(19)国家知识产权局

(12)发明专利

(10)授权公告号 CN 112016735 B
(45)授权公告日 2023.03.28

(21)申请号 202010693026.4

(22)申请日 2020.07.17

(65)同一申请的已公布的文献号
申请公布号 CN 112016735 A

(43)申请公布日 2020.12.01

(73)专利权人 厦门大学
地址 361000 福建省厦门市思明南路422号

(72)发明人 陈龙彪 蒋之晗 范晓亮 王程 洪虹

(74)专利代理机构 厦门市首创君合专利事务所有限公司 35204
专利代理师 张松亭

(51)Int.Cl.
G06Q 10/047(2023.01)
G06Q 50/26(2012.01)
G08G 1/01(2006.01)
G06F 18/2431(2023.01)
G06F 18/214(2023.01)

(56)对比文件
CN 111161443 A,2020.05.15
CN 108596409 A,2018.09.28
CN 109697852 A,2019.04.30
CN 109191828 A,2019.01.11
CN 105489015 A,2016.04.13
CN 109993498 A,2019.07.09
CN 110197589 A,2019.09.03
CN 107644532 A,2018.01.30
CN 106056162 A,2016.10.26
CN 109101568 A,2018.12.28
CN 107248293 A,2017.10.13
CN 106781524 A,2017.05.31
US 2020051196 A1,2020.02.13
US 2018144634 A1,2018.05.24
US 2018174446 A1,2018.06.21
US 2016232785 A1,2016.08.11
JP 6055916 B2,2016.12.27
Guoyu Ou etc."Real-Time Vehicular Traffic Violation Detection in Traffic Monitoring Stream".《Proceedings of the The 2012 IEEE/WIC/ACM International Joint Conferences on Web Intelligence and Intelligent Agent Technology》.2012,15-19页.

(续)

审查员 陈华惠

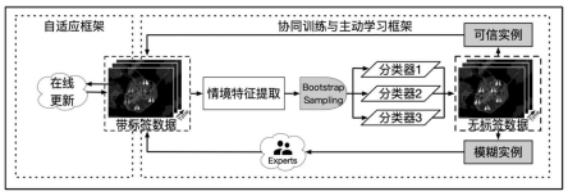
权利要求书3页 说明书8页 附图2页

(54)发明名称

一种基于交通违法热点预测的巡逻路线规划方法、系统和可读存储介质

(57)摘要

本发明公开一种基于交通违法热点预测的巡逻路线规划方法、系统和可读存储介质，方法具体包括如下步骤：将离散的交通违法事件匹配到道路网络，得到交通违法热点候选区域；提取交通违法热点候选区域的相关时空特征组成数据集；采用拔靴抽样法从数据集中抽取三组训练数据集，分别用三组训练数据集通过协同训练与主动学习训练三组预测模型，所述预测模型采用树集成模型；利用所述三组预测模型，对无标签交通数据进行分类，预测未来交通违法热点的分布情况；根据交通违法热点分布情况，进行交警巡逻路线的建模和优化。本发明提供的方法能够低成本、高覆盖、实时的感知预测城市中的交通违法热点并规划交警巡逻路线，帮助城市交通管理部门全面深入地掌握城市中的交通违法热点，并且高效地规划交警巡逻路线。



CN 112016735 B

[转续页]

[接上页]

(56)对比文件

Yunxuan Li etc.“Analyzing traffic violation behavior at urban intersections: A spatiotemporal”.《Elsevier》.2020,1-11.

1.一种基于交通违法热点预测的巡逻路线规划方法，其特征在于，具体包括如下步骤：

S1:将离散的交通违法事件匹配到道路网络，得到交通违法热点候选区域；

S2:得出所述交通违法热点候选区域的交通违法时间序列，计算出交通违法时间序列下的交通违法热点候选区域的违法程度，并提取交通违法时间序列下的交通违法热点候选区域的时空情境特征，将交通违法热点候选区域以及其交通违法时间序列、违法程度、时空情境特征组成数据集；所述违法程度：将每个交通违法热点候选区域在每个小时的违法程度划分为三个等级，分别为轻度，中度，重度；

所述交通违法时间序列为$v_1^i \to v_2^i \to v_3^i \to \cdots \to v_T^i$,其中，$v_t^i, t \in [1,T]$表示$r_i$在第t个小时的交通违法事件数量；$r_i$表示第i个交通违法热点候选区域，$i \in [1,N]$，根据$r_i$在t小时的违法事件数量确定所述交通违法时间序列下$r_i$违法程度$Level_t^i$；

所述时空情境特征包括：基本时间特征、气候特征以及空间特征；

其中基本时间特征包括但不限于该小时是一天中的第几个小时，该小时属于一周中的第几天，该小时属于一年中的第几天，该小时属于一月中的第几天，该小时属于一年中的第几个月，该小时属于一年中第几个周，该小时属于一年中的第几个季度；以及气候特征，包括但不限于该小时该交通违法热点候选区域的温度，该小时该交通违法热点候选区域的湿度和该小时该交通违法热点候选区域的气压；空间特征包括其基本位置特征，基本位置特征包括但不限于该交通违法热点候选区域的经度、纬度，该交通违法热点候选区域的道路等级，该交通违法热点候选区域的道路限速；S3:采用拔靴抽样法从数据集中抽取三组训练数据集，分别用三组训练数据集通过协同训练与主动学习训练三组预测模型，所述预测模型采用树集成模型；

