

引用格式: 李政, 黄勃, 王晨明, 等. 基于异构注意力网络的知识图嵌入元路径推荐算法 [J]. 电子科技大学学报, 2025, 54(5): 776-788.
LI Z, HUANG B, WANG C M, et al. A knowledge graph embedded meta-path recommendation algorithm based on heterogeneous attention networks [J].
Journal of University of Electronic Science and Technology of China, 2025, 54(5): 776-788.

基于异构注意力网络的知识图嵌入元路径 推荐算法



李政, 黄勃*, 王晨明, 刘瑾, 曾国辉

(上海工程技术大学 电子电气工程学院, 上海 201600)

摘要: 针对现有模型在使用知识图嵌入方法时未进行有效约束, 导致噪声信息渗透到底层数据且基于元路径的推荐算法仅采用统一的权重分配策略, 忽视了元路径之间的细微差异, 同时传统模型也面临数据稀疏和冷启动问题。提出了一种基于异构注意力网络的知识图嵌入元路径推荐算法 (MRHAN), 通过定义元路径来捕捉不同类型实体和关系之间的复杂语义信息, 从而更好地利用丰富的异构信息来缓解数据稀疏和冷启动问题。在知识图嵌入过程中, 采用基于节点相关性的约束方法, 将语义相关的高阶实体和关系整合到唯一的元路径中。同时, 使用层次化注意力网络来建模节点对不同邻居和不同元路径的偏好, 学习节点与不同邻居之间的权重关系, 及元路径对不同推荐任务中的权重分配。实验结果表明, 该模型能够有效学习知识图谱的表示以及节点邻居和元路径对节点的重要性, 并进一步缓解数据稀疏和冷启动问题。

关键词: 知识图谱; 推荐系统; 注意力机制; 元路径

中图分类号: TP391

文献标志码: A

DOI: 10.12178/1001-0548.2024129

A knowledge graph embedded meta-path recommendation algorithm based on heterogeneous attention networks

LI Zheng, HUANG Bo*, WANG Chenming, LIU Jin, and ZENG Guohui

(School of Electronic and Electrical Engineering, Shanghai University of Engineering Science, Shanghai 201600, China)

Abstract: Existing models do not impose any constraints when using the knowledge graph embedding methods, which leads to noise permeating the underlying information along with the embedding. Meta-path-based recommendation algorithms only adopt a uniform weight assignment strategy without considering the nuances among meta-paths. Traditional models usually suffer from the problems of data sparsity and cold start. To address these issues, a knowledge graph embedded meta-path recommendation algorithm based on heterogeneous attention network (MRHAN) is proposed. By defining meta-paths to capture the complex semantic information between different types of entities and relationships, the proposed method can better utilize the rich heterogeneous information to alleviate the problems of data sparsity and cold start. Using knowledge graph embedding methods, a constraint approach based on node relevance is employed to model semantically related higher-order entities and relationships into unique meta-paths. A hierarchical attention network is used to model node preferences for different neighbors and different meta-paths, learning the weights of nodes for different neighbors and the weights of meta-paths for different recommendation tasks. Experimental results show that the proposed algorithm can effectively learn the representation of knowledge graphs, as well as the importance of node neighbors and meta-paths, and further alleviate the problems of data sparsity and cold-start.

Key words: knowledge graph; recommender system; attention mechanism; meta-path

近年来, 随着数据的海量产生, 推荐系统已广泛应用于社交媒体、购物平台等领域。推荐系统是一种基于用户历史行为数据和物品属性信息的工

具, 用于预测用户对未来物品的偏好, 以便为用户提供个性化的推荐。传统的推荐系统通常依赖于用户-项目的交互数据^[1], 但这类数据往往存在严重

收稿日期: 2024-05-30

基金项目: 国家重点研发计划 (2020AAA0109300)

作者简介: 李政, 主要从事推荐系统、图神经网络、情感计算等方面的研究。

*通信作者 E-mail: huangbosues@sues.edu.cn

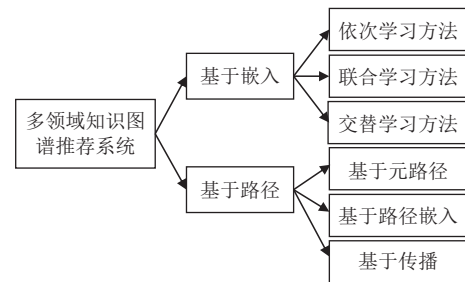
的稀疏性问题^[2], 对于新用户或新项目^[3], 由于缺乏足够的交互数据而难以提供有效的推荐。为了解决这个问题, 研究者们引入了基于元路径的方法。元路径是对象类型之间的一系列关系, 它定义了其起始类型和结束类型之间的新复合关系。不同的元路径不仅可以不同方面描述节点的特征, 还可以捕捉节点的高阶语义。通过分析用户的行为轨迹和交互数据, 它将用户和项目之间的关联信息转化为更丰富的特征表示, 更好地理解用户的兴趣和爱好, 从而提供更精准地推荐。基于元路径的方法在异构信息网络中使用广泛, 通过利用不同类型的关系路径来捕捉异构信息网络中用户和物品之间的复杂关系, 从而能够更好地挖掘潜在的用户行为和物品特征, 有助于解决数据稀疏性和冷启动问题, 提高推荐系统的准确性和覆盖率。但现有基于元路径的算法通常采用统一的权重分配策略, 没有考虑元路径之间的细微差别, 因为不同的元路径对同一节点的影响是不同的。此外, 推荐系统中的数据通常包括多种类型的实体和关系, 传统的推荐方法往往无法充分利用这些异构信息, 基于知识图嵌入的方法可以在不同应用场景中从多个角度提取这些信息的特征。然而, 这种利用整个嵌入知识图而不应用信息相关性监管约束的方法无法阻止噪声渗透到深层信息中。针对上述问题, 本文提出了一种基于异构注意力网络的知识图嵌入元路径推荐算法 (meta-path recommendation algorithm based on heterogeneous attention network, MRHAN), 通过引入知识图谱, 利用丰富的辅助信息来缓解数据稀疏性问题。基于元路径的方法通过预定义元路径来捕捉不同类型实体和关系之间的复杂语义信息, 从而更好地利用异构信息网络 (heterogeneous information network, HIN)。具体来说, 本文利用知识图嵌入方法和基于节点相关性的约束方法, 将语义相关的高阶实体和关系建模到唯一的元路径中。为了区别不同元路径对同一节点的影响, 使用层次化注意力网络来建模节点对不同邻居和不同元路径的偏好。其中节点层次的注意力网络学习节点对不同邻居的权重, 元路径层次的注意力网络学习元路径对推荐任务的权重。该框架能够学习知识图谱的有效表示以及节点邻居和元路径对节点的重要性, 进一步缓解数据稀疏和冷启动等问题。

1 理论知识

基于知识图谱的推荐系统^[4]是一种将知识图谱

中的丰富结构化信息和关联数据应用于推荐场景的方法。相比传统推荐系统, 基于知识图谱的推荐系统能够充分挖掘用户和物品之间的潜在关系, 从而提供更精准和个性化的推荐结果。

基于知识图谱推荐系统根据应用领域的不同, 可以分为多领域和特定领域两种类型。多领域知识图谱推荐系统包括基于嵌入和基于路径的两种推荐方法, 如图1所示。



其中, 基于元路径的方法通过人工或自动学习方式构建元路径, 从而灵活地适应不同的推荐场景和数据集。如果实体之间没有明确的语义关系或者关系较弱, 这种方法可能无法取得理想的推荐效果。HeteRec^[5]通过使用元路径挖掘不同类型实体之间的复杂关系。MetaPath2Vec^[6]形式化了基于元路径的随机游走, 能够学习不同实体和关系的低维表示, 实现对异构网络的结构信息和语义关联的同时建模, 提高推荐性能。MNRec^[7]根据领域知识选择合适的元路径, 有效地从异构信息网络中提取和融合用户偏好和项目属性的有用特征。MDyGSI^[8]设计了一个元路径感知聚合模块, 通过图中特定于行为的元路径, 将元路径感知用户嵌入建模为地理和社会偏好。

1.1 知识图嵌入方法

在知识图谱推荐系统中, 知识表示学习方法主要用于将知识图谱中的实体和关系表示为低维特征向量, 在映射过程中保留了知识图谱的结构信息和实体关系的语义信息, 便于计算和推理。知识表示学习方法可分为3类, 基于距离的翻译模型^[9]、基于语义的匹配模型和基于路径的嵌入学习模型。

