香港中文大學
The Chinese University of Hong Kong

Online Social Network Representation: A Brief Introduction and Its Future Direction

Xiangguo Sun

The Chinese University of Hong Kong

2023.08.29

About Me (Dr. Xiangguo SUN)

- **Research Keywords:**
  - Social Media Analytics, Online Social Networks
  - Graph Learning
- **Homepage:**
  - https://xgsun.mysxl.cn
- **Contact:**
  - WeChat:

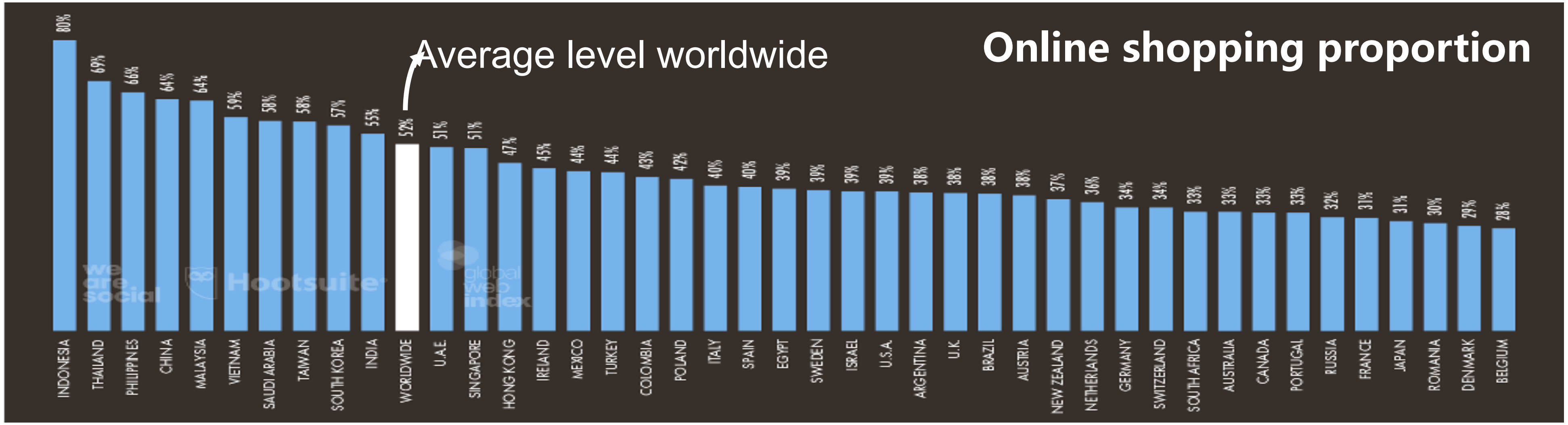
https://xgsun.mysxl.cn

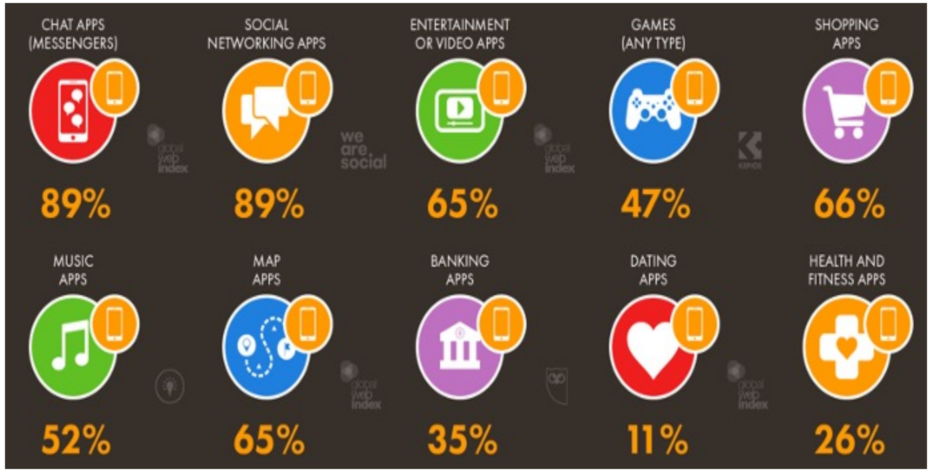
Outline

- A Brief Introduction to Online Social Networks
- Social Relation Analysis
- Sociological Analysis via Deep Learning Technique
- From LLM to LGM: Prompting graphs like ChatGPT

Brief Story about Online Social Networks 1/5

- **status quo**





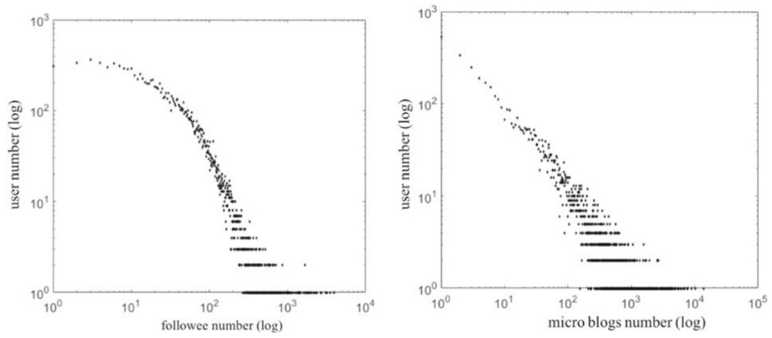
Online action types

“one smartphone, everything done”

Simon Kemp. Digital in 2020: global digital overview. We are Social, Hootsuite, 2020.

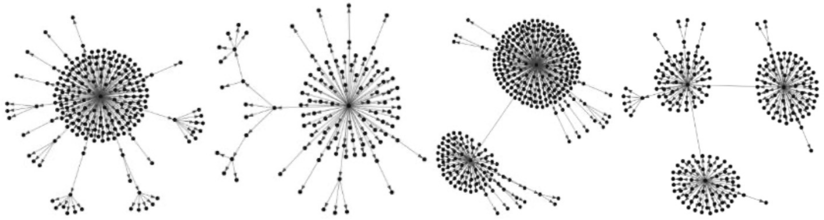
Brief Story about Online Social Networks 2/5

➢ Sparsely Observed but Rich Social Relationships

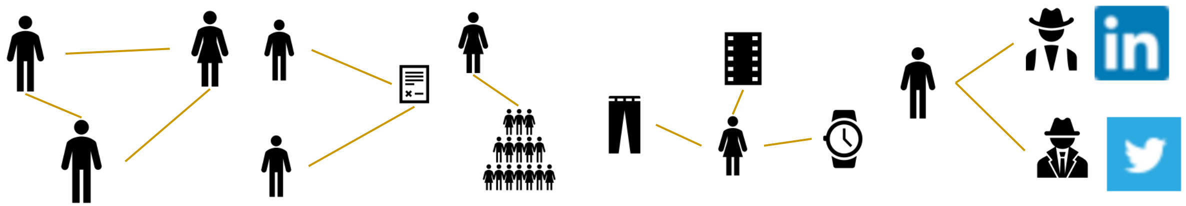


Scale-free Networks

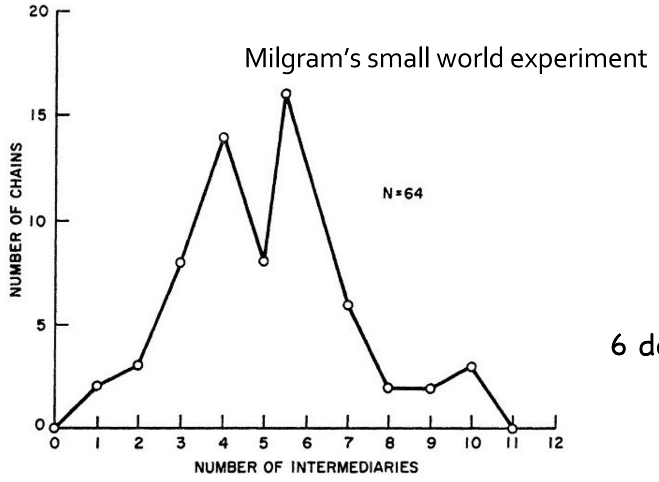
crowdsourcing



media supervision



Heterogeneous Networks

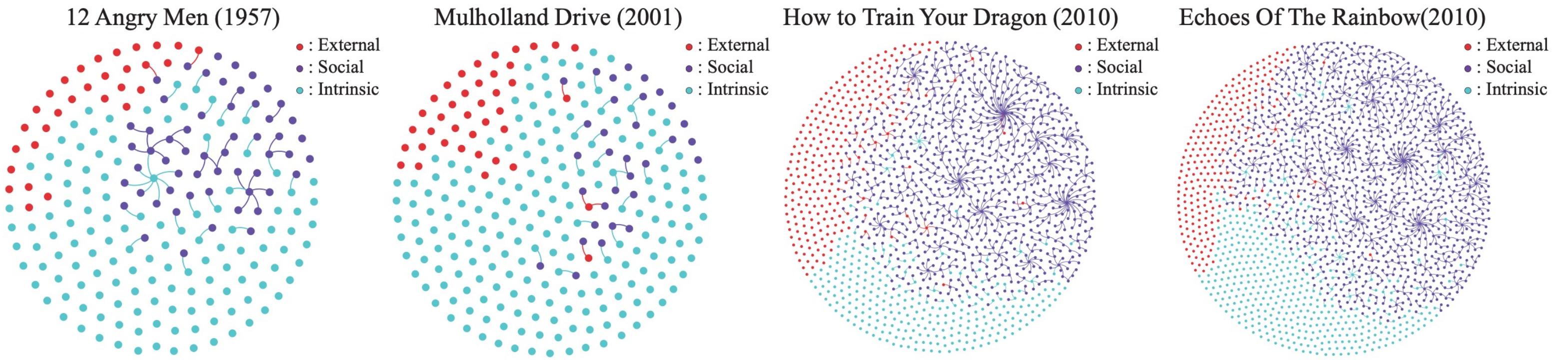


6 degrees of separation

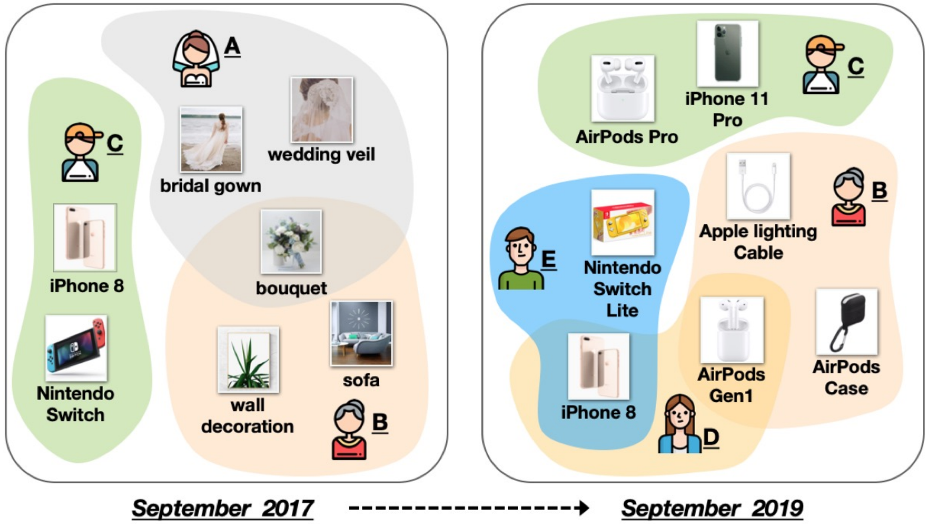
Small-World

Brief Story about Online Social Networks 3/5

➢ Dynamically Evolving driven by Sociology



Yu Rong et al. Why It Happened: Identifying and Modeling the Reasons of the Happening of Social Events. KDD 2015



