

DOI:10.13232/j.cnki.jnju.2023.05.008

融合隐式信任与属性偏好的群组推荐算法

边纪超¹, 庞继芳^{1,2*}, 宋 鹏³

(1. 山西大学计算机与信息技术学院, 太原, 030006; 2. 山西大学计算智能与中文信息处理教育部重点实验室, 太原, 030006; 3. 山西大学经济与管理学院, 太原, 030006)

摘要:随着互联网和推荐系统的不断发展,推荐服务的对象由单一用户扩展为群组成员,获取并融合组内成员的偏好、提升群组推荐效果成为当前推荐领域研究的热点问题.利用用户提供的多属性评分矩阵,提出一种融合隐式信任与属性偏好的群组推荐算法.首先,基于用户共同评分项目数和多属性评分相似度计算用户间的直接隐式信任,并利用信任传递机制获取用户间的间接信任,降低数据稀疏性.然后,通过计算用户各属性评分与总体评分间的距离来挖掘用户的属性偏好,在此基础上,利用注意力机制学习组内用户权重,将用户偏好聚合为群组偏好,进而结合深度学习框架对候选项目进行预测,生成最终的推荐列表.最后,四个数据集上的实验验证了提出的算法的有效性和可行性,实验结果表明,该算法的准确率、 $nDCG$ 等评价指标明显优于对比算法.

关键词:群组推荐,多属性评分矩阵,隐式信任,属性偏好,注意力机制

中图分类号:TP391

文献标志码:A

Group recommendation algorithm combining implicit trust and attribute preference

Bian Jichao¹, Pang Jifang^{1,2*}, Song Peng³

(1. School of Computer and Information Technology, Shanxi University, Taiyuan, 030006, China; 2. Key Laboratory of Computational Intelligence and Chinese Information Processing of Ministry of Education, Shanxi University, Taiyuan, 030006, China; 3. School of Economics and Management, Shanxi University, Taiyuan, 030006, China)

Abstract: With the continuous development of the Internet and recommendation system, the object of recommendation service expands from a single user to group members. How to obtain and integrate the preferences of group members and improve the effect of group recommendation has become a hot issue in the field of recommendation research. This paper makes full use of the user-provided multi-attribute rating matrix, and a group recommendation algorithm combining implicit trust and attribute preference is proposed. Firstly, the direct implicit trust between users is calculated based on the number of items shared by users and the similarity of multi-attribute ratings. In order to reduce the sparsity of data, the trust transfer mechanism is used to obtain indirect trust between users. Then, the user's attribute preference is mined by calculating the distance between each attribute rating and the overall rating. On this basis, the attention mechanism is used to learn the weight of users in the group, and user preferences are aggregated into group preferences. Then the deep learning framework is combined to predict candidate projects and generate the final recommendation list. Finally, experiments conducted on four datasets to verify the effectiveness and feasibility of the proposed algorithm verify its significantly superior to the compared algorithm in accuracy, $nDCG$ and other evaluation indicators.

Key words: group recommendation, multi-attribute rating matrix, implicit trust, attribute preference, attention mechanism

基金项目:国家自然科学基金(62006148,72171137),山西省1331工程

收稿日期:2023-06-27

* 通讯联系人, E-mail:purplepjf@sxu.edu.cn

随着互联网技术的快速发展,推荐系统已融入人们的日常生活^[1]. 传统推荐系统往往以单个用户为目标进行推荐,难以满足群组用户的团体需求^[2]. 近年来,群组推荐逐渐成为活跃的研究方向^[3],现有的群组推荐算法大多基于用户-项目评分矩阵中的总体评分进行推荐^[4-6],对用户偏好的把握不够精准. 为了更好地了解用户偏好,很多电商平台(如携程网、美团网、TripAdvisor等)通过多属性评分模式收集用户的多维评价信息^[4],多属性推荐算法应运而生^[7-13]. 然而,现有的研究忽略了多属性评分中蕴含的隐性信息,对用户偏好的刻画不够全面. 此外,随着信任网络的出现和发展,用户之间的信任关系也被引入推荐系统. 考虑到信任数据通常难以获取,部分学者对基于隐性信任关系的推荐算法进行了初步的探索和研究^[14-19]. 目前,基于信任网络的推荐算法主要利用出入度来表示用户间的信任关系,但是,对隐式信任的计算比较简单、直接,缺乏对信任关系的多角度分析和综合度量. 为了提升群组推荐性能,研究者将深度学习技术融入群组推荐算法,通过注意力机制将用户偏好聚合为群组偏好^[20-21],并利用神经网络来学习用户-项目之间的关联^[22-25],但现有算法没有充分考虑组内用户间的关系,一定程度上限制了群组偏好的聚合效果.

针对上述问题,本文提出一种融合隐式信任与属性偏好的群组推荐算法(Group Recommendation Algorithm Combining Implicit Trust and Attribute Preference, ITAP),其主要贡献是:

(1) 基于多属性评分矩阵从多角度挖掘用户间的隐式信任,并利用信任传递机制对信任矩阵进行填充,最大限度地获取用户间的隐式信任.

(2) 综合考虑用户的总体评分及多属性评分信息,通过分析和度量二者之间的关系来建立组内用户关于属性的偏好矩阵.

(3) 在融合用户隐式信任和属性偏好的基础上,利用注意力机制聚合群组偏好,并结合深度学习框架生成群组推荐列表.

1 相关工作

1.1 基于多属性评分的推荐算法研究现状 为了更加全面、精准地捕获用户偏好,部分学者结合

多属性评分矩阵对推荐算法进行了研究. Zhu et al^[7]基于多属性评分信息,提出一种面向群组的高效提升算法(Efficient Promotion Algorithm in Recommendation, PromoRec). Biswas and Liu^[8]基于多属性评分对用户进行画像,并使用线性回归算法学习用户对各个属性的偏好程度. Papamitsiou and Economides^[9]提出一种基于多属性评分的协同过滤算法,能发现偏好相似的邻居并通过相似邻居完成推荐. 朱文强等^[10]基于多属性评分建立相似度计算公式,通过寻找相似用户来预测用户在每个属性上的评分,再根据预测评分对用户进行推荐. 覃正和李岱峰^[11]建立了一种面向群组的多属性推荐模型,通过模拟退火的方法学习各成员在各属性上的偏好权重,聚合组内所有成员的偏好. Nilashi et al^[12]在使用决策树对多属性评分进行处理的基础上,提出一种多属性协同过滤推荐算法. Majumder et al^[13]发展了一种基于多属性评分的矩阵分解方法来提高推荐的准确率. 但上述算法仅仅分析了用户在各个属性上的偏好,本文充分挖掘用户在不同属性上的评分与总体评分之间的关系,对多属性评分进行更加全面的分析和利用,提高群组推荐效果.

