

基于 CAN 数据的渣土车驾驶行为分析

李洋¹, 苟刚¹, 姚文钦¹, 黄伶俐²

(1. 贵州大学 计算机科学与技术学院 公共大数据重点实验室, 贵阳 550025;

2. 花溪区人民检察院, 贵阳 550025)

摘要: 渣土运输行业普遍存在着一些管理难的现象, 其主要表现在管理者无法有效地管控车辆的运行、安全等情况; 针对渣土车辆管控的要求, 着力于分析渣土车驾驶员的驾驶行为, 以达到提高渣土车驾驶员的驾驶素质、规范其日常驾驶行为和行车安全意识, 减少安全事故发生的目的; 研究通过对渣土车辆的北斗定位数据和 CAN 数据进行采集与解析后; 首先, 利用 K-means 聚类算法对渣土车驾驶员的驾驶倾向性行为进行分析识别, 帮助管理者在管理车辆的同时, 有效地监督驾驶员的驾驶行为规范; 其次, 构建司机驾驶行为评分模型, 使用熵权层次分析法来确定每个指标的权重, 进而由权重来制定指标分值; 最后, 对渣土车驾驶员进行综合评分, 实现对渣土车驾驶员的驾驶行为分析; 研究对渣土车驾驶行为的分析, 有利于驾驶员更有针对性地改善其驾驶行为, 提高驾驶安全, 减少事故。

关键词: 驾驶行为; 聚类; 北斗定位; 评分模型

Driving Behavior Analysis of Muck Truck Based on CAN Data

LI Yang¹, GOU Gang¹, YAO Wenqin¹, HUANG Lingli²

(1. Key Laboratory of Public Big Data, College of Computer Science and Technology, Guizhou University,

Guiyang 550025, China;

2. People's Procuratorate of Huaxi District, Guiyang 550025, China)

Abstract: There are some difficult management problems in muck transport industry, which are mainly manifested in the inability of managers to effectively control the operation and safety of vehicles. Aimed at the requirements of muck vehicle control, the driving behavior of muck vehicle drivers is mainly analyzed to improve the driving quality of muck vehicle drivers, standardize their daily driving behavior and driving safety awareness, and reduce the occurrence of safety accidents. After collecting and analyzing the Beidou positioning data and CAN data of the muck vehicle; Firstly, K-means clustering algorithm is used to analyze and identify the driving propensity behavior of muck truck drivers, which helps managers to effectively supervise drivers' driving behavior norms while managing vehicles. Secondly, the scoring model of driver's driving behavior is constructed, and the weight of each index is determined by the entropy weight analytic hierarchy process, and then the index score is determined by the weight. Finally, a comprehensive score is given to the drivers of the muck truck to realize the analysis of the driving behavior of muck truck drivers. The analysis of muck truck driving behavior is conducive to the driver's more effective improvement of their driving behavior, which improves driving safety and reduces accidents.

Keywords: driving behavior; clustering; beidou positioning; scoring model

0 引言

随着互联网技术的发展与城镇化进程的加快, 城市的发展离不开城市建设工程以及汽车工业的支持。由于人口的发展, 建筑工程会不断增加, 城市的渣土量会日益增多, 近年来渣土车的数量也在逐年攀升。在货运行业中, 一直以来, 渣土运输行业普遍存在着一些管理难的现象, 使管理者无法有效的管控车辆的运行、安全及驾驶员等情况。

2019年9月6日, 交通运输部、我国税务总局发布了《网络平台道路货物运输经营管理暂行办法》, 对网络货运经营者的法律定位、行为规范和对监督管理机关的监管责

任等提出了明确要求, 也为货运行业规范健康发展提供了良好的制度环境。一辆渣土车^[1]在国道上的超限站限重 50 吨, 合法载重 34 吨左右, 在记重收费的高速公路上限重 45 吨, 合法载重 29 吨左右。在渣土车运输的情况下, 其扬尘、遗漏建筑垃圾、驾驶行为不端等不良现象会给道路交通安全^[2]造成一定的威胁。在交通安全中, 有以下因素主要影响着道路交通安全: 交通参与者、交通量、道路线形、自然环境等各方面的影响。其中交通参与者对交通安全的影响程度最大^[3]。同时为了协助驾驶员增强行车安全意识, 降低安全事故的产生, 国内外在这一方向的研究者们针对驾驶行为开展了大量的研究工作, 研究发现, 记录下行车

收稿日期: 2022-08-12; 修回日期: 2022-10-11。

作者简介: 李洋(1996-), 女, 四川广元人, 硕士研究生, 主要从事深度学习、软件工程方向的研究。

苟刚(1976-), 男, 贵州贵阳人, 副教授, 硕士研究生导师, 主要从事深度学习、软件工程方向的研究。

引用格式: 李洋, 苟刚, 姚文钦. 基于 CAN 数据的渣土车驾驶行为分析[J]. 计算机测量与控制, 2023, 31(4): 187-193.

