

智能 Agent 在行车记录分析系统上的应用

Using Agent in Mobile Track Data Analyzing

陈红英 毛革非 (广州华南师范大学计算机系 510631)

摘要: 本文介绍了行车记录仪关键组成技术, 以及智能 Agent 行车记录仪系统轨迹处理上的优化作用。为智能 Agent 在行车记录领域上的应用提供了思路, 是智能交通的一个组成部分。

关键词: 行车记录仪、GIS、GPS、智能 Agent

1 前言

本系统介绍的综合管理型记录行车记录仪, 通过GPS数据采集、报警界限设定、GIS数据拟合、位置计算、司机车辆等基本信息查询、行驶记录查询、分析预测、报表输出等一系列步骤, 实现了对车辆的全程记录、自动报警、后台管理分析预测等功能。

目前, 采用以GPS获得定位信息、以flash为存储介质的行车记录仪, 通常要记录一周以上的数据。以0.1秒记录一条状态数据, 保存10分钟(用于事故分析), 和5秒记录一个轨迹点, 保存一周(用于轨迹分析)来算, 一辆车行驶一周要处理12.68万条数据。而一个中等城市至少要划分100个区域, 每个区域平均用200组坐标表示边界, 那么, 一辆车行驶一周要计算 $100 \times 200 \times 12.68 = 25.36$ 亿次。一个拥有1000台机动车的车队每周要分析全部车辆的运行情况需要计算2.5万亿次。如何采取有效的技术手段, 针对不同车辆运行的“脾气”, 自动制订出可自适应的算法, 使得的行车记录仪轨迹处理速度的瓶颈问题得到改善, 以达到整体处理分析的优化呢? 采用智能Agent技术就是最好的解决办法。

2 系统涉及的关键技术

本系统主要涉及四个关键技术: GPS技术、智能Agent、行车记录仪、地理信息技术(GIS)。

2.1 GPS 技术

GPS源于20世纪初美国的“地基电子导航系统”。GPS全球卫星定位技术是通过距地球2万多公里的24颗人造卫星向地球不断发射定位信号, 由地球上的GPS接收机收到卫星信号并进行计算后就可以得到GPS接收机的经纬度、高度等地理信息及运动的速度和方向。是高

精度、全天候、全球性的无线导航、定位系统。

2.2 Agent 特性

- (1) 代理性: “代理”用户或软件完成某些任务。
- (2) 主动性: 可根据用户需求或当前环境状态触发其主动服务。
- (3) 自主性: 可根据当前动态变化的环境状态, 无需外界干扰, 独立地发现和利用完成任务所需的资源和服务, 独立制订完成任务的方案, 并独立完成任务。
- (4) 智能性: 当用户需求不太明确时, 智能Agent可推测用户意图、爱好或兴趣、规律, 并为其代劳, 从经验中不断学习, 以提高自身处理问题的能力。

2.3 记录分析系统

本系统采用综合管理型记录仪, 通过GPS数据采集、报警界限设定、GIS数据拟合、位置计算、司机车辆等基本信息查询、行驶记录查询、分析预测、报表输出等一系列步骤, 实现了对车辆的全程记录、自动报警、后台管理等功能。

2.4 地理信息技术(GIS)

本系统的电子地图平台采用的是MapInfo, 可视化软件开发使用的是MapInfo公司的MapX。本系统中, 用于轨迹再现的可视化图层共有水塘、水系、街区、主要道路、次要道路、河流等28个, 用于轨迹计算的主要有三个图层: 边界、街区、道路。

