



(12) 发明专利

(10) 授权公告号 CN 112309109 B

(45) 授权公告日 2022. 02. 18

(21) 申请号 201910709193.0

G06K 9/62 (2006.01)

(22) 申请日 2019.08.01

审查员 强宇辰

(65) 同一申请的已公布的文献号

申请公布号 CN 112309109 A

(43) 申请公布日 2021.02.02

(73) 专利权人 中移(苏州)软件技术有限公司

地址 215163 江苏省苏州市高新区昆仑山路58号1幢

专利权人 中国移动通信集团有限公司

(72) 发明人 张蒙澄

(74) 专利代理机构 北京派特恩知识产权代理有限公司 11270

代理人 王姗姗 张颖玲

(51) Int. Cl.

G08G 1/01 (2006.01)

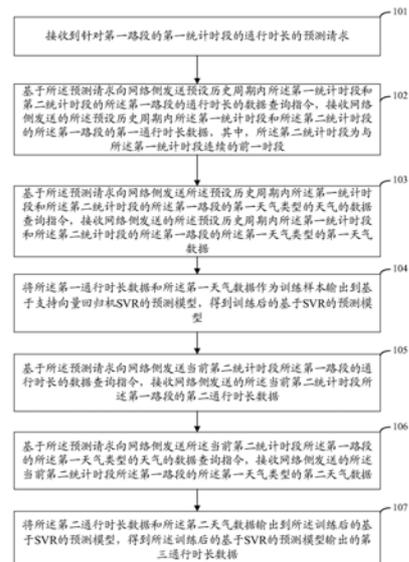
权利要求书4页 说明书23页 附图6页

(54) 发明名称

一种道路通行时间预测的方法、装置及存储介质

(57) 摘要

本发明公开了一种道路通行时长预测的方法,所述方法包括:将第一通行时长数据和第一天气类型的第一天气数据作为训练样本输出到基于支持向量回归机SVR的预测模型,得到训练后的基于SVR的预测模型;将第二通行时长数据和第一天气类型的第二天气数据输出到所述训练后的第二统计时段通行时长预测模型,得到所述训练后的第二统计时段通行时长预测模型输出的第三通行时长数据。通过第一天气类型的天气数据对通行时长数据的波动进行修正,降低通行时长数据的波动对预测结果产生的误差,提高通行时长预测的准确度。本发明还公开了一种道路通行时长预测的装置。



1. 一种道路通行时长预测的方法,其特征在于,所述方法包括:

接收到针对第一路段的第一统计时段的通行时长的预测请求;

基于所述预测请求向网络侧发送预设历史周期内所述第一统计时段和第二统计时段的所述第一路段的通行时长的数据查询指令,接收网络侧发送的所述预设历史周期内所述第一统计时段和所述第二统计时段的所述第一路段的第一通行时长数据,其中,所述第二统计时段为与所述第一统计时段连续的前一时段;

基于所述预测请求向网络侧发送所述预设历史周期内所述第一统计时段和所述第二统计时段的所述第一路段的第一天气类型的天气的数据查询指令,接收网络侧发送的所述预设历史周期内所述第一统计时段和所述第二统计时段的所述第一路段的所述第一天气类型的第一天气数据;

将所述第一通行时长数据和所述第一天气数据作为训练样本输出到基于支持向量回归机SVR的预测模型,得到训练后的基于SVR的预测模型;

基于所述预测请求向网络侧发送当前第二统计时段所述第一路段的通行时长的数据查询指令,接收网络侧发送的所述当前第二统计时段所述第一路段的第二通行时长数据;

基于所述预测请求向网络侧发送所述当前第二统计时段所述第一路段的所述第一天气类型的天气的数据查询指令,接收网络侧发送的所述当前第二统计时段所述第一路段的所述第一天气类型的第二天气数据;

将所述第二通行时长数据和所述第二天气数据输出到所述训练后的基于SVR的预测模型,得到所述训练后的基于SVR的预测模型输出的第三通行时长数据。

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:

基于所述预测请求向网络侧发送所述预设历史周期内所述第一统计时段和所述第二统计时段的所述第一路段的N种天气类型的天气的数据查询指令,接收网络侧发送的所述预设历史周期内所述第一统计时段和所述第二统计时段的所述第一路段的N种天气类型的天气数据,其中,N为正整数;

根据所述第一通行时长数据分别和所述N种天气类型的天气数据建立预测模型;验证建立的预测模型,确定验证结果与实际值偏差最小的预测模型对应的天气类型为所述第一天气类型。

3. 根据权利要求2所述的方法,其特征在于,所述根据所述第一通行时长数据分别和所述N种天气类型的天气数据建立预测模型;验证建立的预测模型,确定验证结果与实际值偏差最小的预测模型对应的天气类型为所述第一天气类型,具体包括:

将所述预设历史周期分为M个子周期,其中,M为正整数;并执行以下步骤:

步骤一:选取第i种天气类型的天气数据,其中,i为正整数, $1 \leq i \leq N$;

步骤二:选取第j个子周期作为测试周期,对应地将其余子周期作为训练周期,其中,j为正整数, $1 \leq j \leq M$;

步骤三:将训练周期的所述第一通行时长数据和第i种天气类型的天气数据作为训练样本输出到所述基于SVR的预测模型,得到训练后的第 $i \times j$ 个基于SVR的预测模型;

步骤四:将测试周期的第一统计时段的所述第一通行时长数据和第i种天气类型的天气数据输出到所述训练后的第 $i \times j$ 个基于SVR的预测模型,得到所述训练后的第 $i \times j$ 个基于SVR的预测模型输出的测试周期的第二统计时段通行时长预测值;

步骤五:根据测试周期的第二统计时段通行时长预测值和测试周期的第二统计时段的所述第一通行时长数据,计算测试周期对应的平均绝对百分比误差MAPE值;

步骤六:再次执行步骤二至步骤五,直至得到所述M个子周期分别作为测试周期时对应的MAPE值,将所述M个子周期分别作为测试周期时对应的MAPE值的平均值,作为第i种天气类型对应的MAPE值;

步骤七:再次执行步骤一至步骤六,直至得到所述N种天气类型对应的MAPE值,将所述N种天气类型中对应的MAPE值最小的天气类型作为所述第一天气类型。

4. 根据权利要求2或3所述的方法,其特征在于,所述方法还包括:

将数据输入基于SVR的预测模型之前,采用数据使用范围为[0,1]的最小最大值归一化Min-Max-scaling算法将数据归一化,将归一化后的数据输入基于SVR的预测模型。

5. 根据权利要求1所述的方法,其特征在于,所述方法还包括:

检验所述第一通行时长数据中是否有部分所述预设历史周期的数据缺失;

当所述第一通行时长数据中有部分所述预设历史周期的数据缺失时,向网络侧发送与所述第一路段有重叠部分的路段的信息查询请求,接收网络侧发送的与所述第一路段有重叠部分的第二路段的信息,其中,所述第二路段为多个;

基于所述第二路段的信息,向网络侧发送部分所述预设历史周期内所述第一统计时段和第二统计时段的所述第二路段的通行时长的数据,接收网络侧发送的部分所述预设历史周期内所述第一统计时段和第二统计时段的所述第二路段的第三通行时长数据;

根据所述第三通行时长数据中所述第二路段与所述第一路段的重叠部分对应的通行时长数据对所述第一通行时长数据中缺失的部分所述预设历史周期的数据进行补充。

6. 根据权利要求5所述的方法,其特征在于,所述方法还包括:

在根据所述第三通行时长数据中所述第二路段与所述第一路段的重叠部分对应的通行时长数据对所述第一通行时长数据中缺失的部分所述预设历史周期的数据进行补充后,再次检验所述第一通行时长数据中是否有部分所述预设历史周期的数据缺失;

当所述第一通行时长数据中仍有部分所述预设历史周期的数据缺失时,根据所述第一通行时长数据获取缺失部分的相邻周期的通行时长数据,计算所述相邻周期的通行时长数据的线性平均值,根据所述相邻周期的通行时长数据的线性平均值补充缺失部分的通行时长数据。

7. 一种道路通行时长预测的装置,其特征在于,所述装置包括:

请求模块,用于接收针对第一路段的第一统计时段的通行时长的预测请求;

第一交互模块,用于基于所述预测请求向网络侧发送预设历史周期内所述第一统计时段和第二统计时段的所述第一路段的通行时长的数据查询指令,接收网络侧发送的所述预设历史周期内所述第一统计时段和所述第二统计时段的所述第一路段的第一通行时长数据,其中,所述第二统计时段为与所述第一统计时段连续的前一时段;

第二交互模块,用于基于所述预测请求向网络侧发送所述预设历史周期内所述第一统计时段和所述第二统计时段的所述第一路段的第一天气类型的天气的数据查询指令,接收网络侧发送的所述预设历史周期内所述第一统计时段和所述第二统计时段的所述第一路段的所述第一天气类型的第一天气数据;

训练模块,用于将所述第一通行时长数据和所述第一天气数据作为训练样本输出到基

于支持向量回归机SVR的预测模型,得到训练后的基于SVR的预测模型;

所述第一交互模块,还用于基于所述预测请求向网络侧发送当前第二统计时段所述第一路段的通行时长的数据查询指令,接收网络侧发送的所述当前第二统计时段所述第一路段的第二通行时长数据;

所述第二交互模块,还用于基于所述预测请求向网络侧发送所述当前第二统计时段所述第一路段的所述第一天气类型的天气的数据查询指令,接收网络侧发送的所述当前第二统计时段所述第一路段的所述第一天气类型的第二天气数据;

预测模块,用于将所述第二通行时长数据和所述第二天气数据输出到所述训练后的基于SVR的预测模型,得到所述训练后的基于SVR的预测模型输出的第三通行时长数据。

8. 根据权利要求7所述的装置,其特征在于,所述第二交互模块,还用于基于所述预测请求向网络侧发送所述预设历史周期内所述第一统计时段和所述第二统计时段的所述第一路段的N种天气类型的天气的数据查询指令,接收网络侧发送的所述预设历史周期内所述第一统计时段和所述第二统计时段的所述第一路段的N种天气类型的天气数据,其中,N为正整数;

所述装置还包括:天气类型选择模块;

所述天气类型选择模块,用于根据所述第一通行时长数据分别和所述N种天气类型的天气数据建立预测模型;验证建立的预测模型,确定验证结果与实际值偏差最小的预测模型对应的天气类型为所述第一天气类型。

9. 根据权利要求8所述的装置,其特征在于,所述天气类型选择模块,具体用于:

将所述预设历史周期分为M个子周期,其中,M为正整数;并执行以下步骤:

步骤一:选取第i种天气类型的天气数据,其中,i为正整数, $1 \leq i \leq N$;

步骤二:选取第j个子周期作为测试周期,对应地将其余子周期作为训练周期,其中,j为正整数, $1 \leq j \leq M$;

步骤三:将训练周期的所述第一通行时长数据和第i种天气类型的天气数据作为训练样本输出到所述基于SVR的预测模型,得到训练后的第 $i \times j$ 个基于SVR的预测模型;

步骤四:将测试周期的第一统计时段的所述第一通行时长数据和第i种天气类型的天气数据输出到所述训练后的第 $i \times j$ 个基于SVR的预测模型,得到所述训练后的第 $i \times j$ 个基于SVR的预测模型输出的测试周期的第二统计时段通行时长预测值;

步骤五:根据测试周期的第二统计时段通行时长预测值和测试周期的第二统计时段的所述第一通行时长数据,计算测试周期对应的平均绝对百分比误差MAPE值;

步骤六:再次执行步骤二至步骤五,直至得到所述M个子周期分别作为测试周期时对应的MAPE值,将所述M个子周期分别作为测试周期时对应的MAPE值的平均值,作为第i种天气类型对应的MAPE值;

步骤七:再次执行步骤一至步骤六,直至得到所述N种天气类型对应的MAPE值,将所述N种天气类型中对应的MAPE值最小的天气类型作为所述第一天气类型。

10. 根据权利要求8或9所述的装置,其特征在于,所述装置还包括:

预处理模块,用于将数据输入基于SVR的预测模型之前,采用数据使用范围为 $[0, 1]$ 的最小最大值归一化Min-Max-scaling算法将数据归一化,将归一化后的数据输入基于SVR的预测模型。

11. 根据权利要求7所述的装置,其特征在于,所述装置还包括:

数据检验模块,用于检验所述第一通行时长数据中是否有部分所述预设历史周期的数据缺失;

第三交互模块,用于当所述第一通行时长数据中有部分所述预设历史周期的数据缺失时,向网络侧发送与所述第一路段有重叠部分的路段的信息查询请求,接收网络侧发送的与所述第一路段有重叠部分的第二路段的信息,其中,所述第二路段为多个;

所述第一交互模块,还用于基于所述第二路段的信息,向网络侧发送部分所述预设历史周期内所述第一统计时段和第二统计时段的所述第二路段的通行时长的数据,接收网络侧发送的部分所述预设历史周期内所述第一统计时段和第二统计时段的所述第二路段的第三通行时长数据;

数据补充模块,用于根据所述第三通行时长数据中所述第二路段与所述第一路段的重叠部分对应的通行时长数据对所述第一通行时长数据中缺失的部分所述预设历史周期的数据进行补充。

12. 根据权利要求11所述的装置,其特征在于,所述数据检验模块,还用于在根据所述第三通行时长数据中所述第二路段与所述第一路段的重叠部分对应的通行时长数据对所述第一通行时长数据中缺失的部分所述预设历史周期的数据进行补充后,再次检验所述第一通行时长数据中是否有部分所述预设历史周期的数据缺失;

