

基于受限随机游走的时间敏感协同过滤算法*

陈斌

(云南大学 旅游文化学院, 云南 丽江 674100)

摘要:【目的】为了解决传统的协同过滤算法在对用户兴趣建模时出现的推荐准确度下降问题,提出改进算法。【方法】在已有算法的基础上提出一种基于受限随机游走的协同过滤算法,分别构建了基于用户和基于项目的信任网络,通过受限随机游走捕捉并反映用户的兴趣变化及项目流行趋势。【结果】与主流的推荐算法相比,本算法能更有效地捕捉项目流行度和用户兴趣的变化趋势。【结论】本算法避免了盲目游走,降低了推荐误差,提高了推荐准确度。

关键词: 社会网络;协同过滤;随机游走;云相似

中图分类号: TP301.6

文献标志码: A

文章编号: 1672-6693(2018)06-0082-06

互联网的快速发展使得信息以爆炸式的方式传播,电子商务逐渐被越来越多的人所接纳,但同时信息过载(Information overload)问题也日益突出。一方面用户往往被淹没在海量的商品中无所适从,另一方面商家由于无法向用户提供个性化的营销而影响业绩。推荐系统能根据用户的历史数据向用户作出有针对性的推荐,被广泛应用于在线旅游、电子商务、新闻网站等多个领域。协同过滤推荐技术是目前应用最广泛的个性化推荐技术,它是对现实生活中熟人之间推荐的模拟,基本思想是如果两个用户对某些项目的评分相似,则说明他们有相似的兴趣爱好,目标用户对未评分项目的喜好程度,可以通过和他有相似爱好的人的评分进行预测^[1-2]。

1 相关工作

1.1 协同过滤算法介绍

协同过滤推荐问题可描述如下:给定用户集 U 和项目集 I ,用户对项目的评分可表示为一个 $|U| \times |I|$ 的

$$\text{用户—项目评分矩阵 } \mathbf{R} = \begin{pmatrix} U_1 \\ \vdots \\ U_i \\ \vdots \\ U_s \end{pmatrix} \times (I_1 \quad \cdots \quad I_j \quad \cdots \quad I_t) = \begin{pmatrix} R_{11} & \cdots & R_{1j} & \cdots & R_{1t} \\ \vdots & & \vdots & & \vdots \\ R_{i1} & & R_{ij} & & R_{it} \\ \vdots & & \vdots & & \vdots \\ R_{s1} & \cdots & R_{sj} & \cdots & R_{st} \end{pmatrix}。 \text{其中 } s \text{ 是用户数, } t \text{ 是项目}$$

数, R_{ij} 是第*i*个用户对第*j*个项目的评分值。评分是用户对项目的喜好程度的数值表示,评分越高,说明用户对该项目的喜好程度越高。协同过滤推荐要解决的问题就是给定 \mathbf{R} 或 \mathbf{R} 的子集,对 \mathbf{R} 中未评分的项目进行评分预测。协同过滤的关键是寻找用户的“最近邻居”,并根据用户的历史数据结合最近邻居向用户作出推荐。目前协同过滤推荐算法主要包括基于近邻的和基于模型的两类^[3-4]。

基于近邻的协同过滤推荐一般应用在数据量较小的应用场景下,根据历史数据计算用户(项目)之间的相似性,然后选出相似度较高的近邻作为最近邻居集,根据最近邻居集的评分来预测用户对特定产品的喜好程度^[5-6];基于模型的协同过滤推荐一般用于离线计算,将用户评分数据分成训练集和测试集,使用训练集数据来训练出推荐模型,然后使用测试集数据来评估模型的精度,当满足特定精度时,可将得到的推荐模型应用于实际线上环境^[7]。

传统的协同过滤推荐算法虽然能够进行相应的推荐,但是随着数据量的日益增大,面临数据稀疏性和冷启

* 收稿日期:2017-09-02 修回日期:2018-09-27 网络出版时间:2018-10-25 10:42

资助项目:云南省教育厅自然科学基金(No. 2015Y358)

