



(12)发明专利

(10)授权公告号 CN 108288096 B

(45)授权公告日 2020.08.21

(21)申请号 201710016048.5

审查员 梁韬

(22)申请日 2017.01.10

(65)同一申请的已公布的文献号

申请公布号 CN 108288096 A

(43)申请公布日 2018.07.17

(73)专利权人 北京嘀嘀无限科技发展有限公司

地址 100193 北京市海淀区东北旺西路8号  
院34号楼

(72)发明人 王征 傅昆

(74)专利代理机构 北京友联知识产权代理事务  
所(普通合伙) 11343

代理人 尚志峰 汪海屏

(51)Int.Cl.

G06N 20/00(2019.01)

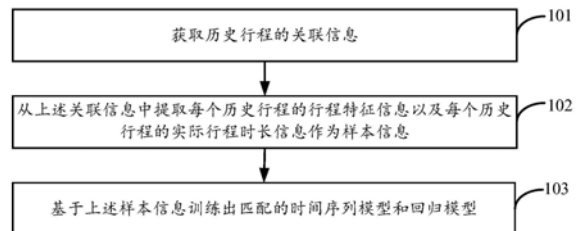
权利要求书6页 说明书12页 附图3页

(54)发明名称

用于估算行程时间、模型训练的方法及装置

(57)摘要

本申请实施例公开了一种用于估算行程时间、模型训练的方法及装置,涉及机器学习技术领域,所述方法的一具体实施方式包括:获取历史行程的关联信息;从所述关联信息中提取每个历史行程的行程特征信息以及所述每个历史行程的实际行程时长信息作为样本信息,所述行程特征信息包括所述每个历史行程全局特征信息以及所述每个历史行程包含的各个路段的路段特征信息;基于所述样本信息训练出匹配的时间序列模型和回归模型。该实施方式训练出了可以用于估算行程时间的模型,由于将历史行程中路段的特征信息与全局特征信息相结合对模型进行训练,将更多的行进速度影响因素考虑在内,因此,提高了训练出的模型估算行程时间的准确度。



1. 一种用于估算行程时间的模型训练方法,其特征在于,所述方法包括:

获取历史行程的关联信息;

从所述关联信息中提取每个历史行程的行程特征信息以及所述每个历史行程的实际行程时长信息作为样本信息,所述行程特征信息包括所述每个历史行程的全局特征信息以及所述每个历史行程包含的各个路段的路段特征信息;

基于所述样本信息训练出匹配的时间序列模型和回归模型;

所述基于所述样本信息训练出匹配的时间序列模型和回归模型,包括:

执行修正的步骤,所述修正的步骤包括:从所述样本信息中提取训练数据,并采用所述训练数据对当前的时间序列模型和回归模型中的参数进行调整;

如果修正之后的时间序列模型和回归模型未满足匹配条件,则继续执行所述修正的步骤;

如果修正之后的时间序列模型和回归模型满足匹配条件,输出所述修正之后的时间序列模型和回归模型作为匹配的时间序列模型和回归模型;

所述采用所述训练数据对当前的时间序列模型和回归模型中的参数进行调整,包括:

将所述训练数据中的路段特征信息输入到当前的时间序列模型中;

将所述当前的时间序列模型输出的隐状态信息以及所述训练数据中的全局特征信息输入到当前的回归模型中;

根据所述当前的回归模型输出的预估时间信息与所述训练数据中实际行程时长信息的差异,调整当前的时间序列模型和回归模型中的参数。

2. 根据权利要求1所述的方法,其特征在于,所述方法还包括:

将行程路段的数量划分为多个范围,根据行程路段的数量对所述历史行程进行分组,使同一组的历史行程的路段数量划分在同一个范围内;

所述从所述样本信息中提取训练数据,包括:

从所述样本信息中提取同一组的多个历史行程的行程特征信息以及实际行程时长信息作为训练数据。

3. 根据权利要求1或2所述的方法,其特征在于,所述时间序列模型模型包括循环神经网络RNN;所述回归模型包括多层感知器MLP。

4. 根据权利要求1所述的方法,其特征在于,所述历史行程的关联信息包括以下任意一项或多项:

所述历史行程的路径信息;

所述历史行程对应的起始时刻信息;

所述历史行程对应的终止时刻信息

所述历史行程对应的日期信息;

所述历史行程对应的路况信息;

完成所述历史行程的实际行程时长;

所述历史行程中的路段的信息;

所述历史行程对应的人员信息;

所述历史行程对应的出行方式的信息。

5. 根据权利要求1所述的方法,其特征在于,所述路段特征信息包括以下任意一项或多

项：

所述路段的标识特征向量；  
所述路段对应的行程的起始时刻特征向量；  
所述路段对应的日期特征向量；  
所述路段对应的路况特征向量；  
所述路段对应的人员特征向量；  
所述路段对应的出行方式的特征向量。

6. 根据权利要求1所述的方法，其特征在于，所述全局特征信息包括以下任意一项或多项：

所述历史行程的路径特征向量；  
所述历史行程对应的起始时刻特征向量；  
所述历史行程对应的日期特征向量；  
所述历史行程对应的路况特征向量；  
所述历史行程对应的人员特征向量；  
所述历史行程对应的出行方式的特征向量。

7. 一种估算行程时间的方法，其特征在于，所述方法包括：

获取目标行程的关联信息；

从所述关联信息中提取所述目标行程的行程特征信息，所述行程特征信息包括行程全程的全局特征信息以及行程中包含的各个路段的路段特征信息；

基于所述行程特征信息采用预先训练的时间序列模型和回归模型估算出所述目标行程的行程时间；

所述基于所述行程特征信息采用预先训练的时间序列模型和回归模型估算出所述目标行程的行程时间，包括：

将所述行程特征信息中的路段特征信息输入到预先训练的时间序列模型；

将所述时间序列模型输出的隐状态信息以及所述行程特征信息中的全局特征信息输入到预先训练的回归模型；

获取所述回归模型输出的结果作为估算出的所述目标行程的行程时间。

8. 根据权利要求7所述的方法，其特征在于，所述时间序列模型包括循环神经网络RNN；所述回归模型包括多层感知器MLP。

9. 根据权利要求7所述的方法，其特征在于，所述目标行程的关联信息包括以下任意一项或多项：

所述目标行程的路径信息；  
所述目标行程对应的起始时刻信息；  
所述目标行程对应的终止时刻信息  
所述目标行程对应的日期信息；  
所述目标行程对应的路况信息；  
所述目标行程中的路段的信息；  
所述目标行程对应的人员信息；  
所述目标行程对应的出行方式信息。

10. 根据权利要求7所述的方法, 其特征在于, 所述路段特征信息包括以下任意一项或多项:

所述路段的标识特征向量;  
所述路段对应的目标行程的起始时刻特征向量;  
所述路段对应的日期特征向量;  
所述路段对应的路况特征向量;  
所述路段对应的人员特征向量;  
所述路段对应的出行方式信息。

11. 根据权利要求7所述的方法, 其特征在于, 所述全局特征信息包括以下任意一项或多项:

所述目标行程的路径特征向量;  
所述目标行程对应的起始时刻特征向量;  
所述目标行程对应的日期特征向量;  
所述目标行程对应的路况特征向量;  
所述目标行程对应的人员特征向量;  
所述目标行程对应的出行方式特征向量。

12. 一种用于估算行程时间的模型训练装置, 其特征在于, 所述装置包括:

获取单元, 被配置为获取历史行程的关联信息;

提取单元, 被配置为从所述关联信息中提取每个历史行程的行程特征信息以及所述每个历史行程的实际行程时长信息作为样本信息, 所述行程特征信息包括所述每个历史行程的全局特征信息以及所述每个历史行程包含的各个路段的路段特征信息;

训练单元, 被配置为基于所述样本信息训练出匹配的时间序列模型和回归模型;

所述训练单元包括:

修正子单元, 被配置为执行修正的步骤, 所述修正的步骤包括: 从所述样本信息中提取训练数据, 并采用所述训练数据对当前的时间序列模型和回归模型中的参数进行调整;

控制子单元, 被配置为在修正之后的时间序列模型和回归模型未满足匹配条件时, 控制所述修正子单元继续执行所述修正的步骤;

输出子单元, 被配置为在修正之后的时间序列模型和回归模型满足匹配条件时, 输出所述修正之后的时间序列模型和回归模型作为匹配的时间序列模型和回归模型;

所述修正子单元通过如下方式采用所述训练数据对当前的时间序列模型和回归模型中的参数进行调整:

将所述训练数据中的路段特征信息输入到当前的时间序列模型中;

将所述当前的时间序列模型输出的隐状态信息以及所述训练数据中的全局特征信息输入到当前的回归模型中;

根据所述当前的回归模型输出的预估时间信息与所述训练数据中实际行程时长信息的差异, 调整当前的时间序列模型和回归模型中的参数。

13. 根据权利要求12所述的装置, 其特征在于, 所述装置还包括:

分组单元, 被配置为将行程路段的数量划分为多个范围, 根据行程路段的数量对所述历史行程进行分组, 使同一组的历史行程的路段数量划分在同一个范围内;

所述修正子单元通过如下方式从所述样本信息中提取训练数据：

从所述样本信息中提取同一组的多个历史行程的行程特征信息以及实际行程时长信息作为训练数据。

14. 根据权利要求12或13所述的装置，其特征在于，所述时间序列模型模型包括循环神经网络RNN；所述回归模型包括多层感知器MLP。

15. 根据权利要求12所述的装置，其特征在于，所述历史行程的关联信息包括以下任意一项或多项：

所述历史行程的路径信息；

所述历史行程对应的起始时刻信息；

所述历史行程对应的终止时刻信息；

所述历史行程对应的日期信息；

所述历史行程对应的路况信息；

完成所述历史行程的实际行程时长；

所述历史行程中的路段的信息；

所述历史行程对应的人员信息；

所述历史行程对应的出行方式的信息。

16. 根据权利要求12所述的装置，其特征在于，所述路段特征信息包括以下任意一项或多项：

所述路段的标识特征向量；

所述路段对应的行程的起始时刻特征向量；

所述路段对应的日期特征向量；

所述路段对应的路况特征向量；

所述路段对应的人员特征向量；

所述路段对应的出行方式的特征向量。

17. 根据权利要求12所述的装置，其特征在于，所述全局特征信息包括以下任意一项或多项：

所述历史行程的路径特征向量；

所述历史行程对应的起始时刻特征向量；

所述历史行程对应的日期特征向量；

所述历史行程对应的路况特征向量；

所述历史行程对应的人员特征向量；

所述历史行程对应的出行方式的特征向量。

18. 一种估算行程时间的装置，其特征在于，所述装置包括：

获取单元，被配置为获取目标行程的关联信息；

提取单元，被配置为从所述关联信息中提取所述目标行程的行程特征信息，所述行程特征信息包括行程全程的全局特征信息以及行程中包含的各个路段的路段特征信息；

估算单元，被配置为基于所述行程特征信息采用预先训练的时间序列模型和回归模型估算出所述目标行程的行程时间；

所述估算单元包括：

第一输入子单元,被配置为将所述行程特征信息中的路段特征信息输入到预先训练的时间序列模型;

第二输入子单元,被配置为将所述时间序列模型输出的隐状态信息以及所述行程特征信息中的全局特征信息输入到预先训练的回归模型;

输出子单元,被配置为获取所述回归模型输出的结果作为估算出的所述目标行程的行程时间。

19. 根据权利要求18所述的装置,其特征在于,所述时间序列模型包括循环神经网络RNN;所述回归模型包括多层感知器MLP。

20. 根据权利要求18所述的装置,其特征在于,所述目标行程的关联信息包括以下任意一项或多项:

所述目标行程的路径信息;

所述目标行程对应的起始时刻信息;

所述目标行程对应的终止时刻信息;

所述目标行程对应的日期信息;

所述目标行程对应的路况信息;

所述目标行程中的路段的信息;

所述目标行程对应的人员信息;

所述目标行程对应的出行方式信息。

21. 根据权利要求18所述的装置,其特征在于,所述路段特征信息包括以下任意一项或多项:

所述路段的标识特征向量;

所述路段对应的目标行程的起始时刻特征向量;

所述路段对应的日期特征向量;

所述路段对应的路况特征向量;

所述路段对应的人员特征向量;

所述路段对应的出行方式信息。

22. 根据权利要求18所述的装置,其特征在于,所述全局特征信息包括以下任意一项或多项:

所述目标行程的路径特征向量;

所述目标行程对应的起始时刻特征向量;

所述目标行程对应的日期特征向量;

所述目标行程对应的路况特征向量;

所述目标行程对应的人员特征向量;

所述目标行程对应的出行方式特征向量。

23. 一种计算机存储介质,所述存储介质中存储有程序指令,其特征在于,所述指令包括:

获取历史行程的关联信息;

从所述关联信息中提取每个历史行程的行程特征信息以及所述每个历史行程的实际行程时长信息作为样本信息,所述行程特征信息包括所述每个历史行程的全局特征信息以

及所述每个历史行程包含的各个路段的路段特征信息；

基于所述样本信息训练出匹配的时间序列模型和回归模型；

所述基于所述样本信息训练出匹配的时间序列模型和回归模型，包括：

执行修正的步骤，所述修正的步骤包括：从所述样本信息中提取训练数据，并采用所述训练数据对当前的时间序列模型和回归模型中的参数进行调整；

如果修正之后的时间序列模型和回归模型未满足匹配条件，则继续执行所述修正的步骤；

如果修正之后的时间序列模型和回归模型满足匹配条件，输出所述修正之后的时间序列模型和回归模型作为匹配的时间序列模型和回归模型；

所述采用所述训练数据对当前的时间序列模型和回归模型中的参数进行调整，包括：

将所述训练数据中的路段特征信息输入到当前的时间序列模型中；

将所述当前的时间序列模型输出的隐状态信息以及所述训练数据中的全局特征信息输入到当前的回归模型中；

根据所述当前的回归模型输出的预估时间信息与所述训练数据中实际行程时长信息的差异，调整当前的时间序列模型和回归模型中的参数。

24. 一种计算机存储介质，所述存储介质中存储有程序指令，其特征在于，所述指令包括：

获取目标行程的关联信息；

从所述关联信息中提取所述目标行程的行程特征信息，所述行程特征信息包括行程全程的全局特征信息以及行程中包含的各个路段的路段特征信息；

基于所述行程特征信息采用预先训练的时间序列模型和回归模型估算出所述目标行程的行程时间；

所述基于所述行程特征信息采用预先训练的时间序列模型和回归模型估算出所述目标行程的行程时间，包括：

将所述行程特征信息中的路段特征信息输入到预先训练的时间序列模型；

将所述时间序列模型输出的隐状态信息以及所述行程特征信息中的全局特征信息输入到预先训练的回归模型；

获取所述回归模型输出的结果作为估算出的所述目标行程的行程时间。

## 用于估算行程时间、模型训练的方法及装置

### 技术领域

[0001] 本申请实施例涉及机器学习技术领域,特别涉及一种用于估算行程时间、模型训练的方法及装置。

### 背景技术

[0002] ETA (Estimated Time of Arrival, 到达时间预估) 技术是根据行程的路径预测行程时间的技术,在LBS (Location Based Service, 基于位置服务) 应用领域中具有非常重要的意义。在此基础上,可以为用户提供更多的服务。目前,通常是根据行程当前的路况信息,直接预测行程时间。例如,可以获取当前行程中的平均行进速度,然后,根据行程的路径计算行进时间,再加上行程中路口的平均耗时,计算出行程的预估时间。但是,由于采用这种方式计算行程的预估时间时,很多影响行进速度的因素并没有考虑在内,因此预测的行程时间的准确度低。

