

基于自然语言增强的签到轨迹与用户匹配方法

王天一, 林友芳, 贡乐天, 陈炜, 郭晟楠, 万怀宇

引用本文

王天一, 林友芳, 贡乐天, 陈炜, 郭晟楠, 万怀宇. [基于自然语言增强的签到轨迹与用户匹配方法](#)[J]. 计算机科学, 2025, 52(2): 99-106.

WANG Tianyi, LIN Youfang, GONG Letian, CHEN Wei, GUO Shengnan, WAN Huaiyu. [Check-in Trajectory and User Linking Based on Natural Language Augmentation](#) [J]. Computer Science, 2025, 52(2): 99-106.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[基于大小语言模型协同增强的中文电子病历依存句法分析](#)

Dependency Parsing for Chinese Electronic Medical Record Enhanced by Dual-scale Collaboration of Large and Small Language Models

计算机科学, 2025, 52(2): 253-260. <https://doi.org/10.11896/jsjcx.231200054>

[辅助判决的案情要素关联与证据提取](#)

Case Element Association with Evidence Extraction for Adjudication Assistance

计算机科学, 2025, 52(2): 222-230. <https://doi.org/10.11896/jsjcx.240600081>

[大语言模型驱动的多元关系知识图谱补全方法](#)

Large Language Model Driven Multi-relational Knowledge Graph Completion Method

计算机科学, 2025, 52(1): 94-101. <https://doi.org/10.11896/jsjcx.240600170>

[一种基于知识图谱的检索增强生成情报问答技术](#)

Retrieval-augmented Generative Intelligence Question Answering Technology Based on Knowledge Graph

计算机科学, 2025, 52(1): 87-93. <https://doi.org/10.11896/jsjcx.240900064>

[SWARM-LLM:基于大语言模型的无人集群任务规划系统](#)

SWARM-LLM: An Unmanned Swarm Task Planning System Based on Large Language Models

计算机科学, 2025, 52(1): 72-79. <https://doi.org/10.11896/jsjcx.241000038>

基于自然语言增强的签到轨迹与用户匹配方法

王天一 林友芳 贡乐天 陈炜 郭晟楠 万怀宇

北京交通大学计算机与信息技术学院 北京 100044

交通数据分析与挖掘北京市重点实验室 北京 100044

(wangtianyi@bjtu.edu.cn)

摘要 随着定位技术和传感器的高速发展,用户移动轨迹数据日渐丰富,但大多分散在不同平台上。为了全面利用这些数据并准确反映用户的真实行为,对轨迹用户匹配的研究变得至关重要。该任务旨在从海量签到轨迹数据中精准关联用户身份。近年来,研究者尝试运用循环神经网络、注意力机制等方法深入挖掘轨迹数据。然而,当前方法在处理用户签到轨迹时面临两大挑战:一是签到数据中有限的时空特征不足以从主观和客观两个角度全面地建模签到点信息,二是用户的签到轨迹往往围绕着一个特定的主题。针对这两点挑战,提出了一种基于自然语言增强的轨迹用户匹配模型(Natural Language Augmented Trajectory User Link, NLATUL)。首先,设计了一套自然语言模板与软提示令牌来描述签到轨迹,并使用语言模型来理解签到点中的主观意图,融合用户的时空状态,提供了一种充分从主观与客观两个方面建模签到点的方法;在此基础上,通过提示学习的方法推理签到轨迹的主题,并对建模的签到点表示的轨迹进行双向编码,通过签到轨迹主题与签到轨迹编码的结合实现对用户签到轨迹的准确理解。在两个真实世界签到数据集上验证的实验结果表明,NLATUL能够更准确地匹配签到轨迹与其对应的用户。

关键词: 轨迹用户匹配;签到序列学习;时空数据挖掘;语言模型;提示学习

中图分类号 TP391

Check-in Trajectory and User Linking Based on Natural Language Augmentation

WANG Tianyi, LIN Youfang, GONG Letian, CHEN Wei, GUO Shengnan and WAN Huaiyu

School of Computer and Information Technology, Beijing Jiaotong University, Beijing 100044, China

Beijing Key Laboratory of Traffic Data Analysis and Mining, Beijing 100044, China

Abstract With the rapid development of positioning technology and sensors, user movement trajectory data is becoming increasingly abundant but scattered on different platforms. In order to fully utilize these data and accurately reflect users' real behavior, the study of trajectory user linking has become crucial. This task aims to accurately correlate user identities from massive check-in trajectory data. In recent years, researchers have tried to use methods such as recurrent neural networks and attention mechanisms to deeply mine trajectory data. However, current methods face two major challenges when processing user check-in sequences. First, the limited spatiotemporal features in the check-in data are insufficient to comprehensively model check-in point information from both subjective and objective perspectives. Second, the topic of the user check-in sequence will affect understanding and modeling check-in sequences. In response to these two challenges, this paper proposes a trajectory user linking model based on natural language augmentation named NLATUL, and designs a set of natural language templates and soft prompt tokens to describe the check-in sequence, and uses the language model to understand the subjective intention in the check-in points, integrating the user's spatiotemporal status, and providing a new perspective and representation that fully models the check-in points from both subjective and objective aspects. On this basis, this paper infer the topic of the check-in sequence through prompt learning, and performs bi-direction encoding on the trajectory represented by the modeled check-in points, so as to achieve an accurate understanding of the user's check-in sequence through the combination of the check-in sequence topic and the check-in sequence encoding, which can link the trajectory with the user more effectively. Verified on two check-in datasets, the experimental results show that proposed method can more accurately link check-in trajectories and their corresponding users.

Keywords Trajectory user link, Check-in sequence learning, Spatiotemporal data mining, Language model, Prompt learning

到稿日期:2024-06-04 返修日期:2024-08-29

基金项目:国家自然科学基金面上项目(62372031)

This work was supported by the National Natural Science Foundation of China(62372031).

