

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2018 年 7 月 5 日 (05.07.2018)



(10) 国际公布号
WO 2018/122803 A1

(51) 国际专利分类号:
G06F 17/30 (2006.01) **G08G 1/01** (2006.01)

(21) 国际申请号: PCT/IB2017/058533

(22) 国际申请日: 2017 年 12 月 30 日 (30.12.2017)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
PCT/IB2016/058105

2016 年 12 月 30 日 (30.12.2016) IB

(71) 申请人 (仅对 CN): 同济大学 (TONGJI UNIVERSITY) [CN/CN]; 上海市曹安公路 4800 号 (CN)。上海智能交通有限公司 (SHANGHAI INTELLIGENT TRANSPORTATION CO., LTD.) ; 上海市上海市杨浦区周家嘴路 1229 号 A 座 4 楼 (CN)。

(71) 申请人 (对除 CN 外的所有指定国): 许军 (XU, Michael Jun) [CN/FR]; 里昂 33 CRÉQUI 路 (FR)。

(72) 发明人: 杜豫川 (DU, Yuchuan); 中国上海市同济大学交通运输工程学院, Shanghai (CN)。邓富文 (DENG, Fuwen); 中国上海市同济大学, Shanghai (CN)。

(72) 发明人 (仅对 CN): 常光照 (CHANG, Guangzhao); 上海市上海智能交通有限公司 (CN)。

(74) 代理人: 许军 (XU, Michael Jun); 中国广西壮族自治区南宁市园湖南路 15 号, Guangxi 530022 (CN)。

(81) 指定国 (除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

(54) Title: SMART ROAD TRAFFIC ANOMALY DETECTION METHOD

(54) 发明名称: 一种智能化道路交通异常检测方法

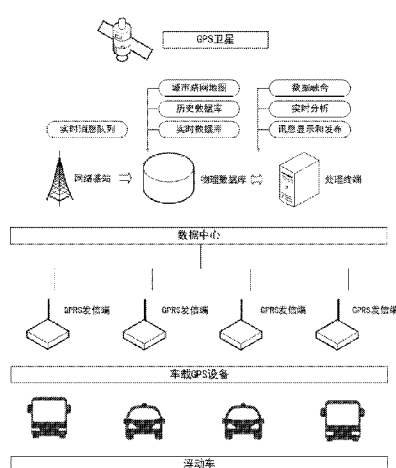


图 1

(57) Abstract: A smart road traffic anomaly detection method, comprising the following steps: establishing time/space sub-segments; preprocessing past tracking data; analyzing past tracking data analysis and training an RNN model; detecting anomalies; indicating the degree of severity of the anomalies; and evaluating system performance. In the method, vehicle tracking information is recorded by means of on-board GNSS positioning devices in floating vehicles driving on roads, and said tracking information is analyzed and mined in order to achieve smart detection of traffic anomalies and incidents on city roads.

(57) 摘要: 一种智能化道路交通异常检测方法, 包括如下步骤: 建立时空子区; 历史轨迹数据的预处理; 历史轨迹数据分析和 RNN 模型训练; 异常检测; 异常严重量化表征; 系统性能评价。利用道路上行驶的浮动车车载 GNSS 定位装置, 记录车辆的轨迹信息, 通过对这些轨迹信息的分析挖掘, 能够智能化检测城市道路交通异常事件。

AA GPS satellite
BB Real-time message queue
CC Map of city road network
DD Database for past data
EE Database for real-time data
FF Data fusion
GG Real-time analysis
HH Information display and distribution
II Network base station
JJ Physical database
KK Processing terminal
LL Data center
MM GPS transmitter
NN On-board GPS devices
OO Floating vehicles

(84) 指定国 (除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

- 包括国际检索报告 (条约第21条(3))。
- 在修改权利要求的期限届满之前进行, 在收到该修改后将重新公布 (细则48.2(h))。
- 黑白: 提交的国际申请包含彩色或灰度, 并且可通过PATENTSCOPE下载

一种智能化道路交通异常检测方法

技术领域

本发明属于交通检测技术领域。特别地，本发明涉及一种城市道路交通异常实时检测方法。通过浮动车的车载 GNSS 定位装置，可获取其不同时刻的空间位置信息，经过数据预处理、地图匹配和数据融合，基于特定时空范围的行程车速历史数据训练 RNN 模型；根据 RNN 模型预测值与实时行程车速真实值间的差异，可有效识别城市道路交通异常事件。

背景技术

交通异常事件检测是城市交通管理的重要组成部分，也是智能交通系统的核心功能之一。交通异常事件主要包括交通事故、车辆抛锚、货车落物、道路交通设施损坏或故障以及其他造成交通流紊乱的特殊事件。该类事件容易造成交通拥堵、路段通行能力降低，严重时影响整个道路交通系统的正常运行。通过交通异常事件检测，可使交通管理者及时了解交通异常信息，并采取适当的诱导和控制措施，降低交通异常事件的不良影响。

交通异常事件检测可分为人工方式和自动方式。人工方式包括巡逻车、紧急电话上报和视频监控等，由于消耗人力物力且实时性差，无法满足交通管理的需要。自动方式依靠自动事件检测（AID，Automated Incidence Detection）算法实现，基本原理是通过检测不同位置道路交通流的变化来识别交通异常事件。目前常用的 AID 算法包括模式识别类算法（如 California 算法、莫妮卡算法）、统计预测类算法（如指数平滑法、卡尔曼滤波算法）、交通流模型算法（如 McMaster 算法）以及智能识别算法（如人工神经网络、模糊逻辑算法）。

但是目前的检测方法存在对设施的要求高、计算复杂度高、无法对异常状况的态势做进一步判断等缺点。本发明利用出租车、公交车车载 GNSS 定位装置回传的轨迹数据，建立历史交通状态数据库和实时交通状态数据库，通过分析两者反映的交通流特征差异，识别交通异常事件。该方法具有实时性好、可并行处理、识别率高以及对检测设施要求低等特点，适用于有实时浮动车定位数据的数据环境下城市道路交通异常事件的检测。

目前，针对交通异常事件监测，有以下代表性技术：

一件美国专利申请，US 20160148512，披露了一种交通异常事件检测和上报系统的组成原理和实施方式。该系统由传感器、通信模块、移动处理模块和用户交互模块组成。传感器用于采集车辆周边的相关数据；通信模块用于发送本车辆数据和接收周边车辆的数据；移动处理模块用于处理和分析相关车辆在某一区域内的数据并生成交通事件报告；用户交互模块能够像用户提供交通事件报告。该方案是一种基于车车和车路通讯网络的交通异常事件检测技术，能够利用传感器采集的各类信息，判别异常事件。然而，由于传感器、通信单元需要单独安装调试，实施难度较大；移动处理单元处理能力受限；同时需要移动和固定的讯息接收端，系统本身存在故障概率，可靠性不佳。

一件中国专利申请，CN 104809878 A，披露了一种利用公交车 GPS 数据检测城市道路交通异常状态的方法。该方案根据 GPS 历史数据获得路段延误时间指数，根据 GPS 当前数据获得瞬时速度、周期平均速度、加权滑动平均速度和多车平均速度，利用规范变量分析算法检测异常。这一方案不需要新

增检测设施，实施便利。但是对于交通态势的表征过于简化，无法分析交通异常状况的特点和成因；对交通场景的划分缺乏依据，未能考虑天气等因素对交通态势变化的影响。

发明内容

为了更清晰地阐述本发明的内容，首先将涉及到的专业术语解释如下：

浮动车：也称探测车。指安装了车载定位装置并行驶在城市道路上的公交汽车和出租车。

GNSS：全球导航卫星系统（Global Navigation Satellite System）。包括 GPS、GLONASS、GALILEO 以及北斗卫星导航系统等。

时空子区：按照时间和空间两个维度划分的片区，反映在一段时间内，一定的空间范围内的情况。将一天划分为若干时间片段，例如 0:00-0:10, 0:10-0:20……，称之为时间子区；将城市道路交通异常检测的实施区域划分为若干空间片段，例如经度 121.58° E-121.59° E，纬度 31.16° N-31.17° N 之间的区域，即空间子区；任意一个时间子区和任意一个空间子区的交集形成的时空片段，称为时空子区，例如经度 121.58° E-121.59° E，纬度 31.16° N-31.17° N 之间的区域在 0:00-0:10 的时空片段。

历史轨迹数据：历史轨迹数据是长时间积累并存储在数据库中的轨迹数据。历史轨迹数据是动态变化的数据，需要及时更新，并定期做重新处理和分析，以保证历史交通特征提取的准确性。每个时空子区的数据可以并行处理以提高效率。本发明中可简称为历史数据。

实时轨迹数据：实时轨迹数据是距离当前时刻最近的一个时间区段内的轨迹数据集合。本发明中可简称为实时数据。

交通态势：一定时间、一定空间内交通运行的综合情况的总称。

交通异常：交通事故、车辆抛锚、货车落物、道路交通设施损坏或故障等事件引发的交通流紊乱的情况。

交通异常严重性：即交通流紊乱的严重性，是正常状态下交通流与交通异常发生后交通流特征的差异。

交通异常指数：交通异常严重性的量度。范围为 0~10，数值越大，交通异常越严重。

交通环境：作用于道路交通参与者的所有外界影响与力量的总和。包括道路状况、交通设施、地物地貌、气象条件，以及其他交通参与者的交通活动。

地图匹配：将地理坐标与城市路网关联的过程。

高峰小时流量：某城市道路断面一日内小时交通流量的最大值。

响应变量：根据自变量发生改变的变量，也称因变量。

RNN：即循环神经网络（Recurrent Neural Network），是一种节点定向连接成环的人工神经网络，这种网络的内部状态可以展示动态时序行为，可以利用它内部的记忆来处理任意时序的输入序列。

Elman-RNN：一种 RNN 网络结构，参见《A RNN that learns to count》。

训练过程：通过迭代计算来优化神经网络参数，使神经网络在训练数据集上的模型误差降低的过程。

本发明的目的是建立一套基于浮动车轨迹记录系统，利用历史 GNSS 定位数据和实时 GNSS 定位数据，结合交通环境信息识别道路交通异常事件的方案。为了达到上述目的，本发明提供了如下技术方案：

本发明的实施前提是：搭载 GNSS 轨迹记录仪的浮动车（出租车、公交车等）；具有大规模存储、

计算、实时任务处理能力的数据中心。

本发明的适用范围是：有上述浮动车经过的城市道路（包括地面道路和高架道路）。

本发明的实施步骤包括：

1) 确定检测的时空范围和建立时空子区。

基于实际的应用需求，确定需要进行交通异常事件检测的时间范围和空间范围。时间范围可以设定为全天，即 0:00-24:00；也可以设定为某一特定的时段，例如要检测 17:00-20:00 这个时段的交通异常时间，则将检测时间范围设定为 17:00-20:00，这里只是列举一个特殊实例，还有很多其他情况，此处不再一一说明。空间范围可以按照行政区划设置为某个市域，例如北京市、上海市、黄浦区等；也可以按照城市空间结构设置为某个城市功能区，例如某市中央商务区、工业区等。

时空子区的建立是指，将检测的时间范围划分为若干个更小的时间片段，将检测的空间范围，即城市道路交通异常检测的实施区域，划分为若干个更小的空间片段。时空子区的建立，可以采用多种经验划分方法，包括等距时空划分法和非等距时空划分法。

