

智能电网多智能体信息自动化融合模型构建

崔跃君, 张浩海, 王昊, 王磊, 顾新桥

(北京中电普华信息技术有限公司, 北京 100000)

摘要: 智能电网多智能体信息具有多源异构性, 有效利用信息的准确性较低, 融合效率不佳。针对上述问题, 构建一种智能电网多智能体信息自动化融合模型。利用数据仓库技术集成智能电网多智能体信息, 并实施完整性处理和异常值处理。为降低信息量、信息维度, 提取信息距离熵特征, 计算匹配系数。然后以匹配系数为输入值, 通过K均值聚类算法将同类别的智能电网多智能体信息划分到一起, 实现信息自动化融合。实验结果表明: 应用所构建的自动化模型, 同类融合的信息间一致性达到最大值, 均在0.8以上; 时间开销均达到最小值, 由此说明所构建的模型融合准确性更高, 融合效率较好。

关键词: 智能电网; 多智能体; 信息集成; 距离熵; K均值聚类算法; 自动化融合

中图分类号: TP410.04 文献标识码: A 文章编号: 1003-7241(2025)01-0110-05

Construction of Intelligent Grid Multi-agent Information Automatic Fusion Model

CUI Yue-jun, ZHANG Hao-hai, WANG Hao, WANG Lei, GU Xin-qiao

(Beijing Zhongdian Puhua Information Technology Co., Ltd., Beijing 100000 China)

Abstract: Smart grid multi-agent information has multi-source heterogeneity, the accuracy of effective use of information is low, and the fusion efficiency is poor. To solve the above problems, a smart grid multi-agent information automatic fusion model is constructed. The data warehouse technology is used to integrate the multi-agent information of smart grid and implement integrity processing and outlier processing. In order to reduce the amount and dimension of information, the information distance entropy feature is extracted and the matching coefficient is calculated. Then, with the matching coefficient as the input value, the smart grid multi-agent information of the same category is divided together by K-means clustering algorithm to achieve automatic information fusion. The experimental results show that the consistency of information among the same kind of fusion reaches the maximum value, which is more than 0.8; The time cost reaches the minimum value, which indicates that the fusion accuracy of the model is higher and the fusion efficiency is better.

Keywords: smart grid; multi agent; information integration; distance entropy; K-means clustering algorithm; automatic fusion

0 引言

电力供应是电力公司的主要业务之一, 在发电厂生产电力后, 借助电网进行电力运输, 将电力输送到客户端。电网在电力系统中起到了重要的连通作用, 逐渐向着智能化、自动化方向发展, 与传统电网相比更为复杂, 因此运行过程中产生了庞大的信息量^[1]。这些信息对于监测电网运行状态、制定电网调度策略均具有重要的参考作用, 在智能电网信息存储中, 由于数据量巨大, 所以通常会储存在多个智能体上, 形成智能电网多智能体信息。该信息虽然分担了电网大数据存储压力, 但是容易导致信息中断, 信息利用较为困难^[2]。面对这种情况, 进行信息融合是十分必要的, 通过信息融合能够让同属性信息聚类到一起, 更加方便人们查询和使用。

关于信息融合问题, 夏伟等人^[3]首先针对多源异构数据进行了预处理, 包括数据纠偏、时间配准以及时序数据填充等三个方面, 然后计算了数据之间的关联性, 最后采用联合卡尔曼滤波算法将同类数据融合到一起, 完成配电网多源异构数据融合。周文文等人^[4]提出一种基于改进D-S证据理论的融合方法, 在研究中首先利用分类算法对数据进行分类, 然后借助D-S证据理论设计合成规则并进行改进, 最后利用改进后合成规则实现数据分类融合。莫慧凌等人^[5]提出了一种基于联邦学习的融合方法, 该研究中针对待融合数据提取其特征, 然后采用平均聚合算法对特征进行聚合, 以保证同一模态的数据提取的特征具有相似性, 最后通过联邦学习系统模型将多源异构数据融合到一起。Vakacharla等人^[6]提出用于太阳能到电网集成的最先进电力电子系统, 该文回顾了为解决温度、辐照度和部分遮光条件的变化而提出的MPPT

*基金项目: 北京市自然科学基金项目 (3202031)

