



计算机科学

COMPUTER SCIENCE

多源异构环境下的车联网大数据混合属性特征检测方法

陈晶, 吴玲玲

引用本文

陈晶, 吴玲玲. 多源异构环境下的车联网大数据混合属性特征检测方法[J]. 计算机科学, 2022, 49(8): 108-112.

CHEN Jing, WU Ling-ling. Mixed Attribute Feature Detection Method of Internet of Vehicles Big Data in Multi-source Heterogeneous Environment[J]. Computer Science, 2022, 49(8): 108-112.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[基于维度属性距离的混合属性近邻传播聚类算法](#)

Mixed Data Affinity Propagation Clustering Algorithm Based on Dimensional Attribute Distance

计算机科学, 2015, 42(Z11): 55-57.

[混合属性数据流的双重 k 近邻聚类算法](#)

Double k-nearest Neighbors of Heterogeneous Data Stream Clustering Algorithm

计算机科学, 2013, 40(10): 226-230.

[基于衰减模型的混合属性数据流离群检测](#)

Outlier Detection Based on the Damped Model in Mixed Data Streams

计算机科学, 2010, 37(5): 157-162.

多源异构环境下的车联网大数据混合属性特征检测方法

陈晶¹ 吴玲玲²

1 中冶赛迪工程技术股份有限公司 重庆 401122

2 重庆交通大学 重庆 400074

(patton_g_2000@163.com)

摘要 现有的车联网大数据特征检测方法忽略了数据属性权重,导致效率偏低,无法在车辆运行中提供高效服务。为此,提出了多源异构环境下的车联网大数据混合属性特征检测方法。该方法利用集成模型集成车联网多源异构数据,并对集成数据进行标准化和属性约简处理;同时,通过加权主成分分析法提取集成数据的属性特征,并利用聚类方法实现特征聚类,完成车联网大数据混合属性特征检测。实验结果表明,与现有方法相比,所提方法在评价指标敏感性指数上取值更高,时间复杂度更低,能更高效地完成车联网大数据混合属性特征提取任务。

关键词:多源异构环境;车联网大数据;混合属性;特征检测方法

中图分类号 TP368.6

Mixed Attribute Feature Detection Method of Internet of Vehicles Big Data in Multi-source Heterogeneous Environment

CHEN Jing¹ and WU Ling-ling²

1 CISDI Engineering Co., Ltd, Chongqing 401122, China

2 Chongqing Jiaotong University, Chongqing 400074, China

Abstract Current feature detection methods for big data of Internet of vehicles ignore the data attribute weight, resulting in low efficiency and fail to provide efficient services in vehicle operation. Therefore, a hybrid attribute feature detection method for big data of Internet of vehicles in multi-source heterogeneous environment is proposed. Middleware is used to build an integration model to integrate multi-source heterogeneous data of the Internet of vehicles, and standardization and attribute reduction of integrated data are completed. With pre-processed data as input, attribute features are extracted by weighted principal component analysis, and feature clustering is realized by clustering method to complete the feature detection of mixed attribute of Internet of vehicles big data. Experimental results show that compared with existing methods, the sensitivity index of the proposed method is higher and the time complexity is lower, which indicates that the proposed feature detection method is more efficient and can more accurately complete the feature extraction task of mixed attributes of the Internet of vehicles big data.

Keywords Multi-source heterogeneous environment, Internet of vehicles big data, Mixed attributes, Feature detection method

1 引言

近年来,为了满足人们的出行需求,车辆保有量逐年增加,汽车行业越发兴盛。随着人们对车辆要求的不断提高,车辆运行逐渐向智能化、自动化方面发展。车联网将汽车与互联网连接,实现人、车、路、云之间数据互通,为驾驶人员驾驶车辆提供帮助,实现智能交通出行。随着车联网越来越智能化,车联网大数据规模越来越大、类型越来越多,如路况数据、天气状况、行驶轨迹等^[1]。这些数据在车辆驾驶中都起到了非常重要的作用,关系到车辆驾驶的安全性和稳定性。然而,随着多源异构数据的不断涌入,现有车联网缺乏处理大量数据的能力,使得在后续数据利用方面面临着巨大挑战。为了

