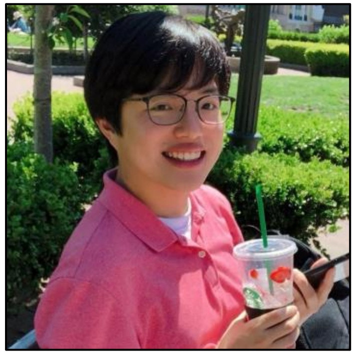
AHP: Learning to Negative Sample for Hyperedge Prediction


Hyunjin Hwang*

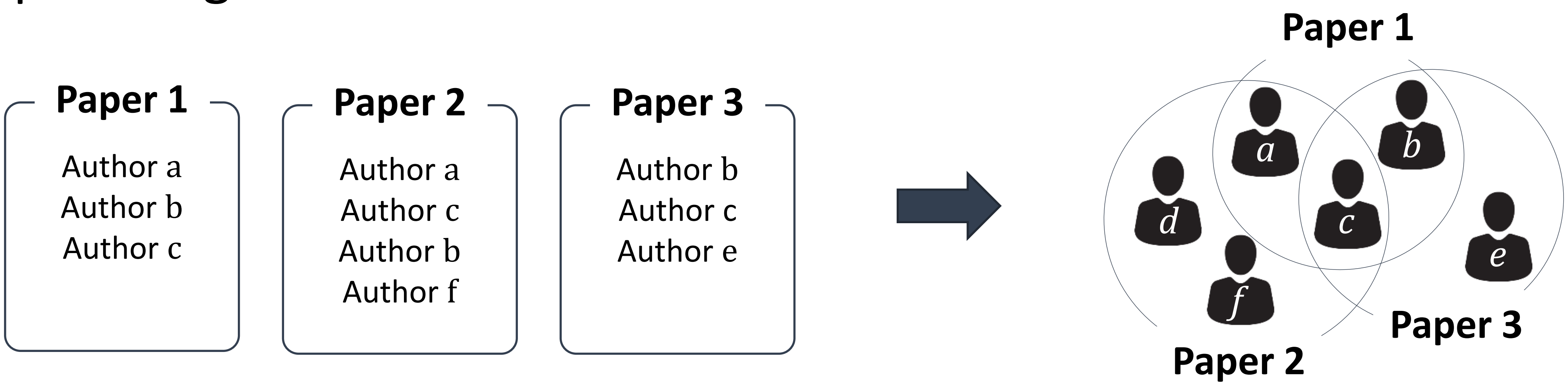

Seungwoo Lee*


Chanyoung Park


Kijung Shin

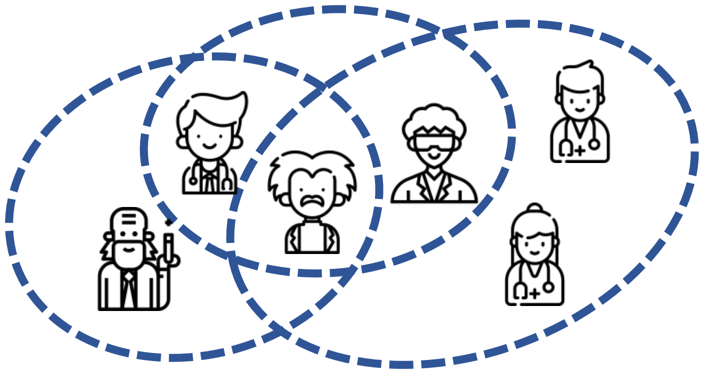
Backgrounds (1) Hypergraphs

- **Hypergraphs, $\boldsymbol{H(V, E)}$**
  - Hypergraphs represent interactions among multiple nodes
  - Each hyperedge $e \in E$ is a subset of node set V
- **Advantages**:
  - Have more expressive power than ordinary graphs
  - Capture higher-order information