过程中的不良驾驶行为并使驾驶员能够及时看见自身驾驶行为的反馈情况,可有效降低驾驶员在驾驶过程中危险性驾驶行为的出险概率。

廖向阳^[4]等人通过 OBD 接口来接入车辆的 CAN (CAN, controller area network) 总线,通过收集和分析 CAN 总线报文,得到了车辆正常行驶时的十二个主要运行特征参数及具体数据。从而实现了对车辆运行状况的数据化、可视化展示以及对异常工作情况的及时预警。许静雯^[5]等人通过因子分析提取与驾驶员行为相关的特征参数,来筛选特定的车辆样本,通过聚类来对车辆驾驶行为危险程度进行分级,探究高速公路上重点营运车辆的驾驶行为规律,以便及时确定驾驶行为具有危险性的车辆。高宇^[6]等人通过对车辆上某保险公司外接的北斗系统内的车辆数据进行分析,来分析车主的驾驶行为和驾驶习惯,并分析其对出险概率的影响,以此降低电力行业工程车的出险概率,提高车辆的利用效率。U. Fugiglando^[7]等人利用无监督学习,将不同群体的司机聚集在一起,运用一种验证方法来检测实验中测试聚类的稳健性,同时还提出了一种利用选定的 CAN 总线信号子集对驾驶员行为进行分析和分类的新方法。S. Arumugam^[8]等人通过对管理驾驶方式模型的类别来进行详细调查,提出一种解决方案,通过考虑驾驶员的行为和情绪因素来发现侵略性驾驶和路怒事件所带来的风险,有助于保险行业更准确地评估驾驶风险。

研究通过对急加速、急减速、超速、平均速度、发动机平均燃油率等九种驾驶员驾驶行为数据进行分析、检测其个人驾驶倾向^[9]以及对其进行驾驶行为进行评分^[10],以达到提高渣土车驾驶员的驾驶素质、规范其日常驾驶行为和行车安全意识,减少安全事故发生的目的。

1 数据处理

1.1 数据来源

在数据来源方面主要分为两部分,一部分为 CAN 总线数据,来自汽车生产商在生产渣土车辆时自主安装的 CAN 总线,这一安装是为了便于后期车辆在使用时能向上级平台传输 CAN 总线数据,其相对于使用车辆的 OBD 接口外接其他设备而言没有更多的外加电路,更具安全性。一部分为车辆的北斗卫星定位数据^[11],这一数据由北斗数据运营商提供。这两部分数据的结合能更好的对渣土车驾驶员的驾驶行为进行分析。通过实际考察某公司 70 位渣土车驾驶员在 3 个月中的 CAN 总线数据和北斗定位数据,其中车辆在启动时将以 3 秒每次的速率发送数据,最终共接收 CAN 数据 3 633 874 条,接收北斗定位数据 3 424 516 条。

1.2 数据解析

渣土车辆北斗定位数据电报的内容包含:时间、经度、纬度、当前时速、方向以及海拔等。在后续工作中,根据报文格式与数据位长度,使用字符串截取函数对报文进行截取,然后将其从 16 进制转换为 10 进制,以供研究其它部分使用。其短报文通信定义来自于:《JT/T808-2011 道路运

输车辆卫星定位系统终端通讯协议及数据格式》,其北斗定位数据解析信息如表 1 所示,数据组列表中由数据块信息组成,CAN 数据解析上传的部分数据块含义如表 2 所示。

表 1 北斗定位数据解析信息

起始位置	数据长度	描述
1	4	自定义通信协议传输方式
5	4	消息 ID
9	4	消息体属性
13	12	终端手机号
25	4	消息流水号
29	2	透传消息类型
31	2	主版本
33	2	次版本
35	4	自定义消息 ID
39	4	自定义消息体长度
43	2	数据组个数
45	16	采集起始时间

表 2 CAN 数据部分数据块信息

块 ID	块名称	块类别
1	报警信息块 1	报警
2	报警信息块 2	报警
3	开关信息块 1	开关量
4	开关信息块 2	开关量
5	车速信息项	数字量
6	油耗信息块	数字量
7	动力信息块	数字量

CAN 数据解析首先需判别此段信息的自定义通讯协议的数据传输方式,再获取消息体 ID、消息体、消息体属性、消息流水号等信息,若自定义消息体 ID 为 0706,则此信息为车速、开关、油耗等信息,需获取块 ID 与块个数,根据块 ID 来区分块的类别,也就是消息属于开关量还是数字量。开关量数据需根据块个数来切分块列表的信息长度,以此获取到偏移时间与数据项列表,通过解析数据项列表来得到需要的信息;数字量数据则需根据块 ID 与数据项位图确定数据块长度,然后调用相应方法来解析数据块列表信息。

若自定义消息体 ID 为 0785,则此信息为超速行驶、空挡滑行、急转弯等行为信息。根据自定义消息体中数据 ID 的信息来判别当前的异常行为是什么状态,进而在解析其当前状态的位置与开始结束时间后,存入数据库。

若自定义消息体 ID 为 0200,则此信息为车辆位置数据。同样可获取到消息体 ID、消息体、消息体属性、消息流水号等信息。还可获取到 8 位的 16 进制报警标识与状态标识,以及经纬度、速度、方向、海拔高度等信息。此时需要对报警标志位进行设定相应内容,拆分消息体报警标志,存入 char [] 类型的 warningMarkChar 中,通过循环来获取相应的报警标志位,并对应标志位内容进行解析。

