

城市道路交通拥堵状态时空相关性分析

张婧^{1,2}, 任刚^{*1,2}

(1. 东南大学城市智能交通江苏省重点实验室, 南京 210096; 2. 现代城市交通技术江苏高校系统创新中心, 南京 210096)

摘要: 城市交通拥堵的形成和消散实际受多周期的交通流冲击波影响, 产生来源非常复杂, 建模也很困难. 正因如此, 有关拥堵时空扩散性的研究多停留在定性层面. 基于直接采集交通数据可进行定量分析, 但数据的细粒度特性使信号不够平稳, 在多时间尺度上甚至表现出相反特征, 缺乏有效的知识挖掘思路和方法. 针对拥堵区域交通参数时空相关性问题, 在皮尔逊相关性指标的基础上提出并采用了一种新的分析方法, 它将道路实测速度轨迹分解为多时间尺度的趋势和细节分量. 应用该方法提出的相关性指标和数据分段算法, 对典型交通拥堵扩散算例进行了分析, 借助相关性状态变化刻画了拥堵在时空中扩散的定量特征.

关键词: 城市交通; 时空相关性; 相关性分析; 交通拥堵; 拥堵扩散

Spatio-temporal Correlation Analysis of Urban Traffic Congestion Diffusion

ZHANG Jing^{1,2}, REN Gang^{*1,2}

(1. Jiangsu Key Laboratory of Urban ITS, Southeast University, Nanjing 210096, China; 2. Jiangsu Province Collaborative Innovation Center of Modern Urban Traffic Technologies, Nanjing 210096, China)

Abstract: The formation and dissipation processes of urban traffic congestions are influenced by shocking waves of different cycles inside the traffic flow. The original factors that lead to traffic congestions are very complicated, the modelling is therefore difficult. This is the main reason that research works about the spatio-temporal dissipation effects for congestions are eventually stopped at the level of qualitative analysis. Some quantitative analysis can be successfully done based on the measured traffic data. However, rare effective knowledge extraction methods can be found to deal with data containing information about multiple time scales and granularities, which however is important to correlation analysis and the direct use of original data leads to unstationary signal features and opposite observation conclusions when putting the data into the discussion of given time scales. Focused on the analysis of spatio-temporal correlation of traffic parameters in congestion areas, a new analyzing method is developed and used based on Pearson's correlation index, which decomposes a measured road speed trajectory into trend and detail components in different time scales. The initial verification and application of this method and the corresponding data segmentation algorithm show the quantitative characteristics of the congestion diffusion in time and space by observing the variation of correlation status.

Key words: urban traffic; spatio-temporal correlation; correlation analysis; traffic congestion; congestion diffusion

收稿日期: 2014-11-20

修回日期: 2015-02-03

录用日期: 2015-03-03

基金项目: 高等学校博士学科点专项科研基金(20120092110043).

作者简介: 张婧(1983-), 女, 重庆人, 博士生.

*通信作者: rengang@seu.edu.cn

1 引言

交通拥堵在时空中具有扩散性,这已得到学界公认.它理应表现为相关时空单元状态参数相关性的规律性变化,从相关性分析入手解析拥堵的机理不失为一个思路.许多研究从空间或时间自相关分析方法出发,希望建立交通状态的时空相关性分析方法.在扩展空间相关性方法上,陈绍宽等在传统的Moran's I指数中引入时间维度^[1],来分析城市道路交通状态的时空变化规律.应用非常广泛的皮尔逊相关系数是针对两组数据之间的相关性分析指标,也被用于时间自相关分析^[2],并被扩展应用于交通时空相关性分析.基于该系数,卡尔曼滤波^[3]和时空ARMA模型^[4]实现了道路状态的预测建模,更能适应突发情况,预测效果较好.皮尔逊相关系数也被直接用于城市网络负荷与交通流的关系^[5],空间邻接关系对伦敦路网的行车时间影响^[6]等道路相关性研究.

皮尔逊相关系数作为一类线性相关性判别指标,表述简单,应用场景具有鲁棒性,这也是它受到广泛欢迎的原因.然而皮尔逊系数用于非平稳、非线性信号的相关性分析时存在较大的局限性.尝试对非平稳信号进行分解,并利用数据分段技术,进一步提高各数据段的线性度,引入扩展相关性

指标,并将其用于拥堵过程的实例分析,从而刻画拥堵扩散的时空特征.

