

# 图神经网络的可扩展性

教材：图深度学习，电子工业出版社  
<https://baike.baidu.com/item/图深度学习>



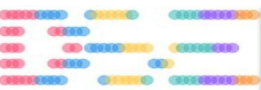


 图神经网络的可扩展性

 逐点采样法

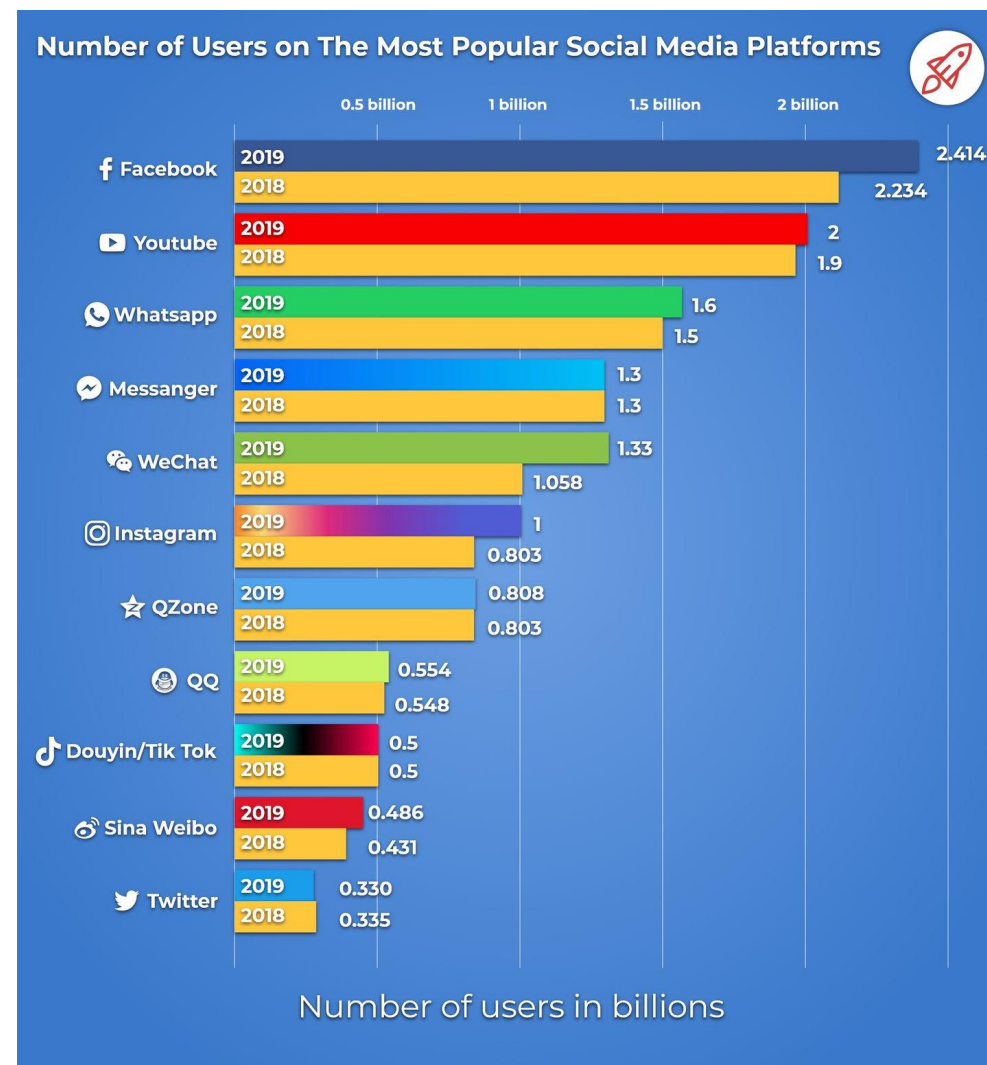
 逐层采样法

 子图采样法



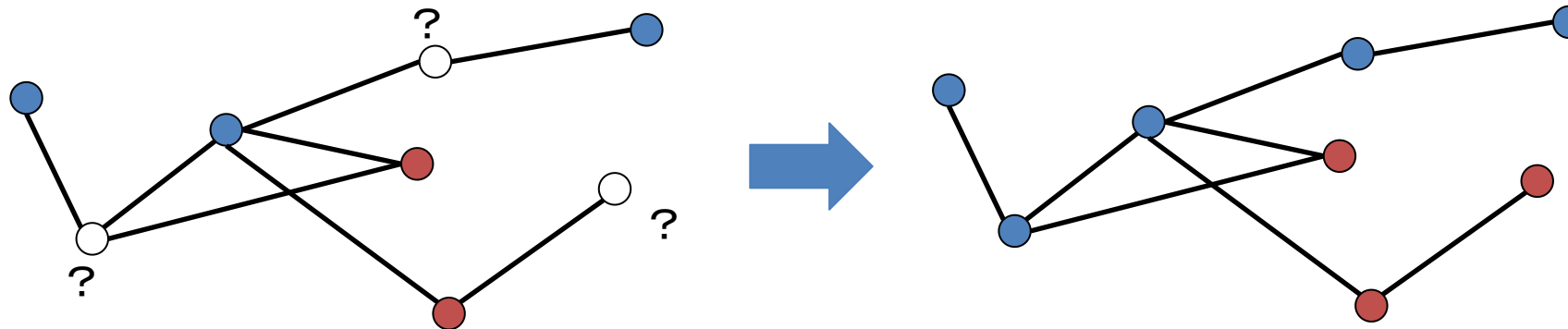


# 图的规模越来越大



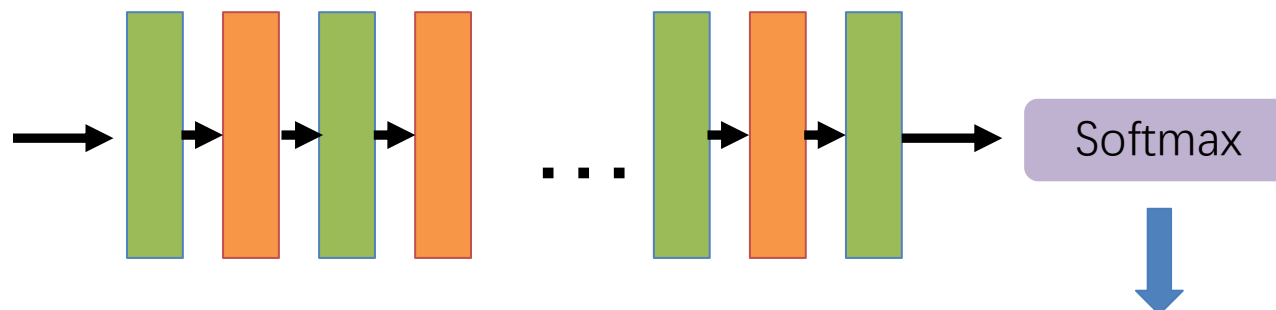


# 利用GNN进行节点分类

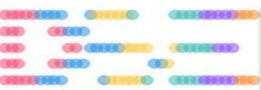


图滤波操作

激活函数

