DOI:10.20033/j.1003-7241.(2025)06-0085-05

基于DAMA体系的智能电网拓扑数据库构建研究

魏 军[1]，王 华[1]，马彦宏[2]，温金梅[1]，王 栋[1]

(1.国网甘肃省电力公司互联网事业部，甘肃 兰州 730050；
2.国网甘肃省电力公司科技部，甘肃 兰州 730050)

摘 要：为缩短传统数据库集中计算时间，提高数据库吞吐能力，增强数据库结果输入的准确率，构建基于数据管理知识体系(data management association International，DAMA)体系的智能电网拓扑数据库。分析智能电网拓扑数据周期，建立数据库模型，根据DAMA体系的完整性和一致性原则，确定电网的拓扑状态，得到拓扑结构，利用不同的连接方式对拓扑网络结构进行划分。建立拓扑数据库结构，完成信息协议同步，使用整体同步并行(bulk synchronous parallel，BSP)计算实现数据查询，使用户在不同模式下可以对电网数据进行集中调度和智能访问。实验结果表明，采用该数据库的输出吻合度与真实值高达98%，集中计算时间更短，得到的无创检测(non-invasive detection，NIND)数量更低，吞吐能力更强，具有很好的管理能力。

关键词：DAMA体系；智能电网；电网拓扑；拓扑数据库；数据库构建

中图分类号：TP311.13 文献标识码：A 文章编号：1003-7241(2025)06-0085-05

Research on the Construction of Smart Grid Topology Database Based on DAMA System

WEI Jun[1], WANG Hua[1], MA Yan-hong[2], WEN Jin-mei[1], WANG Dong[1]

(1. Internet Business Department of State Grid Gansu Electric Power Company, Lanzhou 730050, China;
2. Science and Technology Department of State Grid Gansu Electric Power Company, Lanzhou 730050, China)

Abstract: In order to shorten the centralized computing time of traditional database, improve the database throughput and enhance the accuracy of database result input, a smart grid topology database based on data management association International (DAMA) system is constructed. The topology data cycle of smart grid is analyzed, the database model is established, the topology state of power grid is determined according to the integrity and consistency principle of DAMA system, the topology structure is obtained, and the topology network structure is divided by different connection modes. The topology database structure is established, the information protocol is synchronized, and the data query is realized by BSP computing. Users can conduct centralized scheduling and intelligent access to the power grid data in different modes. The experimental results show that the output coincidence degree of the database is as high as 98% with the real value, the centralized computing time is shorter, the number of non-invasive detection (NIND) is lower, the throughput is stronger, and it has good management ability.

Keywords: data management association International system; smart grid; grid topology; topology database; database construction

0 引言

经济进步的发展，使我国电力系统越来越成熟，能够实现高压、大容量、长距离输电，互联电网范围越来越大，随着新技术的成熟，新能源、可调节负荷等技术都介入了电网，设备的增加使数据越来越多样，这也导致了电网的不确定增加。电力系统内部可调节设备较多，需要不断进行运行调整，通过分析控制，确保电网运行的准确性[1]。

数据库是智能电网的重要组成部分，能够为调控中心提供源源不断的数据，确保数据库能够精准计算，目前电网数据库分为商用数据库和实时数据库，内部包含大量不同的数据，建立数据模型。国家技术的完善使电网监控范围越来越广，在监控过程中，加入了更多先进的中央处理器(central processing unit，CPU)、图形处理单元(graphics processing unit，GPU)、并行计算机等设备，使电网的运行方式是从单机集中式转换成集群分布式，从而更好地满足实时调度的要求。黄华等[2]人提出了基于图数据库的电网标准信息模型(common information model for electric power systems，CIM/E)模型构建及网络拓扑，利用图数据库建立数据模型，实现信息查询，并设计数据模型装载工具，遵照完整性、一致性和高效性的原则，更好地完成电网拓扑，但在构建过程中，缺少对电网内部设备融合的考虑，导致数据库实用性不

