



文章编号:1671-251X(2018)01-0050-05

DOI:10.13272/j.issn.1671-251x.2018.01.2017100022

大数据技术在煤炭工业中的研究现状与应用展望

马小平, 代伟

(中国矿业大学 信息与控制工程学院, 江苏 徐州 221116)

摘要:在总结煤炭大数据大容量、快速性、多样性、真实性、可见性、价值等特征的基础上,从煤炭大数据理论、煤炭大数据与物联网和云计算的关系、煤炭大数据平台建设方面介绍了煤炭大数据研究现状;分析了煤炭大数据的多层面不均匀采样性、多时间尺度特性、不真实数据混杂性等新特点及其对煤炭大数据分析带来的挑战;从大数据采集与管理、大数据分析、大数据共享方面对建设煤炭大数据平台功能进行了展望,并从应用领域探讨了大数据技术在煤炭工业中可能的发展方向。

关键词:煤炭工业;大数据;物联网;云计算;煤炭大数据平台

中图分类号:TD67 文献标志码:A 网络出版时间:2017-12-25 14:14

网络出版地址:<http://kns.cnki.net/kcms/detail/32.1627.TP.20171225.0911.006.html>

Research status and application prospect of big data technology in coal industry

MA Xiaoping, DAI Wei

(School of Information and Control Engineering, China University of Mining and Technology, Xuzhou 221116, China)

Abstract:On the basis of characteristic of coal big data such as volume, velocity, variety, veracity, visibility and value, research status of coal big data was introduced from aspects of theory of coal big data, relationship among coal big data, Internet of things and cloud computing, and construction of coal big data platform. New features of coal big data such as multi-layer non-uniform sampling, multi-time scale characteristic and non-veracity data confounding, and challenge of coal big data analysis caused by the new features were analyzed. Functions of coal big data platform were prospected from aspects of big data collection and management, big data analysis and big data sharing. Possible development direction of big data technology in coal industry was discussed from application field.

Key words:coal industry; big data; Internet of things; cloud computing; coal big data platform

0 引言

随着传感器、计算机、通信、物联网、数据存储等技术的发展,以及企业信息管理系统的不断普及,制

造工业等行业产生并存储了大量数据,且随时间呈指数级增长^[1],工业界已经进入了“大数据”时代^[2-3],煤炭工业就是其中的一个典型代表。煤炭是中国的主体能源,其产业的健康发展对经济社会发

收稿日期:2017-10-16;修回日期:2017-12-12;责任编辑:盛男。

基金项目:国家自然科学基金资助项目(61603393);江苏省重点研发计划资助项目(BE2016046);江苏省自然科学基金资助项目(BK20160275);中国博士后科学基金资助项目(2015M581885);流程工业综合自动化国家重点实验室开放课题基金资助项目(PAL-N201706)。

作者简介:马小平(1961—),男,四川绵阳人,教授,博士研究生导师,博士,主要研究方向为矿山物联网、煤矿自动化、煤矿安全等,E-mail: xpma@cumt.edu.cn。

引用格式: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.

展至关重要,甚至关系着国家能源安全。因此,迫切需要依托物联网、云计算和大数据技术,采集、存储和挖掘海量数据,从数据中探索解决煤炭“采掘机运通排”中若干问题,推动煤炭行业由生产自动化、信息自动化转型升级为知识自动化。

麦肯锡的报告显示,就大数据的数量而言,诸如煤炭等领域的过程工业,其数据产出量及可被接入的设备数量远远超过移动互联网等其他行业,而且增速是其他大数据领域的2倍^[4]。然而,煤炭大数据的利用却远没有像当前移动互联网领域那样普遍和深入,其中数据的价值还有待挖掘。本文阐述了煤炭大数据的特征及研究现状,并在分析煤炭大数据新特点的基础上,对煤炭大数据平台的发展与应用进行了探讨与展望。

1 煤炭大数据特征

大数据分析手段主要是围绕大数据的“4Vs”(Volume——大容量, Velocity——快速性, Variety——多样性, Veracity——真实性)特性去发展与完善的^[5-6]。除“4Vs”特性外,煤炭大数据还应该加上“2Vs”,即 Visibility——可见性(通过大数据分析使以往隐匿的重要因素和信息可见), Value——价值(通过大数据分析得到的信息应该被转换成价值)。这“2Vs”代表了煤炭工业界对于大数据所追求的目的和意义。

(1) 大容量(Volume)体现在数据采样频率高(采样间隔为毫秒级)与采样时间段长(24 h不间断运行)所带来的大容量历史数据。以单台矿井通风机为例,其数据采样频率为50 Hz,每秒产生225 kB数据,按每年7 000 h工作时间计算,仅单台通风机每年生成6 TB数据。

