计算机科学
COMPUTER SCIENCE

基于知识图谱的层次粒化推荐方法

秦琪琦，张月琴，王润泽，张泽华

引用本文

秦琪琦，张月琴，王润泽，张泽华. 基于知识图谱的层次粒化推荐方法[J]. 计算机科学, 2022, 49(8): 64-69.

QIN Qi-qi, ZHANG Yue-qin, WANG Run-ze, ZHANG Ze-hua. Hierarchical Granulation Recommendation Method Based on Knowledge Graph[J]. Computer Science, 2022, 49(8): 64-69.

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)

基于自注意力机制和迁移学习的跨领域推荐算法

Cross-domain Recommendation Algorithm Based on Self-attention Mechanism and Transfer Learning

计算机科学, 2022, 49(8): 70-77. https://doi.org/10.11896/jsjkx.210600011

基于神经架构搜索的点击率预测模型

Click-Through Rate Prediction Model Based on Neural Architecture Search

计算机科学, 2022, 49(7): 10-17. https://doi.org/10.11896/jsjkx.210600009

基于概率元学习的矩阵补全预测融合算法

Fusion Algorithm for Matrix Completion Prediction Based on Probabilistic Meta-learning

计算机科学, 2022, 49(7): 18-24. https://doi.org/10.11896/jsjkx.210600126

基于自适应注意力机制的知识图谱补全算法

Adaptive Attention-based Knowledge Graph Completion

计算机科学, 2022, 49(7): 204-211. https://doi.org/10.11896/jsjkx.210400129

知识图谱推理研究综述

Review of Reasoning on Knowledge Graph

计算机科学, 2022, 49(6A): 74-85. https://doi.org/10.11896/jsjkx.210100122

计算机科学 COMPUTER SCIENCE
http://www.jsjkx.com
DOI:10.11896/jsjkx.210600111

基于知识图谱的层次粒化推荐方法

秦琪琦　张月琴　王润泽　张泽华
太原理工大学信息与计算机学院　太原 030024
(qinqiqi_q@163.com)

摘　要　基于图神经网络的推荐系统是当前数据挖掘应用的研究热点。在异质信息网络(Heterogeneous Information Network,HIN)上结合图神经网络进行推荐,可通过用户的关联信息来学习用户的偏好,从而提升推荐性能。但现有基于HIN的推荐方法大多存在不能有效地解释高阶建模结果及人工设计元路径需要相关领域知识的问题。因此,结合层次粒化思想,在异质推荐过程中引入知识图谱,提出一种基于知识图谱的异质推荐方法(Heterogeneous Recommendation Methods for Knowledge Graphs,HKR)。该方法首先结合知识图谱,对局部上下文和非局部上下文进行层次粒化,分别学习用户特征的粗粒度表示;然后基于门控机制结合局部和非局部的属性节点嵌入,进一步学习用户和项目之间的潜在特征;最后将细粒度的特征融合用于推荐。在真实的大规模数据集上的实验结果表明,所提方法的性能在多方面评测上均优于目前的基于知识图谱的图神经网络推荐方法。

关键词:层次粒化;知识图谱;多粒度融合;推荐系统
中图法分类号　TP391

Hierarchical Granulation Recommendation Method Based on Knowledge Graph

QIN Qi-qi,ZHANG Yue-qin,WANG Run-ze and ZHANG Ze-hua
College of Information and Computer,Taiyuan University of Technology,Taiyuan 030024,China

Abstract　The recommendation system based on graph neural network is the current research hotspot of data mining applications. The recommendation performance can be improved by combining the graph neural network on the heterogeneous information network(HIN). However,the existing HIN-based recommendation methods often have problems that cannot effectively explain the results of high-level modeling,and manual design of meta-paths requires knowledge of related domains. Therefore,this paper combines the idea of hierarchical granulation and proposes a heterogeneous recommendation method(HKR) based on knowledge graphs. The local context and non-local context are hierarchically granulated,and the coarse-grained representation of user characteristics is learned separately. Then based on the gating mechanism, combining local and non-local attribute node embedding,learning the potential features between users and items,and finally fusing fine-grained features for recommendation. The real experimental results show that the performance of the proposed method is better than the current graph neural network recommendation method based on knowledge graph in many aspects.

Keywords　Hierarchical granulation,Knowledge graph,Multi-granularity fusion,Recommendation system

互联网技术的日渐成熟,使网络成为人们获取信息以及消费的主要渠道,如何快速地从海量数据中获取用户感兴趣的信息成为科技领域关注的焦点。个性化推荐技术通过分析用户的历史活动数据来挖掘用户的潜在偏好[1]。现存的推荐方法主要有:基于内容的推荐、协同过滤推荐、基于异质信息网络的推荐。

复杂的异构信息网络能够对各种异构数据进行灵活的建模,因此被用来建模丰富的辅助数据,应用于推荐系统中[2-3]。其中,基于正则化的方法设计了捕获HIN中的附加损失项来正则化推荐器模型学习;CFKG[4]和KTUP[5]通过共享项目嵌入来联合训练推荐任务和关系补全任务,以隐式的方式对高阶关系进行编码,从而优化推荐模型。基于路径的方法提取路径上实体之间的高阶连通性信息,并将它们输入到预测模型,使用路径选择算法来选择显著路径[6-7],或者通过元路径模式约束路径[8-9]。

