

基于 GIS 的公交模型数据库构建及维护

张天然

(上海市城市规划设计研究院, 上海 200040)

摘要: 道路网络的修改直接影响公共交通线路数据的同步更新维护, 大大增加了维护的工作量和复杂性。通过分析常用交通规划软件公交模型数据库的特点, 研究基于动态分段技术的公交模型数据库结构。分别设计和开发了公交模型数据库的自动创建以及在不同路网间采用 GIS 方法进行移植的软件工具。探讨公交线路自动创建和移植的功能、算法流程、软件开发, 并应用于大型工程项目。该软件工具的应用实现了公交模型数据和道路网络模型数据独立编辑, 在提高工作效率的同时大大降低了维护模型和数据库的复杂性。

关键词: 公交模型数据库; 动态分段; 公交线路移植

Development and Maintenance of a GIS-based Transit Model Database

Zhang Tianran

(Shanghai Urban Planning & Design Research Institute, Shanghai 200040, China)

Abstract: Alteration in roadway network directly affects the synchronization of transit line data, and therefore significantly increases the workload and complexity of maintenance. By analyzing the characteristics of transit model database for several commonly used transportation planning software packages, this paper studies the structure of the dynamic segmentation-based transit model database. The paper develops software tools for automatic creation and transfer of transit model database among different roadway networks using the GIS. The functionalities, algorithm flowchart, and software development are discussed. Application of such software tools in large projects can separately edit transit model data and roadway network model data, hence significantly improves the efficiency and reducing the complexity for maintaining the models and databases.

Keywords: transit model database; dynamic segmentation; transfer of transit lines

收稿日期: 2013-09-20

作者简介: 张天然(1980—), 男, 博士, 高级工程师, 主要研究方向: 交通规划和政策、交通模型、交通地理信息系统。E-mail: zhangtianrantj@163.com

城市交通模型在 20 世纪 80 年代引入中国, 经过 30 多年的发展, 逐渐成为诊断交通现状症结、预判未来发展趋势、制定交通发展政策、评估交通与土地利用适应性等的重要技术手段。构建和维护交通模型数据库是研发城市交通模型的重要基础性工作。交通网络数据构建工作量大、技术性强、维护周期长, 一直以来是交通模型快速反馈规划方案的瓶颈。交通网络数据主要包括具有拓扑关系的基础道路网、基础轨道交通网、各种交通方式的换乘连线、形心连线、公共交通线路(含轨道交通列车运行交路)和车站等。公共交通系统的数据库与道路交通网络通过特定的数据库结构相关联, 因此道路网

络的修改直接影响公交线路数据的同步更新维护, 这就大大增加了维护的工作量和复杂性。本文研究在地理信息系统(Geographic Information System, GIS)环境下公交模型数据库的创建和维护技术, 并进行相应的软件工程开发, 以提高数据库创建和维护的效率。

1 问题分析

在公交模型数据库的构建和维护工作中, 交通模型师通常会遇到三个难以避免的问题:

- 1) 城市交通管理部门或公交运营企业提供了公交线路(线层)和公交车站(点层)的 GIS 数据, 如何将其快速地转化为交通规划

软件可用的格式。

2) 交通模型包括现状模型和规划模型, 构建现状公交模型数据库后, 如何将依附于现状道路网络的公交线路应用于规划道路网络、使之与规划网络进行数据关联。

一直以来, 国内外交通模型工作者也尝试了很多方法, 采用现状与规划路网“一张图”管理方法(也称为 Master 网络)^[1]是其中应用较为广泛的一种。Master 网络技术在一定程度上减少了公交模型数据库的维护工作量, 但仅适用于规划路网变化不大的情形。由于中国城市发展变化快, 现状和规划路网往往存在较大差异, 很多时候现状路网的一些路段在规划方案中是不存在的, 这些都对数据关联造成影响。同时, 因为 Master 网络还必须包括衔接现状路网和规划路网的拓扑节点, 这无形中又大大增加了路网维护的复杂程度。

