第 17 卷第 9 期 管 理 科 学 学 报 Vol. 17 No. 9
2014 年 9 月 JOURNAL OF MANAGEMENT SCIENCES IN CHINA Sep. 2014

欧氏平面上的占线中位选址问题分析①

代文强，李仕明
(电子科技大学经济与管理学院，成都 610054)

摘要: 在一般的占线中位选址问题模型的基础上,基于实际选址问题是限制在一个欧氏平面上进行选址决策的现实背景,提出并研究了欧氏平面上的占线中位选址问题. 通过对问题的结构特性的研究,设计了一个多项式时间的竞争算法,证明了该算法具有较好的常数竞争比.

关键词: 选址; 占线中位; 欧氏平面; 竞争比
中图分类号: O221.7 **文献标识码:** A **文章编号:** 1007 -9807(2014)09 -0088 -07

0 引 言

选址问题是组合优化、管理科学等诸多领域中的重要问题,在网络结构分析、企业战略规划等多方面都有着广泛应用[1-4]. k-median(k-中位) 问题是其中的一个最基本的经典选址问题模型. 占线中位选址模型(online median problem②) 是近年来才开始考虑的一类新的重要模型[5],与经典的同时(simultaneous) 选址 k-中位问题模型不同的是占线中位问题是多阶段的选址决策问题模型,考虑的实际背景是首先选址决策者建立初始中位,然后在此基础上增加建设新的设施,同时不拆除已经建立好的设施. 在此背景下,占线中位选址问题模型将实际选址决策考虑成是个多阶段的选址中位序列优化问题,模型的输出是选址中位增加序列,目标是保证该选址序列的各个阶段的费用都同在该阶段的最优选址方案所对应的费用都相差不远. 对每一个给定的待选址中位个数 k,考虑序列的 k 元前序的费用同相应的最优 k-中位模型解费用的比值,定义序列的竞争比为所有这些 k 值中的最大比值. 显然,该竞争比应越小越好. 该模型自从被提出后,受到了越来越多的研究和重视,如文献[6-10].

在以往针对占线中位选址问题模型的研究中,问题的输入是给定的网络,即直接给定离散网络图和图上边的长度. 本文考虑将该网络限制在一个欧氏平面上,即考虑所有的待选址点落在同一个欧氏平面,这样每一个点可赋予一个平面坐标,同时该点的位置可以由坐标数据唯一确定,两点间的距离可以由两点间的欧氏距离决定. 该问题是以往一般性的占线中位问题下具有新的附加特定输入结构的问题. 该问题更加符合现实,这是因为实际的选址问题输入本身就是一个限制在欧氏空间中的优化决策问题,而后者的应用广泛出现在如网络可靠控制、集成电路设计、网络布局等诸多领域[1-2,11-12].

注意到即使是限制在欧氏空间上,欧氏平面上的离散 k-median 仍然是 NP-hard 的[13],因此对欧氏平面上的占线中位问题的研究仍然非常困难. 本文将分析欧氏平面上占线中位问题的结构特性,设计出多项式时间的竞争算法,最后证明该算法具有常数竞争比,该竞争比结果改进了已有的结果.

① 收稿日期: 2012 -06 -28; 修订日期: 2013 -06 -01.
基金项目: 国家自然科学基金重大研究计划资助项目(91224001); 国家自然科学基金资助项目(70901012); 高等学校博士学科点专项科研基金资助项目(200806141084); 中央高校基本科研业务费专项资金资助项目(ZYGX2013J134).
作者简介: 代文强(1978—),男,四川彭州人,副教授,博士. Email: wqdai@ uestc. edu. cn

② 在文献[6]等文献中,将“online median problem”翻译成“占线中心问题”,论文初稿沿用该名词翻译. 审稿人建议将其翻译成“占线中位问题”,这里遵照他们的意见进行了修改,但请读者注意是同一个问题.

