

# 一个基于路网表征的城市交通仿真平台

王力超 陈灿宇 李京昊 王永瑶

100191

摘 要：MANTIS

关键词：

根据中共中央、国务院印发的《国家新型城镇化规划（2014～2020年）》，到2050年，中国城市化率需提高至70%以上。城市化过程中，一个完善的交通运输体系成为城市良好发展的前提，国家新型城镇化规划指出我们要完善综合运输通道和区际交通骨干网络，强化城市群之间交通联系，加快城市群交通一体化规划建设，改善中小城市和小城镇对外交通，发挥综合交通运输网络对城镇化格局的支撑和引导作用。然而，我国部分地区交通发展水平低，道路建设缓慢，路网结构不合理，导致严重的交通拥堵问题；同时，近年来中国经济的飞速发展带动了城市化进程的加速，而城市化进程带来的城市交通拥堵问题也日益严重。据中国交通运输部官网的数据，中国城市交通拥堵问题已经达到了前所未有的严重程度。

总之，上述情况引发人们思考：

（1）针对城市的拥堵现状，是否存在潜在的道路

结构优化方式？

（2）针对乡镇以及城市新区的开发建设，区域的道路该如何规划与建设？

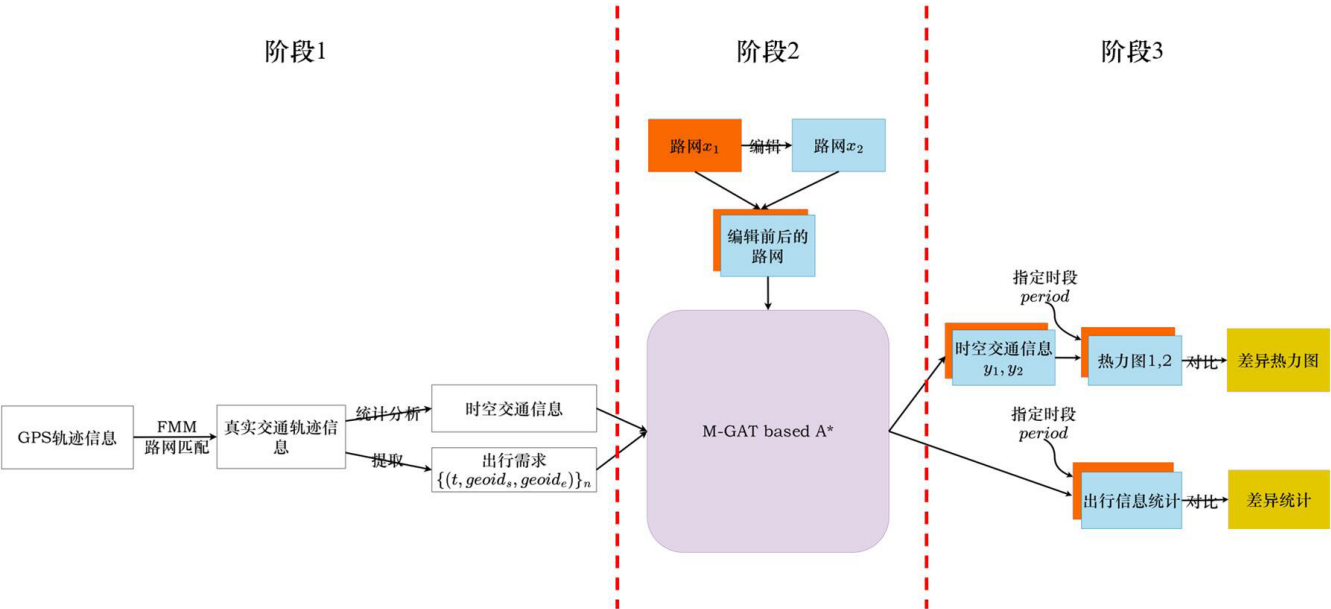
针对这两个问题，笔者设计了一个基于路网表征的城市交通仿真平台：MANTIS（Multi-layer Graph Attention Network for Transportation Simulation，多层图神经网络交通仿真），旨在为交通管理部门合理有效地对道路进行管理和建设提供参考。

## 2 MANTIS平台设计

仿真平台的工作流程如图1所示，分为三个阶段。

### 2.1 数据输入与路网编辑

原始的汽车行驶轨迹信息是通过GPS获取的。GPS获取到的汽车行驶数据通常以经度和纬度的形式呈现。这些数据可以通过GPS设备捕捉并记录车辆位置的变化，形成一系列坐标点，从而形成行驶轨迹。在此将一



