公交线路选择问题

国防科技大学：郭勇、陈项颢、易伟

摘　要

这是一个在城市公交网络中为乘客选择最佳公交线路的优化问题。本文首先总结了公交路径选择问题的研究现状，分析了城市中公交乘客的出行心理，紧紧围绕自主查询系统的实际应用展开讨论，将乘客关心的因素作为优化目标，拟定了各项指标的计算规则，从而建立了侧重于不同指标组合的优化模型，并用基于优先队列的广度扩展算法高效地计算出了问题所要求的线路组合方案。

在问题分析部分，我们先探讨了这个问题的研究现状，发现目前的研究大都以换乘次数最少为第一目标，同时过滤了很多实际情况中的细节，这样不可避免地使模型失真而不能较好地指导城市公交线路查询系统的建立。接下来，我们深入而全面地分析了公交乘客在出行时的心理特征，找出了换乘次数、出行时间、出行费用等优化指标，并说明了将所有指标综合考虑是没有实际意义的。然后讨论说明了“步行时间”和“同名站点”等一些容易被忽略却又影响模型建立的概念。最后说明了现实生活中乘客步行小段距离再转车现象，并拟在模型中加入对这个现象的考虑。

第一问实际意义不大，旨在引导出第二问乃至第三问的模型。故前两问采取相同的算法。针对公交网络的特点很自然地建立基于图论的优化模型，优化目标就是最少换乘、最短时间和最小费用。依据实际情况和给定的参数详细讨论各项指标的计算规则，合理组合上述三个优化目标，例如以最少换乘为第一目标、最短时间为第二目标。然后，用基于优先队列的广度扩展算法求解出最佳线路，算法的特点在于适应公交线路问题中的动态特性并且具有时间高效性。下表是整理后的结果（三项最优不一定同时达到）：

问题一	(1)	(2)	(3)	(4)	(5)	(6)
最少换乘(次)	1	2	1	1	2	1
最短时间(分)	67	102	106	62	105	49
最小费用(元)	3	3	3	2	3	2

问题二	(1)	(2)	(3)	(4)	(5)	(6)
最少换乘(次)	1	2	1	1	2	0
最短时间(分)	65	102	98	56.5	89.5	30
最小费用(元)	3	3	3	2	3	1

注：(1)代表 S3359→S1828，依此类推；具体线路见附录

对于问题三，步行作为一种独立的交通方式影响网络布局，它使任意两个结点都能连通，但实际情况中，乘客步行是有节制的，故引入“步行距离的心理承受上限”的概念来限制步行方式，这样就得到了改进的模型。模型的变更带来了更大的计算量，因而需要改进算法，采用双向搜索的方法提高时间效率。

最后，为了在最佳线路出现拥挤时乘客能有其他选择，本文提出了查询系统应提供多线路方案的改进方向。

[关键词] 出行心理 换乘次数 优先队列 心理承受上限 多线路方案

1 问题重述

我国人民翘首企盼的第 29 届奥运会明年 8 月将在北京举行，届时有大量观众到现场观看奥运比赛，其中大部分人将会乘坐公共交通工具（简称公交，包括公汽、地铁等）出行。这些年来，城市的公交系统有了很大发展，北京市的公交线路已达 800 条以上，使得公众的出行更加通畅、便利，但同时也面临多条线路的选择问题。针对市场需求，某公司准备研制开发一个解决公交线路选择问题的自主查询计算机系统。

为了设计这样一个系统，其核心是线路选择的模型与算法，应该从实际情况出发考虑，满足查询者的各种不同需求。请你们解决如下问题：

1、仅考虑公汽线路，给出任意两公汽站点之间线路选择问题的一般数学模型与算法。并根据附录数据，利用你们的模型与算法，求出以下 6 对起始站→终到站之间的最佳路线（要有清晰的评价说明）。

(1)、S3359→S1828　(2)、S1557→S0481　(3)、S0971→S0485
(4)、S0008→S0073　(5)、S0148→S0485　(6)、S0087→S3676

2、同时考虑公汽与地铁线路，解决以上问题。

3、假设又知道所有站点之间的步行时间，请你给出任意两站点之间线路选择问题的数学模型。

参数设定如下：

相邻公汽站平均行驶时间(包括停站时间)：3 分钟
相邻地铁站平均行驶时间(包括停站时间)：2.5 分钟
公汽换乘公汽平均耗时：5 分钟(其中步行时间 2 分钟)
地铁换乘地铁平均耗时：4 分钟(其中步行时间 2 分钟)
地铁换乘公汽平均耗时：7 分钟(其中步行时间 4 分钟)
公汽换乘地铁平均耗时：6 分钟(其中步行时间 4 分钟)

2 假设

1. 假设附录一、二中的基本参数和线路信息真实可信；
2. 假设乘客出行时都明确知道自己的起始所在点和想要到达的终点；
3. 假设公交站点在整个城区范围内均匀分布；
4. 假设乘客都是以公汽站为起始站；
5. 不考虑路况或其他因素对乘客步行即公汽和地铁行驶的影响；

3 符号说明

D	出行距离，即乘客从起始站到达终到站的线路总长
E	换乘次数，即乘客从起始站到达终到站所换车的次数
T	出行时间，即乘客从起始站到达终到站总共花费的时间
C	出行费用，即乘客从起始站到达终到站乘公汽和地铁的总费用

4　问题分析

4.1　问题背景和研究现状

随着城市的发展，公共交通网络日趋复杂，市民出行时面对庞杂的公交网络时常会感到茫然无措，因此，在大城市中一种用于解决公交线路选择问题的自主查询计算机系统应运而生。下面是北京地铁网站上的查询系统的界面：

图 1：北京市地铁线路查询系统

这种系统一般在公汽站或地铁站设置一个触摸屏终端，乘客只需输入起始站和终到站（乘客想要到达的目标站点），系统很快自动给出可供选择的几组公交线路组合方案。本文的目的正是为这样的系统设计线路选择的模型和算法，也正因为如此，下面的讨论都是紧紧围绕这个实际应用而展开的。