2 拥堵过程的相关性分析

2.1 拥堵过程的机理分析

城市道路交通运行状态可以分为畅通、缓行和拥堵状态等三类.当交通流处于畅通状态时,各车道车辆顺畅行驶,速度较高,道路通行能力最高;缓行时,各车道车辆以相近速度缓慢行驶,速度降低,道路通行能力降低;拥堵时,车辆以较低速度排队行驶,并频繁出现走一停现象,此时速度最低,道路通行能力最低.

基于对道路交通速度、流量、密度等参数的观测和合理推论,拥堵现象同时具有时间和空间的扩散性.从时间上看,拥堵过程可分为4个阶段:拥堵产生、扩散、消散和结束.当到达车流量超过瓶颈通行能力时,拥堵产生并逐渐扩散;当到达车流量小于离去车流量时,拥堵从扩散转为消散.从空间上看,道路出现拥堵后,通行能力下降的截面会随着车辆排队的增加向上游运动^[7].如图1所示,当拥堵的冲击波从瓶颈处向上游扩散时,更多路段的通行能力和速度也随着截面的运动而下降.

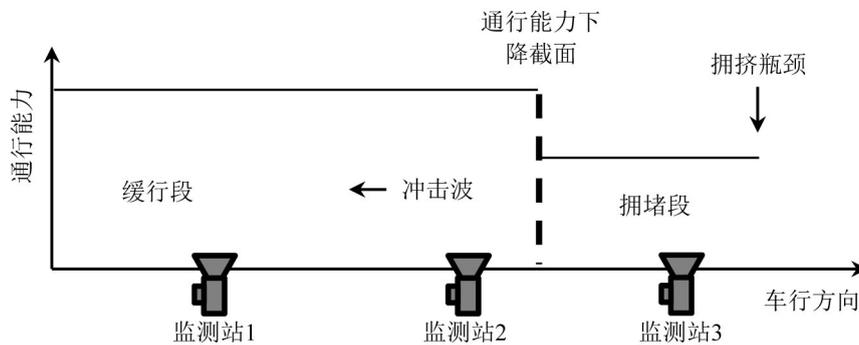


图1 交通拥堵扩散示意图

Fig. 1 The traffic congestion propagation schematic plot

当监测站同时位于图1的缓行段或拥堵段(如监测站1和监测站2)时,道路流量、速度、占有率等参数的相关性较强;反之,当监测站位于通行能力下降截面的两侧(如监测站2和监测站3)时,相关性较弱或不定.随着冲击波的不断产生和减弱、通行能力下降截面不断向上游运动,监测站数据之间的相关关系也处于动态变化中.然而,相关的定

量机理研究还较少.

2.2 基于传统指标的相关性分析

相关性分析是指对两个或多个有序向量,研究其对应元素变化特征的相似性.皮尔逊相关系数(下文简称为 R 系数)被定义为两个向量的协方差和标准差的乘积之比,如式(1)所示.信号的线性相关程度越高, R 值越显著,因此 R 也常被称为线性

相关度指标.

$$R = \frac{\text{cov}(x,y)}{\sigma_x \sigma_y} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (1)$$

式中 x_i 、 y_i 表示随机变量 x 、 y 的测量值; \bar{x} 、 \bar{y} 表示 x 、 y 的均值; n 为序列个数.

显然, R 在 $[-1, 1]$ 内取值, R 的绝对值越大, 线性相关性越强. 当 x 、 y 取 2 条道路同一段时间内的数据序列时, R 系数表示不同道路之间的相关性; 当 x 、 y 取同一道路不同时段的数据时, R 系数亦可用于时间自相关研究.

但是当参数轨迹具有非平稳, 非线性性质的时候, 皮尔逊相关系数并不一定能准确体现参数之间的相关关系. 例如, 图 2 是某 2 条相邻道路实测 6 min 平均车速的轨迹, 在所取的时间窗口内, 上游车速逐渐降低, 下游车速逐渐增加, 两者整体趋势相反, 但波动细节上大体相似. 然而, 根据式 (1) 所计算的 2 条道路的 R 系数为 0.028, 总体趋势和波动细节上呈现出的显著相关性被相互抵消.