Brief Story about Online Social Networks 4/5

Broad Connections to Psychology

- Conformity and Herd Behavior
- Group Polarization

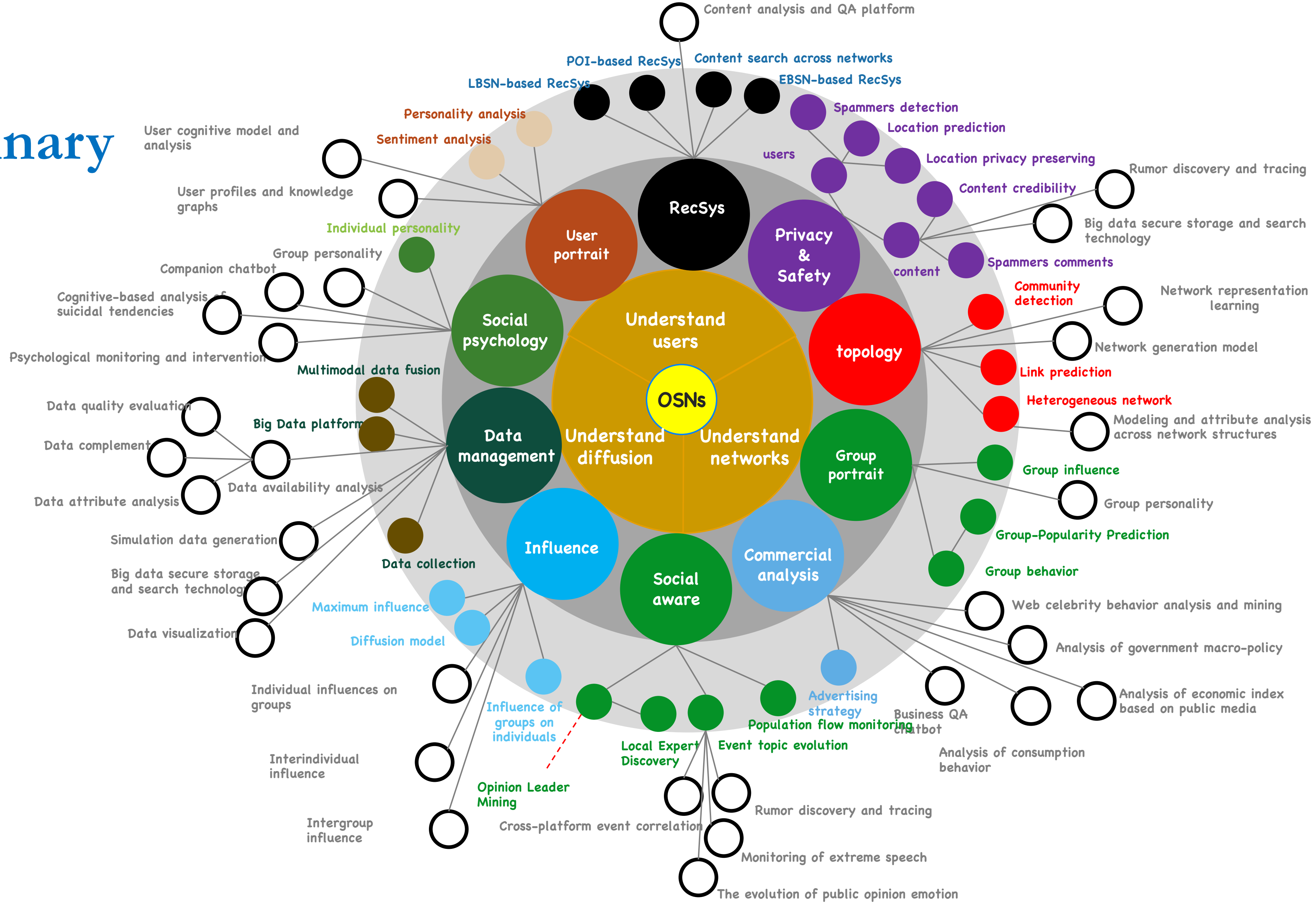
"I don't know why. I just suddenly felt like calling."

2021 United States Capitol attack

- merely on the basis of 10 Facebook "likes." : evaluate a person better than the average work colleague.
- 70 "likes" were enough to outdo what a person's friends knew
- 300 "likes" what their partner knew

Brief Story about Online Social Networks 5/5

- Highly Interdisciplinary Study

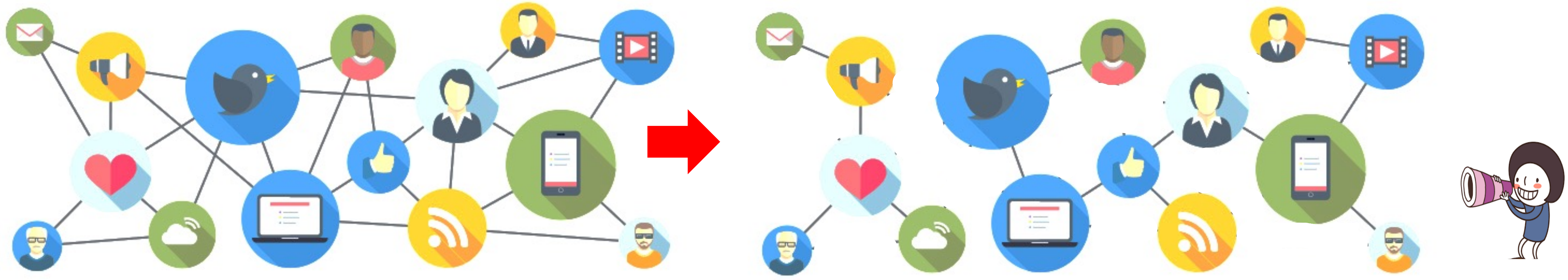


Social Relation Analysis 1/7

Xiangguo Sun et al., Multi-level Hyperedge Distillation for Social Linking Prediction on Sparsely Observed Networks, The Web Conference 2021

Motivations

- Social linking prediction is one of the most fundamental problems in online social networks and has attracted researchers’ persistent attention.
- However, in the real-world, most social links are hidden and the observed network is extremely limited, making the observed links very sparse.



Xiangguo Sun et al., Multi-level Hyperedge Distillation for Social Linking Prediction on Sparsely Observed Networks, The Web Conference 2021

Motivations

- Existing works mostly learn higher-order information via pair-wise learning. However, when the observed links are extremely limited, the learned embeddings might be under-smooth.
- Hypergraphs are better in higher-level learning, but toneless hyperedges lead to over-smooth expressiveness for linking prediction.

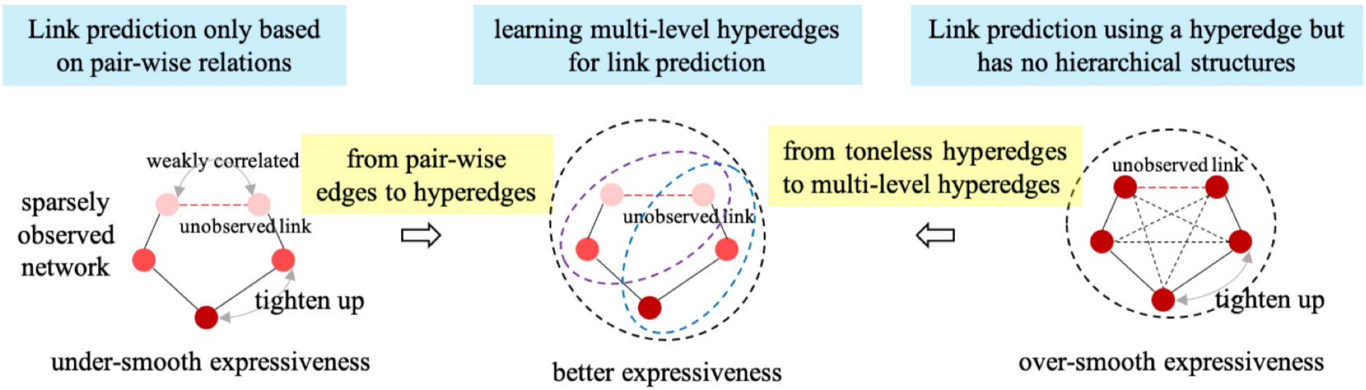
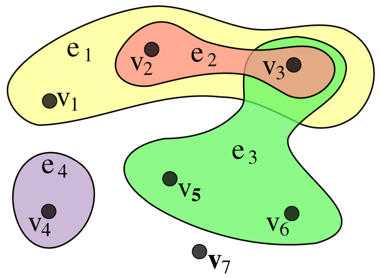


Figure 1: pair-wise relations (left), multi-level hyperedges (middle), and toneless hyperedges (right).



a hypergraph allows one hyperedge to connect multiple nodes

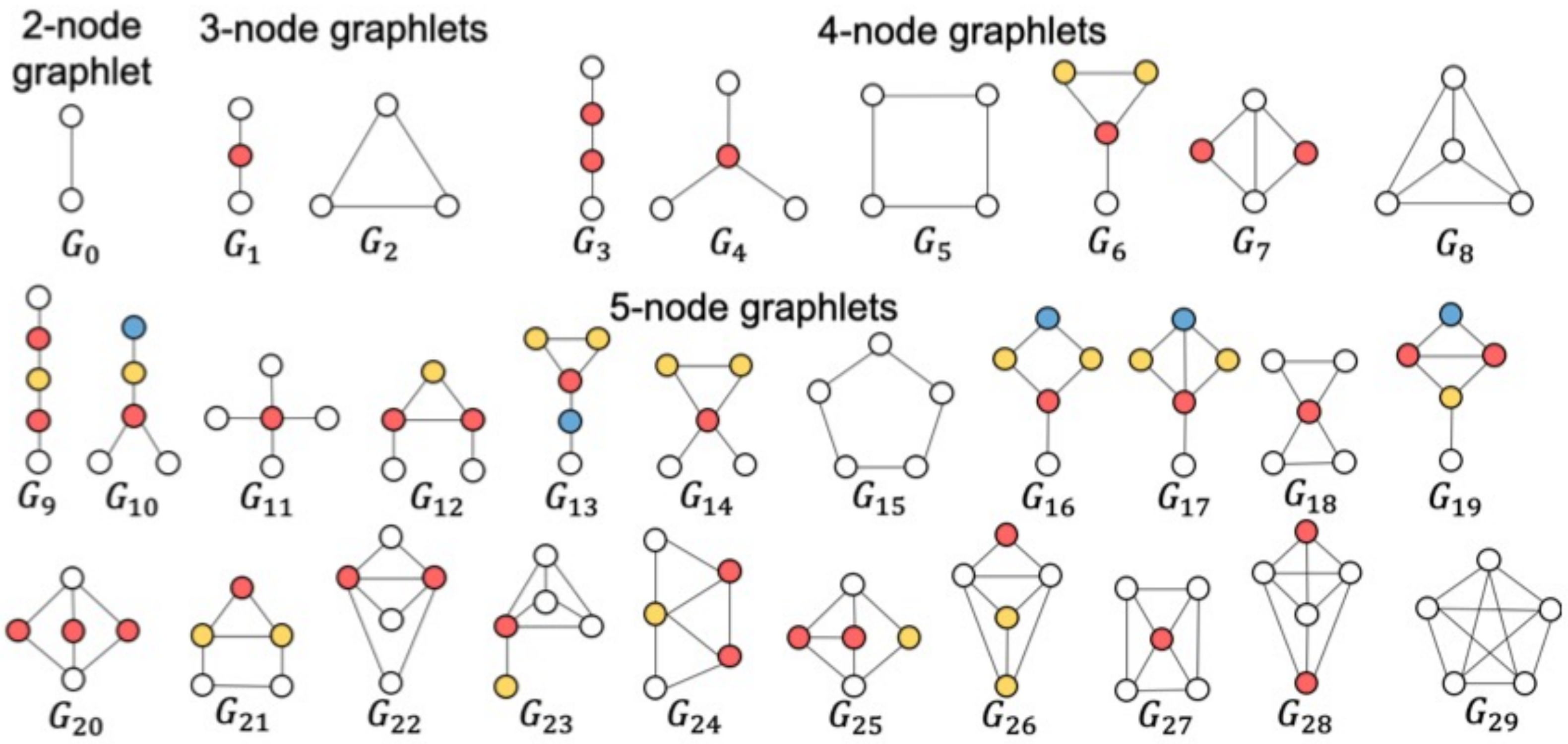
We need to learn multi-level hyperedges automatically from data.

Social Relation Analysis 3/7

Xiangguo Sun et al., Multi-level Hyperedge Distillation for Social Linking Prediction on Sparsely Observed Networks, The Web Conference 2021

Graphlets-driven Hyperedges.

- Merits: structurally complete, hierarchical to some extend, go beyond pair-wise relations.
- Demerits: not sufficient to preserve global structures; A larger network usually contains enormous graphlets; not hierarchical enough



Social Relation Analysis 4/7

Xiangguo Sun et al., Multi-level Hyperedge Distillation for Social Linking Prediction on Sparsely Observed Networks, The Web Conference 2021

- **Hyperedge Expansion.**

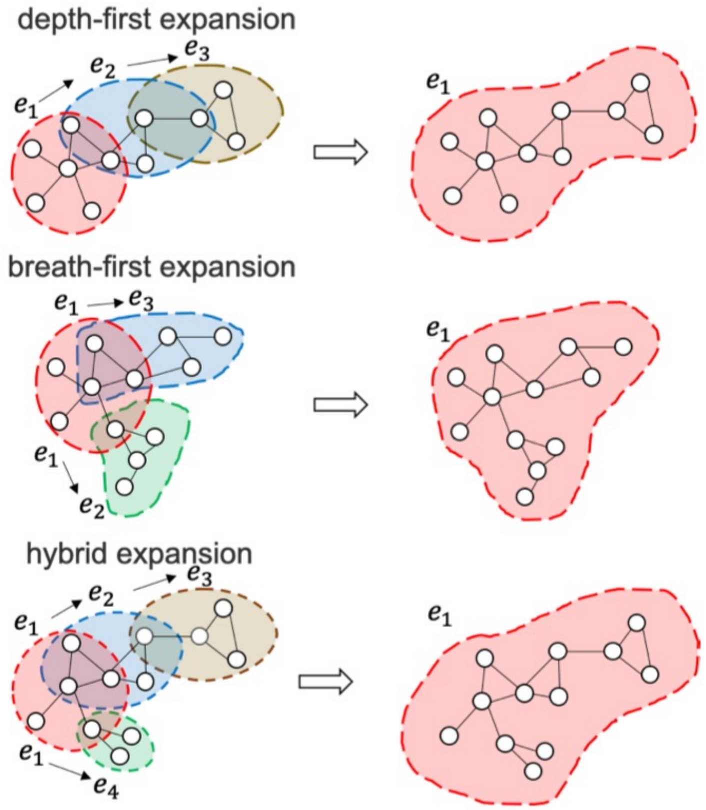


Figure 4: Three hyperedge expansion strategies. top: depth-first expansion; middle: breadth-first expansion; bottom: hybrid expansion.

