

基于改进蚁群算法的数据传输资源优化调度方法研究

黄焯真¹

HUANG Zhuozhen

摘要

为优化数据传输资源调度效果,提高数据传输时效与资源利用率,文章利用改进蚁群算法,开展了数据传输资源优化调度方法研究。首先,采集数据传输资源的基本信息,从数据传输速率、带宽两个维度,对数据传输资源作出评估;其次,建立线性回归模型,基于历史数据和趋势分析,预测未来一段时间内的数据传输需求;最后,利用改进蚁群算法,对数据传输资源进行优化调度。实验结果表明,所提出的方法应用后,能够优化数据传输资源的分配,找到从数据中心到终端用户的最优传输路径,以最小化总体传输时间,且有效地提高了数据传输资源利用率,使其均达到了96%以上,优化调度效果优势显著。

关键词

改进蚁群算法;数据;传输;资源;优化;调度

doi: 10.3969/j.issn.1672-9528.2024.11.041

0 引言

在当今信息化时代,数据传输已成为各行各业不可或缺的一部分,从云计算、物联网到大数据处理,高效、稳定的数据传输对于提升系统性能、优化资源利用及保障用户体验具有至关重要的意义。随着数据量爆炸性增长和传输需求的日益复杂,数据传输过程中面临的带宽限制、延迟问题和资源利用率低等问题愈发凸显。因此,科学合理的数据传输资源调度方法已成为提升数据传输效率和系统整体性能的关键。目前,数据传输资源调度方法的研究已经取得了显著进展。

传统的数据传输资源调度方法在实际应用中仍然存在一定的不足。例如,文献[1]提出的方法依赖于静态配置,即在网络部署初期就根据预期的数据传输需求,对传输资源进行预先分配和配置。该方法难以应对数据传输需求的动态变化,调度准确性得不到保证。文献[2]提出的方法依赖于预设的规则和算法,根据数据的优先级进行调度,在复杂多变的网络环境下,优化效果往往有限,无法充分发挥网络资源的潜力。

蚁群算法是一种模拟自然界中蚂蚁觅食行为的优化算法,由意大利学者 Marco Dorigo 于 1992 年提出。该算法通过模拟蚂蚁在寻找食物过程中释放信息素并相互协作的行为,实现了对复杂问题的优化求解,能够更好地适应复杂多变的数据传输环境^[3]。因此,本文利用改进蚁群算法,开展数据传输资源优化调度方法研究,实现对传输资源的智能管理和高效利用。

1 数据传输资源优化调度方法设计

1.1 数据传输资源评估

收集和整理现有数据传输资源的信息是一个关键步骤,它有助于全面了解当前的数据传输环境。首先,界定需要收集的信息范围,包括但不限于网络设备、传输线路、数据中心、云服务等^[4]。根据范围,列出所有需要收集信息的对象,如路由器、交换机、防火墙、传输光纤、服务器、云服务等。其次,采集数据传输资源的基本信息,数据传输资源基本信息如表1所示。

表1 数据传输资源基本信息

序号	类型	信息
(1)	设备信息	每个设备的型号、序列号、制造商、购买日期、保修状态等。
(2)	网络拓扑	网络设备、连接方式和传输线路。
(3)	配置信息	IP 地址、子网掩码、网关、路由协议、安全设置等。
(4)	性能数据	CPU 使用率、内存占用率、带宽利用率、丢包率等。

利用 Wireshark 网络性能监测工具收集实时或历史的数据传输性能数据。将收集到的数据进行整理,并分类存储,确保数据的准确性和一致性。在此基础上,分别从数据传输速率、带宽两个维度,对数据资源的传输效率作出评估,公式分别为:

$$R = \frac{1}{T} \times \log_2 N \quad (1)$$

$$B = R \times n \quad (2)$$

1. 深圳云天励飞技术股份有限公司 广东深圳 518000

式中： R 表示数据传输速率； T 表示码元脉冲的宽度； N 表示码元所取的有效离散值个数； B 表示带宽； n 表示传输位数^[5]。整理评估结果和分析，编写详细的评估报告。

通过以上流程，可以对数据传输资源进行全面、系统的评估，为后续数据传输需求分析及资源优化调度提供有力支持。

1.2 数据传输需求分析与预测

数据传输资源评估完成后，再深入了解各业务部门的数据传输需求，基于历史数据和趋势分析，预测未来一段时间内的数据传输需求，为后续资源调度提供科学依据。

确定需要分析的数据传输范围，这可能包括特定的网络区域、应用程序、用户群体或数据类型。在该分析范围内，设定分析目标，包括评估当前数据传输能力是否满足业务需求、预测未来数据传输需求以及规划扩容等^[6]。其次，将业务需求分类为不同的应用场景（如文件传输、实时数据流、备份与恢复等），并尝试量化每个场景下的数据传输需求。根据历史数据和业务需求，建立数据传输需求预测模型，通过回归分析来预测数据传输需求。模型表达式为：

