

控制与决策

Control and Decision

考虑商品间适应度的在线零售商分仓选品优化研究

许瑞, 丁子千, 肖巍

引用本文:

许瑞, 丁子千, 肖巍. 考虑商品间适应度的在线零售商分仓选品优化研究[J]. *控制与决策*, 2025, 40(6): 1959–1968.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2024.1067>

您可能感兴趣的其他文章

Articles you may be interested in

[基于粒子群算法的满载需求可拆分车辆路径规划](#)

Split vehicle route planning with full load demand based on particle swarm optimization

控制与决策. 2021, 36(6): 1397–1406 <https://doi.org/10.13195/j.kzyjc.2019.1323>

[基于分解的多目标多因子进化算法](#)

A multiobjective multifactorial evolutionary algorithm based on decomposition

控制与决策. 2021, 36(3): 637–644 <https://doi.org/10.13195/j.kzyjc.2019.0525>

[基于预防维护的单机调度问题](#)

[Single-machine scheduling problem with preventative maintenance activities](#)

控制与决策. 2021, 36(2): 395–402 <https://doi.org/10.13195/j.kzyjc.2019.0626>

[现货市场补充作用下基于总量折扣的运输服务采购问题研究](#)

Transportation service procurement based on total discount under complementary effect of spot market

控制与决策. 2021, 36(11): 2794–2802 <https://doi.org/10.13195/j.kzyjc.2020.0274>

[基于互信息操作变量曲线参数化的间歇过程批内修正优化](#)

Intra-batch correction optimization of batch process with manipulated variable trajectory parameterization based on mutual information

控制与决策. 2021, 36(1): 234–240 <https://doi.org/10.13195/j.kzyjc.2019.0825>

考虑商品间适应度的在线零售商分仓选品优化研究

许瑞, 丁子千, 肖巍[†]

(河海大学商学院, 南京 211100)

摘要: 随着在线零售业的快速发展, 线上订单数量日益庞大, 分仓选品对于高效服务线上订单愈发重要. 现有研究多从降低拆单率的角度优化选品方案, 忽略了拆单造成的额外运输距离差异. 鉴于此, 首先, 构建以最小化拆单率和运输距离为目标的分仓选品问题模型, 提出综合衡量订单商品分布与客户地理分布的商品间适应度指标; 然后, 结合谱聚类方法设计基于固定-优化框架的两阶段分仓选品算法. 数值实验表明: 与直接求解分仓选品模型相比, 所提出算法的固定阶段能够有效缩小搜索空间, 在保证求解质量的前提下能够提升求解效率; 与现有文献算法相比, 所提出算法能够显著降低运输距离和拆单率, 为企业优化分仓选品方案提供决策支持.

关键词: 分仓选品; 商品间适应度; 固定-优化算法; 谱聚类; 拆单

中图分类号: C934 文献标志码: A

DOI: 10.13195/j.kzyjc.2024.1067

引用格式: 许瑞, 丁子千, 肖巍. 考虑商品间适应度的在线零售商分仓选品优化研究 [J]. 控制与决策, 2025, 40(6): 1959-1968.

Optimization of warehouse selection and product allocation for online retailers with consideration of inter-product fitness

XU Rui, DING Zi-qian, XIAO Wei[†]

(School of Business, Hohai University, Nanjing 211100, China)

Abstract: With the rapid growth of online retail, the volume of online orders has surged, making efficient inventory selection across multiple warehouses increasingly crucial. Existing studies predominantly focus on reducing the order splitting rate in inventory selection strategies, often neglecting the impact of increased transportation distances due to order splitting. This paper addresses this gap by constructing a warehouse inventory selection problem model that aims to minimize both the order splitting rate and transportation distance. We introduce a fitness index between products that comprehensively evaluates the distribution of order items and customer geography, providing a more holistic approach to inventory optimization. Moreover, we design a two-stage warehouse inventory selection algorithm based on a fix-and-optimize framework using spectral clustering methods. Numerical experiments demonstrate that, compared to directly solving the warehouse inventory selection model, the fix-stage of the algorithm effectively reduces the search space, enhancing solution efficiency while ensuring solution quality. Compared to existing algorithms in the literature, the proposed approach can significantly reduce transportation distance and order splitting rate, providing valuable decision support for enterprises in warehouse inventory selection.

Keywords: product allocation optimization; inter-product fitness; fix-and-optimize algorithm; spectral clustering; order splitting

0 引言

在线零售业现已成为当前市场主流零售方式之一. 2023 年, 我国网上零售额达 15.43 万亿元, 较上年增长了 11.0%. 相对于传统零售方式, 在线零售方式存在服务客户范围广、线上订单数量庞大、订单时效性要求高以及单个订单包含多种商品 (一单多

品) 等特点^[1]. 当单一仓库库存不足时, 企业需要联合多个区域的仓库共同履行订单, 如在“双十一”期间, 京东和天猫超市等在目标区域仓库的库存无法满足订单需求时, 订单会被拆分为多个发货批次, 从不同区域仓库配送至客户^[2]. 传统仓库选品模式在应对一单多品订单时面临的挑战主要表现为备货不合理,

收稿日期: 2024-09-04; 录用日期: 2024-12-30.

基金项目: 国家自然科学基金项目 (62106098, 42071278, 62272210); 中央高校基本科研业务费专项资金项目 (B240207057, 423180); 广东省类脑智能计算重点实验室项目 (2020B121201001).

责任编委: 刘士新.

[†]通信作者. E-mail: wei.xiao@hhu.edu.cn.

导致订单拆分情况频发,拆单直接影响消费者的购物体验.对于在线零售商而言,拆单情况存在差异,当订单频繁被拆分至距离较远的仓库发货时,额外运输距离造成的运输成本升高不容忽视.因此,减少订单拆分以及降低拆单造成的运输距离,对于在线零售商降本增收具有重要的战略意义.

如何缓解订单拆分情况一直是业界关注的重点,针对拆单相关的优化主要集中于4个方面:分仓选品优化^[3-7]、订单分配优化^[8]、订单分拣以及合并优化^[9-11]、配送路径优化^[12-13].此外,也有部分研究对多个环节联合优化^[14].上述4个步骤在订单履行过程中紧密相关.其中:配送路径的决策受到订单分配以及合并情况的影响,而订单分配情况又取决于仓库商品存储情况,从根本上影响了后续订单分配以及配送路径优化效率,若是基于不合理的商品分仓存储情况,则后续再多的订单分配优化也是徒劳.因此,研究分仓选品优化问题对于降低拆单率、优化配送路径以及降低运输距离均具有重要的推动作用.

