

融合用户社交信息和动态兴趣的推荐算法研究

林雨欣, 吴 冰

同济大学经济与管理学院, 上海

收稿日期: 2025年8月3日; 录用日期: 2025年8月23日; 发布日期: 2025年9月11日

摘 要

个性化推荐系统能够有效缓解信息过载问题, 被广泛运用于电子商务、内容资讯推广等应用领域, 但仍存在冷启动和数据稀疏等问题。本文基于传统的协同过滤算法, 通过融合社交关系和兴趣偏好特征建立用户相似性矩阵, 并基于形如指数函数的时间衰减函数对用户兴趣偏好进行动态加权, 最后通过改进的矩阵分解算法得到预测评分。经过实验表明, 本算法在推荐准确率和精度等方面优于协同过滤等传统推荐算法。

关键词

动态推荐, 协同过滤, 矩阵分解

Research on Recommendation Algorithm Integrating Users' Social Information and Dynamic Interests

Yuxin Lin, Bing Wu

School of Economics and Management, Tongji University, Shanghai

Received: Aug. 3rd, 2025; accepted: Aug. 23rd, 2025; published: Sep. 11th, 2025

Abstract

Personalized recommendation systems can effectively alleviate the problem of information overload and are widely used in application fields such as e-commerce and content information promotion. However, there are still problems such as cold start and data sparsity. This paper is based on the traditional collaborative filtering algorithm. By integrating social relationships and interest preference features, a user similarity matrix is established. The user interest preferences are dynamically weighted based on the time attenuation function in the form of an exponential function.

Finally, the predicted scores are obtained through the improved matrix factorization algorithm. Experiments show that this algorithm is superior to traditional recommendation algorithms such as collaborative filtering in terms of recommendation accuracy and precision.

Keywords

Dynamic Recommendation, Collaborative Filtering, Matrix Factorization

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着互联网、物联网和云计算技术的迅猛发展,现代社会步入信息化、数字化时代,人们面临“信息过载”[1]的问题与挑战。个性化推荐是解决信息过载问题、提高信息检索效率的一个有效措施,通过分析用户的行为数据,如点击记录、浏览记录、评分记录等,可以获取用户的意图信息,从而把用户可能感兴趣的内容推荐给用户,节约用户进行信息检索的时间,同时帮助用户发现潜在兴趣爱好,提高用户体验。然而,随着互联网大数据时代的来临,人们对个性化的服务提出了更高的标准和要求,传统的推荐算法面临着诸多问题,如冷启动、数据稀疏、可拓展性等[2]。如何改善数据稀疏造成的不良影响,提高推荐结果准确性,成为个性化推荐在发展过程中面临的新挑战。

在推断用户偏好的过程中,除了评分等显性特征,社交关系、评论情绪、时间特性等隐性特征往往能挖掘出更多用户潜在的兴趣点[3]。社交网络主要是指以真实社会关系或共同兴趣为基础建立起来的,由一系列社会活动的参与者以及他们之间的二元关系集合构成的社会结构[4]。由社交网络理论可知,人们在社交网络上表现出社群特点和兴趣聚合特点,即社交信息不仅能够体现用户之间的相互影响,还能有效体现出用户之间的兴趣相似性和社交相似性。因此,引入社交信息可以有效改善传统推荐算法中仅能通过目标用户历史行为学习用户兴趣偏好的局限性,通过用户间的互动信息,向推荐系统提供更多可参考信息,从而预测用户可能感兴趣的方向。

同时,社交网络上的信息来源日益增多,且传播速度极快。传统的推荐算法以静态推荐为主,将用户的历史行为视为同等重要的影响因素进行推荐,但这与实际情况并不相符。用户的兴趣和社交关系不是一成不变的,随着时间的流逝,用户对某事物的兴趣可能会减弱甚至消失,用户间的关系也可能会发生不同的改变[5]。因此,动态地追踪用户行为的发展过程也是提高个性化推荐算法准确性的有效方法之一,但用户浏览数据等相关信息每时每秒都在更新,如何在高度动态化的推荐场景下结合时间信息构建动态模型,如何精确模拟时间因素对用户兴趣偏移或社交关系改变的影响,仍然是动态推荐领域的研究重点和难点。