S4:利用所述三组预测模型，对无标签交通数据进行分类，预测未来交通违法热点的分布情况；

S5:根据预测的未来交通违法热点的分布情况，进行交警巡逻路线的建模和优化；

$\mathrm{argmax}_X L \cdot X$　　①

$subject\ to$

$$\sum_{j=1}^{N} x(0,j,1) = \sum_{i=0}^{N}\sum_{j=0}^{N} x(i,j,1) = \sum_{i=0}^{N} x(i,0,T) = \sum_{j=0}^{N}\sum_{t=1}^{T} x(0,j,t) = k \quad ②$$

$$\sum_{i=0}^{N}\sum_{t=1}^{T-1} x(i,0,t) = 0 \quad ③$$

$$\sum_{i=1}^{N} x(i,p,t) = \sum_{j=0}^{N} x(p,j,t+1) \le 1, \forall\, t \in [1,T-1], p \in [1,N] \quad ④$$

$D \cdot X \leqslant \mu$　　⑤

$x(i,j,t) \in \{0,1\}, \forall i \in [0,N], j \in [0,N], t \in [1,T]$　　⑥

其中，①式表示优化目标，X为邻接张量，L为增益张量，邻接张量X使得$L \cdot X$即巡逻路线覆盖的交通违法热点增益最大；限制条件②中，$\sum_{j=1}^{N} x(0,j,1)$表示了从起始点出发的路径数量，$\sum_{i=0}^{N} x(i,0,T)$表示最终回到起始点的路径数量，$\sum_{j=1}^{N} x(0,j,1) = \sum_{i=0}^{N}\sum_{j=0}^{N} x(i,j,1)$表示所有路线都从起始点开始，$\sum_{i=0}^{N} x(i,0,T) = \sum_{j=0}^{N}\sum_{t=1}^{T} x(0,j,t)$表示所有路线都到起始结束，且路径数量为k，t为第t个小时；限制条件③表示没有路径会在工作时间结束前回到起始点；②③两个限制条件共同保证了有向图的基本结构；限制条件④表示到达某个结点的路径数量等于离开该结点的数量，并且路径之间不会存在点交叉；限制条件⑤⑥表示总体的路径长度不超过阈

值μ。

2.根据权利要求1所述的一种基于交通违法热点预测的巡逻路线规划方法，其特征在于，利用所述三组预测模型，对无标签交通数据进行分类，具体包括：

样本被三个预测模型同时分到同一类，则标记为可信实例标签，并加入三个训练集中；

样本被两个预测模型分到同一类，在第三个预测模型中分到不同的类别，则标记为较可信实例标签，并加入第三个预测模型相应的训练集中；

样本被三个预测模型分到了都不相同的类别，则标记为模糊实例标签，加入专家验证集。

3.根据权利要求1所述的一种基于交通违法热点预测的巡逻路线规划方法，其特征在于，所述预测模型定期采用在线更新机制，进行重新训练。

4.根据权利要求1所述的一种基于交通违法热点预测的巡逻路线规划方法，其特征在于，所述步骤S5中交警巡逻路线为一张有向图，图中的每个结点v(i,t)表示在第t个小时的交通违法热点候选区域r_i，属性值Level(i,t)表示r_i在第t个小时的交通违法程度，每条边x(i,j,t)表示交警是否在第t小时开始时从结点v(i,t-1)移动到结点v(j,t)，如果x(i,j,t)＝1则表示存在这样的移动，该式为0则表示不存在这样的移动。

5.根据权利要求4所述的一种基于交通违法热点预测的巡逻路线规划方法，其特征在于，所述步骤S5中还包括：根据交警巡逻路线构建邻接张量X，损失张量D和增益张量L，邻接张量X表示巡逻路径，损失张量D表示某时刻从某点到某点的移动距离成本，增益张量L表示某时刻从某点到某点的交通违法热点覆盖增益。

6.一种基于交通违法热点预测的巡逻路线规划系统，其特征在于，包括：

事件匹配单元：配置于将离散的交通违法事件匹配到道路网络，得到交通违法热点候选区域；

数据集获取单元：配置于得出所述交通违法热点候选区域的交通违法时间序列，计算出交通违法时间序列下的交通违法热点候选区域的违法程度，并提取交通违法时间序列下的交通违法热点候选区域的时空情境特征，将交通违法热点候选区域以及其交通违法时间序列、违法程度、时空情境特征组成数据集；

所述交通违法时间序列为$v_1^i \to v_2^i \to v_3^i \to \cdots \to v_T^i$，其中，$v_t^i, t \in [1,T]$表示ri在第t个小时的交通违法事件数量；$r_i$(i∈[1,N])表示第i个交通违法热点候选区域，根据r_i在t小时的违法事件数量确定所述交通违法时间序列下r_i违法程度$Level_t^i$；

所述时空情境特征包括：基本时间特征、气候特征以及空间特征；

其中基本时间特征包括但不限于该小时是一天中的第几个小时，该小时属于一周中的第几天，该小时属于一年中的第几天，该小时属于一月中的第几天，该小时属于一年中的第几个月，该小时属于一年中第几个周，该小时属于一年中的第几个季度；以及气候特征，包括但不限于该小时该交通违法热点候选区域的温度，该小时该交通违法热点候选区域的湿度和该小时该交通违法热点候选区域的气压；空间特征包括其基本位置特征，基本位置特征包括但不限于该交通违法热点候选区域的经度、纬度，该交通违法热点候选区域的道路等级，该交通违法热点候选区域的道路限速；

预测模型训练单元：配置于采用拔靴抽样法从数据集中抽取三组训练数据集，分别用三组训练数据集通过协同训练与主动学习训练三组预测模型，所述预测模型采用树集成模

型；

模型预测单元：配置于利用所述三组预测模型，对无标签交通数据进行分类，预测未来交通违法热点的分布情况；

巡逻路线获取单元：配置于根据预测的未来交通违法热点的分布情况，进行交警巡逻路线的建模和优化；

$\mathrm{argmax}_X L \cdot X$ ①

$subject to$

$$\sum_{j=1}^{N} x(0,j,1) = \sum_{i=0}^{N}\sum_{j=0}^{N} x(i,j,1) = \sum_{i=0}^{N} x(i,0,T) = \sum_{j=0}^{N}\sum_{t=1}^{T} x(0,j,t) = k \quad ②$$

$$\sum_{i=0}^{N}\sum_{t=1}^{T-1} x(i,0,t) = 0 \quad ③$$

$$\sum_{i=1}^{N} x(i,p,t) = \sum_{j=0}^{N} x(p,j,t+1) \le 1, \forall\, t \in [1,T-1], p \in [1,N] \quad ④$$

