

文章编号: 1000-5471(2008)06-0085-05

LC-检索在城市公交线路查询中的应用研究^①

赵晶璐, 何国斌

西南大学 计算机与信息科学学院, 重庆 400715

摘要: 分析了城市公共交通网络的特点, 利用数据库存储公交线路信息, 给出了相应的数据表结构, 并结合最小成本检索-LC 检索, 提出了一种以交叉路口为结点的改进最短路径算法, 并引入备忘录存储求得的最短路径信息, 以提高查询的效率及灵活性.

关键词: 公交网络; 最短路径; LC-检索; 备忘录

中图分类号: TP301.6

文献标识码: A

公共交通是城市交通网络中的重要组成部分. 随着现代化城市化的发展, 公交线路也在不断增加和更新, 公交线路查询难的问题摆在了市民面前. 为此, 有必要建立一个方便、快捷的公交线路查询系统. 公交查询系统的核心可以归结为求某种形式的最短路径问题.

现有的公交查询系统大多数是采用 Dijkstra 算法或基于该算法的改进算法. 本文根据公共交通网络的特点, 通过对分支限界算法的研究和分析, 提出一种以交叉路口为结点的 LC 检索策略, 并引入备忘录, 记录已找到的线路信息, 以提高查询的效率和灵活性.

1 城市公交线路信息描述

1.1 城市公交网络的分析

城市道路是由许许多多的街道(路段)构成, 多条街道交汇于交叉路口. 以交叉路口为结点, 以各路段为弧段便构成了城市道路网络. 城市公交线路是道路网络中营运的公交行车路线. 每条交通线路途经多个交叉路口和多条路段, 而每个交叉路口或路段可能在多条交通线路上, 这种由众多的交通线路交织而成的网络便构成城市公交网络. 城市道路网的特点决定了公交网络具有连通性、节点性、相异性、网络拓扑复杂性等显著特点.

(1) 连通性: 通过城市公交网络, 乘坐一条或多条公交线路的车, 可以从一个站点到另一个站点, 即站点间具有连通性.

(2) 节点性: 公交线路是若干个站点的有序序列, 表现出图的节点特性.

(3) 相异性: 每条线路都有固定的行车线路和发车频率, 且行驶途中依站点停靠, 因此线路不同的车辆尽管途经的站点可能重叠, 但不可能完全相同. 另外, 由于城市的部分路段采取单行线路或交通管制等原因, 即使是同一路公交车, 其往返的路线也不一定完全一样. 公交线路具有相异性.

(4) 网络拓扑复杂性: 交通网络又是由交通线路交织而成. 以街道交叉路口为结点, 以主要街道为弧, 更可抽象出交通网络的网络拓扑结构. 这样, 在各个交叉路口之间又分布着数量、距离不等的站点. 公交

^① 收稿日期: 2007-10-20

作者简介: 赵晶璐(1983-), 女, 甘肃兰州人, 硕士研究生, 主要从事人工智能及计算机应用.

站点有些位于交叉路口附近,而有些则在某路段内.

1.2 公交线路信息的数据表存储

要实现交通线路的自动查询,应将公交线路的相关信息存储到计算机中. 交通线路信息最基本的是行车线路信息.

利用数据库技术,我们将公交网络的有关信息组织到数据库中,以便数据共享、使用和维护. 通过对公交线路信息的分析,可以采用如下数据表来存储公交网络的相关信息.

表 1 线路信息表(L-Table)

线路编号	首班时间	末班时间	发车间隔	票价	行车站数	站点编号 1	...	站点编号 K1
L0894	6:30	10:30	15	1.00	13	S0231		S0253

表 2 站点信息表(S-Table)

站点编号	站点名称	路口标识	路段名	线路总数	线路编号 1	...	线路编号 N
S0231	天生站	1	人民路	5	L0893		L0894

表 3 站点间的距离(权值)表(W-Table)

站点编号	站点编号	站点间距离(权值)/Km
S0231	S0232	1.5

值得注意的是站点权值可以是两车站间的距离、行车所需时间、交通费用等.

