

DOI: 10. 20040/j. cnki. 1000-7709. 2023. 20220566

基于 Python 数据可视化的水电集控平台主报警信息规则时序匹配与处置模型构建

沙永兵¹, 谌斐鸣¹, 曹德勤¹, 汪 涛¹, 吴 辉¹, 贺 雄²

(1. 五凌电力有限公司, 湖南长沙 410000; 2. 湖南五凌电力科技有限公司, 湖南长沙 410004)

摘要:为实现水电集控平台主报警信息时序匹配和处置,构建了基于 Python 数据可视化的水电集控平台主报警信息规则时序匹配与处置模型,该模型首先利用通信服务器将水电站设备运行故障信息传输至水电集控平台接收端内后生成故障报警序列;然后依据领域专家匹配规则库内的时序匹配规则匹配水电站故障报警主信息后使用状态函数、状态变化函数等对其进行时序特性分析,利用映射表描述主报警信息时序特性;最后对该主报警信息时序特性展开报警选择、报警处理和处理结果分析后,使用基于 Python 编程软件对结果进行可视化并发送至调度员控制台内,实现水电集控平台主报警信息规则时序匹配与处置。试验结果表明,该模型具备较好的主报警信息可视化能力,其爬取的主报警信息覆盖率高达 98.99%;主报警信息规则时序匹配时的皮尔逊相关系数数值较高,处置主报警信息能力较好。

关键词: Python; 数据可视化; 水电集控平台; 信息规则; 时序匹配; 处置模型

中图分类号: [TV-9]; TP391

文献标志码: A

文章编号: 1000-7709(2023)05-0182-05

1 引言

水电集控平台为水电站运维管理的主要途径,其可为用户呈现当前水电站设备运行状态,使用户及时掌握设备故障信息便于调度运维人员执行作业任务^[1]。为使水电集控平台更好地服务用户,对其主报警信息进行匹配和处置尤为必要。李俊杰等^[2]利用关联分析算法将报警时序进行区域划分,以某时间节点为参考对报警关联性进行评价;甘昕艳等^[3]利用卷积神经网络模型对不同时序的报警信息进行迭代训练,得到降维后的报警时序信息,然后利用模糊聚类算法挖掘报警时序特征从而实现其规则匹配。但这两种方法均未对报警时序信息进行管理处置,在应用上存在片面性。Python 为一种解释型的程序设计语言,其具备简单便捷的程序编写操作界面和较为丰富的资源开发库^[4,5],可将 C++、Java 等语言互相互连接生成大型的程序原型。为此,本文构建了基于 Python 数据可视化的水电集控平台主报警信息规则时序匹配与处置模型,并应用 Python 编程软件,将主报警信息可视化,关键时刻辅助、指导

运行值班人员进行应急处置,帮助值班人员更快、更准确地掌握现场工况和事件发展态势。

2 水电集控平台主报警信息规则时序匹配与处置模型构建

2.1 水电集控平台主报警信息处置模型构建

构建水电集控平台主报警信息处置模型,其框架见图 1。

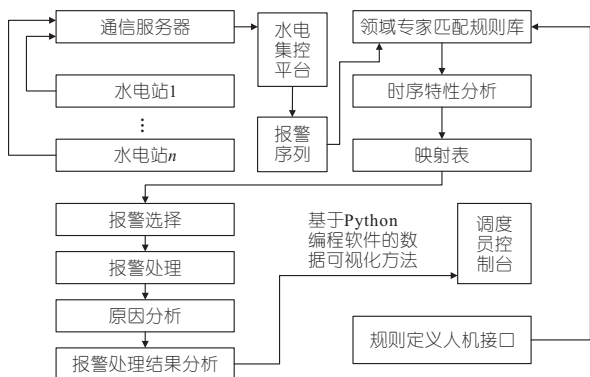


图 1 水电集控平台主报警信息处置模型

Fig. 1 Disposal model of main alarm information of hydropower centralized control platform

收稿日期: 2022-03-23, 修回日期: 2022-08-12

基金项目: 五凌电力远程集控应急处置智能管理平台研发(C99007620G201EC139)

