

基于业务优先级的电网数字化数据调度算法

邓鸿彬¹, 杜洋², 曹泽江¹, 梁田英¹

(1. 南方电网数字电网科技(广东)有限公司, 广东 广州 510000;

2. 佰聆数据服务(广州)有限公司, 广东 广州 510000)

摘要: 针对电网数据调度操作量大、调度效率低等问题, 提出基于业务优先级的电网数字化数据调度算法。建立电网数据流架构, 分析数据特征; 分别从连接性、可靠性和稳定度三个方面评估信道质量, 根据用户优先级设定通信协议, 为数据调度提供良好的操作环境; 通过计算数据有效价值和特征, 确定业务优先级, 判断业务数量是否超出节点处理能力, 将超出部分迁移到其他节点, 减少数据拥堵。实验结果表明, 所提方法能够实现负载均衡, 减少调度时延, 提高吞吐量。

关键词: 业务优先级; 电网数字化数据; 调度算法; 有效价值; 时间特征

中图分类号: TP18 文献标识码: A 文章编号: 1003-7241(2025)03-0066-05

Digital Data Scheduling Algorithm Based on Business Priority

DENG Hong-bin¹, DU Yang², CAO Ze-jiang¹, LIANG Tian-ying¹

(1. China Southern Power Grid Digital Power Grid Technology (Guangdong) Co., Ltd., Guangzhou 510000 China;

2. Brilliant Data Analytics Inc., Guangzhou 510000 China)

Abstract: Aiming at the problems of large amount of operation and low efficiency of power grid data dispatching, a digital data dispatching algorithm based on business priority is proposed. It establishes the grid data flow architecture and analyzes the data characteristics. The channel quality is evaluated from three aspects of connectivity, reliability and stability, and the communication protocol is set according to the user priority to provide a good operating environment for data scheduling. By calculating the effective value and time characteristics of data, it determines the business priority, determines whether the number of services exceeds the processing capacity of nodes, and migrates the excess to other nodes to reduce data congestion. Experimental results show that the proposed method can achieve load balancing, reduce scheduling delay and improve throughput.

Keywords: business priority; digital data of power grid; scheduling algorithm; effective value; time characteristics

0 引言

近年来, 以新能源为主的新型电力系统得到迅猛发展, 电网设备规模大幅增长, 电力设备在运行过程中会生成海量数据, 日积月累后更是给电网数据系统带来不小压力, 这些数据分散在不同模块中, 每个模块都只负责自身的数据维护管理, 并没有形成一体化调用。随着调度系统的改进, 初步构建了多个调度数据通道, 但只能作为调度实时运行的数据池, 难以辅助管理人员完成数据调度任务。为了确保电力系统稳定运行, 提高电网数字化程度, 国家电网提出自动化、信息化和互动化的智能电网发展战略, 并对电力设备操作效率提出明确要求^[1]。因此, 减少数据调度延时、提高数据利用率迫在眉睫。

现阶段, 国内外学者大多从时间方面入手, 选取调度时间最小的数据传输路径, 当该路径上所有窗口都有任

务时, 再选取下一条时间最小的路径继续调度, 直到所有路径都被填满^[2]。国内学者引入了一些智能算法来实现数据调度。例如, 文献[3]研究了基于改进蚁群算法的数据调度策略, 构建 Petri 网的数据调度模型, 将时间、负载和能耗等因素作为约束条件, 利用改进蚁群算法对模型求解, 得到最优调度方案。文献[4]通过改进多目标优化算法实现数据调度。建立分布式数据调度多目标优化模型, 引入自适应交叉策略改进传统遗传算法, 提高算法的勘探能力, 利用交叉与变异算子丰富种群多样性, 确保快速准确地找到调度解集。文献[5]针对不确定的大规模配电网中连接到总线系统可变负载的问题, 设计了统一电能质量调节器补偿器, 以补偿总线系统中的功率流。