1 MANTIS

条车辆的轨迹规范表达为:

$$\{(t_1, long_1, lat_1), \dots, (t_n, long_n, lat_n)\}.$$

其中,  $long_i, lat_i$  表示车辆的经纬度信息,  $t_i$  表示时间戳,  $n$  表示该轨迹的长度。

### 2.1.1 路网匹配

路网匹配 (地图匹配)<sup>[1]</sup>, 是指将 GPS 轨迹点匹配到路网中道路上的过程, 是轨迹预处理的一部分。由于 GPS 误差, 实际采集的 GPS 坐标点往往是在道路附近, 通常车辆只能在路网内行驶, 此时就需要通过路网匹配判断各个轨迹点实际在哪条路上, 既将轨迹序列转化为路段序列, 也起到修正误差的作用。匹配准确率会受到 GPS 误差和轨迹低采样率影响, 因此需要开发更优的匹配算法。

地图匹配是一种将 GPS 轨迹数据与地图上的道路网络进行匹配的技术。它是许多基于位置的服务的基础预处理步骤, 例如导航系统、交通管理和活动识别。最广泛使用的地图匹配算法之一是 Fast Map Match (FMM) 算法<sup>[2]</sup>。

FMM 算法是一种基于概率的地图匹配方法。它假设 GPS 测量数据存在噪声和不确定性, 并使用隐马尔可夫模型 (HMM) 对轨迹和道路网络进行建模。HMM 由一组状态和一组观测值组成。状态表示候选道路段, 观测值表示 GPS 测量数据。FMM 算法使用维特比算法搜索最优状态序列以匹配观测值。

FMM 算法包括两个主要步骤: 初始化和迭代。在初始化步骤中, 算法使用道路网络和 GPS 测量数据构建 HMM。它首先确定距离 GPS 测量点一定距离内的候选道路段。然后根据 GPS 测量误差和方向一致性估计每个候选段的可能性。最后, 用候选段和它们的可能性初始化 HMM。

在迭代步骤中, 算法使用维特比算法找到最优状态序列以匹配观测值。它从初始状态开始, 迭代地计算每个状态的可能性, 基于前一个状态和观测值。然后选择可能性最高的状态作为当前状态, 并重复该过程直到所有观测值匹配。最优状态序列表示与 GPS 轨迹相对应的最可能道路段序列。

通过路网匹配, 得到了轨迹数据集。轨迹数据集中

每一条轨迹数据为带有时间戳的路段 id 序列:

$$\{(t_1, geoid_1), \dots, (t_n, geoid_n)\}.$$

### 2.1.2 收集交通信息

从轨迹数据集中, 统计时空交通信息。将时空交通信息这个概念定义为: 各时段、各路段的车流量、平均车速等。例如, 对于编号为  $id$  的路段, 其对应的时空交通信息  $d_{id}$  定义为:

$$d_{id} = \begin{pmatrix} num_1 & \dots & num_n \\ v_1 & \dots & v_n \end{pmatrix}$$

其中,  $n$  表示将一天 24 小时划分成的时段个数,  $num_i$  为一天中第  $i$  个时段内, 在该路段通过的车辆数;  $v_i$  表示一天内第  $i$  个时段内所有通过车辆在该路段的平均车速。

### 2.1.3 提取 OD 需求

OD 需求指的是人们从某个起点 (Origin) 到达某个终点 (Destination) 的需求。这种需求在交通规划和运营中非常重要, 因为它能帮助决策者了解人们的出行模式、出行时间、出行目的地等信息, 以便更好地规划公共交通路线、调整交通流量、优化道路网络等。

通过统计 OD 需求, 交通决策者才能更好地了解公众出行需求, 提高公共交通的效率和便利性。

从轨迹数据集中, 提取出行 OD 需求集。每一条 OD 需求  $dem$  的形式如下:

$$dem=(t, geoid_o, geoid_d).$$

其中,  $t$  表示出发时间,  $geoid_o$  为起点路段的编号,  $geoid_d$  为终点路段的编号。

## 2 2 2

### 2.2.1 基于图神经网络的路网表征模型 M-GAT

在这一阶段, 设计了多层图注意力网络模型 M-GAT (Multi-layer Graph Attention Network)。

#### 2.2.1.1 模型结构

模型结构图如图 2 所示。

将路网定义为一个有向图  $G$ :

$$G=(V, E, F_v).$$

其中,  $V=\{v_1, \dots, v_{|V|}\}$ , 每一个节点  $v_i$  表示一个路段;

