



移动扫码阅读

张元刚,刘坤,杨林,等.煤炭工业监控大数据平台建设与数据处理应用技术[J].煤炭科学技术,2019,47(3):75-80.doi:10.13199/j.cnki.cst.2019.03.010

ZHANG Yuangang, LIU Kun, YANG Lin, et al. Platform construction and data processing application technology in coal industry monitoring big data [J]. Coal Science and Technology, 2019, 47 (3): 75 - 80. doi: 10.13199/j.cnki.cst.2019.03.010

煤炭工业监控大数据平台建设与数据处理应用技术

张元刚^{1,2},刘坤¹,杨林^{1,2},王磊³

(1.兖矿集团有限公司 信息化中心,山东 邹城 273500;2.北斗天地股份有限公司,陕西 西安 710000;3.兖州煤业股份有限公司 兴隆庄煤矿,山东 兖州 272000)

摘要:为保障煤矿的安全高效生产,及时准确地预测可能发生的故障及灾害,通过分析兖矿集团煤矿大数据的现状及存在的问题,探讨了在大数据时代下,煤炭工业监控大数据平台的建设和使用方法。依托互联网、云计算和大数据技术,采集、整理和存储海量数据。通过数据挖掘技术和视频分析技术来探索发掘现场生产规律,从而提高煤矿应对潜在灾害和发现安全隐患的能力。以煤矿带式输送机工况监控系统大数据和视频监控系大数据分析为例,分别使用流形空间的非平衡处理技术和分水岭等视觉技术对数据进行处理。试验结果表明:流形空间的过采样技术可以提高 AUC 指标 20% 左右,且基于机器视觉的煤流检测也能达到较好的效果。同时,探讨了大数据与煤炭工业监控系统的结合,并提出了报警数据融合应用的具体解决方法。

关键词:大数据;工业监控大数据;视频大数据分析;带式输送机;煤矿智能化

中图分类号:TD67 **文献标志码:**A **文章编号:**0253-2336(2019)03-0075-06

Platform construction and data processing application technology in coal industry monitoring big data

ZHANG Yuangang^{1,2}, LIU Kun¹, YANG Lin^{1,2}, WANG Lei³

(1. Information Center, Yankuang Group, Zoucheng 273500, China; 2. Beidou Tiandi Co., Ltd, Xi'an 710000; 3. Xinglongzhuang Coal Mine, Yanzhou Coal Industry Co., Ltd., Yanzhou 272000)

Abstract: In order to ensure the safe and efficient production of coal mines, and forecast the faults and the disasters timely and accurately, this paper analyzes the current situation and existing problems of coal mines big data in Yankuang Group, and explores the construction and application of big data platform for coal industry monitoring. Massive data is collected, processed and stored by using the internet, cloud computing and big data technology. Through data mining technology and video analysis technology to explore the production rules of the site, the ability of coal mines can be improved in dealing with potential disasters and identifying potential safety hazards. Taking the big data analysis of the coal mine belt conveyor monitoring system and video monitoring system as examples, the data is processed by using the manifold space imbalance processing technology and watershed algorithm. The experimental results show that the manifold space imbalance processing technology can increase the AUC index by about 20%, and the coal flow detection based on machine vision can also achieve better performance. Meanwhile, the combination of big data and coal industry monitoring system is discussed, and the specific solution of alarm data fusion application is proposed.

Key words: big data; industrial monitoring big data; video big data analysis; belt conveyer; coal mine

0 引言

大数据概念最早是在 20 世纪 80 年代提出的,

大数据研究的核心并非数据的量有多大,而是如何有效、有序、系统地处理大数据,包括对数据的访问、收集、存储、管理、分析、挖掘、辅助决策等^[1]。目前

收稿日期:2018-10-12;责任编辑:赵瑞

基金项目:国家重点研发计划资助项目(2017YFC0804409)