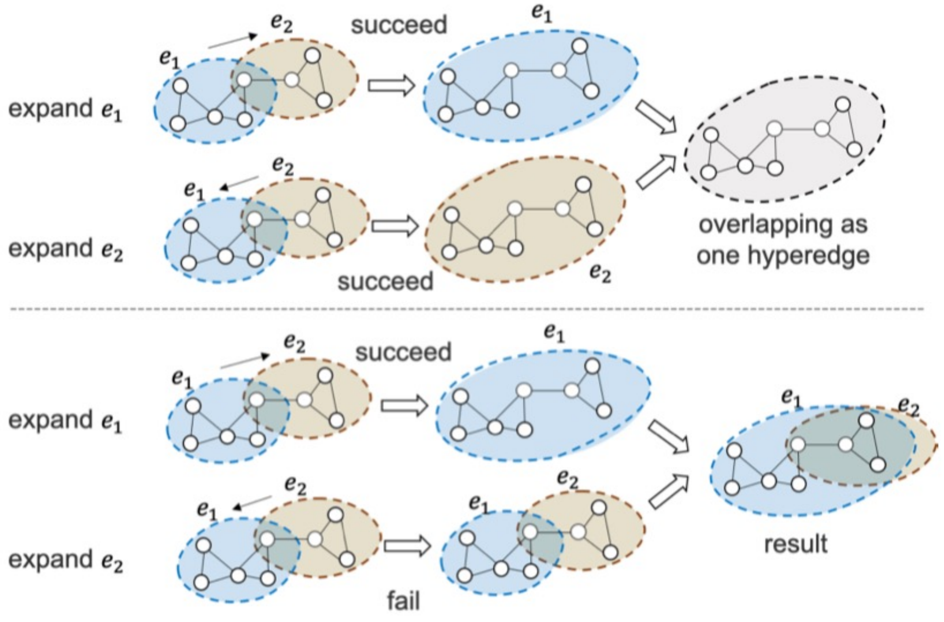


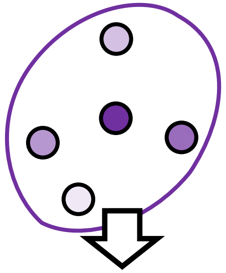
Figure 5: Hyperedges expansion cases. Through hyperedge expansion, we can reduce the number of hyperedges and construct multi-level hyperedges.

Social Relation Analysis 5/7

Xiangguo Sun et al., Multi-level Hyperedge Distillation for Social Linking Prediction on Sparsely Observed Networks, The Web Conference 2021

➢ Hypergraph Neural Network.

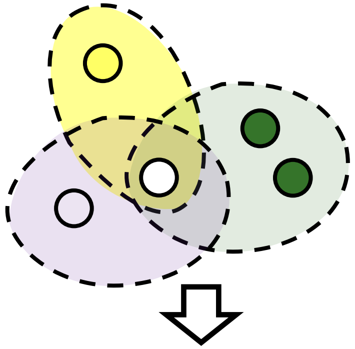
Node-level Attention.



$$p_v^l(i|j,e) = \frac{\exp\left(\sigma\left(\mathrm{H}^l(i,e)\cdot \mathrm{U}_v^l(i,i)\cdot \left[\mathrm{z}_i^l \oplus \mathrm{z}_j^l\right]\cdot \mathrm{p}_l^{\top}\right)\right)}{\sum_{t=1}^{|\mathcal{V}|}\exp\left(\sigma\left(\mathrm{H}^l(t,e)\cdot \mathrm{U}_v^l(t,t)\cdot \left[\mathrm{z}_t^l \oplus \mathrm{z}_j^l\right]\cdot \mathrm{p}_l^{\top}\right)\right)}$$

$$\mathrm{z}_{i|e}^l = \sigma\left(\sum_{j=1}^{|\mathcal{V}|}\frac{\mathrm{U}_v^l(j,j)\,\mathrm{H}^l(j,e)}{\mathrm{D}_e^l(e,e)}\cdot p_v^l(i|j,e)\cdot \mathrm{z}_j^l\right)$$

Hyperedge-level Attention.

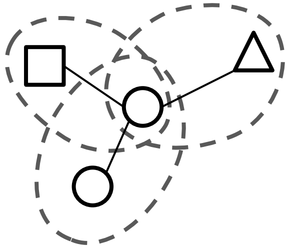


$$\mathrm{m}_e^l = \sum_{i=1}^{|\mathcal{V}|}\mathrm{H}^l(i,e)\cdot \mathrm{z}_{i|e}^l$$

$$\alpha_e^l = \sigma\left(\mathrm{U}_e(e,e)\cdot \tanh\left(\mathrm{m}_e^l\cdot \mathrm{W}_\alpha^l + \mathrm{b}_\alpha^l\right)\cdot \mathrm{q}_l^{\top}\right)$$

$$\mathrm{z}_i^{l|\Phi} = \sum_{t=1}^{|\mathcal{E}_l^{\triangleright|\Phi}|}\frac{\mathrm{H}^{l|\Phi}(i,t)}{\mathrm{D}_v^{l|\Phi}(i,i)}\cdot \alpha_{e_t}^{l|\Phi}\cdot \mathrm{z}_{i|e_t}^{l|\Phi}$$

Semantic-level Attention.



$$\beta_{\Phi_i}^l = \frac{\exp\left(\sum_{j=1}^{|\mathcal{V}|}\tanh\left(\mathrm{W}_\beta^l\cdot \mathrm{z}_j^{l|\Phi_i} + \mathrm{b}_\beta^l\right)\cdot \mathrm{f}_l^{\mathrm{T}}\right)}{\sum_{t=1}^{P}\exp\left(\sum_{j=1}^{|\mathcal{V}|}\tanh\left(\mathrm{W}_\beta^l\cdot \mathrm{z}_j^{l|\Phi_t} + \mathrm{b}_\beta^l\right)\cdot \mathrm{f}_l^{\mathrm{T}}\right)}$$

$$\mathrm{Z}^{l+1} = \sum_{i=1}\beta_{\Phi_i}^l\cdot \mathrm{Z}^{l|\Phi_i}$$

Social Relation Analysis 6/7

Xiangguo Sun et al., Multi-level Hyperedge Distillation for Social Linking Prediction on Sparsely Observed Networks, The Web Conference 2021

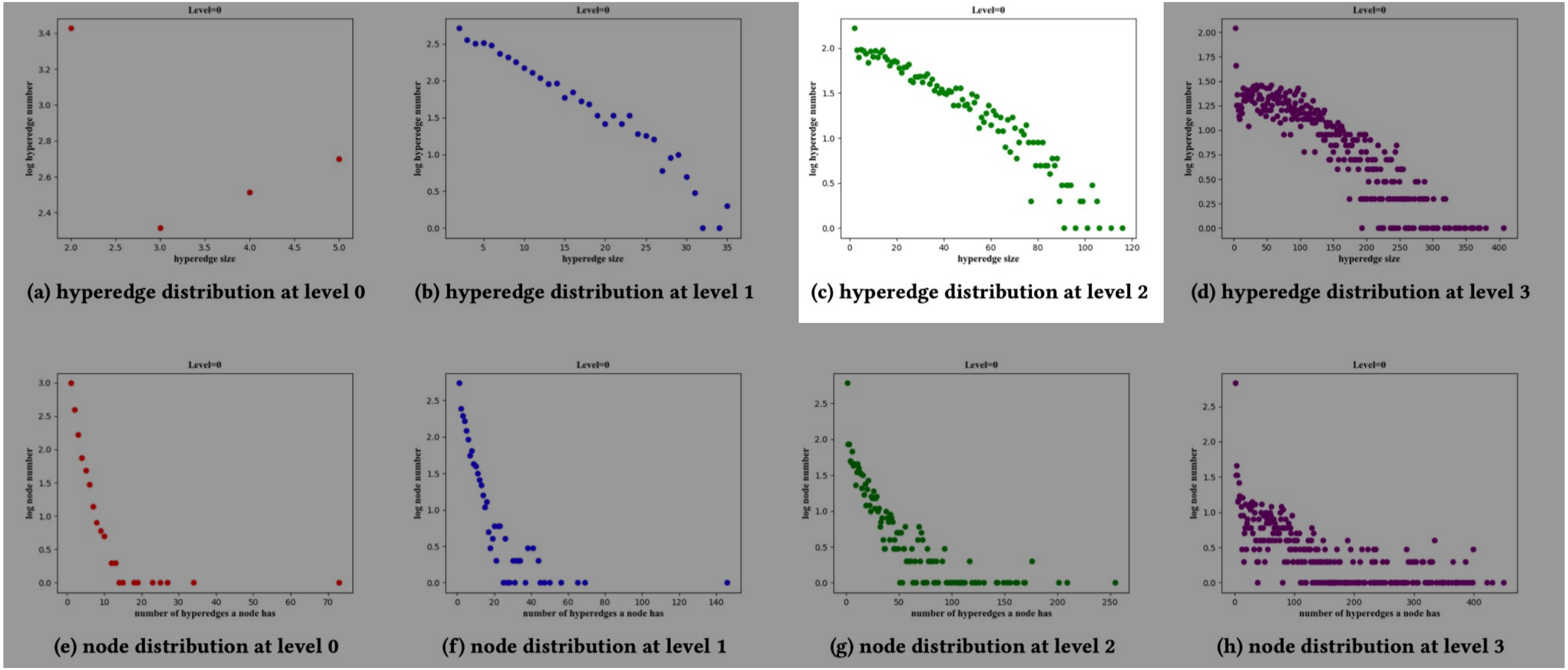
➢ Experiment

Metrics	Methods	Cora (P-P)	DBLP (P-C)	DBLP (P-A)	ACM (P-A)	ACM (P-S)	Yelp (B-U)	Yelp (B-L)
AUC	Ours	0.6288	0.7251	0.6661	0.6284	0.5975	0.5672	0.5924
	Metapath2vec	0.5013	0.5534	0.5429	0.5432	0.5367	0.5016	0.5035
	PME	0.5187	0.6070	0.5859	0.5509	0.5477	0.5054	0.5002
	HAN	0.5564	0.6206	0.6032	0.5673	0.5564	0.5203	0.5326
	HGNN	0.5552	0.6531	0.5678	0.5298	0.5435	0.5191	0.5498
	MGCN	0.5673	0.6904	0.6035	0.5964	0.5975	0.5367	0.5564
AP	Ours	0.1455	0.1606	0.1469	0.1868	0.1438	0.1104	0.1206
	Metapath2vec	0.0993	0.1025	0.1054	0.1003	0.0997	0.0993	0.0919
	PME	0.0979	0.1009	0.1033	0.0941	0.0913	0.0930	0.0914
	HAN	0.1203	0.1326	0.1207	0.1154	0.1207	0.1069	0.1204
	HGNN	0.1133	0.1393	0.1061	0.0943	0.1372	0.1009	0.1165
	MGCN	0.1208	0.1524	0.1203	0.1206	0.1438	0.1096	0.1201
F1	Ours	0.8705	0.8804	0.8742	0.8972	0.8964	0.7506	0.8002
	Metapath2vec	0.5602	0.6932	0.7321	0.5564	0.6037	0.5735	0.5942
	PME	0.6336	0.7348	0.7675	0.6194	0.6857	0.6015	0.6088
	HAN	0.7806	0.8802	0.8697	0.8742	0.8703	0.7562	0.8009
	HGNN	0.7952	0.8697	0.8635	0.8597	0.8713	0.7438	0.7929
	MGCN	0.8079	0.8702	0.8633	0.8864	0.8806	0.7506	0.8002

Social Relation Analysis 7/7

Xiangguo Sun et al., Multi-level Hyperedge Distillation for Social Linking Prediction on Sparsely Observed Networks, The Web Conference 2021

