XINJIANG UNIVERSITY

Rethinking graph anomaly detection: A self-supervised Group Discrimination paradigm with Structure-Aware

Junyi Yan, Enguang Zuo, Chen Chen, Cheng Chen, Jie Zhong, Tianle Li, Xiaoyi Lv*

Xinjiang University, China
Email: yjy@stu.xju.edu.cn

13/7/2023

IEEE International Conference on Multimedia and Expo 2023
Brisbane Convention & Exhibition Centre
10-14 July 2023

*Corresponding author

Introducing a new self-supervised learning paradigm for graph anomaly detection

GDSA

Stronger detection performance and higher computational efficiency

- ✓ Better detection performance
- ✓ Less time complexity
- ✓ Smaller memory usage

Outline

Part1: **(Background)** What is Graph Anomaly Detection **(GAD)** ?

Part2: **Existing Problems** of Graph Anomaly Detection **(GAD)**

Part3: **Rethinking** Graph Anomaly Detection **(GAD)**

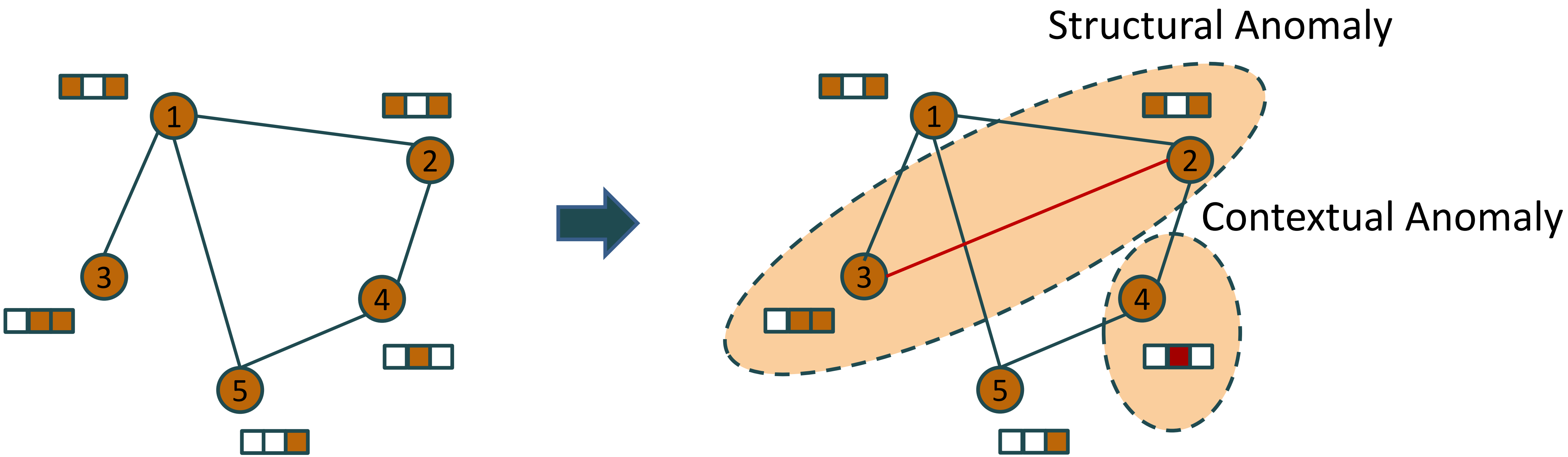
Part4: **Structure Disturbance-**A new approach for Graph Anomaly Detection **(GAD)**

Part5: **Group Discrimination-**Conversion graph anomaly detection **(GAD)**

Part6: **Experiment** Results

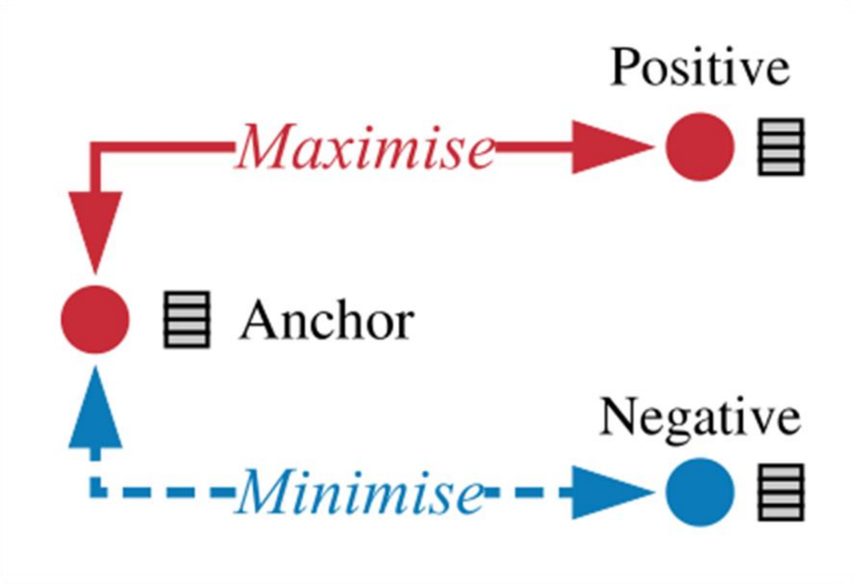
Prat7: **Conculsion** and Future Works

(Background) What is Graph Anomaly Detection (GAD) ?



Existing Problems of Graph Anomaly Detection (GAD) ?

