

结合图对比学习的多图神经网络会话推荐方法

卢敏, 原子婷

引用本文

卢敏, 原子婷. 结合图对比学习的多图神经网络会话推荐方法[J]. 计算机科学, 2024, 51(5): 54-61.

LU Min, YUAN Ziting. [Graph Contrast Learning Based Multi-graph Neural Network for Session-based RecommendationMethod](#) [J]. Computer Science, 2024, 51(5): 54-61.

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)

[面向前提选择的新型图约简表示与图神经网络模型](#)

New Graph Reduction Representation and Graph Neural Network Model for Premise Selection

计算机科学, 2024, 51(5): 193-199. <https://doi.org/10.11896/jsjcx.230300193>

[基于残差网络融合多关系评论特征的虚假评论检测](#)

Fake Review Detection Based on Residual Networks Fusion of Multi-relationship Review Features

计算机科学, 2024, 51(4): 314-323. <https://doi.org/10.11896/jsjcx.230200020>

[图神经网络节点分类任务基准测试及分析](#)

Benchmarking and Analysis for Graph Neural Network Node Classification Task

计算机科学, 2024, 51(4): 132-150. <https://doi.org/10.11896/jsjcx.230200084>

[基于图卷积神经网络的节点分类方法研究综述](#)

Review of Node Classification Methods Based on Graph Convolutional Neural Networks

计算机科学, 2024, 51(4): 95-105. <https://doi.org/10.11896/jsjcx.230600071>

[基于依赖类型剪枝的双特征自适应融合网络用于方面级情感分析](#)

Dual Feature Adaptive Fusion Network Based on Dependency Type Pruning for Aspect-basedSentiment Analysis

计算机科学, 2024, 51(3): 205-213. <https://doi.org/10.11896/jsjcx.230100035>

结合图对比学习的多图神经网络会话推荐方法

卢 敏 原子婷

1 中国民航大学计算机科学与技术学院 天津 300300
2 智慧机场理论与系统民航局重点实验室 天津 300300
(mlu@cauc.edu.cn)

摘 要 会话推荐根据匿名用户短期内的交互数据预测下一个交互物品。针对会话中物品少、物品长尾分布等特性,现有基于图对比学习的会话推荐模型提出对会话内物品采用随机裁剪、扰动等方式构造正负样本。然而,上述随机退出策略进一步缩减较短会话中的可用物品,使得会话更加稀疏,引起会话兴趣学习偏差。为此,提出了结合图对比学习的多图神经网络会话推荐方法。其核心思想是:在物品局部图、物品全局图等上提取融入物品局部和全局的高阶邻域物品表示,并生成物品级的会话表示,然后设计会话-会话图并学习会话级的会话表示,最后递归利用不同级别会话兴趣生成正负样本对,通过对比学习机制增强会话兴趣区分性。与退出策略相比,所提模型保留了完整的会话信息,实现了真正的数据扩充。在两个基准数据集上进行了大量实验,结果表明,该算法的推荐性能远优于主流基线方法。

关键词: 会话推荐;图对比学习;图神经网络;会话兴趣;正负样本
中图分类号 TP391

Graph Contrast Learning Based Multi-graph Neural Network for Session-based Recommendation Method

LU Min and YUAN Ziting

1 College of Computer Science and Technology, Civil Aviation University of China, Tianjin 300300, China
2 Key Laboratory of Smart Airport Theory and System, Civil Aviation Administration of China, Tianjin 300300, China

Abstract Session recommendation predicts the next interaction item based on anonymous user interaction data over a short period of time. Sessions have characteristics such as few items and long-tail distribution of items. Existing session recommendation models based on graph contrast learning construct positive and negative samples by randomly cropping and perturbing the items within a session, etc. However, the above random exit strategy further shrinks the available items in shorter sessions. This makes the sessions more sparse and causes session interest learning bias. To this end, a graph contrast learning based multi-graph neural network for session-based recommendation method is proposed. The core idea is as follows: the model extracts item representations on item local graphs as well as item global graphs, incorporating both local and global higher-order neighborhood information of the items. Based on this, the model generates item-level session representations. Then, Session-level session representations are learned on the session-session graph. Finally, the model recursively generates positive and negative sample pairs using different levels of conversational interest. And the discriminative nature of the session interests is enhanced by the contrast learning mechanism. Compared with the exit strategy, the proposed model preserves the complete session information and achieves true data expansion. Extensive experiments on two benchmark datasets show that the recommendation performance of the algorithm is much better than that of the mainstream baseline approach.

Keywords Session recommendation, Graph contrast learning, Graph neural networks, Session interest, Positive and negative samples

1 引言

基于会话的推荐旨在从给定匿名用户的短期物品访问序列中挖掘动态变化的兴趣偏好,以推荐用户下一个可能点击

的物品^[1]。其由于在电子商务、音乐推荐等领域具有广泛应用,因此受到越来越多的关注^[2]。目前,该领域的大多数研究工作将会话视为有序序列,其中基于循环神经网络(Recurrent Neural Network, RNN)和基于图神经网络(Graph Neu-

到稿日期:2023-03-11 返修日期:2023-11-13

基金项目:中央高校基本科研业务费专项基金(3122021090)

This work was supported by the Fundamental Research Funds for the Central Universities of Ministry of Education of China(3122021090).

通信作者:原子婷(3138194302@qq.com)

ral Network, GNN)的方法表现出了良好性能。

RNN 将会话数据建模为单向序列,按照时间顺序提取会话信息,但是忽略了物品的高阶邻域以及物品间的复杂传递关系,节点转移的潜在关联信息表达不全面^[3]。基于 GNN 的方法做了进一步改进,将会话数据建模为图结构,使用图卷积学习融入高阶邻域的物品表示,并生成相应的会话嵌入向量^[4]。尽管模型性能得到了提升,但是仍然受到数据稀疏问题的影响。由于长期的用户行为数据不可访问,只能利用短期会话内的物品交互关系获得信息。然而,这些数据数量有限,无法准确捕捉用户偏好,导致推荐性能不佳。

图对比学习算法可以通过挖掘自身数据信息为模型提供额外增益,以此缓解数据稀疏问题^[5]。最新研究进展将图对比学习应用于推荐模型,大多通过随机删除、扰动等方式获得正负样本,增强数据表示^[6]。但是会话包含的平均物品数量本身就很少,随机退出策略进一步缩短会话长度,导致会话兴趣捕捉不全面、不准确。