当前针对分仓选品优化问题的研究按照商品存储方式可划分为两种模式:SKU (stock keeping unit) 单点存放^[3-4](一种商品仅存放于一个仓库),SKU 多点存放(一种商品存放于多个仓库).SKU 多点存放模式可基于商品间关联性进行选品,更适用于一单多品订单的场景,相对于单点存放模式,其在零售业应用更广泛.Catalán 等^[5]提出了考虑热销品的贪婪算法,根据历史需求信息,统计各SKU 需求量并计算共现矩阵,选出热销SKU 存放于所有仓库,剩余非热销SKU 则根据共现矩阵贪婪分配存放;李建斌等^[6]考虑数量拆单和品类拆单情况建立了数学模型,并改进热销品算法得到了环形算法,不规定非热销品单点存储而是进行环形分配,相对于热销品算法能够进一步降低拆单率.

目前,多点存放模式下的相关研究大多以最小化拆单率或拆单总量为目标,其局限性在于多种拆单情况均可能造成相同拆单率,但是产生的运输成本却大相径庭.针对在线零售业的部分研究基于仓库在同一城市的假设忽视运输距离差异^[6],而在线零售企业服务于全国范围内的客户,相同拆单率下,选择不同仓库造成的额外运输距离差异不容忽视,将对整体运输距离产生显著影响.Acimovic 等^[8]研究了订单分配问题,以最小化当前和未来运输成本的和为目标选择最佳发货仓库,旨在降低订单履约成本,对于一单多品情况,该研究没有专门针对拆单率进行优化,而是通过优化分批发货和联合运输来减少拆单带来的额外成本,间接减少了拆单;李建斌等^[7]

则聚焦于分仓选品阶段,以最小化运输成本为目标建立了数学模型,针对制造业背景下的生产物流问题,设计了基于最小增量成本的贪婪算法进行选品优化.在分仓选品阶段考虑运输距离的难点在于如何将历史订单包含的商品信息与客户位置分布相结合,通过综合评估分仓选品方案优劣进一步实现运输过程调优.

分仓选品问题的本质是将商品分组并分别存放至合适的仓库,而聚类算法即根据数据相似性对数据进行分组,可通过识别商品数据的内在结构,有效帮助实现商品分组.Chuang 等^[15]针对仓储中的物品分配问题,提出了一个基于物品关联的聚类分配模型,通过聚类将高频共现的物品分配至相邻区域,显著减少了订单拣选距离,后续又在上述模型的基础上提出了双层优化结构^[16];Zhu 等^[3]将聚类算法的基本原理应用于多个仓库间的产品类别分配,构造了一种K-links 方法来优化分仓选品以降低拆单率.此外,聚类算法也被应用于订单拣选^[17]和订单配送优化领域^[18].本文定义商品间适应度指标来量化商品间关联程度,并采用聚类算法对商品进行初步分仓.

本文的主要内容如下:1) 基于在线零售业仓库跨域分布和商品多点存放的现实情况,突破了传统选品方案仅考虑商品共现性的局限,从商品共现程度和目标客户群体分布差异两个角度出发,设计商品共现矩阵和商品变异矩阵,并将二者相结合提出商品间适应度指标,以实现拆单率与运输距离的协同优化.2) 设计基于固定-优化框架的双阶段分仓选品算法,在固定阶段结合谱聚类初步分配SKU 并固定部分决策变量,在优化阶段动态调整剩余SKU 分配,所提出算法能够有效缩小搜索空间,在保持解质量的同时显著提升求解效率.

1 问题建模

1.1 问题描述和基本假设

分仓选品问题可具体描述如下:在客户订单尚未下达时,在线零售商需要根据客户需求分布将商品分配至各仓库存储,决策每种商品应存放在哪一个或哪几个仓库.分仓选品是订单履行的先导阶段,订单下达后,商家将基于分仓选品结果择优匹配满足该订单的仓库,不同选品方案下,订单分配可能的情况如下:基于理想选品方案,订单将交由与客户距离最近的某一仓库满足,以实现低运输距离以及0拆单的最优配送;选品方案欠佳时,订单可能由远距离的某一仓库满足,或被拆分至距离较近的仓库

配送,将导致客户满意度降低,并产生额外运输距离。由此可见,协调多个仓库的商品存储至关重要,需要尽可能减少订单分配过程中的拆单情况,使得客户一次完成收货,进而降低订单运输距离来节省运输成本。

基于前人的一般假设^[6-7],结合本文问题的特性,做出假设如下:1)仓库位置以及库容在未来一段时间内保持不变;2)在线零售商面临的市场环境在未来一段时间内保持稳定;3)每个仓库只能存储有限SKU品类数,没有仓库能够容纳所有SKU品类,但是,所有仓库容量总和能够容纳所有SKU品类;4)本文不考虑数量拆单情况,不对各品类SKU做数量分配。

1.2 模型建立

建模所需的基本符号及其含义如表1所示。

表1 基本符号及其含义

集合和参数	含义
O	表示订单集合,编号为 o
I	表示SKU品类集合,编号为 i
M	表示仓库集合,编号为 m
N	表示客户集合,编号为 n
O_n	表示客户 n 的订单集合, $O_n \subseteq O$
c_m	表示仓库 m 最多可存放的品类数
d_{mn}	表示仓库 m 与客户 n 间的距离
γ_{io}	表示订单 o 中是否包含SKU i ,是为1;否则为0
决策变量	含义
x_{im}	0-1变量,表示SKU i 是否由仓库 m 存放,是为1;否则为0
y_{omn}	0-1变量,表示客户 n 的订单 o 是否由仓库 m 配送,是为1;否则为0
z_{oimn}	0-1变量,表示客户 n 的订单中的SKU i 是否由仓库 m 配送,是为1;否则为0

本文分别以最小化运输距离和最小化拆单率为优化目标,建立如下混合整数规划模型:

$$\min \sum_{m \in M} \sum_{n \in N} \sum_{o \in O_n} y_{omn} d_{mn}, \quad (1)$$

$$\min \frac{\sum_{m \in M} \sum_{n \in N} \sum_{o \in O_n} y_{omn}}{|O|}. \quad (2)$$

$$\text{s.t.} \sum_{m \in M} x_{im} \geq 1, \forall i \in I; \quad (3)$$

$$\sum_{i \in I} x_{im} \leq c_m, \forall m \in M; \quad (4)$$

$$z_{oimn} \leq x_{im} \gamma_{io}, \forall i \in I, m \in M, o \in O_n, n \in N; \quad (5)$$

$$\sum_{m \in M} z_{oimn} \geq \gamma_{io}, \forall i \in I, o \in O_n, n \in N; \quad (6)$$

$$y_{omn} \leq \sum_{i \in I} z_{oimn}, \forall m \in M, o \in O_n, n \in N; \quad (7)$$

$$y_{omn} \geq \frac{\sum_{i \in I} z_{oimn}}{|I|}, \forall m \in M, o \in O_n, n \in N; \quad (8)$$

$$x_{im}, y_{omn}, z_{oimn} \in \{0, 1\}, \forall i \in I, m \in M, o \in O_n, n \in N. \quad (9)$$

