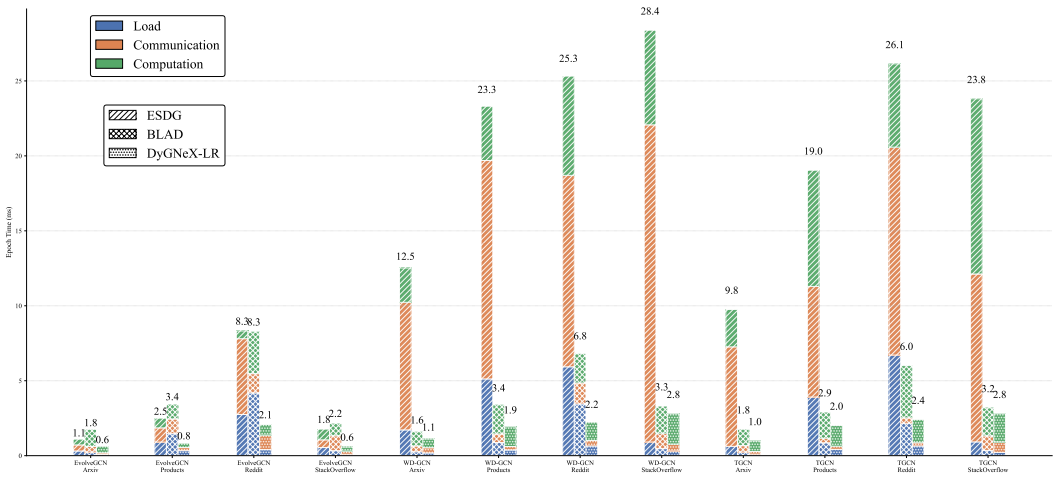

Figure 1: Time breakdown of ESDG, BLAD, and DyGNeX-LR.

Table 1: Time breakdown details

EvolveGCN	Arxiv			Products			Reddit			StackOverflow		
	Load	Comm	Comp	Load	Comm	Comp	Load	Comm	Comp	Load	Comm	Comp
ESDG	0.31	0.40	0.38	0.90	0.95	0.65	2.75	5.06	0.52	0.54	0.51	0.71
BLAD	0.23	0.38	1.15	1.47	0.98	0.99	4.19	1.29	2.80	0.34	0.99	0.84
DyGNeX-LR	0.07	0.11	0.45	0.34	0.20	0.29	0.40	0.97	0.71	0.09	0.21	0.28
WD-GCN	**Arxiv**			**Products**			**Reddit**			**StackOverflow**		
	Load	Comm	Comp	Load	Comm	Comp	Load	Comm	Comp	Load	Comm	Comp
ESDG	1.72	8.51	2.29	5.12	14.57	3.62	5.93	12.77	6.62	0.91	21.16	6.31
BLAD	0.26	0.39	0.95	0.88	0.57	1.95	3.45	1.40	1.96	0.47	1.03	1.81
DyGNeX-LR	0.18	0.35	0.59	0.38	0.23	1.32	0.65	0.36	1.23	0.31	0.46	2.03
TGCN	**Arxiv**			**Products**			**Reddit**			**StackOverflow**		
	Load	Comm	Comp	Load	Comm	Comp	Load	Comm	Comp	Load	Comm	Comp
ESDG	0.62	6.64	2.49	3.91	7.38	7.75	6.71	13.87	5.56	0.92	11.20	11.69
BLAD	0.24	0.43	1.10	0.86	0.30	1.75	2.16	0.36	3.49	0.35	0.97	1.87
DyGNeX-LR	0.10	0.18	0.71	0.42	0.21	1.39	0.65	0.21	1.55	0.22	0.69	1.89

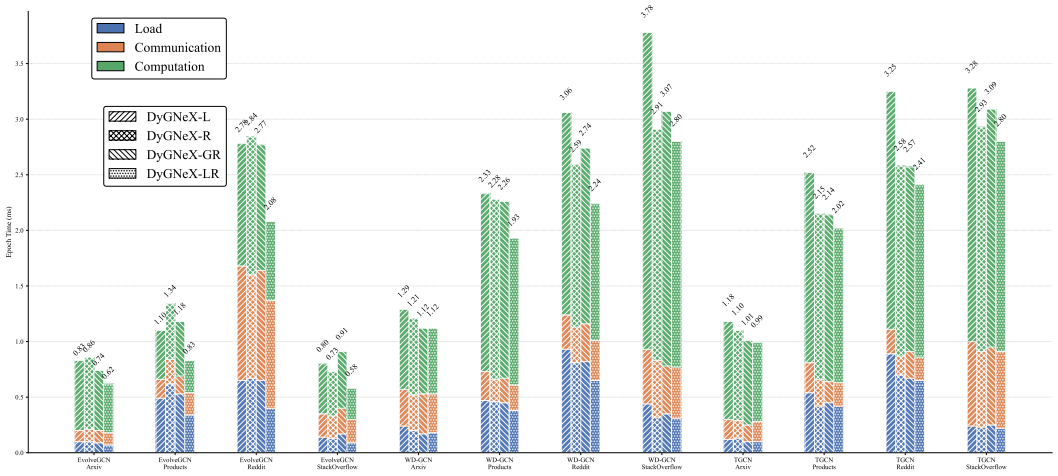

Figure 2: Time breakdown of DyGNeX-L, DyGNeX-R, DyGNeX-GR, DyGNeX-LR.

