

doi:10.19920/j.cnki.jmsc.2025.04.010

考虑演唱能力和社交反馈因素的演唱歌曲推荐方法^①

陈卓华^{1,2}, 刘红岩^{1,2*}, 高歌³

(1. 清华大学现代管理研究中心, 北京 100084; 2. 清华大学经济管理学院, 北京 100084;
3. 西南财经大学工商管理学院, 成都 611074)

摘要:近年来,在线唱歌平台吸引了大量用户在线演唱和发布歌曲。为用户推荐合适的演唱歌曲对于平台优化用户体验、增强用户粘性尤为重要。用户的演唱能力是推荐模型需要考虑的因素,然而如何预测用户演唱能力是具有挑战性的问题。在线唱歌平台同时提供丰富的社交功能,如何利用社交反馈因素提升推荐性能是值得研究的问题。为了解决这些问题,本文提出了新的用户演唱能力预测方法和社交反馈因素建模方法。在此基础上提出了一种新颖的演唱歌曲推荐模型。在真实数据集上的实验结果验证了本文所提方法的有效性以及演唱能力和社交反馈因素对于提升推荐性能的必要性。

关键词:演唱歌曲推荐;演唱能力;社会网络;深度学习

中图分类号:TP391 **文献标识码:**A **文章编号:**1007-9807(2025)04-0159-15

0 引言

唱歌是一种广受欢迎的娱乐消遣方式,近年来在线唱歌平台如全民k歌、唱吧、StarMaker等吸引了大量用户使用。据统计,全民k歌的日活在2020年已经达到了1.7亿用户。在这些平台上,人们不仅可以选取喜欢的歌曲进行演唱,还可以将演唱的作品发布出来供其他人欣赏。同时,这类平台还提供社交网络服务,一个用户可以在平台上关注其他用户,接收其发布歌曲的动态,并对这些作品进行点赞、评论和转发。由于在选取演唱的歌曲时,平台可供用户选择的歌曲类型和数量众多,同时用户对服务的需求往往是个性化的,如何满足用户个性化需求是网络平台服务的重要管理问题,一个精准的唱歌推荐系统对于解决这一管理问题、降低用户的寻找成本、优化用户使用体验、增强用户黏性是十分重要的。

演唱歌曲推荐需要考虑的因素之一是用户的

演唱能力。不同于对收听歌曲或者其他商品或服务的推荐,由于每个人的嗓音条件不同,演唱歌曲的推荐不仅要考虑用户偏好,还要考虑用户是否有能力唱好歌曲。已有的一些研究证明了用户偏向于演唱他们能够唱好的歌曲^[2-6],并提出了一些对用户演唱能力进行评估和建模的方法。但是这些方法存在一些不足,无法很好地应用于现实问题中。例如Guan等^[3,4]提出的方法依赖于两类特殊的音频文件,然而对大部分歌曲来说,这两类文件是不易获取的;Mao等^[5,6]的研究则借助人工对用户的演唱能力进行建模,这限制了该方法的效率。因此如何高效地评估和建模用户演唱能力并将这一信息用于演唱歌曲的推荐是一个有挑战性的研究问题。

演唱歌曲推荐需要考虑的另一个因素是在线唱歌平台上的社会影响,已有研究已经证明了社交媒体平台上的社交网络关系和网络个体的行为特征之间具有较强的相互影响关系^[7]。图1展示了某在线唱歌平台上的“关注”页面,该页面展示

① 收稿日期:2021-11-24; 修订日期:2022-09-22。

基金项目:国家社会科学基金资助项目(20&ZD161)。

通讯作者:刘红岩(1968—),女,山东烟台人,博士,教授,博士生导师。Email:hyliu@tsinghua.edu.cn

了用户好友的动态,图中方框中的数字分别表示该作品收到的播放、点赞、评论和转发的数量,其他在线唱歌平台也都提供类似的功能。根据社会影响理论^[8],用户的行为会受到与其有社交关系的用户(即用户好友)的影响,已有研究也对如何建模社会影响以提升推荐效果进行了大量的研究,但是没有充分考虑在线唱歌平台和歌曲演唱的特点。

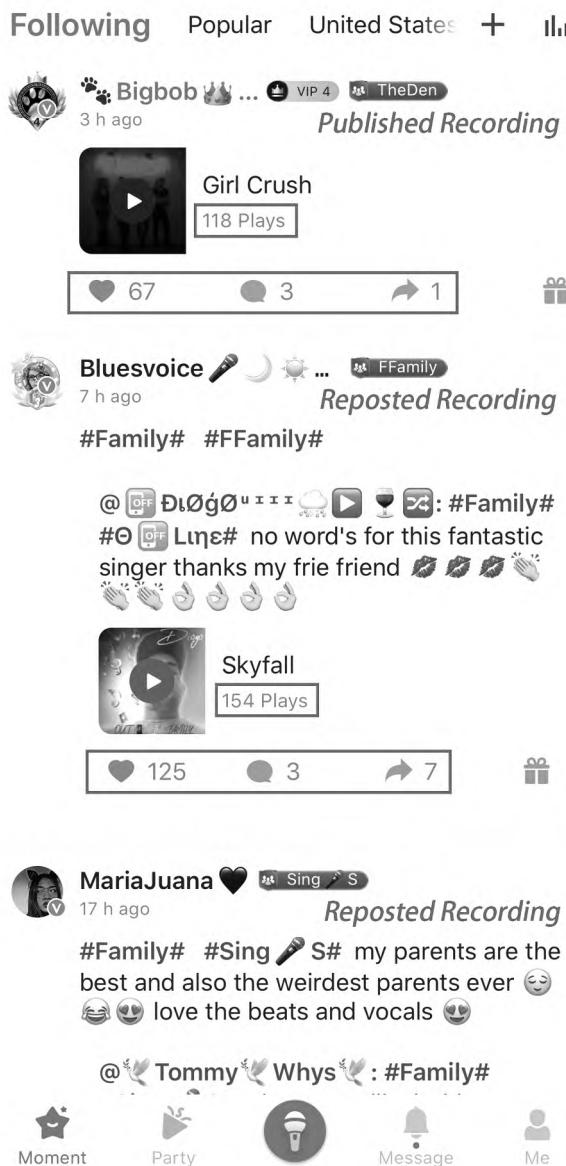


Fig. 1 Square tab of an online karaoke APP

平台上的任何一名用户既可以关注其他用户,也会被其他用户关注,本文将一名用户关注的其他用户构成的社区称为这名用户的“偶像社区”,将关注这名用户的其他用户构成的社区称为其“粉丝社区”,这两类社区中的用户发布的作

品称为对应社区的作品。本文将一个用户的好友定义为其偶像社区中的其他用户。用户发布的作品会在该用户的粉丝社区中收到丰富的社交反馈行为如播放、点赞、评论和转发等,这些社交反馈因素与用户的演唱行为之间存在密切关系。

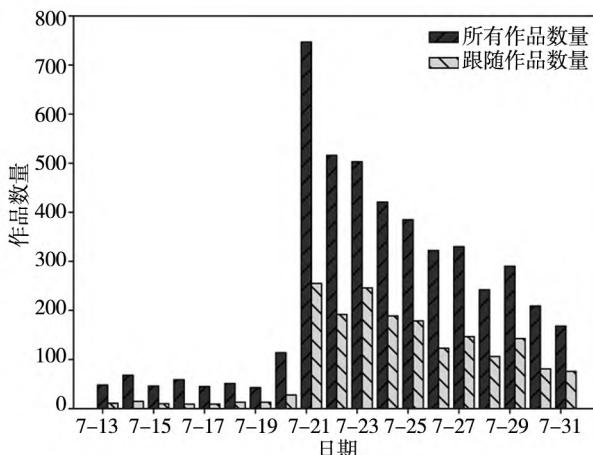


图 2 林肯公园歌曲作品数量与跟随作品数量

Fig. 2 The number of recordings and following recordings of Linkin Park

本文对某在线唱歌平台上用户在 2017 年 7 月 11 日—31 日期间演唱并发布林肯公园乐队歌曲的情况进行了分析。图 2 展示了用户在这段时间内发布林肯公园的歌曲作品的数量与跟随作品的数量。其中“跟随作品”指的是该作品发布者是在其偶像社区出现了该乐队歌曲的作品以后发布的这一作品。从图中可以看到,用户发布的作品数量在 21 日出现显著增长,且之后每天发布的作品中很大一部分属于跟随作品。这一现象说明用户是否发布该乐队歌曲的行为与其关注的用户是否发布过该乐队的歌曲存在着一定的相关关系。可能由于用户的演唱能力不同,大家会选择同乐队不同的歌曲演唱,即不局限于同一首歌,还会延伸到当前流行的歌曲相关联的其他歌曲。图 3 展示了作品收到的播放、点赞和评论次数与该作品对应的跟随作品发布数量的关系,从图中可以看到这些社交反馈的数量与跟随作品的发布数量之间存在正相关关系。

以上分析说明了用户选择演唱并发布作品与三方面因素有关:1) 用户之间的社交网络关系;2) 歌曲在音乐社区中的关联程度;3) 作品收到的社交反馈信息。现有的社会化推荐方法中大部分研究只考虑了第一方面的因素,少数方法同时考虑了第一方面、第二方面的因素,还没有研究同时

考虑这三方面的因素.

