

融合相似度和地理信息的兴趣点推荐

郭晨睿^{1†}, 李平^{1,2}, 郭苗苗³

(1. 长沙理工大学 计算机与通信工程学院, 湖南 长沙 410114;

2. 智能交通大数据处理湖南省重点实验室, 湖南 长沙 410014; 3. 长沙大学图书馆, 湖南 长沙 410022)

摘 要:兴趣点推荐是一种基于上下文信息的位置感知的个性化推荐。由于用户签到行为具有高稀疏性,为兴趣点推荐的精确度带来了很大的挑战。针对该问题,提出了一种融合相似度和地理信息的兴趣点推荐模型,称为SIGFM。首先利用潜在狄利克雷分配(Latent Dirichlet Allocation, LDA)模型挖掘用户相关兴趣特征并进行相似性度量,利用 Louvain Community Detection (LCD) 算法与用户签到数据进行相似性度量,使两种相似度相融合;然后使用地理信息获取用户的签到特征;最后将融合相似度和地理信息结合到一起获得一个新的模型。在真实数据集上的实验结果表明, SIGFM 模型有效解决了数据稀疏性与冷启动问题,优于其他 POIs 的推荐算法。

关键词:潜在狄利克雷分布; Louvain 社区发现; 兴趣点推荐; 地理信息; 相似度

中图分类号: TP311

文献标识码: A

Points-of-interest Recommendation with Similarity and Geographic Information

GUO Chen-rui^{1†}, LI Ping^{1,2}, GUO Miao-miao³

(1. School of Computer and Communication Engineering, Changsha University of Science and Technology, Changsha, Hunan 410114 China;

2. Hunan Provincial Key Laboratory of Intelligent Processing of Big Data on Transportation, Changsha, Hunan 410114 China;

3. Changsha University Library, Changsha University, Changsha, Hunan 410022, China)

Abstract: Point-of-interests (POIs) recommendation is a personalized recommendation based on location-aware with context information. Owing to behavior of check-in from the users is highly sparse, which poses the challenge to the accuracy of the POIs recommendation. In order to solve this problem, this paper propose a new POIs recommendation called Similarity Integration Geography Fusing Model (SIGFM). Firstly, we exploit an aggregated Latent Dirichlet Allocation (LDA) model to learn the interest feature from the users, and then puts the interest feature into similarity measurement. Also, we use the Louvain Community Detection (LCD) and check-in data from the users to calculate the similarity. The similarity measurement utilizing both methods finally merge into the one. Then, a geographical influence measurement is employed to capture the check-in characteristic from the users. Finally, geographical information in conjunction with the similarity forms the new model. Experimental results show that SIGFM can effectively mitigate the sparse-data usually suffered and the cold-start suffer to outperforms other methods.

Key words: latent dirichlet allocation (LDA); Louvain community detection (LCD); point-of-interests (POIs) recommendation; geographic information; similarity

收稿日期: 2019—02—19

基金项目: 国家自然科学基金资助项目(61502054)

作者简介: 郭晨睿(1992—), 男, 河南周口人, 硕士研究生, 研究方向: 数据挖掘。

† 通讯联系人, E-mail: 709245935@qq.com

1 相关工作

随着移动互联网的快速发展,基于位置的社交网络(Location Based Social Networks,LBSNs)应运而生^[1-7](如 Foursquare 等应用),获得了用户们的欢迎。在 LBSNs 中用户可以在目前访问的 POIs(如:餐厅等)以签到的方式发布他们的地理位置。随着 LBSNs 中 POIs 数量的快速增加,POIs 推荐已成为人们发现新位置的首选方式。该方式有效帮助了 LBSNs 中的用户访问 POIs,并以签到的功能发表评论等相关信息,与其他用户分享自己在该 POIs 的访问体验。POIs 推荐旨在帮助用户更好地发现感兴趣的 POIs,为商家提供精准营销策略。这使得 LBSNs 更具有吸引力,吸引了诸多研究^[8,9]。

与传统的推荐问题(如:电影推荐等)相比,POIs 推荐系统更加复杂,面临如下挑战^[1,2,7]:

(1)丰富的上下文。用户的移动偏好受地理位置的影响:用户通常访问频繁活动区域内的 POIs;用户每天可以访问相同的 POIs;用户的偏好依赖于时间;其他的上下文信息包括 POIs 评论等。

(2)数据稀疏。与传统推荐系统相比,POIs 推荐的数据严重稀疏。POIs 推荐实验研究中使用的数据密度通常在 0.1%左右,而 Netflix 电影推荐数据密度为 1.2%^[7]。

