逐边拓扑发散差距：引导组合优化中的搜索

Ilya Trofimov Skoltech 莫斯科，俄罗斯

Daria Voronkova Skoltech、AIRI 莫斯科，俄罗斯

Alexander Mironenko Skoltech 莫斯科，俄罗斯

Anton Dmitriev Skoltech 莫斯科，俄罗斯

Eduard Tulchinskii Skoltech、AI Foundation and Algorithm Lab，俄罗斯莫斯科

Evgeny Burnaev Skoltech、AIRI，俄罗斯莫斯科

Serguei Barannikov Skoltech、CNRS，俄罗斯莫斯科

arXiv:2512.15800v1 [cs.CG] 16 Dec 2025

摘要

我们通过分析旅行商问题（TSP）中巡回路线与最小生成树（MST）之间的散度，引入了一种拓扑反馈机制。我们的主要贡献是提出了一个典范分解定理，将巡回路线与最小生成树之间的差距表示为 RTD-Lite 条形码中的逐边拓扑散度差。基于此，我们为 2-opt 和 3-opt 启发式算法设计了拓扑引导方法，从而提升其性能。我们针对基于热图的方法、TSPLIB 和随机实例获得的巡回路线进行精细优化，并开展了实验。实验表明，在许多情况下，拓扑引导的优化能够取得更好的性能并更快收敛。

1 引言

旅行商问题（TSP）仍然是组合优化的基石，也是机器学习增强求解器的试验场。尽管神经构造策略（Vinyals等人，2015年；Kool 等，2019）和强化学习（RL）改进方法（da Costa 等，2020年；Chen 等，2024年）取得了令人瞩目的进展，当前最先进的流程仍然依赖盲目的事后局部搜索（例如 2-opt），以获得具有竞争力的巡回路线长度。这些移动在缺乏原则性指导的情况下检查 $O(n^2)$个边对，并在中等规模实例上占据了大部分运行时间。

基于近期图的拓扑数据分析方面的进展，尤其是*RTD-Lite* 条形码（Tulchinskii 等，2025 年），我们证明了一个**典范分解定理**：通常的巡回路径-最小生成树长度间隙等于一系列非负的、逐边的拓扑散度差之和。每个差距都是 RTD-Lite 条形码中一条条的长度，并明确指出单条巡回边偏离其唯一关联的最小生成树边的程度。

旅行商问题（TSP）通常被表述为纯粹的成本最小化任务：几乎所有启发式算法、精确求解器和学习方法都仅以巡回路线的总长度来衡量进展。这种视角忽视了随着巡回路线不断构建，城市之间是如何连接的，尽管数十年的几何学与聚类研究表明，优质巡回路线首先会遵循实例的内在连通性，然后才会削减剩余距离（Zahn，1971；Held & Karp，1970）。因此，当代神经求解器耗费大量计算进行盲目的局部搜索，以修正学习过程中未被发现的长程“拓扑错误”（Kool 等，2019；Fan 等，2024）。

拓扑学作为一种缺失的优化信号。 我们认为，明确的拓扑反馈可以弥合这一差距。借助图拓扑学的最新进展，我们采用拓扑散度框架（Tulchinskii 等，2025）来比较候选巡回路线与该实例的最小生成树（MST）。我们的主要定理构造了巡回边与最小生成树边之间的典范双射，将经典的巡回路径-最小生成树长度间隙分解为非负的逐边拓扑发散差距。每个差距都量化了最小生成树聚类通过巡回路线连接时的尺度差异

连接的情况：较大的条表明，巡回边连接相距较远的顶点簇的时间远晚于最小生成树。通过检查这些条，我们可以在构造算法或改进算法的每一步中，知道哪条边对自然连通结构的扭曲最大。

从长度优化迈向拓扑感知优化。 基于这一洞见，我们提出了两种轻量级机制：(i) 一种拓扑引导的*2-opt* 方法，始终优先处理散度最高的边，从而减少局部搜索交换；(ii) 一种间隙塑形的*Q*学习 框架，其奖励结合了巡回路线长度与预测的拓扑间隙，从而加快收敛并获得更优的最终巡回路线。至关重要的是，这两种方法的额外耗时均为 $O(n^2)$，与距离矩阵的构建成本相当。

我们在TSPLIB和随机欧几里得基准上的实验表明，加入这一单一拓扑信号，首次以实证验证了持久同调理论与大规模组合优化之间的联系。我们相信，这一拓扑学视角及其带来的实际收益，将为超越仅以长度为目标的优化开辟一个充满希望的研究方向。

据我们所知，这是首项(i) 提供巡回路线次优性的逐边拓扑刻画，并且(ii) 展示其在经典TSP求解器和基于学习的TSP求解器中均具实用价值的工作。

我们的贡献总结如下：

- **规范巡回路径-最小生成树分解**。我们证明了一个定理，建立巡回边与最小生成树边之间的一一映射，并将巡回路径-最小生成树长度间隙分解为一系列非负的、针对各边的散度值（持久条带长度）之和。这为量化每条巡回边对次优性的贡献提供了合理依据。
- **与 LKH 的 α-Score 的联系**。我们表明，LKH-3 TSP求解器中使用的知名 α-score，在数学上等价于 RTD-Lite 条形码度量的一种特定变体。这将我们的拓扑学视角与一种成功的传统启发式算法联系起来。
- **拓扑引导的局部搜索**。我们提出拓扑引导的2-opt 和 3-opt 启发式算法，优先移除散度差距较大的边。在随机欧几里得实例（最多包含 300 个节点）和 TSPLIB基准上，与标准 2-opt/3-opt 相比，该策略始终能找到更短的巡回路线，并且通常以更少的迭代次数收敛。
- **与神经求解器的集成**。 我们证明，拓扑引导的2-opt能够显著提升近期基于热图的TSP求解器生成巡回路线的精细优化效果，即使面对超大规模实例（最多包含 10,000 个节点）也不例外。值得注意的是，与对这些神经网络输出采用普通的2-opt后处理相比，我们的方法耗时更短，得到的巡回路线也更优。
- **拓扑塑形强化学习**。我们为一个基于Q学习的TSP求解器设计了RTDL间隙塑形奖励，使强化学习智能体能够根据拓扑散度获得逐边反馈。这种奖励塑形能够稳定训练，并且与纯粹基于长度的奖励相比，以更少的回合将最终最优性间隙降低了 25% 以上。

总而言之，我们的工作将拓扑学与大规模组合优化联系起来，提供了一种超越传统基于距离的目标函数的新型优化信号。我们相信，这种拓扑学视角为在TSP及相关问题中开发更具结构性和效率更高的搜索策略开辟了有前景的方向。

2 相关工作

经典旅行商问题启发式算法。 2-opt、3-opt 和 Lin–Kernighan（LK）等局部改进方案仍是高质量 TSP 求解器的核心方法（Croes，1958年；Lin & Kernighan，1973年；Helsgaun，2000年）。LK 和 LKH 依靠巧妙的邻域剪枝来缩小搜索空间。最小生成树和 1-树界长期以来一直用于估算全局成本（Held & Karp，1970），但尚未用于优先选择特定的移动。Aloise等人（2025年）以及 Hansen & Mladenovi´c（2006年）研究了 2-opt 算法中首次改进与最佳改进的比较。

神经网络与强化学习方法。 指针网络（Vinyals 等，2015年）、注意力模型（Kool 等，2019年）、群体强化学习（Team，2023年）和基于扩散的方法（Sanokowski等人，2024年）能够生成近最优巡回路线，但最终的优化工作仍交由经典局部搜索完成，而这占用了大部分

墙钟时间（Fan 等，2024）。直接学习局部移动的尝试（da Costa 等，2020年）仍然缺乏一种有原则的信号，无法判断接下来应更改哪条边。

图的拓扑数据分析（TDA）。持久性条形码捕捉多尺度连通性；RTD（Barannikov 等人）及其高效变体 RTD-Lite 可在 $O(n^2)$ 时间内比较两个加权图（Tulchinskii 等，2025）。以往工作将 RTD 用作标量损失，以在神经表示中保持拓扑结构；我们的工作首次利用其边级信息驱动组合求解器。

