

# 基于图注意力的烟草零售店监管路线优化研究

钱漫<sup>1</sup>, 陈杜勇<sup>1</sup>, 钟培泉<sup>1</sup>, 叶子健<sup>1</sup>, 姜哲<sup>1</sup>,  
刘晓鹏<sup>2</sup>, 胡树波<sup>1</sup>, 钟展兴<sup>1</sup>, 李岱峰<sup>3</sup>, 董佳<sup>1</sup>

- 广东烟草东莞市有限公司, 广东 东莞 523000;
- 广东省烟草专卖局(公司), 广东 广州 510000;
- 中山大学信息管理学院, 广东 广州 510006)

**摘要:** 由于城市烟草零售店较为密集, 传统路径规划算法求解最优监管路径将耗费大量的运算时间, 在规定时间内无法保证运算效果。并且, 现有方法较少考虑求解问题的网络特性及候选子集的可解释性。鉴于此, 提出一种基于图注意力的节点选择及路径优化算法(GA-SGPO), 迭代选择最优坐标节点子集, 在子集上进行求解以减少计算时间。此外, 通过计算节点间的结构相似性, 降低子集样本稀疏性。实验数据包括东莞市4万家零售店的地理坐标。实验结果显示, 所提出的GA-SGPO模型在保证求解精度的同时, 求解时间平均提升48%。GA-SGPO算法可显著节省计算时间, 更贴近实际应用场景。而注意力机制和节点相似度计算, 可为最优节点选择提供可视化依据。

**关键词:** 图注意力; 最短路径优化; 烟草监管; 节点结构相似度; 节点选择

DOI: 10.11907/rjtk.232007

开放科学(资源服务)标识码(OSID):

中图分类号: TP391.41

文献标识码: A

文章编号: 1672-7800(2024)009-0157-06



## Research on Tobacco Retail Store Regulatory Route Optimization Algorithm Based on Graph Attention Mechanism for Node Selection

QIAN Man<sup>1</sup>, CHEN Duyong<sup>1</sup>, ZHONG Peiquan<sup>1</sup>, YE Zijian<sup>1</sup>, JIANG Zhe<sup>1</sup>,  
LIU Xiaopeng<sup>2</sup>, HU Shubo<sup>1</sup>, ZHONG Zhanxing<sup>1</sup>, LI Daifeng<sup>3</sup>, DONG Jia<sup>1</sup>

- Guangdong Tobacco Dongguan Co., LTD., Dongguan 523000, China;
- Guangdong Tobacco Monopoly Bureau(Company), Guangzhou 510000, China;
- School of Information Management, Sun Yat-sen University, Guangzhou 510006, China)

**Abstract:** Since the tobacco retail stores in cities are dense, traditional path planning algorithms for solving the optimal supervision path will consume a lot of time, and cannot guarantee the effect within the specified time. In addition, existing methods seldom consider the network characteristics and the explainability of the candidate subset. This study proposes a graph attention-based node selection and path optimization algorithm (GA-SGPO), which iteratively selects the optimal coordinate node subset and performs calculation on the subset to reduce computation time. In addition, the structural similarity between nodes is calculated to reduce the sparsity of training samples. The experimental data includes the coordinates of 40,000 retail stores in Dongguan City. The experimental results show that the GA-SGPO model ensures the solution accuracy while the solution time is reduced by an average of 48%. The GA-SGPO can significantly save computational time and is closer to practical application scenarios. The attention mechanism and node similarity calculation can provide visualization basis for optimal node selection.

**Key Words:** graph attention; shortest path optimization; tobacco regulation; node structure similarity; node selection

收稿日期: 2023-09-25

基金项目: 广东省烟草专卖局(公司)科技项目专项(粤烟科项202206)

**作者简介:** 钱漫(1978-), 女, 硕士, 广东烟草东莞市有限公司系统分析师, 研究方向为数据挖掘; 陈杜勇(1984-), 男, 硕士, 广东烟草东莞市有限公司烟草专卖管理师(二级), 研究方向为卷烟市场智能监管; 刘晓鹏(1985-), 硕士, 广东省烟草专卖局(公司)政工师, 研究方向为烟草专卖市场监管、电子烟监管; 李岱峰(1982-), 男, 博士, 中山大学信息管理学院教授、博士生导师, 研究方向为情报学、知识工程; 董佳(1986-), 女, 广东烟草东莞市有限公司烟草专卖管理师(二级)、企业培训师(二级), 研究方向为烟草专卖案件查办和市场监管等。本文通讯作者: 董佳。

