Directed Graph Contrastive Learning

Zekun Tong, Yuxuan Liang, Henghui Ding, Yongxing Dai, Xinke Li, Changhu Wang

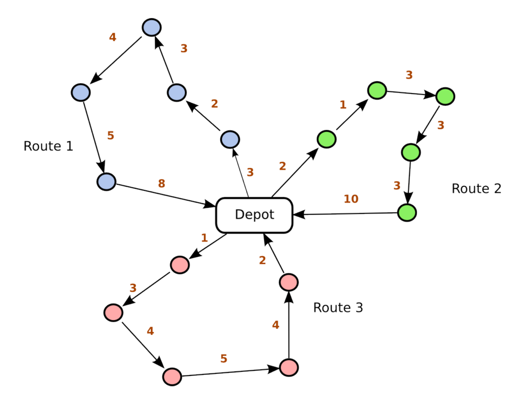
Summary

We present the first contrastive learning framework for learning directed graph representation.

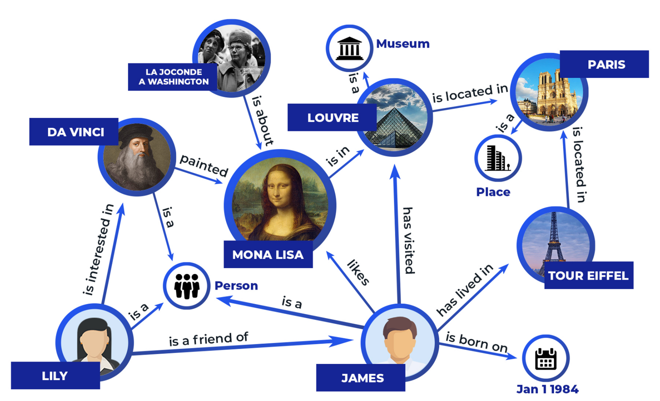
Why we need directed graphs?

Directed structures are everywhere:

- Recommender Systems
- Traffic Forecasting
- Biology (LGT)
- Neuroscience

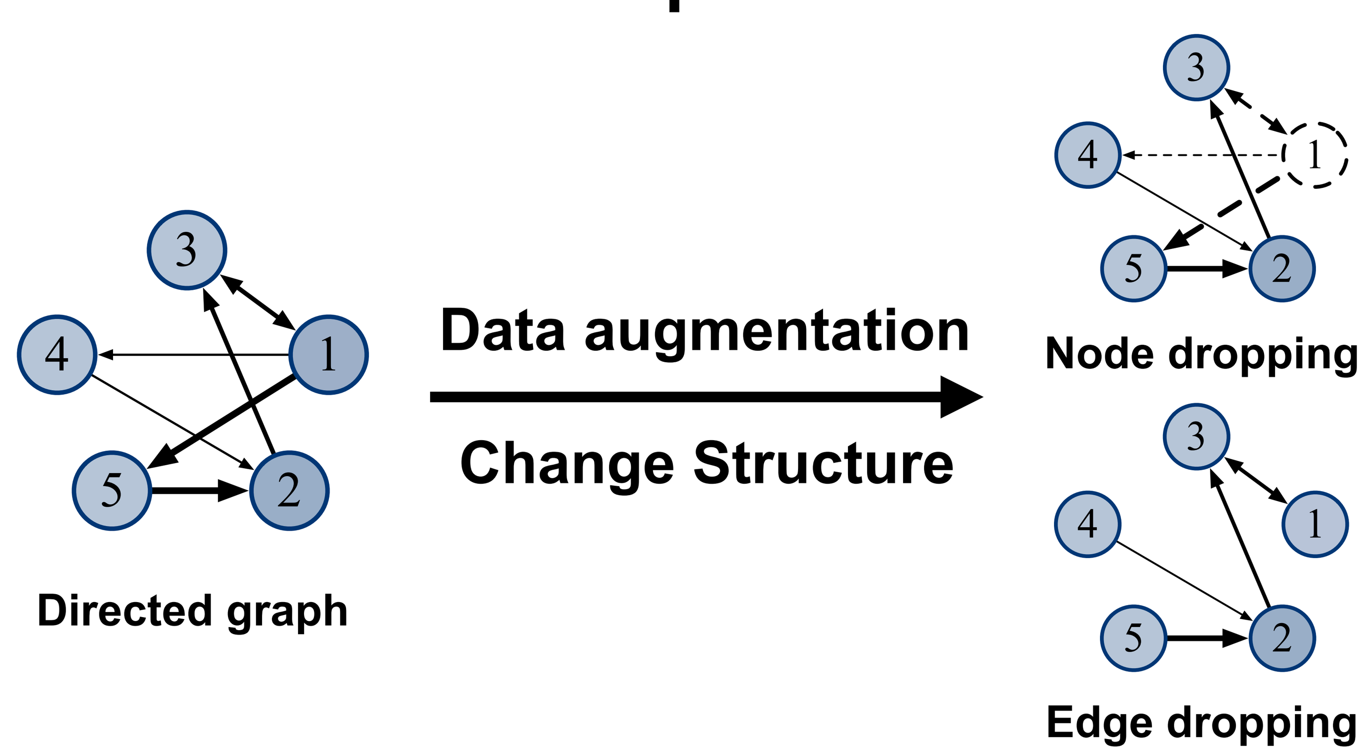


Road network



Knowledge graph

Limitations of Graph Contrastive Learning



Directed graph

Data Augmentation Limits

1. discard distinctive structural information
2. overlooks the discrepancy of nodes and edges

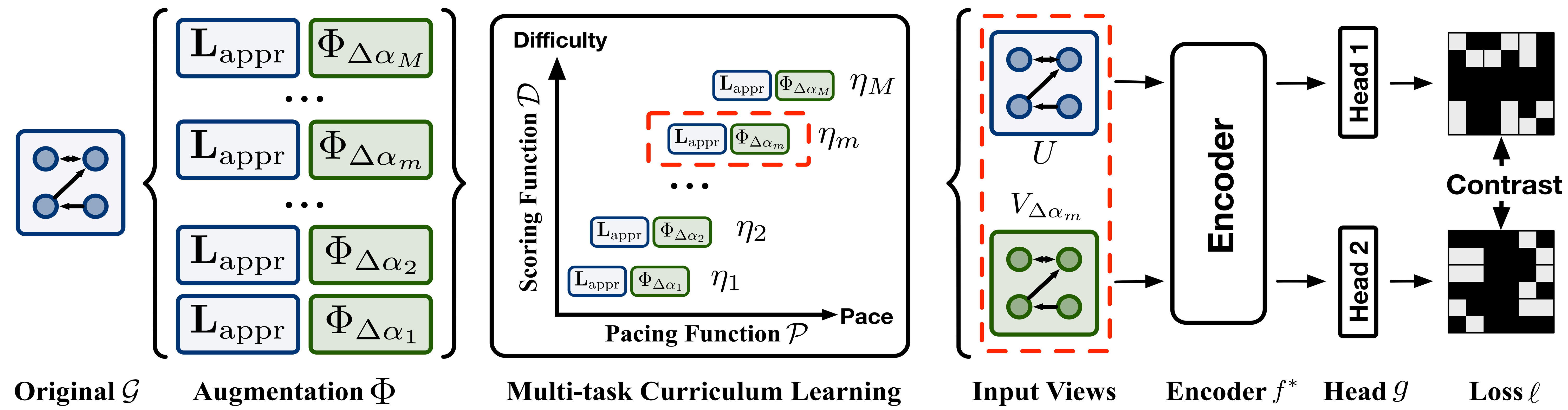
There is a lack of data augmentation methods specifically designed for directed graphs.

Learning Framework Limits

1. inability to take full advantage of data augmentation
2. hand-picking data augmentation parameters

Integrating data augmentation into contrastive learning framework is still at early stage.