2) 数据预处理。

将 GNSS 定位数据进行数据清洗、数据集成、数据转换、数据归约，提高数据的结构化程度。GNSS，即全球导航卫星系统定位系统，是能在地球表面或近地空间的任何地点为提供全天候三维坐标和速度以及时间信息的空基无线电导航定位系统。它主要包括美国的 GPS（Global Positioning System），俄罗斯的 GLONASS（Global Navigation Satellite System）、欧盟的 GALILEO 和中国的北斗卫星导航系统四大全球性导航定位系统，同时还包括日本的 QZSS、印度的 IRNSS 等区域导航定位系统以及美国的 WASS、日本的 MSAS 等卫星定位增强系统。为了在不同的导航定位系统设备中建立统一的数据分发标准，美国国家海洋电子协会制定了统一的 NEMA（National Marine Electronics Association）通讯协议，以规范 GNSS 的数据广播。因此，GNSS 中的各个成员系统，例如 GPS、GLONASS 等，虽然分别由不同国家和机构建立和维护，但是拥有一致的数据分发格式，因此不需要对数据格式进行变换。

选定的空间范围内，有许多安装 GNSS 定位设备的车辆，常见的有出租车、公交车、货运汽车、私家车等。基于当前城市交通数据应用现状，在实际应用当中，通常选用城市出租车为浮动车作为交通异常检测系统的数据来源。

采集的 GNSS 定位信息中包含一些不合理的信息，为了保证交通异常状态检测判别结果的准确性，首先需要进行甄别以提出异常的数据，保证数据的可靠性。这些异常数据包括：落在检测时空范围之外的数据、明显超出合理范围的空间位置跳跃。所谓“明显超出合理范围的空间位置跳跃”，下面举例说明之。若某日 10:30:00 时刻某辆浮动车定位设备上传的定位点记为 A，当日 10:30:30 时刻该浮动车定位设备上传的定位点记为 B，位置 A 与位置 B 的距离为 1500 米，那么据此计算得到该浮动车的行驶速度至少为 180km/h，超出了一般常识，因此是一种异常的空间位置跳跃，数据处理中应当予以剔除。

3) 快速地图匹配。

经过预处理后的 GNSS 定位数据，需要结合城市路网数据，通过地图匹配算法，将 GNSS 定位点投影到城市地图，建立定位点与路段的匹配关系，并修正定位漂移带来的误差。

目前各个地理区域的电子地图都已较为详实，这种电子地图可以来源于城市的地理信息系统，当然也可以来源自其他方式和途径。这些电子地图对城市道路信息进行了详细刻画，通过划分可以得到若干路段。通过借助距离、角度等信息，将定位点匹配到路段上，这样就实现了将定位信息匹配到实际的地理环境中。

4) 浮动车路径的表示和不同车辆路径的匹配。

在给定一组起终点的前提下，车辆的路径可能不是唯一的。复杂的城市交通路网包含了若干路段，将这些不同的路段进行编号，例如，将路段表示为 L1, L2 等。道路可能有两个不同的行驶方向，在这种情况下，应该将两个不同的行驶方向表示为两个不同的路段，给予不同的路段编号。

给定的起点和终点，通常可采用城市路网中路段的交点。已知某浮动车行驶的路径，现需要从其他浮动车已经发送的路径信息中，选择与该浮动车路径相同的路径，从而获得起点和终点间的同路径组。

5) 数据抽样。

浮动车的定位数据中，包含位置坐标、瞬时车速、记录时间等信息。在本专利提出的基于浮动车数据的城市道路交通异常检测方法中，数据抽样是指从全部的浮动车数据中筛选出部分数据进行后续的分析处理，这种筛选是基于数据中心的计算能力以及预先提出的精度要求而进行的。基于不同的计算能力和精度要求，可采用不同的数据抽样方法。例如，当数据中心的计算能力较强，且对检测的精度要求较高时，可以将全部的浮动车定位数据作为处理对象，进行全面的处理分析；而当数据中心的计算能力有限时，假定当前的数据中心能够在 1 分钟内，对每个空间子区处理 500 条数据，而实际情况是在 1 分钟每个空间子区能产生了 2000 条浮动车定位数据，那么可以从 2000 条数据中随机抽取 500 条数据进行分析，从而在数据中心的计算能力范围内，获得精度受限的处理结果。

根据对浮动车数据利用方式的不同，可以针对浮动车数据的不同属性进行采样，例如行程车速和行程时间等。本专利中提出的基于浮动车数据的城市道路交通异常检测方法，采用行程车速作为基础进行城市道路交通异常检测。因此，数据抽样是指对行程车速进行抽样。

6) 历史轨迹数据分析和 RNN 模型训练。

所谓历史轨迹数据，是指在长期的城市道路交通运行中积累下来的浮动车轨迹数据。利用历史浮动车轨迹数据，可以建立城市道路交通特征模型，用来反映城市交通运行的一般特性。这里所说的城市道路交通特征模型，可以指某些特定的指标，例如平均速度、加权平均速度等；也可以指各种某种统计模型，例如行程速度的概率分布。以往的很多模型，采用单一的指标来表示某个路段或区域的交通特征（如历史平均车速），这种方式虽然应用简便，但是精度不高，敏感性差，往往不能在交通异常状态检测中发挥良好的效果。因此，本专利提出对于每个时空子区，用交通特征变量的变化情况进行 RNN 模型的训练。该 RNN 模型可在给出前若干时段交通特征变量真实值时，给出下一时刻交通特征变量的预测值。

可采集的交通特征变量，包括行程车速和行程时间等，本专利采用行程车速反映交通运行特征，故交通特征变量指行程车速。

7) 实时轨迹数据分析和特征提取。

所谓实时轨迹数据，是指距离当前时刻不远的一段时间内的交通运行中浮动车的轨迹数据。利用实时浮动车轨迹数据，可以掌握交通特征的变化动态，用来反映当前交通运行的即时特性。本专利采用当前时空子区的行程速度描述当前交通特征。

8) 异常检测。

系统状态异常检测的思想最早由 Denning 提出，即通过监视系统审计记录上系统使用的异常情况，可以检测出违反安全、可能引发系统异常的事件。Denning 建立的这种模型独立于任何特定的系统、应用环境、系统弱点、故障类型，因而是一种普遍意义上的异常检测模型。该模型包括主体、客体、审计记录、轮廓、异常记录和活动规则 5 个部分。轮廓是用度量和统计模型来表示的主体相对于客体的

正常行为。Denning 的模型定义了 3 种度量，即事件计数器、间隔定时器、资源测量器，并提出了 5 种统计模型，即可操作模型、均值和标准差模型、多变量模型、马尔可夫过程模型和时间序列模型。Denning 提出的模型通过对系统审计数据的分析，建立起系统主体的基于统计的正常行为特征轮廓，检测时，系统中的审计数据与已建立的主体的正常行为特征轮廓相比较，若相异部分超过某个阈值，就认为是一个异常事件。该模型奠定了异常检测的基础，以后发展的许多异常检测方法和系统都是以它为基础而发展起来的。

近几年在异常检测技术的发展过程中，引入了更多人工智能的方法，以提高异常检测的性能。这些人工智能的方法主要包括数据挖掘、人工神经网络、模糊证据理论等。数据挖掘的方法用来确定在大量的数据集合中什么特征是最重要的。该技术用于异常检测中主要是寻求一种正常模式更简洁的定义，而不是像传统的异常检测方法那样简单列举出所有的正常模式。数据挖掘方法的引入使得检测系统能仅通过识别正常模式中的主要特征，就能够概括性地包括训练数据中所未包括的正常模式。人工神经网络异常检测问题可被看作是一个一般的数据分类问题。在前面谈到的统计异常检测中，用户行为数据按照某种统计准则被分为两类：即异常行为和正常行为。由于基于统计的方法在提取、抽象审计实例时存在一定困难，可能造成较大误差，必须依赖于一些概率分布假设，一般需要凭经验和感觉来刻画用户行为的度量，所以引入了人工神经网络的聚类方法。人工神经网络具有自学习自适应能力，用代表正常用户行为的样本点来训练神经网络，通过反复多次学习，神经网络能从数据中提取正常的用户或系统活动的模式，并编码到网络结构中，检测时，将审计数据通过学习好的神经网络，即可判定系统是否正常。由于异常的评判标准具有一定的模糊性，所以模糊证据理论被引入到异常中，如建立一种基于模糊专家系统的入侵检测框架模型，能较好地降低漏警率和虚警率。

本专利提出一种基于循环神经网络的异常检测方案，特别地，是一种采用 RNN（循环神经网络）的异常检测方案，对实时轨迹数据输入基于历史数据的 RNN 模型，得到预测值，再比较预测值与真实值之间的差距。该方案采用了深度学习方法，能够随着时间的推移自动更新模型，具有较强的自适应能力。

9) 异常严重性量化表征及异常信息发布。

交通异常状况的严重性应该通过简洁明了的方式向公众发布，以避免可能的拥堵区域，提高城市交通的运行效率。异常状况的严重程度用交通异常指数表征，范围为 0-10，其中 0 表示无异常，10 表示高度异常。

异常的发生位置投影到电子地图上，并通过智能移动设备 APP 等形式公开发布。

10) 系统性能评价。

系统性能的评价是指评价交通异常状态检测的准确性，其评价指标包括误报率和漏报率。误报率和漏报率越低表明系统的性能越好。

所述步骤 1) 中，时空子区的划分具体可以采用以下方法：

11) 等距时空划分法。确定时间维度的片段尺度，时间片段跨度为固定值，通常取 30min 作为一个时间片段；确定空间维度的片段尺度，空间片段跨度为固定值，通常取 $200\text{m} \times 200\text{m}$ 的空间网格作为一个空间片段。

12) 非等距时空划分法。对于路网密度大于 $2\text{km}/\text{km}^2$ 或高峰小时流量大于 1000 辆/小时的城市中心区，取 30min 的时间片段和 $200\text{m} \times 200\text{m}$ 的空间片段，对于路网密度小于 $2\text{km}/\text{km}^2$ 或高峰小时流量小于 1000 辆/小时的城市郊区，取 30min 的时间片段和 $400\text{m} \times 400\text{m}$ 的空间片段。

所述步骤 3)具体包含以下步骤:

31) 将所需处理的空间区域划分为一定大小的格网, 每个格网区域的范围可表示为 $A_s = \{(x_s, y_s) | x_s \in [x_r, x_{r+1}), y_s \in [y_r, y_{r+1})\}$, 每个格网区域包含若干个路段, 把这些路段的集合表示为 R_s , 所述路段的集合 R_s 中的每条路段表示为 ij , 并为每个路段赋予编号;

32) 判定定位点所在的格网区域, 并利用距离和方位角, 在路段的集合 R_s 中搜索某定位点 A 所在的路段 ij , 匹配方案包括:

321) 单点匹配方案:

搜索距离点 A 最近的路段, 当满足点 A 的行驶方向角与路段 ij 的方向角的差值小于阈值时, 即满足 $|\theta_A - \theta_{ij}| < \delta_\theta$, 完成匹配, 所述的阈值 δ_θ 可取 2.5° , 5° , 10° 等; 若不满足 $|\theta_A - \theta_{ij}| < \delta_\theta$, 在搜索空间中删除路段 ij , 并继续搜索其他路段, 直至满足条件。匹配方法如图 3 所示。