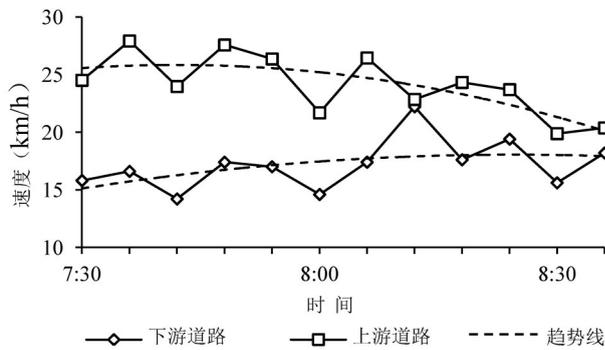


图 2 皮尔逊相关系数无法显著反映相关性的典型场景

Fig.2 A typical scenario that Pearson's correlation index gives insignificant indication

3 基于自适应分段的扩展相关性模型

3.1 扩展相关性模型

3.1.1 道路状态原始数据的归一化处理

在进行拥堵评价时, 道路速度获取容易、拥堵判断较为直观, 是采用频率最高的参数^[1], 本文基于速度数据开展相关性信息的深度挖掘研究.

采用实测速度, 不利于区分不同道路等级、限速、拥堵临界值对测量值的影响. 对速度进行归一化处理, 不同等级道路的测量值将具有可比性. 鉴

于研究者更关心拥堵临界值附近及其以下的情况, 应将自然速度映射到有利于拥堵研究的区间. 在归一化后的映射空间里, 拥堵临界值设为 0, 道路最高限速 (包括高于限速的测量值) 设为 1, 自然速度为 0 对应 -1 . $[-1, 0)$ 和 $(0, 1]$ 采用线性方法完成归一化映射.

3.1.2 城市交通流的相关性指标提取

对动态交通数据的分析既要掌握变化细节, 也要分清变化趋势. 当研究交通拥堵数据的关联关系时, 首先应判断观察窗口内 2 组轨迹点组成趋势的一致性, 此时忽略个别时段的突变点. 然后通过计算细节相关度发现局部相似波形, 以掌握拥堵形成过程更丰富的信息. 在趋势相关性计算中被忽略的突变点通常源于图 1 中体现的冲击波的周期性影响, 乃至多个波之间的衍射影响. 这些影响在细节相关性分析中能得到体现.

考虑到 R 系数更适合处理线性程度较高的信号, 在应用 R 系数进行分析时有必要对包含过多细节信息的数据进行处理, 在考虑总体趋势时应避免波动细节的干扰, 反之亦然. 对任意给定观察窗口中的信号 y , 应借助合适的分解方法将其分解为具有低阶可严格逼近的趋势分量和余下接近采样周期尺度的细节分量, 如式 (2) 所示.

$$y = y_\alpha + y_\beta \quad (2)$$

式中 y_α 表示趋势分量; y_β 表示细节分量.

本文采用二次多项式最小二乘法逼近求取趋势分量, 即:

$$y_\alpha = p_2(x) = a_2 x^2 + a_1 x + a_0 \quad (3)$$

$$\min(I) = \min\left\{\sum_{i=0}^n [p_2(x_i) - y_i]^2\right\} \quad (4)$$

$$\frac{\partial I}{\partial a_j} = 2 \sum_{i=0}^n (\sum_{k=0}^2 a_k x_i^k - y_i) x_i^j = 0 \quad (5)$$

其中, 对于二次型逼近, $j=0, 1, 2$. 在求得趋势分量后, 细节分量即为信号与其趋势分量之差. 将相关性高记作 H, 反之记作 L; 将趋势相关性 R_t 置前, 细节相关性 R_d 置后, 则趋势与细节相关性相互组合, 可以得到 4 种类型: HH、HL、LH、LL.

3.1.3 基于扩展相关性指标的拥堵分析

对拥堵过程的定量分析依赖于对交通瓶颈和冲击波的定量描述, 这两者的模型都不易掌握, 因此构成拥堵研究方面的难题. 如 1.1 节所述, 冲击波

引起的通行能力下降截面(以下简称冲击截面)的移动是造成道路交通参数相关性变化的原因.反过来,对交通参数相关性的分析能揭示冲击波及其截面的部分特征,间接完成了对其建模.

进一步分析,当具有相关性的2条道路交通参数呈现出明显的趋势和细节相关,即处于HH状态时,表示相关路段正受到共同因素制约,形成相似的状态模式,本身具有一定的抗干扰能力,在路段中间没有冲击截面.因此以下将HH状态称作“相关性平衡态”.

