

# 基于时尚图谱增强的个性化互补服装推荐

史金婉, 宋雪萌, 刘子鑫, 聂礼强

山东大学计算机科学与技术学院 青岛 中国 266237

**摘要** 近年来,人们对时尚穿搭有了更高的美学追求。因此,个性化互补服装推荐,即为用户推荐与他/她已购买服装相匹配的互补服装,逐渐吸引了学术界的广泛关注。个性化互补服装推荐不同于一般的推荐任务(如电影推荐),它推荐的服装需要满足两个条件:1)与目标服装搭配;2)满足用户偏好。因此,相关的现有方法主要是基于时尚单品的多模态数据,着力于建模单品与单品之间的兼容性交互和用户与单品之间的偏好交互,以实现个性化互补服装的推荐。这些方法的缺点主要在于它们将每一个单品-单品交互或者用户-单品交互看作一个独立的数据实例,而忽略了单品的属性知识以及时尚实体(即,用户、单品及属性)之间的高阶交互关系。事实上,与一个单品(如,上衣)搭配的所有互补单品(如,下衣)可能会共享某些相同的属性(如,颜色);同时,具有类似品味的用户也可能倾向于选择具有类似属性特征的单品。显然,这些时尚实体之间的高阶关系蕴含了丰富的有关单品兼容性和用户偏好的协同信号,因而能够促进个性化互补服装推荐模型性能的提升。据此,本文构建了一个大规模协同时尚图谱,并基于图卷积神经网络(Graph Neural Networks, GNNs)来探索时尚实体之间的高阶关系,进而更好地实现个性化的互补服装推荐。具体地,本文提出了一个新颖的基于时尚图谱增强的个性化互补服装推荐模型(Fashion Graph-enhanced Personalized Complementary Clothing Recommendation),简称为FG-PCCR。FG-PCCR由两个关键的部分组成:独立的一阶交互建模和协同的高阶交互建模。一方面,独立的一阶交互建模模块基于视觉和文本模态数据,致力于通过神经网络和矩阵分解方法分别对单品-单品搭配交互和用户-单品偏好交互进行综合性建模。另一方面,协同的高阶交互模块基于构建的协同的时尚图谱,通过图神经网络利用信息传播机制来提取高阶的协同信号,进一步丰富用户和单品的向量表示。FG-PCCR模型能够有效整合时尚实体之间的复杂的高阶关系信息,用户和单品的表示学习,进而改进个性化互补推荐的效果。最后,对于给定的用户和目标上衣,我们能够得到推荐的下衣的个性化兼容性分数。另外,在真实的时尚数据集上做的大量实验,充分地验证了本文所提模型FG-PCCR相对于基准方法的优越性。

**关键词** 个性化服装推荐; 时尚兼容性建模; 图神经网络

中图法分类号 TP391.3 DOI号 10.19363/J.cnki.cn10-1380/tn.2021.09.14

## Fashion Graph-enhanced Personalized Complementary Clothing Recommendation

SHI Jinwan, SONG Xuemeng, LIU Zixin, NIE Liqiang

School of Computer Science and Technology, Shandong University, Qingdao 266237, China

**Abstract** Recent years, people have a higher aesthetic pursuit of fashion clothing matching. Therefore, personalized complementary clothing recommendation, that is, to recommend complementary clothes to users for matching their existing garments, has attracted increasing research attention. Different from the general recommendation tasks (e.g., movie recommendation), the recommended items in the context of personalized complementary clothing recommendation must meet two requirements: 1) match with the given item and 2) satisfy with the user's preference. Accordingly, existing research efforts mainly focus on modeling the item-item compatibility interaction and the user-item preference interaction based on the multi-modal data of fashion items to recommend items. One deficiency is that they regard each item-item or user-item interaction as an independent data instance, while overlooking the attribute information of items and the high-order relations among fashion entities, i.e., users, items, and attributes. In fact, items (e.g., bottoms) that go well with one item (e.g., a top) are more likely to share certain underlying attribute patterns (e.g., color), while users with similar tastes tend to choose items with similar attributes. Obviously, these high-order relations among fashion entities convey much implicit collaborative signals towards the item compatibility modeling and user preference modeling, and are critical for personalized complementary clothing recommendation. In light of this, in this work, we build a large-scale collaborative fashion graph to investigate the utility of high-order relations among fashion entities based on the Graph Convolutional Neural Networks(GNNs), in the context of personalized complementary clothing recommendation. In particular, we propose a new fashion graph-enhanced personalized complementary clothing recommendation model, dubbed as FG-PCCR, which consists of two key components: the independent one-order interaction modeling and the collaborative

通讯作者: 宋雪萌, 博士, 副教授, Email: songxuemeng@gmail.com。

本课题得到 CCF-百度松果基金资助。

收稿日期: 2021-04-28; 修改日期: 2021-08-09; 定稿日期: 2021-08-10

high-order interaction modeling. On the one hand, the independent one-order interaction modeling module, based on visual and textual modal data, is dedicated to the comprehensive modeling of item-item matching interaction and user-item preference interaction respectively by using neural network and matrix decomposition methods. On the other hand, the collaborative high-order interaction module is based on the constructed collaborative fashion graph, and distills the higher-order collaborative signals through the Graph Neural Networks using the information transmission mechanism to further enrich the vector representation of users and items. The FG-PCCR can effectively integrate the complex high-order relational information between fashion entities, the representation learning of users and items, and then improve the modeling effect of personalized clothing compatibility. Finally, for a given user and target top, the personalized compatibility score of the matching undergarment is obtained. In addition, extensive experiments on the real-world dataset have demonstrated the superiority of the proposed scheme over the state-of-the-art methods.

**Key words** personalized fashion recommendation; fashion compatibility modeling; graph neural networks

## 1 引言

近年来,随着时尚电商平台的繁荣发展,越来越多的用户选择在线购买服装。然而,享受在线购物便利的同时,用户常常会面临无法在海量服装市场中找到心仪单品的难题。由于人们常常需要为己购买的单品(如,上衣)挑选与之相搭配的互补单品(如,下衣)。因此,个性化互补服装推荐,即为用户推荐与他/她已购买服装相匹配的互补服装,逐渐吸引了诸多学者的关注。然而,对于同一件时尚单品,不同的人可能会给其搭配不同的互补单品。如图 1 所示,对于同一件黄色无袖衬衫来说,第一个用户将它与一条白色阔腿五分裤,一对绿色宝石耳环和一双红色高跟鞋进行搭配;而第二个用户将其与一条黄灰色碎花短裙,一对珍珠耳环和一双绿色凉鞋搭配在一起。这其实反映了一个服装搭配领域中的个性化偏好问题。因此,为了促进电商平台的潜在订单交易,和用户的在线购物体验,亟需开展有关个性化互补服装推荐的研究(personalized complementary clothing recommendation)。该任务的根本问题在于实现个性化的互补服装兼容性建模,即对于给定的用户和上衣,评价下衣的兼容性(匹配程度)。



图 1 IQON 社区平台中用户的套装搭配实例图

Figure 1 Examples of users' outfit compositions on IQON

目前,关于互补服装兼容性建模的研究工作主

要围绕一般性的兼容性建模,而忽略了用户的个性化偏好问题。尽管少数围绕个性化互补服装兼容性建模的研究<sup>[1-2]</sup>已经展开,这些研究工作主要是聚焦于建模一阶的单品-单品兼容性交互以及一阶的用户-单品偏好交互来预测个性化服装的兼容性评分。具体来说,宋等人<sup>[2]</sup>提出了一个新颖的个性化兼容建模框架(GP-BPR)。该框架从视觉和文本两个方面将用户-单品偏好建模整合在一般的单品-单品兼容性建模中。董等人<sup>[1]</sup>通过对单品和用户建模引入了一个新颖的个性化胶囊衣橱创建的方法,来给特定用户进行服装搭配。虽然他们在个性化互补服装推荐中取得了令人钦佩的进步,然而,时尚领域中的实体(如,用户、单品和属性)之间的高阶交互关系(high-order interaction relation)却被他们所忽略。比如,与同一个上衣(top)搭配的下衣(bottom)之间可能共享一些相同的属性(attribute);有着近似品味的用户也可能倾向于选择有相同属性的单品。

事实上,在相关领域中,最近的研究者们已经开始探索时尚实体之间的高阶关系,并提出了 DREP<sup>[3]</sup>和 HFGN<sup>[4]</sup>。具体来说,DREP<sup>[3]</sup>将单品之间的额外连接关系注入到成对的兼容性建模中,进一步增强了对单品的表示学习。值得注意的是,单品之间的这种额外连接关系信息的应用是有限的,只能被用来建模单品-单品之间的兼容性,无法促进对用户偏好的学习。类似于 HFGN<sup>[4]</sup>,DREP<sup>[3]</sup>在其构建的关系型时尚图谱中忽略了单品丰富的属性知识。而单品的属性知识在个性化互补服装推荐中起着至关重要的作用。在一定程度上来说,时尚单品之间的兼容性取决于其属性之间的和谐性,用户对于单品的偏爱程度也依赖于用户对单品属性的喜爱程度。比如,在图 1 中,对于第一个用户所搭配的第四套服装,其上衣和下衣之所以能够搭配兼容主要取决于黑白颜色之间的兼容性。同时,该用户在不同套装中均选择了高跟鞋,这极大可能是因为该用户喜欢高跟这个单品属性。总而言之,虽然时尚实体之间的高阶关

系已经被相关工作所探索,但是其在个性化互补服装推荐中的潜在价值还没有被充分挖掘。

然而,联合探索一阶交互和高阶交互关系来进行个性化互补服装推荐极具挑战性。主要挑战如下:1)缺少统一的时尚图谱。在时尚领域的个性化互补服装推荐中,没有一个统一可用的时尚图谱来表示用户、单品和属性之间的复杂高阶关系。因此,如何去构建一个时尚图谱来表示时尚实体之间的高阶交互关系对我们来说是一个主要的挑战。2)高阶关系的利用。事实上,关键的时尚实体,即用户、单品和属性之间,不仅有用户-单品(user-item)偏好交互关系以及单品-单品(item-item)搭配交互关系,而且还有单品-属性(item-attribute)之间的从属交互关系。因此,如何有效地利用它们之间的高阶交互信息,从而促进对用户和单品的表示学习也是一个至关重要的挑战。3)交互模块的融合。如何将独立的一阶交互建模和协同的高阶交互建模进行有机融合,进而共同促进个性化互补服装推荐的效果构成了另一个艰难的挑战。

为了解决以上提到的多个挑战,我们首次构建了一个协同的时尚图谱,模拟样例图如图2所示,它能够清楚地表示时尚实体之间复杂的高阶交互关系。不同于以往工作<sup>[3-4]</sup>中所使用的时尚图谱,在本工作中,我们将单品的属性知识(即,单品的知识图谱)也考虑在内,并且与用户-单品交互关系以及单品-单品搭配关系共同构成一个大型的协同时尚图谱。同时,将这些单品的属性也看作图谱中的节点。通过这个协同的时尚图谱,我们可以探索更丰富的高阶关系,比如:  $u_m \xrightarrow{r0} t_i \xrightarrow{rm} \{b_1, b_j\} \xrightarrow{r1} a_2$  和  $u_m \xrightarrow{r0} b_j \xrightarrow{r2} a_3 \xrightarrow{r2} b_2 \xrightarrow{r0} \{u_2, u_3\}$ , 其中,符号  $u$ ,  $t$ ,  $b$  和  $a$  分别表示时尚图谱中的用户、上衣、下衣以及单品的属性,它们对应的角标(即  $m, i, j, 1, 2, 3$ )表示它们在相应用户、单品或属性集合中的编号。另外,箭头上的符号表示两个实体之间的关系。其中,  $rm$  指的是单品之间的搭

配关系,  $r0$  表示用户和单品之间的交互关系,  $r1$  表示单品的颜色属性从属关系,  $r2$  表示单品的材质属性从属关系。正如我们所见,这些高阶交互关系对于个性化互补服装推荐呈现了更深层次的信息和理解。比如,从图2可知,对于上衣  $t_i$  来说,它更适合与具有属性  $a_2$  的下衣搭配;用户  $u_m$  和  $u_2$  都喜欢具有属性  $a_3$  的下衣。

