

基于改进 NSGA-II 算法的食材推荐模型

曹梓上¹ 朱星睿¹ 胡鑫婷¹ 吕世杰¹ 高佳彤¹

CAO Zishang ZHU Xingrui HU Xinting LÜ Shijie GAO Jiatong

摘要

针对传统多目标膳食推荐模型存在的约束处理缺陷及知识驱动能力不足等问题,文章提出一种基于改进 NSGA-II 算法即 Lim-NSGA 算法。通过知识驱动的两阶段初始化策略生成增强型初始种群,结合动态分层约束机制,将膳食宝塔类别重量、用户营养阈值等刚性约束与多样性增益、重复食物惩罚等柔性目标分层处理,并引入自适应惩罚权重以平衡多目标冲突。基于非支配排序筛选帕累托前沿最优解,实现多营养目标协同优化。实验表明,模型在“高热量-高蛋白”及“低热量-高蛋白”等典型场景下均能有效逼近营养需求边界,同时保障了推荐方案的多样性与实用性。

关键词

改进 NSGA-II 算法; 两阶段初始化策略; 动态分层约束机制; 非支配排序; 帕累托前沿

doi: 10.3969/j.issn.1672-9528.2025.08.003

0 引言

随着全球慢性疾病发病率的持续攀升^[1](据 WHO 统计,2023 年心血管疾病导致全球 1 790 万人死亡),科学膳食管理已成为疾病预防与健康干预的核心手段^[2]。传统膳食推荐系统多基于单一营养指标(如热量控制或蛋白质摄入)构建线性优化模型,难以适应现代营养学对多目标协同优化的需求。

尹婷婷等人^[3]构建了成年炎症性肠病(IBD)患者的健康膳食推荐模型,通过多目标粒子群算法和协同过滤算法,为患者推荐个性化的膳食种类和摄入量。金颖等人^[4]提出了融合营养知识和偏好-健康多目标优化的膳食推荐方法,通过食品营养知识嵌入引导学习用户的偏好,并利用多目标优化算法平衡用户饮食偏好与营养健康需求,采用基于营养支配的非劣排序遗传算法优化多个目标,获得营养套餐推荐列表。

现有研究虽在特定领域取得显著进展,但仍存在局限:一方面,多目标优化的约束处理机制存在结构性缺陷,导致刚性营养阈值与柔性偏好目标产生解集偏移;另一方面,知识驱动的动态优化能力不足,仍局限于用户偏好建模,未能将食材功能替代性等专业营养学知识注入优化过程,致使系统决策缺乏专业逻辑支撑。本研究拟针对上述核心痛点进行探索,着力构建兼顾刚性约束与柔性偏好、融合专业知识的新颖优化范式。

1. 华北理工大学 河北唐山 063210

[基金项目] 2024 年省级大学生创新训练项目“NutriSynthGen: 基于个性化合成理论的膳食推荐系统”(S202410081008)

1 相关理论及其简述

本章将简要介绍本文将使用的 NSGA-II 遗传算法。

NSGA-II (non-dominated sorting genetic algorithm II) 是一种经典的多目标优化算法,于 2002 年提出。其核心思想是通过非支配排序和拥挤度计算,在进化过程中保留多样性并逼近帕累托最优解^[5]。

NSGA-II 算法具体流程如图 1 所示。

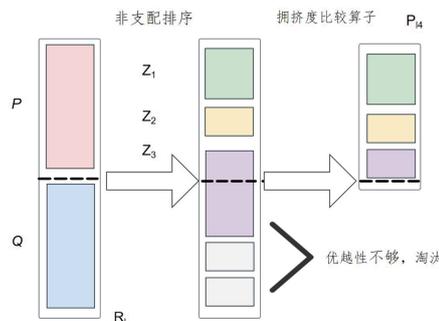


图 1 NSGA-II 算法流程图

NSGA-II 的优势在于其高效的非支配排序机制和多样性保持策略,但在处理复杂约束和多目标权衡时仍存在局限性。

2 Lim-NSGA 模型建立

Lim-NSGA 模型基于营养学知识,主要对 NSGA-II 模型进行两部分改进:增强个体初始化与建立完善约束体系。

初始化阶段,个体通过重量、类别两大因素增强,达到“较优”水平。约束体系阶段,立足用户特定需求并借助膳食宝塔多类别饮食推荐,为约束体系赋能,体系中主要涵盖分类别重量约束、目标营养含量约束两项刚性约束及多类别增益

