



中国科学院大学
University of Chinese Academy of Sciences

博士学位论文

面向位置社交网络的下一个兴趣点推荐方法研究

作者姓名: 谭海宁

指导教师: 毕经平 研究员

中国科学院计算技术研究所

学位类别: 工学博士

学科专业: 计算机系统结构

培养单位: 中国科学院计算技术研究所

2021 年 10 月

Research on Next Point-of-Interest Recommendation

Methods in Location-based Social Networks

A dissertation submitted to

The University of Chinese Academy of Sciences

in partial fulfillment of the requirement

for the degree of

Doctor of Philosophy

in

Computer Systems Organization

by

Haining TAN

Supervisor: Professor Jingping Bi

Institute of Computing Technology,

Chinese Academy of Sciences

October 2021

中国科学院大学
研究生学位论文原创性声明

本人郑重声明：所呈交的学位论文是本人在导师的指导下独立进行研究工作所取得的成果。尽我所知，除文中已经注明引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的研究成果。对论文所涉及的研究工作做出贡献的其他个人和集体，均已在文中以明确方式标明或致谢。

作者签名：

日 期：

中国科学院大学
学位论文授权使用声明

本人完全了解并同意遵守中国科学院有关保存和使用学位论文的规定，即中国科学院有权保留送交学位论文的副本，允许该论文被查阅，可以按照学术研究公开原则和保护知识产权的原则公布该论文的全部或部分内容，可以采用影印、缩印或其他复制手段保存、汇编本学位论文。

涉密及延迟公开的学位论文在解密或延迟期后适用本声明。

作者签名： 导师签名：
日 期： 日 期：

摘要

随着各种具有定位服务的移动设备的普及，基于位置服务的移动社交网络（Location-Based Social Networks, LBSNs）发展迅速，具有位置信息的用户数据呈爆发式增长，比如 Foursquare、Yelp、美团等。在 LBSNs 中，用户可以签到和分享其访问的诸如餐馆、酒吧等兴趣点（Point of Interest, POI），每天会产生数以亿万条签到记录，具有巨大的应用价值。然而，数据量庞大的兴趣点在满足用户需求的同时，也为用户在有效时间内快速地发现兴趣点造成了困难。下一个兴趣点（Next POI, NPOI）推荐作为解决此类问题的主要手段，旨在利用用户的历 史签到序列和上下文信息来推断用户的行为偏好，为用户推荐感兴趣的新地点，从而解决用户的信息过滤和个性化需求问题，还能帮助 LBSNs 服务商实现智能化的位置服务，比如位置广告投放等。因此，下一个兴趣点推荐在 LBSNs 中扮演着越来越重要的角色，受到了学术界的极大关注。

下一个兴趣点推荐在实际应用过程中仍有许多有挑战性的问题值得研究，比如用户社交关系建模不准确、用户连续行为偏好刻画缺失以及数据匮乏环境下模型难以训练等。首先，每个用户的兴趣广泛分布在多个应用领域中，用户之间在多个领域中的兴趣关联也是高度差异的，并且用户社交关系带来的兴趣依赖是不可见的，导致推荐模型难以建模用户间多维度偏好差异，因此如何建模用户社交关系仍然是一个有挑战性的问题；其次，在特定的推荐时刻，时空上下文都会对用户的即时偏好产生重大影响，而现存兴趣点推荐模型未考虑每个兴趣点的时空上下文敏感性，导致推荐模型难以利用连续时间序列建模用户的连续状态变化，因此如何进行用户连续行为偏好刻画也是下一个兴趣点推荐的主要挑战之一；除此之外，由于基础设施建设不完善等原因，很多城市的签到数据匮乏，导致兴趣点推荐模型训练困难，存在严重的冷启动问题，导致推荐模型难以对少样本数据进行有效拟合，因此数据匮乏环境下的模型训练也是一个有挑战性的问题。

针对这些问题，本文围绕位置社交网络中下一个兴趣点推荐的现实需求，利用图表征、元学习等方法在隐含特征提取和知识迁移方面的优势，着力解决现存下一个兴趣点推荐方法面临的用户社交关系建模不准确、用户连续行为偏好刻画缺失以及数据匮乏环境下模型难以训练等极具挑战的问题，本文的创新性主要体现在：

- (1) 提出了一种多维度社交关系感知的下一个兴趣点推荐方法

针对用户社交关系建模不准确导致推荐模型难以建模用户多维度偏好差异的问题，本文提出了一种多维度社交关系感知的 NPOI 推荐方法 GraphSAR。该方法旨在采用维度级的注意力机制来学习用户表征之间的非结构化关联，使得模型可以更好地学习社交关系对用户偏好的影响，提升 NPOI 推荐效果。本文设计了一种基于注意力机制的兴趣偏好通道网络来建模用户之间的社交关系影响，并在影响注意力的引导下获取用户偏好表征。此外，为了有效利用用户-POI 异构交互网络中的丰富语义信息，本文设计了一种基于元路径的语义表征网络，利用图卷积网络对语义网络中具有丰富隐含语义信息的 POI 进行编码，以保证可以充分利用历史签到信息来提取节点间的关系特征。本文使用两个真实签到数据集评估了 GraphSAR 的效果。实验结果表明，对比最好的基线方法，GraphSAR 在 NPOI 推荐任务中的 HR@10 指标上获得了 4% 左右的提升。

（2）提出了一种用户动态偏好感知的下一个兴趣点推荐方法

针对用户连续行为偏好刻画缺失导致推荐模型难以建模用户的连续状态变化的问题，本文提出了一种用户动态偏好感知的 NPOI 推荐方法 PINN。该方法通过整合用户的历史静态偏好和当前动态偏好来进行 NPOI 推荐任务。PINN 包括两个主要部分：偏好记忆网络和偏好动态演化网络。在偏好记忆网络中，本文提出了一种基于用户群体移动行为模式的用户历史偏好表征网络，该网络可以建模用户群体移动模式，并在群体移动模式的指导下获取用户的历史偏好表征；在动态偏好演化网络中，本文设计了一种基于时空神经常微分方程的用户动态偏好感知方法，该方法首次将神经常微分方程用于 NPOI 推荐领域，定义连续的动态过程，可以接受任意时刻的输入数据，生成用户的动态偏好表征，克服了大多数时序推荐模型只能处理静态离散化的时间间隔的局限性，更适用于 NPOI 序列推荐任务。本文使用六个城市的真实签到数据集评估了 PINN 的效果。实验结果表明，对比最好的基线方法，PINN 在 NPOI 推荐任务中的 HR@10 指标上获得了 7% 左右的提升。

（3）提出了一种基于时空元学习的下一个兴趣点推荐方法

针对数据匮乏环境下，NPOI 推荐模型训练困难，导致推荐模型难以对少样本用户进行有效拟合的问题。本文提出了基于时空元学习的 NPOI 推荐方法 MetaRNN。该方法旨在为 NPOI 推荐模型生成高鲁棒性的初始化参数，使得模型仅利用少量的训练样本实现对 NPOI 的精准推荐。本文设计了一种面向 NPOI 推荐的元学习框架，将元学习应用于具有多类时空数据的 NPOI 推荐任务，该框架在建模时空差异性的基础上，通过构建多任务学习具有强泛化能力的初始参数，提升少样本场景下的 NPOI 推荐性能。此外，本文提出了一种基于时空注意力的

POI 序列编码网络，能够建模历史隐状态对当前状态的影响，保证在少样本场景下有效获取用户偏好表征。本文使用五个城市的真实签到数据集验证了 MetaRNN 的效果。实验结果表明，对比最好的基线方法，MetaRNN 方法在 NPOI 推荐任务中的 HR@10 指标上获得了 3% 左右的提升。

关键词：下一个兴趣点推荐，图表征学习，神经常微分方程，元学习

Abstract

With the popularity of various mobile devices with location services, location-based social networks (LBSNs) have developed rapidly. User data with location information has exploded, such as Foursquare, Yelp, and Meituan. In LBSNs, users can check in and share their visits to point of interest (POI) such as restaurant and bar. Tens of millions of check-ins are generated every day, which have huge application values. Although a huge amount of POIs meet the needs of users, it makes it difficult for them to quickly find POIs in an effective time. The next POI (NPOI) recommendation, as the main means to solve such problems, aims to infer the user's behavioral preferences based on their historical check-ins and contextual information. The NPOI recommendation can recommend new POIs to the user, which can solve the user's information filtering and personalized issues. The NPOI recommendation can also help LBSNs service providers to provide intelligent location services, such as location-based advertising. Therefore, the NPOI recommendation plays an increasingly important role in LBSNs and has received great attention from the academia.

The NPOI recommendation is still facing many challenges in the actual application, such as inaccurate modeling of social relationships, lack of characterization of continuous behavior preferences of users and difficulty in training models in the data-poor environment. First of all, the preferences of each user are widely distributed in multiple fields, and the multifaceted preferences between users are also highly different. Besides, the preference's dependency caused by the user's social relationships is invisible, leading to the difficulty of modeling the multifaceted preferences' relations between users using only social relationships. Therefore, how to model user's social relationships is still a challenging problem. Secondly, at a specific recommendation moment, the spatio-temporal context will have a significant impact on the user's momentary preference. However, the existing NPOI recommendation models do not consider the spatio-temporal context sensitivity of each check-in, which makes it difficult to model the user's dynamic preference using continuous time series. Thus, how to characterize the user's dynamic preference is also one of the main challenges for the NPOI recommendation. Finally, due to the imperfect infrastructure, many cities lack

check-ins, which makes it too difficult to train the NPOI recommendation model. In other words, there is a serious cold start problem for cities lacking of check-ins, which makes it difficult for the NPOI recommendation model to effectively fit users with limited check-ins. Therefore, the training of the NPOI recommendation models in a data-poor environment is also a challenging problem.

In response to these problems, towards the actual needs of the NPOI recommendation in LBSNs, this dissertation leverages the advantages of graph representation methods and meta-learning methods in the extraction of the hidden semantic features and knowledge transferring. Specifically, we solve the existing challenges such as inaccurate modeling of user social relationships, lacking of characterization of user continuous behavior preferences, and difficulty in training models in data-poor environments. The novelties of this dissertation are as follows:

(1) A multifaceted social-aware NPOI recommendation method

Towards the problem of inaccurate modeling of user social relationships, which makes it difficult for NPOI recommendation models to use simple relationships between users to model multifaceted preferences correlation between users, we propose a multifaceted social-aware NPOI recommendation method GraphSAR (Graph-based Social Aware Recommendation). This method aims to use a dimensional attention mechanism to learn the unstructured associations between users' representations, so that the model can better learn the influence of social relationships on users' preferences and improve the performance of the NPOI recommendation models. We design an POI preference channel network based on the attention mechanism to model the influence of social relations between users, and get users' preference representations under the social attention. In addition, in order to effectively utilize the rich semantic information in the user-POI heterogeneous information network, we design a multi-channel semantic representation network based on meta-paths, and uses graph convolutional networks to model rich implicit semantic information by the multi-channel semantic network. The POI is semantically encoded to ensure that the historical check-in information can be fully utilized to extract the relationship characteristics between POIs. Experiments on two realworld datasets show that compared with state-of-the-art baselines, GraphSAR achieves nearly 4% on HR@10 in NPOI recommendation.

(2) A dynamic preference intensive NPOI recommendation method

Considering the lack of characterization of user's continuous behavior preference,

which makes it difficult for NPOI recommendation model to use continuous time series to model the user's continuous states, we propose a dynamic preference intensive NPOI recommendation method PINN (Dynamic Preference Intensive Neural Network). PINN integrates the user's historical static preferences and current dynamic preferences to perform the NPOI recommendation. PINN includes two main modules: preference memory network and dynamic preference evolution network. In the preference memory network, PINN proposes a user historical preference-capture method based on the mobile behavior pattern of the user group; In the dynamic preference evolution network, PINN designs a user dynamic preference-aware method based on the spatial-temporal NeuralODE (Neural Ordinary Differential Equations) which is the first time to apply NeuralODE to the field of POI recommendation. By NeuralODE, PINN defines a continuous dynamic process, which can accept input data at any time to overcome the static discretization processing method of the time interval in most NPOI recommendation models, is more suitable for the NPOI recommendation. Experiments on six realworld datasets show that compared with state-of-the-art baselines, PINN achieves nearly 7% on HR@10 in NPOI recommendation.

(3) A spatial-temporal meta-learning based NPOI recommendation method

Taking into account the cold-start problem of NPOI recommendation models in data-poor environments, we propose a spatial-temporal meta-learning-based NPOI recommendation method MetaRNN (Spatial-Temporal Meta-learning Enhanced Recurrent Neural Network). MetaRNN aims to generate highly robust initialization parameters for the NPOI recommendation model, which can make the model to perform the NPOI recommendation with limited samples. We first design a meta-learning framework for NPOI recommendation, which can model the difference between different cities and generate general parameters by transferring the mobile behavior knowledge of multiple source environments to the target environment. In addition, MetaRNN proposes a spatio-temporal attention based sequence encoding network, which is used to solve the problem that the traditional RNN models cannot well model the spatial proximity between check-ins, to generate the effective embedding for each user according to the historial check-ins. Experiments on six realworld datasets show that compared with state-of-the-art baselines, MetaRNN achieves nearly 3% on HR@10 in NPOI recommendation.

Keywords: Next POI Recommendation, Graph Embedding, Neural Ordinary

Differential Equations, Meta-learning

目 录

摘要	I
Abstract	V
目录	IX
图目录	XIII
表目录	XV
符号列表	XVII
第 1 章 绪论	1
1.1 研究背景与意义	1
1.2 论文的贡献	6
1.3 论文的组织	9
第 2 章 下一个兴趣点推荐相关研究综述	11
2.1 兴趣点推荐概述	11
2.1.1 下一个兴趣点推荐定义	11
2.1.2 兴趣点推荐数据集概述	12
2.2 兴趣点推荐国内外研究现状	14
2.2.1 一般兴趣点推荐方法	16
2.2.2 下一个兴趣点推荐方法	18
2.3 小结	24
第 3 章 多维度社交关系感知的下一个兴趣点推荐方法研究	27
3.1 引言	27
3.2 多维度社交关系感知的下一个兴趣点推荐方法	30
3.2.1 问题定义及方法概述	30

3.2.2 基于图卷积神经网络的 POI 隐语义表征网络	32
3.2.3 用户偏好表征向量提取	37
3.2.4 下一个兴趣点推荐	39
3.2.5 复杂度分析	39
3.3 性能评估	41
3.3.1 实验设置	41
3.3.2 实验结果分析	44
3.4 小结	49
第 4 章 用户动态偏好感知的下一个兴趣点推荐方法研究	51
4.1 引言	51
4.2 用户动态偏好感知的下一个兴趣点推荐方法	53
4.2.1 问题定义与方法概述	53
4.2.2 签到行为表征	55
4.2.3 用户偏好记忆网络	56
4.2.4 用户偏好动态演化网络	58
4.2.5 用户动态偏好感知的兴趣点推荐	60
4.2.6 复杂度分析	61
4.3 性能评估	62
4.3.1 实验设置	62
4.3.2 实验结果分析	64
4.4 小结	69
第 5 章 基于时空元学习的下一个兴趣点推荐方法研究	71
5.1 引言	71
5.2 基于时空元学习的下一个兴趣点推荐方法	74
5.2.1 问题定义及方法概述	74
5.2.2 基于时空注意力的签到序列编码网络	76
5.2.3 基于元学习的参数优化	80
5.3 性能评估	83
5.3.1 实验设置	83
5.3.2 实验结果分析	87

5.4 小结.....	93
第 6 章 结论与展望	95
6.1 论文主要贡献和创新	95
6.2 下一步研究工作展望	96
参考文献.....	99
致 谢	109
作者简历及攻读学位期间发表的学术论文与研究成果	111

图目录

图 1.1 位置社交网络结构图.....	1
图 1.2 社交关系对下一个兴趣点推荐的影响.....	3
图 1.3 时空上下文对下一个兴趣点推荐的影响.....	4
图 1.4 下一个兴趣点推荐研究框架.....	5
图 2.1 兴趣点推荐技术分类示意图.....	16
图 3.1 社交关系对用户选择起到关键作用.....	28
图 3.2 某用户与其朋友在不同类别兴趣点上的签到相似度热度图.....	29
图 3.3 GraphSAR 架构示意图.....	31
图 3.4 基于元路径的语义表征示意图.....	32
图 3.5 基于元路径的多通道用户-POI 语义关系抽取示例.....	33
图 3.6 语义归纳表征层.....	34
图 3.7 基于变分自编码机的时间特征编码.....	36
图 3.8 Foursquare (纽约) 中不同元路径引导下的 GraphSAR 性能分析...	46
图 3.9 Foursquare (东京) 中不同元路径引导下的 GraphSAR 性能分析...	47
图 3.10 签到位置 P1->P6 的访问分布图	47
图 3.11 签到位置 P1->P6 表征向量的相似度	48
图 3.12 Foursquare (纽约) 上 dt 和 dr 对命中率和累计损益的影响	49
图 4.1 时间敏感的 NPOI 推荐示例图.....	54
图 4.2 PINN 示意图	54
图 4.3 用户偏好记忆网络.....	56

图 4.4 用户动态偏好演化网络结构.....	58
图 4.5 用户注意力权重热力图.....	67
图 4.6 记忆网络中表征维度 d 对模型的影响 (SF & LA)	68
图 4.7 超参数 γ 对模型的影响 (SF & LA)	68
图 5.1 全球 Foursquare 用户签到 POI 映射图	72
图 5.2 MetaRNN 示意图.....	75
图 5.3 基于时空注意力的编码器 (STA-RNN) 机制示意图	76
图 5.4 基于元学习的参数学习过程框架示意图.....	81
图 5.5 各数据集签到数据分布热力图.....	85
图 5.6 旧金山和洛杉矶数据集上的 HR@N 和 NDCG@N 结果	89
图 5.7 MetaRNN 的训练过程曲线.....	90
图 5.8 MetaRNN 的测试过程曲线.....	91
图 5.9 训练数据大小对 STA-RNN 和 MetaRNN 的影响	92
图 5.10 表征向量维度变化对 MetaRNN 的影响	93

表目录

表 1.1 公开兴趣点推荐数据集汇总.....	12
表 3.1 数据集统计数据.....	41
表 3.2 各类方法在 Foursquare 数据集上的推荐结果统计.....	44
表 3.3 各类方法在 Gowalla 数据集上的推荐结果统计	45
表 4.1 数据集统计数据.....	62
表 4.2 各类方法在 Foursquare (旧金山) 数据集上的推荐结果统计.....	64
表 4.3 各类方法在 Foursquare (洛杉矶) 数据集上的推荐结果统计	65
表 4.4 所选用户访问频次最高的 10 个类别结果统计.....	66
表 5.1 数据集统计数据.....	84
表 5.2 各类方法在 Foursquare (旧金山) 数据集上的推荐结果统计.....	88
表 5.3 各类方法在 Foursquare (洛杉矶) 数据集上的推荐结果统计	88

符号列表

符号	含义解释
S_1, S_2, \dots, S_n	签到数据丰富的源城市 1, 2, ..., n
D_1, D_2, \dots, D_m	签到数据匮乏的目标城市 1, 2, ..., m
U	整体用户集合
u	某目标用户
P	签到兴趣点集合
p	某签到兴趣点
C	签到兴趣点标签类别集合
c	某签到兴趣点标签类别
T	整体签到序列集合
\mathcal{T}_u	目标用户 u 的历史签到序列
r_*	某用户的单次签到记录
E	社交关系集合
t_*	特定时刻
l_*	签到兴趣点的位置
lat_*, lon_*	签到兴趣点位置*的经纬度坐标
p^*	指定待推荐的兴趣点
\mathcal{N}	用户-POI 交互网络
A	访问网络的邻接矩阵
D	访问网络的度矩阵
v_*	交互网络中的节点
$e(*,*)$	交互网络中的边关系
M	用户偏好记忆矩阵
E_c, E_l, E_t	标签类别矩阵、位置矩阵以及时间矩阵

$\mathbf{e}_c, \mathbf{e}_l, \mathbf{e}_t$	标签类别、位置以及时间的表征向量
\mathbf{r}_*	签到记录的表征向量
$\alpha, \beta, \lambda, \theta, b$	推荐模型的超参数
$\mathbf{z}_*, \mathbf{g}_*, \mathbf{h}_t$	时序模型的隐状态
\mathbf{W}_*	模型的权重矩阵
d	表征向量的维度值
$\ \cdot\ _2^2$	矩阵的 L2 范式
D_{KL}	两个隐状态之间的 KL 散度

第1章 绪论

1.1 研究背景与意义

近年来，随着各种具有定位服务的移动设备的普及，人们可以随时随地获取和分享他们的实时位置信息。与地理位置相关的新型在线社交媒体，即基于位置服务的移动社交网络（Location-Based Social Network, LBSN）应运而生并发展迅速，具有位置信息的用户数据呈爆发式增长，比如 Foursquare、Yelp、美团和大众点评等。位置社交网络不仅具有一般社交网络的社交属性，还将地理维度增加到社交网络中使得现实世界中的物理位置与虚拟网络中的数字编码联系起来。在位置社交网络中，用户不仅能够分享基础的位置信息，还可以和其他用户及位置进行实时交互。图 1.1 给出了位置社交网络的结构图，在位置社交网络中，用户可以通过具有定位功能的移动设备在一些位置或者兴趣点（Point-of-Interest, POI），例如著名景点、网红酒吧、博物馆等进行签到、发布内容及分享体验感受。

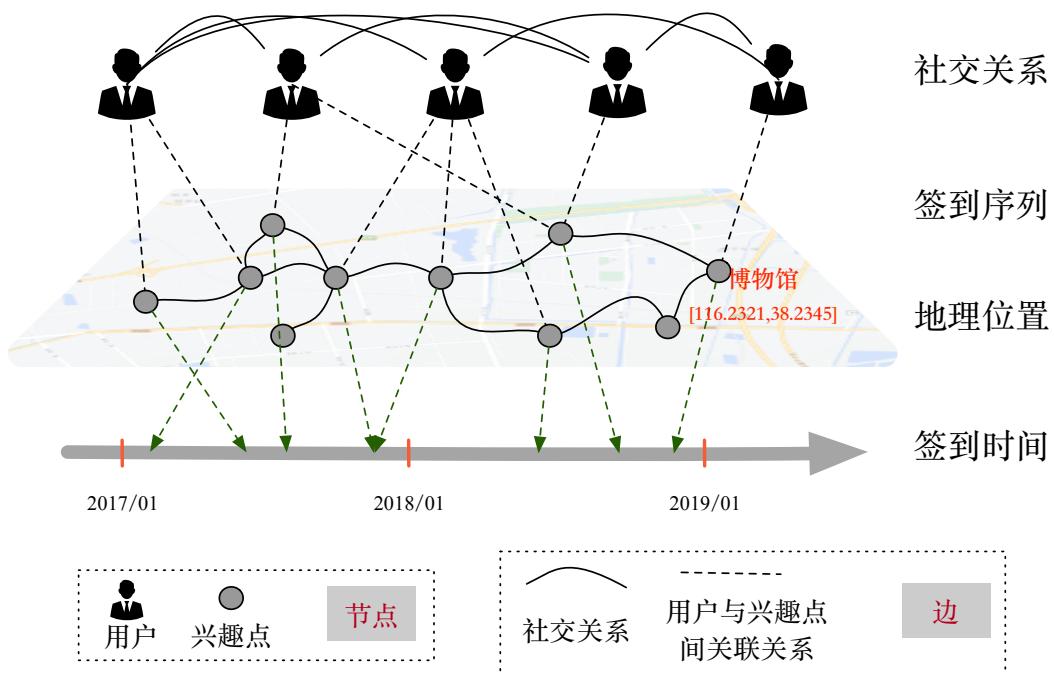


图 1.1 位置社交网络结构图

Figure 1.1 Illustration of location-based social networks

据统计，截止到 2019 年 12 月，Foursquare 的月活用户数已达 5000 万，提供约 105 万个兴趣点的信息查询服务，平台累计存储 120 亿条用户签到记录。截

止 2019 年 12 月 31 日，国内基于位置的社交网络平台美团的年度活跃用户达到 4.4 亿，日活跃用户数接近 7000 万，单季度营业额突破 1000 亿。依靠着庞大的用户基数，各类位置社交网络平台每天收集数以亿万条签到数据，且呈指数级增长。

LBSNs 中海量的兴趣点一方面满足了用户的各类需求，另一方面，也给平台的用户和服务提供商（兴趣点所有者）带来了“选择”难题。对于平台用户来说，海量的数据是过载的，用户越来越难以从中寻找自己感兴趣的商品和服务。对于服务提供商而言，从海量数据中挖掘和理解用户真实意图也变得越来越难。因此，如何精准地为用户推送其感兴趣的 interest point，提升用户体验，成为 LBSNs 中一项至关重要的任务，NPOI 推荐任务应运而生。NPOI 推荐在满足用户个性化需求的同时，也可以帮助 LBSNs 中的服务提供商深入理解用户和群体的基本信息，挖掘其行为偏好，辅助解决实际的商业问题，比如位置敏感的广告投放等。除此之外，NPOI 推荐通过对大量的用户行为数据进行深入分析，可以挖掘其中蕴含的丰富的行为模式和规律，协助进行城市事件的智能化感知和预测，因此，在巨大的商业价值之外，还可以助于智慧城市的发展，产生可观的社会价值。

近年来，围绕着兴趣点推荐任务，研究人员开展了大量的研究工作。按照任务场景将兴趣点推荐技术分为一般兴趣点推荐（General POI recommendation）和下一个兴趣点推荐（Next POI recommendation），这里一般兴趣点推荐^{[33][35][54][101]}是指像电影推荐那样为用户推荐排名靠前的 N 个兴趣点（Top-N），不考虑签到行为的上下文关联，而 NPOI 推荐则是在一般兴趣点推荐的基础上引入对用户最近签到序列的考量^{[15][29][43][55][98][112]}，为用户推荐在后续某时刻可能访问的兴趣点。

NPOI 推荐是一项对上下文信息高度敏感的任务，针对这一复杂的挑战，现有工作在传统的推荐算法的基础上做了很多改进，包括，扩展原始的协同过滤方法来建模兴趣点之间的地理影响和用户签到行为对时间的敏感性^[101]，改进马尔科夫模型来建模用户签到行为的序列化依赖^[104]等。近些年，深度学习技术凭借其在隐含特征提取以及复杂函数近似方面的优势在兴趣点推荐任务上表现优异，在建模上下文信息时能发挥巨大的作用。为了推断用户偏好，研究人员提出了很

多方法来建模多种异质的上下文信息。然而，这些方法存在其局限性，不能良好地捕捉上下文信息对用户偏好的影响。目前，在NPOI推荐领域仍然存在许多具有挑战性的问题：

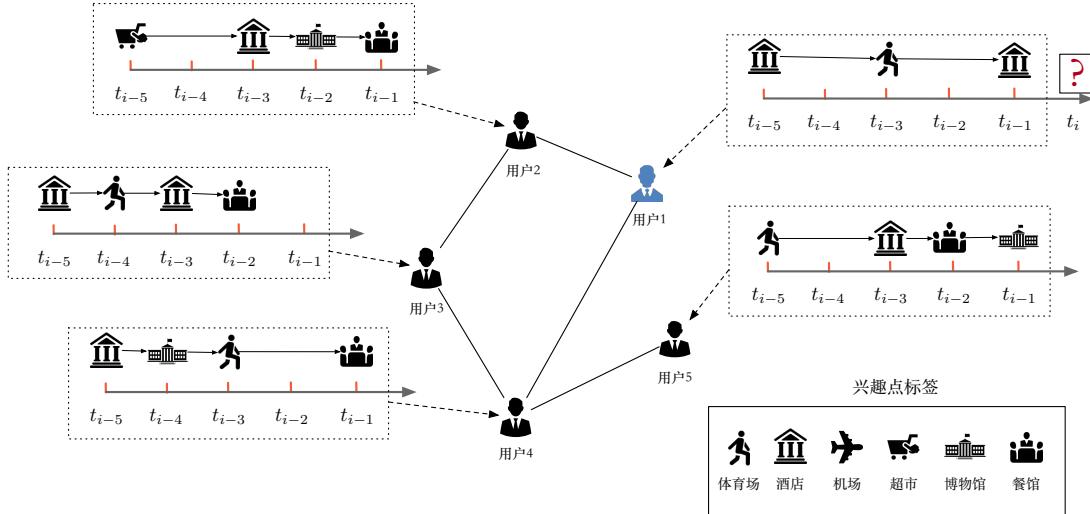


图 1.2 社交关系对下一个兴趣点推荐的影响

Figure 1.2 Social influence on the the next POI recommendation in LBSNs

(1) 用户社交关系建模不准确，导致NPOI推荐模型难以建模用户间多维度偏好差异。

图1.2展示了社交关系对下一个兴趣点推荐影响示意图。给定用户1的历史签到序列（按照时间序列进行排序）：酒店(t_{i-5})、体育场(t_{i-3})、酒店(t_{i-1})，在时刻 t_i 我们应该为用户推荐哪些兴趣点？如果不考虑社交关系对用户NPOI选择的影响，则根据从四个用户签到序列中挖掘的序列模式，则会以等概率推荐酒店和博物馆。离开酒店后，用户2和用户4喜欢参观博物馆，而用户3和用户5更喜欢去餐厅。相反，在考虑社交关系对下一个兴趣点选择的影响情况下，考虑到用户1与用户2和用户4之间的社交关系（见图1中它们之间的边连接），则系统更有可能在时间点 t_i 向用户推荐博物馆。由此可见，社交关系在下一个兴趣点推荐中起到至关重要的作用。

社交关系的建立往往是以一些共同兴趣为基础的，用户在进行兴趣点选择时可能会从朋友那里获得相关的经验，因此，用户的签到行为会受到其朋友签到行为的影响。现有方法将用户之间的社交关系提炼成一个标量值，表达用户之间的偏好相似度。然而，每个用户的兴趣是广泛地分布在多个领域中，且用户在多个

领域中的兴趣关联关系不明显，呈现出高度差异性。由此可见，由单一标量值来刻画社交关系难以表现用户之间的多维度兴趣关联，导致 NPOI 模型难以建模用户间多维度偏好差异，降低模型推荐性能。因此，如何建模用户之间多维度的社交关联仍然是一个有挑战的问题。

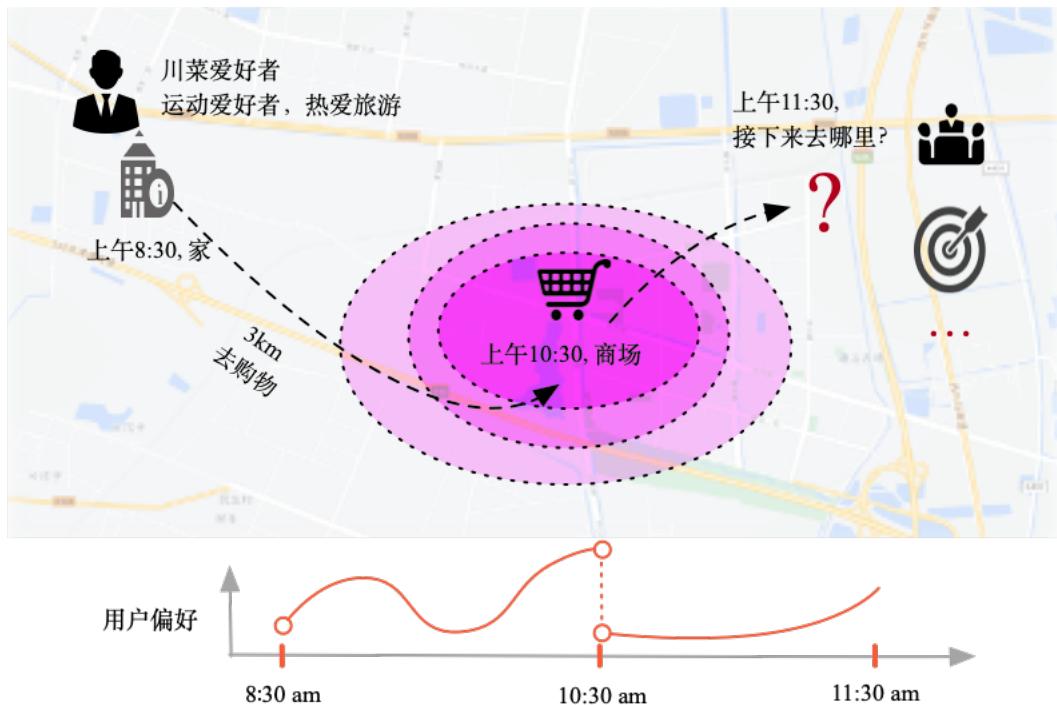


图 1.3 时空上下文对下一个兴趣点推荐的影响

Figure 1.3 Spatio-temporal context's influence on the next POI recommendation in LBSNs

(2) 用户连续行为偏好刻画缺失导致 NPOI 推荐模型难以建模用户动态偏好。

LBSNs 中，用户自身通常有一个长期的稳定兴趣偏好，但是每个具体的推荐时刻，时空上下文（Spatial-temporal context）都会对用户的即时偏好产生重大影响，这里时空上下文是指某签到记录前后签到记录包含的时间和空间地理位置信息，比如，当时间是中午 11:30，某平台得知用户在上午 10:30 刚打卡了某购物广场这一时空信息之后，如图 1.3 所示，考虑当前时间为 11:30（为通常就餐时间段），且用户刚刚打卡了购物广场，则平台在购物广场周边就近推荐餐馆将是不错的选择，推荐的 POI 可能被打卡的概率很大。由此可见，用户的上下文信息给 NPOI 推荐起到了至关重要的作用。然而，现有方法在集成时空上下文的过程中，使用了一种静态的全局加权机制，没有考虑每个兴趣点的强烈的时空上下文敏感

性，故而不能刻画每种上下文动态变化的重要程度，导致推荐模型难以利用连续时间序列建模用户状态的连续变化。因此，如何建模用户偏好变化的动态过程仍然是一个有挑战性的问题。

(3) 数据匮乏环境下 NPOI 推荐模型难以训练，导致模型难以拟合少样本数据。

由于基础设施建设不完善等原因，很多城市的签到数据匮乏，导致兴趣点推荐模型参数训练困难，存在严重的冷启动问题。现有方法通过利用多城市之间的共同用户作为锚链接，进行跨城市兴趣点推荐，由于存在共同用户这一假设约束，不能够很好的进行非共同用户的兴趣点推荐。并且由于城市之间的时空差异性导致用户会产生不同程度的兴趣偏移，加剧了兴趣点推荐模型的冷启动问题，导致推荐模型难以对少样本数据进行有效拟合。因此，如何缓解训练数据匮乏城市的兴趣点推荐冷启动问题仍然是一个有挑战性的问题。

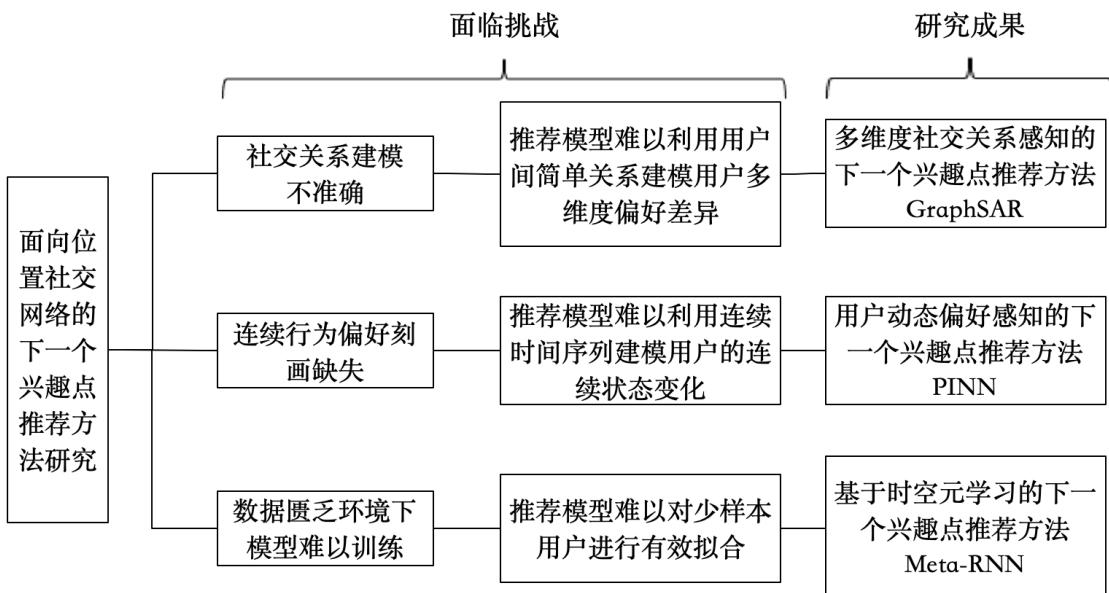


图 1.4 下一个兴趣点推荐研究框架

Figure 1.4 Architecture of next POI recommendation

如图 1.4 所示，该文的研究内容包括建模用户多维度的社交影响、建模用户偏好变化的动态过程和缓解兴趣点推荐冷启动等。尽管它们的目的都是为目标用户提供个性化的 NPOI 推荐，但是每项研究内容服务的场景仍然略有区别：

- 用户多维度的社交影响建模：在用户对兴趣点的签到记录的基础上，引入用户之间的社交关系，为目标用户输出一个其感兴趣却未签到的兴趣

点列表，适用于社交关系感知的 NPOI 推荐场景。

- 用户偏好变化的动态过程建模：在用户对兴趣点的签到记录的基础上，通过建模用户偏好随时间的动态变化过程，为目标用户在目标时间输出一个其感兴趣却未签到的兴趣点列表，适用于时间敏感的 NPOI 推荐场景。
- 训练数据匮乏带来的冷启动问题：在目标城市中用户对兴趣点的签到记录的基础上，引入学习到的多个源城市的移动行为知识，为目标城市中的用户提供一个其感兴趣却未签到的兴趣列表，适用于数据匮乏环境下的 NPOI 推荐场景。