虽然上述方法在一定程度上提升了推荐性能,但是,基于正则化的方法由于缺乏明确的建模,因此不能解释高阶建模的结果,也无法捕捉实体之间多跳的依赖关系;基于路径的

到稿日期:2021-04-06　返修日期:2021-06-13
基金项目:国家自然科学基金(61503273,61702356);教育部产学合作协同育人项目;山西省回国留学人员科研资助项目
This work was supported by the National Natural Science Foundation of China(61503273,61702356),Industry-University Cooperation Education Program of the Ministry of Education and Shanxi Scholarship Council of China.
通信作者:张月琴(zehua_zhang@163.com)

方法对于具有不同类型关系和实体的复杂HIN来说，定义有效的元路径需要领域知识，会耗费大量的精力，此外，当路径稀疏或者混乱时，其很难进行有效的推荐。

2012 年 Google[10]正式提出知识图谱这个术语，知识图谱用于描述真实世界中存在的实体或概念，以及两者之间的关联关系；Dodwad 等[11]基于知识图谱的层次结构对概念进行加权，更加细粒度地分析用户的偏好；Moreno 等[12]完成基于知识图谱和协同过滤的混合推荐系统 Sigtur/E-Destination，90% 以上的用户认为该系统的推荐较为准确；Noia 等[13]利用知识图谱丰富的语义特点来提高基于内容的电影推荐系统的准确率。鉴于知识图谱蕴含丰富的语义关系，可以有效提高推荐系统的准确率，本文将知识图谱引入推荐系统。然而，知识图谱是知识量丰富的异构网络，节点类型和关系类型都很丰富，因此在挖掘相关语义信息时难度较大。粒计算可以通过基于多层次多视角的问题求解方法，使复杂的问题简单化。因此，本文从粒计算的角度出发，提出了层次粒化知识图谱推荐方法，通过将知识图谱中的信息网络粒化为多个独立的粗粒度网络，然后基于门控机制结合局部和非局部的属性节点嵌入，来学习用户和项目之间的潜在特征，最终将细粒度的特征融合用于推荐。

本文的主要工作如下：

(1)在图神经网络的基础上，将中心节点的局部和非局部上下文信息进行层次粒化，学习最粗粒度下用户和项目的潜在特征，有效地避免了使用 HIN 进行推荐时人工设计元路径的问题；

(2)在捕获非局部上下文时，使用了递归嵌入传播，基于节点邻居的嵌入来更新节点的嵌入，并递归地执行这种嵌入传播，从而得到细粒度的特征表示；

(3)通过利用知识图谱的辅助信息进行推荐，提高了 HKR 模型的泛化能力。

1 相关工作

粒计算(Granular Computing)是一种认识客观世界的方法论。信息粒化通过将现实世界抽象化，来简化人类对外部世界的认识。如今这种思想被用在许多人工智能的研究领域，如聚类分析、数据挖掘、机器学习等方面。Zadah 等[14]首次提出了模糊信息粒化理论。Yao 等[15]对粒计算的基本方法、基本问题进行了系统的研究，并将其应用于机器学习、数据挖掘领域。Zhang 等[16]认为人们从极不相同的粒度上观察和分析同一问题是人类智能的表现，并在此基础上提出了商空间理论。Pedrycz[17]在粒计算和模糊数学领域进行了大量有成效的研究，对信息粒化算法、数学框架、不同模型下的应用做了总结。Miao 等[18]将粒计算视为一种粒化的思维方式及方法论，并且是一种独特的基于多层次多视角的问题求解方法。Wu 等[19]基于二元关系为研究 k 阶邻域算子系统和粗糙集近似建立了一个通用框架。Qian 等[20]为解决大数据有限的标记属性、计算效率低下以及属性约简中的过拟合等问题，提出一个局部粗糙集的理论框架，并开发了一系列具有线性时间复杂度的上下文近似和属性约简算法。

基于知识图谱的推荐方法可以分为两大类：基于正则化的方法和基于路径的方法。基于正则化的方法利用知识图谱的结构，将正则化项引入，用于学习实体嵌入的损失函数。协同知识图谱嵌入(Collaborative Knowledge Graph Embedding，CKE)是一种代表性的方法[21]，通过使用 TransR 从知识图谱中导出语义实体表示[22]。项目和实体之间的高阶特征相互作用可以通过交叉压缩单元进一步近似，这些方法非常灵活，然而它们缺乏对语义关系的明确建模。基于路径的方法利用了实体之间的各种连接模式。在解决人工设计元路径的局限性问题上，已有不同的选择规则和传播方法被提出[23]，例如，通过限制路径长度来提取相关路径，然后聚集路径信息[6]。

从本质上讲，知识图谱是一个异构网络，其中，节点代表实体，边代表实体之间的关系。将知识图谱应用于推荐系统的主要挑战是如何有效地利用实体之间的关系及知识图谱的图形结构。Sun 等[6]指出项目之间具有丰富语义关系的知识图谱在提高推荐系统性能方面是有效的。Zhang 等[21]提出使用设计好的正则化损失项来捕获知识图谱的结构，但是其没有在推荐模型中明确考虑知识图谱的语义关系信息。Yu 等[2]提出通过提取路径上实体之间的高阶连通性来探索实体之间的关系，但是这些信息是基于特殊标准手动设计或选择的，严重依赖于相关领域的知识。

