

基于概率矩阵分解算法的社交网络用户兴趣点个性化推荐



张敏军 华庆一

西北大学信息科学与技术学院 西安 710127

(zhangminjun322@163.com)

摘要 在社交网络环境中,传统社交网络用户兴趣点的个性化推荐方法存在网络用户兴趣行为的预测精准性低、用户社交数据覆盖率低的问题,不能充分挖掘用户兴趣点的时空序列特征,为此提出了一种基于概率矩阵分解算法的社交网络用户兴趣点个性化推荐方法。在模型训练的伪代码群中,计算与矩阵概率的变异算子相关的数值结果,实现社交关系网络的物理分割,完成基于概率矩阵分解算法的社交网络节点建模。在此基础上,搭建个性化社交网络框架,按照用户兴趣行为的特征挖掘结果,选择个性化的用户来推荐节点,完成社交网络用户兴趣点个性化推荐方法的建立。实用性检测结果表明,与传统方法相比,应用新型个性化推荐方法后,网络用户兴趣行为的预测精准度最高可达100%,用户社交数据覆盖率约为75%,提高了网络用户兴趣行为的预测精准性和用户社交数据覆盖率,社交网络用户兴趣点的时空序列特征得到了充分挖掘。

关键词: 概率矩阵;分结算法;社交网络用户;兴趣点推荐;伪代码群;变异算子;行为特征;时空序列

中图分类号 TP369

Personalized Recommendation of Social Network Users' Interest Points Based on Probability Matrix Decomposition Algorithm

ZHANG Min-jun and HUA Qing-yi

School of Information Science and Technology, Northwest University, Xi'an 710127, China

Abstract In the social network environment, the traditional personalized recommendation method of social network users' interest points has the problems of low prediction accuracy of network users' interest behavior and low coverage of users' social data, which can not fully mine the temporal and spatial sequence characteristics of users' interest points. Therefore, a personalized recommendation method of social network users' interest points based on probability matrix decomposition algorithm is proposed. In the model training pseudo-code group, the numerical results related to the matrix probability mutation operator are calculated to achieve the physical segmentation of the social network, and the node modeling of the social network based on the probability matrix decomposition algorithm is completed. On this basis, the framework of personalized social network is built, and the results are mined according to the characteristics of users' interest behaviors, and the personalized users are selected to recommend nodes, so as to complete the establishment of personalized recommendation method for users' interest points in social network. The practical test results show that, compared with the traditional method, the new personalized recommendation method can predict the interest behavior of network users with the highest accuracy of 100%, and the coverage rate of social data of users is about 75%, which improve the prediction accuracy of interest behavior of network users and the coverage rate of social data of users, and fully excavate the temporal and spatial sequence characteristics of interest points of social network users.

Keywords Probability matrix, Sub-settlement method, Social network users, Interest point recommendation, Pseudocode group, Mutation operator, Behavioral characteristics, Space-time sequence

在社会化媒体技术的影响下,社交网络内在线用户数量飞速增长,该项应用手段不仅严重影响了人们的在线行为及兴趣方向,也在一定程度上为在线网络用户创建了并行化的信息分享平台,从而使得整个社交网络逐渐成为一个传输性载体,进而为各类参与用户提供基础应用、商务交易、网络服

务、网络金融、媒体信息等物理服务^[1-2]。对于在线用户来说,随着社交网络的应用与推广,整个媒体环境会生成大量的节点数据,且随着节点推荐时间的不断增加,社交网络内用户兴趣指数会呈现规律化增长的变化趋势,进而引发严重的“信息化爆炸”问题。

到稿日期:2019-10-12 返修日期:2019-11-20 本文已加入开放科学计划(OSID),请扫描上方二维码获取补充信息。

基金项目:国家自然科学基金资助项目(61272286);高等学校博士学科点专项科研基金资助项目(20126101110006)

This work was supported by the National Natural Science Foundation of China (61272286) and Specialized Research Fund for the Doctoral Program of Higher Education of China (20126101110006).