针对上述挑战，本文围绕下一个兴趣点推荐任务的现实场景需求，利用图表征学习、元学习等方法在隐含特征提取和知识迁移方面的优势，研究并提出可行的解决方案：(1) 针对用户社交关系建模不准确问题，本文提出多维度社交关系感知的 NPOI 推荐方法，在进行 POI 多通道隐语义信息捕捉的同时，采用维度级的注意力机制建模用户之间多维度的兴趣偏好差异，提升 NPOI 推荐的性能；(2) 针对用户连续行为刻画缺失问题，本文提出用户动态偏好感知的 NPOI 推荐方法，将神经常微分方程用于签到序列表征，定义连续的动态过程，克服了大多数时序推荐模型静态离散化的时间间隔处理方式，提升推荐模型的性能；(3) 针对训练数据匮乏导致的推荐模型训练困难问题，本文提出基于时空元学习的 NPOI 推荐方法，将元学习应用于具有多类时空数据的 NPOI 推荐任务，在建模时空差异性的基础上，通过构建多任务学习具有强泛化能力的初始参数，缓解 NPOI 推荐的冷启动问题，提升少样本场景下的 NPOI 推荐性能。

1.2 论文的贡献

在国家自然基金 62077044, 61702470 与 62002343 等项目的支持下，本文深入分析了兴趣点推荐的国内外研究现状。并在此基础上，解决用户社交关系建模不准确、用户连续行为偏好刻画缺失以及数据匮乏环境下模型难以训练等挑战，提出了多维度社交关系感知的下一个兴趣点推荐方法、用户动态偏好感知的下一个兴趣点推荐方法和基于时空元学习的下一个兴趣点推荐方法。论文的研究成果

弥补了已有相关成果在下一个兴趣点推荐任务上的不足，提高了下一个兴趣点推荐的准确率。

本文的工作和创新性成果体现在以下四个方面：

(1) 深入调研了位置社交网络中下一个兴趣点推荐的背景及研究现状，分析了当前下一个兴趣点推荐面临的挑战

本文对兴趣点推荐领域的背景、内容和研究现状进行了深入的分析和探讨。首先简要介绍了下一个兴趣点推荐的定义、特点及分类，分析了当前下一个兴趣点推荐面临的主要挑战；其次，介绍了位置社交网络中兴趣点推荐的研究现状，并深入调研了已有研究在社交关系感知的下一个兴趣点推荐、时间敏感的下一个兴趣点推荐和数据匮乏环境下的下一个兴趣点推荐领域取得的成果，分析了这些方法存在的不足。

(2) 针对用户社交关系建模不准确问题，提出了多维度社交关系感知的下一个兴趣点推荐方法

在 LBSNs 中，利用用户的社交关系可以更好地建模用户兴趣点偏好，然而每个用户的兴趣广泛分布在多个领域中，且呈现高度差异性，这种差异性给模型建模社交关系的影响带来了很大挑战。此外，在 LBSNs 中，可以通过一张异质信息图来刻画用户的签到行为。在异质图中，除了节点（用户和兴趣点）间通过显式链接表现出来的浅层关系外，通常存在着其他更深层的隐藏关系。利用这些隐藏关系可以更好地进行节点的表征进程，获取有效的表征向量，更好地进行 NPOI 推荐。

针对上述问题，本文提出了多维度社交关系感知的下一个兴趣点推荐方法 GraphSAR。该方法在增强 POI 多通道隐语义信息捕捉的同时，采用维度级的注意力机制建模用户之间多维度的兴趣偏好差异。具体的，本文设计了一种基于注意力机制的兴趣偏好通道网络来建模用户之间的社交关系影响，并在影响注意力的引导下获取用户偏好表征。此外，为了有效利用用户-POI 异构交互网络中的丰富语义信息，本文设计了一种基于元路径的多通道语义表征网络，利用图卷积网络对多通道语义网络中具有丰富隐含语义信息的 POI 进行语义编码，以保证可以充分利用历史签到信息来提取节点间的关系特征。本文使用两个真实签到数据

集评估了 GraphSAR 的效果。实验结果表明，对比最好的基线方法，GraphSAR 在 NPOI 推荐任务中的 HR@10 指标上获得了 4% 左右的提升。

(3) 针对用户连续行为偏好刻画缺失问题,提出了用户动态偏好感知的下一个兴趣点推荐方法

在 LBSNs 中, 签到序列描述了目标用户的位置动态变化, 是用户行为偏好的直观体现。其中, 用户自身的兴趣是一种稳定作用的因素, 可以由用户长期以来的签到行为体现出来。除此之外, 用户在每一次的选择中, 会考虑当前的上下文信息, 这里的上下文信息有多种, 包括时间上下文和空间上下文。因此, 如果可以有效刻画上下文信息的动态作用, 并整合用户自身的静态兴趣偏好, 将极大地提升兴趣点推荐算法的性能。

针对上述问题, 本文提出了用户动态偏好感知的下一个兴趣点推荐方法 PINN。该方法通过整合用户的历史静态偏好和当前动态偏好来进行 NPOI 推荐任务。PINN 包括两个主要部分: 偏好记忆网络和偏好动态演化网络。在偏好记忆网络中, 本文提出了一种基于用户群体移动行为模式的用户历史偏好表征网络, 该网络可以建模用户群体移动模式, 并在群体移动模式的指导下获取用户的历史偏好表征; 在动态偏好演化网络中, 本文设计了一种基于时空神经常微分方程的用户动态偏好感知方法, 该方法首次将神经常微分方程用于 NPOI 推荐领域, 定义连续的动态过程, 可以接受任意时刻的输入数据, 生成用户的动态偏好表征, 克服了大多数时序推荐模型静态离散化的时间间隔处理方式, 更适用于 NPOI 序列推荐任务。本文使用六个城市的真实签到数据集评估了 PINN 的效果。实验结果表明, 对比最好的基线方法, PINN 在 NPOI 推荐任务中的 HR@10 指标上获得了 7% 左右的提升。

(4) 针对数据匮乏环境下模型难以训练问题,提出了基于时空元学习的下一个兴趣点推荐方法

在数据匮乏区域, 样本数据量少, 模型难以收敛, 限制了模型捕捉用户真实行为偏好, 造成推荐模型缺乏对少样本用户的推荐能力。元学习作为知识迁移的有效手段, 旨在从训练数据充足的源城市中学习可迁移的有用知识, 然后在目标城市中仅存在少量训练样本的条件下, 可以经过有限步迭代快速获取有效地适应

目标城市的推荐模型。除此之外，城市之间的时空结构差异也会给兴趣点推荐带来很大的影响，也是需要考量的重要因素。

针对上述问题，本文提出了基于时空元学习的下一个兴趣点推荐方法。该方法旨在为 NPOI 推荐模型生成高鲁棒性的初始化参数，使得模型仅利用少量的训练样本实现对 NPOI 的精准推荐。本文设计了一种面向 NPOI 推荐的元学习框架，将元学习应用于具有多类时空数据的 NPOI 推荐任务，该框架在建模时空差异性的基础上，通过构建多任务学习具有强泛化能力的初始参数，提升少样本场景下的 NPOI 推荐性能。此外，本文提出了一种基于时空注意力机制 POI 序列编码网络，能够建模历史隐状态对当前状态的影响，保证在少样本场景下有效获取用户偏好表征。本文使用五个城市的真实签到数据集验证了 MetaRNN 的效果。实验结果表明，对比最好的基线方法，MetaRNN 方法在 NPOI 推荐任务中的 HR@10 指标上获得了 3% 左右的提升。

1.3 论文的组织

第 2 章对下一个兴趣点推荐的相关背景、研究问题和研究现状进行了综合阐述，包括：简要介绍下一个兴趣点推荐问题的定义、兴趣点推荐的基本方法和面临的挑战，这为后续研究提供指导和方向；深入分析了已有研究在社交关系感知的下一个兴趣点推荐、用户动态偏好感知的下一个推荐和数据匮乏环境下的下一个兴趣点推荐上已经取得的成果，并分析了这些方法存在的不足。

第 3 章针对用户社交关系建模不准确问题，本文提出多维度社交关系感知的下一个兴趣点推荐方法 GraphSAR。详细介绍了 GraphSAR 中时空特征抽取、隐语义表征、下一个兴趣点推荐三个主要部分。然后，在两个真实数据集上进行实验，实验结果证明引入社交关系影响会提升下一个兴趣点推荐模型的推荐效果。

第 4 章针对用户连续行为偏好刻画缺失问题，本文提出用户动态偏好感知的下一个兴趣点推荐方法 PINN。首先，详细介绍了基于神经常微分方程（Neural ODE）的用户偏好表征模型架构，详细说明了每一层的操作步骤；然后，详细阐述了表征模型中的用户记忆网络和用户偏好演化网络两个主要部分；最后，在多个真实数据集上进行实验，实验结果验证了 PINN 在下一个兴趣点推荐任务上的

效果。

第 5 章针对数据匮乏环境下模型难以训练问题，本文提出基于时空元学习的下一个兴趣点推荐方法 MetaRNN。首先，分析了造成兴趣点推荐模型冷启动的原因；然后，详细介绍了 MetaRNN 中用于建模签到时序模式的基于时空注意力的序列编码器和基于时空元学习的参数优化方法；最后，在五个城市的真实签到数据集上进行实验，实验结果验证了 MetaRNN 在数据匮乏环境下的下一个兴趣点推荐任务上的效果。

第 6 章总结了本文的研究工作，并给出了未来的改进方向建议。

第2章 下一个兴趣点推荐相关研究综述

本章首先介绍位置社交网络中下一个兴趣点推荐的定义和当前常用的公开数据集；然后简要概述了当前移动社交网络兴趣点推荐领域国内外研究现状；最后详细介绍了位置社交网络中下一个兴趣点推荐领域中社交关系感知的下一个兴趣点推荐、时间敏感的下一个兴趣点推荐和数据匮乏环境下的下一个兴趣点推荐相关研究工作，并阐述了现存方法的不足之处。

2.1 兴趣点推荐概述

本节首先介绍了位置社交网络中下一个兴趣点推荐问题的定义，然后总结了当前兴趣点推荐领域的开源数据集以及数据特点（比如属性类别等）。

2.1.1 下一个兴趣点推荐基本概念

对位置社交网络中的用户进行 NPOI 推荐时，通常用数值或者向量来表示网络中的实体。例如，在描述用户签到数据时，签到地点用经纬度数值二元组进行记录，时间用数值进行记录，用户的属性信息也会转换成属性向量进行表示。本文研究的问题是如何构建一个合适的用户 POI 兴趣学习模型，并借助用户的历 史签到记录，为其更好地进行 NPOI 推荐。在详细描述 NPOI 推荐算法之前，本节先给出一些基本的符号解释和定义。

定义 1. 兴趣点(Point of Interest)。在位置社交网络(LBSNs)中，一个兴趣点 p_k 表示某一地理位置，由一对地理经纬度坐标以及相关的属性信息进行表示，通常为二元组，即 $p_k = (l_k, c_k)$ ，这里， l_k 表示兴趣点的地理位置， c_k 表示兴趣点的标签类别。

定义 2. 签到记录(Check-in record)。在位置社交网络(LBSNs)中，用户 u 的每一条记录 r_i 可以看成包含位置、兴趣点类别以及时间三个维度的属性描述信息的三元组， r_k ，可以形式化定义为：

$$r_k = (l_k, c_k, t_k)$$

其含义为用户 u 在时刻 t_k 拜访了类别为 c_k 的位置 l_k ，这里，标签类别和位置可以唯一标识一个兴趣点 p_k 。

定义 3. 签到序列 (Check-in sequence)。在位置社交网络(LBSNs)中，某用户 u 的签到序列由该用户的一系列的历史签到行为构成，可以形式化定义为：

$$\mathcal{T}_u = [r_1, r_2, \dots, r_K]$$

在上式中， K 表示用户 u 历史签到记录总数。为了方便，所有用户的历史签到序列可以形式化为： $\mathcal{T} = \mathcal{T}_{u_1}, \mathcal{T}_{u_2}, \mathcal{T}_{u_3}, \dots, \mathcal{T}_{u_{|U|}}$ ，这里， U 表示用户集合。

这里，下一个兴趣点推荐 (Next POI Recommendation) 是指给定签到数据的情况下，对特定目标 u 推荐其在下一个时间点可能访问的兴趣点。

2.1.2 兴趣点推荐数据集概述

兴趣点推荐数据集主要是从大型的基于位置的社交网络平台中采集获得，如 Foursquare、Gowalla、Yelp、Twitter、Facebook、Brightkite、Instagram 等^{[2][3][4][6]}。这些数据集中大多数都是有表格数据组成，主要包含用户和 POI 以及用户和用户之间的关系。数据集中的 POI 语义特征主要包括 POI 的类别以及用户的签到记录信息，其中用户的签到记录信息包括 POI 签到时间、地理位置信息（经度和纬度）、用户签到次数等。然而，用户的语义特征则主要包括签到次数、用户朋友以及签到记录等。

表 2.1 公开兴趣点推荐数据集汇总

Table 2.1 The summary of public datasets for POI recommendation

数据集名称	各数据表包含信息	数据表描述	收集日期
Foursquare	用户 ID、时间、兴趣点 ID、兴趣点名称、兴趣点位置信息、兴趣点类别	描述用户与 POI 的关系，包含所有用户签到记录以及签到时间	2010-2014
	用户 ID、用户朋友 ID	描述用户与用户的关系，包含某用户在移动社交网络中的所有朋友列表	
Gowalla	用户 ID、时间、经度、纬度	描述用户与 POI 的关系，包含每个用户的所有签到记录信息比如签到位置和时间	2009-2010
	用户 ID、用户朋友 ID	描述用户与用户的关系，包	

		含某用户在移动社交网络中的所有朋友列表	
	兴趣点 ID、名称、经纬度、照片数量、签到数量、用户数量、城市信息	描述 LBSN 中的每个 POI，包括用户计数、签到、照片、项目。每个 POI 地图上的名称，城市，日期时间，类别和半径也都包括在内	
	用户 ID、朋友数量、签到记录数量	描述 LBSN 中每个用户的个人详细信息	
Brightkite	用户 ID、签到时间、经度、纬度，兴趣点 ID	描述用户与 POI 关系，每行包含每个用户签到记录的时间和位置信息	2007-2012
	用户 ID、用户朋友 ID	描述用户与用户关系，主要描述了 Brightkite 中用户的朋友关系网络	
Yelp	兴趣点商业 ID、名称、地址、城市、国家、邮编、经纬度、星级、评论数量、开放状态、类别	描述相关 POI 的业务数据包括位置、属性以及类别等，这里属性信息包含该 POI 的是否接受外卖以及停车场设置状况等	2004--
	评论 ID、用户 ID、兴趣点 ID、星级、日期、评论文本	包含完整的评论文本数据，包括撰写评论的用户 ID 和兴趣点 ID	
	用户 ID、昵称、评论数量、开始时间、朋友信息、平均星级、点评信息	描述 Yelp 中用户的基本信息，包括用户的朋友关系以及与其关联的元数据	
	文本、日期、获赞数量、兴趣点 ID、用户 ID	描述用户对 POI 的概要提示，比评论短，并且倾向于传达快速的建议	
	照片 ID、兴趣点 ID、文本描述、标签	描述照片数据，主要包含照片标题描述和分类标签	
Weeplace	用户 ID、兴趣点 ID、时间、	描述用户与 POI 的关系，每	2010-2014

	经度、纬度、城市、类别	行包含用户的签到信息，签到日期和时间等信息	
	用户 ID、用户朋友 ID	描述用户与用户关系，主要描述了 Weeplace 中用户的朋友关系网络	
Instagram	用户 ID、经度、纬度、时间	描述用户与 POI 的关系，每行包含用户签到信息，比如访问的时间以及兴趣点的位置信息	2010--
Twitter	用户 ID、TweetID、经度、纬度、时间、兴趣点 ID、内容信息	描述具有签到内容的推文信息，每行表示一条推文，包含用户、POI 的位置和标识符以及 POI 关联的其他属性	2006--

表 2.1 总结了当前常用的兴趣点推荐公开数据集^[44]，主要分析了各个数据集的数据表格式。可以看出，不同类型的数据集含有的属性信息也有所区别。

2.2 兴趣点推荐国内外研究现状

推荐方法产生后，在短短的 10-20 年内迅速发展，引起了工业界和学术界的广泛关注^{[21][69]}。首先，推荐方法之所以备受关注并快速发展，是因为它能够给企业带来巨大利益。John Hagel 和 Marc Singer（1999）在其早期著作中就指出推荐方法作为“信息中介”对消费者产生的影响能够带来巨大的商业价值。当前，大多数电子商务网站也都部署了各种各样的推荐方法，从简单展示流行的商品到利用数据挖掘技术为用户提供“千人千面”的个性化推荐，这为用户购物提供便利的同时，也为商家增加了利润。

尽管推荐方法已经被许多电子商务网站研究并成功应用，比如 Amazon、Netflix 和淘宝等，但是面向兴趣点的推荐方法近期才被关注，仍处于起步阶段并不成熟。并且与传统的推荐领域不同的是，兴趣点推荐领域具有以下独特特征：（1）地理位置影响，在移动社交网络中，用户更喜欢访问附近的地点，而不是距离较远的地点。地理影响因素是兴趣点推荐领域与传统推荐领域区分开来的重要方面，并且很大程度上会影响用户的访问行为；（2）频率差异与数据稀疏性，

相较于传统的推荐算法，兴趣点推荐算法使用的数据更加稀疏，并且频率差异大。在传统的推荐领域中，用户通常通过所提供项目（例如书、电影、音乐等）的评分等级矩阵来表达自己的偏好。等级的额定值通常设置在一定数值范围内，如[0, 5]，等级越高表示用户的满意度越高。与传统的推荐领域不同的是，在兴趣点推荐领域中，用户的喜好程度常用用户签到频率矩阵来表示。与传统推荐领域中额定值相比，用户访问频次的范围差异更大，比如，一个用户可能在某一个位置签到数千次，而在其他位置只签到几次。除此之外，用户签到频率矩阵的稀疏度远远高于用户评分等级矩阵，这给兴趣点推荐带来了更大的挑战；(3) 社交影响力，通常我们假设朋友倾向于共享共同的兴趣，而用户经常倾向于向朋友寻求建议，因此社交关系对兴趣点推荐是有益的。

随着对兴趣点推荐领域研究的深入，越来越多的兴趣点推荐方法被提出。为了能够更加简明和清晰的对兴趣点推荐方法进行研究，不少学者^{[44][108][113]}按照不同的标准对兴趣点推荐方法进行了详细的分类。可以总结为以下三种方法，如图2.1所示：首先，根据用户签到行为的影响因素（Influencial factors）进行划分，可以分为地理位置影响（Geographical influence）^{[50][52][101][110]}、社交关系影响（Social influence）^{[20][42][95][109]}、时间序列影响（Temporal influence）^{[29][39][43][55][98][112]}以及评论内容影响（Content indications）^{[88][111]}四类；其次，根据推荐方法中数据的来源以及使用方式进行划分，将兴趣点推荐方法分为基于融合模型（Fused model）、基于联合模型（Joint model），比如矩阵分解（Matrix factorization）和图模型（Generative graphical model）^{[45][94][102][104]}、基于深度学习（Deep learning）的兴趣点推荐方法^{[25][28][30][96][107]}；另外，还可以根据任务场景将兴趣点推荐技术分为一般兴趣点推荐（General POI recommendation）和下一个兴趣点推荐（Next POI recommendation），这里一般兴趣点推荐^{[22][33][35][54][101]}是指像电影推荐那样为用户推荐排名靠前的N个兴趣点（Top-N），不考虑签到行为的上下文关联，而下一个兴趣点推荐是在传统兴趣点推荐的基础上引入对用户最近签到序列的考量^{[15][29][43][55][98][112]}，为用户推荐在后续某时刻可能访问的兴趣点。

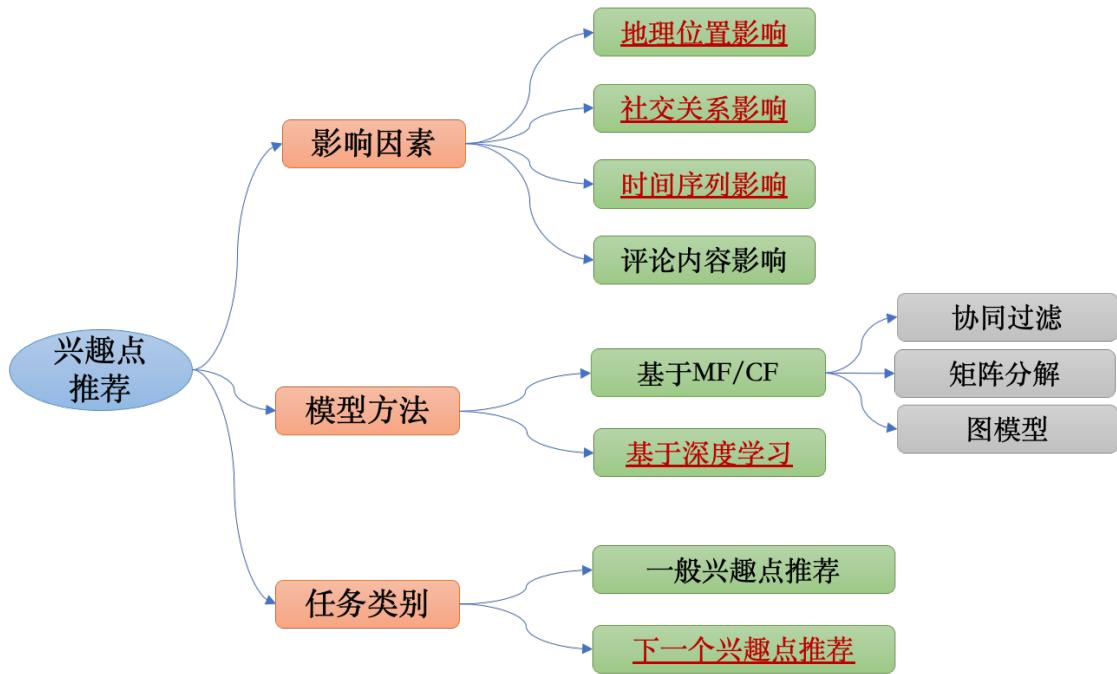


图 2.1 兴趣点推荐技术分类示意图

Figure 2.1 Illustration of taxonomies for POI recommendation

考虑到本文的研究重点是面向位置社交网络的下一个兴趣点推荐任务，因此在接下来将按照任务类别详细介绍现有的兴趣点推荐工作，即传统兴趣点推荐方法与下一个兴趣点推荐方法。

2.2.1 一般兴趣点推荐方法

一般兴趣点推荐方法根据使用数据的来源和方式，可以分为基于 MF/CF (Matrix Factorization or Collaborative Filtering) 的兴趣点推荐方法与基于深度学习的推荐方法。

2.2.1.2 基于 MF/CF 的推荐方法

用户的签到行为数据刻画了用户本身的兴趣点偏好与行为习惯。因此，基于 MF/CF 的兴趣点推荐方法根据获取到的用户签到行为数据，设计基于用户签到行为特征的因子模型或者协同过滤模型，实现对用户兴趣点偏好的捕捉。根据辅助信息的有无，基于 MF/CF 的方法可以进一步分为协同过滤类方法和基于辅助信息的协同过滤类方法。

协同过滤类方法主要是在用户和兴趣点之间学习一个相似性矩阵，然后给具有相似兴趣点偏好的用户推荐兴趣点。协同过滤类方法假设在历史签到记录中具

有相似兴趣的用户，在未来也会具有相似的偏好，广义上来说，协同过滤类方法可以分为两种：基于记忆（memory based）的协同过滤方法和基于模型（model based）的协同过滤方法。基于记忆的协同过滤方法^{[45][94]}通过在用户在兴趣点的签到向量上使用一些相似性度量方式，比如余弦相似度和皮尔逊相关，计算不同用户之间的相似度。至于基于模型的协同过滤方法，例如矩阵因子分解^{[49][51][52]}和张量分解^{[52][62]}，它们学习每个用户和每个兴趣点的潜在表示，然后采用点积计算 POI 与用户之间的相似度。

为了克服数据稀疏性问题，近些年来研究人员常常利用辅助信息（例如，地理信息、社会信息、时间信息等）来提高推荐性能。在本文中，地理影响主要是过渡距离，即特定 POI 到另一个兴趣点之间的实际物理距离。由于物理条件的限制用户通常更喜欢附近的地点而不是遥远的地点。之前大部分研究均是学习一个面向用户签到位置的概率分布，比如采用幂律分布^[101]对同一用户两个签到兴趣点之间的距离进行建模。考虑到用户运动模式的不同， Cheng 等^[20]提出了一种高斯混合模型对每个用户的签到行为得概率分布进行建模。Wang 等^[90]根据用户家庭位置和历史记录将其划分为不同的集群，然后进行兴趣点推荐。受移动社交网络中朋友具有相似的兴趣点偏好这一假设的启发，社交信息被用来提高推荐性能。Li 等^[51]和 Yang 等^[95]提出了一种直接利用社交关系的兴趣点推荐方法，即基于朋友的协同过滤模型（FCF）。该方法将具有相同兴趣点签到偏好的用户视为朋友，在进行推荐时只考虑朋友的喜好。在这里，由于 FCF 中的朋友定义仅基于签到数据，因此它很难揭示用户之间真正的社会关系。作为补充，另外三种类型的朋友关系被定义，即社交好友、位置好友和邻居好友。另一方面，一些其他信息包括时间信息^{[55][56]}、范畴信息^[44]、图像信息^[88]和文本信息^[113]也被用作补充，这是因为它们可以揭示兴趣点的语义。例如，时态信息表示兴趣点的活跃时间、兴趣点的分类信息在推荐时提供优先权、文本信息提供用户对兴趣点的评论和描述。

然而，由于计算用户-用户及兴趣点-兴趣点之间的相似性，需要基于共同的签到记录，因此经常会遇到数据稀疏问题。同时，由于签到数据的稀疏性，导致矩阵分解的估计误差较大，使得这两类方法在推荐任务上效果不佳。另一方面，

基于协同过滤的方法，需要依靠历史签到记录识别用户特别感兴趣的项目，因此会存在冷启动问题。

2.2.1.2 基于深度学习的推荐方法

真实场景中，签到数据往往包含多种类型的数据，比如文本，图片等，具有丰富的语义信息。基于 MF/CF 的推荐方法，难以建模 POI 之间非线性关系，对此，一种有效的解决方案是利用深度学习在建模复杂非线性关系方面的优势，挖掘用户与用户之间以及用户与 POI 之间的相关性，然后进行兴趣点推荐^[25]。

用户兴趣点矩阵的极端稀疏性和冷启动问题严重阻碍了基于 MF/CF 的推荐方法的性能。除此之外，由于地理区域的不同带来的城市构成以及文化的差异，用户偏好也会有很大差异。为了应对这些挑战，Yin 等^[102]立足于深度学习的最新进展，提出了一种空间感知的分层协作深度学习模型 (SH-CDL)。该模型对具有异构特征的 POI 进行深度表示学习，并且后期引入了特征融合策略来处理 POI 的多模态异构特征。Yang 等^[94]提出了一种通用的半监督学习框架 PACE (Preference And Context Embedding)，通过在相邻用户和 POI 之间进行平滑来缓解数据稀缺性，并借助于交互图中的上下文来规范用户偏好表征向量。最后，深度框架 PACE 通过学习得到的用户和 POI 的表征向量来预测用户对 POI 的偏好。除此之外，Ding 和 Chen^[25]提出了一种新的基于深度神经网络的 POI 推荐框架，RecNet。该框架借助于深度神经网络充分挖掘 LBSN 中的各类特征中蕴含的丰富信息并探索它们对用户行为的共同影响机制。RecNet 利用 LBSN 中的共同访问、地理和 POI 标签类别影响来缓解 POI 推荐中的数据稀疏问题，并通过特征表征方法从各类特征中学习得到 POI 和用户的表征向量。最后，POI 和用户表征向量被成对地输入深度神经网络框架中，以自适应地学习特征之间的高阶关联性。

然而，上述方法均面向传统的兴趣点推荐方法，不能够建模用户签到行为的序列化模式，忽略了用户签到行为的上下文依赖，不适用于下一个兴趣点推荐场景。

2.2.2 下一个兴趣点推荐方法

针对本文重点研究的下一个兴趣点推荐方法，本节根据不同的应用场景将下一个兴趣点推荐划分为社交关系感知的下一个兴趣点推荐、时间敏感的下一个兴趣点推荐以及数据匮乏环境下的下一个兴趣点推荐，然后详细分析与总结了不同场景下的研究成果。

2.2.2.1 社交关系感知的下一个兴趣点推荐方法

LBSNs 中具有社交关系的用户之间在签到行为方面会表现出一定的相似性，即朋友喜欢的兴趣点对用户来说，可能是一个不错的选择。并且用户之间的社交关系直观地反映了用户间的相互作用强弱，从而反映了用户朋友社交影响力强弱。因此，有效地利用社交关系信息不仅可以提高推荐算法的性能还可以在一定程度上缓解兴趣点推荐面临的数据稀疏问题。

目前，绝大多数的研究人员通过用户之间的社交关系获得相似性，并将其引入到传统基于模型或者基于内容的协同过滤模型^{[34][35][106][109][110]}，即借助于过滤模型直接集成用户社交关系的相关性。gSCorr^[34]通过整合社交关系和空间地理位置距离两类因素来进行兴趣点推荐，社交关系的引入在一定程度上缓解了兴趣点推荐的冷启动问题。iGSLR^[109]提出了一种整合用户偏好、社交影响以及个性化的地理位置影响等因素的兴趣点推荐框架，在集成的过程中，iGSLR 通过引入用户之间的社交相似性这一标量值完成兴趣点的过滤任务，计算相似度值的过程中共同好友的数量成为衡量相似性的重要指标。另有一部分研究工作通过将用户相似性作为矩阵分解或者张量分解的潜在因子模型的正则项或者权重^{[20][38][97]}，Cheng 等^[20]的工作首次将矩阵分解（Matrix Factorization，MF）与地理和社会关系影响融合到 POI 推荐模型中，首先将用户在某个位置签到的概率建模为多中心高斯模型（Multi-center Gaussian Model，MGM）来捕捉地理位置影响，然后将社交信息通过正则化项的方式引入到一个广义的矩阵分解框架中来建模用户受社交关系的影响程度。Yang 等^[97]提出了基于位置的社交因子分解模型 LBSMF（Location Based Social Matrix Factorization），直接将用户社交影响因素和地理影响因素建模为推荐任务的两个因子构建图模型，学习用户的概率分布来实现对用户的兴趣点推荐。Guo 等^[38]提出了基于信任的因子模型，TrustSVD，通过引入融合评论信息的社交信任因子来建模社交关系对兴趣点推荐的影响。除此之外，部分研究工

作通过构建用户-POI 交互异构图，通过随机游走的方式探索用户与 POI 之间的关联度^{[46][105]}。Ying 等^[105]通过在社交相似性、距离相似性和偏好吻合性三种因子指导下的随机游走命中率来进行兴趣点推荐。Kefalas 等^[46]提出了一种异构时空图中的随机游走策略，RWR-HST，该策略可以充分利用用户-用户之间的社交相似性和用户-位置的相关性进行兴趣点推荐。

近些年，受到 word2vec^{[59][60]}的启发，图表征学习技术发展迅猛。图表征学习（Graph representation learning）是一种将图中的节点表示为低维稠密的向量表示方法，得到的向量在表征空间中具有语义信息以及推理的能力^{[1][5][7][8][10][12]}。图表征学习作为近些年迅速发展的进行网络结构探索的技术，受到了科研界的极大关注，也极大地促进了兴趣点推荐领域的技术革新。图表征学习在发展初期，在进行表征的过程中，通常是将图构建成同构图，即图中仅存在一种类型节点和一类节点边关系^{[13][23][36][65][79][85][86]}。然而真实世界中，一类事物往往与不同种类的事物存在不同类别的交互关系，例如，人会与不同的人成为朋友，人也会拜访打卡不同的地点。这种具有不同类型的节点和关系的图被称为异构图（Heterogeneous information network），具有更加丰富的节点和边关系以及更加丰富的语义信息。为了进一步探索节点之间的深层次语义关系，研究者进一步提出了一系列面向异构图的图表征学习方法^{[11][16][26][32][72][78][87]}。

尽管图表征学习已经取得了较大的发展，然而在兴趣点推荐领域，图表征学习的应用还处于起步阶段^[71]，现存的方法通过处理同构图的方式探索用户与 POI 之间的语义信息，难以捕获复杂的节点关系。并且在进行结构探索的过程中没有考虑社交关系给兴趣点推荐带来的影响，除此之外，考虑用户签到行为的时序关系也起着至关重要的作用。因此，以上方法不能建模用户之间多维度的兴趣点偏好关联，忽略了用户的社交关系对于推荐结果的影响。

2.2.2.2 时间敏感的下一个兴趣点推荐方法

相较于传统的兴趣点推荐模型，时间敏感的下一个兴趣点推荐模型更具挑战性，这是因为需要预测下一个时间戳用户可能访问的兴趣点。为了解决时序推荐问题，马尔科夫模型^[20]被用来进行时序建模，多种上下文信息被用来学习兴趣点之间的转换模式。考虑到不同用户具有不同的移动模式，FPMC^[68]为每一个用户

构建一个转移矩阵，在考虑社交因素的基础上进一步估计了矩阵中的缺失参数。因此，FPMC 在一定程度上缓解了数据稀疏问题。在 FPMC 的基础上，FPMC-LC 进一步利用地理位置信息有效提升了兴趣点推荐效果。Feng 等^[29]则提出了通过度量表征来建模用户转移模式的推荐方法。点过程（Point process）^[58]也被用来进行下一个兴趣点推荐，该分支主要通过学习一个用户-兴趣点签到密度函数来估计用户下一时刻访问兴趣点的兴趣分布，但是该类方法时间开销成为限制其应用的主要瓶颈。上述方法只建模了用户一阶转移模式，缺乏探索更高阶的隐含信息。为了建模高阶时序模式，深度学习被引入进一步探索下一个兴趣点推荐新范式。现存的基于深度学习序列推荐模型大致可以分为两类：递归模型^[98]和卷积模型^[88]。（1）递归模型（RNNs）在建模时序兴趣点方面备受欢迎^[53]，关键在于其可以通过递归反馈来捕捉蕴含在兴趣点签到记录间的长期依赖，并且上下文信息可以很好地集成到递归框架中，例如，SERM^[98]借助于递归模型进行兴趣点时序关系建模，并且通过合并多个上下文信息作为递归模型的输入。集成上下文信息的另一种方法是考虑每层上下文中的信息，为了捕捉时间间隔和过渡距离的影响，ST-RNN^[55]替换了原有的递归模型中的单个过渡矩阵，通过时间矩阵和距离矩阵进行时间约束和空间约束关系的捕捉。在这里，ST-RNN 中的时间窗口是预定义的参数，因此这里需要事先假设用户具有相似的签到频率。（2）卷积网络（CNN）^[40]被广泛的用来提取图片以及文本中的上下文信息，并且可以用来捕捉隐式交互。Wang 等^[88]提出了基于广度学习的上下文感知的兴趣点推荐模型，DCPR，包含三个协同层，其中一个 CNN 层用于捕捉兴趣点的特征信息，一个 RNN 层由于捕捉时序依赖和用户偏好建模，最后通过一个基于矩阵分解的交互层来联合优化推荐模型。Tang 等^[81]提出了一种名为 Caser 的基于 CNN 的模型来学习时序模式。Caser 将一系列最近的签到项目嵌入到“图像”中，然后使用卷积滤波器按照建模局部特征的方式来捕捉时序模式。此外，除了深度模型外，还有其他一些方法可以对高阶过渡模式进行建模，例如，Fossil 用高阶马尔可夫链扩展了 FPMC 和基于成对排序的模型，TransRec 来建模用户、用户先前访问的商品以及下一个商品之间的交互关系。

可见当前主要是将时间划分成时间槽的方式进行时间维度的信息感知。但

是，由于我们采集的 POI 数据的时间点并不是严格规整的，即并不是按照固定的时间间隔进行采集，这就造成 POI 数据之间时间间隔的不确定性。传统的序列化建模方法，将时间间隔分割成不同的时间槽或者直接按照先后顺序进行建模，忽略了 POI 序列之间的连续时间信息，因此序列推荐效果受到严重制约。近期，Chen et al. 等人在其工作 Neural-ODE^[19]中指出了在序列化推荐任务中，连续时间建模对推荐效果的性能有明显的提升，并提出了可以进行连续时间建模的新的深度网络——神经常微分方程。

神经常微分方程（Neural Ordinary Differential Equation）使用神经网络参数化隐藏状态的导数，而不是如往常那样直接参数化隐藏状态。具体地，神经常微分方程在模型训练的过程中，不使用其他 RNN 以及 CNN 等神经网络直接参数化隐藏状态的处理方式，直接使用神经网络参数化隐藏状态的导数。这里参数化隐藏状态的导数就类似构建了连续性的层级与参数，而不再是离散的层级。因此参数也是一个连续的空间，不需要再分层传播梯度与更新参数，这样可以更加高效地进行连续时间建模。目前针对神经常微分方程的工作还相对较少，主要有面向时序序列建模的 Neural-ODE^{[19][70]}， GRU-ODE^[24]等。Neural-ODE 首次提出神经常微分方程(Neural-ODE)，是一种新的深度神经网络，借助于常微分方程，实现对序列数据的连续建模，即 Neural-ODE 不需要在反向传播时存储所有参数和渐变，并且可以通过连续定义的动态来自然地合并在任何时间到达的数据。紧接着，Rubanova 等^[70]提出了基于常微分方程(ODE)的连续时间隐状态的泛化 RNN 模型，ODE-RNN。ODE-RNN 和潜在 ODE (latent ODE) 都可以自然地处理任意时间间隔的观测数据，并且可以使用泊松过程明确地对观测时间的概率进行建模。GRU-ODE^[24]则是结合 Neural-ODE 和门控循环单元 (GRU) 进行时间序列的连续性建模，根据输入离散的时间点观察值，通过 GRU-bayes 更新 Neural-ODE 之间的隐状态向量，达到连续时间状态感知的目的。

