NEURAL INFORMATION PROCESSING SYSTEMS

Revisiting, Benchmarking and Understanding Unsupervised Graph Domain Adaptation

Meihan Liu[1], Zhen Zhang[2], Jiachen Tang[1],
Jiajun Bu[1], Bingsheng He[2], Sheng Zhou[1]

[1]Zhejiang Provincial Key Laboratory of Service Robot, Zhejiang University.
[2]National University of Singapore

浙江大學
ZHEJIANG UNIVERSITY

Background

- GNNs rely heavily on high quality labelled data. However, it could be labor-intensive to annotate sufficient labels.
- Applying a well-trained classifier to another graph can yield inferior performance.

ACM Citation Network

Node Classification ACC: **92%**

DBLP Citation Network

Node Classification ACC: **63%**

Related Work

Unsupervised Graph Domain Adaptation (UGDA) has become an important solution for transferring knowledge from a labeled source graph to an unlabeled target graph.

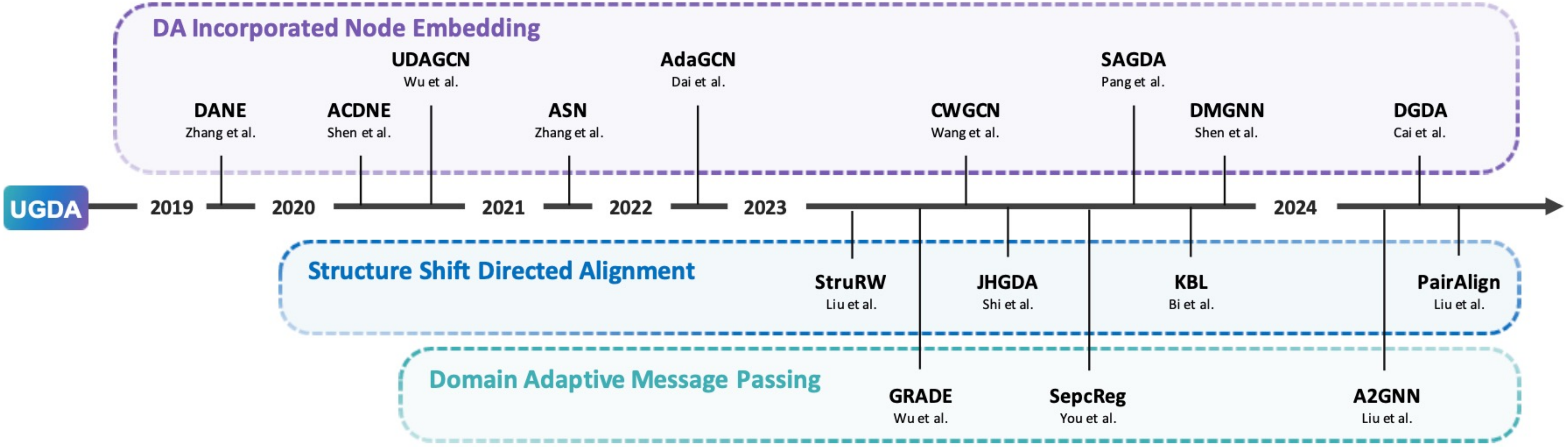

Figure 1: This timeline illustrates the diverse UGDA algorithms revisited in this paper. All of them are incorporated into our PyGDA library. More details are shown in Section 2 and Appendix C

Challenges

- **Inadequate Evaluation of Domain Distribution Discrepancies.**

  The distribution shifts in node attributes, graph structures, and label proportions between graphs will significantly influence the adaptation performance and result in various adaptation scenarios. However, the types and magnitudes of distribution discrepancies among different domains have not been thoroughly evaluated and discussed.

- **Lack of Standard, Fair, and Comprehensive Comparisons.**

  The utilization of distinct datasets, varying data processing methodologies, and divergent data partitioning strategies results in incomparability across different findings. Further more, they are mainly evaluated against limited baselines with constrained scenarios, which lack validation of the model's capability in more diverse or complex applications.

- **Limited Investigation on GNN Inherent Transferability.**

  It is still unclear how data shift impose challenges on GNN and how to unleash the transferabililty power for GNN. Understanding the key components that affect adaptation in GNN will be crucial for enhancing GNN's transferabililty, which is still an open problem.

Datasets

For node classification tasks, we have carefully selected **five** widely used public datasets and compiled a comprehensive collection comprising **74** distinct source-target adaptation pairs.

(1) Wide range of distribution shift. (2) Different scales with variant spans. (3) Various downstream applications.

Table 1: Datasets used in GDABench reflecting a wide range of distribution shifts. '-' indicates no data shift exists. Circles (○, ◑ and ●) represent the degree of the corresponding shift between domains and Airport does not contain node features. The magnitude of shift is directly proportional to the filling area of the circle. The statistic manners and more details are provided in Appendix B.

Dataset	Size	Feature Shift	Structure Shift	Label Shift	# Domains	# Labels	# Homo
Airport	S	-	◑	○	3	4	0.52
Blog	S	○	○	○	2	6	0.40
ArnetMiner	M	◑	◑	○	3	5	0.83
Twitch	M	●	●	◑	6	2	0.59
MAG	L	●	●	●	6	20	0.58

Compared Models

- **Specialized UGDA Methods.**

  This group includes specifically designed algorithms for graph domain adaptation task.

- **SimGDA: Vanilla DA with GNN Variants.**

  To understand the inherent transferability of GNN, we delve into its aggregation process and decouple it from two perspectives: *how to aggregate* and *what to aggregate*. Then, we get 14 models by combining variants with two vanilla DA methods.

- **SimGDA+: SimGDA with Unsupervised Techniques.**

  To further unlock the power of GNN for graph domain adaptation, we enhance SimGDA with unsupervised graph learning techniques on unlabeled target graph, which allows the model to learn meaningful representations without relying on domain-specific labels. We implement three unsupervised techniques in an end-to-end manner.

Compared Models

- **Specialized UGDA Methods.**

  Specifically designed algorithms for graph domain adaptation task.

- **SimGDA: Vanilla DA with GNN Variants.**

  14 models by decoupling aggregation process from two perspectives.

- **SimGDA+: SimGDA with Unsupervised Techniques.**

  Enhance SimGDA with three unsupervised graph learning techniques on unlabeled target graph.