# 0 引言

在城市人口稠密区域,存在大量烟草专卖零售店,一方面能够为市民购买香烟带来便利,另一方面也给烟草质量监管带来挑战。监管人员需要定期设定检查路线,对路线上零售商铺所出售的香烟进行品规、来源、质量及价格方面的检查,以控制出售香烟的质量,保障消费者权益。每次规划都需要耗费大量的时间和人力成本,路径优化求解器是降低监管成本、提升监管效率的有效解决方案。然而,传统的优化求解器在从区域几百家,甚至上千家零售商铺中获取最优路径,在高性能计算卡支持下,也需要10多个小时的计算时间。因此,亟需一种优化的方案以解决上述问题。

路径规划任务的最优解一般只包含网络中的少数节点和边,如果在最优路径求解前将那些不太可能出现在最优解集合中的节点或边删掉,那么删除掉的点及边将不会影响计算准确度。此外,删除无关点越多,网络规模也会越小,计算速度将会显著提升。传统方法如图1(a)所示,直接在原始网络图中,依据节点坐标进行最优闭环路径求解,存在算法求解空间过大的问题。本文模型思想如图1(b)所示,在对整个网络节点求解最优路径前,先使用融合图注意力机制的神经网络判断每个节点及边是否是最优路径上,挑选种子节点及边集合并扩充形成子图,随后利用求解器从子图求解最优闭环路径。

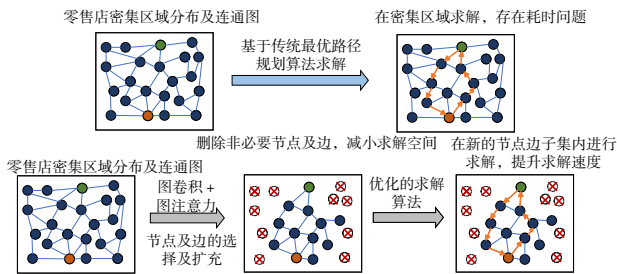


Fig. 1 Comparison of solution processes based on node selection and traditional mode

图1 基于节点选择和传统模式下的求解流程比较

## 1 相关工作

### 1.1 路径规划求解算法

路径规划任务为在一个带坐标位置及通路的节点集合中,计算出任意两点的最短路径。在此基础上,衍生出各种带约束的规划任务,如考虑时间窗、资源约束的最短路径问题等。动态规划及其带松弛问题的约束规划是路径规划求解的重要算法<sup>[1-3]</sup>。学者们在此基础上对算法作持续改进,包括设置下界、通过有限差异搜索提升动态规划算法<sup>[4-6]</sup>。本文利用求解器将动态规划、LP松弛、割平

面法、分支界定<sup>[7-8]</sup>、启发式<sup>[9-10]</sup>等算法进行整合,形成优化求解方案。主要求解思想是在假设原问题的基础上,找到规模更小的求解问题,并通过非基变量判断是否能够进一步优化,不断迭代得到最优解。

### 1.2 机器学习在组合优化问题中的应用

机器学习在组合优化领域的应用较为广泛<sup>[11-12]</sup>。Larsen等<sup>[13]</sup>利用机器学习求解信息不完全情况下的最优解; Kruber等<sup>[14]</sup>利用机器学习判断是否可使用Dantzig-Wolfe分解求解当前整数规划问题; Morabit等<sup>[15]</sup>使用机器学习方法对可能出现在最优路径上的边进行判断,从而减小求解规模。该模型在构建标签识别特征的基础上,采用逻辑斯蒂回归、随机森林和神经网络作出保留边判断。近年来,常运用深度学习理论方法,如卷积网络<sup>[16]</sup>、指针网络<sup>[17]</sup>、深度强化学习<sup>[18-19]</sup>等对旅行商等组合优化问题进行求解,或生成初始可行解。

### 1.3 图注意力机制

图注意力机制是对图卷积神经网络<sup>[20]</sup>和注意力机制<sup>[21]</sup>理论的整合。本文通过图卷积实现网络图中节点的向量编码,该编码包含了节点的网络结构特征,可应用于不同的下游任务。而注意力机制可将图卷积的输出节点向量进行序列输入,在对每个节点标签进行学习的同时,通过注意力机制判断对该节点产生重要影响的其他网络节点。在利用网络结构特性提升模型效果的同时,对标签结果提供可解释支持。而已有研究更侧重于通过人工特征进行构建,如采用节点边的中心度、一阶二阶入出度、随机游走或采样方式获取节点网络结构特性,存在结构信息丢失、在标签训练过程中没有网络关联以及缺少可解释性的问题。

