

引用本文: 赵赞善, 高冠军, 甘维明, 等. 机器学习辅助的多目标优化海底光缆路由规划研究[J]. 光通信技术, 2025, 49(3): 10-15.

机器学习辅助的多目标优化海底光缆路由规划研究

赵赞善^{1,2,3}, 高冠军^{1*}, 甘维明^{2,3}, 王皓宇¹, 段茂生¹, 康达⁴

(1. 北京邮电大学 电子工程学院, 北京 100876; 2. 中国科学院声学研究所 南海研究站, 海口 570105;

3. 陵水海洋信息海南省野外科学观测研究站, 海南 陵水 572423; 4. 哈尔滨工程大学 水声工程学院, 哈尔滨 150000)

摘要: 为提升海底光缆路由规划算法的全局优化能力, 降低累积成本与风险, 同时提升算法效率, 提出一种机器学习辅助 (MLA) 的多目标优化海底光缆路由规划算法。MLA 利用强化学习优势, 自主迭代学习, 同步优化成本与风险, 考虑海底地形、水深等参数, 同时采用 Pareto 前沿面作为收敛评价标准, 并与传统蚁群优化 (ACO) 算法进行对比验证。实验结果表明: 在相同风险水平下, 该算法能降低 27.45% 的铺设成本, 其最优解累积风险仅为 ACO 算法的 25%, 且收敛速度提升了 330 倍以上。此外, 其大部分 Pareto 解位于前沿面, 明显优于 ACO 算法解集的离散分布。

关键词: 海底光缆路由规划; 多目标优化; 强化学习; Q 学习; 海缆系统经济性优化; 海缆系统生存性优化

中图分类号: TN91 **文献标志码:** A **文章编号:** 1002-5561(2025)03-0010-06

DOI: 10.13921/j.cnki.issn1002-5561.2025.03.002

开放科学(资源服务)标识码(OSID):



Research on multi-objective optimization of submarine optical fiber route planning assisted by machine learning

ZHAO Zanshan^{1,2,3}, GAO Guanjin^{1*}, GAN Weiming^{2,3}, WANG Haoyu¹, DUAN Maosheng¹, KANG Da⁴

(1. School of Electronic Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China;

2. Hainan Acoustics Laboratory, Institute of Acoustics, Chinese Academy of Sciences, Haikou 570105, China;

3. Lingshui Ocean Information Hainan Province Field Scientific Observation and Research Station, Lingshui Hainan 572423, China;

4. College of Underwater Acoustic Engineering, Harbin Engineering University, Harbin 150000, China)

Abstract: To improve the global optimization capability of submarine cable routing planning algorithm, reduce cumulative costs and risks, and improve algorithm efficiency, a machine learning assisted (MLA) multi-objective optimization algorithm for submarine cable routing planning is proposed. Utilizing the advantages of reinforcement learning, MLA autonomously iterates learning, synchronously optimizes costs and risks, considers parameters such as seabed topography and water depth, and adopts Pareto frontier as the convergence evaluation criterion. It is compared and verified with traditional ant colony optimization (ACO) algorithm. The experimental results show that under the same risk level, the algorithm can reduce the laying cost by 27.45%, and its optimal solution cumulative risk is only 25% of the ACO algorithm, and the convergence speed is improved by more than 330 times. In addition, most of its Pareto solutions are located at the forefront, which is significantly better than the discrete distribution of the ACO algorithm solution set.

Key words: submarine optical fiber route planning, multi-objective optimization, reinforcement learning, Q-learning, economic optimization of submarine cable system, survivability optimization of submarine cable system

收稿日期: 2024-09-23。

基金项目: 国家重点研发计划项目(2022YFB2903303)资助; 国家自然科学基金项目(62371064)资助。

作者简介: 赵赞善(1986—), 男, 硕士, 副研究员, 主要研究方向为光通信及网络技术。主持完成海南省自然科学基金项目 1 项, 参与海南省重大科技项目 1 项, 荣获海南省电子学会自然科学技术一等奖 1 项。

