Case-Based Interpretability in Graph-Level Anomaly Detection via Contrast with Normal Prototypes

Qiuran Zhao, **Kai Ming Ting**, and **Xinpeng Li**

State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing, China

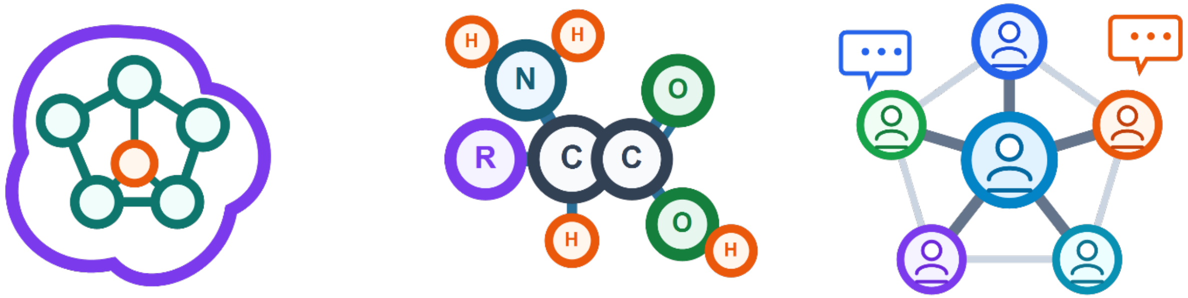
School of Artificial Intelligence, Nanjing University, Nanjing, China

zhaoqr@lamda.nju.edu.cn, tingkm@nju.edu.cn, lixp@lamda.nju.edu.cn

Graph-Level Anomaly Detection

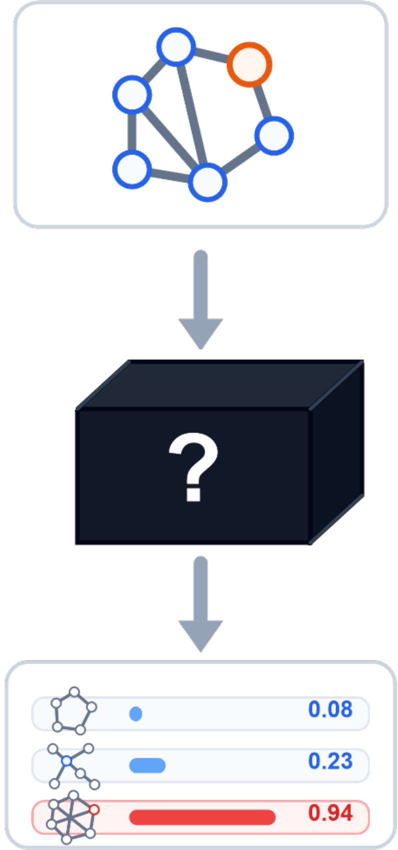
Graph-Level Anomaly Detection (GLAD):

- ➢ Identify graphs that deviate from the majority of graphs in a dataset.



Detection alone is not enough.

The Missing Normal Reference



- **GLAD scores are self-referential:**
  Do not directly show how the anomalous graph differs from normality.

- **Explainable GLAD:**
  - ➢ No normal reference.
  - ➢ Latent prototype only.

