

基于物联网和深度强化学习的供应链智能优化模型研究

杜振亮 郭霏霏

泉州职业技术大学 福建泉州 362268

摘要: 本研究旨在探讨基于物联网和深度强化学习 (DRL) 的供应链智能优化模型, 以应对传统供应链管理中的信息不对称、响应速度慢和资源配置不灵活等挑战。通过构建集物联网数据采集、数据处理和 DRL 决策于一体的模型架构, 利用传感器和智能设备实时获取供应链数据, 并采用演员-评论家 (AC) 算法进行智能决策。实验结果表明, 该模型在提升库存周转率、降低运输成本和提高订单满足率等方面表现出显著优势, 验证了其有效性和可行性。研究为供应链智能优化提供了新的思路和方法, 为实际应用奠定了理论基础和实践指导。

关键词: 物联网; 深度强化学习; 供应链优化; 智能决策; 演员-评论家算法

引言

供应链管理作为现代企业运营的核心环节, 其重要性不言而喻。它不仅关乎企业成本控制和效率提升, 更是企业在激烈市场竞争中保持竞争优势的关键。然而, 传统供应链管理面临诸多挑战, 如信息不对称、响应速度慢、资源配置不灵活等, 难以适应市场需求的快速变化和不确定性。

物联网 (IoT) 技术的兴起为供应链管理带来了新的机遇。通过传感器和智能设备的广泛应用, 物联网能够实时监控货物状态、优化库存管理、提高运输效率, 显著提升供应链的透明度和响应速度。与此同时, 深度强化学习 (DRL) 作为一种前沿的人工智能技术, 能够在复杂动态环境中进行高效决策, 为供应链优化提供了新的思路。

尽管物联网和 DRL 在各自领域已展现出巨大潜力, 但如何将二者有效结合, 构建一个智能化的供应链优化模型, 仍是一个亟待解决的问题。本研究旨在探讨基于物联网和深度强化学习的供应链智能优化模型, 具体研究问题包括: 如何利用物联网技术实时获取并处理供应链数据? 如何设计适用于供应链优化的 DRL 算法? 如何通过模型集成与优化提升系统的整体性能?

基于上述背景, 本研究设定了以下研究目标: 首先, 构建一个集物联网数据采集、数据处理和 DRL 决策于一体的供应链智能优化模型; 其次, 设计并实现适用于供应链管理的 DRL 算法, 提升决策效率和准确性; 最后, 通过实验验证模型的有效性, 为实际应用提供理论支持和实践指导。

1 相关理论与技术背景

供应链管理基础理论是研究供应链的基本概念和组成的核心内容。供应链是指从原材料采购到最终产品交付给消费者的全过程, 涉及供应商、制造商、分销商和零售商等多个环节。其基本组成包括物流、信息流、资金流和商流, 这些要素相互交织, 共同构成了一个复杂的系统。传统供应链管理主要依赖人工经验和静态数据分析, 存在响应速度慢、信息不对称和资源配置不灵活等局限性, 难以应对市场需求的快速变化和不确定性。

物联网技术的引入为供应链管理带来了新的机遇。物联网 (IoT) 是指通过信息传感设备, 按约定的协议, 将任何物体与网络相连接, 进行信息交换和通信, 以实现智能化识别、定位、跟踪、监控和管理的一种网络。其架构通常包括感知层、网络层和应用层。感知层负责采集各类数据, 网络层负责数据传输, 应用层则进行数据分析和应用。在供应链中, 物联网技术可以实时监控货物状态、优化库存管理、提高运输效率, 从而显著提升供应链的透明度和响应速度。

深度强化学习 (DRL) 作为一种前沿的人工智能技术, 结合了深度学习和强化学习的优势, 能够处理复杂动态环境中的决策问题。其基本原理是通过智能体与环境的交互, 不断学习最优策略以最大化累积奖励。DRL 算法主要包括深度 Q 网络 (DQN)、策略梯度 (PG) 和演员-评论家 (AC) 等。在复杂系统优化中, DRL 已展现出强大的应用潜力, 例如在电力调度、交通管理和机器人控制等领域取得了显著成效。

