文章编号:1000-6788(2007)10-0159-06

占线中心选址问题及其竞争算法分析

代文强[1] ,徐寅峰[2,3] ,何国良[4]
(1. 电子科技大学管理学院, 成都 610054; 2. 西安交通大学管理学院, 西安 710049;
3. 机械制造系统国家重点实验室, 西安 710049; 4. 电子科技大学应用数学学院, 成都 610054)

摘要: 基于在建立的设施的个数未知的前提下需要决定如何建立初始设施集,同时要求,当新的设施集建立后,前面已经建立的设施不能被删除的实际选址约束条件下,从占线理论出发考虑了待选址个数不确定的动态选址问题.设计了一个多项式时间的竞争算法,证明了该算法具有的竞争比,该竞争比结果优于已有的结果.

关键词: 选址;占线中心;算法;竞争比
中图分类号: O221.7　　**文献标志码**: A

Online Median Problem and Its Competitive Algorithm Analysis

DAI Wen-qiang[1] , XU Yin-feng[2,3] , HE Guo-liang[4]
(1. School of Management, University of Electronic Science and Technology of China, Chengdu 610054, China; 2 School of Management, Xi'an Jiaotong University, Xi'an 710049, China; 3. State Key Lab for Manufacturing Systems Engineering, Xi'an 710049, China; 4. School of Applied Mathematics, University of Electronic Science and Technology of China, Chengdu 610054, China)

Abstract: Based on the actual constrain in locating facility, that is, the decision-maker must determine where to locate the initial facilities when the final number of facilities is uncertain, and, the constructed facilities can not be removed when the new facility is built, we studies the dynamics facility location problem with the uncertain number of facilities from the online theory view. We present a polynomial competitive algorithm, whose competitive ratio is better than the existed results.

Key words: facility location; online median; algorithm; competitive ratio

1 引言

本文中我们考虑待选址个数不确定的选址问题.在建造诸如学校、医院、工厂、物流配送中心等等设施的时候,实际的选址决策过程是一个长期的过程,选址决策者经常会遇到如下的情形:在全部需要建立的设施的个数不清楚的前提下,我们需要决定在哪里建立初始的设施(或设施集),同时还要求,当新的设施建立后,前面已经建立的设施不能被删除.这是一个待选址个数未知的选址问题.以往的选址问题模型和算法都是建立在假设待建设的设施个数是事先已知道的,或者是将该待选址个数处理成一个模型中待确定的优化变量,这些模型本质上是静态的选址模型.Current 等[1]首先提出并研究了待建立设施个数不确定的前提下的动态选址问题.在他们的论文中,他们假设了可以预先知道的待选址的个数的可能的情况,并使用了两个决策标准:最小化期望损失(minimization of expected opportunity loss)(假设存在一个确定的待选址个数的离散概率分布并且决策者可以精确地获知这个概率分布)以及最小化最大后悔(minimization of maximum regret)来决策建造多少设施以及在哪里建造这些设施.我们这里考虑的动态选址模型放松了上述假设,也即是说,本文并不假设存在对选址个数的可能状态的任何先念认识.我们将把选址过程视为一个逐渐增加的过程并增加一个附加的实际约束:在选址增加的过程中,已经建设好的设施不被删除.

收稿日期:2006-06-27
资助项目:国家自然科学基金(10371094, 70471035, 70602004);国家杰出青年基金(70525004);博士点基金项目(20050698048)

作者简介:代文强(1978-),男,电子科技大学管理学院教师,研究方向:组合最优化与算法设计,E-mail:wqdai@uestc.edu.cn.

理论上,由于最优的选址方案同选址的个数是紧密相关的,不同的选址个数将决定不同的选址点,因此我们只能考虑如何将设计的设施集费用同在同样的选址个数下最优的设施集费用尽量的靠近.这个约束实际上就是考虑如何使设计的选址集合在最坏情形下达到某种最优,可以考虑用占线算法和竞争策略[2,3]的相关理论来建立模型并进行研究.

