

基于时间聚类 and 用户动态相似度的自适应位置推荐算法

朱俊, 韩立新, 宗平, 刘红英, 谢玲, 李景仙

引用本文

朱俊, 韩立新, 宗平, 刘红英, 谢玲, 李景仙. [基于时间聚类 and 用户动态相似度的自适应位置推荐算法](#)[J]. 计算机科学, 2023, 50(12): 113-122.

ZHU Jun, HAN Lixin, ZONG Ping, LIU Hongying, XIE Ling, LI Jingxian. [Adaptive Location Recommendation Based on Time Slots Clustering and User Dynamic Similarity](#) [J]. Computer Science, 2023, 50(12): 113-122.

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)

[面向兴趣点推荐系统的自然噪声过滤算法](#)

Natural Noise Filtering Algorithm for Point-of-Interest Recommender Systems
计算机科学, 2023, 50(11): 132-142. <https://doi.org/10.11896/jsjcx.230400045>

[基于知识增强的命名实体识别方法研究](#)

Study on Named Entity Recognition Method Based on Knowledge Graph Enhancement
计算机科学, 2023, 50(6A): 220700153-6. <https://doi.org/10.11896/jsjcx.220700153>

[基于prompt和知识增强的方面级情感分析](#)

Aspect-based Sentiment Analysis Based on Prompt and Knowledge Enhancement
计算机科学, 2023, 50(6A): 220300279-7. <https://doi.org/10.11896/jsjcx.220300279>

[方面级情感分析综述](#)

Summarization of Aspect-level Sentiment Analysis
计算机科学, 2023, 50(6A): 220400077-7. <https://doi.org/10.11896/jsjcx.220400077>

[命名实体识别任务综述](#)

Overview of Named Entity Recognition Tasks
计算机科学, 2023, 50(6A): 220200119-8. <https://doi.org/10.11896/jsjcx.220200119>

基于时间聚类 and 用户动态相似度的自适应位置推荐算法

朱俊^{1,2} 韩立新² 宗平² 刘红英³ 谢玲³ 李景仙²

1 南京工业职业技术大学计算机与软件学院 南京 210023

2 淮海大学计算机与信息学院 南京 211100

3 南京理工大学紫金学院计算机学院 南京 210023

摘要 位置推荐是位置社交网络中为商家和用户提供服务的一项重要服务,推荐结果易受用户上下文和时空上下文影响。针对当前研究忽略了用户的动态相似度、推荐模型自适应性较弱以及存在严重的数据稀疏问题,提出了一种基于时间聚类 and 用户动态相似度的自适应位置推荐算法(ALRTU)。首先,基于时间槽的签到数据统计特征,对时间进行模糊 C 均值聚类,提取聚类内的时间相似度,利用平滑技术更新原始评分矩阵,以解决数据稀疏问题。分别计算用户在不同时间槽的动态相似度,根据目标时间段所属的时间聚类自适应选择不同的评分数据集,完成用户偏好和时间特征挖掘。其次,根据用户的访问频率特征,为活跃用户和非活跃用户自适应选择核密度估计或幂律分布模型,完成地理特征挖掘。最后,融合用户、时间和空间上下文的综合影响完成位置推荐。在两个真实的位置社交网络数据集 Brightkite 和 Gowalla 中开展准确度评估实验,实验结果表明,与基准方法中最高推荐精度相比,ALRTU 算法在 Brightkite 和 Gowalla 数据集中的准确度仍分别平均提高了 3.74% 和 1.42%。

关键词: 位置推荐; 自适应推荐; 时间聚类; 动态相似度; 空间特征

中图法分类号 TP181

Adaptive Location Recommendation Based on Time Slots Clustering and User Dynamic Similarity

ZHU Jun^{1,2}, HAN Lixin², ZONG Ping², LIU Hongying³, XIE Ling³ and LI Jingxian²

1 School of Computer and Software, Nanjing Vocational University of Industry Technology, Nanjing 210023, China

2 College of Computer and Information Engineering, Hohai University, Nanjing 211100, China

3 College of Computer Science, Nanjing University of Science and Technology Zijin College, Nanjing 210023, China

Abstract Location recommendation is an important service for businesses and users in location-based social networks, and the recommended results are greatly influenced by user preferences and spatial-temporal contexts. Most existing researches ignore the variation in user similarity over time, lack adaptability when making recommendations, and suffer from serious data sparsity problem. To address the issues above, this paper proposes an adaptive location recommendation algorithm(ALRTU) based on time slots clustering and user dynamic similarity. First, time slots clustering based on fuzzy c-means is devised according to the statistics of historical check-in data. Time similarity in each time cluster is calculated, and original ratings are updated further by applying smoothing technology to solve the problem of data sparsity. The dynamic similarities of users are calculated by hour. Different rating subsets are selected adaptively according to time slots clustering and then realize user preferences and temporal influences mining. Second, users are classified on the basis of check-in frequency, and then kernel density estimation or power law distribution algorithm is selected adaptively to mine geographical features. Finally, user preferences and spatial-temporal contexts are effectively fused to produce location recommendations. Extensive offline experiments are conducted on two real-world datasets (Brightkite and Gowalla) to verify the accuracy. Experimental results show that the accuracy of ALRTU on Brightkite and Gowalla datasets is respectively improved by 3.74% and 1.42% on average, compared with the highest recommendation accuracy among the benchmark methods.

Keywords Location recommendation, Adaptive recommendation, Time slots clustering, Dynamic similarity, Spatial features

到稿日期:2023-02-16 返修日期:2023-05-23

基金项目:国家自然科学基金(41771251);江苏省高校自然科学基金项目(21KJB520009);南京工业职业技术大学引进人才科研启动基金(YK23-05-01)

This work was supported by the National Natural Science Foundation of China(41771251), Natural Science Foundation of the Higher Education Institutions of Jiangsu Province, China(21KJB520009) and Start-up Fund for New Talented Researchers of Nanjing Vocational University of Industry Technology(YK23-05-01).

通信作者:朱俊(zj_zijin@163.com)

1 引言

移动通信和定位技术的成熟促进了基于位置的社交网络(Location-based Social Networks, LBSNs)的飞速发展,同时,日益积累的海量信息也增加了用户准确获取目标项目的难度。近年来,致力于解决信息过载问题、以优化位置服务效率为目标的位置推荐系统受到了广泛关注。

与传统推荐系统聚焦于非地理特征的项目(如音乐、电影、笑话等)不同,位置推荐系统的推荐对象是具有经纬度标识的地点,用户对位置的喜爱程度往往会受到距离远近的影响^[1]。例如,用户的活动范围大多数集中在家和工作单位附近;用户在游玩旅游景点时,更愿意去附近的酒店居住。此外,位置推荐是一项对时间特征非常敏感的任务^[2]。从宏观角度来看,用户对位置的喜爱会受到周围大时间环境的影响,例如,美团平台在冬至日为用户推荐水饺店,携程网在夏季为用户推荐水上乐园;从微观角度来看,细粒度的时间影响更能体现用户在特定时间段的签到偏好^[3],例如,餐饮类的位置在12点和18点左右被访问量最大,酒吧的流行度则从21点以后开始上升。

基于以上事实规律,已有部分研究在位置推荐的过程中考虑时空上下文的影响,引入用户签到的时间维度,将位置推荐系统的输入从“用户-位置”二维矩阵拓展为“用户-时间-位置”三阶签到张量,将推荐结果细化为“在什么时间访问什么位置”,从而挖掘时间特征对用户签到行为模式的影响^[4-5]。目前,基于时间感知的位置推荐技术大多存在以下问题:

(1)严重的数据稀疏问题。为了探索用户在目标时间段的行为模式,需要将本就稀疏的签到数据集根据时间轴进一步划分成若干个子集,这无疑会加剧评分矩阵的稀疏程度^[6]。