拓扑感知优化。通过最小生成树边移除进行图聚类可追溯至（Zahn，1971）。近期的"热力图+蒙特卡罗树搜索"流程（Xia 等，2024年；Sun 和 Yang，2023年；Fu 等，2021年；Qiu 等，2022年；Min 等，2023年）虽考虑了结构线索，但仍缺乏对逐边错位的严格度量。我们通过为每条边提供精确且可计算的散度来弥补这一差距，并展示其如何同时驱动局部搜索和强化学习。

总而言之，本文搭建了拓扑数据分析（TDA）与组合优化之间的桥梁：我们提出了一种具有理论依据的逐边误差信号，并证明它能切实提升经典 TSP 求解器和基于学习的 TSP 求解器的性能。

3 方法

考虑无向加权完全图 $G = (V, E)$，边 e 的权重为 $w(e)$。令 T_{tour} 为候选旅行商问题巡回路线。暂且移除最长的巡回边 e_{max}，并将 T_{tour} 视为一条从起始城市 s 到终点城市 t 的哈密顿路径 T，有时称为 (s, t)-巡回路线。令 T_{mst} 为 G 的最小生成树（MST）。我们首先根据边集及其权重，形式化 T_{tour} 与 T_{mst} 之间的关系：

定理1。 我们构造 ϕ，使 T 与 T_{mst} 的边之间存在一一对应，从而将标准的（*TSP*巡回路线）-*MST*差距 $L_{(s,t)-tour} - L_{mst}$ 分解为逐边非负拓扑散度差距的自然和：

$$L_{(s,t)-tour} - L_{mst} = \sum_{e \in E(T_{mst})} w(\phi(e)) - w(e) \tag{1}$$

$$w(\phi(e)) - w(e) \geq 0, \text{ for any } e \in E(T_{mst}). \tag{2}$$

换言之，每条最小生成树边 e 都与一条唯一的巡回边 ϕ（e）配对，其权重不小于 w（e），并且这些权重差之和恰好等于 （s, t）-巡回路线长度与最小生成树长度之间的差距。

证明。为简单起见，假设并集 $E(T_{mst}) \cup E(T)$ 中的所有边具有不同的权重；若有些权重相同，则只需为权重相同的边指定一个顺序。对于$e \in E(T_{mst})$，令 $A, B \in V$ 表示 e 的端点。令 $\tilde{e} \in E(T)$ 表示 T 中权重最小的边，使得将权重小于或等于 $w(\tilde{e})$ 的 $T-$ 条边加入 $T_{mst}^{w<w(e)}$ 后，A 和 B 连通。特别地，A 和 B 由一条路径连接，该路径由权重小于 $w(e)$ 的 $T_{mst}-$ 条边，以及权重大于或等于 $w(e)$ 且小于或等于 $w(\tilde{e})$ 的 $T-$ 条边构成。注意，这样的路径通常不唯一，但根据构造，它总是包含边 $\tilde{e}$。然后定义 $\phi(e) = \tilde{e}$。反过来，给定一条边 $\tilde{e} \in E(T)$，令 A, B 为其端点。令 $e' \in E(T_{mst})$ 表示 T_{mst} 中权重最小的边，使得将权重小于或等于 $w(e')$ 的 $T_{mst}-$ 条边加入 $T^{w<w(\tilde{e})}$ 后，A, B 可由一条路径连接。于是，对应关系 $\psi: \tilde{e} \mapsto e'$ 是 ϕ 的逆映射，见附录B。因此

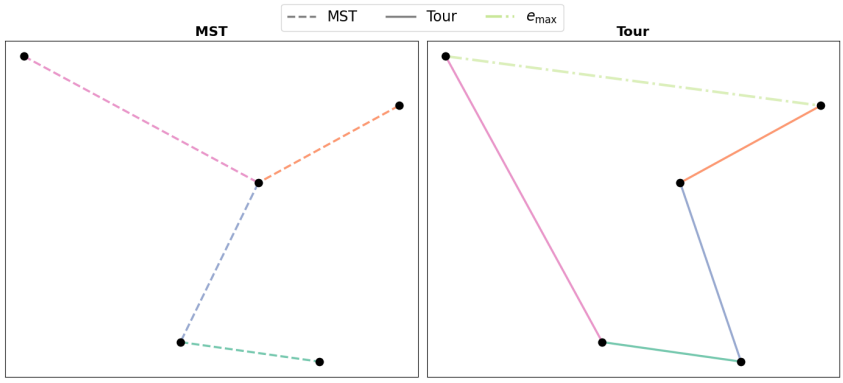
$$\phi : E(T_{mst}) \to E(T), \tag{3}$$

是一一对应，且（2）成立。 □

根据 Tulchinskii 等，2025 对 (T_{tour}, G) 的 RTDL-条形码的定义，可知它是以下区间序列：

$$\text{RTDL-Barcode}(T_{\text{tour}}, G) = \{(w(e), w(\phi(e))) \mid e \in E(T_{mst})\}.$$

哈密顿回路的应用。设 T_{tour} 为一个哈密顿回路。根据定义，RTDL-条形码(A, B)是针对两个具有双射关系的无向加权图 A, B 计算得出的



Figu图 1：哈密顿回路的边 T_{tour} 与最小生成树 T_{mst}之间的映射可视化。每条巡回边（ $e_{\max}$除外）都与一条最小生成树边一一匹配。

顶点。其计算过程（算法1，（Tulchinskii 等，2025））涉及 A，B的最小生成树。为了将拓扑信息纳入旅行商问题求解器，我们计算 RTDL-条形码(T_{tour}，G)。 T_{tour} 的最小生成树为 $T_{\text{tour}} \setminus \{e_{max}\}$，其中 e_{max} 是 T_{tour} 中权重最大的边。定理1蕴含了 ψ 之间的一个双射： $E(T_{\text{tour}})$与 $E(T_{mst})$。对于每个 $e \in E(T_{\text{tour}})$，我们定义惩罚$p(e) = w(e) - w(\psi(e)) \geq 0$，用以衡量其“劣度”，即它比最小生成树边长多少。该过程未定义 e_{max} 的惩罚。我们设$p(e_{max}) = \min_{e \in \{x \in E(T_{tour}) \mid p(x) > 0\}} p(e)$。这一取值在计算实验中表现最佳。图1展示了所述最小生成树边与巡回边之间映射关系的可视化。

RTDL 条形码与 α-分数之间的关系

α-分数用于 LKH-3 算法（Helsgaun，2000）中旅行商问题和车辆路径问题的边选择。 α-分数是衡量边“糟糕程度”的另一种指标。设 G 为一个权重为$w_{i,\ j}$, $G_1 \subset G$ 的加权图，其中包含一个任意顶点“1”及其关联边的子图。对于一条边（$i,\ j$）：

$$\alpha(i,j) = L(R^{+}(i,j)) - L(R), \tag{4}$$

其中， R 是 G的最小1-树，而 $R^{+}(i,\ j)$是 G 的最小1-树，并要求其包含边（$i,\ j$）； $L(\cdot)$是一个长度函数。我们建立了 α-分数与 RTDL 条形码之间的联系，如下所示

命题1。 对于 $i,\ j \neq 1$， $\alpha(i,\ j)$等于 RTDL-条形码($G^{i,\ 1^{j}}$， G_1)中一条对应于边（$i,\ j$）的条的长度。 G_1和 $G^{i,\ 1^{j}}$ 都是由 G 中的所有顶点（不含顶点1）构成的图。 G_1 定义如上，而 $G^{i,\ 1^{j}}$ 仅有一条边（$i,\ j$），其权重为$w_{i,\ j}$.

