

车辆出行轨迹调查分析中的丢点轨迹还原*

王龙飞, 陈红, 李杨, 邓亚娟
(长安大学公路学院交通工程研究所, 西安 710064)

摘要: 针对利用车牌照生成车辆出行轨迹过程中由于车牌遗漏而产生的丢点轨迹问题, 提出一种丢点轨迹还原方法。首先建立一次可达网络模型来描述调查点之间直接可达关系, 并构建轨迹还原解集; 考虑时距相符程度、路径模式匹配程度、路径曲折度指标, 构造轨迹的效用函数, 利用逼近理想解排序的 TOPSIS 法进行轨迹还原解集决策, 得到最优的还原解。通过一个丢点轨迹示例进行了实验分析, 结果表明, 该方法可以得到满意的丢点还原解。

关键词: 交通调查; 车辆出行轨迹; 轨迹还原; 调查点; 车牌照

中图分类号: TP311.1 **文献标志码:** A **文章编号:** 1001-3695(2014)01-0162-04

doi:10.3969/j.issn.1001-3695.2014.01.038

Track patching method for incomplete track in track-oriented traffic survey and analysis

WANG Long-fei, CHEN Hong, LI Yang, DENG Ya-juan

(Traffic Engineering Research Institute, School of Highway, Chang'an University, Xi'an 710064, China)

Abstract: This paper proposed a track patching method for incomplete track in the track-oriented traffic survey by plate number. Firstly, in order to illustrate the work mechanism of track-oriented traffic survey and analysis, it created a direct access network model and based on which constructed the track patching set. Secondly, considering the consistent degree of time distance relation, matching degree of route pattern, tortuous degree of route track, it constructed track utility function. Finally, it made optimal decision to the track patching set using TOPSIS method. An example of is carried out and the results verifies the feasibility of track patching method.

Key words: traffic survey; vehicle traveling track; track patching; survey point; plate number

0 引言

随着交通信息采集设备自动化、智能化水平的不断提升, 全球定位系统(global positioning system, GPS)^[1]、射频识别(radio frequency identification, RFID)^[2]、无线通信^[3]、无线传感器网络^[4]、车牌自动识别^[5]等技术都可用于追踪、识别车辆在路网中的完整出行轨迹, 通过对路网中全部车辆出行轨迹的科学处理, 可得到详实、准确、可靠的路网交通运行现状信息。当前, 我国城市路网和省域公路网每天都在获取超过百万计的车牌照信息, 可利用面向轨迹的思想从网络化、全局化、轨迹化的角度来分析利用车牌照数据, 掌握交通出行特征及分布规律。车辆在路网中运行时, 会因车流密集、车速过快、异常天气以及设备故障或拍摄角度不对等原因而在调查点无法识别车牌或漏掉车牌, 这样, 车辆行驶途经的调查点序列就成了丢点序列, 对应的车辆出行轨迹称为丢点轨迹, 给轨迹分析的准确度带来严重影响。

在国内外的很多具有面向轨迹思想的交通调查分析中, 有的将遗漏车牌数据的情况默认为异常情况不予考虑, 如文献

[6~9], 但这样处理会在遗漏样本量较大时产生较大误差。但也有文献对遗漏车牌进行了处理, 例如文献[10]提出一种基于采用贝叶斯网络的交通流预测方法, 可有效降低遗漏车牌或错误识别车牌导致的失误, 但该方法目前只适合城市小范围的相邻路段; 文献[11]提出以网格的形式对城市交通进行划分, 通过车牌照统计记录进出网格的车辆, 应用数据挖掘获得车辆出行轨迹, 该方法可在遗漏车牌时获取车辆的近似行使路线, 但误差会随着相邻调查点距离变大而增大。本文结合调查点布设方案、路网拓扑结构、车辆行经各调查点的时距关系以及车辆路径选择行为, 提出了一种车辆出行轨迹还原方法, 可适应城市路网下的车辆出行轨迹还原。

1 轨迹调查分析方法机理

由于本文所研究车辆出行轨迹分析的主体目标是获取全部车辆在路网中的出行轨迹, 因此在调查点布设、调查数据处理与分析以及车辆出行轨迹生成的研究中, 始终立足于路网拓扑结构。下面将从路网拓扑结构出发, 建立解析描述调查点之间的可达性的基础网络模型。

收稿日期: 2013-04-03; **修回日期:** 2013-05-06 **基金项目:** 国家自然科学基金资助项目(51208054); 中央高校基本科研基金资助项目(CHD2011JC191)