➢ Experiment



(a) hyperedge distribution at level 0

(b) hyperedge distribution at level 1

(c) hyperedge distribution at level 2

(d) hyperedge distribution at level 3

(e) node distribution at level 0

(f) node distribution at level 1

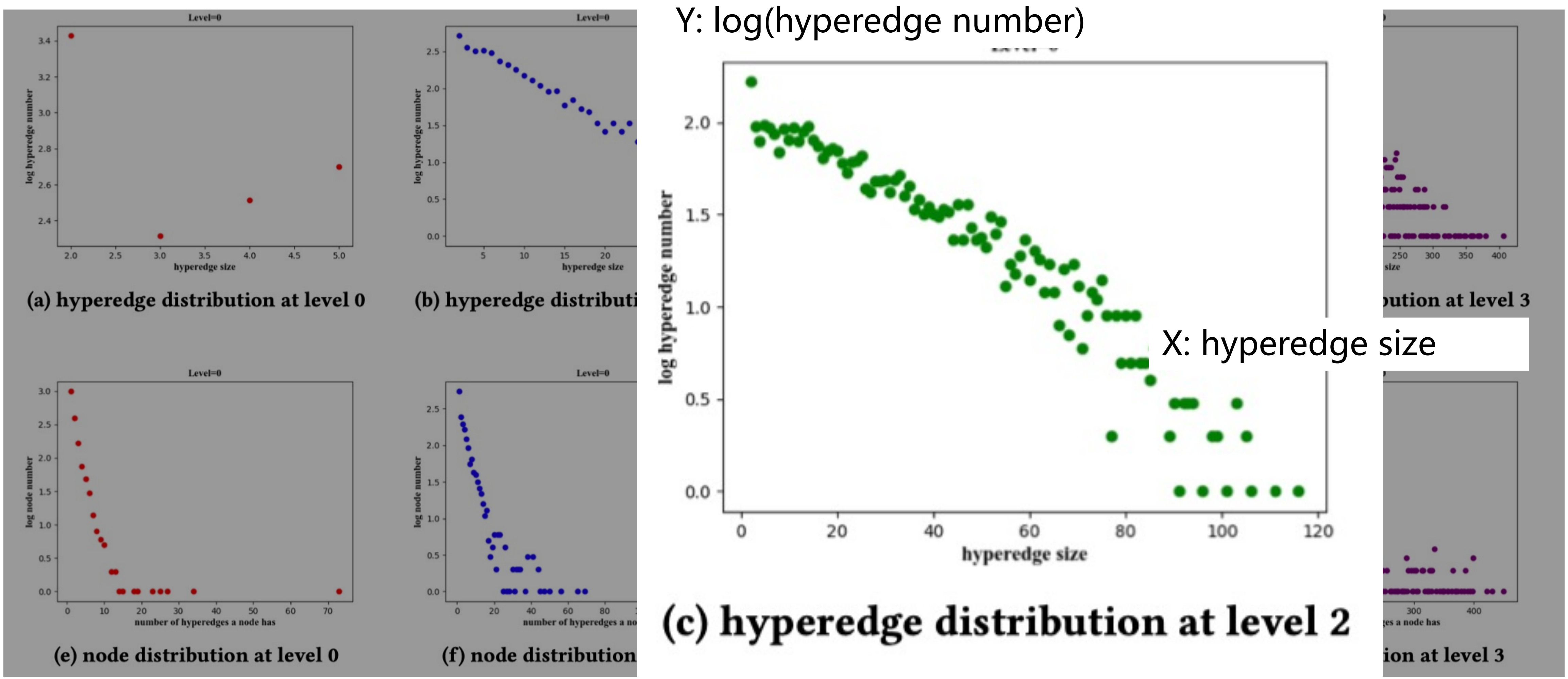
(g) node distribution at level 2

(h) node distribution at level 3

Social Relation Analysis 7/7

Xiangguo Sun et al., Multi-level Hyperedge Distillation for Social Linking Prediction on Sparsely Observed Networks, The Web Conference 2021

➢ Experiment



Xiangguo Sun et al. Self-supervised Hypergraph Representation Learning for Sociological Analysis. IEEE Transactions on Knowledge and Data Engineering (TKDE) 2023

Motivations

Traditional Data Mining Framework

Data → Model → Prediction → Label

Label → Model: Backward gradient

It is difficult to embody the endogenous laws of sociology: the social interaction of the real world does not follow the framework of gradient propagation

Our solution:

The social interaction between users and the environment is simulated, and a sociology-oriented data learning framework is constructed

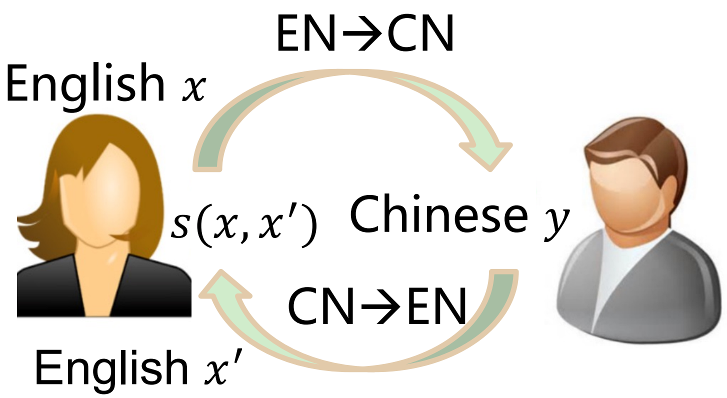
Challenge

- How to simulate the interaction betwe individuals and their environments?

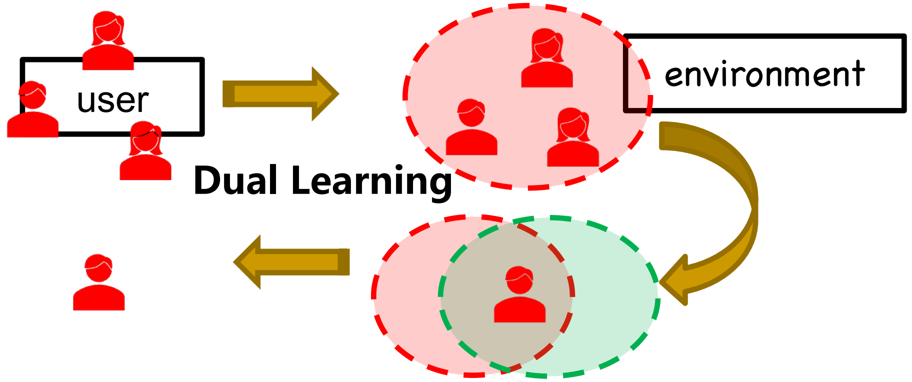
(a) social equivalence. (b) social conformity.

Xiangguo Sun et al. Self-supervised Hypergraph Representation Learning for Sociological Analysis. IEEE Transactions on Knowledge and Data Engineering (TKDE) 2023

- **Dual Learning**
  - A user's role in a social network is determined by the network context (structure) – structural equivalence
  - The network environment is determined by similar group members – homogeneity



Dual learning in machine translation

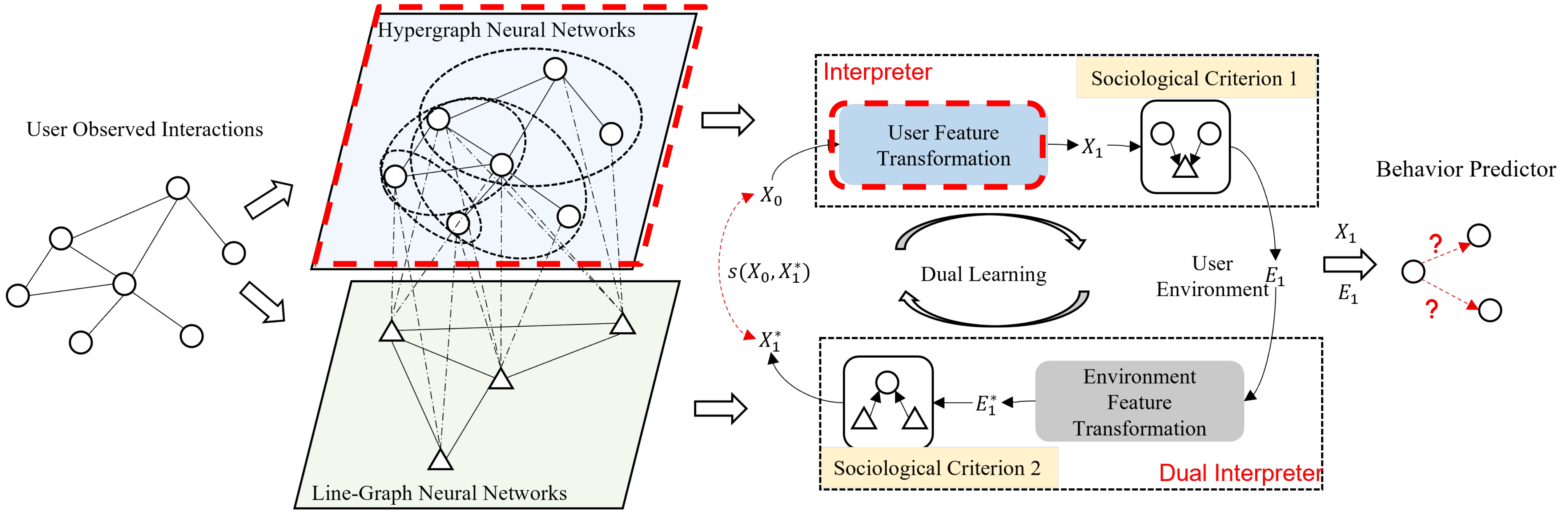


Online roles and online environments

Sociological Analysis via Deep Learning Technique 3/8

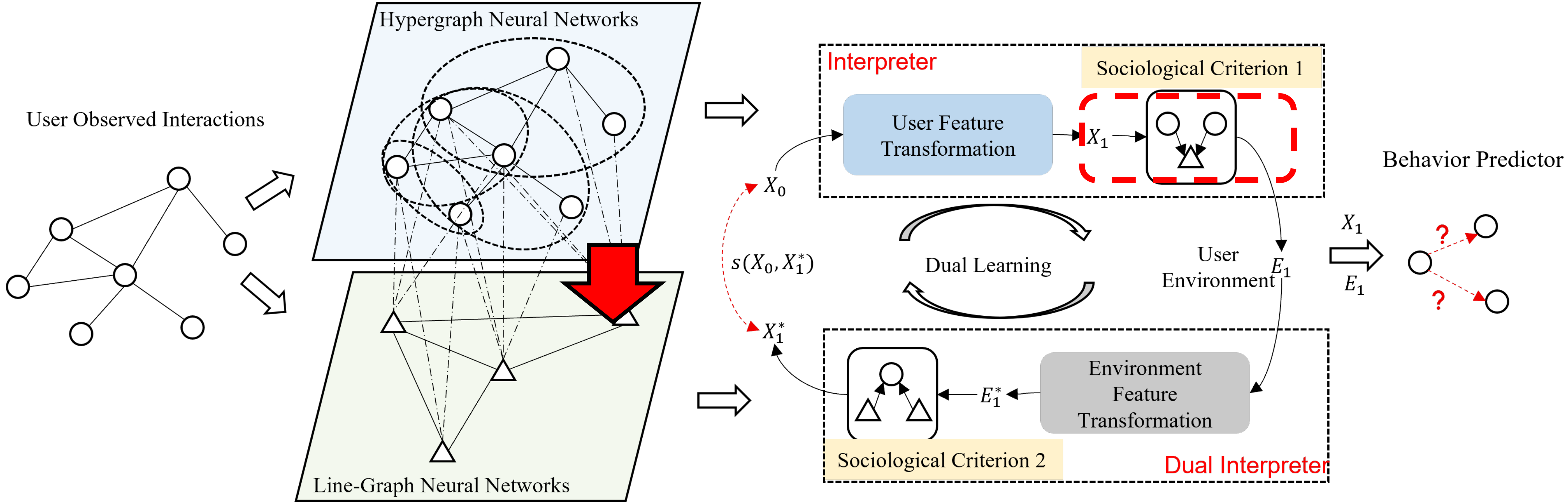
➢ Dual Learning

Xiangguo Sun et al. Self-supervised Hypergraph Representation Learning for Sociological Analysis. IEEE Transactions on Knowledge and Data Engineering (TKDE) 2023



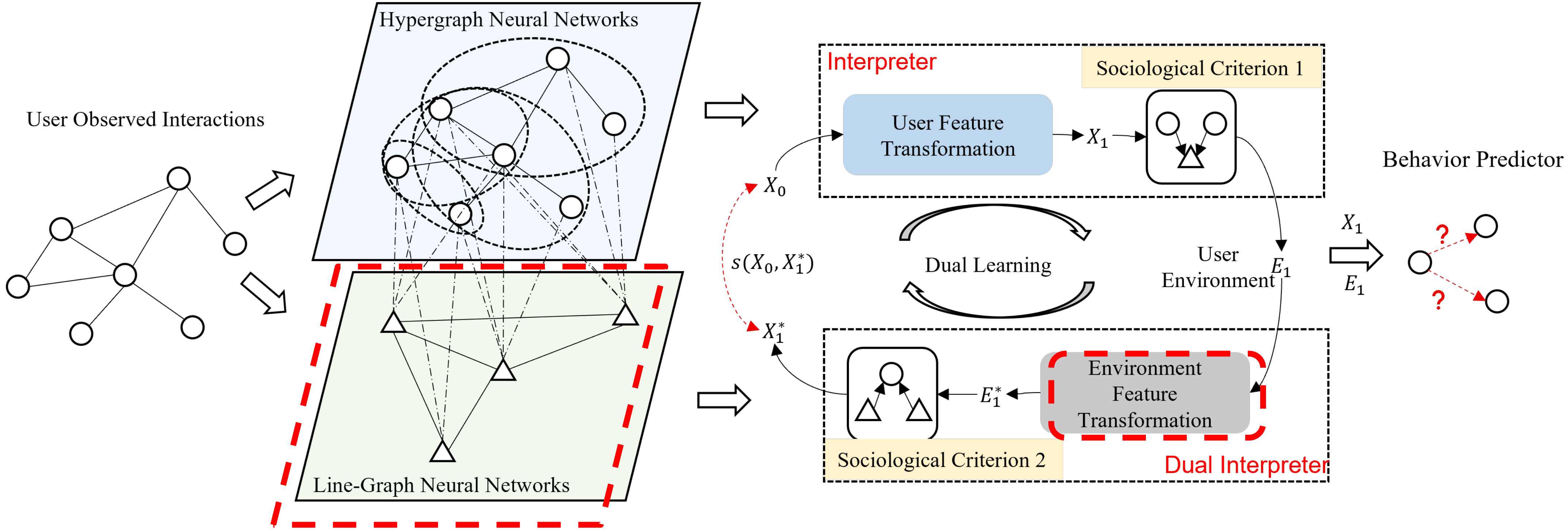
➢ Dual Learning

Xiangguo Sun et al. Self-supervised Hypergraph Representation Learning for Sociological Analysis. IEEE Transactions on Knowledge and Data Engineering (TKDE) 2023