(2) 快速性(Velocity)体现在高频的数据采集。由于井下作业环境恶劣、24 h不间断生产,生产环境安全监测监控系统、各生产环节的自动化系统等实时数据与信息需要安全、准确、实时地传至数据处理中心,所以每秒钟会记录几千兆甚至上万兆字节的数据。

(3) 多样性(Variety)体现在异构性和不同时间尺度、不均匀采样。异构性是指煤炭大数据除“采掘机运通排”控制系统采集的结构化生产过程数据外,还包括以生产环境在线监测为主的视频图像、语音,以及规章制度、应急案例文本等非结构化数据。此外,煤炭工业采用分层次运行方式,采集的时间序列数据既有高维且快速动态采样的压力、流量等过程

数据,又有低速、不均匀采样的灰分、硫分等指标数据。

(4) 真实性(Veracity)体现在真实数据与离群点数据的混杂。由于煤炭掘进、开采和洗选过程生产环境恶劣,测量仪表或变送器受干扰严重,且故障频发,使得测量数据中混杂不真实数据,存在离群点、缺失点等异常样本。如重介质选煤过程中,由于在线灰分仪运行不稳定,经常导致灰分实际测量值出现大偏差,导致历史数据中出现离群点。

(5) 可见性(Visibility)体现在对隐匿性问题的建模和预测。设备性能下降、健康衰退、零部件磨损等问题难以通过测量被量化,而大部分可见的问题都是这些不可见因素积累到一定程度所造成的,因此需要通过大数据分析使以往隐匿的重要因素和信息可见,从而避免可见问题的发生。

(6) 价值(Value)体现在数据密度高、价值小。煤炭生产中,系统经常在某一特定且正常工况下运行,而故障等异常工况只是偶然发生,因此大量数据均是存在冗余的正常运行数据,而反映异常的数据非常少。

2 煤炭大数据研究现状

从20世纪90年代至今,煤矿自动化经历了从单机自动化、综合自动化到煤矿物联网的发展^[7],且随着煤炭企业信息化与自动化的不断提高及两化融合的快速发展,安全监控、人员定位、可视通信、数字化矿山、三维建模、企业资源计划等系统不断应用,产生了海量、不同层次、不同类别的数据资源,为大数据在煤炭工业中的应用奠定了基础^[8]。

由于缺少针对煤炭大数据特征的分析工具及高效的计算平台来提取隐匿的知识,煤炭大数据还未充分利用,目前主要是对数据进行采集、压缩存档、恢复与预测分析。其中预测分析主要集中在煤炭安全生产领域,但尚处于理论研究和起步阶段,没有成功的模式和案例。郑磊^[9]分析了当前中国煤矿安全管理存在的问题,并提出大数据将对煤矿安全管理带来变革。丁振等^[10]讨论了大数据在变革管理思维、增强系统安全观念,提高设备运转可靠度、实现设备健康运行,提供事故分析新视角、实现安全管理关口前移等方面的应用前景。

大数据、物联网与云计算是提升煤矿安全生产水平的3个重要技术手段。胡英^[11]分析了安全生产大数据的特点与煤矿安全生产大数据当前面临的主要问题,提出建设基于物联网、云计算技术的煤矿

安全综合数据库,建设基于专家系统的煤矿安全专家知识库,建设三维虚拟矿井可视化平台,研发煤矿安全动态分析系统的需求。孙继平^[12]分析了煤炭大数据和物联网对煤矿事故分析的重要作用,讨论了大数据在煤与瓦斯突出、冲击地压、水害、火灾等事故预警,煤矿重大关键设备故障诊断,煤炭需求和价格预测等方面的应用。刘玉海^[13]分析了现有的Hadoop 云计算平台及云计算技术用于实现煤矿井下人员定位系统、矿井瓦斯安全预警系统、煤矿安全生产应急系统的可行性。缪建华等^[14]提出了基于大数据和云计算技术的智慧矿山计划,并对淮南煤矿安全高效生产运用前景进行了展望。马小平等^[15]阐述了物联网、大数据及云计算技术的研究现状,并指出3种技术之间的关系,即物联网产生大数据、大数据助力物联网,大数据需要云计算、云计算增值大数据。