证明见附录 C。我们的边惩罚 $p(e)$与α-分数的主要区别在于，我们的惩罚依赖于回路。此外，与 α-分数不同，我们的惩罚将标准的回路长度与最小生成树的差距分解为各巡回边对应的非负差距之和。

4 实验

4.1 RTDL 条形码与最优巡回

RS

一条边的 RTDL 条形码值与旅行商问题解有何关联？为研究这种依赖关系，我们生成了 100 个随机二维旅行商问题，并通过以下方式获得：1）运行具有不同迭代次数的 2-opt 算法，得到非最优巡回路线；2）运行 Concorde[1]，得到最优巡回路线。随后，对于每条非最优巡回路线，我们按照边的 RTDL 条形码升序排列，并估计一条边属于最优巡回路线的经验概率。图 2 显示，具有

[1]https://www.math.uwaterloo.ca/tsp/concorde/

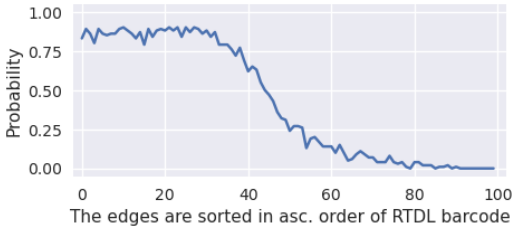


图 2：属于最优巡回路线的概率与边的 RTDL 条形码的关系。

方法	TSP-100		TSP-200		TSP-300	
	长度	时间（秒）	长度	时间（秒）	长度	时间（秒）
LK	8.022	17.49	11.176	111.06	13.672	350.95
2-opt	8.244	1.01	11.465	12.25	13.940	47.51
2-opt+RTDL	8.226	2.84	11.356	18.08	13.798	62.30
2-opt+RTDL+opt(D)	**8.170**	2.77	**11.303**	16.81	**13.654**	57.03
3-opt	8.137	12.30	11.284	160.19	13.747	701.86
3-opt+RTDL	8.082	13.49	11.157	131.50	13.594	545.50
3-opt+RTDL+opt(D)	**8.011**	11.43	**11.126**	104.61	**13.546**	456.67

表1：二维欧几里得旅行商问题。RTDL 条形码对 2-opt 和 3-opt 的改进。

较低条形码值的边以较高概率属于最优巡回路线。更多结果请参见附录 A。

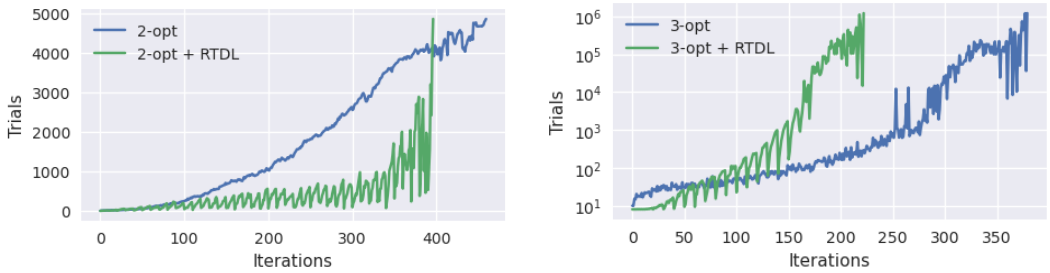
4.2 对 2-opt、3-opt 的改进

我们使用 RTDL-条形码来改进 2-opt 和 3-opt 算法。在这一简化设置中，我们可以单独测试 RTDL 导出的边权重的影响，而不受复杂训练流程等因素的干扰。巡回路线中某条边对应的 RTDL-条形码值定义了该边的“劣度”。对于最优巡回路线的候选方案，一个自然的想法是优先尝试从巡回路线中移除这类边。在原始 2-opt（附录中的算法1）中，算法尝试移除边的选取顺序并未明确规定，通常按顺序选取。我们修改了 2-opt 算法，按照 RTDL 条形码值从高到低选取边（附录中的算法 2）。修改后的算法记为 2-opt+RTDL。

欧几里得旅行商问题。表1展示了二维欧几里得旅行商问题的实验结果（城市均匀采样自一个单位正方形）。我们报告平均巡回路线长度和平均优化时间。所有结果均取自100次试验的平均值。在所有设置下，按RTDL排序都能得到更短的巡回路线，但耗时略长。对于3-opt算法，按RTDL排序既能缩短巡回路线长度，也能减少优化时间。

此外，我们采用了Helsgaun（2000）提出的一种启发式算法。他们指出，给定与顶点相关联的标量 π_i 后，可以修改距离矩阵$d_{i,\ j} \leftarrow d_{i,\ j} + \pi_i + \pi_j$。这种修改会使不同巡回路线的总长度产生相同的常数偏移，因此最优巡回不会改变。通过次梯度优化过程优化 π 值，使1-树中的节点度数达到2，也就是使其更接近巡回路线。显然，经过这种修改后，标准2-opt和3-opt的步骤不会改变，但2-opt/3-opt+RTDL的步骤会改变。表1中将这种启发式算法记为*opt(D)*，结果表明它能进一步提升效果。

我们得出结论，按照 RTDL 条形码排序总能改善目标值，且不会显著增加执行时间。图 3a 显示，对于 2-opt，按照 RTDL 排序可减少每次迭代的试验次数——即巡回路线得到改善前检查的边对数量。对于 3-opt，按照 RTDL 排序可以加快收敛，见图 3b。消融研究见附录 F。



(a) 2-opt，2-opt+RTDL。

(b) 3-opt，3-opt+RTDL。

图3：每次迭代的平均试验次数。

非度量旅行商问题。 我们还针对非度量旅行商问题（最小权重哈密顿环问题）开展了实验，其中边的权重从 (0，1) 中采样。同样，加入 RTDL 总能带来改进，见附录中的表 5。

TSPLib 实例。 我们还对 TSPLib 中的任务进行了实验。对于每项任务，使用不同的初始巡回路线运行优化 100 次，见表3。对于某些初始巡回路线，普通 2-opt 算法难以达到最优解，偶尔还会超过其迭代上限；而在相同的初始条件下，2-opt+RTDL 从未遇到此类问题。更多详情请参见附录D。

4.3 用于热力图引导 2-opt 的 RTDL

对于大规模旅行商问题，有一类独立的求解器并不直接预测最优巡回路线，而是生成热力图。热力图中的每个条目 $\Phi_{i,j}$ $n \times n$都反映了边 (i, j) 与最优解的相关性。为了根据热力图获得可行解，会使用专门的解码策略（Sun & Yang (2023)）。在这项工作中，我们遵循 Sun & Yang (2023)，并使用 2-opt 和 2-opt + RTDL 算法进一步优化通过对生成的热图进行贪心解码而得到的初始巡回路线。通过这项实验，我们旨在验证：当初始巡回路线已经很好地近似最优解时，基于 RTDL 的边遍历能否进一步改进 2-opt。

遵循 Xia 等（2024）的工作，我们使用 DIFUSCO（Sun & Yang (2023)）、ATT-GCN Fu 等（2021）、DIMES（Qiu 等，2022）、UTSP（Min 等，2023）和 SoftDist（Xia 等（2024））方法学习得到的热力图。我们使用所有可用的问题进行测试：128 个规模为 500 的问题、128 个规模为 1000 的问题，以及 16 个规模为 10000 的问题。对于每个模型和问题规模，我们报告优化后巡回路线的平均长度以及巡回路线优化的平均耗时。如表2所示，在大多数情况下，所提出的 2-opt + RTDL 能得到平均巡回路线长度更短、计算效率更高的解。基于热图的方法略逊于 LKH-3 和 Concorde，这反映了这类方法目前的性能。图4进一步证实，2opt + RTDL 能使所有模型更快收敛。为进一步考察，我们还在表6中给出了通过热力图贪心解码得到的初始巡回路线的平均长度。为便于可视化分析，我们在图5中展示了所有模型在同一个 TSP-500 问题上的初始巡回路线，以及经 2-opt 和 2-opt + RTDL 优化后的巡回路线。我们发现，对于 TSP-500，RTDL 每 5 次迭代更新一次；对于 TSP-1000 和 TSP-10000，则每 100 次迭代更新一次。这些设置对所有模型普遍适用。

