

WIPO/PCT

(74) 代理人: 深圳市世联合知识产权代理有限公司(SL INTELLECTUAL PROPERTY CO., LTD.); 中国广东省深圳市福田区沙头街道天安社区泰然四路29号天安创新科技广场一期A座805E, Guangdong 518000 (CN)。

图 2

[见续页]

(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, IT, 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, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW。

(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))。
- 根据申请人的请求, 在条约第21条(2)(a)所规定的期限届满之前进行。

by a joint server corresponding to the joint bus station; and calculating a model update parameter on the basis of the model gradient information and the random number, and transmitting the model update parameter to the target server and each joint server, such that the target server and the joint server iteratively update a regression prediction model by means of the model update parameter. The trained regression prediction model can be stored in a blockchain. By means of the method, the effect of short-time bus passenger flow prediction by means of a regression prediction model is effectively improved.

(57) 摘要: 一种短时公交站客流预测方法、装置、计算机设备及存储介质, 属于人工智能技术领域, 应用于智慧交通领域中, 包括: 接收公交站位置信息和目标公交站, 构建公交站无向网络, 确定公交站无向网络中符合预设条件的公交站, 作为联合公交站; 接收目标公交站对应的目标服务器传输的模型梯度信息和联合公交站对应的联合服务器传输的随机数; 基于所述模型梯度信息和所述随机数计算模型更新参数, 并将所述模型更新参数传输至目标服务器和各个联合服务器, 以使所述目标服务器和联合服务器通过所述模型更新参数迭代更新回归预测模型。其中, 训练后的回归预测模型可存储于区块链中。所述方法有效提高回归预测模型对短时公交车客流的预测效果。

短时公交站客流预测方法、装置、计算机设备及存储介质

本申请要求于2020年10月29日提交中国专利局、申请号为202011182335.1，发明名称为“短时公交站客流预测方法、装置、计算机设备及存储介质”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

技术领域

本申请涉及人工智能技术领域，尤其涉及短时公交站客流预测方法、装置、计算机设备及存储介质。

背景技术

公共交通对提供交通运力，缓解交通拥堵起着重要作用，然而目前国内大部分城市的公共交通运力不足，特别是高峰时段的城市公交运力不足，于是预估公交站上车客流量的各种方法应运而生，通过对公交站上车客流量的预估，实现交通运力的及时调整，甚至可以将预估结果发送至相关用户的客户端，便于人们的出行安排等。

目前，有关于通过回归预测模型预估公交站上车客流量的方法包括基于单个公交站点或者区域公交站进行公交车客流预测。发明人意识到，对于单个公交站点的方案来说，训练样本较为单一，训练效果较差，未能考虑周边公交站的客流对目标公交站的影响。对于区域公交站的方案来说，仅基于区域来确定公交站，确定的公交站与目标公交站之间的相关性较弱，导致预测效果较差。

发明内容

本申请实施例的目的在于提出一种短时公交站客流预测方法、装置、计算机设备及存储介质，提升了回归预测模型对于短时公交车客流预测的准确性。

第一方面，本申请实施例提供一种短时公交站客流预测方法，采用了如下所述的技术方案：

一种短时公交站客流预测方法，应用于中心服务器，包括下述步骤：

接收公交站位置信息和目标公交站，基于所述公交站位置信息构建公交站无向网络，并基于所述目标公交站确定所述公交站无向网络中符合预设条件的公交站，作为联合公交站，其中，目标公交站对应目标服务器，每个所述联合公交站各自对应一个联合服务器；

接收目标公交站对应的目标服务器传输的模型梯度信息和联合公交站对应的联合服务器传输的随机数，并基于所述模型梯度信息和所述随机数计算模型更新参数，并将所述模型更新参数传输至目标服务器和各个联合服务器，以使所述目标服务器和联合服务器通过所述模型更新参数迭代更新回归预测模型。

第二方面，本申请实施例提供一种短时公交站客流预测方法，采用了如下所述的技术方案：

一种短时公交站客流预测方法，应用于联合服务器，包括下述步骤：

接收目标时间窗和对应的联合公交站的乘客乘车信息，基于所述乘客乘车信息和所述目标时间窗构建对应的联合公交站的训练样本集合；

基于所述训练样本集合构建所述联合公交站的回归预测模型，直至达到预设的构建条件，获得构建后的回归预测模型的模型参数；

接收中心服务器传输的对应的联合公交站的边权，计算所述边权与回归预测模型的模型参数之积，作为计算结果，并基于所述计算结果生成多个随机数，将所述随机数分发至各个所述联合服务器；

所有所述联合服务器将接收到的随机数传输至中心服务器，以使所述中心服务器利用所述梯度信息和所述随机数计算模型更新参数；

当目标服务器和各个联合服务器接收到所述中心服务器传输的模型更新参数时，基于所述模型更新参数迭代回归预测模型，直至达到预设的停止条件，获得训练后的回归预测模型。

第三方面，本申请实施例还提供一种短时公交站客流预测装置，采用了如下所述的技术方案：

一种短时公交站客流预测装置，所述短时公交站客流预测装置包括中心服务器，所述中心服务器包括：

网络构建模块，用于接收公交站位置信息和目标公交站，基于所述公交站位置信息构建公交站无向网络，并基于所述目标公交站确定所述公交站无向网络中符合预设条件的公交站，作为联合公交站，其中，目标公交站对应目标服务器，每个所述联合公交站各自对应一个联合服务器；

参数计算模块，用于接收目标公交站对应的目标服务器传输的模型梯度信息，和联合公交站对应的联合服务器传输的随机数，并基于所述模型参数和所述随机数计算模型更新参数，并将所述模型更新参数传输至目标服务器和各个联合服务器，以使所述目标服务器和联合服务器通过所述模型更新参数迭代更新回归预测模型。

第五方面，本申请实施例还提供一种短时公交站客流预测装置，采用了如下所述的技术方案：

一种短时公交站客流预测装置，所述短时公交站客流预测装置包括联合服务器，所述联合服务器包括：

联合样本构建模块，用于接收目标时间窗和对应的联合公交站的乘客乘车信息，基于所述乘客乘车信息和所述目标时间窗构建对应的联合公交站的训练样本集合；

联合模型构建模块，用于基于所述训练样本集合构建所述联合公交站的回归预测模型，直至达到预设的构建条件，获得各自的构建后的回归预测模型的模型参数；

分发模块，用于接收中心服务器传输的对应的联合公交站的边权，计算所述边权与回归预测模型的模型参数之积，作为计算结果，并基于所述计算结果生成多个随机数，将所述随机数分发至各个所述联合服务器；

联合数据传输模块，用于将接收到的随机数传输至中心服务器，以使所述中心服务器利用所述梯度信息和所述随机数计算模型更新参数；

联合迭代更新模块，用于当接收到所述中心服务器传输的模型更新参数时，基于所述模型更新参数迭代回归预测模型，直至达到预设的停止条件，获得训练后的回归预测模型。

第六方面，本申请实施例还提供一种计算机设备，采用了如下所述的技术方案：

一种计算机设备，包括存储器和处理器，所述存储器中存储有计算机可读指令，所述处理器执行所述计算机可读指令时实现如下所述的短时公交站客流预测方法的步骤：

接收公交站位置信息和目标公交站，基于所述公交站位置信息构建公交站无向网络，并基于所述目标公交站确定所述公交站无向网络中符合预设条件的公交站，作为联合公交站，其中，目标公交站对应目标服务器，每个所述联合公交站各自对应一个联合服务器；

接收目标公交站对应的目标服务器传输的模型梯度信息和联合公交站对应的联合服务器传输的随机数，并基于所述模型梯度信息和所述随机数计算模型更新参数，并将所述模型更新参数传输至目标服务器和各个联合服务器，以使所述目标服务器和联合服务器通过所述模型更新参数迭代更新回归预测模型。

一种计算机设备，包括存储器和处理器，所述存储器中存储有计算机可读指令，所述处理器执行所述计算机可读指令时实现如下所述的短时公交站客流预测方法的步骤：

接收目标时间窗和对应的联合公交站的乘客乘车信息，基于所述乘客乘车信息和所述目标时间窗构建对应的联合公交站的训练样本集合；

基于所述训练样本集合构建所述联合公交站的回归预测模型，直至达到预设的构建条件，获得构建后的回归预测模型的模型参数；

接收中心服务器传输的对应的联合公交站的边权，计算所述边权与回归预测模型的模型参数之积，作为计算结果，并基于所述计算结果生成多个随机数，将所述随机数分发至各个所述联合服务器；

所有所述联合服务器均将接收到的随机数传输至中心服务器，以使所述中心服务器利用所述梯度信息和所述随机数计算模型更新参数；

当接收到所述中心服务器传输的模型更新参数时，基于所述模型更新参数迭代回归预测模型，直至达到预设的停止条件，获得训练后的回归预测模型。

第七方面，本申请实施例还提供一种计算机可读存储介质，采用了如下所述的技术方案：

一种计算机可读存储介质，所述计算机可读存储介质上存储有计算机可读指令，所述计算机可读指令被处理器执行时实现如下所述的短时公交站客流预测方法的步骤：

接收公交站位置信息和目标公交站，基于所述公交站位置信息构建公交站无向网络，并基于所述目标公交站确定所述公交站无向网络中符合预设条件的公交站，作为联合公交站，其中，目标公交站对应目标服务器，每个所述联合公交站各自对应一个联合服务器；

接收目标公交站对应的目标服务器传输的模型梯度信息和联合公交站对应的联合服务器传输的随机数，并基于所述模型梯度信息和所述随机数计算模型更新参数，并将所述模型更新参数传输至目标服务器和各个联合服务器，以使所述目标服务器和联合服务器通过所述模型更新参数迭代更新回归预测模型。

一种计算机可读存储介质，所述计算机可读存储介质上存储有计算机可读指令，所述计算机可读指令被处理器执行时实现如下所述的短时公交站客流预测方法的步骤：

接收目标时间窗和对应的联合公交站的乘客乘车信息，基于所述乘客乘车信息和所述目标时间窗构建对应的联合公交站的训练样本集合；

基于所述训练样本集合构建所述联合公交站的回归预测模型，直至达到预设的构建条件，获得构建后的回归预测模型的模型参数；

接收中心服务器传输的对应的联合公交站的边权，计算所述边权与回归预测模型的模型参数之积，作为计算结果，并基于所述计算结果生成多个随机数，将所述随机数分发至各个所述联合服务器；

所有所述联合服务器均将接收到的随机数传输至中心服务器，以使所述中心服务器利用所述梯度信息和所述随机数计算模型更新参数；

当接收到所述中心服务器传输的模型更新参数时，基于所述模型更新参数迭代回归预测模型，直至达到预设的停止条件，获得训练后的回归预测模型。

与现有技术相比，本申请实施例主要有以下有益效果：

中心服务器通过公交站位置信息建立公交站无向网络，便于后续通过公交站无向网络来计算出符合条件的联合公交站。相对于仅是通过区域划分确定联合学习的公交站的方式来说，本申请通过公交站无向网络计算从而确定的联合公交站与目标公交站的相关性较强。联合服务器均通过目标时间窗确定公交站基于目标时间窗的客流量，基于目标时间窗内的客流量构建各自的回归预测模型，实现公交车的短时预测的模型构建。联合服务器将构建后的回归预测模型的参数与联合公交站对应的边权相乘，并生成随机数进行分发，保护了联合公交站数据的隐私性，防止在传输过程中泄露。中心服务器基于目标服务器传输的模型梯度信息和联合服务器传输的随机数综合计算模型更新参数，实现了目标公交站和与目标公交站相关性较强的联合公交站之间的联合学习，通过迭代获得最终训练后的回归预测模型，提升了回归预测模型预测的准确性。

