



(12) 发明专利

(10) 授权公告号 CN 103177570 B

(45) 授权公告日 2015. 03. 04

(21) 申请号 201210572722. 5

(22) 申请日 2012. 12. 25

(73) 专利权人 北京交通发展研究中心

地址 100055 北京市丰台区六里桥南路甲 9
号(首发大厦)A座503室

(72) 发明人 孙建平 温慧敏 郭继孚 高永
张溪

(74) 专利代理机构 北京鼎佳达知识产权代理事
务所(普通合伙) 11348

代理人 王伟锋 刘铁生

(51) Int. Cl.

G08G 1/00(2006. 01)

G06Q 10/04(2012. 01)

审查员 沈紫乐

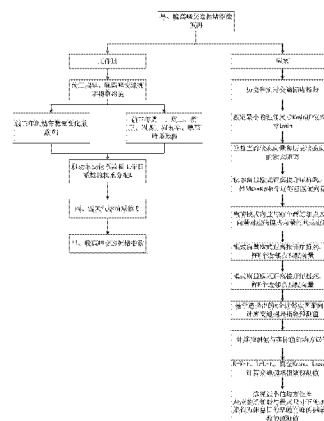
权利要求书2页 说明书8页 附图2页

(54) 发明名称

一种早晚高峰交通拥堵指数的预测方法

(57) 摘要

本发明公开了一种早晚高峰交通拥堵指数的预测方法,包括以下步骤:a. 将预测天分为工作日和休息日;b. 采用因素分解模型对正常天气的工作日的早晚高峰的拥堵指数进行预测,当出现雨、雪、雾的极端天气时,在正常的工作日的早晚高峰的拥堵指数预测的基础上用天气影响系数 ξ 做修正;c. 采用双模式临近模型对正常天气的休息日的早晚高峰的拥堵指数进行预测,当出现雨、雪、雾的极端天气时,在正常的休息日的早晚高峰的拥堵指数预测的基础上用天气影响系数 ξ 做修正。本发明方法可以准确预测交通状况,为未来的交通管理提供依据。



1. 一种早晚高峰交通拥堵指数的预测方法,其特征在于,包括以下步骤:

a. 将预测天分为工作日和休息日;其中

b. 采用因素分解模型对正常天气的工作日的早晚高峰的拥堵指数进行预测,当出现雨、雪、雾的极端天气时,在正常的工作日的早晚高峰的拥堵指数预测的基础上用天气影响系数做修正;

c. 采用双模式临近模型对正常天气的休息日的早晚高峰的拥堵指数进行预测,当出现雨、雪、雾的极端天气时,在正常的休息日的早晚高峰的拥堵指数预测的基础上用天气影响系数做修正。

2. 根据权利要求1所述的早晚高峰交通拥堵指数的预测方法,其特征在于,所述的因素分解模型中以不同工作日和路网中机动车数量为两个影响因素。

3. 根据权利要求2所述的早晚高峰交通拥堵指数的预测方法,其特征在于,采用因素分解模型对正常天气的工作日的早晚高峰的拥堵指数进行预测的过程如下:首先,参考预测周次前一定时期的工作日的早晚高峰交通拥堵指数的数据,并分别计算平均值;然后根据预测的不同工作日及机动车数量变化情况,得到工作日系数和机动车数量系数,并用平均值分别与工作日系数和机动车数量系数相乘,得到两个值;最后结合通过实际数据得到的最优权重比,通过下式1得到早晚高峰交通拥堵指数的预测值 a' :

$$a' = \lambda a \beta_i + (1-\lambda) a \delta_j \quad \text{式 1}$$

式中, a 为参考的交通拥堵指数的平均值,

β_i 为工作日系数, $i = 1, 2, \dots, 5$, i 表示不同的工作日,

δ_j 为机动车数量系数, $j = 1, 2, \dots, 5$, j 表示不同的机动车数量情况,

λ 为工作日影响因素的权重。

4. 根据权利要求3所述的早晚高峰交通拥堵指数的预测方法,其特征在于,极端天气的工作日的早晚高峰交通拥堵指数的预测值 a'' 通过如下公式2得到:

$$a'' = (\lambda a \beta_i + (1-\lambda) a \delta_j) \xi_g \quad \text{式 2}$$

式中, ξ_g 为天气因素系数, $g = 1, 2, 3$, 表示不同种类的天气因素。

5. 根据权利要求3所述的早晚高峰交通拥堵指数的预测方法,其特征在于,所述工作日系数 β_i 通过如下步骤得到,利用预测前一至三年不同工作日的早、晚高峰交通拥堵指数分别求周一至周五早高峰和晚高峰交通拥堵指数的平均值,并分别与周一至周五平均值的总平均值相除,得到周一至周五的工作日系数 β_i 。

6. 根据权利要求3所述的早晚高峰交通拥堵指数的预测方法,其特征在于,所述机动车数量系数 δ_j 通过如下步骤得到,利用预测前一至三年工作日的早、晚高峰交通拥堵指数分别求五种不同机动车数量情况的早高峰和晚高峰交通拥堵指数的平均值,并分别与不同机动车数量情况平均值的总平均值相除,得到机动车数量系数 δ_j 。

7. 根据权利要求1所述的早晚高峰交通拥堵指数的预测方法,其特征在于,所述双模式临近模型采用近邻点法对历史交通拥堵指数数据进行搜索,寻求与预测状态最相近的历史数据进行预测,具体步骤如下:

1): 选定最小的近邻尺寸 k , $k = k_{\min}$, 近邻尺寸为选取与当前状态向量距离最近的向量的个数;

2): 选定最小的模式尺寸 l , $l = l_{\min}$, 模式尺寸为模式向量内元素的个数;

3) :当前的状态向量为 $V(n) = [v(n-1), \dots, v(n-1), v(n)]$, $v(n)$ 代表第 n 个时刻的交通拥堵指数, n 为当前时刻;

模式向量为 $P(n) = [d(n-1+1), \dots, d(n-1), d(n)]$, 模式向量反映状态向量的变化趋势特征;

而 t 时刻的状态向量为 $V(t) = [v(t-1), \dots, v(t-1), v(t)]$, $t \geq 1+1$, 为历史某时刻;

模式向量为 $P(t) = [d(t-1+1), \dots, d(t-1), d(t)]$;

状态向量的差分向量定义为 $S(t) = [s(t-1+1), \dots, s(t-1), s(t)]$, 其中 $s(t) = v(t) - v(t-1)$;

模式向量如下定义

$$d(t) = \begin{cases} -1, & \text{当 } s(t) < 0 \text{ 即 } v(t+1) < v(t) \\ 0, & \text{当 } s(t) = 0 \text{ 即 } v(t+1) = v(t) \\ 1, & \text{当 } s(t) > 0 \text{ 即 } v(t+1) > v(t) \end{cases};$$

对时间序列 $v(1), v(2), \dots, v(n)$ 利用如下状态向量判断准则

$$dsv(t) = \|V(n) - V(t)\| = \sqrt{\sum_{m=0}^l [v(n-m) - v(t-m)]^2}; m \text{ 为 } 0 \text{ 到 } l \text{ 顺序增长的整数};$$

确定当前状态向量与历史状态向量的欧式距离, 并按 $dsv(t)$ 的升序排列, 确定 $M_s (M_s \geq k)$ 个近邻点匹配向量;

4) 由当前状态向量确定的 M_s 个最近邻点匹配向量按如下模式向量判断准则匹配

$$dpv(t) = \|P(n) - P(t)\| = \sqrt{\sum_{m=0}^{l-1} [d(n-m) - d(t-m)]^2}$$

确定当前模式向量 $P(n)$ 与 M_s 个最近邻点匹配向量对应的模式向量的欧式距离, 并按 $dpv(t)$ 的升序排列, 最后确定与当前状态及模式向量均较为匹配的 k 个近邻点;

5) 预测 $n+1$ 时刻的值 $\hat{v}(n+1)$

$$\hat{v}(n+1) = v(n) + s \quad \text{其中 } s = \left(\frac{1}{k}\right) \sum_{h=1}^k S(t, h)$$

在最近邻点中, $S(t, h) = [s(t-1+1, h), \dots, s(t-1, h), s(t, h)]$ 是对应向量 $V(t, h) = [v(t-1, h), \dots, v(t-1, h), v(t, h)]$ 的差分向量, 即 $s(t, h) = v(t, h) - v(t-1, h)$;

6) 计算实际值与预测值间的均方误差

$$RMSE = \sqrt{\frac{1}{N} \sum_{m=1}^N [v(m) - \hat{v}(m)]^2}$$

其中 $\hat{v}(m)$ 是 $v(m)$ 的预测值, N 是预测的总个数;

7) 对于不同的模式尺寸 $1+1, 1+2, \dots, 1 \text{ max}$ 重复步骤 3) 到 6);

8) 对于不同的近邻数 $k+1, k+2, \dots, k \text{ max}$ 重复步骤 2) 到 7);

9) 选择最小的 RMSE 对应的近邻数与模式尺寸下预测结果作为对未来休息日的早晚高峰的拥堵指数的预测值。

一种早晚高峰交通拥堵指数的预测方法

技术领域

[0001] 本发明属于城市道路交通管理中交通状况预测的技术领域,尤其涉及一种早晚高峰交通拥堵指数的预测方法。

背景技术

[0002] 随着社会经济的不断发展,汽车保有量的不断攀升,交通拥堵问题显得日益严重。交通拥堵指数是目前评价城市道路交通运行状况的重要指标。交通拥堵指数是利用路面浮动车数据对道路拥堵进行等级判别的一种方法与手段。目前技术可以根据历史浮动车数据给出现时刻的交通拥堵指数,本发明着重解决对未来天早晚高峰时段的交通拥堵指数进行预测,给出引导未来出行的参考值。

[0003] 对于交通状况的预测,目前已经有一系列的技术方法和手段被应用于该领域。这些预测方法均是基于交通流的预测,大体可以分为两类:一类是以数理统计和微积分等传统数学和物理方法为基础的预测模型;一类是以现代科学技术和方法(如模糊控制、神经网络、模拟技术)为主要研究手段而形成的预测模型。其特点是所采用模型和方法不追求严格意义上的数学推导和明确的物理意义,而更重视对实际交通流现象的拟合效果。前者包括卡尔曼滤波模型、参数回归模型、时间序列-ARIMA模型、指数平滑模型以及由这些模型构成的各种组合模型,后者则包括状态相空间重构模型、非参数回归模型、谱分析法、基于多维分形的方法、基于小波分解与重构的方法和多种神经网络相关的复合预测模型等。以上交通流预测方法是基于实际交通流数据,输出结果也是交通流量。

