

基于安全车辆路径问题的引导式非自回归知识蒸馏

许伟乐, 潘安琪*

东华大学信息与智能科学学院, 上海

收稿日期: 2026年4月23日; 录用日期: 2026年5月21日; 发布日期: 2026年5月28日

摘要

神经组合优化(Neural Combinatorial Optimization, NCO)通过学习问题实例与其对应解之间的映射关系, 在标准路径规划问题中已展现出较强的竞争性能。然而, 当路径规划问题在目标函数、偏好或约束条件等方面发生变化时, 已有训练模型通常难以直接迁移应用。此外, 在此类新问题设定下, 获取高质量最优解作为训练数据往往需要付出较高的计算代价, 从而限制了从头训练深度模型的实际可行性。文章以基于安全优先策略的容量约束车辆路径规划问题(Capacitated Vehicle Routing Problem with Safety-Based Routing Preference, CVRP-SRP)为研究对象, 面向有限数据场景, 提出一种由知识蒸馏驱动的神-启发式优化框架。在已有研究中, 神经辅助自适应邻域搜索(Neural-Assisted Adaptive Neighborhood Search, NANS)算法通过结合轻编码-重解码模型(Light Encoder-Heavy Decoder, LEHD)与自适应大邻域搜索(Adaptive Large Neighborhood Search, ALNS), 能够生成高质量路径解, 并在解的质量与稳定性方面表现出较优性能。然而, 该方法依赖迭代式搜索过程, 计算开销较大, 且在新问题场景下的泛化能力仍存在一定局限。为在有限数据条件下实现高效知识迁移, 文章在上述方法基础上引入引导式非自回归知识蒸馏(Guided Non-Autoregressive Knowledge Distillation, GNARKD)策略, 构建教师-学生学习框架。具体而言, 以在标准CVRP数据上训练得到的LEHD模型作为教师模型, 用于学习通用的路径构造规律与决策模式; 同时构建轻量化学生LEHD模型, 通过知识蒸馏机制对其进行优化, 使其在较少目标数据的条件下即可有效适应CVRP-SRP问题。通过该过程, 实现了从标准路径规划问题到具有安全约束的新问题设定的高效迁移, 并在降低推理复杂度的同时提升模型的应用效率。不同于仅依赖最终解进行学习的传统蒸馏方法, GNARKD同时迁移动作序列与决策概率分布, 从而保留了自回归解码过程中所蕴含的顺序相关决策知识。在社区规模路径实例上的实验结果表明, 与未采用知识蒸馏及直接迁移策略的模型相比, 所提出的框架在解的质量、鲁棒性以及数据利用效率方面均取得了显著提升, 为面向安全感知的无人配送车辆路径规划提供了一种有效且具有实际应用价值的解决方案。

关键词

无人驾驶配送车辆, 知识蒸馏, 车辆路径问题, 自适应大邻域搜索, 小样本学习

*通讯作者。

Guided Non-Autoregressive Knowledge Distillation for Safety-Based Vehicle Routing Problem

Weile Xu, Anqi Pan*

School of Information and Intelligent Science, Donghua University, Shanghai

Received: April 23, 2026; accepted: May 21, 2026; published: May 28, 2026

Abstract

Neural combinatorial optimization (NCO) has emerged as an effective approach for routing problems by learning from instance-solution pairs and achieving competitive performance on standard formulations. However, when routing problems are reformulated with new objectives, preferences, or constraints, existing trained models are often no longer directly applicable. Moreover, under such new problem settings, obtaining optimal solutions as training data usually incurs high computational cost, which limits the practicality of retraining deep models from scratch. This paper studies the Capacitated Vehicle Routing Problem with Safety-Based Routing Preference (CVRP-SRP) as a representative case and proposes a knowledge-distillation-driven neural-heuristic framework for limited-data scenarios. In existing research, the Neural-Assisted Adaptive Neighborhood Search (NANS) algorithm combines the Light Encoder-Heavy Decoder (LEHD) model with Adaptive Large Neighborhood Search (ALNS) to generate high-quality routing solutions, demonstrating superior performance in terms of solution quality and stability. However, this method relies on iterative search processes with substantial computational overhead, and its generalization capability under new problem scenarios remains somewhat limited. To achieve efficient knowledge transfer under limited-data conditions, this paper introduces a Guided Non-Autoregressive Knowledge Distillation (GNARKD) strategy based on the aforementioned method, constructing a teacher-student learning framework. Specifically, an LEHD model trained on standard CVRP data serves as the teacher model to capture general routing construction patterns and decision-making rules; meanwhile, a lightweight student LEHD model is constructed and optimized through the knowledge distillation mechanism, enabling it to effectively adapt to the CVRP-SRP problem with limited target data. Through this process, efficient transfer from standard routing problems to new problem settings with safety constraints is achieved, while reducing inference complexity and improving model application efficiency. Unlike traditional distillation methods that rely solely on final solutions for learning, GNARKD simultaneously transfers action sequences and decision probability distributions, thereby preserving the sequential decision knowledge embedded in the autoregressive decoding process. Experimental results on community-scale routing instances demonstrate that, compared with models without knowledge distillation and direct transfer strategies, the proposed framework achieves significant improvements in solution quality, robustness, and data utilization efficiency, providing an effective and practically valuable solution for safety-aware unmanned delivery vehicle routing.

Keywords

Unmanned Delivery Vehicle, Knowledge Distillation, Vehicle Routing Problem, Adaptive Large Neighborhood Search, Small Sample Learning

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着无人配送车辆(Unmanned Delivery Vehicles, UDV)在社区物流场景中的不断应用, 路径规划方法需更好地满足实际运行需求。不同于传统的配送环境, UDV 通常运行于道路条件复杂、行人与车辆交互频繁且交通状况动态变化的场景中。因此, 其路径规划目标不再是固定不变的, 而是需要根据具体应用需求进行动态调整, 这使得实际 UDV 路径规划问题与现有的传统模型存在显著差异。

