

基于用户潜在特征的社交网络好友推荐方法

肖迎元 张红玉

(天津市智能计算及软件新技术重点实验室 天津 300384)

(天津理工大学计算机科学与工程学院 天津 300384)

摘 要 随着 Facebook、Twitter、微博等社交网站的迅速普及,好友推荐系统逐渐成为各大社交网站的重要组成部分。好友推荐系统通过主动为用户推荐新的潜在好友来有效地扩大用户的社交圈规模并改善用户的社交体验,因而受到了广泛关注。然而,如何针对用户的个性化需求,为用户推荐真正意义上的好友,一直是个性化好友推荐的难点之一。对此,提出一种基于用户潜在特征的社交网络好友推荐方法(SNFRLF)。首先,通过隐语义模型挖掘用户的潜在属性特征;然后,通过用户的潜在特征计算用户间的相似度;最后,将计算得到的相似度引入到随机游走模型中以获得好友推荐列表。实验结果表明,文中所提好友推荐方法较已有的好友推荐方法在性能上有显著提升。

关键词 好友推荐,社交网络,隐语义模型,随机游走

中图法分类号 TP391 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2018.03.034

Friend Recommendation Method Based on Users' Latent Features in Social Networks

XIAO Ying-yuan ZHANG Hong-yu

(Tianjin Key Laboratory of Intelligence Computing and Novel Software Technology, Tianjin 300384, China)

(School of Computer Science and Engineering, Tianjin University of Technology, Tianjin 300384, China)

Abstract With the popularity of social networks, such as Facebook, Twitter and Microblog, friend recommendation systems have gradually become an important part of social networks. Friend recommendation systems effectively expand the scale of user's social circle and improve user's social experience by actively recommending new potential friends for users, thus receiving widespread attention. However, how to personalize the user's needs and recommend real friends to users has been one of the difficulties for personalized friend recommendation. This paper presented a social networking friend recommendation method based on users' latent features, called SNFRLF. SNFRLF first leverages latent factor model to mine users' latent features, and then calculates the similarity between users by means of users' latent features. Finally, the similarity is introduced into the random walk model to get a recommended list. The experimental results show that the proposed method significantly outperforms the existing friend recommendation methods.

Keywords Friend recommendation, Social network, Latent factor model, Random walk

1 引言

随着互联网技术的快速发展以及智能手机的不断普及,社交网络得到了飞速发展。目前,社交网络已成为覆盖互联网用户最广、传播影响力最大的 Web2.0 服务。现在的人们,尤其是新一代年轻人,不仅仅局限于现实生活中的好友社交关系,还通过在线社交网络广交来自世界各地且有共同志趣的新朋友,从而构建起虚拟世界的朋友圈。随着社交网络中用户数量的爆炸式增长,用户要找到真正兴趣相投的朋友变得越来越难,如何有针对性地主动为用户推荐潜在好友,改善用户的社交体验,成为各大社交网站须着力解决的关键问题。

推荐系统是有效解决“信息过载”问题的重要工具,而基于在线社交网络的好友推荐系统则是当前推荐系统研究领域的一个新的热点。

好友推荐系统根据用户目前的好友关系、用户行为记录为用户推荐新的潜在朋友,从而增强用户之间以及用户与社交网站之间的黏性^[1]。传统的协同过滤算法^[2]侧重于记录用户行为来预测和分析用户的喜好,忽略了用户之间的好友关系。本文合理融合用户的兴趣模型和好友关系,首先分析用户的行为记录,利用隐语义模型挖掘用户的潜在兴趣特征以及用户的潜在交友偏好特征;然后通过用户潜在特征计算用户间的相似度;最后将计算得到的相似度引入到随机游走模

收稿日期:2017-07-01 返修日期:2017-08-15 本文受国家自然科学基金重大研究计划(91646117),国家自然科学基金(61170174),天津市自然科学基金(17JCYBJC15200),天津市科技特派员项目(16JCTPJC53600)资助。

肖迎元(1969—),男,教授,博士生导师,主要研究方向为数据库、个性化推荐系统, E-mail: yyxiao@tjut.edu.cn(通信作者);张红玉 女,硕士生,主要研究方向为数据库、个性化推荐系统。

型中,以获取好友推荐列表。与传统的好友推荐方法相比,隐含的用户特征大大降低了用户基本信息的缺失或者伪造虚拟信息等问题对推荐效果的影响。实验结果表明,相比于已有的好友推荐方法,本文所提好友推荐方法在性能上有显著提升。