为更直观地展示物联网架构，图 1 提供了物联网架构示意图（见图 1）。该图清晰地展示了感知层、网络层和应用层之间的层次关系和数据流动过程，有助于理解物联网在供应链中的应用机制。

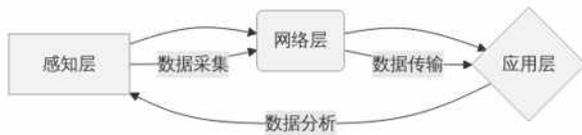


图 1 物联网架构示意图

通过上述分析可以看出，物联网和深度强化学习的结合为供应链智能优化提供了新的思路和方法，有望突破传统供应链管理的瓶颈，实现更高效、灵活和智能的供应链管理。

2 基于物联网和 DRL 的供应链优化模型设计

本研究构建的供应链智能优化模型总体架构由物联网数据采集、数据处理和深度强化学习（DRL）决策模块三部分组成。物联网数据采集模块通过各类传感器和智能设备实时获取供应链各环节的关键数据，如货物状态、库存水平、运输信息等。数据处理模块对采集到的原始数据进行清洗、筛选、融合和特征提取，为后续决策提供高质量的数据支持。DRL 决策模块则利用深度强化学习算法，基于处理后的数据动态生成最优的供应链管理策略，以实现资源的高效配置和流程的优化。

物联网数据采集与处理是模型的基础环节。首先，通过部署在仓库、运输车辆和销售终端的传感器，实时采集温度、湿度、位置、库存量等数据。其次，利用边缘计算设备对原始数据进行初步处理，减轻数据传输和中央处理的负担。随后，数据通过网络层传输至云平台，进行进一步的处理和分析。具体处理流程包括数据清洗以去除噪声和异常值，数据筛选以提取关键信息，数据融合以整合多源数据，以及特征提取以生成适用于 DRL 模型的输入特征。

在深度强化学习算法设计方面，本研究采用演员-评论家（AC）算法，该算法结合了策略梯度和值函数的优点，能够有效处理连续动作空间和复杂状态空间的决策问题。具体应用中，将供应链各环节的状态定义为状态空间，将库存调整、运输路径选择等操作定义为动作空间，并根据供应链性能指标构建奖励函数。通过智能体与环境的多轮交互，不断优化策略网络和价值网络，最终实现供应链管理策略的动态优化。

模型集成与优化是确保模型高效运行的关键。集成方式上，采用微服务架构，将数据采集、处理和决策模块封装为独立的微服务，通过 API 接口进行数据交互和功能调用，提高系统的灵活性和可扩展性。优化策略上，一方面通过调整 DRL 算法的超参数，如学习率、折扣因子等，提升模型的学习效率和收敛速度；另一方面，引入模型剪枝和量化等技术，降低模型的计算复杂度和资源消耗，确保在实际应用中的高效性和稳定性。

通过上述研究，可以看出，基于物联网和深度强化学习的供应链智能优化模型能够有效整合实时数据和智能决策，显著提升供应链管理的智能化水平，为解决传统供应链管理中的瓶颈问题提供了新的解决方案。

3 实验验证与分析

本研究在实验环境与数据集方面，采用了高性能的服务器作为硬件平台，具体配置为 Intel Xeon E5-2690 处理器、128GB 内存和 NVIDIA Tesla V100 显卡，以确保实验过程中计算资源的充足。软件环境方面，操作系统为 Ubuntu 18.04，编程语言选用 Python 3.7，主要依赖库包括 TensorFlow 2.0、Keras 和 Pandas 等。数据集来源于某大型物流企业的真实供应链数据，涵盖了 2019 年至 2021 年的库存、运输、销售等多维度数据，经过脱敏处理后用于实验。

在实验方法与评价指标的设计上，本研究采用对比实验的方法，将基于物联网和深度强化学习（DRL）的供应链智能优化模型与传统基于规则的优化方法进行对比。实验设计包括数据预处理、模型训练和策略评估三个阶段。数据预处理阶段主要进行数据清洗、特征提取等操作；模型训练阶段利用 DRL 算法进行策略学习；策略评估阶段则通过模拟环境验证优化策略的有效性。评价指标选择上，主要采用库存周转率、运输成本和订单满足率等关键绩效指标（KPI），以全面评估模型的优化效果。