基于此,本文提出了结合图对比学习的多图神经网络会话推荐方法(Graph Contrast Learning Based Multi-graph Neural Network for Session-based Recommendation, GCLMNN),在基于多图神经网络的推荐任务上引入一种新的图对比学习方法,充分利用原始会话数据并挖掘更有价值的信息,以此缓解数据稀疏问题,提升模型的推荐性能。

本文主要工作如下:

- 1)构造物品局部图和物品全局图相结合的多图表示网络,结合软注意力机制和反向位置嵌入提取融入局部和全局高阶邻域的物品表示以及会话表示。
- 2)设计会话-会话图提取跨会话信息,从完整会话角度出发学习会话级会话表示。
- 3)递归利用不同级别的会话信息构建正负样本对,通过图对比学习机制不断优化网络模型,提取更有价值的会话数据,缓解数据稀疏问题。
- 4)在真实数据集上进行大量实验,与最先进的基线模型相比,GCLMNN 的性能得到显著提高,证明了所提方法的有效性。

2 相关工作

2.1 会话推荐

会话推荐的早期研究中,矩阵分解和马尔可夫链是常见的两种方法。矩阵分解的基本思想是将物品与用户评分矩阵分解为两个低秩矩阵,分别包含物品和用户的潜在表示^[7]。马尔可夫链将推荐生成视为一个序列优化问题,通过捕捉相邻物品的顺序关系进行下一个物品推荐^[8]。传统方法忽略了物品间复杂的转换关系,难以捕捉会话上下文的信息传递。

随着深度学习的蓬勃发展,基于循环神经网络的会话推荐方法逐渐成为研究热点。文献[3]首次将 RNN 应用于会话推荐,提出了基于门控循环单元的 GRU4Rec(Gated Recurrent Unit for Recommendation)模型。文献[9]提出的 NARM(Neural Attentive Recommendation Machine)模型将 RNN 与注意力机制相结合,以捕获用户的主要意图。文献[10]在 RNN 网络上融入协作邻域信息预测会话兴趣。基于

RNN 的方法侧重于序列建模,根据给定序列的时间顺序预测用户偏好,因此不能对非相邻物品转换等复杂的转换模式进行建模。

近年来,图神经网络被引入到会话推荐上,并表现出了良好性能。文献[4]提出的 SR-GNN(Session-based Recommendation with Graph Neural Network)模型将会话序列建模为图结构,使用门控图神经网络聚合物品信息,充分利用了物品间的转换关系。文献[11]在保留 GNN 对会话数据建模的同时引入自注意力机制,结合全局依赖和局部兴趣偏好进行预测。文献[12]提出 FGNN(Full Graph Neural Network)模型,利用加权注意网络学习会话序列中物品间的隐藏信息。文献[13]利用 GNN 构建搜索空间,以自动搜索最佳适配的会话推荐模型,获得最优推荐结果。文献[14]从用户需求出发,设计需求感知图神经网络,所提取的潜在会话需求为下一个物品推荐提供依据。相比 RNN 模型,基于 GNN 的会话推荐方法能够有效捕捉多对多的物品关系。然而,多层的图卷积网络会纳入与物品没有共同兴趣的高阶邻域,使得会话兴趣无法区分,从而降低性能。同时,由于会话交互数据有限,模型性能依然受数据稀疏问题的影响。

2.2 图对比学习

图对比学习是自监督学习的代表性技术之一,模型在大量无标记数据中挖掘监督信息,通过学习数据间的相似或不同来获取数据集的一般特征^[5]。其早期在计算机视觉和自然语言处理领域得到广泛应用,通过图像旋转、剪切、放大和句子裁剪等操作增强原始数据,获得更有区分性的表示^[15-16]。受其启发,图对比学习算法开始被引入推荐任务中,文献[17]将序列建模为图结构,利用互信息最大化原理学习物品、序列等内在关系,并通过模型预训练增强序列数据表示。文献[18]通过裁剪、遮蔽、重排序等方式实现图数据增强,利用图对比学习策略提取用户行为序列中的自监督信号,提升用户兴趣向量表示。文献[19]提出了 S²-DHCN 模型,将会话数据建模为超图和线图,引入对比学习策略增强会话表示,提取用户兴趣。文献[20]通过超图增强的跨视图对比学习框架捕获全局和局部协作关系,提升推荐系统的表示质量。

虽然上述方法在推荐任务上取得了不错的效果,但是大多采用随机退出策略实现数据增强,造成会话序列中物品数量进一步减少的现象,引起会话兴趣学习偏差。因此,文本尝试在保留完整会话数据的条件下引入一种新的图对比学习策略,结合多图神经网络有效提取用户兴趣,缓解数据稀疏问题对模型的影响。

3 主要算法

本文提出的 GCLMNN 模型如图 1 所示。其核心模块是一个基于多图神经网络的会话编码模块,构建会话推荐任务。在此基础上引入了图对比学习机制作为辅助任务,优化会话兴趣表示,提升准确率。