2 基于知识图谱的粒化推荐系统

定义 1(知识图谱) 知识图谱由实体及实体之间的关系构成，使用 $G=(V,E)$ 来表示，其中 V 代表节点集，E 代表边集。每个“头节点-关系-尾节点”用 (h,r,t) 表示，其中 $h,t\in V$，$r\in R$，每条边代表了从头节点到尾节点之间的关系。使用 e_h,e_r,e_t 分别表示 h,t,r 的嵌入。

定义 2(结构粒化) 结构粒化指给定某异质结构的知识图谱 G，通过引入元路径来处理知识图谱的关系分类问题。元路径指任意一条包含关系类型序列的的路径，这条路径的长度为其连接的节点个数。有了元路径，就可以根据邻居节点与头节点的距离将邻居节点分类，进而将知识图谱分解为多通道网络。元路径的集合在形式上可以表示为：

$$S=\{P_1,P_2,P_3,\cdots,P_{|s|}\}$$

以上对节点进行分类的过程就是对知识图谱粒化的过程。此时，一个知识图谱可以转化为如下的多通道网络：

$$G'=\{G_t'\mid G_t'=(V_1,\varepsilon_{1n}),n=1,\cdots,|S|\}$$

其中，$V_1,V_2,\cdots,V_m$ 代表 m 种不同的节点类型分别组成的集合，本文称之为粒。

由于知识图谱包含众多的异质信息，为了充分挖掘邻居节点包含的丰富信息，本文将知识图谱进行粒化，分别学习不同粒度的向量表示，最后将其融合，得到用户最终的偏好。

3 模型提出

本文提出用于聚集实体上下文的层次粒化知识图谱推荐方法模型 HKR，其整体结构如图 1 所示。图 1(b)由 3 个组件构成：嵌入层、注意力嵌入传播层、聚合层。

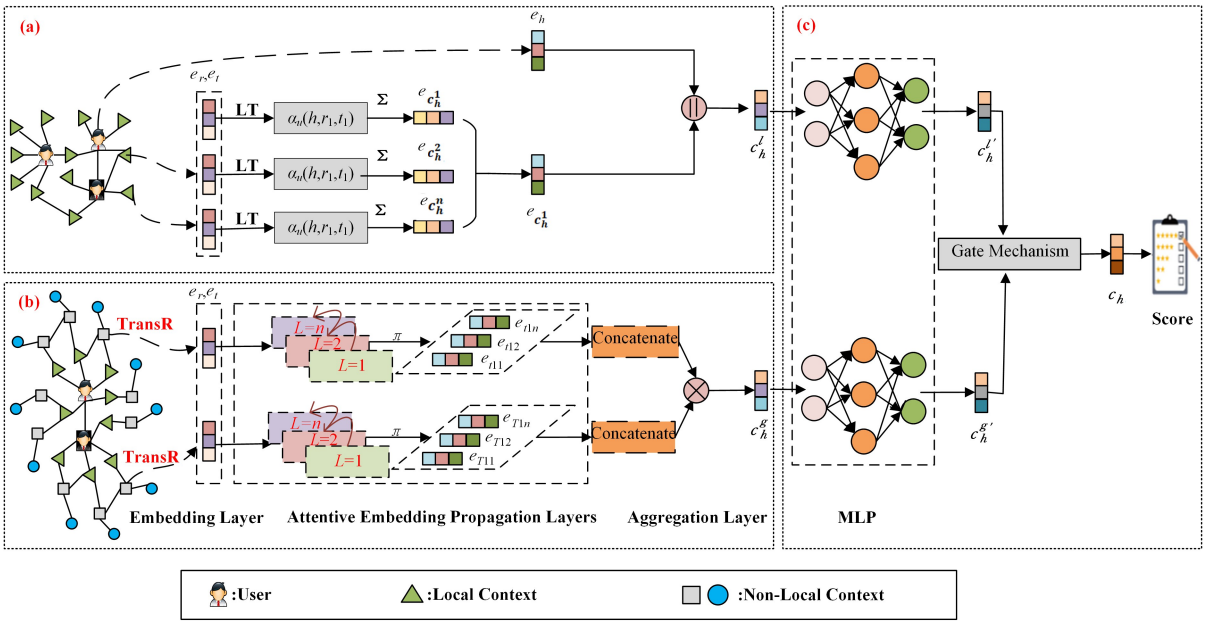

图 1 基于知识图谱的异质推荐方法的整体模型图

Fig. 1 HKR overall model diagram

3.1 局部上下文

对于一个项目实体，它与许多其他实体相连接，这些实体通过知识图谱来丰富信息。出于对用户个性化偏好的考虑，提出了一种特定于用户的知识图谱注意力机制，用于聚合知识图谱中实体的邻域信息。对于不同用户，为同一邻域实体分配不同的权重，基于用户特定的注意力分数聚集邻域实体的嵌入。