敏感性分析见图8。我们可以注意到，对于所有粒度取值，采用 Freq=1 （即每次迭代都更新）的 2-opt+RTDL 通常会导致更高的计算时间。在大多数情况下，采用 Freq=5时，只要粒度设置得当，就能以小得多的计算时间匹配甚至改善 Freq=1 的平均巡回路线长度。总体而言，当 Freq¿1 时，更新越频繁，平均巡回路线长度越短。通常，粒度越大，计算速度越慢，不过频率越高，这种影响越不明显。在某个阈值频率以下，粒度越大，平均巡回路线长度越长。超过该阈值频率后，粒度越大，平均巡回路线长度越短。该阈值频率取决于模型和问题规模。

表2：将 2-opt 和 2-opt + RTDL 算法应用于通过热图贪心解码获得的初始巡回路线时的平均长度和运行时间。

方法	TSP-500		TSP-1000		TSP-10000	
	长度	时间（秒）	长度	时间（秒）	长度	时间（秒）
DIFUSCO						
2-opt	16.98	0.38	24.01	5.78	75.87	7027.63
2-opt +RTDL	**16.88**	0.23	**23.60**	1.76	**74.28**	865.5
ATT-GCN						
2-opt	17.63	1.09	24.72	12.54	77.36	17567.5
2-opt +RTDL	**17.3**	0.55	**24.2**	3.05	**75.54**	1567.69
DIMES						
2-opt	17.74	1.66	24.84	16.05	77.68	25363.25
2-opt +RTDL	**17.35**	0.73	**24.31**	3.56	**75.3**	1797
UTSP						
2-opt	17.54	0.93	24.57	10.20	–	–
2-opt +RTDL	**17.25**	0.56	**24.12**	2.95	–	–
SoftDist						
2-opt	17.41	0.75	24.31	6.43	75.41	5264.88
2-opt +RTDL	**17.24**	0.49	**24.09**	2.65	**74.65**	1549.5

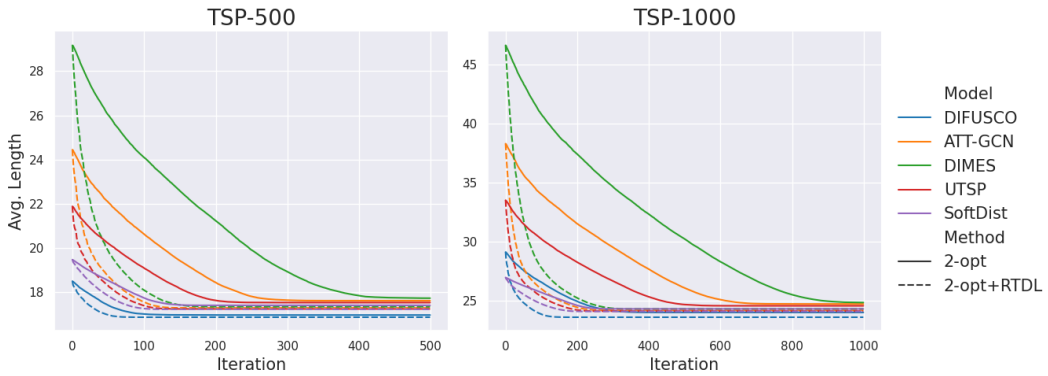


图4：热力图引导的2-opt 和 2opt+ RTDL 在巡回路线优化过程中平均长度的变化。

4.4 深度Q学习的 RTD

为了进一步评估将拓扑洞见融入旅行商问题（TSP）的潜力，我们基于一类流行的强化学习（RL）方法——深度 Q 网络（DQN），设计了一系列实验。我们的目标是评估一种基线学习策略，以及一种经过修改、能够通过奖励调节利用额外拓扑信息的方法。

在这些实验中，我们随机生成城市数量各不相同的旅行商问题任务，然后要求模型针对每个实例，分别多次迭代构建巡回路线。每个模型有 600 个回合从头开始构建巡回路线，并记录其最佳结果。

在 DQN 框架中，核心思想是学习一个 Q 函数 $Q(s,\ a)$，用于近似执行动作 a 于状态 s下、随后继续遵循当前策略所能获得的预期累积奖励。在旅行商问题中，“状态”对应于部分构建的巡回路线，而“动作”则表示选择下一个要访问的城市。通过反复交互来更新 Q 函数，智能体会逐渐学会偏好更有可能带来更短巡回路线的动作，从而在各回合中改进其决策过程。算法详情见附录 E。

在我们的设置中，奖励是稠密的：智能体在每一步都会收到一个负的边权重，其对应于所选边的成本。这确保了累计奖励在一个

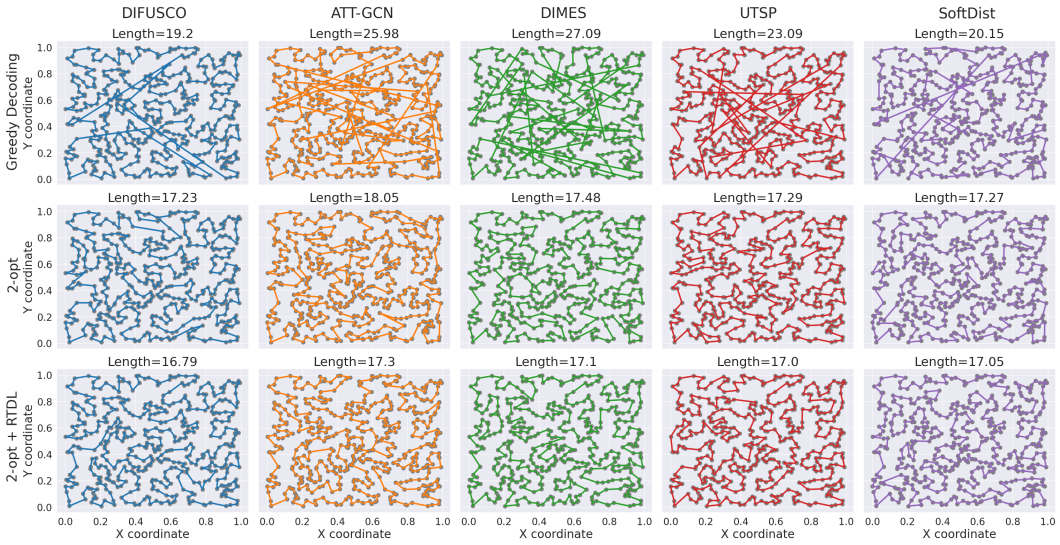


图5：通过不同模型和优化方法求解同一 TSP-500 问题的示例。对于每种模型，我们提供通过热力图贪心解码得到的初始巡回路线，以及经 2-opt 和 2-opt + RTDL 优化后的巡回路线。

表3：2-opt、2-opt+RTDL 和 2-opt+RTDL-Full 在 TSPlib 任务上的实验结果，并与 Concorde 进行比较。所有算法均在相同的时间和迭代上限下运行。我们还报告了相对于 Concorde 所得最佳解的差距（%）。算法详情见附录 C。