针对上述问题,国内外诸多学者提出了多种解决方法。如文献[9]中 Ye 等人基于朴素贝叶斯开发了基于地理影响的协同推荐算法。文献[10]利用 LDA 模型计算用户的兴趣话题,利用用户签到历史数据和 POIs 的类别信息得到每个话题。文献[11]中 Zhuang 等人将用户签到记录和社会影响力综合于一个组合模型中,然后再利用 PageRank 算法产生的影响因子进一步对其进行改进;应用地理影响捕捉用户的物理签到信息。文献[12]中 Liu 等人利用聚合 LDA 模型提取学习用户的兴趣,通过话题分布提取用户和 POIs 的配置文件并结合 PFM 进行兴趣点推荐。文献[13]任等人结合六个上下文信息(兴趣、文本、地理、社会、分类、流行度)提出一种上下文感知的概率矩阵分解模型。文献[14]中 Zhao 等人将用户兴趣、社会关系的兴趣相似度、社会关系的评分行为习惯和人际评分行为相似性四个因素融入矩阵分解中进行 POIs 推荐。在文献[15]中 Wang 等人通过考虑社会关系提出了一种新的协同过滤算法。

本文提出了一种融合相似度和地理信息的模型,该模型命名为 SimilarityIntegration Geography FusingModel(SIGFM)。思想是:首先,使用聚合 LDA 模型^[12,16]挖掘用户的兴趣特征,并计算用户之间的相似性;其次,利用文献[17,18]中的 Louvain Community Detection(LCD)算法根据用户历史签到数据集计算用户之间的相似度;最后,使用地理信息挖掘用户的地理位置。融合以上三种因素,解决了 POIs 推荐的数据稀疏性问题,提高了推荐的准确率和召回率。

总的来说,本文对兴趣点推荐有以下三点贡献:

1) 利用 LCD 算法对 LBSNs 中的用户进行分区,认为在同一社区内的用户彼此之间存在相互影响,结合用户历史签到数据集,对社区内用户与用户之间的相似性进行度量。对新用户进行 POIs 推荐存在严重的冷启动问题,本文中的 LCD 算法将新用户与其好友分到同一个社区 C 内,在社区 C 内进行相似性度量,从而为新用户进行 POIs 推荐,解决了 POIs 推荐系统中的冷启动问题。

2) 提出一种使用指数函数来计算地理影响的方法。

改善了文献[11]中的部分缺点,文献[11]中使用 POIs 之间的物理距离来计算地理影响,却忽略了如果用户访问过的 POIs 与其未访问过的 POIs 之间距离过近会使得地理影响过大,一方面不利于后续的计算;另一方面会导致社交关系的影响可以忽略不计。

3) 将用户的特征向量、社交关系和地理影响融合成 SIGFM 模型,预测每用户签到新的 POIs 的概率。

2 SIGFM 模型基础

本节介绍 SIGFM 模型的基础,表 1 给出了本文使用的主要符号。

表 1 本文中的主要符号

符号	意义
U	在 LBSNs 上所有用户的集合用户
u, v	用户: $u, v \in U$
L	在 LBSNs 上所有 POIs 的集合
l	某 POI: $l \in L, l = (lat, lon)^T$
$r_{u,l,c}$	属于社区 C 的用户 u 在 POI 签到则 $r_{u,l,c} = 1$, 否则 $r_{u,l,c} = 0$

2.1 基于 LDA 模型的相似度量

本文使用文献[12]中的聚合 LDA 模型,如图 1 所示。首先,文档 d_i 聚集同一个 POIs 相关的文档,文档 d_u 聚集同一个用户 u 的标签及其签到过的 POIs 的文本信息。这样我们得到一个文档集合,一个 POIs 或者一个用户对应一个文档^[12,13,16]。 M 是 POIs l 和用户 u 中隐含的主题数, M 维向量 θ_i 和 θ_u 分别表示 POIs 和用户的主题分布,每一维代表该 POIs l 或用户 u 在相应主题下的概率^[16]。

