Graph Anomaly Detection via Multi-Scale Contrastive Learning Networks with Augmented View

Jingcan Duan,1 Siwei Wang,1 Pei Zhang,1 En Zhu,1*
Jingtao Hu,1 Hu Jin,1 Yue Liu,1 Zhibin Dong1
1College of Computer, National University of Defense Technology, Changsha, China
{jingcan duan, yueliu19990731}@163.com, {wangsiwei13, zhangpei, enzhu, hujingtao17, jinhu, dzb20}@nudt.edu.cn

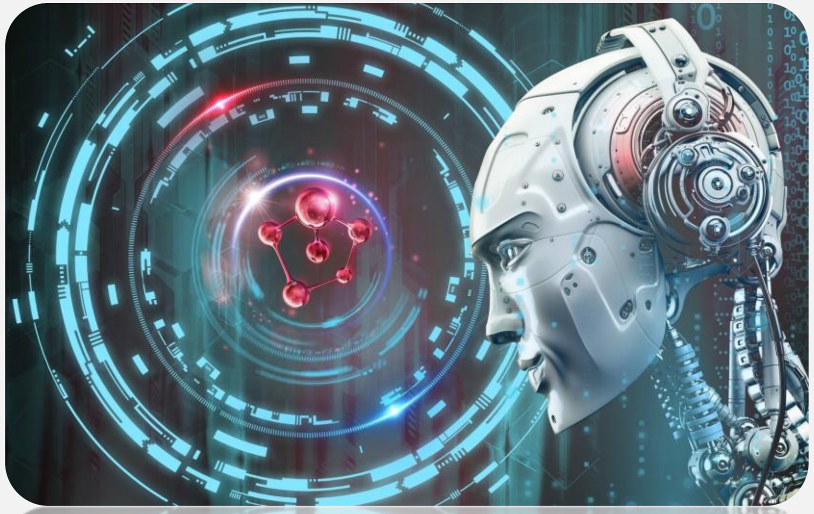
AAAI2023

Code: https://github.com/FelixDJC/GRADATE.

2023.5.13 • ChongQing

Leibniz-Institut für Sozialwissenschaften

Reported by Yang Peng



1.Introduction

2.Method

3.Experiments

Introduction

Task:
graph anomaly detection, which intends to **discern the anomalies** from the majority of nodes. Recent methods have paid attention to various scales of contrastive strategies for GAD, i.e., **node-subgraph** and **node-node contrasts**. However,they neglect the **subgraph-subgraph comparison** information.

Contributions:

1. We introduce subgraph-subgraph contrast to GAD for the first practice and propose a multi-scale contrastive learning networks framework with an augmented view.
2. We investigate the effects of different graph augmentation on subgraph representation learning for the task