从优化角度来看, UDV 路径规划问题可视为容量约束车辆路径问题(Capacitated Vehicle Routing Problem, CVRP)的一种扩展形式。然而, 经典 CVRP 主要关注车辆容量约束下的运输效率, 即通过满足客户需求来最小化运输成本。在实际配送过程中, 路径决策往往需要综合考虑安全性、交通状况以及环境不确定性等因素, 而不仅仅依赖行驶距离。因此, 有必要在传统 CVRP 模型的基础上进行拓展, 构建基于安全优先策略的容量约束车辆路径规划问题(Capacitated Vehicle Routing Problem with Safety-Based Routing Preference, CVRP-SRP), 以使路径规划能够更好地反映 UDV 配送任务的运行特征与优化需求。

目前, 针对 CVRP 已经提出了大量优化算法[1] [2], 总体上可分为启发式方法与基于机器学习的方法。典型的启发式方法包括 Lin-Kernighan-Helsgaun (LKH)算法[3]、蚁群优化(Ant Colony Optimization, ACO)算法[4]、模拟退火(Simulated Annealing, SA)算法[5]以及遗传算法(Genetic Algorithm, GA) [6]等。这些方法通常依赖针对具体实例的迭代搜索过程不断进行优化, 并在标准路径规划问题上表现出较好的性能。然而, 当问题设定发生变化时, 此类方法往往需要重新进行搜索或对算法进行重新设计, 从而限制了其在复杂部署场景环境下的效率。为提升计算效率与模型适应能力, 近年来基于机器学习的路径优化方法逐渐受到关注。Hu 等[7]通过从不同初始解多次重启并进行多方向搜索, 提高了元启发式算法的性能与鲁棒性; Zhao 等[8]利用深度强化学习生成高质量初始路径, 并结合局部搜索进一步优化; Herdianto 等[9]提出了一种将边选择模型与元启发式局部搜索相结合的混合框架, 用于引导路径改进; Cheng [10]等提出了启发式引导的强化学习方法, 将领域启发式知识融入学习过程, 从而加速模型收敛并提升样本利用效率。

尽管上述方法取得了一定的进展, 但基于机器学习的路径规划方法仍然需要依赖大量的高质量“问题-解”的样本来进行训练。对于 CVRP-SRP 这类新构建的问题, 尽管可以通过仿真生成问题实例, 但获取充足的高质量解或最优解仍需较高的计算代价, 从而限制了针对目标问题从头训练深度模型的实际可行性。针对上述问题, 本文构建了一种由知识蒸馏驱动的神经-启发式优化框架, 将在标准 CVRP 上训练得到的教师模型中的有效知识迁移至适用于 CVRP-SRP 的学生模型, 从而降低从头训练所带来的高昂代价。本文的主要贡献如下:

- 1) 提出了一种基于知识蒸馏的学习策略, 用于解决有限数据条件下的社区 UDV 路径优化问题。通过在 CVRP-SRP 模型中引入教师-学生框架, 从而提升模型在小样本场景下的泛化能力与鲁棒性。
- 2) 在结合了 LEHD 模型与 ALNS 的神经启发式优化框架基础上, 引入引导式知识蒸馏策略, 将教师模型生成的高质量决策策略迁移至轻量化学生模型, 从而在降低计算开销的同时实现高质量的路径规划。
- 3) 在社区规模路径规划场景下开展了大量实验验证所提方法的有效性。与未采用知识蒸馏或直接向迁移策略的模型相比, 所提方法在有限训练数据条件下能够获得更优的解质量及更高的数据利用效率。

2. 相关基础理论

2.1. 基于学习的路径问题解决模型

基于学习的方法已经成为求解路径规划及车辆路径问题的重要研究方向, 相较于传统启发式算法, 这类方法通常能够获得更高质量的解, 并展现出更强的泛化能力。现有的大多数方法将路径规划问题建模为一个构造式的决策过程, 即由神经网络按序逐步生成解, 而非采用端到端的直接回归方式。在这一范式下, 编码器-解码器的结构被广泛采用, 其中编码器用于提取问题的全局结构信息, 而解码器则负责在解构造过程中建模复杂的决策依赖关系。近年来的研究进一步表明, 模型结构呈现出向更强解码能力发展的趋势, 通常通过引入并行机制或非自回归策略来提升解的质量与推理效率。这类基于学习的模型在不同规模问题上展现出良好的灵活性、可扩展性与泛化能力。然而, 其性能提升往往会伴随着训练稳定性不足和监督信号有限等问题, 这种问题尤其在复杂解码机制下更为明显。上述问题促使学者进一步探索更有效的模型结构与解码策略。

LEHD 模型[11]已经成为一种基于学习的路径规划模型中的有效设计范式, 其核心目标是在表征能力与计算效率之间取得平衡。不同于传统方法在解码前对所有节点仅进行一次嵌入表示, 该类方法强调在解码阶段对节点表示进行反复且自适应的更新。通过在编码器中控制计算复杂度, 并将主要表达能力集中于解码器, 这类结构能够在解的构造过程中动态刻画节点之间依赖于决策的关系。在解码过程中, 关键节点(如当前位置、目标节点及剩余候选节点)的嵌入表示会在多层注意力机制中被迭代更新, 从而使模型能够随着部分解的逐步构建, 持续捕捉不断变化的结构依赖关系。该动态重嵌入机制有助于学习对问题规模不敏感表示, 从而提升模型在更大规模问题上的泛化能力。

2.2. 混合方法

将基于学习的方法与局部搜索算法结合起来的混合方法已经在路径规划及车辆路径问题中得到了广泛研究[12], 其核心目标在于充分发挥两类方法的互补优势。神经网络模型通常用于捕捉问题的全局结构特征并生成具有潜力的初始解, 而局部搜索方法则通过挖掘邻域结构来对解进行进一步优化, 并保证解的可行性。通过将二者的有机结合, 混合方法在解的质量与鲁棒性方面通常优于单一的基于学习方法或传统启发式方法。另一类具有代表性的混合范式是将神经模型与局部改进策略相结合。不同于“学习增强启发式”方法, 该类研究主要依赖神经网络直接构造完整或部分解, 随后通过局部搜索过程进行优化。在此类框架中, 机器学习模型首先学习全局问题结构以生成高质量初始路径, 随后利用基于邻域的局部改进机制进一步提升解的质量并确保其可行性。相关研究表明, 这类方法能够有效弥合基于学习模型的代表能力与传统局部搜索算法强大优化能力之间的差距。