(2)忽略了用户相似度的时间维动态特征。基于用户的协同过滤算法(User-based Collaborative Filtering, UBCF)是位置推荐中挖掘用户偏好的常用技术^[7-8]。现有研究在计算用户相似度时,不同时间段共享同一个相似度矩阵。然而,用户相似度在一天24小时内会随着时间发生变化。例如,用户中午常与同事一起访问单位附近的餐厅,此时用户与同事之间的相似度高于其与家人之间的相似度,而在下班回家以后,用户常与家人一起访问住址附近的超市,此时用户与家人的相似度更高。因此,在不同时间段使用全局用户相似度是不符合事实规律的。

(3)推荐模型的自适应性较弱。大多数位置推荐方法缺乏自适应机制:首先,在时间维度,无法根据当前时间段的访问特征自适应选择不同的评分数据;其次,在空间维度,忽略了为不同活跃程度的用户自适应选择不同的地理特征挖掘策略。这降低了推荐算法的个性化程度和预测准确度。

针对以上不足,本文提出了一种基于时间聚类 and 用户动态相似度的自适应位置推荐算法。该算法将时间属性与协同过滤推荐模型相结合,在提取用户签到偏好的同时,挖掘时间特征影响;另一方面,根据用户的活跃程度自适应选择不同的签到概率估计模型,挖掘地理特征影响。本文的主要贡献如下:

(1)基于签到数据统计特征对时间槽进行模糊 C 均值

聚类,共享聚类内部评分数据,基于平滑技术更新原始评分,既充分挖掘了聚类内的时间相似性,又缓解了三维评分矩阵的数据稀疏问题。

(2)将时间变量作为用户相似度计算的一个重要的输入参数,按时间槽逐个计算用户的动态相似度;融合时间影响改进传统 UBCF 的评分方式,使其能够根据时间槽自适应选择不同的评分数据,提取用户偏好和时间特征影响。

(3)为活跃用户和非活跃用户分别自适应选择核密度估计和幂律分布模型,挖掘地理特征影响。

(4)在真实的位置社交网络数据集 Brightkite 和 Gowalla 中开展了大量离线实验,实验结果表明,本文算法的推荐准确度优于其他基准算法。

2 相关工作

2.1 时间影响感知

考虑时间影响的位置推荐方法主要包括基于时间槽的方法、基于时间周期的方法以及基于时间序列的方法三大类^[9]。本文采用最常见的基于时间槽的方法提取时间影响。在该类方法中,通常将一天24个小时均匀地划分为若干个时间段,每个时间段称为一个时间槽(Time Slot),通过分析用户在不同时间槽上的偏好差异,为用户提供符合该时间段特征的位置推荐列表。大部分研究将一天平均划分为24个时间槽^[10-12]。Yuan等^[10]基于时间相似度和平滑技术对原评分矩阵进行更新,将时间信息与基于用户的协同过滤技术相结合,量化时间上下文对用户签到行为的影响。Si等^[11]基于用户在各个时间间隔访问相同地址的次数,计算不同时间间隔的访问概率,缓解了数据稀疏问题。Chen等^[12]利用灰色关联分析法度量时间相似性,结合空间地理特征和位置流行度研究用户潜在偏好,填充稀疏矩阵,进而执行矩阵分解算法,提取时间影响。

在基于时间槽的方法中,时间段长度的选择较为灵活,也可以根据数据集的稀疏程度或推荐方法的运行机制而定。例如,Zeng等^[13]将24个小时分为12个时间槽,将时间属性与基于项目的协同过滤相结合进行位置推荐。Yin等^[14]将一天分为4个时间槽,统计用户在不同时间、空间出现的频率,基于用户行为模式提取用户特征向量,构建基于模糊 C 均值聚类的协同过滤推荐模型。

2.2 地理特征挖掘

待推荐项目的地理特征是区别位置推荐与传统推荐的一个重要因素。Tobler提出的地理学第一定律^[15]认为:任何事物都是与其他事物相关的,且距离相近的事物关联更紧密。为了利用地理影响改进位置推荐的准确性,研究者们尝试分析用户历史签到点的经纬度分布,建立概率模型,估计用户在未访问位置上的签到概率。Ye等^[16]通过分析不同用户的签到行为模式,证实了位置社交网络中用户的签到到活动存在地理聚类现象。两个位置的地理距离越近,同一个用户访问它的概率就越大,用户在不同地点的签到概率符合幂律分布(Power-Law Distribution, PLD)。文献^[17-19]认为用户的签到往往围绕几个中心展开,并利用高斯分布(Gaussian Distribution)进行未访问位置的签到概率估计。与幂律分布和高斯

分布对所有用户建立统一的概率模型不同,Zhang等^[20]认为地理特征对每个用户签到行为的影响是个性化的,并提出利用核密度估计(Kernel Density Estimation, KDE)方法实现个性化概率密度建模。由于核密度估计方法能够拟合用户签到的个性化分布,因此,该方法已被很多研究者^[21-23]采用。Si等^[24]在分析地理影响时,将用户聚类为活跃用户和非活跃用户,并分别利用核密度估计和幂律分布为两类用户建立签到概率模型。基于此,本文按照访问频率将用户分为两类:访问频率高于平均水平的用户和访问位置较少的用户,并分别为其自适应选择 KDE 和 PLD 模型,以完成地理特征挖掘。

2.3 用户相似度计算

基于用户的协同过滤算法是位置推荐中挖掘用户偏好的常用技术,其中,用户相似度计算是 UBCF 算法的关键。部分研究^[5,25-27]直接采用传统 UBCF 算法中的相似度计算方法,在用户-位置二维评分矩阵中利用 Jaccard 系数^[5]或余弦相似性^[25-27]计算用户的相似度。部分研究^[28-32]利用时间影响改进传统的相似度计算方法。文献^[28]根据签到到相同位置的数量过滤相似用户集合,根据用户在各时间槽的访问向量计算时间相似度,并基于连续时间槽计算用户相似度。文献^[29]根据用户的总签到次数和时间间隔计算时间相关性,仅对值为 0 的原始评分进行更新,基于更新后的评分矩阵计算用户的时间相似度,基于用户的签到活跃区域与访问频率计算用户的空间相似度。文献^[30]认为不同用户在相同时间对同一位置的访问频率越高,用户之间的相似度就越高,并根据时间序列特征计算用户相似度。文献^[31]基于用户之间的直接信任关系(共同评分)和间接信任关系(信任传播)计算用户相似度,并将其与时间效应、社交关系和地理特征共同融合至矩阵分解框架中进行推荐。文献^[32]在协同过滤框架中同时考虑用户相似度和位置相似度,并综合利用用户偏好、地理距离和社交关系计算用户相似度。

上述研究均未将时间变量作为用户相似度计算的输入参数,不同的时间槽共享同一个全局相似度矩阵,忽略了用户相似度的时间维动态特征。然而,在实际应用中,与目标用户相似的人群会随着时间的发展而变化。因此,使用统一的用户相似度是不符合事实规律的。

本文将一天划分为 24 个时间槽,逐个时间槽计算动态的用户相似度。在时间维度,根据目标时间所属的时间聚类自适应选择评分数据;在空间维度,根据目标用户的历史签到频率自适应选择不同的地理特征挖掘策略,提高推荐算法的预测精度和个性化服务能力。

3 签到数据特征分析

本章基于真实的位置社交网络数据集 Brightkite¹⁾和 Gowalla²⁾对用户的历史签到记录进行数据分析。为了提高数据分析效率,两个数据集的实验观测范围均设定为美国地区签到记录最丰富的前 3 个县。数据预处理的具体流程和数据集的统计信息将在第 5.1 节中详细给出。