作者简介:张元刚(1972—),男,山东章丘人,高级工程师,现任兖矿集团有限公司信息化中心主任。E-mail:zhangyuangang@ykjt.cn

大数据已经逐渐渗透到各个领域,各行业纷纷加快大数据的应用规模,提升企业的竞争力。虽然中国煤矿行业的黄金十年已经过去,但煤炭的安全可靠生产依然是我国经济社会发展的重要保证^[2]。然而煤矿行业的事故和灾害很大程度上影响了人员的安全和企业的生产。并且迄今为止,煤与瓦斯突出、冲击地压等事故发生的规律和机理还没有完全掌握,所以不能通过机理分析准确预警和预防这些灾害^[3]。大数据技术可以在机理不明的情况下,通过对数据的分析,挖掘出事物潜在的机理,因此使用大数据技术,来预警预防煤与瓦斯突出等事故灾害,具有十分重要的理论意义和实用价值。

文献[4]使用云计算对煤矿冲击地压进行研究。文献[5]使用一个半监督过采样的框架来从非平衡数据分类的角度处理冲击地压问题。文献[6]对互联网+煤矿开采大数据技术做了深入的研究和分析。中国煤炭的现状和大数据在煤炭行业应用的现状和展望在文献[7-8]中进行了较为详细的说明。文献[9]研究了煤矿安全生产大数据分析方法与平台设计的方法。上述工作有些是对煤矿行业的大数据发展现状和未来方向做了系统性的说明,有些是针对具体的煤炭生产问题和灾害防治,结合大数据方法做了科学的分析和试验,也有些是对煤炭结合大数据的方式和系统布局进行研究。总之,大数据与煤炭的结合是一个行业的趋势,对煤炭的安全生产有重大的意义。

基于此,笔者针对煤炭生产存在的潜在灾害和故障危险,以兖矿集团有限公司(以下简称兖矿集团)的煤炭安全监控大数据系统为例,结合大数据技术和煤炭安全监控系统,做了一些有益的尝试,来发现潜藏的安全隐患或对可能的故障灾害进行预警。综上,在新时代下,结合大数据技术来提高煤矿企业的管理水平,预防甚至避免瓦斯、水害等事故的发生,减少人员伤亡和财产损失是非常必要的。

1 煤矿大数据

1.1 煤矿大数据含义

煤矿大数据是指大数据在煤炭行业的应用或煤矿与大数据的结合。具体来说,煤矿大数据研究范围很广,例如:如何应用大数据技术解决煤矿行业中关于安全生产的问题、如何结合大数据预测或预警煤矿发生的瓦斯爆炸、水灾等重大灾害;如何使用大数据技术解决煤矿企业的安全生产管理的问题等^[10]。

1.2 兖矿集团煤矿大数据现状

目前,兖矿集团在煤矿大数据方面做了一定的探索,并建设了较好的大数据基础设施,包括矿井的

互联网布置,云平台的设置,监控摄像头和多元传感器的安装等^[11]。

1) 兖矿集团工业信息化建设基础较好,六大系统建设完备。矿井综合自动化系统、工业视频监控系统等早已布置完善。目前,兖矿集团已经建立了覆盖采矿、选矿、机运、通防等方面的监测监控、过程控制的数字化系统。

2) 兖矿集团正在加大基础传感网络和云平台的建设。包括传输网、交换网、数据网等,基本实现了集团内部语音、数据、视频和监控数据的可靠安全传输和存储;并且,部分矿井正在积极推进矿井云系统的建设,调整传统服务器部署方式,提高了计算资源的使用效率。

3) 兖矿集团的经营管理积极应用大数据技术。人力资源、财务、资金、保险、ERP、SRM、OA、综合调度、计量发运等系统已经覆盖全部矿井;除此之外,安全生产调度指挥系统、安全生产实时数据联网系统、工业视频监控等系统也已经实施布置完成,为煤矿提供了人员管理、物资采购、资金计划等业务服务。

1.3 兖矿集团煤矿大数据存在的问题

虽然,兖矿集团在积极应用大数据技术,并且在基础设施建设,人才培养与引进等方面做了很多的努力,但目前依然存在一些问题^[11]。

1) 信息化的基础设施需要进一步升级和改造。与电力、石化等行业相比,煤矿信息化的建设依然落后,相关规范和制度也不健全。

2) 各级部门和工人对大数据的认识依然不足。煤矿的工作人员对煤矿工业大数据的了解不够深入,大都只停留在煤矿总体运营数据上,对煤矿生产和安全监控数据没有较深的研究和思考。

