

基于注意力机制的自监督图卷积会话推荐^①

吴永庆, 朱月, 王钰涵

(辽宁工程技术大学 软件学院, 葫芦岛 125105)

通信作者: 吴永庆, E-mail: yqwuyuwu@163.com



摘要: 为了解决会话推荐系统中数据稀疏性问题, 提出了一种基于注意力机制的自监督图卷积会话推荐模型 (self-supervised graph convolution session recommendation based on attention mechanism, ATSGCN). 该模型将会话序列构建成 3 个不同的视图: 超图视图、项目视图和会话视图, 显示会话的高阶和低阶连接关系; 其次, 超图视图使用超图卷积网络来捕获会话中项目之间的高阶成对关系, 项目视图和会话视图分别使用图卷积网络和注意力机制来捕获项目和会话级别局部数据中的低阶连接信息; 最后, 通过自监督学习使两个编码器学习到的会话表示之间的互信息最大化, 从而有效提升推荐性能. 在 Nowplaying 和 Diginetica 两个公开数据集上进行对比实验, 实验结果表明, 所提模型性能优于基线模型.

关键词: 会话推荐; 自监督学习; 图神经网络; 注意力机制

引用格式: 吴永庆, 朱月, 王钰涵. 基于注意力机制的自监督图卷积会话推荐. 计算机系统应用, 2024, 33(5): 57-66. <http://www.c-s-a.org.cn/1003-3254/9483.html>

Self-supervised Graph Convolution Session Recommendation Based on Attention Mechanism

WU Yong-Qing, ZHU Yue, WANG Yu-Han

(Software College, Liaoning Technical University, Huludao 125105, China)

Abstract: To address the challenge of data sparsity within session recommendation systems, this study introduces a self-supervised graph convolution session recommendation model based on the attention mechanism (ATSGCN). The model constructs the session sequence into three distinct views: the hypergraph view, item view, and session view, showing the high-order and low-order connection relationships of the session. Secondly, the hypergraph view employs hypergraph convolutional networks to capture higher-order pairwise relationships among items within a conversation. The item view and session view employ graph convolutional networks and attention mechanisms respectively to capture lower-order connection details within local conversation data at both item and session levels. Finally, self-supervised learning is adopted to maximize the mutual information between the session representations learned by the two encoders, thereby effectively improving recommendation performance. Comparative experiment on the Nowplaying and Diginetica public datasets demonstrates the superior performance of the proposed model over the baseline model.

Key words: session recommendation; self-supervised learning; graph neural network (GNN); attention mechanism

随着互联网信息量不断爆炸式增长, 网络上的信息变得更加纷繁复杂, 人们对信息的需求也更精细化, 推荐系统 (recommendation system, RS)^[1,2]可以帮助人

们从大量的信息中迅速获取有用信息, 缓解信息过载^[3]对他们的影响. 在许多应用场景中, 用户的明确交互信息无法获取, 从而难以对用户进行准确的推荐, 传统的

① 基金项目: 国家自然科学基金 (52174184)

收稿时间: 2023-11-05; 修改时间: 2023-12-04; 采用时间: 2023-12-20; csa 在线出版时间: 2024-03-22

CNKI 网络首发时间: 2024-03-27

推荐方法主要依靠用户的偏好文件或长期的历史交互信息来进行推荐,但是在现实的音乐推荐或电商平台推荐场景中,有时由于用户通过游客登录或隐私政策的限制等,推荐系统只能通过当前会话信息进行推荐,难以获得用户身份信息和用户的历史交互记录,无法为用户推荐准确的感兴趣的商品或内容.因此基于会话的推荐系统(session-based recommendation system, SBRS)被提出,基于会话的推荐旨在仅通过会话中产生的短期用户交互数据为匿名用户预测下一个最有可能交互的项目.

早期的会话推荐使用基于马尔可夫链的方法,它根据用户上一个行为去预测用户的下一个行为,具有很强的独立性假设,预测精度有限,使得基于马尔可夫链的模型^[4]无法利用长序列行为的依赖性,可能存在数据稀疏性问题.基于协同过滤的推荐^[5,6]利用物品之间的相似性来进行推荐,但忽略了用户的最新偏好,从而限制了推荐的准确性.

近几年,深度学习技术不断发展,许多基于神经网络的方法被提出应用于会话推荐中.一些研究^[7-9]将循环神经网络(recurrent neural network, RNN)应用于基于会话的推荐,取得了比传统推荐方法更好的性能,然而,基于神经网络的方法很难从远距离学习依赖关系.图神经网络逐渐在会话推荐领域成为主流,它将基于会话的序列数据建模为图以关注项目之间的转换关系.但是,现有基于图神经网络的会话推荐算法大多仅针对当前会话提取用户偏好,不能从多维视角提取会话间的复杂高阶成对关系,忽略了跨会话信息对当前会话重要性的影响,导致无法充分挖掘用户偏好.此外,由于会话推荐所采用的短时交互序列包含的信息非常有限,其更容易受到数据稀疏性的影响.

自监督学习(self-supervised learning, SSL)^[10]可以挖掘数据自身的表征特性作为监督信息,提高模型的特征提取能力,解决数据稀疏问题.因此,本文提出了一种基于注意力机制的自监督图卷积会话推荐模型(self-supervised graph convolution session recommendation based on attention mechanism, ATSGCN),将会话推荐与自监督学习相结合来挖掘原始数据中更有价值的信息,以缓解数据稀疏性问题对推荐性能的影响,进而实现推荐性能的提升,本文主要贡献如下.

(1) 将会话序列构建成3种视图:超图视图、项目视图和会话视图,显示会话的高阶和低阶信息,设计了

两种不同的编码器:超图视图通过超图卷积网络获得全局高阶会话表示;项目视图和会话视图使用图卷积网络和注意力机制获得局部低阶会话表示.

(2) 在推荐任务中加入自监督学习,使学习到的两个会话表示互相获取新信息,产生信息量更多的自监督信号,从而弥补短时交互序列数据稀疏性带来的影响.

