

A Real-Time Bus Arrival Time Prediction Algorithm Based on Historical GPS

Jiandong Qiu¹, Shaoxuan Dong², Qiang Li²

¹Shenzhen Urban Transport Planning Center Co., Ltd., Shenzhen Guangdong

²Traffic Information Engineering & Technology Research Center of Guangdong Province, Shenzhen Guangdong
Email: 631581227@qq.com

Received: Jun. 18th, 2018; accepted: Jun. 29th, 2018; published: Jul. 6th, 2018

Abstract

Real-time bus arrival time prediction is an important part of smart transit service. Reliable predicting results can help passengers to reduce waiting time and plan trip reasonably. Common algorithms have many problems, such as many kinds of data needed, complex models, single application scenarios and high complexity of algorithms, this paper presents a real-time bus arrival time prediction algorithm based on historical GPS, providing passengers with remaining travel time, number of stations and distance. And an Android application accessing real-time GPS data in Shenzhen City is developed to verify the accuracy and simplicity of the proposed algorithm to meet the needs of large-scale application.

Keywords

Transit, Real-Time Arrival Time Prediction, GPS

一种基于历史轨迹数据的实时公交到站预测算法

丘建栋¹, 董绍轩², 李 强²

¹深圳市城市交通规划设计研究中心有限公司, 广东 深圳

²广东省交通信息工程技术研究中心, 广东 深圳
Email: 631581227@qq.com

收稿日期: 2018年6月18日; 录用日期: 2018年6月29日; 发布日期: 2018年7月6日

摘要

实时公交到站预测是智慧出行服务的重要组成部分,可靠的预测结果能够帮助乘客合理规划出行方案、减少等车时间。针对常见实时公交到站算法存在的所需数据类型多、模型复杂、应用场景单一、算法复杂度高问题,本文提出了一种基于历史轨迹数据的实时公交到站预测算法,为乘客提供时间、剩余站点数量和剩余距离的预测结果。并开发Android应用,接入深圳市实时公交GPS数据,验证了本文所提出算法的准确性和简单性,可满足大规模应用的需求。

关键词

公交, 实时到站预测, GPS轨迹

Copyright © 2018 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 研究背景

公共交通是城市经济和民众生活中最重要的基础设施之一,发展城市公共交通不仅是缓解城市交通拥堵的有效措施,也是改善城市人居环境,促进城市可持续发展的必然要求。智能交通系统通过实时交通信息发布使得出行者对即将面对的交通环境有足够了解,并据此做出合理出行规划,所以作为公交信息服务的重要组成部分,公交实时到站的可靠预测显得尤为重要[1]。

国外关于公交到站时间的研究可追溯到上世纪70年代,根据公交调度提供的发车信息建立静态预测模型,但是由于当时并没有实时公交数据,所以并不能满足公众的出行服务需求。随着GPS定位技术、通信技术的快速发展,使得实时公交数据采集和传输成为可能,从而可以提供动态的到站预测方法[2]。

常用的实时到站预测方法有历史平均法、神经网络模型、卡尔曼滤波模型、支持向量机模型[3]、基于概率预测模型等[2]。但是这些算法大多存在所需数据种类较多、算法复杂、效率不高、适用场景单一,无法进行大规模结果预测等问题[4]。

本文提出一种基于历史数据的实时到站预测算法,实时轨迹确定车辆位置,根据历史轨迹数据预测到站时间和距离,相对于其他算法,简单、高效,能够满足深圳全市运行车辆的实时到站时间预测需求,同时预测结果也相对准确,满足实际应用需要。

2. 算法思路

相同路段的交通运行状况具有周期性特征,历史数据对于公交实时到站的预测具有非常重要的参考价值。本文提出一种基于历史平均法的实时公交到站预测算法,不同于传统历史平均法,而是用站点按次序将线路打断为物理区间,如图1所示。