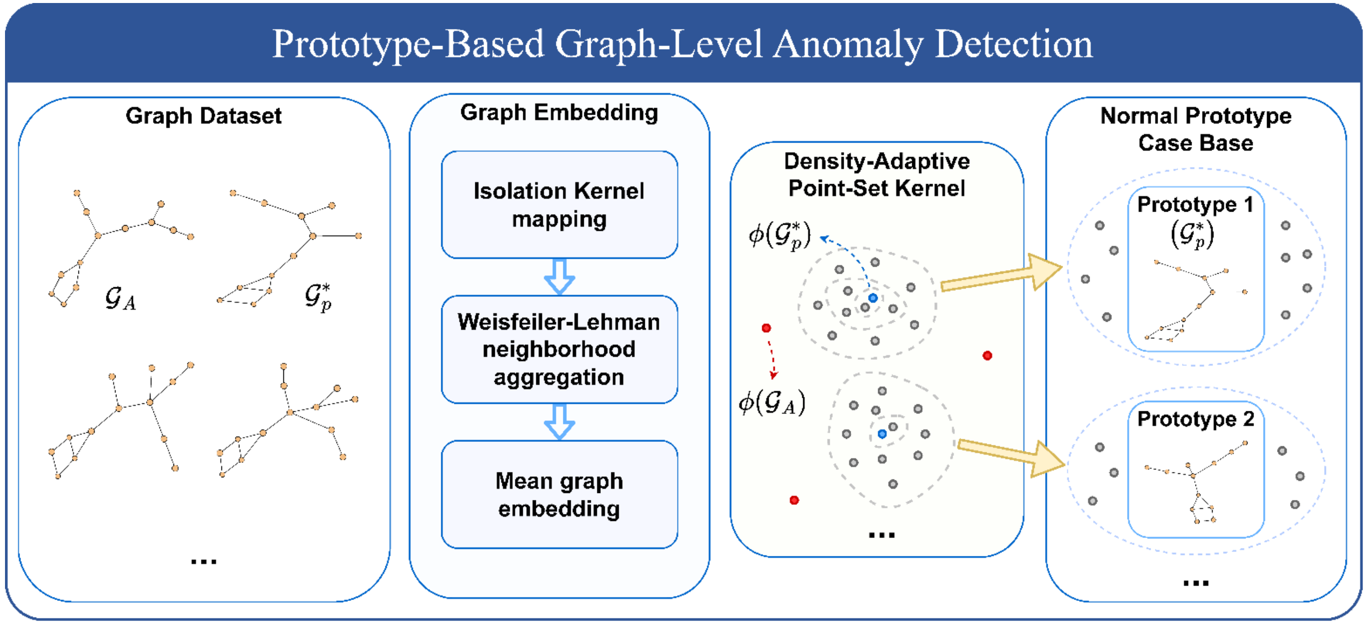
Without a normal reference, an anomaly is hard to verify.

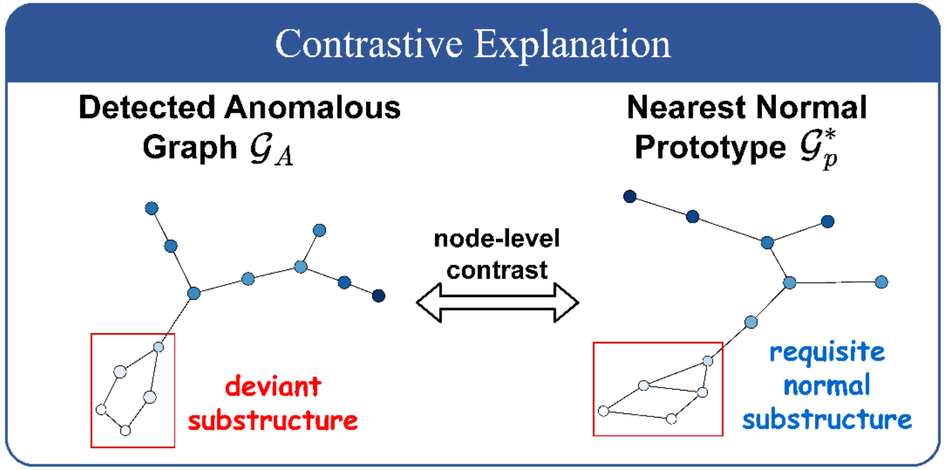
A Case-Based View of GLAD

- **Case base:** Representative normal graphs as prototypes.
- **Retrieve:** Nearest normal prototype & corresponding normal cluster.
- **Compute:** Similarity to normal cluster gives anomaly score.
- **Explain:** Contrast anomaly with prototype.

Anomaly → Dissimilar to all normal cases.