当前,大数据已经成为国家的战略资源和推动产业发展的重要引擎,煤炭工业的安全化、绿色化、智能化发展对煤炭大数据平台的需求迫在眉睫。张茜^[16]利用SWOT分析方法,从优势、劣势、机会和威胁4个方面对煤炭企业建立大数据平台进行了分析。刘香兰^[17]搭建了煤矿安全生产大数据分析模型,研究并设计了煤矿安全生产大数据分析与管理平台的功能架构与技术架构。申琢等^[18]以数据集成、挖掘和可视化展现为主线,设计并开发了由设备层、网络层、数据层、挖掘层和展示层组成的基于数据挖掘的煤炭大数据可视化管理平台。2016年7月19日,煤炭大数据平台V2.0^[19]在“2016夏季全国煤炭交易会”上正式启动,该平台涵盖煤炭GIS、图解煤炭、煤炭物流、信用查询、预测预警、数据分析、能源经济七大功能模块,但仍存在数据体系不完善、采集渠道不畅通、数据时效性不强等问题。

3 煤炭大数据新特点及研究挑战

随着信息化发展和可编程逻辑控制器与集散控制系统的广泛采用,以及安全生产过程管理、设备操作优化需求的不断增加,煤炭生产向大规模、集成化发展,多单元、多产品生产、动态运行。煤炭大数据体现出与广泛过程工业大数据^[20]类似的多层面不均匀采样性、多时间尺度特性、不真实数据混杂性的新特点,对已有研究方法带来挑战。

(1) 多层面不均匀采样性。煤炭生产过程既有高维动态的过程数据,又有不均匀采样的指标数据,比如灰分指标往往难以在线测量,通常通过人工化

验获得,而人工化验时间在一定范围内具有随机性,导致数据具有大延迟和不均匀采样的特点。

现有数据驱动方法多集中在对均匀采样数据的分析与处理,无法对不均匀采样的数据进行分析与处理。

(2) 多时间尺度特性。多时间尺度体现在不同系统根据需求进行分层采样,进而进行优化决策与控制。其中,设备层是整个生产系统的最底层,实时监测设备运行状态,其通常以毫秒为单位采集设备数据(如重介质选煤过程的人料泵变频电流等);上一层为过程控制层,以秒级采集过程控制数据(如重介密度、合介筒液位等),并实施控制,以保证过程的关键变量按照给定的设定值变化,并驱动设备状态按指定逻辑顺序变化,是稳定生产的关键;再上一层为运行优化层,采集从产品质量分析仪或人工化验或统计得到的多种类且不均匀采样的产品质量、产量、能耗等运行指标数据(如灰分、硫分,视频监控带式输送机运煤表面图像数据等),由优化策略进行分析计算,产生新的过程设定值;最上层为生产经营管理层,采集来自市场营销和售后服务与用户的反馈数据(如市场对煤质的定量与定性评价数据),进行市场分析,从而指导和协调生产。

多时间尺度上的数据具有强自相关与互相关关系,以及时间序列相关关系^[21],但不同时间尺度采样使得所获取的数据具有不均匀稀疏性,对现有数据分析方法带来挑战。

(3) 不真实数据混杂性。高维动态历史数据中夹杂与过程特性无关的干扰数据,这是因为受恶劣环境影响,煤炭实际生产过程数据易受采集、传输、存储过程中的异常影响,从而混杂离群点甚至缺失点。这在不同时间尺度采样中均存在,如过程控制层的传感器设备故障,以及运行优化层的化验与人为读数误差等。

现有数据分析方法需要无污染的数据,但不真实数据的存在不可避免。因此,如何从混杂数据中提取出反映过程特性本质的模型,建立鲁棒数据分析方法,将是研究的重点和难点。

4 煤炭大数据展望

4.1 煤炭大数据平台展望

近几年,煤炭大数据技术发展较快,具备了一定基础,但当前中国煤炭数据信息采集渠道仍不畅通,数据传达相对滞后,时效性不强,缺少有效的分析工具及高效的分析方法实现煤炭大数据的价值。因

此,煤炭大数据平台必须重点实现大数据采集与管理、大数据分析、大数据共享功能。

(1) 大数据采集与管理。大数据的基础在于数据,因此需充分利用各类传感器、无线网络等加强对生产过程、监测监控、设备运行、经营管理等各类数据的采集与管理。此外,标准统一也是大数据的基础。没有统一的行业信息标准,就无法整合来自不同领域、企业、类型的数据和信息。按照煤炭大数据建设的核心理念,建立全国煤炭数据实时在线采集系统,对各省市现有煤炭监测系统进行整合,形成基础数据库,并使各系统能够以统一的接口进行数据传输,实现基层煤矿数据的实时更新。同时,依托物联网、大数据技术,将煤炭行业数据全部纳入采集系统,实现各类煤炭数据的全国联网,构建囊括煤炭生产、消费、交易等相关信息的全国煤炭数据库。