1.2 融合信任关系的推荐算法研究现状 随着信任网络的出现和发展,信任关系逐渐被引入推荐系统. Da' u et al^[14]提出一种根据信任网络的出入度来重构信任度的混合推荐算法. Guo et al^[15]提出一种融合项目流行度和用户信任关系的矩阵分解推荐算法. Chen et al^[16]结合信任关系和协同过滤建立了一种随机游走模型来提高推荐性能. Victor et al^[17]将用户交互信息作为信任关系,提出一种社交推荐(Social Recommendation, SoRec)方法. 由于隐私问题,获取用户之间的信任分数通常很困难,当用户之间的信任关系无法直接获知时,可以通过分析用户行为、社交网络等隐含信息来进行推断. 近年来,越来越多的研究开始探索隐性信任的提取方法. 刘辉等^[18]利用用户之间的相似度来计算信任度,提出一种整合用户相似度、地理位置和信任关系的混合推荐算法. 张槟淇等^[19]根据用户的影响力来计算用户信任度,提出一种融合信任关系的群组推荐算法. 上述算法虽然能得到用户之间的信任关系,但缺少对用户

间信任程度的客观度量. 本文基于多属性评分来挖掘用户之间的隐式信任关系, 结合用户在共同评分项目上的数量和多属性评分的相似度两个方面来获得用户之间的信任程度, 能更好地刻画用户信任关系, 获得更精准的群组偏好.

1.3 基于深度学习的推荐算法研究现状 基于深度学习的推荐算法大多是将用户和项目嵌入神经网络, 从用户、群组和项目之间的交互数据中学习用户偏好. Choudhary et al^[20]通过群组和用户组成的双层注意网络来学习群组内成员权重. Cao et al^[21]利用注意力机制, 动态地调整用户在不同群组中的影响力, 较好地解释了群组偏好的集结过程. Wang et al^[22]利用神经注意网络, 以动态方式来学习成员之间的相对影响, 利用组内的成员偏好来表示群组偏好. He et al^[24]利用自注意力机制来挖掘用户与项目间的潜在信息, 进而学习得到群组的潜在偏好. Jia et al^[25]提出一种由成员级偏好网络和组级偏好网络构成的双通道超图卷积网络. 深度学习技术与群组推荐算法的结合有效地促进了群组偏好融合研究, 但现有研究大多根据评分来获取用户与项目之间的交互, 没有充分考虑用户之间的关系, 导致获取的群组偏好不够全面, 影响推荐效果. 本文在利用神经网络学习用户-项目交互的基础上, 引入用户之间的信任关系, 能获得更全面的群组偏好.

2 问题描述

在多属性评分矩阵中, 除了常用的总体评分外, 还包括用户对项目各个属性的细粒度评分, 其中既隐含用户之间的信任程度, 又潜藏着用户对不同属性的重视程度和个体偏好. 充分利用并挖掘用户提供的多属性评分矩阵有助于获得更加细致、全面的群组偏好, 提升群组推荐的效果.

定义1 多属性评分矩阵 设用户集合为 $U = \{u_1, u_2, \dots, u_n\}$, 项目集合为 $I = \{i_1, i_2, \dots, i_m\}$, 属性集合为 $A = \{a_1, a_2, \dots, a_k\}$. 用户 $u_b (1 \leq b \leq n)$ 对项目 $i_j (1 \leq j \leq m)$ 的评分矩阵为 $R_{n \times m} = [r_{bj}]_{n \times m}$, 其中 $r_{bj} = [r_{bj}^0, r_{bj}^1, r_{bj}^2, \dots, r_{bj}^k]$ 表示 u_b 对 i_j 的评分向量, r_{bj}^0 为 u_b 对 i_j 的总体评分, r_{bj}^k 为 u_b 对 i_j 在属性 a_k 下的评分. 多属性评分矩阵的具体形式如下:

$$R_{n \times m} = \begin{pmatrix} r_{11} & \cdots & r_{1m} \\ \vdots & \ddots & \vdots \\ r_{n1} & \cdots & r_{nm} \end{pmatrix} \quad (1)$$

3 ITAP模型

融合隐式信任与属性偏好的群组推荐算法 (ITAP) 由五个部分组成: 数据准备阶段、输入层、池化层、隐藏层、输出层. 在数据准备阶段, 基于多属性评分矩阵获得群组 g_i 、项目 i_j 、混合隐式信任矩阵 $Trust_{n \times n}$ 与属性偏好矩阵 $Prefer_{n \times k}$. 通过对 g_i 、 $Trust_{n \times n}$ 与 $Prefer_{n \times k}$ 进行 one-hot 编码及拆分, 得到群组 g_i 内的用户嵌入向量 $\{u_b, u_c, \dots\}$ 、信任嵌入向量 $\{Trust_b, Trust_c, \dots\}$ 以及属性偏好向量 $\{Prefer_b, Prefer_c, \dots\}$. 输入层在获得数据准备阶段的四类数据 $u_b, i_j, Trust_b$ 与 $Prefer_b$ 后, 利用注意力机制学习用户权重并得到聚合后的群组偏好 $g_i(j)$. 池化层通过对群组偏好 $g_i(j)$ 和项目嵌入 i_j 进行内积运算, 提取二者之间的相互作用. 隐藏层利用多层感知机对池化后的数据进行处理, 捕获群组与项目之间的非线性特征和高阶相关性. 最后, 在输出层生成群组对于候选项目的预测评分. ITAP算法的整体架构如图1所示.