所述数据补充模块,还用于当所述第一通行时长数据中仍有部分所述预设历史周期的数据缺失时,根据所述第一通行时长数据获取缺失部分的相邻周期的通行时长数据,计算所述相邻周期的通行时长数据的线性平均值,根据所述相邻周期的通行时长数据的线性平均值补充缺失部分的通行时长数据。

13. 一种电子设备,其特征在于,所述电子设备包括:处理器和用于存储能够在处理器上运行的计算机程序的存储器;

其中,所述处理器用于运行所述计算机程序时,执行权利要求1至6任一项所述方法的步骤。

14. 一种计算机存储介质,其上存储有计算机程序,其特征在于,所述计算机程序被处理器执行时实现权利要求1至6任一项所述方法的步骤。

一种道路通行时间预测的方法、装置及存储介质

技术领域

[0001] 本发明涉及数据处理技术,尤其涉及一种道路通行时间预测的方法、装置及存储介质。

背景技术

[0002] 现有技术对交通流速进行预测,主要通过对一种来源或多种来源的交通流速数据进行缺失数据的线性填补,去除不合理数据等预处理后,再将数据划分为训练集和测试集,通过一些算法和训练集对交通流速进行建模,从而得出预测值。

[0003] 针对交通流速数据进行研究分析对交通流速进行预测,当交通流速数据的波动较大时,预测的准确度会明显下降。

发明内容

[0004] 本发明实施例提供一种道路通行时长预测的方法、装置及计算机存储介质。

[0005] 本发明实施例提供一种道路通行时长预测的方法,所述方法包括:

[0006] 接收到针对第一路段的第一统计时段的通行时长的预测请求;

[0007] 基于所述预测请求向网络侧发送预设历史周期内所述第一统计时段和第二统计时段的所述第一路段的通行时长的数据查询指令,接收网络侧发送的所述预设历史周期内所述第一统计时段和所述第二统计时段的所述第一路段的第一通行时长数据,其中,所述第二统计时段为与所述第一统计时段连续的前一时段;

[0008] 基于所述预测请求向网络侧发送所述预设历史周期内所述第一统计时段和所述第二统计时段的所述第一路段的第一天气类型的天气的数据查询指令,接收网络侧发送的所述预设历史周期内所述第一统计时段和所述第二统计时段的所述第一路段的所述第一天气类型的第一天气数据;

[0009] 将所述第一通行时长数据和所述第一天气数据作为训练样本输出到基于支持向量回归机SVR的预测模型,得到训练后的基于SVR的预测模型;

[0010] 基于所述预测请求向网络侧发送当前第二统计时段所述第一路段的通行时长的数据查询指令,接收网络侧发送的所述当前第二统计时段所述第一路段的第二通行时长数据;

[0011] 基于所述预测请求向网络侧发送所述当前第二统计时段所述第一路段的所述第一天气类型的天气的数据查询指令,接收网络侧发送的所述当前第二统计时段所述第一路段的所述第一天气类型的第二天气数据;

[0012] 将所述第二通行时长数据和所述第二天气数据输出到所述训练后的基于SVR的预测模型,得到所述训练后的基于SVR的预测模型输出的第三通行时长数据。

[0013] 上述方案中,基于所述预测请求向网络侧发送所述预设历史周期内所述第一统计时段和所述第二统计时段的所述第一路段的N种天气类型的天气的数据查询指令,接收网络侧发送的所述预设历史周期内所述第一统计时段和所述第二统计时段的所述第一路段

的N种天气类型的天气数据,其中,N为正整数;

[0014] 根据所述第一通行时长数据分别和所述N种天气类型的天气数据建立预测模型;验证建立的预测模型,确定验证结果与实际值偏差最小的预测模型对应的天气类型为所述第一天气类型。

[0015] 上述方案中,所述根据所述第一通行时长数据分别和所述N种天气类型的天气数据建立预测模型;验证建立的预测模型,确定验证结果与实际值偏差最小的预测模型对应的天气类型为所述第一天气类型,具体包括:

[0016] 将所述预设历史周期分为M个子周期,其中,M为正整数;并执行以下步骤:

[0017] 步骤一:选取第i种天气类型的天气数据,其中,i为正整数, $1 \leq i \leq N$;

[0018] 步骤二:选取第j个子周期作为测试周期,对应地将其余子周期作为训练周期,其中,j为正整数, $1 \leq j \leq M$;

[0019] 步骤三:将训练周期的所述第一通行时长数据和第i种天气类型的天气数据作为训练样本输出到所述基于SVR的预测模型,得到训练后的第 $i \times j$ 个基于SVR的预测模型;

[0020] 步骤四:将测试周期的第一统计时段的所述第一通行时长数据和第i种天气类型的天气数据输出到所述训练后的第 $i \times j$ 个基于SVR的预测模型,得到所述训练后的第 $i \times j$ 个基于SVR的预测模型输出的测试周期的第二统计时段通行时长预测值;

[0021] 步骤五:根据测试周期的第二统计时段通行时长预测值和测试周期的第二统计时段的所述第一通行时长数据,计算测试周期对应的平均绝对百分比误差MAPE值;

[0022] 步骤六:再次执行步骤二至步骤五,直至得到所述M个子周期分别作为测试周期时对应的MAPE值,将所述M个子周期分别作为测试周期时对应的MAPE值的平均值,作为第i种天气类型对应的MAPE值;

[0023] 步骤七:再次执行步骤一至步骤六,直至得到所述N种天气类型对应的MAPE值,将所述N中天气类型中对应的MAPE值最小的天气类型作为所述第一天气类型。

[0024] 上述方案中,所述方法还包括:

[0025] 将数据输入基于SVR的预测模型之前,采用数据使用范围为 $[0,1]$ 的最小最大值归一化Min-Max-scaling算法将数据归一化,将归一化后的数据输入基于SVR的预测模型。

[0026] 上述方案中,所述方法还包括:

[0027] 检验所述第一通行时长数据中是否有部分所述预设历史周期的数据缺失;

[0028] 当所述第一通行时长数据中有部分所述预设历史周期的数据缺失时,向网络侧发送与所述第一路段有重叠部分的路段的信息查询请求,接收网络侧发送的与所述第一路段有重叠部分的第二路段的信息,其中,所述第二路段为多个;

[0029] 基于所述第二路段的信息,向网络侧发送所述部分预设历史周期内所述第一统计时段和第二统计时段的所述第二路段的通行时长的数据,接收网络侧发送的所述部分预设历史周期内所述第一统计时段和第二统计时段的所述第二路段的第三通行时长数据;

[0030] 根据所述第三通行时长数据中所述第二路段与所述第一路段的重叠部分对应的通行时长数据对所述第一通行时长数据中缺失的所述部分所述预设历史周期的数据进行补充。

[0031] 上述方案中,所述方法还包括:

[0032] 在根据所述第三通行时长数据中所述第二路段与所述第一路段的重叠部分对应

的通行时长数据对所述第一通行时长数据中缺失的所述部分所述预设历史周期的数据进行补充后,再次检验所述第一通行时长数据中是否有部分所述预设历史周期的数据缺失;

[0033] 当所述第一通行时长数据中仍有部分所述预设历史周期的数据缺失时,根据所述第一通行时长数据获取缺失部分的相邻周期的通行时长数据,计算所述相邻周期的通行时长数据的线性平均值,根据所述相邻周期的通行时长数据的线性平均值补充缺失部分的通行时长数据。

[0034] 本发明实施例提供一种道路通行时长预测的装置,所述装置包括:

[0035] 请求模块,用于接收针对第一路段的第一统计时段的通行时长的预测请求;

[0036] 第一交互模块,用于基于所述预测请求向网络侧发送预设历史周期内所述第一统计时段和第二统计时段的所述第一路段的通行时长的数据查询指令,接收网络侧发送的所述预设历史周期内所述第一统计时段和所述第二统计时段的所述第一路段的第一通行时长数据,其中,所述第二统计时段为与所述第一统计时段连续的前一时段;

[0037] 第二交互模块,用于基于所述预测请求向网络侧发送所述预设历史周期内所述第一统计时段和所述第二统计时段的所述第一路段的第一天气类型的天气的数据查询指令,接收网络侧发送的所述预设历史周期内所述第一统计时段和所述第二统计时段的所述第一路段的所述第一天气类型的第一天气数据;

[0038] 训练模块,用于将所述第一通行时长数据和所述第一天气数据作为训练样本输出到基于支持向量回归机SVR的预测模型,得到训练后的基于SVR的预测模型;

[0039] 所述第一交互模块,还用于基于所述预测请求向网络侧发送当前第二统计时段所述第一路段的通行时长的数据查询指令,接收网络侧发送的所述当前第二统计时段所述第一路段的第二通行时长数据;

[0040] 所述第二交互模块,还用于基于所述预测请求向网络侧发送所述当前第二统计时段所述第一路段的所述第一天气类型的天气的数据查询指令,接收网络侧发送的所述当前第二统计时段所述第一路段的所述第一天气类型的第二天气数据;

[0041] 预测模块,用于将所述第二通行时长数据和所述第二天气数据输出到所述训练后的基于SVR的预测模型,得到所述训练后的基于SVR的预测模型输出的第三通行时长数据。

[0042] 上述方案中,所述第二交互模块,还用于基于所述预测请求向网络侧发送所述预设历史周期内所述第一统计时段和所述第二统计时段的所述第一路段的N种天气类型的天气的数据查询指令,接收网络侧发送的所述预设历史周期内所述第一统计时段和所述第二统计时段的所述第一路段的N种天气类型的天气数据,其中,N为正整数;

[0043] 所述装置还包括:天气类型选择模块;

[0044] 所述天气类型选择模块,用于根据所述第一通行时长数据分别和所述N种天气类型的天气数据建立预测模型;验证建立的预测模型,确定验证结果与实际值偏差最小的预测模型对应的天气类型为所述第一天气类型。

[0045] 上述方案中,所述天气类型选择模块,具体用于:

[0046] 将所述预设历史周期分为M个子周期,其中,M为正整数;并执行以下步骤:

[0047] 步骤一:选取第i种天气类型的天气数据,其中,i为正整数, $1 \leq i \leq N$;

[0048] 步骤二:选取第j个子周期作为测试周期,对应地将其余子周期作为训练周期,其中,j为正整数, $1 \leq j \leq M$;

[0049] 步骤三:将训练周期的所述第一通行时长数据和第*i*种天气类型的天气数据作为训练样本输出到所述基于SVR的预测模型,得到训练后的第*i*×*j*个基于SVR的预测模型;

[0050] 步骤四:将测试周期的第一统计时段的所述第一通行时长数据和第*i*种天气类型的天气数据输出到所述训练后的第*i*×*j*个基于SVR的预测模型,得到所述训练后的第*i*×*j*个基于SVR的预测模型输出的测试周期的第二统计时段通行时长预测值;

[0051] 步骤五:根据测试周期的第二统计时段通行时长预测值和测试周期的第二统计时段的所述第一通行时长数据,计算测试周期对应的平均绝对百分比误差MAPE值;

[0052] 步骤六:再次执行步骤二至步骤五,直至得到所述*M*个子周期分别作为测试周期时对应的MAPE值,将所述*M*个子周期分别作为测试周期时对应的MAPE值的平均值,作为第*i*种天气类型对应的MAPE值;

[0053] 步骤七:再次执行步骤一至步骤六,直至得到所述*N*种天气类型对应的MAPE值,将所述*N*中天气类型中对应的MAPE值最小的天气类型作为所述第一天气类型。

[0054] 上述方案中,所述装置还包括:

[0055] 预处理模块,用于将数据输入基于SVR的预测模型之前,采用数据使用范围为[0, 1]的最小最大值归一化Min-Max-scaling算法将数据归一化,将归一化后的数据输入基于SVR的预测模型。

[0056] 上述方案中,所述装置还包括:

[0057] 数据检验模块,用于检验所述第一通行时长数据中是否有部分所述预设历史周期的数据缺失;

[0058] 第三交互模块,用于当所述第一通行时长数据中有部分所述预设历史周期的数据缺失时,向网络侧发送与所述第一路段有重叠部分的路段的信息查询请求,接收网络侧发送的与所述第一路段有重叠部分的第二路段的信息,其中,所述第二路段为多个;

[0059] 所述第一交互模块,还用于基于所述第二路段的信息,向网络侧发送所述部分预设历史周期内所述第一统计时段和第二统计时段的所述第二路段的通行时长的数据,接收网络侧发送的所述部分预设历史周期内所述第一统计时段和第二统计时段的所述第二路段的第三通行时长数据;

[0060] 数据补充模块,用于根据所述第三通行时长数据中所述第二路段与所述第一路段的重叠部分对应的通行时长数据对所述第一通行时长数据中缺失的所述部分所述预设历史周期的数据进行补充。

[0061] 上述方案中,所述数据检验模块,还用于在根据所述第三通行时长数据中所述第二路段与所述第一路段的重叠部分对应的通行时长数据对所述第一通行时长数据中缺失的所述部分所述预设历史周期的数据进行补充后,再次检验所述第一通行时长数据中是否有部分所述预设历史周期的数据缺失;

[0062] 所述数据补充模块,还用于当所述第一通行时长数据中仍有部分所述预设历史周期的数据缺失时,根据所述第一通行时长数据获取缺失部分的相邻周期的通行时长数据,计算所述相邻周期的通行时长数据的线性平均值,根据所述相邻周期的通行时长数据的线性平均值补充缺失部分的通行时长数据。

