Hypergraph Contrastive Learning

2025-08-28

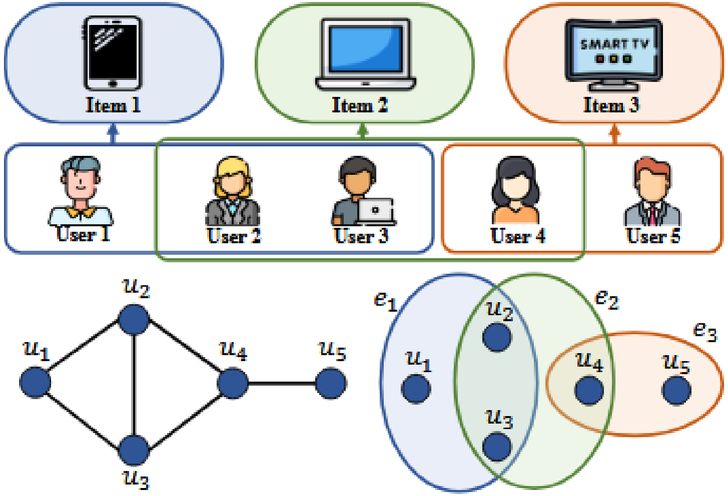
InHyeok Jeong
School of Computer Science and Engineering
Chung-Ang University
MINDS Lab

Content

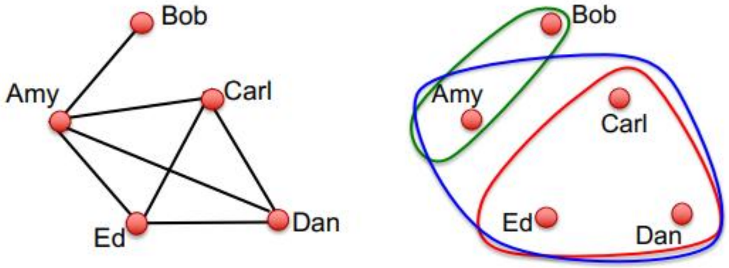
- ☐ Background
- ☐ HyperGCL
- ☐ TriCL
- ☐ CASH
- ☐ Conclusion

Background

- Hypergraphs can naturally model group-wise relations as hyperedges
  - Using graphs can incur information loss



(a) Graph (b) Hypergraph



(a) Graph model (b) Hypergraph model

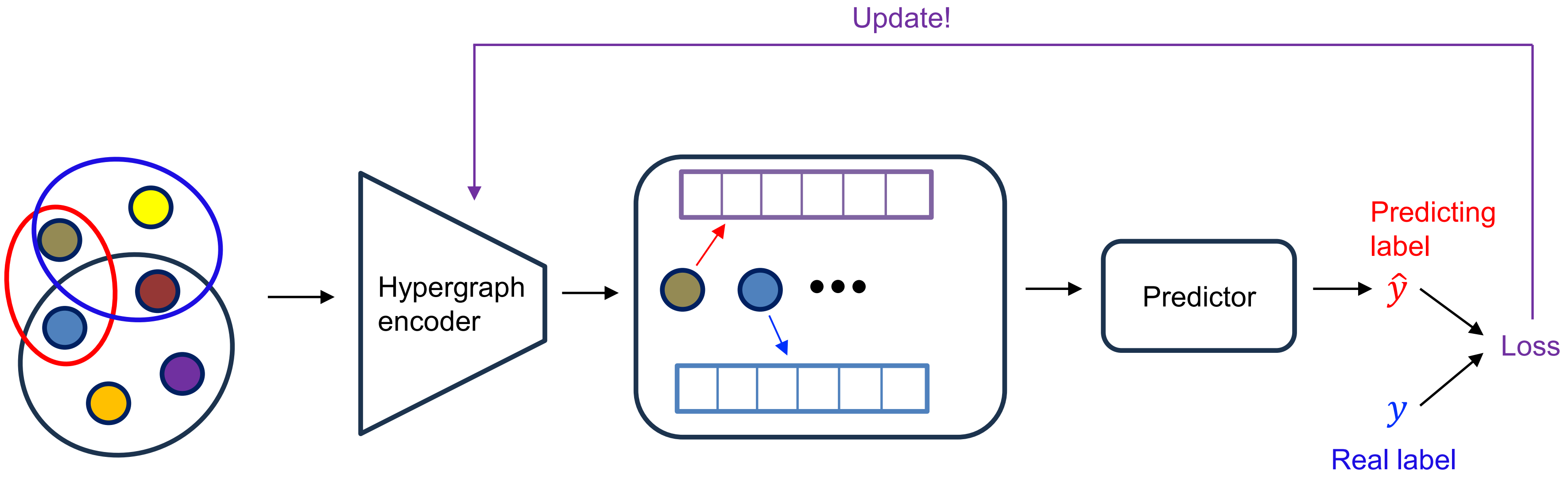
Fig. 2. Graph and hypergraph representations of Fig. 1 data. Colored hyperedges correspond to different email messages.

Ko, Yunyong, Hanghang Tong, and Sang-Wook Kim. "Enhancing hyperedge prediction with context-aware self-supervised learning." *IEEE Transactions on Knowledge and Data Engineering* (2025).

Wolf, Michael M., Alicia M. Klinvex, and Daniel M. Dunlavy. "Advantages to modeling relational data using hypergraphs versus graphs." *2016 IEEE high performance extreme computing conference (HPEC)*. IEEE, 2016.

Background

□ Hypergraph Neural Networks



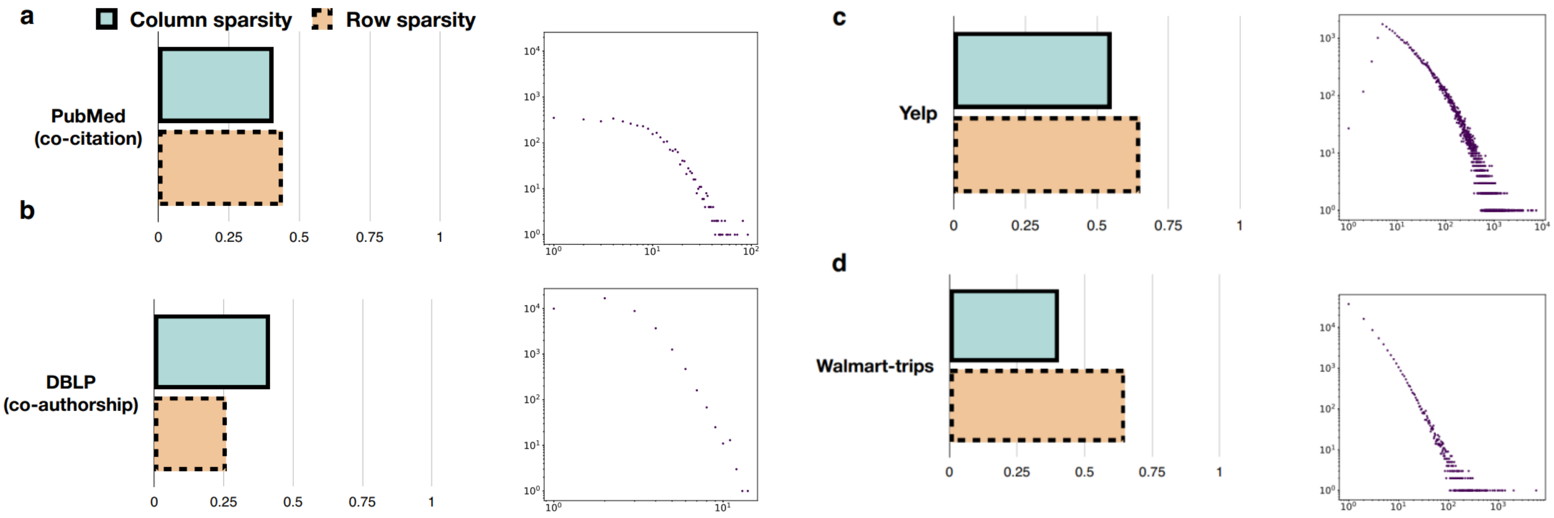
Background

- Hypergraph data labeling is often time, resource, and labor-intensive

Background

- **Real world (hyper)graphs are 'sparse'**
  - Most objects have only a few relationships

Incidence matrix

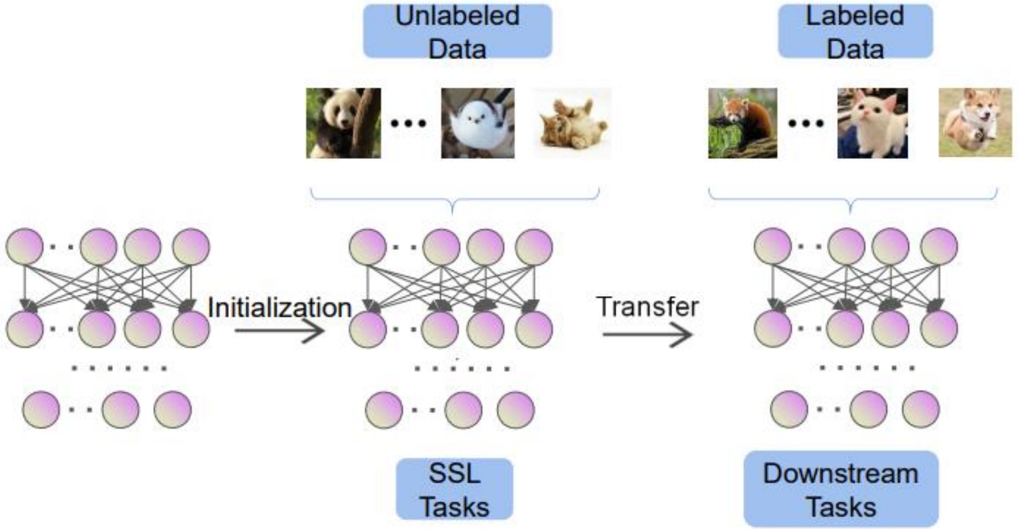


Ye, Bodian, et al. "EasyHypergraph: an open-source software for fast and memory-saving analysis and learning of higher-order networks." *Humanities and Social Sciences Communications* 12.1 (2025): 1-19.

Background

☐ Self-Supervised Learning

- Learn discriminative features from vast quantities of unlabeled instances without relying on human annotations



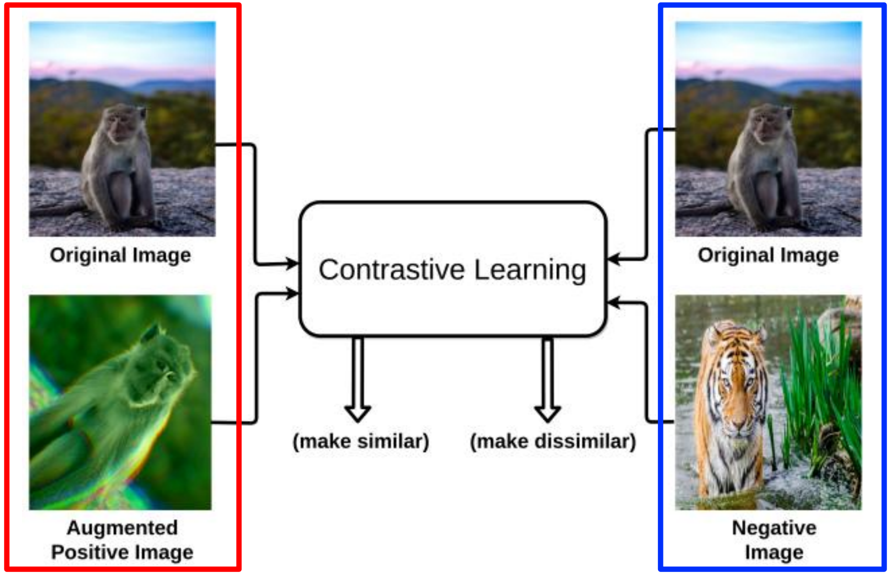
Gui, Jie, et al. "A survey on self-supervised learning: Algorithms, applications, and future trends." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 46.12 (2024): 9052-9071.

Background

☐ Contrastive Learning

- Push original and augmented data closer, push original and negative data away

Promote proximity



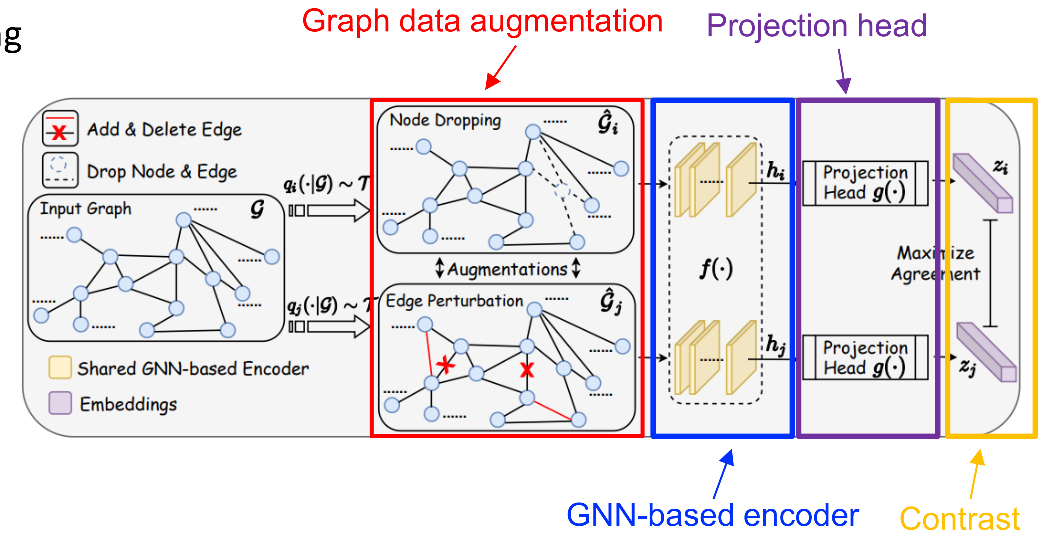
Maximize the separation
between negative pairs

Jaiswal, Ashish, et al. "A survey on contrastive self-supervised learning." *Technologies* 9.1 (2020): 2.

