

# 基于长间隔大规模 数据的地图匹配算法研究

刘彦挺 吴建平 张 鸽

北京交通大学, 电子信息工程学院, 北京 100044

**摘 要:**随着 GPS 定位技术应用的日益推广, GPS 技术被大量应用于公交车辆及出租车辆的定位与预测。公共交通调度中心的数据量和处理负荷将变得非常巨大,而且租用电信无线网络的费用也很高。如何使用尽可能少的数据得到比较好的定位信息和节约计算资源就成为我们面对的一个重要课题。针对杭州交通调度中心的出租车和公交车的大间隔定位信息,本文提出了一种将基于权值,道路拓扑和最优路径选择相结合的综合地图匹配算法,并应用于实际,取得了很好的效果。

**关键词:**城市交通;地图匹配;长间隔数据;全球定位技术;地理信息系统

中图分类号:U491

文献标识码:A

文章编号:1672-4747(2007)02-0069-06

## A Map Matching Algorithm for Public Traffic Control Center Based on Long Space and Limited GPS Data

LIU Yan-ting WU Jian-ping ZHANG Ge

School of Electronics and Information Engineering,  
Beijing Jiaotong University, Beijing 100044, China

**Abstract:** With the development of GPS, the technology of GPS is widely used in public traffic. If we sample the data as one second per time as usual, the public traffic control center must process a lot of data and bear the heavy financial burden of hiring wireless network. How to get the accurate position information with less data and less calculating resource has become a problem that we must solve. The author puts forward a synthetical map matching algorithm, which is based on the weight, topology, optimization theories and aiming at the

收稿日期:2006-06-12.

基金项目:国家“十五”科技攻关 ITS 专项《杭州市智能交通系统应用试点示范工程》项目编号 2005BA414B16。

作者简介:刘彦挺(1981-);男,汉族,山东海阳人,北京交通大学电子信息工程学院硕士生,研究方向为智能交通。

orientation of the taxi and bus in Hangzhou traffic control center. Application showed that it is effective.

Key words : Urban traffic , map matching , long space data , GPS , GIS

## 0 引言

作为 ITS 示范工程的实施城市,杭州市通过装载 GPS 的 4500 辆公共交通车辆的信息,为城市交通管理提供信息支持。地图匹配是整个信息系统的运用基础。由于整个系统采用集中匹配的方式,所需处理的数据量非常巨大。车辆每隔一秒传送一个定位点是不现实的,并且,匹配算法不能过于复杂,否则,也会影响处理速度从而不能达到实时性的要求。

## 1 杭州交通匹配条件

### 1.1 杭州市的交通路网条件

与北京路网比较,杭州整体路网较密,有些地段的路段长度仅有 20 多米,接近 GPS 定位点的误差范围给匹配算法造成很大的困难。

### 1.2 杭州市的原始数据

杭州市的原始数据分为两个部分,即间隔为 10 秒的公交车定位信息和间隔为 30 秒的出租车定位信息。由于公交车行车为固定线路且速度不能反映道路真实的速度情况,所以,主要利用间隔为长达 30 秒的出租车定位信息。

## 2 地图匹配算法的分析

总体来说杭州路网结构密且数据间隔长,导致定位点信息之间的相关性比较差,这就决定了整个匹配算法不能采用要求相关性很强的曲线匹配,并且由于对匹配算法速度的要求,只能采用具有很好独立性和实时性的点匹配。

### 2.1 基于权重的投影算法

下面简要叙述改进过的该算法原理(如图 1 所示)。

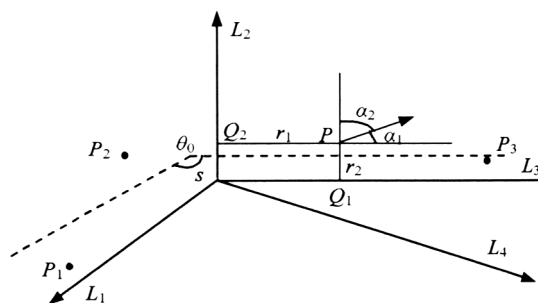


图 1 投影算法原理示意

Fig.1 Projection theory sketch map

图 1 中  $P$  (包括  $P_1, P_2, P_3$ ) 点为待匹配的 GPS 定位点;  $L_1, L_2, L_3, L_4$  表示 GPS 点附近的道路的中心线。在投影匹配算法中,把待匹配的 GPS 定位点向附近所有道路做投影,计算 GPS 点与各道路间的投影距离及车辆行驶方向与道路间的夹角。选出  $r_i, \alpha_i$  值大于给定阈值(剔除一些坏的数据)的所有道路,定位时卫星数少于 4 颗的也不予考虑,根据公式(1)计算各候选道路的权值  $W$ 。

