

基于用户相似度的随机游走社交网络事件推荐算法

马铁民^{1,2}, 周福才¹, 王 爽¹
(1. 东北大学 计算机科学与工程学院, 辽宁 沈阳 110169; 2. 黑龙江八一农垦大学 电气与信息学院, 黑龙江 大庆 163319)

摘 要: 针对基于社交网络的事件推荐覆盖性和准确性不够高的问题,提出了基于用户相似度 Si-user Walker 算法. 该算法利用基于事件的社交网络特征,将线上用户群组数据抽象为图,以重启随机游走算法为基础,改变了传统的完全基于图的拓扑结构进行随机游走的策略. 根据地理位置划分事件类型,提出了新的用户相似度计算方法,然后根据用户相似度矩阵作为随机游走的转移概率,既保留了图的传递性,又保证了图节点游走的真实性. 与其他推荐算法在真实的数据集上实验表明,该算法在均方根误差、准确率及覆盖率上均得到提升.

关 键 词: 事件推荐;社交网络;用户相似度;拓扑结构;重启随机游走

中图分类号: TP 181 **文献标志码:** A **文章编号:** 1005-3026(2019)11-1533-06

Social Network Event Recommendation Algorithms Based on User Similarity Random Walk

MA Tie-min^{1,2}, ZHOU Fu-cai¹, WANG Shuang¹
(1. School of Computer Science & Engineering, Northeastern University, Shenyang 110169, China; 2. College of Electrical and Information, Heilongjiang Bayi Agricultural University, Daqing 163319, China. Corresponding author: ZHOU Fu-cai, E-mail: fczhou@mail.neu.edu.cn)

Abstract: Aimed at the problem of insufficient coverage and accuracy of event recommendation based on social network, a user similarity-based Si-user Walker algorithm was proposed. The algorithm abstracts online user group data into graphs through event-based social network features and changes the traditional strategy of random walk based on graph-based topology structure on the basis of restart random walk algorithm. According to geographical location, event types were divided and a new calculation method of user similarity was proposed. Then, the user similarity matrix works as transition probability of random walk. Transfer probability not only preserves the transitivity of graph, but also guarantees the authenticity of graph node walk. Compared with other recommendation algorithms, it is shown that the proposed algorithm experiments on real data sets can improve root mean square error, accuracy and coverage.

Key words: event recommendation; social network; user similarity; topological structure; restart random walk

基于事件的社交网络(event-based social network, EBSN)为用户提供事件活动组织的平台,大量事件存在于网络上,网站会定期组织一定的事件活动,用户可以选择自己喜欢的事件进行参加,具有相同爱好的用户可以组建群,为用户参加活动提供了信息共享平台,提高了用户参加事

件的参与度,同时也为用户找到合适的事件和朋友提供了便捷. 根据 EBSN 的特点,一般将该网络分为线上网络(online network)与线下网络(offline network). 随着 EBSN 网站的兴起,由于事件的数量庞大,用户众多,且事件的组织具有时效性、地域性限制,用户需要从众多的数据中选择

符合自身需求的事件及其事件参与者. 因此, 基于 EBSN 环境下的事件推荐算法研究得到很多研究者的关注.

协同过滤是推荐领域应用最为广泛的算法^[1], 主要包括 3 种类型: 基于用户的协同过滤^[2]、基于物品的协同过滤^[3]及基于模型的协同过滤^[4]. 基于用户与物品的协同过滤算法通过计算用户或物品之间的相似度完成对目标用户的推荐, 随着用户与物品的增加, 数据稀疏性和冷启动问题制约该算法的推荐质量. 矩阵分解是应用最广泛的基于模型的协同过滤, 其主旨是将用户评分矩阵分解成两个低维矩阵, 分别代表用户隐式特征与物品隐式特征评分, 通过计算低维矩阵进而得到未评分用户的评分, 但用户的评分具有主观性, 且很多用户没有评分, 导致冷启动用户结果精度不高. 为解决冷启动问题与数据稀疏问题, 很多学者将研究的目光转向其他研究方向, 如充分利用隐式反馈数据, 文献[5]中提到了隐式反馈的基本类型及推荐系统中使用隐式反馈的策略. 随着社交网络的兴起, 社交关系成为重要的隐式反馈数据, 如好友关系、信任关系能够为目标用户的推荐提供辅助. 由于用户的关系具有传递性, 因此很多研究将其转换为图的节点传递问题^[6]. 通过对图的节点、边及权重的分析研究, 达到为用户进行推荐的目的. 很多图的推荐算法研究只是考虑用户的直接好友关系, 并没有充分挖掘出用户的隐式社交关系, 即用户有相同的爱好但没有直接的好友关系, 通过图的随机游走可以对用户的隐式社交关系进行挖掘. 图节点之间的随机游走主要依赖两种方式: 一种是基于图的拓扑性质, 如节点的入度与出度, 该方式虽然利用了图的传递性, 但没有将用户实际发生的特征数据融入到节点传递中; 另外一种基于用户之间的相似程度, 传统的用户相似度计算主要有余弦相似度、改进的余弦相似度及皮尔逊相关系数, 但它们依赖用户对物品的评分, 存在评分本身就带有主观性问题, 不是所有的推荐数据都有用户的评分, 例如基于事件的社会网络并没有类似于电影或书籍的评分.