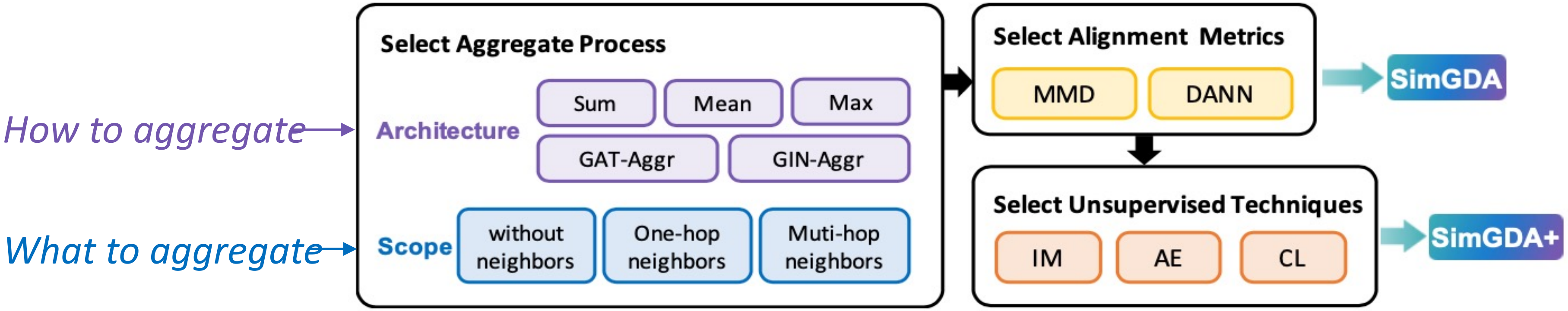

Figure 2: The combination process of SimGDA / SimGDA+.

Overall Comparisons

- **Observation 1:** When facing significant shifts, it is important to design solutions tailored to mitigate structural discrepancies.

Table 2&3: We compared MicroF1 of each model on Airport, Blog, and ArnetMiner. We evaluated Macro-F1 on MAG and AUROC scores on Twitch. Highlighted are the top first, second, and third results. Results for other tasks can be found in Appendix D.

	Models	Airport		Blog		ArnetMiner		Twitch			MAG		
		E→ U	U → E	B1 → B2	B2 → B1	C → A	D → A	DE→ ES	EN → RU	DE → PT	FR → JP	JP → FR	JP → RU
DA Incorporated Node Embedding	DANE	31.18 ±3.26	33.75 ±0.31	32.17 ±3.20	32.77 ±0.66	62.87 ±1.98	59.19 ±1.66	56.71 ±0.57	53.47 ±0.84	55.71 ±1.70	16.16 ±0.24	16.71 ±1.45	12.61 ±0.19
	ACDNE	48.52 ±1.17	45.03 ±0.43	54.30 ±1.42	56.93 ±0.56	72.34 ±0.39	65.37 ±3.50	51.13 ±0.34	50.79 ±0.12	52.47 ±1.83	20.12 ±0.37	18.92 ±1.08	13.91 ±0.35
	UDAGCN	41.62 ±0.58	33.17 ±0.12	27.58 ±0.63	25.46 ±5.96	70.24 ±1.01	62.49 ±1.16	56.48 ±0.18	53.72 ±0.42	54.22 ±2.10	12.22 ±0.31	11.62 ±0.35	11.17 ±0.29
	ASN	46.58 ±0.21	40.85 ±0.54	53.91 ±0.52	56.25 ±0.52	71.70 ±0.38	66.15 ±1.02	53.57 ±0.72	50.29 ±0.15	55.03 ±0.75	11.91 ±1.45	12.04 ±0.70	10.79 ±0.24
	AdaGCN	46.55 ±0.42	49.62 ±0.01	43.06 ±3.03	36.58 ±8.94	67.66 ±0.36	60.47 ±0.99	52.32 ±0.76	51.99 ±1.31	51.09 ±0.61	16.21 ±0.47	14.12 ±0.46	13.05 ±0.08
	DMGNN	45.85 ±1.03	27.82 ±2.95	46.27 ±2.40	44.96 ±1.57	72.91 ±0.44	70.68 ±0.27	54.11 ±0.09	50.42 ±0.03	53.44 ±0.04	12.01 ±0.78	9.93 ±0.40	10.28 ±1.04
	CWGCN	44.68 ±0.42	40.69 ±0.47	31.96 ±3.47	33.46 ±4.96	71.65 ±0.21	68.21 ±0.09	57.62 ±0.64	52.90 ±0.37	58.21 ±0.57	11.01 ±0.48	12.37 ±0.54	12.38 ±0.25
	SAGDA	30.62 ±5.57	35.92 ±1.01	26.51 ±11.12	26.91 ±7.26	65.40 ±4.38	64.60 ±0.89	51.58 ±0.09	51.03 ±0.23	51.96 ±0.70	16.28 ±0.51	3.64 ±5.07	11.40 ±0.59
	DGDA	43.45 ±2.16	43.78 ±2.90	22.10 ±1.45	21.06 ±2.07	52.20 ±4.62	56.31 ±2.01	54.43 ±3.60	51.68 ±1.08	54.29 ±4.28	OOM	OOM	OOM
Structure Shift Directed Alignment	StruRW	45.94 ±0.69	36.09 ±0.01	40.02 ±0.37	42.10 ±1.18	70.59 ±0.15	64.15 ±0.31	59.60 ±0.19	52.04 ±0.36	58.74 ±2.09	22.10 ±0.40	12.89 ±0.85	12.96 ±0.43
	KBL	44.54 ±0.73	32.08 ±0.20	35.14 ±3.97	34.90 ±2.49	70.49 ±0.26	63.34 ±0.53	58.33 ±0.44	55.91 ±0.16	51.66 ±0.08	17.60 ±0.39	6.12 ±0.14	14.49 ±0.30
	JHGDA	36.89 ±0.25	40.85 ±1.68	17.79 ±2.12	23.16 ±6.59	65.53 ±0.94	60.80 ±0.35	62.25 ±0.49	53.75 ±0.15	61.88 ±0.48	20.51 ±0.20	20.46 ±0.57	11.85 ±0.37
	PairAlign	42.38 ±0.77	36.84 ±1.48	32.17 ±10.88	41.16 ±3.02	58.06 ±2.62	56.68 ±0.89	50.78 ±0.22	51.19 ±0.20	52.03 ±0.97	23.29 ±0.49	23.72 ±0.30	12.34 ±0.31
Domain Adaptive Message Passing	GRADE	49.36 ±0.35	48.45 ±1.56	38.64 ±3.73	44.01 ±4.51	69.16 ±0.39	63.47 ±1.10	58.57 ±0.42	53.55 ±0.28	62.12 ±0.17	11.93 ±0.48	10.95 ±0.55	9.35 ±0.25
	SpecReg	37.59 ±2.55	28.91 ±8.77	28.27 ±4.22	30.30 ±1.35	68.90 ±4.78	66.30 ±4.28	51.04 ±0.33	50.17 ±0.06	55.91 ±0.59	19.45 ±0.54	20.17 ±1.35	15.82 ±0.50
	A2GNN	50.64 ±1.47	53.47 ±0.24	22.58 ±0.01	33.04 ±4.12	76.15 ±0.06	74.12 ±0.18	59.41 ±0.34	52.01 ±0.32	61.82 ±0.77	26.20 ±0.74	25.78 ±0.25	16.94 ±0.13
SimGDA+	SimGDA	55.29 ±0.39	54.39 ±0.90	53.35 ±0.69	43.04 ±0.86	70.80 ±0.06	67.04 ±0.15	61.30 ±0.32	53.11 ±0.19	58.27 ±0.16	18.59 ±0.07	15.16 ±0.11	13.27 ±0.04
	SimGDA+	58.11 ±0.40	57.52 ±0.38	57.04 ±0.45	44.17 ±0.02	73.18 ±0.38	71.81 ±2.44	61.53 ±0.08	53.82 ±0.08	61.60 ±0.11	21.94 ±0.18	21.36 ±0.09	15.64 ±0.46