2.3. 应用于组合优化问题中的知识蒸馏

2.3.1. 传统的知识蒸馏框架

知识蒸馏作为一种将高性能教师模型中的知识迁移至轻量化学生模型的有效技术, 已经得到了广泛研究[13], 其目标在于提升模型的学习效率与泛化性能。该方法最初应用于分类任务, 随后逐步扩展至序列建模与结构化预测问题之中, 其在复杂决策场景中展现出良好的应用潜力。在组合优化领域, 知识蒸馏为高容量模型或专家求解器中的有效决策经验向高效学习框架的迁移提供了一种可行途径。传统知识蒸馏方法通常通过引入软化概率分布, 对教师模型与学生模型的输出分布进行对齐。具体而言, 通过最小化预测 logits 或概率向量之间的差异度量, 使学生模型在保持结构紧凑的同时逼近教师模型的决策行为。在此基础上, 知识蒸馏进一步拓展至特征层与中间表示层, 从而实现超越最终输出的深层知识迁移。

2.3.2. 面向组合优化问题的知识蒸馏技术

近期的研究探讨了知识蒸馏在路径规划和组合优化问题中的应用, 在这些问题中, 解决方案通常是在复杂约束条件下进行结构化的决策过程生成的。在车辆路径规划任务中, 知识蒸馏通常在策略决策层面进行, 使学生模型能够模仿由高质量的教师模型或专家求解器所产生的行为分布或路径偏好。通过这一过程, 学生模型可以隐性地学习路径启发式、长期依赖模式以及约束处理策略, 这些是仅通过基于最终解方案的直接监督难以捕捉的。然而, 将知识蒸馏扩展到路径规划问题中的非自回归或并行解码框架仍然具有挑战性。由于路径构建过程中没有明确的逐步反馈, 此类模型往往面临监督信号稀疏以及长距离依赖关系难以建模的问题, 往往会导致决策分布过于平滑以及解决方案的可靠性降低。

针对路径规划与组合优化问题的特点, 近年来已有研究开始探索面向策略学习的知识蒸馏方法。在车辆路径问题中, 解的构造通常涉及一系列在复杂约束条件下进行的顺序决策过程, 因此蒸馏过程往往作用于策略层或决策层, 使学生模型能够模仿教师模型或专家求解器在不同状态下给出的行动分布或路径选择偏好。通过策略级蒸馏, 学生模型不仅能够学习到高质量解对应的最终结果, 还可以隐式地获取路径构造过程中所包含的启发式规则、长期依赖关系以及约束处理策略。这种方式在一定程度上克服了仅依赖最终解进行监督所带来的信息不足问题, 为提升模型的泛化能力与鲁棒性提供了有效途径。然而, 在非自回归并行解码框架下, 知识蒸馏的应用仍面临较大挑战。由于此类模型在路径构造过程中缺乏显式的逐步反馈机制, 学生模型难以完整保留决策序列中的长程依赖关系与约束交互信息, 容易产生过于平滑的决策分布, 从而导致解的稳定性和可靠性下降。因此, 要想在该类框架中实现有效的知识蒸馏, 往往需要设计更加合理的蒸馏策略, 才能在保持推理效率优势的同时, 充分传递序列化路径决策所蕴含的关键信息。在 CVRP 问题中, 知识蒸馏亦得到广泛应用。Bi 等[14]提出自适应多分布知识蒸馏框架, 使学生模型能够从多个不同数据分布下训练得到的教师模型中学习, 从而提升其在多样化 VRP 实例上的泛化能力。Zheng 等[15]进一步引入多任务知识蒸馏, 通过跨任务迁移教师模型知识, 使学生模型在单一模型中具备多任务优化能力并生成高质量路径。Jin 等[16]将知识蒸馏建模为受约束学习过程, 通过引入教师模型的优化轨迹作为引导, 有效提升训练稳定性与知识迁移效果。上述方法在提升收敛速度与缓解过拟合方面表现出良好效果。

3. 数学建模

本节对 CVRP-SRP 问题进行介绍, 给出其完整定义, 并建立相应的数学模型。

本研究聚焦于无人配送车辆(UDV)的物流配送路径优化问题。传统容量约束车辆路径问题(CVRP)以满足所有客户需求为约束, 以最小化车辆总行驶距离为目标, 为车队设计最优配送路径方案。然而, 在实际无人配送场景中, 车辆难以对道路施工、交通拥堵及突发事件等动态路况进行有效感知, 不同路径之间的通行状态存在显著差异。这些因素可能导致仓库与客户之间以及客户之间的通行效率下降, 从而降低路径安全性并影响整体配送效果。因此, 仅以最短行驶距离作为优化目标难以满足 UDV 物流运行的实际需求。针对上述问题, 本文在传统 CVRP 模型基础上, 提出基于安全优先策略的容量约束车辆路径规划问题(CVRP-SRP)。在该模型中, 将道路安全性引入路径决策过程, 不再仅以距离最小化为唯一优化目标。具体而言, 通过对仓库与客户之间以及客户之间各路径的安全性进行评估, 构建安全系数矩阵以量化不同路径的安全水平。在此基础上, 对目标函数进行改进, 将路径距离与安全性进行加权融合, 并引入权重参数 w 以平衡距离优化与安全性能之间的权衡关系。CVRP-SRP 的数学模型如下:

$$f(s) = w \sum_{k \in K} \sum_{i \in V} \sum_{j \in V} d_{ij} x_{ij}^k + (1-w) \sum_{k \in K} \sum_{i \in V} \sum_{j \in V} S_{ij} x_{ij}^k \quad (1)$$

Table 1. Main symbols and their meanings

表 1. 主要符号及含义

符号	含义
i, j	节点索引, 表示配送网络中的节点
$V = \{0, 1, 2, \dots, n\}$	所有节点集合, 其中 0 表示配送中心
$N = \{1, 2, \dots, n\}$	客户需求节点集合
K	配送车辆集合
k	第 k 辆配送车辆
d_{ij}	节点 i 到节点 j 之间的行驶距离
q_i	客户节点 i 的配送需求量
Q	每辆配送车辆的最大载重容量
S_{ij}	节点 i 与节点 j 之间道路的安全系数
w	距离与安全之间的权重系数, $0 \leq w \leq 1$

