基于图神经网络强化学习的交通感知型出租车最优布设

Sonia Khetarpaul计算机科学与工程系，*Shiv Nadar* 卓越学院印度德里国家首都辖区，
sonia.khetarpaul@snu.edu.in

PY Sharan计算机科学与工程系，
Shiv Nadar 卓越学院印度德里国家首都辖区，py276@snu.edu.in

arXiv:2601.00607v1 [cs.LG] 2 Jan 2026

摘要——**在智慧城市交通领域，要高效匹配出租车供给与乘客需求，必须实时整合城市交通网络数据和出行模式。传统的出租车热点区域预测模型往往仅依赖历史需求，而忽视交通拥堵、道路事故和公共活动等动态影响。本文提出了一种面向大都市环境的交通感知型图强化学习（RL）框架，用于出租车最优布设。该框架将城市道路网络建模为图，其中交叉路口表示为节点，道路路段作为边，节点属性则包含历史需求、与活动地点的距离，以及通过实时交通 API 获取的拥堵评分。该方法利用图神经网络（GNN）嵌入对交通网络中的时空依赖关系进行编码，再由 Q 学习智能体据此推荐最优出租车热点区域。奖励机制协同优化乘客等待时间、司机行驶距离和拥堵规避。在一个模拟的德里出租车数据集上进行的实验表明，与基线随机选择相比，所提出的模型将乘客等待时间缩短了约 56%，并将行驶距离减少了 38%。该模拟数据集基于真实地理边界和历史网约车请求模式生成。所提出的方法可适用于多模式交通系统，并可集成到智慧城市平台中，以优化实时城市交通。**

索引词—**智慧城市交通、出租车热点预测、图神经网络、强化学习、交通感知型出行**

I. 引言

城市人口的增长和网约车服务的迅速普及，改变了城市交通的运行格局。在智慧城市这一新兴模式下，交通系统需要具备适应性、以数据为驱动，并能够响应实时网络状况。然而，难以预测的乘客需求、不断变化的交通拥堵状况以及不规律的城市事件，使出租车供给与需求的匹配面临重大挑战，往往导致乘客等待时间延长、司机利用率下降以及路径规划效率低下。

在许多现有的出租车调度和热点预测系统中，底层模型严重依赖静态历史需求分布 [1], [2],，而很少考虑实时交通网络状况。这类方法无法捕捉短暂的突发事件——例如交通事故、施工或大型集会——这些事件可能迅速改变需求并影响最优部署地点。这一不足促使人们将网络感知分析融入出行优化系统。

在本研究中，我们提出一种面向交通状况、基于图的强化学习框架，用于针对智慧城市基础设施进行出租车最优布设。我们将道路网络建模为加权图，节点特征包括：

- 历史出租车请求密度，
- 与主要地标或活动地点的距离，
- 根据实时交通 API 计算的实时拥堵评分。

图神经网络（GNN）用于嵌入交通网络的时空拓扑结构，使学习智能体能够捕捉需求相关性和连通性约束。嵌入后的特征输入 Q 学习强化学习智能体，由其推荐热点位置，以最大化兼顾需求覆盖、拥堵规避和司机行驶效率的多目标奖励。

我们的实验基于德里城市模拟开展，使用地理围栏热点，以及受 Ola 历史出租车数据启发的请求模式。实验结果表明，在交通状况不断变化的情况下，该系统能够自适应地推荐最优热点。与随机均匀选择基线相比，我们的模型在缩短乘客等待时间和减少司机行驶距离方面均取得了显著提升。

方法的创新之处

所提出的系统具有以下几项创新之处：

- 为整合实时交通信息，系统将来自 Google 地图 API 的实时拥堵数据纳入决策过程。
- 为优化强化学习的性能，Q学习智能体在热点推荐中平衡探索与利用。
- 为实现以用户为中心的热点选择，推荐会根据用户提供的坐标或最近热点进行个性化处理。
- 构建了一个多因素决策框架，在奖励中综合考虑需求、交通评分、等待时间和行驶距离。