Overall Comparisons

- **Observation 2:** Domain-adaptive message passing methods demonstrate superior and robust performance across a wide range of datasets and tasks.

Table 2&3: We compared MicroF1 of each model on Airport, Blog, and ArnetMiner. We evaluated Macro-F1 on MAG and AUROC scores on Twitch. Highlighted are the top first, second, and third results. Results for other tasks can be found in Appendix D.

	Models	Airport		Blog		ArnetMiner		Twitch			MAG		
		E→ U	U → E	B1 → B2	B2 → B1	C → A	D → A	DE→ ES	EN → RU	DE → PT	FR → JP	JP → FR	JP → RU
DA Incorporated Node Embedding	DANE	31.18 $\pm$3.26	33.75 $\pm$0.31	32.17 $\pm$3.20	32.77 $\pm$0.66	62.87 $\pm$1.98	59.19 $\pm$1.66	56.71 $\pm$0.57	53.47 $\pm$0.84	55.71 $\pm$1.70	16.16 $\pm$0.24	16.71 $\pm$1.45	12.61 $\pm$0.19
	ACDNE	48.52 $\pm$1.17	45.03 $\pm$0.43	**54.30** $\pm$1.42	**56.93** $\pm$0.56	72.34 $\pm$0.39	65.37 $\pm$3.50	51.13 $\pm$0.34	50.79 $\pm$0.12	52.47 $\pm$1.83	20.12 $\pm$0.37	18.92 $\pm$1.08	13.91 $\pm$0.35
	UDAGCN	41.62 $\pm$0.58	33.17 $\pm$0.12	27.58 $\pm$0.63	25.46 $\pm$5.96	70.24 $\pm$1.01	62.49 $\pm$1.16	56.48 $\pm$0.18	53.72 $\pm$0.42	54.22 $\pm$2.10	12.22 $\pm$0.31	11.62 $\pm$0.35	11.17 $\pm$0.29
	ASN	46.58 $\pm$0.21	40.85 $\pm$0.54	**53.91** $\pm$0.52	**56.25** $\pm$0.52	71.70 $\pm$0.38	66.15 $\pm$1.02	53.57 $\pm$0.72	50.29 $\pm$0.15	55.03 $\pm$0.75	11.91 $\pm$1.45	12.04 $\pm$0.70	10.79 $\pm$0.24
	AdaGCN	46.55 $\pm$0.42	49.62 $\pm$0.01	43.06 $\pm$3.03	36.58 $\pm$8.94	67.66 $\pm$0.36	60.47 $\pm$0.99	52.32 $\pm$0.76	51.99 $\pm$1.31	51.09 $\pm$0.61	16.21 $\pm$0.47	14.12 $\pm$0.46	13.05 $\pm$0.08
	DMGNN	45.85 $\pm$1.03	27.82 $\pm$2.95	46.27 $\pm$2.40	**44.96** $\pm$1.57	**72.91** $\pm$0.44	**70.68** $\pm$0.27	54.11 $\pm$0.09	50.42 $\pm$0.03	53.44 $\pm$0.04	12.01 $\pm$0.78	9.93 $\pm$0.40	10.28 $\pm$1.04
	CWGCN	44.68 $\pm$0.42	40.69 $\pm$0.47	31.96 $\pm$3.47	33.46 $\pm$4.96	71.65 $\pm$0.21	68.21 $\pm$0.09	57.62 $\pm$0.64	52.90 $\pm$0.37	58.21 $\pm$0.57	11.01 $\pm$0.48	12.37 $\pm$0.54	12.38 $\pm$0.25
	SAGDA	30.62 $\pm$5.57	35.92 $\pm$1.01	26.51 $\pm$11.12	26.91 $\pm$7.26	65.40 $\pm$4.38	64.60 $\pm$0.89	51.58 $\pm$0.09	51.03 $\pm$0.23	51.96 $\pm$0.70	16.28 $\pm$0.51	3.64 $\pm$5.07	11.40 $\pm$0.59
	DGDA	43.45 $\pm$2.16	43.78 $\pm$2.90	22.10 $\pm$1.45	21.06 $\pm$2.07	52.20 $\pm$4.62	56.31 $\pm$2.01	54.43 $\pm$3.60	51.68 $\pm$1.08	54.29 $\pm$4.28	OOM	OOM	OOM
Structure Shift Directed Alignment	StruRW	45.94 $\pm$0.69	36.09 $\pm$0.01	40.02 $\pm$0.37	42.10 $\pm$1.18	70.59 $\pm$0.15	64.15 $\pm$0.31	59.60 $\pm$0.19	52.04 $\pm$0.36	58.74 $\pm$2.09	**22.10** $\pm$0.40	12.89 $\pm$0.85	12.96 $\pm$0.43
	KBL	44.54 $\pm$0.73	32.08 $\pm$0.20	35.14 $\pm$3.97	34.90 $\pm$2.49	70.49 $\pm$0.26	63.34 $\pm$0.53	58.33 $\pm$0.44	**55.91** $\pm$0.16	51.66 $\pm$0.08	17.60 $\pm$0.39	6.12 $\pm$0.14	14.49 $\pm$0.30
	JHGDA	36.89 $\pm$0.25	40.85 $\pm$1.68	17.79 $\pm$2.12	23.16 $\pm$6.59	65.53 $\pm$0.94	60.80 $\pm$0.35	**62.25** $\pm$0.49	**53.75** $\pm$0.15	**61.88** $\pm$0.48	20.51 $\pm$0.20	20.46 $\pm$0.57	11.85 $\pm$0.37
	PairAlign	42.38 $\pm$0.77	36.84 $\pm$1.48	32.17 $\pm$10.88	41.16 $\pm$3.02	58.06 $\pm$2.62	56.68 $\pm$0.89	50.78 $\pm$0.22	51.19 $\pm$0.20	52.03 $\pm$0.97	**23.29** $\pm$0.49	**23.72** $\pm$0.30	12.34 $\pm$0.31
Domain Adaptive Message Passing	GRADE	49.36 $\pm$0.35	48.45 $\pm$1.56	38.64 $\pm$3.73	44.01 $\pm$4.51	69.16 $\pm$0.39	63.47 $\pm$1.10	58.57 $\pm$0.42	53.55 $\pm$0.28	**62.12** $\pm$0.17	11.93 $\pm$0.48	10.95 $\pm$0.55	9.35 $\pm$0.25
	SpecReg	37.59 $\pm$2.55	28.91 $\pm$8.77	28.27 $\pm$4.22	30.30 $\pm$1.35	68.90 $\pm$4.78	66.30 $\pm$4.28	51.04 $\pm$0.33	50.17 $\pm$0.06	55.91 $\pm$0.59	19.45 $\pm$0.54	20.17 $\pm$1.35	**15.82** $\pm$0.50
	A2GNN	**50.64** $\pm$1.47	**53.47** $\pm$0.24	22.58 $\pm$0.01	33.04 $\pm$4.12	**76.15** $\pm$0.06	**74.12** $\pm$0.18	59.41 $\pm$0.34	52.01 $\pm$0.32	**61.82** $\pm$0.77	**26.20** $\pm$0.74	**25.78** $\pm$0.25	**16.94** $\pm$0.13
SimGDA+	SimGDA	**55.29** $\pm$0.39	**54.39** $\pm$0.90	53.35 $\pm$0.69	43.04 $\pm$0.86	70.80 $\pm$0.06	67.04 $\pm$0.15	**61.30** $\pm$0.32	53.11 $\pm$0.19	58.27 $\pm$0.16	18.59 $\pm$0.07	15.16 $\pm$0.11	13.27 $\pm$0.04
	SimGDA+	**58.11** $\pm$0.40	**57.52** $\pm$0.38	**57.04** $\pm$0.45	44.17 $\pm$0.02	**73.18** $\pm$0.38	**71.81** $\pm$2.44	**61.53** $\pm$0.08	**53.82** $\pm$0.08	61.60 $\pm$0.11	21.94 $\pm$0.18	**21.36** $\pm$0.09	**15.64** $\pm$0.46