2 LC 检索的基本思想

分支限界算法是一种在问题的解空间树上搜索问题解的方法. 它采用广度优先或最小耗费优先的算法策略搜索解空间树,在搜索过程中,每一个活结点只有一次机会成为扩展结点.

LC 检索又称为 LC 分支限界检索,即最小成本检索,它采用一个“有智力”的排序函数 $C(\cdot)$ 来选择下一个 E-结点,以便加快到答案结构的速度. 检索策略总是选取 $C(\cdot)$ 值最小的活结点作为下一个 E-结点. $C(X)$ 的定义如下:

$$C(x) = \begin{cases} \text{根结点到 } x \text{ 的成本 } s & x \text{ 为答案结点} \\ \infty & \text{否则} \end{cases}$$

由于要得到精确的成本函数 $C(\cdot)$ 一般不现实,因此,在实际应用中,常使用估计结点的成本函数:

$$\hat{C}(x) = f(h(x)) + \hat{g}(x)$$

其中 $f(\cdot)$ 是一个非降函数,以减少算法偏向纵深检索, $h(x)$ 为根结点到 x 的估计成本, $\hat{g}(x)$ 为 x 到答案结点的估计成本函数.

3 LC 检索在交通查询系统的中的应用

3.1 交叉路口结点间的最短路径的 LC-检索

城市中公交站点的数量成千上万,若直接基于每一个站点运用 LC 检索求最短路径,由于数据量大,时间复杂度也相当大的,搜索效率低. 但一个城市中交叉路口数量却远远小于站点数量,在一定时期内,其数量和位置相对固定,又是换乘车机率较大的地方,因此公交查询可视为以交叉路口作为顶点的最短路径问题.

应用 LC 检索寻找交叉路口结点间的最短路径,引入“有智力的”排序函数,即成本估计函数:

$$C(x) = W(x)P_1 + H(x)P_2 + F(x)P_3$$

其中: $C(x)$ 为成本函数, $W(x)$ 为权值成本, $H(x)$ 为根到 x 的级数成本, $F(x)$ 为乘车换线路成本, P_1, P_2, P_3 为各种成本的概率,则有 $\sum p_i = 1$.

若只考虑路径最短,则 $H(x)$ 和 $F(x)$ 取 0, $W(x)$ 为起点到结点 x 的路径长度和,且 $P_1 = 1$.

算法如下:

Procedure LC_Shortpath(T, C)

//Heap 为活结点表, 线路相关信息存保存如上在相关数据表中

E ← T //E—结点

活结点表 Heap 初始化为空

Loop

If E 是答案结点 Then {输出从 X 到 T 的路径; Return; }

For E 的每个儿子 X Do

Call Add(X) // 根据 C(X) 决定是否入活结点表以及加入位置

Parent(X) ← E // 指示到根的路径

Repeat

If Empty(Heap) Then {输入“无答案”; Stop}

Call Least(E) // 对 Heap 按 C 最小选取 E 结点

Repeat

End LC_Shortpath

若某地的公交线路图抽象成如图 1 所示的网络图, 并查询起始点 A 到终点 H 的最短路径. 问题的解空间可以组织成二叉树的形式, 如图 2 所示, 为公交网络图的部分解空间树. 其中每一个结点旁的数字表示该结点所对应的当前路长, 可根据查询需要, 将边权标为速度、停站时间、费用等. 从树根到叶的任一路径表示解空间的一个元素, 即从待查询的起点到终点的行车路线.