2. 文献综述

基于以上事实,已有部分研究在推荐算法中整合用户信任的显式和隐式影响[6],或是考虑用户动态兴趣的影响,在协同过滤算法中引入时间衰减函数表示用户兴趣偏移,调整随时间变化的旧观察值和新观察值的重要性权重[7]。目前,考虑用户社交信息的动态推荐技术大多存在以下问题:

(1) 忽略社交好友与最近邻的不重叠性。例如 Jia 等[8]在计算用户相似度时虽然考虑了基于时间因子的权重来动态模拟用户兴趣随时间的变化,但忽略了用户间的社交关系也会在一定程度上反映用户兴趣

的偏好,在一定程度上降低了推荐算法的预测准确度;

(2) 未考虑用户兴趣的衰减过程。现有研究在用户相似度计算方面,通常未针对用户于不同时间维度下的评分数据进行区别性处理。例如, Sheng Bin [9]将用户的显式社会关系、隐性兴趣相似性与用户自身的评分偏好结合重构用户特征向量,有效缓解了数据稀疏性问题,但未考虑兴趣衰减对推荐效果的负面影响。Bin Suhaim 等[10]提出一种混合方向关系用户相似度模型,结合了作者方向相似度和行为方向相似度来综合计算以匹配用户的上下文,将用户的所有交互行为视为同等重要的因素,忽略了时间因素对于用户兴趣的影响;

(3) 传统矩阵分解算法将特征向量映射到低维空间后导致的相似性差异。把社交信息加入矩阵分解模型后,一般的做法是将评分矩阵与社交矩阵同时分解,共享同一个用户潜在特征空间。例如, Chen 等[11]将用户的信任关系、人际兴趣相似度和项目相似度等社交交互因素融合到概率矩阵分解中,进行多矩阵分解与约束优化; Xu [12]等。但是用户间评分相似度和隐向量相似度之间并没有明显的相关性[13]。基于以上结论,本文改进了同时分解评分矩阵和社交矩阵的方式,通过加入正则项来为矩阵分解的过程施加约束,使用户的潜在特征向量和其社交好友的潜在特征向量相似。从而在矩阵分解时,在相似度更高的社交好友上施加更大的约束,模拟更相似的社交好友对用户的影响也越大的效果。

综上所述,国内外学者不断提出新的推荐算法来分析用户属性并预测用户可能感兴趣的内容,然而在如何从复杂的社交网络中提取用户信任信息、如何精确动态识别用户兴趣变化等方面,仍然是目前个性化推荐领域亟待解决的难点。因此本文在此基础上提出一种基于社交网络的动态推荐算法,融合用户社交信息改进用户相似度的计算,并通过时间衰减函数对用户兴趣进行动态加权,从而提高推荐模型的准确度。

3. 基于社交网络的动态推荐算法构建

本章提出一种基于社交网络的动态个性化推荐算法,算法针对用户社交关系影响与兴趣动态变化等现象,通过构建形如指数函数的时间衰减函数和社交相似度等方式改进传统用户间相似度计算,并在传统矩阵分解模型的基础上,增加社交正则项和最近邻正则项,以提高推荐算法的准确性和稳健性。

3.1. 用户间混合相似度

传统的协同过滤算法利用用户对共同项目的评分数据计算相似度,选择相似度最高的 Top-N 用户组成最近邻集合,再根据最近邻的兴趣为用户进行推荐。本文选择更适合用于稀疏数据的余弦相似性算法来表示用户间评分相似度,如公式(1)所示:

$$S_{uv} = \frac{\sum_{i \in \mathbf{I}_u \cap \mathbf{I}_v} (r_{ui} - \bar{R}_u) \times (r_{vi} - \bar{R}_v)}{\sqrt{\sum_{i \in \mathbf{I}_u \cap \mathbf{I}_v} (r_{ui} - \bar{R}_u)^2} \times \sqrt{\sum_{i \in \mathbf{I}_u \cap \mathbf{I}_v} (r_{vi} - \bar{R}_v)^2}} \quad (1)$$