[0063] 本发明实施例提供一种道路通行时长预测的装置,所述装置包括:处理器和用于存储能够在处理器上运行的计算机程序的存储器;

[0064] 其中,所述处理器用于运行所述计算机程序时,执行上述任一道路通行时长预测的方法的步骤。

[0065] 本发明实施例提供一种计算机存储介质,其上存储有计算机程序,所述计算机程序被处理器执行时实现上述任一道路通行时长预测的方法的步骤。

[0066] 本发明实施例的技术方案中,通过接收针对第一路段的第一统计时段的通行时长的预测请求,基于所述预测请求通过网络侧查询到预设历史周期内所述第一统计时段和所述第二统计时段的所述第一路段的第一通行时长数据和对应的所述第一天气类型的第一天气数据,将所述第一通行时长数据和所述第一天气数据作为训练样本输出到基于SVR的预测模型,得到训练后的基于SVR的预测模型,基于所述预测请求通过网络侧查询到所述当前第二统计时段所述第一路段的第二通行时长数据和对应的所述第一天气类型的第二天气数据,将所述第二通行时长数据和所述第二天气数据输出到所述训练后的基于SVR的预测模型,得到所述训练后的基于SVR的预测模型输出的第三通行时长数据;实现了结合通行时长数据和第一天气类型的天气数据对基于SVR的预测模型训练,利用天气数据波动对通行时长数据波动的影响,减少训练后的基于SVR的预测模型中由于天气数据波动对预测结果的准确度的影响。通过第一天气类型的天气数据对通行时长数据的波动进行修正,降低通行时长数据的波动对预测结果产生的误差,提高通行时长预测的准确度

附图说明

[0067] 附图以示例而非限制的方式大体示出了本文中所讨论的各个实施例。

[0068] 图1为本发明实施例一种道路通行时长预测的方法的流程示意图;

[0069] 图2为本发明实施例一种道路通行时长预测的方法的流程示意图;

[0070] 图3为本发明实施例一种道路通行时长预测的装置的结构示意图;

[0071] 图4为本发明实施例一种道路通行时长预测的装置的结构示意图;

[0072] 图5为本发明实施例中某高速公路部分路段的交通路线图;

[0073] 图6为本发明实施例中通过交叉验证选择最优天气指标的流程示意图;

[0074] 图7为本发明实施例提供的基于硬件实现的道路通行时长预测的设备的结构示意图。

具体实施方式

[0075] 为了能够更加详尽地了解本发明实施例的特点与技术内容,下面结合附图对本发明实施例的实现进行详细阐述,所附附图仅供参考说明之用,并非用来限定本发明实施例。

[0076] 图1为本发明实施例一种道路通行时长预测的方法的流程示意图,如图1所示,本发明实施例的道路通行时长预测的方法包括以下步骤:

[0077] 步骤101,接收到针对第一路段的第一统计时段的通行时长的预测请求。

[0078] 在一些实施例中,所述第一路段包括道路系统中连接指定起点到指定终点的路段,例如:连接某一个高速公路入口到下一个高速公路出口的高速公路路段、一条公路沿途两个公里标之间的路段等等。

[0079] 在一些实施例中,所述接收到针对第一路段的第一统计时段的通行时长的预测请求,具体包括:接收到针对第一路段的第一统计时段中每个窗口时段的平均通行时长的预

测请求。

[0080] 在一些实施例中,所述第一统计时段,可以根据需求进行预设,例如:早高峰时段8:00-10:00、晚高峰时段17:00-19:00等。

[0081] 窗口时段是根据预测的需求来设置时长的时间段,例如:以20分钟为一个窗口时段,则早高峰时段8:00-10:00可以划分为8:00-8:20、8:20-8:40、8:40-9:00、9:00-9:20、9:20-9:40、9:40-10:00六个窗口时段;这里只是举例说明,并非对本发明的限制,在实际应用中,窗口时段的时长可以根据预测的需求确定。

[0082] 将某个窗口时段内通过所述第一路段的全部车辆的平均通行时长作为该窗口时段的平均通行时长。

[0083] 步骤102,基于所述预测请求向网络侧发送预设历史周期内所述第一统计时段和第二统计时段的所述第一路段的通行时长的数据查询指令,接收网络侧发送的所述预设历史周期内所述第一统计时段和所述第二统计时段的所述第一路段的第一通行时长数据,其中,所述第二统计时段为与所述第一统计时段连续的前一时段。

[0084] 在一些实施例中,网络侧包括:交通综合信息平台、交通数据服务器或其它能够查询到道路通行时长历史数据的设备或系统。

[0085] 在一些实施例中,所述基于所述预测请求向网络侧发送预设历史周期内所述第一统计时段和第二统计时段的所述第一路段的通行时长的数据查询指令,具体包括:基于所述预测请求向网络侧发送预设历史周期内所述第一统计时段和第二统计时段中每个窗口时段的平均通行时长的所述第一路段的通行时长的数据查询指令。

[0086] 在一些实施例中,所述第二统计时段根据所述第一统计时段确定,例如:当第一统计时段为8:00-10:00,对应第二统计时段为6:00-8:00;以20分钟为一个窗口时段,对应地,第二统计时段可以划分为6:00-6:20、6:20-6:40、6:40-7:00、7:00-7:20、7:20-7:40、7:40-8:00六个窗口时段。

[0087] 所述接收网络侧发送的所述预设历史周期内所述第一统计时段和所述第二统计时段的所述第一路段的第一通行时长数据,具体包括:接收网络侧发送的所述预设历史周期内所述第一统计时段和所述第二统计时段中每个窗口时段的所述第一路段的第一通行时长数据。

[0088] 步骤103,基于所述预测请求向网络侧发送所述预设历史周期内所述第一统计时段和所述第二统计时段的所述第一路段的第一天气类型的天气的数据查询指令,接收网络侧发送的所述预设历史周期内所述第一统计时段和所述第二统计时段的所述第一路段的所述第一天气类型的第一天气数据。

[0089] 在一些实施例中,网络侧包括:天气信息服务平台、天气数据服务器或其它能够查询到所述第一路段的预设历史周期内所述第一统计时段和所述第二统计时段的天气数据的设备或系统。

[0090] 在一些实施例中,第一天气类型包括:气压、海平面气压、风向、风速、气温、相对湿度或降雨量。在不同的天气检测系统或不同的气候地区,用于描述天气数据的天气类型会有较大差异,这里只是对本发明一些实施例中采用到的天气类型进行举例说明,而非对于本发明的限制,在实际应用中,第一天气类型还可以包括其它种类。

[0091] 在一些实施例中,第一路段的天气数据采用第一路段所在地理区域的天气数据。

[0092] 在实际应用中,天气数据通常是以固定频率间隔的时间点来记录的,因此,在一些实施例中,需要对天气数据进行分析及处理,采用与第一统计时段和第二统计时段最接近的时间点的天气数据作为对应统计时段的天气数据,例如:第一路段所在地区的天气数据是每隔3个小时做一次记录的,即每天记录0点、3点、6点、9点、12点、15点、18点、21点的天气数据,则可以采用6点记录的天气数据代表6:00-8:00的天气数据,采用9点记录的天气数据代表8:00-10:00的天气数据。

[0093] 步骤104,将所述第一通行时长数据和所述第一天气数据作为训练样本输出到基于支持向量回归机SVR的预测模型,得到训练后的基于SVR的预测模型。

[0094] 在一些实施例中,所述将所述第一通行时长数据和所述第一天气数据作为训练样本输出到基于支持向量回归机SVR的预测模型,得到训练后的基于SVR的预测模型,具体包括:

[0095] 所述基于SVR的预测模型的公式为: $f(x) = \omega^T \phi(x) + b$,其中, ω 和 b 为SVR算法的参数, ω^T 表示 ω 的转置,函数 $\phi(x)$ 表示输入特征,函数 $f(x)$ 表示输出标识。

[0096] 所述将所述第一通行时长数据和所述第一天气数据作为训练样本,包括:将所述第一通行时长数据和所述第一天气数据中对应第一统计时段的数据作为训练样本的输出标识,将所述第一通行时长数据和所述第一天气数据中对应第二统计时段的数据作为训练样本的输入特征。

[0097] 步骤105,基于所述预测请求向网络侧发送当前第二统计时段所述第一路段的通行时长的数据查询指令,接收网络侧发送的所述当前第二统计时段所述第一路段的第二通行时长数据。

[0098] 在一些实施例中,所述基于所述预测请求向网络侧发送当前第二统计时段所述第一路段的通行时长的数据查询指令,接收网络侧发送的所述当前第二统计时段所述第一路段的第二通行时长数据,具体包括:基于所述预测请求向网络侧发送当前第二统计时段中每个窗口时段的平均时长的所述第一路段的通行时长的数据查询指令,接收网络侧发送的所述当前第二统计时段中每个窗口时段的平均时长的所述第一路段的第二通行时长数据。

[0099] 步骤106,基于所述预测请求向网络侧发送所述当前第二统计时段所述第一路段的所述第一天气类型的天气的数据查询指令,接收网络侧发送的所述当前第二统计时段所述第一路段的所述第一天气类型的第二天气数据。

[0100] 步骤107,将所述第二通行时长数据和所述第二天气数据输出到所述训练后的基于SVR的预测模型,得到所述训练后的基于SVR的预测模型输出的第三通行时长数据。

[0101] 其中,所述第三通行时长数据为所述第一路段的第一统计时段的通行时长的预测结果。

[0102] 在一些实施例中,所述将所述第二通行时长数据和所述第二天气数据输出到所述训练后的基于SVR的预测模型,得到所述训练后的基于SVR的预测模型输出的第三通行时长数据,具体包括:将所述第二通行时长数据和所述第二天气数据作为输入特征,输出到所述训练后的基于SVR的预测模型,得到所述训练后的基于SVR的预测模型输出的第三通行时长数据。

[0103] 在一些实施例中,将数据输入基于SVR的预测模型之前,采用数据使用范围为[0, 1]的最小最大值归一化Min-Max-scaling算法将数据归一化,将归一化后的数据输入基于

SVR的预测模型。

[0104] 具体地,Min-Max-scaling算法的公式为 $Z = lb + \frac{X - \min(X)}{\max(X) - \min(X)}(ub - lb)$, 其中,

Z为归一化后的变量值,X为实际变量值,max(X)为实际变量的最大值,min(X)为实际变量的最小值,lb是数据使用范围的低边界值,ub是数据使用范围的高边界值。则采用数据使用范围

为[0,1]的Min-Max-scaling算法的公式,可以表示为 $Z = \frac{X - \min(X)}{\max(X) - \min(X)}$ 。

[0105] 对应地,将训练后的基于SVR的预测模型输出的预测结果数据利用公式

$Z = \frac{X - \min(X)}{\max(X) - \min(X)}$ 进行反归一化,得到预测结果的实际变量值。

[0106] 在一些实施例中,上述方法还包括:

[0107] 检验所述第一通行时长数据中是否有部分所述预设历史周期的数据缺失。

[0108] 当所述第一通行时长数据中有部分所述预设历史周期的数据缺失时,向网络侧发送与所述第一路段有重叠部分的路段的信息查询请求,接收网络侧发送的与所述第一路段有重叠部分的第二路段的信息,其中,所述第二路段为多个。

[0109] 基于所述第二路段的信息,向网络侧发送所述部分预设历史周期内所述第一统计时段和第二统计时段的所述第二路段的通行时长的数据,接收网络侧发送的所述部分预设历史周期内所述第一统计时段和第二统计时段的所述第二路段的第三通行时长数据。

[0110] 根据所述第三通行时长数据中所述第二路段与所述第一路段的重叠部分对应的通行时长数据对所述第一通行时长数据中缺失的所述部分所述预设历史周期的数据进行补充。

[0111] 在一些实施例中,上述方法还包括:

[0112] 在根据所述第三通行时长数据中所述第二路段与所述第一路段的重叠部分对应的通行时长数据对所述第一通行时长数据中缺失的所述部分所述预设历史周期的数据进行补充后,再次检验所述第一通行时长数据中是否有部分所述预设历史周期的数据缺失。

[0113] 当所述第一通行时长数据中仍有部分所述预设历史周期的数据缺失时,根据所述第一通行时长数据获取缺失部分的相邻周期的通行时长数据,计算所述相邻周期的通行时长数据的线性平均值,根据所述相邻周期的通行时长数据的线性平均值补充缺失部分的通行时长数据。

[0114] 图2为本发明实施例一种道路通行时长预测的方法的流程示意图,如图2所示,本发明实施例的道路通行时长预测的方法包括以下步骤:

[0115] 步骤201,接收到针对第一路段的第一统计时段的通行时长的预测请求。

[0116] 在一些实施例中,所述第一路段包括道路系统中连接指定起点到指定终点的路段,例如:连接某一个高速公路入口到下一个高速公路出口的高速公路路段、一条公路沿途两个公里标之间的路段等等。

[0117] 在一些实施例中,所述接收到针对第一路段的第一统计时段的通行时长的预测请求,包括:接收到针对第一路段的第一统计时段中每个窗口时段的平均通行时长的预测请求。