由于基于神经常微分方程尚处于萌芽阶段，虽然在时序预测领域取得了不错的研究成果，但是面向兴趣点推荐任务中的高稀疏度语义丰富的时序数据，还没有成型的解决方案，仍需要继续探索如何在兴趣点推荐任务中定义连续的动态过程，使模型能够接受任意时刻的输入数据，并适用于建模连续签到行为序列。

2.2.2.3 数据匮乏环境下的下一个兴趣点推荐方法

训练数据匮乏极大地限制了模型的训练收敛过程，导致模型学习到的数据知识相对有限，限制了模型对用户的兴趣点推荐能力。因此，一部分研究人员提出通过基于知识迁移的兴趣点推荐方法来缓解在训练数据匮乏区域的冷启动问题。该类工作主要包括两类方法：一，直接将数据丰富城市的兴趣点推荐模型迁移到数据匮乏城市进行应用^{[55][56][112][114]}。Liu 等^[55]在循环神经网络的基础上增加了时间敏感性，捕捉在签到行为之间的时间和空间位置信息变化。Zhao 等^[112]通过在长短时间记忆网络（LSTM）中增加了时间和距离两种门限机制来捕捉用户签到行为序列之间的时空关系来进行兴趣点推荐；二，利用多城市之间的共同用户作为锚链接，进行跨城市兴趣点推荐^{[49][103][104]}。Yin 等^[104]提出了一种联合概率模型，来模拟用户在进行签到行为的决策过程，通过集成多种因素可以在一定程度上缓解数据稀疏的问题。Li 等^[49]均假设目标城市和源城市具有共同主题兴趣点，借助于主题模型根据用户在源城市中的目标偏好，进行目标城市中的兴趣点推荐。Yin 等^[103]在联合概率模型中引入对用户的偏好存在城市间偏移这一特性的考量，进而更好地进行兴趣点推荐。

然而，现有下一个兴趣点推荐方法没有考虑兴趣点数据训练数据匮乏这一特性，因此仅适用于进行训练数据丰富环境下的兴趣点推荐。虽然现存一部分工作是进行跨城市兴趣点建模的工作，但同样是建立在存在共同用户的基础上，即需要通过不同城市之间的共同用户建立锚链接，进而挖掘用户在新城市的兴趣点偏好。这类方法受到共同用户这一假设的制约，不能很好地将模型迁移到一个新的不存在共同用户的城市。并且由于新城市中数据匮乏，导致模型训练困难，难以收敛，限制了推荐模型的能力。

元学习（Mate-learning）允许机器通过一些训练示例快速学习新技能或适应新环境，在监督学习（例如小样本图像分类和强化学习场景中）中都取得了成功，是实现知识迁移的有效解决方法。元学习按照迁移方式的不同可以分为四种常见的方法^[100]：（1）使用配备外部或内部记忆存储和查询元知识的循环神经网络^{[61][63]}，比如 Mishra 等^[61]提出了一类简单且通用的元学习器架构，该架构使用了时序卷积和软注意力的新颖组合。前者从过去的经验中收集信息，而后者则用于

确定特定的信息，通过元学习器实现知识迁移；(2) 学习一个可以快速优化模型参数的元优化器^{[50][67]}，例如 Ravi 等^[67]从强化学习的角度处理这个问题，并将任何特定的优化算法表示为策略。本文使用引导策略搜索学习优化算法，并证明所得算法在收敛速度和/或最终目标值方面优于现有的手工设计算法；(3) 学习示例之间的有效距离度量^{[73][77]}，比如 Snell 等^[73]通过神经网络学会一个“好的”映射，将各个样本投影到同一空间中，对于每种类型的样本提取他们的中心点作为原型（prototype）。使用欧几里得距离作为距离度量，训练使得测试样本到自己类别原型的距离越近越好，到其他类别原型的距离越远越好。测试时，通过对到每类原型的距离做 softmax 获得测试样本类别；(4) 学习适当的初始化参数，从中可以在几个梯度步骤内更新模型参数^{[17][31][48]}，其中，Finn 等^[31]提出的模型无关的元学习算法（Model-agnostic meta-learning, MAML）最为经典。和其他元学习模型类似，MAML 同样也通过广泛的任务来进行训练，它可以通过几步梯度变换就可以迅速地适应新的任务。在这一过程中，元学习器会去寻找不仅有利于适应不同的任务，更有利于快速（很少的梯度下降步数）和高效（仅使用少量的样本）地适应特定任务。

尽管元学习方法在监督学习任务中取得了很好的效果，能很好地解决在数据量少的情况下模型训练的问题。但是将元学习应用于时空数据挖掘的工作仍处于探索阶段，Yao 等^[99]提出了将元学习应用于时空预测问题，该方法通过利用来自其他城市的丰富长期数据学习的相关知识进行迁移，解决了数据匮乏城市的时空预测问题。特别地，该文提出的模型被设计为具有元学习范式的时空网络。元学习范式学习时空网络的广义初始化，可以有效地适应目标城市。由于兴趣点推荐任务与时空数据预测不同的是兴趣点签到记录更加稀疏，时空语义更加丰富，因此当前的基于元学习的时空数据预测范式无法应用到训练数据匮乏的兴趣点推荐场景中。

2.3 小结

本章对位置社交网络中兴趣点推荐的研究工作进行了总结：首先，概述了下一个兴趣点的定义及公开数据集的简况；其次，总结了兴趣点推荐技术的相关研

究工作，并重点阐述了现有工作在社交关系感知的下一个兴趣点推荐、时间敏感的下一个兴趣点推荐以及数据匮乏环境下的下一个兴趣点推荐等任务场景下取得的成果和存在的不足。

第3章 多维度社交关系感知的下一个兴趣点推荐方法研究

本章针对用户社交关系建模不准确的问题，提出了一种多维度社交关系感知的下一个兴趣点推荐方法 GraphSAR（**G**raph neural network based **S**ocial-Aware **R**ecommendation）。该方法旨在采用维度级的注意力机制来学习用户表征之间的非结构化关联，使得模型可以更好地学习社交关系对用户偏好的影响，提升下一个兴趣点推荐效果。本文设计了一种基于注意力机制的兴趣偏好通道网络来建模用户之间的社交关系影响，并在影响注意力的引导下获取用户偏好表征。此外，为了有效利用用户-POI 异构交互网络中的丰富语义信息，本文设计了一种基于元路径的语义表征网络，利用图卷积网络对语义网络中具有丰富隐含语义信息的 POI 进行编码，以保证可以充分利用历史签到信息来提取节点间的关系特征。实验使用两个真实的数据集验证了 GraphSAR 在 NPOI 推荐任务上的有效性。

本章的组织如下：3.1 节介绍了本章的研究问题和挑战。3.2 节详细介绍了基于多维度社交关系感知的下一个兴趣点推荐方法，其中，3.2.1 节介绍了下一个兴趣点推荐问题定义及方法概述；3.2.2 节介绍了基于图卷积的 POI 表征网络；3.2.3 节重点介绍了基于多维度社交关系注意力机制的用户偏好表征网络；3.2.4 节介绍了 GraphSAR 的参数优化算法；3.2.5 节对 GraphSAR 进行了复杂度分析。3.3 节通过真实数据集上的实验验证了 GraphSAR 的有效性。3.4 节对本章进行了小结。

3.1 引言

在位置社交网络中，用户可以自主地与其他用户建立社交关系。这些社交关系在用户对兴趣点的反馈之外，又为兴趣点推荐方法提供了一种额外的辅助信息^[96]。从社交关联理论^[80]中诸如同质性和社交影响等理论来看，具备社交关系的两个用户往往存在一定的关联^[91]，这种关联可以帮助我们提升推荐系统的推荐性能^{[34][45]}。更直观地，如图 3.1 所示，用户在面临餐馆选择时，当用户了解到某位朋友最近也打卡过这家餐厅，并且评价不错，那么这家餐馆很可能成为用户的优先选择。总而言之，社交关系提供了用户签到数据之外的另外一种描述用户关系的

独立数据源，在推荐系统中起到了重要的作用，可以用来帮助提升兴趣点推荐方法的性能^[96]。



图 3.1 社交关系对用户选择起到关键作用

Figure 3.1 Illustration of social relationships playing a key role in user selection

近些年，为了建模社交关系给下一个兴趣点推荐方法带来的影响，研究人员已经提出了很多方法。我们可以将这些方法分成三类：(1) 基于记忆的方法，即协同过滤模型^{[34][35][106][109][110]}，借助于过滤模型直接集成用户社交关系的相关性；(2) 基于模型的方法，这部分研究工作通过将用户相似性作为正则项或者权重^{[20][38][97]}引入到矩阵分解或者张量分解的潜在因子模型中，达到将社会关系影响融合到 POI 推荐模型中的目的；(3) 基于图结构的社交关系感知方法，该类研究工作通过构建用户-POI 交互异构图^{[46][105]}，通过随机游走的方式探索用户与 POI 之间的关联度。总的来说，现有方法通过学习一个标量值来进行社交关系的引入，限制了签到数据中丰富信息的表达。

然而，每个用户的兴趣都是分布在多个方面的。对两个用户而言，他们在某个方面的一致性，并不意味着他们在其他方面的一致性。为了证明这一点，我们从 Foursquare 数据集中选择了一个用户作为代表，该用户共有 129 位朋友，我们计算了这个用户和他的每一位朋友在 6 个不同类别上的兴趣点签到相似度，并将每个类别上最大的相似度扩展到 1，其他的等比例缩小，得到了如图 3.2 所示的热度图。从图中可以看出，该用户和他的朋友在 6 个不同类别上的相似度差异很大。当这个用户需要在某个类别上的建议时，他会去参考在该类别上相似度较高的朋友，在其他类别上相似度高的朋友的建议对用户而言并不十分重要。从上述分析来看，在现有方法中使用特定的标量值进行社交关系的引入不能够表达不同的用户多个方面的兴趣关联。然而，要建立多个方面的兴趣关联并不是一件易事，

因为没有任何显式的证据能够直接将用户之间的社交关系直接转化为他们在不用类别上的兴趣点偏好关系。并且在下一个兴趣点推荐任务中，我们还要考虑用户时序关系的建模，给社交关系的引入带来了更大的挑战。

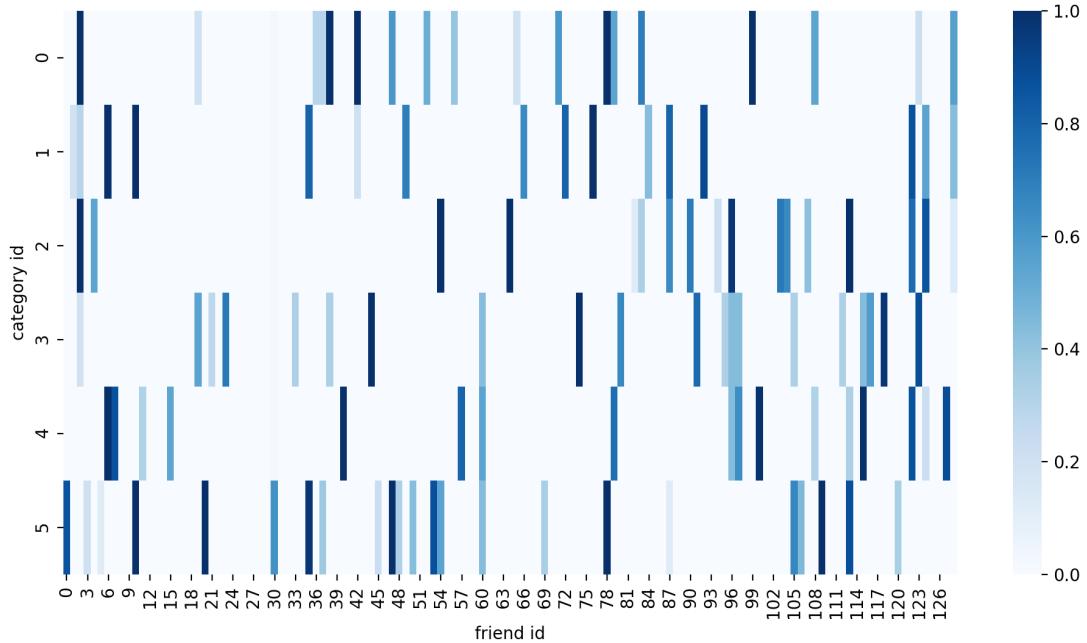


图 3.2 某用户与其朋友在不同类别兴趣点上的签到相似度热度图

Figure 3.2 Heatmap of check-ins similarity between a user and his friends in different categories of POIs in Foursquare

为了解决以上问题，本文提出了多维度社交关系感知的下一个兴趣点推荐方法 GraphSAR。GraphSAR 采用维度级的注意力机制来学习用户表征之间的非结构化关联，使得模型可以更好地学习社交关系对用户偏好的影响，提升下一个兴趣点推荐效果。该方法首先通过元路径对语义网络中具有丰富隐含语义信息的 POI 进行语义编码，以保证可以充分利用历史签到信息来提取节点间的关系特征。然后，GraphSAR 通过基于社交注意力机制的兴趣偏好通道网络进行用户偏好的建模，准确建模用户的社交关系对偏好的影响。因此，GraphSAR 的训练过程可以划分为 POI 表征和社交关系感知的用户偏好表征两个阶段。

在 POI 表征阶段，GraphSAR 设计了适用于异构图结构的图卷积单元，实现对 POI 各类特征的表征聚合，并生成低维度的表征向量。具体地，本文设计了一种基于图卷积的 POI 表征网络，首先借助于元路径进行 POI 关系抽取，然后基于图卷积单元进行特征的传播与聚合过程，生成 POI 表征向量。

在用户偏好表征阶段，GraphSAR 设计了一种基于多维度社交注意力机制的用户兴趣感知方法，从用户的历史签到记录中抽取用户的兴趣偏好表征向量。首先通过多个推理层从用户的历史签到记录中进行特征向量的抽取，然后通过引入多维度社交关系因素的注意力层（Vanilla attention layer）来指导偏好特征向量的优化。

最后，利用上述步骤获得的 POI 表征向量和用户偏好表征向量，GraphSAR 进一步获取特定用户对不同 POI 的偏好分数，根据设定的条件进行筛选获得下一个兴趣点推荐结果。

3.2 多维度社交关系感知的下一个兴趣点推荐方法

本节详细介绍了多维度社交关系感知的下一个兴趣点推荐方法 GraphSAR。首先定义了社交关系感知的下一个兴趣点推荐问题并概述了 GraphSAR 的主要步骤，然后详细阐述了 GraphSAR 的架构设计和算法流程，最后介绍了 GraphSAR 的参数优化方法。

3.2.1 问题定义及方法概述

本节首先定义了下一个兴趣点推荐问题，然后介绍了本文提出的多维度社交关系感知的下一个兴趣点推荐方法 GraphSAR。

3.2.1.1 问题定义

对于 M 个用户 $U = \{u_1, \dots, u_M\}$ 、签到序列集合 \mathcal{T} 、社交关系集合 $E \subseteq U \times U$ 、N 个兴趣点 $P = \{p_1, \dots, p_N\}$ 和 F 个位置语义标签 $C = \{c_1, \dots, c_F\}$ ，每个目标用户 u 都对应一条历史签到记录序列 $\mathcal{T}_u = [r_1, \dots, r_K]$ ，其中每条签到记录 $r_k = (l_k, c_k, t_k)$ 是一个三元组，包含了地理位置 l_k 、兴趣点的标签类别 c_k 以及时间戳 t_k 。根据上述描述，给出社交关系感知的下一个兴趣点推荐的定义：

社交关系感知的下一个兴趣点（NPOI）推荐：给定用户历史签到序列集合 \mathcal{T} 以及社交关系集合 E ，则对于目标 u ，存在一条到时间 t_K 的用户历史签到序列 \mathcal{T}_u ，我们的目标是根据用户的历史签到序列和社交关系为其推荐在下一个时间点 \hat{t} 可能访问的兴趣点。

3.2.1.2 方法概述

如图 3.1 所示，社交关系感知的 NPOI 推荐的目的是估计目标 u 在接下来的

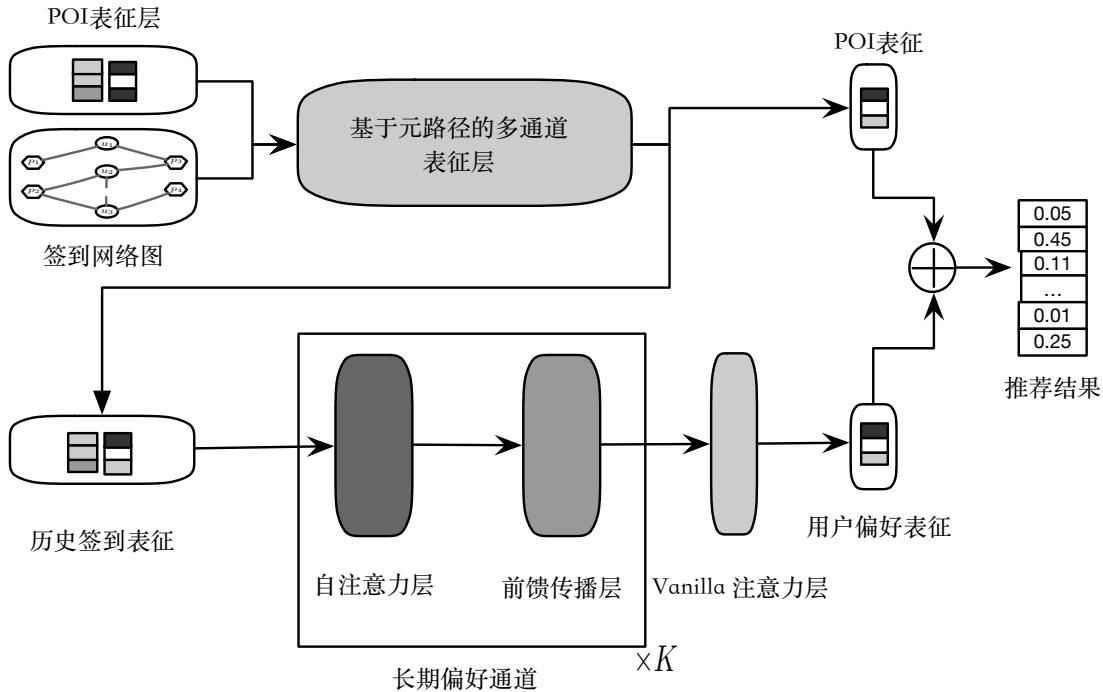


图3.3 GraphSAR架构示意图

Figure 3.3 Architecture of GraphSAR

时间可能感兴趣的兴趣点。利用 GraphSAR 进行兴趣点推荐需要解决三个子问题：第一，如何合理地从大量的历史签到数据中建模 POI 的隐语义特征，获取 POI 表征向量；第二，如何基于社交关系建模用户之间多维度偏好差异，获取用户偏好表征向量；第三，如何利用抽取出的特征向量进行 NPOI 推荐。

为了解决上述三个问题，本文提出了多维度社交关系感知的 NPOI 推荐方法 GraphSAR。如图 3.3 所示，该方法分为三个主要部分：

- 首先，为了建模 POI 的时空特性，本文提出了基于图卷积神经网络的 POI 表征网络。具体地，利用基于差分自编码机的时间编码来获取时间表征向量，并通过 POI 所在经纬度信息的哈希编码来获取空间地理信息表征向量；然后利用图卷积网络来挖掘 POI 之间的关联关系，获取 POI 的表征向量；

- 然后，为了建模用户多维度偏好差异，本文提出了基于注意力机制的兴趣偏好通道网络。需要注意的是，该网络在进行偏好建模的过程中，我们通过多个注意力层和反馈层建模用户的时序特征，通过社交关系注意力层来建模特定用户受社交朋友的影响，获取用户的兴趣偏好表征向量；
- 最后，根据提取的 POI 表征向量和用户偏好表征向量，进行 NPOI 推荐。

3.2.2 基于图卷积神经网络的 POI 隐语义表征网络

在 LBSNs 中，可以通过一张异质信息图来刻画用户的签到行为。在异质图中，除了节点（用户和兴趣点）间通过显式链接表现出来的浅层关系外，通常存在着其他更深层的隐藏关系。利用这些隐藏关系可以更好地进行 POI 的表征，获取有效的 POI 表征向量。由于进行 POI 隐语义关系探索需要估计来自不同用户的影响，因此需要赋予不同的用户-POI 交互边不同的权重，而传统图卷积神经网络不能建模这种变化。为了解决这一问题，本文设计了一种基于自注意力机制的图卷积神经网络用于实现不同类型节点特征的聚合和处理，获取 POI 表征向量。以图卷积神经网络为基础的 POI 表征网络分为三个部分：用户-POI 交互网络图的构建、POI 语义关系表征编码和 POI 时空特征编码。

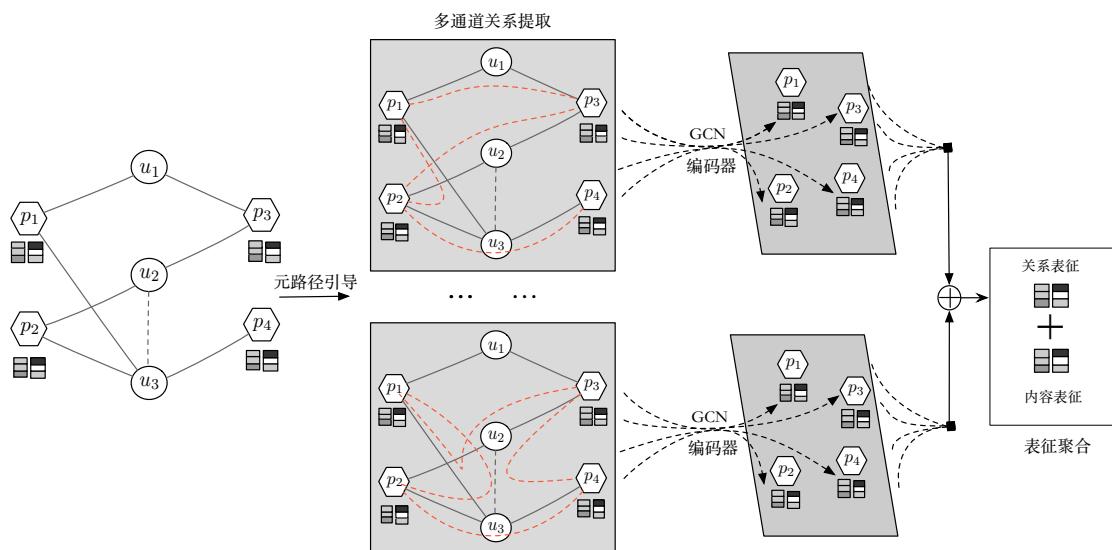


图 3.4 基于元路径的语义表征示意图

Figure 3.4 Architecture of meta-path based embedding model for semantics of POI

3.2.2.1 构建用户-POI 交互网络

获取用户的历史签到记录之后，为进一步探索用户与签到位置之间的复杂语

义关系，本文首先构建用户-POI 交互网络，如图 3.5 所示。用户-POI 网络中每条边的一侧节点为用户 u ，另一侧节点为用户访问的位置 p ，两侧节点之间边的链接关系由公式（3.1）决定，若用户 u_i 访问过位置 p_j ，则 $e(u_i, p_j) = 1$ ，反之， $e(u_i, p_j) = 0$ ：

$$e(u_*, p_*) = \begin{cases} 1, & \text{若 } u_* \text{ 访问过 } p_* \\ 0, & \text{若 } u_* \text{ 未访问过 } p_* \end{cases} \quad (3.1)$$

用户-POI 访问网络图反映了用户与 POI 之间的访问关系与拓扑结构。同时根据用户访问频次的高低赋予不同的边， $e(u_*, p_*)$ ，不同的权重值 w ， w 的计算方式定义为：

$$w_{e(u_*, p_*)} = \frac{f(u_*, p_*)}{\sum_1^n f(u_*, p_i)} \quad (3.2)$$

其中， n 表示用户 u_* 访问过的 POI 的数量， $f(u_*, p_*)$ 表示用户 u_* 访问 POI p_* 的次数。因此，访问网络图 \mathcal{N} 可以形式化为节点数为 $|V|$ ($|V| = N + M$) 的维度为 d 维的特征向量矩阵 $V = \{v_1, v_2, \dots, v_{|V|}, v_1 \in \mathbb{R}^d\}$ ，其元路径引导下的邻接矩阵用 A 表示，其度矩阵可以表示为 D 。

这里元路径（Meta-path）是指连接两个实体的一条特定的路径^{[75][76]}，可以定义为 $\mathcal{P}_i: \mathcal{A} \xrightarrow{r_1} \mathcal{A} \dots \xrightarrow{r_1^{-1}} \mathcal{A}$ ，其中 \mathcal{A} 表示签到网络中一种类型实体的节点，而 r_i 表示节点与节点间的关联关系。换句话说，每条元路径定义了节点之间的唯一语义关系，如图 3.5，展示了 2 条元路径分别表示同一用户访问的不同 POI (PUP) 和朋友之间访问的不同的 POI (PUUP)。

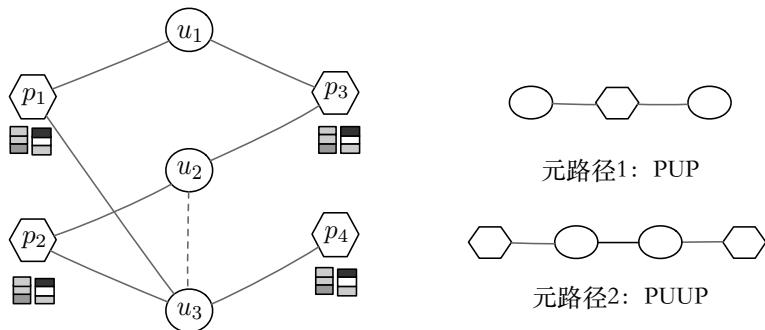


图 3.5 基于元路径的多通道用户-POI 语义关系抽取示例

Figure 3.5 Meta-path based multi-channel semantics extractions

3.2.2.2 POI 关系特征编码

传统的 GCN 网络中，节点受邻居节点的影响是相同的，不做区分。但是实际上，不同邻居节点的影响程度不尽相同，为了表现不同节点之间影响力差异性，这里我们借鉴 Vaswani 等^[83]提出的自注意力机制，对不同邻居节点的影响程度进行区分，可以形式化为：

$$e(v_i, v_j) = \alpha(W_1 v_i, W_1 v_j) \quad (3.3)$$

其中， $e(v_i, v_j)$ 表示节点 v_i 受节点 v_j 的影响程度， $W_1 \in \mathbb{R}^{G \times d}$ 是一个权重矩阵， α 表示映射函数。为了控制卷积神经网络中的注意力分布范围，GraphSAR 将注意力集中在节点的邻居汇总，即 $v_j \in \text{neigh}(v_i)$ ，则节点 v_j 对节点 v_i 的注意力系数为：

$$\alpha(v_i, v_j) = \text{softmax}(e(v_i, v_j)) = \frac{\exp(e(v_i, v_j))}{\sum_{v_k \in \text{neigh}(v_i)} \exp(e(v_i, v_k))} \quad (3.4)$$

这里， $\text{neigh}(v_i)$ 表示节点 v_i 的邻居，这里，注意力 α 通过单层前馈神经网络实现，并且应用非线性函数 *LeakyRelu* 作为激活函数，完全展开后，注意力相关系数 $\alpha(v_i, v_j)$ 可以通过下式进行计算：

$$\alpha(v_i, v_j) = \frac{\exp(\text{LeakyRelu}(\bar{a}^T [W_1 v_i || W_1 v_j])))}{\sum_{v_k \in \text{neigh}(v_i)} \exp(\text{LeakyRelu}(\bar{a}^T [W_1 v_i || W_1 v_k])))} \quad (3.5)$$

其中， $*^T$ 表示矩阵转置， $||$ 表示向量拼接操作， α 表示维度为 $2F$ 的映射函数，即 $\bar{a} \in \mathbb{R}^{2F}$ 。通过对自注意力系数的引入之后，采用非线性激活函数 φ ，经过自注意力蒸馏之后每个节点的输出可以形式化为：

$$h_{1i} = \varphi(\sum_{v_k \in \text{neigh}(v_i)} \alpha(v_i, v_k) W_1 v_k) \quad (3.6)$$

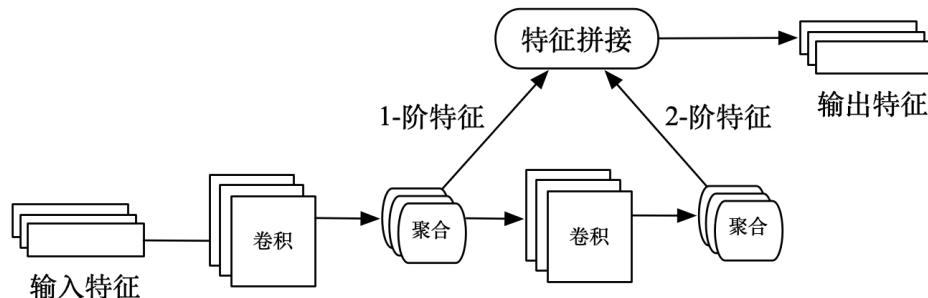


图 3.6 语义归纳表征层

Figure 3.6 Inception layers of Semantics

为了进一步挖掘用户访问网络中的复杂语义关系, GraphSAR 构建了两层图卷积神经网络进行语义信息挖掘, 框架结构如上图所示。一阶用户关系仅可以捕捉用户与其访问的 POI 之间的关系, 而二阶实体关系捕捉则可以扩展到节点与二阶邻居节点之间的关系挖掘, 比如, 两个用户访问同一个 POI 则表示两个用户可能有相似的偏好。根据 GraphSAR 的传播层公式, 第一层基于自注意力机制的图卷积神经网络的输出如下:

$$\mathbf{h}_2 = \varphi(\hat{\mathbf{A}} \mathbf{h}_1 \mathbf{W}_2) \quad (3.7)$$

其中, \mathbf{h}_1 为自注意力层的输出, $\hat{\mathbf{A}} = \tilde{\mathbf{D}}^{-\frac{1}{2}} \tilde{\mathbf{A}} \tilde{\mathbf{D}}^{-\frac{1}{2}}$, $\tilde{\mathbf{A}} = \mathbf{A} + \mathbf{I}$, \mathbf{I} 是单位矩阵, $\tilde{\mathbf{D}}$ 是 $\tilde{\mathbf{A}}$ 的度矩阵 ($\tilde{\mathbf{D}} = \sum_j \tilde{A}_{ij}$)。为了捕捉用户与 POI 之间的二阶语义信息, 我们通过两层图卷积神经网络进行建模, 则二阶语义信息捕捉的输出可以进一步形式化定义为:

$$\ddot{\mathbf{h}}_2 = \varphi(\hat{\mathbf{A}} \varphi(\hat{\mathbf{A}} \mathbf{h}_1 \mathbf{W}_2) \mathbf{W}_3) \quad (3.8)$$

这里, 为了充分利用用户与 POI 的多阶语义信息, 将一阶语义和二阶语义编码进行拼接, 最终的关系语义表征输出为:

$$\mathbf{e}_r^* = [\varphi(\hat{\mathbf{A}} \mathbf{h}_1 \mathbf{W}_2) || \varphi(\hat{\mathbf{A}} \varphi(\hat{\mathbf{A}} \mathbf{h}_1 \mathbf{W}_2) \mathbf{W}_3)] \quad (3.9)$$

3.2.2.3 POI 时空特征编码

本小节介绍了 GraphSAR 中的 POI 时空特征编码部分, 分别详细介绍了 POI 空间语义编码和时间模式的编码过程。

对于目标 \mathbf{u} 的一条签到记录 (l_k, c_k, t_k) , GraphSAR 首先对其中的地理位置坐标进行编码, 针对 POI 地理位置空间属性, $l_k = (lat_k, lon_k)$, 引入矩阵 $\mathbf{E}_l \in \mathbb{R}^{M \times d_l}$, 其中, M 表示 POI 的个数, d_l 表示 POI 位置属性表征向量的维度, 在这里, 我们通过经纬度坐标位置的 GeoHash 编码来初始化 POI 地理位置编码矩阵 E_l , 则每一个 POI 的地理位置编码可以表示为 \mathbf{e}_l 。针对 POI 类别属性, c_k , 引入矩阵 $\mathbf{E}_c \in \mathbb{R}^{N \times d_c}$, 这里, N 表示 POI 的类别数, d_c 表示 POI 类别表征向量的维度, 则针对每一个 POI 的类别编码可以表示为 \mathbf{e}_c 。针对时间属性, 由访问趋势图可以看出, 可见不同的 POI 具有不同的用户-POI 访问时间模式。比如, 针对某 POI 在白天时间访问量比较少, 但是到了凌晨会被瞬间激活, 用户访问量激增, 而另一个 POI

正好相反，白天客流量比较大，到了晚上便沉寂下来。另一方面，经研究表明，用户通常活动规律为七天（即以周为单位），因此通过建立一个 7×24 的时间矩阵来进行时间特征抽取，为了进一步对有效时间特征进行蒸馏，GraphSAR 借助于变分自编码机^[47]进行时间特征向量的生成，框架如下图所示：

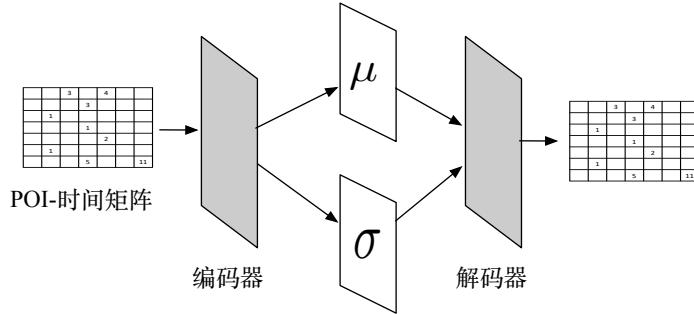


图 3.7 基于变分自编码机的时间特征编码

Figure 3.7 Temporal attributes encoder with VAE

根据输入的时间矩阵 X ，变分自编码机的目的在于输出潜在变量 \mathbf{z} ， \mathbf{z} 可以看成时间矩阵 \mathbf{X} 的特征向量，蕴含了时间特征的表征向量，此编码过程可以形式化为 $p(\mathbf{z}|\mathbf{x})$ 。编码器的损失函数可以表示为：

$$L(\mathbf{x}) = -\text{KL}(q(\mathbf{z}|\mathbf{x}) || p(\mathbf{z})) + \mathbf{E}_{\mathbf{z} \sim q(\mathbf{z}|\mathbf{x})} (\log p(\mathbf{x}|\mathbf{z})) \quad (3.10)$$

其中， $q(\mathbf{z}|\mathbf{x})$ 为识别模型，用于近似后验分布 $p(\mathbf{z}|\mathbf{x})$ ，这里，为了计算方便，GraphSAR 采用重参数化方法，即 $\mathbf{z} = \mu + \sigma \odot \epsilon$, $\epsilon \sim \mathcal{N}(0,1)$ ，让 $p(\mathbf{z})$ 服从一个先验的高斯分布 $\mathcal{N}(0,1)$ ，则 $q(\mathbf{z}|\mathbf{x})$ 其实就可以看成满足均值为 μ ，方差为 σ 的高斯分布，因此最终损失函数可以展开为：

$$L_t = \frac{1}{2} \left(\sum_{j=1}^J \left(1 + \log((\sigma_j)^2) - (\mu_j)^2 - (\sigma_j)^2 \right) \right) + \frac{1}{N} \sum_{i=1}^N \sqrt{(\mathbf{x}_i - \hat{\mathbf{x}}_i)^2} \quad (3.11)$$

其中， \mathbf{x}_i 是输入的时间矩阵， $\hat{\mathbf{x}}_i$ 是经过变分自编码机解码器重构的时间矩阵。经过网络的参数学习之后，提取变分自编码器的隐层输出得到相关 POI 的时间特征编码 e_t 。

3.2.2.4 POI 特征聚合

为了获取有效的 POI 表征向量，本文利用不同元路径进行双通道的关系特征抽取。为了更好地整合不同通道提取的关系特征向量，本文引入权重矩阵 W_r ，则整合关系表征的过程可以形式化：

$$\mathbf{e}_r = \mathbf{W}_r(\mathbf{e}_r^{channel1} || \mathbf{e}_r^{channel2}) \quad (3.12)$$

其中， $||$ 表示向量拼接操作， \mathbf{W}_r 设置为 $d * 2d$ 。为了进一步在POI表征中融入时空特征，本文直接通过向量拼接来获取POI的最终表征向量，可以形式化为：

$$\mathbf{e}_p = \mathbf{e}_r || \mathbf{e}_t || \mathbf{e}_c || \mathbf{e}_l \quad (3.13)$$

3.2.3 用户偏好表征向量提取

为了建模用户的兴趣偏好，本文提出了基于注意力机制的社交关系感知的用户偏好表征网络。该网络借助于注意力机制从用户历史签到记录中提取用户的兴趣偏好，并且考虑社交关系对用户兴趣偏好带来的影响，从而获得高质量的用户特征向量。因此，该注意力网络模型分为两个主要部分：兴趣偏好时序特征生成和社交关系感知校正。

3.2.3.1 时序特征生成

受 Vaswani^[83]人提出的类似的多头自注意力结构的启发，本文根据NPOI推荐场景进行注意力层（Self-attention layer）相关自定义设置。本文利用多头自注意力来计算所有用户签到记录之间的行为相关性。这里给定用户的签到轨迹 $T_u = [r_1, \dots, r_K]$ ，用 $\mathbf{H}_u = [\mathbf{h}_1, \dots, \mathbf{h}_K] \in \mathbb{R}^{d \times M}$ 表示经过 K 个非线性层（Nonlinear layer）生成的隐含特征向量， d 表示隐变量的维度， M 表示用户签到轨迹的长度。因为当前用户偏好可以通过用户轨迹中的签到记录进行体现，这里我们借助于多头注意力机制获取用户偏好表征向量，形式化为：

$$\hat{\mathbf{r}}_k = (head_1^k || head_2^k || \dots || head_H^k) \mathbf{W}^k \quad (3.14)$$

$$head_h^{k,i} = \sum_{j=N-M+1}^N \beta_{ij}^{k,h} \mathbf{r}_j^{k-1,h} \quad (3.15)$$

