

基于会话的双视角社交推荐模型

李佳奇

南京邮电大学理学院, 江苏 南京

收稿日期: 2024年10月25日; 录用日期: 2024年11月14日; 发布日期: 2024年11月21日

摘要

现有的基于会话的社交推荐模型单一的通过图视角学习会话的嵌入表示并做出预测, 而忽略了项目之间的顺序依赖性。针对该问题, 提出了一种基于会话的双视角社交推荐模型(Dual-view approach for session-based social recommendation, DVSSR)。DVSSR分别从图视角以及行为序列视角学习会话嵌入。在会话图视角中, 通过图神经网络传播节点之间的特征信息, 通过注意力机制聚合会话嵌入表示; 在行为序列视角中, 使用自适应旋转位置编码学习项目之间的相对位置关系, 通过多头自注意力机制学习序列嵌入表示, 将两种视角下的嵌入表示融合并做出推荐。然后, 为了确保双重视角下学习到用户偏好的一致性, 设计对比学习模块拉近两种视角下学习嵌入表示, 并通过破坏生成增强视图与原视图对比学习提高模型的鲁棒性和泛化能力。最后, 在三个大型公开数据集上的实验结果验证了DVSSR的有效性和优越性。

关键词

基于会话推荐, 社交网络, 图神经网络, 位置编码, 自注意力机制, 对比学习, 数据增强

Dual-View Model for Session-Based Social Recommendation

Jiaqi Li

School of Science, Nanjing University of Posts and Telecommunications, Nanjing Jiangsu

Received: Oct. 25th, 2024; accepted: Nov. 14th, 2024; published: Nov. 21st, 2024

Abstract

Session-based social recommendation aims to predict the next most likely item of interaction for the current session based on the user's social network and historical session information. Existing session-based social recommendation methods predominantly learn session embeddings from a single graph perspective for making predictions, thereby neglecting the sequential dependencies

among items. To address this issue, we propose a Dual-View approach for Session-Based Social Recommendation (DVSSR). DVSSR learns user and item embeddings through a user-item heterogeneous graph and learns session embeddings from both graph and behavioral sequence perspectives. In the session graph perspective, it propagates feature information between nodes using a graph neural network and aggregates session embeddings via an attention mechanism. In the behavioral sequence perspective, it employs an adaptive rotary position encoder to learn the relative positional relationships among items and captures the sequential dependencies and sequence embeddings using a multi-head self-attention mechanism. The results from these two perspectives are then integrated with user embeddings to form the final session embedding for making predictions. To ensure consistency in user preference learning across the dual perspectives, a contrastive learning module is designed to align the learned embeddings from both views, and the method's robustness and generalization capability are enhanced through contrastive learning between augmented and original views. Finally, experimental results on three large public datasets validate the effectiveness and superiority of DVSSR.

Keywords

Session-Based Recommendation, Social Network, Graph Neural Network, Position Encoding, Self-Attention Mechanism, Contrastive Learning, Data Augmentation

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. 相关工作

进入二十一世纪后，信息技术的不断发展使信息过载日益严重，推荐系统的发展有效缓解该问题。会话推荐模型[1] (Session-based Recommendation, SR)是对用户的短期行为所形成的会话建模，并基于此预测用户的下一个行为。在现实世界中，当用户数量达到一定时，他们之间就会形成社交网络，例如 Facebook, Instagram 和微博等大型社交软件，用户之间可以通过点赞、关注、转发等行为项目连接并构建庞大的社交网络，而这些用户往往会有一定的联系，例如，关注同一用户的用户群体可能存在相似偏好，用户的偏好很可能受到其他用户的影响。因此，目前有研究基于这种特点，对社交网络背景下推荐进行研究，称为基于会话的社交推荐(Session-based Social Recommendation, SSR)。

在目前对 SSR 的研究中，DGRec [2]为了捕获社交网络中用户的相互影响关系，使用图神经网络(Graph Neural Network, GNN)聚合邻居节点的特征信息，使用户的偏好与社交关系联系起来。SERec [3]构建了一张由用户和项目组成的异构图，并通过异构图神经网络(Heterogeneous Graph Neural Network, HGNN)学习所有项目以及用户的嵌入表示，该模型设计了一种个性化会话嵌入模块，结合了当前会话的项目信息以及用户信息生成会话嵌入表示。HF²-Rec [4]在此基础上，通过图注意力网络(Graph Attention Network, GAT)，更加精确的学习到了蕴含用户偏好的会话嵌入，并设计出一个高阶特征融合模块，将用户向量、短期兴趣以及会话嵌入融合后做出推荐，这些模型都取得了良好的推荐性能。

本文在研究中发现，上述的基于会话的社交推荐模型都单一的从图这一种视角出发来解决 SSR 问题，即通过基于图的模型对会话中项目之间的共现关系建模，但这类模型在模拟会话中项目的顺序依赖关系上较弱，因为在图结构中，会话在改变项目出现的先后顺序后所构成的图结构可能与原会话一致。

为了应对上述提出的单一视角和相对位置编码问题，本文提出了一个新的基于会话的社交网络推荐模型：基于会话的双视角社交推荐模型(Dual-view approach for session-based social recommendation,

DVSSR)。该模型首先从项目 - 用户异构图中学习嵌入表示，然后分别从用户的会话图以及行为序列两种视角出发，在会话图视角中，使用 GNN 传播节点之间的特征信息，将节点信息聚合得到会话的嵌入表示。在行为序列视角中，使用一种自适应旋转位置编码，捕获项目之间的相对位置关系，更好地对项目的语义信息以及位置信息融合，通过多头自注意力机制学习行为序列的偏好信息。其次，将两种不同视角下学习到的会话偏好信息融合，作为最终的会话嵌入表示并做出推荐。最后，为了保证两种视角学习到的偏好信息的一致性，通过对比学习拉近两种视角下学习到的嵌入表示，为了提高模型的鲁棒性，对原视角破坏生成增强视角，将原视角下学到的嵌入表示与增强视角下学习到的嵌入表示进行对比学习，以提高抗噪能力。

