DOI:10.20033/j.1003-7241.(2025)08-0033-05

基于多源时控信息与风险因子的梯级水电远程集控预案自动化匹配算法

徐 伟[1]，张 楠[2]

（1.贵州乌江水电开发有限责任公司水电站远程集控中心，贵州 贵阳 550002；
2.广东益泰达科技发展有限公司，广东 广州 530001）

摘 要：常规的梯级水电远程集控预案自动化匹配算法由于构建的梯级水电远程集控预案库的构建不全面，导致算法的匹配效果不佳。为此，设计基于多源时控信息与风险因子的梯级水电远程集控预案自动化匹配算法。通过数据采集器获取梯级输电系统的多源时控信息，清洗和标准化处理数据以识别风险因子，并计算风险指数。基于风险类型分类构建预案库，并利用计算提取系统特征进行自动化匹配。实验结果表明，所提算法在实际应用中的调整兰德指数（adjusted rand index，ARI）值更接近于1，匹配效果更好。

关键词：多源时控信息；风险因子；梯级水电；远程集控；自动化匹配；算法设计
中图分类号：TP273 **文献标识码：**A **文章编号：**1003-7241（2025）08-0033-05

Automatic Matching Algorithm for Remote Control Plan of Cascade Hydropower Based on Multi-source Temporal Control Information and Risk Factors

XU Wei[1], ZHANG Nan[2]

(1. Hydropower Station Remote Centralized Control Center of Guizhou Wujiang Hydropower Development Co., Ltd., Guiyang 550002, China;
2. Guangdong Yitaida Technology Development Co., Ltd., Guangzhou 530001, China)

Abstract: The conventional automatic matching algorithm for remote control plans of cascade hydropower has poor matching performance due to the incomplete construction of the cascade hydropower remote control plan library. Therefore, an automated matching algorithm for remote control plans of cascade hydropower based on multi-source temporal control information and risk factors is designed. It obtains multi-source temporal control information of cascade transmission systems through data collectors, cleans and standardizes data to identify risk factors, and calculates risk indices. It also builds a contingency plan library based on risk type classification, and uses computational extraction system features for automated matching. The experimental results show that the adjusted rand index (ARI) value of the new algorithm is closer to 1 in practical applications, and the matching effect is better.

Keywords: multi source time control information; risk factors; cascade hydropower; remote centralized control; automatic matching; algorithm design

0 引言

梯级水电远程集控作为现代水电能源管理的一种先进模式，能够通过集中监控和智能调度，实现对梯级水电站群的优化运行和协同控制。然而，在实际运行过程中，梯级水电站面临着复杂的运行环境、多变的水文气象条件以及严格的安全运行标准等多重挑战[1]。因此，开发一梯级水电远程集控预案自动化匹配算法，对于提升梯级水电站的运行安全、经济效益以及智能化管理水平具有重要意义[2]。

近年来，已经有很多学者针对这种情况展开了研究，并阐述了自己的观点。文献[3]方法将收集的电力调度系统历史故障数据进行标准化处理，再从中提取出与关键视觉字符相关的特征，由此构建故障模式库，针对不同的故障模式，制定对应的预案，并生成相应的预案匹配算法，实现故障预案的自动化匹配。该方法能够实时获取电力调度系统的状态信息，从而快速识别故障并匹配预案，大大提高了故障处理的效率，但构建的故障模式库导致该方法只能对已有的故障进行匹配，方法的灵活性较低。文献[4]方法首先明确多个业务的特征，并分析网络切片的配置和资源分配情况，筛选出其中的关键特征，利用特征匹配算法，对其进行匹配并对匹配结果进行优