$D \cdot X \leqslant \mu$ ⑤

$$x(i,j,t) \in \{0,1\}, \forall i \in [0,N], j \in [0,N], t \in [1,T] \quad ⑥$$

其中，①式表示优化目标，X为邻接张量，L为增益张量，邻接张量X使得L.X即巡逻路线覆盖的交通违法热点增益最大；限制条件②中，$\sum_{j=1}^{N} x(0,j,1)$表示了从起始点出发的路径数量，$\sum_{i=0}^{N} x(i,0,T)$表示最终回到起始点的路径数量，$\sum_{j=1}^{N} x(0,j,1) = \sum_{i=0}^{N}\sum_{j=0}^{N} x(i,j,1)$表示所有路线都从起始点开始，$\sum_{i=0}^{N} x(i,0,T) = \sum_{j=0}^{N}\sum_{t=1}^{T} x(0,j,t)$表示所有路线都到起始结束，且路径数量为k，t为第t个小时；限制条件③表示没有路径会在工作时间结束前回到起始点；②③两个限制条件共同保证了有向图的基本结构；限制条件④表示到达某个结点的路径数量等于离开该结点的数量，并且路径之间不会存在点交叉；限制条件⑤⑥表示总体的路径长度不超过阈值μ。

7.一种存储有计算机指令的非瞬时计算机可读存储介质，其特征在于，所述计算机指令用于执行权利要求1-5中任一项所述的方法。

一种基于交通违法热点预测的巡逻路线规划方法、系统和可读存储介质

技术领域

[0001] 本发明涉及交通预测领域，特别是指一种基于交通违法热点预测的巡逻路线规划方法、系统和可读存储介质。

背景技术

[0002] 随着经济的发展，城市机动车数量急剧上升，车祸、堵车等各种道路交通问题对人民的生命财产安全造成了巨大的危害。根据世界卫生组织的数据显示，全世界每年有超过一百万人死于交通事故，让绝大多数的国家损失了3％的GDP。根据中国交通运输部发布的数据显示，交通拥堵带来的经济损失占到了城市人口可支配收入的20％，每年达2500亿人民币。而在造成道路交通问题的种种因素中，司机的违法驾驶行为占比在80％以上。因此，减少交通违法行为的发生就成了城市交通管理的关键。通常，交通管理部门会采用在一些交通违法高发地安装设备监测，如监控摄像头、雷达等，并且在一些交通高峰时段安排警力去往热点路口路段值勤。然而，人力物力资源往往是有限而昂贵的，传统的分配方法确实低效而高成本的，这造成了大量的资源浪费。例如，由于监控摄像头的设置不合理导致许多摄像头使用率低，巡逻路线缺乏合理规划给交警带来了巨大的工作量，而工作效率却反而较低。因此，为了能够对这些资源进行合理高效地分配，我们需要对城市中高度动态分布的交通违法热点有全面、及时的感知，并能够据此规划出高效、低成本的交警巡逻路线。

[0003] 在目前已有的相关研究中，对城市中交通违法热点的感知与预测往往是基于对历史交通违法记录的建模和根据交警的经验进行推断，这些方法高度依赖于历史经验的积累。一方面，这导致了我们对一些没有历史数据，比如没有安装摄像头等监控设备或者很少有交警巡逻的地点的缺乏认知。另一方面，随着城市的不断扩张发展，这些方法也无法及时地对新的路口路段进行感知预测。因此，我们需要一个低成本、高覆盖、实时的方法来对城市中的交通违法热点进行感知和预测。而对于巡逻路线规划，大多数研究都致力于缩减对单独的突发事件比如潜在的交通事故的响应，增大对事件覆盖的期望。然而，这些工作只关注对单独事件的响应，却不适用于针对时间长短不一的交通违法热点的路线规划，并且也未能考虑到不同交通违法热点由于其违法程度的不同导致其在路线规划中优先权的不同。另一类研究致力于对有限的警力资源进行分配从而最大化对地点所需警力资源的覆盖，然而，这些工作却只关注从警局到热点地区的分配，而交警的巡逻往往是沿着一个有一系列地点的路线。还有一种研究是采用启发式算法来找出一系列路线使其能够覆盖尽可能多的热点并使其路径总距离尽可能短。然而这种方法往往需要通过调节参数在路径距离和路线覆盖率中进行取舍，无法对实际情况进行自动适应。

发明内容

[0004] 本发明的主要目的在于克服现有技术中的上述缺陷，提出一套基于交通违法热点预测的巡逻路线规划方法和系统，通过对城市中的交通违法热点进行建模预测来规划交警

巡逻路线。该方法和系统能够低成本、高覆盖、实时的感知预测城市中的交通违法热点并规划交警巡逻路线，帮助城市交通管理部门全面深入地掌握城市中的交通违法热点，并且高效地规划交警巡逻路线。

[0005] 本发明采用如下技术方案：

[0006] 一种基于交通违法热点预测的巡逻路线规划方法，其特征在于，具体包括如下步骤：

[0007] S1：将离散的交通违法事件匹配到道路网络，得到交通违法热点候选区域；

[0008] S2：得出所述交通违法热点候选区域的交通违法时间序列，计算出时间序列下的交通违法热点候选区域的违法程度，并提取时间序列下的交通违法热点候选区域的时空情境特征，将交通违法热点候选区域以及其交通违法时间序列、违法程度、时空情景特征组成数据集；