(2) 大数据分析。对大容量数据中知识的挖掘是大数据技术的关键所在,然而,解决问题的知识通常具有隐匿性,其不是直接呈现在数据里,而是呈现在用于揭示数据的模型中。因此,煤炭大数据的核心在于采用降维、分类与聚类、时序数据模式挖掘、数据源融合处理及关联关系分析等数据分析方法,快速获取、分析、处理海量的工业生产过程监控数据,建立不同层次(过程监测、故障诊断、优化决策、控制等)的模型,从而推动煤炭安全管理思维变革,增强系统安全观念,推动安全监察监管信息的共享,加快打通信息源壁垒,推动数据的深度挖掘,为设备运行优化提供理论支撑。对于容量呈指数级增加的大数据,传统数据分析方法无法处理或处理过慢,因此煤炭大数据分析需借鉴已有高速的大数据建模与优化算法,如面向大数据特征的深度学习算法^[22-23]等。

(3) 大数据共享。依靠某个企业单一数据样本很难实现大数据价值,需建立全国煤炭行业各类数据汇总的总平台,使本行业各类信息平台与电力等其他行业相关平台进行数据接入与共享,充分发挥行业协会、煤炭企业、科研院所、信息技术厂商等各方力量。

4.2 煤炭大数据应用展望

煤炭工业应根据自身大数据特点,结合工业大数据在建模、控制与优化技术方面研究成果,利用互联网产业已建立的大数据管理平台,推进煤炭行业各级部门间数据的共享,面向煤炭安全生产、过程优化与决策、控制、故障诊断进行应用实践。

(1) 基于大数据的煤矿智能无人开采。通过积

累各类地质构造、煤岩煤质、大型采掘设备等的大数据样本,构建各类开采条件下的智能开采模型。

(2) 基于大数据的煤矿灾害预警与防治。通过积累煤矿大量监测数据样本,可探索基于大数据的煤矿事故预警方法,将大数据用于煤与瓦斯突出、冲击地压、水害、火灾等煤矿事故预警。

(3) 基于大数据的煤矿设备故障诊断。通过监测煤矿设备振动、声音、温度、功率等设备运行状态数据,研究关键设备异常状态与设备故障的关系,实现基于大数据的煤矿设备故障诊断。

(4) 基于大数据的行业运行监测与预测。通过积累煤炭市场交易数据、供需企业经营数据、港口铁路运输数据、国际大宗商品市场数据、宏观经济及下游行业运行数据等,进行煤炭需求与价格预测。

(5) 基于大数据的煤矿生产全流程一体化。通过不同时间尺度、不均匀采样,可充分挖掘设备层、过程控制层、运行优化层、生产经营管理层不同层级内与层级间的隐匿性知识,设计集全流程生产指标、运行指标、过程设定值与控制规律的优化控制方法,实现生产全流程一体化,从而降低煤炭生产的能耗与物耗,提高产品质量。

5 结语

煤炭大数据还处于理论探索阶段,已建立的煤炭大数据平台尚未能挖掘大数据隐含的价值。同时,煤炭大数据的多层面不均匀采样性、多时间尺度特性、不真实数据混杂性的新特点对大数据分析所带来的挑战性问题,是煤炭大数据技术进一步研究的重点和难点。在煤炭行业开展大数据应用,利用煤炭大数据平台,探寻煤炭产业发展规律,探索煤炭行业科学管理方法,成为煤炭乃至整个能源行业发展的必然要求和必经之路。

参考文献(References):

- [1] SYED A R, GILLELA K, VENUGOPAL C. The future revolution on big data[J]. International Journal of Advanced Research in Computer and Communication Engineering, 2013, 2(6): 2446-2451.
- [2] HALEVI G, MOED H. The evolution of big data as a research and scientific topic: overview of the literature [J]. Research Trends: Special Issue on Big Data, 2012 (30): 3-6.
- [3] 李学龙, 龚海刚. 大数据系统综述[J]. 中国科学: 信息科学, 2015, 45(1): 1-44.
LI Xuelong, GONG Haigang. A survey on big data