收稿日期：2023-11-14

高。王艺霏等人[3]提出了基于随机矩阵理论的智能电网大数据体系结构，利用随机矩阵理论获取电网数据样本，通过训练、学习和模型计算将宏观数据转换成微观数据，从而更好地使用户使用，但这种体系结构对无线通信模式过于依赖，可操作性较差。

为了有效解决上述问题，本文构建了基于DAMA体系的智能电网拓扑数据库。DAMA体系指的是数据管理协会，经过多年实践和总结得到的信息适用于电网大数据时代，相比较于其他模式，DAMA体系能够更好地管理数据，确保数据上传的安全性。综上所述，本文基于DAMA体系建立了智能电网拓扑数据库，并同时检测数据输出值和吞吐性能，判断数据库的可行性。

1 基于DAMA体系的智能电网拓扑数据库模型

DAMA体系是一个标准化的体系，能够管理电网数据周期，分析企业数据，通过多层次的判断实现数据汇总。智能电网拓扑数据周期如图1所示。

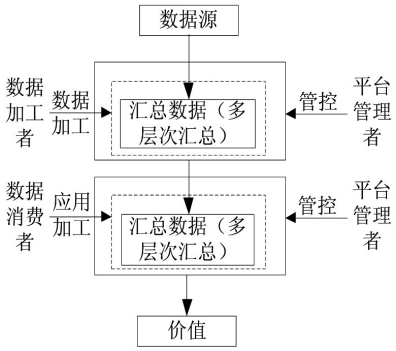


图1 智能电网拓扑数据周期

根据图1可知，智能电网拓扑数据共分为多层次汇总阶段和应用阶段[4]。在多层次汇总阶段，加工者需要确定原始的清单数据，并将清单数据从多层次上进行汇总，利用数据消费者的数据汇总，实现多层次的数据管理，使消费者使用的数据具有直接价值或间接价值。

根据数据汇总的不同阶段，将数据分为汇总层数据和应用层数据。汇总层数据的数据不同，模型也不同，必须要从正确的视角对数据进行合并处理和降维处理，使数据能够得到更好地归一化，利用数据阈值将数据分为业务阈模型和挖掘阈模型，从智能电网的工作业务角度对数据进行提取，将数据的属性和特征汇总，防止数据出现流失现象。应用层能很好地记录数据状态，提取待处理数据的混合特征，并根据混合特性，构成待处理数据的特征向量集；将所述特征向量集输入到业务阈模型中，以便预测数据相应的状态；如果所述状态指示该数据为冷数据，则将数据进行迁移，以实现数据的预测[5]。

在建立数据库模型时，要遵照DAMA体系的完整性和一致性原则，分析电网数据的拓扑状态，并对数据进行表述。采集电网设备上的数据，对数据建模，如果电网数据能够满足IEC 61970−301电力系统公共要求则定义该数据为模型定点，对其他数据定义为边。在数据库内部，顶点为数据库中心，用户只能直接对顶点的数据进行访问，无法对数据边进行访问，如果用户想要对数据边缘进行访问，需要从顶点出发，进行遍历查询。在对数据库访问的过程中，所需要经过的时间与跳数呈正比关系，用户想要访问的顶点和边数越多，需要花费的时间越多。因此，在调整数据库功能时，可以从数据顶点进行优化[6]。首先要建立一个顶点buffer和一个索引buffer，然后在设置了顶点buffer后，每一帧的顶点数据都会发生改变，因此需要进行动态的更新。在每次提交一个顶点数据时，使用glBufferSubData将仅提交实际使用过的数据。这样就不会再触发额外的记忆指派，而且传输的效率也会更高，所需要花费的时间也会更短。