CVRP-SRP 模型具有以下约束条件:

$$\sum_{k \in K} y_i^k = 1, \forall i \in N, \tag{2}$$

$$\sum_{j \in V} x_{ij}^k = \sum_{j \in V} x_{ji}^k, \forall i \in V, k \in K, \tag{3}$$

$$\sum_{i \in N} q_i y_i^k \leq Q, \forall k \in K, \tag{4}$$

$$x_{ii}^k = 0, \forall i \in V, k \in K \tag{5}$$

$$x_{ij}^k S_{ij} \leq 0.8, \forall i, j \in V, k \in K \tag{6}$$

主要符号如表 1 所示, 目标函数见式(1), 主要约束条件见式(2)~(6), 其余约束与传统 CVRP 基本一致。由于 CVRP-SRP 属于以安全性为优先的路径规划问题, 本文引入新的安全系数 S_{ij} , 用于评估仓库与客户之间以及客户与客户之间路径的安全水平。通过调节参数 w , 可以根据实际配送需求对路径进行动态优化。当货物运输需求较大、配送时效要求较高时, 可提高距离项的权重以加快配送速度; 相反, 当运输需求较小、时间较为充裕时, 则可提高安全性的权重, 从而优先选择更安全的路径。此外, 考虑到部分道路可能出现完全阻断或严重拥堵的情况, 导致其安全系数极低, 本文在模型中引入安全约束: 无论参数 w 如何取值, 均不选择安全性最低的 20% 路径, 以保证生成路径的可行性与安全性。基于上述数学模型, 本文旨在路径长度与安全性之间取得平衡, 并通过参数 w 实现对路径的实时调整。

在社区无人配送路径规划过程中, 道路环境的复杂程度会直接影响无人配送车辆的运行安全性。与传统车辆路径规划问题不同, 社区内部道路往往存在行人密集、道路宽度有限以及交叉路口较多等情况, 这些因素都会对无人配送车辆的行驶安全产生影响。因此, 在路径规划过程中, 有必要对不同道路的安全性进行定量评价, 并以安全系数对其进行描述。为更准确地反映社区道路的安全状况, 从多个影响因素出发对道路安全性进行综合评价。具体而言, 道路安全水平主要受到道路结构特征、交通环境以及外部干扰等多方面因素影响。例如, 道路宽度、行人密度以及交叉口数量等因素都会对无人配送车辆的行驶安全产生重要影响。为了综合考虑上述因素, 本文将道路安全水平抽象为安全系数 S_{ij} , 用于表示节点 i 与节点 j 之间道路的安全程度, 其取值范围为 $0 \leq S_{ij} \leq 1$, 数值越大表示道路安全性越高。

道路安全系数主要由道路几何特征、交通参与者动态、环境感知条件以及智能设施与路测条件等多个因素共同决定。通过对这些因素进行综合评估,可以得到对应道路的安全评价结果,并将其作为路径规划模型中的重要参数。通过引入道路安全系数,可以使路径规划模型在优化配送距离的同时,充分考虑道路运行环境对无人配送车辆安全性的影响,从而提高社区无人配送系统的整体运行可靠性。

尽管 CVRP-SRP 模型已经进行了较为完整的建模,但在实际的 UDV 物流场景中,高质量可行解数据的匮乏仍然是一个关键挑战。当前社区规模的无人配送仍处于初步部署阶段,实际运行过程中可获取的轨迹数据与需求数据较为稀缺且不完整,从而限制了数据驱动模型对有效路径模式及安全特性的学习能力。为缓解上述问题,本文引入知识蒸馏(KD)机制,将大容量教师模型中的知识迁移至学生模型。通过利用教师模型生成的高质量解对学生模型进行指导,使其在仅依赖少量真实数据的情况下仍能够获得较好的性能表现,从而提升模型在数据受限场景下的泛化能力与解的质量。

4. 引导式知识蒸馏框架构建

本节对提出的算法进行系统描述,其主要包括两个组成部分:解生成阶段与基于知识蒸馏的模型微调阶段。首先介绍本文的核心机制,即基于知识蒸馏的模型微调方法;其次对整体算法框架进行完整阐述。

4.1. 完整的算法框架

本文提出了一种由知识蒸馏驱动的求解框架,用于解决社区规模物流场景下的 CVRP-SRP 问题。整体算法流程如图 1 所示。该框架由两阶段构成:第一阶段是基于神经-启发式教师模型的解生成阶段,第二阶段是基于引导式知识蒸馏的模型微调阶段。在第一阶段,采用融合自回归 LEHD 模型与 ALNS 的 NANS 框架作为教师模型生成器。该阶段以传统 CVRP 为求解对象,通过神经模型生成初始解,并结合启发式搜索策略对路径进行优化,从而获得高质量路径解。在第二阶段,引入 GNARKD 机制,将教师模型中的路径决策知识迁移至轻量化学生模型。不同于直接使用教师模型进行部署,学生模型通过学习教师模型生成的动作序列与决策概率分布,对 CVRP-SRP 问题进行针对性微调。通过该过程,实现了从传统 CVRP 到安全约束路径规划问题的有效知识迁移,使学生模型在显著降低推理复杂度的同时,仍能够生成高质量且可行的路径解。

在所提出的框架中,结合了 LEHD 模型的 NANS 算法作为教师解生成器,用于为知识迁移提供高质量路径解。在该阶段,首先利用 LEHD 模型基于节点特征构造完整初始解,从而生成结构一致的路径序列;随后,引入自适应大邻域搜索(ALNS)对初始解进行进一步优化,其中通过自适应机制选择破坏-修复算子及移除比例,以提升解的质量。在迭代搜索过程中,接受准则由模拟退火策略控制,从而在避免早熟收敛的同时实现对解空间的有效探索。通过神经模型引导的初始解生成与启发式局部搜索的协同作用,最终获得稳定且高质量的教师解,为后续知识蒸馏阶段提供可靠的监督信号。

在获得教师框架生成的高质量路径解后,本文引入引导式知识蒸馏机制,将路径决策知识迁移至轻量化学生模型。不同于直接模仿完整路径,学生模型学习教师解中蕴含的决策信息,从而在路径构造过程中获得有效指导。教师模型生成的路径被转化为监督信号,用于引导学生模型在每一步选择可行且优质的决策。教师解由结合了 LEHD 的 NANS 框架生成,用以刻画复杂路径决策中的结构特征。在此基础上,引入 GNARKD 机制[17],将教师模型的决策分布蒸馏至学生模型,使其在有限训练数据条件下仍能够保持良好的泛化能力。通过这一过程,学生模型能够提取高质量路径中的关键结构模式。该方法在降低计算开销的同时,实现了高效的路径构造,能够适用于大规模及实时路径规划场景。同时,在保证解