3.1 数据准备阶段 数据准备阶段为本文的主要工作, 在对群组、项目进行 one-hot 编码的基础上增加了两方面的内容. (1) 基于多属性评分矩阵构建混合隐式信任矩阵 $Trust_{n \times n}$. 考虑到用户之间的信任度通常是双向的而且是不对称的, 首先基于用户共同评分项目计算用户置信度, 进而基于多属性评分计算用户相似度, 综合用户置信度与评分相似度得到用户之间的直接隐式信任关系. 在此基础上, 进一步利用信任传递机制^[26]挖掘用户之间的间接隐式信任关系, 对直接信任矩阵进行补全, 得到混合隐式信任矩阵, 减少数据稀疏性的影响. (2) 结合多属性评分与总体评分构建用户属性偏好矩阵 $Prefer_{n \times k}$. 通过距离公式对用户各属性评分与总体评分之间的相关性进行量化分析, 并据此计算用户在不同属性上的偏好程度, 获得更加精准的用户偏好.

3.1.1 混合隐式信任矩阵 混合隐式信任矩阵由直接隐式信任矩阵和间接隐式信任矩阵两部分

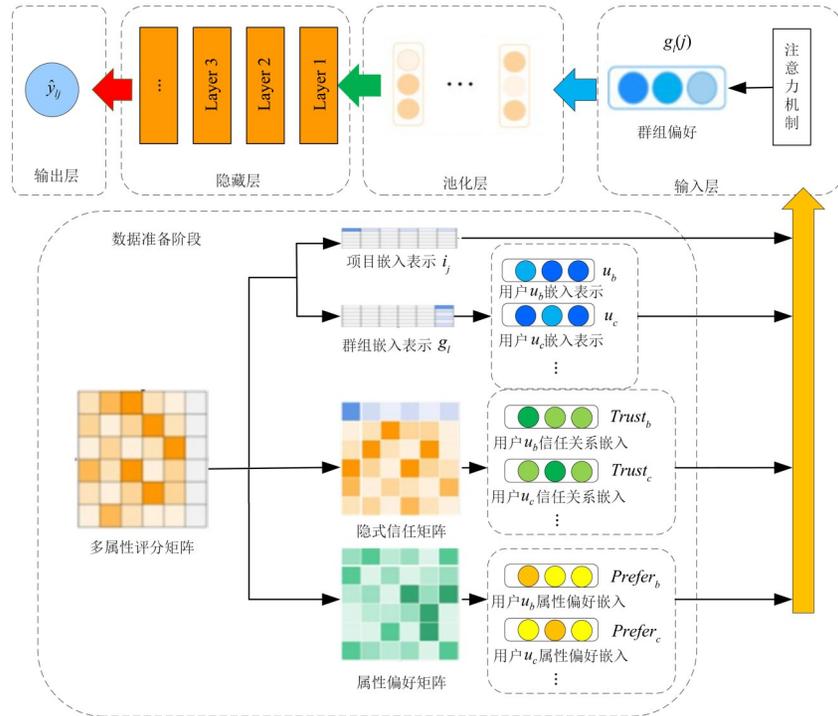


图 1 ITAP算法的框架

Fig. 1 The framework of ITAP algorithm

构成. 其中, 用户间的直接隐式信任度从两个方面进行刻画: (1) 基于共同评分项目的用户置信度; (2) 基于多属性评分的用户相似度.

首先, 通过用户共同评分的项目数量与用户自身评分的项目数量之比来计算基于共同评分项目的用户置信度. 用户 u_b 对 u_c 的置信度为:

$$Con_{bc} = \frac{|I(u_b) \cap I(u_c)|}{|I(u_b)|} \quad (2)$$

其中, $|I(u_b)|$ 为用户 u_b 评分的项目数量, $I(u_b) \cap I(u_c)$ 为用户 u_b 与 u_c 共同评分的项目集合.

然后, 使用皮尔逊相关系数 (Pearson Correlation Coefficient, PCC) 计算^[16] 基于多属性评分的用户相似度. 用户 u_b 与 u_c 在属性 $a_v (1 \leq v \leq k)$ 上的评分相似度 $sim_v(u_b, u_c)$ 的计算如下:

$$sim_v(u_b, u_c) = \frac{\sum_{i_j \in I(u_b) \cap I(u_c)} (r_{bj}^v - \bar{r}_b^v) \cdot (r_{cj}^v - \bar{r}_c^v)}{\sqrt{\sum_{i_j \in I(u_b) \cap I(u_c)} (r_{bj}^v - \bar{r}_b^v)^2} \cdot \sqrt{\sum_{i_j \in I(u_b) \cap I(u_c)} (r_{cj}^v - \bar{r}_c^v)^2}} \quad (3)$$

$$Tr_{bc} = T(Tru_{b,\sigma(1)}, Tru_{\sigma(1),c}) = \frac{Tru_{b,\sigma(1)} \cdot Tru_{\sigma(1),c}}{1 + (1 - Tru_{b,\sigma(1)}) (1 - Tru_{\sigma(1),c})} \quad (6)$$

其中, r_{bj}^v 表示用户 u_b 对于项目 i_j 在属性 a_v 下的评分, $\bar{r}_b^v = \frac{1}{|I(b)|} \sum_{i_j \in I(b)} r_{bj}^v$ 表示用户 u_b 在属性 a_v 下所有评价项目的平均评分.

用户 u_b 与 u_c 在所有属性上的平均相似度为:

$$Sim_{bc} = \frac{1}{k} \sum_{v=1}^k sim_v(u_b, u_c) \quad (4)$$

综合用户置信度与用户相似度可得用户 u_b 对于 u_c 的直接隐式信任度 Tru_{bc} :

$$Tru_{bc} = \frac{1}{2} (Con_{bc} + Sim_{bc}) \quad (5)$$

考虑到现实生活中一些用户的评分项目比较少, 无法直接计算用户间的信任度, 导致得到的直接隐式信任矩阵比较稀疏, 为此, 使用 Victor et al^[16] 的基于 T 范数的信任传播方法来计算间接隐式信任度, 对直接隐式信任矩阵进行补充, 从而获得混合隐式信任矩阵 $Trust_{n \times n}$.