3) 局部道路网络的规划方案修改后, 必须同步修改公交线路以达到联动, 这对工作人员的技术要求非常高。如何能够独立修改两个网络, 并重新创建数据关联是需要攻克的难关。

公交模型数据库的维护工作量十分巨大。例如上海市交通模型有 2 400 余条上下行线路、5.5 万多个车站, 纽约市交通模型有 4 000 余条线路、10 万多个车站。在 GIS 技术快速发展的今天, 特别是其在交通领域的广泛应用, 逐渐形成了交通地理信息系统(GIS-T)学科分支, 使得解决上述三个问题具有可能性。

2 公交模型数据库结构描述

2.1 常用交通规划软件的公交模型数据库结构描述

目前, 国际上较为流行的交通规划软件

包括 TransCAD, VISUM, EMME, CUBE 等。这些软件对公交线路的数据库结构描述具有一定差异。TransCAD 和 VISUM 软件允许公交车站位于路段的任意位置, 位于路网节点上是一种特例, 这种数据库结构来源于 GIS 中的动态分段技术^[2-3]。例如, 对公路的设施(事故)管理需要事先存储设施在路径上的所有分布情况。因此, 动态分段技术可以记录路径上的“要素”和“事件”, 而不是修改路网数据本身。EMME 和 CUBE 软件采用路网节点设置车站, 即通过公共汽车走过的节点来描述线路的走向, 车站只能放在道路网络的节点上。这些节点既包括真实的交叉口, 也包括虚拟的节点。为了提高模型的精度, 位于较长路段中间位置的车站需要打断, 以适应这些软件只支持“车站位于路网节点上”的特点。

2.2 基于动态分段技术的公交模型数据库结构描述

基于动态分段技术的公交车站维护, 不再局限于放到路网的交叉口上, 而是放到路段的任意位置, 使数据模型更加贴近真实道路环境。图 1 显示了动态分段下公交车站的布置情况, 其中公交路径段由两个路径片段组成, 路径片段长度分别占各自所属路段的 70% 和 40%。通过动态分段技术, TransCAD 和 VISUM 软件可以精确地描述公交车站的位置和路径段长度等属性。为了兼顾 EMME 和 CUBE 等软件的使用, 可以根据车站的位置打断路段, 添加虚拟节点, 对于十分靠近节点的车站, 可以挪动至节点。

基于动态分段技术的公交模型数据库具有维护方便、建模精度高且更加通用的优势, 是公交模型数据库构建的首选方式。本文在此技术的 TransCAD 软件平台上进行二次开发, 试图解决前文提到的三个主要问题。

首先, 公交网络的最大特点是与路网具有关联性, 在 GIS 中称为路径系统。路径系统记录了公交线路走过的路段和节点顺序、车站所在的位置, 以及线路和车站的属性。路径系统通过特殊的数据结构存储在计算机系统中, 实现各种关联查询分析。公交模型数据库结构见表 1。

1) 线路表: 与公交线路图层关联, 描述公交线路的一些特性, 其主键为 LineID。

2) 路径段表: 描述公交线路中前后两个车站间的各种属性, 属于软件内部的数据

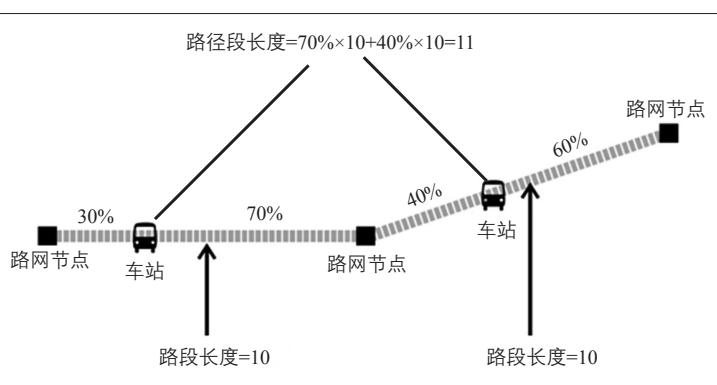


图 1 公交模型数据库中车站的动态分段示意图

Fig.1 Illustration of the dynamic segmentation of stops in the transit model database
资料来源: 文献[4]。

