

基于数据挖掘的火电企业 安全生产和技术监督评价系统

高腾飞, 门凤臣, 刘宝军, 宋敬霞

(国家电投集团电站运营技术(北京)有限公司, 北京 102209)

[摘要] 火电企业传统评价工作模式长期存在共享度低、随意性大、规范性差、缺乏监控、分析困难等问题, 为推进火电企业评价依据标准化、任务执行程序化、问题整改科学化、现场操作便捷化、数据分析简单化, 本文提出了推荐专家规则排序算法, 构建了评价数据分析模型, 利用信息化手段, 研发了基于数据挖掘的火电企业安全生产和技术监督评价系统, 实现了PC端与移动端协同管理。并通过实际应用, 验证了该系统能有效降低人力物力财力, 提高整体工作效率, 保障机组安全、可靠、经济运行, 可为企业决策提供数据支撑。

[关键词] 火电企业; 评价系统; 安全生产; 技术监督; 信息化; 数据挖掘; 协同管理

[中图分类号] TP311.5; TM769 **[文献标识码]** B **[DOI编号]** 10.19666/j.rlf.202102089

[引用本文格式] 高腾飞, 门凤臣, 刘宝军, 等. 基于数据挖掘的火电企业安全生产和技术监督评价系统[J]. 热力发电, 2021, 50(9): 49-54. GAO Tengfei, MEN Fengchen, LIU Baojun, et al. Safety production and technical supervision evaluation system in thermal power enterprises based on data mining[J]. Thermal Power Generation, 2021, 50(9): 49-54.

Safety production and technical supervision evaluation system in thermal power enterprises based on data mining

GAO Tengfei, MEN Fengchen, LIU Baojun, SONG Jingxia

(SPIC Power Plant Operation Technology (Beijing) Co., Ltd., Beijing 102209, China)

Abstract: For a long time, there are many problems in conventional evaluation mode of thermal power enterprises, such as low sharing degree, large randomness, poor standardization, lack of monitoring, and difficult in analysis. In order to promote the standardization of evaluation basis, the procedure of task execution, the scientific problem rectification, the convenience of field operation and the simplification of data analysis for thermal power enterprises, this paper proposes ranking algorithm of recommended expert rules, and constructs an evaluation data analysis model. By using information means, the safety production and technical supervision and evaluation system for thermal power enterprises based on data mining is developed, and the collaborative management between PC terminal and mobile terminal is realized. Through practical application, it verifies that this system can effectively reduce manpower, material and financial resources, improve the overall work efficiency, and ensure the safety, reliability and economic operation of the unit, which can provide data support for enterprise decision-making.

Key words: thermal power enterprise, evaluation system, safety production, technical supervision, informatization, data mining, collaborative management

电力企业安全生产和技术监督是保证电力系统安全、可靠、经济运行和准确计量的重要技术手段^[1-3], 国家电力投资集团有限公司每年定期组织火电企业开展春、秋季安全生产和技术监督评价工作。传统评价工作模式存在信息共享度低、评价随

意性大、报告规范性差、“事中事后”缺乏实时监控、数据综合分析手段不足等弊端, 难以满足当今工作节奏快、效率高、机制新的管理态势, 无法支撑实现自动化、科学化和精细化的管理。

基于上述原因, 拟采用信息化手段, 实现安全

收稿日期: 2021-02-17

基金项目: 国家电力投资集团有限公司项目(100001JX0120190127)

Supported by: Science and Technology Project of SPIC (100001JX0120190127)

第一作者简介: 高腾飞(1987), 男, 硕士, 工程师, 主要研究方向为火力发电技术监督, gaotf1@126.com。

生产和技术监督评价工作的数据采集记录和传输存储。另外,评价准备阶段需要对大量历史数据进行分析处理,以对最优专家和重点评价内容做出正确的预测和评估。数据挖掘技术正是以积累下来的历史数据为研究对象,通过对数据的归类、分析、处理,从而找出隐藏在其中的有用知识^[4-5]。面对数据处理难度的增加,本文研发了基于数据挖掘的火电企业安全生产和技术监督评价系统,为准备、评价、监督和决策提供数据基础。