2 驾驶行为倾向性分析

驾驶行为分析模块包含了周行驶里程统计、周行驶速度统计、驾驶员评分、驾驶员倾向性判别。驾驶行为分析主要分析驾驶员在驾驶过程中车辆的 CAN 总线数据与北斗数据, 以此来实现对驾驶员的驾驶行为倾向性判别与驾驶员评分等功能。其中驾驶员的驾驶行为倾向性判别是在利用车辆行驶中驾驶员的定位数据、超速次数、平均速度、发动机平均油耗率、疲劳驾驶次数等数据的基础上, 使用 K-means 聚类算法对渣土车驾驶员的驾驶倾向性进行判别。其中驾驶员评分功能需先使用熵权层次分析法来构建驾驶行为评分模型, 再确定模型中每个指标的权重, 以此分析每个指标的分值, 最后结合实际来对渣土车驾驶员进行综合评分。

2.1 K-means 聚类算法

聚类分析即根据数据中描述对象与其关系的信息, 将数据对象进行分组。目的是, 组内的对象相互之间是相关的, 而不同组中的对象是不相关的。相关性越大, 组间差距越大, 表明聚类效果越好。聚类分析也是数据挖掘的重要手段, 它也称为群分析, 是研究分类问题的一种统计分析方法。

K-means (K-means clustering algorithm) 聚类算法^[12]是在聚类分析中一个常用的算法, 亦为 K-均值聚类, 是一种由 Macqueen 在 1967 年提出的无监督学习算法, 它也是一种简单的迭代性聚类算法, 使用距离作为它的相似性指标。所谓无监督学习是指无需知道所要寻找的目标, 直接通过算法得到数据的共同特征。K-means 算法需指定数据集中的 K 值, 也就是使数据最终分为几个簇 (类), 迭代时需要使用的每个类的中心也为质心, 它是根据类中所有值各维度取平均值得到的。其目的是让每个样本到每个簇中心点的距离越小越好, 也就是通过连续迭代的方式将 m 个样本点划分为 K 个簇, 使得相似的样本被尽可能的分到同一个簇中。该算法具备原理简单、容易理解、速度快等优势, 可处理数值、文本、图像等数据, 但得事先得确定 K 值的选取。其具体步骤如下:

1) 数据集 D 中有 m 个样本点, 首先指定需要划分的簇数 k , 并随机初始化其各自的质心 ($\mu_1, \mu_2, \mu_3, \dots, \mu_k \in R$)。

2) 用欧式距离作为度量指标, 计算第 i 次迭代时每个样本点 x_t ($t=1 \dots m$) 到质心 $u_k^{(i)}$ 的距离 $dist(t, i)$, 找出样本点 x_t 距离质心 $u_k^{(i)}$ 的最小距离, 将样本点分配在距离质心最小的簇中。

$$dist(t, i) = \sqrt{(x_t - u_k^{(i)})^2} \quad (1)$$

3) 次计算每个簇的中心点作为新的质心。

$$u_k^{(i+1)} = \frac{1}{m_i} \sum_{t=1}^{m_i} x_t \quad (2)$$

4) 算数据集 D 中 m 个样本点的误差平方和 E_i , 如果 $|E_{i+1} - E_i| < \delta$ 则表明质心位置变化不大, 算法停止, 否

则返回步骤 2 再次进行迭代。

$$E_i = \sum_{i=1}^m \sum_{t=1}^{m_i} |x_{it} - u_k^{(i+1)}| \quad (3)$$

2.2 基于 K-means 的驾驶倾向性识别

基于 CAN 数据的解析结果, 获取驾驶倾向性分类中所需的特征参数^[13], 具体包含平均车速、超速次数、行驶里程、急加速、急刹车、发动机转速、超速次数、行驶里程、驾驶时长等特征, 每车 3 秒会发送一次数据, 将 70 辆车每天的数据量进行整合, 成为驾驶员当天的数据总和, 具体的特征参数如表 3 所示。

表 3 部分样本数据

急加速次数	急刹车次数	超速次数	平均速度/(km/h)	时长/h	驾驶疲劳次数	里程	发动机平均燃油消耗率	发动机平均转速
0	3	0	22.39	1.6	0	20.6	16.47	1 157.41
0	0	0	31.04	3.8	0	48	17.42	1 164.44
0	0	0	16.83	1.4	0	11	8.89	924.03
5	0	0	27.57	3.1	0	67	25.90	1 375.12
4	2	2	36.29	15.4	1	180	17.51	998.51

进行 K-Means 聚类后, 把驾驶员的驾驶行为分成稳健型、一般型、激进型 3 种, 经过 K-means 聚类后, 对急加速次数、急刹车次数、超速次数等特征数值较大的数据对象, 可以推测这一类型为激进型驾驶员; 同理, 对于急加速次数、急刹车次数、超速次数、平均车速较低的数据对象, 可推测这一类为稳健型驾驶员; 剩下的数据介于二者之中的为一般型。不同驾驶员不同日期的样本有 2514 条, 其中共有分类类别为“0”的一般型驾驶员的标签 1176 条, 共有分类类别为“1”的激进型驾驶员的标签 867 条, 共有分类类别为“2”稳健型驾驶员的标签 471 条。其统计结果如图 1 所示, 可知, 一般型的驾驶员占比达到 47%, 稳健性的驾驶员占比 19%, 激进型的驾驶员占比高达 34%, 本研究对改善驾驶员行为具有重大意义。

