

2017 年深圳市交通仿真系统日常更新

深圳市规划和国土资源委员会（市海洋局）

深圳市规划国土发展研究中心

二〇一八年九月

目 录

第 1 章 数据更新	1
§1.1 动态交通数据处理	1
1.1.1 原始数据采集	1
1.1.2 数据预处理	2
1.1.3 分类集中存储	3
§1.2 静态 GIS 数据处理	4
1.2.1 现有数据类型	4
1.2.2 道路数据更新	5
1.2.3 公交数据更新	12
1.2.4 其他 GIS 数据更新	19
§1.3 年度更新成果	20
1.3.1 动态交通数据	20
1.3.2 静态 GIS 数据	22

插图目录

1.1	分类集中存储后的目录结构	4
1.2	道路数据更新流程	5
1.3	在已知两个节点之间增加路段	6
1.4	在一个已有节点和一个新增节点之间增加路段	6
1.5	小路段的检测与修复	10
1.6	断头路的检测与修复	10
1.7	相连相同路段的检测与修复	10
1.8	道路立体交叉处的连接关系修复	11
1.9	公交数据更新流程	13
1.10	深圳市交委官方网站提供的公交线路一览表	13
1.11	通过商业地图确定新增公交站点位置	14
1.12	公交线路站点顺序描述表	15
1.13	公交线路站点顺序表	16
1.14	公交线路站点描述表图层设置	16
1.15	生成公交线段图层程序界面	17
1.16	2017 年度更新道路 GIS 数据空间分布图	22
1.17	2017 年度更新公交站点 GIS 数据空间分布图	23
1.18	2017 年度更新公交线路 GIS 数据空间分布图	23

表格目录

1.1	数据预处理工作	3
1.2	数据挖掘主机中分类集中存储文件的目录结构	3
1.3	路段图层 Roadlink 属性信息	7
1.4	节点图层 Roadnode 属性信息	8
1.5	数据挖掘专用道路数据结构	12
1.6	公交站点图层 BusStop 属性数据结构	17
1.7	公交线路图层 BusRoute 属性数据结构	18
1.8	公交线段图层 Roadlink 属性数据结构	19
1.9	车牌识别点位图层 Detector 属性	20
1.10	更新后的公交 GPS 数据成果	21
1.11	更新后的深圳通刷卡数据成果	21
1.12	更新后的车牌识别数据成果	21

1 数据更新

数据更新工作在深圳市交通仿真系统（二期）已有的 2016 年静态 GIS 数据和动态交通数据基础上，根据 2017 年各类数据变化情况对底层数据进行更新，将系统的输入数据时间更新至 2017 年。

数据更新工作是后续数据挖掘、数据查询系统更新以及模型系统更新的前提，是一项重要的基础性工作。

§1.1 动态交通数据处理

动态交通数据指的是带有时间信息的交通数据，可以通过数据本身来反映不同时段交通的特征以及变化趋势。其更新频率高，通常 1 分钟就会更新 2-5 次，可以反映交通的实时状态和动态变化，因此称为动态数据。

深圳市交通仿真系统（二期）中囊括的动态交通数据包括出租车 GPS 数据、公交车 GPS 数据、深圳通 IC 刷卡数据（轨道和常规公交）、车牌识别数据四大类。本项目的目标就是，将这四类数据从原始状态开始经过一系列技术加工，最终更新至符合深圳市交通仿真系统（二期）输入格式的标准化数据。

1.1.1 原始数据采集

动态交通数据来源于不同的单位，因此数据采集工作就是按照一定的时间周期从这些单位获取到最新的原始数据。

本项目按季度分别从市交委、市交警局监控中心、东部公交公司、西部公交公司、巴士集团公司、深圳通公司、地铁 ACC 公司通过现场拷贝、网络传输和光盘邮寄几种方式获取所需数据。本项目已采集的原始动态数据至 2017 年 12 月。

- 公交刷卡数据

数据内容	包括每日 350 万次公交刷卡的线路、站点、时间、车牌以及卡号等信息
数据大小	2017 年全年数据大小约 14GB
数据格式	ASCII 文本格式

数据来源	深圳通公司
采集方式	每季度由深圳通公司将当季度数据刻录光盘，再派指定人员取回

- 轨道刷卡数据

数据内容	包括轨道刷卡和投币每日 650 万次的进出站点、时间和卡号信息
数据大小	2017 年全年数据大小约 15GB
数据格式	ASCII 文本格式
数据来源	地铁 ACC 公司
采集方式	每季度由 ACC 公司将当季度数据刻录光盘，再派指定人员取回

- 车牌识别数据

数据内容	包括全市视频监测点和车牌识别系统提供的车牌识别数据
数据大小	2017 年全年数据大小约 480GB
数据格式	ASCII 文本格式
数据来源	市交警监控中心
采集方式	每季度派指定人员去市交警监控中心现场拷贝

- 公交 GPS 数据

数据内容	包括按每 10 秒/回传频率的出租车运行 GPS 数据
数据大小	2017 年全年数据大小约 350GB
数据格式	ASCII 文本格式
数据来源	市交委
采集方式	每季度由专人去现场拷贝

1.1.2 数据预处理

由于原始数据来源于不同的单位，原始数据的格式存在很大的差异，为达到能够统一输入深圳市交通仿真系统（二期）数据平台的要求，需要按照原系统设计的表格结构对数据进行字段顺序调整、精简和合并等预处理工作实现标准化，本项目中针对每类数据的特点，开发了专门的程序进行处理，实现了自动化作业。表1.1是每类数据处理的工作内容。

表 1.1: 数据预处理工作

数据类型	预处理工作
公交刷卡数据	清理无效字符；按照系统表结构调整顺序
轨道刷卡数据	清理无效字符；按照系统表结构调整顺序
车牌识别数据	清理无效字符；合并成每天一个文件；按照系统表结构调整顺序
公交 GPS 数据	清理无效字符和错误数据；合并成每天一个文件；按照系统表结构调整顺序

1.1.3 分类集中存储

预处理后的动态交通数据将通过 FTP 方式上传到数据挖掘主机（配置见附录），按照预先设定的分类目录进行集中存储，作为后续数据挖掘的输入。

预处理动态交通数据每季度上传一次，包含三个月，每月一周（周一至周日 7 天）的数据。在数据挖掘主机中，建立了以下原始数据目录结构，用于存放这些数据。图1.2是存储文件的目录结构。

表 1.2: 数据挖掘主机中分类集中存储文件的目录结构

根目录	二级目录	三级目录	四级目录	
\$DM_HOME	年份	季度	数据类型	文件

其中，\$DM_HOME 指数据挖掘主机存放数据的根目录，默认为 /home/dm/data。在根目录下，每年数据存放在名称为“4 位年份”的二级目录中，例如 2013、2014、2015。二级目录下创建四个三级目录，目录名称为：s+ 季度（1-4），该子目录用于存放某年一个季度的数据。在三季目录下，按不同种类的交通数据分别建立对应的四级目录：

目录 ic 存储 IC 刷卡数据（包括轨道和巴士刷卡数据）

目录 gps 存储巴士 GPS 数据

目录 lp 存储车牌识别数据

在四级目录下，按照每天一个文件的形式存放相应的数据文件。如果原始数据以压缩文件方式提供，在写入对应目录之间，必须先进行解压缩处理，然后才写入目录。图1.1为分类集中存储后的目录结构截图。