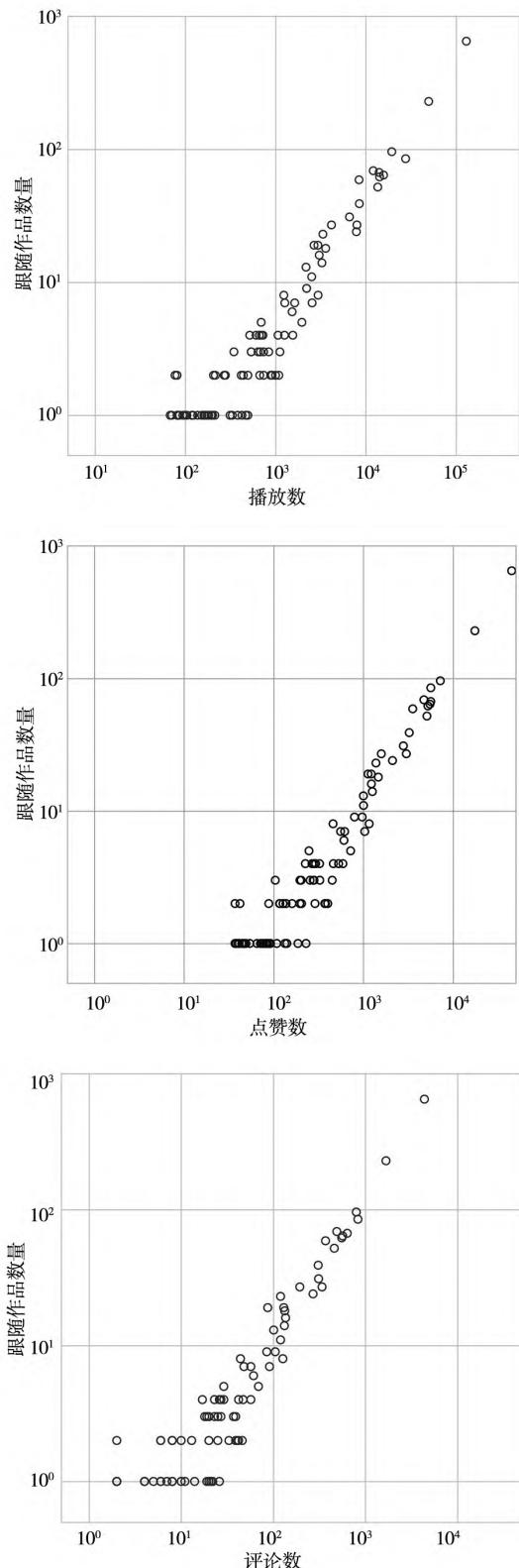


图3 林肯公园歌曲收到的社交反馈数量与跟随作品数量的关系

Fig. 3 The correlations between the number of social feedbacks and the number of following recordings in Linkin Park

在线唱歌平台关注的往往不仅是用户演唱一首歌,更希望用户演唱后发布该作品,因此本文以下部分提到的演唱行为均指用户演唱一首歌曲并发布这一作品.针对在线唱歌平台的演唱歌曲推荐问题,本文提出了一个新颖的推荐方法 VSRec (vocal competence and social influence based recommendation),主要贡献如下:

1) 提出了一种新的用户演唱能力建模方法,该方法不需要关于歌曲本身的额外信息如伴奏信息或专家知识来进行声学建模,只需要用户演唱的结果,因此这种方法通用型更强、可用于大规模的数据.实验结果表明该方法不仅可以准确地衡量用户的演唱能力,还能提升演唱歌曲推荐效果.

2) 提出了新的图神经网络模型 THGCN (Tri-layer Heterogeneous Graph Convolution Network).该模型首先通过一个三层异构网络建模了用户之间、歌曲之间以及作品之间的多元复杂关系,针对这种建模方法,本文提出了充分利用网络结构和节点特征信息的图神经网络学习方法,以更加准确地学习作品和歌曲的特征以及用户的偏好.

3) 提出了一个综合用户对歌曲的喜爱程度、演唱能力和社交反馈因素的影响实现演唱歌曲推荐的端到端模型.在一个大规模真实数据集上的实验证明了该模型相对于已有模型的优势,并证明了模型各模块设计的有效性和必要性.

下文将首先梳理相关研究进展,随后对唱歌推荐问题进行介绍和形式化定义,然后详细描述了本文所提模型的细节,并展示了实验结果和分析,最后对本文的研究进行了总结.

1 相关研究

1.1 演唱歌曲推荐模型

演唱歌曲推荐的目标是为用户推荐其最有可能演唱的歌曲,已有的音乐推荐方法一般基于用户对歌曲的喜爱程度进行收听歌曲的推荐^[9-11],而相比于收听歌曲的推荐,演唱歌曲的推荐情景更为复杂,用户对歌曲的喜爱程度只是需要考虑的因素之一.一些演唱歌曲的推荐直接利用已有方法^[12],没有针对这一具体场景进行优化.而演唱歌曲推荐的一个重要的方面是用户的演唱能

力,现有演唱歌曲推荐方法中对演唱能力的建模方法可以分为以下三类:

第一类方法是 Mao 等^[1]的研究,这类方法基于全体用户的演唱历史构建歌曲演唱难度偏序网络,随后使用重启随机游走算法 (random walk with restart)为用户找到其能较好演唱的歌曲。此研究中存在一个隐含的假设:歌曲之间的演唱难度偏序关系对所有用户是一致的,但这一假设在现实中往往不成立。

第二类方法通过音频评估技术来建模用户的演唱能力。通过这一技术将用户演唱结果的音频与原声作比较,可以得到用户在音量、音准和节奏三个维度上的准确度。Guan 等^[3]构建了用户 - 歌曲、用户 - 音频、歌曲 - 音频三个关系矩阵,使用非负张量分解的方法提取特征,然后将这些特征输入到预测模型中。这类方法所依赖的音频评估技术需要同时用到歌曲的原唱人声和伴奏这两个文件,然而实际应用中所有歌曲同时拥有这两类文件是较难的。

第三类方法通过提取声学特征对用户的演唱能力进行建模。Mao 等^[5]使用专家知识对用户和歌曲构建声学模型,然后从声学模型中提取特征来建模用户的演唱能力。这类方法直接明了,然而在建模过程中需要大量的人工干预,难以被应用于大规模的情景中。

这些研究对用户演唱能力的建模仍存在不足,且在进行歌曲推荐时只考虑了用户的演唱能力,忽略了用户对歌曲的喜爱程度以及音乐社区的影响。因此如何构建一个更有效的用户演唱能力建模方法并结合更多影响因素进行演唱歌曲的推荐是一个值得研究的方向。

1.2 社会化推荐模型

在具有社交属性的平台上,用户的行为往往会受到其好友的影响,用户的在线互动如点赞、评论等是用户行为的一个重要影响方面^[13]。社会化推荐模型在推荐过程中加入社交因素的影响,以提升推荐效果^[14, 15]。

很多社会化推荐模型在矩阵分解模型的基础上加入了社交因素的影响。一些研究^[16-20]认为具有好友关系的用户的隐特征向量应该更相似,因此在模型训练过程中加入了相关的约束。另一种思路^[21-23]则借鉴了贝叶斯个性化排序 BPR^[24]的

做法,在考虑物品之间的排序关系时将用户好友的历史选择考虑在内,以此构建相应的优化目标。还有一些工作^[25-28]对用户与物品的交互矩阵和用户之间的社交关系矩阵进行协同分解,通过这种做法将用户的社交关系考虑在内。也有一些研究基于神经网络模型将用户的偏好和社交关系关联在一起^[30-32]或同时考虑了用户自己对物品的偏好及其好友对这些物品的偏好^[33-37]。

一些模型借助图神经网络^[38, 39]在挖掘图信息上的优势来处理用户的社交网络,从而将其加入到推荐模型中。Wu 等^[40]使用图卷积网络在社交网络上聚合一一个用户的好友的隐特征向量,将聚合的结果作为该用户的表示。何云飞等^[41]使用图卷积神经网络处理异构图以在提升推荐效果的同时缓解冷启动和数据稀疏性问题。

在唱歌推荐任务中,社会化因素的影响体现在音乐社区的影响上,这种影响不仅来自用户与用户之间、歌曲与歌曲之间的关系,还与作品收到的社交反馈有关。已有的研究大部分只聚焦在用户与用户之间的关系,少部分研究考虑了物品与物品之间的关系,但还没有研究物品收到的社交反馈情况带来的影响。

2 研究背景与问题定义

2.1 研究背景

不同的人拥有不同的演唱能力,一首歌曲的 MIDI(musical instrument digital interface)文件保存了音频的音符信息。用户演唱的音符与标准音符的一致性反映了用户的演唱能力。绝大部分在线唱歌平台都会在用户演唱歌曲后对演唱结果进行评估并给出评级,用户可以根据评级决定是否将作品发布出来。一般来讲评级越高,用户越倾向于选择发布。

在线唱歌平台中,用户可以通过点击“关注”来与其他用户建立社交关系,从而形成“偶像社区”和“粉丝社区”。用户发布的作品会在粉丝社区中收到丰富的社交反馈。用户选择歌曲进行演唱的行为与其偶像社区作品及这些作品收到的社交反馈密切相关。因此,在线唱歌平台中,除了用户对歌曲的喜爱程度,演唱能力和音乐社区

的反馈因素都会在用户选择演唱歌曲的行为中产生作用,本文研究如何同时建模这三方面因素,以提升推荐效果.