- Insufficient detection effect (Not directly model the anomaly structure)
- Inefficient calculation (Higher time complexity and Larger memory usage)



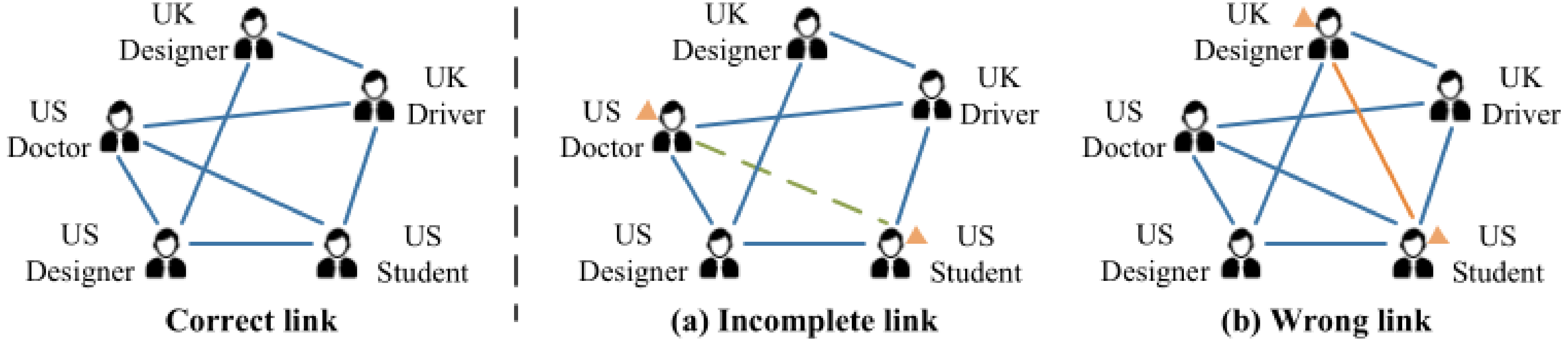
at least O(M)

Rethinking Graph Anomaly Detection (GAD) ?

- **To directly model the topology of the graph**
- **To improve computational efficiency**
- **To improve Generalization and Scalability**

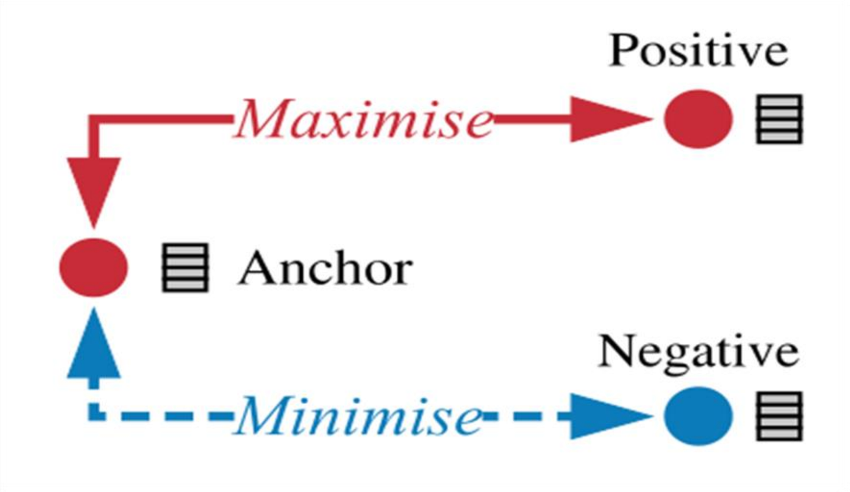
Rethinking Graph Anomaly Detection (GAD) ?

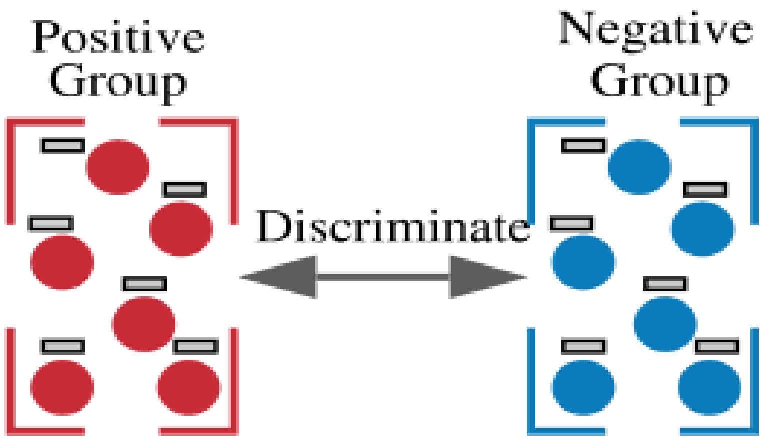
- **Q1: How to explicitly model the abnomal structure in the graph?**



Rethinking Graph Anomaly Detection (GAD) ?