3.1 多图会话编码

本节定义了两种图结构,物品局部图和物品全局图,分别捕获单一会话内相邻物品的一阶依赖关系,以及所有会话上物品的高阶邻域关联。两种图结构均都采用图神经网络编码

物品间的复杂信息,生成不同层次的物品表示向量。最后引入

软注意力机制和反向位置信息生成相应的会话兴趣表示。

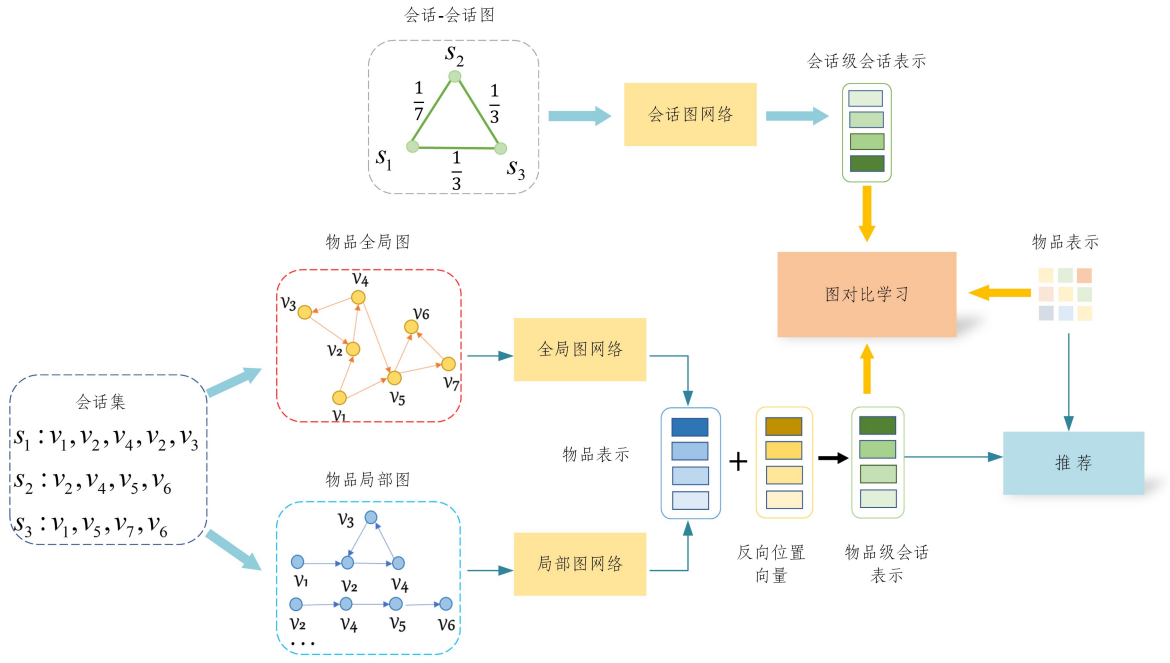


图1 GCLMNN模型的整体框架

Fig.1 Overall framework of GCLMNN model

3.1.1 物品局部表示生成

局部图是对任一会话 $s = [v_1^s, v_2^s, v_3^s, \dots, v_{n_s}^s]$ 构建一个有向图 $G_l = (V_l, E_l)$, 其中 v_i^s 表示匿名用户在会话 s 中第 i 个交互的物品, n_s 为会话 s 的物品长度。节点集 V_l 表示当前会话 s 中包含的物品集合, 边集 E_l 是会话 s 中所有物品转移关系的表示, 其中 $(v_{i-1}^s, v_i^s) \in E_l$ 。为每条边赋予归一化的权值, 计算方法是边出现的次数除以边的度值。

通过门控神经网络 (Gated Graph Neural Network, GGNN) 获取物品局部图上的节点表示。设所有物品的嵌入向量为 $\mathbf{X}^0$ 。对于图 G_l 上的节点 V_l , 其在初始时刻被映射为 $\mathbf{X}_l^0$ 。经 GGNN 更新后节点表示为 $\mathbf{X}_l^1$, 更新函数如下:

$$\mathbf{Z}_l = \mathbf{A}_l (\mathbf{H}_l \mathbf{X}_l^0 + \mathbf{b}_l^0) \quad (1)$$

$$\mathbf{X}_l^1 = \text{GGNN}(\mathbf{W}_l \mathbf{Z}_l + \mathbf{b}_l^1) \quad (2)$$

其中, $\mathbf{A}_l$ 是图 G_l 的关系矩阵, 由入度邻接矩阵和出度邻接矩阵拼接所得。式(1)首先对节点原始嵌入向量进行线性变换, $\mathbf{H}_l$ 和 $\mathbf{b}_l^0$ 为线性变换的参数。之后, 再与矩阵 $\mathbf{A}_l$ 相乘, 得到融入邻域信息的隐向量 $\mathbf{Z}_l$ 。式(2)将 $\mathbf{Z}_l$ 进行线性变换后作为节点初始表示输入 GGNN 中, $\mathbf{W}_l$ 和 $\mathbf{b}_l^1$ 为线性变换过程中需要学习的参数。更新局部图中的所有节点直到收敛后, 可以得到局部图上的物品嵌入 $\mathbf{X}_l^1$ 。其中, 会话 s 中的每一个物品局部表示记为 $\mathbf{x}_l^i \in \mathbf{X}_l^1$ 。

3.1.2 物品全局表示生成

物品全局图汇总训练集中所有会话上的物品局部图。将全局图定义为 $G_g = (V_g, E_g)$, 节点集 V_g 表示所有会话中出现过的物品节点的集合; E_g 表示边集, 会话中任意相邻的两个物品在全局图中用加权有向边连接, 记为 $(v_{i-1}^s, v_i^s) \in E_g$ 。边的权重由成对物品以 $[v_{i-1}^s, v_i^s]$ 相邻形式出现的次数决定。

通过图卷积神经网络获取全局图上节点的表示。对于图

G_g 上的节点 V_g , 其初始表示为所有物品的嵌入向量 $\mathbf{X}^0$, 经过 L 层图卷积网络后更新为 $\mathbf{X}^L$:

$$\mathbf{X}^{l+1} = \hat{\mathbf{D}}_g^{-1} \hat{\mathbf{A}}_g \mathbf{X}^l \mathbf{W}_g^l \quad (3)$$

$$\mathbf{X}^g = \frac{1}{L+1} \sum_{l=0}^L \mathbf{X}^l \quad (4)$$

其中, $\hat{\mathbf{A}}_g = \mathbf{A}_g + \mathbf{I}_g$ 和 $\mathbf{I}_g$ 均为单位矩阵, $\hat{\mathbf{D}}_g = \sum_{p=1}^n \hat{\mathbf{A}}_{g,p}$ 和 $\mathbf{A}_g$ 分别为物品全局图的度矩阵和邻接矩阵。 $\mathbf{X}^l$ 表示在全局图网络中第 l 层的物品嵌入, 对应该层的网络参数矩阵为 $\mathbf{W}_g^l$ 。物品初始映射 $\mathbf{X}^0$ 遍历了 L 个图卷积层后, 将每一层嵌入进行平均, 得到全局图上的物品表示向量 $\mathbf{X}^g$ 。其中, 会话 s 中每一个物品的全局表示记为 $\mathbf{x}_g^i \in \mathbf{X}^g$ 。