1 相关技术

1.1 数据挖掘

数据挖掘指知识发现的过程,通常包括数据清洗、数据集成、数据选择、数据变换、模式发现、模式评估和知识表示。其采用数据源包括数据库、数据仓库、Web、其他信息存储库或动态流入系统的数据^[6-7]。数据挖掘与传统的数据分析(如查询、报表、联机应用分析)的本质区别在于数据挖掘是在没有明确假设的前提下去挖掘信息、发现知识。所得到的信息(或知识)具有先前未知、有效和实用3个特征^[8-9]。

数据挖掘过程是以用户为中心,开展人机交互的探索。包含数据准备(data preparation)、数据挖掘以及结果的解释评估(interpretation and evaluation)3个阶段,数据挖掘可以描述为这3个阶段的反复过程。

1.2 移动协同

移动智能终端包括智能手机、笔记本、PDA智能终端、平板电脑等设备。其拥有独立的操作系统,功能的扩展性极其灵活,当前市场中的移动智能终端支持数据存贮、信息管理、多媒体、GPS、GIS、无线通信等功能,在电力行业有着广泛的研究与应用^[10-12]。

1.3 数据获取技术

传统的火电企业评价工作模式中评价标准往往通过结构化(Excel)或非结构化(PDF)形式存储,不便于评价流程步骤的电子化,不利于指标数据的量化体现。本文通过OCR和POI等相关数据获取技术彻底解决了上述问题,采用OCR技术将扫描文本或基于图像内容的文本转换为可编辑的文本;采用POI技术,提取数据基础信息进行统一管理,通过定义个性化模板,以模板为基准,对不同要素进行量化提取,使数据处理更加精细化、精准化。

2 系统设计与实现

2.1 系统架构设计

火电企业安全生产和技术监督评价系统研发遵循先进、健壮、灵活、稳定、开放、易维护、可扩展、低成本等要求,采用以J2EE为核心的多层分布式架构^[13-14],即在客户端、应用服务层、数据源层3层结构基础上,将应用服务层细分为Web服务层、界面服务层、业务逻辑层、数据访问层。客户端是通过移动智能终端或浏览器进行系统访问,实现人机交互;Web服务层是通过HTTP协议传输数据,为客户端提供对应用程序的访问,实现表现层逻辑封装;界面服务层为Web服务层提供浏览、处理等服务,实现用户界面与各业务功能的隔离;业务逻辑层实现系统业务处理;数据访问层提供对数据库的链接、读写等操作服务;数据源层提供数据存储。这种成熟的B/A/S结构不但有效缩短开发周期、降低系统维护成本、增强企业对象重复可用性,也大大提升了系统的延展性、执行效率、容错能力和负载均衡能力。多层分布式架构如图1所示。

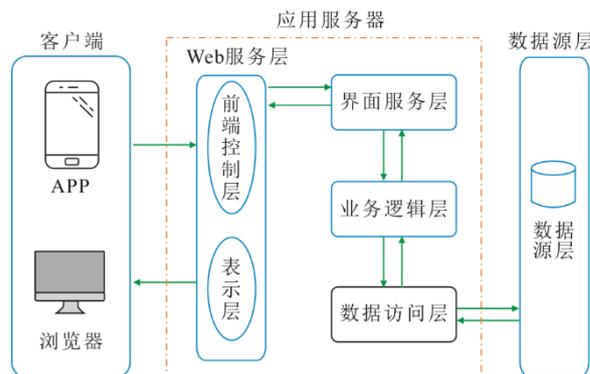


图1 多层分布式架构

Fig.1 The multi-tier distributed architecture

系统后台及服务基于Java开发,支持跨平台部署应用,移动端基于JavaScript和html5开发,Web容器采用Tomcat,数据库采用MySQL。人机交互界面简洁美观、功能命名直观通俗,便于用户自主学习。

2.2 功能设计

火电企业安全生产和技术监督系统以集团公司发布的相关标准为基础,以历次电厂自查评价及集团专家评价实践经验为指导,以业务流转及数据监控为核心,提升公司安全生产和设备健康水平。

