

车辆 GPS 监控数据的自适应采集算法研究

韩小刚, 晏磊*, 陆恣

(北京大学遥感与地理信息系统研究所, 北京大学空间信息集成与 3S 工程应用北京市重点实验室, 北京 100871)

摘要: 在大规模 GPS 车辆监控中, 当精度要求高时终端与中心端的通信负荷及中心端的数据处理负荷都很大, 自适应采集技术可以有效解决这一问题。该文提出了车辆 GPS 监控数据的自适应采集算法, 采用基于端点速度约束和里程约束的拟合多项式表达速度函数, 综合考虑终端有地图和无地图的情况, 采用偏差距离表示速度的方向偏差, 并提出了数据的恢复算法、算法评价指标以及算法的应用模式。最后基于北京市的真实路网和实测 GPS 数据对算法进行了验证, 恢复后的数据在速度和位置上都逼近真实数据, 数据压缩率达 3%, 证明了算法的有效性。

关键词: GPS; 自适应; 采集

中图分类号: P208 文献标识码: A 文章编号: 1672-0504(2011)03-0012-05

0 引言

随着社会经济的发展, 交通运输的规模化、集团化趋势越来越明显, 同时卫星导航定位技术和智能交通技术应用越来越深入, 使得需要管理的动态 GPS 目标数量急剧上升, 其中最典型的应用是车辆监控管理^[1-8], 如何实时地处理这些大规模的 GPS 数据成为一个重要问题。例如, 中国石油天然气股份有限公司华北油田分公司拥有近 11 000 台车辆, 而对于易燃易爆品的运输必须进行精细的管理和监控, 对监控数据的时间间隔要求很高, 这会导致终端与中心端的通信负载过大, 并且中心端的数据管理难度大、占用存储空间大、查询速度慢等。

车辆 GPS 监控数据的自适应采集技术可有效减少数据传输的次數和数据总量, 其基本原理是终端只在特定的条件下(如运动状态变化量超过阈值)触发数据传输操作, 中心端对于两次传输点之间的监控数据(包括速度、方向、位置等)采用运动函数进行推算。学者们对相关问题展开了研究^[9-11]。王生等研究了通讯时间的动态时间点选择算法, 可以较好地减少数据上传次数^[12]。尹建忠等提出了车载端数据流传输控制模型及算法, 探讨了算法中涉及的速度变化量和方向变化量的判断条件, 给出了监控中心数据恢复的方程^[13]。这些研究成果对减少车辆监控中的通信量 and 数据总量起到了积极作用, 但仍存在一些不足, 例如: 运动函数一般假定为线性, 没有考虑终端没有地图的情况, 未考虑运动方向或方向计算的不合理等。本文针对这些问题, 提出了

车辆 GPS 监控数据的自适应采集算法。

1 自适应采集算法设计

1.1 算法设计思想

本文提出的自适应采集算法主要考虑速度函数表达、终端地图匹配、方向偏差度量三方面, 即采用基于端点速度约束和里程约束的拟合 3 次多项式表示速度函数; 综合考虑终端有地图和无地图的情况; 采用偏差距离而非偏差角度表示速度的方向偏差。

1.1.1 速度函数表达 目前自适应采集算法研究中速度函数一般为线性, 即假定两采集点之间的运动状态为匀速或匀加速运动, 两连续采集点之间的速度由线性速度函数推算, 位置由推算的速度和加速度计算得到。线性速度函数引起的误差较大, 因为实际速度变化往往是非线性的, 如图 1 所示的速度变化趋势; 更严重的是线性速度函数会引起速度估算与里程约束的矛盾, 导致位置或速度的不连续。在连续两个采集点中的后一个点处, 其速度和位置为实际的真值, 若以速度真值为基准进行推算, 则依据里程推算出的该点的位置将与位置真值产生较大偏差; 若以位置真值为基准, 则推算出的速度将与速度真值产生较大偏差。

本研究中速度函数的表达采用基于端点速度约束和里程约束的拟合多项式。采用多项式可以较好地表达复杂情况下的速度变化规律, 有效减小速度估算误差。多项式的拟合采用端点速度和里程两个约束条件, 端点速度约束是指多项式曲线必须经过两个端点, 这可以保证估算速度的连续性; 里程约束

收稿日期: 2011-02-12; 修订日期: 2011-04-10

基金项目: 国家 863 计划支持项目(2009AA12Z217)