ProtoGLAD

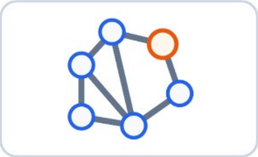




NANJING UNIVERSITY

Case Representation

Node vector $\mathbf{x}_v$ of graph $\mathcal{G}$



Isolation Kernel Feature Mapping

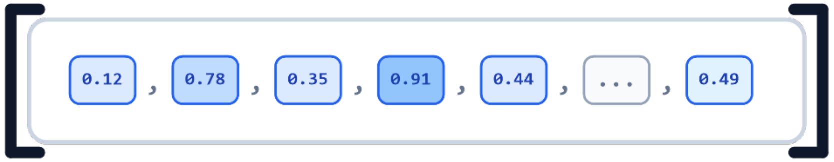
IK feature map $\varphi(\mathbf{x}_v)$

Weisfeiler-Lehman Embedding

WL embedding

$$\varphi^h(\mathbf{x}_v) = [\varphi^{h-1}(\mathbf{x}_v) + \Sigma_{u\in\mathcal{N}(v)}\varphi^{h-1}(\mathbf{x}_u)/|\mathcal{N}(v)|]/2$$

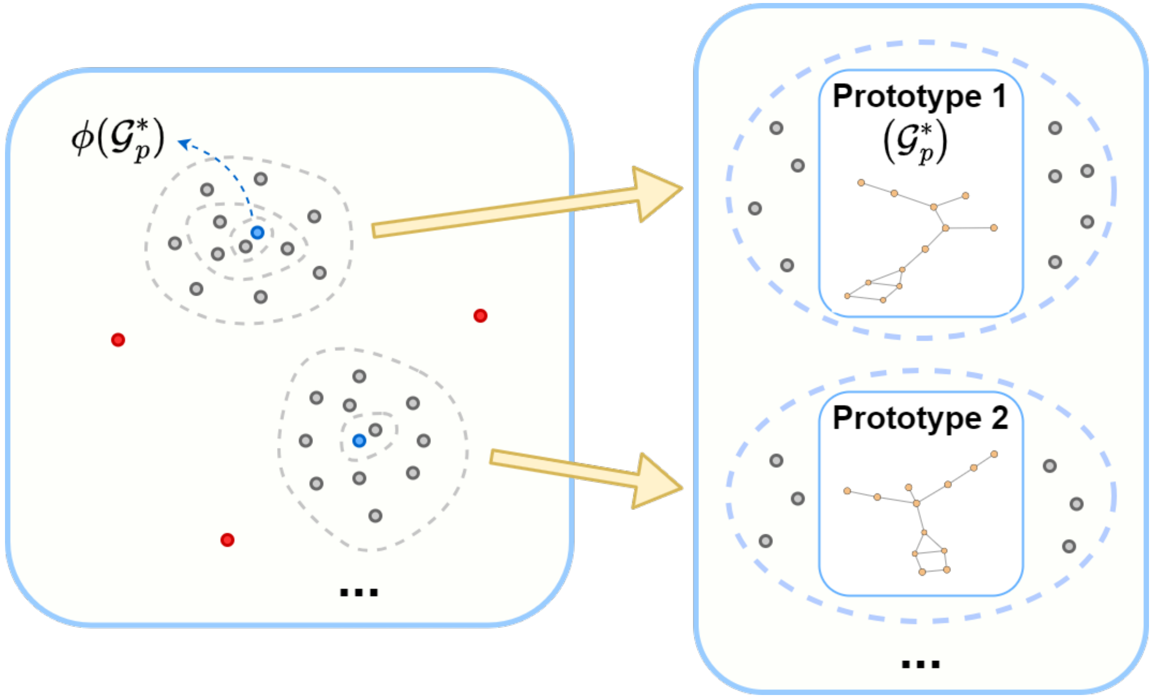
Graph embedding $\phi(\mathcal{G}) = \frac{1}{|\mathcal{V}|}\Sigma_{v\in\mathcal{V}}\,\varphi^h\,(\mathbf{x}_v)$



