计算机研究与发展
Journal of Computer Research and Development
DOI: 10.7544/issn1000-1239.202440137
61(8): 1917−1929, 2024

GraphMLP-Mixer: 基于图-多层感知机架构的高效多行为序列推荐方法

卢晓凯[1,2] 封 军[1,2] 韩永强[1] 王 皓[1] 陈恩红[1]
[1](认知智能全国重点实验室(中国科学技术大学) 合肥 230000)
[2](安徽省优质采科技发展有限责任公司 合肥 230000)
(xiaokailu@mail.ustc.edu.cn)

GraphMLP-Mixer: A Graph-MLP Architecture for Efficient Multi-Behavior Sequential Recommendation Method

Lu Xiaokai[1,2], Feng Jun[1,2], Han Yongqiang[1], Wang Hao[1], and Chen Enhong[1]
[1] (*State Key Laboratory of Cognitive Intelligence* (*University of Science and Technology of China*), *Hefei* 230000)
[2] (*YouZhiCai Science and Technology Development Co., Ltd*, *Hefei* 230000)

Abstract In the domain of multi-behavior sequence recommendation, Graph Neural Networks (GNNs) have been widespreadly adopted, yet they have limitations, notably in terms of adequately modeling the collaborative signals that exist between different sequences and addressing the challenges posed by long-distance dependencies. To bridge these gaps, a novel framework named GraphMLP-Mixer has been introduced. This innovative framework begins by constructing a global item graph, which is designed to bolster the model's capacity to encapsulate the collaborative signals that are present across sequences. It then merges the perceptron-mixer architecture with graph neural networks, resulting in a graph-perceptron mixer model capable of delving deep into the intricacies of user interests. GraphMLP-Mixer stands out for its two principal strengths: It not only succeeds in effectively capturing the global dependencies inherent in user behaviors but also manages to alleviate the issue of excessive information compression. Furthermore, the framework boasts remarkable improvements in terms of time and space efficiency, with its complexity scaling in a linear fashion with the number of user interactions, thus outperforming existing GNN-based models in the realm of multi-behavior sequence recommendation. The robustness and efficiency of GraphMLP-Mixer in tackling the complexities of multi-behavior sequence recommendation have been thoroughly validated through extensive experimentation on three diverse and publicly available datasets.

Key words multi-behavior modeling; sequential recommendation; graph neural network; MLP architecture; global item graph

摘 要 在多行为序列推荐领域,图神经网络(GNNs)虽被广泛应用,但存在局限性,如对序列间协同信号建模不足和处理长距离依赖性等问题.针对这些问题,提出了一种新的解决框架GraphMLP-Mixer.该框架首先构造全局物品图来增强模型对序列间协同信号的建模,然后将感知机-混合器架构与图神经网络结合,得到图-感知机混合器模型对用户兴趣进行充分挖掘.GraphMLP-Mixer具有2个显著优势:一是能够有效捕捉用户行为的全局依赖性,同时减轻信息过压缩问题;二是其时间与空间效率显著提高,其复杂

收稿日期: 2024-03-15; **修回日期**: 2024-05-23
基金项目: 国家自然科学基金项目(62202443, U23A20319);安徽省科技重大专项(2023z020006);CCF-腾讯犀牛鸟研究基金(RAGR20230124)
This work was supported by the National Natural Science Foundation of China (62202443, U23A20319), the Anhui Provincial Science and Technology Major Project (2023z020006), and the CCF-Tencent Rhino-Bird Open Research Fund (RAGR20230124).
通信作者: 王皓(wanghao3@ustc.edu.cn)

度与用户交互行为的数量成线性关系,优于现有基于 GNN 多行为序列推荐模型. 在 3 个真实的公开数据集上进行实验,大量的实验结果验证了 GraphMLP-Mixer 在处理多行为序列推荐问题时的有效性和高效性.

关键词 多行为建模;序列推荐;图神经网络;MLP 架构;全局物品图

中图法分类号 TP183

在当今互联网技术飞速发展的背景下,推荐系统已成为众多网络平台不可或缺的核心组件. 这类系统通过深入挖掘用户与产品间的交互数据,向用户推送量身定制的内容推荐. 特别是序列推荐[1]算法,它通过细致分析用户的历史交互时间序列,显著提高了对用户未来行为的预测精度. 序列推荐的独特之处在于,不仅能够显式地对物品的时间顺序关系进行建模,而且更准确地捕捉了用户动态行为模式和兴趣变化,因此受到了学术界和工业界的广泛重视.

在序列推荐的背景下,模型通常借助物品序列来模拟用户偏好的演变. 基于这一理念,研究者们提出了多种融合了复杂神经网络技术的推荐系统,以捕捉序列内物品间的交互关系,并显著提升了性能. 例如,基于循环神经网络的 GRU4Rec[2]和基于卷积的序列编码器模型 Caser[3]等. 同时,受到 Transformer[4]架构的启发,自注意力机制已被应用于 SASRec[5]和 BERT4Rec[6]模型中,以捕捉物品之间的相互作用. 图神经网络(GNN)[7]在序列推荐系统中的应用也展示了其有效性. 得益于图卷积的能力,一些基于 GNN 的方法如 SRGNN[8]和 GCSAN[9],通过在构建的物品图结构上传递信息来学习物品间的转换关系.

然而,在实际应用场景中用户与商品的互动行为是多种类型的,如点击、浏览、收藏、购买等. 这些多样化的行为揭示了用户的多维偏好,并可作为重要的辅助信息来增强推荐系统对目标行为的预测能力,有效提高针对特定目标的推荐性能. 为了学习这些多样化的信息并有效地利用它们进行用户下一项物品推荐,需要对用户多行为序列进行联合建模[10]. 近年来,受到 GNN 对结构化数据有效建模的启发,一些工作考虑建模用户-物品交互图并提出基于图结构的多行为序列推荐方法. 例如,MB-GCN[11]利用图卷积运算解决多类型用户行为建模中的数据稀疏性问题,并捕捉多样化多行为模式. KHGT[12]采用变换器整合时序信息和辅助知识图谱,引入图注意力网络建模用户行为表征. CML[13]通过元对比编码将对比学习融入多行为推荐,学习个性化行为特征. KMCLR[14]则提出了新颖的训练框架,结合对比学习任务和多个功能模块如多用户行为学习、知识图增强、用户行为粗粒度和细粒度建模等,有效提升了推荐性能. 这些方法为推荐系统的发展提供了新的研究方向和策略.

尽管多行为信息的融入可以进一步有效地探索用户的行为模式和多维用户兴趣,然而,随着多样化用户交互行为的发展,用户行为序列的交互长度在短期时间内急剧增长,这对序列推荐模型的效率带来了严峻的挑战. 在推荐系统的研究与应用中,传统的 GNN 通过其多层局部消息传递机制[15]在用户-物品交互图中传播信息. 这种机制虽然在捕捉节点邻域内的相互作用方面表现出色,但在建模序列间的协同信号以及处理物品间长距离依赖关系方面存在一定的局限性. 全局注意力机制虽然能够一定程度上缓解这些问题,但其计算成本较高,为平方级计算复杂度,这在面临用户行为序列不断增长情况下,对模型效率构成了极大挑战.

对此,本文首先构建一个全局物品图对物品表征进行数据增强,通过建模序列间的协同信号来缓解数据稀疏性问题. 在此基础上,本文提出了一种新颖模型——GraphMLP-Mixer,通过引入 MLP-Mixer[16]架构和 GNN 建模复杂用户兴趣. GraphMLP-Mixer 包含兴趣提取、嵌入、挖掘和用户嵌入 4 个主要模块. 首先兴趣提取模块将用户交互图分割成重叠的兴趣子图,嵌入模块将子图转换为初始兴趣表征,并将生成的表征随后输入兴趣挖掘层,通过多个残差网络高效建模每个兴趣信息和兴趣间的关系以生成兴趣表征. 随后对多兴趣表征进行全局平均池化得到用户表征,并连接到全连接层进行下一个物品的预测. 最终 GraphMLP-Mixer 通过整合全局物品图信息和 MLP-Mixer 兴趣挖掘能力,能够有效实现用户行为序列的深层理解与表示.

本文的主要贡献包括 4 个方面:

1)对多行为序列推荐问题进行了深入探讨,特别关注于高效地捕获用户行为之间的全局依赖关系,以更准确地理解用户的复杂兴趣和行为模式.

2)创新性地提出了一种针对用户行为序列推荐的全局对比数据增强方法,该方法通过深入挖掘物品之间的协同关系,构建了一个全局图结构. 这一结构

有效地增强了用户交互子图的表征能力，从而实现了对用户兴趣和行为模式的更深层次理解，为推荐系统中的数据增强提供了一种全新的视角和解决方案.

3）设计了一种新颖的高效图神经网络结构GraphMLP-Mixer. GraphMLP-Mixer通过融合MLP-Mixer的高效学习能力和GNN的信息挖掘优势，实现了对用户行为全局依赖性的有效捕捉，同时解决了信息过压缩问题，极大地提高了推荐系统的预测精度. 此外，GraphMLP-Mixer计算复杂度与用户交互行为的数量成线性关系，显著优于现有的基于GNN的多行为图推荐模型. 为大规模实时推荐系统和高性能计算环境下的推荐任务提供了理想的解决方案.