质量的前提下, 减少了对大规模训练数据及复杂推理过程的依赖, 为社区级无人配送路径优化提供了一种可行方案。

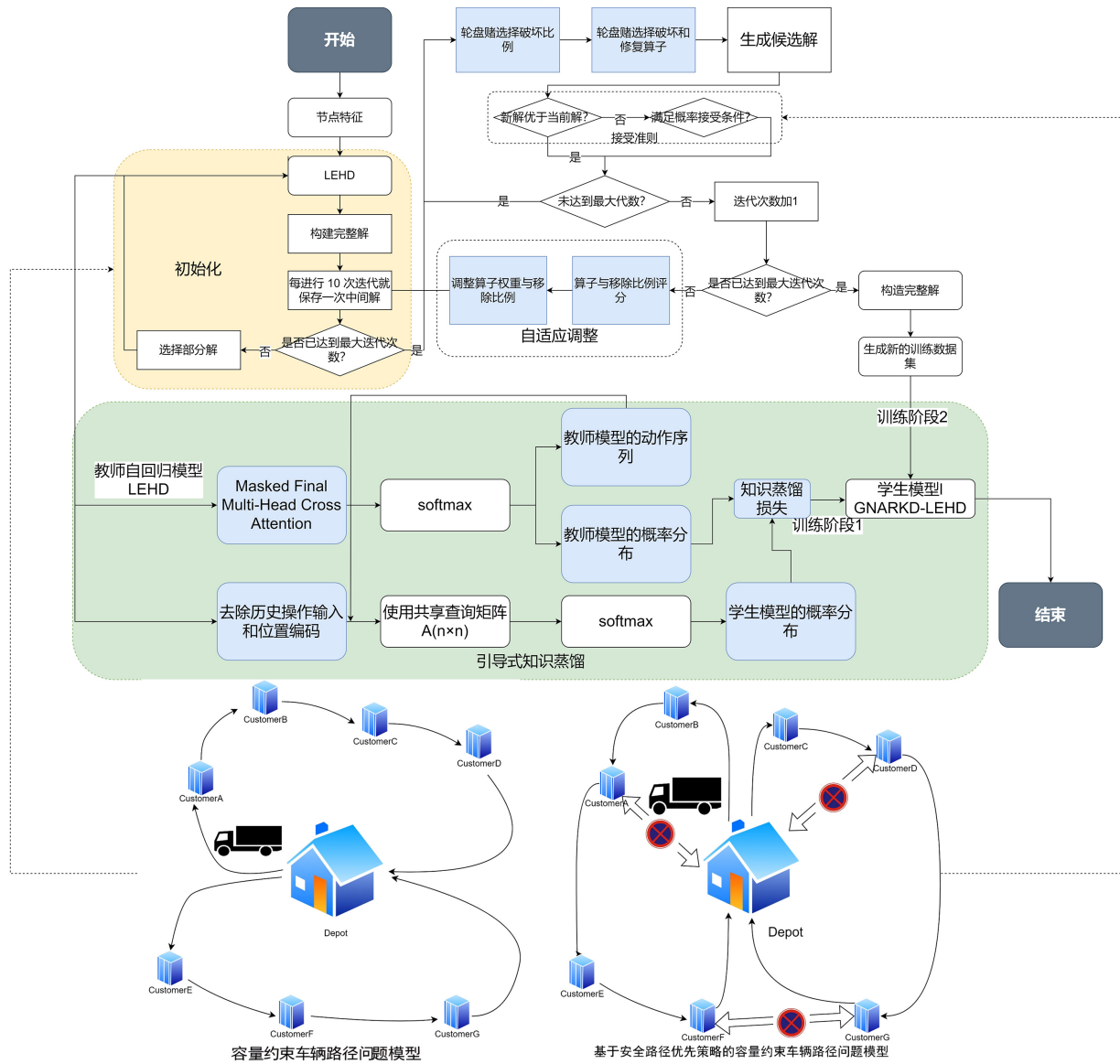


Figure 1. The flowchart of GNARKD
图 1. 引导式知识蒸馏的流程

4.2. 基于知识蒸馏的模型微调

在实际无人配送系统中, 路径求解模型的性能往往受到高质量训练数据匮乏以及计算效率要求的双重制约。在此背景下, 知识蒸馏(KD)为缓解上述问题提供了一种有效途径, 使轻量化模型在降低计算开销的同时, 仍能够学习并保留高性能模型中的关键决策机制。然而, 在车辆路径问题中, 传统非自回归模型由于难以刻画路径构造过程中的序列依赖关系, 通常表现出决策置信度不足及解质量下降的问题。针对这一局限, 本文引入 GNARKD 框架, 该方法是一种具有良好通用性的引导式知识蒸馏策略, 可将任意基于 Transformer 的自回归模型有效转换为高效的非自回归模型。具体而言, GNARKD 在保留教师自

回归模型编码器结构的基础上, 对解码器的输入与输出机制进行重新设计, 以保证序列信息在解码过程中得到有效建模。同时, 在蒸馏过程中引入引导信号, 增强学生模型对教师决策行为的拟合能力, 从而增强了学生模型复制教师决策过程的能力, 使其具有更高的准确性。

Algorithm 1 Fine-tuning of the model

Require: Teacher model LEHD; Student model GNARKD-LEHD; Training epoch E .

Ensure: Trained student model

1: **Phase 1: Knowledge Distillation**

2: **for** $epoch = 1$ to E_1 **do**

3: Generate student sequence based on current model parameters.

4: Align student sequence with teacher sequence.

5: Evaluate distillation loss.

6: Backpropagate loss and update student model parameters.

7: **end for**

8: **Phase 2: Supervised Training**

9: **for** $epoch = 1$ to E_2 **do**

10: Optimize GNARKD-LEHD with supervised loss on the new dataset.

11: **End for**

12: Return the trained SA-GNARKD model.