- **Q2: How to effectively improve the computational efficiency of graph anomaly detection?**





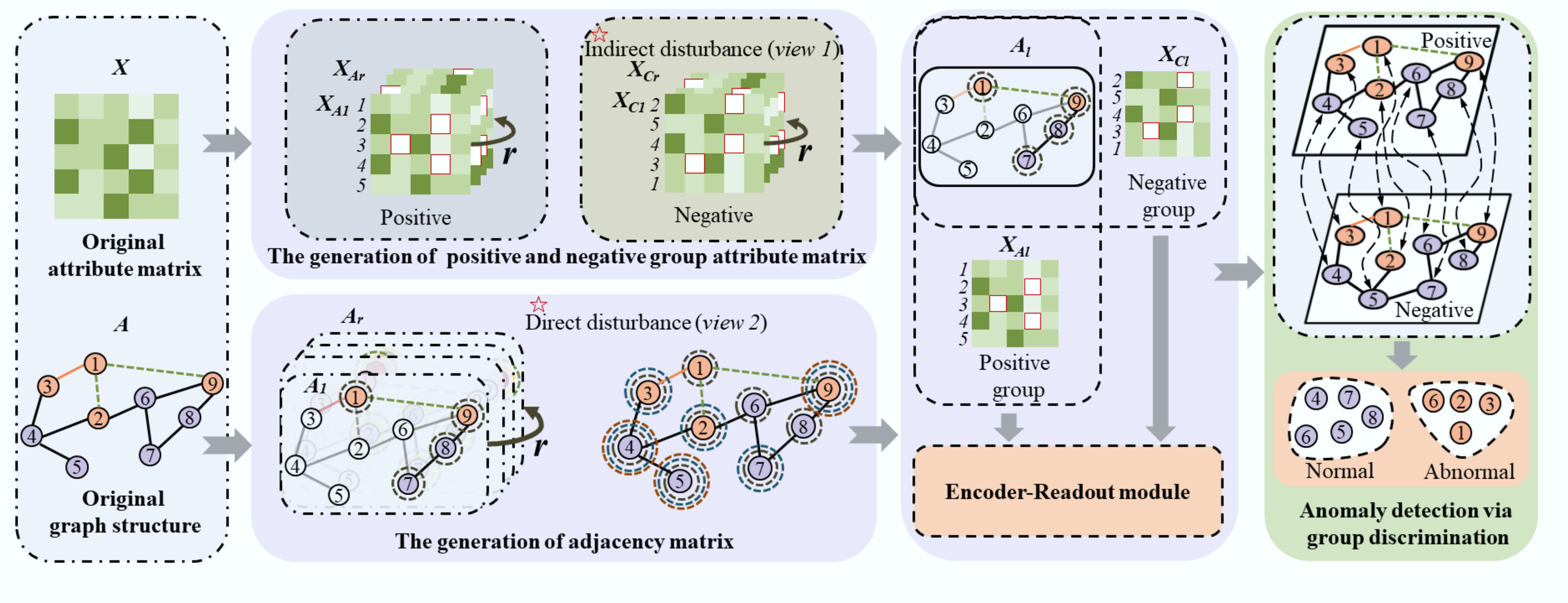
at least O(M) **O(1)**

Rethinking Graph Anomaly Detection (GAD) ?

- **Q3: How to ensure the generalization and scalability of the model?**

✓ **Generalization: Four benchmark datasets of different scales and types (two citation network datasets and two social network datasets)**

✓ **Scalability: The large-scale dataset**

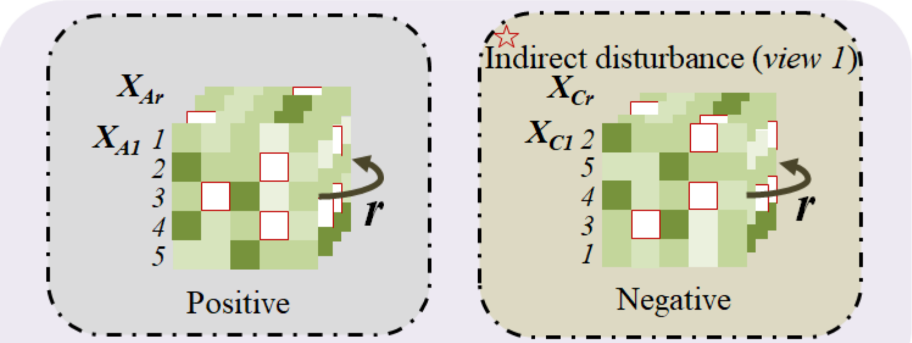


X
Original attribute matrix
A
Original graph structure
Positive
Indirect disturbance (view 1)
Negative
The generation of positive and negative group attribute matrix
Direct disturbance (view 2)
The generation of adjacency matrix
Negative group
Positive group
Encoder-Readout module
Positive
Negative
Normal
Abnormal
Anomaly detection via group discrimination

GDSA

Multi-view Structure Disturbance

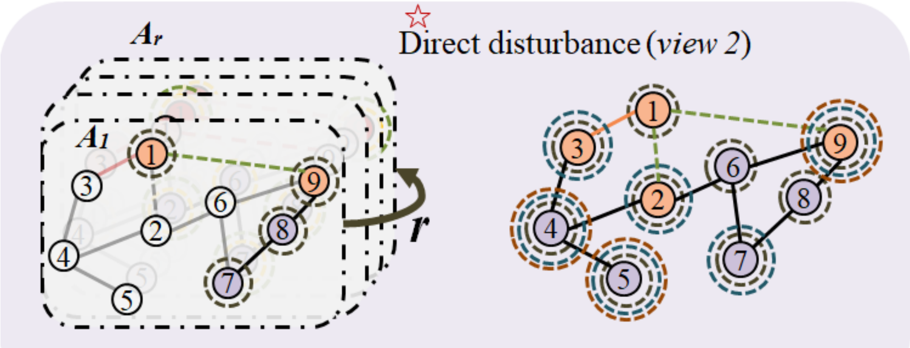
✓ Indirect disturbance (*View 1*):



The generation of positive and negative group attribute matrix

Simulate incomplete links and wrong links

✓ Direct disturbance (*View 2*):



The generation of adjacency matrix

Simulate wrong links to increase the richness of structural disturbance

Multi-view Structure Disturbance

☐ How is the value of the number of edges per round of disturbance *k* determined in *View 2* structure disturbance?