$$W = W_\alpha + W_r + W_n \quad (1)$$

式中,  $W_r, W_\alpha, W_n$  分别是投影距离,方向夹角和星数的权值分子。

在所有候选道路中,选择权值  $W$  最大的作为匹配道路,即认为车辆正在该道路上行驶。算法最后将车辆在匹配道路上的投影点作为车辆的当前位置加以显示。在图 1 中,经过以上计算,选择道路  $L_3$  作为匹配道路,并以  $P$  点在  $L_3$  上的投影点  $Q_1$  作为车辆当前位置。在处理道路交叉点附近车辆转弯时,根据道路的连通性和车辆的转角,在所有接续道路中,选择角度变化最为接近的道路,作为匹配结果。其中定义

$$W_\alpha = A_\omega \cos(\alpha) \quad (2)$$

式中<sup>[1]</sup>,  $\alpha$  为点的速度方向于道路方向的夹角,角度本身的权值  $\cos(\alpha)$  在 0-1 之间波动;  $A_\omega (> 0)$  是扩展因子,  $A_\omega$  值大小决定了此权值分量在综合权

值中的比重系数,当车辆的运行方向变化不大时匹配主要依靠位置投影,角度的加权系数比较小。

定义 1

$$W_r = A_d \left( \frac{2}{\pi} \tan^{-1} \left( \frac{1}{r} \right) \right) \quad (3)$$

式中

$$r = \frac{x_3(y_1 - y_2) - y_3(x_1 - x_2) + (x_1y_2 - x_2y_1)}{\sqrt{(x_1 - x_2)^2 + (y_2 - y_1)^2}} \quad (4)$$

(3) 式中,  $r$  为  $P$  点到线段  $AC$  的距离(图 2); (4) 式<sup>[2]</sup>为  $r$  与  $x, y$  坐标的转换的关系。距离本身的权值  $\frac{2}{\pi} \tan^{-1} \left( \frac{1}{r} \right)$  在 0-1 之间波动  $A_d (> 0)$  是扩展因子,  $A_d$  值大小决定了此权值分量在综合权值中的比重系数。由此可知,  $r$  越小  $W_r$  的值越大,说明线路  $AC$  被选的可能性越大。

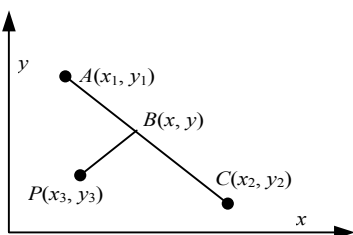


图 2  $r$  的解  
Fig.2 Solution of  $r$

定义 2

$$W_n = A_n \frac{n}{12} \quad (5)$$

式中,在外界条件极好的情况下接受机最多可以接受 12 个卫星的信号,故设星数本身的权值为  $\frac{n}{12}$ ,并在 0-1 之间波动; $A_n$  为扩展因子,  $A_n$  值大小决定了此权值分量在综合权值中的比重系数,卫星数越多权值越大。

在杭州市项目中,我们取三者的权重系数为  $A_w = 1, A_d = 1, A_n = 0.5$ 。 $W$  的取值范围在 0~2.5 之间,这是根据匹配的效果调整地系数,具体如何寻找合适的比例系数还有待研究。

通过调整算法的三个参数,能够正确匹配大多数

点。但是,如距离较近的平行道路中,当根据角度的判决就会失效等的特殊情况下,匹配精度就会降低。考虑到杭州的路网的紧密性,以上算法还需结合道路的拓扑关系才能更精确的匹配路段。

## 2.2 基于拓扑关系匹配法

电子地图足够精确并且具有路网拓扑完整性是地图匹配能够正常工作的前提。由于受车辆速度的限制,因此在一定的时间内,车辆只能驶经有限的路段。可以建立的车辆可能行驶道路实体也是有限的。其中对数字地图的矢量操作提出下列约束条件:

- (1) 路面宽度的不同以“线样式”的宽度加以区分,可根据实际情况对路段进行路宽的等级划分;
- (2) 单行道以特殊线样式表示,且以道路实际走向进行矢量化;
- (3) 对组成交叉路口的所有路段,交叉点应为该组路段的公共节点;而对于立交桥、涵洞等实际上并不相交的特殊情况,该交叉点必不为任一路段的节点。

该算法能够有效减少在复杂路网情况下的计算量并且在定位信息有限的情况下能最大化的利用电子地图信息。表 1 为应用拓扑关系的数据结构,依次分别为数据 ID,原始经度,原始纬度,匹配经度,匹配纬度,匹配路段 ID,相邻路段 ID。

表 1 应用拓扑关系的数据结构  
Tab.2 Data frame of topology

ID	Lon	Lat	Matched lon
Matched lat	Matched road ID	Relating road ID	

## 2.3 异常情况优化处理

通过对杭州道路和 GPS 数据的分析,整个匹配过程主要的异常情况在于如果车辆行驶速度较快或电子地图路段过短,采用 30 秒一个定位数据很可能不会定位在候选队列中引起匹配失败,这需要一个验证环节,如果原始数据点距离候选队列中的路段距离超过某个阈值  $W$  则认为发生异常,并认为该点为新

的起始点重新开始匹配。这样处理的结果就会出现原本一条完整的行车路线被分隔成几个间隔的路段如图3所示。

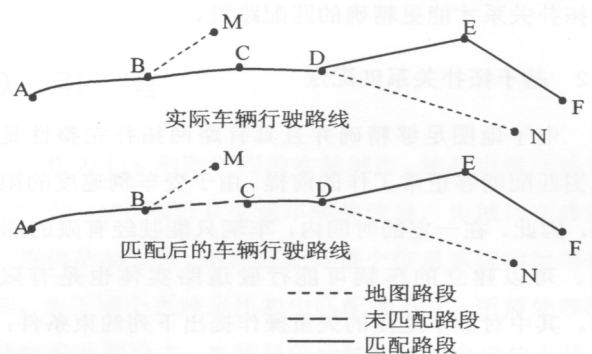


图3 匹配结果与实际行驶路线的差异

Fig.3 Difference between the matching result and real travel path

解决方法

(1) 根据道路本身的拓扑关系

如图3,虽然在BC,DE段没有点匹配到该路段上,但根据道路的连通性可以确定车辆路过BC,DE而不会匹配到BM或DN上

(2) 根据速度信息

如图4所示,已知AB,CD段是已匹配路段可推算出车辆可能走BC段或BFGC段,显然这两段的距离是不同的,因此,可以根据速度来判断车辆的行驶路段。在定位信息中包含车辆速度值 $v$ ,设在AB和CD段之间最近两相邻点过BC的长度为 $s_1$ ,过BFGC的长度为 $s_2$ ,则可求得经由不同路径通过两点的

时间分别为  $t_1 = \frac{s_1}{v}$ ,  $t_2 = \frac{s_2}{v}$ , 并与由GPS数据提供的时间差 $t$ 作对比,取接近的一个则可认为可信。也就是说如果通过时间较长则认为车辆是通过BFGC到达

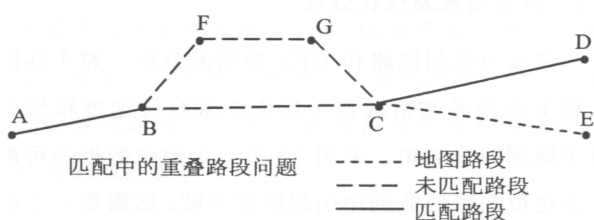


图4 匹配中的重叠问题

Fig.4 Overlap in matching

CD路段的,时间较短则是通过BC到达CD路段。

(3) 通过最优选择

不可避免地,在一些地形特别复杂或定位信息被严重阻隔的路段,能定位的点之间相差好几个路段,这就很难用上述两种方法进行判断。在这里可用两种最优估计法,一种是最优路径选择,一种是最优时间选择。据调查大多数熟悉路网的司机在一般情况下会选择最近的道路到达目的地,因此,在本研究课题中选用最优路径算法。