其中:目标函数(1)为最小化订单配送的总运输距离;目标函数(2)为最小化拆单率;式(3)表示对于任意品类的SKU i ,至少分配给一个仓库存放;式(4)表示仓库 m 存放的SKU品类数不得超过其最大可存放品类数;式(5)为判定约束,当且仅当仓库 m 存放SKU i 时该仓库才可配送;式(6)保证所有订单需求均被满足;式(7)和(8)用于建立决策变量间关系;式(9)为决策变量取值范围约束。

2 商品间适应度指标

分仓选品问题需要将多个类别的商品分配至多个仓库存放,以期实现最小化拆单率和运输距离,其中如何量化商品间的适应度,即衡量两两商品是否适合存放于同一仓库是分仓选品算法有效性的关键。现有研究大多基于历史订单计算商品间共现矩阵,以此作为衡量适应度的唯一标准^[3],但是,本质上均是从订单角度挖掘商品间关联性来设计适应度指标,缺乏对商品目标用户群体间距离差异的衡量。

针对这一局限性,本文从历史订单商品分布和客户地理分布两个角度综合设计商品间适应度指标(CF)。其中:在订单层面采用文献中常见的商品间共现性指标来衡量商品间关联性,其值越高,两两商品在历史订单中越频繁地被同时购买;在客户层面引入商品间变异性指标来度量目标客户群体间的距离差异,其值越高,客户群体分布的距离差异越大。综合以上因素,基于高共现-低变异的原则构建衡量商品间适应度指标,具体如下所示:

$$CF_{i \times i} = CCM_{i \times i} - CVM_{i \times i}, \quad (10)$$

其中 $CCM_{i \times i}$ 和 $CVM_{i \times i}$ 分别为商品间共现矩阵和商品间变异矩阵。

1)商品间共现矩阵(CCM): $CCM_{i \times i} = R_{o \times i}^T \times R_{o \times i}$,表示两两商品在历史订单中同时出现的频率,其中 $R_{o \times i}$ 表示订单 $o \in O$ 对SKU $i \in I$ 的需求矩阵。

2) 商品间变异矩阵 (CVM): 两两商品目标客户群体间距离差异, 其中 μ_i 为SKU $_i$ 目标客户群体质心的位置坐标, 如下所示:

$$CVM_{i \times i} = \begin{bmatrix} \|\mu_1 - \mu_1\| & \dots & \|\mu_1 - \mu_i\| \\ \vdots & \ddots & \vdots \\ \|\mu_i - \mu_1\| & \dots & \|\mu_i - \mu_i\| \end{bmatrix}, \quad (11)$$

$$\mu_i = \frac{(T \cdot L_{n \times 2})_i}{\sum_{n \in N} T_{i \times n}}. \quad (12)$$

这里: $L_{n \times 2}$ 为客户 $n \in N$ 的坐标; $T_{i \times n}$ 为客户 $n \in N$ 购买SKU $_i \in I$ 的频次, 将每个SKU的购买频次与客户坐标相结合, 可得到该SKU目标客户群体的加权平均位置; CVM矩阵中的每个元素 $\|\mu_i - \mu_j\|$ 则表示SKU $_i$ 与SKU $_j$ 的目标客户群体质心位置间的距离差异.

值得注意的是, CCM $_{i \times i}$ 取值范围取决于历史订单数, CVM $_{i \times i}$ 取值范围取决于客户位置以及客户间的距离. 为了消除量纲影响, 使得二者在数值上具有可比性, 本文对CCM $_{i \times i}$ 和CVM $_{i \times i}$ 进行归一化处理.

通过对分仓选品目标要素的分析, 拆单指标关联历史订单的商品共现程度, 距离指标关联目标客户群体的地理分布差异, 商品适应度指标综合以上两个要素, 对分仓选品目标进行综合评估, 为后续算法优化设计提供基础.

显然, 基于商品间适应度进行分仓选品时, 期望存放在同一仓库的商品间适应度尽可能高, 而存放在不同仓库的商品适应度尽可能低^[16, 19]. 本文进一步定义变量cohesion为仓库的内聚度, 变量coupling为仓库间的耦合度, 对存放于同仓库与不同仓库的商品间适应度进行衡量, 具体表示如下所示:

$$cohesion = \frac{1}{2} \sum_{m \in M} \sum_{i \in I} \sum_{i' \in I, i' \neq i} F_{i \times i'} \times x_{im} \times x_{i'm}, \quad (13)$$

$$coupling = \sum_{m \in M} \sum_{m \in M, m \neq m'} \sum_{i \in I} \sum_{i' \in I, i' \neq i} F_{i \times i'} \times x_{im} \times x_{i'm'}. \quad (14)$$

某一仓库的内聚度是指存放在该仓库的所有商品间的适应度的和, cohesion为所有仓库的总内聚度; 仓库 m_1 与仓库 m_2 间的耦合度是指存放在仓库 m_1 的商品与存放在仓库 m_2 的商品间的适应度的和, coupling为所有仓库间的总耦合度. 此外, 由于适应度矩阵 $F_{i \times i}$ 为对称矩阵, 在计算内聚耦合度时需要乘以1/2, 以避免商品间适应度被重复计算.

定理 1 当仓库数量和商品品类数不变时, 将SKU $_i$ 从仓库 m_1 移动至仓库 m_2 , 仓库内聚度与耦合度的和恒为定值, 即最大化仓库内聚度与最小化仓库间耦合度等价.

证明 当SKU $_i$ 从仓库 m_1 移动至仓库 m_2 时, 仓库内聚度和仓库间耦合度的变化如下: 仓库内聚度增加 $\sum_{i \neq i'} F_{i'i} x_{im_2}$, 仓库间耦合度减少 $\sum_{i \neq i'} F_{i'i} x_{im_2}$. 综上, 仓库内聚度变化 $(\sum_{i \neq i'} F_{i'i} x_{im_2} - \sum_{i \neq i'} F_{i'i} x_{im_1})$, 仓库间耦合度变化 $(\sum_{i \neq i'} F_{i'i} x_{im_1} - \sum_{i \neq i'} F_{i'i} x_{im_2})$, 二者的和不变. □

3 分仓选品算法设计

本节设计基于商品间适应度的分仓选品固定-优化 (FFO) 算法. 固定-优化的核心思想是选择模型中部分变量固定取值^[20-21], 剩余变量形成子问题进一步优化, 能够缩小问题规模, 提升求解效率. FFO算法总体由两个模块组成: 模块1为商品间适应度指

