

基于虚拟关系知识图可自适应聚合的推荐算法

李源¹, 杨谋均^{2†}

(1. 湖南省电子信息产业研究院, 湖南 长沙 410001; 2. 中车株洲电力机车研究所有限公司, 湖南 株洲 412001)

摘 要:在信息爆炸的时代,推荐算法成为应对信息过载的有效手段。近年来,图神经网络(GNN)以其强大的建模能力和应对冷启动的优势被广泛应用于推荐算法。本文提出了一种基于深度强化学习与 GNN-R 的联合训练框架,解决 GNN-R 中固定层数和聚合策略的问题,通过间隔经验回放和延迟奖励机制,优化了推荐模型的学习过程。在此基础上,提出了自适应优化 GNN-R 聚合层数和虚拟关系数量的两个优化算法,改进了 VRKG4Rec 模型的性能。实验结果表明,两个优化算法对比 VRKG4Rec 模型都有较好的性能提升。

关键词:推荐算法;图神经网络;深度强化学习;知识图谱

中图分类号:TP391

文献标识码:A

Recommendation Algorithm Based on Adaptive Aggregation of Virtual Relationship Knowledge Graph

LI Yuan¹, YANG Moujun^{2†}

(1. Hunan Electronic Information Industry Institute, Changsha, Hunan 410001, China;

2. CRRC Zhuzhou Institute Co., Ltd., Zhuzhou, Hunan 412001, China)

Abstract: In the era of information explosion, recommendation algorithms have become an effective means to cope with information overload. In recent years, graph neural networks (GNN) have been widely applied in recommendation algorithms due to their powerful modeling capabilities and advantages in addressing cold start issues. A joint training framework based on deep reinforcement learning and GNN-R is proposed in this paper to address the fixed-layer and aggregation strategy issues in GNN-R. By employing interval experience replay and delayed reward mechanisms, the learning process of the recommendation model is optimized. Building upon this, two optimization algorithms for adaptively optimizing the aggregation layers and virtual relation quantities in GNN-R are proposed, enhancing the performance of the VRKG4Rec model. Experimental results demonstrate significant performance improvements of the two optimization algorithms compared to the VRKG4Rec model.

Key words: recommendation algorithm; graph neural network; deep reinforcement learning; knowledge graph

在网络普及率极高的互联网时代,当前社会不管是什么类型的人群,每个类型人群的每类需求都能在这个信息时代得到满足,信息的汪洋包裹住了各类用户。在这种数据量溢出的“倒灌”环境下,用户与服务提供者之间出现了双向选择的不适性,有

效分析出到底用户需要什么样的信息成了一个关键问题。在这样的背景下,推荐系统^[1]就像在茫茫海洋中引路的指南针,它通过智能算法帮助用户在信息的大海中导航,有效缓解了“信息过载”的问题。如今,推荐系统已经成为电子商务、社交媒体、

娱乐流媒体、新闻和广告等领域中至关重要的一部分,它的存在极大地丰富了用户的使用体验,推动了个性化服务的发展。

近年来,知识图谱(Knowledge Graph, KG)^[2]和图神经网络(Graph Neural Networks, GNN)^[3]经常被引入推荐算法中。在基于 GNN 的推荐(GNN-based Recommendation, GNN-R)中,一个合理的聚合策略通常可以使推荐模型获取到有效的来自用户侧或项目侧的高阶邻域信息。但是,目前大多数 GNN-R 在使用 GNN 传播和聚合高阶信息时,使用的都是固定的层数和固定的聚合策略,这限制了模型学习不同项目特征和精确聚合不同子图的能力^[4]。基于此,提出了一种基于 DQN^[5]的双模型联合训练框架,设计了一种间隔经验回放机制,让智能体进行即时和历史经验学习,设计了一种延后于动作的奖励机制,对短期收益和长期收益进行有效的模拟。本文应用这一框架从两个方面优化了 VRKG4Rec 模型^[6],实验结果表明这两个优化算法在 Top-K 推荐中的性能皆优于 VRKG4Rec 模型。