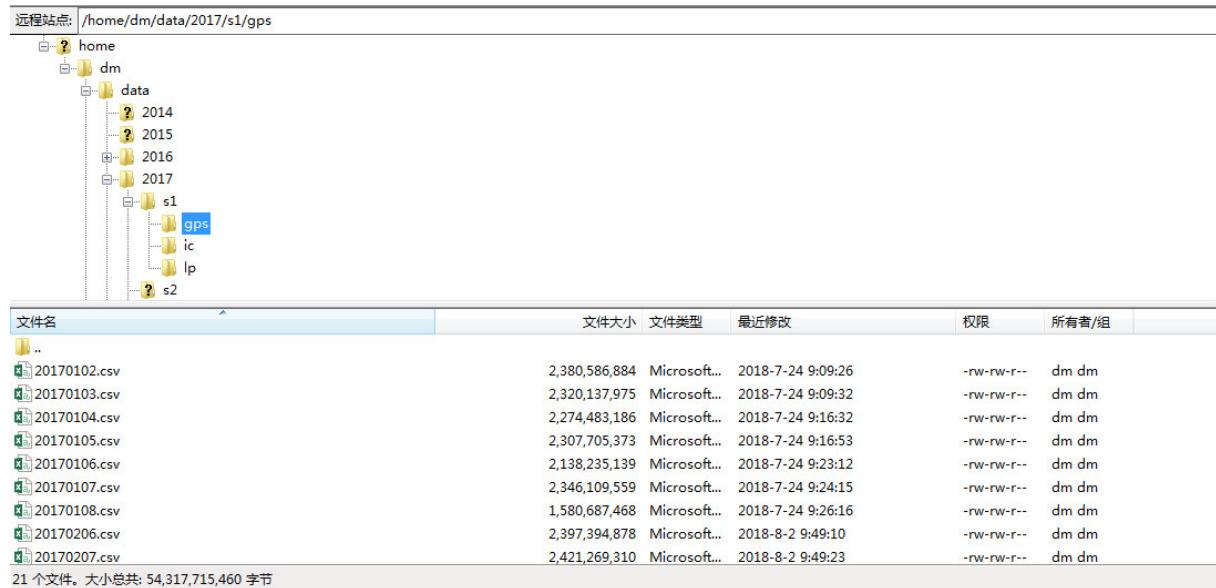


图 1.1: 分类集中存储后的目录结构

§1.2 静态 GIS 数据处理

静态 GIS 数据指用于展示和分析的 GIS 图层数据，包括各类区域边界、道路网络、公交网络站点的形状和属性信息等，主要用于制作各类交通专题图、基础统计以及辅助数据挖掘计算。其变化频率低、更新周期较长，一般为一年更新一次，因此相对于动态交通数据的高更新频率来说属于静态数据。

1.2.1 现有数据类型

深圳市交通仿真系统（二期）中囊括的静态 GIS 数据包括：基础 GIS 数据（行政区、水系、绿地、法定图则、组团、街道），基层路网数据（道路节点、路段、交通小区、主要节点数据），公交和轨道 GIS 数据（公交站点、公交网络、轨道站点、轨道网络），调查数据（居民出行调查、轨道二期开通后调查、交叉口/断面流量调查等），现状和规划土地利用（建筑物普查数据、规划一张图），现状和规划人口岗位数据，POI 数据（医院、学校、口岸、火车站、飞机场、港口、公共停车场、检测器），共 4 大类 26 小类。详细表信息见附表。

其中，道路数据和公交数据涉及后续的数据挖掘系统输入，需要进行年度更新，其他 GIS 数据根据需求不定期进行更新。为了提高工作效率，数据更新工作按照不同 GIS

图层并行开展，更新后的 GIS 数据以 shapefile 格式进行存储，待后续数据挖掘和数据查询系统使用。

1.2.2 道路数据更新

道路数据更新的技术方法是：以市规划国土委（海洋局）信息中心提供的最新版深圳市道路网络数据为底图，按照交通模型和查询的数据结构需求，采用 ArcGIS 软件在深圳市交通仿真系统（二期）已有的 2016 年路网图层（Roadlink）和节点图层（RoadNode）基础上，进行全人工道路形状和属性编辑，将现状道路数据更新至 2017 年 12 月。

