

# 融合全局信息的多图神经网络会话推荐

黄涛,徐贤

(华东理工大学 计算机科学与工程系,上海 200237)

E-mail:760518794@qq.com

**摘要:** 基于会话的推荐旨在根据当前会话预测下一个最可能交互的物品。由于单个会话点击序列较短,仅使用会话本身的信息很难提供准确的推荐。因此,综合考虑其它会话之间的交互信息已成为一种趋势,为了提高推荐性能,本文提出一种融合全局信息的多图神经网络会话推荐模型(GIMGNN)来增强会话推荐的效果。该模型首先通过超图卷积神经网络和门控图神经网络从全局会话超图和局部会话图中学习两个级别的物品表示,然后通过注意力机制将反向位置信息融合到两种表示中,最后利用融合后的表示完成预测。在两个真实数据集 Yoochoose 和 Diginetica 上进行了一系列实验,实验结果表明,对比性能最优的基准模型,GIMGNN 模型在 Yoochoose 上 P@20 和 MRR@20 至少提升了 2.42% 和 4.01%,在 Diginetica 上 P@20 和 MRR@20 至少提升了 6.56% 和 9.11%,验证了模型的有效性。

**关键词:** 会话推荐;超图卷积神经网络;门控图神经网络;注意力机制;位置信息

中图分类号: TP391

文献标识码: A

文章编号: 1000-1220(2024)04-0769-08

## Global Information Multi-graph Neural Network for Session-based Recommendation

HUANG Tao, XU Xian

(Department of Computer Science and Engineering, East China University of Science and Technology, Shanghai 200237, China)

**Abstract:** Session-based recommendation aims to predict the next most likely item to be interacted based on the current session. Given the short duration of a single session and the short sequence of clicks, it is difficult to provide accurate recommendations using only information about the session itself. It has become a trend to comprehensively consider the interaction information between other session sequences. In order to improve the effect of session recommendation, Global information multi-graph neural network for session-based recommendation (GIMGNN) is proposed to enhance the effect of session recommendation. Firstly, the model obtains the global session representation and the local session representation of the item by Hypergraph convolutional neural network and Gate graph neural network respectively, and then integrates the reverse position information into the two representations by attention mechanism, and finally uses the fused representation to complete the prediction. A series of experiments have been carried out on two real data sets, Yoochoose and Diginetica. The experimental results show that compared with the benchmark model with the best performance, GIMGNN model is on Yoochoose P@20 and MRR@20 At least 2.42% and 4.01% were improved on Diginetica P@20 and MRR@20 At least 6.56% and 9.11% were improved, which verify the validity of the model.

**Keywords:** session-based recommendation; hypergraph convolutional neural network; gate graph neural network; attention mechanism; position information

## 0 引言

在互联网搜索、电子商务、流媒体服务等在线应用中,推荐系统都扮演着关键角色,目的在于为用户提供有用的信息,以减轻信息过载对用户造成的冲击。传统方法主要依靠用户的偏好文件或长期的历史交互信息进行推荐。然而,在大多数情况下,由于用户没有任何可用的历史会话交互记录,可用于推荐的唯一资源就是当前会话的信息。因此,基于会话的推荐 SBR (Session-based recommendation) 引起了广泛关注,它根据给定的匿名用户会话序列按时间顺序预测用户下一个最有可能交互的物品<sup>[1]</sup>。

早期对基于会话推荐的研究大多分为两类,分别是基于

马尔科夫链<sup>[2,3]</sup>和协同过滤<sup>[4,5]</sup>的推荐方法。基于马尔科夫链的会话推荐的主要思想是根据用户之前的行为预测下一个行为。这种方法只考虑了相邻物品之间的顺序关系,而没有考虑物品之间的其它关系。通过将矩阵分解和一阶马尔科夫链相结合, Rendle 等人<sup>[6]</sup>提出了 FPMC (Factorized Personalized Markov Chains) 模型,该混合模型通过捕捉序列模式和用户的长期偏好来实现推荐,但忽略了序列之间的潜在表示,因此不能达到满意的效果。基于协同过滤的方法利用物品之间的相似性进行推荐,已经得到了广泛应用。在协同过滤中,许多方法通过计算用户和物品之间的相似矩阵来获得推荐结果。Kabbur 等人<sup>[7]</sup>使用结构方程方法将物品矩阵建模为两个低维潜在因子矩阵的乘积,以完成 Top-K 推荐。近年来,基于神

神经网络的方法也被应用于协同过滤。He等人<sup>[8]</sup>提出了一种基于神经网络的协同过滤框架 NCF (Neural Network-based Collaborative Filtering), 该框架利用多层感知器来学习用户物品交互功能。此外, Chen等人<sup>[9]</sup>提出了一种联合神经协同过滤模型 J-NCF (Joint Neural Network-based Collaborative Filtering), 该模型通过充分挖掘用户和物品的交互信息来学习物品的深层特征。虽然上述协同过滤方法被证明是有效的, 但是忽略了用户的最新偏好和整个点击序列的时间相关性。

近年来, 随着深度学习技术的不断发展, 基于循环神经网络 RNN (Recurrent Neural Network) 的会话推荐方法成为了研究热点。Hochreiter等人<sup>[10]</sup>提出 LSTM (Long Short Term Memory) 的 RNN 变体网络, 以更好地建模序列。Hidasi等人<sup>[11]</sup>将 RNN 结合到会话推荐中, 并利用多层门控循环单元 (GRU) 建模交互式物品序列。Jannach等人<sup>[12]</sup>将 GRU4REC (Gated Recurrent Unit for Recommendation) 与 KNN (K-Nearest Neighbor) 方法相结合, 对会话序列进行采样, 提高了推荐效果。在此基础上, Tan等人<sup>[13]</sup>通过数据增强改进了 RNN, 并考虑了会话数据分布的时间变化。Peng等人<sup>[14]</sup>提出了 HARSAM (A Hybrid Model for Recommendation Supported by Self-Attention Mechanism) 的深度学习模型, 该模型使用自注意力机制对用户交互数据进行建模, 并学习用户的潜在偏好表达。Jing等人<sup>[15]</sup>提出了 NARM (Neural Attentive Recommendation Machine) 注意力模型, 将注意力机制应用于会话特征编码器, 以捕捉用户在当前会话中的主要意图。Qiao等人<sup>[16]</sup>提出了基于注意力机制的短期记忆网络 STAMP (Short-Term Attention/Memory priority) 模型, 有效的捕获了用户的偏好信息。Wang等人<sup>[17]</sup>提出 CSRMM (collaborative session-based recommendation machine) 模型, 将 RNN 网络和注意力机制结合, 同时利用协同信息预测用户偏好。Wang等人<sup>[18]</sup>指出, 基于 RNN 的方法主要通过给对给定历史用户物品交互的序列依赖性建模来实现推荐。虽然基于深度学习的方法 (主要是基于 RNN 的方法) 在会话推荐中取得了一定程度的成功, 但是它们仍然存在一些不足, 具体如下: 1) 它们也只考虑当前会话序列中相邻交互项之间的依赖性, 而忽略了交互项与其它位置项之间的依赖关系; 2) 它们只依赖于当前会话的信息, 而不考虑其它交互会话信息。