4）综合3个真实世界数据集的实验结果，本文GraphMLP-Mixer模型展现出卓越的性能，显著超越了现有的基线模型，top k 指标提升了5%~10%. 此外，模型复杂度分析揭示了GraphMLP-Mixer的线性复杂度特性，相比于基线模型，训练时间减少了约1/3，效率得到显著提升. 通过对图聚类参数的实验分析，本文还证实了模型的鲁棒性，进一步证明了GraphMLP-Mixer在实际推荐系统场景中的巨大潜力.

1 相关工作

1.1 序列推荐

序列推荐通常被视为一个序列预测问题，其关键在于对物品之间的转换规律进行建模. 早期的研究多基于马尔可夫链，通过建模不同物品之间的关联性，例如个性化马尔可夫链分解（FPMC[17]）和结合物品相似度的马尔可夫链（Fossil[18]）. 为了应对序列行为学习中的挑战，基于神经网络的方法被提出，以捕捉复杂的序列信息. 循环神经网络作为建模时间依赖性的直观方式被广泛应用，例如GRU4Rec[2]. 进一步地，注意力机制的引入，使得模型能够识别输入物品序列中的关键部分，有效建模物品之间的长距离依赖性. 例如，SASRec[5]就是通过自注意力机制从用户过去的行为中提取物品相关性. SR-GNN[8]，MTD[19]和GCGNN[20]等方法将序列物品转换为图，并利用GNN对物品进行表示.

然而，这些模型仅考虑了单一类型的序列行为. 在给定用户的交互物品序列的情况下，如何有效利用动态多行为知识来提升推荐准确性，仍是一个亟待解决的重要问题.

1.2 多行为序列推荐

近年来，序列推荐问题在推荐系统中受到了广泛关注. 然而，大多数现有工作都集中在通用序列推荐上，即基于单类型交互序列预测用户下一个目标项，而忽略了真实场景中用户交互的多样性，例如电商平台上的点击、喜欢和购买，这些行为提供了有价值的用户意图洞察. 为了克服这一局限性，研究者们提出了各种处理多行为数据的方法. 在多任务框架方面，一些研究挖掘了不同行为之间的级联关系，例如NMTR[21]. 另一些研究是为不同用户行为分配不同任务，并采用分层注意力机制来提高推荐效率，如DIPN[22]. 其他研究则通过融合多行为数据和使用其他行为作为辅助信号来增强推荐. 例如，MATN[23]使用Transformer和门控网络来捕获行为关系. 受到GNN强大能力的启发，近年来出现了许多通过在连接的用户和物品节点之间进行信息传播来建模用户-物品交互图的多行为推荐方法. MB-GCN[11]旨在解决多类型用户行为数据建模中的数据稀疏性问题. 该方法利用图卷积运算进行消息传递，在构建的用户-物品交互图上捕捉多行为模式. KHGT[12]引入了一种基于变换器的多行为建模方法，重点在于时序信息和辅助知识的融合，该模型利用图注意力网络来捕捉行为嵌入. CML[13]通过提出元对比编码，将对比学习融入到多行为推荐中，使得模型能够学习个性化的行为特征. 此外，KMCLR[14]提出了一种框架，通过2个对比学习任务和3个功能模块（多用户行为学习、知识图增强、用户行为的粗粒度和细粒度建模）来提升推荐系统的性能. 这些方法共同构成了多行为推荐领域的多元化研究路径，为推荐系统的发展提供了新的思路和策略.

尽管整合多种行为信息能够更全面地探索用户的行为模式和多维兴趣，但当行为数据变得更加丰富和复杂时，用户行为序列在短时间内会变得异常冗长. 这对于序列推荐模型来说是一个显著的挑战，因为它需要处理大量的数据点，这不仅增加了计算负担，还可能影响模型的准确性和泛化能力.

2 问题定义

近年来，序列推荐问题已经成为推荐系统领域的一个热点话题. 然而，值得注意的是，大部分现有研究都集中在通用序列推荐上，这种方法基于单一类型的用户交互序列来预测下一个物品，而忽视了用户可能存在的多种行为类型. 在现实世界的应用场景中，例如电子商务平台，用户与物品之间的互动行为是多样化的，可能包括点击、收藏、加入购物车

和实际购买等. 这些行为不仅反映了用户的即时偏好，而且揭示了用户在不同情境下的复杂兴趣和需求.

为了克服现有方法在处理多类型用户行为方面的局限性，本文研究了多行为序列推荐问题. 本文的目标是建立不同用户行为之间的内在联系，深入理解用户在不同行为模式下的行为动机和偏好，并将这些一般性偏好有效地转移到推荐系统所关注的特定目标行为上（目标行为则指的是所预测的行为类型）. 在本文中，我们将其他行为类型（点击、加购物车）视作辅助信息，用以增强模型对目标行为（购买）的预测能力. 通过这种方式，本文的模型能够更全面地捕捉和建模用户的动态行为模式，从而提高推荐系统的准确性和用户满意度.

本文定义多行为序列推荐问题有：

定义 1. 多行为序列推荐. 给定用户集合U、物品集合V和行为类型集合B，对于用户 $u\in U$，其行为感知交互序列S_u由按时间排序的单个三元组(v,b,p)组成. 每个三元组代表在序列位置p下行为类型b交互的物品 v. 因此，对于用户u，问题输入是用户 u的行为感知的$(S_u=\{(v_1,b_1,\boldsymbol{p}_1),(v_2,b_2,\boldsymbol{p}_2),\cdots,(v_L,b_L,\boldsymbol{p}_L)\})$，其中$(v_i,b_i,\boldsymbol{p}_i)$代表序列中第$i$条数据，输出$(v_{L+1},b_t,\boldsymbol{p}_{L+1})$是预测的目标行为$b_t$下一个位置$L+1$的下一个物品$v_{L+1}$.

为了增强模型在序列间的关系建模，本文研究采取了一种与众不同的方法，即不依赖每个序列中物品的局部上下文，而是首先从所有用户的交互序列中抽象出一个全局物品关系图. 图 1 展示了从用户行为序列构建全局物品权重关系图的过程. 在这个图中，每个物品节点通过无向边与其他物品节点相连，形成了一个全局互通的网络结构. 全局物品关系图中的边权重是根据多行为序列中每个物品之间的相邻程度来确定的，这有助于本文利用众多其他用户的行为模式更好地理解目标用户的偏好，从而提高推荐系统的准确性和多样性. 以篮球和足球之间的权重为例，k 表示 2 个对象之间的距离，它们之间的权重为以 $1/k$ 为权重的加权和. 在所有序列中，足球和篮球之间的距离对于 1 来说是 2 次，对于 2 来说是 1 次，n 是 k 的最大值，即 $k=1,2,\cdots,n$. 因此，当 $n=1$ 时，权重 $w=2$；当 $n=2$ 时，权重 $w=2.5$.

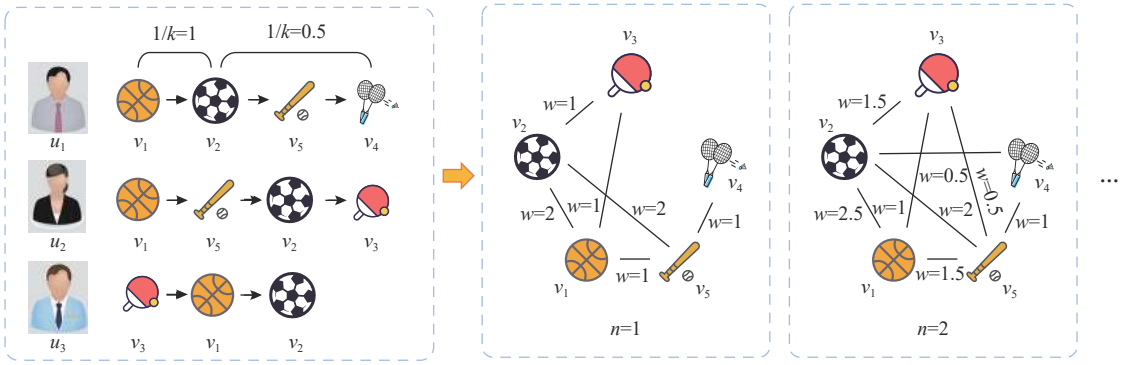


Fig. 1 Example of global item relationship graph

图 1 全局物品关系图示例

定义 2. 物品全局关系图. 给定若干用户交互序列 I_u，$u\in U$代表用户，本文定义一个n 阶全局物品关系图，图的节点为交互序列中所出现的物品节点的集合. 任意 2 个物品节点v_i和v_j之间边的权重 w_{ij}是 2 个节点在所有序列中 $k=\{1,2,\cdots,n\}$阶连接的次数之和，k表示物品的相邻间隔，比如 2 个物品直接相邻即k =1，这 2 个物品间隔 1 个物品即 k=2. 具体公式有：

$$\begin{cases} w_{ij}=\sum\limits_{u\in U}\sum\limits_{k=1}\dfrac{\delta\left(\left|loc(v_i)-loc\left(v_j\right)\right|=k|v_i,v_j\in I_u\right)}{k}, \\ w'_{ij}=Norm\left(w_{ij}\right). \end{cases} \tag{1}$$

本文计算w_{ij}，然后将边权重归一化得到w'_{ij}. w_{ij}表示物品v_i和v_j之间的权重，$loc(v_i)$表示物品v_i在序列中的位置，通过$\left|loc(v_i)-loc(v_j)\right|$将 2 个物品的位置相减计算绝对值即可计算它们之间的距离，δ表示 2 个物品位置在序列中相差k的次数，U是用户集合，I_u是用户u的交互序列. 我们将U中所有的用户序列计算k 阶距离加权求和得到全局图中物品之间的初始权重. 之后，本文设置一个超参数阈值来删除一些边，即将边的权重置为 0，以减少噪声并得到最终的边权重. 这个阈值可以控制全局图的稠密度，在实验中，我们将该阈值统一设置为 0.05. 通过归一化操作后，边数较多的节点，也就是热门的商品节点，每个边的节点权重会相对较低，在根据固定阈值去除边时，边数多的节点会失去更多的边，这种策略有助于缓解

热门节点的影响，从而在一定程度上缓解流行度偏差，使得模型更准确地捕捉商品数据中的关联信息.