实体通过不同的关系与邻居节点相连接，这对用户实体的选择有一定的影响，因此，本文将与用户实体相邻的节点 $t\in c_h^l$ 与对应关系 r 的嵌入进行如下线性变换：

$$e_{rt}=(e_r \parallel e_t)\boldsymbol{W}_0$$

其中，$\parallel$ 为串联操作，$\boldsymbol{W}_0$ 为权重矩阵。

HKR 的局部上下文模型图如图 1(a)所示。将 $C_h^l=\{t\mid(h,r,t)\in D\}$ 作为实体 h 的局部邻域表示，定义 C_h^l 为局部上下文信息。对于用户 u，实体 t 对实体 h 重要性的用户注意得分 $\alpha_u(h,r,t)$ 定义为：

$$\alpha_u(h,r,t)=\frac{\exp[\pi_u(h,r,t)]}{\sum\limits_{(h,\tilde{r},\tilde{t})\in\mathcal{G}}\exp[\pi_u(h,\tilde{r},\tilde{t})]} \tag{1}$$

其中，$\pi_u(h,r,t)$ 由单层前馈神经网络执行，定义如下：

$$\pi_u(h,r,t)=\tanh[(\boldsymbol{e}_h \parallel e_{rt})\boldsymbol{W}_1+\boldsymbol{b}_1]\boldsymbol{m}_u^{\mathrm{T}} \tag{2}$$

其中，$\boldsymbol{e}_h$ 为头节点 h 的向量表示，$\boldsymbol{m}_u^{\mathrm{T}}$ 为用户向量表示 e_u 的非线性变换；$\boldsymbol{m}_u=\mathrm{ReLU}(e_u\widetilde{\boldsymbol{W}}_1+\widetilde{\boldsymbol{b}}_1)$，其中 $\widetilde{\boldsymbol{W}}_1\in\mathbb{R}^{d\times d}$，$\boldsymbol{W}_1\in\mathbb{R}^{2d\times d}$，$\widetilde{\boldsymbol{b}}_1\in\mathbb{R}^{1\times d}$，$\boldsymbol{b}_1$，$\widetilde{\boldsymbol{b}}_1$ 分别表示权重矩阵和偏置向量。

给定 h 的每个相邻实体的系数，计算其嵌入的线性组合，以获得 h 的局部邻域嵌入：

$$e_{c_h^l}=\sum_{t\in c_h^l}\alpha_u(h,r,t)e_t \tag{3}$$

将实体 h 的嵌入和局部邻居嵌入 $e_{c_h^l}$ 聚合，生成 h 的局部上下文信息：

$$c_h^l=\tanh[(\boldsymbol{e}_h \parallel e_{c_h^l}\boldsymbol{W}_2+\boldsymbol{b}_2] \tag{4}$$

其中，$\boldsymbol{W}_2\in\mathbb{R}^{2d\times d}$，$\boldsymbol{b}_2\in\mathbb{R}^{1\times d}$ 分别代表权重矩阵和偏置向量。

3.2 非局部上下文

HKR 探索非局部上下文信息主要包括 3 部分(见图 1(b))：嵌入层、注意力嵌入传播层和信息聚合层。

3.2.1 嵌入层

知识图嵌入是一种将实体和关系参数化为向量表示的有效方法。本文中使用 TransR 方法[22]，将图中的三元组通过 $e_h^r+e_r\approx e_t^r$ 来学习嵌入的每个实体和关系，其中 $e_h,e_t\in\mathbb{R}^d$，$e_r\in\mathbb{R}^k$，e_h^r 和 e_t^r 分别表示 e_h 和 e_t 在 r 空间中的投影表示。因此，对于给定的三元组 (h,r,t)，可信度得分定义如下：

$$g(h,r,t)=\|\boldsymbol{W}_re_h+e_r-\boldsymbol{W}_re_t\|_2^2 \tag{5}$$

其中，$\boldsymbol{W}_r\in\mathbb{R}^{k\times d}$ 是关系 r 的转换矩阵，将实体从 d 维实体空间投影到 k 维关系空间。

3.2.2 注意力嵌入传播层

本文基于图卷积网络架构[25]进行建模，沿着高阶连通性递归传播嵌入；此外，利用图注意力网络[26]的思想，生成级联传播注意力权重来揭示这种连通性的重要性。以描述一个单层为例，其由信息传播层、知识感知注意力层和信息聚合层 3 个部分组成，讨论如何将其推广到多个层。

定义 3(信息传播) 对于一个实体，使用 $N_h=\{(h,r,t),(h,r,t)\in\mathcal{G}\}$ 来表示三元组的集合，其中 h 是头实体。为了刻画实体的一阶连通结构，计算了 h 的自我中心网络的线性组合：

$$e_{\mathcal{N}_h}=\sum_{(h,r,t)\in\mathcal{N}_h}\pi(h,r,t)e_t \tag{6}$$

其中，$\pi(h,r,t)$ 代表了每次在 (h,r,t) 上传播的衰减因子，表示有多少信息经过关系 r 从 h 传播到 t。

定义 4(知识感知注意) 通过关系注意机制来实现 $\pi(h,r,t)$，具体定义如下：

$$\pi(h,r,t)=(\boldsymbol{W}_re_t)^{\mathrm{T}}\tanh((\boldsymbol{W}_re_h+e_r)) \tag{7}$$

使用 tanh[27]作为非线性激活函数，使注意力分数依赖于 $\boldsymbol{e}_h$ 和 $\boldsymbol{e}_t$ 在关系 r 空间中的距离。之后采用 softmax 函数归一化与

h 连接的所有三元组的系数：

$$\pi(h,r,t)=\frac{\exp(\pi(h,r,t))}{\sum_{(h,r',t')\in\mathcal{N}_h}\exp(\pi(h,r',t'))} \tag{8}$$

