



计算机科学

COMPUTER SCIENCE

MB-ATMK:融合属性权重和时序元知识的多行为序列推荐模型

陈毓哲, 曹琼, 黄贤英, 邹世豪

引用本文

陈毓哲, 曹琼, 黄贤英, 邹世豪. MB-ATMK:融合属性权重和时序元知识的多行为序列推荐模型[J]. 计算机科学, 2024, 51(11A): 231100047-9.

CHEN Yuzhe, CAO Qiong, HUANG Xianying, ZOU Shihao. MB-ATMK:Multi-behavior Sequential Recommendation Integrating Attribute Weights and Temporal Meta-knowledge [J]. Computer Science, 2024, 51(11A): 231100047-9.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[基于图神经网络的银行交易欺诈检测方法](#)

Bank Transaction Fraud Detection Method Based on Graph Neural Network

计算机科学, 2024, 51(11A): 240200024-8. <https://doi.org/10.11896/jsjcx.240200024>

[FCTNet:基于双城深度学习的公交车到站时间预测方法](#)

FCTNet:Bus Arrival Time Prediction Method Based on Dual Domain Deep Learning

计算机科学, 2024, 51(11A): 231000180-7. <https://doi.org/10.11896/jsjcx.231000180>

[基于季节分解的混合神经网络的时间序列预测](#)

Time Series Prediction of Hybrid Neural Networks Based on Seasonal Decomposition

计算机科学, 2024, 51(11A): 231200008-7. <https://doi.org/10.11896/jsjcx.231200008>

[基于相对位置编码转换器模块的深度步态识别网络](#)

Deep Gait Recognition Network Based on Relative Position Encoding Transformer

计算机科学, 2024, 51(11A): 240400064-6. <https://doi.org/10.11896/jsjcx.240400064>

[基于改进超像素采样的立体匹配网络](#)

Stereo Matching Network Based on Enhanced Superpixel Sampling

计算机科学, 2024, 51(11A): 231100005-7. <https://doi.org/10.11896/jsjcx.231100005>

MB-ATMK:融合属性权重和时序元知识的多行为序列推荐模型

陈毓哲 曹琼 黄贤英 邹世豪

重庆理工大学计算机科学与工程学院 重庆 400054

(chenyuzhe5233@163.com)

摘要 序列推荐根据用户和项目的交互序列预测用户未来的偏好,然而现有的方法忽略了在现实场景中用户的多行为交互(如浏览、收藏、加入购物车)。其次,用户的偏好有着时序依赖性,同时也受到属性信息的影响。最后,在多行为序列推荐场景中用户的多行为交互存在复杂依赖关系。因此我们提出了一种融合属性权重和时序元知识的多行为序列推荐模型(MB-ATMK)。首先加入用户的多行为交互数据,并基于用户交互的时间戳设计了时序感知编码模块,通过时序感知注意力捕获了用户的动态偏好。其次引入了用户端和项目端丰富的属性信息,设计了属性权重增强的元知识图神经网络。使用元知识提炼了用户的多偏好模式,并基于图神经网络设计了属性权重注意力机制,增强了模型对用户细粒度偏好的捕获。最后提出了包含多行为权重生成模块和偏好迁移网络的元知识预测层,通过生成定制的元知识捕获了用户的跨行为依赖。在两个数据集上进行的大量实验验证了所提模型的有效性和优越性。

关键词: 序列推荐;多行为推荐;图神经网络;注意力机制;属性信息

中图分类号 TP391

MB-ATMK: Multi-behavior Sequential Recommendation Integrating Attribute Weights and Temporal Meta-knowledge

CHEN Yuzhe, CAO Qiong, HUANG Xianying and ZOU Shihao

College of Computer Science and Engineering, Chongqing University of Technology, Chongqing 400054, China

Abstract Sequential recommendation predicts users' future preferences based on the sequence of interactions between users and items. However, existing methods often overlook the multi-behavior interactions (such as page view, favorite, add to cart) in real-world scenarios. Additionally, users' preferences not only depend on temporal sequences but are also influenced by attribute information. Lastly, in the scenario of multi-behavior sequence recommendation, users' multi-behavior interactions exhibit complex dependencies. Therefore, this paper proposes a multi-behavior sequence recommendation model with attribute weights and temporal meta-knowledge (MB-ATMK). Firstly, we incorporate users' multi-behavior interaction data and design a temporal-aware encoding module based on the timestamps of user interactions to capture users' dynamic preferences through temporal-aware attention. Secondly, we introduce rich attribute information on both the user and item sides and design an attribute-weighted meta-knowledge graph neural network. Using meta-knowledge, we refine users' multi-preference patterns and design an attribute-weighted attention mechanism based on graph neural networks to enhance the model's capture of users' fine-grained preferences. Finally, we propose a meta-knowledge prediction layer that includes a multi-behavior weight generation module and a preference transfer network, capturing users' cross-behavior dependencies through generated customized meta-knowledge. Extensive experiments on two datasets validate the effectiveness and superiority of the proposed model.

Keywords Sequential recommendation, Multi-behavior recommendation, Graph neural network, Attention mechanism, Attribute information

1 引言

随着大数据时代的到来,互联网上的数据呈爆炸式增长。推荐系统作为满足用户个性化兴趣和减轻信息过载的工具,在许多在线平台(例如在线广告、在线零售以及视频推荐、音乐推荐)发挥着巨大的作用^[1-2]。传统推荐系统使用协同过滤方法建模用户和项目交互,但忽略了用户和物品交互时的

序列信息。因此序列推荐应运而生,也引起了越来越多研究者的关注^[3]。序列推荐的核心挑战是学习用户交互序列的高阶依赖关系,进而捕获用户的偏好^[4-5]。

近几年的研究中,已经提出了许多方法来增强序列推荐模型对用户偏好的捕获。具体来说,早期的工作旨在在用户复杂的交互序列中捕获用户的整体偏好,进而编码用户的嵌入向量^[6-7]。而随着神经网络技术的快速发展,最近的工作已

基金项目:国家自然科学基金(62141201);重庆市自然科学基金(CSTB2022NSCQ-MSX1672)

This work was supported by the National Natural Science Foundation of China(62141201) and Natural Science Foundation of Chongqing, China (CSTB2022NSCQ-MSX1672).

通信作者:曹琼(jing5589@163.com)

使用了许多神经网络技术去增强序列推荐模型的性能^[8-9]。例如, DGSR^[10]通过动态图结构连接不同用户的顺序,探索具有时间和顺序信息的用户和项目的交互行为;同时设计了动态图注意力神经网络来实现不同用户及其序列在动态图中的信息传播和聚合。MGNN-Spred^[11]通过图神经网络学习项目间的顺序关系,然后使用门控循环单元为用户的交互分配权重。

虽然上述方法已取得了较好的结果,但仍存在挑战。首先,上述方法只考虑了用户单行为下的数据(例如电商平台中的购买行为),而忽略了用户的多种行为(如页面浏览、收藏、加入购物车等)。当只考虑用户的单行为时,由于数据稀疏问题,模型不能有效地反映用户真实的个性化偏好,而用户与商品的多行为交互会在一定程度上反映用户的意图,进一步丰富用户的偏好。如图1所示,用户 u_1 最终的购买行为受到不同行为的影响,包括浏览的电脑、运动鞋、手机,收藏的平板电脑以及加入到购物车的电脑。其次,用户的多行为交互存在复杂的依赖关系,一些用户倾向于浏览商品后进行购买,而另一些用户则倾向于将商品添加到购物车后进行购买,每种行为对用户的重要性不同。为了捕获用户的真实偏好,需要在多行为交互场景下提炼用户个性化多偏好模式的方法。