结构，用户无须维护。其主键是TranLinkID，通过LineID字段与线路表关联。

3) 车站表：与公交车站图层关联，其主键为StopID，描述了车站所在的线路(通过LinkID关联)和路段(通过LinkID关联)，表示线路第几次经过该车站(PassCount)以及其他属性。

4) 路径片段表：是路径段表的组成部分，表示路径段在每个路段上的构成，其主键是SectionID。通过TranLinkID与路径段表关联，LinkID表示该片段属于路网上的哪个路段。

5) 路段表和节点表：即道路网络拓扑关系描述。

此外，TransCAD软件还包括可选的数据类型——真实的公交车站。公交线路的车站一定位于公交物理站台上，同一个公交物理站台可能包含多条线路的车站，若某条线路经过某个物理站台不允许上下客，则认为该线路在这个站台没有车站。

3 公交模型数据库自动创建技术

在数据收集阶段，公交线路的车站数据有多种形式。若收集到的只是栅格图的格式(如JPEG格式的图像文件或传统地图)，则模型师只好看着图纸手工添加，或者对着路网写好数据表再由软件导入，这两种方法的速度相差不大，工作量较大。如果能够收集到GIS格式的矢量数据，就会包含线路图层和车站图层(如ArcGIS软件的SHP格式)。另外，也可以通过一些地图网站进行二次开发来获得公交线路的矢量格式数据，例如采用百度地图的API开发或是Google公司的GTFS数据，都可以获得公交线路的空间和属性数据，产生矢量图层。这时，公交模型数据库就可以通过技术研发自动创建。

这些原始的矢量数据，只是简单的车站“点图层”和线路“线图层”，并没有建立车站和路段ID、路径片段和路段ID等关键的关联属性。交通模型中分析公交(轨道交通)自身的换乘关系、与道路网中其他交通方式之间的接驳关系、道路网中固定线路的公交流量计算等各种功能，均需要通过这种特别的公交模型数据库完成。因此，如何利用原始数据快速地创建公交模型数据库十分重要。

3.1 输入数据

需要输入的数据有以下三种：

1) 公交线路：为线层格式，包括线路名字段，是公交线路唯一的名字，需要区分上下行。

2) 车站层：即公交车站图层，要求线路名字段与线路层中的线路名字段对应；车站层还包括车站序号字段(车站图层一条线路中车站的顺序号)。

3) 公交线路依附的道路网络。

3.2 工作原理

TransCAD软件的GISDK开发语言提供了创建公交线路系统的全套函数。根据上述输入数据创建模型数据库，其关键技术如下：

1) 创建公交线路的缓冲区(Buffer)，对道路网络中缓冲区所包含的路段建立子网集合，公交线路经过的路段优先在子网中查找，以约束线路走向。缓冲区越小，约束的精度越高。若路网坐标匹配程度较好，可以适当加大缓冲区，但一般不超过一个街区的宽度。

2) 车站可位于路段任意位置是研究中最大的难点，一些车站位于交叉口附近，加上位置误差和地图匹配问题，仅从空间位置上难以判别车站属于哪一个路段。因此对一条线路的车站序列进行递归前向探索，以确定车站属于公交线路哪一个路段并判断公交线路在路段上的方向。如图2所示，车站图层中的车站3，在地图中更靠近路网图层中的经二路，但实际该车站位于纬三路上，仅从车站3自身的地理位置无法确定其所属路段。如果缓冲区包含了纬三路而未包括经二路，则可以从路段包含关系中排除经二路。但由于地图匹配偏差，有时候线路创建的缓冲区可能未包含纬三路和经二路，虽然车站比较靠近经二路，但是不能确定车站在该路段上。此时，只能递归前向探索车站4和车站5，才能确定车站3所在路段。一般来