解决数据处理困难的问题,并提高车联网数据的利用效率,对多源异构环境下的车联网大数据进行处理变得十分重要。多源异构环境下的车联网大数据的显著特点是多种属性混合,而多种属性数据混合在一起会造成数据混乱,数据特征不明显,这极易造成车联网给驾驶人员提供错误的指导,从而间接影响驾驶安全。针对上述问题,如何从属性混合数据中检测出特征,成为了车联网数据利用过程中的一个重要环节。因此,本文提出研究多源异构环境下的特征检测方法。

2 相关工作

关于车联网数据处理问题,国内外相关专家和学者都进行了研究。为了实现车联网的仿真研究,Yao等^[2]提出了

到稿日期:2022-03-09 返修日期:2022-05-25

基金项目:国家重点研发计划(2018YFB1601001)

This work was supported by the National Key Research and Development Program of China(2018YFB1601001).

通信作者:吴玲玲(330931228@qq.com)

集成 Vissim 和 Python 的车联网仿真平台构建方法。在采集车联网大数据的基础上,利用 Vissim 和 Python 构建了车联网仿真平台,Vissim 具有动态交通仿真技术,Python 具备具有可视化功能的对象库、增加数据深度挖掘优势、易于实现算法等优点。在此仿真平台中,Python 作为主要控制程序,通过与 Vissim COM 接口通信来实现 Vissim 中的各种交通信息的获取;基于此,基于这些信息实现各种交通控制算法和优化模型;最后将优化结果通过 Vissim COM 接口反馈到 Vissim 中,进而实现对 Vissim 中对象的控制和优化。该研究成果为本文的车联网大数据的仿真实验奠定了基础。Sarmathiy 等^[3]结合 Hadoop,利用机器学习算法对大量非结构化和复杂数据进行特征提取,提取数据相关性的特征以及互信息的特征,这种方法具有较高的灵敏度和特异性,但是效率不高。Tan^[4]将灰色神经网络与残差学习方法相融合,通过神经网络的不断迭代,从混乱的大数据中检测特征数据,这种方法降低了特征提取错误率,但是时间复杂度高。Wu 等^[5]提出了一种基于堆叠去噪自动编码器和批量归一化的方法,提取深度新特征,以选取具有代表性的训练样本,该方法的优点是效率高,但是易受到编码器参数的影响,提取精度存在一定的限制。针对大数据混合特征的检测问题,Zhang 等^[6]提出了结合粗糙集理论的网络大数据混合属性特征检测方法,通过四元组模型的构建,来获取四元组模型的任意邻域信息以测算其长度函数;利用该函数可判断大数据的信息特征相似性;引入粗糙集理论对相似信息特征的邻域熵进行求解,以检测并分类重复数据属性;引入支持向量机分类思想,将大数据混合属性的分类问题变换为线性可分问题,实现网络大数据混合属性特征的检测与分类。

结合已有的研究,针对已有方法存在的问题,本文提出了多源异构环境下的车联网大数据混合属性特征检测方法。该方法的提出以期规范多源异构数据混合属性特征,提高车联网大数据利用率。

3 车联网多源异构大数据混合属性特征检测方法

车联网数据是指导驾驶员完成行驶任务并安全驾驶的重要基础,智能系统给出的很多行驶建议都是依据车联网数据给出的。车联网的数据主要来源于汽车感知数据、驾驶行为数据以及外部周边环境数据。这些数据来自不同的业务系统,因此存在多源异构性,而多源异构性的存在必然使得数据存在多个属性,导致数据混乱,不利于数据利用,因此对车联网多源异构大数据进行混合属性特征检测具有重要的现实意义。关于多源异构环境下的车联网大数据混合属性特征检测方法的研究主要包括 3 个部分,即多源异构车联网数据集成、多源异构数据预处理以及特征检测实现。

3.1 车联网多源异构数据集成

近年来,智慧交通、智能驾驶逐渐成为主流,在此背景下,车联网的应用必不可少。车联网只有与各个业务系统相关联,才能获取所需要的数据,由于这些数据来自不同的业务系统,因此形成了车联网多源异构数据^[7]。为此,首要环节就是车联网多源异构数据集成。

