文章编号: 1005-2542(2003)02-0177-05

加拿大旅行者问题

朱志军[2], 徐寅峰
(西安交通大学 管理学院, 西安 710049)

【摘要】针对加拿大旅行者问题, 分析其主要变形——确定型可恢复的加拿大旅行者问题。考虑堵塞边动态产生, 一个遇到且堵塞边在时间 $l(x,x)$ 后可以自动恢复情况下的道路选择。通常对于在线算法可以从两个方面进行评价: 最坏情形分析和竞争比分析。本文先设计了求解最坏情形下旅行时间最短的标号算法并分析了其计算复杂性。而后在竞争比分析中, 设计了基于贪婪原则的选路策略, 并对其进行了竞争比分析, 证明了该贪婪策略对于确定型可恢复加拿大旅行者问题的竞争比为 $(k+2)/2$。

关键词: 加拿大旅行者问题; 确定型可恢复加拿大旅行者问题; 竞争算法; 竞争比
中图分类号: TB 114.1 文献标识码: A

The Canadian Traveler Problem

ZHU Zhi-jun, *XU Yin-feng*
(School of Management, Xi'an Jiaotong University, Xi'an 710049, China)

【Abstract】The Canadian Traveler Problem and its variations are considered. Devising a travel strategy under the condition that each site is associated with a recovery time $l(x,x)$ to reopen any blocked road that is adjacent to it is the main problem of this paper. The quality of online algorithm is usually measured based on two alternate criteria: competitive ratio and worst-case performance. In the worst-case performance analysis, the paper devised an algorithm to get the optimal worst-case travel time and the algorithm complexity is considered. Greedy Strategy is put forward in the competitive analysis and $\frac{k+2}{2}$ is proved to be the competitive ratio.

Key words: the Canadian traveler problem; the deterministic recoverable Canadian traveler problem; competitive algorithm; competitive Ratio

加拿大旅行者问题[1](Canadian Traveler Problem, CTP)是指一个旅行商希望从 s 点走到 t 点, 事先他知道地形图(地图) $G(V,E)$。其中, V 代表实际中的地点集合, E 是对应于现实各地点之间路径边的集合。同时, 对于每一个 $e\in E$ 都有一个权重 $l(e)$, 可以理解为通过路径 e 所花费的时间。但是, 往往这个地图是不可靠的。因为有时某些路径因为天气的原因(如雪灾等)不能通过, 并且仅当到达了相邻点时才能知道下一条路径是否堵塞。现在的问题就是在这种不确定的情况下, 如何设计出一个行走方案能够使得旅行者在最短的时间内从 s 点走到 t 点。

注意到对于该问题的局外问题, 即所有的堵塞信息如提前知道的话, 这个行走方案就很容易确定, 只需求出去掉这些堵塞边(假定堵塞边不能恢复)后从 s 点走到 t 点的最短路径, 这就是行走方案。然而上述问题的困难就在于它的在线本质, 也就是说必须在只知道局部信息而对未来一无所知的情况下对全局做出决策。未来堵塞边的序列可能对以往甚至以后的决策产生深刻的影响。堵塞边出现不同, 最优行走方案也会随之发生变化。

收稿日期: 2001-12-18
基金项目: 国家自然科学基金资助项目(7980004)
作者简介: 朱志军(1977-), 男, 博士生。主要从事组合优化在线金融问题及竞争比研究。

加拿大旅行者问题首先由 Papadimitriou 等人[1] 提出, Amotz Bar-Noy 等人[2] 分析研究了边的堵塞服从一定概率分布情况下随机型可恢复加拿大旅行者问题, 给出了算法复杂度为 $O(m \log n)$ 的求解算法, 同时还研究了 k-加拿大旅行者问题, 考虑堵塞边的个数为参数情况下的加拿大旅行者问题。最后提出了确定型可恢复加拿大旅行者问题, 并给出了一些定义和性质。本文主要从最坏情况和竞争比分析出发考虑确定型可恢复的加拿大旅行者问题。在该问题中, 假定堵塞的边在一定时间后能够恢复, 重新可以通过, 恢复时间为 $l(x, x)$, 且对于 E 中的任意边都有 $l(x, x) < l(e)$。这样旅行者在顶点 x, 他可以选择走另外一条没有堵塞的路或者是在原地等待 $l(x, x)$ 时间通过边 e(只要下次堵塞的边不是 e)。需要指出的是, 如果边的恢复时间看作是无限的话, 那么可恢复的加拿大旅行者问题就变成了加拿大旅行者问题, 因此可以说加拿大旅行者问题是可恢复的加拿大旅行者问题的一个特例。