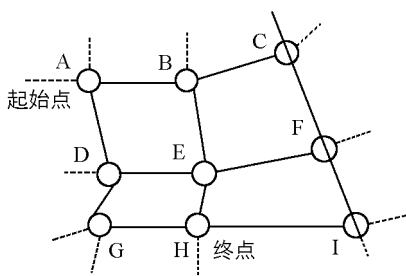


图 1 公交网络图

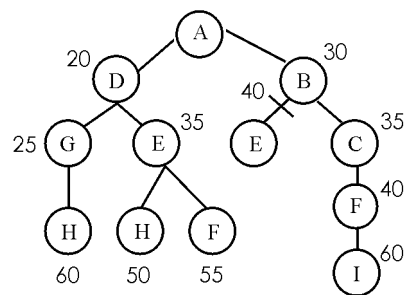


图 2 问题的解空间树

LC 检索解决此问题时, 算法从根结点 A 开始, 初始时活结点队列为空, 结点 A 是当前扩展结点. 结点 A 的 2 个儿子结点 D 和 B 均为可行结点, 将这两个儿子结点成本函数 $C(x)$ 加入最小堆, 并舍弃当前扩展结点 A. 依 LC 检索策略, 先取堆顶结点 D 为下一个扩展结点. 则 G, E 进入最小堆. 再选 G 为扩展结点, 则得到 A 到 H 的一条路线, 长为 60, 这并不是最短路径. 下一个扩展结点为 B, 生成 B 的两个儿子结点 E 和 C, 当前的 E 结点的成本为 40, 而 D 的儿子结点为 35, 利用结点间的控制关系进行剪枝, 放丢当前 E, C 进及最小堆. 如此进行下去, 当扩展到 E 的儿子结点 H, 则找到 A 到 H 的最短路径为 A-D-E-H.

3.2 虚拟结点的 LC 检索算法策略

在实际交通查询中, 检索的起点和终点往往不是交叉路口. 为了提高 LC 算法的运算速度以及算法的通用性, 可把此站点视为交叉路口, 即虚拟结点.

这样有两种求解最短路径算法的策略: 统一处理和分别处理.

(1) 统一处理策略

为了统一应用 LC 检索算法求最短路径, 针对虚拟结点, 应修改相应的站点信息, 将其路口标识置为 True, 然用运用 LC 检索算法求解. 算法过程如下:

① 若查询起始站点或终点是交叉路口结点, 设置虚拟结点路口标识为 True;

- ② 用 LC 检索算法求出最短路径;
- ③ 还原信息, 将虚拟结点路口标识置为 False.

(2) 分别处理策略

若查询的起始点(s)和终点(t)均为虚拟结点, 沿虚拟结点所在的线路至少能找到一对与其相连的交叉路口结点, 即起始点和终点均有一个前交叉点或后交叉点, 车辆行驶中必定要经过一个与起始点相连的交叉点和一个与终点相连的交叉点. 将各种情况进行排列组合最多可形成 4 个子图, 对这 4 个子图分别调用 LC 检索算法, 之后再选出最小值, 必为起点到终点的最短(优)路径.

记 $DIST(I, J)$ 为结点 I 到结点 J 的最短路径长度, $PATH(I, J)$ 为 I, J 之间的最短路径, $L(M, N)$ 为站点 M 到结点 N 的路径长度. 若与起点 S 邻接的交叉路口结点为 A 和 B 与终点 T 邻接的为 C 和 D. 则起点 S 到终点 T 的最短路径长度为:

$$D_{S \rightarrow T} = \min\{L(S, A) + DIST(A, C) + L(C, T), L(S, A) + DIST(A, D) + L(D, T), L(S, B) + DIST(B, C) + L(C, T), L(S, B) + DIST(B, D) + L(D, T)\}$$

通过获取最短路径长度的过程信息, 可找到相应的最短路径.

3.3 备忘录的使用

用户输入起点站和终点站进行查询时, 系统接收用户输入的信息并反馈查询结果. 对众多查询相同的情况, 若系统每次都按照 LC 算法完成一次搜索, 将耗费大量的时间和空间, 因此可以用系统备忘录来存储已得到的查询信息. 可根据需要设置多个备忘录, 每个备忘录可以按查询的频率数或最近查询等原则组织.

备忘录的数据表结构应包括以下几个方面:

起点站	终点站	查询次数	线路信息
-----	-----	------	------

其中, 线路信息即系统反馈的查询结果, 可以用文本方式给出具体的行车路线, 即经过的站点和换乘站点等.

为了将虚拟结点、备忘录的思想用于最短路径算法中, 在查询系统应用之前建立两个备忘录 1 和备忘录 2. 备忘录 1 中记录各个交叉点之间的最短路径即行车路线, 备忘录 2 中记录历史查询结果.

乘客查询时, 如果输入的起点站和终点站恰为交叉点, 就可以通过备忘录 1 直接输出行车线路; 若输入的起点站和终点站不是交叉点, 则在备忘录 2 中进行查询, 然后调用引入虚拟结点的 LC-检索算法搜索. 在系统中添加备忘录后, 查询流程图 3 所示.