(3) 通过在两个真实数据集上进行对比、消融和超参数敏感性实验,实验结果表明,所提出的模型较基线模型取得了一定提升.

1 相关工作

1.1 基于会话的推荐

会话推荐在早期主要采用马尔可夫链来捕获项目交互过程中的序列信息. Rendle 等人^[11]提出了 FPMC 模型,将矩阵分解与马尔可夫链相结合,同时捕获时间信息和长期用户喜好信息. Le 等人^[12]提出了 SEQ* 模型,将马尔可夫链与会话的上下文信息相融合,设计一种隐藏序列偏好的马尔可夫模型. Yang 等人^[13]使用马尔可夫模型进行个性化潜在路线推荐. Chen 等人^[14]将兴趣遗忘特性与马尔可夫模型结合. Wu 等人^[15]提出马尔可夫嵌入模型 PHR,将用户和物品映射到欧几里得空间,根据嵌入结果进行推荐排序.然而基于马尔可夫链的方法忽略了序列信息间的转移模式,难以捕获更高阶的复杂序列,忽略了交互作用整体间的时间依赖性.

随着深度学习的蓬勃发展,深度学习技术被应用于基于会话的推荐模型中,以捕捉项目之间的顺序. Hidasi 等人^[16]提出了 GRU4Rec 模型,首次将循环神经网络 RNN 应用到会话推荐中,并对 GRU 单元进行了优化. Li 等人^[17]提出了 NARM 模型,采用全局编码器和局部编码器捕获用户的主要目的,它同时对用户的顺序行为和主要目的进行建模. Liu 等人^[18]提出了 STAMP 模型,使用新的注意力机制捕获会话中用户的长短期兴趣. Kang 等人^[19]提出了 SASRec 模型,使用多个自注意层来建模用户的历史行为信息. Wang 等人^[20]提出了 CSRMM 模型,将协作邻域信息应用到会话推荐中.

尽管基于深度学习的方法取得了巨大的成功,但基于 RNN 的方法^[21,22]过度依赖相邻项目之间的顺序依赖关系,难以从长距离学习依赖关系.为更好地捕捉项目之间的复杂转移关系,基于图神经网络(graph neural network, GNN)的方法应用于会话推荐中,并显示出良

好的效果^[23]. SR-GNN^[24]将会话数据建模成图结构. 使用 GNN 来捕获复杂的项目转换, 并通过注意机制将会话的局部兴趣和全局偏好组合起来. Yu 等人^[25]提出了 TAGNN 模型, 提出了一种目标感知的注意力捕捉不同项目的不同优先级. Xu 等人^[26]提出 GC-SAN 模型, 将图神经网络与自注意力机制相结合, 增强局部特征的提取能力. Chen 等人^[27]提出 LESSR 模型, 将会话序列的点击顺序作为权重赋值给边, 使图结构能够保留会话的顺序信息, 又使用无权重的捷径图辅助捕获远程依赖, 取得了很好的效果. Qiu 等人^[28]设计了全图神经网络 FGNN, 其中 GAT 用于建模项目之间的转移关系, 后使用读出函数来生成会话表示. Wang 等人^[29]提出 GCE-GNN 模型, 构造全局图和会话图得到所有会话中的项目转移. Pan 等人^[30]提出 SGNN-HN 模型, 利用星形图神经网络来模拟会话中项目间的复杂过渡关系, 并使用高速公路网缓解过拟合问题. Yi 等人^[31]提出邻域增强图转移网络, 对当前会话与其相邻会话之间的复杂交互进行建模. Guo 等人^[32]提出 MIHSG 模型, 构造多粒度意图单元异构会话图从全局角度捕捉完整会话意图. Xie 等人^[33]提出了 PSR-GAT 模型, 它考虑了会话中项目间的连接信息与用户在项目上停留的时间, 能够出色的处理复杂的序列数据. 尽管 GNN 方法取得了良好的效果, 但上述方法主要基于简单图进行构建, 只能捕捉会话内的低阶项目成对信息, 无法捕获会话间的复杂高阶交互关系.

1.2 自监督学习

自监督学习是一种新兴的学习范式, 旨在让模型从原始数据中学习, 提供丰富的自监督信号. 近年来, 一些研究将自监督学习引入到推荐系统中. COTREC^[34]使用两个非对称图编码器, 通过自监督协同训练获得更多有用的信息. S²-DHCN^[35]通过超图神经网络 (hypergraph neural network, HGNN)^[36]学习项目嵌入表示, 利用对比学习^[37]实现两个通道学习到的会话表示之间的互信息最大化. 根据这一研究路线, 本文设计一个 ATSGCN 模型, 将自监督学习与会话推荐模型相结合, 有效缓解了数据稀疏性问题, 增强了推荐效果^[38].

2 ATSGCN 模型

2.1 符号及定义

设项目集合为 $I = \{i_1, i_2, i_3, \dots, i_N\}$, N 为项目数, 每个项目 $i \in I$ 嵌入到同一空间中, $x_i^{(l)} \in R^{d^l}$ 表示在深度

神经网络第 l 层维数 d^l 的项目 i 的向量表示. 每个会话表示为一个集合 $S = [i_{s,1}, i_{s,2}, i_{s,3}, \dots, i_{s,m}], i_{s,k} \in I (1 \leq k \leq m)$ 表示会话中匿名用户的交互项目. ATSGCN 模型的任务是预测给定的会话 s 的下一个交互项目 $i_{s,m+1}$. 将会话序列 s 输入到模型中, 输出得到可能项目的推荐得分 $y = [y_1, y_2, y_3, \dots, y_N]$, 并选取 y 中得分最高的前 k 项 ($1 \leq k \leq N$) 作为会话 s 推荐项目.

定义 1. 超图. 设 $G_h = (V, E)$ 表示一个超图, 其中 V 是一个包含 N 个唯一顶点的集合, E 是一个包含 M 个超边的集合. 每个超边 $\theta \in E$ 包含两个或多个顶点, 并分配一个正权值 $W_{\theta\theta}$, 所有权值构成一个对角矩阵 $W \in R^{M \times M}$. 超图可以用关联矩阵 $H \in R^{N \times M}$ 表示, 对于每个顶点和上边, 其度 D_{ii} 和 $B_{\theta\theta}$ 分别定义为 $D_{ii} = \sum_{\theta=1}^M W_{\theta\theta} H_{i\theta}$; $B_{\theta\theta} = \sum_{i=1}^N H_{i\theta}$, 且 D 和 B 为对角矩阵.

