

工业大数据多集群通信传输节点调度方法

张晓辉¹

ZHANG Xiaohui

摘要

为提高工业大数据多集群通信传输的适应值并减小时隙，文章研究了一种工业大数据多集群通信传输节点调度方法。通过深入分析工业大数据多集群通信传输的需求，建立了相应的通信传输模型。在此基础上，对多集群通信传输链路节点进行了选择，并实现了与时隙分配的同步优化，最终完成了通信传输链路节点资源的调度。实验结果表明，新的调度方法能够有效提高通信传输节点的适应值，减少时隙总数，从而提升了工业大数据多集群通信传输的性能。

关键词

工业大数据；多集群通信传输；节点调度方法；时隙分配；同步优化

doi: 10.3969/j.issn.1672-9528.2025.03.029

0 引言

随着工业 4.0 时代与数字化转型的深入，工业大数据应用已覆盖多领域场景。在工业大数据的传输过程中，多集群通信传输节点调度方法扮演着重要“角色”。这种方法通过利用多个集群进行数据传输和处理，合理分配集群间的通信负载，优化节点调度，从而显著提升数据传输效率和处理速度。其核心理念在于能够动态调整集群间的通信带宽和节点负载，以更好地满足实时数据处理的需求。

在工业大数据环境下，多集群通信传输节点调度方法的应用展现出显著优势^[1]。首先，大幅提高了数据传输和处理效率，进而加快了工业生产和运营的响应速度。其次，通过合理分配集群间的通信负载，进一步优化了节点调度，提升了数据处理效率^[2]。此外，多集群通信传输节点调度方法还具备出色的灵活性，能够根据实际需求调整集群数量和节点配置。基于这些优势，本文将深入研究工业大数据多集群通信传输节点调度方法，旨在进一步优化这一调度方式。

1 建立工业大数据多集群通信传输模型

建立工业大数据多集群通信传输模型涉及多个复杂因素的综合考量。需要对工业大数据多集群通信传输的需求进行深入分析，包括了解数据量、数据类型、传输速率和通信协议等具体要求。同时，还需考虑集群的数量、节点配置以及通信网络的拓扑结构等因素。在建立模型之前，必须对工业大数据进行预处理，包括数据清洗、格式转换和异常值处理等操作，以确保数据的准确性和一致性^[3]。基于需求分析和

数据预处理的结果，确定模型的基本假设和参数。例如，假设各集群间的通信带宽有限且节点处理能力各异。在工业大数据多集群通信网络中，所有节点的发射功率相同，通信距离也一致。将通信距离设置为 T ，在该通信环境中存在一个节点 i 和一个节点 j ，其二者之间的距离可以表示为 $d(i, j)$ 。根据通信要求可知，当存在 $d(i, j) \leq TD$ 时，节点若处于彼此的通信范围内，则可相互传输信息。相邻节点指的是通过一跳就能到达的、在通信范围内的节点。假设所有通信节点均固定于地表，并设定一个统一的坐标系统，因此每个节点都有明确的位置坐标。将网关节点 GW 看作坐标原点，在确定坐标轴后，每个节点的两个坐标变量后，任意两个节点之间的距离可以用公式表示为：

$$d(i, j) = \sqrt{(x(i) - x(j))^2 + (y(i) - y(j))^2} \quad (1)$$

式中： $(x(i), y(i))$ 表示节点 i 的坐标； $(x(j), y(j))$ 表示节点 j 的坐标。工业大数据多集群通信网络以网关与无线路由节点两部分组成。其中，无线路由节点的数量为 N ，利用节点之间的通信距离，得到全局网络的近邻矩阵，用以刻画工业大数据多集群通信网络中节点之间的连通状态。通过 $(N+1) \times (N+1)$ 矩阵来描述