基于以上问题, 本文提出 Si-user Walker 算法, 改变单纯依靠线上基于网络中信任用户相似度的计算, 根据用户的线下事件活动数据, 提出新的用户相似度计算方法, 该方法基于用户活动次数与类别的计算方式, 计算用户的相似度, 形成相似度矩阵. 同时, 基于线上数据, 即用户群组关系表示为图 $G(N, E)$, N 为用户节点集合, E 为节点

边集合. 不再单独利用图的拓扑性质进行节点的概率转移, 而是将相似度矩阵应用于图中, 采用重启随机游走算法计算源用户相似用户列表, 根据用户的新的群组关系给出最终的推荐结果. 通过对真实数据集进行实验, 对比其他推荐方法, 本文方法取得了较好的推荐效果.

1 社会网络

近年来随着社交网络规模不断扩大和新的模式出现, 越来越多的学者对社会网络图进行探索和研究^[7-13]. 文献[7]系统地对社会网络进行了分析和应用, 随着对社会网络的不断分析和挖掘, 逐渐将其应用到推荐系统中, 通过对社会关系等隐式信息的利用, 提高了推荐的精度. 文献[10]以用户作为社会网络图的节点, 通过对节点已有的标签进行学习, 以此完成对没有标签的节点进行概率计算. 文献[14]利用用户之间的强连接与弱连接之间关系进行建模, 提出了 SoDimRec 推荐框架. 文献[15]利用用户-物品、用户-用户关系, 采用贝叶斯概率矩阵分解的方式求取推荐结果.

基于事件的社交网络有其自身的特点, 其模式如图 1 所示. 一般事件由用户创建, 并指出事件的各项属性, 如事件性质、时间、地点、参与人数等, 如一些体育活动参加的人数可能从十几人到几十人不等. 因此, 相同爱好的人在网络上建立群组关系, 其简化形式如图 2 所示, 节点代表用户,

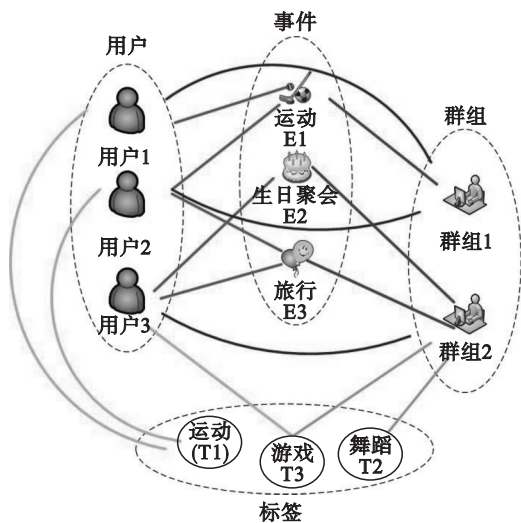


图 1 EBSN 信息结构图
Fig. 1 EBSN structure illustration

边代表用户之间的关系. 节点 $\{1, 2, 3, 4, 5\}$, $\{4, 5, 6, 7\}$, $\{8, 9, 10\}$ 分别构成了不同的群组, 节点 3 与 8 连接了两个群组, 节点 4 与 5 连接了两个群

组. 基于事件的线上群组关系与传统的好友关系有一些区别,从用户参加事件角度考虑,在同一群组中用户可能只是拥有某一方面的爱好,通过对实际的数据集分析,很多参加同一群组的用户的相似程度并不是很高,尤其对于参加事件活动较多的用户.

