

# 一种基于约束的最短路径低频浮动 车数据地图匹配算法

李清泉<sup>2,1</sup> 胡波<sup>1</sup> 乐阳<sup>2</sup>

(1 武汉大学测绘学院,武汉市珞喻路 129 号,430079)

(2 深圳市空间信息智能感知与服务重点实验室,深圳市南山区南海大道 3688 号,518060)

**摘要:**针对精度差、频率低的浮动车数据特点,给出了空间和拓扑约束下的最短路径浮动车数据地图匹配算法,基于不同采样频率的匹配结果证明算法准确度高。基于武汉市浮动车数据的匹配结果表明,算法具有高可靠性,可以用于浮动车数据的交通信息提取与特征挖掘。

**关键词:**地图匹配;浮动车数据;拓扑约束;交通信息提取

**中图分类号:**P208; U495

地图匹配是 GIS 领域的一个经典命题,目前已有大量的研究基础<sup>[1-3]</sup>。但是随着浮动车技术在智能交通领域的推广应用,低频浮动车数据(floating car data, FCD)的地图匹配算法在近年内再次成为 GIS、交通及计算机等领域内的研究热点<sup>[4-6]</sup>。浮动车技术是利用安装在车辆上的 GPS 实时获取车辆的位置、速度和方向等信息。利用浮动车数据进行实时交通状态提取和发布是近几年研究和实践的一个趋势<sup>[6-7]</sup>。国内浮动车以出租车、公交车为主导。浮动车采样频率一般在 30 s~2 min<sup>[8]</sup>,采样数据精度差(城市环境因素影响),信号丢失严重<sup>[1, 9-10]</sup>。因此,高精度的浮动车地图匹配算法研究对智能交通领域具有现实意义,可以提高交通信息获取的准确度。

传统的地图匹配算法,如基于几何匹配度<sup>[11]</sup>、邻近拓扑约束<sup>[12]</sup>、曲线相似等在高频采样下具有较好的匹配结果<sup>[7-9]</sup>。而针对低频数据的主要地图匹配方法可分为有匹配度综合和统计相关两大类,前者包括点到线、曲线相似<sup>[13]</sup>等方法;后者包括神经网络<sup>[14]</sup>、概率统计<sup>[15]</sup>、模糊决策理论<sup>[16]</sup>等方法。匹配度综合算法采用连续采样点与道路曲线形态相关的准则,常忽略拓扑约束,也有部分采用前序点的局部拓扑约束,但易导致连续匹配错误;统计相关算法以采样点间的采样误差存在关联为基本原则,但低频数据的采样点跨

域多个道路,采样点的相关性很弱,以采样误差相关进行连续修正具有不足。总体来说,现有的针对低频浮动车的地图匹配算法的精度依赖数据质量,算法复杂度高。基于此,本文在传统最短路径地图匹配算法的基础上,利用空间和拓扑约束构建浮动车出行轨迹的邻近备选路段集合,以该集合中轨迹起点至终点的最短路径作为匹配结果,提高了计算效率,同时也保证了匹配的准确度。

## 1 最短路径的浮动车地图匹配算法

低频浮动车数据在驾驶行为、复杂道路网等多重因素的影响下存在采样频率不固定、信号丢失、单一采样点邻近路段多样化等现状。根据不同频率下的采样数据观测可知,采样点对间常见的三种路网拓扑关系如下(见图 1):① 相同路段。在高频采样下,当前点与前序点在同一条路段;② 相邻路段。在中频采样下,当前点与前序点在拓扑邻接路段;③ 拓扑缺失。在低频采样下,采样点间跨越多个路段,采样点所在的路段间出现拓扑关系缺失。本文采用 A\* 算法恢复采样点间的最短路径,进行拓扑补全。

本文研究的浮动车以出租车为主,出租车采集的低频浮动车数据受出租车驾驶行为的影响,在出行轨迹的选择上具有随机性,依赖于驾驶人

收稿日期:2013-04-03。

项目来源:国家自然科学基金资助项目(41171348)。

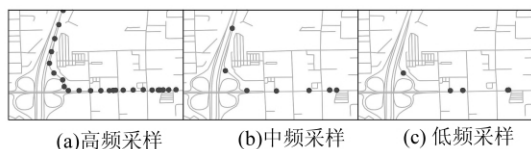


图1 不同频率的采样点空间分布

Fig. 1 Examples of Different GPS Sampling Rates

员对实时交通状态的判断。因此,本文算法采用基于空间和拓扑双重约束准则基于最短路径算法进行浮动车地图匹配。区别于基于全局路网的起/终点最短路径匹配,本算法可以提高匹配结果的准确率。出租车数据记录有载客/空载信息,因此在进行匹配时,以载客或空载行为进行出租车出行轨迹分割,切换点作为轨迹起/终点。邻近备选路段集合是采样点基于空间和路网约束进行邻域搜索的路段集合。具体准则如下:① 空间约束。已有的算法从距离、航向、可达性等进行邻近路段搜索,本文采用距离、航向约束,具体参数以采样数据质量为准。② 拓扑约束。由于存在定位误差,进行拓扑约束的目的是使每一对前后采样点间存在连通路,并最终实现轨迹起点至终点间存在可达路径。