另一方面,虽然部分研究将用户的多行为交互加入到序列推荐中^[12],但他们并未同时考虑用户与项目交互的时序信息和用户与项目丰富的属性信息。首先,用户的偏好有着时序依赖性。用户在不同的时间段与不同的项目发生交互,其偏好会不断发生变化。在图1中,用户 u_2 在 t_1 时刻浏览的裙子、收藏手提包行为(短期行为)以及 t_2 时刻进行的浏览裤子和把深蓝色裙子加入购物车行为(长期行为)共同影响了 t_3 时刻发生的购买行为。其次,用户和项目有着丰富的属性信息。由于用户的性别、职业和年龄的不同,项目的外观、品牌和价格的不同,用户和项目之间也会存在不同的属性依赖关系,因此引入属性信息能够更精确地反映用户的偏好和意图。

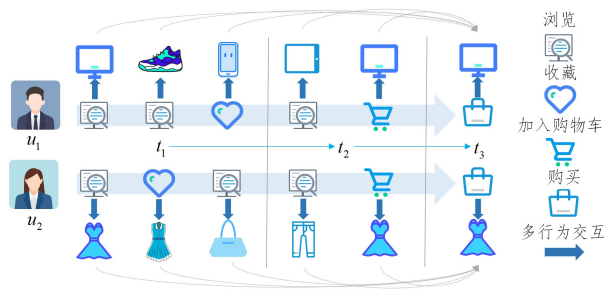


图1 融合时序与属性信息的用户多行为序列推荐示例

Fig. 1 Example of user multi-behavior sequential recommendation integrating time sequence and attribute information

为了应对上述挑战,本文提出了一种融合属性权重和时序元知识的多行为序列推荐模型(Multi-Behavior Sequential Recommendation integrating Attribute Weights and Temporal Meta-Knowledge, MB-ATMK)。首先,加入用户的多行为交互数据以及用户端和项目端丰富的属性信息,缓解了序列推荐中的数据稀疏问题;并基于用户交互的时间戳设计了时序感知编码模块,通过生成初始时序嵌入并使用时序感知注意力机制捕获了用户随时间变化的动态偏好。其次,设计了属性权重增强的元知识图神经网络。使用元知识生成器生成包

含用户时序偏好和多行为语义的元知识,提炼出用户的个性化多偏好模式;并基于图神经网络设计了属性权重注意力机制,通过在每层图神经网络中生成属性权重,增强了模型对属性信息影响用户细粒度偏好的捕获。最后,提出了包含多行为权重生成模块和偏好迁移网络的元知识预测层,通过将多行为嵌入加权聚合生成目标行为嵌入和辅助行为嵌入,进而生成定制的元知识。为了将辅助行为中具有正向影响的用户偏好迁移至目标行为嵌入,我们将元知识转换为用于预测层的可学习参数矩阵,捕获了用户的跨行为依赖,有效提升了模型的性能。

本文的贡献总结如下:

(1)聚焦于用户的个性化偏好,通过加入时序感知编码和属性权重注意,捕获了多行为下用户-项目交互的时序嵌入和属性增强的用户-项目多行为嵌入,反映了细粒度的用户个性化偏好。

(2)提出一种新的 MB-ATMK 模型,将用户的多行为交互数据引入序列推荐中作为辅助信息,多角度丰富了用户的偏好。在用户多行为编码的基础上根据基于优化方法的元学习设计了元知识图神经网络,将包含了用户多偏好模式的元知识作为参数矩阵,用于图神经网络。同时,提出了多行为权重生成模块以及偏好迁移网络,捕获了用户的跨行为依赖。

(3)在两个公共的电商数据集 Tmall 和 Fliggy Trip 进行了一系列的实验,实验表明,所提 MB-ATMK 在许多推荐基线上取得了显著的进步。

2 相关工作

2.1 序列推荐

与一般推荐相比,序列推荐的场景有所不同,它的主要任务被简化为利用用户的历史交互序列来预测用户未来的偏好。例如, Bert4rec^[3]利用距离嵌入将每个 POI(Point of Interest)映射到一个统一的隐含空间中,采用马尔可夫链模型预测 POI 的变化,用两个 POI 之间的距离衡量两者的序列关系。最近,一些研究将序列数据转换为图结构,利用图神经网络来实现序列推荐。例如, TASRec^[13]构建了会话图和时序图,在会话图上使用门控图神经网络捕获会话内的项目交互,在时序图中使用图卷积网络学习时间增强的项目表示。SR-GNN^[14]将序列数据转换为图结构数据,使用图神经网络捕捉复杂项目转移,并使用注意力网络建模用户的全局偏好和当前兴趣偏好。

2.2 多行为推荐

早期的研究使用矩阵分解的方法从多行为交互数据中挖掘用户偏好,例如 BF^[15]构建并分解了多个矩阵,包括非特定行为的用户主题矩阵、单一行为的用户主题矩阵和组合的特定行为主题矩阵,学习隐含的嵌入空间,通过估计用户对主题的偏好来模拟用户偏好。由于图卷积神经网络通过对图结构数据进行高阶特征提取,使用邻居节点加权聚合机制生成高质量的用户项目表征,因此,一些研究人员使用图神经网络模型来捕捉用户和项目之间多种类型的交互。例如 EHCF^[16]通过构建异构行为的关联预测来捕捉不同行为之间的复杂关系,同时采用非采样策略优化模型,能够以较低的时间复杂度从整个异构数据中高效地学习模型参数。但是它们缺乏对用户-项目异构交互的高阶协同信息的编码。GHCF^[17]将节点

(用户和项目)嵌入表示和关系嵌入表示联合嵌入到多关系预测中,探索高阶异构用户-项目交互。

2.3 元学习

元学习的理念是“学会学习”,其致力于训练能够快速适应新任务的模型。通常,元学习的目标是在各种学习任务上训练模型,以便泛化到新的学习任务。元学习主要有3种方法:(1)基于黑盒方法的元学习目标是训练黑盒元学习器(例如神经网络)并通过监督学习^[18-20]输出任务的模型参数。(2)基于优化方法的元学习目标是训练模型的初始参数,使得模型在通过使用新任务的数据计算的一个或多个梯度步骤更新参数后,提升模型的性能。例如MAML^[21]通过训练一组初始化参数,使用少量数据使得模型快速收敛。MetaR^[22]和GANA^[23]利用基于优化的方法来快速适应关系的嵌入。最近也有一些工作将基于优化方法的元学习与图神经网络结合,例如L2P-GNN^[24]使用元学习对GNN进行预训练,使GNN处理下游任务的新图,而MI-GNN^[25]在元学习范式下为每个图定制归纳模型,用于实现跨图的归纳节点分类。(3)基于度量方法的元学习目标是学习可以在所有分类任务中泛化的通用匹配度量^[26]。例如GMatching^[27]通过图神经网络生成度量结构和嵌入。

3 问题定义

在本文的序列推荐场景中,给定用户集合 $U = \{u_1, \dots, u_m, \dots, u_M\}$ 和项目集合 $I = \{i_1, \dots, i_n, \dots, i_N\}$,其中 M (由 m 索引)和 N (由 n 索引)分别表示用户和项目的数量。同时,引入并定义不同行为(如页面浏览、收藏、加入购物车和购买),表示为 $r = \{v, p, c, f\}$,其中 v, p, c, f 表示浏览、购买、加入购物车行为和购买行为。考虑到用户与项目进行了多行为的

交互,我们进一步做出定义:

定义1(行为交互序列 Y_u) 为了反映用户多行为的交互,每个行为的交互序列 Y_u 由一个三元组 (u_m, i_n, r) 组成,其中用户 $u_m \in U$,项目 $i_n \in N$ 。为了更好地捕捉用户的个性化偏好,在多行为序列推荐场景中,我们将要预测的行为(购买)视为目标行为 \bar{T} ,其他行为(如页面浏览、收藏、加入购物车)视为辅助行为 \bar{A} 。