其中, $\mathbf{I}_u$ 表示用户 u 进行过评分行为的项目集合, $\mathbf{I}_v$ 表示用户 v 进行过评分行为的项目集合, r_{ui} 和 r_{vi} 分别表示用户 u 和用户 v 对项目的评分值, $\bar{R}_u$ 和 $\bar{R}_v$ 分别表示用户 u 和用户 v 的平均评分值。

实际中常利用社交网络中的额外信息,如共同好友、共同兴趣等构建强度关系。对于不存在连接关系的两用户,本文使用 Jaccard 相似性算法来衡量用户间社交信任度,如公式(2)所示:

$$J_{uv} = \frac{|\mathbf{F}_u \cap \mathbf{F}_v|}{|\mathbf{F}_u \cup \mathbf{F}_v|} \quad (2)$$

其中, $\mathbf{F}_u$ 和 $\mathbf{F}_v$ 分别表示用户 u 和用户 v 的好友集合,即用户间社交相似度为两用户好友集合的交集数与

并集数之比。

评分相似度和社交相似度从两个不同角度衡量了用户之间的关联, 它们在一定程度上都可以反映用户的潜在兴趣信息, 因此本文选择将两者结合起来的混合相似度表示用户间关系强度, 定义其为:

$$Sim_{u,v} = \lambda \cdot S_{uv} + (1 - \lambda) \cdot J_{uv} \quad (3)$$

其中, λ 表示用户间评分相似度在关系强度中所占的权重, λ 的值越大, 表示用户间评分相似度对用户间关系强度的影响越大。

3.2. 基于遗忘曲线的时间权重

艾宾浩斯遗忘曲线表明, 人的遗忘速度不是均匀的, 而是一开始遗忘速度很快, 逐渐趋于平缓, 最终停留在某一水平位置, 人的兴趣偏好衰减过程也与此相似[14]。基于此记忆衰减的基本模型, 根据时间的推移使用指数型时间衰减函数模拟用户兴趣衰减过程[15]。

假设有 n 个用户, 可以用 $R = [r_{ij}]_{n \times m}$ 表示用户与物品的评分矩阵。若用户 i 在某个时间点对物品进行了评分, 随着时间的推移, 该评分的影响也应该随之减弱, 如公式(4)所示, 定义评分衰减量为:

$$r_{ij}^T = r_{ij} \cdot e^{-\frac{\ln 2}{\theta}(T_m - T)} \quad (4)$$

其中 r_{ij} 表示用户 i 对物品 j 的真实评分分数, T_m 表示所设置的基准时间, T 表示评分行为发生的时间戳信息, $T_m - T$ 表示评分时间点与基准时间之间的时间间隔, 单位为天, θ 表示用以调节衰减速率的参数, 其值越大, 表示时间因素的影响力越弱, 当 $\theta \rightarrow +\infty$ 时, r_{ij}^T 的值就趋于用户对物品的评分值, 即不再考虑时间因素的影响。

根据上述方法得到衰减评分后, 代入公式(1)得到融入时间因素的用户评分相似度, 进一步得到用户动态关系强度。

3.3. 融合社交信息的矩阵分解模型

在实际情况下, 用户可能只对很少一部分物品进行过评分, 或者与整个社交网络相比, 只和少部分是好友关系, 因此评分矩阵和社交矩阵的信息密度其实并不高, 直接利用原始矩阵进行计算的计算量大且效率很低。矩阵分解的本质是考虑到用户对物品的评分是由用户和物品两个维度的少量因素决定的, 因此将用户评分矩阵映射到两个低维空间, 将评分信息转化为用户对一系列影响因素的偏好程度, 可以有效降低计算的复杂程度。

传统矩阵分解的目标函数为:

$$J_{MF} = \sum_{R_{ij} \neq 0} (R_{ij} - U_i^T V_j)^2 + \alpha_r (\|U_i\|_F^2 + \|V_j\|_F^2) \quad (5)$$