GNARKD-LEHD 整体训练过程的伪代码如算法 1 所示, 该算法详细描述了整个训练过程。它被分为两个阶段。在第一阶段, 学生模型会逐步学习模拟教师模型的行为(LEHD)。在每次训练周期中, 学生根据其当前参数生成候选序列, 然后将其与教师的序列进行对齐。两者的差异通过一种蒸馏损失函数来量化, 该损失引导学生模仿教师的序列决策过程。这种迭代优化——包括损失评估、反向传播和参数更新——使学生能够逐渐继承教师模型中嵌入的关键知识, 即使在有限的训练数据下也是如此。在第二阶段, 从第一阶段获得的 GNARKD-LEHD 模型通过监督训练进一步得到完善。在此过程中, 模型被引导在保留第一阶段获得的决策能力的同时整合关于安全性的知识。因此, GNARKD-LEHD 能够生成同时考虑距离和安全性的路线, 生成的解决方案能够在这两项标准之间实现平衡的权衡。

4.3. 引导式知识蒸馏的模型结构设计

在所提出的 NANS 框架中, 采用了 GNARKD 策略, 将路由决策知识从一个自回归(AR)教师模型转移到一个轻量级的非自回归(NAR)学生模型。接下来, 将详细介绍 GNARKD 的架构设计及其相关的学习过程。

教师模型遵循一种传统的基于 Transformer 的自回归解码模式, 在这种模式中, 路由解决方案是按顺序构建的。在每个决策步骤 t 中, 解码器会根据编码后的节点表示以及先前选定节点的动作历史生成一个动作概率分布。教师模型生成的候选动作的概率分布定义为:

$$\lambda_t^{te} = \sigma_{T_t} \left(f^{te} \left(\begin{cases} z, & \text{if } t=1 \\ CE(a_{t-1}), & \text{if } t>1 \end{cases}, h^{te} \right) \right) \quad (7)$$

其中, $\lambda_t^{te} \in R^{1 \times n}$ 表示在第 t 步对所有候选节点的选择概率, σ_{T_t} 表示温度为 T_t 的 Softmax 函数, f^{te} 表示教师解码器的前向传播函数, $z \in R^{d_h \times 1}$ 是一个可学习的参数, 用作输入占位符, 而 $CE(a_{t-1})$ 表示由动作历史 a_{t-1} 构建的上下文嵌入, h^{te} 为编码器输出的节点特征表示。这种自回归形式使教师模型能够明确捕捉高质量路径解决方案中固有的序列依赖性和结构排序信息。因此, 生成的动作序列及其相关的概率分布为后续的知识蒸馏提供了有价值的监督信号。

在非自回归模型中, 学生解码器不再生成逐步的动作概率值。相反, 它旨在估计每个节点与其潜在后继节点之间的成对兼容性, 从而以完全并行的方式建模节点之间的相关性。为此, 学生解码器输出一个亲和矩阵 A , 其中每个元素 A_{ij} 反映从节点 i 转移到节点 j 的关联程度。

为了高效地计算这些成对的关联程度, 所有节点的表示都会通过学生解码器的共享前向传播函数进行同时处理, 从而实现并行解码。具体而言, 将仓库节点与所有客户节点的嵌入表示进行拼接, 并统一映射到解码器的输入空间, 仅仅由解码器通过一次前向传播完成整体计算:

$$A = f^{st} \left(W \left[z, h_1^{st}, \dots, h_n^{st} \right] \right), \quad A \in \mathbb{R}^{(n+1) \times n} \quad (8)$$

其中, $f^{st}(\cdot)$ 表示学生解码器的前向传播函数, 它将输入节点表示映射为节点之间的亲和度得分。向量 $z \in \mathbb{R}^d$ 表示仓库节点的嵌入, 而 $h_i^{st} \in \mathbb{R}^d$ 表示由学生编码器生成的第 i 个客户节点嵌入, 其中 $i \in \{1, \dots, n\}$, 且已嵌入维度为 d 。最终的映射 $W \in \mathbb{R}^{d \times d}$ 是一个学习的线性变换矩阵, 它将连接后的节点嵌入转换为为学生解码器的输入空间。运算符 $[\cdot, \dots, \cdot]$ 表示向量的水平拼接, 从而得到一个维度为 $(n+1) \times d$ 的矩阵。

输出矩阵 $A \in \mathbb{R}^{(n+1) \times n}$ 表示各对查询节点之间的亲和度得分, 其中每一行对应一个查询节点(包括仓库和所有客户节点), 每一列对应一个候选客户节点。每个元素 A_{ij} 表示查询节点 i 与客户节点 j 之间的预测亲和度, 反映了将节点 j 选为路径序列中下一个节点的可能性。在学生解码器中, 会对每个节点嵌入应用共享的线性变换, 以建立节点表示之间的空间联系, 并通过自注意力机制建模节点之间的复杂依赖关系。通过这种方式, 模型能够在一次前向传播中计算所有节点与其潜在后继节点之间的关联程度。同时, 该模型结构通过节点间结构而非解码顺序隐式编码路径信息, 因此有意省略了位置编码。此外, 传统自回归模型中常用的基于掩码的注意力机制也被移除, 因为路径构建的可行性通过解码器后的处理来保证, 而非依赖于严格的顺序约束。

虽然 NAR 的形式能够实现并行传播, 但它也消除了自回归解码中所存在的明确的顺序依赖关系。为了弥补这一损失并保留教师解码过程中所嵌入的基于顺序的决策知识, 在训练过程中引入了一种引导式解码机制。具体而言, 学生模型需要遵循由教师生成的相对动作序列 π_{AR} 。对于每个动作 $a_t \in \pi_{AR}$, 应用温度为 T_2 的 softmax 函数 σ_{T_2} 来对选定节点与每个候选节点 v_i 之间的紧密度 $A_{a_t, i}$ 进行归一化, 从而得到学生在时间 t 的动作概率分布:

$$\lambda_t^{st} = \sigma_{T_2} \left(A_{a_t, \cdot} \odot \begin{cases} -\infty, & v_i \in set_t \\ 1, & \text{otherwise} \end{cases} \right) \quad (9)$$

其中, $\lambda_t^{st} \in \mathbb{R}^{1 \times n}$ 表示学生模型在第 t 步的节点概率选择分布, $A_{a_t, i}$ 表示从当前节点 a_t 转移至候选节点 v_i 的关联强度, σ_{T_2} 为带温度参数的 softmax 函数, $M_t(i)$ 为掩码参数, 用于约束访问不可访问节点, 其定义为:

$$M_t(i) = \begin{cases} -\infty, & v_i \in set_t, \\ 1, & \text{otherwise,} \end{cases} \quad (10)$$

