

# 城市道路间断流行程时间双峰分布

杨帆, 杨晓光, 云美萍

(同济大学 道路与交通工程教育部重点实验室, 上海 201804)

**摘要:** 在传统行程时间单一概率分布研究基础上, 提出了城市间断流行程时间概率分布模型. 以正态分布和对数正态分布为典型分布, 提出了 6 种不同的双峰概率分布函数, 并利用最小二乘法和 Frank-Wolfe 算法进行建模和参数拟合. 研究表明, 双峰概率分布函数能够很好地拟合并反映城市道路间断流行程时间概率分布特征, 突显其与连续流单一概率分布之间的差异性. 研究结果为城市路网路径选择问题提供新的路段阻抗确定方法, 也能够为交通分配提供新的路段行程时间概率函数.

**关键词:** 行程时间概率分布; 双峰分布; 参数拟合

**中图分类号:** U495

**文献标志码:** A

## Urban Road Interrupted Flow Travel Time Bimodal Distribution

YANG Fan, YANG Xiaoguang, YUN Meiping

(Key Laboratory of Road and Traffic Engineering of the Ministry of Education, Tongji University, Shanghai 201804, China)

**Abstract:** Urban interrupted flow travel time probability distribution models are presented based on conventional travel time unimodal distribution. Six types of bimodal probability distributions which apply normal distribution and lognormal distribution as typical distribution are modeled. The parameters of those models are calibrated by Least Square (LS) and Frank-Wolfe algorithm. The study results demonstrate that bimodal distributions, in comparison with unimodal distribution, can fit the real data better and highlight the features of interrupted flow. The research contributions can provide an update link cost definitions for path finding in urban network. Meanwhile, the update link travel time probabilistic function can be studied for traffic assignment based on these contributions.

**Key words:** travel time probability distribution; bimodal distribution; parameter calibration

行程时间是表征路段服务水平的重要特征之一, 城市道路间断流行程时间由于受到信号交叉口延误等因素的影响, 呈现与连续流行程时间不同的特征. 一般而言, 连续流行程时间分布呈现单峰、右偏的分布特征, 多以对数正态分布或 Burr 分布作为连续流的行程时间拟合分布<sup>[1-5]</sup>. 单峰分布模型已经被验证能够很好地描述高速公路及快速路的行程时间特征. 在城市内部道路网中, 定义上游交叉口出口道至下游交叉口出口道为典型路段, 则路段的行程时间会受到道路线型设计, 以及交叉口延误等待而呈现出间断流的特征, 主要体现在: 在相似的外界交通环境条件下, 路段行程时间分布方差较大. 从宏观统计学的角度而言, 必然有一定概率的出行者由于延误而以较长时间通过路段, 因此, 间断流的行程时间通常呈现出双波峰特征. Taylor 等<sup>[6]</sup>进一步对澳大利亚阿德莱德市两条城市主干道的行程时间数据进行分析, 得出该行程时间表征出正偏、长尾的特征, 并存在双波峰(bimodality)的特性, 并在此基础上用 Burr XII 型分布来对路段形成时间和路径行程时间进行了拟合模型研究. 同时, 通过用两个单波峰(unimodal)分布的加权模型来推导出路段行程时间的双波峰分布, 取得了不错的拟合效果. 其研究成果表明, 即使在连续流状态下, 行程时间分布也是有一定概率呈现出双峰分布的特征. Jintanakul 等<sup>[7]</sup>利用贝叶斯混合模型对高速公路行程时间分布进行了研究. 其假设行程时间由两个正态分布的混合分布组成, 用“快速部分”和“慢速部分”来归类快速行程时间样本及慢速行程时间样本. Ji 等<sup>[8]</sup>利用美国 CABS (the campus area bus service) 的公交 AVL (automatic vehicle location) 数据, 在 Jintanakul 等人的研究基础上, 利用分层贝叶斯混合模型对城市路段行程时间进行估计. 结果表明, 该模型可以较好

收稿日期: 2013-04-09

基金项目: 国家“八六三”高技术研究发展计划(2012AA112306); 国家自然科学基金(51178344)

第一作者: 杨帆(1984—), 男, 博士生, 主要研究方向为网络交通流. E-mail: fanyang.tongji@gmail.com

通讯作者: 杨晓光(1959—), 男, 教授, 博士生导师, 工学博士, 主要研究方向为智能交通系统, 交通设计与安全, 智能公交.

E-mail: yangxg@tongji.edu.cn

地描述路段行程时间的双波峰特性. 但总体而言, 对于城市间断流行程时间特征分布的研究目前还较少. 更重要的是, 间断流的行程时间受到如路段交通状况、下游交叉口延误以及高峰平峰等因素的影响, 并非一个固定的双波峰联合分布的情况. 因此, 本文将南京市 RFID (radio frequency identification) 数据为基础, 分析城市间断流路段行程时间宏观分布特征, 针对具体的交通流特征选择合理的联合分布模型. 本文仍以较为公认的正态分布和对数正态分布对行程时间分布进行拟合, 用最小二乘法拟合分布参数, 最后对路段行程时间特性进行定性定量分析.