Method

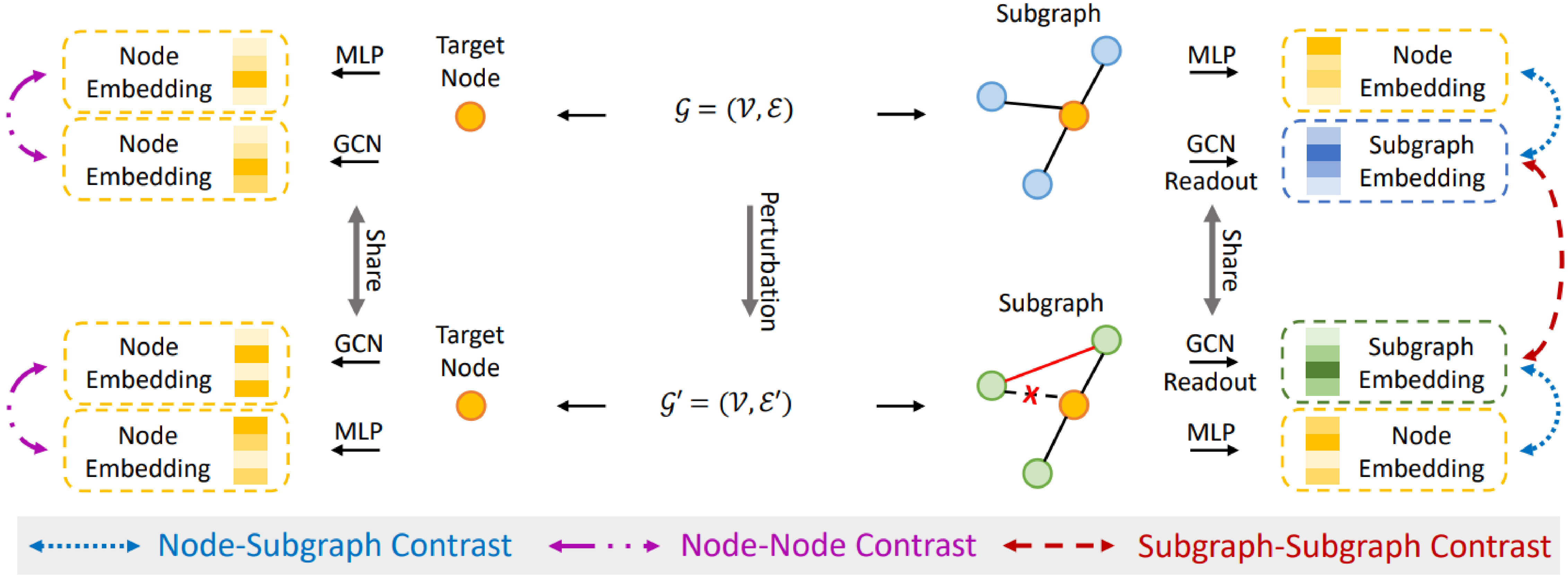
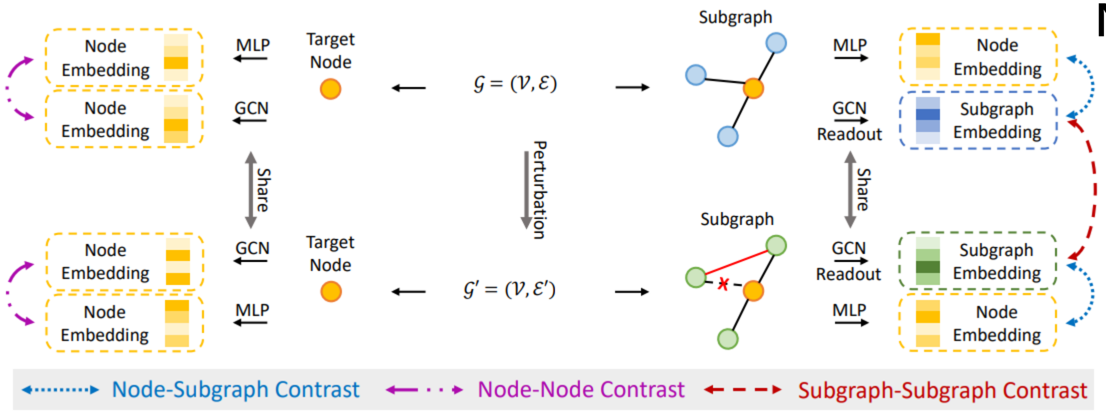
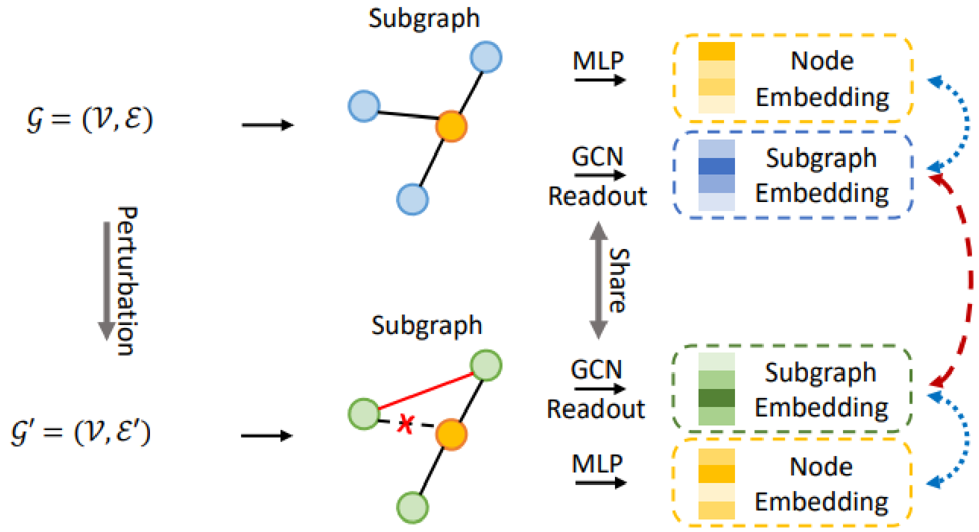