其中, R_{ij} 表示用户 i 对项目 j 的评分, U_i 和 V_j 分别表示用户和物品的潜在特征向量矩阵, $\|U_i\|_F^2$ 和 $\|V_j\|_F^2$ 表示 F2 范数正则项, 用以防止模型过拟合导致算法在测试集表现远不如训练集, α_r 是正则项系数。通过求解这个目标函数得到潜在特征向量矩阵后, 即可相乘得到预测评分矩阵, 从而进行推荐。

考虑融入社交关系的影响, 可以将用户社交关系矩阵与评分矩阵同时分解, 两个矩阵共享用户特征向量矩阵。这意味着默认社交关系矩阵和用户评分矩阵的用户特征表示上是一致的, 但在实际情况中, 用户在不同情境中表现和偏好不一定一致, 如果共享用户特征矩阵, 可能会忽略用户在这些方面的差异, 影响模型的性能。因此本文考虑分解用户评分矩阵得到用户特征矩阵和物品特征矩阵, 同时在目标函数中加入社交正则项以表示社交关系的影响, 从而灵活调整社交信息对用户特征向量的影响程度, 更精确

地捕捉用户在不同情境中的偏好差异, 在降低计算复杂度的同时, 避免因为共享用户特征矩阵造成的信息缺失和扭曲。

本文提出的基于社交网络的动态推荐算法 TimeMF-SN, 在上述传统矩阵分解算法的基础上, 综合考虑了社交关系以及用户兴趣动态偏移的影响, 即用户对某项目的评分由用户自身偏好和社交好友影响共同决定。定义其目标函数为:

$$J = \sum_{u=1}^m \sum_{j=1}^n (r_{ij} - U_u^T V_j)^2 + \alpha \sum_{u=1}^m \sum_{v \in F_u} Sim_{uv} \|U_u - U_v\|_F^2 + \beta \sum_{u=1}^m \sum_{k \in N_u} S_{uk} \|U_u - U_k\|_F^2 + \lambda_1 \|U_u\|_F^2 + \lambda_2 \|V_j\|_F^2 \quad (6)$$

其中, $\sum_{u=1}^m \sum_{j=1}^n (r_{ij} - U_u^T V_j)^2$ 表示用户自身偏好对评分的影响, r_{ij} 表示融入时间权重后用户对物品的评分,

$\sum_{u=1}^m \sum_{v \in F_u} Sim_{uv} \|U_u - U_v\|_F^2$ 表示社交正则项, $\sum_{u=1}^m \sum_{k \in N_u} S_{uk} \|U_u - U_k\|_F^2$ 表示最近邻正则项, α 和 β 分别表示社交正则项和最近邻正则项的正则项系数, Sim_{uv} 表示用户 u 和用户 v 加入时间权重后的混合相似度, S_{uk} 表示用户 u 和用户 k 之间的评分相似度, F_u 表示用户的社交好友集合, N_u 表示用户的最近邻集合。 $U_{k \times m} = [u_1, u_2, \dots, u_m]$ 表示用户潜在特征向量矩阵, $V_{k \times n} = [v_1, v_2, \dots, v_n]$ 表示物品潜在特征向量矩阵。

算法通过随机梯度下降方法迭代目标函数进行求解, 其中控制迭代更新步长的学习率 γ 由交叉验证实验得到。在迭代求解得到用户潜在特征向量和项目潜在特征向量矩阵后, 内积得到用户对项目的预测评分。

4. 实验验证与分析

本章从不同角度对本文所提出的 TimeMF-SN 算法进行性能研究和分析, 并与其他个性化推荐算法和进行了性能对比。

4.1. 实验数据集

本文选择 Epinion 数据集对所提出的 TimeMF-SN 算法进行性能研究。此数据集是评估推荐算法的常用数据集, 包含了用户对项目的评分, 其中评分范围是 1 到 5。同时, 该数据集包括用户的社交关系以及对应的时间信息, 数据集的基本统计量如表 1 所示。

Table 1. Basic statistics of the Epinions dataset

表 1. Epinions 数据集基本统计量

统计量	统计数
用户数	153,265
项目数	755,760
社交关系数	841,372
单个用户评分数最大值	162,169
单个用户社交关系数最大值	2070