更新的流程包括更新道路形状、更新道路节点、更新道路属性、质量检查和更新数据挖掘专用道路数据五个环节（见图 1.2）

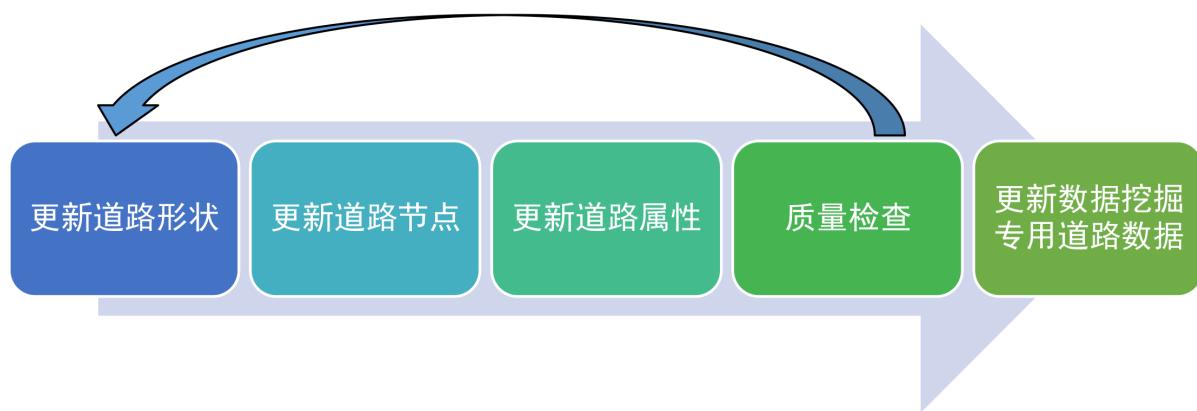


图 1.2：道路数据更新流程

1.2.2.1 更新道路形状

道路形状更新有三种情况：新增路段、删除路段和修改路段。

- **新增路段**

新增路段在年度更新中出现的情况最多，主要工作是把年度新增的道路在 GIS 数据中进行更新，包括以下三种场景：

- (1) 在两个已有节点之间增加路段

在两条道路的交叉口之间新修道路时适用此情况。

操作步骤：打开 ArcMap 的节点捕捉功能，先选择其中一个已有节点；然后再根据信息中心底图数据的形状绘制新增路段，直至另一个已有节点；最后结束形状编辑操作。

如下图所示，在 16258 节点和 16362 节点之间新增路段。

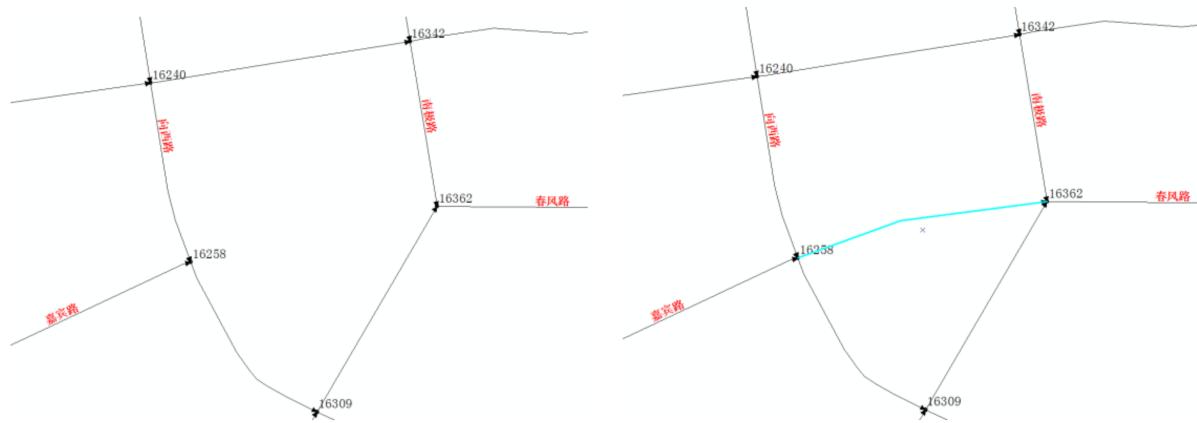


图 1.3: 在已知两个节点之间增加路段

(2) 在一个已有节点和一个新增节点之间增加路段

在一条道路的交叉口与另一条道路新增交叉口之间新修道路时适用此情况。

操作步骤: 首先根据信息中心底图数据的形状, 将已有路段打断; 然后在打断的位置处新增一个节点, 最后按照 (a) 的步骤在已有节点和新增节点绘制新增路段。

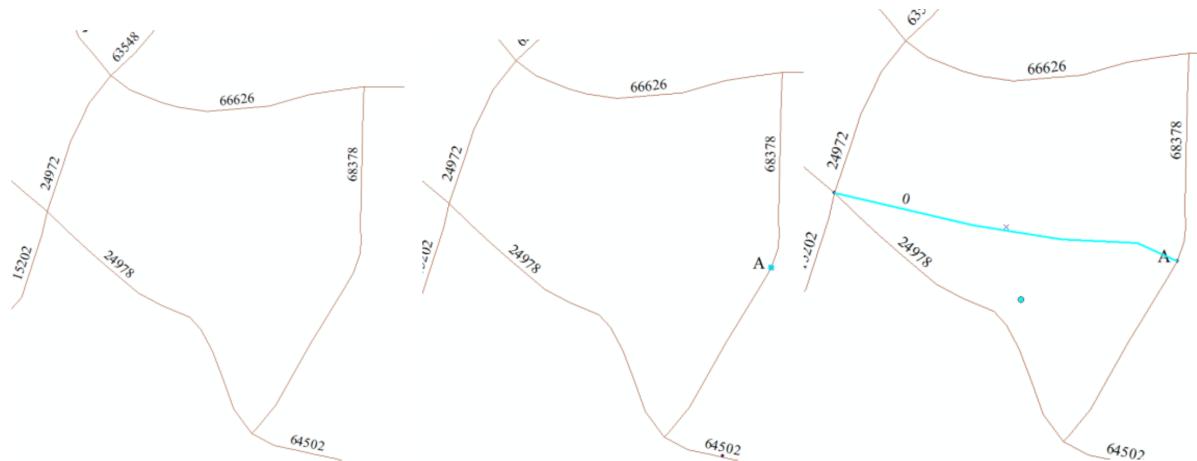


图 1.4: 在一个已有节点和一个新增节点之间增加路段

(3) 在两个新增节点之间增加路段

在两个新增节点之间增加路段的情况较少出现, 适用于在两条道路都新增的交叉口之间新修道路的场景。操作步骤与 (1) 类似, 但是需要在两条道路上都先做打断和新增节点操作。

- **删除路段**

删除路段用于原先绘制的道路错误、道路废弃等情况, 一般较少出现。

操作步骤: 首先选中需要删除的路段, 然后点击删除指令, 最后结束编辑。

- **修改路段**

修改路段主要用于道路大致位置不变, 即首尾节点不变的情况下, 原绘制的路段形状与实际情况存在较大差异时, 需要对道路的形状进行修正。修改路段形状对于道路原有的连接关系没有影响, 不会影响交通仿真模型的运算结果, 但是会影响数据挖掘中 GPS 数据地图匹配的精度。

操作步骤: 选择待修改路段, 打开节点编辑功能, 按照信息中心底图进行形状修改, 最后结束编辑。

1.2.2.2 更新道路节点

道路节点的更新主要是在道路形状发生改变的路段基础上, 新增相应的道路节点, 用于表示道路交叉口; 另外, 还会对原有的道路节点进行删除和修改等修复性工作。按照深圳市交通仿真系统 (二期) 的数据需求, 在所有会发生行驶转向的交叉口都需要有相应的节点数据, 而其他情况的节点都不需要保存。

操作步骤: 打开捕捉节点功能, 按照实际的道路交叉口进行新增节点、删除节点和修改节点的操作, 最后结束编辑。

更新道路节点的工作需要特别注意是高架桥和隧道等区域, 在这些区域由于道路与道路是立体交叉的, 并不会产生行驶转向, 因此不能仅通过图形判断, 必须进行实地踏勘或者从卫星影像图中获得实际的转向信息。

1.2.2.3 更新道路属性

上述两个步骤完成了道路的基本形状: 路段和节点的更新, 但是道路是具有特殊交通属性的线状实体, 因此要在形状更新的基础上进一步更新道路的交通属性信息, 用于描述真实的道路。道路的属性更新工作可以和形状更新同步开展, 也可以在形状更新之后批量修改, 所有的操作都由人工完成。

深圳市交通仿真系统 (二期) 中构建的路段图层 Roadlink 和节点图层 RodeNode 的属性信息如下。

表 1.3: 路段图层 Roadlink 属性信息

属性名称	含义	类型	说明
NO	唯一 ID	Long	
FNO	起节点编号	String	最长 5 位, 与 RodeNode 的 WYID 同

TNO	终节点编号	String	最长 5 位, 与 RoadNode 的 WYID 同
FROM_NAME	起始道路名	String	本图层的 DLMC 字段
TO_NAME	终止道路名	String	本图层的 DLMC 字段
NAME	道路名称	String	中文
CDS	车道数	String	F→T 方向车道数, 数字
LEN	路段长度	Float	单位米
LDKD	路段宽度	String	单位米
F_T	起节点向终节点方向	String	1: 通行; 0: 不通行
T_F	终节点向起节点方向	String	1: 通行; 0: 不通行
DLDJ	道路等级	String	8: 高速公路 7: 快速路 9: 匝道 6: 主干道 4: 次干道 2: 支路 99: 小区道路
SZXZQ	所在行政区	String	行政区图层的 WYID
SZJTXQ	所在交通小区	String	交通小区图层的 WYID
SZJD	所在街道	String	街道图层的 WYID
SZZT	所在组团	String	组团图层的 WYID
SZFDTZ	所在法定图则	String	法定图则图层的 WYID

表 1.4: 节点图层 Roadnode 属性信息

属性名称	含义	类型	说明
WYID	唯一 ID	String	共 8 位, 前两位为主要节点 (Main Node) 的首字母大写 (MN), 后六位与 FID 同, 不足以 0 代替, 如: MN000017
JDMC	节点名称	String	
SZXZQ	所在行政区	String	行政区图层的 WYID
SZJTXQ	所在交通小区	String	交通小区图层的 WYID
SZJD	所在街道	String	街道图层的 WYID
SZZT	所在组团	String	组团图层的 WYID
SZFDTZ	所在法定图则	String	法定图则图层的 WYID
ZBX	坐标 X	Double	
ZBY	坐标 Y	Double	

路段图层包含的属性信息非常多, 其中必须要更新的为路段唯一 ID 字段, 起止节点编号字段, 路段长度字段, 起节点向终节点方向和终节点向起节点方向字段, 道路等级字段, 所在行政区、所在交通小区、所在街道、所在组团、所在法定图则、所属吸引点字段。

其中, 新增加路段的 NO 值, 在原有最大 NO 值的基础上增加 1 或 2, 如为单行线, 则为 1, 如为双行线, 则为 2; 如果删除路段, 原有的 NO 值不再使用, 修改路段不需要修改原有 NO 值。路段长度、起节点向终节点方向和终节点向起节点方向、所在行政区、所在交通小区、所在街道、所在组团、所在法定图则、所属吸引点字段可以在形状更新完成后通过 ArcGIS 自带的数据处理工具自动化辅助更新。