定义 2. 项目图. 设 $G_i = (V, E)$ 表示项目图, 其中 V 表示项目顶点集合, E 表示边集. 在一个会话中连接的任意两个项目 i_a 和 i_b 在项目视图中以有向边 E_{ab} 作为节点连接, 且每条边均带有正权重系数.

定义 3. 会话图. 设 $G_s = (S, E)$ 表示会话图, S 是包含所有会话的节点集, E 为边集. 两个会话连接时代表会话间存在相同物品, 每条边带有一个正权重系数.

2.2 模型描述

本文提出了基于注意力机制的自监督图卷积会话推荐模型 ATSGCN, 整体框架如图 1 所示, 模型分为两部分: 全局表示学习模块和局部表示学习模块. 首先将会话数据构建为超图视图、项目视图、会话视图 3 种不同视图. 通过从不同角度处理会话数据, 能够更全面的理解用户与项目的交互行为, 并提供额外的上下文信息. 其次, 全局表示学习模块中, 超图视图能够捕捉到丰富的语义信息和复杂的关系模式, 通过超图卷积神经网络有效地捕捉超图节点间的复杂关系和重要信息, 通过软注意力机制动态地分配权重给超图中的不同节点, 从而获得更准确的全局高阶会话表示; 局部表示学习模块中, 项目视图和会话视图经过图卷积网络从不同角度进行特征提取, 得到会话中的局部语义信息, 再进一步使用注意力机制对不同节点和边的重要性进行动态分配, 利用会话和项目之间的关系, 捕捉会话中项目之间更多的关联信息和重要性, 从而获得局部低阶会话表示. 最后, 将全局会话表示与局部会话表示进行

自监督学习,学习到更多的语义信息和上下文信息,进一步进行了数据增强,缓解了由于数据稀疏性导致的信息不足问题.通过两个模块的有效协作,联合优化会

话推荐任务与对比学习任务,得到最终的推荐列表,有效地缓解了数据稀疏性问题对推荐效果的影响,提高会话推荐的准确性和模型的鲁棒性与适应性.

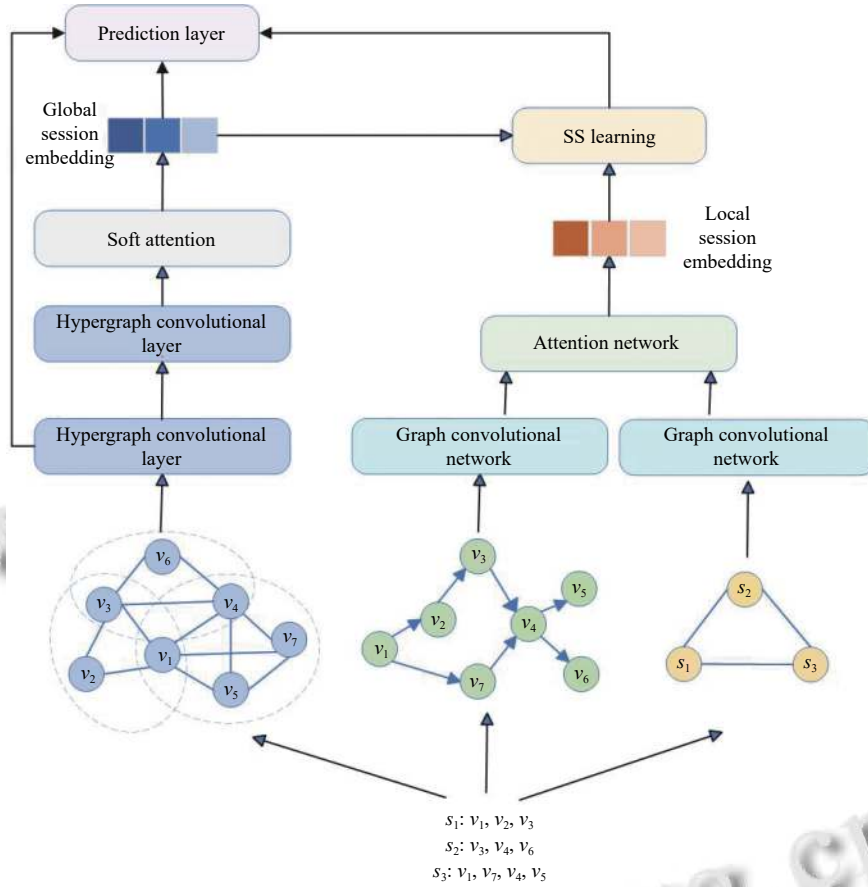


图1 ATSGCN 模型框架

2.2.1 全局表示学习模块

全局表示学习模块由超图经过超图卷积和软注意力模块得到全局高阶会话表示.超图卷积在信息传播过程中以顶点-超边-顶点的方式实现特征转换.超图卷积定义如下:

$$X_h^{(l+1)} = D^{-1}HWB^{-1}H^T X_h^{(l)} \quad (1)$$

其中, $H^T X_h^{(l)}$ 表示从节点到超边的信息聚合,与关联矩阵 H 相乘实现信息由超边聚合到节点.将卷积后得到的嵌入表示进行平均得到全局嵌入表示:

$$X_h = \frac{1}{L+1} \sum_{l=0}^L X_h^{(l)} \quad (2)$$

在会话序列中物品的位置信息^[39]也是关键影响因素,因此本文引入一个可学习的位置矩阵 $P = [p_1, p_2,$

$p_3, \dots, p_m]$, m 为当前会话长度,将反向位置嵌入与全局嵌入连接,在会话中第 k 个物品表示为:

$$x'_k = \tanh(W_1[x_k || P_{m-k+1}] + b) \quad (3)$$

其中, $W_1 \in R^{d \times 2d}$ 和 $b \in R^d$ 为可学参数.在学习会话嵌入时,本文遵循文献[35]的方法,使用软注意机制得到全局高阶会话表示:

$$s_h = \sum_{k=1}^m \alpha_k x'_k \quad (4)$$

其中,项目权重 $\alpha_k = f^T \sigma(W_2 x_s + W_3 x'_k + c)$, $f \in R^d$, x_s 为当前会话嵌入, $W_2 \in R^{d \times d}$, $W_3 \in R^{d \times d}$ 为可学习参数矩阵.