1 数学模型和基本定义

1.1 数学模型

假设交通网络是一个网络图记为 G, 各个地点(路口)对应于图 G 上的各个顶点, $G(V, E)$, 其中, V 代表实际中的地点集合, E 是对应于现实各地点之间路径边的集合。同时, 对于每一个 $e \in E$ 都有一个权重 $l(e)$, 可以理解为通过路径 e 所花费的时间, $|G| = n$。其中: S 为起始点; D 为目的点; SP 为在没有路口堵塞情况下的最短路径; $W(SP)$ 为沿着最短路径所要花费的通过时间。$l(x, x)$ 表示堵塞边恢复的时间, 且对于任意的边 e 有 $l(x, x) < l(e)$。如果 $\delta = (e_1, e_2, \cdots, e_k)$ 表示一个堵塞边序列, 用 $C_A(\delta)$ and $C_{\mathrm{opt}}(\delta)$ 分别表示在事先不知道堵塞边序列的情况下用算法 A 求出的最优解和事先知道堵塞边序列情况下相应问题的最优解, 即所要花费的时间。现在的问题就是如何找到在线算法中 A [3,4], 使得

$$C_A(\delta) = \min_{B \in \mathcal{B}} C_B(\delta), \tag{1}$$

其中, $\mathcal{B}$ 表示所有在线算法集合[5~7]。

1.2 基本定义

用 SP 表示从 s 到 t 没有堵塞边的情况下的最短路径, 如果旅行者到达 x 点后, 发现相邻的一条边 (x, y) 发生了堵塞, 用 SP_1 表示在此情况下从 x 点到 t 点的最优路径。进一步可以令 $SP_{1,1}$ 表示沿着 SP_1 行走在碰到下一个堵塞边之前所经历的路径, $SP_{1,2}$ 表示相应 SP_1 中还没有经历的路径。SP_k、$SP_{k,1}$、$SP_{k,2}$ 的含义如上。

对于在线算法, 以前通常用分析普通优化算法的标准来评价: 最坏情形分析[8] 和平均情况分析。在最坏情形分析中考虑算法在每种特例中的最坏情形; 在平均情况分析中, 假定问题的输入服从一定的分布, 而后考虑算法计算结果的期望值, 在此用期望值的不同来分析各种算法的优劣。事实上, 在线问题的输入往往随着时间和具体实例的不同而变化。假定输入服从一定的概率分布往往很难真正的反映客观实际。正是基于以上的一些观察, Sleator 等人[5] 首次提出了竞争比分析这个思想。在竞争比分析[9] 中考虑的是在线算法和对应最优局外(off-line)算法结果的比值。这两种方法的区别可以用下面简单的例子来说明。假设一个旅行者需要从 Victoria 到 Quebec, 考虑两种情形, 一是离线问题的最优解需要旅行 10 天, 另外一个是离线问题的最优解需要旅行 2 天。在竞争比标准下, 如果在线算法能够保证竞争比为 2, 则对于第 2 种情形在线算法能够保证旅行者在 4 天内到达, 在第 1 种情形下 20 天内可以到达。而对于最坏情形分析标准, 如果另外一种在线算法能保证在任何情形下都能在 15 天内到达, 则旅行者在两种情形下都能在 15 天内到达。显然对于情况 1, 最坏情形分析的结果会好一些, 但对于情况 2, 相对竞争比分析, 结果就差了很远。因此不同的情况, 不同的评价标准会有不同的用处。本文对可恢复加拿大旅行者问题最坏情形和竞争比都作了分析, 并设计了相应的算法。