节点图层必须要更新的字段是唯一 ID 和坐标 X、坐标 Y。其中, 新增加节点的 WYID 值, 在原有最大 WYID 值的基础上增加 1; 坐标 X 和坐标 Y 可以在形状更新完成后通过 ArcGIS 自带的数据处理工具自动化辅助更新。

1.2.2.4 质量检查

由于上述的操作主要由人工完成, 所以在操作过程中难免会出现一些错误, 因此需要通过预设的一些规则自动筛选出可能出现错误的地方, 然后再通过人工判断确定错误的位置, 最后按照上述形状和属性更新的步骤来修复错误。为了保证数据的质量, 质量检查工作同时由多人配合完成。

数据质量检查包括操作过程中的日志记录和后期的自动化检查两部分。其中日志记录是今年新引入的一种工作模式, 其原理是在每一步操作中记录操作内容和操作时间, 目的是为了能够在后期检查中发现错误操作时能够进行有效的回退操作, 减少人工排查的工作量并提高修复精度。

自动化检查部分包括: 路段 ID 的唯一性检查、小路段的检测与修复、断头路的检测与修复、相连相同路段的检测与修复以及道路立体交叉处的连接关系修复。

- **路段 ID 的唯一性检查**

路段的 ID 是后续数据挖掘和数据查询的重要条件, 如果在人工作业中出现了重复 ID, 则会导致后续挖掘和数据查询的错误结果, 因此需要通过程序自动化检查路段 ID 是否出现重复。

- **小路段的检测与修复**

小路段是指路段长度小于一定阈值的路段。小路段的存在大大增加了系统负荷, 降低了系统执行效率。因此, 本项目以 50 米为阈值, 对全路网进行检测, 并人工编辑修复

50 米以下路段。

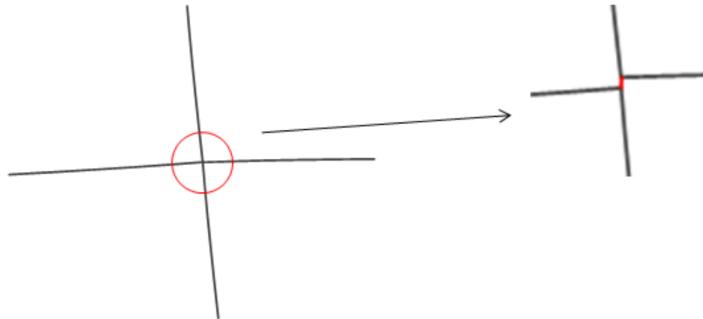


图 1.5: 小路段的检测与修复

- 断头路的检测与修复

断头路是指路段到了终点后，没有后续相连的道路与之连接。断头路分为正常断头路和非正常断头路。非正常断头路事实上是拓扑逻辑存在问题的区域，因此必须予以消除。

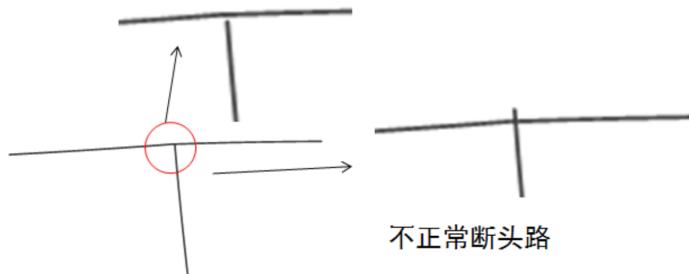


图 1.6: 断头路的检测与修复

- 相连相同路段的检测与修复

相连相同路段是指道路名称 (DLMC)、车道数 (CDS)、道路等级 (DLDJ)、分割类型 (CGLX) 等四个关键属性完全一致的道路，这种类型的道路事实上可以进行合并管理



图 1.7: 相连相同路段的检测与修复

- 道路立体交叉处的连接关系修复

对于匝道、立交桥、隧道等立体交叉位置，采用卫星影像辅助人工判断道路的实际连接关系。

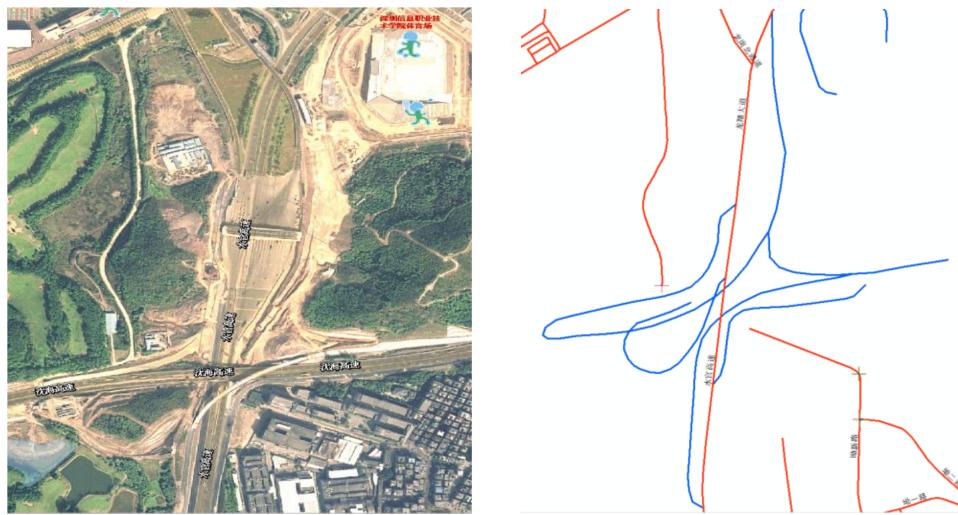


图 1.8：道路立体交叉处的连接关系修复

1.2.2.5 更新数据挖掘专用道路数据

经过前面的道路更新工作后，道路的形状和属性数据都已经更新至 2017 年，可以用于后续的数据查询系统和交通模型系统更新。但是，数据挖掘中的路段车速计算需要对原有的道路数据进行进一步的处理，使其符合数据挖掘系统的输入需要。经过处理后的数据包括：数据挖掘专用路段图层 GIS 数据、地图索引文件和路段编号与小区编号对应文件。具体处理操作如下：

- **坐标转换**

但是，由于数据挖掘系统中的路段车速计算需要将车辆 GPS 数据与道路数据进行地图匹配，而上述所有操作都是在深圳本地坐标系下，和 GPS 数据的 WGS84 坐标不匹配。因此，需要首先通过 ArcGIS 空间纠正工具将原有道路形状纠正到 WGS84 坐标系下。

- **路段图层属性转换**

另外，深圳市交通仿真系统（二期）中采用的路段车速计算程序对道路网络的输入有一套自定义的规则，与表 4 中的字段有差别，所以需要对 RoadLink 图层的字段进行转换处理，使其符合路段车速计算程序的输入要求。下表为路段车速计算程序输入的路段图层字段要求，以及与表 1.3 的转换关系。

表 1.5: 数据挖掘专用道路数据结构

字段名	类型	含义	取值范围	与 Roadlink 图层对应关系
EDGEID	Int(10)	路段编号		NO
FJCID	Int(10)	起点编号		FNO
TJCID	Int(10)	终点编号		TNO
LENGTH	Float	路段长度	单位米	LEN
BYNAMEC	Char(20)	道路名称		NAME
DF	Int(2)	交通流方向	1: 双向可通行 2: 与矢量方向相同 3: 与矢量方向相反 4: 双向不通行	F_T,T_F
NR	Int(2)	管理等级	1: 高速路 2: 国道 3: 快速路 4: 省道 5: 主干道 6: 次干道 7: 支路 8: 支路以下	DLDJ