- Random sampling with put-back of edges in $\mathcal{E}$
- Selecting *k* edges per round, for a total of *r* rounds

Graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathbf{X})$

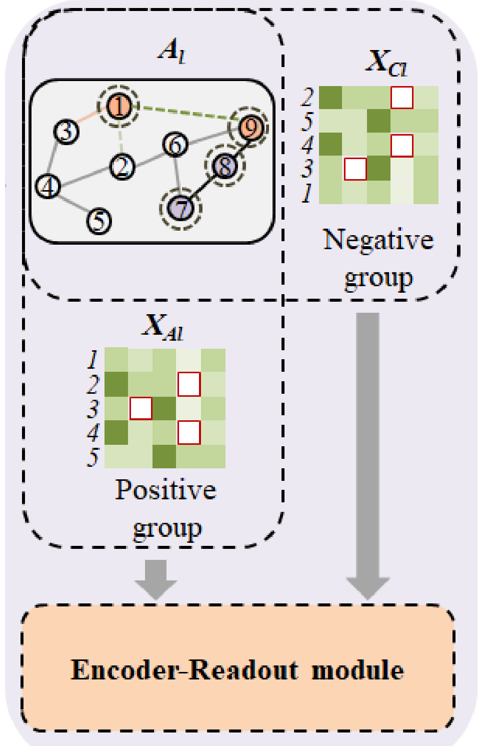
Node set $\mathcal{V} = \{v_1, \dots, v_m\}$
Edge set $\mathcal{E} = \{e_1, \dots, e_n\}$

$$\boldsymbol{E(m) = 1 - \left(\frac{n-k}{n}\right)^r}$$

Group Discrimination

✓ **Feature extraction and transformation**

☐ **Objective: To extract the features of positive and negative groups and convert them into node scalar information for group discrimination.**



✓ **Encoder:** Extract spatial features in the graph

$$\mathbf{E}_l = GCN(\mathbf{A}_l, \mathbf{X}_l)$$

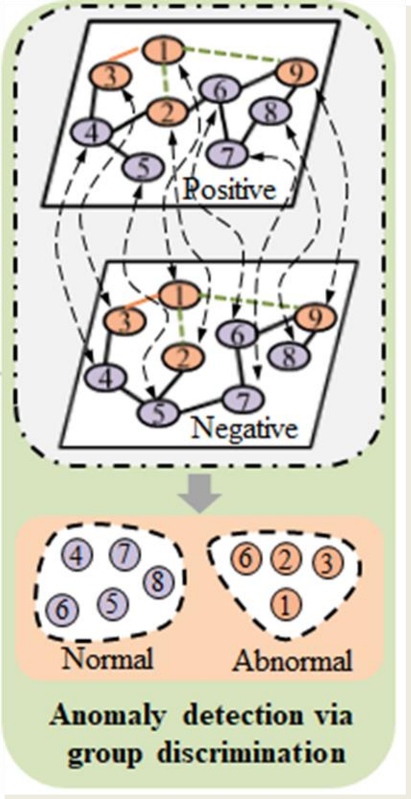
✓ **Readout:** Dimensionality reduction is performed for embedding in the positive and negative groups

$$\begin{cases} \mathbf{e}_l = \sigma(\mathbf{E}_l) \\ s_i = MLP(\mathbf{E}_l) = \sum_{j=1}^{h} \mathbf{e}_l[i,j] \end{cases}$$

Group Discrimination

✓ **Identifying node scalar information to complete graph anomaly detection**

☐ **Time complexity is only O(1)**



✓ BCEloss

$$\mathcal{L}_{BCE} = \frac{1}{2m}\left(\sum_{i=1}^{2m} y_i log\, s_i + (1 - y_i) log(1 - s_i)\right)$$

✓ In an ideal state

s_i of abnormal nodes： **Positive**

s_i of normal nodes： **Negative**

Experiment Results

☐ Generalization-- Detection performance

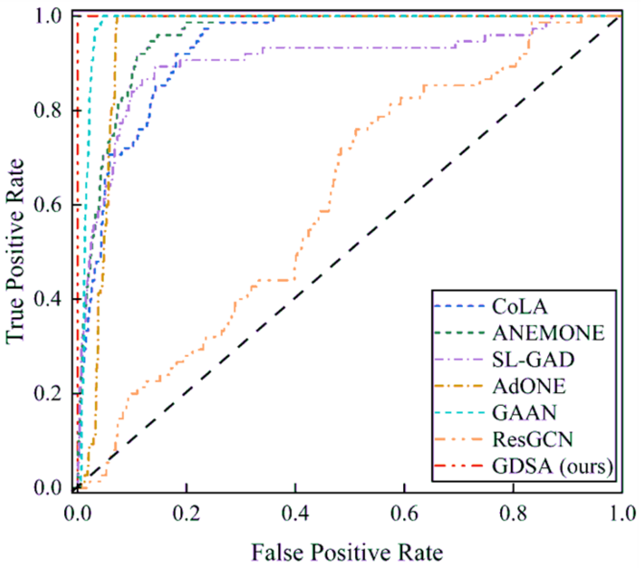
Methods	Cora	Citeseer	BlogCatalog	Flickr
CoLA(TNNLS 2021)	0.9338	0.9055	0.6804	0.6365
ANEMONE(CIKM 2021)	0.9706	0.9655	0.6681	0.6180
SL-GAD(TKDE 2021)	0.9035	0.9127	0.6477	0.6144
AdONE(WSDM 2020)	0.9525	0.9922	0.6144	0.3754
GANN(CIKM 2020)	0.9841	0.9851	0.6051	0.6324
ResGCN(DSAA 2021)	0.6117	0.5135	0.6083	0.6113
GDSA(ours)	1.0000*	1.0000*	0.7163	0.8191