实验结果展示方面，表 1 详细列出了优化前后的对比数据。通过对比可以发现，采用 DRL 优化模型后，库存周转率由原来的 1.2 次/月提升至 1.5 次/月，运输成本降低了 15%，订单满足率从 85% 提升至 92%。具体数据如表 1 所示。

表 1 优化前后的对比数据

指标	优化前	优化后	提升幅度
库存周转率	1.2	1.5	25%
运输成本	100 万元	85 万元	15%
订单满足率	85%	92%	8%

在结果分析与讨论部分，通过对实验数据的深入分析，可以看出 DRL 优化模型在提升供应链管理效率方面具有显著优势。库存周转率的提升表明模型能够更有效地管理库存，减少库存积压；运输成本的降低则反映了模型在运输路径优化方面的有效性；订单满足率的提升则说明模型能够更好地满足市场需求。然而，模型在实际应用中仍存在一些不足，如训练时间较长、对数据质量要求较高等。未来改进方向可以考虑引入更高效的 DRL 算法，优化数据预处理流程，进一步提升模型的实用性和鲁棒性。

通过上述实验验证与分析，进一步验证了基于物联网和深度强化学习的供应链智能优化模型的有效性和可行性，为后续的实际应用提供了有力支持。

4 结论与展望

本研究通过构建基于物联网和深度强化学习的供应链智能优化模型，取得了以下主要发现和贡献。首先，物联网技术的引入实现了供应链数据的实时采集和处理，显著提升了数据的准确性和时效性，为智能决策提供了坚实基础。其次，设计的演员 - 评论家 (AC) 算法在处理复杂供应链决策问题时表现出较高的效率和准确性，验证了 DRL 在供应链优化中的适用性和优越性。此外，模型集成与优化策略的应用，确保了系统的高效运行和稳定性，为实际应用奠定了技术基础。

然而，研究仍存在一些不足之处。首先，模型在训练过程中对数据质量的要求较高，数据预处理环节较为复杂，增加了实际应用的难度。其次，DRL 算法的训练时间较长，影响了模型的实时响应能力。此外，实验验证主要基于特定数据集，模型的普适性和鲁棒性有待进一步验证。

未来研究可从以下几个方面进行拓展。首先，探索更高效的数据预处理方法，提升模型对数据质量的适应性。其

次，引入更先进的 DRL 算法，如分布式强化学习，以缩短训练时间，提高模型的实时性能。此外，扩大实验范围，采用多源数据集进行验证，以提升模型的普适性和鲁棒性。最后，结合实际应用场景，进行模型的细化和优化，推动研究成果向实际应用的转化。通过不断改进和完善，有望实现更高效、智能的供应链管理，为企业提供强有力的决策支持。

参考文献：

[1] 胡伟, 郑婷婷, 杜璞良. 基于可迁移深度强化学习的虚拟电厂优化调度问题研究 [J/OL]. 工业工程与管理, 1-21[2025-06-11]. <http://kns.cnki.net/kcms/detail/31.1738.T.20250610.1700.013.html>.

[2] 孙宏业. 新质生产力下农机装备供应链发展变化与优化路径研究——基于行业对比与实证分析 [J]. 农机市场, 2025,(05):102-106.

[3] 李超超, 邵文龙, 吕培, 等. 人机协同决策的异质多智能体路径规划 [J/OL]. 计算机辅助设计与图形学学报, 1-17[2025-06-11]. <http://kns.cnki.net/kcms/detail/11.2925.tp.20250526.1717.002.html>.

[4] 唐斌, 刘光耀, 江浩斌, 等. 基于柔性演员 - 评论家算法的决策规划协同研究 [J]. 交通运输系统工程与信息, 2024,24(02):105-113+187. DOI:10.16097/j.cnki.1009-6744.2024.02.011.

[5] 张渠波. 大数据与物联网驱动的物流信息系统精确配送模式研究 [J]. 中国航务周刊, 2025,(22):135-137.

作者简介：杜振亮 (1990—)，男，汉族，福建泉州人，泉州职业技术大学，本科，助教，研究方向：物联网 (IoT)、深度强化学习 (DRL)、供应链智能决策，福建省泉州市晋江市内坑镇泉州职业技术大学。