322) 点序列匹配方案:

本方案适用于高频浮动车数据。将浮动车 GNSS 数据采集频率表示为 $f_0 = 1/t_0$, 将时间上与 A 相邻的点 $P(t_A - t_0)$, $P(t_A + t_0)$ 定义为 A 的 1-邻近点, $P(t_A - 2t_0)$, $P(t_A + 2t_0)$ 定义为 A 的 2-邻近点, 以此类推, 则 $P(t_A - kt_0)$, $P(t_A + kt_0)$ 定义为 A 的 k -邻近点。在 $f_0 < 1\text{Hz}$ 时, 取 $k=1$ 或 2。取距离 A 及 A 的 k -邻近点距离最小的路段 ij , 并计算 A 及 A 的 k -邻近点行驶方向角的均值 $\bar{\theta}_{A'}$, 若满足 $|\bar{\theta}_{A'} - \theta_{ij}| < \delta_\theta$, 完成匹配; 否则, 搜索其他路段, 直至满足 $|\bar{\theta}_{A'} - \theta_{ij}| < \delta_\theta$ 。

33) 利用路段的直线方程 (若为曲线路段则近似拆分为直线), 计算 GNSS 定位点在路段上的投影坐标, 减小因 GNSS 定位漂移带来的误差。具体方法采用 GNSS 定位点直线投影法为:

确定路段 ij 的直线方程 (若路段为曲线, 则划分为若干直线路段):

$$y - y_i = k(x - x_i)$$

其中斜率为:

$$k = \left(\frac{y_j - y_i}{x_j - x_i} \right)$$

投影直线方程为:

$$y - y_A = -\frac{1}{k_i}(x - x_A)$$

解出投影坐标 P 为:

$$x_P = \frac{ky_A - ky_i + k^2x_i + x_A}{k^2 + 1}$$

$$y_P = \frac{k^2y_A + y_i + kx_A - kx_i}{k^2 + 1}$$

在地图匹配过程后, 结合定位点坐标的时间戳数据, 将定位点匹配到时空子区。

所述步骤 5)具体可以采用以下方法之一:

51) 速度信息的全样本方案。由一个时空子区内各辆次浮动车的全部行程车速数据, 构成总体。实施方法是计算时空子区 ξ 内每辆车的行程车速: $v_\xi = \frac{d_{1,2} + d_{2,3} + \dots + d_{n-1,n}}{t_n - t_1}$, 其中 $d_{1,2} \dots d_{n-1,n}$ 为时空子区 ξ

内的第 1 个和第 2 个 GNSS 定位点间的距离, ..., 第 $n-1$ 个与第 n 个 GNSS 定位点间的距离, $t_1...t_n$ 为时空子区 ξ 内第 1 个, ..., 第 n 个 GNSS 定位点的时间戳; 将每个时空子区内的数据不做筛选, 构成一个集合 V_ξ , 用于后续处理。

52) 速度信息的时间平滑抽样方案。指定时间片段长度, 同一时间片段数据条数上限; 搜索一个时空子区内时间各时间片段内的速度数据, 若时间片段内速度数据条数超过上限, 随机取上限条数的数据加入待处理数据样本。实施方法是计算时空子区 ξ 内每辆车的行程车速: $v_\xi = \frac{d_{1,2} + d_{2,3} + \dots + d_{n-1,n}}{t_n - t_1}$,

其中 $d_{1,2}...d_{n-1,n}$ 为时空子区 ξ 内的第 1 个和第 2 个 GNSS 定位点间的距离, ..., 第 $n-1$ 个与第 n 个 GNSS 定位点间的距离, $t_1...t_n$ 为时空子区 ξ 内第 1 个, ..., 第 n 个 GNSS 定位点的时间戳; 指定时间片段长度 t_p , 同一时间片段数据条数上限 p_{max} ; 搜索一个时空子区内时间第 i 各时间片段内的速度数据, 若时间片段内速度数据条数超过上限 p_{max} , 随机取 p_{max} 条数据加入 V_ξ 并用于后续处理。

所述步骤 6) 具体包含以下步骤:

61) 确定 RNN 的网络结构, 这里使用 Elman-RNN 神经网络的基本结构, 网络结构如图 4、图 5 所示。网络具体包含以下组成部分:

611) 输入层: 根据神经网络的特征, 输入层为待训练的历史数据的各个实例, 由于这里的输入数据是一维的数据流, 即空间子区内的行程车速数据在时间维度上构成的时间序列数据, 因此, 输入层神经元个数和输出层神经元个数都设置为 1。

612) 隐含层: 在神经网络的设计中, 隐含层的个数设置尚未有定论, 一般都需要通过大量实验来最终确定网络模型的隐含神经元的个数, 建议取值为 5~8。

613) 输出层: 建立神经网络的目的是输出预测值, 即通过前两个时刻的值来预测下一个时刻的值。输出层输出值为预测值, 因此输出层个数设置为 1。

614) 上下文层 (Context 层): 在 Elman-Network 中, 上下文层用于保存前一个时刻隐含层的输出, 因此上下文层神经元个数与隐含层个数取相同值。

615) 偏置节点: 输入层和输出层各含有一个偏置节点, 初始值均设置为 0。

62) 设置 Elman-RNN 神经网络模型的基本参数:

Elman 网络结构中包含两个参数, 分别为隐含层神经元个数和上下文层的层数。隐含层神经元个数通过多次试验选择较好的结果, 建议的设置 5~8; 由于一个上下文层保存前一个时刻的隐含层输出值, 建议设置为 1。

使用反向传播算法进行模型的训练时需要设置参数学习率。学习率的大小决定了神经网络迭代过程中权值的变化程度, 大的学习率能使算法迅速收敛, 但是可能会陷入局部解, 小的学习率使得算法的收敛速度较慢, 但是能保证收敛到全局最小, 因此, 通常情况下, 学习率的取值应该选取较小值, 以确保模型的性能和稳定性, 建议设置为 0.01 到 0.8。

在算法的模拟退火模块中, 需要设置初始温度、终止温度和每个温度值的迭代次数。初始温度的取值对算法性能有着重要影响, 初始温度越大, 算法的迭代次数越多, 收敛到最小值的可能性也越大, 但是时间花销也越大。同样地, 初始温度较低时, 算法的性能将受到影响, 但是时间花销也较少。在实际应用中, 初始温度的设置需要通过多次反复试验进行选择。终止温度是指在温度下降过程中, 终止算法所设置的下限温度。每个温度值的迭代次数越多, 产生可能解的个数越多, 收敛到全局最小值得可能性也就越大。建议将初始温度设置为 10^5 , 将终止温度设置为 10^{-2} , 将每个温度的迭代次数设置

为 100。

63) 进行 RNN 模型的训练。RNN 模型的输入为行程车速构成的时间序列数据 $V=\{v_1, v_2, \dots, v_n\}$ ，输出为具有最优参数的 Elman-RNN 神经网络模型。模型训练的步骤如下：

631) 从数据文件或数据库中读取数据，并将数据组合成（输入，输出）对，即 $(v_1, v_2), (v_2, v_3), \dots, (v_{n-1}, v_n)$ 的形式，以便进行模型训练。

632) 创建 Elman-RNN 神经网络，其中，输入层神经元个数为 1 个，隐藏层神经元个数为 5~8 个，输出层神经元个数为 1 个，上下文层保存上一个时刻隐藏层的输出，神经元个数与隐藏层相同，由于训练数据中的输出都是正数，使用 Sigmoid 激活函数。

633) 初始化 RNN 模型，设置输入层到隐藏层，隐藏层到输出层，神经元之间连接的权重为随机值，分别设置输入层和隐藏层的偏置单元，并初始化为 0。

634) 设置主要训练算法为反向传播算法，备用算法为模拟退火算法，使用这两个算法的混合策略来训练模型，在训练过程中，如果某次训练后模型的错误率没有下降，或者下降的幅度小于设定的最小值，则使用模拟退火算法训练；使用贪心策略，如果某次训练后模型的错误率没有改善，则放弃更新权重，将模型的权重重置为上一次的训练结果。

635) 开始训练前，保存贪心策略需要的当前模型的网络权重和错误率。保存混合策略需要的当前模型的错误率。开始训练，对于上下文层和输入层数据，分别乘以对应的权重，加上偏置，通过激活函数计算出隐藏层的输出。隐藏层数据乘以对应的权重，加上偏置，通过激活函数，得到模型的输出。计算实际输出和模型输出的差值，使用反向传播算法将误差反向传递到隐藏层和输入层，计算梯度。根据算出的梯度计算权重增量，更新权重，保存更新后的权重。复制隐藏层输出到上下文层。

64) 单次训练后，使用以下策略对模型进行处理：

641) 贪心策略：如果本次训练错误率没有改善，恢复权重和错误率为上一次的值。

642) 混合策略：如果本次训练后模型的错误率没有下降，或者下降的幅度小于设定的最小值，则使用模拟退火算法训练。模拟退火算法的一次训练，大致步骤如下：首先计算当前模型的误差得分，之后对当前模型的所有权值和偏置值，添加一个随机数 add ，得到新的权重和偏置值，其中：

$$add = \frac{0.5 - Random}{startTemp} * temp$$

式中，Random 为随机数，startTemp 为初始温度，temp 为当前温度；计算更新后的模型误差得分，如果新的误差得分小于当前误差得分，说明新的权重对模型的性能有改进，则保存新的权值，否则丢弃；将当前温度乘以一个固定的比率 ratio 以降低温度：

$$ratio = \exp\left(\frac{\log(stopTemp / startTemp)}{cycles - 1}\right)$$

式中，stopTemp 为终止温度，cycles 为一次训练的迭代次数。重复以上过程 cycles 次。

65) 停止条件判断：如果对模型的改善程度小于最小值的次数大于设定的阈值，则算法终止。

重复 63)、64)、65) 步骤直到算法终止，得到训练好的 RNN 模型。训练过程参见图 6。

所述步骤 7) 具体可以采用以下方法之一：

71) 时间窗口均值法。使用数据抽样得到的行程车速数据，计算当前空间子区行程车速在当前时间

子区的均值 $\mu_{rt}(t_{\xi,rt}) = \frac{1}{N_{rt}} \sum_{j=1}^{N_{rt}} v_j$ ，作为实时交通态势的表征。

72) 滚动时域均值法。使用数据抽样得到的行程车速数据，计算当前空间子区行程车速在最近 M

个时间子区的均值 $\mu_{rt}(t_{\xi,rt}) = \frac{1}{M \cdot N_{rt}} \sum_{i=1}^M \sum_{j=1}^{N_{rt}} v_{ij}$ ，其中 M 一般取 3~5，作为实时交通态势的表征。

所述步骤 8) 具体包含以下步骤：

81) 将异常点检测所要处理的时间序列数据下标按升序排序的点的序列 $V: v_1, v_2, \dots, v_n$ 。为了对数据进行拟合，将数据以二维点的形式在平面坐标表示为 $(1, v_1), \dots, (n, v_n)$ 。