2.2.2 局部表示学习模块

局部表示学习模块由项目图和会话图两种视图通

过图卷积得到项目级和会话级两个级别的嵌入表示, 将两个嵌入利用注意力网络融合, 通过软注意力机制学习得到局部会话表示。

项目视图捕获项目级连接信息, 经过图卷积层生成项目级嵌入, 项目图卷积定义为:

$$X_I^{(l+1)} = \hat{D}^{-1} \hat{A} X_I^{(l)} W_I^{(l)} \quad (5)$$

其中, $\hat{A}_I = A_I + I$, I 为单位矩阵, A_I 为邻接矩阵, $W_I^{(l)}$ 为参数矩阵. 将每一层图卷积得到的嵌入信息进行平均, 即最终的项目级嵌入:

$$X_I = \frac{1}{L+1} \sum_{l=0}^L X_I^{(l)} \quad (6)$$

会话视图构建会话间结构描述项目与会话关系, 对会话图进行图卷积, 会话图卷积定义为:

$$X_S^{(l+1)} = \hat{D}_S^{-1} \hat{A}_S X_S^{(l)} W_S^{(l)} \quad (7)$$

其中, $\hat{A}_S = A_S + I$, A_S 和 \hat{D}_S 分别表示邻接矩阵和度矩阵, $W_S^{(l)}$ 是参数矩阵. 学习到的会话级嵌入表示为:

$$X_S = \frac{1}{L+1} \sum_{l=0}^L X_S^{(l)} \quad (8)$$

对学习到的项目级、会话级嵌入使用注意机制得到最终项目嵌入, 计算项目 v_i 在这两方面的相对重要性:

$$\mu_I^i = q_1^T \tanh(W_4 X_I + b_1) \quad (9)$$

$$\mu_S^i = q_1^T \tanh(W_4 X_S + b_1) \quad (10)$$

其中, $q_1 \in \mathbb{R}^d$ 为注意向量, $W_4 \in \mathbb{R}^{d \times d}$ 为权重矩阵, $b_1 \in \mathbb{R}^d$ 为偏差向量. X_I 表示项目级嵌入, X_S 表示会话级嵌入. 最终项目嵌入表示为:

$$f_i = \alpha_I^i X_I + \alpha_S^i X_S \quad (11)$$

其中, α_I^i 和 α_S^i 为权重系数, 将最终项目嵌入通过软注意力机制得到最终的局部低阶会话表示:

$$u_i = \sum_{l=1}^{N_i} \beta_{i,l} f_{i,l} \quad (12)$$

其中, $\beta_{i,l}$ 为每个项目的权重系数.

2.2.3 自监督学习

本文使用对比学习进行自监督任务, 通过对比两个表示学习模块得到的会话嵌入进行数据增强, 采用 In-foNCE^[40] 将二元交叉熵损失作为本文的学习目标,

其定义为:

$$L_s = -\log \sigma(f_D(\theta_i^h, \theta_i^u)) - \log \sigma(1 - f_D(\tilde{\theta}_i^h, \theta_i^u)) \quad (13)$$

其中, f_D 为鉴别函数, 对输入向量间的一致性进行评分, 最大化两个不同会话嵌入间的互信息. $\tilde{\theta}_i^h$ 是 θ_i^h 行列变换顺序得到的负样本.

2.2.4 模型预测

将全局表示模块得到的项目嵌入 x_h 与 s_h 内积运算, 再使用 Softmax 函数计算下一个项目的预测概率:

$$\hat{y} = \text{Softmax}(s_h^T x_h) \quad (14)$$

在训练过程中, 损失函数采用交叉熵损失函数:

$$L_r = -\sum_{i=1}^N y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i) \quad (15)$$

其中, y 是 one-hot 向量, 最后推荐任务与自监督学习结合得到总损失:

$$L = L_r + \beta L_s \quad (16)$$

其中, β 为调整自监督任务大小的超参数.

3 实验与分析

3.1 数据集

本文使用 Nowplaying^[41] 和 Diginetica^[42] 两个真实数据集来评估 ATSGCN 模型. Nowplaying 数据集描述了用户的音乐收听行为. Diginetica 数据集来自 2016 年 CIKM 杯, 包含了一个电子商务网站 5 个月内的点击流数据. 为了比较的公平性, 本文参照文献[35]中的工作对数据集进行预处理, 两个数据集的统计数据如表 1 所示.

表 1 数据集统计信息

统计项目	Nowplaying	Diginetica
训练会话	825 304	719 470
测试会话	89 824	60 858
项目数量	60 417	43 097
平均长度	7.42	5.12

3.2 基线模型

为验证模型的有效性, 本文将 ATSGCN 模型与下面的会话推荐模型进行比较.

(1) FPMC^[11]: 是一种基于马尔可夫链的序列方法. 为了适应基于会话的推荐, 我们在计算推荐分数时不考虑用户潜在表示.