基于距离的翻译模型主要是指将实体及关系映射到低维空间, 并通过距离度量来表示实体和关系之间的关联性。文献[10]提出的一种知识表示学习方法 TransE, 它将实体和关系表示为向量, 并通过距离度量来表示实体之间的关联性。TransH^[11]

为每个关系引入一个超平面，将实体投影到这个超平面上， h_{\perp} 和 t_{\perp} 表示投影到超平面的头部实体和尾部实体，然后在投影后的空间中类似于 TransE 进行距离度量。这种方法在处理一对多和多对一关系时表现出较好的性能，公式如下：

$$h_{\perp} = h - w_r^T h w_r \quad (1)$$

$$t_{\perp} = t - w_r^T t w_r \quad (2)$$

式中， w_r 表示变换矩阵； r 表示关系向量；TransE 和 TranH 的算法模型如图 2 和图 3 所示。

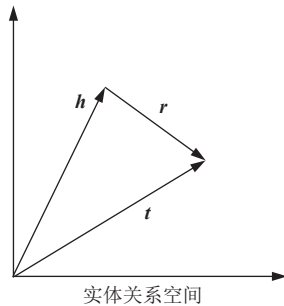


图 2 TransE 模型示意图

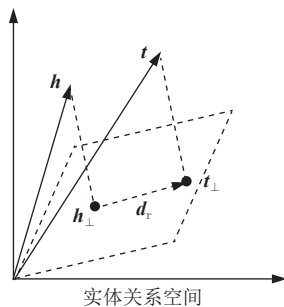


图 3 TransH 模型示意图

TransR^[12] 通过为每个关系引入一个独立的矩阵来映射实体向量到关系空间，从而克服了 TransE 在处理复杂关系方面的局限性，公式如下：

$$h_{\perp} = h M_r \quad (3)$$

$$t_{\perp} = t M_r \quad (4)$$

式中， M_r 表示变换矩阵； r 表示关系向量。

TransD^[13] 同样通过为每个关系引入一个独立的映射矩阵，但这个映射矩阵是通过实体向量和关系向量的外积生成的，使其能够更灵活地处理复杂关系，公式如下：

$$M_{rh} = p \cdot h_p^T + I \quad (5)$$

$$M_{rt} = p \cdot t_p^T + I \quad (6)$$

$$h_{\perp} = h M_{rh} \quad (7)$$

$$t_{\perp} = t M_{rt} \quad (8)$$

式中， M_{rh} 和 M_{rt} 表示由实体和关系共同决定的不同变换矩阵； p 、 h_p^T 、 t_p^T 为投影向量， p 表示如何将实体从实体空间映射到关系空间中， h_p^T 和 t_p^T 分别表示实体和关系的语义； I 表示单位矩阵。TransR 和 TransD 的算法模型如图 4 和图 5 所示。

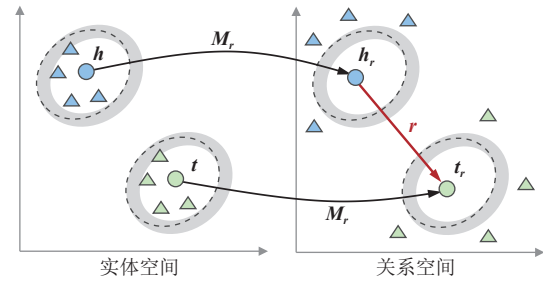


图 4 TransR 模型示意图

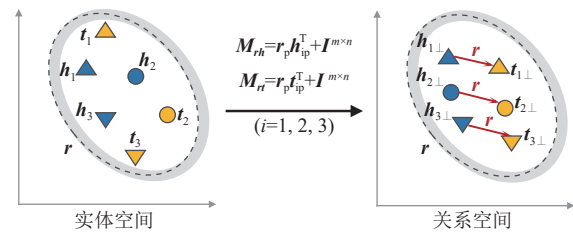


图 5 TransD 模型示意图

基于语义的匹配模型主要关注实体和关系之间的语义信息，并试图在低维空间中捕捉实体和关系之间的语义关联性。RESCAL^[14] 将知识图谱表示为一个三阶张量，并通过张量分解来捕捉实体和关系之间的语义信息。RESCAL 能够表示多种复杂关系，但其计算复杂度较高。Complex^[15] 通过将实体和关系表示为复数向量，以此捕捉实体和关系之间的非对称性。ConvE^[16] 将头实体和关系通过卷积操作来生成尾实体的表示，从而捕捉实体和关系之间的语义关联性。

基于路径的嵌入模型主要关注知识图谱中实体之间的路径信息，以捕捉实体之间的潜在关联性。DeepPath^[17] 利用一个递归神经网络来学习实体之间的潜在路径，并通过强化学习策略来指导路径搜索过程。R-GCN^[18] 在知识图谱中通过考虑关系类型来扩展图卷积网络，并能够学习实体之间的潜在路径信息。

1.2 相似度计算方法

传统基于协同过滤的推荐算法通常利用用户对

物品的评分数据先构建一个用户-物品交互矩阵, 再通过计算用户之间的相似度来预测他们对未交互物品的评分。常用的相似度计算方法有 Pearson 相关系数^[19]、余弦相似度^[20]、Jaccard 相似度^[21]和欧式距离等, 需要根据不同的推荐场景和数据类型灵活选择使用。

Pearson 相关系数是一种用于衡量两个变量之间线性相关程度的方法, 其值的输出范围为 $[-1, 1]$, 可以表示两个变量之间的相关性强度和方向。

Pearson 相关系数在推荐系统中的运用公式为:

$$\text{Sim}_{\text{PCCs}}(i, j) = \frac{\sum_{c \in I_{i,j}} (R_{i,c} - \bar{R}_i)(R_{j,c} - \bar{R}_j)}{\sqrt{\sum_{c \in I_{i,j}} (R_{i,c} - \bar{R}_i)^2} \sqrt{\sum_{c \in I_{i,j}} (R_{j,c} - \bar{R}_j)^2}} \quad (9)$$

式中, $I_{i,j}$ 表示用户 i 和用户 j 都评分过的物品的集合; $R_{i,c}$ 和 $R_{j,c}$ 分别表示用户 i 和用户 j 对物品 c 的评分; \bar{R}_i 和 \bar{R}_j 分别表示用户 i 和用户 j 对 $I_{i,j}$ 中物品评分的均值。

余弦相似度是一种比较两个向量相似度的度量方式, 其值的输出范围为 $[-1, 1]$, 值越大表示两个向量越相似, 值越小则表示它们越不相似。在推荐系统中通过余弦相似性判断用户之间的相似程度时, 利用空间中的一个多维向量表示用户对所有物品的评分, 用户间的相似性可以通过向量间的夹角表示。但由于其存在分数膨胀问题, 即向量维度较高或向量元素的大小差异较大, 会导致计算结果不准确或不可靠。修正的余弦相似度与 Pearson 相关系数很相似, 但区别是 Pearson 相关系数根据两个用户都评分过的物品集合进行计算, 而修正的余弦相似度则对所有物品集合都进行计算, 公式为:

$$\text{Sim}_{\text{Acos}}(i, j) = \frac{\sum_{c \in I_{i,j}} (R_{i,c} - \bar{R}_i)(R_{j,c} - \bar{R}_j)}{\sqrt{\sum_{c \in I_i} (R_{i,c} - \bar{R}_i)^2} \sqrt{\sum_{c \in I_j} (R_{j,c} - \bar{R}_j)^2}} \quad (10)$$

式中, I_i 和 I_j 分别表示用户 i 和用户 j 评分过的物品的集合; $R_{i,c}$ 和 $R_{j,c}$ 分别表示用户 i 和用户 j 对物品 c 的评分; \bar{R}_i 和 \bar{R}_j 分别表示用户 i 和用户 j 对他们所有评分的均值。

Jaccard 相似系数用于衡量两个集合之间的相关程度。在推荐系统中, 将每个用户评分过的物品都视为一个集合, 通过 Jaccard 相似系数计算

不同用户之间对物品评分的相似度, 公式如下所示:

$$\text{Sim}_J(i, j) = \frac{|N_{(u)} \cap N_{(v)}|}{|N_{(u)} \cup N_{(v)}|} \quad (11)$$