图神经网络 GNN<sup>[19-22]</sup> (Graph Neural Network) 能够克服以往方法的局限性, 逐渐成为主流的会话推荐方法。与基于 RNN 的推荐方法不同, 基于 GNN 的方法大多首先将会话序列建模为会话图, 然后使用 GNN 在图中聚合相关邻居节点的信息。Wu等人<sup>[23]</sup>提出 SR-GNN (Session-based Recommendation with Graph Neural Network) 模型, 使用注意力机制计算用户的全局偏好, 将最后一项作为用户的当前偏好, 并使用全局和当前偏好的线性组合生成会话的最终表示。Xu等人<sup>[24]</sup>提出图上下文自注意力模型 GC-SAN (Graph Contextual Self-Attention Model based on Graph Neural Network), 该模型使用图神经网络和自注意力机制来学习会话序列中物品之间的长期依赖关系。Qiu等人<sup>[25]</sup>提出加权注意网络图模型 FGNN (Full Graph Neural Network), 该模型利用多权重图注意层 (WGAT) 来计算会话序列中物品之间的信息流, 从而获得物品表示, 然后通过特征提取器聚合物品表示以捕获会话特征。

Wu等人<sup>[26]</sup>提出了 GARG (Geographical Attentive Recommendation via Graph) 模型, 该模型将应卷积神经网络和注意力机制相结合为用户提供合适的新兴趣点。当使用上述的 GNN 模型构建会话图时, 不同的会话序列可能具有相同的图, 或者相同的图可能映射到不同的会话序列, 将导致信息丢失并影响最终的推荐效果。Chen等人<sup>[27]</sup>提出 LESSR (Lossless Edge-order preserving aggregation and Shortcut graph attention for Session-based Recommendation) 模型, 将会话序列建模为边缘保序图和快捷图, 并设计了边缘保序聚合层和快捷图注意力层, 解决了无效的长期依赖问题, 有效避免信息丢失。尽管上述方法在会话推荐中取得了可接受的推荐结果, 但是存在两个明显的缺点: 1) 基于 GNN 的方法只考虑了当前会话序列, 并没有考虑其它会话对当前会话的影响; 2) 将会话序列建模为一个简单的会话图, GNN 只能捕捉到物品之间的成对传递关系, 但无法获得物品之间复杂的高阶关系。

对于普通图, 一条边只能连接到两个顶点, 而超图是一般图的推广, 一条边可以连接任意数量的顶点。在构造超图时, 将所有会话序列建模为超边。超边上的所有物品都是相互连接的, 不同的超边通过共享项连接。这样构造的超图既考虑了其它会话之间的交互信息又包含了物品之间复杂的高阶信息。因此, 本文提出一种融合全局信息的多图神经网络会话推荐模型 (GIMGNN), 该模型使用超图卷积神经网络 (HGNN) 和门控图神经网络 (GGNN) 分别捕捉物品之间复杂的高阶关系和成对传递关系, 从而有效学习物品全局会话级别和局部会话级别的特征表示, 极大的提高了会话推荐的准确度。

本文的主要工作如下:

1) 提出全局会话超图和局部会话图相结合的多图神经网络会话推荐模型, 充分利用超图神经网络捕捉物品之间复杂的高阶信息能力, 提高会话推荐的性能。

2) 使用求和池化操作融合物品的全局和局部级别上下文信息表示之后, 通过使用注意力机制来处理融合的特征, 学习会话序列的最终表示, 从而得到更为准确的推荐结果。

3) 在两个真实数据集 Yoochoose 和 Diginetica 上进行了一系列实验, 实验结果表明, 对比性能最优的基准模型, GIMGNN 模型在 Yoochoose 上 P@20 和 MRR@20 至少提升了 2.42% 和 4.01%, 在 Diginetica 上 P@20 和 MRR@20 至少提升了 6.56% 和 9.11%, 验证了模型的有效性。

## 1 GIMGNN 模型

在本节中, 首先介绍 GIMGNN 模型总体架构, 其次对会话推荐问题进行描述, 构造两种类型的图模型, 分别是局部会话图和全局会话超图。接着分别使用门控图神经网络 (GGNN) 和超图卷积神经网络 (HGNN) 分别生成物品的局部会话特征表示和全局会话特征表示。然后通过注意力机制融合物品的局部会话特征表示和全局会话特征表示生成会话表示。最后将会话表示和物品全局特征表示进行线性组合得到最终表示, 从而计算每个推荐候选项的排名分数。

### 1.1 模型整体架构

本节将超图卷积神经网络和门控图神经网络应用于基于会话的推荐方法中, 提出了一种融合全局信息的多图神经网络会话推荐模型 (GIMGNN), 如图 1 所示。

该模型主要模块构成如下:

1) 构图模块:根据当前会话序列构建局部会话图,以及根据历史会话序列构建全局会话超图.

2) 局部表示学习模块:将构建的局部会话图通过门控图神经网络(GGNN)生成物品局部会话特征表示.

3) 全局表示学习模块:将构建的全局会话超图通过超图

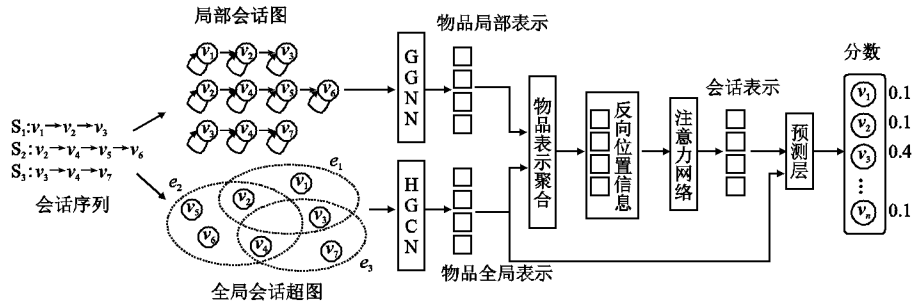


图 1 GIMGNN 模型结构框架

Fig. 1 Model framework of GIMGNN

卷积神经网络(HGCN)生成物品全局会话特征表示.