1　欧氏平面上的占线中位选址问题模型及相关研究

定义 1　定义欧氏平面上的离散度量 k-median 问题为：给定欧氏平面上的一个有限点集 $U = \{x_1, x_2, \cdots, x_n\}$，$x_i \in R^2$，以及相应的正权重 $w(x_1), w(x_2), \cdots, w(x_n) \in R^+$，对该平面上任意两点 A 和 B，设其坐标向量分别是 a 与 b，则其距离定义为 $d(A,B) = \| a - b \| = \sqrt{(a-b)^{\mathrm{T}}(a-b)}$. 给定待选址个数 k，$1 \leqslant k \leqslant n$，在所有元素个数至多为 k 的 U 的子集合中，最小化下述目标函数

$$cost(S) = \sum_{x \in U} d(x,S)\, w(x)$$

这里，点 x 到点集 S 的距离定义为 $d(x,S) = \min_{y \in S} d(x,y)$. 与此相对应的一般的离散度量 k-median 问题的输入是同时给定有限点集 U 及 U 上的边的权重，并假设该边的权重满足度量的条件. 由于欧氏平面上的 k-median 问题边的权重此时计算采用的是欧氏距离函数，它必定构成一个度量，因此欧氏平面上的离散度量 k-median 问题同样也是离散度量 k-median 问题.

定义 2[5-10]　定义一般的度量空间占线中位选址问题为：给定离散度量 k-median 问题的实例，寻求一个待选址集合序列 F_i，$i = 1,2,\cdots,n$，使得 $|F_i| \leqslant i$（定义符号 $|S|$ 表示点集合 S 中的点的个数），$F_1 \subseteq F_2 \subseteq \cdots \subseteq F_n \subseteq U$，且 $cost(F_i) \leqslant r \times opt_i$，其中 opt_i 是 i-median 问题的最优解费用，$r \geqslant 1$ 是常数，称之为竞争比，目标是使得 r 越小越好.

本文考虑的问题即是将定义 2 中待选址集合限制在定义 1 所定义的欧氏平面上，此时定义 2 中的 opt_i 即是欧氏平面上的 i-median 问题的最优解费用. 同样，由于欧氏平面上的相应问题边权重的输入满足度量的要求，因此对一般的度量空间下占线中位选址问题成立的结论，对本文研究的问题也同样成立.

针对一般的度量空间占线中位选址问题，Mettu 和 Plaxton[5] 给出了多项式时间下竞争比是 29.98 的竞争算法，Chrobak 等[7] 和 Lin 等[8] 独立给出了竞争比是 $8c$ 的竞争算法，其中 c 是 k-median 问题能够达到的近似比，代文强等[6] 给出了略好于上述的结果. Lin 等[8] 同时给出了竞争比为 16 的多项式时间竞争算法. 针对欧氏平面上的占线中位问题，以往还没有文献进行专门研究，但根据如上的研究结果可知，由于欧氏平面上的 k-中位问题存在多项式近似方案（PTAS）[14-15]，即对任意给定的 $\varepsilon > 0$，算法能在多项式时间内给出近似比为 $(1+\varepsilon)$ 的近似解，因此，此时的 c 可取为 $1+\varepsilon$，故根据以往的文献，欧氏平面上的占线中位选址问题存在竞争比至多为 29.86、8 和 16 的竞争算法.

本文将对这个问题进行研究，通过对问题的一个结构特征进行分析，给出一个多项式时间下的竞争算法，最后证明算法具有的常数竞争比至多为 $4\sqrt{2} \approx 5.65$，因此极大地改进了已有的结果.

2　一些定义及性质

定义 3　对给定的欧氏平面上离散 k-median 选址问题，分别定义 Δ_e 和 Δ_w 表示给定点集间的最大相对距离和最大相对权重. 即

$$\Delta_e = \frac{\mathrm{maxdistance}}{\mathrm{mindistance}}, \quad \Delta_w = \frac{\mathrm{maxweight}}{\mathrm{minweight}}$$

其中

$$\mathrm{maxdistance} = \max_{u,v \in U} d(u,v),$$

$$\mathrm{mindistance} = \min_{u,v \in U, u \neq v} d(u,v),$$

$$\mathrm{maxweight} = \max_{x \in U} w(x),$$

$$\mathrm{minweight} = \min_{x \in U} w(x)$$

下面的引理 1 来自于文献[6]，其证明这里省略.

引理 1　若 $1 \leqslant i \leqslant n-1$，有 $opt_{i+1} \leqslant opt_i$ 且有 $opt_1 \leqslant (n-1)\Delta_e \Delta_w opt_{n-1}$.

