

针对探测车历史数据的离线地图匹配算法

李 强¹ 曹 鹏¹ 缪立新¹

(1 清华大学深圳研究生院,深圳市丽水路 2279 号, 518055)

摘 要:在现有的离线地图匹配算法基础上,采用两次栅格化的思想设计改进算法,大幅度减少了原有地图匹配算法中点到线段距离的计算次数。试验结果显示,本文算法提升地图匹配效率的同时保持了较高的匹配精度,可快速处理大规模的、数据发送间隔在 10~60 s 之间的探测车历史数据。

关键词:探测车;地图匹配;GPS;历史数据

中图法分类号:U491.1

探测车系统(probe vehicle system,或叫浮动车系统)^[1-2]作为新一代的城市道路交通状况采集方式受到了广泛的瞩目,目前仍处于实验阶段。处理探测车历史数据的第一步是地图匹配。由于探测车历史数据量非常庞大,要求地图匹配算法的精度高并且速度快。本文在文献[3]提出的地图匹配算法的基础上,提出改进算法。

1 离线地图匹配问题

由于 GPS 和电子地图都可能出现误差,探测车提供的位置坐标不一定落在电子地图中的路段上。因此,需要通过地图匹配确定探测车行驶的路段和位置,才能计算车辆在某一路段上的行驶时间、进入时刻等信息,并根据这些信息估计该路段的交通状况。图 1 中探测车提供的初始轨迹为一系列位置坐标点,而地图匹配后其轨迹用一系列连续的路段表示。由于数据来自日本,图中车行轨迹靠左。

根据执行地图匹配的时间,地图匹配可分为在线地图匹配(online map-matching)和离线地图匹配(offline map-matching)^[4]。文献[4]设计了一种基于 Dijkstra 法的离线地图匹配方法,其中各路段的权值为该路段到探测车轨迹之间的距离。文献[5-6]提出了基于探测车轨迹与待匹配路线之间的相似度的离线地图匹配算法。与基于最短路径的算法比较,其算法更加复杂。

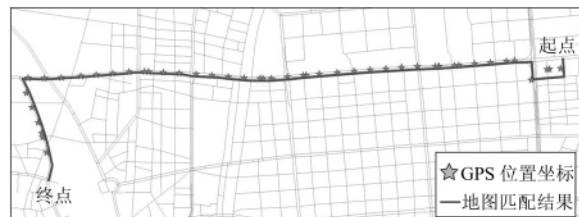


图 1 出行轨迹与地图匹配结果

Fig. 1 Probe Trajectory and Map-Matching Result

Yin 和 Wolfson 在文献[4]中提出的地图匹配算法具有很高的精度,但未对其采用的探测车数据、匹配精度、运行效率等进行描述。相比,文献[3]在论文中详细描述了采用的探测车数据、运行效率验证等内容,且本文已获得文献[3]所采用的探测车数据、电子地图等资料。因此,本文以文献[3]算法为基础进行了改进,并比较了算法效率和精度。

2 改进算法

文献[3]提出的算法直观易懂,匹配精度较高,但非常耗时。究其原因,主要是计算过程中存在大量的点到线之间的距离计算,特别是在抽取探测车轨迹周边路网^[3]时,为了抽取局部路段,路网上的所有路段都需要通过计算来判断其是否在抽取半径之内。

受文献[7]的启发,提出基于栅格化思想的改

收稿日期:2012-12-15。

项目来源:国家自然科学基金资助项目(50808108)。

进算法。栅格化的思想可以不计算点到线的距离而高效地收集一个位置坐标点周边的路段。如图 2 所示,首先计算某个位置坐标点(★)所在栅格的索引 (i, j) ,并通过收集与所在栅格及其周边 8 个栅格相交的路段,可以快速地筛选离位置坐标点距离约等于栅格边长 L 的所有路段。这个过程计算效率远高于计算坐标点到所有路段的距离计算。

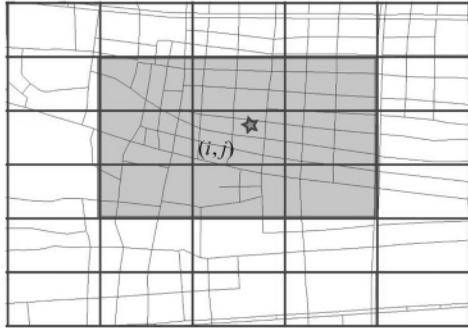


图 2 基于栅格化筛选位置坐标点附近的路段
Fig. 2 Filtering Links Near GPS Coordination
Based on Grid

地图匹配算法步骤如下。

1) 初始化。

把地图划分为两类栅格,第一类栅格用于抽取探测车轨迹周边的路段,生成局部路网。第一类栅格的边长取决于两个连续位置坐标点间距离的最大值,应大于最大值的一半,使其在最坏情况下也能保证局部网络的连通性,避免由于局部路网的不连通导致最短路径搜索失败。第二类栅格用于抽取位置坐标点周边路段,修正其权值。第二类栅格的边长约等于 GPS 误差。