[0118] 在一些实施例中,所述第一统计时段,可以根据需求进行预设,例如:早高峰时段8:00-10:00、晚高峰时段17:00-19:00等。

[0119] 窗口时段是根据预测的需求来设置时长的时间段,例如:以20分钟为一个窗口时段,则早高峰时段8:00-10:00可以划分为8:00-8:20、8:20-8:40、8:40-9:00、9:00-9:20、9:20-9:40、9:40-10:00六个窗口时段;这里只是举例说明,并非对本发明的限制,在实际应用中,窗口时段的时长可以根据预测的需求确定。

[0120] 将某个窗口时段内通过所述第一路段的全部车辆的平均通行时长作为该窗口时段的平均通行时长。

[0121] 步骤202,基于所述预测请求向网络侧发送预设历史周期内所述第一统计时段和所述第二统计时段的所述第一路段的通行时长的数据查询指令,接收网络侧发送的所述预设历史周期内所述第一统计时段和所述第二统计时段的所述第一路段的第一通行时长数据,其中,所述第二统计时段为与所述第一统计时段连续的前一时段。

[0122] 在一些实施例中,网络侧包括:交通综合信息平台、交通数据服务器或其它能够查询到道路通行时长历史数据的设备或系统。

[0123] 在一些实施例中,所述基于所述预测请求向网络侧发送预设历史周期内所述第一统计时段和所述第二统计时段的所述第一路段的通行时长的数据查询指令,包括:基于所述预测请求向网络侧发送预设历史周期内所述第一统计时段和所述第二统计时段中每个窗口时段的平均通行时长的所述第一路段的通行时长的数据查询指令;

[0124] 在一些实施例中,所述第二统计时段根据所述第一统计时段确定,例如:当第一统计时段为8:00-10:00,对应第二统计时段为6:00-8:00;以20分钟为一个窗口时段,对应地,第二统计时段可以划分为6:00-6:20、6:20-6:40、6:40-7:00、7:00-7:20、7:20-7:40、7:40-8:00六个窗口时段。

[0125] 所述接收网络侧发送的所述预设历史周期内所述第一统计时段和所述第二统计时段的所述第一路段的第一通行时长数据,包括:接收网络侧发送的所述预设历史周期内所述第一统计时段和所述第二统计时段中每个窗口时段的所述第一路段的第一通行时长数据。

[0126] 步骤203,基于所述预测请求向网络侧发送所述预设历史周期内所述第一统计时段和所述第二统计时段的所述第一路段的N种天气类型的天气的数据查询指令,接收网络侧发送的所述预设历史周期内所述第一统计时段和所述第二统计时段的所述第一路段的N种天气类型的天气数据,其中,N为正整数。

[0127] 在一些实施例中,网络侧包括:天气信息服务平台、天气数据服务器或其它能够查询到所述第一路段的预设历史周期内所述第一统计时段和所述第二统计时段的天气数据的设备或系统。

[0128] 在一些实施例中,所述N种天气类型包括:气压、海平面气压、风向、风速、气温、相对湿度和降雨量。在不同的天气检测系统或不同的气候地区,用于描述天气数据的天气类型会有较大差异,这里只是对本发明一些实施例中所采用到的天气类型进行举例说明,而非对于本发明的限制,在实际应用中,所述N种天气类型还可以包括其它种类。

[0129] 在一些实施例中,第一路段的天气数据采用第一路段所在地理区域的天气数据。

[0130] 在实际应用中,天气数据通常是以固定频率间隔的时间点来记录的,因此,在一些

实施例中,需要对天气数据进行分析及处理,采用与第一统计时段和第二统计时段最接近的时间点的天气数据作为对应统计时段的天气数据,例如:第一路段所在地区的天气数据是每隔3个小时做一次记录的,即每天记录0点、3点、6点、9点、12点、15点、18点、21点的天气数据,则可以采用6点记录的天气数据代表6:00-8:00的天气数据,采用9点记录的天气数据代表8:00-10:00的天气数据。

[0131] 步骤204,根据所述第一通行时长数据分别和所述N种天气类型的天气数据建立预测模型;验证建立的预测模型,确定验证结果与实际值偏差最小的预测模型对应的天气类型为所述第一天气类型。

[0132] 在一些实施例中,所述根据所述第一通行时长数据分别和所述N种天气类型的天气数据建立预测模型;验证建立的预测模型,确定验证结果与实际值偏差最小的预测模型对应的天气类型为所述第一天气类型,具体包括:

[0133] 将所述预设历史周期分为M个子周期,其中,M为正整数;并执行以下步骤:

[0134] 步骤a:选取第i种天气类型的天气数据,其中,i为正整数, $1 \leq i \leq N$ 。

[0135] 步骤b:选取第j个子周期作为测试周期,对应地将其余子周期作为训练周期,其中,j为正整数, $1 \leq j \leq M$ 。

[0136] 在一些实施例中,为了便于统计,每个子周期为一个星期;在实际应用中,子周期的天数可以根据实际需要进行确定,此处仅为举例说明,而非对于本发明的限制。

[0137] 步骤c:将训练周期的所述第一通行时长数据和第i种天气类型的天气数据作为训练样本输出到所述基于SVR的预测模型,得到训练后的第 $i \times j$ 个基于SVR的预测模型。

[0138] 其中,第 $i \times j$ 个基于SVR的预测模型表示采用第i种天气类型和第j个子周期作为测试集时,对应的训练后的基于SVR的预测模型。

[0139] 所述基于SVR的预测模型的公式为: $f(x) = \omega^T \phi(x) + b$,其中, ω 和b为SVR算法的参数, ω^T 表示 ω 的转置,函数 $\phi(x)$ 表示输入特征,函数 $f(x)$ 表示输出标识。

[0140] 在一些实施例中,所述将训练周期的所述第一通行时长数据和第i种天气类型的天气数据作为训练样本输出到所述基于SVR的预测模型,具体包括:将训练周期的所述第一通行时长数据和第i种天气类型的天气数据中对应第一统计时段的数据作为训练样本的输出标识,将训练周期的所述第一通行时长数据和第i种天气类型的天气数据中对应第二统计时段的数据作为训练样本的输入特征。

[0141] 步骤d:将测试周期的第一统计时段的所述第一通行时长数据和第i种天气类型的天气数据输出到所述训练后的第 $i \times j$ 个基于SVR的预测模型,得到所述训练后的第 $i \times j$ 个基于SVR的预测模型输出的测试周期的第二统计时段通行时长预测值。

[0142] 在一些实施例中,所述将测试周期的第一统计时段的所述第一通行时长数据和第i种天气类型的天气数据输出到所述训练后的第 $i \times j$ 个基于SVR的预测模型,得到所述训练后的第 $i \times j$ 个基于SVR的预测模型输出的测试周期的第二统计时段通行时长预测值,具体包括:将测试周期的第一统计时段的所述第一通行时长数据和第i种天气类型的天气数据作为输入特征,输出到所述训练后的第 $i \times j$ 个基于SVR的预测模型,得到所述训练后的第 $i \times j$ 个基于SVR的预测模型输出的预测标识测试周期的第二统计时段通行时长预测值。

[0143] 步骤e:根据测试周期的第二统计时段通行时长预测值和测试周期的第二统计时段的所述第一通行时长数据,计算测试周期对应的平均绝对百分比误差MAPE值。

[0144] 在一些实施例中,所述根据测试周期的第二统计时段通行时长预测值和测试周期的第二统计时段的所述第一通行时长数据,计算测试周期对应的平均绝对百分比误差MAPE值,具体包括:

[0145] MAPE值的计算公式为 $MAPE = \frac{1}{T} \sum_{t=1}^T \left| \frac{d_t - p_t}{d_t} \right|$,其中,T为测试集的数据周期总数,t

为测试周期的数据周期编号, d_t 为真实值, p_t 为预测值。

[0146] 将测试周期的第二统计时段通行时长预测值作为预测值,将测试周期的第二统计时段的所述第一通行时长数据作为真实值,输入所述MAPE值的计算公式,得到测试周期对应的MAPE值。

[0147] 步骤f:再次执行步骤b至步骤e,直至得到所述M个子周期分别作为测试周期时对应的MAPE值,将所述M个子周期分别作为测试周期时对应的MAPE值的平均值,作为第i种天气类型对应的MAPE值。

[0148] 步骤g:再次执行步骤a至步骤f,直至得到所述N种天气类型对应的MAPE值,将所述N种天气类型中对应的MAPE值最小的天气类型作为所述第一天气类型。

[0149] 步骤205,基于所述预测请求向网络侧发送所述预设历史周期内所述第一统计时段和所述第二统计时段的所述第一路段的第一天气类型的天气的数据查询指令,接收网络侧发送的所述预设历史周期内所述第一统计时段和所述第二统计时段的所述第一路段的所述第一天气类型的第一天气数据。

[0150] 步骤206,将所述第一通行时长数据和所述第一天气数据作为训练样本输出到基于支持向量回归机SVR的预测模型,得到训练后的基于SVR的预测模型。

[0151] 在一些实施例中,所述将所述第一通行时长数据和所述第一天气数据作为训练样本输出到基于支持向量回归机SVR的预测模型,得到训练后的基于SVR的预测模型,具体包括:

[0152] 所述基于SVR的预测模型的公式为: $f(x) = \omega^T \phi(x) + b$,其中, ω 和**b**为SVR算法的参数, ω^T 表示 ω 的转置,函数 $\phi(x)$ 表示输入特征,函数 $f(x)$ 表示输出标识。

[0153] 所述将所述第一通行时长数据和所述第一天气数据作为训练样本,包括:将所述第一通行时长数据和所述第一天气数据中对应第一统计时段的数据作为训练样本的输出标识,将所述第一通行时长数据和所述第一天气数据中对应第二统计时段的数据作为训练样本的输入特征。

[0154] 步骤207,基于所述预测请求向网络侧发送当前第二统计时段所述第一路段的通行时长的数据查询指令,接收网络侧发送的所述当前第二统计时段所述第一路段的第二通行时长数据。

[0155] 在一些实施例中,所述基于所述预测请求向网络侧发送当前第二统计时段所述第一路段的通行时长的数据查询指令,接收网络侧发送的所述当前第二统计时段所述第一路段的第二通行时长数据,具体包括:基于所述预测请求向网络侧发送当前第二统计时段中每个窗口时段的平均时长的所述第一路段的通行时长的数据查询指令,接收网络侧发送的所述当前第二统计时段中每个窗口时段的平均时长的所述第一路段的第二通行时长数据。

[0156] 步骤208,基于所述预测请求向网络侧发送所述当前第二统计时段所述第一路段的所述第一天气类型的天气的数据查询指令,接收网络侧发送的所述当前第二统计时段所

述第一路段的所述第一天气类型的第二天气数据。

[0157] 步骤209,将所述第二通行时长数据和所述第二天气数据输出到所述训练后的基于SVR的预测模型,得到所述训练后的基于SVR的预测模型输出的第三通行时长数据。

[0158] 其中,所述第三通行时长数据为所述第一路段的第一统计时段的通行时长的预测结果。

[0159] 在一些实施例中,所述将所述第二通行时长数据和所述第二天气数据输出到所述训练后的基于SVR的预测模型,得到所述训练后的基于SVR的预测模型输出的第三通行时长数据,具体包括:将所述第二通行时长数据和所述第二天气数据作为输入特征,输出到所述训练后的基于SVR的预测模型,得到所述训练后的基于SVR的预测模型输出的第三通行时长数据。

[0160] 在一些实施例中,将数据输入基于SVR的预测模型之前,采用数据使用范围为[0, 1]的最小最大值归一化Min-Max-scaling算法将数据归一化,将归一化后的数据输入基于SVR的预测模型。

[0161] 具体地,Min-Max-scaling算法的公式为 $Z = lb + \frac{X - \min(X)}{\max(X) - \min(X)}(ub - lb)$,其中,

Z为归一化后的变量值,X为实际变量值,max(X)为实际变量的最大值,min(X)为实际变量的最小值,lb是数据使用范围的低边界值,ub是数据使用范围的高边界值。则采用数据使用范围

为[0,1]的Min-Max-scaling算法的公式,可以表示为 $Z = \frac{X - \min(X)}{\max(X) - \min(X)}$ 。

[0162] 对应地,将训练后的基于SVR的预测模型输出的预测结果数据利用公式

$Z = \frac{X - \min(X)}{\max(X) - \min(X)}$ 进行反归一化,得到预测结果的实际变量值。

[0163] 在一些实施例中,上述方法还包括:

[0164] 检验所述第一通行时长数据中是否有部分所述预设历史周期的数据缺失。

[0165] 当所述第一通行时长数据中有部分所述预设历史周期的数据缺失时,向网络侧发送与所述第一路段有重叠部分的路段的信息查询请求,接收网络侧发送的与所述第一路段有重叠部分的第二路段的信息,其中,所述第二路段为多个。

[0166] 基于所述第二路段的信息,向网络侧发送所述部分预设历史周期内所述第一统计时段和第二统计时段的所述第二路段的通行时长的数据,接收网络侧发送的所述部分预设历史周期内所述第一统计时段和第二统计时段的所述第二路段的第三通行时长数据。

[0167] 根据所述第三通行时长数据中所述第二路段与所述第一路段的重叠部分对应的通行时长数据对所述第一通行时长数据中缺失的所述部分所述预设历史周期的数据进行补充。

[0168] 在一些实施例中,上述方法还包括:

[0169] 在根据所述第三通行时长数据中所述第二路段与所述第一路段的重叠部分对应的通行时长数据对所述第一通行时长数据中缺失的所述部分所述预设历史周期的数据进行补充后,再次检验所述第一通行时长数据中是否有部分所述预设历史周期的数据缺失。