第一作者简介:陈斌,男,彝族,讲师,研究方向为计算机网络、电子商务,Email:chen7552@126.com

网络出版地址: <http://kns.cnki.net/kcms/detail/50.1165.N.20181025.1042.038.html>

动、可扩展性等问题的挑战^[8-10]。特别是只注重用户(项目)间的相似性,而忽略了用户兴趣(项目流行度)的动态变化。为了对用户兴趣和物品流行度变化进行建模,研究者针对传统协同过滤算法提出了许多改进的算法。针对协同过滤推荐中的动态性问题, Ma 等人^[11]提出了一种基于概率矩阵分解框架的社会化推荐方法,将用户的社会关系与评分矩阵进行整合,以体现用户间的社会关系在推荐中的作用。Vermaas 等人^[12]将用户和项目的标签与评分矩阵相结合,利用余弦相似度计算标签的增量,作为用户兴趣变化的衡量指标。Hariri 等人^[13]设计了一个情景变化侦测器,根据用户的持续行为动态地适应情景变化,匹配用户最近的兴趣喜好。郭磊等人^[14]提出了一种结合推荐对象间关联关系的推荐算法,算法使用共享的潜在特征空间对目标函数的求解过程进行约束,提高了预测项目流行度的准确率。Braunhofer 等人^[15]在算法中设计了情景开关,当算法检测到系统添加新用户或新项目时,则会转换到相应的算法进行推荐,以解决冷启动问题,同时该算法可以对老用户最新的兴趣变化进行挖掘。宋伟伟等人^[16]以用户集、时间、标签集及物品资源等4个量形成多维关系,计算出用户和资源之间的标签特征向量,并且依据用户的历史行为来预测用户的偏好,进行推荐。Yehuda^[17]基于因子分解模型构建了时间敏感的因子模型 timeSVD++,该模型使用随机梯度下降算法求解模型中相关参数,构建了用户兴趣变化的曲线模型。Jamali 等人^[18]提出了 TrustWalker 算法,利用社交网络获取用户间的关系并以此作为预测用户兴趣变化的依据,减少由于评分数据稀疏性造成的影响。He 等人^[19]将随机游走与基于项目的协同过滤相结合,根据信任度选择游走方式以缓解冷启动问题。Gavalas 等人^[20]将时间、用户心情等上下文信息用于协同过滤推荐,对推荐结果进行过滤,构建了移动推荐系统模型,该模型可以利用上下文信息预测用户的兴趣变化。Truyen 等人^[21]引入马尔科夫随机场(Markov random fields)来自动评估用户和项目之间的交互关系,并使用稀疏分解(Sparsity-inducing)计算用户-用户和项目-项目之间的相关性。Krishnan 等人^[22]研究了社会影响偏见(Social influence bias)对用户真实评分以及对冲动购买行为的影响,提出设置学习、分析、缓解3个阶段来体现用户最新的兴趣变化,降低社会影响偏见对推荐准确度的影响。何洁月等人^[23]构建了基于实值的状态玻尔兹曼机模型,利用社交网络中的好友信任关系来挖掘用户兴趣的变化趋势,缓解评分数据的稀疏性,提高推荐准确率。郑先荣等人^[24]借鉴心理学遗忘规律,先后提出了线性逐步遗忘推荐算法和非线性逐步遗忘推荐算法,依据评价时间线性逐步减小每项评分的重要性。Zanin 等人^[25]在用户-对象二部图的基础上,把复杂网络的理论应用于基于项目的协同推荐算法,把线性腐蚀时间用于推荐对象的计算,并利用 BA 网络优先扩展的特性来对当前流行的项目进行优先推荐,获得了较好推荐效果。

1.2 已有算法存在的问题及解决方法

以上研究从不同方面对用户兴趣变化进行了建模,但是没有考虑以下3方面的问题:1) 物品的流行度在不断改变,用户的兴趣偏好和评分尺度也在不断改变,但二者变化的时间粒度是不同的。一般来讲,物品的流行度是由大众决定的,变化相对比较缓慢,而用户的兴趣偏好变化更快速,并且存在兴趣反复的现象;2) 现有的社会化推荐算法在构建用户(项目)关系网络时,多采用 Pearson 相关系数、Adamic/Adar 系数等方法计算相似性,但由于用户评分数据的极端稀疏性,严重影响了推荐的准确度;3) 现有的基于网络的推荐算法需要遍历整个网络,导致算法的时间复杂度较高,面对多个拓扑结构相似的结点时,没有根据兴趣相似度区分候选节点的先后次序,对噪声反应过于敏感,影响推荐质量。