在公式中， $||$ 表示拼接操作， H 表示多头自注意机制中设置的头数量， $\mathbf{W}^k \in \mathbb{R}^{d \times d}$ 为参数矩阵， $\mathbf{r}_j^{k-1,h}$ 表示隐表征向量 \mathbf{r}_j^{k-1} 中的第 h 个头向量， $\beta_{ij}^{k,h}$ 表示注意力权重。然后，我们简要概述一下注意力权重 $\beta_{ij}^{k,h}$ 的计算过程，对于给定的面向 k -th层的隐特征向量，比如 $(\mathbf{r}_i^{k-1,h}, \mathbf{r}_j^{k-1,h})$ ，权重 $\beta_{ij}^{k,h}$ 可以通过下式进行计算：

$$\beta_{ij}^{k,h} = \frac{\exp(f(\mathbf{r}_i^{k-1,h}, \mathbf{r}_j^{k-1,h}))}{\sum_{j=N-M+1}^N \exp(f(\mathbf{r}_i^{k-1,h}, \mathbf{r}_j^{k-1,h}))} \quad (3.16)$$

$$f(\mathbf{r}_i^{k-1,h}, \mathbf{r}_j^{k-1,h}) = \frac{\mathbf{r}_i^{k-1,h} (\mathbf{r}_j^{k-1,h})^T}{\sqrt{d}} \quad (3.17)$$

He 等^[40]指出在视觉处理任务中表征层的深度可以有效提升表达效果，受其工作的启发，在 GraphSAR 中借鉴残差网络（residual learning）构建深度学习网络。在本文提出的兴趣偏好提取通道网络中，每一个自注意层都增加了一个残差连接。并且为了能够使网络更加稳定易收敛，GraphSAR 应用了归一化层进行输出规约，可以形式化为：

$$\mathbf{g}_i^k = \text{layer_norm}(\mathbf{r}_j^{k-1}, \hat{\mathbf{r}}_k) \quad (3.18)$$

经过归一化层之后，为了增强 GraphSAR 的表征能力，本文增加了一个全连接层（Fully connected layer）来处理自注意机制层的输出向量，每一个全连接层包含两个转换（Transformation）和一个整流单元（ReLU）。对于给定的输入 \mathbf{g}_i^k ，可以通过下式计算经过全连接层的输出：

$$\mathbf{r}_j^k = \text{ReLU}(\mathbf{g}_i^k \mathbf{W}_1^{k,f}) \mathbf{W}_2^{k,f} \quad (3.19)$$

这里，ReLU() 表示整流激活函数， $\mathbf{W}_1^{k,f}, \mathbf{W}_2^{k,f} \in \mathbb{R}^{d \times d}$ 表示参数矩阵。

经过 K 层非线性层之后，GraphSAR 得到签到行为的抽象表征，其同时考量了序列影响，空间影响以及时间影响等因素。

3.2.3.2 社交关系感知校正

接下来，为了判断用户的签到行为受其历史签到行为的影响程度，GraphSAR 设计了香草注意力机制（Vanilla attention layer）来借助社交关系进一步优化用户表征。在训练阶段，GraphSAR 学习用户与朋友之间的兴趣关联向量，通过引入一个权重矩阵 \mathbf{W}_s 来建模用户和其朋友的自注意变换：

$$\mathbf{c}_{uv} = \text{ReLU}(\mathbf{W}_s^T (\mathbf{h}_{u_i} || \mathbf{h}_{u_j})) \quad (3.20)$$

其中， \mathbf{h}_{u_i} 和 \mathbf{h}_{u_j} 分别是用户 u_i 和其朋友 u_j 的表达。 T 表示矩阵转置操作， $||$ 表示拼接操作，ReLU() 表示整流激活函数。 \mathbf{c}_{ij} 表示朋友 u_j 对用户 u_i 的重要程度的注意力系数，这里为了保证能够表达每个维度上的相关性，本文将 \mathbf{W}_s 的维度设定为 $d * 2d$ 。对每一个用户 u ，会获取一个兴趣关联矩阵 \mathbf{C}_{u_i} ， \mathbf{C}_{u_i} 的每一行代表

用户 u_i 和他的一个朋友的兴趣关联向量。为了使得用户在每个维度上和所有朋友的相似性得以区分，本文对兴趣关联矩阵 \mathbf{C}_{u_i} 按行进行归一化。这里，通过 α_{ij}^d 表示用户 u_i 和其朋友 u_j 的归一化后的兴趣关联向量，每个维度上的值计算方法如下：

$$\alpha_{ij}^d = \frac{\exp(c_{ij}^d)}{\sum_{k \in \mathbb{F}_{u_i}} \exp(c_{ik}^d)} \quad (3.21)$$

最后，利用用户 u_i 的每个朋友的表征来对自身的表征进行平滑，可以形式化为：

$$\tilde{\mathbf{h}}_{u_i} = \sigma(\sum_{k \in \mathbb{F}_{u_i}} \alpha_{ij} \odot \mathbf{h}_{u_j} + \mathbf{h}_{u_i}) \quad (3.22)$$

其中， $\tilde{\mathbf{h}}_{u_i}$ 表示用户平滑之后的表征， \odot 表示两个表征的 Hadamard 乘积， $\sigma(x) = \frac{1}{1+e^{-x}}$ 为非线性函数。经过平滑之后，用户的表征在自身表征的基础上，差异化地集成了所有朋友的表征中的每个维度的值。这种操作使得平滑后的表征能够维持用户自身独特的兴趣偏好，同时也可和朋友共享一些兴趣。

3.2.4 下一个兴趣点推荐

本文基于候选兴趣点表征向量 \mathbf{e}_N^p 以及用户隐表征向量 $\tilde{\mathbf{h}}_N^u$ ，推断用户在接下来可能访问兴趣点的概率为：

$$\mathbf{y} = \text{softmax}((\tilde{\mathbf{h}}_N^u)^T \mathbf{e}_N^p) \quad (3.23)$$

最终，通过 $\text{argmax}_p(\mathbf{y})$ 为用户推荐最有可能访问的 POI \hat{p} 。在训练过程中，本文参照 BPR-Loss 的个性化排名损失训练方式^[74] 进行推荐任务损失函数的构建：

$$\mathcal{L}_{GraphSAR} = - \sum_{p^+} \sum_{p^-} \ln \sigma(y_{p^-} - y_{p^+}) + \mu ||\emptyset|| \quad (3.24)$$

其中， p^+ 表示用户访问过的 POI， p^- 表示到目前为止用户没有访问过的 POI， σ 表示非线性函数，即 $\sigma(x) = \frac{1}{1+e^{-x}}$ ， \emptyset 表示模型的参数集合。

3.2.5 复杂度分析

算法 3.1: GraphSAR NPOI 推荐算法

输入:签到数据集 D , NPOI 推荐模型参数集合 Θ **输出:**

给每个用户生成 NPOI 推荐列表

```

1: for each POI do
2:   获取基于经纬度的位置空间 Geohash 编码和类别编码
3:   根据历史记录获取位置被访问时间编码
4:   根据变分自编码机获取时间特征编码
5: end for
6: 根据公式 (3.6) 构造用户与位置 POI 交互网络
7: while count <  $\eta$  do
8:   根据公式 (3.7) (3.8) 获取 POI 隐语义特征
9: end while
10: 根据公式 (3.13) 获取 POI 表征向量
11: for each user do
12:   通过 K 层偏好网网络获得用户表征
13: end for
14: for each user do
15:   借助于社交注意力层平滑用户表征
16:   根据用户表征和 POI 表征进行 NPOI 推荐
17: end for

```

GraphSAR 中, NPOI 推荐的计算过程如算法 3.1 所示, 主要包括两部分: POI 表征和用户表征。POI 表征过程包括时空特征编码和基于图卷积神经网络的关系表征两个部分: 首先, 时空特征编码部分中地理位置的 Geohash 编码的计算开销与兴趣点的数量成线性关系, 即 $O(N_{POI})$, 而通过差分自编码机进行时间编码部分的时间复杂度为 $O(d^2N_{POI})$, 其中, d 表示初始时间表征向量的维度, 亦是线性关系; 其次, 对于基于图卷积神经网络得到多通道关系特征获取部分, 在一个循环单元中, 图卷积神经网络的计算开销为 $O(dCH)$, 其中, C 表示输入表征向量的维度, H 表示卷积神经网络中隐层的单元节点的数量, d 表示节点特征的维度。用户表征过程主要包括 K 层序列编码网络和社交关系注意力层两个部分: 首先, 序列编码部分的计算开销与签到序列长度是线性关系。在一个循环步骤中,

GraphSAR 序列编码单元处理签到序列中记录的过程仅与当前记录有关，是一个固定的常数。因此，编码部分的计算复杂度是线性的，可以视为 $O(Kd^2N_{user})$ ， d 表示用户签到记录表征向量的维度值， N_{user} 表示用户数量；其次，社交关系注意力层，主要计算复杂度在于计算用户间的兴趣关联向量，这部分计算开销可以视为 $O(d^2N_{user})$ 。由此可见，GraphSAR 的总计算复杂度是线性的，这使得 GraphSAR 适用于大型数据集。

3.3 性能评估

为了评估 GraphSAR 的 NPOI 推荐性能，本文在两个真实移动社交网络数据集上进行实验，并与多种已有的 NPOI 推荐方法进行对比。本节首先简要阐述实验设置情况，然后展示并分析实验结果。

3.3.1 实验设置

本节首先介绍了用于评测 GraphSAR 方法的数据集，然后介绍了对比方法及评价指标，最后阐述了实验中用到的参数设置。

3.3.1.1 实验数据

用于实验的两个公开数据集分别是：Gowalla 和 Foursquare。表 3.1 展示了这两个数据集的数据分布和统计信息。

表 3.1 数据集统计数据

Table 3.1 Statistics of our datasets

数据集	签到数量	用户数量	POI 数量	朋友关系
Gowalla	12,846,151	216,734	1,421,262	736,778
Foursquare	22,809,624	114,324	3,820,891	607,333

Gowalla 是一个基于位置的社交网络，它为用户提供个性化的位置推荐服务。网络中的用户可以发布其对兴趣点的签到信息（Check-ins），并且可以和其他的用户建立社交关系。

Foursquare 也是在位置社交网络领域中的一个大规模的社交网站，允许用户在不同的地理位置进行签到行为，在这里我们使用在 2012 年 9 月到 2013 年 9 月之间的 Foursquare 公开的去隐私化的全球规模用户签到数据集。其中包含用户的

签到信息和社交关系信息。

Gowalla 和 Foursquare 数据集中均带有用户访问的 POI 信息，因此可以直接利用 POI 位置类别标签作为位置语义标签。在实验过程中，我们将所有的标签归为十大类：餐馆、酒店、景点、交通枢纽、商店、教育、娱乐、宗教、公司以及其他类别，协助进行基于 GCN 的 POI 表征网络的参数训练。预处理过程中，为了保证提取有效的时空特征信息，我们过滤了上述两个数据集中签到记录低于 10 条的用户和被访问记录少于 10 条的 POI。对于每个数据集，GraphSAR 随机选择 70% 的签到数据用来构建训练集，并保留 10% 和 20% 的数据分别用于验证和测试。

3.3.1.2 对比方法及评价指标

为了全面评估 GraphSAR 算法的有效性，本文选择了多种现有的 NPOI 推荐方法作为对比方法：

- **SoCF^[37]**: 该算法是一个基于社交关系的协同过滤方法，它直接推荐朋友喜欢的物品给用户；
- **PTPMF^[90]**: 该算法在原有的概率矩阵分解模型中引入强弱连接进行社交关系的建模；
- **STRNN^[55]**: 该算法在循环神经网络的基础上增加了时间敏感性，捕捉在签到行为之间的时间和空间位置信息变化；
- **DeepMove^[115]**: 该算法在循环神经网络的基础上增加了用户兴趣注意力机制，考虑用户签到序列时序关系的基础上考虑了用户的长期兴趣；
- **ST-LSTM^[112]**: 该算法是长短期记忆网络(LSTM)的变式，在 LSTM 中增加了时间和距离两种门限机制来捕捉用户签到行为序列之间的时空关系来进行 POI 推荐；
- **GraphSAR_{DeepWalk}**: GraphSAR 方法的变式，使用图表征方法 DeepWalk^[65]替换 GraphSAR 中的 POI 隐语义表征部分进行兴趣点推荐；
- **GraphSAR_{Node2Vec}**: GraphSAR 方法的变式，使用图表征方法 Node2Vec^[36]替换 GraphSAR 中的 POI 隐语义表征部分进行兴趣点推荐；
- **GraphSAR_{GCN}**: GraphSAR 方法的变式，使用图表征方法 GCN 替换 GraphSAR

中的 POI 隐语义表征部分进行兴趣点推荐;

- GraphSAR_{STRNN}: GraphSAR 方法的变式, 使用序列编码方法 STRNN 替换 GraphSAR 中的序列编码部分进行兴趣点推荐;
- GraphSAR_{sp.}: GraphSAR 方法的变式, 不考虑空间地理特征, 进行位置语义推断;
- GraphSAR_{t.}: GraphSAR 方法的变式, 不考虑位置的时间访问特征, 进行位置语义推断。

评价指标方面, 采用两种推荐指标来评估推荐模型的性能: 命中率 (Hit Rate, HR@N) 和归一化累计增益 (Normalized Discounted Cumulative Gain, NDCG@N):

$$HR@N = \frac{Number\text{Of}Hits@N}{GT} \quad (3.25)$$

GT 表示所有的测试集合, $Number\text{Of}Hits@N$ 表示推荐列表的 Top-N 子集中属于测试集合的个数的总和。比如 $HR@10$ 表示推荐列表的 TOP-10 子集命中真实结果的比率。

$$NDCG@N = \frac{DCG@N}{IDCG} \quad (3.26)$$

这里, DCG 表示折损累计增益, 是在累计增益 CG 的基础上引入了位置影响因素, 计算公式如下:

$$DCG@N = \sum_{i=1}^N \frac{2^{rel_i} - 1}{log_2(i + 1)} \quad (3.27)$$

可以看出, 推荐的相关性越多, DCG 值越大, 相关性强的排在推荐列表前面的话, 推荐效果越好, DCG 值越大。在式(5.17)中, $IDCG$ 表示推荐系统针对某一用户返回的最好推荐列表, 即假设返回结果按照相关性排序, 最相关的结果排在最前面的情况下 DCG 值。由定义可以看出 $NDCG$ 的值介于 $(0, 1]$ 。

3.3.1.3 参数设置

GraphSAR 在 Tensorflow 框架下实现, 其中的关键参数包括: (1) 空间特征表征向量维数设置为 10; (2) 时间特征向量维度为 16; (3) 用户语义表征向量维度 d_{user} , 本实验通过在范围 $\{16, 32, 64, 128, 256\}$ 中进行网格搜索来确定 d_{user}

的值。通常，性能随着 d 增加而在足够大时逐渐稳定，发现在 $d = 128$ 时，实验结果最佳，因此设置 $d = 128$ ；（4）网络的初始学习率设置为 λ ，本实验在 {0.1, 0.05, 0.01, 0.001} 中进行了搜索，衰减权重为 $1e^{-4}$ ，发现最佳效果为 $\lambda = 0.01$ 。小节 3.3 详细介绍参数敏感度分析结果。

3.3.2 实验结果分析

本节分别分析了推荐命中率实验、可视化实验和参数敏感性实验的实验结果。

3.3.2.1 HR@N 和 NDCG@N 指标比较

表 3.2 和表 3.3 是各种对比方法在 HR@5、HR@10、NDCG@5 以及 NDCG@10 指标上的实验结果。可以看出本文的方法与其他基线方法相比在 NPOI 推荐任务上表现更佳。下面对实验结果进行详细分析：

表 3.2 各类方法在 Foursquare 数据集上的推荐结果统计

Table 3.2 Performance comparison for different methods on Foursquare

方法	Foursquare			
	HR@5	NDCG@5	HR@10	NDCG@10
SoCF	0.0298	0.0317	0.0571	0.0637
PTPMF	0.0418	0.0427	0.0912	0.0955
STRNN	0.0559	0.0652	0.1221	0.1307
DeepMove	0.0618	0.0782	0.1255	0.1403
ST-LSTM	0.0662	0.0760	0.1301	0.1457
GraphSAR _(DeepWalk)	0.0728	0.0840	0.1423	0.1576
GraphSAR _(Node2Vec)	0.0782	0.0904	0.1401	0.1622
GraphSAR _(GCN)	0.0803	0.1067	0.1523	0.1618
GraphSAR _(STRNN)	0.0775	0.0920	0.1399	0.1421
GraphSAR _{sp-}	0.0733	0.0856	0.1472	0.1541
GraphSAR _{t-}	0.0815	0.1069	0.1565	0.1645
GaraphSAR _{s-}	0.0822	0.1083	0.1601	0.1723
GraphSAR	0.0970	0.1173	0.1723	0.1826

表 3.3 各类方法在 Gowalla 数据集上的推荐结果统计

Table 3.3 Performance comparison for different methods on Gowalla

方法	Gowalla			
	HR@5	NDCG@5	HR@10	NDCG@10
SoCF	0.0251	0.0285	0.0518	0.0559
PTPMF	0.0390	0.0401	0.0866	0.0933
STRNN	0.0432	0.0552	0.1021	0.1107
DeepMove	0.0522	0.0577	0.1155	0.1163
ST-LSTM	0.0537	0.0630	0.1167	0.1201
GraphSAR _(DeepWalk)	0.0631	0.0672	0.1197	0.1315
GraphSAR _(Node2Vec)	0.0649	0.0770	0.1194	0.1365
GraphSAR _(GCN)	0.0675	0.0907	0.1301	0.1381
GraphSAR _(STRNN)	0.0661	0.0787	0.1184	0.1212
GraphSAR _{sp-}	0.0643	0.0779	0.1127	0.1191
GraphSAR _{t-}	0.0661	0.0969	0.1371	0.1491
GaraphSAR _{s-}	0.0691	0.0934	0.1453	0.1503
GraphSAR	0.0766	0.1093	0.1537	0.1653

表 3.2 和表 3.3 统计了在 Foursquare 和 Gowalla 两个数据集下的 NPOI 推荐结果。对于 Check-in 类型的数据集，如两表所示，GraphSAR 的 NPOI 推荐结果的命中率和归一化累计损益值均高于其他对比的方法。以 Foursquare 数据集为例，GraphSAR 在 HR@10 上可以取得 4% 的提升。所有对比方法中，SoCF 与 PTPMF 方法（为 MF/CF 类推荐方法）在准确率上的表现较差，这说明建模用户签到序列的时序关系对 NPOI 推荐任务起到了至关重要的作用。在所有建模时序关系的推荐方法中，GraphSAR 与 STRNN 方法在准确率上的差距最大，这是因为 POI 语义复杂多样，影响兴趣推荐结果的因素有很多。而 STRNN 方法仅通过设置时空注意力窗口来进行时序特征获取，没有考虑蕴含在 POI-用户交互网络中大量的隐藏语义信息，因此难以获得有效的兴趣点表征向量，难以取得好的 NPOI 推荐效果。对比特征抽取类方法（STRNN、ST-LSTM），图表征学习类方法（Deepwalk、Node2vec、GCN）方法效果较仅利用时序关系的方法要好，但是依然逊于 GraphSAR。这是因为图表征学习类的方法在进行位置语义表征的过程中考虑了用户-POI 图中蕴含的隐语义信息，使得最终的 NPOI 推荐命中率高于特征抽取类方法。与传统的图表征学习类方法相比，GraphSAR 在 NPOI 推荐命中

率上有了较大提升，主要是因为 GraphSAR 借助于异构神经网络中的元路径机制进行了多通道语义考量，并且借助于卷积神经网络进行深层语义信息的探索。除此之外，GaraphSARs-方法明显逊于 GraphSAR，这是因为 GraphSAR 方法在进行 NPOI 推荐过程中建模了用户的社交关系影响，提升了推荐效果。虽然 GraphSAR 只获得了 4%左右的性能提升，但是能够在没有专家知识干预的情况下，通过自动学习用户间的多维度兴趣点偏好关联并获得性能提升，充分体现了 GraphSAR 在 POI 特征提取与社交关系建模方面的能力。另外，所有方法在 Gowalla 数据集上的效果都略逊于 Foursquare，这是因为在 Gowalla 的用户签到行为频率要远高于 Foursquare，且签到位置语义标签分布更加分散，造成其 NPOI 推荐难度加大。

3.3.2.3 元路径有效性分析

本小节通过实验对比不同的元路径引导下的 GraphSAR 的 NPOI 推荐性能，结果如图 3.8 和图 3.9 所示。本实验分别在 Foursquare（东京）和 Foursquare（纽约）数据集上测试两条元路径（ $m1: p \rightarrow u \rightarrow p$ 和 $m2: p \rightarrow u \rightarrow u \rightarrow p$ ）以及其组合的关系表征表现。相较于使用元路径 $m2$ ，使用元路径 $m1$ 语义推断效果更佳，说明通过 $m1$ 通道提取的 POI 关系特征具有更大的区分度。这里， $m1+m2$ 取得了最佳效果，说明整合多通道关系特征有助于提升 NPOI 推荐性能。

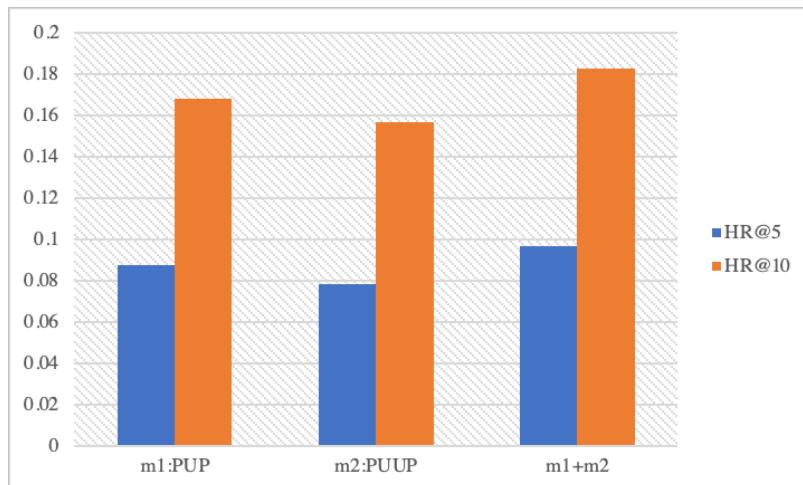


图 3.8 Foursquare（纽约）中不同元路径引导下的 GraphSAR 性能分析

Figure 3.8 The results of GraphSAR under different meta-paths for Foursquare(NYC)

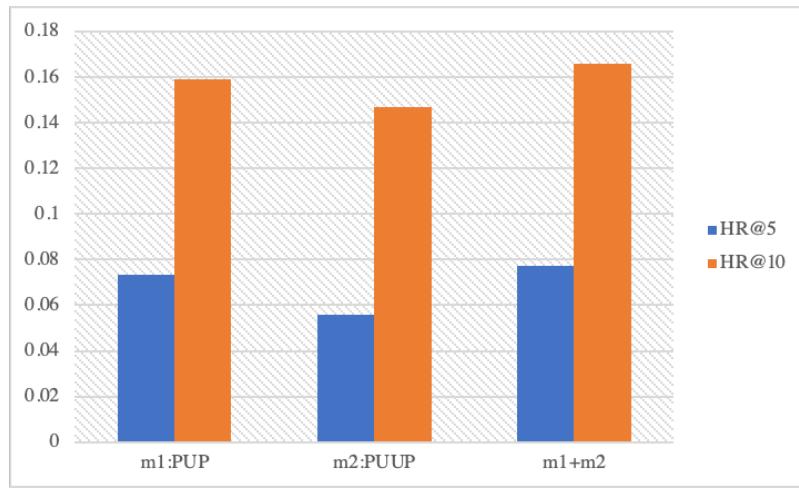


图 3.9 Foursquare (东京) 中不同元路径引导下的 GraphSAR 性能分析

Figure 3.9 The results of GraphSAR under different meta-paths for Foursquare(TKY)

3.3.2.4 位置隐语义表征有效性分析

本实验首先随机选取 6 个签到位置，分别为 P1->P6，并抽取其经过 GraphSAR 表征之后的 POI 表征向量进行可视化分析。图 3.10 表示 6 个位置的用户访问量分布图（注：这里我们将时间统一映射到 0-24 小时），可以看出不同签到位置具有不同的时间访问模式，具有相似类别标签的签到位置时间访问模式类似，比如 P3 和 P5，两个 POI 的类别均为酒吧（Bar），在访问模式上也近似，均在晚上 20 点之后访问量呈现上升趋势，在凌晨 0-1 时达到顶峰。同样的现象也发生在 P2 和 P6 上，其类别标签均为咖啡店（Coffee Shop）。

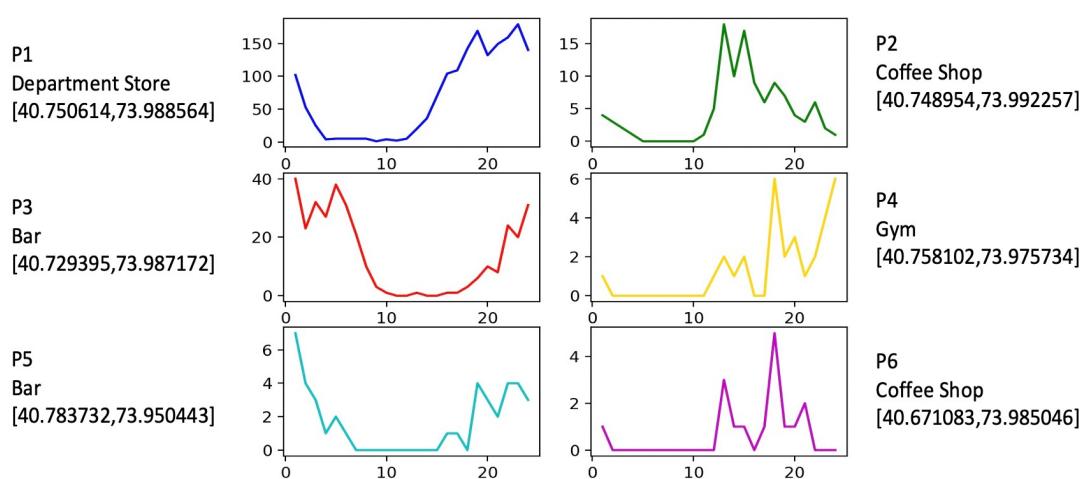


图 3.10 签到位置 P1-P6 的访问分布图

Figure 3.10 The distributions of POI P1-P6

图 3.11 表示位置 P1->P6 语义表征向量之间的余弦相似度，可以看出相同类别标签的 POI，其表征向量的相似度越高，比如 P3 和 P5，表明 GraphSAR 获得的 POI 表征向量能够实现对 POI 语义特征的有效提取。

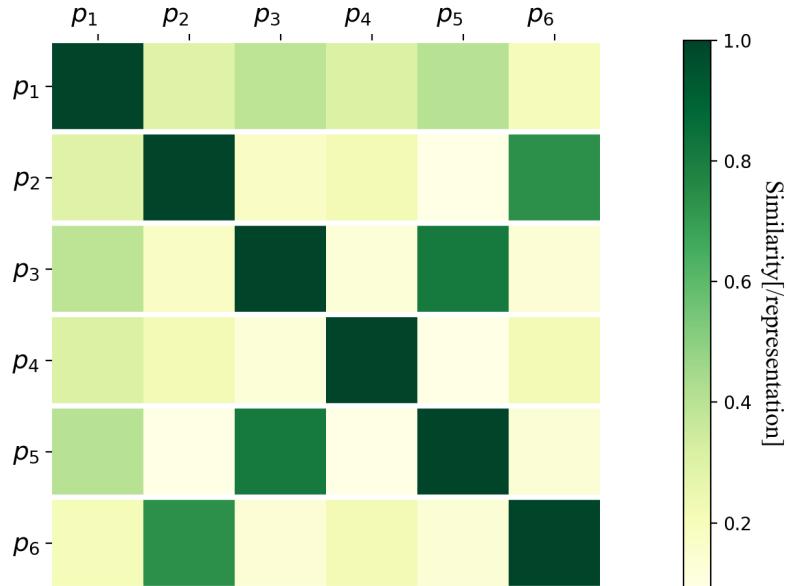


图 3.11 签到位置 P1->P6 表征向量的相似度

Figure 3.11 The similarity of representations in P1->P6

3.3.2.4 参数评估

GraphSAR 兴趣点推荐算法有两个重要的参数：时间表征维度 d_t 和关系表征维度 d_r 。研究这两个参数对 GraphSAR 的影响时，通过改变其中一个参数的值，固定其他参数，来分析其对最终 NPOI 推荐结果的影响。首先对参数 d_t 进行分析，让其在 [2, 4, 8, 16, 32] 范围内取值，图 3.12 展示了 GraphSAR 模型的命中率和累计损益值随着 d_t 的变化趋势，发现均为先上升后下降，在 $d_t = 16$ 时，命中率和累计损益值达到最优值。这说明在维度过低情况下，表征向量对时间特征表达能力受限，导致 NPOI 推荐性能下降；在维度过高情况下，表征向量中存在冗余信息，也会降低 NPOI 推荐结果。除此之外，本文对参数 d_r 也进行了详细分析，如图 3.12 所示，GraphSAR 模型的命中率和累计损益值有着相似变化趋势，随着维度的变大命中率和累计损益值先上升后下降，在 $d_r = 64$ 时，命中率和累计损益值达到最优值。这也说明了合理设置关系维度值对于兴趣点推荐性能的提高是有一定帮助的。

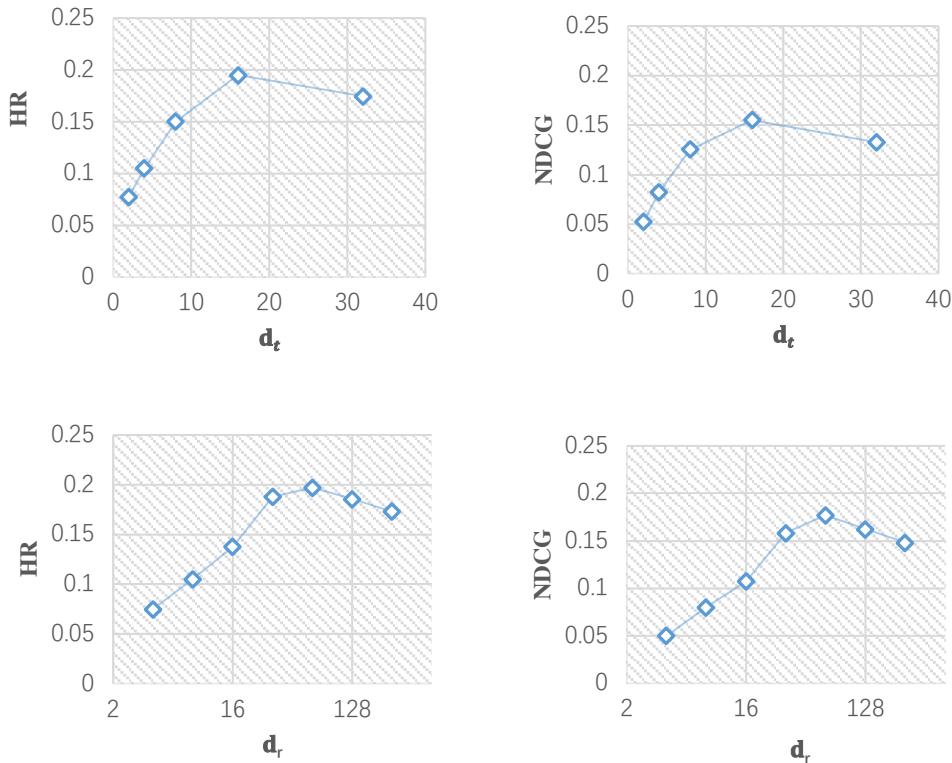


图 3.12 Foursquare (纽约) 上 d_t 和 d_r 对命中率和累计损益的影响

Figure 3.12 The results of HR and NDCG with respect to the value of d_t and d_r on Foursquare(NYC)

3.4 小结

本文针对用户社交关系建模不准确导致模型难以建模用户多维度偏好的问题，提出了一种多维度社交关系感知的下一个兴趣点推荐方法 GraphSAR。GraphSAR 的创新性在于两个方面：其一，设计了一种基于元路径指导的多通道特征提取框架，它能够从异构的用户-POI 交互网络中抽取签到位置（POI）的多通道语义关系；其二，设计了一种基于多维度社交注意力机制的用户偏好表征网络，通过多头注意力机制进行了用户前后签到行为的关系关联，并通过引入多维度社交关系注意力来建模用户的签到行为受社交关系的影响，解决了现有方法难以建模用户多维度偏好差异的问题。实验结果表明，对比基线方法，GraphSAR 在下一个兴趣点推荐任务上的效果更好。

第4章 用户动态偏好感知的下一个兴趣点推荐方法研究

本章针对用户连续行为偏好刻画缺失的问题，提出了用户动态偏好感知的下一个兴趣点推荐方法（Dynamic Preference Intensive Neural Network for NPOI recommendation, PINN）。该方法通过整合用户的历史静态偏好和当前动态偏好来进行下一个兴趣点推荐任务。PINN 包括两个主要部分：偏好记忆网络和偏好动态演化网络。在偏好记忆网络中，提出了一种基于用户群体移动行为模式的用户历史偏好表征网络，该网络可以建模用户群体移动模式，并在群体移动模式的指导下获取用户的历史偏好表征；在动态偏好演化网络中，提出了基于时空神经常微分方程的用户动态偏好感知方法，该方法首次将神经常微分方程用于 POI 推荐领域，定义连续的动态过程，可以接受任意时刻的输入数据，克服了大多数时序推荐模型只能处理静态离散化的时间间隔的局限性，更适用于 POI 序列推荐任务。实验结果证明，PINN 可以显著提升 NPOI 推荐性能。

本章组织如下：4.1 节详细阐述了研究问题的内容和挑战。4.2 节详细介绍了用户动态偏好感知的下一个兴趣点推荐方法，PINN。其中，4.2.1 节介绍了用户动态偏好感知的下一个兴趣点推荐问题定义和方法概述；4.2.2 节介绍了基于偏好记忆网络的用户历史偏好感知方法；4.2.3 节详细介绍了基于时空神经常微分方程的用户动态偏好感知方法；4.2.4 节介绍了用户动态感知的下一个兴趣点推荐方法的推荐模式及训练过程。4.3 节在真实数据集上验证了上述方法的有效性；4.4 节对本章进行了小结。

4.1 引言

兴趣点(POI)表示用户感兴趣的已经访问或者将会访问的地理位置。推荐的潜在兴趣点表示用户在将来某个时间点可能会访问的地理位置。作为面向位置服务的移动社交网络数据挖掘领域的重要分支，下一个兴趣点推荐可以预测用户行为和兴趣演化趋势，对城市交通的规划和商业模式的探索具有极其重要的现实意义。现实应用中，由于影响用户的兴趣点偏好的因素众多，并随着时间动态变化，难以直接进行建模，导致下一个兴趣点推荐模型不能感知用户动态偏好，推荐效

果不佳。

为了解决下一个兴趣点推荐中用户动态偏好的建模问题，现有研究工作提出了很多利用签到行为的时空特征进行用户偏好建模的方法以提升下一个兴趣点推荐效果。根据兴趣点建模用户偏好的不同方式，将现有方法划分为基于划分时间槽的用户偏好建模方法和基于点过程的用户偏好建模方法。基于时间槽的用户偏好建模方法可以进一步细分为两个子类：(1) 将时间划分为时间槽，建模用户在不同时间槽内的行为模式，比如 Liu 等^[55]在循环神经网络的基础上通过时间窗口矩阵增加时间敏感性，捕捉用户签到行为之间的时间和空间位置偏好。该类方法只能建模固定时间槽内的签到序列关系，不能捕捉用户签到行为周期性时序模式；(2) 将时间划分为时间槽，根据签到时间点进行时间编码，作为时间特征，来进一步获取用户的偏好表征向量，比如 Zhou 等^[115]等将签到点按照时间槽进行时间编码，联合用户基本信息作为当前时刻的用户偏好编码。该类方法受限于时间槽的划分，不适用于连续时间建模。基于点过程的用户偏好建模方法，通过将用户偏好视为强度分布函数进行建模，比如 Du 等^[27]提出了移动社交网络中用户偏好时间敏感性问题，并利用 Hawkes^[58]过程解决时间和位置联合预测问题。除此之外，Mei 和 Eisner^[112]提出了通过基于LSTM的点过程（Continuous LSTM）来进行隐式反馈信息的捕捉，弱化强度函数的分布，利用LSTM从数据中学习强度函数的返回值。该类方法受限于随机点过程方法中对强度函数分布的假设，没有考虑数据半隐式反馈特点对时间空间联合预测的影响，仍未能解决用户偏好动态变化问题。

针对上述问题，本文提出了一种用户动态偏好感知的下一个兴趣点推荐方法 PINN。该方法包括用户偏好历史记忆捕捉和动态偏好演化两个阶段。在用户偏好历史记忆捕捉阶段，本文采用记忆网络（Memory Network）对用户表征向量进行处理并学习用户的移动模式表示向量。具体来说，给定待处理的用户原始签到序列，PINN 首先将每一条签到记录转换为固定长度的向量，使得该向量可以有效地编码特定的签到行为。然后将用户的签到行为序列整合获得用户的表征向量，经过对记忆网络的读取之后，获得固定维度的用户移动模式表征向量。在动态偏好演化阶段，PINN 利用神经常微分方程学习用户任意时刻的签到行为，捕

获用户的动态即时偏好。然后整合用户历史偏好和动态偏好来实现对用户偏好的精准感知，提升兴趣点推荐效果。本文使用六个城市的真实签到数据集评估了 PINN 的效果。实验结果表明，对比最好的基线方法，PINN 在 NPOI 推荐任务中的 HR@10 指标上提升了 7%。

4.2 用户动态偏好感知的下一个兴趣点推荐方法

本节详细介绍了用户动态偏好感知的兴趣点推荐方法 PINN。首先定义了兴趣点推荐任务并概述了 PINN 的主要流程。然后针对 PINN 中的主要部分，详细阐述了其网络结构和参数优化过程。