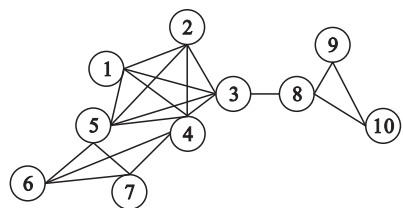


图 2 用户线上群组关系图
Fig. 2 Group diagram on user line

另外,用户参与事件的次数也有很大区别,很多用户参与事件的次数很少,导致 EBSN 网络数据稀疏性非常高. 由于数据的稀疏度较高,考虑将用户参加事件的特征数据应用到群组关系中,通过线上数据与线下数据的结合,既提高推荐的精确度,又在一定程度上解决了冷启动用户的推荐问题.

近十年来许多学者将社会网络转换成图的形式,通过对节点、边及权重的分析,利用图的拓朴性质完成推荐. 其基本思想多为采用随机游走算法,按照一定规则对节点进行随机游走^[6, 16-18]. 随机游走是在齐次图的单变量马尔科夫链上发展起来的,在推荐领域被广泛应用. 充分利用信任网络的作用,先考虑信任网络中用户对物品的评分^[6],然后考虑物品相似性,以此提高推荐的覆盖范围. 在文献[15]中将边的权重值设定为(0, 1),当节点相连接时为 1,否则为 0,将图设定为无向图,即用户之间的影响程度是等价的. 文献[19]中提出的 TidalTrust 在源用户的最短路径内,对所有目标物品进行迭代计算以完成用户对目标物品的评分. 由于其只考虑了最短路径内的其他用户的评分,容易造成评分的欠缺性. Massa 等^[20]提出了 MoleTrust 算法,利用信任传播理论评估信任值的方式,提高了准确性和覆盖率. 用户的群组关系在基于社会维度中将不同的关系混合在一起^[21-22],构成了异构网络,揭示了用户潜在的附属关系,根据社会关系理论,参加相同群组的用户具有相同喜好的概率会更高.

从以上文献可以看到,基于信任的社会网络,如何评估用户之间的信任度是关键问题,而以上文献计算方法虽不同,但都是基于共同好友集合的相似函数进行计算,类似于基于事件的社会网

络中单纯考虑线上的社会群组关系,会产生一定的局限性,没有考虑以下两种情况. 第一,用户共同参加了多个事件,但却没有群组关系;第二,用户通过群组关系成为好友,但参加事件的次数有限,甚至没有. 基于以上两个问题,本文采用将用户线上数据与线下数据相结合,利用线下数据计算用户的相似度,利用线上数据作为社会网络图,为保证用户信任度的传播,采用基于重启的随机游走算法,对目标用户进行随机游走,将用户相似度作为用户节点之间的转移概率. 通过得到与源用户最相似的用户列表,再通过相似用户的事件列表对源用户进行推荐.

2 算法建模

本文算法的核心思想是利用基于事件的社交网络线上数据与线下数据相结合,根据事件的特性进行类别划分,进而进行用户相似度的计算,将用户相似度矩阵融入到线上网络中,最后根据重启随机游走完成源用户的事件推荐.

2.1 事件类别划分

在原始数据集中并没有事件的类别划分特征,但用户参加事件一定有其倾向性,如事件的地点选择、事件类型. 对于事件地点,一般具有一定的指向性,如体育馆一般参加体育活动. 对数据集分析可以发现,很多用户参加的活动较多,但活动的地理位置都为同一地点. 一般而言,这种情况原因有以下两种:第一,该地点举办的活动具有特定性,用户喜欢参加此类活动;第二,用户方便到此地参加活动. 据此,根据事件的地理位置,对事件进行类别划分,定义 $C = \{c_1, c_2, \dots, c_n\}$ 为事件类别划分集合. 其方法描述见算法 1.