3.1 用户相似度的时间维动态特征

在 Brightkite 和 Gowalla 数据集中分别选出访问位置数量最多的 3 个用户,采用余弦相似性方法计算访问位置数最多的用户在各个时间槽上与其余两位用户之间的相似度,如图 1 所示。

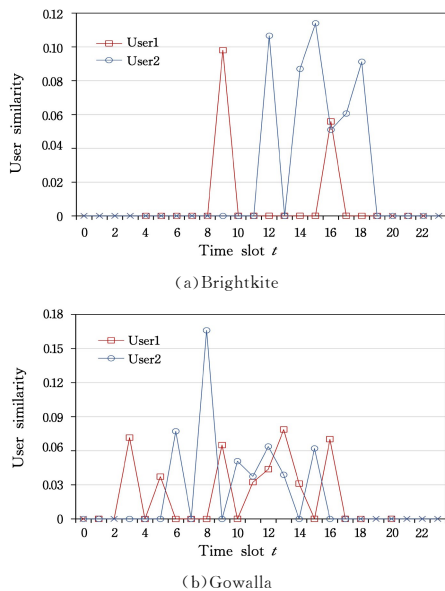


图 1 用户相似度随时间的变化

Fig. 1 User similarity varies over time

图 1 中横坐标轴上的“×”表示某时间槽中两个用户均没有签到记录,无法计算该时间槽的用户相似度。从图 1 中可以看出:(1)用户相似度在不同的时间槽有较大区别,表现为每一幅折线图中的曲线起伏;(2)即使选择了访问位置数最多的 3 位用户进行相似度研究,图中仍有很多时间槽内的用户相似度无法计算。由此可以推断,当为其他用户计算各时间槽中的用户相似度时,评分矩阵的稀疏性带来的难度将会更大。

因此,在挖掘不同时间槽内动态用户相似度的同时,必须尽可能降低稀疏矩阵给相似度计算带来的影响。

3.2 用户签到的空间行为特征

为了探索位置的地理特征给用户签到行为带来的影响,本小节首先计算每个用户在同一天内访问的相邻位置之间的距离,然后统计相邻签到距离的概率分布,如图 2 所示。

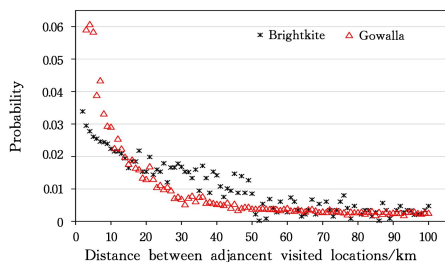


图 2 用户签到的空间行为特征

Fig. 2 Spatial behavior patterns of users' check-ins

从图 2 中可以看出,用户在位置上的签到概率随着地理

¹⁾ <https://snap.stanford.edu/data/loc-Brightkite.html>

²⁾ <http://snap.stanford.edu/data/loc-Gowalla.html>

距离的增大而明显下降。在 Brightkite 数据集中, 25 km 以内的位置签到概率达 51.11%, 超过 50 km 的位置签到概率仅为 16.27%; 在 Gowalla 数据集中, 25 km 以内的位置签到概率达 68.91%, 50 km 以上的位置签到概率仅为 15.66%。由此可见, 地理距离对用户的位置访问有着非常明显的影响。

4 基于时间聚类 and 用户动态相似度的自适应位置推荐算法

基于时间聚类 and 用户动态相似度的自适应位置推荐 (Adaptive Location Recommendation Based on Time Slots Clustering and User Dynamic Similarity, ALRTU) 算法的执行示意图如图 3 所示, 主要包含以下 3 个步骤。

(1) 用户偏好提取 and 时间影响感知。将一天分为 24 个时间槽, 统计每个时间槽内的签到用户数量、被访问位置数量

以及签到次数, 基于以上三阶数据特征对时间进行模糊 C 均值聚类; 提取聚类内的时间相似度, 利用平滑技术更新原始评分矩阵; 计算各时间槽内的用户动态相似度, 基于时间聚类自适应选择不同的评分数据; 对传统 UBCF 算法进行改进, 根据用户的历史签到记录挖掘用户偏好, 生成第一项预评分。

(2) 地理特征挖掘。统计各用户的访问位置个数以及签到次数, 按照访问频率将用户分为两类: 访问频率高于平均值的用户和访问位置较少的用户, 并分别为其自适应选择 KDE 和 PLD 模型, 生成第二项预评分。

(3) 融合用户偏好、时间影响和地理特征的位置推荐。将上述两个步骤输出的预评分相乘, 得到最终预测评分。按最终预测评分对未访问位置进行排序, 为目标用户推荐排名靠前的若干个位置。