当道路交通量的趋势相关性依然很高,但细节相关性陡然降低,即HL状态时,表示局部交通流的稳定性已经被打破,如下游出现了交通状态突变,但上游暂时还没受到影响.因此,以下也将HL状态称为“平衡破坏态”.

基于短时交通流数据,并结合上述相关性分析指标,可进一步实例解读整个拥堵扩散形成过程,包括主要时间节点及拥堵的其它特征.在此之前,仍需解决R系数用于分析非线性信号显著度不高的问题.

3.2 自适应的分段优化算法

日间道路交通数据呈现周期性变化,尤以工作日间为突出,但日内交通数据为非周期或变周期数据.城市路段的日内交通量尤其波动较大,即使其趋势分量的非线性程度也较强.分段过少,段内数据的非线性程度高,不利于相关性知识的挖掘,相关性结论可能不显著;分段过多,有关全局的知识就越少,不利于做出整体判断.因此应尽可能用最少的分段获得最显著的结论.

对一日数据进行分段,并在各分时段内对道路通行情况及其相关性进行分析,可避免不同时段数据之间的干扰.分段后,各段可独立进行相关性分析,相关性指标的显著度也将随分段方法而不同.

采用二次多项式拟合观察窗口中时间序列的变化趋势,当二次函数的2阶系数 a_2 越小时,说明拟合趋势的线性化程度越高.通过调整观察窗口的起止,可能减小待分析的2条复杂轨迹在窗口内段落 $|a_2|$ 的值,但同时也可能增加观察窗口以外段落 $|a_2|$ 的值,还可能增加趋势线与原序列的标准

差.因此,应以最小化各段非线性度 $(|a_2|)$ 及其与原序列标准差的加权和为优化目标,通过摄动分段方案的划分向量,利用灵敏度方法确定搜索方向,迭代搜索目标函数的准优解.应使搜索满足一定的约束,如分段处需保持一定的连续性,预设最小分段长度,并适时合并搜索中的临时分段,以避免分段过多和因人为限制搜索到局部最优解.

式(6)为分段优化算法的目标函数,式(7)为各段拟合函数保持连续性的约束条件:

$$\min(O) = \min \sum_{d=1}^2 \sum_{i=1}^n [|a_{2,i}(d)| + \gamma e_i(d)] \quad (6)$$

$$\text{s.t. } p_2^{(i)}(x_{\text{end}}^{(i)}) = p_2^{(i+1)}(x_{\text{start}}^{(i+1)}), \quad 1 \leq i < n \quad (7)$$

式中 $a_{2,i}(d)$ 表示数据序列 d 分段 i 的趋势线2阶拟合系数 a_2 ; $e_i(d)$ 为序列 d 分段 i 的拟合标准差; γ 为权系数; $p_2^{(i)}$ 为分段 i 趋势分量的二次逼近函数; $x_{\text{start}}^{(i)}$ 、 $x_{\text{end}}^{(i)}$ 分别表示分段 i 的起、止时间.

式(6)、式(7)是带等式约束条件的优化问题,为了简化目标函数的求解,将约束条件以罚函数的形式转化进目标函数,结果如式(8)所示.其中, β 为不同于 γ 的权系数.在下列算例中, β 取0.1、 γ 取1.1.

$$\min(O') = \min \sum_{d=1}^2 (\sum_{i=1}^n [|a_{2,i}(d)| + \gamma e_i(d)] + \beta \cdot \sum_{j=1}^{n-1} [p_2^{(j)}(x_{\text{end}}^{(j)}) - p_2^{(j+1)}(x_{\text{start}}^{(j+1)})]) \quad (8)$$

优化求解所采用的灵敏度搜索算法步骤如下:

步骤1 输入初始分段长度 l_{init} ,基于固定长度分段法确定初始分段方案,输入权系数 γ 、 β 、搜索约束条件最小分段长度 l_{min} 、搜索步长 l_d ,置迭代步 $k=0$;

步骤2 在搜索迭代步 k ,首先将各分段位置记为向量 S_k ,按式(8)计算当前分段方案的目标函数 O'_k .

步骤3 对 S_k 中每个元素按搜索步长 $\pm l_d$ 确定搜索试探量,并按式(8)计算每次试探的目标函数值,以目标函数值最小的试探量作为目标函数最速下降方向,将该方向的划分向量记为 S_{k+1} ,目标函数记为 O'_{k+1} ;

步骤4 迭代终止条件为: $O'_{k+1} > O'_k$ 或达到最大迭代次数,输出最终划分向量 O'_k .如果该条件并不满足,置 $k=k+1$,返回步骤2.

