

海湾大学城汽车行驶工况与交通安全分析

刘宇, 黄俊革, 卢思同, 宋雨晴

(上海应用技术大学, 上海 201418)

摘要: 为更好地评估大学城交通安全情况, 以上海海湾大学城行驶车辆为研究对象, 利用 GPS 设备进行行驶数据采集, 通过数据预处理, 提取出有效的运动学片段; 通过主成分分析法和 K-means 聚类分析法将运动学片段划分成 3 类, 利用皮尔逊相关系数拟合出 1 941 s 的汽车行驶工况。结果表明, 大学城行驶车辆具有平均速度高、匀速行驶比例大、加减速比例和怠速比例相对较小等特点, 与主城区汽车行驶工况存在较大差异, 但符合大学城新建区域的交通现状。

关键词: 交通安全; 大学城交通; 行驶工况; 主成分分析; K-means 聚类分析

中图分类号: U491.2

文献标志码: A

文章编号: 1671-2668(2021)01-0024-05

汽车行驶工况是描述车辆行驶的速度-时间曲线, 能体现汽车行驶的运动学特征, 是车辆燃油消耗和污染物排放测试方法、限制标准和汽车性能指标标定优化的基准, 对新车型开发、认证及道路风险水平评估都有重要作用。起初, 欧洲的 NEDC 行驶工况对中国汽车节能减排技术的发展提供了有力支持, 但随着中国科学技术的发展及人们物质生活条件的提高, NEDC 行驶工况与中国实际情况的偏差越来越大。中国地域辽阔, 自然地理环境差异明显, 城市发展程度也不同, 各地区、各城市间的汽车行驶工况存在明显差异。近年来, 研究人员主要基于主成分分析和聚类分析进行汽车行驶工况构建, 有的还结合神经网络算法进行分析, 以提高聚类的精确性和所构建行驶工况的精度。

随着中国大学城建设的蓬勃发展, 大学城所在区域社会经济迅速发展, 区域建设和规划日益加深, 其中交通建设和规划是一项重点。为评估大学城的交通状况, 预防安全事故的发生, 该文以上海海湾大学城为例, 选取周边教职工、学生乘用车辆为研究对象, 使用 GPS 设备采集汽车行驶数据, 基于数据挖掘技术与运动学片段法进行汽车行驶工况构建, 为该区域交通风险控制提供研究依据。

1 数据采集及预处理

汽车行驶工况构建的数据采集方式主要有车辆追踪、平均车流统计和自主行驶法。考虑到该文研究区域范围小且基本固定, 选取大学城教职工、学生使用的小型车为研究对象, 采用自主行驶法进行数据采集。

1.1 数据模型

基础数据属性包含时间、GPS 车速和经纬度, 同时增加间隔时间、加速度、是否停车 (bool 型)、是否怠速可能 (bool 型)、累计怠速时长、是否怠速及怠速起止标志 (bool 型) 等扩展属性。图 1 为数据表 E-R 模型。

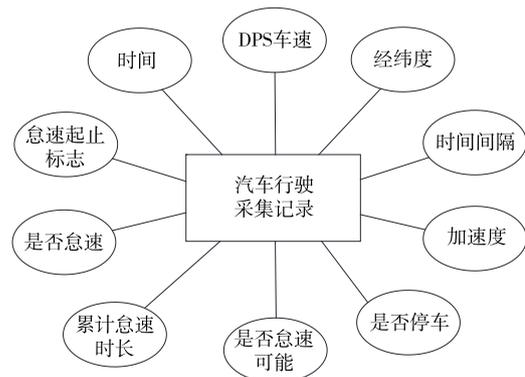


图 1 数据表 E-R 模型

1.2 数据预处理

(1) 记录缺失。GPS 信号丢失造成的数据记录时间不连续, 可理解为数据记录中断, 称作记录缺失。另外, 因预处理删除的部分不良数据, 也归为记录缺失。一般认为怠速时长超过 180 s 为异常情况, 故对缺失时间间隔在 180 s 以内的按步长为 10 s 进行均值处理, 缺失时间间隔在 180 s 以上的则分为两个独立的运动学片段。