上述方法在用户数量较少的时候能够满足调度需求, 但随着数字化进程的加快, 电网用户数量越来越多, 调度任务类型也逐渐增多。这些方法在调度过程中并没有考虑任务的轻重缓急, 容易出现拥塞现象, 难以保证负载均衡。为此, 提出基于业务优先级的电网数字化数据

*基金项目: 2022 年广东省南方电网数研院数字电网分公司调度数据源管理及数据共享服务功能开发项目 (210000HK42220004)

收稿日期: 2023-10-13

调度算法,综合考虑调度数据的价值和时延特征,设置优先级分配策略,确保调度顺序更加合理,在减少延时的基础上满足负载均衡需求。

1 电网数字化数据流架构

1.1 数据流生成流程分析

调度算法是基于电网数字化数据流实现的,数据流产生的过程如图1所示。

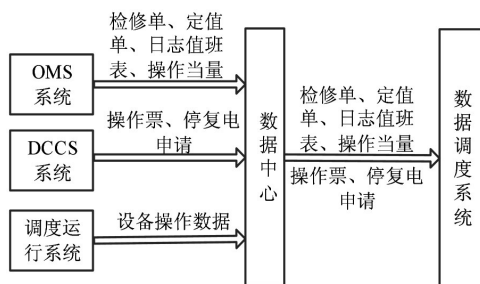


图1 数据流产生过程示意图

1.2 数据流架构特征

明确电网数据流架构特征有利于选择数据通信信道,为数据调度提供良好的环境。

(1) 一体化:这些数据能够全面覆盖不同级别主站和厂站运行需求,支持电网发、输、变、配等环节的一体化管理,协调电网业务与信息的协同与贯通。在模型统一和接口标准一致的情况下完成数据流一体化设计,便于各子系统之间的互联、互通,保证所有功能模块之间的数据资源共享。

(2) 模块化:数字化数据可满足业务功能模块化需求,电网系统提供标准的基础设施与支撑平台,确保电网运行数据以模块化的形式纳入智能系统中,协同完成作业^[6]。

(3) 智能化:电网数据可全面共享,促进各类业务灵活互动,全方位提高电网的系统作业能力,改善工作效率。

电网数据凭借一体化、模块化和智能化等特点,提高了电网业务的智能分析和决策能力,为电力系统安全运行提供有力保障。

2 信道质量评估与通信协议设置

2.1 数据调度通信信道质量评估

在电网数据调度中,信道质量决定了通信性能,影响着数据调度的时效性和稳定性,为此,有必要评估信道质量,便于调度过程中优先选择最佳信道。

在电力通信网络中,评价信道质量的标准通常包括链路稳定性、时延和带宽等。除这些因素外,还应该考虑次级用户SU对主级用户PU的影响。为了获取全方位的评估结果,将以下指标作为衡量信道质量的标准:

(1) 连接性

信道的连接性是指是否在用户端与基站间建立有效连接^[7],当连接度 L_m 等于1时,说明优先级是 q 的 SU_q 用户,其节点 n 所对应的信道 F_i 可用,当 L_m 等于0时,则说明该信道不可用。

(2) 可靠性

评估信道是否满足SU的调度需求,不但要考虑连接性,还要分析信道是否受到各类用户干扰,考察信道切换是否稳定。针对信道 F_i 而言,PU使用的概率记作 Z_i ,则数据调度过程中PU干扰SU的概率也为 Z_i ;优先级更高的用户SU使用该信道的概率表示为 G_i ,也就是此信道内SU受到级别更高的SU干扰概率。假设, SU_q 的优先级表示为 q ,其在信道 F_i 中没有受到级别更高用户干扰的概率如下:

$$1 - G_{iq} = \prod_{p>q} (1 - G_{ip}) \quad (1)$$

式中, p 与 q 的概念相同,表示不同的优先级别, G_{ip} 与 G_{iq} 表示不同的优先级别 p 与 q 下的干扰概率。

假设 E_i 代表传感错误概率, $E_i = P_e + P_l$,其中 P_e 代表错检率, P_l 是漏检率。当 $P_e \langle S_i | S_0 \rangle$ 时,表示出现了检测错误;当 $P_l \langle S_0 | S_i \rangle$ 时,说明出现漏检现象。