	2-opt			2-opt+RTDL-Full			2-opt+RTDL		
Name	路径长度	时间	差距（%）	路径长度	时间	差距（%）	路径长度	时间	差距（%）
ulysses16	74.54 ± 0.48	0.02秒	0.6	74.51 ± 0.41	0.06秒	0.54	**74.49 ± 0.46**	0.06秒	**0.52**
ulysses22	76.28 ± 0.43	0.05秒	0.8	**76.18 ± 0.37**	0.11秒	**0.64**	76.25 ± 0.51	0.13秒	0.78
att48	34842.07 ± 648.93	0.20秒	3.9	**34632.78 ± 608.78**	0.79秒	**3.3**	34848.44 ± 609.44	1s	4.0
eil51	451.66 ± 31.26	0.23秒	4.9	450.84 ± 7.54	0.44秒	4.9	**446.68 ± 7.82**	0.58秒	**3.9**
berlin52	8150.23 ± 240.14	0.27秒	8.0	8133.56 ± 217.95	0.53秒	7.0	**7957.48 ± 237.26**	0.69秒	**5.5**
st70	714.90 ± 46.18	0.63秒	5.5	**705.53 ± 10.45**	1.02秒	**4.1**	709.45 ± 11.09	0.62秒	4.7
eil76	604.50 ± 37.62	1.12秒	11.0	**583.92 ± 11.66**	1.01秒	**7.2**	586.20 ± 11.36	1.83秒	7.6
pr76	113227.56 ± 12746.41	0.81秒	4.7	113392.47 ± 2581.35	1.1s	4.8	**112657.0 ± 2221.49**	1.56秒	**4.2**
gr96	541.12 ± 12.91	1.56秒	5.6	543.32 ± 12.71	1.66秒	6.1	**538.54 ± 12.01**	2.6s	**5.1**
kroC100	22009.93 ± 553.第36卷	1.95秒	6.1	22020年.69 ± 528.93	1.7s	6.1	**21815.72 ± 598.33**	2.79秒	**5.1**

回合奖励恰好等于巡回路线长度的相反数。为纳入拓扑信息，我们计算 RTDL-条形码(T_{tour}, G)，其中 T_{tour} 是构造出的哈密顿回路， G 是完整的城市图。根据定理1，每条巡回边 $e \in E(T_{\text{tour}}) \setminus \{e_{\max}\}$ 都通过双射 ψ与一条最小生成树边配对，而 $e_{\max}$ 对应的配对则单独定义。这样，我们便可以为每条边分配拓扑奖励

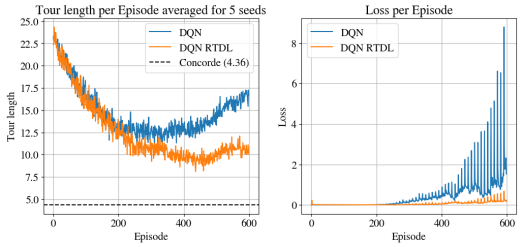
$$r_{\text{RTDL}}(e) = w(e) - w(\psi(e)) \quad \geq 0.$$

选中该边时，会将此值逐步加到环境奖励中。随后，这些惩罚会作为额外的逐步奖励进行分配，从而有效引导智能体生成具有拓扑意义的巡回路线，同时保留标准的基于长度的奖励。

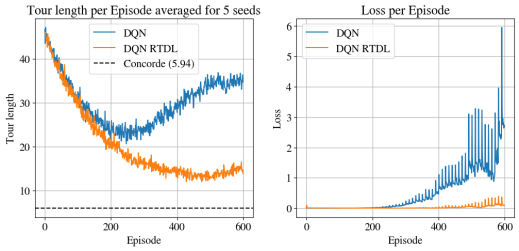
为确保公平，所有模型均在相同的城市配置下进行训练，且起点和终点固定，并在等效条件下接受评估。此外，我们使用 Concorde TSP 求解器计算全局最优解，并以此为基准测试学习得到的解。图 6 所总结的实验表明，基于 RTDL 的奖励塑形显著提升了 DQN 智能体训练的稳定性和效果。基线 DQN 很快便陷入平台期，并在约 300 个回合后甚至出现发散；而 DQN+RTDL 保持稳定，并持续缩短巡回路线长度。总体而言，与基线相比，DQN+RTDL 始终生成更短的巡回路线，且在不同随机种子下的方差更低。表 4 中的数值结果进一步证实了这些发现，且适用于不同的旅行商问题规模。

表 4：比较基线 DQN 与 RTDL 增强变体在不同 TSP 规模下的平均巡回路线长度。

方法	TSP-50	TSP-70	TSP-100
DQN（基线）	8.76	10.31	18.66
DQN +RTDL	**7.65**	**8.46**	**11.75**



（a）50 个城市



（b）100 个城市

图 6：基线 DQN 和 DQN+RTDL 在包含 50 个和 100 个城市的旅行商问题实例上的学习曲线。每条曲线均为五次随机初始化的平均值。

5 结论

在这项工作中，我们提出了旅行商问题的一种全新拓扑学视角，将经典巡回路径-MST总差距分解为逐边散度贡献。我们成功利用拓扑信号构建了拓扑引导的2-opt和3-opt，使随机欧几里得实例和 TSPLib 基准测试中的巡回路线长度得以缩短。此外，拓扑引导的2-opt还改善了基于热图的旅行商问题求解器生成的巡回路线的精细优化，适用于多达 10000 个节点的情形。基于这一发现，我们提出了一种拓扑感知强化学习框架，其中每条最小生成树边都会根据其散度贡献获得相应的局部奖励。实验证明，这种细粒度的奖励分配能够加速收敛，带来更稳定的学习曲线，并最终较传统的整体式奖励方案提升巡回路径质量。此外，我们还建立了 LKH-3 算法中的 α-score 与 RTDL 条形码某一特定变体之间的等价关系，从而强化了我们的拓扑方法与成熟启发式技术之间的联系。通过将持久同调与现代强化学习相结合，我们的工作为在组合优化的深度强化学习中融入全局图结构开辟了新途径。

参考文献

Daniel Aloise、Robin Moine、Celso C Ribeiro 和 Jonathan Jalbert。首次改进还是最佳改进？一项深入的局部搜索计算研究，以阐明一项支配性主张。《欧洲运筹学杂志》，2025年。

Serguei Barannikov、Ilya Trofimov、Nikita Balabin 和 Evgeny Burnaev。《表示拓扑散度：一种比较神经网络表示的方法》。国际机器学习会议，*ICML 2022*。

Qi Chen、Ceyue Liu、Jiahao Sun、Hang Yuan 和 Jian Tang。《一种适用于旅行商问题的通用深度强化学习求解器》。载于第 *12* 届国际学习表征会议（*ICLR*）论文集，维也纳，奥地利，2024年。

George A. Croes。求解旅行商问题的方法。《运筹学》，1958年。

Pedro F. da Costa、Lu´ıs A. A. Rodrigues 和 Lu´ıs M. S. Russo。通过深度强化学习学习旅行商问题的 2-Opt启发式。载于第*34*届*AAAI*人工智能大会论文集，美国纽约州纽约，2020年。

Xinyu Fan、Shunhua Jiang、Cong Xie 和 Xiaoming Sun。重新思考用于大规模旅行商问题的事后搜索型神经方法。载于第*41*届国际机器学习会议论文集，《机器学习研究论文集》，奥地利维也纳，2024年。PMLR。网址：https://proceedings.mlr.press/v238/fan24a.html。

Zhang-Hua Fu、Kai-Bin Qiu 和 Hongyuan Zha。将小型预训练模型泛化到任意大规模的旅行商问题实例。载于*AAAI*人工智能大会论文集，第35卷，第7474–7482页，2021年。

Pierre Hansen 和 Nenad Mladenovi´c。首次改进与最佳改进：一项实证研究。《离散应用数学》，154卷第5期:802–817，2006年。

Michael Held 和 Richard Karp。旅行商问题与最小1-树。《运筹学》，1970年。

Keld Helsgaun。Lin–Kernighan旅行商启发式的有效实现。《欧洲运筹学杂志》，126(1):106–130，2000年。

Roy Jonker 和 Ton Volgenant。将非对称旅行商问题转化为对称旅行商问题。《运筹学快报》，2:161–163，1983年。URL https://api. semanticscholar.org/CorpusID:123004903。

Wouter Kool、Herke van Hoof 和 Max Welling。注意力机制，学会解决路径规划问题！载于国际学习表征会议，2019年。URL https://github.com/ wouterkool/attention-learn-to-route。

