

基于数据驱动的电力安全生产事故 风险预警研究

杨 军¹, 王 东¹, 许 洁², 李朋磊³, 米传民³

(1. 国家电网有限公司安全质量监察部, 北京 西城 100032;

2. 国网江苏省电力有限公司无锡供电分公司, 江苏 无锡 214000;

3. 南京航空航天大学经济与管理学院, 江苏 南京 211106)

摘 要:智能电网的快速发展和广泛应用为电力企业提供了来源复杂、结构多样的海量数据,使得智能电网成为大数据最重要的应用领域之一。如何应用大数据技术实现对电力安全生产数据的采集、存储和挖掘,进而提高电力企业的安全生产水平成为当前重要的研究课题。本文首先阐述了电力大数据和大数据技术的基本概念,然后借鉴传统的“海因里希法则”的思想,从隐患的角度出发,运用大数据技术对我国某省电力安全生产数据进行处理,通过这些数据建立某省电力企业的安全事故比例模型。再通过回归分析,对未来可能存在的隐患数量进行预测。最后根据安全事故比例模型对未来可能发生事故、事件数量做出预测,确定隐患数量的控制目标,形成一套安全生产预警模型,从而达到消除隐患、减少事故发生目的。

关键词:安全事故;隐患;大数据;海因里希法则;预警

文章编号:2096-4633(2019)04-0009-06 **中图分类号:**X93 **文献标志码:**B

智能电网是将传统电网与先进的信息基础设施相结合,实现电网和终端用户之间双向能量流动的現代电网基础设施^[1]。支撑智能电网安全、自愈、绿色和坚强及可靠运行的基础是电网全景实时数据采集、传输和存储,以及累积的海量多源数据快速分析^[2]。先进的通信和数据处理技术给智能电网带来了巨大的好处^[3]。而传统的电网数据来源渠道不足,数据分析挖掘能力欠缺,造成管理过程中面临着“数据海量、信息匮乏”的现象,已经远远不能满足日常安全管理工作的要求。因此,需要采用大数据技术对电力企业数据进行有效的数据管理,进而提高电力数据的综合分析能力,为电力企业的决策和管理提供有价值的信息^[4]。

目前,大数据技术已经成功应用包括金融、制造业、医疗、能源等领域。近年来,国内外许多学者探索了大数据技术在智能电网中的应用,并取得了不错的成果。黄天恩等人提出了一种基于电网安全运行仿真大数据的知识管理方法,将实时电网运行数据与历史数据进行实时匹配,实现了快速、准确的电网超前安全预警^[5]。朱永庆等人

利用大数据分析技术设计,针对不同故障类型设计并完成了一套整体的智能化故障诊断系统,使其具备了良好的处理实际网络故障的能力^[6]。陈碧云等人基于关联规则进行电力生产安全事件大数据关键诱因的选择,挖掘出可靠关联规则,对于辅助电力生产安全态势感知及控制决策的提出具有重要的意义^[7]。上述研究从不同方面利用大数据技术,目的都是为电力企业的安全生产提供有效的指导,提高电力企业安全能力。总的来说,智能电网大数据的研究是一个长期而复杂的工作,目前还处于探索的起步阶段。大数据相关的技术仍在快速发展中,在其研究方法和应用价值方面还有很大的发展空间^[8]。

目前我国安全生产形势总体上稳定向好,但电力安全生产仍然面临着严峻形势。据国家能源局统计,2018 年 5 月至 10 月期间,全国发生电力人身伤亡事故 23 起,死亡 24 人,重伤 3 人。反映了我国电力企业安全生产中隐患大量存在、安全生产法规措施执行不到位、预防措施不到位、预警不及时等问题。从理论上说,只要识别出隐患,并基于隐患建立

预警来控制隐患的数量,就能达到控制和预防事故发生的目的。因此本文以隐患为抓手,以海因里希法则为指导,旨在利用大数据技术实现电力安全生产预警模型的建立。

1 电力大数据

大数据是指无法在一定时间内用常规软件对其内容进行抓取、管理和处理的数据集合,它包括数字、图片、文本、视频、交互记录等。具体地说,电力大数据是指通过传感器、智能化设备、视频监控设备、音频通信设备和移动终端等各种数据采集渠道收集到的结构化、半结构化和非结构化的海量数据,涉及到发电、输电、变电、配电、用电、调度各环节^[9]。结构化的数据主要包括存储在关系数据库中的数据,非结构化数据主要包括视频监控、图形图像处理等产生的数据,非结构化数据在智能电网中占据很大的比重,需要对其进行处理如语音识别、图像识别等以转化为结构化数据^[10]。