4) 聚合模块:通过注意力机制将反向位置嵌入融合到物品局部会话特征表示和物品全局会话特征表示中得到会话表示.

5) 预测模块:根据聚合模块得到的会话表示和物品全局会话特征表示线性组合得到最终表示预测下一个交互物品.

1.2 问题描述

假设  $V = \{v_1, v_2, \dots, v_n\}$  表示系统中所有物品的集合,  $s = [v_1^s, v_2^s, \dots, v_l^s]$  表示当前会话序列,其中  $v_i^s$  表示在会话  $s$  中点击的物品  $v_i$ ,  $l$  表示会话  $s$  的长度. 会话  $s$  中的物品按时间顺序表示,如果  $j > i, v_i^s$  在  $v_j^s$  之前交互.

如前所述,会话推荐的任务是根据当前会话预测下一个最有可能交互的物品. 形式上,本文对会话推荐的定义如下:假设  $S = \{s_1, s_2, \dots, s_n\}$  表示历史会话序列,  $s = [v_1^s, v_2^s, \dots, v_l^s]$  ( $v_i \in V$ ) 表示当前会话序列. 会话推荐的目标是尝试建立函数  $f(\cdot)$  预测一个物品  $v_{i+1}$  作为会话  $s$  中最有可能点击的下一个物品. 例如  $v_{i+1} = f(S, s; \theta)$ , 其中  $\theta$  表示可学习参数集. 为了获得最佳预测结果,将  $S$  中的  $S_i$  序列作为训练集,并且对于每个  $s_i \in S_i$ ,需要学习函数  $f(S, s; \theta)$  中最佳参数  $\theta$ . 公式如下:

$$\theta = \underset{\theta}{\operatorname{argmin}} \frac{1}{|S_i|} \sum_{s \in S_i} (L(v_{m_i}^s, f(S, [v_1^s, \dots, v_{m_i-1}^s]; \theta))) \quad (1)$$

通过公式(1)最小化损失函数  $L(\cdot)$ . 其中  $|S_i|$  是训练集的大小,  $m_i$  是序列  $s_i$  的长度.

1.3 构图模块

本节定义了两个图模型,局部会话图和全局会话超图,以表示当前序列中不同级别之间的物品传递信息. 局部会话图表示当前会话序列中成对物品-物品关系的传递信息,全局会话序列表示所有会话序列中物品-物品关系的复杂高阶信息.

局部会话图是一个有向图,描述给定会话序列中物品的点击顺序. 假设  $s = [v_1^s, v_2^s, \dots, v_l^s]$  是会话序列,  $G_s = (V_s, E_s)$  表示其局部会话图,其中  $V_s$  是由  $s$  中点击的所有物品组成的物品集合.  $E_s = \{e_{i,j}^s\}$  是有向边集,每条边  $(v_i^s \rightarrow v_{i+1}^s) \in E_s$  表示会话  $s$  中节点  $v_i^s$  下一个点击物品是  $v_{i+1}^s$ .

图 2 中显示了单个会话  $s = \{v_1, v_2, v_3, v_2, v_4\}$  的局部会话图. 在图中每个节点都与自身有连接,因此可以在建模中融合自身的消息. 局部会话图由有向图表示,可以很好地表示会话序列中物品之间的顺序相关性. 对于本文构造的局部图的每

一个节点,都可能有 4 种类型的边与其相连,分别表示为  $r_{in}$ ,  $r_{out}$ ,  $r_{in-out}$  和  $r_{self}$ .  $r_{in}$  表示其它节点传递信息到该节点.  $r_{out}$  表示该节点将信息传输到其它节点.  $r_{in-out}$  表示该节点和节点之间存在双向传输信息,  $r_{self}$  表示该节点自身的传递信息.

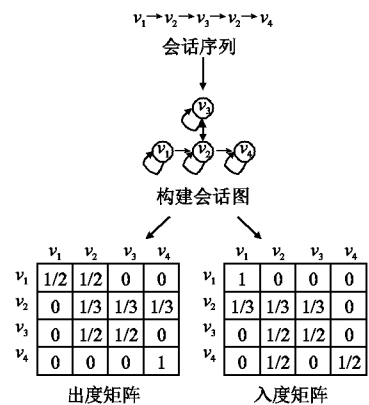


图 2 局部会话图

Fig. 2 Local session graph

全局会话超图用于表示物品的全局信息. 全局会话超图是一个连接任意多个顶点的超边组成的无向超图. 超图  $G$  可以表示为三元组  $G = (V, E, W)$ ,  $V, E$  和  $W$  分别是节点集,超边和超边的权重.  $V$  表示系统中的所有物品,每个  $v \in V$  代表一个物品.  $E$  表示所有历史会话序列,每个  $e \in E$  表示会话序列. 每个  $w(e) \in W$  是超边  $e$  的权重. 在代数中,  $W$  可以定义为对角矩阵  $W \in \mathbb{R}^{|E| \times |E|}$ ,超图可以用关联矩阵  $H \in \mathbb{R}^{|V| \times |E|}$  表示,其条目  $h(v, e)$  定义为:

$$h(v, e) = \begin{cases} 1, & \text{if } v \in e \\ 0, & \text{if } v \notin e \end{cases} \quad (2)$$

根据超图的定义,可以计算超边的度矩阵  $D = \sum_{e \in E} w(e) h(v, e)$  和所有顶点的度矩阵为  $B = \sum_{v \in V} h(v, e)$ ,其中  $D$  和  $B$  都是对角矩阵.