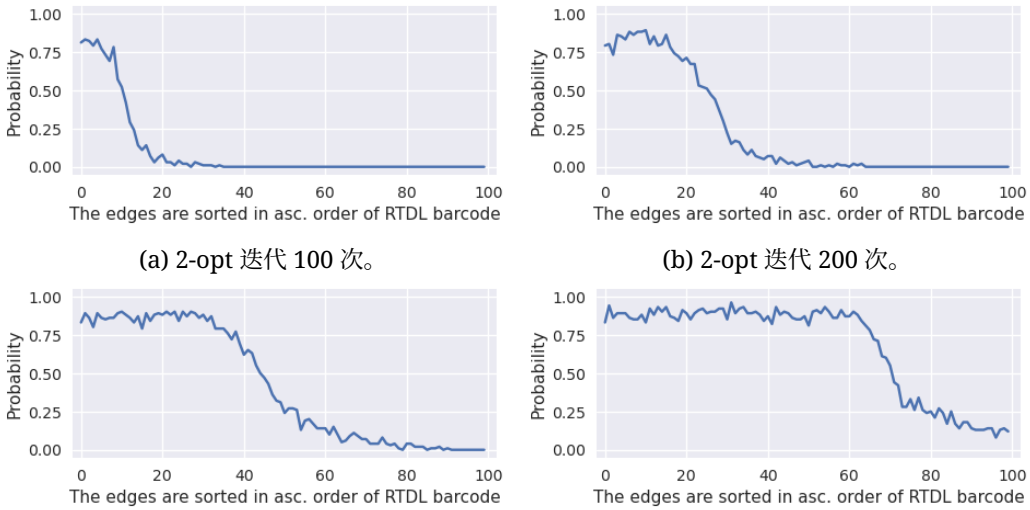
Shen Lin 和 Brian W. Kernighan。《旅行商问题的有效启发式算法》。《运筹学》，1973年。Jiale Ma、Wenzheng Pan、Yang Li 和 Junchi Yan。《ML4CO-Bench-101：图上经典组合问题的基准机器学习》。载于神经信息处理系统大会第*39*届年会数据集与基准测试专题，2025年。网址：https://openreview.net/forum?id=ye4ntB1Kzi。Yimeng Min、Yiwei Bai 和 Carla P Gomes。《用于求解旅行商问题的无监督学习》。《神经信息处理系统进展》，第36卷：47264–47278，2023年。Ruizhong Qiu、Zhiqing Sun 和 Yiming Yang。《DIMES：用于组合优化问题的可微元求解器》。《神经信息处理系统进展》，第35卷：25531–25546，2022年。Sebastian Sanokowski、Sepp Hochreiter 和 Sebastian Lehner。《用于无监督神经组合优化的扩散模型框架（DiffUCO）》。载于第*41*届机器学习国际会议（*ICML*）论文集，《机器学习研究论文集》，奥地利维也纳，2024年。PMLR。网址：https://proceedings.mlr.press/v238/sanokowski24a.html。Zhiqing Sun 和 Yiming Yang。《DIFUSCO：用于组合优化的基于图的扩散求解器》。《神经信息处理系统进展》，第36卷：3706–3731，2023年。InstaDeep研究团队。《训练高性能组合优化强化学习种群》。载于*NeurIPS*，2023年。

Eduard Tulchinskii、Daria Voronkova、Ilya Trofimov、Evgeny Burnaev 和 Serguei Barannikov。《RTD-Lite：用于比较学习任务中加权图的可扩展拓扑分析》。载于第*28*届人工智能与统计国际会议（*AISTATS*）论文集，《机器学习研究论文集》第258卷。PMLR，2025年。网址：https://proceedings.mlr.press/v258/tulchinskii25a.html。

Oriol Vinyals、Meire Fortunato 和 Navdeep Jaitly。《指针网络》。载于 神经信息处理系统进展，2015年。

Yifan Xia、Xianliang Yang、ZichuanLiu、Zhihao Liu、Lei Song 和 Jiang Bian。《立场：重新思考用于解决大规模旅行商问题的事后搜索型神经方法》。*arXiv* 预印本 *arXiv:2406.03503*，2024年。

Charles T. Zahn。《用于检测和描述格式塔聚类的图论方法》。*IEEETransactionsonComputers*，1971年。



(a) 2-opt 迭代 100 次。

(b) 2-opt 迭代 200 次。

(c) 2-opt 迭代 300 次。

(d) 2-opt，已收敛。

图7：属于最优巡回路线的概率与边的 RTDL 条形码之间的关系。

A 补充结果

图7展示了一条边属于最优巡回的概率；这些边取自具有不同迭代次数的 2-opt 优化。表5展示了非度量 TSP 问题中 2-opt 和 3-opt 的结果。

B 定理1证明的详细内容

记 γ^e 为连接边 $e \in E(T_{mst})$ 两个端点 A 和 B 的路径，该路径由权重小于 $w(e)$ 的边 $T_{mst}-$ 以及权重大于或等于 $w(e)$ 且小于或等于 $w(\tilde{e})$ 的边 $T-$ 组成，其中最小权重为 $w(\tilde{e})$。记 $\tilde{\gamma}^{\tilde{e}}$ 为连接边 $\tilde{e} \in T$ 的端点 $\tilde{A}$ 和 $\tilde{B}$ 的路径，该路径由权重小于或等于 $w(e')$ 的边 $T_{mst}-$ 以及权重小于 $w(\tilde{e})$ 的边 $T-$ 组成，其中最小权重为 $w(e')$。如果这样的路径 $w(e')$ 满足 $< w(e)$，那么并集 $\tilde{\gamma}^{\tilde{e}} \cup \gamma^e \setminus \tilde{e}$ 将是一条连接顶点 A、B 的路径，其中包含权重小于 $w(e)$ 的边 $T_{mst}-$，以及权重大于或等于 $w(e)$ 且严格小于 $w(\tilde{e})$ 的边 $T-$，这与 $\tilde{e}$ 的定义矛盾。因此，路径 $e \cup \gamma^e \setminus \tilde{e}$ 满足路径 $\tilde{\gamma}^{\tilde{e}}$ 的条件，且 $w(e')$ 为 $= w(e)$，并且 e' 与 e 重合。因此 $\psi(\phi(e)) = e$，且映射 ϕ、ψ 是双射。

RTDL 条形码与 alpha-score 的关系

设 G 为一个权重为 $w_{i,\ j}$的加权图。考虑一个不包含任意顶点“1”及其关联边的子图 $G_1 \subset G$。

定义。1-树是 G_1 的一个生成树，包含顶点 1，以及连接顶点 1 和 G_1的两条边。

α-score 用于旅行商问题和车辆路径问题中的最先进 LKH 算法 Helsgaun（2000）的边选择。 α-score 是衡量边“劣势”的指标。 α-score 较低的边很可能属于最优巡回路线。对于一条边（$i,\ j$）：

$$\alpha(i,j) = L(R^+(i,j)) - L(R),$$

其中， R 是 G的最小1-树，而 $R^+(i,\ j)$是 G 在限制包含边（$i,\ j$）时的最小1-树， $L(\cdot)$是一个长度函数。

命题。 对于 $i,\ j \neq 1,\ \alpha(i,\ j)$等于RTDL-条形码($G^{i,\ 1^{j}},\ G_1$)中一条对应于边（$i,\ j$）的条的长度。 G_1和 $G^{i,\ 1^{j}}$ 都是图，包含 G 的所有顶点，但不含顶点1。 G_1 定义如上，而 $G^{i,\ 1^{j}}$ 只有一条边（$i,\ j$），其权重为$w_{i,\ j}$.