该系统根据评价工作需要,开发了PC端和移动端,PC端包含系统全部功能;移动端辅助PC端,主要实现生产现场评价工作及业务流程审批工作。

功能设计遵循科学的 PDCA 循环管理思想^[15], 实现全方位、全过程动态监控管理以及闭环管理, 系统

功能模块划分如图 2 所示, 专家评价和问题整改流程如图 3 所示。

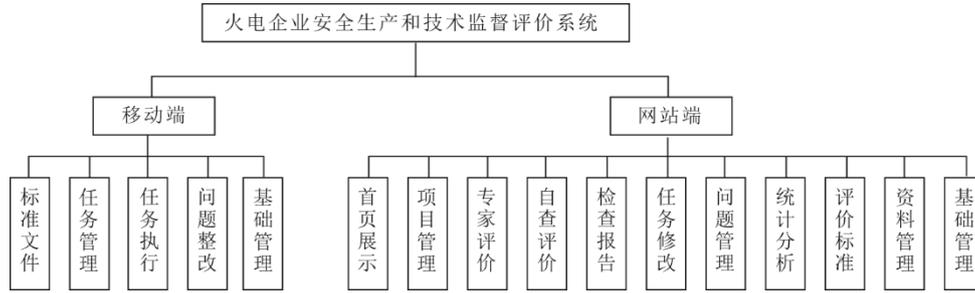


图 2 系统功能模块划分

Fig.2 The division diagram of system function modules

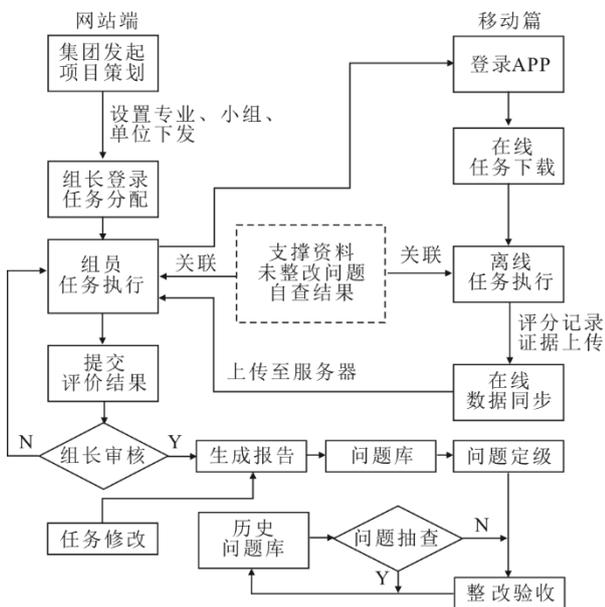


图 3 专家评价及问题整改流程

Fig.3 The process of expert evaluation and problem rectification

实时的数据管理构建了 PC 端和移动端的桥梁, 利用数据挖掘技术提供强大的数据推送、检测推荐、评价匹配功能, 面向管理层提供了实时的数据监控和管理手段, 使管理者可以及时掌握现场工作及问题整改情况。系统的主要功能模块如下。

1) 基础管理 对公司、资产、监督类别、专业、人员、设备、参考文件、职称、版本、参考标准等基础信息进行配置。

2) 评价标准 对各个专业的查评项目、查评标准、评价指标、评价方法、查证方法、查评依据、标准分等评价相关标准进行统一管理。

3) 任务执行 主要包含项目管理、项目执行、专家评价、自查评价、检查报告、任务修改。实现电厂依据标准自行寻找问题、专家开展诊断与指

导、报告自动生成等功能。同时, 支持标准以外问题的录入, 发挥评价人员专业优势、弥补标准不足, 也为后期标准升版奠定基础; 移动端支持在线数据同步、离线数据执行、语音识别、现场照片同步上传、二十五项重点要求条款模糊查询等功能^[16-18], 增强现场实用性, 提高评价效率, 具有权限的管理者可实时跟踪评价进展与质量。

4) 问题整改 依据重要程度, 对问题先定级后整改, 问题分为一、二、三级, 分别对应集团公司、二级单位、三级单位验收; 手动或自动开始相应整改验收流程。单位内部或上级单位对已验收的问题进行抽查, 若不合格, 则需重新整改。

5) 统计分析 对评价报告中的数据进行整理分析, 展现评价数据趋势图, 为监督人员决策提供数据支持。

2.3 算法与模型设计

2.3.1 推荐专家规则排序算法

专家关联关系往往表现的比较模糊, 本文通过对大量专家规则排序进行分析, 发现在准备阶段新建项目设置小组时, 会根据专家的所学专业、工作经验、职称、评价过的单位、历史发现的问题、整改建议、查评数据统计、整改方案有效率等, 推荐最优专家, 便于决策者更准确高效的决策。推荐专家规则排序算法流程如图 4 所示。