$E \subseteq V \times V$  为边集数组, 对任意  $e_{ij} = (v_i, v_j) \in E$ , 表示表示在路网中, 车辆能够从第  $i$  个路段到达第  $j$  个路段;  $F_V$  是路段的特征, 包括路段长度、车道数量、最大限速、路段种类。

接下来, 将建立图神经网络模型。这是一种针对图结构数据的神经网络模型, 它的输入是一个图结构, 输出是对节点或边的特征进行预测或分类的结果。对于没有明确训练任务的情况, 其输出就是一种路网的表征。

本文设计的模型是基于 GAT 网络<sup>[3]</sup>构成的。对于每一层 GAT, 有

$$\alpha_{ij} = \text{softmax} \left( \text{RELU} \left( \vec{a}^T \left[ W \vec{h}_i \parallel W \vec{h}_j \right] \right) \right).$$

其中,  $\alpha_{ij}$  表示节点  $i$  和节点  $j$  之间的注意力系数。权重矩阵  $W \subseteq F \times F'$ 。  $F$  表示输入节点特征的维数, 而  $F'$  表示该层输出节点的维数。其中的  $\vec{h}_i$  和  $\vec{h}_j$  表示, 节点  $i$  和节点  $j$  的节点特征, 如果这层 GAT 为输入层, 那么节点特征直接就是图的原节点特征  $x$ 。  $\parallel$  符号代表表示张量的 concat。例如,  $[[1,2],[3,4]] \parallel [[5,6],[7,8]] = [[1,2],[3,4],[5,6],[7,8]]$ 。

通过 concat, 把两个  $1 \times F'$  的张量粘合成了  $1 \times 2F'$  的大张量。然后乘以一个  $2F' \times 1$  的 attention kernel:  $\vec{a}^T$ 。得到的数值就是未加工的 attention 系数。接着通过激活函数 RELU 函数, 并进行 softmax 归一化。

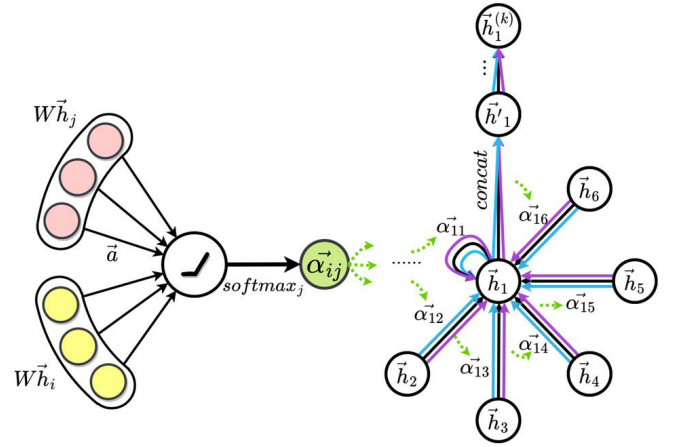
GAT 在聚合邻居节点的表示向量时, 使用注意力权重对其进行加权。这些权重是通过计算当前节点与邻居节点之间的相似度来确定的。

注意力机制可以使模型专注于与当前节点相关的邻居节点, 从而提高模型的效率和准确性。注意力机制还可以在学习节点表示时选择性地强调不同节点之间的关系。

接下来, 就要导出这层 GAT 关于节点  $i$  的输出特征  $\vec{h}_i$ :

$$\vec{h}_i = \text{RELU} \left( \sum_{j \in N_i} \alpha_{ij} W \vec{h}_j \right).$$

其中,  $N_i$  表示节点  $i$  的邻接节点,  $\alpha_{ij}$  是刚刚得到的注意力系数。



2 MGAT

和卷积神经网络 CNN 的操作类似, CNN 中对于每一层特征图的卷积核, 其实可以有多个, 而且每个卷积核相互独立, 从而使得输出特征图具有更多的 channel。在 GAT 中, 可以有  $K$  个独立的 attention 系数, 在训练过程中进行独立训练, 并且有着独立的注意力系数矩阵。此时上式变为:

$$\vec{h}_i' = \parallel_{k=1}^K \sigma \left( \sum_{j \in N_i} \alpha_{ij}^k W^k \vec{h}_j \right).$$