基于会话的社交推荐从本质上讲，是通过对用户的社交关系建模，以提高会话推荐的准确性，因此，在此将对会话推荐模型以及基于会话的社交推荐模型的相关工作进行介绍。

1.1. 会话推荐模型

会话推荐的目标在于，给定用户产生的会话，对会话中项目之间的关系建模，预测出该用户最有可能交互的下一个项目。在神经网络技术得到迅猛发展后，GRU4Rec [5]是首个基于 RNN 架构用来解决会话推荐问题的模型，使用多层次门控递归单元来捕获项目之间的交互关系。NARM [6]和 STAMP [7]等模型在此基础上加入了注意力机制以更好地捕获用户兴趣。SSRM [8]使用矩阵分解对 GRU 进行增强捕获用户的长期兴趣，实现了基于会话流的推荐。Caser [9]指出基于 RNN 的模型可能无法很好地捕获之前几个项目对下一个项目的影响，并通过卷积神经网络学习序列模式，NextItNet [10]引入了具有残差结构的卷积神经网络，捕捉长期和短期的依赖性。由于基于 GNN 的模型能捕获项目之间的高阶交互关系，成为目前主流的推荐模型之一。SR-GNN [11]通过门控 GNN 学习会话图中的项目嵌入。GCE-GNN [12]通过构建全局会话图以缓解数据稀疏问题。上述模型都取得了良好的性能，推动了会话推荐领域的发展，但却不能直接应用于基于会话的社交推荐，因为它们没有考虑到社交网络中用户交互关系产生的影响。

1.2. 基于会话的社交推荐模型

为了充分考虑社交网络中，用户之间的交互行为对结果产生的影响，DGRec [2]通过堆叠 L 层 GNN 聚合邻居用户的特征信息。SERec [3]构建了一张由用户和项目组成的异构图，并通过 HGNN 学习所有项目以及用户的嵌入表示，设计了一种个性化会话嵌入模块，结合了当前会话的项目信息以及用户信息生成会话嵌入表示。HF²-Rec [4]在此基础上，通过 GAT 更加精确地学习到了蕴含用户偏好的会话嵌入，并设计出一个高阶特征融合模块，将用户向量、短期兴趣以及会话嵌入融合后做出推荐。

上述模型都在 SSR 领域中取得了令人满意的成就，但目前大多数的模型都单一的通过图这一视角研究会话中项目之间的共现关系以获得偏好信息，基于图的模型虽然在提取项目之间的高阶转换关系上具有优势，却缺乏建模会话中项目之间的顺序依赖性的能力，因为会话在改变项目出现的先后顺序后所构成的图结构可能与原会话一致。而基于行为序列的模型在此方面具有优势，因此 DVSSR 分别从会话图视角和行为序列视角分别对会话建模，从不同的角度学习会话中的偏好信息，优势互补。

2. 基于会话的双视角社交推荐模型

2.1. 问题阐述

在基于会话的社交网络推荐中，以 $V = \{v_1, v_2, v_3, \dots, V_{|V|}\}$ 和 $U = \{u_1, u_2, u_3, \dots, u_{|U|}\}$ 分别表示数据集中所有项目和用户集合，以 $S = \{s_1^u, s_2^u, s_3^u, \dots, s_{|S|}^u\}$ 表示所有用户的会话集合，集合 S 中的用户 u 的任意会话 s_j^u 都包含着若干项目即 $s_j^u = [v_1, v_2, v_3, \dots, v_n]$ ， n 为该会话的长度。用户之间的社交网络以图 $F = (U, E)$ 表示，

E 表示社交网络图 F 中边的集合，代表着用户之间的交互关系。SSR 的主要目标为，给定社交网络图 F 以及用户 u 的当前会话 $S_j^u = [v_1, v_2, v_3, \dots, v_n]$ ，预测出该用户最有可能交互的下一个项目。

2.2. 模型介绍

本文所提出 DVSSR 的工作流程如下：首先，将项目集合以及用户集合视为节点，构建异构图 K ， K 中的边集由项目 - 项目的交互、项目 - 用户的交互、用户 - 用户的交互组成，通过异构图神经网络学习 K 中所有项目节点以及用户节点的嵌入表示；然后，分别构建会话图视角和行为序列视角学习会话的嵌入表示，在会话图视角中，每一个节点通过 GNN 聚合图中邻居节点的偏好信息，并与用户特征、会话短期特征融合为会话嵌入表示，在行为序列视角中，通过一种自适应旋转位置编码方法将项目嵌入与位置嵌入统一建模，然后使用自注意力网络学习项目之间的顺序依赖关系，得到会话表示；其次，为了保证两种视角学习到的偏好信息的一致性，通过对比学习拉近两种视角下学习到的嵌入表示，为了提高模型的鲁棒性，对原视角破坏生成增强视角，将原视角下学到的嵌入表示与增强视角下学到的嵌入表示进行对比学习，以提高抗噪能力；最后，将两种视角下学习到的会话嵌入表示进行融合，生成最终的嵌入表示，并做出预测，流程图如图 1 所示：

