

刘臣,袁慕婷,周立欣. 基于集合学习的超链接预测[J]. 智能计算机与应用,2024,14(4):162-167. DOI: 10.20169/j.issn.2095-2163.240425

基于集合学习的超链接预测

刘 臣,袁慕婷,周立欣

(上海理工大学 管理学院,上海 200093)

摘要:超图链接预测作为图预测的重要研究方向之一,能够通过预测节点间的高阶相互作用解决许多实际问题。目前大多数链接预测研究多集中于成对关联关系的预测,而实际应用中链接关系的对象往往大于两个。因此,本文提出一种基于集合表示和 transformer 的链接预测模型。该模型通过对集合表示实现链接预测的无序性,并将传统的语言模型拓展应用于链接预测问题。模型首先将数据嵌入编码层对数据特征进行提取,然后使用池化机制对解码层进行解码,并引入评分函数对模型预测结果进行评估。实验表明,本文提出的模型可以有效利用网络结构特征,在6个不同规模的代谢网络数据集上的表现优于多个基准算法。

关键词:超链接预测;注意力机制;集合学习

中图分类号:O157.5

文献标志码:A

文章编号:2095-2163(2024)04-0162-06

Hyperlink prediction based on collection learning

LIU Chen, YUAN Muting, ZHOU Lixin

(Business School, University of Shanghai for Science and Technology, Shanghai 200093, China)

Abstract: Hypergraph link prediction as one of the important research directions of graph prediction, it can solve many practical problems by predicting the high-order interaction between nodes. At present, most of link prediction studies focus on the prediction of pairwise association relations, but in practice, the objects of link relations are often more than two. Therefore, this paper proposes a link prediction model based on set representation and transformer, which realizes the disorder of link prediction through set representation, extends the traditional language model to link prediction problem. The model first embedded the data into the encoding layer to extract the data features, then used the pooling mechanism to decode the decoding layer, introduced the scoring function to evaluate the prediction results of the model. Experiments show that the model we proposed can effectively utilize the structural features of the network and outperform multiple benchmark algorithms on six metabolic network datasets of different sizes.

Key words: hyperlink prediction; attention mechanisms; collection learning

0 引言

超图链接预测作为图预测的重要研究方向之一,通过观察数据结构中已知节点的相关关系预测新链接存在的可能性,从而提供更完整的数据结构信息。目前,现有研究多集中于成对关联关系的预测,然而在现实应用中的关系涉及对象数量往往大于两个,称之为高阶关系。高阶关系是指在3个及3个以上节点之间的相互作用^[1]。例如,烹饪菜谱时所需要的食材大多是复杂的高阶关系;合著网络中合作作者数量往往也大于两个。而超图链接预测

可以通过预测潜在的高阶关系,挖掘作者间合作或文章引用的潜在可能。

超图链接预测的传统方法是基于节点相似性的启发式方法^[2],该方法假设节点倾向于与其相似的节点形成链接,通过对节点间相似性得分的排列,得分更高的节点间更有可能构成缺失的超边^[3]。启发式方法包括 PageRank (PR)^[4]、Katz 指数 (Katz Index, KI)^[5]、SimRank (SR)^[6]、随机游走 (Random Walks)^[7]等等。这类方法实践效果良好,但存在具有较强的人为假设,对于不同类型的数据集,相似性函数往往具有不同定义。例如,在社交网络中,节点