历史行程时间计算以物理区间为基本单元,按工作日、非工作日区分,以一个小时的时间粒度,求得各时间段内物理区间所需的行程时间分布,作为实时公交到站预测的基础。

算法主要包括历史数据处理与实时预测两部分。

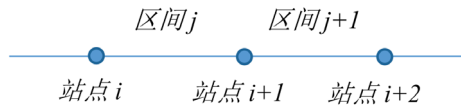


Figure 1. Line interruption
图 1. 线路打断

2.1. 历史数据处理

针对大多数方法所建立的预测模型需要多种数据支持, 本文方法仅需要准备公交线路、站点数据和历史轨迹数据, 对于历史轨迹的处理过程包括以下几步:

1) 数据清洗: 对历史轨迹数据中位置异常和重复、无效记录进行清洗, 将部分有信息缺失的记录根据其他记录进行补齐;

2) 轨迹匹配: 根据轨迹点位置和方向角信息判断车辆运行方向, 并匹配至公交线路最邻近点上(图 2);

3) 到站时间插值: 搜索站点邻近轨迹点, 对历史到站时间进行插值。为了简化估计过程, 本文假设相邻轨迹点间车辆匀速运行, 从而根据距离插值得到站点的到达时间。如图 3 所示, A、B 为相邻两个轨迹点, 记录时间分别为 t_A 和 t_B , 中间经过站点 x , 则到达 x 的时间

$$t_x = t_A + (t_B - t_A) * d_{Ax} / d_{AB}$$

4) 路段行程时间: 计算物理区间内行程时间, 并根据工作日、非工作日, 以 1 个小时为粒度求得各时间段的平均行程时间 t_{AB}^i , i 为第 i 个时间段(图 4)。

2.2. 实时预测

要进行实时公交到站时间预测, 首先要确定车辆所处的线路位置, 非环线的公交大都包含上行和下行两个方向, 而公交 GPS 轨迹信息只包含它所处位置的经纬度、时间、车辆方向角等, 并不包含运行方向。此时, 需要通过位置和方向角判断车辆运行方向。

如图 5 所示, 线路 l 有上行和下行两个方向, 轨迹点 P 的方向角为 α , 与上行方向的夹角更小, 并且 P 点到线路上行方向的最近点 P' 的距离 $|PP'| < 30$ 米, 则将 P 点匹配到线路上行方向上的 P' 点处。此时, 可以推出车辆位于物理区间 $(i, i+1)$ 内, 从而可以预测车辆到达后续站点 $i+1, i+2, i+3, \dots, i+n$ 的剩余时间:

$$t_{P', i+1} = t_{i, i+1} * d_{P', i+1} / d_{i, i+1}$$

$$T_n = t_{P', i+1} + t_{i, i+1} + \dots + t_{n-1, n}$$

3. 算法应用

以本文算法为核心, 以深圳市一个月公交历史轨迹数据为基础进行处理分析, 得到工作日和非工作日各时间段每个物理区间的行程时间分布作为时间预测基础, 开发 Android 应用服务, 接入深圳全市接入实时 11,000 余辆(非全样)GPS 公交数据, 为乘客提供全市范围内约 1 万个公交站点 700 余条线路的实时公交到站预测服务(图 6)。

用户通过应用自动获取的位置信息提供最近站点信息, 或者选择输入需要查询的站点信息, 即可输出经过对应站点的所有线路车辆实时到站信息, 同时可以显示某一条线路上所有车辆的位置信息, 前端 20 秒刷新一次预测结果, 保证信息的时效性。

