



数据分析与知识发现
Data Analysis and Knowledge Discovery
ISSN 2096-3467, CN 10-1478/G2

《数据分析与知识发现》网络首发论文

题目：基于多属性评分隐表征学习的群组推荐算法
作者：张纯金，郭盛辉，纪淑娟，杨伟，伊磊
网络首发日期：2020-10-12
引用格式：张纯金，郭盛辉，纪淑娟，杨伟，伊磊. 基于多属性评分隐表征学习的群组推荐算法. 数据分析与知识发现.
<https://kns.cnki.net/kcms/detail/10.1478.g2.20201012.0958.002.html>



网络首发：在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

出版确认：纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。

基于多属性评分隐表征学习的群组推荐算法

张纯金¹, 郭盛辉², 纪淑娟², 杨伟², 伊磊²

¹(山东科技大学网络安全与信息化办公室 青岛 266590)

²(山东省智慧矿山信息技术重点实验室(山东科技大学) 青岛 266590)

摘要:

[目的] 克服个体用户表征学习受个体用户评分稀疏性影响严重的问题, 提高推荐算法的准确率和反应速度。**[方法]** 提出一种基于神经网络的多属性评分隐表征学习方法, 并应用该方法从用户群组和项目两个维度分别学习多属性评分的隐表征, 最后分别通过用户群组偏好匹配和项目吸引力计算实现两个群组推荐。**[结果]** 基于 TripAdvisor 数据集的实验结果表明: (1) 本文两个算法的准确率、时间性能优于典型的多属性推荐算法、群组推荐算法; 准确率略差于个体推荐算法, 但在线和离线运行时间较个性化推荐算法降低 30% 和 50% 以上。(2) 用户群组的隐表征学习相比项目的隐表征学习对推荐性能的提高作用更明显。**[局限]** 由于真实群组数据难以获取, 现有研究普遍基于某种聚类算法生成虚拟群组。本文亦然。这种方法生成的群组较理想化, 因此群组的偏好比真实群组的偏好可能更易聚合。**[结论]** 基于神经网络学习群组用户的隐表征(也即聚合群组用户的偏好)和项目的隐表征, 可以有效提高群组推荐算法和多属性推荐算法的准确率和召回率, 并使之非常接近最新的个性化推荐算法。

关键词: 群组推荐算法; 评分矩阵; 多属性评分; 隐表征学习; 神经网络

分类号: TP393

DOI: 10.11925/infotech.2096-3467.2020.0264

The Group recommendation algorithms based on implicit representation learning of multi-attribute ratings

Zhang Chunjin¹, Guo Shenghui², Ji Shujuan², Yang Wei², Yi Lei²

¹(Network Security and Information Office, Shandong University of Science and Technology, Qingdao 266590, China)

²(Shandong Provincial Key Laboratory of Wisdom Mine Information Technology, Shandong University of Science and Technology, Qingdao 266590, China)

Abstract:

[Objective] To overcome the serious problem that individual user's representation learning is affected by the sparsity of individual users' ratings and improve the accuracy and response speed of recommendation algorithm. **[Methods]** A neural network-based method is proposed for learning implicit representation from multi-attribute ratings. And then, this method is used to learn the implicit representation of the multi-attribute ratings from two dimensions such as users group and items, respectively. Finally, two group-oriented recommendation algorithms are realized by matching the learned representation of each group and its preference or calculating the attraction of the learned representation of each item to the groups. **[Results]** Experimental results based on TripAdvisor data set show that : (1) the accuracy and time performances of the two algorithms

given in this paper are better than the typical multi-attribute ones and group ones. Moreover, the accuracies of these two algorithms are slightly worse than the personalized one, but the online and offline running time is reduced by more than 30% and 50% respectively compared to the personalized recommendation algorithm. (2) The recommendation performances about implicit representation learning of users groups slightly outperform the ones about implicit representation learning of items. **[Limitations]** Due to the difficulty in obtaining real groups data, existing researches generally generate virtual groups based on some kind of clustering algorithm. The same goes for this paper. However, the groups generated by this method is idealized, so group preferences may be more easily aggregated than real group preferences. **[Conclusions]** Using neural network to learn group users' implicit representation (that is, aggregation of group users' preferences) and items' implicit representation can effectively improve the recommendation accuracy and the recall by making these two metrics close to the recent personalized ones. At the same time, the online running time is more than 40% lower than the personalized ones.

Keywords: Group recommendation algorithms; Rating matrix; Multi-attribute ratings; Implicit representation learning; Neural network

1 引言

推荐系统已广泛应用于社交平台、电商、新闻、旅游、生活服务等网站。随着用户和项目数的指数级增加,用户-项目评分矩阵已成为高维稀疏矩阵。基于高维稀疏矩阵的个性化协同推荐算法设计面临着数据非常稀疏、冷启动、单个用户画像困难、推荐准确率低、处理时间长、扩展性差等一系列问题。基于群组的推荐算法可以较好地解决这些问题^[1,2]。

现有基于群组的推荐算法研究主要集中于两个核心问题,即如何生成群组,如何生成群组推荐列表。群组生成主要有三种方式:基于相似度的生成方式^[3-13]、基于内容的生成方式^[14]以及基于规则的生成方式^[15-18]。群组推荐列表生成方法可以划分为两大类:推荐列表融合法,群体偏好聚合法。前者首先基于群组中的每个个体生成推荐列表,然后融合这些推荐列表得到群组的推荐列表^[19]。后者通过聚合群组中每个成员的偏好得到群组的整体偏好,再做推荐。聚合群体偏好的方法主要有:经典策略(如平均满意度策略^[20,21],最小痛苦策略^[22,23],最大幸福策略^[24-26]),改进的经典策略^[19,27,28,29],基于社会关系和交互的策略^[16,30-32],基于谈判机制的策略^[31,33],基于公平的策略^[34,35]。此外,基于融合注意力和神经协同过滤网络的思想,Cao 等人^[36]提出了一种基于注意力的群组偏好聚合策略。他们利用注意力机制调整群组表示,用神经协同过滤网络学习群组与项目之间的关系。同时,他们还将个体-项目之间的交互模型集成到聚合策略中。然而,现有群组偏好聚合方法大多是在单个评分基础上设计的,基于多属性(也称为多侧面)评分的聚合策略较少^[37]。这是因为多属性评分使得群组偏好聚合算法的数据维度和计算复杂性大幅提高。因此,多属性评分背景下群体偏好聚合算法是面向群组推荐算法研究中需解决的关键问题之一。

鉴于此,本文提出一种基于神经网络的多属性评分隐表征学习方法(NNIIIL模型,详见第3节)。然后,应用该方法学习群组用户的多属性评分隐表征进而聚合群组偏好,并通过群组偏好匹配方法实现面向群组的项目推荐(GMURec算法,详见第4.1节)。为了进一步验证该多属性评分隐表征学习方法的可用性并比较、分析 GMURec 算法的性能,本文应用该方法学习项目多属性评分隐表

征，然后通过项目吸引力计算实现面向群组的推荐（GMIRec 算法，详见第 4.2 节）。同时，设计了两套实验实现 NNIL 模型中最优网络层数和最优的神经元数目求解和 GMURec 算法、GMIRec 算法的推荐性能分析。本文研究与实验框架结构如图 1 所示。

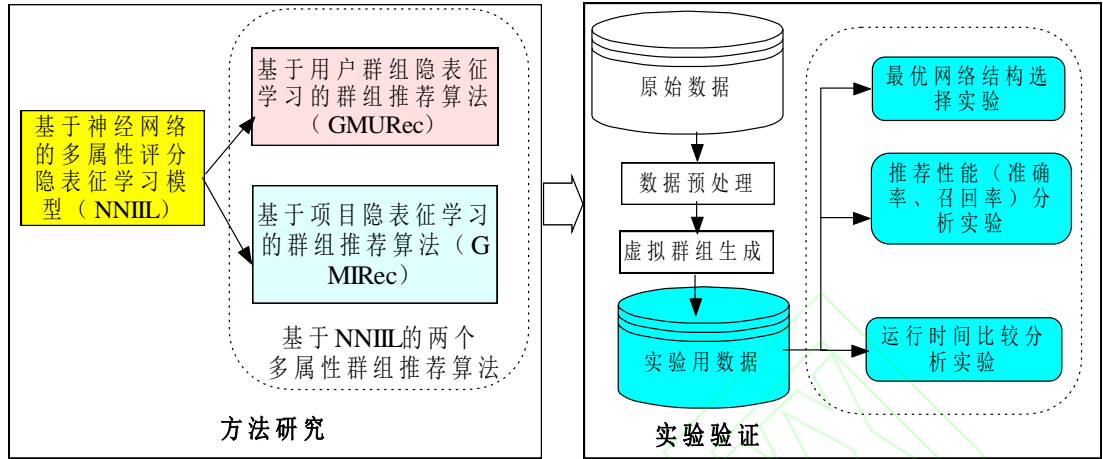


图 1 论文研究与实验框架结构图

Fig. 1. Schematic diagram of the framework

与已有相关研究相比，本文工作的创新点如下：（1）本文面向兴趣相似的虚拟群组实现推荐。与面向个体用户的推荐算法相比，服务对象的粒度更大，且较好地克服了个体用户表征学习受数据稀疏性影响严重的问题。（2）给出了一种基于简单神经网络的多属性评分隐表征学习模型（NNIL），该模型相较于当下流行的深度神经网络^[36]具有简单、调参容易的优点。（3）应用 NNIL 模型对群组用户表征和项目表征分别进行学习，面向群组用户进行推荐。基于 TripAdvisor 真实数据的实验结果表明：本文算法性能优于典型的多属性推荐算法和群组推荐算法；基于用户群组表征的学习比基于项目表征的学习更能提高推荐性能；不论哪种学习方法，推荐性能都非常接近最新的个性化推荐算法，且运行时间降低超过 40%。

论文其余部分组织如下：第 2 节对本文相关工作进行简要综述。第 3 节给出一种基于神经网络的评分隐性信息学习模型（NNIL）；第 4 节给出应用 NNIL 模型实现的两种多属性群组推荐算法（GMURec 和 GMIRec）；第 5 节通过基于真实数据的实验对本文给出的两种推荐算法的性能进行分析。

2 相关工作

本节简要介绍与本文相关的多属性推荐和群组推荐的相关工作。

2.1 基于多属性的推荐算法研究现状

目前，很多电子商务平台（如携程网、美团网、Trip Advisor 等）通过多属性打分模式获得用户的多属性评分。相对于单属性评分，多属性评分中包含着用户对项目各个属性的评分和总评分，可以更精确地反映用户的偏好。