收稿日期: 2023-06-01

技术,该技术有助于概述用于将光伏板集成到电网的最先进电力电子系统。构建了各种接口电力电子架构,包括微型逆变器、中央逆变器和模块化逆变器方法,其工作在低开关频率和高开关频率下,具有谐振和脉宽调制软开关或硬开关。

虽然上述研究取得一定进展,但由于信息量较大,以往融合方法的效率较慢,还容易出现较大融合偏差。针对这一问题,研究一种智能电网多智能体信息自动化融合模型。通过该研究以期能够规范智能电网多智能体信息,提高信息利用率。

1 智能电网多智能体信息自动化融合研究

智能电网多智能体信息量不仅巨大还繁杂,使得其挖掘难度较大,利用率较低。针对这种情况,通过有效融合能够规范信息,让信息更加方便地被利用。本研究主要分为三部分,即智能电网多智能体信息集成与预处理、信息降维、基于特征匹配的信息融合。

1.1 多智能体信息集成与预处理

智能电网多智能体信息自动化融合并不是简单地将信息聚集在一起,聚集仅是融合的第一步,还需进行多智能体信息集成。信息存在多个智能体当中,具有多源异构性,考虑到现阶段的电网已经实现较大规模智能化建设,通过数据仓库技术从智能体中抽取智能电网信息,在此过程中使用大数据信息快速抽取算法找到能够正确分割出有效多智能体信息的标签位置,从而抽取有效智能电网信息,其中大数据信息快速抽取算法是作为海量信息内容解析中最常用的信息处理技术^[7],是信息集成、文本结构化等任务的主要处理算法;在此基础上利用信息集成总线向提供数据的子系统发起连接请求,适配器接到请求后,将所需要处理的信息与需求相匹配,满足调用需求,然后将处理好的信息发送给服务端,最后通过集成器实现集成。通过此方式,在智能电网中引入集成信息,在智能电网标准化融合过程中,能够以实际运行状态的标准实现对多智能体信息集成的准确化判定。基于数据仓库的多智能体信息集成框架如图1所示。

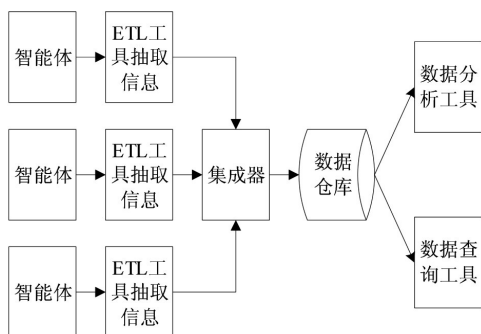


图1 基于数据仓库的多智能体信息集成框架

集成在一起的智能电网多智能体信息不能直接进行融合,多源异构性使数据之间存在差异,因此需要进行预处理,具体步骤包括完整性处理、异常值处理^[8]。

(1) 完整性处理

完整性处理是针对信息缺失而进行填补的一种方法。信息在集成过程中会出现缺失的现象,针对此情况,首先利用标准欧氏距离选取近邻样本,然后结合加权系数来估算缺失值^[9]。估算公式如下:

$$x_i = \sum_{i=1}^M w_i \tilde{x}_i \quad (1)$$

式中, x_i 代表缺失的智能电网多智能体信息; \tilde{x}_i 代表第*i*个近邻样本的多智能体信息; M 代表近邻样本数量; w_i 代表第*i*个最近邻的加权系数。

(2) 异常值处理

异常值处理是处理偏离智能电网多智能体信息正常分布的离群信息,通常通过偏度检验法来实现^[10]。检验公式如下:

$$L_i = \frac{\sqrt{n} \sum_{i=1}^k (x_i - \bar{x})^3}{\left[\sum_{i=1}^k (x_i - \bar{x})^2 \right]^{\frac{3}{2}}} \quad (2)$$

式中, L_i 代表第*i*个智能电网多智能体信息的偏度; \bar{x} 代表信息平均值; x_i 代表第*i*个智能电网多智能体信息。

当 L_i 大于临界值时,认为 x_i 是异常值,直接剔除,完成异常值的处理。

1.2 信息降维

智能电网多智能体信息维度较大,巨大信息量特征使融合工作量处理难度增加。面对这种情况,需要进行降维处理^[11]。具体过程如下:

步骤一:输入高维智能电网多智能体信息。

步骤二:将信息转换为矩阵形式,记为 A_{nm} 。其中, n 代表属性数; m 代表样本数。

步骤三:计算属性 b_i 的信息熵 H_{b_i} ,计算公式如下:

$$H_{b_i} = - \sum_{i=1}^n p_i \log p_i \quad (3)$$

式中, p_i 代表属性 b_i 的概率。

步骤四:判断 H_{b_i} 是否大于信息熵阈值?若大于, b_i 被留下,否则属性 b_i 被剔除,最终剩余 N 个属性。

步骤五:计算属性均值 \bar{c} 。计算公式如下:

$$\bar{c} = \frac{\sum_{i=1}^N a_{ij}}{N} \quad (4)$$

式中, a_{ij} 代表智能电网多智能体信息样本*j*的属性 b_i 的取值。

步骤六:中心化处理。计算公式如下:

$$d_{ij} = a_{ij} - \bar{c} \quad (5)$$

根据公式(5)中心化处理结果,得到中心化矩阵,记为 $D_{Nm} = \{d_{ij}\}_{Nm}$ 。计算不同属性的协方差,构成协方差矩阵,记为 B 。计算协方差矩阵的特征向量,记为 V_{Nk} 。

步骤七:选择前 k 个特征向量组成特征向量矩阵 V 。

步骤八:按照下述公式实现降维结果。

$$E = D \times V^T \quad (6)$$

式中, E 代表降维后的智能电网多智能体信息。降维后,信息量得到了有效缩减,能够有效降低后续融合任务的工作量^[12-14]。

1.3 基于特征提取的信息融合实现

信息融合是将同类信息聚集在一起,按照所提取特征进行信息分类。如何判断是否为同类,需要依据信息特征来判断。基于此,将进行特征提取,然后以此为依据,进行特征匹配,实现信息融合,具体过程如下:

选用距离熵特征作为提取目标,该特征是对距离进行量化的一种数值。距离熵特征进行提取过程如下:

步骤一:假设将每一个智能体来源的信息作为一个单元,共计 m 个单元,记为 $G_j, j=1, 2, \dots, m$ 。每个单元共计 n 个数信息量,记为 $g_i, i=1, 2, \dots, n$,每个信息有 P 个属性。针对单元中每一个信息属性进行归一化处理,将所有属性值都映射到0~1之间。处理公式如下:

$$\hat{u}_{jip} = \frac{u_{jip}}{\sum_{i=1}^n u_{jip}}, p=1, 2, \dots, P \quad (7)$$

式中, u_{jip}, \hat{u}_{jip} 代表归一化前、后的第 j 个单元第 i 个智能电网多智能体信息的第 p 个属性值。

步骤二:计算 \hat{u}_{jip} 与最大属性值 $\max(\hat{u}_{jip})$ 之间的距离 d_{jip} ,公式如下:

$$d_{jip} = \sqrt{\sum_{i=1}^n |\hat{u}_{jip} - \max(\hat{u}_{jip})|^2} \quad (8)$$

步骤三:计算单元距离熵,计算公式如下:

$$h_{ji} = \begin{cases} -\sum_{i=1}^n \left(\frac{d_{jip}}{\sum_{i=1}^n d_{jip}} \right) \cdot \ln \sum_{i=1}^n \frac{d_{jip}}{\sum_{i=1}^n d_{jip}}, d_{jip} \neq 0 \\ 0, \text{ else} \end{cases} \quad (9)$$

式中, h_{ji} 代表第 j 个单元第 i 个智能电网多智能体信息的单元距离熵。

针对上述提取到单元距离熵,下面进行信息融合,具体过程如下:

步骤一:输入上述提取到的局部距离熵,记为 $h_{ji}, j=1, 2, \dots, m; i=1, 2, \dots, n$ 。

步骤二:计算单元信息匹配权重。计算公式如下:

$$v_{ji} = \frac{1 + \frac{h_{ji}}{\ln n}}{n - \sum_{i=1}^n h_{ji}} \quad (10)$$

式中, v_{ji} 代表第 j 个单元第 i 个智能电网多智能体信息的匹配权重。

步骤三:根据 v_{ji} 计算局部匹配系数 Q_{ji} 。计算公式如下:

$$Q_{ji} = \sum_{i=1}^n v_{ji} \cdot d_{jip} \quad (11)$$

由此得到 n 个智能电网多智能体信息的局部匹配系数,记为 $Q_j = \{q_{j1}, q_{j2}, \dots, q_{jn}\}$

步骤四:对 Q_j 进行归一化处理,参考公式(7),处理结果记为 \hat{Q}_j 。

步骤五:基于 \hat{Q}_j 计算整体距离熵,计算参考公式(8)和(9),记为 H_j 。

步骤六:根据 H_j 计算整体信息匹配权重 V_j 。计算参考公式(10)。

步骤七:按照下述公式计算全局匹配系数。

$$F = \sum_{j=1}^m V_j \times Q_j \quad (12)$$

步骤八:基于全局匹配系数,利用K均值聚类算法实现同类别聚类。具体过程如图2所示。

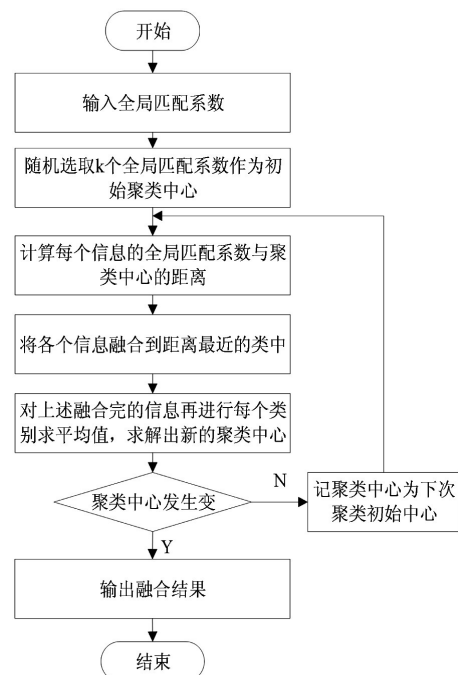


图2 基于K均值聚类算法的融合流程

同类别的智能电网多智能体信息划分到一起,即完成信息融合研究。

2 模型应用与测试

2.1 智能电网多智能体信息准备

利用数据仓库技术从17个智能体数据库中集成智能电网多智能体信息,组成17个单元的信息,以此作为信息融合基础数据。该信息基本情况如表1所示。

表1 智能电网多智能体信息基本情况

信息单元序号	实例个数/个	属性个数/个	信息类别/类
1	5	8	2
2	4	7	1
3	4	6	3
4	3	8	2
5	5	3	2
6	4	5	4
7	6	4	2
8	5	2	3
9	4	1	2
10	3	4	2
11	2	6	2
12	4	5	2
13	6	2	3
14	4	2	2
15	2	4	1
16	4	2	2
17	6	10	3

注:17个智能体数据库中集成智能电网多智能体信息类别共计5个,每个单元中包含的信息类别并不是固定数量的。

2.2 距离熵特征提取结果

基于章节1.3研究,提取单元距离熵以及整体距离熵。其中,单元距离熵以前3个单元中的信息为例,结果如图3和图4所示。

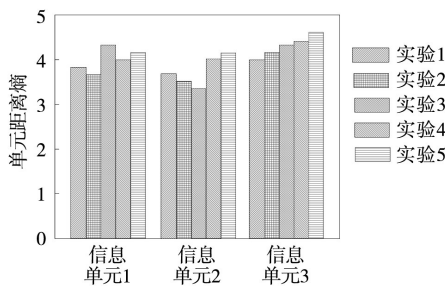


图3 单元距离熵计算结果示例

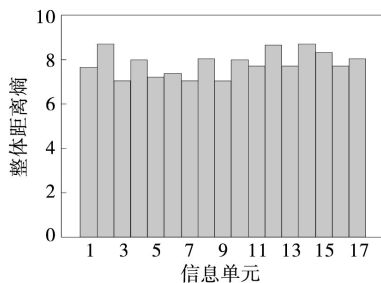


图4 整体距离熵计算结果

从图3、图4中可以看出,所计算出的单元距离熵结果

和整体距离熵结果变化不大,能够维持稳定状态。

2.3 融合效果测试

利用所构建的融合模型以及三种对比模型(基于联合卡尔曼滤波融合模型、基于改进D-S证据理论的融合模型、基于联邦学习的融合模型)对表1中的智能电网多智能体信息进行自动融合。根据融合结果,计算融合信息之间的一致性以及融合时间开销,以此判断融合模型的有效性。其中一致性计算公式如下:

$$\xi_i = \frac{\sum_{j=1}^n (x_{ij} - \bar{x}_i)^2}{n} \quad (13)$$

式中, x_{ij} 代表第*i*类融合结果中第*j*个智能电网多智能体信息; \bar{x}_i 代表第*i*类融合结果的信息平均值。

信息一致性以及时间开销如图5和图6所示。

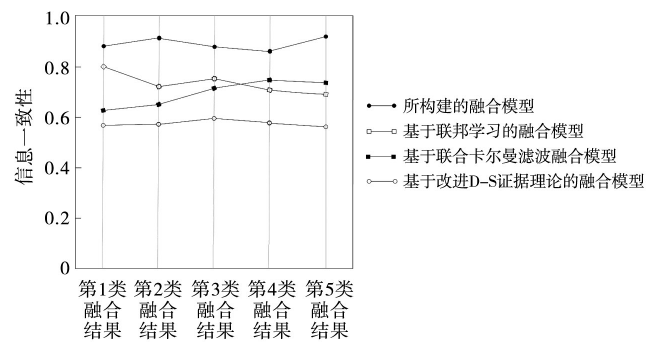


图5 信息一致性测试结果对比图

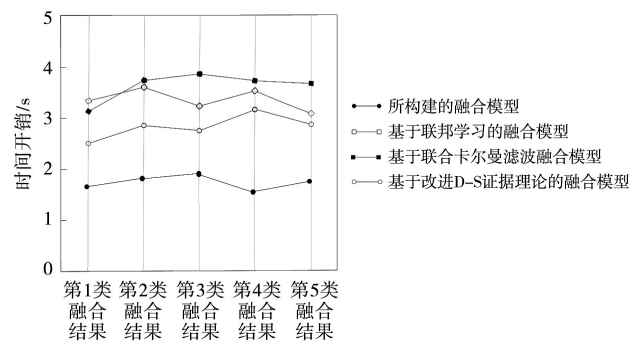


图6 时间开销测试结果对比图

从图5和图6中可以看出,与基于联合卡尔曼滤波融合模型、基于改进D-S证据理论的融合模型、基于联邦学习的融合模型应用效果进行对比,所构建的自动化模型应用下,同类融合的信息间一致性更大,均在0.8以上。此外,时间开销均要更少。由此说明所构建的模型融合准确性更高,融合效率更高。

3 结束语

本文构建一种智能电网多智能体信息自动化融合模型,通过聚类算法实现同类寻找与融合。实验研究得到如下结论:

(1) 所计算出的单元距离熵结果和整体距离熵结果

变化不大,维持稳定状态。

(2) 通过信息一致性对比,本文所构建模型的信息间一致性较大;

(3) 对比时间开销指标,本文模型的时间开销较少,证明了所构建模型的应用性能。

然而,本研究还有不完善之处,例如仅提取了一种特征,单一特征所包含的信息量有限,所以在未来研究中需要进一步提取信息特征,进行更加深入的研究,以期提高融合准确性。

参考文献:

- [1] 侯鑫鑫,朱文佳,朱莉,等.多源异构学术成果大数据的整合与揭示[J].情报理论与实践,2021,44(4):162-168.
- [2] 林裕新,陈楠,蔡建逸.基于多源信息融合的配电网调度数据安全管控技术[J].电测与仪表,2024,61(6):118-125,132.
- [3] 夏伟,蔡文婷,刘阳,等.基于联合卡尔曼滤波的配电网多源异构数据融合[J].电力系统保护与控制,2022,50(10):180-187.
- [4] 周文文,万晓冬,李文.基于改进D-S证据理论的数据融合目标分类[J].半导体光电,2021,42(1):121-126.
- [5] 莫慧凌,郑海峰,高敏,等.基于联邦学习的多源异构数据融合算法[J].计算机研究与发展,2022,59(2):478-487.
- [6] Vakacharla V R, Gnana K, Xuewei P, et al. State-of-the-art power electronics systems for solar-to-grid integration[J]. Solar Energy, 2020, 210(11): 128-148.

