \right\}.$$

在本章，考虑 k-层设施选址问题的特殊情况，2-层设施选址问题，即 $k=2$。注意到，在 2-LFLP中，顾客连接到设施组成的路上，因此，引入路的集合 $P := F_1 \times F_2$。一条路 $p \in P$ 开设当且仅当构成这条路的所有设施都开设。同时，定义顾客 $j \in D$ 到路 $p=(k,i) \in P$ 的距离为

$$c_{ikj} = c_{ik} + c_{kj}.$$

在第一章中曾经介绍过 k-层设施选址问题的最优解由一些不相交的 k 层的树构成的森林，因此可以得到，2-LFLP的最优解中是有一些不相交的 2 层的树 (i,S,J) 构成的森林(见图4.1)。因此，希望可以写出集合覆盖形式的 2-LFLP的线性整数规划。定义由第二层设施 $i \in F_2$，第一层设施集合 $S \subseteq F_1$ 以及顾客集合 $J \subseteq D$ 构成的树相应费用为

$$C_{iSJ} := f_i + \sum_{k \in S} f_k + \sum_{j \in J} \min_{k \in S} c_{ikj}.$$

因此，为了得到2-LFLP的集合覆盖形式的线性整数规划，引入 $0-1$ 变量 X_{iSJ} 表示树 (i,S,J) 是否选入到最优解中。因此，有

$$\begin{aligned} \min \quad & \sum_{i\in F_2}\sum_{S\subseteq F_1}\sum_{J\subseteq D} C_{iSJ}X_{iSJ} \\ \text{s. t.} \quad & \sum_{i\in F_2, S\subseteq F_1, J\subseteq D: j\in J} X_{iSJ} \geq 1, \qquad \forall j\in D, \\ & X_{iSJ}\in\{0,1\}, \qquad \forall i\in F_2, S\subseteq F_1, J\subseteq D. \end{aligned} \tag{4.1}$$

在以上规划中，约束表示对每个顾客 j 都一定存在一个树 (i,S,J) 使得 $j\in J$。

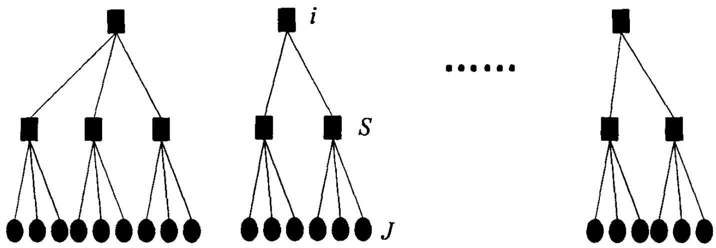


图 4.1　2-LFLP的最优解.

将 $0-1$ 变量松弛成非负变量，可以得到上述规划的线性规划松弛。

$$\begin{aligned} \min \quad & \sum_{i\in F_2}\sum_{S\subseteq F_1}\sum_{J\subseteq D} C_{iSJ}X_{iSJ} \\ \text{s. t.} \quad & \sum_{i\in F_2, S\subseteq F_1, J\subseteq D: j\in J} X_{iSJ} \geq 1, \qquad \forall j\in D, \\ & X_{iSJ}\geq 0, \qquad \forall i\in F_2, S\subseteq F_1, J\subseteq D. \end{aligned} \tag{4.2}$$

在上述规划中，其变量是关于输入的指数。相应的对偶规划为

$$\begin{aligned} \max \quad & \sum_{j\in D}\alpha_j \\ \text{s. t,} \quad & \sum_{j\in J}\alpha_j \leq C_{iSJ}, \qquad \forall i\in F_2, S\subseteq F_1, J\subseteq D, \\ & \alpha_j\geq 0, \qquad \forall j\in D. \end{aligned} \tag{4.3}$$

在对偶中，约束是指数多个，求解其需要利用椭球算法。但是我们不求解对偶规划，而是通过对偶约束构造对偶可行解，同时得到一个原始整数可行解。

在本章中，需要考虑到极大化的设施选址问题(maximized facility location problem，记为 Max-FLP)。该问题可以描述为: 给定设施的集合 F 和顾客的集合

D，每个设施 $i\in F$ 的开设费用为 f_i，每个顾客 $j\in D$ 连接到设施 $i\in F$ 上得到服务即可得到利润 d_{ij}(这里 d_{ij} 不一定满足度量性)。需要开设一些设施，将顾客连接到开设的设施上，使得净利润(连接利润减去设施开设成本)最大，即，

$$\max_{F'\subseteq F}\left\{\sum_{j\in D}\max_{i\in F'}d_{ij}-\sum_{i\in F'}f_i\right\}.$$

由于 Max-FLP可以划归到顶点覆盖问题(见文献[55])，Max-FLP是 NP-困难的。因而，使用近似算法处理 Max-FLP。所谓顶点覆盖问题，给定一个顶点赋权图 $G=(V,E)$，其中 V 是顶点集，E 是边集。每个顶点 $v\in V$ 都具有权重 w_v。需要在顶点中选择一个顶点子集 V' 使得每条边至少有一个顶点在 V' 中，并且使得选择的顶点权重之和最小。

在原始-对偶算法中，与经典的原始-对偶算法类似分为两步。在第一步中，利用近似求解 Max-FLP得到对偶可行解。在第二步中，需要先定义独立的概念，最终选择极大独立集。给定图 $G=(V,E)$，所谓独立集是指集合中任意两点都是独立的，即，独立集中的顶点没有边相连。而极大独立集是指再增加任意一点到集合中，该集合不再是独立集。求解极大独立集问题是 P-问题。事实上，可以通过求最小覆盖集来求最大独立集。

第二节 子问题的算法及分析

在本节，将介绍 2-LFLP的原始-对偶算法的子问题。原始-对偶算法的思想为从对偶规划出发，随时间增长初始值为 0 的对偶变量，得到对偶可行解。再从对偶可行解构造原始整数解。在增长对偶变量的过程中，需要利用对偶约束的分离神谕(separation oracle)。经典的原始-对偶技巧中，分离神谕可以在多项式时间求解，因而可以连续的增长。但不幸的是，2-LFLP对偶规划的分离神谕等价于一个 Max-FLP，因而是 NP-困难的。所以，为了处理这一困难，我们近似求解。接下来，介绍 Max-FLP的近似算法。

首先，考虑对偶的分离神谕，即，给定 α_j, c_{ikj}, f_k 以及 f_i 的取值，需要判定以下不等式何时成立。

$$\sum_{j\in J}\left(\alpha_j-\min_{k\in S}c_{ikj}\right)-\sum_{k\in S}f_k\le f_i,\qquad \forall i\in F_2, S\subseteq F_1, J\subseteq D.$$

任意的树 (i,S,J) 上式均成立，实际上等价于上式左边对任意树 (i,S,J) 求极大，

因而，

$$\max_{i\in F_2,S\subseteq F_1,J\subseteq D}\sum_{j\in J}\left\{\left(\alpha_j-\min_{k\in S}c_{ikj}\right)-\sum_{k\in S}f_k\right\}\leq f_i.$$

需要指出上述不等式左边在多项式时间内可以求解。但是不幸的是，该优化问题的是 NP-困难的，如果 P≠NP，在多项式找不到最优解。

引理 4.1 *优化问题* $\max_{i\in F_2,S\subseteq F_1,J\subseteq D}\left\{\sum_{j\in J}\left(\alpha_j-\min_{k\in S}c_{ikj}\right)-\sum_{k\in S}f_k\right\}$ *是一个NP-困难问题。*

证明. 对于 $i\in F_2$ 的优化，可以通过遍历 F_2 所有设施实现。接下来，考虑对 $S\subseteq F_1$ 和 $J\subseteq D$ 的优化，即，对任意的第二层设施 $i\in F_2$，求解

$$\begin{aligned}&\max_{S\subseteq F_1,J\subseteq D}\sum_{j\in J}\left(\alpha_j-\min_{k\in S}c_{ikj}\right)-\sum_{k\in S}f_k\\ \Leftrightarrow\ &\max_{S\subseteq F_1}\sum_{j\in D}\max_{k\in S}\left(\alpha_j-c_{ikj}\right)^+-\sum_{k\in S}f_k.\end{aligned}\tag{4.4}$$

其中，函数 $(z)^+:=\max\{z,0\}$。事实上，以上优化问题可以视为一个极大化的单层设施选址问题，其中设施的集合为 F_1，顾客的集合为 D，设施的开设费用为 f_k，连接利润为 $\left(\alpha_j-c_{ikj}\right)^+$。

由于 Max-FLP可以归约到顶点覆盖问题，因而该优化问题是 NP-困难的。引理即证。 □

由引理4.1可知对偶问题的分离神谕不是多项式时间可解的。但是，幸运的是，知道这个分离神谕是一个极大化的设施选址问题，而极大化设施选址问题可以近似求解。因此，接下来介绍近似求解这个子问题的算法。

4.2.1 子问题的算法

在本节，将介绍 Max-FLP问题的双因子组合近似算法。

定义 4.2 *一个 Max-FLP的解 S 的双因子为 (r_f,r_c)，如果*

$$r_cC^S_{\max}-r_fF^S_{\max}\geq C^{\mathrm{OPT}}_{\max}-F^{\mathrm{OPT}}_{\max},$$

其中 $C^S_{\max}$ 和 $C^{\mathrm{OPT}}_{\max}$ 分别表示解 S 和最优解中连接利润，$F^S_{\max}$ 和 $F^{\mathrm{OPT}}_{\max}$ 分别表示解 S 和最优解中的设施开设成本。

为了达到目的，首先考虑 Max-FLP的连接利润函数，再讨论利润函数的性质。连接纯利润可以定义为

$$h\left(F'\right)=\sum_{j\in D}\Delta max_{i\in F'}d_{ij}.$$

引理 4.3 *函数 $h(F')$ 是非负的，单调的以及次模的。*

证明. 显然的，函数 $h(F')$ 是非负的，单调的。下面将证明 $h(F')$ 是次模的。

对任意 $T_1, T_2 \subseteq F$，有

$$h(T_1) = \sum_{j \in D} \max_{i \in T_1} d_{ij} \text{ 和 } h(T_2) = \sum_{j \in D} \max_{i \in T_2} d_{ij}.$$

假设 $h(T_1 \cup T_2)$ 中的最远距离在 T_1 中得到，另一种情况类似可得。这样，$h(T_1 \cup T_2) = h(T_1)$。而

$$h(T_2) = \sum_{j \in D} \max_{i \in T_2} d_{ij} \geq \sum_{j \in D} \max_{i \in T_1 \cap T_2} d_{ij} = h(T_1 \cap T_2).$$