具体说, 对于任意一个最多有 k 条堵塞边序列 $\delta = (e_1, e_2, \cdots, e_k)$, e_i 表示第 i 条边发生堵塞, 如果总存在与堵塞边序列无关的常数 α 和 β 满足:

$$C_A(\delta) \leq \alpha C_{\mathrm{opt}}(\delta) + \beta$$

称算法 A 为此在线问题的竞争算法, 且其竞争比为 α, 其中[9] $\alpha = \inf_A C_A^{\mathrm{off}} = \inf\left\{\dfrac{C_A(\delta)}{C_{\mathrm{opt}}(\delta)}\right\}$。

如果用 $E(x)$ 表示从 x 点到 t 点的旅行时间。不失一般性, 可以把旅行者在点 x 做出的决策看成几个阶段组成, 在每一个阶段旅行者都会参考点 x 相邻点的优先次序 $(x_1, x_2, \cdots, x_i)$。在每次碰到堵塞边, 旅行者都会找出没有堵塞且下标 i 最小的 x_i, 然后沿着这条边 (x, x_i) 行走。需要指出的是, 优先次序是根据 $l(x, x_i) + E(x_i)$ 的值进行升序排列形成的[2]。

2 确定型可恢复加拿大旅行者问题

确定型可恢复加拿大旅行者问题是相对随机型

的加拿大旅行者问题而言, 在随机型的加拿大旅行者问题中, 给出了每条边被堵塞的概率 p_i, 且这些堵塞概率相互独立。问题就是求在考虑 p_i 的情况下如何求出期望的旅行时间最短。而对于确定型可恢复加拿大旅行者问题, 不会存在一条边被堵塞的概率 p_i, 因为一条边要么被堵塞, 要么就可以通过。以下仅仅考虑确定型可恢复加拿大旅行者问题。

2 1 最坏情形分析

对于交通网络是一个网络图记为 $G(V, E)$, 每一个 $e \in E$ 都有一个非负权重 $l(e)$, 且对于任意的边 e 有 $l(x, x) < l(e)$。这里所给出的一个行走方案是使得在最坏情形下的旅行时间最短。

对于确定的目的点 t, 如果 $k=0$, 即没有堵塞边发生, 则当旅行者到达点 x, 最优的行走方案就是沿着从 x 到 t 的最短路径行走。

如果 $k>0$, 则可以用递归的方法来算出此时的最优方案, 也就是说在碰到第 k 个堵塞边考虑选路策略时, 对于 $0< i< k$ 的情况都已经计算过。用 $\mathrm{dist}(i, x)$, $0< i< k$ 表示旅行者从 x 点到 t 点考虑最多有 i 条边堵塞的情况下最坏情形下的旅行时间。假设旅行者到达 t 点过程中经过点 x, 可以区分以下两种情况:

(1) 在到达 x 点时, 已经有 $h(1\leq h\leq k)$ 条边发生了堵塞。因为已假设了边恢复时间 $l(x, x)$ 小于通过任何一条边的时间 $l(e)$, 所以当旅行者到达 x 点时, 所有的堵塞边都已恢复可以重新通过。因此, 从 x 点到 t 点, 还将有 $k-h$ 条边会发生堵塞, 于是可以运用 $\mathrm{dist}((k-h), x)$ 的策略。

(2) 如果在到达 x 点时, 还没有边发生堵塞。在此情况下, 旅行者首先检查 x 点相邻边有无堵塞, 如果没有他可以沿着 $(x, x^k_{0,0})$ 边行走, $(x, x^k_{0,0})$ 为 x 点的首选边。

当面对第 $i(l\leq i\leq k)$ 条边发生堵塞时, 旅行者有两种选择: ①在原地等待 $l(x, x)$ 时间, 使得首选边由堵塞变为可以通过, 然后可以运用 $\mathrm{disk}((k-i), x)$ 选路策略; ②旅行者可以选择一条没有堵塞的路径行走。对于这种情况用 $(x, x^k_{i,j})$ 表示可以通过的可选边集合, 其中 $j\geq 2$, 如果可选择边集合少于两条, 则 $(x, x_i) = (x, x)$ 就是该集合中的边。当碰到第 $i(1\leq i\leq k)$ 条堵塞边时, 旅行者可以选择一条 j 的下标最小的不堵塞的边 $(x, x^k_{i,j})$, 如果这条边就是 (x, x), 则旅行者在原地等待, 否则就会沿着选择的未堵塞边行走。