作者简介: 王龙飞(1982-), 男, 黑龙江哈尔滨人, 博士, 主要研究方向为智能交通系统(choseoner002@gmail.com); 陈红(1963-), 女, 湖南湘潭人, 教授, 博导, 博士, 主要研究方向为交通运输规划与管理、交通环境与安全、智能交通系统; 李杨(1983-), 女, 陕西榆林人, 讲师, 博士, 主要研究方向为交通设施安全; 邓亚娟(1979-), 女, 陕西宝鸡人, 副教授, 博士, 主要研究方向为智能交通系统。

定义1 一次可达关系。若车辆从调查点 S_A 出发,途中不经过其他调查点即可到达调查点 S_B ,则称 S_A 与 S_B 之间存在一次可达关系, $R_{A \rightarrow B} = S_A \rightarrow S_B$ 。

定义2 前驱/后继一次可达调查点。若 $S_A \rightarrow S_B$,则定义 S_A 为 S_B 的前驱一次可达调查点, S_B 为 S_A 的后继一次可达调查点。

定义3 一次可达路径。若 $R_{A \rightarrow B} = S_A \rightarrow S_B$,则在路网中用于确定 $R_{A \rightarrow B}$ 的路径称为 $R_{A \rightarrow B}$ 的一次可达路径,记为 $PR_{A \rightarrow B}$ 。

定义4 一次可达网络。以调查点为顶点,以调查点对间一次可达关系为边所构成的有向网络为一次可达网络,记为 $T=(V,E)$,其中 V 为顶点集合, E 为边集合。

定义5 调查点时间序列。车辆在路网中行驶所经过的全部调查点按时序排列后得到的序列 $TS=\{S_1, S_2, \dots, S_N\}$,称为该车辆的调查点时间序列,简称调查点序列。

定义6 连续一次可达性。对于 $TS=\{S_1, S_2, \dots, S_N\}$,若 $R_{i \rightarrow i+1} = S_i \rightarrow S_{i+1}$ ($0 \leq i \leq N-1$),即序列中全部相邻调查点都满足一次可达关系,则称 TS 满足连续一次可达性,记为 $TS \propto (S_1 \rightarrow S_N)$, TS 称为连续一次可达序列;否则, TS 不满足连续一次可达性,记为 $TS \npropto (S_1 \rightarrow S_N)$ 。

车辆通过调查点时会留下车牌号和通过时间,全部调查点上的车牌信息经校核、汇总后可统计出路网中每辆车的调查点时间序列,将各序列用一次可达路径串联即可生成车辆出行轨迹,如图1所示。

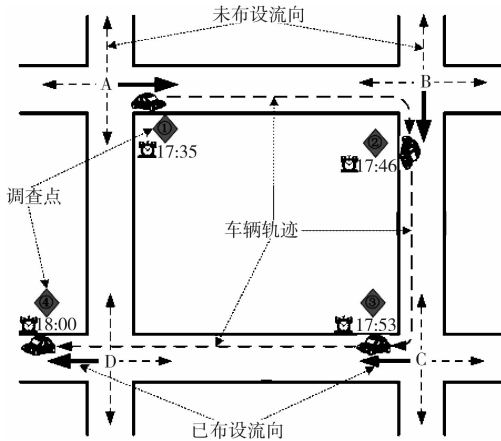


图1 全息信息车辆出行轨迹示意图

2 车辆出行轨迹还原方法

令路网 G 的一次可达网络为 T ,假设丢点序列 $\widehat{TS}=\{\widehat{S}_1, \widehat{S}_2, \dots, \widehat{S}_N\}$,此时 $\widehat{TS} \npropto (\widehat{S}_1 \rightarrow \widehat{S}_N)$,即在 TS 中存在相邻调查点 S_A 与 S_B , $S_A \nrightarrow S_B$,则说明在 S_A 与 S_B 之间丢失了调查点。在丢点还原时,首先按丢失单调点情况获取还原解,若还原不成功,再按丢失连续两个调查点情况获取还原解。因此,车辆出行轨迹还原可分为构造轨迹还原解集和从中决策出最优还原解两步进行。