↑ 28.7%

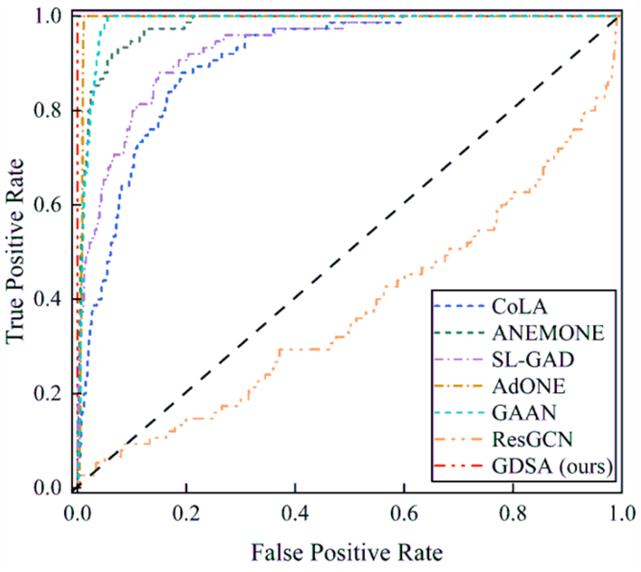
Experiment Results

☐ Generalization-- Detection performance

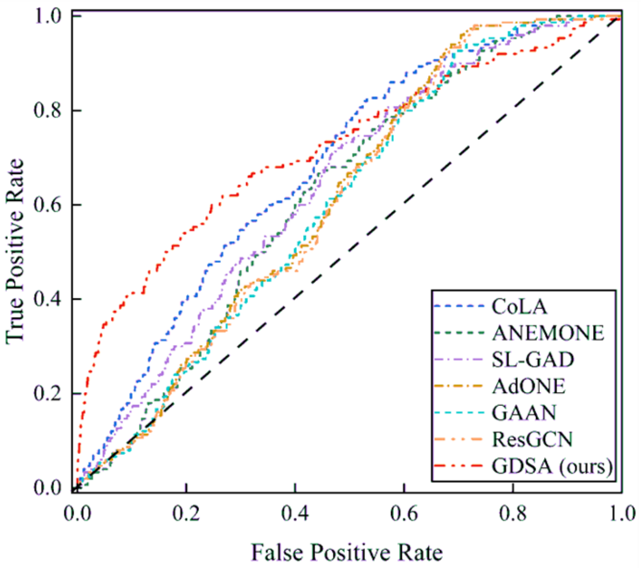
Cora



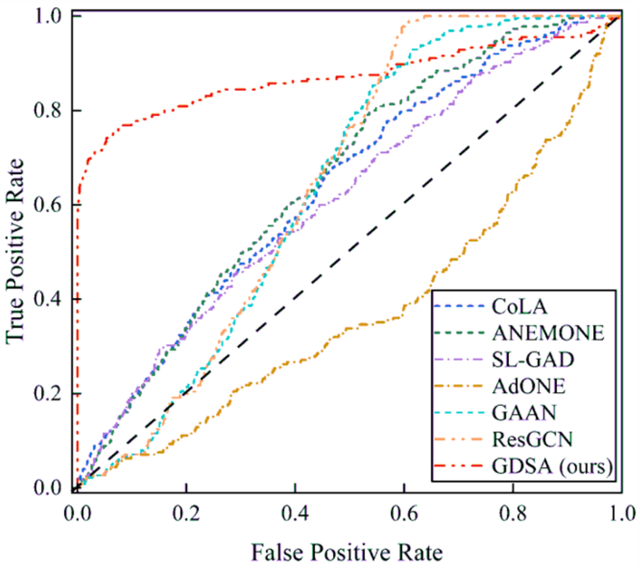
Citeseer



BlogCatalog

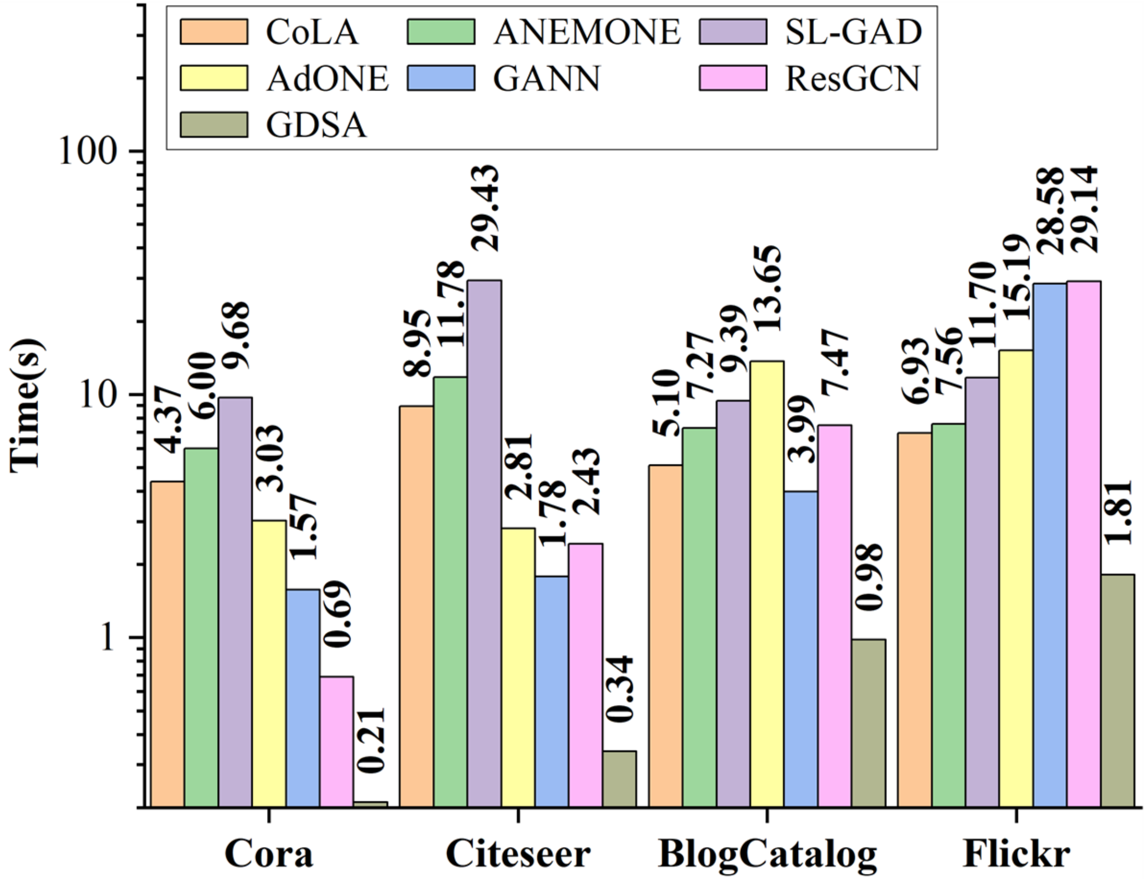


Flickr



Experiment Results

Generalization-- computational efficiency



On the Citeseer dataset

GDSA compared to GANN : improved 5.3×
GDSA compared to SL-GAD : improved 87.8×

Experiment Results

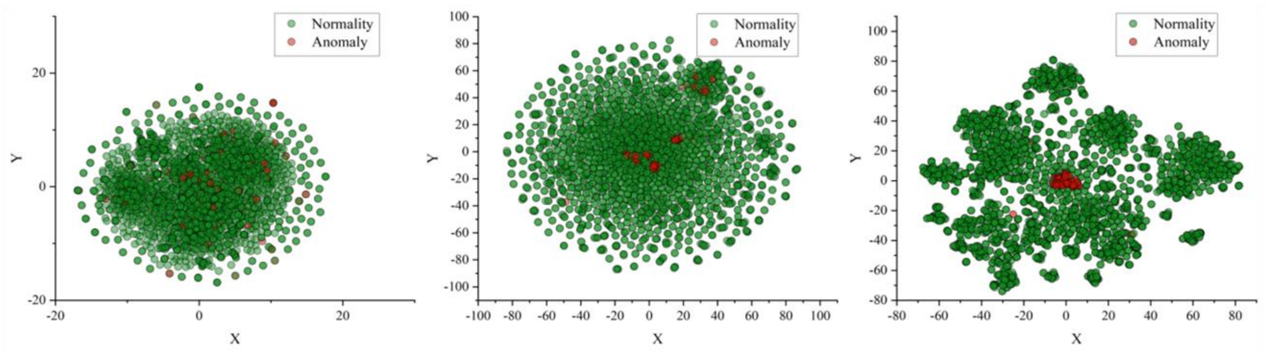