设 $u_b \rightarrow u_{\sigma(1)} \rightarrow u_{\sigma(2)} \rightarrow \dots \rightarrow u_{\sigma(q)} \rightarrow u_c$ 是一条用户 u_b 到 u_c 的路径, 路径长度为 q . 当用户 u_b 到 u_c 的路径长度为 2 时, 用户 u_b 对 u_c 的间接信任度为:

当用户 u_b 到 u_c 的路径长度大于2时,用户 u_b 对 u_c 的间接信任度为:

$$Tr_{bc} = T(Tru_{b,\sigma(1)}, Tru_{\sigma(1),\sigma(2)}, \dots, Tru_{\sigma(q),c}) = \frac{2 \cdot Tru_{b,\sigma(1)} \cdot Tru_{\sigma(q),c} \prod_{p=1}^{q-1} Tru_{\sigma(p),\sigma(p+1)}}{(2 - Tru_{b,\sigma(1)}) (2 - Tru_{\sigma(q),c}) \prod_{p=1}^{q-1} (2 - Tru_{\sigma(p),\sigma(p+1)})} + Tru_{b,\sigma(1)} \cdot Tru_{\sigma(q),c} \prod_{p=1}^{q-1} Tru_{\sigma(p),\sigma(p+1)} \quad (7)$$

若用户 u_b 和 u_c 之间有多条不同的路径,则在每条路径上都存在一个间接信任度. 集结所有路径上的间接信任度可得用户间总的间接信任度.

采用 Victor et al^[17] 的有序加权平均(OWA)算子来集结不同路径的信任度. 由于OWA算子能够灵活反映不同的集结度,其在群体决策的信息集结过程中得到了广泛的应用.

假设从 u_b 到 u_c 有 f 条路径,设 $Tr_{bc}^1, Tr_{bc}^2, \dots, Tr_{bc}^f$ ($Tr_{bc}^1 \geq Tr_{bc}^2 \geq \dots \geq Tr_{bc}^f$) 分别表示 f 条路径的间接信任度,则总的间接信任度 \widetilde{Tru}_{bc} :

$$\widetilde{Tru}_{bc} = OWA(Tr_{bc}^1, Tr_{bc}^2, \dots, Tr_{bc}^f) = \sum_{t=1}^f \rho_t \cdot Tr_{bc}^t \quad (8)$$

其中, Tr_{bc}^t 表示 $\{Tr_{bc}^1, Tr_{bc}^2, \dots, Tr_{bc}^f\}$ 的第 t 条路径.

$\rho = (\rho_1, \rho_2, \dots, \rho_f)^T$ 表示路径权重.

ρ 的取值采用 Victor et al^[17] 的基于语言量词 Q 的量词引导方法计算得到,即:

$$\rho_t = Q\left(\frac{t}{f}\right) - Q\left(\frac{t-1}{f}\right), t = 1, 2, \dots, f \quad (9)$$

利用间接信任度对直接隐式信任矩阵中的缺失值进行填充,可得混合隐式信任矩阵:

$$Trust_{n \times n} = (Tru_{bc})_{n \times n} + (\widetilde{Tru}_{bc})_{n \times n}$$

3.1.2 属性偏好矩阵 用户的属性偏好可通过计算其在某个属性上的评分与总体评分之间的距离来刻画. 用户 u_b 在属性 a_v 上的评分 r_{bj}^v 与总体评分 r_{bj}^0 之间的距离如下:

$$dis(r_b^0, r_b^v) = \sqrt{\frac{1}{|I(u_b)|} \sum_{i_j \in I(u_b)} (r_{bj}^0 - r_{bj}^v)^2} \quad (10)$$

基于距离测度 $dis(r_b^0, r_b^v)$ 可进一步得到用户 u_b 的属性 a_v 与总体评分的相关度:

$$Cor_{bk} = \frac{1}{1 + dis(r_b^0, r_b^v)} \quad (11)$$

通过对相关度 Cor_{bk} 进行归一化得到用户的属性偏好 $Prefer_{bv}$:

$$Prefer_{bv} = Cor_{bk} / \sum_{v=1}^k Cor_{bk} \quad (12)$$

利用式(10)~(12)计算组内所有用户在所有属性上的偏好,得到群组的属性偏好矩阵 $Prefer_{n \times k}$.

3.2 输入层表示 输入层在得到数据准备阶段提供的 $u_b, i_j, Trust_b, Prefer_b$ 后,利用注意力机制获取群组 g_l 对项目 i_j 的偏好 $g_l(j)$,计算如下:

$$g_l(j) = \sum_{u_b \in g_l} \alpha(j, b) u_b \quad (13)$$

其中, $\alpha(j, b)$ 表示群组成员 u_b 在群组 g_l 内的影响权重.

$\alpha(j, b)$ 的计算如下:

$$o(j, b) = h^T \text{ReLU}(P_T i_j + Prefer_b P_P + Trust_b P_T + P_U u_b + bi) \quad (14)$$

$$\alpha(j, b) = \text{Softmax}(o(j, b)) = \frac{\exp(o(j, b))}{\sum_{u_b \in g_l} \exp(o(j, b))} \quad (15)$$

其中, P_T 表示用户 u_b 的隐式信任影响权重矩阵, P_P 表示用户 u_b 的属性偏好影响权重矩阵, P_I 和 P_U 分别表示项目嵌入和用户嵌入的权重矩阵, bi 表示偏置向量, h 为权重向量.

输入层具体流程如图2所示.