作者简介: 沙永兵(1971-),男,高级工程师,研究方向为水电厂发电运行,E-mail:yousi052082@163.com

通讯作者: 谌斐鸣(1981-),男,高级工程师,研究方向为水电厂综合自动化,E-mail:sha_syb@wldl.com.cn

在电集控平台主报警信息处置模型内,若干个水电站设备运行故障信息通过通信服务器传输至水电集控平台接收端内,并生成故障报警序列后,通过规则定义人机接口将领域专家匹配规则导入水电集控平台主报警信息处置模型内。依据领域专家匹配规则库对水电站故障报警主信息进行匹配后,使用状态函数、状态变化函数对其进行时序特性分析。将时序特性分析结果以映射表的形式描述后,通过报警选择、报警处理、原因分析等处理并对报警处理结果进行分析。使用基于 Python 编程软件的数据可视化方法对报警处理分析结果进行可视化处理,然后将可视化处理结果传输到调度员控制台内,实现水电集控平台主报警信息处置。

2.2 主报警信息规则时序匹配分析方法

2.2.1 领域专家匹配规则

当水电集控平台内出现报警信息时,说明当前水电站某个区域运行出现故障^[6],在此可通过领域专家的匹配规则描述报警信息和各种原因之间的关系。令 F 表示水电集控平台可疑故障数据集, c_i 表示可能发生的原因事件, u_i 表示实际的报警信息,则领域专家的匹配规则描述实际报警信息和可能发生的原因事件之间关系形式见图 2。

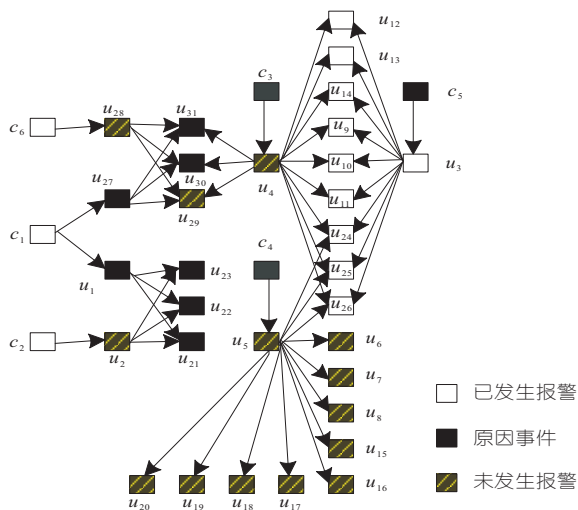


图 2 实际发生报警信息和原因事件关系示意图

Fig. 2 Schematic diagram of the relationship between the actual alarm information and the cause and event

在水电集控平台可疑故障数据集 F 内,依据图 1 的实际发生报警信息和原因事件关系可知,通过正向推理可获取到当前水电站发生不同类型故障的原因事件,并将该原因事件标记为 $c_1 \sim c_6$ 。从可疑故障设备角度出发,以反向推理形式得到发生故障报警信息的线路,即 $u_1 \sim u_{27}$ 。在不考虑报警时序特征时,可利用报警信息和实际发生报警与其原因事件之间的关系分析出水电集控平

台的主报警信息。但由于水电站发生故障报警存在极强的时序特性^[7,8],因此从其时序特性出发考虑水电集控平台主报警信息的匹配规则。

当水电站发生故障报警时,报警信息是依据其故障发生时间向用户发出的,水电集控平台的报警信息表现为离散状态^[9],因此时序特性是呈现报警信息的主要途径。依据水电集控平台报警信息的时序特性,对其进行一致性识别,可有效提升水电集控平台主报警信息的匹配和处置效果^[10]。令 u 表示主报警信息, k 表示时刻, $u \langle k_0 \rangle$ 表示主报警信息 u 在时刻为 k_0 时发生的时序, $u[k_0, k_1]$ 表示主报警信息 u 在时间段 $[k_0, k_1]$ 内的时序,使用时区的方式表达水电集控平台主报警信息的时间特性,以保护水电站基础设备配置为最基础原则,在该设备的维护时间上保留选择性并结合水电站不同设备维护的时间极差。同时以水电站设备故障时刻作为参考,将故障时刻划分不同阶段,如主维护动作时区、失灵保护动作时区及后续维护动作时区。在考虑水电集控平台主报警信息的时序特性时,依据原因事件的时间极差,在水电站设备故障的不同时区内对其展开正向推理和反向推理,获得带有时间标志的水电站设备故障候选原因集合。当获得水电站设备故障候选原因集合后^[11],依据国际上对报警的正确、漏报和误报定义分析水电站故障主报警信息匹配结果是否符合。其中,当水电站运行状态为异常时所发出的报警信息为正确信息,可由 N_{Hit} 表示。当水电站运行状态为异常时,水电集控平台并未接收到报警信息,该情况为漏报,则由 N_{Miss} 表示。当水电站运行正常,而水电集控平台却接收到报警信息时,该情况为错报,由 N_{FA} 表示。