网络中的邻近节点，其中节点 1 代表网关 GW：

$$\begin{cases} d(i, j) \leq T, & i \neq j \rightarrow A(i, j) = 1 \\ d(i, j) > T, & i \neq j \rightarrow A(i, j) = 0 \\ & i = j \rightarrow A(i, j) = 0 \end{cases} \quad (2)$$

式中： $A(i, j)$ 表示邻居节点矩阵。 $A(i, j)$ 为 1 时节点 i 和节点 j 互为邻居节点。否则两个节点不相邻。在此基础上，确定相邻节点，与网关直接通信的节点称为一跳级节点：

$$\forall i \in (2, N+1), d(i, 1) \leq T \quad (3)$$

1. 中冶赛迪信息技术（重庆）有限公司 重庆 401122

可以将节点的跳数级别标记为 $L(i)$ 。首先, 将网关设定为 0 跳级别, 即 $L(1) = 0$ 。随后, 在识别出一跳级别的节点后, 在剩余节点中筛选出与一跳级别节点相邻的节点, 将其定义为二跳级别。这一过程持续进行, 通过逐级扩展, 构建出一个工业大数据多集群通信网络的多跳分级模型。这一模型不仅清晰展现了节点间的层级关系, 还可以作为工业大数据多集群通信传输的有效框架, 支持数据的高效、有序传输。

2 多集群通信传输链路节点选择与时隙分配同步优化

为进一步提升网络的可靠性并解决网络冲突问题, 本文引入了图染色算法对链路进行集中染色^[4]。在链路调度过程中遇到的冲突可以表示为:

$$e(i, j) \cap e(i', j') \neq \emptyset \quad (4)$$

图染色算法的核心思想是使用尽可能少的颜色对整个网络进行着色, 同时确保有冲突的链路使用不同的颜色进行隔离。完成整个网络的着色后, 通过为不同颜色的链路分配不同的时间间隔, 有效解决了网络中的数据传输问题, 为进一步优化奠定了基础。链路选取的目标是确保网络中的各个节点都能与网关进行通信, 从而为各种信息的顺利传递提供必要的条件。

本文专注于探讨数据向网关传输的步骤机制。在网络架构中, 除网关这一核心节点外, 其余所有节点都至少拥有 1 条起始于自身的传输连接, 这一设计确保了数据能在整个网络中顺畅流通。

前期工作中, 已将所有节点的全部传输路径整合为一个集合体系, 即 $L(1), L(2), \dots, L(N)$ 。为保障数据的稳定传输, 必须从每组路径中至少选取 1 条, 以构建覆盖全网的传输链路网。为提高计算效率, 采取简化策略, 在每组中仅选择 1 条路径来构建传输链路, 从而使得整个网络的最终传输链路数与节点数 N 相等。在此构建基础上^[5], 引入矩阵编码 $N \times \text{MAXL}$ 来处理链路选择问题, 同时, 对于时隙分配则采用了十进制编码方法。

时隙是通信网络中用于数据传输的时间段, 为优化时隙分配, 可以采用静态时隙分配和动态时隙分配相结合的方法。静态时隙分配可以预先为每个节点分配固定的时隙, 以确保其数据传输的稳定性^[6]。动态时隙分配则可以根据节点的实时需求和网络状况动态调整时隙, 以实现更高的传输效率。网络链路与时隙分配同步优化算法的详细流程概述如下:

(1) 算法初始化: 首先, 设定遗传算法的关键参数, 其参数包括初始种群的大小、交叉操作的概率、变异操作的概率以及算法的最大迭代次数, 共同决定了算法的运行效率和性能。

(2) 初始种群生成: 随机生成一个初始种群。每个个体在算法中代表一个可能的解决方案, 由 2 个独立的染色体构成链路选择染色体和时隙分配染色体。这 2 个染色体分别通过特定的编码规则来表示网络中的传输路径和通信时隙的分配情况。

(3) 染色体功能: 链路选择染色体负责在网络中确定数据传输的最佳路径, 而时隙分配染色体则负责为这些路径分配适当的通信时隙。在评估每个个体的适应度时, 算法会综合考虑这两个染色体的信息。

