



k -中位问题的近似算法

献给韩继业教授 90 寿辰

吴晨晨¹, 杜东雷², 苗润杰³, 徐大川^{4*}

1. 天津理工大学理学院运筹学与系统工程研究所, 天津 300384;

2. Faculty of Management, University of New Brunswick, Fredericton E3B 5A3, Canada;

3. 郑州大学数学与统计学院, 郑州 450001;

4. 北京工业大学数学统计学与力学学院, 北京 100124

E-mail: wu.chenchen_tjut@163.com, ddu@unb.ca, runjiemiao@hotmail.com, xudc@bjut.edu.cn

收稿日期: 2024-05-26; 接受日期: 2024-09-22; 网络出版日期: 2024-12-30; * 通信作者

国家自然科学基金 (批准号: 12371320) 资助项目

摘要 k -中位问题是理论计算机科学和组合优化领域中的经典问题之一, 其应用场景遍及图像处理、模式识别、供应链管理和数据挖掘等前沿领域. 如今, 随着大数据和机器学习的发展, k -中位问题的实际应用愈发复杂多样, 这也产生了诸多亟待解决的研究课题. 为使相关研究人员快速了解 k -中位问题, 本文综述经典 k -中位问题及其变型问题的基于线性规划舍入、原始-对偶、对偶拟合、局部搜索和 Lagrange 松弛等技巧设计的有效算法. 首先介绍经典 k -中位问题与无容量设施选址问题之间的密切联系; 然后介绍 k -中位问题的若干重要变形, 包括 k -中心问题、 k -设施选址问题、背包中位问题、容错 k -中位问题、带容量的 k -中位问题、带下界的 k -中位问题、带惩罚的 k -中位问题、带异常值的 k -中位问题和在线中位问题等问题; 最后介绍 k -中位问题的近似算法研究成果, 并列出了该领域中的若干公开问题.

关键词 k -中位问题 近似算法 线性规划 局部搜索

MSC (2020) 主题分类 65K05, 65K10, 68W25

1 引言

k -中位问题 (k -median problem) 是组合优化领域中的经典问题, 它的目标是从给定的数据集中选出 k 个中心点, 以最小化所有点到其最近中心点的距离总和. k -中位问题与设施选址问题 (facility location problem) 有着密切联系: 在 k -中位问题中, 用中心点的开设费用取代中心点的开设数目限制, 并将优化目标设为开设费用与距离总和的最小化, 便产生了无容量设施选址问题 (uncapacitated facility location problem). k -中位问题 and 无容量设施选址问题都是经典的 NP (non-deterministic polynomial)-

英文引用格式: Wu C C, Du D L, Miao R J, et al. Survey on approximation algorithms for the k -median problem (in Chinese). Sci Sin Math, 2025, 55: 551–566, doi: 10.1360/SSM-2024-0171

困难问题, 对这类问题设计近似算法是一种重要且有效的解决方法. 本文主要综述 k -中位问题在近似算法领域的研究成果.

对于最小化问题 P , 算法 $\mathcal{A}$ 被称为问题 P 的 ρ -近似算法, 是指对问题 P 的所有实例, 算法 $\mathcal{A}$ 都能在多项式时间内输出一个近似解, 其费用不超过最优解费用的 ρ 倍, 即

$$\text{cost}(\mathcal{A}_I) \leq \rho \cdot \text{cost}(\text{OPT}_I), \quad \text{任取问题 } P \text{ 的实例 } I.$$

由于 k -中位问题和无容量设施选址问题联系紧密, 其近似算法或多或少受到无容量设施选址问题的近似算法的启发. 因此, 先介绍针对无容量设施选址问题的近似算法设计技巧. 粗略地讲, 这些技巧可以分为 4 类, 分别是线性规划舍入、对偶拟合、原始 - 对偶和局部搜索.

线性规划舍入 技巧通常包含 3 个步骤: (1) 将问题建模成整数线性规划; (2) 通过松弛整数变量得到线性规划松弛并求最优解; (3) 将求得的最优解舍入成整数线性规划的可行解. 无容量设施选址问题的第一个常数近似算法便是由 Shmoys 等^[40] 利用线性规划舍入技巧得到的, 其近似比是 3.16. 之后, Chudak 和 Shmoys^[11]、Sviridenko^[41] 以及 Byrka 和 Aardal^[2] 基于线性规划舍入技巧不断改进近似比.

线性规划舍入技巧的优点是, 在设计算法时所求得的最优解可以作为原问题的自然下界, 因此往往能得到较好的近似比. 然而, 由于求解线性规划松弛得到的是分数最优解, 因此在设计算法时难以利用问题的组合结构, 基于线性规划舍入技巧设计的算法也不是组合算法.

对偶拟合 技巧的本质是集合覆盖问题的贪婪算法. 它可以看作将贪婪算法中新集合的选择过程“替换”成线性规划松弛的对偶问题中对偶变量上升过程, 之后根据对偶变量产生一个整数可行解. 通常, 这样得到的对偶解并不可行, 因此需要通过适当的“拟合”进行修正, 而算法的近似比便可通过建立因子揭示线性规划 (factor-revealing linear programming) 来得到. 2003 年, Jain 等^[28] 利用对偶拟合的技巧得到了无容量设施选址问题的 1.61- 近似算法. 之后, Mahdian 等^[38] 通过放缩设施的费用, 将近似比改进到 1.52.

无容量设施选址问题的对偶拟合算法的一个重要的“副产品”是蕴涵其中的双因子近似. 所谓的 (γ_F, γ_C) - 双因子近似是指, 算法 $\mathcal{A}$ 输出的解的费用不超过任意可行解设施费用的 γ_F 倍与连接费用的 γ_C 倍之和, 即

$$\text{cost}(\mathcal{A}) \leq \gamma_F \cdot \text{cost}_F(S) + \gamma_C \cdot \text{cost}_C(S), \quad \text{任取可行解 } S,$$

其中 $\text{cost}_F(S)$ 和 $\text{cost}_C(S)$ 分别表示可行解 S 的设施费用和连接费用. 在设计近似算法时, 可以将双因子近似结果与其他技巧结合得到更好的近似比. 然而, 总体来看, 对偶拟合技巧的“针对性”较强, 而泛化能力较弱.

原始 - 对偶 技巧是通过对线性规划松弛的对偶问题制定对偶上升过程, 得到的对偶可行解. 之后, 再将对偶可行解对应成原问题的整数可行解. 2001 年, Jain 和 Vazirani^[30] 基于原始 - 对偶技巧得到了无容量设施选址问题的 3- 近似算法.

原始 - 对偶技巧的优点是, 它的可移植性强, 易于应用或泛化到不同的拓展问题上. 美中不足的是, 通过原始 - 对偶技巧所得到的近似比一般比较大.

