

图嵌入学习研究综述:从简单图到复杂图

黄苗苗, 王慧颖, 王梅霞, 王业江, 赵宇海

引用本文

黄苗苗, 王慧颖, 王梅霞, 王业江, 赵宇海. 图嵌入学习研究综述:从简单图到复杂图[J]. 计算机科学, 2026, 53(1): 58-76.

HUANG Miaomiao, WANG Huiying, WANG Meixia, WANG Yejiang, ZHAO Yuhai. [Review of Graph Embedding Learning Research:From Simple Graph to Complex Graph](#) [J]. Computer Science, 2026, 53(1): 58-76.

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)

[基于鲁棒分区水印的深度学习模型保护方法](#)

Deep Learning Model Protection Method Based on Robust Partitioned Watermarking
计算机科学, 2026, 53(1): 423-429. <https://doi.org/10.11896/jsjcx.241200005>

[EvR-DETR:融合事件与RGB图像的轻量级端到端目标检测](#)

EvR-DETR:Event-RGB Fusion for Lightweight End-to-End Object Detection
计算机科学, 2026, 53(1): 153-162. <https://doi.org/10.11896/jsjcx.250300021>

[基于深度学习的OCT/OCTA视网膜图像分析方法综述](#)

Review of Retinal Image Analysis Methods for OCT/OCTA Based on Deep Learning
计算机科学, 2026, 53(1): 128-140. <https://doi.org/10.11896/jsjcx.241100047>

[基于图注意力交互的行人轨迹预测方法](#)

Pedestrian Trajectory Prediction Method Based on Graph Attention Interaction
计算机科学, 2026, 53(1): 97-103. <https://doi.org/10.11896/jsjcx.250300132>

[基于KAN的无监督多元时间序列异常检测网络](#)

KAN-based Unsupervised Multivariate Time Series Anomaly Detection Network
计算机科学, 2026, 53(1): 89-96. <https://doi.org/10.11896/jsjcx.241200190>

图嵌入学习研究综述：从简单图到复杂图

黄苗苗¹ 王慧颖² 王梅霞¹ 王业江¹ 赵宇海¹

¹ 东北大学计算机科学与工程学院 沈阳 110819

² 国家电网辽宁省电力有限公司信息通信分公司 沈阳 110004

(huangmiaomiao@stumail.neu.edu.cn)

摘要 图数据作为一种具有强大表达能力的数据类型,因具有复杂的结构而难以高效建模。如何有效捕捉其中的内在信息,成为一个富有挑战性的问题。图嵌入方法将高维稀疏的图映射为低维稠密的特征向量,同时保留图的结构信息,已经引起了广泛关注。然而,现有综述对图嵌入方法的总结不够全面,尤其对复杂图嵌入的关注较少,导致处理多样化图数据的研究现状未能得到系统梳理。对此,从简单图到复杂图,对图嵌入学习方法进行了系统综述。首先,给出了各种类型的图和图嵌入的常见定义;其次,系统地归纳了简单图上的嵌入方法,包括浅层和深度图嵌入方法;然后,按照图种类,总结了复杂图上的嵌入方法,重点介绍深度嵌入技术在动态图、异质图、多重图和超图等复杂图结构中的应用,以弥补现有文献对复杂图结构研究关注较少的不足;最后,讨论了图嵌入技术的实际应用场景,并展望了未来的发展方向。

关键词 图嵌入;图表示;深度学习;神经网络;复杂图

中图分类号 TP181

Review of Graph Embedding Learning Research: From Simple Graph to Complex Graph

HUANG Miaomiao¹, WANG Huiying², WANG Meixia¹, WANG Yejiang¹ and ZHAO Yuhai¹

¹ School of Computer Science and Engineering, Northeastern University, Shenyang 110819, China

² Information and Communication Branch, State Grid Liaoning Electric Power Co., Ltd., Shenyang 110004, China

Abstract Graph data, as a data type with strong expressive power, is difficult to model efficiently due to its complex structure. How to effectively capture its intrinsic information has become a challenging problem. Graph embedding methods have received increasing attention by mapping high-dimensional sparse graphs into low-dimensional dense feature vectors, while preserving the structural information of graphs. However, the existing reviews do not summarize the graph embedding methods comprehensively enough, especially paying less attention to complex graph embedding, which leads to the failure to systematically sort out the current status of research on graph embedding in dealing with diverse graph data. Therefore, this paper presents a systematic review of graph embedding learning methods from simple to complex graphs. Firstly, it gives the common definitions of various types of graphs and graph embedding. Secondly, it systematically summarizes the embedding methods on simple graphs, including shallow and deep embedding methods. Then, it summarizes the embedding methods on complex graphs according to the types of graphs, focusing on the application of deep embedding techniques in complex graph structures such as dynamic graphs, heterogeneous graphs, multiplex graphs, and hypergraphs, to fill the gaps in the existing literature that is insufficiently researched on complex graph structures. Finally, it discusses the practical application scenarios of graph embedding techniques, and looks forward to the future development directions.

Keywords Graph embedding, Graph representation, Deep learning, Neural network, Complex graph

1 引言

近年来,随着互联网技术及其应用的不断发展,产生了海量数据。其中,图结构数据作为一种可以有效存储和表示真实世界中复杂实体之间交互的数据类型,在现实生活中已经无处不在。例如,社交网络对多个用户间的交互关系进行

建模^[1],学术引文网络反映了学术论文之间的引用关系^[2],蛋白质网络表示蛋白质和蛋白质之间的相互作用^[3]等。在这些领域中,用图中的节点表示实体,边表示实体间的关系,可以帮助研究者高效地挖掘实体间的潜在信息。然而,图数据本身的特点为分析和处理带来了一定的挑战性。不同于规则的欧几里德数据(如图像和文本),图是一种非规则的非欧几里

到稿日期:2025-03-17 返修日期:2025-06-09

基金项目:国家自然科学基金(62432003,92267206,62032013)

This work was supported by the National Natural Science Foundation of China(62432003,92267206,62032013).

通信作者:赵宇海(zhaoyuhai@mail.neu.edu.cn)

德数据结构,同时具有数据维度高的特点,如何将其有效地嵌入向量空间中进行学习已经成为一个重要且具有挑战的问题。图嵌入(Graph Embedding, GE)作为一种有前景的表示学习方式,近年来已经引起了学术界的广泛关注。图嵌入旨在将原始的高维稀疏的图数据映射为低维稠密的向量(见图1),以最大程度地保留图的拓扑结构和属性信息,进而更好地应用在不同的下游图分析任务中,如社区检测^[4]、链接预测^[5]和节点分类^[6]等。

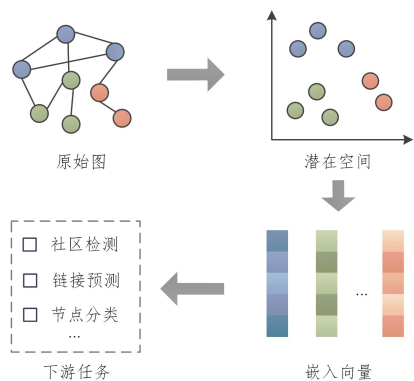


图1 图嵌入的学习过程

Fig. 1 Learning process of graph embedding

早期的工作主要集中于浅层图嵌入方法,特别是基于矩阵分解的方法和基于随机游走的方法。基于矩阵分解的方法^[7-8]关注于用矩阵表示图的属性(如节点对相似性),通常采用拉普拉斯矩阵映射或节点邻近矩阵分解方式将相似性矩阵映射到低维空间中,使得产生的节点表示可以反映出节点的连接关系。基于随机游走的方法^[9-11]利用不同的游走策略对图中的节点进行采样,然后将生成的节点序列输入 SkipGram模型^[12]中进行节点降维嵌入。在随机游走的过程中,模型能够捕获图中的全局和局部结构信息,具有较好的表达能力。这些方法虽然以简单易懂和可解释性强的优点在早期取得了一定的成功,但对图中深层信息的捕获是有限的。

随着深度学习技术的高速发展,尤其是图神经网络(Graph Neural Network, GNN)在图数据处理上实现的重大突破,基于深度学习的图嵌入方法^[13-15]逐渐成为研究热点。这类方法旨在通过应用深度网络模型,如图卷积网络和图注意力网络,来编码图数据的非线性结构,提取更加抽象的高层特征。这类方法很好地保留了图的属性和结构信息,同时克服了浅层方法中人工设计特征费时费力的缺点,很大程度上提高了图嵌入算法的有效性和灵活性,已经在节点分类和图分类等任务中取得了显著的成果。

上述的图嵌入方法尽管已经在简单图上取得了成功,但通常假设图结构是静态的,具有单一类型的节点和边,且边只描述成对节点的关系。实际上,现实世界中的图数据具有更加多样和复杂的特性,例如时间依赖关系、语义异质性、多重关系或高阶关联信息,这些特性使得简单图无法适应。相比之下,复杂图(图2展示了简单图和4种复杂图结构)表现出更强的表达能力:动态图(Dynamic Graph)描述了节点和边随时间演变的动态过程;异质图(Heterogeneous Graph)中包含多种类型的节点和边,如生物信息网络中的药物-疾病-基因

关系;多重图(Multiplex Graph)允许同一节点对之间存在多种语义不同的边;超图(Hypergraph)可以建模连接了两个以上节点的高阶关系,如引文网络中的共同作者关系。这些结构差异使得复杂图更适应于社交网络和生物信息学等实际领域,而简单图嵌入方法难以捕获这些信息。

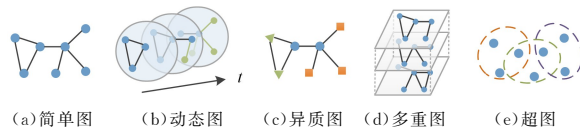


图2 多种类型的图

Fig. 2 Multiple types of graphs

相比于简单图嵌入方法,复杂图嵌入方法显著改变了研究范式,即从静态到动态,从单一类型到多种类型关系,从单一关系到多重关系,从低阶交互到高阶关联。具体来说,动态图嵌入方法^[16]通过学习不同时刻图中的节点和边的嵌入,能够更准确地反映动态图中的演化规律。异质图嵌入方法^[17]结合图中节点和边的异质性以及路径信息,对异质结构和语义进行建模。多重图嵌入方法^[18]同时考虑图中节点之间的多种类型的关系,通过融合不同关系类型的边信息,生成更为全面的嵌入表示。超图嵌入方法^[19]解决了图中的边与多个节点之间形成高阶关联的问题,有效表达了现实世界中多节点的高阶交互。总之,相较于简单图嵌入方法中简单的建模策略,上述这些针对复杂图的嵌入方法具有更强的针对性,能够更加灵活地描述多种真实的应用场景。

目前,已经有一些综述文献^[20-25]对图嵌入学习方法进行了归纳。然而,这些文献仍然存在一些明显的不足。1)当前的综述对简单图嵌入方法归纳得不够全面,尤其在对基于深度学习的图嵌入方法的分类上,许多文献未能清晰划分这些方法之间的边界,导致不同方法的对比不够直观。例如, Wu等^[23]仅通过使用不同类型的模型将深度图嵌入方法分类,没有体现出不同方法的差异;另外,该文献没有包含最新的研究工作,未能充分反映图嵌入学习的最新进展。2)在复杂图嵌入分析方面,现有综述研究较少且总结不够充分。大多数文献主要集中于简单图嵌入方法,缺乏对复杂图嵌入的探讨。尽管部分文献已经对某些复杂图结构类型(如异质图、超图等)进行了讨论,但大多数文献仅关注于单一类型的复杂图,缺乏对多种复杂图的统一和全面的分析。不同类型的复杂图往往在建模方面存在差异,而现有文献未能系统地对这些工作进行综合比较。因此,针对简单图和复杂图嵌入方法的系统梳理和深入分析是有必要的,以填补现有文献的空白,并为后续研究提供更全面的视角。

本文系统地梳理了图嵌入方法,从简单图到复杂图逐步深入,全面探讨了不同类型图的嵌入技术。本文第2章首先给出了不同类型的图和图嵌入的常见定义;第3章详细归纳了在简单图上应用的嵌入方法,包括浅层和深度图嵌入方法,具体涵盖了基于矩阵分解和随机游走的浅层嵌入方法,以及基于卷积操作、注意力机制、编码-解码架构和对抗学习的深度嵌入方法;第4章进一步扩展到复杂图上的嵌入学习,从动态图、异质图、多重图和超图的角度出发,对现有的各种复杂图上的嵌入方法进行系统的总结和归纳;第5章探讨了图嵌

入学习研究在实际中的应用场景和潜在的研究方向;最后总结全文。

2 问题定义

1)简单图:一个简单图可表示为 $G=(V,E)=(\mathbf{X},\mathbf{A})$,其中 $V=\{v_1,\dots,v_n\}$ 和 $E=\{e_1,\dots,e_m\}$ 分别是 G 中 n 个节点和 m 条边的集合。对任意边, $e_{ij}=(v_i,v_j)$ 表示 v_i 和 v_j 之间的连接关系。 \mathbf{X} 和 \mathbf{A} 分别是节点特征矩阵和邻接矩阵, \mathbf{A} 指示图中节点之间的连接关系,即若节点 v_i 和 v_j 之间存在一条边,则 $A_{ij}=1$,否则 $A_{ij}=0$ 。

2)动态图:动态图中节点和边的状态会随着时间进行演变,可以按时间变化表示为一系列的演化图 $G^d=\{G_1,\dots,G_T\}$,其中 T 表示演化图的数量。每个演化图 $G_t=(V^t,E^t)$ 表示 t 时刻节点集 V^t 和边集 E^t 的状态。

3)异质图:异质图 $G^h=\{V^h,E^h,T^v,T^e\}$ 中具有多种类型的节点或边,即 $|T^v|>1$ 或 $|T^e|>1$,其中 V^h 和 E^h 分别是异质图中的节点集和边集, T^v 和 T^e 分别表示节点类型和边类型。在异质图中,不同类型的节点和边的特征处于不同的特征空间中,节点和边的特征集合分别表示为 $\{\mathbf{X}_1,\dots,\mathbf{X}_{|T^v|}\}$ 和 $\{\mathbf{X}'_1,\dots,\mathbf{X}'_{|T^e|}\}$ 。

4)多重图:多重图中节点之间具有多种关系,可表示为 $G^m=(V^m,\{E^m_1,\dots,E^m_r\})$,其中 V^m 表示多重图的节点集合, r 表示不同类型关系的数量。多重图中的 r 种关系也可以表示为 r 个邻接矩阵 $\{A^m_1,\dots,A^m_r\}$,每种关系的邻接矩阵描述了节点之间的连接关系。

5)超图:超图可表示为 $G^p=(V^p,E^p)$,其中 $V^p=\{v^p_1,\dots,v^p_n\}$ 和 $E^p=\{E^p_1,\dots,E^p_m\}$ 分别表示节点集合和边集合。在超图中,一条超边可以同时连接两个以上的节点,即对于任意超边, $E^p_q=\{v^p_1,\dots,v^p_k\}\subseteq V^p$ 。超图结构信息也可以用关联矩阵 $\mathbf{M}\in\mathbb{R}^{n\times m}$ 表示:

$$\mathbf{M}_{ij}=\begin{cases} 1, & v^p_i\in E^p_j \\ 0, & v^p_i\notin E^p_j \end{cases} \quad (1)$$

超图的邻接矩阵 \mathbf{A}^p 可定义为 $\mathbf{A}^p=\mathbf{M}^T\mathbf{M}$,其中 A^p_{ij} 表示超边 i 和超边 j 共有的节点数量。

6)图嵌入:给定任意一个图 G ,图嵌入将其映射到一个低维空间中,同时保留原始图中的结构信息,其映射函数被定义为 $f:v_i\rightarrow\mathbf{Y}_i\in\mathbb{R}^d$,其中 d 表示嵌入向量的维度, \mathbf{Y}_i 表示编码后的节点表示。