本文针对以上问题,主要研究了:1) 分别构建了不同时间粒度的基于用户和基于项目的网络,动态调整推荐结果,以反映用户最新的兴趣变化及项目的流行趋势;2) 采用云相似性在知识层面计算用户(项目)相似度,克服了用户评分数据极端稀疏的情况下推荐质量不高的问题;3) 按照网络结点间云相似度的高低对随机游走过程进行了限制,更注重局部效应,对远离起始节点的网络远端噪声反应相对迟钝。

2 基于受限随机游走的协同过滤算法

2.1 用户兴趣偏好建模

在对时间效应进行建模时,核心点之一就是时间粒度的选择。用户的兴趣偏好可能每天都会有改变,因此对用户偏好的建模需要一个较细的时间粒度,但为了提高预测准确度,又需要每个时间段内包含足够多的评分,因此需要权衡。本文设定用户兴趣偏好时间粒度为 T_u 。根据用户评分矩阵构建用户网络图 $G = \{U, E\}$, $U = \{U_1, U_2, \dots, U_n\}$ 表示 n 个用户集合, $E = \{E_1, E_2, \dots, E_n\}$ 为边的集合, $\forall U_1, U_2$, 若 T_u 时间内 U_1, U_2 有过共同评分,则 $E_{1,2} \in E$ 。边的权重 W 定义为两个用户间的相似性。本文采用云相似性计算用户相似度,云向量表示为

$C_i = (E_{x_i}, E_{n_i}, H_{e_i})$, 其中 $E_{x_i} = \frac{1}{|U_{ij}|} \sum_{u \in U_{ij}} R_{ui}$, $H_{e_i} = \sqrt{\frac{\pi}{2}} \frac{1}{|U_{ij}|} \sum_{u \in U_{ij}} |R_{ui} - E_{x_i}|$, $E_{n_i} = \sqrt{s^2 - \frac{1}{3} H_{e_i}^2}$, $s^2 = \frac{1}{|U_{ij}| - 1} \sum_{u \in U_{ij}} (R_{ui} - E_{x_i})^2$. 用户 i 与 j 的相似度表示为 $Y_{sim}(i, j) = \cos(C_i, C_j) = \frac{C_i \cdot C_j}{\|C_i\| \times \|C_j\|}$.

2.2 随机游走模型

图 G 将用户评分的时序信息和用户间的云相似性相结合, 取代传统的相似度, 能够学习到数据的动态变化, 用户间的云相似性越大, 距离越近, 则关联关系越紧密. 本文使用受限的随机游走来计算用户 U_0 对推荐项目 i 的预测评分. 每次随机游走从目标用户出发, 在用户网络上前进了 k 步后到达某个用户 u . 如果 u 对项目 i 有评分, 则停止此次游走, 返回 R_{ui} 作为此次随机游走的结果. 如果 u 没有评价过 i , 则有 2 个选择:

1) 以概率 $\Phi_{u,i,k}$ 停在用户在用户 u , 选择 u 评价过的与 i 最相似的 m 个项目的评分, 根据 $R_{ui} = \frac{\sum Y_{sim}(i, j) \times R_{uj}}{\sum Y_{sim}(i, j)}$ 计算;

2) 以概率 $1 - \Phi_{u,i,k}$ 继续游走. 将当前用户 u 的 K 个最近邻居按云相似度从大到小排序, 以 $\frac{2m}{n(n+1)}$ 的概率选择第 m 个用户作为下一个结点, 进行下一步游走.

$\Phi_{u,i,k}$ 是游走过程中停止在用户 u 的概率, u 进行过评分的项目与目标项目 i 的相似度越高, 则概率 $\Phi_{u,i,k}$ 越大, 此处分别计算 u 进行过评分的项目与目标项目 i 的相似度, 取最大值. 另外, 游走次数越多, 则用户间偏好的关联性越弱, $\Phi_{u,i,k}$ 应该越大. 为了避免游走过深, 设定游走频数 K 最大值为 6. $\Phi_{u,i,k} = \max_{j \in R_u} \frac{Y_{sim}(i, j)}{1 + e^{-\frac{k}{2}}}$, 其中 R_u 是用户 u 评价过的项目集合.

