

基于多关系感知和自监督学习的超图增强会话推荐模型

王益飞, 翟雨欣

上海理工大学光电信息与计算机工程学院, 上海

收稿日期: 2024年6月6日; 录用日期: 2024年6月30日; 发布日期: 2024年7月8日

摘要

会话推荐旨在基于匿名用户的历史行为序列预测下一个最有可能交互的项目。现有的大多数基于超图解决会话推荐模型没有考虑超边同构, 忽略了会话内可能存在的多种依赖关系, 从而无法捕获会话的潜在信息。同时, 会话匿名性和短序列往往会导致数据稀疏性, 影响了推荐模型的性能。因此本文提出一种基于多关系感知和自监督学习的超图增强会话推荐模型(MS-HGNN), 该模型能够有效捕获会话间用户的复杂高阶关系和会话内顺序及共现依赖关系。具体来说, MS-HGNN模型将会话序列分别建模为超图、共现图和顺序图, 在捕获超图的同时加入位置信息以减轻超边同构, 利用门控机制融合会话内顺序依赖关系和共现依赖关系。此外, 使用自监督学习最大化两个会话表示之间的互信息以缓解数据稀疏性, 有效地融合用户的复杂高阶关系和会话内顺序及共现依赖关系。广泛的实验结果表明, 所提出的MS-HGNN模型在三个真实数据集Diginetica、Tmall和Last.fm上的性能优于目前最先进的模型。

关键词

会话推荐, 超图神经网络, 多关系感知, 自监督学习

A Hypergraph-Enhanced Session Recommendation Model Based on Multi-Relationship Awareness and Self-Supervised Learning

Yifei Wang, Yuxin Zhai

School of Optical Information and Computer Engineering, University of Shanghai for Science and Technology, Shanghai

Received: Jun. 6th, 2024; accepted: Jun. 30th, 2024; published: Jul. 8th, 2024

文章引用: 王益飞, 翟雨欣. 基于多关系感知和自监督学习的超图增强会话推荐模型[J]. 建模与仿真, 2024, 13(4): 4104-4119. DOI: 10.12677/mos.2024.134372

Abstract

Session-based recommendation aims to predict the next most likely item for interaction based on the historical behavior sequence of anonymous users. Most existing hypergraph-based solutions for session recommendation overlook hyperedge isomorphism and ignore the various dependencies that may exist within a session, thus failing to capture the latent information of the session. Meanwhile, the anonymity of sessions and the short sequences often lead to data sparsity, affecting the performance of recommendation models. Therefore, this paper proposes a Multi-Relationship Aware and Self-Supervised Learning-based Hypergraph-Enhanced Session-based Recommendation Model (MS-HGNN), which effectively captures the complex high-order relationships among users across sessions and the sequential and co-occurrence dependencies within sessions. Specifically, the MS-HGNN model models session sequences as hypergraphs, co-occurrence graphs, and sequential graphs, respectively. It incorporates positional information to mitigate hyperedge isomorphism while capturing hypergraphs, and utilizes a gating mechanism to fuse the sequential and co-occurrence dependencies within sessions. Additionally, self-supervised learning is employed to maximize the mutual information between the representations of two sessions, alleviating data sparsity and effectively integrating the complex high-order relationships among users and the sequential and co-occurrence dependencies within sessions. Extensive experimental results demonstrate that the proposed MS-HGNN model outperforms the current state-of-the-art models on three real-world datasets: Diginetica, Tmall, and Last.fm.

Keywords

Session-Based Recommendation, Hypergraph Neural Network, Multi-Relationship Awareness, Self-Supervised Learning

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

推荐系统旨在缓解信息过载[1], 提升用户体验。但传统方法依赖用户长期历史和个人信息, 面临冷启动和隐私问题时效果不佳。基于匿名会话的推荐系统通过分析用户行为来预测兴趣, 重点是如何利用短期数据精准捕捉用户意图, 同时保护隐私。因此, 研究如何在保护隐私的前提下, 通过匿名会话推荐精准捕捉用户意图具有重要意义。

会话推荐(Session-Based Recommendation, SBR)旨在预测当前会话的下一个交互项目。传统方法如协同过滤[2]受冷启动问题影响, 矩阵分解[3]则可能忽视用户真实意图, 而马尔科夫链方法[4]简化过度, 忽略了用户历史交互信息, 这些都可能影响推荐的准确性。

最近, 深度学习技术, 如循环神经网络(RNN) [5]、注意力机制或图神经网络[6] [7]的方法, 已广泛应用于会话推荐。RNN [8] [9]和注意力机制[10]方法能捕捉项目顺序依赖, 但无法学习项目间的共现关系或利用其他会话信息。图神经网络[11]-[13]可利用其他会话信息, 但处理高维图级信息困难。DHCM 模型[14]通过采用超图结构有效捕捉这类信息, 在会话推荐中展现显著优势。实验表明, 基于超图神经网络的模型[15]在会话推荐中表现突出。

基于超图神经网络的模型虽然取得了显著改进, 但仍存在以下局限性: 1) 超图结构往往缺乏时间步长信息, 因此不同的会话序列可能被构造成具有相同结构的图网络, 即超边同构问题。例如, $S_1 = \{v_1, v_2, v_3, v_4, v_1, v_3\}$ 和 $S_2 = \{v_3, v_1, v_2, v_3, v_4, v_1\}$ 是完全不同的会话, 但是在构建超图时却具有相同的图结构, 这不可避免地影响会话表示的准确性。2) 超图神经网络虽然能够有效地捕捉更高维的图级信息, 但是却忽略了会话内多种依赖关系的重要性。如图 1 所示, 购买了 Mac 后再购买 ipad, 这是一种顺序关系, 而购买 ipad 后购买 pencil 2, 亦或者购买 pencil 2 后再购买 ipad 这种顺序关系, 称为共现关系。值得注意的是, 共现项之间的转换关系也存在差异, 购买 ipad 后购买 pencil 2 的可能性明显要大于购买 pencil 2 后再购买 ipad。

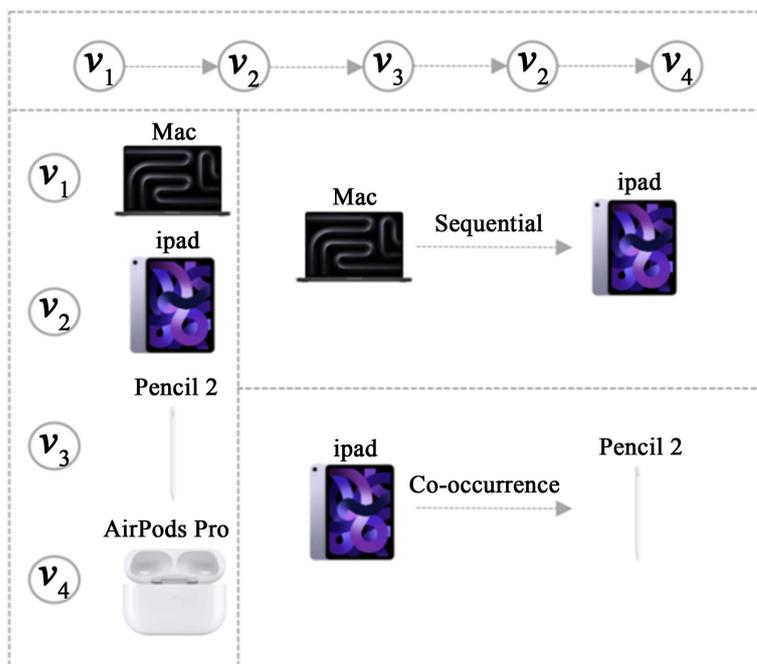


Figure 1. Dependency relationship description
图 1. 依赖关系说明

针对上述问题, 本文提出一种基于多关系感知和自监督学习的超图增强会话推荐模型(MS-HGNN)。该模型通过超图对会话中复杂的多元关系进行建模, 同时保留原始序列信息来缓解超边同构问题。基于绘画序列构建共现图和顺序图结构, 利用门控机制融合会话内顺序依赖关系和共现依赖关系。与此同时, 为了有效地融合会话间的复杂高阶关系和会话内顺序及共现依赖关系, 本文还使用自监督学习最大化两个会话表示之间的互信息, 一定程度上缓解了数据稀疏性, 提高了推荐任务的性能。

