# 第3章　图神经网络

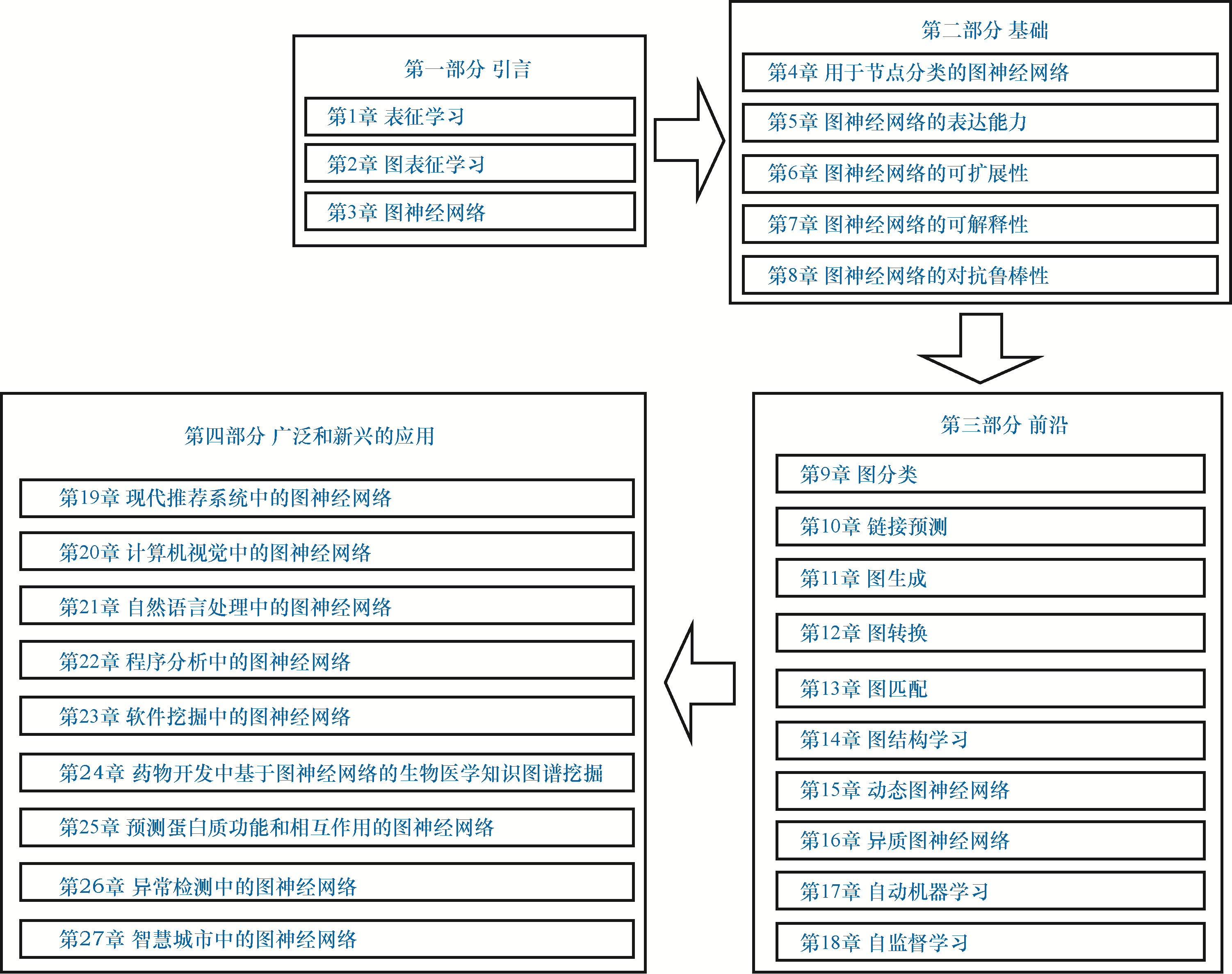


图3.1 本书的组织结构

# 第4章　用于节点分类的图神经网络

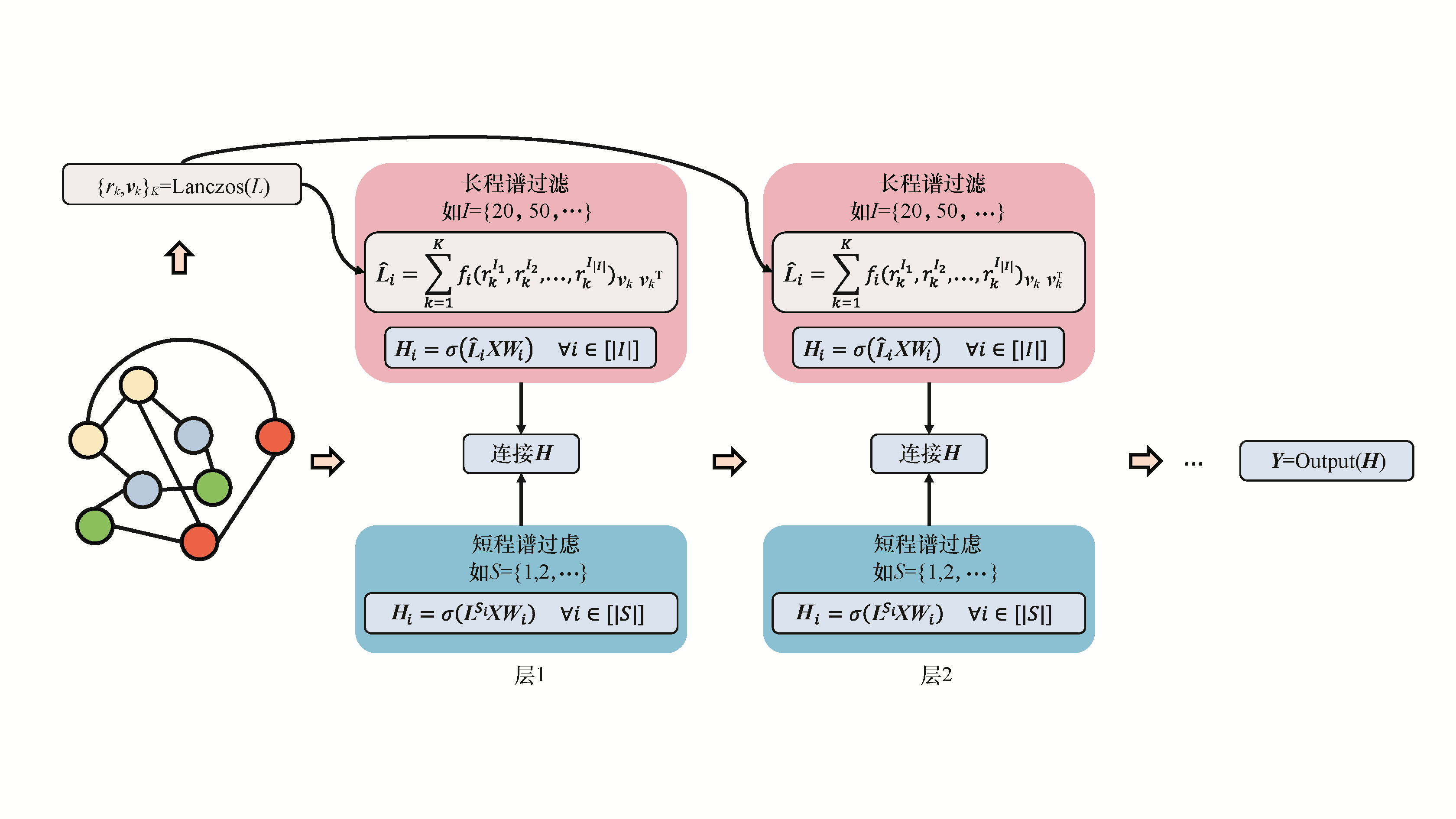


图4.1 Lanczos网络的推理过程。其中，近似的顶部特征值{*rk*}和特征向量{***v****k*}是由Lanczos算法计算得出的。请注意，这个步骤只对每个图执行一次。远程的图卷积（图的顶部）是通过图拉普拉斯的低秩近似来有效计算的。我们可以控制尺度（即特征值的指数）作为超参数。可学习的谱过滤器被应用于近似的顶部特征值{*rk*}。短程的图卷积（图的底部）与GCN相同。图4.1改编自Liao et al（2019b）中的图1

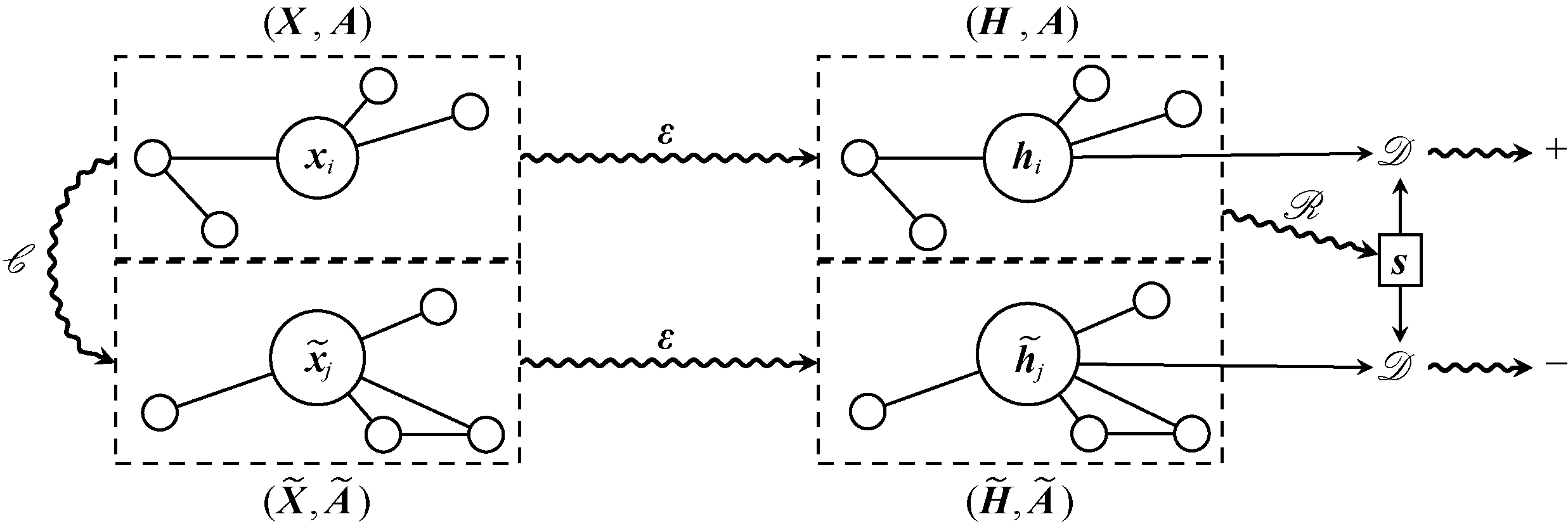


图4.2 深度图信息最大化的整体流程。上方的路径显示的是如何处理正样本，  
下方的路径显示的是如何处理负样本。请注意，正负两方面的图表征是共享的。  
正样本和负样本的子图不一定不同。图4.2改编自Veličković et al（2019）中的图1