最终的注意力得分可以为选择邻居节点提供指导。与 GCN[25] 和 Graph sage[26] 将两个节点之间的损失率设置为 $1/\sqrt{|\mathcal{N}_h||\mathcal{N}t|}$ 或 $1/|\mathcal{N}_t|$ 相比，HKR 不仅利用了图的近邻结构，还指定了不同邻居的重要性。特别地，与仅以节点表示作为输入的图注意力网络不同，本文对节点之间的关系进行建模，在传播过程中聚集更多信息。

3.2.3 信息聚合层

本节将实体表示 e_h 及自我表示注意网络 $e_{\mathcal{N}_h}$ 聚合为实体 h 的新表示 $c_h^g=f(e_h,e_{\mathcal{N}_h})$，通过以下两种聚合器实现。

(1)GCN 聚合器[12] 将两种表示形式相加，并使用了非线性变换：

$$f_{\mathrm{GCN}}=\mathrm{LeakyReLU}(\boldsymbol{W}(e_h+e_{\mathcal{N}_h}) \tag{9}$$

此处激活函数设置为 LeakyReLU；$\boldsymbol{W}\in\mathbb{R}^{d'\times d}$ 是可训练权重矩阵，用于提取传播时有用的信息，d' 用于变换大小。

(2)图形聚合器将两种表示串联起来进行非线性转换：

$$\boldsymbol{f}_{\mathrm{GraphSage}}=\mathrm{LeakyReLU}(\boldsymbol{W}(e_h\parallel e_{\mathcal{N}_h})) \tag{10}$$

其中，$\parallel$ 是串联操作。

由于 Graphsage 在模型训练好之后可以对新加入网络中的节点进行推理，有效避免节点特征需要重新训练的问题[27]，因此本文使用 GraphSage 聚合器实现聚合。之后，在得到节点的局部特征 c_h^l 和非局部特征 c_h^g 后，经过 MLP 函数变换得到最终的局部上下文表示 $c_h^{l'}$ 和非局部上下文表示 $c_h^{g'}$。

定义 5(高阶传播) 堆叠更多传播层以收集更高阶邻居的信息，此处实体可递归地表示为：

$$e_h^{(l)}=f(e_h^{(l-1)},e_{\mathcal{N}_h}^{(l-1)}) \tag{11}$$

其中，实体 h 在自我网络内的信息传播定义为：

$$e_{\mathcal{N}_h}^{(l-1)}=\sum_{(h,r,t)\in\mathcal{N}_h}\pi(h,r,t)e_t^{(l-1)} \tag{12}$$

其中，$e_t^{(l-1)}$ 是之前的信息传播步骤中产生的实体 t 的表示，存储来自于其 $l-1$ 跳的邻居信息，$e_h^{(0)}$ 是 e_h 的初始信息传播迭代，有助于实体 h 在 l 层的表示。因此，嵌入传播过程能捕捉到高阶连通性。

使用与式(4)中相同的聚合器参数，给定 h 在知识图谱中局部和非局部上下文的嵌入，通过学习每个维度的权重，使用门控机制聚合两部分的嵌入。

$$c_h=\sigma(\boldsymbol{\omega})\odot c_h^{l'}+(1-\sigma(\omega))\odot c_h^{g'} \tag{13}$$

其中，$\boldsymbol{\omega}\in\mathbb{R}^{1\times d}$ 是可学习向量，$\sigma(\odot)$ 表示 sigmoid 函数。由于项目是知识图谱中的特殊类型实体，考虑到知识图谱中局部上下文 c_i^l 和非局部上下文 c_i^g，使用式(13)来计算项目 i 的上下文嵌入 c_i。然后将 e_i 和 c_i 连接起来，以此来获取项目 i 的上下文表示 $q_i=(e_i\parallel c_i)$。

3.3 模型优化和预测

本文使用个性化排序常用的优化方法 Bayesian Personalized Ranking(BPR)来学习 HKR 的模型参数，其可以生成更高的排名分数。将 BPR 损失函数定义如下：

$$L_{\mathrm{BPR}}=\sum_{(u,i^+,i^-)\in\tilde{\mathcal{O}}}-\log\sigma(y_{ui^+}-y_{ui^-}) \tag{14}$$

其中，$\tilde{\mathcal{O}}$ 由负采样构成，对于每一个 $(u,i)\in\mathcal{O}$，随机抽取 5 个项目。由于需要学习 KG 中实体和关系的嵌入，设计了基于知识图谱的正则化损失。具体来说，对于每一个三元组 $(h,r,t)\in D$，首先通过定义以下分数描述头实体 h 和尾实体 t 在潜在空间关系 r 中的距离：

$$s_r(h,t)=\|e_h-e_{rt}\|_2^2 \tag{15}$$

然后定义正则化损失函数：

$$\mathcal{L}_{\mathrm{KG}}=\sum_{(h,r,t,t')\in\tilde{D}}\log\sigma(s_r(h,t)-s_r(h,t')) \tag{16}$$