在超图中,如果任意两个超边具有公共顶点,则它们是连通的. 并且超边内的所有顶点都是完全连接的,因此超图包含物品之间的高阶未配对关系,可用于物品全局级物品的特征表示. 图 3 中显示了 3 个会话序列形成的超图,虚线中的 3 个

圆是超边,由它们之间的公共顶点  $v_2, v_3$  和  $v_4$  连接,并计算

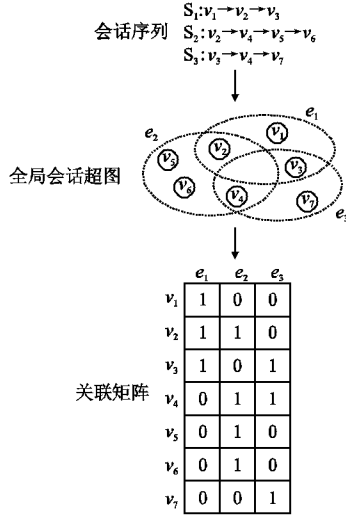


图3 全局会话超图

Fig. 3 Global session hypergraph

其对应的关联矩阵.

#### 1.4 局部表示学习模块

门控图神经网络 (GGNN) 用于获取局部会话序列中成对物品-物品关系的传递信息. 在局部会话图中, 节点向量表示的更新函数如下:

$$\mathbf{a}_{s,i}^t = (\mathbf{A}_{s,i}^{out} | \mathbf{A}_{s,i}^in) [v_1^{t-1}, v_2^{t-1}, \dots, v_n^{t-1}]^T \mathbf{H} + \mathbf{b} \quad (3)$$

$$\mathbf{Z}_{s,i}^t = \sigma(\mathbf{W}_z \mathbf{a}_{s,i}^t + \mathbf{G}_z v_i^{t-1}) \quad (4)$$

$$\mathbf{R}_{s,i}^t = \sigma(\mathbf{W}_r \mathbf{a}_{s,i}^t + \mathbf{G}_r v_i^{t-1}) \quad (5)$$

$$\tilde{\mathbf{h}}_{v_i}^t = \tanh(\mathbf{W}_c \mathbf{a}_{s,i}^t + \mathbf{G}_c (\mathbf{R}_{s,i}^t \odot v_i^{t-1})) \quad (6)$$

$$\mathbf{h}_{v_i}^t = (1 - \mathbf{Z}_{s,i}^t) \odot v_i^{t-1} + \mathbf{Z}_{s,i}^t \odot \tilde{\mathbf{h}}_{v_i}^t \quad (7)$$

其中  $\mathbf{H} \in \mathbb{R}^{d \times 2d}$ ,  $\mathbf{W}_z, \mathbf{W}_r, \mathbf{W}_c \in \mathbb{R}^{2d \times d}$ ,  $\mathbf{G}_z, \mathbf{G}_r, \mathbf{G}_c \in \mathbb{R}^{d \times d}$  表示权重矩阵,  $\mathbf{Z}_{s,i}^t$  和  $\mathbf{R}_{s,i}^t$  分别表示更新门和重置门,  $v_1^{t-1}, v_2^{t-1}, \dots, v_n^{t-1}$  为会话  $s$  中的节点序列,  $\sigma(\cdot)$  表示激活函数,  $\odot$  表示对应元素相乘.  $\mathbf{A}_{s,i}^{out}$  表示局部会话图的出度矩阵,  $\mathbf{A}_{s,i}^in$  表示局部会话图的入度矩阵. 公式(3)用于不同节点之间的信息传播, 在邻接矩阵的作用下提取邻域的潜在向量, 并将其作为输入到神经网络中. 公式(4)和公式(5)通过更新门和重置门分别决定要保留和丢弃的信息. 公式(6)根据前一状态、当前状态和重置门构造候选状态. 公式(7)表示在更新门的控制下, 由前一个隐藏状态和候选状态得到最终状态. 通过对会话图中的所有节点进行更新, 达到收敛之后, 便可以得到最终的节点向量. 之后使用注意力机制计算不同邻居节点对当前节点的影响, 影响程度由注意力机制计算的权重表示. 给定一个节点  $v_i$ , 可以通过元素乘积和非线性变换计算节点  $v_j$  对其的影响权重:

$$e_{ij} = \text{LeakyReLU}(\mathbf{a}_{ij}^T (\mathbf{h}_{v_i} \odot \mathbf{h}_{v_j})) \quad (8)$$

$e_{ij}$  表示节点  $v_j$  对于节点  $v_i$  的重要性,  $\text{LeakyReLU}(\cdot)$  表示激活函数,  $r_{ij}$  表示节点  $v_i$  和  $v_j$  的关系,  $\mathbf{a}_{\cdot} \in \mathbb{R}^d$  是权重矩阵. 通过权重矩阵来表示图中所有节点对于节点  $v_i$  的影响. 为了使不同节点之间的权重具有可比性, 使用 softmax 函数对其进行归一化:

$$\alpha_{ij} = \frac{\exp(\text{LeakyReLU}(\mathbf{a}_{ij}^T (\mathbf{h}_{v_i} \odot \mathbf{h}_{v_j})))}{\sum_{v_k \in \mathcal{N}_{v_i}^s} \exp(\text{LeakyReLU}(\mathbf{a}_{ik}^T (\mathbf{h}_{v_i} \odot \mathbf{h}_{v_k})))} \quad (9)$$

公式(9)注意力系数  $\alpha_{ij}$  是不对称的, 因为节点的邻居是不同的, 意味着对每个节点的贡献是不平等的. 接下来, 通过线性组合计算节点  $v_i$  的特征表示:

$$\mathbf{h}_{v_i}^{local} = \sum_{v_j \in \mathcal{N}_{v_i}^s} \alpha_{ij} \mathbf{h}_{v_j} \quad (10)$$

通过上述计算, 可以将其他节点对当前节点的影响信息和当前节点本身的信息融合在一起, 以表达当前节点的局部特征表示. 并且通过注意力机制, 降低了噪声对局部会话级别物品特征学习的影响.