2.2.2 报警信息时序特性分析方法

依据领域专家匹配规则,分析水电集控平台主报警信息的时序特性。报警信息的时序特性可由状态函数和状态变化函数描述。其中状态函数又由保护状态函数和断路器状态函数组成。令 $r_i(k)$ 为保护状态函数,其表达式为:

$$r_i(k) = \begin{cases} 0 & k < k_{act,ri} \\ 1 & k \geq k_{act,ri} \end{cases} \quad (1)$$

式中, $k_{act,ri}$ 为第 i 个水电站设备保护动作时刻,当 $r_i(k)$ 数值为 0 或 1 时,表明第 i 个保护在时刻为 t 时的状态为无动作或已经发生动作。

令 $c_j(k)$ 为断路器状态函数,其表达式为:

$$c_j(k) = \begin{cases} 1 & k < k_{act,cj} \\ 0 & k \geq k_{act,cj} \end{cases} \quad (2)$$

式中, $k_{act,cj}$ 为第 j 个断路器动作时刻,当 $c_j(k)$

数值为 1 或 0 时,说明第 j 个断路器在时刻为 t 时的状态为闭合或断开。

水电集控平台的主报警信息时序特性的状态变化函数由保护状态变化函数和断路器状态变化函数构成。令 $r_i(k_s, k_e)$ 为保护状态变化函数,其表达式为:

$$r_i(k_s, k_e) = \begin{cases} 0 & k_{act,ri} \notin [k_s, k_e] \\ 1 & k_{act,ri} \in [k_s, k_e] \end{cases} \quad (3)$$

式中, $r_i(k_s, k_e) = 0$ 或 1 时为时间段 $[k_s, k_e]$ 内的第 i 个保护发生动作或未发生动作。

令 $c_j(k_s, k_e)$ 表示断路器状态变化函数,其表达式为:

$$c_j(k_s, k_e) = \begin{cases} 0 & k_{act,cj} \notin [k_s, k_e] \\ 1 & k_{act,cj} \in [k_s, k_e] \end{cases} \quad (4)$$

式中,当 $c_j(k_s, k_e) = 0$ 或 1 时为时间段为 $[k_s, k_e]$ 内的第 j 个断路器状态为发生断开或未发生断开情况。

令 γ_{max} 为水电站设备保护内最大保护动作时间的整定值,其表达式为:

$$\tau_{max} = \max\{\gamma_{ri} \mid 1 \leq i \leq n_r\} \quad (5)$$

式中, γ_{ri} 为第 i 个保护动作时间; n_r 为总保护动作时间。

当在时刻为 k_1 时,水电集控平台收到主报警信息,该信息由 m_1 表示,则主报警信息时序的参考点为 k_0 ,其表达式为:

$$k_0 = k_1 - \gamma_{max} \quad (6)$$

当时间为 k_{N-1} 时,水电集控平台收到主报警信息 m_{N-1} 后,在该报警信息持续时间为 ΔK 内,并未收到其他报警信息,则说明此时水电站设备故障已被排除。

2.3 基于 Python 编程软件的数据可视化方法

2.3.1 信息爬取

在对水电集控平台主报警信息进行数据可视化前,需爬取水电集控平台内的报警实时信息。在此利用 Python 编程软件内的爬虫工具包对水电集控平台内的报警信息进行爬取,即:在了解当前计算机操作系统的前提下,下载适合该操作环境的 Python 版本并安装后,启动 Python 软件打开命令提示符窗口,然后安装 requests 开源 HTTP 库^[12],将该 HTTP 库内的 PiP install requests 添加到 Python 软件命令提示符窗口内,分析水电集控平台网页源代码结构后,利用网址抓取命令抓取并分离网页内容。将抓取分离后的水电集控平台主报警信息存储到 Python 数据库内,将上述过程反复操作后,即实现水电集控平台主报警信息的爬取。