Background

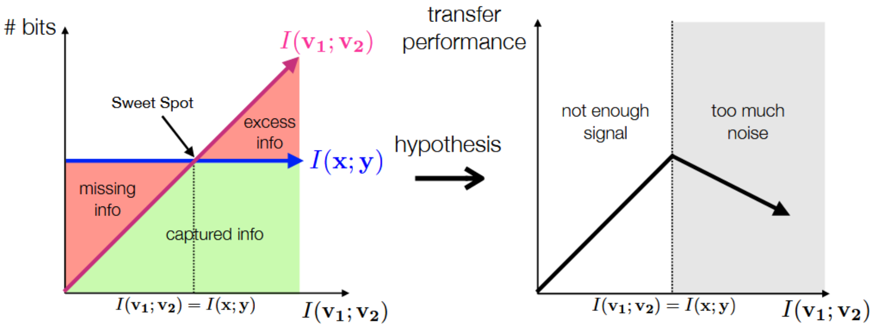
☐ Graph Contrastive Learning

- Graph data augmentation
  - ☐ Dropping, perturbation and masking
- GNN-based encoder
  - ☐ Extract graph-level representation vectors for augmented graphs
- Projection head
- Contrastive
  - ☐ Dependent on task, generally node-level contrast
  - ☐ Maximize the consistency between positive pairs compared with negative pairs

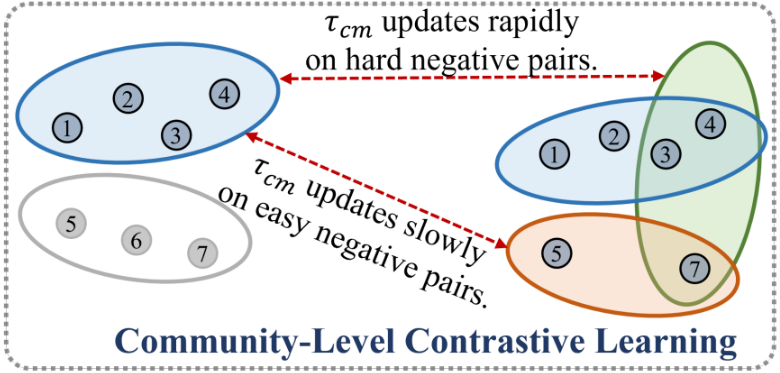


Challenges

- ☐ **How to augment a hypergraph?**
  - The choice of views is what controls the information the representation captures



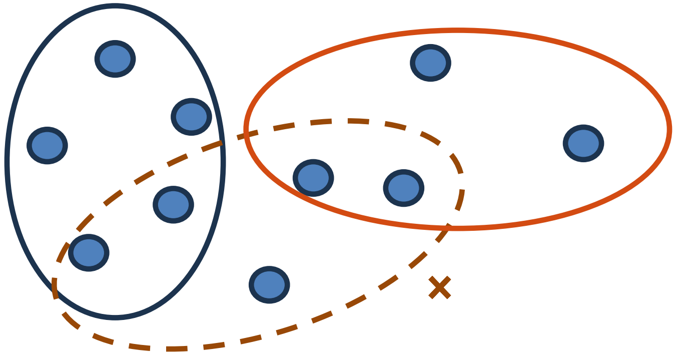
- ☐ **What to contrast?**
  - Node-only contrast cannot reflect higher-order information



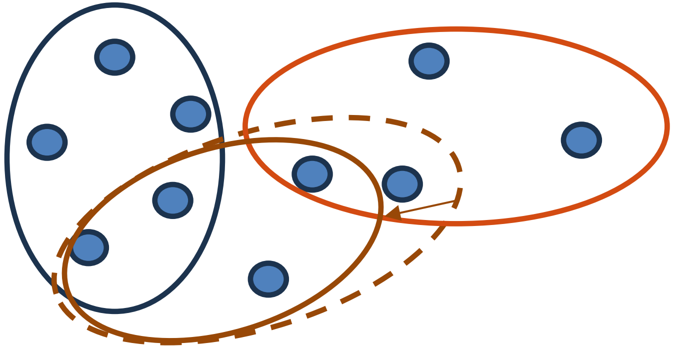
Tian, Yonglong, et al. "What makes for good views for contrastive learning?." *Advances in neural information processing systems* 33 (2020): 6827-6839.
Qian, Yiyue, et al. "Dual-level hypergraph contrastive learning with adaptive temperature enhancement." *Companion Proceedings of the ACM Web Conference 2024*. 2024.

How to augment a hypergraph?

- ☐ **Augment Hyperedges**
  - Hyperedge Perturbation
    - ☐ Partially missing higher-order relations do not significantly affect the semantic meaning of hypergraphs
  - Generalized Hyperedge Perturbation
    - ☐ Randomly kick out vertices from hyperedges



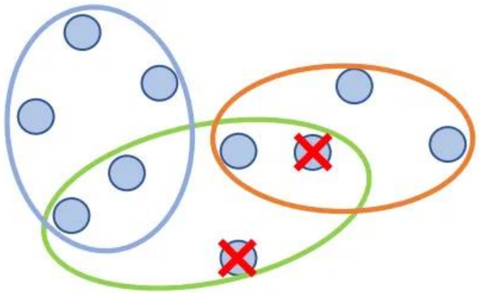
Hyperedge Perturbation



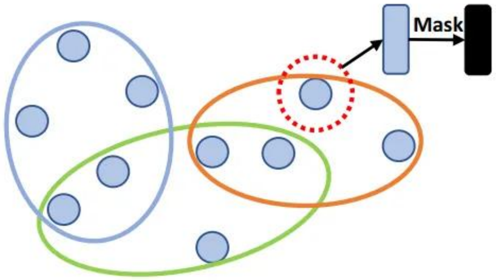
Generalized Hyperedge Perturbation

How to augment a hypergraph?

- ☐ **Augment Vertexes**
  - ■ Vertex dropping
    - ☐ Vertex missing does not alter semantics
  - ■ Attribute masking
    - ☐ Semantic robustness against losing partial attributes



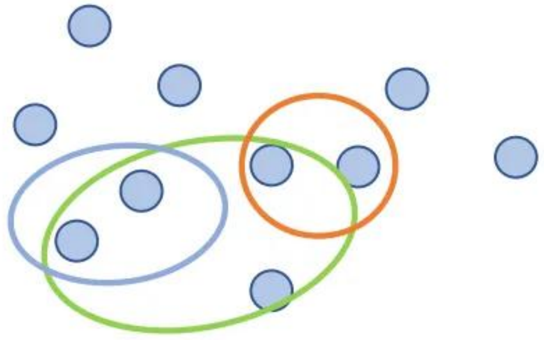
A3: Drop Vertexes



A4: Feature Mask

How to augment a hypergraph?

- **Subgraph**
  - Local structure can hint the full semantics
  - Perform random walk to extract



A5: Subgraph

Augmentations in Hypergraph Contrastive Learning: Fabricated and Generative

Tianxin Wei, Yuning You, Tianlong Chen, Yang Shen, Jingrui He, Zhangyang Wang

University of Illinois Urbana-Champaign, Texas A&M University, University of Texas at Austin

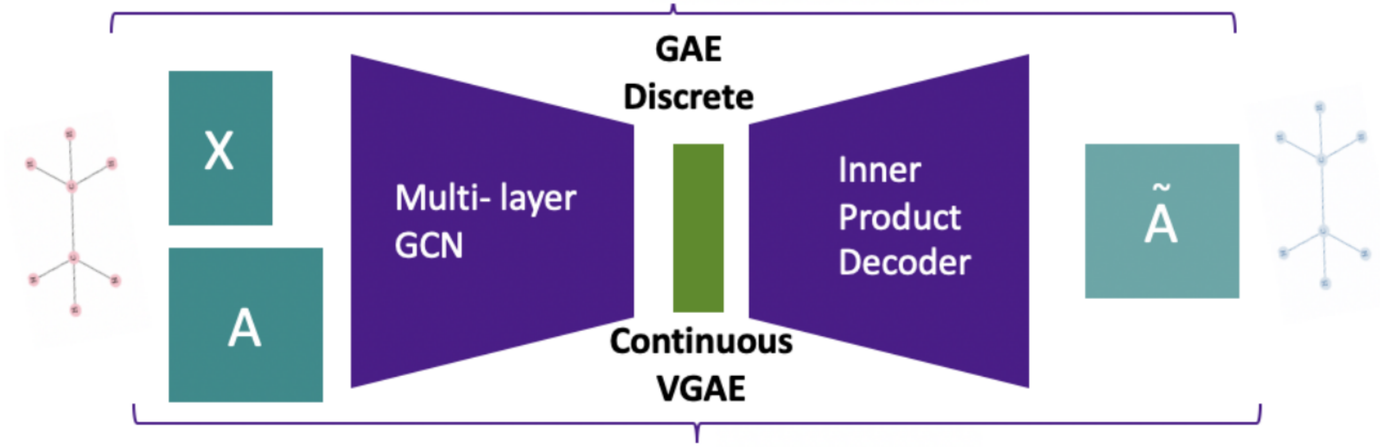
22-NeurIPS

HyperGCL

☐ Motivation

- Augmentation of hypergraph can be learned *during contrastive learning!*

  → Leverage Variational Graph Auto-Encoder!



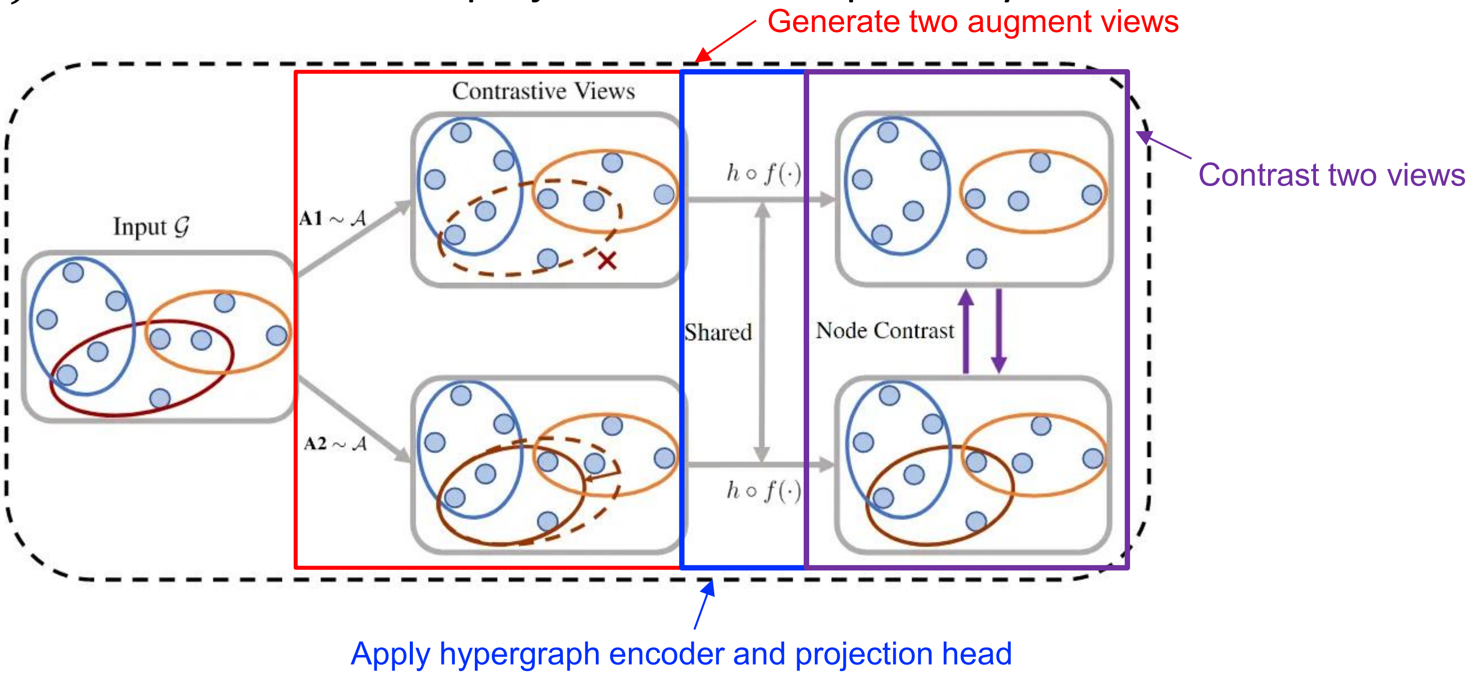
Woodward, Amelia. "Evaluating autoencoder methods for building a molecule graph autoencoder." *Chemistry* 6.12: 1044-1048.

HyperGCL

How to augment a hypergraph?	What to contrast?
O	X

□ Overview

- $f(\cdot), h(\cdot)$: shared encoder and projection head respectively

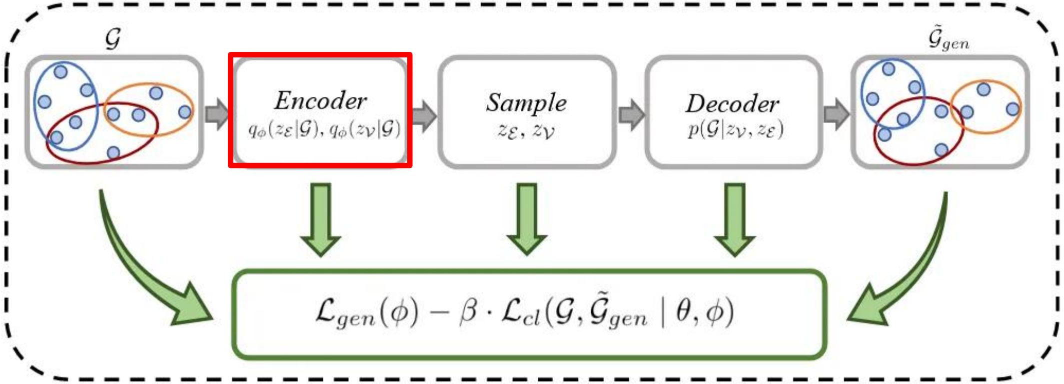


Wei, Tianxin, et al. "Augmentations in hypergraph contrastive learning: Fabricated and generative." *Advances in neural information processing systems* 35 (2022): 1909-1922.