2.1 构造车辆出行轨迹还原解集

构造车辆出行轨迹还原解集步骤如下:

a)建立 S_A 的后继一次可达调查点集合 $\text{Set}_A^{\text{Next}}=\{S_A^{\text{Next},1}, S_A^{\text{Next},2}, \dots, S_A^{\text{Next},M}\}$ 和 S_B 的前驱一次可达调查点集合 $\text{Set}_B^{\text{Pre}}=\{\text{Set}_B^{\text{Pre},1}, \text{Set}_B^{\text{Pre},2}, \dots, \text{Set}_B^{\text{Pre},N}\}$,令 $\text{Set}_{A \cap B} = \text{Set}_A^{\text{Next}} \cap \text{Set}_B^{\text{Pre}}$ 。若 $\text{Set}_{A \cap B} \neq \emptyset$,那么 $\text{Set}_{A \cap B}$ 中每个调查点都可以作为一个还原解加入到还原解集中,转入 c);若 $\text{Set}_{A \cap B} = \emptyset$,转入 b)。

b)遍历 $\text{Set}_A^{\text{Next}}$ 中全部后继调查点(以 $S_A^{\text{Next},1}$ 为例)。若 $S_A^{\text{Next},1}$ 的后继调查点集合与 $\text{Set}_B^{\text{Pre}}$ 的交集 $\text{Set}'_{A \cap B} \neq \emptyset$,则将 $S_A^{\text{Next},1}$ 分别与 $\text{Set}'_{A \cap B}$ 中各调查点组成调查点对作为一个还原解加入到还原解集中;遍历 $\text{Set}_B^{\text{Pre}}$ 中全部前驱调查点(以 $\text{Set}_B^{\text{Pre},1}$ 为例);若 $\text{Set}_B^{\text{Pre},1}$ 与 S_A^{Next} 的交集 $\text{Set}''_{A \cap B} \neq \emptyset$,则分别将 $\text{Set}''_{A \cap B}$ 中各调查点与 $\text{Set}_B^{\text{Pre},1}$ 组成调查点对作为还原解加入到还原解集中,转入 c)。

c)还原解集中的各还原解即可作为 S_A 与 S_B 之间的丢失调查点,供下一步轨迹还原决策。

2.2 TOPSIS 法车辆出行轨迹还原决策

采用逼近理想解排序的 TOPSIS 法进行轨迹拆分解集决策^[12]。设备选轨迹集为 $T=\{T_1, T_2, \dots, T_m\}$, T_i ($i \in M$) 表示第 i 个备选轨迹。轨迹决策考虑三个指标,即时距相符程度、路径模式匹配程度、路径曲折度。对应轨迹四个属性如下:

a)路段数量。设 N_i 为 T_i 中的路段数量,定义其标准化函数如式(1)所示。

$$f_N(T_i) = \begin{cases} \frac{1}{2} \left(\frac{\min_{j \in M} (N_j) - N_i}{\max_{j \in M} (N_j) - \min_{j \in M} (N_j)} \right) & \max_{j \in M} (N_j) \neq \min_{j \in M} (N_j) \\ 1 & \max_{j \in M} (N_j) = \min_{j \in M} (N_j) \end{cases} \quad (1)$$

式(1)中,路段数量越多,函数值越低,反之则越高,路径最少或所有轨迹的路段数量相同时,函数值均为1。

b)时距相符程度。设 $V_{\max}$ 、 $V_{\min}$ 、 V_s 分别为最高速度、最低时速和最佳速度, V_i 为轨迹 T_i 的时速,定义其标准化函数如式(2)所示。

$$f_V(T_i) = \begin{cases} e^{-2 \left(\frac{|V_i - V_s|}{V_{\max} - V_{\min}} \right)} & \max_{j \in M} (V_j) \neq \min_{j \in M} (V_j) \\ 1 & \max_{j \in M} (V_j) = \min_{j \in M} (V_j) \end{cases} \quad (2)$$

式(2)中,时距相符程度越高,函数值就越高,速度过低和过高都会造成函数值偏低,当所有轨迹的平均速度相同时,函数值达到1。

c)路径模式数量。车辆出行路径大体遵循阶梯型模式或直线模式,如图2所示。

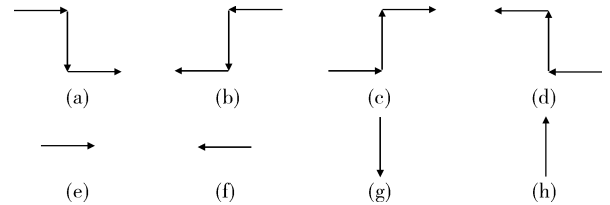


图2 城市路网中车辆出行路径模式

设 P_i 为 T_i 中路径模式的数量,定义其标准化函数如式(3)所示。

$$f_P(T_i) = \begin{cases} \frac{1}{2} \left(\frac{\min_{j \in M} (P_j) - P_i}{\max_{j \in M} (P_j) - \min_{j \in M} (P_j)} \right) & \max_{j \in M} (P_j) \neq \min_{j \in M} (P_j) \\ 1 & \max_{j \in M} (P_j) = \min_{j \in M} (P_j) \end{cases} \quad (3)$$