每一次随机游走都会返回一个预测评分, 为了获得更好的推荐准确性, 可进行多次游走取平均值. 设随机游走需执行 n 次, 定义所有随机游走得到的结果 r_i 的方差为 $\sigma^2 = \frac{1}{n} \sum_{i=1}^n r_i$. 当 $|\sigma_{i+1}^2 - \sigma_i^2| \leq \epsilon$ 时, 游走终止. 一般

取 $\epsilon = 0.0001$, 每次返回 r_i , 则最终结果为 $\hat{r}_{ui} = \frac{1}{n} \sum_{i=1}^n r_i$.

同理可以建立基于项目的网络, 基于项目的网络与基于用户的网络相似, 结点为项目, 项目的流行度时间粒度为 T_i . 与 T_u 相比, T_i 为较粗的时间粒度, 若 T_i 时间内 U_1, U_2 有过共同评分, 则 $E_{1,2} \in E$. 边的权重 W 定义为两个项目间的云相似性.

2.3 两种推荐结果的融合

上面分别介绍了基于用户和基于项目的推荐方法. 但综合几种推荐方法的通常比单一方法更能提高推荐效果. 为了不增加算法时间复杂度, 本文将两种方法进行线性融合, 用 $R_{ij} = \lambda f_u(i, j) + (1 - \lambda) f_i(i, j)$ 进行描述. 其中 $f_u(i, j)$ 表示基于用户的推荐算法计算得出的用户 i 对 j 的预测评分, $f_i(i, j)$ 表示基于项目的推荐算法计算得出的用户 i 对 j 的预测评分, $\lambda \in [0, 1]$ 为融合参数, $\lambda, 1 - \lambda$ 分别代表两种权重值所占的比例, R_{ij} 为融合后的评分. 通过选择合适的 λ 可以将两种方法的优点结合起来, 从而进一步提高推荐质量. 算法模型见图 1.

3 实验结果及分析

3.1 数据集的选取及预处理

本文采用“掘客”网站 Digg 数据集 (http://digg.com) 验证算法有效性. Digg 是一个文章投票评论网站, 它没有职业网站编辑, 编辑全部取决于用户, 范围包括体育、娱乐、政治、科技等. Digg 允许用户进行文章聚合、投票、分

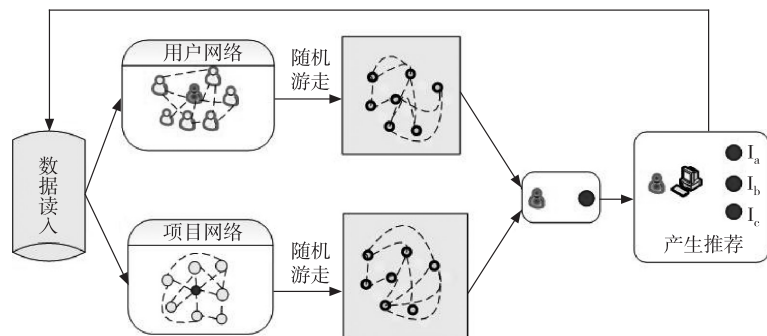


图 1 基于受限随机游走的协同过滤推荐算法框架

Fig. 1 Collaborative filtering recommendation algorithm framework based on constrained random walk

享、交流等。在 Digg 网站上,用户的反馈信息包括 3 种形式。

第 1 种形式为空,表示没有反馈;第 2 种形式是投票(Digg),通过“顶”和“踩”来反馈对新闻的态度;第 3 种是在评论区进行留言回复。与传统的 MovieLens, Jester 等数据集相比,Digg 数据集的社会化网络性更强。