通信作者:万怀宇(hywan@bjtu.edu.cn)

1 引言

随着定位技术和各类传感器技术的飞速进步,用户在线应用平台产生了越来越多的移动数据。然而,单一平台下记录的轨迹往往只是用户真实移动轨迹的一部分。为了更全面地利用用户的移动轨迹数据,有必要将不同账户下的移动数据结合在一起,这样能更全面地反映用户移动的真实情况。

在对用户移动数据识别与整合的高度需求下,越来越多的学者开始关注轨迹用户匹配(Trajectory User Link, TUL)这一任务。这项任务的核心在于解决如何在各类定位服务应用所生成的海量移动轨迹数据(主要为用户的签到轨迹数据)中,准确地将轨迹与其真实的用户身份对应起来。

近年来,深度学习领域的繁荣发展为轨迹用户匹配提供了强大的理论支持。许多学者致力于使用循环神经网络、注意力机制等方法对轨迹数据进行深入挖掘,以提升轨迹用户匹配的精度。在基于用户签到数据的轨迹数据挖掘中,当前主流方法往往关注用户签到数据中在时间与空间的状态信息以及用户在时空状态上的转移规律,以此建模用户签到轨迹中的语义信息。然而,目前对用户签到轨迹的理解还面临着以下两点挑战。

首先,对用户签到点的表示往往来源于签到点的记录,这些记录中往往只包含用户签到时间、地点的描述信息,而这些信息都是时空状态的一种记录,只支持从客观的角度建模签到点。那么,对于在时间和空间上相似的签到点(如图1所示),模型可能会错误地将其视为同一类或者相似的签到行为,从而给后续对轨迹的理解带来偏差。因此,从签到点的角度建模用户的签到行为时,不仅需要考虑用户的客观时空特征,对签到点背后主观意图的理解也至关重要,这能够帮助模型从更多的角度更全面地区分不同的签到行为,从而能够更准确地提供对签到点的理解。

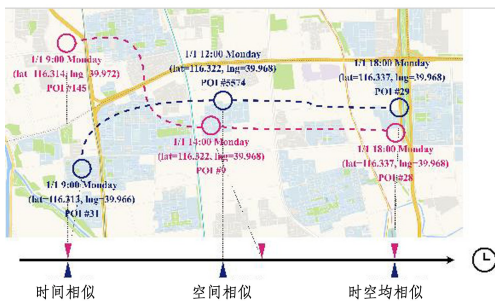


图1 时空特征不能充分描述签到点

Fig. 1 Temporal and spatial features are insufficient to describe check-in point

其次,用户的出行范围和习惯往往围绕着特定的主题。例如图2中的两条签到轨迹分别代表“日常工作”主题和“休闲娱乐”主题,用户从家出发,随后由于不同的主题产生了两条不同的轨迹。在不同主题下,用户有不同的签到活动规律,倾向的出行时间和地点也不同。因此,从签到轨迹的角度出发建模用户的行为模式时,需要考虑到签到轨迹背后的主题,从而真正把握住用户的内在需求和行为动机,实现对用户行为更准确的理解和建模。

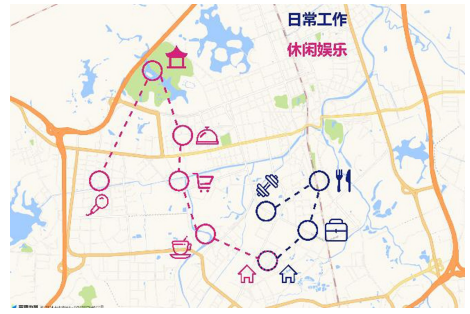


图2 轨迹特征受轨迹主题影响

Fig. 2 Trajectory features are impacted by topic

为了应对上述挑战,本文提出了一种基于自然语言增强的轨迹用户匹配模型。具体来说,通过自然语言这一更具有语义的信息载体来描述签到轨迹,设计了一套自然语言模板与软提示令牌,通过具有强大语义理解能力的语言模型来理解签到点层面的出行意图,并使用合理的方法建模了签到点的时空特征,实现了一种融合用户主观动机与客观状态的签到点建模方法;在此基础上,通过设计针对轨迹主题的提示模板,有效地理解轨迹层面的活动主题,结合一个双向编码器对签到行为转移与规律进行建模。本研究的主要贡献体现在以下几个方面:

1)设计了一套描述签到轨迹的自然语言模板,将结构化的签到数据扩充为非结构化的自然语言,并且注入描述签到节点层面出行意图的软提示令牌,通过语言模型对用户在这些签到点上的主观意图进行理解。随后将时空特征的嵌入与用户的意图相结合,提供了一种理解签到行为的新视角。这一方法能够从签到点的层面同时理解签到行为背后的客观状态与主观意图,从而实现更全面的签到行为建模与理解。

2)基于提示学习对轨迹活动主题进行推理,设计了针对主题的提示模板,有效地建模轨迹层面的活动主题。结合对签到轨迹的双向编码,建模出用户在签到行为上的转移规律,进一步理解整条轨迹,让模型在不同主题下对签到轨迹进行更贴切的理解。

3)基于对签到点的理解和对轨迹主题的推理构造了一个轨迹用户匹配模型,并在两个真实世界签到轨迹数据集中对其进行了验证。实验结果验证了在考虑用户签到意图与出行后,对签到轨迹分析的准确性得到了有效提升。

2 相关工作

2.1 轨迹用户匹配

轨迹用户匹配任务是时空数据挖掘领域中的一个重要研究课题,其目标在于根据用户生成的时空轨迹数据,精确地识别并关联用户与其产生的轨迹。随着大数据和人工智能技术的发展,这一任务的重要性日益凸显,大量学者投入到对该任务的学习中。

由于这项任务是一个针对轨迹的任务,因此,早期学者们采用了一系列基于规则的方法来解决轨迹用户匹配问题,如动态时间规整^[1]、最长公共子序列^[2]和马尔可夫链^[3]等,这些方法被用来衡量轨迹的相似性,实现轨迹与对应用户的匹配。然而,基于规则的方法往往具有一定的局限性,对于复杂多变的用户行为模式适应性有限。