局部搜索 技巧是从任一可行解出发, 在所定义的局部交换规则下, 判断当前可行解能否被改进. 如果能被改进, 则不断更新当前可行解直到找到局部最优解. 不同的局部搜索算法主要通过所定义的局部交换规则进行区分. 2000 年, Korupolu 等^[31] 通过定义增加一个设施、删除一个设施和交换一对设施这 3 种局部交换规则, 得到了无容量设施选址问题的 $(5 + \epsilon)$ - 近似算法, 这里的 ϵ 是任意给定的

表 1 k -中位问题已有的近似算法和近似比下界

技巧	作者	近似比
线性规划舍入	Charikar 等 [7]	$6\frac{2}{3}$
线性规划舍入	Charikar 和 Li [9]	3.25
局部搜索	Arya 等 [1]	$5 + \epsilon$
局部搜索	Arya 等 [1]	$3 + \epsilon$
局部搜索	Cohen-Addad 等 [13]	$2.836 + \epsilon$
原始 - 对偶、两阶段舍入	Jain 和 Vazirani [30]	6
对偶拟合、两阶段舍入	Jain 等 [28]	4
对偶拟合、两阶段舍入	Li 和 Svensson [37]	$1 + \sqrt{3} + \epsilon$
对偶拟合、两阶段舍入	Byrka 等 [4]	$2.675 + \epsilon$
对偶拟合、两阶段舍入	Cohen-Addad 等 [12]	$2.67059 + \epsilon$
对偶拟合、两阶段舍入	Gowda 等 [18]	2.613
近似比下界	Jain 等 [29]	$1 + \frac{2}{e}$

正实数. 在局部搜索算法中, 尾项 ϵ 通常和算法的运行时间有关, 可以看作为保证算法能够在多项式时间内执行完成而损失的近似比. 之后, 在 2004 年, Arya 等 [1] 基于同样的局部交换规则, 分析得到了无容量设施选址问题的 $(3 + \epsilon)$ -近似. 他们也指出可以借助 Charikar 和 Guha [6] 的技巧分析得到 $(1 + \sqrt{2} + \epsilon)$ -近似.

局部搜索技巧的优点是, 算法形式简单, 执行便捷, 便于数值实现且具有强大的泛化能力. 然而, 局部搜索算法分析过程的技巧性较强, 所得到的近似比一般而言也比较大.

综合而言, 以上 4 类技巧各有特色. 为达到更好的性能, 我们时常需要将多种技巧混合使用. 例如, 当前针对无容量设施选址问题的近似比最小的近似算法便是 Li [34] 结合线性规划舍入技巧和对偶拟合技巧得到的, 其近似比是 1.488. 关于无容量设施选址问题的近似比下界, Guha 和 Khuller [19] 指出, 除非 $\text{NP} \subseteq \text{DTIME}(n^{O(\log \log n)})$, 该下界为 1.463. 之后, 在与 Chudak 和 Shmoys [11] 的私人通信中, Sviridenko 将上述前提假设加强到了 $\text{P} \neq \text{NP}$. 关于设施选址问题的更多结果, 可参见综述文献 [48] 和专著 [49].

至此, 基本介绍了设计无容量设施选址问题近似算法的常用技巧. 为使相关研究人员快速了解 k -中位问题, 本文转而综述 k -中位问题的相关结果. 如前所述, k -中位问题的近似算法也用到了上面这 4 类技巧, 相关结果见表 1.

本文较系统地梳理 k -中位问题的近似算法研究结果, 其结构如下. 第 2 节介绍 k -中位问题的重要变型. 第 3 节介绍经典 k -中位问题的研究进展. 第 4 节总结待研究的问题.

2 k -中位问题的重要变型

2.1 再论 k -中位问题和无容量设施选址问题

本小节通过对比 k -中位问题和无容量设施选址问题的线性规划松弛、Lagrange 松弛及相应的对偶问题, 进一步揭示 k -中位问题与无容量设施选址问题之间的密切联系.

给定需求点的集合 C 、待选设施点的集合 F 、输入参数 k 以及每对待选设施点 $i \in F$ 与需求点 $j \in C$ 之间的连接费用 c_{ij} ¹⁾, k -中位问题的目标是从待选设施点的集合 F 中开设不超过 k 个设施以服务所有的需求点, 使得总的连接费用最小. 因此, k -中位问题的目标函数可以写作

$$\min_{S \subseteq F: |S| \leq k} \sum_{j \in C} \min_{i \in S} c_{ij}.$$

如果对于是否开设待选设施点 i 引入 0-1 变量 y_i , 是否将需求点 j 连接到设施点 i 引入 0-1 变量 x_{ij} , 可以得到 k -中位问题的整数线性规划

$$\begin{aligned} \min \quad & \sum_{i \in F, j \in C} c_{ij} x_{ij} \\ \text{s.t.} \quad & x_{ij} \leq y_i, \quad \forall i \in F, \quad j \in C, \\ & \sum_{i \in F} x_{ij} = 1, \quad \forall j \in C, \\ & \sum_{i \in F} y_i \leq k, \\ & x_{ij}, y_i \in \{0, 1\}, \quad \forall i \in F, \quad j \in C, \end{aligned} \quad (2.1)$$

其中, 第一组约束表示, 如果需求点 j 连接到设施点 i 上, 则设施点 i 必须开设; 第二组约束表示每个需求点 j 都必须连接到一个开设的设施点上; 第三组约束表示开设的设施点的总数不超过 k . 值得注意的是, 由于变量 x_{ij} 和 y_i 均为 0-1 变量, 因此整数线性规划 (2.1) 的第二组约束可以等价地修改为

$$\sum_{i \in F} x_{ij} \geq 1, \quad \forall j \in C.$$

如果除去整数约束, 则可以得到 k -中位问题的线性规划松弛

$$\begin{aligned} \min \quad & \sum_{i \in F, j \in C} c_{ij} x_{ij} \\ \text{s.t.} \quad & x_{ij} \leq y_i, \quad \forall i \in F, \quad j \in C, \\ & \sum_{i \in F} x_{ij} \geq 1, \quad \forall j \in C, \\ & \sum_{i \in F} y_i \leq k, \\ & x_{ij}, y_i \geq 0, \quad \forall i \in F, \quad j \in C. \end{aligned} \quad (2.2)$$

如果允许通过在目标函数中支付惩罚来开设多于 k 个中心点, 则将得到 k -中位问题的 Lagrange 松弛

$$\begin{aligned} \min \quad & \sum_{i \in F, j \in C} c_{ij} x_{ij} + \sum_{i \in F} \gamma y_i - \gamma k \\ \text{s.t.} \quad & x_{ij} \leq y_i, \quad \forall i \in F, \quad j \in C, \\ & \sum_{i \in F} x_{ij} \geq 1, \quad \forall j \in C, \\ & x_{ij}, y_i \geq 0, \quad \forall i \in F, \quad j \in C. \end{aligned} \quad (2.3)$$

1) 本文只讨论 c_{ij} 为度量的情形, 即 c_{ij} 满足非负性、对称性和三角不等式.

可以观察到, k -中位问题的线性规划松弛 (2.2) 的可行解也是 Lagrange 松弛 (2.3) 的可行解. 并且, 当 $\gamma \geq 0$ 时, Lagrange 松弛 (2.3) 的最优值可以作为 k -中位问题的线性规划松弛 (2.2) 的最优值的下界. 对于此 Lagrange 松弛, 我们也可以写出它的对偶形式

$$\begin{aligned} \max \quad & \sum_{j \in C} \alpha_j - \gamma k \\ \text{s.t.} \quad & \alpha_j \leq \beta_{ij} + c_{ij}, \quad \forall i \in F, \quad j \in C, \\ & \sum_{j \in C} \beta_{ij} \leq \gamma, \quad \forall i \in F, \\ & \alpha_j, \beta_{ij} \geq 0, \quad \forall i \in F, \quad j \in C. \end{aligned} \quad (2.4)$$

回顾无容量设施选址问题²⁾. 给定需求点的集合 C 、待选设施点的集合 F 、每个待选设施点 i 的开设设施费用 f_i 以及每对待选设施点 $i \in F$ 和需求点 $j \in C$ 之间的连接费用 c_{ij} , 无容量设施选址问题的目标是开设一些设施以服务需求点, 使得开设费用和连接费用的总和最小, 即其目标函数为

$$\min_{S \subseteq F} \left\{ \sum_{i \in S} f_i + \sum_{j \in C} \min_{i \in S} c_{ij} \right\}.$$