表1 公交模型数据库结构

Tab.1 Structure of the transit model database

| 项目 | 线路表 | 路径段表 | 车站表 | 路径片段表 | 路段表 | 节点表 |
|----------------|-----------|------------|------------------|------------|------------|----------|
| | LineID | TranLinkID | LineID StopID | SectionID | LinkID | NodeID |
| 数据 结构 名称 | Frequency | LineID | LinkID | TranLinkID | FromNodeID | LinkFrom |
| | Length | FromStop | PassCount | LinkID | ToNodeID | LinksTo |
| | | ToStop | DwellTime | FromProp | FreeTime | NodeType |
| | | FromMile | FareZone | ToProp | Capacity | |
| | | ToMile | | | | |

说，前向探索2个车站基本足够。

3) 在道路网中，需要对明显不可能的路段进行排除。例如公交车站设置的路段，一般不会是高架道路、隧道和桥梁，可以通过选择集进行限制。例如对于形心连线，其

他模型需要的虚拟连线，也包含在路网图层中，但实际上这些路段是不存在的，可以排除公交线路的通行。再如，轨道交通网络不允许公共汽车通行，可将公共汽车和轨道交通通过不同的路网集合分批创建，应用不同的选择集来排除不合适的路段。

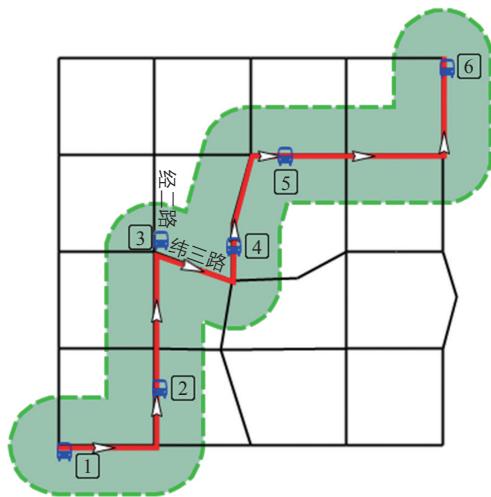


图2 车站所属路段递归分析技术示意

Fig.2 Illustration of recursive analysis of roadway segments with stops



图3 公交模型数据创建校验示例

Fig.3 Examples of transit model data creation and verification



图4 公交模型数据创建

Fig.4 Creation of transit model data

3.3 实际应用中的问题

1) 灵活选择部分线路导入。可以对线路先进行比较严格的条件(缓冲区参数设置)导入，对无法导入或车站错误比较多的线路进行单独选择，放宽条件导入。另外，对于新开行的线路进行添加导入，而不需要完全重新创建。

2) 线路更新和属性数据加载。线路更新主要用于数据的更新，例如公交线路的走向调整。程序可自动删除公交系统中重名的线路，并更新属性数据，自动进行关联加载。

3) 导入数据的检验和修正。公交线路的导入成功率依赖于模型的路网完善程度。例如公交线路经过的地区没有路网，程序可以让线路自动绕行，但若设站的地方没有路网，车站生成只好自动放弃。程序可以自动给出各种错误的报表，以便模型师人工核实少量错误。生成的模型公交线路与原始公交线路的差别检验工作非常重要。附加模块可根据错误报表指定线路名称查询对比，也可逐条浏览。同时，模块还支持批量保存为位图，可用于公交调整线路的查找，便于模型长期维护。公交模型数据创建校验示例见图3。