近年来,随着深度学习等机器学习技术的进步,越来越多的研究者开始探索更加灵活和自适应的学习策略。得益于越来越充分的轨迹数据支持,学者们广泛使用深度学习对用户与轨迹进行建模。TULER^[4]起初用词嵌入的方式建模签到点的时空信息,随后使用RNN^[5],LSTM^[6]与GRU^[7]来建模轨迹的时空转移模式,将其链接到具体的用户;进而,TULVAE^[8]使用变分自编码器^[9]来建模出行语义的层次关系,并将其与多层循环神经网络结合来提升轨迹用户匹配的准确率;DeepTUL^[10]使用注意力机制和循环神经网络来建模用户签到中多周期的规律。研究者还使用其他网络方法利用签到数据的丰富信息,如GNNTUL^[11]通过图神经网络学习用户的活动信息,并通过分类器有效地将轨迹与用户匹配;TULSN^[12]利用孪生网络来建模轨迹间的语义信息;CACSR^[13]通过自监督对比学习来缓解数据稀疏与噪声带来的影响,提出的困难样本生成方法有效地提升了签到轨迹中时空信息的建模能力。

目前轨迹用户匹配通常依赖于对签到轨迹中的时空特征和转移规律的深入挖掘。这些方法通过循环神经网络或图神经网络等模型来模拟用户的移动行为习惯和活动模式,并基于此进行有效的用户身份与轨迹匹配。然而,用户的出行意图在签到数据背后起着至关重要的作用,它包含了目的性、偏好性以及临时决策等因素,而这些因素可能并未充分体现在简单的时空坐标和顺序关系之中。本工作着重关注并尝试解决这个问题,即如何将用户的出行意图融入到轨迹建模过程中,使得系统不仅能够基于时空模式进行预测,更能根据用户的意图做出更为精准判断。

2.2 语言模型在数据挖掘中的应用

语言模型在近几年的发展突飞猛进,从早期的BERT^[14]和GPT^[15]系列的预训练语言模型开始,这些研究不仅极大地提升了机器理解、生成和处理自然语言的能力,还为后续语言模型在更多领域的革新与应用奠定了坚实的基础。

语言模型对自然语言深层次语义结构的理解能力被广泛应用于各类专业领域,比如在时空数据挖掘中,使用语言模型对城市进行理解,然后进行流量预测^[16];在金融市场上,可以利用语言模型来构建市场情绪模型,通过分析海量文本数据捕捉投资者情绪的变化^[17];在推荐系统中,则可通过用户评论、社交网络动态等内容来精细建模用户的个性化偏好^[18]。

本研究充分利用自然语言模型对语义的建模能力,针对轨迹用户匹配这一具体任务,将用户出行行为背后的意图进行深度挖掘和精确表达,以此提升轨迹用户匹配任务中对用户出行意图识别的准确性和精细化程度。

3 问题定义

本章给出形式化的问题定义,描述了问题的输入、输出内容。

定义 1(签到点 (t, l, c)) 签到点通常至少包含一个地点标签 l 和一个时间戳 t ; c 代表签到点中包含的额外信息(如经纬度、地点名称、类别等)。

定义 2(签到轨迹 $S = \langle (t_1, l_1, c_1), (t_2, l_2, c_2), \dots, (t_n, l_n, c_n) \rangle$) 签到轨迹是用户 u 按照时间顺序先后访问的 n 个签到点的轨迹。

定义 3(轨迹用户匹配) 给定一个匿名用户的签到轨迹 S 和一个用户集合 $U = \{u_1, u_2, \dots, u_m\}$, 目标是学习到一个映射规则来匹配轨迹与用户。

4 基于自然语言的签到点表示

本章主要介绍基于自然语言的签到点理解方法,首先介绍将签到数据模板化转换为自然语言的过程;随后阐述如何使用软提示令牌机制,通过语言模型来建模节点层面的意图的方法;最后介绍了签到点中时空特征的建模方法与融合出行意图的签到点表示方法。图3为签到点表示的流程框架图。

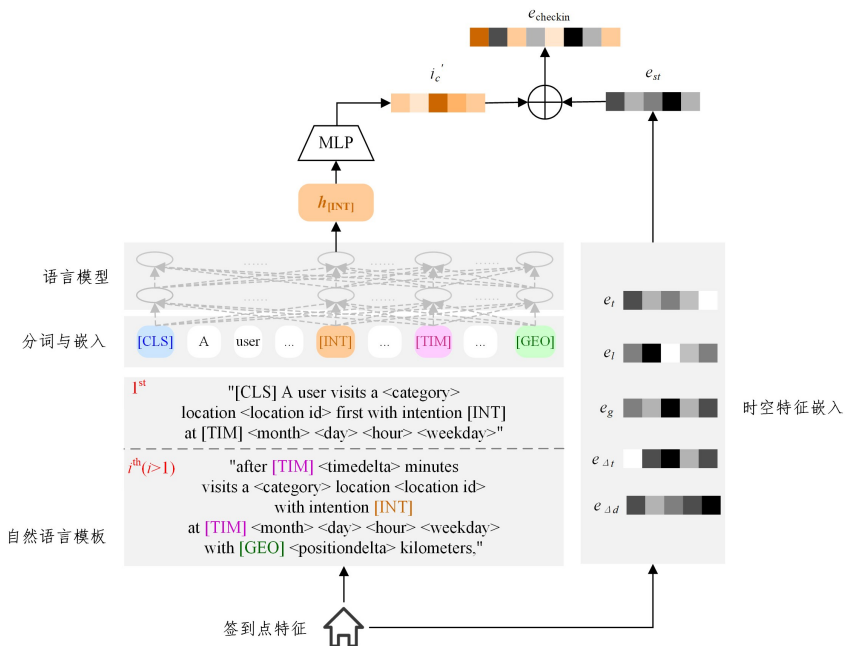


图3 签到点表示流程框架图

Fig. 3 Process framework of check-in point representation

4.1 签到点的自然语言模板构造

首先,用户的签到点数据通常为结构化的表格数据,包括签到点的时间戳、签到地点、位置以及一些额外的描述信息,这些离散或数值型特征如果不加以有效的转换,很难被用来表示用户出行意图,往往会丧失一部分语义表示的精程度。考虑到自然语言模型强大的语义理解能力,将每个签到点按照模板描述为自然语言的一句话,将用户的地理位置、地点类型等富有语义的信息转换为文本描述。这样的形式能够提高语义表达能力,从而易于通过自然语言模型来捕获与理解用户签到数据中的出行意图,建模签到点中的主观信息。