式(3)中,路径模式数量越少,函数值越高,路径模式最少或所有轨迹的路径模式数量相同时,函数值均为1。

d)车辆转弯次数。设 D_i 为 T_i 的转弯次数,定义其标准化函数如式(4)所示。

$$f_D(T_i) = \begin{cases} \frac{1}{2} \left(\frac{\min_{j \in M} (\sqrt{D_j(D_j-1)}) - \sqrt{D_i(D_i-1)}}{\max_{j \in M} (\sqrt{D_j(D_j-1)}) - \min_{j \in M} (\sqrt{D_j(D_j-1)})} \right) & D_j \neq 0, 1 \\ 1 & D_j = 0, 1 \end{cases} \quad (4)$$

$$D = \begin{matrix} & \begin{matrix} X_1 & X_2 & X_3 & X_4 \end{matrix} \\ \begin{matrix} A_1 \\ A_2 \\ A_3 \end{matrix} & \begin{bmatrix} x_{11} & x_{12} & x_{13} & x_{14} \\ x_{21} & x_{22} & x_{23} & x_{24} \\ x_{31} & x_{32} & x_{33} & x_{34} \end{bmatrix} \end{matrix} =$$

$$\begin{matrix} & \begin{matrix} X_1 & X_2 & X_3 & X_4 \end{matrix} \\ \begin{matrix} A_1 \\ A_2 \\ A_3 \end{matrix} & \begin{bmatrix} 1.0000 & 0.7611 & 1.0000 & 0.7769 \\ 1.0000 & 0.7611 & 1.0000 & 1.0000 \\ 0.6065 & 0.5134 & 0.6065 & 0.6065 \end{bmatrix} \end{matrix}$$

a) 计算标准化决策矩阵。令 $r_{ij} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^m x_{ij}^2}}$ ($i = 1, 2, 3, j = 1, 2, 3, 4$)。

$$R = \begin{matrix} & \begin{matrix} r_{11} & r_{12} & r_{13} & r_{14} \\ r_{21} & r_{22} & r_{23} & r_{24} \\ r_{31} & r_{32} & r_{33} & r_{34} \end{matrix} \\ \begin{matrix} A_1 \\ A_2 \\ A_3 \end{matrix} & \begin{bmatrix} 0.6499 & 0.5352 & 0.6499 & 0.5533 \\ 0.6499 & 0.5352 & 0.6499 & 0.7122 \\ 0.3941 & 0.3614 & 0.3941 & 0.4319 \end{bmatrix} \end{matrix}$$

b) 构造权重标准化决策矩阵。令属性值权重向量为 $\omega = (\omega_1, \omega_2, \omega_3, \omega_4) = (0.25, 0.25, 0.25, 0.25)$, 计算标准化决策矩阵, 将矩阵 R 的每列与其相应的权重 ω_j 相乘得到 $V, v_{ij} = \omega_j r_{ij}$ ($i = 1, 2, \dots, m; j = 1, 2, \dots, n$)。

$$V = \begin{matrix} & \begin{matrix} v_{11} & v_{12} & v_{13} & v_{14} \\ v_{21} & v_{22} & v_{23} & v_{24} \\ v_{31} & v_{32} & v_{33} & v_{34} \end{matrix} \\ \begin{matrix} A_1 \\ A_2 \\ A_3 \end{matrix} & \begin{bmatrix} 0.1625 & 0.1338 & 0.1625 & 0.1383 \\ 0.1625 & 0.1338 & 0.1625 & 0.1781 \\ 0.0985 & 0.0904 & 0.0985 & 0.1080 \end{bmatrix} \end{matrix}$$

c) 确定正理想 A^+ 和负理想解 A^- 。

$$A^+ = \{ \max_{i \in M} v_{i1}, \max_{i \in M} v_{i2}, \max_{i \in M} v_{i3}, \max_{i \in M} v_{i4} \} = \{ 0.1625, 0.1338, 0.1625, 0.1781 \};$$

$$A^- = \{ \min_{i \in M} v_{i1}, \min_{i \in M} v_{i2}, \min_{i \in M} v_{i3}, \min_{i \in M} v_{i4} \} = \{ 0.0985, 0.0904, 0.0985, 0.1080 \}。$$