此外,在匹配道路的中还可利用尽可能多的道路管理信息,例如在北京,未载客的出租车辆不能上二环路,以及很多特定的单行线路及昼夜路段通行限制的不同等。在处理异常情况时,可用如图5的流程进行。

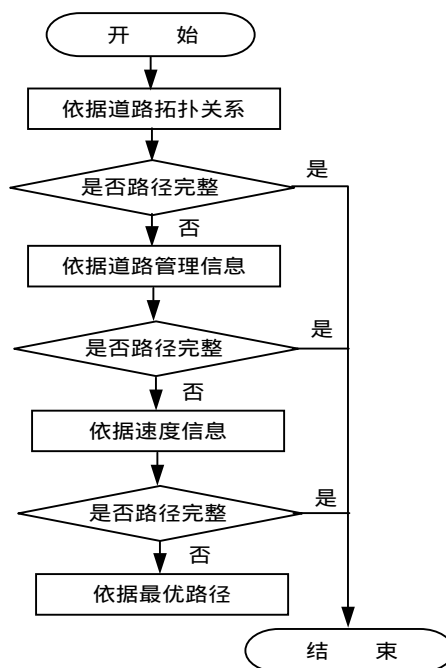


图5 优化处理流程

Fig.5 Procedure of optimization

2.4 算法的整合

2.4.1 算法的整合及总流程图

上面提及了两种地图匹配的算法,在实际匹配过程中要发挥两种算法整合的优势。总的原则是先以改进的投影法确定车辆起始点的位置,在接下来的路段中利用拓扑关系匹配法将车辆可能的下一个路段进行路段筛选,其次再利用投影法在候选路段中确定

车辆实际所在路段,从而降低了计算量也使得匹配精度得到增加。总体流程图如图 6。

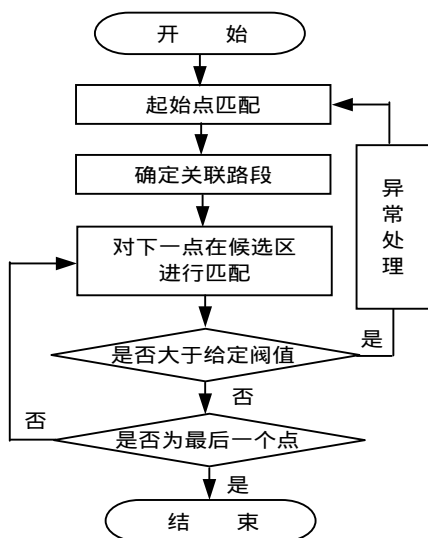


图 6 整合算法的运行框图

Fig.6 Flow chart of the conformity algorithm

如图 7 所示前者采用整合算法,通过拓扑关系将车辆的匹配点限定在一条道路上,是车辆真实行驶路线的反映,而只采用点匹配的算法由于两条路线平行角度一致,当点由于误差倾向于旁边的路段时就会匹

配到旁边的路段,造成如图所示车辆同时行驶在两条道路上的错误匹配。因此,整合算法要明显优于一般点匹配算法。



图 7 整合算法与采用点匹配的效果对比

Fig.7 Fitting contrast between the algorithms of the conformity and matching

### 3 应用实例及结论

对杭州出租车和公交车给出的 GPS 数据和电子地图,采用本文所述的方法进行事后分析和处理,整个正确匹配率为 97.8% 以上。在 Windows 操作系统上,AMD2500+, 256 M 内存配置下单点匹配时间整合算法与一般点算法的平均时间耗用为 0.018 s 和 0.014 s,很好地满足了速度上的要求。算法时间耗用对比如图 8 所示。图 9 为令人满意的匹配效果图。