2.2 问题定义

唱歌推荐问题:给定用户集合 U ,歌曲集合 S ,每个用户的演唱记录 S_i 以及它们对应的 MIDI 序列 $M_{i,j}$,用户的偶像社区中的作品集合 R_i ,针对用户 u_i 和一首他未演唱过的歌曲 s_j ,唱歌推荐问题被定义为如何预测用户 u_i 演唱歌曲 s_j 的可能性,并进而进行歌曲推荐.

令 $U = \{u_1, u_2, \dots, u_n\}$ 表示在线唱歌平台上的用户集合, n 为用户数量. $S = \{s_1, s_2, \dots, s_m\}$ 表示歌曲集合, m 为歌曲数量. 对用户 $u_i \in U$, 令 $S_i \subseteq S$ 表示其演唱过的歌曲集合,对于 S_i 中的每一首歌曲 $s_{i,j}$,其 MIDI 音符序列表示为 $M_{i,j}$. 令 $R_i = \{r_{i,1}, r_{i,2}, \dots, r_{i,m_i}\}$ 表示用户 u_i 的好友发布的作品,其中 m_i 为作品数量. 每一个作品 $r_{i,k} \in R_i$ 都是一个 $(r_{i,k}^u, r_{i,k}^s)$ 对,其中 $r_{i,k}^u$ 代表作品的创作者, $r_{i,k}^s$ 为作品对应的歌曲,使用 $r_{i,k}^f$ 表示这个作品收到的社交反馈信息. 本文用到的主要符号及含义见表 1.

表 1 本文使用的主要符号及其含义

Table 1 The notations and their meanings

| 符号 | 含义 |
|------------------------|--------------------------------|
| U | 所有用户的集合 |
| S, S_i | 所有歌曲的集合, 用户 u_i 演唱过的歌曲集合 |
| R_i | 用户 u_i 的偶像社区中的作品集合 |
| R_i^u, R_i^s | R_i 中的作品对应的发布者和歌曲的集合 |
| $r_{i,k}$ | 用户 u_i 的偶像社区中的一个作品 |
| $r_{i,k}^u, r_{i,k}^s$ | $r_{i,k}$ 对应的发布者和歌曲 |
| $r_{i,k}^f$ | $r_{i,k}$ 收到的社交反馈 |
| $M_{i,j}$ | 用户 u_i 唱过的歌曲 s_j 的 MIDI 序列 |
| U_i | 用户 u_i 的隐特征向量 |
| V_j | 歌曲 s_j 的隐特征向量 |
| $PI_{i,j}$ | 用户 u_i 对歌曲 s_j 的喜爱程度 |
| PU_i | 用户 u_i 的声学模型 |
| PS_j | 歌曲 s_j 的声学模型 |
| $VC_{i,j}$ | 用户 u_i 对歌曲 s_j 的演唱能力 |
| $SI_{i,j}$ | 作品 (u_i, s_j) 受到的社会因素影响 |
| $R_{i,j}$ | 用户 u_i 会演唱歌曲 s_j 的可能性分数 |

3 考虑演唱能力和社交反馈因素影响的推荐模型

在计算用户演唱一首歌曲的可能性时需要考虑三方面因素:用户对歌曲的喜爱程度、用户的演唱能力和社交反馈因素的作用. 在对用户的喜爱程度建模上,本文使用的是经典的矩阵分解模型:对用户 u_i 和歌曲 s_j ,用户对歌曲的喜爱程度 $PI_{i,j}$ 定义为用户的隐特征向量 $U_i \in \mathbb{R}^d$ 与歌曲的隐特征向量 $V_j \in \mathbb{R}^d$ 的内积

$$PI_{i,j} = U_i^T V_j \quad (1)$$

3.1 对用户演唱能力的建模

本文对用户演唱能力的建模借鉴了 Mao 等^[6]的思路,并弥补了其不能自动、高效建模的不足,从而可以在大规模的平台上应用. Mao 等的方法通过专家评估对用户和歌曲构建声学模型,声学模型由两个维度组成,分别代表每一时刻音符的音高和强度. 用户的声学模型代表的是用户在每一个音符上的演唱能力,而歌曲的声学模型则表示标准的演唱结果.

考虑到在线唱歌平台上绝大部分都是非专业的歌手,音符的强度对于评估其演唱结果重要性不大且难以自动识别,因此本文将强度这一维度去掉. 同时加入了对相邻两个音符之间音高关系的度量:一般来讲,人们更容易演唱两个音高差距较小的连续音符,而演唱差距较大的连续音符更难.

为了实现,首先对连续的音频文件采样以获得离散 MIDI 音符特征,这一特征包含了音符的音高信息,采样时间间隔为 200 毫秒,得到的每一维度的特征是取值在 21 ~ 108 的数值,使用 M 表示 MIDI 音符所有可能取值的集合. 对用户 u_i 演唱过的歌曲 s_j ,使用 $M_{i,j}$ 来表示演唱作品的 MIDI 序列

$$M_{i,j} = \langle m_{i,j,1}, m_{i,j,2}, \dots, m_{i,j,L_j} \rangle \quad (2)$$

其中 L_j 表示歌曲经过采样以后得到的音符序列的长度, $m_{i,j,k} \in M, k = 1, 2, \dots, L_j$.

对歌曲 s_j , 使用平台上所有用户的演唱作品来构建其 MIDI 序列 M_j^*

$$\begin{aligned} M_j^* = & \langle med(m_{i,j,1} | u_i \in U_j), \\ & med(m_{i,j,2} | u_i \in U_j), \\ & \dots, \\ & med(m_{i,j,L_j} | u_i \in U_j) \rangle \end{aligned} \quad (3)$$

其中 U_j 为所有演唱过歌曲 s_j 的用户, $med(\cdot)$ 为中位数函数. 由于 M_j^* 考虑了所有用户的演唱结果, 因此本文将其看成是歌曲 s_j 的标准音符序列, 采用中位数是为了减小演唱能力特别强或特别弱的用户对建模结果的影响.

本文通过定义演唱情景来为用户和歌曲构建声学模型, 令 $midi_k$ 表示 M 中的第 k 个音符, 使用 $midi_0$ 表示不需要用户演唱的音符. 演唱情景被定义为音符对 $\langle midi_k, midi_l \rangle$, 代表歌曲中音符 $midi_k$ 后出现的是音符 $midi_l$. 一首歌曲的声学模型 PS_j 可以表示为一个矩阵, 矩阵中的每个元素 $PS_{j,k,l}$ 代表了演唱情景 $\langle midi_k, midi_l \rangle$ 出现的频率

$$PS_{j,k,l} = \frac{\sum_{t=1}^{L_j} \delta(M_{j,t}^* = midi_k \wedge M_{j,t-1}^* = midi_l)}{L_j} \quad (4)$$

其中 $M_{j,t}^*$ 为 M_j^* 序列中的第 t 个音符, $\delta(\cdot)$ 是示性函数

$$\delta(x) = \begin{cases} 1, & \text{若 } x \text{ 为真} \\ 0, & \text{否则} \end{cases} \quad (5)$$

歌曲的声学模型表示的是这首歌包含的演唱情景信息, 用户的声学模型则需要反映用户在不同的演唱情景下的演唱能力. 通过比较用户历史演唱歌曲的音符序列与标准的音符序列可以达到这一目的: 使用 $n_{i,k,l}$ 来表示用户 u_i 在过去的演唱中遇到过演唱情景 $\langle midi_k, midi_l \rangle$ 的次数, 使用 $d_{i,k,l}$ 来表示用户在这一演唱情景中的演唱误差, 通过式(6)~式(9)建立用户的声学模型矩阵 PU_i

$$n_{i,k,l} = \sum_{s_j \in S_i} L_j \times PS_{j,k,l} \quad (6)$$

$$d_{i,k,l} = \sum_{s_j \in S_i} \sum_{t=1}^{L_j} \delta(M_{j,t}^* = midi_k \wedge M_{j,t-1}^* = midi_l) \times |M_{i,j,t} - M_{j,t}^*| \quad (7)$$

$$d^* = \max\left(\frac{d_{i,k,l}}{n_{i,k,l}} \mid u_i \in U, midi_k \in M, midi_l \in M\right) \quad (8)$$

$$PU_{i,k,l} = \frac{d^* - \frac{d_{i,k,l}}{n_{i,k,l}}}{d^*} \quad (9)$$

从式(9)可以看出, 用户声学模型的每一个维度取值都在 $[0, 1]$ 内, 对于用户没有遇到过的演唱情景 $\langle midi_k, midi_l \rangle$, 其对应的 $PU_{i,k,l}$ 取值为 0.