定义2(融合属性的用户项目多行为交互图) 由于属性信息反映了用户和项目的独特属性,如年龄、性别、品牌和价格等。通过加入属性信息,反映了用户的细粒度个性偏好,缓解了推荐系统中的冷启动问题。我们将多行为用户项目交互序列转化为图结构定义为 $G_r = (V, E_r)$,其中 V 表示包括了用户和项目的节点集, E_r 表示包括用户的多种行为的边集。本文的研究问题可以形式化为:

输入:由用户 U 、项目 I 和行为 r 组成的多行为交互序列 Y_u 。

输出:一个用来预测用户 u_m 在目标行为下与项目 i_n 交互的学习函数 $\hat{y}_{u,i}^{\bar{T}}$ 。

4 模型

本文提出的MB-ATMK模型结构如图2所示,由3个关键模块组成:(1)时序感知元知识编码层,生成包含用户多偏好模式的元知识,作为元知识图神经网络的输入;(2)属性权重增强的元知识图神经网络,在图神经网络的迭代聚合中融合属性信息的权重,生成具有高阶连通性的用户和项目的多行为嵌入;(3)融合行为权重的元知识预测层,将用户和项目的多行为嵌入转化为元知识进行模型的预测。

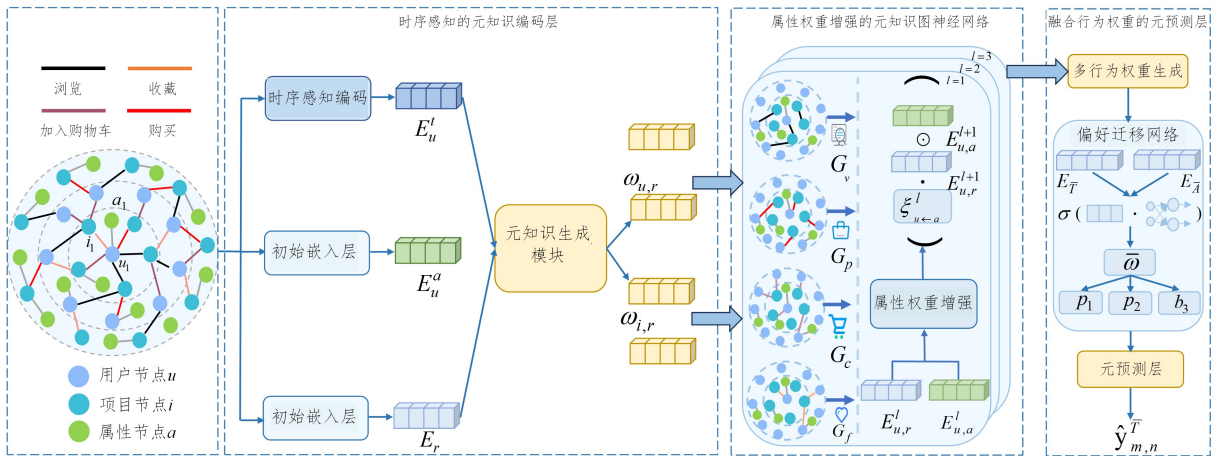


图2 MB-ATMK结构图

Fig. 2 Structure of MB-ATMK

4.1 时序感知元知识编码层

相较于只使用单一行为的序列推荐系统,为了捕获用户的动态偏好,我们设计了时序感知的多行为编码模块并使用元知识生成器提炼了用户的多偏好交互模式。首先通过初始嵌入层得到用户和项目的初始嵌入以及多行为的嵌入,并通过时序感知编码得到时序嵌入,进而使用元知识生成器将多行为交互场景下用户的时序偏好和行为语义注入到用户和项目的初始嵌入中。时序感知元知识编码由3个部分组成:初始嵌入层、时序感知编码、元知识生成器。

4.1.1 初始嵌入层

在基于图神经网络的推荐系统中,每个用户和项目都与一个ID嵌入相关联。因此我们使用初始嵌入层,根据用户和项目的ID、模型的嵌入维度 d 和行为的数量使用Xavier初始化器生成用户初始嵌入序列 $\mathbf{E}_u = \{E_{u_1}, \dots, E_{u_m}, \dots, E_{u_M}\}$,项目初始嵌入序列 $\mathbf{E}_i = \{E_{i_1}, \dots, E_{i_n}, \dots, E_{i_N}\}$ 和多行为嵌入序列 $\mathbf{E}_r = \{E_{r_1}, E_{r_2}, E_{r_3}, E_{r_p}\}$ 。同时,针对引入的属性信息,使用生成用户和项目的初始属性嵌入 $\mathbf{E}_{u,a}, \mathbf{E}_{i,a}$ 。

4.1.2 时序感知编码

在多行为交互场景中,用户的个性化偏好随时间而变化,用户最终的目标行为(购买)受先前时刻的多行为交互共同影响。如图3所示,用户 u_1 在 $t_{u_1,i_1,v}$ 时刻浏览了项目 i_1 ,在 $t_{u_1,i_2,p}$ 时刻购买了项目 i_2 。

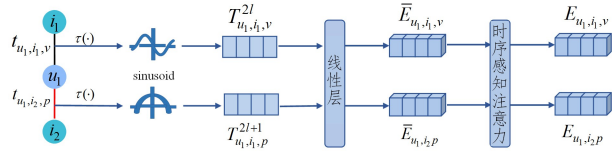


图3 时序感知编码

Fig. 3 Temporal-aware encoding

相较于只使用时间戳生成时间嵌入的KHGT^[28],我们基于用户交互的时间戳和注意力机制设计了时序感知编码模块,首先通过时隙映射函数 $\tau(\cdot)$ 以及正余弦函数生成时序向量 $T_{u,i,r}$,计算如式(1)所示:

$$\begin{aligned} T_{u,i,r}^{(2l)} &= \sin(\tau(t_{u,i,r}) \cdot \bar{w}_1) \\ T_{u,i,r}^{(2l+1)} &= \cos(\tau(t_{u,i,r}) \cdot \bar{w}_2) \\ \bar{w}_1 &= \frac{1}{10000 \frac{2l}{d_t}}; \bar{w}_2 = \frac{1}{10000 \frac{2l+1}{d_t}} \end{aligned} \quad (1)$$

其中, \bar{w}_1, \bar{w}_2 是参数矩阵, d_t 是时序嵌入维度。为了将得到的时序向量与用户的初始嵌入和项目的初始嵌入对齐,将生成的时序向量输入线性层,进而生成特定行为下用户与项目的时间嵌入,计算式如式(2)所示:

$$\begin{aligned} \bar{E}_{u_1,i_1,v} &= \text{Linear}(T_{u_1,i_1,v}) \\ \bar{E}_{u_1,i_2,p} &= \text{Linear}(T_{u_1,i_2,p}) \end{aligned} \quad (2)$$

为了关注用户在不同的时间间隔发生的多行为交互,捕获用户随时间变化的动态偏好,我们基于多头注意力机制设计了时序感知注意力。首先将时序嵌入与用户和项目的初始嵌入拼接: $Z_{u_1} = \bar{E}_{u_1,i_1,v} \parallel E_{u_1}$; $Z_{i_1} = \bar{E}_{u_1,i_1,v} \parallel E_{i_1}$,将拼接后嵌入投影到 query key value 向量中: $Q = Z_u W_Q, K = Z_u W_K, V = Z_u W_V$,其中 $W_{(\cdot)}$ 是可训练的参数矩阵。时序感知注意力的计算式如式(3)所示:

$$Y_T = \text{softmax}\left(\frac{(Z_u Q)(Z_u K)}{\sqrt{D}}\right)(Z_u V) \quad (3)$$