3.1.3 物品级会话表示生成

通过合并物品局部图和全局图上的物品嵌入得到会话 s 内的物品表示向量:

$$\mathbf{x}_s = \mathbf{x}_l^i + \mathbf{x}_g^i \quad (5)$$

其中, $\mathbf{x}_l^i$ 为会话 s 中物品的局部表示, $\mathbf{x}_g^i$ 则是会话 s 中物品的全局表示。当前会话中包含的物品可以表示为: $s = [\mathbf{x}_1^s, \mathbf{x}_2^s, \mathbf{x}_3^s, \dots, \mathbf{x}_{n_s}^s]$ 。虽然局部图和全局图可以很好地捕捉物品之间的复杂关系, 但是忽略了时间顺序对会话兴趣的影响, 为此引入一个可学习的位置矩阵 $\mathbf{P} = [\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_{n_s}]$, 将学到的物品表示与反向位置嵌入相结合。其中, n_s 为当前会话的长度, $\mathbf{p}_j \in \mathbf{R}^d$ 表示序列中第 j 个位置的嵌入。因此, 会话 $s = [v_1^s, v_2^s, v_3^s, \dots, v_{n_s}^s]$ 中第 t 个物品的表示为:

$$\mathbf{x}_s^t = \tanh(\mathbf{W}_1 [\mathbf{x}_t^s \parallel \mathbf{p}_{n_s-t+1}] + \mathbf{b}_1) \quad (6)$$

式(6)的物理含义是将物品的嵌入向量与位置向量拼接后进行线性变换, 经过 $\tanh$ 归一化处理得到最终的物品表示 $\mathbf{x}_s^t$ 。其中, $\tanh$ 是一个正切函数, $\mathbf{W}_1$ 和 $\mathbf{b}_1$ 为线性变换中的参数, 通过训练数据得到。

会话表示通过平均会话序列包含的物品表示获得。考虑到会话中物品的优先级不同,采用软注意力机制来获取物品级的会话表示:

$$\bar{s} = \frac{1}{n_s} \sum_{t=1}^{n_s} \mathbf{x}_s^t \quad (7)$$

$$\gamma_{s,t} = f^T \sigma(\mathbf{W}_2 \bar{s} + \mathbf{W}_3 \mathbf{x}_s^t + \mathbf{b}_2) \quad (8)$$

$$\mathbf{s}^{\text{item}} = \sum_{t=1}^{n_s} \gamma_{s,t} \mathbf{x}_s^t \quad (9)$$

其中, σ 表示 sigmoid 函数, $f, \mathbf{W}_2, \mathbf{W}_3$ 和 $\mathbf{b}_2$ 为学习物品权重 $\gamma_{s,t}$ 的注意力参数,通过训练数据获得。根据不同物品的重要程度聚合物品嵌入,得到会话嵌入 $\mathbf{s}^{\text{item}}$ 。

物品级会话表示 $\mathbf{s}^{\text{item}}$ 不仅融合了局部和全局的上下文信息,而且考虑了物品位置顺序及其重要程度,可以很好地表示会话特征。

3.2 模型优化和推荐生成

给定会话 s , 经过物品局部图和全局图网络后得到一个物品表示 $\mathbf{x}_s$ 和会话表示 $\mathbf{s}^{\text{item}}$, 再通过内积计算每个候选物品在会话中的得分 z_s :

$$z_s = \mathbf{s}^{\text{item}T} \mathbf{x}_s \quad (10)$$

之后,应用 softmax 函数计算该物品在当前会话中成为下一个推荐物品的概率:

$$\hat{y}_s = \text{softmax}(z_s) \quad (11)$$

最后使用交叉熵损失函数作为学习目标,定义为:

$$L_{(\hat{y}_s)} = - \sum_{i=1}^m y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i) \quad (12)$$

3.3 图对比学习增强会话推荐

本节通过辅助任务将对比学习机制引入推荐模型中,以增强会话数据,进一步优化会话兴趣表示。区别于现有会话推荐中随机裁剪、扰动等生成负样本的方式,本文设计了会话-会话图,完整保留会话信息,并生成会话级的会话表示,以缓解数据稀疏性对模型的影响。最后递归地利用不同级别会话兴趣生成正负样本对,通过图对比学习机制增强会话兴趣区分性。

3.3.1 会话级会话表示生成

区别于物品局部图和全局图,会话-会话图结构从另一个角度出发描述了物品和会话的关系,它包括了跨会话信息并强调会话间的连通性。将会话图定义为 $G_e = (V_e, E_e)$, 其节点集 V_e 表示所有会话的集合;边集 E_e 刻画了两会话间存在共享物品的关系。会话-会话图设定为加权无向边结构,边的权值通过两个会话共享物品数量除以两个会话总物品数量得到。

由于会话-会话图中不涉及任何物品,因此首先通过平均每个会话中的物品嵌入来初始化会话嵌入,记为 s^0 , 经过 L 层图卷积学习会话信息,得到会话表示 s^L :

$$\mathbf{s}^{l+1} = \hat{\mathbf{D}}_e^{-1} \hat{\mathbf{A}}_e \mathbf{s}^l \mathbf{W}_e^l \quad (13)$$

$$\mathbf{s}^{\text{sess}} = \frac{1}{L+1} \sum_{l=0}^L \mathbf{s}^l \quad (14)$$

其中, $\hat{\mathbf{A}}_e = \mathbf{A}_e + \mathbf{I}_e$ 和 $\hat{\mathbf{I}}_e$ 均为单位矩阵, $\mathbf{A}_e, \hat{\mathbf{D}}_e = \sum_{p=1}^n \hat{\mathbf{A}}_{e,p}$ 分别为会话图的邻接矩阵和度矩阵。 $\mathbf{s}^l$ 和 $\mathbf{W}_e^l$ 表示第 l 层的会话嵌入和参数矩阵。式(14)计算每一层学习到的会话嵌入的平均

值,并将其作为会话级的会话兴趣表示 $\mathbf{s}^{\text{sess}}$ 。

3.3.2 正负样本选择

本文从两种不同视角学习到物品级的会话表示和会话级的会话表示,它们从原始数据中继承了不同信息,相互补充。将两组会话兴趣对比学习,以充分利用会话蕴含的物品结构信息,进而增强数据。