算法 1 事件类别划分算法

输入:用户参与事件列表 UserList,事件经纬度列表 EventAddress

输出:用户参与事件类别次数字典 UserDic

1. **Initialize** UserDic
2. 将 EventAddress 中相同地理位置事件进行聚类,得到集合 EventDic
3. 生成以 user 为事件的 UserDic 字典
4. **For** user \in UserDic. keys
5. **If** user 对应 event **IN** EventDic 字典
6. 将事件类别信息写入 UserDic. values,对应次数 + 1
7. **End If**
8. **End For**

算法 1 主要完成统计用户参加事件类别的次

数. 为减少数据冗余, 将事件按地理位置重新编码, 形成 EventDic, 地理位置相同即为一类. 为减少计算量, 将用户参加事件按地理位置划分, 然后在 EventDic 中进行迭代, 最终得到图 3a 形式的矩阵.

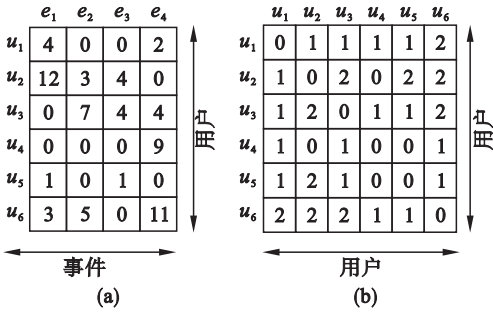


图 3 用户参加事件及类别图示

Fig. 3 User participation events and category diagram

(a)—用户-事件; (b)—用户-用户.

2.2 用户相似度模型建立

无论是余弦相似度还是皮尔逊方法, 都需要用户对物品评分, 而事件数据集不同于影评等数据集, 没有用户的直接评分. 一般对于此种情形, 采用两种方法来计算用户的相似程度.

1) 转换评分. 根据用户购买商品种类的次数进行评分或根据地理位置等因素进行评分, 此种方式可以综合考虑多种因素来完成用户对项目的评分.

2) 基于矩阵的乘积. 图 3a 表示用户参加事件的次数, 图 3b 表示用户之间参加相同事件的数量, 定义矩阵 $M \in \mathbf{R}^{n \times n}$, 矩阵元素 M_{ij} 表示用户 u_i 与 u_j 共同参加事件的次数, 该矩阵为对称矩阵, 其对角线元素为 0. 为了表示用户之间的相似程度, 做运算 $P = M \cdot M^T$, 则新的矩阵 P 的对角线元素不再为 0, 表示用户参与购买物品的程度, 而非对角线元素则表示用户之间的相似程度, 但 P 不是对称矩阵, u_{ij} 与 u_{ji} 的值是不同的, u_{ij} 是以 u_i 为主要参考对象, 而 u_{ji} 是以 u_j 为参考对象.

无论是哪一种方法, 都有一定的不足, 没有充分考虑用户参加事件活动的次数和类别这两个特征属性. 基于此, 本文利用基于用户参加共同类别事件的次数及参加事件的次数这两项指标共同来衡量用户之间的相似程度. 为建立计算用户之间的相似度模型, 提出以下两点假设:

1) 两个用户在单一事件参加的次数越多, 则相似度越高. 如图 3a 所示, 对于事件 e_4 , 用户 $[u_4, u_6]$ 共同参加了事件 9 次, 则用户的相似度较高.

2) 两个用户共同参加的事件类别越多则得分越高. 如 $[u_2, u_3]$ 及 $[u_3, u_6]$ 都分别共同参加了

2 个事件, 则用户的相似度较高.

为满足上述两点假设条件, 定义 $U = \{u_1, u_2, \dots, u_n\}$ 与 $E = \{e_1, e_2, \dots, e_m\}$ 分别为用户集合与事件集合. 定义矩阵 $T \in \mathbf{R}^{n \times n}$, 矩阵元素 t_{ij} 为 $u_i \in U$ 与 $u_j \in U$ 之间相似度评分, 其形式如式 (1) 所示.

$$T = \begin{bmatrix} t_{11} & t_{12} & \cdots & t_{1n} \\ t_{21} & t_{22} & \cdots & t_{2n} \\ \vdots & \vdots & & \vdots \\ t_{n1} & t_{n2} & \cdots & t_{nn} \end{bmatrix}$$