[0009] S3：采用拔靴抽样法从数据集中抽取三组训练数据集，分别用三组训练数据集通过协同训练与主动学习训练三组预测模型，所述预测模型采用树集成模型；

[0010] S4：利用所述三组预测模型，对无标签交通数据进行分类，预测未来交通违法热点的分布情况；

[0011] S5：根据交通违法热点分布情况，进行交警巡逻路线的建模和优化。

[0012] 优选的，所述交通违法时间序列为$v_1^i \to v_2^i \to v_3^i \to \cdots \to v_T^i$，其中，$v_t^i, t \in [1,T]$表示$r_i$在第t个小时的交通违法事件数量；$r_i$(i∈[1,N])表示第i个交通违法热点候选区域，根据r_i在t小时的违法事件数量确定所述时间序列下r_i违法程度$Level_t^i$。

[0013] 优选的，所述时空情景特征包括：时间特征及空间特征。

[0014] 优选的，利用所述三组预测模型，对无标签交通数据进行分类，具体包括：

[0015] 样本被三个预测模型同时分到同一类，则标记为可信实例标签，并加入三个训练集中；

[0016] 样本被两个预测模型分到同一类，在第三个预测模型中分到不同的类别，则标记为较可信实例标签，并加入第三个预测模型相应的训练集中；

[0017] 样本被三个预测模型分到了都不相同的类别，则标记为模糊实例标签，加入专家验证集。

[0018] 优选的，所述预测模型定期采用在线更新机制，进行重新训练。

[0019] 优选的，所述步骤S5中交警巡逻路线为一张有向图，图中的每个结点v(i,t)表示在第t 个小时的交通违法热点候选区域r_i，属性值Level(i,t)表示r_i在第t个小时的交通违法程度，每条边x(i,j,t)表示交警是否在第t小时开始时从结点v(i,t-1)移动到结点v(j,t)，如果x(i,j,t)＝1 则表示存在这样的移动，该式为0则表示不存在这样的移动。

[0020] 优选的，所述步骤S5中还包括：根据交警巡逻路线构建邻接张量X，损失张量D和增益张量L，临界张量X表示巡逻路径，损失张量D表示某时刻从某点到某点的移动距离成本，增益张量L表示某时刻从某点到某点的交通违法热点覆盖增益。

[0021] 优选的，所述步骤S5中交警巡逻路线建模和优化还包括：

[0022] $\mathrm{argmax}_X L \cdot X$ ①

[0023] subjectto

[0024] $$\sum_{j=1}^{N} x(0,j,1) = \sum_{i=0}^{N}\sum_{j=0}^{N} x(i,j,1) = \sum_{i=0}^{N} x(i,0,T) = \sum_{j=0}^{N}\sum_{t=1}^{T} x(0,j \quad ②$$

[0025] $$\sum_{i=0}^{N}\sum_{t=1}^{T-1} x(i,0,t) = 0 \quad ③$$

[0026] $$\sum_{i=1}^{N} x(i,p,t) = \sum_{j=0}^{N} x(p,j,t+1) \le 1, \forall\, t \in [1,T-1], p \in [1,N] \quad ④$$

[0027] $D \cdot X \leqslant \mu$ ⑤

[0028] $x(i,j,t) \in \{0,1\}, \forall i \in [0,N], j \in [0,N], t \in [1,T]$ ⑥

[0029] 其中，①式表示我们的优化目标，邻接张量X使得L·X即巡逻路线覆盖的交通违法热点增益最大；限制条件②中，$\sum_{j=1}^{N} x(0,j,1)$表示了从起始点出发的路径数量，$\sum_{i=0}^{N} x(i,0,T)$表示最终回到起始点的路径数量，$\sum_{j=1}^{N} x(0,j,1) = \sum_{i=0}^{N}\sum_{j=0}^{N} x(i,j,1)$表示所有路线都从起始点开始，$\sum_{i=0}^{N} x(i,0,T) = \sum_{j=0}^{N}\sum_{t=1}^{T} x(0,j,t)$表示所有路线都到起始结束，且路径数量为k；限制条件③表示没有路径会在工作时间结束前回到起始点；②③两个限制条件共同保证了有向图的基本结构；限制条件④表示到达某个结点的路径数量等于离开该结点的数量，并且路径之间不会存在点交叉；限制条件⑤⑥表示总体的路径长度不超过阈值μ。

[0030] 本发明另一方面提供一种基于交通违法热点预测的巡逻路线规划系统，包括：

[0031] 事件匹配单元：配置于将离散的交通违法事件匹配到道路网络，得到交通违法热点候选区域；

[0032] 数据集获取单元：配置于得出所述交通违法热点候选区域的交通违法时间序列，计算出时间序列下的交通违法热点候选区域的违法程度，并提取时间序列下的交通违法热点候选区域的时空情景特征，将交通违法热点候选区域以及其交通违法时间序列、违法程度、时空情景特征组成数据集；

[0033] 预测模型训练单元：配置于采用拔靴抽样法从数据集中抽取三组训练数据集，分别用三组训练数据集通过协同训练与主动学习训练三组预测模型，所述预测模型采用树集成模型；

[0034] 模型预测单元：配置于利用所述三组预测模型，对无标签交通数据进行分类，预测未来交通违法热点的分布情况；

[0035] 巡逻路线获取单元：配置于根据交通违法热点分布情况，进行交警巡逻路线的建模和优化。

[0036] 本发明的再一方面，提供一种存储有计算机指令的非瞬时计算机可读存储介质，所述计算机指令用于执行上述的方法。

[0037] 由上述对本发明的描述可知，与现有技术相比，本发明具有如下有益效果：

[0038] 1、本发明通过对城市中的交通违法热点进行建模预测来规划交警巡逻路线，能够低成本、高覆盖、实时的感知预测城市中的交通违法热点并规划交警巡逻路线，帮助城市交通管理部门全面深入地掌握城市中的交通违法热点，并且高效地规划交警巡逻路线。