定义 4　对任意的一个集合 Y 和点 x，定义 $d_Y(x)$ 表示集合 Y 中距离 x 最近的点，即

$$d(x, d_Y(x)) = d(x,Y) = \min_{y \in Y} d(x,y)$$

定义 5　对给定的 median 问题的解 F，$\forall f \in F$，定义 $\pi_F^{-1}(f)$ 表示在解 F 下，由 f 服务的点的集合，即

$$\pi_F^{-1}(f) = \{x \mid x \in U, d(x,f) < d(x,g), \forall g \in F, f \neq g\}$$

定义 6 记 $cost_X(Y)$ 表示选址点集为 Y，服务点集为 X 时的费用函数，即

$$cost_X(Y) = \sum_{x \in X} w(x) \min_{y \in Y} d(x, y)$$

在定义 6 下，定义 1 中的 $cost(S)$ 即是 $cost_U(S)$.

引理 2 给定欧氏平面上 k-median 问题的输入，给定 $B \subseteq U$ 和任意一个整数 t，设 $t \leqslant |B|$，则存在 $A \subseteq U$，使 $|A| = t$，且 $cost(A) \leqslant cost(B) + \sqrt{2}c \cdot opt_t$，其中 c 是欧氏平面上离散 k-median 问题近似算法能达到的近似比.

证明 给定欧氏平面上 k-median 问题的输入和点集合 B，t，由于 $|B| \geqslant t$，以 B 为点集构造一个新的 t-median 问题的例子如下，距离保持不变，每个点 $f \in B$ 的权重定义为 $\sum_{x \in \pi_B^{-1}(f)} w(x)$. 对这个新的问题寻求 t-median 问题的近似最优解 A. 则有 $|A| = t$，$A \subseteq B$. 接下来只需证明费用不等式. 首先，根据三角不等式知，$\forall x \in U$，$d(x, d_A(x)) \leqslant d(x, d_A(d_B(x))) \leqslant d(x, d_B(x)) + d(d_B(x), d_A(d_B(x)))$，即有 $cost(A) \leqslant cost(B) + cost_B(A)$，且根据 A 的构造方法有 $cost_B(A) \leqslant c \min_{Y \subseteq B, |Y| = t} cost_B(Y)$.

现考虑另一个 t-median 问题(P')，其设施可选集为 U，服务点集仍然保持为 B，设对(P') 的最优解为 Y^*，即 $cost_B(Y^*) = \min_{Y \subseteq U, |Y| = t} cost_B(Y)$. 将服务点集 B 按照设施点集 Y^* 中的各个元素所分别服务的服务点集进行划分，可以分成 t 个不相交子集的并，即有 $B = \bigcup_{f \in Y^*} \pi_{Y^*}^{-1}(f)$. 现任取其中一个子集 $\pi_{Y^*}^{-1}(f)$. $\forall q \in \pi_{Y^*}^{-1}(f)$，则有 $f = d_{Y^*}(q)$. 记 $f_q = d_B(f)$. 若 $f \in B$，则必有 $f = f_q$，则有 $d(q, f) = d(q, f_q)$. 若 $f \notin B$，则 $f \neq f_q$，据假设 $f_q = d_B(f)$，而 $q \in B$，则 $d(f_q, f) \leqslant d(q, f)$. 故总有 $d(f_q, f) \leqslant d(q, f)$. 现对给定的待服务点集合 $\pi_{Y^*}^{-1}(f)$，所有的服务点都在距离圆心为 f，半径为 $d(f_q, f)$ 的圆上或之外，此时待选址点或者在该圆上，或者在该圆外，或者在圆心 f. 根据假设知，选址点取为 f 时费用最小. 但根据对称性，可知道该结论针对圆上也同样成立. 故可以假设该结论是针对所有的欧氏平面上的待选址点都成立. 换句话说，选址点 f 为连续平面上，给定待服务点集合为 $\pi_{Y^*}^{-1}(f)$ 的 Weber 问题[16-17]（即最小化函数 $f(\{P\}) = \sum_{q \in \pi_{Y^*}^{-1}(f)} \| P - q \| w(q)$）这一连续凸优化问题的唯一最优解. 因此凸目标函数 $f(\{P\}$ 在该点 f 的负梯度为零，有(以下为了表述方便，将混合使用点与其对应的坐标向量)

$$\sum_{q \in \pi_{Y^*}^{-1}(f)} \frac{f - q}{\| f - q \|} w(q) = 0 \tag{1}$$

式(1)的左边是一个坐标向量，右边 0 表示一个零向量. 故有

$$\sum_{q \in \pi_{Y^*}^{-1}(f)} d(q, f_q) w(q) = \sum_{q \in \pi_{Y^*}^{-1}(f)} \sqrt{(q - f_q)^{\mathrm{T}}(q - f_q)}\, w(q) \tag{2}$$