据此,我们提出了一个新颖的基于时尚图谱增强的个性化互补服装推荐模型(Fashion Graph-enhanced Personalized Complementary Clothing Recommendation), 简称为 FG-PCCR。FG-PCCR 由两个关键的部分组成:独立的一阶交互建模(Independent One-order Interaction Modeling)和协同的高阶交互建模(Collaborative High-order Interaction Modeling)。其中,一方面,独立的一阶交互建模模块致力于通过神经网络和矩阵分解方法分别对单品-单品搭配交互和用户-单品偏好交互进行建模。同时,为了增强模型效果,我们将时尚单品的视觉和文本模态均整合进去,以综合地实现多模态建模。另一方面,协同的高阶交互模块聚焦于基于协同时尚图谱的高阶交互建模。具体地,利用图神经网络(Graph Neural Networks, GNNs)的信息传播机制来提取高阶的协同信号,进一步丰富用户和单品的向量表示。最后,FG-PCCR 基于贝叶斯个性化排序框架(Bayesian Personalized Ranking, BPR)<sup>[5]</sup>将独立的一阶交互建模和协同的高阶交互建模有机统一在一个端到端的模型中。

总体而言,本文的主要贡献概括为如下三方面:

1) 提出了一个基于时尚图谱增强的个性化互补服装推荐模型 FG-PCCR。FG-PCCR 能够将独立的一阶交互建模和协同的高阶交互建模有效地整合在一个端到端的模型中。该工作的关键创新性在于整合了时尚实体(如,用户、单品和属性)之间的高阶交互关系,进而有效地促进时尚领域的个性化互补服装推荐的效果。

2) 设计了一个协同的高阶交互建模模块。基于图神经网络,该模块能够编译时尚实体之间的复杂高阶关系信息并聚合于用户和单品(如,上衣和下衣),进而促进用户和单品的表示学习,得到其更丰富的向量表示。

3) 在公开的数据集 IQON3000 上进行了全面的实验,评估所提出模型的有效性。大量实验验证了所提模型相对于基准方法的优越性。同时,公布了代码、数据和参数设置,以方便同领域研究人员的对我们工作的复现和拓展。

文章结构安排如下:第2章简单回顾了相关工

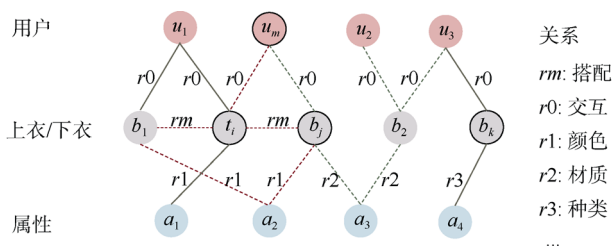


图2 协同时尚图谱的模拟样例图

Figure 2 A toy example of the collaborative fashion graph

作;第3章详细介绍了本文所提出的FG-PCCR模型;第4部分阐明了实验的实施细节以及实验结果并进行了综合性分析;第5部分对本工作进行了总结,并针对本文的不足提出了未来改进的方向。

## 2 相关工作

### 2.1 时尚分析

近年来,随着人们对穿着打扮有了更高的美学追求,时尚领域中的服装兼容性建模问题<sup>[6-8]</sup>逐渐引发了越来越多科研人员的关注。

目前,一些工作着力于评估时尚单品对之间的兼容性。McAuley等人<sup>[9]</sup>提出,将时尚单品映射到一个潜在的风格空间,使可替换或互补单品之间的距离尽可能地近。之后,研究者们开始利用视觉模态进行兼容性建模。比如,Veit等人<sup>[10]</sup>将一对时尚单品的图片输入一个端对端的暹罗网络中,通过学习视觉上的语义信息,进而生成服装套装。He等人<sup>[11]</sup>提出了Monomer模型,通过建模单品视觉上的相似性,来探索单品之间的关系。除了视觉特征,最近的很多工作也在时尚兼容建模相关的任务中强调了探索多模态特征的重要性。例如,李等人<sup>[12]</sup>基于时尚单品的多模态多实例的深度学习,提出了一个套装质量预测器。另外,宋等人<sup>[8]</sup>从时尚社区平台Polyvore收集并创建了一个套装数据集,同时设计了一个基于自动编码器的多模态互补单品兼容性建模框架。之后,杨等人<sup>[13]</sup>提出了一个新颖的神经时尚兼容建模模型TransNFCM,该框架引入了一个多模态的单品编码器,能够有效探索不同模态的互补特征。除此之外,研究人员也探索并采用了一些辅助性信息来促进兼容性建模的效果,比如:单品种类<sup>[14]</sup>和域知识<sup>[8,15]</sup>。具体来说,宋等人<sup>[7]</sup>注意到时尚领域积累的丰富的搭配知识,进而提出了一个基于知识蒸馏的服装兼容性建模模型。该工作清楚地阐释了整合丰富的时尚领域知识来助力数据驱动的时尚兼容性建模的重要性。

另外,时尚领域的个性化推荐也获得了很多研究者的关注<sup>[16-18]</sup>。已有时尚领域中的个性化推荐工作主要采用矩阵分解(Matrix Factorization, MF)<sup>[19]</sup>来建模用户的偏好。MF能够将每一个用户和单品看作一个有ID信息的向量,然后通过内积来建模用户-单品交互。例如,Hu等人<sup>[20]</sup>提出了一个函数张量分解模型以解决个性化套装推荐问题。尽管该方法在整套服装推荐中取得了不错的效果,但是冷启动仍是一个有待探索的问题。为此,He等人<sup>[21]</sup>引入了一个可扩展的矩阵分解模型,将时尚单品的视觉信号整

合于用户偏好预测器中,进而实现推荐任务。

为提高服装兼容性建模的实际应用性,个性化互补服装推荐已成为当今的研究热点<sup>[1-2,20]</sup>。例如,宋等人<sup>[2]</sup>提出了一个关于服装搭配的个性化兼容建模的模型。该模型在一般的兼容性建模中整合了用户的偏好,同时为了获得用户更精准的偏好,考虑了视觉和文本两个方面的偏好。另外,董等人<sup>[1]</sup>设计了一个个性化胶囊衣橱创建的方法。该方法在考虑用户身材的基础上,通过对用户和衣服建模,获得用户的个性化胶囊衣橱进而给用户推荐搭配好的服装。整体而言,对于个性化互补服装推荐任务来说,尽管这些工作已经取得巨大成功,然而他们忽略了一个至关重要的因素:实体(如,用户、单品和属性)之间的高阶交互关系。这种高阶关系能够传递重要的协同信息来丰富实体的表示向量,进而提高个性化互补服装推荐的效果,这也是本工作的研究重点。

### 2.2 知识图谱

知识图谱(Knowledge Graph)作为一种结构化的语义知识库,在挖掘实体之间潜在的复杂关系中有着杰出的能力,被广泛应用于各种研究领域,例如智能语义搜索<sup>[22]</sup>,推荐系统<sup>[23-25]</sup>和智能问答<sup>[26-27]</sup>。在推荐领域,研究人员利用知识图谱已经做了大量相关的工作<sup>[28-31]</sup>,大体上可以分为两类:以路径为基础的方法和以校正为基础的方法。

以路径为基础的方法旨在提取大量携带高阶关系的路径,然后将其输入推荐模型中,对目标用户进行偏好预测。为了处理大量的两个节点之间的路径,研究者一般采用路径选择算法选择最优路径<sup>[32-33]</sup>或定义元路径模式来限制<sup>[34-35]</sup>。以路径为基础的方法存在的主要问题是路径选择对于最后的推荐效果有巨大影响,而且路径不会被优化。另外,定义有效的元路径需要域知识,这对于拥有多样化实体和关系的复杂知识图谱来说,是相当耗费人力的。

以校正为基础的方法旨在联合训练推荐任务和知识图谱构建任务,通过设计构建知识图谱结构的额外损失来校正整体推荐模型的损失,比如,CFKG<sup>[36]</sup>和KTUP<sup>[37]</sup>。由于该方法并不是直接将高阶关系放入到推荐模型中,只是以一种隐含的方式进行编码,所以知识图谱中的高阶关系并不能保证被获取到。之后,为解决以上问题,王等人<sup>[23]</sup>提出了一个知识图谱注意力网络,能够明确有效地利用知识图谱中的高阶关系进行推荐。尽管知识图谱已经被成功地应用在各种任务中,但是在时尚领域个性化互补服装推荐任务中的研究还是十分有限的。

现有时尚领域中涉及的图谱<sup>[3-4]</sup>含有的实体一般



只包括用户和单品,忽略了时尚单品的客观属性知识。事实上,这些属性知识对于单品兼容建模和用户偏好建模能够传递更多隐含的协同信号,进而促进个性化互补服装推荐的效果。因此,本工作在一般的时尚图谱基础上引入了单品的知识图谱(即,单品-属性从属关系),构建了一个协同的时尚图谱,利用用户、单品及属性之间的复杂关系来提高个性化互补服装推荐的效果。

### 2.3 图神经网络

图神经网络(GNNs)在近年来获得了越来越多学者的注意<sup>[23, 38]</sup>。GNNs 已经变成了一种强有力的建模结构化数据的工具,通过图进行信息传递,进而将更丰富细致的信息聚焦于节点,得到更优的向量表示。

图神经网络的概念首次被 Scarselli 等人<sup>[39]</sup>提出。GNNs 将已有的神经网络拓展并应用于图领域表示的数据中,能够直接处理图结构数据。虽然 GNNs 可以被应用于大部分类型的图中,比如循环图,非循环图,有向图以及无向图,但是最初的 GNNs 对于固定点的训练是有一定难度的。之后,图卷积神经网络(Graph Convolutional Network, GCN)<sup>[40]</sup>被提出,该网络提出了适用于图域的卷积操作,并通过聚合来自所有邻居节点的信息来实现节点向量表示的更

新。由于 GCN 需要捕捉图的全局信息来更新节点,Hamilton 等人<sup>[41]</sup>提出了 GraphSAGE,该模型通过统一抽样和聚合局部邻居的特征来更新节点表示,节省了计算开销。尽管图神经网络能够通过传递任意深度的邻居信息,但是长期的信息传播可能产生梯度消失问题。为了纠正这个问题,最近的先进工作<sup>[42]</sup>尝试在传播过程中引入了门控循环单元(Gate Recurrent Units, GRU)。

由于 GNNs 所拥有的杰出的表达能力和模型灵活性,它已经成功地被应用于各种关于复杂关系的任务中,包括检索<sup>[46-47]</sup>和视觉理解<sup>[48]</sup>等。然而,在时尚领域的个性化互补服装推荐任务中,探索 GNNs 应用潜力的研究还相对较少。事实上, GNNs 已经被证明在处理复杂图数据方面,比如社交网络<sup>[41, 43]</sup>,物理系统<sup>[44-45]</sup>,对于高阶交互关系,有着强有力的建模能力。据此,在本工作中,我们提出的 FG-PCCR 模型将以 GNNs 为基础,进而实现协同的高阶交互建模。