* 通信作者: 高冠军(1984—), 男, 博士, 教授, 主要研究方向为光信息通信网络技术。



0 引言

海底光缆是国际通信的重要基础设施, 承载着 99% 以上的国际通信服务^[1], 在全球经济、文化和社会交流中发挥着重要作用。然而, 海底光缆的部署环境恶劣, 维护成本高昂, 且易受人类活动、自然灾害等风险因素影响。因此, 降低海底光缆系统的成本、提升其在复杂风险环境下的生存能力具有重要意义, 而海底光缆路由规划是实现成本与风险优化的有效方法之一。

在传统工程实践中,海底光缆路由规划主要依赖专家经验,需要人工逐段考量多种因素进行设计,这一过程耗时费力。为提升规划效率,近年来研究者开始探索利用优化算法实现海底光缆路由的自动化设计,以优化其成本和风险。例如,文献[2-5]通过加权和法将成本和风险这2个优化目标转化为单个优化目标,利用传统的优化算法(如Dijkstra算法和快速行进法)搜索最优路径。在加权和法中,成本和风险的重要性权重在整个海底光缆路由规划过程中保持固定,这可以视为一种固定策略。然而,固定策略无法根据多因素在局部区域内做出准确决策。例如,在高风险区域,由于风险的权重值较低,可能导致系统低估实际风险而选择该区域作为必经路由。传统多目标优化启发式算法,如蚁群优化(ACO)算法^[6]已应用于海底光缆路由规划,能同时优化多个目标函数。这类算法通过随机搜索增强局部探索能力,具有一定的灵活性。但其存在易陷入次优解和收敛速度慢的固有缺陷。

为此,本文提出一种机器学习辅助(MLA)的多目标优化海底光缆路由规划算法(下文简称MLA算法),以最小化其累积成本和风险。

1 工作原理

1.1 MLA 框架

在MLA中,为智能体(agent)构建学习环境是强化学习中的一个关键步骤。本文以目标海域D为研究对象,通过整合地理信息系统(GIS)数据及与海缆成本和风险相关的数据,基于文献[6]的数值模型建立了

目标海域D内的成本(单位:1 000 CNY/km)和风险(无量纲)二维分布场。该海域采用数字栅格(DS)建模,其地理特征和工程参数可表示为 $M \times N$ 的二维矩阵,其中 M 、 N 分别对应矩阵的行维度和列维度。

MLA的多目标优化海底光缆路由规划示意图如图1所示,该规划方案由数据集、目标海域因素分布图和机器学习三部分组成。首先,系统会采集目标海域内与海底光缆规划相关的数据集,并基于文献[6]的数值模型生成成本分布图和风险分布图,二者分别反映agent从当前位置移动至相邻位置产生的局部成本和局部风险。在数据处理方面,本文采用DS格式表示地理数据,定义Q-learning算法的agent当前位置为节点0,其可移动的8个相邻动作为 $a_1 \sim a_8$ 。为量化多目标优化决策,引入 Q_c 和 Q_r 分别表示成本和风险对agent决策的影响值。以agent从节点0移动至候选点7、8为例,其成本决策 Q 值可分别表示为 $Q_c(s_0, a_7)$ 和 $Q_c(s_0, a_8)$,其计算式(1)所示;风险决策 Q 值 Q_r 采用相同的计算方法。通过这种方式,系统可对目标海域内任意节点的路由选择进行多目标优化评估,为海底光缆规划提供科学依据。

$$\begin{cases} Q_c(s_0, a_7) = \frac{C_0 + C_7}{2} \\ Q_c(s_0, a_8) = \frac{\sqrt{2}(C_0 + C_8)}{2} \end{cases} \quad (1)$$