为了便于理解,表 1 列出了本文常用符号及其含义。

表 1 符号描述

Table 1 Symbol description

符号	含义
$G=(V,E)$	具有节点集 V 和边集 E 的简单图
G^d,G^h,G^m,G^p	动态图,异质图,多重图,超图
$G_t=(V^t,E^t)$	t 时刻节点集 V^t 和边集 E^t 的状态
T^v,T^e	异质图中节点类型和边类型
$\{A^m_1,\dots,A^m_r\}$	多重图的 r 个邻接矩阵
\mathbf{M}	超图关联矩阵
\mathbf{X}	节点特征矩阵
\mathbf{A}	邻接矩阵
\mathbf{Y}_i	节点嵌入
n	节点个数
d	嵌入向量的维度

3 简单图上的嵌入方法

简单图是图结构中最基础且常见的形式,它由一组具有相同类型的节点和连接节点对的无向边组成,且节点对之间至多只有一条边。由于具有直观且易处理的特性,简单图在图嵌入研究中得到了广泛应用,成为许多算法的基础。根据嵌入方式的不同,简单图嵌入方法可以分为浅层和深度图嵌入方法。浅层图嵌入方法较为直接,通常基于矩阵分解或随机游走等传统手段,能够在保留图结构信息的同时生成低维节点表示。相比之下,深度嵌入方法借助深度学习技术,更好地捕捉图中的非线性特征并适应复杂的场景。

3.1 浅层图嵌入方法

早期的图嵌入方法通常采用矩阵分解方式对节点的相似性矩阵进行分解,来获得图中节点的低维表示,但这种方法带来的高计算复杂度使其难以扩展到大规模图中。随机游走通过在图中随机地进行采样,可以将非规则的图数据转变为序列数据,而且在游走过程中能够近似图中的多种属性(如节点中心性和相似性),因此逐渐受到关注。鉴于上述的基于矩阵分解和随机游走的方法均不涉及深层特征的学习,本文将这两种方法归类为浅层图嵌入方法。

3.1.1 基于矩阵分解的方法

基于矩阵分解的方法通常用一个矩阵表示图中节点对之间的关联信息,然后采用矩阵分解技术对该矩阵进行分解,以获得节点的低维嵌入表示。根据学习方式的不同,将这类方法进一步细分为拉普拉斯矩阵映射方法和节点邻近矩阵分解方法。这两种方法的共同目标是构造一个拉近原始图中节点对的相似性与嵌入后节点对的相似性的损失函数,但它们依旧存在本质区别:前者以拉普拉斯矩阵为基础,旨在通过谱分解捕获图的流形结构,以保持节点之间的局部关系;后者直接对局部或高阶邻接矩阵进行非负矩阵分解或奇异值分解,以保留节点间的潜在关系。

1)拉普拉斯矩阵映射方法

拉普拉斯矩阵映射方法源自谱聚类理论,旨在通过节点对之间的相似性来确定节点在低维空间中的嵌入位置,从而有效地反映出原始图中的拓扑结构。具体来说,该方法在节点嵌入过程中遵循以下原则:若两个节点对在图中相似,则它们在低维空间中的距离应该较小,以保留图中节点之间的关系;否则,将对映射后距离变远的相似节点施加惩罚。

经典的拉普拉斯映射方法之一是 LE^[26],其目标是使得相似的节点嵌入在低维空间时尽可能地靠近,即当两个节点的相似性 S_{ij} 较高时,它们的嵌入距离 $\|\mathbf{Y}_i-\mathbf{Y}_j\|^2$ 应当较小。该方法通过最小化流形约束来捕获节点之间的局部关系:

$$\frac{1}{2}\sum_{i,j}\|\mathbf{Y}_i-\mathbf{Y}_j\|^2S_{ij}=\text{tr}(\mathbf{Y}^T\mathbf{L}\mathbf{Y}) \quad (2)$$

其中, $\mathbf{L}=\mathbf{D}-\mathbf{S}$ 是拉普拉斯矩阵, \mathbf{D} 是对角度矩阵。为了避免琐碎的方案,添加约束 $\mathbf{Y}^T\mathbf{D}\mathbf{Y}=\mathbf{I}$ 消除嵌入时的平凡解, \mathbf{I} 是单位矩阵。上述问题可表达为:

$$\min_{\mathbf{Y}^T\mathbf{D}\mathbf{Y}=\mathbf{I}}\mathbf{Y}^T\mathbf{L}\mathbf{Y}=\min_{\mathbf{Y}^T\mathbf{D}\mathbf{Y}=\mathbf{I}}\frac{\mathbf{Y}^T\mathbf{L}\mathbf{Y}}{\mathbf{Y}^T\mathbf{D}\mathbf{Y}} \quad (3)$$

实际上,该方法通过选择节点的 k 个邻居构建一个 kNN

图,并利用拉普拉斯算子 $Lf = \lambda Df$ 来求解特征向量 f ,最后使用最小 k 个特征值对应的特征向量作为节点嵌入。然而,该方法较为关注权重较大的节点对,而忽视权重小的节点对,可能丢失局部拓扑信息。

LE方法尽管可以在嵌入过程中保持节点对之间的局部关系,但依旧存在信息丢失、易受噪声干扰等问题。为此,Meng等^[27]提出修改LE方法的目标函数,通过使用高密度核心节点之间的相似性来表示多组分数据的相对位置,以避免丢失数据的相对位置信息。为了抵抗噪声干扰,Cai等^[28]提出的LRLGL方法结合流形学习和子空间学习来构建低秩的拉普拉斯矩阵,并引入秩约束来构造损失函数,以生成对噪声鲁棒的表示。

2) 节点邻近矩阵分解方法

节点邻近矩阵分解方法通常假设节点之间的相似性与其嵌入向量的内积成正比,即在向量空间中,相似的节点应该具有较大的嵌入内积。基于这一假设,可以直接对节点邻近矩阵进行分解来获得节点的低维表示。具体来说,给定一个节点邻近矩阵 S ,其中每个元素表示节点对之间的相似性,通过最小化矩阵与节点嵌入点积之间的损失 $\|S - YY^T\|$,可以学习到最优的节点嵌入 Y ,其中 Y_c 是上下文节点嵌入。

Graph Factorization(GF)^[29]是一种典型的直接对邻接矩阵进行分解的方法,通过使用内积模型来获得节点嵌入,从而捕获节点对间的关联信息。此外,该方法还引入一个额外的L2正则化项,以保证在数据稀疏时问题的合理性。它的目标函数表示如下:

$$\frac{1}{2} \sum_{(i,j) \in E} (S_{ij} - \langle Y_i, Y_j \rangle)^2 + \frac{\lambda}{2} \sum_i \|Y_i\|^2 \quad (4)$$

其中, λ 是正则化系数。为了降低在大规模图上的应用成本,GF采取分布式框架,将图数据划分并分配到多个机器上进行并行处理。然而,GF方法是通过内积方式构建节点对之间的相似性,较为依赖图的一阶局部结构信息,因此无法充分利用全局信息。

为了捕获图中的全局结构信息,Cao等^[7]提出了GraRep。该方法利用不同阶的概率转移矩阵来表示节点之间的多种关系信息。具体来说,给定图的一阶转移概率 $T = D^{-1}A$,其表示从一个节点到另一个节点之间的转换概率,而节点的 k 阶转移概率 T^k 则能够捕获图的全局结构信息;然后,通过对 T^k 进行奇异值分解来获得节点表示。GraRep能够更全面地捕获图的全局拓扑信息,但由于奇异值分解操作的计算复杂度比较高,其很难扩展到大规模的图数据上。

基于上述方法的思想,Qiu等^[30]提出的NetSMF模型利用谱稀疏化理论,通过生成稀疏矩阵来提高图嵌入学习的效率,并对矩阵进行显式分解以扩展到大规模图数据上。Agibetov^[31]提出了一种基于逐点互信息(Pointwise Mutual Information, PMI)矩阵的低秩分解方法,通过模拟随机游走过程来学习节点的潜在向量表示,从而提高下游任务的精度。Wan等^[32]提出的Z-NetMF方法,将随机游走采样引入矩阵分解,将每一步的随机游走看作从邻居中聚合信号的过程,能够生成具有全局结构信息的表示。

3.1.2 基于随机游走的方法

基于随机游走的方法源于自然语言处理中的单词嵌入

模型^[33]。此类方法通常采用随机游走策略来探索图结构,并生成节点序列,然后将节点序列输入语言模型Skip-Gram^[12],以学习嵌入表示。在这一过程中,此类方法不仅能够捕获图的某些特性(如社区结构和节点中心性),而且在多个图分析任务中的表现都较为出色。

DeepWalk^[9]是一种经典的基于随机游走的算法,其借鉴了词-向量建模思想,通过随机游走将图嵌入转换为词嵌入问题来处理。具体来说,DeepWalk通过随机游走对节点邻居进行采样,以生成定长的节点序列。然后,这些节点序列被输入Skip-Gram模型中(见图3),通过最大化窗口范围内节点间的共现概率,能够保证具有相似邻居的节点在低维空间中具有相似的嵌入表示,从而有效捕捉图的局部结构。其目标函数如式(5)所示:

$$\min_{\Phi} -\log \Pr(\{v_{i-w}, \dots, v_{i-1}, v_{i+1}, \dots, v_{i+w}\} | \Phi(v_i)) \quad (5)$$

其中, w 是窗口大小, Φ 是图嵌入映射函数。尽管该方法能够有效地学习节点间的局部关系信息,但其均匀采样的策略限制了模型对不同邻域信息的捕获能力。

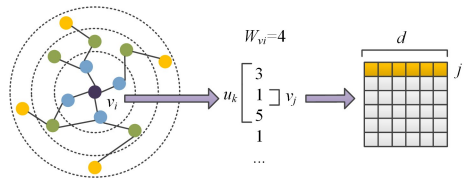


图3 DeepWalk模型架构^[9]

Fig. 3 Architecture of DeepWalk model^[9]

在DeepWalk模型的基础上,node2vec^[10]进一步提出了一种有偏的随机游走策略,以突破其在邻域信息获取上的局限性。具体来说,node2vec通过引入两个参数 p 和 q ,来调节随机游走过程中的广度优先搜索和深度优先搜索之间的平衡,以捕获图中的全局和局部结构。该方法能够依据上一步的节点位置动态调整下一步的跳转概率,增强了邻域探索的灵活性,使得生成的嵌入表示的信息更加丰富。

然而,DeepWalk和node2vec方法在捕获结构等价性方面存在不足。为此,Ribeiro等^[34]提出struc2vec,该方法认为在图中具有相同局部结构的两个节点可能具有相似的嵌入表示,反之则应具有显著不同的表示。基于这一理论,struc2vec构建了一个分层模型,用于衡量结构相似性。具体来说,该方法根据不同的邻域范围来比较节点对之间的局部结构特征,并构建一个加权的多层图以对结构相似性进行编码,然后为节点生成结构上下文来获得节点的嵌入表示。

在上述算法的基础上,研究者针对特定问题提出了一些改进算法,以进一步优化图嵌入效果。例如,Rahman等^[35]延续node2vec的有偏随机游走思想,提出了Fairwalk方法以解决图嵌入中的潜在偏差问题。该方法根据敏感属性值将节点邻居划分为多个组,每组随机采样一个节点进行随机游走,从而生成具有全局性的嵌入表示。为了保留更多的局部和全局结构信息,Zhang等^[36]提出了NRL-RWR算法,该算法在随机游走中引入注意力机制,以自动识别重要性节点,而且可以学习节点的概率分布,从而保留更多的结构信息。此外,为了提高随机游走的收敛速度,Wang等^[37]提出了CNARW

方法,该方法利用最近访问节点与下一跳候选节点的共同邻居来选择下一步访问的节点,并结合无回溯随机游走和基于更多节点访问的无偏采样方案,降低了采样成本并加快了收敛速度。

3.2 深度图嵌入方法

近年来,深度神经网络^[38]在提取数据的深层特征方面取得了显著进展,尤其是图神经网络在图数据上实现了重大突破。因此,基于深度模型架构的嵌入方法已经成为当前的研究热点。与传统的浅层图嵌入方法相比,这类方法直接从原始图中学习高层次特征表示,从而建模实际生活中的复杂问题。根据模型架构的不同,本节将深度图嵌入方法分为4类:基于卷积操作的方法、基于注意力机制的方法、基于编码-解码架构的方法和基于对抗学习的方法。

3.2.1 基于卷积操作的方法

基于卷积操作的方法主要依赖于卷积神经网络(Convolutional Neural Network, CNN)或图卷积网络(Graph Convolutional Network, GCN)模型及其变体在图数据上进行信息传递。通过这一过程,生成的节点嵌入能够描述全局和局部结构信息,从而提升下游任务的性能。CNN^[39]的基本架构包括输入层、隐藏层和输出层,其中隐藏层由卷积层、池化层、激活层和全连接层组成。卷积层通过卷积运算提取数据特征,池化层对数据进行压缩,激活层通过非线性变换增强表达能力,全连接层则用于连接特征。GCN^[40]试图将CNN的卷积思想从欧氏空间扩展到非欧氏空间的图数据上,通过图卷积运算对邻居节点的信息进行聚合。GCN的输入包括节点属性和拓扑结构,在每层网络中聚合节点及其邻居信息以更新节点表示,并将其作为下一层卷积的输入。根据卷积算子构建方式的不同,将此类方法分为谱卷积方法和空间卷积方法。这两种方法都旨在对图结构进行卷积操作,以有效聚合节点特征。它们的不同之处在于:谱卷积方法较为关注谱域信息,在图拉普拉斯特征空间中进行傅里叶变换定义卷积运算;而空间卷积方法则针对特征的迭代更新,从信息传播的角度出发,直接在节点的局部邻域中定义卷积操作。

1) 谱卷积方法

谱卷积方法的理论基础源于图信号处理,通过对图信号进行傅里叶变换,将节点特征转换到谱域上以执行卷积操作。具体来说,图信号的频率通过图的拉普拉斯矩阵 $\mathbf{L} = \mathbf{I} - \mathbf{D}^{-1/2} \mathbf{A} \mathbf{D}^{-1/2}$ 来表示。拉普拉斯矩阵的特征分解表示为 $\mathbf{L} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^T$,其中 \mathbf{U} 是特征向量, $\mathbf{\Lambda}$ 是特征值构成的对角矩阵。通过这一分解,节点特征 \mathbf{x} 可以通过特征矩阵 \mathbf{U} 投影到谱域,从而实现图信号的傅里叶变换,形式为 $\mathcal{F}(\mathbf{x}) = \mathbf{U}^T \mathbf{x}$ 。傅里叶逆变换可以将谱域中的信号转换回去,形式为 $\mathcal{F}^{-1}(\mathbf{x}') = \mathbf{U} \mathbf{x}'$,其中 \mathbf{x}' 是变换后的节点特征。利用卷积定理以及傅里叶变换和逆变换,图域上的卷积定义为:

$$g_\theta * \mathbf{x} = \mathbf{U} g_\theta \mathbf{U}^T \mathbf{x} \quad (6)$$

其中, g_θ 是过滤器。大多数谱方法的卷积操作是基于式(6)进行定义的,核心区别在于过滤器 g_θ 的设计。

最早的谱卷积方法是由Bruna等^[41]提出的Spectral-CNN,该方法将传统的CNN从网格结构数据(如图像)扩展

到不规则的图数据上。利用傅里叶域中卷积的性质,该方法将卷积操作定义在图拉普拉斯的特征向量上。其在第 $l+1$ 层的卷积操作可表示为:

$$\mathbf{X}_j^{(l+1)} = \sigma(\mathbf{U} \sum_{i=1}^{d^{in}} g_{\theta_j}^{(l)} \mathbf{U}^T \mathbf{X}_i^{(l)}), j=1, \dots, d^{out} \quad (7)$$