2 相关工作

协同过滤算法是目前推荐系统中应用得最广泛的推荐算法,主要包括基于内存的方法和基于模型的方法^[3]。其中,基于模型的方法利用概率统计模型或者机器学习方法,从训练集中学习得到用户特征模型(如概率相关模型^[4]、贝叶斯模型^[5]、隐语义模型^[6]等),通过获得的模型预测用户的偏好,并将其推荐给用户,从而解决了用户评分的稀疏性问题。隐语义模型因对用户行为具有较高的预测准确性和良好的扩展性,受到了广泛关注。Koren^[7]通过结合基于邻域的方法与矩阵分解模型提出了SVD++模型,该模型构建了目标用户及其邻域的损失函数,通过多次迭代获取用户与其邻域的潜在偏好,从用户与其邻域两个角度进行混合推荐,显著提高了推荐精度。上述传统的协同过滤方法普遍面临着冷启动问题。

社会化推荐是目前解决冷启动问题的有效方法之一^[8-11]。Ma等^[4]提出了一种将社交网络用户加入到用户-项目评分矩阵的概率矩阵分解方法,其通过加权平均来度量社交网络中受信任用户的偏好对目标用户偏好的影响,实验结果表明,加入社会信息能有效提高推荐的精度;Shen和Jin^[8]认为用户好友中真正有共同兴趣的好友数量较少,而大多数社会化推荐系统中计算了用户全部好友的影响力,这反而降低了推荐的准确性;在此基础上,作者提出了混合专家随机模型,选取活跃度高的好友偏好来度量用户偏好,进而更准确地描述了用户偏好。Chen等^[11]提出了基于排序学习的社交网络好友推荐算法,在基于排序学习的矩阵分解模型中融合了社会关系等网络特征,该算法仅以用户的共同关注对象作为衡量用户之间社交关系影响力的因素,未考虑用户关注行为的偏好特性。

社交网络的好友推荐主要面临两个问题:1)如何度量人的重要性;2)如何度量社交网络中人与人之间的关系。因此,社交关系在好友推荐中有着重要意义。Liben-Howell等人^[12]首次提出对好友推荐进行研究;Jamali等人^[13]提出了基于信任网络的随机游走模型TrustWalker,其将用户间的信任值作为用户社交网络图中的随机游走的转移概率,能有效避免游走过程中噪声数据的影响;Leskovec等人^[14]假设社交网络包括积极关系和消极关系,采用逻辑回归模型来预测积极和消极的在线社交网络的链接,并将社会心理学的平衡理论加入到预测模型中,通过实验证明关联度较高的消极和积极的链接均对用户好友的预测有积极作用;Deng等人^[15]提出了RelevantTrustWalker模型,将用户-项目评分矩阵分解后,利用特征矩阵计算用户间的相似度,并利用用户相似度修正TrustWalker模型,提高了模型的推荐效率。上述方法均从用户的兴趣度角度评价了用户的重要性,但忽略了用户交友行为中用户的潜在交友偏好。

3 基于用户潜在特征的好友推荐(SNFRLF)

3.1 SNFRLF的处理流程

在社交网络中,用户浏览信息、购买商品、关注好友等行为,无一不体现着用户的习惯偏好,通过用户的行为分析挖掘用户的潜在特征可以更好地实现个性化好友的推荐。此外,用户的社交网络结构(用户的社交网络图)也对实现个性化推荐起着重要作用。本文提出的基于用户潜在特征的社交网络好友推荐方法——SNFRLF,主要包括以下处理步骤:1)用户潜在特征的挖掘;2)用户间相似度的计算;3)社交网络图中的随机游走。SNFRLF的具体处理流程如图1所示。

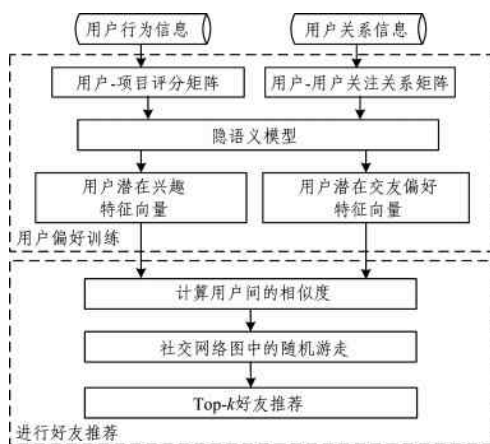


图1 SNFRLF的处理流程

Fig. 1 Processing flow of SNFRLF

其中,社交网络图根据用户在社交网络中的关注关系进行构建,如图2所示,图中的节点代表用户,节点间的有向边代表用户间的关注关系。下文分别对SNFRLF的每一个处理步骤进行详细介绍。