3 方法介绍

3.1 概 述

本文研究首先构建了一个全局物品图进行数据增强，旨在深入建模序列间的协同信号. 在此基础上，本文提出了一种新的模型——GraphMLP-Mixer，它结合了 MLP-Mixer 架构和 GNN，用于建模用户兴趣. 该模型包含兴趣提取、嵌入、挖掘和用户嵌入 4 个模块. 如图 2 所示，兴趣提取模块将用户交互项目图分割成重叠的兴趣子图，嵌入模块将子图转换为初始兴趣表征，这些表征随后被送入兴趣挖掘层生成最终表征. 用户表征通过全局平均池化层获得，并连接到全连接层进行预测. 每个兴趣挖掘模块都是一个残差网络，对每个兴趣信息和兴趣间的关系都进行了高效的建模. 通过整合全局物品图的信息，以及利用 MLP-Mixer 兴趣挖掘能力，实现了对用户行为序列的深层理解和表示.

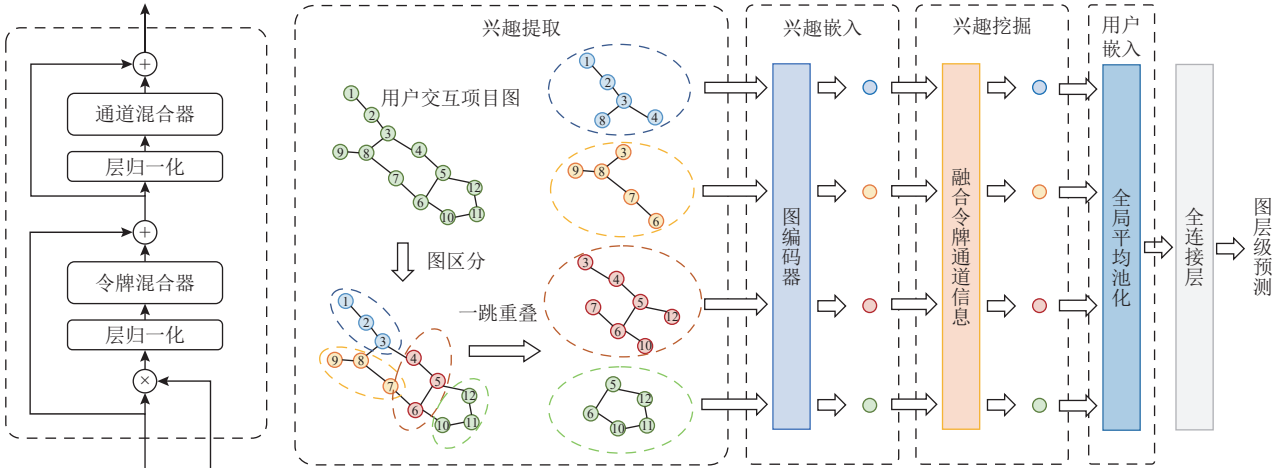


Fig. 2 Illustration of GraphMLP-Mixer

图 2 GraphMLP-Mixer 示意图

具体地，本文通过物品全局关系图，得到各个物品之间边的权重. 对于每个用户，本文在全局图中采集每个用户所相互的物品子图，包括交互过的物品节点及之间的边. 为了方便对后续方法的描述，这里对一些概念进行定义. 令 $G=(\mathcal{V},\mathcal{E})$ 表示每个用户交互序列的子图或用户的兴趣图，其中 $\mathcal{V}$ 是物品节点的集合，$\mathcal{E}$是边的集合. 图有 $N=|\mathcal{V}|$个节点和 $E=|\mathcal{E}|$ 条边. 图的连通性由邻接矩阵 $\boldsymbol{A}\in\mathbb{R}^{N\times N}$表示. 节点$i$的特征表示为$\boldsymbol{h}_i$，而节点$i$和$j$之间的边的特征表示为 $\boldsymbol{e}_{ij}$. 令 $\mathcal{V}_1,\cdots,\mathcal{V}_P$ 是兴趣子图的分块，P是预定义的分块数，$G_i=(\mathcal{V}_i,\mathcal{E}_i)$ 是G的诱导子图，包含所有节点 $\mathcal{V}_i$ 和所有端点属于$\mathcal{V}_i$的边. 令 $\boldsymbol{h}_G$是图级表示，y_G是图级目标.

3.2 兴趣提取模块

在处理用户交互长序列数据时，模型的效率成为了一个关键考量，从用户交互的物品图中高效提取兴趣面临着众多挑战. 由于图的结构各不相同，特别是在节点数量上，这导致了它们无法简单地被均匀分割. 此外，提取过程还必须能够辨识出那些在语义上彼此相似的子图，这些子图代表了用户的特定兴趣点. 因此，兴趣提取模块需要满足要求：1）提取算法需具备普适性，能够适应各种不同的交互图结构. 2）子图内的物品节点之间需要有比与子图外部物品节点之间更紧密的连接，以确保提取的兴趣点是准确和有意义的. 3）提取过程需要保证高效性，其时间复杂度最好能够与图中边的数量成线性关系，即达到 $O(E)$ 的时间复杂度. 这些要求确保了兴趣提取方法不仅能够灵活地处理各种规模的图数据，而且还能够快速且准确地锁定并提取出反映用户兴趣的核心子图.

METIS[24] 是一种高效的图聚类算法，以其卓越的精确度和效率而著称. 该算法能够将图分割成预定的簇数量，确保簇内的连接密度远高于簇间的连接密度，从而有效地捕捉到图中的子图结构. 然而，METIS 的一个局限是它只能识别非重叠的簇. 在推荐系统的问题背景下，例如，一个用户购买了笔记本电脑，这可能同时反映了该用户的学习偏好和娱乐偏好. 如果仅提取非重叠的子图，可能会导致重要信

息的丢失，从而影响预测的准确性. 为了克服这一局限并利用图中的所有边，本文对 METIS 算法进行了改进，使其能够生成重叠的子图. 这种方法保留了用户多元兴趣的完整性，进而提高了推荐系统的预测性能.

具体地，如果边的源节点和目标节点不在同一个子图中，将 2 个节点同时分配给 2 个子图. 这种做法等效于将图分块扩展到块内所有节点的 1 跳邻居. 具体地，本文首先应用 METIS 将图划分为 P 个非重叠拼块 $\mathcal{V}_1,\cdots,\mathcal{V}_P$，使得 $\mathcal{V}_1\cup\cdots\cup\mathcal{V}_P$，且 $\mathcal{V}_i\cap\mathcal{V}_j=\varnothing$，$\forall i\neq j$. 然后，将拼块扩展到它们的 1 跳邻居，以保留拼块间链接的信息并利用所有图边：$\mathcal{V}_i\leftarrow\mathcal{V}_i\cup\{\mathcal{N}_1(j)|j\in\mathcal{V}_i\}$，其中 $\mathcal{N}_k(j)$ 定义节点 j 的 k 跳邻居. 这种方法允许兴趣子图重叠，保留了用户多元兴趣的完整性，进而提高了推荐系统在真实场景中的预测性能.

3.3 兴趣嵌入模块

在对每个用户提取了 P 个兴趣子图之后，对于每个兴趣子图，本文设计兴趣编码器网络来处理复杂的各种子图结构. 其架构设计用于较好地转换图块 G_p 到一个固定大小的表示 x_{G_p}，具体分为 3 个步骤.

1)原始节点和边的线性嵌入. 输入的节点特征 $\boldsymbol{\alpha}_i\in\mathbb{R}^{d_n\times 1}$ 和边特征 $\boldsymbol{\beta}_{ij}\in\mathbb{R}^{d_e\times 1}$ 被线性投影到 d 维隐藏特征：

$$\boldsymbol{h}_i^0=\boldsymbol{U}^0\boldsymbol{\alpha}_i+\boldsymbol{u}^0,\quad \boldsymbol{e}_{ij}^0=\boldsymbol{V}^0\boldsymbol{\beta}_{ij}+\boldsymbol{v}^0, \tag{2}$$

其中 $\boldsymbol{h}_i^0\in\mathbb{R}^d$，$\boldsymbol{e}_{ij}^0\in\mathbb{R}^d$，$\boldsymbol{U}^0\in\mathbb{R}^{d\times d_n}$，$\boldsymbol{V}^0\in\mathbb{R}^{d\times d_e}$ 以及 $\boldsymbol{u}^0,\boldsymbol{v}^0\in\mathbb{R}^d$ 是可学习的参数.