## 2 模型构建

模型整体架构如图2所示。给定初始巡查路径图 $G = (V, A)$ ,其中 $V = \{v_0, v_1, \dots, v_N\}$ 为图 $G$ 中节点的集合(共有 $N$ 个节点), $A = \{a_0, a_1, \dots, a_M\}$ 为节点连边的集合(共有 $M$ 条边),如 $a_0 = (v_1, v_2)$ 表示存在从节点 $v_1$ 到 $v_2$ 的边路径。监管人员从 $v_0$ 点出发,需要在 $G$ 中选择一个最优监管路径。

当 $N$ 较大时,传统方法会产生大量计算时间。现有研究主要考虑通过对原问题的特征构建、深度编码、矩阵分解以寻找初始可行解的方式,并提升求解效果,而较少考虑目标数据集的网络特性,以及通过网络特性进行特征自动编码和最优子集选择。因此,本文引入图注意力机制从 $G$ 中进行全部包含、部分包含子图筛选。最优节点集合 $V^*$ 表示每个节点及边都在最短路径上;全部包含子图即完全包含 $V^*$ ,用 $G' = (V', A')$ 表示;部分包含子图表示子集中的部分节点及边在 $V^*$ 中,用 $G'' = (V'', A'')$ 表示。由于图注意力只能找到较少的部分包含子图,因此采用基于网络结构的相似度计算模型对子图的节点进行扩充,然后采用求解

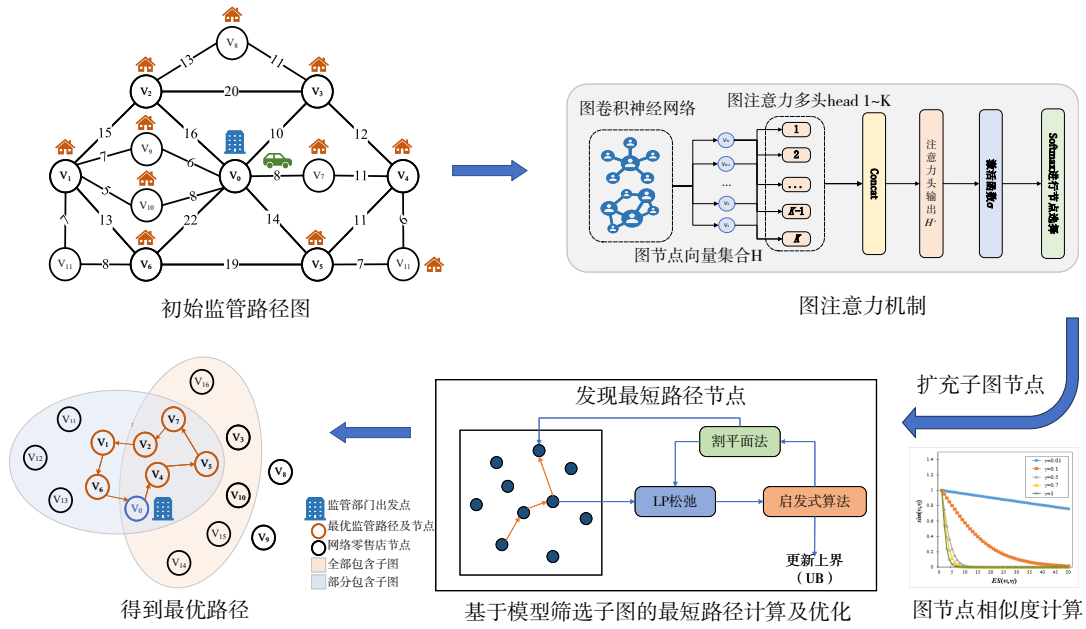


Fig. 2 Model framework

图 2 模型框架

器得到监管最优线路。

### 2.1 图注意力机制

假设一个算例网络包含  $N$  个节点,图注意力机制的输入是经过全量图卷积神经网络 GCN 计算的每个节点向量集合  $H = \{h_1, h_2, \dots, h_i, \dots, h_N\}, h_i \in \mathbb{R}^D, D$  为节点向量空间大小。图注意力机制的输出为  $H' = \{h'_1, h'_2, \dots, h'_i, \dots, h'_N\}, h'_i \in \mathbb{R}^{K \times D'}$ , 其中  $D'$  为每个节点经过多头注意力编码后的特征维度。采用  $W \times h_i$  将节点从原来的  $D$  维变换为  $D'$  维, 其中  $W \in \mathbb{R}^{D' \times D}$ 。随后采用自注意力机制计算节点  $i$  对节点  $j$  的注意力系数  $\alpha_{ij}$ , 如下:

$$e_{ij} = \sigma(\omega_{ij} \times \text{Concat}(W \times h_i, W \times h_j)) \quad (1)$$

$$\alpha_{ij} = \text{softmax}(e_{ij}) = \frac{\exp(e_{ij})}{\sum_{k \in N} \exp(e_{ik})} \quad (2)$$