3.4 应用实例

基于上述工作原理和功能需求，本文开发了相应的软件工具，并应用于具体案例。对上海市公共交通有关部门提供以及某地图网站API开发获得的矢量数据进行数据测试，结果表明上海市2 400余条公交线路可以在1 h左右成功创建，见图4。软件还提供了报错功能(例如车站所在位置的周边地区没有基础道路网络)，错误率在2%以内，可根据错误报告利用检验工具，人工修正出错的少量车站和线路。若按照传统的人工绘制公交线路方式，即使是比较熟练的模型师，也至少需要80个工作日。因此，软件开发大大提高了公交模型数据库的创建效率。

4 公交模型数据库维护技术

使用公交模型数据库创建工具，将公交

模型数据再导出为所开发的软件工具可以输入的数据格式，便可以在新的路网上重新创建公交模型数据库，实现路网的独立编辑。但公交模型数据库包含了更精准的信息，可以开发更为专业化的软件工具实现维护功能。例如上述创建工具有时无法判断公交线路是否通过了重叠的地面和高架道路，正如GPS汽车导航无法完全准确判断车辆在地面道路还是高架道路。而在公交模型数据库中，包含了这些路段和前后关联路段的信息，对于已经建成的公交模型数据库，则可以利用这些信息，进一步提高与路网的匹配度。

4.1 功能特点

模型师往往希望公交系统能从一个网络“移植”到另一个网络，并且不需要数据库字段的对应，只需要两个网络空间位置基本一致。例如将现状道路上的现状公交线路移至规划路网(见图5)，道路网络的编辑脱离公交线路进行自由编辑，多方案路网的公交线路重新生成等，可以大大节省建立公交分配模型时公交线路人工绘制工作。公交线路“搬迁”时应尽量保持原有线路的走向特征，并准确识别公交线路走行高架、地面、地下道路等重叠路段。同时，也可以灵活地选择转换线路集合，或是线路在新网络中可走行的子网集合等，便于选择不同类别的线路进行分批处理。

对比图5a和图5b，规划路网的道路形态有所变化，公交线路则根据新的路网，重新创建了一条与现状线路高度匹配的线路，实现了在现状和规划路网之间的直接移植。实际应用中还需要以下几个功能帮助“移植”：

1) 对既有公交系统，可以创建集合，在“线路集合”界面中选择用户创建的集合，则可以仅移动集合中的公交线路。例如，轨道交通运行交路和公共汽车线路可以分批次地进行移动。

2) 目标公交系统中，可以对路网的节点和路段创建集合，如果使用“排除节点集合”选项，则表示集合中的节点不优先作为线路锚固点，但线路仍旧可能在查找路径时通过集合中的节点。在高级选项设置中勾选“不允许线路通过排除节点集合”选项，则和“排除节点集合”中的节点相连的所有路段都被禁止使用，即线路不能通过集合中的所有节点。

3) 目标公交系统中，如果使用了“排

除路段集合”选项，则表示公交线路不能通过这些路段。“排除路段集合”的应用更为普遍，例如转换公共汽车线路时，可以将形心连线、轨道交通路段等明确公交线路不能通行的路段进行排除。

4.2 工作原理和算法流程

4.2.1 工作原理

移植工具的工作原理与创建公交模型数据工具大体相同，但包含更为详细的参数设置：

1) 线路最短路径缓冲带宽度/m：新路网的路段完全包含在线路缓冲区时，才作为路径选择的对象。移植的公交线路在新的路网中优先在线路缓冲区内的子网中查找路径。

2) 锚固点缓冲带宽度/m：锚固点指新、老路网公交线路走过的节点；“锚固点缓冲带宽度”区域内的新路网节点作为新公交线路的锚固点。

3) 邻近点缓冲带宽度/m：一般大于“锚固点缓冲带宽度”，是指原有路网中公交线路通过的点无法找到锚固点时，用邻近的节点代替。具体查找方法是：在原有路网锚固点缓冲区域内所查找到的路段(去除“排除路段集合”，路段并非完全包含在缓冲区内，接触到的都算)所连节点的最近节点，并且与原有路网中的锚固点的距离小于邻近点缓冲带宽度，即为邻近节点。