HyperGCL

☐ Variational Hypergraph Auto-Encoder (VHGAE)

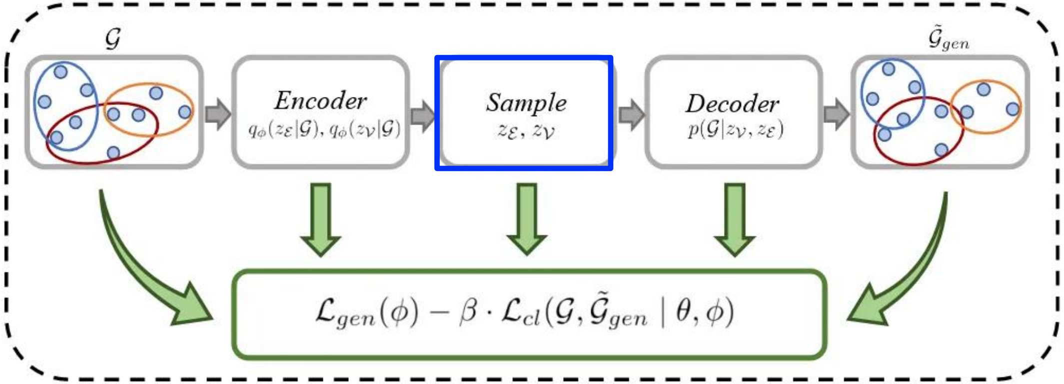
- Embed hypergraphs into latent representations
  - ☐ $z_{\mathcal{V}} \sim q_{\phi}(z_{\mathcal{V}}|\mathcal{G}) = \mathcal{N}(\mu_{\mathcal{V}}, \sigma_{\mathcal{V}}^2);\ \mu_{\mathcal{V}} = HyperGNN_{\mu}^{\mathcal{V}}(\mathcal{G}), \log(\sigma_{\mathcal{V}}) = HyperGNN_{\sigma}^{\mathcal{V}}(\mathcal{G})$
  - ☐ $z_{\mathcal{E}} \sim q_{\phi}(z_{\mathcal{E}}|\mathcal{G}) = \mathcal{N}(\mu_{\mathcal{E}}, \sigma_{\mathcal{E}}^2);\ \mu_{\mathcal{E}} = HyperGNN_{\mu}^{\mathcal{E}}(\mathcal{G}), \log(\sigma_{\mathcal{E}}) = HyperGNN_{\sigma}^{\mathcal{E}}(\mathcal{G})$



Wei, Tianxin, et al. "Augmentations in hypergraph contrastive learning: Fabricated and generative." *Advances in neural information processing systems* 35 (2022): 1909-1922.

HyperGCL

- Variational Hypergraph Auto-Encoder (VHGAE)
  - Apply reparameterization trick
    - $z_{\mathcal{V}} = \mu_{\mathcal{V}} + \sigma_{\mathcal{V}} \odot \delta$
    - $z_{\mathcal{E}} = \mu_{\mathcal{E}} + \sigma_{\mathcal{E}} \odot \delta$
    - $\delta \sim \mathcal{N}(0, I)$



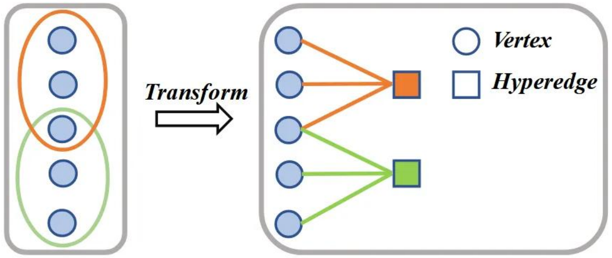
Wei, Tianxin, et al. "Augmentations in hypergraph contrastive learning: Fabricated and generative." *Advances in neural information processing systems* 35 (2022): 1909-1922.

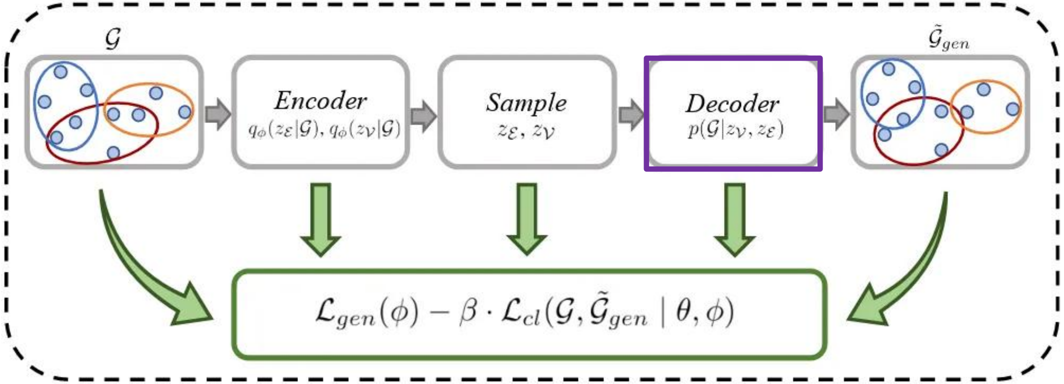
HyperGCL

□ Variational Hypergraph Auto-Encoder (VHGAE)

- Reconstruct the higher-order relations of hypergraphs
- Recover the relations on the converted bipartite graph $\tilde{\mathcal{G}} = \{\tilde{\mathcal{V}}, \tilde{\mathcal{E}}\}$

$$p(\mathcal{G}|z_{\mathcal{V}}, z_{\mathcal{E}}) \approx p(\tilde{\mathcal{G}}|z_{\mathcal{V}}, z_{\mathcal{E}}) = \prod_{e=1}^{|\mathcal{E}|} \prod_{v=1}^{|\mathcal{V}|} p(\tilde{\mathcal{E}}_{v,e}|z_v, z_e) = \prod_{e=1}^{|\mathcal{E}|} \prod_{v=1}^{|\mathcal{V}|} \mathrm{Sigmoid}(z_v^T z_e),$$





Wei, Tianxin, et al. "Augmentations in hypergraph contrastive learning: Fabricated and generative." *Advances in neural information processing systems* 35 (2022): 1909-1922.

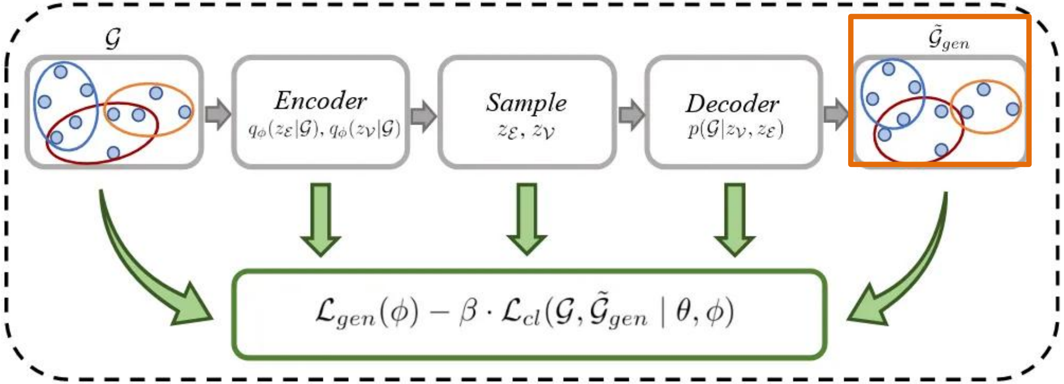
HyperGCL

☐ Variational Hypergraph Auto-Encoder (VHGAE)

- Apply the Gumbel-Softmax trick for the hyperedge distribution

$$
\begin{aligned}
T(\mathcal{G}) &= \text{Gumbel-Softmax}(p(\mathcal{G} \mid z_{\mathcal{V}}, z_{\mathcal{E}})) \\
&= \text{Sigmoid}((w_{\mathcal{V}\mathcal{E}} + \log(\delta) - \log(1-\delta))/\tau)
\end{aligned}
$$

$$
\tilde{\mathcal{G}}_{gen} = T(\mathcal{G}) \circ \mathcal{G},
$$

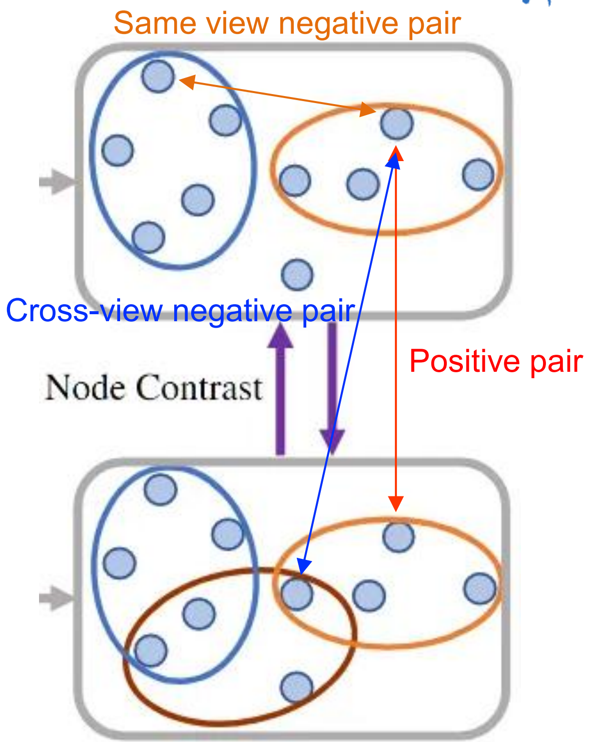


Wei, Tianxin, et al. "Augmentations in hypergraph contrastive learning: Fabricated and generative." *Advances in neural information processing systems* 35 (2022): 1909-1922.

HyperGCL

Objective Function

- Generator loss
  - $\mathcal{L}_{gen}(\phi) = -\mathrm{ELBO}$
- Contrastive loss
  - Node-level contrast
  - $L_{cl}(\tilde{G}_1, \tilde{G}_2) = \frac{1}{2|\mathcal{V}|}\sum_{n=1}^{|\mathcal{V}|}(l(u_n, s_n) + l(s_n, u_n))$
  - $l(u_n, s_n) = -\log \frac{\boxed{e^{\gamma(u_n, s_n)/\tau}}}{e^{\gamma(u_n, s_n)/\tau} + \boxed{\sum_{m \neq n} e^{\gamma(u_n, s_m)/\tau}} + \boxed{\sum_{m \neq n} e^{\gamma(u_n, u_m)/\tau}}}$



$$\rightarrow \min_{\phi} \mathcal{L}_{gen}(\phi) - \beta \cdot L_{cl}(\mathcal{G}, \tilde{\mathcal{G}}_{gen} | \theta, \phi)$$

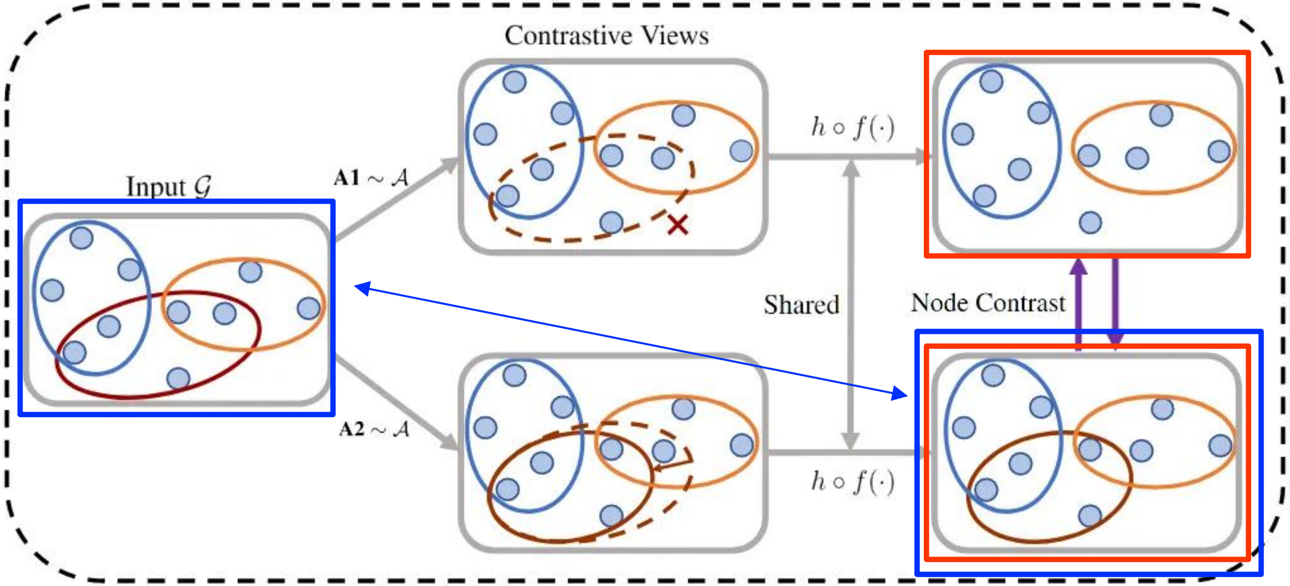
HyperGCL

☐ Training Pipeline

Algorithm 1 Hypergraph Contrastive Learning with Generative Augmentation (A6)

Input: Hypergraph $\mathcal{G}$; HyperGNN and generator parameters θ and ϕ; Multi-task training tradeoff parameters α, β