$$\sum_{q \in \pi_{Y^*}^{-1}(f)} d(q, f_q) w(q) = \sum_{q \in \pi_{Y^*}^{-1}(f)} \sqrt{(q - f)^{\mathrm{T}}(q - f) + 2(q - f)^{\mathrm{T}}(f - f_q) + (f - f_q)^{\mathrm{T}}(f - f_q)}\, w(q) \tag{3}$$

$$\sum_{q \in \pi_{Y^*}^{-1}(f)} d(q, f_q) w(q) \leqslant \sum_{q \in \pi_{Y^*}^{-1}(f)} \sqrt{2(q - f)^{\mathrm{T}}(q - f) + 2(q - f)^{\mathrm{T}}(f - f_q)}\, w(q) \tag{4}$$

$$\sum_{q \in \pi_{Y^*}^{-1}(f)} d(q, f_q) w(q) \leqslant \sqrt{2} \sum_{q \in \pi_{Y^*}^{-1}(f)} \left\{ \sqrt{(q - f)^{\mathrm{T}}(q - f)} + (q - f)^{\mathrm{T}}(f - f_q) 2\sqrt{(q - f)^{\mathrm{T}}(q - f)} \right\} w(q) \tag{5}$$

$$\sum_{q \in \pi_{Y^*}^{-1}(f)} d(q, f_q) w(q) = \sqrt{2} \sum_{q \in \pi_{Y^*}^{-1}(f)} d(q, f) w(q) \tag{6}$$

这里：式(2)是根据距离的定义；式(3)是将该向量点积展开；式(4)根据前面推导的结论 $d(f_q, f) \leqslant d(q, f)$；式(5)是根据若 $a \geqslant 0$，b 任意，$a + b \geqslant 0$，则 $\sqrt{a + b} \leqslant \sqrt{a} + b 2\sqrt{a}$；式(6)是根据式(1)的结论. 从而对 $f \in Y^*$ 求和可得

$$\sum_{q \in B} d(q, f_q) w(q) \leqslant \sqrt{2} \sum_{f \in Y^*} \sum_{q \in \pi_{Y^*}^{-1}(f)} d(q, f) w(q) = \sqrt{2} cost_B(Y^*)$$

另一方面，对问题(P')，考虑问题(P'') 的 t -

median 问题,问题(P'') 中设施可选集保持不变,仍然为 U,但服务点集由 B 变化为 U,该问题即为原始的问题. 则由于此时需要服务的点的个数增加,最优费用必增加,即有

$$cost_B(Y^*) \leqslant \min_{Y\subseteq U,|Y|=t} cost_U(Y) = opt_t$$

综合上述分析,得到

$$cost(A) \leqslant cost(B) + \sqrt{2}c \cdot opt_t$$ 证毕.

3 竞争算法

根据第 3 节建立的结论,可以如下构造竞争算法. 首先注意到当 $k=n$ 时,U 只有一个子集,即 U 本身. 因此只需要构造 $k=1,2,\cdots,n-1$ 时 U 的选址子集合序列.

竞争算法

输入: 欧氏平面上的点集合 U,以及相应赋予的正权重函数 w: $U \to R^+$.

输出: 待选址子集序列 F_i,$i=1,2,\cdots,n-1$,,满足 $F_1 \subseteq \cdots \subseteq F_{n-1} \subseteq F_n = U$.

步骤 1 利用已有的多项式近似方案,对每一个给定的选址点 k ($k=1,2,\cdots,n-1$),求解对应的 k-median 问题,产生近似比为 c 的近似解. 设得到的解的序列为 $\{F_1^*, F_2^*, \cdots, F_{n-1}^*\}$,相应的费用序列是 $W=\{cost(F_1^*), cost(F_2^*), \cdots, cost(F_{n-1}^*)\}$.

步骤 2 令 $T(1)=1$,且对 $k=2,\cdots,m$,令 $T(k+1)$ 是在 W 中满足下述不等式的最小下标 $cost(F^*_{T(k+1)}) \leqslant cost(F^*_{T(k)})q$,其中参数 $q>1$ 在后面证明中确定,并且 m 使得 $T(m)=n-1$. 记得到的下标序列为 $E=\{T(1), T(2), \cdots, T(m)\}$.