为了在自然语言描述中更精细化地标注和表达时空信息,我们通过软提示^[19]的思想,专门设计了一系列附加的软提示令牌(Soft Prompt Tokens)。在轨迹用户匹配这项任务中,首先需要关注用户的时间与空间状态,于是设计了[GEO]与[TIM],后面连接用户的空间与时间特征(时间、地点、时间跨度与距离变化),这样能够在描述用户签到轨迹时让语言模型更好地建模时空语义信息;对于意图与主题,我们充分地利用语言模型优秀的泛化能力,设计了[INT]与[TPC]两个令牌,旨在让模型针对签到点意图和轨迹主题进行理解与提炼对应特征。

软提示令牌的具体构成与用处已在表1中详细列出。这些特殊设计的提示令牌作为一种提示信号,被整合进文本模板的不同部位,旨在明确地标记和引导模型准确领悟每一个句子或词汇组合所承载的特定时空含义和上下文关联。通过这种方式,模型在处理和理解文本数据时,能够更加敏锐地捕捉到其中包含的多种语义信息(如时间、地点等)。

之后,我们设计了一种自然语言模板,将签到点的时空信息、地点类型和转移信息以及软提示令牌相结合,以更准确地表达用户签到轨迹的详细内容。具体来说,对于每一个独立的签到点,都将其特征转换为一句自然语言表述。

表1 附加软提示令牌

Table 1 Additional soft prompt tokens

名称	说明
[GEO]	在模板中距离跨度之前的提示,指示之后是一个距离跨度的标志
[TIM]	在模板中时间特征之前的提示,指示之后是一个时间特征,包括固有时间特征和转移时间特征(时间间隔)
[INT]	对签到点意图的提示,指示这里是一个意图的总结,签到点意图表示就是使用这个令牌在句子中的嵌入向量来表示的
[TPC]	对于签到轨迹主题的提示,指示这里是一个对签到轨迹主题的提炼
[CLS]	提示句子开头的令牌,在构造的语言模板中,这个令牌指示签到轨迹开始
[SEP]	提示两个句子的分隔,在构造的语言模板中,这个令牌分割了签到轨迹描述、轨迹意图提示和任务提示
[MASK]	提示这里是一个被掩盖的令牌,这个令牌需要被语言模型预测

例如,第一个签到点可以表示为“A user visits a <category> location <location id> first with intention [INT] at [TIM] <month> <day> <hour> <weekday>”;之后的签到点,由于与前一个有相对关系,因此将时空转移信息也加入模板,表示为“after [TIM] <timedelta> minutes visits a <category> location

<location id> with intention [INT] at [TIM] <month> <day> <hour> <weekday> with [GEO] <positiondelta> kilometers.”。其中,尖括号(如<category>)表示签到轨迹中的特征占位符,它们会根据实际的签到数据动态填充,以便精确地反映签到点的时间属性、空间位置、时空转移和地点类别等信息。特征占位符的具体含义与示例值如表2所列。方括号(如[GEO])表示预先定义好的软提示令牌,用来引导模型理解上述签到轨迹信息的含义。

表2 特征占位符的说明

Table 2 Description of feature placeholders

名称	说明	举例
<location id>	地点映射的编号	0,1,2
<category>	地点类别说明	“Park”“Gym”
<month>	月份	“March”
<day>	日	25
<hour>	小时	10:00
<weekday>	一周中的第几天	“Thursday”
<timedelta>	与上一个签到记录的时间间隔(分钟)	75.5
<positiondelta>	与上一个签到记录的距离跨度(千米)	8.86

通过这样的方式,可以将每个签到点的信息清晰、详尽地组织成一句话,并按照签到发生的先后顺序串联起来,最终生成一段描述整个签到轨迹的自然语言模板。这一过程不仅有助于提升模型对签到数据的理解深度,而且能够在后续的处理中充分利用自然语言模型强大的语义理解和推理能力,进一步精准地建模用户在出行时的意图,从而提高轨迹用户匹配任务的性能表现。

4.2 签到点意图理解

针对签到轨迹中的每一个签到点,可以通过自然语言将其组织为一段描述语句,并通过软提示令牌来提取用户此处签到点的意图。注意,上文定义了对于意图的软提示令牌[INT]。在根据预先设定的模板构造描述每个签到点的自然语言文本时,对应描述每个签到点的句子都会包含一个[INT]令牌,用于提示这里用户签到的意图信息。随后,将构造好的自然语言模板先经过分词器和嵌入层得到每个单词的嵌入向量,然后通过语言模型的编码机制获得整个句子所有单词的隐向量,从而得到语言模型最后一层对于[INT]令牌的隐向量,得到的结果就视为该签到点节点目的的意图表达。

$$e_{[INT]} = \text{Tokenize}([INT]) \quad (1)$$

$$i_c = h_{[INT]} = LM(e_{[INT]}) \quad (2)$$

其中,Tokenize(·)代表分词与嵌入操作, $e_{[INT]}$ 代表[INT]令牌的嵌入, i_c 为用户在签到点的节点意图表示。

对于意图表示,我们将用户意图作为对用户签到点主观信息的一个特征,通过自然语言处理的方式,使用提示令牌[INT]来提炼总结每个签到点在上下文中可能的意图信息,将其当作从主观角度对签到点建模的增强方法,能够使对签到点的建模更加全面。

4.3 签到点时空特征建模

同时,对于每个签到点,将时空特征分为两种模式。其中,固有时空特征是不受轨迹中其他签到点影响的客观时空属性,包括签到点的时间信息 t 、签到地点的编号 l 和签到点经纬度(lat, lng);而转移时空特征反映了签到点在用户轨迹

中的动态变化,包括当前签到点相对于轨迹中的上一个点的时间间隔 Δt 和距离跨度 Δd 。通过整合与建模这两种时空特征,模型能够更全面地理解用户在时间与空间上的状态。