附图说明

为了更清楚地说明本申请中的方案，下面将对本申请实施例描述中所需要使用的附图作一个简单介绍，显而易见地，下面描述中的附图是本申请的一些实施例，对于本领域普

通技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其他的附图。

图 1 是本申请可以应用于其中的示例性系统架构图;

图 2 是根据本申请的短时公交站客流预测方法的一个实施例的流程图;

图 3 是根据本申请的短时公交站客流预测方法的公交站无向网络的结构示意图;

图 4 是根据本申请的短时公交站客流预测装置的一个实施例的结构示意图;

图 5 是根据本申请的计算机设备的一个实施例的结构示意图。

附图标记: 101、网络构建模块; 102、参数计算模块; 201、目标样本构建模块; 202、目标模型构建模块; 203、目标数据传输模块; 204、目标迭代更新模块; 301、联合样本构建模块; 302、联合模型构建模块; 303、分发模块; 304、联合数据传输模块; 305、联合迭代更新模块。

具体实施方式

除非另有定义,本文所使用的所有的技术和科学术语与属于本申请的技术领域的技术人员通常理解的含义相同;本文中在申请的说明书中所使用的术语只是为了描述具体的实施例的目的,不是旨在于限制本申请;本申请的说明书和权利要求书及上述附图说明中的术语“包括”和“具有”以及它们的任何变形,意图在于覆盖不排他的包含。本申请的说明书和权利要求书或上述附图中的术语“第一”、“第二”等是用于区别不同对象,而不是用于描述特定顺序。

在本文中提及“实施例”意味着,结合实施例描述的特定特征、结构或特性可以包含在本申请的至少一个实施例中。在说明书中的各个位置出现该短语并不一定均是指相同的实施例,也不是与其它实施例互斥的独立的或备选的实施例。本领域技术人员显式地和隐式地理解的是,本文所描述的实施例可以与其它实施例相结合。

为了使本技术领域的人员更好地理解本申请方案,下面将结合附图,对本申请实施例中的技术方案进行清楚、完整地描述。

本申请实施例的短时公交客流预测方法应用于短时公交客流预测系统中。如图 1 所示,系统架构包括中心服务器、目标服务器、联合服务器、第一客户端和第二客户端。目标服务器与中心服务器之间、联合服务器与中心服务器之间、第一客户端与目标服务器之间、第二客户端与联合服务器之间均通过网络进行连接,网络可以包括各种连接类型,例如有线、无线通信链路或者光纤电缆等等。

用户可以使用第一客户端或第二客户端通过网络与目标服务器或联合服务器交互,以接收或发送消息等。第一客户端和第二客户端上可以安装有各种通讯客户端应用,例如网页浏览器应用、购物类应用、搜索类应用、即时通信工具、邮箱客户端、社交平台软件等。

第一客户端和第二客户端可以是具有显示屏并且支持网页浏览的各种电子设备,包括但不限于智能手机、平板电脑、电子书阅读器、MP3 播放器(Moving Picture Experts Group Audio Layer III,动态影像专家压缩标准音频层面 3)、MP4(Moving Picture Experts Group Audio Layer IV,动态影像专家压缩标准音频层面 4)播放器、膝上型便携计算机和台式计算机等等。

应该理解,图 1 中的中心服务器、目标服务器、联合服务器、第一客户端和第二客户端的数目仅仅是示意性的。根据实现需要,可以具有任意数目的中心服务器、目标服务器、联合服务器、第一客户端和第二客户端。

继续参考图 2,示出了根据本申请的短时公交站客流预测方法的一个实施例的流程图。所述的短时公交站客流预测方法,包括以下步骤:

S1: 中心服务器接收公交站位置信息和目标公交站,基于所述公交站位置信息构建公交站无向网络,并基于所述目标公交站确定所述公交站无向网络中符合预设条件的公交站,作为联合公交站,其中,目标公交站对应目标服务器,每个所述联合公交站各自对应一个联合服务器。

在本实施例中,可以根据实际需要进行调整。基于所述目标公交站确定所述公交站无向网络中的联合公交站,提升后续联合公交站的相关数据与目标公交站的数据之间的相关性。其中,目标公交站和每个联合公交站分别对应一个数据中心,后续各公交站基于训练样本的模型构建过程,在各自的数据中心完成,避免数据泄露。

在本实施例中,短时公交站客流预测方法运行于其上的电子设备(例如图1所示的中心服务器)可以通过有线连接方式或者无线连接方式接收公交站位置信息和目标公交站。需要指出的是,上述无线连接方式可以包括但不限于3G/4G连接、WiFi连接、蓝牙连接、WiMAX连接、Zigbee连接、UWB(ultra wideband)连接、以及其他现在已知或将来开发的无线连接方式。

具体的,在步骤S1中,即所述基于所述公交站位置信息构建公交站无向网络,并基于所述目标公交站确定所述公交站无向网络中符合预设条件的公交站,作为联合公交站的步骤包括:

基于所述公交站位置信息将所述公交站作为所述公交站无向网络的节点;

将任意两个公交站均进行连线,作为所述公交站无向网络的边,获得所述公交站无向网络;

计算所述目标公交站与各公交站之间的边权 S_{if} ;

将对应的边权小于预设阈值的公交站作为所述联合公交站。

在本实施例中,公交站为网络的节点,任意两公交站的连线为网络的边,通过边权确定联合公交站,联合公交站即为与目标公交站 i 关系较为密切的公交站,便于后续与目标公交站联合学习,避免了单个公交站点训练数据维度低,预测结果不准确的情况。

如图3所示,图3是根据本申请的短时公交站客流预测方法的公交站无向网络的结构示意图。其中,通过如下公式计算所述边权 S_{if} :

$$S_{if} = a \cdot DIS_{if} + b \cdot |POP_{if}|$$

其中, a , b 均为超参数,且满足 $a+b=1$, DIS_{if} 为所述目标公交站 i 与任意公交站 f 之间的距离, POP_{if} 为目标公交站 i 与公交站 f 在预设范围内的人口总数差值为。

在本实施例中, POP_{if} 为目标公交站 i 与公交站 f 在预设范围内的人口总数差值。其中,所述预设范围为以公交站为圆心, N (km)为半径的圆形覆盖区域, N 为自定义数值。人口总数=人口密度×预设范围的面积,其中,人口密度即表示单位土地面积上的人口数量,可通过调用浏览器查询获得。通过上述公式计算边权,充分考虑了人口密度的情况,而不仅仅只是考虑上车客流量,扩展了后续模型更新参数的维度。

S2:第一客户端通过预设的本地差分隐私算法,扰动本地存储的目标公交站的乘客乘车信息,生成第一扰动数据,并将第一扰动数据发送至目标服务器。

在本实施例中,本申请中乘客乘车信息指乘客搭公交车时的对应的公交站和上车时间,需要加密的内容为乘客是否在当前公交站上车,每个记录都有其对应的时间信息。用户刷卡记录存储在自己手机,银行卡中,第一客户端通过本地差分隐私算法扰动目标公交站的乘客乘车信息,保护了用户的隐私,避免了直接将用户的刷卡记录发送至目标服务器。目标服务器根据扰动数据估算上车总客流量。具体为:对于真实集合中的每一个数据都添加扰动,即添加一个概率计算,形成概率集合,将概率集合发送至目标服务器。目标服务器根据所述概率集合随机生成 $y'_{ij} = 0$ 或者 1 。将随机生成的 y'_{ij} 通过如下公式计算即可得到公交站的总上车人数 $c = \sum_{j=1}^J \frac{y'_{ij} \cdot (e^\epsilon + 1) - 1}{e^\epsilon - 1}$ 。其中, j 指第 j 个用户,通过时间窗可以筛选扰动数据,进而通过上述方式获得时间窗对应的上车人数,即上车客流量。

具体的,通过如下公式计算所述第一扰动数据:

$$P(y'_{ij} = 1) = \frac{1}{e^\epsilon + 1} + y_{ij} \cdot \frac{e^\epsilon - 1}{e^\epsilon + 1}$$

$$P(y'_{ij} = 0) = 1 - P(y'_{ij} = 1)$$

其中, $y_{ij} \in [0, 1]$, 1 表示用户在目标公交站刷卡上车, 0 表示用户未在目标公交站刷卡上车, ε 为超参数。

在本实施例中, 本申请通过上述公式对乘客乘车信息添加扰动, 实现数据中心既无法获知实际数据, 保护了用户的隐私, 又可以根据扰动数据估算上车客流量, 进行后续计算。

S3: 所述目标公交站对应的目标服务器接收目标时间窗和第一客户端发送的第一扰动数据, 基于所述第一扰动数据和所述目标时间窗构建所述目标公交站的训练样本集合。

在本实施例中, 本申请中目标时间窗是自定义的。目标服务器基于目标时间窗构建训练样本集合, 使得训练样本集合与目标时间窗具有相关性, 从而能实现后续的通过训练后的回归预测模型准确预测目标时间窗内的上车客流量。

具体的, 在步骤 S3 中, 即所述基于所述第一扰动数据和所述目标时间窗构建公交站的训练样本集合的步骤包括:

基于所述第一扰动数据, 计算在预设天数中, 所述目标时间窗同时段之前的预设个数 D 的时间窗对应的上车客流量, 作为训练样本的第一输入参数 X , 所述目标时间窗同时段的上车客流量, 作为训练样本的第二输入参数 Y , 生成训练样本集合 $(X_1, Y_1) (X_2, Y_2) (X_3, Y_3) \dots (X_m, Y_m)$, 其中, 所述 X_m 具有的特征维数为 D , D 为整数, m 表示第 m 天。

在本实施例中, 以所述目标公交站点在 n 天中与目标时间窗同时段之前的 D 个时间窗内的公交客流量作为训练样本的特征, 即第一输入参数 X , 以所述目标区域公交站点在 n 天中与目标时间窗同时段的上车客流量作为训练样本的第二输入参数 Y , 构建训练样本 $\{(X_1, Y_1), (X_2, Y_2), \dots, (X_m, Y_m), \dots, (X_n, Y_n)\}$, 每一个样本 (X_m, Y_m) 都包含两个输入参数, 其中, X_m 具有的特征维数为 D , D 为整数。

以下以 D 为 2, 即特征维度数为 2 进行举例说明:

第一天的下午 1 点至 2 点的客流量和下午 2 点至 3 点的客流量组成 $X1$ (2 维向量, 包括两个客流量), 第一天的下午 3 点至 4 点的客流量为 $Y1$, (X_1, Y_1) 组成一对训练样本; 第二天的下午 1 点至 2 点的客流量和下午 2 点至 3 点的客流量组成 $X2$ (2 维向量, 包括两个客流量), 第二天的下午 3 点至 4 点的客流量为 $Y2$, (X_2, Y_2) 组成一对训练样本。第 m 天的下午 1 点至 2 点的客流量和下午 2 点至 3 点的客流量组成 X_m (2 维向量, 包括两个客流量), 第 m 天的下午 3 点至 4 点的客流量为 Y_m , (X_m, Y_m) 组成一对训练样本。

S4: 目标服务器基于所述训练样本集合构建所述目标公交站的回归预测模型, 直至达到预设的构建条件, 获得构建后的回归预测模型的模型参数。

在本实施例中, 构建目标公交站的回归预测模型, 完成回归预测模型的初步训练, 获得构建后的回归预测模型的模型参数, 便于后续的进一步处理。

具体的, 在步骤 S4 中, 所述目标服务器基于所述训练样本集合构建所述目标公交站的回归预测模型, 直至达到预设的构建条件, 获得构建后的回归预测模型的模型参数的步骤包括:

目标服务器随机初始化预设的回归预测模型的模型参数, 其中, 每一天分别对应一个预设的回归预测模型;

基于所述训练样本集合中的训练样本分别训练每天对应的回归预测模型 $\hat{y}_m$, 其中, 所述回归预测模型的公式为:

$$\hat{y}_m = w_0 + \sum_{d=1}^D w_d X_d$$

其中, $\hat{y}_m$ 表示模型预测的第 m 天的目标时间窗同时段的上车客流量; w_0 和 w_d 为所述随机初始化参数; X_d 表示在第 d 个时间窗中的上车客流量;

直至所述回归预测模型达到预设的构建条件, 获得构建后的回归预测模型的模型参数, 其中, 所述预设的构建条件为:

$$\arg \min \frac{1}{2} \sum_{m=1}^n (\hat{y}_m - y_m)^2$$

其中, y_m 表示第 m 天的目标时间窗同时段的上车客流量。

在本实施例中, $\arg \min$ 指使其后面的式子达到最小值时的变量的取值, 本申请通过该构建条件, 实现基于各公交站对应的训练样本对各公交站的回归预测模型的初步训练。

S5: 目标服务器基于所述模型参数计算目标公交站的模型梯度信息, 并将所述模型梯度信息传输至中心服务器。

在本实施例中, 根据如下公式计算所述目标公交站的模型梯度信息:

$$\nabla w_d^{i(t)} = \begin{cases} 1 & d = 0 \\ (\hat{y}_m^t - y_m) x_d & d > 0 \end{cases}$$

$\nabla w_d^{i(t)}$ 是目标公交站 i 在 t 次迭代时候的模型参数梯度, d 表示维度, 即第 d 个时间窗。

S6: 第二客户端分别通过预设的本地差分隐私算法, 扰动本地存储的联合公交站的乘客乘车信息, 分别生成第二扰动数据, 并将第二扰动数据发送至各自对应的联合服务器。

在本实施例中, 本申请中乘客乘车信息指乘客搭公交车时的对应的公交站和上车时间, 需要加密的内容为乘客是否在当前公交站上车, 每个记录都有其对应的时间信息。用户刷卡记录存储在自己手机, 银行卡中, 第二客户端通过本地差分隐私算法扰动目标公交站的乘客乘车信息, 保护了用户的隐私, 避免了直接将用户的刷卡记录发送至联合服务器。联合服务器根据扰动数据估算上车总客流量。具体为: 对于真实集合中的每一个数据都添加扰动, 即添加一个概率计算, 形成概率集合, 将概率集合发送至联合服务器。联合服务器根据所述概率集合随机生成 $y'_{ij} = 0$ 或者 1 。将随机生成的 y'_{ij} 通过如下公式计算即可得到公交站的总上车人数 $c = \sum_{j=1}^J \frac{y'_{ij} \cdot (e^\epsilon + 1) - 1}{e^\epsilon - 1}$ 。其中, j 指第 j 个用户, 通过时间窗可以筛选扰动数据, 进而通过上述方式获得时间窗对应的上车人数, 即上车客流量。

具体的, 通过如下公式生成第二扰动数据:

$$P(y'_{ij} = 1) = \frac{1}{e^\epsilon + 1} + y_{ij} \cdot \frac{e^\epsilon - 1}{e^\epsilon + 1}$$

$$P(y'_{ij} = 0) = 1 - P(y'_{ij} = 1)$$

其中, $y_{ij} \in [0, 1]$, 1 表示用户在当前联合公交站刷卡上车, 0 表示用户未在当前联合公交站刷卡上车, ϵ 为超参数。

在本实施例中, 本申请通过上述公式对乘客乘车信息添加扰动, 实现数据中心既无法获知实际数据, 保护了用户的隐私, 又可以根据扰动数据估算上车客流量, 进行后续计算。

S7: 所述联合公交站对应的联合服务器接收目标时间窗和各自对应的第二扰动数据, 基于所述第二扰动数据和所述目标时间窗构建各自的联合公交站的训练样本集合。

在本实施例中, 本申请中目标时间窗是自定义的。联合服务器分别基于目标时间窗构建各自的训练样本集合, 使得训练样本集合与目标时间窗具有相关性, 从而能实现后续的通过训练后的回归预测模型准确预测目标时间窗内的上车客流量。

具体的, 在步骤 S7 中, 即所述基于所述第二扰动数据和所述目标时间窗构建各自的联合公交站的训练样本集合的步骤包括:

基于各自对应的所述第二扰动数据, 计算在预设天数中, 各自所述目标时间窗同时段之前的预设个数 D 的时间窗对应的上车客流量, 作为训练样本的第一输入参数 X , 所述目标时间窗同时段的上车客流量, 作为训练样本的第二输入参数 Y , 生成各自的联合公交站的训练样本集合 $(X_1, Y_1) (X_2, Y_2) (X_3, Y_3) \dots (X_m, Y_m)$, 其中, 所述 X_m 具有的特征维数为 D , D 为整数, m 表示第 m 天。

在本实施例中, 以所述目标公交站点在 n 天中与目标时间窗同时段之前的 D 个时间窗内的公交客流量作为训练样本的特征, 即第一输入参数 X , 以所述目标区域公交站点在 n

天中与目标时间窗同时段的上车客流量作为训练样本的第二输入参数 Y ，构建训练样本 $\{(X_1, Y_1), (X_2, Y_2), \dots, (X_m, Y_m), \dots, (X_n, Y_n)\}$ ，每一个样本 (X_m, Y_m) 都包含两个输入参数，其中， X_m 具有的特征维数为 D ， D 为整数。

以下以 D 为 2，即特征维度数为 2 进行举例说明：

第一天的下午 1 点至 2 点的客流量和下午 2 点至 3 点的客流量组成 $X1$ (2 维向量，包括两个客流量)，第一天的下午 3 点至 4 点的客流量为 $Y1$ ， (X_1, Y_1) 组成一对训练样本；第二天的下午 1 点至 2 点的客流量和下午 2 点至 3 点的客流量组成 $X2$ (2 维向量，包括两个客流量)，第二天的下午 3 点至 4 点的客流量为 $Y2$ ， (X_2, Y_2) 组成一对训练样本。第 m 天的下午 1 点至 2 点的客流量和下午 2 点至 3 点的客流量组成 X_m (2 维向量，包括两个客流量)，第 m 天的下午 3 点至 4 点的客流量为 Y_m ， (X_m, Y_m) 组成一对训练样本。

S8：联合服务器基于各自的联合公交站的训练样本集合构建所述联合公交站的回归预测模型，直至达到预设的构建条件，获得各自对应的构建后的回归预测模型的模型参数。

在本实施例中，各自构建联合公交站的回归预测模型，完成回归预测模型的初步训练，获得各自对应的构建后的回归预测模型的模型参数，便于后续的进一步处理。

具体的，在步骤 S8 中，即所述联合服务器基于所述训练样本集合构建所述联合公交站的回归预测模型，直至达到预设的构建条件，获得构建后的回归预测模型的模型参数的步骤包括：

联合服务器各自基于所述联合公交站的训练样本集合构建各自联合公交站的回归预测模型，直至达到预设的构建条件，获得各自对应的构建后的联合公交站的回归预测模型的模型参数的步骤包括：

联合服务器各自随机初始化预设的回归预测模型的模型参数，其中，每一天分别对应一个预设的回归预测模型；

联合服务器各自基于所述训练样本集合中的训练样本分别训练每天对应的回归预测模型 $\hat{y}_m$ ，其中，所述回归预测模型的公式为：

$$\hat{y}_m = w_0 + \sum_{d=1}^D w_d x_d$$

其中， $\hat{y}_m$ 表示模型预测的第 m 天的目标时间窗同时段的上车客流量； w_0 和 w_d 为所述随机初始化参数； x_d 表示在第 d 个时间窗中的上车客流量；

直至所述回归预测模型达到预设的构建条件，获得各自对应的构建后的回归预测模型的模型参数，其中，所述预设的构建条件为：

$$\arg \min \frac{1}{2} \sum_{m=1}^n (\hat{y}_m - y_m)^2$$

其中， y_m 表示第 m 天的目标时间窗同时段的上车客流量。

S9：联合服务器接收中心服务器传输的对应的联合公交站的边权，计算所述边权与回归预测模型的模型参数之积，作为计算结果，并基于所述计算结果生成多个随机数，将所述随机数分发至各个所述联合服务器。

在本实施例中，各个联合服务器分别计算对应联合公交站的边权与对应的联合公交站在 t 次迭代时的模型参数之积 $S_{if} \cdot w_d^{f(t)}$ ，并将边权与模型参数之积 $S_{if} \cdot w_d^{f(t)}$ 由密钥分享算法随机生成 k 个随机数分配给各个联合服务器。其中，所述 k 个数的和等于 $S_{if} \cdot w_d^{f(t)}$ ，所有联合服务器接收到的随机数之和即为 $\sum_{f \in N(i)} S_{if} \cdot w_d^{f(t)}$ 。通过随机数的方式保护了联合公交站中的模型参数，提高了数据的隐私性。

S10：所有所述联合服务器将接收到的随机数传输至中心服务器，以使所述中心服务器利用所述模型梯度信息和所述随机数计算模型更新参数。

在本实施例中，所有所述联合服务器将接收到的随机数传输至中心服务器，所有联合

服务器接收到的随机数之和即为 $\sum_{f \in N(i)} S_{if} \cdot w_d^{f(t)}$ 。通过传输随机数的方式提高数据隐私性，避免了传输过程中数据的泄露和篡改。

S11：中心服务器接收目标公交站对应的目标服务器传输的模型梯度信息和联合公交站对应的联合服务器传输的随机数，并基于所述模型梯度信息和所述随机数计算模型更新参数，并将所述模型更新参数传输至目标服务器和各个联合服务器，以使所述目标服务器和联合服务器通过所述模型更新参数迭代更新回归预测模型。