作者简介: 韩小刚(1977-), 男, 博士研究生, 研究方向为基于位置的服务。* 通讯作者 E-mail: liyan@pku.edu.cn

是指速度曲线在开始时间和终止时间之间的定积分等于车辆在两个端点之间的行驶里程, 这可以保证估算位置的连续性。综合考虑算法的复杂度和精度, 本文采用三次多项式进行速度函数拟合(图 1)。

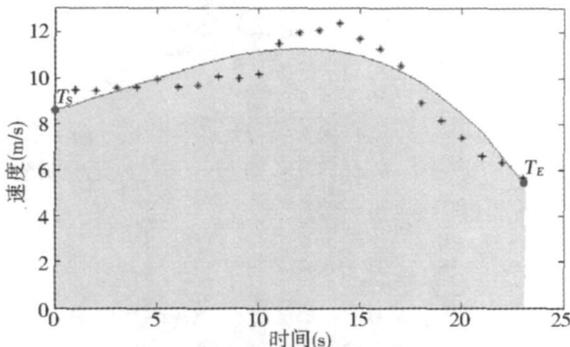


图 1 速度函数示意
Fig. 1 The illustration of speed function

1.1.2 终端地图匹配约束 终端地图匹配约束是指综合考虑终端有地图并成功匹配到地图(以下简称“有地图”)和无地图或无法匹配到地图(以下简称“无地图”)的情况。在实际应用中要考虑成本因素, 所以终端往往没有电子地图; 或者由于电子地图的路网不够详细或更新不及时, 即使安装了电子地图的终端也可能无法有效匹配到路网。因此自适应采集必须综合这两种情况。

有地图和无地图情况下的数据传输触发条件和数据恢复算法都不相同, 数据恢复算法中的里程计算和位置解算也不同, 具体算法将在后面的中心端数据恢复部分详细介绍, 这里重点介绍数据传输触发条件的不同。

当有地图时, 数据传输触发条件为下列两个条件之一: 1) 匹配到路网的路链变化并且在该路链和变化前路链的连接点处存在其它路链, 即车辆行驶过的路网上存在岔路。2) 速度变化超过阈值。当车辆行驶的路上不存在岔路时, 依据开始点和终止点及道路拓扑关系即可得到其行驶轨迹。在终端位置能够匹配到路网的情况下, 匹配路链为单向通行时, 路链前行方向即为车辆的行驶方向, 匹配路链为双向通行时, 可以依据点序列的位置关系推算出车辆行驶方向, 因此触发条件不用考虑方向变化的因素。

当无地图时, 数据传输触发条件为下列两个条件之一: 1) 速度变化超过阈值; 2) 方向偏差(具体度量方法见 1.1.3 节)超过阈值。速度阈值和方向阈值的设定必须合适, 阈值太小会影响数据的压缩比, 阈值太大会影响数据恢复的精度。需要根据实际系统的运行情况进行大量的试验得出经验阈值, 最理想的阈值设定方法是运用自动控制原理对阈值进行

自动调节, 本文不对此进行深入探讨。

1.1.3 方向偏差度量 当终端位置无法匹配到路网时, 车辆行驶方向就成为一个非常重要的因素, 而目前的研究中一般不考虑方向因素或直接用偏差角度度量速度的方向偏差。用偏差角度度量速度的方向偏差的不足可用图 2 说明, P_0 为起始点, 箭头所指为各点的速度方向 D , 在 P_1 点偏差的角度 θ_1 大于 P_2 点偏差的角度 θ_2 , 显然 P_1 点的偏差程度远小于 P_2 点。如果以偏差角度表示速度的方向偏差将会带来大量多余的数据传输, 如一次车道变换都将触发数据传输。所以本研究中用方向偏差距离度量方向偏差, 如图 2 所示的 d_1 和 d_2 。