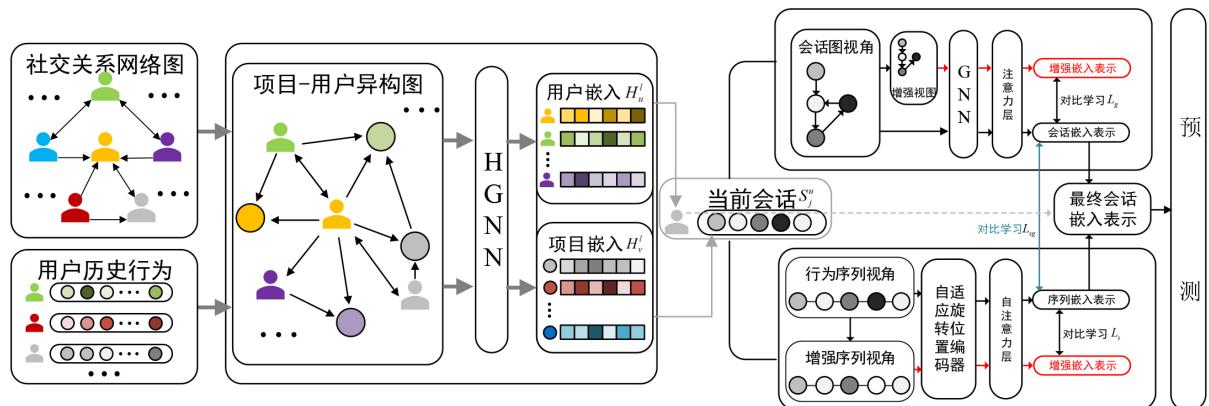


Figure 1. Workflow diagram of DVSSR
图 1. DVSSR 工作流程图

2.3. 用户和项目嵌入表示学习

本节的主要目的是从异构图 K 中学习二者的嵌入表示，异构图 $K = (X, \mathcal{E})$ 中的节点集合为 $X = V \cap U$ ，边集合 \mathcal{E} 由社交网络图的边集 E 、项目 - 项目历史交互以及用户 - 项目历史交互共同构成。

在已有研究中，SERec [3]发现如果直接对异构图中不同种类的节点进行聚合操作，由于用户与项目的交互次数要远大于与其他用户的交互次数，可能会造成信息失衡，本文延续了 SERec 中对该问题的处理方式，通过注意力机制学习不同节点的权重，来缓解信息失衡问题。

首先，计算当前节点与其邻居节点的特征重要性，以用户节点为例：

$$\text{Imp}^l(u, x) = q_{ux}^{l T} \cdot \sigma\left(W_{ux}^l\left(H_u^{l-1} \| H_x^{l-1}\right) + b_{ux}^l\right) \quad (1)$$

其中， $q_{ux}^l \in \mathbb{R}^d$ ， $W_{ux}^l \in \mathbb{R}^{d \times 2d}$ ， $b_{ux}^l \in \mathbb{R}^d$ 是可学习参数， σ 为 sigmoid 激活函数。在异构图神经网络的第 l 层中， $H_u^{l-1} \in \mathbb{R}^d$ 与 $H_x^{l-1} \in \mathbb{R}^d$ 分别为节点 u 与 x 在第 $l-1$ 层学习到的嵌入表示，使用均匀分布对嵌入向量 H_x^0 进行初始化， $\|$ 代表连接操作。然后，为了平衡不同类型节点对目标节点的影响，使用 softmax 函数对特征重要性进行归一化处理，得到每个邻居节点的注意力系数：

$$\alpha_{ux}^l = \text{Softmax}_{x \in N_u} (\text{Imp}^l(u, x)) \quad (2)$$

式(2)中, N_u 为节点 u 的邻居集合。接下来, 将邻居节点 $x \in N_u$ 的节点信息根据注意力系数传播给目标节点 u :

$$\widetilde{H}_u^l = \sum_{x \in N_u} \alpha_{ux} \left(\widetilde{W}^l H_x^{l-1} + \widetilde{b}^l \right) \quad (3)$$

其中, $\widetilde{W}^l \in \mathbb{R}^{d \times d}$, $\widetilde{b}^l \in \mathbb{R}^d$ 是可学习参数。通过公式(3), 将节点 u 的邻居节点 $x \in N_u$ 的上一层状态通过线性变化后乘以它们之间注意力系数, 为目标节点传播特征信息。最后, 将节点 u 在公式(3)中的隐藏层输出与上一层输出相连接, 经过线性变换以及 ReLU 激活函数后, 得到该节点在 L 层的异构图神经网络中学习到的最终的嵌入表示:

$$H_u^l = \text{ReLU} \left(W^l \left(\widetilde{H}_u^l \| H_u^{l-1} \right) + b^l \right) \quad (4)$$

式(4)中, $W^l \in \mathbb{R}^{d \times 2d}$, $b^l \in \mathbb{R}^d$ 是可学习参数。通过这种方式, 使节点 u 捕获到邻居节点的偏好信息。对于项目节点 v , 采用同样的操作得到其嵌入表示 H_v^l 。

2.4. 会话图视角下的会话嵌入表示学习

在上一节中, 通过异构图神经网络学习到了项目与用户的嵌入表示, 但这些嵌入表示不携带当前会话的相关特征信息。因此, 在本节将根据此前学习到的用户以及项目嵌入表示, 提取用户在当前会话中的个性化偏好信息, 并生成该会话的嵌入表示。