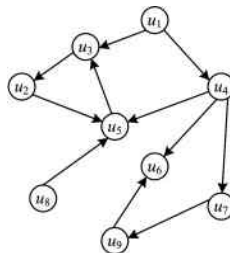


图2 用户间的社交网络图

Fig. 2 Social network diagram between users

3.2 用户潜在特征的挖掘

本文采用隐语义模型^[6]来挖掘用户的潜在特征。该模型能够基于用户的行为对用户与项目进行自动聚类,将用户建模为基于用户潜在特征的向量。传统的基于隐语义模型的推荐方法侧重于用户对项目的评分预测,其目的是将预测评分高的项目推荐给目标用户。本文采用隐语义模型的重点是抽取用户的潜在特征向量,以更精确地计算用户的相似度。本文将用户潜在特征分为用户潜在兴趣特征向量和用户潜在交友偏好特征向量。

3.2.1 用户潜在兴趣特征向量的挖掘

用户对物品的评分行为隐含着用户的兴趣偏好。用户潜

在兴趣特征向量的挖掘基于用户-项目评分矩阵,采用隐语义模型将用户隐含的兴趣映射为一个特征向量,该特征向量的每一个维度代表用户的一个潜在的兴趣属性。

假定 $U = \{u_1, u_2, \dots, u_m\}$ 表示用户集合; $I = \{i_1, i_2, \dots, i_n\}$ 表示项目集合; 矩阵 $R = \{r_{ij}\}_{m \times n}$ 是用户对项目的评分矩阵, r_{ij} 表示用户 u_i 对项目 i_j 的评分。将矩阵 R 分解为表示用户潜在兴趣的特征矩阵 $P \in \mathbb{R}^{m \times a}$ 和表示项目潜在属性的特征矩阵 $Q \in \mathbb{R}^{a \times n}$ 。图 3 给出矩阵 R 被分解成矩阵 P 和矩阵 Q 的一个示例,其中,用户潜在兴趣特征矩阵 P 的每一行即为一个用户的潜在兴趣特征向量。用户潜在兴趣特征向量的抽取被转化为隐语义模型中参数的求解问题,该问题一般通过最优优化损失函数进行求解。本文构建如下损失函数:

$$C = \sum_{(U,I) \in R} (R_{UI} - \sum_{k=1}^F P_{U,k} Q_{k,I})^2 + \lambda \|P_U\|^2 + \lambda \|Q_I\|^2 \quad (1)$$

其中, $\lambda \|P_U\|^2 + \lambda \|Q_I\|^2$ 是用来防止过拟合的正则化项, λ 是正则化参数。基于上述损失函数,本文采用随机梯度下降来进行求解,最终计算得到每个用户的潜在兴趣特征向量。

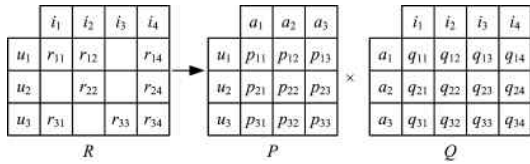


图 3 用户评分矩阵分解的示例

Fig. 3 Example of decomposition of user rating matrix

3.2.2 用户潜在交友偏好特征向量的挖掘

用户的评分行为体现了用户的兴趣偏好特征,而用户的交友行为则体现了用户的交友偏好特征。传统的推荐方法仅考虑了用户的评分行为,本文则综合利用用户的评分行为和交友行为,同时挖掘用户的潜在兴趣特征与交友偏好特征来获得更加准确的用户潜在特征模型,进而更准确地计算用户间的相似度。

假定 $U = \{u_1, u_2, \dots, u_m\}$ 表示用户集合, S 表示用户关注关系的矩阵。将矩阵 S 分解为表示用户潜在交友偏好的矩阵 X 和矩阵 Y 。图 4 给出矩阵 S 被分解成矩阵 X 和矩阵 Y 的一个示例。类似于 3.2.1 节,用户潜在交友偏好特征向量的抽取被转化为隐语义模型中参数的求解问题。