其中, d^{in} 和 d^{out} 分别表示输入和输出维度, σ 是激活函数。

SpectralCNN为图上的卷积操作奠定了理论基础。然而,该方法利用傅里叶基的乘法操作来定义式(7)的卷积,需要进行矩阵分解操作,导致计算复杂度较高。为了解决这一问题,ChebNet^[41]和GCN^[42]相继被提出。

ChebNet^[41]在卷积操作中引入了切比雪夫多项式,以提升计算效率。该方法证明 g_θ 可以近似为切比雪夫多项式的截断展开式,即 $g_\theta = \sum_{i=0}^K \theta_i T_i(\mathbf{A}')$,其中 $\mathbf{A}' = 2\mathbf{A}/\lambda_{\max} - \mathbf{I}$,(λ_{\max} 是最大特征值),避免了直接分解拉普拉斯矩阵的高计算代价。通过利用切比雪夫多项式的递推关系 $T_i(\mathbf{x}) = 2\mathbf{x}T_{i-1}(\mathbf{x}) - T_{i-2}(\mathbf{x})$,其中 $T_0(\mathbf{x}) = 1, T_1(\mathbf{x}) = \mathbf{x}$,定义图卷积操作:

$$g_\theta * \mathbf{x} = \sum_{i=0}^K \theta_i T_i(\mathbf{L}') \mathbf{x} \quad (8)$$

其中, $\mathbf{L}' = 2\mathbf{L}/\lambda_{\max} - \mathbf{I}, \theta_i$ 为系数。式(8)具有K局部化的特性,即每个节点的表示只受到K个邻居的影响。这种特性保证了 g_θ 有效捕获图中的局部信息。

为了进一步降低计算复杂度,GCN^[42]在ChebNet的基础上采用了一阶切比雪夫多项式近似。通过对图卷积进行谱卷积的局部一阶近似,即设定 $K=1$,并令最大特征值 $\lambda_{\max} = 2$,将式(8)的卷积操作简化为 $g_\theta * \mathbf{x} = \theta_0 \mathbf{x} - \theta_1 \mathbf{D}^{-1/2} \mathbf{A} \mathbf{D}^{-1/2} \mathbf{x}$ 。GCN进一步令 $\theta = \theta_0 = \theta_1$,并为GNN提供了一个逐层传播规则,如下所示:

$$\mathbf{H}^{(l+1)} = \sigma(\tilde{\mathbf{D}}^{-1/2} \tilde{\mathbf{A}} \tilde{\mathbf{D}}^{-1/2} \mathbf{H}^{(l)} \mathbf{W}^{(l)}) \quad (9)$$

其中, $\tilde{\mathbf{A}} = \mathbf{A} + \mathbf{I}, \tilde{\mathbf{D}}_i = \sum_j \tilde{\mathbf{A}}_{ij}, \mathbf{W}^{(l)}$ 是权重矩阵, $\mathbf{H}^{(l)}$ 是第 l 层的节点特征矩阵,且 $\mathbf{H}^{(0)} = \mathbf{X}$ 。

GCN^[42]的成功激发了人们对谱方法的深入研究,促使研究者进一步探索如何更有效地在图数据上应用谱卷积。为了提高计算效率,Wu等^[43]提出了SGC模型。SGC通过去除GCN层之间的非线性,将其转换为一个简单的线性变换,从而有效降低了GCN的复杂性。考虑到图中的边特征信息,CensNet^[44]设计了一个边-节点转换的网络,根据输入图构造一个线图来表示图中的边之间的相邻关系,并在输入图和线图上执行图卷积操作以对节点和边同时进行嵌入。CW-GCN^[45]将GCN扩展到领域适应任务中进行图嵌入。该方法研究GCN中的相关熵诱导损失,并引入Wasserstein距离度量图之间的边缘分布差异,以更好地传递图中的丰富知识。

2) 空间卷积方法

谱卷积方法尽管在图嵌入学习中取得了显著的成就,但仍存在一些未解决的问题。例如,谱卷积方法依赖于多个数学假设,无法灵活地迁移到更复杂的图结构数据中。此外,GCN通常采用直推式学习,导致只能处理固定大小的图,难以扩展到大规模图数据上。空间卷积方法的提出有效弥补了这些缺陷,这类方法通过直接在图上定义卷积操作,利用中心

节点的邻居节点特征的聚合来更新中心节点的特征。

Niepert 等^[46]提出的 PATCHY-SAN 模型激发了人们对空间卷积方法的兴趣。该模型的主要思想是将图结构转换成序列结构,从而在序列结构上应用卷积操作。具体来说,该模型首先通过图标注过程对节点进行排序,并在图中遍历选择固定长度的节点序列;然后,使用广度优先搜索策略为每个节点选取固定数量的邻居节点,并对选择的邻域进行归一化;最后,将节点序列输入 CNN 中学习嵌入表示。该模型成功地将 CNN 的思想扩展到图上进行卷积操作,但是依旧难以应用于大规模图数据。

为了提高图卷积方法的可扩展性,GraphSAGE^[47]将 GCN 扩展到归纳式无监督框架中。具体来说,该方法首先对每个节点的邻居节点进行采样,并训练一个聚合函数以聚合节点的邻居信息,然后将节点自身的信息与聚合后的邻居信息进行连接和激活,以此来更新每个节点的特征。其中的图卷积过程可定义为:

$$\mathbf{H}_v^{(l)} = \sigma(\mathbf{W}^{(l)} \cdot f_C(\mathbf{H}_v^{(l-1)}, f_A(\{\mathbf{H}_u^{(l-1)}, \forall u \in N_v\}))$$
(10)

其中, f_C 和 f_A 分别表示连接和聚合操作。

GraphSAGE 不需要将整个图输入来生成节点嵌入,而是通过对节点的邻居进行采样和信息聚合来生成嵌入表示。该模型不仅能够学习每个节点邻居的拓扑结构和节点特征的分布,而且能够很好地扩展到大型图上。

最近,一些新的方法被提出,以进一步提高图嵌入的表示能力。例如, Hu 等^[48]提出的 GraphAIR 模型,通过同时对邻居特征和邻居之间的交互进行建模,能够更好地捕获图结构中复杂的非线性关系。Tailor 等^[49]提出的高效图卷积模型 EGC,通过设计具有空间变化的自适应滤波器来优化卷积操作的效率,同时降低内存消耗。为了同时学习边和节点嵌入, Zhou 等^[50]提出了 CEN-DGCNN 算法,将节点特征和边特征整合到一个消息传递框架中,同时提取节点之间的长距离依赖关系,以捕获局部和全局结构信息。

3.2.2 基于注意力机制的方法

不同于卷积网络将节点的所有邻居节点平等加和,注意力机制为每个邻居节点分配不同的权重系数,从而能自适应地识别重要的邻居节点。基于注意力机制的方法通常采用注意力网络对图中每个节点的邻居信息进行自适应地聚合,或者使用注意力机制在不同级别进行特征组合,以生成具有丰富信息的嵌入表示。按照注意力机制作用在神经网络层内外的方式,这些方法可以分为层内注意力方法和层间注意力方法。

1) 层内注意力方法

层内注意力方法通常将注意力机制引入单层神经网络中的局部邻域上,通过计算局部邻居中不同节点的权重来更新节点表示。最著名的模型之一是图注意力网络 GAT^[51]。GAT 首先通过可学习的注意力机制来获得注意力系数:

$$\alpha_{ij} = \frac{\exp(\text{LeakyReLU}(\vec{\mathbf{a}}^T [\mathbf{W}\vec{\mathbf{h}}_i \parallel \mathbf{W}\vec{\mathbf{h}}_j]))}{\sum_{k \in N_i} \exp(\text{LeakyReLU}(\vec{\mathbf{a}}^T [\mathbf{W}\vec{\mathbf{h}}_i \parallel \mathbf{W}\vec{\mathbf{h}}_k]))}$$
(11)

其中, \parallel 表示连接操作, LeakyReLU 是激活函数, $\vec{\mathbf{a}}$ 是单层前馈神经网络, $\vec{\mathbf{h}}_i$ 是 v_i 的特征, N_i 是 v_i 的邻居节点。

然后,使用多头注意力机制对具有不同重要程度的节点邻居进行有注意力的加权求和,以获得嵌入表示:

$$\vec{\mathbf{h}}_i' = \sigma \left(\frac{1}{K} \sum_{k=1}^K \sum_{j \in N_i} \alpha_{ij}^k \vec{\mathbf{w}}_{hj} \right)$$
(12)

其中, K 表示注意力头的数量。通过多个注意力层的叠加,每个节点可以聚合来自多跳邻居的加权信息,从而更有效地捕获图中的结构信息。

由于注意力模型能够捕获更多的信息,研究者接着对其进行扩展,提出了一系列更为灵活的图嵌入方法。例如, He 等^[52]提出的 ARGAT 将 GAT 应用在生成对抗学习中,使用 GAT 编码邻接矩阵和节点特征,并通过两个解码器进行重构,最后将生成的对抗网络模块作为正则化项,从而学习到更加鲁棒的嵌入表示。针对图中的噪声问题, Ye 等^[53]提出了 SGATs,该方法通过学习稀疏注意力系数,并将其应用于所有 GNN 层以识别噪声边来生成高质量嵌入。He 等^[54]提出的 HGRN 利用注意力机制自适应地加权多跳邻居的节点特征,并根据节点间的相似性更新图结构,从而能够利用高阶信息进行嵌入。

2) 层间注意力方法

对于层间注意力方法,注意力被视为特征融合的一种操作,用于在不同级别或视图的特征空间中选择性地聚合信息,以实现跨层融合操作。Xie 等^[55]提出的 MGAT 模型,将注意力机制应用到多视图图嵌入学习中。具体来说, MGAT 首先对每个视图编码,学习 Q 个视图特定的节点嵌入。由于不同的视图之间存在信息交互,首先通过基于注意力机制的聚合模型学习不同视图在节点 v_i 上的注意力权重:

$$\beta_{ik} = \frac{\exp(\mathbf{t}_k^T \mathbf{c}_i)}{\sum_{j=1}^Q \exp(\mathbf{t}_j^T \mathbf{c}_i)}$$
(13)

其中, \mathbf{t}_k 是第 k 个视图的可学习权重, \mathbf{c}_i 是 v_i 在所有视图上的表示 $\{\mathbf{x}_{i1}, \dots, \mathbf{x}_{iQ}\}$ 的拼接向量。通过此权重,可以将每个视图的节点嵌入集成到全局的节点嵌入中:

$$\mathbf{h}_i = \sum_{k=1}^Q \beta_{ik} \cdot \mathbf{x}_{ik}$$
(14)

除了 MGAT 之外, Huang 等^[56]提出的 SGDMC 模型将注意力机制应用于多视图学习中。SGDMC 利用节点特征的高斯相似性和伪标签的余弦相似性来计算每条边的注意力系数,以增强图注意力层的聚合能力,并利用样本加权策略使得重要节点获得更多关注,从而减轻噪声样本对模型性能的负面影响。Wang 等^[57]提出的多通道 GCN 模型 AM-GCN,基于特征相似性图和原始拓扑图互补的假设,利用注意力机制自适应地学习这两者之间的关联信息,从而生成更具丰富关联信息的嵌入表示。

3.2.3 基于编码-解码架构的方法

基于编码-解码架构的方法旨在利用自编码器、稀疏自编码器或去噪自编码器等模型,或者专门为图结构学习设计的图自编码器、变分图自编码器及其变体,对图中的特征或结构信息进行重建,以捕获图中的潜在信息。根据模型类型的不同,这些方法可分为深度自编码器方法和图自编码器方法。

1) 深度自编码器方法

深度自编码器 (AutoEncoder, AE) 由 Rumelhart 等^[58]

提出,是一种典型的无监督学习模型,其架构主要由编码器和解码器组成。编码器通过对输入数据进行非线性建模,实现数据降维;解码器则利用编码后的低维表示重构出输入数据。整个过程旨在最小化输出表示与输入数据之间的重构误差,具体定义如下:

$$\mathcal{L} = \sum_{i=1}^n \ell(\text{DEC}(\text{ENC}(\mathbf{x}_i)), \mathbf{x}_i) \quad (15)$$

其中,ENC和DEC分别表示编码器和解码器。

Wang等^[59]提出的SDNE是一个经典的基于自编码器的图嵌入模型。为了保留局部和全局结构信息,该模型提出有监督和无监督模块,以联合探索一阶和二阶相似性。在有监督模块中,SDNE引入拉普拉斯特征映射以捕获一阶相似性,确保生成的表示符合平滑性假设,以保留局部信息。其目标函数定义为:

$$\mathcal{L}_{1st} = \sum_{i,j=1}^n \mathbf{A}_{ij} \|\mathbf{Y}_i^{(L)} - \mathbf{Y}_j^{(L)}\|_2^2 \quad (16)$$

无监督模块通过重构节点邻居结构来捕获二阶相似性,从而保留全局信息。其目标函数为:

$$\mathcal{L}_{2nd} = \sum_{i=1}^n \|\hat{\mathbf{x}}_i - \mathbf{x}_i\|_2 \odot B_i \quad (17)$$

其中, \mathbf{x}_i 是邻接矩阵的第*i*行; $\hat{\mathbf{x}}_i$ 是重构向量表示; $B_i = \{B_{ij}\}_{j=1}^n$,若 $A_{ij}=0$,则 $B_{ij}=1$,否则 $B_{ij}=\beta>1$ 。

SDNE尽管能保留图中的一阶和二阶相似性,但在处理大型图时的计算效率较低,实际应用能力受到限制。

Cao等^[13]提出的DNNGR是另一种经典的方法,其通过结合随机游走和去噪自编码器DAE,来捕获图的非线性结构信息。该方法的核心流程包括3个步骤:首先,通过随机冲浪模型生成概率共现矩阵;接着,根据此矩阵计算正点互信息矩阵,以挖掘图的高阶结构信息;最后,利用堆叠的DAE生成节点的低维嵌入。尽管DNNGR在捕获图的高阶结构信息方面表现出色,但该模型忽略了图中节点或边的属性信息。

2) 图自编码器方法

图自编码器(Graph AutoEncoder, GAE)是一种结合GNN和AE的网络模型^[60],被广泛用于无监督图学习任务。GAE包含一个GCN编码器和内积解码器模块。GCN编码器将节点特征和结构信息映射到低维空间中($\mathbf{Z} = \text{GCN}(\mathbf{X}, \mathbf{A})$);内积解码器通过嵌入的内积重构邻接矩阵($\hat{\mathbf{A}} = \sigma(\mathbf{Z}\mathbf{Z}^T)$)。其目标是 minimized 重构邻接矩阵 $\hat{\mathbf{A}}$ 和真实邻接矩阵 \mathbf{A} 的交叉熵损失 $\ell = \sum_{i,j=1}^n -A_{ij} \log \hat{A}_{ij} - (1 - A_{ij}) \log(1 - \hat{A}_{ij})$ 。通过直接对图结构进行编码和解码,GAE能够有效学习图的深层次的拓扑结构和节点属性信息。

为了充分挖掘图中的潜在信息,Sun等^[61]提出了基于双解码器的图嵌入模型DGAE,旨在将图的拓扑结构和节点特征同时嵌入低维表示中。具体来说,DGAE首先利用GCN编码拓扑结构和节点属性,然后通过内积解码器和图卷积解码器分别重构图的邻接矩阵和节点特征矩阵,并最大限度地降低它们的重构损失。DGAE的双解码器结构充分利用了图中隐藏的丰富信息,从而生成更高质量的嵌入。

为了更灵活地学习节点特征,Chen等^[62]提出了DLR-

GAE方法,该方法结合矩阵分解与GAE模型学习图嵌入。具体来说,该方法由两个平行的GAE组成,分别对语义图和拓扑图进行低秩矩阵分解,然后根据因子矩阵和嵌入矩阵重构邻接矩阵,其目标函数为 $\|\mathbf{A} - \mathbf{Y}\mathbf{Y}^T\|_F^2 + g(\mathbf{U})$ 。其中, \mathbf{U} 是潜在因子矩阵, $g(\mathbf{U})$ 是正则化项。通过在两个并行GAE之间共享相同的因子矩阵 \mathbf{U} ,该方法能够衡量两个图之间的一致性。然而,由于矩阵分解过程涉及大量矩阵运算,因此其计算复杂度较高,不适合处理大规模图数据。