对每一个栅格,收集与其相交的所有路段。栅格与路段相交是指在二维平面上,栅格所覆盖的区域与路段的交集存在非空的点集。判断栅格与路段是否存在交集等同于判断线段与长方形是否存在交集,可采用现有成熟几何算法计算。

2) 匹配探测车历史数据。

读取一个探测车车辆轨迹。生成局部网络,对应 Miwa 算法的抽取探测车周边路径步骤^[3]。针对轨迹中的每个位置坐标点,判断坐标点所在的栅格,收集与所在栅格及其周边 8 个栅格相交的路段。汇总探测车轨迹中所有位置坐标所提取的路段作为局部网络,用于最短路径计算。此过程中使用第一类栅格。修正局部路网中的路段权值,对应 Miwa 算法的修正局部路网权值步骤^[4]。

针对轨迹中的每个位置坐标点,判断坐标点所在栅格,收集与所在栅格及其周边 8 个栅格相

交的路段。汇总探测车轨迹中所有位置坐标所提取的路段,修正其权值为 $L_i \times r$,其中 L_i 为路段 i 的长度, r 为缩短率,取值应小于 1.0。此过程中用第二类栅格。

搜索最短路径,作为地图匹配路线。由于缩短了位置坐标点附近的路段的权值,可保证出行轨迹的起终点间的最短路径与地图匹配路线一致。

读取下一个探测车车辆轨迹,从读取探测车轨迹开始再次执行至所有历史数据处理完成。

判断所有路段与所有栅格之间是否有交集是比较耗时的。但是此过程只需在初始化过程中进行一次,不需要在匹配每个探测车轨迹时都进行,所以对处理上百万个探测车轨迹的过程来说,这一步的耗时可忽略不计。

3 算法效率与精度验证

3.1 名古屋试验及其探测车数据简述

2001 年,Internet ITS 共同研究小组在名古屋市开始了探测车信息采集试验,并进行了基于探测车的道路拥挤状况采集、交通信息服务等方面的研发^[1,8]。

为了保证探测车数据的实时性,名古屋实证试验采用了使用探测车的车载设备通过无线通讯网络直接把数据传送给控制中心服务器的传送方式。探测车的车载设备包括 GPS、无线通讯等模块。车载设备每隔一段时间收集一次本车当前的经纬度坐标、时间、满空状态(载客或空驶)等数据,并通过无线网络发送给控制中心的服务器。为了研究最适合的发送频率,Internet ITS 研究小组设计了多种具有不同发送频率和策略的车载终端。表 1 显示了其终端类型。如类型 2 的车载终端以每 300 m 的距离间隔、每 10 s 的时间间隔以及 SS/ST 事件发生时发送位置等信息。其中 SS/ST 事件是车辆启动和停止时发生的。从表中可以看出,两个连续的位置坐标点最远距离为 300 m,如果考虑偶尔发生的发送信息失败,两个连续的位置坐标点最远距离为 600 m。

地图匹配采用日本数字地图协会 2003 年 3 月出版的 DRM 基本道路网 ver. 1500 (http://www.drm.jp/english/drm/e_index.htm)。电子地图具有很高的精度和连通度,其精度误差小于 5 m,且包括了所有路宽大于 5.5 m 的道路。

3.2 算法实现与参数选择

本文采用 Python 语言实现了地图匹配算

表1 探测车车载终端类型
Tab.1 Type of In-vehicle Device

事件		类型					
		1	2	3	4	5	6
距离/m	300	✓	✓	✓	✓		
	100					✓	
	50						✓
时间/s	550				✓	✓	✓
	10		✓				
	5			✓			
SS/ST		✓	✓	✓	✓	✓	

法。算法实现中需要确定第一类和第二类栅格的边长及缩短率 r 。探测车连续两个位置的坐标点最远距离为600 m,因此第一类栅格的边长定为最远距离的一半,即300 m。尽管电子地图的精度很高,由于GPS的定位误差以及道路本身的宽度,探测车提供的GPS坐标与电子地图中的道路中心线之间还是有一定的距离。分析显示,名古屋探测车试验中GPS位置与道路中心线之间的误差绝大部分在30 m以内。因此,第二类栅格的边长定为30 m。缩短率 r 应小于1.0,在算法实现中选择了0.1作为缩短率,大幅度缩小了探测车轨迹位置坐标点附近的路段的权值。

表2 地图匹配算法效率比较

Tab.2 Comparison of Algorithm Efficiency

算法	出行轨迹数	地图匹配用计算机				运行时间/d	平均时耗 (s/出行轨迹)
		CPU	CPU性能	内存/GB	台数		
改进算法	286 841	AMD Athlon X2 Dual-Core QL-64 2.10 GHz	1 037	2	1	0.5	0.150 6
Miwa算法	3 412 952	Intel Pentium D 3.20 GHz	892	2	3	84	6.379 5