给定用户 u 的当前会话 $S_j^u = [v_1, v_2, v_3, \dots, v_n]$, 构建一张会话图 G_j , 图中的节点对应会话中的项目, 边对应会话中项目之间的交互。首先, 将会话 S_j^u 中所有节点的嵌入初始化:

$$\hat{h}_i = H_u^l \| H_v^l \quad (5)$$

其中, $H_u^l \in \mathbb{R}^d$ 以及 $H_v^l \in \mathbb{R}^d$ 分别为该会话的用户和会话中的项目 i 在异构图 K 中学习到的带有社交属性的嵌入表示, 为了强调用户信息的重要性以实现个性化推荐, 将两种嵌入连接生成图 G_j 中该节点的初始化嵌入 $\hat{h}_i \in \mathbb{R}^{2d}$ 。然后, 分别从入度和出度的角度, 为该节点传递其邻居特征信息:

$$\begin{aligned} h_i^{in} &= \frac{1}{\sum_{j \in N_i^{in}} w_{ij}} \sum_{j \in N_i^{in}} w_{ij} W_{in} \hat{h}_j \\ h_i^{out} &= \frac{1}{\sum_{j \in N_i^{out}} w_{ji}} \sum_{j \in N_i^{out}} w_{ji} W_{out} \hat{h}_j \end{aligned} \quad (6)$$

式(6)中, w 是边的权重, 代表着会话中节点 i 与 j 的交互次数, N_i^{in} 和 N_i^{out} 分别为节点 i 的入度邻居以及出度邻居集, W_{in} , $W_{out} \in \mathbb{R}^{d \times 2d}$ 是可学习参数。将该节点聚合到的入度特征与出度特征连接:

$$\tilde{h}_i = h_i^{in} \| h_i^{out} \quad (7)$$

然后, 为了保留节点自身的特征信息, 通过门控网络将原始特征 \hat{h}_i 与 GNN 学习到的特征 \tilde{h}_i 结合:

$$\begin{aligned} \hat{h}_i &= \tanh \left(W_h \left(\tilde{h}_i \| \hat{h}_i \right) + b_h \right) \\ r_i &= \sigma \left(W_r \left(\tilde{h}_i \| \hat{h}_i \right) + b_r \right) \\ h_i &= r_i \odot \hat{h}_i + (1 - r_i) \odot \widehat{W}_h \hat{h}_i \end{aligned} \quad (8)$$

其中 W_h , $W_r \in \mathbb{R}^{d \times 4d}$, $\widehat{W}_h \in \mathbb{R}^{d \times 2d}$ 以及 b_h , $b_r \in \mathbb{R}^d$ 是可学习参数, \tanh 为双曲正切激活函数, \odot 代表哈

达玛积。在学习到节点的嵌入表示后，通过轻量级的注意力机制为每个节点学习其注意力系数，用来聚合会话的嵌入表示：

$$\begin{aligned}\beta_i &= \text{softmax} \left(q_\beta^T \sigma \left(W_\beta \left(h_i \| h_n \| H_u^l \right) + b_\beta \right) \right) \\ S^g &= \sum_{i \in n} \beta_i h_i\end{aligned}\quad (9)$$

式(9)中， $b_\beta, q_\beta \in \mathbb{R}^d$ 以及 $W_\beta \in \mathbb{R}^{d \times 3d}$ 是可学习参数， h_n 是会话中最后一项的嵌入表示，用来代表用户的短期兴趣并融入到会话表示 S_j^g 中。将公式(5)至公式(9)的一系列流程使用个性化图神经网络 PGNN 表示：

$$S_j^g = \text{PGNN}(G_j) \quad (10)$$

2.5. 行为序列视角下的会话嵌入表示学习

目前，在对用户行为序列进行建模的主流模型中，SASRec [13]和BERT4Rec [14]为了识别序列中项目的位置信息，为项目嵌入表示添加了一种可学习的绝对位置编码，虽然有一定效果，但是这种编码方式外推性较差且无法捕捉项目之间的相对位置关系，例如距离相近的项目之间位置编码应保持一定的相似性，距离较远的项目之间位置编码应保持一定的差异性。Eluerfomer [15]框架设计了一种自适应旋转位置编码方式，利用欧拉公式将注意力机制建模到复数向量空间中，实现了语义信息和位置信息的统一。本文受该模型启发，在行为序列视角上通过自适应旋转位置编码提高对项目之间相对位置信息的感知能力，学习到更准确的会话嵌入表示。

首先，根据异构图 K 学习到的项目嵌入表示对会话 $S_j^u = [i_1, i_2, i_3, \dots, i_n]$ 中的项目嵌入信息初始化，得到该行为序列的初始嵌入矩阵 $E_j = [e_1, e_2, e_3, \dots, e_n]$ 。然后为序列中每个位置都生成一个位置嵌入 $P = [p_1, p_2, p_3, \dots, p_n]$ 以及旋转位置嵌入 $R = [r_1, r_2, r_3, \dots, r_n]$ 。将序列嵌入与位置嵌入相加，得到序列嵌入表示 $\tilde{Z} = [\tilde{z}_1, \tilde{z}_2, \tilde{z}_3, \dots, \tilde{z}_n]$ ，其中 $\tilde{z}_i = e_i + p_i$ 。将序列中每一个项目的嵌入根据索引的奇偶拆分为实部复向量和虚部复向量：

$$z_i^\gamma = [t_1, t_3, t_5, \dots, t_{d-1}], z_i^\rho = [t_2, t_4, t_6, \dots, t_d] \quad (11)$$

其中， $z_i^\gamma \in \mathbb{R}^{d/2}$ 和 $z_i^\rho \in \mathbb{R}^{d/2}$ 分别为该项目嵌入的实部复向量和虚部复向量， d 代表该嵌入的维度。根据欧拉公式，对其进行一系列变换，首先，计算其模 λ_i 与相位角 θ_i ：

$$\lambda_i = \sqrt{z_i^\gamma{}^2 + z_i^\rho{}^2}, \theta_i = \tan 2(z_i^\gamma, z_i^\rho) \quad (12)$$

然后，将相位角与旋转位置嵌入组合，转换回直角坐标系中，以实现自适应旋转位置编码：

$$\hat{z}_i = \lambda_i \cdot [\cos(\theta_i + r_i), \sin(\theta_i + r_i)] \quad (13)$$