1 相关工作

推荐系统通常存在数据稀疏性和冷启动问题^[7],为了缓解这两个问题,目前的普遍做法是引入额外的辅助信息,例如上下文、用户标签和知识图谱等。知识图谱是一种包含实体信息以及实体间关系的图形结构化数据,通过将目标项和知识图谱中的实体关联起来,模型能够探索知识图谱中不同项之间的语义关联及高阶连接信息。文献[8]提出了一种新的推荐系统框架,旨在通过结合强化学习和知识图谱推理来提高推荐系统的性能和可解释性。该框架首先利用知识图谱获取深层次的语义信息,然后通过强化学习方法,让模型在知识图谱中自动学习并寻找最佳的推理路径。

图神经网络是卷积神经网络(Convolutional Neural Networks, CNN)^[9]和图表示学习(Graph Representation Learning, GRL)^[10]的结合,它能够解决某些情况下 CNN 难以处理的问题。文献[11]提出了一种新型的基于解耦图神经网络的推荐算法。它能够识别用户、项目以及标签间的相互关系,并据此提供可解释的推荐理由。

在当前的推荐算法研究领域中,深度强化学习技术凭借其结合了深度学习强大的数据处理能力以及强化学习的决策制定能力的优点,已经成为引

领创新的前沿技术。文献[12]提出了一种结合认知心理学的 OJ 习题推荐方法。它通过考虑习题描述、输入输出格式以及题目语言等因素来评估习题难度,基于学生的学习间隔和学习次数来建模学生的记忆-遗忘曲线和学习模式,并使用强化学习和项集挖掘算法提高习题推荐的准确性。

2 基于 DQN 的双模型联合训练框架

文献[4]针对目前大多数基于 GNN 的推荐算法存在的图形邻域感知范围固定的问题,提出了一种能够自适应聚合优化用户和项目高阶连接信息的 DPAO 框架。DPAO 将每个用户和项目的聚合优化问题描述为马尔可夫决策过程,并使用两个 DQN 模型来分别学习用户侧和项目侧 GNN 的自适应聚合策略。基于 DPAO 的设计思想,本文提出了一种更为通用的 GNN-R 联合训练框架,它可以对 GNN-R 模型中某些具有优化空间且在传播聚合过程中较为关键的参数进行优化,而不仅局限于对 GNN 层数进行优化。图 1 为本框架的整体架构。

左侧是输入图,在本文中是一个协作知识图,它的节点包含用户及项目节点,包含项目间的关联关系以及用户-项目交互关系。右侧是 GNN-R 训练、间隔经验收集以及 DQN 训练部分。

2.1 双模型中 DQN 的训练过程

DQN 训练过程如算法 1 所示。

在训练开始前,使用一个局部 GNN-R 来复制全局 GNN-R 模型,目的是缓解双模型训练时, DQN 环境(主模型)参数连续变化带来的连续性和不稳定性问题。获取当前状态与下一状态数据后,使用式(1)来计算动作奖励:

$$r_t = (e_{s_{t+1}}^{a_{t+1}} e_{t+1}^p - e_{s_{t+1}}^{a_t} e_{t+1}^n) - (e_{s_t}^{a_t} e_t^p - e_{s_t}^{a_t} e_t^n) \quad (1)$$

计算延迟奖励之后,立即使用当前收集到的 (s_t, a_t, r_t, s_{t+1}) 数据训练 DQN,并将其存储于经验池中,最后再从经验池中采样训练 DQN。DQN 模型的 Loss 函数如式(2)^[5]:

$$L_{\text{DQN}} = \{Q(s_t, a_t \mid \theta) - [r_t + \gamma \max_{a_{t+1}} Q(s_{t+1}, a_{t+1} \mid \theta^{\text{Target}})]\}^2 \quad (2)$$