如果对于是否开设待选设施点 i 引入 0-1 变量 y_i , 是否将需求点 j 连接到设施点 i 上引入 0-1 变量 x_{ij} , 可以得到无容量设施选址问题的整数线性规划

$$\begin{aligned} \min \quad & \sum_{i \in F} f_i y_i + \sum_{i \in F, j \in C} c_{ij} x_{ij} \\ \text{s.t.} \quad & x_{ij} \leq y_i, \quad \forall i \in F, \quad j \in C, \\ & \sum_{i \in F} x_{ij} \geq 1, \quad \forall j \in C, \\ & x_{ij}, y_i \in \{0, 1\}, \quad \forall i \in F, \quad j \in C. \end{aligned} \quad (2.5)$$

它与 k -中位问题的整数线性规划 (2.1) 的区别在于, 目标函数为设施费用与连接费用的总和, 以及没有开设设施的个数限制. 如果除去整数约束, 则可得无容量设施选址问题的线性规划松弛

$$\begin{aligned} \min \quad & \sum_{i \in F} f_i y_i + \sum_{i \in F, j \in C} c_{ij} x_{ij} \\ \text{s.t.} \quad & x_{ij} \leq y_i, \quad \forall i \in F, \quad j \in C, \\ & \sum_{i \in F} x_{ij} \geq 1, \quad \forall j \in C, \\ & x_{ij}, y_i \geq 0, \quad \forall i \in F, \quad j \in C. \end{aligned} \quad (2.6)$$

而其对称形式为

$$\begin{aligned} \max \quad & \sum_{j \in C} \alpha_j \\ \text{s.t.} \quad & \alpha_j \leq \beta_{ij} + c_{ij}, \quad \forall i \in F, \quad j \in C, \\ & \sum_{j \in C} \beta_{ij} \leq f_i, \quad \forall i \in F, \\ & \alpha_j, \beta_{ij} \geq 0, \quad \forall i \in F, \quad j \in C. \end{aligned} \quad (2.7)$$

2) 不至混淆的情形下, 使用与 k -中位问题相同的符号.

此时, 通过对比 k -中位问题的 Lagrange 松弛 (2.3) 和无容量设施选址问题的线性规划松弛 (2.6), 以及 k -中位问题的 Lagrange 松弛的对偶 (2.4) 和无容量设施选址问题的线性规划松弛的对偶 (2.7), 可以发现除掉目标函数中的 $-\gamma k$ 项, k -中位问题的 Lagrange 松弛与无容量设施选址问题的线性规划松弛几乎一样. 事实上, k -中位问题的 Lagrange 松弛 (2.3) 可以看作设施费用均为 γ 的无容量设施选址问题.

进一步地, 若无容量设施选址问题的近似算法 $\mathcal{A}$ 输出的解 S 满足

$$\text{cost}_C(S) + c \cdot \sum_{i \in S} f_i \leq c \cdot \sum_{j \in C} \alpha_j, \quad (2.8)$$

假设设施开设费用均为 γ , 则可得到

$$\text{cost}_C(S) \leq c \cdot \left(\sum_{j \in C} \alpha_j - \gamma |S| \right).$$

利用 k -中位问题的 Lagrange 松弛的对偶、 k -中位问题的 Lagrange 松弛和 k -中位问题三者最优值之间的大小关系, 可以发现 S 有望作为中位问题的近似解, 通过调整 γ 的取值促使 $|S| = k$ 便得到 k -中位问题的近似解.

因此, 为了设计 k -中位问题的近似算法, 我们也可通过对无容量设施选址问题设计满足 (2.8) 的近似算法, 这样的近似算法通常被称为 Lagrange 乘子保持 (Lagrangian multiplier preserving, LMP)-算法. 已知的 LMP-算法包含 Jain 和 Vazirani^[30] 的 3-近似算法、Jain 等^[28] 的 2-近似算法和 Cohen-Addad 等^[12] 的 1.9973-近似算法. 借助这些 LMP-算法, 人们也得到了 k -中位问题一系列近似结果, 详见表 2 和第 3.2 小节.

2.2 k -中位问题的拓展问题

本小节介绍 k -中位问题的几类拓展问题, 如 k -中心问题、 k -设施选址问题、背包中位问题、带容量的 k -中位问题、带下界的 k -中位问题、带惩罚的 k -中位问题、带异常值的 k -中位问题和在线中位问题等, 并介绍相应的近似算法研究结果.

k -中心问题 (k -center problem) 可以看作 k -中位问题的姊妹问题. 在 k -中心问题中, 我们的输入是需求点集合 C 、待选设施点集合 F 、正整数 k 以及每对待选设施点 $i \in F$ 和需求点 $j \in C$ 之间的连接费用 c_{ij} . k -中心问题的目标是开设不超过 k 个设施使得所有需求点的最大连接费用最小. 因此, k -中心问题可以写作

$$\min_{S \subseteq F: |S| \leq k} \max_{j \in C} \min_{i \in S} c_{ij}.$$

表 2 k -中位问题的两阶段近似算法

作者	第一阶段损失近似比	第二阶段损失近似比	近似比
Jain 和 Vazirani ^[30]	3	2	6
Jain 等 ^[28]	2	2	4
Li 和 Svensson ^[37]	2	$\frac{1+\sqrt{3}+\epsilon}{2}$	$1 + \sqrt{3} + \epsilon$
Byrka 等 ^[4]	2	$1.3371 + \epsilon$	$2.675 + \epsilon$
Cohen-Addad 等 ^[12]	1.9973	$1.3371 + \epsilon$	$2.67059 + \epsilon$
Gowda 等 ^[18]	2	1.3064	2.613

与经典的 k -中位问题对比可以发现两者具有相同的约束条件而目标函数不同. 通俗地讲, k -中位问题关注“整体”费用而 k -中心问题关注“单个”费用, k -中位问题追求总距离最小而 k -中心问题追求半径最小. 关于 k -中心问题的近似算法研究成果见表 3.

k -设施选址问题 (k -facility location problem) 是无容量设施选址问题和 k -中位问题相结合的产物. 在 k -设施选址问题中, 我们的输入包含需求点集合 C 、待选设施点集合 F 、正整数 k 、每个待选设施点 $i \in F$ 的开设费用 f_i 以及每对待选设施点 $i \in F$ 和需求点 $j \in C$ 之间的连接费用 c_{ij} . k -设施选址问题的目标是开设不超过 k 个设施使得开设费用和连接费用的总和最小. 因此, k -设施选址问题可以写作

$$\min_{S \subseteq F: |S| \leq k} \left\{ \sum_{i \in S} f_i + \sum_{j \in C} \min_{i \in S} c_{ij} \right\}.$$

从 k -设施选址问题的模型可以看到, 它继承了无容量设施选址问题的目标函数和 k -中位问题的约束条件. 不难发现, 当设施点开设费用 f_i 均为 0 时, k -设施选址问题退化为 k -中位问题; 当 $k \geq |F|$ 时, k -设施选址问题退化为无容量设施选址问题. 关于 k -设施选址问题的近似算法研究成果见表 4.

背包中位问题 (knapsack median problem) 是 k -中位问题的另一个衍生问题, 在背包中位问题中, 我们的输入包含需求点的集合 C 、待选设施点的集合 F 、预算值 (背包容量) B 、每个待选设施点开设费用为 f_i 以及每个待选设施点 $i \in F$ 和每个需求点 $j \in C$ 之间的连接费用 c_{ij} . 背包中位问题的目标是在保证设施开设费用不超过预算值 B 的前提下开设一些设施使得总连接费用最小, 即背包中位问题可以描述为

$$\min_{S \subseteq F: \sum_{i \in S} f_i \leq B} \sum_{j \in C} \min_{i \in S} c_{ij}.$$

粗略地看, 背包中位问题是 k -中位问题在约束条件 k 上的松弛, 它将每个设施点的“重要性”通过设施费用进行区分因而具有更广阔的应用场景. 不难发现, 当所有设施点的开设费用均为 1 且预算值 B 取为 k 时, 背包中位问题退化为 k -中位问题. 关于背包中位问题的近似算法研究成果见表 5.