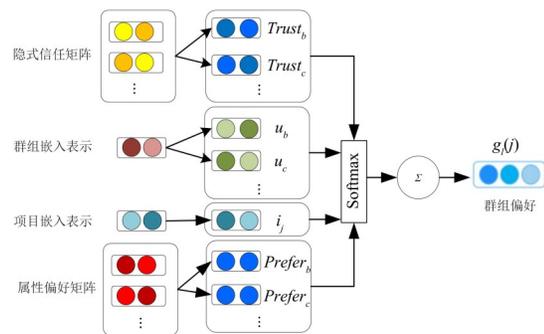


图2 基于注意力机制的群组偏好获取

Fig. 2 Group preference acquisition based on attention mechanism

3.3 池化层表示 池化层主要是对群组偏好表示 $g_l(j)$ 与项目嵌入表示向量 i_j 进行内积运算,其目的是通过在每个嵌入维度上使用乘法来获得两个嵌入向量之间的相互作用^[15]. 为了避免内积运

算中可能出现的原始嵌入信息丢失的情况^[23],可将原始嵌入信息与内积结果进行拼接,得到最终的池化层表示向量 e_0 ,具体计算如下:

$$e_0 = \varphi_{\text{pooling}}(g_l(j), i_j) = \begin{bmatrix} g_l(j) \odot i_j \\ g_l(j) \\ i_j \end{bmatrix} \quad (16)$$

其中, \odot 表示内积运算.

3.4 隐藏层表示 隐藏层的任务是对池化层传入的表示向量 e_0 进行变换,通过多层感知机来捕获群组偏好嵌入 $g_l(j)$ 与项目嵌入 i_j 之间的非线性特征和高阶相关性.计算如下:

$$\begin{cases} e_1 = \text{ReLU}(W_1 e_0 + b_1) \\ e_2 = \text{ReLU}(W_2 e_1 + b_2) \\ \vdots \\ e_h = \text{ReLU}(W_h e_{h-1} + b_h) \end{cases} \quad (17)$$

其中, h 表示隐藏层层数,ReLU为激活函数, b_1, b_2, \dots, b_h 为可训练的偏置向量, W_1, W_2, \dots, W_h 为可训练的隐藏层权重矩阵, e_h 为 e_0 经过多层感知机后的输出向量.

3.5 输出层表示 输出层将由隐藏层传入的最终表示向量 e_h 转换为预测评分 \hat{y}_{ij} ,具体计算如下:

$$\hat{y}_{ij} = w^T e_h \quad (18)$$

其中, w 为可训练的输出层权重矩阵.

通过对不同项目的预测评分进行排序,生成最终的群组推荐列表.

3.6 损失函数 得到预测评分后,利用均方差损失函数对模型进行优化,具体计算如下:

$$\text{Loss} = \frac{1}{M} \sum_{i_j=1}^M (y_{ij} - \hat{y}_{ij})^2 \quad (19)$$

其中, M 表示群组中候选项目数量, y_{ij} 表示训练集中群组 g_l 对项目 i_j 的实际评分, \hat{y}_{ij} 表示 g_l 对 i_j 的预测评分.

在优化策略选择方面,考虑小批量梯度下降(Mini Batch Gradient Descent, MBSD)能比传统的随机梯度下降(Stochastic Gradient Descent, SGD)更快地达到最优解^[27],故本文采用MBSD策略来最小化损失函数.

4 实验分析

4.1 实验数据 在大众点评、Yelp2018、Ratebeer和BeerAdvocate四个数据集上进行实验,数

据集包含用户ID、项目ID、多属性评分、评论时间等信息.随机划分其中80%的数据为训练集,20%的数据为测试集^[28].数据集的信息如表1所示.

表1 实验使用的数据集

Table 1 Datasets used in experiments

| 数据集 | 大众点评 | Yelp2018 | Ratebeer | BeerAdvocate |
|------|-------------|----------------|---------------|---------------|
| 用户数量 | 42706 | 31667 | 40213 | 3389 |
| 项目数量 | 3247 | 8047 | 10419 | 6051 |
| 评分数量 | 422473 | 824349 | 2855232 | 586259 |
| 稀疏度 | 99.69% | 99.67% | 99.31% | 97.14% |
| 评分 | {1,2,3,4,5} | {1,2,3,4,5} | {1,2,3,4,5} | {1,2,3,4,5} |
| 属性集合 | {口味,环境,服务} | {口味,环境,服务,性价比} | {味道,外观,感觉,嗅觉} | {味道,外观,感觉,嗅觉} |

4.2 实验设置及评估标准 群组的形成^[29]是群组推荐的首要环节,通过聚类算法找到一群具有相似兴趣的用户是当前常用的创建群组的方法.K-means作为经典的聚类算法,在以往的研究中获得了很好的效果^[29].因此,本文选择K-means聚类来自动生成群组.

ITAP算法中的主要参数设置如表2所示.

表2 参数设置

Table 2 Parameter settings

| 参数 | 取值 |
|-----------|-----|
| 嵌入向量维数 | 32 |
| 小批量学习样本大小 | 128 |

为了衡量ITAP算法和基线方法的性能,采用标准差(Standard Deviation, STD)^[19]、归一化折损累计增益(Normalized Discounted Cumulative Gain, nDCG)^[20]和均方根误差(Root Mean Square Error, RMSE)^[21]三个评估指标对算法进行比较分析.具体定义如下.

定义2 STD指标 用来衡量数据集中数据点的离散程度,计算如下:

$$\text{STD} = \sqrt{\frac{1}{M} \sum_{j=1}^M (y_{ij} - \bar{y}_{ij})^2} \quad (20)$$

STD越低说明数据越集中,推荐结果越准确.

定义3 nDCG指标 用来衡量生成Top-K推荐列表的命中率,计算如下:

$$\text{nDCG}_K = \frac{\text{DCG}_K}{\max \text{DCG}_K} \quad (21)$$

其中, $DCG_K = \sum_{i_j=1}^K \frac{2^{rel(i_j)} + 1}{\log_2(i_j + 1)}$, $rel(i_j)$ 是项目 i_j 的相关性得分.

$nDCG_K$ 越大, 表示推荐的命中率越高.

定义 4 RMSE 指标 用来衡量用户的真实评分与预测评分之间的误差, 计算如下:

$$RMSE = \sqrt{\frac{1}{M} \sum_{j=1}^M (y_{ij} - \hat{y}_{ij})^2} \quad (22)$$

RMSE 越小, 表示推荐的准确度越高.