1: Randomly initialize θ and ϕ;
2: **while** not converge **do**
3: Obtain view $\tilde{\mathcal{G}}_p$ via fabricated augmentation and view $\tilde{\mathcal{G}}_{gen}$ via generator ϕ;
4: Define HyperGNN loss as: $\mathcal{L}_h = \mathcal{L}_{sup}(\theta) + \alpha \cdot \mathcal{L}_{\text{cl}}(\tilde{\mathcal{G}}_p, \tilde{\mathcal{G}}_{gen} \mid \theta, \phi)$;
5: Define generator loss as: $\mathcal{L}_g = \mathcal{L}_{gen}(\phi) - \beta \cdot \mathcal{L}_{\text{cl}}(\mathcal{G}, \mathcal{G}_{gen} \mid \theta, \phi)$;
6: Update HyperGNN θ to minimize $\mathcal{L}_h$;
7: Update generator ϕ to minimize $\mathcal{L}_g$;
8: **end while**



Wei, Tianxin, et al. "Augmentations in hypergraph contrastive learning: Fabricated and generative." *Advances in neural information processing systems* 35 (2022): 1909-1922.
https://openreview.net/attachment?id=igMc_C9pgYG&name=supplementary_material

Experiments

☐ Datasets

- Split the data into training/validation/test samples using (10%/10%/80%) splitting percentages

	Cora	Citeseer	Pubmed	Cora-CA	DBLP-CA	Zoo	20News	Mushroom	NTU2012	ModelNet40	Yelp	House	Walmart
$\lvert\mathcal{V}\rvert$	2708	3312	19717	2708	41302	101	16242	8124	2012	12311	50758	1290	88860
$\lvert\mathcal{E}\rvert$	1579	1079	7963	1072	22363	43	100	298	2012	12311	679302	341	69906
# feature	1433	3703	500	1433	1425	16	100	22	100	100	1862	100	100
# class	7	6	3	7	6	7	4	2	67	40	9	2	11
max $\lvert e\rvert$	5	26	171	43	202	93	2241	1808	5	5	2838	81	25
min $\lvert e\rvert$	2	2	2	2	2	1	29	1	5	5	2	1	2
avg $\lvert e\rvert$	3.03	3.2	4.35	4.28	4.45	39.93	654.51	136.31	5	5	6.66	34.72	6.59
med $\lvert e\rvert$	3	2	3	3	3	40	537	72	5	5	3	40	5
max d_v	145	88	99	23	18	17	44	5	19	30	7855	44	5733
min d_v	0	0	0	0	1	17	1	5	1	1	1	0	0
avg d_v	1.77	1.04	1.76	1.69	2.41	17	4.03	5	5	5	89.12	9.18	5.18
med d_v	1	0	0	2	2	17	3	5	5	4	35	7	2
h_e	0.86	0.83	0.88	0.88	0.93	0.66	0.73	0.96	0.87	0.92	0.57	0.58	0.75
h_v	0.84	0.78	0.79	0.79	0.88	0.35	0.49	0.87	0.81	0.88	0.26	0.52	0.55

Wei, Tianxin, et al. "Augmentations in hypergraph contrastive learning: Fabricated and generative." *Advances in neural information processing systems* 35 (2022): 1909-1922.

Experiments

☐ Comparison among different hypergraph augmentations

- Generalized hyperedge augmentation works the best among fabricated augmenting operators
- Generative augmentation usually performs the best in all the data sets

Table 3: Results on the test data sets: Mean accuracy (%) ± standard deviation. Bold values indicate the best result. Underlined values indicate the second best. 10% of all vertices are used for training.

	Cora	Citeseer	Pubmed	Cora-CA	DBLP-CA	Zoo	20Newsgroups	Mushroom
SetGNN	67.93 ± 1.27	63.53 ± 1.32	84.33 ± 0.36	72.21 ± 1.51	89.51 ± 0.18	65.06 ± 12.82	79.37 ± 0.35	99.75 ± 0.11
Self	68.24 ± 1.12	62.49 ± 1.48	84.38 ± 0.38	72.74 ± 1.53	89.51 ± 0.23	57.35 ± 18.32	79.45 ± 0.32	95.83 ± 0.23
Con	68.89 ± 1.80	62.82 ± 1.21	84.56 ± 0.34	73.22 ± 1.65	89.59 ± 0.13	61.05 ± 14.54	79.49 ± 0.45	95.85 ± 0.31
A0	68.59 ± 1.33	62.25 ± 2.15	84.54 ± 0.42	71.85 ± 1.62	89.62 ± 0.24	62.57 ± 13.84	79.07 ± 0.46	99.77 ± 0.17
A1	72.39 ± 1.34	66.28 ± 1.27	85.17 ± 0.37	75.45 ± 1.54	89.83 ± 0.21	65.80 ± 13.31	79.47 ± 0.32	99.80 ± 0.14
A2	72.58 ± 1.09	66.40 ± 1.35	85.16 ± 0.38	75.62 ± 1.42	90.22 ± 0.23	66.35 ± 13.26	79.56 ± 0.42	99.80 ± 0.17
A3	72.33 ± 1.23	65.79 ± 1.18	85.24 ± 0.28	75.34 ± 1.40	89.85 ± 0.16	65.79 ± 14.05	79.47 ± 0.34	99.81 ± 0.10
A4	72.95 ± 1.19	66.22 ± 0.95	84.88 ± 0.38	75.29 ± 1.56	90.10 ± 0.18	62.59 ± 12.77	79.45 ± 0.48	99.80 ± 0.14
A5	67.96 ± 0.99	63.21 ± 1.25	84.48 ± 0.40	72.61 ± 1.86	89.75 ± 0.24	62.47 ± 12.39	79.42 ± 0.52	99.79 ± 0.10
A6	**73.12 ± 1.48**	**66.94 ± 1.00**	**85.72 ± 0.38**	**76.21 ± 1.26**	**90.28 ± 0.19**	**66.89 ± 12.44**	**79.78 ± 0.40**	**99.86 ± 0.10**

	NTU2012	ModelNet40	Yelp	House (0.6)	House (1.0)	Walmart (0.6)	Walmart (1.0)	Avg. Rank
SetGNN	73.86 ± 1.62	95.85 ± 0.38	28.78 ± 1.51	68.54 ± 1.89	58.34 ± 2.25	74.97 ± 0.22	59.13 ± 0.20	7.71
Self	73.41 ± 1.65	95.83 ± 0.23	23.49 ± 4.15	67.75 ± 3.29	58.54 ± 2.16	74.76 ± 0.20	58.83 ± 0.21	8.64
Con	73.27 ± 1.53	95.85 ± 0.31	26.14 ± 1.86	68.50 ± 2.52	58.56 ± 2.42	75.17 ± 0.21	59.39 ± 0.20	7.07
A0	73.54 ± 1.93	95.92 ± 0.18	29.43 ± 1.42	67.48 ± 3.21	57.39 ± 2.37	73.14 ± 0.21	56.49 ± 0.60	8.21
A1	74.71 ± 1.81	95.87 ± 0.27	27.18 ± 0.71	68.64 ± 2.99	58.10 ± 3.22	75.42 ± 0.13	60.09 ± 0.25	4.50
A2	74.88 ± 1.66	96.56 ± 0.34	31.39 ± 2.45	69.73 ± 2.60	58.90 ± 1.97	75.50 ± 0.18	60.19 ± 0.20	2.29
A3	74.68 ± 1.74	96.48 ± 0.29	27.57 ± 1.00	67.88 ± 2.90	58.51 ± 2.22	75.29 ± 0.23	60.19 ± 0.20	4.71
A4	74.83 ± 1.75	95.86 ± 0.28	29.64 ± 1.93	69.56 ± 2.89	58.91 ± 2.69	75.43 ± 0.18	59.90 ± 0.24	4.14
A5	74.41 ± 1.86	96.46 ± 0.33	29.24 ± 1.42	68.14 ± 2.97	57.70 ± 2.98	75.26 ± 0.18	59.81 ± 0.22	6.71
A6	**75.34 ± 1.91**	**96.93 ± 0.33**	**34.64 ± 0.39**	**70.96 ± 2.27**	**59.93 ± 1.99**	**75.62 ± 0.16**	**60.46 ± 0.20**	**1.00**

Table 4: Results on the test data sets: Mean accuracy (%) ± standard deviation. Bold values indicate the best result. 1% of all vertexes are used for training.

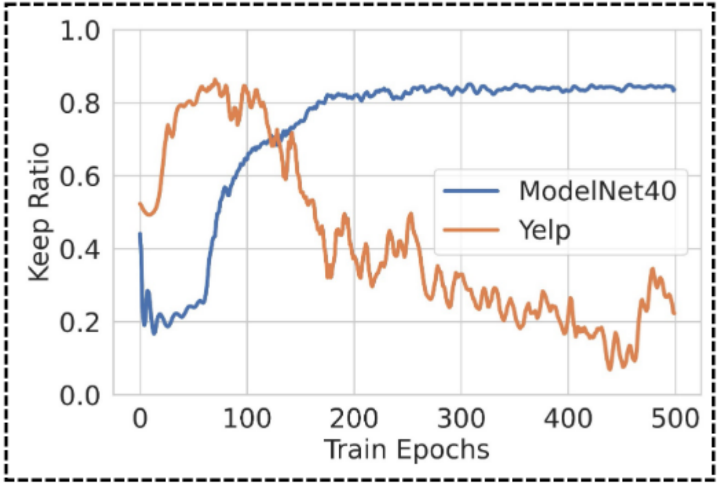
	Cora	Citeseer	Pubmed	Cora-CA	DBLP-CA	20Newsgroups	Mushroom
SetGNN	46.48 ± 3.62	47.01 ± 4.31	76.13 ± 1.19	52.29 ± 4.18	85.52 ± 0.54	73.83 ± 1.40	97.73 ± 1.18
Self	45.79 ± 5.34	44.22 ± 4.43	76.71 ± 0.90	51.64 ± 5.37	84.42 ± 0.37	73.91 ± 0.90	92.25 ± 0.89
Con	49.20 ± 4.38	48.56 ± 4.88	77.51 ± 1.08	52.37 ± 4.41	86.47 ± 0.35	74.39 ± 1.23	92.43 ± 0.87
A0	48.50 ± 4.77	46.43 ± 4.24	78.83 ± 1.79	49.87 ± 5.08	87.34 ± 0.73	74.43 ± 1.11	97.32 ± 1.33
A1	56.42 ± 5.02	55.63 ± 3.96	80.13 ± 1.44	60.86 ± 5.91	87.53 ± 0.30	74.68 ± 1.31	97.95 ± 1.15
A2	56.81 ± 4.49	56.10 ± 2.86	80.22 ± 1.24	60.96 ± 6.31	88.10 ± 0.35	74.72 ± 1.16	98.05 ± 1.18
A3	55.94 ± 3.67	55.82 ± 3.40	80.13 ± 1.02	60.51 ± 4.55	87.47 ± 0.36	74.63 ± 1.00	98.04 ± 0.98
A4	**58.55 ± 5.14**	**57.16 ± 4.62**	80.11 ± 1.02	60.91 ± 5.15	**88.91 ± 0.29**	74.67 ± 1.39	97.72 ± 1.12
A5	46.23 ± 3.44	45.07 ± 4.89	75.95 ± 1.32	53.26 ± 4.86	87.12 ± 0.43	74.81 ± 1.04	97.72 ± 1.25
A6	57.45 ± 5.00	56.23 ± 3.27	**81.10 ± 0.80**	**61.76 ± 4.94**	88.55 ± 0.41	**75.52 ± 0.93**	**98.28 ± 1.03**

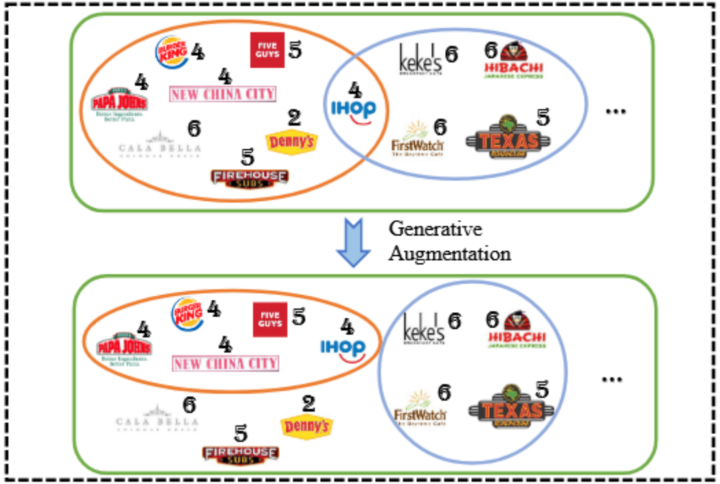
	ModelNet40	Yelp	House (0.6)	House (1.0)	Walmart (0.6)	Walmart (1.0)	Avg. Rank (↓)
SetGNN	88.34 ± 2.69	27.64 ± 1.10	53.69 ± 2.20	51.85 ± 1.64	65.48 ± 0.45	51.15 ± 0.52	7.62
Self	86.85 ± 3.03	20.77 ± 5.15	53.42 ± 2.25	51.14 ± 1.75	65.23 ± 0.43	51.00 ± 0.41	9.69
Con	87.00 ± 2.99	24.23 ± 0.43	53.58 ± 3.04	51.96 ± 1.87	65.47 ± 0.44	51.13 ± 0.46	7.31
A0	88.75 ± 2.78	27.43 ± 0.60	53.60 ± 2.73	51.70 ± 2.13	65.41 ± 0.47	51.10 ± 0.49	7.46
A1	89.34 ± 2.66	26.18 ± 0.51	54.12 ± 3.29	52.23 ± 2.46	65.96 ± 0.36	51.22 ± 0.35	4.08
A2	89.37 ± 2.69	27.67 ± 0.91	54.42 ± 2.83	52.31 ± 1.44	66.01 ± 0.41	51.32 ± 0.30	2.69
A3	89.31 ± 2.62	26.98 ± 0.66	53.71 ± 2.71	52.11 ± 2.24	65.88 ± 0.50	51.35 ± 0.53	4.38
A4	89.03 ± 2.66	27.45 ± 0.81	53.64 ± 2.61	51.77 ± 2.20	65.55 ± 0.51	51.04 ± 0.47	4.54
A5	89.43 ± 2.68	28.09 ± 0.96	54.07 ± 3.09	51.94 ± 1.84	65.52 ± 0.39	50.97 ± 0.47	6.00
A6	**90.22 ± 2.72**	**29.61 ± 0.71**	**56.27 ± 4.18**	**52.55 ± 2.18**	**66.42 ± 0.40**	**51.82 ± 0.39**	**1.23**

Wei, Tianxin, et al. "Augmentations in hypergraph contrastive learning: Fabricated and generative." *Advances in neural information processing systems* 35 (2022): 1909-1922.