其中, θ 是 DQN 的网络参数, θ^{Target} 是目标 DQN 的网络参数, γ 是一个奖励系数折扣因子。

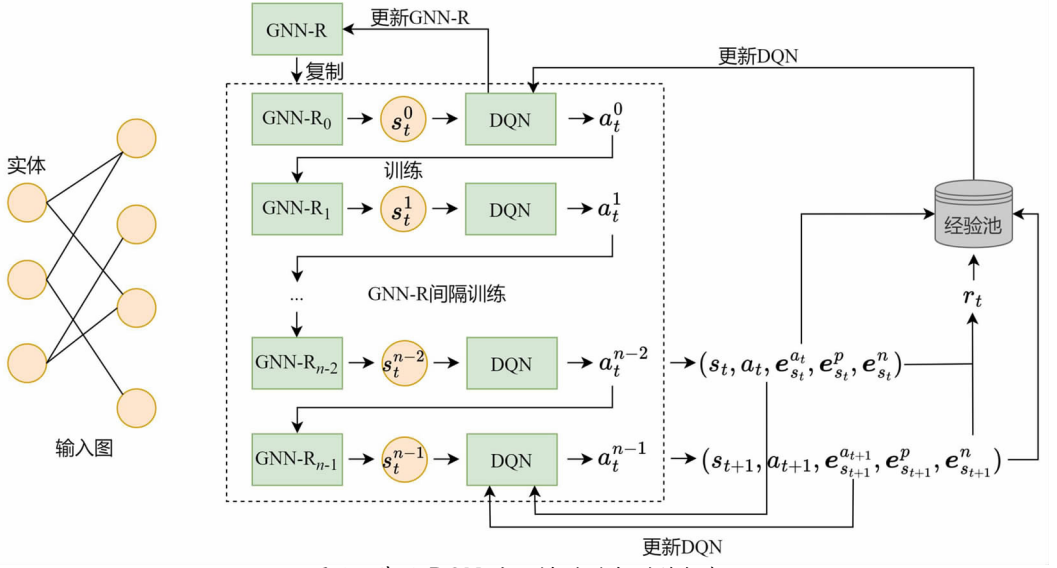


图 1 基于 DQN 的双模型联合训练框架

算法 1 双模型框架 DQN 的训练过程

```

1: 输入: 用户集  $U$ , 项目集  $I$ , 交互集  $T$ , DQN 模型  $Q(s, a)$ , GNN-R 模型
2: 将全局 GNN-R 网络参数复制到局部 GNN-R 模型 GNN -  $R_L$  中
3: 初始化列表:  $L_s, L_a, L_r, L_{s_a}, L_p, L_n$ , 分别用于存储状态  $s_t$ 、动作  $a_t$ 、奖励  $r_t$ 、状态-动作嵌入表示  $e_{s_t}^{a_t}$ 、正向交互嵌入表示  $e_t^p$  和负向交互嵌入表示  $e_t^n$ 
4: 从交互集获取正向交互项目  $i_p$ , 从非交互集随机选取负向交互项目  $i_n$ 
5: for step = 0, 1, ..., n-1 do // 共  $n$  步
6:   for  $u \in U$  do
7:     获取  $s_t = \text{GNN} - R_L^0(u)$ 
8:     以  $\epsilon$  概率随机选择  $a_t$ , 否则获取动作  $a_t = \arg\max_a Q(s_t, a)$ 
9:     获取状态-动作嵌入表示  $e_{s_t}^{a_t} = \text{GNN} - R_{q_L}^0(u)$ ,  $\text{GNN} - R_{q_L}$  表示 GNN -  $R_L$  的  $a_t$  向嵌入表示
10:    获取正向交互嵌入表示  $e_t^p = \text{GNN} - R_L(i_p)$ 
11:    获取负向交互嵌入表示  $e_t^n = \text{GNN} - R_L(i_n)$ 
12:    将  $s_t, a_t, e_{s_t}^{a_t}, e_t^p$  和  $e_t^n$  分别存入  $L_s, L_a, L_{s_a}, L_p, L_n$  列表
13:   end for
14:   使用当前步的  $s_t, a_t, e_t^p$  和  $e_t^n$  训练 GNN -  $R_L$ 
15: end for
16: 从  $L_s, L_{s_a}, L_p, L_n$  获取第  $n-1$  步的数据作为  $s_t, e_{s_t}^{a_t}, e_t^p$  和  $e_t^n$ 
17: 从  $L_s, L_{s_a}, L_p, L_n$  获取第  $n$  步的数据作为  $s_{t+1}, e_{s_{t+1}}^{a_{t+1}}, e_{s_{t+1}}^p$  和  $e_{s_{t+1}}^n$ 
18: 使用式(3.3)获取  $r_t$ 
19: 将  $r_t$  存入  $L_r$ 
20: 立即使用  $(s_t, a_t, r_t, s_{t+1})$  数据训练 DQN
21: 将  $(s_t, a_t, r_t, s_{t+1})$  经验数据存入经验池
22: 从经验池采样批量  $(s_t, a_t, r_t, s_{t+1})$  数据训练 DQN