在 k -中位问题的部分应用场景中, 需求点对于连接的“稳定性”要求不尽相同, 这样的场景可以用容错 k -中位问题 (fault-tolerant k -median problem) 更好地建模. 具体地, 在容错 k -中位问题中, 我们的输入包含需求点的集合 C 、待选设施点的集合 F 、正整数 k 、每对待选设施点 $i \in F$ 和需求点 $j \in C$ 之间的连接费用 c_{ij} , 以及每个需求点 $j \in C$ 对连接的“稳定性”要求 (容错要求) r_j . 其中, 需

表 3 k -中心问题的近似算法

作者	近似比
Garey 和 Johnson ^[16]	NP- 困难性
Gonzalez ^[17]	2
Hochbaum 和 Shmoys ^[26]	2
Hochbaum 和 Shmoys ^[27]	2

表 4 k -设施选址问题的近似算法

技巧	作者	近似比
原始 - 对偶	Jain 和 Vazirani ^[30]	6
对偶拟合	Jain 等 ^[28]	4
局部搜索	Zhang ^[50]	$2 + \sqrt{3} + \epsilon$

求点 j 的容错要求 r_j 是要将点 j 连接到 r_j 个不同的开设设施上. 容错 k -中位问题的目标是开设不超过 k 个设施并在满足每个需求点容错要求的前提下将顾客连接到开设的设施点上, 使得总连接费用最小, 即

$$\min_{S \subseteq F: |S| \leq k} \sum_{j \in C} \sum_{t=1}^{r_j} c_t(j, S),$$

其中 $c_t(j, S)$ 是需求点 j 与开设设施集合 S 中连接费用第 t 小的设施点之间的连接费用. 可以发现, 当所有的需求点的容错要求均为 1 时, 容错 k -中位问题退化为了 k -中位问题. 关于容错 k -中位问题的近似算法研究成果见表 6.

由于资源的有限性, 在许多应用场景中, 设施点的服务能力是有限的, 这便引出了带容量的 k -中位问题 (capacitated k -median problem). 在带容量的 k -中位问题中, 我们的输入包含需求点的集合 C 、待选设施点的集合 F 、正整数 k 、每个待选设施点 $i \in F$ 和每个需求点 $j \in C$ 之间的连接费用 c_{ij} , 以及每个设施点 i 服务能力的上界, 即容量 u_i . 带容量的 k -中位问题的目标是开设不超过 k 个设施并在保证不超过每个开设的设施点的容量限制的前提下将顾客连接到开设的设施点上, 使得总连接费用最小.

与前几种变型问题不同, 在带容量的 k -中位问题中, 我们不仅需要考虑开设设施个数的全局约束, 还需要考虑每个设施点的服务量的局部约束. 因此, 我们实际上需要在 F 中选择子集 $S = \{i_1, i_2, \dots, i_k\}$, 并以 S 中的设施点作为中心将需求点集合进行划分, 即找到 $C_1, C_2, \dots, C_k$ 满足 $\bigcup_{i=1}^k C_i = C$ 和对任意互不相同的 $i, j = 1, 2, \dots, k$ 均有 $C_i \cap C_j = \emptyset$ 且 $|C_i| \leq u_{i_i}$. 则带容量的 k -中位问题可以描述为

$$\min_{S = \{i_1, i_2, \dots, i_k\} \subseteq F: |S| \leq k} \sum_{l=1}^k \sum_{j \in C_l} c_{i_l j}.$$

可以看到, 当每个设施点的容量均为无穷时, 带容量的 k -中位问题便退化到了 k -中位问题. 关于带容量的 k -中位问题的近似算法研究成果见表 7.

带下界约束的 k -中位问题 (lower-bounded k -median problem) 也是 k -中位问题的变型问题. 在带下界约束的 k -中位问题中, 我们的输入包含需求点的集合 C 、待选设施点的集合 F 、正整数 k 、每个待选设施点 $i \in F$ 和每个需求点 $j \in C$ 之间的连接费用 c_{ij} , 以及每个设施点的服务量的下界要求

表 5 背包中位问题的近似算法

技巧	作者	近似比	特殊要求
线性规划舍入	Krishnaswamy 等 [32]	16	预算值增加至多 $f_{\max}$
线性规划舍入	Kumar [33]	2,700	—
线性规划舍入	Charikar 和 Li [9]	34	—
线性规划舍入	Swamy [42]	32	—
线性规划舍入	Byrka 等 [3]	17.46	—

 表 6 容错 k -中位问题的近似算法

技巧	作者	近似比	特殊要求
对偶拟合	Swamy 和 Shmoys [43]	4	容错要求一致
线性规划舍入	Hajiaghayi 等 [22]	93	—

l_i . 带下界约束的 k -中位问题的目标是开设不超过 k 个设施并在保证不低于每个开设的设施点的下界限的前提下将顾客连接到开设的设施点上, 使得总连接费用最小. 与带容量的 k -中位问题类似, 我们需要找到满足设施点下界约束的需求点的划分 $C_1, C_2, \dots, C_k$, 即 $|C_i| \geq l_i$. 因此, 带下界约束的 k -中位问题可以描述为

$$\min_{S=\{i_1, i_2, \dots, i_k\} \subseteq F: |S| \leq k} \sum_{l=1}^k \sum_{j \in C_l} c_{lj}.$$

带下界约束的 k -中位问题可以看作 k -中位问题在需求分配“公平性”方向的变型. 不难发现, 当每个设施点的下界均为 0 时, 带下界约束的 k -中位问题退化为 k -中位问题. 关于带下界约束的 k -中位问题的近似算法研究成果见表 8.

与无容量设施选址问题类似, k -中位问题也有带惩罚及带异常值版本的变型问题. 在带惩罚的 k -中位问题 (k -median problem with penalties) 中, 我们的输入包含需求点的集合 C 、待选设施点的集合 F 、正整数 k 、每个待选设施点 $i \in F$ 和每个需求点 $j \in C$ 之间的连接费用 c_{ij} , 以及惩罚函数 $h: 2^C \rightarrow \mathbb{R}^+$. “惩罚”缘于考虑某些远距离的需求点对解结构的影响, 因此设定某些需求点可以通过支付一定的惩罚费用而不被服务, 这样可能更容易发现问题的本质. 带惩罚的 k -中位问题的目标是开设不超过 k 个设施并选择一些需求点不被服务使得连接费用和惩罚费用总和最小, 即

$$\min_{S \subseteq F: |S| \leq k, P \subseteq C} \left\{ \sum_{j \in C \setminus P} \min_{i \in S} c_{ij} + h(P) \right\}.$$

当惩罚函数 $h(\cdot)$ 为模函数则被称为带线性惩罚的 k -中位问题, 当惩罚函数 $h(\cdot)$ 为次模函数则被称为带次模惩罚的 k -中位问题. 可以看到, 当所有需求点的惩罚费用函数值均为无穷时, 带惩罚的 k -中位问题退化为 k -中位问题. 关于带惩罚的 k -中位问题的近似算法研究成果见表 9.