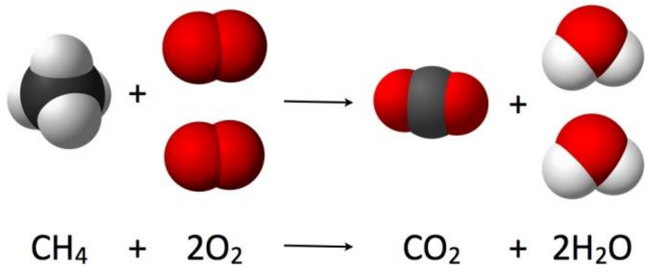


Backgrounds (2) Hyperedge Prediction

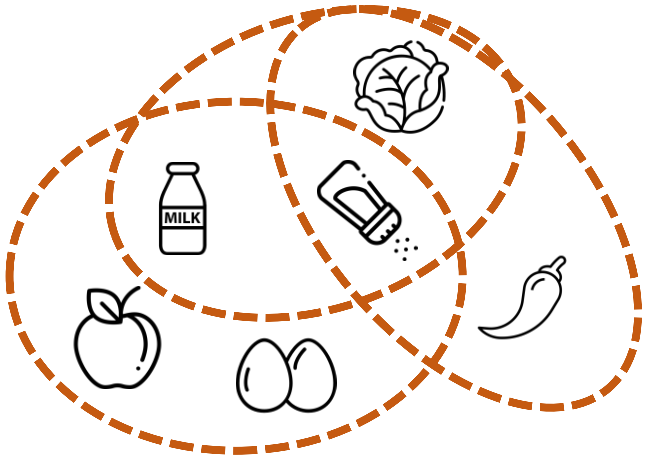
- **Hyperedge prediction:** Predict future or unobserved hyperedges
- **Applications**
  - Collaboration recommendation
  - Chemical reaction prediction
  - Recipe discovery / Drug discovery
  - E-mail recipient recommendation