智能电网利用厂站、母线、开关等不同的设备，采用交流线路行程模型系统拓扑结构，建立的智能电网数据网络拓扑如图2所示。

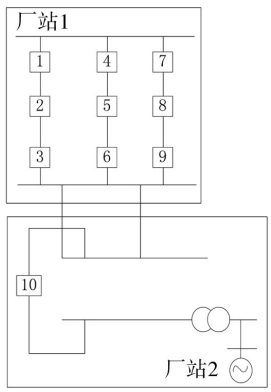


图2 智能电网数据网络拓扑

根据图2可知，在智能电网中，开关等设备是要远远多于线路设备，用户可以使用不同的设备对数据库顶点进行修改，通过修改访问模式，提高访问效率，同时更好地满足不同的准则。智能电网拓扑结构可以通过式(1)表示。

$$M=(Z,L) \tag{1}$$

式中，M表示智能电网拓扑结构；Z表示在拓扑结构内部的所有顶点；L表示拓扑结构的所有边。

利用划分手段，对拓扑结构进行划分，从而得到更多不同的小分量，设定分割了n次，则顶点集合可以表示为：

$$Z=\{Z_1.Z_2,\cdots,Z_n\} \tag{2}$$

边集合可以表示为：

$$L=\{L_1.L_2,\cdots,L_n\} \tag{3}$$

如果分割次数$i\neq j$，则有：

$$M_i\neq M_j \tag{4}$$

即：

$$\{Z_i,L_i\}\neq\{Z_j,L_j\} \tag{5}$$

对拓扑网络进行划分，如图3所示。

观察图3可知，智能电网拓扑网络中任何不同的2个顶点都可以连接，连接方式可以通过一条边或者多条边实现。如果拓扑结构顶点不能连接，则拓扑结构可以形成连通分量，如图4所示。

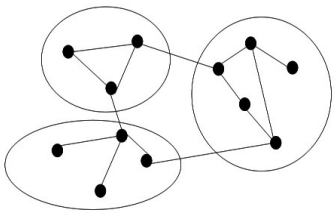


图3 拓扑网络划分子图

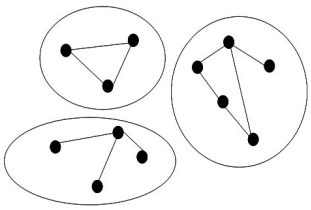


图4 连通分量

在建立数据库时，除了要考虑电网的拓扑关联之外，还要考虑其他关联，如等级关联、基准电压关联等[7]。

2 基于DAMA体系的智能电网拓扑数据库

2.1 数据库架构

本文在综合多种关联后，建立基于DAMA体系的智能电网拓扑数据库架构如图5所示。

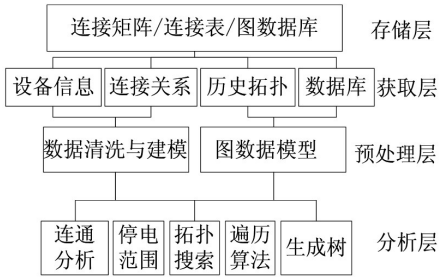


图5 基于DAMA体系的智能电网拓扑数据库架构

根据图5可知，本文设计的基于DAMA体系的智能电网拓扑数据库主要由获取层、预处理层、存储层和分析层四部分组成。获取层能够大规模采集电网数据，提高数据的结构，同时收集不同的电网数据，在确定物理设备的连接关系后，分析拓扑关系，并记录历史数据。预处理层具有很强的处理能力，能够处理电网内部复杂的逻辑关系，并对采集的数据进行较为全面的整合，通过映射方式得到信息，并清洗噪声数据[8]。在完成数据处理后，得到的电网拓扑数据已经具备基本的逻辑，可以进行存储，分析数据结构，采用有效的方法，实现数据存储。分析层是数据库的核心层次，能够为电网数据提供支撑服务，得到的数据以展开的方式分析，从而实现拓扑搜索，确定连通方式，方便后续使用。

2.2 数据库功能

数据库具备查询访问功能，通过DAMA体系描述数据节点和数据边，从而确定数据的运行状态，在数据库内部对数据进行查询访问，数据库内部拥有多个程序接口，可以为数据库提供支撑，同时担当计算引擎的角色。数据库的每个顶点和数据边都具备计算能力，可以以动态的方式实现计算，计算结果会自动进入存储单元，不同顶点的信息通过局部计算实现信息交互，以并行计算的方式得到计算模型。利用BSP的方式对数据进行局部计算，BSP计算原理如图6所示。