(4) 交叉与变异: 针对个体层面的操作进行了精细化的设计。在对 2 个染色体进行处理时, 采用了截然不同的策略。对于交叉操作, 确保两个染色体之间能够按照既定的规则进行有序的基因片段交换, 从而在保持遗传信息多样性的同时, 也促进了优良基因的组合与传承。而对于变异操作, 注重于对染色体中的某些基因进行随机的微调或改变, 以引入新的遗传变异, 增强种群的适应性和探索能力。在整个过程中, 严格把控每一步的操作, 确保交叉与变异在遗传过程中的独立性和有效性。

(5) 新种群生成与评估: 完成交叉和变异操作后, 将生成新一代的种群。然后, 算法会评估新一代种群的适应度, 以筛选出更优个体。

(6) 算法终止条件: 检查是否满足算法的终止条件, 如达到最大迭代次数或找到满足要求的解决方案。一旦满足终止条件, 算法将结束运行, 并输出最终优化后的种群。

通过一系列步骤, 该算法能够同步优化网络链路和时隙分配, 从而显著提升工业大数据多集群通信传输的效率与稳定性。

3 通信传输链路节点资源调度

工业大数据多集群通信网络的链路选择问题本质上是一个路由优化挑战, 其核心在于对网络链路质量进行全面评估, 从而筛选出高质量的链路用于数据传输^[7-9]。为此, 将工业大数据多集群通信网络抽象为图 $G(V, E)$ 的形式, 其中, V 表示节点集合, E 表示传输链路集合。在 2 个节点间传输数据时, 数据会沿着节点之间的传送链路流动, 这种从节点 i 到节点 j 的传输通路被称为链路 $e(i, j)$, 网络中的所有链路共同构成了一个链路集合 E 。

在分层多跳网络模型中, 每个节点 i 都被赋予了一个跳数等级 $LV(i)$, 这反映了它们与网关的相对距离。基于这一模型, 明确, 数据传输的两个基本原则: 一是数据只能从远离网关的节点向靠近网关的节点传输; 二是节点间的通信仅限于相邻节点。此外, 结合数据聚合机制, 节点在接收到数据后, 可以将其与自身数据进行聚合, 形成更大的数据包进

行发送^[10-12]。

在遵循上述条件的前提下，得出了链路 $e(i,j)$ 存在的必要条件，该条件可以通过一个具体的公式来表达，该公式综合考虑了节点的跳数等级、相邻关系以及数据聚合机制等因素，具体公式为：

$$\begin{cases} (L(i)-L(j))^2=1 \\ A_{ij}=1 \end{cases} \quad (5)$$

许多节点处于同一跳数层级并相互邻近，尽管相互之间存在众多连接，但这些连接并未有效传输数据至网关，反而加重了网络负担^[13]。因此，剔除这些冗余链接，尽管可能略微降低网络的连通性，但不会削弱网络的整体效能，反而能显著简化网络结构，为后续网络优化工作扫清障碍，使之更为顺畅。

4 实例应用分析

为构建一个包含3个层级、总共10个节点的工业大数据多集群通信网络，首先需要明确每个层级中的节点数量及其分布情况。假设第一层（最顶层）有1个节点，第二层有3个节点，第三层（最底层）则有6个节点。这样的设计不仅能够保证网络结构的稳定性和扩展性，还能有效支持不同层次间的数据传输需求。

在特定的网络架构下，采用全局干扰模型来模拟实际工作环境中可能存在各种噪声源对信号传输的影响。这意味着，在分析或优化该网络时，必须考虑到所有可能产生干扰的因素，包括但不限于电磁波干扰、物理障碍物阻挡等。此外，只有相邻层级之间的节点才能互为邻居节点，即跳数级别差为1。这一限制条件进一步增加了网络设计与管理的挑战性，因为其要求在规划路由路径时必须严格遵守这一规则，以确保数据能够在正确的链路上高效地流动。