[0170] 当所述第一通行时长数据中仍有部分所述预设历史周期的数据缺失时,根据所述

第一通行时长数据获取缺失部分的相邻周期的通行时长数据,计算所述相邻周期的通行时长数据的线性平均值,根据所述相邻周期的通行时长数据的线性平均值补充缺失部分的通行时长数据。

[0171] 图3为本发明实施例一种道路通行时长预测的装置的结构示意图,如图3所示,本发明实施例的道路通行时长预测的装置包括以下模块:

[0172] 请求模块301,用于接收针对第一路段的第一统计时段的通行时长的预测请求。

[0173] 在一些实施例中,所述第一路段包括道路系统中连接指定起点到指定终点的路段,例如:连接某一个高速公路入口到下一个高速公路出口的高速公路路段、一条公路沿途两个公里标之间的路段等等。

[0174] 在一些实施例中,所述请求模块301,具体用于:接收针对第一路段的第一统计时段中每个窗口时段的平均通行时长的预测请求。

[0175] 在一些实施例中,所述第一统计时段,可以根据需求进行预设,例如:早高峰时段8:00-10:00、晚高峰时段17:00-19:00等。

[0176] 窗口时段是根据预测的需求来设置时长的时间段,例如:以20分钟为一个窗口时段,则早高峰时段8:00-10:00可以划分为8:00-8:20、8:20-8:40、8:40-9:00、9:00-9:20、9:20-9:40、9:40-10:00六个窗口时段;这里只是举例说明,并非对本发明的限制,在实际应用中,窗口时段的时长可以根据预测的需求确定。

[0177] 将某个窗口时段内通过所述第一路段的全部车辆的平均通行时长作为该窗口时段的平均通行时长。

[0178] 第一交互模块302,用于基于所述预测请求向网络侧发送预设历史周期内所述第一统计时段和第二统计时段的所述第一路段的通行时长的数据查询指令,接收网络侧发送的所述预设历史周期内所述第一统计时段和所述第二统计时段的所述第一路段的第一通行时长数据,其中,所述第二统计时段为与所述第一统计时段连续的前一时段。

[0179] 在一些实施例中,网络侧包括:交通综合信息平台、交通数据服务器或其它能够查询到道路通行时长历史数据的设备或系统。

[0180] 在一些实施例中,所述第一交互模块302,具体用于:基于所述预测请求向网络侧发送预设历史周期内所述第一统计时段和第二统计时段中每个窗口时段的平均通行时长的所述第一路段的通行时长的数据查询指令;

[0181] 在一些实施例中,所述第二统计时段根据所述第一统计时段确定,例如:当第一统计时段为8:00-10:00,对应第二统计时段为6:00-8:00;以20分钟为一个窗口时段,对应地,第二统计时段可以划分为6:00-6:20、6:20-6:40、6:40-7:00、7:00-7:20、7:20-7:40、7:40-8:00六个窗口时段。

[0182] 所述第一交互模块302,还具体用于:接收网络侧发送的所述预设历史周期内所述第一统计时段和所述第二统计时段中每个窗口时段的所述第一路段的第一通行时长数据。

[0183] 第二交互模块303,用于基于所述预测请求向网络侧发送所述预设历史周期内所述第一统计时段和所述第二统计时段的所述第一路段的第一天气类型的天气的数据查询指令,接收网络侧发送的所述预设历史周期内所述第一统计时段和所述第二统计时段的所述第一路段的所述第一天气类型的第一天气数据。

[0184] 在一些实施例中,网络侧包括:天气信息服务平台、天气数据服务器或其它能够查

询到所述第一路段的预设历史周期内所述第一统计时段和所述第二统计时段的天气数据的设备或系统。

[0185] 在一些实施例中,第一天气类型包括:气压、海平面气压、风向、风速、气温、相对湿度或降雨量。在不同的天气检测系统或不同的气候地区,用于描述天气数据的天气类型会有较大差异,这里只是对本发明一些实施例中所采用到的天气类型进行举例说明,而非对于本发明的限制,在实际应用中,第一天气类型还可以包括其它种类。

[0186] 在一些实施例中,第一路段的天气数据采用第一路段所在地理区域的天气数据。

[0187] 在实际应用中,天气数据通常是以固定频率间隔的时间点来记录的,因此,在一些实施例中,需要对天气数据进行分析及处理,采用与第一统计时段和第二统计时段最接近的时间点的天气数据作为对应统计时段的天气数据,例如:第一路段所在地区的天气数据是每隔3个小时做一次记录的,即每天记录0点、3点、6点、9点、12点、15点、18点、21点的天气数据,则可以采用6点记录的天气数据代表6:00-8:00的天气数据,采用9点记录的天气数据代表8:00-10:00的天气数据。

[0188] 训练模块304,用于将所述第一通行时长数据和所述第一天气数据作为训练样本输出到基于支持向量回归机SVR的预测模型,得到训练后的基于SVR的预测模型。

[0189] 在一些实施例中,所述训练模块304具体用于:

[0190] 所述基于SVR的预测模型的公式为: $f(x) = \omega^T \phi(x) + b$,其中, ω 和 b 为SVR算法的参数, ω^T 表示 ω 的转置,函数 $\phi(x)$ 表示输入特征,函数 $f(x)$ 表示输出标识。

[0191] 所述将所述第一通行时长数据和所述第一天气数据作为训练样本,包括:将所述第一通行时长数据和所述第一天气数据中对应第一统计时段的数据作为训练样本的输出标识,将所述第一通行时长数据和所述第一天气数据中对应第二统计时段的数据作为训练样本的输入特征。

[0192] 所述第一交互模块302,还用于基于所述预测请求向网络侧发送当前第二统计时段所述第一路段的通行时长的数据查询指令,接收网络侧发送的所述当前第二统计时段所述第一路段的第二通行时长数据。

[0193] 在一些实施例中,所述第一交互模块302,还具体用于:基于所述预测请求向网络侧发送当前第二统计时段中每个窗口时段的平均时长的所述第一路段的通行时长的数据查询指令,接收网络侧发送的所述当前第二统计时段中每个窗口时段的平均时长的所述第一路段的第二通行时长数据。

[0194] 所述第二交互模块303,还用于基于所述预测请求向网络侧发送所述当前第二统计时段所述第一路段的所述第一天气类型的天气的数据查询指令,接收网络侧发送的所述当前第二统计时段所述第一路段的所述第一天气类型的第二天气数据。

[0195] 预测模块305,用于将所述第二通行时长数据和所述第二天气数据输出到所述训练后的基于SVR的预测模型,得到所述训练后的基于SVR的预测模型输出的第三通行时长数据。

[0196] 其中,所述第三通行时长数据为所述第一路段的第一统计时段的通行时长的预测结果。

[0197] 在一些实施例中,所述预测模块305还具体用于:将所述第二通行时长数据和所述第二天气数据作为输入特征,输出到所述训练后的基于SVR的预测模型,得到所述训练后的

基于SVR的预测模型输出的第三通行时长数据。

[0198] 在一些实施例中,所述装置还包括:预处理模块,用于将数据输入基于SVR的预测模型之前,采用数据使用范围为[0,1]的最小最大值归一化Min-Max-scaling算法将数据归一化,将归一化后的数据输入基于SVR的预测模型。

[0199] 具体地,Min-Max-scaling算法的公式为 $Z = lb + \frac{X - \min(X)}{\max(X) - \min(X)}(ub - lb)$,其中, Z 为归一化后的变量值, X 为实际变量值, $\max(X)$ 为实际变量的最大值, $\min(X)$ 为实际变量的最小值, lb 是数据使用范围的低边界值, ub 是数据使用范围的高边界值。则采用数据使用范围为[0,1]的Min-Max-scaling算法的公式,可以表示为 $Z = \frac{X - \min(X)}{\max(X) - \min(X)}$ 。

[0200] 对应地,将训练后的基于SVR的预测模型输出的预测结果数据利用公式 $Z = \frac{X - \min(X)}{\max(X) - \min(X)}$ 进行反归一化,得到预测结果的实际变量值。

[0201] 在一些实施例中,所述装置还包括:

[0202] 数据检验模块,用于检验所述第一通行时长数据中是否有部分所述预设历史周期的数据缺失。

[0203] 第三交互模块,用于当所述第一通行时长数据中有部分所述预设历史周期的数据缺失时,向网络侧发送与所述第一路段有重叠部分的路段的信息查询请求,接收网络侧发送的与所述第一路段有重叠部分的第二路段的信息,其中,所述第二路段为多个。

[0204] 所述第一交互模块302,还用于基于所述第二路段的信息,向网络侧发送所述部分预设历史周期内所述第一统计时段和第二统计时段的所述第二路段的通行时长的数据,接收网络侧发送的所述部分预设历史周期内所述第一统计时段和第二统计时段的所述第二路段的第三通行时长数据。

[0205] 数据补充模块,用于根据所述第三通行时长数据中所述第二路段与所述第一路段的重叠部分对应的通行时长数据对所述第一通行时长数据中缺失的所述部分所述预设历史周期的数据进行补充。

[0206] 在一些实施例中,所述数据检验模块,还用于在根据所述第三通行时长数据中所述第二路段与所述第一路段的重叠部分对应的通行时长数据对所述第一通行时长数据中缺失的所述部分所述预设历史周期的数据进行补充后,再次检验所述第一通行时长数据中是否有部分所述预设历史周期的数据缺失。

[0207] 所述数据补充模块,还用于当所述第一通行时长数据中仍有部分所述预设历史周期的数据缺失时,根据所述第一通行时长数据获取缺失部分的相邻周期的通行时长数据,计算所述相邻周期的通行时长数据的线性平均值,根据所述相邻周期的通行时长数据的线性平均值补充缺失部分的通行时长数据。

[0208] 图4为本发明实施例一种道路通行时长预测的装置的结构示意图,如图4所示,本发明实施例的道路通行时长预测的装置包括以下模块:

[0209] 请求模块401,用于接收针对第一路段的第一统计时段的通行时长的预测请求。

[0210] 在一些实施例中,所述第一路段包括道路系统中连接指定起点到指定终点的路

段,例如:连接某一个高速公路入口到下一个高速公路出口的高速公路路段、一条公路沿途两个公里标之间的路段等等。

[0211] 在一些实施例中,所述请求模块401,具体用于:接收针对第一路段的第一统计时段中每个窗口时段的平均通行时长的预测请求。

[0212] 在一些实施例中,所述第一统计时段,可以根据需求进行预设,例如:早高峰时段8:00-10:00、晚高峰时段17:00-19:00等。

[0213] 窗口时段是根据预测的需求来设置时长的时间段,例如:以20分钟为一个窗口时段,则早高峰时段8:00-10:00可以划分为8:00-8:20、8:20-8:40、8:40-9:00、9:00-9:20、9:20-9:40、9:40-10:00六个窗口时段;这里只是举例说明,并非对本发明的限制,在实际应用中,窗口时段的时长可以根据预测的需求确定。

[0214] 将某个窗口时段内通过所述第一路段的全部车辆的平均通行时长作为该窗口时段的平均通行时长。

[0215] 第一交互模块402,用于基于所述预测请求向网络侧发送预设历史周期内所述第一统计时段和第二统计时段的所述第一路段的通行时长的数据查询指令,接收网络侧发送的所述预设历史周期内所述第一统计时段和所述第二统计时段的所述第一路段的第一通行时长数据,其中,所述第二统计时段为与所述第一统计时段连续的前一时段。

[0216] 在一些实施例中,网络侧包括:交通综合信息平台、交通数据服务器或其它能够查询到道路通行时长历史数据的设备或系统。

[0217] 在一些实施例中,所述第一交互模块402,具体用于:基于所述预测请求向网络侧发送预设历史周期内所述第一统计时段和第二统计时段中每个窗口时段的平均通行时长的所述第一路段的通行时长的数据查询指令;

[0218] 在一些实施例中,所述第二统计时段根据所述第一统计时段确定,例如:当第一统计时段为8:00-10:00,对应第二统计时段为6:00-8:00;以20分钟为一个窗口时段,对应地,第二统计时段可以划分为6:00-6:20、6:20-6:40、6:40-7:00、7:00-7:20、7:20-7:40、7:40-8:00六个窗口时段。

[0219] 所述第一交互模块402,还具体用于:接收网络侧发送的所述预设历史周期内所述第一统计时段和所述第二统计时段中每个窗口时段的所述第一路段的第一通行时长数据。

[0220] 第二交互模块403,用于基于所述预测请求向网络侧发送所述预设历史周期内所述第一统计时段和所述第二统计时段的所述第一路段的N种天气类型的天气的数据查询指令,接收网络侧发送的所述预设历史周期内所述第一统计时段和所述第二统计时段的所述第一路段的N种天气类型的天气数据,其中,N为正整数。

[0221] 在一些实施例中,网络侧包括:天气信息服务平台、天气数据服务器或其它能够查询到所述第一路段的预设历史周期内所述第一统计时段和所述第二统计时段的天气数据的设备或系统。

[0222] 在一些实施例中,所述N种天气类型包括:气压、海平面气压、风向、风速、气温、相对湿度和降雨量。在不同的天气检测系统或不同的气候地区,用于描述天气数据的天气类型会有较大差异,这里只是对本发明一些实施例中采用到的天气类型进行举例说明,而非对于本发明的限制,在实际应用中,所述N种天气类型还可以包括其它种类。