基于GAE的方法虽然能够捕获图的结构信息,但在重构图结构时可能存在过拟合问题。为此,研究者开始转向研究基于变分图自编码器VGAE^[63]的方法,以学习更具鲁棒性的图嵌入表示。在VGAE中,编码器用于学习每个节点对应的潜在变量分布,其形式为:

$$q(\mathbf{Y}|\mathbf{X}, \mathbf{A}) = \prod_{i=1}^N q(\mathbf{Y}_i|\mathbf{X}, \mathbf{A}) \quad (18)$$

$$q(\mathbf{Y}_i|\mathbf{X}, \mathbf{A}) = \mathcal{N}(\mathbf{Y}_i|\boldsymbol{\mu}_i, \text{diag}(\boldsymbol{\delta}_i^2)) \quad (19)$$

其中, $\boldsymbol{\mu}_i$ 和 $\boldsymbol{\delta}_i$ 是均值矩阵 $\boldsymbol{\mu}$ 和方差矩阵 $\boldsymbol{\delta}$ 的第*i*行, $\boldsymbol{\mu} = \text{GCN}_{\boldsymbol{\mu}}(\mathbf{X}, \mathbf{A})$, $\log \boldsymbol{\delta} = \text{GCN}_{\boldsymbol{\delta}}(\mathbf{X}, \mathbf{A})$ 。解码器通过使用嵌入的点积来预测节点之间的连接,其形式为:

$$p(\mathbf{A}|\mathbf{Y}) = \prod_{i=1}^N \prod_{j=1}^N p(A_{ij}|\mathbf{Y}_i, \mathbf{Y}_j) \quad (20)$$

$$p(A_{ij}=1|\mathbf{Y}_i, \mathbf{Y}_j) = \sigma(\mathbf{Y}_i^T \mathbf{Y}_j) \quad (21)$$

VGAE的目标是通过最小化重构损失和变分下界来优化模型,其损失函数定义为:

$$\ell = \mathbb{E}_{q(\mathbf{Y}|\mathbf{X}, \mathbf{A})} [\ln p(\mathbf{A}|\mathbf{Y})] - \text{KL}[q(\mathbf{Y}|\mathbf{X}, \mathbf{A}) \| p(\mathbf{Y})] \quad (22)$$

其中,KL[]是KL散度, $p(\mathbf{Y})$ 是高斯先验分布。

目前,基于VGAE的研究已经提出多种改进方法。例如,Guo等^[64]提出的GC-VGE将聚类引入VGAE中。该模型使用VGAE编码图的拓扑结构和节点特征映射,并构造自学习机制来细化嵌入表示和聚类中心,从而获得更有效的节点表示。考虑到GAE在结构受到干扰时表现较差,Zhang等^[65]提出了自适应图学习模型BAGE和VBAGE,这两种模型分别利用GAE和VGAE对图进行编码,并引入拉普拉斯图结构学习机制以自适应地生成和优化图结构,从而显著提高模型在图结构变化时的鲁棒性。

3.2.4 基于对抗学习的方法

基于对抗学习的方法通常利用生成对抗网络(Generative Adversarial Network, GAN)的对抗学习训练来生成节点嵌入。GAN^[66]是由Goodfellow等在2014年提出的一种生成模型,近些年在图像生成上表现出色。GAN通过生成器G和判别器D的对抗进行学习(见图4),经过不断迭代后,模型可以达到最优状态。生成器从潜在空间中随机采样作为输入,旨在学习真实数据的分布,从而生成与真实样本相似的伪样本;判别器的任务是区分真实样本和生成的伪样本。生成器和判别器通过对抗训练不断优化,最终实现对抗性博弈的平衡,其目标函数可表示为:

$$\min_{G_g} \max_{D_g} V(D, G) = \mathbb{E}_{x \sim P_{\text{data}}(x)} [\log D(x)] + \mathbb{E}_{z \sim P_z(z)} [\log(1 - D(G(z)))] \quad (23)$$

其中, $P_{\text{data}}(x)$ 和 $P_z(z)$ 分别是真实和生成的数据分布。

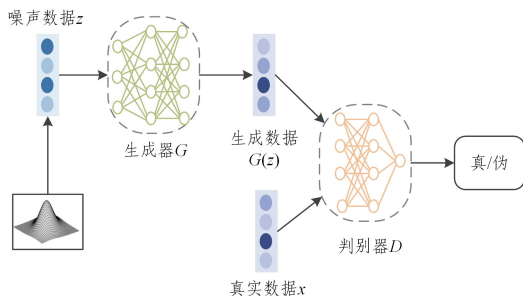


图4 GAN模型架构

Fig. 4 Architecture of GAN model

随着 GAN 的提出,越来越多的研究将其应用于图嵌入学习中。Wang 等^[67]提出的 GraphGAN 模型将图嵌入学习转换为生成器和判别器之间的对抗训练问题。在该模型中,对于每个节点 v_c ,生成器的目标是尽可能地拟合真实的连接分布 $p_{tr}(v|v_c)$,即生成与 v_c 最可能连接的节点来骗过判别器;判别器的任务是区分节点 v 是由生成器生成还是 v_c 的真实邻近节点,并计算 v 和 v_c 间存在边的可能性。其目标函数表示为:

$$\min_{\theta_G} \max_{\theta_D} V(D, G) = \sum_{c=1}^n (\mathbb{E}_{v \sim p_{tr}(\cdot | v_c)} [\log D(v, v_c; \theta_D)] + \mathbb{E}_{v \sim G(\cdot | v_c, \theta_G)} [\log(1 - D(v, v_c; \theta_D))]) \quad (24)$$

除了上述方法外,近年来也提出了一些通过不同策略提高性能的方法。例如, Hong 等^[68]提出的 GANE 模型同样关注于图的连通性问题。该模型的核心思想是通过生成器为节点对生成潜在的边,并通过判别器将生成的边与真实的边区分开,从而学习节点相应的表示;而且 GANE 采用了 Wasserstein-1 距离来优化生成器,以提高训练的稳定性。Yu 等^[15]提出的 NETRA 将 AE 与 GAN 结合。NETRA 使用长短期记忆网络 LSTM 作为编码器,并将生成对抗训练模块视为嵌入过程中的正则化项,以提升模型的泛化性能。为了进一步发挥 GAN 的优势, He 等^[69]提出了 ArmGAN。该方法通过利用编码器和竞争器来学习两种不同的表示,通过两者的相互竞争来生成更好的表示;而判别器对编码器和竞争器的表示机制进行区分。

表 2 对上述简单图嵌入方法的策略、优势和局限性进行了总结。

表 2 简单图嵌入方法
Table 2 Simple graph embedding method

类别	子类	小类	方法	策略	优势	局限性
浅层图嵌入方法	基于矩阵分解的方法	拉普拉斯矩阵映射方法	LE ^[26] , Meng 等 ^[27] , LRLGL ^[28]	利用拉普拉斯矩阵和节点对之间的相似性假设捕获图结构信息	基于流形假设和谱聚类理论,有强理论性,无需复杂调参	忽略全局结构信息,在大规模图上计算成本高
		节点邻近矩阵分解方法	GF ^[29] , GraRep ^[7] , NetSMF ^[30] , Agibetov ^[31] , Z-NetMF ^[32]	构造节点关系矩阵,通过矩阵因子分解学习节点低维嵌入	利用节点间的共现关系,可以捕获局部和高阶结构	高阶矩阵较为稠密,计算复杂度较高
	基于随机游走的方法	—	DeepWalk ^[9] , node2vec ^[10] , struc2vec ^[34] , Fairwalk ^[35] , NRL-RWR ^[36] , CNARW ^[37]	采用随机游走生成节点序列,使用 Skip-Gram 进行嵌入	通过多跳随机游走采样,有效保留节点之间的高阶关系	采样需要大量时间,限制了方法的可扩展性
深度图嵌入方法	基于卷积操作的方法	谱卷积方法	SpectralCNN ^[14] , ChebNet ^[41] , GCN ^[42] , SGC ^[43] , CensNet ^[44] , CW-GCN ^[45]	将节点特征转换到谱域上,进行拉普拉斯分解,以执行卷积操作	基于图的傅里叶变换,具有理论基础	需要计算图拉普拉斯分解,计算复杂度高,灵活性差
		空间卷积方法	PATCHY-SAN ^[46] , GraphSAGE ^[47] , GraphAIR ^[48] , EGC ^[49] , CEN-DGCNN ^[50]	直接在节点域上进行聚合,执行卷积操作	消息传递操作更容易扩展到大规模图上	多层卷积才能获得好的性能,易过平滑
	基于注意力机制的方法	层内注意力方法	GAT ^[51] , ARGAT ^[52] , SGATs ^[53] , HGRN ^[54]	引入注意力机制到单层网络的局部邻域上	根据邻居节点的特征动态计算注意力,自适应捕获邻居的重要性	仅在当前层进行注意力计算,无跨层融合
层间注意力方法	MGAT ^[55] , SGDMC ^[56] , AM-GCN ^[57]	注意力被视为特征融合的一种操作	跨层融合可以聚合不同层次的多样化信息	层间信息差异大时可能影响注意力作用		
基于对抗学习的方法	基于编码-解码架构的方法	深度自编码器方法	AE ^[58] , SDNE ^[59] , DNGR ^[13]	使用编码器和解码器将节点特征编码为潜在向量并重建节点特征	对节点特征进行压缩,可以有效用于下游任务	通常建模节点之间的局部关系,忽略了全局结构
		图自编码器方法	DGAE ^[61] , DLR-GAE ^[62] , VGAE ^[63] , GC-VGE ^[64] , BAGE ^[65] , VBAGE ^[65]	针对图类型设计,使用神经网络编码图结构信息	有效编码图信息,适应于多种图结构学习任务	参数更多,计算成本高
	—	GAN ^[66] , GraphGAN ^[67] , GANE ^[68] , NETRA ^[15] , ArmGAN ^[69]	在生成器和判别器的博弈训练过程中建模图的非线性结构	对抗机制可以提升嵌入的鲁棒性	通常存在训练不稳定的问题	

4 复杂图上的嵌入方法

传统图嵌入方法尽管在处理结构相对简单的图数据,即静态的、单一类型的、具有成对关系的简单图时表现出色,但在面对具有更加多样和复杂特性的图结构时,这些方法的局限性逐渐显现。近年来,随着实际应用场景的日益复杂,图结

构已经呈现出时间依赖性、语义异质性、多重关系或高阶关联等特性。这些复杂属性是简单图所不能表达的,也对如何建模提出了新的挑战。因此,研究人员提出了大量特定于复杂图结构的嵌入方法。这些方法的共同之处包括 3 个方面: 1) 核心目标的一致性,在学习向量表示时,需要保留图中复杂的拓扑结构信息和深层的语义信息; 2) 依赖多种消息传播机

制,复杂图嵌入方法通常依赖不同类型或维度的信息(如时间信息、类型信息)设计不同的消息传播机制;3)不同特征的交互,复杂图中往往涉及多种类型的信息之间的交互,例如多重图中节点对的多重关系的互补性、超图中超边和节点的交互,因此在建模时需要考虑多种交互信息。同时,与传统嵌入方法相比,复杂图嵌入方法显著改变了研究范式,即从静态转变到动态,从单一类型转变到多种类型关系,从单一关系转变到多重关系,从低阶交互转变到高阶关联。这4种不同的研究范式本质上导致了它们的不同之处:动态图嵌入方法主要关注节点和边随时间动态演变的过程,强调时序信息的建模;异质图嵌入方法关注于不同类型的节点与边的语义信息,通常需要捕获远程节点的信息;多重图嵌入方法旨在对同一节点对之间的多种关系进行协同学习,以探索它们的一致性和互补性;超图嵌入方法针对高阶关联关系,需学习一条超边与多个节点之间的多元对应关系。本章将重点探讨复杂图的深度嵌入方法,具体分析基于动态图、异质图、多重图和超图的图嵌入方法如何在复杂场景中发挥作用,并提供更具表现力的节点表示。

4.1 动态图嵌入方法

动态图是一种能够适应结构随时间变化的图类型。与静态的简单图不同,动态图的节点和边的状态以及图的拓扑结构会随着时间发生动态变化。如何捕获这种随时间演变的特性,已经成为图分析领域中的一个重要的研究方向。动态图嵌入方法旨在捕获图中节点在不同时间步长的潜在变化,以生成具有动态演化信息的嵌入表示。与静态图嵌入方法不同,这类方法通常需要引入时间窗口或时间序列信息,以有效地对图中的结构和时间依赖性进行建模。按照处理数据类型的不同,动态图嵌入方法可以分为两类:离散时间型方法和连续时间型方法。

4.1.1 离散动态图方法

离散型动态图通常将动态图按照时间序列表示为一系列的快照图,如图5所示。相邻的快照图之间的图结构存在一些差异,例如,某些节点或边被添加或删除。大多数离散型动态图方法通常随着时间的变化直接聚合所有的快照图,然后采用一些静态模型(如自编码器或GCN模型)对聚合后的图进行编码;或者使用静态模型对单个时间步的快照图进行编码,学习在每个时间戳下的节点表示,然后将这些节点表示进行聚合,以捕获离散的时序信息。由于这类方法的实现相对简单,它们在动态图嵌入中占据主流地位。

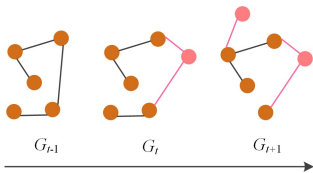


图5 离散型动态图

Fig. 5 Discrete dynamic graph

为了学习离散动态图的嵌入,Goyal等^[16]提出了DynGEM模型。该模型利用自编码器 AE^{t-1} 获得 $t-1$ 时刻的快照图的节点嵌入,同时利用 $t-1$ 时刻的自编码器的参数初始化 t 时刻的自编码器 AE^t ,使得 t 时刻的快照图嵌入能够在 $t-1$ 时刻的嵌入的基础上进行增量学习。由于动态图中的图

结构是不断变化的,该模型根据 $t-1$ 和 t 时刻的快照图的差异调整 t 时刻神经元的数量和隐藏层的层数。最后,该模型在训练过程中使用一阶相似度和二阶相似度来保留图中的局部和全局结构信息。该方法虽然具有较低的复杂度,能够高效地进行动态图嵌入,但是可能无法很好地应对动态图中发生的较大结构变化。

Pareja等^[70]提出的EvolveGCN模型能够在不依赖于节点嵌入的情况下沿着时间维度调整GCN,从而适应在极端情况下不同时间步长的节点集发生变化的情况。具体地,该模型在每个时间步长使用循环神经网络RNN对GCN的参数进行调整,以捕获图序列的动态性。根据GCN权重的演化方式,分为EGCU-H和EGCU-O两个版本。在每个时间步,输入最后一层状态 $\mathbf{H}_i^{(l)}$ 和来自上一时间步长的参数 $\mathbf{W}_i^{(l)}$,EGCU-H通过使用门控循环单元GRU更新当前时间步长的参数 $\mathbf{W}_i^{(l)}$,即 $\mathbf{W}_i^{(l)} = \text{GRU}(\mathbf{H}_i^{(l)}, \mathbf{W}_i^{(l)})$ 。相似地,EGCU-O通过LSTM计算当前时间步长下的GCN参数,此时只输入上一时间步长的参数,即 $\mathbf{W}_i^{(l)} = \text{LSTM}(\mathbf{W}_i^{(l)})$ 。最后,同时使用GNN更新节点嵌入,即 $\mathbf{H}_i^{(t+1)} = \text{GNN}(\mathbf{A}_i, \mathbf{H}_i^{(t)}, \mathbf{W}_i^{(t)})$ 。该方法的一个关键优势是,只训练RNN参数而不训练GCN参数,使得模型的参数数量不再随着时间而增加,更易于管理。