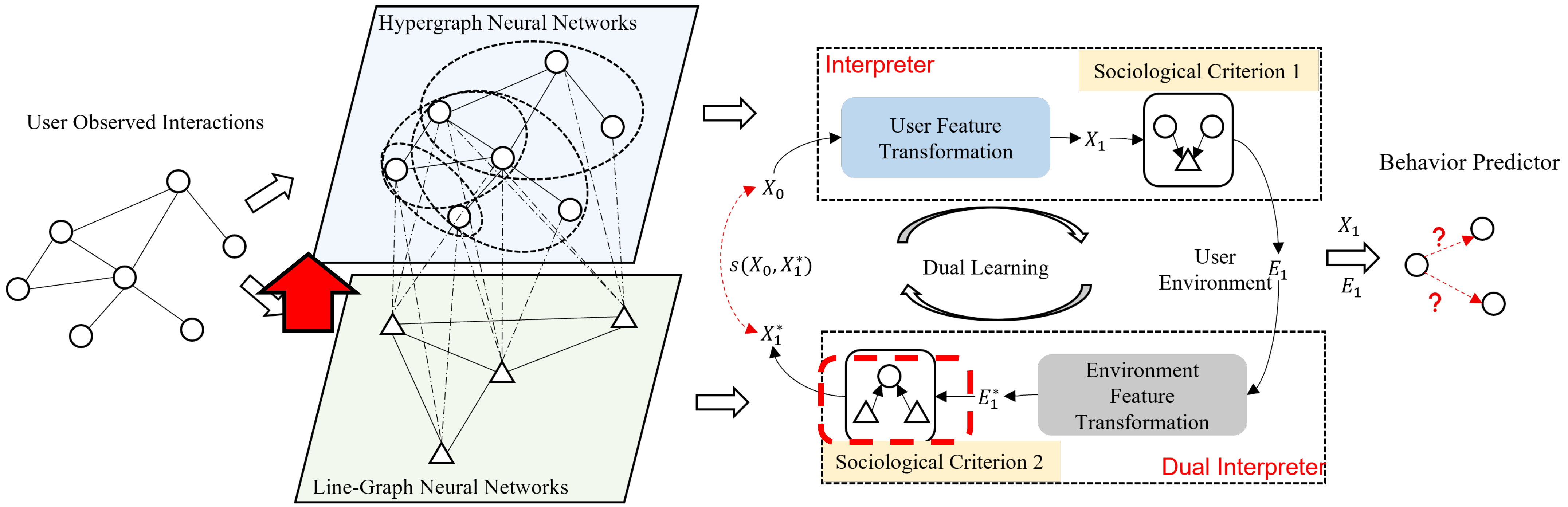
Dual Learning

Xiangguo Sun et al. Self-supervised Hypergraph Representation Learning for Sociological Analysis. IEEE Transactions on Knowledge and Data Engineering (TKDE) 2023



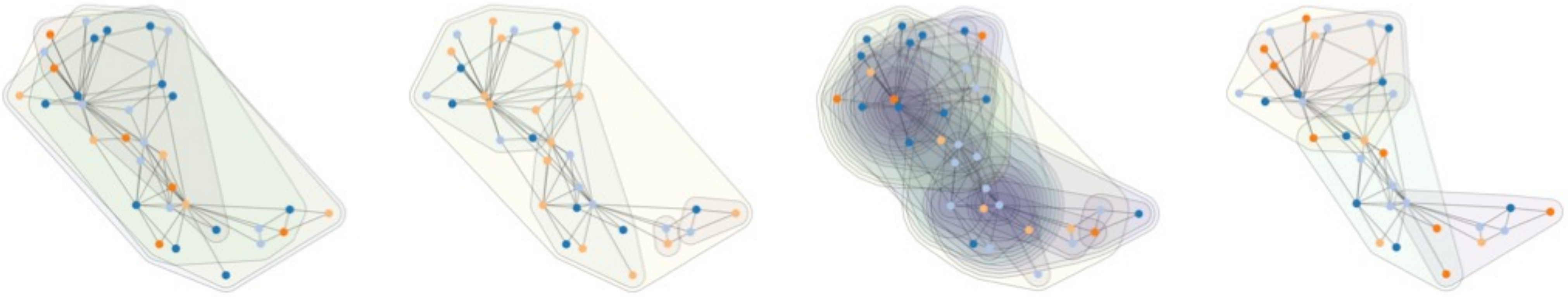
➢ Dual Learning

Xiangguo Sun et al. Self-supervised Hypergraph Representation Learning for Sociological Analysis. IEEE Transactions on Knowledge and Data Engineering (TKDE) 2023



Xiangguo Sun et al. Self-supervised Hypergraph Representation Learning for Sociological Analysis. IEEE Transactions on Knowledge and Data Engineering (TKDE) 2023

- # Experiment

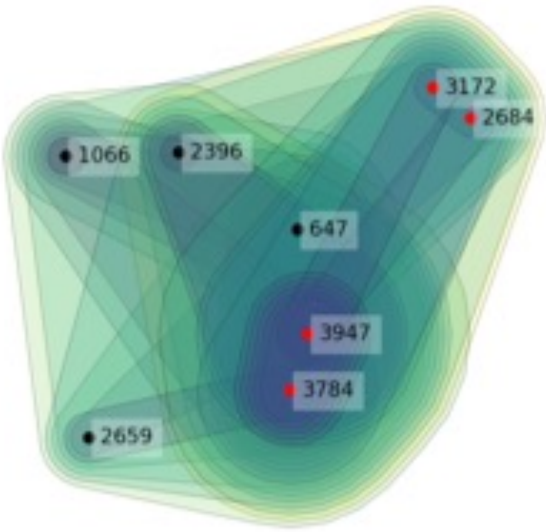


(a) cluster (b) louvain community (c) one-hop neighbour (d) Ours

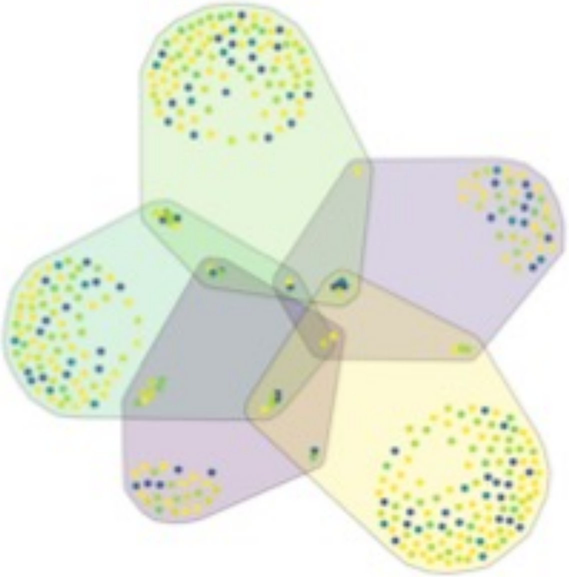
Sociological Analysis via Deep Learning Technique 5/8

➢ Experiment

Xiangguo Sun et al. Self-supervised Hypergraph Representation Learning for Sociological Analysis. IEEE Transactions on Knowledge and Data Engineering (TKDE) 2023



(a) Social Equavelence.



(b) Social Conformity.

those closely connected pairs are usually located in similar environments. users far away from each other usually have different environments.

most members in a found hyperedge intend to be similar w.r.t their personality traits, suggesting the natural social conformity during the model learning.

➢ Experiment

Xiangguo Sun et al. Self-supervised Hypergraph Representation Learning for Sociological Analysis. IEEE Transactions on Knowledge and Data Engineering (TKDE) 2023

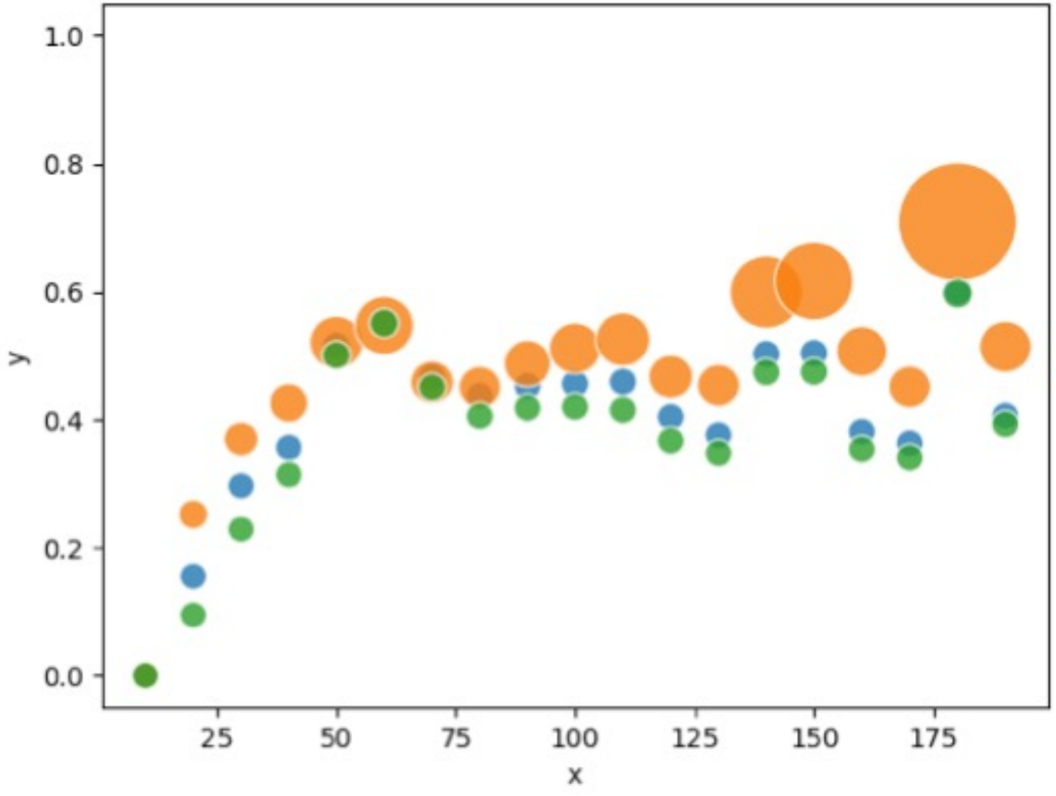
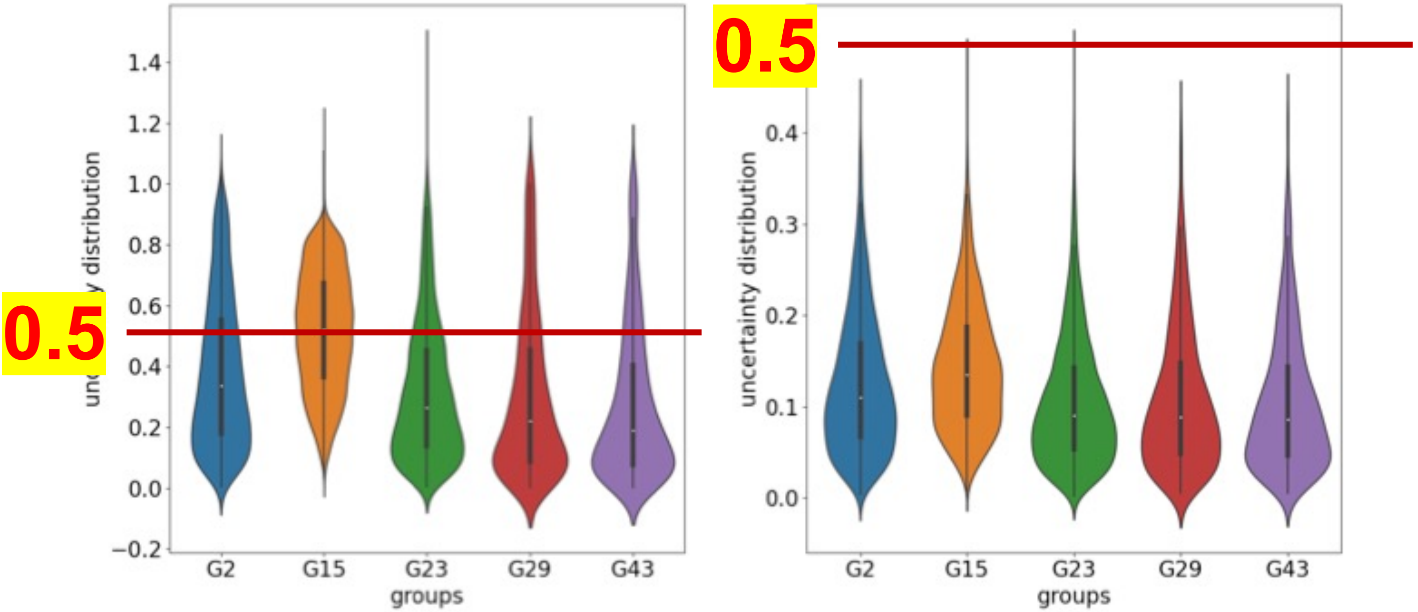


Fig. 11: Simulated Analysis on Group Evolving.

Sociological Analysis via Deep Learning Technique 7/8

- **Experiment**

Xiangguo Sun et al. Self-supervised Hypergraph Representation Learning for Sociological Analysis. IEEE Transactions on Knowledge and Data Engineering (TKDE) 2023



(a) Initial Stage (b) Updated Stage

Fig. 12: Social Polarization: An Entropy Perspective.