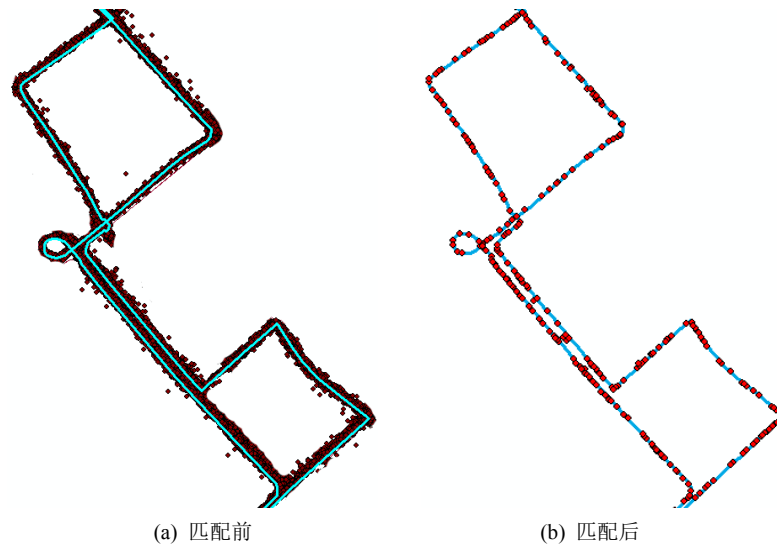


Figure 2. Schematic diagram of trace matching
图 2. 轨迹匹配结果示意图

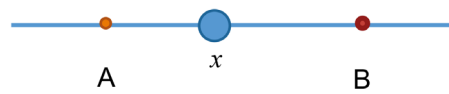


Figure 3. Arrival time interpolation diagram
图 3. 到站时间插值示意图

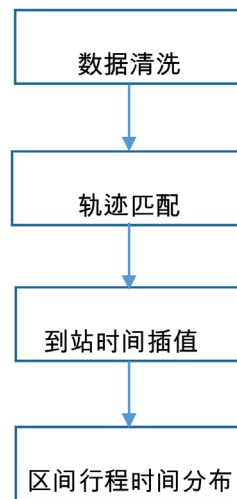


Figure 4. Historical data processing
图 4. 历史数据处理过程

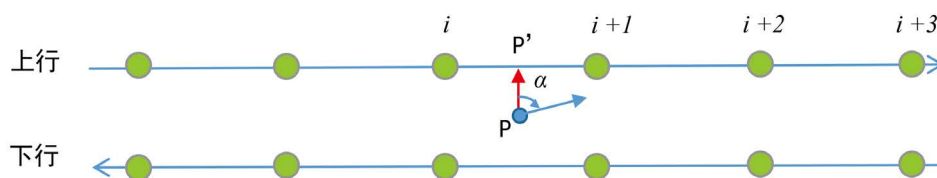


Figure 5. Real-time location acquisition
图 5. 实时位置获取



Figure 6. User interface
图 6. 用户界面

3.1. 数据表设计

实时到站数据表设计如下，包含车辆编号、线路名称、方向名称，每条记录表达当前车辆和公交站点的相对位置和时间预测信息，以 12 路某辆公交车为例，实时获取其位置位于地王大厦站与市委站之间，后续站点的到站预测如表 1 所示。

3.2. 算法效率分析

为满足大规模实时公交到站预测，以深圳市一天实时到站预测算法运行时间为例，测试时间包含从数据接收、处理、计算，到预测结果数据入库全过程，分时间段单次更新所需运行时间分布如图 7 所示。

图 7 可以看出，高峰期运行车辆较多导致计算成本增加，单次更新所需时间大约为 5 秒左右，完全满足实际刷新速度的要求，从而证明本文所提算法的简单、高效性。

3.3. 实测验证

为验证算法有效性，笔者特进行了实地验证，以深圳市竹子林①站为例，通过 7:00~9:00 早高峰时期车辆到站状况与本文预测结果进行对比，实测证明，在实时接入数据缺失的情况下，本文所提方法仍可以实现 70% 以上的准确预测，从而证明了本文所提算法的有效性。

4. 总结与展望

实时公交到站预测是公交信息服务的重要组成部分，本文提出的实时公交到站算法有效解决了当前常见实时公交到站算法存在的所需数据种类多、模型复杂、计算复杂度高等问题。

但是，面对复杂的交通状况，本文算法仍需要从以下几个方面进行改进：

- 1) 对于行程时间预测的精细度有待进一步提高，加入车辆停靠时间、信号灯导致的延迟等因素[5]；
- 2) 对于异常突发事件如事故、台风天气导致的车辆延迟情况，需要在算法中加入异常处理模块。