Table 2: Time breakdown details

EvolveGCN	Arxiv			Products			Reddit			StackOverflow		
	Load	Comm	Comp	Load	Comm	Comp	Load	Comm	Comp	Load	Comm	Comp
DyGNeX-L	0.10	0.10	0.63	0.49	0.17	0.44	0.65	1.03	1.10	0.14	0.21	0.45
DyGNeX-R	0.10	0.11	0.65	0.62	0.22	0.50	0.67	0.93	1.24	0.13	0.20	0.40
DyGNeX-GR	0.09	0.11	0.54	0.53	0.16	0.49	0.65	0.99	1.13	0.17	0.23	0.51
DyGNeX-LR	0.07	0.11	0.44	0.34	0.20	0.29	0.40	0.97	0.71	0.09	0.21	0.28
WD-GCN	**Arxiv**			**Products**			**Reddit**			**StackOverflow**		
	Load	Comm	Comp	Load	Comm	Comp	Load	Comm	Comp	Load	Comm	Comp
DyGNeX-L	0.24	0.33	0.72	0.47	0.26	1.60	0.93	0.31	1.82	0.44	0.49	2.85
DyGNeX-R	0.20	0.32	0.69	0.46	0.20	1.62	0.81	0.32	1.46	0.32	0.51	2.08
DyGNeX-GR	0.17	0.36	0.59	0.45	0.22	1.59	0.82	0.34	1.58	0.35	0.43	2.29
DyGNeX-LR	0.18	0.35	0.59	0.38	0.23	1.32	0.65	0.36	1.23	0.31	0.46	2.03
TGCN	**Arxiv**			**Products**			**Reddit**			**StackOverflow**		
	Load	Comm	Comp	Load	Comm	Comp	Load	Comm	Comp	Load	Comm	Comp
DyGNeX-L	0.12	0.18	0.88	0.54	0.27	1.71	0.89	0.22	2.14	0.24	0.76	2.28
DyGNeX-R	0.13	0.16	0.81	0.42	0.24	1.49	0.70	0.17	1.71	0.23	0.68	2.02
DyGNeX-GR	0.10	0.15	0.76	0.45	0.19	1.50	0.67	0.24	1.66	0.25	0.70	2.14
DyGNeX-LR	0.10	0.18	0.71	0.42	0.21	1.39	0.65	0.21	1.55	0.22	0.69	1.89
GAT-LSTM	**Arxiv**			**Products**			**Reddit**			**StackOverflow**		
	Load	Comm	Comp	Load	Comm	Comp	Load	Comm	Comp	Load	Comm	Comp
DyGNeX-L	0.21	0.20	1.29	0.46	0.29	1.98	0.88	0.36	1.96	0.44	0.29	1.99
DyGNeX-R	0.18	0.21	1.05	0.41	0.33	1.70	0.65	0.44	1.59	0.40	0.28	1.84
DyGNeX-GR	0.17	0.25	0.96	0.44	0.28	1.89	0.80	0.39	1.83	0.45	0.30	2.05
DyGNeX-LR	0.16	0.24	0.93	0.34	0.29	1.46	0.61	0.41	1.44	0.37	0.31	1.66

Table 3: Multi-server multi-GPU (2x2) experimental results. Results obtained using two servers, each equipped with four NVIDIA A100 80GB GPUs. GPUs within each server are connected via NVLink, while the servers are interconnected through a 30Gbps TCP network. Note that BLAD's open source code does not support multi-server training.

Method	Arxiv					Products				
	EvolveGCN	WD-GCN	TGCN	GAT-LSTM	TTGCN	EvolveGCN	WD-GCN	TGCN	GAT-LSTM	TTGCN
ESDG	1.30	13.64	10.33	14.19	19.63	2.65	26.92	25.99	25.31	37.14
BLAD	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
PSG	0.98	1.50	1.46	1.79	4.90	1.54	3.11	3.15	3.80	5.23
DyGNeX-GR	0.89	1.24	1.29	1.69	3.91	1.33	2.68	2.71	3.02	4.69
DyGNeX-LR	0.62	1.10	1.07	1.34	3.37	0.88	2.37	2.44	2.53	3.50
Method	Reddit					Stackoverflow				
	EvolveGCN	WD-GCN	TGCN	GAT-LSTM	TTGCN	EvolveGCN	WD-GCN	TGCN	GAT-LSTM	TTGCN
ESDG	8.86	38.72	40.76	39.40	47.39	1.91	31.09	25.56	29.69	34.42
BLAD	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
PSG	4.34	4.28	4.88	4.90	7.41	1.51	4.15	4.76	4.07	8.92
DyGNeX-GR	3.65	3.27	3.64	4.05	5.98	0.90	3.18	3.71	3.53	7.67
DyGNeX-LR	2.78	2.93	3.12	3.25	4.66	0.58	3.02	3.23	2.96	5.75

Table 4: Multi-server multi-GPU (2x4) experimental results. Results obtained using two servers, each equipped with four NVIDIA A100 80GB GPUs. GPUs within each server are connected via NVLink, while the servers are interconnected through a 30Gbps TCP network. Note that BLAD's open source code does not support multi-server training.