因而引理得证。 □

在 Max-FLP中，如果已知 $F_{\max}^{\text{OPT}}$(不妨设 $F_{\max}^{\text{OPT}} \geq 1$)，那么目标函数中的设施成本可以将目标函数中的设施成本转化为约束 $\sum_{i \in \hat{F}} f_i \leq F_{\max}^{\text{OPT}}$，其中 $\hat{F}$ 是开设设施集合。这样，可以得到等价的规划。但事实上，并不知道真实的 $F_{\max}^{\text{OPT}}$。因此，需要对任意的 $\varepsilon > 0$，在预算 B 的序列 $\{1, 1+\varepsilon, (1+\varepsilon)^2, \cdots, (1+\varepsilon)^L\}$ 中找到 ℓ 使得 $(1+\varepsilon)^\ell \leq F_{\max}^{\text{OPT}} \leq (1+\varepsilon)^{\ell+1}$，其中 $L = \lceil \log_{1+\varepsilon} f_{tot} \rceil$ 和 $f_{tot} := \sum_{i \in F} f_i$。所以，考虑以下在背包约束下非负，单调不减，次模函数极大化问题: 对于一个固定的预算 B,

$$\max_{\hat{F} \subseteq F} \left\{ h(\hat{F}) : \sum_{i \in \hat{F}} f_i \leq B \right\}. \tag{4.5}$$

记取 B 时得到问题的最优解对应的连接利润和设施成本为 $C_{\max}(B)$ 和 $F_{\max}(B)$。

引理 4.4 ([56]) *规划(4.5)存在 $(1-1/e)$-近似算法，运行时间为 $O(n^5)$，其中 $n = |F|$。*

基于以上算法，可以描述处理 Max-FLP的算法。在的子程序算法中，遍历所有可取的 B 运行 Sviridenko[56] 的算法，再在所有的情况取最好的。具体的，子问题的算法可以描绘成以下过程。

算法 4.5

步 1 固定一个常数 $\varepsilon > 0$。令 $f_{tot} := \sum_{i \in F} f_i$ 和 $L := \lceil \log_{1+\varepsilon} f_{tot} \rceil$。

步2 对每个 $\ell=0,1,2,\cdots,L$，令 $B_l:=(1+\varepsilon)^{\ell}$，令 $B=B_\ell$，应用 Sviridenko[56] 近似求解(4.5)。令其最优解为 S_ℓ。

步3 令

$$\ell_0 := \arg\max_{\ell=0,1,\ldots,L}\left\{\frac{e}{e-1}C_{\max}(B_\ell)-\frac{1}{1+\varepsilon}B_\ell\right\}. \tag{4.6}$$

记其相应的连接利润和开设费用分别为 $C_{\max}:=C_{\max}(B_{\ell_0})$ 和 $F_{\max}:=\sum_{i\in S_{\ell_0}}f_i$。

4.2.2 子问题的分析

在本节，将分析算法4.5，得到算法所得解的双因子。规划(4.6)中 ℓ_0 的选择是为了得到以下定理。

定理 4.6 令 $C_{\max}^{\mathrm{FEA}}$ 和 $F_{\max}^{\mathrm{FEA}}$ 分别为Max-FLP的任意可行解FEA的连接利润和设施成本。算法4.5是一个组合算法，并且输出解的纯利润为 $C_{\max}-F_{\max}$ 使得

$$\frac{e}{e-1}C_{\max}-\frac{1}{1+\varepsilon}F_{\max}\geq C_{\max}^{\mathrm{FEA}}-F_{\max}^{\mathrm{FEA}},$$

即，该解的双因子为 $(1/(1+\varepsilon),\ e/(e-1))$。

证明. 由 B 的取值，可以知道存在 ℓ^* 满足 $B_{\ell^*-1}\leq F_{\max}^{\mathrm{FEA}}\leq B_{\ell^*}\leq(1+\varepsilon)F_{\max}^{\mathrm{FEA}}$。这样，当 B 取值 B_{ℓ^*} 时，FEA是规划(4.5)的可行解。由引理4.4和算法4.5，可以得到

$$C_{\max}(B_{\ell^*})\geq\left(1-\frac{1}{e}\right)C_{\max}^{\mathrm{OPT}}(B_{\ell^*})\geq\frac{e-1}{e}C_{\max}^{\mathrm{FEA}}. \tag{4.7}$$

从算法4.5的第3步，规划(4.5)，以及上述不等式，可以得到

$$\begin{aligned}\frac{e}{e-1}C_{\max}-\frac{1}{1+\varepsilon}F_{\max} &\geq \frac{e}{e-1}C_{\max}(B_{\ell_0})-\frac{1}{1+\varepsilon}B_{\ell_0}\\ &\geq \frac{e}{e-1}C_{\max}(B_{\ell^*})-\frac{1}{1+\varepsilon}B_{\ell^*}\\ &\geq C_{\max}^{\mathrm{FEA}}-F_{\max}^{\mathrm{FEA}}.\end{aligned}$$

由于 $F_{\max}$ 是当 $B=B_{\ell_0}$ 得到的设施成本，因而满足 $F_{\max}\leq B_{\ell_0}$。由 ℓ_0 的取法可知，第二个不等式成立。由于 $B_{\ell^*}\leq(1+\varepsilon)F_{\max}^{\mathrm{FEA}}$ 以及不等式(4.7)最后一个不等式成立。 □

第三节 2-层设施选址问题的算法及分析

在本节，给出 2-LFLP的原始-对偶算法。在第4.2.2节中，已近似解决了分离神谕多项式时间求解问题。注意到分离神谕是在给定 α_j 的条件下的问题，但是由于分离神谕是近似求解，不能类似经典的设施选址问题的原始-对偶算法，连续的增长对偶变量。在我们的算法中，给定一个常数 $\varepsilon > 0$ 时离散的时间点 $t \in \{1, (1+\varepsilon), \cdots\}$(不妨设 $f_i \geq 1, c_{ij} \geq 1$)。令 $\bar{F}_2$ 表示在第二层临时开设的设施结合。U 表示未连接的顾客集合。对每个 $i \in F_2$，引入以下符号:

- D_i: 所有连接到设施 i 上的顾客集合。
- $\ell(i)$: i 开设的次数。
- $t_i^{\ell(i)}$: i 第 $\ell(i)$ 次开设的时间。

4.3.1 算法

现在可以描述 2-LFLP的原始-对偶算法。算法可以分为构造对偶可行解和构造原始整数可行解两步。具体的，可以描述为以下过程。

算法 4.7

步*1.* **构造对偶可行解**

步1.0 固定一个常数 $\varepsilon > 0$。考虑离散时间点 $t \in \{1, (1+\varepsilon), (1+\varepsilon)^2, \ldots\}$。初始时，令 $s := 0$，$U := D$，$\bar{F}_2 := \emptyset$，对每个顾客 $j \in D$ 有 $\alpha_j := 0$，对每个设施 $i \in F_2$ 有 $D_i := \emptyset$ 以及 $\ell(i) := 1$。

步1.1 令 $t := (1+\varepsilon)^s$，$F_2(t) := F_2$。

步1.2 如果 $F_2(t) = \emptyset$，运行步*1.6*; 否则，任意选择 $i \in F_2(t)$，构造一个 *Max-FLP*的实例: 对每个顾客 $j \in D$ 和设施 $k \in F_1$，定义单位连接利润为

$$d_{kj} := \begin{cases} \max\{t - c_{ikj}, 0\}, & \text{如果 } j \in U, \\ \max\{\alpha_j - c_{ikj}, 0\}, & \text{否则}. \end{cases}$$

设施 $k \in F_1$ 的开设费用为 f_k。

步1.3 对步*1.2*中实例应用算法*4.5*得到可行解，设其连接利润和设施成本分别为 $C_{\max}$ 和 $F_{\max}$。并且令 $F_2(t) := F_2(t) \setminus \{i\}$。

步1.4 如果

$$\frac{e}{e-1} C_{\max} - \frac{1}{1+\varepsilon} F_{\max} \leq f_i,$$

返回步*1.2*; 否则，运行步*1.5*。

步1.5 令 $t_i^{\ell(i)} := t$。令 $\bar{F}_2 := \bar{F}_2 \cup \{i\}$，同时引入以下符号：

* $S_i^{\ell(i)}$：由算法4.5得到的可行解中，在 F_1 中开设的临时设施集合。
* $\sigma_i^{\ell(i)}(j)$：由算法4.5得到的可行解中，顾客 j 临时连接到 $S_i^{\ell(i)}$ 中的开设设施。
* $\bar{D}_i^{\ell(i)} := \{j \in U | d_{\sigma_i^{\ell(i)}(j)j} > 0\}$：在此次迭代中，由于通过 $S_i^{\ell(i)}$ 连接到 i 而停止增长对偶变量的顾客集合，这样的顾客也称为被冻结的顾客。对每个 $j \in \bar{D}_i^{\ell(i)}$，令 $i(j) := i$ 以及 $\alpha_j := t_i^{\ell(i)}/(1+\varepsilon)$。
* $D_i^{\ell(i)} := \{j \in D | d_{\sigma_i^{\ell(i)}(j)j} > 0\}$：在此次迭代中，由于通过 $S_i^{\ell(i)}$ 连接到 i 的顾客集合。

除此之外，更新 $U := U \setminus \bar{D}_i^{\ell(i)}$，$D_i := D_i \cup D_i^{\ell(i)}$ 和 $\ell(i) := \ell(i)+1$。然后返回到步1.2。

步1.6 如果 $U \neq \emptyset$，令 $s := s+1$，返回步1.1。否则，停止。

步 2. **构造原始整数可行解。**

初始时，一些顾客会同时在若干的 D_i^1 中(见图4.2)。

步2.1 在 $\bar{F}_2$ 中的设施 i 和 i' 是独立的如果存在一个顾客同时在 D_i^1 和 $D_{i'}^1$ 中。在 $\bar{F}_2$ 中选取极大独立集 $\hat{F}_2$。开设 $\hat{F}_2$ 以及 $S_i^1 (\forall i \in \hat{F}_2)$ 中的设施。

步2.2 将每个顾客连接到最近的路径上。

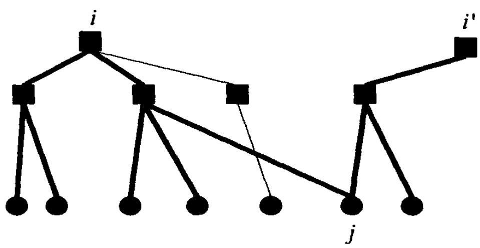


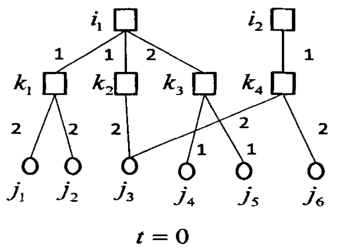
图 4.2 j 落在 D_i^1 和 $D_{i'}^1$ 中

在上述算法中，每个设施 $i \in \bar{F}_2$ 被称为临时开设。对每个 $i \in \bar{F}_2$，称顾客 $j \in \bar{D}_i^{\ell(i)} (\ell(i) \geq 1)$ 碰到 i。