这是个为公交乘客在任意两个给定的站点之间，根据其需求（包括换乘次数、出行距离、出行耗时、出行费用等），选择一套最佳出行线路组合的问题。显然，这个问题的关键技术是求解这套最佳路线的算法。经过一番文献查找后我们发现，对这个问题的研究目前已经有相当多的成果，大多研究将“换乘次数”（4.3 中对这个概念做了解释）做重点考虑。另一方面，这类算法已趋成熟，涉及到 Dijkstra 算法、蚂蚁算法[6]、动态规划[4]、基于宽度优先的算法[7]，等等。这些算法的共同点是都以换乘次数作为首要考虑用图论的模型来求解最佳线路。这些研究将多数人最关心的问题——换乘次数作为第一目标虽然有一定的合理性，但是这样做也有局限性，因为有的人由于赶时间需要将时间最短作为第一考虑，而有的人由于经济拮据需要将费用最少作为第一考虑，等等。这就是说，模型应该要根据乘客的需求不同而给出不同的决策，这也是将这个模型应用在城市中供乘客自主查询公交线路的计算机系统中的生命力所在。

另外，将城市中所有公共交通设施（公交、地铁、步行）全部加入考虑范围尚属比较新的课题，是本文重点的讨论方向。

4.2　对公交乘客的出行心理分析

既然要为公交乘客提供线路建议，首先要了解公交乘客的出行心理。总的来说，乘客希望以最短的时间和最小的代价到达目的地，而文献[5]更具体地给出了乘客出行时所关心的几个因素：换乘次数 E、出行距离 D、出行时间 T、出行费用 C。这些概念已经在符号说明中作了解释，不再赘述。对这些因素的理解是：

第一，最短线路组合不一定是最佳方案。因为最短路线组合很可能会使乘客频繁地换乘不同线路的公交车，这是乘客所不愿意的，因此乘客在出行时，总是希望最好能搭乘一趟车直接到目的地，除非这趟车确实是绕了个大弯浪费了很多时间，或者根本没有这种直达的车，乘客才会考虑在中途换乘其他车次。这就是提出换乘次数这个概念的初衷；

第二，出行距离对应途经站数，即出行距离越长途经站数就越多。本问题中不考虑从起点到上车站台的距离和从下车站台到终点的距离，且第一问不考虑中途换车所步行的距离。这样的话，实际上只要时间上尽量快，乘客是不太会关心出行距离的，因为整个出行过程中乘客基本都在车上（步行换车在 5.5 中单独考虑），车行驶了多长的距离都不会使他太在意，而一般情况下，要时间尽量快就得选择途经站数少的即距离较短的车次，这样距离就与时间统一起来了；

第三，这里讨论的出行时间只包括车行时间、换乘时间和步行时间。特别地，当不考虑步行方式时，只包括前两者；

第四，在一票制的城市中费用与换乘次数是一致的，但是北京的公交采取单一票制与分段计价结合，使两个概念有了区别，但是这也不会对二者的一致性产生太大的影响。

1999 年在南京市 8 个公交站点进行了一次公交乘客出行心理问卷调查 [5]，共得到有效表格 440 份，图 2 是调查结果。

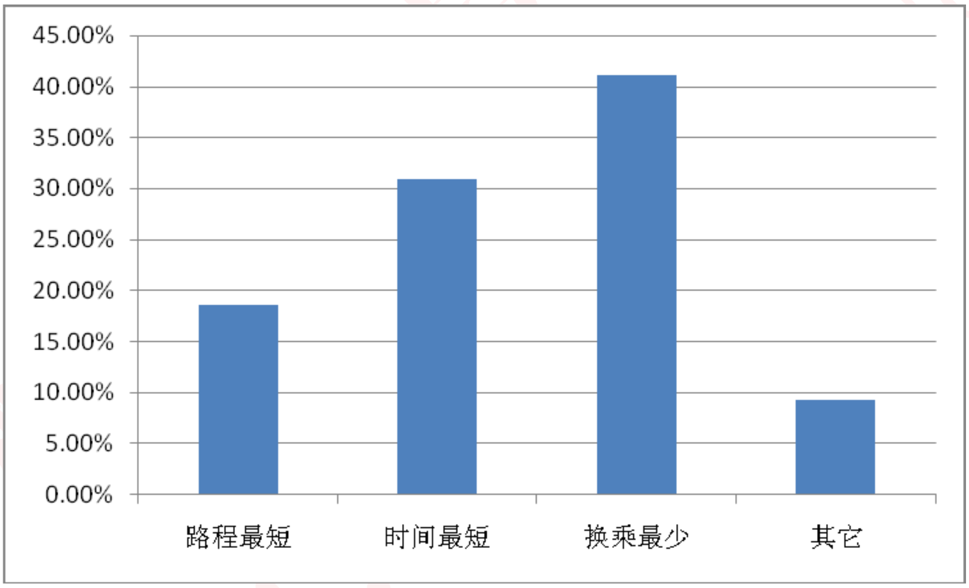

图2：公交乘客出行心理调查

从图中可以看出，有 41.16%的乘客在出行时首先考虑换乘次数最少。这个结论是合理的，因为乘客希望以最短的时间和最小的代价到达目的地，而从一条线路换乘到另一条线路费时费力，且引起不可预知的因素大大增加，如下一趟车到来的时间、路况、第三问中换乘所需的步行距离等，都会使花费的时间增加[3]。这也是目前大量相关论文都以换乘次数为第一目标的原因。其次是考虑时间最短的乘客，占到 30.93%，然后才是考虑路程和费用问题的乘客。

综上所述，为给出线路组合，要考虑不同乘客的不同需求，但特别要指出的是，没有必要对各个指标进行综合考虑，原因是：综合考虑的方法一般是按照各个指标的重要程度对其求加权和作为目标函数，当存在所有指标同时最优解时，

综合考虑与分开考虑在结果并没有区别，而计算却复杂得多；当不存在所有指标同时最优解时，加权得到的结果总是顾此失彼，结果往往对任意一个指标都不是最优，而本文在讨论该问题的研究现状时已经提出，乘客往往有一个自己最为关心的因素，或是时间或是换乘次数，等等。因而这种综合策略对于想要通过这个系统查询线路的乘客来说，并没有实际意义。