4.2.1 问题定义与方法概述

本小节首先定义了基于位置的社交网络下的下一个兴趣点推荐问题，然后简要介绍了本文提出的用户动态偏好感知的下一个兴趣点推荐方法 PINN。

4.2.1.1 问题定义

对于 M 个用户 $U = \{u_1, \dots, u_M\}$ 、签到序列集合 \mathcal{T} 、N 个兴趣点 $P = \{p_1, \dots, p_N\}$ 和 F 个位置语义标签 $C = \{c_1, \dots, c_F\}$ ，每个目标用户 u 都对应一条历史签到记录序列 $\mathcal{T}_u = [r_1, \dots, r_K]$ ，其中每条签到记录 $r_k = (l_k, c_k, t_k)$ 是一个三元组，包含了地理位置 l_k 、兴趣点的标签类别 c_k 以及时间戳 t_k 。根据上述描述，给出时间敏感的下一个兴趣点推荐的定义：

时间敏感的下一个兴趣点（NPOI）推荐：给定用户集合 U，历史签到序列集合 \mathcal{T} 。对于目标 u ，存在一条到时间 t_K 的用户历史签到序列 \mathcal{T}_u ，我们的目标是根据用户历史签到序列为其实现推荐在下一个时间点 t 可能访问的兴趣点。



图 4.1 时间敏感的NPOI推荐示例图

Figure 4.1 Illustration of the time-aware NPOI recommendation

4.2.1.2 方法概述

用户动态偏好感知的 NPOI 推荐流程如图 4.2 所示：

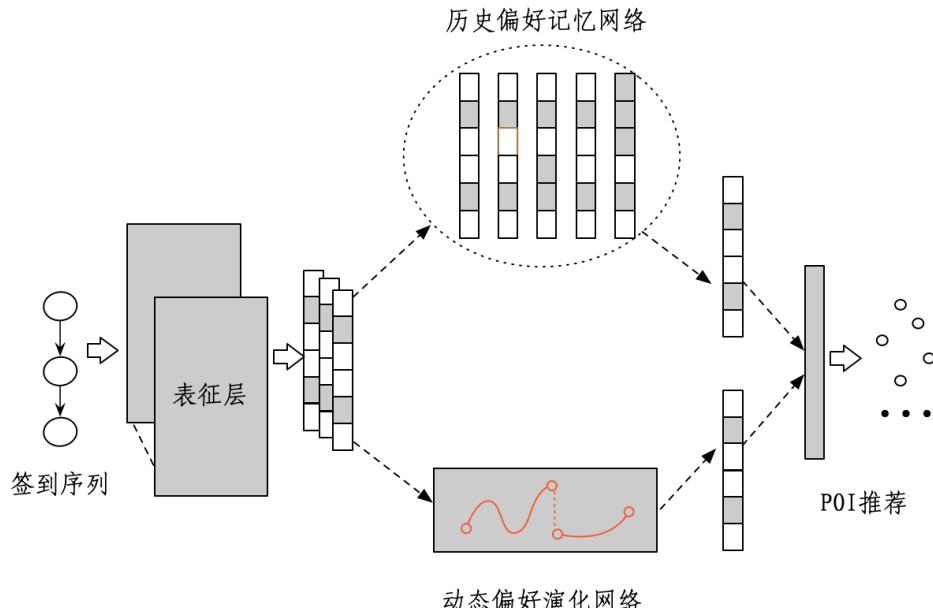


图 4.2 PINN 示意图

Figure 4.2 The framework for POI recommendation

- 签到行为表征层：该层以原始签到行为序列作为输入。在特征层面，我们仅考虑 POI 的类别，POI 位置以及时间戳三类特征，即 (c_k, l_k, t_k) ，这里 c_k 表示该签到行为所在 POI 的所属类别， l_k 表示该 POI 的所在位置，由经纬度坐标表示， t_k 表示发生该签到行为的时间点。
- 用户历史偏好记忆网络：在该层中，输入的签到行为特征序列将被用户偏好特征提取算法处理，并根据记忆网络中存储的移动行为模式获取用户的表征向量。
- 用户动态偏好演化网络：利用基于神经常微分方程的时序关系编码来建模用户的动态偏好，可以接受任意时刻的输入数据，将移动特征序列编码成固定长度的向量，该表征向量能够有效地保存用户的动态偏好信息。
- NPOI 推荐层：以用户历史偏好表征向量和用户动态偏好表征向量为基础，通过注意力机制聚合获取用户的精准偏好表征，并根据表征向量进行 NPOI 推荐。

可以看出 PINN 中有四个核心部分，即签到行为表征、用户偏好记忆网络、用户偏好动态演化网络以及 NPOI 推荐。下面将逐个进行介绍。

4.2.2 签到行为表征

在 POI 表征部分，利用一个嵌入层获取签到行为序列中每一个签到行为的表征向量：

(1) 针对 POI 类别属性， c_k ，引入矩阵 $\mathbf{E}_c \in \mathbb{R}^{F \times d_c}$ ，这里， F 表示 POI 的类别数， d_c 表示 POI 类别表征向量的维度。在这里，矩阵 \mathbf{E}_c 是在多个城市之间进行共享；

(2) 针对 POI 地理位置属性， l_k ，引入矩阵 $\mathbf{E}_l \in \mathbb{R}^{N \times d_l}$ ，这里， N 表示 POI 的个数， d_l 表示 POI 位置属性表征向量的维度。在这里，本文通过经纬度坐标位置的 GeoHash 编码来初始化 POI 矩阵 \mathbf{E}_l ；

(3) 针对时间戳， t_k ，引入矩阵 $\mathbf{E}_t \in \mathbb{R}^{168 \times d_t}$ ，在这里，168 表示将一周分成 168 个时间槽 (7*24 小时)，每一个时间槽的表征向量维度为 d_t 。

然后，针对每一条签到记录，本文将三类特征的表征向量进行聚合，可以获得长度为 d 的表征向量 \mathbf{r}_k ，作为后续时空神经常微分方程的输入，这里聚合过

程可以形式化为：

$$\mathbf{r}_k = \varphi(\mathbf{W}_l \mathbf{e}_k^l + \mathbf{W}_c \mathbf{e}_k^c + \mathbf{W}_t \mathbf{e}_k^t + b) \quad (4.1)$$

其中， \mathbf{e}_k^l 、 \mathbf{e}_k^c 和 \mathbf{e}_k^t 分别表示签到记录中 POI 位置、POI 的类别以及时间戳三类特征的表征向量， \mathbf{W}_l 、 \mathbf{W}_c 和 \mathbf{W}_t 表示相关特征类别的权重矩阵， b 为偏置参数。

4.2.3 用户偏好记忆网络

为了根据用户历史签到序列提取用户历史偏好，本文利用偏好记忆网络（User preference Memory network，User-Mem）进行移动模式的探索和发现。如图 4.3 所示，User-Mem 抽取用户签到序列中的时空不变性的特征用于描述用户的偏好信息，以用户 u 的历史签到序列作为输入，User-Mem 首先将先前访问的 POI 的类别信息嵌入到向量 \mathbf{u}_c 中。然后，它使用一个精心设计的存储网络按照用户组对用户历史偏好进行建模，并产生 u 的最终用户偏好向量 \mathbf{u} 。详细地获取过程如下：

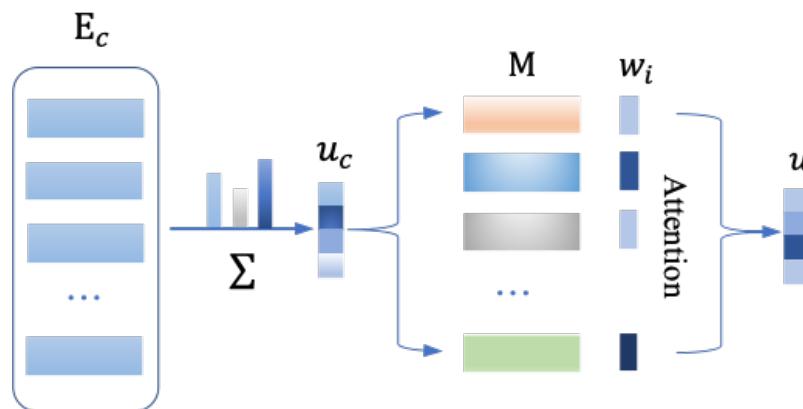


图 4.3 用户偏好记忆网络

Figure 4.3 Illustration of the preference Memory Network

首先，假设签到数据集中存在 N 个 POI 类别标签， S 个用户偏好分组，分别引入矩阵 $\mathbf{E}_c \in \mathbb{R}^{N \times d_c}$ 和 $\mathbf{M} \in \mathbb{R}^{S \times d_m}$ 来分别表示类别标签表征矩阵和用户组表征矩阵。给定用户 u 历史签到序列 $T_u = [r_1, r_2, \dots, r_K]$ ，User-Mem 提取访问过的 POI 的类别标签构成标签访问序列 $[c_1, c_2, \dots, c_K]$ 。通过查询矩阵 \mathbf{E}_c ，获取标签表征序列 $[\mathbf{e}_1^c, \mathbf{e}_2^c, \dots, \mathbf{e}_K^c]$ ，进一步获取用户的表征向量：

$$\mathbf{u}_c = \sum_{k=1}^K w_k^c \cdot \mathbf{e}_k^c \quad (4.2)$$

其中， w_k^c 表示每个类别的权重值，满足 $\sum_{k=1}^K w_k = 1$ 。然后，将用户的类别表征向量作为输入输送到记忆网络中。如架构图中所示，矩阵 $M \in \mathbb{R}^{S \times d_m}$ 中存储所有用户组的移动模式， S 表示记忆矩阵中的槽位数量。这里，将用户的类别标签向量作为查询向量， M 作为上下文矩阵，通过注意力机制生成用户的最终表征向量：

$$w_i^g = \text{softmax}(\mathbf{u}_c^T \mathbf{m}_i) \quad (4.3)$$

$$\mathbf{e}_u = \sum_{i \in 1 \dots S} w_i^g \mathbf{m}_i \quad (4.4)$$

其中， m_i 表示记忆网络中的第*i*个槽位，即第*i*个用户组的表征向量； w_i^g 表示基于用户签到POI标签获取的偏好向量 \mathbf{u}_c 与第*i*个用户组偏好的相似程度。直观地讲，从记忆网络机制可以看出，User-Mem具有一定的鲁棒性，可以借助于用户组的行为模式为签到记录较少的用户进行偏好建模。

为了优化 \mathbf{E} 、 \mathbf{M} 以及在历史记忆网络中的其他参数，本文提出了一种点对优化机制。具体来说，首先通过训练一个评价层去计算任意用户-标签对之间的访问评分值，然后通过运用负采样机制来设计点对优化目标函数。给定一组用户-标签点对 $\langle u, c \rangle$ ，通过User-Mem获取表征向量 \mathbf{e}_u 和 \mathbf{e}_c ，然后计算两个向量的Hadamard乘积 $\mathbf{u}_{\langle u, c \rangle}$ ，即 $\mathbf{u}_{\langle u, c \rangle} = \mathbf{e}_u \odot \mathbf{e}_c$ 。然后，通过一个全连接层获取点对评分：

$$s_{\langle u, c \rangle} = f(\mathbf{u}_{\langle u, c \rangle}) \quad (4.5)$$

其中， f 表示全连接层，将向量转换成数值评分。

针对用户u的历史签到序列 \mathcal{T}_u ，在训练过程中，从 \mathcal{T}_u 中提取的用户-POI类别点对 $\Delta_p = [\langle u, c_1 \rangle, \langle u, c_2 \rangle, \dots, \langle u, c_K \rangle]$ 作为正样本，同时从用户未访问过的类别标签中随机抽取K个点对 $\Delta_n = [\langle u, c'_1 \rangle, \langle u, c'_2 \rangle, \dots, \langle u, c'_K \rangle]$ 为负样本。根据贝叶斯个性化排序规则(BPR^[74])， $s_{\langle u, c'_* \rangle}$ 的值要远小于 $s_{\langle u, c_* \rangle}$ 的值，则User-Mem的目标函数可以形式化为：

$$\mathcal{L}_{mem}(u) = \sum_{\langle u, c \rangle \in \Delta_p} \sum_{\langle u, c' \rangle \in \Delta_n} \max(0, s_{\langle u, c' \rangle} - s_{\langle u, c \rangle} + \eta) \quad (4.6)$$

其中， η 为分离正负样本的决策边界距离。

4.2.4 用户偏好动态演化网络

该网络基于神经常微分方程的序列编码网络进行设计，根据每个用户的签到序列生成动态偏好表征向量。普通的动态偏好建模方法均通过静态离散化的时间间隔处理方式，不能定义连续的动态过程，且不能接受任意时刻的输入数据，因此不能很好建模连续时间间隔的行为序列。为了改善上述的问题，本文使用了基于 GRU 的神经常微分方程，神经常微分方程在模型训练的过程中，不使用其他 RNN 以及 CNN 等神经网络直接参数化隐藏状态的处理方式，直接使用神经网络参数化隐藏状态的导数。因此，神经常微分方程的参数可以看成一个连续的空间，不再需要分层进行梯度传播并更新参数。

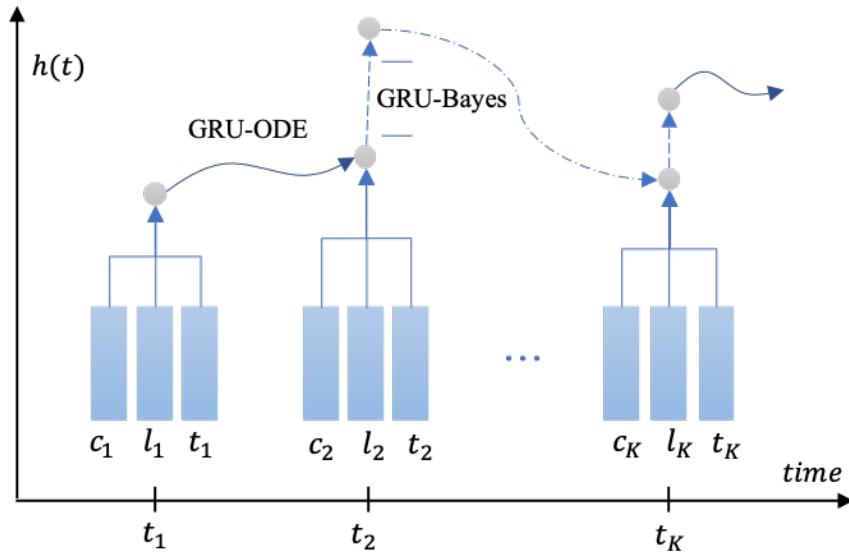


图 4.4 用户动态偏好演化网络结构

Figure 4.4 Architecture of user-preference dynamic evolution network

本文利用 Edward 等^[24]提出的 GRU-ODE-Bayes 思想，设计了面向位置社交网络中用户动态偏好建模的时空神经常微分方程（Spatial-Temporal Neural Ordinary Differential Equations, ST-ODE），学习获得用户的动态偏好表征向量。如图 4.4 所示，用户动态偏好演化网络依次读取用户 u 签到序列 $\mathcal{T}_u = [r_1, r_2, \dots, r_K]$ ，首先通过签到行为表征层获取签到行为表征向量 $[\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_K]$ ，然后经过 ST-ODE 进行偏好建模，输出用户动态偏好隐含状态 \mathbf{h}_k 。从架构图中可以看出，ST-ODE 包含三个主要部分：签到行为表征层、时空神经常微分方程和参数优化更新，下面分别阐述各个部分的方法。

签到行为表征层。在 ST-ODE 中，利用签到表征层获取签到行为表征向量。首先，对于用户 u 的一条签到记录 $r_k = (u, p_k, t_k)$ ，ST-ODE 关注 $[c_k, l_k, t_k]$ 三元组，表征方式如章节 4.2.2；然后，通过引入矩阵 $E_c \in \mathbb{R}^{F \times d_c}$ 、 $E_l \in \mathbb{R}^{N \times d_l}$ 以及 $E_t \in \mathbb{R}^{168 \times d_t}$ 来捕捉三类特征；最后，获得维度为 $d = d_c + d_l + d_t$ 的用户签到行为表征向量。

时空神经常微分方程。本文为了探索用户的动态偏好，借助于 GRU-ODE-Bayes 来探索用户的时空偏好。首先，将从表征层获得签到序列的表征向量作为输入；然后，ST-ODE 通过隐状态传播层（GRU-ODE，控制历史记录对用户隐状态的影响）和隐状态更新层（GRU-Bayes，通过部署一个贝叶斯更新进程来控制当前签到记录对用户隐状态的影响）来捕捉用户的动态偏好信息，最后输出用户动态偏好隐含状态表征向量。如图 4.4 所示，对于输入的用户表征向量 \mathbf{r}_k ，通过基于 GRU 的 ODE 层进行签到行为之间的连续时间状态传播(GRU-ODE)，即将隐状态 \mathbf{h}_{k-1} 从时刻 t_{k-1} 转移到时刻 t_k ，可以形式化为：

$$\mathbf{h}'_k = \text{GRU-ODE}(\mathbf{h}_{k-1}, t_{k-1}, t_k) \quad (4.7)$$

然后，利用隐状态更新层（GRU-bayes）来处理任意时刻输入的签到行为表征 \mathbf{r}_k 来更新隐含状态向量：

$$\mathbf{h}_k = \text{GRU-Bayes}(\mathbf{h}'_k, \mathbf{r}_k) \quad (4.8)$$

这里， \mathbf{h}'_k 和 \mathbf{h}_k 分别表示在根据签到记录 r_k 进行更新前后隐含状态的表征向量。

参数优化方法。为了优化 ST-ODE 模型参数，本文中定义了两部分损失函数：传播损失和更新损失。首先，传播损失来自于在更新前的 GRU-ODE 部分， \mathcal{L}_{ode} ，表示已经观察到的签到记录的负对数似然损失（Negative log-likelihood），即经过传播后的隐状态与观察值之间的分布差异。对于一条签到记录 r_k ， \mathcal{L}_{ode} 可以定义为：

$$\mathbf{p}_{pre} = \text{softmax}((f(\mathbf{h}'_k))) \quad (4.9)$$

$$\mathcal{L}_{ode}(r_k) = -\sum_{i=1}^N \mathbf{p}_k[i] \cdot \log(\mathbf{p}_{pre}[i]) \quad (4.10)$$

其中， f 表示全连接层将 \mathbf{h}'_k 映射为 N 维度的向量， \mathbf{p}_{pre} 表示隐状态向量推

断的该时刻在所有 POI 上的概率分布； $\mathbf{p}_k \in \mathbb{R}^{N \times 1}$ 表示 POI p_k 的 one-hot 表征向量。

其次，更新损失本节使用贝叶斯损失来表示，为了计算贝叶斯损失，假设当前观察到的 POI p_k 先验概率分布为 \mathbf{p}_{obs} ，通过模拟贝叶斯更新过程，获取在该观察点的用户隐状态 \mathbf{p}_{bayes} ，其计算过程可以形式化为：

$$\mathbf{p}_{bayes} \propto \mathbf{p}_{pre} \cdot \mathbf{p}_{obs} \quad (4.11)$$

使用 \mathbf{p}_{post} 表示在签到记录 r_k 发生时刻应用 GRU-Bayes 层更新之后预测的概率分布，即根据隐含状态表征向量 \mathbf{h}_k 预测的用户偏好概率分布。则更新损失 \mathcal{L}_{bayes} 可以定义为：

$$\mathcal{L}_{bayes(r_k)} = D_{KL}(\mathbf{p}_{bayes} || \mathbf{p}_{post}) \quad (4.12)$$

直观上，更新损失使更新前的分布与更新后的分布之间的差异不能太大。

最后，对于有 K 条签到记录的用户 u，ST-ODE 的目标函数 $\mathcal{L}_{st(u)}$ ，可以通过参数 λ 均衡两部分损失函数来定义，即：

$$\mathcal{L}_{st(u)} = \sum_{k=1}^K \mathcal{L}_{ode}(r_k) + \lambda \mathcal{L}_{bayes(r_k)} \quad (4.13)$$

4.2.5 用户动态偏好感知的兴趣点推荐

利用上述偏好记忆网络（User-Mem）和偏好动态演化网络（ST-ODE）获取的用户静态偏好表征向量和用户动态偏好表征向量，本文首先对两个向量进行整合，然后通过一层 softmax 获取预测的用户的 POI 分布：

$$\mathbf{h} = \text{concat}(\mathbf{h}_{\hat{t}}, \mathbf{e}_u) \quad (4.14)$$

$$\mathbf{O} = \text{softmax}(\mathbf{W}_r \mathbf{h} + \mathbf{b}_r) \quad (4.15)$$

其中， \hat{t} 表示某个时刻； $\mathbf{h}_{\hat{t}}$ 表示经过 ST-ODE 获得的用户在时刻 \hat{t} 的偏好隐含状态表征向量； \mathbf{u} 表示经过 User-Mem 获取的用户静态偏好表征向量； $\mathbf{O} \in \mathbb{R}^{N \times 1}$ 表示预测的用户在时刻 \hat{t} 可能访问的 POI 概率分布。

由于 POI 数量规模庞大，在参数更新过程中，直接优化在所有 POI 上的概率分布会变得很棘手并且效率很低。在模型训练过程中，假设用户 u 在时刻 \hat{t} 访问的 POI 为 p ，则在概率向量 \mathbf{O} 中对应的概率值为 y_p 。这里，我们通过负采样策略针对 POI p 随机生成负样本 p^- ，则目标函数可以定义为：

$$\mathcal{L}_{rec} = - \sum_{p \in P} \sum_{p^- \in P^-} \ln (y_{p^-} - y_p) \quad (4.16)$$

则参数优化过程可以形式化为:

$$\Theta \leftarrow \Theta - lr \nabla_{\Theta} \sum_{u \in U} (\mathcal{L}_{rec} + \beta \mathcal{L}_{mem} + \gamma \mathcal{L}_{st}) \quad (4.17)$$

其中, U 表示用户集合, Θ 表示模型 PINN 的参数集合。

4.2.6 复杂度分析

PINN 中, NPOI 推荐的计算过程如算法 4.1 所示, 计算开销主要存在于表征向量获取以及推荐排名两部分: 在表征向量获取方面, 主要包括 User-Mem 和 ST-ODE 两个子模块。针对 User-Mem 部分, 对于特定用户的历史偏好表征向量获取的复杂度为 $O(d^2)$ 。针对 ST-ODE 部分, 主循环在时序列表唯一的时间点上进行迭代, 在传播层 GRU-ODE 步骤中, 针对单个列表, 需要传播所有隐藏状态, 时间复杂度为 $O(|\emptyset_{GRU-ODE}|)$, 这里 $|\emptyset_{GRU-ODE}|$ 表示 GRU-ODE 模块的参数量, 更新层 GRU-Bayes 的更新和损失计算也是仅对在特定具有观测值的时间点执行, 时间复杂度为 $O(|\emptyset_{GRU-bayes}|)$, $|\emptyset_{GRU-bayes}|$ 表示 GRU-Bayes 模块的参数量; 在推荐排名方面, 时间复杂度为 $O(d^2)$, 这里 d 表示表征向量的维度为常量。因此针对特定用户, PINN 的整体时间复杂度为 $O(|\emptyset_{GRU-ODE}| + |\emptyset_{GRU-bayes}| + d^2)$, 这里传播层和更新层中的参数量均为常量, 因此 PINN 的计算开销与用户数量成线性关系, 可以应用于较大规模 NPOI 推荐任务中。

算法 4.1: PINN NPOI 推荐算法

输入:

签到序列: $\mathcal{T}_u = [r_1, r_2, \dots, r_K]$

签到行为三元组: $r_k = (u, v_k, t_k)$

输出:

给每个用户生成 NPOI 推荐列表

- 1: # 获取用户历史偏好表征
- 2: 根据公式 (4.1) 获取用户签到记录表征 r_k
- 3: 根据公式 (4.4) 以及历史签到记录获取历史偏好表征 e_u
- 4: # 获取用户动态偏好表征

```

5:   FOR k = 1 to K DO
6:     #状态传播: 从  $t_{k-1}$  到  $t_k$ 
7:     根据公式 (4.7) 计算用户动态偏好表征  $\mathbf{h}'_k$ 
8:     #状态更新: 在发生签到行为的时刻
9:     根据公式 (4.8) 计算更新后的动态偏好表征  $\mathbf{h}_k$ 
10:    END FOR
11:    #进行 NPOI 推荐
12:    根据公式 (4.14) 整合用户历史偏好表征和用户动态偏好表征
13:    根据用户表征和 POI 表征进行 NPOI 推荐
14:    RETURN 推荐列表

```

4.3 性能评估

基于六个城市的签到数据集，本文通过 NPOI 推荐任务评估用户动态偏好感知的 NPOI 推荐方法（PINN）的性能。

4.3.1 实验设置

本节首先介绍了用于评测 PINN 方法的数据集，然后介绍了对比方法及评价指标，最后介绍了实验中用到的参数设置。

4.3.1.1 实验数据

本文使用的六个城市的数据集均来自于位置社交网络平台 Foursquare。据统计，这个数据集包含 33,278,683 条用户签到数据，涵盖用户量为 266,909，POI 数量为 3,680,126（涉及全球 77 个国家中的 415 个城市）。在这 415 个城市中，每个城市至少包含 10k 以上的签到数据。本节列举了在实验过程中整理的六个城市的签到数据，如表 4.1 所示。

表 4.1 数据集统计数据

Table 4.1 Statistics of our datasets

数据集	用户数量	位置数量	签到数量
New York	17385	66660	581544
Los Angeles	8969	21556	162658
San Francisco	5455	11076	87906
Chicago	6870	21880	184453
Washington	6420	13828	131444
Tokyo	12613	88546	1100216

4.3.1.2 对比方法及评价指标

为了评估 PINN 的性能，将 PINN 与以下九种基线方法进行了对比分析：

传统推荐类方法：

- MF-BPR^[74]: 该算法提出了一种基于贝叶斯的改进矩阵分解模型，主要是通过贝叶斯的个性化排序来优化 MF，通过建模用户访问过的 POI 数据来建模用户与 POI 的隐式关系；
- CML^[41]: 该算法将度量学习(Metric learning)与协同过滤(CF)进行结合来解决用户物品推荐的问题，提出了协同度量学习模型，学习用户和物品之间的潜在关系的同时评估用户和用户以及物品和物品之间的相似性，然后再借助于用户的隐式评价进行 Top-K 推荐。

上下文感知类方法：

- PRME^[29]: 该算法使用针对推荐的排序度量嵌入方法对用户的个性化签到序列进行建模，此模型整合了签到的序列信息、个人偏好和地理影响，以提高推荐性能；
- STRNN^[55]: 该算法在循环神经网络的基础上增加了时间敏感性，捕捉在签到行为之间的时间和空间位置信息变化；
- DeepMove^[115]: 该算法在循环神经网络的基础上增加了用户兴趣注意力机制，考虑用户签到序列时序关系的基础上融合用户的长期兴趣；
- ST-LSTM^[112]: 该算法是长短期记忆网络(LSTM)的变式，在 LSTM 中增加了时间和距离两种门限机制来捕捉用户签到行为序列之间的时空关系来进行 POI 推荐；
- GRU^[14]: 该算法也是循环神经网络的一种，可以看成是 LSTM 变体，结构更加简单，主要解决了 RNN 网络中的长依赖问题。

PINN 变式：

- PINN_{mem-}: 该算法为 PINN 方法的弱化版本，不使用模型的用户偏好记忆网络；
- PINN_{RNN}: 该算法为 PINN 方法的变式，将用户动态偏好演化网络替换为 GRU 序列编码网络。

评价指标方面，本文采用命中率 (HR@N) 和归一化累计增益 (NDCG@N) 来评价 PINN 的 NPOI 推荐效果，指标详细描述见章节 3.3.1.2。

4.3.1.3 参数设置

实验采用 pytorch 的框架实现 PINN。对于 User-Mem，设置用户组的个数为 10，每个类别用户组的表征向量维度为 10。对于 ST-ODE，隐含状态的维度设置为 50，时间维度和 POI 位置维度分别设置为 48 和 40。在参数训练过程中，设定 dropout 的取值范围为 [0, 0.1, 0.2, 0.3]，weight decay 的取值范围为 [0.1, 0.03, 0.01, 0.003, 0.001, 0.0001, 0]，learning rate 的取值范围为 [0.001, 0.003, 0.0001]。本文通过 5 折交叉验证，每次保留验证数据集的 20% 对上述参数训练的 PINN 参数进行评估，模型表现最优时的 dropout 值为 0.2，weight decay 的值为 0.001，learning rate 的值为 0.003。

4.3.2 实验结果分析

本节分别从推荐命中率和归一化累计增益评估、实例可视化分析和参数敏感度分析三个方面分析了 PINN 模型的性能。

4.3.2.1 HR@N 和 NDCG@N

本实验在六个真实城市签到数据集上进行，不同方法得到的推荐结果如表 4.2 和表 4.3 所示（以旧金山和洛杉矶为例）。

表 4.2 各类方法在 Foursquare (旧金山) 数据集上的推荐结果统计

Table 4.2 Performance comparison for different methods on Foursquare (San Francisco)

Method	Foursquare (San Francisco)			
	HR@5	NDCG@5	HR@10	NDCG@10
MF-BPR	0.03456	0.02304	0.06283	0.04409
CML	0.05872	0.04912	0.14871	0.10594
PRME	0.04076	0.03247	0.10189	0.06719
ST-RNN	0.05006	0.03903	0.13649	0.08359
DeepMove	0.06012	0.04997	0.15096	0.10889
ST-LSTM	0.07043	0.05819	0.16762	0.12440
GRU	0.05118	0.04092	0.13765	0.08503
STA-RNN	0.07212	0.0603	0.17641	0.13450
PINN _{mem}	0.07401	0.06198	0.18529	0.14301

PINN _{RNN}	0.07980	0.07021	0.21865	0.16552
PINN	0.09092	0.07568	0.23125	0.17798

表 4.3 各类方法在 Foursquare (洛杉矶) 数据集上的推荐结果统计

Table 4.3 Performance comparison for different methods on Foursquare (Los Angeles)

Method	Foursquare (Los Angeles)			
	HR@5	NDCG@5	HR@10	NDCG@10
MF-BPR	0.03312	0.02801	0.06103	0.04334
CML	0.05541	0.04647	0.14276	0.10012
PRME	0.03972	0.03032	0.09535	0.06197
ST-RNN	0.04718	0.03827	0.12902	0.08143
DeepMove	0.05982	0.04921	0.14872	0.10727
ST-LSTM	0.06801	0.05654	0.16426	0.12026
GRU	0.04831	0.03972	0.13033	0.08209
STA-RNN	0.07104	0.05981	0.16824	0.12903
PINN _{mem}	0.07317	0.06507	0.18023	0.15095
PINN _{RNN}	0.07022	0.06432	0.17839	0.14097
PINN	0.08923	0.07221	0.22767	0.17503

通过上面实验结果可以看出与其他对比方法相比，本文提出的方法在 NPOI 推荐任务方面表现更佳。下面我们将对实验结果进行详细分析：

(1) 连续时间建模的优势 (Contribution of ST-ODE)。由实验结果可以看出，将签到行为序列看成等间隔进行建模的方法 (比如 STRNN 和 DeepMove) 在推荐效果上要略逊于时间敏感的推荐方法 (比如 ST-LSTM 和 ST-ODE)，可以看出签到行为的时序关系也是用户兴趣点推荐不可忽视的因素。实验结果表明，对比最好基线方法，PINN 在多个城市的数据集上取得了 7% 左右的性能提升，表明定义连续的动态过程，能够为用户生成有效的动态偏好表征，表明 PINN 能进一步提升了时间敏感类方法在 NPOI 推荐上的性能。

(2) 用户偏好记忆网络的贡献 (Contribution of User-Mem)。比较 PINN 与 PINN_{mem}，发现，PINN 在多种指标下均有 4% 以上的提升，表明用户的静态历史偏好，对我们理解用户的真实意图具有重要的意义，能够在一定程度上提升兴趣点推荐效果。

4.3.2.2 实例可视化分析

本文对不同城市的一些代表性用户进行了抽样，并分析了在 PINN 中学到的

知识。具体来说，可视化了城市不同用户组表征中用户的注意力权重，并比较了不同用户的喜好类别。在这里，本文从旧金山（SF）和洛杉矶（LA）随机选择 10 个用户，并在表 4.4 中显示了他们最喜欢的 10 个类别。注意， $\{u_1-u_4\}$ 是从 SF 采样的，而 $\{u_5-u_9\}$ 是从 LA 采样的。各个用户经过表征之后的表征向量在各个用户组表征上的关注权重热力图如图 4.5 所示，其中颜色的深浅表示关注权重的强度，每一行代表每个代表用户的关注得分向量。

表 4.4 所选用户访问频次最高的 10 个类别结果统计

Table 4.4 Top10 most frequently visited categories of selected users

UserID	Top-10 categories
u_0	[‘Sculpture Garden’, ‘General Travel’, ‘Clothing Store’, ‘Gourmet Shop’, ‘Tattoo Parlor’, ‘Drugstore / Pharmacy’, ‘Gift Shop’, ‘Residential Building (Apartment / Condo)’, ‘Coffee Shop’, ‘Doctor’s Office’]
u_1	[‘Auditorium’, ‘Bed & Breakfast’, ‘Indian Restaurant’, ‘Parking’, ‘Assisted Living’, ‘Taxi’, ‘Bank’, ‘Government Building’, ‘Plaza’, ‘Fast Food Restaurant’]
u_2	[‘Art Gallery’, ‘Pier’, ‘Chinese Restaurant’, ‘Theater’, ‘Park’, ‘Pizza Place’, ‘Building’, ‘Bar’, ‘Train Station’, ‘Coffee Shop’]
u_3	[‘Rest Area’, ‘Video Store’, ‘Light Rail’, ‘City’, ‘Hot Spring’, ‘Neighborhood’, ‘Skate Park’, ‘Road’, ‘Outdoors & Recreation’, ‘Optical Shop’]
u_4	[‘Donut Shop’, ‘Grocery Store’, ‘Building’, ‘Department Store’, ‘African Restaurant’, ‘Miscellaneous Shop’, ‘General Entertainment’, ‘Bowling Alley’, ‘Paper / Office Supplies Store’, ‘Fried Chicken Joint’]
u_5	[‘Lighthouse’, ‘Travel & Transport’, ‘Video Store’, ‘Scenic Lookout’, ‘Dog Run’, ‘General Travel’, ‘Convenience Store’, ‘Park’, ‘Bridge’, ‘Road’]
u_6	[‘Gym’, ‘Paella Restaurant’, ‘Baseball Field’, ‘Bakery’, ‘Latin American Restaurant’, ‘Sporting Goods Shop’, ‘Fish Market’, ‘Building’, ‘Furniture / Home Store’, ‘Juice Bar’]
u_7	[‘City’, ‘Greek Restaurant’, ‘Baseball Stadium’, ‘Burger Joint’, ‘Office’, ‘Bar’, ‘Train Station’, ‘Coffee Shop’, ‘Gym’, ‘Grocery Store’]
u_8	[‘Auditorium’, ‘Bed & Breakfast’, ‘Indian Restaurant’, ‘Parking’, ‘Assisted Living’, ‘Taxi’, ‘Bank’, ‘Government Building’, ‘Plaza’, ‘Fast Food Restaurant’]
u_9	[‘Art Gallery’, ‘Bookstore’, ‘Music Venue’, ‘Airport’, ‘Office’, ‘Bar’, ‘Train Station’, ‘Coffee Shop’, ‘Gym’, ‘Grocery Store’]

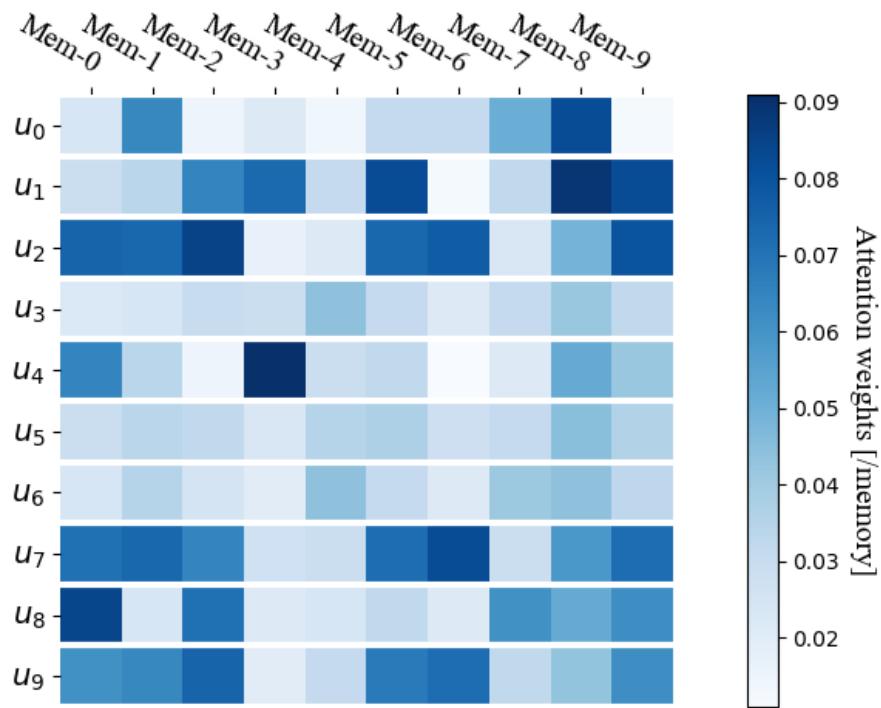


图 4.5 用户注意力权重热力图

Figure 4.5 The heatmap of attention weights of sampled users

如图 4.5 所示，将注意力权重与表 4.5 中所示的每个用户的语义进行比较，可以发现相似的用户具有相似的注意力权重分布。比如，用户 \$u_2\$，\$u_7\$ 和 \$u_9\$ 具有相似的注意力向量，它们的前 10 个类别都包括“酒吧”，“火车站”和“咖啡店”。相反，用户 \$u_1\$ 的注意力权重与其他用户不同，表明的偏好不同于其他用户。尽管用户是从不同的城市抽样的，具有相似喜欢类别的用户也同样具有相似的用户偏好风格，例如 \$u_2\$ 和 \$u_9\$，\$u_3\$ 和 \$u_6\$。这表明不同城市之间的用户群体相对相似。综上所述，用户组嵌入可用于捕获用户的历史偏好信息。