基金项目:上海市哲学社会科学规划课题(2021BTQ003);中国博士后科学基金第69批面上项目(2021M692135)。

作者简介:袁慕婷(1995-),女,硕士研究生,主要研究方向:链路预测;周立欣(1989-),男,博士,主要研究方向:网络分析,深度学习。

通讯作者:刘 臣(1981-),男,博士,副教授,主要研究方向:链路预测,深度学习,数据挖掘。Email:chenliu@usst.edu.cn

收稿日期:2023-03-17

的相似性意味着存在互动关系的可能性更高,然而在化合物反应中并不适用。例如酸性物质和碱性物质会发生化学反应,但相似性极低^[8]。

近年来,链接预测主要方法是基于神经网络的节点嵌入法,按照节点信息提取范围分为两类:第一类以节点的局部邻域信息作为特征提取范围,例如 Node2Vec^[9] 节点嵌入算法,通过学习节点到特征的低维空间映射,使用偏随机游走,最大限度保留节点网络的邻域信息进行链接预测。但由于不是针对链接预测任务设计的,被用于多个下游任务时,无法捕获对链接预测最有效的信息。第二类是从全局角度挖掘每个节点潜在信息的方法,如 Struc2Vec^[10]、CMM^[11] 协调矩阵最小化方法等,将所有超链接投射至邻接空间,从全局角度挖掘每个节点潜在的特征信息,通过节点特征相似性找到缺失链接。但协调矩阵最小化方法优化不足,不能很好的拟合数据,且计算成本较高,难以应用于大规模数据集。

本文提出基于注意力机制的无序语言模型,将超图链接预测转化为集合学习问题,对超图中缺失的链接进行预测。模型通过已知数据间的关联结构,对集合元素所代表的节点、集合所代表的超边特征进行充分提取,更高效、准确地进行链接预测。通过6个超图数据集上的实验表明,研究提出的模型能够有效处理复杂多元关系链接预测问题。

本文将链接预测问题转化为集合学习问题,提出基于注意力机制的链接预测方法,对数据的节点特征进行更为全面的提取和学习。通过模型的置换不变特性实现对不同范围节点关系的捕获,极大提高了数据处理效率。除此之外,模型不对数据结构做任何假设,可适用于多个下游任务,提高效率的同时可以以较低计算成本对大规模数据集进行处理。

1 相关工作

1.1 基于神经网络的方法

图神经网络被证明是一种有效的基于深度学习解决链接预测问题的工具^[12],关系图卷积神经网络 R-GCN^[13] 模型通过图卷积网络进行三元组的嵌入,但是图卷积神经网络无法为每个邻域三元组分配独立的权重。WGNCN^[14] 模型为每个关系分配单独权重,将超图结构拆分成多个关系子图进行学习,Vashishth^[15] 等利用节点及网络结构关系,通过线性或非线性函数来计算三元组权重,为解决三元组权重一致性的问题,提出了图注意力网络。通过图注意力机制生成权重,KBGAT^[16] 模型将图注意力应用

到超图网络上,并使用多阶关系进一步扩大节点聚合邻域信息的范围。但对节点来说,距离的远近对应权重的大小不同,因此 Wang^[17] 等利用衰减机制来分配不同距离之间的权重,但是在节点聚合邻域过程中,三元组相互之间的独立处理,忽略了邻域三元组之间的联系。因此,本研究针对高阶关系提出图注意力机制模型架构,用于增强高阶关系之前的交互,从而提升链接预测效率。

1.2 基于 Transformer 的语言模型

在语言学习领域,应用最为广泛的是基于编码器-解码器架构的 Seq2Seq 模型^[18]。为了在学习过程中不遗漏长序列中的重要信息,在上下文序列构造中加入注意力机制,为输入序列的不同元素赋予相应注意力权重。Seq2Seq 模型的编码器和解码器一般为循环神经网络,但循环神经网络属于序列模型,需要按序列化方式对信息进行处理,时间成本高且随着距离的增加信息不断被稀释,效果并不理想。随后提出的 Transformer 语言模型一次性“学习”所有词汇,利用注意力机制将不同距离的单词进行结合。其优势在于计算位置间关系的成本与顺序无关,可以支持具有大规模参数的数据集训练,在自然语言处理、音频处理、计算机视觉、等多个人工智能领域取得了巨大成功。

1.3 集合学习研究

为了便于解释节点关系间的无序性,本文将链接预测问题转化为集合学习问题,通过集合角度来理解节点关系间的“序”。传统语言模型多为有序的,例如上文中的 Seq2Seq 模型就需要按序列化方式进行学习。本研究中使用无序语言模型进行链接预测,经过多个超图数据集的实验证实,本文所提模型与参照模型相比,在提高链接预测效率的同时取得了相对不错的效果。

