On Generalized Degree Fairness in Graph Neural Networks

Zemin Liu, Trung-Kien Nguyen, Yuan Fang

In Proceeding of 37th AAAI Conference on Artificial Intelligence, February 7-14, 2023

Outline

- **Problem & related work**
- Proposed model: DegFairGNN
- Experiments
- Conclusions

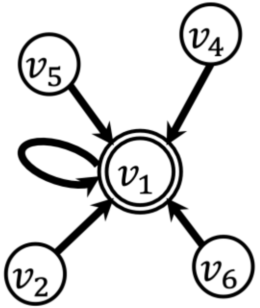
Graph Representation Learning

- Graph embedding approaches
  - DeepWalk [1], node2vec [2], …

- Graph neural networks (GNNs) [3,4,5]

$$\mathbf{h}_v^l = \mathcal{M}(\mathbf{h}_v^{l-1}, \{\mathbf{h}_i^{l-1} : i \in \mathcal{N}_v\}; \theta^l)$$

Message passing function



[1] Perozzi B., et al. 2014. Deepwalk: Online learning of social representations. KDD.
[2] Grover A., et al. 2014. node2vec: Scalable feature learning for networks. KDD.
[3] Kipf, T. N., et al. 2017. Semi-supervised classification with graph convolutional networks. ICLR.
[4] Veličković, P., et al. 2018. Graph attention networks. ICLR.
[5] Hamilton W L., et al. 2017. Inductive representation learning on large graphs. NeurIPS.

Fairness *vs* Degree Fairness in GNNs

- Fairness in GNNs
  - Source
    - Sensitive attributes (e.g., gender, age, or race)
  - Some related methods
    - Filters to debias: CFC [1]
    - Adversarial learning paradigm: [1,2]

- Degree Fairness in GNNs
  - Source
    - Neighborhood structure (degree)
  - Outcome
    - Differential node behaviors and biased outcomes
  - Goal
    - Achieve equitable outcomes for nodes of different degrees

[1] Bose, A., et al. 2019. Compositional fairness constraints for graph embeddings. In ICML, 715-724.
[2] Dai, E., et al. 2019. Say No to the Discrimination: Learning Fair Graph Neural Networks with Limited Sensitive Attribute Information. In WSDM, 680-688.

Degree Fairness in GNNs

- Our goal
  - Degree fairness in graph neural networks
- Non-trivial
  - Neighborhood sampling
    - Cannot address this issue
    - Information loss
  - Degree *vs* sensitive attributes
    - Prior works do not debias neighborhood aggregation

Problem Formulation (1): *generalized degree*

- Generalized degree
  - Source of degree bias
    - Neighborhood & local context
  - Local context

$$\mathcal{C}_r = \{v' \in \mathcal{V} \mid d(v, v') \leq r\}$$

  - Generalized degree

$$\deg_r(v) = [\mathbf{A}^r \mathbf{1}]_v$$

Problem Formulation (2): *generalized degree fairness*

- Generalized degree fairness (Statistical Parity [1] and Equal Opportunity [2])
  - Generalized degree groups
    - *m* groups: $\mathcal{G}_1, \ldots, \mathcal{G}_m$

$$\mathcal{G}_i = \{v \in \mathcal{V} \mid d_i \leq \deg_r(v) < d_{i+1}\}$$

$$d_1 < d_2 < \ldots < d_{m+1}$$

$$d_1 = \min_{v \in \mathcal{V}} \deg_r(v)$$

$$d_{m+1} = \max_{v \in \mathcal{V}} \deg_r(v) + 1$$

  - Generalized degree metrics
    - Degree Statistical Parity (DSP)

$$P(\hat{y}_v = y | v \in \mathcal{G}_i) = P(\hat{y}_v = y | v \in \mathcal{G}_j)$$

    - Degree Equal Opportunity (DEO)