对于每个 $(h,r,t)\in D$，$\tilde{D}$ 是通过随机抽样得到的。在潜在空间中，实体 h 与其直接相连的邻居 t 之间的距离应小于其与 h 不直接相连的实体 t' 之间的距离。通过求解以下目标函数来学习模型参数：

$$\min_{\Theta}\mathcal{L}_{\mathrm{BPR}}+\lambda_1 L_{\mathrm{KG}}+\lambda_2\|\Theta\|_2^2 \tag{17}$$

其中，Θ 代表层次粒化知识图谱推荐算法中的所有参数，λ_1 和 λ_2 是正则化参数。上式由梯度下降算法来解决。

在 HKR 的实现过程中，从一个目标实体 h 的邻域 c_h^l 中随机抽取 S 个邻居，从用户 u 的历史项目 I_u^+ 中抽取 N 个历史项目，计算式(1)的注意力权重。该方法可以使每个小批量的计算模式固定，并提高计算效率。

4 实验设计与分析

本文将所提算法与其他通用的推荐算法在两个真实的大规模数据集上进行比较和横向分析，并为算法设置不同的参数来进行纵向分析，以验证本文所提框架 HKR 的有效性。

4.1 数据集和测评指标

本实验基于 Movielens-1M 和 BookCrossing 两个公共数据集，由于数据具有稀疏性，本文将 BookCrossing 数据集上的所有评级都视为隐式反馈[23,25,30]，在 Movielens-1M 数据集中将大于 4 的评分视为隐式反馈。这些数据集的知识图谱是由 Microsoft Satori 构建的，只保留了整个知识图谱中置信度大于 0.9 的三元组[25]。表 1 列出了这些实验数据集的统计数据。

表 1 实验数据集统计

Table 1 Statistics of experimental dataset

	# Users	# Items	# Interactions	# Density/%	# Entities	# Relations	# Triples
Movielens-1M	6036	2347	376886	2.66	7008	7	20782
BookCrossing	17860	14967	69876	0.026	77903	25	151500

本文在衡量推荐性能时，对于实验中的每个数据集，随机选取 60% 的用户项目交互进行模型训练，选取 20% 的交互进行参数调整，剩下 20% 的交互用作测试数据。在信息检索中，Top-N 推荐任务通常采用类似的评估指标，本实验使用前 K 个精准率($P@K$)、前 K 个召回率($R@K$)和前 K 个点击率($HR@K$)作为评估指标。在实验中，设置 K 的值为 10

和 20。对于每个指标，首先根据测试数据计算每个用户的准确度，然后给出所有用户的平均准确度。

HKR 的潜在空间维度 d 从{8,16,32,64,128}中选择，在模型训练中，随机选取{2,5,8,13,18,36}作为实体本地邻居的数量 S 和用户的历史项目数量 N，正则化参数 λ_1 和 λ_2 从 $\{10^{-6},10^{-5},10^{-4},10^{-3},10^{-2}\}$ 中选择，学习率 η 设定为 10^{-4}。所有方法的最佳超参数由验证数据的性能决定。

4.2 基线模型

本实验使用 tensorflow 实现，使用 Adam 优化器学习模型参数。将 HKR 与以下模型进行比较。

(1)CFKG[4]：将多类型用户行为和项目知识图谱集成到统一的图中，并使用 TransE[28]学习实体嵌入。

(2)RippleNet[23]：利用知识图谱信息，沿着以历史项目为基准的路径，在实体集合上传播用户偏好。

(3)MKR[24]：一种使用知识图谱嵌入来辅助推荐的多任务特征学习方法。

(4)KGNN-LS[29]：将 GCN 应用在知识图谱上，通过传播和聚合项目的邻居信息来计算项目的嵌入。

4.3 与传统推荐算法比较

层次粒化知识图谱推荐方法基于 MovieLens 数据集与 4 种对比算法进行比较。图 2 表明：对于 $P@10$ 和 $R@10$ 两项评测指标而言，HKR 的性能明显优于其他 4 种模型；当评测指标相同时，K 值为 10 的性能优于 K 值为 20 的性能。与 CFKG 相比，层次粒化知识图谱推荐算法具有较好的推荐准确性，这是由于 CFKG 算法主要是通过结合传统协同过滤和基于知识图谱嵌入构建而成，没有融入实体的邻居信息。RippleNet 算法以与用户交互过的历史项目为基准，进行用户偏好传播，由于没有考虑模型中不同组件的不同重要性，因此准确性不高。KGNN-LS 算法是端到端的可训练框架，利用知识图谱捕获用户和项目之间的关系，但是在聚合高阶邻域信息方面没有明显优势。HKR 算法不仅融合了局部和非局部上下文信息，而且在捕获局部上下文信息时加入了注意力机制，在捕捉非局部上下文信息部分通过堆叠多个注意力嵌入传播层来有效地捕获协同信号。HKR 算法在 BookCrossing 数据集上与其他 4 种算法的性能比较如图 3 所示。尽管知识图谱在 BookCrossing 数据集上非常稀疏，HKR 算法仍具有较好的推荐效果。当评测指标相同时，K 值为 10 的性能明显优于 K 值为 20 的性能。