□ Ablation study-- Detection performance

	Cora	Citeseer	BlogCatalog	Flickr
GDSA w/o both	0.0528±0.0069	0.2410±0.0734	0.6303±0.0024	0.7685±0.0176
GDSA w/o *view 1*	0.0550±0.0070	0.3004±0.1339	0.6325±0.0028	0.7739±0.0084
GDSA w/o *view 2*	**1.0000** ±0.0000	**1.0000** ±0.0000	0.7146±0.1343	0.8015±0.0569
GDSA(ours)	**1.0000** ±0.0000	**1.0000** ±0.0000	**0.7163** ±0.0069	**0.8191** ±0.0375

Experiment Results

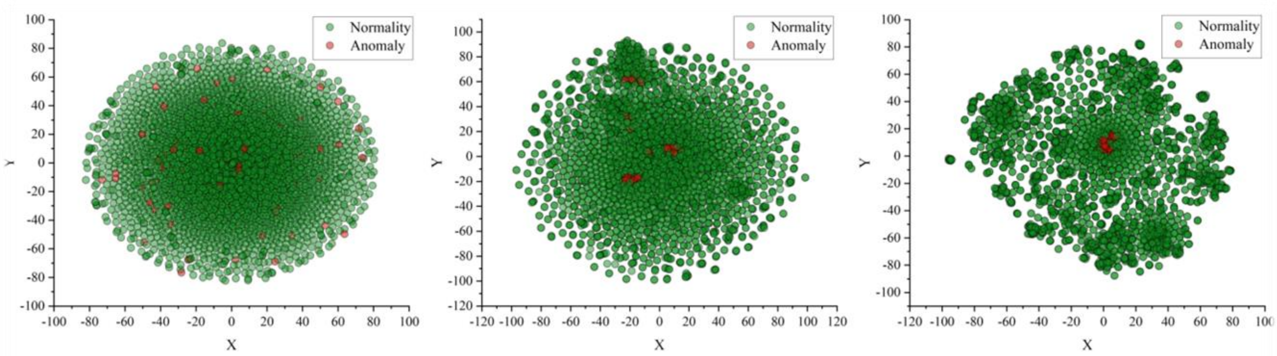
☐ Ablation study-- Latent space distribution