在得到了歌曲和用户的声学模型以后, 本文设计多层感知器 (MLP) 模型来衡量用户的演唱能力. 为了便于计算, 首先将歌曲和用户的声学模型矩阵展开成一维的向量, 然后将其哈达玛 (Hadamard) 乘积输入到模型中来计算用户对歌曲的演唱能力 $VC_{i,j}$, 如式(10)所示

$$VC_{i,j} = \sigma(W_v' \times \text{ReLU}(W_v \times [Unfold(PU_i \otimes PS_j)] + b_v) + b_v') \quad (10)$$

其中 W_v, W_v', b_v, b_v' 均为待学习的参数, \otimes 表示哈达玛乘积, $Unfold(\cdot)$ 是将矩阵展开成向量的操作, $\text{ReLU}(\cdot)$ 是线性整流函数^[43], 其定义如式(11)所示, $\sigma(\cdot)$ 是 Sigmoid 函数, 定义如式(12)所示

$$\text{ReLU}(x) = \begin{cases} x, & \text{若 } x \geq 0 \\ 0, & \text{若 } x < 0 \end{cases} \quad (11)$$

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (12)$$

对用户演唱能力进行建模的模型结构如图 4 所示.

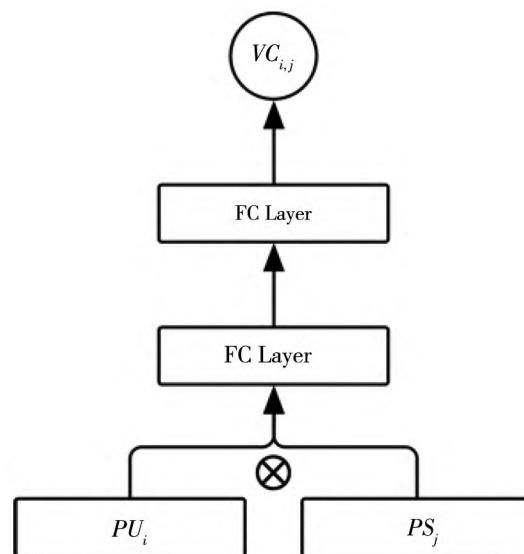


图 4 用户演唱能力建模结构

Fig. 4 The architecture to model vocal ability

3.2 对社交反馈因素的建模

用户发布的作品会在粉丝社区中收到丰富的社交反馈,这些反馈反映了这一作品在粉丝社区中的受欢迎程度,而一个用户的偶像社区中的作品及受欢迎程度与该用户选择演唱歌曲的行为相关。为了对这种社交反馈因素进行建模,对于用户 u_i 的偶像社区中的一个作品 $r_{i,k}$,首先提取其基础的社交反馈特征 $r_{i,k}^f$

$$r_{i,k}^f = [play_{i,k}, like_{i,k}, repost_{i,k}, comment_{i,k}] \quad (13)$$

其中 $play_{i,k}, like_{i,k}, repost_{i,k}, comment_{i,k}$ 分别表示这一作品收到的播放、点赞、转发和评论的次数。本文发现这些社交反馈的次数呈现出幂律分布,因此基于原始的数据构建了更丰富的特征,包括原始取值的对数、平均每次播放可以获得的点赞、转发、评论的次数。最终得到的作品的社交反馈特征 $r_{i,k}^f$ 如式(14)所示

$$\begin{aligned} r_{i,k}^f = & [play_{i,k}, like_{i,k}, repost_{i,k}, comment_{i,k}, \\ & \ln (play_{i,k} + 1), \\ & \ln (like_{i,k} + 1), \\ & \ln (repost_{i,k} + 1), \\ & \ln (comment_{i,k} + 1), \\ & \frac{like_{i,k}}{play_{i,k}}, \frac{repost_{i,k}}{play_{i,k}}, \frac{comment_{i,k}}{play_{i,k}}] \end{aligned} \quad (14)$$

给定用户 u_i 和歌曲 s_j ,用户是否会演唱歌曲 s_j 会受到偶像社区作品 R_i 的影响。令 R_i^u 表示发布这些作品的用户, R_i^s 表示作品相应的歌曲。由于这一社交反馈关系涉及到用户、歌曲、作品三个方面以及它们内部和之间的相互关联关系,相比于已有研究,这种关系具有更复杂、多层次的特点。因此,本文首先构建了一个三层异构网络 $G(O, \varepsilon)$ 来同时表示用户之间的关系、歌曲之间的关系以及作品之间的关系。 $G(O, \varepsilon)$ 的结构如图5所示,整个网络分为三层,第一层是用户层,每个节点都代表了 $R_i^u \cup \{u_i\}$ 中的一个用户,用户 u_i 与 R_i^u 中的用户相连;最底层是歌曲层,每个节点都代表了 $R_i^s \cup \{s_j\}$ 中的一首歌曲,歌曲 s_j 与 R_i^s 中的歌曲相连;中间层是作品层,每个节点都代表了 $R_i \cup \{(u_i, s_j)\}$ 中的一个作品,目标作品 (u_i, s_j) 与 R_i 中的作品相连。每层内部的边表示的是同一类型的节点之间的关系,边的权重表示了关联关

系的强弱,同时中间层的作品节点与上下层中该作品所对应的用户和歌曲相连。

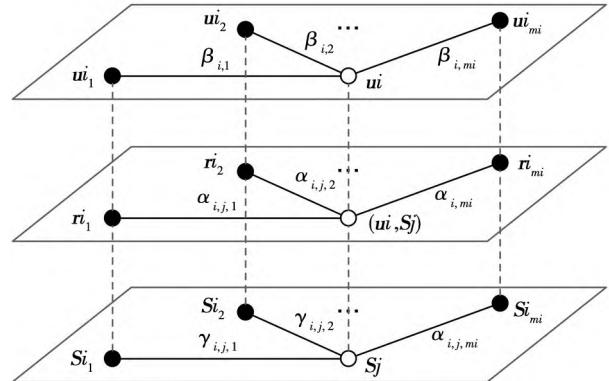


图5 三层异构网络示意图

Fig. 5 The tri-layer heterogeneous graph

随后,对这样一个三层异构网络,本文提出了三层异构图卷积网络THGCN模型来建模集合 R_i 中的作品对用户 u_i 演唱歌曲 s_j 的影响,模型结构如图6所示。

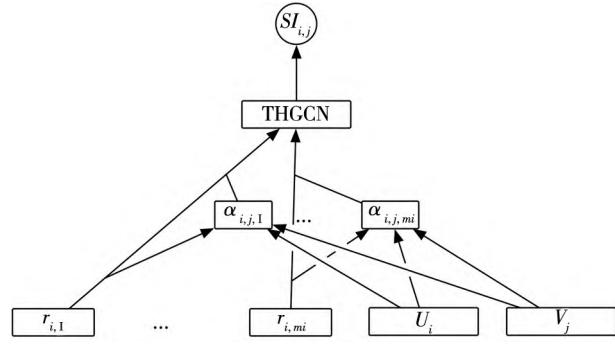


图6 THGCN 模型结构

Fig. 6 The architecture of THGCN

对一个作品 $r_{i,k} \in R_i$,本文使用式(15)~式(16)来分别计算用户 u_i 与 $r_{i,k}^u$ 之间的边的权重 $\beta_{i,k}$ 以及歌曲 s_j 与 $r_{i,k}^s$ 之间的边的权重 $\gamma_{j,i,k}$

$$\beta_{i,k} = \text{ReLU}(\mathbf{U}_{r_{i,k}^u}^\top S_U \mathbf{U}_i) \quad (15)$$

$$\gamma_{j,i,k} = \text{ReLU}(\mathbf{V}_{r_{i,k}^s}^\top S_V \mathbf{V}_j) \quad (16)$$

其中 \mathbf{U}_i 和 $\mathbf{U}_{r_{i,k}^u}^\top$ 分别表示用户的隐特征向量, \mathbf{V}_j 和 $\mathbf{V}_{r_{i,k}^s}^\top$ 分别表示歌曲的隐特征向量, S_U 和 S_V 都是待学习的参数。通过将这两个权重进行结合,得到作品 $r_{i,k}$ 与 (u_i, s_j) 之间的权重 $\alpha_{i,j,k}$,其计算如式(17)所示,本文使用这一权重表示作品 $r_{i,k}$ 对用户 u_i 发布歌曲 s_j 的影响程度

$$\alpha_{i,j,k} = \beta_{i,k} \times \gamma_{j,i,k} \quad (17)$$

以上权重的计算方式使得作品之间的边同时

包含了用户之间的关系和歌曲之间的关系,本文使用与图卷积网络(GCN)^[44]和图注意力网络(GAT)^[45]类似的原理来对作品的社交反馈信息进行聚合以得到社交反馈因素对用户 u_i 发布歌曲 s_j 的影响 $SI_{i,j}$

$$SI_{i,j} = \sigma \left(\sum_{r_{i,k} \in R_i} \alpha_{i,j,k} r_{i,k}^f W \right) \quad (18)$$