4 实例分析

交通状态的量化分析是研究缓堵措施优化的前提.将上述指标和分析算法分别应用于道路交通参数的时间和空间相关性分析,通过相关性的量化进一步揭示拥堵机理并验证所提指标的合理性.

4.1 相邻2日道路拥堵状态的相关性对比

以下算例将本文3.1.2节提出的扩展相关性指

标中的趋势相关性 R_t 应用于2日道路状态的时间相关性分析,并与传统的皮尔逊相关系数 R 进行了比较.

图3为相邻2个工作日同一路段的6 min平滑实测车速数据,直观显示了2日交通拥堵状态具有显著的时间一致性,但拥堵段平均行驶速度和分钟级细节仍存在很大差异.

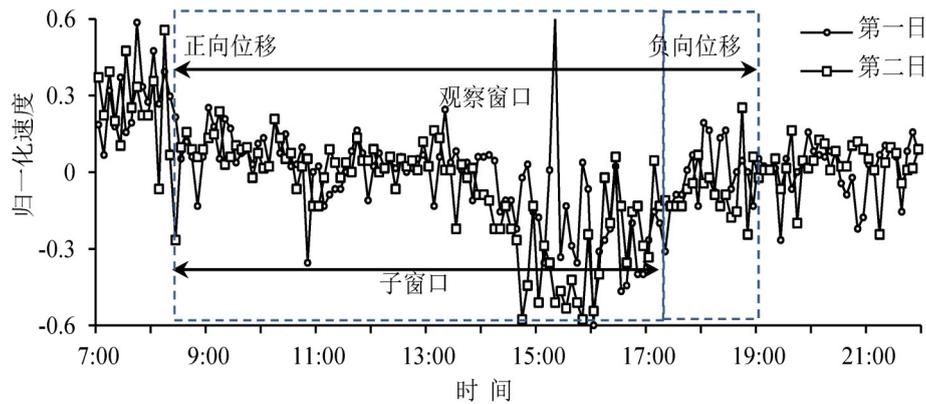


图3 相邻2日的实测行车速度比较

Fig.3 Measured vehicle speed in 2 adjacent days for the same road

选取第二日观测窗口内的一段缓行—拥堵轨迹(窗口段)为研究对象(图3中带方形标记的轨迹),通过平移第一日(带圆形标记)轨迹的观察窗口,研究平移对两者时空相关性的影响.所采用的相关性指标包括皮尔逊相关系数 R 和趋势相关度 R_t ,其中 R_t 根据图3轨迹形态选用2次型拟合,如图4所示.

者相关性最高的结论,最大值为0.5,显著度和位置判断精度均不及 R_t .

4.2 拥堵在相邻道路间扩散的相关性分析

以下算例借助3.1.2节提出的扩展相关性指标和3.2节提出的自适应分段算法,对交通拥堵扩散过程进行了定量分析.

该算例基于某市交通管理部门实测2 min道路均速.考虑到交叉口信号灯对数据波动性的周期性影响,以及道路交通延时给相关性分析可能带来的误差,将原始数据经平滑整理为6 min为间隔的速度信息.在应用自适应分段算法后,各段趋势相关性均为高值.

当交通流未受较大扰动影响时,相邻道路一般体现为正相关关系,此时 R_t 和 R_d 均为正向高值,即相关性平衡态(HH态),如图5中13:25-14:01时段的状态所示.14:01之后,受扰的下游道路拥堵明显加剧,上游道路则似乎未受影响,继续维持缓行状态.从明显下降的 R_d 指标(图5中由0.8降至0.1左右)可知,动态交通流的平衡态已被打破,冲击波断面正位于两者之间,扩散正在进行,即拥堵将经由车辆排队过程向上游道路传递.