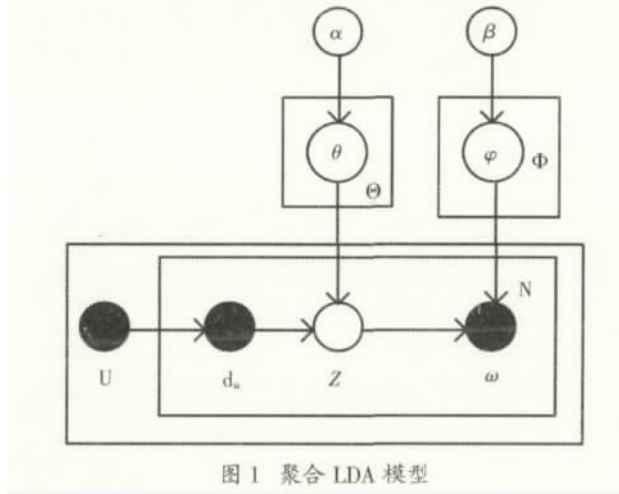


图 1 聚合 LDA 模型

每个用户 u 或 POIs l 的主题特征由聚合 LDA 模型获得, θ_u 表示用户 u_i 的特征向量, θ_l 表示 POIs l_j 的特征向量, $\omega_m^{(u)}$ 和 $\omega_m^{(l)}$ 分别是用户 u_i 和 POIs l_j 归于主题 m 的概率。公式(1)表示基于聚合 LDA 模型计算用户相似度^[16]:

$$\text{sim}_{\text{LDA}}(\theta_u, \theta_l) = \frac{\sum_{m=1}^M (\omega_m^{(u)} * \omega_m^{(l)})}{\sqrt{\sum_{m=1}^M (\omega_m^{(u)})^2} * \sqrt{\sum_{m=1}^M (\omega_m^{(l)})^2}} \quad (1)$$

公式(2)表示基于聚合 LDA 模型计算 POIs 的相似度计算:

$$\text{sim}_{\text{LDA}}(\theta_l, \theta_l) = \frac{\sum_{m=1}^M (\omega_m^{(l)} * \omega_m^{(l)})}{\sqrt{\sum_{m=1}^M (\omega_m^{(l)})^2} * \sqrt{\sum_{m=1}^M (\omega_m^{(l)})^2}} \quad (2)$$

2.2 基于 LCD 算法的用户相似度

文献 [9,11] 中基于用户的相似度仅仅考虑了 LBSNs 中用户与用户之间的直接好友关系,而忽略了可以产生相互影响的两个用户在 LBSNs 中可能不是好友关系。本文使用 LCD 算法^[17]将 LBSNs 中的用户进行分区,认为在同一社区内的用户可以产

生相互影响。

LCD 算法是一种基于多层次的 Modularity^[18]算法,分区的模块化是标量值在-1 和 1 之间,用于衡量社区内部链接密度与社区之间链接的密度。在加权网络中,其定义如下:

$$Q = \frac{1}{2m} \sum_{ij} \left(A_{ij} - \frac{k_i k_j}{2m} \right) \delta(c_i, c_j) \quad (3)$$

其中, A_{ij} 表示节点 i 和节点 j 之间边的权重,网络不是带权图时, $k_i = \sum_j A_{ij}$ 表示所有与节点 i 相连的边的权重之和(度数),若 $u = v$,则 $\delta(u, v) = 1$,否则, $\delta(u, v) = 0$ 。 c_i 表示节点 i 所属的社区, $m = \frac{1}{2} \sum_{ij} A_{ij}$ 表示所有边的权重之和(边的数目)。

如图 2 所示, LCD 算法归纳为两个步骤, 有个 N 节点的加权网络。首先,将网络中的每个节点分配不同的社区;接着,对于每个节点 i 考虑其邻居及节点 j , 并将模块化的收益通过将 i 从它的社区中移除并放置在 j 的社区中来评估;最后将节点 i 放置在该增益最大的社区中^[17]。对所有节点重复并按顺序应用此过程,直到无法实现进一步的改进。通过移动隔离节点 i 获得模块性增益 ΔQ 进入社区 C 可以计算为:

$$\Delta Q = \left[\frac{\sum_{in} + 2k_{i,in}}{2m} - \left(\frac{\sum_{tot} + k_i}{2m} \right)^2 \right] - \left[\frac{\sum_{in}}{2m} - \left(\frac{\sum_{tot}}{2m} \right)^2 - \left(\frac{k_i}{2m} \right)^2 \right] \quad (4)$$

其中, \sum_{in} 是社区 C 内部边权重的总和, \sum_{tot} 表示入射社区 C 内节点 i 的边得权重之和, $k_{i,in}$ 是社区 C 内节点与节点 i 的边权重之和。基于 LCD 算法求出的用户最近邻如公式(5)所示:

$$\text{sim}_{\text{LCD}}(u, v) = \frac{\sum_{l \in L_{u,v}} r_{u,l,c} * r_{v,l,c}}{\sqrt{\sum_{l \in L_{u,v}} (r_{u,l,c})^2} * \sqrt{\sum_{l \in L_{u,v}} (r_{v,l,c})^2}} \quad (5)$$