4.3 对换乘和步行时间、公交站点和线路的理解

问题中有几个概念不太明确。首先是附录 1 中的换乘平均耗时，这个概念本身并不含糊，但是它包含一个“步行时间”，那么，理所当然这个换乘平均耗时分为等车时间和步行时间。而实际中步行有两种情况，一是乘客搭乘公汽下车后穿过马路到达对面的同名站点需要步行，二是乘客为了快捷，从一个站点走到附近的另一个站点转车需要步行。但是，第三问中提到“假设又知道所有站点之间的步行时间”，说明前两问中并不知道“所有站点之间的步行时间”这个参数，也就是说为了简单起见，前两问不需要考虑第二种步行方式。这里又涉及另一个问题，就是如果两条线路有相同的站点，也有两种情况，一是的确在同一个站点相交，二是两条线路经过同名的两个站点，且一般说来，这样的两个站点是隔着马路相对应的，只需穿过人行道或天桥或地下通道就能到达，与附录 1 中公汽换公汽或地铁换地铁的步行时间相近，故本文把附录 1 的“步行时间”理解成同名车站间换乘所花的步行时间。还有一点需要说明的是，经过查询得知，北京公汽和地铁的环线都是双向的。

4.4 步行对线路选择的影响

在前面的模型中，公汽与公汽之间换乘的前提是不同路线之间有公共的站点或者两个站点有公共的地铁站相连，但事实上换乘的方式不只有这一种，因为这样考虑忽略了一个实际情况，就是在现实生活中乘客步行小段距离再转车的现象。乘客在某个站点下车时发现没有直接换乘的车次或者可供换乘的车次需要绕很大的弯，他就会考虑附近是否有更为理想的站点。乘客选择这种方式可以减少换乘次数、节约时间[5]。但步行本身要耗时，更重要的是步行成分在整个出行过程中的增加会使乘客在生理上产生疲劳感和在心理上产生不满情绪，因而这种步行方式必须有节制地采用。

5 模型建立与求解

5.1 优化目标和约束条件

根据对公交乘客出行的心理分析，可以把问题描述为一个多目标优化问题，这里的多目标包括换乘次数 E 最少、出行费用 C 最少、出行时间 T 最短和出行距离 D 最短。

这样，公交路线选择模型就可以描述为：给定起点 V_0（对应起始站）和终点 V_t（对应终到站），将公交网络看成一个图 $G=\{g \mid g=\langle V_i, P_i\rangle, i=0, 1, \cdots, t\}$，则优化目标为：Min　Z(E，C，T，S，D)，由于前面讨论了 S 和 D 对线路选择的影响很小，可以将其简化为 Min　Z(E，C，T)。

但是在实用查询系统中，当乘客给出他最关心的一个或两个指标要求得到线路组合方案时，应该在输出具体的线路组合方案的同时，输出这个方案的总时间、总费用、换乘次数、在哪一站换乘等信息，由于问题没有给出有关车站距离的信息，就不考虑提供最后的出行距离的信息了。

5.2 公汽网络中的最佳线路组合

模型首先应用于一个只有公共汽车网络的理想情况中。为简单起见，假设当乘客需要转车（即换乘）时总是在他下车的车站等车，而不会步行到另一个公汽站搭车。同时也暂不考虑从一个公汽站通过地铁站步行到另一个公汽站的情况，因为如果这个只有公汽模型用于没有地铁的城市，那就根本不存在从一个公汽站通过地铁站步行到另一个公汽站的问题，而只有公汽网络的模型更不可能用于有地铁的城市。

这里需要强调的是，第一问既然已经不考虑地铁的影响，本来就已经与北京等大城市的实际情况不符，故而暂且不讨论这个初步模型与实际情况的贴近程度，而是让模型层进式地一步步贴近实际，第一步就只讨论这个简化模型。

但是，这个简化的情况已经被前人研究得十分透彻了，本文并不仅仅要解决这种简单情况，更重要的是要解决将城市中的各种交通方式——公汽、地铁、步行全部考虑进来的复杂情况，因此虽然模型将朝着这个方向改进，它的雏形也不打算用比较中庸的模型或算法，而是直奔目标。

按照 5.1 的讨论，以所有公汽站为结点集合，凡两个站是同一条公汽线路上的相邻站，则在它们对应的结点之间连一条带权有向边（顺公汽线路方向），边的权重为 3（相邻公汽站平均行驶时间为 3 分钟），当从一条路线换乘到另一条时，会增加时间代价，题设公汽换乘公汽平均耗时是 5 分钟，其中步行时间 2 分钟，这个步行时间应理解为在某站下车后走到马路对面的时间，故而换乘等待时间为 3 分钟，据此假设乘客在第一站的平均等待时间为 3 分钟，这个假设是合理的，因为附录 1 中四个换乘耗时有这样的规律：

5 分钟 — 2 分钟 = 7 分钟 — 4 分钟 = 3 分钟

4 分钟 — 2 分钟 = 6 分钟 — 4 分钟 = 2 分钟

这就说明了在公汽站等车一般平均为 3 分钟，在地铁站等车一般平均为 2 分钟。

定义局部符号如下：

N_{s}　　线路中单一票价的线路条数

N_{m}　　线路中分段计价的线路条数

S_{bus}　　途中公汽行驶的站数

ST_i　　第 i 条分段计价线路的途经站数，其中 i=1，2，…，N_{m}

综上得到时间和费用的表达式：

$$T = 3 + S_{\mathrm{bus}} \times 3 + E \times 5\ ;$$

$$C = 1 \times N_S + \sum_{i=1}^{N_m} f(ST_i)\ ;$$

其中， $$f(x)=\begin{cases}1, 0 < x \le 20\ ; \\ 2, 20 < x \le 40\ ; \\ 3, x > 40\ .\end{cases}$$