式中, $N_{(u)}$ 和 $N_{(v)}$ 分别表示用户 u 和用户 v 评分过的物品的集合。

欧氏距离是一个常用的距离定义, 指在 m 维空间中两个点之间的真实距离或者向量的自然长度 (即该点到原点的距离)。在二维和三维空间中, 欧式距离就是两点之间的真实距离。在推荐系统中, 通过欧式距离可以计算两个用户的评分向量在 n 维空间中的绝对距离, 公式为:

$$\rho(i, j) = \sqrt{\sum (x_i - x_j)^2} \quad (12)$$

式中, $\rho(i, j)$ 表示用户 i 和用户 j 的评分向量在向量空间中的绝对距离, 再经过归一化处理后即可得到两个用户之间的相似性, 公式为:

$$\text{Sim}_{\text{em}}(i, j) = \frac{1}{1 + \rho(i, j)} \quad (13)$$

1.3 基于元路径的异构图学习

在推荐系统中, 元路径是指描述用户和物品之间关系的一系列“路径”, 其中包含了多种不同类型的节点和边。其核心思想是通过定义不同类型的节点和边的组合来描述用户和物品之间的关系。具体来说, 一个元路径可以看作是一个描述用户和物品之间关系的模式, 其中包含了多个节点和边, 每个节点和边都代表一种不同的信息。

基于元路径的方法在异构信息网络中使用广泛, 它通过引入更多的信息和关联性成为分析异构图的有效方法。MAGNN^[22] 考虑了元路径内和元路径间聚合, 以在异构图上融合来自不同元路径的信息。Meta-HGT^[23] 提出一种基于元路径的超图结构, 充分提取原始信息网络中具有不同语义的复合高阶关系。GraphMSE^[24] 设计了高效的元路径采样技术, 并使用多层感知器嵌入元路径实例, 解决了计算成本高和特征异构的问题。MAERec^[25] 从异构信息网络的交互场景中自动挖掘有意义的元路径, 解决了现有基于元路径方法对专业设计的依赖问题。

基于元路径的方法通过构建描述推荐系统中用户和物品之间更为复杂的关系的元路径, 帮助推荐系统更好地理解用户的兴趣和偏好, 从而提供更为个性化的推荐结果。

2 问题定义

定义 $\mathcal{E} = \{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_k\} | k = |\mathcal{E}|$ 表示实体类型集合, $\mathcal{R} = \{r_1, r_2, \dots, r_p\} | p = |\mathcal{R}|$ 表示关系类型集合, $\mathcal{U} = \{u_1, u_2, \dots, u_m\} | m = |\mathcal{U}|$ 表示用户集合, $\mathcal{V} = \{v_1, v_2, \dots, v_n\} | n = |\mathcal{V}|$ 表示项目集合, 且 $\mathcal{U}, \mathcal{V} \subseteq \mathcal{E}$ 。定义 $Y = \{y_{uv} | u \in \mathcal{U}, v \in \mathcal{V}\}$ 表示用户-物品交互记录, 当 $y_{uv} = 1$ 时表示用户 u 与项目 v 之间有过交互记录, 如点击或购买等; $y_{uv} = 0$ 表示用户 u 与项目 v 之间没有过交互记录。

用 $\mathcal{G} = \{E, P, \phi, \psi\}$ 表示一个异构信息图, 其中 E 表示实体集合, P 表示实体之间的路径集合, ϕ 表示实体类型映射函数 $\phi_{f_e} : E \rightarrow \mathcal{E}$, ψ 表示路径类型映射函数 $\psi_{f_p} : P \rightarrow \mathcal{R}$ 。知识图谱中的每个实体元素 $e \in E$ 和路径 $p \in P$ 都通过实体类型映射函数 $\phi_{f_e}(e) \in \mathcal{E}$ 和路径类型映射函数 $\psi_{f_p}(p) \in \mathcal{R}$ 与图 \mathcal{G} 相关联。

元路径 Φ 是表示不同类型节点之间关系的一条有限路径。一个从 ε_a 到 ε_d 的元路径被定义为 $\varepsilon_a \xrightarrow{r_1} \varepsilon_b \xrightarrow{r_2} \varepsilon_c \xrightarrow{r_3} \varepsilon_d$, 表示为 $\Phi(\varepsilon_a, \varepsilon_d) = r_1 \circ r_2 \circ r_3$,

其中 \circ 表示关系复合运算符。

本文定义给定异构图中的节点 i 和元路径 Φ , 节点 i 基于元路径的邻居 N_i^Φ 表示通过元路径 Φ 与节点 i 连接的节点集合。不同元路径可以揭示不同的语义信息, 基于元路径的邻居能够利用异构图中不同方面的结构信息。将用户-物品交互记录 Y 和知识图谱 \mathcal{G} 作为输入, 利用相关性因子和阈值过滤掉不相关的数据, 并利用相关信息构建有影响的子图。通过分层注意力网络对特征进行提取, 通过学习节点对不同邻居的权重, 以及元路径对推荐任务的权重, 体现不同元路径对节点的重要性。将元路径的权重作为系数, 通过内积来融合这些语义特定的嵌入, 最后预测用户可能会感兴趣的物品。

3 模型设计

3.1 模型架构

模型的整个结构如图 6 所示, 主要包括知识图谱嵌入模块、基于节点相关性的约束采样、基于分层注意力网络的特征提取和偏好预测 4 个部分。

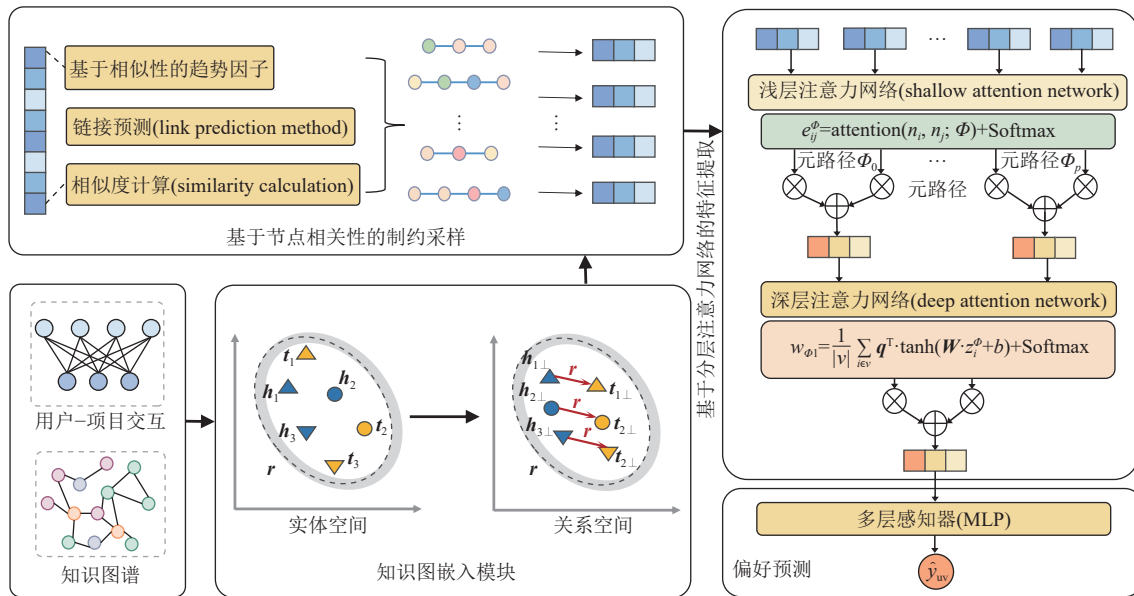


图 6 MRHAN 模型结构图

MRHAN 以用户-项目历史交互信息和知识图谱中与用户-项目相关的知识信息作为输入, 通过 TransD 将实体和关系嵌入独立空间向量, 得到实体和关系的嵌入表示。通过引入受阈值制约的相关性因子, 用来评估当前节点和候选节点之间的相关性程度, 确定语义上与中心节点相关的节点, 将单跳突出的路径统一并创建元路径, 通过唯一标识符识别元路径。

使用多层注意力网络来建模节点对不同邻居和不同元路径的偏好, 通过学习节点对不同邻居的权重, 以及元路径对推荐任务的权重, 体现不同元路径对节点的重要性。将元路径的权重作为系数, 通过内积来融合这些语义特定的嵌入, 最后预测用户的偏好。

3.2 知识图谱嵌入模块

由于异构图的异质性, 实体可以是不同类型的

节点, 如在电影推荐系统中, 实体可以包括电影、演员、导演、用户等。每个实体可以拥有不同的属性(如电影的标题、演员和导演的姓名、用户的性别和年龄等), 这些属性可以被用来描述和区分不同的实体。因此, 不同类型的节点具有不同的特征空间, 将具有相同特征的信息实例放置在相同的向量空间中有助于模型的高效计算, 而传统的基于矩阵分解等方法将实体和关系嵌入在同一空间中, 导致推荐结果不准确。