2)带 MP-GNN[25] 的图卷积层. 本文应用一系列 L 个卷积层，其中节点和边表示通过对每个图块 $G_p=(\mathcal{V}_{G_p},\mathcal{E}_{G_p})$ 应用 MP-GNN 来更新，具体公式为：

$$\boldsymbol{h}_{i,p}^{l+1}=f_{\text{node}}\left(\boldsymbol{h}_{i,p}^l,\left\{\boldsymbol{h}_{j,p}^l|j\in\mathcal{N}(i)\right\},\boldsymbol{e}_{ij,p}^l\right)+g_{\text{patch-node}}\left(\boldsymbol{h}_p^l\right), \tag{3}$$

$$\boldsymbol{e}_{ij,p}^{l+1}=f_{\text{edge}}\left(\boldsymbol{h}_{i,p}^l,\boldsymbol{h}_{i,p}^l,\boldsymbol{e}_{ij,p}^l\right)+g_{\text{patch-edge}}\left(\boldsymbol{e}_p^l\right), \tag{4}$$

其中 $\boldsymbol{h}_{i,p}^{l+1},\boldsymbol{h}_{i,p}^l,\boldsymbol{h}_p^l,\boldsymbol{e}_{ij,p}^{l+1},\boldsymbol{e}_{ij,p}^l,\boldsymbol{e}_p^l\in\mathbb{R}^d$，$l$ 是层索引，p 是图块索引，i,j 表示节点，$\mathcal{N}(i)$ 是节点 i 的邻域. 函数 f_{node} 和 f_{edge}（带有可学习参数）定义任意的 MP-GNN 架构，$\boldsymbol{h}_p^l=\dfrac{1}{|\mathcal{V}_p|}\sum\limits_{i\in\mathcal{V}_p}\boldsymbol{h}_{i,p}^l\in\mathbb{R}^d$，$\boldsymbol{e}_p^l=\dfrac{1}{|\mathcal{E}_p|}\sum\limits_{ij\in\mathcal{E}_p}\boldsymbol{e}_{ij,p}^l\in\mathbb{R}^d$ 分别是图块节点和图块边的平均表示. $g_{\text{patch-node}}$，$g_{\text{patch-edge}}$ 是基于 MLP 的函数，作用于 $\boldsymbol{h}_p^l$ 和 $\boldsymbol{e}_p^l$. 对于图块重叠而覆盖多个图块的每个节点和边，本文通过平均重叠图块中的表示来更新节点或边的表示：

$$\boldsymbol{h}_{i,p}^{l+1}\leftarrow Mean\boldsymbol{h}_{i,k}^{l+1}\{k|i\in\mathcal{V}_k\}, \tag{5}$$

$$\boldsymbol{e}_{ij,p}^{l+1}\leftarrow\boldsymbol{e}_{ij,k}^{l+1}\{k|ij\in\mathcal{E}_k\}, \tag{6}$$

其中 $\{k|i\in\mathcal{V}_k\}$，$\{k|ij\in\mathcal{E}_k\}$ 分别是覆盖节点 i，边 ij 的所有图块的集合.

3)池化和读出. 对 G_p 中的所有节点向量进行平均池化，使得 $\boldsymbol{h}_p=\dfrac{1}{|\mathcal{V}_p|}\sum\limits_{i\in\mathcal{V}_p}\boldsymbol{h}_{i,p}^{l=L}\in\mathbb{R}^d$，并应用一个小型 MLP 以获得固定大小的图块嵌入 $x_{G_p}\in\mathbb{R}^d$. 观察到的图块编码器是一个 MP-GNN，因此具有固有的长距离依赖性不足的限制. 这是否会影响 Graph MLP-Mixer 捕捉长距离交互的能力，这个问题的答案是否定的，因为这个问题仅限于大型图，但是对于小型区块图来说，这个问题实际上不存在或者可以忽略不计.

这种设计的动机在于高效地处理各种复杂的兴趣图结构. 通过线性嵌入，本文初始化节点和边的特征，使其适合于后续的图卷积操作. MP-GNN 的使用则允许网络学习到节点和边之间的相互作用，从而捕捉到物品之间的关系信息. 最后通过池化步骤将学习到的局部物品关系信息聚合为兴趣表示.

3.4 时序信息增强

一般通过图建模用户序列，虽然加强了用户行为的多序列间的关系，但是容易忽略每条序列的时序信息. 然而，对于一般的图，节点和图块的这种排序并不存在. 这种位置信息的缺失降低了模型对用户行为时序性的表达能力[26-28]. 因此，本文使用 2 种显式的位置编码（PE）：一种绝对 PE 用于物品节点，另一种相对 PE 用于兴趣图块.

3.4.1 节点 PE

本文在式(2)中对输入节点特征进行了位置信息增强，通过引入 $\boldsymbol{p}_i\in\mathbb{R}^K$ 和一个可学习的矩阵 $\boldsymbol{T}^0\in\mathbb{R}^{d\times K}$：

$$\boldsymbol{h}_i^0=\boldsymbol{T}^0\boldsymbol{p}_i+\boldsymbol{U}^0\boldsymbol{\alpha}_i+\boldsymbol{u}^0\in\mathbb{R}^d. \tag{7}$$

通过节点 PE，模型能够更好地理解用户行为中的时序信息，从而更准确地预测用户的未来行为.

3.4.2 图块 PE

图块之间的相对位置信息可以从原始图的邻接矩阵 $\boldsymbol{A}\in\mathbb{R}^{N\times N}$ 和 METIS 在 3.2 节中提取的兴趣子图 $\mathcal{V}_1,\cdots,\mathcal{V}_P$ 计算得出. 具体来说，本文通过兴趣子图的粗化邻接矩阵 $\boldsymbol{A}^P\in\mathbb{R}^{P\times P}$ 来捕获相对位置信息：

$$\boldsymbol{A}_{ij}^p=\left|\mathcal{V}_i\cap\mathcal{V}_j\right|=Cut\left(\mathcal{V}_i,\mathcal{V}_j\right), \tag{8}$$

其中，$Cut\left(\mathcal{V}_i,\mathcal{V}_j\right)=\sum\limits_{k\in\mathcal{V}_i}\sum\limits_{l\in\mathcal{V}_j}\boldsymbol{A}_{kl}$ 是标准图切割运算符，它计算兴趣子图 $\mathcal{V}_i$ 与 $\mathcal{V}_j$ 之间的连接边的数量. 本文提取在图块级别的位置编码 $\hat{\boldsymbol{p}}_i\in\mathbb{R}^{\hat{K}}$，其类似于物品节点级别，这将注入（经过线性转换后）到兴趣挖掘模

块的第 1 层：

$$\boldsymbol{x}_i^0 = \widehat{\boldsymbol{T}^0 \boldsymbol{p}_i} + \widehat{\boldsymbol{U}^0 \boldsymbol{x}_i} + \widehat{\boldsymbol{u}^0} \in \mathbb{R}^d, \tag{9}$$

其中，$\boldsymbol{x}_i$ 是图块的嵌入表示. 图块位置编码有助于模型识别不同兴趣图块之间的结构关系，进一步强化了图模型对用户兴趣的时序建模能力[29-31]，从而提高推荐系统的准确性和效果.

3.5 兴趣挖掘模块

对于图像，原始的混合器是一个简单的网络，它交替进行通道（channel）和令牌（token）混合步骤[32]. 在我们的框架中，令牌指的是各个兴趣的表征向量，而通道则是兴趣表征向量的各个维度. 换句话说，令牌和通道是兴趣向量矩阵的 2 个维度. 通道和令牌混合交替的步骤使得不同节点内部和节点之间的信息融合成为可能. 混合器层的简单性导致了计算成本的显著降低，而在性能上的牺牲却很小或没有. 事实上，ViT 中的自注意力机制具有 $O(P^2)$ 的空间复杂度和 $O(P^2)$ 的时间复杂度，而 MLP-Mixer 中的混合器层仅有 $O(P)$ 的空间复杂度和 $O(P)$ 的时间复杂度. 这里将 MLP-Mixer 引入到 GNN 中.

令 $\boldsymbol{X} \in \mathbb{R}^{P \times d}$ 为用户兴趣嵌入，$\boldsymbol{X} = (x_{G_1}, \cdots, x_{G_P})$，兴趣挖掘层可以表示为：

$$\boldsymbol{U} = \boldsymbol{X} + (\boldsymbol{W}_2 \sigma(\boldsymbol{W}_1 LayerNorm(\boldsymbol{X}))) \in \mathbb{R}^{P \times d}, \tag{10}$$

$$\boldsymbol{Y} = \boldsymbol{U} + (\boldsymbol{W}_4 \sigma(\boldsymbol{W}_3 LayerNorm(\boldsymbol{U})^{\mathrm{T}}))^{\mathrm{T}} \in \mathbb{R}^{P \times d}, \tag{11}$$

其中，σ 表示为 GELU[33] 非线性激活函数，$LayerNorm(\cdot)$ 是层归一化[34] 操作. 矩阵 $\boldsymbol{W}_1 \in \mathbb{R}^{d_s \times P}, \boldsymbol{W}_2 \in \mathbb{R}^{P \times d_s}$，$\boldsymbol{W}_3 \in \mathbb{R}^{d_c \times d}, \boldsymbol{W}_4 \in \mathbb{R}^{d \times d_c}$，其中 d_s 和 d_c 分别是令牌混合（兴趣内部）和通道混合（兴趣之间）MLP 中可调整的隐藏层宽度.