带异常值的 k -中位问题 (k -median problem with outliers) 是带惩罚的 k -中位问题的孪生兄弟. 在带异常值的 k -中位问题中, 我们的输入包含需求点的集合 C 、待选设施点的集合 F 、正整数 k 、每个待选设施点 $i \in F$ 和每个需求点 $j \in C$ 之间的连接费用 c_{ij} , 以及异常值 z . 异常值旨在捕获并析出对解的结构影响过大的需求点以更好捕获解的结构特征. 带异常值的 k -中位问题的目标是开设不超

表 7 带容量的 k -中位问题的近似算法

作者	近似比	设施数目违背	设施容量违背	其他特殊要求
Li [36]	$e^{O(1/\epsilon^2)}$	$1 + \epsilon$	—	设施容量一致
Li [35]	$O(1/\epsilon^2) \log(1/\epsilon)$	$1 + \epsilon$	—	—
Byrka 等 [5]	$O(1/\epsilon^2)$	—	$1 + \epsilon$	—
Demirci 和 Li [14]	$O(1/\epsilon^5)$	—	$1 + \epsilon$	—

表 8 带下界约束的 k -中位问题的近似算法

作者	近似比	下界违背	其他特殊要求
Han 等 [24]	$(2 + \sqrt{3} + \epsilon) \frac{1+\alpha}{1-\alpha}$	$\alpha \in [0, 1)$ 倍	—
Han 等 [25]	610	—	下界一致
Han 等 [25]	387	—	下界一致
Guo 等 [20] 和 Wu 等 [47]	$516 + \epsilon$	—	下界一致

表 9 带惩罚的 k -中位问题的近似算法

作者	近似比	特殊要求
Charikar 等 [8]	4	—
Hajiaghayi 等 [23]	$3 + \epsilon$	—
Wu 等 [45]	$1 + \sqrt{3} + \epsilon$	惩罚费用一致
Wu 等 [46]	$3 + \epsilon$	—

 表 10 带异常值的 k -中位问题的近似算法

作者	近似比	设施数目违背	异常值数目违背
Chen [10]	常数	—	—
Friggstad 等 [15]	$3 + \epsilon$	$1 + \epsilon$	—
Gupta 等 [21]	$17 + \epsilon$	—	$O(k \log(n\delta) - \epsilon)$
Wu 等 [46]	$5 + \epsilon$	—	$O(k \log(n\delta) - \epsilon)$
Wu 等 [46]	$3 + \epsilon$	—	$O(k^2 \log(n\delta) - \epsilon)$

过 k 个设施并选择至多 z 个异常点使得总连接费用最小, 即

$$\min_{S \subseteq F: |S| \leq k, P \subseteq C: |P| \leq z} \sum_{j \in C \setminus P} \min_{i \in S} c_{ij}.$$

可以看到, 当异常值的个数为 0 时, 带异常值的 k -中位问题退化为 k -中位问题. 关于带异常值的 k -中位问题的近似算法研究成果见表 10.

k -中位问题的另一种衍生问题是在线中位问题 (online median problem). 在线中位问题中的特点是所能开设的设施数目是提前不知道的, 并将以“在线”形式给出的. 它的目标是确定开设设施的排序, 使得当开设设施数目 (例如为 l) 确定后, 按照所确定的排序开设前 l 个设施所产生的费用与最优解的费用之间的比值最小. 2003 年, Mettu 和 Plaxton [39] 给出了在线中位问题的 3-近似算法.

上面这几种变型问题也可以相互融合以适应更丰富的应用场景, 这也衍生出了非常丰富的研究问题. 相关研究也在如火如荼地进行, 篇幅所限, 不予细释.

3 k -中位问题算法

在文献中, k -中位问题的近似算法可分为直接算法和间接算法 (又称两阶段算法) 两类. 直接算法通常使用局部搜索技巧或线性规划舍入技巧直接设计近似算法, 如 Arya 等 [1] 的 $(5 + \epsilon)$ -近似算法和 $(3 + \epsilon)$ -近似算法、Cohen-Addad 等 [13] 的 $(2.836 + \epsilon)$ -近似算法、Charikar 等 [7] 的 $6\frac{2}{3}$ -近似算法以及 Charikar 和 Li [9] 的 3.25-近似算法. 两阶段算法的第一阶段主要借助无容量设施选址问题构造 k -中位问题的双点解, 之后在第二阶段对双点解进行舍入得到可行解. 因此, 两阶段算法最终的近似比依赖这两个阶段中近似比的“损失”, 如 Jain 和 Vazirani [30] 的 6-近似算法、Jain 等 [28] 的 4-近似算法、Li 和 Svensson [37] 的 $(1 + \sqrt{3} + \epsilon)$ -近似算法、Byrka 等 [4] 的 $(2.675 + \epsilon)$ -近似算法、Cohen-Addad 等 [12] 的 $(2.67059 + \epsilon)$ -近似算法以及 Gowda 等 [18] 的 2.613-近似算法. 这些算法在两个阶段损失的近似比及最终近似比见表 2.

接下来在第 3.1 小节介绍 k -中位问题的直接算法, 在第 3.2 小节介绍 k -中位问题的两阶段算法.

3.1 直接算法

本小节先介绍基于局部搜索技巧设计的近似算法, 再介绍基于线性规划舍入技巧设计的近似算法.

3.1.1 局部搜索算法

如第 1 节所述, 局部搜索算法的执行往往非常简单 (其一般的执行框架见图 1), 而其分析往往技巧性较强. 实际上, 在局部算法的执行中, 为了保证算法能在多项式时间内运行, 我们需要每次执行交换时的改进量大于某个门槛值, 这将损失部分近似比 (也就是近似比中 ϵ 的由来). 因此, 为了叙述简洁, 在本节进一步介绍局部搜索算法时, 我们便不再关注 ϵ 这一部分.

在 Arya 等^[1]设计的 $(5 + \epsilon)$ - 近似算法中, 他们所使用的局部交换规则是单交换, 其含义是在每次执行交换时关掉一个中心点并打开一个新的中心点, 迭代地执行单交换操作直到得到局部最优解 S . 该算法的具体执行过程可以参见算法 1.

算法 1 k - 中位问题的 $(5 + \epsilon)$ - 近似算法

输入: 需求点集合 C 、设施点集合 F 和正整数 k .

输出: 中心点集合 $S \subseteq F$.

1 (构造初始可行解) 从设施点集合 F 中任意选 k 个点作为初始可行解, 记为 S .

2 (局部搜索) 构造邻域 $N_1(S) := \{S - v + u, v \in S, u \in F - S\}$. 计算

$$\tilde{S} := \arg \min_{S' \in N_1(S)} \text{cost}(S').$$

3 (终止准则) 如果 $\text{cost}(\tilde{S}) \geq \text{cost}(S)$, 输出 S , 算法终止. 否则, 更新 $S := \tilde{S}$. 将每个需求点重新分配到最近的开设设施上, 并转回第 2 步.

在分析算法的近似比³⁾时, 需先通过局部最优解 S 和最优解 O 构造适当的交换对集合 $\mathcal{S} = \{\langle s, o \rangle\}$, 其中交换对 $\langle s, o \rangle$ 将 $s \in S$ 关掉并将 $o \in O$ 打开. 由于 S 的局部最优性, 集合 $\mathcal{S}$ 中所有的交换对均不能改进 S , 这意味着

$$\text{cost}(S - s + o) \geq \text{cost}(S), \quad \forall \langle s, o \rangle \in \mathcal{S}.$$

之后, 通过将这些不等式 (全部或部分) 进行整合得到 $\text{cost}(S) \leq \alpha \cdot \text{cost}(O)$, 从而得到最终的近似比.

更具体地, 在 k - 中位问题的 $(5 + \epsilon)$ - 近似算法中, 我们先定义最优解 O 与局部最优解 S 之间的俘获关系 $\sigma^* : O \rightarrow S$, 它将 O 中的每个中心点 o 映到 S 中距离 o 最近的中心点 $\sigma^*(o)$ 上, 由此得到最优解 O 与局部最优解 S 之间的俘获关系. 关于俘获关系的示意图见图 2.