3) 工业大数据平台有待进一步建设和升级。煤炭布置的多元传感器和工业视频系统采集的数据量非常庞大,并且各种数据结构多样,实时性要求极高,所以大数据平台应用面临着许多的问题和挑战。因此,建设高容量、高复杂性、高可靠性的工业大数据平台和先进的分析平台是非常重要的。

4) 大数据的开放共享机制有待完善。目前,煤矿大量的数据在各个专业部门垂直系统、单位应用系统和生产运行子系统中交叉分布,有着很强的信息的耦合。但由于体制、利益等原因,很多的数据只能在本部门访问,没有实现数据的互通和共享。这就导致数据利用率和质量不高,数据浪费的问题。

5) 大数据人才队伍建设急需加强。为了结合大数据解决煤矿现有的问题,需要组建跨学科的复合型人才团队,培养具有煤矿和大数据背景的工作人员。目

前由于兖矿集团工业大数据处于起步阶段,人才培养不足,同时也缺乏相应的人员认定和评价标准。

近几年,随着矿井相关数据的快速增长,煤炭产业大数据已经形成并且发展迅猛,且具备了一定基础,行业发展优势明显,但也存在着数据体系不够完善、采集渠道不够畅通、数据时效性不够强等问题,已成为煤炭行业转型升级待解决的重点和难点问题^[12]。兖矿集团有限公司信息化中心在煤炭企业工业监控大数据的应用方面做了一些有益尝试。

2 基于工业大数据的煤流运输设备在线监测及预警分析系统

2.1 系统简介

随着煤矿生产设备自动化程度的不断提高,生产设备的安全运行已经成为煤矿生产中的关键问题。在兖矿集团兴隆庄煤矿的信息化系统中已经建立了包含监测仪表、数据采集、远程监控、数据集成、分析报告等应用系统,积累了海量的历史数据,包含了生产过程中的参数变化过程,也包含了生产设备在生产过程中的工况变化规律。该系统利用大数据分析,充分利用存储的海量数据,在分析煤矿带式输送机运行状态历史数据的基础上,全面分析故障类型和功能需求,设计较为全面的系统方案,该方案能更好地指导煤矿生产,保障输送带安全运行,具有较强的可实施性。

该系统建立基于设备温度、电动机电流、CST启动装置功率等在线状态监测及运维信息数据库,研究设备故障诊断模型及判据,分析设备历史数据,实现故障诊断和预测预警。根据故障及隐患诊断预警事件,建立具有高实时性的联动机制,可以实现管理人员、维护人员的短信报警,并可与数据发布系统、移动信息化系统及调度门户集成,实现实时机电设备健康状态显示及调度管理。系统总体功能结构如图1所示。

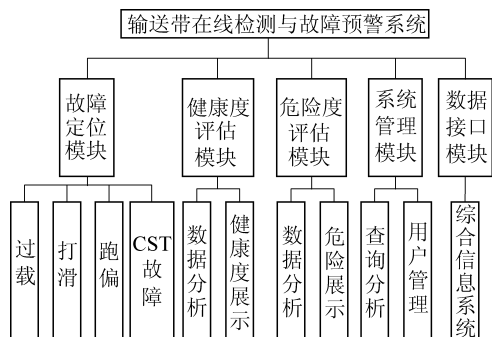


图1 煤流运输设备在线监测及预警分析系统总体功能结构

Fig.1 General function structure early warning of coal flow transportation equipment online monitoring and analysis system

2.2 技术方案

带式输送机作为许多行业的核心运输设备,其稳定安全运行对工业生产流程具有直接的影响。笔者以输送带过载故障检测为例,详细介绍其技术方案和试验过程^[13]。

由于带式输送机空载和轻载情况较少发生,所以主要讨论带式输送机过载这一故障的诊断。观察发现,原始样本中仅有89个过载数据,占总数据量的0.13%。显然,这是一个非常典型的非平衡数据集。因此使用大数据中非平衡学习^[14]的方法来解决这个问题。并且由于工业数据常存在高纬度、强耦合、强噪声的特点,首先使用流形学习^[15]的方法来提取高维数据的低维结构,并在低维流形空间使用过采样方法将非平衡数据集转换为平衡数据集,在平衡数据集上训练分类器。