### 发明内容

[0003] 为了解决上述技术问题,本申请实施例提供了一种用于估算行程时间、模型训练的方法及装置。

[0004] 根据本申请实施例的第一方面,提供一种用于估算行程时间的模型训练方法,包括:

[0005] 获取历史行程的关联信息;

[0006] 从所述关联信息中提取每个历史行程的行程特征信息以及所述每个历史行程的实际行程时长信息作为样本信息,所述行程特征信息包括所述每个历史行程的全局特征信息以及所述每个历史行程包含的各个路段的路段特征信息;

[0007] 基于所述样本信息训练出匹配的时间序列模型和回归模型。

[0008] 根据本申请实施例的第二方面,提供一种估算行程时间的方法,包括:

[0009] 获取目标行程的关联信息;

[0010] 从所述关联信息中提取所述目标行程的行程特征信息,所述行程特征信息包括行程全程的全局特征信息以及行程中包含的各个路段的路段特征信息;

[0011] 基于所述行程特征信息采用预先训练的时间序列模型和回归模型估算出所述目标行程的行程时间。

[0012] 根据本申请实施例的第三方面,提供一种用于估算行程时间的模型训练装置,包括:

[0013] 获取单元,被配置为获取历史行程的关联信息;

[0014] 提取单元,被配置为从所述关联信息中提取每个历史行程的行程特征信息以及所述每个历史行程的实际行程时长信息作为样本信息,所述行程特征信息包括所述每个历史行程的全局特征信息以及所述每个历史行程包含的各个路段的路段特征信息;

[0015] 训练单元,被配置为基于所述样本信息训练出匹配的时间序列模型和回归模型。



- [0016] 根据本申请实施例的第四方面,提供一种估算行程时间的装置,包括:
- [0017] 获取单元,被配置为获取目标行程的关联信息;
- [0018] 提取单元,被配置为从所述关联信息中提取所述目标行程的行程特征信息,所述行程特征信息包括行程全程的全局特征信息以及行程中包含的各个路段的路段特征信息;
- [0019] 估算单元,被配置为基于所述行程特征信息采用预先训练的时间序列模型和回归模型估算出所述目标行程的行程时间。
- [0020] 根据本申请实施例的第五方面,提供一种计算机存储介质,所述存储介质中存储有程序指令,所述指令包括:
- [0021] 获取历史行程的关联信息;
- [0022] 从所述关联信息中提取每个历史行程的行程特征信息以及所述每个历史行程的实际行程时长信息作为样本信息,所述行程特征信息包括所述每个历史行程的全局特征信息以及所述每个历史行程包含的各个路段的路段特征信息;
- [0023] 基于所述样本信息训练出匹配的时间序列模型和回归模型。
- [0024] 根据本申请实施例的第六方面,提供一种计算机存储介质,所述存储介质中存储有程序指令,所述指令包括:
- [0025] 获取目标行程的关联信息;
- [0026] 从所述关联信息中提取所述目标行程的行程特征信息,所述行程特征信息包括行程全程的全局特征信息以及行程中包含的各个路段的路段特征信息;
- [0027] 基于所述行程特征信息采用预先训练的时间序列模型和回归模型估算出所述目标行程的行程时间。
- [0028] 本申请的实施例提供的技术方案可以包括以下有益效果:
- [0029] 本申请的实施例提供的用于估算行程时间的模型训练方法及装置,采用了机器学习的方法,结合历史行程中路段的特征信息以及历史行程全程的全局特征信息,对时间序列模型和回归模型进行训练,从而得到了可以用于估算行程时间的模型,由于历史行程中路段的特征信息与全局特征信息的结合,将更多的行进速度影响因素考虑在内,因此,提高了训练出的模型估算行程时间的准确度。
- [0030] 本申请的实施例提供的估算行程时间的方法及装置,通过结合目标行程中路段的特征信息以及目标行程全程的全局特征信息,采用预先训练的时间序列模型和回归模型,估算出行程时间,由于目标行程中路段的特征信息与全局特征信息的结合,将更多的行进速度影响因素考虑在内,因此,提高了行程时间估算结果的准确度。
- [0031] 应当理解的是,以上的一般描述和后文的细节描述仅是示例性和解释性的,并不能限制本申请。

## 附图说明

- [0032] 此处的附图被并入说明书中并构成本说明书的一部分,示出了符合本申请的实施例,并与说明书一起用于解释本申请的原理。
- [0033] 图1是本申请根据一示例性实施例示出的一种用于估算行程时间的模型训练方法的流程图;
- [0034] 图2是本申请根据一示例性实施例示出的另一种用于估算行程时间的模型训练方

法的流程图；

[0035] 图3是本申请根据一示例性实施例示出的一种估算行程时间的方法的流程图；

[0036] 图4是本申请根据一示例性实施例示出的另一种估算行程时间的方法的流程图；

[0037] 图5是本申请根据一示例性实施例示出的一种用于估算行程时间的模型训练装置框图；

[0038] 图6是本申请根据一示例性实施例示出的一种估算行程时间的装置框图。

### 具体实施方式

[0039] 这里将详细地对示例性实施例进行说明，其示例表示在附图中。下面的描述涉及附图时，除非另有表示，不同附图中的相同数字表示相同或相似的要素。以下示例性实施例中所述的实施方式并不代表与本申请相一致的所有实施方式。相反，它们仅是与如所附权利要求书中所详述的、本申请的一些方面相一致的装置和方法的例子。

[0040] 在本申请使用的术语是仅仅出于描述特定实施例的目的，而非旨在限制本申请。在本申请和所附权利要求书中所使用的单数形式的“一种”、“所述”和“该”也旨在包括多数形式，除非上下文清楚地表示其他含义。还应当理解，本文中使用的术语“和/或”是指并包含一个或多个相关联的列出项目的任何或所有可能组合。

[0041] 应当理解，尽管在本申请可能采用术语第一、第二、第三等来描述各种信息，但这些信息不应限于这些术语。这些术语仅用来将同一类型的信息彼此区分开。例如，在不脱离本申请范围的情况下，第一信息也可以被称为第二信息，类似地，第二信息也可以被称为第一信息。取决于语境，如在此所使用的词语“如果”可以被解释成为“在……时”或“当……时”或“响应于确定”。

[0042] 如图1所示，图1是根据一示例性实施例示出的一种用于估算行程时间的模型训练方法的流程图，该方法可以应用于终端中，也可以应用于服务端中。该方法包括以下步骤：

[0043] 在步骤101中，获取历史行程的关联信息。

[0044] 在本实施例中，历史行程为曾经完成的出行行程。一般来说，在LBS应用(如，行车应用，配送应用，地图应用以及导航应用等)的使用过程中，通常会生成并记录一些和用户的历史行程相关的历史数据。例如，可以通过记录LBS应用的历史订单，或者记录LBS应用下载的历史地图信息等等，从而将和用户的历史行程相关的历史数据记录下来。这些历史数据可以包括用户的标识信息、该历史行程的起点终点以及路径、该历史行程中路况的信息、用户完成该历史行程所花时间、用户采用的出行方式以及用户完成该历史行程时所处的时间段等信息。因此，可以从上述存储的历史数据中获取多个历史行程的关联信息。

[0045] 在本实施例中，历史行程的关联信息可以包括以下一项或多项：该历史行程的路径信息；该历史行程对应的起始以及终止时刻信息；该历史行程对应的日期信息；该历史行程对应的路况信息(如，拥挤状况、平均行进速度、人流量以及车流量等)；完成该历史行程的实际时长；该历史行程中的路段的信息(如，该历史行程中的路段的数量以及每个路段的标识等)；该历史行程对应的人员信息(如完成该历史行程的人员的标识信息等)；该历史行程对应的出行方式的信息(如乘/驾车完成该历史行程、或者骑车完成该历史行程、或者步行完成该历史行程等)。可以理解，历史行程的关联信息还可以包括其它的信息，本申请对历史行程的关联信息的具体内容方面不限定。