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p \quad (3)$$

式中： y 表示预测的数据传输量； $(x_1, x_2, \dots, x_p)$ 表示影响数据传输量的因素（如时间、业务量等）； $(\beta_0, \beta_1, \dots, \beta_p)$ 表示模型的参数。将历史数据输入到预测模型中，执行预测计算，得到未来一段时间内的数据传输需求预测值^[7]。整理分析过程和预测结果，撰写详细的分析报告。

1.3 基于改进蚁群算法的数据传输资源优化调度

在上述数据传输需求分析与预测完成后，利用改进蚁群算法，对数据传输资源进行优化调度。基于改进蚁群算法的数据传输资源优化调度，旨在通过模拟蚁群觅食过程中的信息素机制，实现数据传输资源的智能分配与优化。初始化改进蚁群算法参数设置，初始化参数如表 2 所示。

表 2 初始化参数

序号	项目	参数	说明
(1)	蚂蚁数量	100	决定了搜索的并行性和计算量。
(2)	信息素重要度	1.5	控制信息素对蚂蚁路径选择的影响程度。
(3)	启发式因子重要度	2.0	控制启发式信息对蚂蚁路径选择的影响程度。
(4)	信息素挥发因子	0.3	控制信息素的挥发速度。
(5)	初始信息素量	0.8	确保所有路径在初始时具有相同的吸引力。

在网络的不同节点上随机部署蚂蚁，这些蚂蚁充当数据

传输的起始点或转发枢纽。客户端发起数据传输请求后，该请求被加入到待处理的任务队列中，等待经过优化后的调度执行^[8]。服务器端则针对每个任务进行细化，确定所需的数据分块数量，即分解为多个子任务进行处理^[9]。同时，为网络中各线路资源节点初始化信息素值和启发函数值，具体计算公式分别为：

$$\tau_j(t) = \frac{1}{\gamma_j}; 1 \leq j \leq m \quad (4)$$

$$\eta_j(t) = B_j; 1 \leq j \leq m \quad (5)$$

式中： $\tau_j(t)$ 表示 t 时刻对应传输线路的资源信息素； γ_j 表示传输线路 j 的等待因子； m 表示传输线路资源节点数； B_j 表示传输线路 j 的带宽； $\eta_j(t)$ 表示 t 时刻对应传输线路的启发函数。其次，针对任务队列中的每一个任务，计算其从一个节点（或线路）转移到另一个节点（或线路）的转移概率。这个转移概率的计算不仅考虑了任务的性质与需求，如数据传输量、优先级等，还融入了网络线路当前的负载状况、历史传输效率以及可能存在的瓶颈或优化空间。其公式为：

$$P_{ij}^r = \begin{cases} \frac{\tau_j^\alpha(t) \eta_j^\beta(t)}{\sum_k \tau_k^\alpha(t) \eta_k^\beta(t)}, & j, k \in L \\ 0, & j, k \notin L \end{cases} \quad (6)$$

式中： L 表示所有传输线路的集合； α 表示信息素对应的重要程度； β 表示启发函数对应的重要程度。转移概率的计算公式可能涉及多个变量和参数，需要对因素作出综合考量。通过上述计算，可以更加科学、合理地安排任务的执行顺序与传输路径，以期达到最小化等待时间、最大化资源利用率和整体传输效率的目标。各个子传输任务被看作独立的蚂蚁实体。领头蚂蚁在选择可用的线路资源时，依据计算得出的最高概率来决策，以此确定当前最为适宜的线路进行调度安排^[10]。在数据传输的进程中，一旦监测到有更优的空闲线路出现，蚂蚁会重新评估并计算，以选择新的路径，从而优化任务与线路的匹配关系。一旦任务被成功分配至特定线路，系统将根据任务 i 在该线路 j 上的实际执行时间 T ，运用特定的公式来更新线路上的信息素水平，以此反映线路的当前状态与效率，其公式为：

$$\gamma_j' = \gamma_j + T_{ij} \quad (7)$$

$$\tau_j(t)' = \frac{1 - \rho}{\gamma_j'} \quad (8)$$

式中： ρ 表示信息素的挥发程度； T_{ij} 表示第 i 个传输任务在对应线路 j 上完成任务的时间。任务传输一旦完成，随即触发各线路信息素的更新机制。上述流程将持续迭代，直至所

有客户端的同类蚂蚁（即子任务）均顺利完成数据传输资源的分配任务。随后，对各客户端的数据传输资源调度任务完成情况进行全面评估。当确认所有任务均已完成调度时，算法即告终止，并输出全局范围内的最优调度方案，作为数据传输资源优化调度的最终成果。