2.3 试验分析

首先, Isomap (Isometric Feature Mapping)^[16]、LLE (locally linear embedding)^[17]、MLLE (modified locally linear embedding)^[18] 和 LPP (locally preserving projection)^[19] 四种经典的流形学习方法被使用来构造高维数据的低维嵌入,之后再低维结构上使用 SMOTE (synthetic minority over-sampling technique)^[20]、Borderline SMOTE^[21] 和 ADASYN (adaptive synthetic sampling approach)^[22] 来构造平衡数据集。最后使用粒子群算法优化的支持向量机模型来做数据的分类。评价指标使用非平衡学习中常用的精确率、回归率、F1值、ROC曲线和AUC指标^[20-22]。

试验结果如图2和表1所示。总体而言过采样方法对过载故障诊断性能的影响较小。而分类空间的选择对故障诊断的结果影响较大。相比而言,在 ISOMAP 和 LLE 空间学习到的分类器可以明显提升 F1 和 AUC 两个指标,因此证明在流形空间构造平衡数据集来进行输送带过载故障的诊断是合理且有效的。

3 煤矿输送带视频监控大数据智能识别分析系统

3.1 系统简介

智能分析技术近几年一直得到广泛关注,该技术的大规模应用对于提高安全防范系统及工业视频监控系统的利用效率会起到很大的促进作用,但实际上却没有得到有效的推广。造成这一现象的主要原因,一是现在智能视频分析的应用主要集中在特定的行业;二是针对实际应用中监控视频数量比较