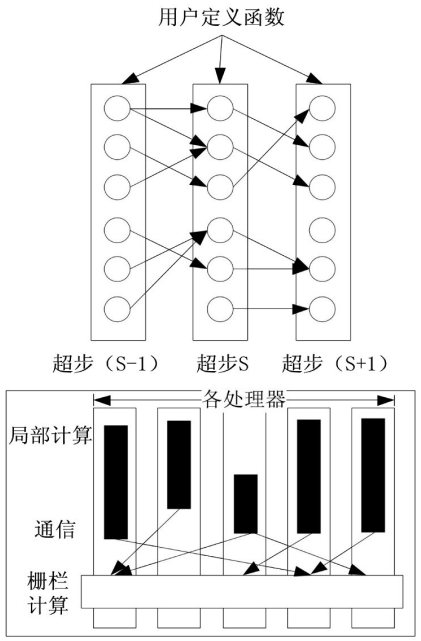


图6 BSP计算原理

根据图6可知，BSP包括三种不同模式，分别是局部计算、通信和栅栏同步，数据库内部的每个不同CPU都具备不同的计算能力，可以独立地计算数据，并实现异步处理，通信方式可以更好地交互数据，利用栅栏同步技术防止在遇到路障时，处理器难以清除。针对不同类型的数据库，使用的查询方法不同。如果数据库架构得到的数据为关系型数据，只需要连接不同的数据表，就可以降低查询的复杂程度，使用户在短时间内完成查询，如果数据为层次型数据库，内部可能存在较多的虚拟节点或冗余数据，则需要对模型进行格式化处理，从而提高查询效率。

3 实验研究

为进一步研究本文构建的基于DAMA体系的智能电网拓扑数据库实际工作效果，本文进行对比实验。选取的实验对象为辽宁省省级电网，电网内部同时拥有多

个设备，如表1所示。

利用本文建立的数据库得到数据库模型，生成的数据共有69 250个顶点和48 233条边，实验过程硬件的CPU型号为Intel XeonE5−26202，能够提供2.5 GHz的频率，数据库内存为128 GB，操作系统为CentOS 6.8.

表1 电网设备

设备	数量
变电站	3 104
母线	2 715
发电机	4 610
负荷	2 013
传输线	2 002
变压器	4 258

根据上述环境和参数设定进行实验，通过多线程的运算方式对不同数据进行局部计算，并通过协调分析改变数据源的时间，在生成不同计算数据后，比较本文数据库输出数值与真实数值，得到的实验结果如图7所示。

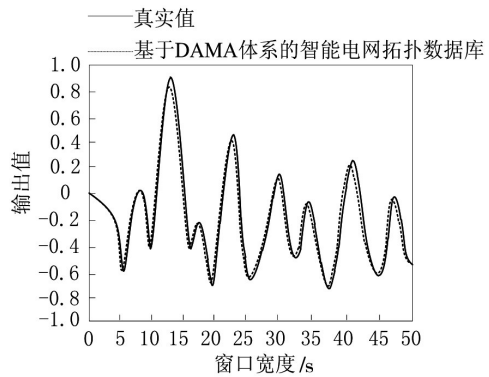


图7 输出数值实验结果

根据图7可知，本文构建的数据库输出数值与真实值吻合度在98%以上，准确度极高，输出的数据结果与真实结果基本一致，数据在输出过程中，不会受到窗口的宽度影响，具有很强的稳定性。造成这种现象的原因是本文设计的电网数据库呈现开放式的模式，通过中间层实现电网数据调度，配合无线射频网络对网络模块中的数据进行上传，数据库内部的微型操作系统(tiny operating system，TinyOS)能够实现接口设计，利用总线技术采集数据，建立存储协议，实现网络传输和智能控制。在存储模块中增加了SRAM和FLASH存储器，能够实现智能控制，对电网内部数据能够快速捕捉，数据库中的内核微处理器能够更好地得到网络通信协议，根据通信装填实现信息监测。并设置FIFO RAM缓冲区从而处理数据特征。上位机模块以远程传输的模式实现控制，从而完成数据实时显示。