[0039] 2、本发明首先将离散的交通违法事件匹配到道路网络，得到交通违法热点候选区域，依据交通违法热点候选区域进行后续建模预测和巡逻路线规划，解决了目前仅仅依据单独的交通违法事件进行预测，以及依据单独的交通违法事件规划巡逻路线的低效以及不准确的问题。

[0040] 3、本发明提取了时间序列下的交通违法热点候选区域的时空情景特征，利用时空

情境特征对城市中的交通违法热点候选区域进行建模并且预测交通违法热点，能够适用于历史记录稀少甚至没有历史记录的地点，并且通过能够通过对各个时空情境特征对交通违法热点预测影响程度的大小分析某地在某时刻频繁发生交通违法事件的原因。

[0041] 4、本发明采用结合协同训练与主动学习构建预测模型，能够使用少量带标签数据准确预测交通违法热点，解决了获取大量带标签数据耗时耗力的问题。

[0042] 5、本发明的模型中采用在线更新机制，即每隔一段时间丢弃一些旧数据，加入一些新数据，对模型进行重新训练，使其能够保持对当前城市道路环境的适应性。

[0043] 6、本发明采用有向图以及约束限制的方式规划巡逻路线，将巡逻路线规划问题建模成为简单的整数线性规划问题，避免资源浪费，使得巡逻路线能够尽可能多的覆盖交通违法热点，且这些违法热点由于其违法程度的不同在路线规划中有不同的优先权，并且能够保证交警巡逻的路线长度需要限制在一个合理的范围内，得到最优的巡逻路径。

[0044] 以下结合附图及实施例对本发明作进一步详细说明；但本发明的一种基于交通违法热点预测的巡逻路线规划方法、系统和可读存储介质不局限于实施例。

附图说明

[0045] 图1是交通违法热点预测模型示意图；

[0046] 图2是本发明交警巡逻路线建模示意图；

[0047] 图3是本发明实施例中某一交警巡逻路线建模示意图；

[0048] 图4是本发明实施例中某一交警巡逻路线建模的邻接张量、损失张量和增益张量示意图；其中(a)邻接张量(b)损失张量(c)增益张量。

具体实施方式

[0049] 实施例，一种基于交通违法热点预测的巡逻路线规划方法，具体包括：

[0050] S1：将离散的交通违法事件匹配到道路网络，得到交通违法热点候选区域；

[0051] 首先，通过公开道路地图OSM(OpenStreetMap)获取道路网络数据，将其分为路口和路段，再将交通违法记录数据匹配离其最近的路口或者路段，我们将这些路口路段称为交通违法热点候选区域。

[0052] S2：得出所述交通违法热点候选区域的交通违法时间序列，计算出时间序列下的交通违法热点候选区域的违法程度，并提取时间序列下的交通违法热点候选区域的时空情景特征，将交通违法热点候选区域以及其交通违法时间序列、违法程度、时空情景特征组成数据集；

[0053] 然后，对于每个交通违法热点候选区域，我们将其存在的交通违法记录按每小时划分并求和，得到每个候区域的违法记录时间特征序列。设我们总共有N个区域的T个小时的数据，则对于某个区域r_i(i∈[1,N])，将其交通违法时间序列记为$v_i^1 \to v_2^i \to v_3^i \to \cdots \to v_T^i$，其中，$v_i^t(t \in [1,T])$表示$r_i$在第t个小时的交通违法事件数量。我们将每个区域在每个小时的违法程度划分为三个等级，分别为轻度(记为0)，中度(记为1)，重度(记为2)。在第t个小时，对于所有违法热点候选区域在该小时的违法事件数量求平均值和标准差，将违法时间数量超过平均值加两倍的标准差的候选区域在该小时标记为重度违法热点，将违法事件数量超过平均值但小于等于平均值加两倍标准差的候选区域该小时标记为重度违法，将违法事件数

量小于等于平均值的候选区域该小时标记为轻度违法。用公式表示如下，

[0054] $$\mathrm{mean}_t = \frac{\sum_{i=1}^{N} v_t^i}{N}, std_t = \sqrt{\frac{\sum_{i=1}^{N}(v_t^i - mean_t)^2}{N}}$$

[0055] $$Level_t^i = \begin{cases} 0, & v_t^i \le mean_t \\ 1, & mean_t < v_t^i \le mean_t + 2 \times std_t \\ 2, & v_t^i > mean_t + 2 \times std_t \end{cases}$$

[0056] 其中，mean_t表示在第t个小时的所有违法热点候选区域的违法事件数量的平均值，std_t表示在第t个小时的所有违法热点区域的违法事件数量的标准差，$Level_t^i$表示在第t个小时的违法热点候选区域r_i的违法程度。

[0057] 接下来，对于每个热点候选区域的每个小时，我们提取其时空情境特征。时间情境特征提取包括其基本时间特征，包括但不限于该小时是一天中的第几个小时，该小时属于一周中的第几天，该小时属于一年中的第几天，该小时属于一月中的第几天，该小时属于一年中的第几个月，该小时属于一年中第几个周，该小时属于一年中的第几个季度；以及气候特征，包括但不限于该小时该区域的温度，该小时该区域的湿度和该小时该区域的气压；空间特征包括其基本位置特征，包括但不限于该区域的经度、纬度，该区域的道路等级，该区域的道路限速)，及兴趣点特征(该区域附近有多少娱乐设施，有多少住宿设施，有多少旅游设施，有多少商务设施，有多少基础设施)。

