

Layer-Neighbor Sampling — Defusing Neighborhood Explosion in GNNs

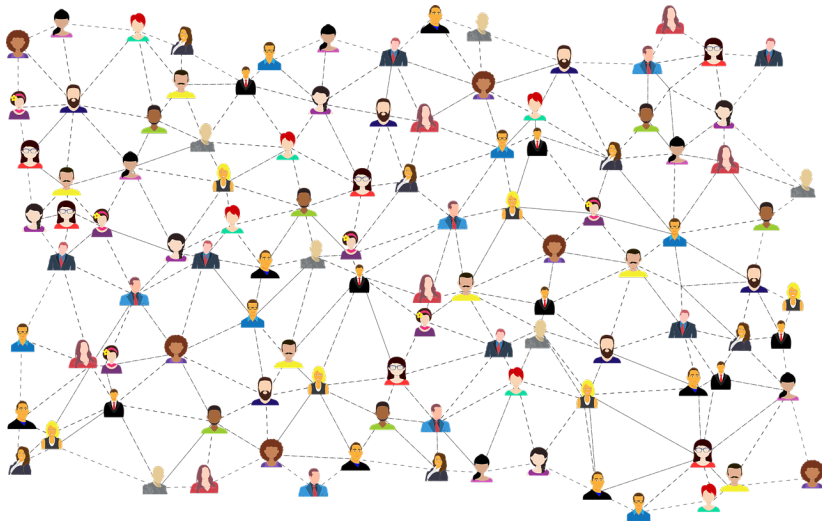
Muhammed Fatih Balın¹ and Ümit V. Çatalyürek¹

Neural Information Processing Systems (NeurIPS'23)
November 13

¹*Computational Science and Engineering, Georgia Institute of Technology*



Applications of GNNs



⊙ Online shopping

⊙ Social media

⊙ Content recommendation

⊙ Showing relevant ads

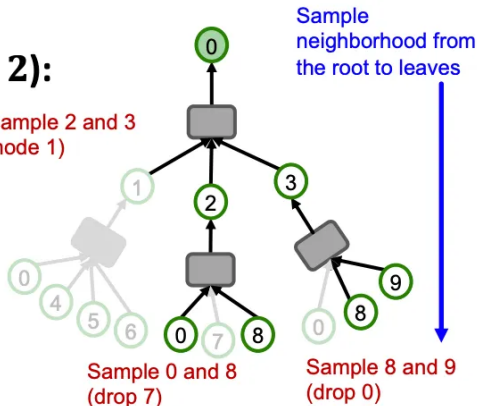
Graph Neural Networks - Neighborhood Explosion

■ Example ($H = 2$):

1st-hop
neighborhood

2nd-hop
neighborhood

First, sample 2 and 3
(drop node 1)



Sampling Goals & Proposed Solution

Goals:

- ⦿ Unbiased sampling

Solution:

Sampling Goals & Proposed Solution

Goals:

- ⊙ Unbiased sampling
- ⊙ Overlapping neighborhoods
(Layer-wise sampling: LADIES)

Solution:

Sampling Goals & Proposed Solution

Goals:

- ⊙ Unbiased sampling
- ⊙ Overlapping neighborhoods
(Layer-wise sampling: LADIES)
- ⊙ Uniformly good approximation
(Node-wise sampling: NS)

Solution:

Sampling Goals & Proposed Solution

Goals:

- ⊙ Unbiased sampling
- ⊙ Overlapping neighborhoods
(Layer-wise sampling: LADIES)
- ⊙ Uniformly good approximation
(Node-wise sampling: NS)

Solution:

- ⊙ Poisson Sampling - flip biased coins: $r \leq \pi, r \sim U(0, 1)$

Sampling Goals & Proposed Solution

Goals:

- ⊙ Unbiased sampling
- ⊙ Overlapping neighborhoods
(Layer-wise sampling: LADIES)
- ⊙ Uniformly good approximation
(Node-wise sampling: NS)

Solution:

- ⊙ Poisson Sampling - flip biased coins: $r \leq \pi, r \sim U(0, 1)$
- ⊙ **Combine NS & LADIES to get best-of-both-worlds LABOR.**

Sampling Goals & Proposed Solution

Goals:

- ⊙ Unbiased sampling
- ⊙ Overlapping neighborhoods
(Layer-wise sampling: LADIES)
- ⊙ Uniformly good approximation
(Node-wise sampling: NS)

Solution:

- ⊙ Poisson Sampling - flip biased coins: $r \leq \pi, r \sim U(0, 1)$
- ⊙ Combine NS & LADIES to get best-of-both-worlds LABOR.
- ⊙ Generalizes to any unbiased sampling method.

Layer-Neighbor Sampling

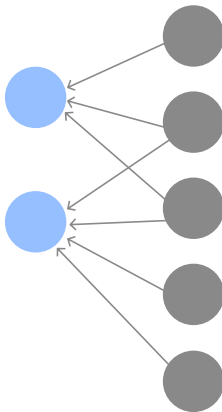
- ⊙ LABOR-0: given s and $r_t \sim U(0, 1)$,
sample $t \rightarrow s$ if $r_t \leq \frac{k}{d_s}$,
 k is the fan-out, d_s is the in-degree.