在 Z（E，C，T）中如何权衡 E、C、T 是直接影响结果的，这个权衡的标准就是乘客对这三个指标的关心程度，或表述为指标对乘客的重要程度。由于数据量庞大，直接对三者综合考虑会使计算过于复杂，故每次取两个指标（分别作为第一和第二目标）来计算。为满足不同需求的乘客，以具有明显区别的三个需求（时间、费用、换乘次数）为第一目标。由于费用与换乘次数的关联性，不考虑以费用与换乘次数的搭配。综上，本文采取如下三种优化方案：

(1)以换乘次数最少为第一目标、时间最短为第二目标。

(2)以费用最少为第一目标、时间最短为第二目标；

(3)以时间最短为第一目标、费用最少为第二目标；

下面的问题就是讨论求解该模型的算法，首先建立这样的数据结构：

(1)换乘线路列表：该列表的每个单元都记录某个结点所有的换乘线路和所有换乘后的下一个车站结点，并与这样的下一个车站结点连一条有向边，另外还记录换乘时间；

(2)运行线路列表：该列表的每个单元都记录经过某个结点的所有线路和所有该结点在这些线路上相应的下一站即后继结点，另外还记录时间和费用；

(3)状态：以“路线，车站，时间，费用，换车次数”表示一个线路组合方案中在某一时刻的一个状态。建立状态后，整个出行过程实际可以看成一个状态的转移过程，算法流程图如下：

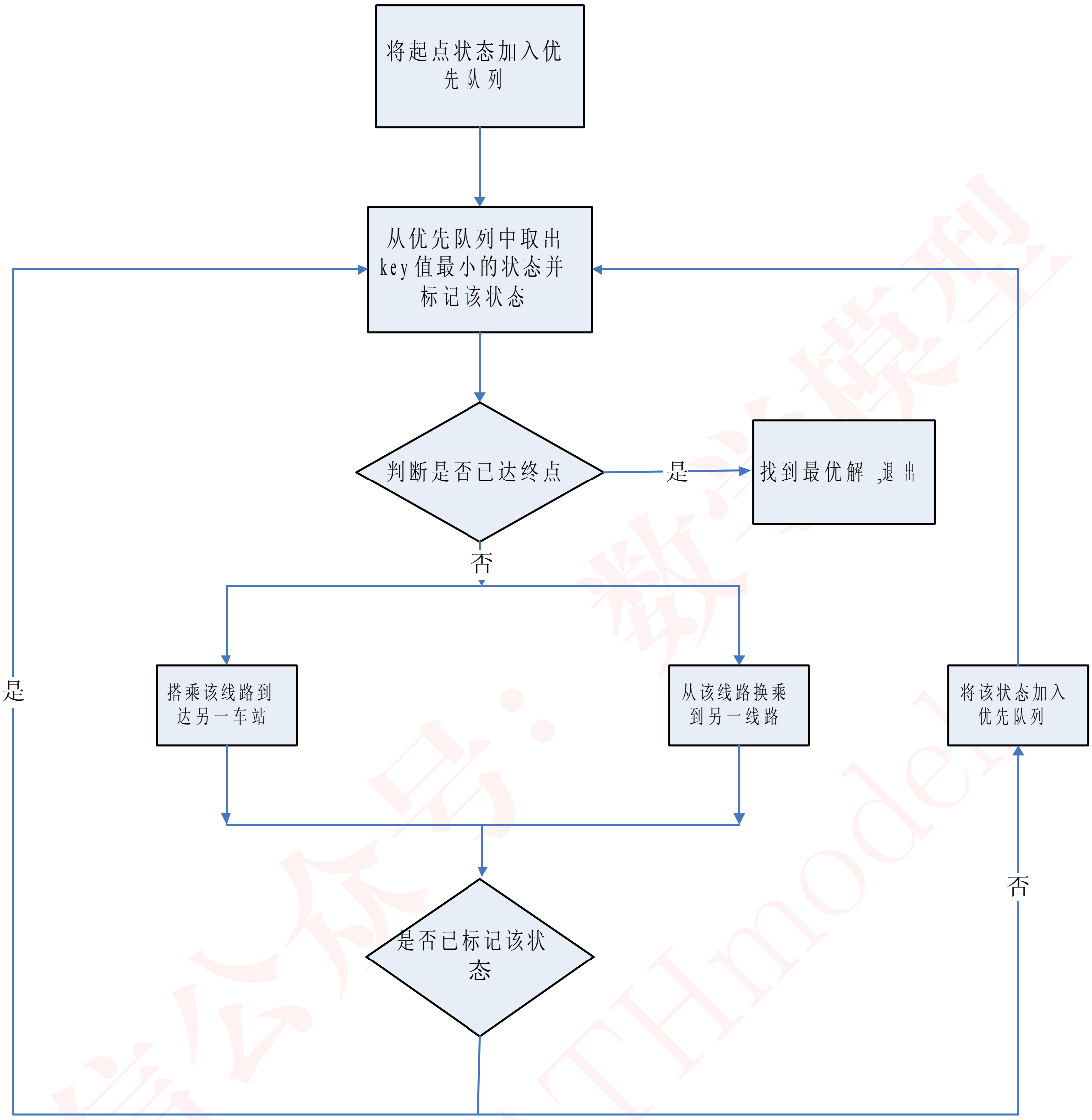

图 2 基于优先队列的广度扩展算法流程图

注：优先队列是一种可动态更新的数据结构 [1]，用优先队列中保存的状态为当前的“路线、车站、费用、时间、换车次数”，每次可以 logN 的时间复杂度获得队列中 key 值最小的状态量，例如，以费用为 key 值，则从队列取出的状态量均为费用最小的状态量。当然，也可以用一个数据组合作为 key 值，以求最小费用、时间为例，先比较费用，费用较小的优先，若费用相同，再比较时间。因此选取不同的 key 值就可获得不同的需求。

算法可行性讨论

经典的 Dijkstra 算法用于解决静态模型的最优解，但是由于这是一个分别以费用、换乘次数、时间作为路径长度的变向最短路问题，需要建立动态模型求解，而算法中引入的状态变化正是一个动态计算过程。

算法高效性讨论

问题中共有公交线路 520 条，地铁线 2 条，公交车站 3957 个，地铁站 39 个，网络规模很大，因而在实用查询系统中对算法的时间效率要求很高。由于公交网络具有稀疏性（有向边的数目远远低于结点的平方），可以使用优先队列，即每次取出一个已经达到最优的结点扩展出其它状态，保证了每个状态只被扩展