2 模型

超图链接预测需要根据复杂数据的相关关系建模,并依据生成的超图特征进行缺失高阶关系预测^[19]。本研究将链接预测问题表示为集合问题,并提出基于自注意力机制的无序链接预测模型,使用传统语言模型中 Encoder-Decoder 框架,在对参数进行训练优化后,模型通过合理利用节点嵌入特征,有效提升链接预测性能。

如图1所示,模型主体分由两部分组成。第一部分为超图网络表征学习,第二部分为超边的评分函数。超图表征学习是在有效提取节点特征的基础

上进行预训练,利用数据自身信息进行自监督学习的模块。将转化为集合问题后的数据作为特征输入,经过3层注意力模块组成的编码层进行特征提取,将训练集中隐藏部分节点的超边嵌入模型进行预训练,有效提取超边节点特征。超边评分函数作

为有监督学习模块,本文先使用池化机制对编码层输出进行解码,随后引入评分函数对预测结果进行衡量。通过评分函数及负样本分类对模型进行微调,以确定符合数据特征的最佳模型。

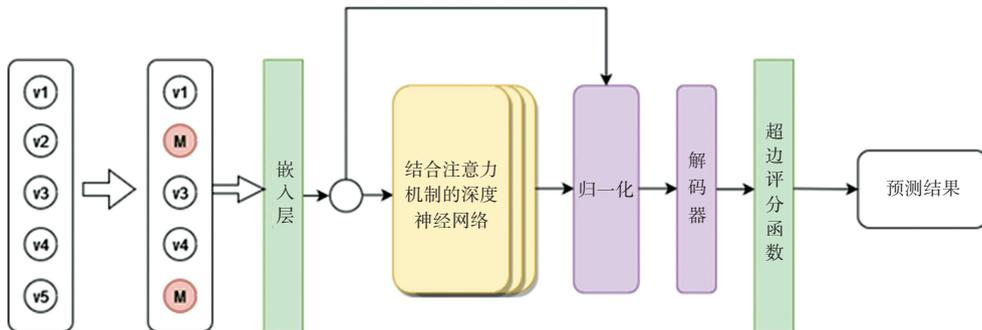


图1 模型的总体结构

Fig. 1 Model structure diagram

2.1 链接预测问题表示方法

模型主体分为两部分,第一部分为超图网络表征学习,第二部分为超边的评分函数。超图表征学习是在有效提取节点特征的基础上进行预训练,利用数据自身信息进行自监督学习的模块。将转化为集合问题后的数据作为特征输入,经过3层注意力模块组成的编码层进行特征提取,将训练集中隐藏部分节点的超边嵌入模型进行预训练,有效提取超边节点特征。超边评分函数作为有监督学习模块,本文先使用池化机制对编码层输出进行解码,随后引入评分函数对预测结果进行衡量。通过评分函数及负样本分类对模型进行微调,以确定符合数据特征的最佳模型。

本文将集合表示运用到超图链接预测问题中,通过集合表示超链接最终构建超图网络。在介绍链接预测模型之前,首先介绍超图网络中相关集合表示的基础概念。

图 $G = (V, E)$ 是由两个集合:顶点集合 V 和边的集合 E 组成。其中, $V(G) = v_1, v_2, \dots, v_n$ 是一个有限非空的集合,集合 $E(G) = \{(v_0, v_1, v_2), \dots, (v_2, v_5, v_7, v_9), \dots, (v_3, v_5, v_7, v_9, v_{10})\}$ 表示超边的集合,本文将超边集合中每个元素都视为超图中的一条超边。根据节点与超边的关系建立关联矩阵 $A = (a_{ij})$:

$$a_{ij} = \begin{cases} 1, & \text{if } v_i \in e_j \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