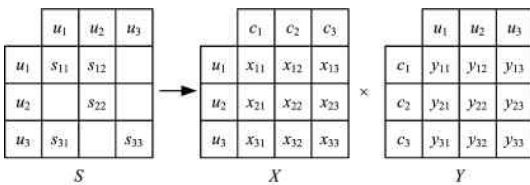


图 4 用户关注关系矩阵分解的示例

Fig. 4 Example of decomposition of users' concern relation matrix

3.3 用户间相似度计算的算

基于 3.2 节获得的用户的潜在特征向量,本节采用带余弦修正的范数来分别计算用户的兴趣相似度和交友偏好相似度。范数和余弦值分别从向量的空间距离和向量方向的一致

性两个不同的角度量化了向量之间的相似度。假定 V_i 和 V_j 表示两个 k 维向量,则 V_i 和 V_j 的范数如式(2)所示:

$$\|V_i - V_j\|_q = (\sum_{l=1}^k |v_{il} - v_{jl}|^q)^{1/q} \quad (2)$$

则特征向量 V_i 和 V_j 对应的两个用户间的相似度可采用范数计算:

$$S_{ij} = \exp^{-\|V_i - V_j\|_q} \quad (3)$$

范数反映了向量的空间距离。进一步,通过向量余弦值(如式(4)所示)引入向量间的方向来修正上述相似度计算公式,得到最终带余弦修正的范数相似度计算公式:

$$\cos(V_i, V_j) = \frac{V_i \cdot V_j}{|V_i| \times |V_j|} \quad (4)$$

$$S_{ij} = \exp^{-\|v_i - v_j\|_q / [0.5 + 0.5 \times \cos(v_i, v_j)]} \quad (5)$$

对于社交网络中的任意两个用户 u_i 和 u_j ,应用上述相似度计算公式分别从用户潜在兴趣特征向量和用户潜在交友偏好特征向量两方面来计算用户的兴趣相似度和交友偏好相似度,如式(6)和式(7)所示:

$$IS_{ij} = S_{ij}(p_i, p_j) = \exp^{-\|p_i - p_j\|_q / [0.5 + 0.5 \times \cos(p_i, p_j)]} \quad (6)$$

$$FS_{ij} = S_{ij}(x_i, y_j) = \exp^{-\|x_i - y_j\|_q / [0.5 + 0.5 \times \cos(x_i, y_j)]} \quad (7)$$

进一步,用户 u_i 和用户 u_j 的综合相似度 $IAFS_{ij}$ 可表示为 IS_{ij} 和 FS_{ij} 的加权平均,如式(8)所示:

$$\begin{aligned} IAFS_{ij} &= \lambda FS_{ij} + (1-\lambda) IS_{ij} \\ &= \lambda \times \exp^{\frac{-\|x_i - y_j\|_q}{[0.5 + 0.5 \times \cos(x_i, y_j)]}} + \\ &\quad (1-\lambda) \exp^{\frac{-\|p_i - p_j\|_q}{[0.5 + 0.5 \times \cos(p_i, p_j)]}} \end{aligned} \quad (8)$$

其中, λ 为权值,用来控制 FS_{ij} 和 IS_{ij} 对用户综合相似度 $IAFS_{ij}$ 的影响。

3.4 基于随机游走的推荐方法

首先基于用户的社交关注关系构建有向加权社交网络图 $G = \langle V, E, W \rangle$, 其中 V 为节点的集合,代表全部用户的集合,假定有 n 个用户 $u_1, u_2, \dots, u_n$, 则 $V = \{u_1, u_2, \dots, u_n\}$; E 代表边的集合,若 u_i 关注 u_j , 则从 u_i 到 u_j 间存在一条边 e_{ij} ; W 为权值的集合,对于任意的 $e_{ij} \in E$, 采用用户综合相似度 $IAFS_{ij}$ 作为 e_{ij} 的权值,即 $W = \{IAFS_{ij} | e_{ij} \in E\}$ 。

本文基于构建的社交网络图 $G = \langle V, E, W \rangle$, 采用随机游走^[19]来计算用户间的最终相似度,并将与目标用户 u_i 最相似的前 k 个用户作为 u_i 的潜在好友推荐给 u_i (Top- k 推荐)。随机游走的主要思想是从图中某个顶点开始的,以某种转移概率决定是选择沿图中与该顶点相邻的边游走到下一个顶点还是直接回到开始的顶点。显然,节点间的转移概率是决定随机游走方法效果的重要因素。本文基于 $G = \langle V, E, W \rangle$ 构建如下转移概率矩阵 $P = \{p_{ij}\}_{n \times n}$, p_{ij} 的计算方式如式(9)所示,其中, m 代表节点 u_i 的出度。

$$p_{ij} = \begin{cases} \frac{IAFS_{ij}}{\sum_{k=1}^m IAFS_{ik}}, & \text{节点 } u_i \text{ 到 } u_j \text{ 间存在一条有向边 } e_{ij} \\ 0, & \text{节点 } u_i \text{ 到 } u_j \text{ 间不存在有向边} \end{cases} \quad (9)$$