TABLE 6: Comparison of Group Entropy

	G2	G15	G23	G29	G43
Initial	963	2820	896	222	218
Updated	443	1166	404	110	113

Sociological Analysis via Deep Learning Technique 8/8

➢ Experiment

Xiangguo Sun et al. Self-supervised Hypergraph Representation Learning for Sociological Analysis. IEEE Transactions on Knowledge and Data Engineering (TKDE) 2023

TABLE 2: Personality Detection (30% labeled)

	Accuracy	Macro F1	Macro Precision	Macro Recall	Micro F1	Micro Precision	Micro Recall
metapath2vec	0.1856	0.7450	0.7155	0.8315	0.7645	0.7106	0.8272
Confluence	0.2201	0.7567	0.7362	0.8003	0.7573	0.7435	0.8066
OCCFRP	0.2366	0.7786	0.7603	0.8102	0.8025	0.7701	0.8238
NCF	0.2515	0.7805	0.7764	0.8035	0.7897	0.7805	0.8105
DeepInf	**0.2937**	0.7703	**0.7903**	0.7685	0.7854	**0.8006**	0.7764
SIDM	0.2891	0.7819	0.7607	0.8127	0.7974	0.7706	0.8260
Our (pluggable)	0.2593	**0.8169**	0.7141	**0.9902**	**0.8244**	0.7197	**0.9920**
Our (unpluggable)	**0.4555**	**0.8337**	**0.8064**	**0.8695**	**0.8464**	**0.8171**	**0.8837**
(bottom line)	(0.0625)	(0.5751)	(0.6885)	(0.4999)	(0.5792)	(0.6885)	(0.4999)

value: top 2 results.

TABLE 3: Rating Prediction

	MAE$\downarrow$	MSE$\downarrow$	R2$\uparrow$	RMSE$\downarrow$	Max Error$\downarrow$
metapath2vec	0.2476	0.0920	0.0889	0.3033	0.9986
Confluence	0.3079	0.1306	0.0065	0.3614	0.9988
OCCFRP	0.2866	0.1049	0.0926	0.3239	0.9255
NCF	0.2376	0.0901	0.1106	0.3002	0.8972
DeepInf	0.2235	0.0860	0.1303	0.2933	**0.8865**
SIDM	0.2011	0.0794	0.1326	0.2818	0.8990
Our (pluggable)	**0.1457**	**0.0371**	**0.2392**	**0.1925**	**0.8681**
Our (unpluggable)	**0.1397**	**0.0347**	**0.2870**	**0.1863**	0.8920

$\uparrow$: higher better; $\downarrow$: lower better; **value**: top 2 results.

TABLE 4: Following Relation Awarenesss

	AUC	AP	MRR	Hits@1	Hits@5
metapath2vec	**0.9322**	0.2122	0.4869	0.2385	0.7588
Confluence	0.8902	0.2164	0.4991	0.2897	0.7536
OCCFRP	0.8864	0.2201	0.5115	0.3006	0.7644
NCF	0.9037	0.2334	0.5564	**0.3537**	0.7967
DeepInf	0.9106	0.2606	**0.5632**	0.3162	0.8036
SIDM	0.9237	0.2564	0.5537	0.3384	**0.8364**
Our (pluggable)	**0.9707**	**0.2678**	0.5576	0.3508	**0.8589**
Our (unpluggable)	0.9297	**0.4158**	**0.6198**	**0.4896**	0.7805

value: top 2 results.

All in One: Multi-Task Prompting for Graph Neural Networks

Xiangguo Sun, Hong Cheng, Jia Li, Bo Liu, Jihong Guan

香港中文大學
The Chinese University of Hong Kong

香港科技大學(廣州)
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)

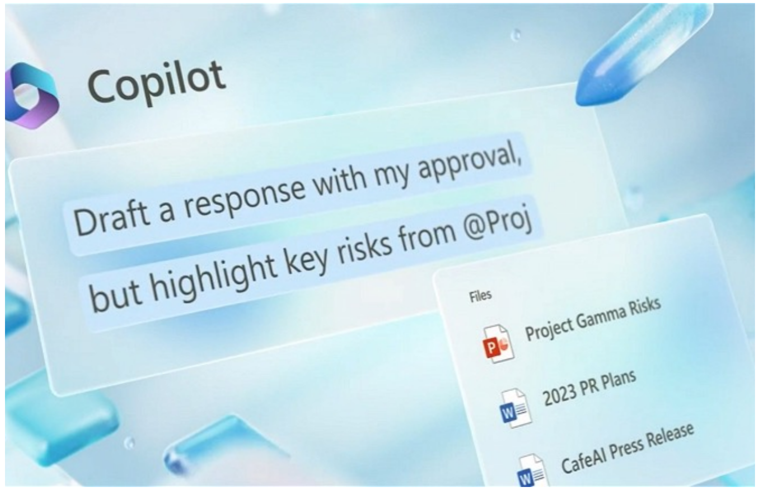
東南大學
SOUTHEAST UNIVERSITY

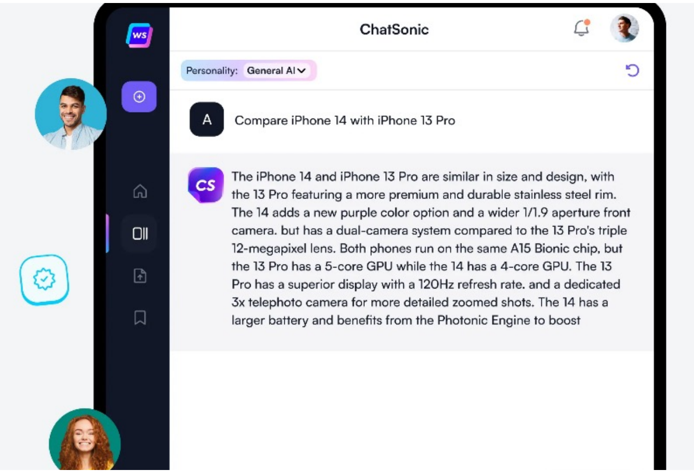
同濟大學
TONGJI UNIVERSITY

KDD2023
AUGUST 6-10
29TH ACM SIGKDD CONFERENCE ON KNOWLEDGE DISCOVERY & DATA MINING

Motivation

- **Prompting large language models (LLM) results in various fantastic achievements.**
  - e.g. ChatGPT, PaLM, LLAMA2 etc.
- **Prompting improves the generalization of LLMs.**





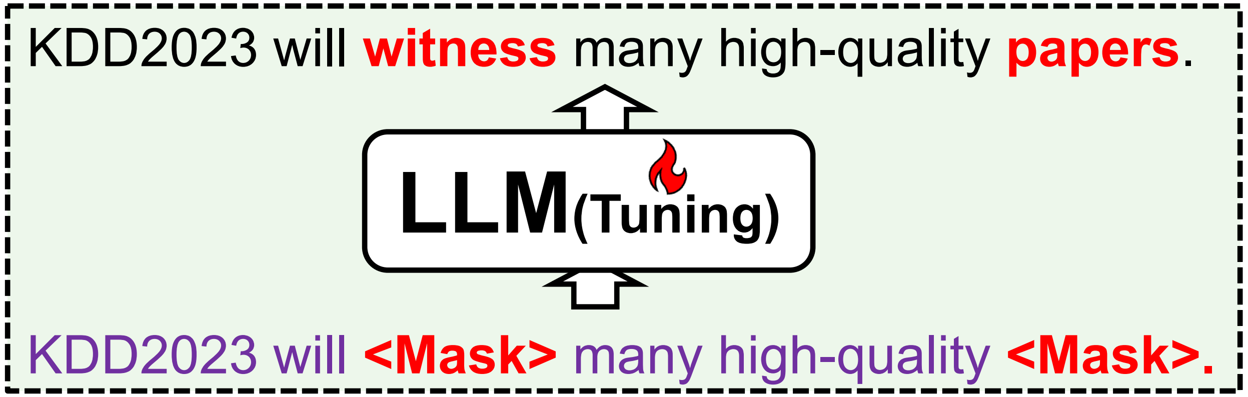
PaLM 2

And More…

ChatGPT

Motivation

- **Prompting a pre-trained LLM**
  - Step1: Pre-training a large language model (LLM).

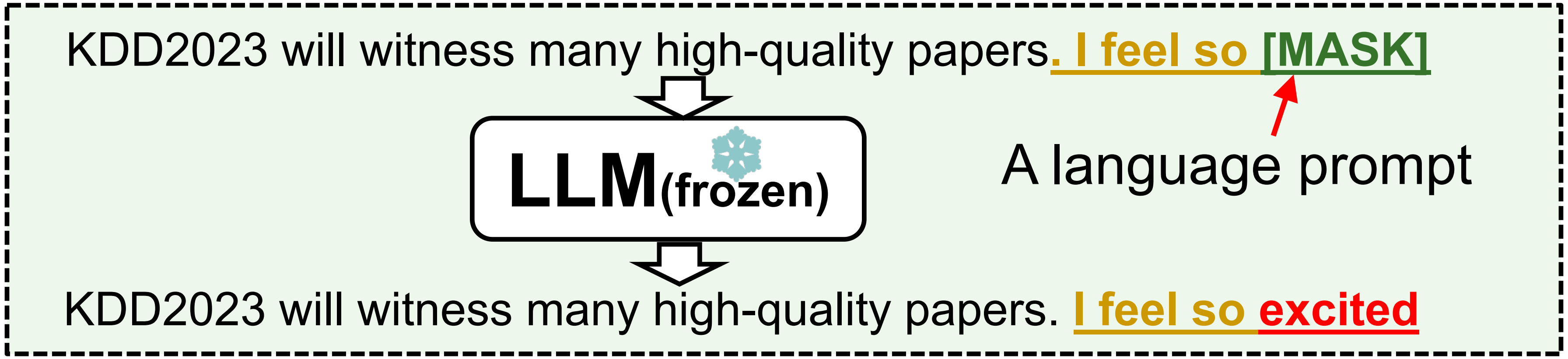


A typical pre-training task for LLM:
Masked Language Modeling (MLM).

Motivation

Prompting a pre-trained LLM

- Step2: Reformulating downstream tasks to the pre-training task by a prompt.

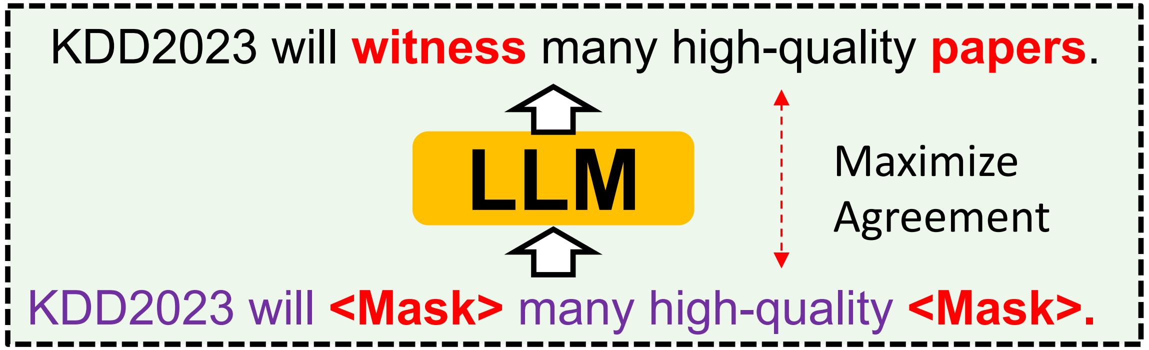


Prompt reformulates the sentiment analysis task to the MLM task.

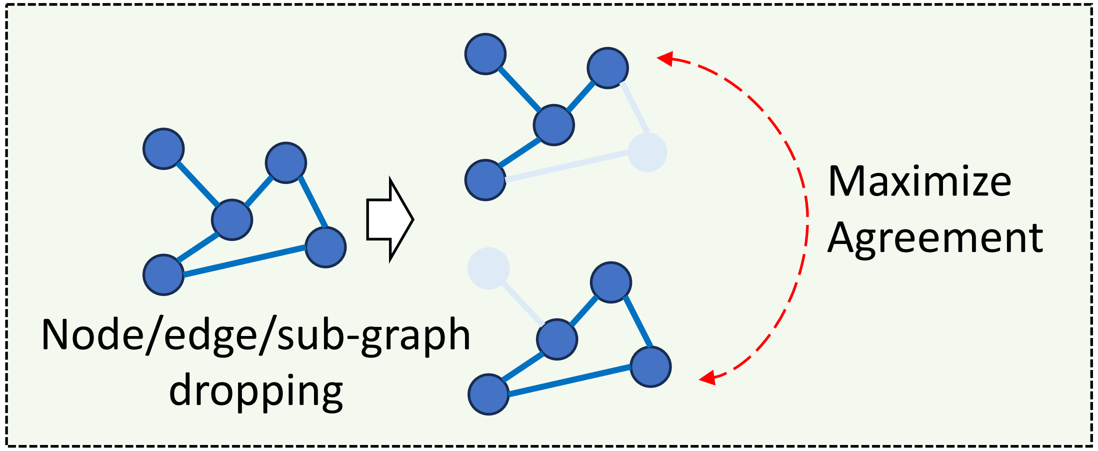
Motivation

➢ Similar insights between LLM and GNN pre-training

- Inspired by NLP Prompt, we wish to **introduce the similar prompt technique in graph domains to improve the generalization of GNNs.**



Aligning two graph views is very similar to predicting some vacant “masks” on graphs.