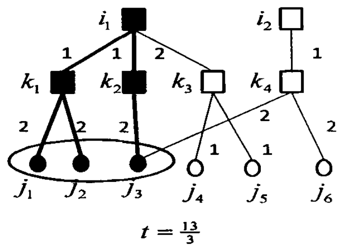
4.3.2 算法的实例

由于算法较为复杂，所以接下来给出一个例子说明算法过程。假设 F_2 中有两个设施 $\{i_1, i_2\}$，其中 $f_{i_1} = 2$ 和 $f_{i_2} = 4$，F_1 中有四个设施 k_1, k_2, k_3, k_4，其中

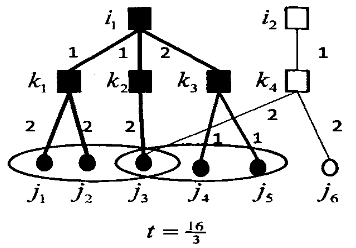
$f_{k_1}=f_{k_2}=1$，$f_{k_3}=3$ 和 $f_{k_4}=2$。连接费用可见图4.4，如果图中两点没有直接的边连接，则该边的连接费用通过最短路得到。随着算法的运行，在第一步中将会发生三个事件。当 $t=13/3$，临时开设 i_1 并且 $D_{i_1}^1=\{j_1,j_2,j_3\}$。当 $t=16/3$，临时开设 i_1 并且 $D_{i_1}^2=\{j_3,j_4,j_5\}$。当 $t=20/3$，临时开设 i_2 并且 $D_{i_2}^1=\{j_3,j_6\}$。然后在第二步，选择极大独立集 $\{i_1\}$。



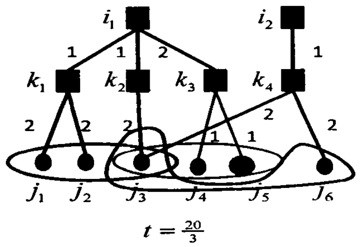
(a) 输入数据



(b) 由于 j_1,j_2,j_3，i_1 临时开设



(c) 由于 j_3,j_4,j_5，i_1 临时开设



(d) 由于 j_3,j_6，i_2 临时开设

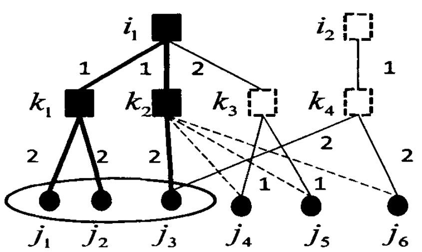


图 4.3 将 j_1,j_2 通过 k_1 连接到 i_1; 将 j_3,j_4,j_5,j_6 通过 k_2 连接到 i_1

4.3.3 算法分析

在本节，将给出子程序算法4.5的分析，为了之后的算法设计，只分析得到

双因子。首先，需要说明由算法得到的 $\{\alpha_j\}$ 是对偶可行解，即，规划(4.3)的可行解。

引理 4.8 *由算法4.7得到的 $\{\alpha_j\}$ 为规划(4.3)的可行解。*

证明. 对每个设施 $i \in F_2$，第一层设施集合 $S \subseteq F_1$，和顾客集合 $J \subseteq D$，需要证明

$$\sum_{j\in J}\alpha_j \le C_{iSJ}, \qquad \forall i \in F_2, S \subseteq F_1, J \subseteq D.$$

对给定的设施 $i \in F_2$，考虑算法4.7第一步结束的时候。此时，连接利润定义为 $d_{kj} := \max\{\alpha_j - c_{ikj}, 0\}$。从定理4.6和算法4.7的步1.4，对任意 Max-FLP的可行解FEA，有

$$C_{\max}^{\text{FEA}} - F_{\max}^{\text{FEA}} \le \frac{e}{e-1}C_{\max} - \frac{1}{1+\varepsilon}F_{\max} \le f_i.$$

现在，构造 Max-FLP的可行解: 开设所有在 S 中的设施，并将 D 中的顾客通过 S 连接到 i。那么，有 $\sum_{j\in D}\max_{k\in S}\max\{\alpha_j - c_{ikj}, 0\} - \sum_{k\in S} f_k \le f_i$。因此，

$$\begin{aligned}
& \sum_{j\in J}\max\left\{\alpha_j - \min_{k\in S} c_{ikj}, 0\right\} - \sum_{k\in S} f_k \\
= & \sum_{j\in J}\max_{k\in S}\max\left\{\alpha_j - c_{ikj}, 0\right\} - \sum_{k\in S} f_k \\
\le & \sum_{j\in D}\max_{k\in S}\max\left\{\alpha_j - c_{ikj}, 0\right\} - \sum_{k\in S} f_k \le f_i.
\end{aligned}$$

在以上不等式中，由于极小化负数等价于求极大则第一个等式成立，由于将对 $J \subseteq D$ 中的顾客求和放大为对 D 中顾客求和因而第二个不等式成立。现在，可以得到

$$\begin{aligned}
\sum_{j\in J}\left(\alpha_j - \min_{k\in S} c_{ikj}\right) & \le \sum_{j\in J}\max\left\{\alpha_j - \min_{k\in S} c_{ikj}, 0\right\} \\
& \le f_i + \sum_{k\in S} f_k.
\end{aligned}$$

等价的有

$$\sum_{j\in J}\alpha_j \le C_{iSJ} = f_i + \sum_{k\in S} f_k + \sum_{j\in J}\min_{k\in S} c_{ikj}.$$

引理即证。 □

证明了对偶可行性之后，接下来，利用 $\{\alpha_j\}$ 需要估计以 i 为根节点的树的费用。

引理 4.9 对任意 $i\in\hat{F}_2$，有

$$f_i+\frac{1}{1+\varepsilon}\sum_{k\in S_i^1} f_k+\frac{e}{e-1}\sum_{j\in D_i^1}\min_{k\in S_i^1} c_{ikj}\le(1+\varepsilon)\frac{e}{e-1}\sum_{j\in D_i^1}\alpha_j.$$

证明. 注意到 D_i^1 是在时刻 t_i^1 连接到 i 的顾客集合，其中 $\bar{D}_i^1$ 是在 t_i^1 之前没有碰到任何设施的顾客集合。考虑任意 $i\in\hat{F}_2$，由算法4.7第一步有

$$\begin{aligned}&\frac{e}{e-1}\left(\sum_{j\in D_i^1\setminus\bar{D}_i^1}\left((1+\varepsilon)\alpha_j-\min_{k\in S_i^1} c_{ikj}\right)+\sum_{j\in\bar{D}_i^1}\left(t_i^1-\min_{k\in S_i^1} c_{ikj}\right)\right)-\frac{1}{1+\varepsilon}\sum_{k\in S_i^1} f_k\\ \ge\ & f_i.\end{aligned}$$

因此，有

$$\begin{aligned}&\frac{e}{e-1}\sum_{j\in D_i^1}\left((1+\varepsilon)\alpha_j-\min_{k\in S_i^1} c_{ikj}\right)-\frac{1}{1+\varepsilon}\sum_{k\in S_i^1} f_k\\ =\ &\frac{e}{e-1}\left(\sum_{j\in D_i^1\setminus\bar{D}_i^1}\left((1+\varepsilon)\alpha_j-\min_{k\in S_i^1} c_{ikj}\right)+\sum_{j\in\bar{D}_i^1}\left((1+\varepsilon)\alpha_j-\min_{k\in S_i^1} c_{ikj}\right)\right)\\ &-\frac{1}{1+\varepsilon}\sum_{k\in S_i^1} f_k\\ \ge\ &\frac{e}{e-1}\left(\sum_{j\in D_i^1\setminus\bar{D}_i^1}\left((1+\varepsilon)\alpha_j-\min_{k\in S_i^1} c_{ikj}\right)+\sum_{j\in\bar{D}_i^1}\left(t_i^1-\min_{k\in S_i^1} c_{ikj}\right)\right)\\ &-\frac{1}{1+\varepsilon}\sum_{k\in S_i^1} f_k\\ \ge\ & f_i.\end{aligned}$$

在以上不等式中，将集合 D_i^1 划分为不相交的 $\bar{D}_i^1$ 和 $D_i^1\setminus\bar{D}_i^1$ 两个集合，第一个等式成立。由 α_j 的定义，可以得到第二个不等式成立。最后一个不等式由算法步1.4得到。引理即证。 □

上述证明估计了在算法所得解中极大独立集相应树的费用。接下来，估计没有连接到 $\hat{F}_2$ 中任何的设施上的顾客对应的连接费用。这类顾客在算法中连接到路径被关闭了，因此需要重新连接到开设的路径上。所以，这类的顾客只需要支付连接费用。

引理 4.10 对任何 $j\in D\setminus\bigcup_{i\in\hat{F}_2} D_i^1$，连接费用不超过 $3(1+\varepsilon)\alpha_j$。

证明. 对每个顾客 $j\in D\setminus\bigcup_{i\in\hat{F}_2}D_i^1$，考虑 j 第一次碰到的开设设施，即，$i(j)$。存在 $\ell\geq 1$ 使得 $j\in\bar{D}_{i(j)}^{\ell}$ 并且 j 通过 $\sigma_{i(j)}^{\ell}(j)\in F_1$ 连接到 $i(j)$。考虑以下情况。

1. 如果 $i(j)$ 开设。
   注意到对任意 $i\in\hat{F}_2$ 都有 $j\notin D_i^1$，这意味着 $j\notin D_{i(j)}^1$。因此，对于任意 $j'\in D_{i(j)}^1$ 有 $\alpha_j\geq\alpha_{j'}$。由三角不等式，任意选择 $j'\in D_{i(j)}^1$，j 通过 $\sigma_{i(j)}^1(j')$ 连接到 $i(j)$ 上的连接费用不超过(见图4.4)

$$
\begin{aligned}
c_{i(j)\sigma_{i(j)}^1(j')}+c_{\sigma_{i(j)}^1(j')j} &\leq c_{i(j)\sigma_{i(j)}^1(j')}+c_{i(j)\sigma_{i(j)}^1(j')}+c_{i(j),\sigma_{i(j)}^{\ell}(j),j}\\
&\leq 2(1+\varepsilon)\alpha_{j'}+(1+\varepsilon)\alpha_j\\
&\leq 3(1+\varepsilon)\alpha_j.
\end{aligned}
$$