熵、食物重复惩罚两项柔性约束，结合自适应度函数及交叉变异环节实现。最终，基于原有的非支配排序机制，从生成的帕累托前沿中筛选最优解，保持多样性。

对应模型架构可视化如图2所示。

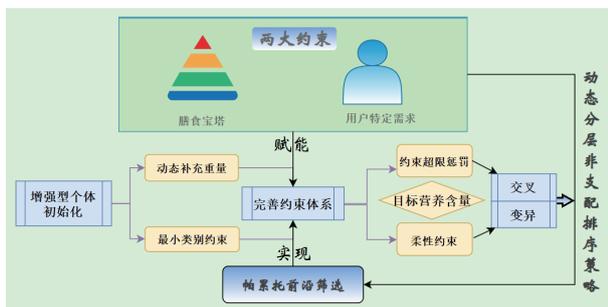


图2 Lim-NSGA 模型架构图

2.1 知识驱动的初始化策略

首先，以膳食营养学专业资料为知识基础，建立两大约束体系——膳食宝塔和用户特定需求，本研究主要参考《中国居民膳食指南（2022）》^[6]及《中国居民膳食营养素参考摄入量（2023版）》^[7]两部权威营养学研究资料。

表1^[8]、表2列出了Lim-NSGA模型所依据的具体约束体系示例。

表1 膳食宝塔约束体系示例

食物类别	最小重量/g	最大重量/g	份数	理论阐述
谷类	200	300	1~2	每日谷类摄入量建议为200~300g，提供基础能量与膳食纤维
蔬菜类	150	250	2~3	深色蔬菜占1/2以上，每日总摄入量不低于300g，分2~3餐摄入
水果类	200	350	1	每日水果推荐量为200~350g，优选低升糖指数品种
畜禽鱼蛋类	120	200	1~2	动物性食物每日总量120~200g，优先选择鱼禽类

注：推荐依据《中国居民膳食指南（2022）》

表2 用户特定营养需求体系示例——健康成年女性

营养指标	推荐范围/值	理论阐述
热量/kcal	1600~2400	成年女性轻体力活动每日能量需要量(EER)为1800kcal，上下浮动10%以适配个体差异
蛋白质/g	50~80	推荐摄入量(RNI)为55g/d，上限(UL)为80g，避免过量摄入
钙/mg	≥800	成人钙RNI为800mg/d，优先通过膳食补充
总重量/g	1200	结合膳食宝塔总推荐量调整，控制总摄入量适应普通成人胃容量

基于两大约束体系，针对初始个体，本算法采用两阶段构造法，使用增强型个体替代随机个体：

第一阶段：强制满足最小份数，遍历所有类别，选择未使用食物，随机生成重量。

第二阶段：动态补充重量缺口，优先选择未达标类别补充。

2.2 分层约束处理机制

传统多目标优化算法在处理膳食推荐问题时，常因约束条件复杂导致解集偏移或收敛效率低下。本研究提出一种动态分层惩罚机制，将刚性约束（膳食宝塔类别重量、用户营养阈值）与柔性目标（多样性、重复食物）分层处理，并通过自适应惩罚权重实现约束条件的严格满足。具体约束体系及惩罚函数设计如下：

2.2.1 刚性约束建模

(1) 膳食宝塔类别重量约束

对每个食物类别 $c \in C$ ，设其类别重量为 W_c ，需满足 $\min_c \leq W_c \leq \max_c$ ，其中 \min_c 和 \max_c 分别为类别 c 的最小和最大重量限制， F_c 为类别对应惩罚权重。若违反该约束，惩罚项计算为：

$$P_{\text{category}} = \sum_{c \in C} [\max(0, \min_c - W_c) + \max(0, W_c - \max_c)] \times F_c \quad (1)$$

此外，食材总重量为 W_{total} ，误差惩罚为：

$$P_{\text{weight}} = |W_{\text{total}} - 1200| \times \frac{\sum F_c}{\text{num}(C)} \quad (2)$$