# 第5章　图神经网络的表达能力

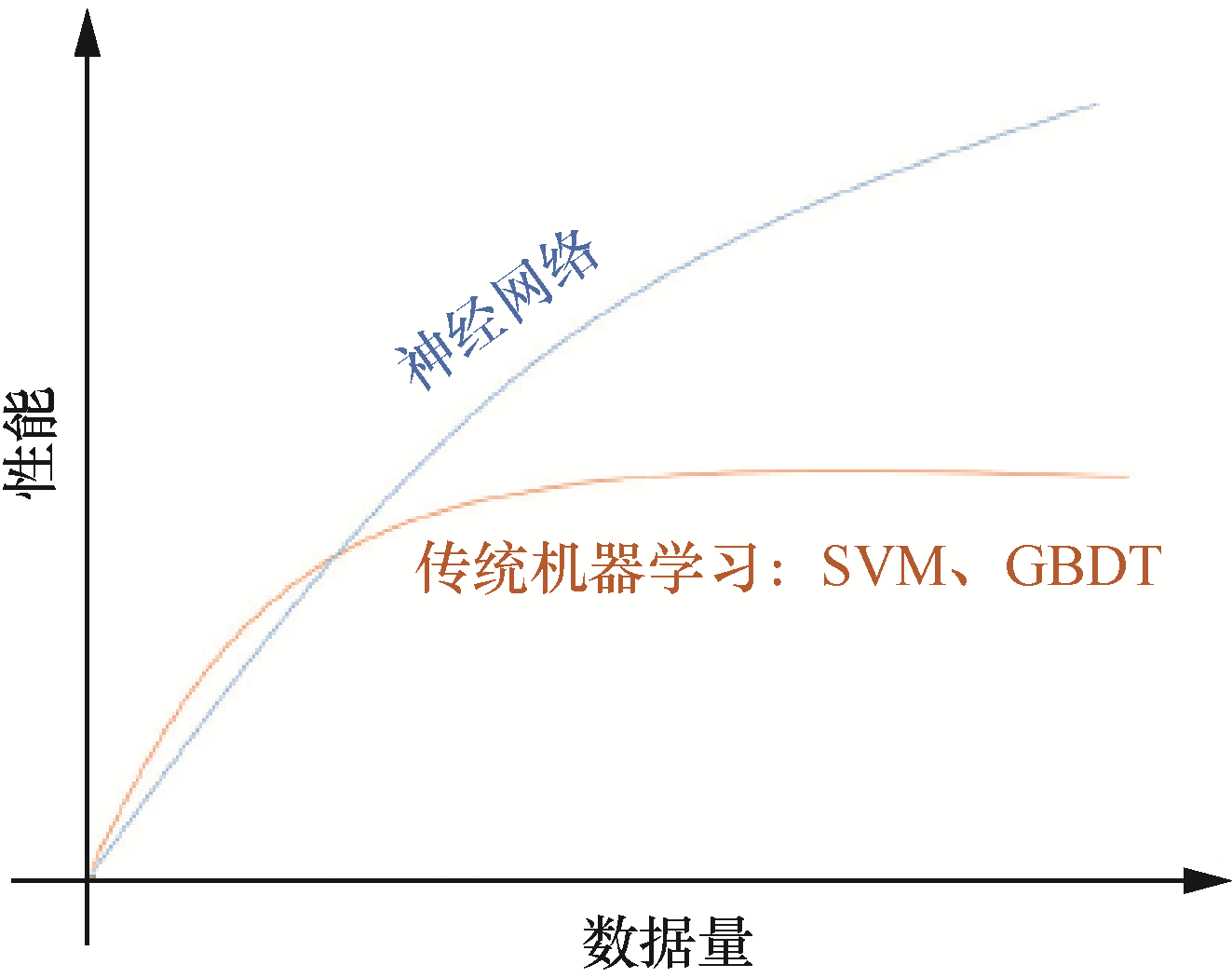


图5.1 数据量与机器学习模型性能的关系

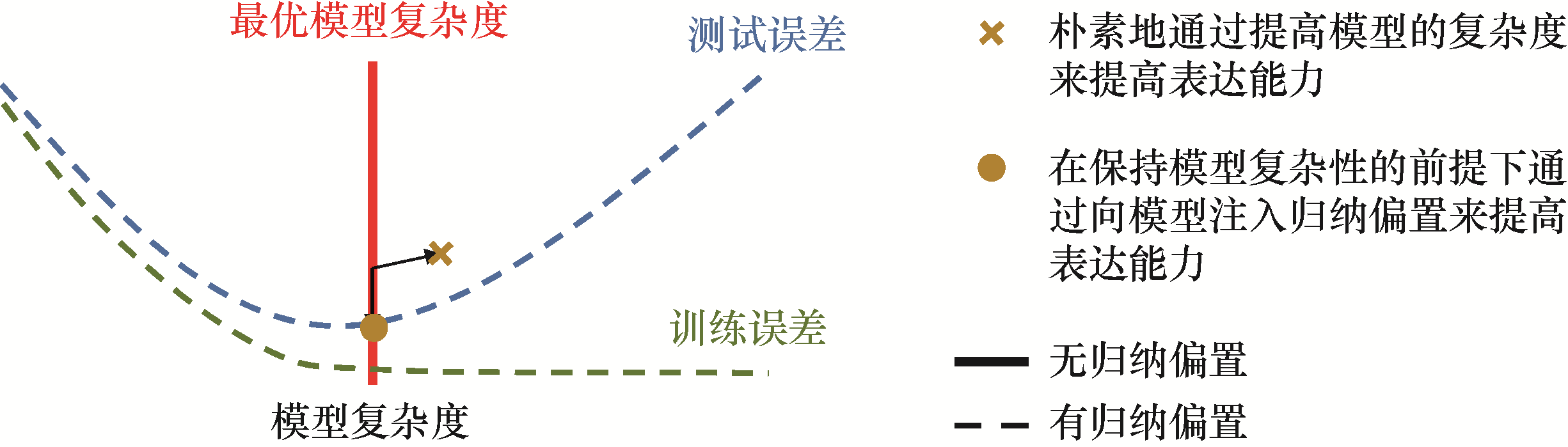


图5.2 有无归纳偏置的训练和测试误差会极大地影响模型的表达能力

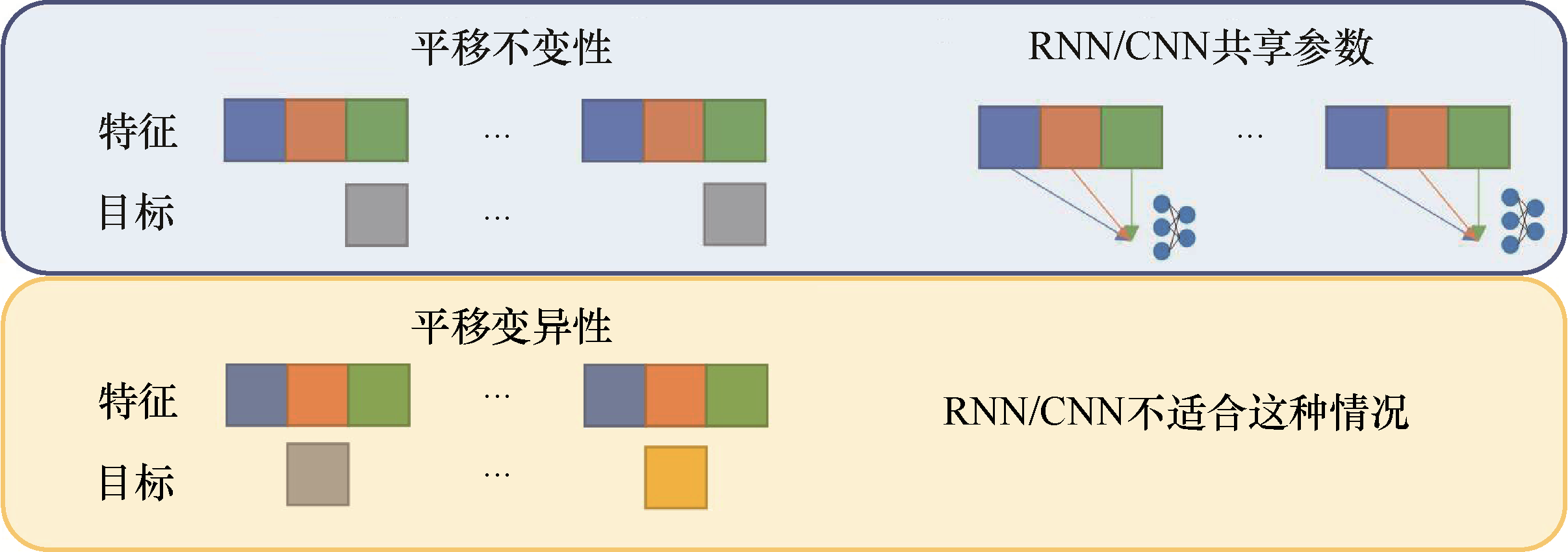


图5.3 一维的平移不变性和平移变异性（RNN/CNN利用平移不变性来共享参数）



图5.4 GNN如何保持排列不变性

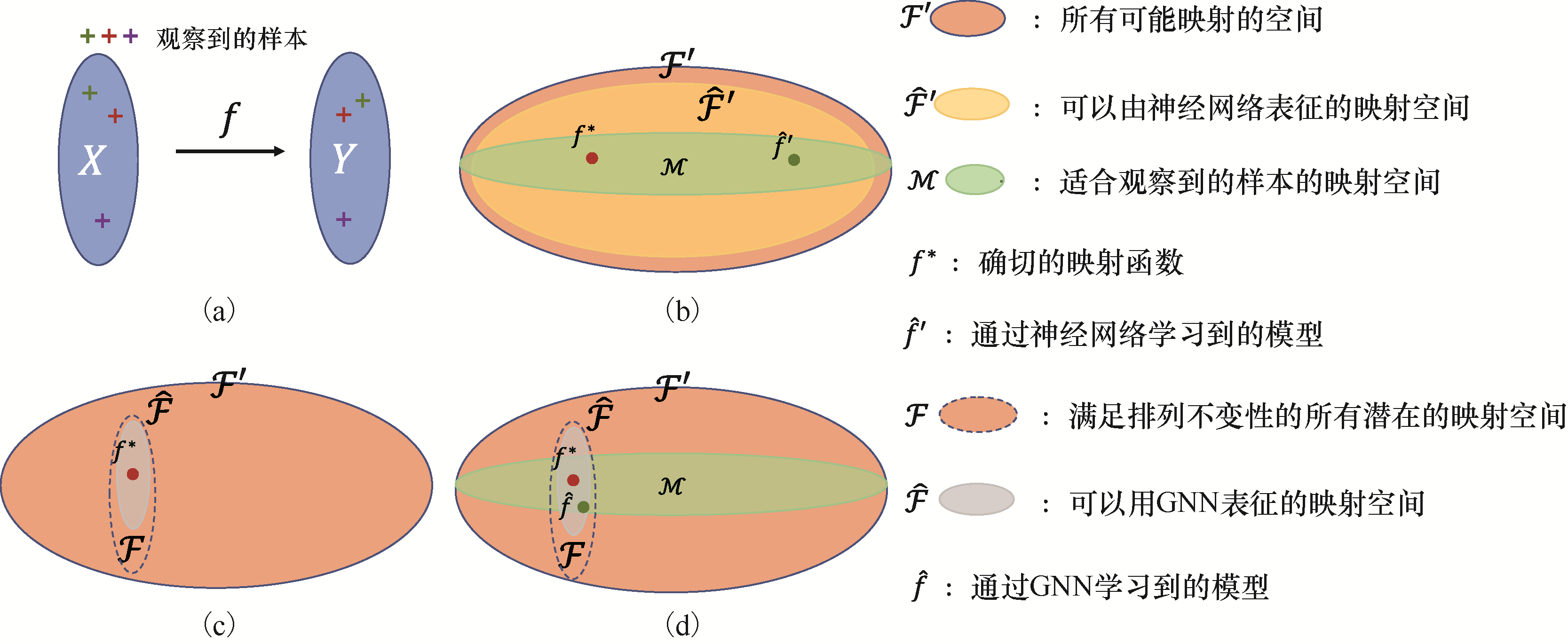


图5.5 神经网络和GNN的表达能力及其对学习模型表现的影响。（a）机器学习问题的目的是根据几个观察到的样本学习从特征空间到目标空间的映射。（b）神经网络的表达能力是指两个空间和之间的差距。尽管神经网络是有表达能力的（在中是密集的），但神经网络会对有限的观察数据过度拟合。  
（c）基于神经网络的学习模型*f'*可能与*f\**有很大的差别。潜在的映射函数的空间将从减少到一个小得多的空间，其中只包括排列不变的函数。如果采用GNN，则同时近似的映射函数的空间就会减少到。（d）尽管GNN的表达能力不如一般的神经网络（），但相对于基于神经网络的模型，基于GNN学习的模型*f*却是比*f\**好得多的近似。因此，对于图结构数据，我们对GNN的表达能力的理解，即和之间的差距，要比对神经网络的表达能力的理解更重要