在本实施例中，通过目标公交站的模型梯度信息和联合公交站的随机数计算模型更新参数。其中，由上述可知，联合公交站是通过边权确定，与目标公交站有较强的关联性，避免了仅通过单个公交站点或区域公交站简单进行模型更新参数的计算，而导致训练结果不佳，预测效果差的情况。本申请通过目标公交站和与目标公交站之间具有符合边权筛选条件的联合公交站联邦训练，综合计算模型更新参数，提升了模型的预测效果。

需要说明的是：本申请的联合公交站区别于区域公交站，区域公交站仅考虑公交站之间的地理位置关系，并未考虑公交站之间在一定区域内的人口数量关系。

具体的，在步骤 S11 中，即所述基于所述模型梯度信息和所述随机数计算模型更新参数，并将所述模型更新参数传输至目标服务器和各个联合服务器的步骤包括：

所述模型更新参数的计算公式为：

$$w_d^{i(t+1)} = \sum_{f \in N(i)} S_{if} \cdot w_d^{f(t)} - \alpha \cdot \nabla w_d^{i(t)}$$

其中， $N(i)$ 为所述联合公交站组成的集合， $\sum_{f \in N(i)} S_{if} \cdot w_d^{f(t)}$ 为接收到的所有随机数之和， α 为学习率， $\nabla w_d^{i(t)}$ 为模型梯度信息；

将所述模型更新参数传输至目标服务器和各个联合服务器。

在本实施例中， $N(i)$ 是与公交站 i 进行联邦学习的公交站集合，即联合公交站组成的集合。其中，公式中的 $\sum_{f \in N(i)} S_{if} \cdot w_d^{f(t)}$ 来自于上文提及的由联合服务器传输的随机数之和。 α 为学习率，本文中设置 α 为 0.001，在实际应用中，可以根据实际需要改变学习率的大小，适用即可。通过此公式可以明确得知模型更新参数的计算方式。

S12：当目标服务器和各个联合服务器接收到所述中心服务器传输的模型更新参数时，基于所述模型更新参数迭代回归预测模型，直至达到预设的停止条件，获得训练后的回归预测模型。

在本实施例中，目标服务器和各个联合服务器均接收中心服务器传回的模型更新参数，根据模型更新参数更新各自的回归预测模型，再根据训练样本进行训练，重复执行上述训练步骤，并将训练结果即目标公交站的模型梯度信息和联合公交站的随机数发送至中心服务器，等待接收中心服务器下一次的模型更新参数。如此循环迭代各自的回归预测模型，直到达到预设的迭代次数 T ，获得训练后的回归预测模型。

需要说明的是，获得的训练后的回归预测模型用于预测目标公交站在目标时间窗同时段的客流量。训练后的回归预测模型可以传输给中心服务器，也可以存储在目标服务器或联合服务器中直接调用。完成模型训练后，通过中心服务器、目标服务器或联合服务器接收构建好的预测样本，将预测样本输入训练后的回归预测模型，即得到目标公交站在目标时间窗的客流量。例如：要预测某天下午 3 点至下午 4 点的客流量 Y ，在维度 D 为 2 时，对应的预测样本 X 为该天下午 1 点至 2 点的客流量和下午 2 点至 3 点的客流量。此时把 X 输入模型即可得到。

中心服务器通过公交站位置信息建立公交站无向网络，便于后续通过公交站无向网络来计算出符合条件的联合公交站。相对于仅是通过区域划分确定联合学习的公交站的方式来说，本申请通过公交站无向网络计算从而确定的联合公交站与目标公交站的相关性较强。联合服务器均通过目标时间窗确定公交站基于目标时间窗的客流量，基于目标时间窗内的客流量构建各自的回归预测模型，实现公交车的短时预测的模型构建。联合服务器将

构建后的回归预测模型的参数与联合公交站对应的边权相乘，并生成随机数进行分发，保护了联合公交站数据的隐私性，防止在传输过程中泄露。中心服务器基于目标服务器传输的模型梯度信息和联合服务器传输的随机数综合计算模型更新参数，实现了目标公交站和与目标公交站相关性较强的联合公交站之间的联合学习，通过迭代获得最终训练后的回归预测模型，提升了回归预测模型预测的准确性。

需要强调的是，为进一步保证上述训练后的回归预测模型的私密和安全性，上述训练后的回归预测模型还可以存储于一区块链的节点中。

本申请所指区块链是分布式数据存储、点对点传输、共识机制、加密算法等计算机技术的新型应用模式。区块链(Blockchain)，本质上是一个去中心化的数据库，是一串使用密码学方法相关联产生的数据块，每一个数据块中包含了一批次网络交易的信息，用于验证其信息的有效性(防伪)和生成下一个区块。区块链可以包括区块链底层平台、平台产品服务层以及应用服务层等。

本申请可应用于智慧交通领域中，从而推动智慧城市的建设。

本领域普通技术人员可以理解实现上述实施例方法中的全部或部分流程，是可以通过计算机可读指令来指令相关的硬件来完成，该计算机可读指令可存储于一计算机可读存储介质中，该计算机可读指令在执行时，可包括如上述各方法的实施例的流程。其中，前述的存储介质可为磁碟、光盘、只读存储记忆体(Read-Only Memory, ROM)等非易失性存储介质，或随机存储记忆体(Random Access Memory, RAM)等。

应该理解的是，虽然附图的流程图中的各个步骤按照箭头的指示依次显示，但是这些步骤并不是必然按照箭头指示的顺序依次执行。除非本文中有明确的说明，这些步骤的执行并没有严格的顺序限制，其可以以其他的顺序执行。而且，附图的流程图中的至少一部分步骤可以包括多个子步骤或者多个阶段，这些子步骤或者阶段并不必然是在同一时刻执行完成，而是可以在不同的时刻执行，其执行顺序也不必然是依次进行，而是可以与其他步骤或者其他步骤的子步骤或者阶段的至少一部分轮流或者交替地执行。

进一步参考图3，作为对上述图2所示方法的实现，本申请提供了一种短时公交站客流预测装置的一个实施例，该装置实施例与图2所示的方法实施例相对应，该装置具体可以应用于各种电子设备中。

如图3所示，本实施例所述的短时公交站客流预测装置包括：中心服务器、目标服务器和联合服务器、第一客户端和第二客户端。

所述中心服务器包括网络构建模块101和参数计算模块102。

网络构建模块101，用于接收公交站位置信息和目标公交站，基于所述公交站位置信息构建公交站无向网络，并基于所述目标公交站确定所述公交站无向网络中符合预设条件的公交站，作为联合公交站，其中，目标公交站对应目标服务器，每个所述联合公交站各自对应一个联合服务器；

参数计算模块102，用于接收目标公交站对应的目标服务器传输的模型梯度信息，和联合公交站对应的联合服务器传输的随机数，并基于所述模型参数和所述随机数计算模型更新参数，并将所述模型更新参数传输至目标服务器和各个联合服务器，以使所述目标服务器和联合服务器通过所述模型更新参数迭代更新回归预测模型。

第一客户端用于通过预设的本地差分隐私算法，扰动本地存储的目标公交站的乘客乘车信息，生成第一扰动数据，并将第一扰动数据发送至目标服务器。

所述目标服务器包括：目标样本构建模块201、目标模型构建模块202、目标数据传输模块203和目标迭代更新模块204。

目标样本构建模块201，用于接收目标时间窗和第一客户端发送的第一扰动数据，基于所述第一扰动数据和所述目标时间窗构建目标公交站的训练样本集合；

目标模型构建模块202，用于基于所述训练样本集合构建所述目标公交站的回归预测模型，直至达到预设的构建条件，获得构建后的回归预测模型的模型参数；

目标数据传输模块 203, 用于基于所述模型参数计算目标公交站的模型梯度信息, 并将所述模型梯度信息传输至中心服务器;

目标迭代更新模块 204, 用于当接收到所述中心服务器传输的模型更新参数时, 基于所述模型更新参数迭代回归预测模型, 直至达到预设的停止条件, 获得训练后的回归预测模型;

第二客户端用于分别通过预设的本地差分隐私算法, 扰动本地存储的联合公交站的乘客乘车信息, 分别生成第二扰动数据, 并将第二扰动数据发送至各自对应的联合服务器。

所述联合服务器包括: 联合样本构建模块 301、联合模型构建模块 302、分发模块 303、联合数据传输模块 304 和联合迭代更新模块 305。

联合样本构建模块 301, 用于接收目标时间窗和各自对应的各自对应的第二扰动数据, 基于所述第二扰动数据和所述目标时间窗构建各自的联合公交站的训练样本集合;

联合模型构建模块 302, 用于基于所述训练样本集合构建所述联合公交站的回归预测模型, 直至达到预设的构建条件, 获得各自的构建后的回归预测模型的模型参数;

分发模块 303, 用于接收中心服务器传输的对应的联合公交站的边权, 计算所述边权与回归预测模型的模型参数之积, 作为计算结果, 并基于所述计算结果生成多个随机数, 将所述随机数分发至各个所述联合服务器;

联合数据传输模块 304, 用于将接收到的随机数传输至中心服务器, 以使所述中心服务器利用所述模型梯度信息和所述随机数计算模型更新参数;

联合迭代更新模块 305, 用于当接收到所述中心服务器传输的模型更新参数时, 基于所述模型更新参数迭代回归预测模型, 直至达到预设的停止条件, 获得训练后的回归预测模型。

中心服务器通过公交站位置信息建立公交站无向网络, 便于后续通过公交站无向网络来计算出符合条件的联合公交站。相对于仅是通过区域划分确定联合学习的公交站的方式来说, 本申请通过公交站无向网络计算从而确定的联合公交站与目标公交站的相关性较强。联合服务器均通过目标时间窗确定公交站基于目标时间窗的客流量, 基于目标时间窗内的客流量构建各自的回归预测模型, 实现公交车的短时预测的模型构建。联合服务器将构建后的回归预测模型的参数与联合公交站对应的边权相乘, 并生成随机数进行分发, 保护了联合公交站数据的隐私性, 防止在传输过程中泄露。中心服务器基于目标服务器传输的模型梯度信息和联合服务器传输的随机数综合计算模型更新参数, 实现了目标公交站和与目标公交站相关性较强的联合公交站之间的联合学习, 通过迭代获得最终训练后的回归预测模型, 提升了回归预测模型预测的准确性。

网络构建模块 101 包括节点子模块、连线子模块、边权子模块和阈值子模块。节点子模块用于基于所述公交站位置信息将所述公交站作为所述公交站无向网络的节点; 连线子模块用于将任意两个公交站均进行连线, 作为所述公交站无向网络的边, 获得所述公交站无向网络; 边权子模块用于计算所述目标公交站与各公交站之间的边权 S_{if} ; 阈值子模块用于将对应的边权小于预设阈值的公交站作为所述联合公交站。

在本实施例的一些可选的实现方式中, 上述边权子模块进一步用于: 通过如下公式计算所述边权 S_{if} : $S_{if} = a \cdot DIS_{if} + b \cdot |POP_{if}|$, 其中, a, b 均为超参数, 且满足 $a+b=1$, DIS_{if} 为所述目标公交站 i 与任意公交站 f 之间的距离, POP_{if} 为目标公交站 i 与公交站 f 在预设范围内的人口总数差值。