82) 为了检测异常点数据，需计算 RNN 模型预测的数据与真实数据的差值，具体实现过程如下：

821) 已知根据历史数据拟合的模型 M_{RNN} ，计算拟合模型的预测值 $v_{pr} = M_{RNN}(V)$ ；

822) 计算模型预测值与真实值的差异： $\text{diff}[v_{rt}, v_{pr}] = |v_{pr} - \mu_{rt}|$ 。

所述步骤 9) 具体包含以下步骤：

91) 将各个时空子区的速度分布差异标准化为 0~1 的规范化数值 a_{ξ} ：

$$a_{\xi_i} = \frac{\text{diff}_{\xi_i} - \min_{\xi}(\text{diff})}{\max_{\xi}(\text{diff}) - \min_{\xi}(\text{diff})}$$

92) 计算各个时空子区的交通异常指数 $A_{\xi} = a_{\xi} \times 10$ ；

93) 将异常指数高于 5 的区域位置投影到电子地图上，并智能移动设备 APP 等形式向社会公开发布，以使司机避开潜在拥堵点，提高城市道路交通的通行效率。

所述步骤 10) 具体包含以下步骤：

101) 计算交通异常状态的漏报率：

$$\alpha_1 = \frac{n_l}{n_a} \times 100\%$$

102) 计算交通异常状态的误报率：

$$\alpha_2 = \frac{n_e}{n_a} \times 100\%$$

以上两式中， n_l 为单位时间内漏报事件总数， n_e 为单位时间内误报事件总数， n_a 为单位时间内实际发生的异常时间总数。

本发明相较于同一领域的相似技术，具有以下优点：

(1) 充分利用现有的浮动车运营数据（GNSS 轨迹数据），通过历史交通特征提取和实时交通态势分析，检测交通状态发生的变化，可以实现城市道路交通异常事件实时性、低成本、智能化检测；

(2) 将交通特征的变化情况建立 RNN 模型，能够更加细致全面地反映交通运行规律，避免了利用单一指数表征交通特征的片面性、不稳定性，检测的可靠性更高；

(3) 能够自适应地调整模型参数，反映交通运行规律在大时间尺度下的细微变化；

(4) 经实际数据的检验，本发明提出的基于浮动车数据的城市道路交通异常检测技术，能够实现准确度较高的异常事件检测，检测率超过 90%，漏报率低于 15%，误报率低于 20%，取得了良好的检测效果，可以应用于城市交通智能化管理、服务。

附图说明

本发明的具体内容及优点结合以下附图将变得明晰和易于理解，其中：

图 1 示出了本发明的组成要素和基本原理示意图；

图 2 示出了本发明在实施过程中的总体流程示意图；

图 3 示出了本发明快速地图匹配算法实施方式示意图；

图 4 示出了本发明中 RNN 的结构简图；

图 5 示出了本发明中 RNN 的结构详图；

图 6 示出了进行 RNN 训练的流程示意图；

图 7 示出了本发明通过 RNN 模型与真实数据的对比进行异常检测的流程示意图。

具体实施方案

为了更加清晰明确地表述本发明的目的、技术方案和优势，下面对本发明的具体实施方案进行详细描述。

如附图 1 所示，本发明的整体系统构架包括：浮动车搭载的车载 GNSS 轨迹记录仪、数据中心、GNSS 卫星以及通信系统。此处的 GNSS 包括 GPS、GLONASS、GALILEO、北斗、IRNSS、QZSS 等任何类似的导航卫星定位系统。出租车、公交车等浮动车搭载的 GNSS 轨迹记录仪，以一定的采样频率 f （一般要求 $f > 0.1\text{Hz}$ ）记录车辆在行驶中各时点的位置信息，并通过 GPRS 移动通信网络（亦可采用 WCDMA、TD-LTE 等无线网络通信技术，但成本将相应提高）将位置信息实时发送至数据中心。数据中心通过数据预处理、数据融合，并通过特定算法建立历史道路交通特征数据库；对于最近接收的实时数据，建立实时交通特征数据库；通过历史数据库和实时数据库的映射关系，判别当前交通特征是否异常，并通过处理终端进行可视化展示并生成交通异常事件报告。

方案的总体流程参见图 2，包括采集和存储 GNSS 轨迹数据，建立时空子区，历史交通特征提取，实时交通特征提取，异常识别等步骤。采集和存储 GNSS 轨迹数据，是整个方案的数据基础，由于数据量级巨大，应采用分布式存储方案，对于分布式存储目前已有成熟的技术，不是本发明的内容。建立时空子区，其基本假设是在某一特定区域、特定时段内，有着相同的交通特征，这一假设，经过长期观测，是普遍适用的。历史交通特征提取，其原理是利用 GNSS 轨迹数据，计算得到行程车速，利用同一时空子区大量的历史行程车速数据，训练 Elman-RNN 模型，利用该神经网络模型表征交通特征的变化规律。实时交通特征提取，其原理是将当前时间段内的行程车速数据进行处理分析，获得实施交通特征统计量。异常识别是将历史行程车速训练的 Elman-RNN 模型得到的预测值域真实值进行比较，确定是否出现交通异常事件。

根据发明内容所述实施方法的组合，给出实施例如下。

实施例一

步骤 11、采用等距时空划分法，确定时间维度的片段尺度，时间片段跨度为固定值，通常取 30min

作为一个时间片段；确定空间维度的片段尺度，空间片段跨度为固定值，通常取 200m×200m 的空间网格作为一个空间片段。

步骤 12、进行数据预处理，将 GNSS 定位数据进行数据清洗、数据集成、数据转换、数据归约，提高数据的结构化程度。

步骤 13、将所需处理的空间区域划分为一定大小的格网，每个格网区域的范围可表示为 $A_s = \{(x_s, y_s) | x_s \in [x_r, x_{r+1}), y_s \in [y_r, y_{r+1})\}$ ；判定定位点所在的格网区域，并利用距离和方位角，搜索定位点所在的路段；搜索距离点 A 最近的路段，取阈值 $\delta_\theta = 2.5^\circ$ ，当满足点 A 的行驶方向角与路段 ij 的方向角的差值小于阈值 δ_θ 时，即满足 $|\theta_A - \theta_{ij}| < \delta_\theta$ ，完成匹配；若不满足 $|\theta_A - \theta_{ij}| < \delta_\theta$ ，在搜索空间中删除路段 ij ，并继续搜索其他路段，直至满足条件；利用路段的直线方程（若为曲线路段则近似拆分为直线），计算 GNSS 定位点在路段上的投影坐标，减小因 GNSS 定位漂移带来的误差，具体方法为：

确定路段 ij 的直线方程（若路段为曲线，则划分为若干直线路段）：

$$y - y_i = k(x - x_i)$$

其中斜率为：

$$k = \left(\frac{y_j - y_i}{x_j - x_i} \right)$$

投影直线方程为：

$$y - y_A = -\frac{1}{k_i}(x - x_A)$$

解出投影坐标 P 为：

$$x_P = \frac{ky_A - ky_i + k^2x_i + x_A}{k^2 + 1}$$

$$y_P = \frac{k^2y_A + y_i + kx_A - kx_i}{k^2 + 1}$$

在地图匹配过程后，结合定位点坐标的时间戳数据，将定位点匹配到时空子区。

步骤 14、由一个时空子区内各辆次浮动车的全部行车车速数据，构成总体。计算时空子区 ξ 内每辆车的行程车速： $v_\xi = \frac{d_{1,2} + d_{2,3} + \dots + d_{n-1,n}}{t_n - t_1}$ ，其中 $d_{1,2} \dots d_{n-1,n}$ 为时空子区 ξ 内的第 1 个和第 2 个 GNSS 定位点间的距离，……，第 $n-1$ 个与第 n 个 GNSS 定位点间的距离， $t_1 \dots t_n$ 为时空子区 ξ 内第 1 个，……，第 n 个 GNSS 定位点的时间戳；将每个时空子区内的数据不做筛选，构成一个集合 V_ξ ，用于后续处理。

步骤 15、建立由输入层、隐含层、输出层、上下文层、偏置节点组成的 Elman-RNN 神经网络，输入层神经元个数设置为 1，隐含层神经元个数设置为 5，输出层神经元个数设置为 1，上下文层神经元个数与隐含层个数取相同值，也设置为 5，偏置节点的初始值设置为 0；设置各个组成部分的初始参数，其中学习率设置为 0.1，模拟退火模块中初始温度设置为 10^5 ，终止温度设置为 10^2 ，每个温度的迭代次数设置为 100；将 V_ξ 作为输入数据，进行神经网络的训练，最终获得训练好的 RNN 模型。

步骤 16、将实时交通数据在当前时间窗口上取均值 $\mu_n(t_{\xi,n}) = \frac{1}{N_n} \sum_{j=1}^{N_n} v_j$ ，作为实施交通态势的表征。

步骤 17、将异常点检测所要处理的时间序列数据下标按升序排序的点的序列 $V: v_1, v_2, \dots, v_n$ ，已知根据历史数据拟合的模型 M_{RNN} ，计算拟合模型的预测值 $v_{pr} = M_{RNN}(V)$ ；计算模型预测值与真实值的差异

$$diff[v_n, v_{pr}] = |v_{pr} - \mu_n|。$$

步骤 18、将各个时空子区的速度分布差异标准化为 0~1 的规范化数值 $a_{\xi i}$:

$$a_{\xi i} = \frac{\text{diff}_{\xi i} - \min_{\xi}(\text{diff})}{\max_{\xi}(\text{diff}) - \min_{\xi}(\text{diff})}$$

计算各个时空子区的交通异常指数 $A_{\xi} = a_{\xi} \times 10$ 。

实施例二

步骤 21、采用等距时空划分法，确定时间维度的片段尺度，时间片段跨度为固定值，通常取 30min 作为一个时间片段；确定空间维度的片段尺度，空间片段跨度为固定值，通常取 200m×200m 的空间网格作为一个空间片段。

步骤 22、进行数据预处理，将 GNSS 定位数据进行数据清洗、数据集成、数据转换、数据归约，提高数据的结构化程度。

步骤 23、将所需处理的空间区域划分为一定大小的格网，每个格网区域的范围可表示为 $A_s = \{(x_s, y_s) | x_s \in [x_r, x_{r+1}), y_s \in [y_r, y_{r+1})\}$ ；判定定位点所在的格网区域，并利用距离和方位角，搜索定位点所在的路段；搜索距离点 A 最近的路段，取阈值 $\delta_\theta = 2.5^\circ$ ，当满足点 A 的行驶方向角与路段 ij 的方向角的差值小于阈值 δ_θ 时，即满足 $|\theta_A - \theta_{ij}| < \delta_\theta$ ，完成匹配；若不满足 $|\theta_A - \theta_{ij}| < \delta_\theta$ ，在搜索空间中删除路段 ij ，并继续搜索其他路段，直至满足条件；利用路段的直线方程（若为曲线路段则近似拆分为直线），计算 GNSS 定位点在路段上的投影坐标，减小因 GNSS 定位漂移带来的误差，具体方法为：

