

华中科技大学

图神经网络实验报告

专业: 计算机科学与技术

班级: CS2501

学号: M202573989

姓名: 张昭骏

电话: 15713716198

邮箱: 1357299736@qq.com

目 录

1 相关工作与背景	1
1.1 图结构数据与图表示学习	1
1.2 图神经网络模型及其扩展	2
1.3 图神经网络应用	4

1 相关工作与背景

近年来，面向图结构数据的表示学习与图神经网络（Graph Neural Networks, GNN）得到了广泛关注。本节对相关工作进行简要综述，首先回顾图表示学习的基础，其次讨论图神经网络模型的主要思想与扩展，最后重点介绍其在代码智能等领域的应用，为后续实验奠定背景。

1.1 图表示学习

图（Graph）是刻画实体间关系的通用数据结构，在社交网络、知识图谱、生物分子及程序代码等领域普遍存在。图表示学习（Graph Representation Learning）旨在将图中的节点、边或整图映射到低维向量空间，以便于后续机器学习模型处理。

早期的图表示学习主要依赖两条思路。一是基于谱图论的方法，利用图拉普拉斯矩阵的谱（特征值与特征向量）来刻画图的结构，将节点嵌入到能保持图“平滑性”的低维空间。二是基于随机游走的方法，通过在图上采样节点序列，再借鉴自然语言处理中的词向量技术（如 word2vec）学习节点表示，从而保持邻近关系与高阶共现信息。这些方法为图数据的向量化提供了基础，但通常与下游任务解耦，缺乏端到端的优化能力，这催生了图神经网络的出现。

1.2 图神经网络

图神经网络将深度学习中的“局部聚合”思想推广到图结构上，形成了可端到端训练的模型框架。根据卷积定义方式，GNN 可分为在图傅立叶域进行滤波的谱域方法，和直接在空间邻域上传播信息的空间域（或消息传递）方法，后者是当前应用的主流。根据任务粒度，又可分为节点级、边级和图级 GNN。尽管模型各异，但其核心思想是在每一层网络中，每个节点聚合其邻居节点的信息来更新自身表示，通过堆叠多层来扩大感受野，捕获更大范围的结构信息。

随着 GNN 的发展，研究者进一步从鲁棒性和可扩展性等方面进行了扩展。为应对图数据中的噪声或对抗攻击，鲁棒 GNN 通过对抗训练、结构正则化等方式提升模型的稳定性。为处理超大规模图，可扩展 GNN 通过邻居采样、子图训

练等策略解决了“邻域爆炸”带来的计算与存储瓶颈。

1.3 图神经网络应用

GNN 已被广泛应用于具有图结构的数据场景。在代码智能（Code Intelligence）领域，程序语言严格的语法和语义为构建高质量的程序图（如抽象语法树 AST、控制流/数据流图等）提供了便利。在这些图上应用 GNN，能够有效捕捉代码的结构化特征，完成代码漏洞检测（图级分类）、变量误用检测与类型推断（节点级预测）等多种静态分析任务。本报告的实验将聚焦于类型推断，将其视为在程序图上进行节点属性预测的一个实例。

除代码外，GNN 在自然语言处理（如基于依存句法树进行关系抽取）、生物信息学（如基于分子图预测理化性质）和知识图谱（如进行知识补全与推理）等领域也取得了显著成功。

近年来，图神经网络与大语言模型（LLM）的结合成为前沿热点。一方面，可利用 GNN 将图结构信息编码为向量或文本，作为 LLM 的输入以增强其结构化推理能力；另一方面，也可利用 LLM 为图的节点和边生成更丰富的语义特征。二者的结合为基于程序图的类型推断等任务提供了新的扩展思路。