本文的主要贡献总结如下:

- a) 构建顺序图和共现图捕获项目间的顺序及共现依赖关系, 通过门控机制融合两种依赖关系以此捕获更有效的会话级信息。
- b) 使用位置编码向量解决超边同构问题。将可学习的位置编码信息引入超图会话信息中, 同时使用自监督学习改进了模型推荐性能。
- c) 通过在 Diginetica、Tmall 和 Last.fm 三个数据集上的实验表明, 本文提出的模型在会话推荐工作中优于对比模型。

本文的其余部分组织如下: 第一节介绍了本文相关工作; 在第二节中首先给出问题定义, 并对本文

提出的 MS-HGNN 模型的框架及实现方法给出详细的描述; 第三节展示了实验结果及其分析; 在第四节中对全文进行了总结, 并展望了未来工作。

2. 相关工作

在本节中, 首先回顾了会话推荐领域的研究进展, 包括传统方法、深度学习和图神经网络方法三个方面。随后, 对于本文使用到的自监督学习做了简要概述。

2.1. 传统方法用于会话推荐

传统的基于会话的推荐方法主要有基于协同过滤[2]、基于矩阵分解[3]或基于马尔科夫链[4]的方法。Sarwar B 等人[2]通过探索项目间的相关性和项目向量间的余弦相似度, 来提高基于物品推荐任务的可扩展性。Koren Y 等人[3]通过从项目评分推断出的因子向量来表征项目和用户, 对项目因子和用户因子的高相关性促成了推荐。Rendle S 等人[4]结合了矩阵分解和马尔可夫链, 利用马尔可夫链为每个用户学习个性化的转移矩阵, 来捕获顺序模式和用户长期的偏好。基于传统的 SBR 方法能够有效的利用会话中的信息, 但是只能学习到低维的潜在表示, 限制了推荐的准确性。

2.2. 深度学习方法用于会话推荐

随着深度学习的迅猛发展, 基于该技术的推荐系统方法取得了显著进步。Hidasi 等人[8]提出的 GRU4Rec 方法首次采用门控循环单元对物品交互序列进行建模。随后, Li 等人[9]提出的 NARM 在 GRU4Rec 基础上引入了注意力机制的混合编码器, 旨在同时捕捉顺序行为和当前会话的主要意图。Liu 等人[10]则通过 STAMP 模型, 利用短期注意力/记忆优先机制, 有效捕捉用户的一般和当前兴趣。而 Song 等人[16]的 ISLF 方法结合循环神经网络(RNN)和变分自编码器(VAE), 旨在提取会话中的用户偏好并考虑其兴趣的动态变化。基于深度学习的方法, 主要是基于 RNN 的方法, 通常只对连续项目之间的单向转换关系进行建模, 这可能导致无法捕获整个会话序列中复杂的项目转换关系。

2.3. 图神经网络方法用于会话推荐

近年来, 图神经网络方法在推荐系统领域备受关注。Wu 等人[11]提出的 SR-GNN 是一种基于 GNN 的创新方法, 将会话转换为有向无权图, 并利用门控 GNN 生成会话表示。Xu 等人[13]提出的 GC-SAN 通过自注意力机制改进 SR-GNN, 以捕获项目之间的长期依赖关系。Pan 等人[17]提出的 SGNN-HN 为每个会话图引入一个星型节点, 以捕获项目之间的长距离关系。虽然这些研究已取得了令人鼓舞的成果, 但仅从短会话中独立学习可能固有地不足以准确揭示用户的真实意图。因此, 研究方法开始转向于结合其他会话的协作信息来为当前会话提供服务, 开始将超图引入推荐任务当中。

Wang 等人[18]提出 SHARE 将会话序列构建成超图, 并使用上下文滑动窗口来建模项目之间的相关性, 并且利用 HGAT 来区分具有不同重要性的项目。Xia 等人[14]提出的 DHCN 引入了超图卷积网络, 利用超图卷积网络来捕获个体会话中的高阶项关系, 并使用自监督学习来增强会话表示。Li 等人[19]提出的 HIDE 利用超图来捕获高阶关系以及物品转换中的潜在意图, 并学习每个物品点击的分解意图块。Peng 等人[20]提出的 GC-HGNN 使用基于超图的神经网络来捕获高阶非成对的关系, 使用 GAT 来捕获局部信息, 通过和池化融合来增强推荐性能。Zhao 等人[21]提出 AC-HCN 通过自适应嵌入式的超图卷积网络来学习上下文信息, 同时使用 α -entmax 函数消除无关项目的影响, 最后将学习到的项目表征与时间间隔和反位置嵌入相结合生成增强项目表示。

上述方法虽然取得了不错的性能, 但它们都没有考虑超边同构的问题, 并且没有充分捕捉到会话内多种依赖关系。本文试图利用超图神经网络来捕获会话序列中项目之间的复杂高阶关系, 并且在项目嵌

中加入位置嵌入, 减轻超边同构问题。此外, 还捕获会话内顺序依赖关系和共现依赖关系, 在捕获共现关系的同时考虑成对项之间的不同转换关系, 通过门控机制融合两种依赖关系, 再通过自监督学习有效地融合会话间的复杂高阶关系和会话内顺序及共现依赖关系, 缓解了数据稀疏性, 从而实现更加有效的推荐。

2.4. 对比自监督学习

近年来, 对比学习结合自监督学习(SSL)在一些领域取得了不错的成绩, 包括计算机视觉(CV)、自然语言处理(NLP)、图嵌入和推荐系统[22]。对比学习基于判别思想, 其核心在于拉近相似样本间的距离, 同时让不同样本远离。在对比自监督学习中, 编码器通过最大化同一实例的两个增强视图间的一致性, 实现了对未标记数据的有效训练。这一学习过程主要涵盖两大步骤: 首先, 进行数据增强, 即对原始训练数据进行变换, 以增强其多样性。在同一训练批次中, 将原始样本的增广版本视为正样本, 而批次中的其他样本则视为负样本。其次, 进行对比训练, 即将原始样本及其增强版本作为模型输入, 得到对应的表示。随后, 建立对比损失函数, 并优化这些表示, 以提高模型的判别能力。