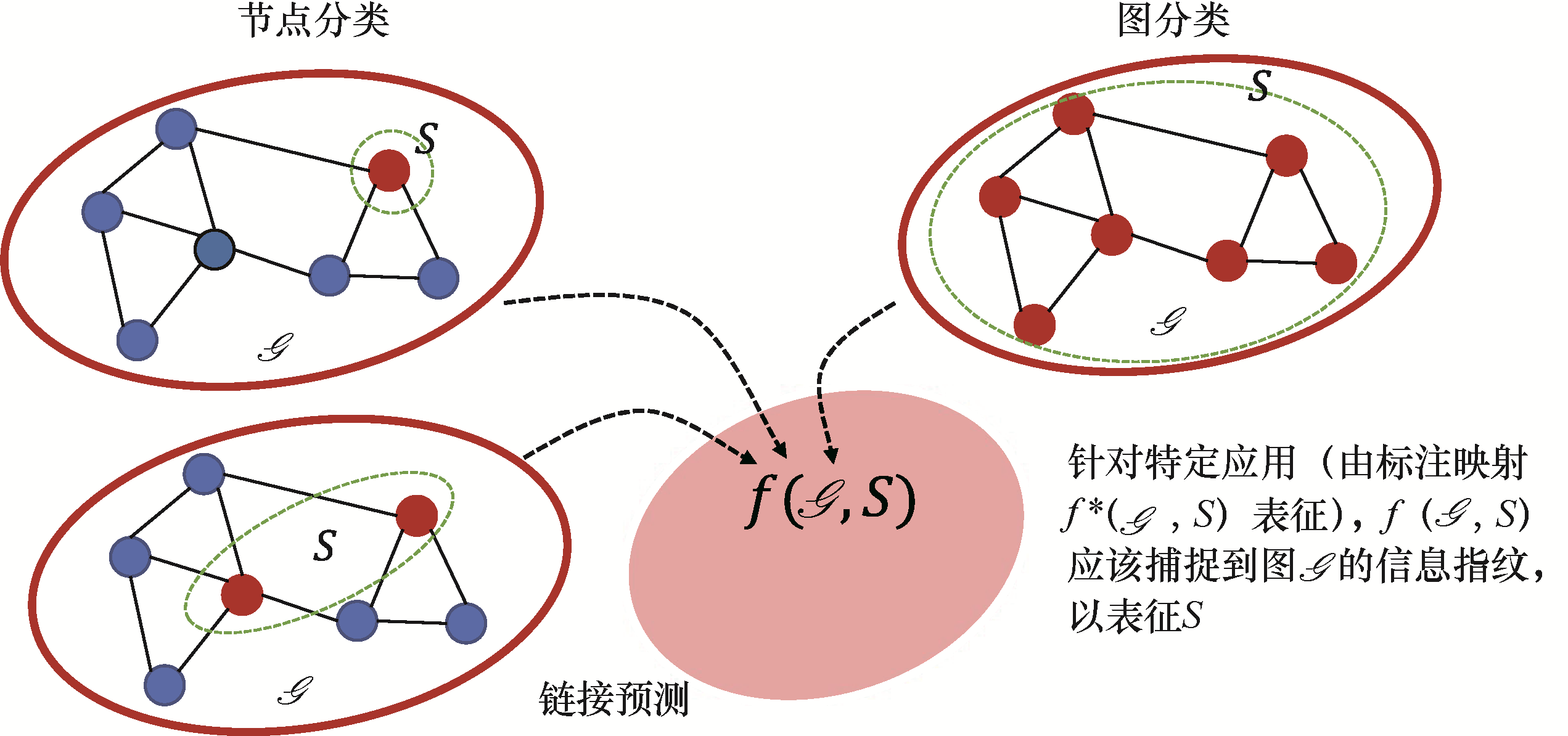


图5.6 近期的文献中经常讨论的图表征学习问题

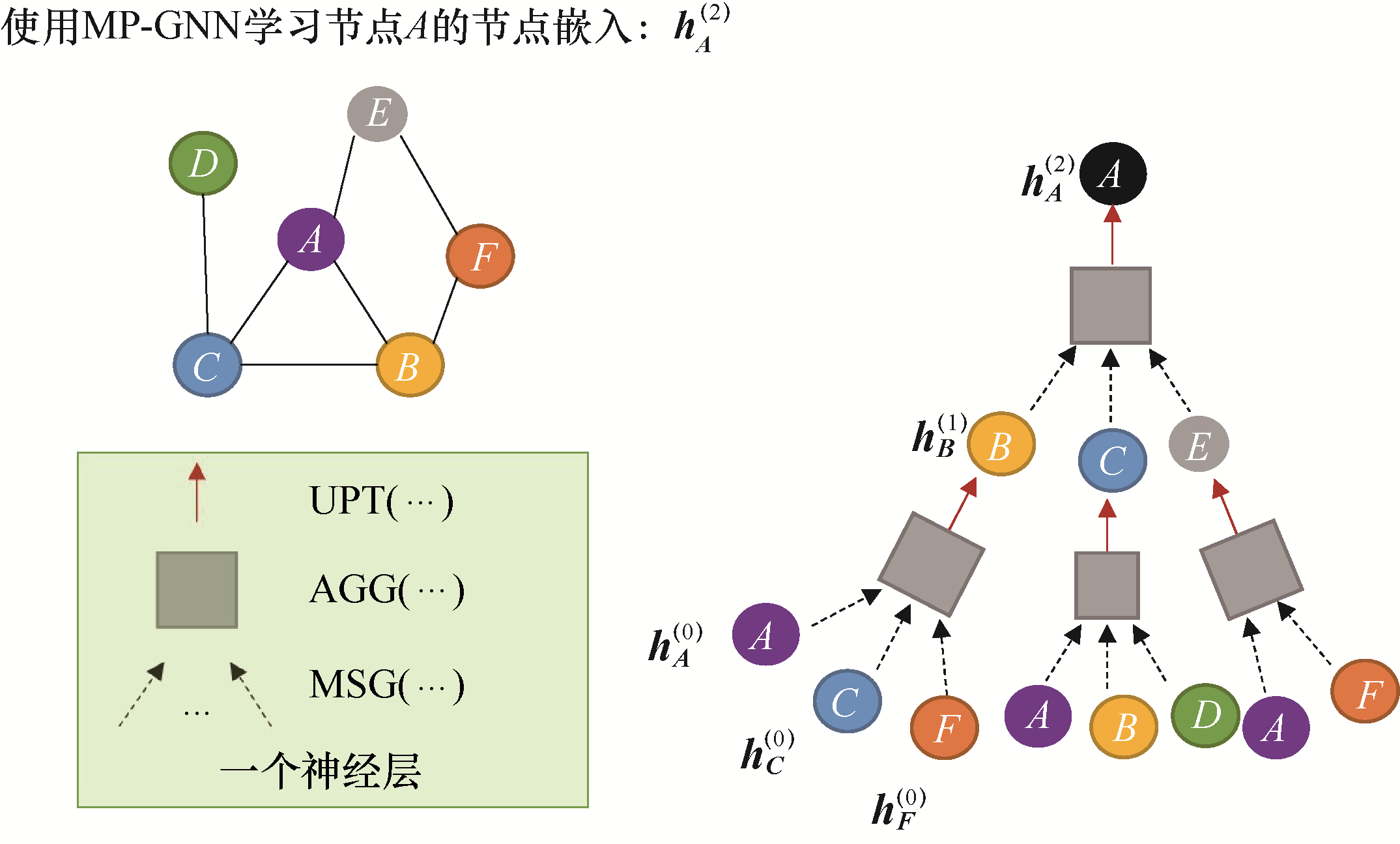


图5.7 使用MP-GNN计算节点表征的流程

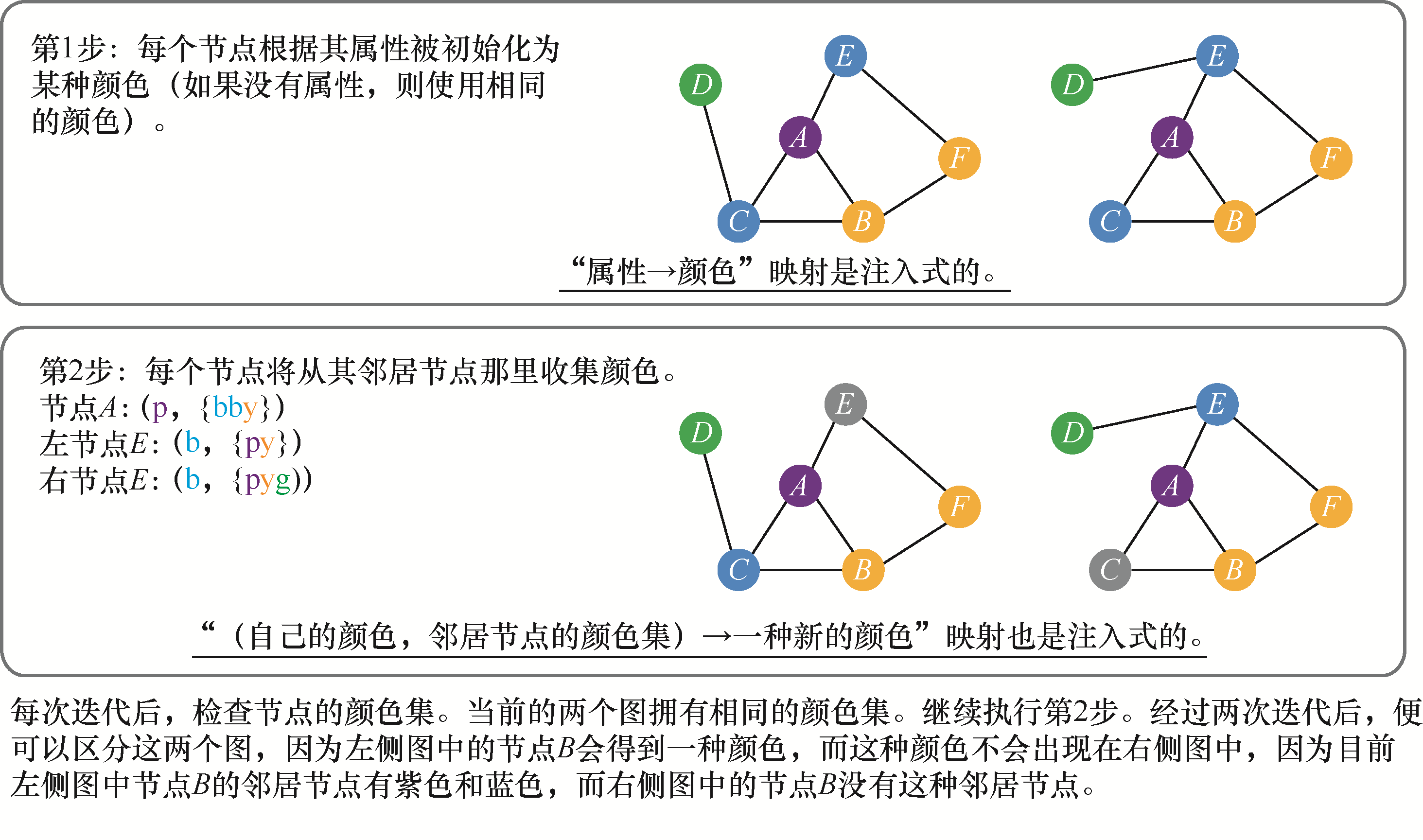


图5.8 通过1-WL测试区分两个图的步骤说明（MP-GNN通过遵循类似的步骤，也可以将两个图区分开）

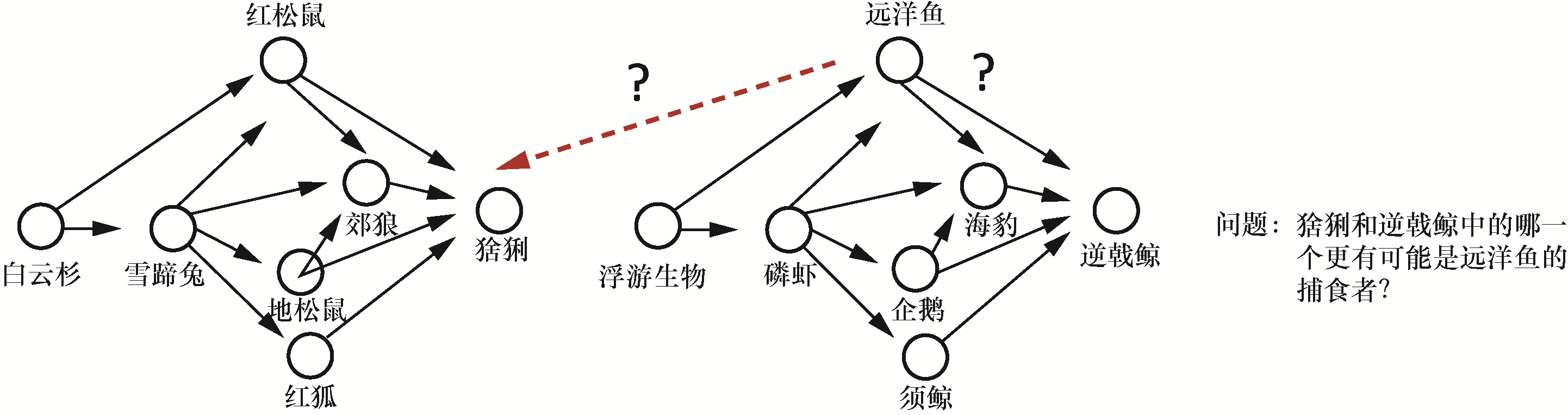


图5.9 用于展示MP-GNN局限性的食物链示例（Srinivasan and Ribeiro，2020a）。MP-GNN将猞猁和逆戟鲸（即虎鲸）关联到相同的节点表征，即，因为这两个节点具有相同的有根子树。请注意，我们不考虑节点特征。因此，MP-GNN不能预测到底是猞猁还是逆戟鲸更有可能成为远洋鱼的捕食者  
（这是一个链接预测任务）

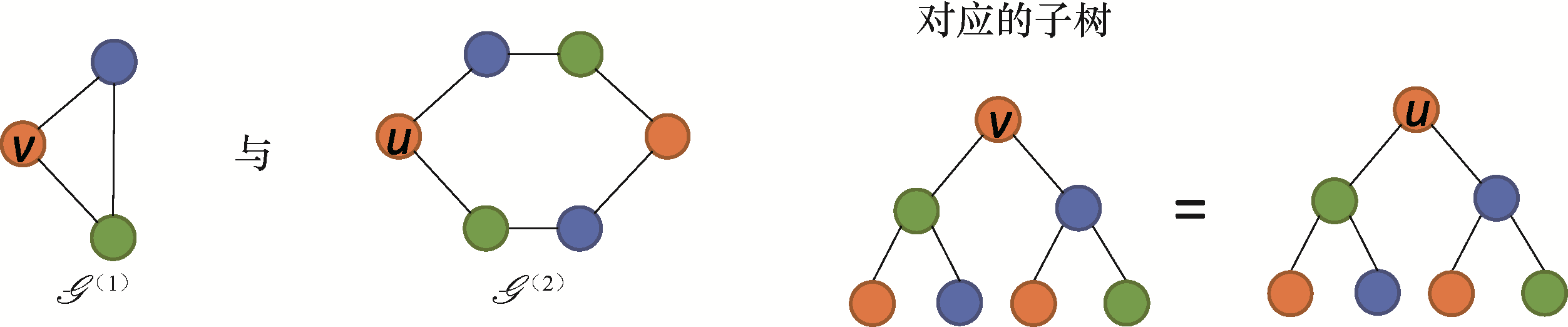


图5.10 MP-GNN给出的节点表征和是相同的，  
尽管它们分别属于不同的环—3节点环和6节点环

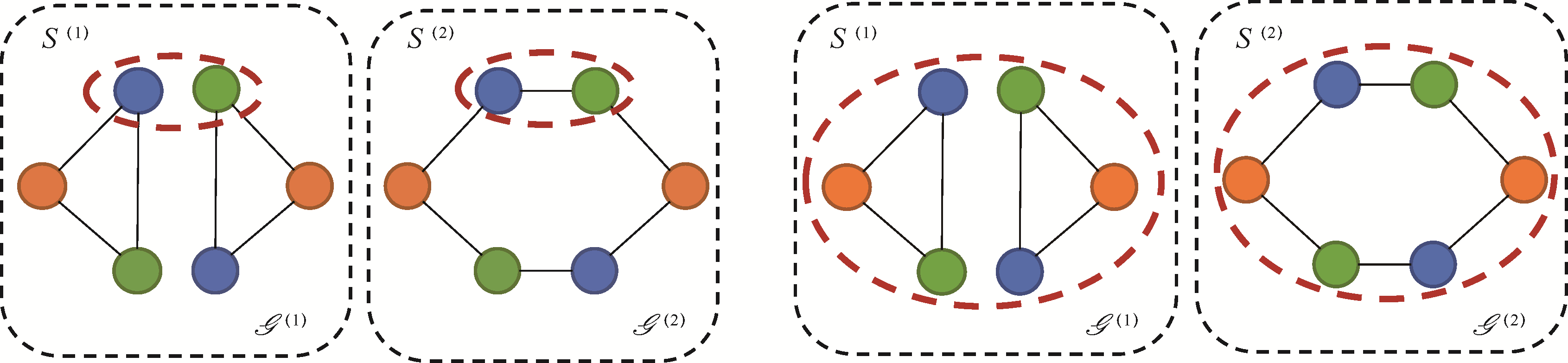


图5.11 一对配置相同的有属性正则图和以及适当的选择和：  
MP-GNN和1-WL测试无法区分和



图5.12 通过注入随机节点属性可以提高GNN的表达能力。不同的研究采用了不同类型的随机节点属性，  
一些随机节点属性包含节点位置信息（即一个节点相对于图中其他节点的位置）

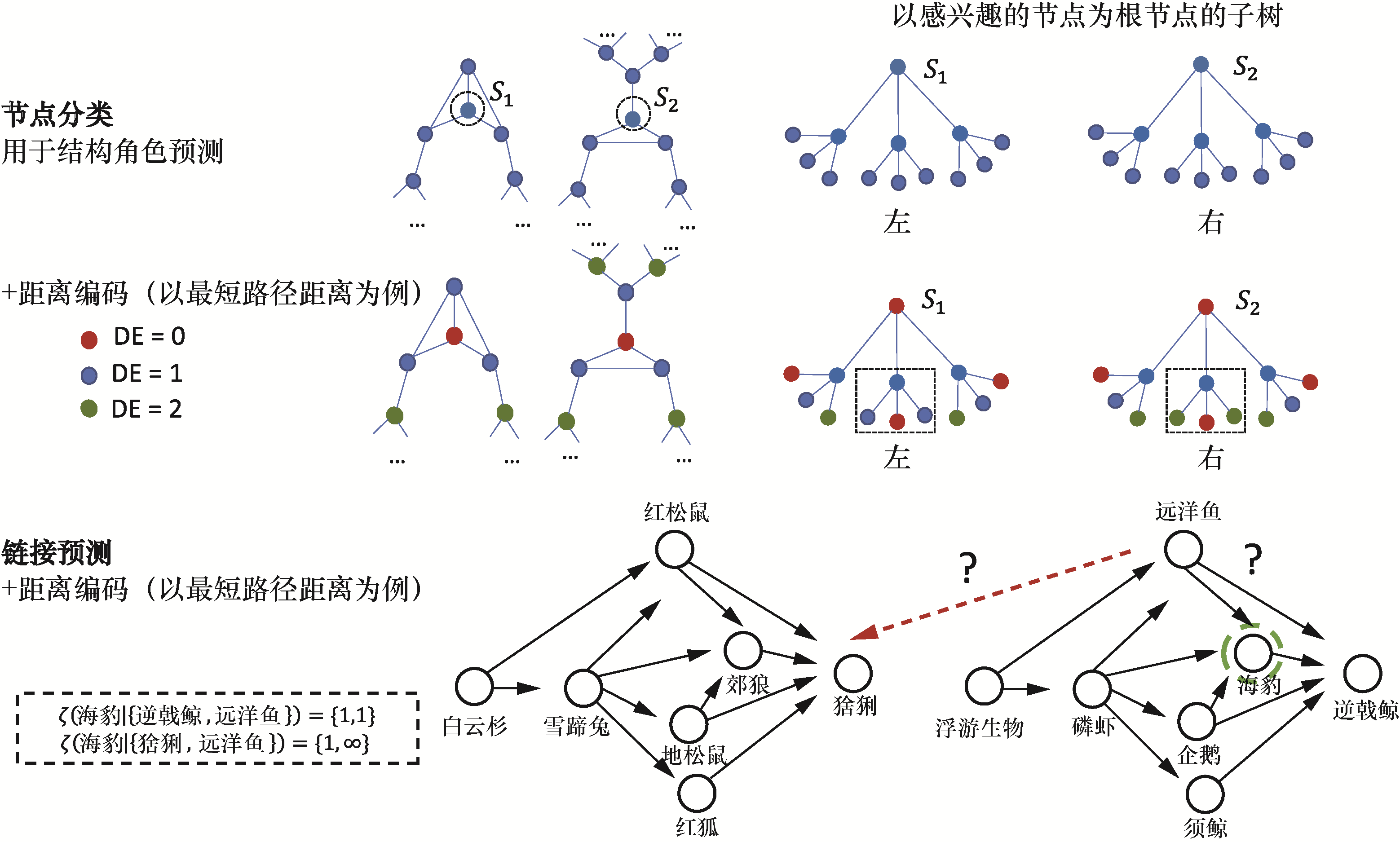


图5.13 距离编码可用于区分非同构图结构的示例。在节点分类的示例中，我们考虑根据节点在其上下文结构中的角色进行分类，称为结构角色（Henderson et al，2012）。*S*1和*S*2中的节点具有不同的结构角色。然而，具有两个层的MP-GNN会混淆这两个节点，而具有距离编码的DE-GNN却可以区分它们。在链接预测的示例中，尽管两个节点{猞猁, G }和{逆戟鲸, G }是同构的（在这里我们忽略节点的身份），但海豹节点上的距离编码使我们能够区分节点对{逆戟鲸, 远洋鱼}和{猞猁, 远洋鱼}