1.2.3 公交数据更新

公交数据更新的技术方法是: 基于交委发布的现状公交站点名称数据, 并结合商业地图中的现状公交站点位置数据, 以深圳市交通仿真系统 (二期) 原有的公交站点图层 (BusStop) 为基础, 采用半自动化方式进行新增和修改操作; 然后再通过专门开发的程序, 按照站点位置和线路站点顺序, 重新生成公交线路, 直接更换原有数据。

更新的流程包括更新公交站点位置和名称、公交网检查、生成公交线路和更新线路和站点属性四个环节。

1.2.3.1 更新公交站点位置和名称

深圳市交委每个月都会在官方网站上提供最新的公交线路和途经站点文件, 文件主要包括线路编号、起点站、终点站、上行途经站点、下行途经站点、上行途经道路和下行途经道路字段。



图 1.9: 公交数据更新流程

但是, 文件中途经站点只有中文名称, 并没有位置数据。因此, 首先通过程序在文件中自动化搜索出在原有文件中没有出现过的站点名称, 然后再通过商业地图, 通过人工方式依次在 BusStop 图层中新增公交站点位置。其中, 在进行新增公交站点的操作过程中, 也同时对 WYID 字段进行更新。WYID 字段是按“BS” + 五位数字规则定义的, 新增的公交站点在原有最大五位数字的基础上增加 1。

深圳市公交线路一览表 (11月份)					
业户名称	线路编号	起点站	终点站	批文上行途经站点	批文下行途经站点
深圳巴士集团股份有限公司	65	香蜜湖总站	东方尊峪总站	香蜜红荔立交、香蜜湖北、市政大厦②、鲁班大厦、莲花山西②、莲花山公园、少年宫、关山月美术馆②、人才大厦、中心公园、振华西路、中航地产、振中路、红岭中学、森园小学、少儿图书馆、红桂大厦、天地大厦、西湖宾馆、桂园中学、笋岗桥、儿童公园①、木头龙、华丽环岛、荷花市场、峰景台②、罗湖体育馆、广岭、西岭下、聚宝路口、莲塘、长岭	长岭沟、长岭、莲塘、聚宝路口、仙湖路口、西城下、广岭、罗湖体育馆、峰景台②、荷花市场、华丽环岛、木头龙、儿童公园①、笋岗桥、宝尚路南、笋岗仓库、西湖宾馆、红桂大厦、少儿图书馆、森园小学、建艺大厦②、华发大厦②、华强路口、华富路口、振华西路、中心公园、人才大厦、武警支队、关山月美术馆②、少年宫、莲花山公园①、莲花山西②、鲁班大厦②、市政大厦、香蜜二村、香梅深南路口、香蜜湖②、深南香蜜立交②
深圳巴士集团股份有限公司	M300	深圳北站	平湖总站	樟坑、潜龙花园、石龙路口、滢水山庄、万家灯火、星光之约①、家和花园、南坑村、雅园路口、中海月朗苑、五园小学、珠江旭景佳园、富豪花园、杨美村、宏扬学校、丽湖花园、中海怡翠、三联桥头、布吉汽车站、大芬地铁站②、布澜路口、康桥紫郡、布吉交警中队、联创科技园、第三人民医院、李朗国际珠宝园、第三人民医院、联创科技园、布吉交警中队、康桥紫郡、布澜路口、桂芳园、大芬地铁站②、布吉汽车站、中海怡翠、丽湖花园、宏扬学校、杨美村、富豪花园、珠江旭景佳园、万科五园幼儿园、中海月朗苑、南坑村、家和花园、星光之约①、万家灯火、滢水山庄②、惠鑫公寓、石龙路口、滢龙花园	平湖广场西、平湖派出所、平湖中心小学、华南城、华南城西门、上木古村、华美嘉配送中心、金积嘉科技园、左右家私、宝福珠宝园、李朗国际珠宝园、第三人民医院、联创科技园、布吉交警中队、康桥紫郡、布澜路口、桂芳园、大芬地铁站②、布吉汽车站、中海怡翠、丽湖花园、宏扬学校、杨美村、富豪花园、珠江旭景佳园、万科五园幼儿园、中海月朗苑、南坑村、家和花园、星光之约①、万家灯火、滢水山庄②、惠鑫公寓、石龙路口、滢龙花园
深圳巴士集团股份有限公司	B607	松坪山总站	深大地铁站	松坪山路口、朗山路中、朗山路口、科苑立交、长园新村、万和制药、万基、赛百诺、科技中一路、帝景园、凯丽花园、维用科技、深南科苑立交	高科技创新中心、维用科技、凯丽花园、帝景园、科技园中一路、赛百诺、万基、万和制药、长园新村、科苑立交、朗山路口、朗山路中、松坪山路口
深圳巴士集团股份有限公司	B727	月亮湾公交总站	南山中心区总站	月亮湾花园三号门、月亮湾花园、荔湾路西、南山花园、深大附中、青青世界、前海沿湖路口、太子山庄、信诺公司、南山党校、城市山林西、城市山林东、东滨路中、东滨路口、南油大厦、南山邮局、佳嘉豪苑、创业东、观海台、后海地铁站、臻日湾畔	后海地铁站、观海台、创业东、佳嘉豪苑、南山邮局、南油大厦、东滨路口、东滨路中、城市山林东、城市山林西、南山党校、信诺公司、太子山庄、前海沿湖路口、青青世界、深大附中、南山花园、荔湾路西、月亮湾花园、月亮湾花园三号门

图 1.10: 深圳市交委官方网站提供的公交线路一览表

更新后的公交站点图层导出为一个文本文件, 包含公交站点唯一 ID (WYID)、站点名称 (ZDMC)、站点 X 坐标 (ZBX) 和站点 Y 坐标 (ZBY) 四个字段。其中需要特别注意的是:

☞ WYID 字段不能出现重复, 且每条记录非空



图 1.11：通过商业地图确定新增公交站点位置

- ⇒ ZDMC 字段建议删除“?”等非法字符
- ⇒ ZBX、ZBY 为深圳本地坐标系, 且每条记录非空
- ⇒ 合法符号转换, 建议将合法的符号全部转换为全角字符

1.2.3.2 更新公交线路数据

根据图1.10的公交线路一览表以及上节生成的公交站点文本文件, 对公交线路途经站点按照中文名称与编号的对应关系进行标准化, 生成公交线路站点顺序描述表, 包括线路和途经站点两个字段, 中间用英文符号“,”分隔; 途经站点之间用中文符号“、”分隔。这个步骤可以编写程序自动化完成。

另外, 还需要按照线路途经的每个站点一条记录的形式生成公交线路站点描述表文件, 包括线路 ID(RouteID)、公交站点 ID(StopID)、公交站点名称 (StopName) 和公交站点坐标 (X 和 Y) 五个字段, 中间以英文符号“,”分隔。

1.2.3.3 生成公交线路 GIS 数据

按照深圳市交通仿真系统 (二期) 的设计, 公交线路 GIS 数据包括公交线路图层 (BusRoute) 和公交线段图层 (BusLink)。公交线路图层是按照每条公交线路一个完整 polyline 实体进行存储的, 而公交线段图层是对公交线路图层在每个公交站点处打断生成的, 一条公交线路中的每两个连续公交站点之间形成一个 polyline 实体。因此, 首先生成公交线路图层, 再生成公交线段图层。