(上接第104页)

例关系。其中在并发用户数量为500个之前时,本文系统的数据管理覆盖度均为100%。当并发用户数量超过500个后,本文系统的数据管理覆盖程度数值呈现小幅度下降趋势,在并发用户数量为900个时,本文系统的数据管理覆盖程度依然高达97.5%左右。该数值表明:本文系统在实际应用中可为用户提供良好的数据采集和传输能力的同时,使用户针对于钩缓检修数据的管理更为准确和精细。

3 结束语

本文设计基于无线网络的钩缓检修数据采集和传输系统,在该系统内使用RFID射频芯片采集钩缓检修数据。RFID射频芯片采集数据的形式具备较高的精确性和安全性。经过实验验证:本文系统具备较为精确的数据采集能力和较强的通信传输能力,为用户管理钩缓检修数据提供了有效的途径,具备较强的应用效果。

参考文献:

- [1] 海英,马温喜,马新生.PC-MEMS地震数据采集与预警信息传输系统研究[J].地震工程与工程振动,2021,41(5):38-46.
- [2] 李锦明,刘梦欣,成乃朋,等.基于FPGA的多路无线数据采集

[7] 于晓翠,陈亮,林泽源.基于人工智能的大数据信息快速抽取算法研究[J].电子设计工程,2021,29(5):149-153.

[8] 李展鹏,邹孝付,苏雍贺,等.基于FPGA的多源异构数据并行可配置采集方法[J].计算机集成制造系统,2021,27(4):1008-1020.

[9] 刘庆利,王文广,张学翠,等.基于改进DS证据理论的指挥控制系统信息融合方法[J].火力与指挥控制,2022,47(2):82-87.

[10] 龙北平,吴家杰,郑勤华,等.一种基于权重相似性模型的不动产数据整合方法[J].测绘通报,2021(6):122-126.

[11] 聂庆科,孙广,郝永攀,等.多源异构监测数据融合方法及应用[J].科学技术与工程,2022,22(13):5348-5357.

[12] 孙文成,李健,彭宇辉,等.基于样本不均衡和特征优选多源融合的输电线路故障类型辨识[J].电测与仪表,2024,61(12):79-89.

[13] 高瑞翔,徐腾寅,房鹤飞,等.基于多传感器信息融合的机床测量数据自动补偿系统[J].自动化技术与应用,2024,43(6):60-63.

[14] 李雪冬,黄莹,胡勇,等.基于PAU-LSSVM模型融合多特征变量的电力负荷预测[J].电子器件,2021,44(2):474-479.

作者简介:崔跃君(1983-),男,本科,中级工程师,研究方向:信息化建设及研究。

集系统设计[J].中国测试,2019,45(12):102-105,118.

[3] 王立文,郭贤,陈斌,等.多传感器跑道积冰预警无线数据采集系统设计(英文)[J].机床与液压,2019,47(6):38-45.

[4] 马鑫龙,艾冬生,孙发鱼,等.基于MQTT的遥测数据无线传输[J].探测与控制学报,2020,42(2):61-64.

[5] 吴洋,徐广平,耿林,等.基于FPGA的激光回波数据采集与传输系统[J].激光与红外,2020,50(11):1345-1351.

[6] 王红亮,王柳明.基于MLVDS和USB3.0的多节点数据传输系统设计与实现[J].电子技术应用,2019,45(1):42-45,50.

[7] 李炳臻,李杰,胡陈君,等.箭载测控系统高速数据采集传输电路的设计[J].中北大学学报(自然科学版),2020,41(6):571-576.

[8] 张美燕,蔡文郁.基于Dubins曲线的无线传感网聚类移动数据采集算法[J].传感技术学报,2019,32(4):603-609.

[9] 肖时辉,梅敏彰,唐孟雄,等.基于LoRa的城市立交桥远程监测数据采集系统[J].公路,2021,66(2):87-93.

[10] 闫静纯,富帅,苏浩航,等.基于nA级电流信号的多通道高精度采集系统设计[J].红外技术,2021,43(5):417-421.

[11] 刘士兴,宋亚杰,刘光柱,等.新型软X射线相机多通道数据采集系统设计[J].核电子学与探测技术,2019,39(5):587-591.

作者简介:王喆波(1979-),男,硕士,工程师,研究方向:自动控制。