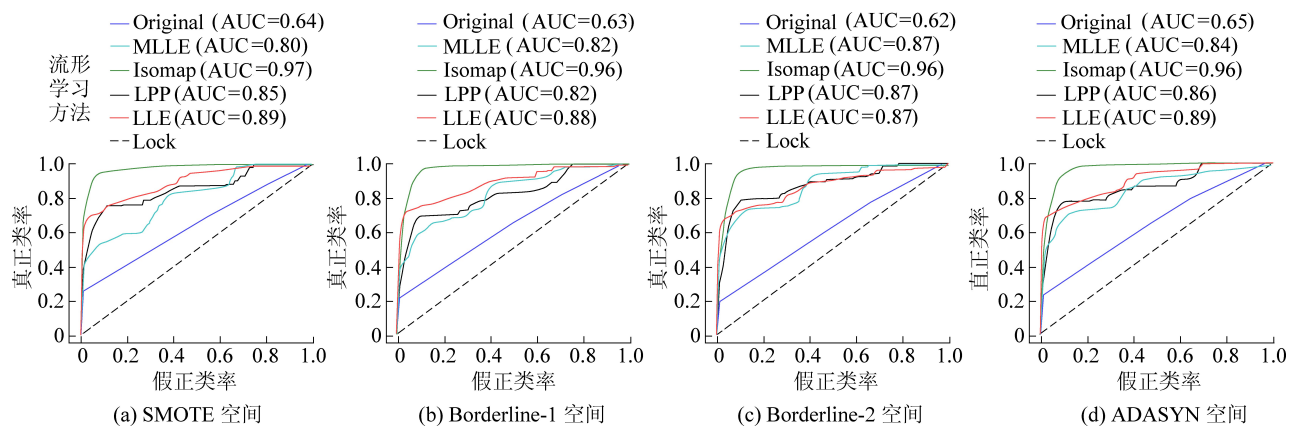


图2 带式输送机过载故障 ROC 曲线

Fig.2 ROC curve of overload fault in belt conveyor

表1 带式输送机过载故障的分类性能

Table 1 Overload fault classification performance of belt conveyor

过采样	空间	总正确率	精确率	回归率	F1 值	AUC 值
SMOTE	观测空间	0.937 5	0.987 3	0.317 8	0.475 6	0.644 8
	ISOMAP	0.976 5	0.841 6	0.928 9	0.879 1	0.972 2
	LLE	0.972 3	0.893 4	0.794 4	0.836 1	0.889 8
	MLLE	0.951 4	0.742 3	0.804 0	0.755 3	0.800 5
Bor1	观测空间	0.931 3	0.969 7	0.253 3	0.395 0	0.632 2
	ISOMAP	0.970 8	0.791 9	0.948 9	0.858 8	0.964 6
	LLE	0.967 9	0.849 3	0.803 3	0.819 8	0.883 1
	MLLE	0.930 8	0.651 6	0.782 7	0.692 5	0.819 0
Bor2	观测空间	0.926 8	0.971 9	0.201 1	0.327 1	0.618 0
	ISOMAP	0.959 0	0.718 7	0.947 8	0.811 0	0.960 6
	LLE	0.965 4	0.855 0	0.775 9	0.802 4	0.866 0
	MLLE	0.936 5	0.661 2	0.808 0	0.710 4	0.867 8
ADASYN	观测空间	0.930 3	0.985 6	0.237 9	0.378 3	0.645 5
	ISOMAP	0.970 7	0.771 9	0.977 8	0.860 6	0.964 4
	LLE	0.960 1	0.836 9	0.808 9	0.810 9	0.894 1
	MLLE	0.939 4	0.711 0	0.748 7	0.694 2	0.839 7
	LPP	0.939 5	0.622 4	0.873 3	0.724 8	0.861 1

大的工业项目,智能分析系统的构架无论从实际经验上,还是技术上都显得比较缺乏,无法完全消除误报的影响,行为判断能力弱,且视频数据帧难以达到特征识别技术要求。

目前国内大多数中大型煤矿的视频监控系统已经实现了全矿井上全覆盖,井下重点区域覆盖。综合显示系统也已经实现了视频数据的存储和显示功

能,但是现在视频监控系统和综合显示系统依赖值班人员的观察,在煤矿出现预警后,才去调取视频,并没有充分发挥视频监控系统的真正作用,从而让显示大屏变成了煤矿用来展示的面子工程。通过对视频图像进行分析,将捕捉的异常画面投放到显示大屏提示值班人员,发现重大问题时则启动报警联动,以广播、短信方式通知调度员和矿上的重要领导,并通过煤矿安全生产系统对现场的设备进行操作,第一时间减少损失,这无疑将为煤矿安全生产提供强有力的保障。智能视频分析系统总体功能结构如图3所示。

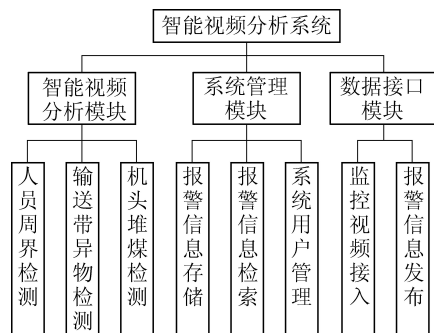


图3 智能视频分析系统总体功能结构

Fig.3 General structure of intelligent video analysis system

究矿集团通过数据整合可以集成企业现有的分布、异构的多个业务应用系统,该系统实现了对市场上主流工业摄像机的驱动程序整合,实现对外提供一致的JSON数据接口,实现系统其他模块的调用。并且可以实现智能异常识别分析报警,包括:

1) 周界报警:实现对单向和双向跨越周界时的行为自动侦测捕获并产生报警。

2) 进入识别:实现当监控视频画面中有人、车辆、器材等物体进入特定区域中时,立刻捕获并发出报警信号。

3) 离开识别:实现当人、车辆、器材等物体从某

个特定区域中离开时,立刻发现并产生离开识别报警。

4)场景变化识别:实现当摄像机被转动位置时,场景发生变化,能够自动捕获这类行为,重新识别背景画面,并发出报警信息。

5)滞留识别:实现当人、车辆、器材在某个危险区域内一定时间停留不走时,进行滞留识别报警,防止人员车辆在危险区域内长时间停留。

另外该系统还可以进行输送带异物检测,实现实时高效地对煤矿输送带上大块矸石、大块煤块、锚杆、钢钎等异物的自动捕获、识别、分析检测和标定。实现在监控画面上进行标识并发出报警信号,同时在数据库中存储原始关键帧、标定关键帧、报警时间戳、报警信息描述、异物标定数据等。除此之外,系统可以对异常信息进行存储,实现在线视频/图像智能分析识别算法处理,将发生异常的关键标注后的数据帧、关键原始数据帧、异常产生时间戳、异常描述等进行信息存储并为后续的异常处理和事后回溯提供数据支持。