图5.14 使用ID-GNN与DE-GNN预测一对节点。左图是Shrikhande图，右图是4×4的Rook图。ID-GNN（为黑色节点附加了身份属性）不能区分节点对{*a*, *b*}和{*c*, *d*}。DE-GNN可以学习节点对{*a*, *b*}和{*c*, *d*}的不同表征。在这两个子图中，每个节点都用它的DE来着色，DE是目标节点对{*a*, *b*}或{*c*, *d*}中任一节点的SPD集合。请注意，DE={1, 1}的节点（虚线框）被红色的椭圆包围，这表明这两个节点的邻居节点具有不同的DE。因此，经过一层后，这两个子图中DE={1, 1}的节点的中间表征是不同的，利用另一层，DE-GNN  
可以区分出节点对{*a*, *b*}和{*c*, *d*}

图5.15 使用*k*-FLW测试和*k*-WL测试区分和的过程

# 第6章　图神经网络的可扩展性

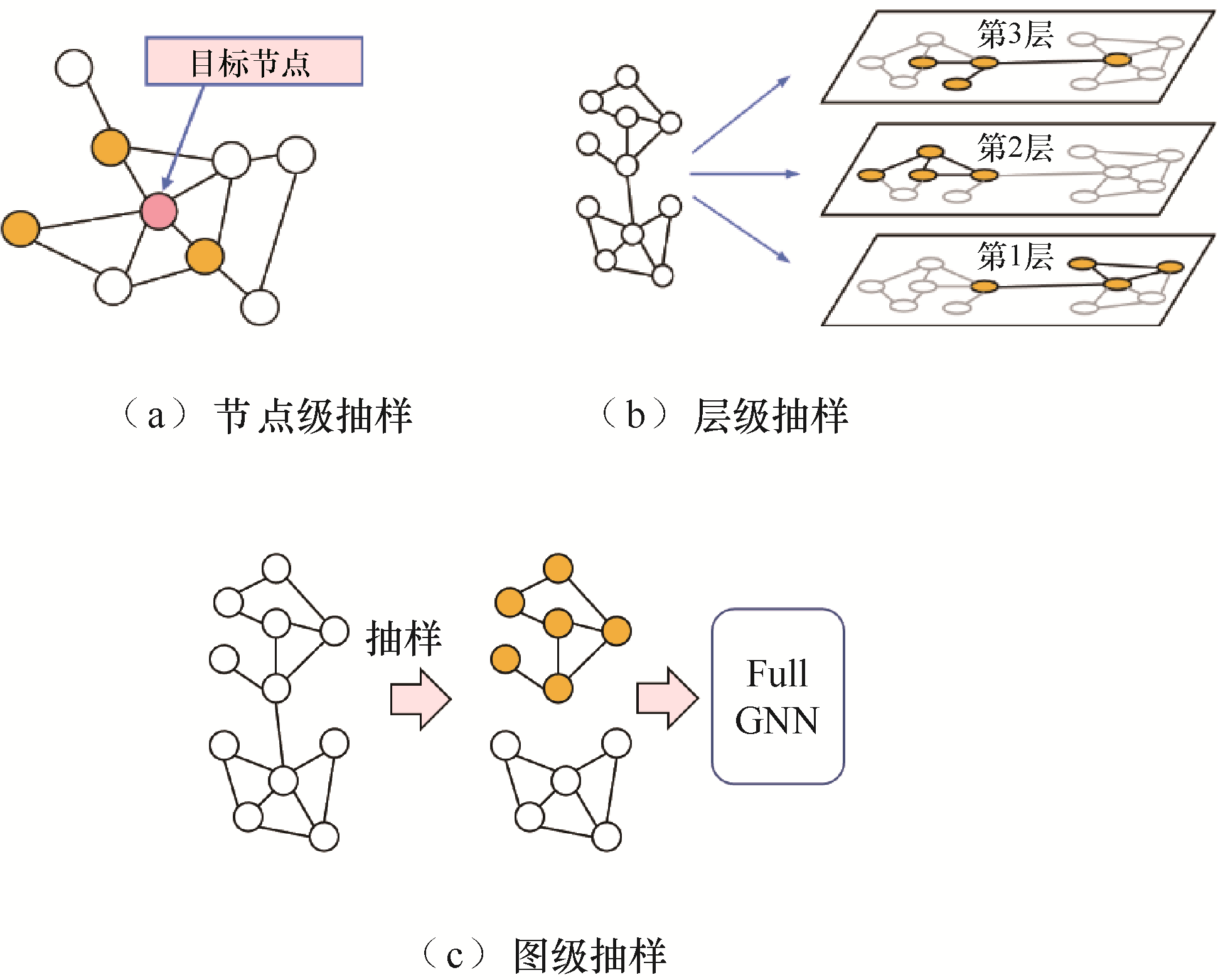


图6.1 面向大规模图的GNN的三种抽样范式

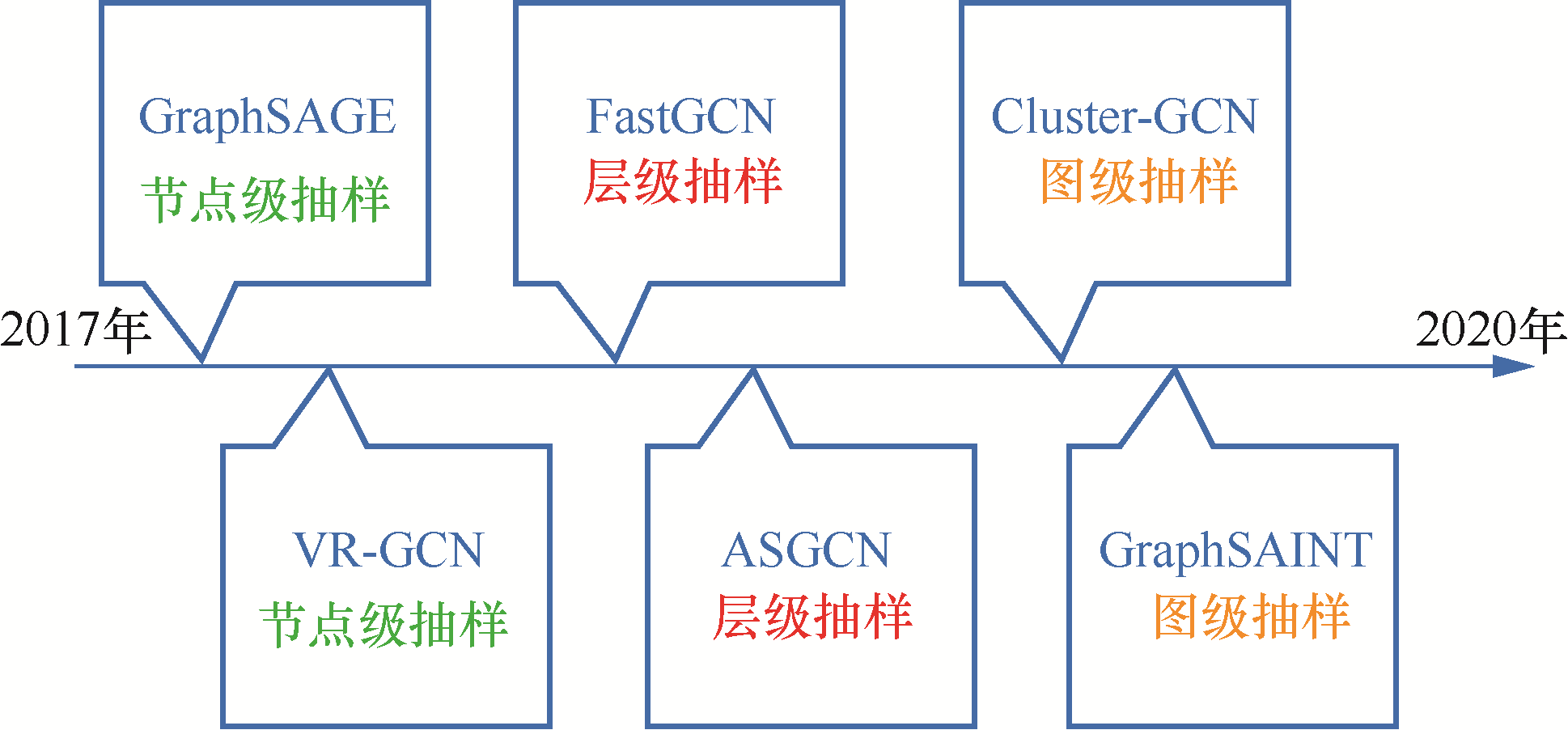


图6.2 面向大规模图的GNN的代表性研究的时间线

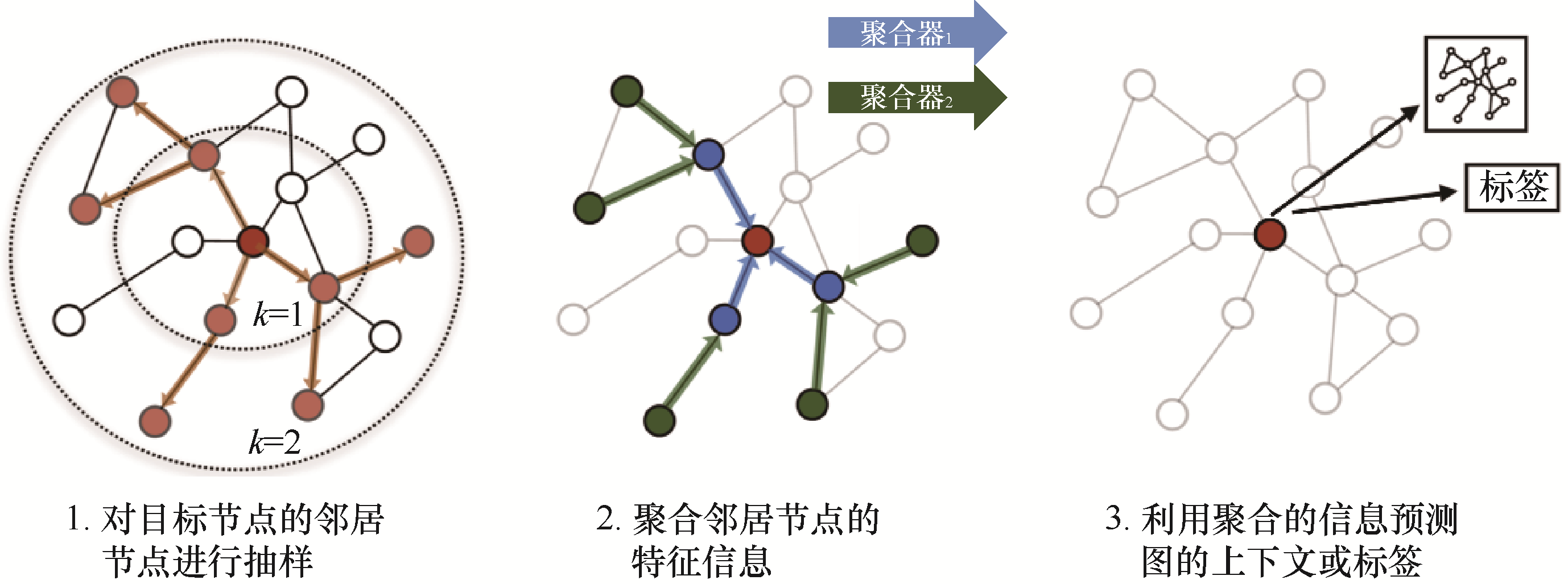


图6.3 GraphSAGE架构概述，摘自（Hamilton et al，2017b）

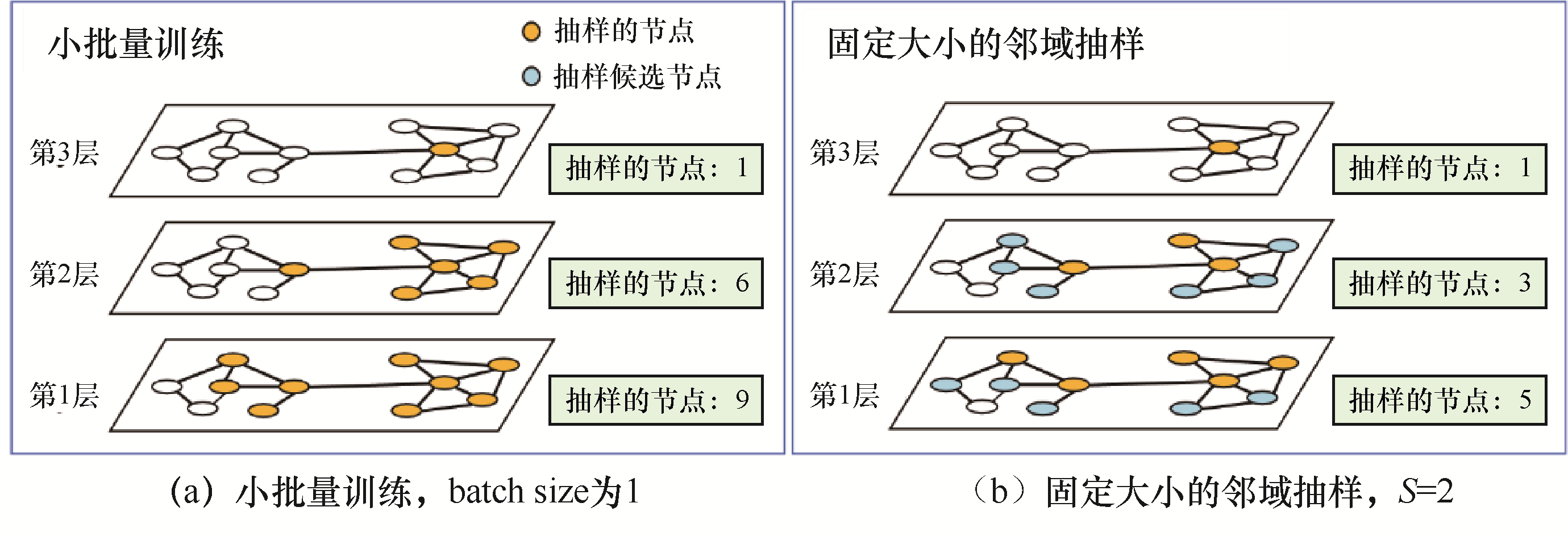


图6.4 比较小批量训练和固定大小的邻域抽样

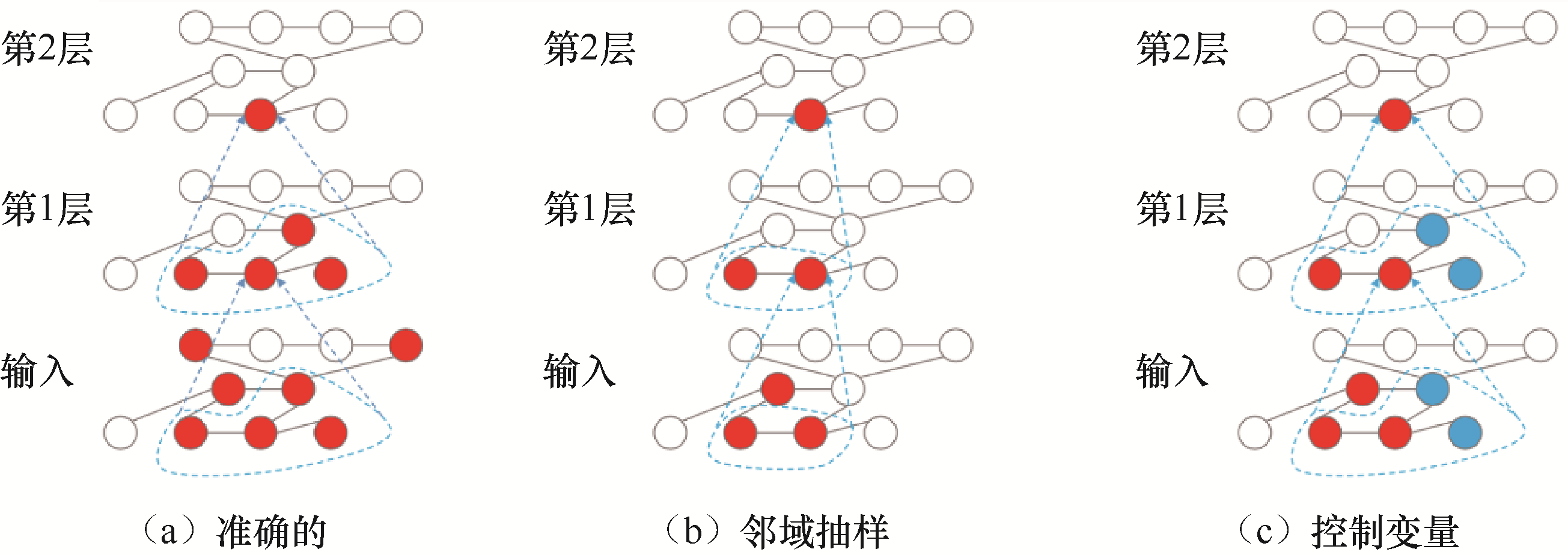


图6.5 比较对两层的图卷积神经网络使用不同抽样策略时目标节点的感受野。  
红色圆圈表示最新的激活，蓝色圆圈表示历史激活。摘自（Chen et al，2018d）