### 3 基于时尚图谱增强的个性化互补服装推荐模型

在本章中,我们首先将研究的问题公式化,然后详细地介绍了本文提出的基于时尚图谱增强的个性化互补服装推荐模型 FG-PCCR,模型结构如图 3

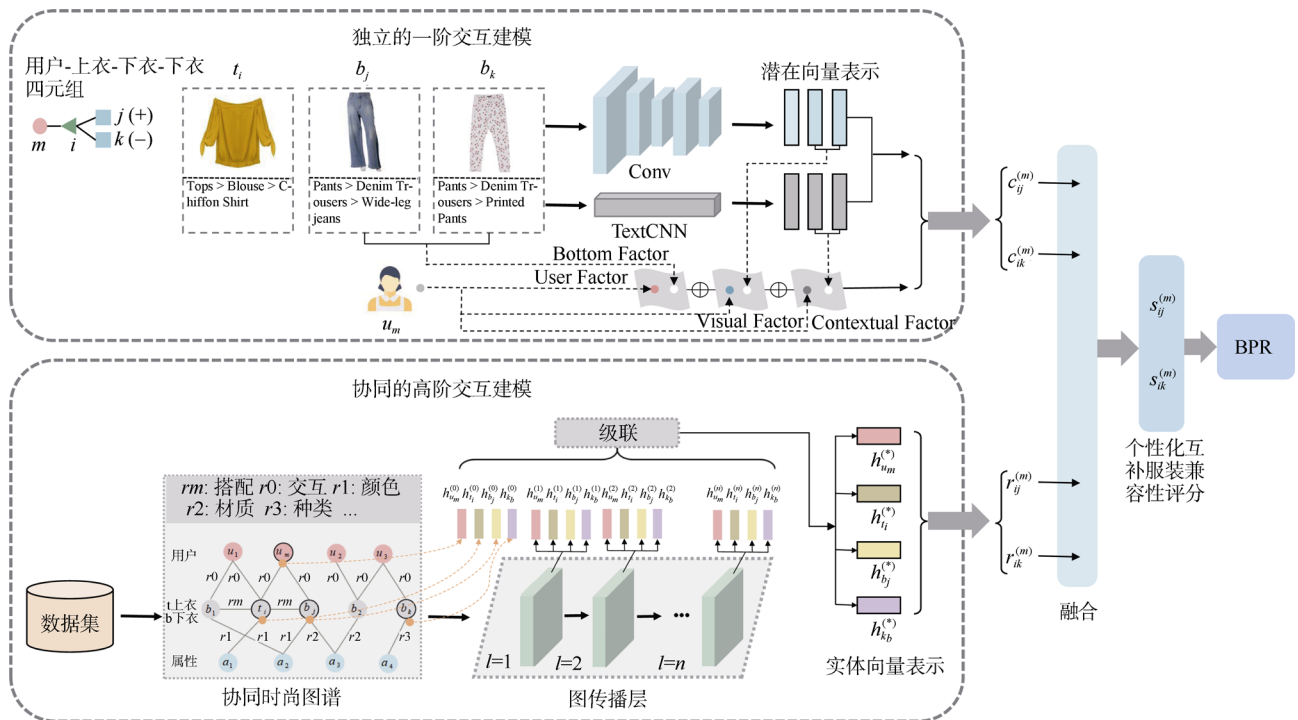


图 3 本文提出的 FG-PCCR 模型结构, 由两个关键组件组成: 1)独立的一阶交互建模 和 2)协同的高阶交互建模  
Figure 3 Illustration of the proposed FG-PCCR, comprising two key components: 1) the independent one-order interaction modeling and 2) the collaborative high-order interaction modeling

所示, 其中包括两个关键组件: 独立的一阶交互建模模块和协同的高阶交互建模模块。最后介绍了整体模型的目标与优化。

### 3.1 问题定义

不失一般性地, 本文以上衣和下衣之间的兼容性建模举例。假设我们有 4 个实体集合: 用户集合  $\mathbf{U} = \{u_1, u_2, \dots, u_M\}$ , 上衣集合  $\mathbf{T} = \{t_1, t_2, \dots, t_{N_t}\}$ , 下衣集合  $\mathbf{B} = \{b_1, b_2, \dots, b_{N_b}\}$ , 还有属性集合  $\mathbf{A} = \{a_1, a_2, \dots, a_{N_a}\}$ , 其中,  $M, N_t, N_b$  和  $N_a$  分别表示用户, 上衣, 下衣还有属性的总数。对于每个用户  $u_m$  来说, 都有一个与之对应的历史套装集合  $\mathbf{O}_m^+ = \{(t_{i_1^m}, b_{j_1^m}), (t_{i_2^m}, b_{j_2^m}), \dots, (t_{i_{N_m}^m}, b_{j_{N_m}^m})\}$ , 其中  $i_{N_m}^m \in \{1, 2, \dots, N_t\}$  和  $j_{N_m}^m \in \{1, 2, \dots, N_b\}$  分别指上衣和下衣的索引,  $N_m$  表示第  $m$  个用户的套装总数。对于每个单品  $t_i(b_j)$ , 我们分别使用  $\mathbf{v}_i^v(\mathbf{v}_i^b) \in \mathbf{R}^{D_v}$  和  $\mathbf{c}_i^t(\mathbf{c}_i^b) \in \mathbf{R}^{D_c}$  来表示它的视觉和文本的嵌入向量。 $D_v$  和  $D_c$  分别表示对应的嵌入向量的维度。

在上衣和下衣之间的个性化兼容性建模任务中, 其核心问题在于如何针对给定的用户, 评价一件上衣与一件下衣的兼容性。因此, 我们意在设计一个准确的个性化互补服装兼容性建模模型  $\psi$ 。我们定义的模型  $\psi$  可以用如下的数学公式来表示:

$$S_{ij}^{(m)} = \psi(u_m, t_i b_j | \Theta) \quad (1)$$

公式中,  $\Theta$  是将要被训练的模型参数集。 $S_{ij}^{(m)}$  表示上衣-下衣对  $(t_i b_j)$  对于用户  $u_m$  来说的个性化互补服装兼容性评分。基于该得分, 对于给定的用户和目标上衣, 我们可以生成一个个性化的下衣排序列表, 从而完成个性化的互补服装推荐。

### 3.2 独立的一阶交互建模模块

由于目标下衣不仅要和上衣能够搭配和谐, 还要满足用户的个人偏好, 我们将独立的一阶交互建模进一步分成了两部分: 客观的单品-单品搭配交互建模和主观的用户-单品偏好交互建模。前者旨在挖掘单品之间的一阶交互, 而后者旨在探索用户与单品之间的一阶交互。由于时尚单品通常都会存在视觉模态(即, 图片)和文本模态(即, 类别元数据和简单的描述), 因此, 在本工作中, 我们同时探索这两种模态来综合衡量时尚单品对(上衣-下衣)的兼容性以及用户的个人偏好。

类似于已有工作<sup>[8]</sup>, 我们认为存在一个潜在的、

非线性的兼容空间, 能够连接来自互补类别的时尚单品, 使得我们可以建模互补单品之间的兼容性。由于多层感知机(Multi-Layer Perception, MLP)在各种表示学习任务中取得了引人注目的成功, 我们在客观的单品-单品交互建模中采用了多层感知机, 来学习时尚单品的向量表示。假设我们采用的多层感知机一共有  $K$  个隐藏层。令  $\mathbf{x}_i^y, \mathbf{x} \in \{\mathbf{v}, \mathbf{c}\}, y \in \{t, b\}$  表示时尚单品视觉(文本)模态的嵌入向量表示, 其中  $\mathbf{v}$  表示视觉模态,  $\mathbf{c}$  表示文本模态,  $t$  表示上衣,  $b$  表示下衣。将单品各模态的嵌入向量表示  $\mathbf{x}_i^y$  输入多层感知机, 我们有如下公式:

$$\begin{cases} \mathbf{z}_{i1}^y = \sigma(\mathbf{W}_1^y \mathbf{x}_i^y + \mathbf{b}_1^y), \\ \mathbf{z}_{ik}^y = \sigma(\mathbf{W}_k^y \mathbf{z}_{i(k-1)}^y + \mathbf{b}_k^y), k = 2, \dots, K \end{cases} \quad (2)$$

公式中,  $\mathbf{z}_{ik}^y$  表示隐藏层的表示向量,  $\mathbf{W}_k^y$  和  $\mathbf{b}_k^y, k = 2, \dots, K$ , 分别指的是权重矩阵和偏差矩阵。 $\sigma: \mathbf{R} \mapsto \mathbf{R}$  是非线性激活函数。在本工作中,  $\sigma$  被指定为 sigmoid 函数  $\sigma(x) = \frac{1}{1 + e^{-x}}$ 。我们将第  $K$  层的输出当作时尚单品的潜在表示向量, 即  $\tilde{\mathbf{x}}_i^y = \mathbf{z}_{ik}^y \in \mathbf{R}^{D_0^y}$ , 其中  $D_0^y$  指的是潜在兼容空间的维度。按此方式, 我们可以得到上衣  $t_i$  和下衣  $b_j$  的潜在视觉和文本向量, 可以分别表示为  $\tilde{\mathbf{v}}_i^t, \tilde{\mathbf{c}}_i^t, \tilde{\mathbf{v}}_j^b$  和  $\tilde{\mathbf{c}}_j^b$ 。最终, 我们按照如下公式定义上衣  $t_i$  和下衣  $b_j$  之间的客观兼容性分数  $o_{ij}$ :

$$o_{ij} = (1 - \alpha)(\tilde{\mathbf{v}}_i^t)^T \tilde{\mathbf{v}}_j^b + \alpha(\tilde{\mathbf{c}}_i^t)^T \tilde{\mathbf{c}}_j^b \quad (3)$$

公式中,  $\alpha$  表示非负的平衡参数, 用来校正不同模态在兼容性建模中的贡献。

考虑到不同用户对于下衣的偏好有所不同。比如, 有的用户喜欢穿牛仔裤, 而有的用户更偏爱裙子。因此, 除了客观的单品-单品兼容性建模, 我们也整合了主观的用户-单品偏好交互建模。鉴于矩阵分解<sup>[19]</sup>已经在大量的个性化推荐任务<sup>[2, 20-21, 49]</sup>中展示出了优越的效果, 在本工作中, 我们也选用了矩阵分解来对用户-单品的偏好交互进行建模。具体地, 我们将用户  $u_m$  对于下衣  $b_j$  的偏好分数定义为如下数学表达式:

$$\begin{aligned} p_{mj} = & \delta + \beta[\theta_m + \theta_j + (\mu_m)^T \mu_j] \\ & + \gamma[(\xi_m^v)^T \tilde{\mathbf{v}}_j^b + (\xi_m^c)^T \tilde{\mathbf{c}}_j^b] \end{aligned} \quad (4)$$

公式中,  $\delta$  代表用户对下衣偏好的统一偏移量。

$[\theta_m + \theta_j + (\mu_m)^T \mu_j]$  和  $[(\xi_m^v)^T \tilde{v}_j^b + (\xi_m^c)^T \tilde{c}_j^b]$  分别表示用户  $u_m$  对于下衣  $b_j$  的整体偏好和面向内容(包括视觉和文本)偏好。 $\theta_m$  和  $\theta_j$  分别指得是用户  $u_m$  和下衣  $b_j$  的偏移量。 $\mu_m$  和  $\mu_j$  分别指得是用户  $u_m$  和下衣  $b_j$  的潜在的整体偏好因子向量。 $\xi_m^v$  和  $\xi_m^c$  分别表示用户  $u_m$  潜在的视觉和文本偏好因子向量。另外,  $\beta$  和  $\gamma$  指的是非负的平衡超参数, 用来平衡用户  $u_m$  的整体偏好和面向内容偏好这两部分的重要性。

最后, 根据独立的一阶交互建模, 对于给定的用户  $u_m$ , 我们根据如下公式得到上衣  $t_i$  和下衣  $b_j$  的个性化兼容分数:

$$c_{ij}^{(m)} = (1 - \lambda)o_{ij} + \lambda p_{mj} \quad (5)$$

公式中,  $\lambda$  是非负的平衡参数。 $o_{ij}$  指的是上衣  $t_i$  和下衣  $b_j$  之间的客观兼容性得分, 而  $p_{mj}$  指的是用户  $u_m$  对下衣  $b_j$  的主观偏好得分。

### 3.3 协同的高阶交互建模模块

如前所述, 大部分已有的个性化互补服装推荐研究<sup>[1-2]</sup>致力于利用时尚单品(比如, 上衣或下衣)的多模态数据(比如, 视觉和文本模态)来建模独立的单品-单品交互和用户-单品交互, 而忽略了实体之间的协同高阶关系。事实上, 与一个给定的上衣相搭配的下衣可能共享某些特定的属性模式(如, 颜色、种类、样式等)。同时, 有着相似穿搭品味的用户们也可能偏好具有类似属性特征的时尚单品。据此, 本工作在一般的时尚图谱中又整合了时尚单品的属性知识, 打开了探索类似的用户以及类似的时尚单品之间的高阶交互关系的大门。作为本工作的一个主要新颖性贡献, 我们充分利用时尚实体(即, 用户, 单品和属性)之间复杂的高阶交互关系来增强用户和时尚单品的向量表示, 这种做法能够使我们对于单品对之间的兼容性建模以及用户对单品的偏好建模, 挖掘更深层次的协同信息, 进而促进整体的个性化互补服装推荐的效果。