   以上不等式中，第一个不等式是由三角不等式得到。最后一个不等式成立由 $\alpha_j\geq\alpha_{j'}$ 保证。

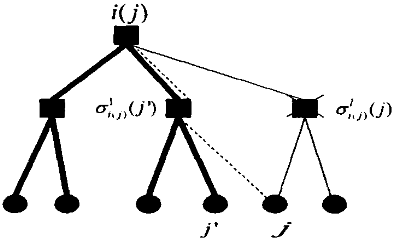


图 4.4 $i(j)$ 开设

2. 如果 $i(j)$ 不开设。
   这种情况下，必然存在一个设施 $i\in\hat{F}_2$ 和顾客 j' 使得 $j'\in D_i^1\cap D_{i(j)}^1$ (见图4.5)。注意到 $\alpha_{j'}\leq\min\{t_i^1,t_{i(j)}^1\}$。因此，$j$ 连接到 i 上的连接费用为

$$
\begin{aligned}
c_{i\sigma_i^1(j')j} &\leq c_{i\sigma_i^1(j')j'}+c_{i(j)\sigma_{i(j)}^1(j')j'}+c_{i(j)\sigma_{i(j)}^{\ell}(j)j}\\
&\leq 2(1+\varepsilon)\alpha_{j'}+(1+\varepsilon)\alpha_j.
\end{aligned}
$$

   第一个不等式成立是由于连接费用满足三角不等式。注意到 $i(j)$ 是 j 第一次连接到的开设设施，这样 $\alpha_j\geq t_{i(j)}^1\geq\alpha_{j'}$。最终，可以得到 $c_{i\sigma_i^1(j')j}\leq 3(1+\varepsilon)\alpha_j$。

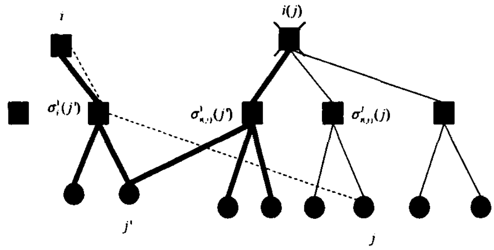


图 4.5 $i(j)$ 不开设

□

现在，可以给出近似比的分析。

定理 4.11 给定常数 $\varepsilon > 0$，令 Cost_f 和 Cost_c 分别表示算法4.7得到的设施费用和连接费用，有

$$\frac{3}{1+\varepsilon}\left(1-\frac{1}{e}\right)\mathrm{Cost}_f+\mathrm{Cost}_c \leq 3\,(1+\varepsilon)\sum_{j\in D}\alpha_j. \tag{4.8}$$

证明. 由于 $\hat{F}_2$ 中的设施是相互独立的，即 $D_i^1 \cap D_{i'}^1 = \emptyset$ $(i,i' \in \hat{F}_2)$。从引理4.9可以得到

$$\frac{3}{1+\varepsilon}\left(1-\frac{1}{e}\right)\sum_{i\in\hat{F}_2}\left(f_i+\sum_{k\in S_i^1} f_k\right)+3\sum_{i\in\hat{F}_2}\sum_{j\in D_i^1}\min_{k\in S_i^1} c_{ikj} \leq 3\,(1+\varepsilon)\sum_{i\in\hat{F}_2}\sum_{j\in D_i^1}\alpha_j.$$

再结合引理4.10，有

$$\begin{aligned}&\frac{3}{1+\varepsilon}\left(1-\frac{1}{e}\right)\sum_{i\in\hat{F}_2}\left(f_i+\sum_{k\in S_i^1} f_k\right)+3\sum_{i\in\hat{F}_2}\sum_{j\in D_i^1}\min_{k\in S_i^1} c_{ikj}+\sum_{j\in D\setminus\cup_{i\in\hat{F}_2}D_i^1}\min_{i\in\hat{F}_2}\min_{k\in\cup_{i\in\hat{F}_2}S_j^1} c_{ikj}\\ \leq\ & 3\,(1+\varepsilon)\sum_{j\in D}\alpha_j.\end{aligned}$$

引理即证。 □

在规划(4.8)中，只有当 Cost_f 前的系数大于 1，才可以得到想要的近似比。

推论 4.12 限制 $\varepsilon \leq 3(1-1/e)-1 \approx 0.8964$，算法4.7是一个原始-对偶 $3(1+\varepsilon)$-近似算法。

注释 4.13 算法4.7同样适用于顾客的服务需求为 d_j 的 2-*LFLP*上。只需在算法中修改 $\alpha_j = d_j t$ 即可。

第四节　算法改进

在本节，将讨论“加”运算对算法4.7的改进，并分析得到2-LFLP的近似比可以改进到$2.172(1+\varepsilon)^2$。具体的说，给定以个2-LFLP的可行解，令c_j是在该解中顾客j的连接费用。类似于算法4.7，对每个在$F_2\setminus\hat{F}_2$中的设施i，构造以下Max-FLP: 设施集合和顾客集合分别为F_1和D。每个设施$k\in F_1$的开设成本为$\hat{f}_k$，其中当k已经开设时，$\hat{f}_k=0$，否则，$\hat{f}_k=f_k$。每个顾客$j\in D$的连接利润为$d_{kj}=\max\{c_j-c_{ikj},0\}$。近似求解以上Max-FLP，可以得到顾客$j$连接$\sigma_i(j)$。“加”运算可以描述为: 一旦由$i$构造的Max-FLP相应近似解的纯利润超过$f_i$，开设设施$i$以及$\sigma_i(j)(\forall j\in D)$。当$d_{\sigma_i(j)j}>0$将$j$通过$\sigma_i(j)$连接到设施$i$上。

算法 4.14

步1 给定常数$\delta\in(1-\frac{1}{e},1)$，并将每个设施的费用都按比例降低$\delta$得到新的实例，即，$f_i:=\delta f_i$。

步2 对新实例应用算法4.7，得到2-LFLP的一个整数可行解。

步3 将设施费用还原为f_i，运行“加”运算直到F_2中没有能够开设的设施，即，“加”运算不能改进解的质量。

为方便起见，记算法4.14所得解的设施费用和连接费用分别为Cost_f和Cost_c，2-LFLP最优解的设施费用和连接费用分别为F^{OPT}和C^{OPT}。以下引理由Zhang[26]得到。

引理 4.15 ([26]) *如果任何的“加”运算不能改进现行解，那么*

$$\mathrm{Cost}_c\le C^{\mathrm{OPT}}+\frac{e}{e-1}F^{\mathrm{OPT}}. \tag{4.9}$$

基于以上引理，得到改进算法的近似比。

定理 4.16 *令$\delta=0.9001$，算法4.14的近似比为$2.172(1+\varepsilon)^2$。*

证明. 由不等式(4.9)可以得到

$$3(1+\varepsilon)^2\frac{1-\delta}{\delta\frac{e}{e-1}-1}\mathrm{Cost}_c\le 3(1+\varepsilon)^2\left(\frac{1-\delta}{\delta\frac{e}{e-1}-1}C^{\mathrm{OPT}}+\frac{1-\delta}{\delta\frac{e}{e-1}-1}\frac{e}{e-1}\delta F^{\mathrm{OPT}}\right).$$

结合不等式(4.8)，可以得到

$$\begin{aligned}&3\left(1-\frac{1}{e}\right)\delta\mathrm{Cost}_f+\left(1+3(1+\varepsilon)^2\frac{1-\delta}{\delta\frac{e}{e-1}-1}\right)\mathrm{Cost}_c\\ \le\ &3(1+\varepsilon)^2\left(\frac{1-\delta}{\delta\frac{e}{e-1}-1}+1\right)\left(C^{\mathrm{OPT}}+F^{\mathrm{OPT}}\right).\end{aligned}$$

选择 $\delta = 0.9001$ 可以得到

$$3\left(1-\frac{1}{e}\right)\delta = 1+3\frac{1-\delta}{\delta\frac{e}{e-1}-1},$$

即得到近似比 $2.172(1+\varepsilon)^2$。 $\square$

第五节 结论

在本章，我们考虑了 2-层设施选址问题。基于经典的原始-对偶算法过程，近似求解 NP-困难的分离神谕得到 $3(1+\varepsilon)$-近似算法，其中 $\varepsilon>0$。这一结论与单层设施选址问题，即，无容量限制的设施选址问题的原始-对偶算法得到的结果基本相同，而一般 k-层设施选址问题的原始-对偶算法近似比为 6。从下界的角度，这两类问题的难度不同，2-层设施选址问题近似比的下界为 1.539 而无容量限制的设施选址问题近似比的下界为 1.463。再通过引入"加"运算，可以将近似比改进为 $2.172(1+\varepsilon)^2$。

由于原始-对偶算法的良好可移植性，可以将算法4.7应用到 2-层设施选址问题的各种变形，例如，带惩罚的 2-层设施选址问题，包括带线性惩罚以及带次模惩罚，带软容量的 2-层设施选址问题，随机 2-层设施选址问题等等。

此外，本章提出的近似分离神谕的思想不能推广到 k-层设施选址问题。这是因为，当考虑 k-层设施选址问问题时，其分离神谕是极大化的 $(k-1)$-层设施选址问题，但这类问题的近似比定义(见1.2.5节)与传统的近似比不同，这种近似比不能应用到近似求解分离神谕过程中。同时，极大化的 $k-1$-层设施选址问题不能视为在背包约束下的次模函数极小化问题。因此，未来如何改进 k-层设施选址问题的近似比是一个有趣的问题。

此页不缺内容

第五章　带软容量限制的 k-层设施选址问题

第一节　问题介绍

在本章，我们考虑带软容量限制的 k-层设施选址问题(k-level facility location problem with soft capacities，简记 k-LFLPWSC)。在 UFLP中，假设每个设施能提供的服务是无限的。但实际中由于资源限制，这种假设并不实际，每个设施都有一个其所能够提供服务的容量上界 u_i。所谓软容量就是每个设施可以通过多次开设，支付多次开设费用来获得多个容量。具体的说，k-LFLPWSC可以描述为: 给定 $k+1$-部完全图 $(F_1\cup F_2\cup\cdots\cup F_k\cup D,E)$，其中，设施是一些具有层次关系的，$F_l(l=1,2,\cdots,k)$ 是第 l 层设施的集合，D 是顾客的集合。这些具有层次关系的设施构成路以为顾客服务，给定 $P:=F_1\times F_2\times\cdots\times F_k$ 为路的集合。同时，令 $F=F_1\cup F_2\cup\cdots\cup F_k$ 为所有设施的集合。图中任意两点 s,t 的距离为 c_{st}，因此 j 到 $p=(i_1,i_2,\cdots,i_k)\in P$ 的距离定义为: $c_{pj}=c_{i_1j}+\sum_{l=1}^{k-1}c_{i_li_{l+1}}$。每个设施 i 可以以开设费用 f_i 获得 u_i 的容量。若设施 i 可以开设了 y 次，那么则需要支付 yf_i，得到的容量为 yu_i。需要在每层设施中开设一些设施，构成开设路，将顾客连接到开设的路上，并且不超过每个设施的容量，使得开设费用和连接费用总和最小。注意到由于图是度量的，因而对任意顾客 $j,j'\in D$ 以及路径 $p,p'\in P$ 都满足 $c_{pj}\le c_{pj'}+c_{p'j'}+c_{p'j}$。