#### 1.5 全局表示学习模块

超图卷积神经网络 (HGNN) 用于获取所有会话序列中物品之间的高阶关系. 在超图上定义卷积运算的主要挑战是如何传播相邻节点的信息. 参考 Feng 等人<sup>[28]</sup> 提出的谱图卷积, 在本文的模型中, HGNN 的定义如下:

$$\mathbf{h}_v^{l+1} = \mathbf{D}^{-1} \mathbf{H} \mathbf{W} \mathbf{B}^{-1} \mathbf{H}^T \mathbf{h}_v^l \quad (11)$$

超图卷积可以看作是顶点-超边-顶点 (V-H-V) 传递. 公式(11)中的  $\mathbf{H}^T \mathbf{h}_v^l$  表示从顶点到超边的聚合信息, 然后与  $\mathbf{H}$  相乘以从相反方向进行聚合信息. 最后, 通过 HGNN 获得节点在每一层的嵌入. 对于给定的节点  $v_i$ , 从第 0 层到第  $L$  层, 每个层次计算  $v_i$  的表示. 最终对  $\mathbf{h}_{v_i}^l$  的结果执行以下平均运算, 得到节点的全局级别表示:

$$\mathbf{h}_{v_i}^{global} = \frac{1}{L+1} \sum_{l=0}^L \mathbf{h}_{v_i}^l \quad (12)$$

#### 1.6 聚合模块

对于每个物品, 本文通过合并其局部表示和全局表示获得物品的最终表示:

$$\begin{aligned} \mathbf{h}_{v_i}^{local} &= \text{dropout}(\mathbf{h}_{v_i}^{local}) \\ \mathbf{h}_{v_i}^{global} &= \text{dropout}(\mathbf{h}_{v_i}^{global}) \\ \mathbf{h}_{v_i}^* &= \text{SumPooling}(\mathbf{h}_{v_i}^{local}, \mathbf{h}_{v_i}^{global}) \end{aligned} \quad (13)$$

公式(13)对两个层次的物品特征表示做求和池化操作, 从而得到物品的最终表示  $\mathbf{h}_{v_i}^*$ . 过程中增加 dropout 层, 防止提取高阶物品信息可能带来的过拟合问题. 在将会话序列输入给 GGNN 和 HGNN 之后, 可以获得会话中涉及的物品最终表示, 即  $\mathbf{H} = \{\mathbf{h}_{v_1}^*, \mathbf{h}_{v_2}^*, \dots, \mathbf{h}_{v_l}^*\}$ .

时间信息对于更好的推荐结果也是不可避免的. 在本文方法中, 通过使用可学习位置矩阵  $\mathbf{p} = [\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_l]$  将反向位置嵌入与物品最终表示相结合, 其中  $\mathbf{p}_l$  表示第一个位置的向量,  $l$  是会话序列的长度. 对于给定会话中的物品  $v_i$ , 在连接两个级别的信息  $\mathbf{h}_{v_i}^*$  和位置矩阵  $\mathbf{p}$  后, 然后执行非线性变换操作来计算  $v_i$  的位置嵌入, 其表示为:

$$\mathbf{z}_i = \tanh(\mathbf{W}_1 [\mathbf{h}_{v_i}^* \parallel \mathbf{p}_{l-i+1}] + \mathbf{b}) \quad (14)$$

其中参数  $\mathbf{W}_1 \in \mathbb{R}^{d \times 2d}$  和  $\mathbf{b} \in \mathbb{R}^d$  是一组可训练的参数,  $\parallel$  表示串联操作. 由于会话序列的长度不固定, 所以选择反向位置嵌入. 相对于前向位置信息, 预测物品和当前物品之间的距离包含更有效的信息. 因此, 反向位置信息可以更准确地表示每个物品的重要性.

会话的表示与学习到的物品信息密切相关. 为了获得会话的表示, 本文平均了会话序列中涉及的信息:

$$\mathbf{s}^* = \frac{1}{l} \sum_{i=1}^l \mathbf{h}_{v_i}^* \quad (15)$$

基于位置嵌入信息  $\mathbf{z}_i$  和会话表示  $\mathbf{s}^*$ , 本文采用软注意力机制来计算权重:

$$\beta_i = \mathbf{q}^T \sigma(\mathbf{W}_2 \mathbf{z}_i + \mathbf{W}_3 \mathbf{s}^* + \mathbf{c}) \quad (16)$$

其中,  $\mathbf{W}_2, \mathbf{W}_3 \in \mathbb{R}^{d \times d}$  和  $\mathbf{q}, \mathbf{c} \in \mathbb{R}^d$  都是可学习的参数. 因此, 通过线性组合操作获得会话的最终表示:

$$\mathbf{S} = \sum_{i=1}^l \beta_i \mathbf{h}_{v_i}^* \quad (17)$$

前面的计算过程表明, 会话表示  $\mathbf{S}$  表示不仅融合了全局和局部级别上下文信息, 而且还包含了所有物品的位置和顺序信息, 因此会话表示  $\mathbf{S}$  可以很好地表示会话特征.

### 1.7 预测模块

与物品  $v_i$  的局部特征表示  $\mathbf{h}_{v_i}^{local}$  相比, 全局特征表示  $\mathbf{h}_{v_i}^{global}$  包含更丰富的高阶信息. 因此, 本文将会话的最终表示与  $\mathbf{h}_{v_i}^{global}$  相结合, 并使用 softmax 函数计算会话序列中交互物品的相应概率, 其公式为:

$$\hat{y}_i = \text{softmax}(\mathbf{S}^T \mathbf{h}_{v_i}^{global}) \quad (18)$$

$\hat{y}_i \in \hat{y}$  表示当前会话中下一个点击物品  $v_i$  的概率. 损失函数是每个会话的真值向量和预测概率向量的交叉熵, 其定义为:

$$L(\hat{y}_i) = - \sum_{i=1}^m y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i) \quad (19)$$

其中,  $y_i$  表示为 one-hot 向量.

## 2 实验与分析

在本节中, 首先介绍实验的数据集和预处理、评估指标、基准方法和实验参数设置, 然后设计了一些列对比实验, 以回答下列 4 个问题:

问题 1. 与现有的会话推荐基准方法相比, GIMGNN 表现如何?

问题 2. 融合全局信息是否增强了会话推荐效果? GIMGNN 模型中每个模块的性能如何?

问题 3. 反向位置嵌入对 GIMGNN 的性能影响如何?

问题 4. 不同超参数(dropout)的设置对 GIMGNN 的性能影响如何?

### 2.1 数据集和预处理

本文实验将基于两个著名的基准数据集 Yoochoose 和 Diginetica 进行. 其中 Yoochoose 数据集来自于 2015 年 RecSys 挑战赛, 通过 <http://2015.recsyschallenge.com/challenge.html> 获取, 数据集包括一个网站的点击历史记录. Diginetica 数据集通过 <https://competitions.codalab.org/competitions/11161> 获取, 来自 2016 年 CIKM 杯, 数据集由典型交易数据组成.