本文通过中间件构建异构数据集成模型,如图 1 所示。

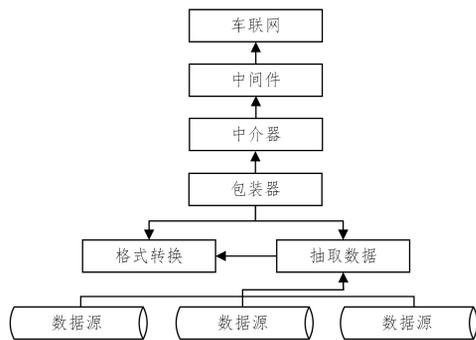


图 1 基于中间件的车联网多源异构数据集成模型

Fig. 1 Multi-source heterogeneous data integration model for Internet of vehicles based on middleware

中间件是一种沟通应用软件和系统软件之间连接的软件,主要从业务系统中,抽取数据,然后将数据输送到车辆用户端,用于后续分析^[8]。

基于中间件的车联网多源异构数据集成模型中,包装器和中介器是关键。其中,包装器是一个数据访问枢纽,主要作用是响应上层的数据集成请求,然后提取源数据并将其转化,转换为车联网系统可以接受的形式^[9];中介器的主要作用是传递上层请求,然后在下层包装器完成数据抽取和转换后,将结果反馈给上层。本文通过集成模型收集车联网多源异构数据,用于后续分析。

3.2 车联网多源异构数据预处理

基于集成数据,采用常见的数据清洗和标准化技术将其转换为适应于特征提取模型的格式,具体步骤如下^[10]。

(1) 集成数据标准化

数据标准化处理的目标是解决多源的车联网异构数据量纲不一致的问题,使所有属性具有量纲统一性^[11],具体过程如下。

步骤 1 计算数据算术平均值,计算式如下:

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} \quad (1)$$

其中, \bar{x} 表示算术平均值; x_1, x_2, \dots, x_n 表示多源的车联网异构数据; n 表示数据数量。

步骤 2 计算数据标准差,计算式如下:

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}} \quad (2)$$

其中, s 表示数据标准差; x_i 表示第 i 个原始车联网多源异构数据。

步骤 3 按照式(3)进行标准化处理。

$$x_i' = \frac{x_i - \bar{x}}{s} \quad (3)$$

其中, x_i' 表示标准化后的数值。

步骤 4 将逆指标前的正负号对调。

(2) 属性约简

针对混合属性存在的缺陷,即属性扰乱性,本节进行属性约简,以减少混合属性导致的数据混乱性^[12]。属性约简过程如下。

步骤 1 给出决策表 $S=(U, A=C \cup D, V, f)$,如表 1 所列。

表1 决策表
Table 1 Decision table

论域	条件 属性 1	条件 属性 2	条件 属性 3	条件 属性 4	决策 属性
1	1	1	0	0	2
2	1	1	1	1	1
3	1	0	1	1	2
4	1	1	1	0	1
5	0	1	0	1	0
6	0	1	1	1	0
7	1	0	0	1	1
8	1	0	0	0	2
9	0	1	1	0	3
10	0	1	1	0	3

$S=(U, A=C \cup D, V, f)$ 中, U 表示论域; A 表示属性集; C 表示条件属性; D 表示决策属性; V 表示属性值的集; f 表示信息函数。

步骤 2 计算条件属性集 C 的等价类簇, 记为 $X=\{X_1, X_2, \dots, X_r\}$ 。

步骤 3 求出决策属性集 D 的等价类簇, 记为 $Y=\{Y_1, Y_2, \dots, Y_k\}$ 。

步骤 4 对给出的决策表进行简化。

步骤 5 构建相关矩阵 M 。

步骤 6 判断矩阵 M 的某一行是否只有一个值, 且该值 m_i 是否为 1, 若是, 则认为 m_i 的属性为核属性, 并将其添加到约简属性集中。

步骤 7 将所有对应值 $m_i=1$ 的行中的数据置为 0。

步骤 8 将条件属性集 C 的列中值为 1 的个数最多的属性 c_k 加入约简。

步骤 9 判断 c_k 是否唯一, 当不是唯一时, 将所有对应 c_k 值为 1 的行的全部数据置换为 0。

步骤 10 判断 M 中所有值是否都为 0, 当都为 0 时, 输出最小约简值。

通过上述约简后, 数据的属性特征更加清晰、明确, 方便后续的特征检测^[13]。

3.3 车联网大数据混合属性特征检测实现

基于经过预处理的车联网多源异构数据进行混合属性特征检测, 检测过程包括特征提取与特征聚类两部分。通过前者提取的特征, 将同类型的特征聚合以实现特征检测^[14]。