Cora



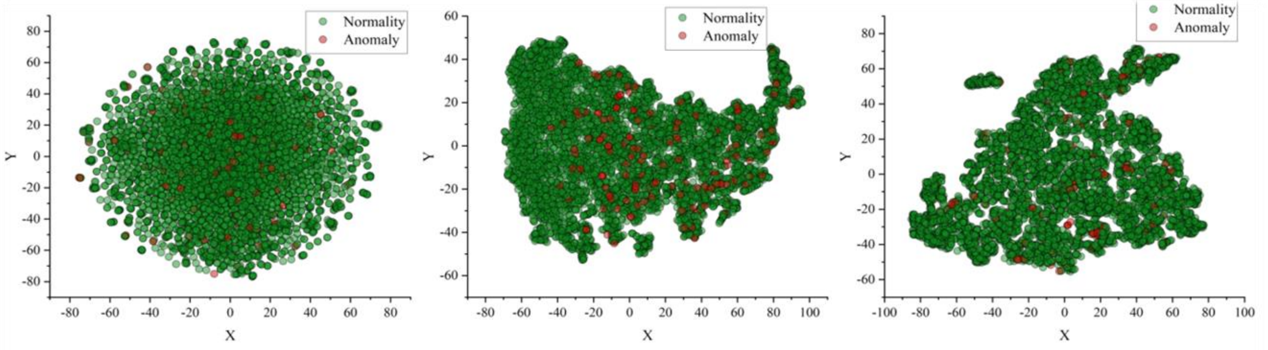
(a) The initial feature. (b) GDSA w/o both. (c) GDSA.

Citeseer



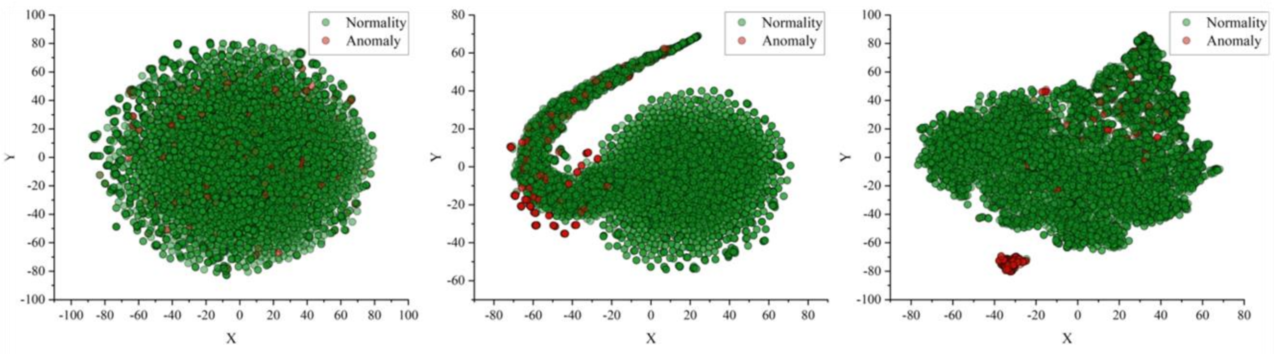
(a) The initial feature. (b) GDSA w/o both. (c) GDSA.

BlogCatalog



(a) The initial feature. (b) GDSA w/o both. (c) GDSA.

Flickr

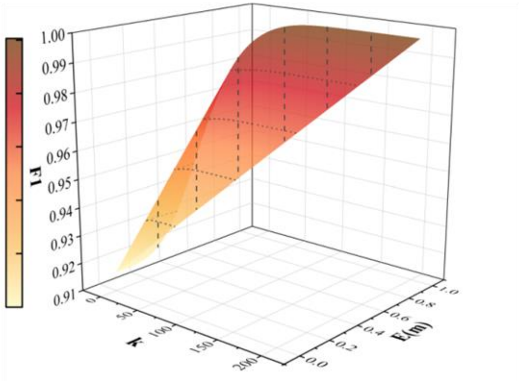


(a) The initial feature. (b) GDSA w/o both. (c) GDSA.