[0004] 现有的预测方法的有缺点如下:历史平均模型具有操作简单,预测速度快的优点,但难以反映不确定性与非线性特性,尤其无法克服随机干扰因素的影响,不能应付突发事件。时间序列-ARIMA模型具有技术成熟,精确度高的优点,但模型容易造成数据遗失,模型参数初始化复杂,而且参数不能移植。卡尔曼滤波模型既能处理平稳数据,也能处理非平稳数据,预测精度高。但不适用于随机性和非线性较强的状态,参数估计计算量大,难于在线预测。非参数回归模型无需假设基本关系,但具有寻找近邻的复杂性的缺点。神经网络模型方法成熟,广泛应用于交通预测领域,但训练过程复杂,训练结果推广能力差,同时,神经网络的学习算法采用经验风险最小化原理,不能使期望风险最小化。相空间重构模型适合非线性条件、预测精确度高,但需要大量的历史数据,计算量大。

[0005] 交通拥堵指数是用于评价路网运行状况的评价指数,可展现和评估实时的路网状况,目前还未有适合的方法用于预测未来的交通拥堵指数。

[0006] 鉴于上述现有预测技术不适用于交通拥堵指数的预测,本发明人基于多年的实践经验和丰富的专业知识,积极加以研究创新,最终发明了一种早晚高峰交通拥堵指数的预测方法,以准确预测交通状况,为未来的交通管理提供依据。

发明内容

[0007] 为了解决现有技术中存在的上述问题,本发明提供了一种早晚高峰交通拥堵指数

的预测方法,可以准确预测交通状况,为未来的交通管理提供依据。

[0008] 为了解决上述技术问题,本发明采用了如下技术方案:

[0009] 一种早晚高峰交通拥堵指数的预测方法,包括以下步骤:

[0010] a. 将预测天分为工作日和休息日;其中

[0011] b. 采用因素分解模型对正常天气的工作日的早晚高峰的拥堵指数进行预测,当出现雨、雪、雾的极端天气时,在正常的工作日的早晚高峰的拥堵指数预测的基础上用天气影响系数 ξ 做修正;

[0012] c. 采用双模式临近模型对正常天气的休息日的早晚高峰的拥堵指数进行预测,当出现雨、雪、雾的极端天气时,在正常的休息日的早晚高峰的拥堵指数预测的基础上用天气影响系数 ξ 做修正。

[0013] 进一步,所述的因素分解模型中以不同工作日和机动车数量为两个影响因素。

[0014] 进一步,采用因素分解模型对正常天气的工作日的早晚高峰的拥堵指数进行预测的过程如下:首先,参考预测周次前一定时期的工作日的早晚高峰交通拥堵指数的数据,并分别计算平均值;然后根据预测的不同工作日及机动车数量变化情况,得到工作日系数和机动车数量系数,并用平均值分别与工作日系数和机动车数量系数相乘,得到两个值;最后结合通过实际数据得到的最优权重比,通过下式 1 得到早晚高峰交通拥堵指数的预测值 a' ;

[0015] $a' = \lambda a \beta_i + (1 - \lambda) a \delta_j$ 式 1

[0016] 式中, a 为参考的交通拥堵指数的平均值,

[0017] β_i 为工作日系数, $i = 1, 2, \dots, 5$, i 表示不同的工作日,

[0018] δ_j 为机动车数量系数, $j = 1, 2, \dots, 5$, j 表示不同的机动车数量情况,

[0019] λ 为工作日影响因素的权重。

[0020] 进一步,极端天气的工作日的早晚高峰交通拥堵指数的预测值 a'' 通过如下公式 2 得到:

[0021] $a'' = (\lambda a \beta_i + (1 - \lambda) a \delta_j) \xi_g$ 式 2

[0022] 式中, ξ_g 为天气因素系数, $g = 1, 2, 3$, 表示不同种类的天气因素。

[0023] 进一步,所述工作日系数 β_i 通过如下步骤得到,利用预测前一至三年不同工作日的早、晚高峰交通拥堵指数分别求周一至周五早高峰和晚高峰交通拥堵指数的平均值,并分别与周一至周五平均值的总平均值相除,得到周一至周五的工作日系数 β_i 。

[0024] 进一步,所述机动车数量系数 δ_j 通过如下步骤得到,利用预测前一至三年工作日的早、晚高峰交通拥堵指数分别求五种不同机动车数量情况的早高峰和晚高峰交通拥堵指数的平均值,并分别与不同机动车数量情况平均值的总平均值相除,得到机动车数量系数 δ_j 。

[0025] 进一步,所述双模式临近模型采用近邻点法对历史交通拥堵指数数据进行搜索,寻求与预测状态最相近的历史数据进行预测,具体步骤如下:

[0026] 1): 选定最小的近邻尺寸 k , 即 $k = k_{\min}$; 近邻尺寸为选取与当前状态向量距离最近的向量的个数;

[0027] 2): 选定最小的模式尺寸 l , 即 $l = l_{\min}$; 模式尺寸为模式向量内元素的个数;

[0028] 3): 当前的状态向量为 $V(n) = [v(n-1), \dots, v(n-1), v(n)]$, $v(n)$ 代表第 n 个时刻

的交通拥堵指数, n 为当前时刻;