然而, 将自监督学习与推荐系统相结合的研究尚处于探索阶段, 存在诸多改进空间。因此, 本文将对比自监督学习与会话推荐任务相结合, 旨在通过引入对比自监督学习作为辅助任务, 共同优化推荐任务的学习过程, 从而进一步提升推荐性能。

3. 模型结构

在本节中, 首先介绍了问题定义, 然后提出一种基于多关系感知和自监督学习的超图增强会话推荐模型, MS-HGNN。该模型的框架如图 2 所示, 模型由三个模块组成: 1) 全局信息提取层: 利用 HGNN 网络来捕获项目之间的高阶特征信息, 同时融合位置信息, 使用注意力机制学习以得到会话的全局信息表示; 2) 局部信息提取层: 捕获会话间的顺序依赖和共现依赖, 通过门控操作融合两种依赖来得到项目的局部信息表示; 3) 自监督学习层: 通过自监督学习辅助任务最大化会话全局高阶信息和项目间的复杂信息, 得到最终预测结果。

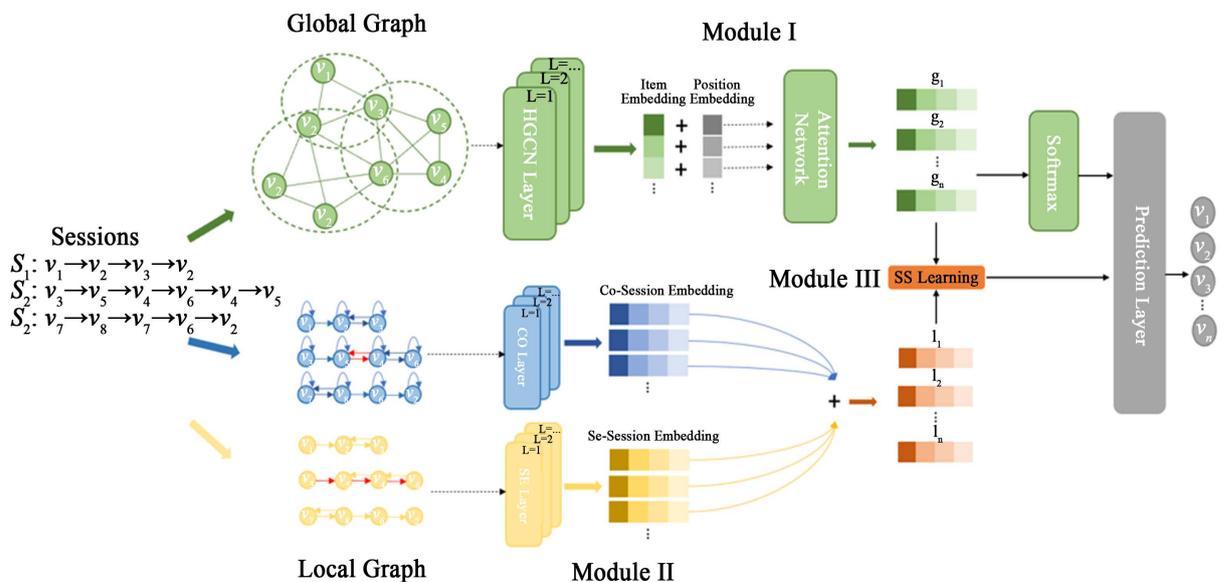


Figure 2. Structure of MS-HGNN model

图 2. MS-HGNN 模型结构

3.1. 问题定义

基于会话的推荐旨在根据匿名用户的会话序列向其推荐下一个可能感兴趣的项目。每个匿名会话序列由一组按时间顺序由与用户交互的项目组成。 $V=[v_1, v_2, \dots, v_n]$ 表示所有唯一项目的集合, 其中 n 是项目的总数。当前会话用 $S=\{v_1^s, v_2^s, \dots, v_l^s\}$ 表示, 其中 l 为会话中条目的数量, $v_i^s \in V$ 表示在 i 时刻与 S 中的匿名用户交互的项目。基于会话的推荐的目的是基于 S 预测下一个项目 v_{l+1}^s 。构建函数 $P(\cdot)$ 来预测匿名用户序列 S 中最有可能与之交互的 \hat{v}_{l+1}^s 的概率, 记 $\hat{v}_{l+1}^s = P(S, s; \theta)$, 其中 θ 是可学习参数的集合, 假设损失函数为 $L(\cdot)$, 学习使模型平均损失最小的最优参数 θ 来得到最佳的预测结果, 表示如下:

$$\theta = \arg \min_{\theta} \frac{1}{|S_i|} \sum_{i=1}^{|S_i|} L(v_{l+1}^s, \hat{v}_{l+1}^s) \quad (1)$$

其中, $|S_i|$ 表示训练集 S_i 的大小, 模型根据最优参数 θ 向用户推荐 top-K 项。

3.2. 全局信息提取层

3.2.1. 全局图的构建

全局图是基于超图构建的, 其中节点表示项目, 超边表示项目之间的关联关系。超图的一条超边可以连接 N 个节点, 超图可以定义为三元组 $G=(V, E, W)$ 表示的集合。 V 是节点集合, $|V|$ 表示节点的数量, 每个节点表示一个项目 $v \in V$ 。 E 是超边集合, $|E|$ 表示超边的数量, 每个超边表示一个会话 $e \in E$ 。 W 是超边上的权重集合, 每个超边分配一个权重 $W(e) \in W$ 。超图中的 W 可以定义为一个对角矩阵, 超图被定义为关联矩阵 T , 且 $r(v, e) \in T$, 如果 $v \in e$, 则 $r(v, e)=1$, 否则为 0 , 并且可以计算得到所有超边的度矩阵 $D = \sum_{e \in E} W(e)T(v, e)$ 和所有的顶点度矩阵 $B = \sum_{v \in V} T(v, e)$ 。

如图3所示, 在超图中如果任意两个超边存在相同的节点, 就会连接起来。在一个超边内部所有节点都是通过无向边全连接起来, 这样就构建出项目之间的高阶非成对关系用于全局级项目的特征表示。

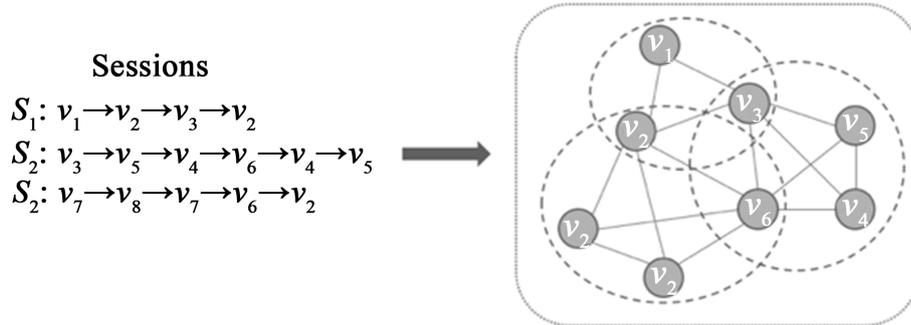


Figure 3. An example of the hypergraph construction process
图3. 超图构建过程的实例

3.2.2. 全局信息提取

本文利用超图卷积网络来学习会话序列的节点向量表示, 以捕捉图结构中节点之间复杂的高阶关系。参考[14]等人提出的超图卷积定义为:

$$h_v^{l+1} = D^{-1}TWB^{-1}T^T h_v^l \quad (2)$$

其中, h^{l+1} 表示第 $l+1$ 层项目 v 的超图卷积结果。超图卷积可以看作是在超图结构上执行“节点-超边-节点”特征转换的两阶段细化。 $T^T h_v^l$ 代表信息从节点传递到超边, 然后和 T 相乘表示信息从超边到节点。