(2) 加、减速度异常数据。按照普通小型车的加减速约束, 从零加速至 100 km/h 的加速时间需大于 7 s, 即加速度不能大于 3.97 m/s^2 ; 紧急刹车的最大减速度为 $7.5 \sim 8 \text{ m/s}^2$ 。对于孤立的不满足约

束的数据记录(记录速度的改变不会引起下一个单位时间的加速度指标异常)进行 GPS 车速修正处理,对于连续的不满足约束的多条数据记录进行删除处理,删除后归为记录缺失异常,作进一步处理。

(3) 长期停车。对于停车不熄火,即车速为零、GPS 定位经纬度无变化但发动机转速不为零的长时间过程,及停车熄火但仍有采集记录,即车速为零、GPS 定位经纬度无变化且发动机转速、瞬时油耗等机动车属性指标为零的记录,作删除处理。

(4) 长时间怠速。对长时间堵车、断断续续低于 10 km/h 行驶的情况,按怠速情况处理,标记为怠速可能。由于怠速时间最长可按 180 s 处理,超过则可能存在行驶异常,将大于 180 s 的怠速部分划为下一个运动学片段。若整个运动学片段只有 180 s 的怠速状态,则作删除处理。

(5) 原始数据中 GPS 定位偏离等不良数据。对于时间连续周期内经纬度瞬时偏离的情况,对数据作定位修复处理。若不仅定位偏离,速度也不在上述加减速约束范围内,则按上述不良数据的情况进行预处理。

通过算法编程处理,从 496 464 条数据中剔除不良数据 22 799 条,剩余有效数据 473 665 条。

2 运动学片段提取

运动学片段是指汽车从怠速状态开始至下一个

怠速状态开始之间的车速区间。经过数据预处理后,选取时间、间隔时间、GPS 车速、间隔加速度 4 个属性进行运动学片段提取和划分。虽然已对不良数据进行了预处理,为保证运动学片段提取的有效性,在进行片段划分时进行迭代处理和无效片段数据再处理。设置中间变量为是否怠速(bool 型标志),当 GPS 车速等于零时,视为怠速。通过怠速标志,可得到怠速起始、结束标志,一个怠速起始点到下一个怠速起始点即为一个运动学片段。考虑到长时间停车、断断续续低速行驶的情况,计算各运动学片段的最大速度,最大速度小于 10 km/h 的运动学片段按与相邻记录的时间间隔大小归为前后的运动学片段。共提取有效运动学片段 2 880 个。

3 行驶数据分析

车辆行驶过程中,工况非常复杂,从各种特征综合考虑,大致可分为加速工况、减速工况、怠速工况和匀速工况。其中怠速工况是指汽车停止运动但发动机保持最低转速运转的状态或最高车速低于 10 km/h 的长时间堵车、断断续续低速行驶的情况。

3.1 汽车运动特征

划分运动学片段后,不难发现采集到的数据形成运动学片段具有一定的特征关系,大量运动学片段里一定数量的片段可组成汽车行驶工况。各有效运动学片段的特征值见表 1。

表 1 各运动学片段的特征值

运动学片段	最大速度 $v_{max}/(km \cdot h^{-1})$	平均速度 $v_m/(km \cdot h^{-1})$	平均行驶速度 $v_{mr}/(km \cdot h^{-1})$	速度标准差 $v_{sd}/(km \cdot h^{-1})$	最大加速度 $a_{max}/(m \cdot s^{-2})$	平均加速度 $a_m/(m \cdot s^{-2})$	最小加速度 $a_{min}/(m \cdot s^{-2})$	平均减速度 $a_d/(m \cdot s^{-2})$	加速度标准差 $a_{sd}/(m \cdot s^{-2})$	怠速时间比 $P_i/\%$	加速时间比 $P_a/\%$	减速时间比 $P_d/\%$	匀速时间比 $P_c/\%$
1	17.40	7.02	7.02	5.24	1.25	0.16	-1.86	0.14	0.47	10.45	8.81	6.87	23.88
2	56.50	26.18	26.18	18.47	1.08	0.11	-2.19	0.12	0.40	26.09	30.16	23.37	20.38
3	45.20	20.43	20.43	14.19	1.19	0.18	-2.22	0.18	0.56	19.49	40.68	28.81	11.02
...
2 878	67.20	51.66	51.66	14.76	1.86	0.08	-1.47	0.08	0.30	2.68	29.48	21.86	45.98
2 879	39.00	20.99	20.99	13.25	1.28	0.13	-0.89	0.05	0.27	20.21	41.49	20.21	18.09
2 880	36.00	27.45	27.45	12.58	0.25	0.03	-1.44	0.27	0.47	10.26	17.95	41.03	30.77