在本实施例的一些可选的实现方式中, 上述第一客户端进一步用于通过如下公式计算所述第一扰动数据: $P(y'_{ij} = 1) = \frac{1}{e^\epsilon + 1} + y_{ij} \cdot \frac{e^\epsilon - 1}{e^\epsilon + 1}$, $P(y'_{ij} = 0) = 1 - P(y'_{ij} = 1)$, 其中, $y_{ij} \in [0, 1]$, 1 表示用户在目标公交站刷卡上车, 0 表示用户未在目标公交站刷卡上车, ϵ 为超参数。

在本实施例的一些可选的实现方式中, 上述目标样本构建模块 201 进一步用于: 基于所述第一扰动数据, 计算在预设天数中, 所述目标时间窗同时段之前的预设个数 D 的时间

窗对应的上车客流量，作为训练样本的第一输入参数 X ，所述目标时间窗同时段的上车客流量，作为训练样本的第二输入参数 Y ，生成训练样本集合 $(X_1, Y_1) (X_2, Y_2) (X_3, Y_3) \dots (X_m, Y_m)$ ，其中，所述 X_m 具有的特征维数为 D ， D 为整数， m 表示第 m 天。

目标模型构建模块 202 包括第一初始化子模块、第一训练子模块和第一构建子模块。第一初始化子模块用于随机初始化预设的回归预测模型的模型参数，其中，每一天分别对应一个预设的回归预测模型；第一训练子模块用于基于所述训练样本集合中的训练样本分别训练每天对应的回归预测模型 $\hat{y}_m$ ，其中，所述回归预测模型的公式为： $\hat{y}_m = w_0 + \sum_{d=1}^D w_d x_d$ ，其中， $\hat{y}_m$ 表示模型预测的第 m 天的目标时间窗同时段的上车客流量； w_0 和 w_d 为所述随机初始化参数； x_d 表示在第 d 个时间窗中的上车客流量；第一构建子模块用于直至所述回归预测模型达到预设的构建条件，获得构建后的回归预测模型的模型参数，其中，所述预设的构建条件为： $\arg \min \frac{1}{2} \sum_{m=1}^n (\hat{y}_m - y_m)^2$ ，其中， y_m 表示第 m 天的目标时间窗同时段的上车客流量。

在本实施例的一些可选的实现方式中，上述目标迭代更新模块 204 进一步用于：根据如下公式计算所述目标公交站的模型梯度信息：

$$\nabla w_d^{i(t)} = \begin{cases} 1 & d = 0 \\ (\hat{y}_m^t - y_m) x_d & d > 0 \end{cases}$$

其中， $\nabla w_d^{i(t)}$ 是目标公交站 i 在 t 次迭代时候的模型参数梯度， d 表示维度，即第 d 个时间窗。

在本实施例的一些可选的实现方式中，上述第二客户端进一步用于：通过如下公式生成第二扰动数据： $P(y'_{ij} = 1) = \frac{1}{e^\varepsilon + 1} + y_{ij} \cdot \frac{e^\varepsilon - 1}{e^\varepsilon + 1}$ ， $P(y'_{ij} = 0) = 1 - P(y'_{ij} = 1)$ ，其中， $y_{ij} \in [0, 1]$ ，1 表示用户在当前联合公交站刷卡上车，0 表示用户未在当前联合公交站刷卡上车， ε 为超参数。

在本实施例的一些可选的实现方式中，上述联合样本构建模块 301 进一步用于基于各自对应的所述第二扰动数据，计算在预设天数中，各自所述目标时间窗同时段之前的预设个数 D 的时间窗对应的上车客流量，作为训练样本的第一输入参数 X ，所述目标时间窗同时段的上车客流量，作为训练样本的第二输入参数 Y ，生成各自的联合公交站的训练样本集合 $(X_1, Y_1) (X_2, Y_2) (X_3, Y_3) \dots (X_m, Y_m)$ ，其中，所述 X_m 具有的特征维数为 D ， D 为整数， m 表示第 m 天。

联合模型构建模块 302 包括第二初始化子模块、第二训练子模块和第二构建子模块。第二初始化子模块用于随机初始化预设的回归预测模型的模型参数，其中，每一天分别对应一个预设的回归预测模型；第二训练子模块用于基于所述训练样本集合中的训练样本分别训练每天对应的回归预测模型 $\hat{y}_m$ ，其中，所述回归预测模型的公式为： $\hat{y}_m = w_0 + \sum_{d=1}^D w_d x_d$ ，其中， $\hat{y}_m$ 表示模型预测的第 m 天的目标时间窗同时段的上车客流量； w_0 和 w_d 为所述随机初始化参数； x_d 表示在第 d 个时间窗中的上车客流量；第二构建子模块用于直至所述回归预测模型达到预设的构建条件，获得各自对应的构建后的回归预测模型的模型参数，其中，所述预设的构建条件为： $\arg \min \frac{1}{2} \sum_{m=1}^n (\hat{y}_m - y_m)^2$ ，其中， y_m 表示第 m 天的目标时间窗同时段的上车客流量。

在本实施例的一些可选的实现方式中，上述参数计算模块 102 进一步用于：通过如下公式计算所述模型更新参数：

$$w_d^{i(t+1)} = \sum_{f \in N(i)} S_{if} \cdot w_d^{f(t)} - \alpha \cdot \nabla w_d^{i(t)}$$

其中， $N(i)$ 为所述联合公交站组成的集合， $\sum_{f \in N(i)} S_{if} \cdot w_d^{f(t)}$ 为接收到的所有随机数之和， α 为学习率， $\nabla w_d^{i(t)}$ 为模型梯度信息。

为解决上述技术问题，本申请实施例还提供计算机设备。具体请参阅图 5，图 5 为本实施例计算机设备基本结构框图。

所述计算机设备 400 包括通过系统总线相互通信连接存储器 401、处理器 402、网络

接口 403。需要指出的是，图中仅示出了具有组件 401-403 的计算机设备 400，但是应理解的是，并不要求实施所有示出的组件，可以替代的实施更多或者更少的组件。其中，本技术领域技术人员可以理解，这里的计算机设备是一种能够按照事先设定或存储的指令，自动进行数值计算和/或信息处理的设备，其硬件包括但不限于微处理器、专用集成电路 (Application Specific Integrated Circuit, ASIC)、可编程门阵列 (Field-Programmable Gate Array, FPGA)、数字处理器 (Digital Signal Processor, DSP)、嵌入式设备等。

所述计算机设备可以是桌上型计算机、笔记本、掌上电脑及云端服务器等计算设备。所述计算机设备可以与用户通过键盘、鼠标、遥控器、触摸板或声控设备等方式进行人机交互。

所述存储器 401 至少包括一种类型的可读存储介质，所述可读存储介质包括闪存、硬盘、多媒体卡、卡型存储器（例如，SD 或 DX 存储器等）、随机访问存储器 (RAM)、静态随机访问存储器 (SRAM)、只读存储器 (ROM)、电可擦除可编程只读存储器 (EEPROM)、可编程只读存储器 (PROM)、磁性存储器、磁盘、光盘等。所述计算机可读存储介质可以是非易失性，也可以是易失性。在一些实施例中，所述存储器 401 可以是所述计算机设备 400 的内部存储单元，例如该计算机设备 400 的硬盘或内存。在另一些实施例中，所述存储器 401 也可以是所述计算机设备 400 的外部存储设备，例如该计算机设备 400 上配备的插接式硬盘，智能存储卡 (Smart Media Card, SMC)，安全数字 (Secure Digital, SD) 卡，闪存卡 (Flash Card) 等。当然，所述存储器 401 还可以既包括所述计算机设备 400 的内部存储单元也包括其外部存储设备。本实施例中，所述存储器 401 通常用于存储安装于所述计算机设备 400 的操作系统和各类应用软件，例如短时公交站客流预测方法的计算机可读指令等。此外，所述存储器 401 还可以用于暂时地存储已经输出或者将要输出的各类数据。

所述处理器 402 在一些实施例中可以是中央处理器 (Central Processing Unit, CPU)、控制器、微控制器、微处理器、或其他数据处理芯片。该处理器 402 通常用于控制所述计算机设备 400 的总体操作。本实施例中，所述处理器 402 用于运行所述存储器 401 中存储的计算机可读指令或者处理数据，例如运行所述短时公交站客流预测方法的计算机可读指令。

所述网络接口 403 可包括无线网络接口或有线网络接口，该网络接口 403 通常用于在所述计算机设备 400 与其他电子设备之间建立通信连接。

在本实施例中，实现了目标公交站和与目标公交站相关性较强的联合公交站之间的联合学习，提升了回归预测模型预测的准确性。

本申请还提供了另一种实施方式，即提供一种计算机可读存储介质，所述计算机可读存储介质存储有计算机可读指令，所述计算机可读指令可被至少一个处理器执行，以使所述至少一个处理器执行如上述的短时公交站客流预测方法的步骤。

在本实施例中，实现了目标公交站和与目标公交站相关性较强的联合公交站之间的联合学习，提升了回归预测模型预测的准确性。

通过以上的实施方式的描述，本领域的技术人员可以清楚地了解到上述实施例方法可借助软件加必需的通用硬件平台的方式来实现，当然也可以通过硬件，但很多情况下前者是更佳的实施方式。基于这样的理解，本申请的技术方案本质上或者说对现有技术做出贡献的部分可以以软件产品的形式体现出来，该计算机软件产品存储在一个存储介质（如 ROM/RAM、磁碟、光盘）中，包括若干指令用以使得一台终端设备（可以是手机，计算机，服务器，空调器，或者网络设备等等）执行本申请各个实施例所述的方法。

权 利 要 求 书

1、一种短时公交站客流预测方法，应用于中心服务器，包括下述步骤：

接收公交站位置信息和目标公交站，基于所述公交站位置信息构建公交站无向网络，并基于所述目标公交站确定所述公交站无向网络中符合预设条件的公交站，作为联合公交站，其中，目标公交站对应目标服务器，每个所述联合公交站各自对应一个联合服务器；

接收目标公交站对应的目标服务器传输的模型梯度信息和联合公交站对应的联合服务器传输的随机数，并基于所述模型梯度信息和所述随机数计算模型更新参数，并将所述模型更新参数传输至目标服务器和各个联合服务器，以使所述目标服务器和联合服务器通过所述模型更新参数迭代更新回归预测模型。

2、根据权利要求1所述的短时公交站客流预测方法，其中，所述基于所述公交站位置信息构建公交站无向网络，并基于所述目标公交站确定所述公交站无向网络中符合预设条件的公交站，作为联合公交站的步骤包括：

基于所述公交站位置信息将所述公交站作为所述公交站无向网络的节点；

将任意两个公交站均进行连线，作为所述公交站无向网络的边，获得所述公交站无向网络；

计算所述目标公交站与各公交站之间的边权 S_{if} ；

将对应的边权小于预设阈值的公交站作为所述联合公交站。

3、根据权利要求1所述的短时公交站客流预测方法，其中，所述基于所述模型梯度信息和所述随机数计算模型更新参数，并将所述模型更新参数传输至目标服务器和各个联合服务器的步骤包括：