Understanding the inherent power of GNN

- **Observation 3:** SimGDA achieves competitive performance compared with UGDA methods.

Table 2&3: We compared MicroF1 of each model on Airport, Blog, and ArnetMiner. We evaluated Macro-F1 on MAG and AUROC scores on Twitch. Highlighted are the top first, second, and third results. Results for other tasks can be found in Appendix D.

	Models	Airport		Blog		ArnetMiner		Twitch			MAG		
		E→ U	U → E	B1 → B2	B2 → B1	C → A	D → A	DE→ ES	EN → RU	DE → PT	FR → JP	JP → FR	JP → RU
DA Incorporated Node Embedding	DANE	31.18 ±3.26	33.75 ±0.31	32.17 ±3.20	32.77 ±0.66	62.87 ±1.98	59.19 ±1.66	56.71 ±0.57	53.47 ±0.84	55.71 ±1.70	16.16 ±0.24	16.71 ±1.45	12.61 ±0.19
	ACDNE	48.52 ±1.17	45.03 ±0.43	54.30 ±1.42	56.93 ±0.56	72.34 ±0.39	65.37 ±3.50	51.13 ±0.34	50.79 ±0.12	52.47 ±1.83	20.12 ±0.37	18.92 ±1.08	13.91 ±0.35
	UDAGCN	41.62 ±0.58	33.17 ±0.12	27.58 ±0.63	25.46 ±5.96	70.24 ±1.01	62.49 ±1.16	56.48 ±0.18	53.72 ±0.42	54.22 ±2.10	12.22 ±0.31	11.62 ±0.35	11.17 ±0.29
	ASN	46.58 ±0.21	40.85 ±0.54	53.91 ±0.52	56.25 ±0.52	71.70 ±0.38	66.15 ±1.02	53.57 ±0.72	50.29 ±0.15	55.03 ±0.75	11.91 ±1.45	12.04 ±0.70	10.79 ±0.24
	AdaGCN	46.55 ±0.42	49.62 ±0.01	43.06 ±3.03	36.58 ±8.94	67.66 ±0.36	60.47 ±0.99	52.32 ±0.76	51.99 ±1.31	51.09 ±0.61	16.21 ±0.47	14.12 ±0.46	13.05 ±0.08
	DMGNN	45.85 ±1.03	27.82 ±2.95	46.27 ±2.40	44.96 ±1.57	72.91 ±0.44	70.68 ±0.27	54.11 ±0.09	50.42 ±0.03	53.44 ±0.04	12.01 ±0.78	9.93 ±0.40	10.28 ±1.04
	CWGCN	44.68 ±0.42	40.69 ±0.47	31.96 ±3.47	33.46 ±4.96	71.65 ±0.21	68.21 ±0.09	57.62 ±0.64	52.90 ±0.37	58.21 ±0.57	11.01 ±0.48	12.37 ±0.54	12.38 ±0.25
	SAGDA	30.62 ±5.57	35.92 ±1.01	26.51 ±11.12	26.91 ±7.26	65.40 ±4.38	64.60 ±0.89	51.58 ±0.09	51.03 ±0.23	51.96 ±0.70	16.28 ±0.51	3.64 ±5.07	11.40 ±0.59
	DGDA	43.45 ±2.16	43.78 ±2.90	22.10 ±1.45	21.06 ±2.07	52.20 ±4.62	56.31 ±2.01	54.43 ±3.60	51.68 ±1.08	54.29 ±4.28	OOM	OOM	OOM
Structure Shift Directed Alignment	StruRW	45.94 ±0.69	36.09 ±0.01	40.02 ±0.37	42.10 ±1.18	70.59 ±0.15	64.15 ±0.31	59.60 ±0.19	52.04 ±0.36	58.74 ±2.09	22.10 ±0.40	12.89 ±0.85	12.96 ±0.43
	KBL	44.54 ±0.73	32.08 ±0.20	35.14 ±3.97	34.90 ±2.49	70.49 ±0.26	63.34 ±0.53	58.33 ±0.44	55.91 ±0.16	51.66 ±0.08	17.60 ±0.39	6.12 ±0.14	14.49 ±0.30
	JHGDA	36.89 ±0.25	40.85 ±1.68	17.79 ±2.12	23.16 ±6.59	65.53 ±0.94	60.80 ±0.35	62.25 ±0.49	53.75 ±0.15	61.88 ±0.48	20.51 ±0.20	20.46 ±0.57	11.85 ±0.37
	PairAlign	42.38 ±0.77	36.84 ±1.48	32.17 ±10.88	41.16 ±3.02	58.06 ±2.62	56.68 ±0.89	50.78 ±0.22	51.19 ±0.20	52.03 ±0.97	23.29 ±0.49	23.72 ±0.30	12.34 ±0.31
Domain Adaptive Message Passing	GRADE	49.36 ±0.35	48.45 ±1.56	38.64 ±3.73	44.01 ±4.51	69.16 ±0.39	63.47 ±1.10	58.57 ±0.42	53.55 ±0.28	62.12 ±0.17	11.93 ±0.48	10.95 ±0.55	9.35 ±0.25
	SpecReg	37.59 ±2.55	28.91 ±8.77	28.27 ±4.22	30.30 ±1.35	68.90 ±4.78	66.30 ±4.28	51.04 ±0.33	50.17 ±0.06	55.91 ±0.59	19.45 ±0.54	20.17 ±1.35	15.82 ±0.50
	A2GNN	50.64 ±1.47	53.47 ±0.24	22.58 ±0.01	33.04 ±4.12	76.15 ±0.06	74.12 ±0.18	59.41 ±0.34	52.01 ±0.32	61.82 ±0.77	26.20 ±0.74	25.78 ±0.25	16.94 ±0.13
	SimGDA	55.29 ±0.39	54.39 ±0.90	53.35 ±0.69	43.04 ±0.86	70.80 ±0.06	67.04 ±0.15	61.30 ±0.32	53.11 ±0.19	58.27 ±0.16	18.59 ±0.07	15.16 ±0.11	13.27 ±0.04
	SimGDA+	58.11 ±0.40	57.52 ±0.38	57.04 ±0.45	44.17 ±0.02	73.18 ±0.38	71.81 ±2.44	61.53 ±0.08	53.82 ±0.08	61.60 ±0.11	21.94 ±0.18	21.36 ±0.09	15.64 ±0.46