根据上述公式，得到了该序列统一语义信息和位置信息的自适应旋转位置编码嵌入矩阵 $\hat{Z} = [\hat{z}_1, \hat{z}_2, \hat{z}_3, \dots, \hat{z}_n]$ ，将其作为多头自注意力层[16]的输入以捕获项目之间成对的转换关系，其中每一个头的自注意力层输出为：

$$Z_i = \text{softmax} \left(\frac{Q_i K_i^T}{\sqrt{d_k}} \right) V_i = \text{softmax} \left(\frac{(\hat{Z} W_i^Q)(\hat{Z} W_i^K)^T}{\sqrt{d_k}} \right) \hat{Z} W_i^V \quad (14)$$

其中， $W_i^Q, W_i^K, W_i^V \in \mathbb{R}^{d \times d_k}$ 是可学习的参数，用来将嵌入 \hat{Z} 线性映射到查询、键和值空间中。 $d_k = d/h$ 是多头注意力层中每个头的维度，然后将不同的头连接后线性变换，经过前馈层后，得到序列最终的嵌入表示：

$$\begin{aligned}\widehat{Z} &= \text{Concat}(\widehat{Z}_1, \widehat{Z}_2, \widehat{Z}_3, \dots, \widehat{Z}_h)W^o \\ Z &= \text{ReLU}(W_{F1}\widehat{Z} + b_{F1})W_{F2} + b_{F2}\end{aligned}\quad (15)$$

式(15)中, $W^o \in \mathbb{R}^{(h \times d_k) \times d}$, $W_{F1}, W_{F2} \in \mathbb{R}^{d \times d}$, $b_{F1}, b_{F2} \in \mathbb{R}^d$, 是可学习参数, h 是头的数量。将序列嵌入矩阵 $Z = [z_1, z_2, z_3, \dots, z_n]$ 中的最后一项 z_n 作为行为序列视角中学习到嵌入表示, 即 $S_j^s = z_n$ 。将公式(11)至公式(15)的一系列流程以欧拉变换网络表示:

$$S_j^s = \text{ELF}(E_j) \quad (16)$$

2.6. 对比学习

在前两节中, 分别从会话图视角以及行为序列视角学习到了用户当前会话的嵌入表示, 为了确保两种视角下学习到的嵌入表示的一致性, 本文设计了一种损失函数, 视同一个会话的两个视角下的嵌入表示为正样本, 拉近它们之间的距离, 同一批次内的其他会话两种视角下的嵌入表示均为负样本, 疏远它们之间的距离, 损失函数如下:

$$L_{sg} = -\log \frac{\exp(\cos(S_j^s, S_j^s))}{\exp(\cos(S_j^s, S_j^s)) + \sum_{S_j \in S^-} \exp(\cos(S_j^s, S_j)) + \sum_{S_j \in S^-} \exp(\cos(S_j^s, S_j))} \quad (17)$$

式(17)中, \cos 为余弦相似度函数, 用来衡量两个嵌入表示的相似程度, S^- 为会话 S_j 的负样本集合。通过该损失函数, 可以最大化会话在会话图视角以及行为序列视角下的学习到的嵌入之间的互信息, 确保学习到偏好信息的一致性。

数据增强能够提高模型的鲁棒性和泛化能力, 在用户的会话中, 并非所有项目之间都存在强相关性, 用户在不经意间的操作可能将不相关的项目引入会话中, 因此产生了噪声, 为了提高模型的鲁棒性和泛化能力, 参考 CL4SRec [17]以及 GRACE [18]中提出的数据增强方法, 为会话生成增强视图并与原视图进行对比学习。

在行为序列视角中, 将序列中每个项目的初始嵌入以 20% 的概率设置为掩码嵌入, 以此得到该序列的增强后的初始嵌入矩阵 \widehat{E}_j , 将 \widehat{E}_j 作为欧拉变换网络的输入, 学习该序列中项目之间的转移关系, 得到增强序列的嵌入表示 $\widehat{S}_j^s = \text{ELF}(\widehat{E}_j)$ 。然后, 视同一个会话的行为序列视角下学习到的原嵌入与增强序列后的嵌入表示为正样本对 (S_j^s, \widehat{S}_j^s) , 同一批次下训练的其他会话的嵌入为负样本, 得到损失函数:

$$L_s = -\log \frac{\exp(\cos(S_j^s, \widehat{S}_j^s))}{\exp(\cos(S_j^s, \widehat{S}_j^s)) + \sum_{k=1, k \neq i}^N \exp(\cos(S_i^s, S_k^s))} \quad (18)$$

在会话图视角中, 生成一个与会话图 G_j 邻接矩阵 $A_j \in \mathbb{R}^{n \times n}$ 大小相同的掩码矩阵 $O_j \in \mathbb{R}^{n \times n}$, 矩阵中的所有元素独立地从伯努利分布中抽取 $O_{ij} \sim B(0, 0.2)$, 将邻接矩阵 A 与掩码矩阵 O 进行元素乘法, 以达到随机丢弃掉边的目的, 生成增强会话图 \widehat{G}_j 。将其作为个性化图神经网络 PGNN 的输入, 为项目节点聚合邻居的特征信息, 生成增强会话图 $\widehat{S}_j^g = \text{PGNN}(\widehat{G}_j)$ 的嵌入表示, 与公式(18)做法类似, 得到损失函数:

$$L_g = -\log \frac{\exp(\cos(S_j^g, \widehat{S}_j^g))}{\exp(\cos(S_j^g, \widehat{S}_j^g)) + \sum_{k=1, k \neq i}^N \exp(\cos(S_j^g, S_k^g))} \quad (19)$$