[0058] 城市中的交通违法事件的分布具有高度的动态性，离散地分布在城市道路网络中。如果仅仅对单独的交通违法事件进行预测，并根据单独的交通违法事件规划巡逻路线是不现实且十分低效的。因此，我们需要提取交通违法事件的时空特征并将其聚集成交通违法热点候选区域。然而，传统的提取热点区域的方法，如将城市划分为细粒度网格，对每个网格中的交通违法事件进行统计等，是低效的。因为城市中的很多区域都不是道路网路。除此之外，简单地将城市划分网格难以确定合适的网格大小，且容易导致许多路口路段被划分到不同的网格中。因此，本技术方案创新性地基于城市道路网络将城市中离散的交通违法事件匹配到道路网络中，从而得到城市中的交通违法热点候选区域。

[0059] S3：采用拔靴抽样法从数据集中抽取三组训练数据集，分别用三组训练数据集通过协同训练与主动学习训练三组预测模型，所述预测模型采用树集成模型；

[0060] 如图1，为交通违法热点预测模型，根据交通违法热点候选区域以及提取时空情境特征，首先利用拔靴抽样法从中抽取三份训练数据集，分别用这三个训练集训练三个分类器，分类器的算法采用树集成模型(XGBoost)。

[0061] S4：利用所述三组预测模型，对无标签交通数据进行分类，预测未来交通违法热点的分布情况；

[0062] 利用这三个预测模型(分类器)分别对剩下的无标签数据进行分类。如果一个实例被三个分类器同时分到同一类，则将其标记为可信实例加上该标签加入三个训练集中；如果一个样本被两个分类器分到同一类，但是在第三个分类器中分到了不同类别，则将其加上两个分类器分到的类别加入第三个分类器的训练集中；如果一个样本被三个分类器都分到了不同的类别，则将其标记为模糊实例，加入一个专家验证集，等待专家进行判断，根据专家给出的标签加入三个分类器的训练集中，专家的标记能力有限，如果超过合理范围则不会再将模糊数据加入专家验证集。在每一轮协同训练中，对训练集产生变化的分类器进

行重新训练，直到所有分类器都不再变化。最后利用这三个分类器对无标签数据进行投票分类，得到未来的交通违法热点分布情况。

[0063] 另一方面，根据交警执法过程中收集到的新数据，我们每隔一段时间丢弃一部分旧数据，加入新数据对模型进行重新训练。比如每过半年，我们收集到了x条数据，则丢弃原有数据集中最旧的x条数据，加入新的x条数据，使其能够保持对当前城市道路环境的适应性。

[0064] 传统的交通违法热点预测方法往往都是基于对历史的交通违法记录的建模和根据交警的经验来推断，这些方法高度依赖于历史经验的积累。一方面，这导致了我们对一些没有历史数据，比如说没有安装摄像头等监控设备或者很少有警力过去巡逻的地点的缺乏认知。另一方面，随着城市的不断扩张发展，这些方法也无法及时地对新的路口路段进行感知预测。因此，本方案创新性地利用交通违法热点候选区域的时空情境特征对其进行建模，并据此预测交通违法热点。使之能够适用于历史记录稀少甚至没有历史记录的地点，并且通过能够通过对各个时空情境特征对交通违法热点预测影响程度的大小分析某地在某时刻频繁发生交通违法事件的原因。

[0065] 此外，通常在训练预测模型时，往往需要提供大量带标签的数据；然而，由于交通违法热点的判断高度依赖于人工和经验积累，获取大量带标签数据十分耗时耗力，本发明创新性地结合了协同训练与主动学习构建预测模型，使其能够使用少量的带标签数据即刻得到高准确度的预测模型。

[0066] 而且，在该模型中加入了在线更新机制，即每隔一段时间丢弃一些旧数据，加入一些新数据，对模型进行重新训练，使其能够保持对当前城市道路环境的适应性。

[0067] S5:根据交通违法热点分布情况，进行交警巡逻路线的建模和优化。

[0068] 如图2为交警巡逻路线建模示意图；首先将交警巡逻路线建模成一张有向图，其中，起点和终点两个Origin表示交警站点，即交警从这里该点出发，在工作时间结束后又回到该点。交警的调度以小时为单位，图中的每个结点v(i,t)表示在第t个小时的交通违法热点候选区域r_i，其具有一个属性Level(i,t)表示该热点候选区域r_i在第t个小时的交通违法程度。每条边x(i,j,t)表示交警是否在第t小时开始时从结点v(i,t-1)移动到结点v(j,t)，如果x(i,j,t)＝1则表示存在这样的移动，该式为0则表示不存在这样的移动。

[0069] 如图3是某一交警巡逻路线的建模实例示意图；该实例存在3个交通违法热点候选区域和2条巡逻路径，工作时长为3，实线表示的x(i,j,t)为1，而虚线则为0，则该图中一条路径从交警起始站点开始，在第一个小时于热点候选区域r_1执勤，并且在第二个小时停留在该区域，最后回到起始站点。另一条路径从起始站点开始，到r_3区域执勤，再转移到r_2区域，最后回到交警起始站点。

[0070] 由此，我们构建3个张量邻接张量X，损失张量D和增益张量L，如图4所示；临界张量X表示巡逻路径，x(i,j,t)表示是否存在从结点v(i,t-1)到v(i,t)的移动，存在则为1，否则为 0；损失张量D表示某时刻从某点到某点的移动距离成本，d(i,j,t)，表示从v(i,t-1)到v(i,t)的移动距离；增益张量L表示某时刻从某点到某点的交通违法热点覆盖增益，l(i,j,t)等于 Level(j,t)表示v(j,t)的交通违法程度。

[0071] 通过以上定义，我们将巡逻路线优化问题转化为整数线性规划问题，假设我们要规划k 条路径，N个交通违法热点区域，T个工作时长，则我们要求的整数线性规划问题可以

写成如下形式：

[0072] $\mathrm{argmax}_X L \cdot X$ ①

[0073] subjectto

[0074] $$\sum_{j=1}^{N} x(0,j,1) = \sum_{i=0}^{N}\sum_{j=0}^{N} x(i,j,1) = \sum_{i=0}^{N} x(i,0,T) = \sum_{j=0}^{N}\sum_{t=1}^{T} x(0,j \quad ②$$