本文使用知识图嵌入方法 TransD 将知识图谱中实体和关系嵌入到独立向量空间中, 并使用以下矩阵保持它们的隔离:

$$M_{ph} = p \cdot n_h^T + I^{k \times d} \quad (14)$$

$$M_{pt} = p \cdot n_t^T + I^{k \times d} \quad (15)$$

$$h_{\perp} = h M_{ph} \quad (16)$$

$$t_{\perp} = t M_{pt} \quad (17)$$

式中, 映射矩阵由实体和关系定义, h 、 t 表示头实体和尾实体; p 、 n_h 、 n_t 表示嵌入信息; h_{\perp} 、 t_{\perp} 表示头尾实体的投影向量。实体属于 E^d 和关系属于 P^k , $I^{k \times d}$ 为单位矩阵; M_{ph} 和 M_{pt} 分别用于将头部和尾部实体的向量投影到相关关系空间。与经典的 Trans 系列知识图谱嵌入算法一样, 实体之间的距离被用来表示分数, 给定三元组 (n_h, p, n_t) 合理性因子 f_p :

$$f_p(n_h, p, n_t) = -\|M_{ph} \cdot n_h + p - M_{pt} \cdot n_t\|_2^2 \quad (18)$$

3.3 基于节点相关性的制约采样

通过基于元路径的随机游走采样方法, 根据预定义的元路径在异构信息网络中进行节点的学习, 可以捕捉不同类型实体间的潜在语义关系。然而, 现有基于元路径的随机游走采样方法^[26]通常对信息相关性的关注没有进行限制, 可能导致对推荐没有帮助的数据作为噪声被关注, 造成推荐效果下降以及不必要的计算开销。针对这些方法存在的问题, 本节提出基于节点相关性的制约采样方法, 通过确定语义相关信息的唯一元路径, 突出和维护嵌入空间中的突出路径来细化数据。

为了过滤掉不太相关的数据实例, 本章引入一个受阈值 ϑ 制约的相关性因子 Rel , 用来评估当前节点 n_c 和候选节点 n_j 之间的相关性程度 $\text{Rel}_{c,j}$, 如果相关性程度过低则舍弃该候选节点 n_j , 并将 $\text{Rel}_{c,j}$ 赋

值为 0, 如果相关性程度满足阈值则赋值为 1, 并保留符合要求的候选节点 n_j 。

基于节点相关性的制约采样结合了基于传播的思想, 首先确定语义上相关的节点, 维护单跳突出路径, 将单跳突出的路径统一并创建元路径。随机将一个节点 n_c 作为中心节点, 计算中心节点 n_c 与候选节点 $n_j (j = 1, 2, \dots)$ 之间的相关性, 并选择与 n_c 相似度最高的节点作为游走下一步的目标节点。

首先计算中心节点 n_c 和候选节点 n_j 基于相似性的趋势因子:

$$\text{Tensim}(n_c, n_j) = \frac{[(\mathcal{N}(n_c) \cap \mathcal{N}(n_j)) + 1] + f(\mathcal{T})_{n_c, n_j}}{[(\mathcal{N}(n_c) \cup \mathcal{N}(n_j)) + 1] \sum_{\forall \Gamma | \Gamma \in \mathcal{N}(n_c)} f(\mathcal{T})_{n_c, \Gamma}} \times 100 \quad (19)$$

式中, $\mathcal{N}(n_c)$ 和 $\mathcal{N}(n_j)$ 分别表示中心节点 n_c 和候选节点 n_j 的邻居; $f(\cdot)$ 表示计算相关节点之间趋势的函数, 如下所示:

$$f(\mathcal{T})_{n_c, n_j} = x Y_r(\mathcal{T}) + (1 - x) Y_b(\mathcal{T}) \quad (20)$$

式中, x 表示范围在 (0,1) 的系数; \mathcal{T} 表示趋势参数; Y_r 和 Y_b 表示最近和之前的用户交互记录, 趋势参数与用户交互记录相关。

本文选择使用 Adamic Adar 链接预测算法和 Jaccard 相似度计算方法计算 n_c 和 n_j 之间的局部相关性, 以及 n_c 和 n_j 在给定路径 p 之间的相似性。给定路径 p , 对于节点 n_c 和节点 n_j , 其相关性计算公式为:

$$\text{Sim}_{AJ}((n_c, n_j) + (n_c, n_j, p)) = 0.25 \left(\sum_{y \in (\mathcal{N}(n_c) \cap \mathcal{N}(n_j))} \frac{2}{\log |\mathcal{N}(y)|} + 2X \right) \quad (21)$$

式中, $X = \frac{[(\mathcal{N}(n_c) \cap \mathcal{N}(n_j)) + 1] |p|}{[(\mathcal{N}(n_c) \cup \mathcal{N}(n_j)) + 1] |p|}$ 表示候选节点在所有邻居节点中所占的比例。

根据式 (19) 和式 (21) 可得当前节点 n_c 和候选节点 n_j 之间的相关性程度为:

$$\text{Rel}(n_c, n_j) = \text{Tensim}(n_c, n_j) + \text{Sim}_{AJ}((n_c, n_j) + (n_c, n_j, p)) \quad (22)$$

式中, $\text{Rel}(n_c, n_j)$ 值越大, 则候选节点 n_j 成为下一个节点的可能性越大。异构图 \mathcal{G} 中节点通过彼此的相关性因子在图中相互关联, 当相关性程度大于或等于阈值 ϑ , 则表示两节点间存在有影响的连接, 否则两节点间不相连。定义如下:

$$\text{Rel}(n_c, n_j) = \begin{cases} 1 & \text{Rel}(n_c, n_j)Y \geq \vartheta \\ 0 & \text{Rel}(n_c, n_j)Y < \vartheta \end{cases} \quad (23)$$

式中, Y 表示用户-物品交互记录是否存在, 同样作为约束条件。

在评估了中心节点 n_c 所有的邻居节点后, 将其相关性最高的邻居节点作为下一个中心节点 n_c , 并重复该过程, 直到模型对所有的节点都执行了相同的方法, 元路径按照其分配的编号统一存储。最终, 基于元路径的编号组合各条元路径信息以构建局部子图。

基于节点相关性的制约采样保留整个图中具有高相关性的节点和它们的语义关联路径, 从而更好地捕捉节点之间的语义关联。通过选择相关性最高的邻居节点作为下一个中心节点, 并将元路径按照其分配的编号统一存储, 可以确保生成子图具有一定的一致性和相关性, 有助于后续基于分层注意力网络的特征提取。

3.4 基于分层注意力网络的特征提取

在元路径上每个节点的邻居在节点嵌入中都有着不同的重要性, 现有基于元路径的算法通常没有考虑元路径之间的细微差别。针对这一问题, 本节提出基于分层注意力网络的特征提取, 使用层次化注意力网络来建模节点对不同邻居和不同元路径的偏好, 通过学习节点对不同邻居的权重, 以及元路径对推荐任务的权重, 体现不同元路径对节点的重要性。

注意力机制是一种加强特定部分信息处理能力的方法, 其核心思想是为输入数据的不同部分分配不同的权重, 这些权重决定了模型在处理数据时对每个部分的关注程度。通过动态地调整权重, 模型可以更好地捕捉输入数据中的关键信息, 从而提高学习效果。注意力机制示意图如图7所示。

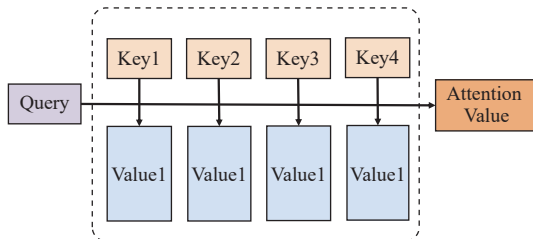


图7 注意力机制示意图

首先通过浅层注意力网络学习每个节点与邻居节点之间的权重。用 $e_{i,j}^\Phi$ 表示通过元路径 Φ 连接的节点对 (n_i, n_j) 中, 节点 n_j 对节点 n_i 的重要程度。

$$e_{i,j}^\Phi = \text{attention}(n_i, n_j; \Phi) \quad (24)$$

给定元路径 Φ , 所有基于元路径的节点对共享 $e_{i,j}^\Phi$ 。然后通过softmax函数对其进行归一化, 获得权重系数 $a_{i,j}^\Phi$:

$$a_{i,j}^\Phi = \text{softmax}(e_{i,j}^\Phi) = \frac{\exp(\sigma(\mathbf{a}_\Phi^\top [n_i \| n_j]))}{\sum_{k \in N_i^\Phi} \exp(\sigma(\mathbf{a}_\Phi^\top [n_i \| n_k]))} \quad (25)$$