[0223] 在一些实施例中,第一路段的天气数据采用第一路段所在地理区域的天气数据。

[0224] 在实际应用中,天气数据通常是以固定频率间隔的时间点来记录的,因此,在一些实施例中,需要对天气数据进行分析及处理,采用与第一统计时段和第二统计时段最接近的时间点的天气数据作为对应统计时段的天气数据,例如:第一路段所在地区的天气数据是每隔3个小时做一次记录的,即每天记录0点、3点、6点、9点、12点、15点、18点、21点的天气数据,则可以采用6点记录的天气数据代表6:00-8:00的天气数据,采用9点记录的天气数据代表8:00-10:00的天气数据。

[0225] 天气类型选择模块404,用于根据所述第一通行时长数据分别和所述N种天气类型的天气数据建立预测模型;验证建立的预测模型,确定验证结果与实际值偏差最小的预测模型对应的天气类型为所述第一天气类型。

[0226] 在一些实施例中,所述天气类型选择模块404具体用于:

[0227] 将所述预设历史周期分为M个子周期,其中,M为正整数;并执行以下步骤:

[0228] 步骤a:选取第i种天气类型的天气数据,其中,i为正整数, $1 \leq i \leq N$ 。

[0229] 步骤b:选取第j个子周期作为测试周期,对应地将其余子周期作为训练周期,其中,j为正整数, $1 \leq j \leq M$ 。

[0230] 在一些实施例中,为了便于统计,每个子周期为一个星期;在实际应用中,子周期的天数可以根据实际需要进行确定,此处仅为举例说明,而非对于本发明的限制。

[0231] 步骤c:将训练周期的所述第一通行时长数据和第i种天气类型的天气数据作为训练样本输出到所述基于SVR的预测模型,得到训练后的第 $i \times j$ 个基于SVR的预测模型。

[0232] 其中,第 $i \times j$ 个基于SVR的预测模型表示采用第i种天气类型和第j个子周期作为测试集时,对应的训练后的基于SVR的预测模型。

[0233] 所述基于SVR的预测模型的公式为: $f(x) = \omega^T \phi(x) + b$,其中, ω 和b为SVR算法的参数, ω^T 表示 ω 的转置,函数 $\phi(x)$ 表示输入特征,函数 $f(x)$ 表示输出标识。

[0234] 在一些实施例中,所述将训练周期的所述第一通行时长数据和第i种天气类型的天气数据作为训练样本输出到所述基于SVR的预测模型,得到训练后的第 $i \times j$ 个基于SVR的预测模型,具体包括:将训练周期的所述第一通行时长数据和第i种天气类型的天气数据中对应第一统计时段的数据作为训练样本的输出标识,将训练周期的所述第一通行时长数据和第i种天气类型的天气数据中对应第二统计时段的数据作为训练样本的输入特征。

[0235] 步骤d:将测试周期的第一统计时段的所述第一通行时长数据和第i种天气类型的天气数据输出到所述训练后的第 $i \times j$ 个基于SVR的预测模型,得到所述训练后的第 $i \times j$ 个基于SVR的预测模型输出的测试周期的第二统计时段通行时长预测值。

[0236] 在一些实施例中,所述将测试周期的第一统计时段的所述第一通行时长数据和第i种天气类型的天气数据输出到所述训练后的第 $i \times j$ 个基于SVR的预测模型,得到所述训练后的第 $i \times j$ 个基于SVR的预测模型输出的测试周期的第二统计时段通行时长预测值,具体包括:将测试周期的第一统计时段的所述第一通行时长数据和第i种天气类型的天气数据作为输入特征,输出到所述训练后的第 $i \times j$ 个基于SVR的预测模型,得到所述训练后的第 $i \times j$ 个基于SVR的预测模型输出的预测标识测试周期的第二统计时段通行时长预测值。

[0237] 步骤e:根据测试周期的第二统计时段通行时长预测值和测试周期的第二统计时段的所述第一通行时长数据,计算测试周期对应的平均绝对百分比误差MAPE值。

[0238] 在一些实施例中,所述根据测试周期的第二统计时段通行时长预测值和测试周期

的第二统计时段的所述第一通行时长数据,计算测试周期对应的平均绝对百分比误差MAPE值,具体包括:

[0239] MAPE值的计算公式为 $MAPE = \frac{1}{T} \sum_{t=1}^T \left| \frac{d_t - p_t}{d_t} \right|$,其中,T为测试集的数据周期总数,t为

测试周期的数据周期编号, d_t 为真实值, p_t 为预测值。

[0240] 将测试周期的第二统计时段通行时长预测值作为预测值,将测试周期的第二统计时段的所述第一通行时长数据作为真实值,输入所述MAPE值的计算公式,得到测试周期对应的MAPE值。

[0241] 步骤f:再次执行步骤b至步骤e,直至得到所述M个子周期分别作为测试周期时对应的MAPE值,将所述M个子周期分别作为测试周期时对应的MAPE值的平均值,作为第i种天气类型对应的MAPE值。

[0242] 步骤g:再次执行步骤a至步骤f,直至得到所述N种天气类型对应的MAPE值,将所述N种天气类型中对应的MAPE值最小的天气类型作为所述第一天气类型。

[0243] 所述第二交互模块403,还用于基于所述预测请求向网络侧发送所述预设历史周期内所述第一统计时段和所述第二统计时段的所述第一路段的第一天气类型的天气的数据查询指令,接收网络侧发送的所述预设历史周期内所述第一统计时段和所述第二统计时段的所述第一路段的所述第一天气类型的第一天气数据。

[0244] 训练模块405,用于将所述第一通行时长数据和所述第一天气数据作为训练样本输出到基于支持向量回归机SVR的预测模型,得到训练后的基于SVR的预测模型。

[0245] 在一些实施例中,所述步骤405具体包括:

[0246] 所述基于SVR的预测模型的公式为: $f(x) = \omega^T \phi(x) + b$,其中, ω 和**b**为SVR算法的参数, ω^T 表示 ω 的转置,函数 $\phi(x)$ 表示输入特征,函数 $f(x)$ 表示输出标识。

[0247] 所述将所述第一通行时长数据和所述第一天气数据作为训练样本,包括:将所述第一通行时长数据和所述第一天气数据中对应第一统计时段的数据作为训练样本的输出标识,将所述第一通行时长数据和所述第一天气数据中对应第二统计时段的数据作为训练样本的输入特征。

[0248] 所述第一交互模块402,还用于基于所述预测请求向网络侧发送当前第二统计时段所述第一路段的通行时长的数据查询指令,接收网络侧发送的所述当前第二统计时段所述第一路段的第二通行时长数据。

[0249] 在一些实施例中,所述第一交互模块402,还具体用于:基于所述预测请求向网络侧发送当前第二统计时段中每个窗口时段的平均时长的所述第一路段的通行时长的数据查询指令,接收网络侧发送的所述当前第二统计时段中每个窗口时段的平均时长的所述第一路段的第二通行时长数据。

[0250] 所述第二交互模块403,还用于基于所述预测请求向网络侧发送所述当前第二统计时段所述第一路段的所述第一天气类型的天气的数据查询指令,接收网络侧发送的所述当前第二统计时段所述第一路段的所述第一天气类型的第二天气数据。

[0251] 预测模块406,用于将所述第二通行时长数据和所述第二天气数据输出到所述训练后的基于SVR的预测模型,得到所述训练后的基于SVR的预测模型输出的第三通行时长数据。

[0252] 其中,所述第三通行时长数据为所述第一路段的第一统计时段的通行时长的预测结果。

[0253] 在一些实施例中,所述预测模块406还具体用于:将所述第二通行时长数据和所述第二天气数据作为输入特征,输出到所述训练后的基于SVR的预测模型,得到所述训练后的基于SVR的预测模型输出的第三通行时长数据。

[0254] 在一些实施例中,所述装置还包括:预处理模块,用于将数据输入基于SVR的预测模型之前,采用数据使用范围为[0,1]的最小最大值归一化Min-Max-scaling算法将数据归一化,将归一化后的数据输入基于SVR的预测模型。

[0255] 具体地,Min-Max-scaling算法的公式为 $Z = lb + \frac{X - \min(X)}{\max(X) - \min(X)}(ub - lb)$,其中,

Z为归一化后的变量值,X为实际变量值,max(X)为实际变量的最大值,min(X)为实际变量的最小值,lb是数据使用范围的低边界值,ub是数据使用范围的高边界值。则采用数据使用范围

为[0,1]的Min-Max-scaling算法的公式,可以表示为 $Z = \frac{X - \min(X)}{\max(X) - \min(X)}$ 。

[0256] 对应地,将训练后的基于SVR的预测模型输出的预测结果数据利用公式

$Z = \frac{X - \min(X)}{\max(X) - \min(X)}$ 进行反归一化,得到预测结果的实际变量值。

[0257] 在一些实施例中,所述装置还包括:

[0258] 数据检验模块,用于检验所述第一通行时长数据中是否有部分所述预设历史周期的数据缺失。

[0259] 第三交互模块,用于当所述第一通行时长数据中有部分所述预设历史周期的数据缺失时,向网络侧发送与所述第一路段有重叠部分的路段的信息查询请求,接收网络侧发送的与所述第一路段有重叠部分的第二路段的信息,其中,所述第二路段为多个。

[0260] 所述第一交互模块402,还用于基于所述第二路段的信息,向网络侧发送所述部分预设历史周期内所述第一统计时段和第二统计时段的所述第二路段的通行时长的数据,接收网络侧发送的所述部分预设历史周期内所述第一统计时段和第二统计时段的所述第二路段的第三通行时长数据。

[0261] 数据补充模块,用于根据所述第三通行时长数据中所述第二路段与所述第一路段的重叠部分对应的通行时长数据对所述第一通行时长数据中缺失的所述部分所述预设历史周期的数据进行补充。

[0262] 在一些实施例中,所述数据检验模块,还用于在根据所述第三通行时长数据中所述第二路段与所述第一路段的重叠部分对应的通行时长数据对所述第一通行时长数据中缺失的所述部分所述预设历史周期的数据进行补充后,再次检验所述第一通行时长数据中是否有部分所述预设历史周期的数据缺失。

[0263] 所述数据补充模块,还用于当所述第一通行时长数据中仍有部分所述预设历史周期的数据缺失时,根据所述第一通行时长数据获取缺失部分的相邻周期的通行时长数据,计算所述相邻周期的通行时长数据的线性平均值,根据所述相邻周期的通行时长数据的线性平均值补充缺失部分的通行时长数据。

[0264] 基于上述道路通行时长预测的方法,本发明实施例提供一种道路通行时长预测方

法,包括以下步骤:

[0265] 步骤501,在获取一组历史行驶时间数据,先按照道路规划将路线分段,图5为某高速公路部分路段的交通路线图,例如如图5所述的交通路线上自2016年7月19号到10月17号的交通行驶时间数据,首先按照岔路口位置及收费站位置将上述交通路线分成10个路段,即道路A-2,A-3,B-1,B-3,C-1,C-3,1-A,3-A,3-B,3-C。

[0266] 步骤502,将历史行驶时间数据所覆盖的时间范围按照固定间隔划分为时间段,例如将自2016年7月19号到10月17号的交通行驶时间数据所有长度超过两个小时的时间段都平分为20分钟的时间点,例如6点到6点20为第一个时间点(例如[2016-7-19 6:00:00, 2016-7-19 6:20:00]),则6点到8点共有六个时间窗口,同样8点到10点,15点到17点和17点到19点都有六个时间窗口。

[0267] 步骤503,计算在每个时间窗口内进入道路,例如在2016-7-19 6:00:00至2016-7-19 6:20:00之间进入道路A-2的所有车辆的平均行驶时间。

[0268] 步骤504,如果存在部分数据缺失,则采取道路互补法来做数据填补。道路互补法的具体操作步骤:当道路R($R \in [B-1, B-3, C-1, C-3]$)的一个20分钟时间窗口内平均行驶时间速度为空,这部分的缺失数据使用它毗邻道路的平均行驶时间数据进行补充。

[0269] 例如:如图5所示,道路C-1的毗邻道路为C-3和B-1,C-3的毗邻道路为C-1和B-3,B-1的毗邻道路为B-3和C-1,B-3的道路为B-1和C-3。如果道路C-3的[2016-09-18 07:00:00, 2016-09-18 07:20:00)时间窗口内平均行驶时间数据缺失,可取道路C-1的岔路C到点p和B-3的点p到收费站3的[2016-09-18 07:00:00, 2016-09-18 07:20:00)时间窗口的数据来填补道路C-3的这部分缺失数据。类似的,如果道路C-1的[2016-09-18 07:00:00, 2016-09-18 07:20:00)时间窗口内平均行驶时间数据缺失,可取道路C-3的岔路C到点p和道路B-1的点p到收费站1的[2016-09-18 07:00:00, 2016-09-18 07:20:00)时间窗口的数据来填补道路C-1的这部分缺失数据。再对B-1和B-3进行类似操作。

[0270] 步骤505,对仍然不能被填补的缺失数据用线性插值法进行填充。

[0271] 步骤506,划分每段道路时间窗口位置和前两个小时的行驶时间。以上面的数据为例,时间窗口位置指的是目标预测时间08:00-10:00和17:00-19:00,这两段时间都被分为20分钟为一个时间窗口的六个时段,例如8:00-8:20为第一个时间窗口位置,8:20-8:40为第二个时间窗口位置,以此类推。则前两个小时的行驶时间指的是6:00-8:00和15:00-17:00的每20分钟内的该段道路内汽车平均行驶时间。将所得到的时间窗口位置前两个小时内的平均行驶时间转化为特征向量。