其中, Y_T 为每层时序感知注意力的权重, D 为嵌入对齐后的维度。在这种方式下,通过前馈神经网络联合得到多头注意力层,进而生成用户的时间嵌入,相同的操作应用于项目端,计算式如式(4)所示:

$$E_{u,r} = \text{FFN}(Y_T \cdots \parallel E_u \cdots \parallel Y_T^K) \quad (4)$$

其中, $\text{FFN}(\cdot) = xW_0 + b_0, W_0$ 和 b_0 为可训练参数。

4.1.3 元知识生成器

元知识生成器使用经过上述操作得到的两种嵌入(时序嵌入和多行为嵌入),生成了融合时序性和多行为语义的元知识,计算式如式(5)所示:

$$\omega_{u,r} = E_{u,r} \parallel E_r \quad (5)$$

$$\omega_{i,r} = E_{i,r} \parallel E_r$$

最后,我们在生成的用户元知识 $\omega_{u,r}$ 和项目元知识 $\omega_{i,r}$ 上加入一个线性层,使得最后生成的元参数矩阵 \mathcal{F}_u 和 \mathcal{F}_i 可用于元知识图神经网络,计算式如式(6)所示:

$$\mathcal{F}_u = \omega_{u,r} \cdot w_1, \mathcal{F}_i = \omega_{i,r} \cdot w_2 \quad (6)$$

其中, w_1 和 w_2 是可训练的参数矩阵。

4.2 属性权重增强的元知识图神经网络

在多行为推荐场景中,用户的多行为交互存在复杂的依赖关系,每个行为都有自身特点。同时,用户和项目包含丰富的属性信息,用户的多行为交互受属性信息所影响。图卷积神经网络(GCN)能够捕捉用户和项目之间的复杂交互模式,通过在图结构中聚合邻居节点的信息,可以有效地传递和利用图中的用户的多行为交互数据,从而缓解数据稀疏问题。同时,利用GCN的迭代聚合操作将属性信息更有效地融合到节点的初始特征中,从而获得更丰富的用户和项目嵌入表示。

为此,我们基于图神经网络设计了属性权重增强的元知识图神经网络,将包含用户多偏好模式的元知识应用于去除了非线性激活的轻量级图神经网络。在4个行为交互图(G_u, G_f, G_i, G_p)中迭代聚合生成具有高阶连通性的用户-项目的多行为嵌入。具体来说,属性权重增强的元知识图神经网络首先基于第3节的多行为用户项目交互图结构,使用第4.1节的时序感知元知识编码层得到第4.2节的属性权重增强的元知识图神经网络的输入,使用第4.2.1节的属性权重注意力生成用户和属性的权重。然后在第4.2.2节的权重增强的多行为编码中进行GCN中的迭代聚合操作,更新用户和项目的多行为嵌入。最后,在第4.2.3节的高阶多行为嵌入聚合中得到GCN的输出。

4.2.1 属性权重注意力

在互联网平台中,用户-项目有着丰富的属性信息,如用户的年龄、性别,项目的品牌、外观。不同属性的用户在面对相同的项目中会反映不同的偏好,为了充分挖掘用户受属性信息所影响的个性化偏好,我们设计了属性权重注意力机制。将元知识参数矩阵注入到用户和项目的初始嵌入中,得到元知识图神经网络中用户和项目的初始嵌入,计算式如式(7)所示:

$$\begin{aligned} E_{u,r}^l &= E_u \parallel E_u \cdot \mathcal{F}_u \\ E_{i,r}^l &= E_i \parallel E_i \cdot \mathcal{F}_i \end{aligned} \quad (7)$$

其中, l 为图神经网络的层数且 $l=0$ 。在得到图神经网络初始嵌入后,使用属性权重注意力学习不同属性的重要性,进而在每一层图神经网络中生成当前层的用户属性权重 $\xi_{u \leftarrow a}^l$ 和项目属性权重 $\xi_{i \leftarrow a}^l$,计算式如式(8)所示:

$$\begin{aligned} \xi_{u \leftarrow a}^l &= \text{softmax}(\sigma(\mathbf{W}^l(E_{u,r}^l \parallel E_{u,a}^l + b^l))) \\ \xi_{i \leftarrow a}^l &= \text{softmax}(\sigma(\mathbf{W}^l(E_{i,r}^l \parallel E_{i,a}^l + b^l))) \end{aligned} \quad (8)$$

其中, ξ^l 为在第 l 层图卷积时属性信息的权重, $u \leftarrow a$ 表示用户的属性权重, σ 为激活函数, \mathbf{W}^l 和 b^l 为第 l 层图卷积时的权重矩阵和偏置。

4.2.2 权重增强多行为嵌入编码

基于当前图神经网络层的用户属性权重和项目属性权重,在融合属性的用户项目多行为交互图中迭代聚合生成具有高阶连通性的用户和项目的多行为嵌入,计算式如式(9)所示:

$$\begin{aligned} E_{u,r}^{(l+1)} &= \sum_{a \in N_u^a} \xi_{u \leftarrow a}^l (E_{u,r}^l \odot E_{u,a}^l, E_{i,r}^l) \\ E_{i,r}^{(l+1)} &= \sum_{a \in N_i^a} \xi_{i \leftarrow a}^l (E_{i,r}^l \odot E_{i,a}^l, E_{u,r}^l) \end{aligned} \quad (9)$$

其中, N_u^a 为用户的邻居属性节点。在融合属性的用户项目多行为交互图中,每个用户和项目的节点都与其属性节点相连。

为此,在每个元知识图神经网络层更新用户和项目多行为嵌入的同时,通过聚合属性节点相连的用户节点嵌入或项目的节点嵌入来更新用户和项目的属性嵌入,计算式如式(10)所示:

$$\mathbf{E}_{u,a}^{(l+1)} = \frac{1}{|N_u^a|} \sum_{u \in N_u^a} \mathbf{E}_{u,r}^l; \mathbf{E}_{i,a}^{(l+1)} = \frac{1}{|N_i^a|} \sum_{i \in N_i^a} \mathbf{E}_{i,r}^l \quad (10)$$

4.2.3 高阶多行为嵌入聚合

通过 L 次迭代传播后,获得了基于属性权重增强的用户多行为嵌入序列 $\{\mathbf{E}_{u,r}^1, \dots, \mathbf{E}_{u,r}^L, \dots, \mathbf{E}_{u,r}^L\}$ 和项目的多行为嵌入序列 $\{\mathbf{E}_{i,r}^1, \dots, \mathbf{E}_{i,r}^L, \dots, \mathbf{E}_{i,r}^L\}$ 。用户和项目最终的多行为嵌入的计算式如式(11)所示:

$$\mathbf{E}_{u,r} = \sum_l \alpha_l \mathbf{E}_{u,r}^l; \mathbf{E}_{i,r} = \sum_l \alpha_l \mathbf{E}_{i,r}^l \quad (11)$$

其中, α_l 是一个超参数,我们将它设置为 $1/L+1$ 。受 LightGCN^[29] 的启发,相较于传统的图卷积神经网络,我们在图卷积过程中去除了非线性激活,提升了模型预测的精确性。通过属性权重增强的元知识图神经网络模块,获得了通过属性增强的用户和项目多行为嵌入序列 $\{\mathbf{E}_{u_1,r}, \dots, \mathbf{E}_{u_m,r}, \dots, \mathbf{E}_{u_m,r}\}; \{\mathbf{E}_{i_1,r}, \dots, \mathbf{E}_{i_n,r}, \dots, \mathbf{E}_{i_n,r}\}$ 。