3.2 主成分分析

构建汽车行驶工况时,若选用 13 个特征参数展开分析,数据量大且重复性高,需选取具有代表性的参数。主成分分析法通过对大量变量进行线性重组,组合成新的等量变量,或者说将大量成分信息重组为新的等量成分,然后通过每个原本变量(成分)与新变量(成分)之间的相关性程度,在新变量(成

分)里确定主要信息,即确定主成分,从而实现降维和简化。

设共有 n 条数据作为待处理项目,每条数据有 m 个特征参数作为评价指标,构成相应的 $n \times m$ 维矩阵。使用 MATLAB 对已划分好的 2 880 个运动学片段的 13 个特征进行主成分分析,结果见表 2。

由表 2 可知: M_1 、 M_2 、 M_3 的主成分方差大于 1,

表2 各主成分的贡献率和累积贡献率

主成分	主成分方差	贡献率/%	累积贡献率/%
M ₁	5.050 846	38.852 659	38.852 659
M ₂	2.862 293	22.017 639	60.870 297
M ₃	1.627 794	12.521 494	73.391 791
M ₄	0.979 686	7.536 043	80.927 834
M ₅	0.797 424	6.134 032	87.061 866
M ₆	0.613 407	4.718 513	91.780 379
M ₇	0.482 201	3.709 238	95.489 617
M ₈	0.286 361	2.202 778	97.692 395
M ₉	0.221 156	1.701 197	99.393 592
M ₁₀	0.052 194	0.401 491	99.795 083
M ₁₁	0.026 552	0.204 242	99.999 326
M ₁₂	0.000 083	0.000 642	99.999 967
M ₁₃	0.000 004	0.000 033	100.000 000

累计贡献率达 73.39%，故选取 M₁、M₂、M₃ 作为主成分进行进一步分析。表 3 为主成分载荷矩阵。

由表 3 可知：主成分 M₁ 主要反映平均行驶速度、平均速度、最大速度、怠速时间比和速度标准差等特征，主成分 M₂ 主要反映加速度标准差和平均减速度等特征，主成分 M₃ 主要反映最大加速度等特征。荷载系数的绝对值越大，特征参数与主成分的相关系数越高。3 个主成分在各运动学片段中的得分见表 4。

3.3 聚类分析

基于以上主成分分析结果进行聚类分析。采用 K-means 聚类法，该方法是一种动态聚类算法，在选取聚类中心后，不断迭代，逐渐使聚类集中每一样本点到该类中心的距离平方和最小。设置迭代次数

表3 主成分载荷矩阵

主成分	最大速度	平均速度	平均行驶速度	速度标准差	最大加速度	平均加速度	最小加速度	平均减速度	加速度标准差	怠速时间比	加速时间比	减速时间比	匀速时间比
	v_{\max}	v_m	v_{mr}	v_{sd}	度 a_{\max}	度 a_m	度 a_{\min}	度 a_d	A_{sd}	P_i	P_d	P_a	P_c
M ₁	0.875 6	0.921 4	0.923 2	0.738 7	0.251 8	0.576 7	-0.631 0	0.521 8	0.508 4	-0.852 2	0.699 1	0.464 6	0.411 8
M ₂	-0.537 1	-0.479 9	-0.477 8	-0.409 9	0.072 7	0.679 4	-0.017 6	0.695 4	0.775 1	-0.131 0	0.469 6	0.501 7	-0.472 9
M ₃	0.176 8	-0.066 9	-0.068 5	0.379 4	0.698 7	0.205 2	-0.383 4	-0.218 4	0.305 3	0.538 9	-0.043 2	-0.446 9	-0.463 1