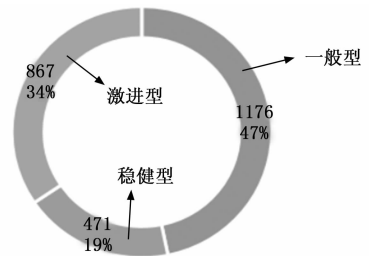


图 1 驾驶行为倾向性聚类分布图

3 评分模型构建

3.1 熵权法

熵是一种热力学的概念, 是用于衡量体系的无序与混乱程度的, 熵越大, 表明事务越混乱, 它广泛应用与工程技术、社会经济等领域^[14]。信息熵这一概念最先由信息奠基人香农 (Shannon) 引入的, 他把信源所含有的信息量称

为信息熵。熵权法是计算权重指标的经典算法之一，是用来判断某个指标的离散（混乱）程度的，它通过指标的变异性大小来判断客观权重，即信息熵越小，离散程度越大，信息量越大，标对综合评价的影响（权重）也就越大，反之亦然^[15]。熵权法的客观性与适应性较强，精度高，能够很好的分析权重结果。因此，可以使用计算熵值来确定某一事件的随机性和无序程度，也可以通过它判断某些指标的离散程度，熵权法的赋权具体步骤如下：

1) 数据归一化，由于各指标的计量单位不同，因此在计算权重时需要对它们进行归一化处理，使数值在 0-1 之间浮动。如给定了 n 个指标 $X_1, X_2, X_3, \dots, X_n$ ，转化后的指标值为： $Y_1, Y_2, Y_3, \dots, Y_n$ ，则：

$$Y_{ij} = \frac{X_{ij} - \min\{x_{1j}, \dots, x_{nj}\}}{\max\{x_{1j}, \dots, x_{nj}\} - \min\{x_{1j}, \dots, x_{nj}\}} \quad (4)$$

2) 计算指标的比重 P_{ij} ，其中 i 为行， j 为列。

$$P_{ij} = \frac{Y_{ij}}{\sum_1^n Y_{ij}} \quad (5)$$

3) 计算指标的熵值 e_j ，若 $P_{ij}=0$ ，则 $\lim_{j \rightarrow 0} P_{ij} \ln P_{ij} = 0$ ，($j = 0, 1, 2, \dots, n$)。

$$e_j = -\frac{1}{\ln(n)} * \sum_{i=1}^n P_{ij} * \ln(P_{ij}) \quad (6)$$

4) 计算信息熵冗余度 d 。

$$d_j = 1 - e_j \quad (7)$$

5) 计算各项指标权重

$$W_j = \frac{d_j}{\sum_1^n d_j} \quad (8)$$

3.2 层次分析法

层次分析法 AHP (analytic hierarchy process) 是美籍运筹学家萨蒂在 20 世纪 70 年代初期根据网络系统理论和多目标综合评价方法而提出的一种层次权重决策分析方法^[16]。它是利用多因素分析方法来判断指标权重的一种主观赋权方式，能够对多目标决策问题进行灵活处理。所确定的指标权重相对合理，所以在各个领域都得以广泛的运用。但层次分析法对定量因素的运用还有不足，其判断矩阵大多依赖决策者经验判断，主观因素往往对权重计算的结果产生了较大的影响。它的基本思想是把所有需要加以分析的问题层次化处理，按照问题的基本特性及其最后要达到的目标，来把问题划分为不同的构成因素。层次分析法的具体步骤如下：

1) 建立层次结构，要把决策的目标、对象和影响因素按照它们之间的关系分为目标层、指标层、方案层。目标层为决策的目的，指标层为需要考虑的因素，方案层为决策的方案，设评分模型有 m 个标准层、 n 个指标层、指标内对应 n_1, n_2, \dots, n_k 个指标。

2) 构建判断矩阵。这里使用了一致矩阵法，即将所有因素两两相互比较，尽可能减少性质不同的因素相互比较的困难，以此增加准确度。判断矩阵 A_{ij} 的标度方法见表 4

所示。

3) 计算各层要素对应权重。这里采用算数平均法计算权重，先求出判断矩阵 A 每列的和，之后再按列归一化（计算该值所占该列比例）得矩阵 B ，最终对每行求算数平均值得到每行的权重 W_i ($i=1, 2, 3, \dots, n$)。

$$W_i = \frac{1}{n} \sum_{i=1}^n B_{ij} \quad (9)$$

表 4 判断矩阵标度

标度	含义
1	两因素相比具有相同重要性
3	两因素相比，一个因素稍微重要于另一因素
5	两因素相比，一个因素明显重要于另一因素
7	两因素相比，一个因素强烈重要于另一因素
9	两因素相比，一个因素极端重要于另一因素
2, 4, 6, 8	上述两相邻判断的中间值
倒数	两个元素的反比较

