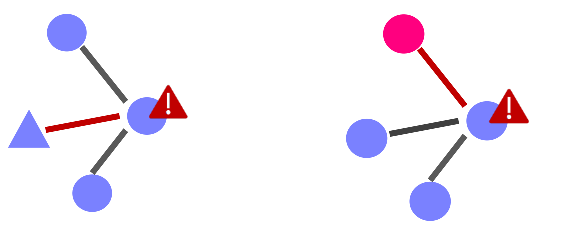
ANEMONE: Graph Anomaly Detection with Multi-Scale Contrastive Learning

MONASH University

Ming Jin, Yixin Liu, Yu Zheng, Lianhua Chi, Yuanfang Li and Shirui Pan

CIKM 2021 1-5 NOVEMBER

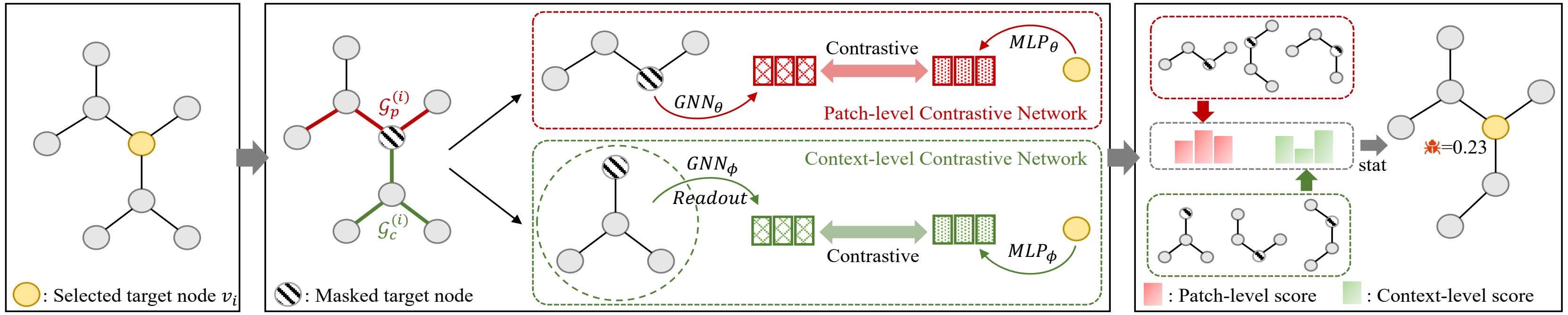
Motivation



We are motivated to evaluate the abnormality of each node according to the degree of agreement from multiple perspectives

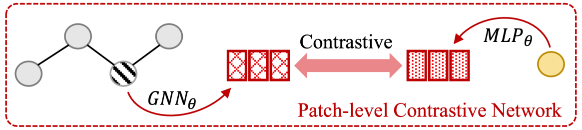
Framework

(a). Input Graph (b). Multi-Scale Contrastive Learning Model (c). Statistical Anomaly Estimator



Methodology

Patch-Level Contrasting



- Firstly, we obtain the embedding of the masked target node

$$\mathbf{H}_p^{(i)} = \sigma\left(\widetilde{\mathbf{D}_p^{(i)}}^{-\frac{1}{2}}\widetilde{\mathbf{A}_p^{(i)}}\widetilde{\mathbf{D}_p^{(i)}}^{-\frac{1}{2}}\mathbf{X}_p^{(i)}\mathbf{\Theta}\right)$$

- Then, we calculate the representation of this node

$$\mathbf{z}_p^{(i)} = MLP_\theta\left(\mathbf{x}^{(i)}\right) = \sigma\left(\mathbf{x}^{(i)}\mathbf{\Theta}\right)$$

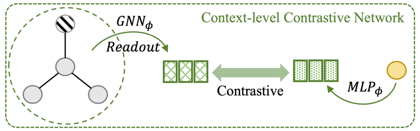
- Finally, we maximize their mutual information via a JSD estimator

$$s_p^{(i)} = Bilinear\left(\mathbf{h}_p^{(i)}, \mathbf{z}_p^{(i)}\right) = \sigma\left(\mathbf{h}_p^{(i)}\mathbf{W}_p\mathbf{z}_p^{(i)\top}\right)$$

$$\tilde{s}_p^{(i)} = Bilinear\left(\mathbf{h}_p^{(j)}, \mathbf{z}_p^{(i)}\right) = \sigma\left(\mathbf{h}_p^{(j)}\mathbf{W}_p\mathbf{z}_p^{(i)\top}\right)$$

$$\mathcal{L}_p = -\frac{1}{2n}\sum_{i=1}^{n}\left(\log\left(s_p^{(i)}\right) + \log\left(1 - \tilde{s}_p^{(i)}\right)\right)$$