步骤 3 令 $F_{T(m)}=F^*_{n-1}$,对 $t=m-1,\cdots,1$,F_t 是按照引理 2 的方法得到的集合 A,其中 B 取为 $F^*_{T(t+1)}$.

步骤 4 对 $t \in \{1,2,\cdots,n\}\backslash E$,设 $T(k)<t<T(k+1)$,此时已计算得到 $F_{T(k)}$ 与 $F_{T(k+1)}$,且 $F_{T(k)} \subset F_{T(k+1)}$. 对 $t=T(k+1)-1,\cdots,T(k)+1$,令 $F_t=F_{T(k)}$.

步骤 5 算法结束.

根据算法,上述方法得到的序列 $\{F_1, F_2, \cdots, F_{n-1}, F_n\}$ 满足 $F_1 \subseteq F_2 \subseteq \cdots \subseteq F_{n-1} \subseteq F_n$,且根据引理 3 知 $\forall t=1,2,\cdots,n$,$\|F_t\| \leqslant t$,即得到的序列必是一个可行的占线中位序列. 下面首先分析算法的时间复杂度. 根据文献 [15],对任意的 $\varepsilon>0$,最好的多项式近似方案最多可在 $O(2^{O(\log 1\varepsilon\varepsilon)} n \log^8 n)$ 时间内产生近似比为 $c=1+\varepsilon$ 的近似解. 算法第 1 步需要对每一个 k 调用该程序,故该步最多需要 $O(2^{O(\log 1\varepsilon\varepsilon)} n^2 \log^8 n)$. 算法的第 2 步总共需要 $O(n)$ 时间即可完成. 根据引理 2 的证明过程,算法第 3 步中需要计算 t-median问题的近似最优解,同样,该步骤在最坏情形下至多为 $O(2^{O(\log 1\varepsilon\varepsilon)} n^2 \log^8 n)$. 算法第 4 步仅需要 $O(n)$ 时间即可完成. 故算法总的时间复杂度为 $O(2^{O(\log 1\varepsilon\varepsilon)} n^2 \log^8 n)$,对于固定的 ε,该时间是一个多项式时间. 因此得到下面的定理 1.

定理 1 竞争算法能在多项式时间内完成.

下面只需考察费用不等式. $\forall t \in \{1,2,3,\cdots,n\}$,分两种情况讨论.

情况 1 若 $t \in E$,设 $T(k)=t$,则根据引理 2 得到

$$\begin{aligned} cost(F_{T(k)}) &\leqslant cost(F_{T(k+1)}) + \sqrt{2}c \cdot opt_{T(k)} \\ &\leqslant cost(F_{T(k+1)}) + \sqrt{2}c \cdot cost(F^*_{T(k)}) \end{aligned}$$

则有

$$\begin{aligned} cost(F_t) &= cost(F_{T(k)}) \leqslant cost(F_{T(k+1)}) + \\ &\quad \sqrt{2}c \cdot cost(F^*_{T(k)}) \\ &\leqslant cost(F_{T(k+2)}) + \sqrt{2}c \cdot cost(F^*_{T(k+1)}) + \\ &\quad \sqrt{2}c \cdot cost(F^*_{T(k)}) \leqslant \cdots \\ &= cost(F^*_{T(m)}) + \sqrt{2}c \sum_{j=0}^{m-k-1} cost(F^*_{T(k+j)}) \end{aligned}$$

但根据算法中下标集 E 的取法规则有,$cost(F^*_{T(k+j)}) \leqslant cost(F^*_{T(k)})q^j$,因此有

$$\begin{aligned} cost(F_t) &\leqslant \frac{cost(F^*_{T(k)})}{q^{m-k}} + \sqrt{2}c \sum_{j=0}^{m-k-1} (1/q)^j cost(F^*_{T(k)}) \\ &= \left(\frac{1}{q^{m-k}} + \sqrt{2}c\,\frac{1-1/q^{m-k}}{1-1/q}\right) cost(F^*_{T(k)}) \end{aligned}$$

即对此种情况的竞争比为 $\left(1/q^{m-k} + \sqrt{2}c\,\frac{1-1/q^{m-k}}{1-1/q}\right)c$.