电力大数据满足大数据的五个特性,即数据量大(Volume)、处理速度快(Velocity)、数据类型多(Variety)、价值密度低(Value)以及精确性高(Veracity)。此外电力大数据还应满足易用性和安全性。大数据技术的本质是从关联复杂的数据中挖掘知识,获取有效信息,提升数据价值。通过大数据技术可以提高数据采集效率、挖掘数据价值,能够为电力企业预警工作的实现发挥巨大作用^[11]。

1.1 数据采集和预处理技术

设备故障一直是引起电网事故的主要源头,全面、及时、准确掌握电力设备运行状态是保障电网安全运行的首要问题。因此智能电网中数据量最大的应属于电力设备状态监测数据,它主要包括电网运行数据、故障和缺陷记录、气象信息、带点检测和在线监测数据、技术参数等。这些数据来源广泛,类型丰富,并且数据采集点越来越多,采集频率也越来越高,导致数据量爆发式增长。传统的以人工录入为主的数据采集方式已不能满足企业的发展要求。而数据采集又称数据获取,是指从传感器和其他待测设备以及数字被测单元中自动采集信息的过程,它能够完成异构多源的海量数据的采集,在智能电网中被广泛应用。

电力应用系统中存在不同的平台、应用系统

以及不同的数据格式,导致电力公司内部不同的系统信息资源分散,横向不能共享、上下级之间纵向贯通困难,形成“信息孤岛”的局面。加之在采集的数据中存在大量如视频、图像、报告文本等形式的非结构化和半结构化数据,更难以实现异构多源数据的融合。因此,必须要对采集的数据进行预处理,包括使大量的非结构化和半结构化数据转化为结构化的格式、状态数据质量评估和数据清洗等过程,实现数据的同质性,提高数据处理的效率^[12]。

1.2 数据存储技术

在大数据时代中,数据量的处理已经成为考量数据库最重要的原因之一。过去很长一段时间内,关系型数据库一直是占据着很重要的位置。但是多源异构海量数据的环境下,关系型数据库已经不再是处理问题的最佳选择。比如关系型数据库不善于处理非结构化的数据、关系型数据库无法大规模扩展、存储数据的花费较高等。大数据存储管理中的一个重要技术是NoSQL数据库技术,它采用分布式数据存储方式,去掉了关系型数据库的关系型特性,数据存储被简化且更加灵活,具有良好的可扩展性,解决了海量数据的存储难题^[13]。此外基于Hadoop文件系统的分布式文件处理系统存储框架和并行编程框架Hadoop MapReduce擅长智能电网海量数据的存储和利用,能够使PB、ZB级的数据存储成为可能,并为应用程序提供高吞吐量的数据访问^[14]。

1.3 数据挖掘和可视化技术

电力大数据的数据挖掘技术就是指从大量的、模糊的、有噪声的、随机的数据中找到里面隐藏的信息,这些信息是人们无法直接从原始数据中发现的。从庞大复杂的数据中快速而有效地分析提取出有用的信息,成为智能电网应用中的一个关键技术难点。当今的云计算和云存储模型是数据挖掘和知识发现成果积累的基础,能让数据实现更大的价值^[15]。

数据可视化主要是借助图形化技术,清晰有效地传达与沟通信息。电力大数据可视化是电力大数据价值直观展现的有效方式。例如智能电网中的绝大部分数据都与空间位置有直接联系,开发以地理信息图为中心的数据可视化展示,能更好地实现对

电网数据与网架结构的综合展示,帮助人们更直观准确地了解数据表达的意义^[16]。

2 我国电力安全事故比例模型

利用大数据技术将日常工作中涉及人员的视频、图像等数据转换成结构化数据,结合电力设备状态监测数据、电网运行数据、故障和缺陷记录、气象信息等数据,对这些数据进行上文提到的数据采集、存储和分析处理,找出其中存在的隐患、事件和事故,基于“海因里希法则”得到电力安全事故比例模型。