本文将 3 种反馈信息分别对应于五分制中的 1,3,5 分,并从数据集中随机选取 300,500,900 个用户的评分记录作为样本。样本的稀疏程度见表 1。

表 1 Digg 数据集情况
Tab. 1 Overview of Digg data sets

样本	用户数	项目数	评分个数	稀疏程度/%
S1	300	1 260	32 810	91
S2	500	1 530	47 031	93
S3	900	1 490	87 024	94

3.2 评价标准

衡量推荐系统推荐效果的指标有很多,其中最常用的指标是平均绝对误差(Mean absolute error, MAE),本文用 σ_{MAE} 表示。 σ_{MAE} 是用实际评分值和预测评分值的差值的大小来衡量准确性,即 $\sigma_{MAE} = \frac{\sum_{i=1}^N |p_i - q_i|}{N}$,其中 N 为用户已评分的项目数量, p_i 为推荐系统对项目 i 的预测评分, q_i 为用户的真实评分。 σ_{MAE} 值越小,说明推荐质量越高。

3.3 对比实验

实验一,推荐准确度对比。本文选取了线性逐步遗忘(Gradual forgetting collaborative filtering, GFCE)、线性腐蚀时间(Complex networks collaborative filtering, CNCF)2 种算法作为本文算法(Restricted random walk, RW)的对照算法,在 Digg 数据集上,对不同时间粒度下的推荐准确度进行对比。

从图 2、图 3 可以看出,RW 相比于 GFCE 和 CNCF 两种算法,在推荐精度上有明显的提高。在 Digg 数据集上,时间粒度 T 以天为单位,基于用户的 RW 算法在 $T=4$ 左右达到最佳推荐效果,基于项目的 RW 算法在 $T=10$ 左右达到最佳推荐效果。时间粒度过大或过小都会影响推荐精度,这是由于当时间粒度过小时,有效评分不足,推荐有很大的偶然性,而当时间粒度过大时,较久远时间的评分过多,最新评分所占比重降低,无法追踪及体现用户最新的兴趣及物品流行趋势。

实验二,算法运行时间对比。该实验比较了 3 种算法的运行时间,从实验结果可以看出,算法的运行所耗时间为 CNCF 最多,RW 次之,GFCE 时间最少。CNCF 由于需要大量的训练数据,训练过程消耗了过多时间。GFCE 采用线性逐步遗忘,时间复杂度最低。从图 4 可以看到,RW 由于考虑了时间粒度,因此时间复杂度比 GFCE 略高,但总体来看,RW 时间复杂度在可接受的范围内。

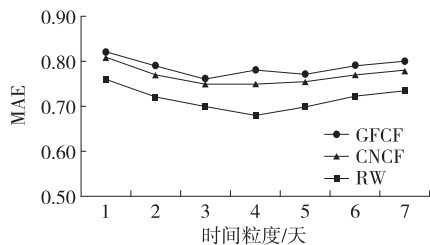


图 2 不同时间粒度下基于用户的推荐准确度

Fig. 2 Recommendation accuracy based on users at different time granularity

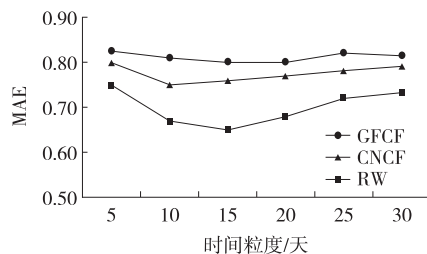


图 3 不同时间粒度下基于项目的推荐准确度

Fig. 3 Recommendation accuracy based on items at different time granularity

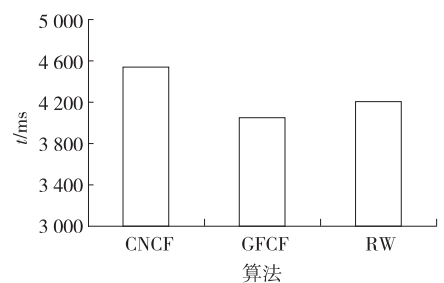


图 4 推荐时间对比

Fig. 4 Comparison of recommendation time

4 结束语

本文将受限的随机游走融入协同过滤推荐中,通过云相似性寻找最近邻居集合,构建了不同时间粒度的用户(项目)网络,从而提高评分预测的准确度。其中基于用户的方法利用评分的时序信息,计算用户间的云相似性,并利用线性衰减对随机游走过程进行限制,从而可以更精确地预测用户偏好,基于项目的方法可以有效缓解物品流行偏置问题。在社会化数据集上的实验表明,该方法与目前的推荐算法相比有更好的推荐效果。算法为