将上述的三种损失函数结合，得到对比学习模块的总损失函数：

$$L_c = L_{sg} + L_s + L_g \quad (20)$$

2.7. 预测

在分别从会话图视角以及行为序列视角学习到会话 S_j^u 的嵌入表示后，将会话图视角下的会话嵌入表示 S_j^g ，行为序列视角下的序列嵌入表示 S_j^s 与该用户的嵌入表示 H_u^l 结合，生成该会话的最终嵌入表示

$$S_j = W_f (S_j^g \| S_j^s \| H_u^l) \quad (21)$$

其中， $W_f \in \mathbb{R}^{d \times 3d}$ 是可学习参数。然后，将最终嵌入表示 S_j 与项目 v_i 的嵌入表示相乘，以得到该项目作为会话下一个项目的概率：

$$\hat{y}_i = \text{softmax}(S_j^T v_i) \quad (22)$$

根据概率排名生成前 top-K 项目的候选项目列表。将损失函数定义为预测和真实标签的交叉熵，与对比学习损失函数结合，作为总损失函数

$$\begin{aligned} L &= -\sum_{i=1}^N y_i \log(\hat{y}_i) \\ L_{total} &= L + \varphi L_c \end{aligned} \quad (23)$$

3. 实验结果

3.1. 数据集介绍

本文使用了三个大型的公开数据集 Gowalla [19]、Delicious [20]、Foursquare [21]，验证所提出模型的性能。在对数据集的预处理方面，参考 SEREC 中的模型。在数据集划分上，前 60% 作为训练集，20% 作为验证集，最后 20% 作为测试集。数据集的主要信息如表 1 所示：

Table 1. Dataset information

表 1. 数据集信息表

数据集	Gowalla	Foursquare	Delicious
用户 - 项目交互次数	1,218,599	3,627,093	266,190
会话数量	258,732	888,798	60,397
用户数量	33,661	39,302	1313
项目数量	41,229	45,595	5793
用户 - 用户交互次数	283,778	304,030	9130
平均会话长度	4.71	4.08	4.41

3.2. 主要参数

本文所提出的 DVSSR 主要参数设置如下：嵌入维度大小为 256，异构图神经网络层数为 1，批次大小为 128，学习率为 0.0002，多头注意力网络头的数量为 2，对比学习损失函数的权重 φ 为 0.01，异构图中邻居节点的采样数量为 10。为了衡量模型的性能，采用推荐系统中的主流评估指标命中率 HIT@K 以及平均倒数排名 MRR@K 衡量生成的候选推荐列表，K 取 10 和 20。

3.3. 对比实验

为了评价所提出模型的性能, 将 DVSSR 在三个数据集上的性能与会话推荐模型: NARM [6]、STAMP [7]、NextItNet [10]、SR-GNN [11]、SSRM [8], 以及基于会话的社交推荐模型: DGRec [2]、SERec [3]、HF²-Rec [4] 进行对比实验, 由于会话推荐模型不能使用数据集中的社交网络信息, 为了实验公平性, 将异构图神经网络所学习到的项目嵌入表示作为这类模型的项目的初始化嵌入。实验结果如表 2 所示:

Table 2. Comparison of experimental results

表 2. 对比实验结果

数据集	Gowalla				Delicious				Foursquare			
	H10	M10	H20	M20	H10	M10	H20	M20	H10	M10	H20	M20
s-NextItNet	45.57	24.91	53.08	25.43	39.00	20.86	48.56	21.52	61.11	33.77	69.53	34.36
s-NARM	44.44	24.09	51.95	24.60	39.69	21.07	49.25	21.83	59.66	32.30	68.05	32.90
s-STAMP	45.26	24.65	52.69	25.18	38.78	20.52	48.06	21.11	61.22	33.73	69.54	34.31
s-SRGNN	45.40	25.07	52.79	25.58	39.97	21.30	49.41	21.98	61.10	33.67	69.55	34.26
s-SSRM	45.07	24.75	52.56	25.26	39.73	21.02	49.28	21.63	61.18	33.83	69.62	34.43
DGRec	42.14	23.08	49.97	23.61	37.69	20.15	47.28	20.80	57.23	31.65	65.91	32.26
SERec	45.98	25.22	53.58	25.74	40.04	21.41	49.36	22.06	61.72	34.10	70.12	34.69
HF ² -Rec	46.80	25.99	54.28	26.52	41.02	22.16	50.34	22.78	63.28	35.65	71.55	36.23
DVSSR	48.35	27.43	56.15	27.97	43.50	23.91	53.40	24.57	64.90	37.85	72.94	38.42
性能提升	3.31%	5.54%	3.45%	5.47%	6.05%	7.90%	6.08%	7.86%	2.56%	6.17%	1.94%	6.04

根据表 2 的对比实验结果, 从整体上看, 会话推荐模型在使用异构图 K 学习到的嵌入表示作为其模型中项目的嵌入初始化后, 效果仍然不如基于会话的社交网络推荐, 因为除异构图外, 基于会话的社交推荐模型比较重视用户的特征信息, SERec、HF²-Rec 以及本文提出的 DVSSR 在会话嵌入表示学习中都将用户嵌入信息以不同的方式融入了会话表示中, 以强调用户这一重要社交属性, 但会话推荐模型没有有效的利用此信息, 因此造成性能上的次优。在所有的基线模型与本文所提出模型的对比中, DVSSR 在所有数据集上的表现都优于其他模型。其中, DVSSR 在 Delicious 数据集上的提升较其他两个数据集更大, 可能的原因是, Delicious 数据集的体量要明显小于其他两个数据集, 会话数量、项目数量、交互次数均在三者中最小, 因此 Delicious 的数据稀疏程度更高。在这种情况下, 如果能采用双重视角, 从图和序列两个视角共同解决问题, 而不是其他基线模型中只关注单一的图视角, 那么也许能够从不同的模式中学习到更多的特征信息, 得到更准确地会话嵌入表示, DVSSR 在最大的数据集 Foursquare 上的性能提升较小也从反面验证了这一观点。因此, 本文所提出的 DVSSR 模型能在数据稀疏的场景中通过结合两种模式得到更好的表现。