其中 W 是待学习的参数。在模型的实现中,本文选取用户 u_i 的偶像社区中最新发布的 K 个作品构成 R_i ,其中 K 为模型的超参数。

3.3 模型整体框架与训练方法

给定用户 u_i 和歌曲 s_j ,用户 u_i 会发布歌曲 s_j 的可能性 $R_{i,j}$ 的计算方式如下

$$R_{i,j} = PI_{i,j} + \tau_1 VC_{i,j} + \tau_2 SI_{i,j} \quad (19)$$

其中 τ_1 和 τ_2 是两个平衡参数,为了降低人为设定参数带来的额外工作和误差,这两个平衡参数的取值随模型的训练来得到, $PI_{i,j}$ 、 $VC_{i,j}$ 、 $SI_{i,j}$ 的计算分别如式(1)、式(10)、式(18)所示。图7展示了本文提出的VSRec模型的整体结构,○表示向量内积。

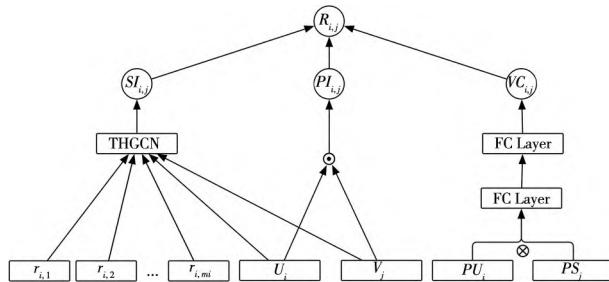


图 7 VSRec 模型整体结构

Fig. 7 The whole model VSRec

BPR 优化模型^[24]是常用的学习模型参数的方法。给定用户 u_i 和他演唱过的歌曲集合 S_i ,本文认为该用户对 S_i 中的歌曲的演唱偏好大于不在该集合中的其他歌曲,并通过采样的方法得到形如 $(u_i, s_j, s_{j'})$ 的三元组来构建训练集 D ,其中 $s_{j'}$ 为不在 S_i 中的歌曲

$$D = \{(u_i, s_j, s_{j'}) \mid u_i \in U, s_j \in S_i, s_{j'} \in S \setminus S_i\} \quad (20)$$

基于数据集 D ,通过最小化损失函数式(21)来得到最优的参数取值

$$L(D \mid \Theta) = \sum_{u_i, s_j, s_{j'} \in D} -\ln \sigma(\hat{R}_{i,j} - \hat{R}_{i,j'}) + \lambda_U \|U\|_F^2 + \lambda_V \|V\|_F^2 \quad (21)$$

其中 Θ 是模型所有参数的集合, $\sigma(\cdot)$ 是 Sigmoid 函数, λ_U 和 λ_V 是正则化超参数, $\|\cdot\|_F^2$ 是 Frobenius 范数。

4 实验结果与分析

为了验证模型的效果,本文设计了四组实验:1)比较了 VSRec 与其他推荐模型的效果来验证模型的推荐准确度;2)消融实验验证对用户演唱能力建模和社交反馈因素建模在推荐过程中起到的作用;3)衡量了对用户演唱能力的建模与标准的评估方法得到结果的一致性;4)对模型的超参数 K ,即选取的用户音乐社区的作品 R_i 的数量,进行了敏感性分析。

4.1 数据与实验设定

本文使用的数据来自某在线唱歌平台,训练数据集包括了 2018 年 2 月 9 日—2018 年 3 月 10 日这段时间内 97 293 名用户演唱 9 933 首歌曲的情况,共包括 1 593 554 条发布的作品,用户之间存在 2 350 901 对关注关系。对于每个作品,数据集还收集了截止至 2018 年 3 月 10 日作品收到的社交反馈情况。验证集和测试集则是收集了相关用户在 2018 年 3 月 10 日之后最新发布的 5 首作品,随机选取了 25% 的用户,将其关联数据作为验证集,其余部分作为测试集。实验部分所有模型的超参数调节都是在验证集上进行的,模型的评估则基于测试集。

在模型评估方面,本文首先对每个用户的推荐效果进行评估,然后对这些结果求均值作为模型的最终效果。考虑到对所有歌曲进行排序的计算开销较大,本文使用了文献[46]中的方法:随机采样 495 首用户在训练集、验证集和测试集中未演唱过的歌曲,与原验证集或测试集中的用户演唱的 5 首歌曲进行合并,依据模型的输出对这 500 首歌曲的排序和评估。

在评估指标方面,本文采用精准率 Precision@ k ,平均精准率 MAP(mean average precision) 和召回率 Recall@ k

$$\text{Precision}@k = \frac{1}{|U|} \sum_{u_i \in U} \frac{\sum_{l=1}^k r_{i,l}}{k} \quad (22)$$

$$MAP = \frac{1}{|U|} \sum_{u_i \in U} \frac{1}{t_i} \sum_{k=1}^{n_i} \frac{\sum_{l=1}^k r_{r_{i,l}}}{k} \times r_{r_{i,l}} \quad (23)$$

$$Recall@k = \frac{1}{|U|} \sum_{u_i \in U} \frac{\sum_{l=1}^k r_{r_{i,l}}}{t_i} \quad (24)$$

其中 n_i 表示为用户 u_i 生成的推荐列表的长度, t_i 表示推荐列表中用户演唱的歌曲的数量, $r_{r_{i,l}} \in \{0,1\}$ 表示为用户 u_i 生成的推荐列表中的第 l 首歌是否被该用户演唱.

表 2 模型对比结果
Table 2 Comparisons of the recommendation models

| Model | MAP | Rec@1 | Rec@10 | Pre@1 | Pre@10 |
|--------|----------------|----------------|----------------|----------------|----------------|
| BPR | 0.362 3 | 0.087 9 | 0.483 7 | 0.437 6 | 0.240 8 |
| PME | 0.323 5 | 0.084 5 | 0.428 1 | 0.420 7 | 0.213 1 |
| DO | 0.148 9 | 0.075 1 | 0.134 9 | 0.374 0 | 0.067 2 |
| SBPR | 0.347 9 | 0.085 5 | 0.467 8 | 0.425 4 | 0.232 9 |
| OLA | <u>0.372 1</u> | <u>0.089 9</u> | <u>0.495 3</u> | <u>0.447 4</u> | <u>0.246 5</u> |
| DSCF | 0.299 8 | 0.074 6 | 0.403 7 | 0.371 5 | 0.200 9 |
| DANSER | 0.331 2 | 0.077 0 | 0.451 9 | 0.405 7 | 0.231 1 |
| VSRec | 0.407 3 | 0.105 7 | 0.525 7 | 0.526 8 | 0.267 8 |
| Impv. | 9.44 % | 17.56 % | 6.15 % | 17.75 % | 9.04 % |

5) OLA^[20]: 社会化推荐模型, 只考虑用户的部分好友对用户的影响.

6) DSCF^[37]: 社会化推荐模型, 同时考虑用户的直接和间接好友对用户的影响.

7) DANSER^[42]: 社会化推荐模型, 使用图神经网络提取社交网络信息.

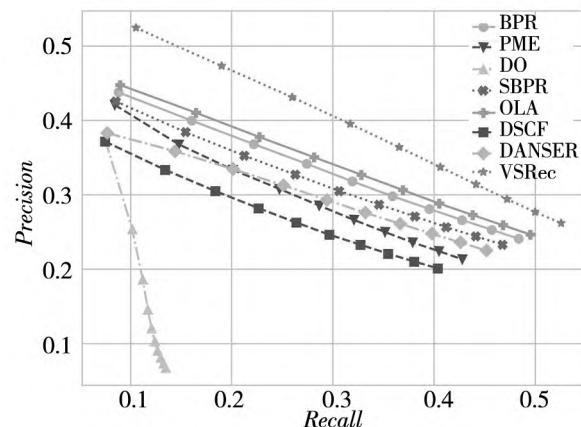


图 8 Precision@k 和 Recall@k 曲线, 其中 k 分别取 [1,10] 内的整数

Fig. 8 Precision@k vs Recall@k curve, k ranges from 1

to 10 from the left to the right

以上所有模型的超参数均选取在验证集上效

4.2 模型效果对比

本文选取了以下几个模型作为对比模型:

1) BPR^[24]: 使用 BPR 优化方法的矩阵分解模型.

2) PME^[12]: 唱歌推荐模型, 未考虑用户演唱能力.

3) DO^[1]: 唱歌推荐模型, 考虑了歌曲演唱难度和用户演唱能力.

4) SBPR^[21]: 社会化推荐模型, 在构建物品间的偏序关系时考虑用户好友的偏好.

果最好的取值. 表 2 列出了 VSRec 与对比模型在测试集上的评估结果, 对比模型的最好结果用下划线标识, 最后一行展示了 VSRec 相对于其他模型最好效果的提升, 图 8 展示了精准率和召回率曲线. 结果显示:

1) VSRec 模型在所有模型中取得了最好的推荐效果, 这说明 VSRec 在解决在线唱歌平台的推荐问题上更有效.