(2) GRU4Rec^[16]: 利用会话并行小批量训练过程,

采用基于排序的损失函数对用户序列进行建模。

(3) NARM^[17]: 基于 RNN 的最先进的模型, 利用注意力机制捕捉用户的主要目的, 并结合顺序行为生成推荐。

(4) STAMP^[18]: 采用注意层替代前面工作中的所有 RNN 编码器, 并采用自注意机制来提高基于会话的推荐性能。

(5) SR-GNN^[24]: 应用门控图卷积层获取项嵌入, 同时采用软注意机制计算会话嵌入。

(6) FGNN^[28]: 使用加权注意层学习下一项目推荐的物品和会话的嵌入。

(7) S²-DHCN^[35]: 构造两种超图学习会话间信息和会话内信息, 并利用自监督学习增强基于会话的推荐。

3.3 评价指标

本文使用 $P@K$ (准确率) 和 $MRR@K$ (平均倒数排名) 来评价推荐结果。

$P@K$ (准确率): $P@K$ 被用作预测精度的衡量标准。它代表了前 K 个项目中正确推荐项目的比例。

$$P@K = \frac{n_{hit}}{N} \quad (17)$$

其中, N 为总样本数, n_{hit} 为排名前 K 项的项目中正确推荐的项目个数。

$MRR@K$ (平均倒数排名): $MRR@K$ 强调正确项目的排名。用来衡量推荐排名的顺序。

$$MRR@K = \frac{1}{N} \sum_{i=0}^N \frac{1}{rank(i)} \quad (18)$$

其中, N 是测试集样本的总数, $rank(i)$ 表示物品 i 在生成的推荐列表中的排名。

3.4 实验参数设置

本文模型将嵌入大小和批处理大小设置为 100, L2 正则化的权重设置为 10^{-5} 。使用标准差为 0.1 的高斯分布初始化参数, 并将初始学习率设为 0.001, 采用 Adam 作为优化器, 训练迭代次数设定为 30, 为比较的公平性, 所有实验均保持上述参数设置在所有模型中一致。

3.5 实验结果分析

3.5.1 对比实验

本文模型与基线模型在两个公开数据集上的实验结果如表 2 所示。根据实验结果可以看出, 基于 RNN 的推荐模型 GRU4Rec、NARM、STAMP 的推荐性能

优于传统推荐方法 FPMC, 说明了项目间的顺序依赖关系在会话推荐中的重要作用。NARM、STAMP 模型中使用注意力机制学习会话中项目的重要程度, 推荐性能优于 GRU4Rec, 说明使用注意力机制有助于更准确获取会话中用户的兴趣偏好。基于 GNN 的推荐模型 SR-GNN、FGNN、S²-DHCN 利用 GNN 进一步捕获会话中复杂的项目转换, 其推荐性能优于基于 RNN 的推荐模型, 说明了将会话序列建模为图可以更好地捕获项目间的成对转换关系。S²-DHCN 在会话中引入自监督任务, 并将会话序列建模为超图, 捕获项目间的复杂高阶关系, 其性能优于 SR-GNN、FGNN。实验结果显示, 基于注意力机制的自监督图卷积会话推荐模型 ATSGCN 性能均优于其他基线模型, 在数据集 Nowplaying 和 Diginetica 上, ATSGCN 与基线模型中性能最好的 S²-DHCN 相比, 在 $P@20$ 上分别提高了 1.17%、0.75%, 在 $MRR@20$ 上分别提高了 0.96%、1.98%。这表明将会话建模为多个视图分别捕捉不同级别的转换关系能更准确地捕获用户意图, 有效提升推荐性能。

表 2 不同模型在两个数据集上的实验结果 (%)

Model	Nowplaying		Diginetica	
	$P@20$	$MRR@20$	$P@20$	$MRR@20$
FPMC	7.36	2.82	26.53	6.95
GRU4Rec	7.92	4.48	29.45	8.33
NARM	18.59	6.93	49.70	16.17
STAMP	17.66	6.88	45.64	14.32
SR-GNN	18.87	7.47	50.73	17.59
FGNN	18.78	7.15	50.58	16.84
S ² -DHCN	22.98	8.04	53.08	18.18
ATSGCN	23.25	8.10	53.59	18.54

3.5.2 消融实验

为研究 ATSGCN 模型中各个模块的有效性, 设计了 4 个不同变体进行消融实验, 其中 ATSGCN-H 删除了超图编码器模块, ATSGCN-I 删除了项目视图模块, ATSGCN-S 删除了会话视图模块, ATSGCN-L 删除了自监督学习任务。实验结果如图 2、图 3, ATSGCN 与其他变体模型相比, 始终保持最佳推荐性能, 这表明模型的每个组件都是必要的。

3.6 参数敏感性分析

3.6.1 自监督学习参数影响

为了研究加入自监督学习对会话推荐模型的帮助和有效性, 本文使用一个自监督学习超参数 β 来控制自监督学习的程度, 保持其他参数不变, 将超参数 β 取

值范围设置为 $\{0, 0.001, 0.002, 0.005, 0.01, 0.02, 0.05\}$, 实验结果如图4、图5所示, 根据实验结果可以看出, 在 Nowplaying 数据集上, β 为 0.001 时性能最佳, 在

Diginetica 数据集上, β 为 0.01 时达到最佳性能. 当模型性能达到峰值后, 随着 β 的增大, 模型性能有所下降. 这表明较大的 β 值会给推荐任务带来负面影响.