其中 $l \in L_{u,v}$ 表示 POIs l 已被用户 u 和用户 v 签到过, $r_{u,l,c} = 1$ 表示在社区 C 内的用户 u 已访问过 POIs l , 否则 $r_{u,l,c} = 0$ 。

2.3 相似度融合:

将上述 LDA 模型计算的相似度与 LCD 算法计算的用户相似度线性融合。由于 LBSNs 中新用户访问,评论过的 POIs 很少,且与其有直接好友关系的用户数量相对较少,因此,对新用户进行 POIs 推荐存在严重的冷启动问题。本文中的 LCD 算法将新用户与其好友分到同一个社区 C 内,利用公式(5)计算新用户与社区 C 内其他用户的相似度,

依次对新用户进行 POIs 推荐。解决了 POIs 推荐的冷启动问题。

用户 u 和用户 v 的相似度权重表示为:

$$\text{sim}(u, v) = \alpha \text{sim}_{\text{LDB}}(u, v) + (1 - \alpha) \text{sim}_{\text{LDA}}(u, v) \quad (6)$$

其中 $\alpha (0 < \alpha < 1)$ 为参数。

给定一个用户 u , 用户 u 将在 POIs l 签到的概率表示为:

$$\hat{r}_{u,l} = \frac{\sum_{v \in U} \text{sim}(u, v) * r_{v,l,c}}{\sum_{v \in U} \text{sim}(u, v)} \quad (7)$$

2.4 地理影响

文献[11]中采用用户 u 已访问过 POIs l^u 到其为访问过的 POIs l 的物理距离来计算用户 u 将访问该 POIs l 的概率, 而忽略了若 POIs l^u 与 POIs l 的物理距离很小会导致地理信息对 POIs 推荐的影响过大, 使得用户与社会的影响可以忽略不计。因此本文提出了一种新的计算地理影响的方法, 即用指数函数来计算 POIs 推荐中的地理影响。

对文献[9, 20]的数据集的分析表明: 一方面人们更喜欢探索经常访问的 POIs 附近的 POIs; 另一方面, 人们更喜欢去他们常在地附近的 POIs。文献[21, 22]表明, POIs 越受欢迎, 用户随时间变化签到的频率越高。因此, 假设: (1) 人们倾向于访问他们已访问过 POIs 附近的 POIs, 并随着距离的增加他们愿意访问的兴趣下降。(2) 具有很少签到到历史记录的新用户更喜欢去热门的 POIs。

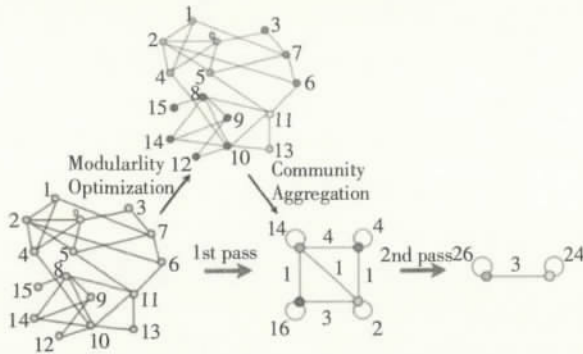


图2 算法过程图解^[17, 19]

用户 u 签到过的 POIs 表示 $L_u = \{l_1, l_2, \dots, l_i\} \in L$, 其中 $l_i = (lat_i, lon_i)^T$ 是纬度 (lat_i) 和经度 (lon_i) 的二维列向量^[22]。用户将在新的 POI 签到的概率为:

$$p(L_u | l) = \frac{1}{\min_{l' \in L_u} \text{dis}(l, l')} \quad (8)$$

其中 $\text{dis}(l, l')$ 表示 POI l 与用户 u 历史签到过的 POI l' 之间的距离。若 $\text{dis}(l, l')$ 不使用指数函数, 那么融合相似度和地理信息的兴趣点推荐 $\text{dis}(l_A, l_B) \approx \text{dis}(l_A, l_C) \approx 0$ 。因此本文采用指数函数来计算

POIs 之间的距离, 计算方法如公式(9):

$$\text{dis}(l, l') = \exp(l - l') (l - l')^T \quad (9)$$

为了将地理影响纳入到 POIs 推荐中, 使用以下方法来产生用户签到他之前未签到过的 POI 的意愿。每个候选 POI 的概率分数为^[9, 11, 19]:

$$p(l | l_u) = p(l) p(L_u | l) \quad (10)$$

其中 $p(l)$ 表示 POI l 的流行度^[24]。

3 SIGFM 模型

本节中, 将上节中三个因素融合成 SIGFM 来预测用户在其未签到过的 POIs l 签到的概率。SIGFM 模型融合了 POIs 推荐中用户的特征向量、社交关系及地理信息三个因素的影响, 提高了 POIs 推荐的精度, 在一定程度上解决了 POIs 推荐系统中的冷启动问题和数据稀疏问题。