Collaborations of Researchers



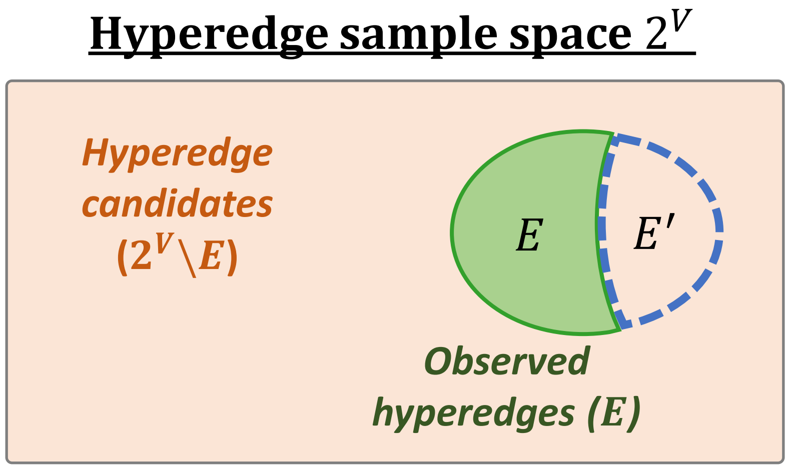
Chemical Reactions



Ingredients in Recipes

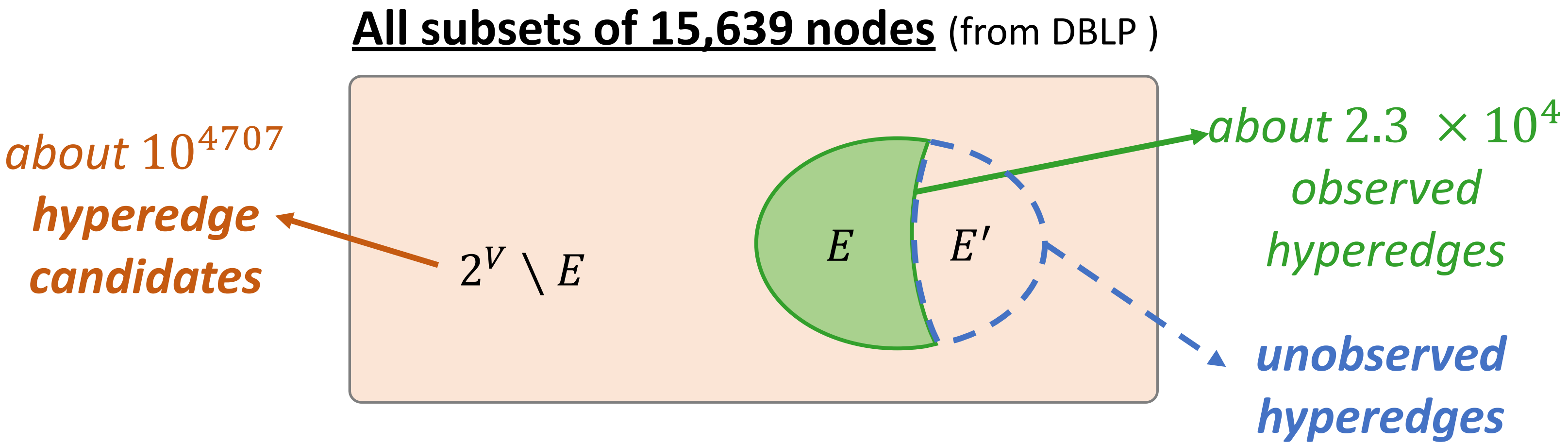
Problem Definition

- **True hyperedge set $E \cup E'$**
  - E: observed hyperedge set
  - **E': unobserved hyperedge set (target set)**
- **Hyperedge candidate, c**
  - A subset of nodes that can be a hyperedge, but currently not observed ($c \notin E$)
- **Problem definition. (Hyperedge prediction)**
  - **Given**: a hypergraph $G(V, E)$ and a hyperedge candidate c
  - **Classify**: **candidate** c whether it would be in the **target set** E' or not



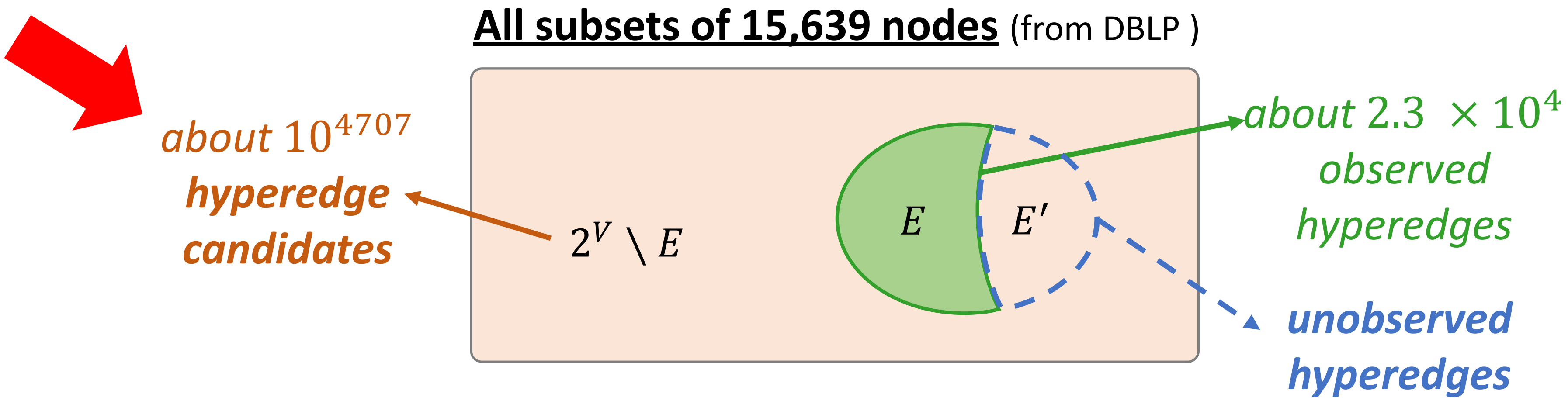
Challenges and Motivation

- **Vast sample space**
  - The number of possible hyperedge candidates is $2^{|V|}$
  - Sampling hyperedge candidates for classification task is inevitable
- **Poor generalization ability**
  - NN models should generalize on the sampled hyperedge candidates to all
  - Heuristic sampling schemes in previous works generalize poorly