通信作者:华庆一(hua-qy@163.com)

概率矩阵分解是一种普遍应用的社交网络预测手段,可根据互联网环境中用户行为数据的生成总量,选择性地分配个性化节点在社交网络中的连接形式,进而解决现存的“信息化爆炸”问题。时空序列是时空数据挖掘中的重要任务,其造成了具有复杂性、多样性以及固有的时空属性特征,然而在概率矩阵分解算法的支持下,模型不能按照社交网络用户兴趣点的时空序列特征,为其分配合理的个性化推荐节点。文献[3]提出了基于用户-内容主题模型的兴趣点联合推荐算法,在推荐过程中同时考虑了用户的偏好分布和兴趣点的主题分布,并根据用户新的兴趣点进行推荐。但是,该方法中用户兴趣行为的预测精准度较低。文献[4]提出了一种基于上下文感知的个性化度量嵌入推荐算法,同时考虑了用户签到的上下文情境信息,以丰富有效数据,缓解数据稀疏性问题,进而提高推荐的准确率,进一步优化算法,降低时间复杂度。但是该方法中用户的社交数据覆盖率较低。

为避免上述情况的发生,本文设计了一种基于概率矩阵分解算法的社交网络用户兴趣点个性化推荐方法。该方法在个性化社交网络中挖掘用户兴趣行为的特征数据,再根据推荐节点的分配现状,扩充概率矩阵内的数据限度条件,最后通过检测对比的方式,突出新型个性化推荐方法的实际应用价值。

1 基于概率矩阵分解算法的社交网络节点建模

基于概率矩阵分解算法的社交网络节点建模是新型推荐方法搭建的基础环节,其具体操作处理流程如下。

1.1 模型训练集群建立

模型训练集群是建立概率矩阵分解算法的基础,能够有效协调社交网络环境内的用户节点条件,按照在线兴趣行为的分布形式 $y_j (j=1,2,\dots,d)$,妥善安排概率矩阵 \mathbf{W} 内可分解数据组织所处的具体位置为:

$$\mathbf{W}=[y_1, y_2, \dots, y_d] \quad (1)$$

假设在训练集群内,每个模型节点 S_b 都具备独立的信息编码方式,且无论社交网络内用户行为 λ_j 出现何种变化,与兴趣点相关的模型训练集群 $J(\bar{X}_j)$ 的位置都不会发生改变,则训练集群的分布为:

$$J(\bar{X}_j)=\frac{S_b y_j}{\lambda_j W}, j=1,2,\dots,l \quad (2)$$

其中, j 表示数据信息总量。在标准的模型训练集群中,头节点定义为 α_i ,可作用于社交网络的各级起始位置,当在线用户的兴趣行为出现明显改变时,头节点会离开原始连接位置,建立与中间节点的物理连接。在此情况下,待分解的矩阵信息 n 大量累计,并以集群的形式快速向社交节点传输。中间节点定义为 Q_j ,可作用于社交网络的各级中心位置,当在线用户的兴趣行为出现明显改变时,中间节点出现明显的两级趋近行为,直至实现与头节点及尾节点的完整对接。尾节点定义为 β_j ,可作用于社交网络的各级结束位置,当在线用户的兴趣行为出现明显改变时,尾节点直接获取社交网络内部的中间节点信息,并与其形成独立的概率训练代码实体,则模型训练集群 $\alpha^T Q \beta$ 的表达式为:

$$\alpha^T Q \beta = \sum_{j=1}^n \alpha_j \beta_j \frac{Q_j}{J(X_j)} \quad (3)$$

具体的模型训练集群的建立标准如表 1 所列。

表 1 模型训练集群的建立标准

Table 1 Establishment criteria of model training pseudo code group establishment criteria

伪代码节点名称	头节点	中间节点	尾节点
训练编码方式	.TF node	.IT node	.ED node
编码作用位置	社交网络的各级起始位置	社交网络的各级中心位置	社交网络的各级结束位置
节点集群行为	离开原始连接位置、与中间节点连接	出现两级趋近行为、与头(尾)结点完整对接	获取社交网络中间节点信息、建立概率训练代码实体