其中,  $e_{ij}$  表示节点  $j$  对节点  $i$  的重要性,  $\text{Concat}$  表示向量连接操作,  $\omega_{ij} \in \mathbb{R}^{2D'}$  为权重向量,  $\sigma$  为激活函数。  $e_{ij}$  通过  $\text{Softmax}$  归一化操作得到  $\alpha_{ij}$ 。本文采用多头注意力机制提升训练效果, 并保证模型稳定性。对于节点  $i$ , 模型对  $K$  个注意力头  $\text{attn-head}_i^k, k \in K$  同时训练并将拼接结果输出, 如下:

$$\text{attn-head}_i^k = \sigma\left(\sum_{j \in N} \alpha_{ij}^k W^k h_j\right), \text{ where } (v_i, v_j) \in A \quad (3)$$

$$h'_i = \text{Concat}(\text{attn-head}_i^1, \dots, \text{attn-head}_i^k, \dots, \text{attn-head}_i^K) \quad (4)$$

其中,  $\text{attn-head}_i^k$  为第  $i$  个节点的第  $k$  个注意力头;  $\alpha_{ij}^k$  为第  $k$  个注意力头中节点  $i$  对  $j$  的注意力系数,  $W^k$  为第  $k$  个注意力头  $\text{attn-head}_i^k$  的可学习权重矩阵。经简化后网络的边向量同样可以采用类似思路。

### 2.2 网络数据标签生成

#### 2.2.1 基于原始求解的训练样本数据标签生成

基于图注意力机制的节点选择, 本质上是在算例网络数据中为每个节点计算分类标签。假设一个监管路径优化问题有  $N$  个节点, 则节点标签输出为  $P = \{p(v_i \in V^*), i \in [1, N]\}$ , 表示每个节点属于最优子集  $V^*$  的概率。对于该路径优化问题, 可以提前通过优化求解器得到最优解  $V^*$  作为种子训练样本,  $V^*$  中的节点标记为 1, 其他节点标记为 0。将  $T$  个具有同样节点数  $N$  的优化问题构建训练样本, 图注意力机制对该训练样本进行训练后, 可对  $V^*$  未知算例中的节点进行 0-1 分类预测。本文对最优路径的边进行了同样标注。

#### 2.2.2 基于图网络空间相似度的训练样本数据标签生成

上述训练预测方法存在预测最优子集稀疏的问题。假设预测的最优子集为  $V^*$ , 结合网络空间结构相似度的思想对预测子集进行扩充。借鉴已有前沿研究思想<sup>[22]</sup>, 与预测子集  $V^*$  中的节点空间结构相似度越高, 其处于最优子集  $V^*$  中的概率会相对较高。基于欧式距离, 结合  $i$  和  $j$  两个节点的位置坐标、图注意力向量计算两个节点的距离为  $ES(v_i, v_j)$ , 则两个节点的相似度  $\text{sim}(v_i, v_j)$  公式如下:

$$\text{sim}(v_i, v_j) = \frac{2}{1 + e^{\gamma \times ES(v_i, v_j)}} \quad (5)$$

其中,  $\gamma$  用于控制  $\text{sim}(v_i, v_j)$  随着两节点空间距离增加而减小的速率(见图 3), 本文将  $\gamma$  设置为 0.1。

节点  $i$  属于最优节点子集的分值  $\text{score}(v_i \in V^*)$  为该节点与预测子集  $V^*$  中所有节点  $j$  之间的  $\text{sim}(v_i, v_j)$  之和。由于  $\text{score}(v_i \in V^*)$  存在大于 1 的情况, 故对其进行归一化得到节点  $v_i$  属于最优子集的概率  $p(v_i \in V^*)$ 。