[0272] 步骤507,在将以上步骤所得到的交通时间特征向量导入SVR模型前,本方案同时还使用天气数据对所得到的交通时间特征向量进行优化,天气数据包括日期,小时,气压,海平面气压,风向,风速,气温,相对湿度和降雨量。天气数据是每隔三个小时做一次记录的,即每天记录0点、3点、6点、9点、12点、15点、18点、21点的天气状况数据。6点和15点的天气数据可以代表在6点到8点和15点的天气到17点的区间内的天气,因此可以将6点和15点的天气数据对该段时间内的交通时间特征向量进行优化,即消除因为天气原因导致的交通时间意外降低或升高对数据准确性的影响。

[0273] 由于天气数据种类较多,同时不同天气数据对交通流量的影响度也有高有低,这主要是因为不同季节内产生主要影响的天气种类不同,为保证预测精度的同时缩短预测时

间,本方案以周为单位,为每周的历史交通行驶时间数据选择一个最优天气指标,即分别将天气数据(气压、海平面气压、风向、风速、气温、相对湿度和降雨量)作为特征向量进行实验,通过交叉验证选择最优天气指标的流程如图6所示,交叉验证是指以随机的一周作为测试集,其余周作为训练集,得出所有实验的MAPE值后取平均值。MAPE值最小的即作为当周的历史交通行驶时间数据的最优天气指标,当使用该周的历史交通行驶时间数据时,需要同时代入该最优天气指标。

[0274] 例如,在一个实施例中,根据计算,MAPE值如表1所示,2016年7月19号至25日温度作为特征向量时,MAPE值最小,则温度为2016年7月19号至25日的最优天气指标,在纳入预测模型时,对于2016年7月19号至25日的历史交通行驶时间数据,需要同时使用2016年7月19号至25日的温度数据对该段时间内的交通行驶时间数据进行优化。

	天气类型	验证结果	测试数据的预测
	无	0.2276	0.1935
	气压	0.2241	0.1934
	海平面气压	0.2241	0.1934
[0275]	风向	0.2275	0.1924
	风速	0.2278	0.1936
	气温	0.2228	0.1894
	相对湿度	0.2280	0.1900
	降雨量	0.2280	0.1936

[0276] 表1

[0277] 步骤508,将优化后的交通行驶时间数据特征向量及最优天气指标特征向量导入采用SVR算法的预测模型,对未来时段的交通行驶时间进行预测,利用交叉验证法在鲁棒归一化(Robust-scaling),标准归一化(Standard-scaling)和Min-Max-scaling中确定对数据使用范围为[0,1]的Min-Max-scaling方法对数据进行预处理,并通过SVR算法最终得到目标时间段的交通行驶时间预测结果。

[0278] 为了实现本发明实施例的道路通行时长预测的方法,图7为本发明实施例提供的基于硬件实现的道路通行时长预测的设备的结构示意图,如图7所示,本发明实施例的道路通行时长预测的设备710包括:处理器71和用于存储能够在处理器上运行的计算机程序的存储器72,其中,

[0279] 所述处理器71用于运行所述计算机程序时,执行:

[0280] 接收到针对第一路段的第一统计时段的通行时长的预测请求;

[0281] 基于所述预测请求向网络侧发送预设历史周期内所述第一统计时段和第二统计

时段的所述第一路段的通行时长的数据查询指令,接收网络侧发送的所述预设历史周期内所述第一统计时段和所述第二统计时段的所述第一路段的第一通行时长数据,其中,所述第二统计时段为与所述第一统计时段连续的前一时段;

[0282] 基于所述预测请求向网络侧发送所述预设历史周期内所述第一统计时段和所述第二统计时段的所述第一路段的第一天气类型的天气的数据查询指令,接收网络侧发送的所述预设历史周期内所述第一统计时段和所述第二统计时段的所述第一路段的所述第一天气类型的第一天气数据;

[0283] 将所述第一通行时长数据和所述第一天气数据作为训练样本输出到基于支持向量回归机SVR的预测模型,得到训练后的基于SVR的预测模型;

[0284] 基于所述预测请求向网络侧发送当前第二统计时段所述第一路段的通行时长的数据查询指令,接收网络侧发送的所述当前第二统计时段所述第一路段的第二通行时长数据;

[0285] 基于所述预测请求向网络侧发送所述当前第二统计时段所述第一路段的所述第一天气类型的天气的数据查询指令,接收网络侧发送的所述当前第二统计时段所述第一路段的所述第一天气类型的第二天气数据;

[0286] 将所述第二通行时长数据和所述第二天气数据输出到所述训练后的基于SVR的预测模型,得到所述训练后的基于SVR的预测模型输出的第三通行时长数据。

[0287] 在一些实施例中,所述处理器71还用于运行所述计算机程序时,执行:

[0288] 基于所述预测请求向网络侧发送所述预设历史周期内所述第一统计时段和所述第二统计时段的所述第一路段的N种天气类型的天气的数据查询指令,接收网络侧发送的所述预设历史周期内所述第一统计时段和所述第二统计时段的所述第一路段的N种天气类型的天气数据,其中,N为正整数;

[0289] 根据所述第一通行时长数据分别和所述N种天气类型的天气数据建立预测模型;验证建立的预测模型,确定验证结果与实际值偏差最小的预测模型对应的天气类型为所述第一天气类型。

[0290] 在一些实施例中,所述处理器71还用于运行所述计算机程序时,执行:

[0291] 将数据输入基于SVR的预测模型之前,采用数据使用范围为[0,1]的最小最大值归一化Min-Max-scaling算法将数据归一化,将归一化后的数据输入基于SVR的预测模型。

[0292] 上述实施例提供的道路通行时长预测的装置与道路通行时长预测的方法实施例属于同一构思,其具体实现过程详见方法实施例,这里不再赘述。

[0293] 当然,实际应用时,如图7所示,该道路通行时长预测的装置还可以包括至少一个通信接口73。道路通行时长预测的装置中的各个组件通过总线系统74耦合在一起。可理解,总线系统74用于实现这些组件之间的连接通信。总线系统74除包括数据总线之外,还包括电源总线、控制总线和状态信号总线。但是为了清楚说明起见,在图7中将各种总线都标为总线系统74。

[0294] 其中,通信接口73,用于与其它设备进行交互。

[0295] 具体来说,所述处理器71可以通过通信接口73向对应所述被调用方应用的应用服务器发送操作结果查询请求,获取所述应用服务器发送的所述被调用方应用的操作结果。

[0296] 本领域技术人员可以理解,存储器72可以是易失性存储器或非易失性存储器,也

可包括易失性和非易失性存储器两者。其中,非易失性存储器可以是只读存储器(ROM,Read Only Memory)、可编程只读存储器(PROM,Programmable Read-Only Memory)、可擦除可编程只读存储器(EPROM,Erasable Programmable Read-Only Memory)、电可擦除可编程只读存储器(EEPROM,Electrically Erasable Programmable Read-Only Memory)、磁性随机存取存储器(FRAM,ferromagnetic random access memory)、快闪存储器(Flash Memory)、磁表面存储器、光盘、或只读光盘(CD-ROM,Compact Disc Read-Only Memory);磁表面存储器可以是磁盘存储器或磁带存储器。易失性存储器可以是随机存取存储器(RAM,Random Access Memory),其用作外部高速缓存。通过示例性但不是限制性说明,许多形式的RAM可用,例如静态随机存取存储器(SRAM,Static Random Access Memory)、同步静态随机存取存储器(SSRAM,Synchronous Static Random Access Memory)、动态随机存取存储器(DRAM,Dynamic Random Access Memory)、同步动态随机存取存储器(SDRAM,Synchronous Dynamic Random Access Memory)、双倍数据速率同步动态随机存取存储器(DDRSDRAM,Double Data Rate Synchronous Dynamic Random Access Memory)、增强型同步动态随机存取存储器(ESDRAM,Enhanced Synchronous Dynamic Random Access Memory)、同步连接动态随机存取存储器(SLDRAM,SyncLink Dynamic Random Access Memory)、直接内存总线随机存取存储器(DRRAM,Direct Rambus Random Access Memory)。本发明实施例描述的存储器72旨在包括但不限于这些和任意其它适合类型的存储器。

[0297] 本发明实施例中,还提供了一种计算机可读存储介质,用于存储上述实施例中提供的计算程序,以完成前述方法所述步骤。计算机可读存储介质可以是FRAM、ROM、PROM、EPROM、EEPROM、Flash Memory、磁表面存储器、光盘、或CD-ROM等存储器;也可以是包括上述存储器之一或任意组合的各种设备,如移动电话、计算机、智能家电、服务器等。

