Augmentations in Hypergraph Contrastive Learning: Fabricated and Generative

Tianxin Wei[1]*, Yuning You[2]*, Tianlong Chen[3], Yang Shen[2], Jingrui He[1], Zhangyang Wang[3]

[1]University of Illinois at Urbana-Champaign, [2]Texas A&M University, [3]University of Texas at Austin

{twei10,jingrui}@illinois.edu}, {yuning.you,yshen}@tamu.edu}, {tianlong.chen,atlaswang}@utexas.edu

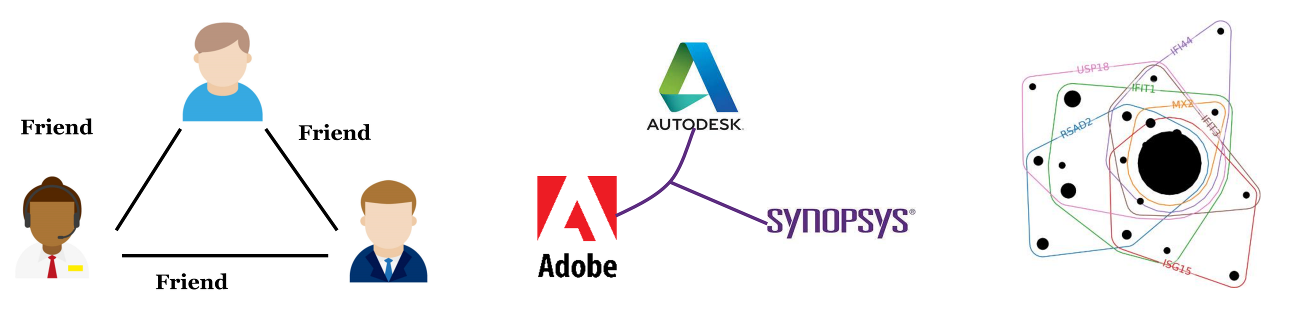
*Equal Contribution

AIFARMS Artificial Intelligence for Future Agricultural Resilience, Management, and Sustainability

NEURAL INFORMATION PROCESSING SYSTEMS

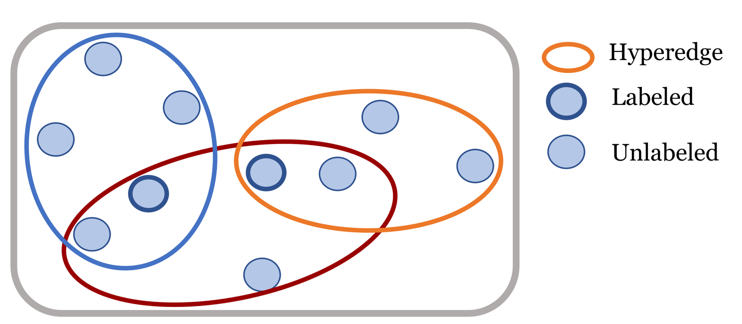
Background

- **Hypergraphs** have raised a surge of interest in the research community



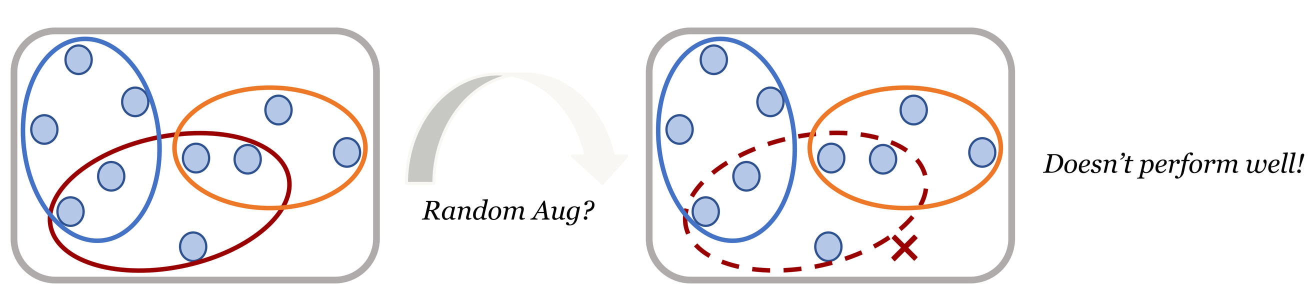
Recommender System — Financial Analyses — Bioinformatics

- **Label scarcity scenarios:** ubiquitous in real-world applications of hypergraphs



Restrict the **generalizability** of HyperGNNs!