Pre-training graph models by contrastive learning.

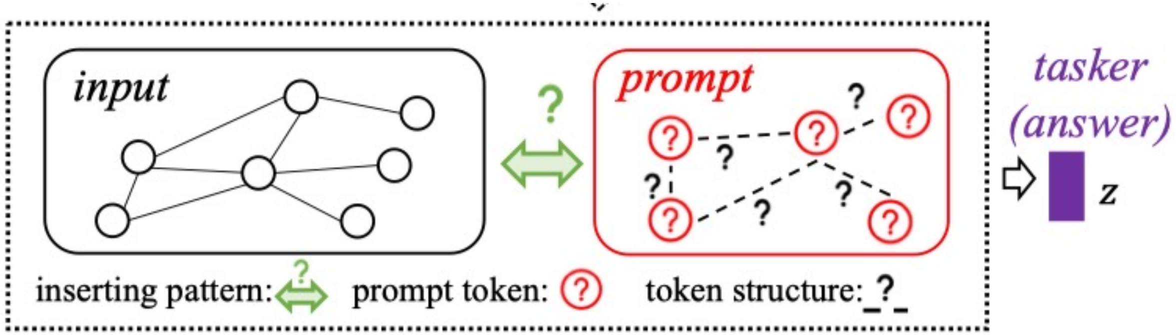
Challenge 1

- **Designing the graph prompt is more intractable than language prompts**
  - NLP prompts are usually some preset tokens, whereas the graph prompt needs to know how to organize these tokens and how to insert the prompt into the original graph.

KDD2023 will witness many high-quality papers. **I feel so [MASK]**

A language prompt

language prompt

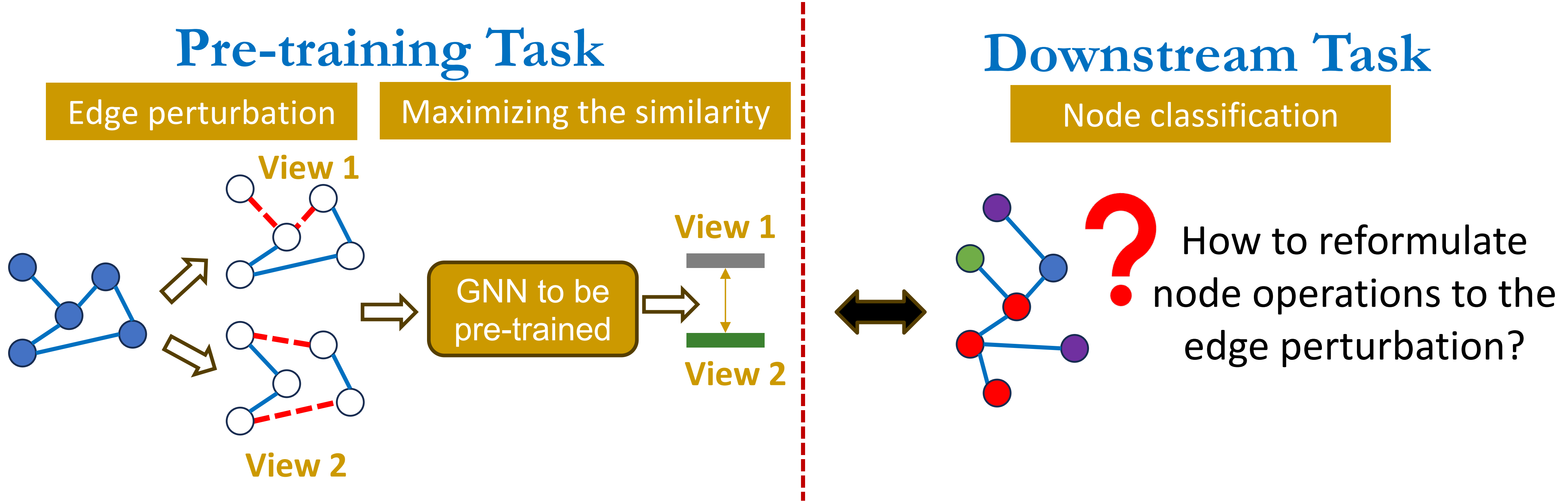


insert the prompt to the input graph

graph prompt

Challenge 2

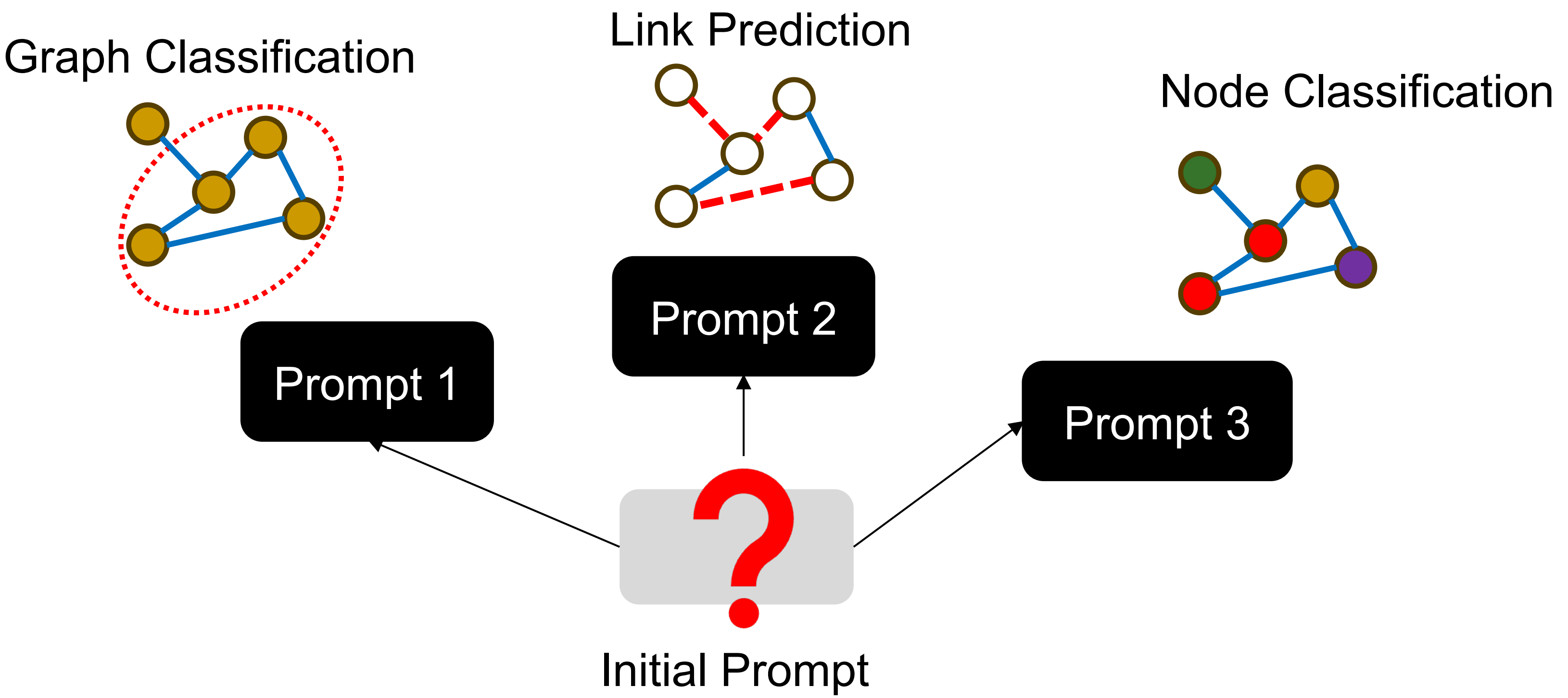
- **Reconciling downstream problems to the pre-training task is more difficult in graph domains**
  - Graph tasks with node level, edge level, and graph level are far diversified.



Challenge 3

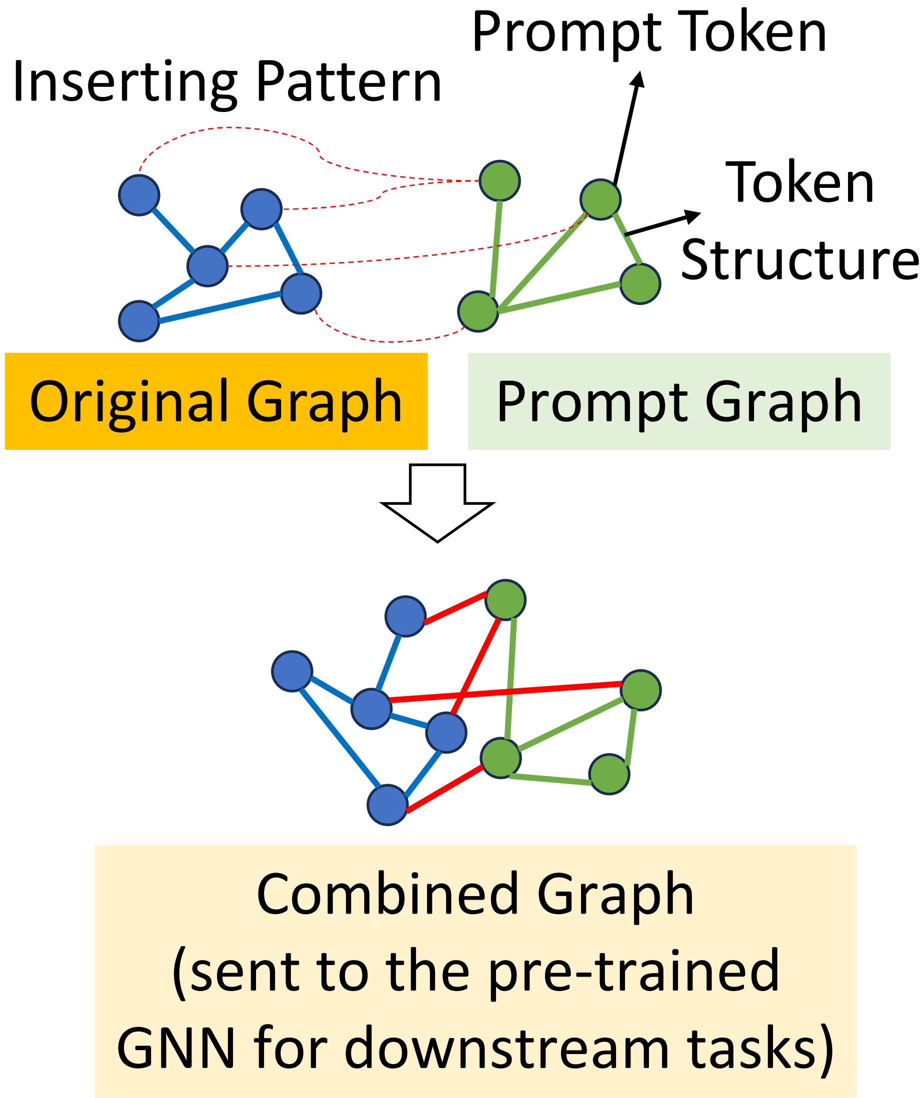
➢ **Learning reliable prompts is more difficult in the multi-task setting**

- Hand-crafted prompts are usually task-bounded, which is far from sufficient for multiple tasks.



Unified Prompt for Graphs

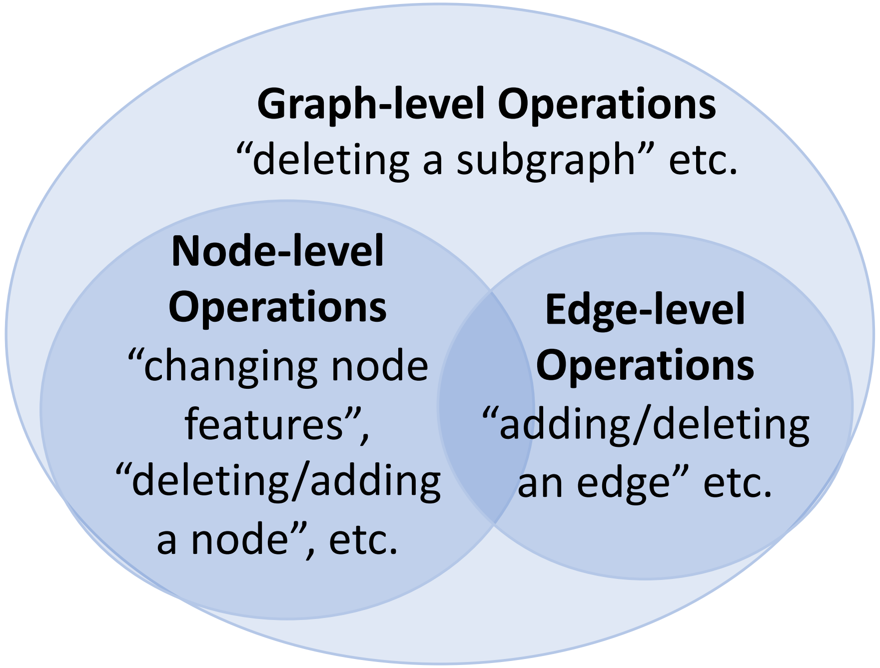
- **Prompt Token**
  - Vectorized information with the same size as node features.
- **Token Structure**
  - Inner connections among different tokens.
- **Inserting Pattern**
  - Cross links between prompt tokens and the original graph.