$A = (a_{ij})$ 是 $n \times d$ 的关联矩阵,其中 n 为节点数量, d 为超边数量,将超图链接预测问题转化为集合

问题后,嵌入模型进行学习。

2.2 超图网络表征学习

超图网络的表征学习通过注意力模块进行节点特征提取,在此基础上进行自监督预训练。本文使用 set transformer 的自注意力模块用于连接预测及预训练, set transformer 作为处理集合型数据的无序模型,主要用于自然语言处理。本研究首次将其拓展于链接预测领域,极大提高了链接预测处理效率。

对于给定超图将其构建为集合问题嵌入注意力模块,使用多头注意力机制对节点特征进行全面提取,通过将 Q, K, V 分别投影到不同 d_q^M, d_k^M, d_v^M 维度的向量上,输出注意力的线性变换:

$$\text{Multihead}(Q, K, V, \lambda, \omega) = \text{concat}(O_1, \dots, O_h) W^o \quad (2)$$

$$O_j = \text{Att}(Q W_j^Q, K W_j^K, V W_j^V; \omega_j) \quad (3)$$

多头注意力可学习参数为 $\lambda = \{W_j^Q, W_j^K, W_j^V\}_j^h = 1, O_j$ 是一次注意力之后的输出结果,其中 $W_j^Q, W_j^K \in R^{dq \times dq}$, $W_j^V \in R^{dv \times dv}$, $W^o \in R^{hd^M \times d}$ 。在此基础上定义无序链接预测模型,超图链接的预测结果不依赖于节点间位置关系。对于给定矩阵 $X, Y \in R^{n \times d}$,多头注意力模块 T 为:

$$T(X, Y) = \text{LayerNorm}(L + \text{rFF}(L)) \quad (4)$$

$$L = \text{LayerNorm}(X + \text{Multihead}(X, Y, Y; \omega)) \quad (5)$$

其中, L 是输入特征与多头自注意力变换后的特征和经过归一化层的结果; rFF 代表前馈层, T 是将经过前馈层的 H 与其自身相加后经过归一化层的结果。在这里归一化层的设置是为了保证训练的稳定性。通过 Z 结合数据特征,可以定义本模型的

注意力模块 TOD 为

$$TOD(\mathbf{X}) = T(\mathbf{X}, \mathbf{X}) \quad (6)$$

当将节点集嵌入 TOD 后,各节点之间执行自注意力机制,输出一个与输入等规模大小的节点间相关性信息。此外,可以通过多个 TOD 叠加来增强编码的高阶相互作用,本文模型中使用3层 TOD 叠加作为模型的编码器。将训练集中超边节点按一定比例掩码后,输入模型进行预训练,利用数据本身提供的信息进行自监督学习。

2.3 超边的评分函数

超边评分函数作为依赖于负样本的有监督学习模块,本文设置了评分函数并对预测结果进行衡量,参照评分结果对模型进行微调,以训练出符合数据特征的最佳模型参数。考虑到超边元素间的复杂关系,首先使用多头注意力对编码层的输出进行解码:

$$PMA_k(Z) = T(\mathbf{S}, \text{rFF}(Z)) \quad (7)$$

$$\text{Decoder}(Z) = \text{rFF}(TOD(PMA_k(Z))) \quad (8)$$

式中: $Z \in R^{n \times d}$ 是编码器输出的特征集, $\mathbf{S} \in R^{k \times d}$ 是可训练的种子向量, PMA_k 的输出是 k 个元素的集合。通过评分函数 I_e , 对预测出的高阶关系进行评定,将常用于定义超图拉普拉斯算子的 maxmin 函数拓展至超链接排名,提出用于连接预测的自适应学习模块:

$$I_e = \sigma(W \cdot \text{maxmin}\{h_v^e\}_{v \in e} + b) \quad (9)$$

$$\text{maxmin}\{x_j; j \in [k]\} = *(\max x_{sl} - \min x_{il}) \quad (10)$$

$$\mathcal{L} = \frac{1}{|E|} \sum_{e \in E} \Lambda\left(\frac{1}{|F|} \sum_{f \in F} I_f\right) - I_e \quad (11)$$