将 h^0 通过 L 层超图卷积层后, 将每层得到的项嵌入进行平均后, 得到最终的全局级项目表示:

$$h_{v_i}^g = \frac{1}{L+1} \sum_{l=0}^L h_{v_i}^l \quad (3)$$

在图结构中, 时间序列位置信息经常缺失, 会导致不同的会话形成相同的图结构, 位置嵌入是 Transformer [23] 中提出的一种用来记忆项目的位置信息的技术。本文采用反位置编码, 使用一个可学习的位置矩阵 $P = [P_1, P_2, \dots, P_l]$, P_i 是第 i 个项目的嵌入向量, 为了使位置信息和项目信息结合, 分别进行如下计算:

$$h_i^{g*} = \tanh(W_1(h_{v_i}^g \parallel P_{l-i+1}) + b) \quad (4)$$

$$\rho_i = d^T(W_2 h_i^{g*} + W_3 h_{v_i}^g + c) \quad (5)$$

$$S_H = \sum_{i=1}^l \rho_i h_{v_i}^g$$

其中, $W_1 \in R^{d \times 2d}$ 、 $W_2, W_3 \in R^{d \times d}$ 和 $b, c \in R^d$ 均为可学习的参数, h_i^{g*} 表示当前会话第 i 个融合位置编码后的会话信息。模型先通过一个软注意力机制来学习相对应的权重 ρ_i , 再通过线性组合操作得到最终的全局会话信息表示 S_H 。

使用全局会话信息表示 S_H 和全局级项目表示 $h_{v_i}^g$ 进行点积, 再采用 softmax 函数进行归一化, 得到下一个点击项目的概率 \hat{y} :

$$\hat{y}_H = \text{softmax}(S_H^T h_{v_i}^g)$$

y 表示 ground-truth 项目的 one-hot 编码向量, 本文使用广泛使用的交叉熵函数作为最终的损失函数, 全局提取层的损失函数定义如下:

$$\ell_H = -\sum_{i=1}^n y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)$$

3.3. 局部信息提取层

3.3.1. 局部图的构建

局部图用来提取当前会话项目中的重要特征。对于提取会话内的共现依赖关系, 将其构建成有向无加权图 $G_s = (V_s, E_s)$, 这里 $V_s \subseteq V$ 表示在局部图中的节点, 也是会话 S 中被点击的项目。 $E_s = \{e_{i,j}^s\}$ 表示边的集合, 每条边 $e_{i,i+1}^s \in E_s$ 意味着在会话 S 中用户点击完 v_i^s 之后点击 v_{i+1}^s 。此外, 为了能聚合项目自身的信息, 为会话图中的每个节点添加了一个自循环。图 4 是一个共现依赖局部图构建示例。

对于提取会话间的顺序依赖关系, 将其构建成有向加权图。此外, 在一个会话中同一个项目可能会出现多次, 本文对会话中边的权值进行了归一化处理, 使得出度矩阵和入度矩阵每一行和为 1。图 5 是一个顺序依赖局部图构建示例。

3.3.2. 局部信息提取

(1) 共现依赖关系的提取: 假设给定两个序列 (v_j, v_i) 和 (v_i, v_j) , 其中 $v_i, v_j \in V$, 每个初始 v 的嵌入表示: $h_v^0 = \text{Embedding}(v_i)$ 。在 CO Layer 中, 将两个共现项的潜在向量拼接起来表示转换关系, 然后使用可训练的参数 w_i 和 b_i 来学习节点传出和传入边之间的差异, 并使用注意力机制将邻居节点的特征聚合到目标节点:

$$e_{out}^{i,j} = A_{out}^{i,j} \tanh\left(\left(h_{v_i}^0 \parallel h_{v_j}^0\right) w_t + b_t\right)$$

$$e_{in}^{i,j} = A_{in}^{i,j} \tanh\left(\left(h_{v_j}^0 \parallel h_{v_i}^0\right) w_t + b_t\right)$$

$$e^{i,j} = \frac{1}{2}\left(e_{out}^{i,j} + e_{in}^{i,j}\right)$$

其中 \parallel 代表拼接操作, \tanh 作为激活函数。 $e_{out}^{i,j}$ 和 $e_{in}^{i,j}$ 表示节点 v_j 对节点 v_i 的重要性, 这样就得到一个节点 v_i 在共现关系上两个不同的注意力系数, $e^{i,j}$ 表示节点 v_j 对节点 v_i 的整体重要性, 然后再考虑到非共现关系的其他邻居节点的信息和自身的节点信息, 即将所有与节点 v_i 有关的相邻系数归一化, 采用 softmax 函数实现:

$$\phi^{i,j} = \frac{\exp\left(e^{i,j}\right)}{\sum_{k:A_{out}^{i,k}+A_{in}^{i,k} \neq 0} \exp\left(e^{i,k}\right)}$$

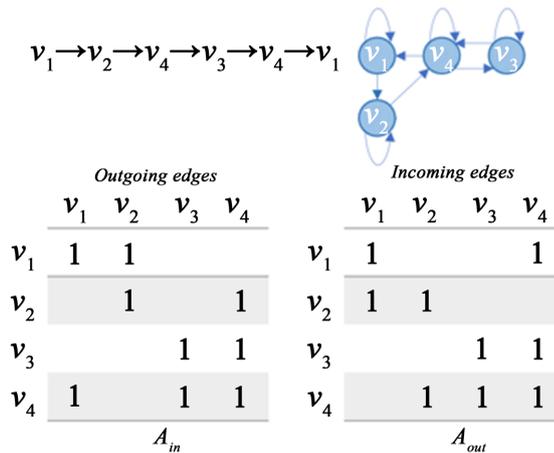


Figure 4. Co-occurrence dependent local graph building process instance
图 4. 共现依赖局部图构建过程实例

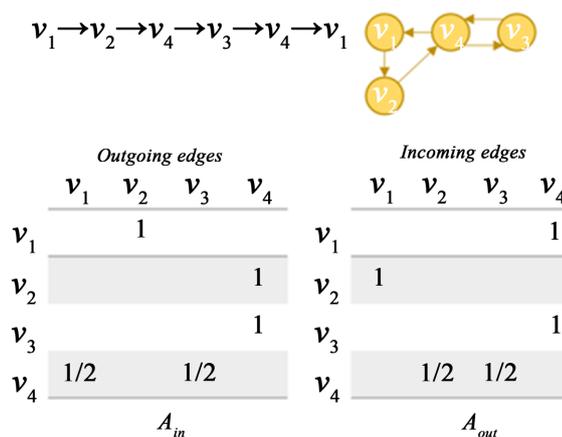


Figure 5. The sequence depends on the local graph building process instance
图 5. 顺序依赖局部图构建过程实例