4) 绕行路段缓冲带宽度/m：一般大于“线路最短路径缓冲带宽度”，在线路缓冲带内找不到路径时，才到更大的缓冲区间查找路径(仅为局部锚固点对间的绕行查找，并非整条线路扩大缓冲区域)。若仍然找不到路径，则放弃移植该线路。

5) 线路局部段落最大距离差异/m：是指新路网中锚固点或邻近节点对之间的距离

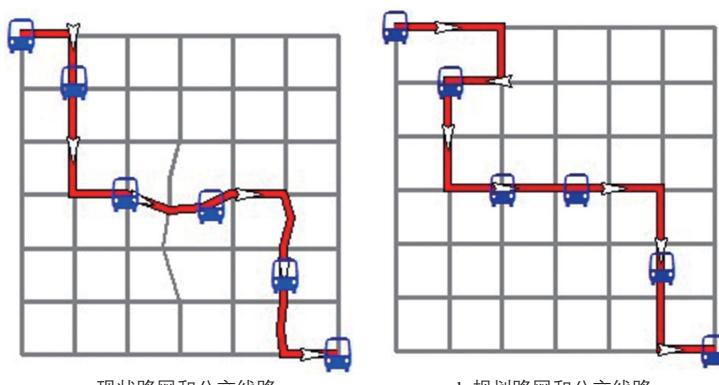


图5 公交模型数据在现状和规划路网上的“移植”

Fig.5 “Transfer” of transit model data between existing and planned roadway networks

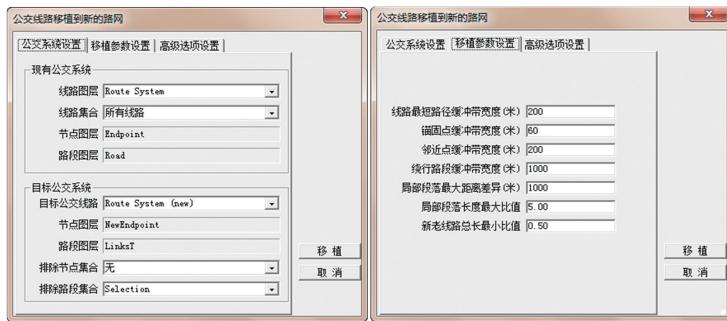


图6 公交模型数据在不同路网上移植的工具用户界面

Fig.6 User interface of tools for transferring transit model data among different roadway networks