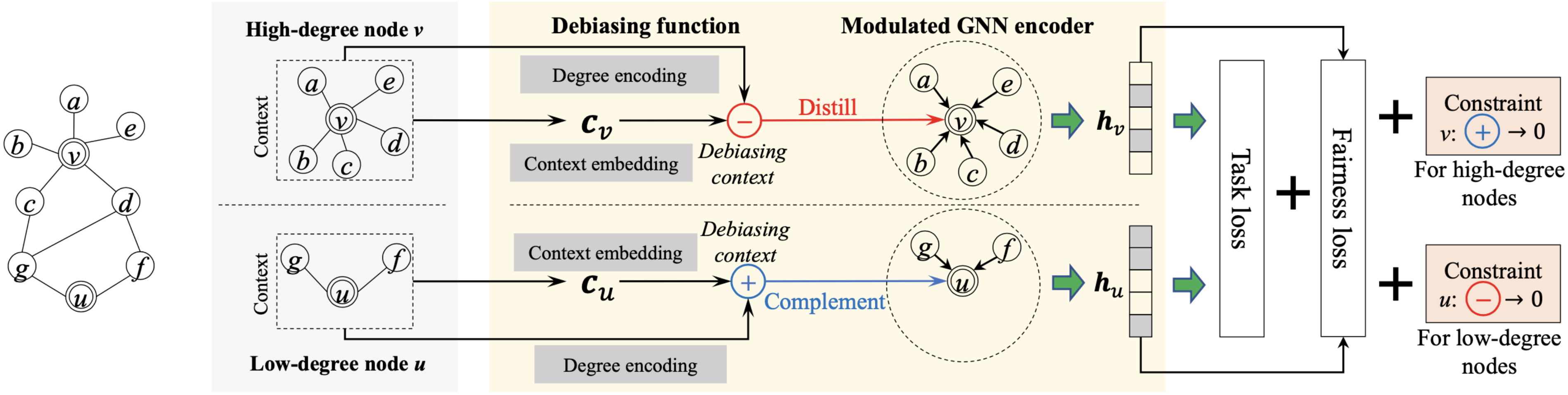
$$P(\hat{y}_v = y | y_v = y, v \in \mathcal{G}_i) = P(\hat{y}_v = y | y_v = y, v \in \mathcal{G}_j)$$

[1] Dwork, C., et al. 2012. Fairness through awareness. In ITCS, 214-226.
[2] Hardt, M., et al. 2016. Equality of Opportunity in Supervised Learning. In NeurIPS, 3315-3323.

Outline

- Problem & related work
- **Proposed model: DegFairGNN**
- Experiments
- Conclusions

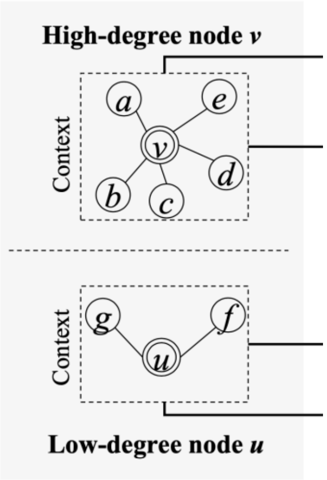
DegFairGNN: overall framework



(a) Toy graph (b) Structural contrast (c) Debiasing neighborhood aggregation (d) Overall objective

Structural Contrast

- Contrastive strategy
  - Low-degree nodes $\mathcal{S}_0$ and high-degree nodes $\mathcal{S}_1$
  - Neighborhood aggregation
    - Only access its one-hop contexts in each layer
  - Definition of the two groups



(b) Structural contrast

$$\mathcal{S}_0 = \{v \in \mathcal{V} \mid \deg_1(v) \leqslant K\}$$

threshold

$$\mathcal{S}_1 = \mathcal{V} \setminus \mathcal{S}_0$$

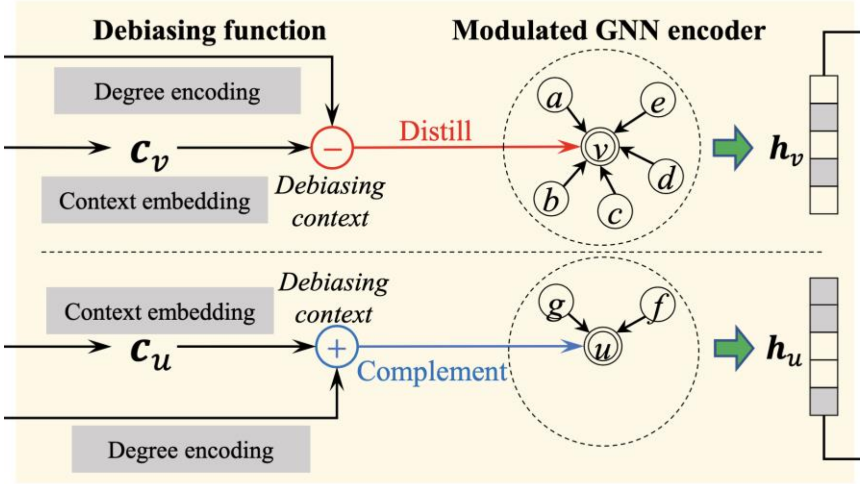
Debiasing Neighborhood Aggregation (1)