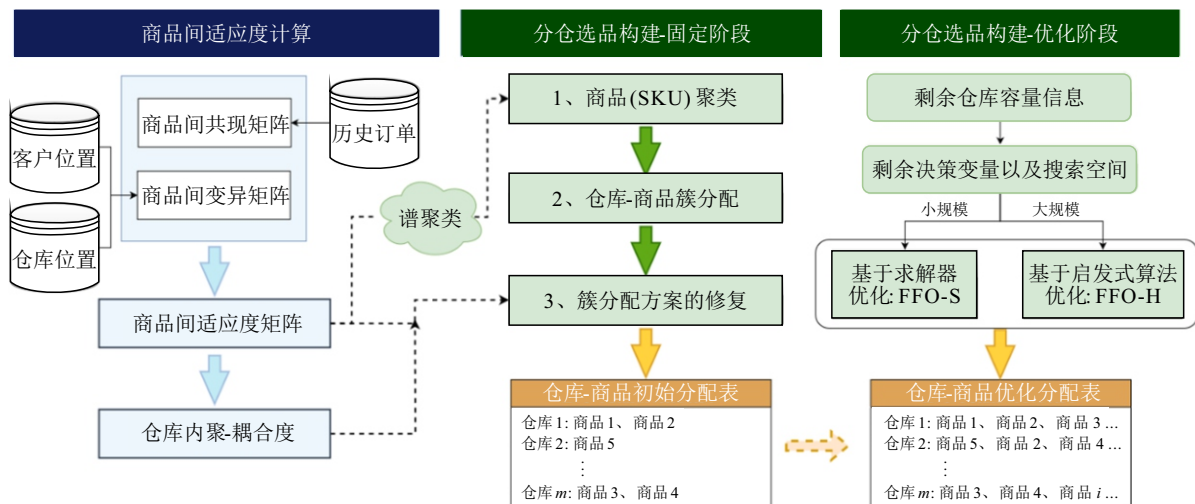


图1 FFO 算法结构

标计算:通过对历史订单以及仓库客户间距离的分析,挖掘商品间共现矩阵以及商品间变异矩阵两类经验,以获得商品间适应度矩阵(详见第2节)。模块2为分仓选品方案构建:通过固定-优化的方式进行两阶段选品。阶段1为固定阶段,基于谱聚类对商品SKU进行单点分配,并固定选品决策;阶段2为优化阶段,基于仓库品类约束对商品SKU进行多点再分配。FFO算法结构如图1所示。

3.1 固定阶段-分仓选品初始方案构建

FFO分仓选品算法的固定阶段通过对各仓库分配合适的商品SKU来确保固化的SKU分配方案属于高质量解集。分仓选品即选取适应度高的商品存放在同一仓库,而适应度低的商品尽可能分配至不同仓库,该过程本质上可视为对商品节点的聚类分配。本文采用谱聚类方法对SKU进行初始聚类分配,并固定SKU初始分配结果。

谱聚类以谱图理论为基础,将样本数据视为空间中的点,其连边权重大小表示样本间的关联程度。该方法利用节点间连边的权重值对节点进行聚类,将数据的聚类转化为图的最优划分问题。所提出商品间适应度将SKU相互联系形成网络,具体可描述为无向图 $G=(V, E)$ 。其中: $V=\{0, 1, \dots, n\}$ 为各SKU节点; E 为SKU间的连边,连边权重即为商品间适应度,对商品聚类即对无向图 G 中的节点进行聚类。规范割集准则方法适用于具有加权连边的图结构,能够有效平衡簇间的分离度与簇内的紧密度,本文使用规范割集准则切图方法将SKU划分为簇,旨在最小化两个簇之间的连边权重和,使得划分在同一簇内的商品内聚度更大,不同簇间的商品耦合度更小,并将划分后的SKU簇和仓库进行分配。

1) SKU聚类。本文将商品间适应度矩阵 F 作为谱聚类方法的自定义相似矩阵,仓库数量 $|M|$ 为划分簇的数量,谱聚类得到所有品类SKU被划分为 $|M|$ 个簇的划分结果。

2) 仓库-SKU簇分配。谱聚类将SKU划分为 $|M|$ 个簇,但是,仍然需要考虑SKU与仓库间的关联,对每个SKU簇和仓库进行匹配。本文从不同SKU簇的目标客户群体角度出发,对于每个SKU簇,计算簇中所有SKU目标客户群体质心的平均位置,将簇中所有SKU分配至距离该平均值位置最近的仓库。

3) 簇分配方案修复。由于每个仓库具有可容纳最大品类数限制,初步分配结果可能会超出仓库容量,导致解不可行。本文基于仓库内耦合度,采用贪婪的方式对可能出现的不可行解进行修复:遍历超

容量仓库的SKU,删除其中使得仓库内聚度降低程度最小的SKU;将被删除SKU插入至其他有剩余容量的仓库,使得仓库耦合度增加程度最小;重复上述步骤直至分配结果均不超出仓库可容纳最大品类数限制。

基于上述初始分仓选品方案构建步骤,设计具体流程如算法1所示。

算法1 基于谱聚类的初始方案构建。

输入: SKU列表 I , 仓库列表 M , SKU间适应度矩阵 F , SKU目标客户群体质心 μ , 每个仓库可容纳品类数限制 c_m ;

输出: 所有SKU至仓库的SKU-仓库分配表 $Table_m$ 。

Begin

1. 对于所有SKU聚类,得到 $|M|$ 个SKU簇 $Cluster_m$
2. foreach SKU簇 $Cluster_m$ do
3. 计算所有 $SKU_i \in Cluster_m$ 目标客户群体质心的平均位置 $\bar{\mu}_m$
4. 将SKU簇 $Cluster_m$ 中的所有SKU分配至距离 $\bar{\mu}_m$ 最近的仓库
5. 计算得到初始SKU-仓库分配表 $Table_m$
6. foreach 仓库 m do
7. if 仓库 m 当前容纳SKU品类数 $skuNum_m > c_m$ then
8. while $skuNum_m > c_m$ do
9. foreach $SKU_{i'}$ do
10. 计算将 $SKU_{i'}$ 从仓库 m 删除,造成仓库 m 内聚度的变化 $\Delta cohesion_{i',m}$
11. 选择 $SKU_{i'} = \arg \min(\Delta cohesion_{i',m'})$ 从仓库 m 删除
12. $skuNum_m = skuNum_m - 1$
13. foreach 有剩余容量的仓库 m' do
14. 计算将 $SKU_{i'}$ 分配至仓库 m' ,造成仓库间耦合程度的变化 $\Delta coupling_{i',m'}$
15. 将被删除的 i' 分配至仓库
 $m' = \arg \min(\Delta coupling_{i',m'})$
16. 更新SKU-仓库分配表 $Table_m$
17. else
18. continue
19. end

3.2 优化阶段-多点存放SKU再分配

基于谱聚类的构建过程对全部SKU进行一次分配得到可行解,但是部分仓库可能仍有选品空间,对其进行合理利用有助于进一步优化选品方案。FFO算法首先固定谱聚类选品方案确定的决策变量,然后在优化环节针对剩余仓容进行多点SKU选品分配。

FFO算法效率的关键在于优化环节,仓库容量越小或SKU种类数越少,剩余搜索空间越小。为了

应对不同搜索空间范围的影响,本文分别提出了基于求解器的 FFO-S (FFO-Solver) 和基于启发式算法的 FFO-H (FFO-Heuristic) 两种算法,并在实验部分对比分析二者的优势和适用性.