构建时间敏感的用户兴趣(项目流行)模型提供了一个思路,也为如何扩展随机游走在推荐算法中的应用提供了一定的借鉴。

未来的工作包括以下方面:1)对算法中参数的自适应确定方法做进一步的研究;2)面对拓扑结构相似的结点时,随机游走概率根据大小进行了线性衰减,下一步将研究非线性方法。

参考文献:

- [1] GEORGE G, HAAS M R, PENTLAND A. Big data and management[J]. *Academy of Management Journal*, 2014, 57(2):321-326.
- [2] 张玉洁,杜雨露,孟祥武. 组推荐系统及其应用[J]. *计算机学报*, 2016, 39(4):745-764.
ZHANG Y J, DU Y L, MENG X W. Research on group recommender systems and their applications[J]. *Chinese Journal of Computers*, 2016, 39(4):745-764.
- [3] 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.
MENG X W, LIU S D, ZHANG Y J, et al. Research on social recommender systems[J]. *Journal of Software*, 2015, 26(6):1356-1372.
- [4] 王立印,张辉,陈勇. 一种基于 Dice-Euclidean 相似度计算的协同过滤算法[J]. *计算机应用研究*, 2015, 32(10):2891-2895.
WANG L Y, ZHANG H, CHEN Y. Collaborative filtering recommendation algorithm based on dice-euclidean similarity function[J]. *Application Research of Computers*, 2015, 32(10):2891-2895.
- [5] JAVARI A, GHARIBSHAH J, JALILI M. Recommender systems based on collaborative filtering and resource allocation[J]. *Social Network Analysis & Mining*, 2014, 4(1):1-11.
- [6] WANG Z, SUN L F, ZHU W W, et al. Joint social and content recommendation for user-generated videos in online social network[J]. *IEEE Trans on Multimedia*, 2013, 15(3):698-710.
- [7] 孙光福,吴乐,刘淇,等. 基于时序行为的协同过滤推荐算法[J]. *软件学报*, 2013, 24(11):2721-2733.
SUN G F, WU L, LIU Q, et al. Recommendations based on collaborative filtering by exploiting sequential behaviors[J]. *Journal of Software*, 2013, 24(11):2721-2733.
- [8] TRUYEN T, DINH P, SVETHA V. Collaborative filtering via sparse Markov random fields[J]. *Information Sciences*, 2016(369):221-237.
- [9] 王立才,孟祥武,张玉洁. 上下文感知推荐系统[J]. *软件学报*, 2012, 23(1):1-20.
WANG L C, MENG X W, ZHANG Y J. Context-aware recommender systems[J]. *Journal of Software*, 2012, 23(1):1-20.
- [10] 贾冬艳,张付志. 基于双重邻居选取策略的协同过滤推荐算法[J]. *计算机研究与发展*, 2013, 50(5):1076-1084.
JIA D Y, ZHANG F Z. A collaborative filtering recommendation algorithm based on double neighbor choosing strategy[J]. *Journal of Computer Research and Development*, 2013, 50(5):1076-1084.
- [11] MA H, ZHOU D, LIU C, et al. Recommender systems with social regularization[C]//*Proceedings of the 4th ACM International Conference on Web Search and Data Mining*. Hong kong, China. New York: ACM, 2011:287-296.
- [12] VERMAAS R, VANDIC D, FRASINCAR F. Incremental cosine computations for search and exploration of tag spaces[M]. Berlin:Springer, 2012:156-167.
- [13] HARIRI N, MOBASHER B, BURKE R. Context adaptation in interactive recommender systems[C]//*Proceedings of ACM Conference on Recommender Systems*, October 6-10, 2014, Silicon Valley, USA. New York: ACM, 2014:41-48.
- [14] 郭磊,马军,陈竹敏,等. 一种结合推荐对象间关联关系的社会化推荐算法[J]. *计算机学报*, 2014, 37(1):219-227.
GUO L, MA J, CHEN Z M, et al. Incorporating item relations for social recommendation[J]. *Chinese Journal of Computers*, 2014, 37(1):219-227.
- [15] BRAUNHOFER M, CODINA V, RICCI F. Switching hybrid for cold-starting context-aware recommend systems[C]//*Proceedings of ACM Conference on Recommender Systems*, October 6-10, 2014, Silicon Valley, USA. New York: ACM, 2014:349-352.
- [16] 宋伟伟,杨德刚,郑敏. 基于时间加权标签的协同过滤推荐算法研究[J]. *重庆师范大学学报(自然科学版)*, 2016, 33(5):113-120.
SONG W W, YANG D G, ZHENG M. The research in collaborative filtering recommendation algorithm based on time-tags[J]. *Journal of Chongqing Normal University (Natural Science)*, 2016, 33(5):113-120.
- [17] YEHUDA K. Collaborative filtering with temporal dynamics[C]//*Proceedings of ACM Conference on Recommender Systems*, October 22-25, 2009, New York, USA. New York: ACM, 2009:447-456.
- [18] JAMALI M, ESTER M. A matrix factorization technique with trust propagation for recommendation in social networks[C]//*Proceedings of ACM Conference on Recommender Systems*, September 26-30, 2010, Barcelona, Spain. New York: ACM, 2010:135-142.

