DOI: 10.13873/J.1000-9787(2021)02-0019-04

交通流多时间尺度特性分析与参数模型研究*

曹　洁[1,2,3], 孙兆军[1], 张　红[1,2], 陈作汉[1,2], 侯　亮[1,2]

(1. 兰州理工大学 计算机与通信学院,甘肃 兰州 730050;
2. 甘肃省城市轨道交通智能运营工程研究中心,甘肃 兰州 730050;
3. 甘肃省制造业信息化工程研究中心,甘肃 兰州 730050)

摘　要: 为揭示交通流的内在动态特性,利用符号动力学中的 Lempel-Ziv 算法计算不同时间尺度下的交通流时间序列复杂度,并基于相关性理论分析方法,探索了交通流在时间和空间上的相关性。针对经典的单段函数模型无法准确表征不同交通状态下交通流参数关系的问题,提出了一种二阶段交通流参数关系模型。采用 PeMS 系统采集的交通流数据,运用非线性最小二乘法对交通流量-密度关系进行拟合。实验结果表明: 本文模型的拟合性能均优于其他四种模型。从根均方误差(RMSE) 可见,模型的拟合标准误差较单段 3PL 模型平均降低了 7.6%,表明模型能较好地表示不同交通状态的参数关系,建立的参数模型对交通状态判别和短时预测具有实际的应用价值。

关键词: 交通流; Lempel-Ziv 算法; 相关性理论; 特征分析; 参数模型

中图分类号: TP391　　**文献标识码**: A　　**文章编号**: 1000-9787(2021)02-0019-04

Multi-time scale characteristic analysis and parameter model study of traffic flow*

CAO Jie[1,2,3], SUN Zhaojun[1], ZHANG Hong[1,2], CHEN Zuohan[1,2], HOU Liang[1,2]

(1. College of Computer and Communication, Lanzhou University of Technology, Lanzhou 730050, China;
2. Gansu Urban Rail Transit Intelligent Operation Engineering Research Center, Lanzhou 730050, China;
3. Gansu Manufacturing Information Engineering Research Center, Lanzhou 730050, China)

Abstract: In order to reveal the inherent dynamic characteristics of traffic flow, the Lempel-Ziv algorithm in symbol dynamics is used to calculate the time series complexity of traffic flow at different time scales. Based on the analysis method of correlation theory, the correlation of traffic flow in time and space is explored. Aiming at the problem that the classical single-regime function model can't accurately characterize the relationship of traffic flow parameter under different traffic condi-tions, and a two-stage traffic flow parameter relationship model is proposed. The traffic flow data collected by the PeMS system is used to fit the relationship of traffic flow-density with the nonlinear least squares method. The experimental results show that the fitting performance of our model is better than the other four models. It can be seen from the RMSE that the standard error of the model is 7.6% lower than that of single-regime 3PL model. Which indicates that the model can better represent the parameter relationship of different traffic states. The established parameter model has value of practical application for traffic state discrimination and short-term prediction.

Keywords: traffic flow; lempel-Ziv algorithm; theory of correlation; characteristics analysis; parameter model

0　引　言

在智能交通系统中,建立准确稳定的交通流参数模型是实现智能交通诱导与控制的关键[1]。交通流参数模型通常以描述交通流运行特性的三个参数(速度、流量和密度) 为基础,考虑到交通流数据统计时间尺度的不同,交通流序列因受到各种复杂因素的影响,表现出较强的非线性和时空特性。因此,分析交通流在不同时间尺度下的特性,揭示交通流特性参量间的相互关系和内在时空演变规律,可为完善交通流理论、进行交通状态识别提供科学依据和理论支撑。

对于交通流时间序列的分析主要有: 数理统计方法、机器学习方法、可视图方法以及复杂网络方法[2]。针对交通

收稿日期: 2019-09-09

* **基金项目**: 国家自然科学基金资助项目(61263031) ; 甘肃省高校科研项目(2016B-031)

(C)1994-2021 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

流参数模型的研究,国外起步较早,从格林希尔治提出的第一个速度-密度关系模型开始,学者基于此模型进行了大量的探索研究。这些研究大体上可分为三类:1) 研究基于“车辆跟驰模型”建立 Aw-Rascle 交通流模型,并对模型的关键参数进行了估计[3,4];2) 研究混合、异质交通条件下的交通流模型[5~7];3) 研究约束条件下的交通流参数模型[8,9]。这些研究发现,已有的关系模型无法准确表示所有的交通状态,例如交通流在自由流状态和拥挤态下的运行特性以及参数关系均存在较大差异。