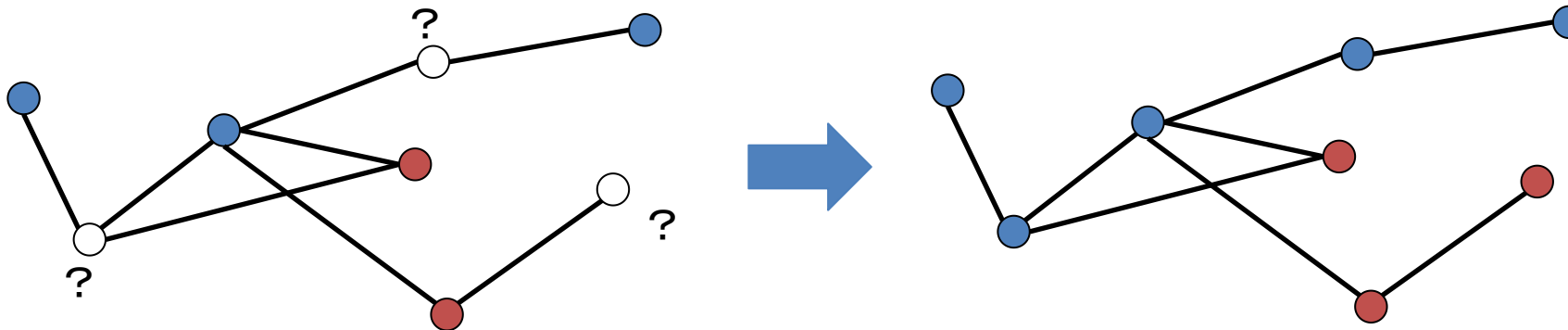


对于第*i*个节点  $f_{GCN}(\mathbf{A}, \mathbf{F}; \Theta)$





## 利用GNN进行节点分类



训练节点集合  $\mathcal{V}_{tr}$

$$\mathcal{L}_{\text{train}} = \sum_{v_i \in \mathcal{V}_{tr}} \ell(f_{GCN}(\mathbf{A}, \mathbf{F}; \Theta)_i, y_i)$$

$$\mathbf{F}^{(l)} = \hat{\mathbf{A}} \mathbf{F}^{(l-1)} \Theta^{(l-1)}, \quad l = 1, \dots, L$$