[0046] 在步骤102中,从上述关联信息中提取每个历史行程的行程特征信息以及每个历史行程的实际行程时长信息作为样本信息。

[0047] 在本实施例中,以一个历史行程为单位,获取多个历史行程的关联信息。针对每个历史行程,可以从上述关联信息中提取出该历史行程的行程特征信息以及实际行程时长信息作为样本信息。具体来说,该历史行程的实际行程时长信息为完成该历史行程所花的实际时间。该历史行程的行程特征信息可以包括该历史行程的全程对应的全局特征信息,以及该历史行程包括的多个路段中每个路段的路段特征信息。其中,该路段为地图中预先被划分出的路段(link)。一般来说,地图中的道路通常会被预先划分成多个路段(link),每个路段对应一个唯一性标识信息,这些路段的划分一般来说是固定的不变的。

[0048] 在本实施例中,针对每个路段,路段特征信息可以包括以下任意一项或多项:该路段的标识特征向量;该路段对应的行程的起始时刻特征向量;所述路段对应的日期特征向量;该路段对应的路况特征向量;该路段对应的人员特征向量;该路段对应的出行方式的特征向量。可以理解,路段特征信息还可以包括其它的信息,本申请对路段特征信息的具体内容方面不限定。

[0049] 在本实施例中,全局特征信息包括以下任意一项或多项:该历史行程全程的路径特征向量;该历史行程全程对应的起始时刻特征向量;该历史行程全程对应的日期特征向量;该历史行程全程对应的路况特征向量;该历史行程全程对应的人员特征向量;该历史行程全程对应的出行方式的特征向量。可以理解,全局特征信息还可以包括其它的信息,本申请对全局特征信息的具体内容方面不限定。

[0050] 在步骤103中,基于上述样本信息训练出匹配的时间序列模型和回归模型。

[0051] 在本实施例中,可以采用上述获取到的多个历史行程的行程特征信息以及实际行程时长信息,训练出匹配的时间序列模型和回归模型。具体来说,首先可以执行修正的步骤,该修正的步骤可以包括从上述样本信息中提取训练数据,并采用该训练数据对当前的时间序列模型和回归模型中的参数进行调整。参数调整完成后,判断该修正的步骤产生的损失函数是否收敛到收敛阈值。如果收敛到收敛阈值,则输出经过修正后的时间序列模型和回归模型作为匹配的时间序列模型和回归模型。如果未收敛到收敛阈值,则继续执行上述修正的步骤,即,再从上述样本信息中提取训练数据,并采用该训练数据继续对当前的时间序列模型和回归模型中的参数进行调整。直至该修正的步骤产生的损失函数收敛到收敛阈值为止,停止上述修正的步骤。

[0052] 在本实施例中,时间序列模型可以是RNN(Recurrent Neural Networks,循环神经网络),也可以是CNN(Convolutional neural networks,卷积神经网络),还可以是HMM(Hidden Markov Model,隐马尔可夫模型)等,可以理解,时间序列模型还可以是其它的模型,本申请对此方面不限定。回归模型可以是MLP(Multilayer Perceptron,多层神经网络),也可以是boosting算法等,可以理解,回归模型还可以是其它的模型,本申请对此方面不限定。

[0053] 本申请的上述实施例提供的用于估算行程时间的模型训练方法,采用了机器学习的方法,结合历史行程中路段的特征信息以及历史行程全程的全局特征信息,对时间序列模型和回归模型进行训练,从而得到了可以用于估算行程时间的模型,由于历史行程中路段的特征信息与全局特征信息的结合,将更多的行进速度影响因素考虑在内,因此,提高了

训练出的模型估算行程时间的准确度。

[0054] 图2是根据一示例性实施例示出的另一种用于估算行程时间的模型训练方法的流程图,该实施例详细描述了基于样本信息训练出匹配的时间序列模型和回归模型的过程,该方法可以应用于终端中,也可以应用于服务端中。该方法可以包括以下步骤:

[0055] 在步骤201中,获取历史行程的关联信息。

[0056] 在步骤202中,从上述关联信息中提取每个历史行程的行程特征信息以及每个历史行程的实际行程时长信息作为样本信息。

[0057] 在步骤203中,从样本信息中提取训练数据。

[0058] 在本实施例中,从上述关联信息中提取的样本信息涉及历史行程中路段的特征信息,并且不同的历史行程所包括的路段数量可能不同。因此,在一个训练阶段(执行一次修正步骤),如果从样本信息中选取的训练数据所对应的历史行程的路段数量(即训练数据的序列长度)相差过大,则会造成计算资源的浪费,降低了模型训练的效率。所以,在一个训练阶段,需要尽可能使选取的训练数据的序列长度比较接近。

[0059] 在一种实现方式中,在一个训练阶段,可以从上述样本信息中提取多个历史行程的行程特征信息以及实际行程时长信息作为训练数据,这多个历史行程之间行程路段的数量差值不超过预设阈值。例如,提取的训练数据所对应的历史行程的路段数量最大为a,最小为b,则a和b相差的值不能超过c。其中,上述预设阈值可以是一个经验值,预设阈值可以取任意合理的数值,本申请预设阈值的具体取值方面不限定。

[0060] 在另一种实现方式中,还可以预先将行程路段的数量划分为多个范围,根据行程路段的数量对历史行程进行分组,使同一组的历史行程的路段数量划分在同一个范围内。例如,可以将路段数量划分为A、B、C三个范围,分别为 $[1, 10)$ ,  $[11, 20)$ ,  $[21, 30)$ ,将行程路段的数量在A范围内的历史行程划分到A组,将行程路段的数量在B范围内的历史行程划分到B组,行程路段的数量在C范围内的历史行程划分到C组。在一个训练阶段,可以从样本信息中提取同一组的多个历史行程的行程特征信息以及实际行程时长信息作为训练数据。

[0061] 在步骤204中,采用上述训练数据对当前的时间序列模型和回归模型中的参数进行调整。

[0062] 在本实施例中,可以将上述训练数据中的路段特征信息依次输入到当前的时间序列模型中,并获取该时间序列模型输出的隐状态信息。将该隐状态信息以及上述训练数据中的全局特征信息输入到当前的回归模型中,并获取该回归模型输出的预估时间信息。根据该预估时间信息与上述训练数据中实际行程时长信息的差异,调整当前的时间序列模型和回归模型中的参数。如果该预估时间信息与实际行程时长信息的差异越大,对上述参数的调整幅度越大,如果该预估时间信息与实际行程时长信息的差异越小,对上述参数的调整幅度越小。

[0063] 在步骤205中,判断该修正之后的时间序列模型和回归模型是否满足匹配条件。

[0064] 在步骤206中,如果未满足匹配条件,重新执行步骤203-步骤205,如果满足匹配条件,输出修正之后的时间序列模型和回归模型作为匹配的时间序列模型和回归模型。

[0065] 在一种实现方式中,可以在修正的步骤之后,判断该修正的步骤产生的损失函数是否收敛到收敛阈值,若收敛到收敛阈值,则说明满足匹配条件,若未收敛到收敛阈值,则说明未满足匹配条件。

[0066] 在另一种实现方式中,也可以在修正的步骤之后,采用从样本信息中抽取的验证数据,对修正之后的时间序列模型和回归模型进行验证,以获得验证误差。判断该验证误差是否小于预设误差,若小于预设误差,则说明满足匹配条件,若大于或等于预设误差,则说明未满足匹配条件。

[0067] 在又一种实现方式中,还可以在修正的步骤之后,即判断该修正的步骤产生的损失函数是否收敛到收敛阈值,又采用从样本信息中抽取的验证数据,对修正之后的时间序列模型和回归模型进行验证,以获得验证误差。当上述损失函数收敛到收敛阈值,并且上述验证误差小于预设误差时,说明满足匹配条件,否则说明未满足匹配条件。