将基于DAMA体系的智能电网拓扑数据库与传统基于图数据库的电网CIM/E模型、基于随机矩阵理论的智能电网大数据体系进行实验对比，不同方法的NIND实验结果如表2所示。

表2 NIND实验结果

方法	NIND
基于DAMA体系的智能电网拓扑数据库	80
基于图数据库的电网CIM/E模型	30
基于随机矩阵理论的智能电网大数据体系	10

数据库的吞吐量实验结果如图8所示。

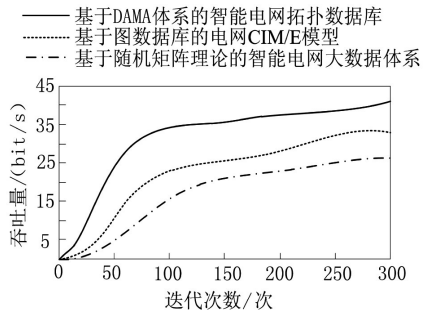


图8 数据库吞吐量实验结果

根据图8可知，随着NIND数量的增加，数据库的吞吐量在逐渐降低，基于随机矩阵理论的智能电网大数据体系结构的NIND数量最高，因此数据吞吐量最低，本文建立的系统NIND数量最低，因此数据吞吐量也在逐渐减少。在迭代过程，本文构建的数据库吞吐量始终高于传统的数据库，由此可见，本文构建的数据库信息融合能力更强，能够在短时间内实现数据的均衡调度。

三种数据库的网络拓扑行为分析过程花费的时间如表3所示。

表3 数据库花费时间实验结果

方法	执行时间/s
基于DAMA体系的智能电网拓扑数据库	0.025 1
基于图数据库的电网CIM/E模型	3.485 1
基于随机矩阵理论的智能电网大数据体系	1.584 4

根据表3可知，本文提出的数据库花费的时间最短，通过加快拓扑分析速度估计运算状态，从而使用户能够得到毫秒级的基础计算模型。而传统的数据库计算模式对数据进行多重分析，需要重设阈值，因此计算结果较差，花费时间较长。

4 结束语

DAMA能够更加全面准确地描述电网拓扑结构，从而实现数据库构建，本文利用DAMA构建了智能电网拓扑数据库，创建数据库模型，分析数据库之间的关联，在得到拓扑架构后，建立数据库架构，并设置了查询功能。实验结果表明，与传统的电网拓扑数据库相比，本文的数据库具有很强的工作效率，能够提供精准的计算服务，确

保电力系统的工作效率。本文在建立数据库过程中，考虑电网拓扑关系，但是对其他关联，如等级关联、基准电压关联的分析描述较少，未来需要更进一步研究这一问题，从而提高数据库的实用效果。

参考文献：

[1] 董威，李永刚.基于复杂网络的智能电网网络攻击影响分析研究[J].信息网络安全，2020(1)：52-60.

[2] 黄华，戴江鹏，王毅，等.基于图数据库的电网 CIM/E 模型构建及网络拓扑[J].电力系统自动化，2019，43(22)：122-129.

[3] 王艺霏，李贤，来骥，等.基于随机矩阵理论的智能电网大数据体系结构设计[J].计算技术与自动化，2019，38(3)：17-21.

[4] 李广野，李伟，王丽霞，等.基于图数据库的电网拓扑搜索引擎研究[J].信息技术，2019，43(5)：155-158.

[5] 孙伟，丁震，王建平.基于多跳无线传感器网络的智能电网延时优化研究[J].电子测量与仪器学报，2019，33(8)：140-146.

[6] 李澄，陈颢，刘恢，等.基于多智能体共享信息的低压配电网拓扑与数据建模技术研究[J].电子测量技术，2020，43(12)：18-25.

[7] 石智永，管晓峰，侯尽然，等.基于数据挖掘的智能电网调度技术研究[J].电子设计工程，2019，27(22)：170-174.

[8] 肖飞，杨国健，邓祥力，等.基于电网故障拓扑分析及多数据综合的复杂故障诊断方法[J].水电能源科学，2020，38(2)：189-192.