本文推荐专家规则排序算法大致包括以下几个步骤。

1) 数据预处理 首先对文本进行关系词标注, 其次对专家、任务、关系词进行识别, 存入专家属性标签和任务属性标签中。

2) 特征提取 建立特征库, 将专家和任务的显著特征提取出来, 便于后续规则匹配。

3) 关联关系 首先建立关联关系规则库,如相似经验匹配规则、工作年限匹配规则、专业匹配规则等;其次,利用关联关系规则对特征提取过程产生的显著特征进行匹配;专家显著特征包含工作年限、岗位、专业、学历、经验等,任务显著特征包含任务重要性、紧急性、难易度、进度、专业、地域等,实现专家关系的初步匹配。

4) 相似度匹配 对专家关系的初步结果放入容器中,并对该容器中的显著特征加权平均进行二次匹配,选取适合的专家。

5) 最优专家确定 通过大规模的数据抽取,得到专家之间的所有关系后,最终通过权重占比,计算得分排序,确认最优专家。

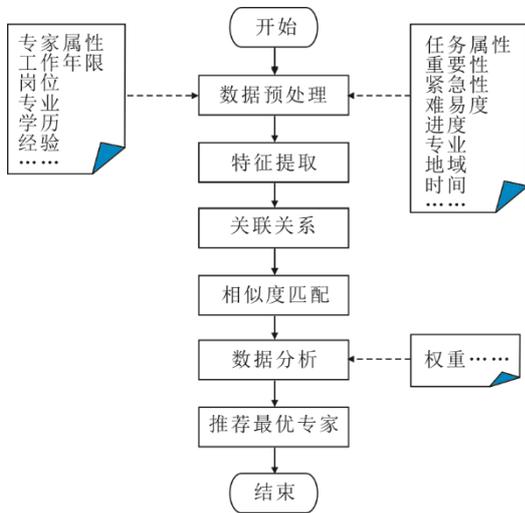


图4 推荐专家规则排序算法流程
Fig.4 The process of recommended expert rule sorting algorithm

2.3.2 基于数据挖掘技术的评价数据分析

基于数据挖掘技术的评价数据分析模型如图5所示。使用时,首先建立大数据体系,以系统积累的基础数据作为输入。如:评价类型、评价单位、监督专业、监督标准、查评得分、问题数量、问题重要程度、问题产生原因、整改情况、风险项等,经过评价数据分析模型,输出安全监督通病、关键评价单位、重点评价专业、专家服务能力评价、评价周期频次等信息。根据这些信息,在下一轮次监督评价工作中,可优先推荐重点关注的问题、重点关注的单位、重点关注的专业等信息,便于有效指导后续安全生产和技术监督评价工作。

基于数据挖掘技术的评价数据分析模型,将基础数据从单位评价水平、专业评价水平和其他相关水平3个维度进行划分。

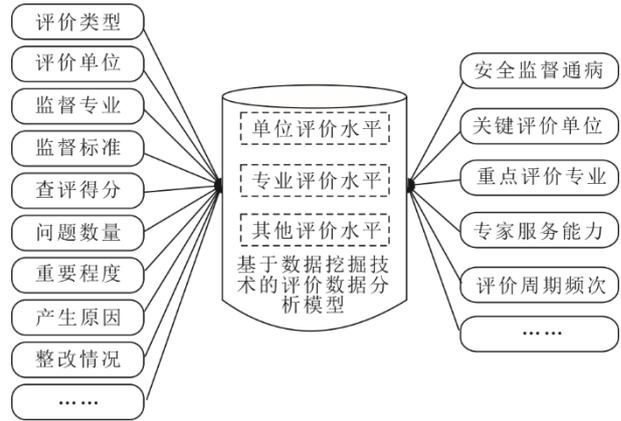


图5 基于数据挖掘技术的评价数据分析模型
Fig.5 The evaluation data analysis model based on data mining technology

1) 单位评价水平 (U)