若二者之间没有关系则 t_{ij} 的值为 0, t_{ij} 计算如式 (1) 所示:

$$t_{ij} = c_{ij} k_{ij}. \quad (1)$$

其中: $c_{ij} = \frac{c_i \cap c_j}{c_i \cup c_j}$, c_{ij} 为 u_i 与 u_j 参加类别项目的

类别权重; $k_{ij} = \frac{n_{ij}}{n_i + n_j}$, k_{ij} 为 u_i 与 u_j 参加共同项目次数的权重. 对 T 进行归一化处理, 将 T 中元素除以该列的元素之和, 得到新的矩阵 T' :

$$t'_{ij} = \frac{t_{ij}}{\sum_{i=1}^n t_{ij}}. \quad (2)$$

2.3 用户群组网络的重启随机游走

本文参考基于重启的随机游走算法, 将用户相似度计算结果作为图节点之间的转移概率. 传统的转移概率计算主要以节点的入度与出度进行计算, 只是单纯从图的拓扑结构考虑, 忽略了节点数据的特征属性. 重启随机游走算法是对随机游走算法的一种改进, 该方法类似于 PageRank 算法, PageRank 是佩奇提出的自动网页排序算法, 通过计算网页之间的链接数目和链接网页的重要程度来计算网页的最终排名, 其计算公式:

$$\text{PageRank}(p_i) = \frac{1-q}{N} + q \sum_{p_j} \frac{\text{PageRank}(p_j)}{L(p_j)}. \quad (3)$$

式中: q 称为跳转因子或阻尼因子, 表示到达某一页面后继续向后浏览的概率, 通过实验将其定义为 0.85; $L(p)$ 表示网页 p 的网页集合. 在计算 PageRank 时, 一般采用幂迭代法进行计算. 基于该算法, 又提出了一些改进算法, 如 Personal PageRank、重启随机游走等.

本文采用基于用户相似度的重启随机游走算法, 对图 $G(N, E)$ 进行用户节点随机游走, 其中 N 为节点集, E 为节点之间的边. 传统的随机游走算法利用网络的拓扑结构, 只考虑图的出度和入度完成节点的转移概率, 如节点的出度为 5, 则转到下一节点的概率为 $1/5$, 这与实际的用户节点转移情况并不相符. 在重启随机游走中, 用户节点游

走到下一个节点有两种选择,以一定的概率 α 返回到上一节点,或随机游走到下一节点.其数学表达如公式(4)所示.

$$r_n = cT'r_{n-1} + (1 - c)r_0. \tag{4}$$

其中: r_0 表示用户的列向量;元素 r_j 表示用户 u_j 被访问的概率; r_0 表示各节点初始概率分布,当 $j = i$ 时,取值为 1,否则为 0; r_n 表示第 n 步到达各用户节点的概率; c 为下一步游走到其最近的邻居的概率; T' 为归一化的用户相似度矩阵.根据 PageRank 算法思想,公式(4)最终会收敛到一个平稳的状态,通过递推的方式可以得到公式(5).

$$r_{\text{Last}} = (1 - c)(I - cT)^{-1}r_0. \tag{5}$$

其中, r_{Last} 为假定经过 n 步随机游走后所得到的平稳概率值.因为 $\|cT\|_1 < 1$,故 $(I - cT)^{-1}$ 存在.收敛到稳态后,向量 r 中的每一项值代表着从目标用户 u_i 经过 n 步游走后到达各个节点的概率值,将概率值进行排序,得到 top- k 的用户最相邻用户.具体算法为

算法 2 基于相似度的重启随机游走

输入: 用户列表 UserList、用户参与事件类别字典 UserDic、跳转因子 c

输出: 所有用户游走后的 top- k 节点

1. 根据 UserDic 与公式(1)计算用户相似度矩阵

$$T = (t_{ij})_{n \times n}$$

2. For $u_i \in \text{UserList}$

3. 构建用户列向量 $r_0 = [0, 0, \dots, 1, \dots, 0, 0]^T$

其中 u_i 的值为 1

4. 计算 $(1 - c)(I - cT)^{-1}r_0$

5. End For

2.4 冷启动用户的处理

在推荐系统中,对冷启动用户的推荐一直都是难题,尤其是基于协同过滤算法中,因为没有数据进行参考,所以不能计算用户或物品的相似度.隐式反馈数据是解决冷启动问题的有效手段,如通过好友关系、信任关系以求获得用户的需求.本文利用线下网络数据计算用户相似度,将相似度应用到线上群组关系中.但存在以下两种情况,使得线下用户与基于线上网络的图节点不能相对应.本文将冷启动问题归结为前文提到的两类:

1) 用户参与了线上群组关系,但没有参与任何事件,即冷启动问题.