最后根据上面得到的注意系数 $\phi^{i,j}$ 就可以使目标节点自适应聚合其他邻居节点的信息和自身的节点信息:

$$h_{v_i}^1 = \sum_{j: A_{v_i v_j}^i + A_{v_j v_i}^i \neq 0} \varphi^{i,j} h_{v_j}^0$$

再进一步, 通过堆叠多层图卷积来聚合高阶邻居信息, 能有效区分成对共现项之间的关系。这样, 第 k 层的 v_i 的项嵌入可表示为:

$$h_{v_i}^k = CR-GNN(h_{v_i}^{k-1})$$

和之前的工作一致[24], 使用 W_Q 、 W_K 和 W_V 来对项目编码 $h_{v_i}^k = [h_{v_1}^k, h_{v_2}^k, \dots, h_{v_m}^k]$ 进行变换, 得到对应的查询向量、键向量和值向量, 然后生成最终的共现依赖会话表示 S_{CO} , 具体操作如下:

$$\xi_{mi} = \frac{S(W_Q h_{v_m}^k, W_K h_{v_i}^k)}{\sum_{j < m} S(W_Q h_{v_m}^k, W_K h_{v_j}^k)}$$

$$S_{CO} = \sum_{i \leq m} \xi_{mi} W_V h_{v_i}^k$$

(2) 顺序依赖关系的提取: 在 SE Layer 中采用图注意力网络学习相邻节点之间的权重, 并获取图中各个节点的特征, 具体计算过程如下所示:

$$s_{ij} = f\left(a_{r_{ij}}^\top (h_{v_i}^0 \odot h_{v_j}^0)\right)$$

$$\alpha_{ij} = \text{softmax}(s_{ij})$$

$$h_{v_i}^1 = \sum_{v_j \in N} \alpha_{ij} h_{v_j}^0$$

其中, s_{ij} 表示节点 v_i 对节点 v_j 的重要性, r_{ij} 为 v_i 与 v_j 之间的边关系, $a \in \mathbb{R}^d$ 表示权重, $f(\cdot)$ 表示激活函数, 本文使用 LeakyRelu 函数, h^k 为 k 次多层扩展处理的结果。

反位置编码在缓解不同的会话形成相同的图结构的问题外, 同样可以缓解图神经网络在对项目编码时不能充分学习项目间顺序依赖关系的问题。同样使用一个可学习的位置矩阵 $P = [P_1, P_2, \dots, P_l]$, P_i 是第 i 个项目的嵌入向量, 为了使位置信息和项目信息结合, 分别进行如下计算:

$$h_i^{se} = \tanh\left(W_1 \left(h_{v_i}^k \parallel P_{l-i+1}\right) + b\right)$$

$$\rho_i = d^T \left(W_2 h_i^{se} + W_3 h_{v_i}^k + c\right)$$

$$S_{se} = \sum_{i=1}^l \rho_i h_{v_i}^k$$

其中, $W_1 \in \mathbb{R}^{d \times 2d}$ 、 $W_2, W_3 \in \mathbb{R}^{d \times d}$ 和 $b, c \in \mathbb{R}^d$ 均为可学习的参数, h_i^* 表示当前会话第 i 个融合位置编码后的会话信息。模型先通过一个软注意力机制来学习相对应的权重 ρ_i , 再通过线性组合操作得到最终的顺序依赖会话表示 S_{se} 。

(3) 共现关系与顺序关系的融合: 使用门控操作将共现依赖会话表示和顺序依赖会话表示结合, 得到更准确的局部级会话表示, 即:

$$f = \sigma(W_{co} S_{co} + W_{se} S_{se})$$

$$S_{LC} = f \cdot S_{se} + (1-f) \cdot S_{co}$$

其中, W_{co} 和 W_{se} 为可学习参数, f 为门控单元。

3.4. 自监督学习层

通过 3.2 与 3.3 小节运算得到会话表示 S_H 和 S_{LC} , 由于每个会话只描述用户的复杂高阶关系和会话内顺序及共现依赖关系, 各会话间表示信息是不同的。自监督学习可以用于标签增强, 两组信息可以作为彼此的 ground-truth, 如果两个会话嵌入表示相同的会话, 则将这一对会话嵌入标记为 ground-truth, 否则将其标记为 negative。在此基础上, 对比两种学习到的会话嵌入:

$$\ell_L = -\log \sigma \left(f_D \left(S_H^i, S_{CO}^i \right) \right) - \log \sigma \left(1 - f_D \left(\tilde{S}_H^i, S_{CO}^i \right) \right)$$

其中 \tilde{S}_H^i 是对 S_H^i 按照行列变换打乱顺序获得的负样本。 $f_D(\cdot)$ 是鉴别器函数, 用于对两个向量的一致性进行评分。因此, 最终目标函数可表示为:

$$\ell = \ell_H + \beta \ell_L$$

β 用于控制自监督学习任务的大小。

4. 实验

本节中, 首先描述了数据集和预处理方法; 然后介绍评估机制、参与对比的基线模型、模型参数设置; 最后, 给出了实验结果和分析。回答如下问题:

- RQ1: MS-HGNN 模型是否优于基线模型?
- RQ2: MS-HGNN 模型的各个组成部分是否对推荐性能提高做出了贡献?
- RQ3: MS-HGNN 模型在局部级信息对比使用不同融合函数的表现如何?
- RQ4: 采用不同的参数设置对模型的最终推荐结果有什么影响?

4.1. 数据集和预处理

为了提升实验结果的说服力, 本文选取了来自不同应用领域的三个数据集: Diginetica、Tmall 和 Last.fm。其中, Diginetica 源自 CIKM Cup 2016, 包含丰富的交易数据; Tmall 则来自 IJCAI-15 竞赛, 涵盖天猫平台上的匿名用户购物日志; Last.fm 则是一个广泛应用于推荐任务的数据集。在数据预处理阶段, 本文沿用了 [11] [13] 相似的方法, 对长度为 1 的会话和出现频率低于 5 次的项目进行了过滤, 并对训练集和测试集的会话序列进行了数据增强, 具体操作为提取每个会话序列的前缀子序列。表 1 详细展示了预处理后数据集的统计信息, 其中 ratio.bi 表示具有双向边的会话在数据集中的占比。

Table 1. Datasets statistics

表 1. 数据集的统计结果

数据集	train	test	items	avg.len	ratio.bi
Diginetica	719,470	60,858	43,097	5.12	12.95%
Tmall	351,268	25,898	40,728	6.69	27.90%
Last.fm	2,837,330	672,833	38,615	11.78	20.14%

4.2. 评估机制

本文选择 P@K 和 MRR@K 作为评估指标来评估推荐系统。

P@K (Precision calculated over top-K items)。P@K 主要用于所有目标物品, 它计算系统的前 K 个物

品中的被正确推荐的物品比例。

$$P@K = \frac{1}{S_i} \sum_{s \in S_i} \frac{n_{s,hit}}{K} \quad (6)$$

其中, $|S_i|$ 表示测试集 S_i 中测试序列的数量, $n_{s,hit}$ 表示为序列 s 推荐的前 K 个物品集中被正确预测的物品数量。

MRR@K (Mean Reciprocal Rank calculated over top-K items)。MRR@K 是指对于一个推荐系统, 给定一个用户交互序列 s 和推荐列表, 计算在推荐列表排名前 K 个物品中, 第一个与用户序列 s 实际交互的物品相一致的被推荐物品位置 $rank_s$ 的倒数平均值。

$$MRR@K = \frac{1}{|S_i|} \sum_{s=1}^{|S_i|} \frac{1}{rank_s} \quad (7)$$

在实验中, 为了便于与基线模型进行比较, 本文将 K 设为 20, 即用 $P@20$ 和 $MRR@20$ 对模型进行评价。

4.3. 基线模型

在实验中, 选取了 10 个会话推荐模型作为基线模型, 分别是基于传统的方法(IKNN、FPMC)、基于深度学习的方法(GRU4Rec、NARM、STAMP)基于图神经网络的方法(SR-GNN、GCE-GNN)以及基于超图神经网络的方法(DHCN、GC-HGNN、AC-HCN), 具体描述如下:

IKNN [2]: 基于当前会话与其他会话的相似度推荐项目。

FPMC [4]: 结合了矩阵分解和一阶马尔科夫链来捕获用户偏好。通过遵循以前的工作, 本文在计算推荐分数时也忽略了用户的潜在表示。

GRU4Rec [8]: 通过门控神经网络将会话序列建模为最终的会话表示。

NARM [9]: 使用门控神经网络和注意力机制学习会话序列的主要目的和序列特征。

STAMP [10]: 用迭代多层注意力层来取代先前工作中的 RNN 编码器, 完全依靠当前会话中的最后一项的自注意力来捕获用户的短期兴趣。

SR-GNN [11]: 使用门控图神经网络来获得项目的编码表示, 后续和 STAMP 类似, 通过自注意力机制最后一个项目来计算会话级的嵌入做推荐。

GCE-GNN [25]: 使用双层图模型从局部和全局上下文中捕获项目过渡关系, 并考虑反向位置信息, 最终生成会话表示。

DHCN [14]: 使用基于超图的神经网络和基于线图的神经网络, 来捕获高阶非成对的关系, 并将自监督学习融入到网络建模中, 作为改进推荐任务的辅助性方法来增强超图建模。

GC-HGNN [20]: 使用基于超图的神经网络来捕获高阶非成对的关系, 使用 GAT 来捕获局部信息, 通过和池化融合来增强推荐性能。

AC-HCN [21]: 提出自适应嵌入式的超图卷积网络来学习上下文信息, 同时使用 α -entmax 函数消除无关项目的影响, 最后将学习到的项目表征与时间间隔和反位置嵌入相结合生成增强项目表示。

4.4. 参数设置

在实验中, embedding size 设置为 128, mini-batch 的大小设置为 512。所有的参数使用均值为 0 且标准差为 0 的高斯分布进行初始化。学习率设置为 0.001, 并使用 Adam 优化器。实验中将控制自监督学习任务的超参数 β 进行网格探索, 值为{0.1, 0.2, 0.3, 0.4, 0.5}, 基线模型均使用相同的数据集和评估设置, 实验过程中通过微调超参数来获得它们的最佳结果。

4.5. 实验结果与分析

通过考察模型在上述四个问题上的推荐表现, 来验证本文提出的 MS-HGNN 模型的性能。下面给出实验结果及其分析。

4.5.1. 总体比较(RQ1)

表 2 给出了模型在三个数据集上的实验结果, 并将每列的最好结果加粗表示。由于在推荐的过程中存在随机操作, 表格中的数值均为多次实验后的平均值结果。表中数据表明, MS-HGNN 的整体性能最好。

Table 2. Performance comparison between MS-HGNN and baseline models

表 2. MS-HGNN 与基线模型性能比较

Model	Diginetica		Tmall		Last.fm	
	P@20	MRR@20	P@20	MRR@20	P@20	MRR@20
IKNN	35.75	11.75	9.15	3.31	14.90	4.04
FPMC	22.14	6.66	16.06	7.32	12.86	3.78
GRU4Rec	30.79	8.22	10.93	5.89	17.90	5.39
NARM	48.32	16.00	23.30	10.70	22.40	7.42
STAMP	46.62	15.13	26.47	13.36	22.01	7.98
SR-GNN	51.26	17.78	27.57	13.72	22.33	8.27
GCE-GNN	54.22	<u>19.04</u>	33.42	15.42	22.49	<u>8.45</u>
DHCN	53.66	18.51	31.42	15.05	22.10	7.44
GC-HGNN	<u>54.10</u>	18.64	36.83	17.37	<u>22.95</u>	8.31
AC-HCN	51.73	18.71	<u>37.42</u>	<u>17.51</u>	22.91	8.22
MS-HGNN	55.04	19.20	39.98	18.83	24.50	8.78
%Gain	1.74%	0.95%	6.85%	7.54%	6.76%	3.91%

基于深度学习的模型(GRU4Re, NARM, STAMP)显著优于传统模型(IKNN, FPMC), 这表明会话序列中的物品顺序对于会话推荐的重要性。会话推荐中用户偏好是动态变化, NARM 和 STAMP 采用了循环单元对用户行为进行编码, 同时引入注意力机制使得模型性能得到大幅度提升, 这表明在对会话建模过程中对不同项目分配不同注意力权重是十分有效的。

基于图神经网络的模型(SR-GNN、GCE-GNN)对比基于深度学习的模型展示了更好的推荐性能。SR-GNN 通过将每个会话建模为子图并使用门控图神经网络学习嵌入表示, GCE-GNN 则使用双层图模型捕获全局级和局部级信息, 探索了除当前会话以外的物品转换信息。它们将会话序列建模成图结构, 并使用 GNN 来捕获项目之间的成对转换关系, 这也证明了在推荐时考虑项目之间依赖关系的重要性。

基于超图神经网络的模型(DHCN、GC-HGNN、AC-HCN)相较于一般的图神经网络又取得了进一步提升。DHCN 使用超图来捕获高阶复杂关系信息, GC-HGNN 使用超图捕获高阶复杂关系的同时融入了人 GAT 来捕获局部级关系, AC-HCN 则在超图捕获高阶复杂信息后加入了 α -entmax 函数来消除无关项目的

影响。这说明捕获会话序列的高阶复杂关系可以提高模型的推荐性能。

从表2可以看出, 与上述模型相比, 本文提出的MS-HGNN模型在三个数据集上的P@20和MRR@20的性能最佳。原因可能在于MS-HGNN模型从三个方面进行了改进: 一是在捕获高阶复杂信息后加入了位置信息编码, 减轻了超边同构问题。二是捕获了会话内顺序及共现依赖关系, 并通过门控机制融合这两种关系。三是MS-HGNN模型集成了自监督学习, 最大化全局级会话信息和局部级会话信息之间的互信息。

4.5.2. 消融实验(RQ2)

消融实验用来验证模型主要模块对推荐性能的贡献。通过移除 MS-HGNN 不同模块形成的模型变体, 考察它们在三个数据集上推荐性能来评价各模块对结果的影响:

MS-HGNN-H: 从 MS-HGNN 中将超图通道中位置信息模块去掉, 只使用超图来捕获全局级会话表示。

MS-HGNN-CO: 从 MS-HGNN 中将共现依赖关系提取层去掉, 用顺序依赖关系提取层来获取局部级会话表示。

MS-HGNN-SE: 从 MS-HGNN 中将顺序依赖关系提取层去掉, 用共现依赖关系提取层来获取局部级会话表示。

MS-HGNN-S: 去除了自监督学习模型, 只使用交叉熵损失函数模型。

表3给出了实验结果, 它对比了五个模块在 Diginetica、Tmall 和 Last.fm 三个数据集上的性能, 评估参数分别为 P@20 和 MRR@20。