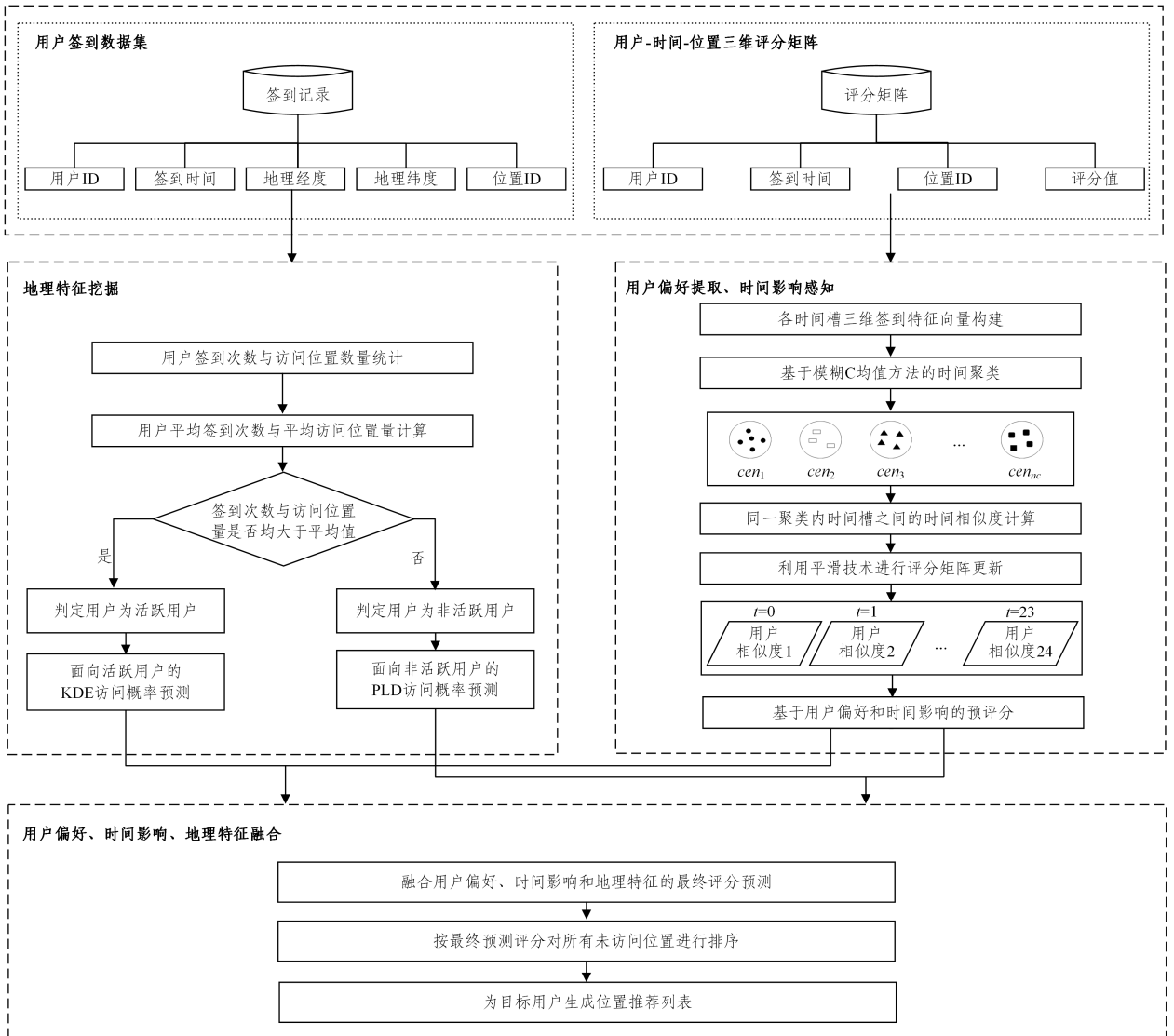


图 3 ALRTU 算法的执行流程
Fig. 3 Workflow of ALRTU algorithm

4.1 用户偏好提取 and 时间影响感知

4.1.1 基于签到数据统计特征的时间聚类

假设签到数据集 C 中共有 n 条签到记录, 记作 $C = \{c_1, c_2, \dots, c_n\}$, 其中, 每一条签到记录均是由用户 ID、签到时间、

地理纬度、地理经度以及位置 ID 五元组表示。 U 是数据集中所有用户的集合, L 是所有位置的集合, NU 和 NL 分别是用户和位置的数量。对签到记录中的所有时间点取整, 将一天的时间划分为 24 个离散的时间槽, 例如, 签到时间 15:13:23

对应的时间槽为 $t=15$, 时间槽集合表示为 $\mathbf{T}=\{0,1,2,\dots,23\}$ 。如果用户 u 在时间槽 t 去过位置 l , 则评分 $r_{u,t,l}=1$; 反之, $r_{u,t,l}=0$ 。由原始签到记录转换生成的用户-时间-位置三维评分矩阵记作 $\mathbf{R}=\{r_{u,t,l}\}$, 其中 $u \in [1, NU]$, $t \in [0, 23]$, $l \in [1, NL]$ 。

统计每个时间槽 $t(t \in [0, 23])$ 的三维签到特征向量 $\mathbf{x}_t = \{Unum_t, Lnum_t, Cnum_t\}$, 其中 $Unum_t$ 表示在时间槽 t 内发生了签到行为的用户数量, $Lnum_t$ 表示在时间槽 t 内被访问的位置数量, $Cnum_t$ 表示在时间槽 t 内发生的签到次数。构建时间槽签到数据特征集合 $\mathbf{X}=\{x_0, x_1, \dots, x_{23}\}$ 。在此基础上, 采用模糊 C 均值方法^[33] 对时间槽进行聚类, 生成 nc 个聚类中心 $\mathbf{Cen}=\{cen_1, cen_2, \dots, cen_{nc}\}$, $nc \in [2, 23]$ 。

4.1.2 基于时间相似度的评分矩阵更新

考虑到同一聚类内的各时间槽具有较为类似的访问特征和关联性, ALRTU 算法逐个计算各时间聚类内任意两个时间槽之间的相似度。设 t 和 t' 均属于聚类 cen_j , 该聚类中的时间槽数量为 n_j , 则 t 和 t' 的时间相似度为:

$$\rho(t, t') = \left(\frac{\sum_{u \in U} \sum_{l \in L} r_{u,t,l} \cdot r_{u,t',l}}{\sum_{u \in U} \sum_{l \in L} r_{u,t,l}^2 \cdot \sum_{u \in U} \sum_{l \in L} r_{u,t',l}^2} \right) / NU \quad (1)$$

为了缓解三维评分矩阵的数据稀疏问题, 基于时间聚类内的时间相似度和平滑技术对原评分矩阵进行更新, 更新后的评分为:

$$r'_{u,t,l} = \frac{\sum_{a=1}^{n_j} r_{u,cen_j[a],t} \cdot \rho(t, cen_j[a])}{\sum_{a=1}^{n_j} \rho(t, cen_j[a])} \quad (2)$$

4.1.3 用户动态相似度计算

从图 1 的分析结论中可以得知, 用户之间的相似度会随着时间的变化而改变。因此, ALRTU 算法将时间变量作为用户相似度计算的一个重要的输入参数, 挖掘 24 个时间槽内用户的动态相似度。为了提高相似度计算的可行性与有效性, ALRTU 基于更新后的评分矩阵执行以下计算。

用户 u 和用户 v 在时间槽 t 时的相似度为:

$$sim(u, v, t) = \frac{\sum_{l \in L} r'_{u,t,l} \cdot r'_{v,t,l}}{\sqrt{\sum_{l \in L} r_{u,t,l}^2} \cdot \sqrt{\sum_{l \in L} r_{v,t,l}^2}} \quad (3)$$

按时间槽逐个计算用户的动态相似度之后, 将产生 24 个 NU 行 \times NU 列的用户相似度矩阵, 即用户 u 和用户 v 在 24 个时间槽内有相应的 24 个不同的相似度值。

4.1.4 基于用户偏好和时间感知的预评分

根据聚类内相似度高、聚类间相似度低的基本原理, 本文合理利用同一聚类内其他时间槽中的评分信息, 改进传统的 UBCF 算法, 根据近邻用户在同一聚类内其他时间槽中的签到历史提取用户偏好, 自适应生成基于时间感知的预评分。在该阶段的推荐策略中, 适应性体现在根据目标时间段的访问特征及其所属的时间聚类, 选择不同的评分数据生成推荐结果。

假设时间槽 t 属于聚类 cen_j , 该聚类中的时间槽数量为 n_j 。根据聚类内的时间相似度以及用户的动态相似度, 目标用户 u_a 在 t 时对位置 l 的预评分为:

$$\hat{r}_1(u_a, t, l) = \frac{\sum_{v \in U, v \neq u_a} sim(u_a, v, t) \cdot \sum_{i=1}^{n_j} r'_{v,cen_j[i],t} \cdot \rho(t, cen_j[i])}{\sum_{v \in U, v \neq u_a} sim(u_a, v, t)} \quad (4)$$

4.2 地理特征挖掘

基于签到数据集 $\mathbf{C}=\{c_1, c_2, \dots, c_n\}$, 分别统计每个用户 u 的签到次数 CN_u 和访问位置个数 LN_u , 计算所有用户的平均签到次数 ACN 与平均访问位置个数 ALN 。若 $CN_u > ACN$ 且 $LN_u > ALN$, 则将用户 u 划分为活跃用户, 否则, 判定用户 u 为非活跃用户。在该阶段的推荐策略中, 适应性体现在为不同活跃程度的用户选择不同的地理影响挖掘算法, 即分别为活跃用户和非活跃用户自适应选择核密度估计 (KDE) 方法和幂律分布 (PLD) 模型^[24]。

4.2.1 面向活跃用户的 KDE 访问概率预测

KDE 方法是一种非参数估计, 仅需收集用户已访问位置的经纬度信息, 无须学习用户历史签到数据的概率分布, 也无需提前设定统一的计算函数, 完全通过一个用户自身丰富的签到数据特征拟合个性化概率密度模型, 挖掘地理特征对用户签到行为的个性化影响^[20-23]。在用户历史访问记录较多的情况下, 核函数相对稳定。因此, 为了充分利用大量的历史签到数据中的位置经纬度信息, 本文为活跃用户选择 KDE 方法进行未访问地址的签到概率预测^[24, 34]。