Layer-Neighbor Sampling

- ⊙ LABOR-0: given s and $r_t \sim U(0, 1)$,
sample $t \rightarrow s$ if $r_t \leq \frac{k}{d_s}$,
 k is the fan-out, d_s is the in-degree.
- ⊙ k sampled items in expectation, matching NS.

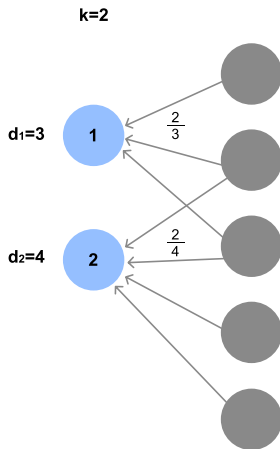
LABOR-0 example

Blue: seed vertices, green: sampled, gray: not sampled



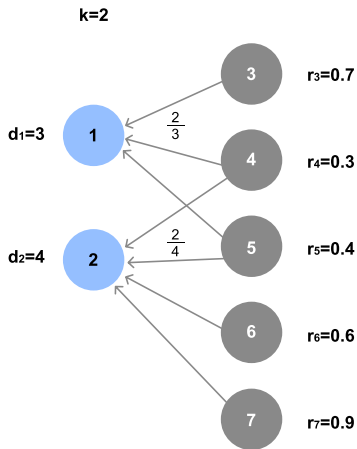
LABOR-0 example

Blue: seed vertices, green: sampled, gray: not sampled



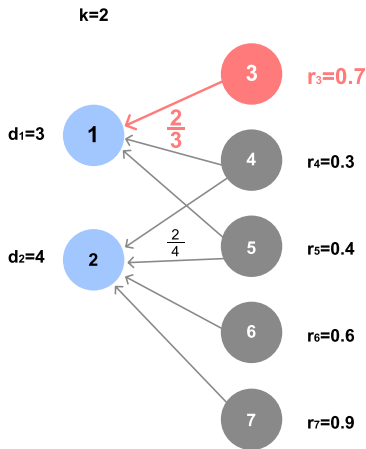
LABOR-0 example

Blue: seed vertices, green: sampled, gray: not sampled



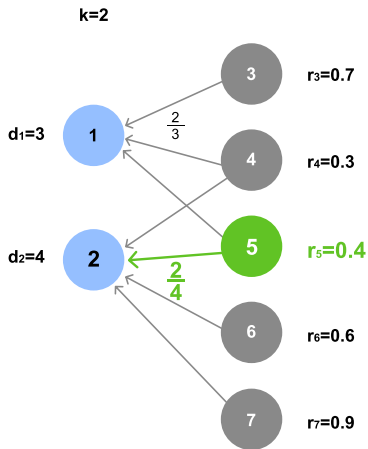
LABOR-0 example

Blue: seed vertices, green: sampled, gray: not sampled



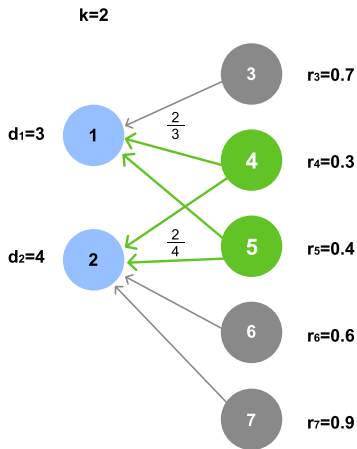
LABOR-0 example

Blue: seed vertices, green: sampled, gray: not sampled



LABOR-0 example

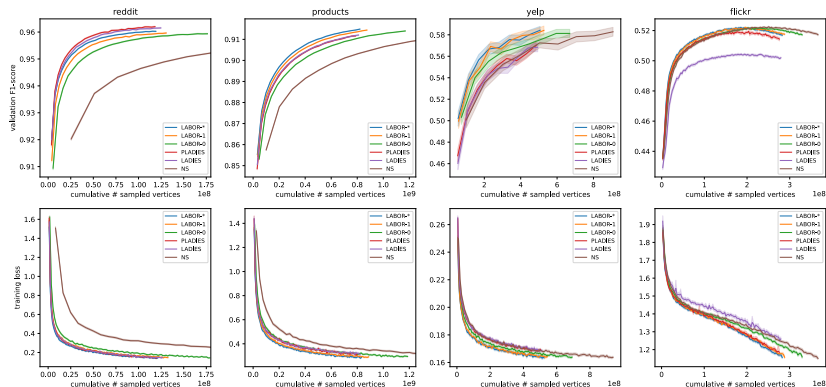
Blue: seed vertices, green: sampled, gray: not sampled



Properties of the datasets used in experiments.