不同于第四章提出的集合覆盖模型，对设施 i 引入整数变量 y_i，$y_i=1$ 表示设施 i 开设，否则不开设。对顾客 j 是否连接到路 p 引入0-1变量 x_{pj}，$x_{pj}=1$ 表示顾客 j 连接到路径 p 上，否则表示不连接。因此，可以得到以下 k-LFLPWSC的整数线性规划:

$$
\begin{aligned}
\min\quad & \sum_{i\in F} f_iy_i+\sum_{p\in P}\sum_{j\in D}c_{pj}x_{pj} \\
\text{s. t.}\quad & \sum_{p\in P}x_{pj}\ge 1, \qquad \forall j\in D, \\
& \sum_{p\in P:i\in p}x_{pj}\le y_i, \qquad \forall p\in P, j\in D, \\
& \sum_{p\in P:i\in p}\sum_{j\in D}x_{pj}\le u_iy_i, \qquad \forall i\in F,
\end{aligned}
\tag{5.1}
$$

$$
\begin{aligned}
& x_{pj} \in \{0,1\}, \qquad \forall p \in P, j \in D, \\
& y_i \in \mathbb{N}.
\end{aligned}
$$

在上述规划中，$i \in p$ 表示路径 p 通过的设施 i。第一组约束表示顾客 j 必须连接到一条路径上。第二组约束表示如果顾客 j 连接到经过设施 i 上的路 p 上，那么设施 i 必须开设。第三组约束表示所有连接通过设施 i 的路径的顾客不能超过设施 i 的容量限制 u_i。其相应的线性规划松弛为

$$
\begin{aligned}
\min \quad & \sum_{i \in F} f_i y_i + \sum_{p \in P} \sum_{j \in D} c_{pj} x_{pj} \\
\text{s. t.} \quad & \sum_{p \in P} x_{pj} \geq 1, \qquad \forall j \in D, \\
& \sum_{p \in P: i \in p} x_{pj} \leq y_i, \qquad \forall p \in P, j \in D, \\
& \sum_{p \in P: i \in p} \sum_{j \in D} x_{pj} \leq u_i y_i, \qquad \forall i \in F, \\
& x_{pj}, y_i \geq 0 \qquad \forall i \in F, p \in P, j \in D.
\end{aligned} \tag{5.2}
$$

定义 5.1 *一个 k-层设施选址问题的算法被称为 (γ_f, γ_c)-近似算法，如果由算法所得解的设施费用 Cost_f 和 Cost_c 总和不超过 $\gamma_f F^* + \gamma_c C^*$，其中 F^* 和 C^* 为(5.2)的分数最优解中的设施费用和连接费用，即，其中*

$$\mathrm{Cost}_f + \mathrm{Cost}_c \leq \gamma_f F^* + \gamma_c C^*$$

(γ_f, γ_c) 被称为算法的双因子。

在本章，将带软容量限制的 k-层设施选址问题归约到 k-层设施选址问题，并且的得到两个问题双因子近似比的关系。因此，再对 k-层设施选址问题设计基于线性规划舍入技巧的算法，得到 k-LFLPWSC的整数可行解，最终分析得到近似比为 5.5053。

第二节 k-层设施选址问题的双因子近似算法

如果带软容量限制的 k-层设施选址问题中，所有设施的容量限制均为无穷，则该问题是 k-层设施选址问题(k-level facility location problem，简记 k-LFLP)。类似的，可以得到 k-LFLP的整数线性规划。

$$\min \quad \sum_{i \in F} f_i y_i + \sum_{p \in P} \sum_{j \in D} c_{pj} x_{pj}$$

$$\begin{aligned}\text{s. t.}\quad & \sum_{p\in P} x_{pj}\geq 1, && \forall j\in D, \\ & \sum_{p\in P: i\in p} x_{pj}\leq y_i, && \forall p\in P, j\in D, \\ & x_{pj}, y_i\in\{0,1\}, && \forall i\in F, p\in P, j\in D.\end{aligned} \tag{5.3}$$

对应的，其松弛为

$$\begin{aligned}\min\quad & \sum_{i\in F} f_i y_i+\sum_{p\in P}\sum_{j\in D} c_{pj}x_{pj} \\ \text{s. t.}\quad & \sum_{p\in P} x_{pj}\geq 1, && \forall j\in D, \\ & \sum_{p\in P: i\in p} x_{pj}\leq y_i, && \forall p\in P, j\in D, \\ & x_{pj}, y_i\geq 0, && \forall i\in F, p\in P, j\in D.\end{aligned} \tag{5.4}$$

在算法中，首先引入参数 $\alpha\in(0,1)$，将规划(5.4)的最优解改造成另一解。新的解可以平衡设施费用和连接费用。之后，类似UFLP中选择中心。最后，将所有顾客聚类，以及在每个聚类中只开设一条路径。

算法 5.2

步0 给定参数 $\alpha\in(0,1)$。

步1 求解规划(5.4)的最优分数解 $(\tilde{x},\tilde{y})$。

步2 对每个顾客 $j\in D$，定义分数连接到顾客 j 的路径集合为 $P_j:=\{j:\tilde{x}_{pj}>0\}$。将 P_j 中的路径按 c_{pj} 递增的次序得到 P_j 中路径的排序 π。计算

$$s^*:=\min\left\{s:\sum_{s'\leq s}\tilde{x}_{p(\pi(s'))j}\geq\alpha\right\}.$$

令 $c_j(\alpha):=c_{p()\pi(s^*)j}$ 以及 $\beta_j^{\alpha}=\sum_{p:x_{pj}\leq c_j(\alpha)}\tilde{x}_{pj}$。按以下规则将 $(\tilde{x},\tilde{y})$ 转化成另一个解 $(\bar{x},\bar{y})$。

$$\bar{x}_{pj}:=\begin{cases}\frac{\tilde{x}_{pj}}{\beta_j^{\alpha}}, & \text{如果 } c_{pj}\leq c_j(\alpha), \\ 0, & \text{否则};\end{cases}\qquad\qquad \bar{y}_i:=\min\left\{\frac{\tilde{y}_i}{\alpha},1\right\}.$$

在解 $(\bar{x},\bar{y})$ 下，顾客 j 的分数连接费用记为 $\bar{C}_j:=\sum_{p\in P}c_{pj}\bar{x}_{pj}$。

步3 将分数解 $(\bar{x},\bar{y})$ 舍入成整数可行解 $(\hat{x},\hat{y})$。

步3.0 对每个顾客 j，定义分数连接的路径集合为邻居，即，$N_j:=\{p\in P:\bar{x}_{pj}>0\}$。$N_j$ 中第 l 层($l=1,2,\cdots,k$)的设施集合为 $N_j^l:=\cup_{p\in N_j}\{i\in F_l:$

$i \in p\}$。令 D' 和 U 分别表示中心集合和未被聚类的顾客的集合。初始时，$D' := \emptyset$ 以及 $U := D$。

步3.1 在 U 中选择最小的 $c_j(\alpha)+\bar{C}_j$ 作为中心，即，选择 $j' := \arg\min_{j\in U}\{c_j(\alpha)+\bar{C}_j\}$。令分数连接到 $N_{j'}$ 的顾客集合为 $B'_j := \{j \in U : \exists l \in \{1,2,\cdots,k\}, N_j^l \cap N_{j'}^l \neq \emptyset\}\}$。更新中心集合 $D' := D' \cup \{j'\}$，未被聚类集合 $U := U \setminus B'_j$。称 $(j', N_{j'}, B_{j'})$ 为以 j' 中心的聚类。

步3.2 重复步3.1直到 $U = \emptyset$。

步3.3 在每个聚类 $(j', N_{j'}, B_{j'})$ 中，以 $\bar{x}_{pj}$ 的相关概率随机开设路径 $p \in N_{j'}$ 上的设施。再将每个顾客连接到开设的最近路径。

注意到，$c_j(\alpha)$ 是 N_j 中距离 j 最远路径的距离。接下来，将分别分析设施费用和连接费用的期望以获得近似比。首先，分析设施费用的期望。

引理 5.3 *算法(5.2)得到的设施费用期望至多为* $\frac{1}{\alpha}\sum\limits_{i\in F} f_i\bar{y}_i$。

证明. 考虑以 j' 为中心的聚类。注意到路径 $p \in N_{j'}$ 开设的概率为 $\bar{x}_{pj'}$，而 $\sum_{p\in N_{j'}}\bar{x}_{pj'} = 1$，因此，在这个聚类中设施费用的期望为

$$
\begin{aligned}
\sum_{p\in N_{j'}}\left(\sum_{i\in p} f_i\right)\bar{x}_{pj'} &= \sum_{i\in\bigcup_l N_{j'}^l}\left(\sum_{p\in P: i\in p}\bar{x}_{pj'}\right) f_i \\
&\leq \sum_{i\in\bigcup_l N_{j'}^l} f_i\bar{y}_i \\
&\leq \frac{1}{\alpha}\sum_{i\in\bigcup_l N_{j'}^l} f_i\bar{y}_i.
\end{aligned}
$$

第一个等式由求和符号交换得到。第二个不等式由 $(\bar{x},\bar{y})$ 是可行解得到，即，$\bar{x}_{pj} \leq \bar{y}_i$。最后一个不等式由 $\bar{y}_i$ 的定义可得。

最终，对所有的聚类中心求和可以得到引理。 □

其次，考虑连接费用的期望。

引理 5.4 *对每个顾客* $j \in D$，*有*

$$c_j(\alpha) \leq \frac{1}{1-\alpha}\sum_{p\in P} c_{pj}\tilde{x}_{pj}.$$

证明. 由 $c_j(\alpha)$ 的定义，有

$$\sum_{p:c_{pj}\leq c_j(\alpha)} \tilde{x}_{pj} \geq \alpha \qquad \text{和} \qquad \sum_{p:c_{pj}<c_j(\alpha)} \tilde{x}_{pj} < \alpha.$$

因此，

$$\sum_{p:c_{pj}\geq c_j(\alpha)} \tilde{x}_{pj} \geq 1-\alpha.$$

这意味

$$c_j(\alpha) \leq \frac{1}{1-\alpha} \sum_{p:c_{pj}\geq c_j(\alpha)} c_{pj}\tilde{x}_{pj} \leq \frac{1}{1-\alpha} \sum_{p\in P} c_{pj}\tilde{x}_{pj}.$$

引理即证。 □

现在，可以估计算法得到连接费用。

引理 5.5 *算法5.2得到的连接费用期望不超过*

$$\left(1+\frac{2}{1-\alpha}\right) \sum_{p\in P} \sum_{j\in D} c_{pj}\tilde{x}_{pj}.$$

