

Pre-Training of GNN

刘闯

chuangliu@whu.edu.cn

2020.03.09



Pre-Training

NLP 和 CV 中的预训练结构: BERT , VGG-Nets

简化下游任务的训练过程

- 1. 从未标注的数据中学习输入数据的通用表征**
- 2. 并为模型提供更合理的初始化参数**



Pre-Training of GNN

例如：

- * 用图神经网络做图上节点的分类，然而有标签的节点很少，设计一些其他任务，比如利用图神经网络预测节点的度，节点的度信息可以简单的统计得到
- * 通过这样的学习，图神经网络能够学习到每个节点在图结构中的局部信息，而这些信息对于最终的节点分类任务是有帮助的

上述预训练得到：

1. 节点度预测的 embedding
2. 节点度预测权重矩阵

继而：

1. embedding 接到分类器，用 label 数据学习
2. 使用有标签数据继续训练，调整矩阵

PS: GNNs 中需要预测测试集中的一些图，这些图具有不同于训练集中所包含图的图结构，导致容易陷入 out-of-distribution 预测问题。例如预测新合成的，与训练集分子结构不同的分子的化学属性，或者来自新物种的蛋白质的功能，其具有与先前研究的物种不同的 PPI 网络结构。



Reference

1. **Strategies for Pre-training Graph Neural Networks**, ICLR 2020, Stanford , **Jure Leskovec** , (P5 - P7)
2. **Pre-Training Graph Neural Networks for Generic Structural Feature Extraction**, UCLA , **Yizhou Sun** , (P8 - P10)



Pre-Training GNN

Pre-Training Graph Neural Networks for Generic Structural Feature Extraction

1. **GNNs 是否能够从预训练中受益?**
2. **设置哪几种预训练任务比较合理?**



Pre-Training Framework

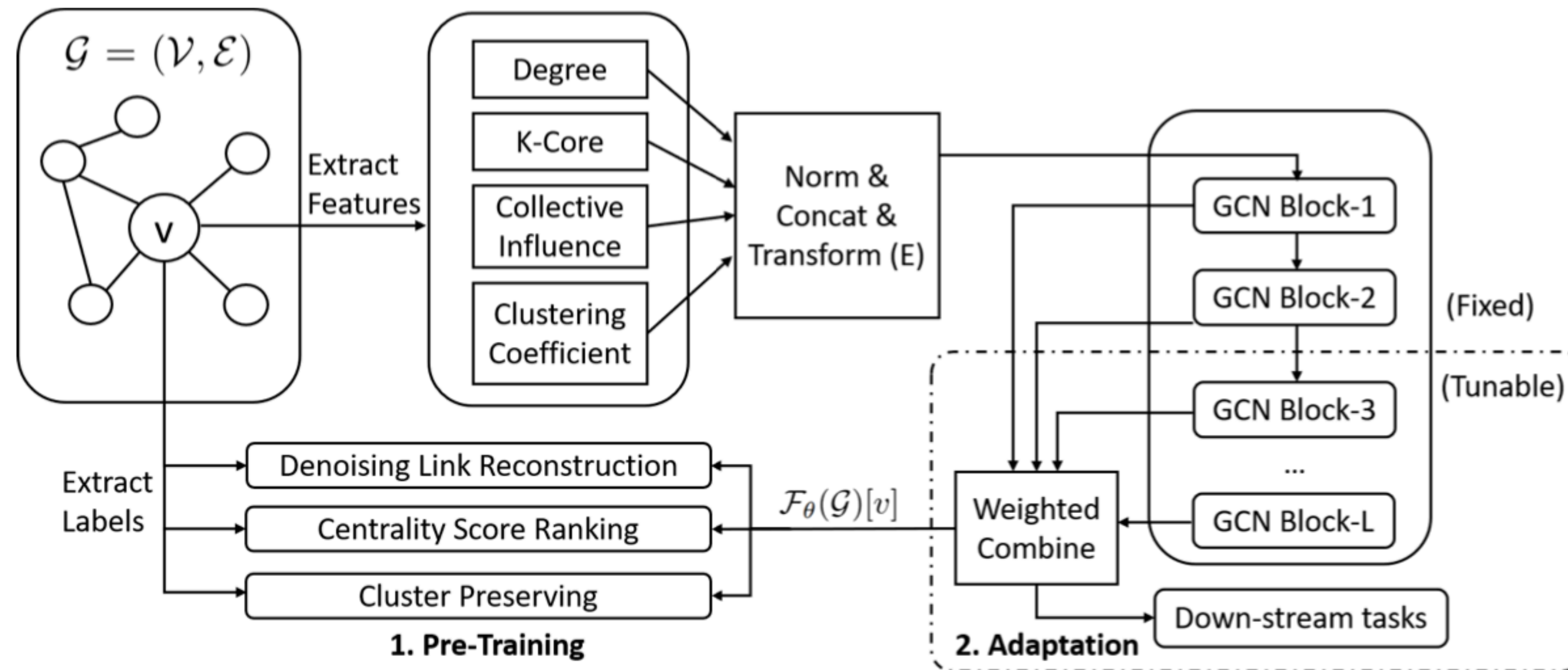
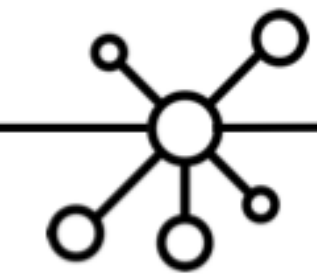