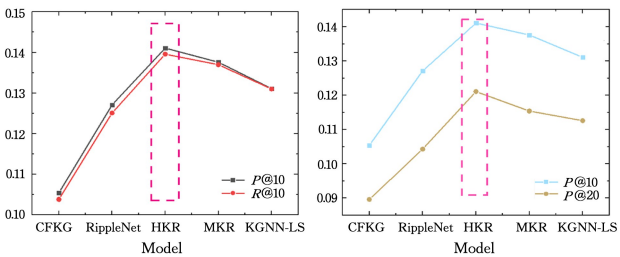

图 2 Movielens 数据集上不同模型的性能比较

Fig. 2 Performance comparison of different models on Movielens dataset

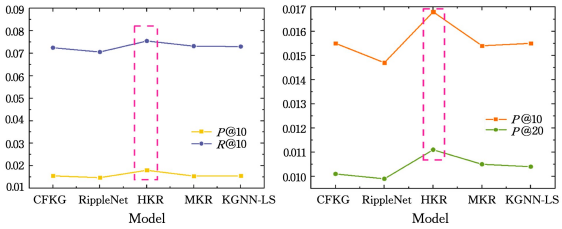

图 3 BookCrossing 数据集上不同模型的性能比较

Fig. 3 Performance comparison of different models on BookCrossing dataset

4.4 消融分析

通过消融研究分析层次粒化知识图谱推荐算法变体的性能。1)$HKR_{w/dL}$ 代表从原始模型中删除了局部上下文嵌入，仅将非局部上下文嵌入作为最终嵌入，即将式(13)中的 $\sigma(\omega)$ 设置为 0；2)$HKR_{w/dN}$ 代表从原始模型中删除了非本地上下文嵌入；3)$HKR_{w/dU}$ 代表删除了局部上下文信息中用户的嵌入，即删除式(2)中的 $\boldsymbol{m}_u$。

表 2 列出了 $HR@10$ 的消融分析准确率。以 $HR@10$ 的推荐准确率为例，HKR 的表现始终优于 $HKR_{w/dL}$ 和 $HKR_{w/dN}$，这表明知识图谱的本地和非本地上下文对于推荐结果都很重要。HKR 取得了比 $HKR_{w/dU}$ 更好的性能，表明用户特定的图形注意力机制比普通的注意力机制更适合于个性化推荐。$HKR_{w/dN}$ 的性能略优于 $HKR_{w/dL}$，表明在推荐中非局部上下文信息对局部上下文信息起着补充作用，即在提高推荐系统准确性方面，局部上下文信息更重要。

表 2 消融分析准确率

Table 2 Ablation analysis accuracy

Dataset	$HKR_{w/dL}$	$HKR_{w/dN}$	$HKR_{w/dU}$	HKR
Movielens-1M	0.6538	0.6602	0.6613	**0.6654**
BookCrossing	0.1384	0.1395	0.1326	**0.1477**

4.5 重要参数对模型的影响

HKR 包含一些重要的参数，在 Movielens-1M 数据集上进行实验，验证了 3 个参数的性能影响。

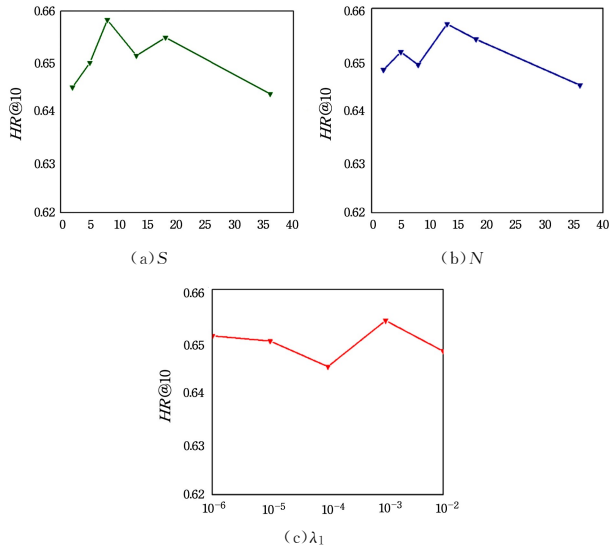

图 4 在 Movielens-1M 数据集上 HKR 的参数调整

Fig. 4 HKR parameter adjustment on Movielens-1M dataset

如图 4(a)所示，控制其他参数为定值，改变实体本地邻居的数量 S，当 S 为 8 时，$HR@10$ 取得了最好的性能。

同理，控制其他参数为定值，改变历史项目数量 N 和正则化参数λ_1，当 N 和λ_1分别取 13 和 10^{-3}时，$HR@10$ 取得最好的性能，实验结果如图 4(b)和图 4(c)所示。

结束语 本文针对知识图谱应用于推荐系统，提出一种新的层次粒化知识图谱推荐方法。该方法首先对知识图谱的局部上下文和非局部上下文进行层次粒化，分别学习两部分细粒度的用户特征表示；然后基于门控机制结合局部和非局部的属性节点嵌入，来学习用户和项目之间的潜在特征；最后，将细粒度的特征融合用于推荐。大量的实验结果证明了HKR 在推荐效果方面的有效性。

此外，层次粒化知识图谱推荐算法可以有效地学习不同层次粒度间的注意力权重，但是由于多层的注意力机制融合会提高算法复杂度，因此，未来的工作将进一步探索知识图谱应用于推荐系统的相关理论知识以及利用图神经网络等来构造异质网络的表征学习方法进行特征抽取。

参 考 文 献

[1] FOGELMAN-SOULIE F, MEI L, J ZHANG J, et al. Recommender Systems and Attributed Networks[M]. John Wiley & Sons, Ltd, 2020: 4.