2.1 安全生产数据统计

根据《国家电网公司安全事故调查规程》规定,安全事故(事件)共分八级,依次为特别重大事故(一级事件)、重大事故(二级事件)、较大事故(三级事件)、一般事故(四级事件)、五级事件、六级事件、七级事件、八级事件。统计数据量越大,得到的规律越能反映整体的趋势,在实际应用中才能越程度地发挥指导作用。理论上若能统计到全国电网中各省的隐患、事故数据,就能得到整个电力行业的宏观比例关系。本文以某省的统计数据作为研究分析对象,将其按照上述规定分为隐患、事件(5-8级事件)和事故(1-4级事件),表1为2013-2017年隐患、事件及事故数量,事故总量为0,事件总量为454,隐患总量为44 379。

表1 某省2013-2017年事故、事件及隐患数量
Tab.1 Number of hidden dangers, incidents and accidents in a province from 2013 to 2017

年份	事故	事件	隐患
2013	0	102	10 320
2014	0	104	10 821
2015	0	86	8 845
2016	0	97	8 912
2017	0	68	7 481
合计	0	454	44 379

2.2 安全事故比例模型

海因里希事故致因理论和海因里希法则在安全生产事故研究方面起着重要的作用。海因里希事故致因理论的主要内容是从大量典型事故的本质原因中提炼出事故机理和事故模型。这些机理和模型反

映了事故发生的规律性,能够为事故原因的定性、定量分析以及事故的预测预防从理论上提供科学、完整的依据。事故致因理论还指出造成事故和人员伤亡的直接原因是人的不安全行为和物的不安全状态^[17]。“海因里希法则”则是在统计分析5 000起事故的基础上得到了无伤、轻伤、重伤事故300:29:1的比例,具体的含义是在330次事故中,按照伤害程度进行统计,其中会有1次严重伤害、29次轻微伤害和300次无伤害事故。在安全生产过程中人的不安全行为和物的不安全状态都可以称之为“隐患”,那么海因里希法则用比例关系表明了隐患和事故之间的关系^[18]。

通过对表1中的数据进行统计分析之后可得,得到某省2013-2017年的安全事故模型如图1所示,可以看出:某省事故:事件:隐患=0:454:44 379,表明近五年来没有发生事故,发生了454起事件,44 379个隐患。

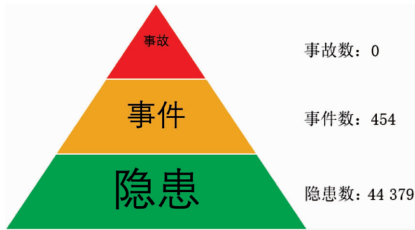


图1 安全事故比例模型
Fig.1 Safety accident ratio model

3 电力生产安全隐患回归分析预测

回归分析是一种统计分析的方法,目前在各个领域应用的比较广泛,它通过对观测数据进行观察研究,建立起各变量之间的某种依赖关系,从而能对数据的内在规律进行分析,在此基础上还可以对未来的发展变化情况进行预测。如图2所示为2013-2017年隐患总量散点图。其中自变量是年份,因变量是隐患数量。

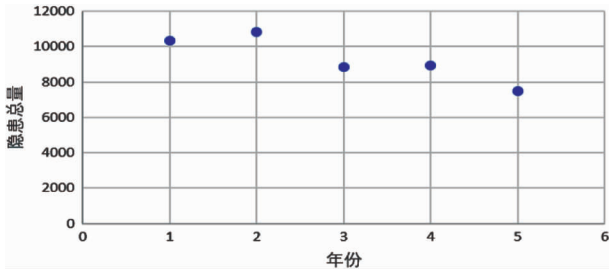


图2 隐患总量散点图
Fig.2 Total scatter diagram of hidden dangers

从图中可以看出,近几年来隐患的数量总体上呈现下降的趋势。利用 SPSS 工具对其进行函数拟合计算,如表 2 所示得到 3 种拟合度较高的拟合方式,分别为线性拟合、对数拟合和指数拟合。线性拟合方程为,拟合度为 0.820 3;对数拟合方程为,拟合度为 0.689 2;指数拟合方程为,拟合度为 0.826 4。拟合度 R² 越高表明回归方程的拟合度越好,可见指数拟合的拟合度最高,因此最终确定求得的回归方程为。利用求出的指数回归方程,x 的取值为 6,代入公式可预测某省 2018 年的隐患总量为 7144,约占上年隐患总量的 95%,由上述的安全事故比例模型推测 2018 年可能发生的事件数量为 73。