确定路段 ij 的直线方程（若路段为曲线，则划分为若干直线路段）：

$$y - y_i = k(x - x_i)$$

其中斜率为：

$$k = \left(\frac{y_j - y_i}{x_j - x_i} \right)$$

投影直线方程为：

$$y - y_A = -\frac{1}{k_i}(x - x_A)$$

解出投影坐标 P 为：

$$x_P = \frac{ky_A - ky_i + k^2x_i + x_A}{k^2 + 1}$$

$$y_P = \frac{k^2y_A + y_i + kx_A - kx_i}{k^2 + 1}$$

在地图匹配过程后，结合定位点坐标的时间戳数据，将定位点匹配到时空子区。

步骤 24、计算时空子区 ξ 内每辆车的行程车速： $v_{\xi} = \frac{d_{1,2} + d_{2,3} + \dots + d_{n-1,n}}{t_n - t_1}$ ，其中 $d_{1,2} \dots d_{n-1,n}$ 为时空子

区 ξ 内的第 1 个和第 2 个 GNSS 定位点间的距离，……，第 $n-1$ 个与第 n 个 GNSS 定位点间的距离， $t_1 \dots t_n$ 为时空子区 ξ 内第 1 个，……，第 n 个 GNSS 定位点的时间戳；指定时间片段长度 t_p ，同一时间片段数据条数上限 p_{max} ；搜索一个时空子区内时间第 i 各时间片段内的速度数据，若时间片段内速度数据条数超过上限 p_{max} ，随机取 p_{max} 条数据加入 V_{ξ} 。

步骤 25、建立由输入层、隐含层、输出层、上下文层、偏置节点组成的 Elman-RNN 神经网络，输入层神经元个数设置为 1，隐含层神经元个数设置为 5，输出层神经元个数设置为 1，上下文层神经元

个数与隐含层个数取相同值,也设置为 5,偏置节点的初始值设置为 0;设置各个组成部分的初始参数,其中学习率设置为 0.1,模拟退火模块中初始温度设置为 10^5 ,终止温度设置为 10^2 ,每个温度的迭代次数设置为 100;将 V_ξ 作为输入数据,进行神经网络的训练,最终获得训练好的 RNN 模型。

步骤 26、将实时交通数据在当前时间窗口上取均值 $\mu_{rt}(t_{\xi,rt}) = \frac{1}{N_{rt}} \sum_{j=1}^{N_{rt}} v_j$, 作为实施交通态势的表征。

步骤 27、将异常点检测所要处理的时间序列数据下标按升序排序的点的序列 $V: v_1, v_2, \dots, v_n$, 已知根据历史数据拟合的模型 M_{RNN} , 计算拟合模型的预测值 $v_{pr} = M_{RNN}(V)$; 计算模型预测值与真实值的差异 $diff[v_{rt}, v_{pr}] = |v_{pr} - \mu_{rt}|$ 。

步骤 28、将各个时空子区的速度分布差异标准化为 0~1 的规范化数值 a_{ξ_i} :

$$a_{\xi_i} = \frac{diff_{\xi_i} - \min_{\xi}(diff)}{\max_{\xi}(diff) - \min_{\xi}(diff)}$$

计算各个时空子区的交通异常指数 $A_{\xi} = a_{\xi} \times 10$ 。

实施例三

步骤 31、采用非等距时空划分法,对于路网密度大于 $2\text{km}/\text{km}^2$ 或高峰小时流量大于 1000 辆/小时的城市中心区,取 30min 的时间片段和 $200\text{m} \times 200\text{m}$ 的空间片段,对于路网密度小于 $2\text{km}/\text{km}^2$ 或高峰小时流量小于 1000 辆/小时的城市郊区,取 30min 的时间片段和 $400\text{m} \times 400\text{m}$ 的空间片段。

步骤 32、进行数据预处理,将 GNSS 定位数据进行数据清洗、数据集成、数据转换、数据归约,提高数据的结构化程度。

步骤 33、将所需处理的空间区域划分为一定大小的格网,每个格网区域的范围可表示为 $A_s = \{(x_s, y_s) | x_s \in [x_r, x_{r+1}), y_s \in [y_r, y_{r+1})\}$:

将浮动车 GNSS 数据采集频率表示为 $f_0 = 1/t_0$, 将时间上与 A 相邻的点 $P(t_A - t_0)$, $P(t_A + t_0)$ 定义为 A 的 1-邻近点, $P(t_A - 2t_0)$, $P(t_A + 2t_0)$ 定义为 A 的 2-邻近点,以此类推,则 $P(t_A - kt_0)$, $P(t_A + kt_0)$ 定义为 A 的 k-邻近点。在 $f_0 < 1\text{Hz}$ 时,取 $k=1$ 或 2。取距离 A 及 A 的 k-邻近点距离最小的路段 ij , 并计算 A 及 A 的 k-邻近点行驶方向角的均值 $\bar{\theta}_{A'}$, 取阈值 $\delta_{\theta} = 5^\circ$, 若满足 $|\bar{\theta}_{A'} - \theta_{ij}| < \delta_{\theta}$, 完成匹配; 否则, 搜索其他路段, 直至满足条件。

利用路段的直线方程(若为曲线路段则近似拆分为直线), 计算 GNSS 定位点在路段上的投影坐标, 减小因 GNSS 定位漂移带来的误差。具体方法为:

确定路段 ij 的直线方程(若路段为曲线, 则划分为若干直线段): $y - y_i = k(x - x_i)$

其中斜率为:

$$k = \left(\frac{y_j - y_i}{x_j - x_i} \right)$$

投影直线方程为:

$$y - y_A = -\frac{1}{k_i}(x - x_A)$$

解出投影坐标 P 为:

$$x_p = \frac{ky_A - ky_i + k^2x_i + x_A}{k^2 + 1}$$

$$y_p = \frac{k^2y_A + y_i + kx_A - kx_i}{k^2 + 1}$$

在地图匹配过程后, 结合定位点坐标的时间戳数据, 将定位点匹配到时空子区。

步骤 34、计算时空子区 ξ 内每辆车的行程车速: $v_{\xi} = \frac{d_{1,2} + d_{2,3} + \dots + d_{n-1,n}}{t_n - t_1}$, 其中 $d_{1,2} \dots d_{n-1,n}$ 为时空子

区 ξ 内的第 1 个和第 2 个 GNSS 定位点间的距离, ..., 第 $n-1$ 个与第 n 个 GNSS 定位点间的距离, $t_1 \dots t_n$ 为时空子区 ξ 内第 1 个, ..., 第 n 个 GNSS 定位点的时间戳; 指定时间片段长度 t_p , 同一时间片段数据条数上限 p_{max} ; 搜索一个时空子区内时间第 i 各时间片段内的速度数据, 若时间片段内速度数据条数超过上限 p_{max} , 随机取 p_{max} 条数据加入 V_{ξ} 。

步骤 35、建立由输入层、隐含层、输出层、上下文层、偏置节点组成的 Elman-RNN 神经网络, 输入层神经元个数设置为 1, 隐含层神经元个数设置为 5, 输出层神经元个数设置为 1, 上下文层神经元个数与隐含层个数取相同值, 也设置为 5, 偏置节点的初始值设置为 0; 设置各个组成部分的初始参数, 其中学习率设置为 0.05, 模拟退火模块中初始温度设置为 10^5 , 终止温度设置为 10^{-2} , 每个温度的迭代次数设置为 200; 将 V_{ξ} 作为输入数据, 进行神经网络的训练, 最终获得训练好的 RNN 模型。

步骤 36、采用滚动时域均值法, 使用数据抽样得到的行程车速数据, 计算当前空间子区行程车速

在最近 M 个时间子区的均值 $\mu_{rt}(t_{\xi,rt}) = \frac{1}{M \cdot N_{rt}} \sum_{i=1}^M \sum_{j=1}^{N_{rt}} v_{ij}$, 其中 $M=3$, 作为实时交通态势的表征。

步骤 37、将异常点检测所要处理的时间序列数据下标按升序排序的点的序列 $V: v_1, v_2, \dots, v_n$, 已知根据历史数据拟合的模型 M_{RNN} , 计算拟合模型的预测值 $v_{pr} = M_{RNN}(V)$; 计算模型预测值与真实值的差异

$$diff[v_{rt}, v_{pr}] = |v_{pr} - \mu_{rt}|。$$

步骤 38、将各个时空子区的速度分布差异标准化为 0~1 的规范化数值 a_{ξ_i} :

$$a_{\xi_i} = \frac{diff_{\xi_i} - \min_{\xi}(diff)}{\max_{\xi}(diff) - \min_{\xi}(diff)}$$

计算各个时空子区的交通异常指数 $A_{\xi} = a_{\xi} \times 10$ 。

权利要求书

1. 一种智能化道路交通异常检测方法，包括如下步骤：

- 1) 建立时空子区：将一天划分为若干时间片段，每个时间片段称为一个时间子区；将城市道路交通异常检测的实施区域划分为若干空间片段，每个空间片段称为一个空间子区；任意一个时间子区和任意一个空间子区的交集称为时空子区；实施方法是：确定时间维度的片段尺度，时间片段跨度为固定值，取 30min 作为一个时间片段；确定空间维度的片段尺度，空间片段跨度为固定值，取 200m×200m 的空间网格作为一个空间片段；
- 2) 历史轨迹数据的预处理：将浮动车 GNSS 定位历史数据处理为历史轨迹的抽样车速数据；实时轨迹数据的预处理：将浮动车 GNSS 定位实时数据处理为实时轨迹的抽样车速数据；
- 3) 历史轨迹数据分析和 RNN 模型训练：将所述历史轨迹的抽样车速数据进行组织规整，训练 RNN 模型，得到历史车速特征模型 M_{RNN} ；
实时轨迹数据分析和特征提取：利用所述实时轨迹的抽样车速数据，计算能够反映实时交通特征的参数；
- 4) 异常检测：通过差异性指标衡量所述历史车速特征模型与所述实时交通特征的差异，得到历史与实时交通特征差异值；
- 5) 异常严重性量化表征：利用所述历史与实时交通特征差异值计算交通状况异常指数；
- 6) 系统性能评价：评价交通异常状态检测的准确性，衡量系统运行的稳定程度。

2. 如权利要求 1 所述的智能化道路交通异常检测方法，其特征在于，步骤 2)所述的历史轨迹数据的预处理包括：

- 2a) 数据结构化：将浮动车 GNSS 定位历史数据进行数据清洗、数据集成、数据转换、数据归约，得到结构化 GNSS 定位历史数据；
- 2b) 快速地图匹配：结合城市路网数据，通过地图匹配算法，将结构化 GNSS 定位历史数据投影到城市路网，建立所述结构化 GNSS 定位历史数据中的定位点与路段的匹配关系，得到所述结构化 GNSS 定位历史数据中的定位点与所述路段的匹配关系表，并修正定位漂移带来的误差；
- 2c) 交通运行特征参数计算和抽样：根据所述结构化 GNSS 定位历史数据计算交通运行特征参数，得到历史轨迹的交通特征数据，并对所述历史轨迹的交通特征数据进行数据抽样，得到抽样历史轨迹的抽样交通特征数据。

3. 如权利要求 2 所述的智能化道路交通异常检测方法，其特征在于，步骤 2b)所述的快速地图匹配包括：

- 2b1) 将所需处理的空间区域划分为一定大小的格网，每个格网区域的范围可表示为 $A_s = \{(x_s, y_s) | x_s \in [x_r, x_{r+1}), y_s \in [y_r, y_{r+1})\}$ ，每个格网区域包含若干个路段，把这些路段的集合表示为 R_s ，所述路段的集合 R_s 中的每条路段表示为 ij ，并为每个路段赋予编号；
- 2b2) 判定定位点所在的格网区域，并利用距离和方位角，在路段的集合 R_s 中搜索某定位点 A 所在的路段 ij ；
- 2b3) 利用 GNSS 定位点直线投影法，计算 GNSS 定位点在路段上的投影坐标。