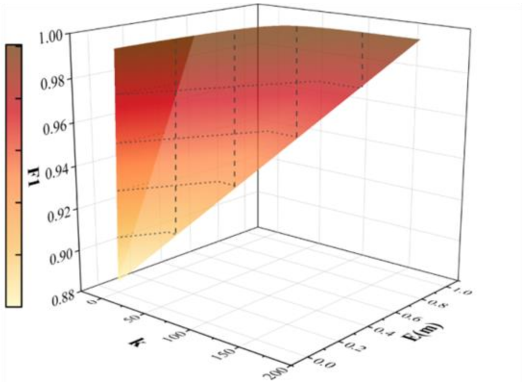
Experiment Results

Parameter analysis

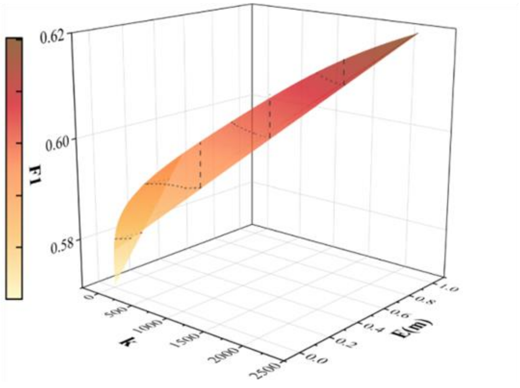
Cora



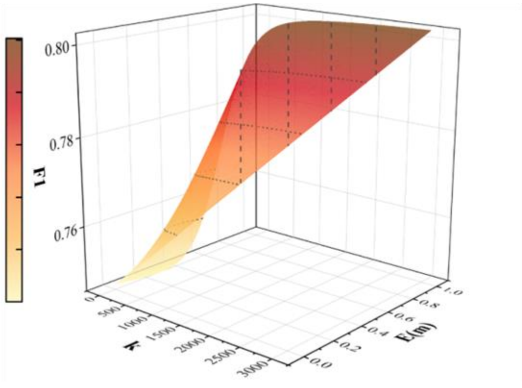
Citeseer



BlogCatalog



Flickr



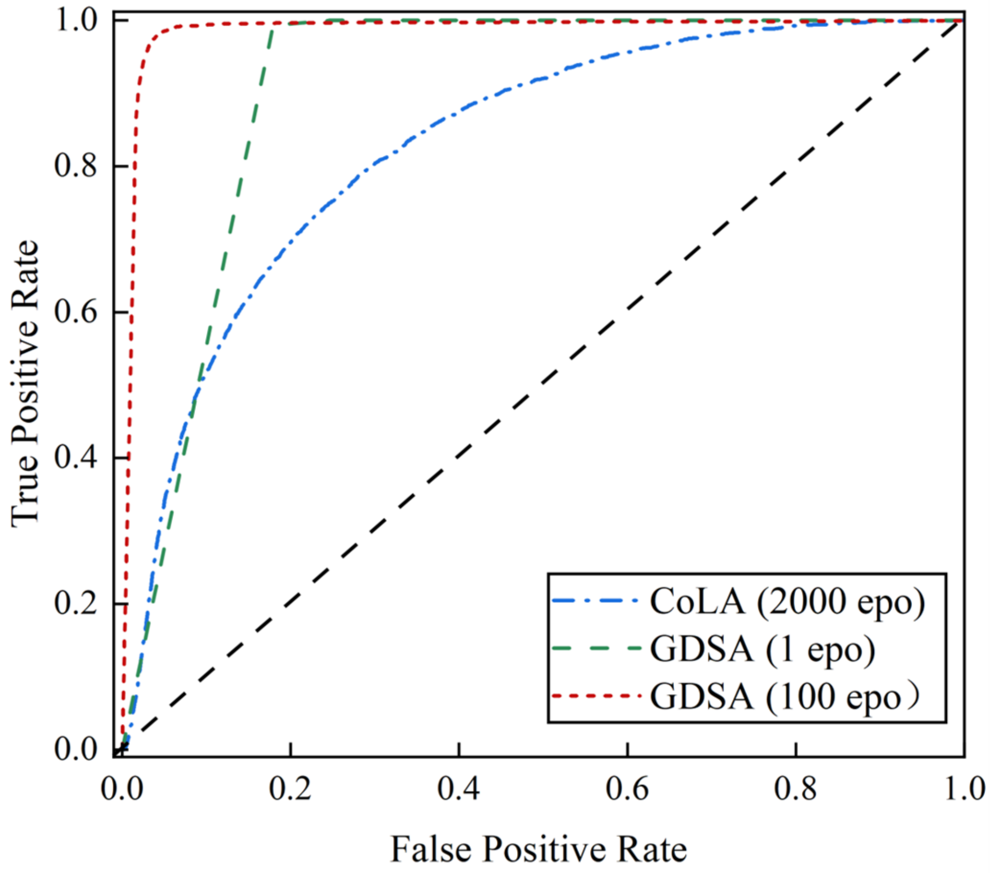
Experiment Results

□ Scalability-- Detection performance and Computational efficiency (ogbn-arxiv)

Methods	AUC	Memory usage	Training time	Testing time
CoLA(2000 epo)	0.8291	8961MB	78h24m25s	7h18m43s
GDSA(1 epo)	0.9096 ↑9.7%	3414MB|61.9%	3m59s|1181.0×	5.1973s|5064.7×
GDSA(100 epo)	0.9867 ↑19.0%	4850MB|45.9%	8h11m32s|9.6×	5.1973s|5064.7×

Experiment Results

□ Scalability-- Detection performance and Computational efficiency (ogbn-arxiv)



Conculsion and Future Works

- ✓ Algorithmic ideas: An efficient self-supervised **group discrimination** paradigm with **structure -Aware** for graph anomaly detection is proposed.
- ✓ Practical significance: GDSA breaks through the barrier of graph anomaly detection that it is difficult to **directly model anomalous structures**.
- ✓ Future Works:
  - **Rethinking the graph anomaly detection** problem from more angles.
  - Explore frameworks that **can detect different types of anomalies in graph anomaly detection** more comprehensively and efficiently.

Rethinking graph anomaly detection: A self-supervised Group Discrimination paradigm with Structure-Aware

Junyi Yan, Enguang Zuo, Chen Chen, Cheng Chen, Jie Zhong, Tianle Li, Xiaoyi Lv*

闫俊依
邮箱：yjy@stu.xju.edu.cn
stu.xju.edu.cn
长按识别二维码，添加联系人至通讯录

Xinjiang University, China
Email: yjy@stu.xju.edu.cn

*Corresponding author

IEEE International Conference on Multimedia and Expo 2023
Brisbane Convention & Exhibition Centre
10-14 July 2023