证明. 对每个顾客 $j\in D$，考虑两种情况。

情况1. $j\in D'$。

这种情况下，N_j 中一定存在一个开设路径。将 j 连接到这个路径上。因此，j 的连接费用期望至多为

$$\bar{C}_j = \sum_{p\in N_j} c_{pj}\tilde{x}_{pj} \leq \sum_{p\in P} c_{pj}\tilde{x}_{pj}.$$

情况2. $j\notin D'$。

这种情况下，一定存在一整个中心 $j'\in D'$ 以及 $l\in\{1,2,\cdots,k\}$ 使得 $N^l_{j'}\cap N^l_j\neq\emptyset$，即，存在路径 $p\in N_j$ 和 $p''\in N_{j'}$ 中经过相第 l 层相同的设施(见图5.1)。将 j 连接到以 j' 为中心的聚类中开设的路径 $p'(j')$ 上。由三角不等式，有

$$\begin{aligned}
E[c_{p'(j')j}] &\leq c_{pj}+c_{p''j'}+E[c_{p'(j')j'}] \\
&\leq c_j(\alpha)+c_{j'}(\alpha)+\bar{C}_{j'} \\
&\leq 2c_j(\alpha)+\bar{C}_j \\
&\leq \left(1+\frac{2}{1-\alpha}\right)\sum_{p\in P} c_{pj}\tilde{x}_{pj}.
\end{aligned}$$

综合以上两种情况，引理即证。 □

最终，由引理5.4和5.5可得到 k-层设施选址问题的双因子近似比。

定理 5.6 *对任意的 $\alpha\in(0,1)$，算法5.2是 k-LFLP的 $\left(\frac{1}{\alpha},1+\frac{2}{1-\alpha}\right)$-近似算法。*

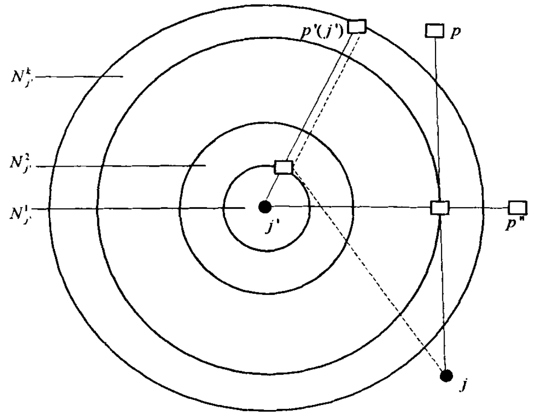


图 5.1 $j \notin D'$

在算法中，α 是固定的常数。事实上，可以通过随机选取参数 α 得到更好的双因子算法。

算法 5.7

步0 给定参数 $\beta \in (0,1)$。
步1 取 α 为在 $(\beta,1)$ 中均匀分布的参数，以 α 为参数运行算法5.2。

类似文献[8]，可以估计 $c_j(\alpha)$。

引理 5.8

$$\int_0^1 c_j(\alpha)d\alpha = \sum_{p\in P} c_{pj}\tilde{x}_{pj}.$$

证明. 将 P_j 中的路径按 c_{pj} 递增的顺序排序得到 π。定义 $r := |\{p : \tilde{x}_{pj} > 0\}|$ 为在分数解 $(\bar{x},\bar{y})$ 中分数连接的路径个数。很容易验证到对任意 $r' = 1,2,\cdots,r$，当 $\alpha \in \left(\sum_{s=1}^{r'-1} x_{p_s j}, \sum_{s=1}^{r'} x_{p_s j}\right]$ 时，$c_j(\alpha) = c_{p_{r'}j}$。因此 $c_j(\alpha)$ 为一个阶梯函数(见图5.2)，引理即证。

□

从引理5.3和5.8，可以得到设施的费用和连接费用的期望分别不超过 $\frac{1}{\alpha}\sum_{i\in F} f_i\bar{y}_i$ 和 $2c_j(\alpha) + \sum_{p\in P}\sum_{j\in D} c_{pj}\tilde{x}_{pj}$。因此，可以得到

定理 5.9 *算法5.7对 k-层设施选址问题的一个 $\left(\frac{\ln\frac{1}{\beta}}{1-\beta}, 1+\frac{2}{1-\beta}\right)$-近似算法。*

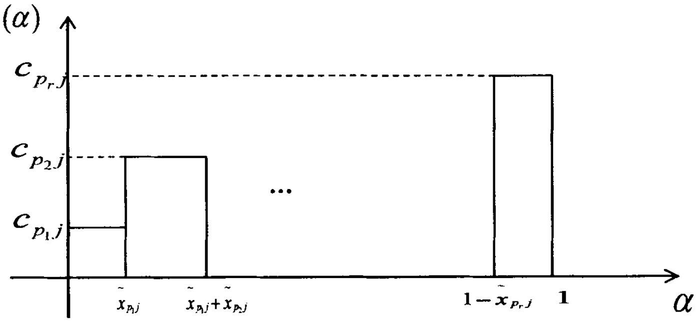


图 5.2 $c_j(\alpha)$

证明. 由引理5.3，设施开设费用的期望不超过

$$E\left[\frac{1}{\alpha}\sum_{i\in F} f_i\tilde{y}_i\right] = \left(\int_{\beta}^{1}\frac{1}{1-\beta}\frac{1}{\alpha}d\alpha\right)\sum_{i\in F} f_i\tilde{y}_i = \frac{\ln\frac{1}{\beta}}{1-\beta}\sum_{i\in F} f_i\tilde{y}_i.$$

由引理5.8，连接费用的期望不超过

$$\begin{aligned} E\left[\sum_{j\in D}\bar{C}_j + 2\sum_{j\in D} c_j(\alpha)\right] &= \sum_{j\in D}\bar{C}_j + 2\sum_{j\in D}\int_{\beta}^{1}\frac{1}{1-\beta}c_j(\alpha)d\alpha \\ &\le \sum_{j\in D}\bar{C}_j + \frac{2}{1-\beta}\sum_{j\in D}\int_{0}^{1} c_j(\alpha)d\alpha \\ &= \left(1+\frac{2}{1-\beta}\right)\sum_{p\in P}\sum_{j\in D} c_{pj}\tilde{x}_{pj}. \end{aligned}$$

□

第三节 带软容量限制的 k-层设施选址问题的归约

在本节，我们将带软容量的 k-层设施选址问题归约至 k-层设施选址问题，并分析出 k-层设施选址问题的双因子近似比可以转化为带软容量 k-层设施选址问题近似比。首先，对给定的常数 $\lambda \in [0,1]$，从规划(5.2)的第三组约束 $\sum_{p\in P: i\in p}\sum_{j\in D} x_{pj} \le u_i y_i$，以及假设 (x,y) 是(5.2)的最优解，可以得到

$$\sum_{i\in F}\lambda f_i \frac{\sum_{p\in P:i\in p}\sum_{j\in D} x_{pj}}{u_i} \le \sum_{i\in F}\lambda f_i y_i.$$

因此，将第三组约束乘以朗格朗日乘子 λ 作为惩罚项放到目标中，则以下规划是(5.1)的下界。

$$\begin{array}{lll}\min & \sum_{i\in F}(1-\lambda)f_iy_i+\sum_{p\in P}\sum_{j\in D}\left(c_{pj}+\lambda\cdot\sum_{i\in p}\frac{f_i}{u_i}\right)x_{pj} & \\ \text{s. t.} & \sum_{p\in P}x_{pj}\geq 1, \qquad \forall j\in D, & \\ & \sum_{p\in P: i\in p}x_{p,j}\leq y_i, \qquad \forall i\in F, j\in D, & (5.5)\\ & x_{pj}\in\{0,1\}, \qquad \forall p\in P, j\in D, & \\ & y_i\in\mathbb{N}, \qquad \forall i\in F. & \end{array}$$

事实上，规划(5.5)可以看做 k-LFLP的实例，其中设施 $i\in F$ 的开设费用为 $1-\lambda f_i$，顾客 $j\in D$ 到第1层设施 $i_1\in F_1$ 的连接费用为 $\bar{c}_{i_1j}=c_{i_1j}+\lambda\frac{f_{i_1}}{u_{i_1}}$，第 $l-1$ 层设施 $i_{l-1}\in F_{l-1}$ 到第 l 层设施 $i_l\in F_l$ 这间的连接费用为 $\bar{c}_{i_{l-1}i_l}:=c_{i_{l-1}i_l}+\lambda\frac{f_{i_l}}{u_{i_l}}$。注意到 $\bar{c}_{pj}:=c_{pj}+\lambda\cdot\sum_{i\in p}\frac{f_i}{u_i}$。首先 $\bar{c}_{pj}$ 是非负的，对称的。下面说明 $\bar{c}_{pj}$ 仍然满足三角不等式。

引理 5.10 *连接费用 $\bar{c}_{pj}$ 仍满足三角不等式。*

证明. 对任意的顾客 $j,j'\in D$ 以及设施路径 $p,p'\in P$，

$$\bar{c}_{pj}=c_{pj}+\lambda\cdot\sum_{i\in p}\frac{f_i}{u_i}.$$

由 c_{pj} 满足三角不等式，有

$$c_{pj}\leq c_{pj'}+c_{p'j'}+c_{pj'}.$$

因此，

$$\begin{aligned}\bar{c}_{pj} &\leq c_{pj'}+c_{p'j'}+c_{pj'}+\lambda\cdot\sum_{i\in p}\frac{f_i}{u_i}\\ &\leq c_{pj'}+\lambda\cdot\sum_{i\in p}\frac{f_i}{u_i}+c_{p'j'}+\lambda\cdot\sum_{i\in p'}\frac{f_i}{u_i}+c_{pj'}+\lambda\cdot\sum_{i\in p}\frac{f_i}{u_i}\\ &= \bar{c}_{pj'}+\bar{c}_{p'j'}+\bar{c}_{pj'}.\end{aligned}$$

引理即证。 □

现在，将给出从 k-层设施选址问题的解转化为带软容量问题的可行解问题。

算法 5.11

步0 给定参数 $\lambda \in [0,1]$ 以及 $\delta \in (0,+\infty)$。

步1 输入 k-*LFLPWSC*的实例，按照上述方法将其转化为 k-*LFLP*。

步2 将设施的费用按比例提高 δ 倍，应用 k-层设施选选址问题的 (γ_f,γ_c)-近似算法，得到 (x,y)。

步3 令

$$\hat{x}_{ij} := x_{ij}, \qquad \hat{y}_i := \left\lceil \frac{\sum\limits_{p\in\mathscr{P}:i\in p}\sum\limits_{j\in D} x_{ij}}{u_i} \right\rceil, \qquad \forall i \in F, j \in D.$$

输出 $(\hat{x},\hat{y})$ 作为 k-*LFLPWSC*的解。

算法5.11给出了一种从 k-LFLP的解构造 k-LFLPWSC解的方法。

引理 5.12 令 $\lambda = \frac{\gamma_c}{\gamma_f+\gamma_c}$ 以及 $\delta := \frac{\gamma_c}{\gamma_f}$，算法5.11是一个 k-*LFLPWSC*的 $(\gamma_f+\gamma_c)$-近似算法。