(1) 特征提取

特征提取, 即从所有特征中抽取出具有代表性的特征。本文通过加权主成分分析法来进行特征提取。加权主成分分析法是在原始主成分分析法的基础上优化而来的, 通过为不同属性类型数据赋予不同权值来提高特征的代表性^[15], 具体过程如下。

步骤 1 假设车联网大数据样本为 N 个属性混合在一起的样本, 记为 Z 。

步骤 2 将 Z 进行转置, 组成矩阵形式, 描述如下:

$$Z = \begin{bmatrix} z_{11} & z_{12} & \dots & z_{1n} \\ z_{21} & z_{22} & \dots & z_{2n} \\ \dots & \dots & \dots & \dots \\ z_{m1} & z_{m2} & \dots & z_{mn} \end{bmatrix} \quad (4)$$

步骤 3 对矩阵进行归一化处理, 处理公式如下:

$$z'_{ij} = \frac{z_{ij} - \min(z_{ij})}{\max(z_{ij}) - \min(z_{ij})} \quad (5)$$

其中, z'_{ij} 表示归一化处理后的车联网数据; z_{ij} 为原始车联网大数据; $\min(z_{ij})$ 和 $\max(z_{ij})$ 表示原始车联网大数据中的最小值和最大值^[16]。

步骤 4: 计算归一化处理后的矩阵 M' 中每列的均值, 进行均值化处理^[17]。

$$s_{ij} = \frac{z'_{ij}}{u} \quad (i=1, 2, \dots, m; j=1, 2, \dots, n) \quad (6)$$

其中, u 表示均值。

步骤 5: 计算车联网大数据样本的权值, 计算式如下:

$$\omega_i = \frac{\left(\sum_{j=1}^m \frac{z'_{ij}}{\sum_{j=1}^m z'_{ij}} \right)}{n}, i=1, 2, \dots, n \quad (7)$$

其中, ω_i 表示权值; z'_{ij} 表示矩阵 Z 的元素 z_{ij} 经过归一化后的值; n 表示数据数量^[18]。

步骤 6 将步骤 4 的结果与步骤 5 的结果相乘, 构建新的数据样本, 组成新的矩阵 H 。

步骤 7 对新的数据样本进行主成分分析。

步骤 8 计算矩阵 H 的协方差, 并组成协方差矩阵 V 。

步骤 9 计算特征值和对应的特征向量。

步骤 10 计算累积贡献率, 通过累积贡献率选择前 k 个数, 并将其作为主成分^[19]。

步骤 11 将选出的前 k 个主成分作为数据特征。

(2) 特征聚类

基于上文提取出的车联网大数据混合属性特征, 本节对特征进行聚类, 将具有一致性的特征聚集在一起, 具体过程如下。

步骤 1 从提取的特征集中随机选择一个特征作为聚类中心。

步骤 2 依据聚类中心进行模糊聚类。

步骤 3 进行特征一致性度量, 特征一致性度量方式如下:

$$Sim(x_1, x_2) = \sum_{i=1}^c \sum_{j=1}^n \|x_1 - x_2\| \quad (8)$$

其中, $Sim(x_1, x_2)$ 表示特征一致性度量结果; x_1 和 x_2 表示提取的车联网大数据混合属性特征; c 表示聚类中心; n 表示特征数量。

步骤 4 判断特征是否冗余, 即判断特征一致性度量 $Sim(x_1, x_2)$ 值是否小于 0.5, 若小于, 则认为存在冗余, 并执行以下步骤。

步骤 4.1 对特征从大到小进行排序;

步骤 4.2 确定存在冗余特征;