1.2 矩阵概率的变异算子计算

矩阵概率的变异算子指从用户偏好向量空间到物品偏好向量空间的映射,是社交网络用户兴趣点个性化推荐系统的基础计算步骤之一,也是既定分解算法中的必要性参数指标,可根据社交网络中用户节点文本的相似性,将个性化兴趣行为聚合分类。在社交网络用户节点文本中,为满足概率矩阵分解算法的应用需求,通常将一个节点文档 \vec{m}_i 看成由若干个兴趣主题 i 组成,而每一个主题则代表一系列用户关联词 \vec{m} ,通过概率化处理的方式就可以建立由用户节点到兴趣行为主题之间的映射关系 S_b ,其表达式为:

$$S_b = \sum_{i=1}^n \alpha^T Q \beta (\vec{m}_i - \vec{m}) (\vec{m}_i - \vec{m})^T \quad (4)$$

掌握社交网络环境内个性化主题的概率分布情况,通过计算节点向量间的相似性水平,就可以知道矩阵概率的变异算子的具体结果^[5-6]。在既定分解算法中,设 $\hat{\mathbf{y}}$ 代表用户的潜在偏好向量, δ 代表物品的潜在属性向量,利用 $\hat{\mathbf{y}}$ 和 δ 可将社交网络分解算法的矩阵概率的变异算子表示为:

$$\mathbf{P} = \frac{\sum_{\epsilon=1}^{\omega} \hat{\mathbf{y}}_{\delta} S_b}{\sum_{\epsilon=1}^{\omega} |Q_{\delta}|} \quad (5)$$

其中, ω 和 ϵ 分别代表社交网络用户节点概率性变换的上、下限积分数值, $\hat{\mathbf{y}}_{\delta}$ 代表物品的潜在属性向量为 δ 时的维度, φ 和 ϵ 分别代表矩阵概率指标的上、下限积分数值, Q_{δ} 代表物品的潜在属性向量 δ 的用户后验概率。

1.3 社交关系网络的分割

社交关系网络的分割是概率矩阵分解算法的实际应用表现形式,可按照在线用户兴趣点的分配方式,选择性地普及数据信息传输所必经的个性化节点。按照相邻层次间的在线用户行为,物理社交网络可以划分为 7 部分:理论层、节点层、行为层、分解层、传输层、实践层、累计层,如图 1 所示。

由图 1 可知,理论层处于整个社交网络的最底层,负责处理与概率矩阵分解算法相关的变异算子数据。节点层中分布了大量的社交用户兴趣节点,可与理论层保持定向连接,进而疏导网络数据信息在传输过程中所面临的物理压力^[7-8]。行为层负责执行节点层的节点连接指令,能够有效分析社交网络用户的兴趣应用行为。分解层可妥善平衡社交网络内的矩

阵概率的变异算子信息,并将其整合成差异化参量,以供后续的兴趣点个性化推荐所应用。传输层连接了整个社交网络结构,实现了由一个用户节点到另一个用户节点的兴趣行为数据传输。实践层与累计层附属在社交网络的同一侧结构,能够按照用户兴趣行为的个性化推荐需求,统计概率变异算子在分解矩阵中的分布形式。