式中, σ 表示激活函数; $\|$ 表示连接操作; \mathbf{a}_Φ 是元路径 Φ 的节点关注向量。由于节点连接的顺序不同, 以及每个节点具有不同的邻居, 所以归一化后的权重系数 $a_{i,j}^\Phi$ 是不对称的, 即它们对彼此的重要性不同。通过聚合邻居节点 n_j 的投影特征与相应的系数, 得到节点 n_i 基于元路径的嵌入:

$$z_i^\Phi = \sigma\left(\sum_{j \in N_i^\Phi} a_{i,j}^\Phi \cdot n_j\right) \quad (26)$$

式中, z_i^Φ 是元路径 Φ 中节点 n_i 的学习嵌入表示, 每个节点嵌入都由其邻居聚合。由于浅层注意力权重 $a_{i,j}^\Phi$ 是由单个元路径生成的, 所以它能够代表一种特定的语义结构。

对于给定元路径集 $\{\Phi_0, \Phi_1, \dots, \Phi_P\}$, 其节点特征通过浅层注意力得到 P 组语义特定的节点嵌入, 表示为 $\{Z_{\Phi_0}, Z_{\Phi_1}, \dots, Z_{\Phi_P}\}$, 每个元路径的深层权重表示为 $\{\beta_{\Phi_0}, \beta_{\Phi_1}, \dots, \beta_{\Phi_P}\} = \text{attention}\{Z_{\Phi_0}, Z_{\Phi_1}, \dots, Z_{\Phi_P}\}$ 。通过非线性变换得到特定于语义的嵌入表示:

$$w_{\Phi_i} = \frac{1}{|\mathcal{V}|} \sum_{i \in \mathcal{V}} \mathbf{q}^\top \cdot \tanh(\mathbf{W} \cdot z_i^\Phi + \mathbf{b}) \quad (27)$$

式中, \mathbf{q} 表示其注意力权重; \mathbf{W} 是可学习权重参数矩阵; \mathbf{b} 是偏差向量。获取每个元路径的特征后, 对 w_{Φ_i} 通过softmax函数做归一化处理:

$$\beta_{\Phi_i} = \text{softmax}(w_{\Phi_i}) = \frac{\exp(w_{\Phi_i})}{\sum_{i=1}^P \exp(w_{\Phi_i})} \quad (28)$$

由此得到元路径 Φ_i 中节点的深层注意力权重, β_{Φ_i} 值越大, 表示该元路径上的节点越重要。由于元路径对不同任务有着不同的权重, 可以将元路径的权重作为系数, 通过内积来融合这些语义特定的嵌入, 得到最终的嵌入表示:

$$\mathbf{Z} = \sum_{i=1}^P \beta_{\Phi_i} \cdot Z_{\Phi_i} \quad (29)$$

3.5 偏好预测

隐式反馈的推荐任务是根据用户与候选集里的项目进行偏好程度估计, 然后按照偏好程度对项目进行降序排序, 使得用户即将发生的交互尽可能地有更靠前的排名。

将用户视为目标节点, 一组嵌入向量 $\{Z_{u,1}, Z_{u,2}, \dots, Z_{u,k}\}$ 代表用户的不同方面。类似的, 将项目视为目标节点, 一组嵌入向量 $\{Z_{v,1}, Z_{v,2}, \dots, Z_{v,k}\}$ 代表项目的不同方面。

对于Top-K推荐任务, 本节将用户的各方面嵌入向量与候选项目的各方面嵌入向量进行内积, 再通过sigmoid激活函数将其约束在0~1的范围内, 运算结果作为该用户对该项目的偏好。用户 u 对项目 v 的偏好得分:

$$S_{u,v} = \sum_{k=1}^K \mathbf{Z}_{u,k}^T \cdot \mathbf{Z}_{v,k} \quad (30)$$

$$\hat{y}_{u,v} = \text{sigmoid}(S_{u,v}) = \frac{1}{1 + e^{-S_{u,v}}} \quad (31)$$

在所有的训练集中, 为了解决类别不平衡问题, 通过平衡正负样本的数量从而保证模型训练的效果, 为每个节点提取与正样本数量相同的负样本, 完整的损失函数为:

$$L = - \sum_{(u,v) \in \mathcal{Y}^+} \log \hat{y}_{u,v} - \sum_{(u,v^-) \in \mathcal{Y}^-} \log(1 - \hat{y}_{u,v^-}) \quad (32)$$

式中, \mathcal{Y}^+ 表示通过控制采样率和有交互的正样本; \mathcal{Y}^- 表示通过没有交互的样本中均匀采样的负样本。

4 实验

4.1 实验设定

为了评估本章提出的模型性能, 实验选用了Yelp、Amazon-electricity、MovieLens这3个公共数据集作为基准数据集。

Yelp: 选择其中一个被广泛采用的本地商业推荐子数据集。将餐馆等本地企业视为项目, 将每条评论视为用户和项目之间的互动。为了确保数据集的质量, 将至少有10次互动的用户和项目, 以及至少有5个朋友的用户作为数据内容。

Amazon-electricity: 该数据集中的评分数据通常较为稀疏, 这意味着大部分用户和产品之间没有互动, 为了确保数据集的质量, 将确保每个用户和物品至少有8个互动, 每个物品至少有2个相关物品。

MovieLens: 该数据集由MovieLens网站上的

明确评级组成, 将电影视为项目, 用户和项目之间的每一次交互都被标记为0或1, 表明用户是否对该项目进行了评分。为了确保数据集的质量, 将确保每个用户和项目至少有5个。

对于每个数据集, 按照时间顺序对交互进行排序。具体来说, 首先确定数据集中最早和最晚的时间戳, 并根据时间戳信息对交互数据进行排序, 选择前80%的历史交互作为训练集, 接下来的10%作为验证集, 最后的10%作为测试集。

为了验证本章提出模型的有效性和通用性, 选取了BPRMF、NEUMF、MCRec、NeuACF、GC-MC、NGCF、HAN、DGCF、CG-KGR、Meta-HRP、IHDT等模型参与了实验。

BPRMF^[27]: 一种基于贝叶斯理论的个性化排序矩阵分解模型, 使用了隐式反馈数据, 将用户和物品向量相乘预测用户对物品的兴趣度, 是否存在交互行为(如点击或购买)。通过最大化后验概率优化模型, 利用贝叶斯理论来处理不确定性, 采用高斯先验来约束用户和物品向量的值, 并使用随机梯度下降算法进行优化。

NEUMF^[28]: 一种基于神经协作过滤的推荐模型, 利用神经网络的灵活性和非线性, 代替矩阵分解的点积, 提高模型的表达能力, 具有隐式反馈。

MCRec^[29]: 利用基于元路径的上下文和顶部的共同关注机制进行推荐。

NeuACF^[30]: 考虑用户和项目的多个方面, 并使用深度神经网络进行推荐。

GC-MC^[31]: 采用用户项二分图中的GCN编码器来生成用户和项的表示, 其中只考虑一阶邻居。

NGCF^[32]: 使用嵌入传播层来丰富用户和项目表示, 以利用用户-项目二分图中的高阶连通性。

HAN^[33]: 在异构图神经网络中引入注意力机制, 通过不同的元路径聚合邻居信息。

DGCF^[34]: 将用户复杂的意图解耦, 以能更好地建模细粒度的用户-项目关系。

CG-KGR^[35]: 利用知识图谱补充辅助信息, 提出协作指导机制对知识图谱和用户项目交互进行充分和连贯的学习。

Meta-HRP^[36]: 利用历史交互数据学习先验知识, 并基于最近的一些用户评分记录快速适应新项目, 及时捕捉用户偏好转移, 缓解了冷启动问题。

IHDT^[37]: 通过在双塔神经网络之间增加交互性和高阶特征学习来改进现有模型。

实验中按照8:1:1的比例将数据集划分为训

训练集、验证集和测试集。公平起见，所有基线算法均采用 Adam 优化器进行训练，学习率设置为 0.005，所有模型的嵌入大小都固定为 100，并应用网格搜索来优化超参数。