4. 如权利要求3所述的智能化道路交通异常检测方法,其特征在于,步骤2b2)采用下述方法之一:

2b21) 单点匹配方法: 搜索距离某定位点A最近的路段,实施方法是: 对于路段的集合 R_s 中的某一路段 ij , 当满足点A的行驶方向角 θ_A 与路段 ij 的方向角 θ_{ij} 的差值小于阈值 δ_θ 时, 即满足

$$|\theta_A - \theta_{ij}| < \delta_\theta \text{ 时, 完成匹配; 若不满足 } |\theta_A - \theta_{ij}| < \delta_\theta, \text{ 继续搜索路段集合 } R_s \text{ 中的其他路段, 直至}$$

$$\text{满足 } |\theta_A - \theta_{ij}| < \delta_\theta;$$

2b22) 点序列匹配方法: 本方案适用于高频浮动车数据; 将每两条相邻时间的浮动车GNSS数据时间间隔表示为 t_0 , 将浮动车GNSS数据采集频率表示为 $f_0=1/t_0$, 将某定位点A的时间记录表示为 t_A , 将时间上与所述定位点A相邻的点 $P(t_A-t_0)$, $P(t_A+t_0)$ 定义为A的1-邻近点, $P(t_A-2t_0)$, $P(t_A+2t_0)$ 定义为某定位点A的2-邻近点, 以此类推, 则 $P(t_A-kt_0)$, $P(t_A+kt_0)$ 定义为某定位点A的 k -邻近点; 在 $f_0 < 1\text{Hz}$ 时, 取 $k=1$ 或2; 取距离某定位点A及该定位点的 k -邻近点距离最小的一个路段 ij , 并计算该定位点及其 k -邻近点行驶方向角的均值 $\bar{\theta}_{A'}$, 若满足 $|\bar{\theta}_{A'} - \theta_{ij}| < \delta_\theta$,

完成匹配; 否则, 搜索其他路段, 直至满足 $|\bar{\theta}_{A'} - \theta_{ij}| < \delta_\theta$ 。

5. 如权利要求4所述的智能化道路交通异常检测方法, 其特征在于, 步骤2c)所述的历史轨迹数据的预处理采用下述方法之一:

2c1) 全样本方法: 由一个时空子区内 ξ 各辆次浮动车的全部行程车速数据, 构成总体, 实施方法是计算时空子区 ξ 内每辆车的行程车速: $v_\xi = \frac{d_{1,2} + d_{2,3} + \dots + d_{n-1,n}}{t_n - t_1}$, 其中 $d_{1,2} \dots d_{n-1,n}$ 为时空子区

ξ 内的第1个和第2个GNSS定位点间的距离, ..., 第 $n-1$ 个与第 n 个GNSS定位点间的距离, $t_1 \dots t_n$ 为时空子区 ξ 内第1个, ..., 第 n 个GNSS定位点的时间戳; 将每个时空子区内的行程车速数据不做筛选, 构成一个集合 V_ξ , 用于后续处理;

2c2) 时间平滑的抽样方法: 指定时间片段长度, 设置同一时间片段数据条数上限; 搜索一个时空子区各时间片段内的速度数据, 若时间片段内速度数据条数超过上限, 随机取上限条数的数据用于后续处理, 实施方法是计算时空子区 ξ 内每辆车的行程车速: $v_\xi = \frac{d_{1,2} + d_{2,3} + \dots + d_{n-1,n}}{t_n - t_1}$,

其中 $d_{1,2} \dots d_{n-1,n}$ 为时空子区 ξ 内的第1个和第2个GNSS定位点间的距离, ..., 第 $n-1$ 个与第 n 个GNSS定位点间的距离, $t_1 \dots t_n$ 为时空子区 ξ 内第1个, ..., 第 n 个GNSS定位点的时间戳; 指定时间片段长度 t_p , 设置同一时间片段数据条数上限 p_{max} ; 搜索一个时空子区内时间第 i 各时间片段内的速度数据, 若时间片段内速度数据条数超过上限 p_{max} , 随机取 p_{max} 条数据加入 V_ξ 并用于后续处理。

6. 如权利要求1至5之一所述的智能化道路交通异常检测方法, 其特征在于, 步骤3)所述的历史轨迹数据分析和RNN模型训练采取以下步骤:

3a) 建立Elman-RNN神经网络的基本结构, 包括设置输入层的神经元个数, 设置隐含层的神经元个数, 设置输出层的神经元个数, 设置上下文层的神经元个数, 设置输入层和输出层的偏置

节点并赋初始值；

- 3b) 设置 Elman-RNN 神经网络模型的基本参数, 包括反向传播算法进行模型训练时的参数学习率, 模拟退火的初始温度和终止温度及每个温度的迭代次数;
 - 3c) 利用所述历史轨迹的抽样车速数据, 进行 RNN 模型的训练;
 - 3d) 单次训练后对模型进行处理, 包括更新神经网络的权重, 模型的错误率评估, 及错误率未下降时的更新策略;
 - 3e) 进行模型训练的停止条件判断, 如果对模型的改善程度小于最小值的次数大于设定的阈值, 则算法终止, 得到 RNN 模型 M_{RNN} ;
- 步骤 3)所述的实时轨迹数据分析和特征提取采用下述方法之一:
- 3f) 时间窗口均值法: 利用所述实时轨迹的抽样车速数据, 计算当前空间子区行程车速在当前时间

子区的均值 $\mu_{rt}(v_{\xi,rt}) = \frac{1}{N_{rt}} \sum_{j=1}^{N_{rt}} v_j$, 作为实时交通特征的表示参数;

- 3g) 滚动时域均值法: 利用所述实时轨迹的抽样车速数据, 计算当前空间子区行程车速在最近 M

个时间子区的均值 $\mu_{rt}(v_{\xi,rt}) = \frac{1}{M \cdot N_{rt}} \sum_{i=1}^M \sum_{j=1}^{N_{rt}} v_{ij}$, 其中 M 取 3~5, 作为实时交通特征的表示参数。

- 7. 如权利要求 6 所述的智能化道路交通异常检测方法, 其特征在于, 步骤 3a)包含以下步骤:

- 3a1) 将输入层的神经元个数设置为 1;
- 3a2) 将隐含层的神经元个数设置为 5~8;
- 3a3) 将输出层的神经元个数设置为 1;
- 3a4) 将上下文层的神经元个数设置为与隐含层神经元个数相同;
- 3a5) 为输入层与输出层各设置一个偏置节点, 初始值均设置为 0;

步骤 3b)包含以下步骤:

- 3b1) 将反向传播算法进行模型训练时的参数学习率设置为 0.01~0.8;
- 3b2) 将模拟退火的初始温度设置为 10^5 , 模拟退火的终止温度设置为 10^{-2} ;
- 3b3) 将每个温度的迭代次数设置为 100。

- 8. 如权利要求 6 所述的智能化道路交通异常检测方法, 其特征在于, 步骤 3c)包含以下步骤:

- 3c1) 将每个空间子区的历史轨迹的抽样车速数据按时间排序, 并将排序后的历史轨迹的抽样车速数据两两组合成(输入, 输出)对, 即 $(v_1, v_2), (v_2, v_3), \dots, (v_{n-1}, v_n)$ 的形式;
- 3c2) 创建 Elman-RNN 神经网络, 其中, 输入层神经元个数为 1 个, 隐藏层神经元个数为 5、6、7 或 8 个, 输出层神经元个数为 1 个, 上下文层保存上一个时刻隐藏层的输出, 神经元个数与隐藏层相同; 使用 Sigmoid 激活函数;
- 3c3) 设置输入层到隐藏层、隐藏层到输出层神经元之间连接的权重为 0~1 之间的随机值; 分别设置输入层和隐藏层的偏置单元并初始化为 0;
- 3c4) 采用反向传播算法和模拟退火两种算法, 组成混合策略来训练模型。

9. 如权利要求 6 所述的智能化道路交通异常检测方法, 其特征在于, 步骤 3d)采用以下方法之一:

3d1) 贪心策略: 如果某一次训练后模型的错误率没有下降, 恢复权重和错误率为训练前的值;

3d2) 混合策略: 如果某一次训练后模型的错误率没有下降, 或者下降的幅度小于设定的最小值, 则使用模拟退火算法训练; 模拟退火算法的一次训练步骤如下: 首先计算当前模型的误差得分, 之后对当前模型的所有神经元之间连接的权重和偏置单元的值, 添加一个随机数 add , 得到新的权重和偏置单元的值, 其中 $add = (0.5 - Random) / startTemp * temp$, 式中, $Random$ 为随机数, 范围为大于 0 小于 1, $startTemp$ 为初始温度, 取 10^5 , $temp$ 为当前温度; 计算更新后的模型误差得分, 如果新的误差得分小于当前误差得分, 说明新的权重对模型的性能有改进, 则保存新的权重, 否则丢弃; 将当前温度乘以一个固定的比率 $ratio$ 以降低温度: $ratio = \exp(\log(stopTemp / startTemp) / (cycles - 1))$, 式中, $stopTemp$ 为终止温度, 取 10^{-2} , $cycles$ 为一次训练的迭代次数, 取 100; 重复以上过程 $cycles$ 次。

10. 如权利要求 1 至 5 之一所述的智能化道路交通异常检测方法, 其特征在于, 步骤 4)异常检测包括以下步骤:

4a) 将异常检测所要处理的历史轨迹的抽样车速数据的下标按升序排序的点的序列 $V: v_1, v_2, \dots, v_n$;

将数据以二维点的形式在平面坐标表示为 $(1, v_1), \dots, (n, v_n)$;

4b) 计算 RNN 模型的预测值 $v_{pr} = M_{RNN}(V)$, 并计算模型预测值与真实值的差异

$$diff[v_n, v_{pr}] = |v_{pr} - \mu_n|;$$

步骤 5)异常严重性量化表征包括:

5a) 将各个时空子区的速度分布差异标准化为 0~1 的规范化数值 a_{ξ} :

$$a_{\xi} = \left[diff_{\xi} - \min_{\xi}(diff) \right] / \left[\max_{\xi}(diff) - \min_{\xi}(diff) \right]$$