情况 2 若 $t \notin E$,则根据算法步骤 4,必有 $T(k) \in E$,使 $T(k)<t<T(k+1)$,且有 $cost(F_t)=cost(F_{T(k)})$. 根据情况 1 的结果,得到

$$cost(F_t) \leqslant \left(\frac{1}{q^{m-k}} + \sqrt{2}c\,\frac{1-1/q^{m-k}}{1-1/q}\right) cost(F^*_{T(k)})$$

另一方面,由于 $T(k+1)$ 是满足 $cost(F^*_{T(k+1)}) \leqslant cost(F^*_{T(k)})/q$ 的最小下标,因此有 $cost(F^*_t) \geqslant$

$cost(F^*_{T(k)})/q$，故得到

$$cost(F_t) \leqslant \left(\frac{1}{q^{m-k-1}} + \sqrt{2}c\,\frac{q-1/q^{m-k-1}}{1-1/q}\right)cost(F_t^*)$$

即这种情况下的竞争比至多为

$$\left(\frac{1}{q^{m-k-1}} + \sqrt{2}c\,\frac{q-1/q^{m-k-1}}{1-1/q}\right)c$$

综合上述两种情况，算法的竞争比为上述两种情况下的相应竞争比的最大值，注意到 $q > 1$ 及 $1 \leqslant k \leqslant m$，因此得到竞争比为

$$c\left(\frac{1}{q^{m-k-1}} + \sqrt{2}c\,\frac{q-1/q^{m-k-1}}{1-1/q}\right) = c\left(\frac{\sqrt{2}qc + 1/q^{m-k-1}(1-1/q-\sqrt{2}c)}{1-1/q}\right) \leqslant c\,\frac{\sqrt{2}qc + \frac{1}{q^{m-2}}(1-1/q-\sqrt{2}c)}{1-1/q}$$

更进一步，由于 $cost(F^*_{T(k+j)}) \leqslant cost(F^*_{T(k)})q^j$，即有 $q^{m-1} \leqslant cost(F_1^*)\,cost(F^*_{T(m-1)}) \leqslant c \cdot opt_1 opt_{n-1}$. 根据引理 1，令 $M = (n-1)\Delta_e\Delta_w$，则有 $q^{m-1} \leqslant cM$，故竞争比至多为

$$g(q) = c\,\frac{\sqrt{2}qc + \frac{q}{cM}(1-1/q-\sqrt{2}c)}{1-1/q}$$

于是问题即为求 $g(q)$ 的最小值.

对 q 求导并令其为零，得到当 $q = 1 + \sqrt{\frac{c^2M-c}{c^2M-c+\sqrt{2}/2}}$ 时，竞争比取得最小值. 但由于该竞争比的表达式比较复杂，转而寻求竞争比比较简洁的形式. 注意到 $\sqrt{2}c > 1$，故

$$g(q) < c\,\frac{\sqrt{2}qc + \frac{q}{cM}\left(-\frac{1}{q}\right)}{1-1/q} = c\,\frac{\sqrt{2}q^2c - \frac{q}{cM}}{q-1}$$

因此得到，当 $q = 1 + \sqrt{1 - \frac{1}{\sqrt{2}Mc^2}}$ 时，竞争比为 $2\sqrt{2}c + 2\sqrt{2}\sqrt{c^2 - \frac{\sqrt{2}}{2}\frac{1}{M} - \frac{1}{cM}}$. 因此，得到下述的定理.

定理 2　欧氏平面上的占线中位选址问题存在一个常数竞争算法，其竞争比至多为 $2\sqrt{2}c + 2\sqrt{2}\sqrt{c^2 - \frac{\sqrt{2}}{2}\frac{1}{M} - \frac{1}{cM}}$，其中 $M = (n-1)\Delta_e\Delta_w$，当问题输入给定后它是常数，$c$ 为欧氏平面上的离散 k-median 问题近似算法能达到的近似比.

根据这个定理，结合文献[15]中给出的多项式时间近似方案的相应结果，有如下的推论.

推论 1　欧氏平面上的占线中位选址问题存在一个常数竞争比算法，其竞争比至多为 $2\sqrt{2}c + 2\sqrt{2}\sqrt{c^2 - \frac{\sqrt{2}}{2}\frac{1}{M} - \frac{1}{cM}}$，其中 $M = (n-1)\Delta_e\Delta_w$ 当问题输入给定后它是常数. 并且任意给定 $\varepsilon > 0$，该常数竞争比可以在多项式时间内以 $1+\varepsilon$ 倍任意接近.

特别地，由于 $M > 1$，有下面的推论.