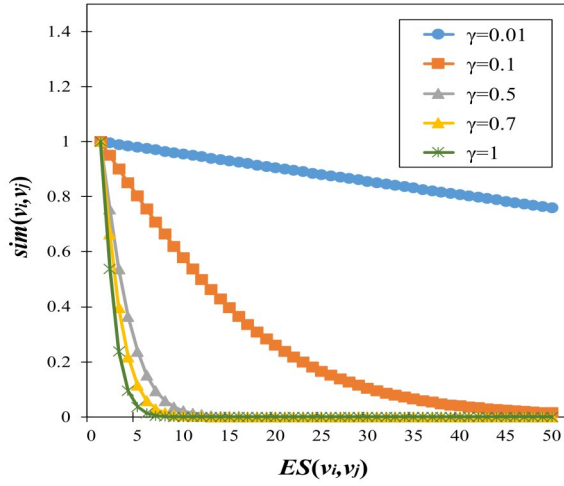


Fig. 3 Similarity curves under different settings

图3 不同设置下的相似度曲线

$$\text{score}(v_i \in V^*) = \sum_{v_j \in V^*} \text{sim}(v_i, v_j) \quad (6)$$

$$p(v_i \in V^*) = \frac{\text{score}(v_i \in V^*) - \min_{k \in V} \text{score}(v_k \in V^*)}{\max_{k \in V} \text{score}(v_k \in V^*) - \min_{k \in V} \text{score}(v_k \in V^*)} \quad (7)$$

### 2.3 基于近优子集的求解器求解

假设在监管路径规划图中,  $s(v_i, v_j)$  表示图中节点  $v_i$  到  $v_j$  是否直接连通;  $s(v_i, v_j) = 1$  为连通, 而为 0 时则是不连通;  $c(v_i, v_j)$  表示两个节点之间到达的时间成本。则求解器目标函数为:

$$\sum_{v_i \in V} \sum_{v_j \in V} c(v_i, v_j) \times s(v_i, v_j) - \sum_{v_i \in V} \sum_{v_j \in V} s(v_i, v_j) \quad (8)$$

$$\text{s.t.} \quad \sum_{v_i \in V} s(v_i, v_k) = \sum_{j \in V} s(v_k, v_j) = 1, \quad \forall v_i, v_j \in V, v_k \notin \{v_0, v_{m+1}\} \quad (9)$$

$$\sum_{v_i \in V} s(v_0, v_i) = 1 \text{ and } \sum_{v_i \in V} s(v_i, v_{m+1}) = 1 \quad (10)$$

其中, 式(8)为监管优化目标函数, 第一项表示总路径最小, 第二项表示路径经过尽可能多的节点; 式(9)为流量平衡性约束, 每个非起点和终点的节点进入及离开流量均为 1; 式(10)表示路径起点为 0, 路径终点为  $m+1$ 。求解器采用主流的 Bellman-Ford 算法进行最短路径计算, 该算法通过动态规划解决最短路径问题。此外, 该算法在求解过程中, 使用启发式搜索、LP 松弛、割平面法等方法缩小求解空间、加速收敛并提升精度。

## 3 实验与结果分析

### 3.1 实验设置

#### 3.1.1 数据集

本文选用东莞市 4 万余家烟草零售店的地理坐标作为实验对象, 所有零售店按照监管范围可划分为 100 多个区域, 平均每个区域的零售店超过 400 家。对每个区域依据坐标进一步划分为 20、10、5、2 个子区域, 对每个子区域

使用上文所述方法对数据进行标注, 获得训练数据集, 所有数据集基础信息如表 1 所示。不同节点数的数据集算例数量相同(如 20 个子区域平均每个区域的节点数设置为 20)。

Table 1 Experimental dataset information

表 1 实验数据集信息

区域节点数	训练算例数量	验证算例数量	测试算例数量
20	1 600	200	200
40	800	100	100
80	300	100	100
100	200	100	100
200	0	0	100
合计	2 900	500	600

#### 3.1.2 实验参数设置

本文通过参数网格实验对图注意力网络的超参数进行设置, 将图注意力头的数量设定为 4, 每个注意力头中神经元数量设置为 16, Dropout 设置为 0.6, 优化器设置为 Adam (Adaptive Moment Estimation)。为了评估模型在特定任务下的效果, 本文采用“全部包含子图占比”指标, 记为 RNS (Ratio of All Optimal Nodes Included In A Subgraph):

$$\text{RNS} = \frac{N_{an}}{N_{test}} \times 100\% \quad (11)$$

其中,  $N_{test}$  表示所有测试样本数量;  $N_{an}$  表示在测试样本数据中, 模型挑选出来的节点全部属于最优节点子集(全部包含子图)中的测试样本数量。与召回率、正确错误比率等指标的不同之处在于, RNS 更关注模型结果中全部包含子图的占比, 因为部分包含子图会缺失部分最优节点, 同时包含许多非最优节点, 进而导致结果值下降。实验的硬件环境为: CPU 32 核, GPU GeForce RTX 3090 Ti, 内存 100 G。