4.3 对比算法 为了评价 ITAP 算法的综合性能, 将其与 AGREE, COM, PromoRec, TrustSVD, HCR 和 HHGR 六种算法进行对比分析. 表 3 展示了七种算法的异同.

表 3 七种算法的比较

Table 3 Comparison of seven algorithms

| 算法名称 | 简称 | 群组推荐 | 总体评分 | 多属性评分 | 信任关系 |
|--------------------|--------------------------|------|------|-------|------|
| 基于注意力机制的群组推荐算法 | AGREE ^[21] | ✓ | ✓ | | |
| 基于群组共识的群组推荐算法 | COM ^[30] | ✓ | ✓ | | |
| 基于群体偏好的群组推荐算法 | PromoRec ^[7] | ✓ | | ✓ | |
| 融合用户信任的协同过滤推荐算法 | TrustSVD ^[15] | | ✓ | | ✓ |
| 基于超图卷积网络的群组推荐算法 | HCR ^[25] | ✓ | ✓ | | |
| 基于自监督超图学习的群组推荐算法 | HHGR ^[27] | ✓ | ✓ | | |
| 融合隐式信任与属性偏好的群组推荐算法 | ITAP | ✓ | ✓ | ✓ | ✓ |

4.4 参数敏感性分析

4.4.1 嵌入维度 对于嵌入维度的选取, 一般需要通过具体的任务来进行评测, 维度过低则模型的代表能力不够, 维度过高则容易导致过拟合. 为了说明嵌入维度对 ITAP 性能的影响, 将小批量学习样本的大小固定为 128, 迭代次数固定为 10, TOP-K 值固定为 10, 嵌入维度分别设为 [8, 16, 32, 64, 128] 进行实验. 实验结果如图 3 所示.

由图可见, 当嵌入维度在 8~32 依次递增时, RMSE 逐渐降低, nDCG 逐渐提高; 当嵌入维度在 32~128 时, RMSE 逐渐升高, nDCG 逐渐降低. 因此将 ITAP 的嵌入维度设为 32.

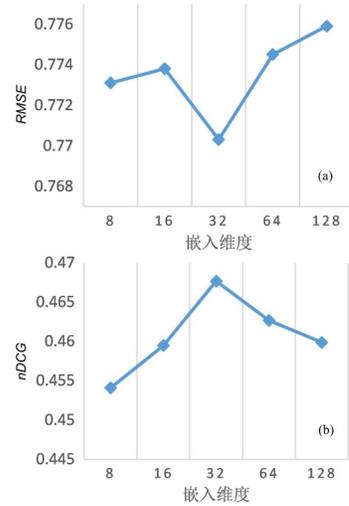


图 3 不同嵌入维度下的实验结果

Fig. 3 Experimental results with different embedding dimensions

4.4.2 批量学习样本大小 批量学习样本数越大, 其确定的下降方向越准确, 训练所需的迭代次数越少, 但当批量学习样本数增大到一定程度时, 其确定的下降方向已经基本不再变化. 为了说明批量学习样本数对 ITAP 性能的影响, 将嵌入维度固定为 32, 迭代次数固定为 10, TOP-K 值固定为 10, 批量学习样本数分别设为 [32, 64, 128, 256] 进行实验. 实验结果如图 4 所示.

由图可见, 当批量学习样本数为 32~128 时, RMSE 逐渐降低, nDCG 逐渐提高, 这是因为随着批量学习样本数的增加, 训练效率更高. 当批量

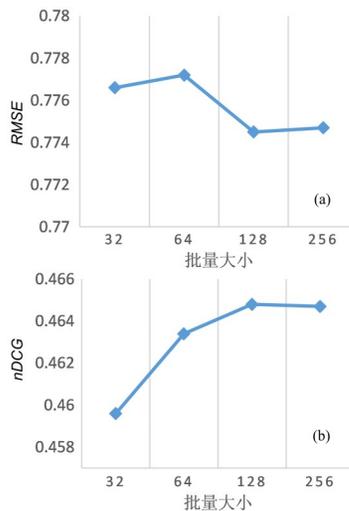


图 4 不同批量学习样本下的实验结果

Fig. 4 Experimental results with different batch sizes

学习样本数在 128~256 时, $RMSE$ 与 $nDCG$ 趋于稳定,但在相同精度下,运行时间大大增加. 因此选择 128 作为 ITAP 的批量学习样本的大小.

4.5 对比实验及分析 图 5 展示了七种算法在四个数据集上的 STD 与 $RMSE$. 由图可见,随着 TOP-K 的扩大,七种算法的 STD 与 $RMSE$ 逐渐降低. 在绝大多数情况下,ITAP 算法的 STD 与 $RMSE$ 更小,表现出更高的推荐准确性.

图 6 展示了七种算法在不同数据集上的 $nDCG$. 由图可见,随着 TOP-K 的扩大,七种算法在不同数据集上的 $nDCG$ 均逐渐增大,其中,ITAP 算法的 $nDCG$ 均比其他对比算法更高,这主要是因为 ITAP 算法考虑了隐式信任与属性偏好,能够更加全面、细致地刻画群组偏好,所以表现出更高的推荐准确率.

通过上述对比实验可知,ITAP 算法的三种评价指标均表现良好,充分说明在总体评分的基础上,融合用户间的隐式信任与属性偏好可以获得更加精确的群组偏好,能有效增强推荐效果.