证明。计算 RTDL-Barcode(A, B) 的算法如下（Tulchinskii 等，2025）。进行 RTDL 评估的前提是图 A、B 的顶点之间存在双射。构造辅助图 C，其中包含 A、B 的边的并集，以及边 $w_e^C = min(w_e^A, w_e^B)$ 上的权重。当缺少某条边时，其权重视为 $+\infty$。随后，执行用于求图 C 的最小生成树的克鲁斯卡尔算法。设 e^C 是图 C 的最小生成树中的一条边。边 e^C 连接两个连通分量 C_1、C_2。在 RTDL 计算算法中，按权重递增顺序添加 A 的最小生成树中的边。添加某条边 e^A 后，会出现一条连接 C_1 和 C_2（或它们各自任一顶点）的路径。因此， e^C 与 e^A 配对。它们共同构成 RTDL-条形码中的一条（e^C, e^A）。注意，始终有 $w_e^C \leq w_e^A$。如果 $e^A \neq e^C$，那么将 e^A 添加到 $C_1 \cup C_2$ 后会出现一个环。如（Helsgaun，2000）所示，1-树 $R^+(i,j)$（若 $i,j > 1$）可由 $R \cup e^A$ 删除上述环中最长的边得到。显然，该边就是 e^C。因此， $\alpha(i,j) = w_e^A - w_e^C$。

□

在表3和附录D中，我们还评估了 2-opt+RTDL 的两个变体：

- **2-opt+RTDL-Full**- 如算法 2 所述，但在每个优化阶段应用于所有边。
- **2-opt+RTDL** - 如算法 2 所述，每个优化周期按大小为 10 的批次选择边。

算法1： 2-opt 算法

初始化巡回路线;

重复
    开始;
    改进的 ← false;
    对巡回路线中的每个 e_1 中巡回路线 **执行**
        对巡回路线中的每个 e_2 中巡回路线 **执行**
            巡回路线 new ←巡回路线;
            从新的巡回路线中移除 e_1,← e_2 并重新连接它;
            如果*Len(*新巡回路线*) < Len(*巡回路线*)* **则**
                巡回路线 ←新巡回路线;
                改进 ← 真;
                跳转至 START:
            end
        end
    end
直到 改进;
返回 巡回路线

D 2-opt 和 RTDL 变体的鲁棒性分析

In t本节中，我们使用 2-opt、2-opt+RTDL 和 2-opt+RTDL-Fu 进行了大量实验， ll to e以评估它们从不同初始巡回路线收敛的能力。

我们发现，普通 2-opt 有时无法从某些初始巡回路线收敛到全局最优解，而 2-opt+RTDL 始终能避免此类情况。图 9 展示了 76 城市 Christofides/Eilon 实例的收敛时间分布及最终巡回路线长度，表 7 则汇总了更广泛问题集中极端情况的出现频率。

具体而言，我们报告了两类失败情况：(i) 巡回路线长度超过全局最优解 10% 以上的运行（“差距 $> 10\%$”），以及 (ii) 超过 20 秒时间预算的运行（“时间

算法 2:2-opt+RTDL 算法

初始化巡回路线;
$N \leftarrow$ min(10, |巡回路线|)

重复
  开始;
  已改进 ←false;
  E ←巡回路线[0中的边：N] 按 RTDL 条形码降序排序;
  对每个 e_1 属于 E 的**执行**
    对每个 e_2 属于 E 的**执行**
      新巡回路线 ←巡回路线;
      从新巡回路线中移除 e_1, e_2 并重新连接;
      如果*Len(*巡回路线 *new) < Len(*巡回路线*)* **那么**
        巡回路线 ←巡回路线_new;
        改进 ←真;
        转到 START:
      end
    end
  end
  如果 $N < |tour|$**那么**
    $N \leftarrow$ min($N + 10$, |巡回路线|);
    改进 ←true;
  end
直到 改进;
返回巡回路线

方法	TSP-100		TSP-200		TSP-300	
	长度	时间（秒）	长度	时间（秒）	长度	时间（秒）
2-opt	1.871	0.21	2.340	1.60	2.732	4.02
2-opt+RTDL	1.631	2.31	1.866	9.65	2.061	32.90
2-opt+RTDL+opt(D)	**1.581**	4.16	**1.832**	18.53	**2.037**	50.65
3-opt	1.161	28.05	1.234	346.36	1.304	1482.76
3-opt+RTDL	1.156	29.00	1.210	329.49	1.282	1602.79
3-opt+RTDL+opt(D)	**1.144**	22.03	**1.203**	310.40	**1.278**	1327.23

表5：非度量旅行商问题。RTDL 条形码对 2-opt 和 3-opt 的改进。

表6：通过热图贪心解码获得的初始巡回路线的平均长度。

模型	TSP-500	TSP-1000	TSP-10000
DIFUSCO	18.51	29.12	113.74
ATT-GCN	24.45	38.29	205.84
DIMES	29.17	46.61	331.73
UTSP	21.89	33.51	–
SoftDist	19.48	26.97	84.20

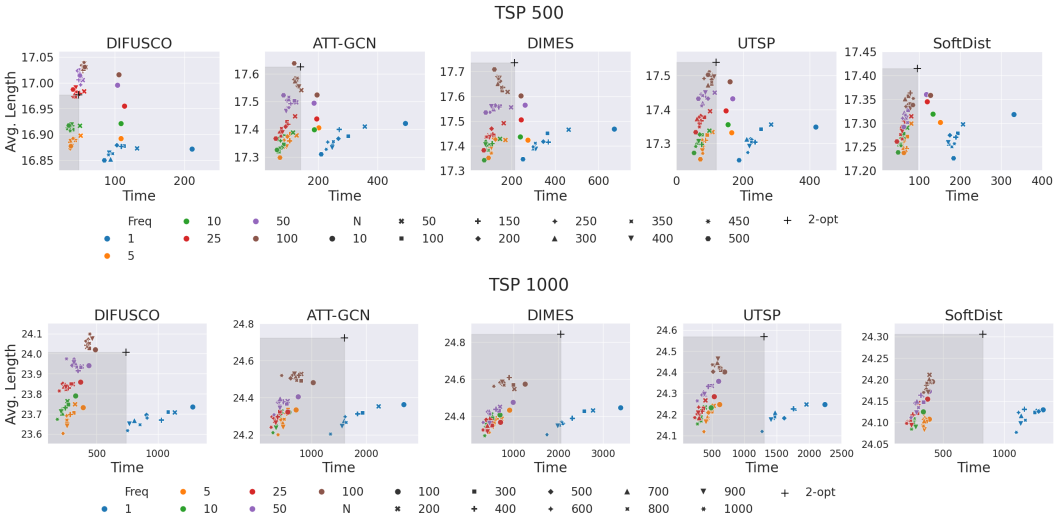


图8：热力图引导的2-opt + RTDL 的敏感性分析。*Freq* 表示 RTDL 更新频率，*N* 表示要优化的顶点数量。灰色区域表示优于2-opt解的结果，即平均长度更短、计算速度更快。

表 7：不同 TSP 求解器在 TSPLib 实例上的失败案例比较。对于每个实例，我们报告构造的巡回路线长度比全局最优解高出 10% 以上的运行所占百分比（“Len. $>$ 10% opt.”），以及在分配的 20 秒时间预算内未能收敛的运行所占百分比（“时间限制”）。

	2-opt		2-opt+RTDL-Full		2-opt+RTDL	
Name	差距 $>$ 10% ↓	时间限制 (%) ↓	差距 $>$ 10% ↓	时间限制 (%) ↓	差距 $>$ 10% ↓	时间限制 (%) ↓
ulysses16	0.0	0.0	0.0	0.0	0.0	0.0
ulysses22	0.0	0.0	0.0	0.0	0.0	0.0
att48	0.6	0.0	0.7	0.0	**0.4**	0.0
eil51	0.7	0.0	0.9	0.0	**0.5**	0.0
berlin52	27.1	0.0	23.5	0.0	**20.7**	0.0
st70	5.6	4.0	5.9	**0.0**	**4.0**	0.0
eil76	7.7	2.7	9.7	**0.2**	**2.3**	1.5
pr76	1.6	0.0	2.7	0.0	**1.2**	0.0
gr96	4.9	0.0	6.4	0.0	**3.0**	0.0
kroC100	9.2	0.0	11.8	0.0	**7.8**	0.0
rd100	12.25	0.0	19.0	0.0	**9.0**	0.0
kroD100	6.75	0.0	6.75	0.0	**3.75**	0.0
eil101	45.25	39.0	4.75	**1.5**	**7.25**	7.75
lin105	12.25	0.5	19.0	**0.25**	**12.0**	0.5
ch130	9.0	0.0	7.5	0.0	**2.0**	0.0
ch150	11.2	0.0	19.2	0.0	**5.0**	0.0