4.3 融合行为权重的元知识预测层

属性权重增强的元知识图神经网络实现了用户与项目之间的信息传播与聚合,捕获了受时序和属性信息共同影响的用户个性化偏好。为了进一步捕获用户跨行为之间的依赖,我们设计了多行为权重生成模块和偏好迁移网络,进而利用定制的元知识进行多行为用户项目的交互预测。具体来说,属性权重增强的元知识图神经网络由 3 个部分组成:多行为权重生成、偏好迁移网络、元预测层。

4.3.1 多行为权重生成

在多行为场景中,不同的行为对预测用户偏好的贡献不同,例如在电商平台中,购买行为的权重应比浏览行为的权重要大,并且在同一行为中,不同用户的权重也不相同。为了利用多行为信息更准确地捕获用户的偏好,使模型自动学习不同行为的重要性,我们提出了一种多行为权重生成方法,计算式如式(12)所示:

$$\alpha_r = \frac{\omega_r \cdot n_u^r}{\sum_{s \in N_R} \omega_s \cdot n_u^s} \quad (12)$$

$$\mathbf{E}_{u,r} = \mathbf{E}_{u,r} \cdot \alpha_r$$

其中, ω_r 表示用户 u 在行为 r 下的权重,对所有用户都是相同的; n_u^r 表示用户 u 在行为 r 下交互项目的次数; n_u^s 表示用户 u 总共的交互数量。进而将权重赋予用户嵌入,反映用户的个性化偏好。

4.3.2 偏好迁移网络

用户的多行为交互信息提供了有效的外部知识,用户辅助行为的交互(浏览、收藏、加入购物车)可以进一步丰富自身的个性化偏好,但不同行为对用户有着不同的重要性。从图 1 中用户 u_1 的交互序列可以观察到,用户整体的偏好倾向于数码产品,但却在 t_1 时刻浏览了运动鞋。为了将辅助行为中正向影响的偏好迁移至目标行为中,我们设计了偏好迁移网络,首先基于融合了行为权重的多行为嵌入生成目标行为嵌入和辅助行为嵌入,计算式如式(13)所示:

$$\begin{aligned} \mathbf{E}_{\bar{T}} &= \sigma(\mathbf{w}_1 \cdot (\mathbf{E}_{u,\bar{T}} \parallel \mathbf{E}_{i,\bar{T}})) \\ \mathbf{E}_{\bar{A}} &= \sigma(\mathbf{w}_2 \cdot (\mathbf{E}_{u,\bar{A}} \parallel \mathbf{E}_{i,\bar{A}})) \end{aligned} \quad (13)$$

其中, $\sigma(\cdot)$ 表示激活函数 ReLU; $\mathbf{w}_{(\cdot)}$ 为学习的参数矩阵; \bar{T} 为目标行为(购买); \bar{A} 为辅助行为, $\mathbf{E}_{u,\bar{A}} = \mathbf{E}_{u,v} \parallel \mathbf{E}_{u,f} \parallel \mathbf{E}_{u,c}$ 。

相较于以往只对单一行为建模的序列推荐系统,MB-ATMK 通过对目标行为和辅助行为的建模,有效捕获了多行为之间的依赖关系。偏好迁移网络基于目标行为嵌入和辅助行为嵌入生成定制的元知识 $\bar{\omega}$,计算式如式(14)所示:

$$\bar{\omega} = \sigma(\mathbf{w} \cdot (\mathbf{E}_{\bar{T}} \parallel \mathbf{E}_{\bar{A}})) \quad (14)$$

基于元知识 $\bar{\omega}$ 生成用于元预测层的可学习参数矩阵 $\mathbf{p}_1, \mathbf{p}_2, \mathbf{b}_3$,计算式如式(15)所示:

$$\begin{aligned} \mathbf{p}_1 &= \omega_1 \cdot \bar{\omega} + \bar{b}_1 \\ \mathbf{p}_2 &= \omega_2 \cdot \bar{\omega} + \bar{b}_2 \\ \mathbf{b}_3 &= \omega_3 \cdot \bar{\omega} + \bar{b}_3 \end{aligned} \quad (15)$$

其中, $\omega_{(\cdot)}$ 是偏好迁移网络的参数, $\bar{b}_{(\cdot)}$ 是偏置项。

4.3.3 元预测层

在元预测层中,为了预测用户目标行为下的交互,我们基于目标行为嵌入和参数矩阵 \mathbf{p}_1 和 \mathbf{b}_3 生成特征矩阵 \mathbf{F} ,计算式如式(16)所示:

$$\mathbf{F} = \sigma(\mathbf{p}_1 \cdot \ell(\mathbf{E}_{u,\bar{T}}, \mathbf{E}_{i,\bar{T}}) + \mathbf{b}_3) \quad (16)$$

其中 σ 为激活函数 ReLU, $\ell(\cdot)$ 为嵌入运算函数, $\ell(\mathbf{E}_{u,\bar{T}}, \mathbf{E}_{i,\bar{T}}) = \mathbf{E}_{u,\bar{T}} \odot \mathbf{E}_{i,\bar{T}} \parallel \mathbf{E}_{u,\bar{T}} \parallel \mathbf{E}_{i,\bar{T}}$ 。基于偏好迁移网络生成的可学习参数 \mathbf{p}_2 和特征矩阵 \mathbf{F} ,计算 MB-ATMK 最终的预测 $\hat{y}_{m,n}^{\bar{T}}$,计算式如式(17)所示:

$$\hat{y}_{m,n}^{\bar{T}} = \mathbf{F} \cdot \mathbf{p}_2 \quad (17)$$

4.3.4 损失函数

我们使用边际成对损失作为训练过程中的损失函数,计算式如式(18)所示:

$$\mathcal{L} = - \sum_{m=1}^M \sum_{s=1}^S \sum_{i=1}^{\bar{T}+1} \sum_{\bar{a}=1}^{\bar{T}+1} \max(0, 1 - \hat{y}_{m,p_s,\bar{i}}^{\bar{a}} + \hat{y}_{m,n_s,\bar{i}}^{\bar{a}}) \quad (18)$$

其中, p_s 和 n_s 为正样本和负样本, m 和 n 为用户和项目的索引。负样本随机选择用户未交互的项目。使用随机梯度下降的 Adam 优化器来训练网络模型。

5 实验设置

5.1 数据集

为了验证所提方法的有效性,我们在 Tmall 和 Fliggy Trip 两个基准数据集上对所提模型的有效性进行了评估。数据集中包含用户多行为类型交互下的用户 ID 以及与该用户所交互的项目 ID。其中 ID 为用户和项目的唯一标识。用户或项目的交互数量较少导致数据不足以反映用户的真实偏好,因此过滤掉了交互数量少于 5 的用户和项目。表 1 列出了两个数据集的数据分布情况。

表 1 Tmall 和 Fliggy Trip 数据集划分
Table 1 Datapartition of Tmall and Fliggy Trip

Dataset	User	Item	Interactions
Tmall	7775	5953	60087
Fliggy Trip	17307	4923	161316

其中 Tmall 是由天猫商店提供的推荐数据集,其中包含的用户多行为交互类型有浏览、收藏、加入购物车、购买。此外,该数据集还包含了用户和项目的属性信息,用户的属性信息包括年龄、性别,项目的属性信息包括类别、品牌、项目售卖方。

Fliggy Trip 是由飞猪平台提供的推荐数据集。其中包含的用户多行为交互类型有浏览、加入购物车、购买。此外,该数据集还包含了用户和项目的属性信息以及用户和项目交