由于在一个视角下有价值的样本在另外一个视角下同样发挥着重要作用,因此给定会话级视角下的会话 s , 使用物品级的会话兴趣向量 $\mathbf{s}^{\text{item}}$ 来预测其下一个物品:

$$\text{score}_s = \mathbf{X}_s \mathbf{s}^{\text{item}} \quad (15)$$

$$y_s = \text{softmax}(\text{score}_s) \quad (16)$$

其中, y_s 表示每个物品被推荐给会话 s 的预测概率。

根据计算出的概率值,选择分值最高的前 K 个物品作为正样本:

$$c_{s^+} = \text{top-K}(y_s) \quad (17)$$

而负样本的提取是图对比学习方法中的难点所在,如果选择分值最低的物品作为负样本,并不能带给模型很高的贡献度。因此,本文在分值排名前 10% 的物品中除去前 K 个物品后,再随机选择 K 个物品作为负样本 c_{s^-} , 这些难负样本可以为模型提供最大价值的信息增益。

同理,给定物品级视角下的会话 s , 使用会话级的会话表示 $\mathbf{s}^{\text{sess}}$ 预测其下一个物品。

$$\text{score}_e = \mathbf{X}_e^0 \mathbf{s}^{\text{sess}} \quad (18)$$

$$y_e = \text{softmax}(\text{score}_e) \quad (19)$$

使用同种方式生成物品级视角下的正负样本。每批次利用另一级别的会话表示重新构建正负样本,会获得信息更加丰富的样本表示,网络模型也会不断优化,递归提取到更有价值的信息。

3.3.3 会话兴趣的对比学习

对于会话型推荐,假设最后点击的物品与下一物品最相关,以此构建正负样本对。通过最大化(最小化)最后点击物品与预测正(负)样本的表示之间的一致性,拉近嵌入空间中锚点与正样本的距离,拉远锚点与负样本的距离。最后采用具有标准二元交叉熵损失的 InfoNCE 损失函数作为优化目标。

$$L_s = - \log \frac{\sum_{i \in c_{s^+}} \varphi(\mathbf{x}_s^{\text{last}}, \mathbf{s}^{\text{item}}, \mathbf{x}_s^i)}{\sum_{i \in c_{s^+}} \varphi(\mathbf{x}_s^{\text{last}}, \mathbf{s}^{\text{item}}, \mathbf{x}_s^i) + \sum_{j \in c_{s^-}} \varphi(\mathbf{x}_s^{\text{last}}, \mathbf{s}^{\text{item}}, \mathbf{x}_s^j)} - \log \frac{\sum_{i \in c_{e^+}} \varphi(\mathbf{x}_e^{\text{last}}, \mathbf{s}^{\text{sess}}, \mathbf{s}_e^i)}{\sum_{i \in c_{e^+}} \varphi(\mathbf{x}_e^{\text{last}}, \mathbf{s}^{\text{sess}}, \mathbf{s}_e^i) + \sum_{j \in c_{e^-}} \varphi(\mathbf{x}_e^{\text{last}}, \mathbf{s}^{\text{sess}}, \mathbf{s}_e^j)} \quad (20)$$

其中, $\mathbf{x}_s^{\text{last}}$ 是给定会话的最后点击物品表示, $\varphi(x_1, x_2, x_3) = \exp(f_D(x_1 + x_2, x_2 + x_3)/\tau)$, $f_D(\cdot): \mathbb{R}^d \times \mathbb{R}^d \mapsto \mathbb{R}$ 为鉴别器函数,评判两个向量的相近程度,实验中选择余弦运算; τ 是超参数,设定为 0.2。通过上述对比学习,物品级和会话级的会话兴趣迭代交换信息,相互补充,优化会话兴趣学习。

最后,将推荐任务与对比学习辅助任务统一起来,模型的总损失函数定义为:

$$L = L_{(\hat{y}_s)} + \beta L_s \quad (21)$$

其中, β 为控制对比学习强度的超参数。使用反向传播算法对模型进行训练。

4 实验与分析

4.1 数据集和预处理

本文使用公开的 Tmall 和 Dijinetica 数据集进行实验并对所提模型进行评估。Tmall 数据集来自 IJCAI-15 竞赛,包括天猫网站的用户购物记录。Dijinetica 数据集通过 CIKM Cup 2016 获取,包括搜索引擎的用户会话记录。本文只使用其交易数据。

在模型训练之前,首先对两个数据集进行预处理,过滤掉长度为 1 的会话,太短的会话记录不具备参考价值,同时还删除了出现次数小于 5 的物品。将最后一周的数据设置为测试集,将其余数据设置为训练集。此外,对每个会话 $s=[v_1^s, v_2^s, v_3^s, \dots, v_{n_s}^s]$ 进行分割,生成序列和相应的标签 $([v_1^s], v_2^s)$, $([v_1^s, v_2^s], v_3^s), \dots, ([v_1^s, v_2^s, v_3^s, \dots, v_{n-1}^s], v_n^s)$ 。其中,会话中最后一个点击的物品是该序列的标签。预处理后的数据统计结果如表 1 所列。

表 1 实验数据集统计信息
Table 1 Statistics of experimental datasets

数据集	Tmall	Diginetica
训练集会话数	351268	719470
测试集会话数	25898	60858
物品数	40728	43097
平均会话长度	6.69	5.12