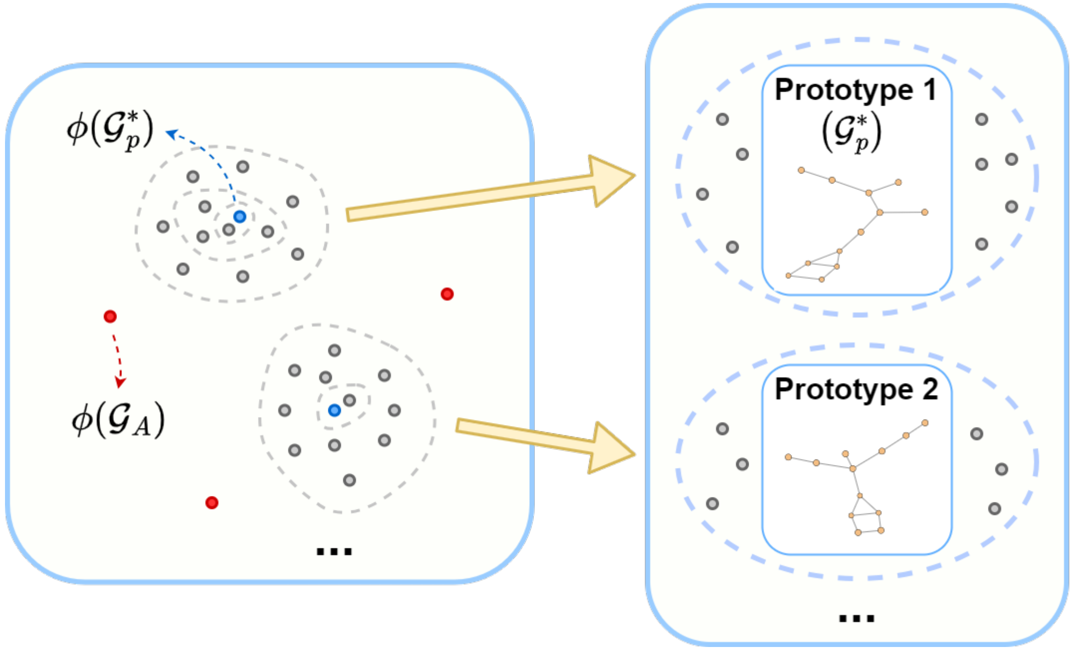
Case Base Construction

Normal cluster & prototype discovery via a density-adaptive point-set kernel

Case base of normal prototype graphs



Anomaly Detection

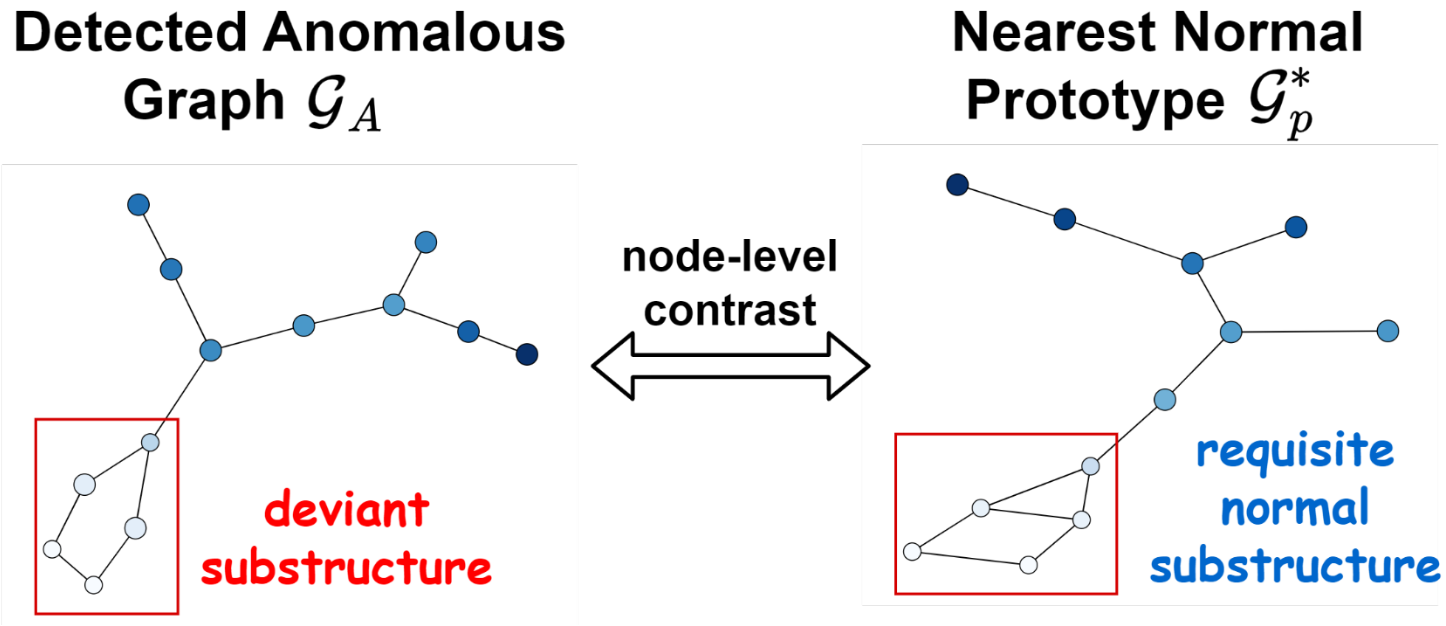


- Anomaly score:

$$s_i = -\max_{j=1,\ldots,k} K\left(\phi(\mathcal{G}_i), C^j\right)$$

- **Normal**: Close to at least one normal cluster.
- **Anomalous**: Dissimilar to all normal clusters.

Contrastive Case-Based Explanation



Node normality score: $c(u) = \langle \varphi^h(u), \phi(\mathcal{G}_p^*) \rangle$

Experimental Setup

- **Datasets:**
  - 10 real-world datasets from the TUDataset benchmark.
  - 3 synthetic datasets with ground-truth anomalous motifs.
- **Baselines:** 8 unsupervised GLAD methods
  - Two-step embedding-then-detection methods: WL-iForest and MUSE.
  - End-to-end deep learning methods: GLADC, GLocalKD, OCGTL, SIGNET*, GLADPro*, and KDGAE.
- **Evaluation metrics:**
  - Graph-level AUC for detection.
  - Node-level and edge-level AUC for explanation.