Table 1. Data display example
表 1. 数据表示例

| 车辆编号 | 站点名称 | 线路名称 | 方向名称 | 剩余站数 | 剩余时间(秒) | 剩余距离(米) | 更新时间 |
|---------|-------|------|--------------|------|---------|---------|------------------------|
| 粤 XXXXX | 市委 | 12 路 | 火车站 - 福田农批市场 | 0 | 72 | 362 | 2017-12-24 14:41:14 |
| 粤 XXXXX | 市总工会 | 12 路 | 火车站 - 福田农批市场 | 1 | 208 | 933 | 2017-12-24 14:41:14 |
| 粤 XXXXX | 荔园小学 | 12 路 | 火车站 - 福田农批市场 | 2 | 310 | 1278 | 2017-12-24 14:41:14 |
| 粤 XXXXX | 妇儿医院 | 12 路 | 火车站 - 福田农批市场 | 3 | 513 | 1775 | 2017-12-24 14:41:14 |
| 粤 XXXXX | 圣延苑酒店 | 12 路 | 火车站 - 福田农批市场 | 4 | 598 | 2144 | 2017-12-24 14:41:14 |
| 粤 XXXXX | 海馨苑 | 12 路 | 火车站 - 福田农批市场 | 5 | 716 | 2856 | 2017-12-24 14:41:14 |

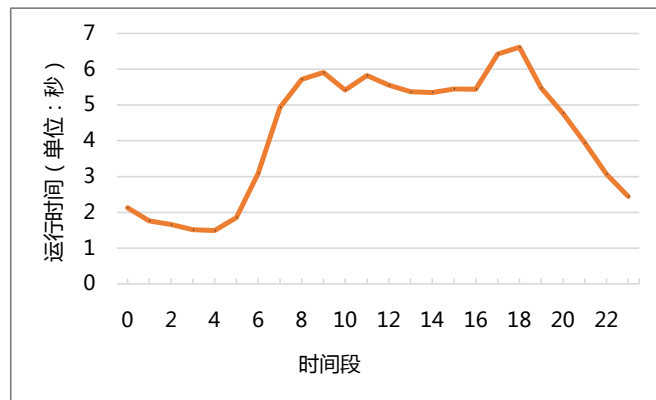


Figure 7. Algorithm efficiency analysis
图 7. 算法效率分析

致 谢

感谢深圳市科技计划项目(项目编号 GGF2016033017241891, 项目名称“深圳市交通大数据公共技术服务平台”)和深圳市战略性新兴产业发展专项资金 2017 年第一批扶持计划(项目名称: 深圳市交通碳排放工程实验室, 批复文号: 深发改(2017)550 号)的资助。

基金项目

深圳市科技计划项目(GGF2016033017241891), 深圳市战略性新兴产业发展专项(深发改(2017)550 号)。

参考文献

- [1] 沈旗. 基于浮动公交车的交通信息实时检测和公交车到站时间预测研究[D]: [硕士学位论文]. 杭州: 浙江大学, 2007.
- [2] 向红艳, 彭学文. 公交到站时间预测研究现状与发展趋势[J]. 交通信息与安全, 2014, 32(4): 57-61.

- [3] 王锐, 杨宏业, 张海鹏. 基于浮动公交车的 SVM 实时交通状态识别研究[J]. 内蒙古工业大学学报, 2016, 35(1): 18-23.
- [4] 胡华, 高云峰, 刘志钢. 基于 AVL 数据的公交到站时间实时预测模型[J]. 重庆交通大学学报(自然科学版), 2012, 31(5): 1014-1017.
- [5] 孙棣华, 赖云波, 廖孝勇, 等. 公交浮动车辆到站时间实时预测模型[J]. 交通运输工程学报, 2011(2): 84-89.

知网检索的两种方式:

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>
下拉列表框选择: [ISSN], 输入期刊 ISSN: 2326-3431, 即可查询
2. 打开知网首页 <http://cnki.net/>
左侧“国际文献总库”进入, 输入文章标题, 即可查询

投稿请点击: <http://www.hanspub.org/Submission.aspx>
期刊邮箱: ojtt@hanspub.org