You are AllSet: A Multiset Learning Framework for Hypergraph Neural Networks

Eli Chien[1] Chao Pan[1] Jianhao Peng[1] Olgica Milenkovic[1]

University of Illinois Urbana-Champaign, Illinois, USA[1]

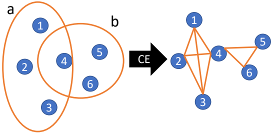
ICLR 2022

Our code is available online: `https://github.com/jianhao2016/AllSet`

Machine learning on hypergraphs

- Hypergraph, where a hyperedge can contain more than two nodes, is better in modeling higher-order relations.

- Examples: subspace clustering, hierarchical species classification of a FoodWeb.



- Naive approach: clique-expansion (CE).
  - Issue: known to lose information.[1,2,3]

[1] The total variation on hypergraphs-learning on hypergraphs revisited, Hein et al. NeurIPS 2013.

[2] Submodular hypergraphs: p-laplacians, cheeger inequalities and spectral clustering, Li et al. ICML 2018.

[3] HS^2: Active learning over hypergraphs with pointwise and pairwise queries, Chien et al. AISTATS 2019.

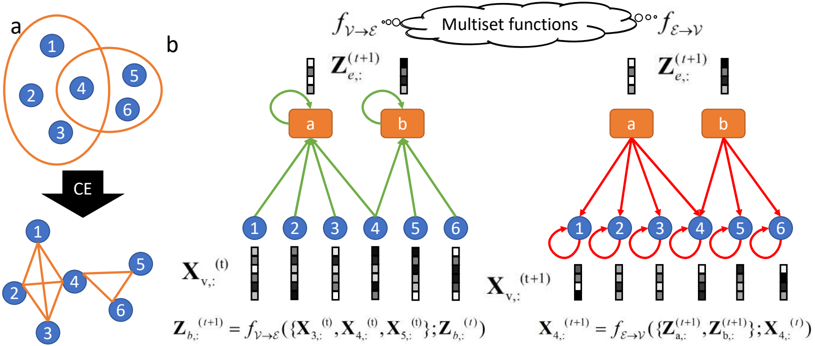
Hypergraph learning beyond CE

- Many sophisticated propagation rules directly applicable on hypergraphs, and related to tensor eigenproblems, have been studied as well.
  - Multilinear PageRank[4]: related to Z eigenproblem in spectral hypergraph theory.
  - Tudisco et al. also show that CE-based propagation can be suboptimal on several tasks[5].
- Natural questions:
  - Is there a general framework that includes CE-based, Z-based and other propagations on hypergraphs?
  - Can we learn propagation schemes for hypergraph neural networks suitable for different datasets and different learning tasks?

[4] Multilinear pagerank, Gleich et al. SIMAX 2015.
[5] Nonlinear higher-order label spreading, Tudisco et al. WWW 2021.

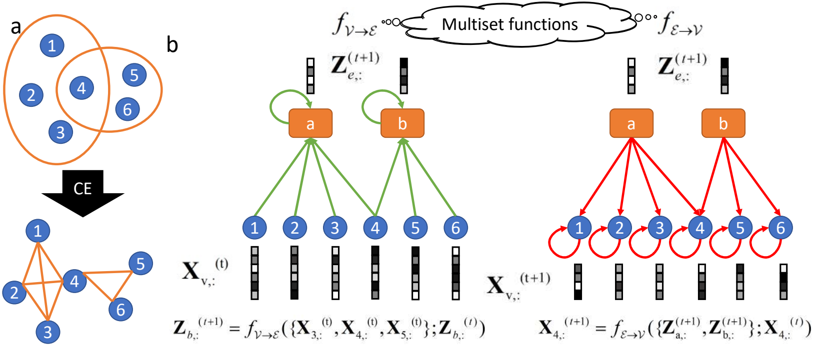
The answer: AllSet framework

- We give affirmative answers to both questions.
- AllSet: a general multiset function framework.