[0029] 模式向量为 $P(n) = [d(n-1+1), \dots, d(n-1), d(n)]$, 模式向量反映状态向量的变化趋势特征;

[0030] 而 $t(t \geq 1+1)$ 时刻的状态向量为 $V(t) = [v(t-1), \dots, v(t-1), v(t)]$, t 为历史某时刻;

[0031] 模式向量为 $P(t) = [d(t-1+1), \dots, d(t-1), d(t)]$;

[0032] 状态向量的差分向量定义为 $S(t) = [s(t-1+1), \dots, s(t-1), s(t)]$, 其中 $s(t) = v(t) - v(t-1)$;

[0033] 模式向量如下定义

[0034]

$$d(t) = \begin{cases} -1, & \text{当 } s(t) < 0 \text{ 即 } v(t+1) < v(t) \\ 0, & \text{当 } s(t) = 0 \text{ 即 } v(t+1) = v(t) \\ 1, & \text{当 } s(t) > 0 \text{ 即 } v(t+1) > v(t) \end{cases};$$

[0035] 对时间序列 $v(1), v(2), \dots, v(n)$ 利用如下状态向量判断准则

[0036] $dsv(t) = \|V(n) - V(t)\| = \sqrt{\sum_{m=0}^t [v(n-m) - v(t-m)]^2}$; m 为 0 到 1 顺序增长的整数;

[0037] 确定当前状态向量与历史状态向量的欧式距离, 并按 $dsv(t)$ 的升序排列, 确定 $M_s (M_s \geq k)$ 个近邻点匹配向量;

[0038] 4) 由当前状态向量确定的 M_s 个最近邻点匹配向量按如下模式向量判断准则匹配

[0039] $dpv(t) = \|P(n) - P(t)\| = \sqrt{\sum_{m=0}^{t-1} [d(n-m) - d(t-m)]^2}$

[0040] 确定当前模式向量 $P(n)$ 与 M_s 个最近邻点匹配向量对应的模式向量的欧式距离, 并按 $dpv(t)$ 的升序排列, 最后确定与当前状态及模式向量均较为匹配的 k 个近邻点;

[0041] 5) 预测 $n+1$ 时刻的值 $\hat{v}(n+1)$

[0042] $\hat{v}(n+1) = v(n) + s$ 其中 $s = (\frac{1}{k}) \sum_{h=1}^k S(t, h)$

[0043] 在最近邻点中, $S(t, h) = [s(t-1+1, h), \dots, s(t-1, h), s(t, h)]$ 是对应向量 $V(t, h) = [v(t-1, h), \dots, v(t-1, h), v(t, h)]$ 的差分向量, 即 $s(t, h) = v(t, h) - v(t-1, h)$;

[0044] 6) 计算实际值与预测值间的均方误差

[0045] $RMSE = \sqrt{\frac{1}{N} \sum_{m=1}^N [v(m) - \hat{v}(m)]^2}$

[0046] 其中 $\hat{v}(m)$ 是 $v(m)$ 的预测值, N 是预测的总个数;

[0047] 7) 对于不同的模式尺寸 $1+1, 1+2, \dots, 1 \text{ max}$ 重复步骤 3) 到 6);

[0048] 8) 对于不同的近邻数 $k+1, k+2, \dots, k \text{ max}$ 重复步骤 2) 到 7);

[0049] 9) 选择最小的 RMSE 对应的近邻数与模式尺寸下预测结果作为对未来休息日的早晚高峰的拥堵指数的预测值。

[0050] 本发明的交通拥堵指数的预测方法开创性的基于历史交通拥堵指数对未来交通拥堵指数值展开预测,是对交通拥堵指数的应用拓展,同时,也在引导公众出行方面起到了更大的作用。本发明以交通拥堵指数数据为实施基础,通过对数据特性的分析,确定对工作日及周末早晚高峰的交通拥堵指数预测采用不同模型,通过预测为未来交通管理提供依据。预测的准确度较高,能准确预测交通状况。

附图说明

[0051] 图 1 为本发明的流程图;

[0052] 图 2 为本发明方法与其他几种预测方法早高峰交通拥堵指数预测效果对比图;

[0053] 图 3 为本发明方法与其他几种预测方法晚高峰交通拥堵指数预测效果对比图。

具体实施方式

[0054] 下面结合具体实施例对本发明作进一步详细描述,但不作为对本发明的限定。

[0055] 北京市现有 4 万辆浮动车数据用于监测路网交通拥堵指数,每 15 分钟计算该时段内的交通拥堵指数,每天早高峰 7:00-9:00 交通拥堵指数的平均值作为早高峰交通拥堵指数,晚高峰 17:00-19:00 交通拥堵指数的平均值作为晚高峰交通拥堵指数。现需要对未来一周的早、晚高峰交通拥堵指数进行预测。

[0056] 以北京市为例,早晚高峰交通拥堵指数的预测方法,包括以下步骤:

[0057] a. 将预测天分为工作日和休息日;其中

[0058] b. 采用因素分解模型对正常天气的工作日的早晚高峰的拥堵指数进行预测,当出现雨、雪、雾的极端天气时,在正常的工作日的早晚高峰的拥堵指数预测的基础上用天气影响系数 ξ 做修正;因素分解模型中以不同工作日和机动车数量为两个影响因素。不同工作日分为周一、周二、周三、周四、周五,机动车数量因素是不同工作日的机动车数量变化带来的影响,北京主要是不同的尾号限行组合因素,分别为机动车尾号 1 和 6、2 和 7、3 和 8、4 和 9、5 和 0,不同的尾号组合机动车数量是有差异的,如尾号为 4 和 9 的机动车数量比其他组合少 20 多万辆机动车。工作日的早晚高峰交通拥堵指数的预测包括如下步骤:

[0059] 首先,参考需要预测周次前一定时期(一般为前三周)工作日的早晚高峰交通拥堵指数的数据,并分别计算平均值 a ;

[0060] 其次,利用预测前一至三年(当然也可少于一年如六个月,或多于三年,如四年等)不同工作日(周一至周五)的早、晚高峰交通拥堵指数分别求周一至周五早高峰和晚高峰交通拥堵指数的平均值,并分别与周一至周五平均值的总平均值相除,得到周一至周五的工作日系数 β_i , $i = 1, 2, \dots, 5$, 分别表示周一至周五;同理,利用预测前一至三年工作日的早、晚高峰交通拥堵指数分别求五种不同机动车数量情况(本实施例为限行的 5 种尾号组合)的早高峰和晚高峰交通拥堵指数的平均值,并分别与不同机动车数量情况平均值的总平均值相除,得到机动车数量系数 δ_j , $j = 1, 2, \dots, 5$, 分别表示五种机动车数量情况,本实施例中为限行的五种尾号组合,根据要预测的不同工作日及机动车尾号限行情况,然后用前三周工作日的早晚高峰交通拥堵指数的平均值 a 分别乘上相应的工作日系数和机动车数量系数,得到两个值 $a\beta_i$ 和 $a\delta_j$;

[0061] 最后,根据实际数据回归分析,得到工作日系数与机动车数量(限行)系数的最优

权重比 λ , 求得最后的交通拥堵指数的预测值 a' : $a' = \lambda a \beta_i + (1-\lambda) a \delta_j$ 式 1

[0062] 上面是正常天气时的工作日的交通拥堵指数的预测, 没有考虑天气因素的影响。然而在实际生活中, 天气因素有时会显著影响交通情况, 故而在上面的正常天气时的工作日的交通拥堵指数的预测的基础上进行修正, 得到极端天气 (雨, 雪, 雾) 时的工作日的交通拥堵指数的预测数据。经过天气因素系数 $\xi_g (g = 1, 2, 3)$, 分别表示不同的天气因素) 的修正后, 可以得到极端天气的工作日的交通拥堵指数的预测值 a'' :

[0063] $a'' = (\lambda a \beta_i + (1-\lambda) a \delta_j) \xi_g$ 式 2

[0064] c. 采用双模式临近模型对正常天气的休息日的早晚高峰的拥堵指数进行预测, 当出现雨、雪、雾的极端天气时, 在正常的休息日的早晚高峰的拥堵指数预测的基础上用天气影响系数 ξ 做修正。

[0065] 作为本实施的优选, 双模式临近模型采用近邻点法对历史交通拥堵指数数据进行搜索, 寻求与预测状态最相近的历史数据进行预测。具体实施方法如下:

[0066] 1) 选定最小的近邻尺寸 k , 即 $k = k_{\min}$; 近邻尺寸为选取与当前状态向量距离最近的向量的个数; 本实施例中 $k_{\min} = 3$, 即按照距离最近排序至少取前 3 个;

[0067] 2) 选定最小的模式尺寸 l , 即 $l = l_{\min}$; 模式尺寸为模式向量内元素的个数; 本实施例中 $l_{\min} = 2$, 即连续取历史交通拥堵指数的个数至少为 2;

[0068] 3) 当前的状态向量为

[0069] $V(n) = [v(n-1), \dots, v(n-1), v(n)]$, $v(n)$ 代表第 n 个时刻的交通拥堵指数, n 为当前时刻;

[0070] 模式向量反映状态向量的变化趋势特征, 模式向量为

[0071] $P(n) = [d(n-1+1), \dots, d(n-1), d(n)]$, $P(n)$ 为当前状态向量的模式向量, 由状态向量差分的正负号决定的, 因此模式向量比状态向量少一维。

[0072] 而历史某时刻 $t (t \geq 1+1)$ 时刻的状态向量为

[0073] $V(t) = [v(t-1), \dots, v(t-1), v(t)]$, $V(t)$ 为历史状态向量,

[0074] 模式向量为

[0075] $P(t) = [d(t-1+1), \dots, d(t-1), d(t)]$, $P(t)$ 为历史状态向量的模式向量。

[0076] 状态向量的差分向量定义为 $S(t) = [s(t-1+1), \dots, s(t-1), s(t)]$, 其中 $s(t) = v(t) - v(t-1)$; 模式向量如下定义

[0077]

$$d(t) = \begin{cases} -1, & \text{当 } s(t) < 0 \text{ 即 } v(t+1) < v(t) \\ 0, & \text{当 } s(t) = 0 \text{ 即 } v(t+1) = v(t) \\ 1, & \text{当 } s(t) > 0 \text{ 即 } v(t+1) > v(t) \end{cases}$$

[0078] 模式向量是由状态向量差分的正负号决定的, 因此模式向量比状态向量少一维;