Experiments

- **Training dynamics of keep ratio**
  - Highly related to the dataset homophily
    - The homophily of ModelNet40 : (0.92/0.88)
    - The homophily of Yelp : (0.57/0.26)





I’m Me, We’re Us, and I’m Us: Tri-directional Contrastive Learning on Hypergraphs

Dongjin Lee and Kijung Shin

School of Electrical Engineering, Kim Jaechul Graduate School of AI, KAIST, South Korea

23-AAAI

TriCL

☐ Motivation

Node-only Contrast

- Cannot capture higher-order relations
- Lead to limited expressiveness

Tri-directional Contrast

- Node + Group + Membership contrast
- Preserve higher-order structural information
- Produce richer, more generalizable embeddings

TriCL

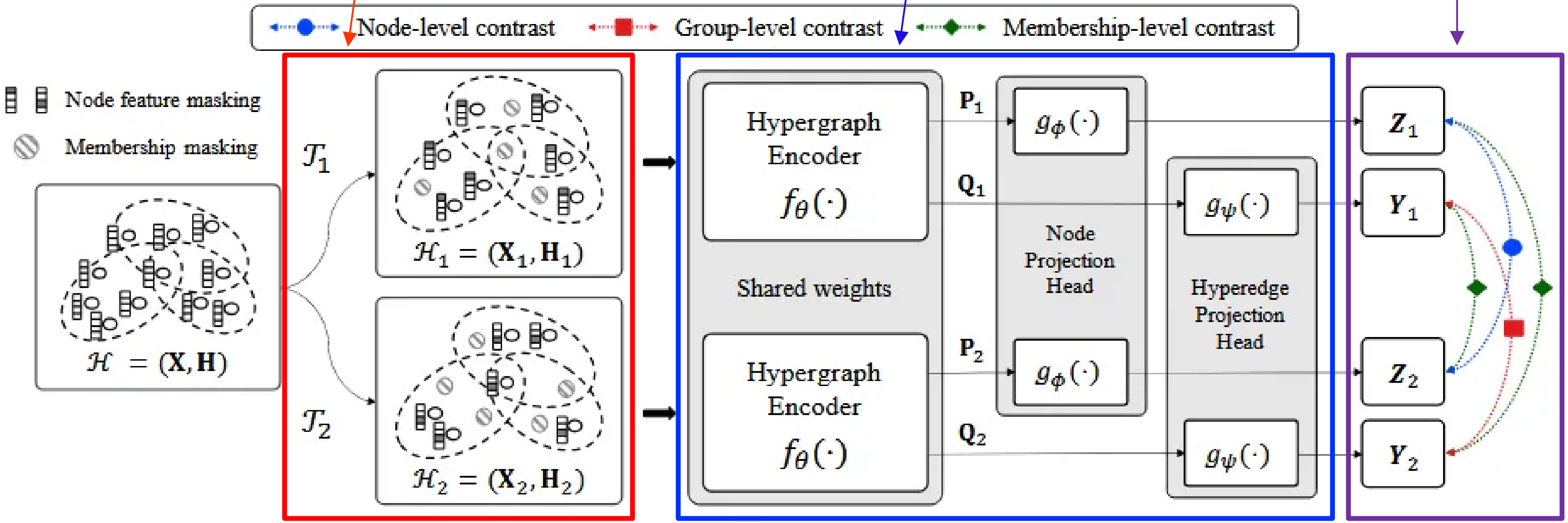
How to augment a hypergraph?	What to contrast?
Δ	O

☐ Overview

Generalized Hyperedge Perturbation + Attribute masking

Form node and hyperedge representations

Apply three forms of contrast



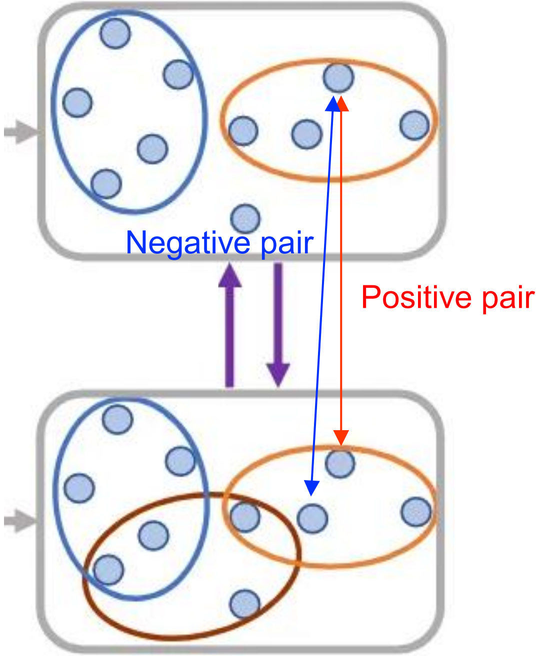
TriCL

☐ Node-level contrast

- Discriminate the representations of the same node in the two augmented views from other node representations

$$\ell_n(z_{1,i}, z_{2,i}) = -\log \frac{e^{s(z_{1,i}, z_{2,i})/\tau_n}}{\sum_{k=1}^{|V|} e^{s(z_{1,i}, z_{2,k})/\tau_n}},$$

$$\mathcal{L}_n = \frac{1}{2|V|} \sum_{i=1}^{|V|} \{\ell_n(z_{1,i}, z_{2,i}) + \ell_n(z_{2,i}, z_{1,i})\}.$$



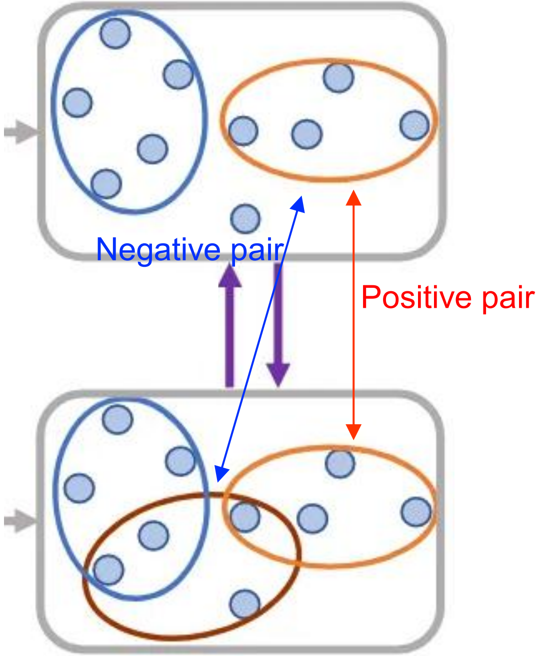
TriCL

Group-level contrast

- Distinguish the representations of the same hyperedge in the two augmented views from other hyperedge representations

$$\ell_g(\boldsymbol{y}_{1,j}, \boldsymbol{y}_{2,j}) = -\log \frac{e^{s(\boldsymbol{y}_{1,j}, \boldsymbol{y}_{2,j})/\tau_g}}{\sum_{k=1}^{|E|} e^{s(\boldsymbol{y}_{1,j}, \boldsymbol{y}_{2,k})/\tau_g}},$$

$$\mathcal{L}_g = \frac{1}{2|E|} \sum_{j=1}^{|E|} \left\{ \ell_g(\boldsymbol{y}_{1,j}, \boldsymbol{y}_{2,j}) + \ell_g(\boldsymbol{y}_{2,j}, \boldsymbol{y}_{1,j}) \right\}.$$

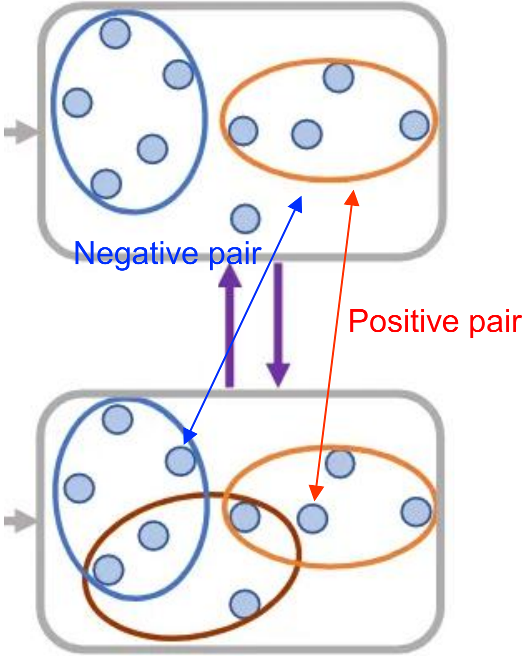


TriCL

- Membership-level contrast
  - Learn to represent if relationships actually exist and to move away if they don't exist

$$\ell_m(\boldsymbol{z}_i, \boldsymbol{y}_j) = -\log \underbrace{\frac{e^{\mathcal{D}(\boldsymbol{z}_i,\boldsymbol{y}_j)/\tau_m}}{e^{\mathcal{D}(\boldsymbol{z}_i,\boldsymbol{y}_j)/\tau_m} + \sum_{k:i\notin k} e^{\mathcal{D}(\boldsymbol{z}_i,\boldsymbol{y}_k)/\tau_m}}}_{\text{when } \boldsymbol{z}_i \text{ is the anchor}} - \log \underbrace{\frac{e^{\mathcal{D}(\boldsymbol{z}_i,\boldsymbol{y}_j)/\tau_m}}{e^{\mathcal{D}(\boldsymbol{z}_i,\boldsymbol{y}_j)/\tau_m} + \sum_{k:k\notin j} e^{\mathcal{D}(\boldsymbol{z}_k,\boldsymbol{y}_j)/\tau_m}}}_{\text{when } \boldsymbol{y}_j \text{ is the anchor}},$$

$$\mathcal{L}_m = \frac{1}{2K} \sum_{i=1}^{|V|} \sum_{j=1}^{|E|} \mathbb{1}_{[h_{ij}=1]} \{\ell_m(\boldsymbol{z}_{1,i}, \boldsymbol{y}_{2,j}) + \ell_m(\boldsymbol{z}_{2,i}, \boldsymbol{y}_{1,j})\}.$$



Experiments

☐ Datasets

Table 6: Statistics of datasets used in our experiments.

	Cora-C	Citeseer	Pubmed	Cora-A	DBLP	Zoo	20News	Mushroom	NTU2012	ModelNet40
# Nodes	1,434	1,458	3,840	2,388	41,302	101	16,242	8,124	2,012	12,311
# Hyperedges	1,579	1,079	7,963	1,072	22,363	43	100	298	2,012	12,311
# Memberships	4,786	3,453	34,629	4,585	99,561	1,717	65,451	40,620	10,060	61,555
Avg. hyperedge size	3.03	3.20	4.35	4.28	4.45	39.93	654.51	136.31	5	5
Avg. node degree	3.34	2.37	9.02	1.92	2.41	17.00	4.03	5.00	5	5
Max. hyperedge size	5	26	171	43	202	93	2241	1808	5	5
Max. node degree	145	88	99	23	18	17	44	5	19	30
# Features	1,433	3,703	500	1,433	1,425	16	100	22	100	100
# Classes	7	6	3	7	6	7	4	2	67	40

Experiments

Performance on Node Classification

- Graph contrastive learning methods show significantly lower accuracy compared to TriCL

Table 1: Node classification accuracy and standard deviations. Graph methods, marked as ⋆, are applied after converting hypergraphs to graphs via clique expansion. For each dataset, the best and the second-best performances are highlighted in **boldface** and <u>underlined</u>, respectively. A.R. denotes average rank, OOT denotes cases where results are not obtained within 24 hours, and OOM indicates out of memory on a 24GB GPU. In most cases, TriCL outperforms all others, including the supervised ones.