其中, $C_0 \sim C_8$ 表示相邻位置的成本, $Q_c(s_0, a_k)$ 表示agent根据成本考虑选择部署海底光缆的概率。

在机器学习部分,本文采用Q-learning算法进行路由优化。该算法以预先计算的目标海域各节点的成

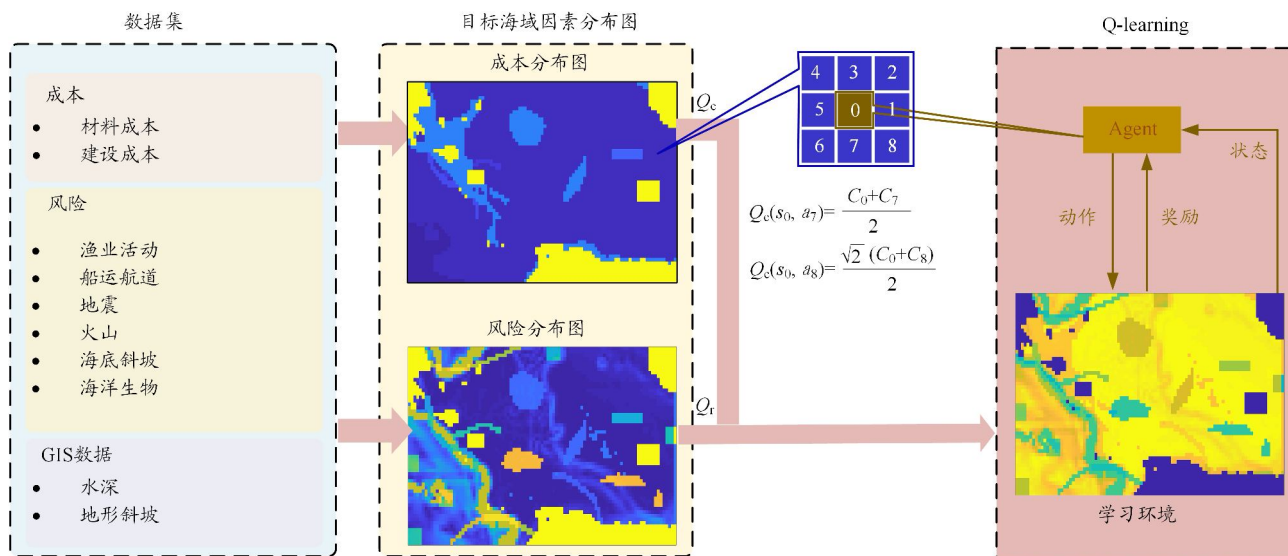


图1 MLA的多目标优化海底光缆路由规划的示意图

赵赞善,高冠军,甘维明,等: 机器学习辅助的多目标优化海底光缆路由规划研究

本决策值 Q_c 和风险决策值 Q_r 作为先验数据, 初始化 agent 的状态-动作空间(即 Q 表)。agent 基于当前状态的动作值, 运用贪心策略(ϵ -greedy)从其 8 个相邻候选位置中选择最优节点作为光缆路由的下一跳。当 agent 执行选定动作后, 学习环境会相应更新系统状态, 并根据动作质量给予奖励反馈, 同时将最新环境状态返回给 agent。通过这种持续的环境交互机制, agent 能够从起始点逐步探索直至目标节点, 每一条成功连接的路径即构成一条可行光缆路由方案。这种基于强化学习的优化方法具有两大优势: 首先, 通过持续的环境交互和策略优化, agent 能够有效提升在多目标约束下(成本与风险)寻找最优路径的能力; 其次, 该方法的自适应学习特性显著提高了路径搜索的时间效率, 为大规模海域的光缆路由规划提供了可行的解决方案。

1.2 先验数据初始化

为实现多因素决策, 需将各类优化目标的分布特征编码至 agent 的动作空间(Q 表)。设 agent 当前处于网格 (i, j) , 其对应的成本为 C_0 , 8 个相邻位置的成本为 C_{1-8} 。在动作空间中, agent 的当前状态为 s_0 , agent 的下一个候选状态用 s_{1-8} 表示。

本文以成本因素的量化为例进行说明。基于成本计算的当前位置状态-动作值可由 $Q_c(s_0, a_k) = \Gamma(C_0 + C_k)/2$ 计算得出, $k=1, 2, \dots, 8$ 。 Γ 表示与距离相关的因子, C_k 表示节点 k 的成本值, 如果下一个候选状态位于由候选状态组成正方形的对角线位置, 则设为 $\sqrt{2}$, 否则设为 1。本文以式(1)中 $Q_c(s_0, a_7)$ 和 $Q_c(s_0, a_8)$ 为例进行说明。若网格分辨率较低, 需将 DS 坐标系转换为墨卡托投影坐标, 并使用实际距离比替代理论值 $\sqrt{2}$ 。通过全域网格节点的遍历计算, 即可构建基于成本优化的完整状态-动作空间。