一次，简略估计算法复杂度如下：

总的状态数 = 总的线路数×平均每条线路车站

转移量 = 平均每个车站能换乘的方式数 + 平均每条线路的长度。

总的状态数 $\le 522\times(3957+39)\div 522=3957$，转移量远远小于 522，因此算法最坏情况的复杂度为：状态数×转移复杂度<5000000,任意两点之间的路径方案都可在 0.5 秒内计算获得，满足实际中对时间效率的要求。

经过编程计算，得三种计算方案的结果（具体路线见附录）如下表：

表 1：只考虑公汽线路时各线路的部分查询结果

起始站→终到站	第一目标	换乘次数	出行时间(分钟)	出行费用(元)
(1)S3359→S1828	换乘最少	1	104	3
	时间最短	2	67	3
	费用最少	2	67	3
(2)S1557→S0481	换乘最少	2	109	3
	时间最短	3	102	4
	费用最少	2	109	3
(3)S0971→S0485	换乘最少	1	131	3
	时间最短	2	106	3
	费用最少	2	106	3
(4)S0008→S0073	换乘最少	1	86	2
	时间最短	4	62	5
	费用最少	1	86	2
(5)S0148→S0485	换乘最少	2	109	3
	时间最短	3	105	4
	费用最少	2	109	3
(6)S0087→S3676	换乘最少	1	68	2
	时间最短	2	49	3
	费用最少	1	68	2

5.3 公汽与地铁混合网络中的最佳线路组合

上面得到了简单情况下的模型及算法，但是研究现状中能看到，这个简化的情况目前已经被研究得比较透彻了，故更重要的是考虑将地铁加入到整个公交网络中来，形成公汽与地铁的混合网络。这时，问题较为贴近实际也因此变得复杂很多。一是前面讨论过的从一个公汽站通过地铁站步行到另一个公汽站的问题，在这里需要加以考虑了，二是由地铁换乘公汽或由公汽换乘地铁都需要步行且平均耗时不一样，三是地铁与公汽平均行驶时间和计价都不一样。然而问题还是转换成一个网络模型，这个网络比第一问中要复杂，也就使得算法须得修改，而且求解的计算量大大增加。

计算规则如下：

(1)如果刚一开始选择搭乘公汽，则等第一趟车需要 3 分钟，而若选择搭乘地铁，则等第一趟车需要 2 分钟；

(2)从公交站通过地铁站步行到另一个公交站转车是公汽转公汽的一种情况，不做单独考虑；

(3)考虑在最后一站时的特殊情况，从公汽或地铁下车后步行到达终到站，分别为 2 分钟和 4 分钟。

由于参数有变化，定义局部符号如下：

S_{bus}　　途中公汽行驶的站数

S_{Sub}　　途中地铁行驶的站数

E_{bb}　　公汽换乘公汽次数

E_{bs}　　公汽换乘地铁次数

E_{sb}　　地铁换乘公汽次数

E_{ss}　　地铁换乘地铁次数

N_{S}　　公汽线路中单一票价的线路条数

N_{m}　　公汽线路中分段计价的线路条数

N_{Sub}　　地铁线路条数

ST_i　　第 i 条分段计价线路的途经站数，其中 i=1，2，…，N_{m}

各个指标的算式表达为：

$$T = w_0 + S_{\text{bus}} \times 3 + S_{\text{Sub}} \times 2.5 + E_{\text{bb}} \times 5 + E_{\text{bs}} \times 6 + E_{\text{sb}} \times 7 + E_{\text{ss}} \times 4 + w_{\text{t}} \ ;$$

$$w_0 = \begin{cases} 2\text{，刚一开始选择搭乘地铁} \\ 3\text{，刚一开始选择搭乘公汽} \end{cases}$$

$$w_{\text{t}} = \begin{cases} 2\text{，从公汽站走到终到站} \\ 4\text{，从地铁站走到终到站} \end{cases}$$

$$C = 1 \times N_S + \sum_{i=1}^{N_{\text{m}}} f(ST_i) + 3 \times N_{\text{Sub}} \ ;$$

$$\text{其中，}\ f(x) = \begin{cases} 1, 0 < x \le 20 \ ; \\ 2, 20 < x \le 40 \ ; \\ 3, x > 40 \ . \end{cases}$$

与前面采取同样的算法和同样的指标搭配方式算得结果（具体路线见附录）如下表：

表 2：考虑公汽和地铁线路时各线路的部分查询结果

起始站→终到站	第一目标	换乘次数	出行时间(分钟)	出行费用(元)

(1)S3359→S1828	换乘最少	1	104	3
	时间最短	**4**	**65**	**7**
	费用最少	2	67	3
(2)S1557→S0481	换乘最少	2	109	3
	时间最短	3	102	4
	费用最少	2	109	3
(3)S0971→S0485	换乘最少	1	131	3
	时间最短	**3**	**98**	**6**
	费用最少	2	106	3
(4)S0008→S0073	换乘最少	1	86	2
	时间最短	**3**	**56.5**	5
	费用最少	1	**83**	2
(5)S0148→S0485	换乘最少	2	**90.5**	**5**
	时间最短	3	**89.5**	**6**
	费用最少	2	**106**	3
(6)S0087→S3676	换乘最少	**0**	**30**	**3**
	时间最短	**0**	**30**	3
	费用最少	**0**	**37**	**1**

注：字体加粗部分为与表 1 中不同的结果。

5.5　结果分析

问题一和问题二的结果在某些指标上发生了变化，这个变化主要有这么几个特点：

第一，时间在第二问中相对于第一问全都是减少的，而换乘次数和费用有升有降。这个结果是合理的，因为在城市中建立地铁体系不仅可以缓解公汽的压力，更重要的是地铁的速度较公汽快，而且不存在地面上交通工具需要面临的交通阻塞或路况异常等问题，能够使乘客更快到达目的地，因此，将地铁加入到城市公交网络中后，就应该能使乘客的时间花费降低；