3.2 技术方案

以输送带载煤测试为例,介绍机器视觉技术方案和试验过程。在煤矿现场,实时检测输送带上的负载等级和煤流宽度,可以避免输送带长时间在超负载的情况下运行,减少事故的发生。使用的方法可以分为以下7个步骤:

1)使用负载掩膜对输送带区域进行提取,并将提取之后的灰度图转换为BGR格式的彩色图。

2)对BGR格式的输送带区域彩色图像进行分水岭算法提取。

3)计算煤流区域占据像素个数,用像素点的个数代表煤流的面积。

4)使用煤流面积除以输送带区域的高度,计算出输送带区域的平均像素宽度,计算每个像素标识的近似实际距离。

5)计算煤流的实际宽度,更新历史输送带负载宽度,并更新标定数据。

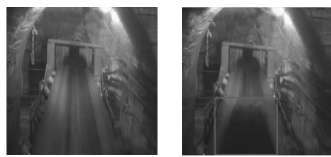
6)根据煤流实际宽度的大小,将输送带负载分为重载、正常、轻载和空载。

7)对前后两帧灰度图像相减之后的图像进行二值化,并计算输送带区域白色部分的像素个数,如果白色部分像素个数小于某一阈值,则前后两帧灰度图像基本相同,输送带处于停机状态。

3.3 试验分析

通过应用煤矿输送带视频监控大数据智能识别分析系统对输送带进行载煤测试,结果如图4所示。

机器视觉分析方法可以较好地将大堆煤炭煤流标记出来。从图4b可以看出,当煤流出现时,该算法可以较好地提取出煤堆。并且在此基础上,进一步可以通过视觉的方法来计算一定时间的煤量。



(a) 输送带无煤 (b) 输送带有煤

图4 输送带载煤测试结果

Fig 4 Coal detection result on coal mine belt

4 结 论

1)工业监控大数据是在工业领域信息化应用中产生的海量数据,工业监控大数据不但具有大数据的特点;还具有多源异构的结构化数据和非结构化数据并存,工业监控大数据持续采样频率高,数据价值密度大,数据间关联性强,具有动态多时空时间序列性,数据处理实时性要求高等特征。

2)笔者介绍了大数据分析和工业监控大数据分析,探讨了工业监控大数据分析的基本方法。

3)以煤炭行业大数据应用为例,通过有效整合和利用煤炭企业的多个数据集,建立起基于工业大数据的煤流运输设备在线监测、预警分析系统和煤矿输送带智能视频识别分析系统,初步形成了煤炭企业的工业监控大数据典型应用案例。

参考文献(References):

- [1] 吴亚坤,郭海旭,王晓明.大数据技术研究综述[J].辽宁大学学报:自然科学版,2015,42(3):236-242.
WU Yakun, GUO Haixu, WANG Xiaoming. Research overview on big data technology[J]. Journal of Liaoning University: Natural Science Edition, 2015, 42(3): 236-242.
- [2] 孙继平.煤矿事故分析与煤矿大数据和物联网[J].工矿自动化,2015,41(3):1-5.
SUN Jiping. Accident analysis and big data and internet of things in coal mine[J]. Industry and Mine Automation, 2015, 41(3): 1-5.
- [3] 孙继平.煤与瓦斯突出报警方法[J].工矿自动化,2014,40(11):1-5.
SUN Jiping. Alarm methods of coal and gas outburst[J]. Industry and Mine Automation, 2014, 40(11): 1-5.
- [4] 姜福兴,曲效成,王颜亮,等.基于云计算的煤矿冲击地压监控预警技术研究[J].煤炭科学技术,2018,46(1):199-206,244.
JIANG Fuxing, QU Xiaocheng, WANG Yanliang, et al. Study on monitoring & control and early warning technology of mine pressure bmp based on cloud computing[J]. Coal Science and Technology, 2018, 46(1): 199-206, 244.
- [5] 程健,杨凌凯,王全魁,等.基于半监督过采样非平衡学习的