同样地, 基于风险因素的状态-动作空间可通过相同方法构建。将成本和风险 2 个优化目标整合, 定义

综合状态-动作价值函数为 $Q(s, a) = \sum_{f \in c, r} [1 - \text{norm}(Q_f)]$,

其中 $\text{norm}(\cdot)$ 表示归一化, f 表示考虑因素的类别。假如有更多的考虑因素, 则可以通过求和公式扩展更多优化目标。agent 从 $Q(s, a)$ 中学习最优策略, 最终输出满足多目标优化要求的缆线路径。

1.3 状态更新与奖励规则

Q-learning 算法特别适用于离散状态空间和动作空间的优化问题。鉴于本文采用栅格化方法对环境进

行建模, 且 agent 的可选动作(8 个相邻节点)构成有限离散集, 因此选用 Q-learning 算法进行海底光缆路径规划具有理论合理性和工程适用性。Q-learning 算法的状态更新公式^[7]如下:

$$Q(s_t, a_t) = Q(s_t, a_t) + \alpha \{ r'_t + \gamma \max_a [Q(s_{t+1}, a)] - Q(s_t, a_t) \} \quad (2)$$

其中, $Q(s_t, a_t)$ 表示在状态 s_t 采取动作 a_t 的价值函数, t, α, r'_t 和 γ 分别表示时刻、学习率、 t 时刻的奖励和奖励折扣率。

本文使用 r' 表示 agent 的奖励, 其设置规则可以表示为

$$r' = \begin{cases} 5, & \text{agent 达到终点} \\ -1, & \text{agent 超出边界} \\ 0, & \text{其它} \end{cases} \quad (3)$$

1.4 算法及复杂度分析

本文使用 L, Ω 和 R 分别表示第 L 次循环、最大循环次数、Pareto 解集、Pareto 解集对应的路径集合。在迭代过程中, agent 以概率 $1-\epsilon$ 选择 Q 值最大的动作, 同时以等概率 ϵ 从 8 个候选节点中随机选取一个作为下一个节点。MLA 算法如算法 1 所示。

算法 1 MLA 算法

输入: 成本、风险分布地图, 海底光缆系统的起点和终点

输出: 输出 Pareto 解集及对应的路径

- 1) 使用成本、风险分布地图, 初始化 $Q_c(s, a)$ 和 $Q_r(s, a)$
- 2) 利用 $Q_c(s, a)$ 和 $Q_r(s, a)$ 计算 $Q(s, a)$
- 3) $R \leftarrow []$, $\Omega \leftarrow []$
- 4) While $L \leq L$
- 5) for each episode
- 6) agent 利用 $Q(s, a)$ 策略选下一步动作
- 7) 采取动作 a , 观察奖励 r' 、新状态 S^*
- 8) 利用式(1)更新 $Q(s, a)$
- 9) if agent 超出地图范围
- 10) $r' = -1$, Break
- 11) else
- 12) 直到 S^* 为终点, 获取可行路径 R^*
- 13) end if
- 14) end for
- 15) if R^* 是最新 Pareto 解
- 16) 更新 R 和 Ω and $r' = 5$

17) end if

18) $k \leftarrow k+1$

19) end while

20) 返回 Ω 和 R

在仿真中,每个 episode 的循环最大次数设置为地图规模的 10%。状态空间和动作空间的大小分别为 $N \times M$ 和 8。MLA 算法的时间复杂度为 $O[0.1 \times 8 \times (N \times M)^2]$,即简化为 $O[(N \times M)^2]$ 。尽管仿真过程中存在额外的条件限制,但这些附加条件所需的复杂度通常与地图规模呈线性关系。因此,MLA 算法的空间复杂度为 $O(N \times M)$ 。