表1提供了关于每个节点的具体信息，包括所在的层级位置以及每个周期内预计发送的数据量。

表1 通信网络节点信息记录表

序号	V 通信网络节点	Q 节点周期数据量 (packet)	H 通信网络节点负载 (packet/s)
(1)	N1	0	0
(2)	N2	4	1
(3)	N3	6	1
(4)	N4	3	1
(5)	N5	2	2
(6)	N6	3	2
(7)	N7	5	2
(8)	N8	4	2
(9)	N9	4	3
(10)	N10	2	3

这些详细数据对于后续进行性能评估、资源分配乃至故障诊断都至关重要。通过对这些参数的深入分析，可以更好地理解整个系统的工作状态，并据此制定出更加科学合理的维护策略和技术方案，从而提高整体运营效率与服务质量。同时，考虑到未来可能出现的新需求或者技术进步，这种模块化且具有一定灵活性的网络布局也为后续升级留下了足够的空间。

在实验设计阶段，设定初始种群的数量为60个，以确保遗传算法具有足够的多样性来探索解空间。同时，将交叉率固定为0.8，以促进种群中优秀基因的交换与组合。对于变异率 $P_m(t)$ ，采用随迭代次数自适应变化的策略，其公式为：

$$P_m(t)=[F_{\max}(t)-F_{\text{mean}}(t)]/F_{\max}(t) \quad (6)$$

式中： $F_{\max}(t)$ 和 $F_{\text{mean}}(t)$ 分别表示当前种群中最优个体的适应值和平均适应值。通过动态调整变异率，算法能够在迭代过程中更好地平衡探索与开发的能力，从而提高求解效率和质量。根据本文上述提出的调度方法，针对该工业大数据多集群通信网络，给出表2所示的分配方案。

表2 工业大数据多集群通信节点调度分配方案

序号	节点	下一跳节点	传输周期性数据量 (packet)	分配到时隙顺序
(1)	N10	N5	2	2
(2)	N9	N5	1	1
(3)	N8	N4	1	2
(4)	N7	N3	5	2
(5)	N6	N2	3	3
(6)	N5	N4	2	3
(7)	N4	N1	3	1、5
(8)	N3	N1	6	2、4
(9)	N2	N1	4	3

根据上述详尽的调度分配方案，执行通信传输任务，确保每一步都严格按照预定计划进行。在传输过程中，密切监控各项关键指标。当成功完成100次传输后，对通信网络中的各个节点进行了全面的数据分析。详细记录了每个节点的适应值，这一数据反映了节点在当前网络环境下的性能；同时，统计总时隙数，以评估网络资源的利用效率。此外，为了了解节点间的均衡性，计算节点间的差异平均值，这一指标有助于识别网络中可能存在的瓶颈或不平衡现象。最后，为了评估网络传输的实时性，测量并记录了平均延迟，即数据在网络中传输所需的平均时间。完成所有数据的收集后，将其仔细整理，并清晰地记录在表3中，以便后续的分析 and 参考。

表3 节点调度效果记录表

序号	项目	调度前		调度后	
		适应值	时隙总数	适应值	时隙总数
(1)	最大值	0.261 8	10.791 2	0.380 4	7.000
(2)	最小值	0.246 8	10.343 8	0.348 0	5.000
(3)	平均值	0.256 9	10.554 6	0.368 0	5.537 4
(4)	方差	8.165 0e-04	0.354 0	0.004 3	22.171 0

根据表3中记录的数据可以观察到,调度前节点适应值的最大值、最小值、平均值和方差均显著低于调度后。而时隙总数的最大值、最小值、平均值和方差则明显大于调度后。节点适应值反映了节点在处理数据时的性能表现,其值越大,说明节点的处理能力越强。时隙总数则代表了网络中可用的通信时隙数量,直接影响数据传输的效率和实时性。