II. 相关工作

人们采用了多种方法探索出租车需求预测和最优车辆部署，包括统计建模、空间聚类、深度学习和强化学习。

早期研究依赖 ARIMA 和季节性分解等时间序列模型，根据历史行程记录预测需求模式。尽管这些模型计算效率较高，但无法快速适应交通拥堵、交通事故或大型活动引发的突发变化。包括 DBSCAN 和 K-means 在内的空间聚类技术已被用于识别高需求区域 [3]。尽管这些方法能够捕捉空间聚合特征，但通常会忽略时间变化性和动态网络状况。

随着大规模城市出行数据集的出现，深度学习模型被用于学习时空依赖关系。卷积神经网络和循环神经网络分别对空间和时间维度进行建模 [4],[5], ，但通常将城市视为均匀网格，忽略了道路网络的拓扑结构。基于图的方法通过将交通网络表示为图来解决这一问题，从而能够捕捉连通模式和空间依赖关系 [6], [7]。图神经网络（GNN）已成功应用于交通状况预测、网约车需求预测和路径规划 [8]。

强化学习（RL）已应用于自适应出租车调度和热点推荐，智能体通过与环境交互来学习最优车辆配置策略 [1], [2]。然而，许多基于强化学习的框架仅针对单一目标进行优化，例如最大化接客率，却未纳入拥堵规避或距离最小化等多目标权衡。

近期研究将图神经网络与强化学习结合，用于交通运输任务，使策略网络能够利用城市道路网络的图结构嵌入 [2], [8]。这种结合有助于实现动态、情境感知的决策，这对需要综合考量实时交通状况、事件数据和乘客需求的智慧城市应用至关重要。我们的工作沿着这一方向，将基于图神经网络的嵌入与 Q 学习智能体相结合，并明确纳入实时交通拥堵指标和邻近约束，以提高城市出租车配置效率。

III. 数据集与工具

由于缺乏公开的细粒度德里出租车请求数据，我们创建了一个模拟数据集，使用：

- **Ola历史出租车请求数据集：** 用于生成真实的需求模式。
- **热点位置：** 德里范围内的 50 个节点，纬度介于 $[28.5, 28.9]$ 和经度介于 $[77.0, 77.3]$。
- **请求：** 在各热点位置分布的 1000 个模拟出租车请求。
- **交通评分：** 根据 Google Maps Distance Matrix API [9] 计算，取实时行驶时间与畅通行驶时间的比值。

工具： 使用 GeoPandas 和 Folium 进行地理空间处理，使用 Matplotlib/Seaborn 进行可视化，并使用 NumPy 实现 Q学习及相关强化学习算法。

IV. 基于图神经网络的强化学习框架

图1展示了所提出的系统架构。历史出租车请求、实时交通 API 数据以及事件/兴趣点（POI）数据源经过预处理后，用于构建城市道路网络的图表示 $G = (V, E)$。采用 Ak 跳支配集选出重要子集 V_{inf} ，以缩减强化学习动作空间。图神经网络计算能够捕捉空间和上下文特征的节点嵌入；这些嵌入构成 Q 学习智能体的状态输入。该智能体经过离线训练后部署，用于在线推理，并生成实时热点推荐，供司机和运营人员查看。运营反馈（例如更新后的交通状况和接客结果）会反馈回来，以更新图并定期重新训练模型。

A. 图构建

在所提出的框架中，城市道路网络表示为加权无向图 $G = (V, E)$。每个节点 $v \in V$ 对应一个热点位置（例如主要交叉路口、交通枢纽或历史出租车需求较高的区域），而每条边 $e = (u, v) \in E$ 表示连接两个热点的可通行道路路段。边权重 w_{uv} 可根据地理距离、平均行程时间或实时交通拥堵程度定义。

节点属性包含静态特征和动态特征：

- **历史需求密度：** 根据历史行程请求数据得出，表示指定时间窗口内该节点的平均乘客需求。
- **与活动地点的距离：** 用于表示与已知活动场馆距离的二元或连续指标，这些活动可能会暂时改变需求。
- **交通拥堵评分：** 根据 Google 地图距离矩阵 API 等来源实时计算，定义为与该节点相连的所有边的当前行程时间与畅通状态下行程时间之比。