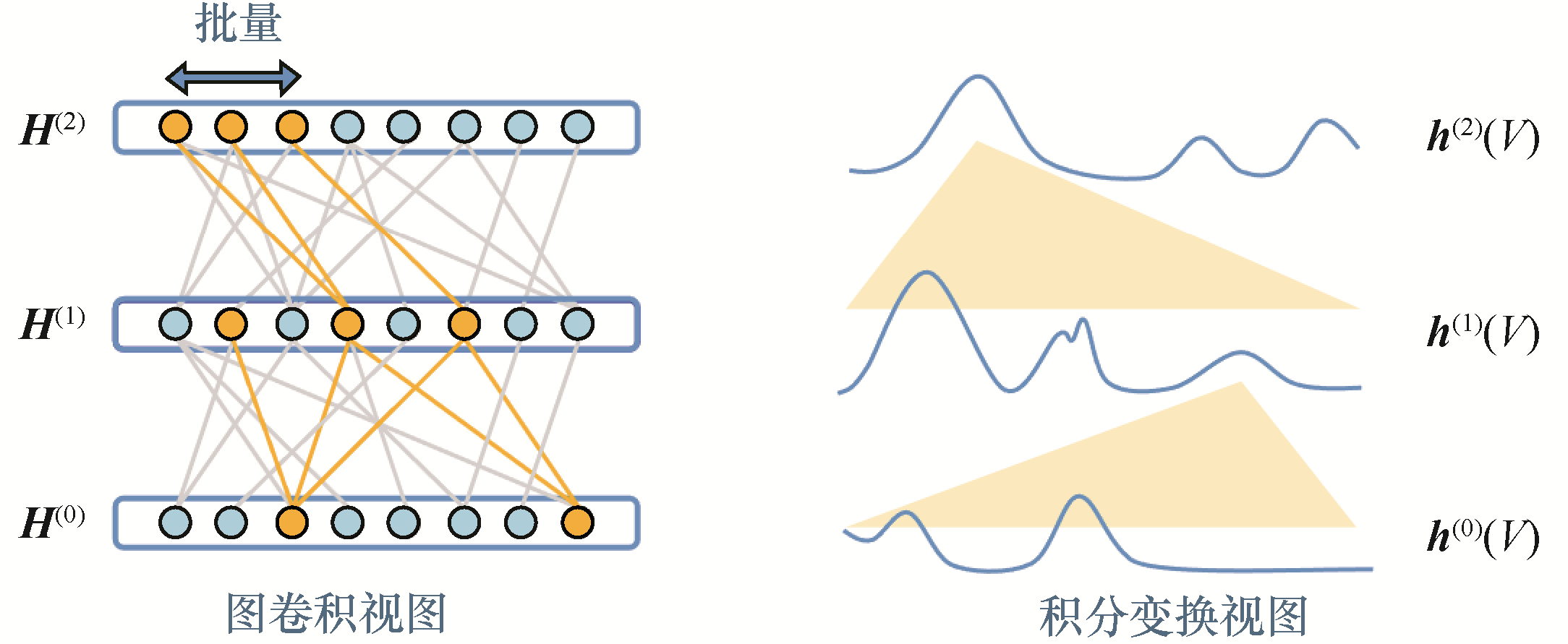


图6.6 GCN的两个视图：圆圈表示图中的节点，黄色的圆圈表示抽样的节点，线表示节点之间的连接

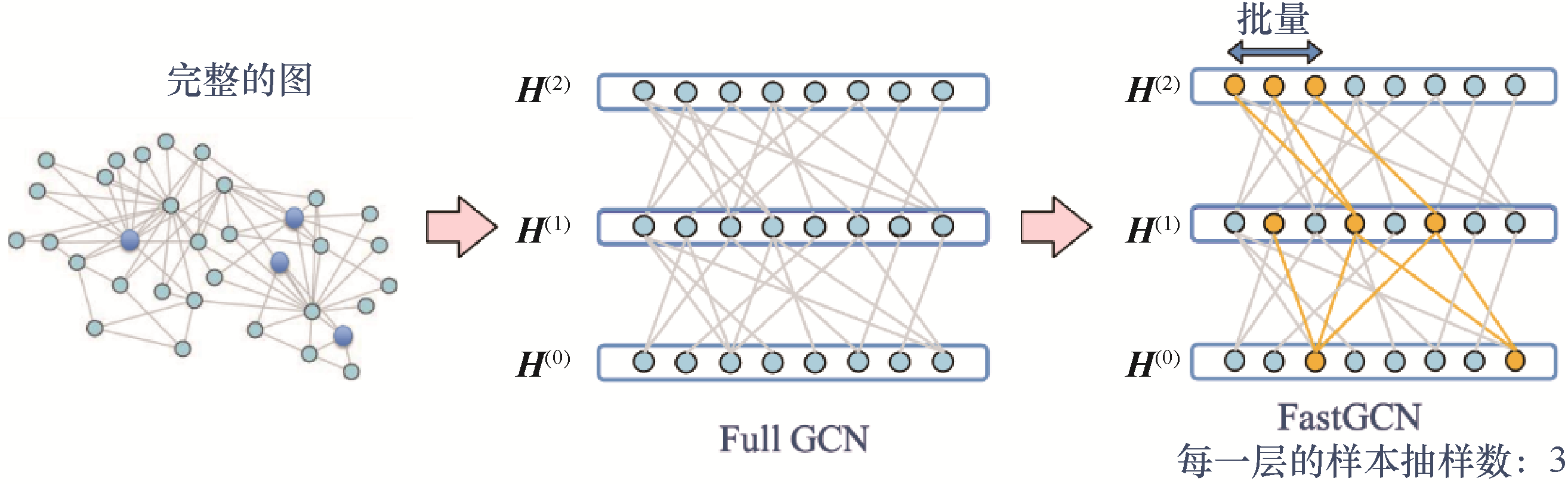


图6.7 比较Full GCN和FastGCN的抽样差异

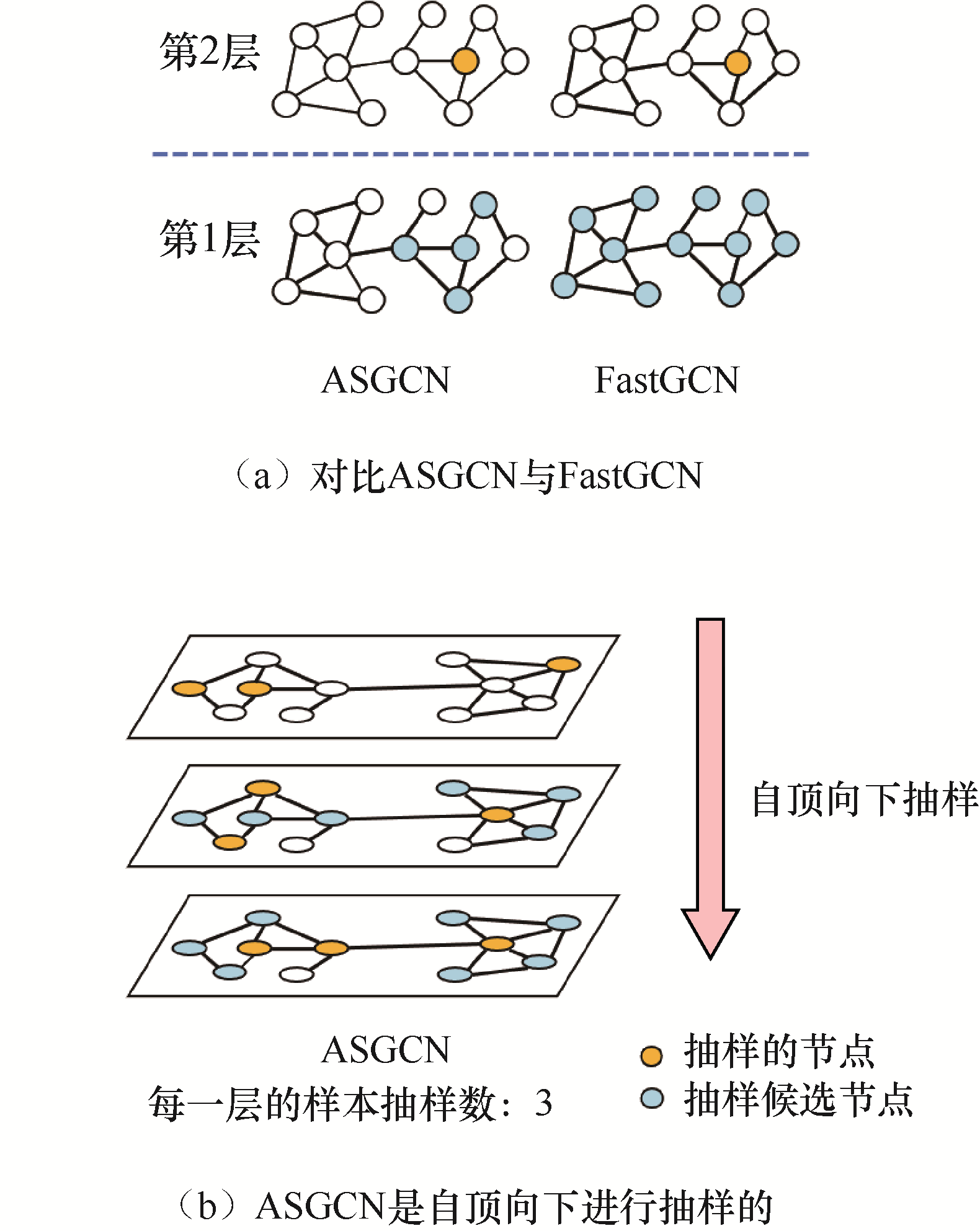


图6.8 ASGCN采用的抽样策略

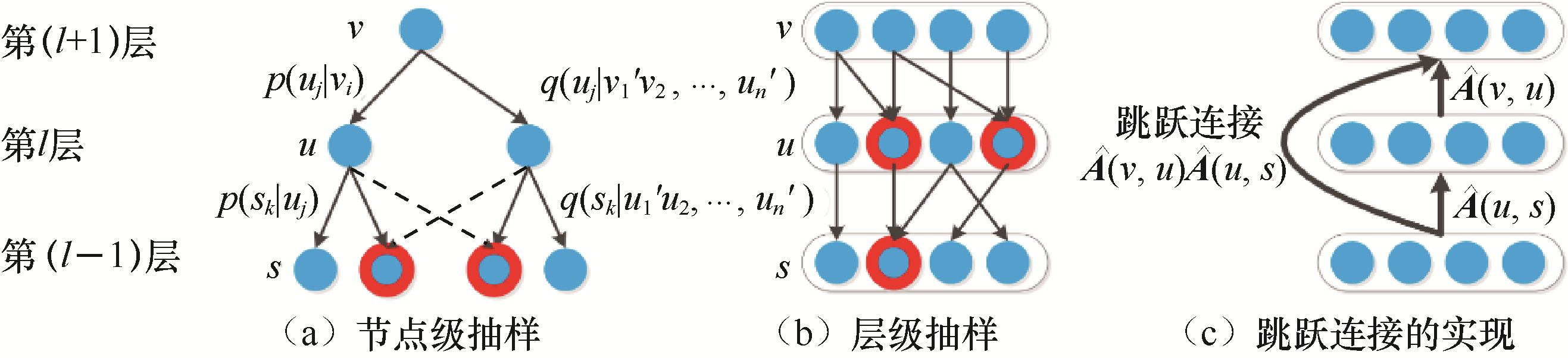


图6.9 网络构建示例。摘自（Huang et al，2018）

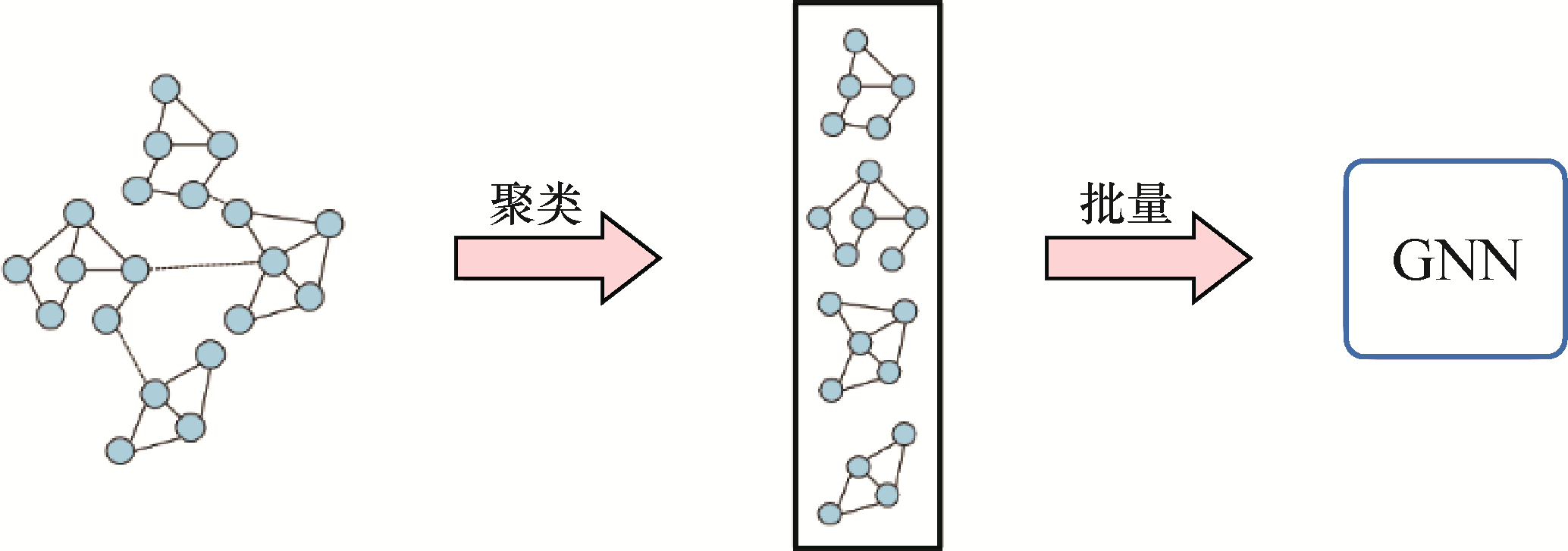


图6.10 大规模图上的图级抽样