- **Solution**: Develop **contrastive self-supervision** on hypergraphs, which relies on **augmentation construction**



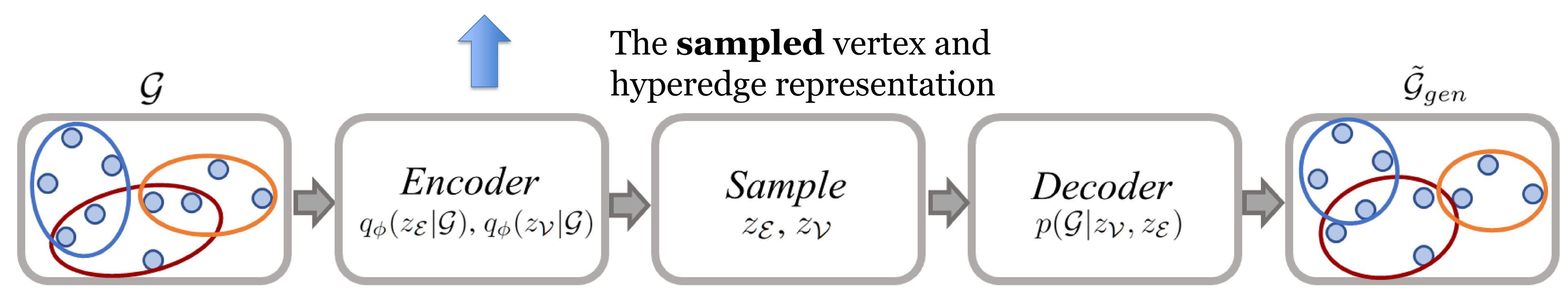
- **Bad construction** would result in negative transfer
- It's also **non-trivial** to build hypergraph views due to their overly intricate topology

$\sum_{e=1}^{N}\binom{N}{e}$ **possibilities** for one hyperedge on N vertices !

- **Question**: How to construct augmentations on hypergraph?
- **Contribution:** Propose to generate better augmented views in a **data-driven manner:**
  - ➢ Novel hypergraph generator → **parameterize** a certain **augmentation space** of hypergraphs
  - ➢ End-to-end pipeline → **jointly learn** hypergraph augmentations and model parameters

Method: Hypergraph Generative Models

- **Propose** a novel variational hypergraph auto-encoder architecture (**VHGAE**) to parametrize the augmentation space of **edge perturbation**
- **Encoder** embeds the **vertex** and **hyperedge** representation with **variational distribution**



- **Decoder** attempts to reconstruct the higher-order relations of hypergraph

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

Relation Reconstruction

- Optimize the evidence lower bound (ELBO):

Variational Regularization

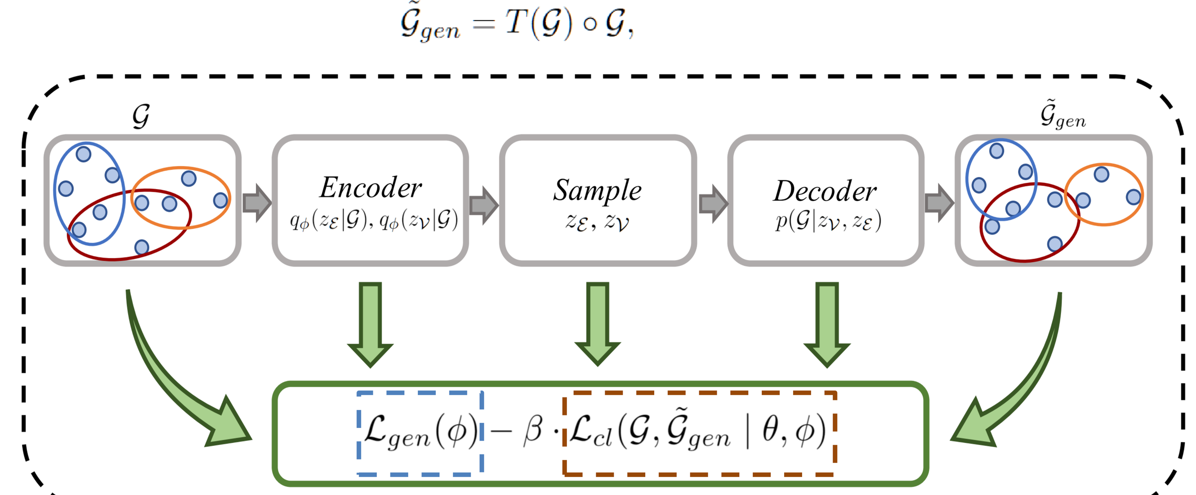
$$\text{ELBO} = \mathbb{E}_{q_\phi(z_{\mathcal{E}}|\mathcal{G})}\mathbb{E}_{q_\phi(z_{\mathcal{V}}|\mathcal{G})}\left[\log p_\theta(\mathcal{G}|z_v, z_e)\right] - \boxed{\text{KL}[q_\phi(z_{\mathcal{V}} \mid \mathcal{G}) \mid p(z_{\mathcal{V}})] - \text{KL}[q_\phi(z_{\mathcal{E}} \mid \mathcal{G}) \mid p(z_{\mathcal{E}})]},$$