1-up, BS008634、BS005408、BS005409、BS005340、BS005334、BS005319、BS005312、
 BS005283、BS005268、BS005265、BS004870、BS004864、BS004841、BS010037
 1-down, BS009297、BS004859、BS004865、BS004867、BS005377、BS005282、
 BS005280、BS005254、BS005315、BS005318、BS005338、BS005339、BS005407、
 BS008634
 高峰专线1号-up, BS009501、BS005453、BS005261、BS005246、BS005242、BS009371、
 BS009527
 高峰专线1号-down, BS009527、BS005161、BS005243、BS005260、BS005262、
 BS008615、BS005348、BS009501
 高快巴士1号-up, BS008960、BS009084
 福田保税区1号-up, BS010734、BS004641、BS004642、BS004636、BS004637、
 BS004635、BS004675、BS004677、BS011184
 福田保税区1号-down, BS011183、BS004707、BS004676、BS004629、BS009075
 观光线(观光1线)-down, BS010292、BS005597、BS006403、BS005608、BS005548、
 BS005544、BS005530、BS005517、BS005493、BS005457、BS005460、BS007846、
 BS005380、BS005374、BS005287、BS005274、BS005217、BS004855、BS005134、
 BS004813、BS004797、BS004786、BS004802、BS005056、BS008496、BS008494、
 BS008484、BS005929、BS005925、BS005908、BS005820、BS009433、BS000255、
 BS000256、BS000257、BS006562、BS000258、BS000287、BS000136

图 1.12: 公交线路站点顺序描述表

• 生成公交线路图层

生成公交线路图层工作利用 ArcGIS Network Analysis 模块进行自动化完成。具体步骤如下:

- (1) 将上节得到的公交线路站点描述表文件导入 ArcMap, 使用“add xy data”功能转换为 shapefile 点文件, 坐标为 ZBX 和 ZBY;
- (2) 在 ArcCatalog 中, 选择更新后的路段 GIS 文件 RoadLink, 选择“Create Network dataset”功能创建路网的网络数据集;
- (3) 在 ArcMap 中加载上一步中创建的网络数据集和 (a) 中的公交线路站点描述表图层, 打开 Network Analysis 模块, 新增一个线路工程 (add new route);
- (4) 采用 ArcToolbox→Network analysis 工具 → 添加位置功能, 输入相应的数据; 其中, “输入网络分析图层”中输入 (b) 中创建的网络数据集, “子图层”输入 (a) 中的公交线路站点描述表图层, “字段映射”的 RouteName 输入公交线路站点描述表图层中的线路编号字段 RouteID;
- (5) 最后, 使用 Network Analysis 模块中的 Solve 功能, 自动生成线路。

生成后的公交线路图层命名为 BusRoute。需要特别注意的是, BusRoute 图层的唯一 ID 字段需要和深圳市交通仿真系统(二期)中使用的 BusRoute 图层的唯一 ID 字段, 根据线路名称字段进行连接操作 (join), 保证唯一 ID 在不同年份的连续性和唯一性, 使得在公交线路相关指标历史查询的正确性; 当线路唯一 ID 无法与原有数据 join 时 (无法 join 说明此线路为 2017 年新增线路), 唯一 ID 按照最大 ID 值增加 1 来取值。

• 生成公交线段图层

1.txt - 记事本					
文件(F)	编辑(E)	格式(O)	查看(V)	帮助(H)	×
RouteID,StopID,StopName,X,Y					
1-up, BS005459, 太宁小学, 123110.630484, 22290.704141					
1-up, BS005460, 百仕达东郡, 122972.465338, 22375.2310546					
1-up, BS005391, 百仕达花园, 122702.039185, 22416.9412002					
1-up, BS005389, 五方广场, 122382.81811, 22275.4935433					
1-up, BS005369, 田贝四路, 121756.87246, 22067.7243323					
1-up, BS005367, 田贝②, 121439.674645, 21738.2098278					
1-up, BS005334, 东门中②, 121502.203204, 20446.9543822					
1-up, BS005319, 东门①, 121270.766188, 19800.7406726					
1-up, BS005316, 金城大厦②, 121322.876257, 19449.0352001					
1-up, BS004921, 金光华广场, 121040.071255, 19157.0658179					
1-up, BS004915, 罗湖小学, 120959.005187, 19003.8948374					
1-up, BS004891, 京湖酒店, 120953.206423, 18609.0566054					
1-down, BS004891, 京湖酒店, 120953.206423, 18609.0566054					
1-down, BS004915, 罗湖小学, 120959.005187, 19003.8948374					
1-down, BS004917, 国贸, 120960.977631, 19236.6583015					
1-down, BS005319, 东门①, 121270.766188, 19800.7406726					
1-down, BS005334, 东门中②, 121502.203204, 20446.9543822					
1-down, BS005331, 儿童公园①, 121288.054598, 20973.5198361					
1-down, BS005304, 洪湖沃尔玛, 121061.108525, 21322.8765076					
1-down, BS005366, 田贝新村, 121225.643235, 21716.8951903					
1-down, BS005369, 田贝四路, 121756.87246, 22067.7243323					
1-down, BS005389, 五方广场, 122382.81811, 22275.4935433					
1-down, BS005391, 百仕达花园, 122702.039185, 22416.9412002					
1-down, BS005458, 百仕达东郡, 122934.461709, 22359.6778618					
2-up, BS005464, 布心花园, 123224.890104, 22891.9966784					
2-up, BS005413, 市慢病防治中心, 122659.472891, 23072.8288802					
2-up, BS007816, 君逸华府①, 122261.732283, 23013.063328					

图 1.13: 公交线路站点顺序表

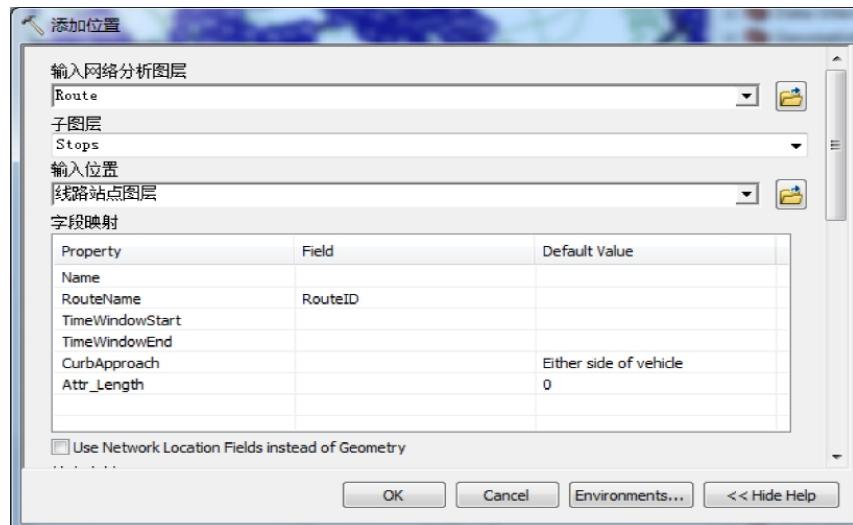


图 1.14: 公交线路站点描述表图层设置

由上述公交线路图层生成公交线段图层的工作由基于 ArcMap 组件开发的程序自动完成, 程序的主界面如下:

操作步骤如下:

- (1) “公交站点” 输入 1.2.3.1 节更新后的公交站点图层 BusStop;
- (2) “公交线路” 输入 1.2.3.2 节生成的公交线路图层 BusRoute;
- (3) “公交线路站点顺序描述表” 输入 2.2.3.2 节生成的公交线路站点顺序描述表;

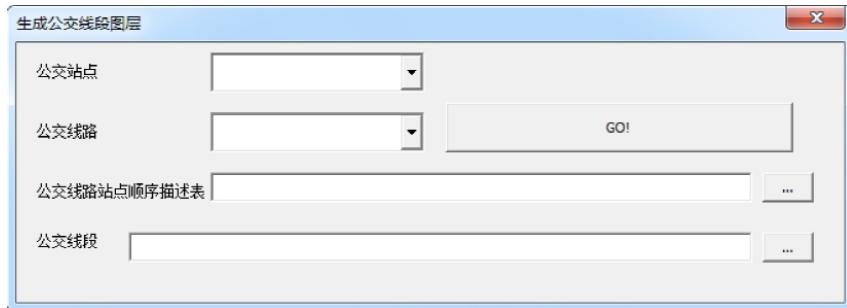


图 1.15: 生成公交线段图层程序界面

- (4) “公交线段”输入需要导出公交线段图层的路径位置;
- (5) 点击“GO!”按钮进行运算, 运算结束后会有对话框提示。

生成后的公交线路图层命名为 BusLink。另外, 运算程序不仅会自动打断并生成公交线段, 而且会在运算过程中将 BusRoute 图层的起止公交站点编号字段 (QDZH 和 ZDZH) 自动更新。