对于两个数据集的预处理工作参考了文献[23], 首先过滤掉长度为 1 的会话和两个数据集中出现次数少于 5 次的物品, 然后将上周的会话(最新数据)设置为测试数据集, 剩余的历史数据用于训练集. 由于 Yoochoose 数据集非常大, 本文对 Yoochoose 数据集进行了额外的处理. 此外, 已有研究证明使用接近测试集的训练集进行模型训练可以提高推荐性能. 因此, 实验中仅使用 Yoochoose1/64 作为数据集. 表 1 列出了

数据预处理后两个数据集的统计信息.

表 1 预处理数据统计

Table 1 Preprocessed data statistics

统计项	Yoochoose1/64	Diginetica
点击数量	557248	982961
训练集会话数	369859	719470
测试集会话数	55898	60858
物品数量	16766	43097
平均会话长度	6.16	5.12

### 2.2 评价指标

为了便于与基线模型进行比较, 本文选择了常用的精度(P)和平均倒数排名(MRR)作为评估指标. 在实际推荐中, 系统通常同时推荐多个物品. 为了评估不同物品数量的推荐效果, 使用 P@K 和 MRR@K 以测量模型的性能, 其中 K 表示推荐物品的数量.

P@K 被计算为推荐排名列表中测试用例前 K 位的正确项目, 并定义为:

$$P@K = \frac{n_{hit}}{N} \quad (20)$$

其中 N 是测试集中的序列数,  $n_{hit}$  是在排名列表中前 K 个物品中正确推荐的物品数.

MRR@K 是对正确推荐的物品  $v_i$  出现在包含 K 个物品的推荐列表 I 中的位置进行评分, 其具体值等于  $v_i$  在 I 中的排名倒数. 如果  $v_i$  位于 I 的第 1 位时 MRR@K 为 1, 当  $v_i$  未出现在 I 中时 MRR@K 为 0. 假设测试集的大小为 N, 取平均值 MRR@K 作为评估的度量:

$$MRR@K = \frac{1}{N} \sum_{v_i \in \mathcal{S}_{test}} \frac{1}{rank(v_i)} \quad (21)$$

其中  $rank(v_i)$  是  $v_i$  在推荐列表中的排名.

### 2.3 对比模型

为了验证 GIMGNN 模型的性能, 实验用以下 9 种推荐模型作为对比模型:

1) POP: 推荐训练数据集中频率最高的前 N 项.

2) Item-KNN: 通过余弦距离来衡量两个物品之间的相似度, 并根据相似度推荐物品.

3) FPMC<sup>[6]</sup>: 同时考虑了矩阵分解和一阶马尔可夫链, 是一种混合方法, 但在计算推荐分数时, 忽略了用户的潜在表示.

4) GRU4REC<sup>[12]</sup>: 通过门控神经网络(GRU)将会话序列建模为最终会话表示, 并使用排名损失来训练模型.

5) NARM<sup>[15]</sup>: 基于 GRU4Rec 的模型, 该模型扩展了注意力层, 并结合了 RNN 的编码状态, 使模型能够密切关注会话序列的重要特征.

6) STAMP<sup>[16]</sup>: 将先前工作中的 RNN 编码器替换为多重注意力, 并且将当前会话中最后一项的自我注意力作为用户的短期兴趣.

7) CSR<sup>[17]</sup>: 假设历史会话中包含与目标会话相似的用户偏好, 并将这种协作信息应用于推荐任务.

8) SR-GNN<sup>[23]</sup>: 使用 GRU 获得物品的嵌入表示, 与 STAMP 类似, 它通过关注最后一个物品来计算会话的特征表示.

9) GC-SAN<sup>[24]</sup>: 使用自注意力网络来学习会话中物品之间的全局和局部依赖信息.

## 2.4 参数设置

在本文的实验中,将隐向量维度大小设置为 100,批处理大小设置为 100.所有可学习参数均使用高斯分布初始化,平均值为 0,标准偏差为 0.1.学习率设置为 0.001,并由 Adam 优化器优化.本文在两个数据集上将 epoch 设置为 20.对于 GIMGNN 的层数设置,不同的数据集对应不同的层数.通过实验发现,对于数据集 Yoochoose1/64 和 Diginetica,当层数设置为 3 时,该模型的性能最佳.对于基线,如果原始论文的实验评估机制和数据集与本文相同,本文将采用最佳实验结果进行比较.此外,为了公平起见,还将参数设置为模型性能最佳时的参数.

## 2.5 结果分析(问题 1)

通过表 2 的实验结果可以看出,与基线模型相比,本文提出的 GIMGNN 模型在两个数据集的指标上都优于基线模型.

表 2 不同模型在两个数据集上性能实验结果%

模 型	Yoochoose1/64		Diginetica	
	P@ 20	MRR@ 20	P@ 20	MRR@ 20
POP	6.72	1.65	0.88	0.21
Item-KNN	51.61	21.80	35.75	11.57
FPMC	45.62	15.01	26.52	6.95
GRU4REC	60.63	22.88	29.45	8.33
NARM	68.24	28.63	49.70	16.17
STAMP	68.72	29.66	45.65	14.32
CSRM	69.91	29.54	50.82	17.16
SR-GNN	70.58	30.95	50.73	17.68
GC-SAN	70.09	29.65	50.75	17.43
GIMGNN	<b>72.29</b>	<b>32.19</b>	<b>54.08</b>	<b>19.29</b>
Improve(%)	2.42	4.01	6.56	9.11

传统的 POP 模型实验结果最差,因为只考虑了训练数据集中频率最高的前 N 项,而没有考虑其它交互信息.FPMC 方法通过结合矩阵分解和一阶马尔可夫链来捕获用户偏好,表现出比 POP 更好的性能,证明了用户偏好在推荐中的重要性.在传统方法中,Item-KNN 模型在 Yoochoose1/64 和 Diginetica 数据集上显示了最佳结果.该模型根据当前会话和其他会话之间的相似性推荐物品,证明了会话之间存在一定的依赖性.然而,由于传统推荐模型无法捕捉同一会话中物品之间的顺序关系或顺序相关性,所以推荐的准确性明显受到影响.