互的时间戳。其中用户的属性信息包括年龄、职业、喜欢的城市,项目的属性信息包括类别、城市。

5.2 实施细节

本文提出的模型在 Tensorflow 框架上实施。超参数设置如下:MB-ATMK 中多头注意力头的个数为 6(即 $k = 6$);Tmall 和 Fliggy Trip 中的 dropout 都为 0.7;Tmall 中的学习率为 $\times 10^{-3}$, Fliggy Trip 中的为 $\times 10^{-2}$;batch_size 在两个数据集中设置为 128;嵌入向量维度为 32,时序嵌入向量维度为 16。每个训练和测试过程都在单个 RTX 3090 GPU 上完成。实验结果数据都是基于测试集上 5 次随机运行的平均分数。所有的超参数都是通过基于测试性能的网格搜索来确定的。

5.3 评估指标

模型使用了推荐任务中广泛使用的两个评价指标:归一化折损累计增益(NDCG)和命中率(HR)。其中 NDCG 衡量排序的结果,强调推荐列表的顺序性,NDCG 值越大,说明排序的结果越符合用户的偏好;HR 强调推荐的准确性。本实验采用了在以往的研究中广泛使用的 left-one-out 评价方法。

5.4 基线模型

(1)单行为推荐模型

NeuMF^[1]:一个典型的基于深度学习推荐算法。它结合了传统矩阵分解和多层感知机,可以同时抽取低维和高维的用户-项目交互信号。

IMPGCN^[30]:具有相似偏好的用户聚合为用户项目交互图,并使用图神经网络编码用户和项目的嵌入。

NGCF^[31]:设计了一种神经网络方法来递归地在图中传播嵌入,而不是将交互图进行扩展;并设计了一个分层传播层,通过聚合交互项目的嵌入来细化用户的嵌入。

(2)多行为推荐模型

GHCF^[17]:基于用户的多行为交互,在图卷积操作中加入行为的权重,即使用注意力机制和记忆单元。

MBGCN^[32]:提出了融合多行为的用户项目异构图,使用用户嵌入包含了特定的行为的权重,执行行为感知的嵌入传播。

MBHT^[33]:设计了一个多行为超图增强型 Transformer 来捕获用户短期和长期的跨类型行为依赖性。

(3)序列推荐模型

SASREC^[34]:提出了一个基于自注意的顺序模型来捕获长期语义信息。

Bert4rec^[3]:通过利用用户历史行为,来提取用户的历史兴趣进行用户关注商品的预测。

SURGE^[35]:基于度量学习将松散的项目序列重构为紧密的项目-项目兴趣图,将长期用户行为中不同类型的偏好整合到图中的簇中。

(4)融合外部知识的推荐模型

KGAT^[36]:结合了用户项目交互图和知识图谱,并使用了注意力机制区分不同邻居节点的重要性。

AGCN^[37]:将属性信息融入用户项目交互图中,并提出了一种用于联合项目推荐和属性推理的自适应图卷积网络。

6 结果和分析

本章首先讨论了所提模型和基线模型的实验结果,其次进行了消融实验来研究时序感知编码、元知识生成器、属性权重增强模块和元预测层的贡献,接下来进行了不同的实验来

验证行为对模型的影响以及 MB-ATMK 模型框架的性能。

6.1 与其他基线模型比较

表 2 列出了本文提出的 MB-ATMK 和基线在 Tmall 和 Fliggy Trip 数据集上的实验性能。我们将基线分为只使用单行为的推荐模型、使用多行为的推荐模型、序列推荐模型以及使用外部知识增强用户偏好的推荐模型。可以观察到:

(1)与只使用单行为的推荐模型(NeuMF^[1], IMPGCN^[29], NGCF^[31])相比,MB-ATMK 的性能明显提升,这是因为我们引入多行为作为辅助信息,有效捕获了用户的个性化偏好。

(2)MB-ATMK 与多行为推荐模型(GHCF^[17], MBGCN^[32], MBHT^[33])以及序列推荐模型(Bert4rec^[3], SASREC^[34], SURGE^[35])之间性能的差距,证明了编码用户的时序信息以及使用属性信息增强用户偏好建模的有效性。首先用户最终时刻的交互受先前时刻多行为交互的共同影响,通过编码用户的时序嵌入有效捕获了用户随时间变化的动态偏好。同时,丰富的属性信息可以影响用户的多行为交互,本文设计的属性权重注意力,在每一层元知识图神经网络的迭代聚合过程中根据不同属性的权重增强用户和项目的嵌入编码,捕获了用户的个性化偏好。此外,通过时序角度和属性角度共同编码用户的个性化偏好,显著提升了模型的性能。

(3)MB-ATMK 与使用融合外部信息的推荐模型(KGAT^[36], AGCN^[37])相比,MB-ATMK 性能提升明显,证明了我们根据基于优化的元学习设计的元知识图神经网络的优势。首先通过元知识生成器提炼了用户的多样化交互和跨行为之间的依赖关系,将元知识转换为参数矩阵应用于图神经网络编码了具有高阶连通性的用户和项目的多行为嵌入。同时,我们设计的多行为权重生成模块和偏好迁移网络在多行为交互场景中捕获了个性化的多行为语义,编码了用户的跨行为依赖,刻画了细粒度的用户偏好。

表 2 在 Tmall 和 Fliggy Trip 数据集上的实验结果

Table 2 Experimental results on Tmall and Fliggy Trip datasets

Model	Tmall		Fliggy Trip	
	HR	NDCG	HR	NDCG
NeuMF	0.021	0.013	0.057	0.036
NGCF	0.024	0.017	0.064	0.039
IMPGCN	0.038	0.025	0.091	0.044
GHCF	0.161	0.078	0.071	0.065
MBGCN	0.187	0.084	0.083	0.074
MBHT	0.204	0.121	0.105	0.097
SASREC	0.137	0.056	0.098	0.042
Bert4rec	0.164	0.059	0.128	0.056
SURGE	0.217	0.121	0.127	0.097
AGCN	0.131	0.049	0.076	0.042
KGAT	0.039	0.021	0.105	0.064
MB-ATMK	0.237	0.152	0.164	0.115

6.2 消融实验

为了研究 MB-ATMK 中设计模块的不同贡献,我们对两种数据集进行了消融研究。考虑以下设置:

w/o TAE(时序感知编码):在元知识编码层中去除了时序感知编码模块。

w/o MKG(元知识生成模块):去除提炼了用户的多偏好模式的元知识,只使用用户和项目的初始嵌入以及多行为嵌入进行图卷积神经网络的迭代聚合操作。

w/o AWE(属性权重注意力):去除了属性权重注意力

机制,不使用属性信息编码用户的个性化偏好。

w/o MKP(元知识预测):在模型预测模块阶段去除了多行为权重生成模块以及偏好迁移网络,只使用通过属性权重增强的元知识图神经网络生成具有高阶连通性的用户和项目的多行为嵌入进行预测。

表3列出了消融实验的结果,可以发现:(1)通过将MB-ATMK与w/o TAE相比,我们发现将时序感知编码引入模型,使用时序感知注意力,模型可以捕获随着时间的发展用户的动态偏好,从而提升模型的推荐性能。

表3 MB-ATMK在两个数据集上的消融实验

Table 3 Ablation experiments of MB-ATMK on two datasets

	Tmall		Fliggy Trip	
	HR	Improve/%	HR	Improve/%
MB-ATMK	0.237	—	0.164	—
w/o TAE	0.225	5	0.147	10
w/o MKG	0.195	17	0.132	19
w/o AWE	0.217	8	0.139	15
w/o MKP	0.231	2	0.158	3