本文采用 Lempel-Ziv 算法计算不同时间尺度下交通流时间序列的复杂度,利用相关性分析方法,探索了交通流在时间和空间上的相关性。并针对现有的单段函数模型无法准确表征不同交通状态下交通流参数关系的问题,提出了一种二阶段交通流参数关系模型。

1 交通流多时间尺度特性分析

1.1 实验数据描述

实验数据来源于美国加州交通局运行监测系统(performance measurement system,PeMS),选取地点为美国加州洛杉矶市编号为 US101-N 的高速公路,选取检测站编号为 VDS769403,VDS 717484 和 VDS 769388,数据的时间跨度为 2018 年 7 月到 10 月,实时采集的时间间隔为 30 s,经过 PeMS 系统整合得到时间间隔为 5,15,30,60 的交通流数据,每个检测站记录了该高速公路 4 个车道的交通流量、速度和占有率等数据。

1.2 交通流时间序列复杂度计算

本文将 Lempel-Ziv 算法[10]引入到交通流时间序列复杂度计算中,Lempel-Ziv 算法的步骤如下:

Step1 给定一个包含 n 个元素的时间序列($y_1,y_2,\cdots,y_n$),则可以按照以下规则重构一个符号序列($s_1s_2\cdots s_n$):如果 $y_i>y$(其中,y 表示时间序列的平均值),$s_i=1$;否则,$s_i=0$,因此,序列$\{s_1s_2\cdots s_n\}$是一个 0-1 符号序列。

Step2 定义变量 $c(n)$ 为符号序列($s_1s_2\cdots s_n$)的复杂度,S 和 Q 分别为两个不同的符号序列,SQ 表示 S,Q 两个符号序列相加组成的总符号序列,$SQ\pi$ 表示删去 SQ 中最后一个字符所得的符号序列,$\nu(SQ\pi)$ 定义为 $SQ\pi$ 的所有子序列集合。

Step3 初始化 $c(n)=1,\{s_1\}=1,\{s_2\}=1,S=1$ 和 $Q=1$,因此 $SQ\pi=\{s_1\}$。假设 $S=\{s_1s_2\cdots s_r\}$,$Q=\{s_{r+1}\}$,若 $Q\in\nu(SQ\pi)$,则符号序列 Q 是$\{s_1s_2\cdots s_r\}$的一个子序列,因此,复杂度值保持不变,只将 Q 更新为$\{s_{r+1}s_{r+2}\}$,再判断 Q 是否属于 $\nu(SQ\pi)$(符号序列 $SQ\pi$ 也被更新),重复上面的步骤,直到 $Q\notin\nu(SQ\pi)$,然后,将 $c(n)$ 值加 1,读取下一个字符并取 $Q=\{S_{r+3}\}$。

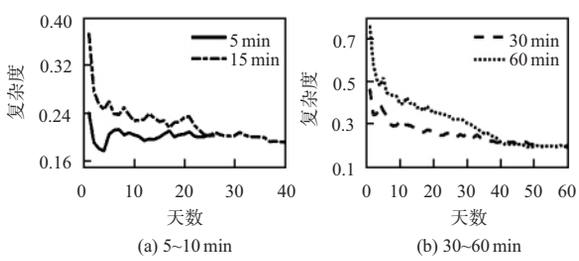
Step4 重复以上步骤,直到符号序列$\{s_1s_2\cdots s_r\}$中的所有元素都被计算到,则得到的 $c(n)$ 是给定时间序列的复杂度。

Step5 根据 Lempel-Ziv 算法,定义归一复杂测度为

$$C_N(n)=c(n)/b(n)\in[0,1] \tag{1}$$

$$b(n)=\lim_{n\to\infty}c(n)\approx\frac{n}{\log_2 n} \tag{2}$$

使用 L-Z 算法计算不同时间尺度下的交通流时间序列复杂度值。如图 1 所示,5 min 数据的样本总容量为 8 640,其 C_N 值在 21 天后大约在 0.210;15 min 数据的样本总容量为 3 840,其 C_N 值在 36 天后大约为 0.195;30 min 数据的样本总容量为 2 496,其 C_N 值在 42 天后大约为 0.200;60 min 数据的样本总容量为 1 440,其 C_N 值在 55 天后大约为 0.190;可见,5 min 时间尺度的交通流序列的复杂度高于其他时间尺度的复杂度,对于交通流时间序列,不同时间尺度的交通流数据具有不同的复杂度值。通过以上复杂度计算的分析表明,小时间尺度的数据表现出较为明显地波动性和随机性,其复杂度较大;随着时间尺度变大,数据逐渐平滑,其削弱了随机性、降低了复杂度。



