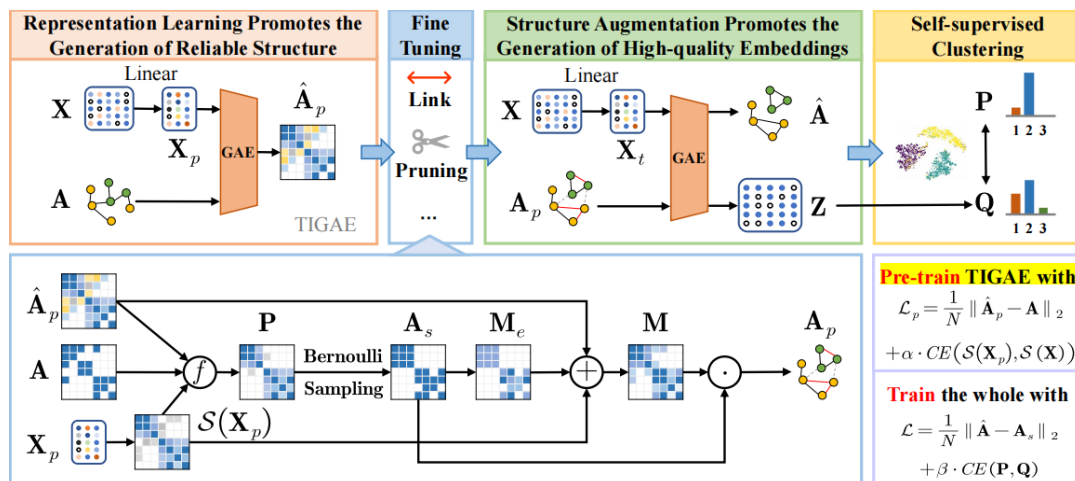


Synergistic Deep Graph Clustering Network(SDGCN)

1. 题目解释:

Synergistic--->协同的，文中指表示学习 (**representation learning**) 和结构增强 (**structure augmentation**)，两者之间互相促进，稳定的结构能够获得好的嵌入表达，而好的嵌入能引导结构的增强。

2. 模型图:



该结构分为两部分，分别是TIGAE和Sync。

TIGAE: 首先对属性进行线性变换，加入偏差bias，这样得到的Ap是n*n维的，相比于GAE，由于加了偏差，解码以后的Z会包含更多的信息。并且会有一个损失函数得到较为稳定的参数和嵌入。（这个也算是预训练的过程，先编码后解码）

$$\mathbf{Z} = \tilde{\mathbf{L}}\text{ReLU}(\tilde{\mathbf{L}}(\mathbf{X}\mathbf{W}_a^\top + \mathbf{W}_b)\mathbf{W}_1)\mathbf{W}_2,$$

$$\hat{\mathbf{A}}_p = \text{Sigmoid}(\mathbf{Z}\mathbf{Z}^\top).$$

$$\mathcal{L}_p = \frac{1}{N} \|\hat{\mathbf{A}}_p - \tilde{\mathbf{A}}\|_2 + \alpha \cdot CE(\mathcal{S}(\mathbf{X}_p), \mathcal{S}(\mathbf{X})),$$

Fine-tuning微调策略: 该策略从三个角度对上面模型得到的Ap进行修正：选取边预测概率，从P中选取概率最高的边，相当于增大了重要边的影响力；由于边预测邻接矩阵是非对称的，因此需要转置后相加除以2得到对称，然后用该矩阵与度矩阵进行运算，这样做的目的是增加重要节点的影响力；选取属性信息进行余弦相似度计算；最后将上面三个方面的矩阵相加均分得到微调后的增强邻接矩阵，相比于原始的邻接矩阵，增强后的加入了重要边信息、重要节点信息以及属性信息。

Sync: 协同交互模型，对TIGAE进行预训练得到较为稳定的参数和预测图，然后用参数不计算梯度得到Ap，将Ap微调后再次喂入TIGAE中，这次计算梯度，通过损失函数2来进行迭代优化

$$\mathcal{L} = \frac{1}{N} \|\hat{\mathbf{A}} - \mathbf{A}_s\|_2 + \beta \cdot CE(\mathbf{P}, \mathbf{Q}),$$

使用的核心模型是TIGAE，Sync更像是一种策略，先获得较为稳定的表达，然后用得到的表达带入模型中进行迭代。（这里要注意为什么不是减Ap，而是减As）

Clustering: 聚类方式选用的是自监督聚类（软聚类），将每个嵌入与聚类中心进行计算

$$q_{ij} = \frac{(1 + \|\mathbf{z}_i - \boldsymbol{\mu}_j\|^2)^{-1}}{\sum_{j'} (1 + \|\mathbf{z}_i - \boldsymbol{\mu}_{j'}\|^2)^{-1}},$$

3. 数据集：

Dataset	#Samples	#Classes	#Dimension	#Edges	R_h
UAT	1190	4	239	13599	0.70
CORA	2708	7	1433	5278	0.81
ACM	3025	3	1870	13128	0.82
CITE	3327	6	3703	4552	0.74
DBLP	4057	4	334	3528	0.80
AMAP	7650	8	745	119081	0.83
Wisconsin	251	5	1703	515	0.20
Texas	183	5	1703	325	0.11

没用大数据集，反而选取了少样例多维度的数据集，可能该方法在大数据集表现不好。但是用了软聚类后结果比传统的Kmeans要好，可以借鉴

4. 结果：

SynC KMeans	SynC SSL
91.51±0.12	92.73±0.04
70.01±0.26	73.58±0.22
76.43±0.30	79.58±0.11
91.49±0.12	92.74±0.04
77.62±0.55	83.48±0.13
45.62±1.54	55.11±0.24
50.17±1.18	61.70±0.27
77.19±0.54	82.90±0.17
66.93±0.80	71.77±0.27
40.49±0.53	46.37±0.42
41.52±0.88	48.09±0.45
62.97±0.36	65.72±0.36*
77.22±0.40*	78.58±0.38
57.26±0.57	58.13±0.52*
55.34±0.64	57.90±1.06
76.31±0.41*	77.65±0.30
80.94±0.11	82.48±0.04
68.92±0.18	69.70±0.23
61.85±0.54	65.02±0.11*
79.51±0.40*	80.69±0.11
60.39±0.46	57.33±0.13*
30.26±0.62	28.58±0.24*
30.09±0.76	26.60±0.17*
58.91±1.02	57.34±0.23*

5. 总结：

本论文的亮点是微调策略，将重要边信息、重要节点信息和属性信息都加入到增强邻接矩阵中，并且能够通过是否计算梯度来获得更好的表达。可以学习的地方还有论文排版和撰写，条例清晰，在有限的页数中消融实验、参数分析、结果图、时间复杂度都写上了。