具体地,为给用户 u 推荐潜在好友,在图 G 上从目标节

点 u 开始,基于上述转移概率矩阵进行随机游走。游走到任意一个节点时,首先按照一个指定的概率 α 决定是继续游走还是停止这次游走并从节点 u 重新开始游走,若决定继续游走,则从当前节点指向的节点中按照均匀分布随机选择一个节点作为下次游走经过的节点,这样经过多次随机游走后,每个节点(用户)被访问到的概率会收敛到一个确定的值。最终,推荐列表就按照用户被访问概率的大小依次排序,最大的前 k 个用户将被推荐给目标用户 u 。

假定需要给用户 u 推荐潜在好友(即用户 u 为目标用户); w 为图 G 中的任意节点(任意用户); $S_u(w)$ 表示在图 G 中从目标节点 u 开始进行随机游走,节点 w 被访问到的概率,其计算方法如式(10)所示:

$$S_u(w) = (1 - \alpha) \cdot p_{vw} \cdot S_u(v) + \alpha \cdot r_u \quad (10)$$

其中, v 表示上一次随机游走经过的节点; p_{vw} 表示从 v 到 w 的转移概率; $S_u(v)$ 表示 u 访问到 v 的概率; α 为指定的重启概率,表示当前游走有两种选择:以概率 α 停止游走并返回到出发节点 u 或以概率 $1 - \alpha$ 走向下一个用户节点 w ; r_u 取值 0 或 1,当决定停止游走并返回目标节点时, r_u 取值 1,否则取值 0。

对于有向非周期的社交网络图,经过有限次的上述随机游走过程,即重复上述的迭代式(10),目标用户到达图中每个节点的概率值将达到平稳状态,再次游走时图中节点的概率分布将不会改变,此时图中每个用户节点的概率值可被看作是该节点与目标用户节点的相似度。将用户节点的相似度从大到小排序,并将相似度最大的 Top- k 个用户推荐给目标用户。

4 实验评估

4.1 实验数据集

本文使用公开数据集 Epinions^[17] 对提出的 SNFRLF 方法进行实验分析。该数据集主要包含用户对项目的评分记录和用户间的好友关系两个文本文件,具体信息如表 1 所列。

表 1 Epinions 数据集的具体信息

Table 1 Specific information of Epinions data set

文件名称	Rating_data.txt	Trust_data.txt
文件大小/MB	8.24	6.06
包含的属性	user_id	source_user_id
	items_id	target_user_id
	rating_value	trust_statement_value

本实验对该数据集中的 49290 名用户的行为记录进行实验分析,其中包含 49290 名用户的 664823 条项目评分记录和 487182 条用户间的好友关系记录。实验过程中记录用户的所有项目评分。借鉴已有研究的经验,本文也将用户好友关系记录以 4:1 的比例分为训练集和测试集进行实验。

4.2 评价指标

本实验采用准确率(Precision)、覆盖率(Coverage)和 F 指标(F)作为性能评价标准。Precision 用于衡量预测的精度,表示推荐的好友中正确的好友数目所占的比例,其值越大,预测的精度越高;Coverage 反映推荐算法发掘长尾的能力,覆盖率越高,说明越能将长尾中的用户推荐为好友;F 用于评估推荐方法的整体性能,该值越大,算法的准确性和完备

性就越高。假定所有用户集合为 U ,给目标用户 u 推荐 k 个好友,记为 $R(u)$,用户 u 在测试集上的好友集合为 $T(u)$,则各项评测标准的定义如下:

$$Precision = \frac{\sum_u |R(u) \cap T(u)|}{\sum_u |R(u)|}$$

$$Coverage = \frac{|\bigcup_{u \in U} R(u)|}{|U|}$$

$$F = \frac{2 \times Precision \times Coverage}{Precision + Coverage}$$

4.3 实验结果与分析

在 Epinions 数据集上将本文提出的 SNFRLF 与以下几种典型的好友推荐方法进行比较:1)基于社交网络图的好友推荐算法(Pop)^[18],该方法主要依靠用户间的好友关系将朋友的朋友推荐给目标用户;2)基于用户的协同过滤好友推荐算法(User-based CF)^[19];3)基于协同排序的好友推荐算法(MF+LTR)^[20],该算法考虑了用户对两个待推荐好友偏好的排序关系,进而提升了推荐精度;4)基于信任的随机游走好友推荐算法(RelevantTrustWalker)^[15]。