Understanding the inherent power of GNN

- **Observation 4:** The benefit of multi-hop neighbors depends on the degree of label shift and graph heterophily.

The impact of structural information varies depending on the dataset with diverse degree of label shift.

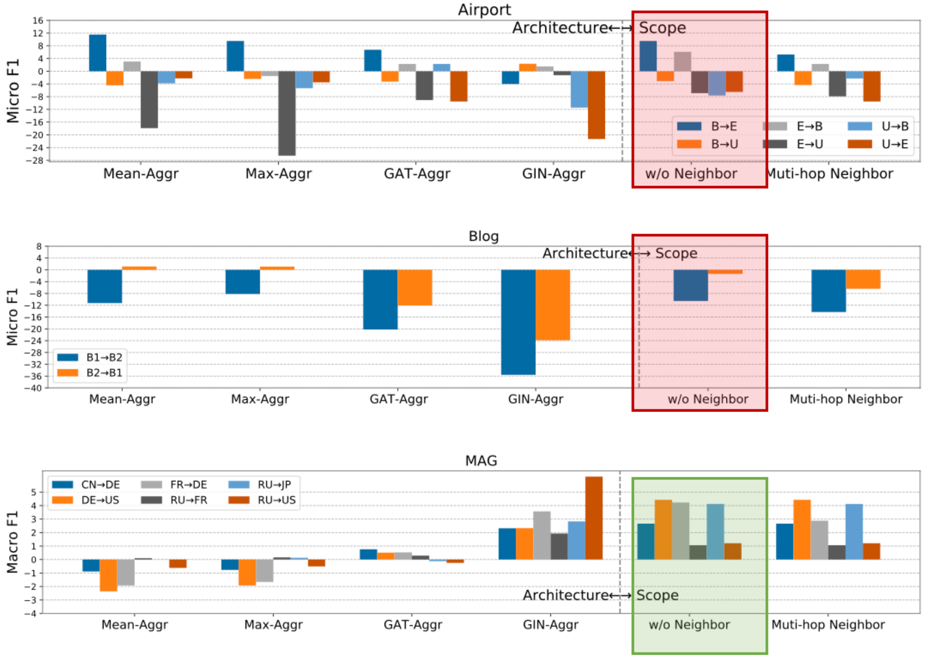

Figure 5: The compared performance of vanilla DA with 6 GNN variants.

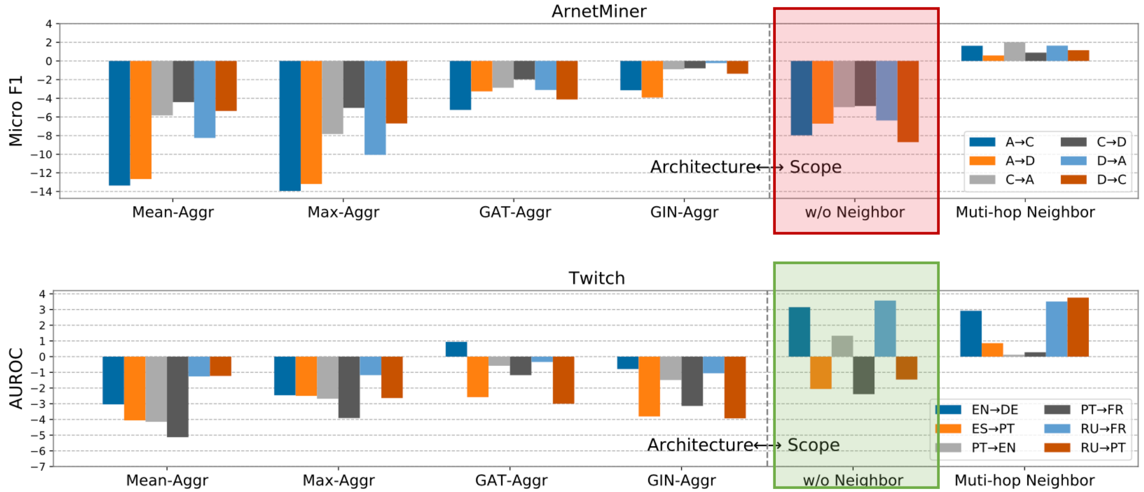

Figure 3: SimGDA: the compared performance of vanilla DA with 6 GNN variants.

Understanding the inherent power of GNN

- **Observation 4:** The benefit of multi-hop neighbors depends on the degree of label shift and graph heterophily.