Figure 1: Overview of the GRADATE model. It is composed of two main modules: (1) Graph augmentation. We take the original graph as the first view and the edge-modified graph as the second view. Subgraphs used in the later module are generated by random walk with restart. (2) Graph contrastive network. The network captures various anomalous information from multiple scales under two views by building node-subgraph, node-node, and subgraph-subgraph contrasts. Then we comprehensively calculate the anomaly score for each node.

Method





Node-Subgraph Contrast

The subgraph hidden-layer representation

$$\mathbf{H}_i^{(\ell+1)} = \sigma\left(\widetilde{\mathbf{D}}_i^{-\frac{1}{2}} \widetilde{\mathbf{A}}_i \widetilde{\mathbf{D}}_i^{-\frac{1}{2}} \mathbf{H}_i^{(\ell)} \mathbf{W}^{(\ell)}\right), \quad (1)$$

$$\boldsymbol{z}_i = Readout\left(\mathbf{Z}_i\right) = \sum_{j=1}^{n_i} \frac{\left(\mathbf{Z}_i\right)_j}{n_i}. \quad (2)$$

The node hidden-layer representation

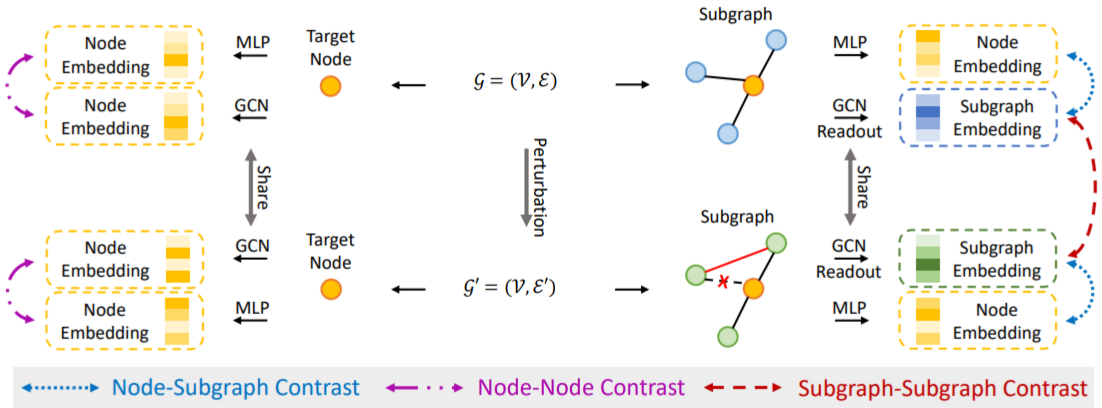
$$\boldsymbol{h}_i^{(\ell+1)} = \sigma\left(\boldsymbol{h}_i^{(\ell)} \mathbf{W}^{(\ell)}\right), \quad (3)$$