### 3.2 实验基准模型设置

本文设计了 5 个基准模型进行对比: ① PO (Path Optimization): 直接使用优化求解器对全量样本进行计算, 求得最优路径; ② R-SGPO (Randomly Subgraph Generation and Path Optimization): 从每个原始算例网络中选择节点, 将选择的节点及边形成子图代替原始算例, 使用优化求解器计算最优路径; ③ RF-SGPO (Random Forest Based Subgraph Generation and Path Optimization): 借鉴已有研究思路<sup>[15, 23]</sup>, 构建网络节点和边特征集, 并采用 Random Forest 进行训练测试以选择最优节点集合, 最后采用优化求解器求解; ④ MLP-SGPO (Multi-layer Perceptron Based Subgraph Generation and Path Optimization): 采用 MLP 替换 RF 进行训练测试以选择最优节点集合; ⑤ GA-SGPO (Graph Attention Based Subgraph Generation and Path Optimization): 使用本文提出的图注意力机制对每个算例的节点及边进行选择, 形成子图作为新的算例, 使用优化求解器计算最优监管路径。

### 3.3 实验结果

本文首先设置实验对最优注意力头以及注意力头中

的神经元数量进行选择。使用节点数为 20 的数据集对模型进行训练,部分实验结果如表 2 所示。将注意力头数量设置为 4,隐藏层神经元数量为 16,最优的 RNS 为 64.19%。而 RF-SGPO 和 MLP-SGPO 的 RNS 分别是 23% 和 18%,远低于 GA-SGPO。

**Table 2 The impact of different hyperparameters on model performance**

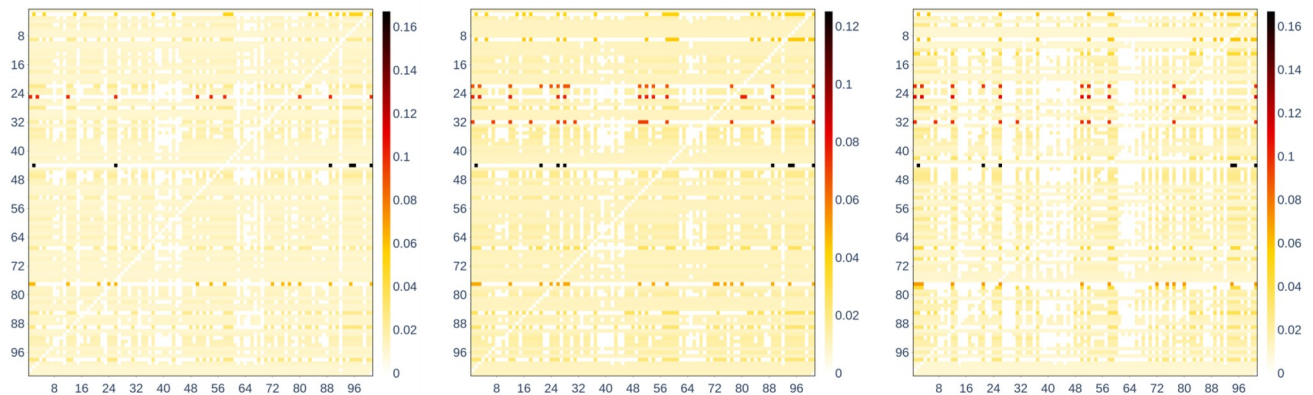
**表 2 不同超参数对模型性能的影响**

模型	参数设置	RNS/%
GA-SGPO 1	注意力头: 1;隐藏层神经元: 16	58.56
GA-SGPO 2	注意力头: 2;隐藏层神经元: 16	60.63
GA-SGPO 3	注意力头: 4;隐藏层神经元: 8	53.33
GA-SGPO 4	注意力头: 4;隐藏层神经元: 16	64.19
RF-SGPO	树的数量 100;树的深度: 5;	23.00
MLP-SGPO	隐藏层 3 层;每层神经元: 32, 16, 8	18.00

对节点数为 100 的部分测试样本进行注意力矩阵的热力图可视化,实验结果如图 4 所示。矩阵中某些行存在较为明显的水平深色点分布,说明某些节点和图中其他节点之间存在显著注意力关联,该关联对于预测最优节点具有

显著性作用。此外,热力图中部分节点与其他节点的注意力分布具有显著的模式关联,即只针对部分节点存在注意力值,而且不同节点的注意力系数分布也存在较显著的差异,说明模型能够学习不同节点的模式关联,该模式有助于预测节点分类标签。