值得注意的是在备忘录 2 中查询应遵循一定的搜索策略, 因为当备忘录中存有相当数量的查询结果时, 其查找仍然具有较高的计算时间复杂度. 对按某种方式排序的备忘录, 可以使用的策略是设定一阈值, 如全部查询的 10% 或前、后 500 条. 根据该阈值, 进行正序和逆序查找. 当未找到时, 则调用引入虚拟结点的 LC 检索算法搜索最短路径, 以提高系统查询的效率.

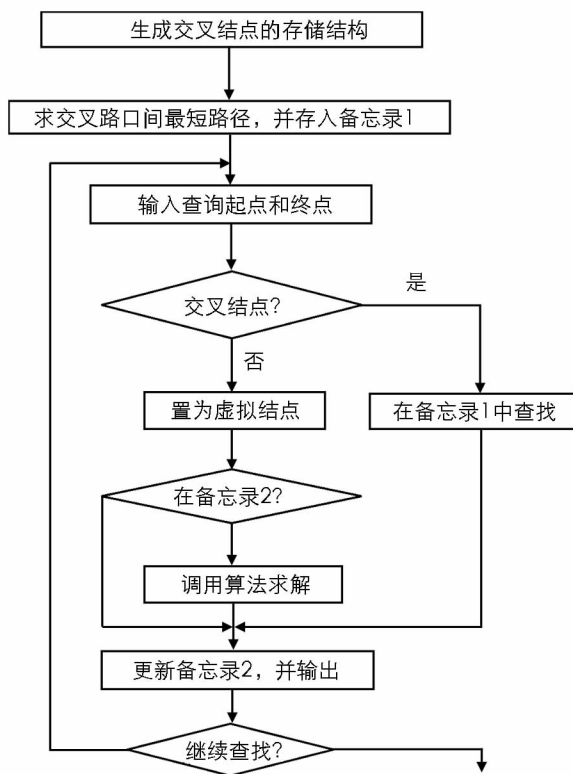


图 3 查询流程图

4 结 论

城市公交线路网络是一个复杂的网络图形,为获取最优路线,本文探讨了公交查询系统中的最短路径算法.针对公交网络的特点,结合 LC 检索的“有智力”的排序函数,提出了一种以交叉路口为结点的 LC-检索算法,并引入备忘录存储最短路径信息,以提高查询的效率及灵活性.

参考文献:

- [1] 王世祥,饶维亚.大中城市公交线路查询的数据结构及其算法的实现 [J]. 计算机系统应用, 2007, 9.
- [2] 袁小群,刘蕊.一种新的公交数据模型与公交查询系统实现 [J]. 计算机工程与应用, 2007, 3.
- [3] 董鑫,郑逢斌,李莘莘. Dijkstra 算法的改进及其在警用 GIS 中的实现 [J]. 郑州轻工业学院学报(自然科学版), 2007, 5.
- [4] 陈立潮,刘玉树,张永梅.城市交通智能咨询系统的设计与实现 [J]. 计算机工程. 2003, 29: 132-340.
- [5] Michel Mainguenaud. A Query Resolution Engine to Handle a Path Operator With Multiple Paths [J]. Transportation Research Part C, 2000, (8): 109-127.
- [6] 陆忠,钱翔东,张登荣.基于最短路径查询的城市公交网络拓扑建模研究 [J]. 遥感信息, 2002, 1: 11-14.

Research on the Application of LC - search in Urban Public Transportation Query System

ZHAO Jing-lu, HE Guo-bin

Southwest University, College of Computer and Information Science, 400715

Abstract: After analyzing the characteristic of urban public traffic network and making use of a data base to store the information of public transit line, in this paper, corresponding data sheet structure was given. Founded on Least Cost search-LC search, an improvement shortest path algorithm was produced which takes the crossway corners as crunodes and introduced the memorandum which stores the shortest-path information, as to enhance the efficiency and flexibility of the query.

Key words: public traffic network; shortest path; least cost-search; memorandum

责任编辑 陈绍兰