[0298] 需要说明的是:本发明实施例所记载的技术方案之间,在不冲突的情况下,可以任意组合。

[0299] 以上所述,仅为本发明的较佳实施例而已,并非用于限定本发明的保护范围。

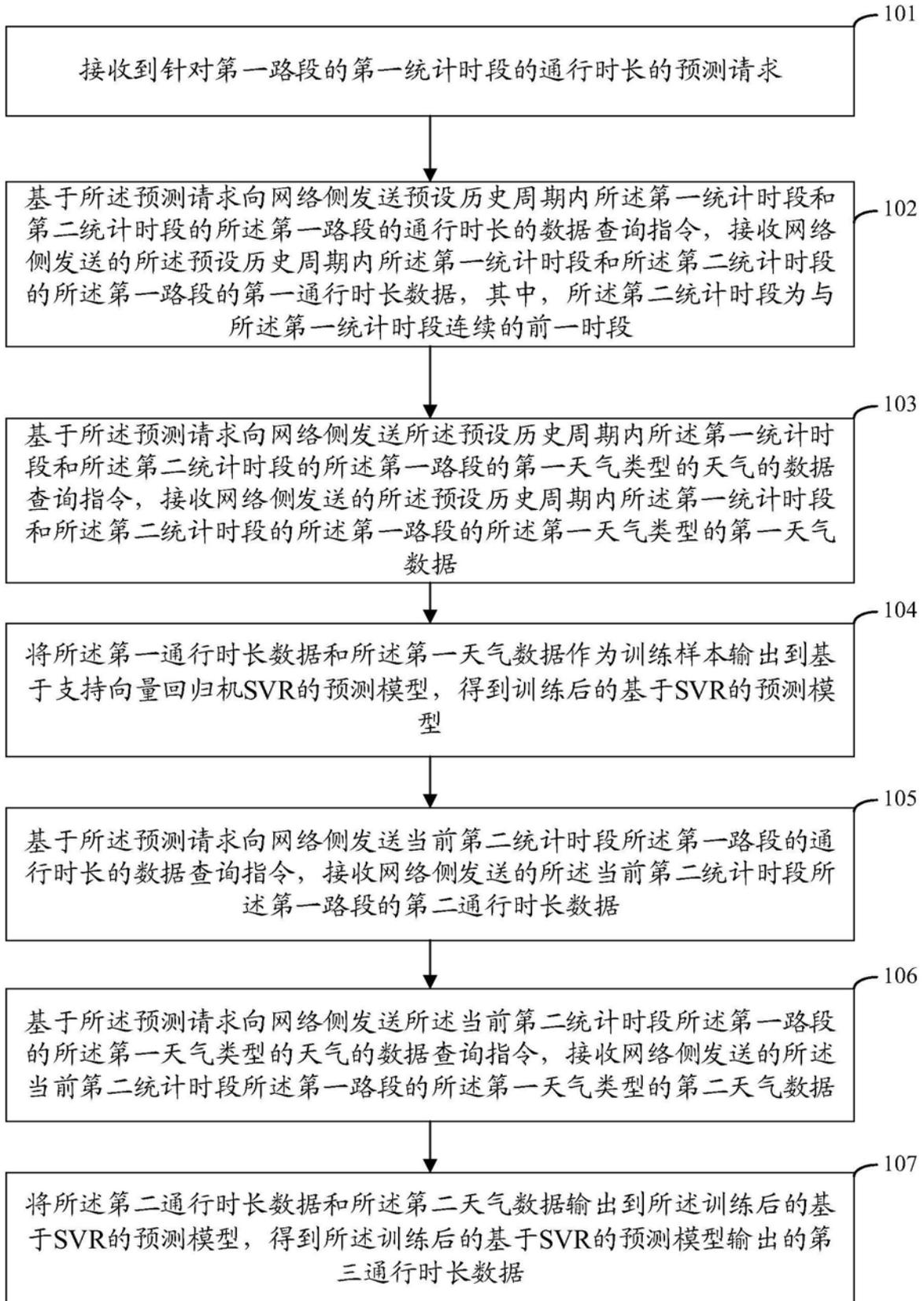


图1

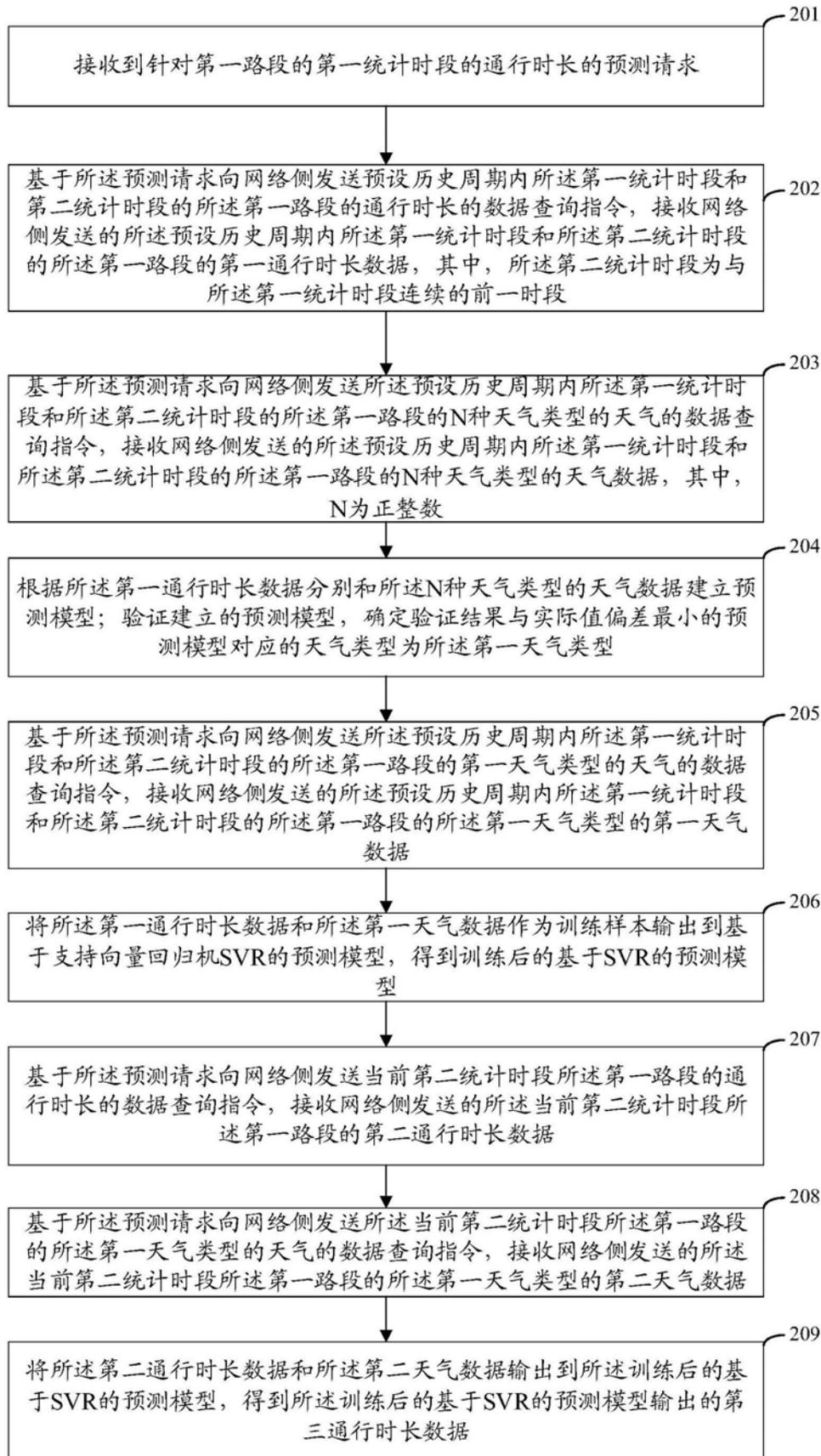


图2

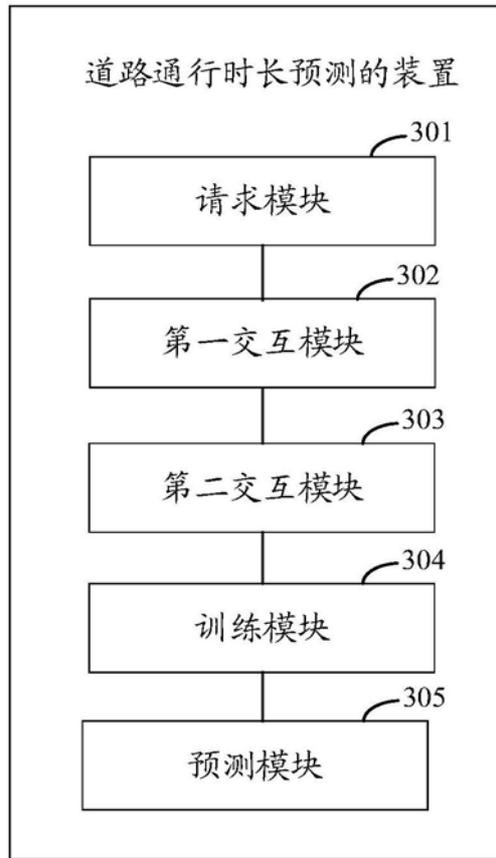


图3

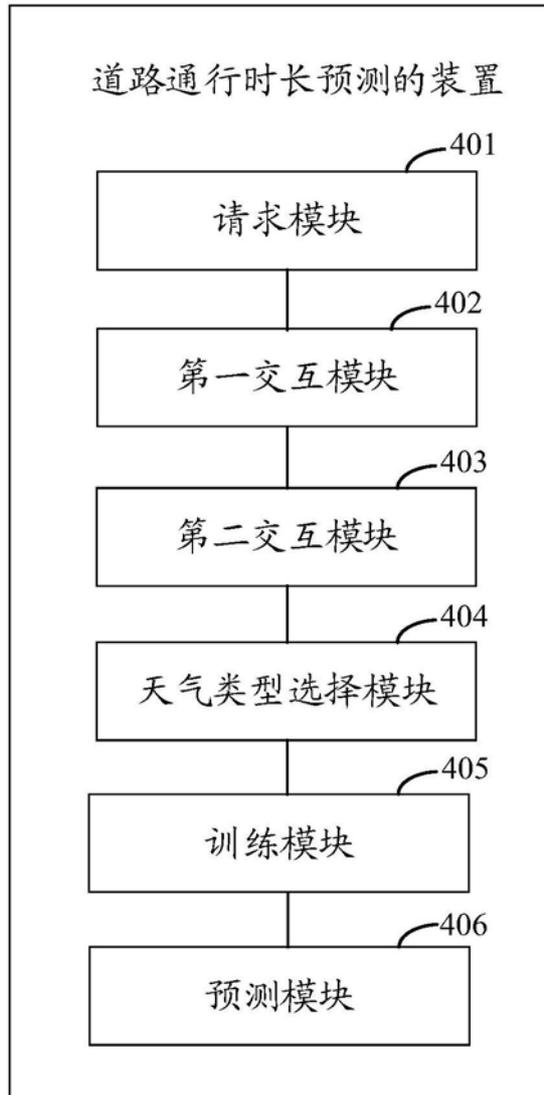


图4

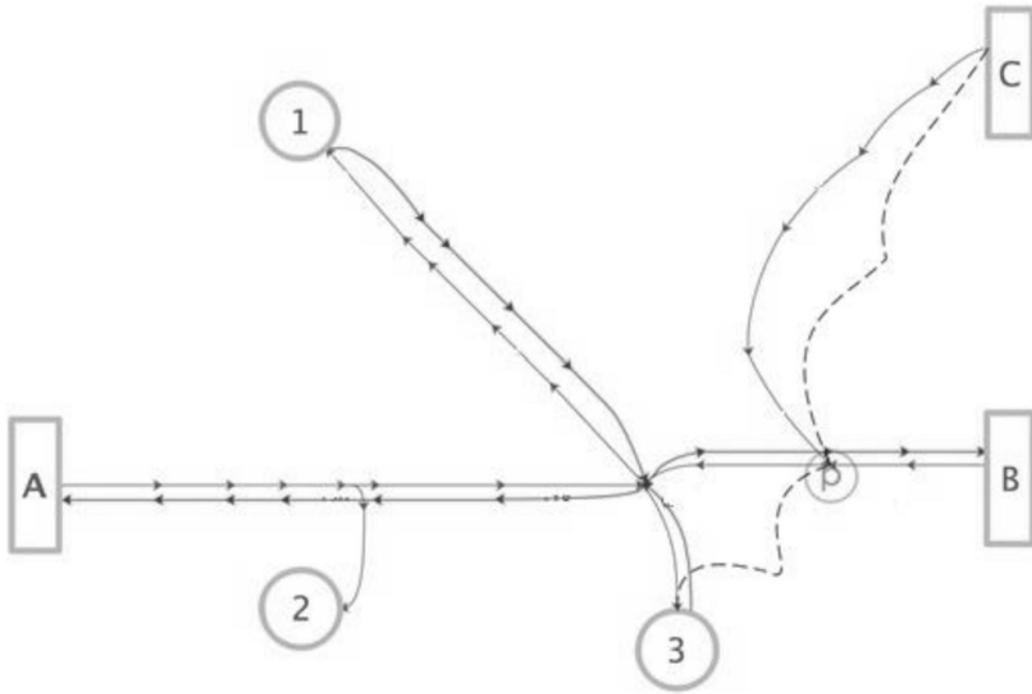


图5

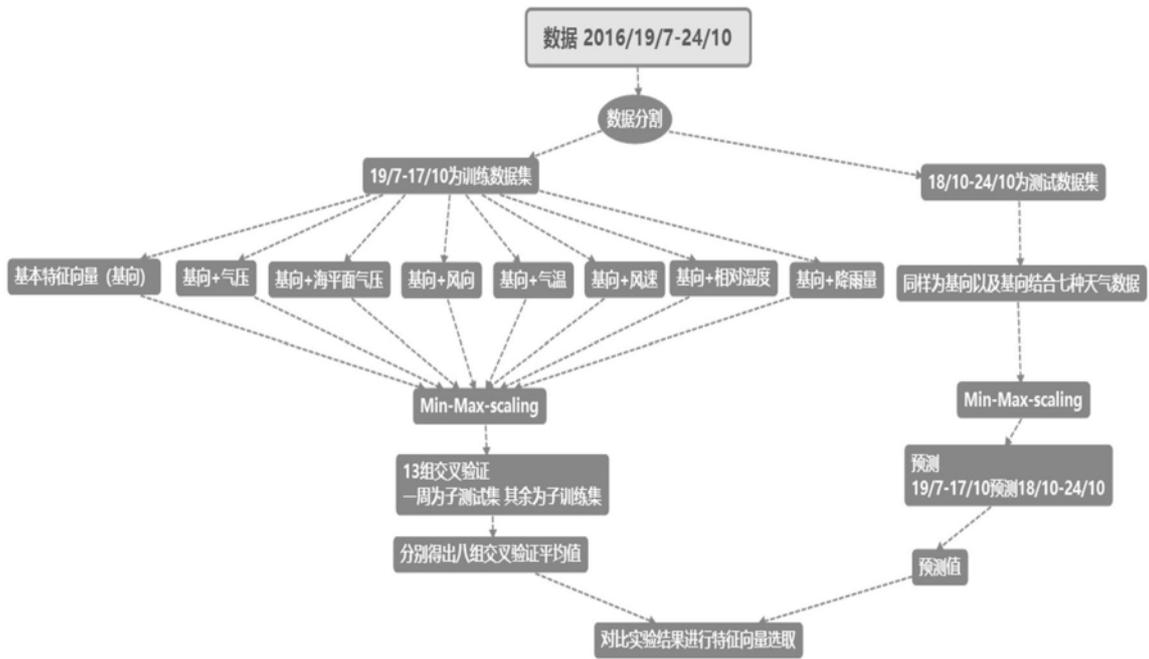


图6

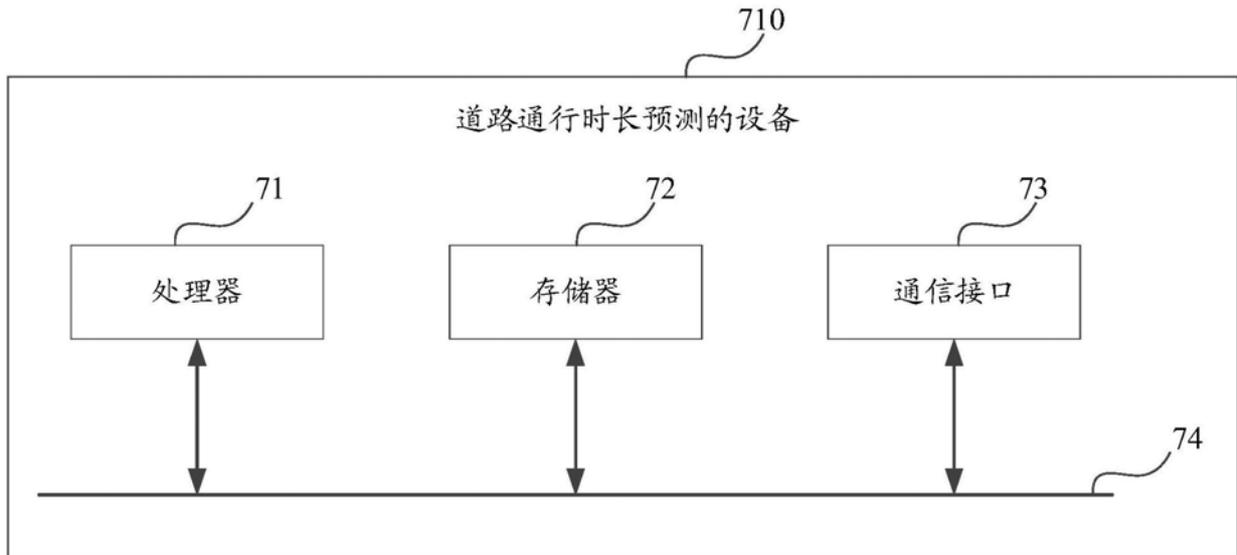


图7