FFO-S 算法通过固定阶段已分配至仓库 m' 的 SKU 品规 i' , 添加线性约束 $x_{i'm'} = 1$ 形成新的整数规划模型, 使用 Gurobi 求解器进行求解. 由于本文问题考虑了两个优化目标, 实验部分分别以最小化拆单率和最小化运输距离为目标进行求解 (详见后文第 4.3.1 节). FFO-H 算法通过贪婪启发式策略进行选品优化, 即对于每个仍有剩余容量的仓库 m , 选取与当前存放于该仓库的 SKU 适应度最高的 $c_m - \sum_{i \in I} x_{im}$ 个 SKU 补充存放于该仓库, 以实现仓容的负载均衡.

定理 2 设 s_m 为仓库 m 已分配并固定的 SKU 数, 集合 $M' = \{m | c_m > s_m\}$ 为初始分配后仍有剩余容量的仓库集合, FFO 算法可缩减搜索空间为 $2^{|I| \times |M|} - 2^{\sum_{m \in M'} (|I| - s_m)}$.

4 实验分析

实验环境为 Windows 10 操作系统, 内存为 16 GB, 处理器为 Intel(R) Core(TM) i7-8550U CPU@1.80 GHz. 算法在 Anaconda 中使用 Python 3.7 编写实现, 采用 Gurobi 10.0.2 求解器实现模型求解.

4.1 数据来源

通过分析国内某在线零售企业一年内的订单数据发现: 70% 左右的订单为多品类 SKU 订单; 每个订单包含的品类数通常在 2~6 个, 其中占比 82.5% 的订单包含 2~4 种商品, 剩余 17.5% 订单包含商品品类大于 4 种; 总订单量与 SKU 数量比约为 5:1. 按照上述分布经验生成随机算例: 客户订单数取值分布在 [100, 12 000] 中; 每个订单所包含的 SKU 品类在数 [2, 6] 中, 基于幂律分布生成; 订单包含的 SKU 品类随机生成; 总仓库按照企业现状设为 4 个; 客户和仓库的位置随机分布在 [-100, 100] 区域内. 后续实验部分采用 SKU 品类数 I 和订单数 O 表示算例规模, 如 I20O100 表示 SKU 品类数为 20, 历史订单数为 100 的算例.

4.2 商品间适应度指标有效性分析

商品间适应度指标由共现矩阵和变异矩阵组成, 考虑商品在历史订单中共现频率的基础上, 加入考虑目标客户群体差异的变异矩阵. 为了验证商品间适应度指标的有效性, 即变异矩阵的加入对算法效果的影响, 在 4 种规模的实例上分别进行 10 次实验,

将仅保留共现矩阵的 FFO-H 算法与有变异矩阵的

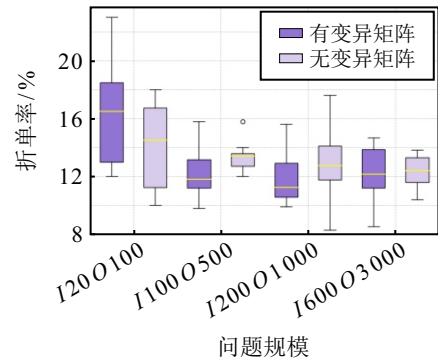
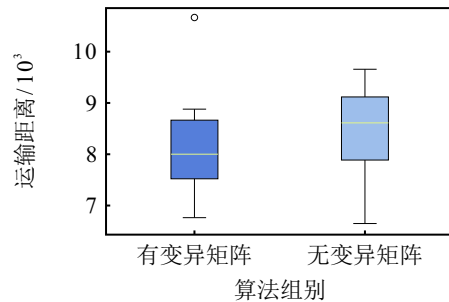
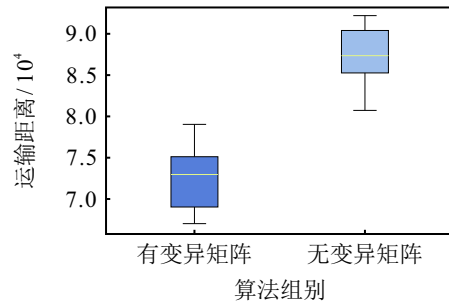


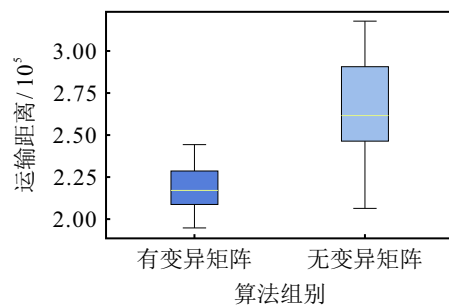
图2 拆单率箱型图对比



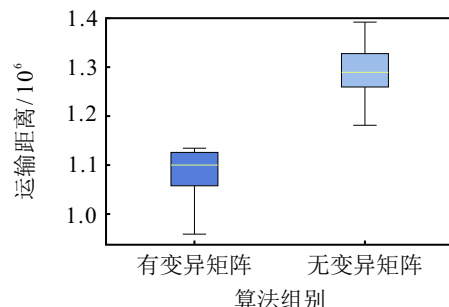
(a) I20O100



(b) I100O500



(c) I200O1000



(d) I600O3000

图3 运输距离箱型图对比

表2 Wilcoxon 秩和检验-有无变异矩阵

指标	中位数±标准差		Z值	P值	Cohen's d值
	有变异矩阵	无变异矩阵			
拆单率	12.15 % ± 2.89 %	12.90 % ± 2.24 %	0.753	0.459	0.037
运输距离	136926.57 ± 433232.43	149337.75 ± 515008.40	5.309	< 0.001	0.742

FFO-H 算法进行对比, 根据实验结果绘制的箱型图如图 2 和图 3 所示, 并使用 Wilcoxon 秩和检验来评估有无变异矩阵下指标是否存在显著差异, 结果如表 2 所示。

箱型图展示了 4 种问题规模下 10 组实验结果的最小最大值、上下四分位数、中位数以及异常值。由图 2 可知, 两种情况下算法在拆单率指标上的求解质量各有优劣, 对于 I200O100 规模的实例, 无变异矩阵的算法在拆单率上的表现显著优于原 FFO-H 算法, 但是, 随着问题规模增大该优势逐渐减弱。由表 2 中的 Wilcoxon 秩和检验结果可知, 在 99 % 置信水平下, 两种情况下算法拆单率指标无显著差异 ($P=0.459$), 具有统计学意义。由图 3 可知, 在 4 种问题规模下, 原 FFO-H 算法的运输距离指标均显著低于无变异矩阵的算法, 运输距离平均降低 16.25 %, 且在较大的实例上降低程度更为明显。表 2 显示, 在 99 % 置信水平下两种情况下算法的运输距离指标具有显著差异 ($P < 0.001$)。综上, 所提出商品间适应度指标保留了共现矩阵对拆单率的优化效果, 同时, 加入变异矩阵可以实现运输距离的进一步降低, 表明商品间适应度指标具有合理性和有效性。