作者简介：魏军（1987—），女，硕士，高级工程师，研究方向：数据管理。

（上接第 75 页）

表 2　基于知识图谱的故障根因定位方法的准确性评估结果

故障类型	准确率	召回率	F1 得分
主机 CPU 陡增故障	0.928 6	0.764 7	0.838 7
主机 CPU 爬升故障	0.803 4	0.836 5	0.819 6
主机内存过高故障	0.914 3	0.905 6	0.909 9
主机磁盘不足故障	0.895 3	0.792 3	0.840 6
主机磁盘读 IO 故障	0.862 7	0.892 7	0.877 4
主机磁盘写 IO 故障	0.792 3	0.721 8	0.755 4
容器 CPU 陡增故障	0.913 2	0.752 3	0.824 9
容器 CPU 爬升故障	0.819 2	0.870 1	0.843 8
容器内存过高故障	0.759 1	0.829 1	0.792 5
容器进程中止故障	0.863 9	0.719 1	0.784 8

3　结束语

综上所述，提出一种基于知识图谱的微服务架构的故障根因定位方法。首先采用多维关联分析算法，对 M 商城中多种维度的原始数据进行特征提取，构建可以直接用于人工智能分类算法的数据集，进而训练模型，得到成熟的系统是否发生故障的分类模型。结合知识图谱的 PageRank 算法，对 M 商城微服务架构的各种拓扑节点和边的重要性行排序，以此构建出最有可能发生故障的服务组件，达到根因定位的目的。通过使用自研的故障注入工具，随机地触发多种不同类型的故障的方式，测试本故障根因定位算法的准确性，对 10 种类型的故障进行测试，可以发现，该算法具有较高的准确性与普适性，可以实现微服务架构故障根因的精准定位。虽然本算法取得了一定成绩，但还有一定的改进空间，比如由于测试范围有限，故障类型仅限于主机类和容器类。对进程类与服务类的故障测试仍然有所欠缺，后续会继续对这些类型的故障进行测试。

参考文献：

[1] 邹丹丹，丁建兵，王希栋，等.基于微服务调用链双向搜索的故障根因定位方法[J].通信技术，2022，55(11)：1515-1522.

[2] 于庆洋，白晓颖，李明杰，等.基于调用链控制流分析的大型微服务系统性能建模与异常定位[J].软件学报，2022，33(5)：1849-1864

[3] 王萌，王昊奋，李博涵，等.新一代知识图谱关键技术综述[J].计算机研究与发展，2022，59(9)：1947-1965.

[4] 庄卫金，张鸿.基于深度学习的微服务故障检测研究[J].计算机工程与应用，2022，58(16)：326-332.

[5] 殷康璘，杜庆峰.基于混沌工程的微服务韧性风险识别和分析[J].软件学报，2021，32(5)：1229-1255.

[6] 韩建兴，李卓桐，井音吉，等.联合网络拓扑与知识图谱的光网络多故障定位[J].光通信研究，2022，(4)：27-51.

[7] 周刚，闫龙川，吴小华，等.面向电力传输信息系统故障定位的知识图谱构建与应用[J].激光杂志，2021，42(12)：190-196.

[8] 王骏东，杨军，裴洋舟，等.基于知识图谱的配电网故障辅助决策研究[J].电网技术，2021，45(6)：2101-2112.

[9] 林凌云，陈青，金磊，等.基于知识图谱的变电站告警信息故障知识表示研究与应用[J].电力系统保护与控制，2022，50(12)：90-99.

[10] 叶欣智，尚磊，董旭柱，等.面向配电网故障处置的知识图谱研究与应用[J].电网技术，2022，46(10)：3739-3748.

[11] 熊学海，万春竹，杨波，等.继电保护定值模型建立及语义识别技术研究[J].自动化技术与应用，2023，42(5)：128-132.

[12] 朱延廷，陈锦荣，李响.一种基于熵值法的配网变压器负荷不平衡故障诊断[J].自动化技术与应用，2023，42(4)：17-20，40.

作者简介：刘劲（1974—），男，硕士研究生，高级工程师，研究方向：计算机技术。