模型参数的设置(包括学习速率、正则化参数、随机游走的重启概率)对上述推荐方法的性能有很大影响。为了客观、公正地进行比较,对每种方法都选取了其相对最优的参数设置,据此来比较上述 5 种方法的性能。通过多次的实验,确定了本文提出的 SNFRLF 方法的参数设置:潜在特征的维数为 150,学习速率为 0.005,正则化参数 λ 为 0.004,随机游走的重启概率 α 为 0.02;其他 4 种方法也采用相应文献中给出的最优参数设置。实验对比了上述 5 种方法在不同的推荐长度($k=10 \sim 100$)下的准确率、覆盖率和 F 值。

本文提出的 SNFRLF 与其他 4 种方法在不同推荐长度下的准确率对比情况如图 5 所示。从图中展示的实验结果可以看出,各推荐方法的准确率随着推荐人数的增加逐渐降低。当推荐人数在区间 10~40 时,准确率下降得较快。推荐人数在 40 以上时(尤其是 40~60 之间),下降趋势明显减弱,可以初步说明,对于该数据集,当推荐人数在 50 左右时,各方法生成的推荐结果的准确性较好。由实验结果还可以明显看出,本文提出的 SNFRLF 在不同的推荐人数下均获得了最高的准确度;其次是 RelevantTrustWalker 方法,该方法综合考虑了用户兴趣和社交关系,因此精确度也比较高;MF+LTR 方法由于未考虑社交关系,精确度明显较低;User-based CF 和 Pop 方法考虑因素单一,受到数据稀疏性的影响较大,精确度大大降低。

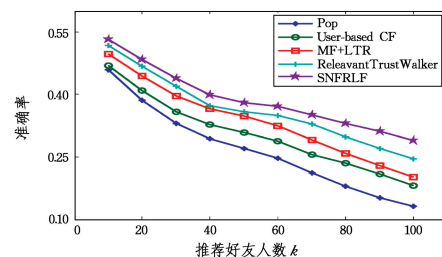


图 5 5 种方法的准确率比较

Fig. 5 Comparison of accuracy rate of five methods

5 种方法在不同推荐人数下的好友推荐覆盖率的实验结果对比情况如图 6 所示。覆盖率反映了推荐方法挖掘长尾的能力。由实验结果可以看出,各推荐方法的覆盖率均随着推荐人数的增加而逐渐增加。当推荐人数 k 在 10~40 之间时,各方法的覆盖率增长得较快;推荐人数大于 40 之后,覆盖率的增长趋势明显减缓。在不同的推荐人数下,SNFRLF 的覆盖率均取得最大值,说明本文方法具有较强的长尾挖掘能力。

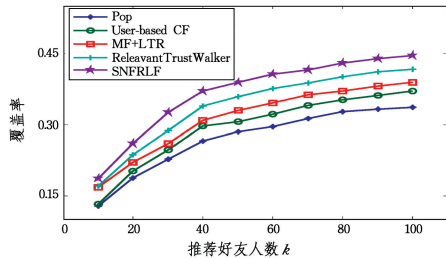


图 6 5 种方法的覆盖率比较

Fig. 6 Comparison of coverage rate of five methods

评价指标 F 在不同推荐人数下的实验结果对比情况如图 7 所示。通过 F 值,能更加全面地评估各推荐方法的综合性能。实验结果表明,该指标的值先随着推荐人数的增加而增大,达到一定值后又随着推荐人数的增加而减小。SNFRLF 的 F 值在不同推荐人数下与其他 4 种方法相比均有较大的提高,且当推荐人数 k 取 50 时, F 值取得最大值,此时准确率和覆盖率达到综合最佳。

