

# 图神经网络在动态图上的链路预测综述

王天赐  
计算机学院  
湖北工业大学  
武汉, 湖北, 中国

**摘要**—图神经网络 (GNNs) 正迅速成为在图形结构数据上学习的主导方式。链接预测是新的 GNN 模型的一个几乎通用的基准。许多先进的模型, 如动态图神经网络 (DGNNs) 专门针对动态链接预测。然而, 这些模型, 特别是 DGNNs, 很少与其他模型或现有启发式方法进行比较。不同的工作以不同的方式评估他们的模型, 因此人们无法直接比较评估指标。受此启发, 我们进行了一次全面的比较研究。我们比较了链接预测启发式方法、GNNs、离散 DGNNs 和连续 DGNNs 对动态链接预测的作用。我们发现, 简单的链接预测启发式方法往往比 GNN 和 DGNN 表现更好, 而在所有被考察的图神经网络中, DGNN 的表现一直优于静态 GNN。

**Index Terms**—动态图神经网络; 启发式; 链接预测;

## I. 引言

### A. 研究背景

在本文中主要的研究内容是在动态图上的链接预测, 与静态图的链接预测不同的是, 动态图的链接预测增加了时间维度, 预测难度增加, 但是在现实中的应用更加广泛。

### B. 研究方法

在近年来的研究中, 很多链路预测方法取得了发展, 在静态图上研究者们提出了很多高精度的链接预测方法。比如基于启发式算法的共同邻居算法 [1], 以及首次将 GNN 应用在图的链接预测的 SEAL 模型 [1], SEAL 模型

### C. 研究问题

近年来, 图神经网络 (GNNs) 作为一个新兴的研究领域, 得到了长足的发展, 提出并发展了多种体系结构。但是在动态图神经网络 (DGNN) 领域, 这些问题因以下原因而进一步加剧:

- 1) 数据的动态性质:
- 2) 缺乏通用术语:

3) 缺乏既定的强大基线 (大多数研究不与其他 DGNN 比较性能):

4) 离散和连续 DGNN 之间的鸿沟:

5) 大量的实验设计选择: 这些选择包括: 如何表示动态网络 (如快照、时间窗口、连续、边的生存时间等), 包括哪些节点特征, 如何将数据分成训练-验证-测试集, 用哪些指标来评价结果, 如何在报告的指标中使用负采样率, 以及如何选择/优化神经网络参数 (如学习率、早期停止准则、嵌入空间维度等)。所有这些都意味着, 通过阅读研究论文来比较方法的性能是不可能的, 除非他们清楚地说明他们所有的设计选择, 而且这些设计选择在不同的论文中是相同的。这些选择包括: 如何表示动态网络 (如快照、时间窗口、连续、边的生存时间等), 包括哪些节点特征, 如何将数据分成训练-验证-测试集, 用哪些指标来评价结果, 如何在报告的指标中使用负采样率, 以及如何选择/优化神经网络参数 (如学习率、早期停止准则、嵌入空间维度等)。所有这些都意味着, 通过阅读研究论文来比较方法的性能是不可能的, 除非他们清楚地说明他们所有的设计选择, 而且这些设计选择在不同的论文中是相同的。DGNNs 是建立网络动态模型的一个很有前途的途径, 因为它们既能通过 GNNs 编码空间模式, 又能通过时间序列组件 (如循环神经网络 (RNN) 或自我注意) 编码时间模式。然而, 迄今为止提出的 DGNNs 已经在少数数据集上进行了测试, 并且很少与其他 DGNNs 进行比较。不同的研究在不同的数据集上进行比较, 因为在 DGNN 基准测试中使用哪种数据集的问题上没有共识。

## II. 相关研究模型与数据集

### A. 相关数据集

### B. 相关研究模型

## 参考文献

- [1] J. Zhang, J. Tang, B. Liang, Z. Yang, S. Wang, J. Zuo, and J. Li, “Recommendation over a heterogeneous social network,” in *Web-Age Information Management, 2008. WAIM'08. The Ninth International Conference on.* IEEE, 2008, pp. 309–316.
- [2] O. Ronneberger, P. Fischer, and T. Brox, *U-Net: Convolutional Networks for Biomedical Image Segmentation.* Springer International Publishing, 2015.