	Method	Cora-C	Citeseer	Pubmed	Cora-A	DBLP	Zoo	20News	Mushroom	NTU2012	ModelNet40	A.R.↓
Supervised	MLP	60.32 ± 1.5	62.06 ± 2.3	76.27 ± 1.1	64.05 ± 1.4	81.18 ± 0.2	75.62 ± 9.5	79.19 ± 0.5	99.58 ± 0.3	65.17 ± 2.3	93.75 ± 0.6	12.5
	GCN⋆	77.11 ± 1.8	66.07 ± 2.4	82.63 ± 0.6	73.66 ± 1.3	87.58 ± 0.2	36.79 ± 9.6	OOM	92.47 ± 0.9	71.17 ± 2.4	91.67 ± 0.2	11.7
	GAT⋆	77.75 ± 2.1	67.62 ± 2.5	81.96 ± 0.7	74.52 ± 1.3	88.59 ± 0.1	36.48 ± 10.0	OOM	OOM	70.94 ± 2.6	91.43 ± 0.3	11
	HGNN	77.50 ± 1.8	66.16 ± 2.3	83.52 ± 0.7	74.38 ± 1.2	88.32 ± 0.3	78.58 ± 11.1	80.15 ± 0.3	98.59 ± 0.5	72.03 ± 2.4	92.23 ± 0.2	8.1
	HyperConv	76.19 ± 2.1	64.12 ± 2.6	83.42 ± 0.6	73.52 ± 1.0	88.83 ± 0.2	62.53 ± 14.5	79.83 ± 0.4	97.56 ± 0.6	72.62 ± 2.6	91.84 ± 0.1	9.8
	HNHN	76.21 ± 1.7	67.28 ± 2.2	80.97 ± 0.9	74.88 ± 1.6	86.71 ± 1.2	78.89 ± 10.2	79.51 ± 0.4	99.78 ± 0.1	71.45 ± 3.2	92.96 ± 0.2	8.9
	HyperGCN	64.11 ± 7.4	59.92 ± 9.6	78.40 ± 9.2	60.65 ± 9.2	76.59 ± 7.6	40.86 ± 2.1	77.31 ± 6.0	48.26 ± 0.3	46.05 ± 3.9	69.23 ± 2.8	15.1
	HyperSAGE	64.98 ± 5.3	52.43 ± 9.4	79.49 ± 8.7	64.59 ± 4.3	79.63 ± 8.6	40.86 ± 2.1	OOT	OOT	OOT	OOT	14.7
	UniGCN	77.91 ± 1.9	66.40 ± 1.9	<u>84.08 ± 0.7</u>	77.30 ± 1.4	90.31 ± 0.2	72.10 ± 12.1	**80.24 ± 0.4**	98.84 ± 0.5	73.27 ± 2.7	94.62 ± 0.2	5.9
	AllSet	76.21 ± 1.7	67.83 ± 1.8	82.85 ± 0.9	76.94 ± 1.3	90.07 ± 0.3	72.72 ± 11.8	79.90 ± 0.4	99.78 ± 0.1	75.09 ± 2.5	96.85 ± 0.2	6.2
Unsupervised	Node2vec⋆	70.99 ± 1.4	53.85 ± 1.9	78.75 ± 0.9	58.50 ± 2.1	72.09 ± 0.3	17.02 ± 4.1	63.35 ± 1.7	88.16 ± 0.8	67.72 ± 2.1	84.94 ± 0.4	15.6
	DGI⋆	78.17 ± 1.4	68.81 ± 1.8	80.83 ± 0.6	76.94 ± 1.1	88.00 ± 0.2	36.54 ± 9.7	OOM	OOM	72.01 ± 2.5	92.18 ± 0.2	9.3
	GRACE⋆	79.11 ± 1.7	68.65 ± 1.7	80.08 ± 0.7	76.59 ± 1.0	OOM	37.07 ± 9.3	OOM	OOM	70.51 ± 2.4	90.68 ± 0.3	10.4
	S^2-HHGR	78.08 ± 1.7	68.21 ± 1.8	82.13 ± 0.6	78.15 ± 1.1	88.69 ± 0.2	80.06 ± 11.1	79.75 ± 0.3	97.15 ± 0.5	73.95 ± 2.4	93.26 ± 0.2	6.8
	Random-Init	63.62 ± 3.1	60.44 ± 2.5	67.49 ± 2.2	66.27 ± 2.2	76.57 ± 0.6	78.43 ± 11.0	77.14 ± 0.6	97.40 ± 0.6	74.39 ± 2.6	96.29 ± 0.3	11.9
	TriCL-N	80.23 ± 1.2	70.28 ± 1.5	83.44 ± 0.6	81.94 ± 1.1	90.88 ± 0.1	79.94 ± 11.1	<u>80.18 ± 0.2</u>	99.76 ± 0.2	75.20 ± 2.6	97.01 ± 0.2	3.4
	TriCL-NG	<u>81.45 ± 1.2</u>	<u>71.38 ± 1.2</u>	<u>83.68 ± 0.7</u>	<u>82.00 ± 1.0</u>	<u>90.94 ± 0.1</u>	<u>80.19 ± 11.1</u>	<u>80.18 ± 0.2</u>	<u>99.81 ± 0.1</u>	**75.25 ± 2.5**	<u>97.02 ± 0.1</u>	<u>2</u>
	TriCL	**81.57 ± 1.1**	**72.02 ± 1.2**	**84.26 ± 0.6**	**82.15 ± 0.9**	**91.12 ± 0.1**	**80.25 ± 11.2**	80.14 ± 0.2	**99.83 ± 0.1**	<u>75.23 ± 2.4</u>	**97.08 ± 0.1**	**1.5**

Experiments

☐ Performance on Node Classification

- Considering the different types of contrast together can help improve performance

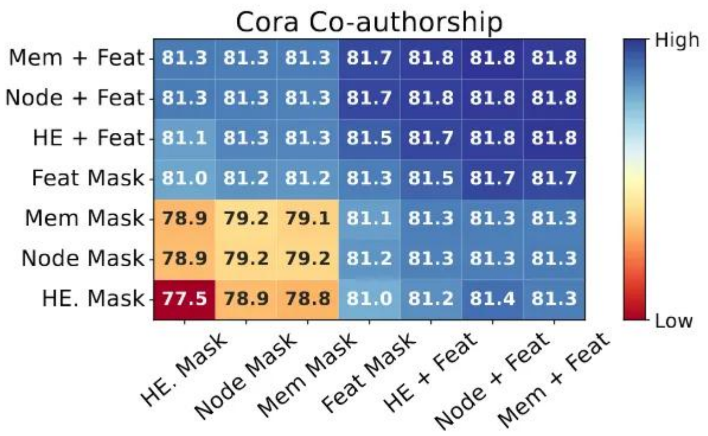
Table 2: Comparison of node classification accuracy according to whether or not to use each type of contrast (i.e., $\mathcal{L}_n$, $\mathcal{L}_g$, and $\mathcal{L}_m$). Using all types of contrasts (i.e., node-, group-, and membership-level contrast) achieves the best performance in most cases as they are complementarily reinforcing each other.

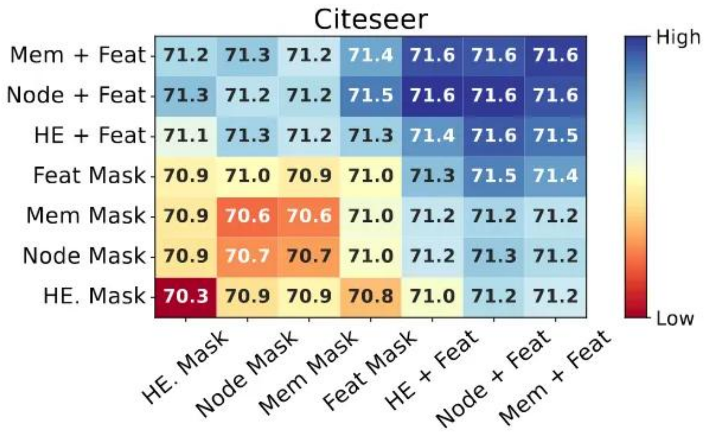
$\mathcal{L}_n$	$\mathcal{L}_g$	$\mathcal{L}_m$	Cora-C	Citeseer	Pubmed	Cora-A	DBLP	Zoo	20News	Mushroom	NTU2012	ModelNet40	A.R.↓
✓	-	-	80.23 ± 1.2	70.28 ± 1.5*	83.44 ± 0.6	81.94 ± 1.1	90.88 ± 0.1	79.94 ± 11.1	**80.18 ± 0.2**	99.76 ± 0.2	75.20 ± 2.6	97.01 ± 0.2	3.8
-	✓	-	79.69 ± 1.6	71.02 ± 1.3*	80.20 ± 1.3	78.98 ± 1.4	88.60 ± 0.2	79.31 ± 10.7	79.35 ± 0.4	99.13 ± 0.3	74.41 ± 2.6	96.66 ± 0.2	5.7
-	-	✓	76.76 ± 1.8	63.98 ± 2.0	79.86 ± 0.9	76.77 ± 1.1	63.95 ± 7.2	79.80 ± 11.0	79.27 ± 0.3	94.87 ± 0.7	73.11 ± 2.8	96.57 ± 0.2	6.9
✓	✓	-	81.45 ± 1.2	71.38 ± 1.4	83.68 ± 0.7	82.00 ± 1.0	90.94 ± 0.1	80.19 ± 11.1	**80.18 ± 0.2**	99.81 ± 0.1	**75.25 ± 2.5**	97.02 ± 0.1	2.3
✓	-	✓	80.49 ± 1.3	70.46 ± 1.5	83.98 ± 0.7	81.62 ± 1.0	90.75 ± 0.1	80.19 ± 11.1	80.15 ± 0.2	99.74 ± 0.2	75.12 ± 2.5	97.03 ± 0.1	3.6
-	✓	✓	80.80 ± 1.1	71.73 ± 1.4	82.81 ± 0.7	80.24 ± 1.0	90.17 ± 0.1	80.20 ± 11.1	79.29 ± 0.2	99.82 ± 0.1	73.76 ± 2.5	96.74 ± 0.1	4.1
✓	✓	✓	**81.57 ± 1.1**	**72.02 ± 1.4**	**84.26 ± 0.6**	**82.15 ± 0.9**	**91.12 ± 0.1**	**80.25 ± 11.2**	80.14 ± 0.2	**99.83 ± 0.1**	75.23 ± 2.4	**97.08 ± 0.1**	**1.4**

Experiments

☐ Hypergraph augmentation

- Using the structural and attribute augmentations together always yields better performance than using just one





Lee, Dongjin, and Kijung Shin. "I'm me, we're us, and i'm us: Tri-directional contrastive learning on hypergraphs." *Proceedings of the AAAI conference on artificial intelligence*. Vol. 37. No. 7. 2023.

Enhancing Hyperedge Prediction with Context-Aware Self-Supervised Learning

Yunyong Ko, Hanghang Tong, Sang-Wook Kim

Chung-Ang University (CAU), University of Illinois at Urbana-Champaign (UIUC), Hanyang University

25-TKDE

CASH

How to augment a hypergraph?	What to contrast?
O	Δ

Context-aware hyperedge prediction

Form node representations using original hypergraph

☐ Overview

Generate representation for hyperedge candidate and predict

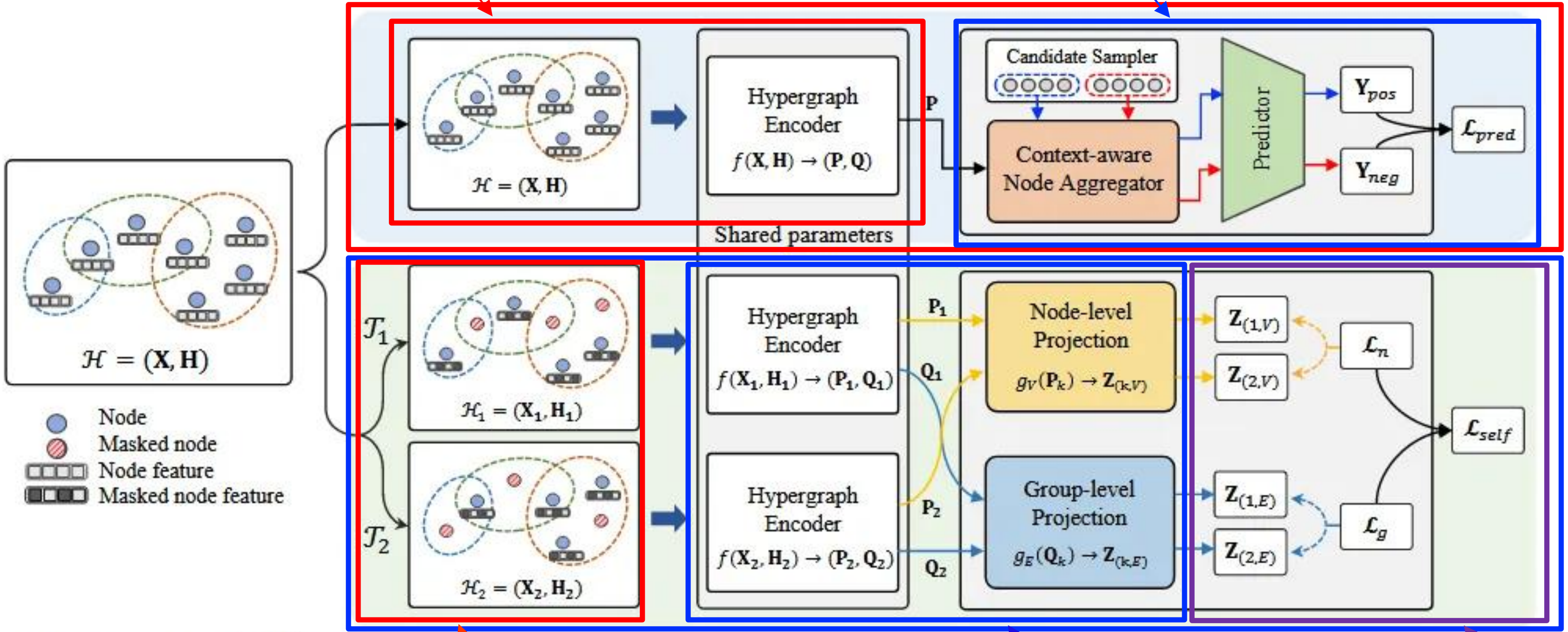


Fig. 2. The overview of CASH: (1) Context-aware hyperedge prediction (upper) and (2) Self-supervised contrative hypergraph learning (lower).

Hyperedge-aware augmentation

Form node and hyperedge representations

Apply dual contrast

Self-supervised contrastive hypergraph learning

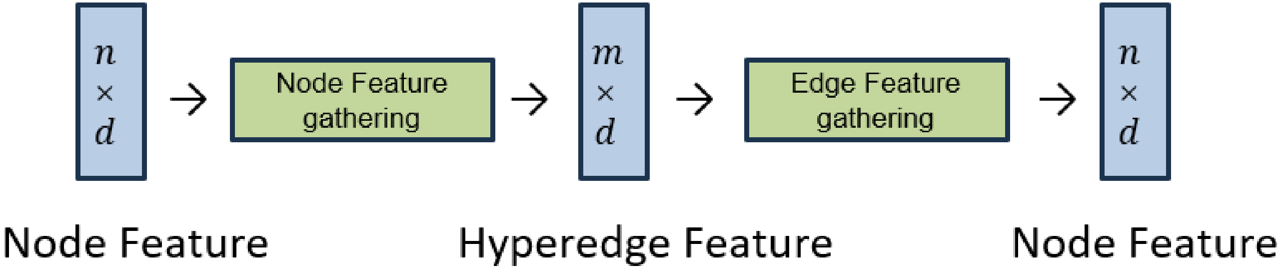
Ko, Yunyong, Hanghang Tong, and Sang-Wook Kim. "Enhancing hyperedge prediction with context-aware self-supervised learning." *IEEE Transactions on Knowledge and Data Engineering* (2025).