表 2 拟合程度表
Tab. 2 Fitting degree scale

拟合方式	拟合方程	拟合度 R ² %
线性拟合	$y = -758.7x + 11\,552$	82.03
对数拟合	$y = -1\,730 \ln(x) + 10\,933$	68.92
指数拟合	$y = 1\,1826 e^{-0.084x}$	82.64

4 电力生产安全预警模型

我国电网动态安全评估和预警系统(dynamic security assessment system,PDSA)项目在 2004 年已经启动。PDSA 的主要目标是实现电网运行的在线检测与预警,及时系统的安全隐患,并提出解决措施^[19-20]。PDSA 系统改变传统的基于典型方式进行离线稳定分析的模式,实现了在线整合数据进行稳定分析,并在分析隐患的基础上给出预警方案。但是 PDSA 系统并没有涉及隐患和运行事故之间的内在联系,无法对由于隐患存在可能导致的事故数量做出判断和预测。海因里希法则恰恰说明了隐患和事故之间的关系,所以本文建立了基于海因里希法则的电力生产安全预警模型。

将当前某省隐患数量按月累计,与去年同期隐患累计总量进行比较,并根据比较结果设定预警程度。根据预测可知 2018 年隐患总量为 7 144,约占去年隐患总量的 95%,那么规定若当前隐患累计总量不超过去年同期累计隐患数量的 95% 处于正常区,在 95% - 100% 之间处于警告区,超过 100% 处于危险区。如图 3 所示,绿色区

间表示正常区,黄色区间表示警告区,红色区间表示危险区。

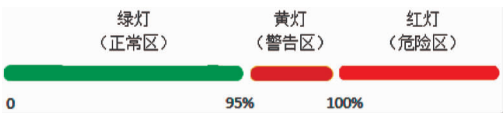


图 3 安全生产预警指示图

Fig. 3 Warning indicator diagram of safety production

5 结果分析

表 3 某省 2017 年 1 - 5 月份隐患数量
Tab. 3 Number of hidden dangers from January to December in 2017

月份	1	2	3	4	5	总量
隐患	573	567	627	588	584	2 939

表 4 某省 2018 年 1 - 5 月隐患表
Tab. 4 Hidden danger list of a province from January to May 2018

月份	1	2	3	4	5	总量
隐患	557	576	634	591	572	2 930

某省 2017 年统计的 1 - 12 月份的隐患数量如表 3 所示,2018 年 1 - 5 月份的隐患数量如表 4 所示,2018 年 1 - 5 月份隐患统计总量为 2 930,超过 2017 年同期隐患总量 2 939 的 95%,由图 3 可知某省现在处于警告区,说明某省的安全生产情况不太乐观。由上述的安全事故比例模型可知,要消除多个隐患才可能避免一起事件的发生,在目前累计隐患数量偏多的情况下,企业需强化监督检查力度,加大隐患排查力度,从源头控制隐患,减少事故的发生次数。

6 小结

本文在海因里希法则研究的基础上,通过统计近几年的隐患和事故数据,建立了电力安全事故比例模型。但由于目前采集的数据还不够全面,模型在实践中的参考价值有待提升。未来研究的重点主要有三个方面:一是建立隐患标准知识库,对历来的典型隐患以及随着智能电网发展新出现的隐患进行统一整理,以便快捷地在工作中发现尽可能多的隐患;二是采用大数据相关技术提高企业数据采集、数据挖掘的效果;三是模型的优化,如从季节、月份甚至更小的时间单位去统计数据,或从隐患所引发的

不同事故类型方面去统计数据,建立精确度更高的模型。让企业安全生产预警真正成为安全管理的重要工具。

参考文献:

- [1] KAUR, D; AUJLA, GS; KUMAR, N; et al. Tensor-based big data management scheme or dimensionality reduction problem in smart grid systems: SDN perspective [J]. IEEE Transactions on Knowledge and Data Engineering, 2018, 30(10): 1985–1998.
- [2] 宋亚奇, 周国亮, 朱永利. 智能电网大数据处理技术现状与挑战[J]. 电网技术, 2013, 37(04): 927–935.
SONG Yaqi, ZHOU Guoliang, ZHU Yongli. Present status and challenges of big data processing in smart grid [J]. Power System Technology, 2013, 37(04): 927–935.
- [3] JUN WU, KAORU OTA, MAINXIONG DONG, et al. Big data analysis-based security situational awareness for smart grid [J]. IEEE Transactions on Big Data, 2018, 4(03): 408–417.
- [4] 冷喜武, 陈国平, 蒋宇, 等. 智能电网监控运行大数据分析系统的数据规范和数据处理[J]. 电力系统自动化, 2018, 42(19): 169–176.
LENG Xiwu, CHEN Guoping, JIANG Yu, et al. Data specification and processing in big-data analysis system for monitoring and operation of smart grid [J]. Automation of Electric Power Systems, 2018, 42(19): 169–176.
- [5] 黄天恩, 孙宏斌, 郭庆来, 等. 基于电网运行仿真大数据的知识管理和超前安全预警[J]. 电网技术, 2015, 39(11): 3080–3087.
HUANG Tianen, SUN Hongbin, GUO Qinglai, et al. Knowledge management and security early warning based on big simulation data in power grid operation [J]. Power System Technology, 2015, 39(11): 3080–3087.
- [6] 朱永庆, 黄新宇, 唐宏, 等. 利用大数据分析技术的智能化网络故障诊断系统[J]. 电讯技术, 2018, 58(10): 1115–1120.
ZHU Yongqing, HUANG Xinyu, TANG Hong, et al. An intelligent network fault diagnosis system via big data analysis technique [J]. Telecommunication Engineering, 2018, 58(10): 1115–1120.
- [7] 陈碧云, 丁晋, 陈绍南. 基于关联规则挖掘的电力生产安全事故事件关键诱因筛选[J]. 电力自动化设备, 2018, 38(04): 68–74.
CHEN Biyun, DING Jin, CHEN Shaonan. Selection of key incentives for power production safety accidents based on association rule mining [J]. Electric Power Automation Equipment, 2018, 38(04): 68–74.
- [8] 张东霞, 苗新, 刘丽平, 等. 智能电网大数据技术发展研究[J]. 中国电机工程学报, 2015, 35(01): 2–12.
ZHANG Dongxia, MIAO Xin, LIU Liping, et al. Research on development strategy for smart grid big data [J]. Proceedings of the CSEE, 2015, 35(01): 2–12.
- [9] 鄢斌, 陈宾, 杨春麟, 等. 浅谈大数据管理在基层供电企业的应用与发展[J]. 电力大数据, 2018, 21(01): 32–34.
- YAN Bin, CHEN Bin, YANG Chuilin, et al. Talking about the application and development of big data management in grassroots power supply enterprises [J]. Power System and Big Data, 2018, 21(01): 32–34.
- [10] 王钦, 蒋怀光, 文福拴, 等. 智能电网中大数据的概念、技术与挑战[J]. 电力建设, 2016, 37(12): 1–10.
WANG Qin, JIANG Huaiguang, WEN Fushuan, et al. Concept, technology and challenges of big data in smart grids [J]. Electric Power Construction, 2016, 37(12): 1–10.
- [11] 薛禹胜, 赖业宁. 大能源思维与大数据思维的融合(一)大数据与电力大数据[J]. 电力系统自动化, 2016, 40(01): 1–8
XUE Yusheng, LAI Yening. Integration of macro energy thinking and big data thinking part one big data and power big data [J]. Automation of Electric Power Systems, 2016, 40(01): 1–8.
- [12] 江秀臣, 盛戈倬. 电力设备状态大数据分析的研究和应用[J]. 高电压技术, 2018, 44(04): 1041–1050.
JIANG Xiuchen, SHENG Gehao. Research and application of big data analysis of power equipment condition [J]. High Voltage Engineering, 2018, 44(04): 1041–1050.
- [13] 彭小圣, 邓迪元, 程时杰, 等. 面向智能电网应用的电力大数据关键技术[J]. 中国电机工程学报, 2015, 35(03): 503–511.
PENG Xiaosheng, DENG Diyu, CHENG Shijie, et al. Key technologies of electric power big data and its application prospects in smart grid [J]. Proceedings of the CSEE, 2015, 35(03): 503–511.
- [14] 宋亚奇, 周国亮, 朱永利, 等. 云平台下输变电设备状态监测大数据存储优化与并行处理[J]. 中国电机工程学报, 2015, 35(02): 255–267.
SONG Yaqi, ZHOU Guoliang, ZHU Yongli, et al. Storage optimization and parallel processing of condition monitoring big data of transmission and transforming equipment based on cloud platform [J]. Proceedings of the CSEE, 2015, 35(02): 255–267.
- [15] YIYING ZHANG, SUXIANG ZHANG. The power big data applications for intelligent community in smart grid [J]. Journal of Internet Technology, 2016, 17(07): 1491–1500.
- [16] 曲朝阳, 熊泽宇, 颜佳, 等. 基于空间分割的电力大数据三维全景可视化场景管理方法[J]. 华北电力大学学报(自然科学版), 2016, 43(02): 23–29.
QU Zhaoyang, XIONG Zeyu, YAN Jia, et al. Scene management method of three-dimensional panoramic visualization of electric power big data based on space partition [J]. Journal of North China Electric Power University, 2016, 43(02): 23–29.
- [17] 李杰, 陈伟炯. 海因里希安全理论的学术影响分析[J]. 中国安全科学学报, 2017, 27(09): 1–7.
LI Jie, CHEN Weijiong. Analyzing academic effect of Heimlich's safety theory [J]. China Safety Science Journal, 2017, 27(09): 1–7.
- [18] 刘国愈, 雷玲. 海因里希事故致因理论与安全思想因素分析