2.3.2 水电集控平台主报警信息提取

获得水电集控平台报警信息后,使用 requests 库提取 Python 数据库内的水电集控平台报警信息。并利用 BeautifulSoup 解析器解析水电集控平台报警信息,获得主报警信息内每个数据对应的元素和名称,再利用 find、find_all 等函数对主报警信息内的数据进行标签定位和提取。利用正则表达式描述定位提取后的水电集控平台主报警信息,并逐个爬取主报警信息的字段名称和提取规则。

依据水电集控平台主报警信息的提取规则爬取得到主报警信息的字段名后,此时的水电集控平台主报警信息较为混乱,需对其进行数据整理^[13,14]。利用 Python 软件内的 pandas 库内的 mean、loc、merge 等模块对水电集控平台主报警信息进行清洗、统计处理后^[15],并使用 DataFrame 表格形式呈现处理后的水电集控平台主报警信息。

2.3.3 数据可视化输出

对水电集控平台主报警信息进行处理后,使用 Python 编程软件内的 matplotlib 库内的散点图、饼状图、柱状图等形式对数据进行表述并输出,即实现水电集控平台主报警信息的可视化过程。

3 试验分析

以某大型水电站的水电集控平台为例,使用 Matlab 仿真软件搭建仿真环境,应用本文模型对该水电集控平台的主报警信息进行规则匹配和处置。其中 Python 编程软件版本为 V3.9.0,其运行环境为 windows10 旗舰版操作系统,浏览器 IE6.0 版本。

3.1 信息爬取测试

在仿真软件内设置不同水电站设备运行故障,且每种故障次数不同,应用本文模型对该故障的主报警信息进行爬取,分析本文模型的信息爬取能力,结果见表 1。

由表 1 可知,在不同类型的主报警信息和故障信息条数情况下,本文模型爬取的主报警信息数量和正确爬取数量与实际主报警信息条数相差不大。当实际主报警信息条数低于 300 条时,本文模型爬取的主报警信息数量与实际主报警信息条数完全相同。当实际主报警信息条数超过 300 条后,本文模型爬取的主报警信息条数与实际主报警信息出现轻微偏差,但最大偏差数值仅为 4 条,其爬取的主报警信息覆盖率高达 98.99%。

表 1 信息爬取测试结果

Tab. 1 Information crawling test results

主报警信息类型	实际主报警信息/条	爬取数量/条	正确爬取/条
主控单片机故障	282	281	280
励磁开关跳闸	306	305	304
绝缘故障	117	117	117
单相弧光接地故障	32	32	32
调速器拒动	399	395	391
流程中断	138	138	138
流信号器异常	156	156	156
水压超标	216	216	214
介质损耗异常	212	212	211
发热功率超标	87	87	87
局部过热	411	403	400

从正确爬取数量角度分析,当实际主报警信息条数低于 200 条时,本文正确爬取数量和实际主报警信息完全相同。当实际主报警信息为 411 条时,本文模型正确爬取数量为 400 条,其爬取的精度高达 97.3%。

3.2 数据可视化输出

以水电集控平台主报警信息的事件总数、待处理数量等为试验对象,绘制数据可视化的饼状图和位置分布图,结果见图 3。由图 3 可知,本文模型可较为清楚地呈现当前水电站集控镜头主报警信息详情,对水电集控平台主报警信息可视化效果较好。

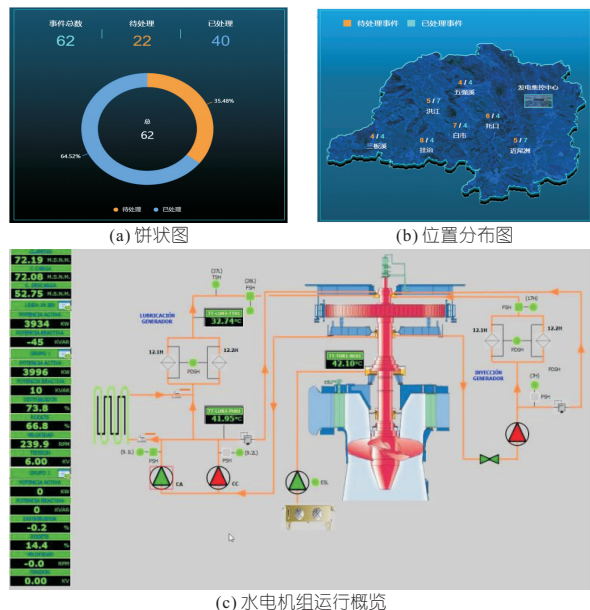


图 3 报警信息可视化测试结果

Fig. 3 Alarm information visualization test results

3.3 处置结果可靠性分析

以判定系数 R^2 作为衡量本文模型对水电集控平台主报警信息处置的可靠性。在 Matlab 仿真软件内模拟不同数量的主报警信息,使用本文模型对该主报警信息展开处置并绘制本文模型的