1.2.3.4 更新线路和站点属性

上述几节的工作生成了公交站点图层 BusStop、公交线路图层 BusRoute 和公交线段图层 BusLink 的 GIS 图形数据以及部分属性数据, 本节需要继续对三个图层的剩余属性数据进行补全。

- 公交站点 (BusStop) 图层属性更新

BusStop 图层的属性信息如表 1.6 所示。

表 1.6: 公交站点图层 BusStop 属性数据结构

属性名称	含义	类型	说明
WYID	唯一 ID	String	BS+5 位数字, 例如 BS00001
ZDMC	站点名称	String	
ZBX	坐标 X	Double	
ZBY	坐标 Y	Double	
SZXZQ	所在行政区	String	行政区图层的 WYID
SZJTXQ	所在交通小区	String	交通小区图层的 WYID
SZJD	所在街道	String	街道图层的 WYID
SZZT	所在组团	String	组团图层的 WYID
SZFDTZ	所在法定图则	String	法定图则图层的 WYID
SZLDBH	所在路段编号	String	即路段图层 RoadLink 的 NO 字段
GNGW	关内关外	String	1: 关内; 2: 关外

SSXYD	所属吸引点	String	
-------	-------	--------	--

其中，唯一 ID (WYID)、站点名称 (ZDMC)、坐标 (ZBX 和 ZBY) 字段在[1.2.3.1](#)节的更新工作中已经完成，剩余的所在行政区、所在交通小区、所在街道、所在组团、所在法定图则、所在路段编号、关内关外、所属吸引点字段可以通过将 BusStop 图层与相应图层进行简单的空间关系判断来补全，全部操作都使用 ArcMap 的空间连接功能完成。

- **公交线路 (BusRoute) 图层属性更新**

BusRoute 图层的属性信息如表[1.7](#)所示。

表 1.7：公交线路图层 BusRoute 属性数据结构

属性名称	含义	类型	说明
ID	唯一 ID	String	BR+5 位数字, 例如 BR00001
NAME	线路名称	String	由 GJXLMC 和上下行关键字组成, 如 1-up 表示 1 路上行
GJXLMC	公交线路名称	String	如 M241, 115
SXX	上下行	String	0: 上行;1: 下行
XLCD	线路长度	Double	单位米
YXGS	运行公司	String	
QDZM	起点站名	String	与 BusStop 图层的 ZDMC 字段相同
QDZH	起点站号	String	与 BusStop 图层的 WYID 字段相同
ZDZM	终点站名	String	与 BusStop 图层的 ZDMC 字段相同根据 QDZM 和 ZDZM 就可以区分上下行
ZDZH	终点站号	String	与 BusStop 图层的 WYID 字段相同根据 QDZH 和 ZDZH 就可以区分上下行
YTCZ	沿途车站	String	首末车站的中间车站的 WYID, 按顺序输入, 由逗号分隔
GNGW	关内关外	Integer	1: 全部关内;2: 全部关外;3: 关内关外

其中，唯一 ID (WYID)、线路名称 (NAME) 字段、起点站号 (QDZH) 和终点站号 (ZDZH) 字段在[1.2.3.2](#)节的更新工作中已经完成；公交线路名称 (GJXLMC) 字段和上下行 (SXX) 字段可以根据线路名称字段进行简单的字符分隔处理自动化获得；线路长度 (XLCD) 字段在图形编辑完成后可以由 ArcMap 自动计算获得；起点站名 (QDZM) 和终点站名 (ZDZM) 字段可以采用 ArcMap 中的连接功能 (join)，按照起

点站号（QDZH）和终点站号（ZDZH）字段和公交线路站点描述表中的公交站点编号和公交站点名称对应关系自动补全；关内外（GNGW）字段可以通过将 BusRoute 图层与关内外区域图层进行简单的空间关系判断来补全；运行公司（YXGS）字段，可以根据实际信息补全；沿途车站（YTCZ）字段由于 shapefile 文件字段不得超过 255 个字符的限制，因此不直接在 shapefile 文件更新，而是将 shapefile 文件的属性导出为 csv 文件后，再由线路名称字段与公交线路站点顺序描述表进行匹配后自动补全。

• 公交线段（BusLink）图层属性更新

BusLink 图层的属性信息如表1.8所示。

表 1.8: 公交线段图层 Roadlink 属性数据结构

属性名称	含义	类型	说明
WYID	唯一 ID	String	共 8 位, 前两位为公交线段 (Bus Link) 的首字母大写 (BL), 后六位与 FID 同, 不足以 0 代替, 如:BL000017
GJXLMC	公交线路名称	String	与 BusRoute 图层的 GJXLMC 相同
SXX	上下行	String	0: 上行;1: 下行
NAME	公交线路	String	与 BusRoute 图层的 NAME 字段相同
XLCD	线路长度	Double	该段路线的长度, 单位米
YXGS	运行公司	String	
QDZM	起点站名	String	与 BusRoute 图层的 QDZM 相同
QDZH	起点站号	String	与 BusRoute 图层的 QDZH 相同
ZDZM	终点站名	String	与 BusRoute 图层的 ZDZM 相同
ZDZH	终点站号	String	与 BusRoute 图层的 ZDZH 相同

BusLink 图层的属性与 BusRoute 图层完全相同，除了唯一 ID 字段需要在线路打断处理后重新生成。

1.2.4 其他 GIS 数据更新

根据 1.2.1 节显示的深圳市交通仿真系统（二期）所有静态 GIS 数据，除了需要进行年度更新的道路数据和公交数据之外，其他的 GIS 数据可以根据年度的实际情况进行更新，更新的范围也包括图形和属性更新两部分。例如，某些区域图层的边界发生变化，或者新开通了轨道线路和站点等情况。2017 年由于交警新增了一部分车牌识别监测点，因此本项目也同步更新了点位的 GIS 图层（Detector）。其主要属性如表1.9所示。

表 1.9: 车牌识别点位图层 Detector 属性

属性名称	含义	类型	说明
WYID	唯一 ID	String	两种编码方式: 1. 前两位为检测的拼音首字母大写 (JC), 后六位与 FID 同, 不足以 0 代替, 如:JC000017; 2. 根据规土委给的测点 ID(8 位数字)
JCQLX	检测器类型	String	视频、流量、牌照
WZMS	位置描述	String	检测器所在的道路、交叉口等位置描述, 从 CAD 文件中输入 JCFX 检测方向 String 东行、西行、南行、北行、出、入还有部分为空, 代表没有明确指定检测方向
ZBX	坐标 X	Double	
ZBY	坐标 Y	Double	
SZXZQ	所在行政区	String	行政区图层的 WYID
SZJTXQ	所在交通小区	String	交通小区图层的 WYID
SZJD	所在街道	String	街道图层的 WYID
SZJT	所在组团	String	组团图层的 WYID
SZFDTZ	所在法定图则	String	法定图则图层的 WYID
SZLD	所在路段	String	路段图层的 NO 字段 (以 50m 为半径判断所在道路)
JCFXO	方向逻辑描述	Integer	文字 JCFX 字段 (如南行、北行), 与所在道路的 FNO、TNO 字段在空间位置上的一致性。相同为 1, 相反为 0