在一定程度上来说, 受文献[23]的启发, 我们首次基于公开可用的数据集创建了一个大规模的双向的协同时尚图谱  $G$ , 该时尚图谱可以表示为  $\{(h, r, t) | h, t \in E, r \in R\}$ , 其中, 符号  $E$  表示一个包含用户, 时尚单品(即, 上衣和下衣)以及单品属性的实体集合; 符号  $R$  表示头实体和尾实体之间的关系集合, 包括用户-单品偏好交互, 单品-单品搭配交互和单品-属性从属交互。此外, 在本工作中, 我们采用了

图卷积神经网络 GCNs, 其最初被设计是用于半监督的节点分类任务, 而我们用其循环获取并聚合来自更高跳的邻居节点的信息, 进而促进用户、上衣和下衣的表示学习。特别地, 我们首先利用高斯分布来初始化协同时尚图谱  $G$  中的每一个节点, 并且将初始化后的节点向量表示为  $h_v^0$ 。然后, 每个节点  $v \in E$  的传播机制可以用如下所示的数学表达描述:

$$\begin{cases} h_{N_v}^{(l-1)} = A_v [h_1^{(l-1)^T} \dots h_{|E|}^{(l-1)^T}] \\ h_v^l = \text{LeakyReLU}(W^l (h_v^{(l-1)} + h_{N_v}^{(l-1)})) \\ l = 1, \dots, L \end{cases} \quad (6)$$

公式中, 符号  $N_v$  表示所有三元组  $(v, r, t)$  的集合, 且在时尚图谱  $G$  中将节点  $v$  看作头实体。符号  $h_{N_v}^{(l-1)}$  表示在第  $(l-1)$  层的传播中, 节点  $v$  从其所有邻居节点获取的潜在信息。 $[h_1^{(l-1)^T} \dots h_{|E|}^{(l-1)^T}] \in \mathbb{R}^{|E| \times D_{(l-1)}}$  表示由所有节点在第  $(l-1)$  层卷积操作时的向量表示组成的矩阵, 其中  $|E|$  表示时尚图谱  $G$  中所有节点的总个数,  $D_{(l-1)}$  表示节点向量的维度大小。符号  $A_v \in \mathbb{R}^{|E|}$  表示节点  $v$  的正则化邻接矩阵, 该矩阵是根据图谱中的三元组生成的。具体而言, 对于  $A_v$  中第  $(a, b)$  个元素来说, 即第  $a$  行  $b$  列的值。如果第  $a$  个实体和第  $b$  个实体之间存在一条关系的话, 那么这个值是  $\frac{1}{|N_t|}$ , 其中  $N_t$  指的是节点  $v$  所有邻居节点的个数; 相反, 如果第  $a$  个实体和第  $b$  个实体之间不存在任何关系的话, 那么这个值就是 0。最后, 符号  $h_v^l$  是节点  $v$  在第  $(l-1)$  层卷积操作之后更新所得的向量表示, 它融合了节点本身的信息以及其所有邻居节点的信息。符号  $W^l$  指的是第  $(l-1)$  层的权重矩阵。另外, 在本工作中, 我们采用了 LeakyReLU<sup>[49]</sup> 作为激活函数。

假设本文一共采用  $L$  个图卷积层, 那么时尚图谱  $G$  中每个节点均可以获得一系列向量表示。例如, 用户  $u_m$  的所有向量表示为  $\{h_{u_m}^{(0)}, \dots, h_{u_m}^{(L)}\}$ , 上衣  $t_i$  的所有向量表示为  $\{h_{t_i}^{(0)}, \dots, h_{t_i}^{(L)}\}$ , 下衣  $b_j$  的所有向量表示为  $\{h_{b_j}^{(0)}, \dots, h_{b_j}^{(L)}\}$ 。类似于文献[23], 本工作为了能够积累来自每一个传播层之后获得的单品兼容性以及用户偏好信息, 我们进行了聚合操作, 将各层的向量表示级联成了一个综合向量, 其数学表达公式如下所示:

$$\begin{cases} \mathbf{h}_{u_m}^* = \mathbf{h}_{u_m}^{(0)} \parallel \mathbf{h}_{u_m}^{(1)} \dots \parallel \mathbf{h}_{u_m}^{(L)} \\ \mathbf{h}_{t_i}^* = \mathbf{h}_{t_i}^{(0)} \parallel \mathbf{h}_{t_i}^{(1)} \dots \parallel \mathbf{h}_{t_i}^{(L)} \\ \mathbf{h}_{b_j}^* = \mathbf{h}_{b_j}^{(0)} \parallel \mathbf{h}_{b_j}^{(1)} \dots \parallel \mathbf{h}_{b_j}^{(L)} \end{cases} \quad (7)$$

公式中, 符号  $\mathbf{h}_{u_m}^*$ ,  $\mathbf{h}_{t_i}^*$  和  $\mathbf{h}_{b_j}^*$  分别表示用户  $u_m$ , 上衣  $t_i$  以及下衣  $b_j$  的最后的向量表示。符号  $\parallel$  表示向量的级联操作。

最后, 我们从高阶交互建模的角度定义了个性化互补服装兼容性分数  $r_{ij}^{(m)}$ 。它反映了用户  $u_m$  对于给定的上衣  $t_i$ , 对搭配的下衣  $b_j$  的偏好程度, 其数学表达式如下所示:

$$r_{ij}^{(m)} = (1 - \zeta)(\mathbf{h}_{t_i}^{*T} \mathbf{h}_{b_j}^*) + \zeta(\mathbf{h}_{u_m}^{*T} \mathbf{h}_{b_j}^*) \quad (8)$$

公式中, 第一项  $(\mathbf{h}_{t_i}^{*T} \mathbf{h}_{b_j}^*)$  和第二项  $(\mathbf{h}_{u_m}^{*T} \mathbf{h}_{b_j}^*)$  分别表示基于高阶交互建模的上衣  $t_i$  和下衣  $b_j$  之间的兼容性和用户  $u_m$  对于下衣  $b_j$  的偏好。另外, 符号  $\zeta$  是一个非负的平衡参数。

### 3.4 整体模型的目标与优化

基于上述独立的一阶交互建模和协同的高阶交互建模这两个关键的组件, 可以为个性化互补服装推荐任务得出最后综合的个性化互补服装兼容性评分  $s_{ij}^{(m)}$ , 其数学表达公式如下所示:

$$s_{ij}^{(m)} = (1 - \tau)c_{ij}^{(m)} + \tau r_{ij}^{(m)} \quad (9)$$

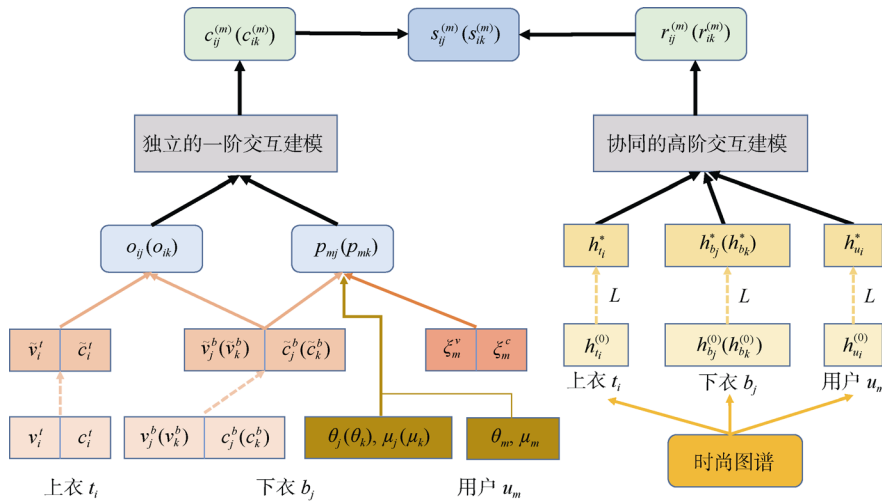


图 4 基于时尚图谱增强的个性化互补服装推荐模型的工作流程图

Figure 4 The workflow of fashion graph-enhanced personalized complementary clothing recommendation model

## 4 实验

为了评估本文提出的基于时尚图谱增强的个性化互补服装推荐模型 FG-PCCR, 我们在真实的数据

公式中, 符号  $\tau$  是一个非负的平衡参数, 来平衡独立的一阶交互建模和协同的高阶交互建模这两个关键组件的重要性。

鉴于贝叶斯个性化排序框架已经在许多成对的隐含偏好建模工作<sup>[2, 53-54]</sup>中取得了杰出效果, 在本工作中, 我们同样采用了 BPR 框架进行目标优化。根据文献<sup>[21]</sup>, 我们所提出的 FG-PCCR 模型的最后的函数, 可以被定义为如下的数学表达式:

$$\begin{aligned} \Theta^* = \arg \min & \left( \sum_{(m,i,j,k) \in S_{train}} [-\ln(\sigma(s_{ij}^{(m)} - s_{ik}^{(m)}))] \right. \\ & \left. + \rho \|\Theta\|_2^2 \right) \end{aligned} \quad (10)$$

公式中,  $S_{train} = \{(m, i, j, k) | u_m \in u \wedge (t_i, b_j) \in O_m^+ \wedge b_k \in B \setminus b_j\}$  表示训练数据集。每个四元组  $(m, i, j, k)$  表示针对上衣  $t_i$ , 用户  $u_m$  更加偏好于下衣  $b_j$ , 而不是下衣  $b_k$ 。符号  $\sigma(\cdot)$  指的是 sigmoid 函数。符号  $\Theta$  指的是所提模型 FG-PCCR 的模型参数集, 即  $\Theta = \{\mathbf{W}_k^y, \mathbf{b}_k^y, \delta, \theta_m, \theta_j, \mu_m, \mu_j, \xi_m^y, \xi_m^c, \mathbf{H}, \mathbf{W}^{(l-1)}\}$  其中, 符号  $\mathbf{H}$  是由时尚图谱  $\mathbf{G}$  中所有实体的向量表示构成的矩阵。为了防止模型过拟合, 本工作采用了通过参数  $\rho$  校正的 L2 范数  $\|\Theta\|_2^2$ 。整个模型的优化算法被简明地总结在算法 1 中, 以及图 4 详细描述了我们所提模型 FG-PCCR 的工作流程。

集 IQON3000 上做了大量实验。整个实验的开展旨在回答以下四个研究问题:

问题一: 我们提出的 FG-PCCR 模型是否能超越最新的基准方法?



问题二: FG-PCCR 中协同的高阶交互建模模块是如何影响整体模型效果的?

问题三: 不同的模态信息对于模型效果的贡献是怎样的?

问题四: FG-PCCR 在互补时尚单品检索应用中的效果是怎样的?