单位评价水平指待评单位存在隐患数量的多少、整改情况好坏的程度。其参考的基础输入要素有查出问题数量、查评得分、整改情况,由查出问题数量评分、得分率、整改率等数据进行衡量。查出问题数量计算公式为

$$y = ax + b \quad (x \leq c, -1 \leq a \leq 0) \quad (1)$$

$$y = 0 \quad (x > c) \quad (2)$$

式中, x 为问题数量, y 为问题数量对应的评分, 系数 b 为评价满分标准, 系数 c 为问题数量上限。其中系数 a 的绝对值越大, 表示问题数量与得分的相关性越强。当 $x > c$ 时, 不得分。如可以设定评价标准为 100 分, 问题数量上限为 50 个, 当问题数量为 0 个时, 得分最高 100 分, 依次递减, 问题数量超过 50 个时, 不得分。

得分率定义为所选评价项得分情况与所选评价项总分之比。整改率定义为累计整改问题项数与累计查出问题项数之比。单位评价水平高低由查出问题数量评分、得分率、整改率最终确定, 计算公式为

$$U = \frac{y}{100} \times 30\% + \frac{y}{b} \times 30\% + \frac{d}{x} \times 40\% \quad (3)$$

式中 d 为累计整改问题项数。

单位评价水平 U 按照数值大小进行等级评定: $\geq 85\%$ 为 A, $75\% \sim 85\%$ 为 B, $< 75\%$ 为 C。

2) 专业评价水平 (P)

专业评价水平指评价结果好坏的程度。依据问题的重要程度, 评价结果可分为严重问题(含告警问题)、普通问题、轻微问题(及时整改)等。专业评价水平以累计查出问题数量评分为主要依据, 以严重问题数量评分和普通问题数量评分稍作权重,

共同决定。其计算公式为

$$P = \frac{t}{100} \times 50\% + \frac{g}{100} \times 30\% + \frac{n}{100} \times 20\% \quad (4)$$

式中, t 为累计查出问题数量评分, g 为严重问题数量评分, n 为普通问题数量评分。

按照数值大小进行等级评定: $\geq 85\%$ 为 A, $75\% \sim 85\%$ 为 B, $< 75\%$ 记为 C。

3) 其他相关水平 (O)

其他相关水平高低由专家对此次整体情况评价确定, 如被评价单位对本次评价准备程度、配合程度等, 属于灵活调整偏差值, 同样设置 A、B、C 3 个等级。

4) 综合评价水平 (S)

综合评价水平由单位评价水平、专业评价水平、其他相关水平三要素共同决定。设定三要素权重比为 5:3:2, 综合评价水平计算公式为

$$S = U \times 50\% + P \times 30\% + O \times 20\% \quad (5)$$

综合评价水平等级表现形式设定为 AAA, 其中, 第 1 个字母表示单位评价水平等级为 A, 第 2 个字母表示专业评价水平等级为 A, 第 3 个字母表示其他评价水平等级为 A, 以此类推。综合评价水平如图 6 所示: 横向维度(从左往右), 当单位评价水平等级相同时, 综合评价水平随着专业评价水平的降低而降低; 纵向维度(从上往下), 当专业评价水平等级相同时, 综合评价水平随着单位评价水平的降低而降低。

从大量的历史评价数据中, 可以挖掘出相关评价规则, 评价数据分析模型可依据这些属性, 对评价工作进行分类和预测, 系统智能推送安全监督通病、关键评价单位、重点评价专业、评价周期频次等, 无需投入大量精力进行人工分析, 加强了评价工作的针对性, 有效提升了评价效率, 提高了运行设备的安全性和可靠性。

综合评价水平	专业评价水平不断降低		
	AAA	ABA	ACA
单位评价水平不断降低	AAB	ABB	ACB
	AAC	ABC	ACC
	BAA	BBA	BCA
	BAB	BBB	BCB
	BAC	BBC	BCC
	CAA	CBA	CCA
	CAB	CBB	CCB
	CAC	CBC	CCC

图 6 综合评价水平

Fig.6 The comprehensive evaluation level

3 系统应用

本系统已在多家火电企业开展自查评价和专家评价, 指导生产运行, 也成功支撑了国家电力投资集团有限公司 2019 年秋季安全生产和技术监督评价工作。涉及 93 家火电企业三级单位、20 个监督专业、60 名专家成员, 报告审批由线下改为线上, 评价耗时由 6~7 d/厂缩短至 4~5 d/厂, 效率提升 20% 以上, 错误率低于 1%, 形成自查评价与专家评价报告近百份。