Heterophilous graphs exhibit a larger degree of conditional shift, and aggregation process may help to mitigate this situation by providing a more comprehensive view of the node's context.

Figure 5: The compared performance of vanilla DA with 6 GNN variants.

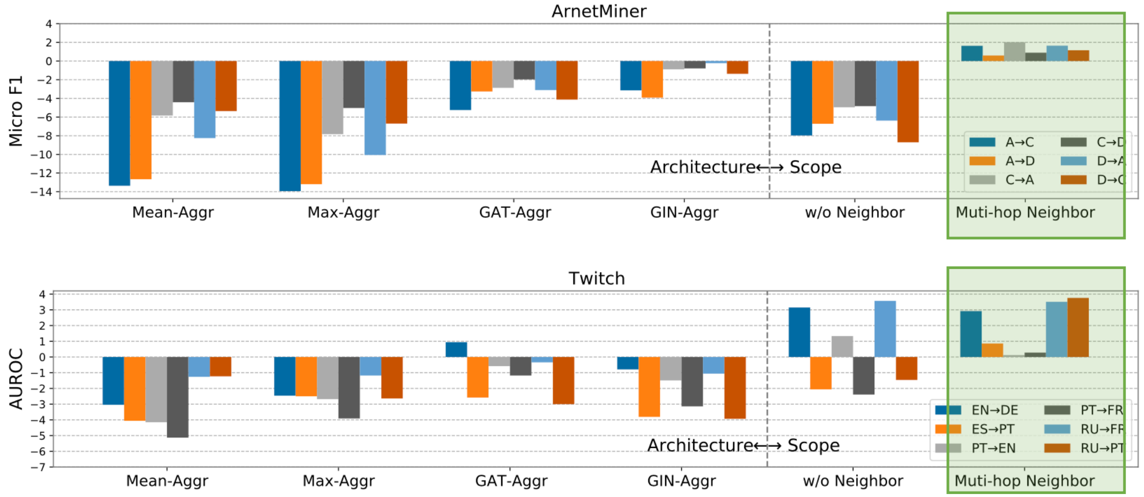

Figure 3: SimGDA: the compared performance of vanilla DA with 6 GNN variants.

Understanding the inherent power of GNN

- **Observation 5:** A source-unbiased discriminative aggregation mechanism is needed.

Superiority of the GCN aggregator over mean and max emphasizes *the necessity of a discriminative aggregation operator with highly expressive power*.

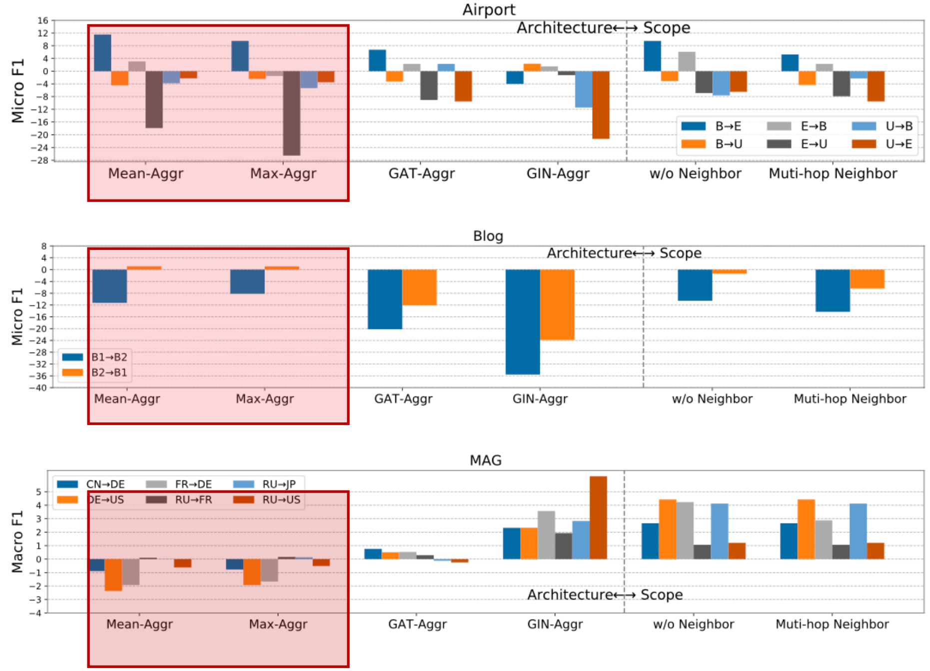

Figure 5: The compared performance of vanilla DA with 6 GNN variants.

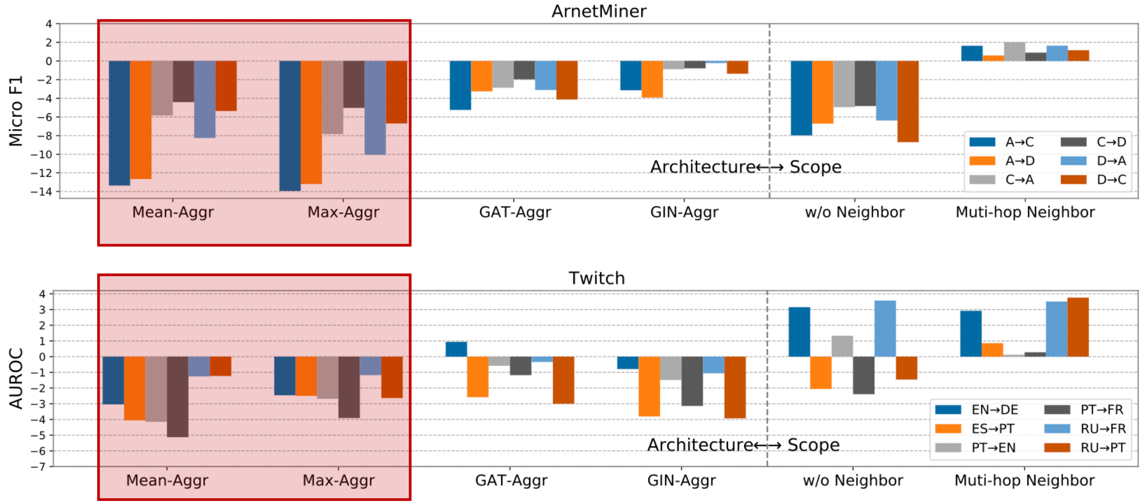

Figure 3: SimGDA: the compared performance of vanilla DA with 6 GNN variants.

Understanding the inherent power of GNN

- **Observation 5:** A source-unbiased discriminative aggregation mechanism is needed.