本文考虑的动态选址模型是由 Mettu 和 Plaxton 在 2003 年所建立并研究的占线中心(online median)问题[4].他们的模型的输出是一个待选址序列,并保证序列的每一个 k 元前序的费用,都同相应的 k-median 问题的最优解费用接近.具体的算法度量是用算法竞争比来度量的,竞争比越小越好.自从该模型被提出后,受到了越来越多的管理学家、运筹学家以及计算机学家的重视,但是进一步的研究成果很少.本文将构造一个常数竞争算法,并证明该算法的竞争比,我们构造的算法的竞争比结果要优于以往的结果.

2 度量空间占线中心问题及已有的结果

度量空间占线中心选址问题(一般就直接称为占线中心问题)可以数学描述如下:

给定一个点集 U 及定义在 U 上的非负距离函数 $d: U \times U \to R^+$ 和一个正的权重函数 $w: U \to R^+$.我们一般假设距离函数 d 是一个度量,即 d 非负,对称和满足三角不定式,以及 $d(x,y)=0$ 当且仅当 $x=y$.同时定义一个点 x 到一个点集 S 的距离是 $d(x,S)=\min_{y\in S} d(x,y)$,且 $|S|$ 表示点集 S 中的点的个数.记 $n=|U|$ 表示所有的顾客数目,对任意子集 S,$S\subseteq U$,定义其费用是 $\mathrm{cost}(S)=\sum_{x\in U} d(x,S)w(x)$.度量空间 k-median 问题(一般直接称为 k-中心问题)的目标就是寻求一个元素个数是 k 的子集 $S\subseteq U$,使得它在所有元素个数是 k 的子集中最小化 $\mathrm{cost}(S)$.占线中心问题的数学描述如下:寻求子集序列 F_i, $i=1,2,\cdots,n$,使得 $|F_i|=i$, $F_1\subset F_2\subset\cdots\subset F_n=U$,且 $\mathrm{cost}(F_i)\le \alpha \times \mathrm{opt}_i$,其中 opt_i 是 i-median 问题的最优解费用,α 是常数,称为竞争比,α 越小越好.

占线中心选址问题是一个非常困难的问题,这体现在如下两点上:(1)问题的离线形式是-中心问题,这是一个是非常困难的问题,已经证明该问题是 NP 难[5,6]的.(2)我们同时需要确定 n 个待修建的设施的序列,要求其满足 $F_i\subset F_{i+1}$.

对于这个问题,2003 年 Mettu 和 Plaxton 给出了一个竞争比是 29.98 的竞争算法[4],2006 年 Chrobak 等[7]在不要求 $|F_i|=i$,只要求 $|F_i|\le i$ 的条件给出了竞争比是 24 的竞争算法,本文将在不放松 $|F_i|=i$ 的条件下给出一个比上述结果都要好的竞争算法.

占线中心问题的离线问题 k-中心问题是一个非常基础的选址模型[6],现在已经存在了很多的研究工作,其中最重要的是 2003 年 Arya 等[8]给出近似比为 3 的近似算法,这也是迄今最好的近似算法,他们的算法同时具有如下很好的性质,即对于 k-median 问题其输出的近似解的势(点的个数)恰好就是 k.

3 度量空间占线中心问题竞争算法

3.1 一些定义及性质

引理 1 对 k-median 问题,若 $1\le i\le n-1$,我们有 $\mathrm{opt}_{i+1}\le \mathrm{opt}_i$.

证明 设集合 $X\subseteq U$,$|X|=i$ 使得 i-median 问题达到最优,$\forall y\notin X$,我们有

$$\mathrm{cost}(X)=\sum_{z\in U} d(z,X)w(z),\quad \mathrm{cost}(\{y\}\cup X)=\sum_{z\in U} d(x,\{y\}\cup X)w(z)$$

但由于 $\forall z\in U$,成立 $d(z,X)\ge d(z,X\cup\{y\})$,因此,

$\mathrm{opt}_i=\mathrm{cost}(X)\ge \mathrm{cost}(X\cup\{y\})\ge \mathrm{opt}_{i+1}$,命题得证.

定义 对任意给定的选址问题,定义

$$\text{Max distance}=\max_{u,v\in U} d(u,v),\quad \text{Min distance}=\min_{u,v\in U,\,u\neq v} d(u,v)$$

且定义

$$\Delta_e=\frac{\text{Max distance}}{\text{Min distance}}$$

即 Δ_e 表示输入度量空间中最大的相对距离.