*基金项目：贵州省重点科技项目（6017D202300203）
收稿日期：2024-04-30

化。该方法可以灵活地匹配多维特征业务与网络切片，满足不同业务的不同需求，但其设计的匹配算法过于复杂，导致方法的实时性不高。文献[5]方法先对本体库进行初始化，再从中筛选出属性类型等多种特征，在周期学习理论的作用下，构建本体匹配模型，利用交互式草蜢优化算法对其匹配模型的求解结果进行优化，由此得到最优的匹配结果。该方法的匹配效率较高，但考虑本体匹配问题多样，而该方法难以针对实际情况进行参数调整，导致方法的适应性不强。文献[6]方法先对网络节点进行分析，再定义相应的匹配规则，由此构建对应的匹配模型，并利用优化算法对其进行优化，提高匹配结果准确性。该方法能够根据实际需求调整匹配规则，适应性较强，但其对匹配规则和匹配模型构建要求较高，如果构建不合理，容易导致匹配结果不稳定。

为了解决上述方法中存在的问题，本文设计了基于多源时控信息与风险因子的梯级水电远程集控预案自动化匹配算法。在多源时控信息与风险因子的作用下，获取梯级水电站的多角度数据信息，并对获取的数据信息进行量化处理，由此构建对应的梯级水电远程集控预案库，并设计相应的匹配算法。通过引入多源时控信息和风险因子，提高了预案匹配的准确性和实时性，也填补了现有技术的空白[7]。同时，解决了传统预案匹配方法存在的问题，提升了梯级水电站的安全运行水平和智能化管理水平，为水电能源的可持续发展贡献力量。

1 梯级水电远程集控预案自动化匹配算法设计

1.1 基于多源时控信息与风险因子的梯级水电数据获取及量化处理

为实现对梯级水电远程集控预案自动化匹配算法的设计，需要先明确数据需求，结合梯级水电站的运行情况，获取多源时控信息。多源时控信息包含梯级水电系统在运行过程中，不同来源和时间点收集到的与水电运行相关的各种数据，其数据类型多样，包括水电站内部的实时运行数据、环境监测数据等[8]。多源时控信息具有多样性和时间敏感性等特点，导致在获取数据中存在部分重复数据和错误数据，数据的质量较差。为此，对其进行数据清洗。其数据清洗的具体过程如图1所示。

数据清洗的过程中，考虑数据的特性，先去除数据中的重复数据，减小后续设计中的计算量。对于数据中的缺失值，如果缺失值较多，可以直接将缺失值所在数据集去除，如果缺失值较少，可以通过均值、插值等多种方法对数据进行填充，保证数据的完整性[9]。对于数据中的异常值，采用四分位距法对其进行判断，根据判断结果，去除其中的异常值。对于数据中的错误数据，先判断数据的有效性，再对其进行替换和删除。在上述基础上，对处理过的数据进行标准化处理，统一数据的格式[10]。其标准化处理的具体过程如下所示：

$$y' = \frac{y - y_{\min}}{y_{\max} - y_{\min}} \tag{1}$$

式中，y'表示经过标准化处理的梯级水电系统多源时控信息，y表示原始的梯级水电系统多源时控信息，$y_{\min}$表示多源时控信息的最小值，$y_{\max}$表示多源时控信息的最大值。