SIGFM 模型计算方法如公式(11)所示:

$$p_{u,l} = \hat{r}_{u,l} * p(l | L_u) \quad (11)$$

$p_{u,l}$ 表示用户 u 访问其未签到过的 POI l 的概率。

4 实验结果与分析

4.1 数据集

本文实验数据来自文献[7]中的 Foursquare 数据集, 如表 2 所示。Foursquare 是一个 LBSNs, 用户可以在不同的位置签到。为了移除异常数据, 过滤访问少于 10 人的 POIs 以及少于 10 次签到的用户。本文实验所使用的数据集包括 24941 个用户, 28593 个 POIs 和 1196248 条签到或者评价, 用户-POIs 签到矩阵的稀疏度为 99.900%^[7]。在实验数据集中, 随机选择 20% 作为测试集, 余下的 80% 数据集作为训练集。

表2 使用的数据集说明

Foursquare 数据集:	
用户	24941
POIs	28593
签到或评价	1196248
稀疏度	99.900%

4.2 评价指标

在性能对比实验中, 采用两种 Top-N 指标评估推荐的质量: 准确率 Precision@N 和召回率 Recall@N。衡量指标如式(12), (13)所示:

$$Precision@N = \frac{\sum_u R(u) \cap T(u)}{N} \quad (12)$$

$$Recall@N = \frac{|\sum_u R(u) \cap T(u)|}{|T(u)|} \quad (13)$$

其中 N 是推荐给用户的 POIs 的数量, $T(u)$ 是用户在测试集上的行为列表, $R(u)$ 是根据用户在训练集上的行为给用户做出的推荐列表^[24]。

4.3 实验结果分析

本文对以前工作做了对比试验, 分别对比了 Use-based Collaborative Filtering (UBCF)^[25]、Social-based Collaborative Filtering (SBCF)^[26]、Geo-Measured Friendship-based Collaborative Filtering (GM-FCF)^[27]、Geographical Information Algorithm (GI)^[28] 四种算法的有效性来验证本模型的有效性。对比模型如表 3 所示:

表 3 本文模型和对比实验模型

缩写	全称
SIGFM	
UBCF	融合相似度和地理影响
基于用户的协同过滤	
SBCF	基于社会的协同过滤
GM-FCF	基于地理与用户的协同过滤
GI	地理信息

4.3.1 参数分析

本节中, 将分析参数在数据集上对实验性能的影响, 该参数在平衡社会影响和用户签到信息方面起着重要的作用。图 3 和图 4 给出了 SIGFM 模型在参数的范围在 $[0, 1]$ 范围内变化时对准确率和召回率的影响。从图 3 和图 4 中可以看出, SIGFM 模型的性能随着参数的增加而增加, 当参数 α 等于 0.6 时, 准确率和召回率均达到最高值, 因此本文选取参数 α 的值为 0.6。