4.3.2.4 参数敏感度分析

本实验测试了 PINN 的两个主要参数：(1) 记忆网络中用户组偏好向量表示维度 d ；(2) 训练过程中超参数 γ 。下面评估不同 d 值和不同超参数 γ 的值对 HR@5 指标的影响。本实验过程中，设置 d 的范围为 $[2, 4, 6, \dots, 18, 20]$ ， γ 的范围为 $[10^{-6}, 10^{-5}, \dots, 10^{-1}, 0.2, \dots, 0.5]$ ，实验结果如图 4.6 和图 4.7 所示：

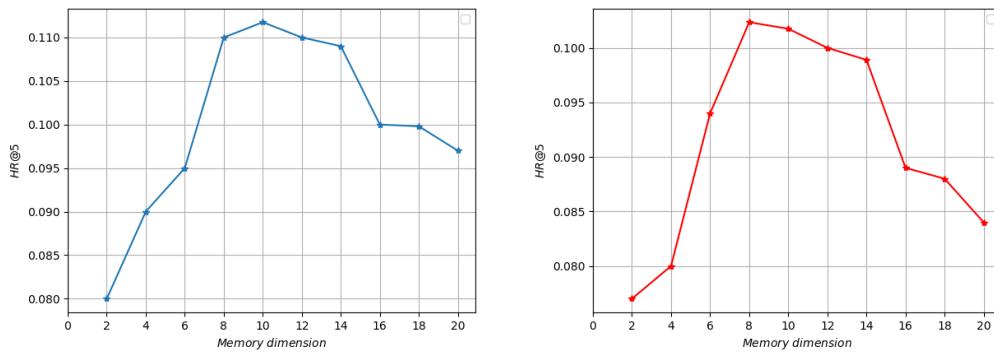
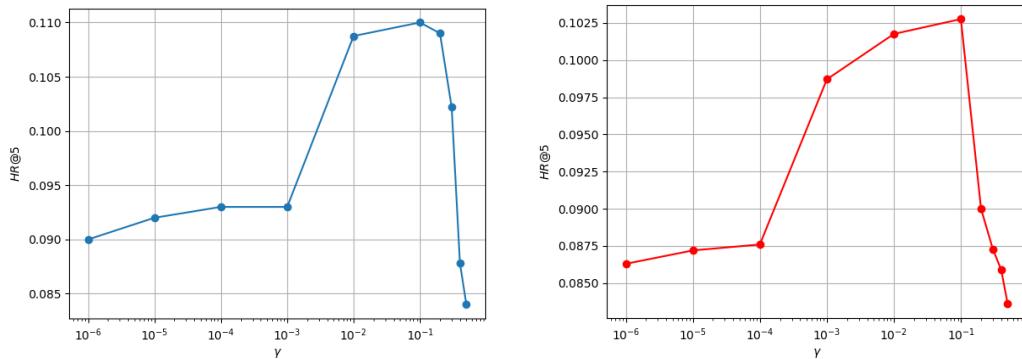


图 4.6 记忆网络中表征维度 d 对模型的影响 (SF & LA)

Figure 4.6 The influence of memory dimension d for PINN (SF & LA)

图 4.7 超参数 γ 对模型的影响 (SF & LA)Figure 4.7 The influence of the hyper-parameter γ for PINN (SF & LA)

记忆网络中表征维度 d 对模型的影响结果如图 4.6 所示。总的来说，随着 d 的增加，PINN 推荐性能首先会上升，随后出现下降。主要是因为在到达顶峰之前随着维度的上升，记忆网络对用户偏好模式的表达能力随着维度的上升而增强，之后出现下降主要是因为维度继续变大会导致训练模型出现过拟合，推荐效果下降。两个城市中呈现相似的变化趋势，PINN 均在维度值为 10 的时候达到了最佳效果，故在实验对比的过程中将记忆表征维度设置为 10。超参数 γ 对模型的影响结果如图 4.7 所示。PINN 推荐性能随着权重因子 γ 的逐渐变大，呈现上升趋势，随后出现下降。表明 User-Mem 和 ST-ODE 都能对 PINN 推荐性能产生积极影响，权重参数 γ 在达到 0.1 的时候，PINN 取得最佳推荐效果。在对比实验中也进行了相关设置。

4.4 小结

本文针对用户连续行为偏好刻画缺失导致推荐模型难以建模用户的连续状态变化的问题，提出了一种用户动态偏好感知的 NPOI 推荐方法 PINN。该方法的创新性主要体现在：(1) 提出了一种基于时空神经常微分方程的用户动态偏好感知方法，该方法首次将神经常微分方程用于 NPOI 推荐领域，定义连续的动态过程，生成用户的动态偏好表征，克服了大多数时序推荐模型静态离散化的时间间隔处理方式，更适用于 NPOI 推荐任务。(2) 提出了一种基于用户群体移动行为模式的用户历史偏好表征网络，该网络可以建模用户群体移动模式，并在群体移动模式的指导下获取用户的历史偏好表征。本文在六个真实城市的签到数据集上验证 PINN 的推荐效果。实验结果表明，与其他基线方法相比，PINN 在 NPOI 推荐任务上可以获得更高的命中率和归一化累计增益。

第5章 基于时空元学习的下一个兴趣点推荐方法研究

本章针对数据匮乏环境下，下一个兴趣点推荐模型训练困难，导致推荐模型难以对少样本用户进行有效拟合的问题，提出了基于时空元学习的下一个兴趣点推荐方法 MetaRNN（**Spatial-temporal Meta-learning Enhanced Recurrent Neural Network for NPOI recommendation**）。该方法旨在为 NPOI 推荐模型生成高鲁棒性的初始化参数，使得模型可以仅利用少量的训练样本实现对 NPOI 的精准推荐。MetaRNN 通过设计一种面向 NPOI 推荐的元学习框架，将元学习应用于具有多类时空数据的 NPOI 推荐任务。具体地，该元学习框架可在建模时空差异性的基础上，通过构建多任务学习具有强泛化能力的初始参数，提升少样本场景下的 NPOI 推荐性能。此外，本文设计了一种基于时空注意力机制 POI 序列编码网络，能够建模历史隐状态对当前状态的影响，保证在少样本场景下有效获取用户偏好表征。实验验证阶段利用五个真实的城市签到数据集验证了 MetaRNN 的推荐性能。实验结果显示，MetaRNN 在 NPOI 推荐任务上的效果要明显优于其他基线方法。

本章的组织如下：5.1 节详细阐述了本章节研究问题的内容和挑战。5.2 节详细介绍了基于时空元学习的下一个兴趣点推荐方法的主要内容，其中，5.2.1 节定义了数据匮乏环境下的下一个兴趣点推荐问题，并概述了 MetaRNN 的流程；5.2.2 节详细介绍了基于时空注意力机制的时序关系建模；5.2.3 节详细介绍了在元学习框架的引导下 MetaRNN 的优化过程。5.3 节通过命中率和累计损益等方面实验验证了 MetaRNN 算法在下一个兴趣点推荐任务上的效果。5.4 节对本章进行了小结。

5.1 引言

推荐的潜在的下一个兴趣点指的是为用户推荐在将来某个时间点其可能会访问的兴趣点。作为面向位置服务的移动社交网络数据挖掘领域的重要分支，POI 推荐可以预测用户行为和兴趣演化趋势，对城市交通的规划和商业模式的探索具有很重要的现实意义。POI 推荐有着广泛的应用，例如，大众点评网站

的餐厅推荐、旅游路线规划以及城市交通事件预警等。

当前大多数的 POI 推荐方法都假设 POI 数据是同分布的，将 POI 数据看成一个整体进行用户 POI 兴趣的挖掘工作，比如 STRNN^[55], DEEPMOVE^[115], ST-LSTM^[112]等。但是，这些方法将用户的签到 POI 数据看成一个整体，且并没有考虑轨迹数据的城市不均衡特性，缺乏对城市之间共性信息和城市本身特殊性信息进行区分的个性化建模，这样也就导致了在数据匮乏的城市，由于 POI 历史记录少，对用户进行 NPOI 推荐的难度大，且效率低。图 1 为 Foursquare 收集的截止到 2014 年 1 月的用户签到 POI 数据的全球位置映射分布图，其中涵盖用户数量为 2,733,324, POI 数量为 11,180,160, 签到条目数为 90,048,627。从图中可以看出受到采集手段、隐私保护等因素的影响，用户 POI 签到数据在城市之间分布极其不均衡。北美、欧洲、各大洲沿海发达城市 POI 数据分布比较集中，其他城市 POI 数据分布非常稀疏。

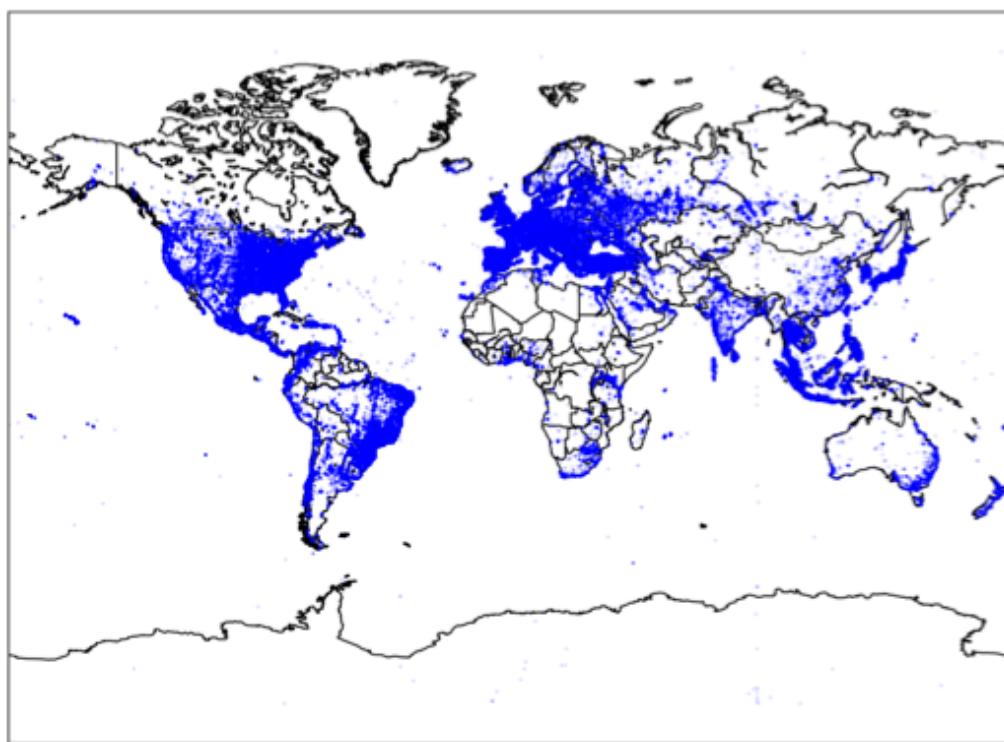


图 5.1 全球 Foursquare 用户签到 POI 映射图

Figure 5.1 Global Map of check-ins on Foursquare

因此，亟需探索如何在数据匮乏环境下进行 POI 推荐的有效方法。数据匮乏环境下的 NPOI 推荐的难点有两方面：第一个难点是如何对签到行为的时序关系

进行建模。因为用户签到序列存在大量的隐藏时序关系，因此如何在考虑时空特征的情况下进行时序关系探索极具挑战性；第二个难点在于不同城市的 POI 分布和城市结构不一致，导致若直接将学习到的模型参数应用到新的数据匮乏环境中，则会严重影响 NPOI 推荐效果。

目前，已有众多工作尝试通过知识迁移的方式进行数据匮乏城市的 POI 推荐。这些工作可以分为两大类：一类工作是直接将训练好的单域的兴趣点推荐模型，如 STRNN、ST-LSTM 等，直接迁移应用到目标城市。但是，由于没有区分源城市和目标城市结构差异，直接进行跨城市推荐效果不佳；其次，有少量的工作通过利用多域之间的共同用户作为锚链接，进行跨域兴趣点推荐，比如 JIM^[104]，JFT^[93]。这些工作都是基于源城市和目标城市之间存在共同用户的前提下，将关联用户在源城市的行为信息进行迁移，完成关联用户在目标城市的 POI 兴趣发现和推荐任务。但是，这类方法受限于共同用户这一假设制约，导致推荐方法的应用范围受限，不能够很好地进行非共同用户的 NPOI 推荐。

为了缓解数据匮乏给 NPOI 推荐带来的冷启动问题，本文借助于元学习的思想，提出了基于元学习的 NPOI 推荐方法 MetaRNN。该方法是一种知识迁移学习^[64]算法，通过对数据丰富城市的泛化共性移动知识进行学习，并将其迁移应用到数据匮乏城市的 NPOI 推荐任务中。具体地，MetaRNN 首先通过基于时空注意力机制 ST-ATT 的时序感知进行签到行为时序关系建模，然后通过多个源城市来构建多任务进行初始参数的更新。总的来说，MetaRNN 利用元学习思想进行移动行为知识的迁移，在建模时空差异性的基础上，通过构建多任务学习具有强泛化能力的初始参数，来缓解数据匮乏区域的推荐模型的冷启动问题。MetaRNN 的训练过程包括数据表征、时序关系建模和参数优化三个阶段：

在数据表征阶段，MetaRNN 从签到数据中进行特征抽取，并获得签到记录的表征向量；在时序关系建模阶段，MetaRNN 利用 RNN 将变长的用户签到行为序列编码为固定长度的表征向量来建模用户偏好。然而传统的 RNN 及其变体（GRU，LSTM）虽然能建模签到记录内部前后签到行为间的序列关系，但是在建模长序列的过程中容易出现梯度弥散，导致模型过多依赖短期行为偏好，缺乏对历史偏好信息的捕捉。为了解决这一问题，本文提出了基于时空注意力机制的

序列编码网络，该网络通过增加一个外部记忆层（External Memory Layer）来存储已经处理过的用户历史偏好信息，并利用时空注意力机制来捕获下一次用户签到行为与过去签到行为之间的关联，来有效地描述用户的兴趣偏好；在参数优化阶段，MetaRNN 利用多个数据丰富的源城市的签到数据来优化模型参数。该阶段的难点在于如何在签到数据分布不一致的城市之间进行移动行为知识迁移。为了解决这一难题，本文提出了一种面向 NPOI 推荐的元学习框架，在建模时空差异性的基础上，通过构建多任务学习具有强泛化能力的初始参数，提升少样本场景下的 NPOI 推荐性能。

5.2 基于时空元学习的下一个兴趣点推荐方法

本节详细介绍了基于时空元学习的下一个兴趣点推荐方法 MetaRNN。首先定义了数据匮乏环境下的下一个兴趣点推荐问题，并概述了 MetaRNN 的主要流程，然后详细阐述了 MetaRNN 各个步骤的计算过程和模型结构，最后分析了 MetaRNN 推荐算法的复杂度。

5.2.1 问题定义及方法概述

本节首先定义了数据匮乏下的下一个兴趣点推荐问题，即跨城市 NPOI 推荐，然后简要介绍了本文提出的基于元学习的 NPOI 推荐方法 MetaRNN。

5.2.1.1 问题定义

存在用户签到数据丰富的城市集合 $S = \{S_1, S_2, \dots, S_n\}$ 和用户签到数据匮乏的城市集合 $D = \{D_1, D_2, \dots, D_m\}$ ，签到数据包括用户集合 U 、签到序列集合 \mathcal{T} 、兴趣点集合 P 和语义标签集合 C ，这里，针对每个目标用户 u 都对应一条历史签到记录序列 $\mathcal{T}_u = [r_1, \dots, r_K]$ ，其中每条签到记录 $r_k = (l_k, c_k, t_k)$ 是一个三元组，包含了地理位置 l_k 、兴趣点的标签类别 c_k 以及时间戳 t_k 。根据上述描述，给出数据匮乏环境下的下一个兴趣点推荐的定义：

数据匮乏环境下的 NPOI 推荐：给定签到数据丰富的城市集合 $S = \{S_1, S_2, \dots, S_n\}$ 和用户签到数据匮乏的城市集合 $D = \{D_1, D_2, \dots, D_m\}$ ，满足：

$$|\mathcal{T}^{U_{S_i}}| \gg |\mathcal{T}^{U_{D_j}}|, i \in 1, 2, \dots, n, j \in 1, 2, \dots, m$$

我们的目标是为数据匮乏城市的目标城市 D_j 中的用户 u 推荐其下一个可能感兴趣的 POI p_{K+1} , 这里我们可以将推荐过程定义为:

$$p_{K+1} = \operatorname{argmax}_{p_{K+1}} f(p_{K+1} | p_{t_1}, p_{t_2}, \dots, p_{t_K}, \theta)$$

在上式中, f 表示 NPOI 推荐模型, θ 表示模型相关参数。模型的整体框架图如图 5.2 所示。

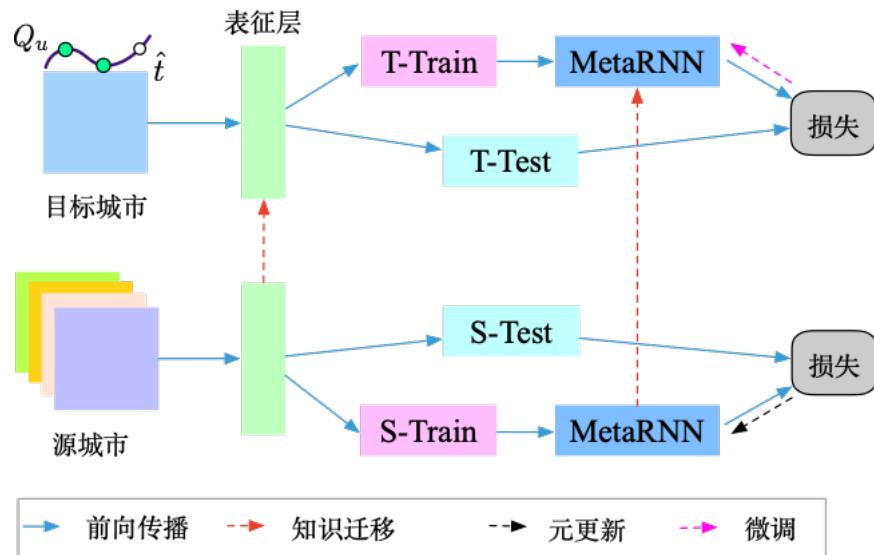


图 5.2 MetaRNN示意图

Figure 5.2 Architecture of MetaRNN

5.2.1.2 方法概述

MetaRNN 是一个基于元学习的 NPOI 推荐方法。它首先通过一个表征层进行签到记录的时空特征抽取, 形成固定维度的表征向量。以表征向量为输入, MetaRNN 学习一个签到序列编码网络。该网络将任意长度的签到序列编码为固定长度的向量, 并将其作为目标用户的移动行为偏好。具体来说, 对于任意一条签到序列 $T_u = [r_1, \dots, r_K]$, MetaRNN 会将其投影到 d 维表征向量, 然后根据表征向量进行 NPOI 推荐任务。图 5.2 展示了 MetaRNN 的架构, 可以看出 MetaRNN 主要由两部分组成: 基于时空注意力的签到行为编码网络和基于元学习的参数优化。

基于时空注意力的签到行为编码网络。 MetaRNN 利用 RNN 对轨迹进行建模, 并将 RNN 的最后隐含状态作为轨迹的表征向量。然而, 传统 RNN 及其变体

(GRU, LSTM) 仅能独立地对每个序列前后之间的关系进行建模，不能区分不同历史签到记录对最终表征向量的影响。在 MetaRNN 中，通过引入时空注意力来控制不同隐状态对最终表征向量的影响。它采用空间记忆张量来存储已经处理过的签到记录的隐状态，使得 RNN 能够克服梯度弥散的问题，更好地利用历史签到信息。

基于时空元学习的参数优化。 MetaRNN 利用元学习框架，指导编码网络参数的优化。优化 MetaRNN 时，通过对数据丰富区域的源城市构建多任务，根据多任务来维护更新 MetaRNN 的初始参数，并进行不同城市间的初始参数共享，这意味着所有城市的初始参数都是相同的，这种处理方式会在一定程度上增强模型初始参数的泛化能力，让其在新的数据匮乏城市能够更好更快地收敛到最佳状态。

5.2.2 基于时空注意力的签到序列编码网络

本节介绍了基于时空注意力的签到行为编码网络，STA-RNN，时空注意力可以增强传统 RNN 模型时序表征能力，使其具备在缓解梯度弥散问题的基础上建模长序列签到行为的能力。下面首先介绍编码网络的输入表征层，然后介绍注意力增强的签到行为编码网络 STA-RNN。

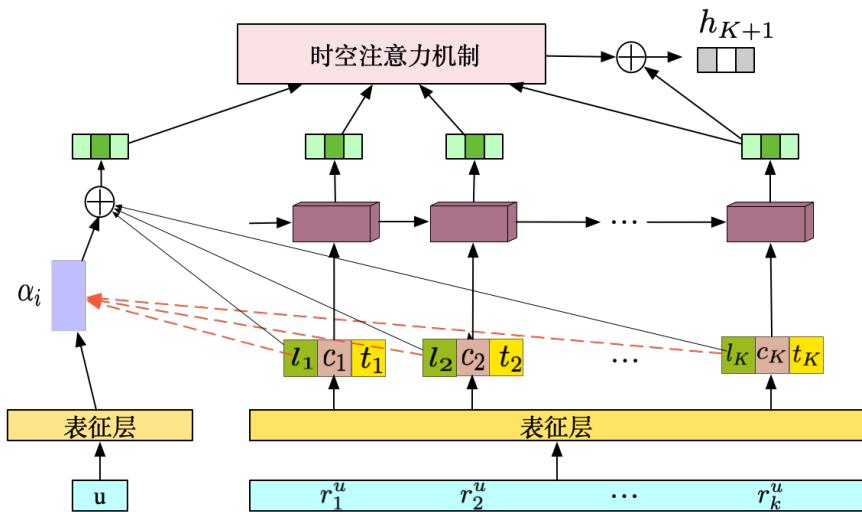


图 5.3 基于时空注意力的编码器 (STA-RNN) 机制示意图

Figure 5.3 Illustration of the spatio-temporal attention based encoder (STA-RNN)

5.2.2.1 签到行为表征

在 POI 表征中，首先，利用一个嵌入层获取签到行为序列中每一个签到行为的表征向量。如图 5.3 所示，在特征层，仅考虑 POI 的类别、POI 位置以及时间戳三类特征，即 (c_k, l_k, t_k) ，这里 c_k 表示该签到行为所在 POI 的所属类别， l_k 表示该 POI 的所在位置，由经纬度坐标表示， t_k 表示发生该签到行为的时间点。

(1) 针对 POI 类别属性， c_k ，引入矩阵 $\mathbf{E}_c \in \mathbb{R}^{N \times d_c}$ ，在这里， N 表示 POI 的类别数， d_c 表示 POI 类别表征向量的维度。在这里，矩阵 \mathbf{E}_c 是在多个城市之间进行共享。

(2) 针对 POI 地理位置属性， l_k ，引入矩阵 $\mathbf{E}_l \in \mathbb{R}^{M \times d_l}$ ，在这里， M 表示 POI 的个数， d_l 表示 POI 位置属性表征向量的维度。在这里，通过经纬度坐标位置的 GeoHash 编码来初始化 POI 矩阵 \mathbf{E}_l 。

(3) 针对时间戳， t_k ，引入矩阵 $\mathbf{E}_t \in \mathbb{R}^{168 \times d_t}$ ，在这里，168 表示将一周分成 168 个时间槽 (7*24 小时)，每一个时间槽的表征向量维度为 d_t 。

然后，针对每一条签到记录，将三类特征的表征向量进行聚合，可以获得长度为 d 的表征向量 \mathbf{r}_k ，作为后续时空神经常微分方程的输入，这里聚合过程可以形式化为：

$$\mathbf{r}_k = \varphi(\mathbf{W}_l \mathbf{e}_l^k + \mathbf{W}_c \mathbf{e}_c^k + \mathbf{W}_t \mathbf{e}_t^k + \mathbf{b}) \quad (5.1)$$

其中， \mathbf{e}_l^k 、 \mathbf{e}_c^k 和 \mathbf{e}_t^k 分别表示签到记录 r_k 中 POI 位置、POI 的类别以及时间戳三类特征的表征向量， \mathbf{W}_l 、 \mathbf{W}_c 和 \mathbf{W}_t 表示相关特征类别的权重矩阵， \mathbf{b} 为偏置参数。

为获取用户表征向量，本文采用了隐注意力机制。首先根据用户的表征向量，进行不同签到记录的注意权重的获取：

$$\alpha_k = \frac{\exp(\mathbf{u}^T \mathbf{r}_k)}{\sum_i \exp(\mathbf{u}^T \mathbf{r}_k)} \quad (5.2)$$

然后，我们根据不同签到记录的权重值获取用户的最终表征向量：

$$\mathbf{e}_u = \sum_{k=1}^K \alpha_k [\mathbf{e}_l^k || \mathbf{e}_c^k || \mathbf{e}_t^k] \quad (5.3)$$

在这里， $[\mathbf{e}_i^l || \mathbf{e}_i^c || \mathbf{e}_i^t]$ 表示将三元组 (l, c, t) 表征向量进行拼接， \mathbf{e}_u 表示用户最终的表征向量。

5.2.2.2 时序关系建模

图 5.3 展示了时序关系建模的单元架构。通过 GRU 进行签到序列中时序关系建模，捕捉用户的兴趣偏好。如图所示，在每个 GRU 循环步骤 k 中，GRU 单元以 \mathbf{r}_k 和前一步骤的隐含状态 \mathbf{h}_{k-1} 作为输入，输出 \mathbf{h}_k ，并进行下一个循环。其计算过程为：

$$\mathbf{u}_k = \sigma(\mathbf{W}^u \cdot \mathbf{r}_k + \mathbf{U}^u \cdot \mathbf{h}_{k-1} + \mathbf{b}^u) \quad (5.4)$$

$$\mathbf{z}_k = \sigma(\mathbf{W}^r \cdot \mathbf{r}_k + \mathbf{U}^r \cdot \mathbf{h}_{k-1} + \mathbf{b}^r) \quad (5.5)$$

$$\widetilde{\mathbf{h}}_k = \tanh(\mathbf{W}^h \mathbf{r}_k + \mathbf{z}_k \circ \mathbf{U}^h \mathbf{h}_{k-1} + \mathbf{b}^h) \quad (5.6)$$

$$\mathbf{h}_k = (\mathbf{1} - \mathbf{u}_k) \circ \mathbf{h}_{k-1} + \mathbf{u}_k \circ \widetilde{\mathbf{h}}_k \quad (5.7)$$

其中， $\mathbf{W} \in \mathbb{R}^{d \times d}$ ， $\mathbf{U} \in \mathbb{R}^{d \times d}$ ， $\mathbf{b} \in \mathbb{R}^{d \times 1}$ ， d 是隐含状态的维度值。所有的门向量 $(\mathbf{u}_k, \mathbf{z}_k)$ 和隐含状态 $(\mathbf{h}_k, \mathbf{h}_{k-1})$ 有着相同的维度： $\mathbb{R}^{d \times 1}$ 。 \mathbf{r}_k 为 GRU 单元的输入，表示第 k 条签到记录表征向量。

为了获得隐含状态 \mathbf{h}_k ，该单元执行以下操作步骤：

- 更新门。如公式 5.4 所示，其中 \mathbf{r}_k 为第 k 个时刻的输入向量，即输入序列的第 k 条签到记录对应的分量。然后它通过与权重矩阵 \mathbf{W}^u 相乘来进行线性变换。 \mathbf{h}_{k-1} 保存的是前一个时刻 t_{k-1} 的隐含状态，同样也经过了线性变换。最后更新门，将两部分信息以及偏置向量相加之后输入 sigmoid 激活函数 σ 获得更新门的激活结果；
- 重置门。如公式 5.5 所示，该单元主要决定了信息遗忘的权重，且其处理方式与更新门类似，将整合后的信息经过 sigmoid 激活函数 σ 来作为关系遗忘权重；
- 当前隐状态操作。如公式 5.6 所示，该单元使用重置门存储过去相关的信息。重置门是由范围为 0~1 的值组成的向量，衡量门控开启的大小，例如，某个元素对应的重置门控值为 0，则表示这个元素的信息被完全遗忘掉；
- 最终隐含状态操作。如公式 5.7 所示，该单元产生 \mathbf{h}_k ，并将其输出到下一个循环步骤。在这个过程中，通过更新门 \mathbf{u}_k 决定当前记忆内容 $\widetilde{\mathbf{h}}_k$ 和前一时刻的记忆内容 \mathbf{h}_{k-1} 中需要收集的信息内容权重。

5.2.2.3 时空注意力

签到行为编码网络通过时空注意力来控制历史隐状态对当前状态的影响权

重。从直觉上来说，并非所有的历史签到行为都与用户的下一步行为密切相关，即需要更加注意用户信息偏好。但是，标准的 RNN 网络无法检测其输入的哪一部分对 NPOI 推荐更加重要。因此本章节引入时空注意力，并提出了一种基于时空注意力的 RNN 网络(Spatial-Temporal Attention Recurrent Neural Network, STA-RNN)来解决上述问题。时空注意力旨在捕获下一步用户行为与过去签到行为之间的不同关联度。通过使用该时空注意力，STA-RNN 还可以帮助选择代表用户喜好的代表性签到行为，并为他们分配不同的权重。这些都使 STA-RNN 可以更好地集成用户的历史签到行为的表征，可有效地描述用户的兴趣偏好。

如图 5.3 所示，我们首先引入矩阵 $H \in \mathbb{R}^{d \times K}$ ，它包含 STA-RNN 在所有评估点的隐状态向量 $\{\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_K\}$ ，这里 d 表示隐状态向量的维度， k 表示用户签到序列的长度。STA-RNN 借助于注意力权重向量 $\boldsymbol{\alpha}$ 来进行历史轨迹隐状态向量 $\{\mathbf{h}_k\}$ 的整合：

$$\hat{\mathbf{h}}_{K+1} = \sum_{k=1}^K \boldsymbol{\alpha}_k \mathbf{h}_k \quad (5.8)$$

在这里，对于第 k 条签到记录， $\boldsymbol{\alpha}_k$ 表示第 k 条记录与下一步的签到行为的匹配程度，可以通过下式进行权重矩阵的计算：

$$\boldsymbol{\alpha}_k = \frac{\exp(g(\mathbf{h}_k, \mathbf{e}_u))}{\sum_{k=1}^K \exp(g(\mathbf{h}_k, \mathbf{e}_u))} \quad (5.9)$$

式中， $g(\mathbf{h}_{t_k^+}, \mathbf{e}_u)$ 为注意力函数。另外，在本文中使用点积注意力来进行注意力的计算：

$$g(\mathbf{h}_{t_k^+}, \mathbf{e}_u) = \frac{\mathbf{h}_{t_k^+}(\mathbf{e}_u)^T}{\sqrt{d}} \quad (5.10)$$

式中， \mathbf{e}_u 表示用户 u 的上下文向量。在本文中，首先根据用户访问过的 POI 的类别(category)信息进行 \mathbf{e}_u 的初始化， $\mathbf{e}_u = \sum_{k=1}^K \mathbf{w}_k \mathbf{e}_{k,c}$ ， \mathbf{w}_i 表示类别权重， $\mathbf{e}_{k,c}$ 表示在第 k 条签到行为所在 POI 的类别表征向量，在训练过程中，将其和参数进行同步更新。将 STA-RNN 模型整体算法总结如下：

算法 1 STA-RNN：基于时空注意力机制的签到序列编码网络

输入：

签到序列: $\mathcal{T}_u = [r_1, \dots, r_K]$

签到行为三元组: $r_k = (l_k, c_k, t_k)$

输出:

用户隐状态表征: \mathbf{h}_k

损失值: $loss$

INITIALIZE $time = 0, loss = 0$, final time T .

FOR $k = 1$ to K DO:

#状态表征: 获取 \mathbf{r}_k

根据公式(5.1)获取用户在 t_k 的表征向量

#当前状态更新: 从 t_{k-1} 到 t_k

根据公式(5.6)计算当前隐含状态

根据公式(5.7)计算向后传播时刻 t_k 的隐含状态

END FOR

#用户历史签到隐含状态整合

根据公式(5.8)进行历史隐状态向量整合

RETURN (\mathbf{h}_k)

5.2.3 基于元学习的参数优化

在本节中, 首先概述了 STA-RNN 网络的目标函数, 然后阐述了元学习框架 MetaRNN 下参数学习过程。

5.2.3.1 STA-RNN 目标函数

图 5.3 展示了 STA-RNN 方法, 该方法使用 RNN 将变长签到行为序列映射到一个 d 维向量, 并使时空注意力进行用户历史签到偏好的捕捉, 最终获得用户的最终隐状态向量 \mathbf{h}_{k+1} 。可通过下式计算用户下一时刻会访问某候选 POI v_i 的概率:

$$\mathbf{y}_{p_i} = (\mathbf{h}_{k+1})^T f(c_i, l_i, t_{k+1}) \quad (5.11)$$

这里, $f()$ 表示表征函数, 获取 POI 的相关特征的表征向量, \mathbf{y}_{p_i} 表示用户在 POI p_i 上的概率值。

在推荐过程中，通过 $\text{argmax}_p(\mathbf{y})$ 为用户推荐最有可能访问的POI。在训练过程中，参照BPR-Loss的个性化排名损失训练方式^[74]进行推荐任务损失函数的构建：

$$\mathcal{L}_{rec} = - \sum_{p^+} \sum_{p^-} \ln \sigma (\mathbf{y}_{p^-} - \mathbf{y}_{p^+}) \quad (5.12)$$

其中， p^+ 表示用户访问过的POI， p^- 表示到目前为止用户没有访问过的POI， σ 表示非线性函数，即 $\sigma(x) = 1/(1 + e^{-x})$ 。因此在目标城市 D_j 的参数训练过程中，损失函数可以重新定义为：

$$\mathcal{L}_{D_j} = \mathcal{L}_{rec} + \mu ||\emptyset|| \quad (5.13)$$

5.2.3.2 参数优化

城市的移动行为模式会随着时间的推移而发生变化，并且城市与城市之间的具体情况也各不相同。比如，在旧金山人们可能更习惯于打卡高科技企业或者著名的金门大桥，而在拉斯维加斯人们可能更喜欢打卡赌场或者红石公园。但同时，不同城市之间的用户也存在着诸多的共性，比如大家在上午都习惯于打卡咖啡店，在打卡完运动场之后可能更喜欢打卡餐厅。因此，可将来自于多个城市的共性行为知识，比如空间邻近性和时间依赖性等，迁移到目标城市。这也意味着需要在城市之间数据分布不同以及结构不一致情况下进行兴趣点推荐模型的更新。

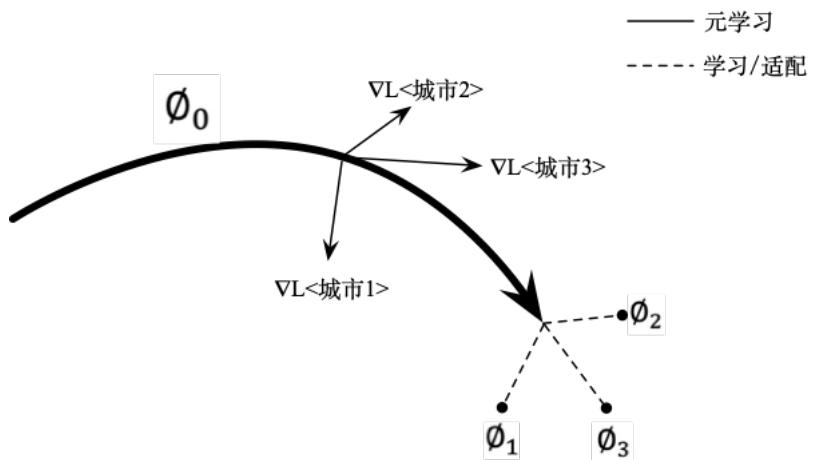


图 5.4 基于元学习的参数学习过程框架示意图

Figure 5.4 Illustration of the learning process of meta learning based methods

为了解决上述问题，受 Chelsea^[17]提出的域间自适应机制（MAML）的启发，

本章节提出了基于元学习的参数优化方法 MetaRNN。它借助于多个数据丰富的源城市来辅助 NPOI 推荐模型在目标城市的参数更新。使用 \emptyset 表示基于注意力机制的序列编码器 (STA-RNN) 的参数集合，因此 \emptyset 的相关参数中便隐含了时空移动知识。所以只需要传递更新参数 \emptyset 便可以实现共性行为知识的迁移。这里沿用元学习的思想来进行 STA-RNN 的参数学习，针对不同的源城市，MetaRNN 构建了多任务来共同更新一套初始参数，损失函数沿用公式(5.13)，在一次迭代过程中可以通过下式进行初始参数的更新：

$$\emptyset_0 = \min_{\emptyset_0} \sum_{S_i \in S} \mathcal{L}(f_{\emptyset_0} - \lambda \nabla_{\emptyset} \mathcal{L}_{S_i}(f_{\emptyset})) \quad (5.14)$$

在上式中， \emptyset_0 为 STA-RNN 的参数初始值， $\mathcal{L}_{S_i}(f_{\emptyset})$ 表示 STA-RNN 在源城市数据集 S_i 上的训练损失值。如在框架图中描述的，在每个源城市中，将会迭代地更新参数，比如在数据集 S_i 中，按照下式进行参数更新：

$$\emptyset_{S_i} = \emptyset_0 - \lambda \nabla_{\emptyset} \mathcal{L}_{S_i}(f_{\emptyset}) \quad (5.15)$$