步骤 4.3 删除冗余特征。

若特征一致性度量 $Sim(x_1, x_2)$ 值大于 0.5, 则认为不存在冗余, 继续进行下一步骤。

步骤 5 进行聚类完全性判断。当聚类不完全时, 执行步骤 4.1—步骤 4.3, 否则输出聚类结果。

本文通过上述特征提取与特征聚类两个部分, 实现了车联网大数据混合属性特征检测, 为车联网大数据提供了极大的便利^[20-22]。

4 仿真实验设计

为了验证所提方法在多源异构环境下的车联网大数据混合属性特征提取中的应用效果, 本文将基于机器学习算法的特征提取方法、基于灰色神经网络与残差学习的特征提取

方法、基于堆叠去噪自动编码器和批量归一化的特征提取方法作为对比方法,由 Matlab2017 编程实现,进行仿真测试。

4.1 样本数据

多源异构环境下的车联网大数据类型和属性多样,大致可以分为3类,即汽车感知数据、驾驶行为数据以及外部周边环境数据,这3类又包括多个属性。为保证测试仿真结果的准确性,本文构建了基础样本数据,结果如表2所列。

表2 仿真测试样本数据
Table 2 Simulation test sample data

类型	内容	属性特征数量	样本数量
汽车感知数据	车况状态	12	5656
	行驶位置	2	3455
	行驶轨迹	24	3210
驾驶行为数据	车辆故障记录	20	4540
	里程数	25	5440
	油耗情况	4	4547
	行驶时间	6	5634
周边环境数据	事故违章记录	30	7545
	道路类型	23	6440
	地形条件	22	5520
	天气状况	12	4878
	行人情况	5	6440

4.2 样本数据预处理

按照3.2节中的处理过程,对表2所列的多源异构环境下的车联网大数据进行预处理,处理结果如图2所示。

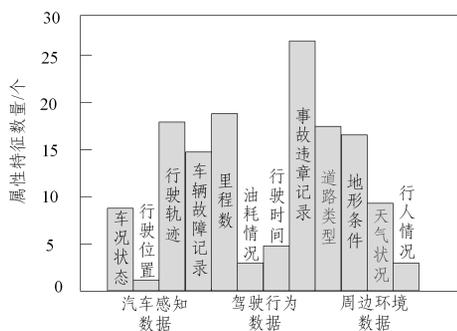


图2 样本数据属性特征约简处理结果

Fig. 2 Reduction results of sample data attribute feature

4.3 特征检测评价指标

选择两个指标对特征提取方法进行评测,这两个指标分别用于检测特征提取方法的准确性和效率。下面针对这两个指标进行具体分析。

(1) 敏感性指数

敏感性指方法正确检出特征的样本数量与总数量之间的比例开方与错误检出数量之间的比值。敏感性越大,检测方法对特征越敏感,准确性越高。

$$\alpha = \sqrt{\frac{x}{X}} / y \quad (9)$$

其中, α 表示敏感性指数; x 表示正确检出特征的样本数量; X 表示总样本数量; y 表示错误检出样本数量。

(2) 时间复杂度

时间复杂度指执行算法进行车联网大数据混合属性特征检测所需要的计算工作量。时间复杂度越小,该方法进行特征检测的效率越高。

$$\beta = m^2 n + m^2 S^3 + \frac{x}{m + P + Nm - NP} \quad (10)$$

其中, β 表示时间复杂度; m 表示样本中属性的总个数; n 表示数据集中的样本数量; S 表示方法的运行时间; P 表示对应样本中属性的数量; N 表示属性值个数的最大值; x 表示正确检出特征的样本数量。