首先, 为目标用户 u_a 进行距离样本收集, 根据位置经纬度计算 u_a 访问过的每对位置之间的地理距离 d , 形成 u_a 的距离样本集合 \mathbf{X}_a , 通过距离 d 上的概率密度函数 f 估计距离分布。

$$f(d) = \frac{1}{|\mathbf{X}_a| \sigma^d \int_{\mathbf{X}_a} K\left(\frac{d-d'}{\sigma}\right)} \quad (5)$$

其中, σ 是平滑系数, 也称为带宽, $K(\cdot)$ 是高斯核函数。

$$K(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad (6)$$

给定目标用户 u_a 已访问过的位置集合 \mathbf{L}_a , 目标用户 u_a 访问候选位置 l 的概率为:

$$pr(l|\mathbf{L}_a) = \frac{1}{|\mathbf{L}_a|} \sum_{l' \in \mathbf{L}_a} f(d_{ll'}) \quad (7)$$

4.2.2 面向非活跃用户的 PLD 访问概率预测

与活跃用户相比, 非活跃用户的已访问位置相对较少, 因此, 通过分析单个用户为数不多的签到数据特征进行个性化访问概率估计变得难以实现。对于非活跃用户来说, 他们适合将所有已访问位置作为参考点, 而位置之间的距离变得尤为重要。一维的 PLD 模型正是一种基于地理距离的参数估计方法。该方法在假设每次签到位置都是相互独立的基础上, 利用用户的历史签到数据作为先验知识, 学习地理距离的概率分布与幂律参数。此外, 考虑到标度不变性是幂律分布的一个重要属性, 不同数据量的样本会影响函数系数, 但不会改变函数性质。因此, 即使非活跃用户的签到记录样本量较少, 利用幂律分布函数构建的预测模型仍能代表用户的整体访问特征^[34]。基于此, 本文为非活跃用户选择 PLD 模型完成签到概率预测^[24, 34], 既能充分利用有限的所有签到记录 (已访问位置间的距离关系), 又能深刻挖掘用户的整体签到

行为特征。

假设目标用户 u_a 已访问过的位置集合为 L_a , 则目标用户访问候选位置 l 的概率可以用朴素贝叶斯方法表示:

$$\begin{aligned} pr(l|L_a) &= \frac{pr(l \cup L_a)}{pr(L_a)} \\ &= \frac{pr(L_a) \times \prod_{l' \in L_a} pr(l|l')}{pr(L_a)} \\ &= \prod_{l' \in L_a} pr(l|l') \end{aligned} \quad (8)$$

其中 $pr(l|l')$ 表示当用户目前在位置 l' 时, 用户访问候选位置 l 的条件概率。该概率值与两个位置之间的地理距离 $dist(l, l')$ 密切相关。

$$pr(l|l') = \frac{pr(dist(l, l'))}{\sum_{l_k \in L, l_k \neq l'} pr(dist(l', l_k))} \quad (9)$$

$pr(dist(l, l'))$ 是基于地理距离的幂律函数:

$$pr(dist(l, l')) = a \times dist(l, l')^b \quad (10)$$

其中, a 和 b 是幂律函数的两个参数, 可以通过极大似然估计方法^[16]得到这两个参数的数值。

4.3 融合用户偏好、时间影响和地理特征的位置推荐

在获得步骤 1(第 4.1 节)和步骤 2(第 4.2 节)中生成的预评分后, 综合考虑用户偏好、时间影响和地理特征, 计算目标用户 u_a 在 t 时访问位置 l 的最终评分:

$$\hat{r}(u_a, t, l) = \hat{r}_1(u_a, t, l) \times pr(l|L_a) \quad (11)$$

将所有未访问位置按最终预测评分进行排序, 为目标用户推荐排名靠前的若干个位置。

5 实验仿真与结果分析

5.1 实验数据集

本文在两个真实的位置社交网络数据集 Brightkite 和 Gowalla 中进行离线仿真实验。为了将源数据集中的签到记录转换为用户-时间-位置三维评分矩阵, 需要基于原始的签到数据实行一系列必要的预处理操作, 包括观测数据范围的设定、不活跃用户和冷门位置的删除。

首先, 考虑到用户的大多数签到到活动都限定在一定的地理区域内, 过分稀疏的源数据和不足量的数据规模会导致实验结果出现偏差^[21], 因此, 本文将观测范围设定为美国地区签到数据量排名前三的县(county)。在对每个县的访问位置数量和签到次数进行统计和排序后, 分别汇总 Brightkite 数据集中的 Los Angeles, San Francisco, New York 地区和 Gowalla 数据集中 Travis, San Francisco, Los Angeles 地区的签到记录并将其作为观测数据集。

接下来进一步过滤不活跃的用户和访问量较少的位置^[35]。在 Brightkite 的观测数据集中筛选出访问地址数量超过 3 个的用户和至少被 3 个用户访问过的位置, 在 Gowalla 的观测数据集中筛选出访问地址数量超过 20 个的用户和至少被 15 个用户访问过的位置, 形成本文最终的实验数据集。在下文的实验分析和结果展示中, 被标注为“Brightkite”和“Gowalla”的数据集均是指执行完以上预处理操作的数据集。两个实验数据集中的各类统计信息如表 1 所列。

表 1 数据预处理后的数据集统计

Table 1 Statistics of datasets after pre processing

统计项	Brightkite	Gowalla
用户数量	1336	2248
位置数量	2806	2615
签到记录数量	49744	182229
平均每个用户的签到次数	37.23	81.06
平均每个用户的访问位置数量	9.53	46.68
平均每个位置被访问的次数	17.73	69.69
平均每个位置被多少个用户访问	4.54	40.78
评分矩阵的稀疏度	0.996	0.982

5.2 基准方法

考虑到本文的研究方法是针对位置社交网络中的用户偏好、时间影响和地理特征进行挖掘与利用, 因此, 本文选择了以下 7 种从不同角度挖掘用户、位置、时间上下文中一种或几种信息的推荐算法作为基准方法(见表 2), 其中 UBCF, KDE 和 PLD 算法是经典的提取用户偏好和地理特征的单一推荐算法, 是位置推荐研究领域重要的研究基础。其余 4 种算法均是近几年发表的与本文研究方法相关的混合推荐算法。

表 2 各推荐算法利用的上下文种类

Table 2 Contexts utilized by each recommendation algorithm

推荐算法	用户上下文	时间上下文	位置上下文
UBCF ^[36]	✓ (用户偏好)	×	×
KDE ^[21]	×	×	✓ (核密度估计)
PLD ^[16]	×	×	✓ (幂律分布)
UTE+SE ^[10]	✓ (用户偏好)	✓ (时间相似度) (位置流行度)	✓ (幂律分布)
UPT ^[11]	✓ (用户偏好)	✓ (时间间隔内的访问概率) (位置流行度)	×
APRA-SA ^[24]	×	✓ (位置流行度)	✓ (核密度估计或幂律分布)
TeSP-TMF ^[12]	✓ (用户偏好)	✓ (时间相似度) (位置流行度)	✓ (核密度估计)
本文算法 ALRTU	✓ (用户偏好)	✓ (时间聚类) (时间相似度) (用户动态相似度)	✓ (核密度估计或幂律分布)

7种基准算法的具体实现原理如下:

(1)UBCF^[36]:该方法是推荐系统领域经典的协同过滤技术,常常用于用户偏好挖掘。其基于的假设是具有相同历史偏好的用户对同一个位置的兴趣往往类似。

(2)KDE^[21]:该方法是一种无参估计,其利用核密度估计方法实现个性化概率密度建模,能够挖掘地理特征对用户签到行为的个性化影响。

(3)PLD^[16]:该算法是一种有参估计,其利用幂律分布模型挖掘地理特征,通过线性曲线拟合方法学习幂律参数,基于已访问过的位置集合推测目标用户访问候选地址的似然概率。