兴趣挖掘模块通过将 GNN 与 MLP-Mixer 网络结合，迭代通道和令牌混合步骤，实现了兴趣内部和兴趣间的信息融合. 这种方法大幅度降低了计算复杂度，同时保持了模型性能，使其能够更高效地挖掘和利用用户兴趣信息.

3.6 用户嵌入模块

在构建推荐系统的深度学习模型时，本文引入了一个兴趣挖掘模块，该模块专门设计用于提取用户的兴趣表征. 为了从这些表征中得出一个全面且精确的用户兴趣概览，采用了平均池化技术. 平均池化是一种在多个特征上取平均值的操作，它能够有效地融合来自不同部分的兴趣信息，从而得到一个综合的用户表征. 具体地，本文定义了用户表征的生成公式为：

$$\boldsymbol{h}_G = \sum_p m_p x_{G_p} / \sum_p m_p \in \mathbb{R}^d, \tag{12}$$

其中 m_p 是一个二元变量，对于非空区块值为 1，对于空区块值为 0（由于图的大小可变，小图可能产生空区块）. 最后，本文应用一个小型 MLP 来获取用户表征 $\boldsymbol{u}$. 对于单目标多行为推荐问题，最大化 $Q(u)$ 是本文的优化目标：

$$Q(u) = \frac{\exp\left(\boldsymbol{u}\boldsymbol{e}_{v_t}\right)}{\sum_{v \in \mathcal{V}} \exp\left(\boldsymbol{u}\boldsymbol{e}_v\right)}, \tag{13}$$

其中 $v_t \in v_{L+1}$ 表示待预测的以目标行为交互的物品. 通过这种优化策略，本文不仅能够确保用户表征能够准确地反映出用户的兴趣，而且还能够保证推荐系统能够有效地预测用户可能感兴趣的目标物品.

3.7 数据增强

MLP-Mixer 架构虽然在多个领域展现出了强大的表示能力，但它也被观察到存在较强的过拟合[35] 倾向. 为了缓解这一问题，本文在 METIS 图分区算法中引入了随机扰动策略. 具体而言，本文首先定义了一个原始图 $G = (\mathcal{V}, \mathcal{E})$，然后创建了一个扰动图 $G' = (\mathcal{V}, \mathcal{E}')$，其中 $\mathcal{E}'$ 是通过随机删除原始边集 $\mathcal{E}$ 中的一小部分边得到的.

在每次迭代中，本文将在 G' 上应用 METIS 图分区算法，以获得略微不同的物品节点分区 $\mathcal{V}_1, \cdots, \mathcal{V}_P$. 这些分区是基于扰动图的结构特性生成的，因此它们能够引入一定的随机性，从而有助于打破模型可能陷入的局部最优. 然后，本文提取图块 $G_1, \cdots, G_P$，其中 $G_i = (\mathcal{V}_i, \mathcal{E}_i)$ 是原始图 G 的诱导子图，而不是修改后的 G'. 从而，本文可以在每次迭代中生成不同的图块，使其保留原始图的所有节点和边的信息.

通过这种方式，本文在每次迭代中都能够生成一组不同的图块，这些图块在保留了原始图结构的基础上引入了随机性. 这种策略不仅增加了模型的泛化能力，而且还提高了其在面对未见数据时的鲁棒性. 随机扰动策略有效地打破了模型可能过度依赖的数据中的特定模式，从而有助于减轻过拟合问题，使模型能够在更广泛的真实应用场景中保持稳定而高效的性能.

3.8 模型复杂度分析

整体而言，该模型在兴趣提取、嵌入、挖掘和用户嵌入等模块中的时间复杂度都具有线性的特征. 首先，在兴趣提取模块中，模型的时间复杂度主要受 METIS 聚类方法的影响，而 METIS 的时间复杂度通常在 $O(|E| + |V| \log |V|) \sim O(|V| \log^2 |V|)$，其随着图的规模

和划分的精度而变化. 接着，在兴趣嵌入模块中，MP-GNN 的图卷积层是线性的，因为每个节点的表示更新只涉及到其邻居节点的简单加权和聚合. 兴趣挖掘模块中的 MLP-Mixer 层也具有线性的时间复杂度，因为混合器层的计算时间与兴趣子图的数量成正比. 最后，在用户嵌入模块中，平均池化操作的时间复杂度也是线性的，因为它只涉及到对输入节点或向量的简单平均操作. 因此，整体来看，GraphMLP-Mixer 模型的时间复杂度为线性增长，效率较高.

4 实验结果与分析

4.1 实验设置

4.1.1 实验数据

为了精确地衡量和展示本文推荐系统的性能，本文选择了 3 个业界公认的大规模真实世界数据集进行了深度实验. 这些数据集在多行为序列推荐领域具有极高的权威性，并已被广泛采用作为性能评估的基准.

这些数据集涵盖了用户与在线平台互动时产生的多样化的行为模式，包括点击浏览、收藏、加入购物车以及最终的购买行为. 在本文实验中，我们的任务是单目标的多行为序列推荐，模型输入是用户的多行为购买、点击、加购物车、收藏交互序列，输出是目标行为(购买)下的推荐结果.

具体来说，IJCAI 数据集是由 IJCAI 竞赛 2015 发布的，用于预测重复购买者的任务. CIKM 和 Taobao 数据集是由中国的电子商务公司阿里巴巴和淘宝发布的，它们都是中国的知名的电商平台，拥有亿级的用户量和海量的交易数据. 通过在这样具有代表性的数据集上进行实验，从而能够确保本文的推荐方法不仅在理论上是创新的，而且在实际应用中也是有效的，能够满足用户的需求并提升他们的购物体验.

在数据处理方面，本文遵循了与先前研究类似的方法，移除了交互记录少于 20 条的用户和物品，确保了数据集中的用户和物品有足够的交互数据来准确捕捉用户的偏好和行为. 接下来，本文专注于购买行为，因为它能直接为平台带来财务上的利益. 为了确保每个用户序列中这种行为有足够的代表性，本文要求它至少出现 5 次. 除了上述步骤之外，本文还进行了常规的数据去重和清理工作，以增强实验结果的稳定性和可复现性. 处理后的数据集的详细统计信息总结在表 1 中.

Table 1 Statistics of the Experimental Datasets

表 1 实验数据集的统计信息

统计条目	数据集		
	CIKM	IJCAI	Taobao
用户数量	254 356	324 859	279 052
物品数量	521 900	331 064	731 517
交互数量	18 824 670	46 694 666	32 758 555
数据密度	1.4e–4	4.3e–4	1.6e–4
平均序列长度	74.01	143.74	117.39

4.1.2 对比基线

为了全面验证 GraphMLP-Mixer 方法的有效性，本文进行了一系列详尽的实验，并将其与一系列当前最先进的基准方法进行了深入的比较. 这些基准方法涵盖了从传统的序列推荐到多行为序列推荐的多个领域，以确保本文的比较是全面和详尽的.

在传统的序列推荐领域，本文选择了 5 个具有代表性的方法进行比较：

1) SASRec[5]. 该方法采用了基于变换器的编码器来学习序列和物品表示，能够有效地建模不同物品之间的相关性. 变换器的自注意力机制使其能够捕捉到序列中的长距离依赖关系，从而提高推荐的准确性. SASRec 在推荐系统领域被广泛应用，其基于变换器的架构能够有效地处理序列数据，捕捉到序列中的复杂关系.

2) BERT4Rec[6]. BERT4Rec 采用双向编码器和变换器来建模序列信息，并通过预测掩码目标进行优化. 这种方法的双向性质使其能够同时利用序列的前向和后向信息，从而更好地理解用户的兴趣和偏好. BERT4Rec 在推荐系统中的应用取得了显著的成果，其双向编码器和变换器架构能够有效地捕捉到序列中的长距离依赖关系.

3) CL4SRec[36]. CL4SRec 将对比自监督学习与基于变换器的 SR 模型相结合，通过裁剪、遮罩和重排等增强运算符来提高模型的泛化能力. 这种方法在不依赖外部监督信号的情况下，能够有效地学习到序列中的有用信息，其对比自监督学习和变换器架构能够有效地捕捉到序列中的复杂关系.

4) FMLP-Rec[37]. FMLP-Rec 结合了快速傅里叶变换 (FFT) 和逆 FFT 过程，以最小化噪声的影响并改进序列表示. 这种方法在频域中处理序列数据，能够更有效地捕捉到序列中的周期性模式，其在推荐系统中的应用取得了良好的效果，能够高效捕捉到序列中的周期性模式.

5）FEARec. FEARec 在频域中改进了原始时间域的自注意力，采用斜坡结构显式地学习低频和高频信息. 其频域自注意力和斜坡结构能够有效地捕捉到序列中的长期和短期模式，从而高效建模用户的长短期兴趣，从而提高推荐的准确性.