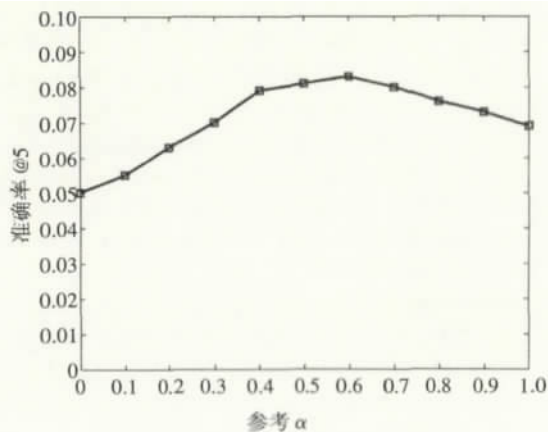


图 3 参数对准确率的影响

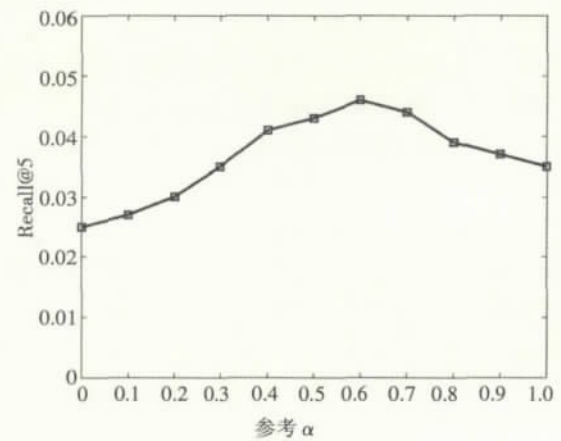


图 4 参数对准确率的影响

图 5 和图 6 给出了不同主题数相应的准确度和召回率曲线。当主题个数少于 30 时, 推荐的准确率和召回率都相对较低。由于项目与主题之间的相关性较强, 在主题个数少于 30 的实验中, 许多项目都被分为同一种类, 从而使得实验结果的准确率和召回率都相对较低。当主题数量增加, 准确率和召回率也在增加, 随后趋于稳定。在具有 60 个主题数时, 准确率和召回率最高。因此, 我们将 60 选作实验部分主题个数的预设参数。

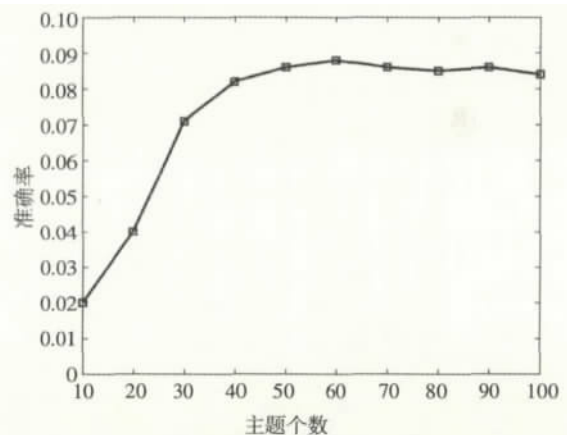


图 5 不同主题个数对应的准确率曲线

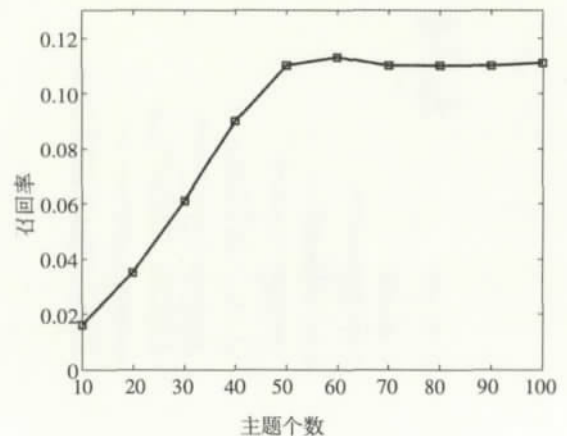


图 6 不同主题个数对应的召回率曲线

4.3.2 推荐模型的比较与分析

本节中,根据4.3.1参数分析对参数进行设置:将参数 α 与主题个数分别设置为0.6和60。图7和图8显示了本文SIGFM模型与其他四种模型在准确率和召回率上的对比结果。如图7和图8所示:随着Top-N增加,准确率减少,召回率增加。给用户推荐的POIs增加时,用户发现自己愿意签到的POIs随着增加,但有些被推荐的POIs被访问的可能性就减少。对比算法中,SBCF和GI模型只考虑了一种影响因素的协同过滤,稀疏的数据集对它们影响相对较大,因此它们在稀疏数据集的POIs推荐上具有很大的不足,准确率和召回率都相对较低;UBCF算法是基于用户的协同过滤,主要考虑了LBSNs用户与其现实生活朋友之间的关系,因此其推荐结果相对较好,但是由于其影响因素单一,在稀疏数据集及冷启动问题上还有很大不足,因此其推荐结果不如SIGFM。SIGFM充分考虑了用户特征向量、用户关系及地理影响,解决了POIs推荐中的数据稀疏与冷启动问题。在不同Top-N的情况下,SIGFM模型与其他方法相比,准确率和召回率都有明显的提高。