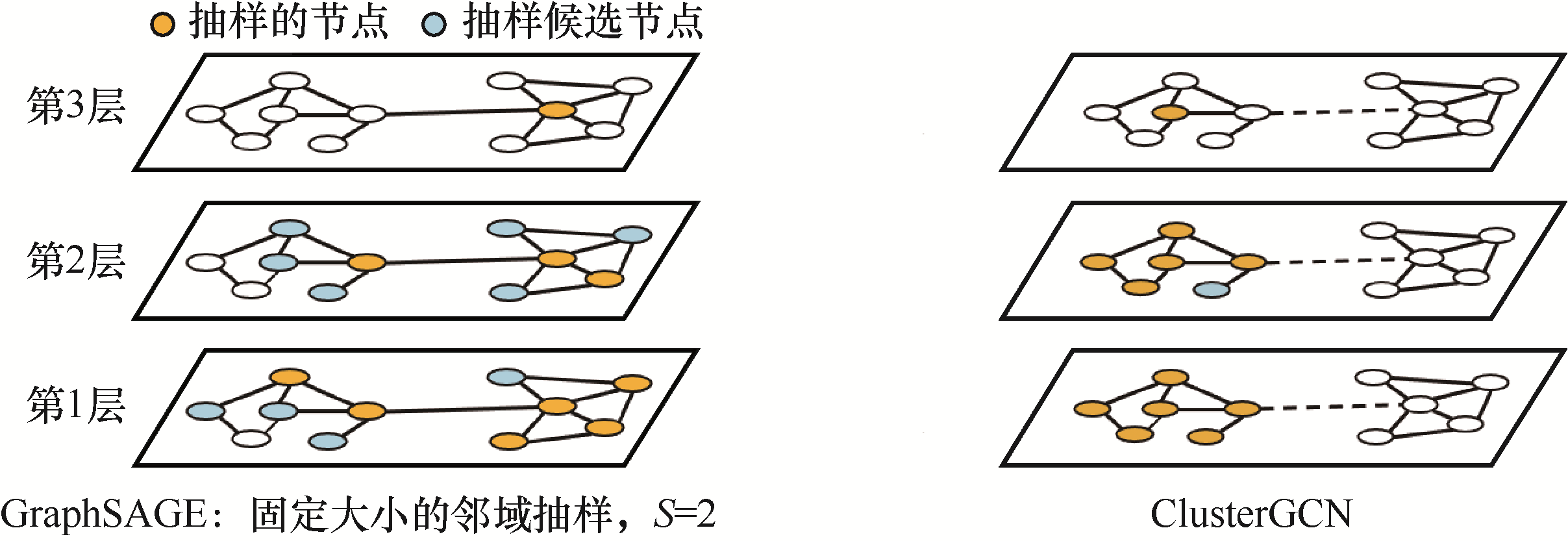


图6.11 比较GraphSAGE和Cluster-GCN，Cluster-GCN只对每个子图中的节点进行抽样

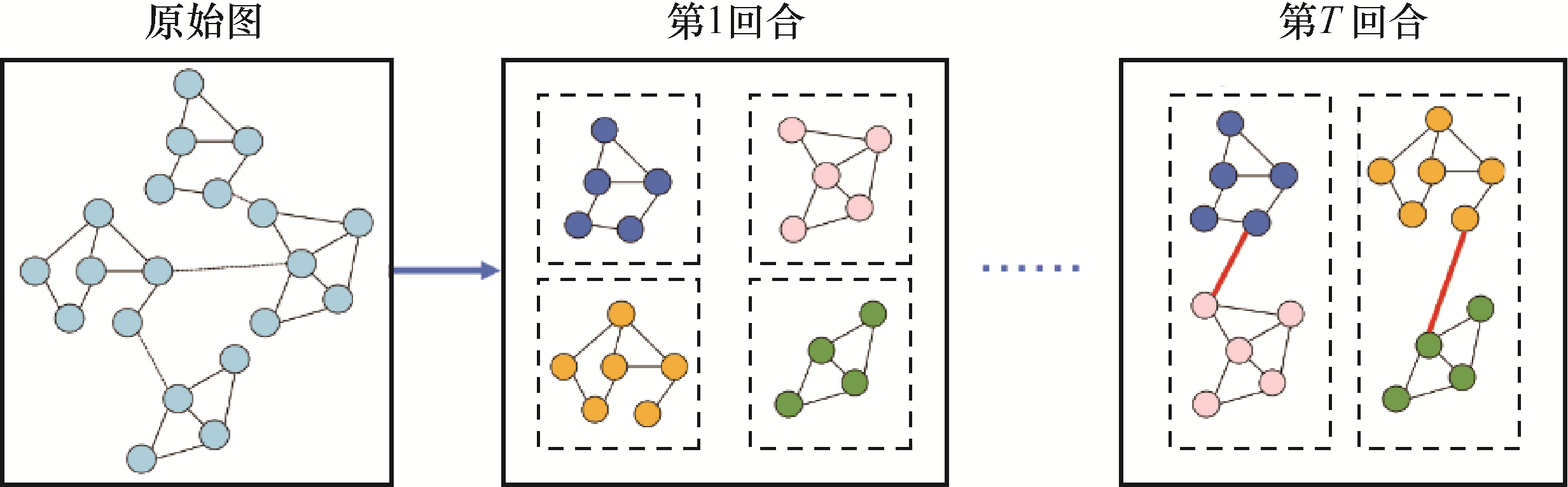


图6.12 随机多划分方案

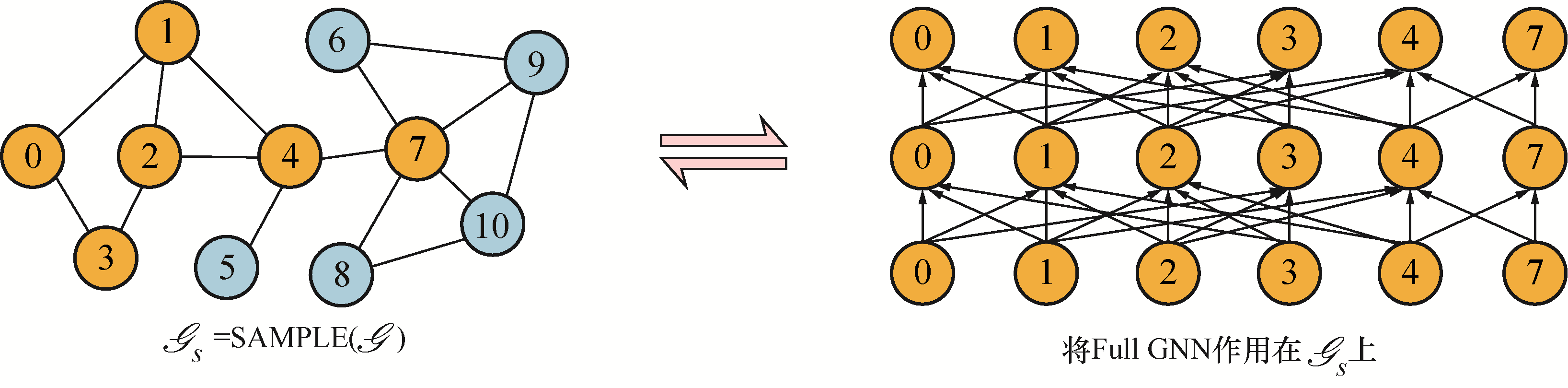


图6.13 GraphSAINT训练算法的示意图，黄色的圆圈表示抽样的节点

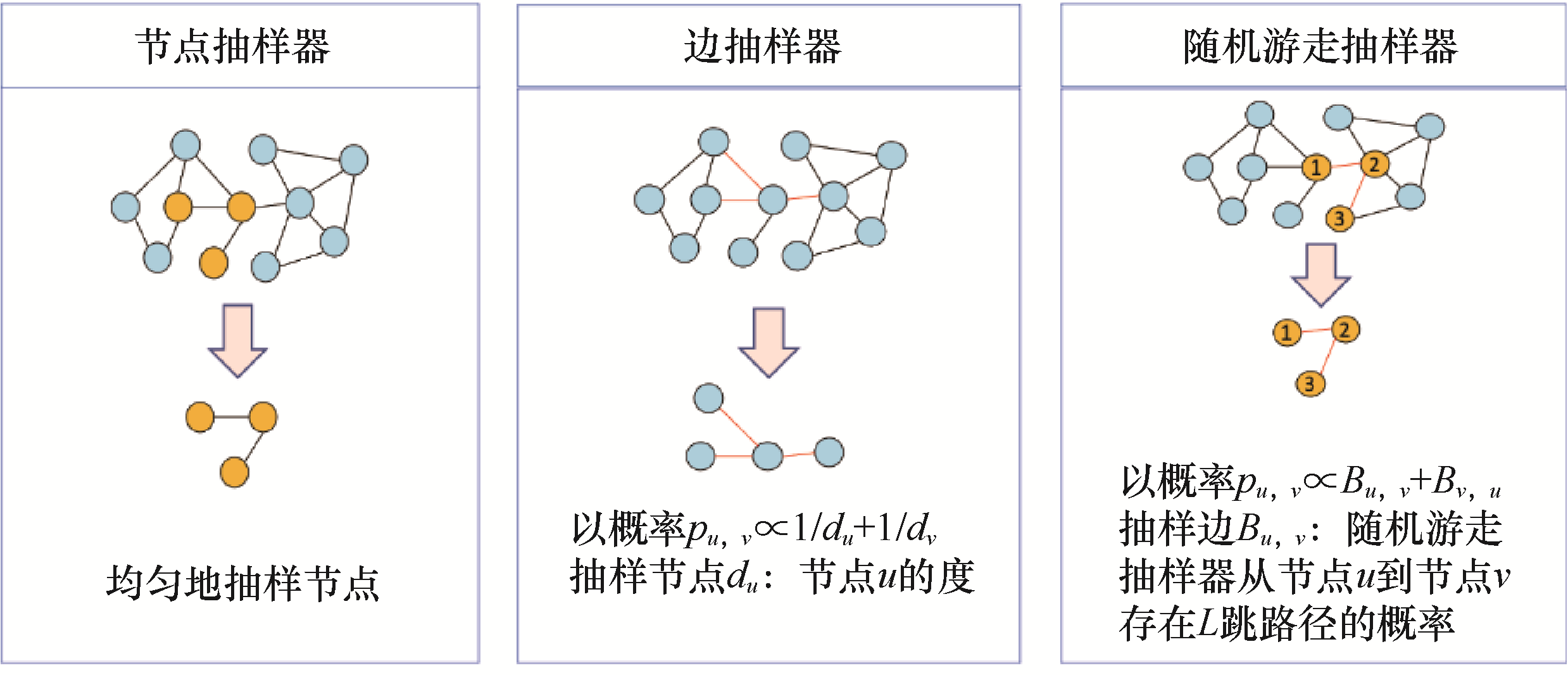


图6.14 三种子图抽样器

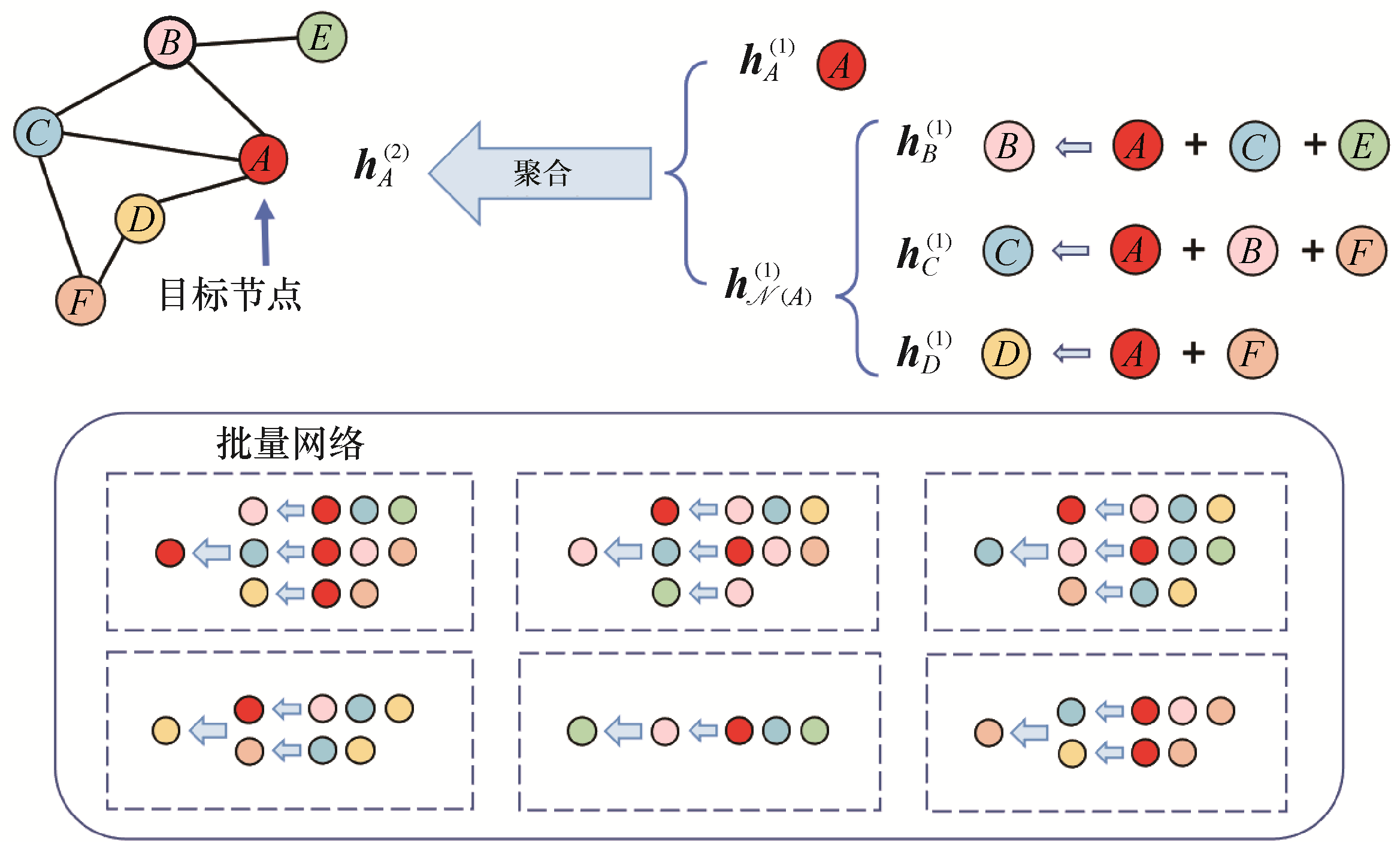


图6.15 PinSage架构通过节点的不同颜色来说明图卷积的构造