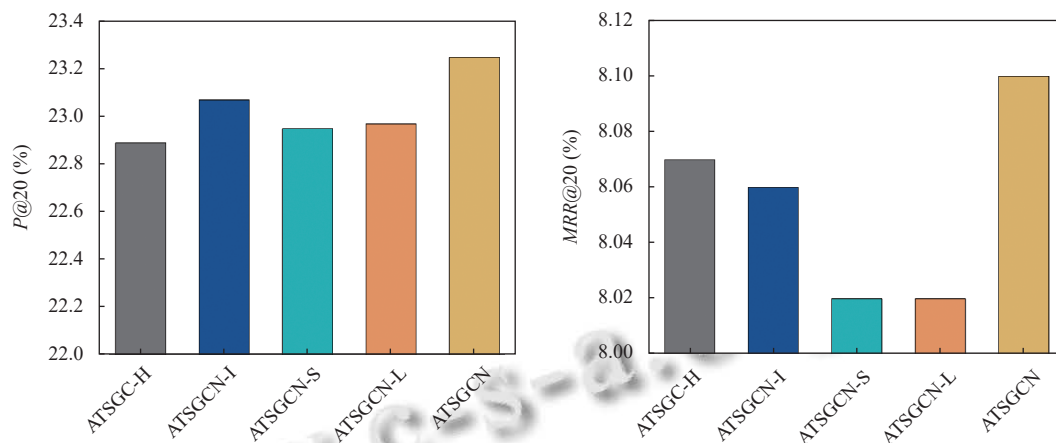


图2 数据集 Nowplaying 中的消融实验

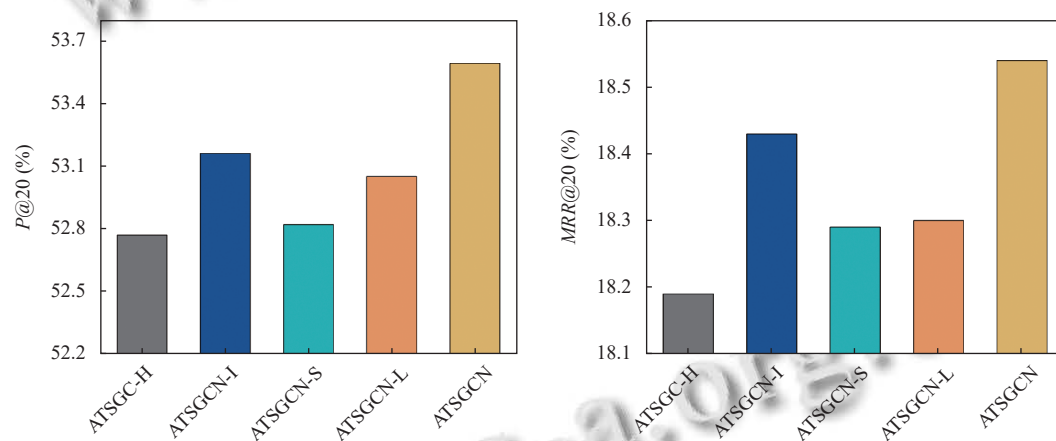


图3 数据集 Diginetica 中的消融实验

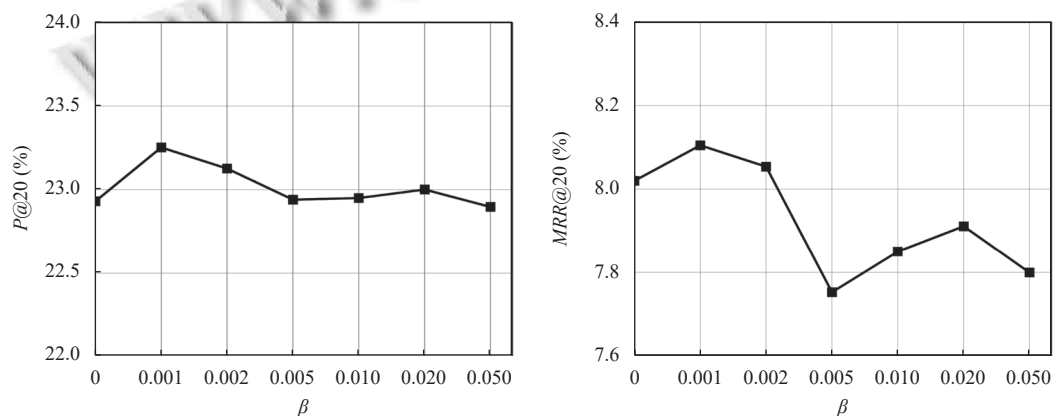


图4 Nowplaying 中自监督学习参数大小的影响

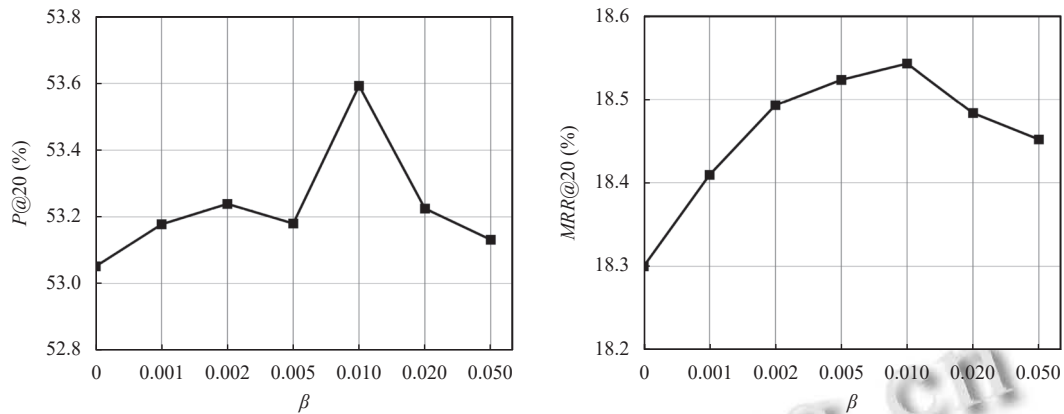


图5 Diginetica 中自监督学习参数大小的影响

3.6.2 模型深度影响

为了研究图卷积网络深度对模型推荐性能的影响,保持其他参数不变,本文将模型层数取值范围设置为{1,

2, 3, 4, 5}. 实验结果如图6、图7所示,在 Nowplaying 和 Diginetica 两个数据集上,模型层数设置为3层性能最好.随着层数的增加,会出现过平滑问题,导致性能逐渐下降.