本文采用最小二乘法对城市间断流行程时间分布进行拟合, 建立凸规划问题, 利用 Frank-Wolfe 算法得到路段双峰行程时间分布函数及各参数. 在此基础上, 通过对这些参数的分析研究对路段交通状

态进行评估.

## 1 间断流行程时间概率分布参数拟合

### 1.1 数据来源

RFID 数据是以无线射频技术为基础, 以射频标签与路侧接收器的通讯来获取车辆的信息数据. 目前在南京主城区已经有超过 55 个主干道路段布设有 RFID 检测器, 全市有超过 70 万辆汽车安装有 RFID 标签(南京市汽车保有量为 117 万辆), 对于每个路段的车辆采样率均在 80% 以上. 图 1 为 RFID 设备布置的示意图及 RFID 行程时间样本分布图. RFID 设备可以采集通过该断面的所有装有 RFID 标签的车辆的信息, 包括加密后的车牌信息、通过该断面的时间、车辆信息等. 对于本研究而言, 车辆连续通过上下游两个断面的时间是主要参数, 通过该

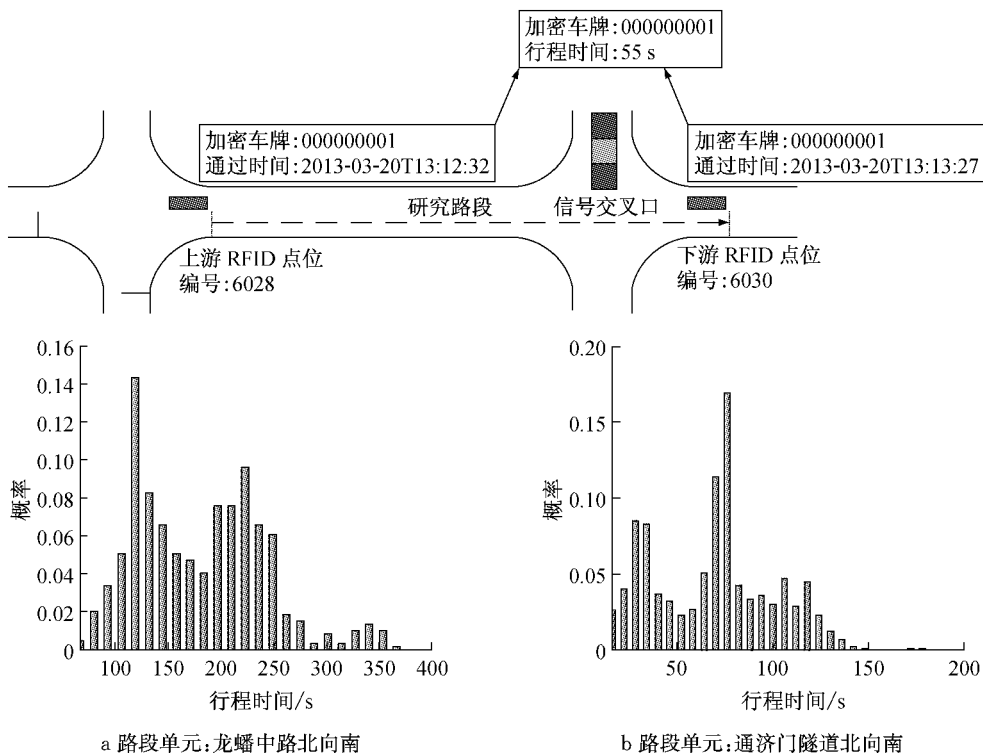


图 1 南京 RFID 数据行程时间概念图及概率分布图

Fig.1 Figures of interrupted travel time and possibility distributions based on RFID data in Nanjing

数据可以推出该车辆通过该路段所花费的行程时间. 行程时间样本分布也明显体现出双波峰特性, 并且呈现出不同的双峰分布形状. 本文对不同的分布进行参数拟合, 以更好地表述每一个路段的行程时间分布函数.

根据行程时间分布的不同形状特征, 以及双峰模型的概念, 城市间断流行程时间的概率分布通式如式(1)所示.

$$f(x) = \lambda f_1(x) + (1 - \lambda) f_2(x)$$

$$F(x) = \lambda F_1(x) + (1 - \lambda) F_2(x) \quad (1)$$

式中:  $f_1(x)$  为第一概率密度函数(简称第一子分布);  $f_2(x)$  为第二概率密度函数(简称第二子分布);  $f(x)$  为双峰分布的概率分布函数;  $F_1(x)$  为第一累积分布函数;  $F_2(x)$  为第二累积分布函数;  $F(x)$  混合分布的累积分布函数;  $\lambda$  为权重系数.

本文将具体地通过 6 种情形的分析, 对行程时

间分布参数进行拟合. 6 种情形分别为:正态分布(用 N 表示)、对数正态分布(用 LogN 表示)、对数正态分布+对数正态分布(用 LogN\_LogN 表示)、对数正态分布+正态分布(用 LogN\_N 表示)、正态分布+对数正态分布(用 N\_LogN 表示)和正态分布+正态分布(用 N\_N 表示). 参数拟合的目标是为了让拟合后的函数与实测的样本之间误差最小,因此以最小二乘法概念建立各拟合模型.