同时针对多行为序列推荐，本文也选择 6 个具有代表性的方法进行比较：

1）MATN[23]. MATN 使用注意力网络和记忆单元来学习多种行为类型之间的交互依赖关系. 它只编码用户与物品之间的局部连接模式，无法捕捉高阶依赖关系. 这种方法在处理多行为数据时，能够有效地捕捉到不同行为之间的关系.

2）NMTR[21]. NMTR 将神经协同过滤和多任务学习范式结合起来，以预定义的相关性建模用户的行为依赖性. 这种方法能够同时处理多个任务，从而提高推荐的准确性和效率.

3）MB-GCN[20]. MB-GCN 旨在解决多类型用户行为数据建模中的数据稀疏性问题. 这种方法采用图卷积运算进行消息传递，在构建的用户-物品交互图上捕捉多行为模式. 图卷积网络能够有效地捕捉到图结构数据中的局部信息和全局信息，从而提高推荐系统的准确性.

4）KHGT[12]. KHGT 引入了一个基于变换器的多行为建模方法，重点是进行时序信息和辅助知识的融合. 该模型利用图注意力网络来捕捉行为嵌入. 这种方法能够同时利用时序信息和辅助知识，从而提高推荐的准确性和解释性.

5）CML[13]. CML 通过提出元对比编码，将对比学习融入到多行为推荐中，这使得模型能够学习个性化的行为特征. 对比学习能够有效地捕捉到不同行为之间的差异和相似性，从而提高推荐的准确性.

6）KMCLR[14]. KMCLR 是一种新颖的框架，通过 2 个对比学习任务和 3 个功能模块（多用户行为学习、知识图增强、用户行为的粗粒度和细粒度建模）来提升推荐系统的性能. 这种方法能够同时处理多个任务，并利用知识图来增强推荐的可解释性.

通过与这些先进对比方法进行充分比较，本文旨在全面展示 GraphMLP-Mixer 方法在用户多行为建模的优异兴趣建模能力以及预测性能.

4.1.3 评估指标

为了深入探索推荐系统的性能评估领域，本文精心选择了 2 个广泛认可的评估指标：命中率（hit ratio, *HR@K*）和归一化折损累积增益（normalized discounted cumulative gain, *NDCG@K*）.

HR@K 是一个直观且直接的评估指标，它测量的是在 top-*K* 推荐列表中，用户实际交互过的物品被成功推荐的比率. 具体来说，如果至少 1 个相关物品出现在 top-*K* 推荐列表中，那么这个查询的 *HR@K* 值就被记为 1；如果没有相关物品出现在列表中，则记为 0. 最终，本文计算所有查询的 *HR@K* 值的平均值，以此来衡量各个推荐算法的整体性能.

NDCG@K 则是一个更为精细的评估指标，它不仅考虑了推荐物品的相关性，还考虑了物品在推荐列表中的位置. *NDCG@K* 的计算基于累积增益（CG），但引入了折损因子，使得排名靠前的物品对最终得分贡献更大. 此外，通过归一化处理，*NDCG@K* 能够使得不同查询之间的得分具有可比性. *NDCG@K* 的值越接近 1，表明推荐系统的性能越好.

为了确保本文的评估结果具有说服力和可比性，本文采取了一系列措施来保证评估过程的公平性和有效性. 首先，本文确保每个正实例，即用户实际交互过的物品都与 99 个随机选择的非交互物品配对，这样的设置能够模拟真实的推荐场景，并确保评估结果不会受到数据集中特定交互模式的影响. 这种做法与近期最先进的研究保持一致，从而使得本文的评估结果能够与其他研究直接比较.

在评估方法上，本文采用了留一法（leave-one-out），这是一种常用的评估策略，特别适用于序列推荐任务的性能评估. 具体来说，本文将每个用户时间上有序的最后一次购买作为测试样本，即作为本文需要预测的目标物品. 而前一次购买则作为验证样本，用于调整模型的超参数和避免过拟合. 这种方法的优点在于它能够充分利用用户的历史行为数据，同时保证了测试样本的独立性，从而能够更准确地评估推荐系统的实际性能.

4.1.4 实验细节

为了确保不同模型之间的比较在一个公平和统一的标准下进行，本文采取了一系列严格的实验设置. 这些设置旨在消除由于参数配置不同而导致的性能差异，从而保证评估结果的准确性和可比性. 在所有参与比较的模型中，本文首先统一了嵌入大小，将其设定为 128 维. 这一决策基于嵌入维度对模型性能和计算效率之间平衡的考量，批量大小被统一设定为 512. 这个选择是基于在模型的优化过程中多次实验的经验，本文采用了 Adam 优化算法，这是因为 Adam 优化算法在许多推荐系统任务中表现出了优异的性能. 它能够自适应地调整每个参数的学习率，从而加快收敛速度，并提高模型在训练数据上的表现.

为了寻找每个模型的最佳超参数，本文采用了网格搜索(grid search)方法. 这种方法通过系统地遍历超参数的所有可能组合，帮助本文找到最优的超参数配置. 本文还设置了一个严格的停止条件，以避免不必要的计算资源浪费，并确保模型不会过拟合. 具体来说，本文将最大迭代次数设定为 1 000. 如果在验证数据集上的 *NDCG@K* 指标在连续 20 个迭代周期内没有观察到改进，训练过程将自动停止.

对于基线模型，本文严格遵循了它们原始论文中推荐的超参数设置. 这一决定是为了确保本文的比较结果尽可能地与原始研究保持一致，从而提高比较的公正性和可信度. 最后，本文使用 Xavier 初始化方法来初始化所有模型的参数. Xavier 初始化是一种流行的参数初始化策略，它能够帮助模型更快地收敛.

4.2 总体性能

在表 2 中，本文全面展示了各种方法在不同数据集上的性能比较，并总结出观察结果.

Table 2 Overall Performance Comparison of All Methods in Terms of *HR@K* and *NDCG@K*

表 2 所有方法在 *HR@K* 和 *NDCG@K* 方面的整体性能比较

评估方法	CIKM				Taobao				IJCAI			
	H@10	*N*@10	*H*@20	*N*@20	*H*@10	*N*@10	*H*@20	*N*@20	*H*@10	*N*@10	*H*@20	*N*@20
SASRec	0.305 5	0.186 1	0.398 1	0.228 5	0.299 5	0.176 6	0.383 4	0.212 9	0.350 5	0.187 8	0.455 3	0.234 6
BERT4Rec	0.335 0	0.214 7	0.423 5	0.233 7	0.285 6	0.167 0	0.367 9	0.199 1	0.374 8	0.206 0	0.489 0	0.265 8
CL4Rec	0.342 2	0.214 0	0.439 2	0.247 2	0.313 2	0.184 0	0.402 2	0.220 6	0.387 7	0.210 4	0.504 7	0.267 2
FMLP-Rec	0.376 4	0.235 4	0.483 1	0.271 9	0.344 6	0.202 4	0.442 4	0.242 7	0.426 5	0.231 5	0.555 2	0.294 0
FEARec	0.339 2	0.209 3	0.438 6	0.249 3	0.321 3	0.189 1	0.412 0	0.227 4	0.386 6	0.208 5	0.502 7	0.262 6
MATN	0.387 7	0.242 4	0.497 6	0.280 1	0.354 9	0.208 4	0.455 7	0.250 0	0.419 2	0.223 4	0.471 8	0.262 8
NMTR	0.394 6	0.246 5	0.506 8	0.289 7	0.368 6	0.213 3	0.468 6	0.258 5	0.424 8	0.229 9	0.512 8	0.262 0
MB-GCN	0.388 9	0.213 6	0.495 1	0.289 4	0.362 8	0.201 6	0.451 0	0.249 8	0.362 7	0.196 9	0.472 1	0.250 2
KHGT	0.401 4	0.230 5	0.516 0	0.299 3	0.382 4	0.218 1	0.481 5	0.267 0	0.410 4	0.221 4	0.533 7	0.279 1
CML	0.423 4	0.246 6	0.540 0	0.306 0	0.4212	0.2420	0.5311	0.2942	0.434 4	0.236 6	0.565 8	0.301 8
KMCLR	0.4344	0.2494	0.5584	0.3239	0.358 7	0.207 0	0.456 3	0.250 3	0.4441	0.2397	0.5769	0.3051
GraphMLP-Mixer	**0.456 3**	**0.251 9**	**0.581 5**	**0.339 7**	**0.432 6**	**0.248 1**	**0.545 3**	**0.302 1**	**0.464 4**	**0.252 3**	**0.594 6**	**0.321 1**

注：*K*=10, 20；*p* 值<0.05；黑体数值表示最优值，下划线数值表示次优值；*H@K* 为 *HR@K* 的简写，*N@K* 为 *NDCG@K* 的简写.

首先，多行为方法通常优于通用序列推荐方法，这凸显了本文研究基础问题的意义. 通过使用多种行为数据作为辅助信息来建模不同用户行为之间的复杂关系，本文有效地将通用偏好转换为特定行为以提高性能.

其次，基于 GNN 的方法与传统方法的结果进行比较，可以得出基于图网络的方法通常比基于简单注意力的方法表现更好的结论，这进一步证明了先前工作的结论，即 GNN 具有更强大的建模能力，它们可以表示用户和物品更高阶的信息.

