Stage-Aware Dual Alignment for Covariate Shift in Graph Domain Adaptation

Hongwei Wen · Can Zhang · Haoyu He · Xintao Zhao · Hanyuan Hang · Minglong Lei

Problem & Motivation

Graph domain adaptation transfers from labeled source graphs to unlabeled target graphs. Covariate shift arises in the joint feature–structure distribution P(X, A).

Conventional final-layer alignment entangles distinct sources of discrepancy.

Stage-wise Bias: Key Insight

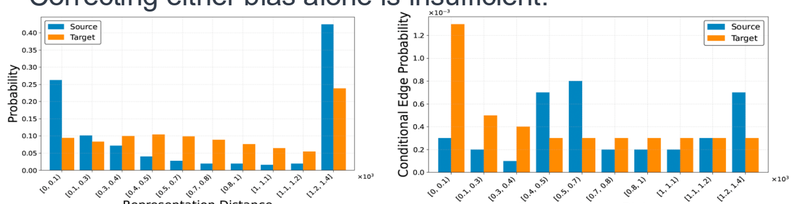
$$P(X, A) = P(X)\,P(A \mid X)$$

FS: $P_s(X) \neq P_t(X)$ **FCSS: $P_s(A \mid X) \neq P_t(A \mid X)$**

Feature Shift (FS) distorts initial representations before message passing.

Feature-Conditional Structure Shift (FCSS) biases structure-dependent propagation.

Correcting either bias alone is insufficient.



(a) FS (b) FCSS

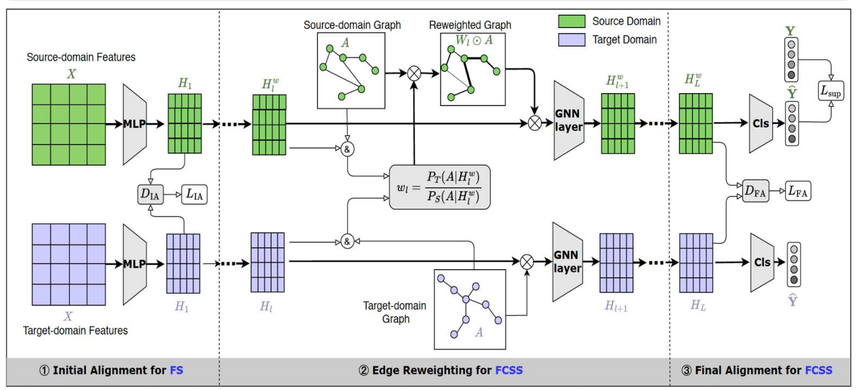
Our Contributions

① Identify two non-interchangeable bias sources induced by graph covariate shift.

② Propose DACS: initial adversarial alignment, layer-wise edge reweighting, and final adversarial alignment.

③ Establish a target-risk decomposition and validate on synthetic and real-world shifts.

DACS Framework

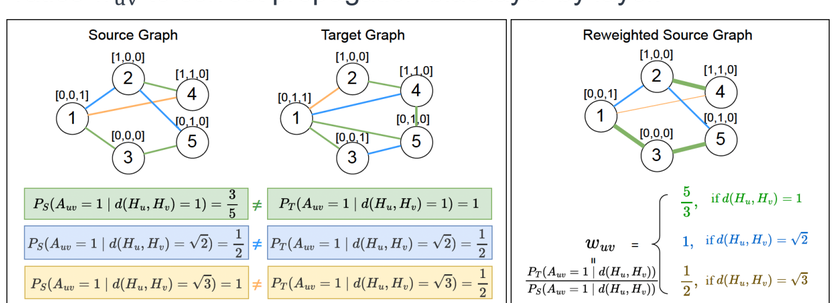


Three interventions (IA, RW, and FA) match the computational stages at which discrepancy is introduced.

Stage-Aware Dual Alignment

① Initial Alignment (IA): adversarially match $P_s(H^1)$ and $P_t(H^1)$ to mitigate representation bias.

② Reweighting (RW): use feature-conditioned edge probability ratios w_{uv} to correct propagation bias layer by layer.



③ Final Alignment (FA): adversarially align reweighted source and target final representations to reduce residual mismatch.