为了充分利用这些数据，学者们做了很多相关工作。例如 Bilge 等人^[38]将多属性评分应用到协同过滤算法中，提出了一种利用多属性评分发现偏好相似的邻居并进行推荐的算法。Adomavicius 等人^[39]提出了一种基于多属性评分的相似度计算方法，该算法利用多属性评分来寻找相似用户并预测用户对每个属性的评分。

根据用户的预测评分,对用户进行推荐。Nilashi 等人^[40]利用决策树和 EM 算法提高多属性协同过滤推荐系统的准确率。Majumder 等人^[41]利用矩阵分解的方法对多属性评分进行处理,结果表明该算法可以提高推荐的准确率。陈冬林等人^[42]将项目的多属性评分引入协同过滤算法中得到一种多属性推荐算法。该算法充分利用用户的浏览内容、浏览时间和项目的属性等影响因素,对各个属性的评分进行预测。Zhang 等人^[43]利用项目的多属性特征来对用户偏好进行建模,并将其与协同过滤算法相结合进行推荐。

上述算法仅利用多属性评分得到用户对各个属性的偏好,却忽略了其中蕴含的隐性信息。为了学习多属性评分中蕴含的隐性信息,Jannach 等人^[44]利用支持向量回归挖掘用户对各个属性评分和总评分的关系,从而学习用户更看重的属性并进行推荐。Jannach 等人^[45]通过挖掘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。Zheng 等人^[46]提出一种标准链(Criteria Chains)算法,该算法不仅学习用户对各属性评分和总评分之间的关系,还学习各个属性之间关系。李岱峰等人^[47]提出一种面向群组的多属性推荐模型,该模型利用模拟退火的方法学习各成员在各属性上的偏好权重,再通过偏好聚合策略聚合群组内所有成员的偏好。以上算法证明,通过学习单个用户的多属性评分和总评分中蕴含的隐性信息可以提高推荐的准确率。

2.2 群组偏好聚合策略的研究现状

经典的群组偏好聚合策略有平均满意度策略^[20,21],最小痛苦策略^[22,23],最大幸福策略^[24-26]等。为了提高群组成员偏好聚合策略的性能,一些学者通过改进经典的偏好聚合策略对群组成员偏好进行聚合。例如,刘荣荣等人^[27]在经典偏好聚合策略基础上引入时间因素,在产生群体推荐列表之前先用时间衰减函数对数据评分进行预处理,从而提高推荐质量。Ma, Ji 等人^[19]将平均聚合策略和 Borda 策略相结合提出了一种名为 HAaB 的混合群组偏好聚合策略。郭均鹏等人^[28]综合了平均满意度策略^[20,21]和最小痛苦策略^[22,23]的优势给出一种新策略,并引入差异度对该模型进行调优,最后分析了差异度对群组推荐算法的影响。融合最小痛苦策略和投票聚合思想,Roy 等人^[29]给出了一种群组偏好聚合和生成方法。

与经典聚合策略不同,Quijano-Sanchez 等人^[16]综合考虑多个社会信任关系的影响因素(共同的好友数目、相识时间、相互交流频繁度、相互共享的信息数目等),提出一种策略对群组成员偏好进行聚合。梁昌勇等人^[31]根据群组中所有成员之间的相互作用确定整个用户群组的偏好,从而对用户群组进行推荐。赵梦楠等人^[30]首先用中心度作为分析指标计算每个成员在群组中的重要度,然后根据重要度对各个成员的偏好进行聚合得到群组对项目的预测偏好,最后根据预测偏好对用户群组进行推荐。Shang 等人^[32]利用社会影响理论对群组中所有成员的偏好进行聚合,并基于该偏好对群体进行推荐。此外,Garcia 等人^[48]基于群体所有成员的偏好、人口的统计信息及其历史行为信息,利用投票机制与谈判机制聚合群组中所有成员的偏好从而为群体用户推荐旅游活动。李汶华等人^[33]利用案例推理与协商方法对成员的偏好进行聚合并基于得到的偏好对目标群体推荐。此外,基于公平的群组偏好聚合策略^[34,35]近两年也一直是群组推荐算法的一个热点问题。近几年,基于神经网络的群组推荐已成为新的热点。特别值得注意的是,Cao 等人^[36]提出了一种基于注意力的群组偏好聚合策略。该策略融合了注意力机制和深度神经协同过滤网络的思想,利用注意力机制调整群组表示,用神经协同过滤网络学习群组与项目之间的关系。同时,他们还将个体-项目之间的交互模型集成

到聚合策略中。

以上群组聚合方法全部是基于总评分的，没有考虑用户在多个属性（如宾馆的洁净程度、位置、安全性等）上的偏好。Zhu 等人^[37]提出一种面向群组的推荐算法 PromoRec。该算法属于一个典型的多属性群组推荐算法，可以在稀疏数据中获取群体的偏好，从而提高推荐性能。基于多属性评分的群组推荐算法虽然可以较好地提高推荐性能，但是多个属性使得群组偏好聚合方法的维度更大，计算复杂性更高，已成为面向大数据推荐算法设计的难点问题之一。

3 基于神经网络的多属性评分隐表征学习模型

3.1 多属性评分定义

在多属性评分电子商务系统中，用户 u 对项目 i 的评价可以表示为如下的矩阵形式。

定义 1: (评分矩阵) \vec{r}_{ui} 表示用户 u 对项目 i 的评分，用户的集合和项目的集合分别记为 $User = \{u_1, u_2, \dots, u_m\}$ 和 $Item = \{i_1, i_2, \dots, i_n\}$ 。用户对项目的评分矩阵 $R_{m \times n} = [\vec{r}_{ui}]_{m \times n}$ 可以表示为公式 (1) 的形式。

$$R_{m \times n} = \begin{pmatrix} \vec{r}_{11} & \cdots & \vec{r}_{1n} \\ \vdots & \ddots & \vdots \\ \vec{r}_{m1} & \cdots & \vec{r}_{mn} \end{pmatrix} \quad (1)$$

其中 \vec{r}_{ui} 是用户 u 给项目 i 的评分向量，它可以表示为 $\vec{r}_{ui} = [a_1, a_2, \dots, a_k, a_{num}]$ ，其中 a_k 是用户 u 对项目 i 第 k 个属性的评分， a_{num} （也记做 tr_{ui} ）是用户 u 对项目 i 的总评分。

由于一个用户可能多次给某个项目评分，因此对用户多次评分的聚合方法定义如下。

定义 2: (用户多次评分聚合) 用户 u 对项目 i 所有评价的聚合可定义为一个新向量，记为 $\vec{r}_{ui} = [\bar{a}_{ui}^1, \bar{a}_{ui}^2, \dots, \bar{a}_{ui}^{num}]$ ，其中 \bar{a}_{ui}^k 是用户 u 对项目 i 第 k 个属性多次评价的平均值，可根据公式 (2) 计算得到， a_{ui}^k 是用户 u 对项目 i 第 k 个属性的评价， h 表示用户 u 对项目 i 的历史评价总次数。

$$\bar{a}_{ui}^k = \frac{1}{h} \sum_{l=1}^h a_{ui}^{k,l} \quad (2)$$

3.2 基于神经网络的多属性评分隐表征学习模型

由于评分矩阵中存在用户、项目和多个属性的评分值、总评分值多个维度的信息，为了深度挖掘评分矩阵中的隐表征，本文给出一个基于神经网络的多属性评分隐表征学习模型 (NNIIL)，其结构如图 2 所示。

该模型为一个前向神经网络，一个输入层，一或多个隐含层，一个输出层。模型的执行步骤为：(1) 将用户对各个属性的评分数据（定义 1 给出的 \vec{r}_{ui} 中的前 $num-1$ 项，即 $[a_1, a_2, \dots, a_{num-1}]$ ）通过输入层载入到 NNIIL 网络结构中；(2) 利用初始化的权重矩阵和偏置矩阵将数据正向传播，得出预测的总评分；(3) 以真实的总评分（定义 1 给出的 \vec{r}_{ui} 中的第 num 项，即 $[a_{num}]$ ）为标准，利用梯度下降算法^[49]不断对权重矩阵和偏置矩阵进行调整直至误差阈值或迭代次数达到可接受范围为止。该模型通过监督学习方法得到多属性评分之间、多属性评分与总评分之间的隐性关系。

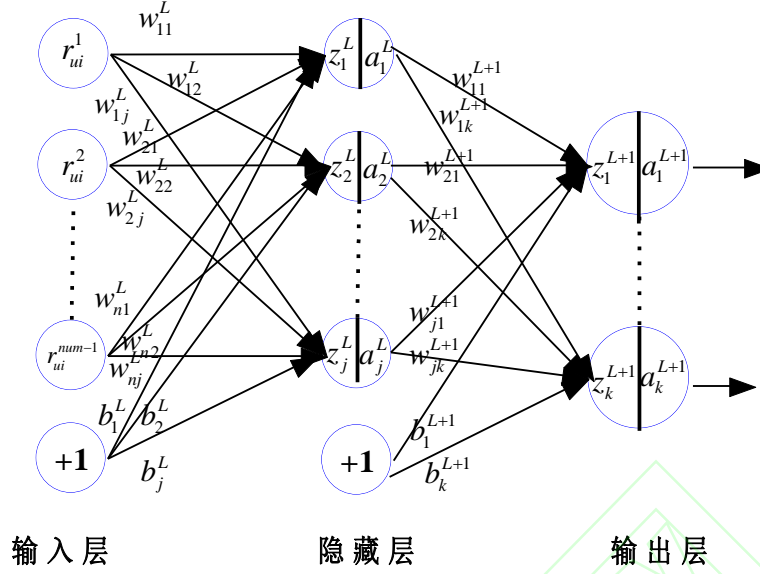


图 2 NNIL 模型

Fig. 2. NNIL model

4 基于隐表征学习的多属性群组推荐算法

在第 3 节给出的隐表征学习模型（NNIL）的基础上，本节给出一个基于用户群组隐表征学习的多属性群组推荐算法（GMURec，详见 4.1）和一个基于项目隐表征学习的多属性群组推荐算法（GMIRec，详见 4.2）。在此之前，由于本文的推荐对象为群组，本节先给出群组定义如定义 3 所示。需注意，本文所涉及群组不是家庭等真实群组，而是兴趣相似的虚拟群组。