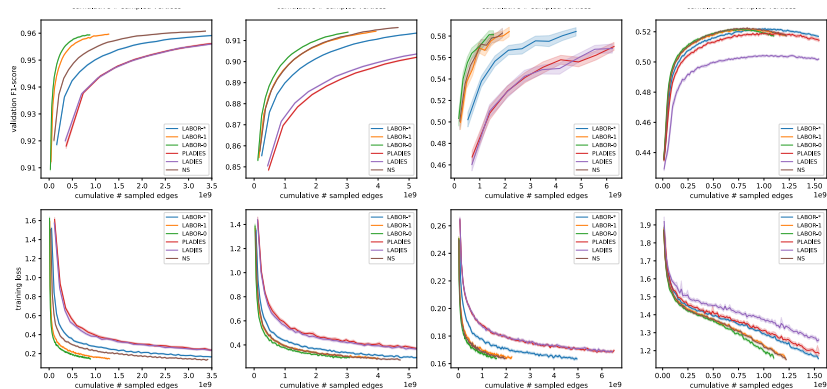
Dataset	$ V $	$ E $	$\frac{ E }{ V }$	# feats.	train - val - test (%)
reddit	233K	115M	493.56	602	66.00 - 10.00 - 24.00
products	2.45M	61.9M	25.26	100	8.00 - 2.00 - 90.00
yelp	717K	14.0M	19.52	300	75.00 - 10.00 - 15.00
flickr	89.2K	900K	10.09	500	50.00 - 25.00 - 25.00

Experiments - Vertex efficiency



The validation F1-score and training loss curves. The x-axis is scaled w.r.t. # cumulative sampled vertices.

Experiments - Edge efficiency



The validation F1-score and training loss curves. The x-axis is scaled w.r.t. # cumulative sampled edges.

Experiments - PLADIES and LABOR Evaluation

Dataset	Algo.	V ³	E ²	V ²	E ¹	V ¹	E ⁰	V ⁰	it/s	test F1-score
reddit	PLADIES	24	2390	14.1	927	6.0	33.2	1	1.7	96.21 ± 0.06
	LADIES	25	2270	14.5	852	6.0	32.5	1	1.8	96.20 ± 0.05
	LABOR-*	24	1070	13.7	435	6.0	26.9	1	4.1	96.23 ± 0.05
	LABOR-1	27	261	14.4	116	6.1	16.7	1	24.8	96.23 ± 0.06
	LABOR-o	36	177	17.8	67	6.8	9.6	1	37.6	96.25 ± 0.05
	NS	167	682	68.3	100	10.1	9.7	1	14.2	96.24 ± 0.05
products	PLADIES	160	2380	51.2	293	9.7	11.7	1	4.1	78.44 ± 0.24
	LADIES	165	2230	51.8	270	9.7	11.5	1	4.2	78.59 ± 0.22
	LABOR-*	166	1250	51.8	167	9.8	10.6	1	6.2	78.59 ± 0.34
	LABOR-1	178	799	53.4	136	9.8	10.5	1	21.3	78.47 ± 0.26
	LABOR-o	237	615	62.4	100	10.1	9.9	1	32.5	78.76 ± 0.26
	NS	513	944	95.4	106	10.6	9.9	1	24.6	78.48 ± 0.29
yelp	PLADIES	100	1300	29.5	183	6.2	6.9	1	5.1	61.55 ± 0.87
	LADIES	102	1280	29.7	182	6.2	6.9	1	5.3	61.89 ± 0.66
	LABOR-*	105	991	30.7	158	6.1	6.8	1	13.3	61.57 ± 0.67
	LABOR-1	109	447	31.0	96	6.2	6.8	1	27.3	61.71 ± 0.70
	LABOR-o	138	318	35.1	54	6.2	6.3	1	27.2	61.55 ± 0.85
	NS	188	392	42.5	55	6.3	6.3	1	23.0	61.50 ± 0.66
flickr	PLADIES	55	309	24.9	85	6.2	6.9	1	10.2	51.52 ± 0.26
	LADIES	56	308	25.1	85	6.2	6.9	1	10.5	50.79 ± 0.29
	LABOR-*	57	308	25.6	85	6.3	6.9	1	20.3	51.67 ± 0.27
	LABOR-1	58	242	25.9	73	6.3	6.9	1	32.7	51.66 ± 0.24
	LABOR-o	66	219	29.1	52	6.4	6.7	1	33.3	51.65 ± 0.26
	NS	73	244	32.8	52	6.4	6.7	1	31.7	51.70 ± 0.23

Importance Sampling Experiments - Fixed point iterations

Number of vertices (in thousands) in 3rd layer w.r.t # fixed point iterations (its). * denotes applying the fixed point iterations until convergence, i.e., LABOR-*, 1 its stands for LABOR-1 etc.

Dataset	NS	0	1	2	3	*
reddit	167	36	27	25	25	24
products	513	237	178	170	169	166
yelp	188	138	109	106	105	105
flickr	73	66	58	57	57	56

$$\pi_t^{(i+1)} = \pi_t^{(i)} \max_{t \rightarrow s} c_s(\pi^{(i)})$$

Thanks!

⦿ For more information, scan:



- Email: balin@gatech.edu
- Visit: mfbal.in
- Visit: tda.gatech.edu

⦿ Acknowledgement of Support:



THANK YOU