3.4. 消融实验

为了验证本文所提出的 DVSSR 模型中各模块的有效性以及对性能的影响, 分别设计三种模式进行消融实验。模式一: 在该模式中, 只采用单一的会话图视角, 通过 PGNN 学习会话嵌入表示, 以验证通过双重视角学习会话嵌入表示的有效性。模式二: 在该模式中, 去掉设计的对比学习模块, 不通过对对比学习约束两种模式下学习到的偏好的一致性, 忽略增强视图为模型带来的影响。模式三: 在该模式中, 使用普通的可学习的绝对位置编码与嵌入表示相加, 而非自适应旋转位置编码, 作为多头自注意力网络的输入, 验证学习相对位置信息对性能产生的影响。实验结果如表 3 所示:

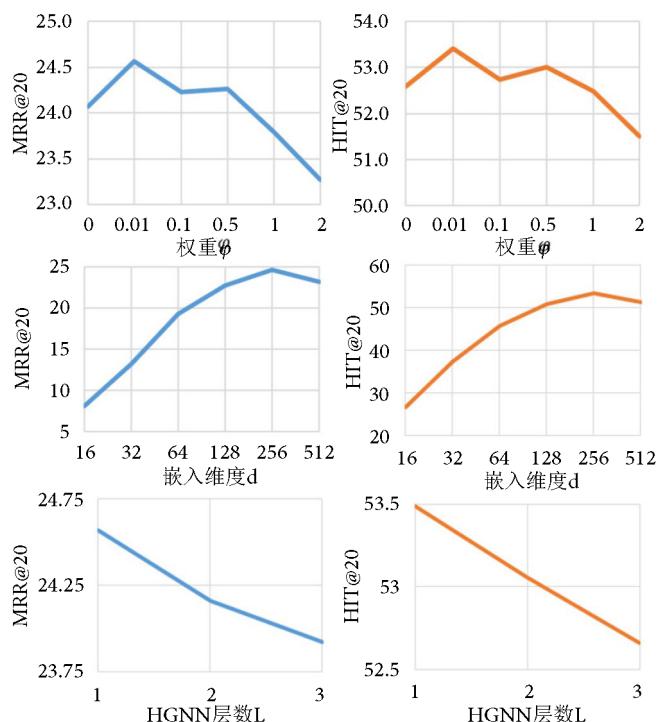
Table 3. Results of ablation experiment**表 3. 消融实验结果**

数据集	Gowalla		Delicious		Foursquare	
模式	H20	M20	H20	M20	H20	M20
模式一	53.79	26.07	49.76	22.31	70.68	35.43
模式二	55.64	27.52	52.59	24.07	72.38	37.67
模式三	55.89	27.75	53.00	24.32	72.66	38.05
DVSSR	56.15	27.97	53.40	24.57	72.94	38.42

根据表 3 消融实验结果来看，模式一由于只通过会话图视角学习项目之间的共现关系，未能从项目顺序依赖性的角度丰富会话的嵌入表示，导致性能下滑最为严重，与完整的 DVSSR 模型相比在不同的数据集上下降了约 4%~8% 的性能。由于模式一中不包含行为序列视角，因此性能下滑的原因可能包括模式二中的未使用对比学习以及模式三中的未使用自适应位置编码方法，但毫无疑问，去掉行为序列视角对 DVSSR 模型的影响最大，由此验证了双重视角框架的有效性。在模式二中，去掉了对比学习损失函数，两种视角下学习的用户偏好可能带有偏差，并且鲁棒性以及抗噪能力均低于 DVSSR，因此造成了约 0.7%~2% 的性能下跌。在模式三种，使用可学习的绝对位置编码，未能有效的捕获到项目之间的相对位置关系，损失了 0.5%~1% 的性能。

3.5. 参数实验

为了研究 DVSSR 模型中参数的变化对性能的影响，本节将在 Delicious 数据集上对 $\varphi \in \{0, 0.01, 0.1, 0.5, 1, 2\}$, $d \in \{16, 32, 64, 128, 256, 512\}$, $L \in \{1, 2, 3\}$ 的取值展开实验，实验结果如图 2 所示。

**Figure 2.** Parameter experiment results**图 2. 参数实验结果**

权重 φ 影响了对比学习损失函数 L_c 在总损失函数 L_{total} 的占比，从图 2 中可以看出，当 φ 的取值超过 0.01 时，对比学习任务逐渐主导了学习过程，因此可能会影响预测任务的重要性，造成了性能的大幅下降。嵌入维度 d 从 16 逐渐上升到 256 的过程中，性能不断提升，这说明合理的提升嵌入维度能够增加模型容量，获得更准确的嵌入表示，但是当 d 增加到 512 时，可能产生了过拟合现象，造成了性能下滑，因此要合理选择嵌入维度。HGNN 的层数 L 决定着异构图中的节点能够得到 L 跳范围的邻居信息，但从实验结果来看，增加 L 的层数可能会引起过渡平滑现象，即节点的嵌入表示趋近相同，造成了性能下降，因此将层数 L 设定为 1。