- Debiasing function
  - Low-degree node

$$u \in \mathcal{S}_0,\ \mathcal{D}(u;\theta_0^l) \in \mathbb{R}^{d_l}$$

  - High-degree node

$$v \in \mathcal{S}_1,\ \mathcal{D}(v;\theta_1^l) \in \mathbb{R}^{d_l}$$



(c) Debiasing neighborhood aggregation

- Requirement of debiasing function
  - Comprehensiveness
  - Adaptiveness

Debiasing Neighborhood Aggregation (2)

- Comprehensiveness
  - Context embedding

$$\mathbf{c}_v^l = \textsc{Pool}(\{\mathbf{h}_u^{l-1} \mid u \in \mathcal{C}_r(v)\})$$

  - Debiasing function

$$\mathcal{D}(v; \theta_*^l) = f(\mathbf{c}_v^l; \theta_{c,*}^l)$$

- Adaptiveness

$$\mathcal{D}(v; \theta_*^l) = (\gamma_v^l + \mathbf{1}) \odot f(\mathbf{c}_v^l; \theta_{c,*}^l) + \beta_v^l$$

$$\gamma_v^l = \phi_\gamma(\delta^l(v); \theta_\gamma^l), \quad \beta_v^l = \phi_\beta(\delta^l(v); \theta_\beta^l)$$

Degree encoding

  - Degree encoding

$$[\delta^l(v)]_{2i} = \sin(\deg_1(v)/10000^{2i/d_l})$$
$$[\delta^l(v)]_{2i+1} = \cos(\deg_1(v)/10000^{2i/d_l})$$

Debiasing Neighborhood Aggregation (3)

- Modulated GNN Encoder

$$\mathbf{h}_v^l = \sigma\Bigg(\text{AGGR}(\{\mathbf{h}_u^{l-1} \mid u \in \mathcal{N}_v\}; \omega^l)$$
$$+\, \epsilon \cdot \big(\underbrace{I(v \in \mathcal{S}_0)\mathcal{D}(v; \theta_0^l)}_{\text{complement low-deg. group}} + \underbrace{I(v \in \mathcal{S}_1)\mathcal{D}(v; \theta_1^l)}_{\text{distill high-deg. group}}\big)\Bigg)$$

Control the impact

Training Constraints and Objective

- Node classification loss $\mathcal{L}_1 = -\sum_{v\in\mathcal{V}^{\text{tr}}}\sum_{y=1}^{|\mathcal{Y}|}[\mathbf{y}_v]_y \ln[\mathbf{h}_v^{\ell}]_y$
- Fairness loss

$$\mathcal{L}_2 = \left\|\frac{1}{|\mathcal{S}_0^{\text{tr}}|}\sum_{v\in\mathcal{S}_0^{\text{tr}}}\mathbf{h}_v^{\ell} - \frac{1}{|\mathcal{S}_1^{\text{tr}}|}\sum_{v\in\mathcal{S}_1^{\text{tr}}}\mathbf{h}_v^{\ell}\right\|_2^2$$

- Constraints on debiasing contexts

$$\mathcal{L}_3 = \sum_{l=1}^{\ell}\left(\sum_{v\in\mathcal{S}_0^{\text{tr}}}\|\mathcal{D}(v;\theta_1^l)\|_2^2 + \sum_{v\in\mathcal{S}_1^{\text{tr}}}\|\mathcal{D}(v;\theta_0^l)\|_2^2\right)$$

- Constraints on scaling and shifting

$$\mathcal{L}_4 = \sum_{l=1}^{\ell}\sum_{v\in\mathcal{V}^{\text{tr}}}(\|\gamma_v^l\|_2^2 + \|\beta_v^l\|_2^2)$$

- Overall loss

$$\mathcal{L} = \mathcal{L}_1 + \mu\mathcal{L}_2 + \lambda(\mathcal{L}_3 + \mathcal{L}_4)$$