受EvolveGCN思想的启发,近年来有大量的离散动态图嵌入方法使用不同的模型或策略进行GNN参数的更新。例如,Gao等^[71]提出的DGCN模型通过最大化局部和全局图表示之间的互信息,并利用LSTM更新GCN的权重参数,分别捕获动态图中特定和全局时间步长的全局信息。Zhang等^[72]提出的DynGNN方法将LSTM嵌入GraphSAGE的每个隐藏层中,利用GraphSAGE来聚合邻居信息,并将每层的聚合特征输入循环单元,从而学习到细粒度的动态演化过程。Zhu等^[73]提出的结合GNN和元学习策略以及随机梯度聚合机制的WinGNN模型,将元学习策略融入GNN中来计算当前快照的逐帧损失,并将损失梯度传递到下一个快照。与此同时,引入随机滑动窗口来获取连续快照的梯度,最终通过两种梯度的聚合和更新,提高了模型的鲁棒性。

4.1.2 连续动态图方法

连续型动态图通过为每条边赋予多个时间戳来捕捉节点间的交互信息,如图6所示。当节点之间发生交互时,新的边将被创建并标注时间戳,以此来保留节点间的状态变化。因此,连续型动态图可以从更细粒度的角度捕获图中的时序信息。然而,传统的离散动态图方法不能处理这种类型的动态图。最近,一些连续型动态图嵌入方法被提出以处理这种场景,这些方法通常将动态图看作一个整体,重点关注边的时间戳信息,从而更精准地捕获图的动态演化过程。

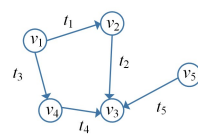


图6 连续型动态图

Fig. 6 Continuous dynamic graph

Trivedi等^[74]提出的DyRep模型通过捕获图的动态拓扑演变和图上的节点活动的信息生成节点表示。具体来说,DyRep以关联(拓扑演变)和通信(节点活动)事件的形式捕获

动态图中的信息,从自我传播、外源驱动和局部嵌入传播 3 个角度更新节点表示。在自我传播中,节点嵌入是相对于先前位置的变化,而不是随机方式进行更新;在外源驱动中,通过时间间隔更新节点特征;在局部嵌入传播中,两个节点形成能够从一个节点邻域传播到另一节点的临时或永久路径。对于节点 v 的第 p 个事件,节点表示更新为:

$$\mathbf{z}^v(t_p) = \sigma(\mathbf{W}^r \mathbf{z}^v(\bar{t}_p) + \mathbf{W}^l(t_p - \bar{t}_p) + \mathbf{W}^s \mathbf{h}_s^u(\bar{t}_p)) \quad (25)$$

其中, t_p 是当前事件的时间戳; \bar{t}_p 是先前事件的时间戳; $\mathbf{z}^v(\bar{t}_p)$ 是节点 v 从先前时间中获取的特征; \mathbf{h}_s^u 是从节点 u 的邻居聚合获得的表示; $\mathbf{W}^r, \mathbf{W}^l$ 和 \mathbf{W}^s 分别是自我传播、外源驱动和局部嵌入传播的参数。最后,利用时间注意力机制对结构和时间信息进行耦合,使得节点嵌入可以学习到非线性的动态信息。

另外,使用记忆模块存储节点状态的策略在动态图学习中也受到关注。例如, Rossi 等^[75]提出了一种高效的 TGN 模型,该模型由嵌入模块和记忆模块组成。嵌入模块利用时间图和节点的内存生成嵌入,对批量交互进行预测,并计算损失;记忆模块通过利用这些交互更新内存,从而能够捕获节点的长期依赖关系。Souza 等^[76]提出的 PINT 模型将 Weisfeiler-Leman 测试引入动态图学习中,利用内射时间消息传递来更新节点特征和相对位置特征,增强记忆状态,在新的时间层上建立模型,以提高 TGN 的表达能。力。

4.2 异质图嵌入方法

在社交网络等现实应用中,图数据通常由不同类型的节点和关系构成。然而,简单图无法捕获这一结构特性。异质图显示建模了多种类型的节点和边的信息,能够准确地表达复杂的语义关系。传统的图嵌入方法通常依赖于强同质性假设,即所有节点和边属于同一类型,无法有效地处理这种复杂结构。为了解决这一问题,研究者提出了大量针对异质图的嵌入方法。这些方法通常利用随机游走策略或者 GNN 来传播和整合不同类型的节点和边信息,从而捕获图中的异质结构和语义信息。因此,本节将此类方法划分为随机游走方法和深度网络方法。

4.2.1 随机游走方法

随机游走方法通常使用元路径随机游走策略以捕获不同类型的节点之间的依赖关系,从而使得生成的嵌入能够描述复杂的语义关联。经典的模型包括 metapath2vec^[11]和 HIN2Vec^[77]。

metapath2vec^[11]通过元路径随机游走构建异质图的节点邻域,并利用扩展的 Skip-Gram 模型对图中的异质节点进行建模,以最大程度地保留结构和语义信息。具体来说,给定一个元路径方案 $\mathcal{P}: V_1 \xrightarrow{R_1} V_2 \xrightarrow{R_2} \dots V_i \xrightarrow{R_i} V_{i+1} \dots \xrightarrow{R_{l-1}} V_l$, 其中 $R = R_1 \circ R_2 \circ \dots \circ R_l$ 表示节点类型 V_1 和 V_l 之间的复合关系,节点的下一步的转移概率为:

$$p(v_{i+1} | v_i, \mathcal{P}) = \begin{cases} 1/|N^{v_i}(v_i)|, & (v_{i+1}, v_i) \in E, \phi(v_{i+1}) = t+1 \\ 0, & (v_{i+1}, v_i) \in E, \phi(v_{i+1}) \neq t+1 \\ 0, & (v_{i+1}, v_i) \notin E \end{cases} \quad (26)$$

其中, $v_i \in V_i$, $N^{v_i}(v_i)$ 表示节点 v_i 的 V_i 类型的邻域节点集合。考虑到负采样时节点类型的不同,metapath2vec++^[11]在

该方法的基础上引入了异质负采样策略,进一步建模异质图中的结构和语义相关性。

Fu 等^[77]提出的 HIN2Vec 方法考虑了不同类型的节点以及它们之间多种类型的关系,从而有效捕获异质图中的丰富信息。该方法主要包括数据准备和表示学习两个阶段。具体来说,该方法首先结合随机游走和负采样,设计了训练数据采样策略以生成训练数据;然后设计了二进制分类器,通过预测节点对的关系概率,将节点之间的关系学习转换为二分类问题。

除了上述方法之外,考虑到不同关系类型间的平衡, Lee 等^[78]提出了 BHIN2vec,将异质图嵌入转换为多任务学习问题来进行处理。该模型通过调整不同任务的训练比例来平衡不同关系对节点嵌入的影响,从而提高模型的表达能力。为了不依赖于元路径, Liu 等^[79]提出的 HeteEdgeWalk 设计了一种动态调整的双向边采样游走策略,使用边采样和最近选择的边类型来决定如何采样,能够有效地保留异质结构。

4.2.2 深度网络方法

此外,许多深度网络方法也被提出,用于对异质图中的不同类型的节点和边进行建模。例如, Wang 等^[80]提出了专门为异质图嵌入设计的 HAN 模型,该模型结合了节点级和语义级的注意力机制,旨在从不同层次上学习图中信息。在节点级注意力方面, HAN 通过学习节点和基于元路径的邻居之间的重要性来学习节点嵌入,具体如下:

$$\alpha_{ij}^p = \frac{\exp(\sigma(\alpha_p^T [\mathbf{h}_i' \| \mathbf{h}_j']))}{\sum_{k \in N_i^p} \exp(\sigma(\alpha_p^T [\mathbf{h}_i' \| \mathbf{h}_k']))} \quad (27)$$

$$\mathbf{z}_i^p = \sum_{k=1}^K \sigma(\sum_{j \in N_i^p} \alpha_{ij}^p \cdot \mathbf{h}_j) \quad (28)$$

其中, N_i^p 是通过元路径 \mathcal{P} 与 v_i 相连的节点邻居, \mathbf{h}_i' 是映射后的节点特征, α_p 是与 \mathcal{P} 相关的注意力, \mathbf{z}_i^p 是 v_i 基于 \mathcal{P} 的节点特征。给定元路径集合 $\{\mathcal{P}_1, \mathcal{P}_2, \dots, \mathcal{P}_{|\mathcal{P}|}\}$, 得到 $|\mathcal{P}|$ 组语义特定的节点特征 $\{\mathbf{Z}_{\mathcal{P}_1}, \mathbf{Z}_{\mathcal{P}_2}, \dots, \mathbf{Z}_{\mathcal{P}_{|\mathcal{P}|}}\}$ 。在语义级注意力上, HAN 通过学习不同元路径之间的重要性来获得最终的节点嵌入:

$$\beta_p = \frac{\exp\left(\frac{1}{|\mathcal{V}|} \sum_{i \in \mathcal{V}} \mathbf{q}^T \cdot \tanh(\mathbf{W} \cdot \mathbf{z}_i^p + \mathbf{b})\right)}{\sum_{\rho=1}^{|\mathcal{P}|} \exp\left(\frac{1}{|\mathcal{V}|} \sum_{i \in \mathcal{V}} \mathbf{q}^T \cdot \tanh(\mathbf{W} \cdot \mathbf{z}_i^\rho + \mathbf{b})\right)} \quad (29)$$

$$\mathbf{Z} = \sum_{\rho=1}^{|\mathcal{P}|} \beta_\rho \cdot \mathbf{Z}_\rho \quad (30)$$

其中, $|\mathcal{P}|$ 是元路径种类数, \mathbf{q} 表示语义注意力向量。通过节点级和语义级层次的注意力, HAN 能够以分层的方式聚合基于元路径的邻居特征,以生成丰富的节点嵌入,从而更好地捕获异质图中的语义和结构信息。

然而, HAN 只能学习到元路径两端的节点信息,忽略了中间节点信息。为此, Fu 等^[81]提出 MAGNN 模型,通过设计元路径内和元路径间聚合,来学习元路径中间的语义节点和来自多个元路径的信息。具体来说,该模型首先将不同类型的节点投影到同一潜在空间中;接着对每个元路径进行元路径内聚合,通过实例编码器将元路径上的所有节点的特征转换为单个向量;最后,与 HAN 类似,该模型使用节点级注意力计算节点重要性以及节点特征,同时设计元路径外部注意力机制聚合元路径间的信息。

HAN 和 MAGNN 的成功,激发了人们设计多种 GNN 模型进行异质图嵌入学习,催生了一系列方法。例如,He 等^[82]提出的 STHGE 模型引入了 HSIC 准则来调整元路径特征空间之间的独立性,并使用 GAT 学习不同元路径的节点嵌入,提高模型对节点关系的表达能力。Yang 等^[83]提出了一个高效的 SeHGNN 模型,该模型使用 GCN 预先计算邻居聚合,避免了每个训练时期的重复邻居聚合,并采用具有长元路径的单层结构来扩展感受野,以充分利用语义信息。

4.3 多重图嵌入方法

现实世界中的实体之间的交互通常具有多样性。例如,在视频共享网站中,用户可以通过订阅或评论与其他用户进行互动。而简单图只能描述节点之间的单一关系,因此,多重图被提出以构建这种多重关系。然而,多重图中关系的多样性也为节点嵌入学习带来了一些困难。尤其是,不同类型的关系可能具有完全不同的特征,导致难以提取它们的一致性信息并保留差异性。近年来,研究者已经提出了多种多重图嵌入方法,旨在揭示多重图中隐藏的复杂信息。这些方法主要集中于学习所有类型(维度)上的通用的节点表示,以及针对每种类型的特定的节点表示。

较早关注多重图嵌入问题的是 Liu 等^[84],他们提出了 3 种模型,即网络聚合、结果聚合和层协同分析。具体来说,如图 7 所示,网络聚合模型将多重图整合为一个简单图,然后应用 node2vec 方法进行图嵌入;结果聚合模型认为不同类型的关系具有完全不同的边,对每种类型的图分别应用 node2vec,然后聚合生成的多种节点嵌入来获得最终的节点嵌入;层协同分析模型进一步考虑了不同关系类型的图之间的交互,通过使用随机游走策略在多个类型关系的图之间进行遍历,从而学习到能够捕获不同类型的图间交互的节点表示。

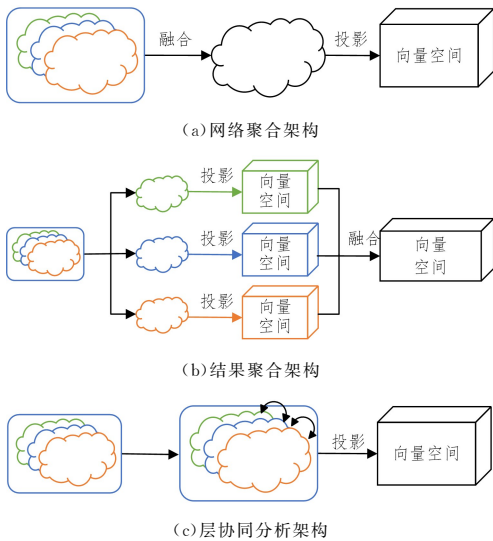


图 7 多重图的 3 种框架^[84]

Fig. 7 Three types of frameworks for multiplex graph^[84]

Park 等^[18]提出一种多重图嵌入方法 DMGI,通过衡量节点和图表示之间的互信息学习节点嵌入。具体来说,该方法首先使用 GCN 编码每种类型的图来获得关系类型特定的节点表示 $\mathbf{H}^r = \text{GCN}(\mathbf{X}, \mathbf{A}^r)$,并进行池化来获得图级表示 $\mathbf{S}^r = \text{Readout}(\mathbf{H}^r)$ 。同时,该方法定义了关系类型特

定的损失函数:

$$L^r = \sum_{i=1}^n \log D(\mathbf{h}_i^r, \mathbf{s}^r) + \sum_{j=1}^n \log(1 - D(\tilde{\mathbf{h}}_j^r, \mathbf{s}^r)) \quad (31)$$

其中, $\mathbf{h}_i^r \in \mathbf{H}^r$ 是类型 r 的 v_i 的嵌入, $\tilde{\mathbf{h}}_j^r$ 是负样本节点嵌入, $D(\mathbf{h}_i^r, \mathbf{s}^r)$ 是评分函数。为了集成不同类型的嵌入,该方法引入一致性矩阵 \mathbf{Z} ,通过最小化 \mathbf{H}^r 和 \mathbf{Z} 之间的差距并最大化 $\tilde{\mathbf{H}}^r$ 和 \mathbf{Z} 之间的差距,得到正则化项:

$$L_{cs} = \left[\mathbf{Z} - \frac{1}{|R|} \sum_{r \in R} \mathbf{H}^r \right]^2 - \left[\mathbf{Z} - \frac{1}{|R|} \sum_{r \in R} \tilde{\mathbf{H}}^r \right]^2 \quad (32)$$

其中, R 是所有的关系类型。最终的目标函数为 $J = \sum_{r \in R} L^r + \alpha L_{cs} + \beta \|\mathbf{W}\|^2$ 。

上述方法的研究思想对后续工作产生了深远影响,近期出现了多种多重图嵌入方法。例如, Jing 等^[85]提出的 HDMI 方法通过最大化节点特征、节点表示和图表示之间的高阶互信息,来同时捕获节点嵌入与全局嵌入和节点属性之间的相互依赖性,并使用注意力机制融合多个关系类型的节点嵌入。Mo 等^[86]提出的 MGDCR 方法关注了图内和图间的去相关性损失,通过减小每个图中噪声的影响,有效捕获不同类型的图之间的公共信息。为了进一步提高公共信息的质量,Mo 等^[87]提出了 DMG 模型来研究多重图的解耦表示学习。DMG 使用 GCN 和 MLP 编码器为每个图生成公共表示和私有表示,并提出匹配损失和对比损失来获得干净的公共信息和私有信息。