4) 进行一致性检验。计算判断矩阵 A 的最大特征值 λ_{\max} 与一致性指标 CI ，其中 n 为指标层个数。 CI 越大，不一致越严重，当 CI 越接近 0，表示有越满意的一致性。

$$\lambda_{\max} = \frac{1}{n} \sum_{i=1}^n \frac{(AW)_i}{W_i} \quad (10)$$

$$CI = \frac{\lambda_{\max} - n}{n - 1} \quad (11)$$

5) 定义一致性比率 CR 。为了衡量 CI 的大小，引入随机一致性指标 RI ，其值如表 5 所示。当 $CR < 0.1$ 时，认为构建的判断矩阵通过一致性检验，构造合理，可用其权重 W_i 。

$$CR = \frac{CI}{RI} \quad (12)$$

表 5 随机一致性指标 RI 值

n	1	2	3	4	5	6	7	8	9
RI	0	0	0.58	0.90	1.12	1.24	1.32	1.41	1.45

3.3 评分模型指标体系构建

熵权法为客观赋权法，层次分析法为主观赋权法，主观赋权法通常是由专家依据自身的知识经验和问题实际状况确定权重，主观性较强。客观赋权法是通过与原始数据之间的关联来判断权重，客观性较强。为此结合两种方法使用熵权层次分析法来实现驾驶行为评分模型的构建，使用熵权法来削弱层次分析法主观性的影响，同时平衡了熵权法结果的突出与局部差异^[17]。熵权层次分析法的主要方法为：首先使用熵权法与层次分析法分别得出各自权重，然后把两种方法计算的权重通过组合赋权法^[18]来加以综合，从而得出综合权重 W_j ，其中 W_{js} 为通过熵权法计算的权重，而 W_{jc} 则为通过层次分析法计算的权重。

$$W_j = \frac{W_{js} * W_{jc}}{\sum_{i=1}^n W_{js} * W_{jc}} \quad (13)$$

根据解析的数据, 统计出车辆超速、疲劳驾驶、急刹车、急加速等次数。根据熵权层次分析法将车辆行驶情况、出行情况、驾驶行为表现、发动机情况列为标准层即一级指标; 将平均车速、超速次数、每日里程、每日驾驶总时长、急加速次数、急刹车次数、驾驶疲劳次数、每日发动机平均油耗率、每日发动机平均转速这 9 个因素列为指标层即二级指标, 具体信息如表 6 所示。

表 6 驾驶行为指标分层

目标层	标准层	指标层
驾驶行为评分	行驶情况 θ_1	平均车速 θ_{11}
		超速次数 θ_{12}
	出现情况 θ_2	每日里程 θ_{21}
		每日驾驶总时长 θ_{22}
	驾驶行为表现 θ_3	急加速次数 θ_{31}
		急刹车次数 θ_{32}
		驾驶疲劳次数 θ_{33}
	发动机情况 θ_4	每日平均油耗率 θ_{41}
		每日平均转速 θ_{42}

3.4 评分模型指标权重

结合文献 [19-20] 中的驾驶行为调查结果和专家建议来对各项指标进行评定, 得出了驾驶行为的评分模型判断矩阵, 具体见表 7。通过驾驶行为评分模型判断矩阵中构建的值, 通过使用式 (9) ~ (12) 来进行一致性检验, 经过运算后得出, 标准层的一致性比率 CR 的数值为: 0.003 9, 其值小于 0.1, 即建立的判断矩阵通过了一致性检验, 其构造合理。同理, 指标层也依次进行一致性检验, 结果构造合理。

表 7 驾驶行为评分模型判断矩阵

标准层					指标层					
θ	θ_1	θ_1	θ_2	θ_3	θ_4	θ_1	θ_{11}	θ_{12}		
						θ_{11}	1	$\frac{1}{4}$		
						θ_{12}	4	1		
	θ_2	θ_2	θ_1	θ_2	θ_3	θ_4	θ_2	θ_{21}	θ_{22}	
							θ_{21}	1	2	
							θ_{22}	$\frac{1}{2}$	1	
	θ_3	θ_3	θ_1	θ_2	θ_3	θ_4	θ_3	θ_{31}	θ_{32}	θ_{33}
							θ_{31}	1	$\frac{1}{2}$	$\frac{1}{4}$
							θ_{32}	2	1	$\frac{1}{2}$
	θ_4	θ_4	θ_1	θ_2	θ_3	θ_4	θ_4	θ_{41}	θ_{42}	
							θ_{41}	1	$\frac{1}{2}$	
							θ_{42}	2	1	

标层权重; 再根据判断矩阵的值, 使用式 (9) 来分别计算出层次分析法中标准层与指标层权重。最终通过式 (13) 使用熵权层次分析法来将权重加以综合, 得出熵权层次分析法指标层的权重 W_j , 最终结果如表 8 所示。