**算法 1.** FG-PCCR 模型的训练过程算法.

输入:  $S_{train}$ ,  $\mathbf{G}$ ,  $\rho$ ,  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\lambda$ ,  $\zeta$  和  $\tau$

输出: FG-PCCR 中的模型参数  $\Theta$

- 1: 获取协同时尚图谱  $\mathbf{G}$  中每个节点的正则化邻接矩阵  $\mathbf{A}_i$
- 2: 初始化 FG-PCCR 中的模型参数  $\Theta$
- 3: FOR
- 4: 从  $S_{train}$  中获取四元组  $(m, i, j, k)$
- 5: (a)根据公式(5)计算  $c_{ij}^{(m)}$   
(b)根据公式(8)计算  $r_{ij}^{(m)}$   
(c)根据公式(9)计算  $s_{ij}^{(m)}$
- 6: FOR  $\theta$  IN  $\Theta$
- 7: 更新  $\theta \leftarrow \theta + \eta(\sigma(-s_{ij}^{(m)})\frac{s_{ij}^{(m)}}{\theta} - 2\rho\theta)$
- 8: END FOR
- 9: UNTIL 模型收敛
- 10: END FOR

## 4.1 数据集

为了验证本工作提出的 FG-PCCR 模型的有效性, 我们采用了时尚领域中公开可用的基准数据集 IQON3000<sup>[2]</sup>。此数据集由 3,568 个用户创建的 308,747 套时尚套装组成, 一共包含了 672,335 个时尚单品(比如, 上衣、下衣、外套、鞋子等)。每个时尚单品都具有一个视觉图片, 一条标题描述, 和相应的类别元数据。表 1 汇总了 IQON3000 数据集的详细统计数据。

**表 1 基准数据集 IQON3000 的统计数据**  
**Table 1 Data statistics of the benchmark data set IQON3000**

种类	数量
外套	35765
下衣	77813
连衣裙	25816
上衣	119895
鞋子	106598
配饰	306448

本工作以上衣和下衣之间的互补推荐举例, 因此只保留了该数据集中各套装的上衣和下衣数据。最后一共获得由 3,236 个用户搭配而成的 216,791 套仅由上衣和下衣组成的套装。对于用户  $u_m$  的每一个上衣-下衣对  $(t_i, b_j)$ , 我们随机从整个下衣的数据集中获取一个负例, 即下衣  $b_k$ , 进而构成四元组  $(m, i, j, k)$ , 其中  $m, i, j, k$  分别表示的是用户, 上衣, 正例下衣, 负例下衣。据此, 我们获得了一个四元组数据集  $S$ , 该数据集被随机分成三部分: 80%作为训练集  $S_{train}$ , 10%作为验证集  $S_{validate}$ , 10%作为测试集  $S_{test}$ 。

## 4.2 实验设置

### 4.2.1 评估指标

为了对本工作所提出的 FG-PCCR 模型效果进行全面彻底的评估, 本文采用广泛使用的 ROC 曲线面积(Area Under Curve, AUC)<sup>[53]</sup>以及平均倒数排序(Mean Reciprocal Rank, MRR)<sup>[54]</sup>作为衡量个性化互补服装推荐效果的评估指标。首先, 我们定义的 AUC 数学表达式如下所示:

$$AUC = \frac{1}{|\mathbf{U}|} \sum_m \frac{1}{|\mathbf{E}(m)|} \sum_{(i,j,k) \in \mathbf{E}(m)} f(s_{ij}^{(m)} - s_{ik}^{(m)}) \quad (11)$$

公式中,  $\mathbf{E}(m) := \{(i, j, k) | (m, i, j) \in S_{test} \wedge (m, i, k) \notin S\}$ 。  $f(a)$  是一个指示函数。若参数  $a$  大于 0, 返回 1; 否则, 返回 0。其次, MRR 评估指标的数学表达式如下所示:

$$MRR = \frac{1}{|S_{test}|} \sum_{i=0}^{|S_{test}|} \frac{1}{R_i^{(m)}} \quad (12)$$

公式中,  $R_i^{(m)}$  指的是对于给定的用户  $u_m$  和上衣  $t_i$ , 与其搭配的正例下衣的排序位置。

### 4.2.2 实施细节

**时尚图谱。**区别于文献[22], 在本文中, 我们引入了三种实体(即, 用户、单品和属性)作为协同时尚图谱的节点。特别地, 在本文的个性化互补服装推荐的场景中, 单品指的是一个上衣或者下衣。在协同时尚图谱中, 实体之间的关系包括用户-单品偏好交互关系, 单品-单品搭配交互关系以及单品-属性从属交互关系。值得注意的是, 在我们的工作中首次在时尚图谱中引入单品-属性从属交互关系。根据属性的类型, 本工作做了以下预处理: 我们手动将一些属性(比如, 印花、袖长和衣长等)统一成一种类型, 即样式; 同时, 通过删除一些冗余的属性(比如, category\*color), 我们将时尚单品的所有属性汇总成五种类型, 即颜色(color), 材质(material), 品牌(brand), 种类(category)

和样式(pattern)。值得注意的是, 不同于研究工作[4], 在构建的协同时尚图谱中, 我们引入了互补的时尚单品之间的搭配关系, 同时将时尚单品的属性看作图谱中节点, 进而促进实体的表示学习。

同样值得一提的是, 在我们的个性化互补服装推荐工作中, 我们采用了直推式策略来构建协同时尚图谱。其中, 我们将整个数据集  $S$  中所有的用户、单品以及属性都看作实体节点。与此同时, 整个数据集  $S$  中所有的单品-属性从属交互都看作图谱中的关系。至于用户-单品交互关系和单品-单品搭配关系的数据则只取自于训练集  $S_{train}$ 。在表 2 和图 5 中, 我们分别展示了构建的协同时尚图谱的数据统计以及部分可视化图。

表 2 协同时尚图谱的数据统计表

Table 2 Data statistics of the collaborative fashion graph

协同时尚图谱			数量
实体	#用户		3236
	#单品	上衣	99655
		下衣	43086
		#属性	
	关系		7
三元组			1225377

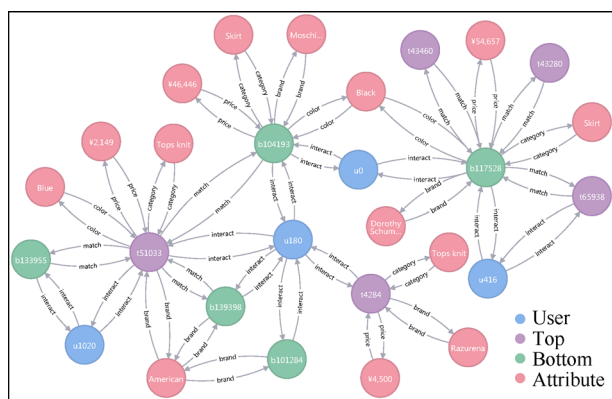


图 5 协同时尚图谱的部分可视化图

Figure 5 Visualization of partial collaborative fashion graph

**文本表示。**在个性化互补服装推荐工作中, 我们将每个时尚单品的标题和类别元数据均看作文本信息。由于数据集是从日本的电商网站所爬取, 其文本信息是日语, 本文并利用日本的形态学分析器

Kuromoji<sup>①</sup>进行分词处理。本文没有采用传统的语言特征<sup>[55-56]</sup>, 而是鉴于卷积神经网络结构(Convolutional Neural Network, CNN)在各种自然语言处理任务<sup>[57-58]</sup>中已取得令人瞩目的成就, 采用 CNN 来提取单品的文本特征。特别地, 我们首先将单品的文本模态信息表示成一个矩阵。矩阵的每一行是一个 300 维的词向量表示, 对应相应的单词。这个 300 维的词向量是通过 NINJAL Web 日本语料库<sup>[61]</sup>提供的 word2vec Nwjc2vec 工具获取的。之后, 我们设置了单通道的 CNN, 它由一个卷积层和一个最大池化层组成。特别地, 我们采用了四种卷积核, 大小分别为 2, 3, 4, 5。每一个卷积核都对应 100 个特征图。另外, 我们选用的激活函数是 Rectified Linear Unit(ReLU)。最终, 我们通过处理每个时尚单品的文本模态, 得到了一个 400 维的表示向量。

**视觉表示。**由于深度卷积神经网络在各种图片表示学习任务<sup>[60-62]</sup>中已经取得了非常出色的效果, 因此, 本工作采用高级的深度卷积神经网络来获得时尚单品的视觉表示。特别值得注意的是, 本文采用的是 50 层的残差网络(ResNet50)<sup>[63]</sup>。将每一个时尚单品的图片输入网络中, 然后将网络中平均池化层的输出作为视觉模态的表示向量。因此, 我们将时尚单品的视觉模态表示为一个 2048 维的向量。

**参数设置。**适应性估计方法(Adam)<sup>[64]</sup>在非凸优化问题中含有诸多优势: 计算高效, 占用内存较少, 调参方便, 适用于大规模数据问题, 能解决高噪音等问题。因此本工作采用 Adam 算法进行模型优化, 最小化之前提到的损失函数来学习所有参数。而且, 我们采用网格搜索策略来决定校正参数  $\rho$  以及平衡参数  $(\alpha, \beta, \gamma, \zeta, \tau)$  在本模型中的最优值。另外, 批处理大小、独立交互建模中的隐藏单元的大小以及学习率  $\eta$  分别在  $[256, 512, 1024]$ ,  $[256, 512, 1024]$  和  $[0.05, 0.01, 0.005, 0.001]$  中搜索。在协同的高阶交互建模部分, 协同时尚图谱中所有实体节点的初始化的表示向量大小设置为 64。同时, 我们不仅调整协同的高阶交互建模模块的深度  $L$  在  $[1, 2, 3, 4, 5, 6]$  范围内, 即 GCNs 的传播层数。我们根据经验将独立的一阶交互建模中的隐藏层的大小设置为  $K=1$ 。与此同时, 我们提出的 FG-PCCR 模型通过 60 个 epoch 进行了微调, 并且在测试集上的效果都被一一展示出来。值得注意的是, 本工作提出的 FG-PCCR 模型在 Pytorch 工具上实现。

类似于其他深度学习方法<sup>[7, 8, 65]</sup>中所描述的收敛

① <http://www.atilika.org/>.

性, 本文也首先验证了所提模型 FG-PCCR 的收敛性。具体来说, 我们在图 6 中展示了所有实例在训练过程中的平均损失(Loss)以及 FG-PCCR 训练一轮的 AUC 变化过程。正如我们从图中看到的, 在前几个 epoch 内, Loss 和 AUC 这两个值急剧下降, 然后倾向于稳定, 这表明我们提出的 FG-PCCR 模型具有良好的收敛性。

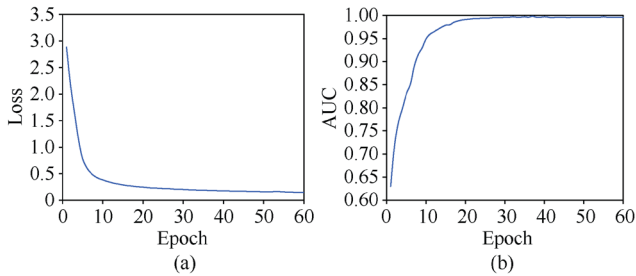


图 6 FG-PCCR 模型中训练损失和准确度收敛性的描述图

Figure 6 Illustration of training loss and accuracy convergence of the proposed FG-PCCR

### 4.3 模型比较

为了验证我们所提出的基于时尚图谱增强的个性化互补服装推荐模型的有效性, 我们选择了如下所述的前沿基准方法:

**POP:** 为了衡量上衣和下衣之间的兼容性, 我们采用“popularity”这个术语来表示某个下衣的受欢迎程度。“popularity”被定义为, 在训练集中与这个下衣搭配的所有上衣的数量。

**RAND:** 随机分配了个性化互补服装的兼容性评分  $s_{ij}^{(m)}$  和  $s_{ik}^{(m)}$ 。

**Bi-LSTM:** 根据文献[6], 我们采用了一个双向的 LSTM 来发现套装兼容性序列, 根据上一个单品预测下一个单品。其中, 我们将双向的 LSTM 的损失看作一个套装兼容性的指示器。值得注意的是, 在我们研究个性化互补服装推荐的背景下, 其中的每个套装只有两个时尚单品组成: 一个上衣和一个下衣。

**BPR-DAE:** 我们采用了文献[8]中引入的内容为基础的神经模型, 它通过一个双向的自动编码器能够联合建模时尚单品的不同模态之间的一致关系, 以及单品对之间的隐藏偏好。然而, 值得注意的是, BPR-DAE 在互补服装推荐中忽略了用户因子。

**BPR-MF:** 我们使用了研究文献[63]中引入的成对排序方法。其中, 用户和时尚单品之间潜在的偏好关系通过矩阵分解的方式来获取。

**KGAT:** KGAT<sup>[23]</sup>作为一个个性化推荐的方法,

它仅仅使用到了实体之间的偏好关系以及属性知识, 却忽略了单品之间内部的兼容性关系以及其他模态信息, 比如, 视觉和文本模态。在我们的工作中, 为了公平起见, 我们意在给定的用户和时尚单品(比如, 上衣)来推荐另一个时尚单品(比如, 下衣)。

**HFGN:** HFGN<sup>[4]</sup>是一种关于个性化套装推荐的方法。在本工作中, 我们通过构建一个分层的图神经网络来获得用户和套装的表示, 然后进行个性化推荐。然而, 该模型没有利用时尚单品的多模态信息(如, 文本信息)。