4.4 超图嵌入方法

简单图中的每条边仅能连接两个节点对,以表达成对的依赖关系,无法建模复杂的协同关系。而超图对简单图的结构进行了扩展,它允许一条超边连接两个以上的节点,从而可以捕获节点间的高阶关联。正是由于这些特性,超图的嵌入学习更具有挑战性。因此,超图嵌入成为近年来的研究热点,为解决非传统图结构问题提供了新的思路。超图嵌入方法旨在将超图中的节点映射到低维空间中,同时尽可能地保留节点之间的高阶关联关系,从而挖掘超图的深层信息。与以成对关系为中心的简单图嵌入方法相比,超图嵌入方法更关注于超边形成的结构,能够更好地保留多元关系的整体信息。本节按照是否将超图展开为传统图结构处理或是直接在超图上进行学习,将这些方法分为展开式和非展开式方法。

4.4.1 展开式方法

展开式方法将超图按照某种特定模式转换为简单图结构,然后将简单图嵌入方法迁移到超图上进行学习。其中, HGNN^[19]是一种被广泛应用的方法。受 GCN 中卷积运算的启发,该方法对超图卷积运算重新进行定义,使得在超图结构中编码高阶的数据相关性。其中,基于超图拉普拉斯算子的超图卷积层为:

$$\mathbf{H}^{(l+1)} = \sigma(\mathbf{D}_v^{-1/2} \mathbf{M} \mathbf{W} \mathbf{D}_v^{-1} \mathbf{M}^T \mathbf{D}_v^{-1/2} \mathbf{H}^{(l)} \Theta^{(l)}) \quad (33)$$

其中, $\mathbf{H}^{(l)}$ 是第 l 层的超图的节点特征矩阵, \mathbf{D}_e 和 \mathbf{D}_v 分别表示超边和顶点度的对角矩阵, Θ 是训练参数。

HGNN 是一个将 GNN 迁移到超图学习的通用框架,传统的 GCN 可以看作 HGNN 的特例。受 HGNN 的启发,学者提出了更多面向超图的嵌入方法。例如, Ji 等^[88]提出的

DHCF 模型利用超图结构对用户和项目信息进行建模,并利用一个跳跃超图卷积模型提取超图结构中的高阶关联信息。Yang 等^[89]在超图学习中引入了一个线扩展策略,该策略能够在保留原始高阶信息的同时,将超图映射为一个简单图,从而使传统 GCN 模型可以应用于超图嵌入学习。

4.4.2 非展开式方法

虽然展开式超图方法受到了广泛应用,但在将超图转换的过程中可能会丢失部分语义信息,从而未能充分利用超图中的丰富信息。非展开式超图方法直接在超图上进行建模,从而能够更好地保证超图信息的完整性。Arya 等^[90]受空间卷积方法通过图结构进行信息传递思想的启发,设计了一个归纳式模型 HyperSAGE。该模型使用两级聚合策略在超图上进行信息传播。具体来说,首先对同一超边中的节点特征进行聚合,生成超边特征表示,然后对包含中心节点的超边特征进行聚合,过程分别表示如下:

$$\mathbf{x}_{i,t}^{(e)} \leftarrow \mathcal{F}_1(\{\mathbf{x}_{j,t-1} \mid v_j \in N(v_i, e; \alpha)\}) \quad (34)$$

$$\mathbf{x}_{i,t} \leftarrow \mathbf{x}_{i,t-1} + \mathcal{F}_2(\{\mathbf{x}_{i,t}^{(e)} \mid v_i \in E(v_i)\}) \quad (35)$$

其中, $N(v_i, e; \alpha)$ 表示超边 e 内采样的 α 个节点, $E(v_i)$ 表示包

含中心节点 v_i 的超边集合, \mathcal{F}_1 和 \mathcal{F}_2 分别表示节点特征和超边特征聚合函数。 \mathcal{F}_1 和 \mathcal{F}_2 的具体定义如下:

$$\mathcal{F}_1(v_i) = \left(\frac{1}{|N(v_i, e)|} \sum_{v_j \in N(v_i, e)} \mathbf{x}_j^e \right)^{\frac{1}{p}} \quad (36)$$

$$\mathcal{F}_2(v_i) = \left(\frac{1}{|E(v_i)|} \sum_{e \in E(v_i)} \frac{N(v_i, e)}{N(v_i)} (\mathcal{F}_1(v_i))^e \right)^{\frac{1}{p}} \quad (37)$$

基于上述方法的思想,为了进一步衡量不同特征在聚合时的重要性,Hyper-SAGNN^[91]方法在超图建模中引入了自注意力机制。该模型首先生成与超边连接无关的静态嵌入,同时使用多头注意力层对超图建模来获得与超边连接有关的动态嵌入。通过结合这两种嵌入,该方法能够保留更多的信息。同样,Bai 等^[92]通过注意力机制衡量不同节点在聚合时的重要程度,从而生成更具有判别力的嵌入表示。此外,Wu 等^[93]提出了一个双视图超图模型 DHGNN,其根据拓扑结构和节点属性构造对偶超图,通过超图卷积层和注意力机制建模和融合对偶超图的节点嵌入。

表 3 对以上复杂图嵌入方法的策略、优势和局限性进行了总结。

表 3 复杂图嵌入方法

Table 3 Complex graph embedding method

类别	子类	方法	策略	优势	局限性
动态图嵌入方法	离散动态图方法	DynGEM ^[16] , EvolveGCN ^[70] , DGCN ^[71] , DynGNN ^[72] , WinG-NN ^[73]	将动态图视为一系列的快照图,使用静态模型编码	使用静态图模型,实现较为简单,效率更高	忽略了快照之间的细节,不能很好地学习演变过程
	连续动态图方法	DyRep ^[74] , TGN ^[75] , PINT ^[76]	将动态图看作一个整体,直接对边的时间戳信息建模	精确捕获任意时间的信息,保持时间连续性	实时更新的计算和空间复杂度较高
异质图嵌入方法	随机游走方法	metapath2vec ^[11] , metapath2vec++ ^[11] , HIN2Vec ^[77] , BHIN2vec ^[78] , HeteEdgeWalk ^[79]	利用不同的元路径随机游走策略来生成具有不同语义的节点序列	元路径反映了节点间的语义关系,可解释性强	较为依赖于设定的元路径
	深度网络方法	HAN ^[80] , MAGNN ^[81] , STHGE ^[82] , SeHGNN ^[83]	利用深度学习框架自动学习不同类型的节点和边之间的关系	利用深度学习模型可以自动学习节点关系,容易扩展	深度学习模型的可解释性较差
多重图嵌入方法	—	Liu 等 ^[84] , DMG ^[18] , HDM ^[85] , MGDCR ^[86] , DMG ^[87]	学习节点的通用或类型特定的节点表示,以捕获图中不同的关系	可以捕获多个图之间的一致性和互补性	若多个结构之间差异较大,需要设计复杂的学习策略
超图嵌入方法	展开式方法	HGNN ^[19] , DHCF ^[88] , Yang 等 ^[89]	将超图转换为简单图,使用简单图的嵌入模型建模	可以利用已有的简单图模型,易于实现	将超图转为简单图会损失一部分语义信息
	非展开式方法	HyperSAGE ^[90] , Hyper-SAGNN ^[91] , Bai 等 ^[92] , DHGNN ^[93]	扩展各种深度模型,以适应超图结构	有效保留高阶关联信息	实现较难,复杂度较高

5 图嵌入的应用场景和未来研究方向

图嵌入通过将高维的图数据映射到低维的向量空间中,极大地提升了对图数据的处理效率,因此,图嵌入学习已经在机器学习中占据了重要地位,并在多个应用场景中带来了显著的改进。例如,通过将图编码为低维向量后,图嵌入方法能够有效应用于社交网络分析、推荐系统和生物信息学等领域的各种问题。同时,随着数据规模、模型复杂度和应用需求的不断增长,图嵌入技术也面临一系列新的挑战,需要进一步完善和创新。因此,本章简要总结了图嵌入学习的主要应用场景,并探讨了未来的研究方向。

5.1 应用场景

1) 社交网络分析

随着在线社交媒体网站和移动应用程序的迅速发展,从

中收集到的用户生成内容也迅速增加,这使得社交网络的结构愈发复杂,因此产生了多种具有挑战性的任务^[94]。例如,社交网络中存在显著的结构特征,包括社区结构和中心性等;在用户行为分析中需要模型具有强可解释性,以个性化操作等。在面对这些挑战时,图嵌入技术为社交网络数据的学习和分析提供了有效的解决方式。针对社交网络的结构建模问题,Wang 等^[95]提出了一种基于边的图嵌入方法 edge2vec。该方法利用深度自编码器和随机游走模型来保留全局和局部的结构信息,有效保留了社交网络结构中的丰富信息。为了提高用户行为建模的可解释性,一些方法通过识别有影响力的节点来发现用户行为背后的特性。Zhang 等^[96]受残差网络和常微分方程的启发,设计了一种具有二阶连续的 GNN 模型,旨在识别社交网络中有影响力的邻居节点,以提高对用户行为的理解。类似地,Kumar 等^[97]提出了 SGNN 模型,旨

在识别有影响力的用户,以在网络上最大化特定信息。该模型将问题转换为伪回归问题,利用 `struc2vec` 进行节点嵌入并送入基于 GNN 的回归器中,从而预测网络中节点的影响。

2) 推荐系统

推荐系统作为一种个性化的信息过滤工具,能够帮助用户在海量数据中迅速定位目标信息。为了表示用户和项目之间的交互信息,通常会构造一个用户-项目关联图。而异质图能够自然地描述这些信息^[98]。通过编码异质图中的用户与项目的交互信息,可以更好地理解用户的兴趣,从而生成精准的个性化推荐。然而,在推荐系统中,用户与项目之间的直接交互较少,单一的局部关系无法准确反映用户偏好。为此,一些研究结合多跳的上下文信息来捕获更多的关联信息。Sang 等^[99]提出的基于元路径的异构图对比学习模型 IHGCL 构建了一个双对比机制,将意图整合到推荐中,以构建意图-意图对比和意图-交互对比,能够捕获细粒度的关系信息。Yu 等^[100]提出了利用高阶用户关系来增强推荐系统的多通道超图卷积模型 MHCN,该模型为异质图上的每个通道都构建了一个超图来描述用户的高阶关系,并通过学习每个超图的嵌入来获得全面的用户表示。Zhang 等^[101]提出的 HKGAT 模型利用异构知识图谱来捕获多跳的实体和关系的交互信息,以结合上下文信息产生准确的用户偏好。

3) 生物信息学

在生物医学大数据时代,随着基因数据的迅速增加,研究者需要从基因数据中挖掘出有价值的信息,从而推动生物学的发展。生物信息学整合了数学、计算机科学和生命科学等学科的方法,已经成为研究和分析生物学问题的重要工具。然而,生物数据的结构通常较为复杂,维度较高,而且实际实验的困难使得数据中交互关系往往较为稀疏。图嵌入技术能够有效地捕获生物分子网络的结构信息,这对于加快药物的研发进程有重要意义。针对药物-靶标交互的稀疏性,Li 等^[102]利用多种生物医学数据构建了一个关于药物-靶标的知识图谱,通过对三元组进行编码,能够捕获潜在的语义信息,有助于识别新开发的潜在药物。为了捕获蛋白质之间的交互信息,Baranwal 等^[103]提出了一种基于蛋白质 3D 结构的方法 Struct2Graph,该方法用于从折叠蛋白球的结构数据中识别蛋白质-蛋白质相互作用,从而提高对蛋白质功能的理解。Su 等^[104]提出一个异质分子网络 MTV-PPI,通过同时学习视图间的蛋白质序列和视图内分子间的相互作用来提取蛋白质表示,有助于理解相关疾病的形成机制。

4) 计算机视觉

由于图嵌入学习在特征提取方面具备优势,它已经被广泛应用于计算机视觉的多个实际任务,如图像分类和动作识别。然而,这些任务也面临诸多挑战,例如,图像分类中的样本标注不足会导致小样本分类问题,动作识别中要如何同时建模空间结构和时间依赖关系。针对小样本学习,Garcia 等^[105]通过利用输入图像的上下文信息来构造一个图结构,图中的节点包括有标签和无标签的样本,然后使用 GNN 的消息传递机制捕获图中的拓扑结构以学习语义信息。在动作识别中,人体关节通过骨骼连接,它们可以自然地表示为一个

图,提取这些图的表示有助于识别人物的动作。Xie 等^[106]提出一种空间-时间图卷积网络 DS-STGCN,将人体动作数据表示为图,并利用多层时空图卷积联合学习空间和时间维度的动作,以保留时空特性。Ahmad 等^[107]对身体的远距离关节之间的协作关系进行编码,并对长时间依赖关系进行建模,以更好地捕获人体动作中的潜在信息。

5) 自然语言处理

在自然语言处理领域,句子中的词语之间通常具有高度的上下文依赖性,而序列数据无法表达这种结构依赖性。为克服这一问题,图嵌入方法通过将文本序列中的句子结构表示为各种图结构,可以更自然地捕获词与词之间的关联关系。例如,Yao 等^[108]根据单词共现和文档单词关系构建了一个文本图,并使用 GCN 编码文本图来进行文本分类任务。Ragesh 等^[109]提出一种异质图卷积方法 HETEGCN,该方法旨在使用不同的跨层的图进行非线性变换来联合学习特征嵌入和文档嵌入。Yang 等^[110]提出一种异质图嵌入方法,该方法利用节点级和类型级双重注意力机制的异质 GAT 模型来对短文本进行编码,可以解决短文本中的语义稀疏问题。针对分层多标签文本分类任务,Ma 等^[111]提出一种混合嵌入方法,该方法通过图嵌入模型对层次结构中的全局和局部结构信息进行编码,并结合词嵌入模型对层次结构中的每个类标签的单词进行编码,这种方式有助于提高文本分类的准确性。

6) 知识图谱补全

知识图谱补全是异质图嵌入的一个重要应用。知识图谱本质上是一种特殊的异质图,通常以三元组(即实体-关系-实体)表示。通过利用各种异质图嵌入方法,可以将知识图谱中的实体和关系映射到低维的向量空间中,以帮助预测缺失的实体或关系信息。然而,图嵌入方法在知识图谱补全中依旧面临多关系建模、跨语义对齐困难、模态异质性等挑战。为了解决这些问题,Schlichtkrull 等^[112]引入一种关系图卷积网络,利用知识图谱的不同关系类型之间的消息传播,建模多种关系对实体的作用。Bi 等^[113]提出 Triple2Seq 模型解决节点和边的语义异质性问题,旨在将上下文子图作为输入序列进行采样,以保留相关信息。Liang 等^[114]提出的超节点关系图注意力模型,引入视觉模态、数值特征等信息,协同学习多模态知识图谱中的结构信息和多模态信息,以更好地进行知识图谱补全。

5.2 未来研究方向

1) 可扩展性

尽管现有的图嵌入工作在处理小规模图数据上取得了显著的效果,但是随着社交网络和生物医学等领域的快速发展,图数据的规模不断增加,现有的方法仍面临诸多挑战。例如,大规模数据导致巨大的存储和交互开销,模型训练效率低下。目前已经有一些图嵌入方法通过分块、粗化和分布式计算来处理大规模图数据。例如,Chiang 等^[115]提出的利用图聚类和 GCN 的方法 Cluster-GCN,利用图聚类算法识别密集子图,在每一步对与子图相关的节点块进行采样,并限制邻域搜索范围,从而能够将训练图的规模扩展到百万节点级别。Liang 等^[116]提出的 MILE 方法通过将图反复粗化为更小的