4.6 消融实验 通过消融实验进一步说明隐式

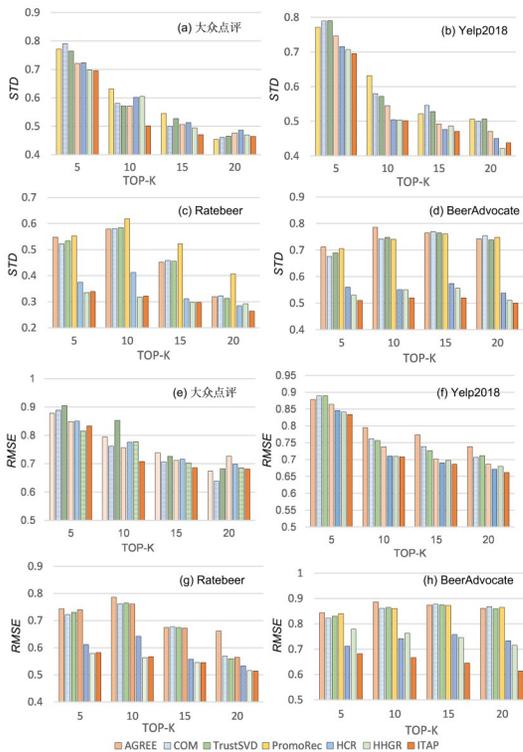


图 5 七种算法在四个数据集上的 STD 和 $RMSE$ 的比较
Fig.5 STD and $RMSE$ of seven algorithms on four datasets

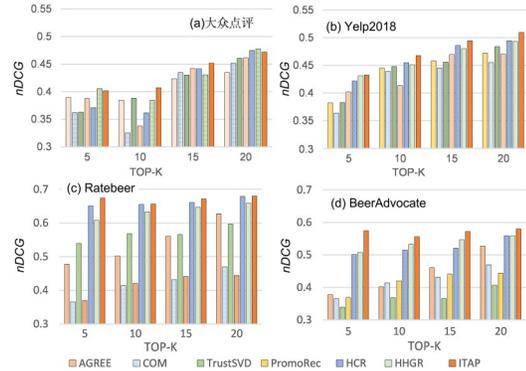


图 6 七种算法在四个数据集上的 $nDCG$ 比较
Fig.6 $nDCG$ of seven algorithms on four datasets

信任和属性偏好在学习群组偏好中的作用. 表 4 展示了三种变体算法 ITAP-T, ITAP-A 和 ITAP-M 与本文算法 ITAP 之间的关系.

表 4 变体算法的介绍

Table 4 The introduction of variant algorithms

| 算法名称 | 隐式信任 | 注意力机制 | 属性偏好 |
|--------|------|-------|------|
| ITAP-T | × | √ | √ |
| ITAP-A | √ | × | √ |
| ITAP-M | √ | √ | × |
| ITAP | √ | √ | √ |

图 7 为四种算法在不同数据集上的 $RMSE$ 和 $nDCG$ 的比较. 由图可见,随着 TOP-K 的扩大,四种算法在不同数据集上的 $RMSE$ 均逐渐下降, $nDCG$ 均逐渐上升,其中,ITAP 算法的 $RMSE$ 均比其他算法低,而 $nDCG$ 均比其他算法高.

通过实验结果对比分析可知,随着推荐列表长度的增加,隐式信任与属性偏好在整个推荐过程中所起的作用越显著,所以推荐效果更好.

5 结论

本文提出的融合隐式信任与属性偏好的群组推荐算法(ITAP)综合考虑了用户的总体评分和多属性评分,从多个角度挖掘组内用户间的混合隐式信任关系,建立了属性偏好矩阵,通过注意力机制融合各类数据,可以获得更准确的群组偏好. 本文算法具有较强的客观性和综合性,在降低数据稀疏性的同时有效提高了推荐准确率,在四个数据集上的多个对比实验和消融实验验证了本文算法的有效性.

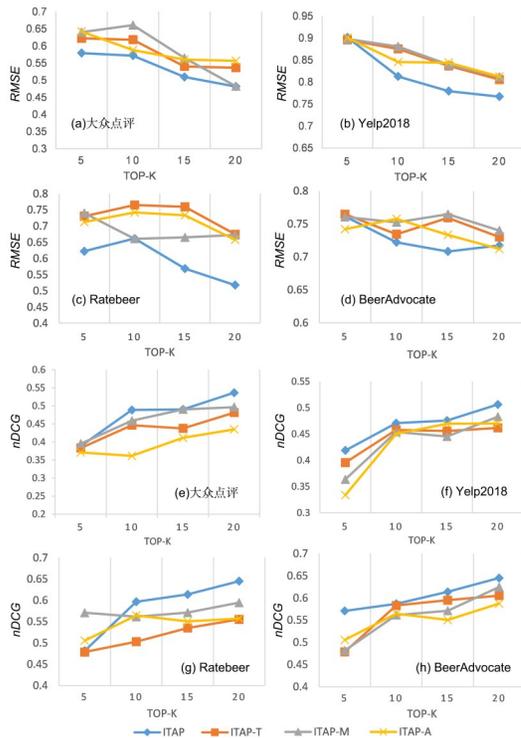


图7 四种算法在四个数据集上的RMSE和nDCG的比较
Fig.7 RMSE and nDCG of four algorithms on four datasets

未来将考虑群组推荐过程中用户属性(如年龄、性别、工作等)、项目属性(如价格、服务等)及地理位置等辅助信息,通过聚合多维用户偏好来实现群组画像的多兴趣提取,获得更好的推荐效果。