[0068] 可以理解,还可以通过其它的方式判断该修正之后的时间序列模型和回归模型是否满足匹配条件,本申请对此方面不限定。

[0069] 在本实施例中,收敛阈值以及预设误差均可以是经验值,本申请对收敛阈值以及预设误差的具体取值方面不限定。在本实施例中,匹配的时间序列模型和回归模型即为训练好的模型,可以用于估算行程时间。

[0070] 需要说明的是,对于与图1实施例中相同的步骤,在上述图2实施例中不再进行赘述,相关内容可参见图1实施例。

[0071] 本申请的上述实施例提供的用于估算行程时间的模型训练方法,采用了机器学习的方法,结合历史行程中路段的特征信息以及历史行程全程的全局特征信息,对时间序列模型和回归模型的参数进行调整,如果调整时的损失函数未收敛到收敛阈值,则继续对时间序列模型和回归模型的参数进行调整,直至损失函数收敛到收敛阈值,停止上述调整。从而得到了可以用于估算行程时间的模型,由于历史行程中路段的特征信息与全局特征信息的结合,将更多的行进速度影响因素考虑在内,因此,提高了训练出的模型估算行程时间的准确度。

[0072] 如图3所示,图3是根据一示例性实施例示出的一种估算行程时间的方法的流程图,该方法可以应用于终端中,也可以应用于服务端中。该方法可以包括以下步骤:

[0073] 在步骤301中,获取目标行程的关联信息。

[0074] 在本实施例中,目标行程为待估算的行程,目标行程的关联信息可以包括以下任意一项或多项:目标行程的路径信息;目标行程对应的起始和终止时刻信息;所述目标行程对应的日期信息;目标行程对应的路况信息(如,拥挤状况、平均行进速度、人流量以及车流量等);目标行程中的路段的信息(如,该目标行程中的路段的数量以及每个路段的标识等);目标行程对应的人员信息(如需要完成该目标行程的人员的标识信息等);目标行程对应的出行方式信息(如需要乘/驾车完成该目标行程、或者需要骑车完成该目标行程、或者需要步行完成该目标行程等)。可以理解,目标行程的关联信息还可以包括其它的信息,本申请对目标行程的关联信息的具体内容方面不限定。

[0075] 在步骤302中,从上述关联信息中提取目标行程的行程特征信息。

[0076] 在本实施例中,可以从上述关联信息中提取出该目标行程的行程特征信息。该目标行程的行程特征信息可以包括该目标行程的全程对应的全局特征信息,以及该目标行程包括的多个路段中每个路段的路段特征信息。其中,该路段为地图中预先被划分出的路段(link)。

[0077] 在本实施例中,针对每个路段的路段特征信息可以包括以下任意一项或多项:该

路段的标识特征向量;该路段对应的目标行程的起始时刻特征向量;所述路段对应的日期特征向量;该路段对应的路况特征向量;该路段对应的人员特征向量;该路段对应的出行方式信息。可以理解,路段特征信息还可以包括其它的信息,本申请对路段特征信息的具体内容方面不限定。

[0078] 在本实施例中,该目标行程的全局特征信息包括以下任意一项或多项:该目标行程全程的路径特征向量;该目标行程全程对应的起始时刻特征向量;该目标行程全程对应的日期特征向量;该目标行程全程对应的路况特征向量;该目标行程对应的人员特征向量;该目标行程全程对应的出行方式特征向量。可以理解,该目标行程的全局特征信息还可以包括其它的信息,本申请对该目标行程的全局特征信息的具体内容方面不限定。

[0079] 在步骤303中,基于上述行程特征信息采用预先训练的时间序列模型和回归模型估算出目标行程的行程时间。

[0080] 在本实施例中,可以基于上述行程特征信息采用预先训练的时间序列模型和回归模型估算出目标行程的行程时间。具体来说,首先,将行程特征信息中的路段特征信息输入到预先训练的时间序列模型中,并获取该时间序列模型输出的隐状态信息。接着,将该隐状态信息以及行程特征信息中的全局特征信息输入到预先训练的回归模型中,并获取该回归模型输出的结果,该结果即为估算出的目标行程的行程时间。

[0081] 在本实施例中,预先训练的时间序列模型可以是预先训练的RNN,也可以是预先训练的CNN,还可以是预先训练的HMM等,可以理解,预先训练的时间序列模型还可以是其它的模型,本申请对此方面不限定。预先训练的回归模型可以是预先训练的MLP,也可以是预先训练的boosting模型,可以理解,预先训练的回归模型还可以是其它的模型,本申请对此方面不限定。

[0082] 本申请的上述实施例提供的估算行程时间的方法,通过结合目标行程中路段的特征信息以及目标行程全程的全局特征信息,采用预先训练的时间序列模型和回归模型,估算出行程时间,由于目标行程中路段的特征信息与全局特征信息的结合,将更多的行进速度影响因素考虑在内,因此,提高了行程时间估算结果的准确度。

[0083] 如图4所示,图4是根据一示例性实施例示出的另一种估算行程时间的方法的流程图,该实施例详细描述了基于上述行程特征信息采用预先训练的时间序列模型和回归模型估算出目标行程的行程时间的过程,该方法可以应用于终端中,也可以应用于服务端中。该方法可以包括以下步骤:

[0084] 在步骤401中,获取目标行程的关联信息。

[0085] 在步骤402中,从上述关联信息中提取目标行程的行程特征信息。

[0086] 在步骤403中,将上述行程特征信息中的路段特征信息输入到预先训练的时间序列模型。

[0087] 在步骤404中,将上述时间序列模型输出的隐状态信息以及上述行程特征信息中的全局特征信息输入到预先训练的回归模型。

[0088] 在步骤405中,获取上述回归模型输出的结果作为估算出的目标行程的行程时间。

[0089] 需要说明的是,对于与图3实施例中相同的步骤,在上述图4实施例中不再进行赘述,相关内容可参见图5实施例。

[0090] 本申请的上述实施例提供的估算行程时间的方法,通过获取目标行程的关联信

息,从上述关联信息中提取目标行程的行程特征信息,将上述行程特征信息中的路段特征信息输入到预先训练的时间序列模型,将上述时间序列模型输出的隐状态信息以及上述行程特征信息中的全局特征信息输入到预先训练的回归模型,并获取上述回归模型输出的结果作为估算出的目标行程的行程时间。由于目标行程中路段的特征信息与全局特征信息的结合,将更多的行进速度影响因素考虑在内,因此,提高了行程时间估算结果的准确度。

[0091] 应当注意,尽管在附图中以特定顺序描述了本申请方法的操作,但是,这并非要求或者暗示必须按照该特定顺序来执行这些操作,或是必须执行全部所示的操作才能实现期望的结果。相反,流程图中描绘的步骤可以改变执行顺序。附加地或备选地,可以省略某些步骤,将多个步骤合并为一个步骤执行,和/或将一个步骤分解为多个步骤执行。

[0092] 与前述用于估算行程时间、模型训练的方法实施例相对应,本申请还提供了用于估算行程时间、模型训练的方法的装置的实施例。

[0093] 如图5所示,图5是本申请根据一示例性实施例示出的一种用于估算行程时间的模型训练装置框图,该装置包括:获取单元501,提取单元502以及训练单元503。

[0094] 其中,获取单元501,被配置为获取历史行程的关联信息。

[0095] 提取单元502,被配置为从上述关联信息中提取每个历史行程的行程特征信息以及每个历史行程的实际行程时长信息作为样本信息,该行程特征信息包括每个历史行程的全局特征信息以及每个历史行程包含的各个路段的路段特征信息。

[0096] 训练单元503,被配置为基于上述样本信息训练出匹配的时间序列模型和回归模型。

[0097] 在一些可选实施方式中,训练单元503可以包括:修正子单元,控制子单元和输出子单元(图5未示出)。

[0098] 其中,修正子单元,被配置为执行修正的步骤,修正的步骤包括:从上述样本信息中提取训练数据,并采用该训练数据对当前的时间序列模型和回归模型中的参数进行调整。