Outline

- Problem & related work
- Proposed model: DegFairGNN
- **Experiments**
- Conclusions

Experimental setup

Datasets

Dataset	Nodes	Edges	Features	Classes
Chameleon	2,277	31,371	2,325	5 (traffic volume)
Squirrel	5,201	198,353	2,089	5 (traffic volume)
EMNLP	2,600	7,969	8	2 (citation count)

Baselines

- Degree-specific models:
  - DSGCN [4], Residual2Vec [5], Tail-GNN [6]
- Fairness-aware models:
  - FairWalk [7], CFC [8], FairGNN [9], FairAdj [10], FairVGNN [11]

Base GNN models

- GCN [1]
- GAT [2]
- GraphSAGE [3]

[1] Kipf, T. N., et al. 2017. Semi-supervised classification with graph convolutional networks. ICLR.
[2] Veličković, P., et al. 2018. Graph attention networks. ICLR.
[3] Hamilton W L., et al. 2017. Inductive representation learning on large graphs. NeurIPS.
[4] Tang, X., et al. 2020. Investigating and Mitigating Degree-Related Biases in Graph Convoltuional Networks. CIKM.
[5] Kojaku, S., et al. 2021. Residual2Vec: Debiasing graph embedding with random graphs. NeurIPS.
[6] Liu, Z., et al. 2021. Tail-GNN: Tail-Node Graph Neural Networks. KDD.
[7] Rahman, T. A., et al. 2019. Fairwalk: Towards Fair Graph Embedding. IJCAI.
[8] Bose, A., et al. 2019. Compositional fairness constraints for graph embeddings. In ICML, 715-724.
[9] Dai, E., et al. 2019. Say No to the Discrimination: Learning Fair Graph Neural Networks with Limited Sensitive Attribute Information. In WSDM, 680-688.
[10] Li, P., et al. 2021. On dyadic fairness: Exploring and mitigating bias in graph connections. ICLR.
[11] Wang, Y., et al. 2022. Improving Fairness in Graph Neural Networks via Mitigating Sensitive Attribute Leakage. KDD.

Experimental setup

Evaluation Metrics for fairness

$$\Delta_{\text{DSP}} = \frac{1}{|\mathcal{Y}|} \sum_{y \in \mathcal{Y}} \Big| P(\hat{y}_v = y | v \in \mathcal{G}_0) - P(\hat{y}_v = y | v \in \mathcal{G}_1) \Big|,$$

$$\Delta_{\text{DEO}} = \frac{1}{|\mathcal{Y}|} \sum_{y \in \mathcal{Y}} \Big| P(\hat{y}_v = y | y_v = y, v \in \mathcal{G}_0) - P(\hat{y}_v = y | y_v = y, v \in \mathcal{G}_1) \Big|.$$

Model Accuracy and Fairness

Table 2: Comparison with baselines ($r = 1$, 20% Top/Bottom).

Henceforth, tabular results are in percent with standard deviation over 5 runs; the best fairness result is **bolded** and the runner-up is underlined.