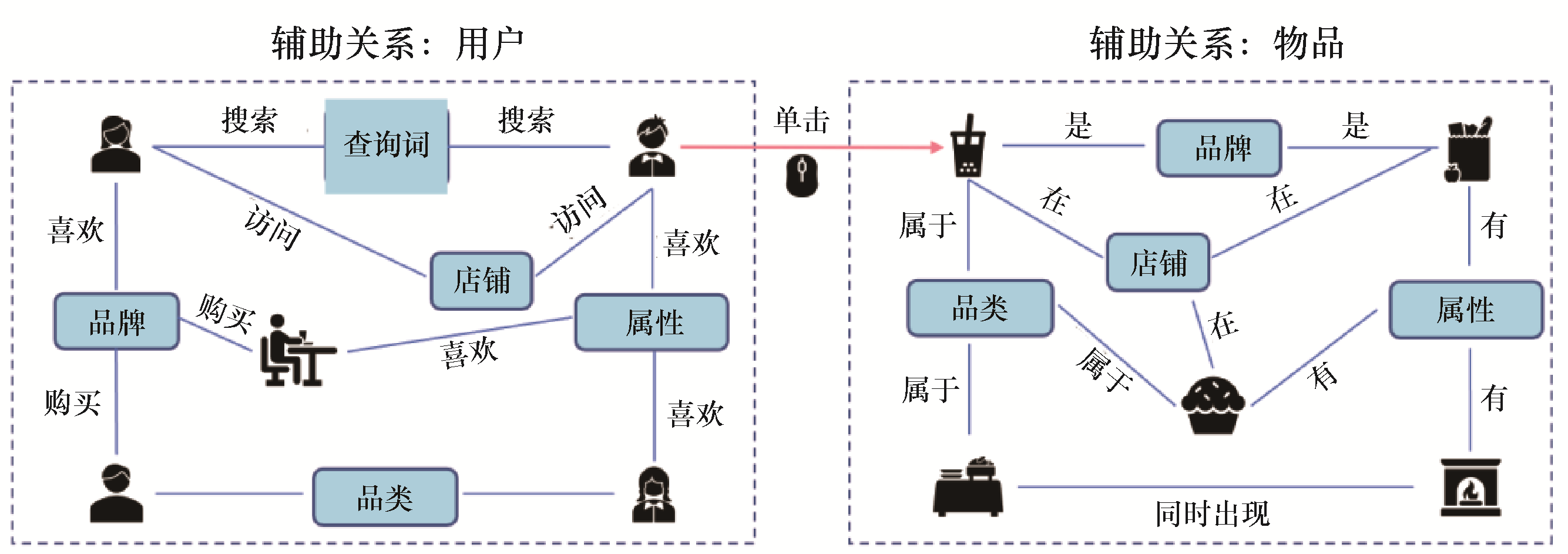


图6.16 电子商务网站上的异质性辅助关系的示例

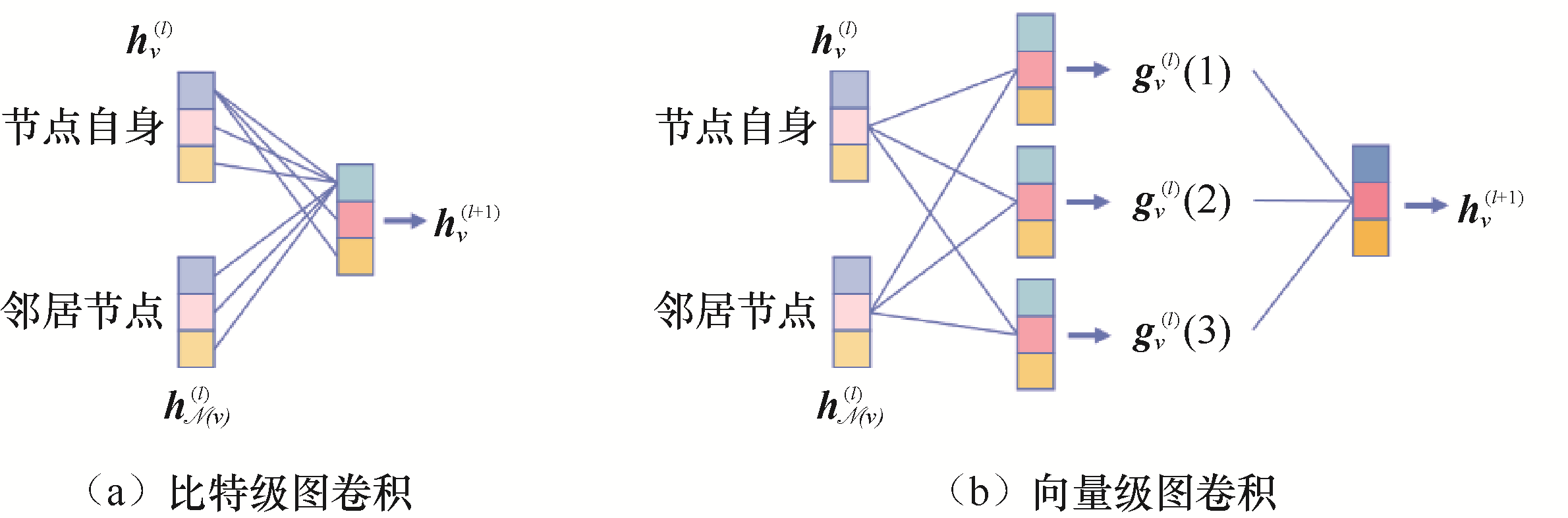


图6.17 比特级和向量级图卷积的比较

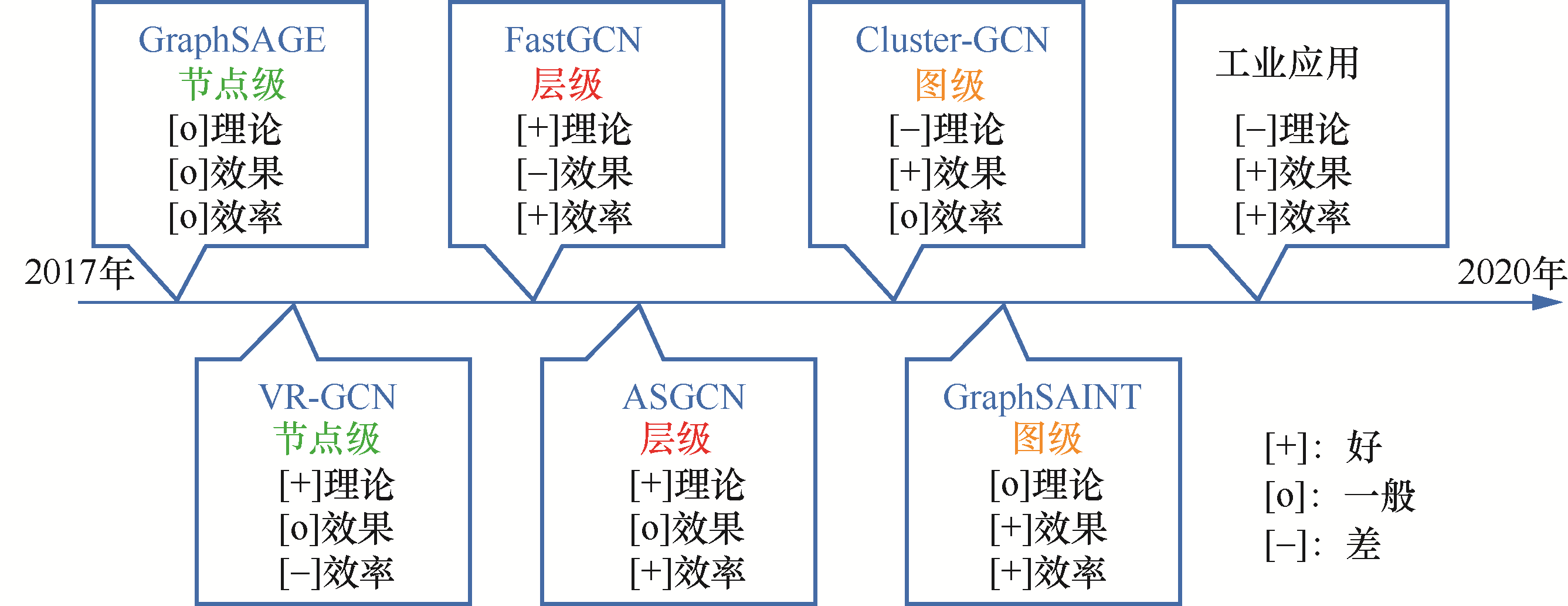


图6.18 大规模图神经网络的研究进展

# 第7章　图神经网络的可解释性

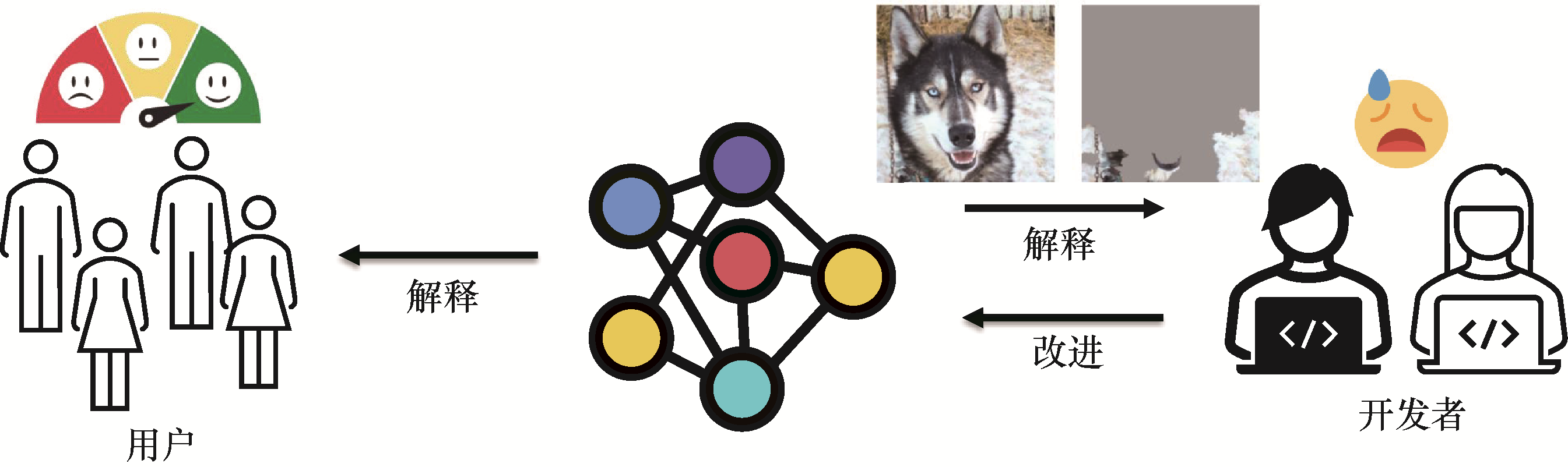


图7.1 左图：解释可以使用户在与模型的交互中受益。右图：通过解释，我们可以识别出  
人类不希望看到的模型行为，并致力于改进模型（Ribeiro et al，2016）