4.3 分仓选品算法实验分析

为了验证 FFO 算法的有效性, 本文分别对小规模和大规模问题进行实验分析, 并分别讨论 FFO-S 算法和 FFO-H 算法的适用性。其中: 小规模算例为 I200O1000 及以下规模算例, 对比 FFO 算法与不同目标下 Gurobi 求解器的结果; 大规模算例为 I400O2000 及以上规模算例, 对比所提出 FFO-H 算法与文献中具有代表性的分仓选品算法。

选品结果服务于后续订单分配过程, 其优劣需要经过订单分配衡量。本文通过调研在线零售企业实际订单分配方式, 综合考虑拆单率和运输距离因素, 采用以下步骤进行订单分配。

step 1: 当一个订单到达时, 首先检测是否有单个仓库能够完全覆盖该订单需求的 SKU。若有一个仓库可满足, 则完成该订单分配; 若有多个仓库可满足, 则将该订单分配给距离客户最近的仓库。

step 2: 若该订单不可由唯一仓库满足, 则将订单拆分。选择能够最大程度覆盖该订单 SKU 品规的

仓库作为发货仓之一; 删除订单中已被满足的 SKU 品规, 余下品规视为新的子订单; 迭代选择最大程度覆盖该子订单的仓库, 直至原订单中所有 SKU 均被满足。

4.3.1 小规模问题对比分析

对于在 I200O1000 以下的小规模算例, 将所提出 FFO-S 算法和 FFO-H 算法与 Gurobi 求解器进行对比。由于本文问题考虑了两个优化目标, 基于求解器的 FFO-S 算法和 Gurobi 算法在求解时将分别以

表3 FFO-S、FFO-H 与 Gurobi 拆单率指标对比 %

问题规模	Gurobi	FFO-S	FFO-H	Gap1	Gap2
I200O100	30.0	30.0*	34.0	0.0	13.3
I600O300	23.7	24.3	27.3	2.5	15.2
I1000O500	22.8	21.8	25.6	-4.4	12.3
I1600O800	37.4	22.3	27.0	-40.4	-27.8
I2000O1000	41.3	18.7	24.2	-54.7	-41.4
均值	31.0	23.4	27.6	-24.5	-11.0

注: Gurobi和FFO-S以拆单率为优化目标;

$$Gap1 = \frac{FFO-S - Gurobi}{Gurobi} \times 100\%$$

$$Gap2 = \frac{FFO-H - Gurobi}{Gurobi} \times 100\%$$

表4 FFO-S、FFO-H 与 Gurobi 运输距离指标对比

问题规模	Gurobi	FFO-S	FFO-H	Gap1/%	Gap2/%
I200O100	1.19e+04*	1.19e+04*	1.33e+04	0.0	11.8
I600O300	3.12e+04*	3.12e+04*	3.40e+04	0.0	9.0
I1000O500	5.07e+04*	5.11e+04*	5.83e+04	0.8	15.0
I1600O800	8.19e+04	8.42e+04*	9.10e+04	2.8	11.1
I2000O1000	9.30e+04*	9.41e+04*	1.05e+05	1.2	12.9
均值	5.37e+04	5.55e+04	6.03e+04	3.4	12.3

注: Gurobi和FFO-S以拆单率为优化目标;

$$Gap1 = \frac{FFO-S - Gurobi}{Gurobi} \times 100\%$$

$$Gap2 = \frac{FFO-H - Gurobi}{Gurobi} \times 100\%$$

表5 FFO-S、FFO-H 与 Gurobi 运行时间对比 s

问题规模	Gurobi		FFO-S		FFO-H
	最小化拆单率	最小化运输距离	最小化拆单率	最小化运输距离	
I200O100	7200	33.2	59.4	1.9	0.03
I600O300	7200	8.5	7200	3.3	0.05
I1000O500	7200	33.4	7200	9.7	0.09
I1600O800	7200	7200	7200	25.0	0.18
I2000O1000	7200	3264.2	7200	35.9	0.27
均值	7200	2107.9	5771.9	15.2	0.12

拆单率和运输距离作为优化目标. 由于分仓选品属于战略层决策, 对于实时性要求较低^[22], 本文将最大求解时间均限定为 7 200 s. 此外, 为了确保对比的公平性, 所有指标计算均基于相同的订单分配规则, 实验结果如表 3 ~ 表 5 所示. 其中: 符号*表示算法或求解器在 7 200 s 内求得最优解, 粗体表示横向对比中的最好结果.

表 3 为小规模算例下 3 种算法在拆单率指标上的对比结果. 由表 3 可见: Gurobi 在任意算例中均无法在 7 200 s 内获得最优解; 相比之下, FFO-S 与 FFO-H 在拆单率指标上均有提升, 但是 FFO-S 的效果更加显著. 具体来看: FFO-S 的拆单率平均改进 24.5%, 且仅在 I200O100 规模算例上略劣于 Gurobi (约 2.5%); 随着问题规模增大, FFO-S 优势更加显著, 在 I200O1 000 规模上拆单率下降 22.6%. FFO-H 的拆单率平均改进 11.0%, 在 I100O500 以下规模算例均劣于 Gurobi (约 13.6%), 但是, 在 I160O800 及以上规模问题上的表现优于 Gurobi (约 34.6%), 主要因为问题规模更小时 Gurobi 能够探索更多情况, 但是随着问题规模的增大, 问题复杂度以及搜索空间均增加, 使得 Gurobi 更难以在合理时间内获得优质解. 此外, FFO-H 相比于 FFO-S 仍有 3.3% 的差距, 但是, 由表 5 可知, FFO-H 的平均运行时间均在秒级, 显著小于以拆单率为优化目标下的 FFO-S 算法.

表 4 为小规模算例下 3 种算法在运输距离指标上的对比结果. 由表 4 可见: Gurobi 在 4/5 的算例中求得最优解; 相比之下, FFO-S 能够在全部算例中求得最优解, 与 Gurobi 的平均 Gap 为 3.4%; FFO-H 与 Gurobi 的平均 Gap 为 12.3%. 尽管在运输距离指标上与 Gurobi 有所差距, 但是, 由表 5 可知: Gurobi 的平均求解时间为 2 107.9 s, 而 FFO-S 和 FFO-H 的平均求解时间分别为 15.2 s 和 0.12 s, 表明 FFO 算法能够大大降低求解时间, 在短时间内获得优质解.