图,在最粗糙的图上应用现有的图嵌入方法,可以将图嵌入扩展到大型图上。Ji等^[117]通过批处理方式对异质图上的节点邻域进行采样,提升了在异质图嵌入上的学习效率。这些方法尽管能够在一定程度上提升模型效率,但在处理超大规模数据时依旧难以解决扩展性问题。随着图规模的扩大和神经网络层数的增加,模型训练的时间将迅速增长,参数量也将大幅增加。因此,设计具有扩展性的图嵌入模型以应对现实生活中不断增长的大规模数据需求,仍然是当前图嵌入领域面临的一个重要挑战。

2) 可解释性

最近的图嵌入模型主要是基于深度神经网络模型而设计的,已经在处理和图结构数据方面取得了显著的成果。然而,这些方法普遍具有可解释性不足的问题,尤其在一些实际的应用场景中。例如,在药物筛选中,图嵌入方法能够有效学习分子表示,以筛选出与靶标更具有相互作用的药物,但如何揭示模型在学习过程中捕获的生物信息仍不清楚。目前,有少数工作对可解释性问题进行了研究。例如,Feng等^[118]提出的KerGNNs方法将图核集成到GNN的消息传递中,采用可训练的隐藏图作为图滤波器,并结合子图更新节点嵌入,从而提高了模型的可解释性。Yang等^[119]设计了一种具有可解释性且高效的异质图卷积网络ie-HGCN,该模型利用对象级和类型级的分层聚合架构,自动提取有用的元路径,增强了模型的可解释性。Huang等^[120]提出的GraphLIME方法从每个节点的 N 跳邻居生成一个非线性可解释模型,并使用HSIC准则选择 K 个最具代表性的特征作为预测解释,以增强模型的可解释性。尽管这些研究为提高图嵌入模型的可解释性提供了初步的探索,但目前在这一领域的工作较为有限,仍旧存在泛化性不足、不能与实际应用对齐等挑战。因此,如何在保证模型性能的同时增强其可解释性是一个非常有意义的研究方向。

3) 鲁棒性

在真实世界的应用中,图数据常常受到噪声的影响。这些噪声可能通过修改图拓扑或节点属性来干扰图嵌入的学习过程,进而影响下游任务的表现。因此,提升图嵌入在噪声干扰下的鲁棒性是一个重要的方向。为了应对该挑战,一些工作致力于从异常检测出发来研究图嵌入在噪声环境的鲁棒性。例如,Kiouche等^[121]在图聚类的基础上,提出基于子结构和图编辑距离的图嵌入方法来检测异质图中的异常图,以提高模型的鲁棒性。Duan等^[122]将表示学习和异常检测整合到一个框架中,设计异常链接选择器识别异常的边,并将节点嵌入解耦合为正常嵌入和异常嵌入,能够有效识别图中的异常信息。Han等^[123]提出的融合稀疏自编码器和图网的FuSAGNet方法,通过联合优化重建和预测任务来建模多变量时间序列中的关系并进行异常检测。这些研究尽管在一定程度上提高了图嵌入的鲁棒性,但在应对多样化的噪声和更复杂的扰动时,仍存在一定的问题。因此,如何设计更加鲁棒性的图嵌入学习模型依旧是一个充满挑战的研究方向。

4) 可迁移性

自然语言处理和生物信息学等领域的研究表明,预训练

嵌入模型的迁移性学习有助于提高模型在新任务中的性能。近年来,图嵌入模型也在借鉴这种理念,以提升模型在不同任务中的表现。然而,与规则的图像或序列数据相比,图数据的结构是不规则的,这为迁移学习带来了困难。为此,Xie等^[124]提出了结合大语言模型和GNN的框架GALM,该框架通过在大型图语料库上进行图感知语言模型预训练,能够在不同的下游任务中进行有效的迁移和微调。Li等^[125]提出一个分子预训练的图学习方法MPG,该方法利用自监督策略在节点级和图级预训练模型,能够从大规模未标记分子图中学习分子表示,从而移植到药物挖掘的各种任务上。Rong等^[126]提出了一个基于Transformer的模型GROVER,该模型通过在节点级、边级和图级设计自监督任务,可以从大量未标记的分子图中学习分子的丰富结构和语义信息。尽管这些方法展示了迁移学习的潜力,但避免语义偏移以及过拟合的问题是较为困难的。因此,设计具有可迁移性的学习模型是一个充满潜力的研究方向。

5) 组合复杂性

现阶段的图嵌入学习的主要关注点依然集中在简单图上,尽管一些嵌入模型已经面向异质图、动态图、多重图和超图进行了研究,但现实世界中的图往往更加复杂,它们可能同时具备异质性、时序性、多重关系或高度关联等多种特性。这种更复杂的图结构能够全面地表达复杂的关系,但也导致了结构更为复杂,语义难对齐,为图嵌入的建模带来了困难。目前有一些工作关注于这个问题。Sun等^[127]提出了一种异质超图嵌入方法,该方法将异质超图投影到一系列快照中,并利用局部超图卷积运算生成嵌入表示,从而探索异质超图的多重性质。Liu^[128]提出Meta-HGT进行异质超图嵌入学习,该方法通过扩展元路径来提取高阶关系,并构造具有不同复合语义的超图,从而增强模型对复杂关系的捕获能力。Fu等^[129]提出一种多重异质图嵌入方法MHGCN+,其可以将拓扑结构、节点特征和多种关系映射到同一潜在空间。随着图嵌入学习研究的不断深入,设计更灵活的嵌入模型来探索复杂图的结构、语义、时序等多类型的性质,仍然是一个具有前景的研究方向。

结束语 图嵌入作为一种有前景的图特征提取技术,通过将非线性的图结构数据转换为低维且稠密的向量表示,能够有效地保留节点之间的邻接关系。本文围绕着图嵌入学习方法的研究现状,讨论和分析了在简单图和复杂图上的图嵌入方法的研究进展。对于简单图嵌入方法,根据模型嵌入方式的不同,将其分为浅层和深度图嵌入方法。浅层图嵌入方法主要关注于矩阵分解和随机游走等传统手段,通常具有良好的可解释性。深度嵌入方法则借助深度学习技术,能够捕捉更复杂的非线性特征。在复杂图嵌入方法中,本文讨论了在动态图、异质图、多重图和超图上的嵌入方法。这些方法通过使用不同的图嵌入技术,在生成节点表示的同时,能够有效保留复杂图中的时序性、异质性、多重关系或高度关联等信息。最后,总结了图嵌入学习的实际应用场景,并展望了未来有潜力的研究方向。本文从简单图到复杂图,系统地梳理了图嵌入学习的研究进展,旨在为未来的发展方向提供有益的

参考。未来,随着计算能力的不断提升和算法的持续优化,图嵌入技术将在各个领域发挥更加重要的作用。

参 考 文 献

- [1] SHARMA K, LEE Y C, NAMBI S, et al. A survey of graph neural networks for social recommender systems[J]. *ACM Computing Surveys*, 2024, 56(10): 1-34.
- [2] CUI G Q, ZHOU J, YANG C, et al. Adaptive graph encoder for attributed graph embedding[C]// *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. New York: ACM, 2020: 976-985.
- [3] ZHOU B X, ZHENG L R, WU B H, et al. Protein engineering with lightweight graph denoising neural networks[J]. *Journal of Chemical Information and Modeling*, 2024, 64(9): 3650-3661.
- [4] FU N, NI W W, HOU L H, et al. Community detection in decentralized social networks with local differential privacy[J]. *Information Sciences*, 2024, 661: 120164.
- [5] ZHANG Q J, XU Y D. Knowledge graph embedding with inverse function representation for link prediction[J]. *Engineering Applications of Artificial Intelligence*, 2024, 127: 107225.
- [6] LI S, ZAIDI N A, DU M J, et al. Property graph representation learning for node classification[J]. *Knowledge and Information Systems*, 2024, 66(1): 237-265.
- [7] CAO S S, LU W, XU Q K. GraRep: Learning graph representations with global structural information[C]// *Proceedings of the 24th ACM International Conference on Information and Knowledge Management*. New York: ACM, 2015: 891-900.
- [8] OU M D, CUI P, PEI J, et al. Asymmetric transitivity preserving graph embedding[C]// *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. New York: ACM, 2016: 1105-1114.
- [9] 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 & Data Mining*. New York: ACM, 2014: 701-710.
- [10] GROVER A, LESKOVEC J. node2vec: Scalable feature learning for networks[C]// *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. New York: ACM, 2016: 855-864.
- [11] DONG Y, CHAWLA N V, SWAMI A. metapath2vec: Scalable representation learning for heterogeneous networks[C]// *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. New York: ACM, 2017: 135-144.
- [12] MIKOLOV T. Efficient estimation of word representations in vector space[J]. *arXiv:1301.3781*, 2013.
- [13] CAO S S, LU W, XU Q K. Deep neural networks for learning graph representations[C]// *Proceedings of the 30th AAAI Conference on Artificial Intelligence*. Palo Alto, CA: AAAI, 2016: 1145-1152.
- [14] BRUNA J, ZAREMBA W, SZLAM A, et al. Spectral networks and locally connected networks on graphs[J]. *arXiv:1312.6203*, 2013.
- [15] YU W C, ZHENG C, CHENG W, et al. Learning deep network representations with adversarially regularized autoencoders [C]// *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. New York: ACM, 2018: 2663-2671.
- [16] GOYAL P, KAMRA N, HE X, et al. DynGEM: Deep embedding method for dynamic graphs[J]. *arXiv:1805.11273*, 2018.
- [17] ZHANG C X, SONG D J, HUANG C, et al. Heterogeneous graph neural network [C]// *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. New York: ACM, 2019: 793-803.
- [18] PARK C, KIM D, HAN J, et al. Unsupervised attributed multiplex network embedding [C]// *Proceedings of the 34th AAAI Conference on Artificial Intelligence*. Palo Alto, CA: AAAI, 2020: 5371-5378.
- [19] FENG Y F, YOU H X, ZHANG Z Z, et al. Hypergraph neural networks [C]// *Proceedings of the 33rd AAAI Conference on Artificial Intelligence*. Palo Alto, CA: AAAI, 2019: 3558-3565.
- [20] XU M J. Understanding graph embedding methods and their applications[J]. *SIAM Review*, 2021, 63(4): 825-853.
- [21] GOYAL P, FERRARA E. Graph embedding techniques, applications, and performance: A survey [J]. *Knowledge-Based Systems*, 2018, 151: 78-94.
- [22] YUE X, WANG Z, HUANG J G, et al. Graph embedding on biomedical networks: Methods, applications and evaluations[J]. *Bioinformatics*, 2020, 36(4): 1241-1251.
- [23] WU Y Z, CHEN Y K, YIN Z S, et al. A survey on graph embedding techniques for biomedical data: Methods and applications [J]. *Information Fusion*, 2023, 100: 101909.
- [24] ANTELM I A, CORDASCO G, POLATO M, et al. A survey on hypergraph representation learning [J]. *ACM Computing Surveys*, 2023, 56(1): 1-38.
- [25] WANG X, BO D Y, SHI C, et al. A survey on heterogeneous graph embedding: Methods, techniques, applications and sources [J]. *IEEE Transactions on Big Data*, 2022, 9(2): 415-436.
- [26] BELKIN M, NIYOGI P. Laplacian eigenmaps and spectral techniques for embedding and clustering[J]. *Advances in Neural Information Processing Systems*, 2001, 14: 585-591.
- [27] MENG H, ZHANG H, DING Y, et al. Component preserving laplacian eigenmaps for data reconstruction and dimensionality reduction[J]. *Applied Intelligence*, 2023, 53(23): 28570-28591.
- [28] CAI M, SHEN X, ABHADIOMHEN S E, et al. Robust dimensionality reduction via low-rank Laplacian graph learning [J]. *ACM Transactions on Intelligent Systems and Technology*, 2023, 14(3): 1-24.
- [29] AHMED A, SHERVASHIDZE N, NARAYANAMURTHY S, et al. Distributed large-scale natural graph factorization [C]// *Proceedings of the 22nd International Conference on World Wide Web*. New York: ACM, 2013: 37-48.
- [30] QIU J Z, DONG Y X, MA H, et al. NetSMF: Large-scale network embedding as sparse matrix factorization [C]// *Proceedings of the World Wide Web Conference*. New York: ACM, 2019:

- 1509-1520.
- [31] AGIBETOV A. Neural graph embeddings as explicit low-rank matrix factorization for link prediction[J]. *Pattern Recognition*, 2023, 133: 108977.
- [32] WAN L, FU Z, LING Y, et al. Z-Laplacian matrix factorization: Network embedding with interpretable graph signals[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2023, 36(8): 4331-4345.
- [33] GOLDBERG Y. Word2vec Explained: Deriving Mikolov et al. 's negative-sampling word-embedding method [J]. *arXiv*: 1402.3722, 2014.
- [34] 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 & Data Mining*. New York: ACM, 2017: 385-394.
- [35] RAHMAN T, SURMA B, BACKES M, et al. Fairwalk: Towards fair graph embedding[C]// *Proceedings of the 28th International Joint Conference on Artificial Intelligence*. San Francisco, CA: Morgan Kaufmann, 2019: 3289-3295.
- [36] ZHANG Y, SHEN J, ZHANG R, et al. Network representation learning via improved random walk with restart[J]. *Knowledge-Based Systems*, 2023, 263: 110255.
- [37] WANG R, LI Y K, LIN S, et al. Common neighbors matter: Fast random walk sampling with common neighbor awareness[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2022, 35(5): 4570-4584.
- [38] GAWLIKOWSKI J, TASSI C R N, ALI M, et al. A survey of uncertainty in deep neural networks[J]. *Artificial Intelligence Review*, 2023, 56: 1513-1589.
- [39] CONG S, ZHOU Y. A review of convolutional neural network architectures and their optimizations[J]. *Artificial Intelligence Review*, 2023, 56(3): 1905-1969.
- [40] REN H T, LU W, XIAO Y, et al. Graph convolutional networks in language and vision: A survey[J]. *Knowledge-Based Systems*, 2022, 251: 109250.
- [41] DEFFERRARD M, BRESSON X, VANDERGHEYNST P. Convolutional neural networks on graphs with fast localized spectral filtering[J]. *Advances in Neural Information Processing Systems*, 2016, 29: 3837-3845.
- [42] KIPF T N, WELING M. Semi-supervised classification with graph convolutional networks[J]. *arXiv*: 1609.02907, 2016.
- [43] WU F, AMAURI H, SOUZA J, et al. Simplifying graph convolutional networks[C]// *Proceedings of the 36th International Conference on Machine Learning*. New York: PMLR, 2019: 6861-6871.
- [44] JIANG X D, ZHU R H, LI P S, et al. Co-embedding of nodes and edges with graph neural networks[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, 45(6): 7075-7086.
- [45] WANG W, ZHANG G W, HAN H Y, et al. Correntropy-induced Wasserstein GCN: Learning graph embedding via domain adaptation[J]. *IEEE Transactions on Image Processing*, 2023: 32:3980-3993.
- [46] NIEPERT M, AHMED M, KUTZKOV K. Learning convolutional neural networks for graphs[C]// *Proceedings of the 33rd International Conference on Machine Learning*. New York: PMLR, 2016: 2014-2023.
- [47] HAMILTON W, YING Z, LESKOVEC J. Inductive representation learning on large graphs[C]// *Advances in Neural Information Processing Systems*. 2017: 1024-1034.
- [48] HU F Y, ZHU Y Q, WU S, et al. GraphAIR: Graph representation learning with neighborhood aggregation and interaction[J]. *Pattern Recognition*, 2021, 112: 107745.
- [49] TAILOR S A, OPOLKA F L, LIO P, et al. Do we need anisotropic graph neural networks? [J]. *arXiv*: 2104.01481, 2021.
- [50] ZHOU Y, HUO H, HOU Z, et al. Co-embedding of edges and nodes with deep graph convolutional neural networks[J]. *Scientific Reports*, 2023, 13(1): 16966.
- [51] VELIČKOVIĆ P, CUCURULL G, CASANOVA A, et al. Graph attention networks [J]. *arXiv*: 1710.10903, 2017.
- [52] HE J Y, WANG J M, YU Z Z. Attention based adversarially regularized learning for network embedding[J]. *Data Mining and Knowledge Discovery*, 2021, 35(5): 2112-2140.
- [53] YE Y, JI S H. Sparse graph attention networks[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2021, 35(1): 905-916.
- [54] HE L C, BAI L, YANG X, et al. High-order graph attention network[J]. *Information Sciences*, 2023, 630: 222-234.
- [55] XIE Y, ZHANG Y Q, GONG M G, et al. MGAT: Multi-view graph attention networks[J]. *Neural Networks*, 2020, 132: 180-189.
- [56] HUANG Z M, REN Y Z, PU X, et al. Self-supervised graph attention networks for deep weighted multi-view clustering[C]// *Proceedings of the 37th AAAI Conference on Artificial Intelligence*. Palo Alto, CA: AAAI, 2023: 7936-7943.
- [57] WANG X, ZHU M Q, BO D Y, et al. AM-GCN: Adaptive multi-channel graph convolutional networks[C]// *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. New York: ACM, 2020: 1243-1253.
- [58] RUMELHART D E, HINTON G E, WILLIAMS R J. Learning representations by back-propagating errors[J]. *Nature*, 1986, 323(6088): 533-536.
- [59] WANG D X, CUI P, ZHU W W. Structural deep network embedding[C]// *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. New York: ACM, 2016: 1225-1234.
- [60] LI J, LU G Q, WU Z T, et al. Multi-view representation model based on graph autoencoder[J]. *Information Sciences*, 2023, 632: 439-453.
- [61] SUN D D, LI D S, DING Z L, et al. Dual-decoder graph autoencoder for unsupervised graph representation learning [J]. *Knowledge-Based Systems*, 2021, 234: 107564.
- [62] CHEN Z L, WU Z H, WANG S P, et al. Dual low-rank graph autoencoder for semantic and topological networks[C]// Pro-