### 1.2 单一正态分布和对数正态分布拟合模型

用正态分布或对数正态分布对行程时间分布进行拟合较为简单. 其模型如式(2)和式(3)所示.

$$\begin{aligned} \min Z &= \sum_i [N(x_i, \mu, \sigma^2) - p(x_i)]^2 \\ \text{s. t } T_{\min} &\leq \mu \leq T_{\max} \\ \sigma^2 &\geq 0 \end{aligned} \quad (2)$$

$$\begin{aligned} \min Z &= \sum_i [\ln N(x_i, \mu, \sigma^2) - p(x_i)]^2 \\ \text{s. t } \ln T_{\min} &\leq \mu - \sigma^2 \leq \ln T_{\max} \\ 0 &\leq \sigma^2 \leq 1 \end{aligned} \quad (3)$$

式中:  $T_{\max}$  和  $T_{\min}$  分别表示行程时间样本中的最大值和最小值;  $\mu$  和  $\sigma^2$  为概率分布的两个参数, 根据特定分布具有特定的含义;  $x_i$  为行程时间样本;  $p(x_i)$  为  $x_i$  所在的行程时间区间值在全样本中所占的比例.

单一正态分布和对数正态分布模型是在式(1)中令  $\lambda=0$  (或  $\lambda=1$ ) 而得的特殊情况. 区别在于, 在对数正态分布模型中, 为了满足概率密度函数为右偏曲线,  $\sigma^2$  为 0 到 1 之间的数. 另外, 对数正态分布中的均值  $\mu$  为行程时间样本统一取对数以后得到的正态分布样本的均值, 因此在约束集中需要在上下界处分别取自然对数值作为其约束范围.

### 1.3 LogN\_N 拟合模型

LogN\_N 分布的拟合模型如式(4)所示.

$$\begin{aligned} \min Z &= \sum_i [\gamma Z_1(x_i, \mu_1, \sigma_1) + (1 - \gamma) \cdot \\ &Z_2(x_i, \mu_2, \sigma_2) - p(x_i)]^2 = \\ &\sum_i \left[ \gamma \frac{1}{\sigma_1 \sqrt{2\pi}} e^{-\frac{(\ln x_i - \mu_1)^2}{2\sigma_1^2}} + \right. \\ &\left. (1 - \gamma) \frac{1}{\sigma_2 \sqrt{2\pi}} e^{-\frac{(x_i - \mu_2)^2}{2\sigma_2^2}} - p(x_i) \right]^2 \\ \text{s. t. } \ln T_{\min} &\leq \mu_1 - \sigma_1^2 \leq \ln \mu_2 \\ e^{\mu_1} &\leq \mu_2 \leq T_{\max} \\ 0 &\leq \sigma_1^2 \leq 1 \\ 0 &\leq \sigma_2^2 \leq \left( \frac{T_{\max} - C_2}{3} \right)^2 \\ 0 &\leq \gamma \leq 1 \end{aligned} \quad (4)$$

式中:  $\gamma$  为权重系数;  $C_2$  为双峰模型的第二个波峰

值. 约束 4 是由正态分布的特性所得, 如图 2 所示. 如样本服从正态分布, 则约 68.3% 数值分布在距离平均值有 1 个标准差之内的范围, 约 95.4% 数值分布在距离平均值有 2 个标准差之内的范围, 以及约 99.7% 数值分布在距离平均值有 3 个标准差之内的范围. 称为“68-95-99.7 法则”或“经验法则”. 因此, 以此约束来界定方差的范围.

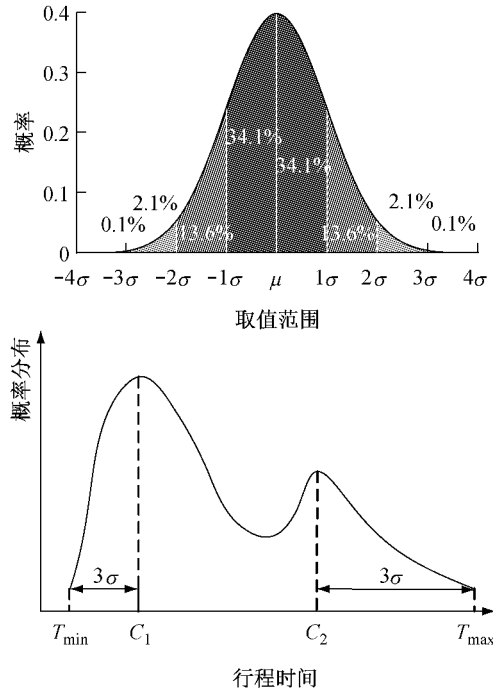


图 2 正态分布特性及其在行程时间分布中的应用

Fig. 2 Normal distribution features and its application to travel time distribution