- systems[J]. *Scientia Sinica Informationis*, 2015, 45(1):1-44.
- [4] 周兴社. 工业大数据特点、价值及其计算[EB/OL]. (2017-10-13) [2017-10-15]. <https://wenku.baidu.com/view/e68d87d6e518964bce847c0e.html>.
- [5] BARWICK H. IIIS: the 'four Vs' of big data[EB/OL]. (2011-08-05) [2017-10-15]. https://www.computerworld.com.au/article/396198/iiis_four_vs_big_data/.
- [6] 苏鑫, 吴迎亚, 裴华健, 等. 大数据技术在过程工业中的应用研究进展[J]. *化工进展*, 2016, 35(6): 1652-1659.
- SU Xin, WU Yingya, PEI Huajian, et al. Recent development of the application of big data technology in process industries [J]. *Chemical Industry and Engineering Progress*, 2016, 35(6): 1652-1659.
- [7] 丁恩杰, 赵志凯. 煤矿物联网研究现状及发展趋势[J]. *工矿自动化*, 2015, 41(5): 1-5.
- DING Enjie, ZHAO Zhikai. Research advances and prospects of mine Internet of things[J]. *Industry and Mine Automation*, 2015, 41(5): 1-5.
- [8] 石宏杰. 基于大数据时代下煤矿管理信息化系统规划[J]. *通讯世界*, 2017(14): 254-255.
- SHI Hongjie. Plan of coal mine management information system in the era of big data[J]. *Telecom World*, 2017(14): 254-255.
- [9] 郑磊. 大数据在煤矿安全管理中的应用研究[J]. *无线互联科技*, 2016(11): 133-135.
- ZHENG Lei. Study on application of big data in coal mine safety management [J]. *Wireless Internet Technology*, 2016(11): 133-135.
- [10] 丁振, 张骥. 浅析大数据技术助力煤矿安全管理[J]. *中国煤炭*, 2015, 41(10): 121-123.
- DING Zhen, ZHANG Lin. Brief analysis of big data technology contributing to coal mine safety management[J]. *China Coal*, 2015, 41(10): 121-123.
- [11] 胡英. 大数据与煤矿安全生产浅析[J]. *现代工业经济和信息化*, 2015, 5(21): 92-93.
- HU Ying. Analysis on the big data and coal mine safety production[J]. *Modern Industrial Economy and Informationization*, 2015, 5(21): 92-93.
- [12] 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.
- [13] 刘玉海. 云计算及煤矿大数据应用研究[J]. *信息系统工程*, 2016(9): 136.
- LIU Yuhai. Application research of cloud computing and coal mine big data[J]. *China CIO News*, 2016(9): 136.
- [14] 缪建华, 李俊斌. 大数据和云计算在煤矿安全高效开采运用前景展望[J]. *煤炭与化工*, 2017, 40(8): 134-136.
- MIAO Jianhua, LI Junbin. Prospect of applying big data and cloud computing in safe and efficient mining of coal mine[J]. *Coal and Chemical Industry*, 2017, 40(8): 134-136.
- [15] 马小平, 胡廷军, 缪燕子. 物联网、大数据及云计算技术在煤矿安全生产中的应用研究[J]. *工矿自动化*, 2014, 40(4): 5-9.
- MA Xiaoping, HU Yanjun, MIAO Yanzi. Application research of technologies of Internet of things, big data and cloud computing in coal mine safety production [J]. *Industry and Mine Automation*, 2014, 40(4): 5-9.
- [16] 张茜. 煤炭企业大数据平台建设的 SWOT 分析[J]. *现代经济信息*, 2016(19): 107.
- ZHANG Qian. SWOT analysis of big data platform construction in coal enterprises[J]. *Modern Economic Information*, 2016(19): 107.
- [17] 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 production[J]. *Coal Engineering*, 2017, 49(6): 32-35.
- [18] 申琢, 谭章禄. 基于数据挖掘的煤矿大数据可视化管理平台研究[J]. *中国煤炭*, 2016, 42(12): 86-89.
- SHEN Zhuo, TAN Zhanglu. Research on big data visual management platform of coal mine based on data mining[J]. *China Coal*, 2016, 42(12): 86-89.
- [19] 中国煤炭市场网. 煤炭大数据平台 V2.0 全新上线[EB/OL]. (2016-07-19) [2017-09-23]. <http://www.mycoal.cn/news/21/100118.html>.
- [20] QIN S J. Process data analytics in the era of big data [J]. *AIChE Journal*, 2014, 60(9): 3092-3100.
- [21] 刘强, 秦泗钊. 过程工业大数据建模研究展望[J]. *自动化学报*, 2016, 42(2): 161-171.
- LIU Qiang, QIN Sizhao. Perspectives on big data modeling of process industries[J]. *Acta Automatica Sinica*, 2016, 42(2): 161-171.
- [22] 张清辰. 面向大数据特征学习的深度计算模型研究[D]. 大连: 大连理工大学, 2015.
- [23] LYU Yisheng, DUAN Yanjie, KANG Wenwen, et al. Traffic flow prediction with big data: a deep learning approach [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2015, 16(2): 865-873.