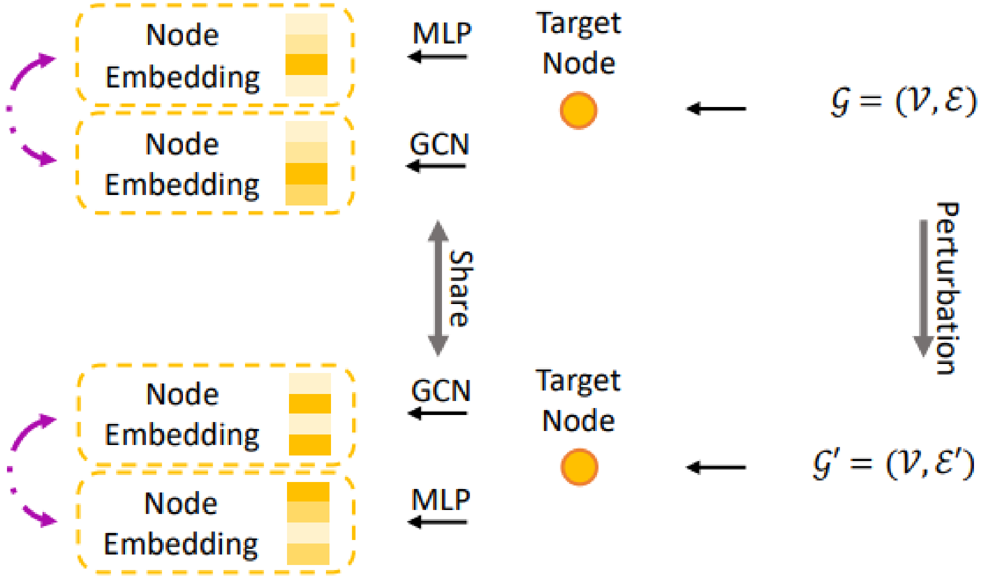
$$s_i^1 = Bilinear\left(\boldsymbol{z}_i, \boldsymbol{e}_i\right) = sigmoid\left(\boldsymbol{z}_i \mathbf{W} \boldsymbol{e}_i^{\top}\right). \quad (4)$$

$$\mathcal{L}_{NS}^1 = -\sum_{i=1}^{N} \left(y_i \log\left(s_i^1\right) + \left(1 - y_i\right) \log\left(1 - s_i^1\right)\right), \quad (5)$$

$$\mathcal{L}_{NS} = \alpha \mathcal{L}_{NS}^1 + (1 - \alpha) \mathcal{L}_{NS}^2, \quad (6)$$

Method





Node-Node Contrast

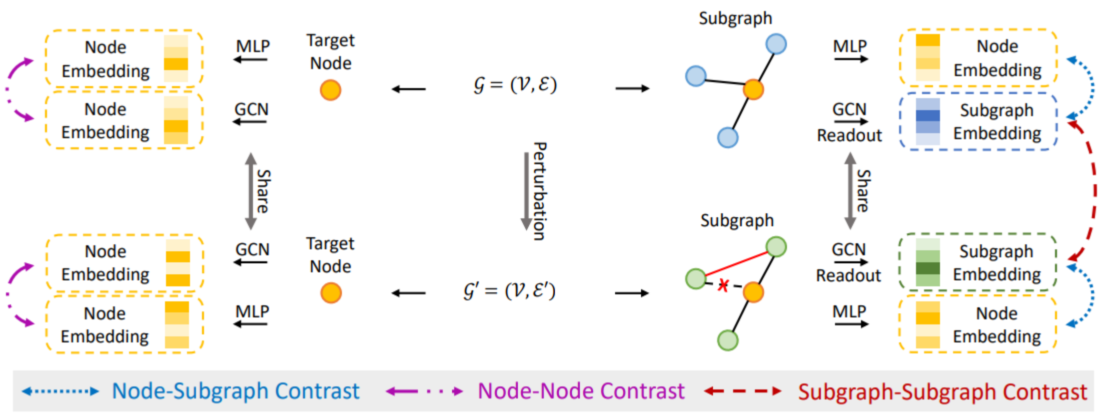
$$\mathbf{H}_i^{\prime(\ell+1)} = \sigma\left(\widetilde{\mathbf{D}'}_i^{-\frac{1}{2}} \widetilde{\mathbf{A}'}_i \widetilde{\mathbf{D}'}_i^{-\frac{1}{2}} \mathbf{H}_i^{\prime(\ell)} \mathbf{W}^{\prime(\ell)}\right), \quad (7)$$