定义 3: (用户群组) 一个用户群组 $g_c = [u_1, u_2, \dots, u_p]$ (p 是第 c 组的用户数目) 是一组具有相同兴趣的用户，他们对项目的各属性的评分和总评分的值都类似。多个用户群组构成一个群组集合，记为 $G = [g_1, g_2, \dots, g_s]$ 。

4.1 基于用户群组隐表征学习的群组推荐算法（GMURec）

图 3 给出了 GMURec 算法的框架图。该框架由四大模块组成，即用户群组数据模块、项目数据模块、训练模块和推荐模块。下面详细介绍后两个模块的实现细节。

(1) 基于 NNIL 模型的群组用户特征学习

基于 NNIL 模型学习用户群组特征时，首先以每组用户（记做 g_c ）对各个属性的评分（记做 $r_{g_c,i}$ ）作为输入，把与之对应的总体评分（记做 $r_{g_c,i}^{tr}$ ）作为标准输出，进行有监督的训练。NNIL 模型可以抽象表示为公式(3)。为了简化计算，本文选择 sigmoid 函数 ($f(x) = \frac{1}{1+e^{-x}}$) 作为激活函数。

$$\hat{r}_{g_c,i}^{tr} = f(w_{g_c}^T \cdot r_{g_c,i} + b_{g_c}) \quad (3)$$

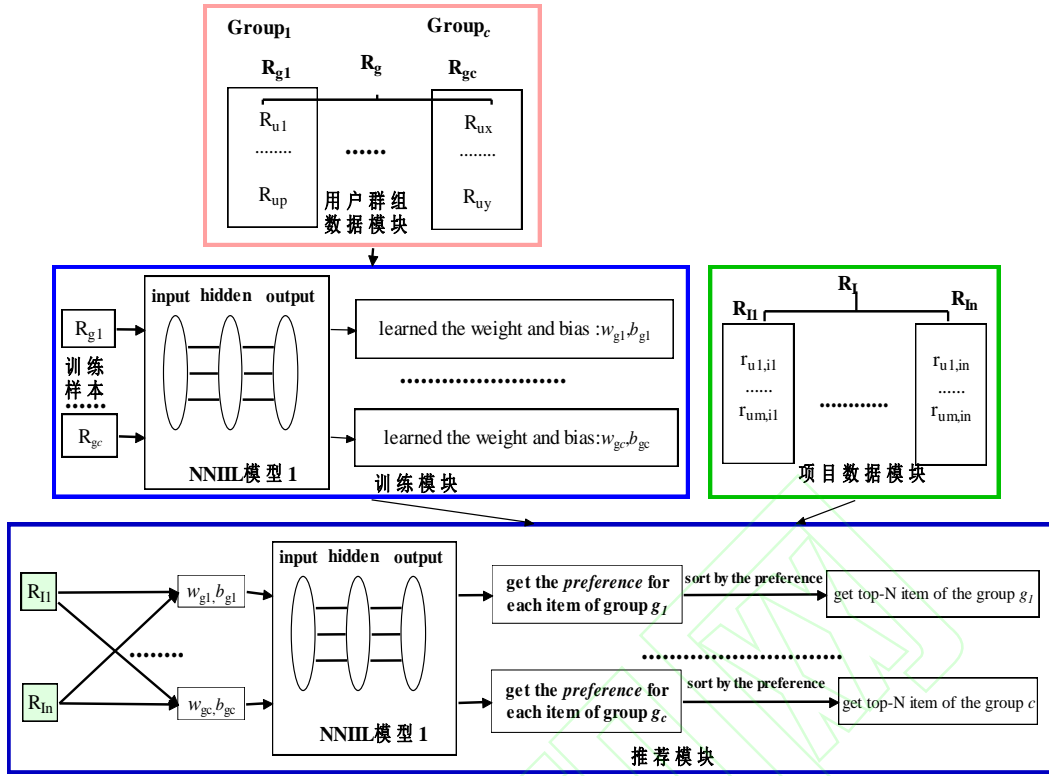


图3 GMURec 算法框架图

Fig. 3. GMURec algorithm framework diagram

在训练过程中，采用梯度下降算法^[49]来优化公式(3)中所使用的权重 ($w_{g_c}^T$ 和 b_{g_c})。具体优化过程^[49]为：首先，利用公式(4)定义每个用户群组的损失函数；然后，利用公式(5)和公式(6)计算该群组在每个参数上的梯度；最后，利用公式(7)和公式(8)对该用户群组在每个属性上的偏好权重和偏置矩阵进行迭代更新。

$$error(g_c) = \sum_{i=1}^{|I|} (\hat{r}_{g_c,i}^{tr} - r_{g_c,i}^{tr})^2 \quad (4)$$

$$\nabla_w error^j(g_c) = -\frac{1}{|I|} \sum_{i=1}^{|I|} (\hat{r}_{g_c,i}^{tr} - r_{g_c,i}^{tr}) \cdot r_{g_c,i}^j \quad (5)$$

$$\nabla_b error^j(g_c) = -\frac{1}{|I|} \sum_{i=1}^{|I|} (\hat{r}_{g_c,i}^{tr} - r_{g_c,i}^{tr}) \quad (6)$$

$$w_{g_c}^j \leftarrow w_{g_c}^j + \alpha \nabla_w error^j(g_c) \quad (7)$$

$$b_{g_c} \leftarrow b_{g_c} + \alpha \nabla_b error^j(g_c) \quad (8)$$

其中， $\hat{r}_{g_c,i}^{tr}$ 是根据 NNIL 模型训练得到的用户群组 g_c 对项目 i 预测总评分， w_{g_c} 和 b_{g_c} 分别是根据用户群组 g_c 的数据训练出的 NNIL 模型中的权重和偏置矩阵， $r_{g_c,i}$ 是用户群组 g_c 对各个属性的评分向量， $f(x)$ 是 sigmoid 函数， $error(g_c)$ 是用户群组 g_c 的评分预测误差， $r_{g_c,i}^{tr}$ 是用户群组 g_c 对项目 i 真实总评分， $\nabla_w error^j(g_c)$ 是群组 g_c 在属性 j 偏好权重 w_{g_c} 上的梯度， $\nabla_b error^j(g_c)$ 是群组 g_c 在属性 j 偏好偏置 b_{g_c} 上的梯度， I 是项目集合， α 是梯度下降算法中的学习率。

(2) 偏好匹配与推荐

基于用户群组的评分数据训练得到的 NNIL 模型中蕴含着每个群组用户对

各属性的偏好、多个属性的评分与总评分之间的隐性关系等信息。为了充分利用这些隐性的表征信息，本文提出直接基于训练好的 NNIL 模型选择适合的项目进行推荐。为了达到这个目的，需要知道两类信息：(a) 用户群组对每个项目的喜欢程度是多少？(b) 哪些项目是用户群组比较喜欢的？第一个问题最重要。如果第一个问题得以解决，通过对偏好程度进行排序并选择 Top-N 个项目进行推荐即可。为了计算每个用户群组对每个项目的喜欢程度，本节定义了群组偏好程度 (*preference*) 的概念如下。

定义 4: $preference(g_c, i)$ 代表用户群组 g_c 对项目 i 的偏好程度，它的计算方法如公式(9)和公式(10)所示。

$$preference(g_c, i) = \frac{\left| \bigcup_{g_c \in GI_{i-p}} \{r_{g_c i}^{tr} \mid |\hat{r}_{g_c i}^{tr} - r_{g_c i}^{tr}| \leq 1\} \right|}{\left| \bigcup_{g_c \in GI_i} \{r_{g_c i}^{tr} \mid |\hat{r}_{g_c i}^{tr} - r_{g_c i}^{tr}| \leq 1\} \right| + \sigma} \quad (9)$$

其中， $\hat{r}_{g_c i}^{tr}$ 是基于用户群组 g_c 中数据训练出的 NNIL 模型预测用户群组 g_c 对项目 i 的总评分，其计算方法已在公式 (3) 中给出； $r_{g_c i}^{tr}$ 是与 $\hat{r}_{g_c i}^{tr}$ 对应的真实总评分， GI_i 为对项目 i 评分的用户群组集合， GI_{i-p} 为对项目 i 的积极评分的用户群组集合； w_{g_c} 和 b_{g_c} 分别是根据用户群组 g_c 的数据训练出的 NNIL 模型中的权重和偏置矩阵， σ 为调整因子是为了避免出现分母为 0 的情况而设置。

GMURec 算法实现的详细步骤如下（见算法 1）：(1) 利用 NNIL 模型对每组用户的评分数据进行隐表征学习（如图 3 中的训练模块），即算法 1 中的第 1) 步。(2) 由于基于用户群组数据训练得到的 NNIL 模型中蕴含着用户对各个属性的偏好程度，利用公式 (9) 定义的 *preference* 计算用户群组 g_c 对每个项目的偏好程度，即算法 1 中的第 2)-4) 步。(3) 基于用户群组 g_c 对每个项目的偏好程度，实现面向群组用户的推荐（算法 1 中的第 5) 步），即图 3 中的推荐模块。

算法 1: GMURec

Input: The data set of the g_c and the data set of each item

Output: The recommendation list for g_c

- 1) Learn the NNIL model (w_{g_c} and b_{g_c}) with the dataset of g_c ;
 - 2) for i in I do
 - 3) Compute the $preference(g_c, i)$ according to (9);
 - 4) end for
 - 5) Select Top-N items as the recommendation list to g_c
-

4.2 基于项目隐表征学习的群组推荐算法 (GMIRec)

为进一步比较、分析 GMURec 算法利用 NNIL 模型聚合群组偏好的方法对推荐性能的提高程度，本节提出基于项目数据应用 NNIL 模型实现项目表征学习，并通过项目吸引力计算实现基于项目隐表征学习的群组推荐算法 (GMIRec)。算法的实现框架如图 4 所示，其包含两个关键步骤：(1) 基于 NNIL 模型的项目特征学习；(2) 项目对用户群组的吸引力计算及基于吸引力的推荐。基于 NNIL 模型的项目特征学习与 4.1 节中基于 NNIL 模型学习群组用户的特征类似。NNIL 模型也可以类似地抽象表示为公式(10)的形式，且仍然选择 sigmoid 函数作为激活函数。

$$\hat{r}_{g,i}^{tr} = f(w_i^T \cdot r_{g,i} + b_i) \quad (10)$$

与 GMIRec 算法不同的是，模型输入和输出不是每个用户群组的评分数据（图 3 中的红色方框）而是每个项目的所有评分数据（图 4 中的绿色方框）。由于篇幅限制，不再赘述基于 NNIL 模型的项目特征学习过程。