所述模型更新参数的计算公式为：

$$w_d^{i(t+1)} = \sum_{f \in N(i)} S_{if} \cdot w_d^{f(t)} - \alpha \cdot \nabla w_d^{i(t)}$$

其中， $N(i)$ 为所述联合公交站组成的集合， $\sum_{f \in N(i)} S_{if} \cdot w_d^{f(t)}$ 为接收到的所有随机数之和， α 为学习率， $\nabla w_d^{i(t)}$ 为模型梯度信息；

将所述模型更新参数传输至目标服务器和各个联合服务器。

4、一种短时公交站客流预测方法，应用于联合服务器，包括下述步骤：

接收目标时间窗和对应的联合公交站的乘客乘车信息，基于所述乘客乘车信息和所述目标时间窗构建对应的联合公交站的训练样本集合；

基于所述训练样本集合构建所述联合公交站的回归预测模型，直至达到预设的构建条件，获得构建后的回归预测模型的模型参数；

接收中心服务器传输的对应的联合公交站的边权，计算所述边权与回归预测模型的模型参数之积，作为计算结果，并基于所述计算结果生成多个随机数，将所述随机数分发至各个所述联合服务器；

所有所述联合服务器均将接收到的随机数传输至中心服务器，以使所述中心服务器利用所述梯度信息和所述随机数计算模型更新参数；

当接收到所述中心服务器传输的模型更新参数时，基于所述模型更新参数迭代回归预测模型，直至达到预设的停止条件，获得训练后的回归预测模型。

5、根据权利要求4所述的短时公交站客流预测方法，其中，所述基于所述乘客乘车信息和所述目标时间窗构建对应的联合公交站的训练样本集合的步骤包括：

基于所述乘客乘车信息，计算各自的联合公交站在预设天数中，所述目标时间窗同时段之前的预设个数D的时间窗对应的上车客流量，作为训练样本的第一输入参数X，所述目标时间窗同时段的上车客流量，作为训练样本的第二输入参数Y，各自生成联合公交站的训练样本集合 (X_1, Y_1) (X_2, Y_2) (X_3, Y_3) (X_m, Y_m) ，其中，所述 X_m 具有的特征维数为D，D为整数，m表示第m天。

6、根据权利要求5所述的短时公交站客流预测方法，其中，所述所述基于所述训练样本集合构建所述联合公交站的回归预测模型，直至达到预设的构建条件，获得构建后的回归预测模型的模型参数的步骤包括：

随机初始化预设的回归预测模型的模型参数，其中，每一天分别对应一个预设的回归预测模型；

基于所述联合公交站的训练样本集合中的训练样本分别训练每天对应的回归预测模型 $\hat{y}_m$ ，其中，所述回归预测模型的公式为：

$$\hat{y}_m = w_0 + \sum_{d=1}^D w_d x_d$$

其中， $\hat{y}_m$ 表示模型预测的第m天的目标时间窗同时段的上车客流量； w_0 和 w_d 为所述随机初始化参数； x_d 表示在第d个时间窗中的上车客流量；

直至所述回归预测模型达到预设的构建条件，获得构建后的回归预测模型的模型参数，其中，所述预设的构建条件为：

$$\arg \min \frac{1}{2} \sum_{m=1}^n (\hat{y}_m - y_m)^2$$

其中， y_m 表示联合公交站第m天的目标时间窗同时段的上车客流量。

7、一种短时公交站客流预测装置，所述短时公交站客流预测装置包括中心服务器，所述中心服务器包括：

网络构建模块，用于接收公交站位置信息和目标公交站，基于所述公交站位置信息构建公交站无向网络，并基于所述目标公交站确定所述公交站无向网络中符合预设条件的公交站，作为联合公交站，其中，目标公交站对应目标服务器，每个所述联合公交站各自对应一个联合服务器；

参数计算模块，用于接收目标公交站对应的目标服务器传输的模型梯度信息，和联合公交站对应的联合服务器传输的随机数，并基于所述模型参数和所述随机数计算模型更新参数，并将所述模型更新参数传输至目标服务器和各个联合服务器，以使所述目标服务器和联合服务器通过所述模型更新参数迭代更新回归预测模型。

8、一种短时公交站客流预测装置，所述短时公交站客流预测装置包括联合服务器，所述联合服务器包括：

联合样本构建模块，用于接收目标时间窗和对应的联合公交站的乘客乘车信息，基于所述乘客乘车信息和所述目标时间窗构建对应的联合公交站的训练样本集合；

联合模型构建模块，用于基于所述训练样本集合构建所述联合公交站的回归预测模型，直至达到预设的构建条件，获得各自的构建后的回归预测模型的模型参数；

分发模块，用于接收中心服务器传输的对应的联合公交站的边权，计算所述边权与回归预测模型的模型参数之积，作为计算结果，并基于所述计算结果生成多个随机数，将所述随机数分发至各个所述联合服务器；

联合数据传输模块，用于将接收到的随机数传输至中心服务器，以使所述中心服务器利用所述梯度信息和所述随机数计算模型更新参数；

联合迭代更新模块，用于当接收到所述中心服务器传输的模型更新参数时，基于所述模型更新参数迭代回归预测模型，直至达到预设的停止条件，获得训练后的回归预测模型。

9、一种计算机设备，包括存储器和处理器，所述存储器中存储有计算机可读指令，所述处理器执行所述计算机可读指令时实现如下所述的短时公交站客流预测方法的步骤：

接收公交站位置信息和目标公交站，基于所述公交站位置信息构建公交站无向网络，并基于所述目标公交站确定所述公交站无向网络中符合预设条件的公交站，作为联合公交站，其中，目标公交站对应目标服务器，每个所述联合公交站各自对应一个联合服务器；

接收目标公交站对应的目标服务器传输的模型梯度信息和联合公交站对应的联合服

务器传输的随机数，并基于所述模型梯度信息和所述随机数计算模型更新参数，并将所述模型更新参数传输至目标服务器和各个联合服务器，以使所述目标服务器和联合服务器通过所述模型更新参数迭代更新回归预测模型。

10、根据权利要求9所述的计算机设备，其中，所述基于所述公交站位置信息构建公交站无向网络，并基于所述目标公交站确定所述公交站无向网络中符合预设条件的公交站，作为联合公交站的步骤包括：

基于所述公交站位置信息将所述公交站作为所述公交站无向网络的节点；

将任意两个公交站均进行连线，作为所述公交站无向网络的边，获得所述公交站无向网络；

计算所述目标公交站与各公交站之间的边权 S_{if} ；

将对应的边权小于预设阈值的公交站作为所述联合公交站。

11、根据权利要求9所述的计算机设备，其中，所述基于所述模型梯度信息和所述随机数计算模型更新参数，并将所述模型更新参数传输至目标服务器和各个联合服务器的步骤包括：

所述模型更新参数的计算公式为：

$$w_d^{i(t+1)} = \sum_{f \in N(i)} S_{if} \cdot w_d^{f(t)} - \alpha \cdot \nabla w_d^{i(t)}$$

其中， $N(i)$ 为所述联合公交站组成的集合， $\sum_{f \in N(i)} S_{if} \cdot w_d^{f(t)}$ 为接收到的所有随机数之和， α 为学习率， $\nabla w_d^{i(t)}$ 为模型梯度信息；

将所述模型更新参数传输至目标服务器和各个联合服务器。

12、一种计算机设备，包括存储器和处理器，所述存储器中存储有计算机可读指令，所述处理器执行所述计算机可读指令时实现如下所述的短时公交站客流预测方法的步骤：

接收目标时间窗和对应的联合公交站的乘客乘车信息，基于所述乘客乘车信息和所述目标时间窗构建对应的联合公交站的训练样本集合；

基于所述训练样本集合构建所述联合公交站的回归预测模型，直至达到预设的构建条件，获得构建后的回归预测模型的模型参数；

接收中心服务器传输的对应的联合公交站的边权，计算所述边权与回归预测模型的模型参数之积，作为计算结果，并基于所述计算结果生成多个随机数，将所述随机数分发至各个所述联合服务器；

所有所述联合服务器均将接收到的随机数传输至中心服务器，以使所述中心服务器利用所述梯度信息和所述随机数计算模型更新参数；

当接收到所述中心服务器传输的模型更新参数时，基于所述模型更新参数迭代回归预测模型，直至达到预设的停止条件，获得训练后的回归预测模型。

13、根据权利要求12所述的计算机设备，其中，所述基于所述乘客乘车信息和所述目标时间窗构建对应的联合公交站的训练样本集合的步骤包括：

基于所述乘客乘车信息，计算各自的联合公交站在预设天数中，所述目标时间窗同时段之前的预设个数 D 的时间窗对应的上车客流量，作为训练样本的第一输入参数 X ，所述目标时间窗同时段的上车客流量，作为训练样本的第二输入参数 Y ，各自生成联合公交站的训练样本集合 $(X_1, Y_1) (X_2, Y_2) (X_3, Y_3) \dots (X_m, Y_m)$ ，其中，所述 X_m 具有的特征维数为 D ， D 为整数， m 表示第 m 天。

14、根据权利要求13所述的计算机设备，其中，所述基于所述训练样本集合构建所述联合公交站的回归预测模型，直至达到预设的构建条件，获得构建后的回归预测模型的模型参数的步骤包括：

随机初始化预设的回归预测模型的模型参数，其中，每一天分别对应一个预设的回归预测模型；

基于所述联合公交站的训练样本集合中的训练样本分别训练每天对应的回归预测模型 $\hat{y}_m$ ，其中，所述回归预测模型的公式为：

$$\hat{y}_m = w_0 + \sum_{d=1}^D w_d x_d$$

其中， $\hat{y}_m$ 表示模型预测的第 m 天的目标时间窗同时段的上车客流量； w_0 和 w_d 为所述随机初始化参数； x_d 表示在第 d 个时间窗中的上车客流量；

直至所述回归预测模型达到预设的构建条件，获得构建后的回归预测模型的模型参数，其中，所述预设的构建条件为：

$$\arg \min \frac{1}{2} \sum_{m=1}^n (\hat{y}_m - y_m)^2$$

其中， y_m 表示联合公交站第 m 天的目标时间窗同时段的上车客流量。

15、一种计算机可读存储介质，所述计算机可读存储介质上存储有计算机可读指令，所述计算机可读指令被处理器执行时实现如下所述的短时公交站客流预测方法的步骤：

接收公交站位置信息和目标公交站，基于所述公交站位置信息构建公交站无向网络，并基于所述目标公交站确定所述公交站无向网络中符合预设条件的公交站，作为联合公交站，其中，目标公交站对应目标服务器，每个所述联合公交站各自对应一个联合服务器；

接收目标公交站对应的目标服务器传输的模型梯度信息和联合公交站对应的联合服务器传输的随机数，并基于所述模型梯度信息和所述随机数计算模型更新参数，并将所述模型更新参数传输至目标服务器和各个联合服务器，以使所述目标服务器和联合服务器通过所述模型更新参数迭代更新回归预测模型。