表 8 3 种方法下的权重对比

标准层	指标层	熵权法权重 W_j		层次分析法权重 W_j		熵权层次分析法权重 W_j	
行驶情况 θ_1	平均车速 θ_{11}	0.224	0.04	0.045 4	0.23	0.011 3	0.04
	超速次数 θ_{12}	0.161		0.181 7		0.032 5	
出行情况 θ_2	每日里程 θ_{21}	0.004 2	0.13	0.075 7	0.23	0.035	0.21
	每日驾驶总时长 θ_{22}	0.120 8		0.151 5		0.203 3	
驾驶行为表现 θ_3	急加速次数 θ_{31}	0.417 8	0.64	0.060 4	0.42	0.280 3	0.65
	急刹车次数 θ_{32}	0.184 1		0.120 9		0.247 2	
	驾驶疲劳次数 θ_{33}	0.044 3		0.241 8		0.119 0	
发动机情况 θ_4	每日平均油耗率 θ_{41}	0.153 5	0.19	0.040 8	0.12	0.069 5	0.10
	每日平均转速 θ_{42}	0.036 8		0.081 7		0.033 4	

由表 8 中的数据可知熵权法、层次分析法和熵权层次分析法这 3 种权重计算方法对于驾驶行为评分模型的权重计算值, 其中熵权法计算的各项标准层指标权重为: 0.04、0.13、0.64、0.19。层次分析法中的标准层权重分别为 0.23、0.23、0.42、0.12。其中, 熵权法中权重的结果值, 体现了数据在局部特征之间的不同情况, 它削弱了指标之间重要程度对最终结果的影响, 但对于层次分析法来说, 它所计算的权重结果值和判断矩阵的参数设置有很大关系, 决策者对指标重要程度的知识经验对最后权重的结果值影响比较大, 而熵权层次分析法则融合了熵权法和层次分析法, 最终的权重值也融合了两者的权重, 可以更好的体现各个指标对于驾驶行为的影响。

对上表权重的计算结果进行分析后, 将采用熵权层次分析法计算出来的权重作为驾驶行为评分模型的指标分数, 其中驾驶行为评分采用百分制, 按照不同指标的权重值来确定相应的指标分数, 具体指标分数见表 9 所示。

表 9 驾驶行为指标分数

指标	急加速次数	急刹车次数	超速次数	平均速度/(km/h)	时长/h	驾驶疲劳次数	里程/km	发动机平均燃油消耗率/(L/h)	发动机平均转速
分数	28	25	3	1	20	12	1	7	3

通过使用式 (4) ~ (8) 来计算熵权法中标准层与指

3.5 驾驶行为综合评分

在确定了各个指标的分值后，还需要根据实际情况制定详细的评分标准。研究结合了文献 [21] 与车辆保险行业车联网解决方案白皮书以及根据现有数据的实际分布情况，采用了如表 9 所示的驾驶行为评分指标分数，制定了如表 10 所示的具体评分细则。其中 a 为指标所占分值， S 为该指标实际的最终分值， x 为指标实际值。

表 10 驾驶行为评分细则

指标	指标分值	评分细则
平均车速/ (km/h)	1	$S = \begin{cases} a, & 0 < x < 35 \\ 0.9a, & 35 < x \leq 50 \\ 0.75a, & 50 < x \leq 60 \end{cases}$
超速次数	3	$S = \begin{cases} a - \frac{x}{10}, & 0 \leq x \leq 20 \\ 1, & x > 20 \end{cases}$
每日里程/ km	1	$S = \begin{cases} a - \frac{x}{400}, & 0 \leq x \leq 400 \\ 0, & x > 400 \end{cases}$
驾驶时长	20	$S = \begin{cases} a - \frac{x}{2}, & 0 \leq x \leq 20 \\ 10, & x > 20 \end{cases}$
急加速次数	28	$S = \begin{cases} a - \frac{x}{2}, & 0 \leq x \leq 30 \\ 13, & x > 30 \end{cases}$
急刹车次数	25	$S = \begin{cases} a - \frac{x}{2}, & 0 \leq x \leq 30 \\ 10, & x > 30 \end{cases}$
驾驶疲劳次数	12	$S = \begin{cases} a - 3x, & 0 \leq x \leq 3 \\ 3, & x > 3 \end{cases}$
每日平均油耗率/(L/h)	7	$S = \begin{cases} a - \frac{x}{9}, & 0 \leq x \leq 54 \\ 1, & x > 54 \end{cases}$
每日平均转速	3	$S = \begin{cases} a - \frac{x}{600}, & 0 \leq x \leq 1200 \\ 1, & x > 1200 \end{cases}$

将通过驾驶行为倾向性识别后的 70 辆车在不同日期下产生的 2 514 条数据样本，按照倾向性识别类型：一般型、稳健型、激进型分类，再进行驾驶行为评分，得到如图 2~4 所示的不同倾向驾驶员的分值分布。