5b) 计算各个时空子区的交通异常指数 $A_{\xi} = a_{\xi} \times 10$ 。

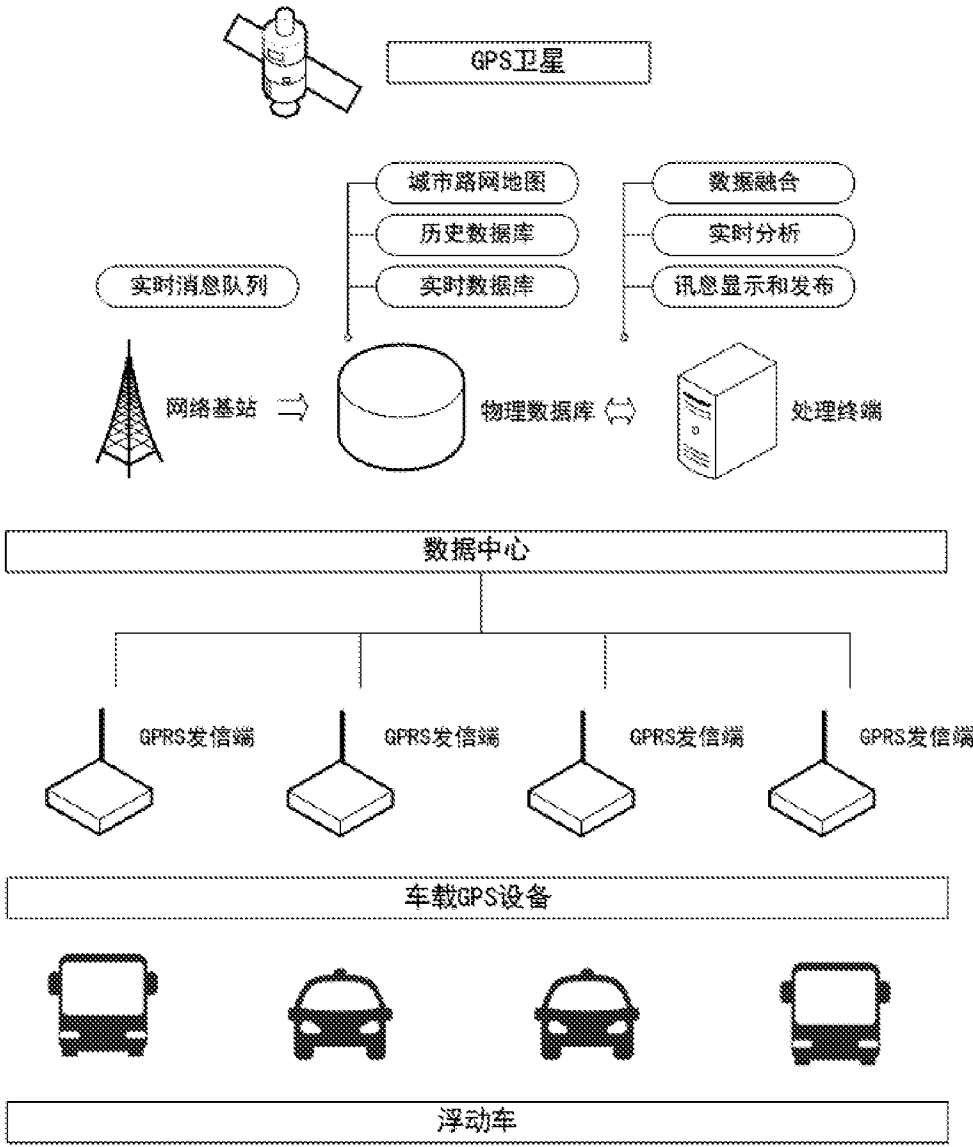


图 1

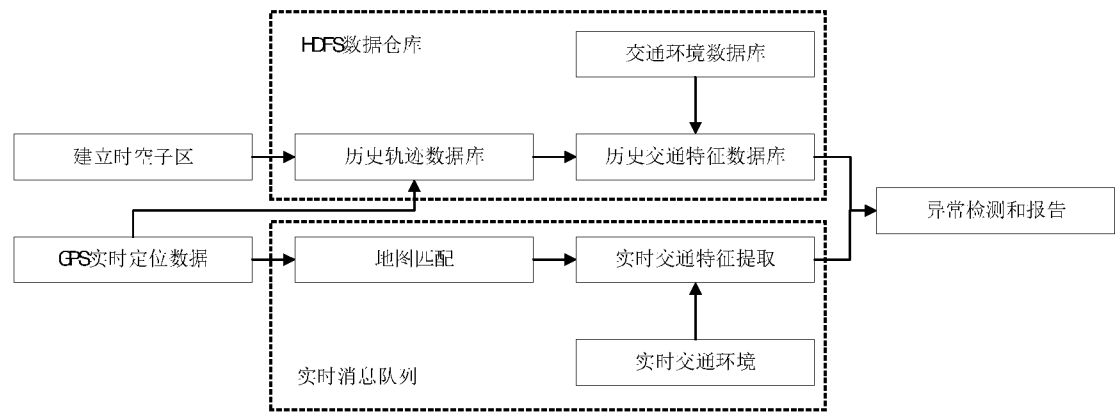


图 2

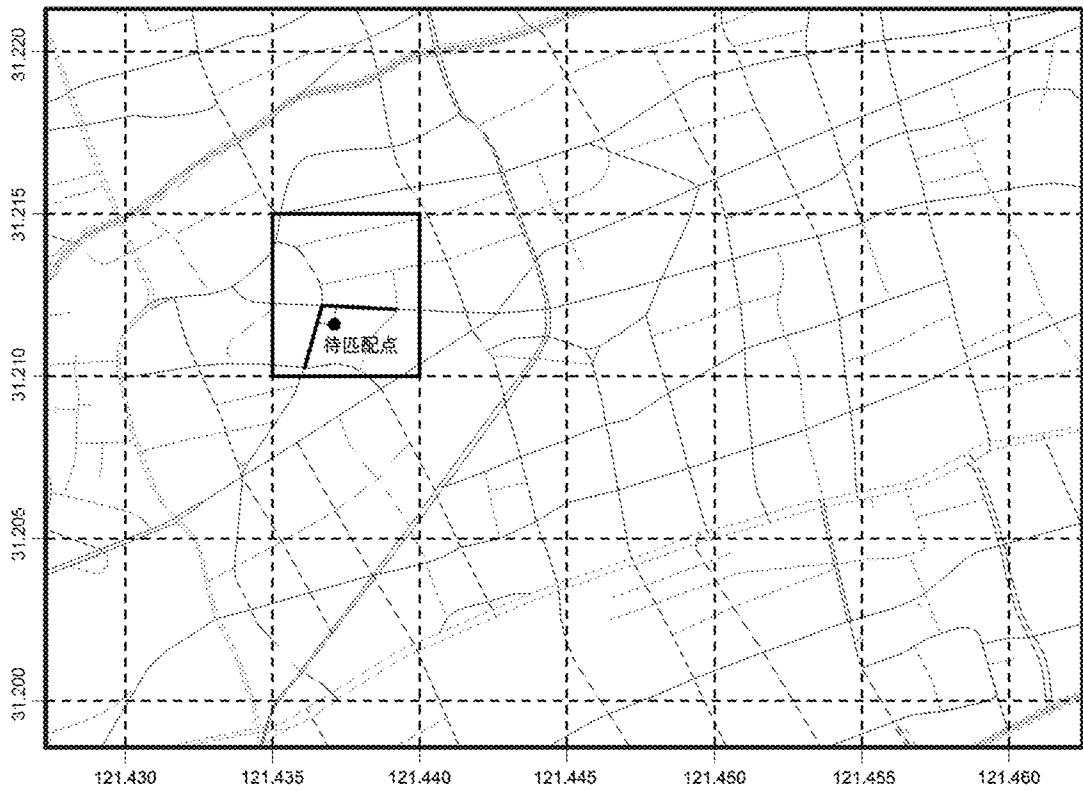


图 3

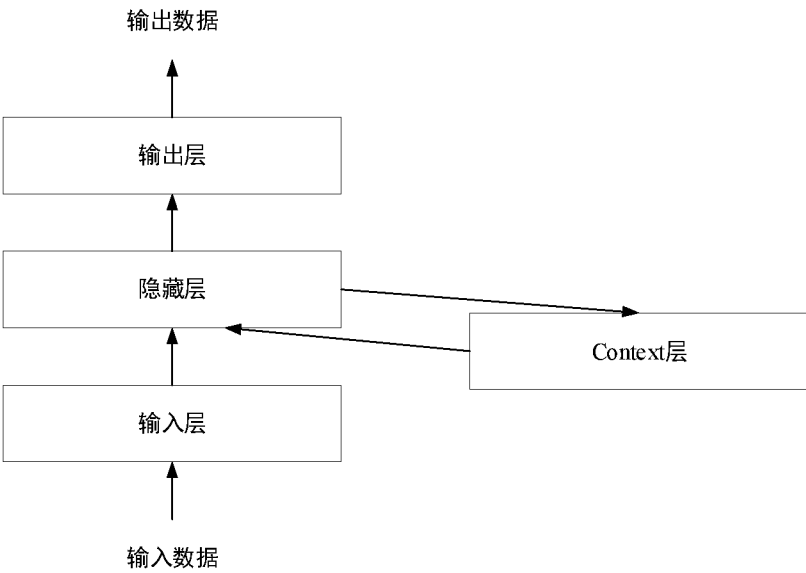


图 4

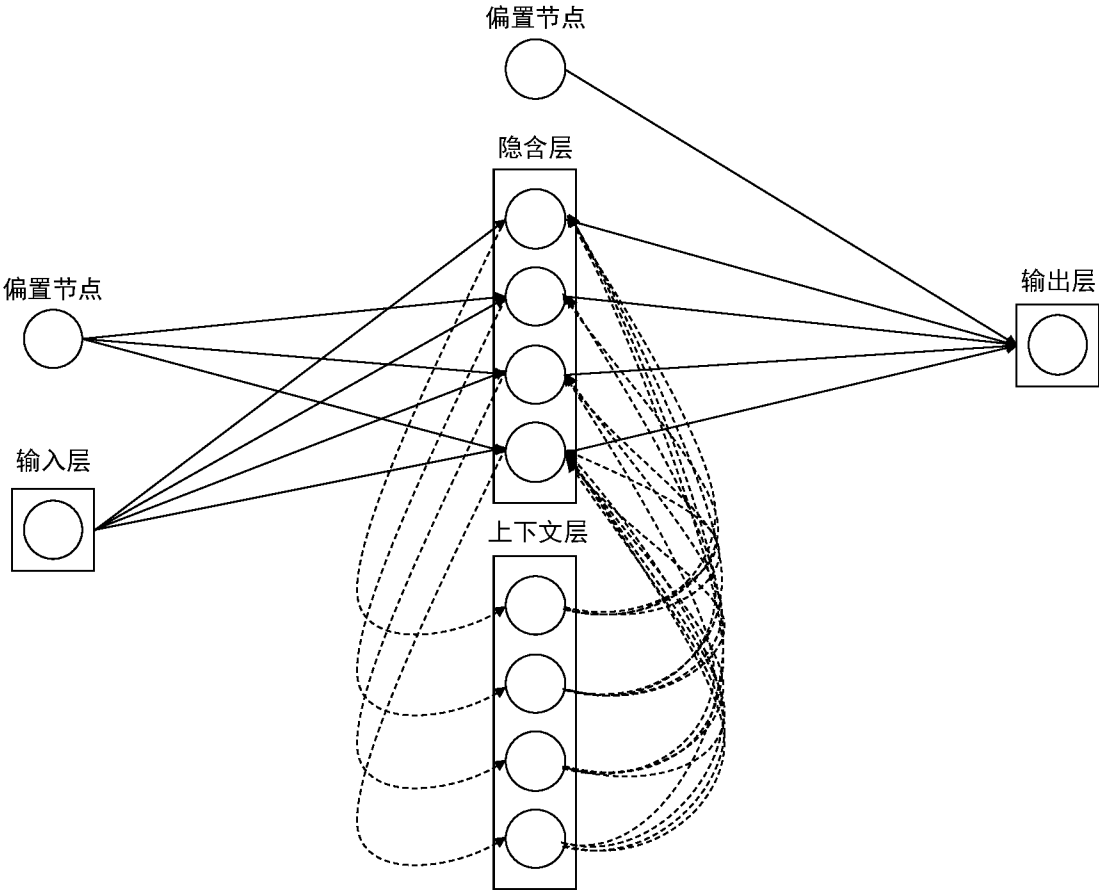


图 5

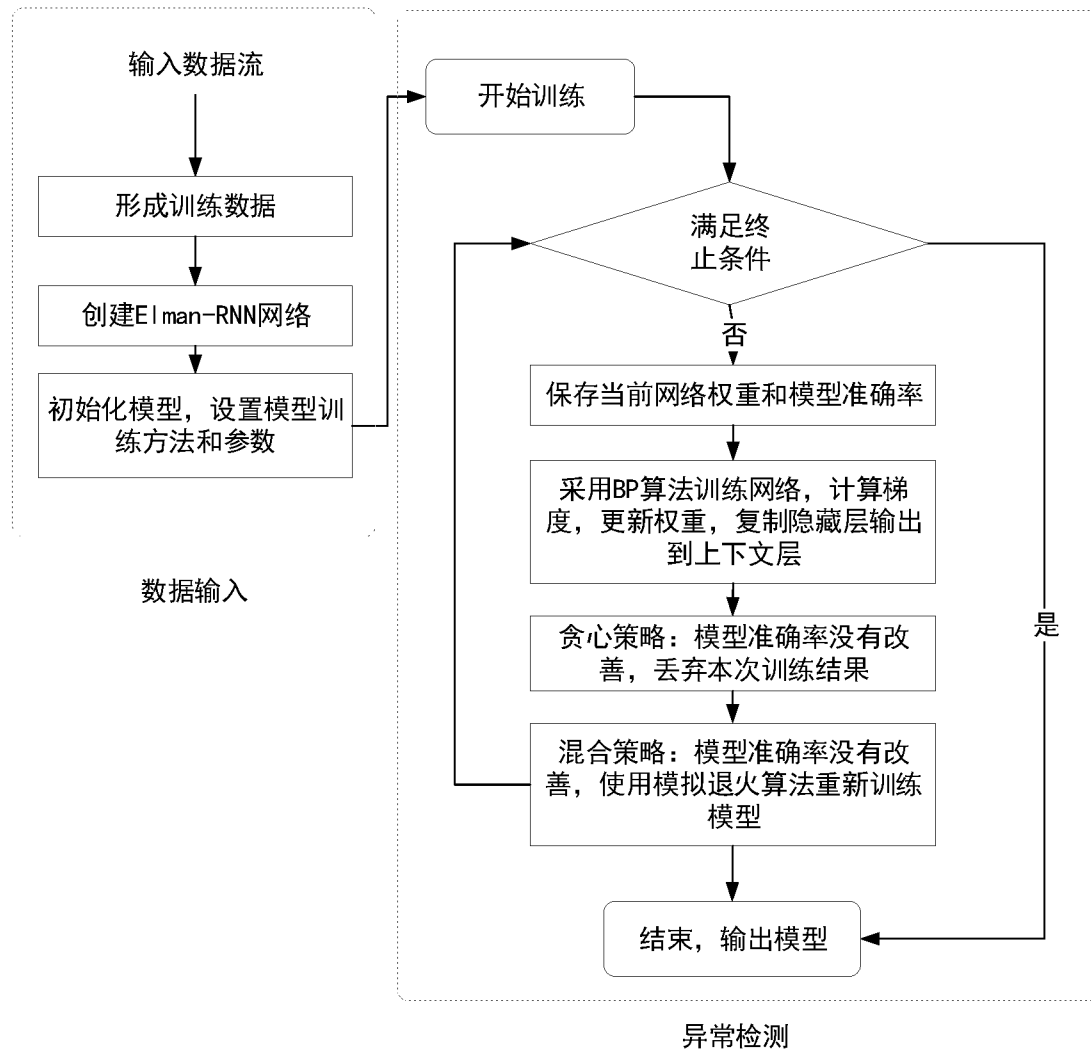


图 6

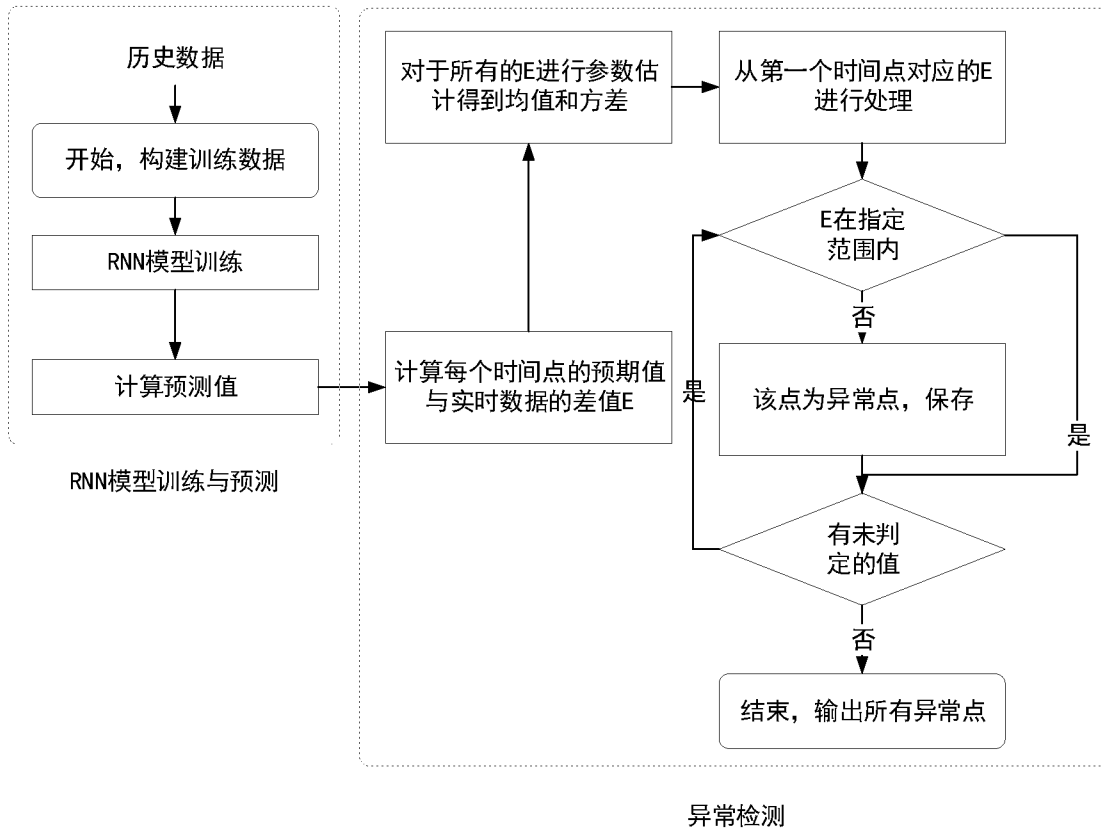


图 7

INTERNATIONAL SEARCH REPORT

International application No.
PCT/IB2017/058533

A. CLASSIFICATION OF SUBJECT MATTER

G06F 17/30 (2006.01) i; G08G 1/01 (2006.01) i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06F; G08G

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNABS, CNTXT, VEN, SIPOABS: 智能化, 道路, 交通, 异常, 检测, 时空子区, 历史轨迹数据, intelligentize, road, roadway, traffic, unconventionality, detect+, space time, history, data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 104408924 A (SHENZHEN BEIHANG NEW INDUSTRIAL TECHNOLOGY RESEARCH INSTITUTE), 11 March 2015 (11.03.2015), claims 1-2	1-10
A	CN 201927174 U (XIAN INSTITUTE OF POSTS AND TELECOMMUNICATIONS), 10 August 2011 (10.08.2011), entire document	1-10
A	CN 201946111 U (XIAN INSTITUTE OF POSTS AND TELECOMMUNICATIONS), 24 August 2011 (24.08.2011), entire document	1-10
A	CN 104504901 A (ZHEJIANG ENJOYOR RESEARCH INSTITUTE CO., LTD.), 08 April 2015 (08.04.2015), entire document	1-10
A	JP 2003228788 A (FUJI XEROX CO., LTD.), 15 August 2003 (15.08.2003), entire document	1-10

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:	“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
“A” document defining the general state of the art which is not considered to be of particular relevance	“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
“E” earlier application or patent but published on or after the international filing date	“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	“&” document member of the same patent family
“O” document referring to an oral disclosure, use, exhibition or other means	
“P” document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search 11 April 2018	Date of mailing of the international search report 28 April 2018