```

2.2 间隔经验回放与延后动作奖励机制

基于上述介绍的 DQN 训练过程,本文提出了两个新的学习机制:

1. 间隔经验回放机制

双模型联合训练时,本文使用一个专用于为 DQN 收集经验的局部 GNN-R 模型,基于 DQN 模型给出的优化指导,训练局部的 GNN-R 模型 n 步,并使用第 $n-1$ 步和第 n 步的实体嵌入作为经验数据。这种方式可以在一定程度上缓解 GNN-R 模型连续训练导致的参数振动问题。

2. 延后动作奖励机制

当 DQN 给出动作优化指导时,并不立即基于当前状态和动作计算奖励,而是做出动作后训练主 GNN-R 模型,使其进入下一个状态才延迟计算奖励。这时缓存列表里存在两种不同的状态和动作,将这两种不同状态得分的差值作为当前状态做出动作后的奖励。与 DPAO 只优化基于 GNN 层数的聚合不同,这种奖励机制更具有普适性,可以优化主模型在训练过程中的关键参数。

3 基于 VRKG 可自适应聚合优化的推荐算法

VRKG4Rec 模型对知识图中具有潜在相关性的关系进行聚类以生成虚拟关系,通过无监督学习方案构建虚拟关系知识图 (Virtual Relational Knowledge Graph, VRKG) 来缓解长尾分布问题。然而,在 VRKG4Rec 中需要手动设置虚拟关系的

数量,基于 DQN 的双模型联合训练框架可以解决这一问题。同时,该联合训练框架也能够让 VRKG4Rec 在 GNN 层(LWS 层)上进行信息传播和聚合时能够自适应地选择某一层上的用户嵌入表示,优化 VRKG4Rec 的推荐性能。本节将应用基于 DQN 的双模型联合训练框架来优化 VRKG4Rec 中的 LWS 层数和 VRKG 的数量,为便于介绍,将两个优化模型简称为 LWS 模型和 VR 模型。模型训练整体流程如图 2 所示:

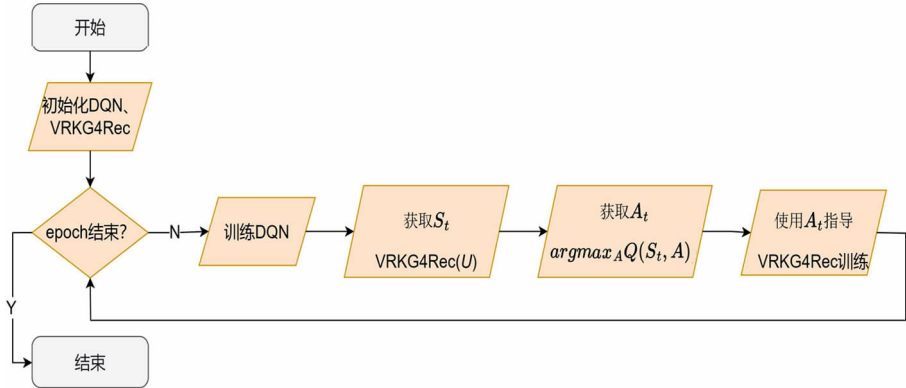


图 2 双模型框架整体训练流程

其中,DQN 模型的训练过程已在 2.1 小节中做出详述。为了弥合初始状态与各层状态间的数量级方面可能存在的差距,LWS 模型使用第 0 层的实体嵌入作为实体的初始状态,在第 0 层 LWS 操作中跳过传播聚合过程,仅仅保留 LWS 的平滑操作。VR 模型的初始状态为项目在所有虚拟关系数量聚合嵌入表示中的均值。之后以实体的状态(S_t)从 DQN 中获取动作指导,并在后续计算损失时使用对应的动作层数聚合得到的实体表示。VRKG4Rec 模型的 Loss 函数使用 BPR 损失式^[13],如式(3):

$$L_{BPR} = \sum_{i \in I(u), j \notin I(u)} - \ln \sigma[\text{Score}(e_u^*, e_i^*) - \text{Score}(e_u^*, e_j^*)] \quad (3)$$

其中, σ 为激活函数,Score 是一个内积函数, (u, i) 是交互集 I 中的交互对,而 (u, j) 不在 I 中。

4 实验与分析

4.1 实验数据集

本文实验使用以下两个公开数据集:
1. Last-fm:数据来自 Last. fm 在线音乐系统,包

含用户社交网络、用户应用标签和音乐收听信息。数据集收集了大约 2000 名用户的音乐收听记录。

2. MovieLens:是一个基于电影评分构成的数据集,从 MovieLens 网站收集,包含用户信息、电影信息及用户对电影的评分(0~5 分)。此数据集由 GroupLens 研究组在明尼苏达大学中组织收集,包含 1 M、10 M 和 20 M 等规格,本文使用 MovieLens-1M 数据集。

上述两个公开数据集的基本数据量统计结果如表 1 所示。

4.2 评价指标

本文采用的评价指标为 Top-K 推荐中被广泛使用的召回率、归一化折损累计增益和命中率^[1,14]。

召回率定义如式(4):

$$\text{Recall@K} = \frac{\sum_{u=1}^N |R(u) \cap T(u)|}{\sum_{u=1}^N |T(u)|} \quad (4)$$

其中, N 为用户数, $R(u)$ 是推荐系统给用户 u 推荐的项目列表, $T(u)$ 是用户 u 在测试集上实际存在交互的项目列表。

表 1 数据集数据量统计

数据集	用户-项目交互数据			知识图数据		
	用户数	项目数	交互数	实体数	关系数	三元组数
Last-fm	1872	3846	42346	9366	60	15518
MovieLens-1M	6036	2347	753772	6729	7	95

归一化折损累计增益定义如式(5):

$$\text{NDCG@K} = \frac{\text{DCG@K}}{\text{IDCGV@K}} \tag{5}$$

其中,DCG 定义如式(6):

$$\text{DCG@K} = \sum_{i=1}^K \frac{\text{rel}_i}{\log_2(i+1)} \tag{6}$$

其中,rel_{*i*} 为第 *i* 个推荐项目的真实相关性分数。IDCG 定义如式(7):

$$\text{IDCG@K} = \sum_{i=1}^{|\text{REL}|} \frac{\text{rel}_i}{\log_2(i+1)} \tag{7}$$

其中,IDCG 是指 Ideal DCG,它会对推荐项目按相关性评分排序后再进行计算。

命中率定义如式(8),当推荐列表中至少包含

一个在测试集上与用户存在交互的项目时,视为命中。

$$\text{hr}(u) = \begin{cases} 1, & R(u) \cap T(u) \neq \emptyset \\ 0, & R(u) \cap T(u) = \emptyset \end{cases}$$
$$\text{HR@K} = \frac{\sum_{u=1}^N \text{hr}(u)}{N} \tag{8}$$