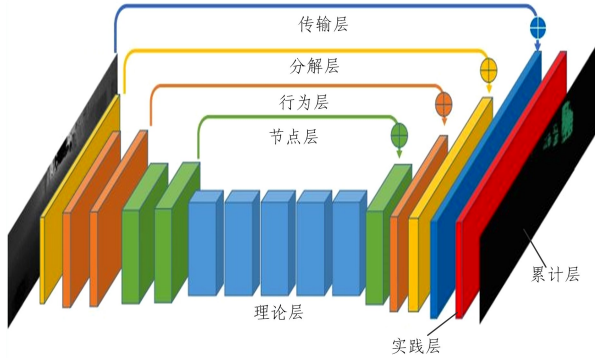


图1 社交关系网络的分割原理

Fig.1 Principles of social relations network segmentation

2 社交网络用户兴趣点个性化推荐方法

2.1 个性化社交网络的框架

个性化社交网络的框架是在线用户兴趣点推荐处理所必备的传输条件,能够根据概率矩阵分解算法中变异算子的依附状态,选择性地传输暂存于用户节点内的社交数据信息。初始情况下,概率矩阵内的变异算子只能按照个性化推荐模型传输,并会随着社交网络内用户节点的迁移行为,逐渐形成独立的个性化推荐链接组织。但在这种情况下,在线用户的兴趣行为不能得到有效满足,且残留于节点内的社交信息很难在既定时间内传输至相关网络结构内部^[9-10]。而在个性化社交网络框架的促进下,概率矩阵内的变异算子首先将信息传输至社交用户兴趣点数据集内;其次以基本推荐变量的形式,建立单独的用户兴趣搜索行为;然后按照社交网络内节点的搜索习惯,将已经完成迁徙行为的节点再次聚合成独立的集合形式;最后在既定分解算法的支持下,形成完整的个性化推荐链接组织结构^[11]。图2反映了完整的个性化社交网络框架的结构。

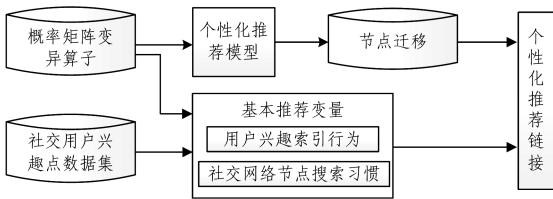


图2 个性化社交网络的框架

Fig.2 Framework of personalized social network

2.2 用户兴趣行为的特征挖掘

用户兴趣行为的特征挖掘指用户通过社交网络系统获取其所需求的信息时,所表现出来的需求表达、信息获取、信息利用等行为,如点击率、网页停留时间、访问量等。计算机通过用户的行为得到用户的偏好特征,再根据这些偏好特征从

大量的信息中对用户感兴趣的内容进行挖掘,最后将挖掘结果推荐给用户^[12-15]。用户兴趣行为的特征挖掘利用个性化社交网络框架的传输特点,安排各类社交节点在个性化推荐过程中的属性分布条件。通常情况下,社交网络内用户节点必须时刻保持良好的签到行为,且随着兴趣点内数据信息的大量累计,待挖掘行为特征开始出现集群性的分布趋势,并以此带动概率矩阵内变异算子的快速连接^[16-17]。在不出明显外力影响的情况下,用户兴趣行为的特征挖掘与概率矩阵内数据信息总量、社交数据签到时长保持着定向的影响关系。概率矩阵内数据信息与用户兴趣行为的特征保持着相同的变化趋势,可在兴趣点行为保持不变的情况下,促进社交网络内的数据分布趋于均衡;社交数据签到时长常表示为 t ,与用户兴趣行为的特征保持相反的变化趋势,可在兴趣点行为保持不变的情况下,确定特征数据的阶段性传输能力,则用户兴趣行为特征的挖掘结果表示为:

$$l = \frac{P \sum_{s=1}^j (\mu + \kappa) j}{t \sqrt{\lambda g^2 + \omega f^2}} \quad (6)$$

其中, s 代表挖掘处理的下限值参量, μ 和 κ 分别代表两个不同的用户兴趣行为应用参量, λ 和 ω 分别代表两个不同的概率矩阵常性分解系数, g 代表概率矩阵中的最大化分解参量条件, f 代表概率矩阵中的最小化分解参量条件。