$$\hat{\boldsymbol{h}}_i^{(\ell+1)} = \sigma\left(\hat{\boldsymbol{h}}_i^{(\ell)} \mathbf{W}^{\prime(\ell)}\right), \quad (8)$$

$$\mathcal{L}_{NN}^1 = -\sum_{i=1}^{N} \left(\hat{y}_i \log\left(\hat{s}_i^1\right) + (1-\hat{y}_i) \log\left(1-\hat{s}_i^1\right)\right). \quad (9)$$

$$\mathcal{L}_{NN} = \alpha \mathcal{L}_{NN}^1 + (1-\alpha)\, \mathcal{L}_{NN}^2, \quad (10)$$

Method



Subgraph-Subgraph Contrast

$$\mathcal{L}_{SS} = -\sum_{i=1}^{n} \log \frac{\exp\left(\boldsymbol{z}_i^1 \cdot \boldsymbol{z}_i^2\right)}{\exp\left(\boldsymbol{z}_i^1 \cdot \boldsymbol{z}_j^1\right) + \exp\left(\boldsymbol{z}_i^1 \cdot \boldsymbol{z}_j^2\right)}, \quad (11)$$

$$\mathcal{L} = \beta \mathcal{L}_{NS} + (1-\beta)\mathcal{L}_{NN} + \gamma \mathcal{L}_{SS}, \quad (12)$$

Anomaly Score Calculation

$$s_i = s_i^n - s_i^p, \quad (13)$$

$$s_i = \alpha s_i^1 + (1-\alpha) s_i^2,$$
$$\hat{s}_i = \alpha \hat{s}_i^1 + (1-\alpha)\hat{s}_i^2, \quad (14)$$
$$S_i = \beta s_i + (1-\beta)\hat{s}_i,$$

$$\bar{S}_i = \frac{1}{R}\sum_{r=1}^{R} S_i^{(r)},$$
$$S_i = \bar{S}_i + \sqrt{\frac{1}{R}\sum_{r=1}^{R}\left(S_i^{(r)} - \bar{S}_i\right)^2}, \quad (15)$$

Experiments

Table 2: The statistics of datasets.

Datasets	Nodes	Edges	Attributes	Anomalies
EAT	399	5993	203	30
WebKB	919	1662	1703	60
UAT	1190	13599	239	60
Cora	2708	5429	1433	150
UAI2010	3067	28311	4973	150
Citation	8935	15098	6775	450

Experiments

Table 3: Performance comparison for AUC. The bold and underlined values indicate the best and runner-up results, respectively.

Methods	EAT	WebKB	UAT	Cora	UAI2010	Citation
LOF (Breunig et al. 2000)	0.5255	0.2903	0.4906	0.3538	0.7052	0.3059
ANOMALOUS (Peng et al. 2018)	0.4109	0.3417	0.3356	0.3198	0.5026	0.5656
DOMINANT (Ding et al. 2019)	0.6023	0.7787	0.6503	0.8929	0.7698	0.7748
CoLA (Liu et al. 2021)	0.6762	0.8175	0.6538	0.8847	0.7949	0.7296
ANEMONE (Jin et al. 2021a)	0.7726	0.8208	0.8087	0.9122	0.8731	0.8028
SL-GAD (Zheng et al. 2021)	0.6974	0.8678	0.6851	0.9192	0.8454	0.8095
HCM (Huang et al. 2021)	0.4536	0.5064	0.3262	0.6276	0.5210	0.5414
Sub-CR (Zhang, Wang, and Chen 2022)	0.6672	0.8423	0.6788	0.9133	0.8571	0.7903
GRADATE	**0.7980**	**0.8740**	**0.8451**	**0.9237**	**0.9262**	**0.8138**