第二，以费用为第一目标时，六套线路都没有包含地铁，因为地铁较之公汽票价应该说是贵一些；以换乘次数为第一目标时，包含地铁不多，原因在于地铁的站点少，覆盖面小；以时间为第一目标时，六套线路中有五套线路包含地铁，这又体现了地铁的快速优点。

第三，两个费用上升的方案都是对应地缩短了时间，这说明为了使时间尽量短，可以选择搭乘票价稍贵的地铁；

第四，线路 S0087→S3676 的换乘次数全都变为 0，就是说在有了地铁以后，以 S0087 为起点站、以 S3676 为终到站的乘客在整个过程中只需搭乘地铁，且这种直达线路的时间花费明显减少这是地铁对城市公交带来方便的最集中体现。

第五，总体来说，地铁较之公汽占的比例并不很大，再随机给定 20 组起点和终点求解得到的线路发现只有三组通过地铁改善目标，更说明了地铁的劣势地位，原因在于公汽站有 3957 个，遍布整个城市，而地铁站却只有 39 个，覆盖面小，这就是说，北京还需要修建几条地铁线，事实上，目前北京的地铁线已经有四条而不是问题中的两条，车站也已经是 70 个了，但是为满足奥运会期间的需求，这还是远远不够的。

5.6 考虑步行后的修正模型

步行方式的引入对原先模型产生这样的影响：可以在两个没有边相连的结点之间连上一条边，以减少换乘次数和时间等代价。单纯地这样考虑会使网络变为完全图，计算量太大。实际上，乘客并不愿意频繁地步行，也就是这样的边不能随便连接，因为前面说明了步行会使乘客产生疲劳感和不满情绪，故提出乘客"步行距离的心理承受上限"的概念，经验表明它与步行距离大致呈线性相关。

从乘客心理上来说，步行的最主要目的是减少换乘次数，故先讨论以换乘为第一目标的路线选择。统计表明，在公交网络中通过 3 次以上换车到达目的地的情况是很少发生的[5]，且本文前两问的计算结果也显示换车次数很少超过 3 次，因此，在本问中不考虑换乘超过 3 次的情况。

为了证明改进的模型的正确性，须得找一组模拟数据求解出合理的结果，但是由于不清楚北京公交站点的分布，不可能假设出任意两个站点之间步行需要的时间，也没找到这方面的数据资料，故而无法求解。若构造一个公交网络用来测试，一方面毕竟不是真实的网络缺乏说服力，另一方面，无法短时间内构造出足够大的网络，而在一个小网络中做优化其实意义不大，因此没有必要这样做。另一方面，我们可以给出完整的算法实现细节，如果给出所有站之间的步行时间，就可以算出最优解。

定义局部符号如下：

V　　　　正常人步行的速率，一般为 1.5 米/秒；

L　　　　乘客在换乘时步行距离的上限；

D(A，B)　　站点 A 和站点 B 的距离；

T(A，B)　　站点 A 到站点 B 的步行时间。

若 D(A，B) $\le$ L，即两站之间距离没有超过步行距离的心理承受上限，则乘客有可能从站点 A 步行到站点 B，称 A 和 B 可达。由附件 2 可知同一地铁站对应的任意两个公汽站之间可以通过地铁站换乘，故可认为这样两个公汽站互为可达站点。

由于 T(A，B)的加入，公交网络图中边数剧增，运算量太大，使得前面算法中利用原图稀疏性的优点得不到发挥。因此，在此问中本文采用了从起点和终点同时搜索的方法，即双向搜索法，这种方法可以使搜索量减少一半

给定起点站 A 和终到站 B，以换乘次数的递增顺序进行搜索[5]（见图 3）：

a). 换乘次数为 0：即从 A 可以乘车直达 B，首先搜索出所有经过 A 的线路 RA，再搜索出所有经过 B 的线路 RB，若有相同的线路，则该线路为从 A 到 B 的直达线路（图 3，①）。若没有相同的线路，则考虑换乘次数为 1 的情况。

b). 换乘次数为 1：找出 RA 和 RB 通过的所有站点 M 和 N，若站点有相同，可通过此站点换乘一次到达 B（图 3，②）；若 M 和 N 中存在两个站点 X，Y 使 D(X，Y)$\leqslant$ L，也可通过从 X 步行到 Y 来实现换乘（图 3，③）；若这两个条件都不能满足，考虑换乘次数为 2 的情况。

c). 换乘次数为 2：找出所有过任意 M 和 N 及它们可达站点的线路 SA 和 SB，若 SA 和 SB 中有相同的线路，可从 M 乘车到达 N（图 3，④）或从 M 步行至可达站点 X，再从 X 乘车到达 Y，Y 步行至 N，从N 乘车到达 B（图 3，⑤），但这两种路线只是二次换乘的部分方式，因为可从 M、N、X、Y 中任选三点构成满足要求的线路。若没有相同的线路，考虑换乘次数为 3 的情况。

d). 换乘次数为 3：找出 SA 和 SB 通过的所有站点 P 和 Q，若站点有相同，

从 M 可通过此站点换乘一次到达 N（图 3，⑥），线路总的换乘次数为 3；若 M 和 N 中存在两个站点 X，Y 使 D(X，Y)≤ L，也可通过从 M 乘车到 X，X 步行到 Y，Y 乘车到 X 来实现（图 3，⑦）；由于这两条线路是在换乘次数为 2 的线路上得来的，所以它们也只是 3 次换乘的部分方式。若这两个条件都不能满足，由于此时已到达考虑换乘次数的上限，得不到可行的线路，算法结束。