其中, set_t 表示在第 t 步不可访问的节点集合, 包括已访问节点或不满足约束条件的节点。通过该掩码机制, 可有效保证生成路径的可行性。

学生模型并非直接模仿最终的路径解决方案, 而是在解决方案构建过程中通过训练使其决策分布与教师模型的决策分布相一致。具体而言, 教师的动作概率分布 λ_t^{te} 被用作监督信号, 以构建一个保留了序列决策中隐含顺序依赖信息的代理分布。具体而言, 通过最小化教师模型与学生模型在整个路径生成过程中的概率分布差异, 使学生模型能够学习教师模型在不同决策阶段的行为特征。为此, 采用 KL 散度作为蒸馏损失函数, 其表达式为:

$$\begin{aligned}
L_{KD} &= E_{\pi_{AR} \sim P_{AR}(\cdot|s)} \left[KL(P_{AR}(\pi_{AR} | s) \| P_{NAR}(\pi_{AR} | s, \theta)) \right] \\
&= \frac{1}{B} \sum_{t=1}^B \sum_{i=1}^l \sum_{j=1}^n \lambda_{t,i}^{te} \log(\lambda_{t,i}^{te} - \lambda_{t,i}^{st})
\end{aligned} \tag{11}$$

其中, B 表示批量大小, l 表示路径长度, $\lambda_{t,i}^{te}$ 与 $\lambda_{t,i}^{st}$ 分别表示教师模型与学生模型在第 t 步选择节点 v_j 的概率。上述蒸馏损失函数通过对齐教师与学生每一步决策中的概率分布, 使学生模型能够学习到教师模型在路径构造过程中所蕴含的序列结构信息与决策偏好。与仅对最终路径进行监督的方式相比, 该方法能够更加细粒度地刻画路径生成过程, 从而有效提升学生模型的表达能力与泛化性能。

5. 实验结果与性能对比

在本节中, 对 NANS 框架进行了评估, 并将其与基于学习的算法和启发式算法进行了比较。此外, 还对知识蒸馏机制的有效性进行了评估。

5.1. 实验设置

所有实验所使用的数据集均源自文献[11]。训练集包含一百万个 CVRP 问题实例, 而测试集则包含 10,000 个实例。所有实验均在配备 NVIDIA RTX 4090 GPU 的工作站上进行。LEHD 模型训练了 40 个周期, 总训练时间约为 10 小时。除非另有说明, 所有报告的结果都是基于多次独立运行的平均值得出的。

5.2. 初始化模型对比

为了验证所提出的将机器学习与启发式优化相结合的混合框架的有效性, 以及通过 LEHD 模型生成多个初始解的优势, 在初始化阶段使用了几个替代的神经组合优化模型来取代 LEHD。具体而言, 选择了 BQ [18]、MDAM [19] 和 POMO [20] 作为基准模型。

表 2 展示了在三种不同的权重设置下, 具有 100 个客户节点的 CVRP-SRP 实例的实验结果。所有方法均采用了相同的 ALNS 阶段进行公平比较。

Table 2. Experimental results of NANS with different neural combinational optimization engines

表 2. 使用不同神经组合优化模型的实验结果

Model	$w = 0.2$	$w = 0.5$	$w = 0.8$
LEHD	26.24	27.59	20.08
BQ	27.02	27.70	22.46
MDAM	27.63	28.45	22.94
POMO	27.72	28.49	25.88
Greedy Generation	33.12	29.49	25.97

从结果中可以看出, 在生成多个初始解时, LEHD 总是比其他神经模型表现更出色。这证实了 LEHD 在 CVRP-SRP 中具有特别出色的表现, 并且所提出的机器学习辅助初始化策略既有效又稳健。

5.3. 知识蒸馏的有效性

为了进一步评估知识蒸馏的效果, 还进行了一个额外的实验, 以比较蒸馏前后的表现。结果(总结于表 3 中)表明, 尽管蒸馏后的模型在解决方案质量方面略逊于教师模型, 但其计算效率有了显著提升。

Table 3. Comparison results of knowledge distillation experiments
表 3. 知识蒸馏实验的比较结果

Model	Reasoning Method	Learning Gap	Acceleration Ratio
AM	Greedy	12.61%	≈42×
AM	Sampling (100)	4.87%	≈37×
POMO	Greedy (No Augment)	8.80%	≈35×
POMO	Greedy (×8 Augment)	5.62%	≈30×
LEHD	Greedy	11.10%	≈22×

为了进一步评估 GNARKD-LEHD 的性能, 将其与 LEHD + NANS 以及一些经典的启发式算法进行了比较, 这些算法包括遗传算法(GA)和蚁群优化算法(ACO)。由于原始的 GA 和 ACO 是以标准的 CVRP 设计的, 因此这两种算法都被扩展以纳入安全系数和 CVRP-SRP 的修改目标函数。

表 4 汇总了在三种权重设置下的比较结果。对于每种算法, 所报告的数值包括最终的目标值和平均运行时间。

Table 4. Experimental comparison results with heuristic algorithms
表 4. 与启发式算法的实验对比结果

Algorithm	$w = 0.2$	$w = 0.5$	$w = 0.8$
NANS	26.24/9.97 s	27.59/6.07 s	21.08/7.63
GNARKD	26.43	28.92	22.55
GA	26.47/1 m 40 s	30.72/1 m 45 s	26.32/1 m 43 s
ACO	41.39/58 s	33.93/56 s	24.23/56 s

6. 结论

在本文中, 提出了一种基于知识蒸馏的神经启发式框架, 用于解决在有限数据条件下基于社区的 UDV 路由优化问题。并采用 CVRP-SRP 模型来描述具有安全意识和容量限制的社区物流场景。在此基础上, 通过将 LEHD 模型和 ALNS 融合, 引入引导式知识蒸馏策略, 进一步扩展了神经启发式优化框架。通过将决策策略从高质量的教师模型转移到轻量级的学生模型, 能够在小样本设置中实现方案质量、鲁棒性和数据效率的提升。实验结果表明, 所提出的方法优于未进行蒸馏和直接转移的模型, 表明其在实际的社区物流中具有应用价值。在未来的工作中, 将进一步探索改进措施, 以提高所提框架的适应性和计算效率。知识蒸馏过程可以扩展到更适应性的重构或学习机制, 以进一步降低计算成本。此外, 还将研究基于生成的高质量路线构建新训练数据集以及重新训练路由模型的可行性。这些研究方向有望进一步推动神经启发式优化方法在数据稀缺的无人配送路由问题中的应用。