Experiments

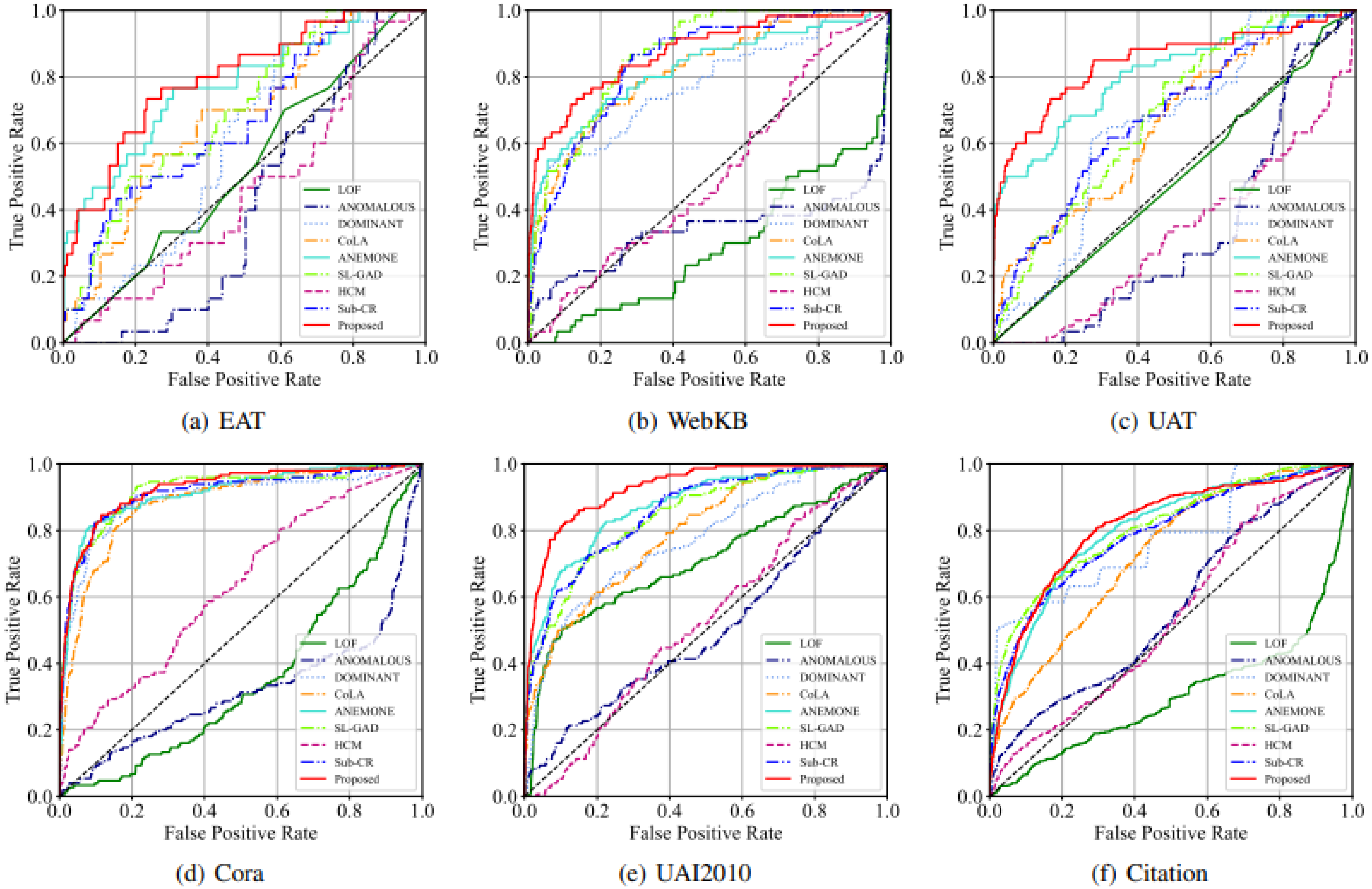


Figure 2: ROC curves on six benchmark datasets. The area under the curve is larger, the anomaly detection performance is better. The black dotted lines are the "random line", indicating the performance under random guessing.