2) 用户在线下参加事件,但没有参与线上的群组关系.

在本文框架下,采用以下两种方式解决:

1) 用户没有参与事件,但存在于群组中,采用传统随机游走方法,利用图的拓扑性质计算转

移概率.

2) 用户只参加了线下事件,没有线上的群组关系,此用户虽然与其他用户有相似度计算,但在以群组关系为基础的图中不存在该节点,即为悬浮节点,为保证图的连通性,文献[22]假设悬浮节点与所有节点都相连,但该方法只是保证了图的连通性,与实际情况并不相符合.根据相似度计算结果,选取与悬浮节点相似度数值的前 6 个节点作为连接节点,既保证了连通性,又与实际情况相符.

3 实验结果分析

为了验证算法的准确性和实用性,本文以真实的 Meetup 数据集作为实验对象,对数据集进行了数据分析,对用户的事件进行预测,与其他一些常用推荐算法从多个指标进行了对比,得到了较好的推荐效果.

3.1 数据集介绍

Meetup 网站主要由各种各样的群组构成,群组中的成员具有相同的爱好,群组的组织者可以组织、策划、发起各种各样的活动事件并将它们发布到网站上,吸引人们来参与、评论、分享.当一个事件被发布后,任何用户都可以做出回应.实验中采用的数据集包含用户 940 582,发布事件 267 126.图 4 中展示了用户参加事件次数的分布情况.通过对数据集分析可以看到,用户参加事件的次数很不均衡,参加一次事件的用户占的比例是最大的,达到了 38.7%.

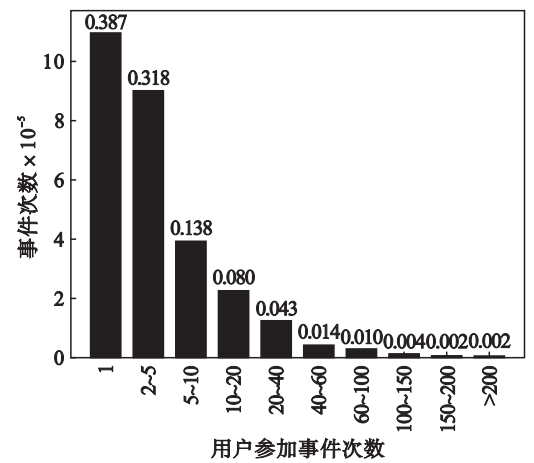


图 4 用户参加事件比例图
Fig. 4 User participation event scale map

3.2 性能对比

实验将数据集分为训练集与测试集,训练集占 80%,测试集占 20%.为分析本文提出算法的

推荐质量,采用均方根误差、精确率、覆盖率作为评价指标,对比了基于用户协同过滤的推荐算法 (User-based CF)、基于用户信任度的推荐算法 (Mole Trust) 及基于信任的随机游走推荐算法 (Relevant Trust Walker). 其结果如表 1 所示.

表 1 算法性能对比
Table 1 Performance comparison of algorithms

指标	User-bas ed CF	Mole Trust	Relevant Trust Walker	Si-user Walker
均方差	1. 389	1. 341	1. 143	1. 121
精确率/%	66. 45	70. 48	71. 53	74. 48
覆盖率/%	71. 32	86. 21	87. 32	89. 17

从表 1 中可以明显观察到,采用基于用户相似度随机游走的事件推荐算法的表现要明显优于其他算法. 本文所提的算法更适合用户基于事件的网络推荐,融合线上与线下数据,通过群组关系的传递弥补了协同过滤算法只依靠事件推荐的弊端,提高了准确率,同时减小了均方根误差. 同时可以看到,基于信任传递的算法由于其传递性,使得覆盖率明显高于协同过滤,而本文的算法通过线下数据计算用户相似度取代了硬性的拓扑传递,在各项指标均优于其他算法.

4 结 语

本文研究基于社交网络中事件的推荐问题,通过对线下活动的特征数据进行计算,将其融合到线上网络的群组分析中,利用基于重启的随机游走算法对用户重新进行社团划分,从而达到为用户推荐事件的目的,尤其是冷启动用户的推荐,充分使用了社交网络中的显式和隐式反馈数据,提高了推荐系统的精确性. 在未来的工作中,将进一步利用社交网络中的上下文数据,如地理位置、标签等数据,提高算法的高效性和可行性. 其次,推荐必将为用户提供更多的服务,如为用户推荐一系列活动,节省用户的规划时间,将成为日后研究的一个方向.