2) VSRec 模型在效果上优于 PME 和 DO 这两个模型. 其中 PME 模型仅考虑用户对歌曲的喜爱程度而忽略了用户演唱能力和社交反馈因素的影响, DO 模型仅考虑用户的演唱能力而忽略了用户对歌曲的喜爱程度和社交反馈因素的影响. 从结果来看, DO 模型效果最差, 这说明只考虑用户演唱能力是远远不够的, 而 VSRec 比 PME 和 DO 都好则证明了同时考虑这三方面因素的必要性.

3) 除了 VSRec 模型以外, 在其余几个社会化推荐模型中, 只有 OLA 模型的效果优于通用的 BPR 模型. OLA 模型只考虑了用户一部分好友的

影响,DSCF 和 DANSER 则考虑了所有好友甚至间接好友的影响,本文推测其效果不好是因为这些多余的好友信息属于噪音信息. VSRec 只考虑用户的好友最新发布的部分作品这种做法也对这一噪音进行了有效的过滤.

4.3 消融实验

为了验证对用户演唱能力和社交反馈因素的建模在唱歌推荐问题中的作用,本文构建了两个模型进行消融实验: VRec 模型去掉了 VSRec 中对社交反馈因素建模的部分; SRec 模型去掉了 VSRec 中对用户演唱能力建模的部分. 同时去掉 VSRec 模型中的用户演唱能力部分和社交反馈因素部分后模型退化为 BPR 模型,因此消融实验比较了这四个模型,实验结果如表 3 和图 9 所示.

从表 3 和图 9 可以看到,VRec 模型和 SRec 模型都比 BPR 模型的效果好,这一点证明了考虑用户的演唱能力和社交反馈因素对于提升推荐效

果都是有作用的,并且相比于用户的演唱能力,考虑社交反馈因素带来的提升更大. 另外,VSRec 模型比 VRec 模型和 SRec 模型效果都更好,证明了同时考虑两方面因素进行推荐是更有效的.

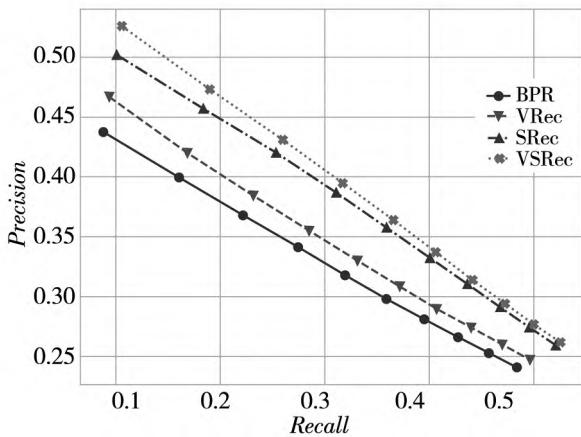


图 9 消融实验的 $Precision@k$ 和 $Recall@k$ 曲线

Fig. 9 $Precision@k$ vs $Recall@k$ curve of ablation study

表 3 消融实验结果

Fig. 3 The results of the ablation study

| Model | MAP | Rec@1 | Rec@10 | Pre@1 | Pre@10 |
|--------------|---------|---------|---------|---------|---------|
| BPR | 0.362 3 | 0.087 9 | 0.483 7 | 0.437 6 | 0.240 8 |
| VRec | 0.375 8 | 0.093 8 | 0.496 3 | 0.472 3 | 0.247 0 |
| <i>Impv.</i> | 3.74 % | 6.66 % | 2.61 % | 7.93 % | 2.57 % |
| SRec | 0.399 3 | 0.100 9 | 0.521 0 | 0.498 6 | 0.253 1 |
| <i>Impv.</i> | 10.22 % | 14.71 % | 7.71 % | 13.94 % | 5.11 % |
| VSRec | 0.407 3 | 0.105 7 | 0.525 7 | 0.526 8 | 0.267 8 |
| <i>Impv.</i> | 12.43 % | 20.19 % | 8.69 % | 20.38 % | 11.21 % |

表 4 子模块与其他模型的对比

Fig. 4 Comparison with competitive models

| Model | MAP | Rec@1 | Rec@10 | Pre@1 | Pre@10 |
|----------------------|---------|---------|---------|---------|---------|
| DO | 0.148 9 | 0.075 1 | 0.134 9 | 0.374 0 | 0.067 2 |
| OLA | 0.372 1 | 0.089 9 | 0.495 3 | 0.447 4 | 0.246 5 |
| VRec | 0.375 8 | 0.093 8 | 0.496 3 | 0.472 3 | 0.247 0 |
| <i>Impv. overDO</i> | 152 % | 24.9 % | 268 % | 28.3 % | 268 % |
| SRec | 0.399 3 | 0.100 9 | 0.521 0 | 0.498 6 | 0.253 1 |
| <i>Impv. overOLA</i> | 3.71 % | 12.2 % | 5.19 % | 11.44 % | 2.68 % |
| VSRec | 0.407 3 | 0.105 7 | 0.525 7 | 0.526 8 | 0.267 8 |
| <i>Impv. overOLA</i> | 9.44 % | 17.56 % | 6.15 % | 17.75 % | 9.04 % |

本文还对 VRec 模型与考虑了用户演唱能力的 DO 模型进行了比较,对 SRec 模型与考虑了社交因素中表现最好的 OLA 模型进行了比较,以说明本文所提出的对用户演唱能力和社交反馈因素建模的有效性. 从表 4 的结果可以看到本文对演唱能力的建模方法和使用 THGCN 来建模社交反

馈因素的方法均优于已有模型.

4.4 对用户演唱能力建模的准确性

用户在唱歌平台上演唱一首歌曲之后,平台会针对用户的演唱结果进行打分,本文对模型得到的用户演唱能力的准确性进行了评估. 在在线唱歌平台上,平台会对用户的演唱结果给出 A ++ 、

A+、A、B、C 和 D 六个等级, 通过比较这些评级, 可以得到用户演唱不同歌曲的能力。而本文使用式(10)对用户的演唱能力进行建模, 可以通过比较 $VC_{i,j}$ 的相对大小来得到用户对不同歌曲的演唱能力高低。通过比较模型结果与真实的评级关系的一致性, 可以得到模型对用户演唱能力建模的准确性。

表 5 演唱评级对应评分表

Table 5 Performance level mapping

| 演唱结果评级 | 演唱能力评分 |
|--------|--------|
| A++ | 100 |
| A+ | 90 |
| A | 80 |
| B | 60 |
| C | 50 |
| D | 30 |

本文还尝试了使用 KNN 模型进行用户演唱能力评估: 给定用户 u_i 和歌曲 s_j , 找到用户 u_i 在训练阶段唱过的前 k 首与歌曲 s_j 最相似的歌曲, 相似度为式(4)所得到的歌曲声学模型的余弦相似度, 将这 k 首歌曲的演唱评级使用表 5 所示的对应关系将等级映射到得分后, 取它们的均值作为预测的用户 u_i 对歌曲 s_j 的演唱能力评分。

对测试集的数据, 使用式(10)和 KNN 分别得到预测的用户演唱能力得分后, 根据得分对歌曲进行排序, 使用 NDCG (normalized discounted cumulative gain) 来评估预测结果与真实排序的一致性, 其计算如式(25)~式(26)所示

$$DCG = \sum_{i=1}^N \frac{2^{rel_i} - 1}{\ln(i + 1)} \quad (25)$$

$$NDCG = \frac{DCG}{IDCG} \quad (26)$$

其中 N 为测试集中演唱记录的数量, rel_i 是根据预测分数排序后第 i 个作品的评级对应的相关性数值: 6 个评级由低到高分别对应 1~6 这 6 个取值。IDCG 为 DCG 最大的可能取值。

表 6 展示了实验结果, 从结果可以看到本文提出的建模用户演唱能力的方法与真实的结果一致性较高, 且相比于使用 KNN 模型能够取得更好的效果。

表 6 用户演唱能力建模的准确性

Table 6 Evaluation of vocal ability

| Model | NDCG |
|---------------|---------|
| VC | 0.961 0 |
| KNN, $k = 1$ | 0.943 6 |
| KNN, $k = 5$ | 0.932 4 |
| KNN, $k = 10$ | 0.930 4 |