[0099] 控制子单元,被配置为在上述修正的步骤产生的损失函数未收敛到收敛阈值时,控制修正子单元继续执行上述修正的步骤。

[0100] 输出子单元,被配置为在上述修正的步骤产生的损失函数收敛到收敛阈值时,输出经过修正后的时间序列模型和回归模型作为匹配的时间序列模型和回归模型。

[0101] 在另一些可选实施方式中,修正子单元通过如下方式从样本信息中提取训练数据:

[0102] 从样本信息中提取多个历史行程的行程特征信息以及实际行程时长信息作为训练数据,该多个历史行程之间行程路段的数量差值不超过预设阈值。

[0103] 在另一些可选实施方式中,该装置进一步还包括:分组单元。

[0104] 其中,分组单元,被配置为将行程路段的数量划分为多个范围,根据行程路段的数量对上述历史行程进行分组,使同一组的历史行程的路段数量划分在同一个范围内。

[0105] 修正子单元通过如下方式从样本信息中提取训练数据:

[0106] 从上述样本信息中提取同一组的多个历史行程的行程特征信息以及实际行程时长信息作为训练数据。

[0107] 在另一些可选实施方式中,修正子单元通过如下方式采用训练数据对当前的时间

序列模型和回归模型中的参数进行调整：

[0108] 将训练数据中的路段特征信息输入到当前的时间序列模型中；

[0109] 将当前的时间序列模型输出的隐状态信息以及训练数据中的全局特征信息输入到当前的回归模型中；

[0110] 根据当前的回归模型输出的预估时间信息与训练数据中实际行程时长信息的差异，调整当前的时间序列模型和回归模型中的参数。

[0111] 在另一些可选实施方式中，时间序列模型可以包括循环神经网络RNN；回归模型可以包括多层感知器MLP。

[0112] 在另一些可选实施方式中，该历史行程的关联信息包括以下任意一项或多项：

[0113] 该历史行程的路径信息；

[0114] 该历史行程对应的起始时刻信息；

[0115] 该历史行程对应的终止时刻信息；

[0116] 所述历史行程对应的日期信息；

[0117] 该历史行程对应的路况信息；

[0118] 完成该历史行程的实际行程时长；

[0119] 该历史行程中的路段的信息；

[0120] 该历史行程对应的人员信息；

[0121] 该历史行程对应的出行方式的信息。

[0122] 在另一些可选实施方式中，路段特征信息包括以下任意一项或多项：

[0123] 该路段的标识特征向量；

[0124] 该路段对应的行程的起始时刻特征向量；

[0125] 该路段对应的日期特征向量；

[0126] 该路段对应的路况特征向量；

[0127] 该路段对应的人员特征向量；

[0128] 该路段对应的出行方式的特征向量。

[0129] 在另一些可选实施方式中，全局特征信息包括以下任意一项或多项：

[0130] 该历史行程的路径特征向量；

[0131] 该历史行程对应的起始时刻特征向量；

[0132] 该历史行程对应的日期特征向量；

[0133] 该历史行程对应的路况特征向量；

[0134] 该历史行程对应的人员特征向量；

[0135] 该历史行程对应的出行方式的特征向量。

[0136] 应当理解，上述装置可以预先设置在终端或服务端中，也可以通过下载等方式而加载到终端或服务端中。上述装置中的相应单元可以与终端或服务端中的单元相互配合以实现用于估算行程时间的模型训练方案。

[0137] 如图6所示，图6是本申请根据一示例性实施例示出的一种估算行程时间的装置框图，该装置包括：获取单元601，提取单元602以及估算单元603。

[0138] 其中，获取单元601，被配置为获取目标行程的关联信息。

[0139] 提取单元602，被配置为从关联信息中提取目标行程的行程特征信息，该行程特征



信息包括行程全程的全局特征信息以及行程中包含的各个路段的路段特征信息。

[0140] 估算单元603,被配置为基于该行程特征信息采用预先训练的时间序列模型和回归模型估算出目标行程的行程时间。

[0141] 在一些可选实施方式中,估算单元603可以包括:第一输入子单元,第二输入子单元和输出子单元(图6未示出)。

[0142] 第一输入子单元,被配置为将行程特征信息中的路段特征信息输入到预先训练的时间序列模型。

[0143] 第二输入子单元,被配置为将时间序列模型输出的隐状态信息以及行程特征信息中的全局特征信息输入到预先训练的回归模型。

[0144] 输出子单元,被配置为获取回归模型输出的结果作为估算出的目标行程的行程时间。

[0145] 在另一些可选实施方式中,时间序列模型包括循环神经网络RNN;回归模型包括多层感知器MLP。

[0146] 在另一些可选实施方式中,目标行程的关联信息包括以下任意一项或多项:

[0147] 该目标行程的路径信息;

[0148] 该目标行程对应的起始时刻信息;

[0149] 该目标行程对应的终止时刻信息;

[0150] 该目标行程对应的日期信息;

[0151] 该目标行程对应的路况信息;

[0152] 该目标行程中的路段的信息;

[0153] 该目标行程对应的人员信息;

[0154] 该目标行程对应的出行方式信息。

[0155] 在另一些可选实施方式中,该路段特征信息包括以下任意一项或多项:

[0156] 该路段的标识特征向量;

[0157] 该路段对应的目标行程的起始时刻特征向量;

[0158] 该路段对应的日期特征向量;

[0159] 该路段对应的路况特征向量;

[0160] 该路段对应的人员特征向量;

[0161] 该路段对应的出行方式信息。

[0162] 在另一些可选实施方式中,该全局特征信息包括以下任意一项或多项:

[0163] 该目标行程的路径特征向量;

[0164] 该目标行程对应的起始时刻特征向量;

[0165] 该目标行程对应的日期特征向量;

[0166] 该目标行程对应的路况特征向量;

[0167] 该目标行程对应的人员特征向量;

[0168] 该目标行程对应的出行方式特征向量。

[0169] 应当理解,上述装置可以预先设置在终端或服务端中,也可以通过下载等方式而加载到终端或服务端中。上述装置中的相应单元可以与终端或服务端中的单元相互配合以实现估算行程时间的方案。

[0170] 对于装置实施例而言,由于其基本对应于方法实施例,所以相关之处参见方法实施例的部分说明即可。以上所描述的装置实施例仅仅是示意性的,其中所述作为分离部件说明的单元可以是或者也可以不是物理上分开的,作为单元显示的部件可以是或者也可以不是物理单元,即可以位于一个地方,或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部模块来实现本申请方案的目的。本领域普通技术人员在不付出创造性劳动的情况下,即可以理解并实施。

[0171] 本申请实施例可采用在一个或多个其中包含有程序代码的存储介质(包括但不限于磁盘存储器、CD-ROM、光学存储器等)上实施的计算机程序产品的形式。

[0172] 相应的,本申请实施例还提供一种计算机存储介质,该存储介质中存储有程序指令,该指令包括:

[0173] 获取历史行程的关联信息;

[0174] 从所述关联信息中提取每个历史行程的行程特征信息以及所述每个历史行程的实际行程时长信息作为样本信息,所述行程特征信息包括所述每个历史行程的全局特征信息以及所述每个历史行程包含的各个路段的路段特征信息;

[0175] 基于所述样本信息训练出匹配的时间序列模型和回归模型。

[0176] 相应的,本申请实施例还提供一种计算机存储介质,该存储介质中存储有程序指令,该指令包括:

[0177] 获取目标行程的关联信息;

[0178] 从所述关联信息中提取所述目标行程的行程特征信息,所述行程特征信息包括行程全程的全局特征信息以及行程中包含的各个路段的路段特征信息;