(a) 5~10 min (b) 30~60 min

图 1 不同时间尺度下的交通流复杂度

1.3 交通流时空特性分析

交通系统是一个相对较封闭的复杂系统,交通流数据是典型的时空大数据,具有不同于其他时间序列的特征。表现在时间上,下一时刻的交通流可以看作是前一时刻数据在特定规律下的延续;表现在空间上,交通流数据受上下游和相邻车道交通状态的影响而呈现出较强的相关性,如某一车道的交通流数据不仅与同一车道上、下游截面的数据有关,还与同一截面不同车道的数据存在相关性。通常用相关性理论中的皮尔逊相关系数(下面简称 R 系数)来描述交通流的时空相关性,如式(3)所示

$$R=\frac{\mathrm{cov}(x,y)}{\sigma_x\sigma_y}=\frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x-\bar{x})^2\sum_{i=1}^{n}(y-\bar{y})^2}} \tag{3}$$

式中 x_i,y_i 为随机变量 x,y 的测量值;$\bar{x},\bar{y}$ 为 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,线性相关性就越高。

交通流时间序列包括横向时间序列和纵向时间序列,纵向时间序列相关性反映了交通流数据的长期趋势性,通

(C)1994-2021 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

常特定区域内具有较稳定的活动模式和规律性,导致不同周内相同时间的交通流具有较强的相关性。

采用编号 VDS 717484 检测站连续 5 个周一 5:00 ~ 17:00的交通数据,时间间隔为 15 min,分析交通流的纵向时间相关性,交通流数据趋势图如图 2 所示。

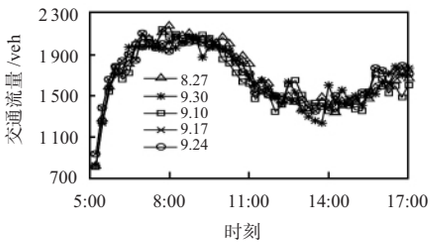


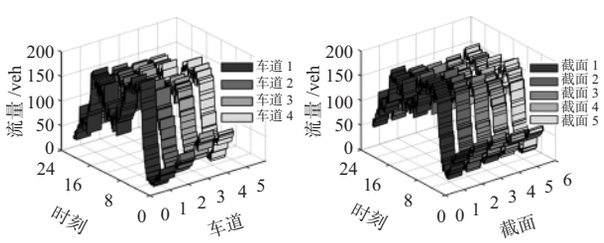
图 2　VDS717484 检测站连续 5 个周一交通数据趋势

由图可见,交通流纵向时间序列数据具有较强的相关性,同一时刻不同日期的交通流数据相差不大且时间序列的变化趋势大致一致。为定量描述交通流纵向时间序列数据的相关性,表 1 给出了连续 5 个周一交通流数据的相关性矩阵,由表可见,时间序列之间的相关系数均在 0.96 以上,说明两者之间具有较强的相关性。

表 1　连续 5 个周一交通流数据的相关性

R	8.27	9.30	9.10	9.17	9.24
8.27	1	0.977	0.980	0.977	0.982
9.30	0.977	1	0.971	0.968	0.971
9.10	0.980	0.971	1	0.972	0.977
9.17	0.977	0.968	0.972	1	0.975
9.24	0.982	0.971	0.977	0.975	1

交通流数据的空间相关性包括横向空间相关性和纵向空间相关性。图 3 分别给出了同一检测截面不同车道的交通流数据趋势图和同一车道不同检测截面的交通流量趋势图。



(a) 同一检测器不同车道交通流量趋势图　　(b) 同一车道不同检测截面交通流量趋势图

图 3　交通流空间趋势

2　交通流参数关系模型

2.1　交通流单段函数模型

典型的交通流参数关系模型主要可分为两类: 单段模型和多段模型,单段模型是指使用单一的函数来表示参数之间的关系; 而多段模型主要是考虑在不同的交通运行状态下,交通流各参数关系之间表现出不同的函数形式。经典的单段模型主要有: Greenshields 模型、Underwood 模型、Pipes 模型和三参数逻辑(three-parameter logistic,3PL) 模型等[11],其函数表达式如表 2 所示。