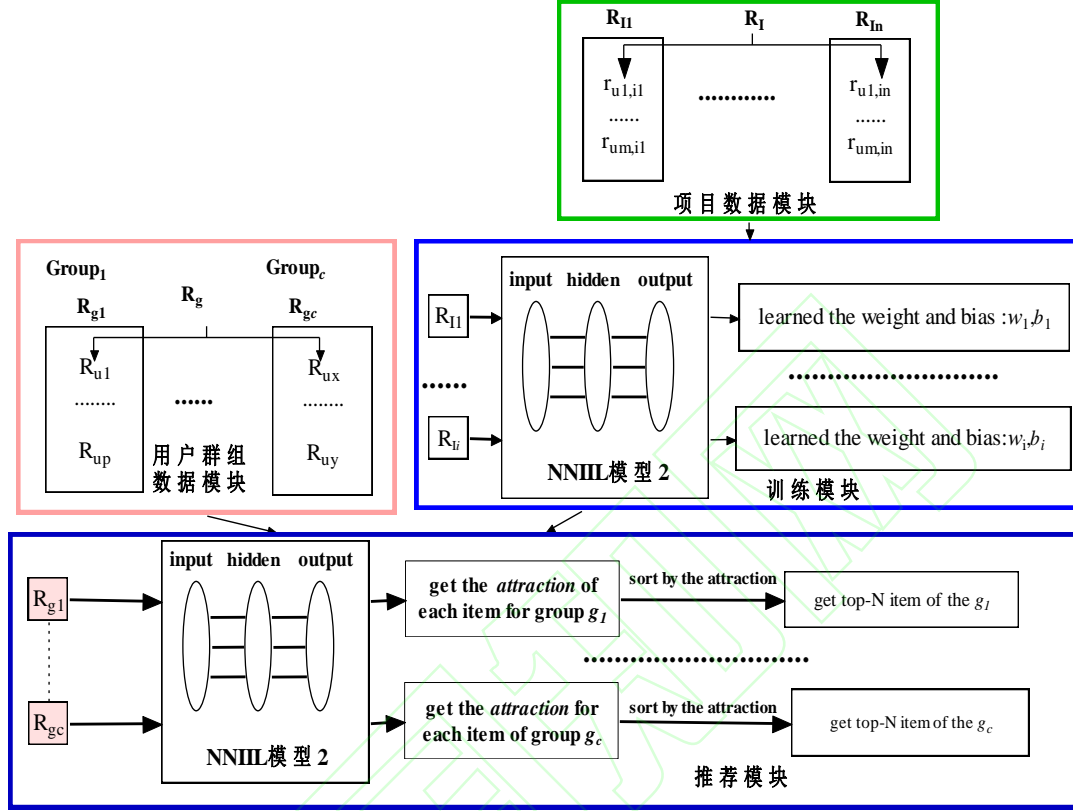


图 4 GMIRec 算法框架图

Fig. 4. GMIRec algorithm framework diagram

基于每个项目的评分数据训练得到的 NNIL 模型中蕴含着该项目的优点、缺点等隐特征，可利用这些特征选择适合的群组进行推荐。为实现该目标，需要得到 2 类信息：(a) 目标项目对每个用户群组的吸引力是多少？(b) 该项目吸引哪些用户群组？第一个问题更重要。如果第一个问题得以解决，通过对吸引程度进行排序并选择 Top-N 个项目即可实现推荐。为了计算项目对每个用户群组的吸引力，本节定义吸引力 (*attraction*) 的概念，如下所示。

定义 5: $attraction(i, g_c)$ 代表项目 i 对群组 g_c 的吸引力，如公式(11)。

$$attraction(i, g_c) = \frac{\left| \bigcup_{i \in I_{g_c-p}} \{r_{g,i}^{tr} \mid |\hat{r}_{g,i}^{tr} - r_{g,i}^{tr}| \leq 1\} \right|}{\left| \bigcup_{i \in I_{g_c}} \{r_{g,i}^{tr} \mid |\hat{r}_{g,i}^{tr} - r_{g,i}^{tr}| \leq 1\} \right| + \sigma} \quad (11)$$

其中， $\hat{r}_{g,i}^{tr}$ 是基于项目 i 中数据训练出的 NNIL 模型预测用户群组 g_c 对项目 i 的总评分，其值根据公式 (10) 学习得到； $r_{g,i}^{tr}$ 是与 $\hat{r}_{g,i}^{tr}$ 对应的真实总评分， I_{g_c} 为用户群组 g_c 进行评分的项目集， I_{g_c-p} 为用户群组 g_c 进行积极评分的项目集； w_i 和 b_i 分别是根据用项目 i 的数据训练出的 NNIL 模型中的权重和偏置矩阵； σ 为调整因子，旨在避免出现分母为 0 的情况。

算法 2: GMIRec

Input: The data set of I and the data set of g_c **Output:** The recommendation list for g_c

- 1) for i in I do:
 - 2) Learn the NNIL model (w_i and b_i) with the dataset of i ;
 - 3) Compute the *attraction* (i, g_c) according to (11);
 - 4) end for
 - 5) Select Top-N items as the recommendation list to g_c
-

GMIRec 算法的详细步骤如下 (算法 2): (1) 利用 NNIL 模型对每个项目的评分数据进行隐表征学习 (如图 4 中的训练模块), 即算法 2 中的第 2) 步; (2) 由于训练得到的 NNIL 模型中蕴含着项目的隐表征, 利用公式(11)定义的 *attraction*, 计算每个项目对用户群组 g_c 的吸引力, 即算法 2 中的第 3) 步; (3) 基于每个项目对用户群组 g_c 的吸引力进行推荐 (算法 2 中的第 5) 步), 即图 4 中的推荐模块。

5 实验与结果分析

为了验证 GMURec 和 GMIRec 算法的性能, 本文设计并实现了两套实验。第一套实验采用插值法得到 NNIL 模型中最优网络层数和最优的神经元数目。第二套实验以 Precision、Recall 和 Time 为评估标准比较、分析了本文算法和一些基线算法的性能。在具体实验之前, 5.1 节介绍本文使用的数据集, 5.2 节介绍了虚拟群组生成方法, 5.3 节介绍实验设置和评估标准, 5.4 节对实验结果进行详细分析。

5.1 数据预处理

本文数据集是从 Tripadvisor 网站上爬取的关于宾馆的数据集 (<https://pan.baidu.com/s/1GdUAFf8NP1dleSRFeVvGWA>)。该数据集含多个属性评分, 例如价格 (Price)、交通 (Traffic)、洁净 (Clean)、位置 (Location)、服务 (Service)、环境 (Surrounding) 和总评 (Overall)。原始数据集十分稀疏, 一共有 796958 条数据, 536952 个用户和 3945 个项目。为了缓解数据稀疏问题, 本文将不活跃的用户和项目 (即评论数据少于 10 条的用户和评论数据少于 100 条的项目) 删除, 得到预处理后数据 (详见表 1)。

表 1 初步数据预处理

Table1 Data preprocessing		
属性	原始数据	预处理后数据
用户	536952	20443
项目	3945	1755
数据	796958	262300
稀疏度	99.96%	99.3%

5.2 虚拟群组生成

面向群组推荐算法面临着一个重要问题——生成虚拟群组。为实现该目的, 本文基于 K 均值聚类算法^[50]设计了一个用户虚拟群组生成方法。具体包括以下步骤:

第一步, 计算用户之间的相似度。首先, 利用 Pearson 相关系数^[51]计算用户

在各个属性上的评分相似度（公式(12)），从而得到用户之间在各个属性上的评分相似度向量，如公式(13)。其次，利用公式(14)计算得到用户之间的直接相似度。用户之间的相似度不仅仅和共评论数据有关，用户共评论数据占有数据的比率也会相似度产生影响。如果两个人共评论数据所占总数据的比率比较大，那么他们的偏好可能很相似。本节用 Jaccard 系数^[51]来评估共评论数据占比带来的影响，如公式(15)。

$$S_{PCC}^j(u, v) = \frac{\sum_{i \in I_u \cap I_v} (r_{ui}^j - \bar{r}_u^j)(r_{vi}^j - \bar{r}_v^j)}{\sqrt{\sum_{i \in I_u \cap I_v} (r_{ui}^j - \bar{r}_u^j)^2} \sqrt{\sum_{i \in I_u \cap I_v} (r_{vi}^j - \bar{r}_v^j)^2}} \quad (12)$$

$$Sim^U(u, v) = [S_{PCC}^1(u, v), \dots, S_{PCC}^j(u, v), \dots, S_{PCC}^{num}(u, v)] \quad (13)$$

$$Sim^{direct}(u, v) = \sqrt{\frac{1}{num} \sum_{j=1}^{num} (S_{PCC}^j(u, v))^2} \quad (14)$$

其中， $\bar{r}_u^j = \frac{1}{n} \sum_{i=1}^n r_{ui}^j$ 是用户 u 对所有项目在属性 j 上评分的平均值， n 是用户 u 评论数据的数目； $S_{PCC}^j(u, v)$ 是用户 u 和用户 v 在属性 j 上评分的皮尔逊相似度， $S_{PCC}^{num}(u, v)$ 是用户 u 和用户 v 在总评分上的皮尔逊相似度， r_{ui}^j 和 r_{vi}^j 分别是用户 u 和用户 v 对项目 i 在属性 j 上的评分， I_u 和 I_v 分别是用户 u 和用户 v 评论的项目集； $Sim^U(u, v)$ 是用户 u 和用户 v 在各属性上的评分相似度向量。

$$Sim^{Jaccard}(u, v) = \frac{|I_u \cap I_v|}{|I_u \cup I_v|} \quad (15)$$

其中， I_u 和 I_v 分别是用户 u 和用户 v 评论的项目集。

根据 Sim^{direct} 和 Jaccard 相似度，可以得到用户之间的最终相似度，如公式(16)。

$$sim(u, v) = Sim^{direct}(u, v) \cdot Sim^{Jaccard}(u, v) \quad (16)$$

第二步，根据计算得到的用户之间的相似度，应用 K 均值聚类算法^[50]发现偏好相似的用户群组。利用公式（17）计算每个群组的轮廓系数^[52]。选择轮廓系数最大的分类群组作为最优群组。本步中，引入轮廓系数旨在克服传统的 K 均值聚类算法必须人为给定用户群组的数目的局限性。

$$S(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \quad (17)$$

其中， $a(i)$ 为 i 向量到同一簇内其他点不相似程度的平均值， $b(i)$ 为 i 向量到其他簇的平均不相似程度的最小值。

图 5 列出了求最优群组时的部分实验结果。由图可知，当群组数为 1120 时轮廓系数最大为 0.826，效果最好。因此，本文将划分得到的数目为 1120 的群组作为后面实验用群组。实验过程中，本文将数据集随机分为 10 份，选择其中的 8 份作为训练集，另外 2 份作为测试集，并利用五折交叉验证进行实验。