Method	Arxiv					Products				
	EvolveGCN	WD-GCN	TGCN	GAT-LSTM	TTGCN	EvolveGCN	WD-GCN	TGCN	GAT-LSTM	TTGCN
ESDG	1.04	8.18	6.51	9.22	12.56	1.78	18.84	18.19	16.71	25.99
BLAD	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
PSG	0.83	0.94	0.87	1.14	3.20	1.15	2.34	2.21	2.69	3.78
DyGNeX-GR	0.63	0.77	0.70	0.82	2.14	0.81	1.80	1.69	2.11	2.58
DyGNeX-LR	0.41	0.68	0.61	0.75	1.88	0.65	1.45	1.31	1.73	1.95
Method	Reddit					Stackoverflow				
	EvolveGCN	WD-GCN	TGCN	GAT-LSTM	TTGCN	EvolveGCN	WD-GCN	TGCN	GAT-LSTM	TTGCN
ESDG	6.20	23.23	26.08	26.00	28.91	1.32	21.45	17.89	19.59	23.06
BLAD	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
PSG	3.27	2.94	2.83	3.51	4.52	0.96	2.74	2.96	3.29	5.44
DyGNeX-GR	2.14	2.09	2.20	2.41	3.64	0.70	2.15	2.14	2.23	4.30
DyGNeX-LR	1.64	1.67	1.66	1.79	2.80	0.41	1.64	1.88	1.84	3.25

Table 5: Evaluation on 500 snapshots dataset.

Method	**Model**				
	EvolveGCN	WD-GCN	TGCN	GAT-LSTM	TTGCN
1×4 Cluster					
ESDG	4.90	37.46	43.84	42.56	69.43
PSG	4.86	7.15	7.96	7.47	19.75
DyGNeX-GR	3.45	5.67	5.85	5.96	12.82
2×4 Cluster					
ESDG	3.24	25.96	30.87	28.37	43.94
PSG	2.97	4.34	4.64	4.26	10.91
DyGNeX-GR	1.93	3.12	3.28	3.24	6.91

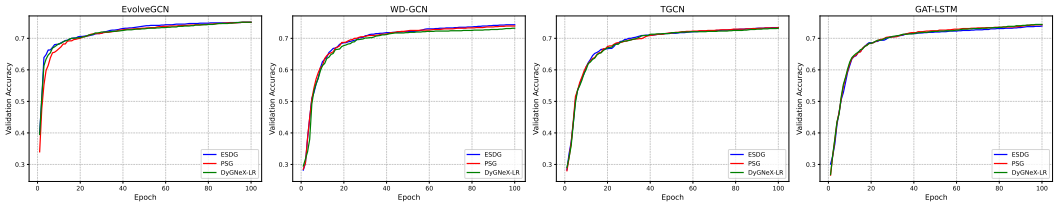

Figure 3: Validation accuracy on Arxiv dataset

Table 6: Test accuracy details of Figure 4 in original paper.

	Arxiv				Products				Reddit			
	EG	WG	TG	GL	EG	WG	TG	GL	EG	WG	TG	GL
ESDG	0.739	0.723	0.715	0.725	0.813	0.802	0.723	0.752	0.945	0.915	0.917	0.903
PSG	0.734	0.722	0.718	0.723	0.813	0.795	0.712	0.76	0.940	0.913	0.912	0.900
DyGNeX-GR	0.733	0.719	0.715	0.721	0.808	0.792	0.717	0.753	0.945	0.909	0.914	0.898
DyGNeX-LR	0.733	0.717	0.714	0.718	0.813	0.790	0.719	0.762	0.946	0.914	0.916	0.894

Table 7: Attributes of the Papers100M-Sample Datasets. The symbols $|V|$ and $|E|$ denote the total number of nodes and edges. $\overline{|V|}$ and $\overline{|E|}$ represent the average number of nodes and edges per snapshot. The term d_v represents the dimension of the node features. The parameters β and γ indicate the average degree and the number of snapshots, respectively.

Dataset	$\lvert V\rvert$	$\lvert E\rvert$	$\overline{\lvert V\rvert}$	$\overline{\lvert E\rvert}$	d_v	β	γ
Papers100M-Sample	1.1M	11.2M	1.0M	10.0M	128	9.9	30

Table 8: Epoch Time (s) for Different Methods Across Various Models on Papers100M-Sample.

	EvolveGCN	WD-GCN	TGCN	GAT-LSTM
ESDG	7.98	31.87	26.4	34.68
BLAD	OOM	OOM	OOM	N/A
PSG	5.80	6.47	8.34	8.02
DyGNeX-GR	4.44	6.05	6.65	6.76
DyGNeX-LR	3.70	4.56	5.89	6.12