下面解释如何来计算对任意 $i(1\leq i\leq k)$ 可选边集合 $\{(x, x^k_{i,0}), \cdots, (x, x^k_{i,j})\}$。如果知道了这些可选边集合, 就可找出点 x 此时的首选边。假设 (x, y) 是第一条可选边, 如果顺着这条边行走, 则从 x 到 t 最坏情形下的旅行时间为 $l(x, y) + \mathrm{dist}((k-i), y)$, 因此可选边中的第 1 条边 $(x, x^k_{i,0})$ 满足:

$$\min\{l(x, y) + \mathrm{dist}((k-i), y)\}$$

其中, $y\in\{x\}\cup Adj(x)$, $Adj(x)$ 表示与 x 点相邻点的集合。如果 $(x, x^k_{i,0}) = (x, x)$, 则旅行者就会在原地等待。假设对于任意的都有 $(x, x^k_{i,a}) \neq (x, x)$, 则第 j 条可选边是由下式决定:

$$\min\{l(x, y) + \mathrm{dist}((k-i), y)\}$$

其中, $y\in\{x\}\cup Adj(x) - \bigcup_{a=0}^{j-1}\{x^k_{i,a}\}$。需要指出的是上面对于 i 的可选边集合在递归中都已经计算过, 为了说明这个问题, 给出如下引理。

引理 1 对于 $1\leq j\leq i-1$, 边 $(x, x^k_{i,j})$ 和边 $(x, x^{k-1}_{i-1,j})$ 是一条边。

根据整个递归的思想, 引理 1 显然成立。

以下主要介绍如何来计算首选边。最坏情形: ①首选边不堵塞, 让旅行者沿着首选边行走; ②堵塞其中的一些关键中间点 x, 让旅行者不得不沿着 x 的某条替代可选边行走或留在原地。通过上面的分析, 不难知道从 x 到 t 最坏情形下的旅行时间 $\mathrm{dist}(k, x)$ 满足下面等式:

$$\mathrm{dist}(k, x) = \max_{i=0}^{k}\{l(x, x^k_{i,i}) + \mathrm{dist}((k-i), x^k_{i,i})\}$$

即对于 i 条堵塞边的最坏情况就是对于 $1\leq j\leq i-1$, 所有的边 $(x, x^k_{i,j})$ 都被堵塞。而首选边就是使得上式最小的边, 可表示为

$$\min \mathrm{dist}(k, x) = \min\max_{i=0}^{k}\{l(x, x^k_{i,i}) + \mathrm{dist}((k-i), x^k_{i,i})\}$$

不难发现所有的首选边组成了一个类似于最短树的树状图。如果对于每个 $x\in V$, $\mathrm{primary}(x)$ 表示 x 的首选边, 则 $\mathrm{primary}(x)$ 和 $\mathrm{dist}(k, x)$ 可以用一个类似于 Dijkstra's [10] 算法的方法得出。在该算法中, L 表示已标记点的集合; 最初 $L = \varnothing$。对于任意 $x\in L$, $\mathrm{primary}(x)$ 和 $\mathrm{dist}(k, x)$ 都已算出。对于 $x\in V-L$, $\mathrm{tentative\text{-}primary}(x)$ 和 $\mathrm{tentative\text{-}dist}(k, x)$ 分别表示 $\mathrm{primary}(x)$ 和 $\mathrm{dist}(k, x)$ 变量的中间值。最初 $\mathrm{tentative\text{-}disk}(k, t)=0$ 和任意 $x\in V-L$ 有 $\mathrm{tentative\text{-}dist}(k, x)$ 为无穷大。每次循环包括两个部分: 标记阶段和更新阶段。

标记阶段 从所有的 $x\in V-L$ 中找出 $\mathrm{tentative\text{-}dist}(k, x)$ 最小的 x, 并将它加入到集合 L 中且将 $\mathrm{dist}(k, x)$ 的值改为 $\mathrm{tentative\text{-}dist}(k, x)$。

更新阶段　对每一个 $y \in Adj(x) - L$，检查如果让边 (y, x) 成为首选边，y 的 tentative-dist (k, y) 是否会有所改进；如果是，则修改 y 的中间值。需要指出的是，对于每个 y，只是检查它由首选边 (y, x) 导致的最坏情形下的旅行时间。即

$$\max\{l(y,x)+\mathrm{dist}(k,x), \max_{i=1}^{k}\{l(y,y_{i,i}^{k})+\mathrm{dist}((k-i),y_{i,i}^{k})\}\}$$