限制”）。普通的**2-opt** 基线出现此类失败的比例最高，而**2-opt+RTDL-Full**则显著减少了此类失败。**2-opt+RTDL**变体的整体表现最为稳健，在不同的 TSPLib 实例中，失败率始终较低。

E DQN 算法

本节介绍应用于旅行商问题的标准深度Q学习（DQN）算法3。

F 消融研究

在消融研究中，我们比较了按 RTDL 条形码对边进行排序与按距离降序排序的方法，见表 8。

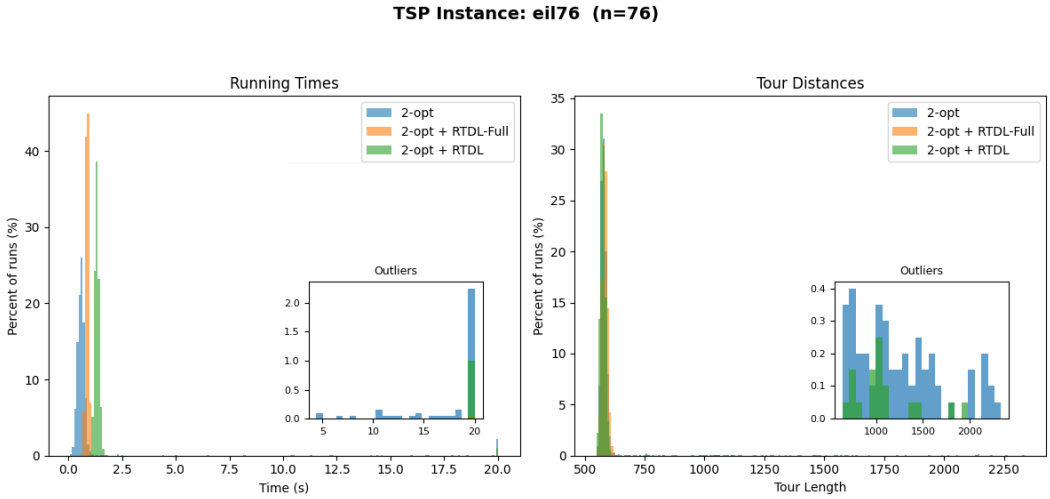


图 9：不同版本的 2-opt 的计算时间分布及其产生的巡回路径长度。

算法3： 用于旅行商问题的深度Q学习（DQN）

初始化经验回放缓冲区 D;
使用随机权重 θ 初始化 Q网络;
使用权重 $\theta_{\text{target}} \leftarrow \theta$ 初始化目标网络;
使用固定的城市坐标初始化环境 env;
初始化最优巡回路线和最优路径长度 $\leftarrow \infty$;
for 回合 $= 1$ **到** M **do**
  $state \leftarrow$env.reset();
  for $t = 1$ **到** T **do**
    以概率 ε 选择随机动作 a，否则 $a \leftarrow \arg\max_a Q(state, a; \theta)$;
    $state'$，奖励，完成 $\leftarrow$env.step(a);
    将 ($state$，a，奖励 $state'$，完成) 存入 D;
    如果 $|D| \geq$ 批次大小**则**
      从 D 中随机抽取一个状态转移小批量;
      $y_j \leftarrow$ 奖励$_j + \gamma \cdot (1 -$ 完成$j) \cdot$ 最大值$_{a'}\, Q(state'_j,\ a'; \theta_{\text{target}})$;
      $L(\theta) \leftarrow \frac{1}{B} \sum_j \left(Q(state_j,\ a_j;\ \theta) - y_j\right)^2$;
      更新 $\theta \leftarrow \theta - \alpha \cdot \nabla_\theta L(\theta)$;
    end
    如果 t 模 $C = 0$**则**
      $\theta_{\text{target}} \leftarrow \theta$;
    end
    如果 完成 **则**
      如果巡回路线长度 $<$ 最优路径长度 **则**
        更新最优巡回路线和最优路径长度;
      end
      跳出;
    end
    $state \leftarrow state'$;
  end
end
返回最优巡回路线

方法	TSP-100		TSP-200		TSP-300	
	长度	时间（秒）	长度	时间（秒）	长度	时间（秒）
2-opt+距离。	8.474	0.72	11.780	6.32	14.291	26.64
2-opt+α	8.471	0.76	11.75	6.70	14.31	29.65
2-opt+RTDL	**8.226**	2.84	**11.356**	18.08	**13.798**	62.30

表 8：二维欧几里得旅行商问题。RTDL 条形码的消融研究。

问题集	最优长度	2-opt		2-opt +RTDL	
		长度	Time	长度	Time
ATSP 均匀-50	1.55 ± 0.18	2.08 ± 0.26	0.01	1.93 ± 0.24	0.14
ATSP 均匀-100	1.57 ± 0.13	2.11 ± 0.2	0.06	1.98 ± 0.18	0.63
ATSP 均匀-200	1.56 ± 0.11	2.16 ± 0.15	0.35	1.90 ± 0.14	4.01
ATSP 均匀-500	1.57 ± 0.06	2.22 ± 0.10	3.7	2.05 ± 0.08	38.4
ATSP HCP-50	0.0	4.27 ± 1.84	0.01	2.89 ± 1.54	0.1
ATSP HCP-100	0.0	4.01 ± 1.85	0.05	2.44 ± 1.48	0.43
ATSP HCP-200	0.0	3.37 ± 1.48	0.27	1.69 ± 1.09	2.7
ATSP HCP-500	0.0	3.4 ± 1.18	2.43	1.57 ± 1.06	9.44

表 9：2-opt 和 2-opt +RTDL 算法应用于非对称旅行商问题、哈密顿回路问题时的平均长度和运行时间。

G 局限性

尽管我们的实验主要集中于合成实例和基准测试 TSPLIB 实例，但现实世界的路径规划通常涉及其他约束（时间窗、容量、动态更新），而这些约束与拓扑奖励方案之间的相互作用仍有待探索。

H 关于重复权重的说明

如果加权图中的某些权重重复，则最小生成树可能不唯一。只有当相应的边在 $G^{w<w(e)}$ 中连接相同的连通分量时，才会发生这种情况。此时，（s,t）巡回与最小生成树之间的双射也不唯一。然而，边惩罚 $p(e) = w(e) - w(\psi(e))$ 仍然是唯一定义的，因为 $w(\psi(e))$ 不会改变。

I 应用于其他组合优化问题

我们针对其他组合优化问题提供了额外评估——ATSP（非对称旅行商问题）和 HCP（哈密顿回路问题，可视为 ATSP 问题的一种特例）。我们利用了这样一个事实：通过将顶点数量加倍并转换成本矩阵，可以将非对称旅行商问题归约为对称旅行商问题（参见 Jonker 与 Volgenant（1983年））。为进行公平比较，我们将 2-opt 和 2-opt + RTDL 应用于归约后的 ATSP。我们采用了基准测试中的 ATSP 和 HCP 问题（Ma 等人，2025年）。表 9 报告了平均巡回路线长度和平均时间。所得结果证实，2-opt + RTDL 能够提供平均长度更小的巡回路线。我们提供这些结果，以证明 2-opt + RTDL 可以应用于其他组合优化问题。与 2-opt 相比，当前实现会产生计算开销，其改进工作留待后续研究。

J 大语言模型（LLMs）的使用

我们使用大语言模型润色了“引言”和“摘要”部分的文字。