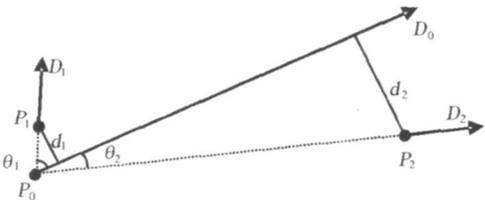


图 2 方向偏差度量示意
Fig. 2 The illustration of direction deviation

1.1.4 算法流程 依据算法的设计思想可得到算法流程, 如图 3 表示。

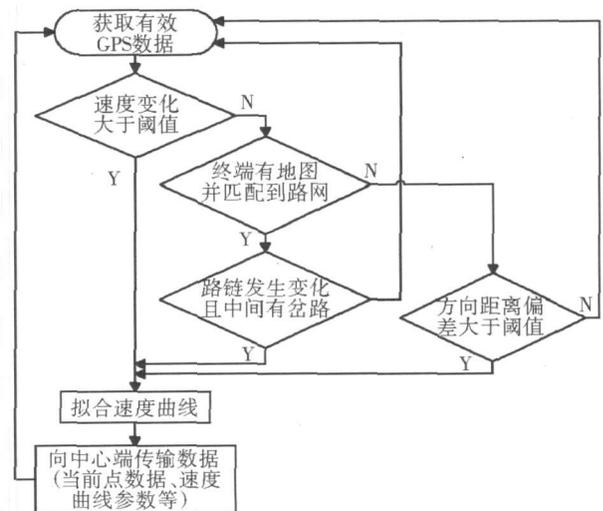


图 3 算法流程
Fig. 3 Algorithm flowchart

1.2 速度函数参数解算

拟合速度曲线是自适应采集算法的重点, 速度曲线用 3 次多项式表示, 设为 $\alpha_3 T^3 + \alpha_2 T^2 + \alpha_1 T + \alpha_0$, 速度曲线拟合即速度多项式系数的解算。

速度多项式满足端点速度约束, 将两个端点的时间和对应速度值(开始时间 T_s 、开始速度 V_s 、结束时间 T_E 、结束速度 V_E) 分别代入多项式, 即得到式(1)和式(2)。速度多项式满足里程约束, 将速度函数在开始时间 T_s 和结束时间 T_E 之间定积分, 即可

得到 T_S 和 T_E 时段内车辆行驶的里程 S , 由此可导出式(3)。当终端位置能够匹配到路网时, S 为两个端点间路链的长度; 当终端位置无法匹配到路网时, S 为两个端点间的直线距离。在端点速度和里程约束条件下, 根据两个端点间其它点的速度数据对速度多项式进行最小二乘拟合, 可以得到式(4)。

将这 4 个方程联立成速度函数参数解算方程组, 解该方程组即得到多项式的系数, 进而得到速度函数。

$$\begin{cases} V_S = \alpha_3 T_S^3 + \alpha_2 T_S^2 + \alpha_1 T_S + \alpha_0 & (1) \\ V_E = \alpha_3 T_E^3 + \alpha_2 T_E^2 + \alpha_1 T_E + \alpha_0 & (2) \\ S = \int_{T_S}^{T_E} (\alpha_3 T^3 + \alpha_2 T^2 + \alpha_1 T + \alpha_0) dT & (3) \\ \alpha_3, \alpha_2, \alpha_1, \alpha_0 \left| \sum_{i=2}^{N-1} [V_i - (\alpha_3 T_i^3 + \alpha_2 T_i^2 + \alpha_1 T_i + \alpha_0)]^2 \right. \text{最小} & (4) \end{cases}$$

1.3 中心端数据恢复

经过自适应采集得到的 GPS 数据从终端传输到中心端后, 采样点之间的大量数据被过滤, 中心端使用轨迹数据时需要进行数据恢复: 先根据速度函数求出待求点的速度, 然后求取速度函数定积分得到待求点处的里程, 最后根据里程和方向信息解算待求点位置坐标。

(1) 速度解算。设待求点时间为 T_X , 则待求点速度:

$$V_X = \alpha_3 T_X^3 + \alpha_2 T_X^2 + \alpha_1 T_X + \alpha_0 \quad (5)$$

(2) 里程解算。如图 4 所示, 待求点里程 S_X 为阴影部分的面积, 即:

$$S_X = \int_{T_S}^{T_X} (\alpha_3 T^3 + \alpha_2 T^2 + \alpha_1 T + \alpha_0) dT \quad (6)$$

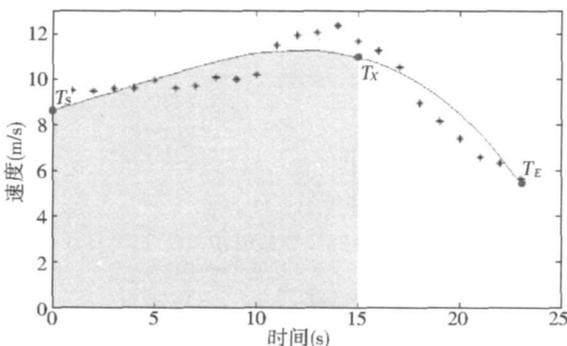


图 4 里程解算示意
Fig. 4 Travel distance calculation

(3) 位置解算。如果所有点在终端都匹配到地图上, 根据路链坐标点序列和 P_X 点处的里程 S_X , 即可解算出 P_X 点坐标, 如图 5a 所示。当点集中有任意点匹配不到路网时, 则速度方向为从起点 P_S 指向终点 P_E 的方向, 求解直线方程即可得到 P_X 点的坐标, 如图 5b 所示。