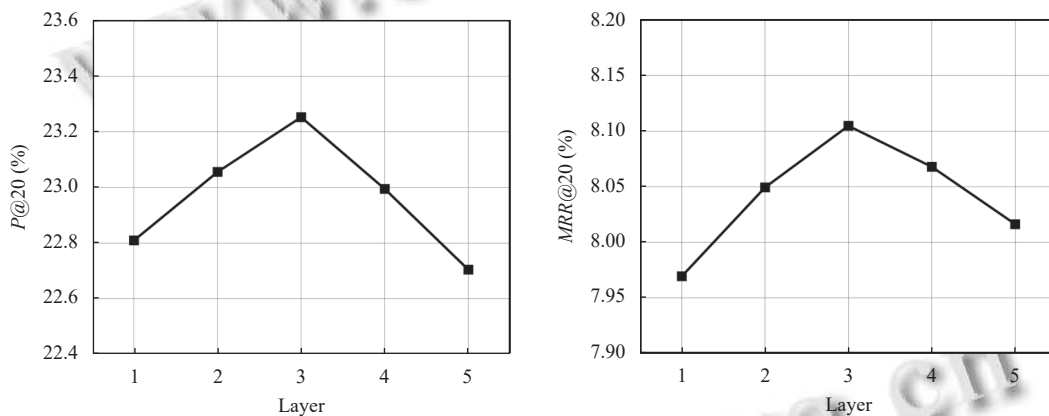


图6 Nowplaying 中模型深度的影响

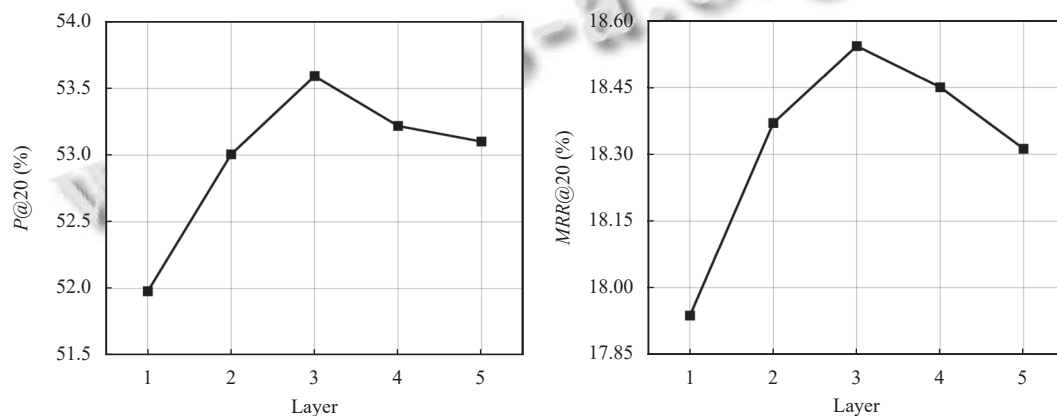


图7 Diginetica 中模型深度的影响

4 结束语

为了解决会话推荐系统的数据稀疏等问题,本文提出了基于注意力机制的自监督图卷积会话推荐模型

(ATSGCN),模型构建了3种视图利用图卷积和注意力机制从全局表示学习模块和局部表示学习模块学习不同级别的物品间的转换关系,更全面准确地捕获用户

的真实兴趣偏好,并通过在模型中加入自监督任务产生自监督信号实现数据增强,缓解了会话推荐系统存在的数据稀疏性问题.与基线模型进行对比实验,实验结果证明了本文模型的有效性.在今后的研究中,将考虑设计一个更有效的辅助任务来提升推荐性能.

参考文献

- 1 倪美玉,曹为刚.基于深度神经网络的个性化混合商品推荐模型.计算机系统应用,2021,30(5):184-189.[doi:10.15888/j.cnki.csa.007898]
- 2 张栩翔,汤玉祺,赵文,等.基于知识图谱和图嵌入的个性化学习资源推荐.计算机系统应用,2023,32(5):180-187.[doi:10.15888/j.cnki.csa.009068]
- 3 刘君良,李晓光.个性化推荐系统技术进展.计算机科学,2020,47(7):47-55.
- 4 Jannach D, Ludewig M. When recurrent neural networks meet the neighborhood for session-based recommendation. Proceedings of the 11th ACM Conference on Recommender Systems. Como: ACM Press, 2017. 306-310.
- 5 Chen H, Qian FL, Chen J, et al. Attribute-based neural collaborative filtering. Expert Systems with Applications, 2021, 185: 115539. [doi:10.1016/j.eswa.2021.115539]
- 6 刘锦涛,谢颖华.基于图注意力网络表示学习的协同过滤推荐算法.计算机系统应用,2022,31(4):273-280.[doi:10.15888/j.cnki.csa.008443]
- 7 Hidasi B, Karatzoglou A, Baltrunas L, et al. Session-based recommendations with recurrent neural networks. Proceedings of the 4th International Conference on Learning Representations. San Juan: ICLR Press, 2016. 1-10.
- 8 Hidasi B, Quadrana M, Karatzoglou A, et al. Parallel recurrent neural network architectures for feature-rich session-based recommendations. Proceedings of the 10th ACM Conference on Recommender Systems. Boston: ACM, 2016. 241-248.
- 9 Feng CQ, Shi CY, Liu CM, et al. Context-aware item attraction model for session-based recommendation. Expert Systems with Applications, 2021, 176: 114834. [doi:10.1016/j.eswa.2021.114834]
- 10 Liu X, Zhang FJ, Hou ZY, et al. Self-supervised learning: Generative or contrastive. IEEE Transactions on Knowledge and Data Engineering, 2023, 35(1): 857-876.
- 11 Rendle S, Freudenthaler C, Schmidt-Thieme L. Factorizing personalized Markov chains for next-basket recommendation. Proceedings of the 19th International Conference on World Wide Web. Raleigh: ACM Press, 2010. 811-820.
- 12 Le DT, Fang Y, Lauw HW. Modeling sequential preferences with dynamic user and context factors. Proceedings of the 2016 European Conference on Machine Learning and Knowledge Discovery in Databases. Riva del Garda: Springer International Publishing, 2016. 145-161.
- 13 Yang YD, Xiao P, Yao X, et al. PHR: A personalized hidden route recommendation system based on hidden Markov model. Proceedings of the 4th International Joint Conference on Web and Big Data. Tianjin: Springer International Publishing, 2020. 535-539.
- 14 Chen J, Wang CK, Wang JM. A personalized interest-forgetting Markov model for recommendations. Proceedings of the 29th AAAI Conference on Artificial Intelligence. Austin: AAAI Press, 2015. 16-22.
- 15 Wu X, Liu Q, Chen EH, et al. Personalized next-song recommendation in online Karaoke. Proceedings of the 7th ACM Conference on Recommender Systems. Hong Kong: ACM Press, 2013. 137-140.
- 16 Hidasi B, Karatzoglou A. Recurrent neural networks with top-k gains for session-based recommendations. Proceedings of the 27th ACM International Conference on Information and Knowledge Management. Torino: ACM, 2018. 843-852.
- 17 Li J, Ren PJ, Chen ZM, et al. Neural attentive session-based recommendation. Proceedings of the 2017 ACM on Conference on Information and Knowledge Management. Singapore: ACM Press, 2017. 1419-1428.
- 18 Liu Q, Zeng YF, Mokhosi R, et al. 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: ACM, 2018. 1831-1839.
- 19 Kang WC, McAuley J. Self-attentive sequential recommendation. Proceedings of the 2018 IEEE International Conference on Data Mining (ICDM). Singapore: IEEE, 2018. 197-206.
- 20 Wang MR, Ren PJ, Mei L, Chen ZM, et al. A collaborative session-based recommendation approach with parallel memory modules. Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval. Paris: ACM, 2019. 345-354.
- 21 Zhang ZZ, Wang B. Graph spring network and informative anchor selection for session-based recommendation. Neural Networks, 2023, 159: 43-56. [doi:10.1016/j.neunet.2022.12.003]
- 22 孟志鹏,成卫青.基于自注意力网络的时间感知序列化推