### 1.4 LogN\_LogN, N\_LogN 以及 N\_N 模型

类似地, 其他 3 类联合分布模型, LogN\_LogN, N\_LogN 以及 N\_N 模型也可以以此方式建模, 如式(5)~(7)所示. 如图 3 所示,  $C_1$  为双峰分布的第一个波峰,  $C_2$  为双峰分布的第二个波峰,  $p_1, p_2$  分别是双峰分布两个模数处对应的概率密度函数的概率值,  $M_1, M_2$  分别两个模数处对应的累积分布函数的概率值. 可以看出, 此 6 种情形的模型均为凸规划问题. 目标函数为一个凸函数, 而约束集为线性约束集 (需适当将变量进行调整, 如在式(6)中, 需令  $y_1 = \ln \mu_1$ , 则可以将约束转化为线性约束). 因此, 可以用凸规划求解算法对其进行求解.

$$\begin{aligned} \min Z &= \sum_i [\gamma Z_1 + (1 - \gamma) Z_2 - p(x_i)]^2 \\ \text{s. t. } \ln T_{\min} &\leq \mu_1 - \sigma_1^2 \leq \mu_2 - \sigma_2^2 \\ \mu_1 - \sigma_1^2 &\leq \mu_2 - \sigma_2^2 \leq \ln T_{\max} \\ 0 &\leq \sigma_1^2 \leq 1 \\ 0 &\leq \sigma_2^2 \leq 1 \end{aligned}$$

$$\begin{aligned}
 & 0 \leq \gamma \leq 1 & (5) \\
 \min Z = & \sum_i [\gamma Z_1 + (1 - \gamma) Z_2 - p(x_i)]^2 \\
 \text{s. t. } & \ln T_{\min} \leq \ln \mu_1 \leq \mu_2 \\
 & \ln \mu_1 \leq \mu_2 - \sigma_2^2 \leq \ln T_{\max} \\
 & 0 \leq \sigma_1^2 \leq (\frac{C_1}{3})^2 \\
 & 0 \leq \sigma_2^2 \leq 1 \\
 & 0 \leq \gamma \leq 1 & (6)
 \end{aligned}$$

$$\begin{aligned}
 \min Z = & \sum_i [\gamma Z_1 + (1 - \gamma) Z_2 - p(x_i)]^2 \\
 \text{s. t. } & T_{\min} \leq \mu_1 \\
 & \mu_1 \leq \mu_2 \leq \ln T_{\max} \\
 & 0 \leq \sigma_1^2 \leq (\frac{C_1}{3})^2 \\
 & 0 \leq \sigma_2^2 \leq (\frac{T_{\max} - C_2}{3})^2 \\
 & 0 \leq \gamma \leq 1 & (7)
 \end{aligned}$$

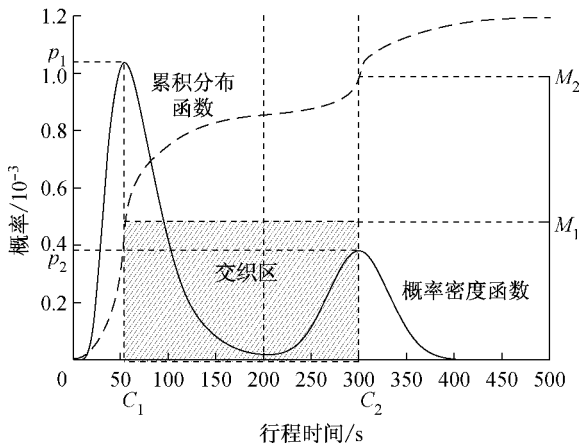


图 3 双峰分布模型及其参数

Fig.3 Bimodal distribution and its parameters

## 2 模型求解

Frank-Wolfe 算法是经典的凸规划问题求解算法,其基本思路是从一个初始解出发,寻找目标函数最优下降方向,并沿此最优方向寻找最优步长,进行最优解的优化,直到目标函数达到最优. Frank-Wolfe 算法多用于交通规划中的用户均衡分配问题,而本问题为普适性的凸规划问题,因此求解过程与交通规划中的 Frank-Wolfe 算法略有不同. 以 LogN\_N 模型为例,介绍具体的 Frank-Wolfe 算法的求解过程,其他模型求解与其基本相同.

步骤 1 (初始化):初始化目标函数. 令  $n=0$ ,  $(\sigma_1^2)^{(0)}$  为 0 到 1 之间的随机数,  $\mu_1^{(0)} = \ln C_1$ ;  $\mu_2^{(0)} =$

$C_2$ ,  $(\sigma_2^2)^{(0)} = (T_{\max} - C_2)/3$ ;  $\lambda^{(0)} = 0.8$ ; 令  $x^{(0)} = \{\mu_1^{(0)}, (\sigma_1^2)^{(0)}, \mu_2^{(0)}, (\sigma_2^2)^{(0)}, \lambda^{(0)}\}$  计算初始目标函数值为  $Z(x^{(0)})$ .