- [19] HE M, WEI Z. A random walk method using trust factor in collaborative filtering[J]. *The Open Cybernetics & Systemics Journal*, 2014, 8: 843-852.
- [20] GAVALAS D, KNETERIS M. A web pervasive recommendation system for mobile tourist guides[J]. *Personal and Ubiquitous Computing*, 2011, 15(7): 759-770.
- [21] TRUYEN T, DINH P, SVETHA V. Collaborative filtering via sparse Markov random fields[J]. *Information Sciences*, 2016, 369: 221-237.
- [22] KRISHNAN S, PATEL J, FRANKLIN M J, et al. A methodology for learning, analyzing, and mitigating social influence bias in recommender systems[C]//*Proceedings of ACM Conference on Recommender Systems*, October 6-10, 2014, Silicon Valley, USA. New York: ACM, 2014: 137-144.
- [23] 何洁月, 马贝. 利用社交关系的实值条件受限玻尔兹曼机协同过滤推荐算法[J]. *计算机学报*, 2016, 39(1): 183-194.
- HE J Y, MA B. Based on real-valued conditional restricted boltzmann machine and social network for collaborative filtering[J]. *Chinese Journal of Computers*, 2016, 39(1): 183-194.
- [24] 郑先荣, 曹先彬. 线性逐步遗忘协同过滤算法的研究[J]. *计算机工程*, 2007, 33(6): 72-82.
- ZHENG X R, CAO X B. Research on lineal gradual forgetting collaborative filtering algorithm[J]. *Computer Engineering*, 2007, 33(6): 72-82.
- [25] ZANIN M, CANO P, BULDU J M, et al. Complex networks in recommendation systems[C]//*Proc 2nd WSEAS Int Conf on Computer Engineering and Applications*, January 25-27, 2008, Acapulco, Mexico. Wisconsin: World Scientific and Engineering Academy and Society, 2008: 120-124.

A Collaborative Filtering Recommendation Algorithm Based on Restricted Random Walk

CHEN Bin

(College of Tourism and Culture, Yunnan University, Lijiang Yunnan 674100, China)

Abstract: [Purposes] Traditional collaborative filtering disregards the granularity of users' preference drifting and item popularity bias in modeling, thus diminished the accuracy of recommendation. [Methods] A new collaborative filtering algorithm is proposed based on Restricted Random Walk. Two new trust network: user-based and item-based are proposed, with Restricted Random Walk to adaptively track the change of users' preference drifting and item popularity bias. [Findings] Experimental results on social dataset show that the proposed algorithm could capture the popularity of items and users' preference drifting compared with other algorithms. [Conclusions] The proposed algorithm avoids blind walking, reduces the recommendation error and improves the accuracy of recommendation.

Keywords: social networks; collaborative filtering; random walk; cloud similarity

(责任编辑 黄 颖)