5.3 实验设置和评估标准

本文算法中的主要参数设置如表 2 所示。为了衡量 GMURec、GMIRec 算法和对比算法的性能，本文采用推荐算法常用的两个评价指标准确率^[53]（Precision）、召回率^[54]（Recall）和一种评估算法效率的运行时间（Time）作为指标，对算法进行评估。具体定义如下。

定义 6： Precision（准确率）^[53]旨在评估推荐的准确率，即推荐列表中用户喜欢项目的占比，如公式(18)。

$$Precision = \frac{L_u \cap B_u}{L_u} \quad (18)$$

定义 7: Recall（召回率）^[54]旨在评估测试集中有多少用户喜欢的项目被推荐，如公式(19)。

$$Recall = \frac{L_u \cap B_u}{B_u} \quad (19)$$

其中， L_u 是用户 u 的推荐列表， B_u 是测试集中用户喜欢的项目集合， T_{end} 是算法结束的时间， T_{begin} 是算法开始的时间。

定义 8: Time（运行时间）旨在评估算法的运行效率，如公式(20)。

$$Time = T_{end} - T_{begin} \quad (20)$$

其中， L_u 是用户 u 的推荐列表， B_u 是测试集中用户喜欢的项目集合， T_{end} 是算法结束的时间， T_{begin} 是算法开始的时间。

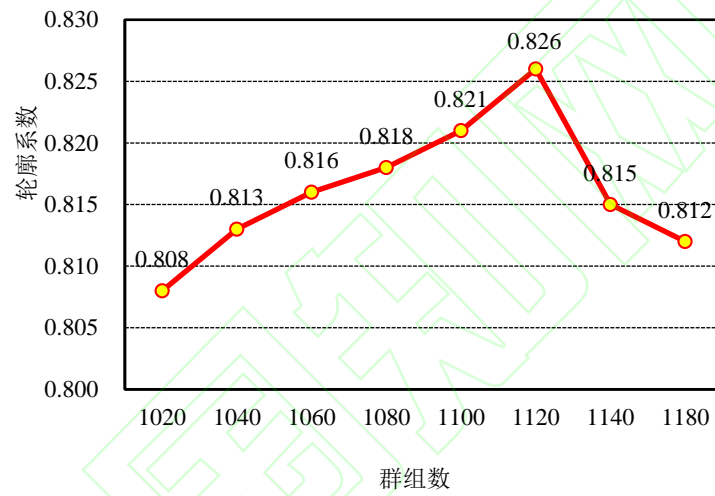


图 5 不同数目群组下的轮廓系数

Fig. 5. Silhouette coefficient in different numbers of groups

表 2 参数设置
Table2 Parameter Settings

参数	值	出现位置
梯度下降算法中的学习率 α	0.15	公式(6)和(7)
梯度下降中误差阈值和迭代次数	0.01 和 100000	梯度下降算法
σ	0.8	公式(9)和(11)

除了以上标准，为了评估神经网络隐特征学习方法的性能，本文将 $marcoF1$ 值^[48]作为指标评估多分类算法（因为 TripAdvisor 数据中的评分是五级分类数据，所以神经网络中的预测问题可看作五分类问题）中的准确率和召回率的宏比例，具体定义如公式(21)所示。

定义 9: $marcoF1$ 值^[55]定义为如下值：

$$marcoF1 = \frac{2 * macroP * macroR}{macroP + macroR} \quad (21)$$

$$macroP = \frac{1}{N} \sum_{i=1}^n \frac{TP_i}{FP_i + TP_i} \quad (22)$$

$$marcoR = \frac{1}{N} \sum_{i=1}^n \frac{TP_i}{FN_i + TP_i} \quad (23)$$

其中, TP_i 是数据中将属于 i 类的数据归为 i 类的数目, FP_i 为数据中将不属于 i 类的数据归为 i 类的数目, FN_i 为数据中将属于 i 类的数据归为 i 类之外其它类的数目, $marcoP$ 是整个算法分类的准确率, $marcoR$ 是整个算法分类的召回率, N 为分类数目。

5.4 实验结果与分析

(1) 最优网络结构选择结果与分析

基于用户群组和项目数据, 利用插值法学习 GMURec 和 GMIRec 算法中的最优 NNILL 模型结构, 得到结果如图 6 和图 7 所示。图中横轴为神经元个数, 纵轴为依据公式 (21) 计算得到的评估标准 $marcoF1$ 。由图 6 可知, GMURec 中最优的 NNILL 模型结构 (即 $macroF1$ 值最大的模型结构) 为 3 层神经网络结构 (1 层隐藏层), 隐藏层中神经元个数为 24。由图 7 可知, GMIRec 中最优的 NNILL 模型结构为四层神经网络结构 (2 层隐藏层), 隐藏层中神经元个数为 27。

基于得到的 GMURec 和 GMIRec 算法中最优的 NNILL 模型结构, 本文可以更加精确地学习用户群组评分数据和项目评分数据中的隐表征, 从而进行更加准确的推荐。因此, 后面的实验结果都是基于最优网络结构得到的推荐结果。

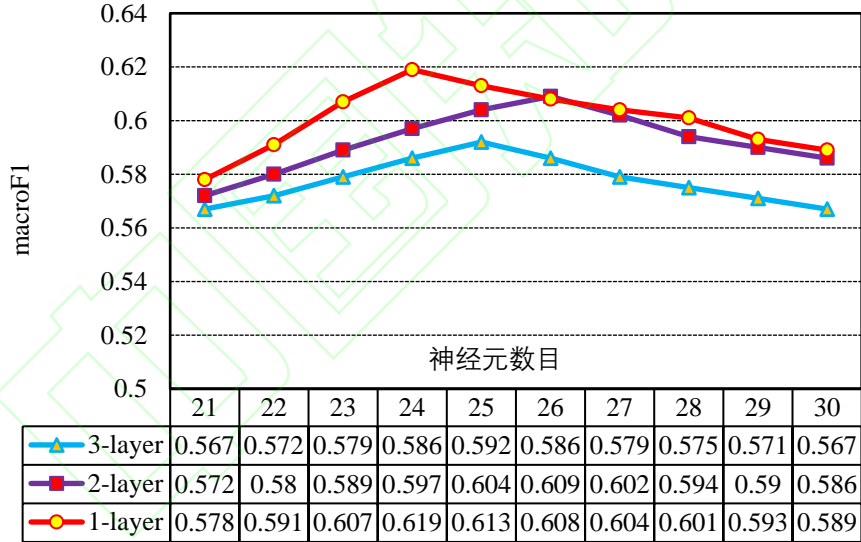


图 6 GMURec 算法中神经网络结构的 $marcoF1$ 值

Fig. 6. $marcoF1$ of the neural network structure in GMURec

(2) 推荐性能分析

为了分析 GMURec、GMIRec 算法的性能, 本文选择五个算法作为对比算法, 如表 3 所示。ATTGRecs^[36]是一个融合了注意力机制和神经协同过滤网络的群组推荐算法。PromoRec 算法^[37]是一个典型的多属性群组推荐算法, 它通过在稀疏数据中探索群体的偏好提高推荐性能。该算法是与本文给出的算法最接近的比较算法。LP 算法^[22]是一个典型的单属性群组推荐算法, 它采用最小痛苦策略对群组偏好进行聚合, 没有考虑用户对各个属性的偏好, 只考虑了用户对项目的整体偏好; PCCCF 算法^[51]是一个典型的基于协同过滤的多属性算法, 该算法采用皮

尔逊相关系数寻找相似邻居；HyCF 算法^[56]是一个最新的面向个体用户的协同过滤多属性推荐算法，该算法采用一种混合的相似度寻找相似邻居。PCCCF^[51]和HyCF^[56]算法需要设定邻居数目，本文将其设置为 10（也是这些算法所在文献的最优值）。

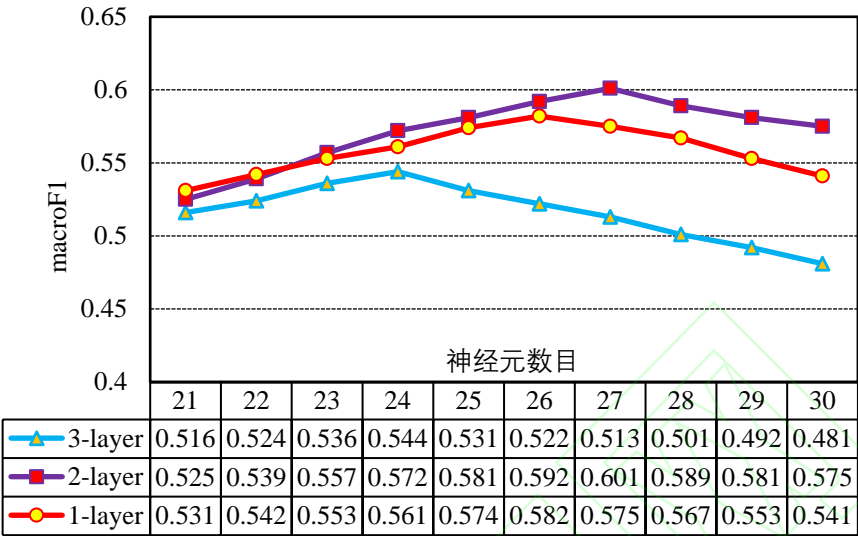


图 7 GMIRec 算法中神经网络结构的 marcoF1 值

Fig. 7. marcoF1 of the neural network structure in GMIRec

表 3 对比算法

Table3 Baselines

算法名称	简称	面向个体用户	面向群体用户	单属性	多属性
基于神经网络的群组推荐算法	ATTGRec ^[36]		√	√	
基于最小痛苦策略的群组推荐算法	LP ^[22]		√	√	
PromoRec 算法	PromoRec ^[37]		√		√
基于皮尔逊相似度的协同过滤算法	PCCCF ^[51]	√			√
基于一种混合相似度的协同过滤算法	HyCF ^[56]	√			√
本文算法	GMURec、GMIRec		√		√