- [J]. 安全与环境工程,2013,20(01):138-142.
- LIU Guoyu, LEI Ling. Analysis of heinrich accident-causing theory and the factors of safety ideology [J]. Safety and Environmental Engineering,2013,20(01):138-142.
- [19] 严剑峰,于之虹,田芳,等. 电力系统在线动态安全评估和预警系统[J]. 中国电机工程学报,2008(34):87-93.
- YAN Jianfeng, YU Zhihong, TIAN Fang, et al. Dynamic security assessment & early warning system of power system [J]. Proceedings of the CSEE,2008(34):87-93.
- [20] 杨玥,张胜军,康琪. 基于电网运维数据的智能预警系统设计[J]. 内蒙古电力技术,2017,35(04):20-23.

YANG Yue, ZHANG Shengjun, KANG Qi. Design of intelligent early warning system based on grid operation and maintenance data[J]. Inner Mongolia Electric Power,2017,35(04):20-23.

收稿日期:2019-01-12

作者简介:



杨 军(1974),男,硕士,高级工程师。主要从事电力大数据分析、电力安全信息化建设的研究工作。

(本文责任编辑:范 斌)

Research on the risk early warning of power production safety accident based on data drive

YANG Jun¹, WANG Dong¹, XU Jie², LI Penglei³, MI Chuanmin³

(1. Safety and Quality Supervision Department of State Grid Co., Ltd., Beijing 100032, China;;

2. State Grid Jiangsu Electric Power Co., Ltd., Wuxi Power Supply Branch, Wuxi 214000 Jiangsu, China;

3. School of Economics and Management, Nanjing University of Aeronautics and Astronautics, Nanjing 211106 Jiangsu, China)

Abstract: The rapid development and wide application of smart grid provide massive data with complex sources and diverse structures for power enterprises, making smart grid become one of the most important application fields of big data. How to apply big data technology to achieve data collection, storage and mining, and then improve the level of safe production of power enterprises has become an important research topic. This paper first expounds the basic concepts of big data and big data technology of electric power, then draws lessons from the traditional "Heinrich rule" and proceeds from the perspective of hidden trouble, uses big data technology to process the massive data generated in the safe production of power in a province of China, through which a safety accident proportion model of electric power enterprises in a province is established. Then, through regression analysis, the number of potential dangers in the future is predicted. Finally, according to the proportion model of safety accidents, the number of possible accidents and accidents in the future is predicted, and determine the control target of the number of hidden dangers, and in the end a set of production safety early warning model is formed in order to eliminate hidden dangers and reduce accidents.

Key words: safety accident; hidden danger; big data; early warning