因此,优化节点适应值和时隙总数对于提升网络传输性能至关重要。通过对比调度前后的数据,可以清晰地看到,本文提出的调度方法对工业大数据多集群通信网络传输节点进行了有效优化。该方法通过合理分配节点资源和通信时隙,显著提升了节点的处理能力和网络传输性能。具体来说,调度方法优化了节点适应值的分布,使得最大值、最小值、平均值得到显著提升。同时,该方法也优化了时隙总数的分布,减少了方差,使网络传输的实时性得到提高,这意味着数据能够更快地被处理和传输,从而提高了整个工业生产或运营的效率。综上所述,本文提出的调度方法实现了对工业大数据多集群通信网络传输节点的优化,显著提升了传输性能和网络实时性,为工业生产和运营带来了巨大的优势和效益。

5 结语

总体而言,多集群通信传输节点调度方法是工业大数据处理中一种极为关键的方法,通过合理分配集群间的通信负载和优化节点调度,能够显著提高数据传输和处理的效率。此外,该方法展现出高度的灵活性与良好的扩展潜力,能够轻松适应各种实际需求,并便于后续的调整与优化。随着工业大数据应用的不断深入,多集群通信传输节点调度方法将发挥越来越重要的作用,为工业生产和运营带来更多的便利和效益。在未来的发展过程中,期待看到更多创新性的研究和技术突破,以进一步提升这种方法的性能和应用范围,从而更好地服务于现代工业的发展需求。这不仅有助于提升企业的竞争力,也将推动整个行业向着更加智能化、高效化的方向迈进。

参考文献:

[1] 葛晓琳,王云鹏,朱肖和,等.计及差异化能量惯性的电—热—气综合能源系统日前优化调度[J].电网技术,2021,45(12):4630-4642.

[2] 陈晓娜.基于改进粒子群算法的无线网络节点安全传输链路调度方法[J].信息与电脑(理论版),2023,35(14):106-108.

[3] 王恒,段思懿,谢鑫.基于信息年龄优化的多信道无线网络调度方法[J].电子与信息学报,2022,44(2):702-709.

[4] 严宇平,洪雨天,陈守明,等.基于参数估计的配电网载波通信异常信号识别方法[J].电测与仪表,2022,59(10):123-129.

[5] 叶恒舟,郝薇,黄凤怡.AoI与样本挤压感知的通信调度算法仿真[J].计算机仿真,2022,39(6):192-196.

[6] 国强,崔玉强,王勇.无线传感器网络中基于动态簇的节点调度算法[J].吉林大学学报(工学版),2022,52(6):1466-1476.

[7] 张媛,李文娟,高鹏.基于数据挖掘的复杂光通信网络节点调度方法[J].激光杂志,2022,43(2):139-143.

[8] 罗伟华.面向5G的车载无线通信低时延信道节点调度方法[J].贵阳学院学报(自然科学版),2021,16(4):12-15.

[9] 司恩波,王晶,靳其兵,等.工业无线网络链路选择与时隙分配的同步优化[J].浙江大学学报(工学版),2016,50(6):1203-1213.

[10] 钟耀霞,苏华,翟书娟.物联网环境下复杂光纤通信传输网络节点调度[J].激光杂志,2024,45(4):196-201.

[11] 廖伟国,文明瑶.基于偏高免疫网络的低功耗节点调度方法仿真[J].计算机仿真,2023,40(11):365-369.

[12] 郭勤,曾辉,李慧玲.基于数据驱动的光纤网络冗余节点状态调度方法[J].激光杂志,2023,44(10):162-166.

[13] 徐志光,林晓康,陈励凡,等.面向电力调度数据网的节点重要性评估方法[J].计算机工程与应用,2023,59(16):330-336.

【作者简介】

张晓辉(1987—),男,山西运城人,硕士,研究方向:人工智能、工业物联网、工业大数据等。

(收稿日期:2024-11-13)