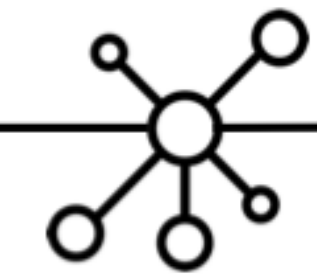
Figure 1: The overall framework for pre-training and adaption. First, a multi-layer GNN is trained on three structure-guided tasks. Then, the lower layer of the GNN is fixed and the upper layer of the GNN is fine-tuned on the given downstream task.

- 边重建：首先 mask 一些边得到带有噪声的图结构，训练图神经网络预测 mask 掉的边
- Centrality Score Ranking：通过对每个节点计算不同的 Centrality Score，其中，包括：Eigencentrality, Betweenness, Closeness和 Subgraph Centrality；然后，通过各个 Centrality Score 的排序值作为label训练 GCN；
- 保留图簇信息：计算每个节点所属的子图，然后训练 GNNs 得到节点特征表示，要求这些节点特征表示仍然能保留节点的子图归属信息。



Adaptation Procedure

**Feature-based
Fine-Tuning**



Pre-Training of GNN

Strategies for Pre-training Graph Neural Networks

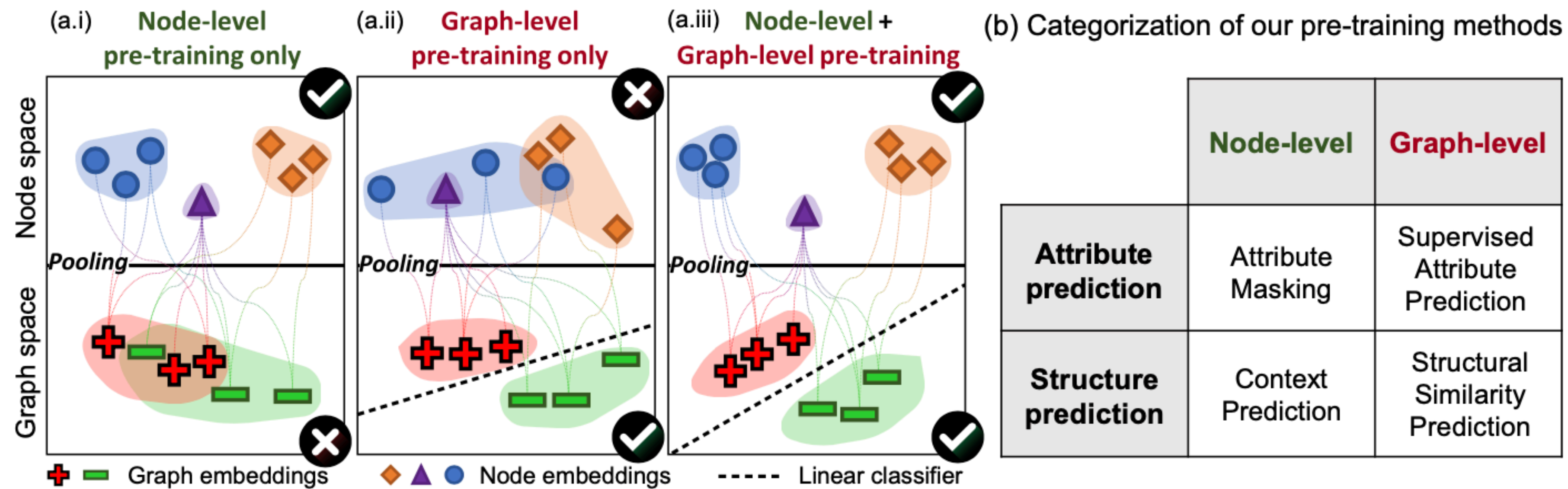
Challenge