空间复杂度:  $O(L \cdot |V| \cdot d)$



## 分批训练

$$\mathcal{L}_{\text{train}} = \sum_{v_i \in \mathcal{V}_{tr}} \ell(f_{GCN}(\mathbf{A}, \mathbf{F}; \Theta)_i, y_i)$$

抽取一个批次的节点，仅在这个批次上进行训练

$\mathcal{B}$

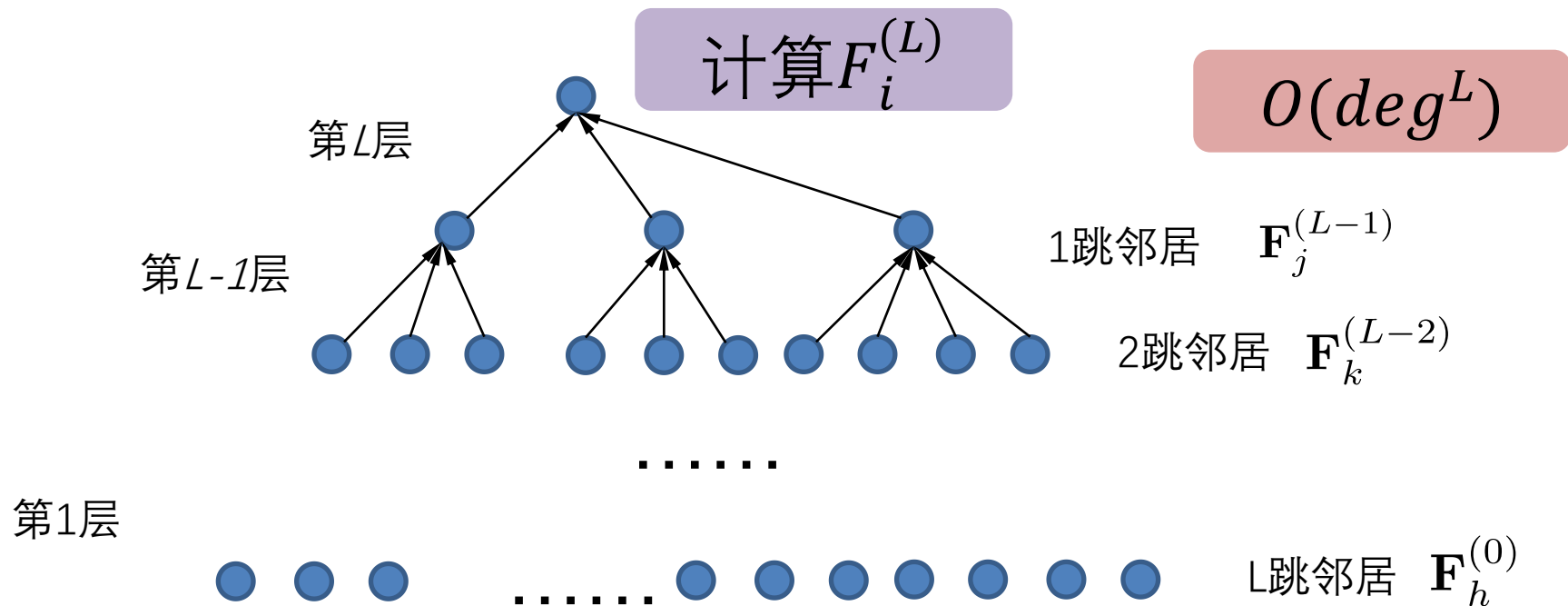
$$\mathbf{F}^{(l)} = \hat{\mathbf{A}} \mathbf{F}^{(l-1)} \Theta^{(l-1)}, \quad l = 1, \dots, L$$

$$\mathbf{F}_i^{(l)} = \sum_{v_j \in \tilde{\mathcal{N}}(v_i)} \hat{\mathbf{A}}_{i,j} \mathbf{F}_j^{(l-1)} \Theta^{(l-1)}, \quad l = 1, \dots, L$$