本文在邻近搜索备选路段时,不采用匹配度准则确定当前点的匹配路段,而是确定一个依据空间和拓扑约束的备选路段集合  $G'(V, E)$ , 其中,  $V$  是备选路段的节点集,  $E$  是备选路段边集。从  $G'$  中寻找轨迹出行起点  $P_0$  到终点  $P_n$  的最短路径作为出行轨迹的匹配结果。当  $P_0$  至  $P_n$  间不存在可达路径时,则地图匹配失败;当存在多条可达路径时,以其中的最短路径作为匹配结果。该方法相对于局部拓扑关联采用空间和拓扑约束,实现了轨迹相关的全路径拓扑约束,可以提高匹配的准确度。匹配流程见图2。

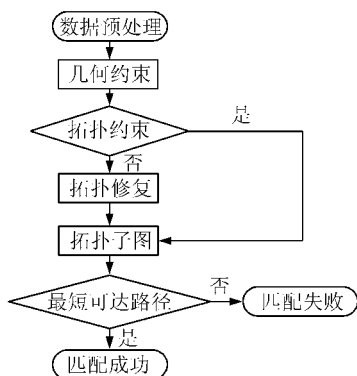


图2 算法核心流程

Fig. 2 Flowchart of Algorithm

## 2 算法验证

### 2.1 精度验证

为了验证该算法的准确度,采用 ACM SIGSPATIAL GIS (ACMGIS) 2012 挑战杯所提供的地图匹配数据<sup>[17]</sup>进行精度论证。对 ACM-GIS 2012 大赛提供的 1 s 高频采样数据,本文进行重采样,分别获得 5 s、10 s、30 s 的采样数据。此数据无采样点的速度及方位信息,因此在进行空间约束的航向检查时,以前后点对间的几何方向为准。参照会议提供的匹配结果,得到不同采样频率下的匹配准确率(见表1)。

表1 不同采样率下的匹配准确率

Tab. 1 Matching Accuracy Under Different Sampling Frequencies

采样率/s	样本数	准确率/%
1	1 320	98.4
5	264	98.5
10	131	98.5
30	44	97.7

表1中的结果表明,采样率降低时,匹配准确率也有所降低,但在4种不同采样频率下的准确率均优于97.7%,说明该算法的匹配准确率高。插页III彩图4中蓝色为采样点,红色为匹配点。从插页III彩图4可知,在不同采样率下,本文算法的匹配结果较为一致,在关键路段(立交桥)的匹配准确度也比较理想。

### 2.2 真实低频数据验证

利用 ACMGIS 大赛提供的高频数据验证本文算法的匹配精度。为了充分验证本文算法的实用性,使用真实浮动车数据进行地图匹配。现有的浮动车均以出租车为主要构成,采样的主要信息见表2。由于受城市环境等多种因素的影响,真实的采样频率一般在 30 s~2 min,也存在信号连续丢失的情况(如穿越地下通道)。

表2 浮动车数据格式

Tab. 2 Data Format of Flowing Car

车辆ID	经度/(°)	纬度/(°)	速度/(m·s <sup>-1</sup> )	车头角/(°)	时间/s	载客
25048	114.336 0	30.515 7	6.24	234.22	4 699	0
25048	114.334 3	30.514 4	5.06	234.79	4 745	1
25048	114.331 1	30.513 2	11.7	248.27	4 785	1
25048	114.328 3	30.512 8	0.00	0.00	4 824	1
25048	114.325 6	30.511 0	11.8	193.69	4 908	0

本文算法采用基于空间和拓扑双重约束准则构建邻近备选路段集合,区别于基于全局路网的

起/终点最短路径匹配,提高了匹配结果的准确率。选取空间邻域路段时,需要确定采样点的缓冲半径和判断航向准确度。本文对浮动车采样点不同缓冲半径内的备选道路数量和不同采样速度下的车头角精度进行了统计分析,见图 3。图 3(a)中,当缓冲半径为 40 m 时,缓冲区内的道路数量趋于稳定。图 3(b)中,当出租车速度高于 20 m/s 时,车头角与道路方向一致(平行关系);在 10~20 m/s,存在 4 个显著峰值(平行或垂直)。因此,本文认为空间约束采用的具体参数如下:缓冲半径为 40 m;车速高于 10 m/s,以车头角为航向;车速低于 10 m/s,以采样点对的几何方向为航向;车速低于 1 m/s,不进行航向限制。