Theoretical Insight

Theorem 4.1 (Target Risk Decomposition). *Let $H_1 = f(X)$ denote the initial representation and $H_L = \phi(H_1, A)$ denote the final representation after message passing. Under CS, the target-domain classification risk satisfies*

$$\mathcal{R}_T(g) \le \mathcal{R}_S(g) + \sqrt{2}\cdot D_{\mathrm{KL}}^{1/2}\big(P_S(H_1), P_T(H_1)\big) + \sqrt{2}\cdot \mathbb{E}_{H_1\sim P_T} D_{\mathrm{KL}}^{1/2}\big(P_S(H_L \mid H_1), P_T(H_L \mid H_1)\big), \tag{12}$$

where $D_{\mathrm{KL}}(P_S, P_T)$ is the KL divergence between the distributions P_S and P_T.

Target risk is controlled by source risk together with two distinct discrepancy terms on initial representation and conditional propagation.

Neither term can generally be eliminated by controlling only the other. Dual alignment follows from the bias decomposition.

Real-world Results

Table 2: Performance on DBLP/ACM and Airport datasets.

Models	DBLP/ACM		Airport					
	A-D	D-A	U-E	E-U	B-E	E-B	B-U	U-B
UDAGCN	68.86±1.08	63.91±1.16	48.87±1.45	43.41±0.52	50.77±0.70	47.62±1.13	49.78±0.55	61.22±1.25
ASN	72.70±0.38	71.62±0.78	46.45±0.34	46.25±2.65	49.62±1.04	59.03±1.56	49.86±2.37	51.91±0.89
AdaGCN	68.59±1.12	66.75±1.59	43.64±0.82	45.57±0.41	48.52±0.63	56.17±1.22	50.30±1.45	47.24±1.20
CWGCN	76.74±0.11	72.75±0.45	47.03±0.85	45.74±0.95	48.77±0.52	63.96±0.89	49.82±1.46	62.29±0.78
JHGDA	75.58±1.65	73.22±0.68	50.75±3.01	52.27±6.25	56.64±1.14	**73.13±1.37**	50.20±0.54	69.27±3.32
GAA	78.37±7.43	74.76±1.57	55.96±3.69	53.33±1.24	56.64±0.75	68.30±2.08	53.22±1.25	69.81±4.92
StruRW	70.19±2.10	66.57±0.42	53.77±0.98	49.67±2.88	56.06±1.31	65.65±0.15	52.19±2.01	62.84±0.11
PairAlign	75.24±1.06	74.77±1.57	55.39±0.94	54.28±2.07	55.72±0.97	52.90±1.48	52.78±1.62	67.86±0.50
Adalign	78.49±7.70	73.47±1.08	48.62±2.17	49.38±0.33	53.08±0.89	60.15±0.57	50.81±0.91	70.99±5.29
DACS	**86.21±0.24**	**79.14±0.31**	**56.58±0.36**	**55.72±0.08**	**57.39±0.15**	70.76±0.19	**54.09±0.13**	**74.34±0.30**

Table 3: Performance on Arxiv dataset.

Models	1950-2007		1950-2009		1950-2011	
	2014-2016	2016-2018	2014-2016	2016-2018	2014-2016	2016-2018
UDAGCN	38.10 ± 1.62	OOM	42.85 ± 2.09	OOM	53.13 ± 0.31	OOM
AdaGCN	37.69 ± 1.19	43.01 ± 1.87	45.50 ± 0.68	47.53 ± 1.12	51.37 ± 1.12	53.34 ± 1.60
CWGCN	32.53 ± 0.52	35.10 ± 1.95	42.05 ± 0.37	36.97 ± 2.31	46.08 ± 0.99	41.56 ± 0.59
StruRW	38.56 ± 0.77	37.17 ± 2.75	43.55 ± 2.37	43.55 ± 2.37	53.19 ± 0.45	53.64 ± 0.65
PairAlign	39.98 ± 0.77	41.14 ± 2.07	44.60 ± 0.42	44.43 ± 0.34	53.75 ± 0.48	52.83 ± 0.98
Adalign	41.14 ± 0.22	45.07 ± 1.63	43.65 ± 0.48	48.15 ± 0.69	49.95 ± 2.62	52.91 ± 0.69
DACS	**42.94 ± 2.25**	**45.33 ± 4.26**	**45.87 ± 0.97**	**50.35 ± 2.03**	**54.80 ± 0.75**	**55.01 ± 1.68**

The best and second-best performances are marked as **bold** and underline, respectively.