Directed Graph Contrastive Learning Framework

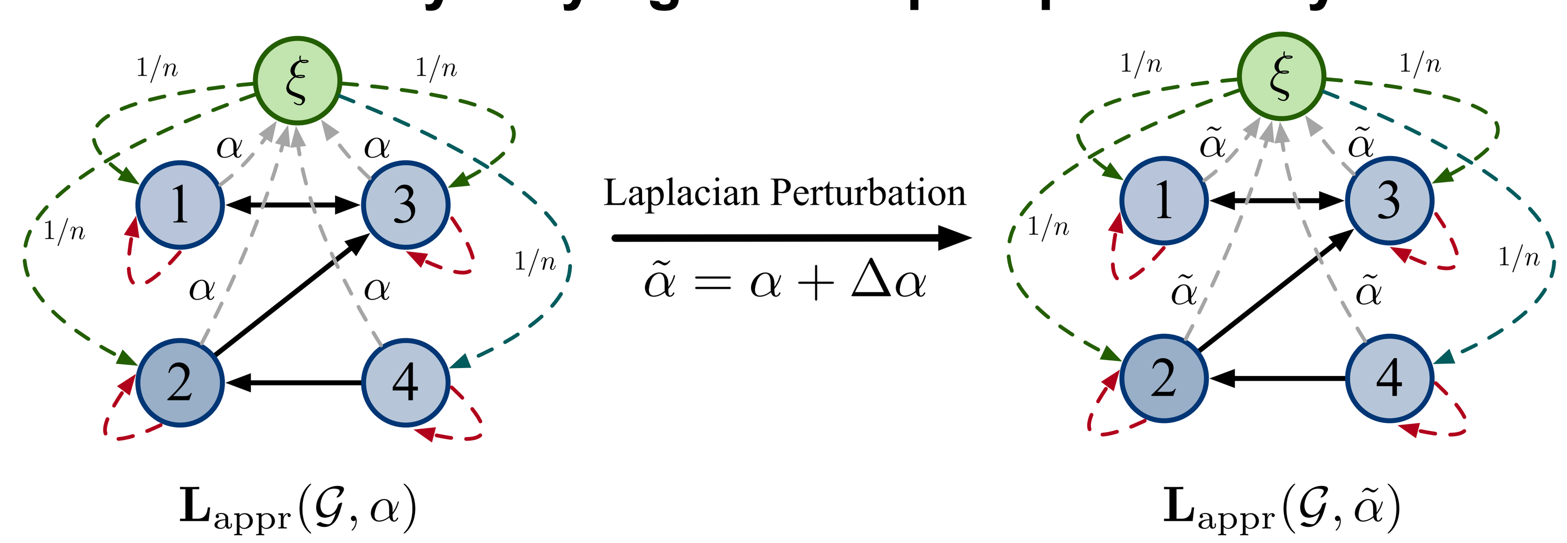


Original $\mathcal{G}$ Augmentation Φ Multi-task Curriculum Learning Input Views Encoder f^* Head g Loss ℓ

1. We first generate M different pairs of contrastive views by **Laplacian perturbation**.
2. The different contrastive view pairs are then scored by a **scoring function** and mapped to different training paces by a **pacing function**.
3. Finally, the arranged contrastive view pairs are input into a **shared encoder** to progressively learn the unsupervised graph representation with contrastive loss.

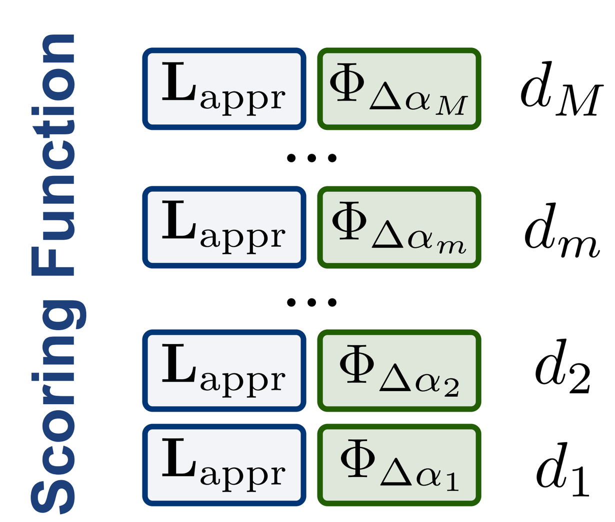
Directed Graph Data Augmentation

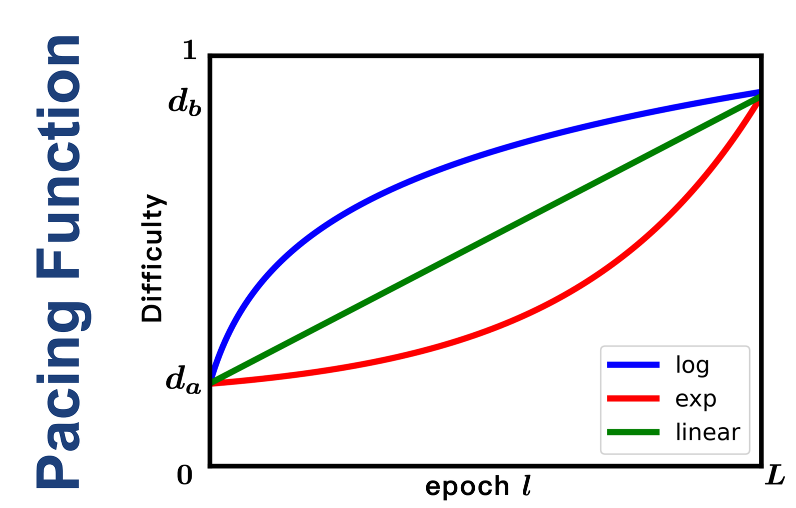
Laplacian Perturbation: changing the Laplacian matrix by varying the teleport probability.



Multi-task Curriculum Learning

Utilizes **prior knowledge** about the difficulty of the learning tasks to learn from **easy-to-difficult** contrastive views.