最后，GraphMLP-Mixer 在所有数据集上始终优于其他基线，展示了其有效性和泛化能力. GraphMLP-Mixer 能够有效捕捉用户行为的全局依赖性，并减轻信息过压缩的问题. 通过整合全局物品图的信息，以及利用 MLP-Mixer 多维度上的信息挖掘能力，该模型实现了对用户行为序列的深层次理解和表示. 这种设计使得 GraphMLP-Mixer 在处理多行为序列推荐问题上具有显著的优势.

4.3 消融实验

为了深入分析 GraphMLP-Mixer 模型中各个组件的作用和性能，本文研究精心设计了一系列对比实验，以评估模型中关键组件的有效性. 这些实验包括 5 种变体：

1)不使用全局对比数据增强方法的 w/o global 变体. 在这个变体中，本文移除了全局数据增强方法，以此来单独评估这一组件对模型性能的影响. 全局数据增强通过构建物品全局图，能够帮助模型更好地捕捉序列间的关系信息.

2)在兴趣挖掘模块中不使用令牌混合的 w/o tm 变体. 在这个变体中，本文移除了兴趣挖掘模块中的令牌混合(token mixing)功能.

3)在兴趣挖掘模块中不使用通道混合的 w/o cm 变体. 在这个变体中，本文移除了兴趣挖掘模块中的通道混合(channel mixing)功能.

4)在模型中不使用时序信息增强的 wo/ se 变体. 在这个变体中，本文移除了时序信息增强策略.

5)在模型中不使用数据增强策略的 wo/ de 变体. 在这个变体中，本文移除了数据增强模块.

通过对比这些变体与原始模型的性能，本文能够更准确地识别出 GraphMLP-Mixer 模型中哪些组件对整体性能至关重要，哪些组件对提升推荐系统的准确性和效率具有显著影响. 实验结果如表 3 所示.

Table 3 Ablation Study on Key Modules of GraphMLP-Mixer

表 3 GraphMLP-Mixer 中关键模块的消融实验

评估方法	CIKM		IJCAI	
	H@20	*N*@20	*H*@20	*N*@20
w/o global	0.528 0	0.302 7	0.506 3	0.280 6
w/o tm	0.552 2	0.320 3	0.515 2	0.285 6
w/o cm	0.549 2	0.315 0	0.493 7	0.272 2
w/o de	0.534 2	0.303 9	0.505 4	0.280 0
w/o te	0.565 6	0.329 9	0.524 0	0.289 8
GraphMLP-Mixer	0.581 5	0.339 7	0.545 3	0.302 1

注：*H*@20 为 *HR*@20 的简写，*N*@20 为 *NDCG*@20 的简写.

w/o global 实验结果表明，本文提出的针对多行为序列推荐的全局对比数据增强方法，通过构建全局物品图、利用物品间的协同关系，有效地增强了用户交互子图的表征能力. 这一结果凸显了全局数据增强在提升模型性能中的关键作用. w/o tm 和 w/o cm 的实验结果进一步证实了 GraphMLP-Mixer 中令牌融合和通道融合的重要性. 令牌混合关注于兴趣间的信息挖掘，而通道混合则专注于兴趣内的信息挖掘. 这 2 种混合方式共同作用，使得 GraphMLP-Mixer 能够更全面地捕捉用户行为序列中的多维度偏好，从而提高推荐系统的准确性. w/o te 变体证明了时序信息编码有助于模型识别不同兴趣图块之间的关系，进一步强化了图模型对用户兴趣的时序建模能力，从而提高推荐系统的准确性和效果. w/o de 变体证明了随机扰动策略能有效地打破模型可能过度依赖数据中的特定模式，从而有助于减轻过拟合问题，使模型能够在更广泛的真实应用场景中保持稳定而高效的性能. 综上所述，本文研究的实验结果充分验证了 GraphMLP-Mixer 模型中各个组件的有效性，尤其是全局对比数据增强方法、令牌混合和通道混合在提升模型性能方面的重要作用.

4.4 效率实验

在实验环境中，本文对 GraphMLP-Mixer 模型在不同数据集上训练一轮(epoch)所需的时间进行了详细的比较分析. 为了确保公平地比较，本文选取了性能最佳的基线模型 CML 和 KMCLR 作为参照，对其时间效率进行了全面的对比评估.

实验中，本文采用了 Linux 操作系统，并使用了 cudatoolkit 的 10.1 版本. 实验所使用的 GPU 设备为搭载 12 GB 显存的高性能 TITAN V 显卡. 此外，为了衡量不同模型在处理大规模数据时的性能，本文特别将批处理大小设置为 4 096. 这一设置旨在模拟真实世界应用中可能遇到的大规模数据处理场景，从而更准确地评估模型在实际应用中的性能表现.

表 4 展示了 GraphMLP-Mixer 模型在不同数据集上的时间效率. 本文将其与表现最佳的基线模型 CML 和 KMCLR 进行了时间效率的对比. 从表 4 中可以明显看出，GraphMLP-Mixer 在每个数据集上的训练时间都显著少于 KMCLR 和 CML 模型.

Table 4 Time Efficiency of GraphMLP-Mixer

表 4 GraphMLP-Mixer 的时间效率 min/epoch

数据集	GraphMLP-Mixer	KMCLR	CML
CIKM	55.87	166.08	216.516
Taobao	82.29	239.31	314.127
IJCAI	85.01	255.98	332.394

这些数据揭示了 GraphMLP-Mixer 模型在训练时间上的显著优势，这得益于其高效的模型架构和较低的时间复杂度. GraphMLP-Mixer 采用了 MLP-Mixer 架构与 GNN 的结合，通过全局物品图增强和对比学习策略，实现了对用户行为序列的深层次理解和表示. 这种结合不仅保留了 GNN 在图结构数据上的强大建模能力，还引入了 MLP-Mixer 在处理序列数据时的优势，如高效的标记混合和通道混合能力. 此外，GraphMLP-Mixer 的时间效率与空间效率得到了显著提升，其复杂度与用户交互行为的数量呈线性关系，这使其在处理日益增长的用户行为序列时，相较于基于 GNN 的传统方法具有更高的效率.

总体而言，GraphMLP-Mixer 模型在时间效率方面展现出了显著的优势，显示出了其在实际应用中的巨大潜力.

4.5 图聚类参数实验

为了验证兴趣聚类对模型效果的影响，我们使用了不同的聚类簇数进行超参数实验. 如表 5 所示，我们对 3 个数据集在不同的聚类簇数下进行了统计，以评估参数对最终用户推荐结果的影响，实验评价指标为 *NDCG*@10.

Table 5 Cluster Parameter Experiment of GraphMLP-Mixer

表 5 GraphMLP-Mixer 的聚类簇参数实验

聚类簇数	CIKM	Taobao	IJCAI
4	0.242 5	0.223 4	0.236 9
8	0.251 9	0.248 1	0.249 8
12	0.250 6	0.244 8	0.252 3
16	0.246 4	0.237 3	0.246 4

结果显示, 模型对聚类簇数并不是非常敏感. 聚类簇数过少可能会限制模型对用户兴趣的建模能力, 而过多则容易导致模型过拟合. 因此, 我们在实验中发现, 当聚类数为 8 或 12 时, 模型通常能够取得最佳的结果. 这表明在保持模型兴趣建模能力的同时, 避免了过度拟合的情况, 使模型在性能和泛化能力之间取得了良好的平衡.

5 结 论

本文针对用户多行为序列推荐问题进行了深入研究, 并提出了一种新的全局数据增强方法, 通过构建全局物品图, 增强了用户交互子图的表征. 在此基础上, 本文设计了一种新颖高效图神经网络结构——GraphMLP-Mixer, 其具有 2 个显著优势: 一是能够有效捕捉用户行为的全局依赖性, 同时减轻信息过压缩问题; 二是其时间与空间效率显著提高, 其复杂度与用户交互行为的数量成线性关系, 优于现有的基于 GNN 的多行为序列推荐模型. 实验结果表明, GraphMLP-Mixer 在 3 个真实公开数据集上的表现优于现有基于 GNN 的多行为推荐方法, 在实现显著的性能提升同时有效降低了模型复杂度, 显示出了其在实际应用中的巨大潜力.

虽然 GraphMLP-Mixer 在处理多行为序列推荐问题方面取得了显著的性能提升, 但其计算复杂度仍然有进一步的提升空间, 比如在兴趣提取模块中采用更高效的图聚类算法, 也可以尝试设计更高效的 GNN 结构, 以进一步降低模型的计算复杂度. 同时, 在全局图中, 我们未考虑用户与商品交互过程中的多种行为, 未来研究可以尝试将用户的行为多样性信息融入全局图的建模中, 通过多行为异构图更准确地捕捉用户的行为模式和兴趣变化. 最后, 在未来的研究工作中, 也可以进一步探索其他类型的用户行为数据, 如评论、评分等以进一步提高序列推荐性能.

作者贡献声明: 卢晓凯提出了算法思路和实验方案、完成实验并撰写论文; 封军参与论文算法的讨论并协助论文写作; 韩永强负责论文修改; 王皓、陈恩红提出指导意见并修改论文.