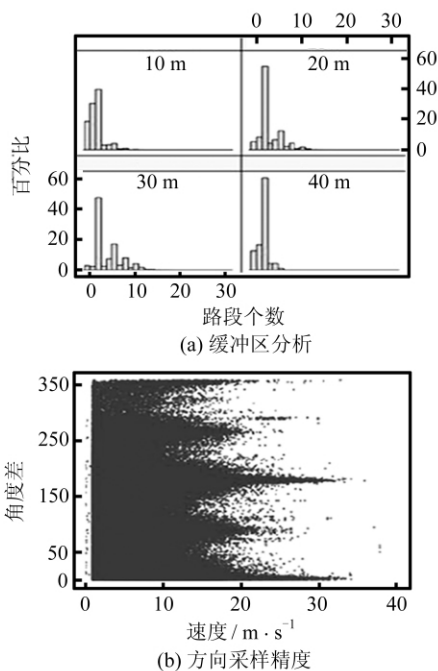


图 3 浮动车数据质量分析

Fig. 3 Data Quality Analysis of Flowing Car

对武汉市 2009 年一个星期 11 000 辆出租车的采样数据进行地图匹配。在浮动车采样存在各种误差的影响下,一个星期的匹配结果统计值如下:99.8%的车辆轨迹恢复率高于 80%,93%的车辆轨迹恢复率高于 90%,轨迹恢复成功率高。本文算法在出租车连续跨越多个交叉路口进行高架桥行驶,以及简单直线道路行驶情况下的出行轨迹均实现了高准确度的匹配(见图 4)。

### 3 算法比较

以一辆浮动车一天的采样数据为例,共包括 1 729 个采样点,分为 89 个轨迹段,将本文算法与采用基于点到线以及全局路网的最短路径匹配算

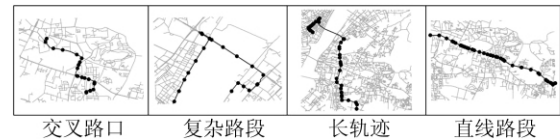


图 4 匹配结果示例

Fig. 4 Match Result Example

法进行对比匹配精度分析。对比结果见表 3。

表 3 匹配算法对比分析结果

Tab. 3 Performance Comparison of Different Map-Matching Algorithms

算法	点到线	全局路网	约束条件
准确率/%	81.2	83.1	94.2
耗时/s	2.1	2.5	4.1

实验分析结果表明,本文算法的匹配准确率相比传统方法提高了 10% 以上,但是以计算时间为代价,消耗在匹配约束条件的判断。点到线的匹配算法在复杂道路网情况下,匹配精度有所下降。全局路网的最短路径并未考虑驾驶习惯,而基于约束条件的最短路径在局部范围进行最短路径匹配,同时考虑了路网的空间约束和拓扑约束,匹配准确度相对较高。

### 4 结 语

本文针对现有匹配算法在低频浮动车数据上的不足,提出了一种基于约束条件的最短路径地图匹配算法。对 ACMGIS 大赛提供的轨迹数据进行了匹配,结果表明,30 s 及以下采样频率的匹配准确度均优于 97.7%;针对武汉市浮动车数据的匹配准确度也优于 90%。这说明在城市道路复杂路网的情况下,该算法虽然需要耗费时间进行备选道路的空间和拓扑约束,在时间效率上有所降低,但相比经典匹配算法,准确率可提升 10% 以上。本文针对特殊情况的处理还有待完善,需要进行更精细的轨迹切分方法和局部搜索准则的改进。

本文现有的最短路径计算以道路长度定权,浮动车的驾驶行为也会考虑到路网交通状态,因此后续研究还可以综合考虑道路的交通状况、通行效率、道路等级等相关因素进行权重确定,以提高最短路径匹配的准确度。

#### 参 考 文 献

- [1] 苏洁,周东方,岳春生. GPS 车辆导航中的实时地图匹配算法[J]. 测绘学报, 2001, 30(3):252-256
- [2] 陈南,高扬. 数字地图在车辆导航中应用的若干技