判定系数曲线,结果见图 4。由图 4 可知,本文模型的判定系数 R^2 数值随主报警信息数量的增加呈激增情况,当激增到一定程度后保持不变的状态。在主报警信息数量为 100 条之前时,本文模型的判定系数 R^2 数值涨幅较大,在主报警信息数量约为 100~400 条之间时,判定系数 R^2 数值虽仍呈上涨趋势,但涨幅明显降低。当主报警信息数量超过 400 条后,本文模型的判定系数 R^2 数值始终保持为 1。该数值表明本文模型处置水电集控平台主报警信息时的可靠性极高,其具备良好的实际应用性。

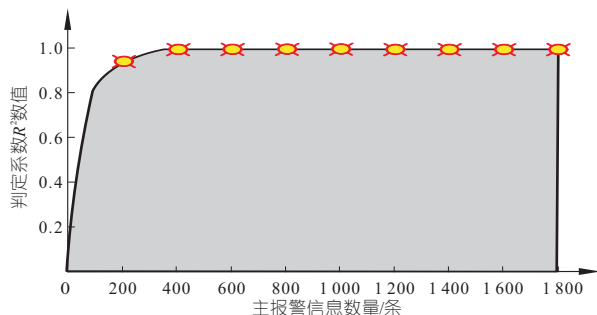


图 4 处置可靠性结果

Fig. 4 Disposal reliability results

3.4 主报警信息时序匹配测试

皮尔逊相关系数是衡量两个变量之间相关程度的指标,利用皮尔逊相关系数描述本文模型对水电主控平台的主报警信息时序匹配能力,结果见图 5。由图 5 可知,本文模型在对主报警信息进行时序匹配时的正向推理和反向推理的皮尔逊相关系数在主报警信息数量约为 6 000 条之前时,其数值分别约为 1、0.99。当主报警信息超过 6 000 条后,本文模型下的正向推理和反向推理的皮尔逊相关系数数值呈下降趋势,但降幅不大。综上可知,本文模型对水电集控平台主报警信息时序匹配效果较好。

3.5 主报警信息处置测试

给出本文模型对水电集控平台的主报警信息处置页面,分析其对主报警信息的处置能力,结果见图 6。由图 6 可知,在本文模型的主报警信息处置界面内可设置事件库、经验规则库、设备管

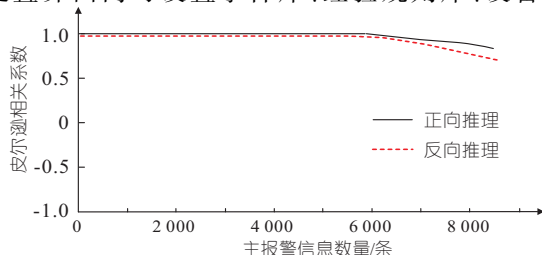


图 5 主报警信息时序匹配测试结果

Fig. 5 The main alarm information timing matching test results

理、用户管理、导入及修改和删除等功能。可见本文模型处置水电集控平台主报警信息功能较为全面,具备较强的普适性。

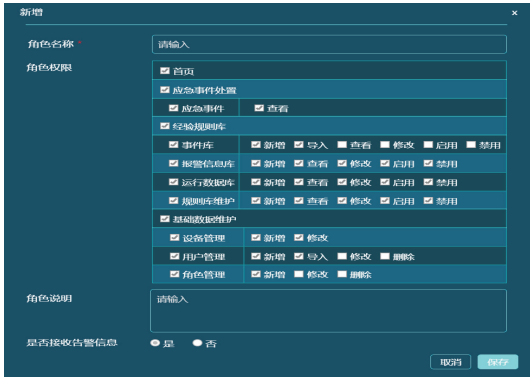


图 6 主报警信息处置测试结果

Fig. 6 The main alarm information disposal test results

4 结论

本文构建了水电集控平台主报警信息规则时序匹配与处置模型,在该模型中使用 Python 可编软件实现主报警信息的可视化。通过对本文模型进行验证可知,该模型具有良好的可靠性和主报警信息时序匹配效果,且涵盖主报警信息处置范围较广,具备较好的应用性和普适性。