§1.3 年度更新成果

1.3.1 动态交通数据

按照1.1节进行预处理后的动态交通数据, 通过 FTP 方式上传到数据挖掘主机上, 按照1.1.3节的文件夹结构进行统一存储, 下面列出更新后的文件夹包含文件的信息。

表 1.10: 更新后的公交 GPS 数据成果

路径	第一季度:/home/dm/data/2017/s1/gps 第二季度:/home/dm/data/2017/s2/gps 第三季度:/home/dm/data/2017/s3/gps 第四季度:/home/dm/data/2017/s4/gps
文件名	日期.csv, 如 20170101.csv, 共 84 个
说明	标准化预处理并以日为单位合并后的公交 GPS 数据, 每天一个文件存储
大小	

表 1.11: 更新后的深圳通刷卡数据成果

路径	第一季度:/home/dm/data/2017/s1/ic 第二季度:/home/dm/data/2017/s2/ic 第三季度:/home/dm/data/2017/s3/ic 第四季度:/home/dm/data/2017/s4/ic
文件名	ic_bus_ 月份.csv (公交刷卡数据)、ic_metro_ 月份.csv (轨道刷卡数据), 如 ic_bus_01.csv、ic_metro_12.csv, 共 24 个
说明	标准化预处理并以日为单位合并后的公交 GPS 数据, 每天一个文件存储
大小	

表 1.12: 更新后的车牌识别数据成果

路径	第一季度:/home/dm/data/2017/s1/lp 第二季度:/home/dm/data/2017/s2/lp 第三季度:/home/dm/data/2017/s3/lp 第四季度:/home/dm/data/2017/s4/lp
文件名	sbbak+ 日期.txt, 如 sbbak2017-01-06.txt, 共 84 个
说明	标准化预处理并以日为单位合并后的车牌识别数据, 每天一个文件存储
大小	

1.3.2 静态 GIS 数据

本年度已核查全市 40142 条路段，28101 个节点，修正和合并部分次干路和支路并删除部分支路和小区道路共计 45 条路段和 72 个节点。下图为年度更新道路 GIS 数据的空间分布情况。



图 1.16: 2017 年度更新道路 GIS 数据空间分布图 (红色为年度更新路段)

更新前的深圳市交通仿真系统 (二期) 公交站点 GIS 数据共有个 11615 站点, 本年度更新后共有 12664 个。图1.17为年度更新公交站点 GIS 数据的空间分布情况。

更新前的深圳市交通仿真系统 (二期) 公交线路 GIS 数据共有 1877 条 (包括上下行), 本年度更新后共有 2658 条 (包括上下行)。图1.18为年度更新公交线路 GIS 数据的空间分布情况。

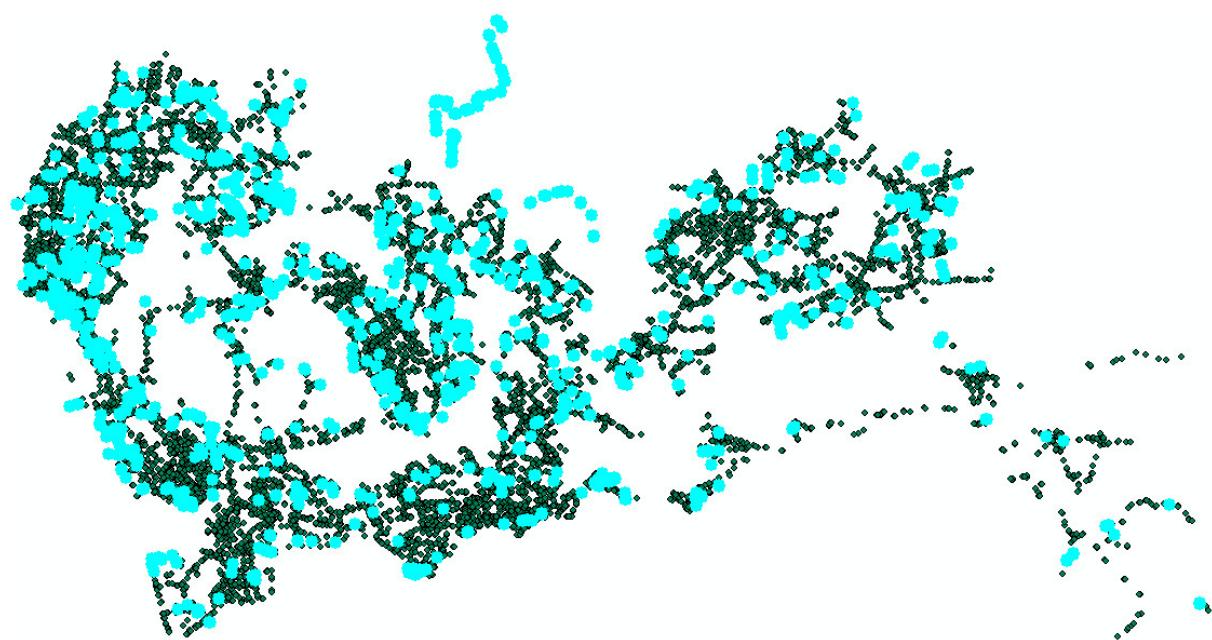


图 1.17: 2017 年度更新公交站点 GIS 数据空间分布图 (蓝色为年度更新站点)

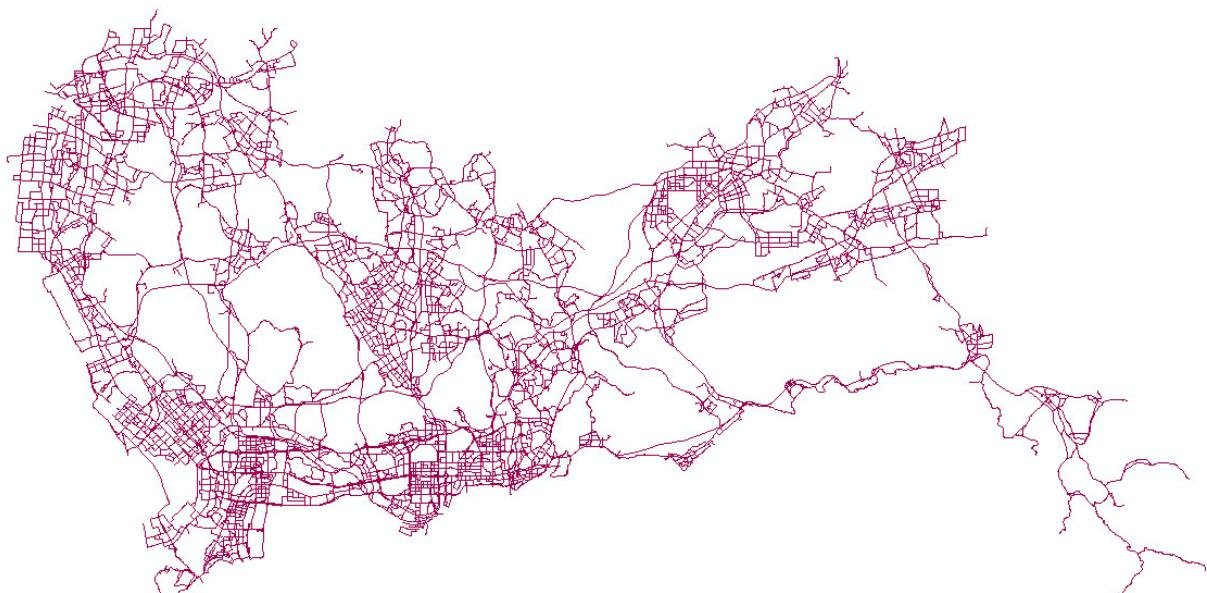


图 1.18: 2017 年度更新公交线路 GIS 数据空间分布图