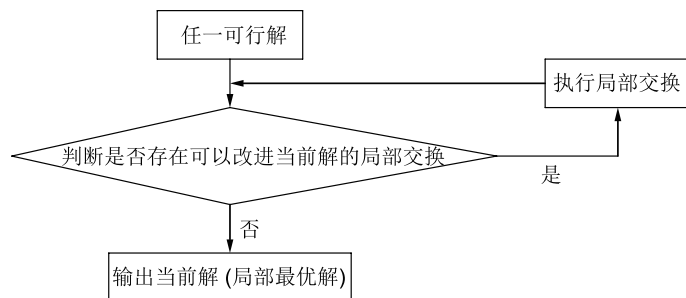


图 1 局部搜索算法执行框架

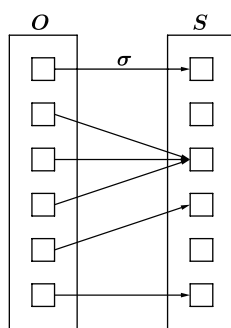


图 2 俘获关系示意图

根据俘获关系, 可以将 S 中的点分成 3 种类型: 俘获 O 中零个中心点、恰好俘获 O 中一个中心点和俘获 O 中至少两个中心点, 分别记为 S_0 , S_1 和 S_2 . 在构造和讨论交换对时, 我们将利用 S_1 中的中心点恰好一次、 S_0 中的中心点至多两次, 最终将相应的不等式累加便可得到

$$\text{cost}(S) \leq 5 \cdot \text{cost}(O).$$

进一步地, 通过考虑多交换 (即每次可以交换多个中心点⁴⁾), Arya 等^[1] 得到了 k -中位问题的 $(3 + \epsilon)$ -近似算法.

k -中位问题目前近似比最小的局部搜索算法是 Cohen-Addad 等^[13] 给出的 $(2.836 + \epsilon)$ -近似算法, 他们的关键创新点是引入势函数

$$\Phi(F') = \sum_{c \in C} (d(c, F') + \beta \cdot \min\{\alpha \cdot d(c, F'), d_2(c, F')\})$$

来代替 k -中位问题原有的目标函数 $\text{cost}(F') = \sum_{c \in C} d(c, F')$, 其中 $d_2(c, F')$ 是需求点 c 与集合 F' 中距离次近点之间的距离.

在 Cohen-Addad 等的算法中, 他们以势函数为新的目标函数, 并针对势函数执行局部搜索算法得到局部最优解 S . 最后, 将此局部最优解 S 作为 k -中位问题的近似解输出. 与 Arya 等^[1] 的 $(3 + \epsilon)$ -近似算法一样, Cohen-Addad 等的算法所使用的局部交换规则也是多交换.

在分析近似比时, 他们也是在上述局部最优解 S 与 k -中位问题的最优解 O 之间构造交换对进行对比. 通过构造并讨论简单交换 (simple swap) 和树交换 (tree swap), 最终证得 $\text{cost}(S) \leq 2.836 \cdot \text{cost}(O)$.

3.1.2 线性规划舍入算法

与无容量设施选址问题一样, k -中位问题的第一个常数近似算法也是通过线性规划舍入技巧得到的, 该算法的执行包含如下 4 步:

- (1) 求解线性规划松弛得到最优解;
- (2) 对需求点类似“聚类”进行合并;
- (3) 得到线性规划松弛的半整解 (决策变量取值于 $\{0, 1, \frac{1}{2}\}$);
- (4) 将半整解舍入成整数可行解.

Charikar 等^[7] 基于此框架给出了 k -中位问题的 $6\frac{2}{3}$ -近似算法. 之后, 在 2012 年, Charikar 和 Li^[9] 利用相关舍入得到了 k -中位问题的 3.25-近似算法.

3) 此处采用 Williamson 和 Shmoys^[44] 的分析.

4) 需要保证交换后仍是 k -中位问题的可行解.

3.2 两阶段算法

顾名思义, k -中位问题的两阶段算法的特点是由相对独立的两个阶段整合而成的, 其一般框架如图 3 所示.

在算法的第一阶段, 通常是借助无容量设施选址问题的原始 - 对偶算法或对偶拟合算法, 利用 k -中位问题的 Lagrange 松弛与无容量设施选址问题之间的联系 (如第 2.1 小节所述), 得到 k -中位问题的双点解. 所谓的双点解是通过设置不同的设施开设费用 λ_1 和 λ_2 , 先得到两个解 S_1 和 S_2 , 它们满足

$$|S_1| \leq k \leq |S_2|, \text{ 并且可取 } \alpha_1, \alpha_2 \in [0, 1] \text{ 使得 } \alpha_1|S_1| + \alpha_2|S_2| = k.$$

之后, 令 $S = \alpha_1 S_1 + \alpha_2 S_2$. 此时, 可以发现 $|S| = k$. 然而, 由于这样得到的 S 未必是整解, 因此需要在两阶段算法的第二阶段中进一步将它舍入成整解.

2001 年, Jain 和 Vazirani^[30] 设计得到了 k -中位问题的第一个两阶段近似算法, 他们算法的近似比是 6. 在他们算法的第一阶段, Jain 和 Vazirani 利用无容量设施选址问题原始 - 对偶 3-近似算法 $\mathcal{A}_{JV}$ 生成 k -中位问题的双点解 S . 之后, 在第二个阶段, 他们将双点解舍入成 k -中位问题真正的可行解, 其具体执行过程如算法 2 所示.

算法 2 k -中位问题的 6-近似算法

输入: 需求点集合 C 、设施点集合 F 和正整数 k .

输出: 中心点集合 $S \subseteq F$.

1 (初始化) 令 $\lambda_1 = \sum_{i \in F, j \in C} c_{ij}$, $\lambda_2 = 0$.

2 (产生双点解) 分别以 λ_1 和 λ_2 作为设施开设费用构造无容量设施选址问题实例并运行算法 $\mathcal{A}_{JV}$, 得到双点解 S_1 和 S_2 .

3 (二分法更新双点解) 如果 $|S_1| = k$ 或 $|S_2| = k$, 则算法终止, 将其作为近似解 S 并输出. 否则, 通过二分法更新 λ_1 或 λ_2 的取值为 $\lambda = \frac{1}{2}(\lambda_1 + \lambda_2)$, 转回第 2 步, 重新生成双点解.

4 (构造映射) 将 S_1 中的中心点映到 S_2 中离它最近的中心点上, 并将所得到的集合补至基数恰好为 $|S_1|$, 记所得到的集合为 S'_2 .

5 (舍入双点解) 以 α_1 的概率选择 S_1 , 若未选中 S_1 , 则选择 S'_2 ; 对 $S_2 \setminus S'_2$ 中的每个中心点, 以概率 $\frac{k - |S_1|}{|S_2| - |S_1|}$ 随机进行选取. 由此得到中心点集合 S .

算法 2 的第 1-3 步是两阶段算法的第一阶段, 在此阶段构造了 k -中位问题的双点解; 第 4 和 5 步是两阶段算法的第二阶段, 在此阶段将双点解舍入为 k -中位问题的可行解. 通过讨论连接费用, Jain 和 Vazirani 证明了算法 2 中舍入过程损失的近似比是 2, 因此该算法总的近似比是 6.

在 Jain 和 Vazirani 之后, Jain 等^[28] 通过对偶拟合技巧设计了无容量设施选址问题的 Lagrange 乘子保持 2-近似算法 $\mathcal{A}_{JMS}$. 用 $\mathcal{A}_{JMS}$ 算法代替算法 2 中的 $\mathcal{A}_{JV}$, Jain 等将生成 k -中位问题双点解过程损失的近似比改进到 2, 尔后利用与 Jain 和 Vazirani 相同的舍入过程得到了 k -中位问题的 4-近似算法.