- 矿山微震信号识别[J].煤炭科学技术,2018,46(2):213-218,202.
- CHENG Jian, YANG Lingkai, WANG Quankui, *et al.* Mine micro-seismic detection based on semi-supervised over-sampling framework[J]. Coal Science and Technology, 2018, 46(2): 213-218, 202.
- [6] 张科利,王建文,曹豪.互联网+煤矿开采大数据技术研究与实践[J].煤炭科学技术,2016,44(7):123-128.
- ZHANG Keli, WANG Jianwen, CAO Hao. Study and practice on big data technology of internet plus coal mining[J]. Coal Science and Technology, 2016, 44(7): 123-128.
- [7] 孙杰,陈美英,唐朝苗,等.我国煤炭资源勘查现状跟踪研究[J].中国煤炭地质,2017,29(11):1-8.
- SUN Jie, CHEN Meiyang, TANG Chaomiao, *et al.* Tracking study of coal resources exploration status quo in China[J]. Coal Geology of China, 2017, 29(11): 1-8.
- [8] 马小平,代伟.大数据技术在煤炭工业中的研究现状与应用展望[J].工矿自动化,2018,44(1):50-54.
- MA Xiaoping, DAI Wei. Research status and application prospect of big data technology in coal industry[J]. Industry and Mine Automation, 2018, 44(1): 50-54.
- [9] 刘香兰.煤矿安全生产大数据分析与管理平台设计研究[J].煤炭工程,2017,49(6):32-35.
- LIU Xianglan. Design and research of big data analysis and management platform for coal mine safety[J]. Coal Engineering, 2017, 49(6): 32-35.
- [10] 王海军,武先利.“互联网+”时代煤矿大数据应用分析[J].煤炭科学技术,2016,44(2):139-143.
- WANG Haijun, WU Xianli. Analysis on application of coal mine big data in age of ‘Internet +’ [J]. Coal Science and Technology, 2016, 44(2): 139-143.
- [11] 杨林,刘坤.兖矿集团煤矿工业大数据发展应用刍议[C].济南:中国煤矿信息化与自动化高层论坛,2017.
- [12] 李飒.大数据挖掘技术在煤炭企业安全管理中的应用[J].煤炭技术,2016,35(3):323-324.
- LI Sa. Application of big data mining technology in coal enterprise safety management[J]. Coal Technology, 2016, 35(3): 323-324.
- [13] 程健,杨凌凯,崔宁,等.基于流形嵌入过采样的非平衡数据分类方法[J].中国矿业大学学报,2018,47(6):1331-1339.
- CHENG Jian, YANG Lingkai, CUI Ning, *et al.* A novel classification method for imbalanced data set based on manifold embedded over-sampling [J]. Journal of China University of Mining & Technology, 2018, 47(6): 1331-1339.
- [14] GALAR M, FERNANDEZ A, BARRENECHEA E, *et al.* A review on ensembles for the class imbalance problem: bagging, boosting, and hybrid-based approaches [J]. IEEE Transactions on Systems Man & Cybernetics Part C Applications & Reviews, 2012, 42(4): 463-484.
- [15] LAW M H C, JAIN A K. Incremental nonlinear dimensionality reduction by manifold learning [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2006, 28(3): 377-391.
- [16] TENENBAUM J B, DE S V, LANGFORD J C. A global geometric framework for nonlinear dimensionality reduction [J]. Science, 2000, 290: 2319-2322.
- [17] ROWEIS S T, SAUL L K. Nonlinear dimensionality reduction by locally linear embedding [J]. Science, 2000, 290: 2323-2326.
- [18] SCHOLKOPF B, PLATT J, HOFMANN T. MLL: modified locally linear embedding using multiple weights [J]. Advances in Neural Information Processing Systems, 2006, 19: 1593-1600.
- [19] HE X, NIYOGI P. Locality preserving projections [C] // Advances in neural information processing systems. MIT Press, British Columbia, 2004: 153-160.
- [20] CHAWLA N V, BOWYER K W, HALL L O, *et al.* SMOTE: synthetic minority over-sampling technique [J]. Journal of Artificial Intelligence Research, 2011, 16(1): 321-357.
- [21] HAN H, WANG W Y, MAO B H. Borderline-SMOTE: a new over-sampling method in imbalanced data sets learning [J]. Lecture Notes in Computer Science, 2005(5): 878-887.
- [22] HE H, BAI Y, GARCIA E A, *et al.* ADASYN: adaptive synthetic sampling approach for imbalanced learning [C]. 2008 International Joint Conference on Neural Networks, Hongkong, China, 2008.