对不同节点数的测试样本算例进行对比实验,对节点数为 25、40、80 时间设置为 800 s,即当求解时间超过 800 s 时不再迭代。对节点数 100、200 时间设置为 1 200 s。实验结果比较如表 3 所示(其中,黑色加粗数值表示在相应指标下所有模型所能达到的最优值,加下划线的数值表示次优值),评估指标包括运行时长 T,目标函数最小值 OBJ 和最优路径的长度 PL,目标函数的 OBJ 值越小,效果越好。结果表明,本文提出的 GA-SGPO 模型在不降低计算结果精准度的前提下,可显著降低计算时间,证明了所提出模型的有效性。模型的 OBJ 和 PL 指标显著优于 RF-SGPO、MLP-SGPO,进一步证明所提出的图注意力算法在最优节点选择上的有效性。



**Fig. 4 Attention matrix heatmap**

**图 4 注意力矩阵热图**

**Table 3 Comparison of experiment results**

**表 3 实验结果比较**

数据集	PO			R-SGPO			RF-SGPO			MLP-SGPO			GA-SGPO		
	T	OBJ	PL	T	OBJ	PL	T	OBJ	PL	T	OBJ	PL	T	OBJ	PL
20	0.69	<u>3.33</u>	3.23	0.12	9.93	4.98	<u>0.14</u>	2.27	16.86	0.12	7.43	10.68	0.25	3.87	<u>3.64</u>
40	88.73	<u>-38.57</u>	8.65	18.53	-17.34	11.44	66.88	-19.24	23.38	122.41	-25.19	26.03	<u>41.22</u>	<u>-31.29</u>	<u>9.94</u>
80	695.05	<u>-53.05</u>	9.43	<u>343.56</u>	-39.51	10.15	<b>308.56</b>	-40.09	34.19	579.12	-40.41	27.78	384.28	<b>-56.77</b>	<u>9.45</u>
100	1047.59	<b>-67.08</b>	<u>10.84</u>	<b>463.61</b>	-57.63	12.79	<u>465.07</u>	-33.46	35.97	823.51	-59.22	36.47	522.35	<u>-62.50</u>	<b>10.58</b>
200	1200.00	<u>-106.5</u>	<u>17.48</u>	775.31	-95.37	20.65	<u>798.56</u>	-51.2	41.13	1200.00	-71.25	39.85	824.32	<b>-118.47</b>	13.27

## 4 结语

针对城市烟草零售店较为密集的特点,本文提出了基于图注意力的节点、边选择及路径优化算法,通过剔除大概率不在最优路径上的节点以获取更小的求解空间,进而提高求解效率。现有前沿方法主要通过深度学习进行直

接求解或获得可行初始解,少数筛选节点边的思路主要是基于人工特征构建,较少考虑网络结构特性。实验表明,经过图注意力编码的节点,可以提升最优子集节点的预测精准度。本文方法能够用更短的时间找到优化的监管路径,更贴近烟草智慧监管的实际应用场景。