推论 2　欧氏平面上的占线中位选址问题存在多项式时间下的竞争比至多为 $4\sqrt{2}$ 的常数竞争比算法.

4　数值算例

下面以一个简单算例来验证竞争算法的有效性. 假设某个城市给定的点集如图 1 所示，包含 A，B，C，D，O 共 5 个点，其中 A，B，C，D 的需求量都为 100，点 O 的需求量为 1，点 O 的坐标是(0，0)，点 A，B，C，D 4 者的坐标分别是(0，1)，(−1，0)，(1，0)和(0，−1). 决策者考虑选址问题的标准是加权运输距离和最小，即是以欧氏平面上的离散 k-median 问题为基本选址模型. 现在需要决策该图上的分阶段的选址服务方案.

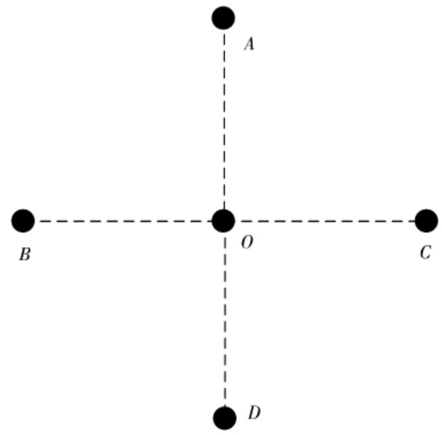

图 1　某城市待服务点示意图

Fig. 1 Illustration for a city with serving points

通过简单计算与比较，可以知道

$$opt_1 = cost(\{O\}) = 400,$$
$$\begin{aligned} opt_2 &= cost(\{A,B\}) = cost(\{A,C\}) \\ &= cost(\{A,D\}) = cost(\{B,C\}) \\ &= cost(\{B,D\}) = cost(\{C,D\}) \\ &= 200\sqrt{2} + 1, \end{aligned}$$
$$\begin{aligned} opt_3 &= cost(\{A,B,C\}) = cost(\{A,B,D\}) \\ &= cost(\{A,C,D\}) = cost(\{B,C,D\}) \\ &= 100\sqrt{2} + 1, \end{aligned}$$
$$opt_4 = cost(\{A,B,C,D\}) = 1,$$
$$opt_5 = 0.$$

在算法第 1 步中，为了简单起见，假设算法直接输出的是最优解($c = 1$)．同时根据对称性和为了说明算法的竞争比，在考虑 F_2^* 和 F_3^* 时，假设输出为最坏情形：F_2^* 不是 F_3^* 的真子集．即第 1 步输出得到解的序列集合

$$\{F_1^*, F_2^*, F_3^*, F_4^*\} = \{\{O\}, \{A,B\}, \{B,C,D\}, \{A,B,C,D\}\}$$

对应的费用序列集合为

$$W = \{400, 200\sqrt{2} + 1, 100\sqrt{2} + 1, 1\}$$

在算法第 2 步中，$T(1) = 1$．为了计算简单起见，取 $q = 2$，此时有 $T(2) = 3$，$T(3) = 4 = n - 1$．

在算法第 3 步中，有 $F_{T(3)} = F_4 = F_4^* = \{A,B,C,D\}$，$F_{T(2)} = F_3 = \{B,C,D\}$，$F_{T(1)} = F_1 = \{B\}$(这里用到对称性)．在算法第 4 步中，对 $t \in \{1,2,3,4\} \setminus E = \{2\}$，得到 $F_2 = F_{T(1)} = F_1 = \{B\}$．因此算法针对此特例问题的输出是 $\{\{B\}, \{B\}, \{B,C,D\}, \{A,B,C,D\}\}$．经验证，$cost(F_1) = cost(\{B\}) = 1.2096 opt_1$，$cost(F_2) = cost(\{B\}) \leqslant 1.7047 opt_2$，$cost(F_3) = opt_3$，$cost(F_4) = opt_4$．即竞争比为 1.704 7．事实上，在第 4 节中理论上已经保证算法在最坏情况下能够达到竞争比最多为 $4\sqrt{2} \leqslant 5.65$．

5　结束语

本文在占线理论的基础上，基于实际选址问题给定点集限制在欧氏平面上的实际背景，研究了欧氏平面上的占线中位选址问题，通过对离线问题内在的结构性质的研究，本文给出了比已有竞争性能比要好的竞争算法．本文的研究及得到的结论不仅丰富了相关的研究，而且对现有的结论具有互补作用．在未来的研究中，可以考虑如何获得更多的结构特征以改进算法．

参 考 文 献：

[1] Drezner Z, HamacherH W. Facility Location: Applications and Theory [M]. Berlin: Springer, 2002.