Name and mailing address of the ISA State Intellectual Property Office of the P. R. China No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088, China Facsimile No. (86-10) 62019451	Authorized officer ZHAO, Guiqin Telephone No. (86-10) 010-62085093

International application No.
PCT/IB2017/058533

Form PCT/ISA/210 (patent family annex) (July 2009)

国际检索报告

国际申请号

PCT/IB2017/058533

A. 主题的分类 G06F 17/30(2006.01)i; G08G 1/01(2006.01)i 按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类		
B. 检索领域 检索的最低限度文献(标明分类系统和分类号) G06F; G08G 包含在检索领域中的除最低限度文献以外的检索文献 在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用)) CNABS, CNTXT, VEN, SIPOABS, 智能化, 道路, 交通, 异常, 检测, 时空子区, 历史轨迹数据, intelligentize, road, roadway, traffic, unconventionality, detect+, space time, history, data		
C. 相关文件		
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
A	CN 104408924 A (深圳北航新兴产业技术研究院) 2015年 3月 11日 (2015 - 03 - 11) 权利要求1-2	1-10
A	CN 201927174 U (西安邮电学院) 2011年 8月 10日 (2011 - 08 - 10) 全文	1-10
A	CN 201946111 U (西安邮电学院) 2011年 8月 24日 (2011 - 08 - 24) 全文	1-10
A	CN 104504901 A (浙江银江研究院有限公司) 2015年 4月 8日 (2015 - 04 - 08) 全文	1-10
A	JP 2003228788 A (FUJI XEROX CO LTD) 2003年 8月 15日 (2003 - 08 - 15) 全文	1-10
☐ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。		
* 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件 “T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件 “X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 “Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 “&” 同族专利的文件		
国际检索实际完成的日期		国际检索报告邮寄日期
2018年 4月 11日		2018年 4月 28日
ISA/CN的名称和邮寄地址		受权官员
中华人民共和国国家知识产权局(ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088 传真号 (86-10)62019451		赵桂芹 电话号码 (86-10)010-62085093

国际检索报告
关于同族专利的信息

国际申请号

PCT/IB2017/058533

检索报告引用的专利文件			公布日 (年/月/日)	同族专利	公布日 (年/月/日)
CN	104408924	A	2015年 3月 11日	CN 104408924 B	2016年 6月 1日
CN	201927174	U	2011年 8月 10日	无	
CN	201946111	U	2011年 8月 24日	无	
CN	104504901	A	2015年 4月 8日	CN 104504901 B	2016年 6月 8日
JP	2003228788	A	2003年 8月 15日	JP 4200747 B2	2008年 12月 24日