类似地,我们定义 $\text{Max weight} = \max_{x \in U} w(x)$,$\text{Min weight} = \min_{x \in U} w(x)$

且定义

$$\Delta_w = \frac{\text{Max weight}}{\text{Min weight}}$$

即 Δ_w 表示问题输入度量空间中最大相对权重.

在上述定义下,我们有下面的重要性质.

引理 2　$\text{opt}_1 \leq (n-1)\Delta_e\Delta_w \text{opt}_{n-1}$

证明　设给定的空间 $U = \{x_1, x_2, \cdots, x_n\}$,且不妨假设$\{x_1\}$使得 1-median 问题达到最优. $U \setminus \{x_n\}$使得$(n-1)$-median 问题达到最优,有

$$\text{opt}_1 = \sum_{i=2}^{n} d(x_i, x_1) w(x_i) \leq (n-1)\text{Max distance} \cdot \text{Max weight}$$

$$\text{opt}_{n-1} = d(x_n, U \setminus \{x_n\}) w(x_n) \geq \text{Min distance} \cdot \text{Min weight}$$

于是,我们得到

$$\frac{\text{opt}_1}{\text{opt}_{n-1}} \leq \frac{(n-1)\text{Max distance} \cdot \text{Max weight}}{\text{Min distance} \cdot \text{Min weight}} = (n-1)\Delta_e\Delta_w$$

因此 $\text{opt}_1 \leq (n-1)\Delta_e\Delta_w \text{opt}_{n-1}$. 命题成立.

3.2　占线中心问题竞争算法

下面我们给出设计的竞争算法,首先我们注意到由于当 $k = n$ 时, U 的子集只有一个,即 U 本身,因此我们只需要考虑 $k = 1, 2, \cdots, n-1$ 时 U 的子集(设施)的选择问题. 下面我们首先写出竞争算法,后面证明算法的有关性质并证明算法的竞争比.

占线中心问题竞争算法

输入: 度量空间(U, d) ,及在每一个点 $x \in U$ 上定义的正权重函数 $w(x) > 0$

输出: 子集序列 $F_i, i = 1, 2, \cdots, n-1$,使得$|F_i| = i$,且 $F_1 \subset F_2 \subset \cdots \subset F_n = U$

Step 1　对每一个 $k(k = 1, 2, \cdots, n-1)$,由 Arya 等[8]的局部搜索近似算法产生近似比为 $c = 3$ 的近似解,同时,该算法对于 k-median 问题其输出的近似解的势(点的个数)恰好就是 k. 设得到的解的序列集合为 $F^* = \{F_{n-1}^*, \cdots, F_2^*, F_1^*\}$,对应的费用序列集合为 $W = \{\text{cost}(F_{n-1}^*), \cdots, \text{cost}(F_2^*), \text{cost}(F_1^*)\}$.

Step 2　考察 W 是否满足 $\text{cost}(F_{k+1}^*) \leq \text{cost}(F_k^*)$. 假若不是,从 $n-1$ 到 1 依次选取使得 $\text{cost}(F_{k+1}^*) > \text{cost}(F_k^*)$ 的 k ,并令 $F_{k+1}^* = F_k^*$. 这样处理后,仍记得到的序列集合为 $W = \{\text{cost}(F_{n-1}^*), \cdots, \text{cost}(F_2^*), \text{cost}(F_1^*)\}$. 经过处理后序列将满足以非降序排列,同时仍满足条件$|F_k^*| \leq k$, $\text{cost}(F_k^*) \leq c \cdot \text{opt}_k$. (简单证明如下:如果 $\text{cost}(F_{k+1}^*) > \text{cost}(F_k^*)$,则按照处理方法,新获得的设施集 $F_{k+1}^*(\text{new})$满足$|F_{k+1}^*(\text{new})| = |F_k^*| \leq k < k+1$,且 $\text{cost}(F_{k+1}^*(\text{new})) = \text{cost}(F_k^*) < \text{cost}(F_{k+1}^*) \leq c \cdot \text{opt}_{k+1}$.)