步骤 2 (寻找最优下降方向): 令  $n=1$ . 分别对  $\mu_1, \sigma_1, \mu_2, \sigma_2$  以及  $\lambda$  求偏导,寻找解  $x^{(n-1)}$  处的梯度,所得梯度为  $\mathbf{g}^{(n)} = \left\{ \frac{\partial Z}{\partial \mu_1}, \frac{\partial Z}{\partial \sigma_1}, \frac{\partial Z}{\partial \mu_2}, \frac{\partial Z}{\partial \sigma_2}, \frac{\partial Z}{\partial \lambda} \right\} = \{c_1, c_2, c_3, c_4, c_5\}$ .

步骤 3 (寻找最优步长): 建立线性模型如式 (10) 所示,用单纯性法求该模型的最优解为  $y^{(n)}$ . 该步骤与传统交通规划中的步骤不同. 在传统交通规划中,在初始流量加载的基础上,寻求到的最短路径即为该线性模型的最优解. 但更广泛意义上来说, Frank-Wolfe 算法需要对  $x^{(n)}$  处进行广义线性规划问题求解,以得到最优解的优化方向.

$$\begin{aligned}
 \max Y = & \sum_i (-c_i) y_i \\
 (\mu_1 = & y_1, \sigma_1^2 = y_2, \ln \mu_2 = y_3, \sigma_2^2 = y_4, \gamma = y_5; \\
 & b_1 = \ln T_{\min}, b_3 = \ln T_{\max}, b_5 = 1, \\
 & b_7 = (\frac{T_{\max} - C_2}{3})^2, b_9 = 1) \\
 \text{s. t. } & y_1 - y_2 - y_6 = b_1, \\
 & y_1 - y_2 - y_3 + y_7 = 0, \\
 & y_3 + y_8 = b_3, \\
 & y_2 - y_9 = 0, \\
 & y_2 + y_{10} = b_5, \\
 & y_4 - y_{11} = 0, \\
 & y_4 + y_{12} = b_7, \\
 & y_5 - y_{13} = 0, \\
 & y_5 + y_{14} = b_9 & (10)
 \end{aligned}$$

用黄金分割法求最优步长  $\alpha$ . 求最优步长  $\alpha$  使得目标函数  $Z(x^{(n)} + \alpha y^{(n)})$  最大.

步骤 4 (迭代更新). 如果  $|Z(x^{(n-1)} + \alpha y^{(n-1)}) - Z(x^{(n)} + \alpha y^{(n)})| < \epsilon$  (收敛阈值),则停止迭代,得到最优解及最优目标函数;否则更新最优解,令  $x^{(n)} = x^{(n-1)} + \alpha y^{(n-1)}$ ,返回步骤 2,重新更新计算.

## 3 数值分析

选取南京市若干个主要路段进行数据拟合分析,结果见表 1,表 2. 表 1 为路段基本信息,其中 B\_ID 是路段上游 RFID 点位 ID, E\_ID 为路段下游 RFID 点位 ID. 表 2 为最优拟合模型及参数,其中用 SECTION\_ID 来表述某一具体的路段,如“62456243”代表“中山南路南向北(金沙井到府西街)”. 图 4 是 6 种组合分布的具体拟合情况.

表 1 路段基本信息

Tab.1 Section basic information

B_ID	E_ID	长度/m	说明	B_ID	E_ID	长度/m	说明
6245	6243	431	中山南路南向北(金沙井到府西街)	6338	6339	602	白下路东向西(洪武路到鼎新路)
6244	6246	348	中山南路北向南(金沙井到集庆路)	6424	6255	629	长江路西向东(洪武北路到太平北路)
6260	6258	551	中山东路西向东(洪武路到太平南路)	6292	6290	559	珠江路西向东(中山路到洪武北路)
6258	6256	332	中山东路西向东(太平南路到长白路)	6290	6288	813	珠江路西向东(洪武北路到花红园)
6256	6254	641	中山东路西向东(长白街到龙蟠中路)	6288	6286	716	珠江路西向东(花红园到龙蟠中路)
6254	6252	563	中山东路西向东(龙蟠中路到解放路)	6286	6284	480	珠江路西向东(龙蟠中路到黄浦路)
6252	6250	679	中山东路西向东(解放路到明故宫路)	6285	6287	466	珠江路东向西(黄浦路到龙蟠中路)
6250	6249	312	中山东路西向东(明故宫路到清溪路)	6287	6289	1 071	珠江路东向西(成贤街到丹凤街)
6251	6253	390	中山东路东向西(明故宫路到黄浦路)	6289	6291	664	珠江路东向西(成贤街到丹凤街)
6033	6031	638	龙蟠中路南向东(象房村路到大光路)	6151	6149	731	明故宫路北向南(珠江路到中山东路)
6028	6030	851	龙蟠中路北向南(中山东路到常府街)	6148	6150	834	洵道街南向北(瑞金路到中山东路)
6030	6032	706	龙蟠中路北向南(常府街到大光路)	6150	6152	679	明故宫路南向北(中山东路到珠江路)
6028	6032	1 556	遇济门隧道北向南	6144	6172	1 301	光华路西向东(街道街到蓓园大街)
6332	6330	568	户部街西向北(洪武路到太平南路)	6143	6141	808	大明路北向南(光华路到秦虹路)
6330	6328	297	常府街西向东(太平南路到长白街)	6141	6139	980	大明路北向南(秦虹路到育仁路)
6328	6326	797	瑞金路西向东(长白街到龙蟠中路)	6139	6137	958	大明路北向南(育仁路到永乐路)
6321	6323	1 705	后标营路东向西(首蓓园大街到街道街)	6136	6138	612	大明路南向北(明匙路到大明西路)
6325	6325	484	瑞金路东向西(街道街到解放路)	6138	6140	1 113	大明路南向北(大明西路到应天大街)
6325	6327	506	瑞金路东向西(解放路到龙蟠中路)	6140	6142	504	大明路南向北(应天大街到秦虹路)
6327	6329	645	常府街东向西(龙蟠中路到长白街)	6266	6264	278	汉中路西向东(虎踞路到罗廊巷)
6329	6331	340	户部街东向西(长白街到太平南路)	6264	6262	513	汉中路西向东(罗廊巷到莫愁路)
6171	6169	581	首蓓园大街南向北(光华路到紫金路)	6222	6220	758	中山北路南向北(鼓楼到云南北路)
6343	6344	553	升州路西向东(仙鹤路到中山南路)	6220	6218	845	中山北路南向北(云南北路到湖南路)
6344	6345	315	升州路西向东(中山南路到中华路)	6231	6229	1 113	中央路南向北(湖南路到新模范马路)
6345	6346	575	健康路西向东(中华路到平江府路)	6272	6270	377	北京东路西向东(丹凤街到进香河路)
6336	6337	857	白下路东向西(健康路到太平南路)	6269	6271	377	北京东路东向西(市政府门前到进香河路)
6337	6338	425	白下路东向西(太平南路到洪武路)	6271	6273	379	北京东路东向西(进香河路到丹凤路)
6268	6052	1 030	北京东路西向东(太平北路到太平门路)				