2.3 个性化推荐节点的选择

个性化推荐节点指社交网络下对用户特定偏好进行推荐的服务器。个性化推荐节点的选择是新型社交网络中用户兴趣点个性化推荐的末端处理环节,可按照概率矩阵分解算法的具体应用需求,为相关用户节点分配数据传输所需的具体参量信息^[18-19]。个性化节点的质量由概率矩阵分解结果决定,从实用性角度来看,社交网络内用户参量的分布频度与定向位移宽度都直接影响着推荐节点的选择结果^[20]。推荐闭锁系统推荐受到阻碍,假设社交网络内用户参量分布频度始终保持为 d ,且随着概率矩阵内变异算子总量的不断增加,相邻推荐节点间的物理间距会随之减小,为避免因节点拥塞而造成的推荐闭锁问题,用户参量分布频度 d 、定向位移宽度 h 间的制约常数始终保持为 $|\tilde{\chi}|$ 。在上述物理变量的支持下,根据式(2)可将个性化推荐节点的选择结果表示为:

$$M = \frac{\sum_{\sigma=1}^{\chi} |h(d - \overset{\dots}{d}) / v'|}{\omega \cdot |\tilde{\chi}|} \quad (8)$$

其中, χ, σ 分别代表社交网络选择推荐积分的上、下限积分数值, $\overset{\dots}{d}$ 代表用户参量分布频度, v' 代表概率矩阵的常性分解系数, ω 代表用户兴趣点的个性化推荐周期常量。

至此,完成所有数值参量的整合处理,实现了基于概率矩阵分解算法的社交网络用户兴趣点个性化推荐方法的设计^[21-22],推荐方法的具体流程如图3所示。

首先建立模型训练的伪代码群,并计算概率矩阵概率的变异算子,根据计算结果对社交关系网络进行物理分割;然后根据用户偏好特征对用户感兴趣的内容进行挖掘,如果兴趣点行为保持不变,则根据挖掘结果选择个性化节点;最后将其推荐给用户。

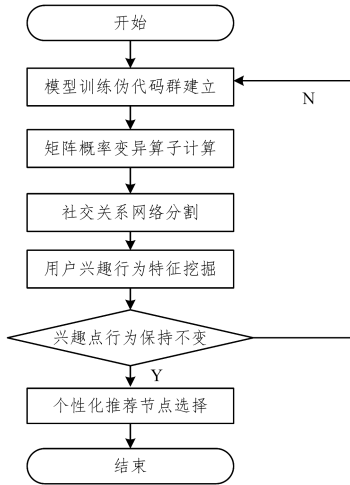


图3 推荐方法的实现流程

Fig. 3 Recommended method of implementation process

3 实践应用检测

3.1 实验环境与数据来源

为验证基于概率矩阵分解算法的社交网络用户兴趣点个性化推荐方法的实效性,在 Matlab7.0, VS2010 + OpenCV 2.4.13, Windows10, Intel (R) Xeon (R) CPU E5-2603v4 @ 2.20 GHz 操作系统, 32 GB 内存下,采用本文方法、文献[3]中的方法和文献[4]中的方法进行仿真实验。在 HWiNFO64 软件的支持下,设置虚拟化社交平台,通过调节用户接入参数的方式,测定相关数值指标的变化情况。

采用 Epinions 数据集作为实验的数据集, Epinions 网站数据集中有 210 739 条数据,包括用户数、商品数、评分数、信任关系数和兴趣标签数。该数据集的统计信息如表 2 所列。

表2 Epinions 数据集的统计信息

Table 2 Statistics of Epinions dataset

数据	数量/条
用户数	52310
商品数	40527
评分数	67526
信任关系数	50203
兴趣标签数	173

3.2 用户签到时间分布

通过社交网络个性化推荐系统,采用本文方法分析用户签到的时间分布规律,结果如图 4 所示。

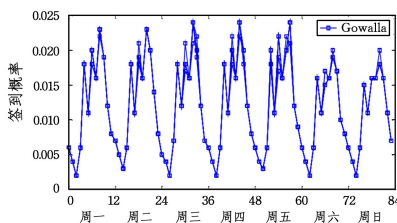


图4 用户一周的签到时间分布

Fig. 4 Weekly check-in time distribution for users

由图 5 的结果可知,周一到周五每天有 3 个峰值,分别为 8 点、12 点和 18 点,这 3 个时间段是人们活动的高峰期。

3.3 用户兴趣行为的预测精准度对比

用户兴趣行为的预测精准度是评估推荐算法最基本的指标,也是最为重要的指标。采用用户兴趣行为的预测精准度指标对本文方法、文献[3]中的方法和文献[4]中的方法进行对比分析,结果如图 5 所示。