CASH

☐ Hypergraph encoding

- Produce the node and hyperedge embeddings via a 2-stage aggregation strategy

$$Q^{(k)} = \sigma\left(D_E^{-1} H^T P^{(k-1)} W_E^{(k)} + b_E^{(k)}\right)$$
$$P^{(k)} = \sigma\left(D_V^{-1} H Q^{(k)} W_V^{(k)} + b_V^{(k)}\right)$$

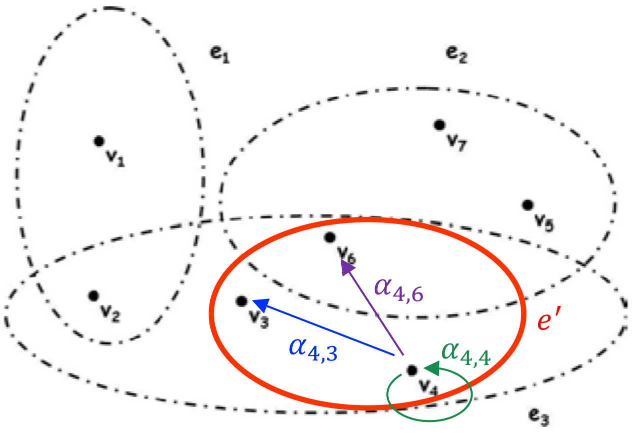


CASH

Context-Aware Hyperedge Prediction

- Hyperedge candidate scoring
  - Calculate the relative degrees of influences of nodes by using the attention mechanism

$$\alpha_{i,j} = \frac{exp(p_{v'_i}\mathbf{W}''_{agg} \cdot x^\top)}{\sum_{v'_j \in e'} exp(p_{v'_j}\mathbf{W}''_{agg} \cdot x^\top)}$$

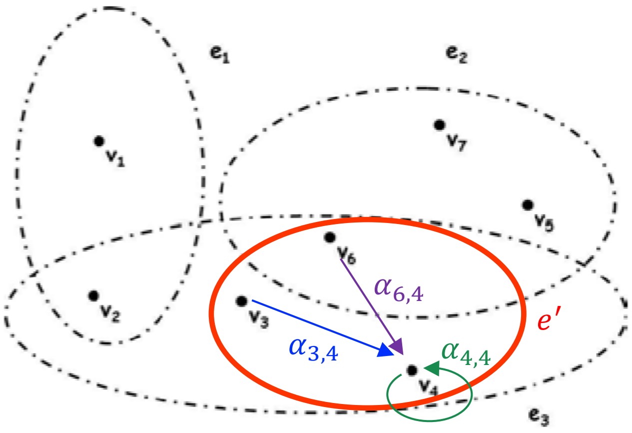


Zhou, Dengyong, Jiayuan Huang, and Bernhard Schölkopf. "Learning with hypergraphs: Clustering, classification, and embedding." *Advances in neural information processing systems* 19 (2006).

Context-Aware Hyperedge Prediction

- Hyperedge candidate scoring
  - Update each node embedding based on the relative degrees of influences

$$p^*_{v'_j} = \sum_{v'_i \in e'} \alpha_{i,j} \cdot p_{v'_i} \mathbf{W}'_{agg}$$



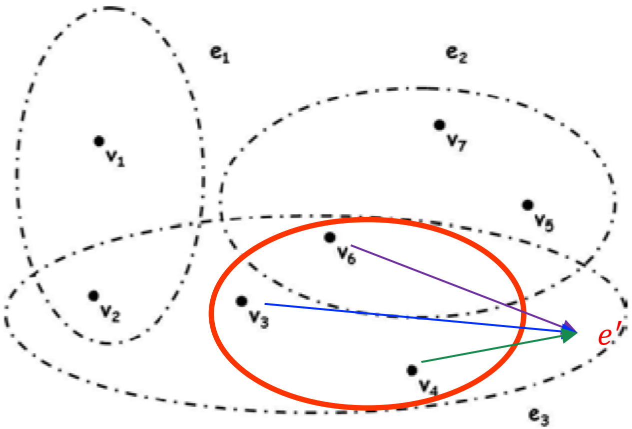
Zhou, Dengyong, Jiayuan Huang, and Bernhard Schölkopf. "Learning with hypergraphs: Clustering, classification, and embedding." *Advances in neural information processing systems* 19 (2006).

Context-Aware Hyperedge Prediction

Hyperedge candidate scoring

- Aggregate the influence-reflected embeddings of the nodes via *element-wise max pooling*
- $pred(\cdot)$: a hyperedge predictor (a fully-connected layer $(d \times 1)$)

$$\hat{y}_{e'} = pred(q^*_{e'}), \qquad q^*_{e'} = MaxPool(P^*[e', :])$$



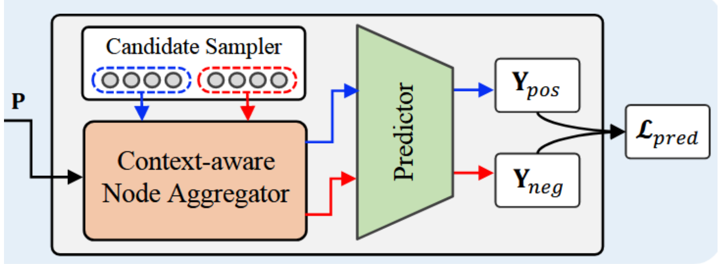
Zhou, Dengyong, Jiayuan Huang, and Bernhard Schölkopf. "Learning with hypergraphs: Clustering, classification, and embedding." *Advances in neural information processing systems* 19 (2006).

CASH

☐ Context-Aware Hyperedge Prediction

■ Model training

- ☐ Use heuristic negative sampling methods
  - Sized NS (SNS) : sampling k random nodes (easy)
  - Motif NS (MNS) : sampling a k-connected component in a clique-expanded hypergraph (difficult)
  - Clique NS (CNS) : selecting a hyperedge e and replacing one of its incident nodes $u \in e$ with a node $v \notin e$ (most difficult)



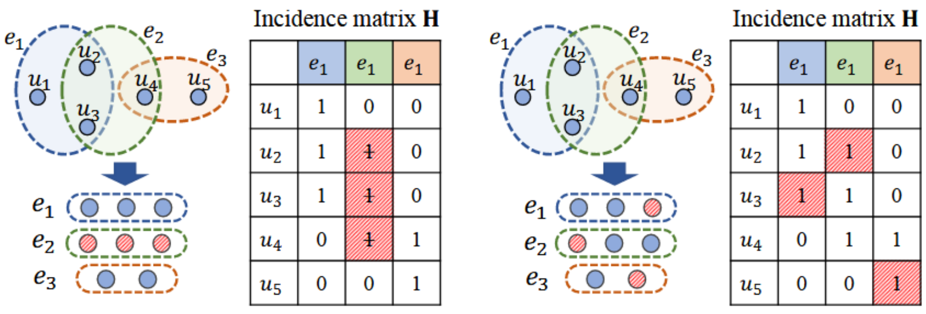
- ☐ Prediction loss

$$\mathcal{L}_{pred} = -\frac{1}{|E'|} \sum_{e' \in E'} \underbrace{y_{e'} \cdot \log \hat{y}_{e'}}_{\text{positives}} + \underbrace{(1 - y_{e'}) \cdot \log (1 - \hat{y}_{e'})}_{\text{negatives}}$$

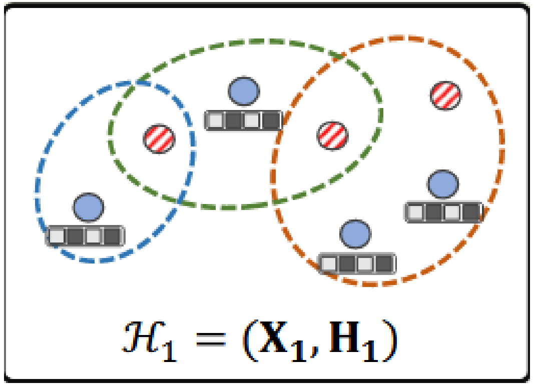
CASH

Self-Supervised Contrastive Hypergraph Learning

- **Hypergraph augmentation**
  - Mask random p_m% members of each hyperedge *individually*
  - Mask random p_f% dimensions of node features



(a) Random masking (b) Hyperedge-aware masking



CASH

Self-Supervised Contrastive Hypergraph Learning

- Contrastive loss

$$\mathcal{L}_{con} = -\underbrace{\log sim(\mathbf{Z}_{(1,V)}, \mathbf{Z}_{(2,V)})}_{\text{node-level contrast}} - \underbrace{\log sim(\mathbf{Z}_{(1,E)}, \mathbf{Z}_{(2,E)})}_{\text{group-level contrast}},$$

- Unified loss of **CASH**

$$\mathcal{L} = \mathcal{L}_{pred} + \beta L_{con}$$

Experiments

☐ Datasets

STATISTICS OF HYPERGRAPH DATASETS

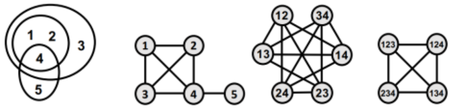
Dataset	$\|V\|$	$\|E\|$	# Features	Type
Citeseer	1,457	1,078	3,703	Co-citation
Cora	1,434	1,579	1,433	Co-citation
Pubmed	3,840	7,962	500	Co-citation
Cora-A	2,388	1,072	1,433	Authorship
DBLP-A	39,283	16,483	4,543	Authorship
DBLP	15,639	22,964	4,543	Collaboration

Experiments

☐ Hyperedge prediction accuracy

- Consistently outperform all competitors

Dataset	Metric / Test set	AUROC SNS	AUROC MNS	AUROC CNS	AUROC MIX	AUROC Average	AP SNS	AP MNS	AP CNS	AP MIX	AP Average
Citeseer	Expansion	0.663	0.781	0.331	0.588	0.591 ± 0.011	0.765	0.817	0.498	0.630	0.681 ± 0.001
	HyperSAGNN	0.540	0.410	0.473	0.478	0.475 ± 0.019	0.627	0.455	0.497	0.507	0.512 ± 0.015
	NHP	**0.991**	0.701	0.510	0.817	0.751 ± 0.009	**0.990**	0.731	0.520	0.768	0.751 ± 0.011
	AHP	0.943	0.881	0.651	0.820	0.824 ± 0.020	0.952	0.870	0.660	0.795	0.819 ± 0.022
	CASH	0.925	**0.921**	**0.720**	**0.857**	**0.856 ± 0.011**	0.928	**0.919**	**0.701**	**0.831**	**0.845 ± 0.009**
	Improvement (%)	-6.65%	+4.54%	+10.60%	+4.51%	+3.88%	-6.26%	5.63%	+6.21%	+4.53%	+3.17%
Cora	Expansion	0.470	0.707	0.256	0.476	0.477 ± 0.009	0.637	0.764	0.454	0.563	0.607 ± 0.009
	HyperSAGNN	0.617	0.527	0.494	0.540	0.545 ± 0.021	0.687	0.574	0.508	0.566	0.584 ± 0.019
	NHP	0.943	0.641	0.472	0.774	0.703 ± 0.015	0.949	0.678	0.509	0.744	0.718 ± 0.020
	AHP	**0.964**	0.860	0.572	0.799	0.799 ± 0.019	**0.961**	0.837	0.552	0.740	0.772 ± 0.035
	CASH	0.923	**0.867**	**0.671**	**0.824**	**0.822 ± 0.011**	0.915	**0.854**	**0.644**	**0.789**	**0.801 ± 0.016**
	Improvement (%)	-4.25%	+0.81%	+17.31%	+3.13%	+2.88%	-4.79%	+2.03%	+16.67%	+6.62%	+3.76%
Pubmed	Expansion	0.520	0.730	0.241	0.497	0.497 ± 0.015	0.675	0.755	0.440	0.565	0.612 ± 0.010
	HyperSAGNN	0.525	0.686	0.546	0.580	0.584 ± 0.066	0.534	0.680	0.529	0.561	0.576 ± 0.050
	NHP	**0.973**	0.694	0.524	0.745	0.733 ± 0.004	**0.973**	0.656	0.513	0.678	0.707 ± 0.004
	AHP	0.917	0.840	0.553	0.763	0.763 ± 0.009	0.918	0.834	0.526	0.717	0.749 ± 0.007
	CASH	0.805	**0.871**	**0.640**	**0.772**	**0.772 ± 0.009**	0.810	**0.880**	**0.644**	**0.765**	**0.775 ± 0.008**
	Improvement (%)	-17.26%	+3.69%	15.73%	+1.18%	+1.18%	-16.75%	+5.52%	+21.74%	+6.69%	+3.47%
Cora-A	Expansion	0.690	0.842	0.434	0.658	0.656 ± 0.011	0.690	0.876	0.577	0.672	0.706 ± 0.020
	HyperSAGNN	0.386	0.591	0.542	0.505	0.506 ± 0.019	0.532	0.643	0.545	0.563	0.571 ± 0.009
	NHP	0.909	0.672	0.550	0.773	0.723 ± 0.015	0.925	0.720	0.585	0.766	0.748 ± 0.019
	AHP	0.958	0.924	0.782	0.887	0.888 ± 0.014	0.957	0.898	0.796	0.878	0.882 ± 0.014
	CASH	**0.971**	**0.975**	**0.833**	**0.931**	**0.927 ± 0.011**	**0.969**	**0.973**	**0.832**	**0.926**	**0.925 ± 0.011**
	Improvement (%)	+1.36%	+5.52%	+6.52%	+4.96%	+4.39%	+1.25%	+8.35%	+4.52%	+5.47%	+4.88%
DBLP-A	Expansion	0.634	0.826	0.350	0.603	0.603 ± 0.006	0.730	0.852	0.512	0.641	0.687 ± 0.004
	HyperSAGNN	0.548	0.791	0.563	0.636	0.634 ± 0.007	0.686	0.805	0.552	0.655	0.675 ± 0.004
	NHP	**0.966**	0.623	0.555	0.721	0.716 ± 0.005	**0.965**	0.604	0.534	0.663	0.693 ± 0.007
	AHP	0.916	0.926	0.668	0.838	0.837 ± 0.004	0.928	0.928	0.707	0.836	0.850 ± 0.003
	CASH	0.929	**0.957**	**0.747**	**0.877**	**0.877 ± 0.003**	0.933	**0.955**	**0.741**	**0.863**	**0.873 ± 0.005**
	Improvement (%)	-3.83%	+3.35%	+11.83%	+4.65%	+4.78%	-3.32%	+2.91%	+4.81%	+3.23%	+2.71%
DBLP	Expansion	0.645	0.801	0.366	0.607	0.607 ± 0.005	0.751	**0.856**	0.518	0.655	0.698 ± 0.004
	HyperSAGNN	0.448	0.574	0.572	0.530	0.531 ± 0.018	0.562	0.602	0.586	0.577	0.582 ± 0.016
	NHP	0.663	0.540	0.503	0.572	0.569 ± 0.003	0.608	0.523	0.501	0.542	0.544 ± 0.002
	AHP	**0.946**	0.820	0.568	0.778	0.778 ± 0.002	**0.947**	0.815	0.561	0.735	0.764 ± 0.007
	CASH	0.875	**0.836**	**0.708**	**0.807**	**0.807 ± 0.015**	0.874	0.832	**0.696**	**0.793**	**0.799 ± 0.011**
	Improvement (%)	-7.50%	+1.95%	+23.78%	+3.73%	+3.73%	-7.70%	-2.80%	+18.77%	+7.89%	+4.58%



(a) Hypergraph (b) 2-pg (c) 3-pg (d) 4-pg

Figure 1: Hypergraph and its n-projected graphs (n-pgs). Edge weights, dropped for simplicity, reflect n-way interactions. For example, edge weight between the 3-pg nodes {1, 2} and {2, 3} is the number of times {1, 2, 3} interacted as a group.