3 行车记录分析系统的实现

3.1 系统流程

如图1所示, 本系统主要由数据采集、实时报警、轨迹处理、数据挖掘四个部分组成。

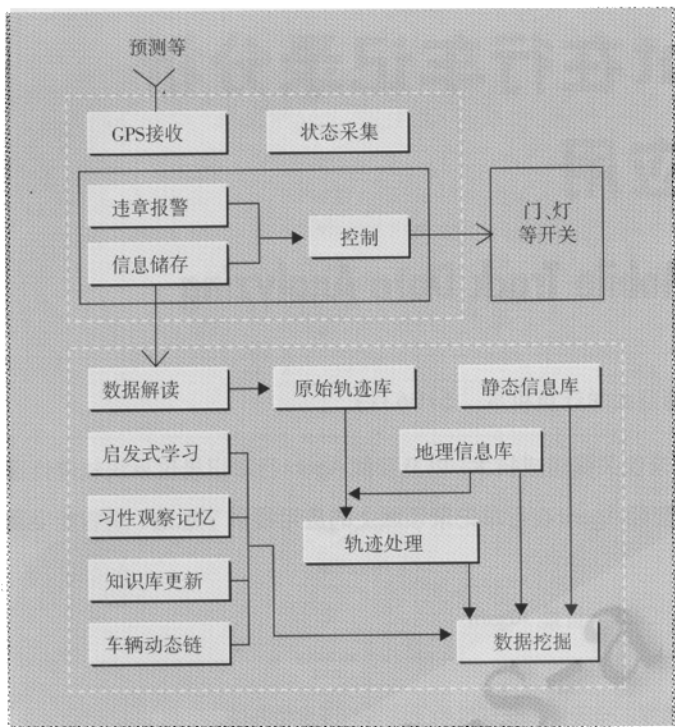


图1 系统流程图

顺序,各区域名间用“.”隔开;此外还有一个每日时间片字段,用于记录出发的正向和逆向(如,从家里出发为正向,返回家里为逆向)。如图2,假设地图上分1、2、3...共七个区域,车辆“京A 88888”经常走的路线是1-4-5-3,那么,计算其轨迹时,如轨迹点的时间为正向,则首先判断区域1,若干点后,当轨迹点离开区域1时,直接判断区域4,以此类推,最后判断区域3,逆向时反之。这样,每个轨迹点不用遍历近7个区域,而是基本上只判断一个区域。

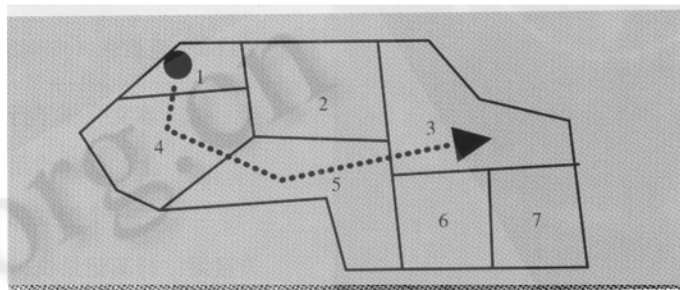


图2 区域路线图

每次轨迹的计算后,根据区域权值频度库,我们进行知识更新,自动更新路线规律列表的顺序。

(2) 基于区域权值、信息频度的学习方法。对于出租车、货车等行驶路线随机性很大的车辆,因为一段时间内,它们在各个区域中出现的频率趋于稳定。对此,我们建立区域权值频度库,用于记录各个车辆用户在各个区域内出现的次数。其结构是:车牌号、区域频度有序列表,其中区域频度有序列表结构如下:区域名1+次数, ...区域名n+次数,名称和次数用“+”连接,不同名称用“,”分隔。处理时优先判断是否具有路线规律,当路线规律分值很低时,首先判断是否在出现次数最多的区域,超出这个区域时,优先判断这个区域的邻接区域,如不在,则判断出现次数第二的区域,等等以此类推。知识更新:每次轨迹的计算,自动更新路线规律列表的顺序。但在最坏情况下,可能退化为 $m \cdot n \cdot k$ 。其中m为车辆轨迹记录数,n为城市区域个数,k为每个区域的平均边界点数。

(3) 基于多维检索的空间信息查找方法。GIS库的数据量很大,仅一个中等城市的一个图层(如区域),一辆车一周最坏情况下就要计算25亿次,采用常规的一维检索进行地理位置的判断不能满足要求。在此,我们采用多维检索方法,以区域图层为例,按照区域类别、区域名称、经度、纬度等分量组合查找,对于每个轨迹点,首先确定其类别,然后以这个点为中心,以此类别的范围为半径,查找一个圆形覆盖内的地理特性,从中取出区域名变量值,即完成了这个轨迹点的位置判断。