[0179] 基于所述行程特征信息采用预先训练的时间序列模型和回归模型估算出所述目标行程的行程时间。

[0180] 描述于本申请实施例中所涉及到的单元模块可以通过软件的方式实现,也可以通过硬件的方式来实现。所描述的单元模块也可以设置在处理器中,例如,可以描述为:一种处理器包括获取单元,提取单元以及训练单元。其中,这些单元模块的名称在某种情况下并不构成对该单元模块本身的限定,例如,获取单元还可以被描述为“用于获取历史行程的关联信息的单元”。

[0181] 作为另一方面,本申请还提供了一种计算机可读存储介质,该计算机可读存储介质可以是上述实施例中所述装置中所包含的计算机可读存储介质;也可以是单独存在,未装配入终端或服务器中的计算机可读存储介质。该计算机可读存储介质存储有一个或者一个以上程序,该程序被一个或者一个以上的处理器用来执行描述于本申请的用于估算行程时间的方法以及模型训练的方法。

[0182] 计算机可用存储介质包括永久性和非永久性、可移动和非可移动媒体,可以由任何方法或技术来实现信息存储。信息可以是计算机可读指令、数据结构、程序的模块或其他数据。计算机的存储介质的例子包括但不限于:相变内存(PRAM)、静态随机存取存储器(SRAM)、动态随机存取存储器(DRAM)、其他类型的随机存取存储器(RAM)、只读存储器(ROM)、电可擦除可编程只读存储器(EEPROM)、快闪记忆体或其他内存技术、只读光盘只读存储器(CD-ROM)、数字多功能光盘(DVD)或其他光学存储、磁盒式磁带,磁带磁盘存储或其他磁性存储设备或任何其他非传输介质,可用于存储可以被计算设备访问的信息。

[0183] 本领域技术人员在考虑说明书及实践这里公开的发明后,将容易想到本申请的其它实施方案。本申请旨在涵盖本申请的任何变型、用途或者适应性变化,这些变型、用途或者适应性变化遵循本申请的一般性原理并包括本申请未公开的本技术领域中的公知常识或惯用技术手段。说明书和实施例仅被视为示例性的,本申请的真正范围和精神由下面的权利要求指出。

[0184] 应当理解的是,本申请并不局限于上面已经描述并在附图中示出的精确结构,并且可以在不脱离其范围进行各种修改和改变。本申请的范围仅由所附的权利要求来限制。