## 计算一个节点的第L层表示

$$\mathbf{F}_i^{(l)} = \sum_{v_j \in \tilde{\mathcal{N}}(v_i)} \hat{\mathbf{A}}_{i,j} \mathbf{F}_j^{(l-1)} \Theta^{(l-1)}, \quad l = 1, \dots, L$$



 图神经网络的可扩展性

 逐点采样法

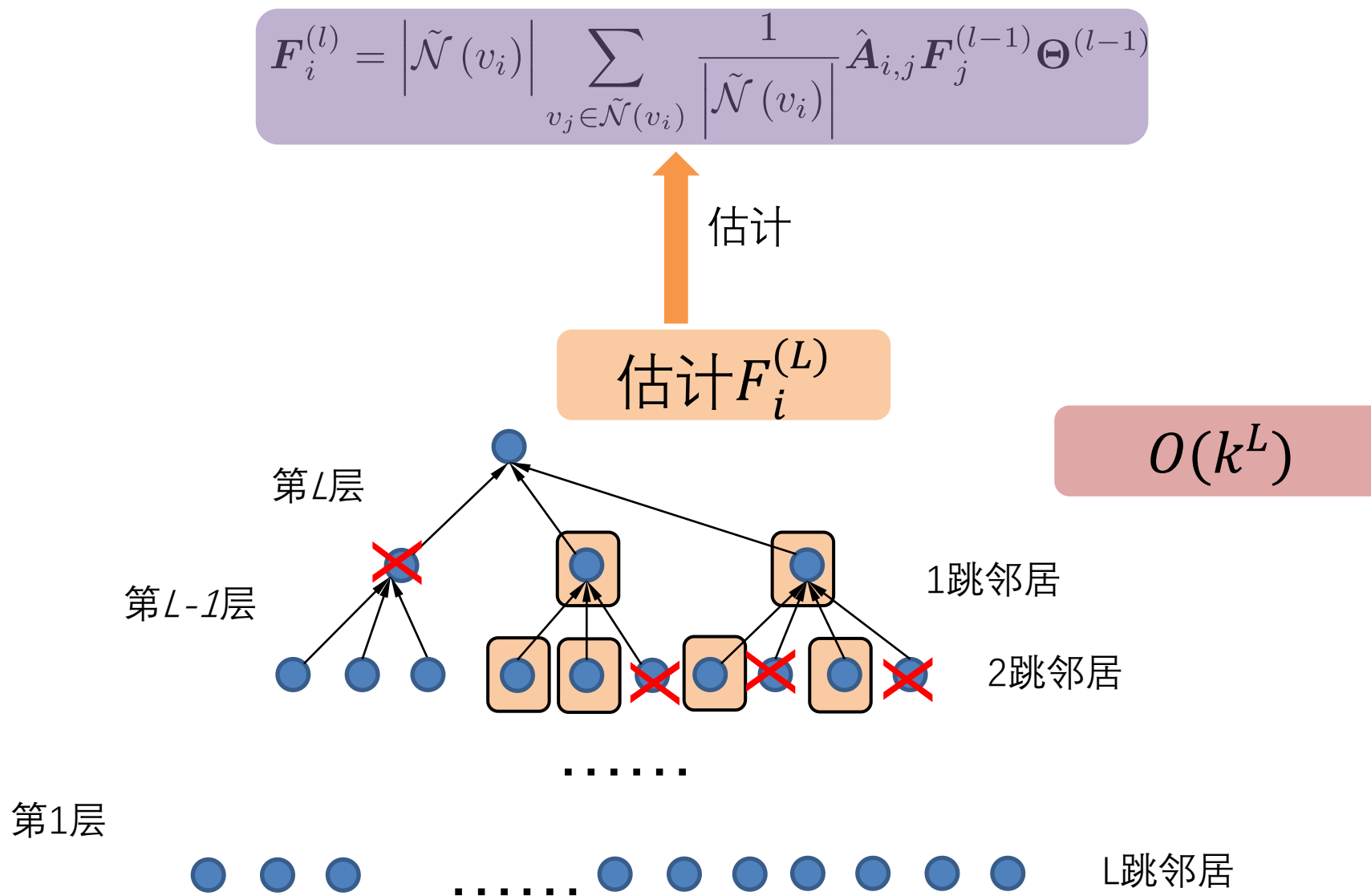
 逐层采样法

 子图采样法





# 逐点采样法



 图神经网络的可扩展性

 逐点采样法

 逐层采样法

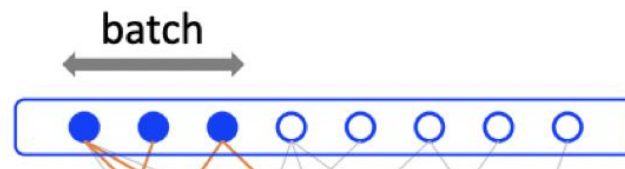
 子图采样法



## 逐层采样法

$$F_i^{(l)} = \left| \tilde{\mathcal{N}}(v_i) \right| \sum_{v_j \in \tilde{\mathcal{N}}(v_i)} \frac{1}{\left| \tilde{\mathcal{N}}(v_i) \right|} \hat{A}_{i,j} F_j^{(l-1)} \Theta^{(l-1)}$$