本文研究虽然发现了图注意力编码和最优子集节点的关联,但在实验中仍存在一些不能很好解释的算例。后

续将会构建理论框架对关联进行深入分析,并在此基础上设计更优的节点编码方式,以进一步增强模型的实际应用价值。

#### 参考文献:

- [1] DESROCHERS M, DESROSIERS J, SOLOMON M. A new optimization algorithm for the vehicle routing problem with time windows [J]. *Operations Research*, 1992, 40(2): 342-354.
- [2] ROUSSEAU L M, GENDREAU M, PESANT G, et al. Solving VRPTWs with constraint programming-based column generation [J]. *Annals of Operations Research*, 2004, 130(1): 199-216.
- [3] WANG L B, LUAN Y Q, XU L, et al. A method for path planning of unmanned aerial vehicle swarms based on dynamic cluster particle swarm optimization [J/OL]. *Journal of Computer Applications*, 2023-11-08. <http://kns.cnki.net/kcms/detail/51.1307.TP.20230428.1458.006.html>.  
王龙宝, 梁茵琪, 徐亮, 等. 基于动态簇粒子群优化的无人机集群路径规划方法 [J/OL]. *计算机应用*, 2023-11-08. <http://kns.cnki.net/kcms/detail/51.1307.TP.20230428.1458.006.html>.
- [4] BANI A, HALLAOU I, CORREA A, et al. Solving a real-world multi-depot multi-period petrol replenishment problem with complex loading constraints [J]. *European Journal of Operational Research*, 2023, 311(1): 154-172.
- [5] JI B, ZHOU S, ZHANG D, YU S. A branch-and-price-based heuristic for the vehicle routing problem with two-dimensional loading constraints and time windows [J]. *International Transactions in Operational Research*, 2023(early view): 1-33.
- [6] RIGHINI G, SALANI M. New dynamic programming algorithms for the resource constrained elementary shortest path problem [J]. *Networks: An International Journal*, 2008, 51(3): 155-170.
- [7] LI J X, HAN X. Revised pulse algorithm for elementary shortest path problem with resource constraints [C]//*Proceedings of Seventh International Symposium on Computing and Networking*, 2019: 37-44.
- [8] LODI A, ZARPELLON G. On learning and branching: a survey [J]. *Top*, 2017, 25(2): 207-236.
- [9] LI J X, LIU R C, WANG R N. Elastic strategy-based adaptive genetic algorithm for solving dynamic vehicle routing problem with time windows [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2023(early access): 1-18.
- [10] CHEN J W, BAO C C, ZHAO Z H. Research on route planning of logistics UAV based on improved A\* algorithm [J/OL]. *Software Guide*, 2023-09-05. <http://kns.cnki.net/kcms/detail/42.1671.TP.20230904.1059.008.html>.  
陈继伟, 包长春, 赵子恒. 基于改进A\*算法的物流无人机航迹规划研究 [J/OL]. *软件导刊*, 2023-09-05. <http://kns.cnki.net/kcms/detail/42.1671.TP.20230904.1059.008.html>.
- [11] BENGIO Y, LODI A, PROUVOST A. Machine learning for combinatorial optimization: a methodological tour d'horizon [J]. *European Journal of Operational Research*, 2021, 290(2): 405-421.
- [12] FONSECA G J C, SURITA G C, NETO J M, et al. A multi-agent system for solving the dynamic capacitated vehicle routing problem with stochastic customers using trajectory data mining [DB/OL]. <https://arxiv.org/abs/2009.12691>, 2022.
- [13] LARSEN E, LACHAPELLE S, BENGIO Y, et al. Predicting tactical solutions to operational planning problems under imperfect information [DB/OL]. <https://arxiv.org/abs/1807.11876v3>, 2018.
- [14] KRUBER M, LUBBECKE M E, PARMENTIER A. Learning when to use a decomposition [C]//*Proceedings of International Conference on AI and OR Techniques in Constraint Programming for Combinatorial Optimization Problems*, 2017: 202-210.
- [15] MORABIT M, DESAULNIERS G, LODI A. Machine-learning-based arc selection for constrained shortest path problems in column generation [J]. *Informatics Journal on Optimization*, 2023, 5(2): 191-210.
- [16] ZHANG R X, WU C X, SUN T C, et al. Deep reinforcement learning and research progress in path planning [J]. *Computer Engineering and Applications*, 2021, 57(19): 44-56.  
张荣霞, 武长旭, 孙同超, 等. 深度强化学习及在路径规划中的研究进展 [J]. *计算机工程与应用*, 2021, 57(19): 44-56.
- [17] VINYALS O, FORTUNATO M, JAITLY N. Pointer networks [DB/OL]. <https://arxiv.org/abs/1506.03134v2>, 2015.
- [18] FANG W, JIE Z B, LU H Y. Solving the traveling salesman problem with deep reinforcement learning algorithm based on multi-start and mask strategy [J/OL]. *Control and Decision Making*, 2023-11-06. <https://doi.org/10.13195/j.kzyjc.2022.0755>.  
方伟, 接中冰, 陆恒杨, 等. 基于多起点和Mask策略的深度强化学习算法求解覆盖旅行商问题 [J/OL]. *控制与决策*, 2023-11-06. <https://doi.org/10.13195/j.kzyjc.2022.0755>.
- [19] HILDEBRANDT F D, THOMAS B W, ULMER M W. Opportunities for reinforcement learning in stochastic dynamic vehicle routing [J]. *Computers & Operations Research*, 2023, 150: 106071.
- [20] HAMILTON W L, YING R, LESKOVEC J. Inductive representation learning on large graphs [C]//*Long Beach: Proceedings of Advances in Neural Information Processing Systems*, 2017.
- [21] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [C]//*Long Beach: Proceedings of Advances in Neural Information Processing Systems*, 2017.
- [22] QI M, LIN W H, NAN L, et al. A spatiotemporal partitioning approach for large-scale vehicle routing problems with time windows [J]. *Transportation Research Part E*, 2012, 48(1): 248-257.
- [23] KOOL W, BLIEK L, NUMEROSO D, et al. The euro meets NeurIPS 2022 vehicle routing competition [C]//*Proceedings of Machine Learning Research (NeurIPS'22)*, 2022: 35-49.

(责任编辑:孙娟)