*Explainable GLAD methods.

Detection Performance

Dataset	WL-iForest	MUSE	GLADC	GLocalKD	SIGNET	OCGTL	GLADPro	KDGAE	ProtoGLAD
AIDS	61.4	**99.8**	96.7	96.9	97.6	97.5	52.6	83.3	<u>98.7</u>
BZR	52.8	68.0	61.1	68.1	<u>77.9</u>	61.8	60.6	55.3	**85.6**
COX2	51.2	<u>67.3</u>	57.7	60.2	64.5	57.9	63.7	55.2	**70.9**
DD	70.4	<u>80.3</u>	57.2	75.3	71.2	78.5	76.6	73.6	**81.1**
MUTAG	71.1	85.1	73.2	83.3	76.6	79.5	<u>89.1</u>	67.1	**89.8**
PROTEINS	62.1	72.3	75.8	<u>76.1</u>	73.3	70.7	73.6	58.2	**79.4**
NCI1	50.4	70.1	68.3	65.3	<u>71.3</u>	58.1	59.2	59.1	**72.1**
DHFR	51.7	<u>62.5</u>	60.4	61.8	**63.4**	51.0	60.9	52.4	62.4
IMDB-B	59.9	63.3	**65.4**	49.9	63.0	60.1	62.1	59.2	<u>64.9</u>
REDDIT-B	55.7	74.4	**82.1**	77.9	71.4	73.9	63.1	58.0	<u>79.1</u>
B-TYPE	78.1	72.6	53.0	75.9	<u>78.6</u>	74.7	69.9	54.8	**86.2**
B-COUNT	<u>88.7</u>	87.7	82.6	75.8	70.0	87.9	66.7	74.0	**91.2**
B-SIZE	76.2	60.0	64.3	74.9	65.7	<u>79.6</u>	69.5	63.6	**88.3**
Avg. Rank	7.00	<u>3.77</u>	5.46	4.54	4.23	5.23	5.69	7.69	**1.38**