**VTBPR:** 我们将文献[21]中引入的 VBPR 方法扩展到了 VTBPR。该模型不仅利用了时尚单品的视觉模态, 还将其文本模态也考虑进去, 通过视觉和文本模态综合地对用户的偏好进行建模。

**GP-BPR:** 该方法<sup>[2]</sup>是近年来服装领域中一个杰出的个性化兼容建模的模型, 该模型将一般的兼容性建模和个人偏好建模有机地整合在了一起。特别值得注意的是, GP-BPR 忽略了实体之间的高阶关系, 然而这些高阶关系却能够挖掘在单品兼容性和用户偏好上的协同信息。

表 3 中详细地展示了不同方法的实验比较结果。为了验证模型的效果改进是显著性的, 我们在所提出模型 FG-PCCR 和基准方法之间做了统计显著性检验。从表 3 中, 我们可以得到如下的观察结果:

第一, 和朴素的方法(即, POP 和 RAND)相比较, Bi-LSTM 通过充分探索多模态的数据提高了实验效果, 这表明在时尚单品之间的兼容性建模中, 多模态数据确实起到了至关重要的作用。

第二, BPR-DAE 的效果优于 Bi-LSTM。一个可能的原因是对于一般的只有两个时尚单品的套装的兼容性建模来说, 独立的一阶交互建模比起序列建模来说, 起到的建模效果是更显著的。

第三, GP-BPR 相对于 BPR-MF 和 VTBPR 来说, 其实验结果展示了巨大的优越性, 这证实了在个性化互补服装推荐中, 联合建模时尚单品之间的兼容性以及用户对单品的偏好是很有必要的。

第四, 对于 GP-BPR 和 VTBPR 方法来说, 其实验效果均优于 KGAT 和 HFGN, 这在一定程度上表明, 在个性化服装推荐中, 考虑单品的文本和视觉特征比起只考虑用户单品之间的高阶关系来说, 获得的实验效果更好。

第五, 整体而言, FG-PCCR 和以上提到的所有基准方法相比较, 它取得了最好的实验效果。特别值得注意的是, FG-PCCR 和结果最好的基准方法

GP-BPR 相比, 它的评估指标 AUC 相比于 GP-BPR 提高了 2.39%, 这表明在个性化互补服装推荐任务中, 整合协同时尚图谱来探索时尚实体之间的高阶交互关系是十分有必要的。这背后的原因在于基于时尚图谱的高阶交互建模能够在单品兼容性和用户偏好上传播协同信号, 进而促进用户和时尚单品的表示学习。

**表 3 不同方法在 AUC 上的效果比较表。符号\*表示我们的模型和所有基线方法相比, 效果改进是统计显著性的( $p < 0.01$ )**

**Table 3 Performance comparison among different approaches in terms of AUC. The symbol \* denotes that the performance improvement of our model is statistically significant with  $p < 0.01$  compared against all the baselines**

方法	AUC
POP	0.6039
RAND	0.5002
Bi-LSTM	0.6676
BPR-DAE	0.6886
BPR-MF	0.8005
KGAT	0.8037
HFGN	0.8155
VTBPR	0.8287
GP-BPR	0.8340
<b>FG-PCCR</b>	<b>0.8579*</b>

#### 4.4 高阶交互建模对模型的影响

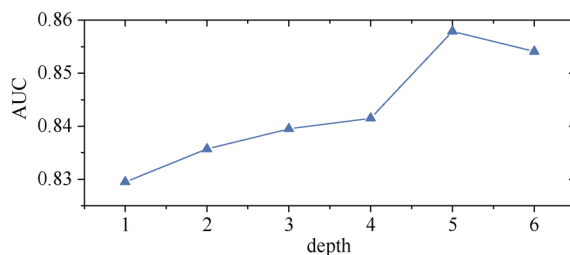
为了对本工作提出的 FG-PCCR 中协同的高阶交互建模模块有更深理解, 本文做了以下两部分实验。一方面, 我们探索了 GCNs 的深度在协同的高阶交互建模部分的影响。另一方面, 我们也做了消融实验来验证其在整个模型中的影响。

##### 4.4.1 GCNs 深度在高阶交互建模中的影响

由于 GCNs 的深度可能会影响协同的高阶交互建模并进而影响个性化互补服装推荐的效果, 所以, 我们做了对应的敏感性分析。对于模型来说, 本文设置了不同的传播层个数, 范围是从 1~6。相应地, 我们得到了 6 个衍生模型, 分别表示为 FG-PCCR-1、FG-PCCR-2、FG-PCCR-3、FG-PCCR-4、FG-PCCR-5、FG-PCCR-6。

如图 7 所示, 我们观察到了模型采用不同 GCNs 深度时, 在 ACU 评估指标上的实验效果。从图中我们可以看到, 随着协同的高阶交互建模模块中 GCNs 深度从 1 到 5 的不断增加, 评估指标 AUC 的大小也

在明显增大。这说明 GCNs 深度的增加在一定程度上能够逐渐增强个性化互补服装推荐的效果。然而, 继续增加 GCNs 深度, 即,  $L=6$  的时候, 模型效果却降低了。显而易见, FG-PCCR-5, 即  $L=5$  时, 本文所提模型取得了最好的实验效果。这表明通过五阶的交互建模来捕获有关单品兼容性和用户偏好上的协同信号是足够充分的。另一个可能的解释是, 当使用的 GCNs 有着更多传播层数的时候(比如,  $L=6$ ), 有关单品兼容性和用户偏好的传播信息会得到累积, 这可能会导致过平滑, 因此进一步恶化实验的效果<sup>[66-67]</sup>。



**图 7 不同深度 GCNs 对模型 FG-PCCR 的效果比较图**  
**Figure 7 Performance comparison of FG-PCCR with different depth of GCNs**

##### 4.4.2 消融实验

为了验证协同时尚图谱在个性化互补服装推荐任务中的影响, 我们将本工作所提出的 FG-PCCR 模型与其两个变体进行了实验比较: FG-PCCR-woI 和 FG-PCCR-woH。具体来说, FG-PCCR-woI 表示通过设置平衡参数  $\tau=1$ , 将独立的一阶交互建模模块从整个的 FG-PCCR 模型中删除之后得到的模型; FG-PCCR-woH 指的是通过设置平衡参数  $\tau=0$ , 将协同的高阶交互建模模块从整个的 FG-PCCR 模型中删除之后得到的模型。由于 FG-PCCR-5 在以上 4.4.1 的实验中已经展示出了最好的效果, 因此这部分的实验我们是在 FG-PCCR-5 模型基础上进行的。

表 4 展示了消融实验在 AUC 评估指标上的效果比较。从表 4 中, 我们可以清晰地观察到, FG-PCCR-woH 方法的效果很大程度上超越了 FG-PCCR-woI 方法。这样的实验结果表明, 独立的一阶交互建模在个性化互补服装推荐中起到了主导的作用。然而, 缺少协同的高阶交互建模模块的 FG-PCCR-woH 方法的实验效果劣于我们提出的 FG-PCCR 模型。这说明协同时尚图谱可以看作时尚领域中个性化互补服装推荐任务的一种强有力的辅助信息, 它能够有效地建模用户, 单品和属性之间的高阶交互关系, 进而捕捉隐含的兼容信号和偏好信号, 从而提高个性化互补服装推荐的效果。

为了直观地展示协同时尚图谱在我们所提出的



表 4 消融实验在 AUC 上的效果比较表

Table 4 Performance comparison on ablation study in terms of AUC

方法	AUC
FG-PCCR-woH	0.8235
FG-PCCR-woI	0.7363
<b>FG-PCCR</b>	<b>0.8579</b>

FG-PCCR 模型中起到的重要作用, 我们在图 8 中可视化了一些实例。具体地, 为了客观的分析, 我们将 FG-PCCR 模型与 FG-PCCR-woH 模型的实际效果在测试四元组中进行了比较。由于展示清晰, 本文中只展示了部分时尚图谱来进行分析。其中所有的四元组满足条件:  $\{u_m, t_i\}: b_j > b_k$ , 即, 对于给定的用户  $u_m$  和上衣  $t_i$ , 正例下衣  $b_j$  比负例下衣  $b_k$  搭配的效果更好。同时, 我们只是在时尚图谱中可视化了一些关键的高阶交互关系, 直观地强调彩色线标识的高阶交互关系的影响。正如我们从图 8 中所观察到的, FG-PCCR 模型在前两个例子中给出了正确的评估结果, 同时, 在第 3 个例子中得到了错误的评估结果, 然而, FG-PCCR-woH 模型却在这 3 个例子中均得到了与 FG-PCCR 相反的结果。通过观察图 8 中所展示的

部分协同时尚图谱, 我们注意到, 关于用户 user1, FG-PCCR 模型获得正确的评估结果的原因, 可以从以下两个角度进行总结。一方面, 对于用户 user1 所在的部分时尚图谱来说, 其中红色线标识的高阶关系表明, 用户 user1 曾经对于给定的上衣  $t_i$  搭配的下衣与本工作给定的正例下衣  $b_j$  共享属性“chiffon”(雪纺), 这可能促进了单品对的兼容性建模。另一方面, 图中绿色的交互表明, 正例下衣  $b_j$  被另一个用户所偏爱, 这个用户和用户 user1 因为都喜欢裙子所以有着类似的时尚品味, 故正例下衣  $b_j$  也很可能被用户 user1 所偏爱, 所以, 这可能促进了用户偏好建模。因此, 总体而言, 对于用户 user1 来说, 为搭配上衣  $t_i$ , 相比下衣  $b_k$  来说, FG-PCCR 模型给下衣  $b_j$  分配了更高的个性化互补服装兼容性评分。同理, 第二个例子也可以得到类似的观察结果。

然而, 除了成功的评估结果, 协同的时尚图谱对于整个 FG-PCCR 模型来说, 也可能导致失败的评估结果。正如图 8 中第 3 个样例所示, 下衣  $b_j$  和  $b_k$ , 均与用户 user3 所购买的上衣搭配过的下衣共享类似的属性(如红色和黄色线标注的关系)。而负例  $b_k$  另外通过绿色线标注的关系来促进用户的

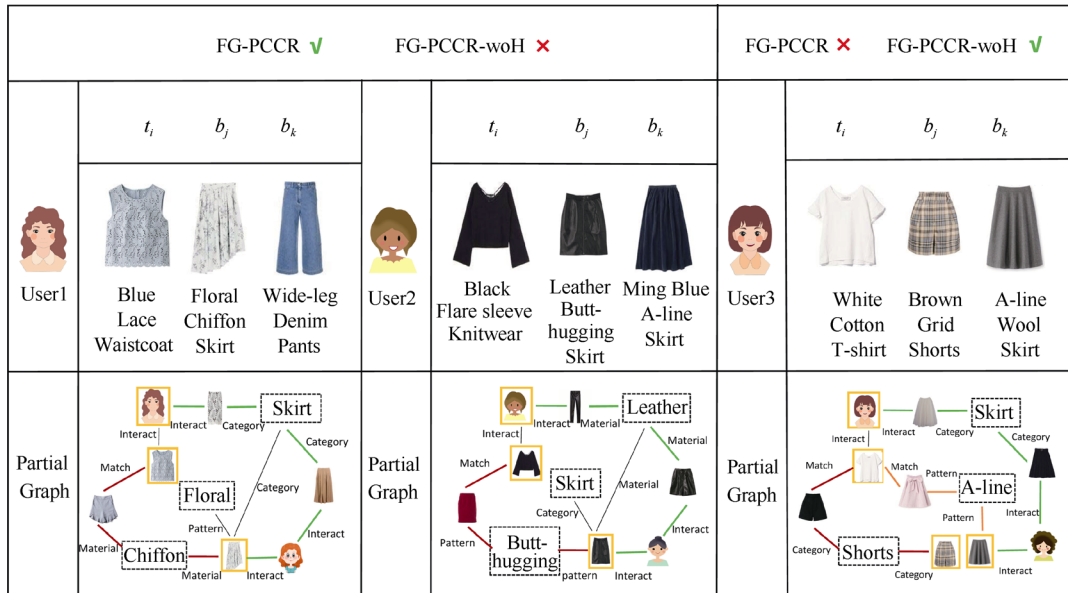


图 8 协同的高阶交互建模的影响描述图。所有的四元组都满足条件  $\{u_m, t_i\}: b_j > b_k$  我们只是在协同时尚图谱中可视化了一些关键的高阶关系并用线的颜色来直观地展示