估计

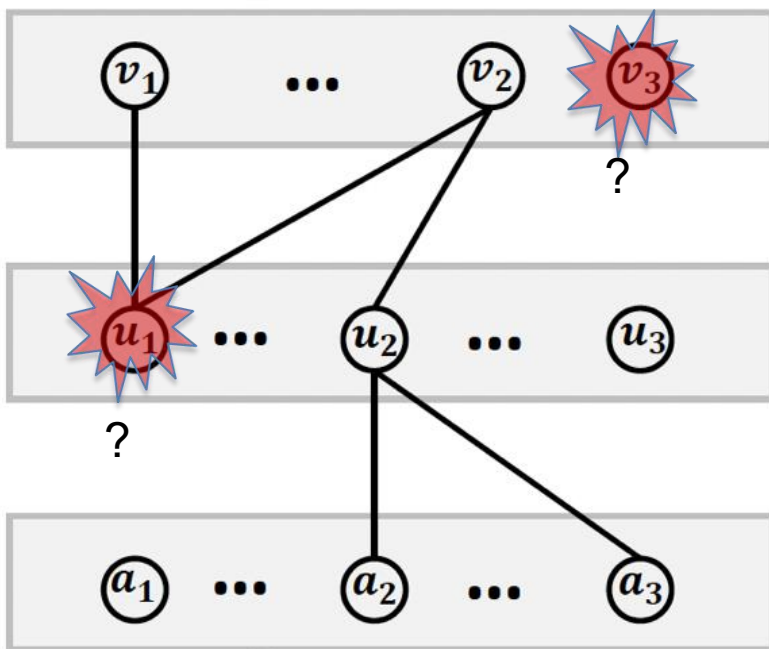


Minibatch

$O(kL)$



## 逐层采样法可能存在的问题



- ❑ 有的节点可能没有连接
- ❑ 无法计算它们的表示

采样时需要让层与层之间有较为紧密的连接

 图神经网络的可扩展性