Experiments

Table 4: Ablation study for contrast scale w.r.t. AUC.

	EAT	WebKB	UAT	Cora	UAI2010	Citation
NS	0.6762	0.7949	0.6538	0.8847	0.8175	0.7296
NS+SS	0.6800	0.8310	0.6603	0.8956	0.9055	0.6978
NS+NN	0.7726	0.8208	0.8087	0.9122	0.8731	0.8028
NS+NN+SS	**0.7980**	**0.8740**	**0.8451**	**0.9237**	**0.9262**	**0.8138**

Table 5: Ablation study for graph augmentation w.r.t. AUC.

	EAT	WebKB	UAT	Cora	UAI2010	Citation
GNF	0.7548	0.8183	0.8327	0.9031	0.9193	0.7902
FM	0.7782	0.8148	0.8256	0.8924	0.9171	0.8034
GD	0.7618	0.8062	0.8143	0.9026	0.9161	0.8030
EM	**0.7980**	**0.8740**	**0.8451**	**0.9237**	**0.9262**	**0.8138**

Experiments

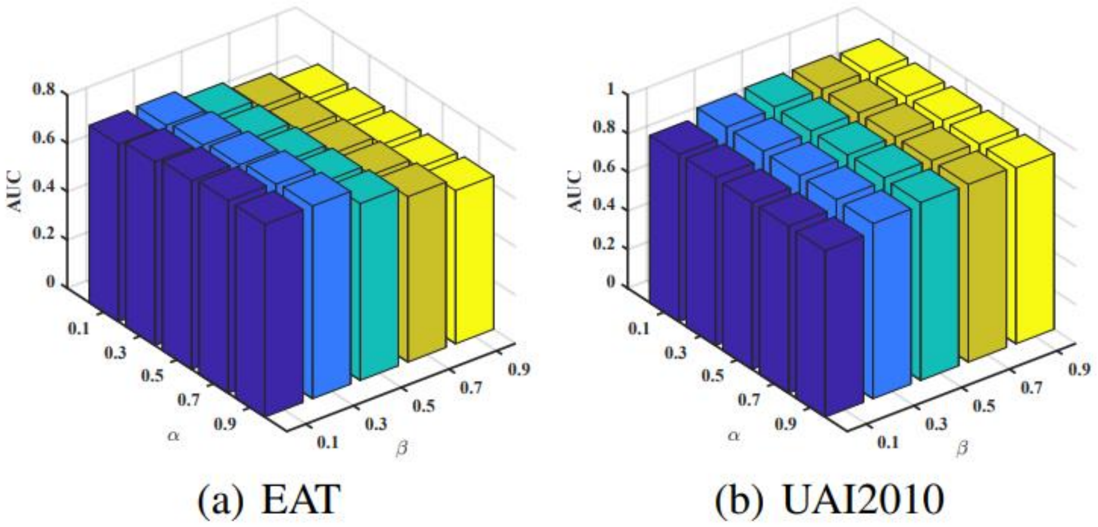


Figure 3: Sensitivity analysis for the balance parameters α and β w.r.t. AUC on EAT and UAI2010.

Figure 4 illustrates the performance variation of GRADATE when γ varies from 0.1 to 0.9. From the figure, we observe that GRADATE tends to perform well by setting γ to 0.1 across all benchmarks.