4.2. 算法性能比较

本文选择了几种常见的个性化推荐算法与本文所提出的 TimeMF-SN 算法进行对比, 包括矩阵分解算法 SVD++, 融合社交信息的矩阵分解算法 SoRec, 动态协同过滤算法 TimeSVD++, 具体说明如下:

- (1) SVD++: 在数据稀疏时用户显式数据量小, SVD++在传统矩阵分解算法的基础上融入用户对物品的隐式行为, 即认为评分 = 显式兴趣 + 隐式兴趣 + 偏见, 有效缓解了数据稀疏时推荐算法性能下降。
- (2) SoRec: 融合社交信息的矩阵分解算法, 在原有概率矩阵分解推荐模型的基础上增加用户社交关系矩阵, 同时将其分解为用户特征矩阵与社交特征矩阵, 迭代求解得到用户与项目特征向量后内积得到预测评分。
- (3) TimeSVD++: 动态协同过滤算法, 基于矩阵分解算法, 在预测评分时加入时间偏置以拟合时间信息对用户的影响。

对比实验中, 设置超参数 $\alpha = 0.01$, $\beta = 0.01$, $\mu = 0.5$, $\lambda_1 = 0.05$, $\lambda_2 = 0.05$, D 表示潜在特征向量的维度。进行 5 折交叉实验确定梯度下降每次更新迭代步长, 每个算法都进行了 30 次实验取平均值, 实验结果如表 2, 表 3, 表 4 所示。

Table 2. Comparison of TimeMF-SN with RMSE of other recommendation algorithms
表 2. TimeMF-SN 与其余推荐算法 RMSE 对比

	RMSE			
	F = 5	F = 10	F = 15	F = 20
SVD++	0.502	0.478	0.431	0.538
SoRec	0.653	0.607	0.673	0.702
TIMESVD++	0.471	0.453	0.449	0.421
TimeMF-SN	0.441	0.456	0.463	0.467

Table 3. Accuracy comparison between TimeMF-SN and other recommendation algorithms
表 3. TimeMF-SN 与其余推荐算法准确率对比

	Accuracy			
	F = 5	F = 10	F = 15	F = 20
SVD++	0.712	0.732	0.731	0.724
SoRec	0.795	0.798	0.792	0.801
TIMESVD++	0.843	0.823	0.854	0.801
TimeMF-SN	0.85	0.8675	0.85	0.783

Table 4. The recall rate comparison between TimeMF-SN and other recommendation algorithms
表 4. TimeMF-SN 与其余推荐算法召回率对比

	Recall			
	F = 5	F = 10	F = 15	F = 20
SVD++	0.658	0.689	0.683	0.642
SoRec	0.735	0.743	0.749	0.738
TIMESVD++	0.837	0.85	0.863	0.872
TimeMF-SN	0.918	0.880	0.858	0.795

由实验结果可以看出, 融合社交信息或融合时间信息的推荐算法表现优于传统矩阵分解算法 (SVD++), 本文提出的融合社交网络的动态推荐算法相较于只考虑社交信息 (SoRec) 或只考虑时间信息 (TimeSVD++) 的两种算法而言具有更低的 RMSE, 说明融合社交信息的动态推荐算法的确改善了推荐系统的性能, 得到的预测评分的准确性更高。

5. 总结与展望

随着 5G 时代的来临, 更多新技术不断涌现, 个性化推荐也迎来新的发展机遇, 需要不断更新迭代以适应大数据时代对推荐系统提出的更高要求[16]。随着互联网社交平台的发展, 融合社交信息的推荐算法在推荐领域的重要性与日俱增。与此同时, 大数据时代背景下, 用户每时每刻都会接收到大量的数据信息, 与之对应的是用户兴趣的变化也会随之加快, 如果将用户的所有历史行为数据视为同等重要的地位, 可能导致推荐结果反复出现用户已经失去兴趣的项目。