则针对 SU_q 节点, F_i 的可靠性计算公式如下:

$$K_{iq} = (1 - Z_i)(1 - G_{iq})(1 - E_i)(1 - T_i) \quad (2)$$

式中, T_i 代表信道切换时产生的干扰。

(3) 稳定度

稳定度表示单位时间内,信道没有出现干扰的时间与总时间之比。将信道状态模型表示为ON-OFF,该模型描述交替更新过程,ON表示被用户占用,OFF则为空闲状态。ON与OFF的时长满足参数是 λ 和 μ 的指数分布,分别利用 T_{on} 和 T_{off} 两个随机变量描述。在该模型中,将主用户PU的交替更新过程当作一个周期,则 F_i 的稳定度在该周期内的计算公式如下:

$$W_i = \frac{E(\mu_i) + E(\lambda_i)}{R} = \frac{\mu_i + \lambda_i}{\mu_i \lambda_i R} \quad (3)$$

式中, R 代表周期长度参考值。

由上述可知,针对用户 SU_q 的节点 n 而言, F_i 的质量参数 H_{inq} 的计算公式如下:

$$H_{inq} = L_m [\alpha K_{iq} + (1 - \alpha) W_i] \quad (4)$$

式中, α 属于权重系数。

结合上述指标对调度信道进行综合评价,分析信道质量,为数据调度挑选质量更高的传输信道。

2.2 信道通信协议设置

由于所提方法考虑了业务优先级,因此在调度过程中,也应分析用户的优先级。在相同时间段中,信道会同时存在主用户和次用户,如果主用户没有数据调度任务,

则将次用户接入信道。

在通信协议中,用户的退避时间区间表示为 $(0, W_j-1)$, W_j 描述退避阶段 j 的窗口尺寸。如果 $j=m$,则有 $CW_{\max}=2^m CW_{\min}$, W_{\min} 为最小竞争窗口。

为防止主、次用户间的调度任务发生碰撞,确保主用户的优先级^[8-9],则对上述退避机制做出如下改进。将主用户的退避时间区间表示为 $[0, W]$,次用户退避时间为 $[W', W_j-1]$,其中, $W_j = \left\lceil W_j \frac{M}{M+N} \right\rceil$, M 和 N 分别代表主、次用户数量。

如果出现碰撞,次用户的窗口将扩大一倍;反之,窗口大小不变。

3 基于业务优先级的电网数字化数据调度

3.1 业务优先级计算

选择最佳调度信道后,计算业务的优先级别,确保价值较高的任务被优先调度。

(1) 有效价值

因数据类型不同,其有效价值也会存在很大差异。例如,设备报警数据的价值远远高于常规数据。任务的优先级别不但与数据状态种类相关,还和其来源在电网中的重要性有关^[10]。

有效价值主要体现在如下两个方面,一是数据的关键性,表示为 g ,二是数据种类,包括常规、异常和报警数据,记作 k 。上述两方面的合称为有效价值,利用 W 表示。

任意一部分的有效价值对整体价值产生的影响如下:

$$W_i = \frac{H_{inq}}{g_i + k_i} \quad (5)$$

(2) 时间特征

不同电力子系统的的数据对于调度时间有着不同需求,因此在调度过程中,必须研究时间需求。在相同时段内,级别较高的任务需优先处理,确保时效性。

大部分电网数据均有自身的时间属性,通过下述公式表示时间特征:

$$T_i = \{s_i, t_i, w_i, d_i, a_i, b_i\} \quad (6)$$

式中, T_i 代表任务 i 的时间特征, s_i 是数据生成时间, t_i 表示预期调度时间, w_i 描述等待时间, d_i 和 a_i 分别代表截止和开始时间, b_i 为完成时间。