Figure 8 Illustration of the effect of the collaborative high-order interaction modeling. All the quadruplets satisfy the condition that  $\{u_m, t_i\}: b_j > b_k$ . We only visualize some key higher-order relations in the collaborative fashion graph and intuitively highlight their effects with line colors

偏好建模, 导致负例  $b_k$  的个性化兼容性评分高于正例  $b_j$ , 得到了失败的评估结果。可能的原因是基于协同时尚图谱的用户偏好预测只是基于单品的种类信息进行的, 对于我们的个性化互补服装推荐任务会产生一定程度上的影响。

#### 4.5 模态比较

为了表明在时尚领域的个性化互补服装推荐中, 充分利用多模态数据的优势, 我们设计了模型 FG-PCCR 的 3 个变体: FG-PCCR-V, FG-PCCR-C 以及 FG-PCCR-G。上述这 3 个变体是在我们的模型 FG-PCCR 的基础上, 分别只考虑视觉模态, 文本模态和图的结构化模态所得到的模型。类似于 4.4.2 小节中的消融实验, 这部分的实验我们也是基于 FG-PCCR-5 所做, 因为它在前面部分的实验中展示出了最优的实验结果。

表 5 展示了不同的模态在 AUC 评估指标上的比较结果。从此表中, 我们可以获得以下观察结果: 1) 本文提出的 FG-PCCR 模型, 利用多模态数据信息的基础上, 在个性化互补服装推荐中实现了最好的实验效果。具体来说, 从 AUC 评估指标来看, 我们的方法得到的实验效果比 FG-PCCR-V, FG-PCCR-C 以及 FG-PCCR-G 的效果分别提高了 3.85%、3.48% 及 12.16%。这充分表明视觉信息、文本信息和图结构信息之间相互互补, 并联合促进个性化互补服装推荐的效果。2) 只考虑文本模态的 FG-PCCR-C 比其他两种分别只考虑视觉模态和图模态的模型, 即 FG-PCCR-V 和 FG-PCCR-G, 取得的效果更优。这说明在时尚领域的个性化互补服装推荐任务中, 文本模态的贡献更多。一个可能的解释是因为单品的文本信息通常更加具体, 能够清晰明了地传递出时尚单品的关键特征, 从而提升个性化互补服装推荐性能。

表 5 不同模态在 AUC 上的效果比较表

Table 5 Performance comparison on different modality

方法	AUC
FG-PCCR-V	0.8194
FG-PCCR-C	0.8231
FG-PCCR-G	0.7363
<b>FG-PCCR</b>	<b>0.8579</b>

#### 4.6 互补时尚单品检索应用

为了有效地评估我们所做工作的实际应用性,

我们通过互补的时尚单品检索任务来评估提出的 FG-PCCR 模型。类似于文献[68], 我们将每一个用户-上衣对  $(u_m, t_i)$  都被看作为一个查询。然后, 为每一个查询, 我们随机地选取  $T$  个下衣作为候选下衣集合。该集合包含一个正例下衣和  $T-1$  个通过随机选取的负例下衣。之后, 我们将这些候选下衣输入经过训练的模型之中, 然后计算得到其与给定的查询  $(u_m, t_i)$  的个性化互补服装兼容性评分  $s_{ij}^{(m)}$ 。最后, 根据该分数可以获得一个关于候选下衣的排序列表。在本工作的设置中, 我们聚焦于真实结果(即, 正例下衣)在所有候选下衣排序列表中的平均位置, 因此采用了 MRR 评估指标。

图 9 展示了不同方法在时尚单品检索中的 MRR 效果。整体而言, 在不同的下衣候选数目实验中, 本文提出的 FG-PCCR 模型相比较于其他的基准方法, 展示出了杰出的优越性。这样的结果再次证明我们提出的 FG-PCCR 方法在互补的时尚单品检索任务中的鲁棒性和有效性。

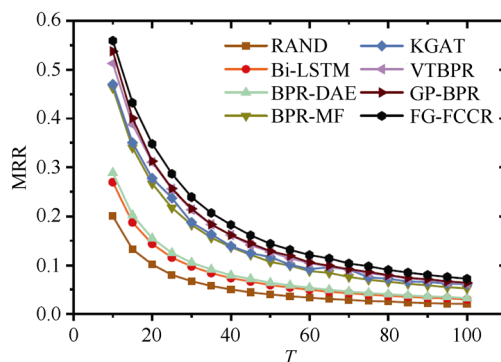


图 9 不同方法关于不同的下衣候选数量的 MRR 效果比较图

Figure 9 Performance of different methods in regard to MRR at different numbers of the bottom candidates.

另外, 协同的高阶交互建模模块对独立的一阶交互建模模块起到了强大的辅助作用, 为了对其有更深和更直观的理解, 本文可视化了互补的时尚单品检索中的部分实例的排序结果, 如图 10 所示。正如我们从图中所观察到的, 与忽略高阶交互建模的 FG-PCCR-woH 模型相比, 本文提出的 FG-PCCR 模型拥有巨大的优势, 它能够将互补的正例单品尽可能地排在前列。这再次验证了在时尚领域的个性化互补服装推荐中, 时尚图谱中的高阶交互关系起到了举足轻重的作用。

Query		1	2	3	4	5	6	7	8	9	10
 User1	FG-PCCR-woH										
	FG-PCCR										
 User2	FG-PCCR-woH										
	FG-PCCR										
 User3	FG-PCCR-woH										
	FG-PCCRF										

图 10 排序结果描述。绿色框中标出的下衣为正例

Figure 10 Illustration of the ranking results. The bottoms highlighted in the green boxes are the ground truth

5 总结与展望

5.1 总结

在本工作中，我们提出了一个新颖的基于时尚图谱增强的个性化互补服装推荐模型 FG-PCCR，该模型能够通过联合独立的一阶交互建模和协同的高阶交互建模促进个性化互补服装推荐效果。在我们研究任务中，目标是对给定的用户和目标上衣，搭配一件下衣。搭配的目标下衣不仅要和上衣搭配和谐，而且还要满足用户的偏好。

首先，独立的一阶交互建模涉及了用户-单品交互或者单品-单品交互这样的一阶关系，它包括主观的用户-单品偏好交互建模以及客观的单品-单品搭配交互建模。同时，时尚单品一般都有多种模态的信息，在本工作中，为了提高个性化服装搭配效果，我们通过神经网络和矩阵分解的方法对两种模态信息(即，视觉模态和文本模态)进行一阶交互建模。

其次，我们基于公开可用的数据集 IQON3000 构建了一个协同的时尚图谱并将其引入协同的高阶交互建模模块中。这是本文提出的一个主要创新点。该协同时尚图谱的实体节点包括用户、时尚单品以及单品的属性，关系边包括用户-单品偏好交互关系，单品-单品搭配交互关系以及单品-属性从属交互关

系。整体而言，构建的时尚图谱是一个协同图谱，既包括客观的知识(如，单品的属性)，又包括变化的交互(如单品-单品交互关系)。基于此协同时尚图谱，我们可以得到用户、单品及属性之间的高阶交互关系，同时，结合卷积神经网络 GCNs 进行了协同的高阶交互建模，进一步挖掘潜在的用户隐藏偏好及单品兼容性。

最后，我们在真实的数据集上做了大量的实验，对比了本工作所提出的FG-PCCR模型与最新的基准方法的效果，探究了协同的高阶交互建模模块对整个个性化互补服装推荐模型的影响，研究了不同模态对整个模型的作用，探索了FG-PCCR在时尚互补单品检索中的实际应用价值。总体而言，实验表明了个性化互补服装推荐中考虑时尚图谱的必要性以及验证了FG-PCCR模型的有效性。

5.2 展望

本文设计的基于时尚图谱增强的个性化互补服装推荐模型取得初步成效，但是仍然有诸多问题和挑战有待解决和探索，我们将持续跟进以下相关研究：

(1) 在本工作中，构建的协同时尚图谱只考虑了单品的属性信息，而忽略了用户的属性信息，比如用户的身高、体重、年龄、肤色、身材等信息。其

实, 用户属性也可以看作图谱中的实体节点, 使得构建的协同时尚图谱中的知识更丰富, 有助于模型学到更精准的用户偏好关系以及单品之间的兼容性关系。本工作未涉及的根本原因是已有的服装领域中可用于个性化互补服装推荐的相关数据集均没有用户的属性信息。在未来工作中, 我们打算进一步融入用户属性信息, 构建一个包括用户和时尚单品的属性信息的完整大规模数据集, 基于此数据集构建关系更丰富的协同时尚图谱, 从而提高个性化互补服装推荐的效果。

(2) 在本工作中, 我们将协同的高阶交互建模中的实体节点的所有邻居节点均平等对待, 从不同邻居获取的信息重要性都看作是一样的。然而实际上, 不同的邻居对于实体节点的贡献应该是不同的, 这样才能准确了解个性化服装搭配的本质原因。因此, 在未来的工作中, 可以增加注意力机制来区分不同邻居的贡献大小, 为协同时尚图谱中不同的交互关系分配重要权重, 可以确定最相关的方面, 从而解释个性化互补服装推荐的根本原因。故可解释性的基于时尚图谱增强的个性化互补服装推荐是本工作的另一个非常值得研究的方向。

**致 谢** 在此向对本文工作提出指导的老师、同学表示感谢, 以及对提出建议的评审专家表示感谢。

## 参考文献

- [1] Dong X, Song X M, Feng F L, et al. Personalized Capsule Wardrobe Creation with Garment and User Modeling[C]. *The 27th ACM International Conference on Multimedia*, 2019: 302-310.
- [2] Song X M, Han X J, Li Y K, et al. GP-BPR: Personalized Compatibility Modeling for Clothing Matching[C]. *The 27th ACM International Conference on Multimedia*, 2019: 320-328.
- [3] Yang X, Du X Y, Wang M. Learning to Match on Graph for Fashion Compatibility Modeling[J]. *The AAAI Conference on Artificial Intelligence*, 2020, 34(1): 287-294.
- [4] Li X C, Wang X, He X N, et al. Hierarchical Fashion Graph Network for Personalized Outfit Recommendation[C]. *The 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2020: 159-168.
- [5] Rendle S, Freudenthaler C, Gantner Z, et al. BPR: Bayesian Personalized Ranking from Implicit Feedback [C]. *In Conference on Uncertainty in Artificial Intelligence. AUAI Press*, 2009: 452-461.
- [6] Han X T, Wu Z X, Jiang Y Q, et al. Learning Fashion Compatibility with Bidirectional LSTMS[C]. *The 25th ACM international conference on Multimedia*, 2017: 1078-1086.
- [7] Song X M, Feng F L, Han X J, et al. Neural Compatibility Modeling with Attentive Knowledge Distillation[C]. *SIGIR '18: The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*, 2018: 5-14.
- [8] Song X M, Feng F L, Liu J H, et al. NeuroStylist: Neural Compatibility Modeling for Clothing Matching[C]. *The 25th ACM International Conference on Multimedia*, 2017: 753-761.
- [9] McAuley J, Targett C, Shi Q F, et al. Image-Based Recommendations on Styles and Substitutes[C]. *The 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2015: 43-52.
- [10] Veit A, Kovacs B, Bell S, et al. Learning Visual Clothing Style with Heterogeneous Dyadic Co-Occurrences[C]. *2015 IEEE International Conference on Computer Vision*, 2015: 4642-4650.
- [11] He R N, Packer C, McAuley J. Learning Compatibility across Categories for Heterogeneous Item Recommendation[C]. *2016 IEEE 16th International Conference on Data Mining*, 2016: 937-942.
- [12] Li Y C, Cao L L, Zhu J, et al. Mining Fashion Outfit Composition Using an End-to-End Deep Learning Approach on Set Data[J]. *IEEE Transactions on Multimedia*, 2017, 19(8): 1946-1955.
- [13] Yang X, Ma Y S, Liao L Z, et al. TransNFCM: Translation-Based Neural Fashion Compatibility Modeling[J]. *The AAAI Conference on Artificial Intelligence*, 2019, 33: 403-410.
- [14] Yang X, Ma Y S, Liao L Z, et al. TransNFCM: Translation-Based Neural Fashion Compatibility Modeling[J]. *The AAAI Conference on Artificial Intelligence*, 2019, 33: 403-410.
- [15] Zhou Z Z, Di X, Zhou W, et al. Fashion Sensitive Clothing Recommendation Using Hierarchical Collocation Model[C]. *The 26th ACM International Conference on Multimedia*, 2018: 1119-1127.
- [16] Chen C M, Tsai M F, Liu J Y, et al. Using Emotional Context from Article for Contextual Music Recommendation[C]. *The 21st ACM International Conference on Multimedia - MM'13*, 2013: 649-652.
- [17] He X N, Zhang H W, Kan M Y, et al. Fast Matrix Factorization for Online Recommendation with Implicit Feedback[C]. *The 39th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2016: 549-558.
- [18] Tran T, Lee K, Liao Y M, et al. Regularizing Matrix Factorization with User and Item Embeddings for Recommendation[C]. *The 27th ACM International Conference on Information and Knowledge Management*, 2018: 687-696.
- [19] Koren Y, Bell R, Volinsky C. Matrix Factorization Techniques for Recommender Systems[J]. *Computer*, 2009, 42(8): 30-37.
- [20] Hu Y, Yi X, Davis L S. Collaborative Fashion Recommendation: A Functional Tensor Factorization Approach[C]. *The 23rd ACM International Conference on Multimedia*, 2015: 129-138.
- [21] He R N, McAuley J. VBPR: Visual Bayesian Personalized Rank-