对于不同种类的时空特征,需要设计合理的嵌入与编码方法。对于时间信息 t ,取月 t_m 、日 t_d 、小时 t_h 和星期中的第几天 t_w ,分别经过嵌入并取它们的和,得到一个嵌入向量 e_t ;对于地点编号 l ,通过嵌入得到向量 e_l ;对于经纬度,采用 Geo-hash 编码^[20],将经纬度联合通过网格划分并编号,然后通过奇偶位交叉编码的方式将其转换为一条向量 e_g ;对于转移时空特征,采用场嵌入(Field Embedding)^[21]的方法,对时间间隔和距离跨度分别构造可学习的场向量,然后通过归一化时间间隔(秒)和距离跨度(米)与场向量相乘得到嵌入向量 $e_{\Delta t}$ 和 $e_{\Delta d}$ 。

$$e_{\sigma} = [e_t \parallel e_l \parallel e_g \parallel e_{\Delta t} \parallel e_{\Delta d}] \quad (3)$$

其中, e_{σ} 代表最终构造出来的时空嵌入向量。

4.4 用户签到点的表示

上文得到了描述用户主观意图的 i_c 和描述客观时空状态

的 e_{σ} ,这两者分别来自对非结构化的自然语言的理解和对结构化时空特征的建模,处于不同的语义空间。因此这里使用一个非线性投影层将二者统一至同一语义空间,这样能够消除后续对签到点理解的偏差。

$$e_{\text{checkin}} = e_{\sigma} \parallel i_c' \quad (4)$$

$$i_c' = W_2(\text{ReLU}(W_1 i_c + b_1)) + b_2 \quad (5)$$

其中, W_1, W_2, b_1, b_2 分别代表投影层中两个线性层的权重与偏置。经过语义空间对齐操作后,最终能够得到一个融合用户主观与客观信息的签到点表示 e_{checkin} 来整体、全面地描述签到点的特征,以此提升后续对签到点分析的准确性。

5 签到轨迹建模与结果映射

本章首先介绍基于提示学习的签到轨迹主题推理与表示方法;然后阐述如何运用第3章中得到的签到点表示,通过双向编码捕捉用户签到轨迹中各个签到点之间的关联关系;最后描述如何将编码得到的签到轨迹映射到用户身份的结果中。这部分的整体框架如图4所示。

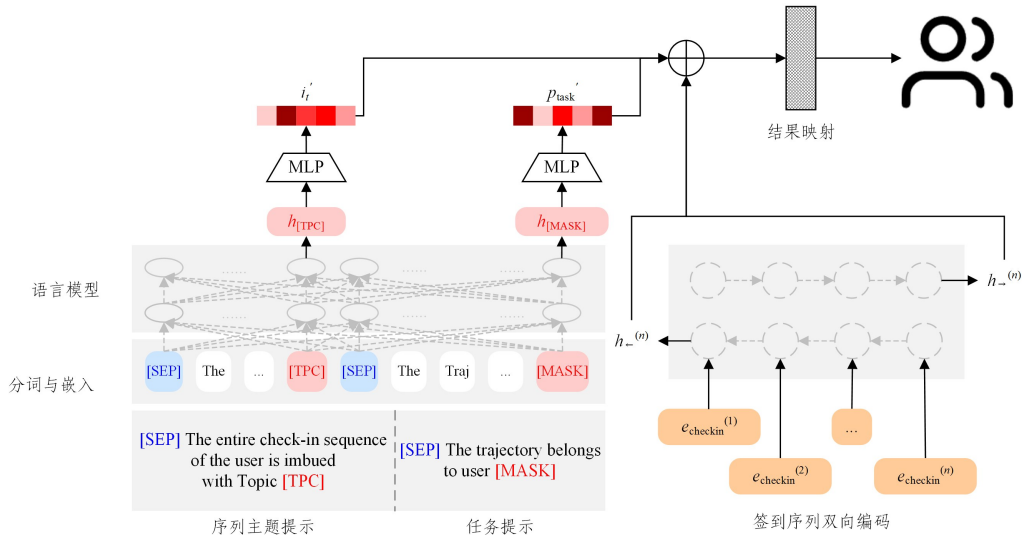


图4 签到轨迹建模与结果映射框架图

Fig. 4 Framework of check-in sequence modeling and result mapping

5.1 签到轨迹主题推理与表示

为了引导模型对整个签到过程中的主题进行推理,我们借鉴了预训练语言模型中用于句子摘要的[CLS]令牌以及提示学习(Prompt Learning)^[19]方法的思想。

首先,根据签到轨迹学习中的特点,结合我们关注的轨迹主题,在句子中设计“check-in sequence”“user”以及“topic”几个关键词,以指导语言模型提炼对于用户签到序列的主题;其次,对于整个提示模板的设计,需要既考虑提示的表达能力,也考虑提示模板的长度,定义设计提示模板时应该简洁而精炼,在尽量控制句子长度的同时,也充分考虑句子的通顺性。另外,对于语言模型,其已经经历了在大规模的数据集上的预训练。基于自然语言学习领域的认知,语言模型已经初步具备归纳总结和阅读理解的能力,并且拥有一定的知识储备,对主题的语义有一定的理解。在本工作中,[TPC]令牌的嵌入会在签到数据中进行微调,以适应签到序列中不同主题的分布,有效地理解并提取轨迹主题。

基于此,我们设计了一段具有引导性的提示文本——“[SEP] The entire check-in sequence of the user is imbued with topic [TPC]”,用以指导语言模型准确捕获并提炼出该签到轨迹的主题。同样,经过分词器、嵌入层、语言模型就可以得到由语言模型提炼的这条签到轨迹的主题,以 i_c 表示。

5.2 签到轨迹的双向编码

基于第3章中构造的一个融合用户主观与客观特征的签到点表示方法,将签到点表示输入至一个双向GRU来编码用户签到轨迹的转移模式。双向编码的好处在于它能够同时考虑签到轨迹中前向和后向的依赖关系,既考虑当前签到点的信息对其前后签到点的影响,也能反映前后签到点对当前点的作用。通过这种方式,模型能够更有效地学习用户在整条轨迹上的行为模式,更好地理解整条签到轨迹。对于一条长度为 n 的签到轨迹,取正向编码最后一个隐藏状态 $h_{+}^{(n)}$ 和反向编码最后一个隐藏状态 $h_{-}^{(n)}$ 的拼接:

$$h = h_{+}^{(n)} \parallel h_{-}^{(n)} \quad (6)$$

$$\mathbf{h}_*^{(n)} = \text{GRUCell}(e_{\text{checkin}}, \mathbf{h}_*^{(n-1)}) \quad (7)$$

$$\mathbf{h}_*^{(n)} = \text{GRUCell}(e_{\text{checkin}}, \mathbf{h}_*^{(n-1)}) \quad (8)$$