Source-bised discriminative aggregation mechanisms deteriorate the model’s transfer capability.

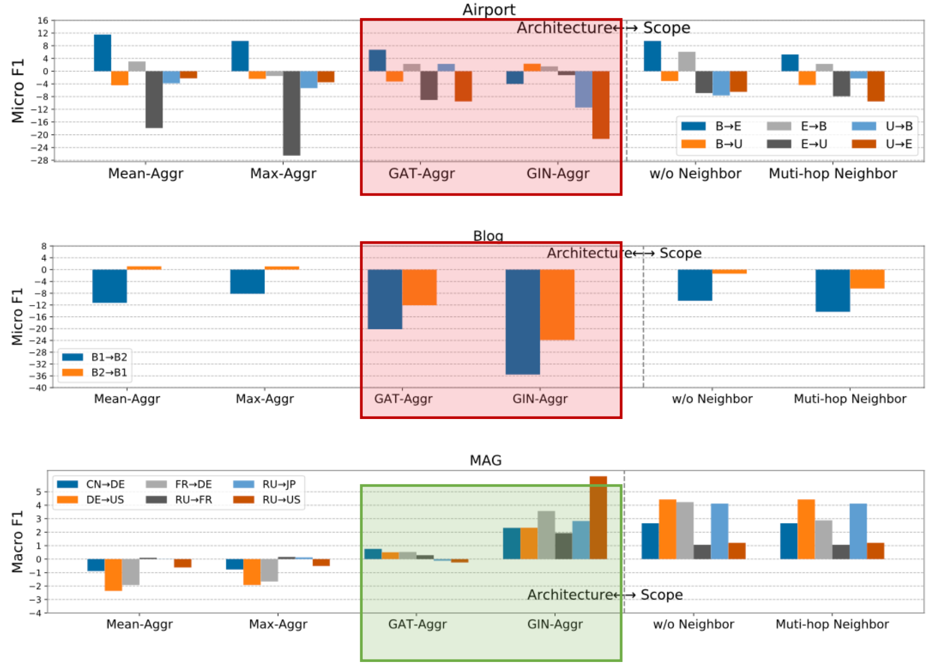

Figure 5: The compared performance of vanilla DA with 6 GNN variants.

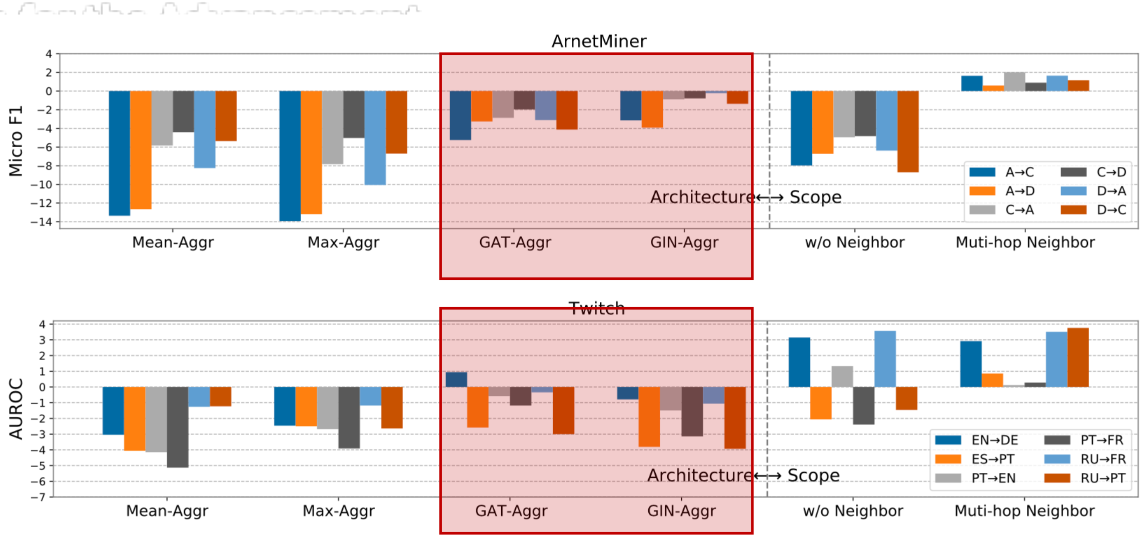

Figure 3: SimGDA: the compared performance of vanilla DA with 6 GNN variants.

Unlocking the inherent power of GNN

- **Observation 6:** GNNs can serve as a powerful graph domain adaptor with appropriate aggregators, careful selection of neighbor hops, and unsupervised learning techniques.

Table 2&3: We compared MicroF1 of each model on Airport, Blog, and ArnetMiner. We evaluated Macro-F1 on MAG and AUROC scores on Twitch. Highlighted are the top first, second, and third results. Results for other tasks can be found in Appendix D.