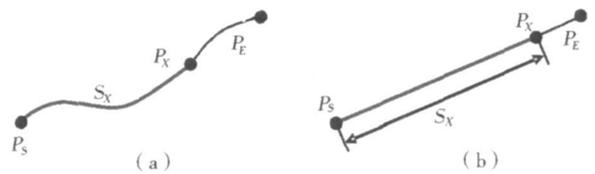


图 5 位置解算示意
Fig. 5 Position calculation

1.4 算法性能评价

在阐述自适应采集算法性能评价指标前, 首先引入速度变化指数 I_v , 用来衡量车辆行驶速度变化的复杂度, 其计算方法为:

$$I_v = w_0 \frac{N_0}{N} + w_1 \frac{N_1}{N} \quad (7)$$

其中, N 为行驶路段内获取的有效 GPS 点数, N_0 为与前一点相比速度变化超过 0.3 m/s 的点的总数量, w_0 为其对应的权重, 一般取 0.4 ; N_1 为与前一点速度变化不相同的点的总数量, w_1 为其对应的权重, 一般取 0.6 。

自适应采集算法性能的主要指标包括位置精度 A_p 、速度精度 A_v 、数据压缩比 J 。

$$A_p = \left(\sum_{i=1}^N D_i \right) / N \quad (8)$$

其中, D_i 为估算点坐标与实际坐标的距离。

$$A_v = \frac{\sum_{i=1}^N |V_{c_i} - V_{r_i}|}{N} \quad (9)$$

其中, V_{c_i} 为估算速度, V_{r_i} 为实际速度。

数据压缩比 J 为自适应采集后的数据量与原始数据量的比值。自适应采集算法性能指标之间相互影响、相互制约, 并且都受到速度变化指数的影响。

1.5 算法应用模式

算法的应用模式可以用图 6 表示。终端接收服务端传送的阈值, 在数据传输被触发后, 对数据进行处理并上传到服务器端, 上传数据帧可表示为 { 车辆 ID(10), 经度(8), 纬度(8), 速度(4), 方向(4), 时间(8), 车辆状态(2), 拟合系数 α_3 (4), 拟合系数 α_2 (4), 拟合系数 α_1 (4), 拟合系数 α_0 (4), 是否与地图匹配(1), 起始匹配路链 ID(4), 终止匹配路链 ID(4) }, 其中小括号内为对应数据项的长度, 单位为字节。数据帧通过 GPRS/CDMA/3G 等无线通信方式以二进制字节流的方式进行传输。

服务器端将数据帧的内容存储在车辆信息数据库中, 在查询车辆轨迹数据时, 根据查询的参数, 依据前文的数据恢复方法得到车辆轨迹的详细数据。

当需要对车辆进行实时监控时, 对于上一个上传点以前的所有数据, 可用数据恢复的方法得到实际轨迹数据; 对于上一个上传点到当前时间的轨迹数

据,可用上传点处的速度和方向进行估计,最好用另外的颜色显示这段轨迹,表明其数据的精度较差;当有新的上传点数据到达时,对这段轨迹进行更新,同时显示为正常颜色。当需要知道监控车辆当前确切的数据时,可发送指令进行数据的强制上传。

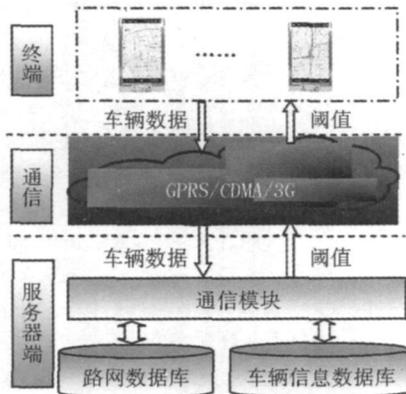


图 6 算法应用模式
Fig. 6 Application mode of the algorithm

2 试验与分析

2.1 试验方法

依据 1.5 节所示的算法应用模式构建试验系统。终端系统的开发与试验终端选用 HTC Diamond2, 操作系统为 Window Mobile6. 5, 采用 Visual C# 2008 语言开发。终端通过 GPRS 采用 Soker 的方式与中心端进行通信。地图数据选用北京市 2007 年冬版本的导航电子地图, 路链数大于 17 万条。