表 2　单段函数模型关系式

模型	函数
Greenshields	$q = kv_f(1 - k/k_j)$
Underwood	$q = kv_f\exp(-k/k_m)$
Pipes	$q = kv_f[1 - (k/k_j)^n]$
3PL	$q = \dfrac{kv_f}{1 + \exp(\dfrac{k - k_m}{\theta})}$

2.2　二阶段交通流参数模型的建立

单段函数模型结构简单,且与实际的交通流状态存在较大的差异,主要不足有两个方面: 1) 仅用单一函数形式描述交通流参数之间的整体关系,当交通状态发生转变时,通常不能较准确地表示各状态下的交通流参数关系。2) 对于大量交通数据表现出的时间序列趋势特征,单段模型通常显示出无限大的自由流速度和阻塞密度,这与实际的交通情况是不相符的。基于此,本文基于 3PL 模型提出一种二阶段交通流参数模型,模型的函数表达式如式(4) 所示,运用交通流基本参数关系式 $q = k \cdot v$ 可以推导出速度-密度和流量-速度关系式

$$\begin{cases} q = \dfrac{kv_f}{1 + \exp(\dfrac{k - k_m}{\theta})}, 0 \leqslant k \leqslant k_m \\ q = \dfrac{v_m k}{(k_j/k_m - 1)}(k_j/k - 1), k_m \leqslant k \leqslant k_j \end{cases} \tag{4}$$

式中　v_f 为自由流速度,k_j 为阻塞密度,k_m 为临界密度,v_m 为临界速度。

2.3　二阶段交通流参数模型的拟合

为了验证本文提出的二阶段交通流模型的性能,选择 VDS769403、VDS 717484 和 VDS 769388 三个检测站采集到的数据进行模型有效性测试。选取道路各车道流量最大时所对应的密度作为临界密度,则临界密度 k_m 依次为30.18, 27.80,34.28 veh/km,采用非线性最小二乘法分别对 Greenshields 模型(GS 模型) 、Underwood 模型(UW 模型) 、Pipes 模型、3PL 模型和本文模型拟合流量-密度关系,拟合结果如图 4 所示。

2.4　二阶段交通流参数模型性能评价

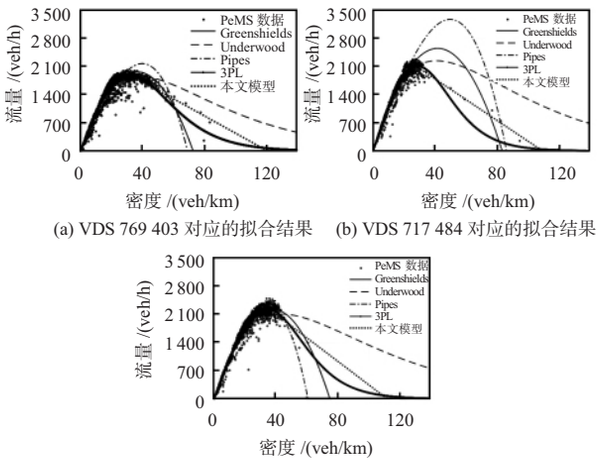
在模型性能评价时,通常采用相关系数(R^2) 和标准误差(RMSE) 对模型的适用性进行评价,二者的计算公式为

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(x_i - y_i)^2}{\sum_{i=1}^{N}(x_i - \hat{x})^2} \tag{5}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{N}(y_i - x_i)^2} \tag{6}$$

式中　x_i 和 $\hat{x}$ 分别为参数的实际值及实际值的平均值,N 为实际值的数目,y_i 为估计值。通常 R^2 越大,模型的适

(C)1994-2021 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net



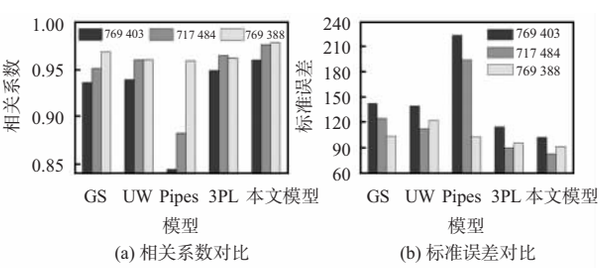
(a) VDS 769 403 对应的拟合结果 (b) VDS 717 484 对应的拟合结果