其中，$(y, y_{1,1}^{k}), \cdots, (y, y_{i,i}^{k})$ 表示 y 的可选边集合。

当所有的点都标记完了，也就是 $L = V$ 时，整个算法结束。通过对算法的分析有以下引理成立（证明略）。

引理 2　如果对于每个 $x \in V$，primary (x) 表示 x 的首选边，则算法所得出的 dist (k, x) 就表示了从 x 点到 y 点考虑最多会有 k 条堵塞边最坏情形下的旅行时间。

引理 3　对于每个 $x \in V$，不管首选边如何，dist (k, x) 是所有且最多会有 k 条堵塞边最坏情形下旅行时间的最小值。

计算复杂性分析　如果最多会有 k 条堵塞边，对于一个指定的目的点 t，在所有 $1 \le i \le k-1$ 的输出已知的情况下，求 $i = k$ 时的输出，可以在 $O(km + n\log n)$ 时间内完成。因此，整个时间复杂度为 $O(k^2 m + kn\log n)$。

2.2　竞争比分析

行走策略：如果考虑在旅行者每次遇到堵塞边时，他都会找到一条去掉堵塞边情况下的最优路径，然后和原地等待堵塞边恢复后沿着原先既定路线行走相比较，总是保证沿着两种情况下的最优路径行走。本文称这种行走策略为一种贪婪策略，因为它每次保证沿着在已知情况下的最优路径行走。

令 $C_B(\delta)$ 表示这种贪婪策略最后总的行走时间。为了分析简单，假设每次的最优路径都不是原地等待堵塞边恢复而是重新选择一条新路径，根据第 1 节的定义，在最坏情况下，也就是堵塞边恰好都出现在每次新选择的最优路径上，可以有以下等式成立：

$$C_B(\delta) = W(SP_{1,1}) + W(SP_{1,2}) + \cdots + WP(SP_{1,k}) + W(SP_{k+1})$$

因为在面对第 i 条堵塞边时，应选择一条新的路径而不是等待第 $i-1$ 条堵塞边恢复，因此应有如下不等式成立：

$$W(SP_{i+1}) \le W(SP_{i,2}) + l(x,x)$$

这样对于所有的 k 条边，应有：

$$W(SP_1) \le W(SP_{0,2}) + l(x,x) \tag{2}$$

$$W(SP_2) \le W(SP_{1,2}) + l(x,x) \tag{3}$$

$$W(SP_3) \le W(SP_{2,2}) + l(x,x) \tag{4}$$

$$\vdots$$

$$W(SP_{k+1}) \le W(SP_{k,2}) + l(x,x) \tag{5}$$

其中：$SP_{0,1}$ 表示在第 1 次遇到堵塞边时，已经沿着最短路径走过的路径，而 $SP_{0,2}$ 表示剩下的路径。

将式(1)～(5)相加，可得

$$W(SP_1) + W(SP_2) + W(SP_3) + \cdots + W(SP_{k+1}) \le W(SP_{0,2}) + W(SP_{1,2}) + W(SP_{2,2}) + \cdots + W(SP_{k,2}) + (k+1) * l(x,x)$$

事实上，就算最优的选路策略是在原地等待堵塞边恢复，重复 k 次选择，上式依然成立。因为对于任意 i 有

$$W(SP_i) = W(SP_{i,1}) + W(SP_{i,2})$$

因此上式可以表示为

$$W(SP_{1,1}) + W(SP_{1,2}) + \cdots + WP(SP_{1,k}) + W(SP_{k+1}) \le W(SP_{0,2}) + (k+1) * l(x,x)$$

即有　$C_B(\delta) \le W(SP_{0,2} + (k+1) * l(x,x)$

因为，$C_{opt}(\delta) \ge W(SP) = W(SP_{0,1} + W(SP_{0,2})$，$W(SP_{0,1} \ge l(x,x)$，所以，$C_{opt}(\delta) \ge W(SP_{0,2}) + l(x,x)$。