对应这四种换乘得到的线路解，往往都是多条线路，这时可以通过乘客的要求建立第二目标，在所得到这些线路中搜索符合乘客要求的线路。

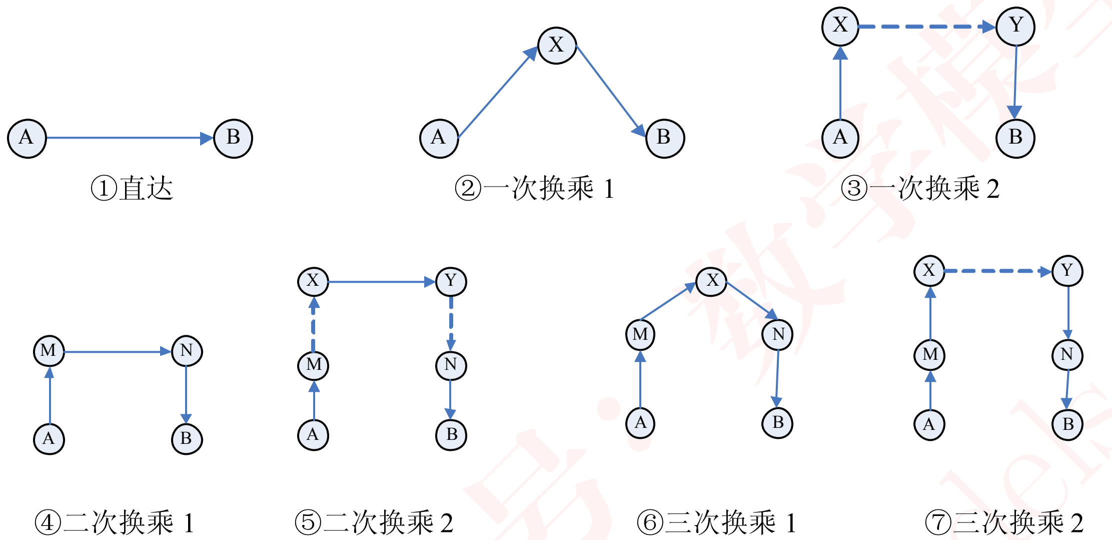

图 3：各换乘次数对应的部分换乘方式图

依据以上的算法，只要给定所有站点之间步行需要的时间就可以很快求解出结果。此算法经过改进，还可以用来求解以时间最短或费用最少为第一目标的最佳线路。基本思路为:搜索出所有三次以内换乘的线路并存储，再根据第一目标的要求检索这些线路，得到最优解。

6 模型的进一步讨论

上面的讨论遗漏了一个问题：实际中不能只提供一套最佳线路。原因有二：一是如果只提供一套方案，就会使始终点相同的乘客全都搭乘同一线路，导致该线路上的公交车过于拥挤，二是当最佳路线所在的线路出现交通拥塞或者路况异常，又或最佳路线上的车辆内出现拥挤，就必须考虑提供其他路线。

通过上面的计算结果可以看出，次优的线路在各项代价上往往与最佳线路相差无几。因此，在实际应用中有必要提供多套路线方案和实时的线路拥挤程度以供乘客选择，实时的线路拥挤程度不是本模型的讨论范围，我们只须考虑提供多套路线方案。

求解最佳线路的同解或次优解是比较复杂的。大体思路是将已经求出的最优解屏蔽，但不能屏蔽任何一个结点，用原来的算法重新计算即可得到同解或次优解。但实施的过程中有很多困难。另外，次优解的路线往往与最优解重叠很多，不能解决上述实际问题，因而本文没有对此进行实现。

本文主要是针对自主查询系统的实际应用来展开讨论的，故模型可以在“GIS”、“乘客舒适度”、“查询结果地图化”，“系统智能化”等方面继续进行改进。

7 模型评价

第一，本文紧贴实际应用，在已有的研究成果的基础上，深入分析城市中公交乘客的出行心理，全面考虑乘客的不同需求，针对不同需求提供不同的线路建议，解决了传统的算法只考虑了仅以换乘次数最少为第一目标的问题，真正做到了与实际应用相结合；

第二，本文在建立优化目标过程中，深入讨论细节问题，例如费用计算中的特殊情况等，使得目标函数建立得比较完善，保证了模型求解出的最优解是合乎实际的名副其实的最优解；

第三，模型求解的算法是该问题中的关键技术之一，本文中采取的基于优先队列的广度扩展算法适应公交线路问题中的动态特性，以较好的时间效率计算出了结果，进一步提高了模型的实用性；

第四，为了更加贴近实际，本文考虑了乘客以步行方式减少时间代价的情况，根据乘客心理状况引入了步行距离上限，得到比较完善的改进模型；

第五，由于问题三是“假设知道所有站点之间的步行时间”，而没有提供这个数据，并没有像前两问一样算出结果，不能不说是个不足指出，但是模型和算法已经给出，一旦给定所有站点之间的步行时间，就可以马上计算出结果。同样地，模型进一步讨论中提出的多路线问题具有非常实际的应用前景，但是本文出于时间考虑没能实现，这也是整个模型的一个缺憾。

参考文献

[1] Thomas H.Cormen etc，INTRODUCTION TO ALGORITHMS，The MIT，2002
[2]杨新苗等，基于 GIS 的公交乘客出行路径选择模型，东南大学学报，2000 年第 30 卷第 6 期
[3]徐兵、谢仕义，基于站点优先级的公交换乘算法实现，计算机时代，2005年第7期
[4]赵巧霞等，以最小换乘次数和站数为目标的公交出行算法，计算机应用，2004 年第 24 卷第 12 期
[5]王建林，基于换乘次数最少的城市公交网络最优路径算法，经济地理，2005 年第 25 卷第 5 期
[6]张帅等，蚂蚁算法在公交查询最短路径求法中的应用，华中科技大学学报，2003 年第 31 卷
[7]孙湧，基于宽度优先遍历树的公交线路换乘算法，深圳职业技术学院学报，2004 年第 4 期