最终得到整条轨迹的双向编码表示 h , 它整合了签到轨迹的上下文信息, 提高了对用户签到行为复杂性和多样性的理解能力。

5.3 结合主题与签到轨迹编码的结果映射

在得到对于签到轨迹主题的推理与签到轨迹转移的双向编码后, 需要结合这二者中的信息, 将对签到轨迹的理解匹配到用户的概率分布中。

在实施最终的映射步骤之前, 为了进一步提升模型对当前任务的理解和辨别能力, 在向语言模型输入数据构建自然语言模板时, 我们特意插入了一条关于识别签到轨迹所属用户任务的提示信息, 即“[SEP] the trajectory belongs to user [MASK]”, 以此提升语言模型对目前在做任务的辨别能力。与提取意图理解的步骤相似, 将表示任务提示的[MASK]令牌取出来, 然后把把这个令牌的表示记作 \mathbf{p}_{task} , 将其当作一个附加的任务提示信息。

由于签到轨迹主题表示 \mathbf{i}_i 和任务提示 \mathbf{p}_{task} 都是来自自然语言的, 因此我们采用在签到点意图介绍中的投影层进行语义空间对齐。

$$\mathbf{i}'_i = W_2(\text{ReLU}(W_1 \mathbf{i}_i + b_1)) + b_2 \quad (9)$$

$$\mathbf{p}'_{\text{task}} = W_2(\text{ReLU}(W_1 \mathbf{p}_{\text{task}} + b_1)) + b_2 \quad (10)$$

$$\mathbf{O}_i = \text{Softmax}(W_o(\mathbf{h} \parallel \mathbf{i}'_i \parallel \mathbf{p}'_{\text{task}}) + b_o) \quad (11)$$

$$\hat{u} = \text{argmax}(\mathbf{O}_i) \quad (12)$$

其中, W_o 和 b_o 分别为输出映射层的可学习权重和偏置, $\text{Softmax}(\cdot)$ 用于将签到轨迹经过输出映射的结果转换至用户编号的概率分布空间。

最终通过优化交叉熵损失, 并使用反向传播来更新模型参数, 得到一个能够通过签到轨迹特征匹配用户的模型, 从而达到通过签到轨迹特征有效匹配相应用户的目的。目标函数为:

$$L = \frac{1}{M} \sum_{k=1}^M \sum_{k=1}^N -u_k \log(\hat{u}_k) \quad (13)$$

6 实验

6.1 数据集与设定

为了切实检验 NLATUL 模型在实际场景中的表现和泛化能力, 我们选取了两个真实世界的签到数据集进行实验。这两个数据集均包含了大量用户在现实生活中产生的签到记录, 涵盖了多元化的用户行为、丰富的时空信息以及多变的用户意图。

Gowalla 数据集^[22]包含了用户在 Gowalla 这一地理社交网络上的签到行为记录。Gowalla 曾在 2011 年前后流行, 用户可以在各个地点签到并与朋友分享他们的足迹。

Weeplaces 数据集^[23]是一个来源于同名签到服务的用户签到记录数据集。该数据集包含了用户在全球各地的签到历史、时间戳, 以及社会关系网络数据。

本工作对 Gowalla 和 Weeplaces 两个数据集进行了一定的筛选, 通过一些条件筛选出相对高质量的签到轨迹, 最终得到了两个样本集合, 如表 3 所列。

表 3 数据集统计

Table 3 Statistics of datasets

数据集	用户数量	签到轨迹数量	签到点数量
Gowalla	6838	94012	1103485
Weeplaces	4834	128300	1511335

本工作采用深度学习框架 PyTorch 来实现所提出的模型架构。对于模型参数, 我们最终确定了以下设置: 嵌入维度为 256, 编码器隐藏层维度为 256, 语言模型使用 HuggingFace 平台中的 Tinybert-uncased, 这是一个使用 312 维度词嵌入的预训练语言模型; 在训练时, Batch Size 设定为 32, 学习率为 1×10^{-5} , 以 5 步的早停忍耐度训练 80 轮。所有的实验都可以在一块 24 GB 的 Nvidia A40 显卡中完成。

6.2 评价指标与对比方法

我们采用了两个常见的多分类评价指标来评估性能: k -命中率 (Hit Rate at k , HR@ k) 和平均倒排排名 (Mean Reciprocal Rank, MRR)。

k -命中率 (HR@ k) 是一项衡量在前 k 个预测结果中找到正确用户的概率的指标。具体来说, 对于每一条待匹配的轨迹, 预测出最有可能与之相关的前 k 名用户, 如果这前 k 名中包含了实际对应的真实用户, 则记为本次预测命中。将第 i 次中前 k 个预测结果的命中情况记作 $hit@k(i)$, HR@ k 可以记作:

$$HR@k = \frac{1}{N} \sum_{i=1}^N hit@k(i) \quad (14)$$

平均倒排排名 (MRR) 则是计算模型预测结果中所有查询的平均倒排排名得分。对于每一次查询, 如果模型将正确的用户预测为第一位, 则得分最高为 1; 若预测为第二位, 则得分是 1/2; 依此类推。然后将所有查询的得分求平均, 即可得到 MRR 值。将第 i 次正确的结果位于预测结果中的位次记作 $rank(i)$, 则 MRR 可以记作:

$$MRR = \frac{1}{N} \sum_{i=1}^N \left\{ \frac{1}{rank(i)} \right\} \quad (15)$$

为了验证模型的能力, 我们选择了一些经典与先进的轨迹用户匹配模型进行对比。

TULER^[4]使用了词嵌入的方式建模签到点的时空信息, 随后通过循环神经网络来建模轨迹的时空转移模式。根据循环神经网络的类型, 有 3 种变体: TULER-G, TULER-L 和 Bi-TULER。