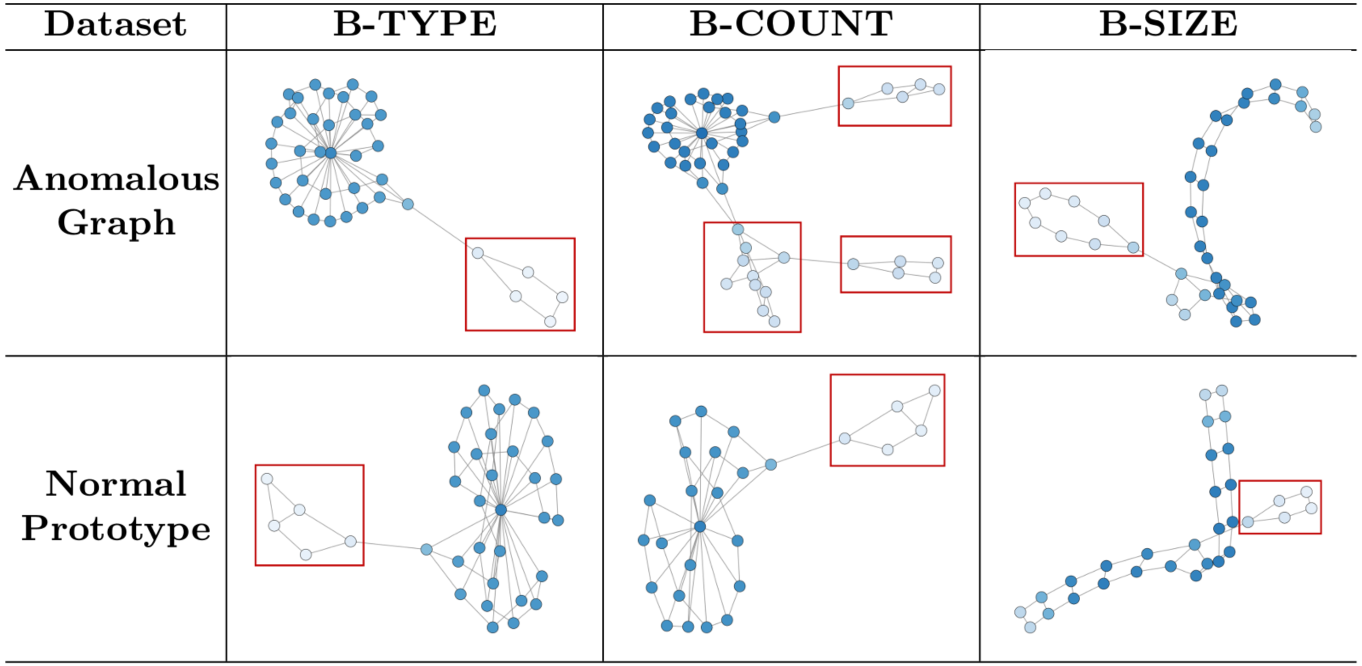
* Graph-level AUC.