(4) 基于知识库的快速推理机制。知识库中,除了车辆轨迹、区

(1) 数据采集包括:经度、纬度、速度、方向、时间等基本信息,通过gps定位系统采集。点火、熄火等状态信息,左右灯、温度等控制信息,通过传感器采集。

(2) 实时报警包括:超速、超时、怠速等报警。报警界限设定:超速、超时、怠速。

(3) 轨迹处理包括:每个轨迹点的位置判断、超速超时等违章判断、事故前10分钟的数据分类等。

(4) 数据挖掘包括:行驶区域分析、常规路线分析、物流匹配分析、运营重点。

3.2 采用的 Agent 技术

整个系统中,时间开销关键由轨迹计算部分决定。对此,我们采用了知识库技术,建立了常规形势路线库、行驶区域频率库、节假日动态库等,利用智能Agent的代理性、自主性、智能性、机动性等功能力,自动搜集、整理、学习、感知各个车辆用户的兴趣和规律。从而使得行车记录的每个轨迹点,不用遍历整个地理信息库,而是带智能的只查找几个地理区域,并且在查找了第一个区域后,还能判断第二个、第三个...待查区域的顺序,大大缩小了查找范围,提高了效率。

(1) 基于路线观察记忆的学习方法。私家车、客运车、长途货运车、单位用车等,每天的行驶路线具有较大的规律性,据此,增加常规行驶路线库,用于记录各个车辆用户的行驶路线规律。此库以车牌号(具有唯一性)为主索引,以路线规律列表为次索引。路线规律列表的结构为:区域1, ...区域n,为有序表,存储区域遍历

域权值频度、路线规律等动态信息和地理信息外,还有司机情况,车辆情况,油耗、吨公里等参数情况等静态信息。要保持知识库的快速反应能力,必须实时地对知识库进行知识更新、知识遗忘、知识关联和知识表述。

(5) 基于逻辑推理的电子地图自动切换。GIS信息库通常做法是:一个行政区域(如一个城市及其所属的县、乡)做成一个图集,每个图集又包含若干个图层。由于受硬件环境的限制,一次打开的图集不能太多,那么,当轨迹点不在当前打开的图集上时,电子地图要有自动图集切换功能。在此,我们采用树型结构,如图3。根节点是全国略图,第二层为各个省略图,再往下是一些地理位置相邻的城市组成的片图,最后节点是各个城市详图。轨迹判断时,当连续多个点超出在现有图集时(因为GPS在某些情况下存在“飞点”,所以区域改变时要多考虑几个坐标位置),首先在全国略图中找出它所在的省,然后在这个省的略图上找出它所在的片,第三步是在片图上找出城市,最后在某市详图上定位到所在的区域。

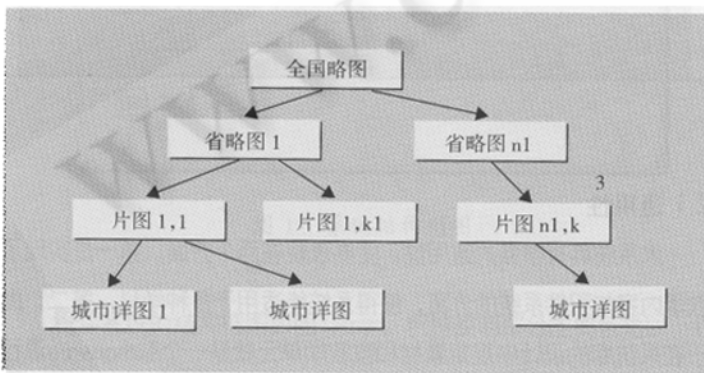


图3 电子地图搜索采用树形结构

(6) 基于用户兴趣的主动服务。利用智能Agent的代理性、主动性、交互性等特点,我们增加了用户兴趣模块。主要包括:

固定信息服务:根据所有车辆、驾驶员的档案记录,定时提醒车辆该年审了、驾照该年审了、汽车该大修了等;

变动信息服务:建立节假日动向库,每次轨迹操作时,自动更新此库的信息,根据这个库的信息和常规路线库,节假日快到时,自动提醒用户是否到某地方去,并可给出路线。

(7) 基于责任分析的操作员记录反馈。任何操作和车辆行驶都具有记录,以进行责任分析、成本分析、事故分析、发展分析。责任分析是针对软件使用者的一种记录分析,操作员每进行一种操作,都进行操作合法性判断,并在事件库中加以记录。如进行行车记录仪初始化操作时,若将超速值设在110公里/小时以上,则软件首先会自动提示:“超速设置太大,是否继续?”,继续,则谁为那台车设置的这