(4)UTE+SE^[10]:该方法利用平滑技术更新原评分矩阵,将时间影响与UBCF技术相结合来提取时间感知下的签到偏好,利用幂律分布和位置流行度提取空间特征影响。

(5)UPT^[11]:该方法基于时间间隔特征和平滑技术改进UBCF算法,同时考虑位置的总体流行度和时间感知流行度,生成最终预测评分。

(6)APRA-SA^[24]:该方法采用模糊C均值聚类方法将用户划分为活跃用户和非活跃用户两大类,分别采用KDE和PLD模型为活跃用户和非活跃用户挖掘地理影响,结合基于时间的位置流行度得出最终评分。

(7)TeSP-TMF^[12]:该方法利用灰色关联分析法计算时间相似度,基于平滑后的评分矩阵和UBCF技术学习用户潜在偏好;利用KDE方法和位置流行度提取空间地理特征;将以上学习结果(预评分)填入稀疏矩阵,执行矩阵分解算法,建立融入时间特征的推荐模型。

5.3 评估指标

与传统的准确度评估方式不同,本文方法在衡量基于时间感知的推荐算法的准确性时,需要按时间槽逐个较预测评分和真实评分的匹配程度^[10]。

假设 $TestU$ 是所有目标用户的集合,推荐算法在 t 时刻提供给某目标用户 u 的推荐列表长度为 N , $TP(u,t)$, $FP(u,t)$ 和 $FN(u,t)$ 分别是推荐列表中正例对分、负例错分和正例错分的位置数量,则推荐算法在时间槽 t 内的精确率为:

$$precision(t) = \frac{\sum_{u \in TestU} TP(u,t)}{\sum_{u \in TestU} (TP(u,t) + FP(u,t))} \quad (12)$$

推荐算法在时间槽 t 内的召回率为:

$$recall(t) = \frac{\sum_{u \in TestU} TP(u,t)}{\sum_{u \in TestU} (TP(u,t) + FN(u,t))} \quad (13)$$

推荐算法的总体精确率和召回率是各个时间槽中评估指标的平均值:

$$precision = \frac{\sum_{t \in T} precision(t)}{|T|} \quad (14)$$

$$recall = \frac{\sum_{t \in T} recall(t)}{|T|} \quad (15)$$

在得到推荐算法的总体精确率和召回率之后,可按传统方法计算推荐系统的综合准确度F1值:

$$F1 = \frac{2 \times precision \times recall}{precision + recall} \quad (16)$$

5.4 结果对比与分析

本文采用离线方法进行实验分析,将每个数据集随机

划分为训练集和测试集,其中,训练集中包含90%的数据,测试集中包含10%的数据。测试集中的每个用户都是推荐系统的服务对象,本文利用无放回的随机重采样方法选择测试数据集。设定推荐列表长度 N 的取值范围为10~50,所有推荐算法都独立执行50次,最终的准确性度量结果是50次评估结果的平均值。

本文采用模糊C均值方法对时间槽进行聚类,为了确定聚类簇集个数 nc 的取值,在两个数据集中基于 $nc=2 \sim 23$ 运行聚类算法,采用常见的聚类内部评价指标轮廓系数(Silhouette Coefficient,SC)^[37]对聚类结果进行分析比较。图4给出了不同簇集数量对应的SC值曲线。可以看出,在Brightkite和Gowalla数据集中,SC分别在簇集数量为2和3时取得最大值。因此,将Brightkite和Gowalla数据集中的簇集个数分别设置为 $nc=2$ 和 $nc=3$ 。

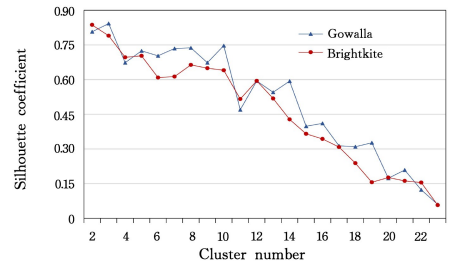


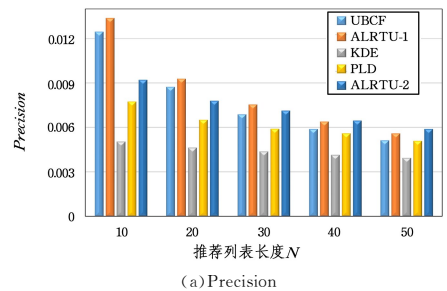
图4 不同簇集数量的轮廓系数值

Fig. 4 Values of SC with different numbers of clusters

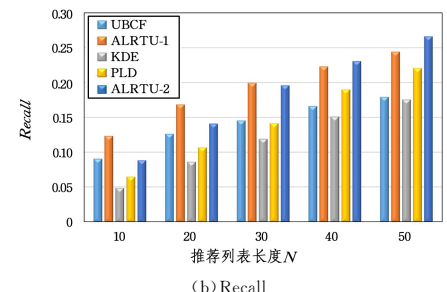
5.4.1 与单一推荐算法的比较

本小节主要对比3种单一算法UBCF,KDE,PLD以及ALRTU算法的两个子步骤的推荐准确度。将本文算法的第一步(用户偏好提取与时间影响挖掘,记作ALRTU-1)与经典的UBCF算法进行比较,将本文算法的第二步(地理特征挖掘,记作ALRTU-2)与KDE和PLD模型进行比较。

图5和图6分别展示了5种推荐算法在Brightkite和Gowalla数据集上的精确率和召回率。



(a) Precision



(b) Recall

图5 第一组算法在Brightkite数据集中的精确率和召回率

Fig. 5 Precision and recall values of the first group of algorithms on Brightkite

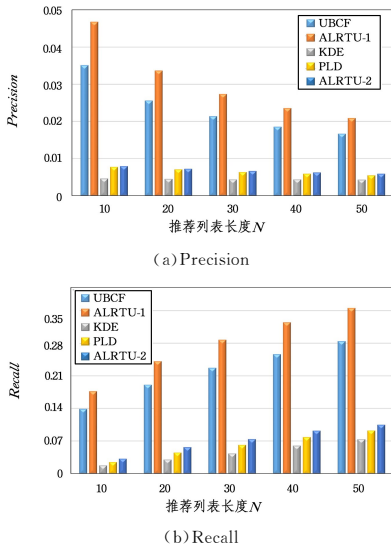


图6 第一组算法在 Gowalla 数据集中的精确率和召回率

Fig. 6 Precision and recall values of the first group of algorithms on Gowalla

表3和表4详细列出了第一组算法在两个数据集上的F1值,其中加粗数值表示该行F1指标的最大值。

表3 第一组算法在 Brightkite 数据集中的 F1 值

Table 3 F1 values of the first group of algorithms on Brightkite

N	UBCF	ALRTU-1	KDE	PLD	ALRTU-2
10	0.0219	0.0240	0.0091	0.0138	0.0166
20	0.0163	0.0176	0.0088	0.0122	0.0148
30	0.0131	0.0145	0.0084	0.0113	0.0138
40	0.0114	0.0124	0.0081	0.0109	0.0126
50	0.0100	0.0109	0.0077	0.0100	0.0115

表4 第一组算法在 Gowalla 数据集中的 F1 值

Table 4 F1 values of the first group of algorithms on Gowalla

N	UBCF	ALRTU-1	KDE	PLD	ALRTU-2
10	0.0561	0.0740	0.0074	0.0118	0.0128
20	0.0452	0.0591	0.0079	0.0122	0.0129
30	0.0391	0.0500	0.0080	0.0116	0.0122
40	0.0346	0.0440	0.0082	0.0110	0.0117
50	0.0316	0.0395	0.0081	0.0104	0.0112

从图表中可以看出:

(1)精确率 Precision 一般随着推荐列表长度 N 的增大而降低,但召回率 Recall 一般都与 N 成正比。

(2)ALRTU-1 算法的 Precision, Recall 和 F1 值均高于同样挖掘用户偏好的 UBCF 算法。与 UBCF 相比,ALRTU-1 在 Brightkite 和 Gowalla 数据集中的 F1 值分别平均提高了 9.40% 和 28.59%。这是因为 UBCF 算法在任何时间给出的位置推荐列表都是一样的,它无法根据具体时间点为目标用户提供不同的、符合时间特征的推荐结果,而基于时间聚类的 ALRTU-1 算法提取了用户的动态相似度,研究了用户在不同时间段的签到模式,因此能够明显提高推荐准确度。

(3)与 KDE 和 PLD 相比,ALRTU-2 算法的推荐准确度更优。在 Brightkite 数据集中,ALRTU-2 算法的 F1 结果在 N 的不同取值下分别比 KDE 和 PLD 平均提高了 63.89% 和 18.88%,在 Gowalla 数据集中,ALRTU-2 的 F1 值则平均提高了 54.28% 和 6.79%,这得益于 ALRTU-2 算法为活跃用户和非活跃用户自适应选择不同的地理影响挖掘算法,个性化的推荐策略提高了推荐结果的准确度。

(4)在 Brightkite 数据集中,当 $N \leq 30$ 时,ALRTU-1 算法的准确度高于 ALRTU-2,但随着 N 的增大,ALRTU-1 的准确率优势逐渐减小,最终略逊于 ALRTU-2 算法,这说明 Brightkite 数据集中地理特征的影响逐渐凸显,当推荐列表长度超过 40 时,地理特征的影响程度比用户偏好和时间的更大。然而,在 Gowalla 数据集中,ALRTU-1 算法的准确度始终明显高于 ALRTU-2,说明用户偏好和时间上下文对用户签到行为的影响大于地理特征。这是因为 Gowalla 数据集的稀疏程度没有 Brightkite 数据集严重,较为丰富的历史访问数据为挖掘不同时间段用户的签到偏好提供了良好的数据基础。

5.4.2 与混合推荐算法的比较

接下来,将 ALRTU 算法与 4 个基于时间感知的混合推荐算法 (APRA-SA, UTE+SE, UPT 和 TeSP+TMF) 进行准确度比较。

图7和图8分别展示了5种推荐算法在 Brightkite 和 Gowalla 数据集上的精确率和召回率。

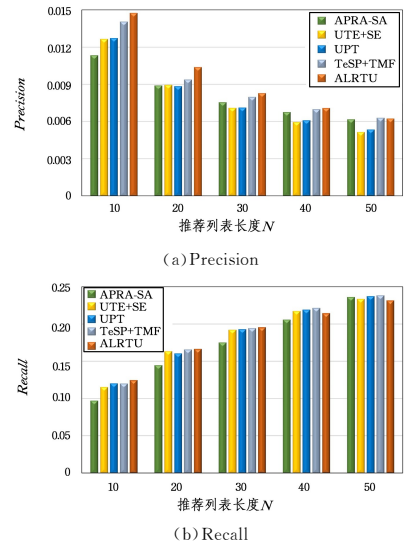


图7 第二组算法在 Brightkite 数据集中的精确率和召回率

Fig. 7 Precision and recall values of the second group of algorithms on Brightkite

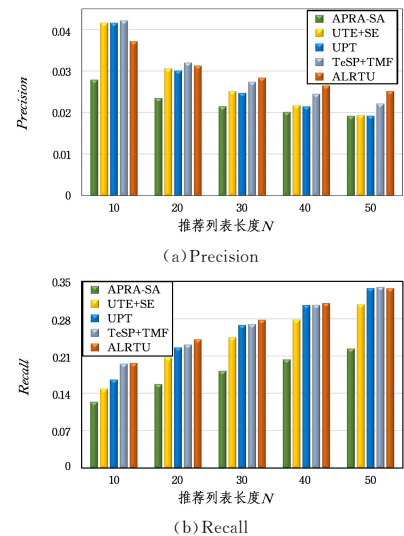


图8 第二组算法在 Gowalla 数据集中的精确率和召回率

Fig. 8 Precision and Recall values of the second group of algorithms on Gowalla

表5和表6详细列出了第二组算法在两个数据集集中的F1值,其中加粗数值表示该行F1指标的最大值。

表5 第二组算法在Brightkite数据集集中的F1值

Table 5 F1 values of the second group of algorithms on Brightkite

N	APRA-SA	UTE+SE	UPT	TeSP-TMF	ALRTU
10	0.0203	0.0228	0.0230	0.0252	0.0264
20	0.0168	0.0169	0.0168	0.0178	0.0196
30	0.0145	0.0137	0.0137	0.0153	0.0159
40	0.0131	0.0116	0.0119	0.0136	0.0137
50	0.0121	0.0101	0.0105	0.0123	0.0121

表6 第二组算法在Gowalla数据集集中的F1值

Table 6 F1 values of the second group of algorithms on Gowalla

N	APRA-SA	UTE+SE	UPT	TeSP-TMF	ALRTU
10	0.0456	0.0650	0.0665	0.0694	0.0626
20	0.0409	0.0533	0.0532	0.0563	0.0554
30	0.0384	0.0456	0.0453	0.0498	0.0516
40	0.0367	0.0403	0.0401	0.0454	0.0489
50	0.0354	0.0365	0.0364	0.0416	0.0469

从结果中可以看出:

(1)除了Gowalla数据集集中的 $N \leq 20$ 之外,其他情况下ALRTU的F1值均高于子算法ALRTU-1和ALRTU-2,这说明用户上下文、时间上下文和位置上下文的融合能够有效提升推荐性能。

(2)在Brightkite数据集中,当推荐列表长度 $N \leq 30$ 时,ALRTU的召回率最高,当 N 超过40时,ALRTU在召回率方面表现不佳;ALRTU的精确率在 $N \leq 40$ 的情况下始终最高,当 $N=50$ 时,TeSP-TMF算法的精确率略超过本文算法。与APRA-SA,UTE+SE,UPT和TeSP-TMF相比,ALRTU的F1值在 N 的5种取值下分别平均提高了12.34%,17.31%,15.88%和3.74%。

(3)在Gowalla数据集中,ALRTU的召回率在推荐列表长度 $N \leq 40$ 时优势明显,当 $N=50$ 时,略逊于TeSP-TMF算法;当 $N=10$ 时,ALRTU的精确率低于UTE+SE,UPT和TeSP-TMF算法,但仍高于APRA-SA算法。随着 N 的增大($N \geq 30$),ALRTU的精确率优势逐渐凸显,最终高于其他算法。与APRA-SA,UTE+SE,UPT和TeSP-TMF相比,ALRTU的F1值在 N 的5种取值下分别平均提高了34.57%,12.52%,12.53%和2.46%。

(4)与两组实验中9种算法(UBCF,KDE,PLD,ALRTU-1,ALRTU-2,APRA-SA,UTE+SE,UPT和TeSP-TMF)的最高准确度相比,ALRTU算法在Brightkite数据集和Gowalla数据集集中的F1值在 N 的不同取值下仍分别平均提高了3.74%和1.42%。

结束语 针对当前位置推荐系统研究忽略了用户相似度的时间维动态特征、推荐模型自适应性较弱以及存在的严重的数据稀疏问题,本文以有效融合用户偏好、时间影响和地理特征等上下文信息为目标,提出了一种基于时间聚类 and 用户动态相似度的自适应位置推荐算法。该算法采用模糊C均值方法对时间槽进行聚类,有效缓解了数据稀疏问题。在时间维度,本文挖掘了用户在不同时间槽的动态相似度,并将其融入协同过滤框架,根据目标时间段的访问特征自适应选择不同的评分数据,有效提取用户偏好和时间影响。在空间