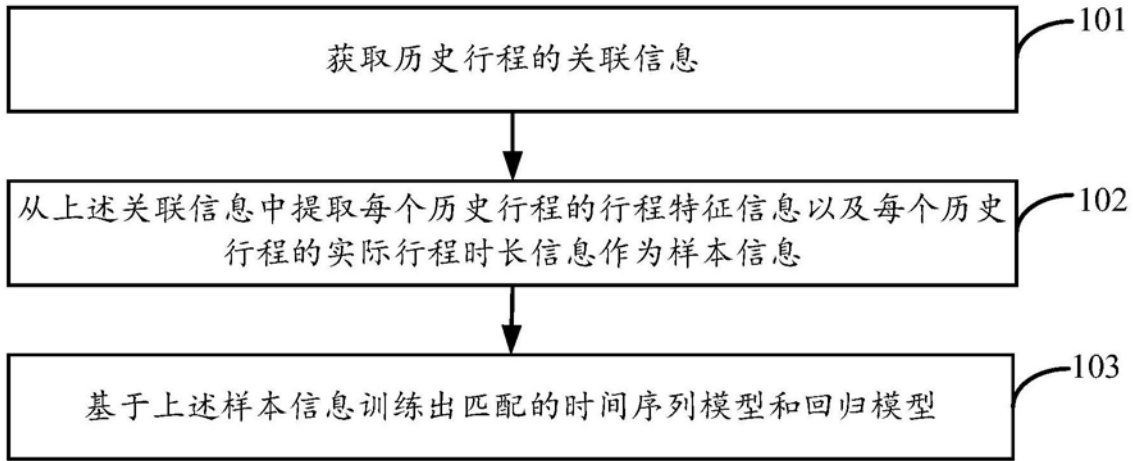


图1

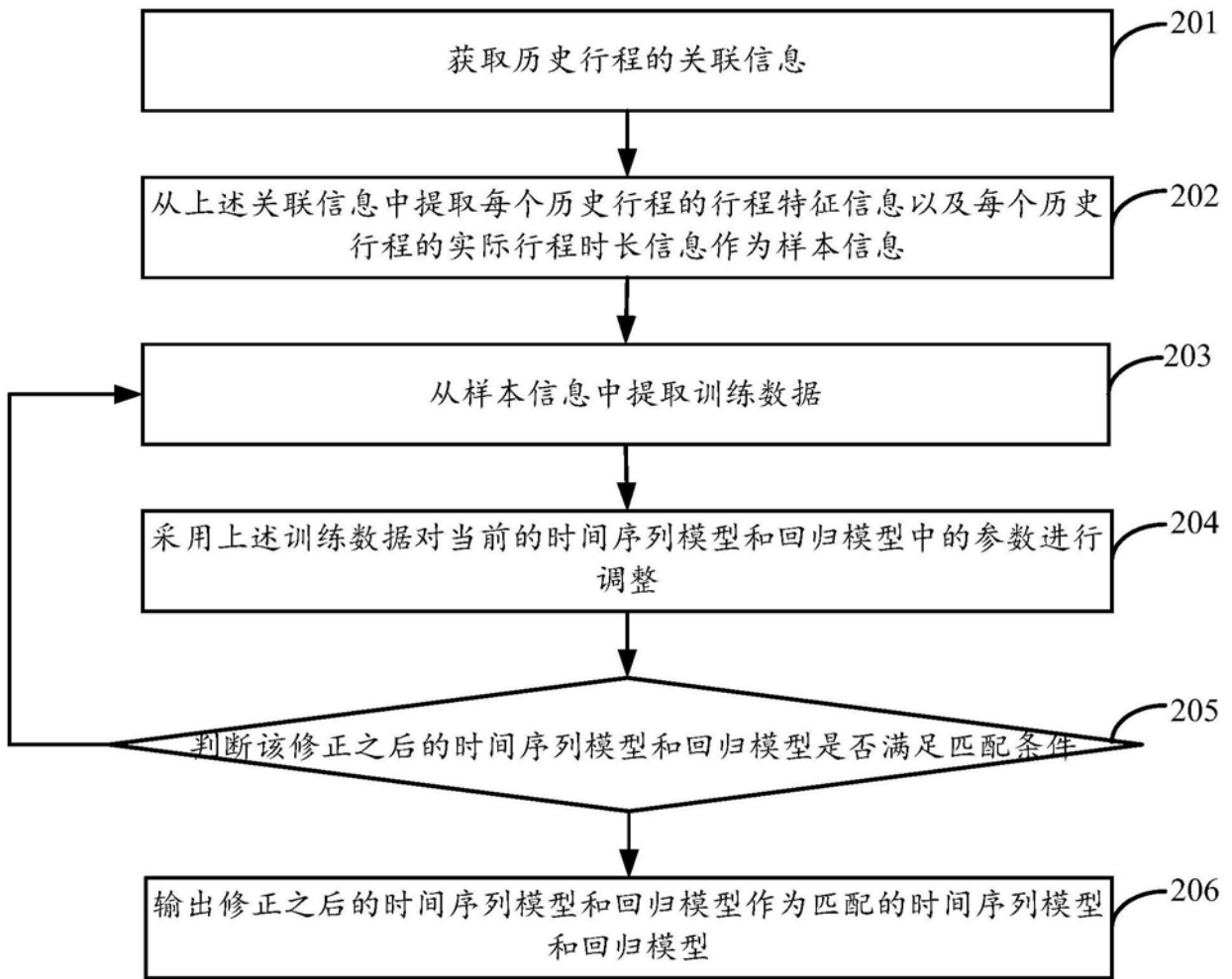


图2

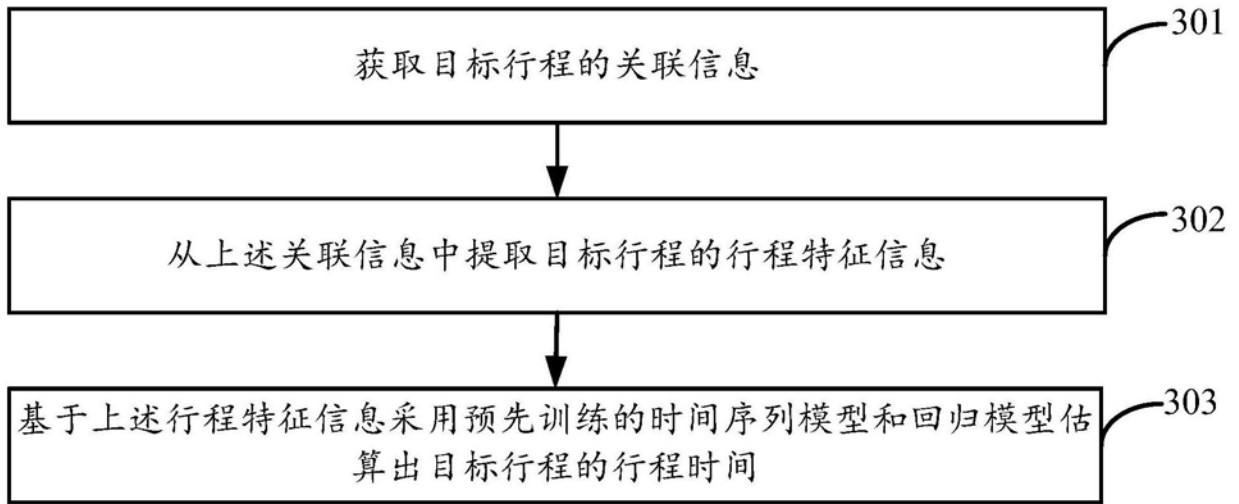


图3

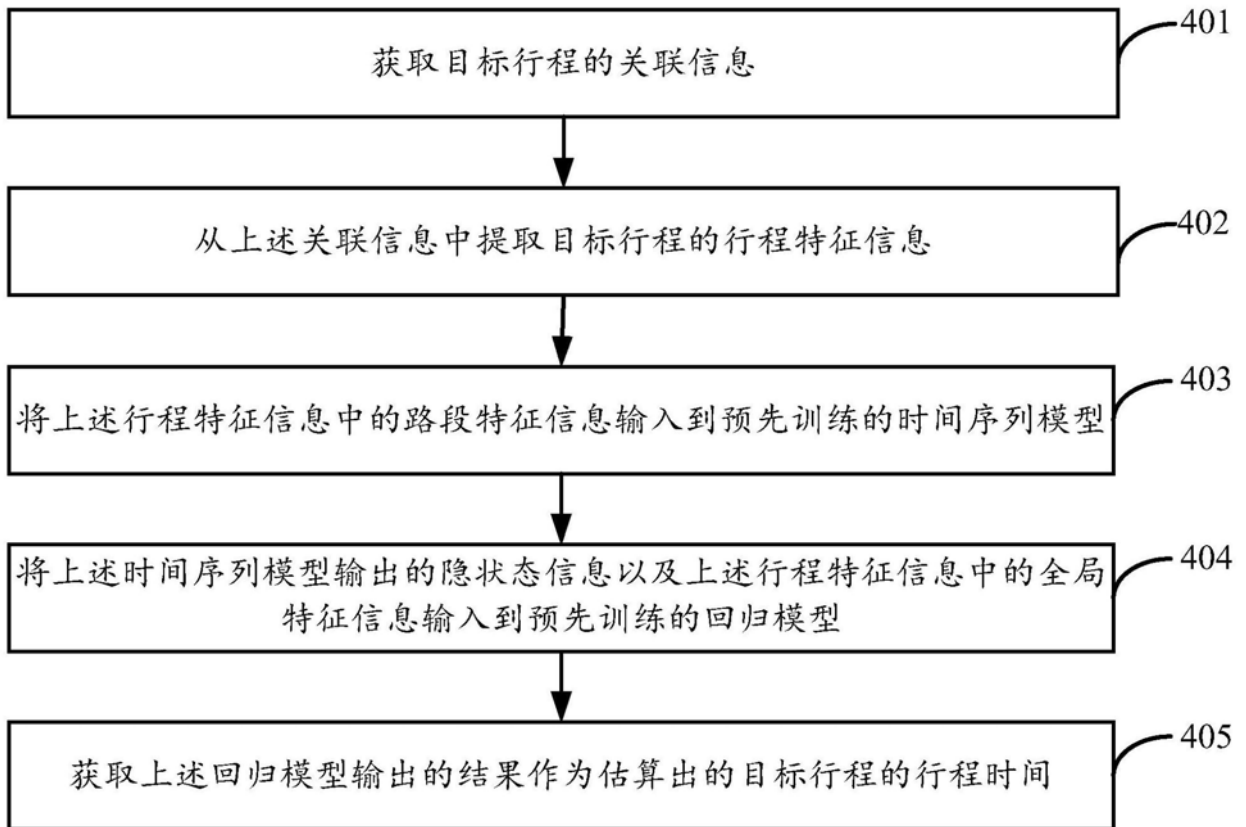


图4

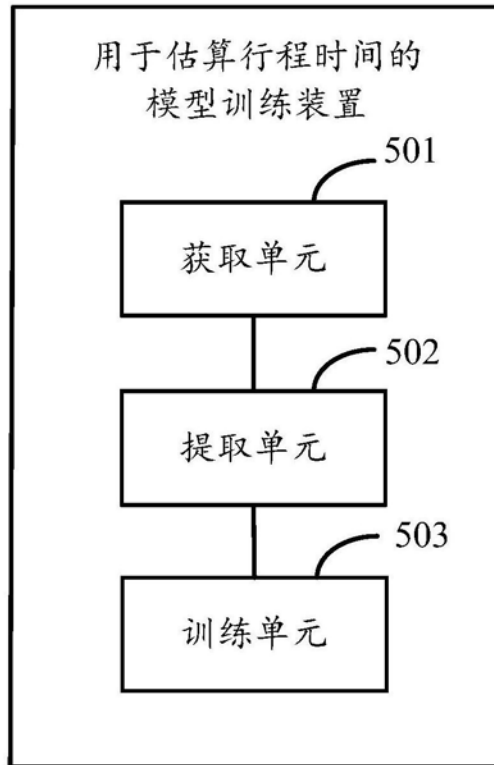


图5

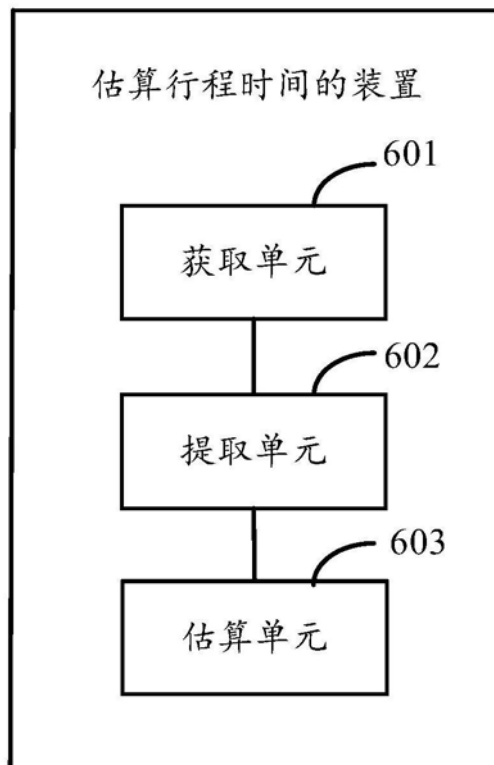


图6