4.3 实验结果

表 2 和表 3 为 VRKG4Rec 和 LWS 模型在 Last-fm 和 MovieLens 数据集上的性能对比。

表 2 LWS 模型 Top-1 和 Top-5 实验结果对比

数据集	模型	Top-1/%			Top-5/%		
		Recall	NDCG	HR	Recall	NDCG	HR
Last-fm	VRKG4Rec	6.83	16.70	16.70	20.16	17.76	40.14
	Ours	7.49	17.66	17.66	20.36	18.25	40.75
	Improv.	+9.66	+5.75	+5.75	+0.99	+2.76	+1.52
MovieLens	VRKG4Rec	4.32	36.66	36.66	15.02	31.26	70.23
	Ours	4.41	37.24	37.24	15.08	31.67	70.43
	Improv.	+2.08	+1.53	+1.53	+0.33	+1.31	+0.31

表 3 LWS 模型 Top-10 和 Top-20 实验结果对比

数据集	模型	Top-10/%			Top-20/%		
		Recall	NDCG	HR	Recall	NDCG	HR
Last-fm	VRKG4Rec	28.82	21.25	51.72	38.83	24.54	62.21
	Ours	28.81	21.66	51.60	39.19	25.03	62.33
	Improv.	−0.03	+1.93	−0.23	+0.93	+2.00	+0.19
MovieLens	VRKG4Rec	23.25	30.43	80.79	34.10	31.87	88.19
	Ours	23.36	30.67	80.99	34.36	32.15	88.34
	Improv.	+0.47	+0.79	+0.25	+0.76	+0.88	+0.17

由表 2 和表 3 可知,相比于 VRKG4Rec,LWS 模型在数据集 Last-fm 和 MovieLens 上的大部分推荐性能皆占优势。尤其是在 Last-fm 数据集的 Top-1 推荐上,Recall 提升了 9.66%、NDCG 和 HR 提升了 5.75%。这表明 LWS 模型在对用户

进行精确推荐时存在较大优势,当推荐列表长度增大时,推荐效果有所下降,这是由于 DQN 在对实体进行嵌入学习指导时引入了一定的噪声干扰。

VR 模型在 Last-fm 和 MovieLens 数据集上的实验结果如表 4 和表 5 所示:

表 4 VR 模型 Top-1 和 Top-5 实验结果对比

数据集	模型	Top-1/%			Top-5/%		
		Recall	NDCG	HR	Recall	NDCG	HR
Last-fm	VRKG4Rec	6.83	16.70	16.70	20.16	17.76	40.14
	Ours	7.23	17.30	17.30	20.92	18.49	41.35
	Improv.	+5.86	+3.59	+3.59	+3.77	+4.11	+3.01
MovieLens	VRKG4Rec	4.32	36.66	36.66	15.02	31.26	70.23
	Ours	4.41	37.61	37.61	15.21	31.83	70.49
	Improv.	+2.08	+2.59	+2.59	+1.26	+1.82	+0.37

表 5 VR 模型 Top-10 和 Top-20 实验结果对比

数据集	模型	Top-10/%			Top-20/%		
		Recall	NDCG	HR	Recall	NDCG	HR
Last-fm	VRKG4Rec	28.82	21.25	51.72	38.83	24.54	62.21
	Ours	29.19	21.83	51.66	39.00	25.08	62.03
	Improv.	+1.28	+2.73	−0.12	+0.44	+2.20	−0.29
MovieLens	VRKG4Rec	23.25	30.43	80.79	34.10	31.87	88.19
	Ours	23.36	30.81	80.74	34.19	32.16	88.24
	Improv.	+0.47	+1.25	−0.06	+0.26	+0.91	+0.06

由表 4 和表 5 可知,VR 模型在两个数据集上的推荐性能表现大部分都优于 VRKG4Rec 模型。与 LWS 模型实验结果类似,VR 模型在 Last-fm 数据集的 Top-1 推荐上效果提升最明显,Recall 提升了 5.86%、NDCG 和 HR 提升了 3.59%。