① 准确率的比较与分析

GMURec、GMIRec 算法和对比算法在准确率（Precision）上的对比结果如图 8 所示。由图可知，LP 算法^[22]的准确率随着推荐列表长度的增加而降低，最终趋于稳定；其它算法的准确率都随着推荐列表长度的增加先升高后降低最终趋于稳定。在所有的算法中，无论推荐列表的长度为多少，HyCF 算法^[56]的准确率最大。那是因为 HyCF 算法^[56]是一种多属性的个性化推荐算法，它可以针对每个用户的偏好进行推荐，因此它的准确率最大。本文提出算法（GMURec 和 GMIRec 算法）的准确率比 HyCF 算法小，当推荐列表长度大于 11 时，GMURec 和 GMIRec 算法的准确率非常接近个性化推荐算法 HyCF。那是因为 GMURec 和 GMIRec 算法是群组推荐算法，它们需要对群组偏好进行聚合，所以需要推荐列表长度相对较大时性能才能较好。PCCCF 算法^[51]也是一种多属性个性化推荐算法，但是它没有考虑多属性评分中蕴含的隐性信息，因此它的准确率比 GMURec 和 GMIRec 算法小很多。ATTGRec 算法^[36]与 LP 算法^[22]的准确率较差是因为他们都只考虑了总评分，没有考虑多个属性的评分。同时也说明最小痛苦策略无法有效

聚合群体的偏好。PromoRec 算法^[37]是一种多属性群组推荐算法，它的准确率比 GMURec 和 GMIRec 算法小，原因是 PromoRec 算法^[37]虽然考虑了对项目各个属性的评分，却没有考虑这些评分蕴含的隐性信息。通过比较，还可以发现：在大部分情况下 GMURec 算法的准确率比 GMIRec 算法大。这说明聚合群组用户的偏好要比挖掘项目在各属性上的吸引力更有利于提高推荐的准确率。通过以上结果和分析，可以得出以下结论：

结论 1：通过学习群组用户的隐表征（或称聚合群组用户的偏好）以及学习项目的隐表征，可以有效提高群组推荐算法的准确率，并使之非常接近个性化推荐算法的准确率。

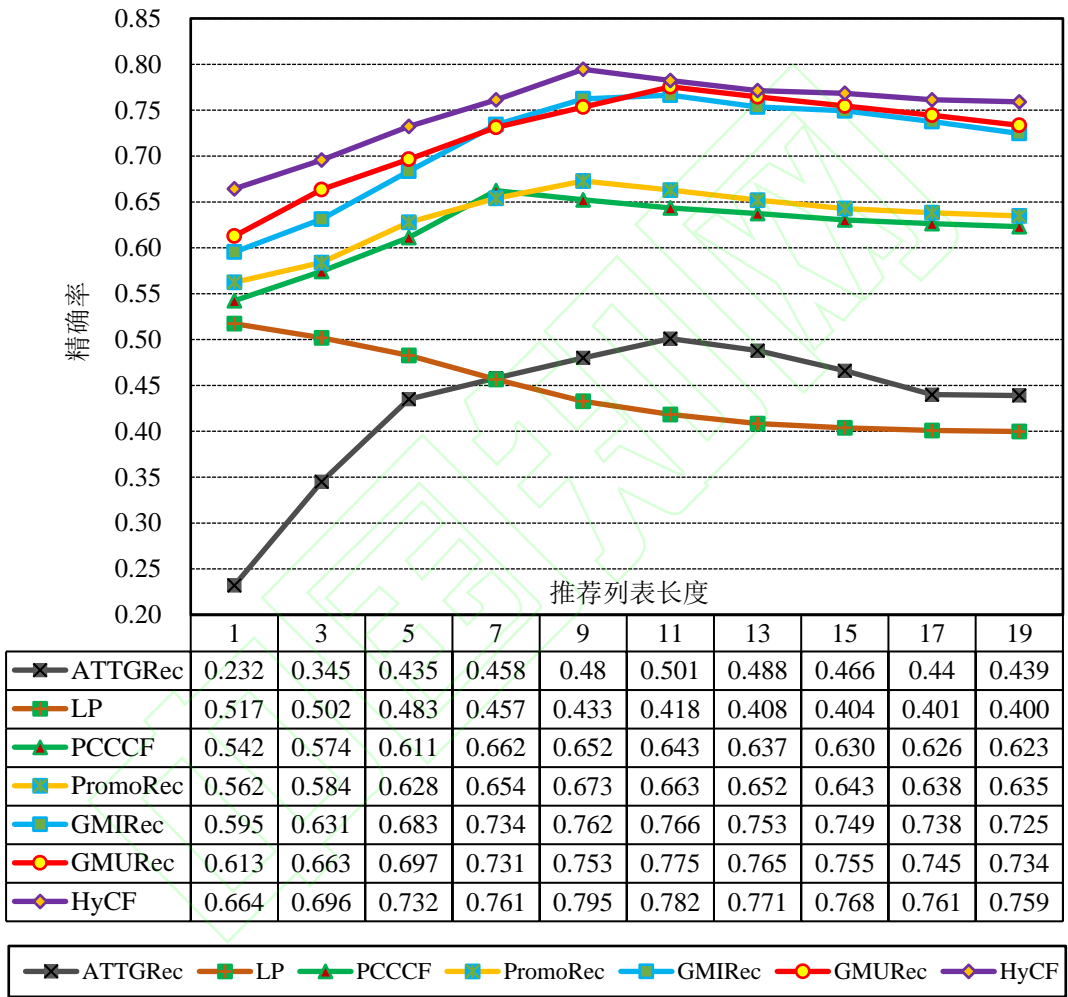


图 8 对比算法的准确率

Fig. 8. Accuracy of the algorithm and baselines

② 召回率的比较与分析

GMURec、GMIRec 算法和对比算法在召回率（Recall）上的对比结果如图 9 所示。由图 9 可知，除 LP 算法^[22]之外其它算法的召回率相差不大。那是因为召回率和测试集中用户喜欢的项目数成反比和推荐列表中用户喜欢的项目数成正比。对所有算法而言，测试集中用户喜欢的项目数是相同的，只是每个算法形成的推荐列表中用户喜欢的项目数不同。相对于测试集中用户喜欢的项目数，每个算法形成的推荐列表中用户喜欢的项目数之间的差异很小。所以，所有算法的召回率相差不大。LP 算法^[22]采用最小痛苦策略聚合群组偏好，它并不能考虑到所

有用户的偏好，所以它的召回率就相对较小。根据以上分析，可以得出结论：

结论 2：除 LP 算法^[22]之外，所有比较算法的召回率相差无几。

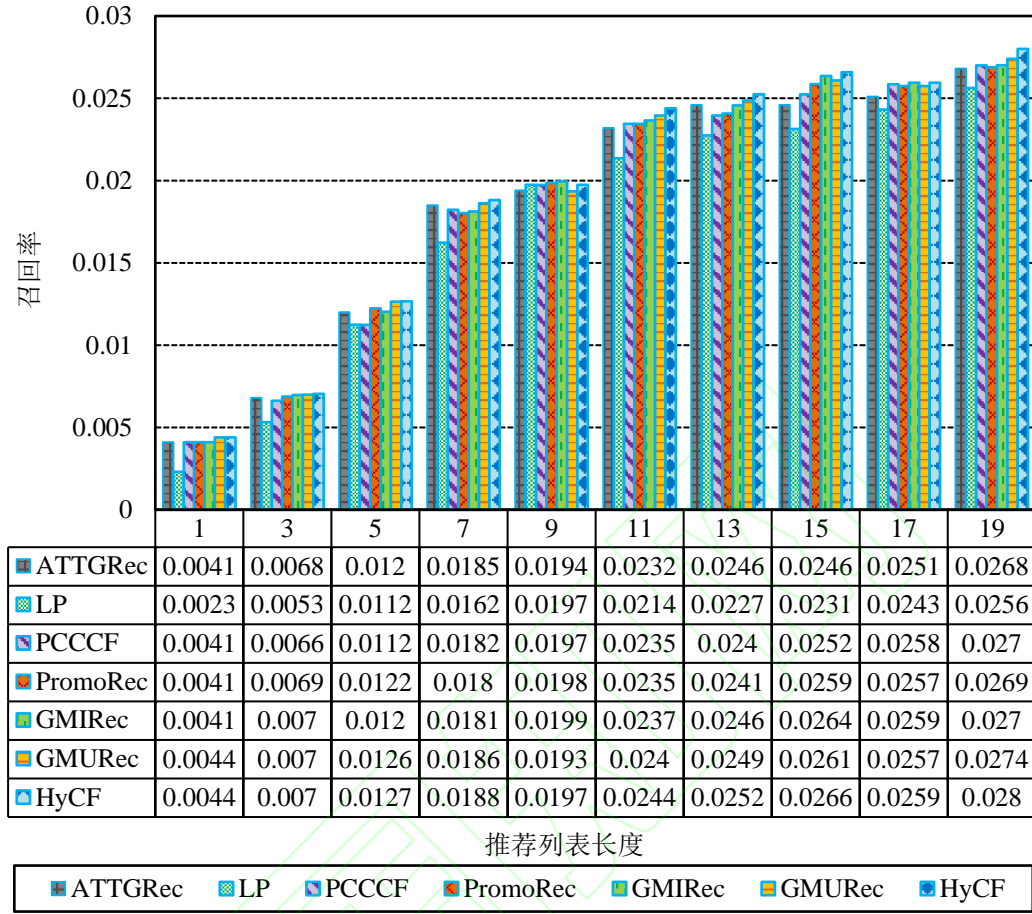


图 9 对比算法的召回率
Fig. 9. Recall of the algorithm and baselines

(3) 运行时间的比较与分析

本节从离线和在线两个阶段分析了 GMURec、GMIRec 算法和对比算法在运行时间上的差异。在线阶段的任务是对测试集中的用户进行实时推荐。除此之外的其它任务都是在离线阶段完成。由于各个算法的绝对时间相差较大，所以本节将所有算法中耗时最长的算法的运行时间设为 1，并将其它算法的运行时间归一化得到所有算法的相对运行时间（记做 Time ratio）作为图 10 的纵轴。由图 10 可知 HyCF 算法^[56]的离线时间最长，那是因为它需要计算四种用户相似度并将其进行融合。在该算法中，仅仅用户相似度耗时就超过了 100 个小时。ATTGRec^[36]、GMURec 和 GMIRec 算法的离线时间相差不大，并且都要比 HyCF 算法^[56]的离线时间少，那是因为它们都需要花费大量的时间训练神经网络。PCCCF 算法^[51]的离线时间最少，那是因为在离线阶段它只需要计算用户之间的皮尔逊相似度。LP^[22]和 PromoRec 算法^[37]是群组推荐算法，它们需要花费时间发现用户群组，并聚合他们的偏好。在线运行时间方面，从图 10 可知 ATTGRec^[36]、LP^[22]、GMURec、GMIRec 和 PromoRec 算法^[37]的在线时间相差不多，那是因为这四种算法都是群组推荐算法，不需要对所有用户形成推荐列表。而且，这四种群组推荐算法仅仅在群组偏好聚合的策略不同。PCCCF^[51]和 HyCF 算法^[56]消耗的在线时间最多，那