根据竞争比的定义，有

$$\alpha = \inf_A c_A^{\mathrm{off}} = \inf\left\{\frac{C_A(\delta)}{C_{opt}(\delta)}\right\} = \inf\left\{\frac{C_B(\delta)}{C_{opt}(\delta)}\right\} = \frac{k+2}{2}$$

即对于任意的堵塞序列 δ，总存在和堵塞边位置无关的常数 α 和 β 满足：

$$C_A(\delta) \le \alpha C_{opt}(\delta) + \beta$$

其中，$\alpha = (k+2)/2$，$\beta = 0$。

通过以上分析不难得出以下引理成立：

引理 4　贪婪策略对于确定型可恢复加拿大旅行者问题是竞争比为 $(k+2)/2$ 的确定性竞争算法。

3　结　语

加拿大旅行者问题可以看成是很多网络选线问题的一个简单抽象。在很多通信网络中，信息通常是根据局部网络结构顺着最短路传播的[11,12]。然而网络拓扑结构会随着时间而变化。如果在传输过程中，网络拓扑结构没有发生变化，当然这种传输策略是最优的，然而就算只有一条边发生错误不能通过，结果可能都会很差。可恢复加拿大旅行者问题是解决此问题的一个很好思路。

在本文中，首先提出了一个加拿大旅行者问题的一个变形，即可恢复加拿大旅行者问题。而后分别

从最坏情形和竞争比分析两个方面, 对该问题进行了深一步的分析。找到了计算最坏情形下最短旅行时间的算法和给出了解决此在线问题的一个贪婪策略, 并分析证明了贪婪策略是竞争比为 $(k+2)/2$ 的确定性竞争算法。本文还将有许多有待进一步深入研究的理论问题, 如没有 $l(x, x) < l(e)$ 限制, 每次事先知道后 l 步信息等情形, 这些都是今后进一步研究的方向和问题。

参考文献:

[1] Papadimitriou C H, Yannakakis M. Shortest paths without a map [C]. In Proc 16th ICALP, Lect Notes in Comp Sci, 1989, 372: 610-620.

[2] Amotz Bar-Noy, Schieber Baruch. The Canadian traveller problem [C]. Proc of The Second Annual ACM-SIAM Symposium on Discrete Algorithms, 1991. 261-270.

[3] Xu Y, Wang K. The on-line taxi problem and its competitive algorithm [J]. Journal of Xi'an Jiaotong University, 31 (1): 56-61.

[4] Koutsoupians E, Papadimitriou C. On the k-server conjecture [C]. Proc of STOC Montreal, 1994. 507-511.

[5] Sleator D D, Taijan R E. Amortized efficiency of list update and paging rules [J]. Communications of the ACM, 28: 202-208.

[6] David S B, Borodin A, Karp R, *et al*. On the power of randomization in online algorithms [C]. In Proc 22nd Symposium on Theory of Algorithms, 1990. 379-386.

[7] Alon N, Karp R, Peleg D, *et al*. A graph-theoretic game and its application to the k-server problem [J]. In DIMACS Series in Discrete Mathematics and Theoretical Computer Science, 1992, 7: 1-9.

[8] David S B, Borodin A. A new measure for the study of online algorithms [C]. Manuscript, 1990.

[9] Hochbaum D S. Approximation algorithms for NP-hard problems [M]. PWS Publishing Company, 1997.

[10] Dijkstra E. A note on two problems in connexion with graphs [J]. Numer Math, 1959, 1: 269-271.

[11] McQuillan J M. Richer I, Rosen E C. The new routing algorithm for the ARPANET [J]. IEEE Transactions on Communications, May 1980, COM-28: 711-719.

[12] Rosen E C. The updating protocol of AEPANET's new routing algorithm [J]. Computer Networks, February 1980, 4: 11-19.

下期发表文章摘要

二级供应链模型及其在服务销售问题中的应用

黄小原, 卢　震

(东北大学 工商管理学院, 沈阳 110004)

摘　要: 本文研究仅具有制造商和销售商的二级供应链模型, 这类模型广泛适用于服务销售系统。在文献[1-5]基础上建立了不但具有选址及市场顾客配置作用, 还具有对于市场销路调控的供应链混合整数规则模型, 设计了适用于这种混合整数规划供应链管理决策的遗传算法。最后对于服务销售中的实例——奶品零售问题进行了供应链仿真。