4.5 超参数敏感性分析

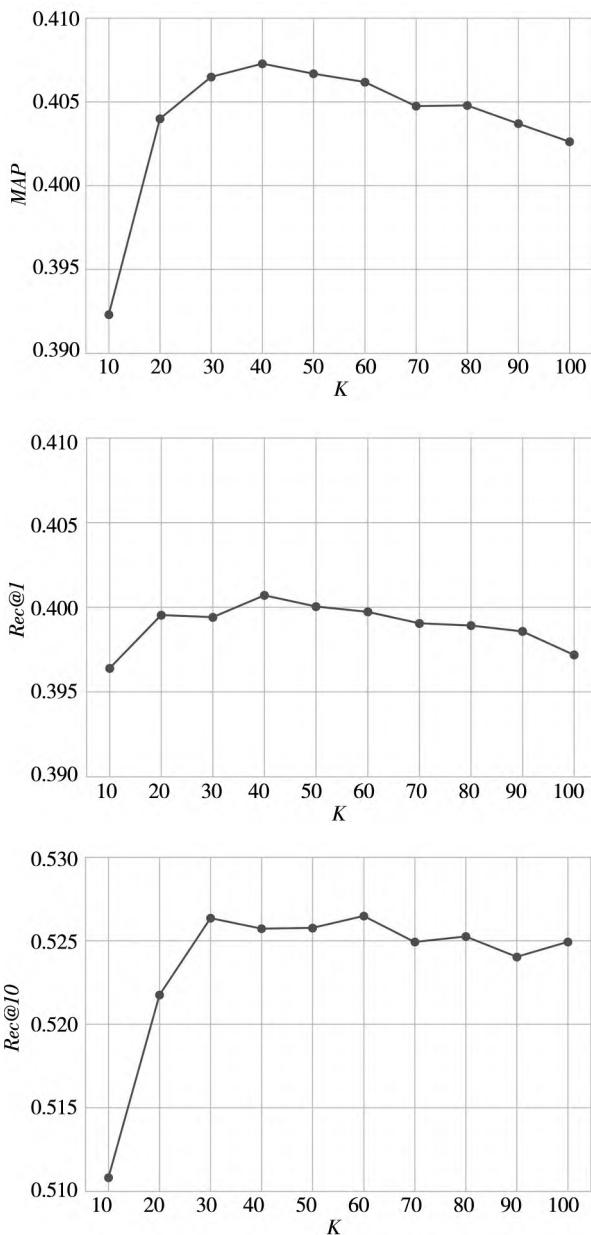


图 10 超参数的灵敏度分析结果

Fig. 10 Parameter sensitivity study

在对用户的社交反馈因素进行建模时, 选取

了用户的偶像社区最新的 K 首歌曲。为了验证模型对这一超参数的敏感性,本实验从 $10 \sim 100$ 以 10 为步长选取了不同的 K 值进行实验。图 10 展示了 K 取不同的值时模型的推荐效果。

从结果可以看到,当 $K \leq 40$ 时,模型的效果随着 K 的增加而上升;当 K 取值在 $40 \sim 70$ 之间时,模型的效果几乎不变;当 $K > 70$ 时,模型的效果会稍有下降。出现这种结果的原因可能是:一个用户的注意力是有限的,用户往往只能关注到其偶像社区最新的部分作品。但总体而言,模型的效果受 K 取值的影响不大。

5 结束语

本文聚焦于在线唱歌平台上的演唱歌曲推荐方法研究。现有方法虽然考虑了用户的演唱能力,但是所提方法需要人工干预或其他较难获取的资源,导致适用性差。为此,提出了一种新的用户演唱能力建模方法,该方法仅仅依赖于基础的数据且可实现自动化的评估和建模,因此有更强的实用性。除此之外,本文还发现用户的好友发布的作品及其收到的社交反馈信息与用户的演唱行为存在较强的相关性,但现有方法忽略了社交反馈信息对推荐的作用。本文提出构建一种新的建模方法,设计了一个三层异构网络来表示用户、歌曲、作品三者之间的关系,并提出一个新的图神经网络模型 THGCN 从该网络中提取特征。综合用户对歌曲的喜爱程度,本文提出了面向在线唱歌平

台的演唱歌曲推荐模型 VSRec。

为了验证 VSRec 模型的效果,本文在真实的数据集上进行了实验。首先通过与其他具有代表性的一些推荐模型的比较,显示 VSRec 模型在演唱歌曲推荐这一任务上具有更好的推荐效果。其次,还通过消融实验证明了考虑用户的演唱能力和社交反馈因素均能提升最终的推荐效果,且本文所提方法对这两部分因素的建模相比于已有方法更有效。同时结果还表明,相较于唱歌能力,社交反馈因素的作用更大,这对于平台运营也有指导意义,例如可提供更丰富的社交反馈方式以提升用户活跃度等。除此之外,本文还通过实验评估了所提出的演唱能力评估方法的有效性。最后,对模型的超参数的敏感性进行分析,分析结果同时显示用户行为往往只会受到最新的社交反馈信息的影响,这对于平台的机制设计和后续研究的关注点都具有一定的意义。

本文也存在一些不足,为后续的研究提供了一些可能的研究方向。在进行用户演唱能力建模的时候,本文使用相邻两个音符之间的关系构建了演唱情景来构建歌曲和用户的声学模型,然而更高阶的关系可能带来更多的提升,未来可以进一步在这个方向上进行研究。在本文对社交反馈信息的处理中,仅仅用到了社交反馈信息的播放、点赞等行为的数量,然而社交反馈信息还包含了更丰富内容。例如社交反馈来源的用户不同,对该用户的影响也可能是不同的,本文对这一因素的影响并没有做更深入的挖掘,这也是未来的研究中可以关注的方向。

参 考 文 献:

- [1] 华中生. 网络环境下的平台服务及其管理问题[J]. 管理科学学报, 2013, 16(12): 1–12.
Hua Zhongsheng. Platform service and its management problems in the network environment[J]. Journal of Management Sciences in China, 2013, 16(12): 1–12. (in Chinese)
- [2] Mao K, Fan J, Shou L, et al. Song Recommendation for Social Singing Community[C]. Proceedings of the 22nd ACM International Conference on Multimedia, Association for Computing Machinery, New York: 2014: 127–136.
- [3] Guan C, Fu Y, Lu X, et al. Vocal Competence Based Karaoke Recommendation: A Maximum-Margin Joint Model[C]. Proceedings of the 2016 SIAM International Conference on Data Mining, Society for Industrial and Applied Mathematics, Miami, Florida: 2016: 135–143.
- [4] Guan C, Fu Y, Lu X, et al. Efficient Karaoke song recommendation via multiple kernel learning approximation[J]. Neurocomputing, 2017, (254): 22–32.
- [5] Mao K, Luo X, Chen K, et al. My DJ: Recommending Karaoke Songs from One's Own Voice[C]. Proceedings of The 35th International ACM SIGIR Conference on Research and Development in Information Retrieval, Association for Computing Ma-

- achinery, New York: 2012: 1009 – 1009.
- [6] Mao K, Shou L, Fan J, et al. Competence-based song recommendation: Matching songs to one's singing skill[J]. IEEE Transactions on Multimedia, 2015, 17(3): 396 – 408.
- [7] 李永立, 陈杨, 樊宁远, 等. 考虑个体效用因素的社会网络演化分析模型[J]. 管理科学学报, 2018, 21(3): 41 – 53.
LI Yongli, Chen Yang, Fan Ningyuan, et al. Utility-based model for interpreting evolution patterns of social networks[J]. Journal of Management Sciences in China, 2018, 21(3): 41 – 53. (in Chinese)
- [8] Cialdini R B, Goldstein N J. Social influence: Compliance and conformity[J]. Annu. Rev. Psychol, 2004, (55): 591 – 621.
- [9] 李博, 陈志刚, 黄瑞, 等. 基于LDA模型的音乐推荐算法[J]. 计算机工程, 2016, 42(6): 175 – 179, 184.
Li Bo, Chen Zhigang, Huang Rui, et al. Music recommendation algorithm based on LDA model[J]. Computer Engineering, 2016, 42(6): 175 – 179, 184. (in Chinese)
- [10] 李琳, 唐守廉. 基于多层注意力表示的音乐推荐模型[J]. 电子学报, 2020, 48(9): 1672 – 1679.
Li Lin, Tang Shoulian. Hierarchical attention representation model for music recommendation[J]. Acta Electronica Sinica, 2020, 48(9): 1672 – 1679. (in Chinese)
- [11] 张全贵, 张新新, 李志强. 基于注意力机制的音乐深度推荐算法[J]. 计算机应用研究, 2019, 36(8): 2297 – 2299.
Zhang Quangui, Zhang Xinxin, Li Zhiqiang. Music recommendation algorithm based on attention mechanism[J]. Application Research of Computers, 2019, 36(08): 2297 – 2299. (in Chinese)
- [12] Wu X, Liu Q, Chen E, et al. Personalized Next-Song Recommendation in Online Karaoke[C]. Proceedings of the 7th ACM Conference on Recommender Systems, Association for Computing Machinery, New York: 2013: 137 – 140.
- [13] 金晓玲, 周中允, 尹梦杰, 等. 在线用户点赞与评论行为的产生机理差异研究——以医疗健康类企业微信公众号为例[J]. 管理科学学报, 2021, 24(4): 54 – 68.
Jin Xiaoling, Zhou Zhongyun, Yin Mengjie, et al. Understanding antecedent differences across online users' like and comment behaviors: The case of healthcare enterprise WeChat public platform[J]. Journal of Management Sciences in China, 2021, 24(4): 54 – 68. (in Chinese)
- [14] 王刚, 蒋军, 王含茹. 社会化推荐研究综述[J]. 计算机科学, 2018, 45(S2): 37 – 42, 62.
Wang Gang, Jiang Jun, Wang Hanru. Review of social recommendation[J]. Computer Science, 2018, 45(S2): 37 – 42, 62. (in Chinese)
- [15] 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 Xiangwu, Liu Shudong, Zhang Yujie, et al. Research on social recommender systems[J]. Journal of Software, 2015, 26(6): 1356 – 1372. (in Chinese)
- [16] Jamali M, Ester M. A Matrix Factorization Technique with Trust Propagation for Recommendation in Social Networks[C]. Proceedings of the fourth ACM conference on Recommender systems, Association for Computing Machinery, New York: 2010: 135 – 142.
- [17] Ma H, Zhou D, Liu C, et al. Recommender Systems with Social Regularization[C]. Proceedings of the Fourth ACM International Conference on Web Search and Data Mining, Association for Computing Machinery, New York: 2011: 287 – 296.
- [18] Yang X, Steck H, Liu Y. Circle-Based Recommendation in Online Social Networks[C]. Proceedings of The 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Association for Computing Machinery, New York: 2012: 1267 – 1275.
- [19] Wang X, Hoi S C H, Ester M, et al. Learning Personalized Preference of Strong and Weak Ties for Social Recommendation[C]. Proceedings of The 26th International Conference on World Wide Web, International World Wide Web Conferences Steering Committee, Geneva: 2017: 1601 – 1610.
- [20] Wang X, Zhu W, Liu C. Social Recommendation with Optimal Limited Attention[C]. Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, Association for Computing Machinery, New York: 2019: 1518 – 1527.
- [21] Zhao T, McAuley J, King I. Leveraging Social Connections to Improve Personalized Ranking for Collaborative Filtering[C]. Proceedings of The 23rd ACM International Conference on Conference on Information and Knowledge Management, Association for Computing Machinery, New York: 2014: 261 – 270.