参考文献

- [1] Shrivastava R, Sisodia D S, Nagwani N K. Deep neural network-based multi-stakeholder recommendation system exploiting multi-criteria ratings for preference learning. *Expert Systems with Applications*, 2023(213):119071.
- [2] Wang F Y, Lu X Q, Lyu L. CGSNet: Contrastive graph self-attention network for session-based recommendation. *Knowledge-Based Systems*, 2022(251):109282.
- [3] Pham P, Nguyen L T T, Nguyen N T, et al. A hierarchical fused fuzzy deep neural network with heterogeneous network embedding for recommendation. *Information Sciences*, 2023(620):105-124.
- [4] 李琳,王培培,杜佳,等.融合纳什均衡策略和神经协同过滤的群组推荐方法. *模式识别与人工智能*, 2022, 35(5):412-421. (Li L, Wang P P, Du J, et al. Group recommendation method with Nash equilibrium strategy and neural collaborative filtering. *Pattern Recognition and Artificial Intelligence*, 2022, 35(5):412-421.)
- [5] Guan B Y, Hu L T, Liu P C, et al. dpSmart: A flexible group based recommendation framework for digital repository systems//2019 IEEE International Congress on Big Data. Milan, Italy: IEEE, 2019: 111-120.
- [6] 宋玉龙,马文明,刘彤彤.融合用户信任度的概率矩阵分解群组推荐算法. *计算机工程*, 2022, 48(1):105-111. (Song Y L, Ma W M, Liu T T. Group recommendation algorithm incorporating user trust with probability matrix factorization. *Computer Engineering*, 2022, 48(1):105-111.)
- [7] Zhu Q, Zhou M X, Liang J F, et al. Efficient promotion algorithm by exploring group preference in recommendation//Proceedings of 2016 IEEE International Conference on Web Services. San Francisco, CA, USA: IEEE, 2016:268-275.
- [8] Biswas P K, Liu S L. A hybrid recommender system for recommending smartphones to prospective customers. *Expert Systems with Applications*, 2022(208):118058.
- [9] Papamitsiou Z, Economides A A. Recommendation of educational resources to groups: A game-theoretic approach//2018 IEEE Global Engineering Education Conference. Santa Cruz de Tenerife, Spain: IEEE, 2018:760-767.
- [10] 朱文强,徐军.融合社交信任的多属性元路径好友推荐方法. *小型微型计算机系统*, 2020, 41(10):2036-2044. (Zhu W Q, Xu J. Friend recommendation method based on multi-attribute meta-path with social trust. *Journal of Chinese Computer Systems*, 2020, 41(10):2036-2044.)
- [11] 覃正,李岱峰.一种基于资源多属性分类的群组推荐模型. *统计与决策*, 2010(8):153-155. (Qin Z, Li D F. A Group recommendation model based on multi-attribute classification of resources. *Statistics and Decision*, 2010(8):153-155.)
- [12] Nilashi M, Esfahani M D, Roudbaraki M Z, et al. A Multi-criteria collaborative filtering recommender system using clustering and regression techniques. *Journal of Soft Computing and Decision Support Systems*, 2016, 3(5):24-30.

- [13] Majumder G S, Dwivedi P, Kant V. Matrix factorization and regression-based approach for multi-criteria recommender system//Proceedings of the International Conference on Information & Communication Technology for Intelligent Systems. Ahmedabad, India: Springer Cham, 2017: 103–110.
- [14] Da'u A, Salim N, Idris R. An adaptive deep learning method for item recommendation system. Knowledge-Based Systems, 2021(213): 106681.
- [15] Guo G B, Zhang J, Yorke-Smith N. TrustSVD: Collaborative filtering with both the explicit and implicit influence of user trust and of item ratings//Proceedings of the 29th AAAI Conference on Artificial Intelligence. Austin, Texas, USA: AAAI Press, 2015: 123–129.
- [16] Chen Y L, Cheng L C, Chuang C N. A group recommendation system with consideration of interactions among group members. Expert Systems with Applications, 2008, 34(3): 2082–2090.
- [17] Victor P, Cornelis C, De Cock M, et al. Practical aggregation operators for gradual trust and distrust. Fuzzy Sets and Systems, 2011, 184(1): 126–147.
- [18] 刘辉, 曾斌, 刘子恺. 融合邻居选择策略和信任关系的兴趣点推荐. 计算机工程与科学, 2020, 42(2): 365–372. (Liu H, Zeng B, Liu Z K. Point-of-interest recommendation based on neighbor selection strategy and trust relationship. Computer Engineering & Science, 2020, 42(2): 365–372).
- [19] 张滨淇, 任丽芳, 王文剑. 融合信任隐性影响和信任度的推荐模型. 模式识别与人工智能, 2022, 35(4): 374–385. (Zhang B Q, Ren L F, Wang W J. Recommendation model combining implicit influence of trust with trust degree. Pattern Recognition and Artificial Intelligence, 2022, 35(4): 374–385.)
- [20] Choudhary N, Minz S, Bharadwaj K K. Circle-based group recommendation in social networks. Soft Computing, 2021, 25(22): 13959–13973.
- [21] Cao D, He X N, Miao L H, et al. Attentive group recommendation//Proceedings of the 41st International ACM SIGIR Conference on Research & Development in Information Retrieval. Ann Arbor, MI, USA: ACM, 2018: 645–654.
- [22] Wang P P, Li L, Wang R, et al. Socially-driven multi-interaction attentive group representation learning for group recommendation. Pattern Recognition Letters, 2021(145): 74–80.
- [23] Du J, Li L, Gu P, et al. A group recommendation approach based on neural network collaborative filtering//2019 IEEE 35th International Conference on Data Engineering Workshops. Macao, China: IEEE, 2019: 148–154.
- [24] He X N, Liao L Z, Zhang H W, et al. Neural collaborative filtering//Proceedings of the 26th International Conference on World Wide Web. Perth, Australia: International World Wide Web Conferences Steering Committee, 2017: 173–182.
- [25] Jia R Q, Zhou X F, Dong L H, et al. Hypergraph convolutional network for group recommendation//2021 IEEE International Conference on Data Mining. Auckland, New Zealand: IEEE, 2021: 260–269.
- [26] Zhang Z, Gao Y, Li Z L. Consensus reaching for social network group decision making by considering leadership and bounded confidence. Knowledge-Based Systems, 2020(204): 106240.
- [27] Zhang J W, Gao M, Yu J L, et al. Double-scale self-supervised hypergraph learning for group recommendation//Proceedings of the 30th ACM International Conference on Information & Knowledge Management. Queensland, Australia: ACM, 2021: 2557–2567.
- [28] 张晓阳, 秦贵和, 邹密, 等. 基于 LDA 模型的餐厅推荐方法研究. 计算机科学, 2017, 44(7): 180–184, 214. (Zhang X Y, Qin G H, Zou M, et al. Research on recommendation method of restaurant based on LDA model. Computer Science, 2017, 44(7): 180–184, 214.)
- [29] 陈建, 王子磊, 奚宏生. 基于情境感知的广播电视群组发现策略. 计算机工程, 2018, 44(5): 140–145. (Chen J, Wang Z L, Xi H S. TV group detection strategy based on context awareness. Computer Engineering, 2018, 44(5): 140–145.)
- [30] Quan Y, Gao C, Chin-Yew L. COM: A generative model for group recommendation//Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York, NY, USA: Association for Computing Machinery, 2014: 163–172.

(责任编辑 杨可盛)