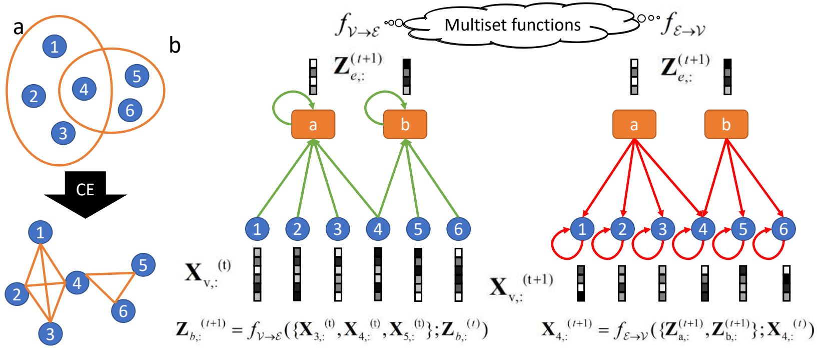
Expressive power of AllSet framework

- We prove that many existing hypergraph neural network layers have strictly less expressive power compare to AllSet framework.
- Idea: AllSet can model both these layers and Z-based propagation, but these layers cannot model Z-based propagation.



Expressive power of AllSet framework

- We also show that AllSet framework is a generalization of Message Passing Neural Network (MPNN).



Designing learnable AllSet layers

- While being powerful, it is still unclear how to design learnable AllSet layers with the same expressive power?
- Equipped with AllSet methodology, an intuitive way is to replace both $f_{\mathcal{V}\rightarrow\mathcal{E}}$ and $f_{\mathcal{E}\rightarrow\mathcal{V}}$ with deep set function learners.
  - DeepSets[6] and Set Transformer[7]
- Both are provably universal set function approximators.
- Hence, we get two hypergraph neural network layers, AllDeepSets and AllSetTransformer, which has the same expressive power as the general AllSet framework.

[6] Deep Sets, Zaheer et al. NeurIPS 2017

[7] Set Transformer: A framework for attention-based permutation-invariant neural networks, Lee et al. ICML 2019.

Experiments

- We focus on node classification task in this work.
- Graph case: OGB or PyG library including many benchmark datasets.
- Hypergraph learning mostly focus on citation and coauthor networks only.
- In our codebase, we not only collect 8 different datasets from literatures, but also curate 3 new benchmark datasets.
- Hence, our work also make a first step toward benchmarking hypergraph neural networks.

Experiments

Table 2: Results for the tested datasets: Mean accuracy (%) $\pm$ standard deviation. Boldfaced letters shaded grey are used to indicate the best result, while blue shaded boxes indicate results within one standard deviation of the best result. NA indicates that the method has numerical precision issue. For HAN*, additional preprocessing of each dataset is required (see the Section 6 for more details).

	Cora	Citeseer	Pubmed	Cora-CA	DBLP-CA	Zoo	20Newsgroups	mushroom
AllSetTransformer	78.59 ± 1.47	73.08 ± 1.20	88.72 ± 0.37	83.63 ± 1.47	91.53 ± 0.23	**97.50 ± 3.59**	81.38 ± 0.58	**100.00 ± 0.00**
AllDeepSets	76.88 ± 1.80	70.83 ± 1.63	**88.75 ± 0.33**	81.97 ± 1.50	91.27 ± 0.27	95.39 ± 4.77	81.06 ± 0.54	99.99 ± 0.02
MLP	75.17 ± 1.21	72.67 ± 1.56	87.47 ± 0.51	74.31 ± 1.89	84.83 ± 0.22	87.18 ± 4.44	**81.42 ± 0.49**	**100.00 ± 0.00**
CECGN	76.17 ± 1.39	70.16 ± 1.31	86.45 ± 0.43	77.05 ± 1.26	88.00 ± 0.26	51.54 ± 11.19	OOM	95.27 ± 0.47
CEGAT	76.41 ± 1.53	70.63 ± 1.30	86.81 ± 0.42	76.16 ± 1.19	88.59 ± 0.29	47.88 ± 14.03	OOM	96.60 ± 1.67
HNHN	76.36 ± 1.92	72.64 ± 1.57	86.90 ± 0.30	77.19 ± 1.49	86.78 ± 0.29	93.59 ± 5.88	81.35 ± 0.61	**100.00 ± 0.01**
HGNN	79.39 ± 1.36	72.45 ± 1.16	86.44 ± 0.44	82.64 ± 1.65	91.03 ± 0.20	92.50 ± 4.58	80.33 ± 0.42	98.73 ± 0.32
HCHA	79.14 ± 1.02	72.42 ± 1.42	86.41 ± 0.36	82.55 ± 0.97	90.92 ± 0.22	93.65 ± 6.15	80.33 ± 0.80	98.70 ± 0.39
HyperGCN	78.45 ± 1.26	71.28 ± 0.82	82.84 ± 8.67	79.48 ± 2.08	89.38 ± 0.25	N/A	81.05 ± 0.59	47.90 ± 1.04
UniGCNII	78.81 ± 1.05	73.05 ± 2.21	88.25 ± 0.40	83.60 ± 1.14	**91.69 ± 0.19**	93.65 ± 4.37	81.12 ± 0.67	99.96 ± 0.05
HAN (full batch)*	**80.18 ± 1.15**	74.05 ± 1.43	86.21 ± 0.48	**84.04 ± 1.02**	90.89 ± 0.23	85.19 ± 8.18	OOM	90.86 ± 2.40
HAN (mini batch)*	79.70 ± 1.77	**74.12 ± 1.52**	85.32 ± 2.25	81.71 ± 1.73	90.17 ± 0.65	75.77 ± 7.10	79.72 ± 0.62	93.45 ± 1.31