NANJING UNIVERSITY

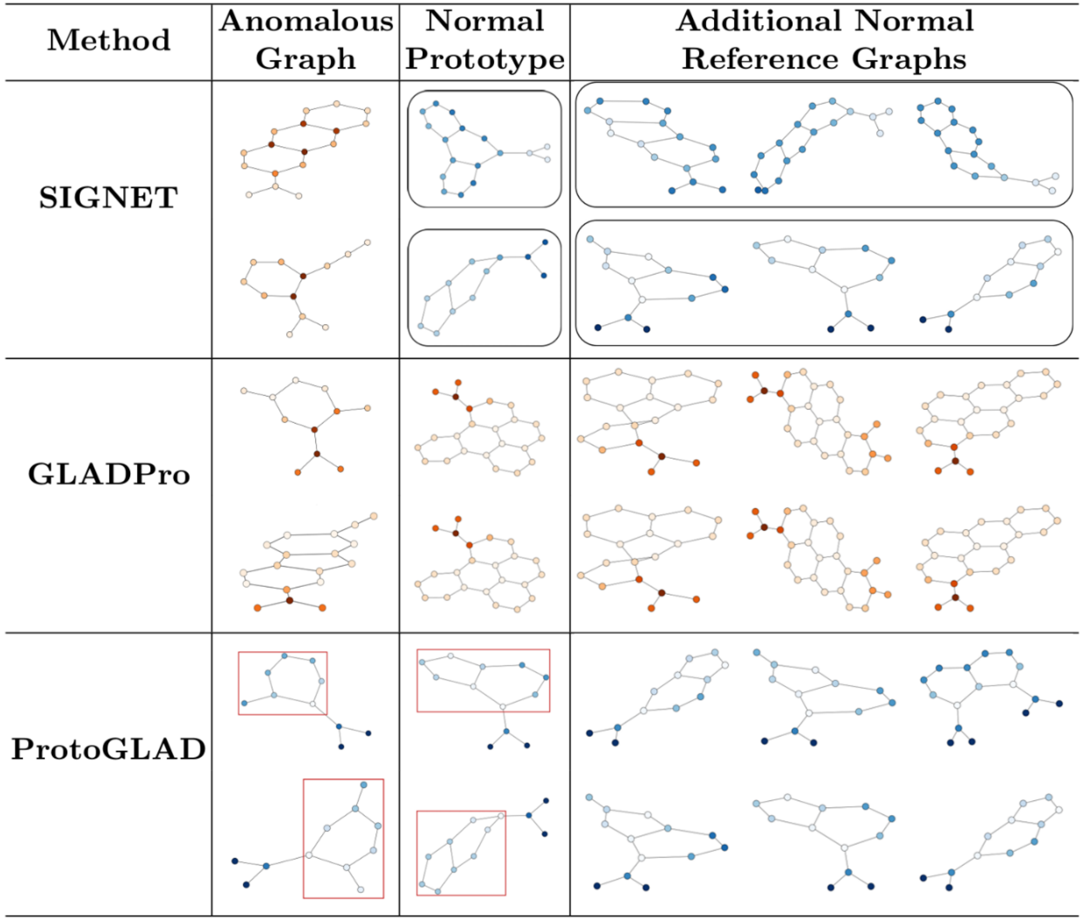
Explanation Quality on Synthetic Data

Method	B-TYPE N-AUC	B-TYPE E-AUC	B-COUNT N-AUC	B-COUNT E-AUC	B-SIZE N-AUC	B-SIZE E-AUC
SIGNET	81.3	83.2	60.9	60.7	82.1	86.4
GLADPro	63.3	58.5	63.0	59.7	68.8	67.1
ProtoGLAD	**88.4**	**91.6**	**64.4**	**67.2**	**91.5**	**94.2**

* Node-level AUC (N-AUC) and edge-level AUC (E-AUC).



Real-World Case-Based Explanations



* **SIGNET:** No native normal reference.
* **GLADPro:** Latent prototype, indirect visualization.
* **ProtoGLAD:** Real prototype, direct contrast.

* For SIGNET which does not identify prototypes, ProtoGLAD is used to provide the normal reference; for GLADPro, the closest graph to its latent prototype is shown.

Ablation Study

Variant	AIDS	BZR	COX2	DD	MUTAG	PROT.	NCI1	DHFR	IMDB	REDDIT	B-TYPE	B-CNT	B-SIZE	Avg.Δ
w/o IK	95.9	67.8	55.9	78.8	86.7	77.2	67.4	60.6	62.1	72.3	85.2	83.7	85.5	−5.4
w/o WL	61.1	79.8	69.7	77.2	72.6	69.6	62.7	61.7	59.5	42.6	80.3	87.7	84.6	−10.8
w/o Proto.	98.7	55.8	46.3	20.3	12.2	32.3	31.7	59.0	51.1	76.8	80.9	87.1	81.5	−24.3
Full	**98.7**	**85.6**	**70.9**	**81.1**	**89.8**	**79.4**	**72.1**	**62.4**	**64.9**	**79.1**	**86.2**	**91.2**	**88.3**	—

* Graph-level AUC.

- The multi-prototype case base is the most critical component; a single center fails to cover multiple normal clusters.

Conclusions

- ProtoGLAD brings case-based reasoning to graph-level anomaly detection.
- Strong detection performance with human-interpretable, contrastive explanations.
- Anomaly detection should ground its judgment in a concrete contrast.

Thanks!

Full paper

Contact me

Presented by **Xinpeng Li**