试验中, 在不同时刻不同路段携带终端进行车

辆行驶试验, 设定不同的阈值进行数据采集传输, 并在终端存储所有的原始数据, 在中心端进一步分析。通过分析试验过程和恢复数据的精度, 验证算法的有效性, 并分析算法的影响因素。

2.2 试验结果

(1) 算法应用有效性验证。按设定的速度阈值, 携带终端进行典型路段车辆行驶试验及数据的自适应采集, 并在中心端进行数据恢复, 验证算法应用的有效性。将速度阈值设为 4, 在典型路段进行车辆行驶试验, 速度精度为 0.75, 位置精度为 4.6, 压缩比为 3.0%。其中某一路链上各估算点的速度偏差(为显示方便只取前 25 个点)如图 7 所示。

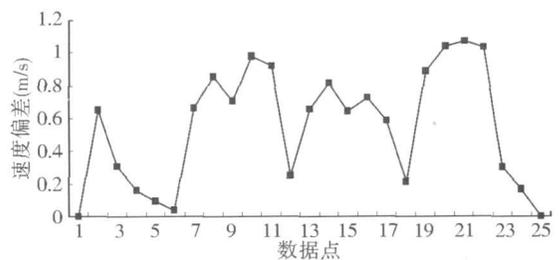


图 7 速度偏差
The illustration of speed deviation

图 8 为数据采集与恢复的试验轨迹。图 8a 为车辆轨迹原始数据, 共 160 个点, 每隔 10 个点标注编号; 160 个原始坐标点采用自适应采集算法后只需要上传 5 个点(图 8b); 图 8c 为恢复后的轨迹图, 恢复轨迹逼近原始轨迹。由以上数据可知算法在实际应用中具有可行性。

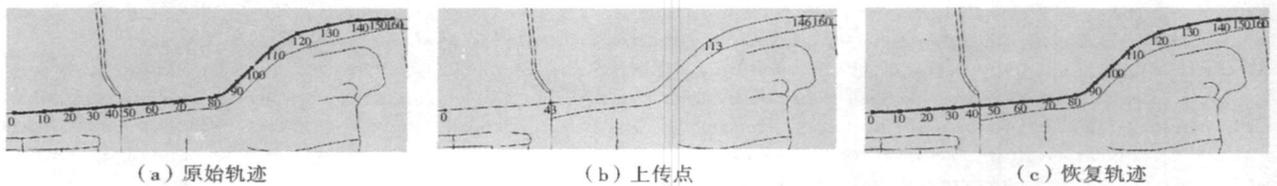


图 8 数据采集与恢复试验轨迹

Fig. 8 Paths in data collection and recovery experiment

(2) 地图约束的影响。分析同等精度要求时有地图和无地图两种情况下数据压缩比的差异。对无地图时的计算, 将路网约束去掉即可。图 9 是当精度要求相同时, 在不同路段有地图和无地图情况下的数据压缩比的对比分析图, 路段的速度变化指数 $I_A < I_B < I_C$, 精度要求取位置精度 4.6 m。

由图 9 可知, 终端安装地图对降低数据压缩比的效果非常明显。随着速度变化指数的增大, 数据压缩比也明显变大。一般车辆导航用 GPS 模块的位置测量精度为 15 m, 测速精度为 0.5 m/s, 但在相对较小的地区相对短的时间内, 影响 GPS 测量精度

的条件基本相同, 各测量点之间的相对误差很小, 如果有地图, 经过地图匹配后可以进一步减小位置误差。所以试验中位置、速度和方向的测量误差对最后

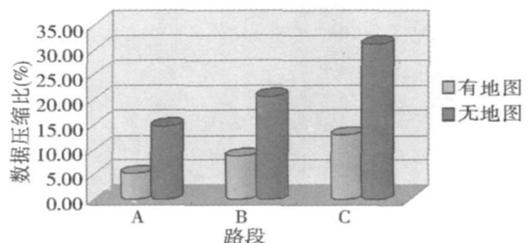


图 9 地图约束对数据压缩比的影响
Map constraints impact on data compression rate

的试验结果影响并不大,得到了较好的速度精度和位置精度。

3 结语

本文综合考虑速度函数表达、终端地图匹配与否、方向偏差度量等因素提出车辆 GPS 监控数据的自适应采集算法,并研究了算法的性能评价指标。通过试验证明了算法具有较高的精度和数据压缩比指标,同时通过数据分析对比可知,终端安装地图对减少数据压缩比的效果非常明显,随着速度变化指数的增大,数据压缩比也明显增大。下一步的研究中应进行速度函数的优化,探索采用更高阶的多项式或其它函数表达速度,以进一步提高精度和数据压缩率指标。