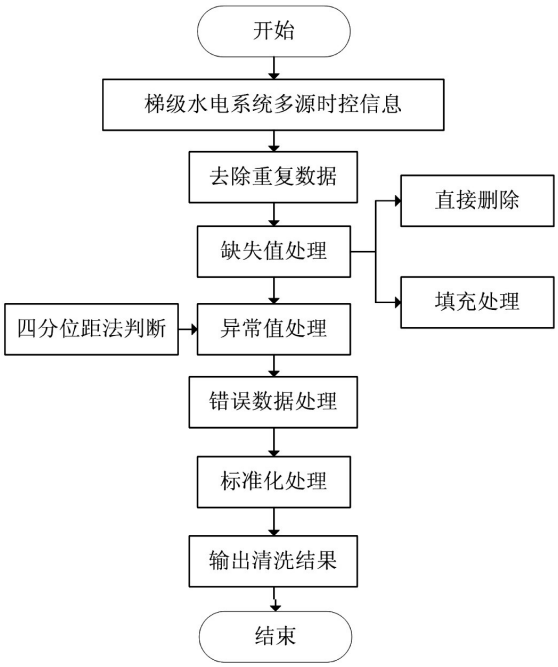


图1 数据清洗的具体过程

利用上述标准化处理，完成对多源时控信息的数据清洗，提高多源时控信息的质量，以此为基础，为后续识别数据中的风险因子，并对风险因子进行量化处理的设计提供数据支持。

根据上述处理过的梯级水电系统多源时控信息，从中筛选出对梯级水电系统远程集控影响较大的风险因子。这些风险因子包含多种因素，如环境因素、设备状态、操作情况等。在本文的设计中，通过计算不同风险因子的风险指数，对梯级水电系统中存在的风险进行量化处理。其具体量化过程如下所示：

$$\begin{cases} E_l = y' \cdot \sum_{i=1}^{n} P_i \cdot L_i \\ P_i = y' \cdot \dfrac{K_i}{K_t} \end{cases} \tag{2}$$

式中，E_l表示风险因子的风险指数，P_i表示风险因子出现的概率，L_i表示该风险因子出现时，梯级水电系统的损失值，K_i表示该风险因子在梯级水电系统运行过程中出现的次数，K_t表示所有风险因子出现的次数。

利用上述公式，对风险因子进行量化处理，根据风险

指数的大小，判断当前风险因子对梯级水电系统远程集控的影响，为后续构建梯级水电系统远程集控预案库奠定基础[11]。至此，基于多源时控信息与风险因子的梯级水电数据获取及量化处理的设计完成。

1.2 梯级水电远程集控预案库构建

将上述获取的多源时控信息和识别的风险因子作为基础，构建梯级水电远程集控预案库。在构建预案库的过程中，先确定梯级水电集控的核心目标，如安全金控和故障预警，确保构建的预案库能够为梯级水电远程集控系统的运行提供支持。在确定目标后，结合上述确定的风险因子，针对不同的风险类型，设计相应的应急预案，并计算设计预案的可行性和可操作性，保证其自身的切实可行[12]。本文设计的梯级水电远程集控预案部分如表1所示。

表1 梯级水电远程集控预案库部分内容

预案分类	具体内容	预案内容
设备故障预案	发电机组故障 变压器故障 开关设备故障	先手动隔离故障部分，再启用备用设备，并对故障设备进行及时修复
自然灾害预案	洪水 地震 雷电	先关闭较为关键的阀门和开关，保证设备安全，并对损坏程度进行评估
人为失误预案	操作错误 操作疏忽	立即采取相应的补救措施，并分析出现错误的原因
网络安全预案	黑客攻击 病毒入侵	加强安全防护措施，并启动应急响应机制，对受影响的系统进行隔离和修复，避免漏洞进一步扩大

确定相应的预案内容后，需要根据实际情况，对预案内容进行分析和及时更新，保证预案的切实有效性，为后续设计梯级水电远程集控预案自动化匹配算法奠定基础[13]。至此，完成梯级水电远程集控预案库的构建。

1.3 梯级水电远程集控预案自动化匹配算法设计

将上述构建的梯级水电远程集控预案库作为基础，设计梯级输电远程集控预案自动化匹配算法。在设计匹配算法的过程中，根据上述风险因子量化处理的结果，提取出对应的风险特征，及构建的预案库，提取出对应的预案特征[14]。两种特征的提取具体过程如下所示：

$$\begin{cases} T_f = \mathrm{sqrt}\left(\sum_{i=1}^{n}\dfrac{\left(P_i-\overline{P}\right)^2}{w_i}\right) \\ T_y = \sum_{j=1}^{m}\dfrac{t_j}{k_h} \end{cases} \tag{3}$$