参考文献:

[1] 汪悦颀,焦在滨. 基于继电保护同步时序信息特征的配电网故障诊断方法[J]. 南方电网技术,2019,13(4):73-79.
 [2] 李俊杰,杨博,李宏光. 基于报警时序信息挖掘的报警关联分析方法[J]. 信息与控制,2019,48(1):29-34.
 [3] 甘昕艳,唐晓年. 基于 CNN 的时序数据关联规则挖掘模型[J]. 计算机仿真,2021,38(3):282-285,326.

[4] 许奕杰,王嵘,万永菁,等. 基于 AE-LSTM 网络模型的机场周界入侵报警及分类算法[J]. 华东理工大学学报(自然科学版),2021,47(3):323-330.
 [5] LI ZHIGUO, ZHONG JIANG, ZHONG LUMAN. Study on processing technology for complex event management based on multivariate time series data [J]. Computer science, 2019,46(6):55-63.
 [6] 张泽生,李宏光,杨博,等. 基于置信规则库推理的过程报警预测方法[J]. 控制工程,2019,26(4):765-772.
 [7] 陈丛,周力臻. 基于 Python 爬虫技术的虚假数据溯源与过滤[J]. 计算机仿真,2021,38(3):346-350.
 [8] 徐传福,王曦,刘舒,等. 基于 Python 的大规模高性能 LBM 多相流模拟[J]. 计算机科学,2020,47(1):17-23.
 [9] 王五妹. DCS 信息监视与控制系统报警卡转换工具研制[J]. 核科学与工程,2021,41(3):644-648.
 [10] 刘天鹏,牛锋,张什永,等. 基于时序匹配及变长滑窗的雷达信号分选[J]. 雷达科学与技术,2020,18(5):562-567,578.
 [11] 于浩,贾清泉,李珍国,等. 基于时间序列模式匹配的电能质量区域化治理[J]. 中国电机工程学报,2019,39(13):3788-3799.
 [12] 严智. 基于关联规则的网络信息节点风险评估系统设计[J]. 电子设计工程,2021,29(13):124-128.
 [13] WANG XUECHUN, HUANG SHAOPING, XU AIDONG, et al. A similarity matching based data privacy-protection scheme for edge terminal in smart grid[J]. Southern power system technology, 2020,14(1):80-85.
 [14] 滕志军,郭力文,吕金玲,等. 基于时序信息分析的 WSN 贝叶斯信誉评价模型[J]. 郑州大学学报(工学版),2019,40(1):38-43.
 [15] 顾佩月,刘峥,李云,等. 基于时滞特征的时序依赖情节发现[J]. 计算机应用,2019,39(2):421-428.

Construction of Time Sequence Matching and Disposal Model of Main Alarm Information Rules of Hydropower Centralized Control Platform Based on Python Data Visualization

SHA Yong-bing¹, CHEN Fei-ming¹, CAO De-qin¹, WANG Tao¹, WU Hui¹, HE Xiong²

(1. Wuling Power Corporation Ltd., Changsha 410000, China;

2. Hunan Wuling Power Technology Co., Ltd., Changsha 410004, China)

Abstract: In order to realize the timing matching and disposing of the main alarm information of the hydropower centralized control platform, a timing matching and disposing model was constructed based on Python data visualization. The model transmitted the fault information of hydropower station equipment to the receiver of hydropower centralized control platform by using communication server, and the fault alarm sequence was generated. After matching the main fault alarm information according to the time sequence matching rules in the domain expert matching rule base, state function and state change function were used to analyze the time sequence characteristics. Then the mapping table was used to describe the timing characteristics of the main alarm information. After analyzing the timing characteristics of the main alarm information, the timing matching and disposal of the main alarm information in the hydropower centralized control platform were realized. Experimental results show that the model has good visualization ability of main alarm information, and the coverage rate of main alarm information crawled by the model is up to 98.99%. The model has higher Pearson correlation coefficient and better ability to deal with the main alarm information.

Key words: Python; data visualization; hydropower centralized control platform; information rules; timing matching; disposal model