		GCN	DSGCN	Residual2Vec	Tail-GNN	FairWalk	CFC	FairGNN	FairAdj	FairVGNN	DegFairGCN
Chamel.	Acc. $\uparrow$	62.45 ± 0.21	63.90 ± 1.28	49.04 ± 0.01	66.08 ± 0.19	56.36 ± 0.75	63.02 ± 0.84	70.70 ± 0.52	51.71 ± 1.13	72.32 ± 0.50	69.91 ± 0.19
	$\Delta_{\text{DSP}} \downarrow$	9.68 ± 1.37	8.81 ± 1.15	14.52 ± 0.69	8.51 ± 1.72	8.18 ± 0.93	10.12 ± 1.28	7.33 ± 1.09	9.79 ± 1.91	8.86 ± 1.11	**5.85** ± 0.32
	$\Delta_{\text{DEO}} \downarrow$	36.08 ± 2.65	25.14 ± 2.67	37.31 ± 1.99	26.09 ± 3.25	22.89 ± 2.75	29.54 ± 1.95	26.83 ± 1.95	27.48 ± 2.06	26.02 ± 2.39	**21.60** ± 0.71
Squirrel	Acc. $\uparrow$	47.85 ± 1.33	40.71 ± 2.17	28.47 ± 0.01	42.62 ± 0.06	37.68 ± 0.65	45.64 ± 2.19	57.29 ± 0.77	35.18 ± 1.22	46.97 ± 0.48	59.21 ± 0.97
	$\Delta_{\text{DSP}} \downarrow$	13.37 ± 2.83	16.08 ± 0.86	25.11 ± 0.48	18.91 ± 0.26	**7.94** ± 0.36	12.40 ± 0.48	12.96 ± 1.03	16.63 ± 1.56	26.67 ± 0.52	9.54 ± 1.02
	$\Delta_{\text{DEO}} \downarrow$	27.00 ± 3.79	32.61 ± 3.74	34.49 ± 0.72	33.60 ± 0.72	17.12 ± 1.50	21.60 ± 2.69	17.62 ± 2.40	27.54 ± 1.73	35.80 ± 1.76	**16.42** ± 1.38
EMNLP	Acc. $\uparrow$	78.92 ± 0.43	82.19 ± 0.77	80.69 ± 0.01	83.72 ± 0.28	82.23 ± 0.18	80.15 ± 1.13	86.81 ± 0.22	76.50 ± 1.55	84.03 ± 0.34	79.92 ± 0.77
	$\Delta_{\text{DSP}} \downarrow$	44.55 ± 1.90	50.00 ± 2.98	12.90 ± 0.15	41.18 ± 1.58	33.52 ± 1.46	56.60 ± 1.95	58.23 ± 1.44	40.38 ± 4.64	43.92 ± 1.43	**12.38** ± 3.72
	$\Delta_{\text{DEO}} \downarrow$	34.05 ± 3.56	46.92 ± 2.91	11.26 ± 0.67	36.76 ± 1.48	30.67 ± 1.42	45.21 ± 2.27	51.56 ± 1.38	41.89 ± 4.78	40.95 ± 1.71	**8.52** ± 2.26

Table 3: Comparison to baselines ($r = 2$, 20% Top/Bottom).

		GCN	FairWalk	FairGNN	DegFairGCN
Chamel.	Acc. $\uparrow$	62.45 ± 0.21	56.36 ± 0.75	70.70 ± 0.52	69.91 ± 0.19
	$\Delta_{\text{DSP}} \downarrow$	5.96 ± 0.89	10.38 ± 0.85	6.70 ± 0.32	**5.25** ± 0.39
	$\Delta_{\text{DEO}} \downarrow$	26.92 ± 2.09	25.46 ± 1.66	23.66 ± 0.93	**19.05** ± 0.74
Squirrel	Acc. $\uparrow$	47.85 ± 1.33	37.68 ± 0.65	57.29 ± 0.77	59.21 ± 0.97
	$\Delta_{\text{DSP}} \downarrow$	14.61 ± 2.63	9.64 ± 0.50	11.11 ± 0.93	**8.26** ± 0.57
	$\Delta_{\text{DEO}} \downarrow$	28.62 ± 3.89	17.37 ± 1.10	16.29 ± 2.07	**14.95** ± 1.22
EMNLP	Acc. $\uparrow$	78.92 ± 0.43	82.23 ± 0.18	86.81 ± 0.22	79.92 ± 0.77
	$\Delta_{\text{DSP}} \downarrow$	45.03 ± 1.77	34.80 ± 1.26	52.88 ± 1.39	**10.87** ± 4.00
	$\Delta_{\text{DEO}} \downarrow$	34.71 ± 3.31	31.11 ± 1.34	45.78 ± 1.36	**8.72** ± 2.17

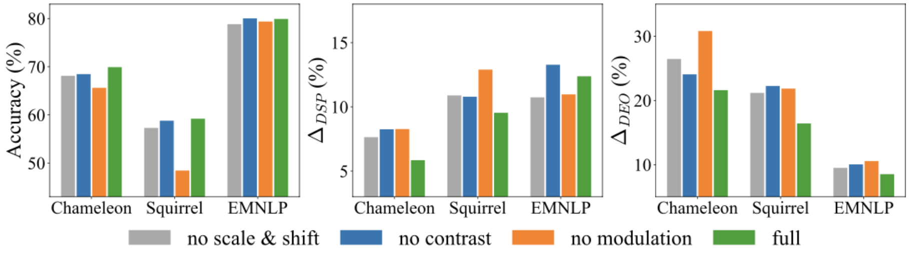
Table 4: Comparison to baselines ($r = 1$, 30% Top/Bottom).