- [22] Rafailidis D, Crestani F. Joint Collaborative Ranking with Social Relationships in Top-N Recommendation [C]. Proceedings of The 25th ACM International Conference on Information and Knowledge Management. Association for Computing Machinery, New York: 2016: 1393 – 1402.
- [23] Rafailidis D, Crestani F. Learning to Rank with Trust and Distrust in Recommender Systems [C]. Proceedings of the Eleventh ACM Conference on Recommender Systems, Association for Computing Machinery, New York: 2017: 5 – 13.
- [24] Rendle S, Freudenthaler C, Gantner Z, et al. BPR: Bayesian Personalized Ranking from Implicit Feedback [C]. Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence, AUAI Press, Arlington, Virginia: 2009: 452 – 461.
- [25] Ma H, Yang H, Lyu M R, et al. Sorec: Social Recommendation Using Probabilistic Matrix Factorization [C]. Proceedings of the 17th ACM Conference on Information and Knowledge Management, Association for Computing Machinery, New York: 2008: 931 – 940.
- [26] Tang J, Hu X, Gao H, et al. Exploiting Local and Global Social Context for Recommendation [C]. Twenty-Third International Joint Conference on Artificial Intelligence, AAAI Press, Washington: 2013: 2712 – 2718.
- [27] Guo G, Zhang J, Yorke-Smith N. Trustsvd: Collaborative Filtering with Both the RxplicitAnd Implicit Influence of User Trust And of Item Ratings [C]. Proceedings of the AAAI Conference on Artificial Intelligence, AAAI Press, Washington, DC, USA, 2015, 29(1).
- [28] Fan W, Derr T, Ma Y, et al. Deep Adversarial Social Recommendation [C]. 28th International Joint Conference on Artificial Intelligence, IJCAI, Darmstadt: 2019: 1351 – 1357.
- [29] Feng C, Liang J, Song P, et al. A fusion collaborative filtering method for sparse data in recommender systems [EB/OL]. Information Sciences, 2020, 521: 365 – 379. DOI: 10.1016/j.ins.2020.02.052.
- [30] Rafailidis D, Crestani F. Recommendation with Social Relationships via Deep Learning [C]. Proceedings of the ACM SIGIR International Conference on Theory of Information Retrieval, Association for Computing Machinery, New York: 2017: 151 – 158.
- [31] Hu Z, Wang J, Yan Y, et al. Neural graph personalized ranking Top-N recommendation [J]. Knowledge-Based Systems, 2020, 213(8): 106426.
- [32] Wu L, Sun P, Hong R, et al. Collaborative neural social recommendation [J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2021, 51(1): 464 – 476.
- [33] Ma H, King I, Lyu M R. Learning to Recommend with Social Trust Ensemble [C]. Proceedings of the 32nd International ACM SIGIR Conference on Research And Development in Information Retrieval, Association for Computing Machinery, New York: 2009: 203 – 210.
- [34] Chaney A J B, Blei D M, Eliassi-Rad T. A Probabilistic Model for Using Social Networks in Personalized Item Recommendation [C]. Proceedings of the 9th ACM Conference on Recommender Systems, Association for Computing Machinery, New York: 2015: 43 – 50.
- [35] Sun P, Wu L, Wang M. Attentive Recurrent Social Recommendation [C]. The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval, Association for Computing Machinery, New York: 2018: 185 – 194.
- [36] Chen J, Feng Y, Ester M, et al. Modeling Users' Exposure with Social Knowledge Influence and Consumption Influence for Recommendation [C]. Proceedings of the 27th ACM International Conference on Information and Knowledge Management, Association for Computing Machinery, New York: 2018: 953 – 962.
- [37] Fan W, Ma Y, Yin D, et al. Deep Social Collaborative Filtering [C]. Proceedings of the 13th ACM Conference on Recommender Systems, Association for Computing Machinery, New York: 2019: 305 – 313.
- [38] 吴 博, 梁 循, 张树森, 等. 图神经网络前沿进展与应用 [J]. 计算机学报, 2022, 45(1): 35 – 68.
Wu Bo, Liang Xun, Zhang Shusen, et al. Advances and applications in graph neural network [J]. Chinese Journal of Computers, 2022, 45(1): 35 – 68. (in Chinese)
- [39] Wu Z, Pan S, Chen F, et al. A comprehensive survey on graph neural networks [J]. IEEE Transactions on Neural Networks and Learning Systems, 2020, 32(1): 4 – 24.
- [40] Wu L, Sun P, Fu Y, et al. A Neural Influence Diffusion Model for Social Recommendation [C]. Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval, Association for Computing Machinery, New York: 2019: 235 – 244.

- [41] 何云飞, 张以文, 吕智慧, 等. 异质信息网络中元路径感知的评分协同过滤[J]. 计算机学报, 2020, 43(12): 2385–2397.
He Yunfei, Zhang Yiwen, Lü Zhihui, et al. Meta path-aware rating collaborative filtering in heterogeneous information network[J]. Chinese Journal of Computers, 2020, 43(12): 2385–2397. (in Chinese)
- [42] Wu Q, Zhang H, Gao X, et al. Dual Graph Attention Networks for Deep Latent Representation of Multifaceted Social Effects in Recommender Systems[C]. The World Wide Web Conference, Association for Computing Machinery, New York; 2019: 2091–2102.
- [43] Nair V, Hinton G E. Rectified Linear Units Improve Restricted Boltzmann Machines[C]. Proceedings of the 27th International Conference on International Conference on Machine Learning, Omnipress, Madison; 2010: 807–814.
- [44] Kipf T N, Welling M. Semi-Supervised Classification with Graph Convolutional Networks[C]. International Conference on Learning Representations, Ithaca; 2017.
- [45] Veličković P, Cucurull G, Casanova A, et al. Graph Attention Networks[C]. International Conference on Learning Representations, Ithaca; 2018.
- [46] Cheng W, Shen Y, Zhu Y, et al. DELF: A Dual-Embedding Based Deep Latent Factor Model for Recommendation[C]. Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, IJCAI, Darmstadt; 2018: 3329–3335.

A song recommendation method considering vocal ability and social feedback factors

CHEN Zhuo-hua^{1, 2}, LIU Hong-yan^{1, 2*}, GAO Ge³

1. Research Center for Contemporary Management, Tsinghua University, Beijing 100084, China;
2. School of Economics and Management, Tsinghua University, Beijing 100084, China;
3. School of Business Administration, Southwestern University of Finance and Economics, Chengdu 611074, China

Abstract: Online singing platforms have attracted many users to sing and publish songs online. Accurate singing-song recommendation systems are essential to enhance user experience and stickiness for such platforms. User's vocal ability is one of the major factors to consider in the recommendation model. However, how to predict a user's vocal ability is challenging. Meanwhile, online singing platforms provide social networking services. How to utilize the social feedback information to improve recommendation performance is worth studying. To address these issues, this paper proposes new methods to predict a user's vocal ability on a song and to model social feedback factors. Further, a novel model for singing-song recommendations is developed. Experiments conducted on a real-world dataset demonstrate the effectiveness of our proposed model and the necessity of modeling vocal ability and social feedback factors to improve recommendation performance.

Key words: singing-song recommendations; vocal ability; social network; deep learning