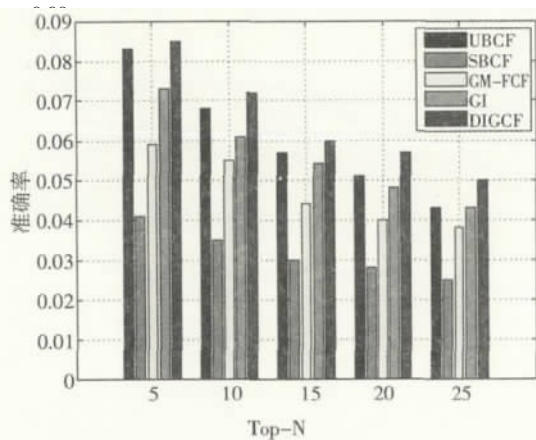


图7 五种算法准确率对比

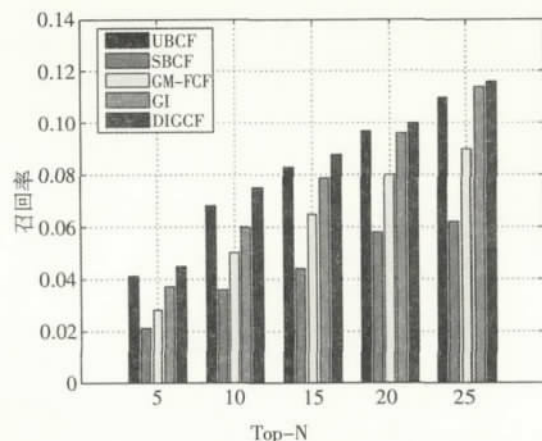


图8 五种算法召回率对比

5 结论与展望

提出的SIGFM模型融合了用户的特征向量、社会关系及地理信息。首先使用LDA模型学习用户的兴趣特征以计算其相似度;然后,使用LCD算法计算同一社区内用户之间的相似度;最后,结合POIs的地理信息对用户进行POIs推荐。在真实的Foursquare的数据集上进行实验,实验结果表明:提出的SIGFM模型比其他的推荐技术在准确率和召回率上有所程度增加,同时一定程度上缓解了POIs推荐系统中的数据稀疏与冷启动问题。

下一步工作,计划结合深度学习中的循环神经网络来挖掘用户与POIs的特征向量,或者融合时间影响,进一步提高兴趣点推荐的质量。

参考文献

- [1] ZHAO S, KING I, LYU M R. A survey of point-of-interest recommendation in location-based social networks [J]. <https://arxiv.org/abs/1607.00647>, 2016.
- [2] BAO J, ZHENG Y, WILKIE D, et al. Recommendations in location-based social networks: a survey [J]. *Geoinformatica*, 2015, 19(3):525—565.
- [3] SALEM Y, HONG J, LIU W. Finder: A cold-start friend finder in large-scale social networks [C]// IEEE International Conference on Big Data. IEEE Computer Society, 2015:687—696.
- [4] WANG B, HUANG J, OU L, et al. A collaborative filtering algorithm fusing user-based, item-based and social networks [C]// IEEE International Conference on Big Data. IEEE, 2015:2337—2343.
- [5] MENG X W, LIU S D, ZHANG Y J, HU X. Research on social recommender systems [J]. *Ruan Jian Xue Bao/Journal of Software*, 2015, 26(6):1356—1372 (in Chinese).
- [6] 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, 38(2):322—336.
- [7] LIU Y, PHAM N, GAO C, et al. An experimental evaluation of point-of-interest recommendation in location-based social networks [J]. *Proceedings of the Vldb Endowment*, 2017, 10(10):1010—1021.
- [8] BARAL R, LI T. MAPS: a multi aspect personalized POI recommender system [C]// ACM Conference on Recommender Systems. ACM, 2016:281—284.
- [9] YE M, YIN P, LEE W C, et al. Exploiting geographical influence for collaborative point-of-interest recommendation [C]// Proceeding of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2011, Beijing, ACM, 2011.
- [10] YIN H, CUI B, SUN Y, et al. LCARS: A spatial item recommender system [J]. *Acm Transactions on Information Systems*, 2014, 32(3):11—11.
- [11] ZHANG D C, LI M, WANG C D. Point of interest recommendation with social and geographical influence [C]// IEEE International Conference on Big Data. IEEE, 2017:1070—1075.