[2] 代　颖，马祖军，朱道立，等．震后应急物资配送的模糊动态定位——路径问题 [J]．管理科学学报，2012，15(7)：60－70.
Dai Ying, Ma Zujun, Zhu Daoli, et al. Fuzzy dynamic location-routing problem in post-earhquake delivery of relief materials [J]. Journal of Management Sciences in China, 2012, 15(7): 60－70. (in Chinese)

[3] 周爱莲，李旭宏，毛海军．一类企业物流中心动态选址模型研究 [J]．系统工程学报，2011，26(3)：360－366.
Zhou Aailian, Li Xuhong, Mao Haijun. Research on a multiple-period dynamic location model of enterprise logistics centers [J]. Journal of Systems Engineering, 2011, 26(3): 360－366. (in Chinese)

[4] 李　波，曾成培．一种逆向物流网络的多期动态选址方法 [J]．管理科学学报，2008，11(5)：76－84.
Li Bo, Zeng Chengpei. Method of multi-period dynamic location in reverse logistic network [J]. Journal of Management Sciences in China, 2008, 11(5): 76－84. (in Chinese)

[5] Mettu R R, Plaxton C G. The online median problem [J]. SIAM Journal on Computing, 2003, 32(3): 816－832.

[6] 代文强，徐寅峰，何国良．占线中心选址问题及其竞争算法分析 [J]．系统工程理论与实践，2007，27(10)：159－164.
Dai Wenqiang, Xu Yinfeng, He Guoliang. Online median problem and its competitive algorithm analysis [J]. Systems Engineering-Theory & Practice, 2007, 27(10): 159－164. (in Chinese)

[7] Chrobak M, Kenyon C, Noga J, et al. Online medians via online bidding [J]. Algorithmica, 2008, 50(2): 455－478.

[8] Lin G, Nagarajan C, Rajamaran R, et al. A general approach for incremental approximation and hierarchical clustering [J]. SIAM Journal on Computing, 2010, 39(8): 3633 – 3669.

[9] Dai W Q, Zeng X J. Incremental facility location problem and its competitive algorithms [J]. Journal of Combinational Optimization, 2010, 20 (3): 307 – 320.

[10] 代文强. 具有建设成本的占线中心选址问题及其竞争算法设计 [J]. 系统工程理论与实践, 2011, 31(12): 2342 – 2347.

Dai Wenqiang. Online median problem with constructive cost and its competitive algorithm analysis [J]. Systems Engineering-Theory & Practice, 2011, 31(12): 2342 – 2347. (in Chinese)

[11] Du D Z, Hwang F K. Computing in Euclidean Geometry [M]. Singapore: World Scientific, 1992.

[12] Weber A. Theory of the Location of Industries [M]. Chicago: University Chicago Press, 1957.

[13] Papadimitriou C H. Worst-case and probabilistic analysis of a geometric location problem [J]. SIAM Journal on Computing, 1981, 10(3): 542 – 557.

[14] Arora S, Raghavan P, Rao S. Approximation schemes for Euclidean k-medians and related problems [C] // Proceedings of the 30th Annual ACM Symposium on Theory of Computing, ACM, New York, 1999: 106 – 113.

[15] Kolliopoulos S, Rao S. A nearly linear-time approximation scheme for the euclidean k-median problem [J]. SIAM Journal Computing, 2007, 37 (3): 757 – 782.

[16] Plastria F. Continuous Location Problems [M] // Drezner Z. Facility Location, A Survey of Applications and Methods, Springer, 1995: 225 – 262.

[17] Kuhn H W. A note on fermat's problem [J]. Mathematical Programming, 1973, 4(1): 98 – 107.

Online median location problem analysis in Euclidean plane

DAI Wen-qiang, *LI Shi-ming*

School of Management and Economics, University of Electronic Science and Technology of China, Chengdu 610054, China

Abstract: Considering that most of the actual location practice is in a Euclidean plane, this paper studied the constraint mathematical model of online median location problem on the Euclidean plane under the general metric model. By analyzing the structural properties of this problem, this paper designed a polynomial-time competitive algorithm with proven good constant competitive ratio.

Key words: facility location; online median; Euclidean plane; competitive ratio