(2)通过将MB-ATMK与w/o MKG相比,我们发现通过使用行为语义以及时序嵌入生成的元知识有效提炼了用户的多样化交互,反映了用户的多偏好模式。相较于使用用户和项目的初始嵌入,将元知识作为图神经网络的输入显著提升了模型的推荐性能。

(3)通过将MB-ATMK与w/o AWE相比,我们发现通过属性权重增强的用户和项目的嵌入,效果比只使用用户和项目的初始嵌入更好。用户的多行为交互受用户和项目丰富的属性信息所影响,如用户的年龄、性别;项目的品牌、外观。我们加入属性信息并设计的属性权重注意力机制通过在每层元知识图神经网络中生成当前层数的属性权重显著增强了用户的细粒度偏好。

(4)通过将MB-ATMK与w/o MKP相比,我们发现本文设计的多行为权重生成模块通过用户的总交互次数以及当前行为的交互次数编码用户多行为权重捕获了用户个性化的多行为语义。同时,偏好迁移网络通过将多行为嵌入转化为目标行为嵌入以及辅助行为嵌入,进而生成用于预测的元知识参数矩阵在多行为交互场景中编码了用户的跨行为依赖。

6.3 多行为影响分析

为了进一步证明将多种行为引入MB-ATMK模型的有效性,我们进行了另一项消融实验。实验结果如图4所示,其中“-View”“-Fav”“-Cart”“-Buy”分别表示在模型中删除了浏览行为、收藏行为、加入购物车行为和购买行为。

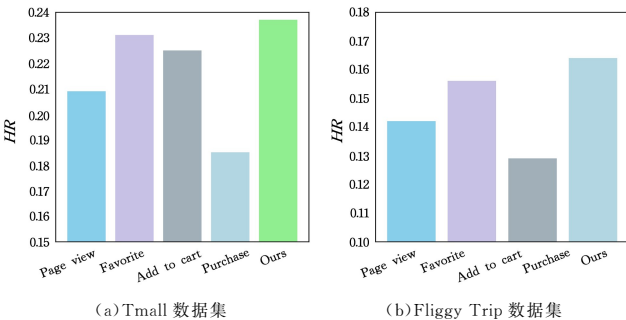


图4 用户多行为消融实验

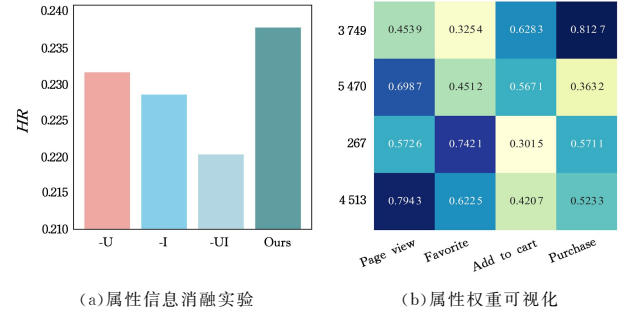
Fig. 4 User multi-behavior ablation experiment

我们可以从两个数据集中观察到购买行为对模型预测的

贡献是最大的,其次是浏览行为。在现实场景中,用户的购买行为和多次的浏览行为可以更好地反映出用户的个性化偏好,结果说明在序列推荐中行为的影响是很大的。本文模型通过引入多种行为有效缓解了数据稀疏问题,从多个角度丰富了用户的偏好,并增强了模型的推荐性能。

6.4 属性信息影响分析

我们对属性信息进行了消融实验,从图5(a)可以看出项目的属性信息对模型性能的影响最大,加入属性信息有效提升了模型的性能。在现实场景中,不同的属性信息对用户的多行为交互有不同的影响,我们基于第二层元知识图神经网络中使用属性权重注意力生成的属性权重 $\xi_{u,a}^2$ 进行了可视化操作,如图6(b)所示。



(a)属性信息消融实验

(b)属性权重可视化

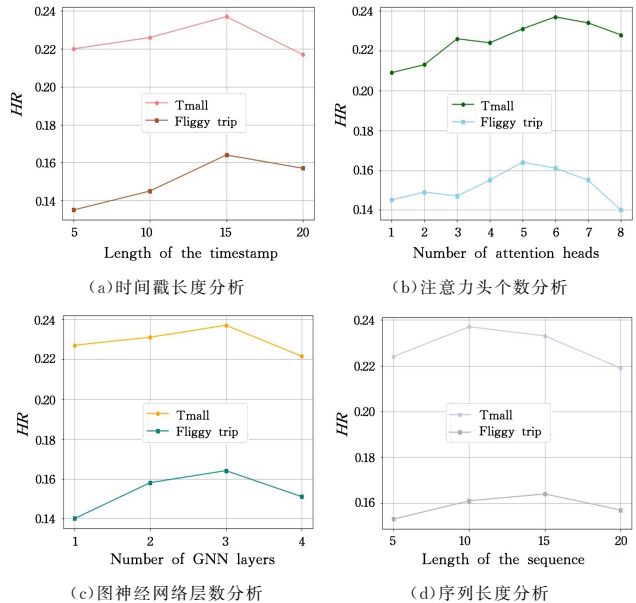
图5 上下文和话语者信息融合模块

Fig. 5 Context and speaker information fusion module

可以看出该用户针对id为5470的项目,其购买行为的属性权重为0.3632,浏览行为的属性权重为0.6987,说明由于项目的属性信息,该用户更倾向于浏览该项目。使用属性权重注意力在每层图神经网络生成的属性权重增强了对用户个性化偏好的捕获。

6.5 参数分析

为了评估本文提出的MB-ATMK模型的性能,我们进行了4项实验,展示了不同参数设置的效果。本模型共有4个关键参数配置,即时间戳的长度 $t_{u,i,r}$ 、时序感知注意力中注意力头的数量 k 、元知识图神经网络的层数 l 和用户交互序列的长度。得出的结论如下:



(c)神经网络层数分析

(d)序列长度分析

图6 参数分析

Fig. 6 Parameter analysis

6.5.1 时间戳长度分析

在提出的时序感知编码模块中,我们使用不同长度的时间戳 $t_{u,i}$ 观察相应的 HR 得分。图 6(a)展示了这种分析结果,其中可以观察到:当时间戳长度为 10 时,时序感知编码模块在两个数据集上获得了更好的性能。然而当时间戳的长度增加,模块的性能减弱。造成这种情况的原因可能是交互时间过久的项目无法有效地反映用户当下的偏好。

6.5.2 注意力头个数分析

在提出的时序感知注意力中,我们使用不同数量的注意力头 k 观察相应的 HR 和 NDCG 得分。从图 6(b)可以观察到,在 Tmall 数据集中,当注意力头的数量为 6,在 Flliggy Trip 数据集中,当注意力头的数量为 5 时,时序感知注意力在两个数据集上获得了更高的定度测量。

6.5.3 图神经网络层数的分析

在 MB-ATMK 模型中,我们使用不同层数的图卷积神经网络观察相应的 HR 和 NDCG 得分。从图 6(c)可以看出,使用 3 层图神经网络可以有效提高模型的性能,证明了有高阶连通性的嵌入具有积极的效果。但随着层数的增加,引入了过多的噪声,进而导致传播层出现过渡平滑问题,因此模型的性能下降。

6.5.4 用户交互序列长度分析

在 MB-ATMK 模型中,我们使用不同的用户交互序列长度观察相应的 HR 和 NDCG 得分。从图 6(d)可以看出,在 Tmall 数据集中,当序列长度为 10,在 Flliggy Trip 数据集中,当序列长度为 15 时,获得了更高的定度测量。当用户交互序列较短时,由于数据稀疏,模型无法获得更好的性能。当用户交互序列过长时,由于引入了数据导致噪声的增加,模型的性能开始下降。