表 2 各路段行程时间最优拟合模型及参数

Tab.2 The optimized fitting model and its parameters of each section travel time fitting

SECTION_ ID	$T_{max}/s$	$T_{min}/s$	最佳拟合参数	SECTION_ ID	$T_{max}/s$	$T_{min}/s$	最佳拟合参数
60286030	472	40	LogN_LogN(4.88,0.53,5.44,0.69,0.61)	62316229	375	93	LogN_N(5.12,0.08,262.92,2723.55,0.75)
60286032	704	49	LogN_N(4.55,0.05,258.23,22535.97,0.78)	62446246	199	16	N_LogN(68.08,641.78,4.41,0.98,0.72)
60306032	385	31	LogN_LogN(4.61,0.26,5.46,0.86,0.8)	62456243	254	20	LogN_LogN(4.61,0.01,5.23,0.82,0.8)
60336031	382	38	LogN_N(4.41,0.16,190.72,4186.84,0.78)	62506249	185	9	N(50.99,782.01)
61366138	334	30	LogN_LogN(4.12,0.2,5.39,0.25,0.83)	62516253	223	17	LogN_LogN(3.26,0.6,4.33,0.03,0.8)
61386140	366	66	LogN_LogN(5.01,0.18,5.99,0.23,0.84)	62526250	371	34	N_LogN(123.55,2146.78,4.99,1,0.71)
61396137	489	48	N_LogN(104.65,1369,5.38,1,0.75)	62546252	294	28	N_LogN(119.03,2433.78,4.5,0.48,0.64)
61406142	299	26	LogN_LogN(4.06,0.2,4.99,0.68,0.69)	62566254	192	15	LogN_N(3.74,0.28,85.24,1387.94,0.73)
61416139	344	57	LogN_N(5.02,0.35,289.44,7086.14,0.86)	62586256	192	15	LogN_N(3.74,0.28,85.27,1387.54,0.73)
61436141	483	49	LogN_LogN(4.83,0.31,5.46,0.66,0.73)	62606258	324	33	N_LogN(90.17,961.4,93,0.74,0.78)
61446172	379	67	LogN_N(4.99,0.12,238.47,2435.83,0.72)	62646262	293	32	N(78.56,1711.33)
61486150	385	58	LogN_LogN(4.81,0.17,5.75,0.21,0.83)	62666264	164	12	N_N(32.21,196.97,29,841,0.61)
61506152	377	42	LogN_LogN(4.43,0.51,5.91,0.32,0.86)	62686052	432	66	N_LogN(135.13,1764,5.63,0.78,0.78)
61516149	318	42	LogN_LogN(4.41,0.2,5,0.25,0.69)	62696271	222	17	LogN_LogN(3.36,0.15,4.75,0.24,0.77)
61716169	342	32	LogN_LogN(4.3,0.42,5.03,0.15,0.67)	62716273	123	17	N_LogN(32,113.78,3.88,1,0.8)
62206218	406	48	LogN_LogN(4.8,0.14,5.58,0.6,0.8)	62726270	210	19	LogN_LogN(4.03,0.64,4.74,0.62,0.8)
62226220	445	34	N(78.11,3765.04)	62856287	276	27	LogN_N(3.69,0.98,120.23,3063.44,0.7)
63296331	195	22	LogN_LogN(4.07,0.27,4.9,0.27,0.63)	62866284	260	30	LogN_LogN(4.14,0.31,5.22,0.24,0.84)
63306328	165	13	LogN_N(3.62,0.74,74.4,1177.78,0.62)	62876289	363	63	LogN(4.93,0.24)
63326330	294	64	LogN_LogN(4.95,0.13,5.71,0.97,0.74)	62886286	358	46	LogN_LogN(4.59,0.34,5.52,0.15,0.83)
63366337	258	66	LogN_LogN(4.93,0.02,5.5,0.06,0.8)	62896291	358	43	N_N(65.57,400,199.03,2916,0.8)
63376338	151	33	N_LogN(42,196,3.53,1,0.8)	62926290	280	79	LogN_N(2.05,0.59,97.34,5873.44,0.34)
63386339	298	29	LogN_LogN(4.22,0.26,5.26,0.25,0.84)	63216323	400	93	N_LogN(251.41,7679.26,5.43,0.9,0.73)
63436344	324	31	N_LogN(67,498.78,5.04,0.03,0.8)	63276329	376	39	LogN_LogN(4.59,0.33,5.36,0.38,0.8)
63446345	166	17	LogN_N(3.5,0.14,119.36,866.13,0.69)	63286326	341	52	LogN_N(2,0.6,128.06,3151.15,0.32)
63456346	335	39	N_LogN(84.77,677.82,5.48,0.19,0.8)	64246255	350	37	N_LogN(71.64,413.44,5.64,0.31,0.81)