Table 3. The impact of each module of the model on the recommendation performance

表 3. 模型各模块对推荐性能的影响

Dataset	Tmall		Diginetica		Last.fm	
	P@20	MRR@20	P@20	MRR@20	P@20	MRR@20
MS-HGNN-H	39.16	18.25	54.39	19.12	23.78	8.09
MS-HGNN-CO	38.63	17.56	54.47	19.14	23.85	8.14
MS-HGNN-SE	39.55	18.52	54.73	19.16	23.85	8.01
MS-HGNN-S	39.46	18.23	54.81	19.18	23.57	8.77
MS-HGNN	39.98	18.83	55.04	19.20	24.50	8.78

从整体上来看, 本文提出的 MS-HGNN 在 P@20 和 MRR@20 上性能最佳。在 Tmall 数据集上, 由于该数据集的 ratio.bi 偏高, MS-HGNN-SE 比 MS-HGNN-CO 性能更佳, 对于 Last.fm 数据集, 该数据集 avg.len 较长, 根据依赖于会话内顺序关系, 所以 MS-HGNN-SE 的性能反而下降更多。此外 MS-HGNN-H 和 MS-HGNN-S 模型在整体上低于 MS-HGNN, 并且 MS-HGNN-H 降低较多, 说明减轻超边同构及引入自监督学习辅助推荐任务能够在一定程度上提升推荐性能。

4.5.3. 融合函数影响(RQ3)

在推荐系统领域中, 融合函数对推荐性能也有着一定影响。实验考察了采用 4 种不同的融合函数在融合信息时聚合信息对推荐效果的影响, 这 4 种函数如下:

- 1) 求和函数(sum pooling), 即将两个表示按相同维度相加;
- 2) 平均池化函数(Average pooling), 即对两个表示的每个维度上求均值;

- 3) 拼接函数(Concat), 即直接将两个表示拼接到一起;
- 4) 融合门控函数(Fusion Gate), 相当于在两个表示之间进行线性插值:

$$f = \sigma(W_f^1 I_v^{local} + W_f^2 I_v^{global} + b_f)$$

$$I = f I_v^{local} + (1 - f) I_v^{global}$$

表 4 详细展示了实验结果, 从中观察到一些明显的趋势。首先, 拼接函数在三个数据集上的性能表现均不尽如人意。这可能是由于不同用户行为之间固有的多样性导致的。拼接函数在处理不同表示时倾向于平滑过渡, 难以精准捕捉共现依赖关系和顺序依赖关系之间的差异, 从而影响了其性能。相比之下, 求和函数和平均池化函数在实验过程中展现出了良好的性能, 并且它们的计算效率也相对较高。这两种函数在处理用户行为数据时, 能够有效地整合信息, 同时保持较高的运算速度。特别值得一提的是, 融合门控函数在三个数据集上都达到了最优性能。这充分证明了在两个表示之间进行线性插值处理能够更有效地融合会话内部的多种依赖关系。这种融合方式不仅考虑到了不同用户行为之间的差异性, 还能够在一定程度上保留原始数据中的关键信息, 从而提升了推荐系统的准确性和效率。

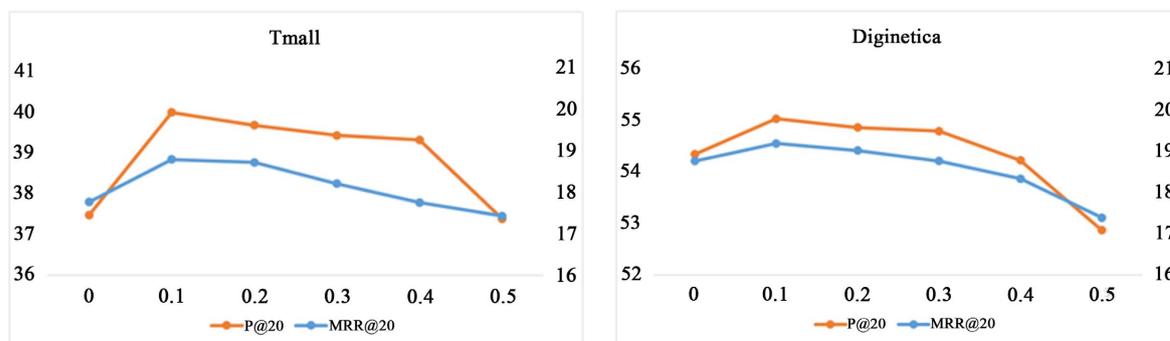
Table 4. Effects of different fusion functions

表 4. 不同融合函数的影响

Dataset	Tmall		Diginetica		Last.fm	
	P@20	MRR@20	P@20	MRR@20	P@20	MRR@20
Function						
Concat	38.92	18.11	54.02	18.56	23.58	7.98
Sum	39.66	18.34	54.27	18.99	24.25	8.64
Average	39.75	18.65	54.75	19.13	24.35	8.57
FG	39.98	18.83	55.04	19.20	24.50	8.78

4.5.4. 参数 β 对推荐性能的影响(RQ4)

在本文的 2.4 小节中, 提出了一个总损失函数, 其中引入了超参数 β 来控制对比自监督学习任务在联合优化训练中的比例。为了研究 β 的大小对推荐效果的影响, 本文将 β 的数值控制在 {0.1, 0.2, 0.3, 0.4, 0.5} 范围内, 来考察模型推荐性能的表现。由图 6 可见, 对于 Diginetica、Tmall 和 Last.fm 这三个数据集, 当 β 取值在 0.1 附近时, P@20 和 MRR@20 均接近最优性能。总体来说, 随着 β 的增大, 推荐性能呈下降趋势。这表明了适当的 β 值可以平衡会话推荐任务和对学习辅助任务之间的关系, 从而获得较理想的推荐性能。



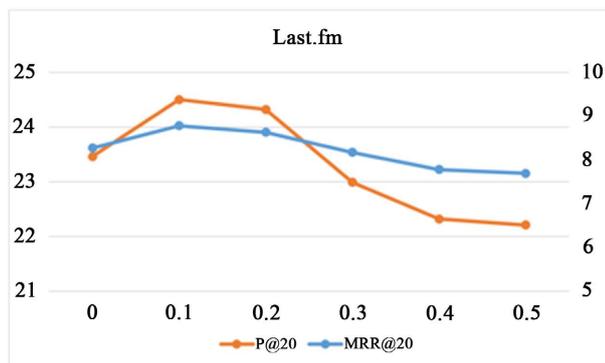


Figure 6. Effect of parameter β on recommended performance
图 6. 参数 β 对推荐性能的影响

5. 结论

现有基于超图的推荐方法建模会话序列时存在局限性, 导致推荐性能次优。本文提出 MS-HGNN 模型, 利用超图神经网络融入位置信息, 缓解超边同构问题, 同时考虑会话内顺序及共现依赖, 更精准捕捉用户意图。通过自监督学习融合, 缓解数据稀疏性。实验显示该模型在三个数据集上性能优越, 未来将进一步研究会话中的其他依赖关系以提升模型性能。