上式代表中间层的输出形式。最后输出层的输出形式为:

$$\vec{h}_i' = \text{softmax} \left( \frac{1}{K} \sum_{k=1}^K \sum_{j \in N_i} \alpha_{ij}^k W^k \vec{h}_j \right).$$

通过使用多头注意力机制, 能够使模型:

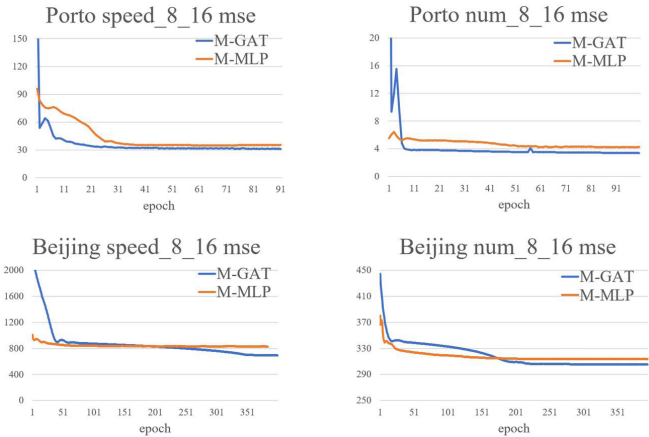
- (1) 关注图中更深度的子空间和位置信息;
- (2) 让模型能够抽取到更加丰富的特征信息;
- (3) 增强模型的稳定性和泛化能力。

#### 2.2.1.2 模型准确性验证

使用了真实世界的两个大规模数据集: 北京和波尔图的路网数据以及出租车轨迹。如表 1 所示。

表 1: 数据集来源及信息

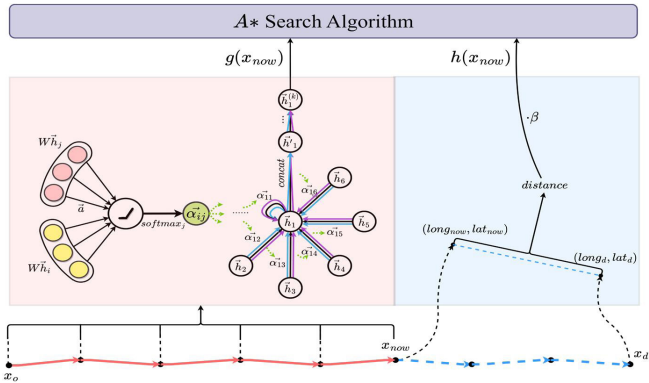
数据集	北京	波尔图
时间范围	2015/11/01-2015/11/30	2013/07/01-2014/07/01
轨迹条数	1018312	695085
路段条数	10903	38479



3



4



5 MGAT based A\*

为了进行更有效的对比，构建了与 M-GAT 模型类似但不蕴含图结构信息的多层感知机模型 modified multi-layer perception，记为 M-MLP。与基线模型 M-MLP 的训练效果如图 3 所示。可见，M-GAT 能够更快、误差更小的完成任务。

此外，仿真了北京三元桥道路施工案例中交通管制对车辆出行的影响。可以看到，仿真结果与实际情况相近。如图 4 所示。

2.2.2 将 OD 需求转化为预测轨迹

这一步的任务是将 OD 需求  $(t, geoid_o, geoid_d)$  转化

为预测轨迹  $\{(t_1, id_1), \dots, (t_n, id_n)\}$ 。这涉及到路径规划问题。选用 A\* 算法。A\* 算法是一种常用于路径搜索的启发式搜索算法。它采用了最优先搜索策略，结合了 BFS 和贪心算法的思想，在减少搜索空间的同时能够保证找到最优解。这种算法的核心在于利用一个估价函数来预估到目标节点的代价，并根据此来决定搜索的方向。因此，A\* 算法在许多领域都得到了广泛应用，如游戏 AI、路线规划、机器人路径规划等。A\* 算法的优点在于它具有高效性和准确性。通过合理地选择估价函数，A\* 算法可以在相对较短的时间内找到最优解，并且在大多数情况下能够减少搜索的节点数。如图 5 所示。