2 仿真结果与讨论

为验证算法的有效性和鲁棒性,本文将其与 ACO 算法进行对比实验,采用文献[6]的目标区域、起点和终点设置进行仿真。目标区域范围为西南(18.49°N , -72.89°W)至东北(18.78°N , -72.29°W),并在其中设置高密度风险因素。区域高程图及风险分布参见文献[6]中图 3,基于该高程图可计算生成海底斜坡分布图。为保持对比实验的一致性,本文采用与 ACO 算法相同的海底光缆路由起止点坐标:起点(18.7326°N , -72.8082°W),终点(18.5436°N , -72.4813°W)。实验数据采用国际公共海域的虚拟路由及世界大洋海底地形图(GEBCO)公开地理数据,相关数据可通过 <https://www.gebco.net> 获取,所有数据仅用于算法性能验证研究。

本文首先采用 Q-learning 算法进行海底光缆路径规划,以最小化累积成本和风险。动作决策采用 ϵ -greedy 策略,其中 $\epsilon=0.1$, $\alpha=0.5$, $\gamma=0.99$ 。算法最大迭代次数设为 50 000 次。奖励函数设置如下:成功到达终点奖励+5,移出目标区域惩罚-1,其它情况奖励 0。

为验证 MLA 算法的优化性能,本文将 Q-learning 与 ACO 算法的 Pareto 解集结果进行对比分析,结果如图 2 所示。可以看出,Q-learning 算法的最大累积成本

低于 ACO 算法;以 Q-learning 算法的最大累积成本为例,Q-learning 算法规划的累积风险仅为 ACO 算法的最小成本对应累积风险的 25%;根据计算可得,当累积风险为 40 时,Q-learning 算法可降低 27.45%的成本。因此,MLA 算法在 Pareto 解集的优化性能上显著优于 ACO 算法,能够更高效地平衡成本与风险的双目标优化问题。此外,ACO 算法解集呈现明显的前沿偏离现象,易陷入局部次优解;而 Q-learning 算法展现出近乎完全的 Pareto 前沿收敛性,体现了更强的全局搜索能力。这是因为 ACO 算法受初始解影响较大,其启发式搜索容易导致局部收敛,而 MLA 算法通过强化学习的策略探索机制,实现了更有效的全局搜索,从而避免了局部停滞问题。上述结果验证了 MLA 算法在全局探索与局部优化 2 个维度均显著优于 ACO 算法,其 agent 交互式学习机制在多目标路由规划中展现出更强的鲁棒性和适应性。

为直观展示 MLA 算法在海底光缆路径规划中的性能优势,将图 2 中的典型可行解(路径 1~3)叠加在成本分布热图上进行可视化分析,结果如图 3 所示。可以看出,所有路径均有效规避最高成本区域(D_1 区黄色高亮部分)。此外,路径 3 虽然穿过 D_1 区域,其穿越这部分区域的长度明显长于路径 1 和路径 2,这导致了路径 3 的累积成本高于路径 1 和路径 2。这一观察结果与图 2 中的分析结果相符。

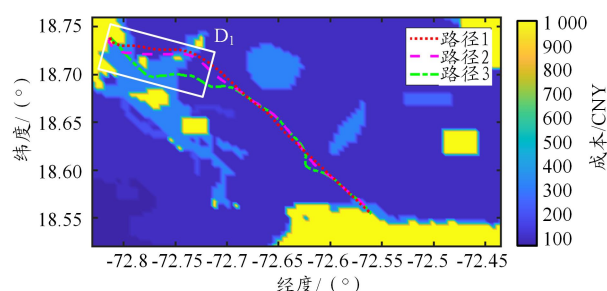


图3 海底光缆路径 1~3 在成本分布图上的效果图

路径 1~3 在风险分布图上的效果如图 4 所示。可以看出,3 条路径均成功规避最高风险区域。与路径 1 和路径 2 相比,路径 3 完全避开了 D_1/D_2 高风险区,因此其累积风险最低。路径 1~3 的风险差异主要源于 D_1/D_2 区域的通过策略差异,这与图 2 中的结果一致。图 2 和图 3 采用的成本/风险指标均为基于 GEBCO 等公开数据构建的仿真模型计算结果,实际工程应用需结合实测数据校准。