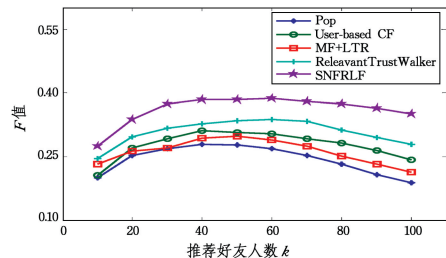


图 7 5 种方法的 F 值比较

Fig. 7 Comparison of F values of five methods

上述实验结果的对比,充分说明了本文所提方法的有效性。综合各实验结果对比图可以看出,当推荐人数 k 在 40~60 之间时,准确率下降的幅度减小,覆盖率值的上升趋势也趋于平缓, F 值取得了最大值,表明推荐人数在该区间时,各算法的推荐效率最高。推荐人数 k 为 50 时,不同算法的综合性能对比如表 2 所列。

表 2 5 种方法的性能对比 ($k=50$)

Table 2 Performance comparison of five methods

算法	准确率	覆盖率	F 值
Pop	0.2701	0.2853	0.2775
User-based CF	0.3086	0.3061	0.3073
MF+LTR	0.3476	0.3296	0.2983
RelevantTrustWalker	0.3586	0.3579	0.3341
SNFRLF	0.3716	0.3886	0.3840

由表 2 数据可以看出,与其他方法相比,本文提出的好友推荐方法在推荐结果的各项评测指标中均有较大的提升,获

得了最好的推荐结果。

结束语 本文提出一种基于用户潜在特征的社交网络好友推荐方法(SNFRLF),该方法将用户的兴趣模型和好友关系进行了合理的融合。首先,对用户的行为记录进行分析,利用隐语义模型挖掘用户的潜在兴趣特征以及用户的潜在交友偏好特征;然后,通过用户潜在特征计算用户间的相似度;最后,将计算得到的相似度引入到随机游走模型中以获得好友推荐列表。与传统的好友推荐方法相比,隐含的用户特征大大降低了用户基本信息缺失或者伪造虚拟信息等问题对推荐效果的影响。实验结果表明,本文提出的好友推荐方法比已有的好友推荐方法在性能上有显著提升。

参 考 文 献

- [1] SHIH S Y, LEE M, CHEN C C. An effective friend recommendation method using learning to rank and social influence[C]// PACIS 2015. 2015.
- [2] CHEN J, GEYER W, DUGAN C, et al. Make New Friends, but Keep the Old: Recommending People on Social Networking Sites [C]// Proceedings of the 27th International Conference on Human Factors in Computing Systems. 2009: 201-210.
- [3] MENG X W, LIU S D, ZHANG Y J, et al. Research on Recommender Systems[J]. Journal of Software, 2015, 26(6): 1356-1372. (in Chinese)
孟祥武, 刘树栋, 张玉洁, 等. 社会化推荐系统研究[J]. 软件学报, 2015, 26(6): 1356-1372.
- [4] MA H, YANG H, LYU M R, et al. SoRec: social recommendation using probabilistic matrix factorization[C]// Proceedings of the 17th ACM Conference on Information and Knowledge Management (CIKM'08). ACM, 2008: 931-940.
- [5] RENDLE S, FREUDENTHALER C, GANTNER Z, et al. BPR: Bayesian personalized ranking from implicit feedback[C]// Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence (UAI'09). ACM, 2009: 452-461.
- [6] HOFMANN T. Latent SEMANTIC Models for Collaborative Filtering[J]. ACM Transactions on Information Systems, 2004, 22(1): 89-115.
- [7] KOREN Y. Factorization meets the neighborhood: a multifaceted collaborative filtering model[C]// Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2008: 426-434.
- [8] SHEN Y, JIN R. Learning personal + social latent factor model for social recommendation[C]// Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'12). ACM, 2012: 1303-1311.
- [9] YE M, LIU X, LEE W C. Exploring social influence for recommendation — a generative model approach[C]// Proceedings of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'12). 2012: 671-680.

中识别中度分化和完全分化的腺管。目前正在尝试首先对 HE 图像进行三分类(由于第三类腺管形态完全消失,因此舍弃这些图片),然后通过细胞质环对第一、二类腺管形态进行腺管识别。