4.2 实验指标

本文实验利用 Pytorch 实现。对于测试数据集中的每个用户，将用户未与之交互的所有项目视为负面项目。Top- K 推荐任务的内容是为用户推荐最感兴趣的 K 个物品，这是推荐系统中最常见的任务之一。采用 3 个广泛使用的指标：Precision、Recall 和 NDCG 对实验结果进行比较和分析。

Precision@ K （精确度）表示分类器判定为正样本的个数中，分类正确的正样本个数所占的比例。在推荐系统中用于衡量前 K 个推荐物品中，实际为用户感兴趣的物品所占的比例，公式如下：

$$\text{Precision@}K = \frac{\sum_{u \in U} |R(u) \cap T(u)|}{\sum_{u \in U} |R(u)|} \quad (33)$$

Recall@ K （召回率）表示在用户实际感兴趣的物品中，被推荐到前 K 个物品中所占的比例，公式如下：

$$\text{Recall@}K = \frac{\sum_{u \in U} |R(u) \cap T(u)|}{\sum_{u \in U} |T(u)|} \quad (34)$$

NDCG@ K （归一化折损累计增益）用来评估排序结果。DCG 考虑了物品在推荐列表中的位置，排名越靠前的物品其权重越高，排名靠后的物品权

重越低。DCG 值越高，表示推荐列表中的相关物品排序越合理，公式如下：

$$\text{DCG@}K = \sum_{i=1}^K \frac{\text{rel}(i)}{\log_2(i+1)} \quad (35)$$

归一化折损累计增益 NDCG 是 DCG 的一个标准化版本，用于比较不同长度的推荐列表。它将 DCG 值除以理想状况中的最大 DCG 值（IDCG），得到的值在 [0,1] 之间，NDCG 值越接近 1，表示推荐系统的推荐性能越好，公式如下：

$$\text{IDCG@}K = \sum_{i=1}^{|\text{REL}|} \frac{2^{\text{rel}(i)} - 1}{\log_2(i+1)} \quad (36)$$

$$\text{NDCG@}K = \frac{\text{DCG@}K}{\text{IDCG@}K} \quad (37)$$

式中，|REL| 表示按照相关性的大小顺序排序得出的 K 个结果的集合。

4.3 实验结果

本章对 MRHAN 和基线模型在 Yelp、Amazon-Electricity 和 Movielens-10M 这 3 个数据集上进行了实验。根据各模型最后给出的推荐列表，返回前 K 个推荐结果用于 Top- K 推荐任务。

表 1 详细列举了不同算法在 Precision@10、Recall@10 和 NDCG@10 指标下的评估数据，图 8 给出部分模型在 Top- K 推荐场景下的 Precision@ K 、Recall@ K 和 NDCG@ K 指标对比。本文模型各指标始终优于其他基线模型，证明了基于异构注意力网络的知识图嵌入元路径推荐方法的有效性。

表 1 不同模型各指标对比

Model	Yelp			Amazon-Electricity			Movielens-10M		
	Prec@10	Recall@10	NDCG@10	Prec@10	Recall@10	NDCG@10	Prec@10	Recall@10	NDCG@10
BPR-MF	0.279 4	0.473 5	0.599 7	0.053 4	0.303 8	0.322 7	0.486 3	0.151 9	0.696 3
NeuMF	0.271 8	0.446 2	0.590 3	0.052 7	0.329 0	0.316 3	0.601 7	0.170 9	0.783 9
MCRec	0.280 9	0.464 1	0.602 2	0.080 1	0.368 1	0.354 9	0.547 8	0.147 8	0.751 0
NeuACF	0.287 3	0.486 0	0.621 6	0.061 4	0.334 7	0.323 2	0.585 6	0.168 1	0.772 0
GC-MC	0.292 9	0.475 1	0.598 2	0.088 3	0.417 9	0.367 7	0.496 3	0.150 9	0.706 0
NGCF	0.281 9	0.480 2	0.607 4	0.078 6	0.375 0	0.345 2	0.553 2	0.166 1	0.753 1
HAN	0.310 1	0.493 8	0.622 8	0.092 7	0.457 1	0.376 7	0.576 9	0.152 3	0.768 0
DGCF	0.298 7	0.489 0	0.615 5	0.079 8	0.398 6	0.354 8	0.567 9	0.158 7	0.725 5
CG-KGR	0.302 1	0.487 1	0.621 5	0.090 5	0.436 2	0.380 3	0.576 3	0.162 2	0.734 9
Meta-HPR	0.293 0	0.478 5	0.620 1	0.091 2	0.462 0	0.375 8	0.580 5	0.160 8	0.770 3
IHDT	0.308 7	0.473 6	0.619 9	0.089 3	0.459 2	0.381 2	0.579 6	0.167 1	0.754 2
MRHAN	0.327 3	0.498 7	0.631 5	0.098 5	0.485 7	0.391 5	0.608 6	0.173 2	0.784 7

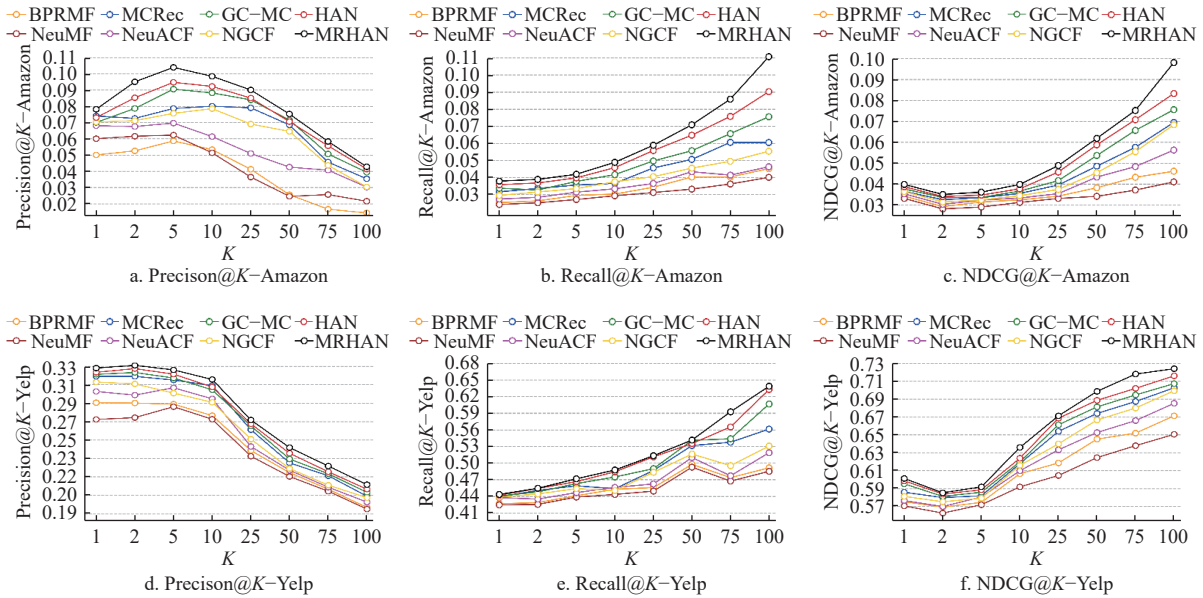


图 8 不同模型在 Top-K 推荐场景下的指标对比

从各模型在 3 个数据集上的综合实验结果来看, 大部分基于元路径的推荐算法效果超过了协同过滤和图神经网络的推荐算法, 这说明基于元路径的方法通过定义和利用元路径来捕捉不同类型节点之间的语义关系, 能够更准确地理解节点之间的关系, 提高推荐的准确性。基于异构信息网络的推荐方法, 如 MCRRec 和 NeuACF, 利用不同的元路径来捕捉用户和项目之间的语义关系, 所以其实验结果优于传统的基于协同过滤的推荐方法, 但相较于其他方法较差。可能的原因是, 这两种模型元路径的选择和设计还不完善, 需要深入研究元路径的选择, 挖掘不同元路径对节点的不同影响。MetaHRP 通过 HIN 上的元路径构建元任务, 然后利用基于多通道图卷积网络和注意力机制的嵌入表示生成器, 为用户和项目生成相应的嵌入, 缓解了数据层面的冷启动问题。MRHAN 使用基于节点相关性的制约采样, 计算中心节点和候选节点的相关程度, 对元路径进行合理的筛选, 提高了推荐性能。