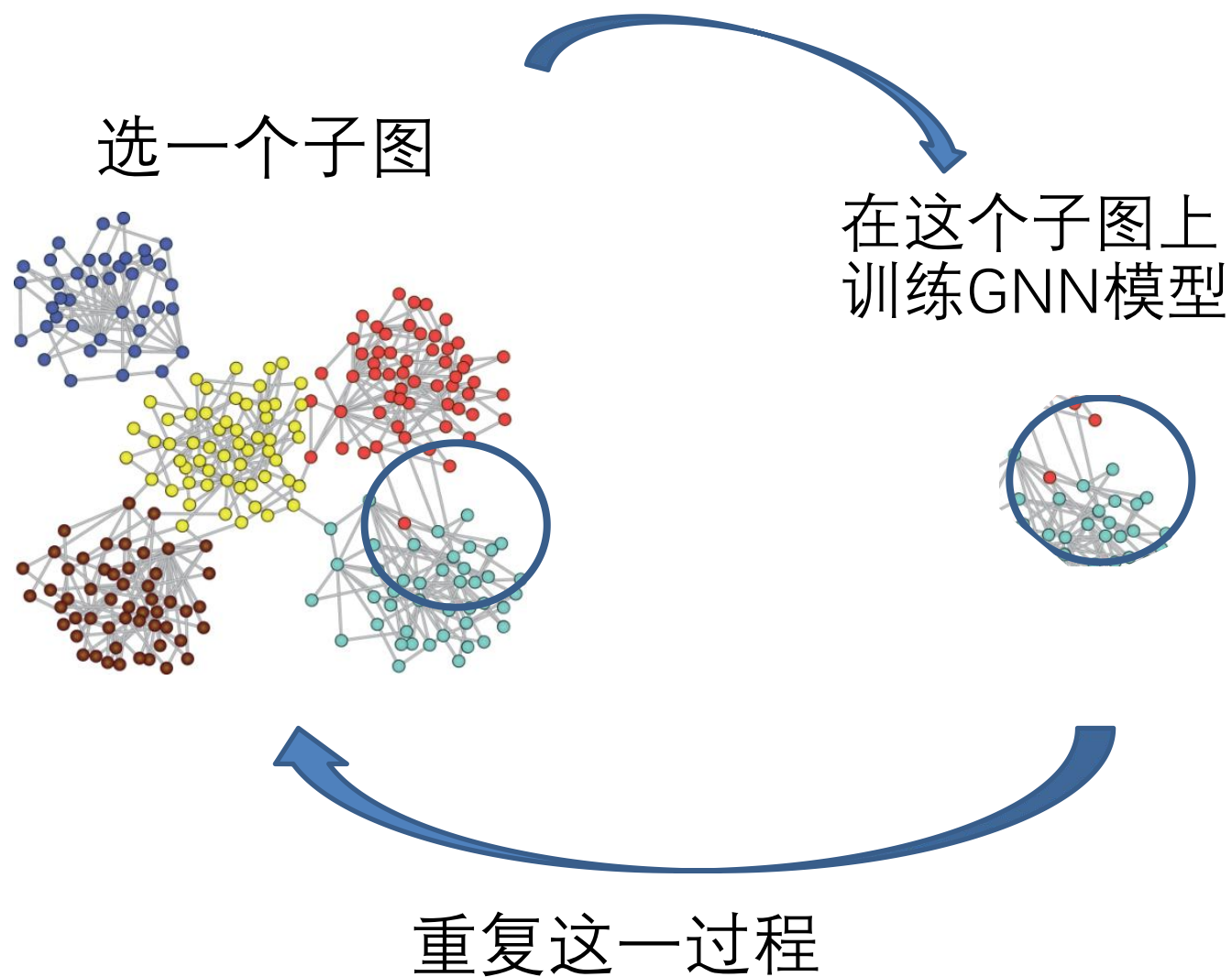
 逐点采样法

 逐层采样法

 子图采样法

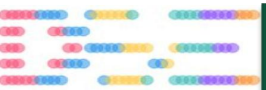
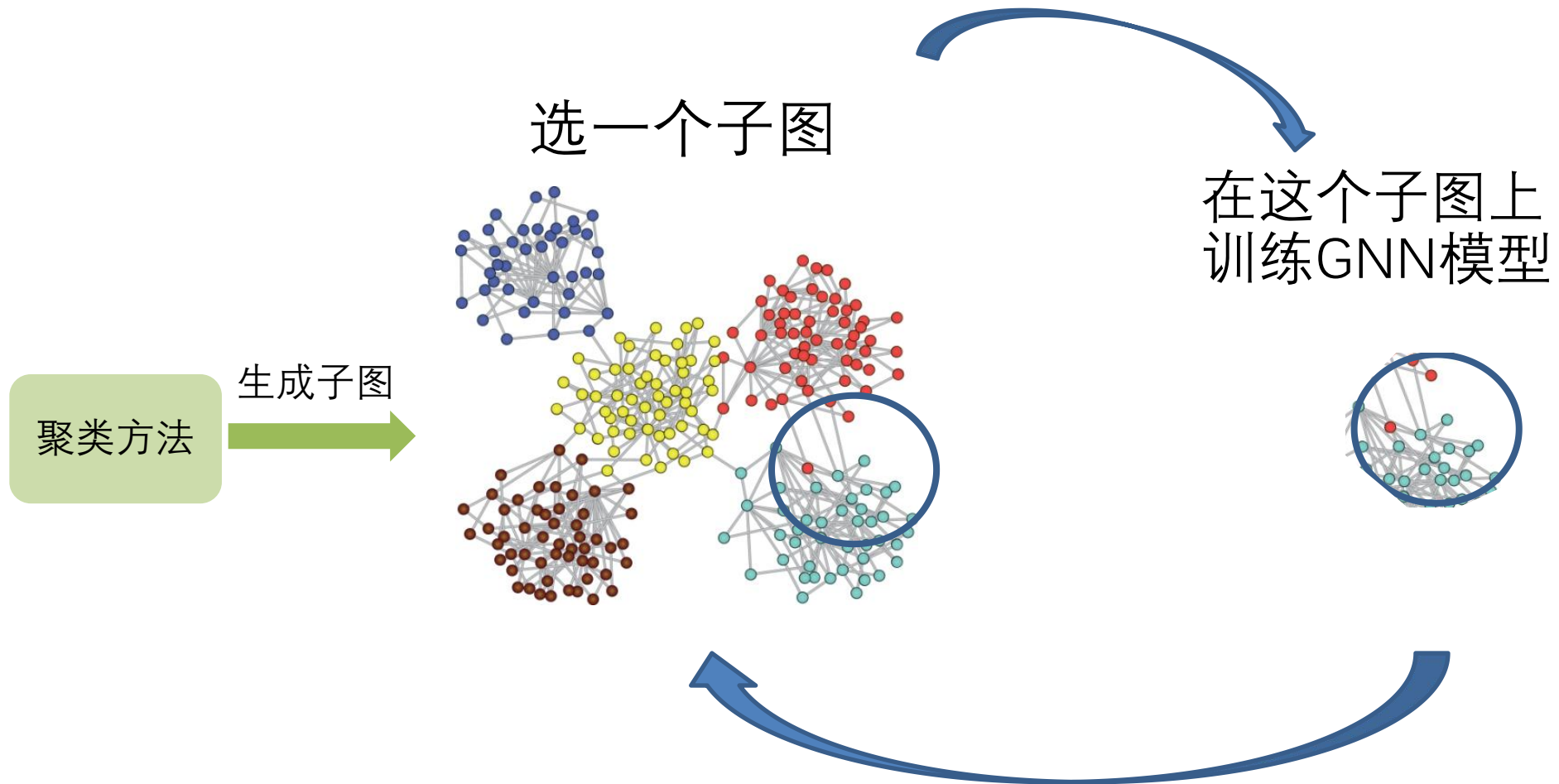