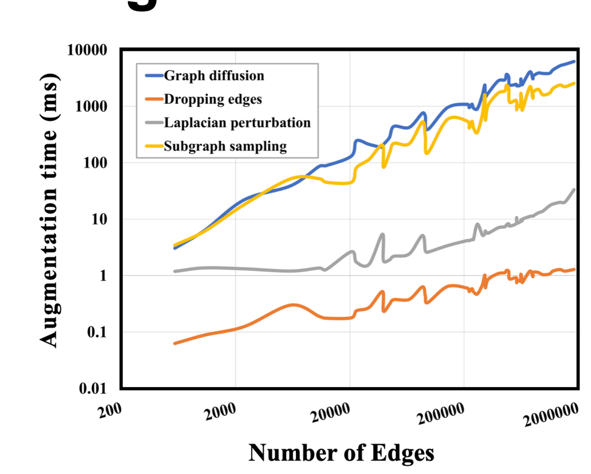


Brief Experimental Results

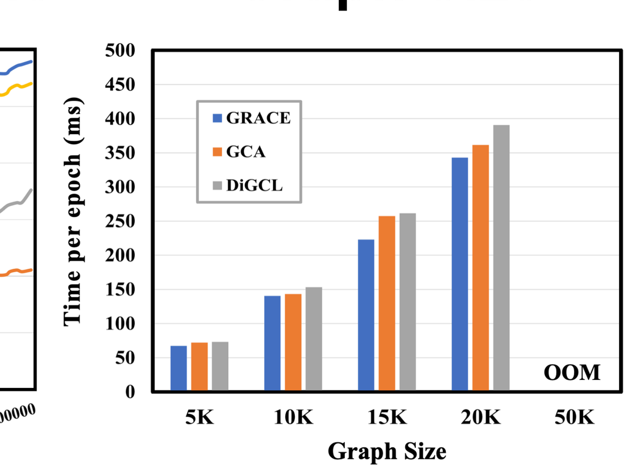
Node classification task in directed graphs

	Method	Directed: Cora-ML	Directed: CiteSeer	Directed: AM-Photo	Undirected: PubMed	Undirected: DBLP
Supervised	GCN [17]	70.92 ± 0.39	63.00 ± 0.45	88.52 ± 0.47	78.78 ± 0.30	73.54 ± 0.77
	GAT [46]	72.22 ± 0.57	63.73 ± 0.57	88.36 ± 1.25	77.49 ± 0.47	76.08 ± 0.54
	APPNP [18]	70.31 ± 0.67	61.63 ± 0.63	87.43 ± 0.98	79.35 ± 0.48	77.92 ± 0.75
	MagNet [64]	76.32 ± 0.10	65.04 ± 0.47	86.80 ± 0.65	74.23 ± 0.46	69.73 ± 0.98
	DiGCN [43]	$\underline{77.03 \pm 0.70}$	64.60 ± 0.60	88.66 ± 0.51	76.79 ± 0.49	73.37 ± 0.72
Unsupervised	DGI[47]	75.21 ± 1.29	64.58 ± 1.78	85.25 ± 0.59	74.11 ± 0.62	76.53 ± 1.24
	GMI[32]	76.59 ± 0.35	63.29 ± 0.70	81.12 ± 0.01	80.27 ± 0.16	76.66 ± 0.48
	MVGRL[13]	76.67 ± 0.12	62.22 ± 0.02	86.15 ± 0.21	79.98 ± 0.04	OOM
	GraphCL [63]	67.34 ± 0.12	57.84 ± 0.11	67.66 ± 0.05	75.29 ± 0.08	77.85 ± 0.22
	GRACE [68]	73.88 ± 0.25	61.20 ± 0.20	87.95 ± 0.32	79.54 ± 0.05	78.03 ± 0.09
	GCA [69]	76.32 ± 0.33	63.25 ± 0.10	87.35 ± 0.27	79.81 ± 0.61	77.83 ± 0.35
	Ours + No Curr	75.86 ± 0.09	66.99 ± 0.54	87.32 ± 0.14	79.57 ± 0.12	78.28 ± 0.05
	Ours + Random	76.52 ± 1.66	$\underline{67.15 \pm 0.82}$	$\underline{89.03 \pm 0.46}$	$\mathbf{80.75 \pm 0.10}$	$\underline{79.58 \pm 0.14}$
	Ours + Anti Curr	76.12 ± 1.04	66.83 ± 1.13	88.83 ± 0.73	80.22 ± 0.37	79.42 ± 0.15
	Ours + Curr	$\mathbf{77.53 \pm 0.14}$	$\mathbf{67.42 \pm 0.14}$	$\mathbf{89.41 \pm 0.11}$	$\underline{80.69 \pm 0.08}$	$\mathbf{79.70 \pm 0.13}$

Augmentation Time



Graph Size



More info