其中, W 是维度为 $1 \times d$ 的权重参数; σ 是 sigmoid 函数对评分结果进行收敛; $\text{maxmin}\{x_j; j \in [k]\}$ 是最大超边中节点数与最小超边节点数的差异。优化函数 \mathcal{L} 中, F 是链接预测超边集合的子集, E 是现有超边集合的子集, 函数 $\Lambda(x) = \log(1 + e^x)$ 为非递减函数。

优化函数通过最大化现有超边 E 的分数,设定现有超边 E 分数高于 F 中预测超边集合的平均分数。为了使模型拟合训练更稳定,在此加入 Relu 将负值收敛为0。在对预测结果进行评分的同时,对参数进行优化以确定最佳模型。

3 实验

3.1 数据集

本研究使用5个物种的6个代谢网络数据集进行了实验,其中包括大肠杆菌、巴氏菌、荣格达利菌、幽门螺杆菌和智人。数据集共计包含5599个节

点,已知超链接7469,详细统计数据见表1。数据收集来自各种生物的79个代谢网络,在此筛选出含有外来代谢物或已经存在于网络中的候选反应,并随机删除一些缺失的超链接,且保留剩余的作为观察数据。根据网络大小,被删除反应的数目在25~200或50~400的范围内。

表1 数据集统计信息

Table 1 Statistics of the datasets

数据集	Vertices	Hyperlinks
iAB_RBC_283	342	469
iIT341	485	554
iAF692	628	690
iHN637	698	785
iAF692	1668	2388
iJO1366	1805	3583

3.2 基准方法

为验证本文模型的有效性,将其与以下几种链接预测学习方法进行比较。

(1) 高效的超图学习算法 $\text{SHC}^{[20]}$ (Spectral Hypergraph Clustering), 最初设计用于利用超链接关系对超图的节点进行分类。将关联矩阵进行转置,将每个节点视为一个超链接,将每个超链接视为一个节点,从而通过 SHC 进行超链接预测。

(2) $\text{HPLSF}^{[21]}$ 作为一种监督学习的超链接预测方法,其沿着每个潜在特征维度计算熵值,以获得固定长度的特征输入。本文根据这些熵训练了一个逻辑回归模型,用来输出预测分数。

(3) $\text{FM}^{[22]}$ (因子分解机) 是一种灵活的因子分解模型,本文使用 FM 的分类函数,其中观察到的关联矩阵的列被作为模型的特征输入。

(4) $\text{CN}^{[23]}$ 将传统的成对公共领域拓展为超链接,其计算方法与 Katz 类似。

3.3 评价指标

实验结果主要使用 $\text{AUC}^{[24]}$ (ROC 曲线下的面积) 进行评估。 AUC 作为二分类重要评价指标,数值越大模型效果越好。通过分别计算模型结果的“真正例率”(True Positive Rate, TPR) 和“假正例率”(False Positive Rate, FPR), 将 TPR 作为纵轴, FPR 作为横轴,即可得到 ROC 曲线。

真正例表示正例被预测为正的个数,假正例表示将负类预测为正的个数,本文中指将虚拟超链接预测为原本存在的真链接个数;假负例表示将真链接预测为虚拟超链接的个数,真返利表示将虚拟超链接预测为虚构的假超链接个数。

3.4 参数设置

首先,将数据集按比例划分为训练正例和测试正例,本模型训练正例占比 80%,测试正例 20%,并生成同等数量的虚拟超链接作为训练负例和测试负例。对训练集中超边节点按 30%的比例随机 Mask 后,嵌入模型进行 Mask 节点预测。训练集批次大小为 512,学习率为 0.001,节点的初始嵌入向量维度为 128,模型注意力机制头数为 8,注意力层数为 3。测试集批次大小为 2 048,学习率为 0.000 1。评分损失权重为 1,并定义如果损失函数连续 4 次没有下降,则当前参数构成的模型为最佳模型。对每个数据集进行 30 次实验取平均值作为最终结果。