参考文献:

[1] Goldberg D, Nichols D, Oki B M, et al. Using collaborative filtering to weave an information tapestry [J]. *Communications of ACM*, 1992, 35 (12) : 61 – 70.

[2] Konstan J A, Miller B N, Maltz D, et al. GroupLens: applying collaborative filtering to Usenet news [J]. *Communications of ACM*, 1997, 40 (3) : 77 – 87.

[3] Sarwar B, Karypis G, Konstan J, et al. Item-based collaborative filtering recommendation algorithms [C] // International Conference on World Wide Web. Hong Kong:

ACM Press, 2001 : 285 – 295.

[4] Salakhutdinov R, Mnih A. Probabilistic matrix factorization [C] // International Conference on Neural Information Processing Systems. Washington D C: ACM Press, 2007 : 1257 – 1264.

[5] Oard D W, Kim J. Implicit feedback for recommender system [C] // The Fifteenth National Conference on Artificial Intelligence. Madison: AAAI Press, 1998 : 81 – 83.

[6] 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 and Data Mining. Paris: ACM Press, 2009 : 397 – 406.

[7] Wasserman S, Faust K. Social network analysis: methods and applications [J]. *Contemporary Sociology*, 1995, 91 (435) : 219 – 220.

[8] Iglesias J A, García-Cuerva A, Ledezma A, et al. Social network analysis: evolving Twitter mining [C] // IEEE International Conference on Systems, Man, and Cybernetics. Lberta: IEEE Press, 2017 : 1809 – 1814.

[9] Bhattacharyya D, Seth S, Kim T. Social network analysis to detect inherent communities based on constraints [J]. *Social Science Electronic Publishing*, 2016, 8 (1) : 385 – 396.

[10] Tang L, Liu H. Relational learning via latent social dimensions [C] // ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Paris: ACM Press, 2009 : 817 – 826.

[11] Sun Y Z, Han J W. Mining heterogeneous information networks: a structural analysis approach [J]. *ACM SIGKDD Explorations Newsletter*, 2012, 14 (2) : 20 – 28.

[12] O' Malley A J, Onnela J P. Introduction to social network analysis [M]. Dartmouth: Springer-Verlag, 2017 : 1 – 44.

[13] Alhaji R, Rokne J. Encyclopedia of social network analysis and mining [M]. Dartmouth: Springer-Verlag, 2014 : 1 – 50.

[14] Tang J L, Wang S H, Hu X, et al. Recommendation with social dimensions [C] // Thirtieth AAAI Conference on Artificial Intelligence. Phoenix: AAAI Press, 2016 : 251 – 257.

[15] Ma H, Yang H, Lyu M R, et al. SoRec: social recommendation using probabilistic matrix factorization [C] // ACM Conference on Information & Knowledge Management. New York: ACM Press, 2008 : 931 – 940.

[16] Backstrom L, Leskovec J. Supervised random walks: predicting and recommending links in social networks [C] // ACM International Conference on Web Search & Data Mining. Hong Kong: ACM Press, 2011 : 635 – 644.

[17] Gao B, Liu T Y, Wei W, et al. Semi-supervised ranking on very large graph with rich metadata [C] // Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Diego: ACM Press, 2011 : 96 – 104.

[18] Lee S, Song S I, Kahng M. Random walk based entity ranking on graph for multidimensional recommendation [C] // ACM Conference on Recommender Systems. Chicago: ACM Press, 2011 : 93 – 100.

[19] Golbeck J A. Computing and applying trust in web-based social networks [D]. Maryland: University of Maryland at College Park, 2005.

[20] Massa P, Avesani P. Trust-aware recommender systems [C] // ACM Conference on Recommender Systems. Minnesota: ACM Press, 2007 : 17 – 24.

[21] Tang L, Liu H. Scalable learning of collective behavior based on sparse social dimensions [C] // Proceedings of the 18th ACM Conference on Information and Knowledge Management. Hong Kong: ACM Press, 2009 : 1107 – 1116.

[22] Pham T A N, Li X, Cong G, et al. A general graph-based model for recommendation in event-based social networks [C] // IEEE International Conference on Data Engineering. Winston-Salem: IEEE Press, 2015 : 567 – 578.