基于图神经网络的推荐方法在大多数情况下比传统的基于协同过滤的推荐方法具有更好的性能。这表明根据节点的不同类型, 建立信息传播模型对于提高推荐性能非常重要。GC-MC 利用一阶邻居来指导信息的聚合, NGCF 利用异构信息网络中没有语义关系的多阶邻居, DGCF 利用细粒度的意图丰富了用户表示并采用迭代更新机制使结果更贴近实际。而 HAN 通过元路径来建模不同类型的邻居节点, 并引入注意力机制, 通过信息传播来捕捉节点之间的协作信号, 其推荐性能优于 GC-MC 和

NGCF, 这验证了在异构图中建立语义关系的重要性。CG-KGR 将历史交互封装起来, 并以此为指导从知识图谱中提取信息, 提供更精确的个性化推荐。IHDT 考虑了双塔模型的速度优势和图神经网络的精度优势, 在基于图神经网络的方法中更具表现力。MRHAN 使用基于分层注意力网络的特征提取, 建模节点对不同邻居和不同元路径的偏好, 提高了推荐的准确性。

4.3.1 消融实验

为进一步验证 MRHAN 模型各组成部分的有效性, 对 3 种 MRHAN 变体进行消融实验。MRHAN/G 表示 MRHAN 框架在知识图嵌入模块不使用 TranD 表示学习方法, MRHAN/S 表示在 MRHAN 框架中不使用基于节点相关性的制约采样, 不对元路径的选择进行任何限制, MRHAN/A 表示在 MRHAN 框架中不使用基于分层注意力网络提取特征, 直接聚合预测。表 2 展示了消融实验的结果。

表 2 MRHAN 框架消融实验

Model	Yelp	Amazon	Movielens
MRHAN/G	0.208 3	0.061 9	0.541 3
MRHAN/S	0.296 5	0.079 3	0.568 3
MRHAN/A	0.288 5	0.076 2	0.577 1
MRHAN	0.327 3	0.098 5	0.608 6

从表 3 中可以看出, 知识图嵌入模块使用过渡矩阵将头部和尾部投影到超平面上, 可将图中节点及其边的关系嵌入低维向量空间, 从而提供更丰富的节点特征表示, 为计算节点间的相似度降低复杂

度；基于节点相关性的制约采样将语义相关的高阶实体和关系建模到唯一的元路径中，通过筛选有效连接的元路径防止噪声信息对底层信息的渗透，提高了推荐的精确度；基于分层注意力网络的特征提取，使用分层注意力网络来建模节点对不同邻居和不同元路径的偏好，通过学习节点对不同邻居的权重，以及元路径对推荐任务的权重，体现不同元路径对节点的重要性。综上所述，MRHAN 框架中的各部分都是必要的，去除其中任一部分后得到的结果都是次优的。

4.3.2 数据稀疏性分布的性能比较

推荐系统普遍面临着数据稀疏问题的挑战。为了研究 MRHAN 模型对于数据稀疏性问题的缓解效果，在 Amazon 数据集中将训练数据按不同比例分为 4 个子数据集，分别为 20%、40%、60% 和 80%，验证不同算法在不同数据集比例情况下的实验效果。通过对比本文模型与基线模型在不同的数据稀疏性分布下的 NDCG@10 性能，尤其是在用户-物品交互记录较少的情况下的性能，来验证基于异构注意力网络的知识图嵌入元路径推荐方法是否能够缓解数据稀疏问题的影响，对比结果如图 9 所示。

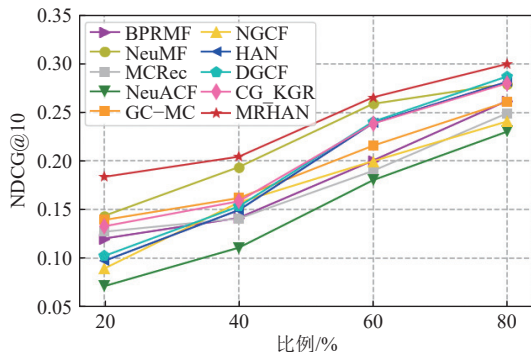


图 9 数据稀疏性分布的 NDCG@K 比较

从图 9 的结果中可以看出，MRHAN 模型在不同的数据稀疏性分布下表现优异，尤其是在交互记录稀少的情况下，可以显著提高推荐性能，而且随着训练数据的增加，改进幅度逐渐减小。这证明了基于异构注意力网络的知识图嵌入元路径推荐方法在缓解数据稀疏性方面的有效性。相比之下，基于 GNN 的推荐方法（如 NGCF），虽然可以捕捉用户和项目之间的高阶连接信息，但不能保证比传统的推荐方法更优，可能是因为他们忽略了用户和项目之间不同方面的语义关系。

4.3.3 几种不同表示学习方法的性能比较

由于异构图的异质性，实体可以是不同类型的

节点，每个实体可以拥有不同的属性，这些属性可以被用来描述和区分不同的实体。因此，不同类型的节点具有不同的特征空间，而传统的基于矩阵分解等方法将实体和关系嵌入在同一空间中，导致推荐结果不准确。

基于知识图嵌入的表示学习方法可以将节点和关系映射到低维向量空间中，从而提供更为丰富的节点特征表示。这些特征不仅包含了节点自身的信息，还考虑了节点周围的语义关系，这有助于提高推荐系统的精度。

在 MRHAN 模型的知识图嵌入模块中，将异构图通过 TranD 方法对节点进行表示学习。为了验证基于知识图嵌入的表示学习方法的有效性。采用 Precision@K 作为评估指标，对基于图嵌入的方法 TransE、TransH、TransR 和 TranD，基于矩阵分解的方法 SVD，基于图嵌入的方法 DeepWalk^[38] 和基于深度学习的方法 CNN，以及经典的图神经网络学习方法 GraphSAGE^[39]，比较其在 MRHAN 中推荐性能的影响，如表 3 所示。

表 3 不同表示学习方法对 Precision@10 的影响

Model	Yelp	Amazon	Movielens
TransE	0.276 7	0.074 1	0.602 8
TransH	0.294 8	0.075 1	0.537 4
TransR	0.308 5	0.079 5	0.591 5
SVD	0.245 1	0.063 1	0.568 7
DeepWalk	0.265 3	0.079 7	0.554 2
CNN	0.208 3	0.061 9	0.541 3
GraphSAGE	0.276 2	0.079 5	0.564 7
TransD	0.327 3	0.098 5	0.608 6

从表 3 中可以看出，在知识图嵌入模块中，使用 TransD 方法对知识图谱实体进行表示学习的效果优于其他方法。TransD 将实体嵌入映射到对应的关系空间，为每个实体和关系学习低维表示向量，以保留知识图谱的结构信息，从而更好地描述知识图谱中实体之间的相互作用和联系。相比其他表示学习方法，TransD 的关系投影矩阵能够将节点投影到对应的空间，捕获不同节点之间的语义关系，提高推荐的准确性和个性化程度。

4.3.4 不同链接预测和相似度方法的性能比较

在基于节点相关性的制约采样中，使用不同的链接预测方法和相似度计算方法对 Precision@10 的影响如表 4 所示。

表4 不同链接预测和相似度计算方法的影响

Method	Yelp	Amazon	Movielens
Pearson相关系数	0.285 4	0.079 5	0.563 7
修正余弦相似度	0.276 3	0.075 1	0.589 2
Jaccard相似度	0.327 3	0.098 5	0.608 6
共同邻居	0.217 5	0.065 3	0.498 1
资源分配	0.301 7	0.087 6	0.596 3
优先链接	0.238 5	0.069 2	0.505 1
Adamic Adar	0.327 3	0.098 5	0.608 6

从表4中可以看出, 基于节点相关性的制约采样中, 使用 Adamic Adar 算法和 Jaccard 相似度的效果优于其他方法。可能的原因是, 共同邻居和优先链接通过共同节点数量计算节点相连概率, 共同的邻居越多连接的可能性越大。

5 结束语

本文提出一种基于异构注意力网络的知识图嵌入元路径推荐算法 MRHAN, 利用知识图嵌入方法和基于节点相关性的约束方法, 将语义相关的高阶实体和关系建模到唯一的元路径中。引入基于节点相关性的约束方法, 使用相关性因子和阈值过滤掉不相关的数据。为了区别不同元路径对同一节点的影响, 使用层次化注意力网络来建模节点对不同邻居和不同元路径的偏好。浅层注意力网络学习节点对不同邻居的权重, 深层注意力网络学习元路径对推荐任务的权重。该框架能够学习知识图谱的有效表示以及节点邻居和元路径对节点的重要性。实验结果表明了该模型在解决稀疏性和冷启动问题方面的有效性, 消融实验也进一步验证了本文模型各部分的重要性。

参考文献

- [1] 刘佳玮, 石川, 杨成, 等. 基于异质信息网络的推荐系统研究综述[J]. 信息安全学报, 2021, 6(5): 1-16.
LIU J W, SHI C, YANG C, et al. Review of research on recommendation systems based on heterogeneous information networks[J]. Journal of Information Security, 2021, 6(5): 1-16.
- [2] YANG L, SHIJIA E, XU S Y, et al. Interactive knowledge graph attention network for recommender systems[C]// Proceedings of the International Conference on Data Mining Workshops. Sorrento, Italy: IEEE, 2020: 211-219.
- [3] MU R H, ZENG X Q, ZHANG J Y. Heterogeneous information fusion based graph collaborative filtering recommendation[J]. Intelligent Data Analysis, 2023, 27(6): 1595-1613.
- [4] ZHANG F Z, YUAN N J, LIAN D F, et al. Collaborative knowledge base embedding for recommender systems[C]// Proceedings of the Proceedings of the 22nd

ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York: ACM, 2016: 353-362.