- 如何生成节点 embedding，以捕获节点和它邻居的在结构的相似性(相邻节点的Embedding在投影空间中相近)；
- 如何生成可组合(composable)的节点 embedding，这些Embedding通过 pooling的方式能够刻画整个图的Embedding；
- 如何生成领域知识(domain-knowledge)相关的节点 embedding

PS: GNNs 中需要预测测试集中的一些图，这些图具有不同于训练集中所包含图的图结构，导致容易陷入 out-of-distribution 预测问题。例如预测新合成的，与训练集分子结构不同的分子的化学属性，或者来自新物种的蛋白质的功能，其具有与先前研究的物种不同的 PPI 网络结构。



Pre-Training Strategies



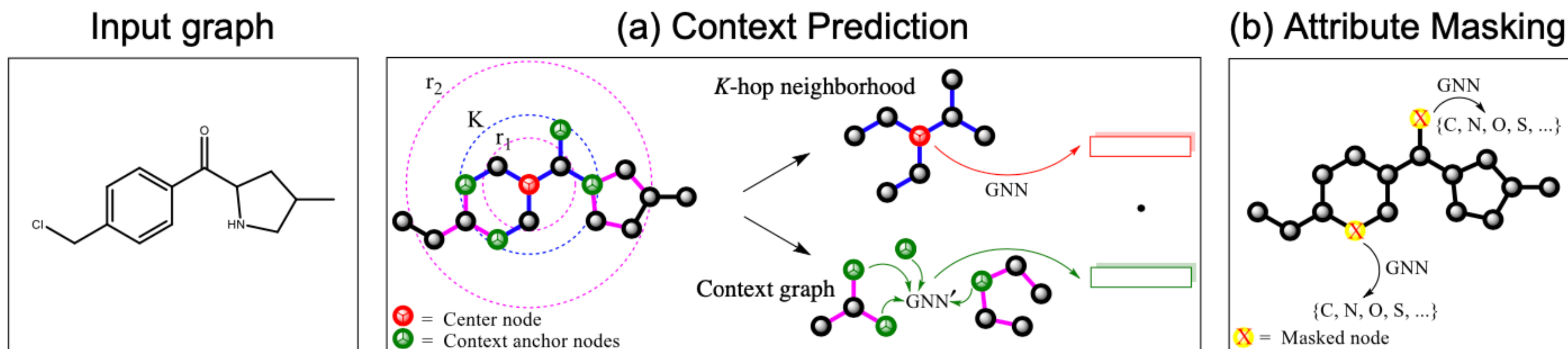
1. 为了属性预测：在节点的层次上，作者设计了masking任务，在图的层次上，作者设计了监督属性预测的任务
2. 为了学习图结构信息，在节点的层次上，作者设计了上下文预测的任务，在图的层次上，作者设计了结构相似性预测的任务。

a图：

如果只考虑节点层次：节点可分，图不可分；

只考虑图层次：图可分，节点语义不一致

Pre-Training Strategies: Node-Level



上下文预测：Word2vec 的 distributional hypothesis

GNN中每个节点的Embedding都是由其邻居节点更新的，因此作者采用节点 v 的 K -hop 子图对 v 进行编码，最终获得节点的embedding

K -hop 作为词， r_1 ， r_2 作为上下文

masking

对节点/边的某些属性进行masking，然后利用学习到的节点Embedding进行预测