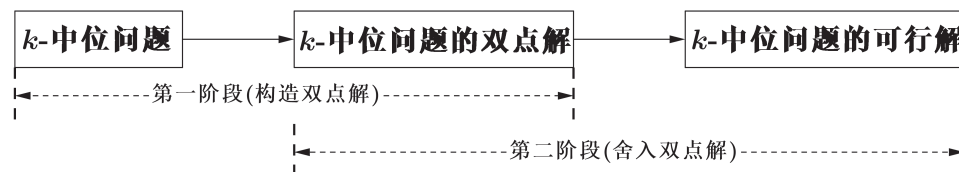


图 3 两阶段算法的一般框架

2023 年, Cohen-Addad 等^[12] 进一步尝试改进第一阶段损失的近似比, 他们给出了 k -中位问题的 $(2.67059 + \epsilon)$ -近似算法. 他们的主要创新点是将对偶拟合技巧和局部搜索技巧相结合得到改进的算法 $\mathcal{A}_{\text{CGLS}}$, 尔后利用算法 $\mathcal{A}_{\text{CGLS}}$ 来改进第一阶段损失的近似比. 具体而言, $\mathcal{A}_{\text{CGLS}}$ 是在执行完 Jain 等^[28] 的对偶拟合算法之后, 将所得到的解作为初始可行解执行局部搜索算法. 为了分析算法 $\mathcal{A}_{\text{CGLS}}$ 的性能, Cohen-Addad 等先在 Jain 等的因子揭示线性规划基础上, 对设施费用和连接费用的比值 (以下简称比值) 增加一个上界约束, 得到此特殊情形下新的因子揭示线性规划. 之后, 通过求解新的因子揭示线性规划, 他们得到此特殊约束下更好的近似比. 对于比值不满足此特殊约束的情形, 他们进一步借助局部搜索算法进行分析也得到了改进的近似比. 基于这样的研究框架, Cohen-Addad 等得到了 k -中位问题的 $(2.67059 + \epsilon)$ -近似算法. 笼统地说, Cohen-Addad 等的算法可以看作是将算法 2 中 $\mathcal{A}_{\text{JV}}$ 替换为 $\mathcal{A}_{\text{CGLS}}$ 得到的.

2016 年, Li 和 Svensson^[37] 引入了伪解的定义, 并证明了可以通过找到具有性能保证的伪解来设计 k -中位问题的近似算法. 这样, 在对 k -中位问题双点解的舍入过程中, 便有了更大的操作空间: 可以通过舍入成伪解来设计算法. Li 和 Svensson 通过将双点解舍入成伪解, 将舍入过程损失的近似比由 2 改进到了 $\frac{1+\sqrt{3}+\epsilon}{2}$. 由此, 他们给出了 k -中位问题的 $(1 + \sqrt{3} + \epsilon)$ -近似算法. 之后, 2017 年, Byrka 等^[4] 进一步利用相关舍入技巧, 将上述近似比损失由 $\frac{1+\sqrt{3}+\epsilon}{2}$ 改进到了 $1.3371 + \epsilon$, 他们得到了 k -中位问题的 $(2.675 + \epsilon)$ -近似算法. 粗略地讲, Li 和 Svensson 的算法以及 Byrka 等的算法与算法 2 的区别主要体现在第二阶段的舍入过程中.

当前, 对于 k -中位问题近似比最小的近似算法由 Gowda 等^[18] 在 2023 年给出, 他们的算法也是基于算法 2 的框架. 他们通过引入“间隔点”将算法 2 中的 S_1 , S'_2 和 $S_2 \setminus S'_2$ 分隔成多个子集, 之后根据这些子集设计逐渐加细的“算法族”来确定开设设施. 借此, Gowda 等将双点解舍入过程中损失的近似比降低到了 1.3064, 由此得到了 k -中位问题的 2.613-近似算法.

4 讨论

根据第 2 和 3 节的研究现状, 总结出下面一些值得进一步研究的问题.

(1) 与无容量设施选址问题的近似间隙 $[1.463, 1.488]$ 相比, 经典的 k -中位问题的近似间隙 $[1 + \frac{2}{e}, 2.613]$ 仍然比较大. 这既说明我们对于 k -中位问题的认识和理解还不够深刻, 也说明 k -中位问题仍存在更广阔的研究空间. 根据我们的认识, 一个可能的改进上界的思路是, 从连续化的角度考虑 Gowda 等的算法, 设计连续化的舍入过程.

(2) 以下在线 k -中位问题仍有待进一步研究. 在在线 k -中位问题中, 需求点是随时间到达的, 每当一个新的需求点到达时需要立即对它进行分配. 在线 k -中位问题要在保证开设设施点不超过 k 个的前提下, 使得总连接费用尽可能小.

(3) 相较于容错无容量设施选址问题, 容错 k -中位问题的结果较少, 其近似比也有待进一步改进.

参考文献

- 1 Arya V, Garg N, Khandekar R, et al. Local search heuristics for k -median and facility location problems. *SIAM J Comput*, 2004, 33: 544–562
- 2 Byrka J, Aardal K. An optimal bifactor approximation algorithm for the metric uncapacitated facility location problem. *SIAM J Comput*, 2010, 39: 2212–2231
- 3 Byrka J, Pensyl T, Rybicki B, et al. An improved approximation algorithm for knapsack median using sparsification. *Algorithmica*, 2015, 80: 1093–1114