表4 主成分的得分

运动学片段	各主成分的得分			运动学片段	各主成分的得分		
	M ₁	M ₂	M ₃		M ₁	M ₂	M ₃
1	-0.859 060	1.493 105	-1.154 500	2 796	1.394 206	-1.939 610	-0.848 720
2	0.859 779	-1.415 740	-0.028 270	2 797	1.869 966	1.904 406	-0.900 240
3	1.152 580	0.993 123	-0.015 360	2 798	2.265 987	-3.469 340	-1.068 830
4	0.803 066	-1.868 600	0.745 082	2 799	-0.523 180	-1.183 740	-0.722 890
5	2.945 549	-1.750 220	1.238 985	2 800	0.442 841	-0.308 380	-2.859 680
...				

为 100，将 2 880 个片段分成 3 类，其中第一类共 685 个片段，第二类共 1 221 个片段，第三类共 974 个片段。每类运动学片段的特征值及总体特征值见表 5。其中：第一类运动学片段的平均速度、平均行

驶速度和最大速度最大，怠速时间比最小，符合高速路况特征；第二类运动学片段的平均速度、平均行驶速度、最大速度和怠速时间比介于其他两类之间，其为中速路况；第三类运动学片段的平均速度、平均行

表5 各类片段的特征值

片段类别	最大速度/(km·h ⁻¹)	平均速度/(km·h ⁻¹)	平均行驶速度/(km·h ⁻¹)	速度标准差/(km·h ⁻¹)	最大加速度/(m·s ⁻²)	平均加速度/(m·s ⁻²)	最小加速度/(m·s ⁻²)	平均减速度/(m·s ⁻²)	加速度标准差/(m·s ⁻²)	怠速时间比/%	加速时间比/%	减速时间比/%	匀速时间比/%
第一类	111.50	41.72	41.71	24.34	10.29	-0.12	-10.29	0.12	0.44	6.35	15.65	11.63	66.35
第二类	65.00	20.14	20.12	13.82	3.86	0.17	-5.94	0.18	0.55	9.82	25.33	19.58	45.26
第三类	55.20	9.20	9.06	11.95	3.89	0.09	-4.06	0.09	0.38	79.51	2.91	2.49	15.08
整体	111.50	27.79	27.75	23.77	10.29	0.13	-10.29	0.13	0.46	59.52	13.81	10.62	16.05

驶速度和最大速度最小,怠速时间比远大于其他两类,其为低速路况。

4 汽车行驶工况构建与分析

为选取更具代表性的运动学片段,按式(1)计算各类运动学片段的总特征参数与该类片段中各片段的特征参数的皮尔逊相关系数,计算结果见表 6。该系数接近于 1,说明该运动学片段在该类片段中更具代表性。

$$\rho_{G,H} = \frac{Cov(G,H)}{\sigma_G \cdot \sigma_H} \quad (1)$$

式中: G 表示每个类别总片段; H 表示每个类别里的单独片段。

表 6 运动学片段与所属类别相关系数

类别	运动学片段	相关系数	类别	运动学片段	相关系数
1	2	0.76	2
	4	0.75		652	0.05
	5	0.84		654	0.42
	6	0.78		661	0.44
		16	0.33
	1205	0.56		17	0.02
	1206	0.72		18	0.08
2	1208	0.76	3	21	-0.01
	1	0.09	
	3	0.61		806	0.00
	10	0.41		807	0.57
	12	0.55		809	0.00