在实际的更新参数过程中，会根据情况进行多次迭代，并最终获得具有普适性的初始参数。参数优化的过程如下所示：

算法 2 MetaRNN：基于元学习的参数优化

输入：

源城市集合： S_1, S_2, \dots, S_n

目标城市集合： D_1, D_2, \dots, D_m

输出：

适配于相关目标城市的元参数： \emptyset_0

NPOI 推荐列表

INITIALIZE \emptyset_0

WHILE model is not convergent DO:

FOR city S_i in $\{S_1, S_2, \dots, S_n\}$ DO:

 根据公式(5.13)评估 $\mathcal{L}_{S_i}(f_{\emptyset_{S_i}})$

 根据公式(5.14)更新参数 \emptyset_{S_i}

END FOR

```
根据公式(5.15)更新参数 $\phi_0$ 
END WHILE
FOR city  $D_j$  in  $\{D_1, D_2, \dots, D_m\}$  DO:
    根据公式(5.15)更新参数 $\phi_{D_j}$ 
    根据训练的模型完成推荐任务
END FOR
```

5.2.3.3 复杂度分析

利用 MetaRNN 进行兴趣点推荐计算过程包括两部分：基于时空注意力的序列编码部分和基于元学习的参数更新部分。序列编码部分的计算开销与签到序列长度是线性关系。在一个循环步骤中，STA-RNN 单元处理签到序列中记录的过程仅与当前记录有关，是一个固定的常数。因此，轨迹编码部分的计算复杂度是线性的，可以视为 $O(d^2 N)$ ， d 表示表征向量的维度值。基于元学习的参数更新部分，可以视为多步 STA-RNN 单元的循环，与训练次数 N_{step} 成正比。因此，STA-RNN 的总计算复杂度也是线性的。

对于 MetaRNN 来说，基于元学习的初始参数更新部分仅需要在目标城市进行参数优化之前计算一次。在后续的 NPOI 推荐过程不需要重复进行，因此对于 NPOI 推荐模型来说，当进行新的目标用户的签到序列输入时，仅需要对新的签到序列进行编码，并基于用户偏好表征向量进行 NPOI 推荐。因此，计算开销与用户签到序列数量大小成线性关系，这使得 MetaRNN 适用于大型数据集。

5.3 性能评估

为了评估 MetaRNN，本文在五个真实城市签到数据集上进行 POI 推荐效果评估实验，用于评估 MetaRNN 在兴趣点推荐任务上的性能。本节首先详细介绍了实验的数据集及相关实验设置，然后展示并分析了各个模型的性能表现。

5.3.1 实验设置

实验设置部分主要包括数据集、对比方法、评估及参数设置，下面进行详细介绍。

5.3.1.1 实验数据

本文使用的六个城市签到数据集均采自于位置社交网络 Foursquare。在实验过程中整理的城市签到数据样例，如表 5.1 所示。图 5.5 展示了数据集统计表中东京(Tokyo)、旧金山(San Francisco)等六个城市签到数据分布热力图，可以看出城市间之间 POI 分布极不均衡：

表 5.1 数据集统计数据

Table 5.1 Statistics of our datasets

数据集	用户数量	位置数量	签到数量	平均数
New York	17385	66660	581544	33.45
Los Angeles	8969	21556	162658	18.14
San Francisco	5455	11076	87906	16.11
Chicago	6870	21880	184453	26.87
Washington	6420	13828	131444	20.47
Tokyo	12613	88546	1100216	87.23

在实验评估过程中，按照用户的平均签到条目数来区分数据丰富城市和数据匮乏城市。在这里，分别指定洛杉矶(Los Angeles)和旧金山(San Francisco)为数据匮乏城市，其余城市划分为数据丰富城市。为了验证方法在数据匮乏区域的效果，在数据匮乏城市，只使用 1 个月的 POI 数据进行实验。除此之外，在实际评估过程中，按照 7:1:2 的比例划分训练集、验证集和测试集。

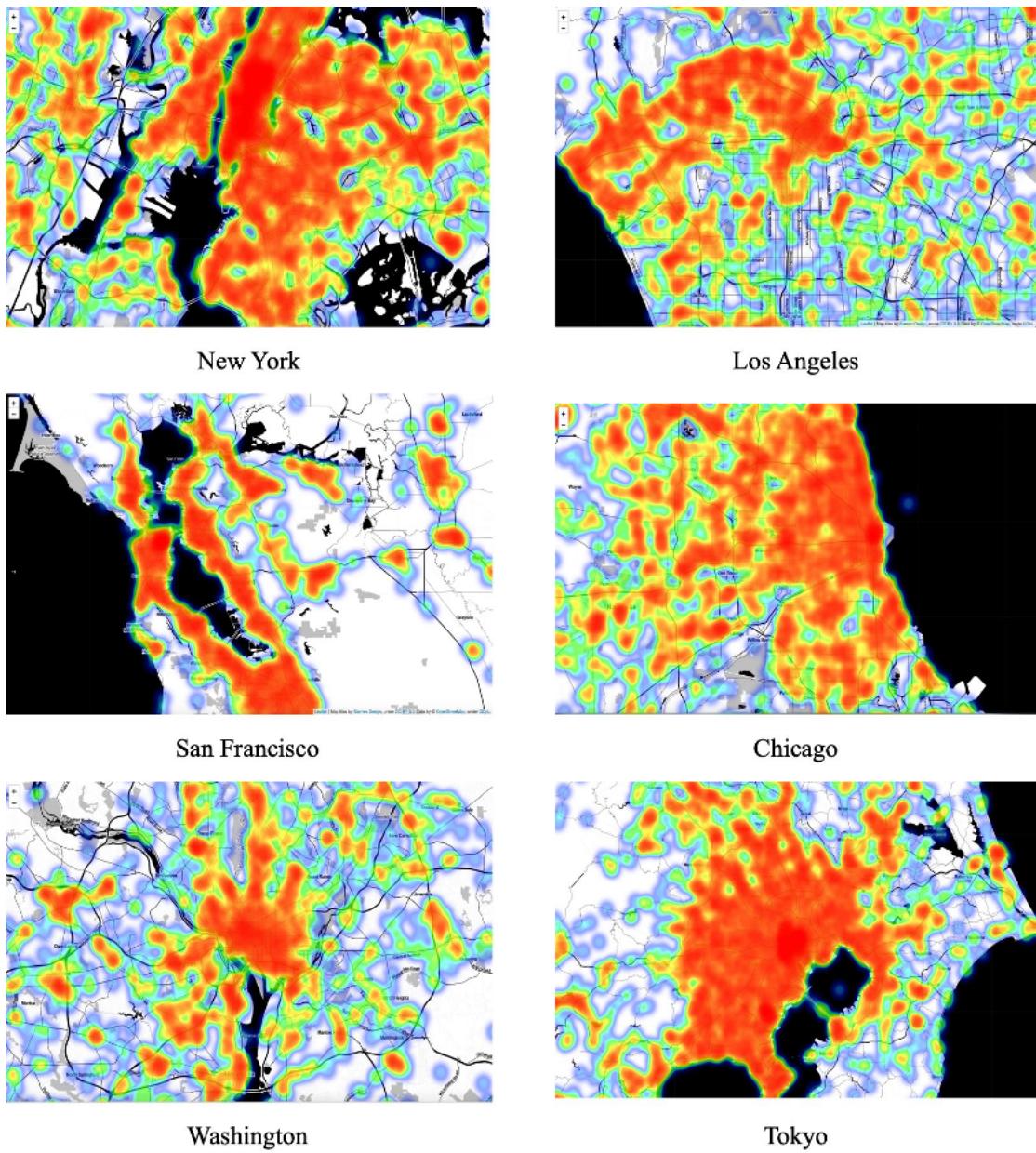


图 5.5 各数据集签到数据分布热力图

Figure 5.5 The heat maps of the check-ins in datasets

5.3.1.2 对比方法及评价指标

为了评估 MetaRNN 的 NPOI 推荐性能，本文将 MetaRNN 与当前九种基线方法进行了对比分析：

- **MF-BPR^[74]**: 该算法提出了一种基于贝叶斯的改进矩阵分解模型，且主要是通过贝叶斯的个性化排序来优化 MF，并通过建模用户访问过的 POI 数据来建模用户与 POI 的隐式关系；

- CML^[41]: 该算法将度量学习(Metric learning)与协同过滤(CF)进行结合来解决用户物品推荐的问题。其提出了协同度量学习模型，并在学习用户和物品之间的潜在关系的同时评估用户和用户，物品和物品之间的相似性，然后再借助于用户的隐式评价进行 Top-K 推荐；
- PRME^[29]: 该算法使用针对推荐的排序度量嵌入方法对用户的个性化签到序列进行建模。此模型整合了签到的序列信息、个人偏好和地理影响，以提高推荐性能；
- STRNN^[55]: 该算法在循环神经网络的基础上增加了时间敏感性，捕捉在签到行为之间的时间和空间位置信息变化；
- DeepMove^[115]: 该算法在循环神经网络的基础上增加了用户兴趣注意力机制，在考虑用户签到序列时序关系的基础上整合了用户的长期兴趣来进行下一个兴趣点推荐；
- ST-LSTM^[112]: 该算法是长短期记忆网络(LSTM)的变式，在 LSTM 中增加了时间和距离两种门限机制来捕捉用户签到行为序列之间的时空关系来进行 POI 推荐；
- GRU^[14]: 该算法也是循环神经网络的一种，可以看成是 LSTM 变体，结构更加简单，主要解决了 RNN 网络中的长依赖问题；
- STA-RNN: 该 A 算法是 MetaRNN 的弱化版本，在不使用元学习机制下，应用基于注意力机制的序列编码网络进行推荐，旨在比较分析元学习机制在该任务中是否有效。

评价指标方面，本章也采用命中率(HR@N)和归一化累计增益(NDCG@N)来评价 MetaRNN 的 NPOI 推荐效果，指标详细描述见章节 3.3.1.2。

5.3.1.3 参数设置

MetaRNN 中的关键参数包括：(1) 签到记录表征向量维数 d ；(2) 序列编码网络的学习率 (learning rate) r 。本实验通过在范围 $\{16, 32, 64, 128, 256\}$ 中进行网格搜索来确定表征向量维度 d 的值。通常，性能会先随着 d 值得变大而增加到达峰值之后逐渐稳定甚至呈下降趋势，发现在 $d=128$ 时，实验效果最佳。对于 r ，本实验在 $\{0.001, 0.002, 0.003, \dots, 0.01\}$ 中进行了搜索，发现在两个数据集上 $r=0.003$

时实验结果最佳。因此设置 $d=128$ 和 $r=0.003$ 。此外，将训练是 dropout 值设置为 0.2。小节 5.3.2.3 中详细阐述了参数敏感度分析的结果。

5.3.2 实验结果分析

本节从 HR@N 推荐命中率、NDCG@N 累计损益、参数敏感度分析、模型收敛性四个方面分析了 MetaRNN 的准确率和鲁棒性。

5.3.2.1 HR@N 和 NDCG@N 指标比较

表 5.2 和表 5.3 是各种对比方法在 HR@5、HR@10、NDCG@5 以及 NDCG@10 指标上的实验结果。可以看出本章节提出的方法与其他对比方法相比在 NPOI 推荐任务上表现更佳。下面对实验结果进行详细分析：

(1) 上下文信息的重要性 (Context-aware)。由实验结果可以看出利用了时空上下文信息的方法(比如 PRMF, STRNN, DeepMove, ST-LSTM 等)在推荐任务上要优于没有利用上下文信息的方法 (MF-BPR 和 CML)，这说明了上下文信息在推荐任务中起着至关重要的作用，可以有助于模型捕捉用户签到行为的时空上的序列关系，从而提升推荐效果。文章[92]也通过大量的实验证明了上下文信息对 POI 推荐的重要性。

(2) 神经网络的优势 (Neural networks)。由实验结果可以看出在基于矩阵分解的方法 (比如 MF-BPR 和 PRME) 在推荐效果上欠佳，而基于神经网络的方法 (比如 STRNN, GRU, DeepMove, ST-LSTM 等) 在推荐任务上的效果要更好。从中，也可以看出单纯的矩阵分解方法在捕捉用户的高阶关系上已经显得力不从心，而神经网络由于可以建模用户更高阶的潜在关系，所以推荐效果要优于传统的矩阵分解方法。

(3) 时空注意力的优势 (Spatial-temporal attention)。由实验结果可以看出，借助于时空注意力的 STA-RNN 在推荐效果上要优于不使用注意力机制的 GRU 方法。其中的主要原因在于时空注意力会对用户偏好的权重进行了更好地建模，并克服了在长时间序列传播过程中的知识损失，从而提升了 POI 推荐效果。

(4) 知识迁移的必要性 (Knowledge transfer)。本文针对知识迁移的必要性进行了实验证明。实验通过 MetaRNN 的几个进阶版本进行对比实验。STA-RNN

是仅利用目标城市的数据进行训练，即没有基于元学习机制进行参数的初始化。Single-FT 是利用单个源城市对目标城市的参数进行更新。由实验结果可以看出，将多个城市视为不同的推荐任务来借助元学习进行初始参数学习的方法 MetaRNN 具有更加优异的表现。

(5) 虽然 HR@10 在两个城市的性能提升只有 3%，相较于其他方法（比如 ST-LSTM）的较低的命中率（以 Foursquare（旧金山）为例，其命中率为 16.76%），3% 的性能提升不容小觑。这说明了元学习框架在知识迁移方面的优势，通过多任务的分解可以更好地学习模型的初始化参数，可以有效表明该方法的有效性。

表 5.2 各类方法在 Foursquare（旧金山）数据集上的推荐结果统计

Table 5.2 Performance comparison for different methods on Foursquare (San Francisco)

Method	Foursquare (San Francisco)			
	HR@5	NDCG@5	HR@10	NDCG@10
MF-BPR	0.03456	0.02304	0.06283	0.04409
CML	0.05872	0.04912	0.14871	0.10594
PRME	0.04076	0.03247	0.10189	0.06719
ST-RNN	0.05006	0.03903	0.13649	0.08359
DeepMove	0.06012	0.04997	0.15096	0.10889
ST-LSTM	0.07043	0.05819	0.16762	0.12440
GRU	0.05118	0.04092	0.13765	0.08503
STA-RNN	0.07212	0.0603	0.17641	0.13450
Single-FT	0.07381	0.06145	0.18302	0.14089
MetaRNN	0.08122	0.07031	0.20872	0.16252

表 5.3 各类方法在 Foursquare（洛杉矶）数据集上的推荐结果统计

Table 5.3 Performance comparison for different methods on Foursquare (Los Angeles)

Method	Foursquare (San Francisco)			
	HR@5	NDCG@5	HR@10	NDCG@10
MF-BPR	0.03312	0.02801	0.06103	0.04334
CML	0.05541	0.04647	0.14276	0.10012
PRME	0.03972	0.03032	0.09535	0.06197
ST-RNN	0.04718	0.03827	0.12902	0.08143
DeepMove	0.05982	0.04921	0.14872	0.10727
ST-LSTM	0.06801	0.05654	0.16426	0.12026
GRU	0.04831	0.03972	0.13033	0.08209

STA-RNN	0.07104	0.05981	0.16824	0.12903
Single-FT	0.07217	0.05807	0.16002	0.13501
MetaRNN	0.08086	0.06769	0.20224	0.16157

除此之外，在进行 Top-N 推荐任务的过程中，还通过将参数 N 在取值范围 [1, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50] 内进行变化，来观察不同模型的表现情况，实验结果如图 5.6 所示。从中，可以看出 MetaRNN 在 N 的取值范围内表现均优于其他对比实验方法。另外，将图 5.6 中实验结果进行横向对比，我们发现，在旧金山数据集上，MetaRNN 的 NDCG@N 的提升效果更佳明显，并且在洛杉矶数据集上也发现了类似现象。这说明在 MetaRNN 的排序结果中，高关联度的结果出现在更靠前的位置，具有更高的累计增益值。

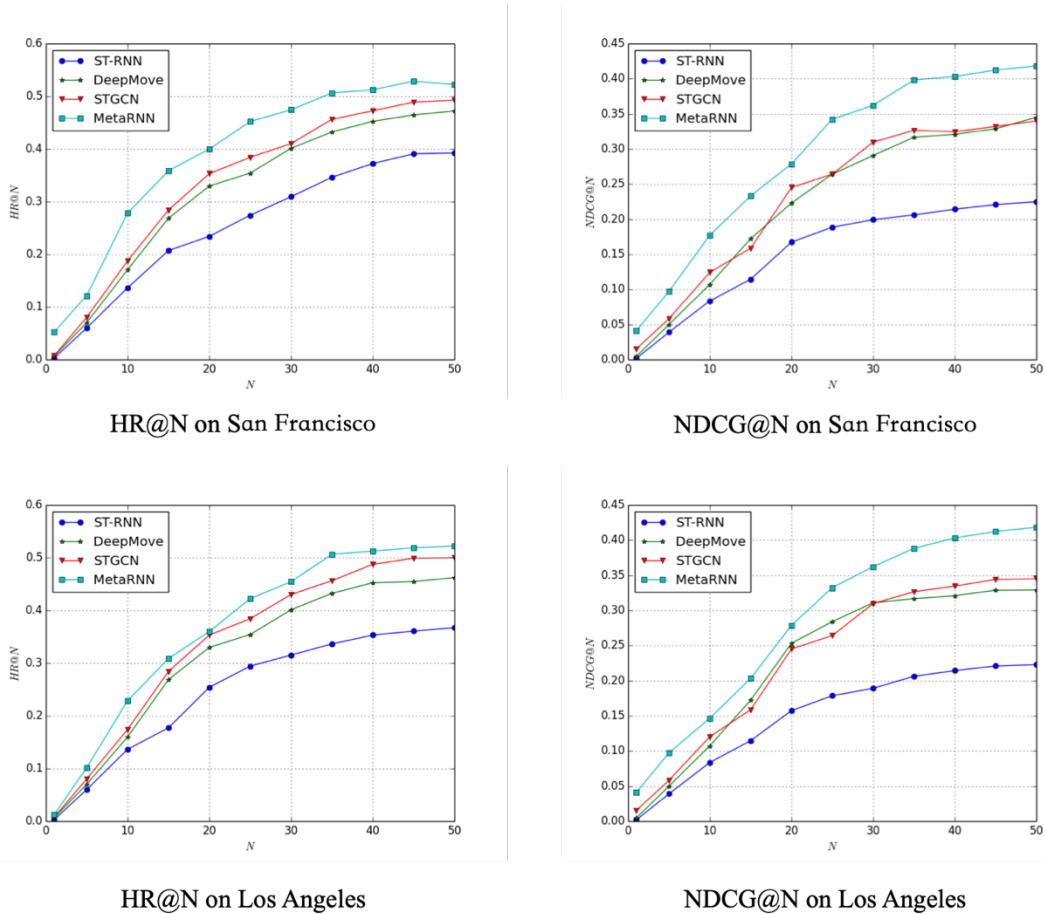


图 5.6 旧金山和洛杉矶数据集上的 HR@N 和 NDCG@N 结果

Figure 5.6 The HR@N and NDCG@N of the San Francisco and Los Angeles, respectively

5.3.2.2 模型收敛性分析

本实验还分析了 MetaRNN 模型的收敛速度，如图 5.7 和 5.8 所示。实验过程中，本文对比了通过单个源城市来进行 MetaRNN 初始参数更新（SFT）和通过多个数据丰富的源城市来构建多任务进行初始参数更新（Multi-FT）两种情况下，MetaRNN 的训练和测试过程的损失变化趋势。从实验结果中可以看出 MetaRNN 模型能够经过很少的迭代次数实现模型的收敛，并且通过多个源城市进行元学习任务在一定程度上缓解了单个城市移动行为知识迁移的局限性，具有更强的泛化能力，有助于加快模型的训练速度和提升模型的推荐效果。

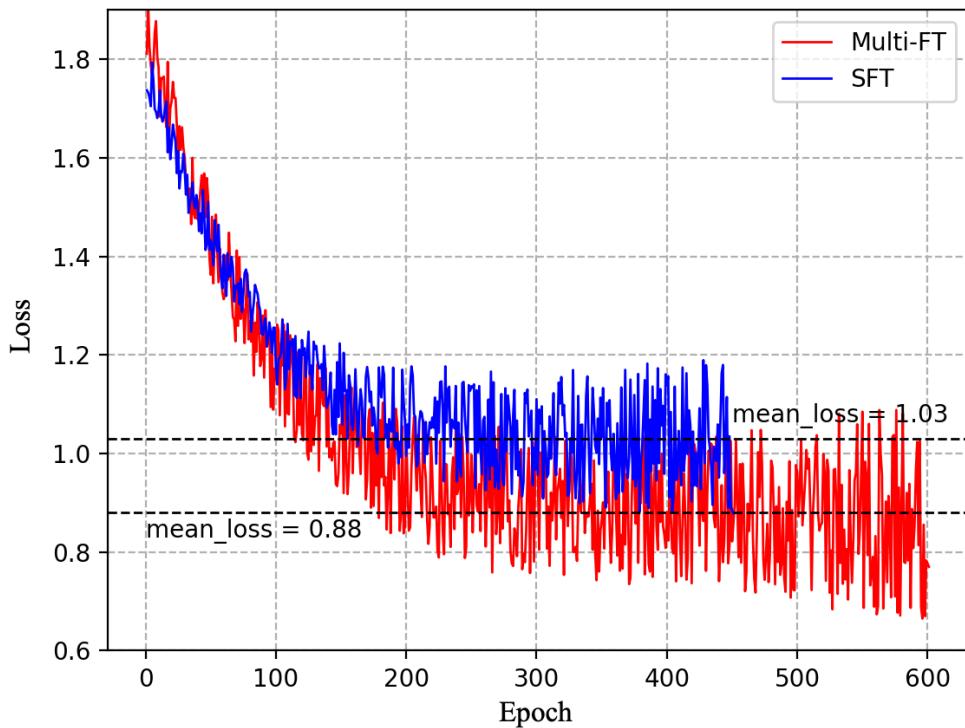


图 5.7 MetaRNN 的训练过程曲线

Figure 5.7 The loss curve of MateRNN in the Mate-training

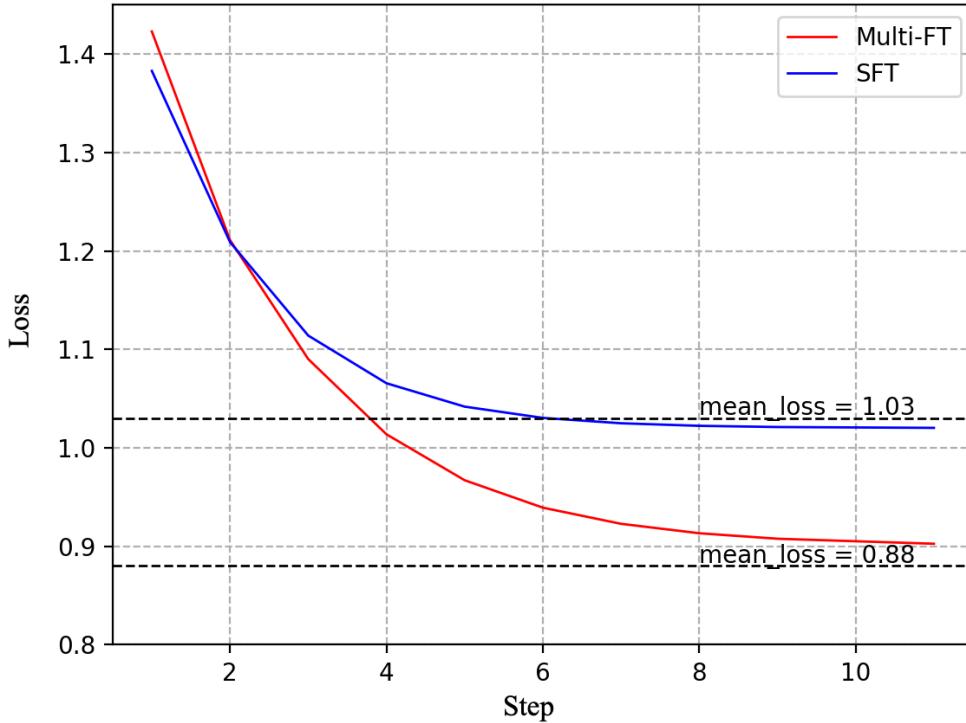


图 5.8 MetaRNN 的测试过程曲线

Figure 5.8 The loss curve of MateRNN in the Mate-testing

5.3.2.3 参数敏感度分析

本节评估了 Foursquare（洛杉矶）和 Foursquare（旧金山）两个数据集上 MetaRNN 对两个关键参数的敏感性：模型训练数据规模和签到记录表征向量维度 d。

模型训练数据规模的敏感性：本实验研究了在目标城市中模型训练数据规模对 MetaRNN 性能的影响。图 5.9 显示了两种指标（HR@10 & NDCG@10）度量下，MetaRNN 及 STA-RNN 对训练数据规模的敏感性。其中，本文将两个数据集上的训练集比例（training ratio）从 0.2 递增至 1.0。发现在少量训练数据下，MetaRNN 比 STA-RNN 更鲁棒。当数据训练比例仅为 0.2 时，MetaRNN 和 STA-RNN 之间的差异很大，这是因为 MetaRNN 经过元学习框架之后获得了泛化能力更强的初始参数，能在数据匮乏环境下取得不错的 POI 推荐效果。另外，当训练集比例超过 0.6 时，MetaRNN 的性能变得相对稳定，依然优于 STA-RNN。

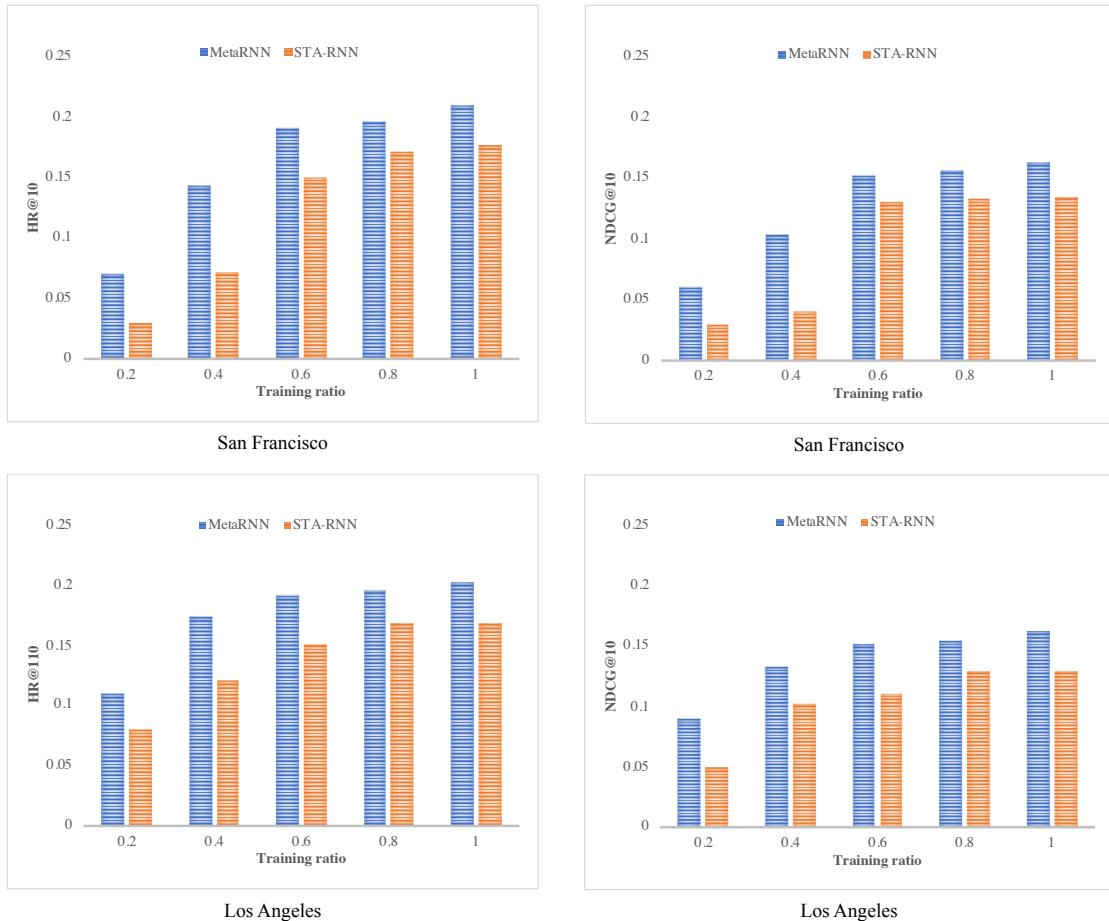


图 5.9 训练数据大小对 STA-RNN 和 MetaRNN 的影响

Figure 5.9 HR@10 of STA-RNN and MetaRNN with varying training ratio

签到行为表征向量维度 d 的灵敏度：本实验研究了表征向量维度 d 对 MetaRNN 性能的影响。图 5.10 显示了 d 从 8 逐步增加到 256，MetaRNN 及对比方法在 HR @ 10 以及 NDCG@10 的值。从图中可以观察到，MetaRNN 及其变体的性能随着 d 的增加先呈上升趋势然后趋于稳定甚至略有下降。其原因在于参数 d 控制了 MetaRNN 的复杂度。当 d 增加时，MetaRNN 具有更强的表达能力能够捕获签到行为序列度量空间的时空内在结构。但当 d 过大时，因训练数据集的大小有限，模型容易出现过拟合，从而导致了模型的推荐效果下降。

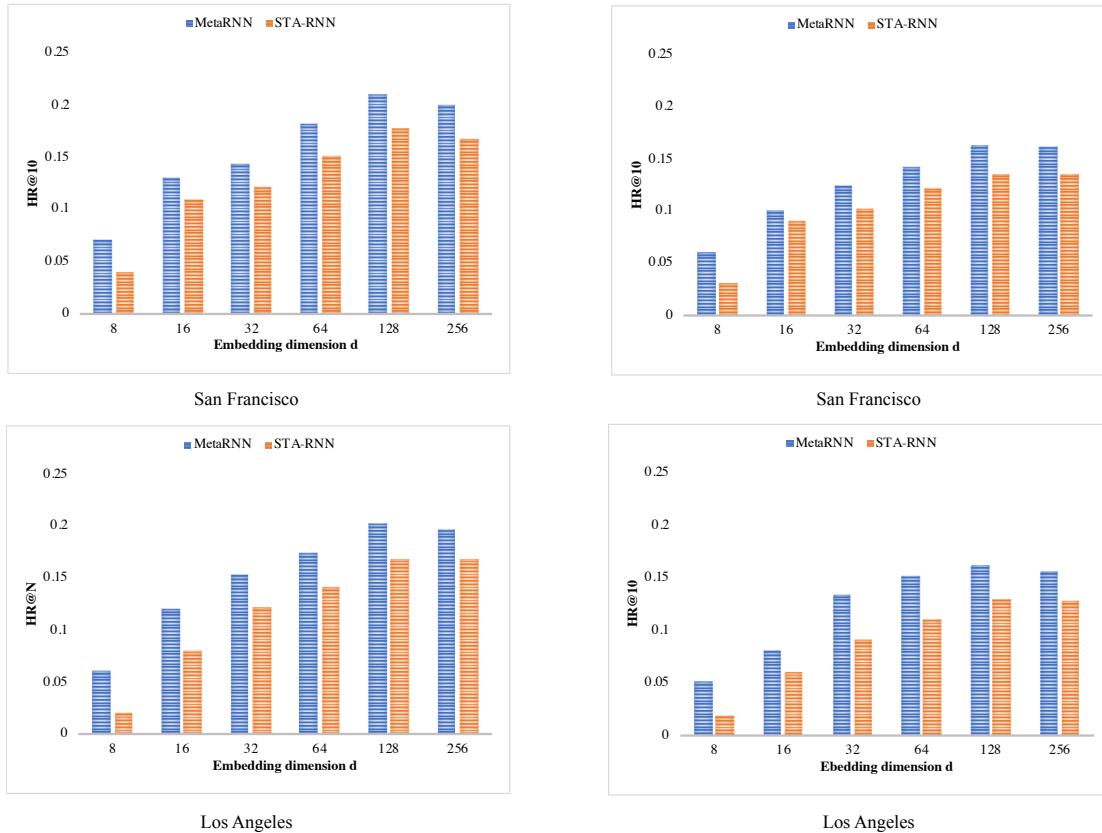


图 5.10 表征向量维度变化对 MetaRNN 的影响

Figure 5.10 HR@10 of MetaRNN with varying embedding size d

5.4 小结

本章提出了一种基于时空元学习 NPOI 推荐方法 MetaRNN，解决了数据匮乏环境下，NPOI 推荐模型训练困难，导致推荐模型难以对少样本用户进行有效拟合的问题。MetaRNN 具有快速收敛、准确和泛化能力强的特点。它的创新性在于两个方面：(1) 设计了一种面向 NPOI 推荐的元学习框架，将元学习应用于具有多类时空数据的 NPOI 推荐任务，该框架能够在建模时空差异性的基础上，通过构建多任务学习具有强泛化能力的初始参数，来提升少样本场景下的 NPOI 推荐性能；(2) 提出了一种基于时空注意力的 POI 序列编码网络，能够建模历史隐状态对当前状态的影响，保证在少样本场景下有效获取用户偏好表征。在多个真实数据集上的实验结果表明，对比基线方法，MetaRNN 在数据匮乏环境下进行 NPOI 推荐的命中率更高。

第6章 结论与展望

6.1 论文主要贡献和创新

本文深入分析和探讨了位置社交网络中下一个兴趣点推荐背景、内容和研究现状。首先简要介绍了兴趣点的定义、特点及分类；然后，介绍了兴趣点推荐的研究框架，并概述了主要的研究工作。最后，深入调研了已有研究在多维度社交关系感知的下一个兴趣点推荐、时间敏感的下一个兴趣点推荐和数据匮乏环境下的下一个兴趣点推荐领域已经取得的成果，分析了这些成果的不足和挑战。

本文围绕用户社交关系建模、用户连续行为偏好刻画和数据匮乏环境下模型训练三项研究内容。利用深度学习方法在隐含特征提取和复杂函数近似方面的优势，提出了多维度社交关系感知的下一个兴趣点推荐方法、用户动态偏好感知的下一个兴趣点推荐方法和基于时空元学习的下一个兴趣点推荐方法，解决了兴趣点推荐面临的用户社交关系建模不准确、用户连续行为偏好刻画缺失、数据匮乏环境下模型难以训练等问题。本文的主要创新性工作和贡献如下：

（1）提出了一种多维度社交关系感知的下一个兴趣点推荐方法

针对用户社交关系建模不准确导致推荐模型难以建模用户多维度偏好差异的问题，本文提出了一种社交关系感知的 NPOI 推荐方法 GraphSAR。该方法旨在采用维度级的注意力机制来学习用户表征之间的非结构化关联，使得模型可以更好地学习社交关系对用户偏好的影响，提升 NPOI 推荐效果。本文设计了一种基于注意力机制的兴趣偏好通道网络来建模用户之间的社交关系影响，并在影响注意力的引导下获取用户偏好表征。此外，为了有效利用用户-POI 异构交互网络中的丰富语义信息，本文设计了一种基于元路径的多通道语义表征网络，利用图卷积网络对多通道语义网络中具有丰富隐含语义信息的 POI 进行语义编码，以保证可以充分利用历史签到信息来提取节点间的关系特征。本文使用两个真实签到数据集评估了 GraphSAR 的效果。实验结果表明，对比最好的基线方法，GraphSAR 在 NPOI 推荐任务中的 HR@10 指标上获得了 4% 左右的提升。

（2）提出了一种用户动态偏好感知的下一个兴趣点推荐方法

针对用户连续行为偏好刻画缺失导致推荐模型难以建模用户的连续状态变

化的问题，本文提出了一种用户动态偏好感知的 NPOI 推荐方法 PINN。该方法通过整合用户的历史静态偏好和当前动态偏好来进行 NPOI 推荐任务。PINN 包括两个主要部分：偏好记忆网络和偏好动态演化网络。在偏好记忆网络中，本文提出了一种基于用户群体移动行为模式的用户历史偏好表征网络，该网络可以建模用户群体移动模式，并在群体移动模式的指导下获取用户的历史偏好表征；在动态偏好演化网络中，本文设计了一种基于时空神经常微分方程的用户动态偏好感知方法，该方法首次将神经常微分方程用于 NPOI 推荐领域，定义连续的动态过程，可以接受任意时刻的输入数据，生成用户的动态偏好表征，克服了大多数时序推荐模型只能处理静态离散化的时间间隔的局限性，更适用于 NPOI 序列推荐任务。本文使用六个城市的真实签到数据集评估了 PINN 的效果。实验结果表明，对比最好的基线方法，PINN 在 NPOI 推荐任务中的 HR@10 指标上获得了 7% 左右的提升。

(3) 提出了一种基于时空元学习的下一个兴趣点推荐方法

针对数据匮乏环境下，NPOI 推荐模型训练困难，导致推荐模型难以对少样本用户进行有效拟合的问题。本文提出了基于时空元学习的 NPOI 推荐方法 MetaRNN。该方法旨在为 NPOI 推荐模型生成高鲁棒性的初始化参数，使得模型仅利用少量的训练样本实现对 NPOI 的精准推荐。本文设计了一种面向 NPOI 推荐的元学习框架，将元学习应用于具有多类时空数据的 NPOI 推荐任务，该框架在建模时空差异性的基础上，通过构建多任务学习具有强泛化能力的初始参数，提升少样本场景下的 NPOI 推荐性能。此外，本文提出了一种基于时空注意力的 POI 序列编码网络，能够建模历史隐状态对当前状态的影响，保证在少样本场景下有效获取用户偏好表征。本文使用五个城市的真实签到数据集验证了 MetaRNN 的效果。实验结果表明，对比最好的基线方法，MetaRNN 方法在 NPOI 推荐任务中的 HR@10 指标上获得了 3% 左右的提升。

6.2 下一步研究工作展望

在现有工作基础上，本文作者计划在未来展开以下几方面的研究：

(1) 面向大规模签到数据的下一个兴趣点推荐方法

本文提出的 GraphSAR 设计的基于图神经网络的 POI 语义表征网络仅合适处理小规模的异构图数据。在实际的位置社交网络中，由签到数据构建的异构图网络通常包含千万级甚至亿级的节点和边。在该场景下，以整个图的邻接关系作为输入，将造成巨大的计算开销和存储开销。在下一步工作中，将继续研究大规模异构签到图的图神经网络方法，通过设计合适的采样策略，在保证效果的同时，提高计算效率。

(2) 下一个兴趣点推荐场景下的关系挖掘可解释性研究

在本文提出的时序关系建模的网络中，签到行为之间时序关系的学习是一个黑盒的过程，使得提取的用户偏好时序表征通常不具有可解释性。而实际应用中，时序关系特征的可解释性对下一个兴趣点推荐具有重要意义。本文所提出的 GraphSAR、PINN 以及 MetaRNN 中的时序关系表征部分，虽然可以在监督信息的指导下建模签到行为间的关联关系，但生成的用户偏好时序表征并不能直观表现用户签到行为间是由何种关系建立的关联。在下一步工作中我们将进行关系挖掘的可解释性研究，重点解决时序关系表征网络提取到的时序关系特征的可解释性问题。

(3) 高效的下一个兴趣点在线推荐方法

本文提出的 PINN 等下一个兴趣点推荐方法的工作原理是基于离线训练学习的方式进行下一个兴趣点推荐。相比与在线下一个兴趣点推荐系统，现存的离线兴趣点推荐模型过多的依赖于用户历史签到数据，因此在推荐模型的冷启动和用户移动行为偏好演化性等方面存在很大的缺陷。而当前基于 web 的在线推荐方法已经在短视频推荐和电子商务等领域取得了不错的研究成果。因此，下一步工作中，将继续研究在实时推荐场景下的模型结构，解决在线下一个兴趣点推荐存在的冷启动以及用户动态偏好捕捉的问题。

参考文献

- [1] 陈丽, 朱裴松, 钱铁云, 等. 基于边采样的网络表示学习模型[J]. 软件学报, 2018, 3: 756-771.
- [2] 廖国琼, 姜珊, 周志恒等. 基于位置社会网络的双重细粒度兴趣点推荐[J]. 计算机研究与发展. 2017, 54(11): 2600-2610.
- [3] 刘树栋, 孟祥武. 基于位置的社会化网络推荐系统[J]. 计算机学报, 2015, 038(002): 322-336.
- [4] 彭宏伟, 靳远远, 吕晓强, 等. 一种基于矩阵分解的上下文感知 POI 推荐算法[J]. 计算机学报, 2019(8):1797-1811.
- [5] 齐金山, 梁循, 李志宇, 等. 大规模复杂信息网络表示学习: 概念, 方法与挑战[J]. 计算机学报, 2017:1-26.
- [6] 任星怡, 宋美娜, 宋俊德. 基于用户签到行为的兴趣点推荐[J]. 计算机学报. 2017. 4 0(1): 28-50.
- [7] 涂存超, 杨成, 刘知远, 等. 网络表示学习综述[J]. 中国科学: 信息科学, 2017, 52(5 2): 1412.
- [8] 温雯, 黄家明, 蔡瑞初, 等. 一种融合节点先验信息的图表示学习方法[J]. 软件学报, 2018, 3: 786-798.
- [9] BELKIN M, NIYOGI P. Laplacian eigenmaps and spectral techniques for embedding and clustering[C]//Advances in neural information processing systems. 2002: 585 -591. <http://www-2.cs.cmu.edu/Groups/NIPS/NIPS2001/papers/psgz/AA42.ps.gz>.
- [10] BENGIO Y, COURVILLE A, VINCENT P. Representation learning: A review and new perspectives[J]. IEEE transactions on pattern analysis and machine intelligence, 2013, 35(8): 1798-1828. <https://arxiv.org/pdf/1206.5538.pdf>.
- [11] CAI X, HAN J, YANG L. Generative adversarial network based heterogeneous bibliographic network representation for personalized citation recommendation[C]//Thirteenth AAAI Conference on Artificial Intelligence. 2018: 5747—5754. <https://www.aaai.org/ocs/index.php/AAAI/AAAI18/paper/viewPDFInterstitial/16310/16136>.
- [12] CAI H, ZHENG V W, CHANG K C C. A comprehensive survey of graph embedding: Problems, techniques, and applications[J]. IEEE Transactions on Knowledge and Data Engineering, 2018, 30(9): 1616-1637. <https://arxiv.org/pdf/1709.07604.pdf>.
- [13] CAO S, LU W, XU Q. Graep: Learning graph representations with global structural information[C]//Proceedings of the 24th ACM International Conference on Information and Knowledge Management. ACM, 2015: 891-900. <https://dl.acm.org/pdf/10.1145/2806416.2806512.pdf>.
- [14] CAO W, WANG D, LI J, et al. Brits: Bidirectional recurrent imputation for time series[J]. arXiv preprint arXiv:1805.10572, 2018. <https://arxiv.org/pdf/1805.10572.pdf>.