2 实验分析

2.1 实验准备

以 R 数据传输网络为此次研究的依托，其由多个数据中心和终端用户组成，网络中存在多个传输路径，每个路径具有不同的带宽、延迟和成本等属性。R 数据传输网络概况介绍如表 3 所示。

表 3 R 数据传输网络概况

序号	项目	概况
(1)	数据中心	3 个，分别标记为 DC1（位于坐标 (0, 0)，代表网络的一个中心节点）、DC2（位于坐标 (100, 0)，偏向网络的一侧）、DC3（位于坐标 (50, 100)，位于 DC1 和 DC2 的斜对角线上）。
(2)	终端用户	100 个，随机分布在网络的各个位置。
(3)	传输路径	每个数据中心与每个终端用户之间有多条可选的传输路径，共计 300 条路径。

随机从 100 个终端用户中选取 3 个用户，标号为 U1 ~ U3，统计 3 个数据中心到 3 个用户进行数据传输资源所需时间与资源利用率，如表 4 所示。

表 4 数据传输时间与资源利用率

标号	传输路径	传输时间 /ms	资源利用率 /%
LJ-01	DC1-U1	18	88.2
LJ-02	DC1-U2	20	89.4
LJ-03	DC1-U3	25	85.6
LJ-04	DC2-U2	22	90.4
LJ-05	DC2-U3	24	88.6
LJ-06	DC3-U3	28	89.3

通过表 4，获取数据传输资源所需的时间与资源利用率。在此基础上，通过改进蚁群算法优化数据传输资源的分配，开展实验测试。

2.2 优化调度结果分析

将上文提出的基于改进蚁群算法的优化调度方法设置为实验组，将文献 [1]、文献 [2] 提出的两种传统调度方法分别设置为对照组 1 与对照组 2。三种方法应用后，数据传输资源所需时间对比结果如图 1 所示。

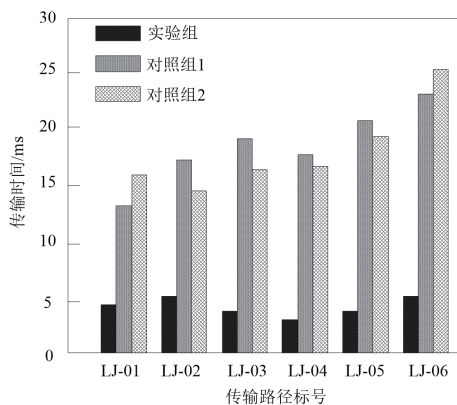


图 1 数据资源传输时间对比结果

通过图 1 的对比结果可以看出，本文提出的优化调度方法应用后，6 组传输路径完成数据资源传输所需的时间均得到了显著降低，且始终低于两种传统调度方法，能够优化数据传输资源的分配，找到从数据中心到终端用户的最优传输路径，以最小化总体传输时间。在此基础上，统计三种方法应用后的数据传输资源利用率，对比结果如图 2 所示。

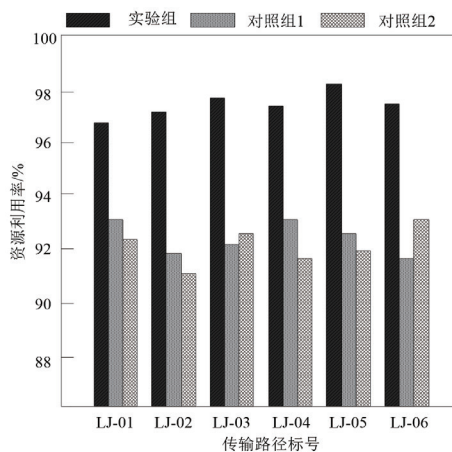


图 2 数据传输资源利用率对比结果

通过图 2 可知，本文提出的基于改进蚁群算法的数据传输资源优化调度方法应用后，有效地提高了数据传输资源利用率，均达到了 96% 以上，远高于另外两种传统调度方法，证明了该方法能够更好地利用网络资源，优化调度效果优势显著。

3 结语

综上所述，本文在深入探讨基于改进蚁群算法的数据传输资源优化调度方法的研究后，不难发现，这一方法不仅给数据传输领域带来了变革，也为未来网络资源的高效利用开辟了新路径。通过引入蚁群算法的自适应性和正反馈机制，显著提升了数据传输的效率和资源利用率，有效缓解了网络拥堵和延迟问题。

（下转第 181 页）

Press, 2018: 5956-5963.

[5] 郑诚, 曹源, 薛满意. 面向方面级情感分类的多层神经网络[J]. 计算机工程与应用, 2020, 56(19): 176-181.