TULVAE^[9]使用变分自编码器来建模出行语义的层次关系, 随后通过循环神经网络将其结合来提升轨迹用户匹配的准确率。

DeepTUL^[11]使用注意力机制来发现用户多周期的语义, 随后使用循环神经网络来建模用户签到中语义的转移规律。

CACSR^[13]采用自监督对比学习的手段来学习一个强大的签到轨迹编码器, 通过构造困难本来增强编码器的能力, 随后在下游任务中微调编码器来实现轨迹用户匹配。

6.3 实验结果与分析

在 Gowalla 和 Weeplaces 两个数据集集中的实验结果如表 4 和表 5 所列。其中, 加粗为最优结果, 下划线为次优结果。

表4 Gowalla 数据集上的实验结果

Table 4 Results on Gowalla dataset

	HR@5	HR@20	MRR
TULER-L	0.4955	0.5312	0.4578
TULER-G	0.4916	0.5324	0.4553
Bi-TULER	0.5009	0.5472	0.4699
TULVAE	0.5147	0.5560	0.4725
DeepTUL	0.5332	0.6075	0.4862
CACSR-TUL	0.5598	0.6353	0.5221
NLATUL(C)	<u>0.5810</u>	<u>0.6784</u>	<u>0.5327</u>
NLATUL	0.5948	0.6792	0.5525

表5 Weeplaces 数据集上的实验结果

Table 5 Results on Weeplaces dataset

	HR@5	HR@20	MRR
TULER-L	0.7955	0.8162	0.7031
TULER-G	0.7972	0.8188	0.7012
Bi-TULER	0.8019	0.8264	0.7088
TULVAE	0.8191	0.8327	0.7410
DeepTUL	0.8102	0.8295	0.7194
CACSR-TUL	0.8158	0.8508	0.7379
NLATUL(C)	<u>0.8315</u>	<u>0.8701</u>	<u>0.7623</u>
NLATUL	0.8376	0.8723	0.7685

从实验结果中可以看出,本文提出的模型 NLATUL 与两个数据集中的最佳基线模型相比都实现了一定的提升,这证明了提出的自然语言增强方法能够更有效地建模用户的签到轨迹。

进一步来说,TULER 系列模型能够建模用户的移动模式并且识别出相应的用户,但其只关注时空信息的转移模式,在语义提取上稍有欠缺;TULVAE 在循环结构的基础上增加了变分自编码器的结构,建模出行语义的层次;DeepTUL 通过自注意力机制发现用户多周期的语义结构,有效地提取了用户周期转移的模式;CACSR 使用对比学习预训练一个编码器,有效地提升了签到轨迹建模的能力。在这些基准模型的时间结果中,可以初步看出对于轨迹用户匹配,不同编码器(如循环神经网络、变分自编码器和注意力机制)对结果的影响,尤其可以看出考虑双向编码后在精度上有一定的提升,这证明对于签到轨迹序列来说,需要从不同方向来考虑签到点之间的转移模式。然而,这些基线模型在某种程度上忽略了用户签到的主观意图和签到轨迹主题对轨迹建模的影响。从实验结果来看,理解签到点的主观信息和签到轨迹的主题确实对签到轨迹建模有很大的帮助,有效地提升了模型的性能。

另外,针对签到轨迹,学术研究中主要使用 FourSquare, Gowalla, Weeplaces 这些从位置社交应用上获得的数据集,其中主要包含经纬度、时间戳、地点编号与描述等信息,这些特征本身与语言弱相关。本工作主要是有效地使用自然语言来将这些信息建模,设计了英文框架下的自然语言模板,得到了良好的结果。我们主要关注在建模签到点与轨迹的时候特征反映出的语义信息,如果针对中文,这些信息在中文语言模型下同样能够被有效地提取。对此,我们尝试将所设计的自然语言模板翻译为中文的一段话,使用本文提出的编码签到点和轨迹的方法,在两个数据集上同样得到了有参考意义的结果 NLATUL(C),这证明本文提出的语言模板能够在不同语言下提供对签到轨迹的有效理解。

6.4 消融实验

为了有效地证明模型中各个组成部分的作用,分别单独消除了模型的某些部分,构建出了3种基于原模型的不同变体,旨在从不同角度和层面揭示各部分在模型运行机制中的作用和价值。

w/o ic: 去掉签到点主观意图的理解 i_c , 只使用时空信息来建模签到点。

w/o it: 去掉签到轨迹主题的推理 i_t , 只对建模的签到点进行双向编码。

w/o st: 去掉签到点中时空信息的编码的 e_s , 只使用签到点中用户的主观意图和轨迹主题来建模。

消融实验结果如图5所示。

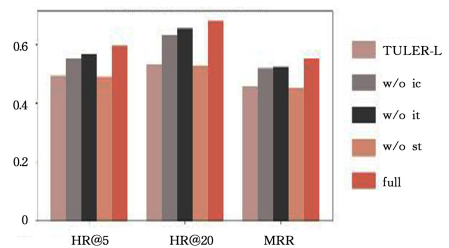


图5 在 Gowalla 上的消融实验结果

Fig. 5 Ablation experiment results on Gowalla

1) 去掉签到点主观信息(w/o ic)

实验结果表明,在去除签到点主观意图理解后,模型对用户行为的匹配精度有了一定程度的下降。我们认为模型此时失去了理解用户在特定地点签到背后的目的和动机的能力,无法有效地区分看似相同但实际意义不同的签到事件。这证明了使用自然语言对签到点建模进行增强的方法能够使模型更有效地理解签到行为,从而提升用户匹配的准确率。

2) 去掉轨迹主题的理解(w/o it)

当移除对签到轨迹主题的理解后,模型的匹配准确率也有了小幅度的降低。我们认为此时模型失去了对用户签到轨迹主题的推理能力,在面对不同主题下的签到轨迹中,模型表现出了较低的适应性和灵活性。这说明在对用户签到轨迹的理解中,准确地推理用户的签到主题能够有效地提升用户匹配的精度。实验结果同时证明了我们提出的基于提示学习的主题理解方法在匹配用户中是有效的。