5 结 论

推荐系统作为连接用户与信息的关键桥梁,其应用面临着诸多挑战,尤其在处理冷启动和数据稀疏性问题时显得力不从心。为解决这些问题,科研和工业界不断引入新的算法和技术。将知识图谱引入推荐算法,不仅能够提高推荐算法的准确性,还能够发掘图上的潜在高阶关联关系,增强推荐的透明度和可解释性。图神经网络的引入使得推荐模型能够更好地挖掘图结构数据中的高阶信息,融合用户信息、项目信息以及交互信息,从而学习到用户和项目的更为复杂的表示。然而,这些新技术的引入也带来了新的问题,如知识图谱中关系的长尾分布和图神经网络中存在的实体嵌入表示学习的粒度限制问题。

在此背景下,提出了一种基于 DQN 的双模型联合训练框架,以缓解双模型联合训练时环境振荡带来的问题。在此框架中提出了新的强化学习方案,能够自适应地优化推荐模型,提高推荐的个性化和准确性。针对 VRKG4Rec 模型存在的问题,提出了能够自适应优化相关参数的推荐算法,使用 DQN 模型自适应地调整 LWS 聚合层数和虚拟关系数量,从而提升 VRKG4Rec 模型的性能。实验结果表明,提出的两个优化模型在数据集 Last-fm 和 MovieLens 上的大部分推荐性能较 VRKG4Rec 模型都有所提升。本文的工作为推荐算法的发展提供了新的思路和方法,而解决本文提出的算法存在的复杂度过高从而导致计算效率低的问题,提升算法的泛化能力以及加强算法的可解释性,将是未来工作的重点研究方向。

参考文献

[1] BOBADILLA J, ORTEGA F, HERNANDO A, et al. Recommender systems survey[J]. Knowledge-based Systems, 2013, 46: 109–132.

[2] SINGHAL A. Introducing the knowledge graph: things, not strings[J]. Official Google Blog, 2012, 5(16): 3.

[3] ZHOU J, CUI G Q, HU S D, et al. Graph neural networks: a review of methods and applications[J]. AI Open, 2020, 1: 57–81.

[4] JUNG H, KIM S, PARK H. Dual policy learning for aggregation optimization in graph neural network-based recommender systems[C]//Proceedings of the ACM Web Conference 2023. 2023: 1478–1488.

[5] MNIH V, KAVUKCUOGLU K, SILVER D, et al. Human-level control through deep reinforcement learning[J]. Nature, 2015, 518(7540): 529–533.

[6] LU L Y, WANG B, ZHANG Z Z, et al. VRKG4Rec: virtual relational knowledge graph for recommendation[C]//Proceedings of the Sixteenth ACM International Conference on Web Search and Data Mining. 2023: 526–534.

[7] SUN Z, GUO Q, YANG J, et al. Research commentary on recommendations with side information: a survey and research directions[J]. Electronic Commerce Research and Applications, 2019, 37: 1–29.

[8] XIAN Y K, FU Z H, MUTHUKRISHNAN S, et al. Reinforcement knowledge graph reasoning for explainable recommendation[C]//Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval. 2019: 285–294.

[9] SZEGEDY C, LIU W, JIA Y Q, et al. Going deeper with convolutions[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015: 1–9.

[10] PEROZZI B, AI-RFOU R, SKIENA S. Deepwalk: online learning of social representations[C]//Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2014: 701–710.

[11] 杜晓宇, 陈正, 项欣光. 基于解耦图神经网络的可解释标签感知推荐算法[J]. 软件学报, 2023, 34(12): 5670–5685.

[12] 金天成, 窦亮, 肖春芸, 等. 记忆与认知融合的个性化 OJ 习题推荐方法[J]. 计算机学报, 2023, 46(1): 103–124.

[13] RENDLE S, FREUDENTHALER C, GANTNER Z, et al. BPR: Bayesian personalized ranking from implicit feedback [C]//Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence. 2009: 452–461.

[14] WU S W, SUN F, ZHANG W T, et al. Graph neural networks in recommender systems: a survey[J]. ACM Computing Surveys, 2022, 55(5): 1–37.