维度,本文为不同活跃程度的用户选择不同的访问概率预测模型,分别挖掘地理特征影响。本文在两个真实的数据集中进行了评估实验,结果表明本文算法在推荐准确度(精确率、召回率、F1值)方面优于其他算法。

下一步我们将继续研究融合多维异构上下文的位置推荐:(1)在推荐算法设计方面,深入挖掘位置流行度及其时效特征,进一步提升推荐准确度;(2)在推荐系统评估方面,将定义并研究位置推荐的时效性评估体系,考虑如何在提高位置推荐准确度的同时,确保位置推荐系统的时效性。

参考文献

- [1] ZHU J, HAN L X, GOU Z N, et al. A robust personalized location recommendation based on ensemble learning [J]. *Expert Systems with Applications*, 2021, 167: 114065.
- [2] XING C Z, ZHU J X, MENG X F, et al. Point-of-interest recommendation: A survey [J]. *Computer Science*, 2021, 48(S2): 176-183.
- [3] WANG Y L, JIANG C C, FENG X N, et al. Time aware point-of-interest recommendation [J]. *Computer Science*, 2021, 48(9): 43-49.
- [4] CHEN W, WAN H Y, GUO S N, et al. Building and exploiting spatial-temporal knowledge graph for next POI recommendation [J]. *Knowledge-Based Systems*, 2022, 258: 109951.
- [5] STEPAN T, MORAWSKI J M, DICK S, et al. Incorporating spatial, temporal, and social context in recommendations for location-based social networks [J]. *IEEE Transactions on Computational Social Systems*, 2016, 3(4): 164-175.
- [6] LIU X, YANG Y J, XU Y B, et al. Real-time POI recommendation via modeling long- and short-term user preferences [J]. *Neurocomputing*, 2022, 467: 454-464.
- [7] NAM L N H. Towards comprehensive approaches for the rating prediction phase in memory-based collaborative filtering recommender systems [J]. *Information Sciences*, 2022, 589: 878-910.
- [8] CHEN Z L, WANG Y, ZHANG S, et al. Differentially private user-based collaborative filtering recommendation based on k-means clustering [J]. *Expert Systems with Applications*, 2021, 168: 114366.
- [9] CHEN J M, ZHANG W D. Review of point of interest recommendation systems in location-based social networks [J]. *Journal of Frontiers of Computer Science and Technology*, 2022, 16(7): 1462-1478.
- [10] YUAN Q, CONG G, MA Z Y, et al. Time-aware point-of-interest recommendation [C] // *Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 2013: 363-372.
- [11] SI Y L, ZHANG F Z, LIU W Y. A time-aware POI recommendation method exploiting user-based collaborative filtering and location popularity [C] // *Proceedings of 2017 2nd International Conference on Communications, Information Management and Network Security*. 2017: 17-25.
- [12] CHEN J M, ZHANG W D. TeSP-TMF: A temporal-aware personalized POI recommendation approach based on potential preferences and grey relational analysis [J]. *Electronic Commerce*

- Research and Applications, 2023, 58:101243.
- [13] ZENG J, LI Y H, LI F, et al. A point-of-interest recommendation method using location similarity[C]//Proceedings of 2017 6th IIAI International Congress on Advanced Applied Informatics (IIAI-AAD). 2017:436-440.
- [14] YIN M H, LIU Y H, ZHOU X, et al. A fuzzy clustering based collaborative filtering algorithm for time-aware POI recommendation [J]. Journal of Physics: Conference Series, 2021, 1746(1):012037.
- [15] TOBLER W R. A computer movie simulating urban growth in the detroit region [J]. Economic Geography, 1970, 46(2):234-240.
- [16] YE M, YIN P F, 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. 2011:325-344.
- [17] CHO E, MYERS S A, LESKOVEC J. Friendship and mobility: user movement in location-based social networks[C]//Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2011:1082-1090.
- [18] CHENG C, YANG H Q, KING I, et al. Fused matrix factorization with geographical and social influence in location-based social networks[C]//Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence. 2012:17-23.
- [19] LIU B, FU Y J, YAO Z J, 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.
- [20] 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.
- [21] GENG B R, JIAO L C, GONG M G, et al. A two-step personalized location recommendation based on multi-objective immune algorithm [J]. Information Sciences, 2019, 475:161-181.
- [22] DAVIS A D, GIANNAKIS D. Graph-theoretic algorithms for Kolmogorov operators: approximating solutions and their gradients in elliptic and parabolic problems on manifolds [J]. arXiv:2104.15124, 2022.
- [23] ZHANG J D, CHOW C Y. CoRe: exploiting the personalized influence of two-dimensional geographic coordinates for location recommendations [J]. Information Sciences, 2015, 293:163-181.
- [24] SI Y L, ZHANG F Z, LIU W Y. An adaptive point-of-interest recommendation method for location-based social networks based on user activity and spatial features [J]. Knowledge-Based Systems, 2019, 163:267-282.
- [25] ZHANG J D, CHOW C Y. TICRec: A probabilistic framework to utilize temporal influence correlations for time-aware location recommendations [J]. IEEE Transactions on Services Computing, 2016, 9(4):633-646.
- [26] CHEN J, ZHANG H, CAO F Y. Study on point-of-interest collaborative recommendation method fusing multi-factors [J]. Computer Science, 2019, 46(10):77-83.
- [27] YIN M H, LIU Y H, ZHOU X, et al. A tensor decomposition based collaborative filtering algorithm for time-aware POI recommendation in LBSN [J]. Multimedia Tools and Applications, 2021, 80:36215-36235.
- [28] SONG Y W, SI Y L, LIU W Y, et al. Point-of-interest recommendation algorithm combining temporal features and collaborative filtering [J]. Journal of Chinese Computer Systems, 2016, 37(6):1153-1158.
- [29] JIANG C Q, SHU D Y, DUAN R. Location recommendation algorithm based on spatial-temporal similarity of user [J]. Computer Engineering, 2018, 44(7):177-182.
- [30] YUAN F Y, LI C, LEI U, et al. POI dynamic recommendation algorithm based on fusion time series [J]. Journal of Chinese Computer Systems, 2020, 41(2):291-295.
- [31] XU CH, DING A S, ZHAO K D. A novel POI recommendation method based on trust relationship and spatial-temporal factors [J]. Electronic Commerce Research & Applications, 2021, 48:101060.
- [32] DAVTALAB M, ALESHEIKH A A. A multi-criteria point of interest recommendation using the dominance concept [J]. Journal of Ambient Intelligence and Humanized Computing, 2023, 14:6681-6696.
- [33] ZHU J, HAN L X, GOU Z N, et al. A fuzzy clustering-based denoising model for evaluating uncertainty in collaborative filtering recommender systems [J]. Journal of the Association for Information Science and Technology, 2018, 69(9):1109-1121.
- [34] SI Y L. Research on adaptive point-of-interest recommendation approaches based on user check-in behaviors [D]. Qinhuangdao: Yanshan University, 2020.
- [35] VALVERDE-REBAZA J C, ROCHE M, PONCELET P, et al. The role of location and social strength for friendship prediction in location-based social networks [J]. Information Processing and Management, 2018, 54(4):475-489.
- [36] HERLOCKER J L, KONSTAN J A, BORCHERS A, et al. An algorithmic framework for performing collaborative filtering [C]//Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. 1999:230-237.
- [37] WANG H Y, WANG J S, WANG G. A survey of fuzzy clustering validity evaluation methods [J]. Information Sciences, 2022, 618:270-297.



ZHU Jun, born in 1987, Ph.D, associate professor, is a member of China Computer Federation. Her main research interests include machine learning and recommender systems.