参考文献:

- [1] 张锦明,付永恒.基于GPS/GPRS的位置信息服务系统的设计与实现[J].测绘通报,2007(11):30-33.
- [2] 韦中亚,田原,刘宇,等.构筑于GIS-GPS-GSM技术集成的120急救系统设计[J].地理学与国土研究,2002,18(1):35-36.
- [3] 陈秀万,杨凯,方裕,等.基于空间信息技术的城市应急救援联动系统研究[J].地理学与地理信息科学,2003,19(4):49-52.
- [4] ZHOU T, WANG L. SINS and GPS integrated navigation system of a small unmanned aerial vehicle[A]. Proceedings of 2008 International Seminar on Future Biomedical Information Engineering[C]. 2008. 465-468.
- [5] ZHENG J H, WINSTANLEY W C, YAN L, et al. Economical LBS for public transport: Real-time monitoring and dynamic scheduling service[A]. The 3rd International Conference on Grid and Pervasive Computing Symposia/ Workshops[C]. 2008. 184-188.
- [6] CHEN Y K, LIU Y C. A new method for GPS-based urban vehicle tracking using Pareto frontier and fuzzy comprehensive judgment[A]. Geoscience and Remote Sensing Symposium[C]. 2007. 683-686.
- [7] AYALA D, LIN J, WOLFSON O, et al. Communication reduction for floating car data based traffic information systems[A]. 2010 Second International Conference on Advanced Geographic Information Systems, Applications, and Services[C]. 2010. 44-51.
- [8] LIU D, LIANG Z W, LI Y C, et al. Research and design of a high performance GPS vehicle monitoring system[A]. 2010 International Conference on Future Information Technology and Management Engineering[C]. 2010. 59-62.
- [9] 李清泉,黄练,谭文霞.基于道路特征的海量GPS监控数据压缩方法[J].武汉大学学报(信息科学版),2008,33(4):337-340.
- [10] BRAKATSOU LAS S, PFOSE D, TRYFONA N. Practical data management techniques for vehicle tracking data[A]. 21st International Conference on Data Engineering[C]. 2005. 324-325.
- [11] 王卫京,翁敬农,樊珂.车辆监控系统中时空数据模型设计与实现[J].计算机工程与设计,2006,27(6):1042-1044.
- [12] 王生生,刘大有.基于移动对象模型的车辆监控系统[J].仪器仪表学报,2004,25(3):509-511.
- [13] 尹建忠,贺奋琴,李清泉,等.导航监控一体化系统中数据流传输控制算法研究[J].武汉大学学报(信息科学版),2008,33(9):900-903.

Study on Adaptive Collecting Algorithm of Vehicle GPS Monitoring Data

HAN Xiao-gang, YAN Lei, LU Min

(Institute of Remote Sensing and Geographic Information System, Peking University, Beijing Key Lab of Spatial Information Integration & Its Applications, Peking University, Beijing 100871, China)

Abstract: In large scale high precision vehicle monitoring with GPS, the communication between the server and terminals is extremely frequent, and data processing load is heavy. In order to solve this problem, an adaptive collecting algorithm of vehicle GPS monitoring data is proposed in this paper. The algorithm is designed considering three aspects: expresses vehicle speed function by polynomial fitting based on constraints of endpoint speed and travel distance, takes both situations of the terminal with a map and without a map into account, and uses deviation distance to indicate the direction deviation of speed. Based on the theories above, the adaptive data collecting algorithm is established so that GPS data is sent adaptively from terminals and recovered in the server automatically. Also performance evaluating indicators of the algorithm is discussed. Finally, the algorithm is validated based on the real road network and measured GPS monitoring data of Beijing. The experiment shows that the recovered data has a good approximation in speed and position, and the compression rate of GPS data reaches 3%, which proves the effectiveness of the algorithm. Further research should be conducted to optimize the speed function, explore higher order polynomial function or other speed expression, to improve the accuracy and data compression rate in adaptive GPS data collection.

Key words: GPS; adaptive; collection