上述实验结果表明了所提出 FFO 算法的有效性: 基于谱聚类的初始构建方案能够对 SKU 合理分配, 能够提供高质量初始解; 通过固定分配的单个 SKU 能够减少决策变量, 在保证解的质量的前提下, 能够有效减少求解时间. 本文在更大规模问题上进行预实验, 发现 FFO-S 均无法在 7 200 s 时限内求得算例最优解, 因此, 后续在大规模问题上的实验仅用 FFO-H 进行对比.

4.3.2 大规模问题对比分析

在线零售场景下商品种类繁多、订单数量庞大, 为了验证所提出 FFO-H 算法在大规模问题上的有

效性和在现实背景下的应用, 本文将 FFO-H 算法与分仓选品启发式算法在订单数为 2 000 ~ 12 000 的大规模问题上进行对比, 每个规模下求解 10 个不同的算例求其均值, 实验结果如表 6 和表 7 所示. 热销品算法^[5]、环形算法^[6]和增量算法^[7]是求解分仓选品问题的代表性算法, 但是, 增量算法针对制造业即时物流进行设计, 与本文在线零售业问题场景有所差异, 因此, 在此仅与环形算法和热销品算法进行对比.

表6 环形算法、热销品算法与 FFO-H 拆单率指标对比 %

问题规模	环形算法	热销品算法	FFO-H	Gap1	Gap2
I400O2 000	33.3	29.4	24.8	-25.5	-15.6
I600O3 000	33.3	29.3	24.9	-25.2	-15.0
I1 000O5 000	35.2	31.0	26.6	-24.4	-14.2
I1 600O8 000	33.7	29.6	24.9	-26.1	-15.9
I2 000O10 000	33.5	29.4	24.7	-26.3	-16.0
I2 400O12 000	32.6	28.4	24.0	-26.4	-15.5
均值	33.6	29.5	25.0	-25.6	-15.4

注: Gurobi和FFO-S以拆单率为优化目标;

$$\text{Gap1} = \frac{\text{FFO-H} - \text{环形算法}}{\text{环形算法}} \times 100\%$$

$$\text{Gap2} = \frac{\text{FFO-H} - \text{热销品算法}}{\text{热销品算法}} \times 100\%$$

表7 环形算法、热销品算法与 FFO-H 运输距离指标对比

问题规模	环形算法	热销品算法	FFO-H	Gap1/%	Gap2/%
I400O2 000	4.93e+05	4.47e+05	2.30e+05	-53.3	-48.5
I600O3 000	6.63e+05	5.99e+05	3.41e+05	-48.6	-43.1
I1 000O5 000	1.21e+06	1.09e+06	5.94e+05	-50.9	-45.5
I1 600O8 000	1.84e+06	1.65e+06	9.15e+05	-50.3	-44.5
I2 000O10 000	2.38e+06	2.14e+06	1.14e+06	-52.1	-46.7
I2 400O12 000	2.59e+06	2.34e+06	1.34e+06	-48.3	-42.7
均值	1.53e+06	1.38e+06	7.60e+05	-50.3	-45.2

注: Gurobi和FFO-S以拆单率为优化目标;

$$\text{Gap1} = \frac{\text{FFO-H} - \text{环形算法}}{\text{环形算法}} \times 100\%$$

$$\text{Gap2} = \frac{\text{FFO-H} - \text{热销品算法}}{\text{热销品算法}} \times 100\%$$

表 6 为大规模问题下不同算法拆单率指标的对比结果. 由表 6 可见: FFO-H 算法在全部实验算例上的平均拆单率为 25.0%, 相比于环形算法与热销品算法拆单率分别平均降低了 25.6% 和 15.4%. 主要原因在于: 本文在设计商品间适应度指标时, 商品间共现矩阵从拆单角度出发进行衡量, 使得高频出现在同一订单的 SKU 倾向于存放于同一仓库; 且 FFO-H 算法的优化阶段对多点存放 SKU 的再分配过程能够充分利用仓库容量; 另一方面, 订单分配优先考虑不拆单的分配情况, 能够协同降低拆单率. 此外, FFO-H 算法的拆单率指标未随着订单规模增大明显增加, I2 400O12 000 规模下 10 组实例的拆单

率不超过 27.5 %。

表 7 为大规模问题下不同算法运输距离指标的对比结果。由表 7 可见: FFO-H 算法在全部实验算例上的平均运输距离为 $7.60e+05$, 相比于环形算法与热销品算法拆单率分别平均优化 45.2 % 和 50.3 %; 且相较于拆单率指标, FFO-H 算法对于运输距离指标的优化程度更为显著。主要原因在于: 环形算法和热销品算法主要从历史订单共现角度进行分仓选品, 未考虑运输距离优化; 而 FFO-H 算法中的商品间变异矩阵使得目标客户群体相近的 SKU 倾向于存放于同一仓库, 能够有效避免远距离拆单造成运输距离大幅度增加。

综合上述实验结果可知: FFO-H 算法能够有效处理大规模问题, 在短时间内(秒级)求得优质解, 相比于其他启发式算法, 进一步实现了拆单率和运输距离的同时优化; 而 FFO-S 算法适用于 $I200O1\ 000$ 以下的小规模问题, 能够在可接受时间范围内求得比 FFO-H 算法质量更好的解。

5 结论

本文研究了在线零售业背景下的分仓选品问题, 以最小化拆单率和最小化运输距离为目标, 建立了混合整数规划模型。从历史订单共现情况和商品目标客户群体差异两个角度评估商品是否适合存放于同一仓库, 提出了商品间适应度指标, 并基于此设计了固定-优化两阶段的 FFO 算法。固定阶段通过谱聚类分配初始 SKU 并固定该部分决策变量; 优化阶段设计了基于求解器和启发式的两种策略, 以对剩余库容进行 SKU 的多点存放再分配。最后, 通过数值实验验证了所提出适应度指标的有效性, 以及 FFO-S 和 FFO-H 算法在不同规模场景下的有效性。

未来研究将进一步探索订单分配对选品方案的影响, 探索分仓选品与订单分配的联合优化; 此外, 分仓选品方案需要更新迭代来适应变动的市场需求, 研究可以综合考虑市场需求, 实现分仓选品动态优化。

参考文献 (References)

- [1] 胡祥培, 丁天蓉, 张源凯, 等. 基于关联网络的机器人移动货架系统货位分配方法[J]. 工程管理科技前沿, 2022, 41(1): 56-64.
(Hu X P, Ding T R, Zhang Y K, et al. A method for storage assignment in robotic mobile fulfillment system based on association networks[J]. *Frontiers of Science and Technology of Engineering Management*, 2022, 41(1): 56-64.)
- [2] 郑思宇, 黄敏, 蒋松辰, 等. 基于 GELTR-WOA 算法的多地点库存动态集成调拨[J]. 控制与决策, 2024, 39(7): 2334-2344.