d) 计算距离。

$$S_{i+} = \sqrt{\sum_{j=1}^4 (v_{ij} - v_j^+)^2} \quad i = 1, 2, 3$$

$$S_{1+} = 0.1589, S_{2+} = 0, S_{3+} = 0.4896$$

$$S_{i-} = \sqrt{\sum_{j=1}^4 (v_{ij} - v_j^-)^2} \quad i = 1, 2, 3$$

$$S_{1-} = 0.4193, S_{2-} = 0.4896, S_{3-} = 0$$

e) 计算与理想解的相对接近程度。

$$C_{1+} = S_{1-} / (S_{1+} + S_{1-}) = 0.7252, C_{2+} = 1, C_{3+} = 0$$

f) 根据 C_{i+} 的大小排列方案偏好顺序为

$$A_2 > A_1 > A_3$$

在本算例中, 方案 A_2 被选中作为最优方案。从对三个备选方案的总体分析来看, A_1 和 A_2 的平均速度为 41.5 km/h, A_3 的平均速度为 54.6 km/h, A_1 和 A_2 与最佳运行速度 (40 km/h) 的偏差最小, 而且 A_1 和 A_2 的轨迹属于典型的阶梯型路径, 而 A_3 的轨迹则属于不规则带有明显绕行的路径, 所以 A_1 和 A_2 要明显优于 A_3 。但就路径的曲折度来讲, A_2 的曲折度比 A_1 更小, 更符合驾驶员的路径选择偏好, 所以 A_2 最终胜出作为最优

方案, 而 A_1 次之, 最差是 A_3 。

4 结束语

本文主要研究了车辆出行轨迹调查分析的丢点轨迹还原问题。首先基于调查点间的一次可达关系构建轨迹还原解集, 并采用 TOPSIS 法进行最优还原解集决策, 应用算例说明了还原方法的可行性。然而车辆出行轨迹以及调查点的车牌数据获取都有很大的不确定性, 本文所提方法只是对轨迹还原问题的初探, 为保证轨迹获取的准确性, 必须将具有针对性调查点布设和区域交通运行特征、规律以及驾驶人路径选择等行为纳入到方法中考虑, 才可确保轨迹还原方法实际应用具有较好的环境适应性。

参考文献:

- [1] 邓中伟, 季民河, 陈雯, 等. 耦合被动式 GPS 与网络调查的居民出行调查[J]. 交通运输系统工程与信息, 2010, 10(2): 178-183.
- [2] HATO E. Development of behavioral context addressable loggers in the shell for travel-activity analysis[J]. Transportation Research Part C: Emerging Technologies, 2010, 18(1): 55-67.
- [3] QIU Zhi-jun, CHENG Peng, RAN Bin. Investigate the feasibility of traffic speed estimation using cell phones as probes[J]. International Journal of Services Operations and Informatics, 2007, 2(1): 53-64.
- [4] LI Xu, SHU Wei, LI Ming-lu, et al. Performance evaluation of vehicle-based mobile sensor networks for traffic monitoring[J]. IEEE Trans on Vehicular Technology, 2009, 58(4): 1647-1653.
- [5] CASTILLO E, JIMÉNEZ P, MENÉNDEZ J M, et al. A Bayesian method for estimating traffic flows based on plate scanning[J]. Transportation, 2013, 40(1): 173-201.
- [6] SHAHPARA A H, AASHTIANIB H Z, FAGHRIA A. Development of a delay model for unsignalized intersections applicable to traffic assignment[J]. Transportation Planning and Technology, 2011, 34(5): 497-507.
- [7] SHERALI H D, DESAI J, RAKHA H. A discrete optimization approach for locating automatic vehicle identification readers for the provision of roadway travel times [J]. Transportation Research Part B, 2006, 40(10): 857-871.
- [8] CASTILLO E, MENÉNDEZ J M, SÁNCHEZ-CAMBRONERO S. Traffic estimation and optimal counting location without path enumeration using Bayesian networks[J]. Computer-Aided Civil and Infrastructure Engineering, 2008, 23(3): 189-207.
- [9] 杨宇伟, 孙剑, 李克平. 城市快速路 AVI 检测设施布局仿真优化[J]. 交通信息与安全, 2010, 28(4): 126-132.
- [10] SÁNCHEZ-CAMBRONERO S, CASTILLO E, MENÉNDEZ J M, et al. Dealing with error recovery in traffic flow prediction using Bayesian networks based on license plate scanning data[J]. Journal of Transportation Engineering, 2011, 137(9): 615-629.
- [11] 赵明, 王寒凝. 基于车牌照识别技术的 OD 调查系统分析与设计[J]. 公路交通科技: 应用技术版, 2008, 48(12): 188-190.
- [12] 徐玖平, 吴巍. 多属性决策的理论与方法[M]. 北京: 清华大学出版社, 2006.
- [13] 王龙飞, 陈红, 李杨, 等. 城市路网车辆出行轨迹调查分析系统[J]. 计算机系统应用, 2011, 20(12): 16-20.