Context-Level Contrasting



- Firstly, we obtain the contextual embedding of the masked target node

$$\mathbf{H}_c^{(i)} = \sigma\left(\widetilde{\mathbf{D}_c^{(i)}}^{-\frac{1}{2}}\widetilde{\mathbf{A}_c^{(i)}}\widetilde{\mathbf{D}_c^{(i)}}^{-\frac{1}{2}}\mathbf{X}_c^{(i)}\mathbf{\Phi}\right)$$

$$\mathbf{h}_c^{(i)} = readout\left(\mathbf{H}_c^{(i)}\right) = \frac{1}{K}\sum_{j=1}^{K}\mathbf{H}_c^{(i)}[j,:]$$

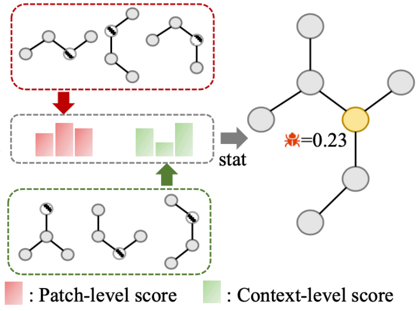
- Similarly, we maximize the mutual information between $z_c^{(i)}$ and $h_c^{(i)}$

$$\mathcal{L}_c = -\frac{1}{2n}\sum_{i=1}^{n}\left(\log\left(s_c^{(i)}\right) + \log\left(1 - \tilde{s}_c^{(i)}\right)\right)$$

The overall training objective:

$$\mathcal{L} = \alpha\mathcal{L}_c + (1-\alpha)\mathcal{L}_p$$

Anomaly Estimator



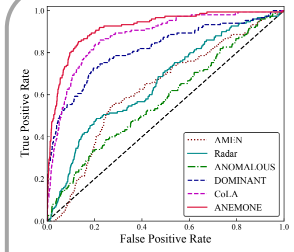
$$b_{view,j}^{(i)} = \tilde{s}_{view,j}^{(i)} - s_{view,j}^{(i)} \qquad \bar{b}_{view}^{(i)} = \sum_{j=1}^{R} b_{view,j}^{(i)}/R$$

$$y_{view}^{(i)} = \bar{b}_{view}^{(i)} + \sqrt{\sum_{j=1}^{R}\left(b_{view,j}^{(i)} - \bar{b}_{view}^{(i)}\right)^2/R}$$

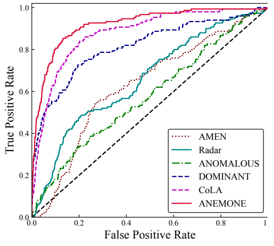
The final anomaly score of a node v_i:

$$y^{(i)} = \alpha y_c^{(i)} + (1-\alpha)y_p^{(i)}$$

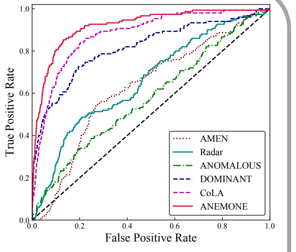
Experiments



ROC curve of Cora



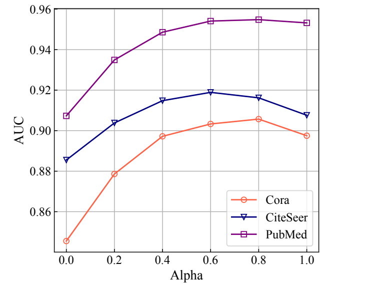
ROC curve of CiteSeer



ROC curve of PubMed

Methods	Cora	CiteSeer	PubMed
AMEN	0.6266	0.6154	0.7713
Radar	0.6587	0.6709	0.6233
ANOMALOUS	0.5770	0.6307	0.7316
DOMINANT	0.8155	0.8251	0.8081
CoLA	0.8779	0.8968	0.9512
CoLA$_{stat}$	0.8869	0.9047	0.9532
ANEMONE$_{mean}$	0.8963	0.9066	0.9524
ANEMONE$_{std}$	0.5402	0.7077	0.7440
ANEMONE	**0.9057**	**0.9189**	**0.9548**

AUC of ANEMONE, its competitors and variants



Trade-off parameter α w.r.t. AUC values