[0075] $$\sum_{i=0}^{N}\sum_{t=1}^{T-1} x(i,0,t) = 0 \quad ③$$

[0076] $$\sum_{i=1}^{N} x(i,p,t) = \sum_{j=0}^{N} x(p,j,t+1) \le 1, \forall\, t \in [1,T-1], p \in [1,N] \quad ④$$

[0077] $D \cdot X \leqslant \mu$ ⑤

[0078] $$x(i,j,t) \in \{0,1\}, \forall i \in [0,N], j \in [0,N], t \in [1,T] \quad ⑥$$

[0079] ①式表示我们的优化目标，即要找到这样的邻接张量X使得L·X即巡逻路线覆盖的交通违法热点增益最大；限制条件②中，$\sum_{j=1}^{N} x(0,j,1)$表示了从起始点出发的路径数量，$\sum_{i=0}^{N} x(i,0,T)$表示最终回到起始点的路径数量，$\sum_{j=1}^{N} x(0,j,1) = \sum_{i=0}^{N}\sum_{j=0}^{N} x(i,j,1)$表示所有路线都从起始点开始，$\sum_{i=0}^{N} x(i,0,T) = \sum_{j=0}^{N}\sum_{t=1}^{T} x(0,j,t)$表示所有路线都到起始结束，且路径数量为k；限制条件③表示没有路径会在工作时间结束前回到起始点；②③两个限制条件共同保证了图2这个有向图的基本结构。限制条件④表示到达某个结点的路径数量等于离开该结点的数量，并且路径之间不会存在点交叉。限制条件⑤⑥表示总体的路径长度不会超过一个合理的阈值μ。该阈值可根据不同情境进行具体设定。该整数线性规划问题，可由Matlab 的凸优化建模系统CVX直接求解，得到最优解，即k条交警巡逻路径。

[0080] 实施例另一方面，还提供一种基于交通违法热点预测的巡逻路线规划系统，包括：

[0081] 事件匹配单元：配置于将离散的交通违法事件匹配到道路网络，得到交通违法热点候选区域；

[0082] 数据集获取单元：配置于得出所述交通违法热点候选区域的交通违法时间序列，计算出时间序列下的交通违法热点候选区域的违法程度，并提取时间序列下的交通违法热点候选区域的时空情景特征，将交通违法热点候选区域以及其交通违法时间序列、违法程度、时空情景特征组成数据集；

[0083] 预测模型训练单元：配置于采用拔靴抽样法从数据集中抽取三组训练数据集，分别用三组训练数据集通过协同训练与主动学习训练三组预测模型，所述预测模型采用树集成模型；

[0084] 模型预测单元：配置于利用所述三组预测模型，对无标签交通数据进行分类，预测未来交通违法热点的分布情况；

[0085] 巡逻路线获取单元：配置于根据交通违法热点分布情况，进行交警巡逻路线的建模和优化。

[0086] 本发明的再一方面，提供一种存储有计算机指令的非瞬时计算机可读存储介质，所述计算机指令用于执行本发明所提供的一种基于交通违法热点预测的巡逻路线规划方法。

[0087] 计算机指令(也称作程序、软件、软件应用或代码)包括可编程处理器的机器指令，并且可以利用高级过程和/或面向对象的编程语言、和/或汇编/机器语言来实施这些计算机指令。术语“机器可读介质”和“计算机可读介质”指的是用于将机器指令和/或数据提供

给可编程处理器的任何计算机程序产品、设备、和/或装置(例如,磁盘、光盘、存储器、可编程逻辑装置(programmable logic device,PLD)),包括,接收作为机器可读信号的机器指令的机器可读介质。

[0088] 本发明提出的基于交通违法热点预测的巡逻路线规划方法和系统,通过对城市中的交通违法热点进行建模预测来规划交警巡逻路线。该方法和系统能够低成本、高覆盖、实时的感知预测城市中的交通违法热点并规划交警巡逻路线,帮助城市交通管理部门全面深入地掌握城市中的交通违法热点,并且高效地规划交警巡逻路线。

[0089] 上述仅为本发明的具体实施方式,但本发明的设计构思并不局限于此,凡利用此构思对本发明进行非实质性的改动,均应属于侵犯本发明保护范围的行为。

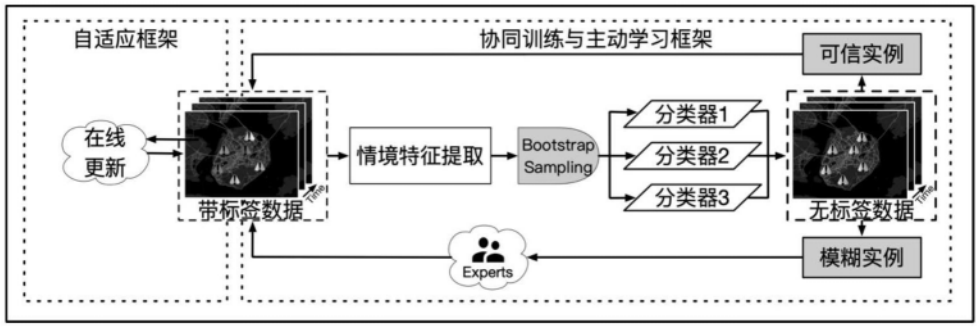

自适应框架
协同训练与主动学习框架
可信实例
在线
更新
带标签数据
情境特征提取
Bootstrap
Sampling
分类器1
分类器2
分类器3
无标签数据
Experts
模糊实例