16、根据权利要求15所述的计算机可读存储介质，其中，所述基于所述公交站位置信息构建公交站无向网络，并基于所述目标公交站确定所述公交站无向网络中符合预设条件的公交站，作为联合公交站的步骤包括：

基于所述公交站位置信息将所述公交站作为所述公交站无向网络的节点；

将任意两个公交站均进行连线，作为所述公交站无向网络的边，获得所述公交站无向网络；

计算所述目标公交站与各公交站之间的边权 S_{if} ；

将对应的边权小于预设阈值的公交站作为所述联合公交站。

17、根据权利要求15所述的计算机可读存储介质，其中，所述基于所述模型梯度信息和所述随机数计算模型更新参数，并将所述模型更新参数传输至目标服务器和各个联合服务器的步骤包括：

所述模型更新参数的计算公式为：

$$w_d^{i(t+1)} = \sum_{f \in N(i)} S_{if} \cdot w_d^{f(t)} - \alpha \cdot \nabla w_d^{i(t)}$$

其中， $N(i)$ 为所述联合公交站组成的集合， $\sum_{f \in N(i)} S_{if} \cdot w_d^{f(t)}$ 为接收到的所有随机数之和， α 为学习率， $\nabla w_d^{i(t)}$ 为模型梯度信息；

将所述模型更新参数传输至目标服务器和各个联合服务器。

18、一种计算机可读存储介质，所述计算机可读存储介质上存储有计算机可读指令，所述计算机可读指令被处理器执行时实现如下所述的短时公交站客流预测方法的步骤：

接收目标时间窗和对应的联合公交站的乘客乘车信息，基于所述乘客乘车信息和所述目标时间窗构建对应的联合公交站的训练样本集合；

基于所述训练样本集合构建所述联合公交站的回归预测模型，直至达到预设的构建条件，获得构建后的回归预测模型的模型参数；

接收中心服务器传输的对应的联合公交站的边权，计算所述边权与回归预测模型的模

型参数之积，作为计算结果，并基于所述计算结果生成多个随机数，将所述随机数分发至各个所述联合服务器；

所有所述联合服务器均将接收到的随机数传输至中心服务器，以使所述中心服务器利用所述梯度信息和所述随机数计算模型更新参数；

- 5 当接收到所述中心服务器传输的模型更新参数时，基于所述模型更新参数迭代回归预测模型，直至达到预设的停止条件，获得训练后的回归预测模型。

19、根据权利要求 18 所述的计算机可读存储介质，其中，所述基于所述乘客乘车信息和所述目标时间窗构建对应的联合公交站的训练样本集合的步骤包括：

- 10 基于所述乘客乘车信息，计算各自的联合公交站在预设天数中，所述目标时间窗同时段之前的预设个数 D 的时间窗对应的上车客流量，作为训练样本的第一输入参数 X，所述目标时间窗同时段的上车客流量，作为训练样本的第二输入参数 Y，各自生成联合公交站的训练样本集合 $(X_1, Y_1) (X_2, Y_2) (X_3, Y_3) \dots (X_m, Y_m)$ ，其中，所述 X_m 具有的特征维数为 D，D 为整数，m 表示第 m 天。

- 15 20、根据权利要求 19 所述的计算机可读存储介质，其中，所述基于所述训练样本集合构建所述联合公交站的回归预测模型，直至达到预设的构建条件，获得构建后的回归预测模型的模型参数的步骤包括：

随机初始化预设的回归预测模型的模型参数，其中，每一天分别对应一个预设的回归预测模型；

- 20 基于所述联合公交站的训练样本集合中的训练样本分别训练每天对应的回归预测模型 $\hat{y}_m$ ，其中，所述回归预测模型的公式为：

$$\hat{y}_m = w_0 + \sum_{d=1}^D w_d x_d$$

其中， $\hat{y}_m$ 表示模型预测的第 m 天的目标时间窗同时段的上车客流量； w_0 和 w_d 为所述随机初始化参数； x_d 表示在第 d 个时间窗中的上车客流量；