Step 3　根据 $W = \{\text{cost}(F_{n-1}^*), \cdots, \text{cost}(F_2^*), \text{cost}(F_1^*)\}$,令(参数 q 在后面证明中确定)

$$b_1 = \text{cost}(F_{n-1}^*),\ b_2 = \begin{cases} qb_1 & qb_1 \in W \\ \text{在 } W \text{ 中不大于 } qb_1 \text{ 的最大元} & qb_1 \notin W \end{cases}, \cdots,$$

$$b_t = \begin{cases} q^{t-1} b_1 & q^{t-1} b_1 \in W \\ \text{在 } W \text{ 中不大于 } q^{t-1} b_1 \text{ 的最大元} & q^{t-1} b_1 \notin W \end{cases}, \cdots,$$

$$b_m = \begin{cases} q^{m-1} b_1 & q^{m-1} b_1 \in W \\ \text{在 } W \text{ 中不大于 } q^{m-1} b_1 \text{ 的最大元} & q^{m-1} b_1 \notin W \end{cases}$$

其中 m 使得 $b_m = \text{cost}(F_1^*)$. 记这样得到的 W 的子集合

$$B_q = \{b_1, b_2, \cdots, b_m\} = \{\text{cost}(F_{T(m)}^*), \text{cost}(F_{T(m-1)}^*), \cdots, \text{cost}(F_{T(2)}^*), \text{cost}(F_{T(1)}^*)\}$$

这里 $T(1) = 1$, $T(m) = n-1$. 同时我们记得到的集合 B_q 的下标序列为(降序排列).

$$E = \{T(m), T(m-1), \cdots, T(1)\}$$

Step 4 令 $F^{**}_{T(m)} = F^{*}_{T(m)}$,对 $k = m-1, \cdots, 1$,令 $F^{**}_{T(k)} = \Gamma(F^{*}_{T(k)}, F^{**}_{T(k+1)})$. 其中,集合函数 $\Gamma(A,B)$ 被定义为如下的集合:$\Gamma(A,B) = \{Z \subseteq B, |Z| \text{最小} | \forall a \in A, d_{aZ} = d_{aB}\}$.

Step 5 对 $t \in \{1,2,\cdots,n\} \setminus E$,取 $F^{**}_t = F^{**}_{t^-}$,其中,$t^- = \max\{i \in E | i \le t\}$.

Step 6 令 $F_n = U$,对 $t = n-1, n-2, \cdots, 1$,依次考察 F^{**}_t 是否满足 $|F^{**}_t| = t$,如果是,则令 $F_t = F^{**}_t$,否则(即 $|F^{**}_t| < t$ 成立),在 $F_{t+1} - F^{**}_t$ 中任意取 $t - |F^{**}_t|$ 个元素放入 F^{**}_t 中,并令得到的设施集为 F_t.

关于算法 Step 4 中的函数 $\Gamma(A,B)$,我们有如下的性质.

引理 3 对于两个集合 A 和 B,定义集合 B 的一个子集函数

$$\Gamma(A,B) = \{Z \subseteq B, |Z| \text{最小} | \forall a \in A, d_{aZ} = d_{aB}\}$$

则有(1) $|\Gamma(A,B)| \le |A|$;

(2) 给定 A 和 B,$|A|$,$|B|$ 有限,$\Gamma(A,B)$ 能够在多项式内计算出来;

(3) $\mathrm{cost}(\Gamma(A,B)) \le 2\mathrm{cost}(A) + \mathrm{cost}(B)$.

证明 1) 根据 $\Gamma(A,B)$ 函数的定义显然.

2) 下面是一个可行的计算方法.

计算子集合 $\Gamma(A,B)$ 的算法:

Step1 初始化,令 $Z = \phi$.

Step2 $\forall a \in A$,计算 $d(a,B)$,设 $x \in B$ 使得 $d(a,B) = d(a,x)$,令 $Z = Z \cup \{x\}$.

Step3 令 $A = A \setminus \{a\}$,如果 $A \ne \phi$,转 Step2,否则输出集合 Z.

如上的算法在 $|A|$,$|B|$ 有限的情况下,$\Gamma(A,B)$ 能够在 $O(|A||B|)$ 时间内计算出来;该算法为一个多项式时间算法.