[2] YU X, REN X, SUN Y Z, et al. Personalized entity recommendation: A heterogeneous information network approach[C]// WSDM. 2014: 283-292.

[3] ZHAO H, YAO Q M, LI J D, et al. Metagraph-based recommendation fusion over heterogeneous information networks[C]// Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2017.

[4] AI Q Y, AZIZI V, CHEN X, et al. Learning Heterogeneous Knowledge Base Embeddings for Explainable Recommendation [J]. Algorithms, 2018, 11(9): 137.

[5] CAO Y X, WANG X, HE X N, et al. Unifying Knowledge Graph Learning and Recommendation: Towards a Better Understanding of User Preferences[C]// WWW'19. ACM, 2019: 151-161.

[6] 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: 100879.

[7] WANG X, WANG D X, XU C R, et al. Explainable Reasoning over Knowledge Graphs for Recommendation[C]// AAAI. 2019.

[8] HU B B, SHI C, ZHAO W X, et al. Leveraging Meta-path-based Context for Top-N Recommendation with A Neural Co-Attention Model[C]// SIGKDD. 2018: 1531-1540.

[9] YU X, REN X, GU Q Q, et al. Collaborative filtering with entity similarity regularization in heterogeneous information networks [C]// IJCAI 27. 2013.

[10] LIU Q, LI Y, DUAN H, et al. Knowledge graph construction techniques[J]. Journal of Computer Research and Development, 2016, 53(3): 582-600.

[11] DODWAD P R, LOBO L. A context-aware recommender system using ontology based approach for travel applications[J]. International Journal of Advanced Engineering and Nano Technology, 2014, 1(10): 8-12.

[12] MORENO A, VALLS A, ISERN D, et al. SigTur/E-destination: Ontology based personalized recommendation of tourism and leisure activities[J]. Engineering Applications of Artificial Intelligence, 2013, 26(1): 633-651.

[13] DI NOIA T, MIRIZZI R, OSTUNI V C, et al. Linked open data to support content-based recommender systems[C]// Proceedings of the 8th International Conference on Semantic Systems. Graz, Austria: ACM, 2012.

[14] ZADAH L A. Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic[J]. Fuzzy Sets and Systems, 1997, 90(2): 111-127.

[15] YAO Y Y, NING Z. Potential Applications of Granular Computing in Knowledge Discovery and Data Mining[C]// Proceedings of World Multiconference on Systemics, Cybernetics and Informatics. 1999: 573-580.

[16] ZHANG B, ZHANG L. The Theory and Application of Problem Solving. [M]. Beijing: Tsinghua University Press, 1990.

[17] PEDRYCZ W. Granular Computing: An Emerging paradigm. Hardcover[M]. Physica-Verlag, 2001.

[18] MIAO D Q, WANG G Y, LIU Q, et al. Granular Computing: Past, Present and Prospects[M]. Beijing: Science Press, 2007.

[19] WU W Z, ZHANG W X. Neighborhood operator systems and approximations [J]. Information Sciences, 2002, 114(1/2/3/4): 201-217.

[20] QIAN Y, LIANG X, WANG Q. et al. Local rough set: a solution to rough data analysis in big data[J]. International Journal of Approximate Reasoning, 2018, 97: 38-63.

[21] ZHANG F Z, YUAN N J, LIAN D, et al. Collaborative Knowledge Base Embedding for Recommender Systems[C]// KDD'16. ACM, 2016: 353-362.

[22] LIN Y K, LIU Z Y, SUN M S, et al. Learning entity and relation embeddings for knowledge graph completion[C]// AAAI. 2015: 2181-2187.

[23] WANG H W, ZHANG F Z, WANG J L, et al. RippleNet: Propagating User Preferences on the Knowledge Graph for Recommender Systems[C]// CIKM. 2018: 417-426.

[24] WANG H W, ZHANG F Z, ZHAO M, et al. Multi-task feature learning for knowledge graph enhanced recommendation[C]// WWW'19. ACM, 2019: 2000-2010.

[25] KIPF T N, WELLING M. Semi-Supervised Classification with Graph Convolutional Networks[C]// ICLR'17. 2017.

[26] VELICKOVIC P, CUCURULL G, CASANOVA A, et al. Graph Attention Networks[C]// ICLR. 2018.

[27] HAMILTON W L, YING R, LESKOVEC J. Inductive Representation Learning on Large Graphs[J]. arXiv: 1706. 02216, 2017.

[28] BORDES A, USUNIER N, GARCIA-DURAN A, et al. Translating embeddings for modeling multi-relational data [C]// NeurIPS. 2013: 2787-2795.

[29] WANG H W, ZHANG F Z, ZHANG M D, et al. Knowledge graph convolutional networks for recommender systems with label smoothness regularization[C]// KDD'19. 2019.

QIN Qi-qi, born in 1996, postgraduate. Her main research interests include recommendation system and so on.

ZHANG Yue-qin, born in 1963, professor, master supervisor, is a member of China Computer Federation. Her main research interests include data mining, intelligent information processing and knowledge discovery on graph.

（责任编辑：柯颖）