从表 2 中不难发现,基于深度学习的模型优于传统的推荐系统方法.在性能方面,GRU4REC 首先使用 RNN 对会话序列进行建模获取特征表示,其性能仍然低于 NARM 和 STAMP 模型.因为 GRU4REC 只考虑顺序关系,而不考虑序列中的其余信息,因此很难获得用户偏好的变化.这意味着,虽然 RNN 非常适合序列建模,但很难解决基于会话推荐中的用户意图可能发生变化的问题.通过考虑会话中不同物品的重要性,NARM 和 STAMP 使用注意机制更准确地表达用户的意图,在一定程度上提高了推荐效果,性能优于 GRU4REC.通过比较 RNN 和注意机制相结合的 NARM 与完全使用注意机制的 STAMP 的性能,本文发现 STAMP 的性能明显优于 NARM.前一种方法通过迭代多个注意层来替换先

前工作中的 RNN 编码器,并将当前会话中的最后一项视为用户的短期偏好,进一步证明了使用 RNN 学习用户的表示可能会导致用户意图的偏离.与 NARM 和 STAMP 相比,CSRM 方法在 Diginetica 数据集上表现出更好的性能.它使用内存网络来研究最近的 m 个会话,以便更好地预测用户在当前会话中的意图.正如 CSRM 的性能所示,为了更好地进行会话推荐,需要考虑其他会话对当前会话的影响.

通过 SR-GNN 与 GC-SAN 实验结果可知,基于 GNN 会话模型的推荐效果明显好于传统推荐方法与深度学习方法.因为将会话序列建模为会话图以及使用 GNN 可以更好的捕获物品之间的成对传递关系,表明了在做推荐时考虑物品之间的独立的重要性.与 RNN 相比,GNN 可以捕获会话序列中物品之间更复杂的依赖信息.两种基于 GNN 的方法将会话序列建模为简单的图,不能够充分捕捉物品的成对传递,也无法学习物品之间的复杂高阶关系,这在一定程度上影响了推荐效果.

与基线模型不同,本文提出的 GIMGNN 模型将历史会话序列建模为超图,并使用超图神经网络(HGCN)学习物品之间的复杂高阶关系.模型同时考虑了局部和全局会话级别的上下文信息,因此可以有效地表示当前会话序列中物品的特征,这也是 GIMGNN 具有优异性能的原因.

## 2.6 全局特征对性能的影响(问题 2)

在两个数据集上进行实验,以评估全局级特征编码器和会话级特征编码器的有效性.本文设计了两种对比模型:

1) GIMGNN-NL:删除局部表示学习模块(GGNN),只留下全局表示学习模块(HGCN)来捕获全局会话级别的上下文信息.

2) GIMGNN-NG:删除全局表示学习模块(HGCN),只留下局部表示学习模块(GGNN)来捕获局部会话级别的上下文信息.

表 3 在 Yoochoose1/64 和 Diginetica 两个数据集上比较了 3 个模型的推荐结果,并分别给出了评估指标 P@20 和 MRR@20.从表中可知,本文提出的 GIMGNN 方法在推荐性能方面显示出最佳结果.在 Diginetica 数据集上,GIMGNN-NL 的性能优于 GIMGNN-NG,表明使用超图神经网络(HGCN)捕捉高阶信息的有效性.在 Yoochoose1/64 数据集上,GIMGNN-NG 的性能优于 GIMGNN-NL,可能是因为 Yoochoose1/64 数据集中的会话平均长度比 Diginetica 数据集中的会话平均长度长,有助于门控图神经网络(GGNN)有效学习会话中物品之间的依赖关系.实验结果表明,通过考虑物品之间复杂高阶信息,可以有效提高会话推荐的性能.

表 3 对比模型在两个数据集上性能实验结果%

Table 3 Experimental results of contrast models on two datasets%

模 型	Yoochoose1/64		Diginetica	
	P@ 20	MRR@ 20	P@ 20	MRR@ 20
GIMGNN-NL	67.62	29.37	53.12	18.78
GIMGNN-NG	71.05	30.02	51.51	17.38
GIMGNN	<b>72.29</b>	<b>32.19</b>	<b>54.08</b>	<b>19.29</b>

## 2.7 位置向量对性能的影响(问题 3)

为了验证位置向量对性能影响,并评估在 GIMGNN 中提

出的反向位置向量的有效性,本节设计了一系列对比模型:

1) GIMGNN-NP:使用前向位置向量替换反向位置向量的 GIMGNN 模型。

2) GIMGNN-SA:使用自注意机制取代位置感知注意力的 GIMGNN 模型。

表 4 显示了不同对比度模型的性能。从表中可以看出 GIMGNN-NP 在两个数据集上表现不佳,因为模型无法捕捉其它物品与预测物品之间的距离。GIMGNN-SA 在 Diginetica 数据集上的表现优于 GIMGNN-NP,表明会话中的最后一项包含最相关的推荐信息。然而,它在 Yoochoose1/64 数据集上表现不佳,因为它缺乏对每个物品贡献的更全面的判断。与这两种变体相比,反向位置嵌入证明了模型的有效性,也验证了反向位置信息可以更准确地表明每个物品的重要性。此外,通过注意力机制,有效过滤了当前会话中的噪声,使得模型的性能更好。

表 4 对比模型在两个数据集上性能实验结果%

Table 4 Experimental results of contrast models on two datasets%

模型	Yoochoose1/64		Diginetica	
	P@20	MRR@20	P@20	MRR@20
GIMGNN-NP	69.35	30.26	50.45	17.65
GIMGNN-SA	68.64	29.42	51.68	17.94
GIMGNN	<b>72.29</b>	<b>32.19</b>	<b>54.08</b>	<b>19.29</b>

## 2.8 超参数对性能的影响(问题 4)

Dropout 作用是防止模型过拟合,原理是在训练期间随机丢弃具有给定可能性的神经元,但将所有神经元用于测试。在实验中,神经元脱落的可能性在 0.0~0.9 范围内变化,间隔为 0.1,评估指标为 P@20。图 4 显示了 GIMGNN 在两个数据集不同 dropout 下的实验性能,说明了模型在两个数据集上的性能在开始时随着 dropout 的增加而增加,当 dropout 增长到

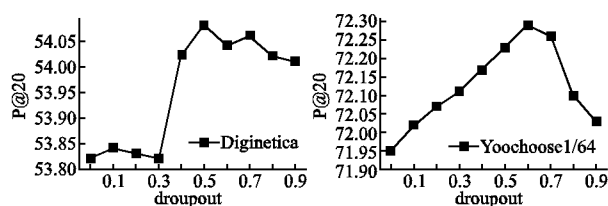


图 4 不同 dropout 对推荐性能的影响

Fig.4 Impact of different dropout on recommended performance

一定值时,模型性能开始下降。也就是说,模型的性能在曲线的拐点处最好。特别是在 Yoochoose1/64 和 Diginetica 数据集中,本文的 GIMGNN 模型在 dropout 分别为 0.6 和 0.5 时表现最佳。因此,本文在两个数据集上的实验中将 dropout 设置为相应的值。