3) $\forall x \in U$,有 $d(x,\Gamma(A,B)) \le d(x,g)$,其中,我们取 $g \in \Gamma(A,B) \subseteq B$ 且使得 $d(u,g) = d(u,\Gamma(A,B))$,$u \in A$ 使得 $d(x,u) = d(x,A)$ 成立. 根据三角不等式,我们得到 $d(x,\Gamma(A,B)) \le d(x,g) \le d(x,u) + d(u,g)$. 又由于 $g \in \Gamma(A,B) \subseteq B$ 使得 $d(u,g) = d(u,\Gamma(A,B))$,同时根据函数 $\Gamma(A,B)$ 的定义,有 $d(u,\Gamma(A,B)) = d(u,B)$,因此,如果令 $f \in B$ 使得 $d(x,f) = d(x,B)$,则有 $d(u,g) \le d(u,f)$. 于是我们得到 $d(x,\Gamma(A,B)) \le d(x,g) \le d(x,u) + d(u,f)$. 最后我们再次根据三角不等式得到 $d(x,\Gamma(A,B)) \le d(x,g) \le d(x,u) + d(u,f) \le 2d(x,u) + d(x,f)$,注意到我们这里对元素 u 和 f 的定义,对不定式两边同时乘上 x 的权重 $w(x)$ 并关于 x 求和我们可以得到结论. 命题证毕.

关于设计算法的时间复杂性分析,我们有下面的定理.

定理 1 我们设计的算法能在多项式时间内完成.

证明 根据文献[8],Arya 等[8]对 k-median 问题给出的近似算法,在最坏情况下所需要花费的时间是 $O(k^5 \log n)$,而设计的算法的 Step1 需要对每一个 k 进行调用,因此 Step1 总共需要花费的时间是 $O\left(\sum_{i=1}^{n} i^5 \log n\right) = O(n^6 \log n)$;Step2 的时间复杂度为 $O(n)$;Step3 的时间复杂度为 $O(n)$;Step4 在最坏情况下需要计算 n 次集合函数 $\Gamma(A,B)$,根据引理 3(2) 得到在最坏情况下该步的时间复杂度为 $O(n^3)$;Step5 和 Step6 的时间复杂度都是 $O(n)$,因此我们设计的算法在最坏情况下的时间复杂度为 $O(n^6 \log n)$,这是一个多项式时间算法,定理得证.

3.3 算法竞争比度量

首先我们可以得到如下的引理.

引理 4 $\forall T \in W$,若 $b_{k-1} < T \le b_k$,则 $T \ge q^{k-2} b_1$.

证明 若 $b_{k-1} = q^{k-2} b_1$,命题成立. 下面设 $b_{k-1} < q^{k-2} b_1$. 假设命题不成立,即 $T < q^{k-2} b_1$,由于 $T \in W$,并且 b_{k-1} 是在 W 中小于 $q^{k-2} b_1$ 的最大元素,于是 $b_{k-1} \ge T$,这与假设 $b_{k-1} < T \le b_k$ 矛盾,故命题成立. 证毕.

引理 5　$\forall T \in W$，设 b_k 是 B_q 中不小于 T 的最小的数，若 $\frac{\mathrm{cost}(F_1^*)}{\mathrm{cost}(F_{n-1}^*)} \le M$，则有

$$\frac{b_1 + b_2 + \cdots + b_k}{T} \le 1 + q + \frac{M-1}{M(q-1)}$$

证明　据 b_k 是 B_q 中不小于 T 的最小数，有 $b_{k-1} < T \le b_k$，根据引理 4，有 $T \ge q^{k-2} b_1$

$$\frac{b_1 + b_2 + \cdots + b_k}{T} \le \frac{1 + q + \cdots + q^{k-1}}{q^{k-2}} = 1 + q + \frac{1}{q-1} - \frac{1}{q^{k-2}(q-1)}$$

根据算法有 $b_1 = \mathrm{cost}(F_{n-1}^*)$ 及 $q^{m-1} b_1 \ge \mathrm{cost}(F_1^*) = b_m \ge q^{m-2} b_1$，从而 $q^{m-2} < M \le q^{m-1}$，我们得到

$$\frac{b_1 + b_2 + \cdots + b_k}{T} \le 1 + q + \frac{1}{q-1} - \frac{1}{q^{k-2}(q-1)} \le 1 + q + \frac{M-1}{M(q-1)}$$

引理 6　当 $q = 1 + \frac{\sqrt{M^2 - M}}{M}$ 时，$1 + q + \frac{M-1}{M(q-1)}$ 取得最小值：$2 + 2\sqrt{1 - \frac{1}{M}}$

证明　我们有

$$1 + q + \frac{M-1}{M(q-1)} = 2 + (q-1) + \frac{M-1}{M(q-1)} \ge 2 + 2\sqrt{(q-1) \times \frac{M-1}{M(q-1)}} = 2 + 2\sqrt{1 - \frac{1}{M}}$$

当 $q - 1 = \frac{M-1}{M(q-1)}$ 时，上式取得等号，即 $q = 1 + \sqrt{1 - \frac{1}{M}}$. 证毕.