表 2 在代谢网络数据集上的连接预测结果

Table 2 Results of connection prediction on metabolic network datasets

Dataset	本文方法	SHC	HPLSE	FM	CN
iAB_RBC_283	<u>0.684 3±0.012 7</u>	0.696 3±0.012 2	0.435 1±0.012 6	0.662 0±0.027 5	0.388 1±0.011 6
iIT341	0.671 5±0.011 6	<u>0.597 8±0.011 7</u>	0.521 2±0.047 1	0.569 2±0.018 0	0.439 9±0.010 0
iAF692	0.681 4±0.021 3	<u>0.616 5±0.017 8</u>	0.471 9±0.045 0	0.546 5±0.021 2	0.430 0±0.021 3
iHN637	0.658 9±0.004 7	<u>0.617 0±0.013 8</u>	0.471 1±0.050 0	0.578 6±0.019 8	0.424 0±0.021 4
iAF1260b	0.721 8±0.009 0	<u>0.715 0±0.005 0</u>	0.541 8±0.008 8	0.614 9±0.014 2	0.467 9±0.010 5
iJO1366	0.749 3±0.031 0	<u>0.710 5±0.004 2</u>	0.483 4±0.033 5	0.630 9±0.022 8	0.437 1±0.010 5

由表 2 可以看出,在 6 个代谢网络数据集上本文提出的模型表现整体优于其他链接预测方法。在 iAB_RBC_283 数据集上,本文方法相较于 SHC 表现稍有逊色,这是因为 iAB_RBC_283 数据集中节点及关系数量相对较少,网络关系信息不充分,导致对本文提出的以网络结构关系为主的链接预测相对 SHC 以节点分类进行连接预测的方法而言效率相对较低。

在其他 5 个数据集上本文方法的表现均为最佳,而 SHC 方法的表现均为次优,这主要是因为 SHC 方法的稀疏构造对于数据集计算优势越明显,但在模型的稳定性方面无法与本文提出的模型相比,因为 SHC 是利用超链接关系的节点分类算法,而不是超链接预测算法。本文模型通过节点编码层的嵌入,实现了对数据结构特征的全面提取,并且通过对超链接的集合表示实现了链接预测无序性,使该模型可适用于多个下游任务且有效提高了预测的准确性。

通过数据集实验结果可以看到,代谢网络在本模型的超链接预测 AUC 值较其他方法有较大提升。

4 结束语

本文针对网络中的超链接预测任务提出了一种

对于数据集中的超链接,模型会随机 Mask 部分节点并对 Mask 的节点进行预测。但由于在预训练后的实验阶段会出现模型中未出现的节点,为了缓解预训练和实际实验效果相差较大的问题,本文设定在被随机选中掩码的节点中有 10%保留其本身,另外 10%的节点会被随机替换为其他节点,剩余的 80%节点被掩码。

3.5 实验结果分析

本文提出的模型与其他 4 种经典链接预测方法,在代谢网络数据集上对应的 AUC 对比结果见表 2。其中,加粗的 AUC 为最优,下划线 AUC 数据为次优。

基于集合表示并结合 Transformer 机制的复合模型。借助集合表示的置换不变性有效提高数据处理效率,在传统语言模型框架基础上进行拓展优化全面提取数据结构特征。可适用于多个下游任务,从而实现对大规模数据集的高效处理。经在 6 个代谢网络数据集上的评估实验结果表明,本文提出的方法有效改进了超链接预测在代谢网络中的应用,在关系数量较多和节点特征单一的网络上具有显著优势。

下一步的研究将从两个方面展开,一是尝试在节点特征多样性的网络中进行链接预测;二是能够根据网络结构特征自动调整取值,或给出恰当的取值范围从而进一步提高链接预测性能。

参考文献

- [1] BATTISTON F, CENCETTI G, IACOPINI I, et al. Networks beyond pairwise interactions: structure and dynamics[J]. Physics Reports, 2020, 874: 1-92.
- [2] LI D, XU Z, LI S, et al. Link prediction in social networks based on hypergraph [C]//Proceedings of the 22nd International Conference on World Wide Web. 2013: 41-42.
- [3] MARTÍNEZ V, BERZAL F, CUBERO J C. A survey of link prediction in complex networks [J]. ACM Computing Surveys (CSUR), 2016, 49(4): 1-33.
- [4] LANGVILLE A N, MEYER C D. Google's PageRank and Beyond