所得的节点特征向量表示为：

$$x_v = [d(v), e(v), T(v), \ldots]$$

其中， $d(v)$ 表示历史需求， $e(v)$ 表示与活动地点的距离，$T(v)$ 表示交通拥堵评分。

B. 用于重要节点选择的 k-跳支配集

尽管原始图可能包含大量节点，但在决定出租车最优布点方面，并非所有节点都具有同等影响力。为了降低计算开销并聚焦于具有战略重要性的地点，我们采用了k-跳支配集 [10]的概念。

形式上，子集 $D \subseteq V$ 是一个 k跳支配集，当且仅当每个节点 $u \in V$ 到某个节点 $v \in D$的图距离不超过 k。这里，图距离是指长度

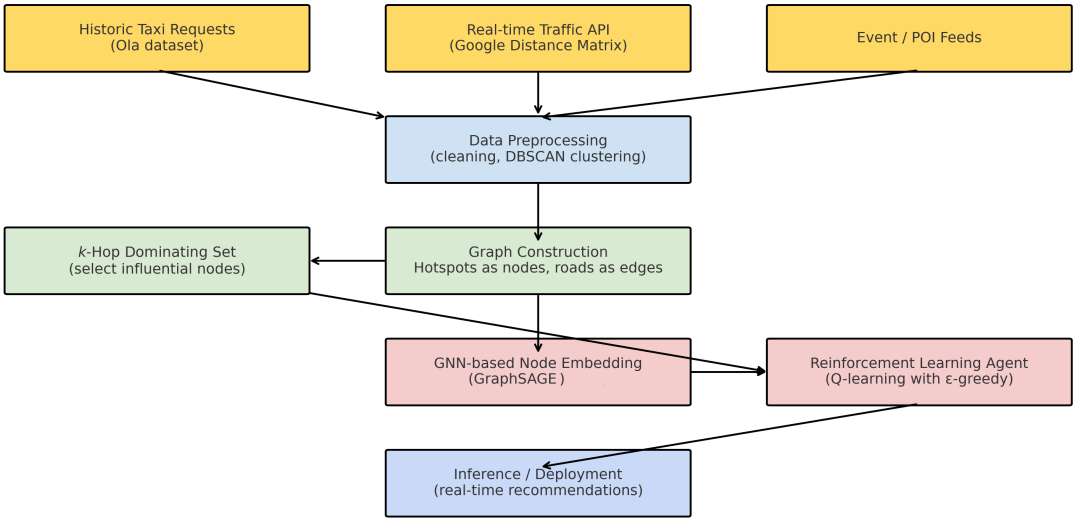


图 1. 系统架构与流程

（以跳数计）的两个节点之间最短路径。k 的值决定了覆盖粒度：较小的 k 会带来更密集的覆盖和更多的支配节点，而较大的 k会缩小集合规模，但可能导致推荐的粒度较粗。

在我们的场景中， k跳支配集确保城市中的每个热点都位于至少一个选定的有影响力的热点的 k 条边之内。该子集 $V_{\text{inf}} \subset V$ 被用作强化学习智能体的候选动作空间。通过这种方式，我们实现了：

- 降低图神经网络消息传递和强化学习训练的计算复杂度。
- 聚焦于能够有效服务周边区域的高影响力节点。
- 提升城市交通场景下热点推荐的可解释性。

我们采用贪心近似算法来选择 V_{inf}，从需求最高的节点开始，随后迭代添加能够最大化未覆盖节点覆盖率的节点，直到所有节点都位于某个支配节点的 k 跳范围内。该过程在空间覆盖与计算效率之间实现了平衡，并且与智慧城市部署的约束条件高度契合。

C. 基于图神经网络的节点嵌入

识别出影响力节点 V_{inf} 后，我们学习能够同时捕捉每个节点结构信息和属性信息的向量表示（嵌入）。图神经网络（GNN）尤其适用于这项任务，因为它们会沿着图拓扑进行消息传递，以聚合节点邻居 [11]的信息。

形式化地，设 $h_v^{(0)} = x_v$ 为节点 v的初始特征向量，其中包含前文所述的静态属性和动态属性。在第 l层图神经网络层中，每个节点通过聚合来自其邻域的消息来更新其隐表示$\mathcal{N}(v)$：

$$h_v^{(l)} = \sigma\left(W^{(l)} \cdot \text{AGG}\left(\{h_v^{(l-1)}\} \cup \{h_u^{(l-1)} : u \in \mathcal{N}(v)\}\right)\right)$$