[0079] 对时间序列 $v(1), v(2), \dots, v(n)$ 利用如下状态向量判断准则

[0080] $dsv(t) = \|V(n) - V(t)\| = \sqrt{\sum_{m=0}^t [v(n-m) - v(t-m)]^2}$

[0081] 确定当前状态向量与历史状态向量的欧式距离, 并按 $dsv(t)$ 的升序排列, 确定 $M_s (M_s \geq k)$ 个近邻点匹配向量;

[0082] 4) 由当前状态向量确定的 M_s 个最近邻点匹配向量按如下模式向量判断准则匹配

$$[0083] \quad dpv(t) = \|P(n) - P(t)\| = \sqrt{\sum_{m=0}^{t-1} [d(n-m) - d(t-m)]^2}$$

[0084] 确定当前模式向量 $P(n)$ 与 M_s 个最近邻点匹配向量对应的模式向量的欧式距离，并按 $dsv(t)$ 的升序排列，最后确定与当前状态及模式向量均较为匹配的 k 个近邻点；

[0085] 5) 预测 $n+1$ 时刻的值 $\hat{v}(n+1)$

$$[0086] \quad \hat{v}(n+1) = v(n) + s \text{ 其中 } s = \left(\frac{1}{k}\right) \sum_{h=1}^k S(t, h)$$

[0087] 在最近邻点中， $S(t, h) = [s(t-1+1, h), \dots, s(t-1, h), s(t, h)]$ 是对应向量 $V(t, h) = [v(t-1, h), \dots, v(t-1, h), v(t, h)]$ 的差分向量，即 $s(t, h) = v(t, h) - v(t-1, h)$ ；

[0088] 6) 计算实际值与预测值间的均方误差

$$[0089] \quad RMSE = \sqrt{\frac{1}{N} \sum_{m=1}^N [v(m) - \hat{v}(m)]^2}$$

[0090] 其中 $\hat{v}(m)$ 是 $v(m)$ 的预测值， N 是预测的总个数；

[0091] 7) 对于不同的模式尺寸 $1+1, 1+2, \dots, 1_{\max}$ 重复步骤 3) 到 6)，本实施例中 $1_{\max} = 96$ ；

[0092] 8) 对于不同的近邻数 $k+1, k+2, \dots, k_{\max}$ 重复步骤 2) 到 7)，本实施例中 $k_{\max} = 20$ ；

[0093] 9) 选择最小的 RMSE 对应的近邻数与模式尺寸下预测结果作为对未来休息日的早晚高峰的拥堵指数的预测值。

[0094] 在步骤 3) 与 4) 中，可以根据实际情况赋予不同的权重 ϖ_m 和 θ_m ，因为历史数据对当前数据的影响会随着间隔的增加而减小，因此一般来说 ϖ_m 和 θ_m 升序排列，即距离当前时刻越近权重越大，距离当前时刻越远权重越小。

[0095]

$$wdsv = \|V(n) - V(t)\| = \sqrt{\sum_{m=0}^l \varpi_m [v(n-m) - v(t-m)]^2} \left(\sum_{m=0}^l \varpi_m = 1 \right)$$

$$[0096] \quad wdpv = \|P(t-1) - P(n-1)\| = \sqrt{\sum_{m=0}^l \theta_m [p(n-m) - p(t-m)]^2} \left(\sum_{m=0}^l \theta_m = 1 \right)$$

[0097] 同样可以训练 ϖ_m 和 θ_m 搜索更加匹配的历史数据，进而更好的预测。

[0098] 对比本实施例结果与滑动平均模型、双指数平滑模型、ARIMA 模型等方法一周早高峰数据预测结果。如下表 1 所示。

[0099] 表 1

[0100]

	周一	周二	周三	周四	周五	周六	周日
实际值	1.28	7.5	5.4	5.1	5.4	5	1.36

滑动平均	4.11	4.08	4.85	4.58	4.58	4.59	4.49
双指数平滑	4.42	4.42	4.43	4.43	4.43	4.43	4.43
ARIMA 模型	1.99	6.93	5.18	4.94	5.31	4.68	2.21
本发明方法	1.34	6.09	6.08	4.93	5.52	4.52	1.38

[0101] 下表 2 为早高峰各种预测统计量对比：

[0102] 表 2

[0103]

	滑动平均	双指数平滑	ARIMA 模型	本发明方法
平均绝对误差 (MAE)	1.67	1.78	0.42	0.42
平均绝对百分比 误差(MAPE)	77.2%	81.8%	20.1%	7.5%
均方误差 (MSE)	4.21	4.32	0.24	0.36
均方根误差 (RMSE)	2.05	2.08	0.49	0.60

[0104] 通过上表可以看出对于早高峰预测结果,本发明方法预测值与真实值较为接近(MAE 和 MAPE 较小),但预测的波动性较 ARIMA 模型的预测结果稍微大些(RMSE 较大些)。

[0105] 对比本实施例结果与滑动平均模型、双指数平滑模型、ARIMA 模型等方法一周晚高峰数据预测结果。如下表 3 所示。

[0106] 表 3

[0107]

	周一	周二	周三	周四	周五	周六	周日
实际值	5.78	8.1	7.5	7.3	8.3	8.9	7.5
滑动平均	7.52	7.59	7.52	7.56	7.49	7.53	7.62
双指数平滑	7.69	7.59	7.7	7.7	7.7	7.7	7.71
ARIMA 模型	5.67	8.5	7.36	7.7	7.94	8.37	7.32