直至所述回归预测模型达到预设的构建条件，获得构建后的回归预测模型的模型参数，其中，所述预设的构建条件为：

$$\arg \min \frac{1}{2} \sum_{m=1}^n (\hat{y}_m - y_m)^2$$

- 25 其中， y_m 表示联合公交站第 m 天的目标时间窗同时段的上车客流量。

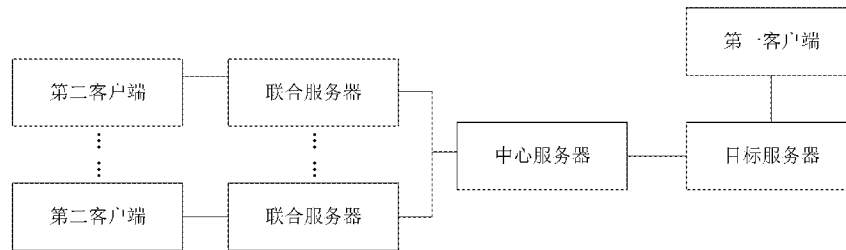


图 1

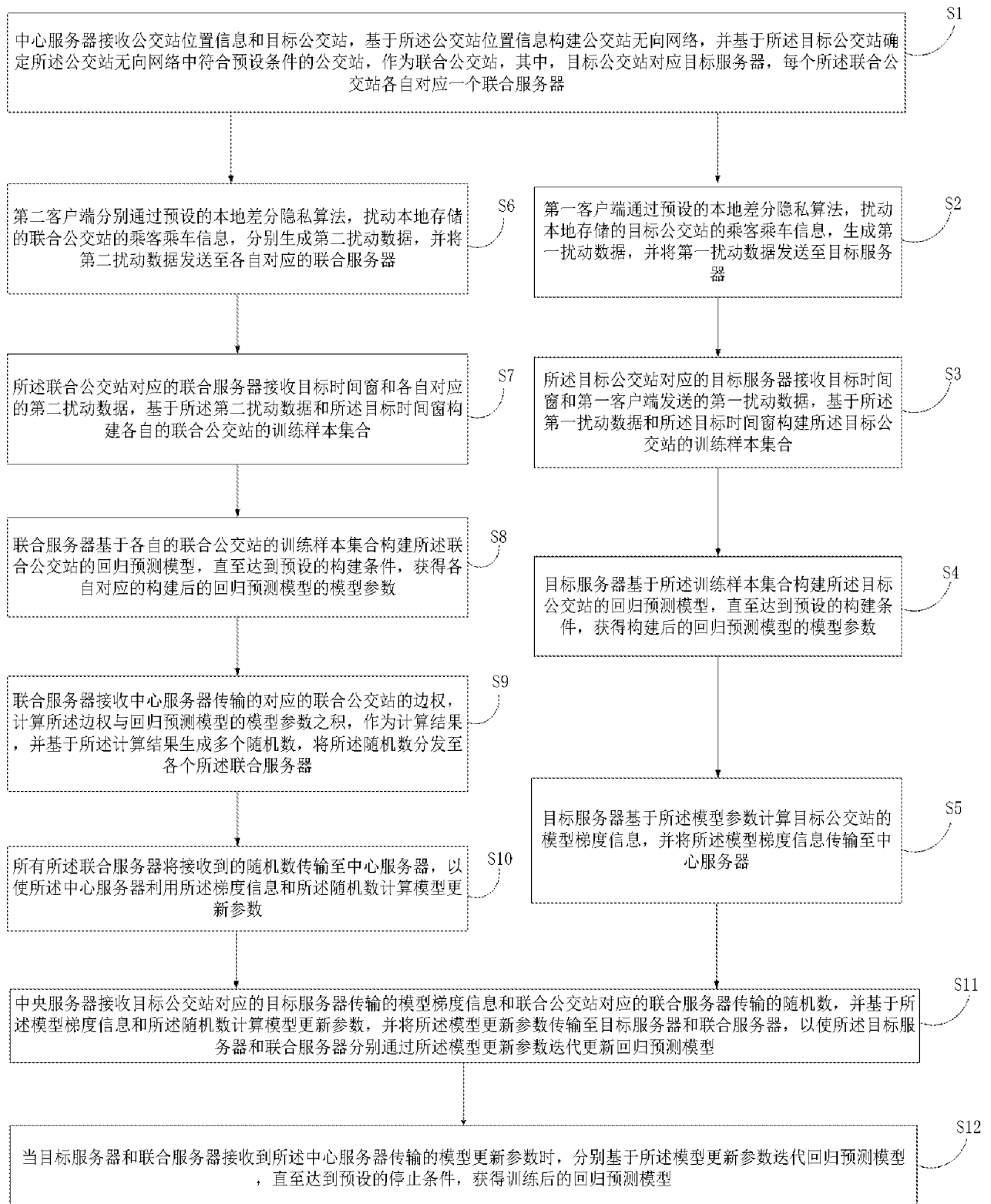


图 2

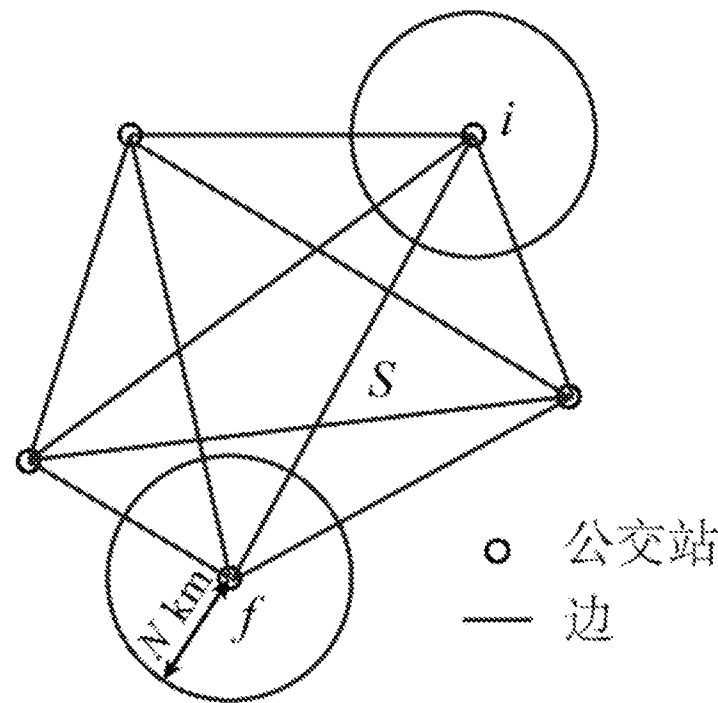


图 3

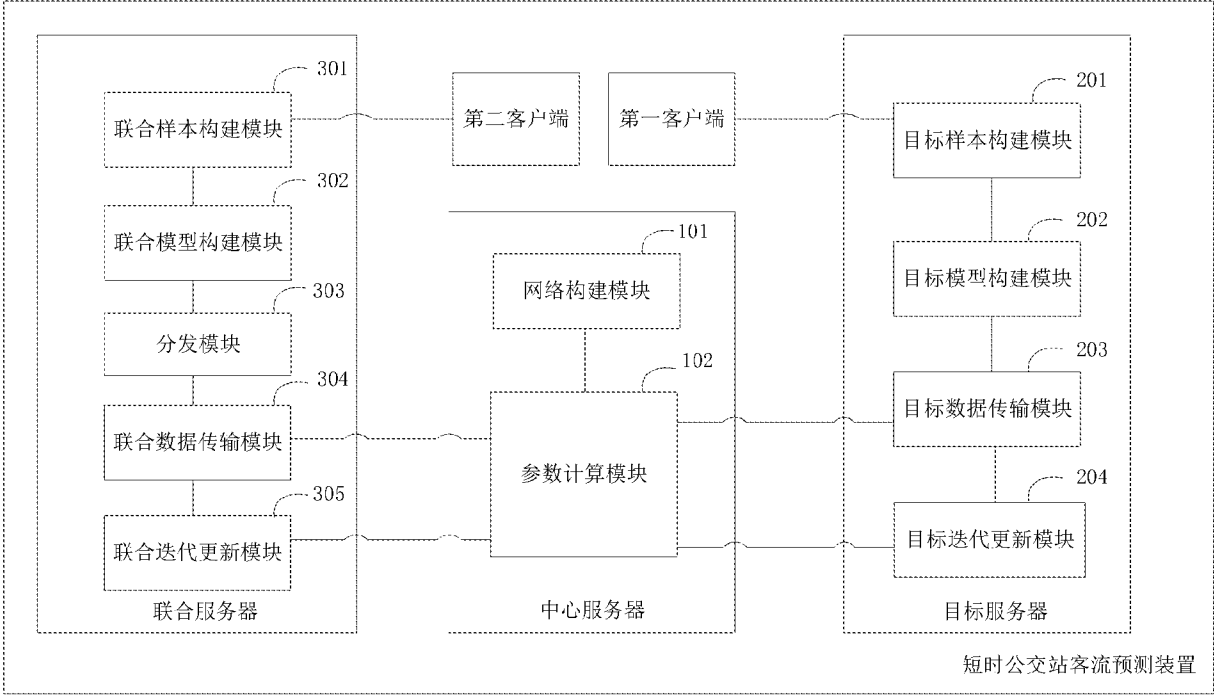


图 4

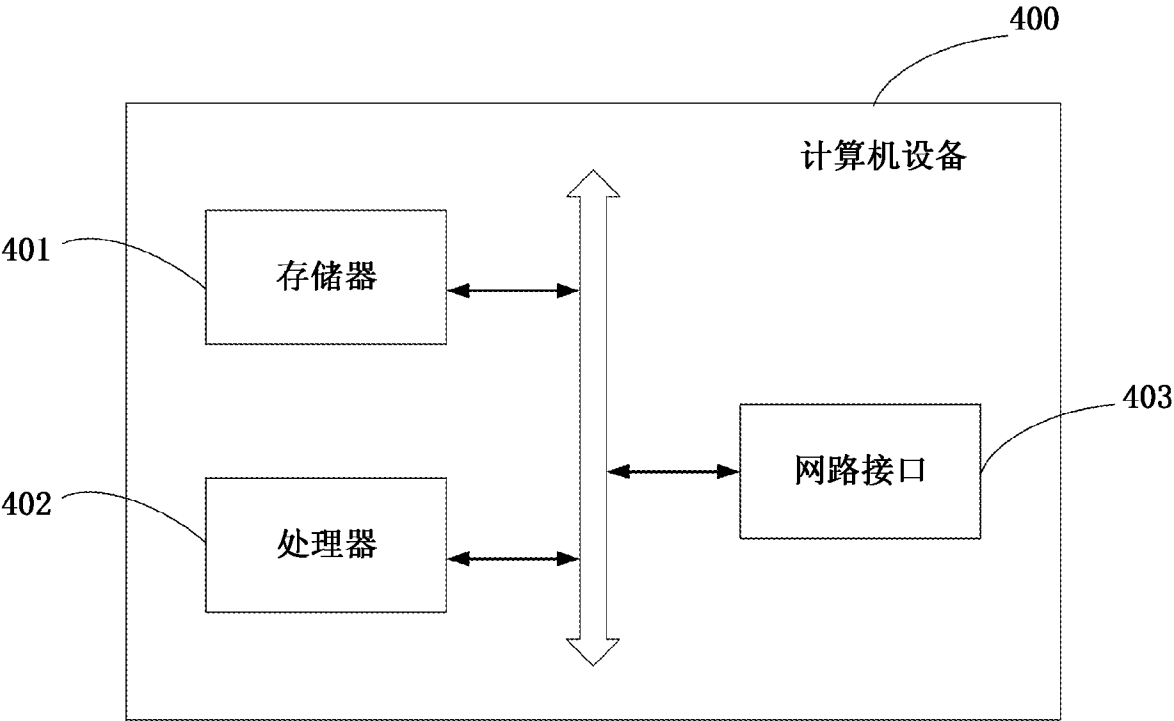


图 5

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2020/134955

A. CLASSIFICATION OF SUBJECT MATTER

G06Q 10/04(2012.01)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)

G06Q

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)

CNPAT, CNKI, WPI, EPODOC, IEEE: 公交, 站点, 关联, 联合, 客流, 预测, 预估, 回归, 梯度, 随机数, 模型, 参数, 更新, 迭代, bus, station, relat+, associate+, passenger, flow+, predict+, forecast+, regressive, grads, random w number, model, parameter, update, iterative

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 107563540 A (CENTRAL SOUTH UNIVERSITY) 09 January 2018 (2018-01-09) description paragraphs 8-16,27-44	1-20
A	CN 105095993 A (JINAN MUNICIPAL ENGINEERING DESIGN & RESEARCH INSTITUTE (GROUP) CO., LTD.) 25 November 2015 (2015-11-25) entire document	1-20
A	CN 111754025 A (WENZHENG COLLEGE OF SOOCHOW UNIVERSITY) 09 October 2020 (2020-10-09) entire document	1-20
A	CN 102436603 A (BEIHANG UNIVERSITY) 02 May 2012 (2012-05-02) entire document	1-20
A	CN 102169606 A (TONGJI UNIVERSITY) 31 August 2011 (2011-08-31) entire document	1-20
A	US 2015356458 A1 (BERENGUERES, Jose Oriol Lopez et al.) 10 December 2015 (2015-12-10) entire document	1-20



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“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)

“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

“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

“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

“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

“&” document member of the same patent family

Date of the actual completion of the international search

15 July 2021

Date of mailing of the international search report

26 July 2021

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/
CN)
No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing
100088
China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2020/134955

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	107563540	A	09 January 2018	None			
CN	105095993	A	25 November 2015	None			
CN	111754025	A	09 October 2020	None			
CN	102436603	A	02 May 2012	None			
CN	102169606	A	31 August 2011	None			
US	2015356458	A1	10 December 2015	WO	2015189768	A1	17 December 2015

国际检索报告

国际申请号

PCT/CN2020/134955

A. 主题的分类 G06Q 10/04 (2012.01) 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) G06Q 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)) CNPAT, CNKI, WPI, EP000C, IEEE: 公交, 站点, 关联, 联合, 客流, 预测, 预估, 回归, 梯度, 随机数, 模型, 参数, 更新, 迭代, bus, station, relat+, associate+, passenger, flow+, predict+, forecast+, regressive, grads, random w number, model, parameter, update, iterative																							
C. 相关文件 <table border="1"> <thead> <tr> <th>类 型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>A</td> <td>CN 107563540 A (中南大学) 2018年 1月 9日 (2018 - 01 - 09) 说明书第8-16, 27-44段</td> <td>1-20</td> </tr> <tr> <td>A</td> <td>CN 105095993 A (济南市市政工程设计研究院集团有限责任公司) 2015年 11月 25日 (2015 - 11 - 25) 全文</td> <td>1-20</td> </tr> <tr> <td>A</td> <td>CN 111754025 A (苏州大学文正学院) 2020年 10月 9日 (2020 - 10 - 09) 全文</td> <td>1-20</td> </tr> <tr> <td>A</td> <td>CN 102436603 A (北京航空航天大学) 2012年 5月 2日 (2012 - 05 - 02) 全文</td> <td>1-20</td> </tr> <tr> <td>A</td> <td>CN 102169606 A (同济大学) 2011年 8月 31日 (2011 - 08 - 31) 全文</td> <td>1-20</td> </tr> <tr> <td>A</td> <td>US 2015356458 A1 (BERENGUERES, Jose Oriol Lopez等) 2015年 12月 10日 (2015 - 12 - 10) 全文</td> <td>1-20</td> </tr> </tbody> </table>			类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求	A	CN 107563540 A (中南大学) 2018年 1月 9日 (2018 - 01 - 09) 说明书第8-16, 27-44段	1-20	A	CN 105095993 A (济南市市政工程设计研究院集团有限责任公司) 2015年 11月 25日 (2015 - 11 - 25) 全文	1-20	A	CN 111754025 A (苏州大学文正学院) 2020年 10月 9日 (2020 - 10 - 09) 全文	1-20	A	CN 102436603 A (北京航空航天大学) 2012年 5月 2日 (2012 - 05 - 02) 全文	1-20	A	CN 102169606 A (同济大学) 2011年 8月 31日 (2011 - 08 - 31) 全文	1-20	A	US 2015356458 A1 (BERENGUERES, Jose Oriol Lopez等) 2015年 12月 10日 (2015 - 12 - 10) 全文	1-20
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求																					
A	CN 107563540 A (中南大学) 2018年 1月 9日 (2018 - 01 - 09) 说明书第8-16, 27-44段	1-20																					
A	CN 105095993 A (济南市市政工程设计研究院集团有限责任公司) 2015年 11月 25日 (2015 - 11 - 25) 全文	1-20																					
A	CN 111754025 A (苏州大学文正学院) 2020年 10月 9日 (2020 - 10 - 09) 全文	1-20																					
A	CN 102436603 A (北京航空航天大学) 2012年 5月 2日 (2012 - 05 - 02) 全文	1-20																					
A	CN 102169606 A (同济大学) 2011年 8月 31日 (2011 - 08 - 31) 全文	1-20																					
A	US 2015356458 A1 (BERENGUERES, Jose Oriol Lopez等) 2015年 12月 10日 (2015 - 12 - 10) 全文	1-20																					
☐ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。																							
<table border="0"> <tr> <td style="vertical-align: top;"> * 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 </td> <td style="vertical-align: top;"> “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 </td> </tr> </table>			* 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																			
* 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 2021年 7月 15日		国际检索报告邮寄日期 2021年 7月 26日																					
ISA/CN的名称和邮寄地址 中国国家知识产权局 (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		受权官员 曹妹妹 电话号码 86-(10)-53961315																					

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2020/134955

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	107563540	A	2018年 1月 9日	无		
CN	105095993	A	2015年 11月 25日	无		
CN	111754025	A	2020年 10月 9日	无		
CN	102436603	A	2012年 5月 2日	无		
CN	102169606	A	2011年 8月 31日	无		
US	2015356458	A1	2015年 12月 10日	WO	2015189768 A1	2015年 12月 17日