- [12] LIU B, XIONG H. Point-of-Interest recommendation in location based social networks with topic and location awareness [M]// Proceedings of the 2013 SIAM International Conference on Data Mining, 2013.
- [13] 任星怡, 宋美娜, 宋俊德. 基于位置社交网络的上下文感知的兴趣点推荐[J]. 计算机学报, 2017, 40(4): 824—841.
- [14] ZHAO G, QIAN X, XIE X. User-service rating prediction by exploring social users' rating behaviors[J]. IEEE Transactions on Multimedia, 2016, 18(3): 496—506.
- [15] WANG H, TERROVITIS M, MAMOULIS N. Location recommendation in location-based social networks using user check-in data [C]// ACM Sigspatial International Conference on Advances in Geographic Information Systems. ACM, 2013: 374—383.
- [16] 李心茹, 夏阳, 张硕硕. 基于相似度融合和动态预测的兴趣点推荐算法[J]. 计算机工程与应用, 2018(10): 105—109.
- [17] BLONDEL V D, GUILLAUME J L, LAMBIOTTE R, et al. Fast unfolding of communities in large networks [J]. Journal of Statistical Mechanics, 2008, 2008(10): 155—168.
- [18] CLAUSET A, NEWMAN M E, MOORE C. Finding community structure in very large networks[J]. Physical Review E Statistical Nonlinear & Soft Matter Physics, 2004, 70(2): 066111.
- [19] LALWANI D, SOMAYAJULU D V L N, KRISHNA P R. A community driven social recommendation system [C]// IEEE International Conference on Big Data. IEEE, 2015: 821—826.
- [20] YUAN Q, CONG G, MA Z, et al. Time-aware point-of-interest recommendation [C]// International ACM SIGIR Conference on Research and Development in Information Retrieval. ACM, 2013: 363—372.
- [21] YAO L, SHENG Q Z, QIN Y, et al. Context-aware Point-of-interest recommendation using tensor factorization with social regularization[J]. 2015: 1007—1010.
- [22] LI X, JIANG M, HONG H, et al. A time-aware personalized point-of-interest recommendation via high-order tensor factorization [J]. ACM Transactions on Information Systems, 2017, 35(4): 1—23.
- [23] ZHANG J D, LI Y, LI Y. LORE: exploiting sequential influence for location recommendations [C]// ACM Sigspatial International Conference on Advances in Geographic Information Systems. ACM, 2014: 103—112.
- [24] 项亮. 推荐系统实践[M]. 北京: 人民邮电出版社, 2012.
- [25] SARWAR B, KARYPIS G, KONSTAN J, et al. Item-based collaborative filtering recommendation algorithms [C]// International Conference on World Wide Web. ACM, 2001: 285—295.
- [26] YE M, YIN P, LEE W C, et al. Exploiting geographical influence for collaborative point-of-interest recommendation [C]// Proceeding of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2011, Beijing, ACM, 2011.
- [27] YE M, YIN P, LEE W C. Location recommendation for location-based social networks [C]// ACM Sigspatial International Symposium on Advances in Geographic Information Systems, ACM-GIS 2010, DBLP, 2010: 458—461.
- [28] WANG H, TERROVITIS M, MAMOULIS N. Location recommendation in location-based social networks using user check-in data [C]// ACM Sigspatial International Conference on Advances in Geographic Information Systems. ACM, 2013: 374—383.

(上接第66页)

由图 6(a)可知,风机网侧电压实际测量值和额定值两者之间比值维持在 1 左右,表明发电机工作情况稳定,且当风力速度在 5 秒、10 秒时出现约 1 秒的大幅度变化。图(d)表明,双馈电机电磁转矩图形变化是负数,表明发电机处于发电状态,当发电机刚运行以及风力变化较大时,电机转矩产生了明显波动。发电机频率如图 6(f)所示,满足电能质量对于频率的要求,能够实现并网。

由上可知,添加含有 Smith 模型的 PI 策略,可以实现对变流器更为优化的控制,确保双馈风力发电保持优良的稳定特征,进一步提高电力系统电能质量。

3 结 论

当风速出现变化时,对双馈风力发电系统中变流器的控制策略进行分析,提出一个添加有 Smith 模型的 PI 策略的逻辑控制器,避免了非线性被控对象线性化过程中的不准确,提高了系统的鲁棒

性,并在 PSCAD 软件平台上进行模型搭建与仿真,仿真结果表明:变流器在这种运行模式时,能够保证双馈风力发电机正常工作,保证风机输出的电能可以达到系统所要求的电能质量,进行并网。

参考文献

- [1] 张崇巍, 张兴. PWM 整流器及其控制[M]. 北京: 机械工业出版社, 2003.
- [2] 刘中亮. 双馈变速恒频风力发电机 P-Q 解耦控制技术研究[D]. 沈阳: 沈阳工业大学, 2011.
- [3] 贺子倩. 变速恒频双馈风机发电系统改进直流功率控制策略[D]. 杭州: 浙江大学, 2012.
- [4] 李士勇. 模糊控制、神经控制和智能控制论[M]. 哈尔滨: 哈尔滨工业大学出版社, 1998.
- [5] 王建辉, 齐昕, 顾树生. 一类纯滞后系统模糊 Smith 控制策略的研究[J]. 控制与决策, 1998(3): 141—145.
- [6] 文定都, 何玲. 基于 Smith 模糊 PID 控制算法的炉温控制系统[J]. 仪表技术与传感器, 2009(4): 107—117.
- [7] 风力发电机组规范[S]. 北京: 中国船级社, 2008.