本发明方法	5.78	8.1	7.65	7.05	8.55	8.1	7.53
-------	------	-----	------	------	------	-----	------

[0108] 下表 4 为晚高峰各种预测统计量对比：

[0109] 表 4

[0110]

	滑动平均	双指数平滑	ARIMA 模型	本发明方法
平均绝对误差 (MAE)	0.69	0.72	0.30	0.21
平均绝对百分比误差(MAPE)	9.6%	10.1%	3.8%	2.5%
均方误差(MSE)	0.84	0.85	0.11	0.11
均方根误差(RMSE)	0.92	0.92	0.34	0.34

[0111] 对于晚高峰预测结果,本发明方法预测值与真实值较为接近 (MAE 和 MAPE 较小),但预测的波动性较 ARIMA 模型的预测结果类似,整体来说本发明方法预测结果最好。

[0112] 图 2 和图 3 分别为本发明方法与其他几种预测方法晚高峰交通拥堵指数预测效果的对比图。

[0113] 经对比,本发明方法的预测精度综合来看最高,平均绝对误差 (MAE) 和平均绝对百分比误差 (MAPE) 比滑动平均和双指数平滑低很多,与之接近的 ARIMA 模型虽拥有较高的预测精确度,但是需要复杂的参数估计,而且计算出的参数不能移植。在实际情况中,经常由于各种各样的原因容易造成数据遗漏,导致模型精确度降低,而且实时计算效率低。另外,ARIMA 模型特别适用于稳定的交通流,交通状态变化急剧时,由于计算量过大,该模型将在预测延迟方面暴露出明显的不足,针对突发的恶劣天气也无法快速反映其影响。因此,本发明方法适用于实时进行预测应用。

[0114] 以上实施例仅为本发明的示例性实施例,不用于限制本发明,本发明的保护范围由权利要求书限定。本领域技术人员可以在本发明的实质和保护范围内,对本发明做出各种修改或等同替换,这种修改或等同替换也应视为落在本发明的保护范围内。