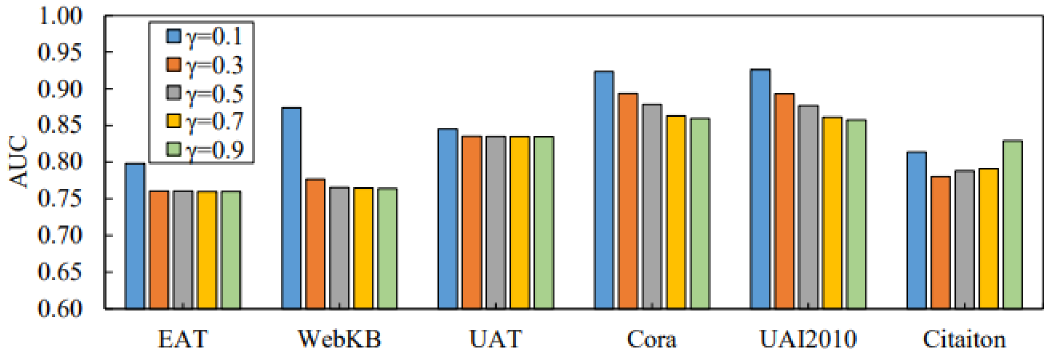


Figure 4: Trade-off parameter γ w.r.t. AUC.

Experiments

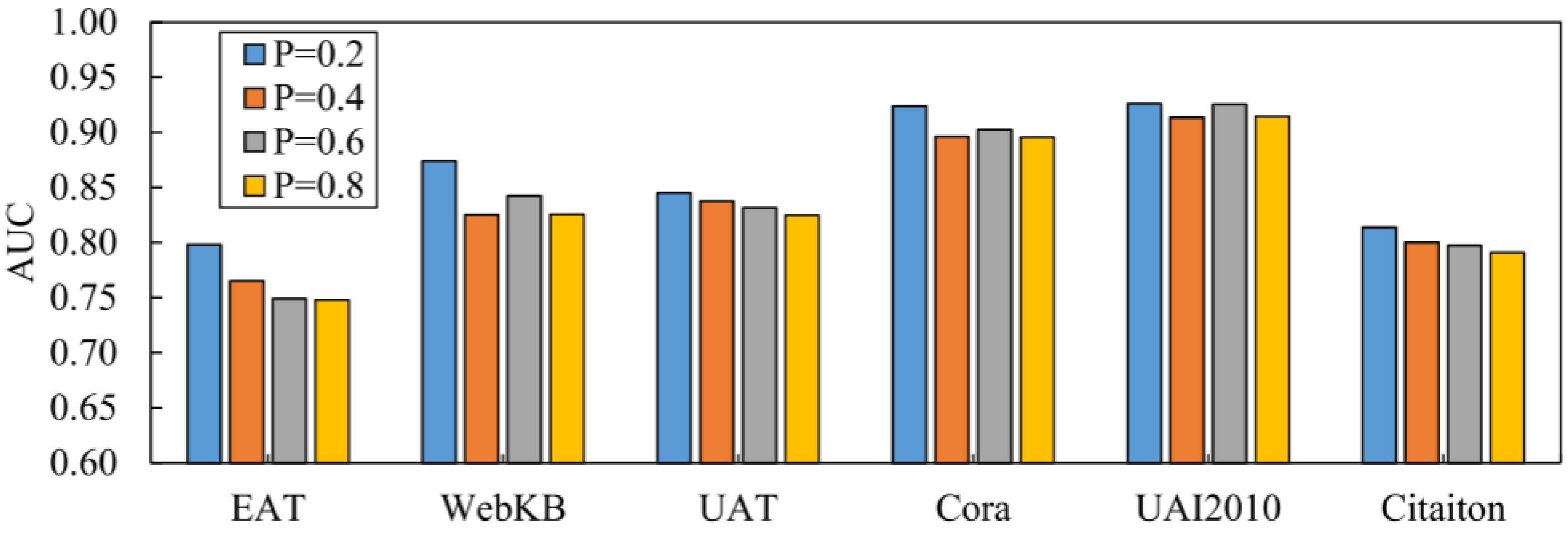


Figure 5: Perturbation proportion P w.r.t. AUC.

Thank you!