(3) 查全率

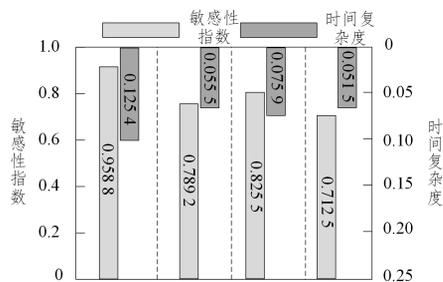
查全率(Recall Ratio)指从数据库中检出的相关的信息量与总量的比率。查全率的绝对值很难计算,只能根据数据库内容、数量来估算。查全率越高,该大数据混合属性特征检测的全面性越理想,检测精度越高。

$$Recall Ratio = \frac{K}{L} \times 100\% \quad (11)$$

其中, K 表示检索出的大数据混合属性特征检测量, L 表示大数据混合属性特征总量。

4.4 结果分析

在相同测试条件下,利用本文方法与对比方法根据检测结果统计敏感性指数和时间复杂度来评判检测方法的有效性,结果如图3所示。



注:从左往右依次为本文方法、基于机器学习算法的特征提取方法、基于灰色神经网络与残差学习的特征提取方法、基于堆叠去噪自动编码器和批量归一化的特征提取方法的应用结果

图3 本文方法与各对比方法的敏感性指数和时间复杂度
Fig. 3 Comparison of sensitivity index and time complexity

从图3可以看出,本文方法的敏感性指数为0.9588,时间复杂度为0.1254,与基于机器学习算法的特征提取方法、基于灰色神经网络与残差学习的特征提取方法、基于堆叠去噪自动编码器和批量归一化的特征提取方法的应用结果相比,本文方法表现更好。

在相同测试条件下,利用本文方法以及文献[3]提出的基于机器学习算法的特征提取方法、文献[4]提出的基于灰色神经网络与残差学习的特征提取方法进行特征检测查全率测试,实验结果如图4所示。

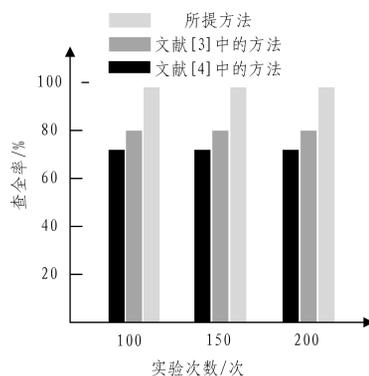


图4 不同方法的查全率对比

Fig. 4 Comparison of recall rates of different methods

由图4可知,文献[3]中的方法和文献[4]中的方法的查全率为70%~80%,在此查全率区间的特征抽取方法实现了基本的特征抽取功能,但无法满足当下高精度特征的应用要求。本文通过仿真实验验证了本文方法的有效性,其查全率始终保持在95%以上。

结束语 本文针对因车联网数据规模大而处理较困难的问题,提出了多源异构环境下的多属性特征检测方法。该方法通过对车联网大数据混合属性进行处理和提取,使得数据规范存储,以便应用。最终,仿真测试也证明了所研究的检测方法更高效、准确地完成车联网大数据混合属性特征提取任务。

本文提出的多源异构环境下的车联网大数据混合属性特征检测方法虽然取得了一定的成果,具有一定的优越性,但仍存在不足。在仿真测试中所使用的数据为小型且理想的数据,与实际数据集相比质量更好,且所提方法的运行速率还有待提高,因此有待在实际环境中进一步检验该方法的实际应用性。未来将针对网络大数据混合属性特征的约简速率进行深入研究,从约简速率提升的角度来提高属性特征的检测速度,在保证检测准确性的基础上进一步提高检测效率。