参 考 文 献

[1] Wang Shoujin, Hu Liang, Wang Yan, et al. Sequential recommender systems: Challenges, progress and prospects[J]. arXiv preprint, arXiv: 2001.04830, 2019

[2] Hidasi B, Karatzoglou A, Baltrunas L, et al. Session-based recommendations with recurrent neural networks[J]. arXiv preprint, arXiv: 1511.06939, 2015

[3] Tang Jiaxi, Wang Ke. Personalized top-*n* sequential recommendation via convolutional sequence embedding[C]//Proc of the 11th ACM Int Conf on Web Search and Data Mining. New York: ACM, 2018: 565−573

[4] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[J]. Advances in Neural Information Processing Systems, 2017, 30

[5] Kang W, McAuley J. Self-attentive sequential recommendation[C]// Proc of 2018 IEEE Int Conf on Data Mining (ICDM). Piscataway, NJ: IEEE, 2018: 197−206

[6] Sun Fei, Liu Jun, Wu Jian, et al. BERT4Rec: Sequential recommendation with bidirectional encoder representations from transformer[C]//Proc of the 28th ACM Int Conf on Information and Knowledge Management. New York: ACM, 2019: 1441−1450

[7] Wu Zonghan, Pan Shirui, Chen Fengwen, et al. A comprehensive survey on graph neural networks[J]. IEEE Transactions on Neural Networks and Learning Systems, 2020, 32(1): 4−24

[8] Wu Shu, Tang Yuyuan, Zhu Yanqiao, et al. Session-based recommendation with graph neural networks[C]//Proc of the AAAI Conf on Artificial Intelligence. Palo Alto, CA: AAAI, 2019, 33(1): 346−353

[9] Xu Chengfeng, Zhao Pengpeng, Liu Yanchi, et al. Graph contextualized self-attention network for session-based recommendation[C]//Freiburg: IJCAI. 2019, 19: 3940−3946

[10] Zhao Zhe, Cheng Zhiyuan, Hong Lichan, et al. Improving user topic interest profiles by behavior factorization[C]//Proc of the 24th Int Conf on World Wide Web. New York: ACM, 2015: 1406−1416

[11] Jin Bowen, Gao Chen, He Xiangnan, et al. Multi-behavior recommendation with graph convolutional networks[C]//Proc of the 43rd Int ACM SIGIR Conf on Research and Development in Information Retrieval. New York: ACM, 2020: 659−668

[12] Xia Lianghao, Huang Chao, Xu Yong, et al. Knowledge-enhanced hierarchical graph transformer network for multi-behavior recommendation[C]//Proc of the AAAI Conf on Artificial Intelligence. Palo Alto, CA: AAAI, 2021, 35(5): 4486−4493

[13] Wei Wei, Huang Chao, Xia Lianghao, et al. Contrastive meta learning with behavior multiplicity for recommendation[C]//Proc of the 15th ACM Int Conf on web search and data mining. New York: ACM, 2022: 1120−1128

[14] Xuan Hongrui, Liu Yi, Li Bohan, et al. Knowledge enhancement for contrastive multi-behavior recommendation[C]//Proc of the 16th ACM Int Conf on Web Search and Data Mining. New York: ACM, 2023: 195−203

[15] Xu Keyulu, Hu Weihua, Leskovec J, et al. How powerful are graph neural networks?[J]. arXiv preprint, arXiv: 1810.00826, 2018
[16] Tolstikhin I O, Houlsby N, Kolesnikov A, et al. MLP-Mixer: An all-MLP architecture for vision[J]. Advances in Neural Information Processing Systems, 2021, 34: 24261−24272
[17] Rendle S, Freudenthaler C, Schmidt-Thieme L. Factorizing personalized Markov chains for next-basket recommendation[C]// Proc of the 19th Int Conf on World Wide Web. New York: ACM, 2010: 811−820
[18] He Ruining, McAuley J. Fusing similarity models with Markov chains for sparse sequential recommendation[C]//Proc of 2016 IEEE 16th Int Conf on Data Mining (ICDM). Piscataway, NJ: IEEE, 2016: 191−200
[19] Huang Chao, Chen Jiahui, Xia Lianghao, et al. Graph-enhanced multi-task learning of multi-level transition dynamics for session-based recommendation[C]//Proc of the AAAI Conf on Artificial Intelligence. Palo Alto, CA: AAAI, 2021, 35(5): 4123−4130
[20] Wang Ziyang, Wei Wei, Cong Gao, et al. Global context enhanced graph neural networks for session-based recommendation[C]//Proc of the 43rd Int ACM SIGIR Conf on Research and Development in Information Retrieval. New York: ACM, 2020: 169−178
[21] Gao Chen, He Xiangnan, Gan Dahua, et al. Neural multi-task recommendation from multi-behavior data[C]//Proc of 2019 IEEE 35th Int Conf on Data Engineering (ICDE). Piscataway, NJ: IEEE, 2019: 1554−1557
[22] Guo Long, Hua Lifeng, Jia Rongfei, et al. Buying or browsing?: Predicting real-time purchasing intent using attention-based deep network with multiple behavior[C]//Proc of the 25th ACM SIGKDD Int Conf on Knowledge Discovery & Data Mining. New York: ACM, 2019: 1984−1992
[23] Xia Lianghao, Huang Chao, Xu Yong, et al. Multiplex behavioral relation learning for recommendation via memory augmented transformer network[C]//Proc of the 43rd Int ACM SIGIR Conf on Research and Development in Information Retrieval. New York: ACM, 2020: 2397−2406
[24] Karypis G, Kumar V. A fast and high quality multilevel scheme for partitioning irregular graphs[J]. SIAM Journal on Scientific Computing, 1998, 20(1): 359−392
[25] Kipf T N, Welling M. Semi-supervised classification with graph convolutional networks[J]. arXiv preprint, arXiv: 1609.02907, 2016
[26] Bresson X, Laurent T. Residual gated graph convnets[J]. arXiv preprint, arXiv: 1711.07553, 2017
[27] Hu Weihua, Liu Bowen, Gomes J, et al. Strategies for pre-training graph neural networks[J]. arXiv preprint, arXiv: 1905.12265, 2019
[28] Dwivedi V P, Bresson X. A generalization of transformer networks to graphs[J]. arXiv preprint, arXiv: 2012.09699, 2020
[29] Rampášek L, Galkin M, Dwivedi V P, et al. Recipe for a general, powerful, scalable graph transformer[J]. Advances in Neural Information Processing Systems, 2022, 35: 14501−14515
[30] Dwivedi V P, Luu A T, Laurent T, et al. Graph neural networks with learnable structural and positional representations[J]. arXiv preprint, arXiv: 2110.07875, 2021
[31] Dwivedi V P, Joshi C K, Luu A T, et al. Benchmarking graph neural networks[J]. Journal of Machine Learning Research, 2023, 24(43): 1−48
[32] Hendrycks D, Gimpel K. Gaussian error linear units (GELUS)[J]. arXiv preprint, arXiv: 1606.08415, 2016
[33] Ba J L, Kiros J R, Hinton G E. Layer normalization[J]. arXiv preprint, arXiv: 1607.06450, 2016
[34] Liu Hanxiao, Dai Zihang, So D, et al. Pay attention to MLPS[J]. Advances in Neural Information Processing Systems, 2021, 34: 9204−9215
[35] Xie Xu, Sun Fei, Liu Zhaoyang, et al. Contrastive learning for sequential recommendation[C]//Proc of 2022 IEEE 38th Int Conf on Data Engineering (ICDE). Piscataway, NJ: IEEE, 2022: 1259−1273
[36] Zhou Kun, Yu Hui, Zhao Xin, et al. Filter-enhanced MLP is all you need for sequential recommendation[C]//Proc of the ACM Web Conf 2022. New York: ACM, 2022: 2388−2399
[37] Du Xinyu, Yuan Huanhuan, Zhao Pengpeng, et al. Frequency enhanced hybrid attention network for sequential recommendation[C]// Proc of the 46th Int ACM SIGIR Conf on Research and Development in Information Retrieval. New York: ACM, 2023: 78−88



Lu Xiaokai, born in 1980. PhD candidate, senior engineer. His main research interests include intelligent recommendation, data mining, knowledge graph, large model of industrial products, and intelligent supply chain.

卢晓凯, 1980 年生. 博士研究生, 高级工程师. 主要研究方向为智能推荐、数据挖掘、知识图谱、工业品大模型、智慧供应链.



Feng Jun, born in 1983. Master candidate, senior engineer. His main research interests include data mining, machine learning, and digital supply chain intelligence. (fengjun@youzhicai.com)

封　军, 1983 年生. 硕士研究生, 高级工程师. 主要研究方向为数据挖掘、机器学习、数智化供应链.



Han Yongqiang, born in 1999. Master. Member of CCF. His main research interests include information retrieval and recommendation algorithms. (harley@mail.ustc.edu.cn)

韩永强, 1999 年生. 硕士. CCF 会员. 主要研究方向为信息检索、推荐算法.



Wang Hao, born in 1993. PhD, special associate researcher. His main research interests include representation learning, graph mining, and large language models. (wanghao3@ustc.edu.cn)

王　皓, 1993 年生. 博士, 特任副研究员. 主要研究方向为表征学习、图挖掘、大语言模型.



Chen Enhong, born in 1968. PhD, professor. His main research interests include machine learning and knowledge discovery. (cheneh@ustc.edu.cn)

陈恩红, 1968 年生. 博士, 教授. 主要研究方向为机器学习、知识发现.