## 3 结束语

基于图神经网络的会话推荐模型是当前学术研究的热点,物品的高阶交互信息可以提高基于图神经网络的会话模型的推荐效果。为了充分考虑物品的局部和全局会话上下文

信息,本文的工作首先将当前会话序列转换为局部会话图,其次将所有会话序列转换为全局会话超图,然后使用超图卷积神经网络(HGCN)和门控图神经网络(GGNN)捕捉物品的全局会话特征表示和局部会话特征表示。并且,通过使用注意力机制处理融合特征来学习会话序列的最终表示。大量实验结果表明,本文提出的 GIMGNN 模型始终优于最先进的方法。然而,使用超图卷积神经网络(HGCN)捕获全局信息也有一个缺点,可能会将不相关的信息融合到当前会话中。因此,在未来的工作中,将研究如何构建基于超图的模型,以缓解无关信息对会话推荐结果的影响。

## References:

- [1] ZENG Y F, MU Q L, ZHOU L, et al. Graph embedding based session perception model for next-click recommendation [J]. Journal of Computer Research and Development, 2020, 57(3): 590-603.
- [2] Shani G, Heckerman D, Brafman R I. An MDP-based recommender system [J]. Journal of Machine Learning Research, 2005: 1265-1295, doi:10.1007/s10846-005-9023-3.
- [3] Zimdars A, Chickering D M, Meek C. Using temporal data for making recommendations [C]//Proceedings of the 17th Conference in Uncertainty in Artificial Intelligence, 2001: 580-588.
- [4] Ekstrand M D, Riedl J T, Konstan J A. Collaborative filtering recommender systems [J]. Foundations and Trends in Human-Computer Interaction, 2011, 4(2): 81-173.
- [5] Koren Y, Bell R, Volinsky C. Matrix factorization techniques for recommender systems [J]. Computer, 2009, 42(8): 30-37.
- [6] Rendle S, Freudenthaler C, Schmidt-Thieme L. Factorizing personalized markov chains for next-basket recommendation [C]//Proceedings of the 19th International Conference on World Wide Web, 2010: 811-820.
- [7] Kabbur S, Ning X, Karypis G. Fism: factored item similarity models for top-n recommender systems [C]//Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2013: 659-667.
- [8] Rendle S, Krichene W, Zhang L, et al. Neural collaborative filtering vs. matrix factorization revisited [C]//14th ACM Conference on Recommender Systems, 2020: 240-248.
- [9] Chen W, Cai F, Chen H, et al. Joint neural collaborative filtering for recommender systems [J]. ACM Transactions on Information Systems (TOIS), 2019, 37(4): 1-30.
- [10] Hochreiter S, Schmidhuber J. Long short-term memory [J]. Neural Computation, 1997, 9(8): 1735-1780.
- [11] Hidasi B, Karatzoglou A, Baltrunas L, et al. Session-based recommendations with recurrent neural networks [C]//Proceedings of the 4th International Conference on Learning Representations, 2016, doi:10.48550/arXiv.1511.06939.
- [12] Jannach D, Ludewig M. When recurrent neural networks meet the neighborhood for session-based recommendation [C]//Proceedings of the 11th ACM Conference on Recommender Systems, 2017: 306-310.
- [13] Tan Y K, Xu X, Liu Y. Improved recurrent neural networks for session-based recommendations [C]//Proceedings of the 1st Workshop on Deep Learning for Recommender Systems, 2016: 17-22.
- [14] Peng D, Yuan W, Liu C. HARSAM: a hybrid model for recommendation supported by self-attention mechanism [J]. IEEE Access,

- 2019,7:12620-12629,doi:10.1109/ACCESS.2019.2892565.
- [15] Li J, Ren P, Chen Z, et al. Neural attentive session-based recommendation[C]//Proceedings of the ACM on Conference on Information and Knowledge Management,2017:1419-1428.
- [16] Liu Q, Zeng Y, 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,2018:1831-1839.
- [17] 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.
- [18] Wang S, Hu L, Wang Y, et al. Sequential recommender systems: challenges, progress and prospects [C]//28th International Joint Conference on Artificial Intelligence(IJCAI),2019:6332-6338.
- [19] Zhou J, Cui G, Hu S, et al. Graph neural networks: a review of methods and applications [J]. AI Open,2020,1:57-81,doi:10.1017/9781108924184.009.
- [20] Yu J, Yin H, Li J, et al. Enhance social recommendation with adversarial graph convolutional networks[J]. IEEE Transactions on Knowledge & Data Engineering, 2020, (1), doi: 10.1145/3397271.3401198.
- [21] Qiu R, Huang Z, Li J, et al. Exploiting cross-session information for session-based recommendation with graph neural networks[J]. ACM Transactions on Information Systems( TOIS ),2020,38(3):1-23.
- [22] BAI B, LIU Y T, MA C C, et al. Graph neural network[J]. Science in China: Series A,2020,50(3):367-384.
- [23] 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.
- [24] Xu C, Zhao P, Liu Y, et al. Graph contextualized self-attention network for session-based recommendation [C]//International Joint Conferences on Artificial Intelligence Organization, 2019:3940-3946.
- [25] Qiu R, Li 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,2019:579-588.
- [26] Wu S, Zhang Y, Gao C, et al. Garg: anonymous recommendation of point-of-interest in mobile networks by graph convolution network [J]. Data Science and Engineering,2020,5(4):433-447.
- [27] Chen T, Wong R C W. Handling information loss of graph neural networks for session-based recommendation [C]//Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining,2020:1172-1180.
- [28] Feng Y, You H, Zhang Z, et al. Hypergraph neural networks [C]//Proceedings of the AAAI Conference on Artificial Intelligence, 2019,33(1):3558-3565.

#### 附中文参考文献:

- [1] 曾义夫,牟其林,周乐,等.基于图表示学习的会话感知推荐模型[J].计算机研究与发展,2020,57(3):590-603.
- [22] 白铂,刘玉婷,马驰骋,等.图神经网络[J].中国科学:数学,2020,50(3):367-384.