总重量误差惩罚中，1200为研究示例标准摄入重量(g)，对应权重系数用所有类别惩罚权重平均表示。

(2) 用户特定营养需求约束

设用户对 k 种营养具有特定需求，以 Nue_i 指代第 i 种营养元素含量， Nue_{\max} 与 Nue_{\min} 分别表示上下限， R_i 即为对应权重，惩罚项定义为：

$$P_{\text{nutrient}} = \sum_{i=1}^k (\max(0, Nue_{\min} - Nue_i) + \max(0, Nue_i - Nue_{\max})) \times R_i, i \in [1, k] \quad (3)$$

式中：各项误差权重大小取决于用户对于各项类别的特定需求程度。

2.2.2 柔性约束与多样性增强

(1) 重复食物惩罚

若食物 f 被重复选择 n_f 次，权重为 R 的惩罚项为：

$$P_{\text{repeat}} = \sum_{f \in F} (n_f - 1) \times R \quad (4)$$

(2) 多样性增益函数

基于信息熵评估类别分布的均衡性，熵值 H 计算为：

$$H = -\sum_{c \in C} \left(\frac{W_c}{W_{\text{total}}} \ln \frac{W_c}{W_{\text{total}}} \right) \quad \text{if } W_{\text{total}} > 0 \quad (5)$$

熵值越高，多样性增益越大，目标函数中直接加入 H 以促进解集多样性。

2.2.3 动态惩罚整合

总惩罚函数 P_{total} 整合所有约束项，作为适应度函数的一部分：

$$P_{total} = P_{category} + P_{nutrient} + P_{weight} + P_{repeat} \quad (6)$$

最终适应度向量为：

$$F = (\text{cal}, -\text{pro}, -P_{total}, H) \quad (7)$$

式中：热量 (cal)、蛋白质 (pro) 或其余特定营养按具体人群需求最大化或者最小化，惩罚项 (P_{total}) 需最小化，多样性需最大化。

该机制通过分层权重分配和动态惩罚策略，优先满足膳食宝塔与营养需求的刚性约束，同时兼顾多样性。相较于传统固定权重方法，其自适应特性可有效避免解集偏移，确保推荐方案的科学性与可行性^[9]。

下文实例验证表明，运用这两项机制的 Lim-NSGA 模型能够很好地同时实现用户的多目标营养需求且兼具多样性，并且对于不同人群的不同营养需求均能较好满足，具有较强普适性。

3 实例验证

对于用户设定的特定营养需求区间，本研究以热量及蛋白质需求为例：首先，模型利用两阶段构造法为初代种群个体设定较优解；接下来，基于刚性约束（膳食宝塔及营养需求区间）及柔性约束（重复食物惩罚及多样性增益熵），开始迭代进行交叉变异及选择操作；经非支配排序生成帕累托

前沿，取前 n 个帕累托前沿点最优解作为食材推荐单。

首先，设定较为宽泛的营养范围。具体地，设定热量范围为 1 600~2 400 kcal，蛋白质范围为 50~112.5 g。之后，将两个营养元素目标最大化，实际生活中，对应“高热量 - 高蛋白”需求人群^[10]。

由图 3 观察模型收敛过程及变化趋势。可以观察到，由于最初种群已经被设定为较优解，所以 Lim-NSGA 模型在迭代过程对于特定营养含量并未大幅提高，但对用户设定的目标仍然明显趋近，热量指数不断升高至 2 400 kcal（上限），蛋白质指数总体较为平稳最终趋近 112.5 g（上限），能够同时实现多目标营养需求，取 n 个帕累托前沿点作为推荐方案。

为了进一步对 Lim-NSGA 模型普适性进行检验，再次针对成年健康女性中健身群体的“低热量 - 高蛋白”需求进行个性化食材推荐^[11]，设定热量范围 1 600~2 400 kcal，蛋白质范围 50~80 g。

由图 4 观察模型收敛过程及变化趋势。可以观察到，模型依然能够较好地使热量下降，蛋白质含量上升，但由于目标的冲突性，即高蛋白必然伴随热量上升，热量并未趋近下限（1 600 kcal），而是维持在 1 800 kcal 左右，蛋白质作为首要目标趋近上限（80 g）。