Method: Jointly Augmenting and Contrasting

- Main barrier for applying VHGAE to augmentation and optimization: the **discrete sampling** of hyperedges which is **non-differentiable**

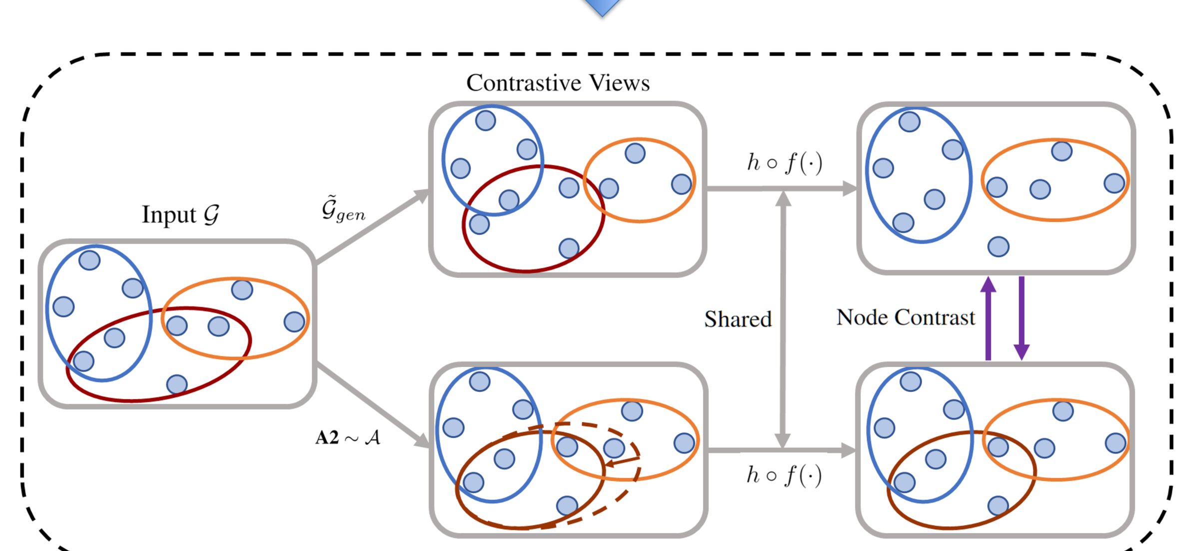
$$T(\mathcal{G}) = \boxed{\text{Gumbel-Softmax}}(p(\mathcal{G} \mid z_{\mathcal{V}}, z_{\mathcal{E}}))$$
$$= \text{Sigmoid}((w_{\mathcal{V}\mathcal{E}} + \log(\delta) - \log(1-\delta))/\tau)$$
$$\tilde{\mathcal{G}}_{gen} = T(\mathcal{G}) \circ \mathcal{G},$$

Differential Sampling



VHGAE Optimization

- ➢ Generator loss (-ELBO) to be minimized
- ➢ Maximizing CL loss in VHGAE to avoid capturing redundant information



HyperGNN Optimization

- ➢ Optimize HyperGNN with generated augmentation
- ➢ HyperGNN is eventually used for evaluation

Experiments

- Evaluate vertex classification performance on thirteen data sets
- Baselines
  - ➢ Existing hypergraph self-supervised learning methods
  - ➢ Fabricated Augmentation Operations
  - ➢ Our generative augmentation