是因为这两种算法需要对所有用户形成推荐列表。根据以上分析,可以得出结论:

结论 3: 本文提出的群组推荐算法的在离线时间和在线时间都比个性化推荐算法要少很多。

综合以上结果可以得出以下结论:

结论 4: 基于神经网络学习群组用户的隐表征 (或称聚合群组用户的偏好) 和项目的隐表征,可以有效提高群组推荐算法和多属性推荐算法的准确率和召回率,并使之非常接近个性化推荐算法。同时,在线运行时间较个性化推荐算法降低 30%以上, 离线时间较个性化推荐算法降低 50%以上。

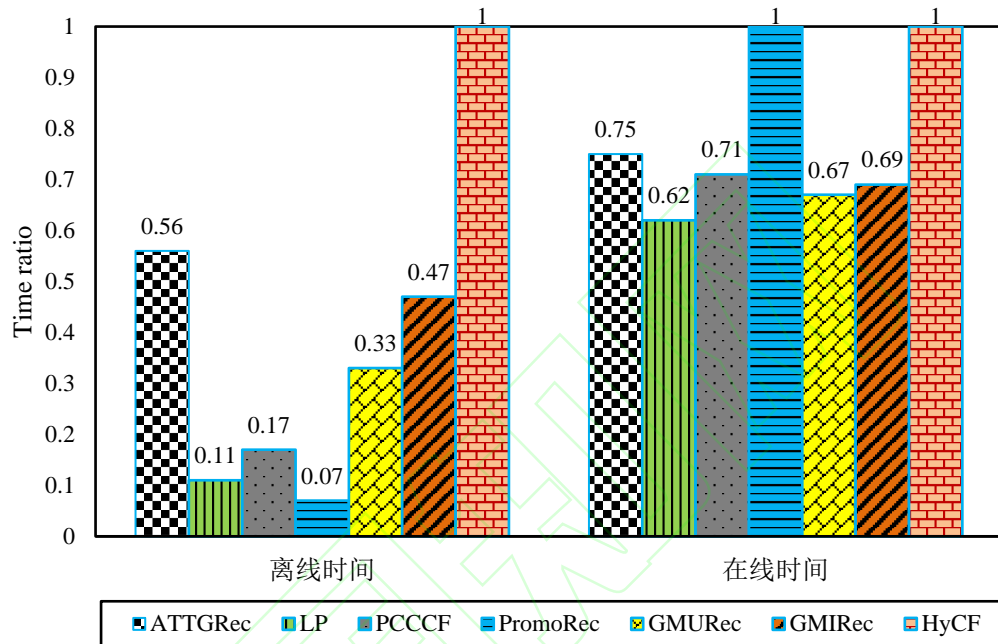


图 10 对比算法的运行时间

Fig. 10. Running time of the algorithm and baselines

6 结语

为了克服个体用户表征学习受个体用户评分稀疏性影响严重、推荐算法预测准确率和反应速度较低等问题,本文给出了一种基于神经网络的多属性评分隐表征学习方法;应用该方法从用户群组和项目两个维度分别学习多属性评分的隐表征,并分别通过用户群组偏好匹配和项目吸引力计算实现了两个群组推荐算法 GMURec、GMIRec。基于 TripAdvisor 数据集的实验结果表明:(1)通过学习群组用户的隐表征 (或称聚合群组用户的偏好) 以及学习项目的隐表征,可以有效提高群组推荐算法的准确率,并使之非常接近个性化推荐算法的准确率。(2)大幅提高推荐准确率的同时,使得在线运行时间较个体推荐算法降低 30%以上, 离线运行时间较个体推荐算法降低 50%以上。

自 Cao 等人^[36]的工作之后,已有很多人^[57,58]沿着深度神经网络的路线研究群组推荐算法。下一步我们将选择多个典型的深度神经网络群组推荐算法,比较一下本文算法及相关工作的性能。此外,本文实验中群组是给定的静态群组。然而,现实生活中群组是动态的。受文献^[29]的启发,未来我们将进一步探索面向动态群组的用户表征学习方法。

参考文献:

- [1] 朱成纯, 张谧. 基于活动的社交网络中的群组推荐算法设计[J]. 计算机系统应用, 2017, 26(9): 103-108. (Zhu Chengchun, Zhang Yan. Design of Group Recommendation Algorithm in Activity-Based Social Networks [J]. Journal of Computer System Applications, 26 (9): 103-8.)
- [2] 李珊, 姚叶慧, 厉浩, 等. 基于 ISA 联合聚类的组推荐算法研究 [J]. 数据分析与知识发现, 2019, 3(8):77-87. (Li Shan, Yao Yehui, Li Hao, et al. Research on group recommendation algorithm based on ISA joint clustering [J]. Data Analysis and Knowledge Discovery, 2019, 3 (8): 77-87.)
- [3] Sarwar B M, Karypis G, Konstan J, et al. Recommender Systems for Large-scale E-Commerce: Scalable Neighborhood Formation Using Clustering[J]. Communications, 2002, 50(12):158-167.
- [4] Lu J, Shambour Q, Xu Y, et al. A Web-Based Personalized Business Partner Recommendation System Using Fuzzy Semantic Techniques [J]. Computational Intelligence, 2013, 29(1):37-69.
- [5] Ortega F, Hernando A, Bobadilla J, et al. Recommending Items to Group of Users Using Matrix Factorization Based Collaborative Filtering [J]. Information Sciences, 2016, 345: 313-324.
- [6] 黄国言, 李有超, 高建培, 等. 基于项目属性的用户聚类协同过滤推荐算法[J]. 计算机工程与设计, 2010, 31(5): 1038-1041. (Huang Guoyan, Li Youchao, Gao Jianpei, et al. Collaborative Filtering Recommendation Algorithm Based on User Clustering of Item Attributes[J]. Computer Engineering and Design, 2010, 31(5): 1038-1041.)
- [7] 陈克寒, 韩盼盼, 吴健. 基于用户聚类的异构社交网络推荐算法[J]. 计算机学报, 2013, 36(2): 349-359. (Chen Kehan, Han Panpan, Wu Jian. User Clustering Based Social Network Recommendation [J]. Chinese Journal of Computers, 2013, 36(2): 349-359.)
- [8] 王晓军. 推荐系统中分布式混合协同过滤方法[J]. 北京邮电大学学报, 2016, 39(2): 25-29. (Wang Xiaojun. A Distributed Hybrid Collaborative Filtering Method in Recommender Systems [J]. Journal of Beijing University of Posts and Telecommunications, 2016, 39(2): 25-29.)
- [9] 黄贤英, 李沁东, 熊李媛. 结合拓扑势用户聚类的协同过滤推荐算法[J]. 计算机工程与设计, 2018, 39(1): 90-95. (Huang Xianying, Li Qindong, Xiong Liyuan. Collaborative Filtering Recommendation Algorithm with Topology Potential Combined User Clustering [J]. Computer Engineering and Design, 2018, 39(1): 90-95.)
- [10] 王兴茂, 张兴明, 吴毅涛, 等. 基于启发式聚类模型和类别相似度的协同过滤推荐算法[J]. 电子学报, 2016, 44(7): 1708-1713. (Wang Xingmao, Zhang Xingming, Wu Yitao, et al. A Collaborative Recommendation Algorithm Based on Heuristic Clustering Model and Category Similarity[J]. Acta Electronica Sinica, 2016, 44(7): 1708-1713.)
- [11] 张峻玮, 杨洲. 一种基于改进的层次聚类的协同过滤用户推荐算法研究[J]. 计算机科学, 2014, 41(12): 176-178. (Zhang Junwei, Yang Zhou. Collaborative Filtering Recommendation Algorithm Based on Improved User Clustering [J]. Computer Science, 2014, 41(12): 176-178.)
- [12] 李华, 张宇, 孙俊华. 基于用户模糊聚类的协同过滤推荐研究[J]. 计算机科学, 2012, 39(12): 83-86. (Li Hua, Zhang Yu, Sun Junhua. Research on Collaborative Filtering Recommendation Based on User Fuzzy Clustering [J]. Computer Science, 2012, 39(12): 83-86.)
- [13] 李贵, 陈召新, 李征宇, 等. 基于谱聚类群组发现的协同过滤推荐算法[J]. 计算机科学, 2014, 41(11A): 354-358. (Li Gui, Chen Zhaoxin, Li Zhengyu, et al. Collaborative Filtering Recommendation Algorithm Based on Spectral Clustering Subgroups Discovering[J]. Computer Science, 2014, 41(11A): 354-358.)
- [14] Zheng N, Bao H. Flickr Group Recommendation Based on User-Generated Tags and Social Relations via Topic Model[C]. In: Proceedings of the 10th International Symposium on Neural Networks. Springer, 2013: 514-523.
- [15] 陈婷, 朱青, 周梦溪, 等. 社交网络环境下基于信任的推荐算法[J]. 软件学报, 2017, 28(3): 721-731. (Chen Ting, Zhu Qing, Zhou Mengxi, et al. Trust-Based Recommendation Algorithm in Social Network [J]. Journal of Software, 2017, 28(3): 721-731.)