4. 结束语

现有的基于会话的社交推荐模型单一的从图视角解决问题，忽略了项目之间顺序依赖性带来的影响，针对这一问题，本文提出了一种基于会话的双视角社交推荐模型。首先，从用户 - 项目异构图中学习项目嵌入以及用户嵌入表示。然后分别从用户的会话图以及行为序列两种视角出发，在会话图视角中，使用 GNN 传播节点之间的特征信息，将节点信息聚合得到会话的嵌入表示。在行为序列视角中，使用一种自适应旋转位置编码，捕获项目之间的相对位置关系，更好地对项目的语义信息以及位置信息融合，通过多头注意力机制学习行为序列的偏好信息。其次，将两种不同视角下学习到的会话偏好信息融合，作为最终的会话嵌入表示并做出推荐。接着，为了保证两种视角学习到的偏好信息的一致性，通过对比学习拉近两种视角下学习到的嵌入表示，为了提高模型的鲁棒性，对原视角破坏生成增强视角，将原视角下学到的嵌入表示与增强视角下学习到的嵌入表示进行对比学习，以提高抗噪能力。最后，在三个大型公开数据集上的对比实验表明了提出的 DVSSR 模型在性能上优于现有的会话推荐模型以及基于会话的社交推荐模型，在不同数据上性能约提升了 2%~8%，并通过消融实验分别验证了所提出框架中不同模块的有效性。

参考文献

- [1] Ninichuk, M. and Namiot, D. (2023) Survey on Methods for Building Session-Based Recommender Systems. *International Journal of Open Information Technologies*, **11**, 22-32.
- [2] Song, W., Xiao, Z., Wang, Y., Charlin, L., Zhang, M. and Tang, J. (2019) Session-Based Social Recommendation via Dynamic Graph Attention Networks. *Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining*, Melbourne, 11-15 February 2019, 555-563. <https://doi.org/10.1145/3289600.3290989>
- [3] Chen, T. and Wong, R.C. (2021) An Efficient and Effective Framework for Session-Based Social Recommendation. *Proceedings of the 14th ACM International Conference on Web Search and Data Mining*, Virtual Event, 8-12 March 2021, 400-408. <https://doi.org/10.1145/3437963.3441792>
- [4] Wang, L., Li, M. and Zheng, H. (2023) Rethinking Rule-Based Approaches in Session-Based Recommendation. *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Rhodes Island, 4-10 June 2023, 1-5. <https://doi.org/10.1109/icassp49357.2023.10094851>
- [5] Hidasi, B., Karatzoglou, A., Baltrunas, L., et al. (2016) Session-Based Recommendations with Recurrent Neural Networks. *Proceedings of the 4th International Conference on Learning Representations*, Puerto Rico, May 2016, 1-10.
- [6] 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>
- [7] 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>
- [8] Guo, L., Yin, H., Wang, Q., Chen, T., Zhou, A. and Quoc Viet Hung, N. (2019) Streaming Session-Based Recommendation. *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, Anchorage, 4-8 August 2019, 1569-1577. <https://doi.org/10.1145/3292500.3330839>
- [9] Tang, J. and Wang, K. (2018) Personalized Top-N Sequential Recommendation via Convolutional Sequence Embedding.

- Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining*, Marina, 5-9 February 2018, 565-573. <https://doi.org/10.1145/3159652.3159656>
- [10] Yuan, F., Karatzoglou, A., Arapakis, I., Jose, J.M. and He, X. (2019) A Simple Convolutional Generative Network for Next Item Recommendation. *Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining*, Melbourne, 11-15 February 2019, 582-590. <https://doi.org/10.1145/3289600.3290975>
- [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] 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*, Virtual Event, 25-30 July 2020, 169-178. <https://doi.org/10.1145/3397271.3401142>
- [13] 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>
- [14] Sun, F., Liu, J., Wu, J., Pei, C., Lin, X., Ou, W., et al. (2019) BERT4Rec: Sequential Recommendation with Bidirectional Encoder Representations from Transformer. *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*, Beijing, 3-7 November 2019, 1441-1450. <https://doi.org/10.1145/3357384.3357895>
- [15] Tian, Z., Zhao, W.X., Zhang, C., Zhao, X., Ma, Z. and Wen, J. (2024) EulerFormer: Sequential User Behavior Modeling with Complex Vector Attention. *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval*, Washington, 14-18 July 2024, 1619-1628. <https://doi.org/10.1145/3626772.3657805>
- [16] 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.
- [17] Xie, X., Sun, F., Liu, Z., Wu, S., Gao, J., Zhang, J., et al. (2022) Contrastive Learning for Sequential Recommendation. 2022 IEEE 38th International Conference on Data Engineering (ICDE), Kuala Lumpur, 9-12 May 2022, 1259-1273. <https://doi.org/10.1109/icde53745.2022.00099>
- [18] Zhu, Y., Xu, Y., Yu, F., et al. (2020) Deep Graph Contrastive Representation Learning. <https://doi.org/10.48550/arXiv.2006.04131>
- [19] Cho, E., Myers, S.A. and Leskovec, J. (2011) Friendship and Mobility. *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Diego, 21-24 August 2011, 1082-1090. <https://doi.org/10.1145/2020408.2020579>
- [20] Cantador, I., Brusilovsky, P. and Kuflik, T. (2011) Second Workshop on Information Heterogeneity and Fusion in Recommender Systems (HetRec2011). *Proceedings of the Fifth ACM Conference on Recommender Systems*, Chicago, 23-27 October 2011, 387-388. <https://doi.org/10.1145/2043932.2044016>
- [21] Yang, D., Zhang, D., Yu, Z. and Yu, Z. (2013) Fine-Grained Preference-Aware Location Search Leveraging Crowdsourced Digital Footprints from LBSNs. *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, Zurich, 8-12 September 2013, 479-488. <https://doi.org/10.1145/2493432.2493464>