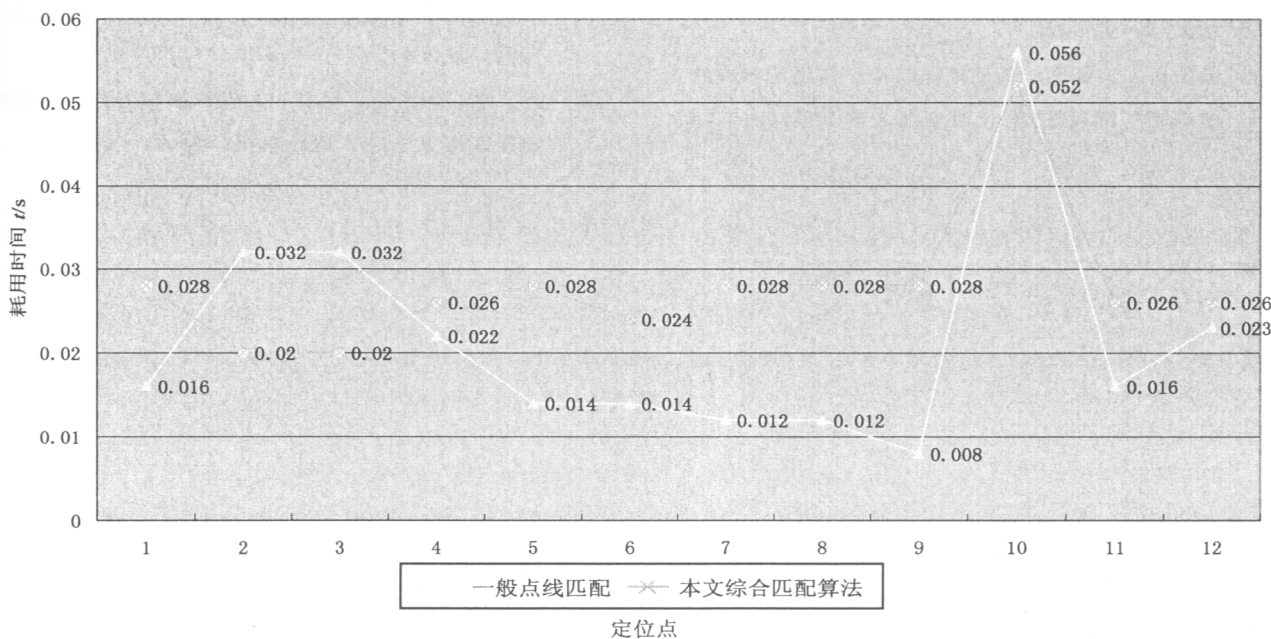


图 8 不同算法的时间耗用对比

Fig.8 Time contrast between the two algorithms

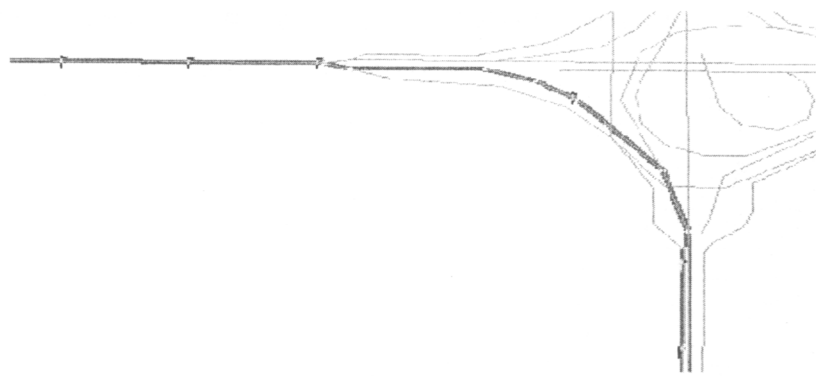


图9 在实际杭州路网中的应用

Fig.9 Application results in Hangzhou road network

采用权重法,投影法,拓扑结构法相结合的综合方法,应用于杭州的实际路网不但拥有较好的匹配效

果,而且运算速度快,能够满足杭州市大容量,高实时的匹配要求。取得了良好的社会效益。

参考文献

[1] Kim J S ,Lee J H ,Kang T H ,et al. Node based map matching algorithm for car navigation system [C]. Proceedings of the 29th ISATA Symposium. Florence : 1996. 21-26.

[2] White Christopher E., Bernstein David, Kornhauser Alain L. Some map matching algorithms for personal navigation assistants [J]. Transportation Research C, 2000, 8 : 91-108.

[3] 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(8).

[4] 陈 锋.一种改进的点到线的地图匹配方法[J].内蒙古科技与经济,2004,(6).

[5] 杨东凯,寇艳红,吴今培,张其善.智能交通系统中的地图匹配定位方法[J].交通运输系统工程与信息,2003,(8).

[6] 王 楠,王勇峰,刘积仁.一个基于位置点匹配的地图匹配算法[J].东北大学学报(自然科学版),1999,20(4).

[7] 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,(8).