结束语 本文通过引入多行为作为辅助信息并从多个角度丰富了用户在项目交互中的多种偏好,同时加入时序感知编码和属性权重增强模块,有效提取了用户和项目丰富的属性信息和时序信息。此外,我们通过基于优化的元学习方法设计了元知识生成器,使用元知识综合表示用户的多偏好模式,并将含有行为语义和时序偏好的元知识转换为参数,用于轻量级图神经网络层和预测层,进而设计了序列推荐模型,即 MB-ATMK。在两个数据集上进行了对比实验,实验结果表明,该模型在许多推荐基线上取得了显著的进步。

然而,MB-ATMK 仍有一些不足之处。例如,在现实世界的推荐系统中,不仅用户和项目,它们之间的行为都在随时间变化。为了反映最新的偏好,下一步我们将考虑将传统的用户-项目交互图转换为超图,进而编码用户的动态偏好。

参考文献

- [1] HE X, LIAO L, ZHANG H, et al. Neural collaborative filtering [C] // Proceedings of the 26th International Conference on World Wide Web. 2017:173-182.
- [2] WANG X, HE X. Neural graph collaborative filtering[C] // Proceedings of the 42nd international ACM SIGIR Conference on Research and Development in Information Retrieval. 2019:165-174.
- [3] SUN F, LIU J, WU J, et al. Bert4rec: Sequential recommendation with bidirectional encoder representations from transformer [C] // Proceedings of the 28th ACM International Conference on Information and Knowledge Management. 2019.
- [4] WU S, TANG Y, ZHU Y, et al. Session-based recommendation with graph neural networks [C] // Proceedings of the AAAI Conference on Artificial Intelligence. 2019, 33(1):346-353.
- [5] ZHOU H, TAN Q, HUANG X, et al. Temporal augmented graph neural networks for session-based recommendations [C] // Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval. 2021:1798-1802.
- [6] TANG J, WANG K. Personalized top-n sequential recommendation via convolutional sequence embedding [C] // Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining. 2018:565-573.
- [7] MA C, MA L, ZHANG Y, et al. Memory augmented graph neural networks for sequential recommendation [C] // Proceedings of the AAAI Conference on Artificial Intelligence. 2020, 34(4):5045-5052.
- [8] TAN Q, ZHANG J, YAO J, et al. Sparse-interest network for sequential recommendation [C] // Proceedings of the 14th ACM International Conference on Web Search and Data Mining. 2021:598-606.
- [9] CHANG J, GAO C, ZHENG Y, et al. Sequential recommendation with graph neural networks [C] // Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval. 2021:378-387.
- [10] ZHANG M, WU S, YU X, et al. Dynamic graph neural networks for sequential recommendation [J]. IEEE Transactions on Knowledge and Data Engineering, 2022, 35(5):4741-4753.
- [11] WANG W, ZHANG W, LIU S, et al. Beyond clicks: Modeling multi-relational item graph for session-based target behavior prediction [C] // Proceedings of the Web Conference 2020:3056-3062.
- [12] YANG Y, HUANG C, XIA L, et al. Multi-behavior hypergraph-enhanced transformer for sequential recommendation [C] // Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. 2022:2263-2274.
- [13] ZHOU H, TAN Q, HUANG X, et al. Temporal augmented graph neural networks for session-based recommendations [C] // Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval. 2021:1798-1802.
- [14] JIANG N, ZENG Z, WEN J, et al. Incorporating multi-interest into recommendation with graph convolution networks [J]. International Journal of Intelligent Systems, 2022, 37(11):9192-9212.
- [15] ZHAO Z, CHENG Z, HONG L, et al. Improving user topic interest profiles by behavior factorization [C] // Proceedings of the 24th International Conference on World Wide Web. 2015:1406-1416.
- [16] CHEN C, ZHANG M, ZHANG Y, et al. Efficient heterogeneous collaborative filtering without negative sampling for recommendation [C] // Proceedings of the AAAI Conference on Artificial Intelligence. 2020:19-26.
- [17] CHEN C, MA W, ZHANG M, et al. Graph heterogeneous multi-relational recommendation [C] // Proceedings of the AAAI Conference on Artificial Intelligence. 2021:3958-3966.

- [18] JAMAL M A, QI G J. Task agnostic meta-learning for few-shot learning[C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019:11719-11727.
- [19] JIN B, CHENG K, QU Y, et al. Fast sparse connectivity network adaption via meta-learning[C] // 2020 IEEE International Conference on Data Mining(ICDM). IEEE, 2020:232-241.
- [20] SANTORO A, BARTUNOV S, BOTVINICKM, et al. Meta-learning with memory-augmented neural networks[C] // International Conference on Machine Learning. PMLR, 2016:1842-1850.
- [21] FINN C, ABBEEL P, LEVINE S. Model-agnostic meta-learning for fast adaptation of deep networks[C] // International Conference on Machine Learning. PMLR, 2017:1126-1135.
- [22] CHEN M, ZHANG W, ZHANG W, et al. Meta relational learning for few-shot link prediction in knowledge graphs[J]. arXiv: 1909.01515, 2019.
- [23] NIU G, LI Y, TANG C, et al. Relational learning with gated and attentive neighbor aggregator for few-shot knowledge graph completion[C] // Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval. 2021:213-222.
- [24] LU Y, JIANG X, FANG Y, et al. Learning to pre-train graph neural networks[C] // Proceedings of the AAAI Conference on Artificial Intelligence. 2021:4276-4284.
- [25] WEN Z, FANG Y, LIU Z. Meta-inductive node classification across graphs[C] // Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval. 2021:1219-1228.
- [26] ACHILLE A, LAM M, TEWARI R, et al. Task2vec: Task Embedding for meta-learning[C] // Proceedings of the IEEE/CVF International conference on Computer Vision. 2019:6430-6439.
- [27] XIONG W, YU M, CHANG S, et al. One-shot relational learning for knowledge graphs[J]. arXiv:1808.09040, 2018.
- [28] XIA L, HUANG C, XU Y, et al. Knowledge-enhanced hierarchical graph transformer network for multi-behavior recommendation[C] // Proceedings of the AAAI Conference on Artificial Intelligence. 2021:4486-4493.
- [29] HE X, DENG K, WANG X, et al. Lightgcn: Simplifying and powering graph convolution network for recommendation[C] // Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval. 2020:639-648.
- [30] LIU F, CHENG Z, ZHU L, et al. Interest-aware message-passing gcn for recommendation[C] // Proceedings of the Web Conference 2021. 2021:1296-1305.
- [31] WANG X, HE X, WANG M, et al. Neural graph collaborative filtering[C] // Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval. 2019:165-174.
- [32] JIN B, GAO C, HE X, et al. Multi-behavior recommendation with graph convolutional networks[C] // Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval. 2020:659-668.
- [33] YANG Y, HUANG C, XIAL, et al. Multi-behavior hypergraph-enhanced transformer for sequential recommendation[C] // Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. 2022:2263-2274.
- [34] KANG W C, MCAULEY J. Self-attentive sequential recommendation[C] // 2018 IEEE International Conference on Data Mining(ICDM). IEEE, 2018:197-206.
- [35] WANG M, REN P, MEI L, et al. A collaborative session-based recommendation approach with parallel memory modules[C] // Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval. 2019:345-354.
- [36] WANG X, HE X, CAO Y, et al. Kgat: Knowledge graph attention network for recommendation[C] // Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. 2019:950-958.
- [37] PENG Z, LIU H, JIA Y, et al. Attention-driven graph clustering network[C] // Proceedings of the 29th ACM International Conference on Multimedia. 2021:935-943.



CHEN Yuzhe, born in 1999, master. His main research interesetes include recommendation system and so on.



CAO Qiong, born in 1979, master, lecturer. Her main research interesetes include data science and deep learning.