Distance-wise Prototypical Graph Neural Network for Node Imbalance Classification

Yu Wang[1]   Charu Aggarwal[2]   Tyler Derr[1]

Personal website   Wechat

Imbalanced Node Classification and Our Framework

**(a) Class prototype-driven balanced training**

(b) Distance Metric Learning

(c) Imbalanced Label Propagation

If $t = \frac{0.35 - 0.23}{0.29} = 0.41 > \eta$?

Assign pseudo label

(d) Self-supervised Learning

$$\mathcal{L}_{\text{self}_p} = \sum_{i,j} \mathbf{SIM}_{ij} - \text{tr}(\mathbf{SIM})$$

$$\mathcal{L}_{\text{self}_s} = \mathbf{G}^T \mathbf{L} \mathbf{G}$$
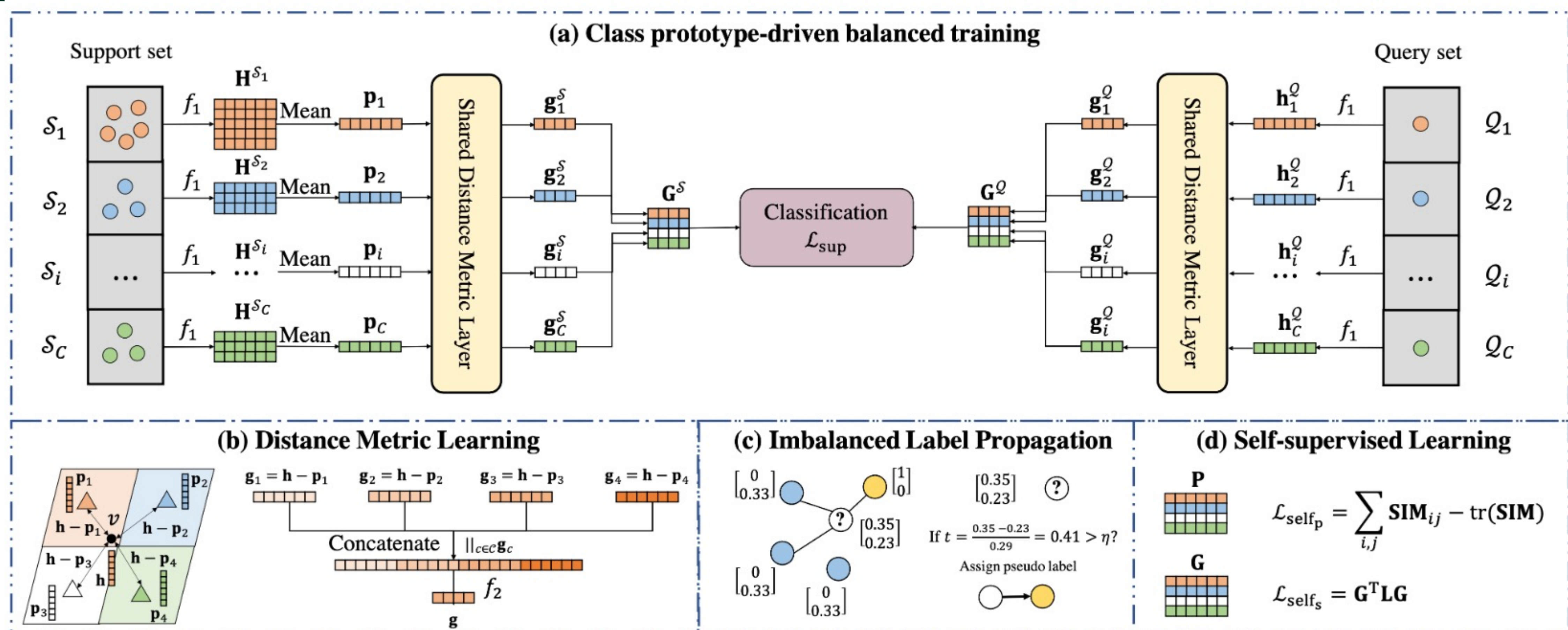
Experimental Verification