参考文献

- [1] Yang, C.C., Chen, H. and Hong, K. (2003) Visualization of Large Category Map for Internet Browsing. *Decision Support Systems*, **35**, 89-102. [https://doi.org/10.1016/s0167-9236\(02\)00101-x](https://doi.org/10.1016/s0167-9236(02)00101-x)
- [2] Sarwar, B., Karypis, G., Konstan, J. and Riedl, J. (2001) Item-Based Collaborative Filtering Recommendation Algorithms. *Proceedings of the 10th International Conference on World Wide Web*, Hong Kong, 1-5 May 2001, 285-295. <https://doi.org/10.1145/371920.372071>
- [3] Koren, Y., Bell, R. and Volinsky, C. (2009) Matrix Factorization Techniques for Recommender Systems. *Computer*, **42**, 30-37. <https://doi.org/10.1109/mc.2009.263>
- [4] Rendle, S., Freudenthaler, C. and Schmidt-Thieme, L. (2010) Factorizing Personalized Markov Chains for Next-Basket Recommendation. *Proceedings of the 19th International Conference on World Wide Web*, Raleigh, 26-30 April 2010, 811-820. <https://doi.org/10.1145/1772690.1772773>
- [5] Cho, K., van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., et al. (2014) Learning Phrase Representations Using RNN Encoder-Decoder for Statistical Machine Translation. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Doha, October 2014, 1724-1734. <https://doi.org/10.3115/v1/d14-1179>
- [6] Scarselli, F., Gori, M., Tsoi, A.C., Hagenbuchner, M. and Monfardini, G. (2009) The Graph Neural Network Model. *IEEE Transactions on Neural Networks*, **20**, 61-80. <https://doi.org/10.1109/tnn.2008.2005605>
- [7] Yu, B., Zhang, Y., Xie, Y., Zhang, C. and Pan, K. (2021) Influence-Aware Graph Neural Networks. *Applied Soft Computing*, **104**, Article ID: 107169. <https://doi.org/10.1016/j.asoc.2021.107169>
- [8] Hidasi, B., Karatzoglou, A., Baltrunas, L., et al. (2016) Session-Based Recommendations with Recurrent Neural Networks.
- [9] Li, J., Ren, P., Chen, Z., Ren, Z., Lian, T. and Ma, J. (2017) Neural Attentive Session-Based Recommendation. *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, Singapore, 6-10 November 2017, 1419-1428. <https://doi.org/10.1145/3132847.3132926>
- [10] Liu, Q., Zeng, Y., Mokhosi, R. and Zhang, H. (2018) STAMP: Short-Term Attention Memory Priority Model for Session-Based Recommendation. *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, London, 19-23 August 2018, 1831-1839. <https://doi.org/10.1145/3219819.3219950>
- [11] Wu, S., Tang, Y., Zhu, Y., Wang, L., Xie, X. and Tan, T. (2019) Session-Based Recommendation with Graph Neural Networks. *Proceedings of the AAAI Conference on Artificial Intelligence*, **33**, 346-353. <https://doi.org/10.1609/aaai.v33i01.3301346>

-
- [12] Qiu, R., Li, J., Huang, Z. and Yin, H. (2019) Rethinking the Item Order in Session-Based Recommendation with Graph Neural Networks. *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*, Beijing, 3-7 November 2019, 579-588. <https://doi.org/10.1145/3357384.3358010>
- [13] Xu, C., Zhao, P., Liu, Y., Sheng, V.S., Xu, J., Zhuang, F., et al. (2019) Graph Contextualized Self-Attention Network for Session-Based Recommendation. *Proceedings of the 28th International Joint Conference on Artificial Intelligence*, Macao, 10-16 August 2019, 3940-3946. <https://doi.org/10.24963/ijcai.2019/547>
- [14] Xia, X., Yin, H., Yu, J., Wang, Q., Cui, L. and Zhang, X. (2021) Self-Supervised Hypergraph Convolutional Networks for Session-Based Recommendation. *Proceedings of the AAAI Conference on Artificial Intelligence*, **35**, 4503-4511. <https://doi.org/10.1609/aaai.v35i5.16578>
- [15] Rao, Y., Mu, T., Zeng, S., Xue, J. and Liu, J. (2023) Multi-Session Aware Hypergraph Neural Network for Session-Based Recommendation. *Multimedia Tools and Applications*, **83**, 12757-12774. <https://doi.org/10.1007/s11042-023-15894-w>
- [16] Song, J., Shen, H., Ou, Z., Zhang, J., Xiao, T. and Liang, S. (2019) ISLF: Interest Shift and Latent Factors Combination Model for Session-Based Recommendation. *Proceedings of the 28th International Joint Conference on Artificial Intelligence*, Macao, 10-16 August 2019, 5765-5771. <https://doi.org/10.24963/ijcai.2019/799>
- [17] Pan, Z., Cai, F., Chen, W., Chen, H. and de Rijke, M. (2020) Star Graph Neural Networks for Session-Based Recommendation. *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*, 19-23 October 2020, 1195-1204. <https://doi.org/10.1145/3340531.3412014>
- [18] Wang, J., Ding, K., Zhu, Z., et al. (2021) Session-Based Recommendation with Hypergraph Attention Networks. *Proceedings of the 2021 SIAM International Conference on Data Mining (SDM)*, 29 April-1 May 2021, 82-90.
- [19] Li, Y., Gao, C., Luo, H., Jin, D. and Li, Y. (2022) Enhancing Hypergraph Neural Networks with Intent Disentanglement for Session-Based Recommendation. *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*, Madrid, 11-15 July 2022, 1997-2002. <https://doi.org/10.1145/3477495.3531794>
- [20] Peng, D. and Zhang, S. (2022) GC-HGNN: A Global-Context Supported Hypergraph Neural Network for Enhancing Session-Based Recommendation. *Electronic Commerce Research and Applications*, **52**, Article ID: 101129. <https://doi.org/10.1016/j.elerap.2022.101129>
- [21] Zhao, C., Cao, H., Lv, P., Chu, Y., Wang, F. and Liao, T. (2023) Adaptive Context-Embedded Hypergraph Convolutional Network for Session-Based Recommendation. *Information Technology and Control*, **52**, 111-127. <https://doi.org/10.5755/j01.itc.52.1.32138>
- [22] Liu, Z., Chen, Y., Li, J., et al. (2021) Contrastive Self-Supervised Sequential Recommendation with Robust Augmentation.
- [23] Vaswani, A., Shazeer, N., Parmar, N., et al. (2017) Attention Is All You Need. *Proceedings of the 31st International Conference on Neural Information Processing Systems*, Long Beach, 4-9 December 2017, 6000-6010.
- [24] Kang, W. and McAuley, J. (2018) Self-Attentive Sequential Recommendation. 2018 *IEEE International Conference on Data Mining (ICDM)*, Singapore, 17-20 November 2018, 197-206. <https://doi.org/10.1109/icdm.2018.00035>
- [25] Wang, Z., Wei, W., Cong, G., Li, X., Mao, X. and Qiu, M. (2020) Global Context Enhanced Graph Neural Networks for Session-Based Recommendation. *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, 25-30 July 2020, 169-178. <https://doi.org/10.1145/3397271.3401142>