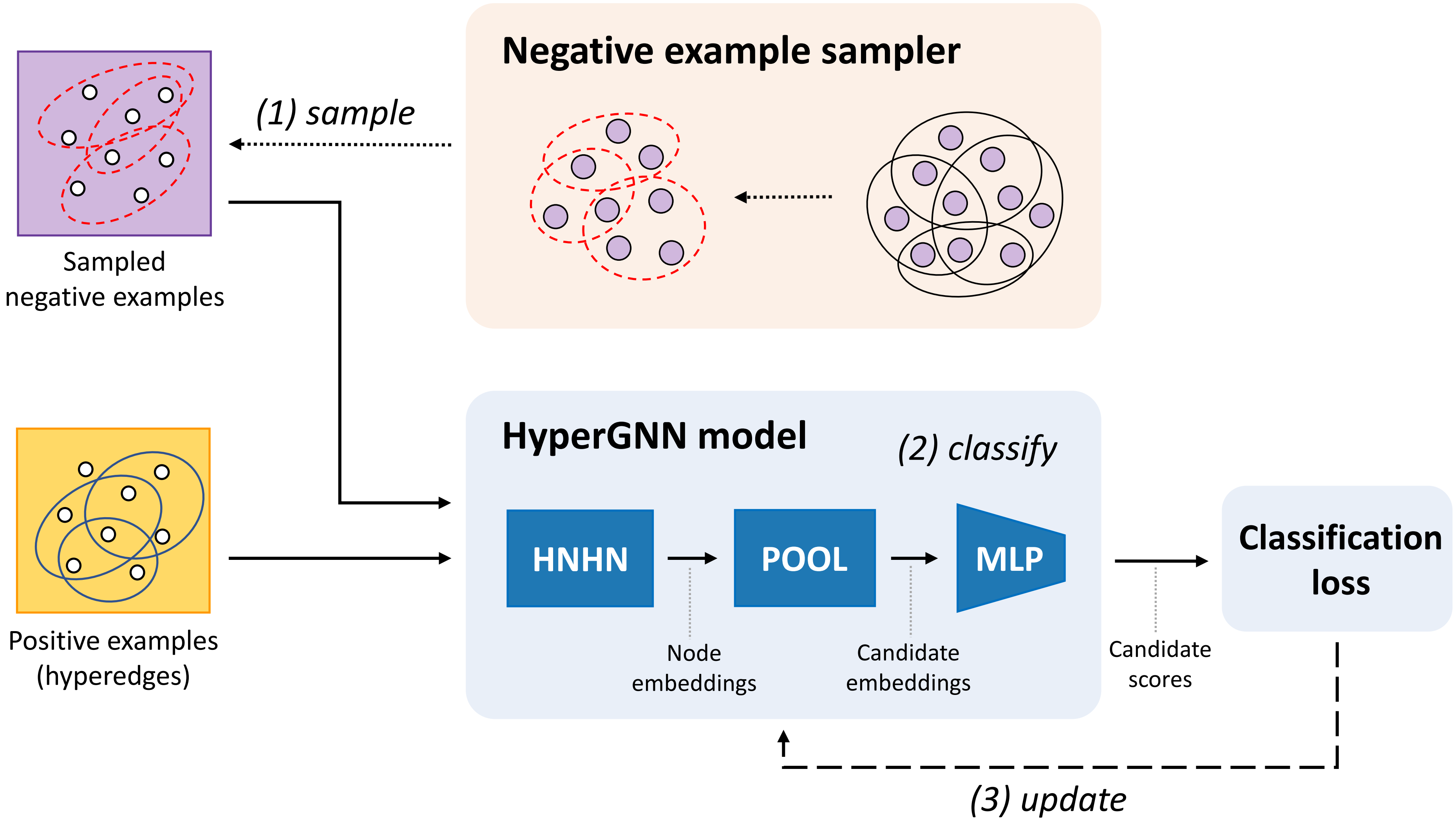


Challenges and Motivation

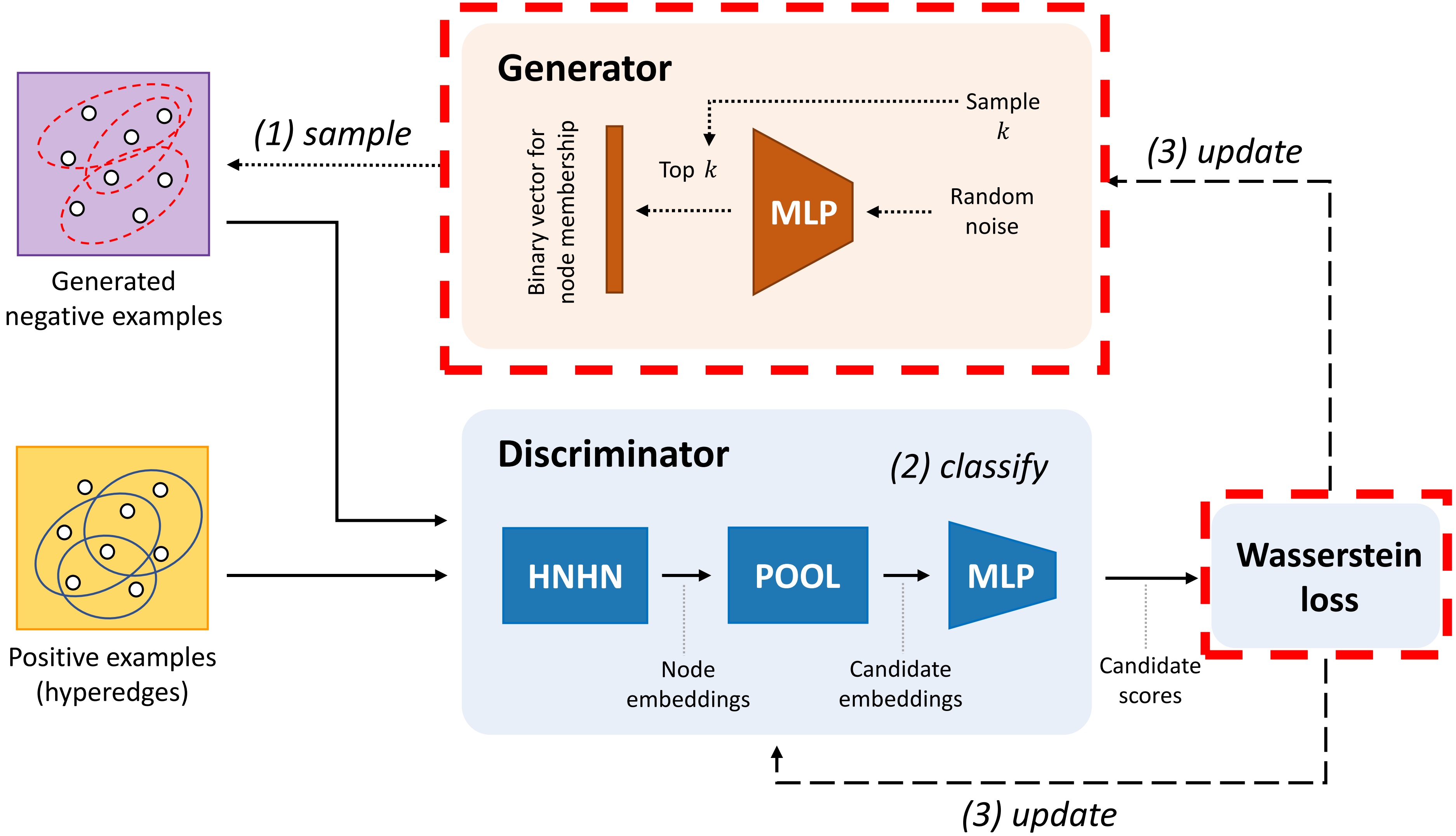
- **Vast sample space**
  - The number of possible hyperedge candidates is $2^{|V|}$
  - Sampling hyperedge candidates for classification task is inevitable
- **Poor generalization ability**
  - NN models should generalize on the sampled hyperedge candidates to all
  - Heuristic sampling schemes in previous works generalize poorly



Previous Approaches



Proposed: Adversarial Hyperedge Prediction (AHP)