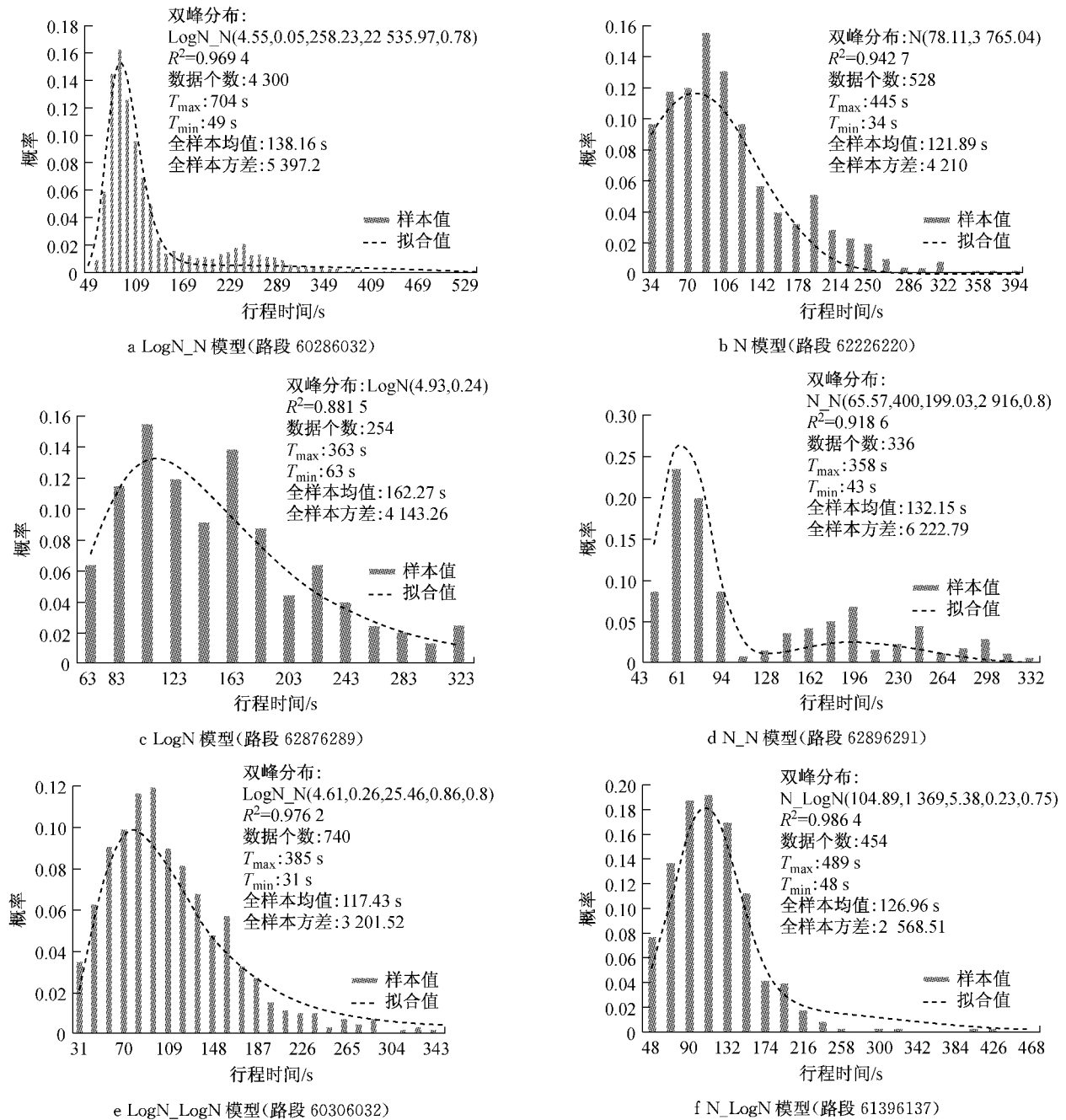


图 4 6 种不同类型的模型拟合曲线及样本分布图

Fig.4 Fitting curves and sample distribution of six models

试验结果表明,不同的路段存在不同的单一分布或双峰模型.以路段 62896291 为例(图 4d),其由两个正态分布组成(图 5),由于第一子分布的权重很大(0.8),同时第二子分布的方差很大(2916),导致第二波峰虽不明显,却与单峰分布有了很大的区别,也更能反映出该路段上以较长时间通过该路段的车辆比例较大.