参考文献

- [1] Gendreau, M., Hertz, A. and Laporte, G. (1994) A Tabu Search Heuristic for the Vehicle Routing Problem. *Management Science*, **40**, 1276-1290. <https://doi.org/10.1287/mnsc.40.10.1276>
- [2] Clarke, G. and Wright, J.W. (1964) Scheduling of Vehicles from a Central Depot to a Number of Delivery Points. *Operations Research*, **12**, 568-581. <https://doi.org/10.1287/opre.12.4.568>
- [3] Helsgaun, K. (2017) An Extension of the Lin-Kernighan-Helsgaun TSP Solver for Constrained Traveling Salesman and Vehicle Routing Problems. Roskilde University.

-
- [4] Lee, C.Y., Lee, Z.J., Lin, S.W., *et al.* (2010) An Enhanced Ant Colony Optimization (EACO) Applied to Capacitated Vehicle Routing Problem. *Applied Intelligence*, **32**, 88-95. <https://doi.org/10.1007/s10489-008-0136-9>
- [5] Tavakkoli-Moghaddam, R., Safaei, N. and Gholipour, Y. (2006) A Hybrid Simulated Annealing for Capacitated Vehicle Routing Problems with the Independent Route Length. *Applied Mathematics and Computation*, **176**, 445-454. <https://doi.org/10.1016/j.amc.2005.09.040>
- [6] Jasim, A.N. and Chaari Fourati, L. (2024) Guided Genetic Algorithm for Solving Capacitated Vehicle Routing Problem with Unmanned-Aerial-Vehicles. *IEEE Access*, **12**, 106333-106358. <https://doi.org/10.1109/access.2024.3438079>
- [7] Hu, K.C., Tsai, C.W. and Chiang, M.C. (2020) A Multiple-Search Multi-Start Framework for Metaheuristics for Clustering Problems. *IEEE Access*, **8**, 96173-96183. <https://doi.org/10.1109/access.2020.2994813>
- [8] Zhao, J., Mao, M., Zhao, X. and Zou, J. (2021) A Hybrid of Deep Reinforcement Learning and Local Search for the Vehicle Routing Problems. *IEEE Transactions on Intelligent Transportation Systems*, **22**, 7208-7218. <https://doi.org/10.1109/tits.2020.3003163>
- [9] Herdianto, B., Billot, R., Lucas, F., *et al.* (2025) Edge-Selector Model Applied for Local Search Neighborhood for Solving Vehicle Routing Problems. arXiv:2508.14071.
- [10] Cheng, C.A., Kolobov, A. and Swaminathan, A. (2021) Heuristic-Guided Reinforcement Learning. *Advances in Neural Information Processing Systems*, **34**, 13550-13563.
- [11] Lin, X., Liu, F., Luo, F., Wang, Z. and Zhang, Q. (2023) Neural Combinatorial Optimization with Heavy Decoder: Toward Large Scale Generalization. *Advances in Neural Information Processing Systems* 36, New Orleans, 10-16 December 2023, 8845-8864. <https://doi.org/10.52202/075280-0387>
- [12] Vidal, T. (2022) Hybrid Genetic Search for the CVRP: Open-Source Implementation and SWAP Neighborhood. *Computers & Operations Research*, **140**, Article 105643. <https://doi.org/10.1016/j.cor.2021.105643>
- [13] Das, S., Camporese, G., Cheng, S. and Ballan, L. (2024) Distilling Knowledge for Short-to-Long Term Trajectory Prediction. 2024 *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Abu Dhabi, 14-18 October 2024, 13001-13008. <https://doi.org/10.1109/iros58592.2024.10801315>
- [14] Bi, J., Cao, Z., Chee, Y.M., Chen, J., Ma, Y., Sun, Y., *et al.* (2022) Learning Generalizable Models for Vehicle Routing Problems via Knowledge Distillation. *Advances in Neural Information Processing Systems* 35, New Orleans, 28 November-9 December 2022, 31226-31238. <https://doi.org/10.52202/068431-2264>
- [15] Zheng, Y., Luo, F., Wang, Z., Wu, Y. and Zhou, Y. (2025) MTL-KD: Multi-Task Learning via Knowledge Distillation for Generalizable Neural Vehicle Routing Solver. arXiv:2506.02935.
- [16] Jin, X., Peng, B., Wu, Y., Liu, Y., Liu, J., Liang, D., *et al.* (2019) Knowledge Distillation via Route Constrained Optimization. 2019 *IEEE/CVF International Conference on Computer Vision (ICCV)*, Seoul, 27 October-2 November 2019, 1345-1354. <https://doi.org/10.1109/iccv.2019.00143>
- [17] Xiao, Y., Wang, D., Li, B., Wang, M., Wu, X., Zhou, C., *et al.* (2024) Distilling Autoregressive Models to Obtain High-Performance Non-Autoregressive Solvers for Vehicle Routing Problems with Faster Inference Speed. *Proceedings of the AAAI Conference on Artificial Intelligence*, **38**, 20274-20283. <https://doi.org/10.1609/aaai.v38i18.30008>
- [18] Andreoli, J., Drakulic, D., Mai, F., Michel, S. and Sors, A. (2023) BQ-NCO: Bisimulation Quotienting for Efficient Neural Combinatorial Optimization. *Advances in Neural Information Processing Systems* 36, New Orleans, 10-16 December 2023, 77416-77429. <https://doi.org/10.52202/075280-3385>
- [19] Xin, L., Song, W., Cao, Z. and Zhang, J. (2021) Multi-Decoder Attention Model with Embedding Glimpse for Solving Vehicle Routing Problems. *Proceedings of the AAAI Conference on Artificial Intelligence*, **35**, 12042-12049. <https://doi.org/10.1609/aaai.v35i13.17430>
- [20] Kwon, Y.D., Choo, J., Kim, B., *et al.* (2020) Pomo: Policy Optimization with Multiple Optima for Reinforcement Learning. *Advances in Neural Information Processing Systems*, **33**, 21188-21198.