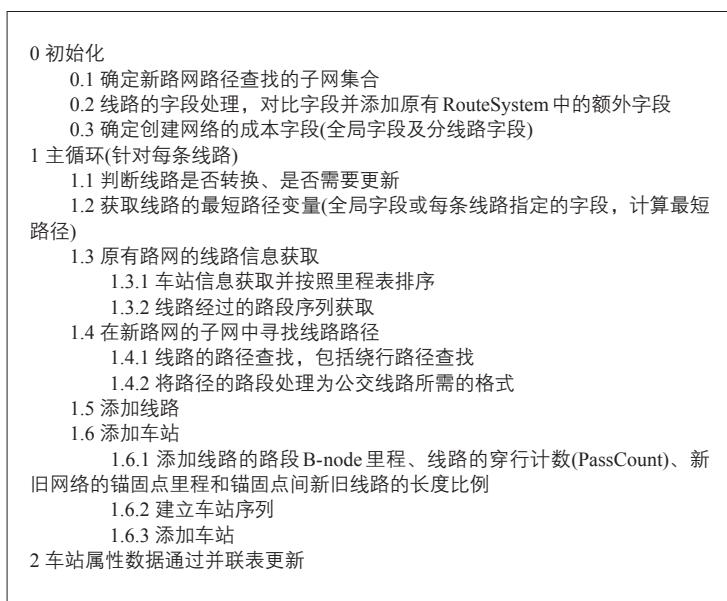


图7 算法流程

Fig.7 Algorithm flowchart

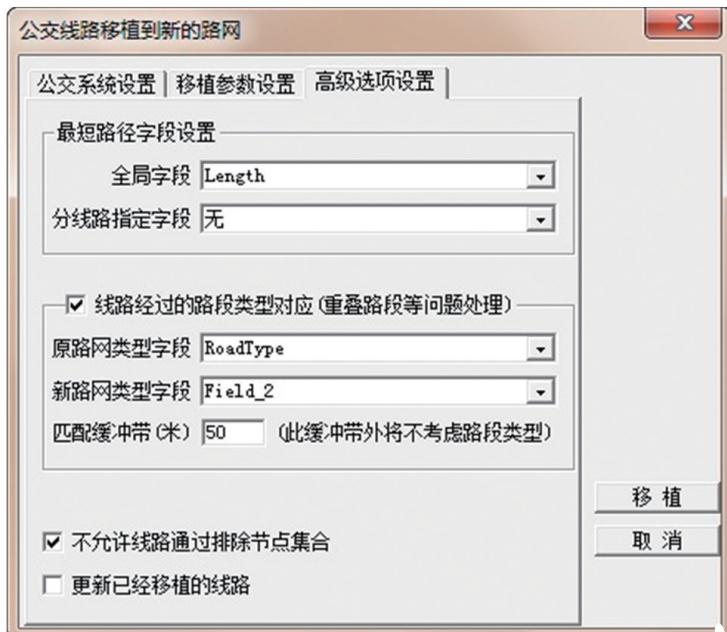


图8 公交模型数据在不同路网上移植的高级功能开发

Fig.8 Development of advanced functionalities for transferring transit model data among different roadway networks

与原有公交线路对应点之间的长度差的绝对值。如果存在某个段落大于这一值，即放弃移植该线路。

6) 线路局部段落长度最大比值：是指新路网中锚固点或邻近节点对之间的距离与原有公交线路对应点之间的长度相对比值。如果存在某个段落大于这一值，即放弃移植该线路。

7) 新旧线路总长最小比值。如果新的线路比原有线路短很多，则放弃移植该线路。

参数集4)~7)可以看作是放弃自动移植的阈值，软件将记录所放弃的线路或线路段，以便进一步手工甄别处理。公交模型数据在不同路网上移植的工具用户界面见图6。

上述参数设置后进行第一次移植，如有失败的线路(一般只在路网改动很大的情况出现)，可以加载新旧两个公交系统，并逐步将参数设置适当放宽，重新对失败的线路进行移植。

4.2.2 算法流程

移植工具的计算主流程如图7所示。

算法流程创建数据库的维护关键技术如下：

1) 通过缓冲区的创建，进行线路走向约束，在车站信息的基础上，提高公交线路在路网上通行的空间范围精度。

2) 采用线路绕行技术，提高线路对不同路网的适应度。线路优先在缓冲区内查找路径，查找失败后则可以在绕行范围内查找，甚至在全网查找。

3) 针对线路重复迂回地经过某些路段，建立 PassCount 字段，从而清晰地描述线路走向及经过路段的计数。

4) 采取分段缓冲区的办法，对公交线路走过的路段类型进行匹配，解决了二维空间的重叠问题，提高了公交线路在地面道路和高架、地下等重叠道路走行路径的识别率。

5) 灵活地应用了选择集合和参数设置，可以提高软件工具的使用灵活度和效率。

4.3 公交模型数据库维护的高级功能

1) 根据不同线路类型设置最短路径字段。最短路径字段是指线路查找路径的成本字段，全局默认字段一般选择为长度，也可以为其他字段或分线路指定字段(原公交系统中建立一个成本字段，则对应查找新路网中的相应字段名作为成本查找最短路径)。