### 4 结语

本文基于对数正态分布和正态分布,提出了6

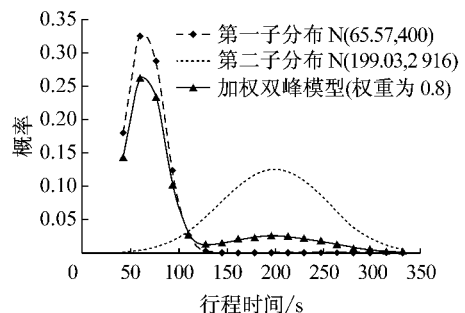


图 5 双正态分布及对应的双峰分布

Fig.5 Two normal distributions and bimodal distributions

类不同的行程时间概率分布的单峰模型和双峰模型.以南京市 RFID 数据为基础,对实际行程时间分布进行了模型和参数的拟合,并分析了双峰模型对于城市行程时间描述的内涵.从广义的角度而言,本文提出的是一种理念和方法,并使用了正态分布和对数正态分布作为具体的分布函数进行拟合,目的是利用这两种合理并且拟合度较高的分布来印证本文对于间断流行程时间双峰分布的假设和研究,但并不表示所有的路段都可以用这两种分布来描述.因此,在后续研究中,可以利用其他分布,如 Burr 分布等,寻求其他的组合,以提高拟合的精度.

间断流双峰分布特性为间断流行程时间可靠性分析提供了新的思路.传统的可靠性指标,如 BI(buffer index)等,均是基于连续流行程时间样本独立同分布的特性.在间断流条件下,则需要区分快速流和慢速流部分的行程时间样本,并重新定义符合间断流行程时间可靠性的一些指标,用以评估城市间断流路段通行质量.此外,本研究成果也验证了在城市道路路径选择、流量分配等环节中,不能简单地以传统最短路径算法来搜索路径,而应该基于城市间断流行程时间概率函数来寻找可行路径集,这将是本文的后续研究方向之一.

#### 参考文献:

- [1] Wardrop J G. Some theoretical aspects of road traffic research [J]. Proceedings of the Institution of Civil Engineers, 1952, 1 (2): 325.
- [2] Herman R, Lam T. Trip time characteristics of journeys to and from work[C]//Proceedings of the 6th International Symposium on Transportation and Traffic Theory. Sydney: University of New South Wales, 1974: 57-86.
- [3] Richardson A J, Taylor M A P. Travel time variability on commuter journeys [J]. High Speed Ground Transportation Journal, 1978, 12(1): 77.
- [4] Polus A. A study of travel time and reliability on arterial routes [J]. Transportation, 1979, 8(2): 141.
- [5] Al-Deek H, Emam E B. New methodology for estimating reliability in transportation networks with degraded link capacities[J]. Journal of Intelligent Transportation Systems, 2006, 10(3): 117.
- [6] Taylor M A P, Somenahalli S. Travel time reliability and the bimodal travel time distribution for an arterial road [J]. Road & Transport Research, 2010, 19 (4): 37.
- [7] Jintanakul K, Chu L, Jayakrishnan R. Bayesian mixture model for estimating freeway travel time distributions from small probe samples from multiple days [J]. Transportation Research Record, 2009, 2136: 37.
- [8] Ji Y, Zhang H M. Travel time distributions on urban streets: their estimation with a hierarchical Bayesian mixture model and application to traffic analysis using high-resolution bus probe data[C/CD]// 92th Transportation Research Board. Washington D C: TRB, 2013.

#### • 下期文章摘要预报 •

### 基于贝叶斯克里金的地下空间多源数据建模

李晓军,李培楠,朱合华,刘俊

利用贝叶斯克里金方法集成钻孔数据与地震解译数据估计地质层面的高程.将线性贝叶斯理论运用于克里金估计,把用于空间估计的数据分为2类,即把钻孔数据视为硬数据,把数据不确定性较大的地震解译数据视为软数据,构建了一个考虑贝叶斯方法的地统计模型,用区域性变量理论研究这2类数据的空间变化特征.通过对某煤矿煤层表面高程的估计以及与普通克里金方法的估计值与误差的比较证明了贝叶斯克里金法能较好地考虑地质剖面和地震解译数据的不确定性,在地质层面高程的估计方面具有一定的优越性,是一种可行的考虑数据全局不确定性的多源数据集成建模方法.