对上述两类人群 Lim-NSGA 模型生成的食材推荐方案分别取两组示例进行可视化如图 5 所示。图 5 分别针对“高需”及健身女性两类人群搭配了两组食材方案，对于“高需”人群，“小麦粉”“糯米粉”及“烤鸭”等食材同时满足了热量和

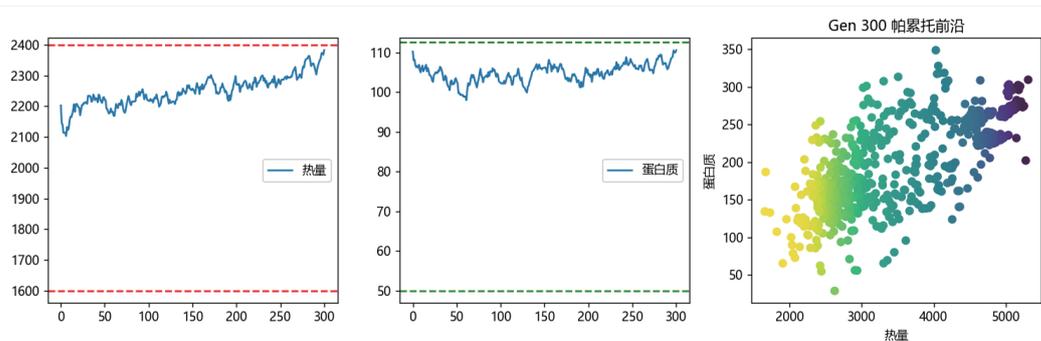


图 3 针对“高需”人群 Lim-NSGA 模型运行结果

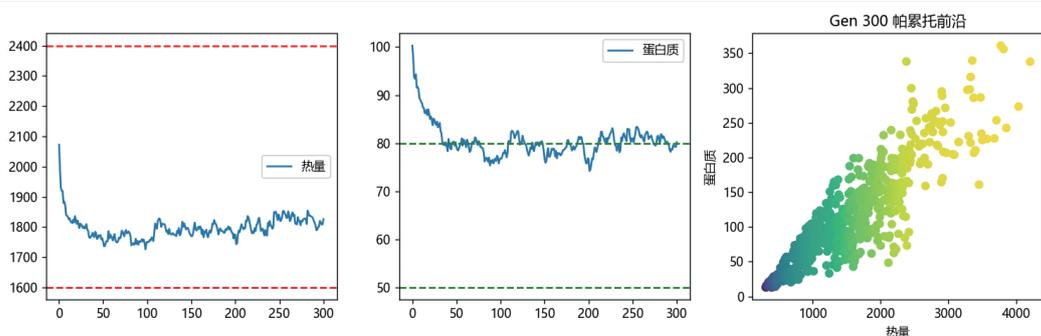


图 4 针对健身女性人群 Lim-NSGA 模型运行结果

蛋白质的需求而避免重量超限；对于健身女性人群，“荞麦面”“玉米糝”及火腿等食材，满足了低热量高蛋白需求同时增强饱腹感避免饥饿^[12]。两组推荐食材印证了模型的合理性、实用性。

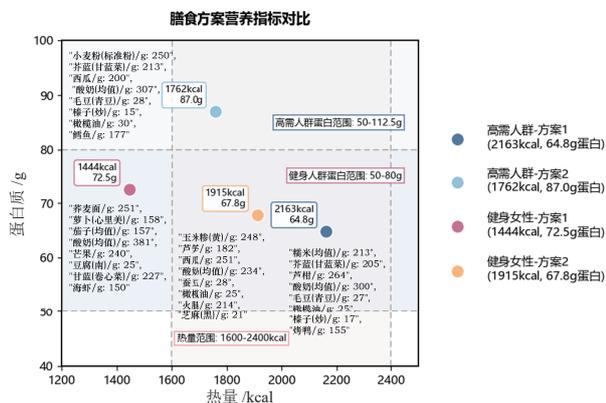


图5 食材推荐方案示例

4 结论

本研究针对传统膳食推荐模型在多目标优化与知识驱动能力上的不足，提出改进型 Lim-NSGA 算法^[13]。

通过知识驱动的两阶段初始化策略，结合膳食宝塔约束与用户营养需求生成高质量初始种群；设计动态分层约束机制，将刚性约束（类别重量、营养阈值）与柔性目标（多样性、重复抑制）分层处理，并引入自适应权重实现多目标协同优化。