Experiments (1) Datasets

- **Collaboration dataset** (node: author, hyperedge: authors in a paper)
- **Authorship dataset** (node: paper, hyperedge: papers by an author)
- **Co-citation dataset** (node: paper, edge: cited papers in a paper)

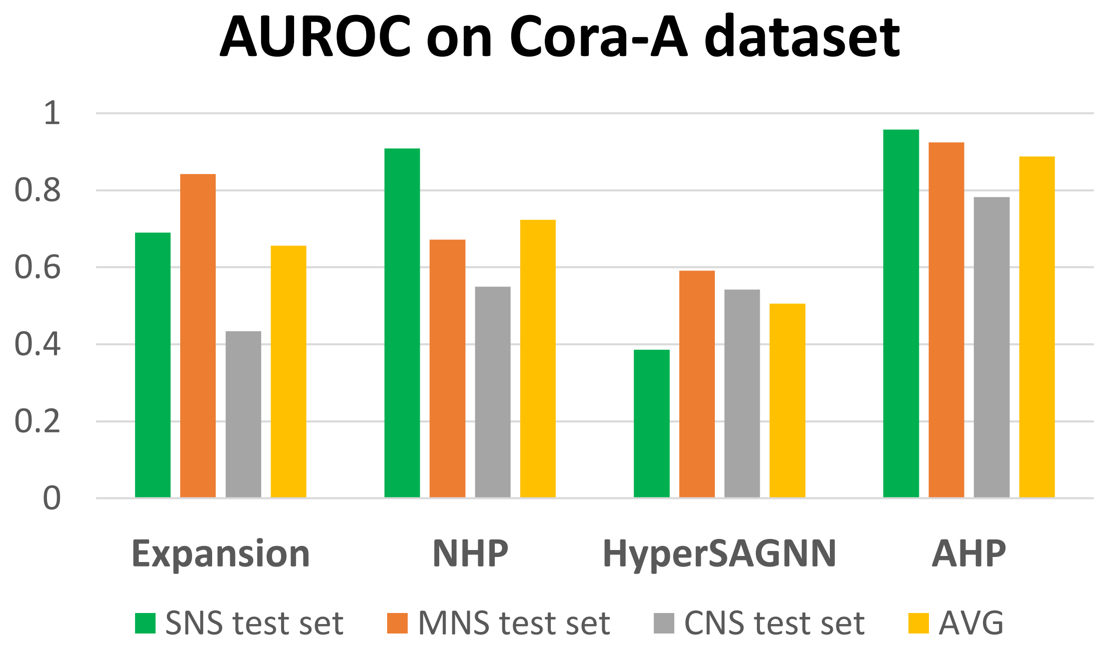
Property	Citeseer	Cora	Cora-A	Pubmed	DBLP-A	DBLP
Domain	Co-citation	Co-citation	Authorship	Co-citation	Authorship	Collaboration
Number of nodes	1,457	1,434	2,388	3,840	39,283	15,639
Number of edges	1,078	1,579	1,072	7,962	16,483	22,964
Average size of hyperedges	3.2	3	4.3	4.3	4.5	2.7
Maximum size of hyperedges	26	5	43	171	80	18
Minimum size of hyperedges	2	2	2	2	2	2
Dimension of node feature	3,703	1,433	1,433	500	4543	4543