- [16] Quijano-Sanchez L, Recio-Garcia J A, Diaz-Agudo B. Personality and Social Trust in Group Recommendations[C]. In: Proceedings of the IEEE International Conference on Tools with Artificial Intelligence. IEEE Computer Society, 2010: 121-126.
- [17] Lai CH, Liu DR, Lin CS. Novel Personal and Group-Based Trust Models in Collaborative Filtering for Document Recommendation [J]. Information Sciences, 2013, 239: 31-49.
- [18] Kagita V R, Pujari A K, Padmanabhan V. Virtual User Approach for Group Recommender Systems Using Precedence Relations[J]. Information Sciences, 2015, 294: 15-30.
- [19] Ma Y, Ji S, Liang Y, et al. A Hybrid Recommendation List Aggregation Algorithm for Group Recommendation[C]. In: Proceedings of the 2015 IEEE / WIC / ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT). ACM, 2015: 405-408
- [20] Chao D L, Balthrop J, Forrest S. Adaptive Radio: Achieving Consensus Using Negative Preferences[C]. In: Proceedings of the International ACM Siggroup Conference on Supporting Group Work. ACM, 2005: 120-123.
- [21] Masthoff J. Group Modeling: Selecting a Sequence of Television Items to Suit a Group of Viewers [J]. User modeling and user-adapted interaction, 14(1): 37-85.
- [22] Deventer, M. O. van, et al. Group Recommendation in a Hybrid Broadcast Broadband Television Context [C]. In: Proceedings of the 21st Conference on User Modeling, Adaptation, and Personalization, 2013.
- [23] McCarthy, Kevin, et al. CATS: A Synchronous Approach to Collaborative Group Recommendation[C] In: Proceedings of the Nineteenth International Florida Artificial Intelligence Research Society Conference, 2006: 86-91.
- [24] Ntoutsis I, Stefanidis K, Norvag K, et al. gRees: A group recommendation system based on user clustering[C] In: Proceedings of the International Conference on Database Systems for Advanced Applications. Springer-Verlag, 2012: 299-303.
- [25] Quijano-Sanchez L, Recio-Garcia J A, Diaz-Agudo B. An architecture and functional description to integrate social behaviour knowledge into group recommender systems [J]. Applied Intelligence, 2014, 40(4): 732-748.
- [26] McCarthy JF. Pocket Restaurantfinder a situated recommender system for groups[C] In: Proceedings of the Workshop on Mobile Ad hoc Communication at the 2002 ACM Conference on Human Factors in Computer Systems, 2002
- [27] 刘荣荣. 考虑时间情境的群体推荐算法研究 [J]. 武汉理工大学学报(信息与管理工程版), 2016, 38(1): 597-600. (Liu Rongrong. Research on Group Recommendation Algorithm Considering Time Context [J]. Journal of Wuhan University of Technology (Information and Management Engineering Edition), 2016, 38(1): 597-600.)
- [28] 郭均鹏, 赵梦楠. 面向在线社区用户的群体推荐算法研究 [J]. 计算机应用研究, 2014, 31(3): 696-699. (Guo Junpeng, Zhao Mengnan. Research on Group Recommendation Algorithms for Online Community Users [J]. Journal of Computer Applications, 2014, 31(3): 696-699.)
- [29] Roy S B, Lakshmanan L V, Liu R, et al. From Group Recommendations to Group Formation[C] In: Proceedings of the International Conference on Management of Data, 2015: 1603-1616.
- [30] 赵梦楠. 基于 SNA 的群体推荐系统的研究[J]. 信息技术, 2000(2): 199-202. (Zhao Mengnan. Research on SNA-based group recommendation system [J]. Information Technology, 2000 (2): 199-202.)
- [31] 梁昌勇, 冷亚军, 王勇胜, 等. 电子商务推荐系统中群体用户推荐问题研究 [J]. 中国管理科学, 2013, V21(3): 153-8. (Liang Changyong, Leng Yajun, Wang Yongsheng, et al. Research on group user recommendation in e-commerce recommendation system [J]. China Management Science, 2013, V21 (3): 153-8.)
- [32] Shang, S., Hui P., Kulkarni, S. R., Cuff, P. W. Wisdom of the Crowd: Incorporating Social Influence in Recommendation Models[C] In: IEEE International Conference on Parallel & Distributed Systems. IEEE, 2012: 835-840.

- [33] 李文华, 熊晓栋, 郭均鹏. 一种基于案例推理和协商的群体推荐算法 [J]. 系统工程, 2013, 11: 93-98. (Li Wenhua, Xiong Xiaodong, Guo Junpeng. A group recommendation algorithm based on case-based reasoning and negotiation [J]. Systems Engineering, 2013, 11: 93-98.)
- [34] Sacharidis D. Top-N group recommendations with fairness[C] In: Proceedings of the ACM Symposium on Applied Computing, 2019: 1663-1670.
- [35] Maria S, Jyrki N, Evaggelia P, et al. Fair sequential group recommendations[C] In: Proceedings of the 35th Annual ACM Symposium on Applied Computing, 2020: 1443-1452.
- [36] Cao D, He X, Miao L, et al. Attentive Group Recommendation[C] In: Proceedings of the International ACM Sigir Conference on Research and Development in Information Retrieval, 2018: 645-654.
- [37] Q. Zhu, M. Zhou, J. Liang, T. Yan, S. Wang. Efficient Promotion Algorithm by Exploring Group Preference in Recommendation[C] In: Proceedings of the IEEE International Conference on Web Services. IEEE, 2016: 268-275.
- [38] Bilge A, Kaleli C. A multi-criteria item-based collaborative filtering framework [C] In: Proceedings of the IEEE 2014 11th International Joint Conference on Computer Science and Software Engineering (JCSSE), 2014: 18-22.
- [39] Adomavicius G, Kwon Y O. New Recommendation Techniques for Multicriteria Rating Systems [J]. IEEE Intelligent Systems, 22(3): 48-55.
- [40] Nilashi M, Esfahani M D, Roudbaraki M Z, et al. A Multi-Criteria Collaborative Filtering Recommender System Using Clustering and Regression Techniques[C] In: Proceedings of the Decision Support Systems, 2016: 24-30.
- [41] Majumder G S, Dwivedi P, Kant V. Matrix Factorization and Regression-Based Approach for Multi-Criteria Recommender System[C] In: Proceedings of the International Conference on Information & Communication Technology for Intelligent Systems, 2017: 103-110.
- [42] 陈冬林, 聂规划. 基于商品属性隐性评分的协同过滤算法研究 [J]. 计算机应用研究, 2006, 26(4): 966-968. (Chen Donglin, Nie Planning. Research on collaborative filtering algorithm based on recessive scoring of commodity attributes [J]. Journal of Computer Applications, 2006, 26 (4): 966-968.)
- [43] Zhang J, Peng Q, Sun S, et al. Collaborative filtering recommendation algorithm based on user preference derived from item domain features [J]. Physica A Statistical Mechanics & Its Applications, 2014, 396: 66-76.
- [44] Jannach D., Karakaya Z., Gedikli F., Accuracy improvements for multi-criteria recommender systems[C]. In: Proceedings of the ACM Conference on Electronic Commerce, ACM 2012: 674-689.
- [45] D. Jannach, M. Zanker, M. Fuchs, Leveraging multi-criteria customer feedback for satisfaction analysis and improved recommendations[J] Information Technology & Tourism. 2014, 14(2): 119-149.
- [46] Y. Zheng, Criteria Chains: A Novel Multi-Criteria Recommendation Approach[C] In: Proceedings of the International Conference on Intelligent User Interfaces. ACM 2017: 29-33.
- [47] 覃正, 李岱峰. 一种基于资源多属性分类的群组推荐模型[J]. 统计与决策, 2010, (8): 153-155. (Qin Zheng, Li Daifeng. A group recommendation model based on multi-attribute classification of resources [J]. Statistics and decision-making, 2010, (8): 153-155.)
- [48] Garcia I, Sebastia L, Onaindia E, et al. A Group Recommender System for Tourist Activities[C] In: Proceedings of the Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics, 2009: 26-37
- [49] Liang, Geng. Neuron Adaptive and Neural Network Based on Gradient Descent Searching Algorithm for Diagonalization of Relative Gain Sensitivity Matrix Decouple Control for MIMO System. [C] In: Proceedings of the IEEE International Conference on Networking. IEEE, 2008: 368-373.

- [50] 哈林顿, P.). 机器学习实战[M]. 人民邮电出版社, 2013. 15-31. (Harrington, P.). Machine Learning in Action [M]. People's Posts and Telecommunications Press, 2013. 15-31.)
- [51] 范永全, 杜亚军. 基于加权相似度的用户协同过滤方法[J]. 计算机工程与应用, 2016, 52(22):222-225. (Fan Yongquan, Du Yajun. User collaborative filtering method based on weighted similarity [J]. Computer Engineering and Applications, 2016, 52 (22): 222-225.)
- [52] Sai L N, Shreya M S, Subudhi A A, et al. Optimal k-means clustering method using silhouette coefficient[J]. International Journal of Applied Research on Information Technology and Computing, 2017, 8(3): 335-344.
- [53] Yuan Zhengwu, Chen Ran. Collaborative filtering recommendation algorithm based on multi-level hybrid similarity[J]. Journal of Computer Applications, 2018, 38(3): 633-638.
- [54] Chao C, Qu S, Du T. Research of Collaborative Filtering Recommendation Algorithm for Short Text[J]. Journal of Computer & Communications, 2014, 2(14):59-66.
- [55] 丁晟春, 王小英, 刘梦露. 基于本体和加权朴素贝叶斯的网络舆情主题分类[J]. 现代情报, 2018, 38(8):12-17, 34. (Ding Shengchun, Wang Xiaoying, Liu Menglu. Classification of Internet Public Opinion Topics Based on Ontology and Weighted Naive Bayes [J]. Modern Information, 2018, 38(8):12-17, 34.)
- [56] Wang Y, Deng J, Gao J, et al. A Hybrid User Similarity Model for Collaborative Filtering [J]. Information Sciences, 2017, 418-419(12): 102-118.
- [57] Tran L V, Pham T A N, Tay Y, et al. Interact and Decide: Medley of Sub-Attention Networks for Effective Group Recommendation[C] In: Proceedings of the 42nd International ACM SIGIR Conference. ACM, 2019: 255-264.
- [58] Delic A, Ricci F, Neidhardt J, et al. Preference Networks and Non-Linear Preferences in Group Recommendations[C] In: Proceedings of the Web Intelligence, 2019: 403-407.

(通讯作者: 纪淑娟, ORCID: 0000-0003-2650-0161, E-mail: jsjsuzie@sina.com。)

基金项目: 本文系青岛社会科学规划研究项目 (QDSKL1801138)、国家自然科学基金项目 (71772107, 71403151, 62072288)、山东省自然科学基金项目 (ZR2018BF013)、山东省研究生质量提升计划项目 (2016)、山东科技大学山海英才计划项目 (2019)、泰山学者攀登计划项目 (2014) 的研究成果之一。

作者贡献声明:

张纯金: 提出研究思路, 设计研究方案;

郭盛辉: 算法设计与实验;

纪淑娟: 论文起草与实验结果分析;

杨伟: 采集、清洗和分析数据;

伊磊: 论文草稿排版与最终版本修订。

利益冲突声明:

所有作者声明不存在利益冲突关系。

支撑数据:

支撑数据由作者自存储, <https://pan.baidu.com/s/1GdUAFf8NP1dleSRFeVvGWA>;