A\* 路径搜索启发式算法的核心是构建估价函数  $f(x)$  来计算每个节点的优先级：

$$f(x)=g(x)+h(x).$$

其中， $x$  表示当前所在的位置； $g(x)$  表示历史代价，使用前述模型预测路段通行时间  $t_i$ ， $t_i$  作为有向图的 weights 送入 A\* 算法； $h(x)$  表示未来代价，通过计算当前位置与目标位置的距离来得到。为保证  $g(x)$  和  $h(x)$  的影响相当， $\beta$  在实际操作中取 300。

A\* 算法在运算过程中，每次从优先队列中选取  $f(x)$  值最小（优先级最高）的节点作为下一个待遍历的节点。

在轨迹搜索中，采用多种其它路径搜索算法进行了实验。如表 2 所示，相比 BFS、Dijkstra，选用的 A\* 算法在算法执行耗时，生成轨迹的通行用时两个指标的综合考量中表现最好。

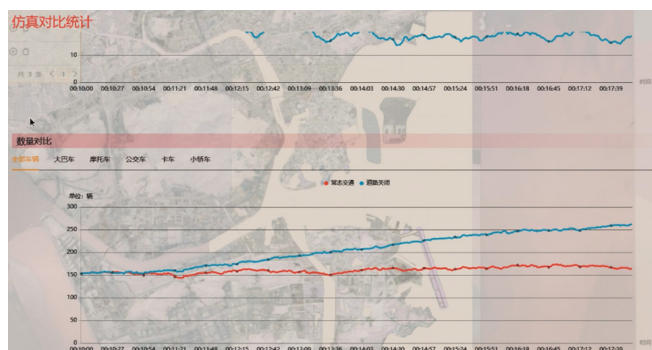
表 2：A\* 与其他路径搜索算法的对比

		算法执行 平均耗时 /s	生成轨迹的 平均通行用时 /s
Porto	BFS	0.00992	1585.8
	Dijkstra	2.94	817.2
	A*	0.146	892.9
Beijing	BFS	0.0416	1066.5
	Dijkstra	41.8	421.8
	A*	0.179	792.5

2 3 3

基于前两个阶段的成果，开发了可交互的仿真平台名为 MANTIS (Multi-layer Graph Attention Network for





6



7

Transportation Simulation)。

MANTIS 的核心工作原理为：在已有路网以及历史出行数据的训练基础上，在编辑后的新路网下，按照历史的 OD 需求，预测每条 OD 需求的轨迹，并利用这些轨迹进行统计，从而得到在新路网下的时空交通信息。

MANTIS 包括了路网导入编辑等功能，并搭载了 UCAPS 智慧交通模块以及 SUMO 仿真系统。这部分的具体效果可参见视频附件。

### 2.3.1 单路网（原始路网）时空交通分析

这一部分搭载了新兴好用、功能完善的 UCAPS 智慧交通模块，实现了对于单个路网的时空交通分析。在这一部分，MANTIS 具体涵盖了以下功能：

- (1) 道路数据统计显示；
- (2) 拥堵排名统计；
- (3) 拥堵影响关联分析；
- (4) 拥堵时空热力图展示。

### 2.3.2 路网编辑

用户可以进行新增路段以及删除道路的操作，通过这种方式模拟交通管制、道路改道。

### 2.3.3 编辑前后双路网宏观对比

在路网编辑模块的基础上，平台能够对编辑前后的路网下的交通状况进行宏观以及微观的对比分析。如图

6 和图 7 所示。

宏观分析：分析的粒度在整个路网。展示的是各个时段下的交通信息的数据对比。如图 6 所示。

微观分析：在路段路段下进行可视化展示。如图 7 所示。这部分通过对接 SUMO 仿真软件的功能模块来实现展示。

## 3 总结展望

交通仿真技术的应用可以帮助交通管理部门进行交通流量的优化和交通规划的设计，从而提高城市道路交通的效率和安全性。MANTIS 平台作为一种基于图神经网络的路网表征模型，能够模拟城市道路交通系统的各种情况，并通过可视化平台对模拟结果进行直观的展示，能有效为城市交通管理部门的管理和建设提供工作参考。

## 参考文献

- [1] , , . [J]. , 2018, 29(2): 225-250.
- [2] Yang C, Gidófalvi G. Fast map matching: an algorithm integrating hidden Markov model with precomputation [J]. International Journal of Geographical Information Science, 2017: 1-24.
- [3] Velikovi P, Cucurull G, Casanova A et al. Graph Attention Networks [C]. Proceedings of the International Conference on Learning Representations, 2018.

## 作者简介

2001-

2002-

2001-

2001-