	Models	Airport		Blog		ArnetMiner		Twitch			MAG		
		E → U	U → E	B1 → B2	B2 → B1	C → A	D → A	DE → ES	EN → RU	DE → PT	FR → JP	JP → FR	JP → RU
DA Incorporated Node Embedding	DANE	31.18 ±3.26	33.75 ±0.31	32.17 ±3.20	32.77 ±0.66	62.87 ±1.98	59.19 ±1.66	56.71 ±0.57	53.47 ±0.84	55.71 ±1.70	16.16 ±0.24	16.71 ±1.45	12.61 ±0.19
	ACDNE	48.52 ±1.17	45.03 ±0.43	**54.30** ±1.42	**56.93** ±0.56	72.34 ±0.39	65.37 ±3.50	51.13 ±0.34	50.79 ±0.12	52.47 ±1.83	20.12 ±0.37	18.92 ±1.08	13.91 ±0.35
	UDAGCN	41.62 ±0.58	33.17 ±0.12	27.58 ±0.63	25.46 ±5.96	70.24 ±1.01	62.49 ±1.16	56.48 ±0.18	53.72 ±0.42	54.22 ±2.10	12.22 ±0.31	11.62 ±0.35	11.17 ±0.29
	ASN	46.58 ±0.21	40.85 ±0.54	**53.91** ±0.52	**56.25** ±0.52	71.70 ±0.38	66.15 ±1.02	53.57 ±0.72	50.29 ±0.15	55.03 ±0.75	11.91 ±1.45	12.04 ±0.70	10.79 ±0.24
	AdaGCN	46.55 ±0.42	49.62 ±0.01	43.06 ±3.03	36.58 ±8.94	67.66 ±0.36	60.47 ±0.99	52.32 ±0.76	51.99 ±1.31	51.09 ±0.61	16.21 ±0.47	14.12 ±0.46	13.05 ±0.08
	DMGNN	45.85 ±1.03	27.82 ±2.95	46.27 ±2.40	**44.96** ±1.57	**72.91** ±0.44	**70.68** ±0.27	54.11 ±0.09	50.42 ±0.03	53.44 ±0.04	12.01 ±0.78	9.93 ±0.40	10.28 ±1.04
	CWGCN	44.68 ±0.42	40.69 ±0.47	31.96 ±3.47	33.46 ±4.96	71.65 ±0.21	68.21 ±0.09	57.62 ±0.64	52.90 ±0.37	58.21 ±0.57	11.01 ±0.48	12.37 ±0.54	12.38 ±0.25
	SAGDA	30.62 ±5.57	35.92 ±1.01	26.51 ±11.12	26.91 ±7.26	65.40 ±4.38	64.60 ±0.89	51.58 ±0.09	51.03 ±0.23	51.96 ±0.70	16.28 ±0.51	3.64 ±5.07	11.40 ±0.59
	DGDA	43.45 ±2.16	43.78 ±2.90	22.10 ±1.45	21.06 ±2.07	52.20 ±4.62	56.31 ±2.01	54.43 ±3.60	51.68 ±1.08	54.29 ±4.28	OOM	OOM	OOM
Structure Shift Directed Alignment	StruRW	45.94 ±0.69	36.09 ±0.01	40.02 ±0.37	42.10 ±1.18	70.59 ±0.15	64.15 ±0.31	59.60 ±0.19	52.04 ±0.36	58.74 ±2.09	**22.10** ±0.40	12.89 ±0.85	12.96 ±0.43
	KBL	44.54 ±0.73	32.08 ±0.20	35.14 ±3.97	34.90 ±2.49	70.49 ±0.26	63.34 ±0.53	58.33 ±0.44	**55.91** ±0.16	51.66 ±0.08	17.60 ±0.39	6.12 ±0.14	14.49 ±0.30
	JHGDA	36.89 ±0.25	40.85 ±1.68	17.79 ±2.12	23.16 ±6.59	65.53 ±0.94	60.80 ±0.35	**62.25** ±0.49	**53.75** ±0.15	**61.88** ±0.48	20.51 ±0.20	20.46 ±0.57	11.85 ±0.37
	PairAlign	42.38 ±0.77	36.84 ±1.48	32.17 ±10.88	41.16 ±3.02	58.06 ±2.62	56.68 ±0.89	50.78 ±0.22	51.19 ±0.20	52.03 ±0.97	**23.29** ±0.49	**23.72** ±0.30	12.34 ±0.31
Domain Adaptive Message Passing	GRADE	49.36 ±0.35	48.45 ±1.56	38.64 ±3.73	44.01 ±4.51	69.16 ±0.39	63.47 ±1.10	58.57 ±0.42	53.55 ±0.28	**62.12** ±0.17	11.93 ±0.48	10.95 ±0.55	9.35 ±0.25
	SpecReg	37.59 ±2.55	28.91 ±8.77	28.27 ±4.22	30.30 ±1.35	68.90 ±4.78	66.30 ±4.28	51.04 ±0.33	50.17 ±0.06	55.91 ±0.59	19.45 ±0.54	20.17 ±1.35	**15.82** ±0.50
	A2GNN	**50.64** ±1.47	**53.47** ±0.24	22.58 ±0.01	33.04 ±4.12	**76.15** ±0.06	**74.12** ±0.18	59.41 ±0.34	52.01 ±0.32	**61.82** ±0.77	**26.20** ±0.74	**25.78** ±0.25	**16.94** ±0.13
SimGDA+	SimGDA	**55.29** ±0.39	**54.39** ±0.90	53.35 ±0.69	43.04 ±0.86	70.80 ±0.06	67.04 ±0.15	**61.30** ±0.32	53.11 ±0.19	58.27 ±0.16	18.59 ±0.07	15.16 ±0.11	13.27 ±0.04
	SimGDA+	**58.11** ±0.40	**57.52** ±0.38	**57.04** ±0.45	44.17 ±0.02	**73.18** ±0.38	**71.81** ±2.44	**61.53** ±0.08	**53.82** ±0.08	61.60 ±0.11	21.94 ±0.18	**21.36** ±0.09	**15.64** ±0.46

Conclusions

- The performance of current UGDA models varies greatly across different datasets and adaptation scenarios;
- It is crucial to develop tailored strategies to address graph structural shifts, especially when the distribution discrepancies are significant;
- The GNN’s transferability in UGDA heavily relies on two factors: aggregation scope and aggregation architecture, which are influenced by the severity of label shift and the level of graph heterophily, etc;
- The inherent adaptability of GNNs is largely underestimated by existing methods, which motivates the investigate of a simple yet effective model that relies heavily on GNN’s property.

PyGDA

Quickly Build with Only 4 Lines

run.py

```python
from pygda.models import A2GNN

# choose a graph domain adaptation model
model = A2GNN(num_classes=num_classes, device=args.device)

# train the model
model.fit(source_data, target_data)

# evaluate the performance
logits, labels = model.predict(target_data)
```

- **Consistent API and detailed documentation.**
- **Covering all GDA baseline models.**
- **Scalable architecture for batch processing of large graph data.**
- **Fully compatible with PyG data structures.**

PyGDA

Quickly Build with Only 4 Lines

run.py

```python
from pygda.models import A2GNN

# choose a graph domain adaptation model
model = A2GNN(num_classes=num_classes, device=args.device)

# train the model
model.fit(source_data, target_data)

# evaluate the performance
logits, labels = model.predict(target_data)
```

Border Impacts and Limitations

Our benchmark fosters innovation and advances research in graph domain adaptation by providing a standardized evaluation platform, leading to the development of more effective algorithms. This standardization helps researchers compare methods more fairly, driving progress and collaboration within the field.

However, benchmark datasets may introduce limitations that could impact the generalization of findings to real-world scenarios. This risk includes the potential for unrealistic performance expectations if the benchmark does not adequately represent the diversity and complexity of real-world data. We plan to enhance GDABench by including more settings such as source-free and open-set scenarios. This expansion will help to cover a wider range of domain adaptation challenges, thereby fostering the development of algorithms that are not only more robust but also versatile enough to navigate the complexities of diverse and dynamic real-world scenarios. This trajectory in research will be pivotal in advancing the capabilities of domain adaptation techniques, ensuring their applicability and efficacy across various domains and evolving data landscapes.

NEURAL INFORMATION PROCESSING SYSTEMS

Thanks!

Contact Me

Meihan Liu

Zhejiang University

lmh_zju@zju.edu.cn

浙江大學

ZHEJIANG UNIVERSITY