图1

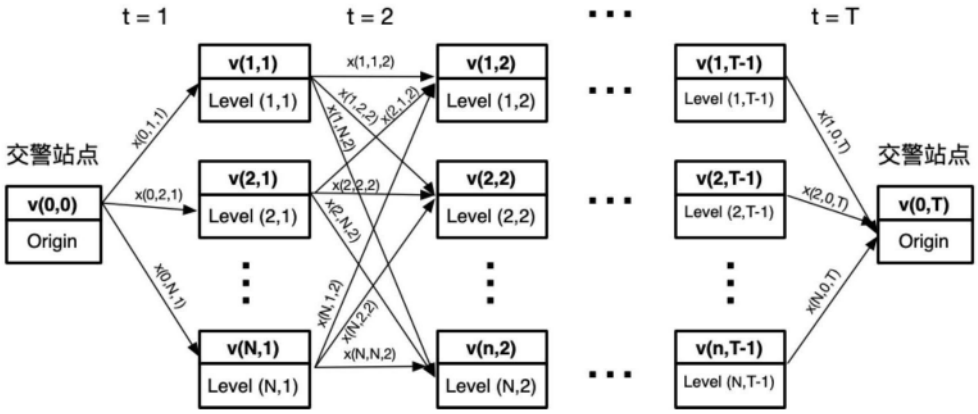

t = 1
t = 2
t = T
交警站点
v(0,0)
Origin
v(1,1)
Level (1,1)
v(2,1)
Level (2,1)
v(N,1)
Level (N,1)
v(1,2)
Level (1,2)
v(2,2)
Level (2,2)
v(n,2)
Level (N,2)
v(1,T-1)
Level (1,T-1)
v(2,T-1)
Level (2,T-1)
v(n,T-1)
Level (N,T-1)
v(0,T)
Origin
交警站点

图2

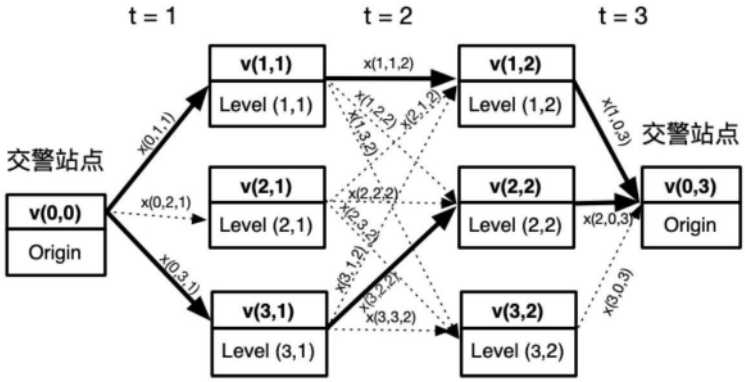
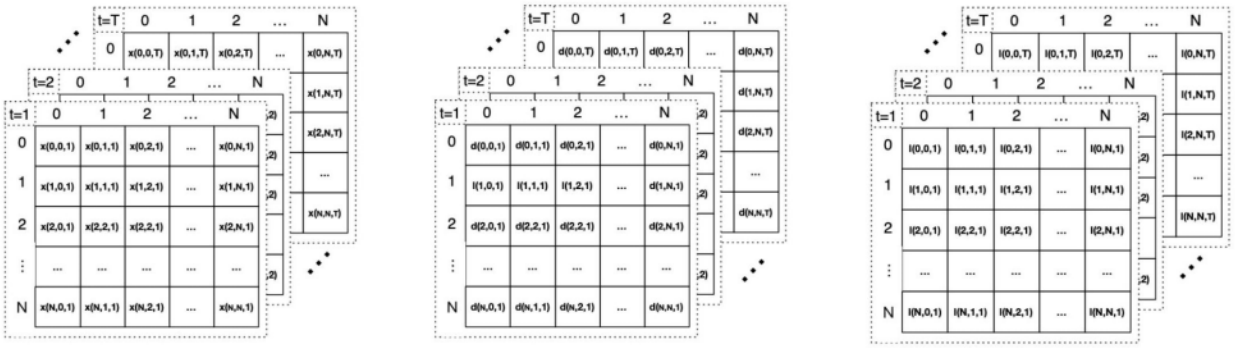

t = 1
t = 2
t = 3
交警站点
v(0,0)
Origin
v(1,1)
Level (1,1)
v(2,1)
Level (2,1)
v(3,1)
Level (3,1)
v(1,2)
Level (1,2)
v(2,2)
Level (2,2)
v(3,2)
Level (3,2)
v(0,3)
Origin
交警站点

图3


t=T
t=2
t=1
x(0,0,T)
x(0,1,T)
x(0,2,T)
x(0,N,T)
x(1,N,T)
x(2,N,T)
x(N,N,T)
x(0,0,1)
x(0,1,1)
x(0,2,1)
x(0,N,1)
x(1,0,1)
x(1,1,1)
x(1,2,1)
x(1,N,1)
x(2,0,1)
x(2,2,1)
x(2,N,1)
x(N,0,1)
x(N,1,1)
x(N,2,1)
x(N,N,1)
d(0,0,T)
d(0,1,T)
d(0,2,T)
d(0,N,T)
d(1,N,T)
d(2,N,T)
d(N,N,T)
d(0,0,1)
d(0,1,1)
d(0,2,1)
d(0,N,1)
d(1,0,1)
d(1,1,1)
d(1,2,1)
d(1,N,1)
d(2,0,1)
d(2,2,1)
d(2,N,1)
d(N,0,1)
d(N,1,1)
d(N,2,1)
d(N,N,1)
I(0,0,T)
I(0,1,T)
I(0,2,T)
I(0,N,T)
I(1,N,T)
I(2,N,T)
I(N,N,T)
I(0,0,1)
I(0,1,1)
I(0,2,1)
I(0,N,1)
I(1,0,1)
I(1,1,1)
I(1,2,1)
I(1,N,1)
I(2,0,1)
I(2,2,1)
I(2,N,1)
I(N,0,1)
I(N,1,1)
I(N,2,1)
I(N,N,1)

(a) (b) (c)

图4