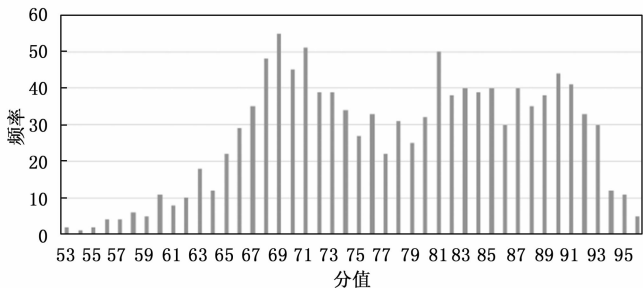


图 2 一般型驾驶员分值分布情况

由图 2、图 3、图 4 所示，激进型驾驶员的分值分布在 60~75 分之间的次数比较多；稳健型驾驶员的分值分布在

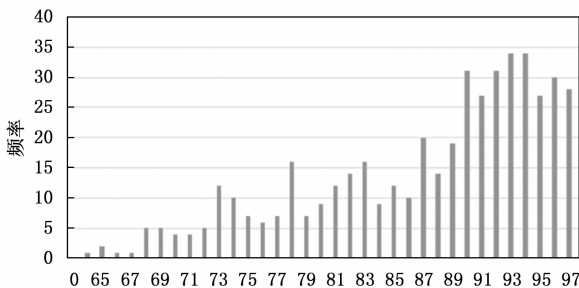


图 3 稳健型驾驶员分值分布情况

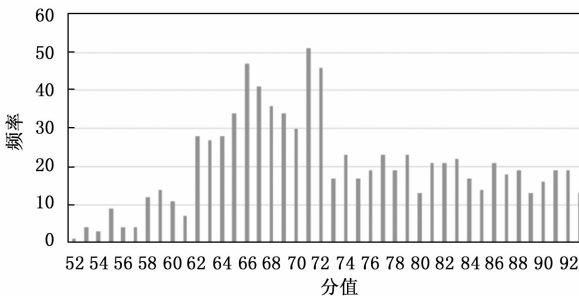


图 4 激进型驾驶员分值分布情况

85~100 分之间的次数比较多，而一般型驾驶员的分值主要分布在 65~95 分之间，由此，研究所构建的驾驶评分模型具有合理性，可以为管理者和驾驶员提供驾驶行为信息，为后续驾驶员的行车操作做出警示，对于分值较高的驾驶员需要继续保持，对于分值较低的驾驶员需要加以改善。

4 结束语

研究分析了基于 CAN 数据的渣土车驾驶行为。首先，介绍了 CAN 数据的解析方案，进而获取到了数据，对 CAN 数据与北斗定位数据进行处理分析；其次，采用聚类的 K-Means 算法对渣土车司机的驾驶行为倾向性进行分析；随后介绍了熵权法与层次分析法，在计算出二者在指标层上的权重后利用熵权层次分析法集成权重；最后通过文献结合实际情况的方式制定评分细则，计算得出渣土车驾驶员的评分，以便帮助驾驶员更有针对性的改善其驾驶行为。通过实验表明，改进方法能够满足用户的需求，有利于管理员管理的同时提升驾驶员安全行车水平，能有效提升驾驶员的安全。

参考文献：

[1] 母洪铤, 苟刚, 姚文钦, 等. 基于 NSGA-II 的渣土车调度系统的研究与设计 [J]. 计算机测量与控制, 2022, 30 (8): 183-188, 196.

[2] 许治琦, 姚远, 周兴社, 等. 基于模糊综合评价的驾驶安全性评价研究 [J]. 计算机测量与控制, 2019, 27 (7): 261-265.

[3] 马阳. 城市道路交通安全影响因素综合分析研究 [J]. 中华建设, 2021 (10): 144-145.