轨道交通和公共汽车的运行线路可以分开设置最短路径字段。

2) 线路经过的路段类型。主要针对平面坐标上重叠的路段进行处理,例如高架道路、地面道路、地下道路等。有些公交线路既走地面道路又走高架道路,不能通过前面排除路段集合区分重叠区域的路段。这种情况可在新旧路网中各设置一个路段类型字段,程序将原公交线路按照经过的道路类型分段绘制缓冲区,每个路段在新路网中查找相应的匹配类型。同时用户需要设置匹配缓冲带,缓冲带内的路段考虑其类型并优先查找路径,缓冲带外的路段将不考虑类型。

3) 有时部分线路移植效果不理想,可根据需要选择更新已经移植成功的线路。

公交模型数据在不同路网上移植的高级功能开发界面见图8。

4.4 公交模型数据库维护的工作效率

根据对上海、天津、广州等多个大城市公交网络数据移植分析,该移植工具工作效率较高,一般可以在1 h左右移植2 000多条公交线路。路网变化不大的情况下,通常准确率在99%以上;路网变化较大时,则可以通过逐次放大移植参数来完成,一般移植成功率约为95%。通过该软件工具的应用,实现了公交模型数据和道路网络模型数据独立编辑,在提高工作效率的同时,大大降低了维护模型和数据库的复杂度。

5 结语

全文介绍了在GIS环境下的公交模型数据库创建和维护技术。通过TransCAD交通

(上接第77页)

- [9] Wang H, Shen S J. Induced Traffic and Forecast Model[J]. Journal of Highway and Transportation Research and Development, 2003, 20(3): 147-150.
- [10] Goodwin P B. Car Ownership and Public Transport Use: Revisiting the Interaction[J]. Transportation, 1993, 20(1): 21-33.
- [11] 国务院. 国务院关于实施成品油价格和税费改革的通知[EB/OL]. 2008[2013-05-15]. http://www.gov.cn/xxgk/pub/govpublic/mrlm/200812/t20081219_33048.html.
- [12] Meyer M D, Miller E J. Urban Transportation

规划软件的二次开发,实现了相关功能并可以应用于实际项目,较大程度地提高了公交模型数据库创建和维护的效率,为交通模型快速反馈多种交通规划方案提供了技术基础。由于TransCAD软件的公交数据结构具有通用性,因此,可以同步开发VISUM,EMME和CUBE等软件的数据接口工具,实现数据共享。

参考文献:

References:

- [1] 贺崇明, 马小毅. 基于GIS的广州市交通规划模型的特点与发展[C]//上海市城市综合交通规划研究所. 城市交通模型技术与应用: 城市交通模型研讨会. 上海: 同济大学出版社, 2007.
- [2] 刘勋, 宣国富, 陈方才. 动态分段技术在公交查询系统中的应用[J]. 北京测绘, 2009(3): 50-52.
Liu Xu, Xuan Guofu, Chen Fangcai. The Application of Section Dynamicly Technique in the Public Traffic Query System[J]. Beijing Surveying and Mapping, 2009(3): 50-52.
- [3] 肖娟, 高光明, 彭涛, 尹超, 林腾. 基于动态分段技术的公交查询系统的设计与实现[J]. 城市勘测, 2010(2): 58-61.
Xiao Juan, Gao Guangming, Peng Tao, Yin Chao, Lin Teng. Design and Realization of Public Transport Inquiry System Based on Dynamic Segmentation Technology[J]. Urban Geotechnical Investigation & Surveying, 2010(2): 58-61.
- [4] Caliper Corporation. TransCAD User's Guide[R]. USA: Caliper Corporation, 2013.

Planning[M]. USA: McGraw-Hill Higher Education, 2001.

[13] 中华人民共和国国家统计局. 国家统计数据库 [DB/OL]. 2012[2013-05-15]. <http://www.stats.gov.cn/>.

[14] Small K A, Van Dender K. Fuel Efficiency and Motor Vehicle Travel: the Declining Rebound Effect[J]. Energy Journal, 2007, 28(1): 25-51.

[15] Litman V T. Generated Traffic and Induced Travel Implications for Transport Planning [J]. ITE Journal, 2009, 71(4): 38-47.