本文通过 MLA 算法获得了 Pareto 前沿解集,在当前仿真条件下已达到多目标优化的理论极限。路径 1~

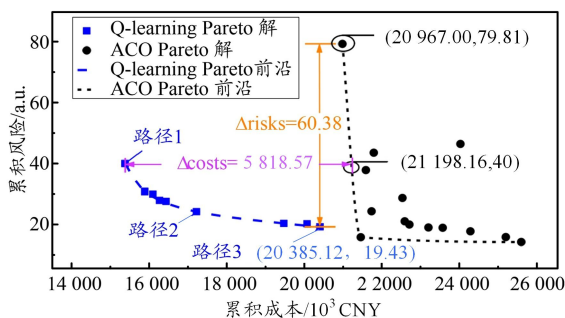


图2 Q-learning 和 ACO 算法的 Pareto 解

赵赞善,高冠军,甘维明,等: 机器学习辅助的多目标优化海底光缆路由规划研究

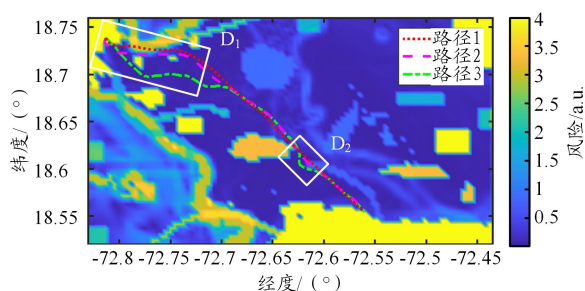


图4 海底光缆路径1~3在风险分布图上的效果图

3 在成本或风险方面均展现出了各自的优势。然而,在实际工程应用中,仅需要选择一条海底光缆路径作为最终的最优方案。因此,本文采用文献[6]中的综合评价方法,通过加权综合评价从 Pareto 前沿解集中遴选出最优工程实施方案。3 条路径的综合评分结果如表 1 所示。综合评分最高的路径更适合实际工程应用。因此,路径 3 被选定为最终路由方案。

表 1 位于 Pareto 前沿的可行解及其综合评分结果

路径名称	成本/ 10^3 CNY	风险/a.u.	综合评分/a.u.
路径 1	15 379.59	40.00	1.85
路径 2	17 203.94	24.23	2.18
路径 3	20 385.12	19.43	2.30

本文采用动态更新的 Pareto 前沿作为算法收敛性分析的评估基准。当 agent 完成一次从起点到终点的搜索后,将获得一条具有特定累积成本和累积风险的海底光缆路径。该路径将与当前 Pareto 解集中的所有路径进行非支配比较:若不存在任何路径能在累积成本和风险 2 个目标上同时优于当前路径(当前路径不被任何现有解支配),则该路径将被纳入 Pareto 解集作为最新 Pareto 解。通过持续记录这些非支配解的演化过程,可以直观地展现算法的收敛特性。

MLA 算法最新 Pareto 解的累积成本和累积风险随着迭代次数的变化曲线如图 5 所示。可以看出,与

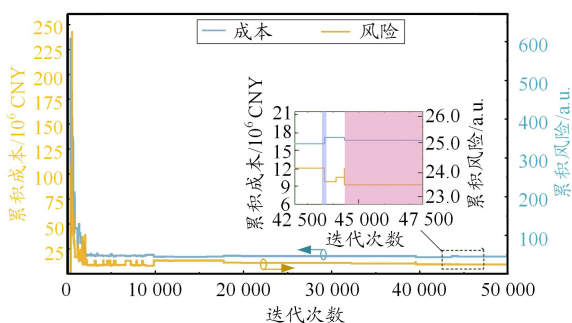


图5 成本和风险区域的收敛曲线

传统优化算法类似,MLA 算法的 Pareto 解累积成本和累积风险随迭代次数增加逐渐下降。但由于多目标优化的特性,两者无法收敛至单一值。累积成本与累积风险变化趋势相反:在放大图中的蓝色区域,累积成本上升时累积风险下降。这是因为海底光缆的成本与风险目标存在内在冲突,如避开高风险海域虽能降低风险,但会导致路径延长或施工难度增加,从而提高成本。因此,算法收敛曲线会在一定范围内波动以寻求平衡。在放大图中的红色区域,随着迭代的进行,累积成本和风险曲线基本不变,表明算法已达到收敛,难以发现新 Pareto 解。证明了 MLA 算法能有效处理冲突多目标优化并实现收敛。