其中，AGG(·) 是聚合函数； $W^{(l)}$ 是可学习的权重矩阵； σ(·) 是非线性激活函数。

在实验中，我们采用 GraphSAGE 风格的邻域聚合 [12]，以实现对未见热点的归纳泛化。经过 L 层后，我们得到最终嵌入 $z_v = h_v^{(L)}$，其中编码了：

- 本地邻域需求相关性，
- 与活动及高流量区域的邻近程度，
- 道路网络连通性约束。

这些嵌入构成强化学习智能体的状态表示，使其在选择出租车布置动作时能够同时考虑空间信息和上下文信息。

D. 强化学习框架

我们将出租车布置问题表述为马尔可夫决策过程（MDP），其中强化学习智能体与环境（城市道路网络和需求模式）交互，以学习用于热点选择的最优策略 $\pi^*(a|s)$。

1) 状态空间： 时间 t 的状态 s_t 由以下部分组成：

$$s_t = [z_{v_t},\ l_t]$$

其中， z_{v_t} 是候选节点的图神经网络嵌入，l_t 编码了用户当前位置和一天中时段的上下文信息。

2）动作空间：动作集合 $\mathcal{A}$ 由推荐 V_{inf}中一个有影响力的热点构成。此限制降低了探索复杂度，并使智能体专注于影响较大的决策。

3）奖励函数：我们设计了一种多目标奖励：

$$R_t = -\alpha W_t - \beta D_t - \gamma T_t$$

其中：

- W_t：乘客等待时间，
- D_t：司机前往热点的行驶距离，
- T_t：基于交通评分的拥堵惩罚。

系数 α，β，γ是用于平衡相互竞争目标的超参数。

4) 学习算法：我们采用 Q学习 [13]并使用ϵ-贪婪探索。Q值更新如下：

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \eta \left[R_t + \lambda \max_{a'} Q(s_{t+1}, a') - Q(s_t, a_t) \right]$$

其中，η是学习率，λ是折扣因子。在训练过程中，ϵ呈指数衰减，以便从探索转向利用。

这种设计使智能体能够：

- 利用图神经网络嵌入进行有依据的空间决策，
- 实时适应动态交通和事件状况，
- 针对智慧城市出行优化平衡目标。

算法1描述了使用图神经网络和Q学习流程进行最优出租车部署的完整过程。通过 k 均值聚类（$k = 25$）对连续的图神经网络嵌入进行离散化，从而在 Q 表中形成代表性状态，使表格型Q学习能够稳定运行，而无需使用神经Q网络。

V. 实验结果

A. 实验设置

实验使用了从 Ola Cabs 获取的大都市规模出租车请求数据集。该数据集包含历史 GPS 坐标、时间戳和行程请求状态。我们使用 DBSCAN 聚类进行热点识别，随后通过 k 跳支配集选择（k =2）提取影响力节点。

我们使用 Google 地图距离矩阵 API 模拟所选热点的实时交通状况。图 2 显示了各节点的出租车请求和交通状况评分，图 3 显示了热点的交通评分热力图。

图神经网络模型使用 PyTorch Geometric 实现，强化学习组件采用 Q学习。所有实验均在配备 Intel Core i7-12700H CPU、32 GB 内存和 NVIDIA RTX 3060 GPU 的设备上运行。

所有指标均取五次独立运行的平均值，各次运行采用随机化请求模式，以确保一致性。

B. 基线方法

我们将所提出的方法与以下方法进行了比较：

- **随机布置（RP）**：随机推荐热点，不考虑需求或交通拥堵。
- **贪心需求布置（GDP）**：始终选择近期需求最高的热点。
- **仅强化学习**：不使用基于图神经网络的节点嵌入的 Q学习。