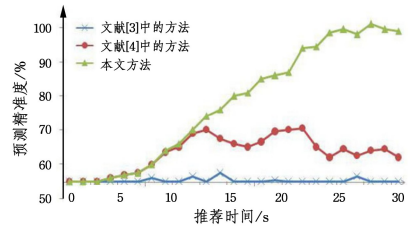


图5 3种方法的预测精准度对比

Fig. 5 Comparison of prediction accuracy of three methods

由图 5 可知,本文方法的用户兴趣行为的预测精准度最高可达 100%,而文献[3]中的方法和文献[4]中的方法的用户兴趣行为的预测精准度最高分别为 58% 和 72%。综上所述,使用基于概率矩阵分解算法的社交网络用户兴趣点个性化推荐方法,可以准确地预测出用户的签到频率和兴趣,针对其兴趣进行个性化推荐,可实现提升用户兴趣行为的预测精确定度的目的。

3.4 覆盖率对比

覆盖率为推荐系统推荐的对象占总对象集合的比例,覆盖率越大,推荐的对象的质量就越高。为了进一步验证本文方法的有效性,使用覆盖率指标对本文方法、文献[3]中的方法和文献[4]中的方法进行对比分析,结果如图 6 所示。

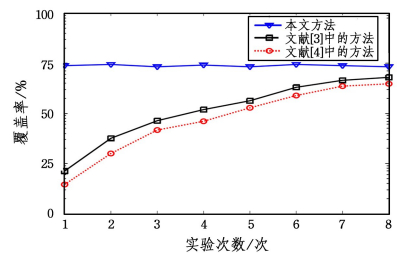


图6 3种方法的覆盖率对比图

Fig. 6 Coverage comparison diagram of three methods

由图 6 可知,随着实验次数的增加,应用新型个性化推荐方法后,用户社交数据的覆盖率呈现平稳趋势,为 75% 左右,而文献[3]中的方法和文献[4]中的方法的覆盖率最高只有 70% 和 65%。本文方法的覆盖率较高,说明其为用户推荐的对象质量较高。

结束语 社交网络用户兴趣点个性化推荐方法在概率矩阵分解算法的支持下,根据模型训练集群的建立标准,计算矩阵内的概率变异算子,进而实现社交关系网络的分割处理,再联合基础框架元件,选择最为合理的个性化推荐节点信息。从实用性的角度来看,随着该推荐方法的应用,传统方法遗留的用户兴趣行为预测精确度过低、用户社交数据覆盖率不达标等问题得到了解决,因此该方法具备较强的实际应用意义。本文方法也可以应用在社交网络推荐系统中,但是对于用户的隐私问题还有研究,在未来的工作中,我们会将社交网络用户隐私数据加密作为研究重点。