4.2 评价指标

本文采用会话推荐中常用的两个评价指标:准确率 $P@K$ (Precision) 和平均倒数排名 $MRR@K$ (Mean Reciprocal Rank, MRR)。 $P@K$ 是对预测准确性的度量,表示预测正确的物品在前 K 个物品中所占的比例,计算式为:

$$P@K=\frac{n_{hit}}{N} \tag{22}$$

其中, N 为测试集中的样本数量, n_{hit} 为推荐列表前 K 个物品中包含正确物品的数量。

$MRR@K$ 表示正确推荐项目的倒数排名的平均值,该值越大,表明目标物品在推荐列表中的位置越靠前。其计算式为:

$$MRR@K=\frac{1}{N}\sum_{i\in S_o}\frac{1}{rank(i)} \tag{23}$$

其中, S_o 为测试集包含的所有目标物品, $rank(i)$ 是目标物品

i 在列表中的排名。本文实验中 K 的取值为 10 和 20。

4.3 对比模型

将所提方法与以下代表性基线模型进行对比。

- 1)POP^[21]:根据训练集中物品的流行度排名推荐。
- 2)Item-KNN^[22]:根据余弦距离计算两个物品间的相似度,推荐下一个可能点击的物品。
- 3)FPMC^[8]:同时利用马尔可夫链和矩阵分解算法进行会话推荐。
- 4)GRU4Rec^[3]:首次在会话推荐中使用 RNN 建模序列,使用门控神经网络(GRU)生成会话表示。
- 5)NARM^[23]:使用 RNN 的同时加入注意力机制,考虑用户的主要偏好。
- 6)STAMP^[24]:采用注意力机制捕获用户的长期偏好,将最后一次点击的物品作为用户短期偏好,结合两种偏好共同进行会话推荐。
- 7)SR-GNN^[4]:首次在会话推荐中使用 GNN 建模序列,结合用户的一般兴趣和当前兴趣做推荐。
- 8)FGNN^[12]:使用图卷积网络捕获会话和物品间的复杂关系,并使用加权注意力机制获得物品表示。
- 9)S²-DHCN^[19]:通过构造超图和线图网络学习物品以及会话嵌入,结合自监督学习加强会话兴趣表示。
- 10)Disen-GNN^[25]:通过独立因素表示每个物品嵌入,并根据因子级注意力推断会话中的用户兴趣。

4.4 参数设置

实验设置隐向量维度固定为 100,训练批次大小固定为 100。所有可学参数均以高斯分布进行初始化,其均值为 0,标准差为 0.1。学习率在实验中设置为 0.001,正则化系数 $L_2=10^{-5}$ 。使用 Adam 算法对模型中的参数进行优化。对于基线模型,由于其原始论文的数据集和评估机制与本文相同,因此本文采用对应论文的最佳实验结果进行比较。

4.5 整体性能比较与分析

表 2 列出了基线模型和本文所提模型 GCLMNN 在两个真实数据集上的实验结果,其中黑体字标记了每列的最佳结果。可以观察到,本文模型在这两个数据集的两个指标($K=10$ 和 20)上都实现了最佳性能,证明了本文方法的有效性。

表 2 基于 Tmall 和 Diginetica 数据集的算法性能对比
Table 2 Performance comparison of algorithms based on Tmall and Diginetica datasets

模型	Tmall				Diginetica			
	$P@10$	$M@10$	$P@20$	$M@20$	$P@10$	$M@10$	$P@20$	$M@20$
POP	1.67	0.88	2.00	0.90	0.76	0.26	1.18	0.28
Item-KNN	6.65	3.11	9.15	3.31	25.07	10.77	35.75	11.57
FPMC	13.10	7.12	16.06	7.32	15.43	6.20	26.53	6.95
GRU4REC	9.47	5.78	10.93	5.89	17.93	7.33	29.45	8.33
NARM	19.17	10.42	23.30	10.70	35.44	15.13	49.70	16.17
STAMP	22.63	13.12	26.47	13.36	33.98	14.26	45.64	14.32
SR-GNN	23.41	13.45	27.57	13.72	36.86	15.52	50.73	17.59
FGNN	20.67	10.07	25.24	10.39	37.72	15.95	50.58	16.84
S ² -DHCN	26.22	14.60	31.42	15.05	40.21	17.59	53.66	18.51
Disen-GNN	26.58	14.82	31.56	15.31	38.13	16.39	53.49	18.99
GCLMNN	30.87	17.53	36.75	18.88	41.65	18.04	54.37	19.02

从表2中的实验结果来看,POP的性能最差,因为它只推荐前 n 个出现最频繁的物品。与POP相比,FPMC利用了矩阵分解以及马尔可夫链,显现出了其有效性。此外,Item-KNN模型也取得了不错的效果,但是它只考虑了物品之间的相似性,并没有考虑会话中物品之间的顺序转换关系。

GRU4Rec,NARM,STAMP等模型的性能相比传统方法有了很大提升,这说明深度学习能够有效捕捉会话信息,准确提取用户兴趣。NARM和STAMP模型的高性能证实了注意力机制对学习会话表征的有效性。同时,会话中物品的时间信息和邻域节点信息包含了用户长短期的兴趣偏好,用户近期的物品交互行为更能代表用户当前兴趣。

通过实验结果不难发现,基于图的模型更优于基于RNN的方法。SR-GNN和FGNN将每个会话序列建模为一个子图,通过图神经网络挖掘和利用物品间多对多的关系,模型性能有了很大提高。Disen-GNN进一步丰富了图网络结构,对当前会话以外的物品转换关系进行建模,取得了不错的效果。此外, S^2 -DHCN模型在图网络基础上引入自监督学习增强会话数据,获得的良好性能表明了自监督学习对于会话推荐任务的有效性。

而本文模型在表2所列的所有基线方法和所有数据集中表现最优,原因如下:1)在基于多图神经网络的会话推荐任务上引入对比学习机制作为辅助任务,增加了会话兴趣区分性,提升了推荐准确率;2)相比现有的对比学习方法,本文没有采用删除、扰动等随机退出策略,而是新增了会话-会话图,从完整会话角度出发,递归利用两个视角的会话表示构建正负样本对,实现了真正的数据扩充,缓解了数据稀疏问题对模型的影响。

4.6 图对比学习策略分析

为了进一步研究图对比学习在GCLMNN模型中的有效性,本文将其与4种变体进行对比。

1)GCLMNN-G:移除会话-会话图以及图对比学习模块,只利用多图神经网络进行会话推荐。

2)GCLMNN-S:在物品局部图、物品全局图的基础上保留会话-会话图,将两个级别的会话表示相加做推荐。

3)GCLMNN-R:正样本是同一会话的不同级别表示,负样本则通过对物品级或会话级会话表示进行随机行(列)变换生成。

4)GCLMNN-F:采取物品退出策略,在每个会话中随机丢弃一些物品,生成的新会话作为正样本,其他会话作为负样本。

从图2可以看出,添加了图对比学习策略的模型GCLMNN-R,GCLMNN-F以及GCLMNN的性能远优于其他模型,证明了图对比学习的有效性。GCLMNN-S模型相比GCLMNN-G模型有一定提高,原因是加入会话-会话图可以聚合邻域会话表示,利用其他会话信息丰富当前会话表示。而本文所提模型GCLMNN的性能远优于其他模型,表明了对比学习方法的重要性。相比随机删除、扰动等方法,保留完整会话信息更能给模型带来信息增益。递归生成正负样本的方法最大限度地利用了两个级别的会话信息,有效缓解了数据稀疏问题。