参 考 文 献

- [1] LATSON L, SEBEK B, POWELL K A. Automated cell nuclear segmentation in color images of hematoxylin and eosin-stained breast biopsy[J]. *Analytical and Quantitative Cytology and Histology/the International Academy of Cytology and American Society of Cytology*, 2003, 25(6): 321-331.
- [2] PETUSHI S, GARCIA F U, HABER M M, et al. Large-scale computations on histology images reveal grade-differentiating parameters for breast cancer[J]. *BMC Medical Imaging*, 2006, 6(1): 1-11.
- [3] BAMFORD P, LOVELL B. Unsupervised cell nucleus segmentation with active contours[J]. *Signal Processing*, 1998, 71(2): 203-213.
- [4] XU J, JANOWCZYK A, CHANDRAN S, et al. A weighted mean shift, normalized cuts initialized color gradient based geodesic active contour model: Applications to histopathology image segmentation[C]// *Medical Imaging: Image Processing*, 2010.
- [5] SHI J, MALIK J. Normalized cuts and image segmentation[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000, 22(8): 888-905.
- [6] CHAN T F, VESE L A. Active contours without edges[J]. *IEEE Transactions on image processing*, 2001, 10(2): 266-277.
- [7] BASAVANHALLY A, YU E, XU J, et al. Incorporating domain knowledge for tubule detection in breast histopathology using O'Callaghan neighborhoods[C]// *Proceedings of SPIE—The International Society for Optics and Photonics*, 2011.
- [8] MAQLIN P, THAMBURAJ R, MAMMEN J J, et al. Automatic detection of tubules in breast histopathological images[M]// *Proceedings of Seventh International Conference on Bio-Inspired Computing: Theories and Applications (BIC-TA 2012)*. India: Springer, 2013: 311-321.
- [9] NGUYEN K, BARNES M, SRINIVAS C, et al. Automatic glandular and tubule region segmentation in histological grading of breast cancer[C]// *SPIE Medical Imaging. International Society for Optics and Photonics*, 2015.
- [10] SIRINUKUNWATTANA K, PLUIM J P W, CHEN H, et al. Gland segmentation in colon histology images: The glas challenge contest[J]. *Medical Image Analysis*, 2016, 35: 489-502.
- [11] JANOWCZYK A, MADABHUSHI A. Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases[J]. *Journal of Pathology Informatics*, 2016, 7(1): 29.
- [12] PENG Y, JIANG Y, EISENGART L, et al. Segmentation of prostatic glands in histology images[C]// *2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro. IEEE*, 2011: 2091-2094.
- [13] NGUYEN K, SARKAR A, JAIN A K. Structure and context in prostatic gland segmentation and classification[M]// *Medical Image Computing and Computer-Assisted Intervention. Springer Berlin Heidelberg*, 2012: 115-123.
- (上接第 222 页)
- [10] WAN S, LAN Y, GUO J, et al. Informational friend recommendation in social media[C]// *Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval(SIGIR'13)*. ACM, 2013: 1045-1048.
- [11] CHEN C C, SHIH S Y, MENG L, et al. Who should you follow? Combining learning to rank with social influence for informative friend recommendation[J]. *Decision Support Systems*, 2016, 90: 33-45.
- [12] LIBEN-NOWELL D, KLEINBERG J. The link-prediction problem for social networks[J]. *Journal of the American Society for Information Science and Technology*, 2007, 58(7): 1019-1031.
- [13] JAMALI M, ESTER M. Trustwalker: a random walk model for combining trust-based and item-based recommendation[C]// *ACM Sigkdd International Conference on Knowledge Discovery & Data Mining*, 2009: 397-406.
- [14] LESKOVEC J, HUTTENLOCHER D, KLEINBERG J. Predicting positive and negative links in online social networks[C]// *Proceedings of the 19th International Conference on World Wide Web(WWW'10)*. ACM, 2010: 641-650.
- [15] DENG S G, HUANG L T, XU G H. Social network-based service recommendation with trust enhancement[J]. *Expert Systems with Applications*, 2014, 41(18): 8075-8084.
- [16] SARUKKAI R R. Link prediction and path analysis using Markov chains[J]. *Computer Networks*, 2000, 33(1-6): 377-386.
- [17] MASSA P, AVESANI P. Trust-aware recommender systems[C]// *Proceedings of the 2007 ACM Conference on Recommender Systems*, 2007: 17-24.
- [18] ARMENTANO M G, GODOY D, AMANDI A. Topology-based recommendation of users in micro-blogging communities[J]. *Journal of Computer Science and Technology*, 2012, 27(3): 624-634.
- [19] WANG J, DE VRIES A P, REINDERS M J. Unifying user-based and item-based collaborative filtering approaches by similarity fusion[C]// *Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM, 2006: 501-508.
- [20] CHEN K, CHEN T, ZHENG G, et al. Collaborative personalized tweet recommendation[C]// *Proceedings of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval(SIGIR'12)*. ACM, 2012: 661-670.