	NTU2012	ModelNet40	Yelp	House(1)	Walmart(1)	House(0.6)	Walmart(0.6)	avg. ranking (↑)
AllSetTransformer	88.69 ± 1.24	**98.20 ± 0.20**	**36.89 ± 0.51**	69.33 ± 2.20	**65.46 ± 0.25**	83.14 ± 1.92	**78.46 ± 0.40**	**2.00**
AllDeepSets	88.09 ± 1.52	96.98 ± 0.26	30.36 ± 1.57	67.82 ± 2.40	64.55 ± 0.33	80.70 ± 1.59	**78.46 ± 0.26**	4.47
MLP	85.52 ± 1.49	96.14 ± 0.36	31.96 ± 0.44	67.93 ± 2.33	45.51 ± 0.24	81.53 ± 2.26	63.28 ± 0.37	6.27
CECGN	81.52 ± 1.43	89.92 ± 0.46	OOM	62.80 ± 2.61	54.44 ± 0.24	64.36 ± 2.41	59.78 ± 0.32	9.66
CEGAT	82.21 ± 1.23	92.52 ± 0.39	OOM	69.09 ± 3.00	51.14 ± 0.56	77.25 ± 2.53	59.47 ± 1.05	8.80
HNHN	89.11 ± 1.44	97.84 ± 0.25	31.65 ± 0.44	67.80 ± 2.59	47.18 ± 0.35	78.78 ± 1.88	65.80 ± 0.39	5.87
HGNN	87.72 ± 1.35	95.44 ± 0.33	33.04 ± 0.62	61.39 ± 2.96	62.00 ± 0.24	66.16 ± 1.80	77.72 ± 0.21	5.73
HCHA	87.48 ± 1.87	94.48 ± 0.28	30.99 ± 0.72	61.36 ± 2.53	62.45 ± 0.26	67.91 ± 2.26	77.12 ± 0.26	6.40
HyperGCN	56.36 ± 4.86	75.89 ± 5.26	29.42 ± 1.54	48.31 ± 2.93	44.74 ± 2.81	78.22 ± 2.46	55.31 ± 0.30	9.87
UniGCNII	**89.30 ± 1.33**	98.07 ± 0.23	31.70 ± 0.52	67.25 ± 2.57	54.45 ± 0.37	80.65 ± 1.96	72.08 ± 0.28	3.87
HAN (full batch)*	83.58 ± 1.46	94.04 ± 0.41	OOM	**71.05 ± 2.26**	OOM	**83.27 ± 1.62**	OOM	6.73
HAN (mini batch)*	80.77 ± 2.36	91.52 ± 0.96	26.05 ± 1.37	62.00 ± 9.06	48.57 ± 1.04	82.04 ± 2.68	63.10 ± 0.96	7.60

Thanks for your attention!

Please also check our work on improving GNNs with raw data[8]!

[8] Node Feature Extraction by Self-Supervised Multi-scale Neighborhood Prediction, Chien et al. ICLR 2022