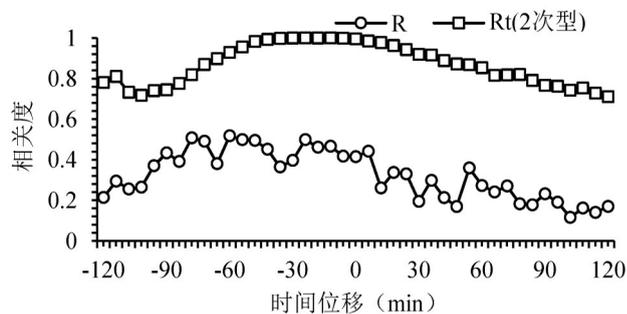


图4 相邻2日交通数据的相关性分析

Fig.4 The correlation of traffic data in 2 adjacent days

图4中,横坐标表示第二日轨迹相对于第一日轨迹的时间平移量,第二日窗口平移方向如图3所示.与皮尔逊相关系数相比,当第二日窗口向右平移0-30 min范围内,两者的 R_t 最高,这与直观判断匹配得更好; R 系数得出向右平移78 min时两

随着排队过程的进行,从14:31左右开始,上下游道路被重新“关联”在一起,在图5的算例中呈现接近-1的负相关性,对应冲击波断面已通过2条道路,可能继续向更远路段扩散传播.之所以表现为显著负相关性,与严重拥堵导致的走—停周期交替发生有关.

拥堵消散过程则与此类似,15:25-16:07,2条道路 R_d 指标的绝对值再次明显下降(相关度从-1变至0.2),处于新的“平衡破坏态”,道路速度在达到最低点(15:37)后,逐渐回升.在16:07之后,2条道路的 R_d 指标恢复到0.5左右,回到相关性平衡

态,再次进入正相关.

依靠直接观察,一般仅能识别拥堵扩散或消散结束后的阶段,此时速度已稳定在明显的低(高)值,无法觉察“平衡破坏态”,而恰巧该状态刻画了拥堵扩散的时空定量特征.采用本文方法计算的指标更加显著和平稳,可用于拥堵扩散和消散过程的分析.在扩展相关性指标显示的“平衡破坏态”所处时段(即算例中的14:01-14:37和15:25-16:07),此时趋势相关度较高、细节相关度较低.当拥堵瓶颈引起的冲击波停止产生或扩散过程已通过观测路段后,细节相关度的绝对值将显著升高.

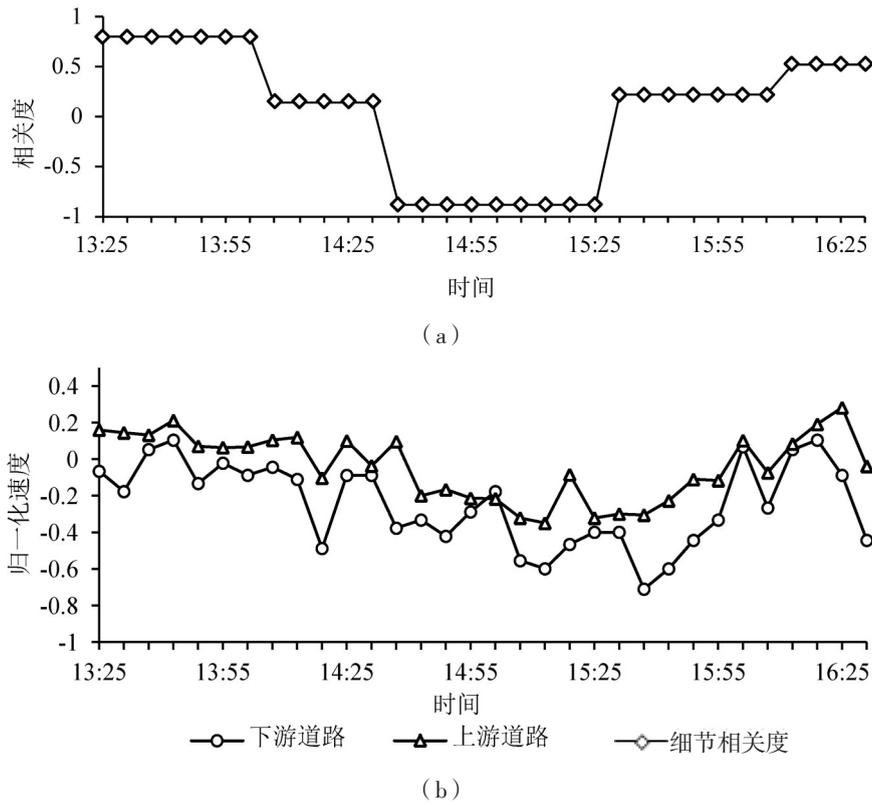


图5 拥堵扩散和消散过程与细节相关度的关系

Fig.5 The relationship between the phases of congestion diffusion and dissipation and the detail correlation degree