- 术问题[J]. 武汉大学学报·信息科学版, 2002, 27(2): 143-147
- [3] 李清泉, 徐敬海, 郑年波, 等. 基于功能的导航数据模型[J]. 武汉大学学报·信息科学版, 2007, 32(3): 266-270
- [4] Jin Juntang, 曹凯. 一种自适应轨迹曲线地图匹配算法[J]. 测绘学报, 2008, 37(3): 308-315
- [5] 宋洁, 李国燕, 李娜娜, 等. 基于模糊逻辑的 GPS/DR 地图匹配算法[J]. 计算机工程与科学, 2008, 30(10): 30-32
- [6] 刘春, 杨超, 范业明. 基于流动车数据的道路车速匹配与实时发布[J]. 武汉大学学报·信息科学版, 2010, 35(4): 394-398
- [7] Lorkowski S, Mieth P, Schafer R P. New ITS Applications for Metropolitan Areas Based on Floating Car Data[C]. ECTRI Young Researcher Seminar, Den Haag, 2005
- [8] Lou Yin, Zhang Chengyang, Zheng Yu, et al. Map-Matching for Low-Sampling-Rate GPS Trajectories[C]. The 17th ACM SIGSPATIAL Conference on Geographic Information System, Seattle, USA, 2009
- [9] Wang Yin, Zhu Yanmin, He Zhaocheng, et al. Challenges and Opportunities in Exploiting Large-Scale GPS Probe Data[EB/OL]. <http://www.hpl.hp.com/techreports/2011>, 2011
- [10] 王美玲, 程林. 浮动车地图匹配算法研究[J]. 测绘学报, 2012, 41(1): 133-138
- [11] 付梦印, 李杰, 邓志红. 一种适于车辆导航系统的快速地图匹配算法[J]. 北京理工大学学报, 2005, 25(3): 225-229
- [12] 张小国, 王庆, 万德钧. 基于路网拓扑特性及先验知识的地图匹配算法[J]. 东南大学学报(自然科学版), 2006, 36(4): 625-629
- [13] 李清泉, 黄练. 基于 GPS 轨迹数据的地图匹配算法[J]. 测绘学报, 2010, 39(2): 207-212
- [14] 苏海滨, 王光政, 王继东. 基于模糊神经网络的地图匹配算法[J]. 北京科技大学学报, 2012, 34(1): 43-47
- [15] 唐进君, 刘芳. 基于路径预测的不确定性推理组合地图匹配算法[J]. 测绘学报, 2010, 39(5): 546-550
- [16] 杨新勇, 黄圣国. 基于拓扑结构/自适应模糊决策的地图匹配算法[J]. 上海交通大学学报, 2004, 38(8): 1 389-1 393
- [17] Ali M, Krumm J, Teredesai A. ACM SIG Spatial GIS 2012[EB/OL]. <http://acmgis2012.cs.umd.edu/>, 2012

第一作者简介: 李清泉, 教授, 博士, 博士生导师。现主要从事智能交通系统、3S 集成、基于位置服务、测量工程等方面的研究。  
E-mail: qqli@whu.edu.cn

## Flowing Car Data Map-Matching Based on Constrained Shortest Path Algorithm

LI Qingquan<sup>2,1</sup> HU Bo<sup>1</sup> YUE Yang<sup>2</sup>

(1 School of Geodesy and Geomatics, Wuhan University, 129 Luoyu Road, Wuhan 430079, China)

(2 Shenzhen Key Laboratory of Spatial Smart Sensing and Services, 3688 Nanshai Avenue, Nanshan District, Shenzhen 518060, China)

**Abstract:** As the floating car data is often with a low precision and frequency, this paper proposes a shortest path based map-matching algorithm for floating car data under spatial and topological constraints, the experimental results on different sampling frequencies data show the algorithm have a promising accuracy. Moreover, the experimental result on floating car data from Wuhan city demonstrate that the algorithm has high reliability which can be used for traffic state extraction and feature analysis.

**Key words:** map-matching; floating car data; topological matching; traffic information extraction

**About the first author:** LI Qingquan, professor, Ph. D, Ph. D supervisor. His researches are focused on intelligent transportation system, integration of GIS, GPS and RS, location based service, surveying engineering and so on.