- 4 Byrka J, Pensyl T, Rybicki B, et al. An improved approximation for k -median and positive correlation in budgeted optimization. *ACM Trans Algorithms*, 2017, 13: 1–31
- 5 Byrka J, Rybicki B, Uniyal S. An approximation algorithm for uniform capacitated k -median problem with $1 + \epsilon$ capacity violation. In: *Proceedings of the 18th International Conference on Integer Programming and Combinatorial Optimization*. Liège: Springer, 2016, 262–274
- 6 Charikar M, Guha S. Improved combinatorial algorithms for the facility location and k -median problems. In: *Proceedings of the 40th Annual Symposium on Foundations of Computer Science*. New York: IEEE, 1999, 378–388
- 7 Charikar M, Guha S, Tardos É, et al. A constant-factor approximation algorithm for the k -median problem. *J Comput System Sci*, 2002, 65: 129–149
- 8 Charikar M, Khuller S, Mount D M, et al. Algorithms for facility location problems with outliers. In: *Proceedings of the 12th Annual ACM-SIAM Symposium on Discrete algorithms*. Philadelphia: SIAM, 2001, 642–651
- 9 Charikar M, Li S. A dependent LP-rounding approach for the k -median problem. In: *Proceedings of the 39th International Colloquium on Automata, Languages, and Programming*. Berlin: Springer, 2012, 194–205
- 10 Chen K. A constant factor approximation algorithm for k -median clustering with outliers. In: *Proceedings of the 19th Annual ACM-SIAM Symposium on Discrete Algorithms*. Philadelphia: SIAM, 2008, 826–835
- 11 Chudak F A, Shmoys D B. Improved approximation algorithms for the uncapacitated facility location problem. *SIAM J Comput*, 2003, 33: 1–25
- 12 Cohen-Addad V, Grandoni F, Lee E, et al. Breaching the 2 LMP approximation barrier for facility location with applications to k -median. In: *Proceedings of the 34th Annual ACM-SIAM Symposium on Discrete Algorithms*. Philadelphia: SIAM, 2023, 940–986
- 13 Cohen-Addad V, Gupta A, Hu L, et al. An improved local search algorithm for k -median. In: *Proceedings of the 33rd Annual ACM-SIAM Symposium on Discrete Algorithms*. Philadelphia: SIAM, 2022, 1556–1612
- 14 Demirci G, Li S. Constant approximation for capacitated k -median with $(1 + \epsilon)$ -capacity violation. In: *Proceedings of the 43rd International Colloquium on Automata, Languages, and Programming*. Rome: Springer-Verlag, 2016, 73:1–73:14
- 15 Friggstad Z, Khodamoradi K, Rezapour M, et al. Approximation schemes for clustering with outliers. *ACM Trans Algorithms*, 2019, 15: 1–26
- 16 Garey M R, Johnson D S. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. San Francisco: W. H. Freeman, 1979
- 17 Gonzalez T F. Clustering to minimize the maximum intercluster distance. *Theoret Comput Sci*, 1985, 38: 293–306
- 18 Gowda K N, Pensyl T, Srinivasan A, et al. Improved bi-point rounding algorithms and a golden barrier for k -median. In: *Proceedings of the 34th Annual ACM-SIAM Symposium on Discrete Algorithms*. Philadelphia: SIAM, 2023, 987–1011
- 19 Guha S, Khuller S. Greedy strikes back: Improved facility location algorithms. *J Algorithms*, 1999, 31: 228–248
- 20 Guo Y T, Huang J Y, Zhang Z. A constant factor approximation for lower-bounded k -median. In: *Proceedings of the 16th Annual Conference on Theory and Applications of Models of Computation*. Berlin: Springer, 2020, 119–131
- 21 Gupta S, Kumar R, Lu K F, et al. Local search methods for k -means with outliers. In: *Proceedings of the 43rd International Conference on Very Large Data Bases*. Munich: VLDB Endowment, 2017, 757–768
- 22 Hajiaghayi M, Hu W, Li J, et al. A constant factor approximation algorithm for fault-tolerant k -median. *ACM Trans Algorithms*, 2016, 12: 1–19
- 23 Hajiaghayi M, Khandekar R, Kortsarz G. Local search algorithms for the red-blue median problem. *Algorithmica*, 2012, 63: 795–814
- 24 Han L, Hao C L, Wu C C, et al. Approximation algorithms for the lower-bounded k -median and its generalizations. In: *Proceedings of the 26th International Computing and Combinatorics Conference*. Berlin: Springer, 2020, 627–639
- 25 Han L, Hao C L, Wu C C, et al. Approximation algorithms for the lower-bounded knapsack median problem. In: *Algorithmic Aspects in Information and Management. Lecture Notes in Computer Science*, vol. 12290. Cham: Springer, 2020, 119–130
- 26 Hochbaum D S, Shmoys D B. A best possible heuristic for the k -center problem. *Math Oper Res*, 1985, 10: 180–184
- 27 Hochbaum D S, Shmoys D B. A unified approach to approximation algorithms for bottleneck problems. *J ACM*, 1986, 33: 533–550
- 28 Jain K, Mahdian M, Markakis E, et al. Greedy facility location algorithms analyzed using dual fitting with factor-revealing LP. *J ACM*, 2003, 50: 795–824
- 29 Jain K, Mahdian M, Saberi A. A new greedy approach for facility location problems. In: *Proceedings of the 34th Annual ACM Symposium on Theory of Computing*. Montreal: ACM, 2002, 731–740
- 30 Jain K, Vazirani V V. Approximation algorithms for metric facility location and k -median problems using the primal-dual schema and Lagrangian relaxation. *J ACM*, 2001, 48: 274–296
- 31 Korupolu M R, Plaxton C G, Rajaraman R. Analysis of a local search heuristic for facility location problems. *J Algorithms*, 2000, 37: 146–188
- 32 Krishnaswamy R, Kumar P, Nagarajan V, et al. The matroid median problem. In: *Proceedings of the 22nd Annual*

- ACM-SIAM Symposium on Discrete Algorithms. Philadelphia: SIAM, 2011, 1117–1130
- 33 Kumar A. Constant factor approximation algorithm for the knapsack median problem. In: Proceedings of the 23rd Annual ACM-SIAM Symposium on Discrete Algorithms. Philadelphia: SIAM, 2012, 824–832
- 34 Li S. A 1.488 approximation algorithm for the uncapacitated facility location problem. Inform and Comput, 2013, 222: 45–58
- 35 Li S. Approximating capacitated k -median with $(1+\epsilon)k$ open facilities. In: Proceedings of the 27th Annual ACM-SIAM Symposium on Discrete Algorithms. Arlington Philadelphia: SIAM, 2016, 786–796
- 36 Li S. On uniform capacitated k -median beyond the natural LP relaxation. ACM Trans Algorithms, 2017, 13: 1–18
- 37 Li S, Svensson O. Approximating k -median via pseudo-approximation. SIAM J Comput, 2016, 45: 530–547
- 38 Mahdian M, Ye Y Y, Zhang J W. Improved approximation algorithms for metric facility location problems. In: Proceedings of the 5th International Workshop on Approximation Algorithms for Combinatorial Optimization. Berlin: Springer, 2002, 229–242
- 39 Mettu R R, Plaxton C G. The online median problem. SIAM J Comput, 2003, 32: 816–832
- 40 Shmoys D B, Tardos É, Aardal K. Approximation algorithms for facility location problems. In: Proceedings of 29th Annual ACM Symposium on Theory of Computing. New York: ACM, 1997, 265–274
- 41 Sviridenko M. An improved approximation algorithm for the metric uncapacitated facility location problem. In: Proceedings of the 9th International Conference on Integer Programming and Combinatorial Optimization. Cambridge: Springer, 2002, 240–257
- 42 Swamy C. Improved approximation algorithms for matroid and knapsack median problems and applications. ACM Trans Algorithms, 2016, 12: 1–22
- 43 Swamy C, Shmoys D B. Fault-tolerant facility location. ACM Trans Algorithms, 2008, 4: 1–27
- 44 Williamson D P, Shmoys D B. The Design of Approximation Algorithms. Cambridge: Cambridge Univ Press, 2011
- 45 Wu C C, Du D L, Xu D C. An approximation algorithm for the k -median problem with uniform penalties via pseudo-solution. Theoret Comput Sci, 2018, 749: 80–92
- 46 Wu C C, Möhring R H, Wang Y S, et al. Approximation algorithms for robust clustering problems using local search techniques. In: Proceedings of the 18th Annual Conference on Theory and Applications of Models of Computation. Berlin: Springer, 2024, 197–208
- 47 Wu X L, Shi F, Guo Y T, et al. An approximation algorithm for lower-bounded k -median with constant factor. Sci China Inf Sci, 2022, 65: 140601
- 48 Xu D C, Du D L, Wu C C. A Survey on approximation algorithms for facility location problems (in Chinese). Adv Math (China), 2014, 43: 801–816 [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, 2014, 43: 801–816]
- 49 Xu D C, Zhang J W. Approximation Algorithms for the Facility Location Problem (in Chinese). Beijing: Science Press, 2014, 43: 801–816 [徐大川, 张家伟. 设施选址问题的近似算法. 北京: 科学出版社, 2013]
- 50 Zhang P. A new approximation algorithm for the k -facility location problem. Theoret Comput Sci, 2007, 384: 126–135

Survey on approximation algorithms for the k -median problem

Chenchen Wu, Donglei Du, Runjie Miao & Dachuan Xu

Abstract The k -median problem is one of the classical topics in the fields of theoretical computer science and combinatorial optimization, which is widely applied to image processing, pattern recognition, supply chain management, and data mining. Nowadays, with the development of big data and machine learning techniques, the practical applications of the k -median problem become more and more complex. This leads to a variety of new research topics that need to be solved urgently. In this paper, we introduce effective algorithms based on linear programming rounding, primal-dual, dual-fitting, local search, Lagrange relaxation, and other techniques for the classical k -median problem and its variants. We begin with the close relation between the k -median problem and the uncapacitated facility location problem. Then we introduce several important variants of the k -median problem including the k -center problem, k -facility location problem, knapsack median problem, fault-tolerant k -median problem, capacitated k -median problem, k -median problem with lower bounds, k -median problem with penalties, k -median problem with outliers, and online median problem. Finally, we survey the approximation algorithms for the k -median problem and propose several open problems in this area.

Keywords k -median problem, approximation algorithm, linear program, local search

MSC(2020) 65K05, 65K10, 68W25

doi: 10.1360/SSM-2024-0171