根据 2019 年秋季评价工作数据分析, 系统应用前后的各项指标对比见表 1。系统投入使用后报告自动生成, 评价问题自动生成问题库且实时跟踪, 加强了问题的闭环管理; 实现监督过程流程化、现场操作电子化、数据管理信息化、问题治理闭环化, 提升了工作效率, 成为安全生产和技术监督评价工作有效辅助工具, 受到使用人员的一致好评。

表 1 系统应用前后的各项指标对比
Tab.1 Comparison of indicators before and after the system application

系统应用前	系统应用后
报告格式多样、纷乱, 查询不便	报告格式统一、规范, 可以快速查询检索
评价标准执行欠缺, 缺少监督手段	规定评价内容, 促进集团相关标准落地, 规范日常工作; 可查看标准以外问题, 为标准升版奠定基础
评价过程中, 标准、支撑资料、历次未整改问题查阅或核实不便; 打印纸质版不易修改, 浪费纸张	标准文件 PC 端动态维护, 规范统一; 问题管理模块可直接查看历次未整改问题; 任务执行界面可查看该条目支撑资料, 促进无纸化办公
现场检查需先行记录及拍照, 到办公室进行整理汇总	实现任务离线执行, 不过分依赖网络; 现场直接进行问题记录及拍照上传, 可充分利用零碎时间进行问题完善, 语音输入提高评价效率
报告需手动编写及汇总, 工作量大, 且易出错	报告自动生成, 需要较少的线下调整; 预览报告生成功能体现人性化
评价进度及问题整改缺乏实时监控, 人工进行整改情况统计	“事中监控、事后分析”, 实现评价进度质量及问题整改情况实时监控, 数据分析简单, 结果展示形象直观, 支持切换与导出
线下整改流程流转较慢, 整改资料难以保存; 线上整改需问题录入, 三级单位自行整改闭环	整改资料便于保存、整改流程流转便捷、可移动端填报审批、问题实现分级管控
评价问题回溯困难、查询不便	可快速查询历次评价情况
检查耗时 6~7 d	检查耗时 4~5 d, 提升效率 20%

4 结 语

本文利用数据挖掘、移动协同、数据获取技术, 设计推荐专家规则排序算法, 构建评价数据分析模

型,开发了火电企业安全生产和技术监督评价系统。该系统有效规范了火电企业安全生产和技术监督评价工作,实现了提质增效,确保评价任务新建、任务执行、报告自动生成、问题整改验收的全过程监控及闭环管理,相关数据查询、统计分析工作更为方便,形式更为灵活、直观,能够为火电企业的安全生产和技术监督评价工作管理、企业决策提供有力数据支撑。同时,有效降低了人力物力财力,保障了机组安全、可靠、经济运行,产生了较大的直接与间接经济效益、安全效益。