针对时间特征而言,有效时间与等待时间较为关键^[11],这些因素对整体造成的影响如下:

$$T'_i = \frac{w_i}{d_i - s_i} \quad (7)$$

式中, $d_i - s_i$ 即为有效时间。

3.2 动态优先级调度策略设计

将上述数据优先级计算结果作为调度依据,在其约束下设计如下动态调度策略。

当海量数据在相同时间达到时,特别是这些数据均为故障与报警数据情况下,利用下述公式完成调度:

$$P_i = \alpha W_i + \beta T'_i \quad (8)$$

式中, α 和 β 分别代表数据的有效价值与时间特征权重系数。

当产生海量数据^[12]时,在各类因素影响下,某些数据调度任务很难及时完成,但这些数据的价值依旧很重要,当到达截止时间时,这些数据不会被丢弃。对于预警数据来说,即便超过期限,优先级也最大,而其他常规数据到达截止时间后,优先级会降低,变化过程如下:

$$P'_i = P_{\max} - \frac{P_i(z_i - d_i)}{w_i}(T_{\max} + k_i) \quad (9)$$

式中, P_{\max} 代表优先级的极大值, T_{\max} 是时间特征对优先级的影响。

3.3 调度流程

基于上述调度策略,设置如下调度流程:

步骤1:在队列中加入任务,若任务数量 l_j 满足 $l_j \leq K_j$, K_j 是设定的阈值,则重新建立队列,将其他任务加入到此队列中;

步骤2:如果某节点上待调度的任务数 $l_j \geq K_j$,则要另外选择一个节点协助处理任务;

步骤3:如果节点任务数量没有达到设定阈值,在该队列中比较所有调度数据的优先值,根据优先值大小完成调度。

调度算法整体流程如图2所示。

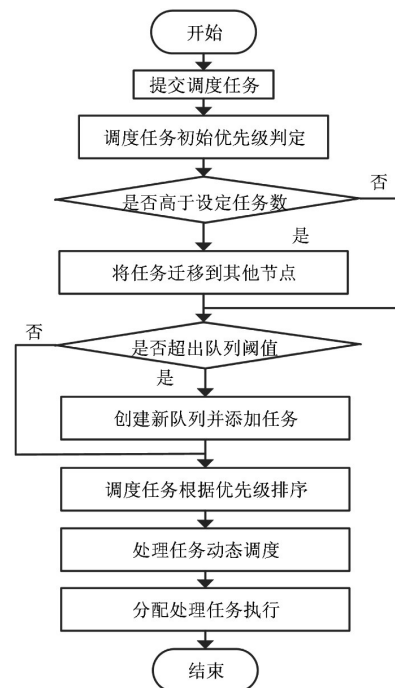


图2 调度策略流程示意图

4 实验过程与结果分析

为证明所提调度算法的实用性,将南网数据总调度系统作为实验目标,该系统在DCCS、OMS和运行数据等源端子系统基础上获取检修单、停复电、设备运行时间等数据,整体架构如图3所示。