证明. 记规划(5.4)的最优值为 $F_{k-LFLP}+C_{k-LFLP}$，其中 F_{k-LFLP} 是设施费用，C_{k-LFLP} 是连接费用。注意到算法5.11步2所得 k-LFLP解的设施开设费用为

$$F_{SOL}^{k-LFLP} := \sum_{i\in F}(1-\lambda)f_i y_i$$

以及连接费用为

$$C_{SOL}^{k-LFLP} := \sum_{p\in P}\sum_{j\in D}\left(c_{pj}+\lambda\sum_{i\in p}\frac{f_i}{u_i}\right)x_{pj}.$$

因此，假设由算法5.11步2所得解 SOL的开设费用(未按比例提高)和连接费用费用分别为 F_{SOL}^{k-LFLP} 和 C_{SOL}^{k-LFLP}，有

$$\delta F_{SOL}^{k-LFLP}+C_{SOL}^{k-LFLP} \quad = \quad \delta\sum_{i\in F}(1-\lambda)f_i y_i+\sum_{p\in P}\sum_{j\in D}\left(c_{pj}+\lambda\sum_{i\in p}\frac{f_i}{u_i}\right)x_{pj}.$$

假设由算法5.11所得 k-LFLPWSC的解的设施费用为

$$F := \sum_{i\in F} f_i \left\lceil \frac{\sum\limits_{p\in P:i\in p}\sum\limits_{j\in D}\hat{x}_{pj}}{u_i} \right\rceil \le \sum_{i\in F} f_i\left(\hat{y}_i+\frac{\sum\limits_{p\in P:i\in p}\sum\limits_{j\in D}\hat{x}_{pj}}{u_i}\right),$$

连接费用为

$$C := \sum_{p\in P}\sum_{j\in D} c_{pj}\hat{x}_{pj}.$$

因此，有

$$\begin{aligned}
&\delta F_{SOL}^{k-LFLP} + C_{SOL}^{k-LFLP} \\
= &\ \frac{\gamma_c}{\gamma_f+\gamma_c}\sum_{i\in F}\left(\frac{\sum_{p\in P:i\in p}\sum_{j\in D}x_{pj}}{u_i}+\hat{y}_i\right)f_i+\sum_{p\in P}\sum_{j\in D}c_{pj}x_{pj} \\
\geq &\ \frac{\gamma_c}{\gamma_f+\gamma_c}F+C.
\end{aligned}$$

由调用了 k-LFLP的 (γ_f,γ_c)-近似算法得到 SOL，可以得到

$$\begin{aligned}
\delta F_{SOL}^{k-LFLP} + C_{SOL}^{k-LFLP} &= \frac{\gamma_c}{\gamma_f}F_{SOL}^{k-LFLP}+C_{SOL}^{k-LFLP} \\
&\leq \gamma_f\cdot\frac{\gamma_c}{\gamma_f}F_{k-LFLP}+\gamma_c\cdot C_{k-LFLP} \\
&= \gamma_c(F_{k-LFLP}+C_{k-LFLP}) \\
&\leq \gamma_c OPT_{k-LFLPSC},
\end{aligned}$$

其中 $OPT_{k-LFLPSC}$ 是 k-LFLPWSC的最优值。注意到最后一个不等式成立是由于按构造方式，k-LFLP的解是 k-LFLPWSC的下界。因此，

$$\begin{aligned}
F+C &\leq F+\frac{\gamma_f+\gamma_c}{\gamma_c}C \\
&\leq \frac{\gamma_f+\gamma_c}{\gamma_c}\left(\delta F_{SOL}^{k-LFLP}+C_{SOL}^{k-LFLP}\right) \\
&\leq (\gamma_f+\gamma_c)OPT_{k-LFLPWSC}.
\end{aligned}$$

□

结合算法5.7和算法5.11，可以得到改进的带软容量 k-LFLPWSC的近似算法。

定理 5.13 *当在算法5.7中取值 $\beta=0.2222$，算法5.11是 k-LFLPWSC的* 5.5053-*近似算法。*

证明. 从引理5.9和引理5.12可以得到，k-LFLPWSC的近似比为

$$\min_{\beta\in(0,1)}\left\{\frac{\ln\frac{1}{\beta}}{1-\beta}+1+\frac{2}{1-\beta}\right\}.$$

令 $\beta := 0.2222$ 引理即证。 □

第四节 结论

在本章，我们考虑了带软容量限制的 k-层设施选址问题。为了改进带软容量限制的 1-层设施选址问题的近似比，将 Mahdian等 [29] 关于 1-LFLPWSC的归约的结论推广到多层。使用这种归约，建立带软容量限制的 k-层设施选址问题和 k-层设施选址问题之间的联系，并得到带软容量限制的 k-层设施选址问题的 5.5053-近似算法。由于 1.61 是 k-层设施选址问题的近似比下界，因而也可视为带软容量限制的 k-层设施选址问题的近似比下界。未来，如何缩小带软容量 k-层设施选址问题近似比 5.5053 与 1.61 之间的间隙是一个值得研究的问题。另一个具有挑战性的问题是如何降低 k-层设施选址 3-近似算法，这个问题的解决有助于处理带软容量限制的问题。

此页不缺内容

第六章 总结

本论文主要研究了两阶段随机设施选址问题，带线性惩罚的两阶段随机设施选址问题，2-层设施选址问题，和带软容量限制的 k-层设施选址问题。

这些问题在供应链管理,生产决策等方面有重要应用，是经典计算机理论中的问题。由于上述问题都是 NP-困难的，因而利用近似算法研究这些问题。本论文利用线性规划舍入，线性规划原始-对偶方法，归约等技巧设计了不同的近似算法，得到的主要结果如下:

1. 对两阶段随机设施选址问题利用线性规划舍入的技巧得到 2.3613 的单场景界，这是目前为止随机设施选址问题最好的单场景界。
2. 对带线性惩罚的两阶段随机设施选址问题基于线性规划舍入设计了第一个常数单场景界的近似算法，分析得到近似比为 3.0294。
3. 对 2-层选址问题，基于近似分离神谕的设计了原始-对偶的 3-近似算法。基于原始-对偶技巧良好的可移植性，可以将这种原始-对偶的技巧应用到 2-层设施选址问题的各种变形。
4. 对带软容量限制的 k-层设施选址问题，利用双因子归约的技巧可以得到 5.5053-近似算法，这是目前为止该问题最好的近似比。

重要待研究的问题 无容量设施选址问题的近似比下界为 1.463，这也自然是无容量设施选址变形问题的下界。目前为止已知的无容量设施选址问题近似比的上界是 1.488，近似比的间隙约为 0.025，表明人们对无容量设施选址问题的近似可解性有了很好的理解。本论文分别对随机优先设施选址问题和随机容错设施选址问题设计了 1.8526 和 5 近似算法，这与下界 1.463 之间还有较大的间隙，研究能否缩小这些问题的近似比间隙将是有意义的问题。

此页不缺内容

参考文献

[1] Papadimitriou C H. Computational Compexity[M]. Addison-Weslsy, 1994.

[2] Karp R M. Reducibility among combinatorial problems[M]//Miller R E, Thatcher J W, eds. Complexity of Computer Computations. New York: Plenum, 1972: 85-103.

[3] Williamson D P, Shomys D B. The Design of Approximation Algorithms[M]. Cambridge University Press, New York, 2011.

[4] 堵丁柱, 葛可一, 胡晓东. 近似算法的设计与分析[M]. 北京: 高等教育出版社, 2011.

[5] Vazirani V V. Approximation Algorithms[M]. Berlin: Springer-Verlag, 2001.

[6] 徐大川, 张家伟. 设施选址问题的近似算法[M]. 北京: 科学出版社, 2013.

[7] Hochbaum D S. Heuristics for the fixed cost median problem[J]. Mathematical Programming, 1982, 22: 148-162.

[8] Shmoys D B, Tardös E, and Aardal K I. Approximation algorithms for facility location problems[C]// Proceedings of the 29th annual ACM symposium on Theory of computing, NY, USA: ACM New York, 1997: 265-274.

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

[10] Sviridenko M. An Improved Algorithm for the metric uncapacited facility location problem[C]//Proceedings of the 9th International Conference of Integer Programming and Combinatorial Optimization, Springer-Verlag Berlin Heidelberg: Springer Berlin Heidelberg, 2002: 240-257.

[11] Jain K, Vazirani V V. Approximation algorithms for metric facility location and k-median problems using the primal-dual schema and Lagrangian relaxation[J]. Journal of the ACM, 2001, 48: 274-296.

[12] Charilar M, Guha S. Improved combinatorial algorithms for facility location problems[J]. SIAM Jounral on Computing, 2005, 34: 803-824.

[13] Jain K, Mahdian M, Markakis E, Saberi A, Vazirani V V. Greedy facility location

algorithm analyzed using dual fitting with factor-revealing LP[J]. Jounal of the ACM, 2003, 50: 795-824.

[14] Mahdian M, Ye Y Y, Zhang J W. Approxiamtion Algorithms for metric facility location problems[J]. SIAM Journal on Computing, 2006, 36: 411-432.

[15] Arya V, Garg N, KHandekar R, Meyerson A, Munagala K, Pandit V. Local search heuristics for k-median and facility location problems[J]. SIAM Journal on Computing, 2004, 33: 544-562.

[16] Byrka J, Aardal K. An optimal bifactor approximation algorithm for the metric uncapacitated facility location problem[J]. SIAM Journal on Computing, 2010, 39: 2212-2231.

[17] Li S. A 1.488-approximation algorithm for the uncapacitated facility location problem[J]. Information and Computation, 2013, 222: 45-58.

[18] Guha S, Khuller S. Greedy strike back: improved facility location algorithms[J]. Journal of Algorithms, 1999, 31: 228-248.

[19] Sviridenko M. Cited as personal communication in [9], July, 1998.

[20] Meyerson A, Munagala K, Plotkin S. Cost-distance: two-metric network design[C]//Proceedings of 41st Annual Symposium on Foundations of Computer Science, IEEE Xplore, 2000. 624-630.

[21] Ageev A A, Ye Y Y, Zhang J W. Improved combinatorial approximation algorithms for the k-level facility location problem[J]. SIAM Journal on Discrete Mathematics, 2004, 18: 207-217.

[22] Aardal K I, 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.

[23] Gabor A, van Ommeren J. A new approximation algorithm for the multilevel facility location problem[J]. Discrete Applied Mathematics, 2010, 158: 453-460.

[24] Byrka J, Rybicki B. Improved LP-rounding approximation algorithm for k-level uncapacitated facility location[C]// Proccedings of the 39th International Colloquium on Automata, Languages and Programming, Springer-Verlag Berlin Heidelberg: Springer Berlin Heidelberg, 2012: 157-169.