Model	Cora (0.81)			Citeseer (0.74)			Pubmed (0.80)			Cora-ML (0.79)		
	F1-macro	F1-weight	F1-micro	F1-macro	F1-weight	F1-micro	F1-macro	F1-weight	F1-micro	F1-macro	F1-weight	F1-micro
GCN	0.5205	0.5195	0.5212	0.3870	0.4169	0.4692	0.5501	0.5569	0.5928	0.5205	0.5195	0.5212
GCN$_{us}$	0.5631	0.5659	0.5727	0.4503	0.4822	0.5220	0.6272	0.6323	0.6451	0.5656	0.5516	0.5611
GCN$_{rw}$	0.5609	0.5660	0.5724	0.4457	0.4800	0.5156	0.6169	0.6178	0.6327	0.5609	0.5660	0.5724
GCN$_{st}$	0.5488	0.5398	0.5519	0.4462	0.4794	0.5127	0.5861	0.5964	0.6186	0.5488	0.5398	0.5519
GraphSMOTE	0.5845	0.6026	0.5820	0.4236	0.4774	0.5020	0.6122	0.5998	0.6110	0.6233	0.6450	0.6130
RECT	0.5234	0.5025	0.5448	0.4002	0.4243	0.4549	0.5713	0.5597	0.6002	0.5530	0.5560	0.6026
DR-GCN	0.5513	0.5362	0.5520	0.3924	0.4414	0.4880	0.5628	0.5730	0.5559	0.5412	0.4716	0.5060
DPGCN	**0.7115**	**0.7029**	**0.7111**	**0.4838**	**0.5180**	**0.5397**	**0.7018**	**0.7176**	**0.7189**	**0.7273**	**0.7278**	**0.7305**

Visualizing the performance of our model under different imbalanced ratio
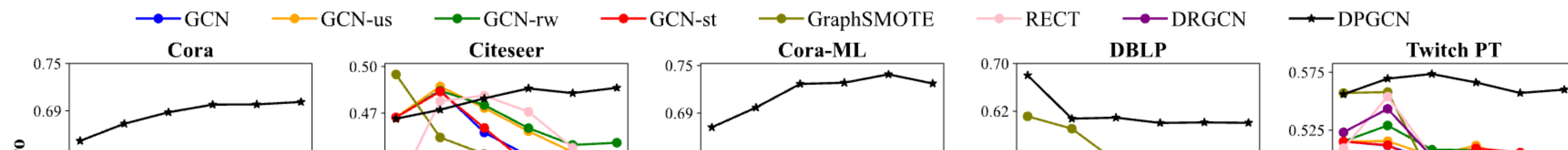
Ablation Study

Imbalanced Label Propagation

(a) Cora   (b) Citeseer
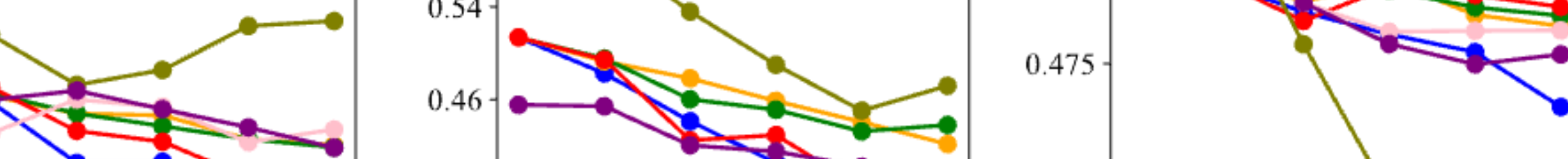
Recent/Future Work

Mitigating Sensitive Leakage (KDD 22 Poster-143)

Structural Explanation for Bias (KDD 22 Poster-79)

Bias Explainer h_Φ → Bias Explanation with edge set $\bar{\varepsilon}_i$

Mitigating Imbalance (CIKM 22)

Graph SSL (Springer 22)

Fairness and Explanation (Coming soon!)