- [15] CHANG B, PARK Y, PARK D, et al. Content-Aware Hierarchical Point-of-Interest Embedding Model for Successive POI Recommendation[C]//IJCAI. 2018, 2018: 27 th. <https://www.ijcai.org/Proceedings/2018/0458.pdf>.
- [16] CHANG S, HAN W, TANG J, et al. Heterogeneous network embedding via deep architectures[C]//Proceedings of the 21th ACM SIGKDD International Conference o n Knowledge Discovery and Data Mining. ACM, 2015: 119-128. <https://dl.acm.org/doi/pdf/10.1145/2783258.2783296>.
- [17] CHELSEA FINN, PIETER ABBEEL, and SERGEY LEVINE. Model-agnostic meta - learning for fast adaptation of deep networks. In Proceedings of the 34th Internat ional Conference on Machine Learning-Volume 70. JMLR. org, 2017. 1126–1135. <http://proceedings.mlr.press/v70/finn17a/finn17a.pdf>.
- [18] CHEN T, SUN Y. Task-guided and path-augmented heterogeneous network embedd ing for author identification[C]//Proceedings of the Tenth ACM International Confer ence on Web Search and Data Mining. 2017: 295-304. <https://dl.acm.org/doi/pdf/10.1145/3018661.3018735>.
- [19] CHEN R T Q, RUBANOVA Y, BETTENCOURT J, et al. Neural ordinary differe ntial equations[C]//Advances in neural information processing systems. 2018: 6571- 6583. <https://arxiv.org/pdf/1806.07366.pdf%20http://arxiv.org/abs/1806.07366>.
- [20] CHENG C, YANG H, KING I, et al. Fused matrix factorization with geographical and social influence in location-based social networks[C]//Proceedings of the AAA I conference on artificial intelligence. 2012, 26(1). <https://ojs.aaai.org/index.php/AA AI/article/view/8100/7958>.
- [21] CHENG H T, KOC L, HARMSEN J, et al. Wide & deep learning for recommend er systems[C]//Proceedings of the 1st workshop on deep learning for recommender systems. 2016: 7-10. <https://dl.acm.org/doi/pdf/10.1145/2988450.2988454>.
- [22] CUI Y, DENG L, ZHAO Y, et al. Hidden poi ranking with spatial crowdsourcing [C]//Proceedings of the 25th ACM SIGKDD International Conference on Knowledg e Discovery & Data Mining. 2019: 814-824. <https://dl.acm.org/doi/pdf/10.1145/3292500.3330844>.
- [23] DAI Q, LI Q, TANG J, et al. Adversarial network embed-ding[C]//Thirty-Second AAAI Conference on Artificial Intelligence. 2018. <https://ojs.aaai.org/index.php/AAA I/article/download/11865/11724>.
- [24] DE BROUWER E, SIMM J, ARANY A, et al. GRU-ODE-Bayes: Continuous mod eling of sporadically-observed time series[C]//Advances in Neural Information Proce ssing Systems. 2019: 7379-7390. <https://arxiv.org/pdf/1905.12374>.
- [25] DING R, CHEN Z. RecNet: A deep neural network for personalized POI recomme ndation in location-based social networks[J]. International Journal of Geographical I nformation Science, 2018, 32(8): 1631-1648. <https://www.tandfonline.com/doi/abs/10.1080/13658816.2018.1447671>.
- [26] DONG Y, CHAWLA N V, SWAMI A. metapath2vec: Scalable representation learn

- ing for heterogeneous net-works[C]//Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2017: 135-144. <https://dl.acm.org/doi/pdf/10.1145/3097983.3098036>.
- [27] DU N, WANG Y, HE N, et al. Time-sensitive recommendation from recurrent user activities[C]//Advances in Neural Information Processing Systems. 2015: 3492-3500. <http://www.cc.gatech.edu/~lsong/papers/DuWanHeSon15.pdf>.
- [28] FENG J, LI Y, ZHANG C, et al. Deepmove: Predicting human mobility with attentional recurrent networks[C]//Proceedings of the 2018 world wide web conference. 2018: 1459-1468. <https://dl.acm.org/doi/pdf/10.1145/3178876.3186058>.
- [29] FENG S, LI X, ZENG Y, et al. Personalized ranking metric embedding for next new poi recommendation[C]//IJCAI'15 Proceedings of the 24th International Conference on Artificial Intelligence. ACM, 2015: 2069-2075. <https://research.tees.ac.uk/ws/files/4382450/592759.pdf>.
- [30] FENG S, CONG G, AN B, et al. Poi2vec: Geographical latent representation for predicting future visitors[C]//Proceedings of the AAAI Conference on Artificial Intelligence. 2017, 31(1). <https://ojs.aaai.org/index.php/AAAI/article/view/10500/10359>.
- [31] FINN C, ABBEEL P, LEVINE S. Model-agnostic meta-learning for fast adaptation of deep networks[C]//International Conference on Machine Learning. PMLR, 2017: 1126-1135. <http://proceedings.mlr.press/v70/finn17a/finn17a.pdf>.
- [32] FU T, LEE W C, LEI Z. Hin2vec: Explore meta-paths in heterogeneous information networks for representation learning[C]//Proceedings of the 2017 ACM on Conference on Information and Knowledge Management. ACM, 2017: 1797-1806. <https://dl.acm.org/doi/pdf/10.1145/3132847.3132953>.
- [33] GAO H, TANG J, HU X, et al. Exploring temporal effects for location recommendation on location-based social networks[C]//Proceedings of the 7th ACM conference on Recommender systems. 2013: 93-100. <https://dl.acm.org/doi/pdf/10.1145/2507157.2507182>.
- [34] GAO H, TANG J, LIU H. gSCorr: Modeling geo-social correlations for new check-ins on location-based social networks[C]//Proceedings of the 21st ACM international conference on Information and knowledge management. 2012: 1582-1586. <https://dl.acm.org/doi/pdf/10.1145/2396761.2398477>.
- [35] GAO H, TANG J, LIU H. Addressing the cold-start problem in location recommendation using geo-social correlations[J]. Data Mining and Knowledge Discovery, 2015, 29(2): 299-323. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.587.1763&rep=rep1&type=pdf>.
- [36] GROVER A, LESKOVER J. node2vec: Scalable feature learning for networks[C]//Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2016: 855-864. <https://dl.acm.org/doi/pdf/10.1145/2939672.2939754>.
- [37] GUI I, ZWERDLING N, CARMEL D, et al. Personalized recommendation of soci

- al software items based on social relations[C]//Proceedings of the third ACM conference on Recommender systems. 2009: 53-60. <https://dl.acm.org/doi/pdf/10.1145/1639714.1639725>.
- [38] GUO G, ZHANG J, YORKE-SMITH N. A novel recommendation model regularized with user trust and item ratings[J]. *ieee transactions on knowledge and data engineering*, 2016, 28(7): 1607-1620. <http://www.logicsystems.org.in/Base%20Papers/2016%20Java/LSJ1627%20-%20A%20Novel%20Recommendation%20Model%20Regularized.pdf>.
- [39] HAN H, ZHANG M, HOU M, et al. STGCN: A Spatial-Temporal Aware Graph Learning Method for POI Recommendation[C]//2020 IEEE International Conference on Data Mining (ICDM). IEEE, 2020: 1052-1057. <https://ieeexplore.ieee.org/abstract/document/9338281/>.
- [40] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 770-778. https://openaccess.thecvf.com/content_cvpr_2016/papers/He_Deep_Residual_Learning_CVPR_2016_paper.pdf.
- [41] HSIEH C, YANG L, CUI Y, et al. Collaborative Metric Learning[C]//the web conference, 2017: 193-201.<https://dl.acm.org/doi/pdf/10.1145/3038912.3052639>.
- [42] HUANG L, MA Y, LIU Y, et al. DAN-SNR: A Deep Attentive Network for Social-aware Next Point-of-interest Recommendation[J]. *ACM Transactions on Internet Technology (TOIT)*, 2020, 21(1): 1-27. <https://dl.acm.org/doi/pdf/10.1145/3430504>.
- [43] HUANG L, MA Y, WANG S, et al. An attention-based spatiotemporal lstm network for next poi recommendation[J]. *IEEE Transactions on Services Computing*, 2019. <https://ieeexplore.ieee.org/abstract/document/8723186/>.
- [44] ISLAM M, MOHAMMAD M M, DAS S S S, et al. A Survey on Deep Learning Based Point-Of-Interest (POI) Recommendations[J]. *arXiv preprint arXiv:2011.10187*, 2020. <https://arxiv.org/pdf/2011.10187>.
- [45] JIANG M, CUI P, WANG F, et al. Scalable recommendation with social contextual information[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2014, 26(11): 2789-2802. <https://ieeexplore.ieee.org/abstract/document/6714549/>.
- [46] KEFALAS P, SYMEONIDS P, MANOLOPOULOS Y. Recommendations based on a heterogeneous spatio-temporal social network[J]. *World Wide Web*, 2018, 21(2): 345-371. <https://datalab-old.csd.auth.gr/wp-content/uploads/publications/WWW17ksm.pdf>.
- [47] KINGMA D P, WELLING M. Auto-encoding variational bayes[J]. *arXiv preprint arXiv:1312.6114*, 2013. <https://arxiv.org/pdf/1312.6114.pdf>.
- [48] LEE Y, CHOI S. Gradient-based meta-learning with learned layerwise metric and subspace[C]//International Conference on Machine Learning. PMLR, 2018: 2927-2936. <http://proceedings.mlr.press/v80/lee18a/lee18a.pdf>

- [49] LI D, GONG Z, ZHANG D. A common topic transfer learning model for crossing city POI recommendations[J]. IEEE transactions on cybernetics, 2018, 49(12): 428 2-4295. <https://ieeexplore.ieee.org/abstract/document/8451956/>.
- [50] LI K, MIALIK J. Learning to optimize[J]. arXiv preprint arXiv:1606.01885, 2016. [https://arxiv.org/pdf/1606.01885](https://arxiv.org/pdf/1606.01885.pdf).
- [51] LI X, CONG G, LI X L, et al. Rank-geofm: A ranking based geographical factorization method for point of interest recommendation[C]//Proceedings of the 38th international ACM SIGIR conference on research and development in information retrieval. 2015: 433-442. <https://dl.acm.org/doi/pdf/10.1145/2766462.2767722>.
- [52] LIAN D, ZHAO C, XIE X, et al. GeoMF: joint geographical modeling and matrix factorization for point-of-interest recommendation[C]//Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining. 2014: 831-840. <https://dl.acm.org/doi/pdf/10.1145/2623330.2623638>.
- [53] LIPTON Z C, KALE D, WETZEL R. Directly modeling missing data in sequences with rnns: Improved classification of clinical time series[C]//Machine learning for healthcare conference. PMLR, 2016: 253-270. <http://proceedings.mlr.press/v56/Lipton16.pdf>.
- [54] LIU B, FU Y, YAO Z, et al. Learning geographical preferences for point-of-interest recommendation[C]//Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining. 2013: 1043-1051. <https://dl.acm.org/doi/pdf/10.1145/2487575.2487673>.
- [55] LIU Q, WU S, WANG L, et al. Predicting the next location: A recurrent model with spatial and temporal contexts[C]//Proceedings of the AAAI Conference on Artificial Intelligence. 2016, 30(1). <https://ojs.aaai.org/index.php/AAAI/article/view/9971/9830>.
- [56] LIU X, LIU Y, LI X. Exploring the Context of Locations for Personalized Location Recommendations[C]//IJCAI. 2016: 1188-1194. <https://www.ijcai.org/Proceedings/16/Papers/172.pdf>.
- [57] MAATEN L, HINTON G. Visualizing data using t-SNE[J]. Journal of machine learning research, 2008, 9(Nov): 2579-2605. <https://www.jmlr.org/papers/volume9/vandermaaten08a/vandermaaten08a.pdf?fbclid=IwA>.
- [58] MEI H, EISNER J. The neural hawkes process: A neurally self-modulating multivariate point process[J]. arXiv preprint arXiv:1612.09328, 2016. <https://arxiv.org/pdf/1612.09328.pdf>.
- [59] MIKOLOV T, CHEN K, CORRADO G, et al. Efficient estimation of word representations in vector space[J]. arXiv preprint arXiv:1301.3781, 2013. <https://arxiv.org/pdf/1301.3781.pdf>.
- [60] MIKOLOV T, SUTSKEVER I, CHEN K, et al. Distributed representations of words and phrases and their compositionality[C]//Advances in neural information processing systems. 2013: 3111-3119. <http://papers.nips.cc/paper/5021-distributed-representations.pdf>.

- tions-of-words-and-phrases-and-their-com.pdf.
- [61] MISHRA N, ROHANINEJAD M, CHEN X, et al. A simple neural attentive meta-learner[J]. arXiv preprint arXiv:1707.03141, 2017. <https://arxiv.org/pdf/1707.03141.pdf>.
- [62] MNIIH, SALAKHUTDUNOV R. Probabilistic matrix factorization[C]//Advances in neural information processing systems. 2008: 1257-1264. <http://www.cs.utoronto.ca/~amnih/papers/pmf.pdf>.
- [63] MUNKHDALAI T, YUAN X, MEHRI S, et al. Rapid adaptation with conditionally shifted neurons[C]//International Conference on Machine Learning. PMLR, 2018: 3664-3673. <http://proceedings.mlr.press/v80/munkhdalai18a/munkhdalai18a.pdf>.
- [64] PAN S J, YANG Q. A survey on transfer learning[J]. IEEE Transactions on knowledge and data engineering, 2009, 22(10): 1345-1359. https://www.cse.ust.hk/~qyang/Docs/2009/tkde_transfer_learning.pdf.
- [65] PEROZZI B, AL-RFOU R, SKIENA S. Deepwalk: Online learning of social representations[C]//Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2014: 701-710. <https://dl.acm.org/doi/pdf/10.1145/2623330.2623732>.
- [66] RAJKOMAR A, OREN E, CHEN K, et al. Scalable and accurate deep learning with electronic health records[J]. NPJ Digital Medicine, 2018, 1(1): 1-10. <https://www.nature.com/articles/s41746-018-0029-1%22>.
- [67] RAVI S, LAROCHELLE H. Optimization as a model for few-shot learning[J]. 2016. <https://openreview.net/pdf?id=rJY0-Kcll>.
- [68] RENDLE S, FREUDENTHALER C, GANTNER Z, et al. BPR: Bayesian personalized ranking from implicit feedback[J]. arXiv preprint arXiv:1205.2618, 2012. <https://arxiv.org/pdf/1205.2618.pdf>.
- [69] RESNICK P, VARIAN H R. Recommender systems[J]. Communications of the ACM, 1997, 40(3): 56-58. <https://dl.acm.org/doi/pdf/10.1145/245108.245121>.
- [70] RUBANOVA Y, CHEN R T Q, DUVENAUD D K. Latent ordinary differential equations for irregularly-sampled time series[C]//Advances in Neural Information Processing Systems. 2019: 5320-5330. <https://openreview.net/pdf?id=HygCYNSILB>.
- [71] QIAN T, LIU B, NGUYEN Q V H, et al. Spatiotemporal representation learning for translation-based POI recommendation[J]. ACM Transactions on Information Systems (TOIS), 2019, 37(2): 1-24. <https://dl.acm.org/doi/pdf/10.1145/3295499>.
- [72] SHI C, HU B, ZHAO W X, et al. Heterogeneous information network embedding for recommendation[J]. IEEE Transactions on Knowledge and Data Engineering, 2019, 31(2): 357-370. <https://arxiv.org/pdf/1711.10730.pdf%7D>.
- [73] SNELL J, SWERSKY K, ZEMEL R S. Prototypical networks for few-shot learning[J]. arXiv preprint arXiv:1703.05175, 2017. <https://arxiv.org/pdf/1703.05175.pdf>.
- [74] STEFFEN RENDLE, CHRISTOPH FREUDENTHALER, ZENO GANTNER, and LARS SCHMIDT-THIEME. BPR: Bayesian personalized ranking from implicit feedb

- ack. arXiv preprint. arXiv:1205.2618. <https://arxiv.org/pdf/1205.2618.pdf>.
- [75] SUN Y, HAN J. Mining heterogeneous information networks: principles and methodologies[J]. Synthesis Lectures on Data Mining and Knowledge Discovery, 2012, 3(2): 1-159. <https://www.morganclaypool.com/doi/abs/10.2200/s00433ed1v01y201207dmk005>.
- [76] SUN Y, HAN J, YAN X, et al. Pathsim: Meta path-based top-k similarity search in heterogeneous information networks[J]. Proceedings of the VLDB Endowment, 2011, 4(11): 992-1003. <https://dl.acm.org/doi/pdf/10.14778/3402707.3402736>.
- [77] SUNG F, YANG Y, ZHANG L, et al. Learning to compare: Relation network for few-shot learning[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 1199-1208. https://openaccess.thecvf.com/content_cvpr_2018/papers/Sung_Learning_to_Compare_CVPR_2018_paper.pdf.
- [78] TAN H, TANG W, FAN X, et al. SERL: Semantic-Path Biased Representation Learning of Heterogeneous Information Network[C]. knowledge science, engineering and management, 2018: 287-298. https://link.springer.com/chapter/10.1007/978-3-319-99365-2_26.
- [79] TANG J, QU M, WANG M, et al. Line: Large-scale information network embedding[C]//Proceedings of the 24th International Conference on World Wide Web. International World Wide Web Conferences Steering Committee, 2015: 1067-1077. <https://dl.acm.org/doi/pdf/10.1145/2736277.2741093>.
- [80] TANG J, HU X, LIU H. Social recommendation: a review[J]. Social Network Analysis and Mining, 2013, 3(4): 1113-1133. <http://www.cse.msu.edu/~tangjili/publication/socialrecommendationreview.pdf>.
- [81] TANG J, WANG K. Personalized top-n sequential recommendation via convolutional sequence embedding[C]//Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining. 2018: 565-573.
- [82] TZENG E, HOFFMAN J, SAENKO K, et al. Adversarial discriminative domain adaptation[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 7167-7176. https://openaccess.thecvf.com/content_cvpr_2017/papers/Tzeng_Adversarial_Discriminative_Domain_CVPR_2017_paper.pdf.
- [83] VASWANI A, SHAZER N, PARMAR N, et al. Attention is all you need[C]//Advances in neural information processing systems. 2017: 5998-6008. <http://papers.nips.cc/paper/7181-attention-is-all-you-%0Aneed.pdf>.
- [84] WANG C, SONG Y, LI H, et al. Knowsim: A document similarity measure on structured heterogeneous information networks[C]//2015 IEEE International Conference on Data Mining. IEEE, 2015: 1015-1020. <https://www.ncbi.nlm.nih.gov/pmc/article/s PMC4811603>.
- [85] WANG D, CUI P, ZHU W. Structural deep network embedding[C]//Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2016: 1225-1234. <https://dl.acm.org/doi/pdf/10.1145/2939672.2939672>.

753.

- [86] WANG H, WANG J, WANG J et al. GraphGAN: graph representation learning with generative adversarial nets[C] // Thirty-Second AAAI Conference on Artificial Intelligence. 2018. <https://ojs.aaai.org/index.php/AAAI/article/download/11872/11731>.
- [87] WANG H, ZHANG F, HOU M, et al. Shine: Signed heterogeneous information network embedding for sentiment link prediction[C]//Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining. 2018: 592-600. <https://dl.acm.org/doi/pdf/10.1145/3159652.3159666>.
- [88] WANG S, WANG Y, TANG J, et al. What your images reveal: Exploiting visual contents for point-of-interest recommendation[C]//Proceedings of the 26th international conference on world wide web. 2017: 391-400. <https://dl.acm.org/doi/pdf/10.1145/3038912.3052638>.
- [89] WANG L, GENG X, MA X, et al. Crowd flow prediction by deep spatio-temporal transfer learning[J]. arXiv preprint arXiv:1802.00386, 2018. <https://www.researchgate.net/Crowd-Flow-Prediction-by-Deep-Spatio-Temporal-Transfer-Learning.pdf>.
- [90] WANG X, HOI S C H, ESTER M, et al. Learning personalized preference of strong and weak ties for social recommendation[C]//Proceedings of the 26th International Conference on World Wide Web. 2017: 1601-1610. <https://dl.acm.org/doi/pdf/10.1145/3038912.3052556>.
- [91] WANG Y, SHEN H, LIU S, et al. Learning user-specific latent influence and susceptibility from information cascades[C]//Twenty-Ninth AAAI Conference on Artificial Intelligence. 2015. <https://www.aaai.org/ocs/index.php/AAAI/AAAI15/paper/viewPDFInterstitial/9805/9283>.
- [92] WEI Y, ZHENG Y, YANG Q. Transfer knowledge between cities[C]//Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2016: 1905-1914. <https://dl.acm.org/doi/pdf/10.1145/2939672.2939830>.
- [93] XIAO L, MIN Z, YONGFENG Z. Joint factorizational topic models for cross-city recommendation[C]//Asia-Pacific Web (APWeb) and Web-Age Information Management (WAIM) Joint Conference on Web and Big Data. Springer, Cham, 2017: 591 -609. https://drive.google.com/file/d/1_fwddZtzzLVM7kgRyQGgtmJBy_XgXMcp/view.
- [94] YANG C, BAI L, ZHANG C, et al. Bridging collaborative filtering and semi-supervised learning: a neural approach for poi recommendation[C]//Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2017: 1245-1254. <https://dl.acm.org/doi/pdf/10.1145/3097983.3098094>.
- [95] YANG D, QU B, YANG J, et al. Revisiting user mobility and social relationships in lbsns: a hypergraph embedding approach[C]//The World Wide Web Conference. 2019: 2147-2157. <https://dl.acm.org/doi/pdf/10.1145/3308558.3313635>.
- [96] YANG D, QU B, YANG J, et al. Lbsn2vec++: Heterogeneous hypergraph embedding for location-based social networks[J]. IEEE Transactions on Knowledge and Data Engineering, 2020. <https://research.tudelft.nl/files/87409438/09099985.pdf>.

- [97] YANG D, ZHANG D, YU Z, et al. A sentiment-enhanced personalized location recommendation system[C]//Proceedings of the 24th ACM conference on hypertext and social media. 2013: 119-128. <https://dl.acm.org/doi/pdf/10.1145/2481492.2481505>.
- [98] YAO D, ZHANG C, HUANG J, et al. Serm: A recurrent model for next location prediction in semantic trajectories[C]//Proceedings of the 2017 ACM on Conference on Information and Knowledge Management. 2017: 2411-2414. <https://dl.acm.org/doi/pdf/10.1145/3132847.3133056>.
- [99] YAO H, LIU Y, WEI Y, et al. Learning from multiple cities: A meta-learning approach for spatial-temporal prediction[C]//The World Wide Web Conference. 2019: 2181-2191. <https://dl.acm.org/doi/pdf/10.1145/3308558.3313577>.
- [100] YAO H, WEI Y, HUANG J, et al. Hierarchically structured meta-learning[C]//International Conference on Machine Learning. PMLR, 2019: 7045-7054. <http://proceedings.mlr.press/v97/yao19b/yao19b.pdf>.
- [101] YE M, YIN P, LEE W C, et al. Exploiting geographical influence for collaborative point-of-interest recommendation[C]//Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval. 2011: 325 -334. <https://dl.acm.org/doi/pdf/10.1145/2009916.2009962>.
- [102] YIN H, WANG W, WANG H, et al. Spatial-aware hierarchical collaborative deep learning for POI recommendation[J]. IEEE Transactions on Knowledge and Data Engineering, 2017, 29(11): 2537-2551. https://drive.google.com/file/d/1kxy4_hu3dE2byj_m_ojxMaYrkCzei3XOR/view.
- [103] YIN H, ZHOU X, CUI B, et al. Adapting to user interest drift for poi recommendation[J]. IEEE Transactions on Knowledge and Data Engineering, 2016, 28(10): 2566-2581. <http://net.pku.edu.cn/daim/hongzhi.yin/papers/TKDE-2016.pdf>.
- [104] YIN H, ZHOU X, SHAO Y, et al. Joint modeling of user check-in behaviors for point-of-interest recommendation[C]//Proceedings of the 24th ACM International on Conference on Information and Knowledge Management. 2015: 1631-1640. <https://dl.acm.org/doi/pdf/10.1145/2806416.2806500>.
- [105] YING J J C, KUO W N, TSENG V S, et al. Mining user check-in behavior with a random walk for urban point-of-interest recommendations[J]. ACM Transactions on Intelligent Systems and Technology (TIST), 2014, 5(3): 1-26. <https://dl.acm.org/doi/pdf/10.1145/2523068>.
- [106] YING J J C, LU E H C, KUO W N, et al. Urban point-of-interest recommendation by mining user check-in behaviors[C]//Proceedings of the ACM SIGKDD International Workshop on Urban Computing. 2012: 63-70. <https://dl.acm.org/doi/pdf/10.1145/2346496.2346507>.
- [107] YU F, CUI L, GUO W, et al. A category-aware deep model for successive poi recommendation on sparse check-in data[C]//Proceedings of the web conference 2020. 2020: 1264-1274. <https://dl.acm.org/doi/pdf/10.1145/3366423.3380202>.
- [108] YU Y, CHEN X. A survey of point-of-interest recommendation in location-based s

- ocial networks[C]//Workshops at the Twenty-Ninth AAAI Conference on Artificial Intelligence. 2015. <https://www.aaai.org/ocs/index.php/WS/AAAIW15/paper/viewPDFInterstitial/10132/10253>.
- [109] ZHANG J D, CHOW C Y. iGSLR: personalized geo-social location recommendation: a kernel density estimation approach[C]//Proceedings of the 21st ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems. 2013: 334-343. <https://dl.acm.org/doi/pdf/10.1145/2525314.2525339>.
- [110] ZHANG J D, CHOW C Y, LI Y. iGeoRec: A personalized and efficient geographical location recommendation framework[J]. IEEE Transactions on Services Computing, 2014, 8(5): 701-714. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.723.2593&rep=rep1&type=pdf>.
- [111] ZHANG Y, DAI H, XU C, et al. Sequential click prediction for sponsored search with recurrent neural networks[J]. arXiv preprint arXiv:1404.5772, 2014. <https://ojs.aaai.org/index.php/AAAI/article/download/8917/8776>.
- [112] ZHAO P, ZHU H, LIU Y, et al. Where to go next: A spatio-temporal gated network for next POI recommendation[C]//Proceedings of the International Conference on Artificial Intelligence, 2019: 5877–5884. <https://ojs.aaai.org/index.php/AAAI/article/view/4537/4415>.
- [113] ZHAO S, KING I, LYU M R. A survey of point-of-interest recommendation in location-based social networks[J]. arXiv preprint arXiv:1607.00647, 2016. <https://arxiv.org/pdf/1607.00647>.
- [114] ZHOU X, MASCOLO C, ZHAO Z. Topic-enhanced memory networks for personalised point-of-interest recommendation[C]//Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. 2019: 3018-3028. <https://dl.acm.org/doi/pdf/10.1145/3292500.3330781>.
- [115] ZHOU Y, HUANG Y. Deepmove: Learning place representations through large scale movement data[C]//2018 IEEE International Conference on Big Data (Big Data). IEEE, 2018: 2403-2412. <https://arxiv.org/pdf/1807.04241>.

致 谢

时光流转，岁月更迭，蓦然回首，读博已经七个年头。读博期间有感恩、有收获、有心酸无助、有恋恋不舍。感谢这段人生经历，让我在苦苦探索的过程中不断成长，逐渐蜕变，由一个初出茅庐的毛头小子，渐渐成长为一个能够有能力有信心去探索未知的世界，去接受更难挑战的成年人。在此，我要向各位老师、同学和家人表达我最衷心的感谢。

首先，我要衷心感谢我的指导老师，毕经平研究员！感谢她为我提供了良好的科研环境；感谢她在更高的层次和角度上给我立身立业的指导。能够成为她的学生，是我一辈子的荣幸！还记得，每次在组会上做完报告，毕老师都会对我的近期工作进行悉心点评和综合梳理，让我的科研思路更加清晰，使得科研工作能够顺利开展。感谢毕老师在我迷茫的时候提供科研思路并给予学术资源的支持，这些都让我心存感激。尤其是在读博后期，从研究工作的初期调研到科研问题具体的研究以及最终论文撰写的各个阶段，都离不开她的悉心指导。没有她这些年来在学术研究上对我的激励和指导，我不可能取得现在的成绩。同时，也为我未能早些理解她传授的科研、做人的道理而遗憾，同时，为我未能达到她的预期而感到自责。在之后的人生道路上，我会继续以毕老师严谨的治学态度和积极乐观的人生态度作为工作和生活的指引，砥砺前行。在此谨向毕老师致以最诚挚的谢意！

其次，感谢计算所的领导和老师！在计算所这个温馨的大家庭中，所领导的关怀和教育让我充满信心，大师的风采和指导让我受益匪浅，老师和同学们的帮助也给我足够勇气克服各种困难。感谢冯刚老师、周世佳老师和研究生部其他老师给我在读博期间提供的支持和帮助。感谢课题组的黄建辉博士、范鑫鑫博士、景全亮老师、刘妍妍老师，感谢他们在学术研究和工作上对我的指导和帮助。特别感谢课题组的姚迪师兄，从论文工作的方向探索以及后期的工作开展，都离不开姚师兄的细心指导，甚是感激！感谢课题组的于金萍师姐、丁自旋师姐、朱志华师弟、王保利师弟、黄涛师弟和所有的师弟师妹们，有幸与你们相识，何其幸

运，感谢期间对我工作和生活上的支持，陪我走过这一段难忘的求学经历。

最后，感谢我的家人，感谢我的父亲和母亲，感谢你们把我带到这纷繁热闹的人世间，感谢你们的言传身教使我不断的进步。感谢我的姐姐，相爱相杀近三十载，给我无尽的关怀，并带来了姐夫、外甥女以及外甥，带给我太多的欢乐。感谢你们在背后默默地奉献和支持，让我没有后顾之忧的完成我的博士学位。谢谢你们，你们辛苦了！

站在离别的车站，即将踏上新的起点，我心中思绪万千，但是太多的话语终归有结束的一刻。请各位老师亲友放心，我会在新的道路上披荆斩棘。最后的最后，祝各位老师、同学、朋友和家人身体健康，一切顺利！

作者简历及攻读学位期间发表的学术论文与研究成果

作者简历:

2010 年 09 月——2014 年 07 月, 在北京科技大学计通学院学习, 获学士学位。
2014 年 09 月——至今, 在中国科学院计算技术研究所攻读博士学位。

获奖情况:

- [1] 2019 年 博士研究生国家奖学金
- [2] 2019 年 “智能争锋” 信息智能处理大赛亚军
- [3] 2017 年 中科院计算所北纬通信优秀科研博士奖学金
- [4] 2016 年 中国科学院大学三好学生
- [5] 2016 年 中国科学院大学优秀学生干部
- [6] 2016 年 海淀区优秀志愿者标兵、“义工之星”
- [7] 2016 年 中国科学院职工运动会 铅球季军

已发表（或正式接受）的学术论文:

- [1] **Tan H**, Tang W, Fan X, et al. SERL: Semantic-Path Biased Representation Learning of Heterogeneous Information Network[C]/International Conference on Knowledge Science, Engineering and Management. Springer, Cham, 2018: 287-298. (*KSEM-18*)
(Oral presentation, Best paper runner up)
- [2] 谭海宁, 姚迪, 毕经平, 向徐, 杨啸: 面向数据匮乏环境的下一个 POI 推荐方法, 高技术通讯, 2020, 已录用
- [3] **Tan H**, Yao D, Huang T, et al. Meta-Learning Enhanced Neural ODE for Citywide Next POI Recommendation[C]/2021 22nd IEEE International Conference on Mobile Data Management. IEEE, 2021: 89-98. (*MDM-21*)
(Oral presentation, Best paper runner up)
- [4] **Tan H**, Yao D, Bi J. Deep Transfer Learning for Successive POI Recommendation [C]/International Conference on Spatial Data and Intelligence. Springer, Cham, 2021: 134-

140. (*Spatial-DI-2I*)

- [5] 赵瑜, 谭海宁, 刘志方, 武超; 基于生成对抗模型的异质信息网络语义表征方法研究, 中文信息学报, 2019-0017
- [6] Jing Q, Yao D, Fan X, Tan H, et al. TRANSFAKE: Multi-task Transformer for Multimodal Enhanced Fake News Detection[C]//2021 International Joint Conference on Neural Networks. IEEE, 2021: 1-8. (*IJCNN-2I*)
- [7] Wang B, Fan X, Jing Q, Tan H, et al. AdvCGAN: An Elastic and Covert Adversarial Examples Generating Framework[C]//2021 International Joint Conference on Neural Networks. IEEE, 2021: 1-8. (*IJCNN-2I*)
- [8] Jing Q, Yao D, Fan X, Tan H, et al. TRAJCROSS: Trajectory Cross-Modal Retrieval with Mult-level Contrastive Learning[C]//2021 IEEE International Conference on Big Data. (*BigData-2I*)

参加的研究项目:

- [1] 2018.10-至今: 时空轨迹数据分析挖掘研究 (预研项目), 技术骨干
- [2] 2015.4-2017.10: 卫星大数据监控及目标识别系统 (预研项目), 技术骨干
- [3] 2016.9-2017.12: 域间路由分析系统, 核心开发人员
- [4] 2014.10-2016.10: 路由流量监测系统, 核心开发人员
- [5] 国家自然基金: 基于教育大数据的认知过程建模方法研究
- [6] 国家自然基金: 边缘计算弹性网络架构及协同交互安全可信机制研究
- [7] 国家自然基金: 弱监督下异常轨迹在线检测方法研究