- ing from Implicit Feedback [C]. *The International Joint Conference on Artificial Intelligence. AAAI*, 2016:144–150.
- [22] Yao Z J, Sun Y F, Ding W C, et al. Dynamic Word Embeddings for Evolving Semantic Discovery[C]. *The Eleventh ACM International Conference on Web Search and Data Mining*, 2018: 673–681.
- [23] Wang X, He X N, Cao Y X, et al. KGAT: Knowledge Graph Attention Network for Recommendation[C]. *The 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2019: 950–958.
- [24] Wang H W, Zhang F Z, Xie X, et al. DKN: Deep Knowledge-Aware Network for News Recommendation[C]. *The 2018 World Wide Web Conference on World Wide Web - WWW'18*, 2018: 1835–1844.
- [25] Wang H W, Zhang F Z, Zhao M, et al. Multi-Task Feature Learning for Knowledge Graph Enhanced Recommendation[C]. *WWW'19: The World Wide Web Conference*, 2019: 2000–2010.
- [26] Qu Y Q, Liu J, Kang L Y, et al. Question Answering over Freebase via Attentive RNN with Similarity Matrix Based CNN[EB/OL]. 2018.
- [27] Shen Y, Deng Y, Yang M, et al. Knowledge-Aware Attentive Neural Network for Ranking Question Answer Pairs[C]. *SIGIR '18: The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*, 2018: 901–904.
- [28] Wang H W, Zhang F Z, Wang J L, et al. RippleNet: Propagating User Preferences on the Knowledge Graph for Recommender Systems[C]. *The 27th ACM International Conference on Information and Knowledge Management*, 2018: 417–426.
- [29] Xiao Yu, Xiang Ren, Yizhou Sun, et al. Personalized Entity Recommendation: a Heterogeneous Information Network Approach[C]. *In WSDM*, 2014:283–292.
- [30] Huang J, Zhao W X, Dou H J, et al. Improving Sequential Recommendation with Knowledge-Enhanced Memory Networks[C]. *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*, 2018: 505–514.
- [31] Zhang F Z, Yuan N J, Lian D F, et al. Collaborative Knowledge Base Embedding for Recommender Systems[C]. *The 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016: 353–362.
- [32] Sun Z, Yang J, Zhang J, et al. Recurrent Knowledge Graph Embedding for Effective Recommendation[C]. *The 12th ACM Conference on Recommender Systems*, 2018: 297–305.
- [33] Wang X, Wang D X, Xu C R, et al. Explainable Reasoning over Knowledge Graphs for Recommendation[J]. *The AAAI Conference on Artificial Intelligence*, 2019, 33: 5329–5336.
- [34] Hu B B, Shi C, Zhao W X, et al. Leveraging Meta-Path Based Context for Top- N Recommendation with a Neural Co-Attention Model[C]. *The 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2018: 1531–1540.
- [35] Xiao Yu, Xiang Ren, Quanquan Gu, et al. Collaborative Filtering with Entity Similarity Regularization in Heterogeneous information networks[C]. *IJCAI 27 (2013)*, 2013.
- [36] Cao Y X, Wang X, He X N, et al. Unifying Knowledge Graph Learning and Recommendation: Towards a Better Understanding of User Preferences[C]. *WWW'19: The World Wide Web Conference*, 2019: 151–161.
- [37] Ai Q Y, Azizi V, Chen X, et al. Learning Heterogeneous Knowledge Base Embeddings for Explainable Recommendation[J]. *Algorithms*, 2018, 11(9): 137.
- [38] Cao Y X, Liu Z Y, Li C J, et al. Multi-Channel Graph Neural Network for Entity Alignment[C]. *The 57th Annual Meeting of the Association for Computational Linguistics*, 2019: 1452–1461.
- [39] Scarselli F, Gori M, Tsoi A C, et al. The Graph Neural Network Model[J]. *IEEE Transactions on Neural Networks*, 2009, 20(1): 61–80.
- [40] Kipf T N, Welling M. Semi-Supervised Classification with Graph Convolutional Networks[EB/OL]. 2016.
- [41] William L. Hamilton, Zitao Ying, et al. Inductive Representation Learning on Large Graphs[C]. *In Annual Conference on Neural Information Processing Systems*, 2017:1024–1034.
- [42] Li Y J, Tarlow D, Brockschmidt M, et al. Gated Graph Sequence Neural Networks[EB/OL]. 2015.
- [43] Chen X L, Song X M, Ren R Y, et al. Fine-Grained Privacy Detection with Graph-Regularized Hierarchical Attentive Representation Learning[J]. *ACM Transactions on Information Systems*, 2020, 38(4): 1–26.
- [44] Battaglia P W, Pascanu R, Lai M, et al. Interaction Networks for Learning about Objects, Relations and Physics [C]. *In Annual Conference on Neural Information Processing Systems*, 2016: 4502–4510.
- [45] Sanchez-Gonzalez A, Heess N, Springenberg J T, et al. Graph Networks As Learnable Physics Engines for Inference and Control[EB/OL]. 2018.
- [46] Hong R C, Li L, Cai J J, et al. Coherent Semantic-Visual Indexing for Large-Scale Image Retrieval In the Cloud[J]. *IEEE Transactions on Image Processing*, 2017, 26(9): 4128–4138.
- [47] Hong R C, Yang Y, Wang M, et al. Learning Visual Semantic Relationships for Efficient Visual Retrieval[J]. *IEEE Transactions on Big Data*, 2015, 1(4): 152–161.
- [48] Chen L, Zhang H W, Xiao J, et al. Counterfactual Critic Multi-Agent Training for Scene Graph Generation[C]. *2019 IEEE/CVF International Conference on Computer Vision*, 2019: 4612–4622.
- [49] Sun G L, Cheng Z Q, Wu X, et al. Personalized Clothing Recommendation Combining User Social Circle and Fashion Style Consistency[J]. *Multimedia Tools and Applications*, 2018, 77(14): 17731–17754.
- [50] Andrew L Maas, Awni Y Hannun, and Andrew Y Ng. Rectifier

- Nonlinearities Improve Neural Network Acoustic Models[C]. In *Proc. ICML*, Vol. 30. 3.2013.
- [51] Cao D, Nie L Q, He X N, et al. Embedding Factorization Models for Jointly Recommending Items and User Generated Lists[C]. *The 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2017: 585-594.
- [52] Liu S Y, Wu Q, Miao C Y. Personalized Recommendation Considering Secondary Implicit Feedback[C]. *2018 IEEE International Conference on Agents*, 2018: 87-92.
- [53] Zhang H W, Zha Z J, Yang Y, et al. Attribute-Augmented Semantic Hierarchy: Towards Bridging Semantic Gap and Intention Gap In Image Retrieval[C]. *The 21st ACM International Conference on Multimedia*, 2013: 33-42.
- [54] Jiang L, Yu S I, Meng D Y, et al. Fast and Accurate Content-Based Semantic Search In 100M Internet Videos[C]. *The 23rd ACM International Conference on Multimedia*, 2015: 49-58.
- [55] Song X M, Nie L Q, Zhang L M, et al. Multiple Social Network Learning and Its Application In Volunteerism Tendency Prediction[C]. *The 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2015: 213-222.
- [56] Xuemeng Song, Liqiang Nie, Luming Zhang, et al. Interest Inference via Structure Constrained Multi-Source Multi-Task Learning[C]. In *International Joint Conference on Artificial Intelligence. AAAI Press*, 2015:2371-2377.
- [57] Hu Z T, Ma X Z, Liu Z Z, et al. Harnessing Deep Neural Networks with Logic Rules [C]. In *Annual Meeting of the Association for Computational Linguistics. The Association for Computer Linguistics(1)*.2016.
- [58] Severyn A, Moschitti A. Twitter Sentiment Analysis with Deep Convolutional Neural Networks[C]. *The 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2015: 959-962.
- [59] Shinnou H, Asahara M, Komiya K, et al. Nwjc2vec: Word Embedding Data Constructed from NINJAL Web Japanese Corpus[J]. *Journal of Natural Language Processing*, 2017, 24(5): 705-720.
- [60] Khosla A, Das Sarma A, Hamid R. What Makes an Image Popular? [C]. *The 23rd International Conference on World Wide Web - WWW'14*, 2014: 867-876.
- [61] Lim J H, Japar N, Ng C C, et al. Unprecedented Usage of Pre-Trained CNNs on Beauty Product[C]. *The 26th ACM International Conference on Multimedia*, 2018: 2068-2072.
- [62] Liu M, Nie L Q, Wang M, et al. Towards Micro-Video Understanding by Joint Sequential-Sparse Modeling[C]. *The 25th ACM International Conference on Multimedia*, 2017: 970-978.
- [63] He K M, Zhang X Y, Ren S Q, et al. Deep Residual Learning for Image Recognition[C]. *2016 IEEE Conference on Computer Vision and Pattern Recognition*, 2016: 770-778.
- [64] Kingma D P, Ba J L. Adam: A Method for Stochastic Optimization [C]. In *International Conference on Learning Representations*. 2015:arXiv preprint arXiv:1412.6980.
- [65] Liu J H, Song X M, Chen Z M, et al. MGCM: Multi-Modal Generative Compatibility Modeling for Clothing Matching[J]. *Neurocomputing*, 2020, 414: 215-224.
- [66] Chen Z M, Wei X S, Wang P, et al. Multi-Label Image Recognition with Graph Convolutional Networks[C]. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019: 5172-5181.
- [67] Li Q M, Han Z C, Wu X M. Deeper Insights into Graph Convolutional Networks for Semi-Supervised Learning[EB/OL]. 2018
- [68] He X N, Liao L Z, Zhang H W, et al. Neural Collaborative Filtering[C]. *The 26th International Conference on World Wide Web*, 2017: 173-182.



**史金婉** 于 2018 年在青岛科技大学计算机科学与技术专业获得学士学位。现在山东大学计算机技术专业攻读硕士学位。研究兴趣包括: 信息检索和多媒体处理等领域。Email: shijinwann@gmail.com



**宋雪萌** 于 2016 年在新加坡国立大学计算机专业获得博士学位。现任山东大学计算机科学与技术学院副教授、硕士生导师。研究兴趣包括信息检索和多媒体处理等领域。Email: songxuemeng@gmail.com



**刘子鑫** 于 2018 年在西安邮电大学计算机科学与技术专业获得学士学位。现在山东大学计算机技术专业攻读硕士学位。研究兴趣包括: 信息检索和多媒体处理等领域。Email: zixin34728@mail.sdu.edu.cn



**聂礼强** 于 2013 年在新加坡国立大学计算机专业获得博士学位。现任山东大学计算机科学与技术学院教授、博士生导师。研究兴趣包括信息检索和多媒体处理等领域。Email: nieliqiang@gmail.com