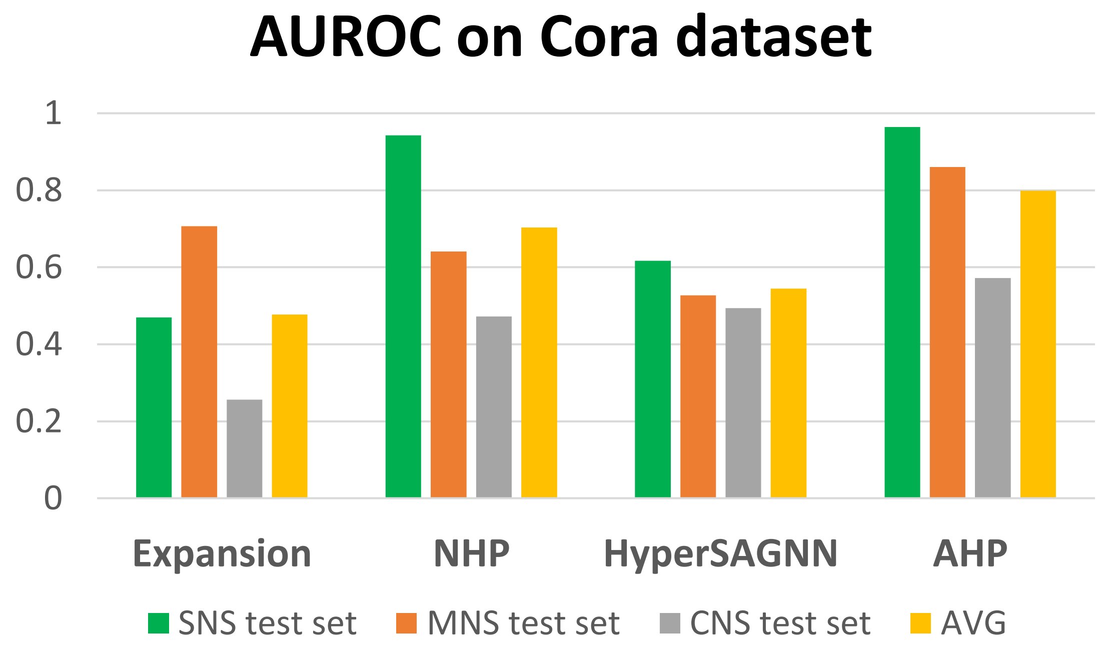
Experiments (2) Baselines

Model	Node embedding	Edge pooling	Negative sampling
NHP	NHP	Maxmin	Half corruption
HyperSAGNN	Random walk	HyperSAGNN	MNS
N-order expansion	Six features from given candidate		Stars in clique expansion
AHP-S	HNHN	Maxmin	SNS
AHP-M	HNHN	Maxmin	MNS
AHP-C	HNHN	Maxmin	CNS
AHP-S+M+C	HNHN	Maxmin	SNS+MNS+CNS
AHP	HNHN	Maxmin	Generator

Experiments (3) Results: Generalization Ability

- State of the art methods have **poor generalization ability**





Experiments (3) Results: Overall Performance

- Our proposed method, **AHP, outperforms on average**

Predictive Performance	Cora-A		Cora		Citeseer		Pubmed		DBLP		DBLP-A	
	AUROC	AP	AUROC	AP	AUROC	AP	AUROC	AP	AUROC	AP	AUROC	AP
Expansion	0.656	0.706	0.477	0.607	0.591	0.681	0.497	0.612	0.607	0.698	0.603	0.687
NHP	0.723	0.748	0.703	0.718	0.751	0.751	0.733	0.707	0.569	0.544	0.716	0.693
HyperSAGNN	0.506	0.571	0.545	0.584	0.475	0.521	0.584	0.576	0.531	0.582	0.634	0.675
AHP-S	0.872	0.869	0.777	0.764	0.781	0.787	0.759	0.740	0.773	0.757	0.813	0.825
AHP-M	0.873	0.874	0.777	0.758	0.748	0.756	0.751	0.738	0.774	**0.764**	0.813	0.826
AHP-C	0.876	0.878	0.767	0.749	0.777	0.786	0.727	0.724	0.736	0.723	0.812	0.829
AHP-S+M+C	0.874	0.878	0.771	0.753	0.774	0.779	0.760	0.748	0.767	0.759	0.814	0.832
AHP (Proposed)	**0.888**	**0.882**	**0.799**	**0.772**	**0.824**	**0.819**	**0.763**	**0.749**	**0.778**	**0.764**	**0.837**	**0.850**

Conclusions

- **Observation:** **Heuristic sampling schemes limit the generalization ability** of deep learning-based hyperedge-prediction.

- **Algorithm**: **AHP learns to sample negative examples** by adversarial training for better generalization

- **Results:** In terms of AUROC, **AHP is up to 28.2% better than best existing methods** and up to 5.5% better than variants with sampling schemes tailored to test sets

Code & Datasets: https://github.com/HyunjinHwn/SIGIR22-AHP

Thank you!