Ko, Yunyong, Hanghang Tong, and Sang-Wook Kim. "Enhancing hyperedge prediction with context-aware self-supervised learning." *IEEE Transactions on Knowledge and Data Engineering* (2025).

Yoon, Se-eun, et al. "How much and when do we need higher-order information in hypergraphs? A case study on hyperedge prediction." *Proceedings of the web conference 2020*. 2020.

Experiments

☐ Ablation study

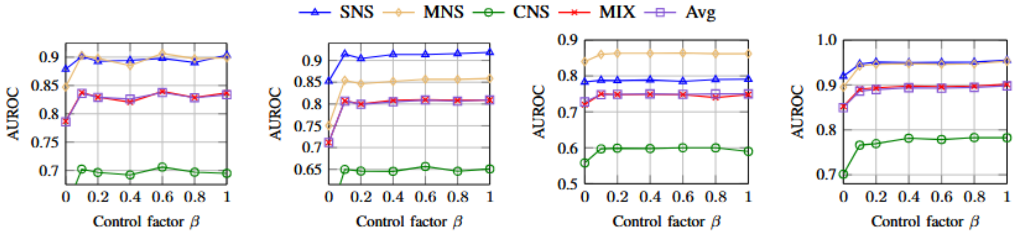
- Each of strategies is always beneficial

Dataset	Metric / Test set	AUROC SNS	AUROC MNS	AUROC CNS	AUROC MIX	AUROC Average	AP SNS	AP MNS	AP CNS	AP MIX	AP Average
Citeseer	CASH-No	0.878	0.847	0.630	0.786	0.786 ± 0.003	0.890	0.841	0.653	0.775	0.790 ± 0.007
	CASH-CL	0.907	0.890	0.679	0.832	0.827 ± 0.019	0.905	0.874	0.675	0.815	0.817 ± 0.013
	CASH-HCL	0.908	0.897	0.691	0.839	0.833 ± 0.013	0.909	0.880	0.691	0.824	0.826 ± 0.005
	CASH-ALL	**0.925**	**0.921**	**0.720**	**0.857**	**0.856 ± 0.011**	**0.928**	**0.919**	**0.701**	**0.831**	**0.845 ± 0.009**
	Improvement (%)	+5.35%	+8.74%	+14.29%	+9.03%	+8.91%	+4.27%	+9.27%	+7.35%	+7.23%	+6.96%
Cora	CASH-No	0.852	0.750	0.532	0.711	0.712 ± 0.019	0.856	0.759	0.531	0.684	0.707 ± 0.021
	CASH-CL	0.895	0.837	0.600	0.782	0.779 ± 0.015	0.873	0.809	0.566	0.727	0.744 ± 0.019
	CASH-HCL	0.893	0.835	0.600	0.780	0.777 ± 0.017	0.879	0.816	0.565	0.730	0.747 ± 0.018
	CASH-ALL	**0.923**	**0.867**	**0.671**	**0.824**	**0.822 ± 0.011**	**0.915**	**0.854**	**0.644**	**0.789**	**0.801 ± 0.016**
	Improvement (%)	+8.33%	+15.60%	+26.13%	+15.89%	+15.45%	+6.89%	+12.52%	+21.28%	+15.35%	+13.30%
Pubmed	CASH-No	0.782	0.844	0.558	0.727	0.728 ± 0.007	0.802	0.852	0.555	0.708	0.730 ± 0.007
	CASH-CL	0.806	0.845	0.562	0.735	0.737 ± 0.010	0.817	0.847	0.552	0.708	0.731 ± 0.007
	CASH-HCL	**0.814**	0.848	0.562	0.739	0.741 ± 0.008	**0.823**	0.851	0.547	0.708	0.732 ± 0.006
	CASH-ALL	0.805	**0.871**	**0.640**	**0.772**	**0.772 ± 0.009**	0.810	**0.880**	**0.644**	**0.765**	**0.775 ± 0.008**
	Improvement (%)	+2.94%	+3.20%	+14.70%	+6.19%	+6.04%	+1.00%	+3.29%	+16.04%	+8.05%	+6.16%
Cora-A	CASH-No	0.949	0.894	0.701	0.852	0.849 ± 0.020	0.951	0.906	0.738	0.857	0.863 ± 0.017
	CASH-CL	0.943	0.934	0.756	0.883	0.879 ± 0.030	0.944	0.936	0.772	0.881	0.884 ± 0.026
	CASH-HCL	**0.972**	0.949	0.833	0.921	0.919 ± 0.008	**0.972**	0.919	**0.845**	0.908	0.911 ± 0.007
	CASH-ALL	0.971	**0.975**	**0.833**	**0.931**	**0.927 ± 0.011**	0.969	**0.973**	0.832	**0.926**	**0.925 ± 0.011**
	Improvement (%)	+2.32%	+9.06%	+18.83%	+9.27%	+9.19%	+1.89%	+7.40%	+12.74%	+8.05%	+7.18%

Experiments

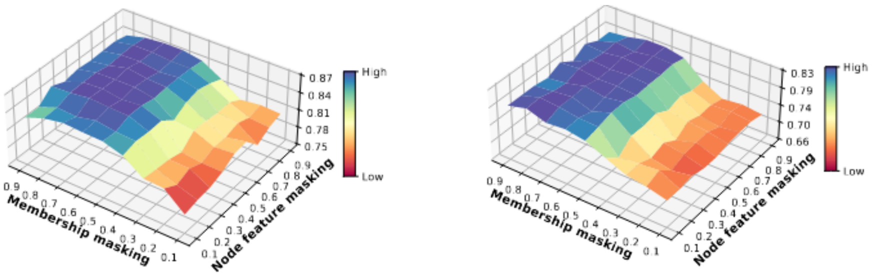
☐ Hyperparameter sensitivity

- ■ Control parameter β
  - ☐ Insensitive to its hyperparameter β



(a) Citeseer (b) Cora (c) Pubmed (d) Cora-A

- ■ Augmentation parameter p_m, p_f
  - ☐ p_m is more important than p_f

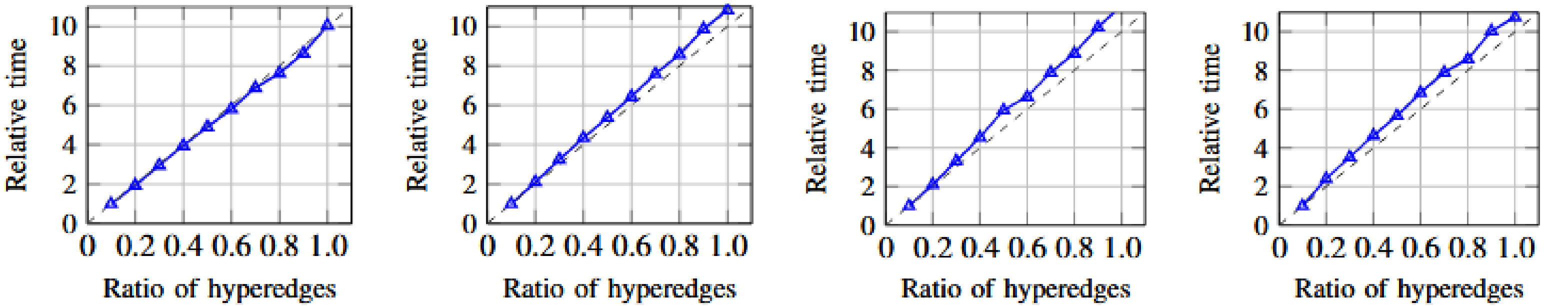


(a) Averaged AUROC on Citeseer (b) Averaged AUROC on Cora

Ko, Yunyong, Hanghang Tong, and Sang-Wook Kim. "Enhancing hyperedge prediction with context-aware self-supervised learning." *IEEE Transactions on Knowledge and Data Engineering* (2025).

Experiments

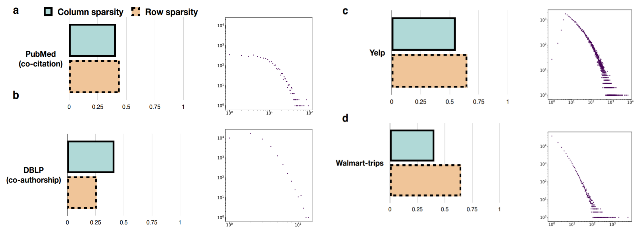
- Efficiency
  - Scale up linearly



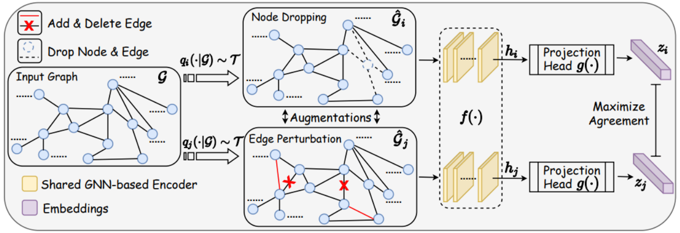
(a) Citeseer (b) Cora (c) DBLP-A (d) DBLP

Conclusion

- ☐ **Hypergraph data labeling is difficult because of its sparsity**



- ☐ **Graph Contrastive Learning**
  - ■ Maximize the agreement between the two views

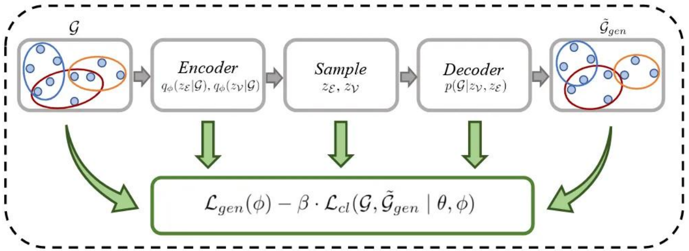


- ☐ **Challenges**
  - ■ How to augment a hypergraph?
  - ■ What to contrast?

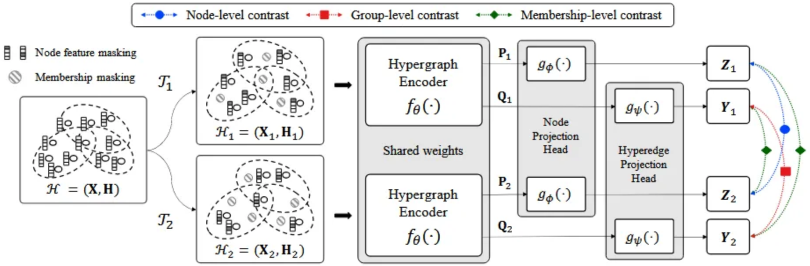
You, Yuning, et al. "Graph contrastive learning with augmentations." *Advances in neural information processing systems* 33 (2020): 5812-5823.

Conclusion

- ☐ **HyperGCL**
  - ■ Augmentation of hypergraph can be learned during contrastive learning using variational hypergraph auto-encoder



- ☐ **TriCL**
  - ■ Node + Group + Membership contrast



- ☐ **CASH**
  - ■ Context-Aware Hyperedge Prediction
  - ■ Hyperedge-aware masking

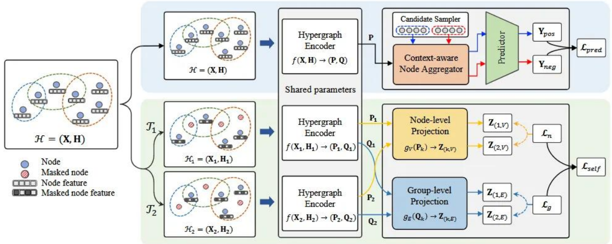


Fig. 2. The overview of CASH: (1) Context-aware hyperedge prediction (upper) and (2) Self-supervised contrative hypergraph learning (lower).

Wei, Tianxin, et al. "Augmentations in hypergraph contrastive learning: Fabricated and generative." *Advances in neural information processing systems* 35 (2022): 1909-1922.

Lee, Dongjin, and Kijung Shin. "I'm me, we're us, and i'm us: Tri-directional contrastive learning on hypergraphs." *Proceedings of the AAAI conference on artificial intelligence*. Vol. 37. No. 7. 2023.