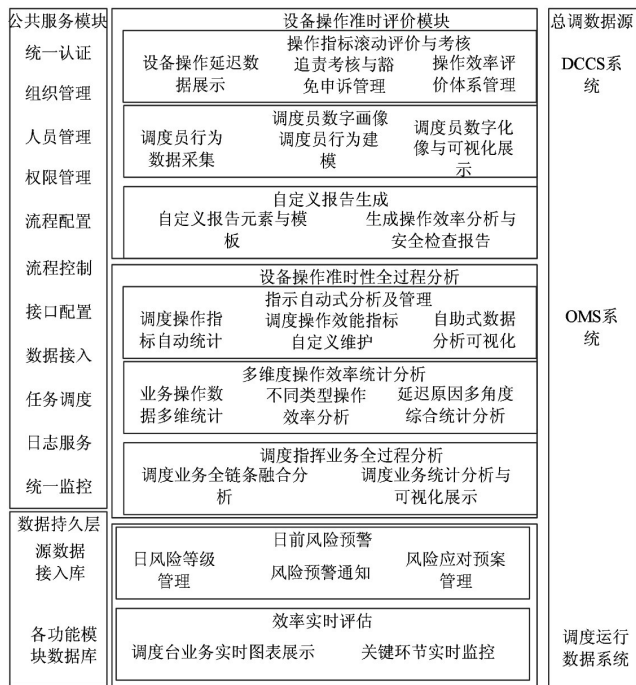


图3 实验目标整体架构图

测试网络中共包含12个源节点,网络半径设置为0.5 km,节点之间的最大距离是250 m,所有传输链路的最大带宽都可设为200 MHz,各节点都有缓存区域,节点位置如图4所示。

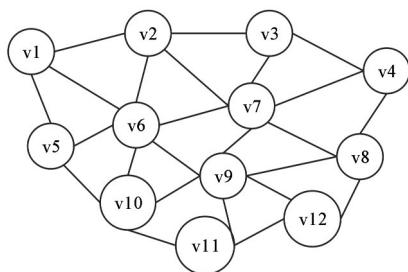


图4 测试网络中源节点位置示意图

图4中,每个节点上的数据包数值根据业务种类与任务优先级随机生成。除上述12个源节点外,还包含海量的目的节点。

(1) 节点利用率

该指标能够体现出调度过程中节点是否被有效利用,如果利用率高,表明负载均衡性良好。指标计算公式如下:

$$U_q = \frac{\lambda_q}{n_q \mu_q} \quad (10)$$

式中, λ_q 代表调度数据到达节点的实际速率, n_q 是参与调

度操作的用户数量, μ_q 表示平均服务速率。

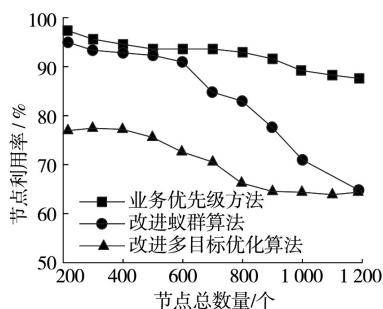


图5 不同方法节点利用率测试结果图

由图5可知,随着参与调度任务的节点数量逐渐增多,三种算法的节点利用率均出现下降趋势,但业务优先级方法始终都能保持在90%左右,没有大幅下降;改进蚁群算法在节点较少时利用率高,但是当节点数量超过600个时,利用率曲线急剧下降;改进多目标优化方法的节点利用率始终在65%—80%之间。这是因为,在业务优先级方法中,通过信道质量评估和通信协议的设定,确保节点接收到的任务量相差较小,提高利用率,保持负载均衡,避免出现任务堆积到同一个节点的情况。

(2) 调度时延

该指标表示数据从源节点传输到目的节点所用的时间,表达式如下:

$$T = n't' + t_d \quad (11)$$

式中, n' 代表源节点与目的节点之间的链路数量, t' 表示网络响应时间, t_d 是队列等待时长。

上述三种方法的测试结果如图6所示。

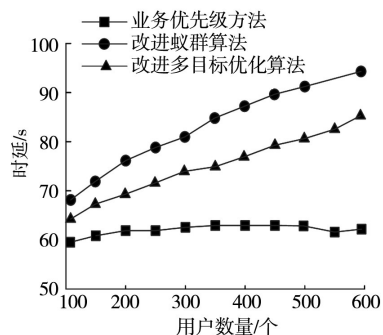


图6 不同方法的调度时延测试结果图

由图6可知,只有业务优先级方法的调度时延没有受到用户数量的影响,其他两种方法随用户数量的增加,调度所需时间都有所提高。这是因为计算了任务的时间特征与价值特征后,通过动态调度策略减少传输路径的堵塞,确保每个调度任务都能顺利完成,提高数据调度效率。

(3) 吞吐量

是指一段时间内成功调度数据的数量,表达式如下:

(下转第131页)