(c) VDS 769 388 对应的拟合结果

图4 各检测站流量-密度关系拟合结果

用性就越强，当 R^2 大于 0.7 时，认为模型的准确性较高，相反，当 RMSE 越大时，模型的适用性就越差。

分别使用 5 种模型对采集的交通流数据进行测试，对于每一个检测站采集的数据，不同的模型产生不同的拟合结果。通过统计分析，图 5 给出了 5 种参数模型的相关系数和标准误差对比。模型的标准差越小，表明模型参数标定鲁棒性越高，模型的适用性也越好。由图可见，本文的流量-密度关系模型拟合三个检测站交通数据所对应的 R^2 均最大，分别为 0.960，0.976 和 0.978，标准误差相比 3PL 模型有所减小，表明本文模型的拟合性能优于其他 4 种模型。



(a) 相关系数对比 (b) 标准误差对比

图5 5 种模型性能对比

3 结 论

1) 采用动力学中的 Lempel-Ziv 算法计算了不同时间尺度交通流序列的复杂度值，结果表明，小时间尺度的交通流时间序列具有较强的波动性和随机性；在一定的时间序列长度内，其复杂度趋于一个稳定的值。

2) 将相关性理论引入到交通流时空特性分析中，采用皮尔逊相关系数计算方法，定量描述了交通流在时间相关性，揭示了交通流序列具有较强的时空相关特性。

3) 基于 3PL 模型提出了一种二阶段交通流参数关系模型。采用非线性最小二乘法，运用 5 种交通流参数模型对流量-密度关系进行拟合和标定参数，对比发现，本文模型的拟合标准误差最小，表明本文模型可实现对交通流参数关系的准确拟合。

在后续的研究中，需考虑多区域路网的交通流数据进行特性分析，验证模型在不同环境、不同数据集下的适用性和鲁棒性，为交通运行状态分析和短时预测提供理论支撑。

参考文献：

[1] 董春娇，邵春福，诸葛承祥，等. 基于实测数据的快速路交通流参数模型[J]. 交通运输系统工程与信息，2013，13(3)：46－52，59.

[2] 邢雪，于德新，田秀娟，等. 结合可视图的多状态交通流时间序列特性分析[J]. 物理学报，2017，66(23)：57－65.

[3] GOATIN P，LAURENT-BROUTY N. The zero relaxation limit for the Aw-Rascle-Zhang traffic flow model[J]. Zeitschrift für angewandte Mathematik und Physik，2019，70(1)：30－53.

[4] THANKAPPAN A，VANAJAKSHI L. De-velopment and application of a traffic stream model under heterogeneous traffic conditions[J]. Journal of the Institution of Engineers，2015，96(4)：267－275.

[5] 杨达，苏刚，吴丹红. 基于社会力模型的无车道划分异质交通流研究[J]. 交通运输系统工程与信息，2018，18(3)：94－100.

[6] MAHAPATRA G，MAURYA A K. Dynamic parameters of vehicles under heterogeneous traffic stream with non-lane discipline: An experimental study[J]. Journal of Trafic and Transportation Engineering，2018，5(5)：386－405.

[7] 李黎山，李冰，成卫. 基于空间比和感知密度的混合自行车交通流模型[J]. 交通运输系统工程与信息，2019，19(1)：104－110，150.

[8] ZHANG W W，LI R M，SHANG P，et al. Impact analysis of rainfall on traffic flow characteristics in Beijing[J]. International Journal of Intelligent Transportation Systems Research，2019，17(2)：150－160.

[9] 张洪海，许炎，张哲铭，等. 终端区空中交通流参数模型与仿真[J]. 交通运输系统工程与信息，2014，14(6)：58－64.

[10] TANG J J，WANG Y H，WANG H，et al. Dynamic analysis of traffic time series at different temporal scales: A complex networks approach[J]. Physica A Statistical Mechanics Its Applications，2014(405)：303－315.

[11] ZHANG J，QU X B，WANG S A. Reproducible generation of experimental data sample for calibrating traffic flow fundamental diagram[J]. Transportation Research: Part A，2018(111)：41－52.

作者简介：

曹 洁(1966－)，女，教授，博士研究生导师，研究领域为智能交通，智能信息处理，信息融合。

孙兆军(1992－)，男，硕士研究生，研究方向为智能交通，E-mail: 1689515250@qq.com。

张 红(1977－)，女，副教授，硕士生导师，研究领域为大数据，智能交通。

陈作汉(1979－)，男，博士研究生，研究领域为智能信息处理。

侯 亮(1976－)，男，博士研究生，研究领域为智能信息处理，智能交通。

(C)1994-2021 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net