E-mail: qqli@whu.edu.cn

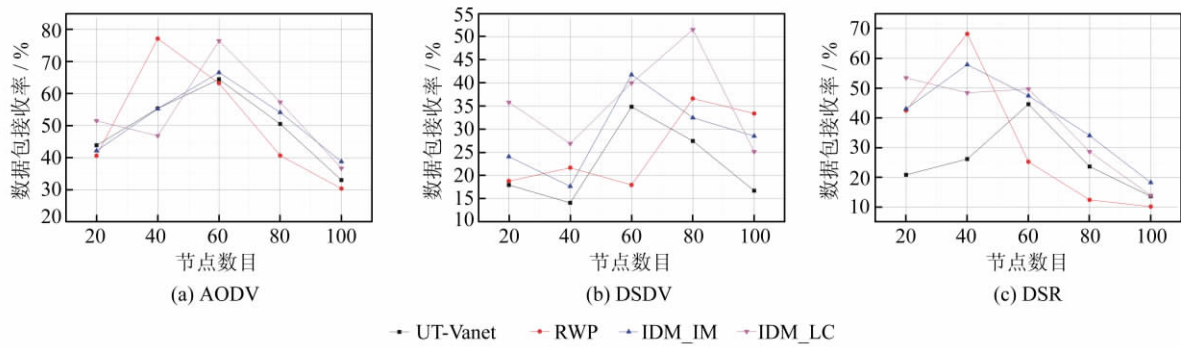


图 1 不同节点数目下的数据包接收率

Fig. 1 Packets Delivery Ratio as a Function of Node Number

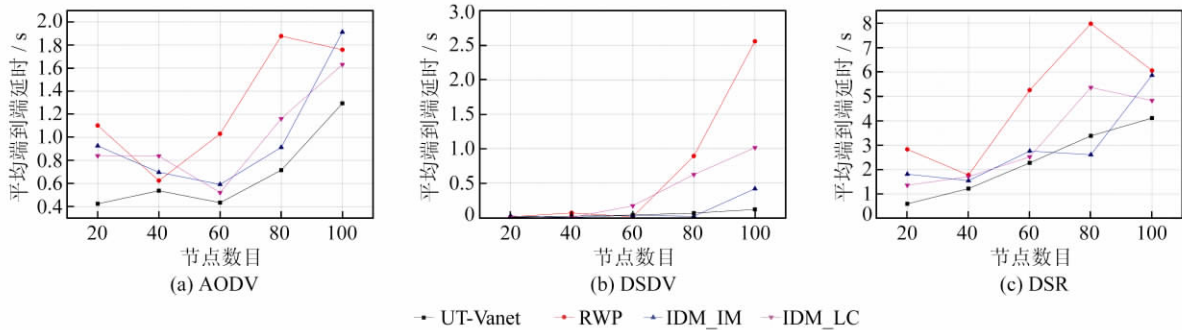


图 2 不同节点数目下的平均端到端延时

Fig. 2 Average End to End Delay as a Function of Node Number

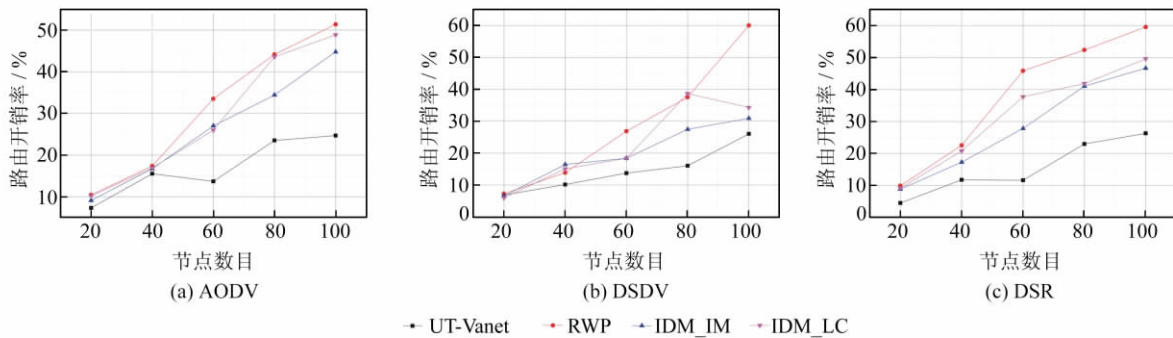


图 3 不同节点数目下的归一化路由开销

Fig. 3 Normalized Routing Load as a Function of Node Number

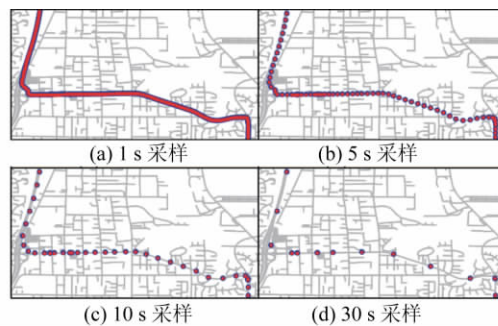


图 4 不同采样频率的匹配结果

Fig. 4 Matching Results Under Different Sampling Frequency