- [M]. Princeton:Princeton University Press, 2011.
- [5] KATZ L. A new status index derived from sociometric analysis [J]. *Psychometrika*, 1953, 18(1): 39–43.
- [6] JEH G, WIDOM J. Simrank: A measure of structural–context similarity [C]//Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2002; 538–543.
- [7] LIU W, LÜ L. Link prediction based on local random walk [J]. *EPL (Europhysics Letters)*, 2010, 89(5): 58007.
- [8] YADATI N, NITIN V, NIMISHAKAVI M, et al. NHP: Neural hypergraph link prediction [C]//Proceedings of the 29th ACM International Conference on Information and Knowledge Management. 2020; 1705–1714.
- [9] GROVER A, LESKOVEC J. Node2vec: Scalable feature learning for networks [C]//Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2016; 855–864.
- [10] RIBEIRO L F R, SAVERESE P H P, FIGUEIREDO D R. Struc2vec: Learning node representations from structural identity [C]//Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2017; 385–394.
- [11] ZHANG M, CUI Z, JIANG S, et al. Beyond link prediction: Predicting hyperlinks in adjacency space [C]//Proceedings of the AAAI Conference on Artificial Intelligence. 2018.
- [12] SCARSELLI F, GORI M, TSOI A C, et al. The graph neural network model [J]. *IEEE Transactions on Neural Networks*, 2008, 20(1): 61–80.
- [13] SCHLICHTKRULL M, KIPF T N, BLOEM P, et al. Modeling relational data with graph convolutional networks [C]//Proceedings of European Semantic Web Conference. Cham: Springer, 2018; 593–607.
- [14] SHANG C, TANG Y, HUANG J, et al. End-to-end structure-aware convolutional networks for knowledge base completion [C]//Proceedings of the AAAI Conference on Artificial Intelligence. 2019; 3060–3067.
- [15] VASHISHTH S, SANYAL S, NITIN V, et al. Interact: Improving convolution–based knowledge graph embeddings by increasing feature interactions [C]//Proceedings of the AAAI Conference on Artificial Intelligence. 2020; 3009–3016.
- [16] NATHANI D, CHAUHAN J, SHARMA C, et al. Learning attention–based embeddings for relation prediction in knowledge graphs [J]. *arXiv preprint arXiv:1906.01195*, 2019.
- [17] WANG R, LI B, HU S, et al. Knowledge graph embedding via graph attenuated attention networks [J]. *IEEE Access*, 2019, 8: 5212–5224.
- [18] CHO K, VAN MERRIËNBOER B, GULCEHRE C, et al. Learning phrase representations using RNN encoder–decoder for statistical machine translation [J]. *arXiv preprint arXiv:1406.1078*, 2014.
- [19] CHEN H, YIN H, SUN X, et al. Multi-level graph convolutional networks for cross–platform anchor link prediction [C]//Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2020; 1503–1511.
- [20] ZHOU D, HUANG J, SCHÖLKOPF B. Learning with hypergraphs: Clustering, classification, and embedding [J]. *Advances in Neural Information Processing Systems*, 2006, 19: 1601–1608.
- [21] XU Y, ROCKMORE D, KLEINBAUM A M. Hyperlink prediction in hypernetworks using latent social features [C]//Proceedings of International Conference on Discovery Science. Berlin Heidelberg: Springer, 2013; 324–339.
- [22] RENDLE S. Factorization machines with libFM [J]. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 2012, 3(3): 1–22.
- [23] LIBEN–NOWELL D, KLEINBERG J. The link prediction problem for social networks [C]//Proceedings of the Twelfth International Conference on Information and Knowledge Management. 2003; 556–559.
- [24] HUANG J, LING C X. Using AUC and accuracy in evaluating learning algorithms [J]. *IEEE Transactions on Knowledge and Data Engineering*, 2005, 17(3): 299–310.