5 研究结论

针对交通拥堵的动态参数轨迹的非平稳特征,本文提出了扩展相关性指标,将参数轨迹分解为趋势和细节分量,引入了趋势和细节相关度概念.通过对相邻2日道路状态进行的相关性分析,其结果显示趋势相关度相较于传统指标,对时间维度上相关性信息的表达更为显著.针对皮尔逊相关系数应用于非平稳交通参数轨迹分析上的局限

性,本文进一步提出了自适应分段优化算法,将一日交通数据分为具有较高趋势相关性的多段,利用趋势和细节相关度的不同组合状态对道路拥堵发展过程进行分析.在此基础上,利用细节相关度对拥堵向上游扩散的现象进行了实例分析,发现了相关性突变前存在的平衡破坏态.揭示了其机理,给出了扩散中由冲击波引起的通行能力下降的时间节点,实现了拥堵时空相关性的定量分析,

对深入研究拥堵形成机理、实现拥堵预警和早期干预都将具有重要意义。

参考文献:

- [1] 陈绍宽, 韦伟, 毛保华, 等. 基于改进时空 Moran's I 指数的道路交通状态特征分析[J]. Acta Phys. 2013, 62 (14). [CHEN S K, WEI W, MAO B H, et al. Analysis on urban traffic status based on improved spatio-temporal Moran's I[J]. Acta Phys, 2013, 62(14).]
- [2] YANG Y, AGO Yeh. Spatio temporal traffic-flow dependency and short-term traffic forecasting[J]. Environment and Planning, 2008, 35(5): 762-771.
- [3] Okutani I, Stephanades Y. Dynamic prediction of traffic volume through kalman filtering theory[J]. Transportation Research, 1984, 18(1): 1-11.
- [4] Pfeifer P E, Deutsch S J. Identification and interpretation of first-order space-time ARMA models[J]. Technometrics, 1980, 22 (3): 397-403.
- [5] 赵婷婷, 张毅, 周彧, 等. 城市路网交通流的空间互相关性[J]. 清华大学学报(自然科学版), 2011, 51(3): 313. [ZHAO T T, ZHANG Y, ZHOU Y, et al. Spatial cross correlations of traffic flows on urban road network[J]. Journal of Tsinghua University (Science. & Technology.), 2011, 51(3): 313.]
- [6] 赵玲, 王佳璆, 邓敏, 等. 基于行车时间的道路网络时空自相关模型研究[J]. 中南大学学报(自然科学版), 2012, 43(10): 4114-4122. [ZHAO L, WANG J Q, DENG M, et al. Spatial-temporal autocorrelation model of road network based on travelling time[J]. Journal of Central South University (Science and Technology), 2012, 43(10): 4114-4122.]
- [7] 姜桂艳. 道路交通状态判别技术与应用[M]. 北京: 人民交通出版社, 2004. [JIANG G Y. Technologies and applications of the identification of road traffic conditions[M]. Beijing: China Communications Press, 2004.]
-
- 上接第174页
- [9] Cristian E C, Sergio J D, Alejandro T. Intergrating short turning and deadheading in the optimization of transit services[J]. Transportation Research Part A, 2011, 45, 419-434.
- [10] 巫威眺, 靳文舟, 魏明, 等. 配合区间车的单线公交组合调度模型[J]. 华南理工大学自然科学版, 2012, 40 (11): 147-154. [WU W T, JIN W Z, WEI M, et al. Mixed scheduling model of single-line bus for zonal vehicles [J]. Journal of South China University of Technology (Natural Science Edition), 2012, 40 (11): 147-154.]
- [11] 巫威眺, 靳文舟, 任然. 单线公交车辆组合调度与购车计划的双层规划模型[J]. 吉林大学学报(工学版), 2013, 43(05): 1196-1203. [WU W T, JIN W Z, REN R. Bi-level programming model for transit mixed scheduling and procurement scheme[J]. Journal of Jilin University (Engineering and Technology Edition), 2013, 43(05): 1196-1203.]
- [12] 窦慧丽, 刘好德, 杨晓光. 基于站点上下客人数的公交客流 OD 反推方法研究[J]. 交通与计算机, 2007, 25(2): 79-83. [DOU H L, LIU H D, YANG X G. OD matrix estimation method of public transportation flow based on passenger boarding and alighting[J]. Computer and Communications, 2007, 25(2): 79-83.]