本文提出的 TimeMF-SN 算法从社交关系和时间信息两个方面对传统个性化推荐算法进行改进, 构建融合社交信息的动态推荐模型。通过实验验证, TimeMF-SN 算法在一定程度上缓解了数据稀疏导致的推荐准确度不高的问题, 但在算法时间复杂度、对社交好友有效信息的提取方面仍然具有一定的提升空间。除此之外, 算法在迭代求解用户和项目潜在特征向量时使用随机梯度下降算法, 虽然每次只需要选择一个样本进行迭代计算, 但由于梯度方向是随机的, 找到最优解的时间较长, 在未来的研究中可以寻找更有效的算法计算潜在特征向量的最优解。

参考文献

- [1] Eppler, M.J. and Mengis, J. (2004) The Concept of Information Overload: A Review of Literature from Organization Science, Accounting, Marketing, MIS, and Related Disciplines. *The Information Society*, **20**, 325-344. <https://doi.org/10.1080/01972240490507974>
- [2] 刘君良, 李晓光. 个性化推荐系统技术进展[J]. 计算机科学, 2020, 47(7): 47-55.
- [3] 任秋臻, 陈红梅, 周丽华. 基于时间信息表示学习的个性化推荐方法[J]. 计算机技术与发展, 2023, 33(1): 34-41.
- [4] Zeng, F., Tang, R. and Wang, Y. (2022) User Personalized Recommendation Algorithm Based on GRU Network Model in Social Networks. *Mobile Information Systems*, **2022**, Article ID: 1487586. <https://doi.org/10.1155/2022/1487586>
- [5] Yang, X. and Esquivel, J.A. (2024) Time-Aware LSTM Neural Networks for Dynamic Personalized Recommendation on Business Intelligence. *Tsinghua Science and Technology*, **29**, 185-196. <https://doi.org/10.26599/tst.2023.9010025>
- [6] Guo, G., Zhang, J. and Yorke-Smith, N. (2015) TrustSVD: Collaborative Filtering with Both the Explicit and Implicit Influence of User Trust and of Item Ratings. *Proceedings of the AAAI Conference on Artificial Intelligence*, **29**, 123-129. <https://doi.org/10.1609/aaai.v29i1.9153>
- [7] Koren, Y. (2010) Collaborative Filtering with Temporal Dynamics. *Communications of the ACM*, **53**, 89-97. <https://doi.org/10.1145/1721654.1721677>
- [8] 贾俊康, 李玲娟. 结合贡献度与时间权重的协同过滤推荐算法[J]. 计算机技术与发展, 2023, 33(3): 167-172.
- [9] Bin, S. (2023) An E-Commerce Personalized Recommendation Algorithm Based on Multiple Social Relationships. *Sustainability*, **16**, Article 362. <https://doi.org/10.3390/su16010362>
- [10] Bin Suhaim, A. and Berri, J. (2022) Directional User Similarity Model for Personalized Recommendation in Online Social Networks. *Journal of King Saud University-Computer and Information Sciences*, **34**, 10205-10216. <https://doi.org/10.1016/j.jksuci.2022.10.017>
- [11] Chen, R., Chang, Y., Hua, Q., Gao, Q., Ji, X. and Wang, B. (2020) An Enhanced Social Matrix Factorization Model for Recommendation Based on Social Networks Using Social Interaction Factors. *Multimedia Tools and Applications*, **79**, 14147-14177. <https://doi.org/10.1007/s11042-020-08620-3>
- [12] 徐上上, 孙福振, 王绍卿, 等. 基于社交信任的概率矩阵因子分解推荐算法[J]. 计算机应用与软件, 2023, 40(11): 254-258+301.
- [13] 马玲. 基于社交网络挖掘的个性化推荐算法研究[D]: [硕士学位论文]. 天津: 天津大学, 2019.
- [14] 窦润亮, 孟繁松, 南国芳, 等. 基于信任度与遗忘函数的个性化产品服务配置[J]. 管理科学学报, 2025, 28(2): 102-114.
- [15] 冯浩源. 动态用户兴趣模型构建及推荐算法研究[D]: [博士学位论文]. 天津: 天津大学, 2017.
- [16] 张晓娟, 刘怡均, 刘杰, 等. 个性化学术论文推荐研究综述[J]. 情报学报, 2024, 43(1): 106-126.