		GCN	FairWalk	FairGNN	DegFairGCN
Chamel.	Acc. $\uparrow$	62.45 ± 0.21	56.36 ± 0.75	70.70 ± 0.52	69.91 ± 0.19
	$\Delta_{\text{DSP}} \downarrow$	5.95 ± 1.02	8.16 ± 0.38	6.92 ± 0.29	**4.15** ± 0.02
	$\Delta_{\text{DEO}} \downarrow$	18.00 ± 1.76	16.65 ± 1.32	14.52 ± 1.09	**8.39** ± 0.37
Squirrel	Acc. $\uparrow$	47.85 ± 1.33	37.68 ± 0.65	57.29 ± 0.77	59.21 ± 0.97
	$\Delta_{\text{DSP}} \downarrow$	10.34 ± 2.15	**6.17** ± 0.36	9.27 ± 0.68	7.39 ± 0.63
	$\Delta_{\text{DEO}} \downarrow$	22.62 ± 3.10	**14.97** ± 1.12	17.42 ± 1.11	17.71 ± 1.05
EMNLP	Acc. $\uparrow$	78.92 ± 0.43	82.23 ± 0.18	86.81 ± 0.22	79.92 ± 0.77
	$\Delta_{\text{DSP}} \downarrow$	42.87 ± 1.40	34.19 ± 0.91	48.25 ± 1.97	**14.46** ± 3.35
	$\Delta_{\text{DEO}} \downarrow$	37.89 ± 3.27	34.49 ± 0.91	48.83 ± 1.97	**10.92** ± 2.87

Model Analysis

Table 5: With other base GNNs ($r = 1$, 20% Top/Bottom).

		GAT	DegFairGAT	GraphSAGE	DegFairSAGE
Chamel.	Acc. ↑	63.15 ± 0.40	69.64 ± 0.44	53.15 ± 0.56	60.95 ± 0.84
	Δ_{DSP} ↓	9.35 ± 1.61	**7.88** ± 1.30	10.86 ± 0.74	**8.22** ± 1.22
	Δ_{DEO} ↓	29.59 ± 1.43	**26.12** ± 2.06	29.42 ± 1.57	**26.40** ± 2.32
Squirrel	Acc. ↑	41.44 ± 0.21	45.55 ± 1.44	34.39 ± 0.62	34.63 ± 1.31
	Δ_{DSP} ↓	12.60 ± 0.77	**12.03** ± 0.63	5.39 ± 0.66	**3.76** ± 0.23
	Δ_{DEO} ↓	24.89 ± 0.69	**20.64** ± 3.06	17.13 ± 2.86	**14.91** ± 1.35
EMNLP	Acc. ↑	70.42 ± 0.77	81.57 ± 1.14	83.96 ± 0.31	83.57 ± 0.44
	Δ_{DSP} ↓	24.40 ± 3.06	**14.11** ± 6.28	56.33 ± 1.12	**28.43** ± 3.79
	Δ_{DEO} ↓	**8.36** ± 1.29	12.28 ± 6.19	51.71 ± 0.88	**24.65** ± 3.35



(a) Accuracy ↑ (b) Δ_{DSP} ↓ (c) Δ_{DEO} ↓

Figure 2: Ablation study on the effect of each module.

Additional Base GNNs

– Outperform their corresponding base GNNs

Ablation Study

– Without scaling and shifting
  - Worse accuracy and fairness
– Without structural contrast
  - Fairness generally become worse
– Without the modulation of aggregation
  - Fairness become worse in most cases

Outline

- Problem & related work
- Proposed model: DegFairGNN
- Experiments
- **Conclusions**

Conclusions

- **Problem**
  - Degree fairness in graph neural networks
- **Proposed model: DegFairGNN**
  - Target the root of degree bias
    - modulating the core operation of neighborhood aggregation through a structural contrast
  - Flexibly work with most modern GNNs
- **Experiments**

Thanks!

Paper, code, data…
www.yfang.site

On Generalized Degree Fairness in Graph Neural Networks

Zemin Liu, Trung-Kien Nguyen, Yuan Fang

In Proceeding of 37th AAAI Conference on Artificial Intelligence, February 7-14, 2023