[6] 李锦, 夏鸿斌, 刘渊. 基于 BERT 的双特征融合注意力的方面情感分析模型[J]. 计算机科学与探索, 2024, 18(1): 205-216.

[7] 汪红松, 李嘉展, 叶浩贤, 等. 结合语法增强与噪声削减的方面级情感分析模型[J]. 计算机工程与应用, 2024, 60(13): 152-161.

[8] 刘鲁, 刘志明. 基于机器学习的中文微博情感分类实证研究[J]. 计算机工程与应用, 2012, 48(1): 1-4.

[9] 王红斌, 沈强, 钱岩团. 融合迁移学习的中文命名实体识别[J]. 小型微型计算机系统, 2017, 38(2): 346-351.

[10] FU X H, SUN X D, WU H Y, et al. Weakly supervised topic sentiment joint model with word embeddings[J]. Knowledge-Based Systems, 2018, 147: 43-54.

[11] ALMEIDA M D, MAIA V M, TOMMASSETTI R, et al. Sentiment analysis based on a social media customised dictionary[J]. MethodsX, 2021, 8: 101449.1-101449.6.

[12] 栗雨晴, 礼欣, 韩煦, 等. 基于双语词典的微博多类情感分析方法[J]. 电子学报, 2016, 44(9): 2068-2073.

[13] HUNG C, CHEN S J. Word sense disambiguation based sentiment lexicons for sentiment classification [J]. Knowledge-Based Systems, 2016, 110(15): 224-232.

[14] YANG H, ZHANG C, LI K. PyABSA: a modularized framework for reproducible aspect-based sentiment analysis[C]// Proceedings of the 32nd ACM International Conference on Information and Knowledge Management. New York: ACM, 2023: 5117-5122.

[15] KUMAR S, GAHALAWAT M, ROY P P, et al. Exploring impact of age and gender on sentiment analysis using machine learning[J]. Electronics, 2020, 9(2): 374.1-374.14.

【作者简介】

戴薇(1996—), 通信作者(email: 1540046908@qq.com), 女, 四川泸县人, 硕士, 研究方向: 大数据分析、文本挖掘。

刘继(1974—), 男, 四川达州人, 博士, 教授, 研究方向: 数据智能分析、文本挖掘。

(收稿日期: 2024-08-05)

(上接第 176 页)

未来, 随着网络技术的不断进步和应用场景的不断拓展, 基于改进蚁群算法的数据传输资源优化调度方法有望在更多领域发挥重要作用, 推动社会信息化进程的加速发展。期待在后续的研究中, 能够进一步完善和优化这一方法, 使其更加适应复杂多变的网络环境, 为数据传输领域带来更多的创新和突破。

参考文献:

[1] 郑萌, 张雷, 梁炜. 面向 LQG 控制的 IEEE 802.11ax 网络资源调度方法[J]. 通信学报, 2024, 45(3): 81-91.

[2] 戴翠琴, 卞梦玥, 杜涛, 等. 星地融合边云协同网络下的资源调度研究[J]. 移动通信, 2024, 48(1): 25-32.

[3] 柴蓉, 陈米铃, 李锦红. 基于效用优化的星地融合网络联合用户关联及资源块调度算法[J]. 电子学报, 2023, 51(12): 3483-3495.

[4] 任勃, 朱志诚, 宋建国, 等. 低轨大型星座分布式数据接收资源调度方法[J]. 电讯技术, 2024, 64(6): 887-892.

[5] 王玉, 张海民. 多径信道下网络多媒体 P2P 资源的自适应调度方法[J]. 湖北理工学院学报, 2023, 39(4): 47-51.

[6] 汤小春, 朱紫钰, 毛安琪, 等. 数据密集作业在 GPU 集群上的调度算法研究[J]. 软件学报, 2022, 33(12): 4429-4451.

[7] 安宁, 张之栋. 基于深度强化学习的 D2D 通信网络抗干扰资源调度方法[J]. 电力信息与通信技术, 2022, 20(9): 108-114.

[8] 秦生辉, 赵卓峰, 杨中国. 面向 IoT 数据工作流的分割与调度方法[J]. 南京大学学报(自然科学版), 2022, 58(1): 153-162.

[9] 刘润滋, 吴伟华, 张文柱, 等. 基于图学习的密集空间网络传输资源调度方法[J]. 电子学报, 2021, 49(11): 2133-2137.

[10] 付培玉, 伍军, 张小飞. 基于边缘计算和稳态遗传算法的 AIoT 资源调度研究[J]. 湘潭大学学报(自然科学版), 2020, 42(5): 71-83.

【作者简介】

黄焯真(1989—), 男, 广东东莞人, 本科, 高级工程师, 研究方向: 警务、公共安全、人工智能、人脸识别。

(收稿日期: 2024-07-09)