附录

问题一：

一、换乘次数为第一目标、时间最短为第二目标的路线方案：

(1)、S3359→S1828

L436 S3359 L436 S1784 L167 S1784 L167 S1828

表示从起始站S3359乘坐L436号公汽线到S1784车站，换乘L167号公汽线到达终到站S1828

Time=104 Cost=3 Change=1

(2)、S1557→S0481

L084 S1557 L084 S1919 L189 S1919 L189 S3186 L460 S3186 L460 S0481

Time=109 Cost=3 Change=2

(3)、S0971→S0485

L013 S0971 L013 S2184 L417 S2184 L417 S0485

Time=131 Cost=3 Change=1

(4)、S0008→S0073

L159 S0008 L159 S0291 L058 S0291 L058 S0073

Time=86 Cost=2 Change=1

(5)、S0148→S0485

L308 S0148 L308 S0036 L156 S0036 L156 S3351 L417 S3351 L417 S0485

Time=109 Cost=3 Change=2

(6)、S0087→S3676

L454 S0087 L454 S3496 L209 S3496 L209 S3676

Time=68 Cost=2 Change=1

一、 时间最短为第一目标、费用最少为第二目标的路线方案：

(1)、S3359→S1828

L484 S3359 L484 S3727 L485 S3727 L485 S1784 L217 S1784 L217 S1828

Time=67 Cost=3 Change=2

(2)、S1557→S0481

L363 S1557 L363 S1919 L189 S1919 L189 S3186 L091 S3186 L091 S0903 L516 S0903 L516 S0481

Time=102 Cost=4 Change=3

(3)、S0971→S0485

L013 S0971 L013 S2517 L290 S2517 L290 S2159 L469 S2159 L469 S0485

Time=106 Cost=3 Change=2

(4)、S0008→S0073

L198 S0008 L198 S1691 L476 S1691 L476 S2085 L017 S2085 L017 S0604 L328 S0604 L328 S0525 L103 S0525 L103 S0073

Time=62 Cost=5 Change=4

(5)、S0148→S0485

L308 S0148 L308 S3604 L081 S3604 L081 S2361 L156 S2361 L156 S3351 L417 S3351 L417 S0485

Time=105 Cost=4 Change=3

(6)、S0087→S3676

L021 S0087 L021 S0088 L231 S0088 L231 S0427 L462 S0427 L462 S3676

Time=49 Cost=3 Change=2

三、费用最少为第一目标、时间最短为第二目标的路线方案：

(1)、S3359→S1828

L484 S3359 L484 S3727 L485 S3727 L485 S1784 L167 S1784 L167 S1828

Time=67 Cost=3 Change=2

(2)、S1557→S0481

L084 S1557 L084 S1919 L189 S1919 L189 S3186 L460 S3186 L460 S0481

Time=109 Cost=3 Change=2

(3)、S0971→S0485

L013 S0971 L013 S2517 L290 S2517 L290 S2159 L469 S2159 L469 S0485

Time=106 Cost=3 Change=2

(4)、S0008→S0073

L159 S0008 L159 S3614 L058 S3614 L058 S0073

Time=86 Cost=2 Change=1

(5)、S0148→S0485

L308 S0148 L308 S0036 L156 S0036 L156 S3332 L417 S3332 L417 S0485

Time=109 Cost=3 Change=2

(6)、S0087→S3676

L454 S0087 L454 S3496 L209 S3496 L209 S3676

Time=68 Cost=2 Change=1

问题二：

一、 换乘次数为第一目标、时间最短为第二目标的路线方案：

(1)、S3359→S1828

L436 S3359 L436 S1784 L167 S1784 L167 S1828

Time=104.0 Cost=3 Change=1

(2)、S1557→S0481

L363 S1557 L363 S1919 L189 S1919 L189 S3186 L460 S3186 L460 S0481

Time=109.0 Cost=3 Change=2

(3)、S0971→S0485

L013 S0971 L013 S2184 L417 S2184 L417 S0485

Time=131.0 Cost=3 Change=1

(4)、S0008→S0073

L159 S0008 L159 S0400 L474 S2633 L474 S0073

Time=83.0 Cost=2 Change=1

(5)、S0148→S0485

L024 S0148 L024 S1487 T1 D02 T1 D21 L051 S0466 L051 S0485

Time=90.5 Cost=5 Change=2

(6)、S0087→S3676

L454 S0087 T2 D27 T2 D36 L326 S3676

Time=30.0 Cost=3 Change=0

二、时间最短为第一目标、费用最少为第二目标的路线方案：

(1)、S3359→S1828

L132 S3359 L132 S2903 L201 S2903 L201 S0609 T2 D12 T2 D37 L428 S1961 L428 S1671 L041 S1671 L041 S1828

Time=65.0 Cost=7 Change=4

(2)、S1557→S0481

L363 S1557 L363 S1919 L189 S1919 L189 S3186 L091 S3186 L091 S0903 L239 S0903 L239 S0481

Time=102.0 Cost=4 Change=3

(3)、S0971→S0485

L094 S0971 L094 S0567 T1 D01 T1 D15 L156 S2534 L156 S3332 L417 S3332 L417 S0485

Time=98.0 Cost=6 Change=3

(4)、S0008→S0073

L200 S0008 L200 S2534 T1 D15 T1 D12 T2 D12 T2 D25 L103 S0525 L103 S0073

Time=56.5 Cost=5 Change=3

(5)、S0148→S0485

L024 S0148 L024 S1487 T1 D02 T1 D15 L156 S2534 L156 S3351 L417 S3351 L417 S0485

Time=89.5 Cost=6 Change=3

(6)、S0087→S3676

L206 S0087 T2 D27 T2 D36 L326 S3676

Time=30.0 Cost=3 Change=0

三、费用最少为第一目标、时间最短为第二目标的路线方案：

(1)、S3359→S1828

L132 S3359 L132 S2903 L485 S2903 L485 S1784 L217 S1784 L217 S1828

Time=67.0 Cost=3 Change=2

(2)、S1557→S0481

L363 S1557 L363 S1919 L189 S1919 L189 S3186 L460 S3186 L460 S0481

Time=109.0 Cost=3 Change=2

(3)、S0971→S0485

L013 S0971 L013 S2517 L290 S2517 L290 S2159 L469 S2159 L469 S0485

Time=106.0 Cost=3 Change=2

(4)、S0008→S0073

L159 S0008 L159 S0400 L474 S2633 L474 S0073

Time=83.0 Cost=2 Change=1

(5)、S0148→S0485

L308 S0148 L308 S3604 L454 S3604 L454 S1919 L417 S1920 L417 S0485

Time=106.0 Cost=3 Change=2

(6)、S0087→S3676

L454 S0087 L231 S0088 L231 S0427 L326 S3676

Time=37.0 Cost=1 Change=0