个超速值自动记录到事件库。这样,可规范操作员行为,保障各项关键参数的设置的正确性。

(8) 基于深度优先遍历的区域查找。当车辆的行驶毫无规律时,我们无需将所有区域都一一遍历来判断轨迹点的位置。针对地理信息库的详细图层,我们建立了如下拓扑结构:以图的邻接表的方式存储各个区域间的结构。如针对图2所示的区域分布,我们形成如图4的拓扑结构,建立如图5的邻接表结构。通常电子地图上,每个区域的直接邻居不超过四个,那么采用图的邻接表结构、图的深度优先遍历方法,除了第一个轨迹点外,每个轨迹点的位置判断不超过四个区域,即时间复杂度趋于 $O(m*4*k)$,对于拥有100个分区的图层的计算效率将提高 $100/4=25$ 倍。

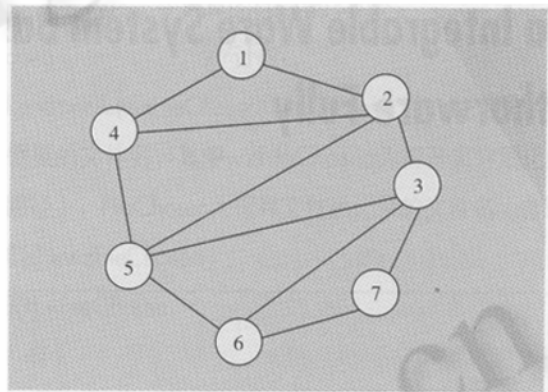


图4 图2的区域拓扑图

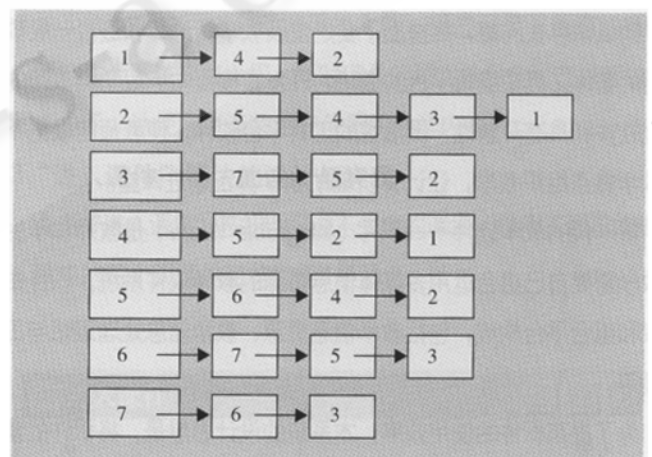


图5 图4的邻接表结构

(9) 启发式学习方法。通过观察各个车辆用户在不同区域、不同时间片的出现频率、行驶路线规律和电子地图的拓扑结构相结合,程序自动调整一个轨迹的计算策略,同时学习车辆用户的规律,用的时间越长,和车辆用户的“理解”就深,从而达到时间效率的最优。

4 效率分析和程序运行

经测试,对于行驶具有一定规律的车辆用户来说,遍历时间可由 $O(m*n*k)$ 改善为 $O(m*k)$,其中 m 为车辆轨迹记录数, n 为城市区域个数, k 为每个区域的平均边界点数,效率提高了 n 倍。

如前言,采用智能Agent以前,一辆车行驶一周要记录12.68万条数据,计算的时间复杂度是 $O(25.36$ 亿次),耗时约10分钟。一个拥有1000台机动车的车队每周要分析全部车辆的运行情况需要计算2.5万亿次,耗时约166小时,时间开销巨大。采用智能Agent以后,同样是12.68万条记录,平均两个区域即可定位,那么计算的时间复杂度为 $O(0.5$ 亿次),耗时约12秒,1000台机动车的车队仅需3小时即可完成分析计算,极大地提高了效率(注:以上计算采用的机器是奔腾3-866M)。

5 结束语

本系统在行车记录分析系统中引入了智能Agent技术,使得轨迹处理的瓶颈问题得到改善,使大型车队,大容量轨迹处理计算成为可能。希望本文能起到抛砖引玉的作用,为智能Agent在行车记录领域上的应用提供思路。

参考文献

- 1 GIS在交通信息管理系统中的应用,霍宏、胡福乔,计算机工程,2002,5。
- 2 计算机与GPS/OEM板通讯的研究与实现,林世究、杨春金、赵昂,计算机应用研究,2002,6。