- 荐. 计算机系统应用, 2023, 32(1): 197–205. [doi: [10.15888/j.cnki.csa.008887](https://doi.org/10.15888/j.cnki.csa.008887)]
- 23 Chu F, Jia CY. Self-supervised global context graph neural network for session-based recommendation. *PeerJ Computer Science*, 2022, 8: e1055. [doi: [10.7717/peerj-cs.1055](https://doi.org/10.7717/peerj-cs.1055)]
- 24 Wu S, Tang YY, Zhu YQ, *et al.* Session-based recommendation with graph neural networks. *Proceedings of the 33rd AAAI Conference on Artificial Intelligence*. Honolulu: AAAI Press, 2019. 346–353.
- 25 Yu F, Zhu YQ, Liu Q, *et al.* TAGNN: Target attentive graph neural networks for session-based recommendation. *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*. New York: ACM Press, 2020. 1921–1924.
- 26 Xu CF, Zhao PP, Liu YC, *et al.* Graph contextualized self-attention network for session-based recommendation. *Proceedings of the 28th International Joint Conference on Artificial Intelligence*. Macao: IJCAI.org, 2019. 3940–3946.
- 27 Chen TW, Wong RCW. Handling information loss of graph neural networks for session-based recommendation. *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. New York: ACM, 2020. 1172–1180.
- 28 Qiu RH, Li JJ, Huang Z, *et al.* 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: ACM, 2019. 579–588.
- 29 Wang ZY, Wei W, Cong G, *et al.* 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*. New York: ACM Press, 2020. 169–178.
- 30 Pan ZQ, Cai F, Chen WY, *et al.* Star graph neural networks for session-based recommendation. *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*. New York: ACM, 2020. 1195–1204.
- 31 Yi ZJ, Song R, Li J, *et al.* Neighbor-enhanced graph transition network for session-based recommendation. *International Journal of Machine Learning and Cybernetics*, 2023, 14(4): 1317–1331. [doi: [10.1007/s13042-022-01702-8](https://doi.org/10.1007/s13042-022-01702-8)]
- 32 Guo JY, Yang YM, Song XC, *et al.* Learning multi-granularity consecutive user intent unit for session-based recommendation. *Proceedings of the 15th ACM International Conference on Web Search and Data Mining*. New York: ACM Press, 2022. 343–352.
- 33 Xie YQ, Li ZR, Qin T, *et al.* Personalized session-based recommendation using graph attention networks. *Proceedings of the 2021 International Joint Conference on Neural Networks (IJCNN)*. Shenzhen: IEEE, 2021. 1–8.
- 34 Xia X, Yin HZ, Yu JL, *et al.* Self-supervised graph co-training for session-based recommendation. *Proceedings of the 30th ACM International Conference on Information & Knowledge Management*. ACM, 2021. 2180–2190.
- 35 Xia X, Yin HZ, Yu JL, *et al.* Self-supervised hypergraph convolutional networks for session-based recommendation. *Proceedings of the 35th AAAI Conference on Artificial Intelligence*. Vancouver: AAAI Press, 2021. 4503–4511.
- 36 Feng YF, You HX, Zhang ZZ, *et al.* Hypergraph neural networks. *Proceedings of the 33rd AAAI Conference on Artificial Intelligence*. Honolulu: AAAI Press, 2019, 33 (1): 3558–3565.
- 37 Zhou K, Wang H, Zhao WX, *et al.* S³-Rec: Self-supervised learning for sequential recommendation with mutual information maximization. *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*. New York: ACM, 2020. 1893–1902.
- 38 王永贵, 赵晓暄. 结合自监督学习的图神经网络会话推荐. *计算机工程与应用*, 2023, 59(3): 244–252. [doi: [10.3778/j.issn.1002-8331.2204-0245](https://doi.org/10.3778/j.issn.1002-8331.2204-0245)]
- 39 Gan MX, Zhang XT. Integrating community interest and neighbor semantic for microblog recommendation. *International Journal of Web Services Research*, 2021, 18(2): 54–75. [doi: [10.4018/IJWSR.2021040104](https://doi.org/10.4018/IJWSR.2021040104)]
- 40 van den Oord A, Li YZ, Vinyals O. Representation learning with contrastive predictive coding. *arXiv:1807.03748*, 2018.
- 41 Zangerle E, Pichl M, Gassler W, *et al.* #Nowplaying music dataset: Extracting listening behavior from Twitter. *Proceedings of the 1st International Workshop on Internet-scale Multimedia Management*. Association for Computing Machinery. Orlando: ACM, 2014. 21–26.
- 42 Ren PJ, Chen ZM, Li J, *et al.* RepeatNet: A repeat aware neural recommendation machine for session-based recommendation. *Proceedings of the 33rd AAAI Conference on Artificial Intelligence*. Honolulu: AAAI Press, 2019. 4806–4813.

(校对责编: 孙君艳)