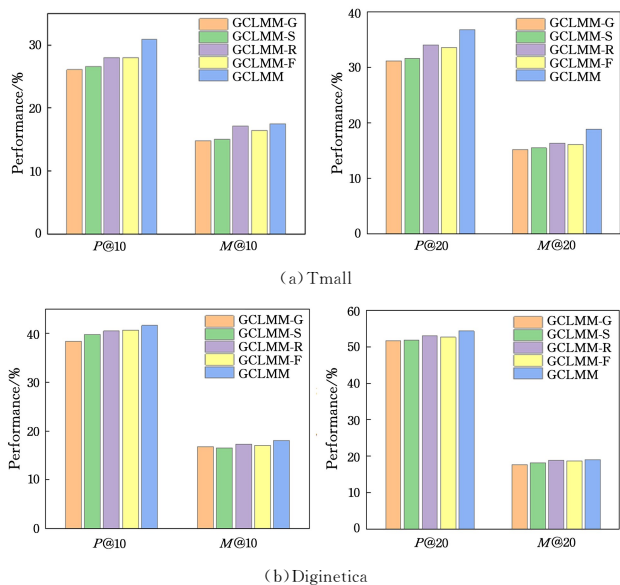


图2 不同图对比学习策略的性能比较结果

Fig. 2 Performance comparison results of different graph contrastive learning strategies

4.7 消融实验

本节中,为了研究模型各个模块的实际价值,用以下3种模型作为对比模型。

1)GCLMNN-P:去掉反向位置嵌入。

2)GCLMNN-N:去掉软注意力机制,计算会话中每个物品表示的平均值,获得物品级会话表示。

3)GCLMNN:本文模型,保留全部模块。

所得实验结果如图3所示,在两个不同的数据集上,两个模块的贡献度是不同的。

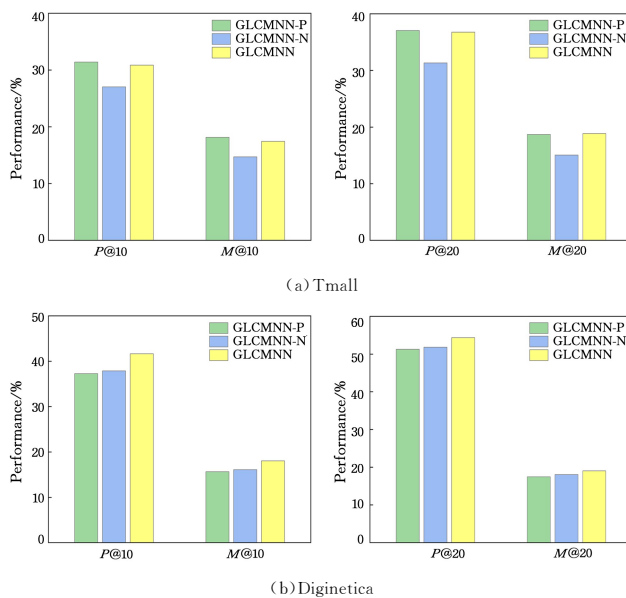


图3 消融实验

Fig. 3 Ablation experiments

相比本文模型GCLMNN,去掉反向位置嵌入的GCLMNN-P在Diginetica数据集上的性能下降,但在Tmall数据集上反而提升了,性能优于完整模型,这说明严格的顺序建模可能会对模型造成一定负面影响。而GCLMNN-N模型

在两个数据集上的性能都明显下降,证明了软注意力机制对模型的贡献,不同物品在会话中的优先级不同,其重要性高于物品表示平均值对模型的影响。

4.8 超参数影响分析

本文引入了一个超参数 β 来控制图对比学习强度。为了权衡推荐和对比学习这两个任务,避免两任务由于梯度冲突而对模型造成影响,本文在两个数据集上进行了大量实验,并

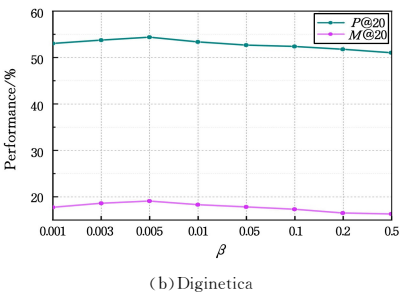
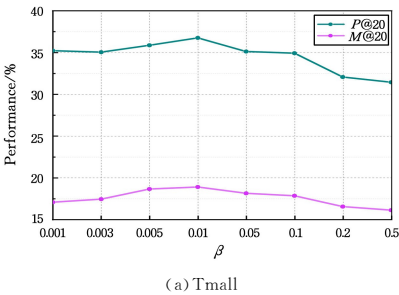
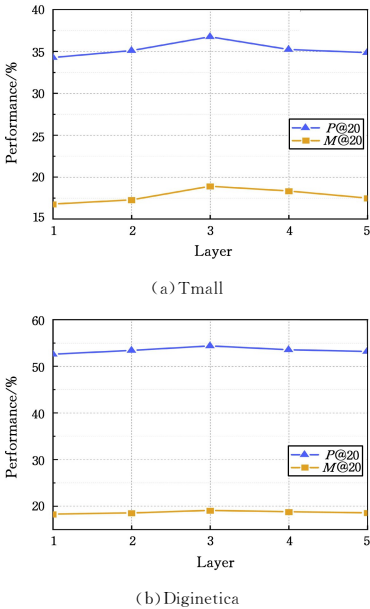


图 4 超参数 β 分析
Fig. 4 Analysis on hyperparameter β

4.9 模型层数的影响

为了研究图卷积网络中层数对模型的影响,本文将层数范围定为 $\{1, 2, 3, 4, 5\}$ 。从图 5 可观察到,对于 Tmall 和 Diginetica 数据集,都在 3 层网络设置时获得最佳性能。层数继续变多会导致物品表示过度平滑,不利于模型学习,性能反而会下降。



兴趣表示向量递归生成正负样本对,最后通过图对比学习方法增强数据表示,进一步增加会话兴趣区分性,提升推荐性能。与退出策略相比,本文方法更加全面地利用了会话信息,真正实现了数据的扩充,缓解了数据稀疏问题对模型的影响。广泛的实验进一步验证了模型的有效性。在后续的工作中,将尝试从其他角度挖掘难负样本,探索新的图对比学习方法,进一步提升推荐性能。

参 考 文 献

[1] WANG S,CAO L,WANG Y,et al. A survey on session-based recommender systems [J]. ACM Computing Surveys, 2021, 54(7):1-38.