【参考文献】

- [1] 钟世明,邵锐,张胜,等.基于实时预警的火电厂技术监督管理系统[J].继电器,2005,33(18):61-66.
ZHONG Shiming, SHAO Rui, ZHANG Sheng, et al. Technical supervision management system based on real-time early warning[J]. Relay, 2005, 33(18): 61-66.
- [2] 王威,胡逸铭,陶敏.电力技术监督信息管理系统的设计与开发[J].华东电力,2005,33(12):86-89.
WANG Wei, HU Yiming, TAO Min. Design and development of electric technology supervision information management system[J]. East China Electric Power, 2005, 33(12): 86-89.
- [3] 董立勉,刘岩,卢宁.基于信息内外网的电力技术监督管理系统设计与应用[J].河北电力技术,2012,31(3):21-23.
DONG Limian, LIU Yan, LU Ning. Design and application of electric power technical supervision management system on corporation intranet and extranet[J]. Hebei Electric Power, 2012, 31(3): 21-23.
- [4] 刘歆一.基于数据挖掘的电厂设备健康状态评价系统研究[D].北京:华北电力大学,2008:1.
LIU Xinyi. Research of power plant equipment health state evaluation system based on data mining[D]. Beijing: North China Electric Power University, 2008: 1.
- [5] 李延强.基于数据挖掘技术的火电厂设备预测检修研究[D].保定:华北电力大学,2006:1.
LI Yanqiang. Research on predictive maintenance of thermal power plant equipment based on data mining technology[D]. Baoding: North China Electric Power University, 2006: 1.
- [6] LNMOM W H. Building the data warehouse[M]. New York: John Wiley & Sons, 1996: 56-58.
- [7] 周党生.大数据背景下数据预处理方法研究[J].山东化工,2020,49(1):110-111.
ZHOU Dangsheng. Research on data preprocessing method under the background of big data[J]. Shandong Chemical Industry, 2020, 49(1): 110-111.
- [8] MEHTA M, AGRAWAL R, RISSANEN J. A fast scalable classifier for data mining[R]. San Jose, California: IBM Lmaden Research Center, 1995: 34-36.
- [9] 张金良,李光泉,杨忠直,等.数据挖掘在证券交易系统分析中的应用[J].地质技术经济管理,2002,24(6):31-33.
ZHANG Jinliang, LI Guangquan, YANG Zhongzhi, et al. Application of data mining on security bargain analysis system[J]. Geological Technoeconomic Management, 2002, 24(6): 31-33.
- [10] 罗睿,徐衍安,王毅,等.电厂厂级监控信息系统移动端开发及应用[J].热力发电,2018,47(6):137-142.
LUO Rui, XU Yan'an, WANG Yi, et al. Development and application of mobile terminal in plant-level supervisory information system for power plant[J]. Thermal Power Generation, 2018, 47(6): 137-142.
- [11] 薛晗光,刘聪睿,赵轩,等.智能电站工作票移动办理的设计与实现[J].热力发电,2017,46(12):40-43.
XUE Hanguang, LIU Congrui, ZHAO Xuan, et al. Design and implementation of work order mobile management: practice of intelligent power plant[J]. Thermal Power Generation, 2017, 46(12): 40-43.
- [12] 武环宇,陈国军,李勇,等.发电企业“互联网+”移动安全管理平台的开发[J].电力安全技术,2019,21(6):38-42.
WU Huanyu, CHEN Guojun, LI Yong, et al. Development of internet plus(mobile) mobile security management platform in power generation company[J]. Electric Safety Technology, 2019, 21(6): 38-42.
- [13] 陈垒.电厂化学监督信息化的研究及应用[D].保定:华北电力大学,2012:15.
CHEN Lei. Research and application of informatization in chemical supervision for power plant[D]. Baoding: North China Electric Power University, 2012: 15.
- [14] 刘正,张书锋,赵鹤鸣.MVC模式下多层分布式软件系统架构设计[J].现代电子技术,2018,41(7):135-139.
LIU Zheng, ZHANG Shufeng, ZHAO Heming. Design of multi-layer distributed software system architecture based on MVC mode[J]. Modern Electronics Technique, 2018, 41(7): 135-139.
- [15] 董中凯,田东坡.基于Web的电力集团技术监督管理系统设计与应用[J].电力信息与通信技术,2013,11(9):115-118.
DONG Zhongkai, TIAN Dongpo. Design and application of a web-based electric power group technical supervision and management system[J]. Electric Power Information and Communication Technology, 2013, 11(9): 115-118.
- [16] 朱翰超,马蕊.考虑需求侧管理的冷热电联供微电网优化配置方法[J].电力系统保护与控制,2019,47(2):139-146.
ZHU Hanchao, MA Rui. Optimal configuration method of CCHP microgrid considering demand side management [J]. Power System Protection and Control, 2019, 47(2): 139-146.
- [17] 邹军,谭华,敖光辉,等.火电厂技术监督管理系统研发[J].热力发电,2019,48(9):90-93.
ZOU Jun, TAN Hua, AO Guanghui, et al. Research and development of technical supervision and management system for thermal power plant[J]. Thermal Power Generation, 2019, 48(9): 90-93.
- [18] 郑国太,李昊,赵宝国,等.基于供需能量平衡的用户侧综合能源系统电/热储能设备综合优化配置[J].电力系统保护与控制,2018,46(16):8-18.
ZHENG Guotai, LI Hao, ZHAO Baoguo, et al. Comprehensive optimization of electrical/thermal energy storage equipment for integrated energy system near user side based on energy supply and demand balance[J]. Power System Protection and Control, 2018, 46(16): 8-18.

(责任编辑 杜亚勤)