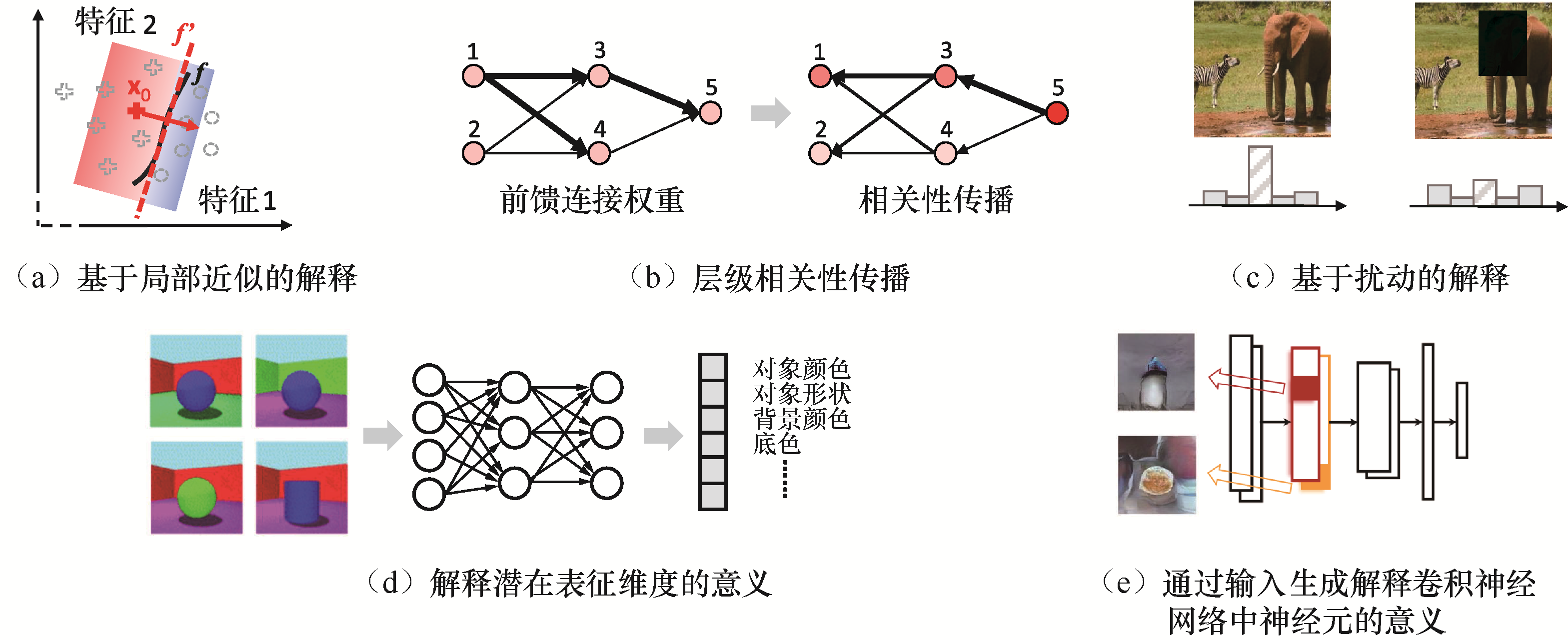


图7.2 关于事后解释方法的说明

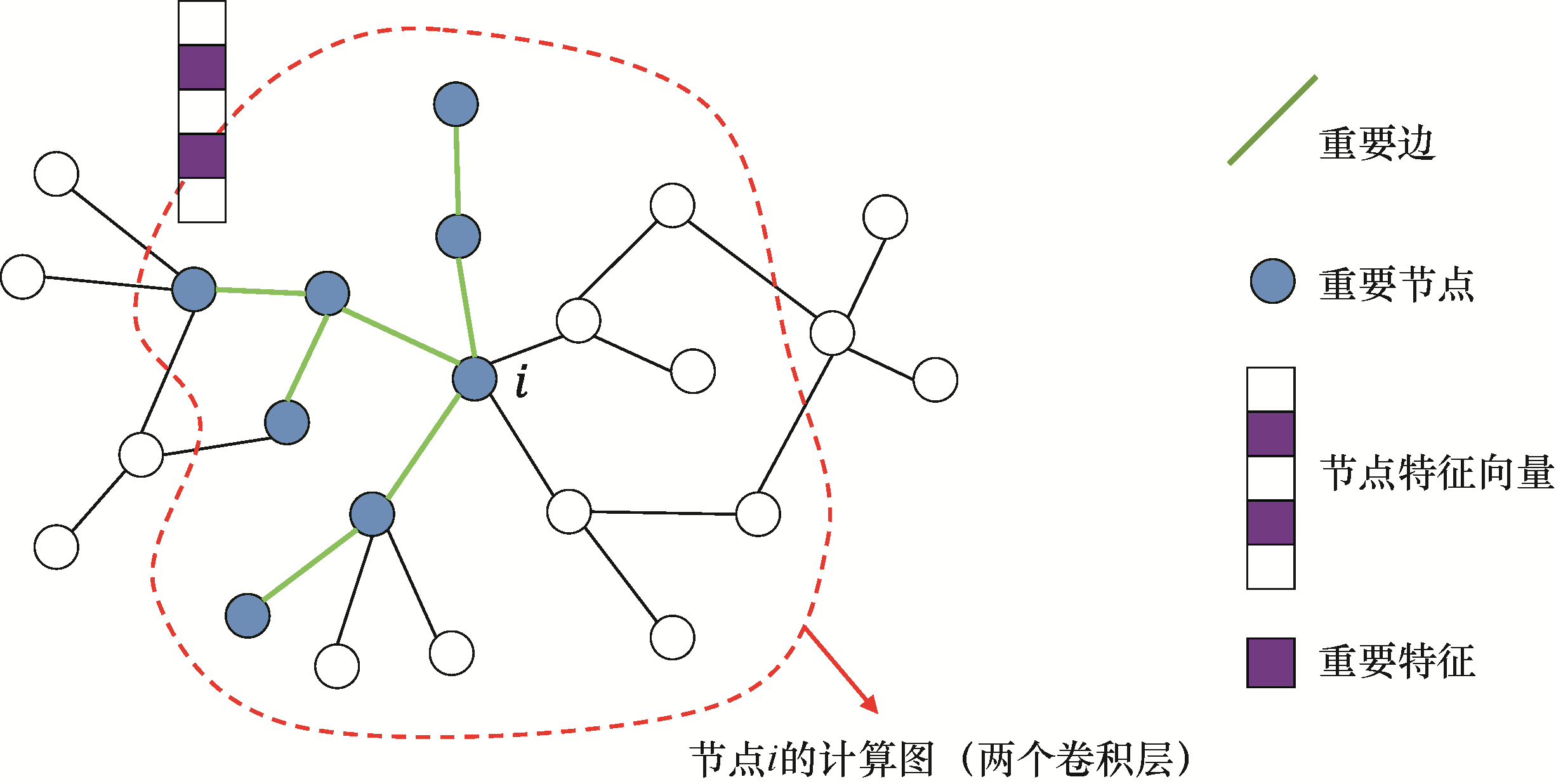


图7.3 解释结果的格式。图神经网络的解释结果可以是重要节点、  
重要边、重要特征等。解释方法可以返回多种类型的结果

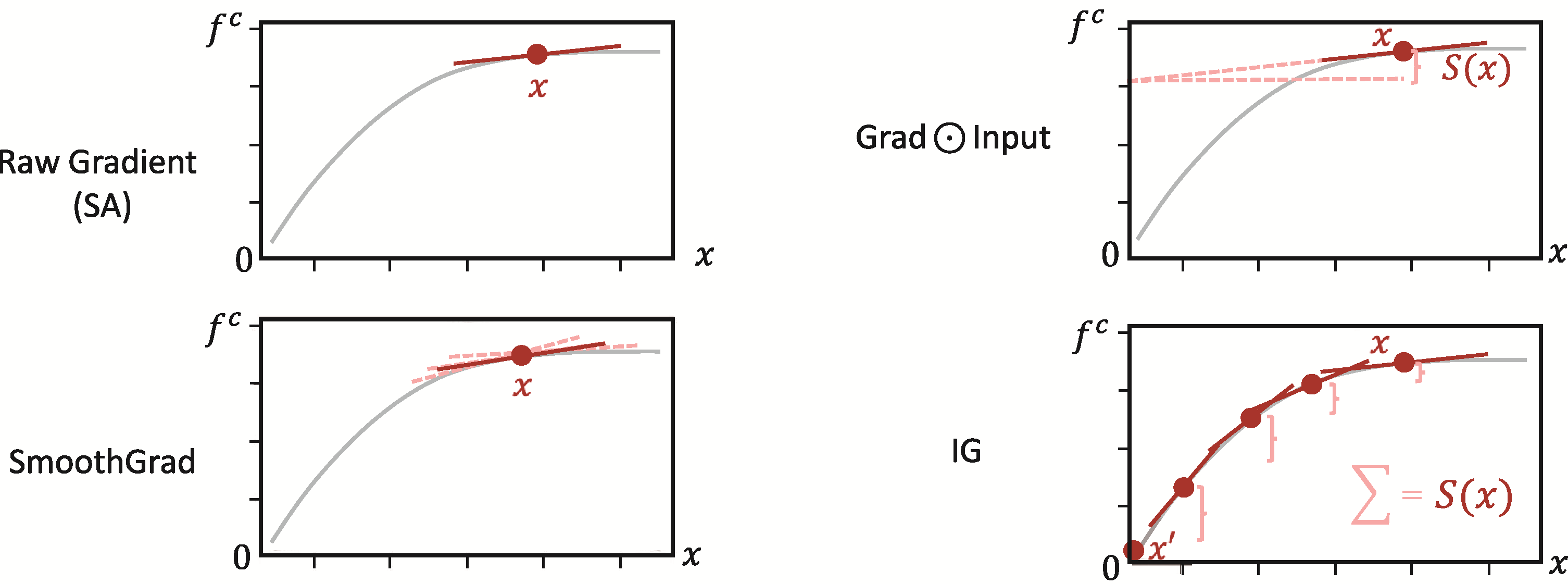


图7.4 几种基于梯度的解释方法。依靠局部梯度的解释方法可能会受到饱和问题或输入噪声的影响，  
换言之，一个特征的局部敏感性与它产生的整体贡献并不一致。SmoothGrad通过平均化附近点的  
多个解释来消除解释中的噪声。在测量特征贡献方面，IG相比Grad⊙Input更准确

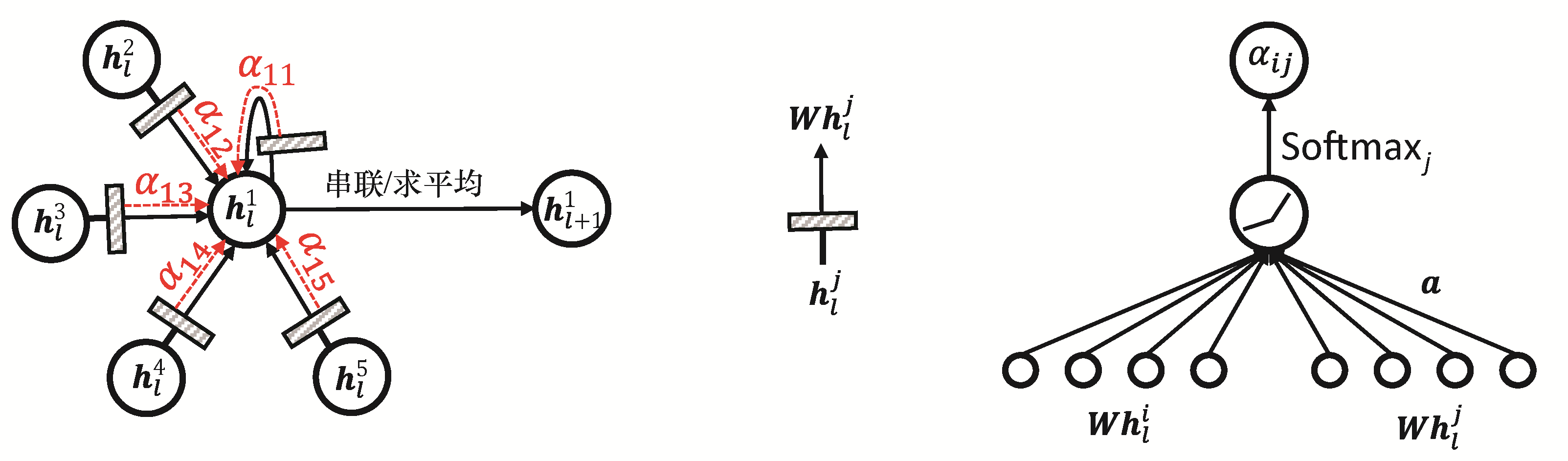


图7.5 左图：图卷积。中间图：具有共享参数矩阵的线性变换。  
右图：(Veličković et al，2018）中采用的注意力机制

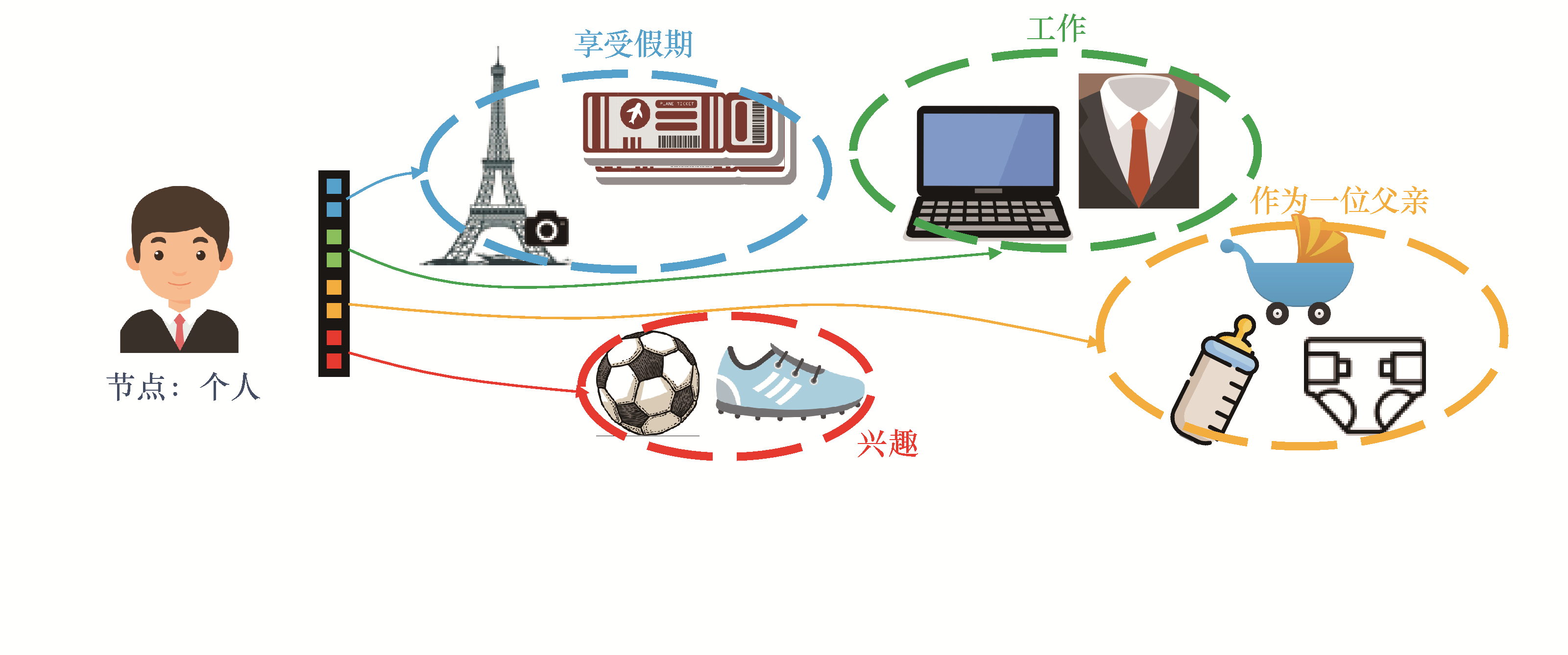


图7.6 使用多个嵌入来表征用户的兴趣，每个嵌入对应数据中的一个方面（Liu et al，2019a）

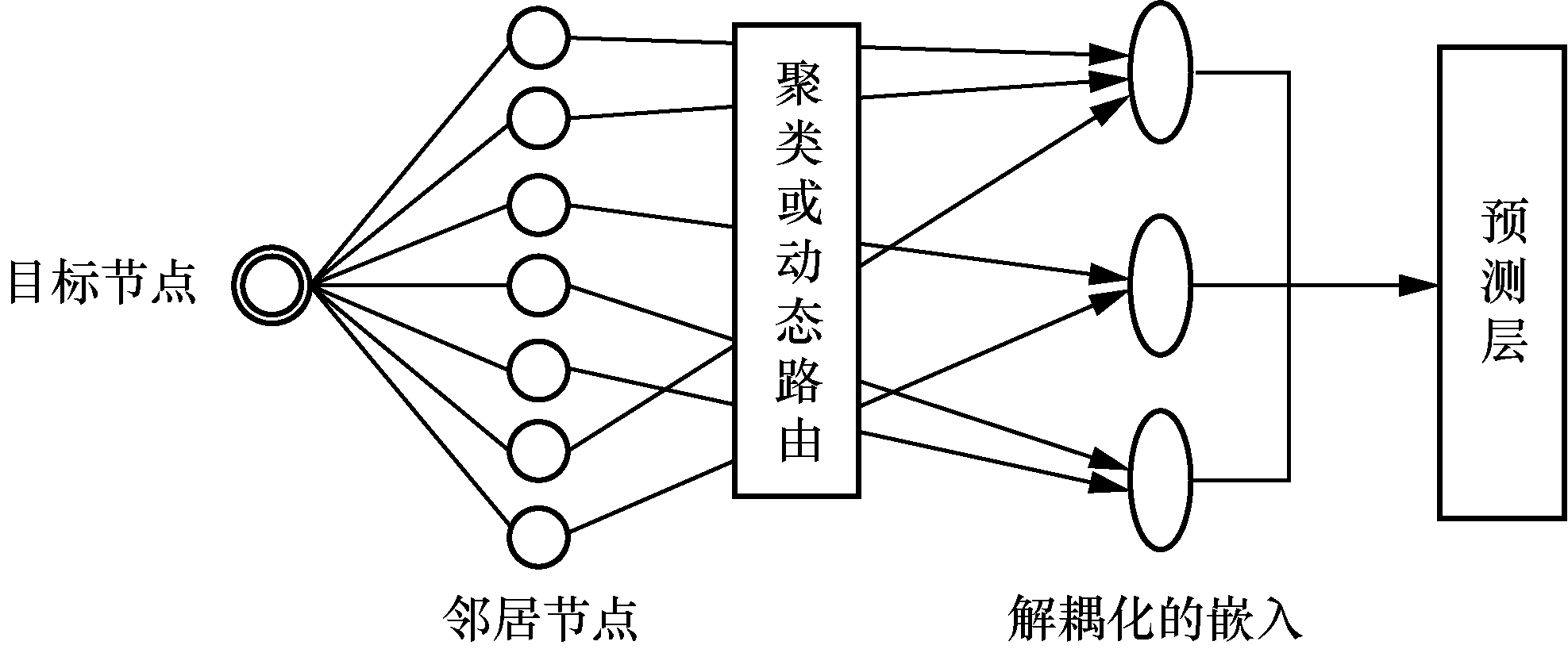


图7.7 通过使用聚类或动态路由为目标节点学习解耦化节点嵌入的高层思路