表 7 与其他典型汽车行驶工况的比较

行驶工况	平均速度/ ($\text{km} \cdot \text{h}^{-1}$)	平均行驶速度/ ($\text{km} \cdot \text{h}^{-1}$)	平均加速度/ ($\text{m} \cdot \text{s}^{-2}$)	平均减速度/ ($\text{m} \cdot \text{s}^{-2}$)	怠速时间 比/%	加速时间 比/%	减速时间 比/%	匀速时间 比/%
海湾大学城	26.09	26.05	0.06	0.15	18.02	18.98	14.54	48.44
上海主城区	16.20	24.91	0.41	0.69	37.67	24.29	26.26	11.78
NEDC	33.60	44.05	0.51	0.71	23.81	23.81	17.63	34.75

高,与主城区汽车行驶工况和欧洲 NEDC 工况存在较大差异,但符合地处市郊的大学城交通现状。虽然大学城地处城市郊区,但在特定时段大学城生活区内的人口活动密度与频率堪比城市。根据《道路交通安全法》及其实施条例,没有限速标志、标线的普通城市道路的最高速度为 70 km/h,而大学城道路行驶最大速度为 111.5 km/h,远超过城市郊区道路的限制,拟合的代表性工况中车速大于 70 km/h 的超速行驶占 21.74%。超速行驶会使驾驶人的反

各类别运动学片段在最终拟合工况里的持续时间计算公式如下:

$$t_i = \frac{t_{\text{drivingcycle}}}{t_{\text{overall}}} \sum_{j=1}^{n_i} t_{i,j} \quad (2)$$

式中: t_i 为类别 i 在工况中的持续时间; $t_{\text{drivingcycle}}$ 为工况的总持续时间; t_{overall} 为所有记录数据的总时间; $t_{i,j}$ 为类别 i 中运动学片段 j 的时间; n_i 为类别 i 中的运动学片段数。

按照式(2)计算,得 t_1 、 t_2 、 t_3 分别占 $t_{\text{drivingcycle}}$ 的 34.39%、52.13%和 13.48%。按照该比例从第一类运动学片段中选取序号为 938、1138、561 的片段,从第二类中选取序号为 121、893、972、1192 的片段,从第三类中选取 667 和 1097 片段组成 1 941 s 的代表性工况(见图 2)。结合表 5 中各类运动学片的特征及上述所占比例与其他典型汽车行驶工况进行对比,结果见表 7。

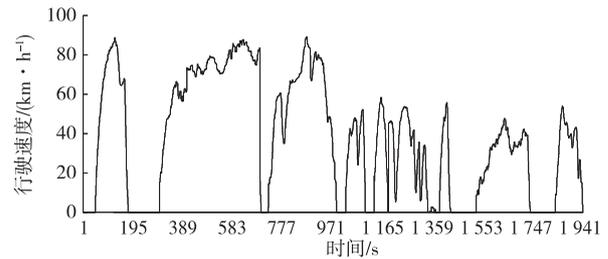


图 2 汽车行驶工况(GPS 车速)

由表 7 可知:大学城汽车行驶的匀速行驶比例较大、加速和减速比例及怠速比例小,说明大学城交通状况较好;平均速度虽不及 NEDC 地区,但也较

应能力降低,增加交通事故发生概率,加大事故的严重程度,暗示着大学城的交通存在一定安全隐患。另外,虽然大学城工况的加、减速比例较低,但两者之和占到整个行驶过程的 1/3,高速行驶与加、减速相结合会大大增加交通安全风险。此外,大学城的公交、大巴也存在速度较快和频繁加、减速的现象。加上大学城行人和车辆活动的时间较集中,随着师生出行活动的频繁,周边的交通开始混乱,而且近年来外卖配送规模与日俱增,骑手“飞速”穿梭于大学

城的大街小巷,超速、闯红灯、闯机动车道,甚至一边开车一边打电话,更是增加了大学城的交通安全隐患。目前已有不少大大小小的交通事故发生,部分学校已对进入校园的外来车辆进行了管控,但校园外围公路的管制措施、设备和设施还有待进一步加强。如在道路上增设限速标志、在路口增设测速设备,必要时在行人活动高峰期设置志愿者及交通管理人员,相关部门还可以和高校联合进行宣传教育活动,从而减少交通事故的发生。