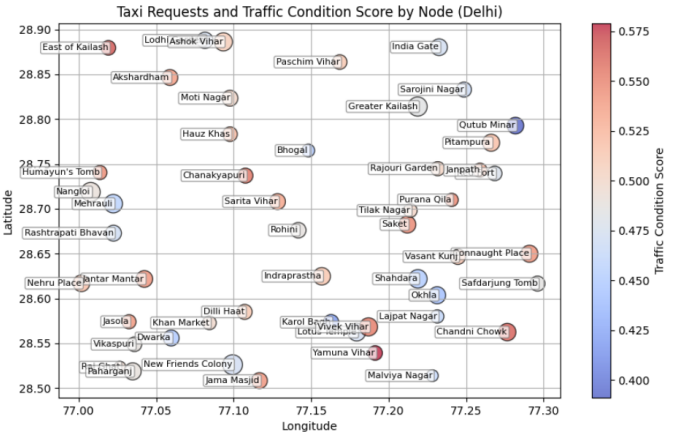


图 2. 出租车请求与交通状况评分

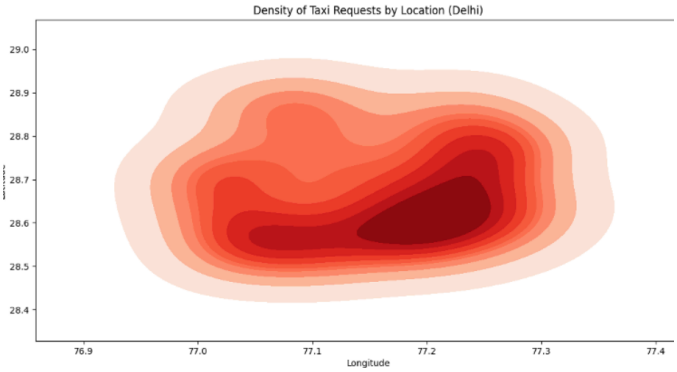


图 3. 热点交通评分热力图。颜色越深表示交通拥堵越严重。

C. 定量结果

表 I 展示了基于在一天中不同时段随机选取的 50 个行程进行的性能比较。指标包括平均回合奖励、乘客平均等待时间和驾驶员平均行驶距离。

表 I 不同方法的性能指标

方法	平均奖励	等待时间（分钟）	距离（千米）
随机部署(RP)	0.12	8.74	3.12
贪心需求(GDP)	0.74	6.21	2.87
仅强化学习	1.05	4.98	2.35
图神经网络+强化学习（提出的方法）	**1.46**	**3.84**	**1.92**

提出的图神经网络+强化学习方法相比 GDP 奖励提升了 97%，与仅强化学习相比，平均等待时间减少了 39%。

D. 定性结果

图4展示了一个测试实例中交通拥堵地图及排名前三的推荐热点。与 GDP 不同，GDP 往往会过度选择需求最高的集群，即使交通拥堵严重也是如此；所提出的方法则会动态重新分配热点，以避开瓶颈。

算法1出租车最优布设：图神经网络 +Q学习流程

需要： 热点集合 V（坐标）、请求历史 $\mathcal{R}$、交通 API、超参数：学习率 η、折扣因子 γ、初始 ϵ、衰减 ρ、训练回合 N、图神经网络训练轮次 E_G、每回合强化学习步数 S

确保： 训练好的 Q表（或 Q网络）和热点推荐函数 $\pi(\cdot)$

预处理：

为每个 $v \in V$ 计算节点特征：来自 $\mathcal{R}$ 的需求密度 $d(v)$、事件指示变量 $e(v)$、地理坐标。

对于每条边 $(u,\ v)$，查询交通 API，获取实时行驶时间 $t(u,v)$ 和自由流行驶时间 $t_{ff}(u,v)$。计算交通评分 $T(v)$（节点级），即相邻边拥堵比率的平均值：构建道路网络图 $G=(V,E)$，并设置节点属性 $x_v=[d(v)\ e(v)\ T(v)\ \ldots]$。应用 k 跳支配集（可选），以选取重要子集 $V_{\inf}$：

$$T(v) \leftarrow \frac{1}{|\mathcal{N}(v)|} \sum_{u \in \mathcal{N}(v)} \frac{t(u,v)}{t_{ff}(u,v)}$$