[2] QIAO S,ZHOU W,WEN J,et al. Multi-perspective enhanced representation for effective session-based recommendation[J]. Knowledge-Based Systems,2023,263(12):110284.

[3] HIDASI B,KARATZOGLOU A,BALTRUNAS L,et al. Session-based recommendations with recurrent neural networks [J]. arXiv:1511. 06939,2015.

[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. Palo Alto, CA: AAAI, 2019,33(1):346-353.

[5] JAISWAL A,BABU A R,ZADEH M Z,et al. A survey on contrastive self-supervised learning[J]. Technologies, 2020,9(1): 2-23.

[6] ZHU Y Q,XU Y C,YU F,et al. Graph contrastive learning with adaptive augmentation[C]// Proceedings of the Web Conference 2021. New York: ACM,2021:2069-2080.

[7] KOREN Y,BELL R,VOLINSKY C. Matrix factorization techniques for recommender systems[J]. Computer, 2009,42(8): 30-37.

[8] RENDLE S,FREUDENTHALER C,SCHMIDT-THIEME L. Factorizing personalized markov chains for next-basket recommendation[C] // Proceedings of the 19th International Confe-

图 5 模型层数的影响

Fig. 5 Impacts of the number of model layer

结束语 针对会话中物品数量少、物品长尾分布等问题,本文提出了一种结合图对比学习的多图神经网络会话推荐模型 GCLMNN。首先在由物品局部图、物品全局图组成的多图神经网络中提取融入物品高阶邻域关联的物品嵌入,并设计网络注意机制和反向位置嵌入生成会话表示向量。区别于现有基于对比学习的会话推荐方法中利用随机裁剪、扰动等构建正负样本对的方式,本文从保留完整会话的角度出发设计会话-会话图,以提取跨会话信息,根据物品级和会话级的

rence on World Wide Web, 2010;811-820.

[9] LI J,REN P,CHEN Z,et al. Neural attentive session-based recommendation[C]// Proceedings of the 2017 ACM on Conference on Information and Knowledge Management, 2017;1419-1428.

[10] 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, ACM,2019;345-354.

[11] XU C,ZHAO P,LIU Y,et al. Graph Contextualized Self-Attention Network for Session-based Recommendation [C] // Proceedings of the 28th International Joint Conference on Artificial Intelligence, Macao,PEOPLES R China, IJCAI. 2019,19;3940-3946.

[12] QIU R H,LI J J,HUANG Z,et al. Rethinking the item order in session-based recommendation with graph neural networks [C]//Proceedings of the 28th ACM International Conference on Information and Knowledge Management, New York: ACM, 2019;579-588.

[13] CHEN J F,ZHU G H,HOU H J,et al. AutoGSR: Neural Architecture Search for Graph-based Session Recommendation [C]//Proceedings of The 45th International ACM SIGIR Conference on Research and Development in Information Retrieval, Madrid;SPAIN,2022;1694-1704.

[14] YANG L Q,LUO L H,XIN L F,et al. DAGNN: Demand-aware Graph Neural Networks for Session-based Recommendation[J]. arXiv:2105.14428v1,2022.

[15] HE K,FAN H,WU Y,et al. Momentum contrast for unsupervised visual representation learning [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020;9729-9738.

[16] DEVLIN J,CHANG M W,LEE K,et al. Bert: Pre-training of deep bidirectional transformers for language understanding[J]. arXiv:1810.04805,2018.

[17] ZHOU K,WANG H,ZHAO W X,et al. S³-rec: Self-supervised learning for sequential recommendation with mutual information maximization[C]// Proceedings of the 29th ACM International Conference on Information & Knowledge Management, New York: ACM,2020;1893-1902.

[18] XIE X,SUN F,LIU Z Y,et al. Contrastive learning for sequential recommendation [J]. arXiv:2010.14395,2020.

[19] XIA X,YIN H Z,YU J L,et al. Self-supervised hypergraph convolutional networks for session-based recommendation [C] // Proceedings of the AAAI Conference on Artificial Intelligence, Palo Alto,CA: AAAI,2021;4503-4511.

[20] XIA L H,HUANG C,XU Y,et al. Hypergraph Contrastive Collaborative Filtering[C]// Proceedings of The 45th International ACM SIGIR Conference on Research and Development in Information Retrieval, Madrid;PAIN,2022;70-79.

[21] PHUONG T M,THANH T C,BACH N X. Combining user-based and session-based recommendations with recurrent neural networks[C]// International Conference on Neural Information Processing, New York: Springer, Cham,2018;487-498.

[22] SARWAR B,KARYPIS G,KONSTAN J,et al. Item-based collaborative filtering recommendation algorithms [C] // Proceedings of the 10th International Conference on World Wide Web, New York: ACM,2001;285-295.

[23] LI J,REN P,CHEN Z,et al. Neural attentive session-based recommendation[C]// Proceedings of the 2017 ACM on Conference on Information and Knowledge Management, New York: ACM,2017;1419-1428.

[24] LIU Q,ZENG Y F,MOKHOSI R,et al. STAMP: short-term attention/memory priority model for session-based recommendation[C]// Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, New York: ACM,2018;1831-1839.

[25] LI A S,CHENG Z Y,LIU F,et al. Disentangled Graph Neural Networks for Session-based Recommendation[J]. IEEE Transactions on Knowledge and Data Engineering, 2022,35(8): 1-13.



LU Min, born in 1985, Ph.D, is a member of CCF(No. 23698S). His main research interests include web mining and information retrieval.



YUAN Ziting, born in 1996, postgraduate. Her main research interests include recommender system and contrastive learning.

(责任编辑:喻藜)