3.4 精度验证

为了验证改进算法的匹配精度,本文随机抽取了1 000条出行轨迹,并人工检查了地图匹配路线与出行轨迹是否一致。验证结果显示92.2%的出行轨迹的地图匹配结果完全一致。文献[3]显示Miwa算法的路段匹配正确率为92.5%。路段匹配正确率是指正确匹配路段占有所有匹配路段的比例。作为精度衡量标准,出行轨迹正确率比路段匹配正确率更加严格,因为只有一条轨迹的所有路段全部匹配正确才认为此轨迹的匹配路线正确。因此,本文提出的地图匹配算法的精度高于Miwa算法。

参 考 文 献

[1] 姜桂艳,常安德,张玮. 基于GPS浮动车采集交通

3.3 效率验证

为验证算法的效率,本文选择了286 841条探测车轨迹(2002年10月~2003年3月期间,工作日早7:00~10:00)进行了地图匹配,并与Miwa的算法进行了比较,如表2所示。Miwa采用的数据为2002年1~3月和2002年10月~2003年3月之间(共9个月)所有数据,包括工作日和休息日0:00~24:00的数据,其探测车轨迹数为3 412 952条。本文选择的数据集为Miwa的数据集的子集。如表2所示,改进算法处理一条出行轨迹的平均时耗约为Miwa算法的1/42,效率远远高于Miwa所提出的算法。由于日本轨道交通非常方便,且出租车的费用较高,出租车出行的轨迹较短。70%的探测车轨迹长度小于4 km,且平均长度为4 km左右^[4]。

需要注意的是,两个算法的开发语言不同,地图匹配时采用的计算机也不同。本文采用了Python语言,Miwa采用了C++,而Python的性能弱于C++。Miwa采用的计算机的CPU为Intel Pentium D 3.20 GHz,本文采用的CPU为AMD Athlon X2 Dual-Core QL-64,CPU的性能评测分别为892和1 037,本文采用的CPU的性能略高。即使考虑这些不同因素,本文提出的地图匹配算法还是明显优于Miwa算法。

信息的路段划分方法[J]. 武汉大学学报·信息科学版, 2010,35(1):42-45

[2] 李清泉,尹建忠,贺奋琴. 面向道路网的GPS浮动车覆盖率模型研究[J]. 武汉大学学报·信息科学版, 2009,34(6):715-718

[3] Miwa T. Travel Behavior Analysis in Road Networks Based on Probe Vehicle Data[D]. Japan:Nagoya University, 2005

[4] Huabei Y. A Weight-based Map Matching Method in Moving Objects Databases; Scientific and Statistical Database Management[C]. The 16th International Conference on Scientific and Statistical Database Management, Greece, 2004

[5] Alt H, Efrat A, Rote G, et al. Matching Planar Maps[J]. Journal of Algorithms, 2003,49(2):262-283

[6] Brakatsoulas S, Pfoser D, Salas R, et al. On Map-matching Vehicle Tracking Data[C]. The 31st International Conference on Very Large Data Bases, Norway, 2005

[7] 张振辉, 崔铁军. 车辆导航系统中地图匹配算法的研究[J]. 地理信息世界, 2007, 2:55-59

[8] 唐克双, 姚恩建. 日本 ITS 开发和运用的实例——

名古屋基于浮动车信息的 P-DRGS 简介[J]. 城市交通, 2006, 3:74-76

第一作者简介:李强,博士,讲师,研究方向为交通状态估计和探测车系统。

E-mail:liqiang@sz.tsinghua.edu.cn

Offline Map-matching for Archived Probe Vehicle Data

LI Qiang¹ CAO Peng¹ MIAO Lixin¹

(1 Graduate School at Shenzhen, Tsinghua University, 2279 Lishui Road, Shenzhen 518055, China)

Abstract: Matching a trajectory of probe vehicle to the roads on the map is necessary to calculate link travel time from probe data. Based on existing offline map-matching algorithms, a map-matching algorithm aiming at archived probe vehicle data is developed. Idea of twice gridding is incorporated into the algorithm, which brings greatly decrease of counts for computing the distance from point to link. Experimental results show that the proposed algorithm significantly improves the efficiency and holds a relative high level accuracy for map-matching. This algorithm is capable of dealing with large-scale archived probe vehicle data having data-sending-rate of one point every 10-60 s.

Key words: probe vehicle; map matching; GPS; archived data

About the first author: LI Qiang, Ph.D, lecturer, majors in traffic performance estimation and probe vehicle system.

E-mail: liqiang@sz.tsinghua.edu.cn

(上接第 156 页)

were complicated and various, under different conditions, water invasion models were different leading to the analysis uncertain results. Floor water invasion model was built by compound analysis technology and its uncertainties were analyzed in this paper. The risk of floor water invasion was zoned and influencing factors were sorted by sensitive analysis technology. Experiment results show that equivalent thickness was the most important factor for reference to make decision of flooding prevention and cure.

Key words: uncertainty; floor water invasion; compound analysis; sensitive analysis

About the first author: QI Chunyan, master, majors in spatial analysis and uncertainty of GIS data.

E-mail: qcy-2001@163.com