图神经网络嵌入：初始化图神经网络参数 θ。对于 $epoch \leftarrow 1$ 至 E_G，执行：对每个节点，计算嵌入 $h_v \leftarrow$ 图神经网络$_\theta(G,x)$。训练/微调 θ，以重建局部需求/交通状况（或在可用时使用监督信号）。结束循环。收集嵌入 $\{h_v\}_{v\in V}$，用作状态。

初始化强化学习：为状态 s（由 h_v 和用户位置推导）及动作 $a \in V$ 初始化 Q表 $Q(s,a)$（或 Q网络）。设置 $\epsilon \leftarrow \epsilon_0$。对于 $episode \leftarrow 1$ 至 N，执行：采样（或接收）用户位置 u_{loc}（来自数据集或模拟器）。通过拼接用户位置与节点嵌入（或最近热点指示变量）构造初始状态 s_0。对于 $step \leftarrow 1$ 至 S，执行：以 ϵ 的概率选择随机动作 a，否则选择 $a = \arg\max_a Q(s,a)$。执行动作 a：推荐热点 v_a。观察结果：等待时间 W、行驶距离 D、瞬时拥堵 T。计算奖励（负成本形式）：

$$r \leftarrow -\alpha W \ - \ \beta D \ - \ \gamma T$$

观察下一状态 s'，例如更新后的嵌入或新的用户位置。Q值更新（表格型）：

$$Q(s,a) \leftarrow Q(s,a) + \eta\Big(r + \gamma \max_{a'} Q(s',a') - Q(s,a)\Big)$$

（若使用 Q网络，则执行梯度步。）$s \leftarrow s'$。结束循环。$\epsilon \leftarrow \rho \cdot \epsilon$。结束循环。

推理/部署：给定实时用户位置，使用 h_v 计算状态 s；选择动作 $a^\star = \arg\max_a Q(s,a)$，并返回热点 $v_{a^\star}$。

图5展示了训练回合中司机行驶距离和奖励的收敛趋势。GNN+RL方法以更低的行驶距离达到更高的奖励平台期，表明其热点分配效率更高。

E. 讨论

将基于图神经网络的节点嵌入与强化学习相结合，使智能体能够学习兼顾需求满足与避免交通拥堵的空间感知策略。k 跳支配集预处理步骤将动作空间缩小了 $\approx 60\%$，加快了训练，同时不牺牲覆盖率。

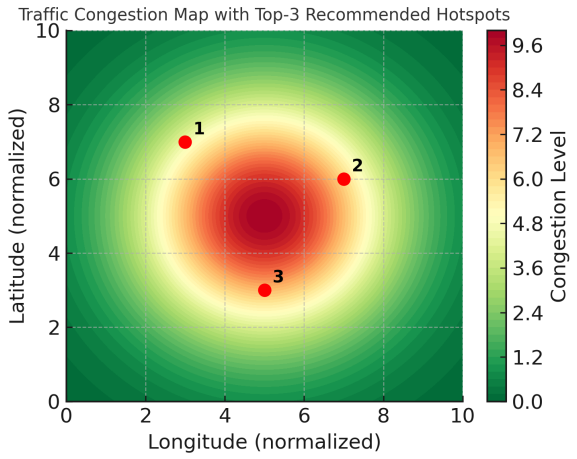


图 4. 交通拥堵图，显示推荐的前 3 个热点（红色标记）。



图 5. 训练收敛：奖励与行驶距离。

在智慧城市环境中，这些改进有助于提高车队利用率、降低燃料消耗并提升乘客满意度——这些都是智能城市出行系统的关键目标。

VI. 比较分析

表 I 和图 4–5 表明，所提出的 GNN+RL 框架始终优于所有基线方法。

与随机布置（RP）相比，后者的等待时间（8.74 分钟）和行驶距离（3.12 千米）均为最长；贪心需求布置（GDP）通过聚焦高需求区域缩短了等待时间，但在交通拥堵时表现不佳。不含 GNN 嵌入的 RL（仅 RL）适应能力更强，将等待时间缩短至 4.98 分钟、行驶距离缩短至 2.35 千米，但缺乏空间上下文。