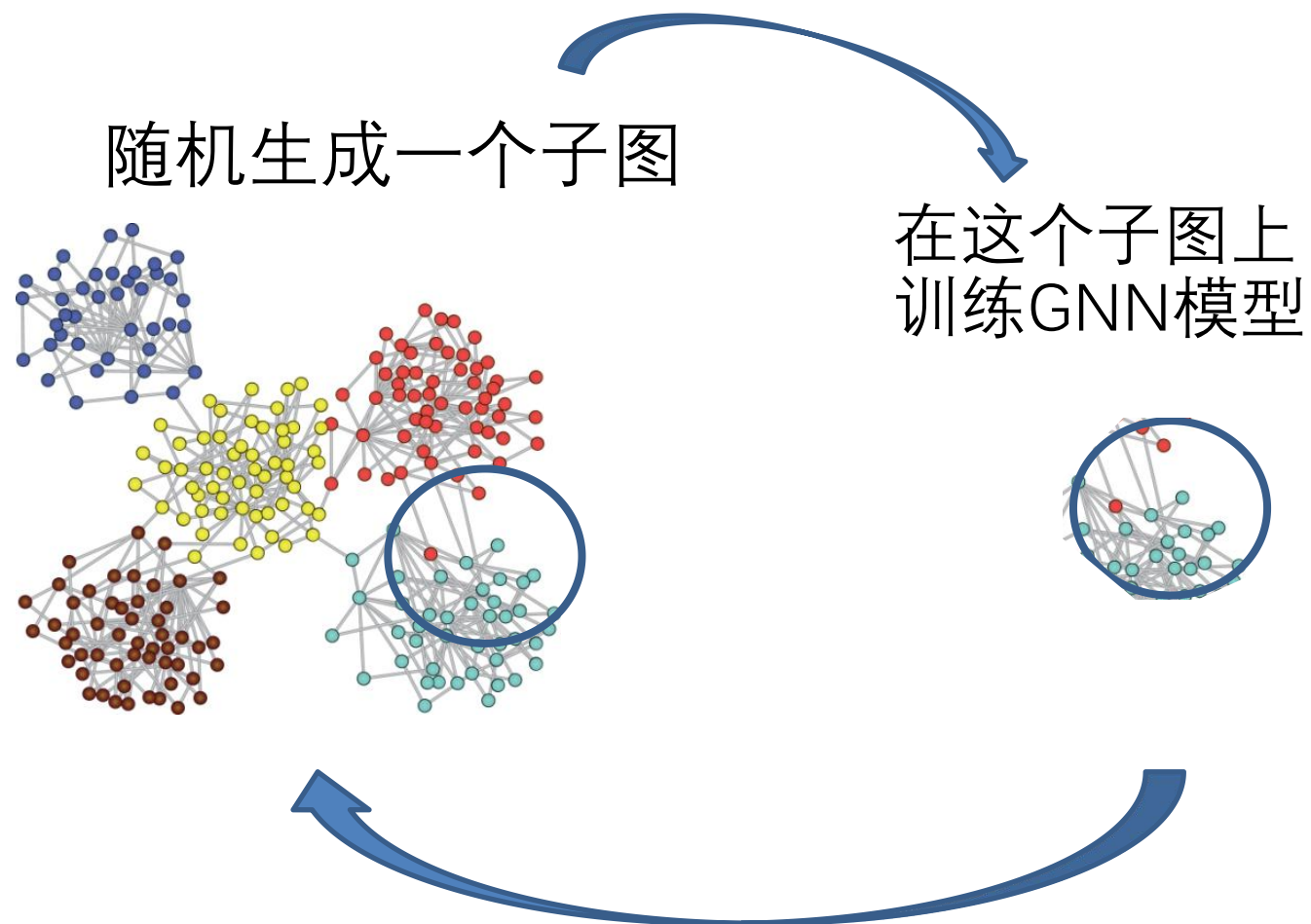
# 子图采样法



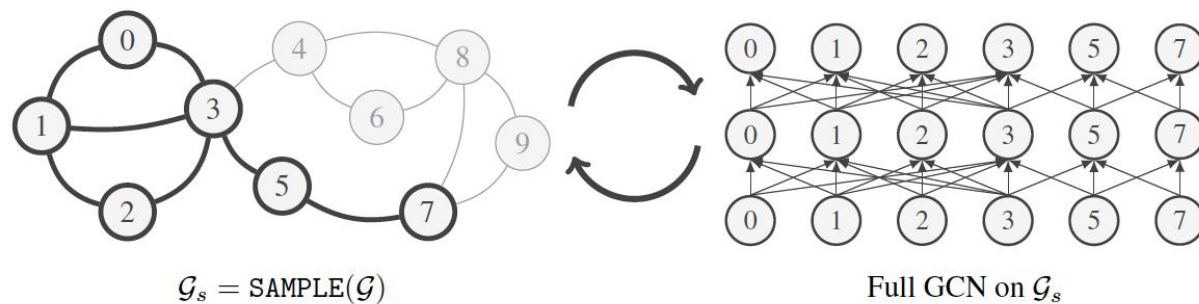


# Cluster-GCN

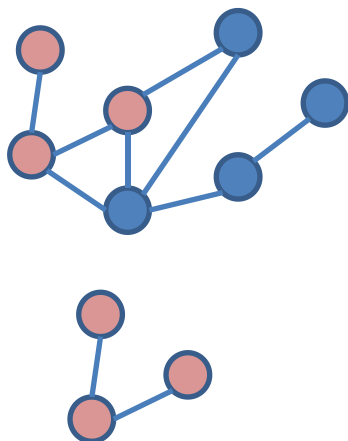




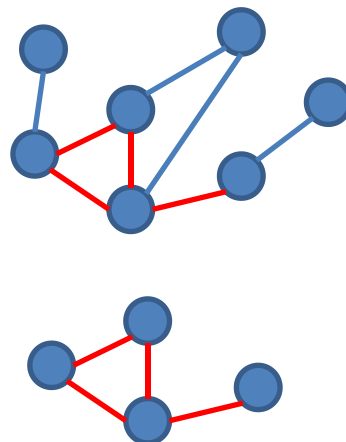




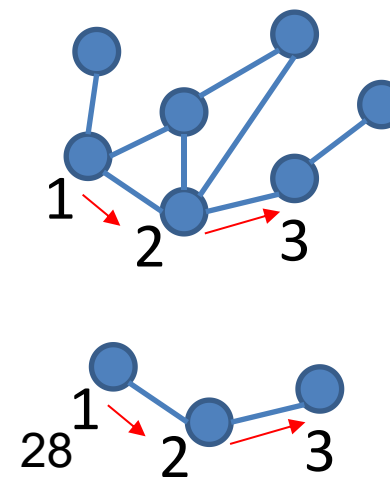
## Random node sampling



## Random edge sampling



## Random walk sampling



 图神经网络的可扩展性

 逐点采样法

 逐层采样法

 子图采样法