- ceedings of the 37th AAAI Conference on Artificial Intelligence. Palo Alto, CA; AAAI, 2023; 4191-4198.
- [63] KIPF T N, WELLING M. Variational graph auto-encoders[J]. arXiv:1611.07308, 2016.
- [64] GUO L, DAI Q. Graph clustering via variational graph embedding[J]. Pattern Recognition, 2022, 122: 108334.
- [65] ZHANG R, ZHANG Y X, LU C J, et al. Unsupervised graph embedding via adaptive graph learning[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 45(4): 5329-5336.
- [66] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets[C]// Advances in Neural Information Processing Systems. 2014: 2672-2680.
- [67] WANG H W, WANG J, WANG J L, et al. GraphGAN: Graph representation learning with generative adversarial nets[C]// Proceedings of the 32nd AAAI Conference on Artificial Intelligence. Palo Alto, CA; AAAI, 2018; 2508-2515.
- [68] HONG H T, LI X, WANG M Z. GANE: A generative adversarial network embedding[J]. IEEE Transactions on Neural Networks and Learning Systems, 2019, 31(7): 2325-2335.
- [69] HE D X, WANG T, ZHAI L, et al. Adversarial representation mechanism learning for network embedding[J]. IEEE Transactions on Knowledge and Data Engineering, 2021, 35(2): 1200-1213.
- [70] PAREJA A, DOMENICONI G, CHEN J, et al. EvolveGCN: Evolving graph convolutional networks for dynamic graphs[C]// Proceedings of the 34th AAAI Conference on Artificial Intelligence. Palo Alto, CA; AAAI, 2020; 5363-5370.
- [71] GAO C, ZHU J Y, ZHANG F, et al. A novel representation learning for dynamic graphs based on graph convolutional networks[J]. IEEE Transactions on Cybernetics, 2022, 53(6): 3599-3612.
- [72] ZHANG C Y, YAO Z L, YAO H Y, et al. Dynamic representation learning via recurrent graph neural networks[J]. IEEE Transactions on Systems, Man, and Cybernetics; Systems, 2022, 53(2): 1284-1297.
- [73] ZHU Y F, CONG F P, ZHANG D, et al. WinGNN: Dynamic graph neural networks with random gradient aggregation window[C]// Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery & Data Mining. New York; ACM, 2023; 3650-3662.
- [74] TRIVEDI R, FARAJTABAR M, BISWAL P, et al. DyRep: Learning representations over dynamic graphs[C]// Proceedings of the International Conference on Learning Representations. 2024.
- [75] ROSSI E, CHAMBERLAIN B, FRASCA F, et al. Temporal graph networks for deep learning on dynamic graphs[J]. arXiv: 2006.10637, 2020.
- [76] SOUZA A, MESQUITA D, KASKI S, et al. Provably expressive temporal graph networks[J]. Advances in Neural Information Processing Systems, 2022, 35: 32257-32269.
- [77] FU T Y, LEE W C, LEI Z. HIN2Vec: Explore meta-paths in heterogeneous information networks for representation learning[C]// Proceedings of the 2017 ACM on Conference on Information and Knowledge Management. New York; ACM, 2017; 1797-1806.
- [78] LEE S, PARK C, YU H. BHIN2vec: Balancing the type of relation in heterogeneous information network[C]// Proceedings of the 28th ACM International Conference on Information and Knowledge Management. New York; ACM, 2019; 619-628.
- [79] LIU Z, ZHANG S, ZHANG J, et al. HeteEdgeWalk: A heterogeneous edge memory random walk for heterogeneous information network embedding[J]. Entropy, 2023, 25(7): 998.
- [80] WANG X, JI H Y, SHI C, et al. Heterogeneous graph attention network[C]// Proceedings of the 19th World Wide Web Conference. New York; ACM, 2019; 2022-2032.
- [81] 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.
- [82] HE Y F, YAN D C, ZHANG Y W, et al. Semantic tradeoff for heterogeneous graph embedding[J]. IEEE Transactions on Computational Social Systems, 2022, 10(3): 1263-1276.
- [83] YANG X C, YAN M Y, PAN S R, et al. Simple and efficient heterogeneous graph neural network[C]// Proceedings of the 37th AAAI Conference on Artificial Intelligence. Palo Alto, CA; AAAI, 2023; 10816-10824.
- [84] LIU W Y, CHEN P Y, YEUNG S L, et al. Principled multilayer network embedding[C]// 2017 IEEE International Conference on Data Mining Workshops (ICDMW). IEEE, 2017; 134-141.
- [85] JING B Y, PARK C Y, TONG H H. HDMI: High-order deep multiplex infomax[C]// Proceedings of the Web Conference 2021. New York; ACM, 2021; 2414-2424.
- [86] MO Y J, CHEN Y H, LEI Y J, et al. Multiplex graph representation learning via dual correlation reduction[J]. IEEE Transactions on Knowledge and Data Engineering, 2023, 35(12): 12814-12827.
- [87] MO Y J, LEI Y J, SHEN J L, et al. Disentangled multiplex graph representation learning[C]// Proceedings of the 40th International Conference on Machine Learning. New York; PMLR, 2023; 24983-25005.
- [88] JI S Y, FENG Y F, JI R R, et al. Dual channel hypergraph collaborative filtering[C]// Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. New York; ACM, 2020; 2020-2029.
- [89] YANG C Q, WANG R J, YAO S C, et al. Semi-supervised hypergraph node classification on hypergraph line expansion[C]// Proceedings of the 31st ACM International Conference on Information & Knowledge Management. New York; ACM, 2022; 2352-2361.
- [90] ARYA D, GUPTA D K, RUDINAC S, et al. HyperSAGE: Generalizing inductive representation learning on hypergraphs[J]. arXiv: 2010.04558, 2020.
- [91] ZHANG R C, ZOU Y S, MA J. HyperSAGNN: A self-attention based graph neural network for hypergraphs[J]. arXiv: 1911.02613, 2019.

- [92] BAI S,ZHANG F H,TORR P H S. Hypergraph convolution and hypergraph attention[J]. *Pattern Recognition*, 2021, 110: 107637.
- [93] WU L C,WANG D L,SONG K S, et al. Dual-view hypergraph neural networks for attributed graph learning[J]. *Knowledge-Based Systems*, 2021, 227: 107185.
- [94] CHE X J,SUN Y P. Graph node classification algorithm based on similarity random walk aggregation[J]. *Journal of Jilin University(Engineering and Technology Edition)*, 2025, 55(6): 2069-2075.
- [95] WANG C P,WANG C K,WANG Z, et al. Edge2vec:Edge-based social network embedding[J]. *ACM Transactions on Knowledge Discovery from Data*, 2020, 14(4): 1-24.
- [96] ZHANG Y F,GAO S Q,PEI J, et al. Improving social network embedding via new second-order continuous graph neural networks[C]// *Proceedings of the 28th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. New York: ACM, 2022: 2515-2523.
- [97] KUMAR S,MALLIK A,KHETARPAL A, et al. Influence maximization in social networks using graph embedding and graph neural network[J]. *Information Sciences*, 2022, 607: 1617-1636.
- [98] LI J K,WANG R J,ZHANG F L, et al. Attribute Heterogeneous Network Embedding Method Combining Attention Mechanisms[J]. *Journal of Chinese Computer Systems*, 2024, 45(6): 1466-1473.
- [99] SANG L,WANG Y,ZHANG Y, et al. Intent-guided Heterogeneous Graph Contrastive Learning for Recommendation[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2025, 37(4): 1915-1929.
- [100] YU J L,YIN H Z,LI J D, et al. Self-supervised multi-channel hypergraph convolutional network for social recommendation [C]// *Proceedings of the Web Conference 2021*. 2021: 413-424.
- [101] ZHANG Y,TIAN J,SUN J, et al. HKGAT: Heterogeneous knowledge graph attention network for explainable recommendation system[J]. *Applied Intelligence*, 2025, 55(6): 549.
- [102] LI N,YANG Z,WANG J, et al. Drug-target interaction prediction using knowledge graph embedding [J]. *Iscience*, 2024, 27(6): 109393.
- [103] BARANWAL M,MAGNER A,SALDINGER J, et al. Struct Graph: A graph attention network for structure based predictions of protein-protein interactions[J]. *BMC Bioinformatics*, 2022, 23(1): 370.
- [104] SU X R,HU L,YOU Z H, et al. Multi-view heterogeneous molecular network representation learning for protein-protein interaction prediction[J]. *BMC Bioinformatics*, 2022, 23(1): 234.
- [105] GARCIA V,BRUNA J. Few-shot learning with graph neural networks[J]. *arXiv: 1711. 04043*, 2017.
- [106] XIE J,MENG Y,ZHAO Y, et al. Dynamicsemantic-based spatial-temporal graph convolution network for skeleton-based human action recognition[J]. *IEEE Transactions on Image Processing*, 2024, 33: 6691-6704.
- [107] AHMAD T,RIZVI S T H,KANWAL N. Transforming spatio-temporal self-attention using action embedding for skeleton-based action recognition[J]. *Journal of Visual Communication and Image Representation*, 2023, 95: 103892.
- [108] YAO L,MAO C S,LUO Y. Graph convolutional networks for text classification[C]// *Proceedings of the 33th AAAI Conference on Artificial Intelligence*. Palo Alto, CA: AAAI, 2019: 7370-7377.
- [109] RAGESH R,SELLAMANICKAM S,IYER A, et al. HeteGCN: Heterogeneous graph convolutional networks for text classification[C]// *Proceedings of the 14th ACM International Conference on Web Search and Data Mining*. New York: ACM, 2021: 860-868.
- [110] YANG T C,HU L M,SHI C, et al. HGAT: Heterogeneous graph attention networks for semi-supervised short text classification[J]. *ACM Transactions on Information Systems*, 2021, 39(3): 1-29.
- [111] MA Y L,LIU X F,ZHAO L J, et al. Hybrid embedding-based text representation for hierarchical multi-label text classification [J]. *Expert Systems with Applications*, 2022, 187: 115905.
- [112] SCHLICHTKRULL M,KIPF T N,BLOEM P, et al. Modeling relational data with graph convolutional networks[C]// *The Semantic Web: 15th International Conference*. 2018: 593-607.
- [113] BI Z,CHENG S,CHEN J, et al. Relphormer: Relational graph Transformer for knowledge graph representations[J]. *Neurocomputing*, 2024, 566: 127044.
- [114] LIANG S,ZHU A,ZHANG J, et al. Hyper-node relational graph attention network for multi-modal knowledge graph completion[J]. *ACM Transactions on Multimedia Computing, Communications and Applications*, 2023, 19(2): 1-21.
- [115] CHIANG W L,LIU X Q,SI S, et al. Cluster-GCN: An efficient algorithm for training deep and large graph convolutional networks[C]// *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. New York: ACM, 2019: 257-266.
- [116] LIANG J,GURUKAR S,PARTHASARATHY S. Mile: A multi-level framework for scalable graph embedding[C]// *Proceedings of the International AAAI Conference on Web and Social Media*. Palo Alto, CA: AAAI, 2021: 361-372.
- [117] JI Y G,YIN M Y,YANG H X, et al. Accelerating large-scale heterogeneous interaction graph embedding learning via importance sampling[J]. *ACM Transactions on Knowledge Discovery from Data*, 2020, 15(1): 1-23.
- [118] FENG A S,YOU C Y,WANG S Q, et al. KerGNNs: Interpretable graph neural networks with graph kernels[C]// *Proceedings of the 36th AAAI Conference on Artificial Intelligence*. Palo Alto, CA: AAAI, 2022: 6614-6622.
- [119] YANG Y M,GUAN Z Y,LI J X, et al. Interpretable and efficient heterogeneous graph convolutional network [J]. *IEEE Transactions on Knowledge and Data Engineering*, 2021, 35(2): 1637-1650.
- [120] HUANG Q,YAMADA M,TIAN Y, et al. GraphLIME: Local

- interpretable model explanations for graph neural networks[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2022, 35(7):6968-6972.
- [121] KIOUCHE A, LAGRAA S, AMROUCHE K, et al. A simple graph embedding for anomaly detection in a stream of heterogeneous labeled graphs [J]. *Pattern Recognition*, 2021, 112: 107746.
- [122] DUAN D, ZHANG C, TONG L, et al. An anomaly aware network embedding framework for unsupervised anomalous link detection[J]. *Data Mining and Knowledge Discovery*, 2024, 38(2):501-534.
- [123] HAN S, WOO S. Learning sparse latent graph representations for anomaly detection in multivariate time series[C]// *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*. New York: ACM, 2022: 2977-2986.
- [124] XIE H, ZHENG D, MA J, et al. Graph-aware language model pre-training on a large graph corpus can help multiple graph applications[C]// *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*. New York: ACM, 2023: 5270-5281.
- [125] LI P Y, WANG J, QIAO Y X, et al. An effective self-supervised framework for learning expressive molecular global representations to drug discovery[J]. *Briefings in Bioinformatics*, 2021, 22(6):bbab109.
- [126] RONG Y, BIAN Y T, XU T Y, et al. Self-supervised graph Transformer on large-scale molecular data[J]. *Advances in Neural Information Processing Systems*, 2020, 33:12559-12571.
- [127] SUN X G, YIN H Z, LIU B, et al. Heterogeneous hypergraph embedding for graph classification[C]// *Proceedings of the 14th ACM International Conference on Web Search and Data Mining*. New York: ACM, 2021: 725-733.
- [128] LIU J, SONG L Y, WANG G T, et al. Meta-HGT: Metapath-aware hypergraph Transformer for heterogeneous information network embedding[J]. *Neural Networks*, 2023, 157: 65-76.
- [129] FU C, YU P, YU Y, et al. MHGCN+: Multiplex heterogeneous graph convolutional network[J]. *ACM Transactions on Intelligent Systems and Technology*, 2024, 15(3): 1-25.



HUANG Miaomiao, born in 1999, Ph.D. Her main research interests include graph learning and drug discovery.



ZHAO Yuhai, born in 1975, Ph.D, professor. His main research interests include data mining and machine learning.

(责任编辑:柯颖)