[25] Bumb A, Kern W. A simple dual ascent algorithm for the multilevel facility lo-

cation problem[C]//Proceedings of the 4th international Workshop on Approximation Algorithms for Combinatorial Optimization Problmes and the 5th Internatonal Workshop on Randomnization and Approximation Techniques in Computer Science, Springer-Verlag Berlin Heidelberg: Springer Berlin Heidelberg, 2001: 55-63.

[26] Zhang J W. Approximating the two-level facility location problem via a quasi-greedy approach[J]. Mathematical Programming, 2006, 108: 159-176.

[27] Krishnaswamy R, Sviridenko M. Inapproximability of the multilevel uncapacitated facility location problem[C]//Proceedings of the 23rd annual ACM-SIAM symposium on Discrete Algorithms. SIAM, 2012: 718-734.

[28] Chudak F A, Shmoys D B. Improved approximation algorithms for the capacitated facility location problem[C]//Proceedings of the 10th annual ACM Symposium on Theory of Computing, NY, USA: ACM New York, 1999: 875-876.

[29] Mahdian M, Ye Y Y, Zhang J. Improved approximation algorithms for metric facility location problems[C]// Proceedings of 5th International Workshop on Approximation Algorithms for Combinatorial Optimization, 2002: 229-242.

[30] Mahdian M, Ye Y Y, Zhang, J. A 2-approximation algorithm for the soft-capacitated facility location problem[C]// Proceedings of the 6th International Workshop on Approximation Algorithms for Combinatorial Optimization, 2003: 149-162.

[31] Korupolu M, Plaxton C, Rajaraman R. Analysis of a local search heuristic for facility locaiton problem[J]. Journal of Algorithms, 2000, 37: 146-188.

[32] Chudak F A, Williamson D P. Improved approximation algorithms for capacitated facility location problems[J]. Mathematical Programming. 2005, 102: 207-222.

[33] Pal M, Tardös E, Wexler T. Facility location with nonuniform hard capacities[C]//Proccedings of 42nd IEEE Symposium on Foundations of Computer Science, IEEE Xplore, 2001: 329-338.

[34] Zhang J, Chen B, Ye Y Y. A multiexchange Local Search Algrorithm for the Capacitated Facility Locaiton Problem[J]. Mathematics of Operations Research, 2005, 30: 389-403.

[35] Ravi R, Sinha A. Hedging uncertainty: approximation algorithhms for stochastic

optimization problems[J]. Mathematical Programming, 2006, 108: 97-114.

[36] Shmoys D B, Swamy C. An approximation scheme for stochastic linear programming and its application to stochastic integer programs[J]. Journal of the ACM, 2006, 53: 978-1012.

[37] Srinivasan A. Approximation algorithms for stochastic and risk-averse optimization[C]// Proceedings of 18th annual ACM-SIAM symposium on Discrete algorithms, PA, USA: Society for Industrial and Applied Mathematics Philadelphia, 2007: 1305-1313.

[38] Byrka J, Ghodsi M, Sirnivasan A. LP-rounding algorithms for facility-location problems[DB]. Conrnell Univerisity Library, http://arXiv.org/abs/1007.3611.

[39] Ye Y Y, Zhang J W. An approximation algorithm for the dynamic facility location problem[M]// Cheng X Y, Li Y S, Du D Z. Combinatorial Optimization in Communication Networks, Kluwer Academic Publishers, 2005: 623-637.

[40] Swamy C. Approximation algorithms for clustering problems[D]. Cornell University, 2004.

[41] Charikar M, Khuller S, Mount D M, Naraasimban G, Algorithms for facility location problems with outliers[C]// Proceedings of the 12th annual ACM-SIAM symposium on Discrete algorithms, PA, USA: Society for Industrial and Applied Mathematics Philadelphia, 2001: 642-651.

[42] Xu G and Xu J. An LP rounding algorithm for approximating uncapacitated facility location problem with penalty[J]. Information Processing Letters, 2005, 94: 119-123.

[43] Geunes J, Levi R, Romeijn H E, Shmoys D B. Approximation algorithms for supply chain planning and logistics problems with market choice[J]. Mathematical Programming, 2011, 130: 85-106.

[44] Xu G, Xu J. An improved approximation algorithm for uncapacitated facility location problem with penalty[J]. Journal of Combinatorial Optimization, 2008, 17: 424-436.

[45] Li Y, Du D L, Xiu N H, and Xu D C, Improved Approximation Algorithms for the Facility Location Problems with Linear/submodular Penalty[C]// Proccedings of the 19th International Computing and Combinatorics Conference, Springer-

Verlag Berlin Heidelberg: Springer Berlin Heidelberg, 2013: 292-303.

[46] Hayrapetyan A, Swamy C, Tardös E. Network design for information networks[C]// Proceedings of the 16th annual ACM-SIAM symposium on Discrete algorithms, PA, USA: Society for Industrial and Applied Mathematics Philadelphia, 2005. 933-942.

[47] Du D L, Lu R X, Xu D C. A primal-dual approximation algorithm for the facility location problem with submodular penalty[J]. Algorithmica, 2012, 63: 191-200.

[48] Li Y, Du D L, Xiu N H, Xu D. A combinatorial 2.375-approximation algorithm for the facility location problem with submodular penalties[J]. Theoretical Computer Science, 2013, 476: 109-117.

[49] Xu D C, Gao D X, Wu C C. A primal-dual 3-approximation algorithm for the stochastic facility location problem with submodular penalties, 2013. DOI: 10.1080/02331934.2013.793326.

[50] Cornuejols, G.,Marshall L. Fisher, Nemhauser, G.L. Exceptional paper—location of bank accounts to optimize float: an analytic study of exact and approximate algorithms[J]. Management Science, 1977, 23: 789-810.

[51] Aardal A, Sviridenko M. An 0.828-approximation algorithm for the uncapacitated facility location problem[J]. Discrete Applied Mathematics, 1999, 93: 289-296.

[52] A. Bumb. An approximation algorithm for the maximization version of the two level uncapacitated facility location problem[J]. Operations Research Letters, 2001, 29: 155-161.

[53] Zhang J W, Ye Y Y. A note on the maximization version of the multi-level facility location problem. Operations Research Letters, 2002, 30: 333-335.

[54] Fortuin C M, Kasteleyn P W, Ginibre J. Correlation inteualities on some partially ordered sets. Communications in Mathematical Physics, 1971, 22: 89-103.

[55] Cornuejols G, Nemhauser G L, Wosely L A. The uncapacitated facility location problem[M]//Mirchandani P, Francis R. Discrete Location Theory. New York: Wiley, 1990: 119-171.

[56] Sviridenko M. A note on maximizing a submodular set function subject to a knapsack constraint[J]. Operations Research Letters, 2004, 32: 41-43.

此页不缺内容

致谢

首先, 我要感谢我的导师杨庆之教授. 在我博士三年期间, 杨教授无论从学习还是生活上都给了我极大的理解和支持. 杨教授严谨的学风对我三年的学习有着深远的影响, 也继续影响着我未来的生活和工作.

我还要特别感谢我的硕士导师和博士期间的合作导师北京工业大学徐大川教授. 徐大川教授淡泊名利, 对学术有执着的追求. 徐老师带领我进入到组合优化的领域直到现在仍然继续指导我的研究工作. 本论文就是在徐大川教授的指导下完成的, 在此我对于徐老师无私的教导表示衷心感谢.

此外, 我还要感谢加拿大新不伦瑞克的杜东雷教授. 基于前期的研究合作, 杜教授邀请我到加拿大访问一年. 在我访问期间, 杜教授在生活和学习上都给予我极大的帮助和关心.

同时, 我还要感谢北京工业大学赵欣苑老师, 李改弟老师对我的帮助和支持, 感谢我在南开大学的以及北京工业大学的同门.

最后，谨以此文献给我挚爱的父母，他们在背后的默默支持是我前进的动力. 在此，祝愿他们身体健康，心情愉快！

此页不缺内容

个人简历

出生年月: 1986年7月11日

教育经历:

时间	学校	所获学位
2004年9月-2008年7月	安庆师范学院	学士
2008年9月-2011年7月	北京工业大学	硕士
2011年9月-2014年7月	南开大学	博士

- **发表文章期刊列表**

1. **Chenchen Wu**, Donglei Du, and Dachuan Xu, A per-scenario bound for the two-stage stochastic facility location problem with linear penalty, Optimization, DOI: 10.1080 /02331934.2013.840626. (SCI刊源)
2. **Chenchen Wu**, Dachuan Xu, Improved approximation algorithms for the k-level facility location problem with soft capacities, accepted by Acta Mathematicae Applicatae Sinica. (SCI刊源)
3. **Chenchen Wu**, Donglei Du, Dachuan Xu, An improved semidefinite programming hierarchies rounding approximation algorithm for maximum graph bisection problems, Journal of Combinatorial Optimization, DOI: 10.1007/s10878-013 -9673-1. (SCI刊源)
4. **Chenchen Wu**, Donglei Du, and Dachuan Xu, An improved per-scenario bound for the two-stage stochastic facility location problem, submitted to SCIENCE CHINA Mathematics. (SCI刊源)
5. **Chenchen Wu**, Dachuan Xu, Donglei Du, Wen-qing Xu, A complex semidefinite programming rounding approximation algorithm for the balanced Max-3-Uncut problem, accepted by Proceedings of COCOON, 2014.
6. Fengmin Wang, Dachuan Xu, **Chenchen Wu**, Approximation algorithms for the priority facility location problem with linear penalties, accepted by Journal of Systems Science and Complexity. (SCI刊源)
7. Dachuan Xu, Dongxiao Gao,**Chenchen Wu**, A 3-approximation algorithm

for the stochastic facility location problem with submodular penalties, Optimization, DOI: 10.1080/02331934.2013.793326 (SCI刊源)

8. 徐大川, 杜东雷, **吴晨晨**, 设施选址问题的近似算法综述. 数学进展,已录用.(国内核心)
9. 徐大川, 万玮, **吴晨晨**, 徐文青, 随机容错设施选址问题的原始对偶近似算法, 运筹学学报, 已录用. (国内核心)

投稿文章:

1. **Chenchen Wu**, Donglei Du, and Dachuan Xu, Primal-dual approximation algorithms for the two-level facility location problem, submitted to Theoretical Computer Science. (SCI刊源)
2. **Chenchen Wu**, Dachuan Xu, and Jiawei Zhang, Semidefinite programming safe approximation for distributionally robust joint chance constrained program, submitted to Asia-Pacific Journal of Operational Research. (SCI刊源)
3. 徐大川, **吴晨晨**, 杜东雷. 二次整数规划的半定规划舍入近似算法. 已投往运筹与管理.(国内核心)