On Generalized Degree Fairness in Graph Neural Networks

SMU Singapore Management University | School of Computing and Information Systems | NUS National University of Singapore

Zemin Liu, Trung-Kien Nguyen, Yuan Fang

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

Motivation

Problem

Degree Fairness in GNNs

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

Problem Formulation

Generalized degree

- Source of degree bias: Neighborhood and local context
- Local context: $\mathcal{C}_r = \{v' \in \mathcal{V} \mid d(v, v') \le r\}$
- Generalized degree $\deg_r(v) = [\mathbf{A}^r \mathbf{1}]_v$

Generalized degree fairness

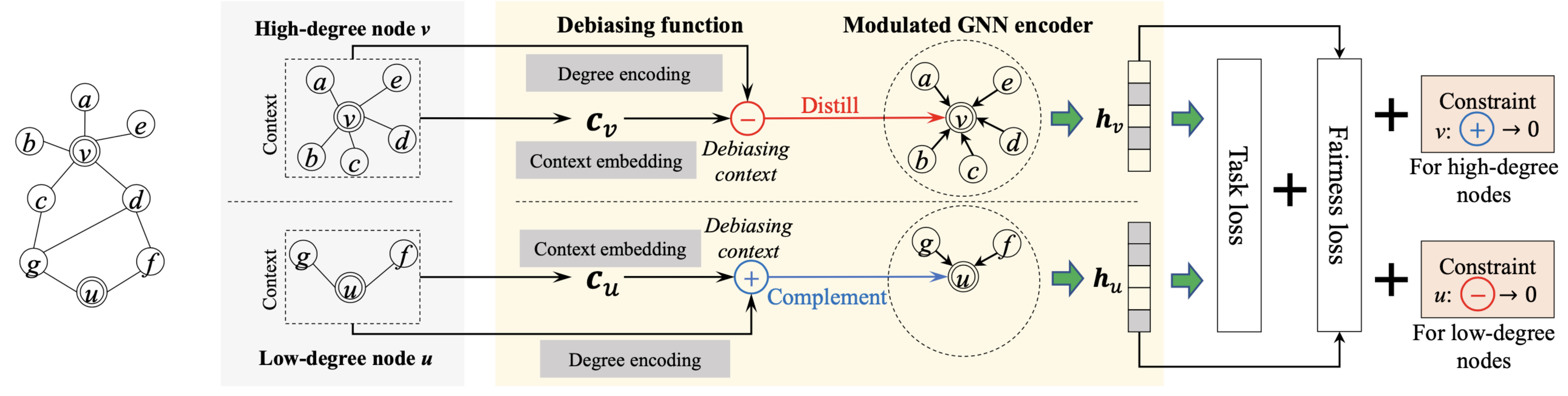
- Groups: $\mathcal{G}_i = \{v \in \mathcal{V} \mid d_i \le \deg_r(v) < d_{i+1}\}$
- Metrics
  - Degree Statistical Parity: $P(\hat{y}_v = y \mid v \in \mathcal{G}_i) = P(\hat{y}_v = y \mid v \in \mathcal{G}_j)$
  - Degree Equal Opportunity: $P(\hat{y}_v = y \mid y_v = y, v \in \mathcal{G}_i) = P(\hat{y}_v = y \mid y_v = y, v \in \mathcal{G}_j)$

Proposed model: DegFairGNN

Structural Contrast

Low-degree group: $\mathcal{S}_0 = \{v \in \mathcal{V} \mid \deg_1(v) \leqslant K\}$

High-degree group: $\mathcal{S}_1 = \mathcal{V} \setminus \mathcal{S}_0$



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

Debias Neighborhood Aggregation

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

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

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

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

Overall Loss

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

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

Experiments

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)

- **Base GNN models**
  - GCN [1]
  - GAT [2]
  - GraphSAGE [3]
- **Baselines**
  - Degree-specific models: DSGCN [4], Residual2Vec [5], Tail-GNN [6]
  - Fairness-aware models: FairWalk [7], CFC [8], FairGNN [9], FairAdj [10], FairVGNN [11]

Metrics

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

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

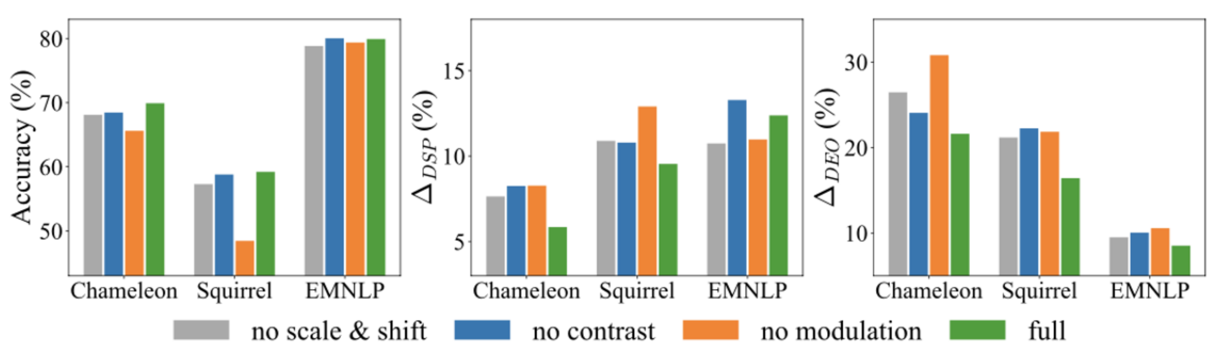
Main results

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. ↑	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
	Δ_{DSP} ↓	9.68 ± 1.37	8.81 ± 1.15	14.52 ± 0.69	8.51 ± 1.72	8.18 ± 0.93	10.12 ± 1.28	<u>7.33</u> ± 1.09	9.79 ± 1.91	8.86 ± 1.11	**5.85** ± 0.32
	Δ_{DEO} ↓	36.08 ± 2.65	25.14 ± 2.67	37.31 ± 1.99	26.09 ± 3.25	<u>22.89</u> ± 2.75	29.54 ± 1.95	26.83 ± 1.95	27.48 ± 2.06	26.02 ± 2.39	**21.60** ± 0.71
Squirrel	Acc. ↑	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
	Δ_{DSP} ↓	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	<u>9.54</u> ± 1.02
	Δ_{DEO} ↓	27.00 ± 3.79	32.61 ± 3.74	34.49 ± 0.72	33.60 ± 0.72	<u>17.12</u> ± 1.50	21.60 ± 2.69	17.62 ± 2.40	27.54 ± 1.73	35.80 ± 1.76	**16.42** ± 1.38
EMNLP	Acc. ↑	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
	Δ_{DSP} ↓	44.55 ± 1.90	50.00 ± 2.98	<u>12.90</u> ± 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
	Δ_{DEO} ↓	34.05 ± 3.56	46.92 ± 2.91	<u>11.26</u> ± 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

Ablation study



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

- (1) No scale and shift: worse in accuracy and fairness
- (2) No constrast: fairness generally becomes worse
- (3) No modulation: fairness becomes worse in most cases

Conclusion

- **Problem**
  - Degree fairness in graph neural networks
- **Proposed model: DegFairGNN**
  - Target the root of degree bias
  - Flexibly work with most modern GNNs
- **Experiments**
  - Extensive experiments on three benchmark datasets shows promising results on both accuracy and fairness metrics

References

[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 Convolutional 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.

Acknowledgments

This research is supported by the Agency for Science, Technology and Research (A*STAR) under its AME Programmatic Funds (Grant No. A20H6b0151).