参 考 文 献

- [1] GUO J, KIM S, WYMEERSCH H, et al. Guest Editorial: Introduction to the Special Section on Machine Learning-Based Internet of Vehicles: Theory, Methodology, and Applications [J]. IEEE Transactions on Vehicular Technology, 2019, 68(5): 4105-4109.
- [2] YAO Z H, JIANG Y S. Integrated Connected Vehicle Simulation Platform of Vissim and Python [J]. Computer Simulation, 2018, 35(12): 143-146, 405.
- [3] SARUMATHY C K, GEETHA K, RAJAN C. Improvement in Hadoop performance using integrated feature extraction and machine learning algorithms [J]. Soft Computing, 2020, 24(1): 627-636.
- [4] TAN X W. Feature Extraction Algorithm Based on Big Data Depth Residual Learning [J]. Bulletin of Science and Technology, 2019, 35(4): 89-92.
- [5] WU Y X, WU Q B, ZHU J Q. Data-driven wind speed forecasting using deep feature extraction and LSTM [J]. IET Renewable Power Generation, 2019, 13(12): 2062-2069.
- [6] ZHANG D Y, LUO Y M. Simulation of Mixed Attribute Feature Detection for Large Network Data Based on Rough Set [J]. Computer Simulation, 2021, 38(1): 460-463, 485.
- [7] WU F F. Feature Extraction of Cosmetic Packaging Symbol Elements Based on Big Data Clustering [J]. China Surfactant Detergent & Cosmetics, 2020, 50(1): 44-48.
- [8] LIN Q Y. Big Data Multi-label Attribute Classification Technology in Cloud Service Environment [J]. Microelectronics & Computer, 2019, 36(2): 101-104.
- [9] LI Y, HU J C. Outlier Detection Based on Neighborhood Granule for Mixed Attribute Data [J]. Journal of Chinese Computer Systems, 2020, 41(4): 855-860.
- [10] LIU W J, JI W X, ZHANG C Y. Big Data Modeling Analysis Method for Intelligent Production Maintenance [J]. China Mechanical Engineering, 2019, 30(2): 37-44.
- [11] YANG P, SHEN H T, TAO P, et al. Parallel permutation entropy feature extraction method for time series data based on cloud platform [J]. Electric Power Automation Equipment, 2019, 39(4): 217-223.
- [12] DONG J W, WU W B, HUANG J X, et al. State of the Art and Perspective of Agricultural Land Use Remote Sensing Information Extraction [J]. Journal of Geo-Information Science, 2020, 22(4): 772-783.
- [13] ZHANG B, XIONG C B. Automatic Point Cloud Registration Based on Voxel Downsampling and Key Point Extraction [J]. Laser & Optoelectronics Progress, 2020, 57(4): 109-117.
- [14] WANG Y, LI J, ZENG H, et al. Real-time Feature Extraction Algorithm Based on Mutual Information [J]. Journal of Chinese Computer Systems, 2019, 40(6): 1242-1247.
- [15] MADDUMALA V, ARUNKUMAR R. Big Data-Driven Feature Extraction and Clustering Based on Statistical Methods [J]. Traitement du Signal, 2020, 37(3): 387-394.
- [16] ZHAO B, LI S, GAO Y, et al. A Framework of Combining Short-Term Spatial/Frequency Feature Extraction and Long-Term IndRNN for Activity Recognition [J]. Sensors, 2020, 20(23): 6984.
- [17] LI Y, CHEN X, YU J. Feature extraction and classification of ship radiated noise based on VMD and SVM [J]. Guofang Keji Daxue Xuebao/Journal of National University of Defense Technology, 2019, 41(1): 89-94.
- [18] HEITMANN A J, GARDINER-GARDEN R S. A Robust Feature Extraction and Parameterized Fitting Algorithm for Bottom-Side Oblique and Vertical Incidence Ionograms [J]. Radio Science, 2019, 54(1/2): 115-134.
- [19] LIANG L, DING X, LIU F, et al. Feature Extraction Using Sparse Kernel Non-Negative Matrix Factorization for Rolling Element Bearing Diagnosis [J]. Sensors, 2021, 21(11): 3680.
- [20] NURMAINI S, TUTUKO B, RACHMATULLAH M N, et al. Machine Learning Techniques with Low Dimensional Feature Extraction for Improving the Generalizability of Cardiac Arrhythmia [J]. IAENG International Journal of Computer Science, 2021, 48(2): 369-378.
- [21] GOVINDARAJAN S, RAGAVAN V, EL-HAG A, et al. Development of Hankel Singular-Hypergraph Feature Extraction Technique for Acoustic Partial Discharge Pattern Classification [J]. Energies, 2021, 14(6): 1564.
- [22] LV Z, LLORET J, SONG H. Guest Editorial Software Defined Internet of Vehicles [J]. IEEE Transactions on Intelligent Transportation Systems, 2021, 22(6): 3504-3510.



CHEN Jing, born in 1982, master, engineer. His main research interests include digitalization and informatization of engineering industry.



WU Ling-ling, born in 1976, Ph.D, associate professor. Her main research interests include transport engineering and so on.