- (Zheng S Y, Huang M, Jiang S C, et al. Multi-location inventory dynamic integrated allocation based on GELTR-WOA algorithm[J]. *Control and Decision*, 2024, 39(7): 2334-2344.)
- [3] Zhu S, Hu X P, Huang K, et al. Optimization of product category allocation in multiple warehouses to minimize splitting of online supermarket customer orders[J]. *European Journal of Operational Research*, 2021, 290(2): 556-571.
- [4] 丁天蓉, 张源凯, 王玉英, 等. 基于关联网络的“人-机”双拣选系统仓储商品选择方法[J]. 中国管理科学, 2023, 31(3): 26-37.
(Ding T R, Zhang Y K, Wang Y Y, et al. An item allocation method for “manual-robotic” dual picking systems based on association networks[J]. *Chinese Journal of Management Science*, 2023, 31(3): 26-37.)
- [5] Catalán A, Fisher M. Assortment allocation to distribution centers to minimize split customer orders[J]. *Social Science Electronic Publishing*, 2013, 215(1): 1-31.
- [6] 李建斌, 李乐乐, 黄日环. 最小化拆单率的在线零售商多仓商品摆放优化策略研究[J]. 管理工程学报, 2017, 31(3): 167-173.
(Li J B, Li L L, Huang R H. Inventory allocation policy to distribution centers with minimum split orders in online retailing[J]. *Journal of Industrial Engineering and Engineering Management*, 2017, 31(3): 167-173.)
- [7] 李建斌, 宋向荣, 郭金晟, 等. 制造业生产物流中多仓协同选品策略研究[J]. 系统工程理论与实践, 2023, 43(6): 1736-1749.
(Li J B, Song X R, Guo J S, et al. Collaborative optimization of assortment allocation among multiple warehouses in manufacturing production logistics[J]. *Systems Engineering — Theory & Practice*, 2023, 43(6): 1736-1749.)
- [8] Acimovic J, Graves S C. Making better fulfillment decisions on the fly in an online retail environment[J]. *Manufacturing & Service Operations Management*, 2015, 17(1): 34-51.
- [9] 黄敏芳, 张源凯, 王颜新. 网上超市拆分订单的合并打包优化决策方法[J]. 系统工程理论与实践, 2021, 41(2): 286-296.
(Huang M F, Zhang Y K, Wang Y X. A decision-making method of package consolidation for split orders in online supermarket[J]. *Systems Engineering — Theory & Practice*, 2021, 41(2): 286-296.)
- [10] 吴仁超, 贺建军, 李欣, 等. 基于混合元启发式算法的订单分批问题[J]. 控制与决策, 2022, 37(8): 2110-2118.
(Wu R C, He J J, Li X, et al. Hybrid metaheuristic algorithm for order batching problem[J]. *Control and Decision*, 2022, 37(8): 2110-2118.)
- [11] 李儒博, 邓旭东, 刘翱, 等. RMFS 补货货品存储分配问题研究[J]. 控制与决策, 2025, 40(2): 528-536.
(Li R B, Deng X D, Liu A, et al. Research on RMFS replenishment items storage assignment problem[J]. *Control and Decision*, 2025, 40(2): 528-536.)

- [12] Huang M F, Zhao Y Y, Wang Y X, et al. Application of a green logistics strategy to fulfil online retailing: Split-order consolidated delivery in large-scale online supermarkets[J]. *International Journal of Logistics Research and Applications*, 2024, 27(5): 746-759.
- [13] 高展, 姜艳萍. 考虑交付期限的订单配送策略[J]. *控制与决策*, 2024, 39(6): 2089-2096.
(Gao Z, Jiang Y P. Requests dispatch strategy with delivery constraints[J]. *Control and Decision*, 2024, 39(6): 2089-2096.)
- [14] Zhang M X, Wu J Y, Wu X, et al. Hybrid evolutionary optimization for takeaway order selection and delivery path planning utilizing habit data[J]. *Complex & Intelligent Systems*, 2022, 8(6): 4425-4440.
- [15] Chuang Y F, Lee H T, Lai Y C. Item-associated cluster assignment model on storage allocation problems[J]. *Computers & Industrial Engineering*, 2012, 63(4): 1171-1177.
- [16] Chuang Y F, Lee H T, Tan S W. A robust heuristic method on the clustering-assignment problem model[J]. *Computers & Industrial Engineering*, 2016, 98: 63-67.
- [17] Jane C C, Lai Y W. A clustering algorithm for item assignment in a synchronized zone order picking system[J]. *European Journal of Operational Research*, 2005, 166(2): 489-496.
- [18] 罗健, 唐加福, 于清雅, 等. O2O 外卖商圈划分及顾客需求分布规律发现[J]. *中国管理科学*, 2023, 31(3): 58-68.
(Luo J, Tang J F, Yu Q Y, et al. The division of O2O takeaway business zone and discovery of customer demand distribution law[J]. *Chinese Journal of Management Science*, 2023, 31(3): 58-68.)
- [19] 徐森, 高婷, 徐秀芳, 等. 一种基于成员选择的簇加权聚类集成算法[J]. *控制与决策*, 2024, 39(12): 4136-4140.
(Xu S, Gao T, Xu X F, et al. A cluster-weighted clustering ensemble algorithm based on member selection[J]. *Control and Decision*, 2024, 39(12): 4136-4140.)
- [20] Fonseca G H G, Figueiroa G B, Toffolo T A M. A fix-and-optimize heuristic for the unrelated parallel machine scheduling problem[J]. *Computers & Operations Research*, 2024, 163: 106504.
- [21] 金爽, 周晶, 胡骞. 考虑灵活弧中断的网络维护调度问题及其算法研究[J]. *中国管理科学*, DOI: 10.16381/j.cnki.issn1003-207x.2022.2617.
(Jin S, Zhou J, Hu Q. Research on the network maintenance scheduling problem with flexible arc outages and its algorithm[J]. *Chinese Journal of Management Science*, DOI: 10.16381/j.cnki.issn1003-207x.2022.2617.)
- [22] 翟梦月, 王征, 李延通, 等. 可移动货架仓储系统中商品储位分配问题研究[J]. *中国管理科学*, 2023, 31(3): 167-176.
(Zhai M Y, Wang Z, Li Y T, et al. Research on the problem of allocating commodity storage locations in a mobile — Rack storage system[J]. *Chinese Journal of Management Science*, 2023, 31(3): 167-176.)

作者简介

许瑞 (1983-), 男, 副教授, 博士, 主要研究方向为路径优化、生产调度、智能优化算法、供应链管理, E-mail: rxu@hhu.edu.cn;

丁子千 (2000-), 女, 硕士生, 主要研究方向为智能优化算法、机器学习, E-mail: 221313030016@hhu.edu.cn;

肖巍 (1994-), 男, 讲师, 博士, 主要研究方向为供应链管理、优化算法, E-mail: wei.xiao@hhu.edu.cn.