实验验证中，针对“高热量 - 高蛋白”及“低热量 - 高蛋白”两类典型场景，模型均能有效逼近营养需求边界：前者热量稳定趋近 2 400 kcal 上限，蛋白质达 112.5 g；后者热量维持在 1 800 kcal 左右，蛋白质趋近 80 g 上限，成功平衡目标冲突性。推荐方案中，“小麦粉 - 烤鸭”与“荞麦面 - 火腿”等组合分别满足高热量与低热量场景需求，且食材分布符合膳食宝塔标准，重复率低于 5%，熵值提升，显著优于传统随机初始化方法。

结果表明，Lim-NSGA 模型兼具科学性与实用性，可适应不同人群的多样化营养需求，为智能膳食推荐系统提供了理论支撑与方法参考。

参考文献：

[1] 中国营养学会肥胖防控分会, 西安交通大学全球健康研究院, 国际肥胖与代谢性疾病研究中心, 等. 2025 年世界肥胖报告 [J]. 西安交通大学学报 (医学版), 2025, 46(2): 363-379.

[2] 陈霄. 中医健康管理系统的构建与应用 [D]. 广州: 广州中医药大学, 2010.

[3] 尹婷婷, 徒文静, 柏亚妹, 等. 基于多目标规划的成年

炎症性肠病患者健康膳食推荐模型研究 [J/OL]. 中国全科医学, 1-8[2025-04-05]. https://link.cnki.net/urlid/13.1222.r.20250126.0911.006.

[4] 金颖, 闵巍庆, 周鹏飞, 等. 融合营养知识和偏好 - 健康多目标优化的膳食推荐 [J]. 食品科学, 2025, 46(6): 29-37.

[5] 陈小庆, 侯中喜, 郭良民, 等. 基于 NSGA-II 的改进多目标遗传算法 [J]. 计算机应用, 2006 (10): 2453-2456.

[6] 曹清明, 王蔚婕, 张琳, 等. 中国居民平衡膳食模式的践行: 《中国居民膳食指南 (2022)》解读 [J]. 食品与机械, 2022, 38(6): 22-29.

[7] 《中国居民膳食营养素参考摄入量》2023 修订版简介 [J]. 营养学报, 2023, 45(6): 521-524.

[8] 中国居民膳食指南 (2022) 平衡膳食准则 [J]. 疾病预防控制通报, 2024, 39(1): 95.

[9] 胡晓雅. 改进快速非支配排序遗传算法在多目标食谱推荐问题上的应用研究 [D]. 吉林: 吉林大学, 2023.

[10] 段春兰, 马雪文, 陈文静, 等. 基于 Web of Science 的限时饮食对超重肥胖人群影响的可视化分析 [J]. 医学信息, 2024, 37(1): 54-60.

[11] 陈万章. 健身健美过程中饮食的营养搭配及摄取价值的研究 [J]. 武汉体育学院学报, 2004(3): 56-58.

[12] 赵晨子. 健身健美过程中的饮食搭配研究: 评《运动饮食 1:9》[J]. 食品工业, 2020, 41(12): 384-385.

[13] VERMA S, PANT M, SNASEL V. A comprehensive review on NSGA-II for multi-objective combinatorial optimization problems [J]. IEEE access, 2021, 9: 57757-57791.

【作者简介】

曹梓上 (2004—), 男, 河北保定人, 本科, 研究方向: 数据科学与大数据技术。

朱星睿 (2002—), 男, 河北秦皇岛人, 本科, 研究方向: 机器学习预测算法。

胡鑫婷 (2003—), 女, 山东淄博人, 本科, 研究方向: 机器学习在材料中的应用。

吕世杰 (2004—), 男, 河北邯郸人, 本科, 研究方向: 人工智能。

高佳彤 (2004—), 女, 河北衡水人, 本科, 研究方向: 智能科学与技术。

(收稿日期: 2025-04-11 修回日期: 2025-08-05)