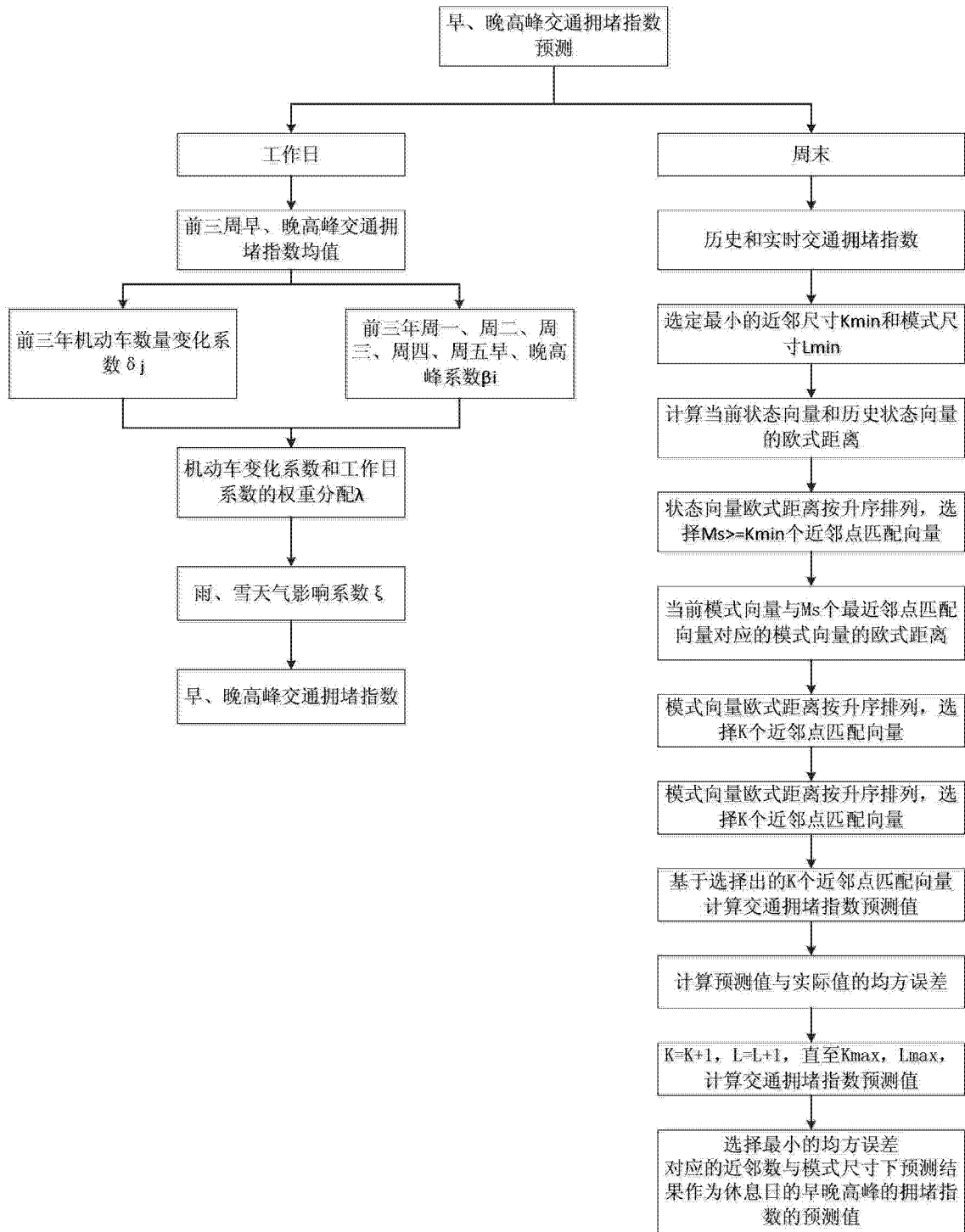


图 1

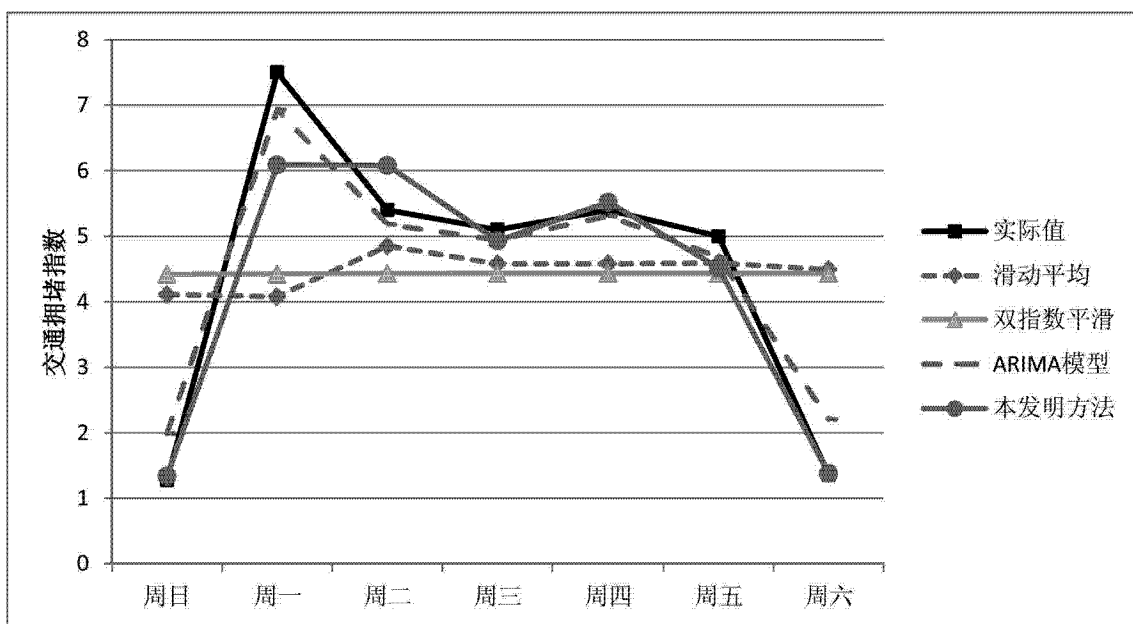


图 2

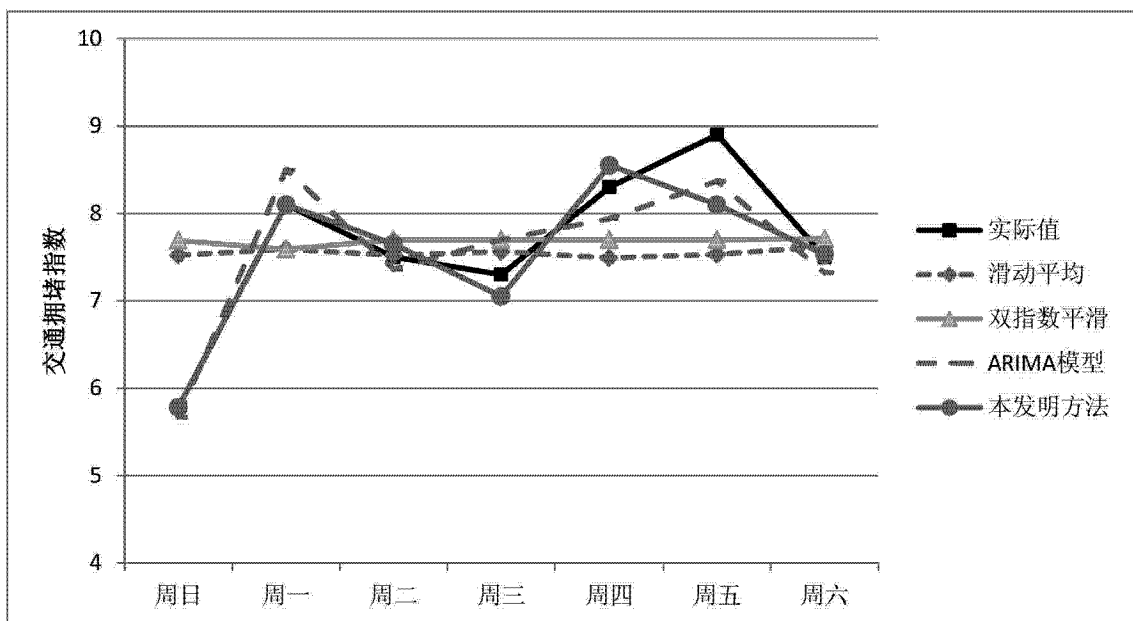


图 3