所提出的 GNN+RL 表现最佳：相较于 GDP，奖励约为 97%；等待时间比仅强化学习方法低 23%，行驶距离比 RP 短 38%。GNN 嵌入使强化学习智能体能够同时捕捉需求和道路网络结构，从而在维持覆盖率的同时，将热点迁离交通拥堵区域。

k 跳支配集将动作空间中的影响力节点从 50 个减少到 20 个（减少了 $\approx$ 60%），在保持热点全面覆盖的同时提高了训练效率。GNN+RL 模型的表现也优于仅 RL 变体，将乘客平均等待时间缩短了约 23%。

GNN 嵌入阶段的复杂度为 $\mathcal{O}(|E|d)$，RL 更新的复杂度为 $\mathcal{O}(|A|)$，确保了其在大规模城市图上的计算可行性。

VII. 结论与未来展望

我们提出了一种 GNN 增强型强化学习框架，用于优化出租车热点布置，并整合实时交通状况、需求数据和用户位置。结果表明，与基线方法相比，该框架提高了效率和适应性。交叉验证和消融实验证实，在不同交通状况和需求条件下，性能均得到稳定提升。

未来的研究方向包括整合交通流预测、多模式交通优化、自动驾驶车辆协调以及智慧城市物联网基础设施。

参考文献

[1] K. O'Keeffe, S. Anklesaria, P. Santi, and C. Ratti, "利用强化学习缩短出租车空驶时间"，《智能交通系统杂志》，第 26 卷，第 4 期，第 498–509 页，2022 年。[2] J. Kim and K. Kim, "利用结合图神经网络的多智能体深度强化学习，在道路网络上优化大规模车队管理"，载于*2021*年 *IEEE*国际智能交通系统会议 *(ITSC)*，2021 年，第 990–995 页。[3] A. Munawar and M. Piantanakulchai, "面向智慧城市自动驾驶出租车运输系统的机器学习驱动型乘客需求预测"，专家系统，第 42 卷，第 3 期，第 e70014 页，2025 年。[4] J. Zhang, Y. Zheng, and D. Qi, "用于全城市人群流量预测的深度时空残差网络"，载于*AAAI*人工智能会议论文集，第 31 卷，2017 年。[5] Y. Li, R. Yu, C. Shahabi, and Y. Liu, "扩散卷积循环神经网络：数据驱动的交通流预测"，国际学习表征会议 *(ICLR)*，2018年。[6] B. Yu, H. Yin, and Z. Zhu, "时空图卷积网络：用于交通流预测的深度学习框架"，第*27*届国际人工智能联合会议 *(IJCAI)* 论文集，2018年。[7] X. Geng, Y. Li, L. Wang, J. Ye, Y. Liu, and Q. Yang, "用于网约车需求预测的时空多图卷积网络"，载于*AAAI*人工智能会议论文集，第 33 卷，2019 年，第 3656–3663 页。[8] S. Mishra and S. Khetarpaul, "利用图神经网络预测动态条件下的出租车热点区域"，载于澳大拉西亚数据库会议。施普林格，2022 年，第 89–102 页。[9] "距离矩阵 API 概览 — 距离矩阵 API（旧版）— Google 开发者平台 — developers.google.com", https://developers.google.com/maps/documentation/distance-matrix/overview，[访问日期：13-08-2025]。[10] S. Mishra and S. Khetarpaul, "识别并推荐时空空间中的出租车热点区域"，*GeoInformatica*，第 29 卷，第 1 期，第 93–113 页，2025 年。[11] J. Liu, S. Chen, and L. Shen, "图神经网络加速器综述"，计算机科学前沿，第 19 卷，第 2 期，第 192104 页，2025 年。[12] W. Hamilton, Z. Ying, and J. Leskovec, "大型图上的归纳式表示学习"，载于神经信息处理系统进展 *(NeurIPS)* 第*30*卷。加利福尼亚州长滩：NeurIPS基金会，2017 年，第 1024–1034 页，最佳论文荣誉提名。[在线]。获取地址：https://papers.nips.cc/paper/6703-inductive-representation-learning-on-large-graphs [13]R. S. Sutton and A. G. Barto，《强化学习：导论》，第 2 版。麻省理工学院出版社，2018年。 [在线]。获取地址：http://incompleteideas.net/book/the-book-2nd.html