式中，T_f表示提取的梯级水电系统远程集控的风险特征，$\overline{P}$表示梯级水电系统远程集控系统风险因子的均值，w_i表示不同风险因子的权重值，n表示风险因子的数量，T_y表示梯级水电系统远程集控的预案特征，t_j表示预案措施的类别，k_h表示预案措施的使用频率，m表示梯级水电远程集控预案的数量。

利用上述公式，分别提取出对应的特征。根据提取结果，计算不同特征之间的相似度，由此实现对预案的自动化匹配。在上述过程中，集控预案自动化匹配的具体过程如下所示：

$$M\left(T_f,T_y\right)=\frac{A\left(T_f\cap T_y\right)}{A\left(T_f\right)+A\left(T_y\right)-A\left(T_f\cap T_y\right)} \tag{4}$$

式中，$M(T_f,T_y)$表示梯级水电远程集控源自动化匹配结果，$A(T_f\cap T_y)$表示两种特征相似度出现的概率，$A(T_f)$表示梯级水电系统运行过程中出现的风险概率，$A(T_y)$表示集控预案库中不同预案出现的概率。

利用上述公式，计算相应的匹配结果，并将计算结果与设定的阈值进行对比，若超过设定的阈值，则匹配成功，反之，则需要重新进行匹配。在上述基础上，再对匹配结果进行优化，进一步提高匹配结果的可信度[15]。在上述过程中，对匹配结果的优化过程如下所示：

$$Q=H\left(M\left(T_f,T_y\right)\right)\cdot\gamma+b \tag{5}$$

式中，Q表示匹配结果的蚁群优化结果，$H(\cdot)$表示对匹配结果的蚁群优化函数，γ表示对应的优化参数，b表示随机常数项。

利用上述公式，进一步提高上述匹配结果的可靠性，保证预案匹配的准确性。至此，基于多源时控信息与风险因子的梯级水电远程集控预案的自动化匹配算法。

2 实验测试

在本次实验中，将传统的自动化匹配算法和本文设计的算法进行对比，增强实验结果的可靠性。其中，将基于多源时控信息与风险因子的梯级水电远程集控预案自动化匹配算法为算法1，传统的预案自动化匹配算法为算法2，基于人工智能与随机森林算法的预案自动化匹配算法为算法3。为对比上述3种算法在实际应用中的效果，设计的实验具体如下。

2.1 实验准备

以某地区的梯级水电系统为实验对象，利用振动传感器、数据采集器、数据记录器、调制解调器等多种设备采集某梯级水电系统的运行状态信息。使用的硬件设备具体如图2所示。

在上述设备中，利用振动传感器、数据采集器采用某梯级水电系统的多种数据，并用数据记录其长期监测和记录采集的数据，再用调制解调器将采集的模拟信号转换成数字信号。在上述基础上，为保证实验的顺利进行，设置对应的实验参数。其具体的实验参数如表2。

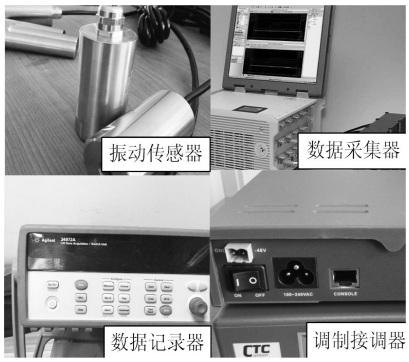


图2 硬件设备

表2 实验参数

序号	实验参数	参数设置
1	风险因子库大小	500 个
2	预案库大小	300 个
3	多源时控信息数据量	1 TB
4	相似度阈值	0.8
5	学习率	0.001
6	时间窗口大小	24 h
7	批次大小	64
8	迭代次数	100