引理 7　$\forall k \in \{1, 2, 3, \cdots, n-1\}$，成立不等式 $\mathrm{cost}(F_k^{**}) \le 2\sum_{l=i,T(i)=k^-}^{m} \mathrm{cost}(F_{T(l)}^*) - \mathrm{cost}(F_{T(m)}^*)$

证明　若 $k \in E$，设 $k = T(i)$，则根据引理 3(3) 和算法的 Step 3，我们得到

$$\mathrm{cost}(F_{T(i)}^{**}) = \mathrm{cost}(\Gamma(F_{T(i)}^*, F_{T(i+1)}^{**})) \le 2\mathrm{cost}(F_{T(i)}^*) + \mathrm{cost}(F_{T(i+1)}^{**})$$

利用递推得到（根据 $k = k^-$，$F_{T(m)}^{**} = F_{T(m)}^*$）：

$$\begin{aligned}\mathrm{cost}(F_k^{**}) &= \mathrm{cost}(F_{T(i)}^{**}) \le 2\mathrm{cost}(F_{T(i)}^*) + \mathrm{cost}(F_{T(i+1)}^{**}) \\ &\le 2\mathrm{cost}(F_{T(i)}^*) + 2\mathrm{cost}(F_{T(i+1)}^*) + \mathrm{cost}(F_{T(i+2)}^{**}) \\ &\le \cdots\cdots \\ &\le 2\sum_{l=i,T(i)=k}^{m-1} \mathrm{cost}(F_{T(l)}^*) + \mathrm{cost}(F_{T(m)}^{**}) \\ &= 2\sum_{l=i,T(i)=k}^{m} \mathrm{cost}(F_{T(l)}^*) - \mathrm{cost}(F_{T(m)}^*)\end{aligned}$$

若 $k \notin E$，则有 $k^- \in E$，且 $F_k^{**} = F_{k^-}^{**}$，从而：

$$\mathrm{cost}(F_k^{**}) = \mathrm{cost}(F_{k^-}^{**}) \le 2\sum_{l=i,T(i)=k^-}^{m} \mathrm{cost}(F_{T(l)}^*) - \mathrm{cost}(F_{T(m)}^*)$$

根据上面建立的引理，我们可以得到本文的主要结论如下.

定理 2　令 $\beta = \frac{1}{(n-1)\Delta_e \Delta_w}$，取 $q = 1 + \sqrt{1 - \frac{\beta}{3}}$，则经过我们的算法，所得到的集合序列 $F = \{F_n, F_{n-1}, \cdots, F_2, F_1\}$ 满足如下性质：

1) $F_1 \subseteq F_2 \subseteq \cdots \subseteq F_{n-1} \subseteq F_n$；

2) $\forall k = 1, 2, \cdots, n$，$|F_k| = k$；

3) $\forall k = 1, 2, \cdots, n$，$\mathrm{cost}(F_k) \le \alpha\,\mathrm{opt}_k$，其中 $\alpha = 12 + 12\sqrt{1 - \frac{\beta}{3}} - \beta$，$\beta = \frac{1}{(n-1)\Delta_e \Delta_w}$ 即算法的竞争比是 α.

证明　由于算法的最后一步 Step 6 只考虑 $|F_k| = k$ 的目的，同时根据引理 6 的证明思路，该步骤不会增加结果的竞争比，我们只需要证明由算法 Step1-5 得到的结果序列 $F^{**} = \{F_{n-1}^{**}, \cdots, F_2^{**}, F_1^{**}\}$ 满足如

下的三条即可.