- [4] 廖向阳, 余剑武, 张诗语, 等. 基于CAN总线的汽车实时数据采集与可视化研究 [J]. 电子元器件与信息技术, 2021, 5 (5): 48-51.
- [5] 许静雯, 朱丹, 董军, 等. 基于北斗数据的重点营运车辆驾驶行为分析 [J]. 交通节能与环保, 2020, 16 (6): 36-38.
- [6] 高宇, 李兰心. 基于北斗导航的用户驾驶行为与用户出行风险相关性分析 [C] // 卫星导航定位与北斗系统应用 2019——北斗服务全球融合创新应用, 2019: 187-193.
- [7] FUGIGLANDO U, et al. Driving behavior analysis through CAN bus data in an uncontrolled environment [J]. IEEE Transactions on Intelligent Transportation Systems, 2019, 20 (2): 737-748.
- [8] ARUMUGAM S, BHARGAVI R. A survey on driving behavior analysis in usage based insurance using big data [J]. Journal of Big Data, 2019, 6 (1): 1-21.
- [9] 邓天民, 朱杰, 朱凯家, 等. 基于 iForest + Biscting K-means 的驾驶风格辨识方法研究 [J]. 重庆交通大学学报 (自然科学版), 2021, 40 (6): 1-6.
- [10] 应海宁, 唐振民, 韩旭. 基于两次聚类的 PWARX 驾驶行为辨识模型 [J]. 重庆交通大学学报 (自然科学版), 2019, 38 (2): 79-85.
- [11] BU J, LI X, ZUO X, et al. Quality comparison and analysis of Beidou/GPS/GLONASS multi system satellite positioning data [J]. Progress in Geophysics, 2018, 33 (1): 1-9.
- (上接第 160 页)
- 信息装备通用测试平台的总体思路、软硬件结构框架、技术途径。硬件平台利用软件无线电技术构建, 采用合成仪器, 收发一体综合测试技术, 硬件利用率高, 集成度高。软件平台采用基于信号的测试资源智能管理技术, 采用先进故障诊断技术, 支持国际标准的数字测试交换格式。实现了提升舰艇信息装备维修保障能力的目标。为其他测试平台的设计提供了新的思路与途径, 具有很大的借鉴意义。
- 参考文献:**
- [1] 吴千, 刘东风. 装备综合诊断技术的研究与应用 [J]. 中国修船, 2008, 21 (6): 50-52.
- [2] 张宝珍. 国外综合诊断, 预测与健康管理体系的发展及应用 [J]. 计算机测量与控制, 2008, 16 (5): 591-594.
- [3] 王发坤, 罗云锋. 舰艇装备综合诊断技术和体系框架 [J]. 舰船电子工程, 2007, 27 (2): 13-15.
- [4] 任安民, 王卫国. 武器装备综合诊断技术的现状与发展 [J]. 舰船电子工程, 2007, 27 (3): 20-22.
- [5] 田雨, 陈光禹. 电子系统通用测试平台的研究与实现 [J]. 测控技术, 2004, 23 (10): 21-24.
- [6] 楼才义, 徐建良, 杨小牛. 软件无线电原理与应用 [M]. 北京: 电子工业出版社, 2014.
- [7] 陶玉柱, 胡建旺, 崔佩璋. 软件无线电技术综述 [J]. 通信技术, 2011 (1): 37-39.
- [8] 陈星, 黄考利, 连光耀, 等. 从 1149.1 标准到 1149.7 标准 [12] MAJHI S K, BISWAL S. Optimal cluster analysis using hybrid K-means and ant lion optimizer [J]. Karbala International Journal of Modern Science, 2018, 4 (4): 347-360.
- [13] 杨光. 基于车辆行驶数据的驾驶员行为模型构建与应用 [D]. 西安: 西安石油大学, 2021.
- [14] TSALLIS C. Entropy [J]. Encyclopedia, 2022, 2 (1): 264-300.
- [15] GAO C, Li S, WANG J, et al. The risk assessment of tunnels based on grey correlation and entropy weight method [J]. Geotechnical and Geological Engineering, 2018, 36 (3): 1621-1631.
- [16] HO W, MA X. The state-of-the-art integrations and applications of the analytic hierarchy process [J]. European Journal of Operational Research, 2018, 267 (2): 399-414.
- [17] 张正. 基于车联网驾驶行为评分的安全驾驶卫士分析与设计 [D]. 南京: 南京邮电大学, 2017.
- [18] ODU G O. Weighting methods for multi-criteria decision making technique [J]. Journal of Applied Sciences and Environmental Management, 2019, 23 (8): 1449-1457.
- [19] 彭江琴. 基于GID的车联网保险UBI费率与驾驶行为评分研究 [D]. 南京: 南京邮电大学, 2016.
- [20] 陈东攀. 基于北斗车辆定位数据的驾驶行为分析 [D]. 郑州: 华北水利水电大学, 2020.
- [21] 陈辉. 面向车联网大数据的驾驶行为分析 [D]. 贵阳: 贵州大学, 2020.
- 分析边界扫描技术的发展 [J]. 计算机测量与控制, 2009, 17 (8): 1460-1462, 1472.
- [9] 杨迪珂, 颜学龙. 边界扫描测试算法的分析与优化 [J]. 计算机测量与控制, 2017, 25 (5): 4-6, 13.
- [10] 李国卿, 黄金杰. 基于故障字典法的雷达故障诊断技术研究 [J]. 科学技术与工程, 2007, 7 (11): 2653-2655.
- [11] 陈如清. 两种基于神经网络的故障诊断方法 [J]. 中国电机工程学报, 2005, 25 (16): 112-115.
- [12] 倪绍徐, 张裕芳, 易宏, 等. 基于故障树的智能故障诊断方法 [J]. 上海交通大学学报: 自然科学版, 42 (8): 1372-1375.
- [13] IEEE STD 1232-2010 [S]. IEEE standard for artificial intelligence exchange and service tie to all test environments. 2010.
- [14] IEEE STD 1232-2002 [S]. IEEE standard for artificial intelligence exchange and service tie to all test environments. 2002.
- [15] 舰船武器装备系统级测试程序集数据规范 [S]. 海军装备技术研究所, 2009.
- [16] 林志文, 贺喆, 刘松风. 基于 ATML 的系统级 TPS 开发及综合诊断应用 [J]. 仪器仪表学报, 2010, 31 (5): 1010-1016.
- [17] 张桂英, 侯倩, 范利花, 等. 基于 ATML 的可移植 TPS 开发技术研究 [J]. 测控技术, 2017, 36 (1): 112-115.
- [18] 刘明, 高海英, 张伟昆. 基于 ATML 的 TPS 可移植性设计 [J]. 电子科技, 2019, 32 (5): 68-74.
- [19] 陈亮, 魏蛟龙, 史慧. 基于 IEEE1445 标准的电路故障诊断技术研究 [J]. 信息与标准化, 2003 (12): 30-33.