3) 去掉时空信息(w/o st)

在舍弃时空信息编码的情况下,模型的性能有了比较明显的下降。这是因为时空信息是用户轨迹中最基本、最直接的线索,缺失这部分信息后,模型无法准确追踪用户的空间移动路径和时间规律,即使有主观意图和轨迹主题信息的帮助,也难以弥补缺失客观特征造成的损失。这也说明,对用户签到行为的理解需要同时从主观观上进行考虑。

结束语 本文提出了一种基于自然语言增强的轨迹用户匹配模型 NLATUL。首先,从主客观同时理解签到点,通过自然语言描述了签到轨迹,使用语言模型理解并提取用户在签到点上的主观意图;接着,通过语义空间对齐实现与签到点客观时空特征的结合,得到了一个同时关注签到点的主客观信息的综合表示;然后,通过提示学习的方法推理用户的签到轨迹的主题,将其与对签到轨迹的双向编码结果进行融合;

最后,将编码的结果特征映射到用户集合的概率分布中,实现了对用户签到轨迹的准确匹配。在真实世界数据集 Gowalla 和 Weeplaces 上进行了验证,结果表明,本文提出的模型展现出了较高的用户匹配准确率和较好的排序效果,证实了所提出的自然语言增强方法的有效性。

参 考 文 献

- [1] CHEN L, NG R. On The Marriage of L_p-norms and Edit Distance[C]//Proceedings 2004 VLDB Conference. 2004;792-803.
- [2] DING H, GOCE T, PETER S, et al. Querying and mining of time series data[C]// Proceedings of the VLDB Endowment. 2008;1542-1552.
- [3] RENDLE S, FREUDENTHALER C, SCHMIDT-THIEME L. Factorizing personalized Markov chains for next-basket recommendation[C]// Proceedings of the 19th International Conference on World Wide Web. 2010.
- [4] GAO Q, ZHOU F, ZHANG K P, et al. Identifying Human Mobility via Trajectory Embeddings[C]// International Joint Conferences on Artificial Intelligence(IJCAI 2017). 2019;1689-1695.
- [5] HOPFIELD J J. Neural networks and physical systems with emergent collective computational abilities[C]// Proceedings of the National Academy of Sciences. 1982;2554-2558.
- [6] HOCHREITER S, SCHMIDHUBER J. Long Short-Term Memory[J]. Neural Computation, 1997, 9(8): 1735-1780.
- [7] CHUNG J, GÜLÇEHRE Ç, CHO K, et al. Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling. [J]. arXiv:1412.3555, 2014.
- [8] ZHOU F, GAO Q, TRAJCEVSKI C, et al. Trajectory-User Linking via Variational AutoEncoder[C]// 27th International Joint Conference on Artificial Intelligence and 23rd European Conference on Artificial Intelligence(IJCAI-ECAI 2018). 2018; 3212-3218.
- [9] XU W D, SUN H Z, DENG C, et al. Variational Autoencoder for Semi-Supervised Text Classification[C]// Proceeding of the Thirty-First AAAI Conference on Artificial Intelligence: Twentieth-Ninth Innovative Applications of Artificial Intelligence Conference and Seventh Symposium on Educational Advances in Artificial Intelligence. 2017; 3358-3364.
- [10] MIAO C C, WANG J, YU H, et al. Trajectory-User Linking with Attentive Recurrent Network[C]// International Conference on Autonomous Agents and Multiagent Systems(AAMAS 2020). 2021; 869-877.
- [11] ZHOU F, CHEN S P, WU J, et al. Trajectory-User Linking via Graph Neural Network[C]// 2021 IEEE International Conference on Communications; IEEE International Conference on Communications(ICC). 2021; 1-6.
- [12] YU Y, TANG H, WANG F, et al. TULSN: Siamese Network for Trajectory-user Linking[C]// 2020 International Joint Conference on Neural Networks(IJCNN). 2020.
- [13] GONG L, LIN Y, GUO S, et al. Contrastive Pre-training with Adversarial Perturbations for Check-In Sequence Representation Learning [C]// Proceedings of the AAAI Conference on Artificial Intelligence. 2023; 4276-4283.
- [14] DEVLIN J, CHANG M W, LEE K, et al. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding [C]// Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics. 2019.
- [15] ALEC R, JEFFREY W, REWON C, et al. Language models are unsupervised multitask learners[EB/OL]. <https://openai.com/research/overview>.
- [16] XUE H, FLORA D, REN Y L, et al. Translating Human Mobility Forecasting through Natural Language Generation [C] // Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining (WSDM '22). Association for Computing Machinery, New York, NY, USA, 2022; 1224-1233.
- [17] HU B L. Sentiment Analysis, Stock Price, Text Mining, Stock Market[J]. Statistics and Application, 2021, 10(6): 957-962.
- [18] ZAMANI H, SHAKERY A. A language model-based framework for multi-publisher content-based recommender systems [J]. Information Retrieval Journal, 2018, 21(5): 369-409.
- [19] LIU P, YUAN W, FU J, et al. Pre-train, Prompt, and Predict: A Systematic Survey of Prompting Methods in Natural Language Processing[J]. arXiv:2107.13586, 2021.
- [20] Chris Veness. Geohash[EB/OL]. <https://www.movable-type.co.uk/scripts/geohash.html>.
- [21] GUO H, CHEN B, TANG R, et al. An Embedding Learning Framework for Numerical Features in CTR Prediction [C] // Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining. 2021.
- [22] CHO E, MYERS S A, LESKOVEC J. Friendship and mobility [C]// Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2011.
- [23] LIU Y, WEI W, SUN A, et al. Exploiting Geographical Neighborhood Characteristics for Location Recommendation [C] // Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management. 2014.



WANG Tianyi, born in 2000, postgraduate. His main research interests include spatio-temporal data mining and deep learning.



WAN Huaiyu, born in 1981, Ph.D, professor, Ph.D supervisor, is a member of CCF(No. 17732D). His main research interests include spatial-temporal data mining, information extraction and social networks mining.