(1) $F_1^{**} \subseteq F_2^{**} \subseteq \cdots \subseteq F_{n-1}^{**}$;

(2) $\forall k = 1, 2, \cdots, n-1$, $|F_k^{**}| \leq k$;

(3) $\forall k = 1, 2, \cdots, n-1$, $\mathrm{cost}(F_k^{**}) \leq \alpha \cdot \mathrm{opt}_k$.

首先我们分析(1)(2). 事实上,(1)是显然的,同时根据算法,若 $k \in E$,设 $k = T(i)$,则 $|F_k^{**}| = |F_{T(i)}^{**}| = |\Gamma(F_{T(i)}^{*}, F_{T(i+1)}^{**})| \leq |F_{T(i)}^{*}| \leq T(i) = k$(这里用到了引理 3),若 $k \notin E$,则据 Step4,$|F_k| = |F_{k^-}| \leq k^- \leq k$. 命题的(1)(2)成立. 下面证明(3),首先由引理 7,我们得到 $\forall k \in \{1, 2, 3, \cdots, n-1\}$,成立 $\mathrm{cost}(F_k^{**}) \leq 2 \sum_{l=i, T(i)=k^-}^{m} \mathrm{cost}(F_{T(l)}^{*} - \mathrm{cost}(F_{T(m)}^{*})$.

又根据引理 5,6 和算法,我们得到对于给定的非降费用序列 W,算法得到的子序列 B_q 满足 $\sum_{l=i, T(i)=k^-}^{m} \mathrm{cost}(F_{T(l)}^{*}) \leq \gamma \mathrm{cost}(F_k^{*})$,其中根据引理 5,$\gamma = 1 + q + \frac{M-1}{M(q-1)}$,又据引理 2,$M = \frac{\cos(F_1^{*})}{\mathrm{cost}(F_{n-1}^{*})} = \frac{c \cdot \mathrm{opt}_1}{\mathrm{opt}_{n-1}} \leq c(n-1)\Delta_e\Delta_w = \frac{c}{\beta}$,从而根据引理 6,得到当 $q = 1 + \sqrt{1 - \frac{\beta}{c}}$ 时,γ 取得最小值,且最小值为 $2 + 2\sqrt{1 - \frac{\beta}{c}}$.

因此,我们得到

$$\mathrm{cost}(F_k^{**}) \leq 2 \sum_{l=i, T(i)=k^-}^{m} \mathrm{cost}(F_{T(l)}^{*}) - \mathrm{cost}(F_{T(m)}^{*}) \leq 2\gamma \mathrm{cost}(F_k^{*}) - \beta \mathrm{opt}_k \leq (2\gamma c - \beta) \cdot \mathrm{opt}_k$$

其中,用到如下的不等式(根据引理 1 和引理 2)

$$\mathrm{cost}(F_{T(m)}^{*}) = \mathrm{cost}(F_{n-1}^{*}) \geq \mathrm{opt}_{n-1} \geq \beta \mathrm{opt}_1 \geq \beta \mathrm{opt}_k$$

因此,我们得到算法的竞争比是 $2\gamma c - \beta = 12 + 12\sqrt{1 - \frac{\beta}{3}} - \beta$,命题成立.

参考文献:

[1] Current J, Ratick S, Revelle C. Dynamic facility location when the total number of facilities is uncertain: a decision analysis approach[J]. European Journal of Operation Research, 1998, 110(3), 597 - 609.

[2] Borodin A, El-Yaniv R. Online Computation and Competitive Analysis[M]. Cambridge University Press, Cambridge, UK, 1998.

[3] Fiat A, Woeginer GJ. Online Algorithms: The State of the Art[M], Springer, 1998.

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

[5] Garey M R, Johnson D S. Computers and Intractability: A Guide to the Theory of NP-Completeness[M]. Freeman, San Francisco, CA, 1979.

[6] Kariv O, Hakimi S L. An algorithm approach to network location problems, Part Ⅱ: The p-medians[J]. SIAM J Appl Math, 1979, 37: 539 - 560.

[7] Chrobak M, Kenyon C, Noga J, Young N E. Oblivious medians via online bidding[C]//Proceedings of Latin American Theoretical Informatics (LATIN06), LNCS 3887, 2006.

[8] Arya V, Garg N, Khandekar R, et al. Local search heuristics for k-median and facility location problems[J]. SIAM Journal on Computing, 2003, 33(3): 544 - 562.