5 结语

该文以上海海湾大学城为对象,采用自主行驶法采集到 496 464 条数据,经过处理得到有效数据 473 665 条,从中提取 2 880 个有效运动学片段,将各片段特征值通过主成分分析、K-means 聚类分析和皮尔逊相关系数法比较得到 1 941 s 的汽车行驶工况。上海海湾大学城私人用车具有平均速度高且匀速行驶比例大、加速和减速比例及怠速比例相对较小等特点,与主城区的汽车行驶工况和欧洲 NEDC 标准有较大差异。结合工况特征与交通安全法分析,大学城的交通存有一定安全隐患,建议除学校以外的交通管理部门加强对大学城周边道路的交通管制,预防交通事故的发生。

参考文献:

- [1] 蔺宏良,余强,张雪莉.市区道路车辆行驶工况研究进展[J].机械科学与技术,2014,33(2):282-288.
- [2] 田宇,朱建军,周博雅.太原市轻型车实际行驶工况构建[J].汽车技术,2018(3):51-55.
- [3] 张翠平,王铁,徐妙侠,等.内燃机排放与控制[M].北京:机械工业出版社,2013.
- [4] 胡志远,秦艳,谭丕强,等.基于大样本的上海市乘用车行驶工况构建[J].同济大学学报(自然科学版),2015,43(10):1523-1527.
- [5] FOTOUHI A, MONTAZERIGH M. Tehran driving cycle development using the K-means clustering method [J]. Scientia Iranica, 2013, 20(2): 286-293.
- [6] 石琴,仇多洋,周洁瑜.基于组合聚类法的行驶工况构建与精度分析[J].汽车工程,2012,34(2):164-169.
- [7] 李宁.城市道路车辆行驶工况的构建与研究[D].保定:河北农业大学,2013.
- [8] 黄友友,程勇,李闯.基于车辆能耗状态的济南市道路行驶工况构建[J].西南交通大学学报,2012,47(6):899-995.
- [9] MATJAZ Knez, TARIQ Muneer, BORUT Jereb, et al. The estimation of a driving cycle for Celje and a comparison to other European cities [J]. Sustainable Cities and Society, 2013, 11: 56-60.
- [10] 杜爱民,步曦,陈礼璠,等.上海市公交车行驶工况的调查和研究[J].同济大学学报(自然科学版),2006,34(7):943-946.
- [11] 仇多洋.汽车行驶工况的构建及波动特性研究[D].合肥:合肥工业大学,2012.
- [12] 郑殿宇,吴晓刚,陈汉,等.哈尔滨城区乘用车行驶工况的构建[J].公路交通科技,2017,34(4):101-107.
- [5] YUAN F, CHEU R L. Incident detection using support vector machines [J]. Transportation Research Part C (Emerging Technologies), 2003, 11(3): 309-328.
- [6] WANG L, NGAN H Y T, YUNG N H C. Automatic incident classification for large-scale traffic data by adaptive boosting SVM [J]. Information Sciences, 2018, 467: 59-73.
- [7] LIU Q, LU J, CHEN S, et al. Multiple naive bayes classifiers ensemble for traffic incident detection [J]. Mathematical Problems in Engineering, 2014, 2014(4): 1-16.
- [8] 张轮,杨文臣,刘拓,等.基于朴素贝叶斯分类的高速公路交通事件检测[J].同济大学学报(自然科学版),2014,42(4):1508-1513.
- [9] SRINIVASAN D, JIN X, CHEU R L. Adaptive neural network models for automatic incident detection on freeways [J]. Neurocomputing, 2005, 64: 473-496.
- [10] LU J, CHEN S, WANG W, et al. A hybrid model of partial least squares and neural network for traffic incident detection [J]. Expert Systems with Applications, 2012, 39(5): 4775-4784.
- [11] 于勇,邓天民,刘龙甘.一种新的快速路交通事件综合检测算法[J].公路与汽运,2009(1):34-37.
- [12] 姜桂艳,温慧敏,杨兆升.高速公路交通事件自动检测系统与算法设计[J].交通运输工程学报,2001,1(1):77-81.

收稿日期:2020-04-01

收稿日期:2020-06-17

(上接第 23 页)

38-42.