- [5] YU X, REN X, SUN Y Z, et al. Personalized entity recommendation: A heterogeneous information network approach[C]// Proceedings of the 7th ACM international conference on Web search and data mining. New York: ACM, 2014: 283-292.
- [6] DONG Y X, CHAWLA N V, SWAMI A. metapath2vec: Scalable representation learning for heterogeneous networks[C]// Proceedings of the Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York: ACM, 2017: 135-144.
- [7] TAN L, GONG D F, XU J M, et al. Meta-path fusion based neural recommendation in heterogeneous information networks[J]. Neurocomputing, 2023, 529: 236-248.
- [8] DING D, YI J, XIE J Y, et al. Meta-path aware dynamic graph learning for friend recommendation with user mobility[J]. Information Sciences, 2024, 666: 120448.
- [9] GAO Y W, YANG Y, JIANG F J. TransRT-RMP: An rotation translation and relationship mapping attribute-based representation learning model[C]// Proceedings of the 3rd International Conference on Computer Science, Electronic Information Engineering and Intelligent Control Technology. New York: IEEE, 2023: 134-138.
- [10] BORDES A, USUNIER N, GARCIA-DURÁN A, et al. Translating embeddings for modeling multi-relational data[EB/OL]. [2023-12-20]. <https://papers.nips.cc/paper/2013/file/1cecc7a77928ca8133fa24680a88d2f9-Paper.pdf>.
- [11] WANG Z, ZHANG J W, FENG J L, et al. Knowledge graph embedding by translating on hyperplanes[C]// Proceedings of the 28th AAAI Conference on Artificial Intelligence. [S.l.]: AAAI, 2014: 1112-1119.
- [12] LIN Y K, LIU Z Y, SUN M S, et al. Learning entity and relation embeddings for knowledge graph completion[C]// Proceedings of the 229th AAAI Conference on Artificial Intelligence. [S.l.]: AAAI, 2015: 2181-2187.
- [13] JI G L, HE S Z, XU L H, et al. Knowledge graph embedding via dynamic mapping matrix[C]// Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing. [S.l.]: Association for Computational Linguistics, 2015: 687-696.
- [14] NICKEL M, TRESP V, KRIEGEL H P. A three-way model for collective learning on multi-relational data[C]// Proceedings of the 28th International Conference on Machine Learning. [S.l.]: ICML, 2011: 809-816.
- [15] TROUILLON T, WELBL J, RIEDEL S, et al. Complex embeddings for simple link prediction[C]// 33rd International Conference on Machine Learning. [S.l.]: ICML, 2016, 5: 3021-3032.
- [16] DETTMERS T, MINERVINI P, STENETORP P, et al. Convolutional 2D knowledge graph embeddings[C]// Proceedings of the AAAI Conference on Artificial Intelligence. [S.l.]: AAAI, 2018: 1811-1818.
- [17] XIONG W H, HOANG T, WANG W Y. DeepPath: A

- reinforcement learning method for knowledge graph reasoning[C]//Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing. Stroudsburg: Association for Computational Linguistics, 2017: 564-573.
- [18] SCHLICHTKRULL M, KIPF T N, BLOEM P, et al. Modeling relational data with graph convolutional networks[M]//Lecture Notes in Computer Science. Cham: Springer International Publishing, 2018: 593-607.
- [19] BENESTY J, CHEN J D, HUANG Y T, et al. Pearson correlation coefficient[M]//Springer Topics in Signal Processing. Heidelberg: Springer, 2009: 1-4.
- [20] XIA P P, ZHANG L, LI F Z. Learning similarity with cosine similarity ensemble[J]. *Information Sciences*, 2015, 307: 39-52.
- [21] JACCARD P. The distribution of the flora in the alpine zone[J]. *New Phytologist*, 1912, 11(2): 37-50.
- [22] FU X Y, ZHANG J N, MENG Z Q, et al. MAGNN: Metapath aggregated graph neural network for heterogeneous graph embedding[C]//Proceedings of The Web Conference 2020. New York: ACM, 2020: 2331-2341.
- [23] LIU J, SONG L, WANG G, et al. Meta-HGT: Metapath-aware hyperGraph transformer for heterogeneous information network embedding[J]. *Neural Networks*, 2023, 157: 65-76.
- [24] LI Y, JIN Y L, SONG G J, et al. GraphMSE: Efficient meta-path selection in semantically aligned feature space for graph neural networks[C]//Proceedings of the AAAI Conference on Artificial Intelligence. [S.l.]: AAAI, 2021, 35(5): 4206-4214.
- [25] ZHANG Y H, LIAO W W, WANG Y L, et al. Meta-path automatically extracted from heterogeneous information network for recommendation[J]. *World Wide Web*, 2024, 27(3): 26.
- [26] SHI C, HU B B, ZHAO W X, et al. Heterogeneous information network embedding for recommendation[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2019, 31(2): 357-370.
- [27] ZHANG F Z, YUAN N J, LIAN D F, et al. Collaborative knowledge base embedding for recommender systems[C]//Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York: ACM, 2016: 353-362.
- [28] HE X N, LIAO L Z, ZHANG H W, et al. Neural collaborative filtering[C]//Proceedings of the 26th International Conference on World Wide Web. Switzerland: International World Wide Web Conferences Steering Committee, 2017: 173-182.
- [29] 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]//Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. New York: ACM, 2018: 1531-1540.
- [30] FERRARI DACREMA M, CREMONESI P, JANNACH D. Are we really making much progress? A worrying analysis of recent neural recommendation approaches[C]//Proceedings of the 13th ACM Conference on Recommender Systems. New York: ACM, 2019: 101-109.
- [31] VAN DEN BERG R, KIPF T, WELLING M. Graph convolutional matrix completion[EB/OL]. [2023-12-25]. <https://www.semanticscholar.org/reader/fdc708aaa0d18c791f878ff2214201410fa52439>.
- [32] WANG X, HE X N, WANG M, et al. Neural graph collaborative filtering[C]//Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval. New York: ACM, 2019: 165-174.
- [33] WANG X, JI H Y, SHI C, et al. Heterogeneous graph attention network[C]//Proceedings of the World Wide Web Conference. New York: ACM, 2019: 2022-2032.
- [34] WANG X, JIN H Y, ZHANG A, et al. Disentangled graph collaborative filtering[C]//Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval. New York: ACM, 2020: 1001-1010.
- [35] CHEN Y K, YANG Y M, WANG Y J, et al. Attentive knowledge-aware graph convolutional networks with collaborative guidance for personalized recommendation[C]//Proceedings of the IEEE 38th International Conference on Data Engineering. New York: IEEE, 2022: 299-311.
- [36] ZHOU M Q, LI K P, DAI K L, et al. HIN-based rating prediction in recommender systems via GCN and meta-learning[J]. *Applied Intelligence*, 2023, 53(20): 23271-23286.
- [37] HE Q, LI X K, CAI B. Graph neural network recommendation algorithm based on improved dual tower model[J]. *Scientific Reports*, 2024, 14(1): 3853.
- [38] PEROZZI B, AL-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. New York: ACM, 2014: 701-710.
- [39] HAMILTON W L, YING R, LESKOVEC J. Inductive representation learning on large graphs[EB/OL]. [2023-12-29]. <http://arxiv.org/abs/1706.02216v4>.