如表2所示，利用上述参数，进行后续实验测试。实验中，利用本文设计的方法计算不同风险因子的风险指数，其风险指数的具体计算结果如表3。

表3 风险因子的风险指数

序号	风险因子	风险指数
1	发电机组故障	0.55
2	变压器故障	0.62
3	开关设备故障	0.35
4	洪水	0.45
5	地震	0.56
6	雷电	0.57
7	操作错误	0.59
8	操作疏忽	0.55
9	黑客攻击	0.33
10	病毒入侵	0.21

如表3所示，将上述计算的风险指数作为基础，构建对应的远程集控预案库，再通过计算对应的相似度，实现对预案的匹配。将上述匹配结果和其他两种算法的匹配结果进行对比，其具体对比结果如下。

2.2 实验结果讨论

为验证上述3种算法在实际应用中的效果，以算法的匹配失败率为评价指标，进行实验测试。实验中，利用三种算法对上述梯级水电系统远程集控预案进行匹配，统计其匹配失败率。其具体统计结果如图3。

如图3所示，在上述实验结果中，算法1的匹配失败率较低，且多次实验中均保持较低水平，算法2和算法3的匹配失败率较高，远高于算法1。因此，本文设计的算法在实际应用中匹配效果较好。两种对比方法匹配失败率较高的原因是本文的实验环境是专门针对梯级水电远程集控预案自动化匹配设置而设计，而其他两种方法在验证时并不是在这一特定环境下进行验证。因此，当实验环境发生改变，会导致两种传统方法的匹配性能下降。

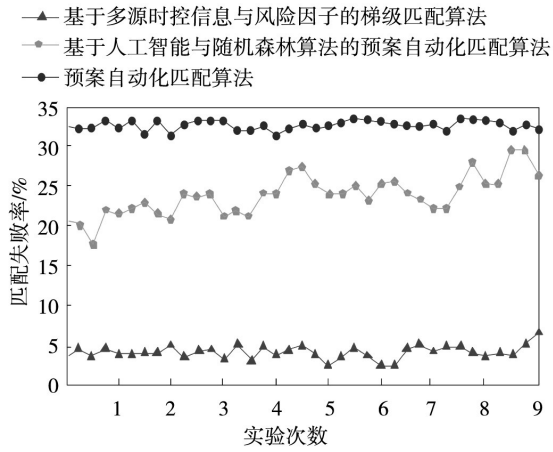


图3 3种算法的匹配结果

为进一步验证上述3种算法在实际应用中的效果，以算法的ARI值为评价指标，对比3种算法的性能。ARI值为调整兰德系数，其数值越接近于1，匹配效果越好。其具体统计结果图4所示。

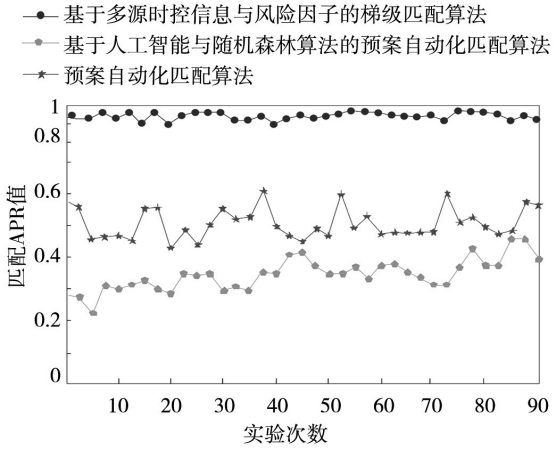


图4 三种算法的匹配ARI值

如图4所示，在上述实验结果中，算法1的ARI值较高，且在多次实验中数值较为接近于1，算法2和算法3的ARI数值较低，远低于算法1。因此，本文设计的算法在实际应用中匹配效果较差。

综上所述，本文设计的基于多源时控信息与风险因子的梯级水电远程集控预案自动化匹配算法在实际应用中匹配失败率较低，ARI数值较高，且更接近于1，匹配效果较好，能够实现对远程集控预案的准确匹配，保证梯级水电系统的稳定运行。

（下转第78页）