为了评估 MLA 算法的时间效率,本文在相同硬件环境下对 MLA 算法和 ACO 算法进行了对比测试。通过 5 次独立重复实验取平均值以消除随机误差,结果如表 2 所示。可以看出,MLA 算法平均耗时仅 40.17 s (标准差 $\sigma=1.23$ s),而 ACO 算法需要 13 296.69 s ($\sigma=387.5$ s),效率提升 330 倍。这一显著优势源于 MLA 独特的强化学习机制:其通过 Q-learning 算法的价值导向搜索和 ϵ -greedy 策略的智能探索-利用平衡,有效避免了传统启发式算法(如 ACO 算法)在解空间中的盲目随机搜索,从而实现了更快的收敛速度。实验数据充分验证了 MLA 算法在多目标路径规划问题中的计算效率优势。

表 2 2 种算法计算时间的比较

实验序号	算法	计算时间/s	平均耗时/s
1	ACO	12 905.83	13 296.69
2		12 354.61	
3		12 334.08	
4		12 897.68	
5		15 991.24	
1	MLA	37.96	40.17
2		38.78	
3		40.54	
4		44.04	
5		39.52	

3 结束语

本文提出了一种 MLA 的多目标优化海底光缆路由规划算法,该算法通过构建与成本和风险目标相关的学习环境,应用 agent 学习最优策略,以最小化累积

成本和累积风险。通过与传统的 ACO 算法对比实验表明:在相同风险条件下,MLA 算法可节省 27.45%的成本;其最优 Pareto 解的累积风险仅为 ACO 算法最小成本解对应风险的 25%;同时,MLA 算法解的 Pareto 前沿偏离更小,表明其全局和局部优化能力显著优于 ACO 算法。在时间效率方面,MLA 算法比 ACO 算法提升 330 倍,且收敛速度更快。此外,MLA 通过结合随机搜索与策略驱动,避免了固定权重带来的局限性。综上所述,MLA 算法在海底光缆路由规划中能显著优化性能,提高效率与收敛速度,于成本-风险多目标优化中表现更强的鲁棒性与工程适用性。

参考文献:

- [1] WANG X, WANG Z, ZUKERMAN M. Application of adaptive parallel fast marching method in automatic submarine cable path planning[J]. IEEE Transactions on Automation Science and Engineering, 2025, 22: 9498–9514.
- [2] ZHAO M, CHOW T W S, TANG P, et al. Route selection for cabling considering cost minimization and earthquake survivability via a semisupervised probabilistic model[J]. IEEE Transactions on Industrial Informatics, 2016, 13(2): 502–511.
- [3] WANG Z, WANG Q, ZUKERMAN M, et al. Multiobjective path optimization for critical infrastructure links with consideration to seismic resilience[J]. Computer-Aided Civil and Infrastructure Engineering, 2017, 32(10): 836–855.
- [4] WANG Z, WANG Q, MORAN B, et al. Terrain constrained path planning for long-haul cables[J]. Optics Express, 2019, 27(6): 8221–8235.
- [5] WANG X, WANG Z, WANG T, et al. Designing cost-effective and reliable submarine communications cable path: lessons from the tonga volcano disaster[J]. IEEE Communications Magazine, 2023, 61(7): 179–185.
- [6] ZHAO Z, WANG J, GAO G, et al. Multi-objective optimization for submarine cable route planning based on the ant colony optimization algorithm[J]. Photonics, 2023, 10(8): 896–1–896–13.
- [7] WATKINS C J C H, DAYAN P. Q-learning[J]. Machine Learning, 1992, 8: 279–292.