参 考 文 献

- [1] ZHU G, CAO X L, SUN Y. Game Strategy Analysis of Privacy Protection input in Social Network Environment-based on the Perspective of Evolutionary Game [J]. *Information Science*, 2017, 25(7): 25-30.
- [2] CHEN Z X, ZENG C, GAO R. A point of interest recommendation model based on location social network fusion of a variety of situational information [J]. *Research on Computer Application*, 2017, 34(10): 2978-2983.
- [3] LU L, ZHU F X, GAO R, et al. Joint recommendation algorithm of interest points based on user-content topic model [J]. *Computer Engineering and its Application*, 2018, 54(4): 154-159.
- [4] XIAN X F, CHEN X J, ZHAO P P, et al. The next point of interest recommendation based on context awareness and personalization metric embedding [J]. *Computer Engineering and Science*, 2018, 280(4): 50-59.
- [5] XIAO Y P, LIU H S, LIU Y B. A social network recommendation scheme based on bipartite graph and node role division [J]. *Electronic Journal*, 2017, 45(10): 2425-2433.
- [6] LIU Y J, HE S, WU Z Q, et al. Recommendation of topic Diversity Reading in University Library based on user Social Network Analysis [J]. *Library and Information Work*, 2018, 62(8): 67-73.
- [7] CHEN J M, LI J G, TANG F Y, et al. Cooperative friend recommendation algorithm based on user-Project-user interest tag Diagram [J]. *Computer Science and Exploration*, 2018, 12(1): 92-100.
- [8] XIAO Y P, SUN H C, DAI T J, et al. A social network recommendation system scoring prediction method based on cloud model [J]. *Journal of Electronics*, 2018, 425(7): 229-234.
- [9] LIU H T, YANG L Q, LING C. A recommendation algorithm for integrating social relations and semantic information in social networks [J]. *Pattern Recognition and Artificial Intelligence*, 2018, 31(3): 236-244.
- [10] WEI Y, LIU G, LI F. Identification and influence evaluation of "invisible" key nodes in online social network information diffusion [J]. *Information science*, 2018, 36(3): 138-143.
- [11] WANG J F, ZHANG P F, GU Z P, et al. An optimized matrix decomposition algorithm with bias probability [J]. *Minicomputer System*, 2017, 38(5): 1081-1085.
- [12] FENG X, ZHAO Z F, ZHAO Y. Cognitive tracking waveform design based on matrix weighted multi-model fusion [J]. *Journal of Harbin University of Technology*, 2018, 50(5): 30-37.
- [13] WANG J F, MIAO Y L, HAN P F. A probability matrix decomposition cooperative filtering recommendation algorithm based on trust mechanism [J]. *Minicomputer System*, 2019, 40(1): 33-37.
- [14] WEN J H, SUN G H, LI S. Research of matrix decomposition recommendation algorithm based on user clustering and mobile context [J]. *Computer Science*, 2018, 45(4): 215-219.
- [15] ZHAO J. Personalized book recommendation algorithm based on improved user interest model [J]. *Machine Tool & Hydraulics*, 2018, 46(6): 193-198.
- [16] MAN T, SHEN H W, HUANG J M, et al. SCMF: A soft constraint matrix decomposition recommendation algorithm for fusion of multi-source data [J]. *Chinese Journal of Information*, 2017, 31(4): 174-183.
- [17] SHAO C C, CHEN P H. Recommendation of interest points that combine social networks and image content [J]. *Computer Application*, 2019, 39(5): 21-28.
- [18] WANG Z Q, LIANG J Y, LI R. Probability matrix decomposition link prediction method based on information fusion [J]. *Computer Research and Development*, 2019, 56(2): 82-94.
- [19] WEN K, ZHU C L. The canonical matrix decomposition recommendation model of the social network and interest is fused [J]. *Computer Application*, 2018(9): 2523-2528.
- [20] DENG M T, LIU X J, LI B. Diversity recommendation method based on user preference and dynamic interest [J]. *Minicomputer System*, 2018(9): 2029-2034.
- [21] DUAN Z T, CAI D D, TANG L, et al. Based on the predicted LBSN interest point recommendation algorithm [J]. *Microelectronics and Computers*, 2019, 36(1): 72-75.
- [22] GAO B, WANG L N, LI L. A social network partitioning method based on clustering ensemble and minimumspanning tree [J]. *Machine Tool & Hydraulics*, 2017, 45(24): 120-125.



ZHANG Min-jun, born in 1979, Ph.D student. His main research interests include intelligent information processing and human-computer interaction engineering.



HUA Qing-yi, born in 1956, Ph.D, professor, Ph.D supervisor, is a member of China Computer Federation. His main research interests include human computer interaction and user interface engineering.