	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**

- Proposed generative augmentation (A6) achieves **substantial improvements**

	Cora			Citeseer			ModelNet40		
	Random	Net	Minmax	Random	Net	Minmax	Random	Net	Minmax
SetGNN	66.87 ± 1.33	66.26 ± 1.54	66.58 ± 1.02	62.89 ± 1.57	62.81 ± 1.32	62.21 ± 1.64	95.74 ± 0.22	95.41 ± 0.28	93.33 ± 0.26
A2	71.90 ± 1.63	71.16 ± 0.92	70.86 ± 1.22	66.41 ± 1.08	65.38 ± 1.47	64.69 ± 0.98	96.09 ± 0.17	95.52 ± 0.24	93.64 ± 0.26
A4	72.11 ± 1.60	70.49 ± 1.29	70.52 ± 1.39	65.94 ± 1.24	65.15 ± 1.70	64.12 ± 1.19	95.79 ± 0.27	95.44 ± 0.25	93.35 ± 0.24
A6	**72.15 ± 1.70**	**71.94 ± 1.48**	**71.98 ± 1.36**	**66.60 ± 1.61**	**65.68 ± 1.09**	**65.51 ± 1.13**	**96.58 ± 0.24**	**96.23 ± 0.23**	**94.82 ± 0.33**
	NTU2012			House (0.6)			House (1.0)		
	Random	Net	Minmax	Random	Net	Minmax	Random	Net	Minmax
SetGNN	73.84 ± 2.18	73.38 ± 1.36	70.71 ± 1.89	67.16 ± 2.55	68.88 ± 2.68	64.78 ± 2.20	56.86 ± 1.93	59.95 ± 1.92	56.52 ± 2.52
A2	74.50 ± 2.03	73.86 ± 1.84	71.40 ± 1.64	67.71 ± 2.94	69.59 ± 2.32	65.23 ± 2.89	57.74 ± 2.70	60.73 ± 2.30	57.00 ± 1.94
A4	73.73 ± 1.59	73.72 ± 1.59	71.06 ± 1.53	67.55 ± 2.41	68.85 ± 1.38	64.97 ± 3.35	57.47 ± 2.72	60.10 ± 1.74	56.65 ± 2.26
A6	**75.06 ± 1.97**	**74.37 ± 1.99**	**72.09 ± 1.98**	**69.88 ± 3.27**	**73.14 ± 2.71**	**68.84 ± 2.71**	**60.06 ± 2.07**	**62.41 ± 1.77**	**58.76 ± 2.24**

data set	Method	AUROC	F1	$\Delta_{SP}(\downarrow)$	$\Delta_{EO}(\downarrow)$
German Credit	SetGNN	59.16 ± 2.51	81.84 ± 0.93	2.65 ± 5.62	4.06 ± 6.76
	A2	59.81 ± 3.00	82.26 ± 0.13	**0.55 ± 0.95**	0.78 ± 0.70
	A4	59.66 ± 3.83	80.54 ± 3.52	3.03 ± 6.54	5.07 ± 7.81
	A6	**59.88 ± 3.04**	**82.36 ± 0.38**	0.95 ± 0.92	**0.47 ± 0.56**
Recidivism	SetGNN	96.51 ± 0.48	89.84 ± 0.97	8.63 ± 0.50	4.16 ± 0.51
	A2	96.34 ± 0.39	**90.09 ± 0.53**	8.53 ± 0.52	3.92 ± 0.68
	A4	96.45 ± 0.35	89.75 ± 0.68	**8.49 ± 0.27**	3.49 ± 0.66
	A6	**96.55 ± 0.54**	89.22 ± 0.55	8.51 ± 0.25	**3.13 ± 0.64**
Credit defaulter	SetGNN	73.46 ± 0.17	87.91 ± 0.27	2.79 ± 0.99	0.98 ± 0.69
	A2	73.43 ± 0.27	87.82 ± 0.24	2.64 ± 1.32	0.93 ± 0.87
	A4	73.58 ± 0.19	87.92 ± 0.25	2.84 ± 1.14	1.38 ± 0.32
	A6	**73.78 ± 0.16**	**88.03 ± 0.14**	**2.58 ± 0.91**	**0.81 ± 0.37**

- First robustness and fairness evaluation for hypergraphs
- **Robust** against adversarial attacks
- **Fair** w.r.t. sensitive attributes

Conclusion

- **Problem**: Label scarcity scenarios of Hypergraph Application
- **Algorithms**: Generative Hypergraph Contrastive Learning (HyperGCL)
  - ➢ Parametrize the augmentation space
  - ➢ Jointly learn augmentation and model
- **Evaluation**: Effectiveness on generalization, robustness, and fairness

Acknowledgment

This work is supported by National Science Foundation under Award No. IIS-1947203, IIS-2117902, IIS-2137468, CCF-1943008; US Army Research Office Young Investigator Award W911NF2010240; National Institutes of General Medical Sciences under grant R35GM124952; and Agriculture and Food Research Initiative grant no. 2020-67021-32799/project accession no.1024178 from the USDA National Institute of Food and Agriculture. The views and conclusions are those of the authors and should not be interpreted as representing the official policies of the government agencies.