Reformulating Downstream Tasks

Why reformulating downstream tasks to graph-level tasks?

- Node/edge-level operations can be treated as some special cases at the graph-level operations.
  - e.g, “deleting a subgraph” is the higher-level operation of “deleting nodes and edges”.

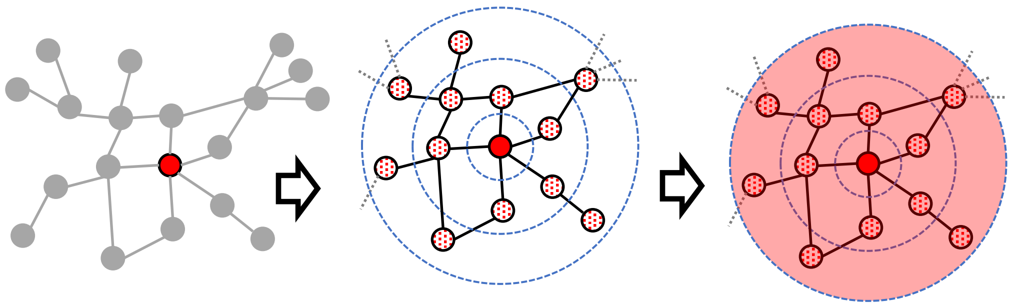


Reformulating Downstream Tasks

KDD2023 AUGUST 6-10 29TH ACM SIGKDD CONFERENCE ON KNOWLEDGE DISCOVERY & DATA MINING

- **Reformulating downstream tasks by induced graphs**
  - Node tasks to graph tasks.
  - Edge tasks to graph tasks.

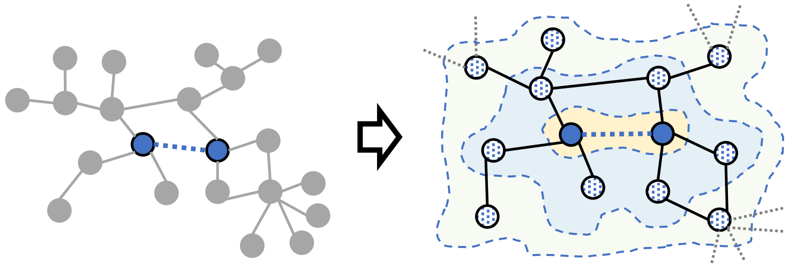
Reformulating node classification to graph classification



Finding a k-hop ego-net for the target node

Assigning the node label to the graph label

Reformulating link prediction to graph classification



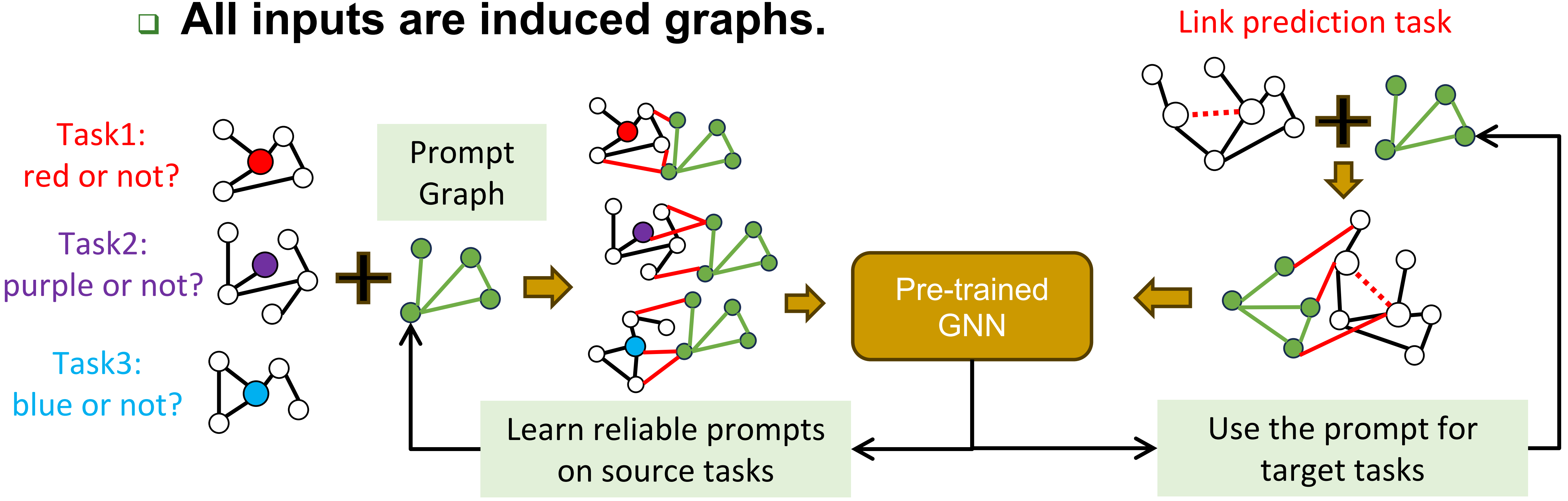
Graph label is positive if the node pair has an edge and vice versa.

Extending a node pair to their k-hop neighbours

Assigning the graph label according to node pair connection

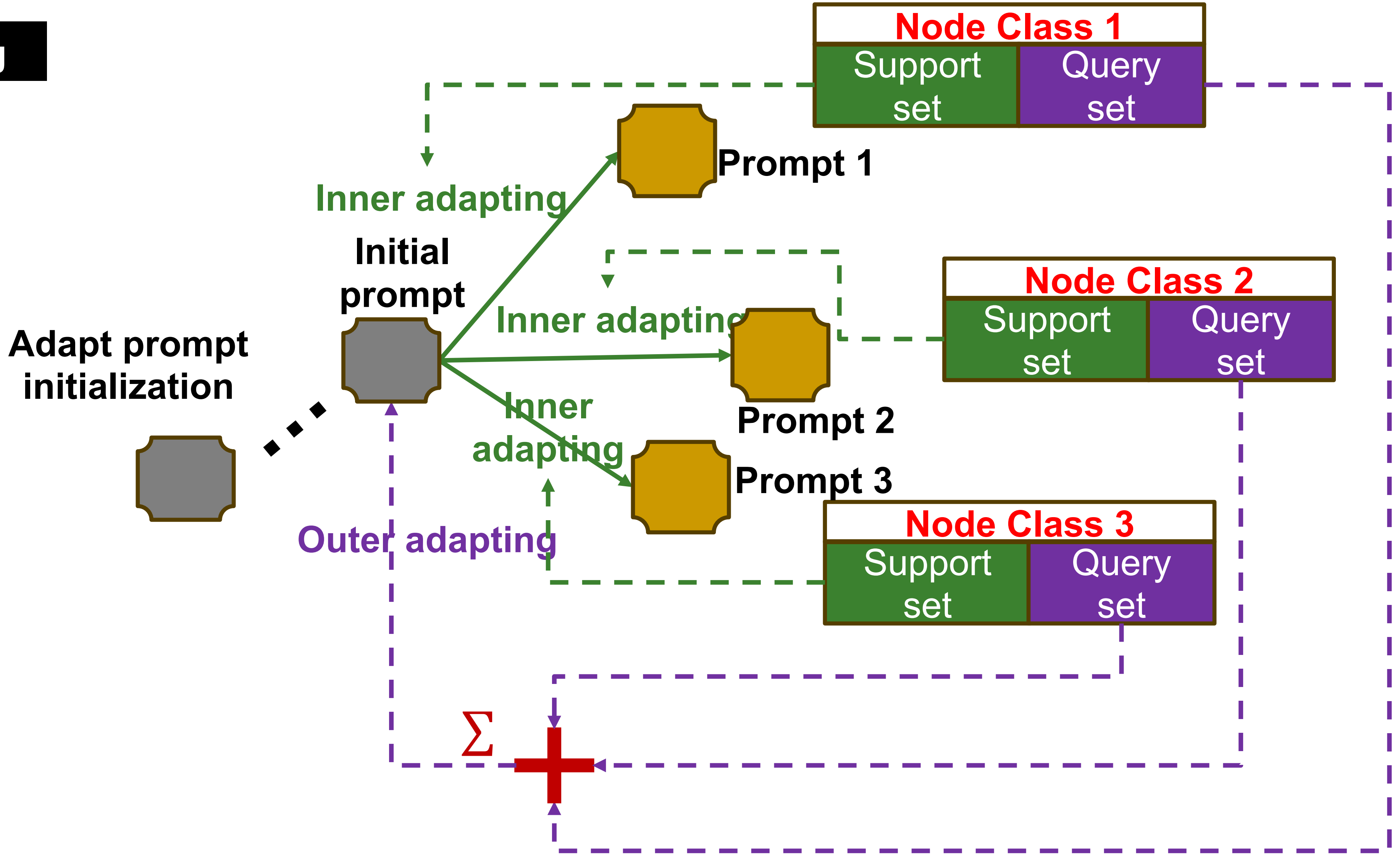
Multi-task Prompting via Meta Learning

- **An Example:**
  - **Target task:** link prediction.
  - **Source tasks:** Node binary classification tasks.
    - Each task corresponds to one node class.
  - **All inputs are induced graphs.**



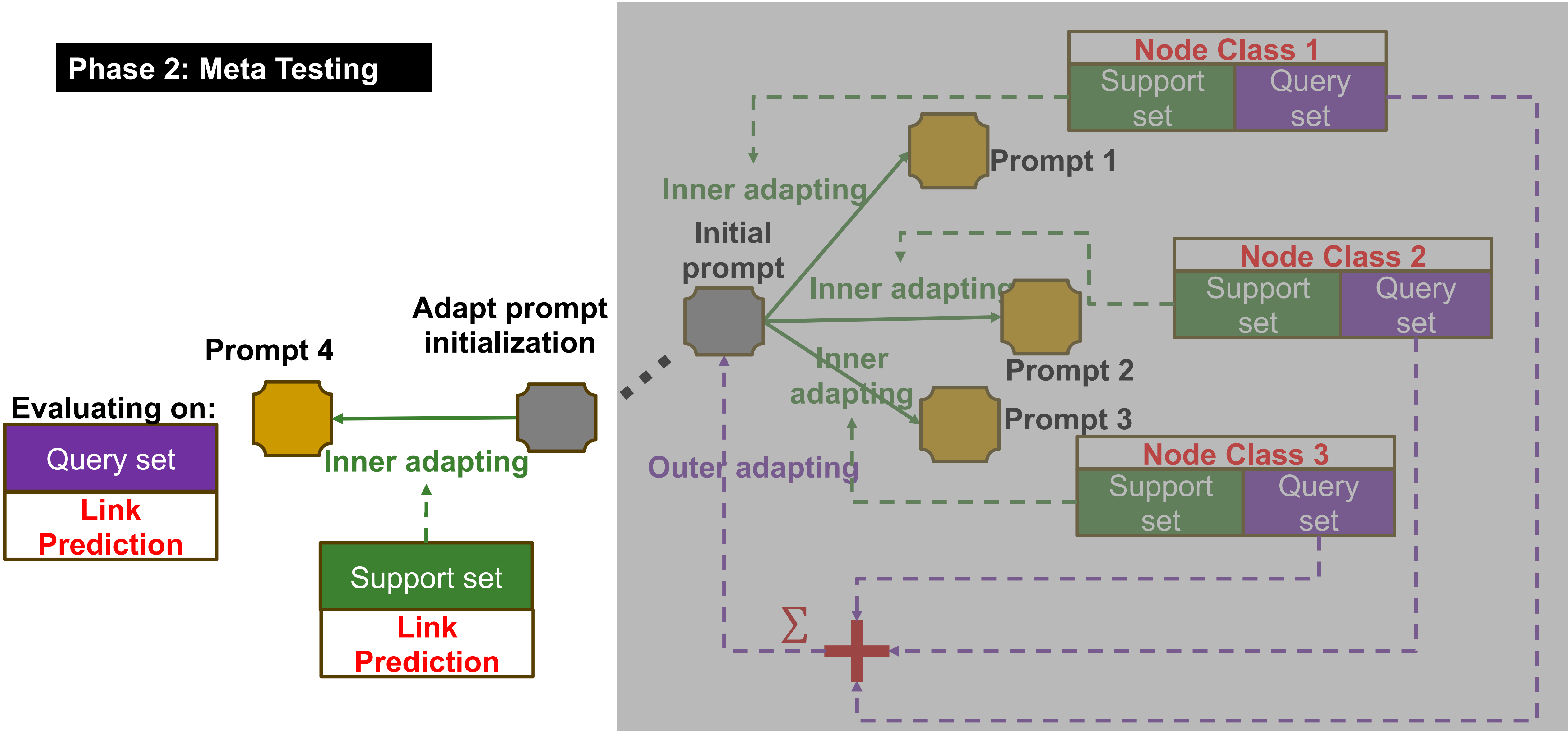
Multi-task Prompting via Meta Learning

Phase 1: Meta Training



Multi-task Prompting via Meta Learning

Phase 2: Meta Testing



Experiments

➢ Multi-Task Performance with Few-shot Learning Settings

Table 2: Node-level performance (%) with 100-shot setting. IMP (%): the average improvement of prompt over the rest.

Node classification

Training schemes	Methods	Cora			CiteSeer			Reddit			Amazon			Pubmed		
		Acc	F1	AUC	Acc	F1	AUC	Acc	F1	AUC	Acc	F1	AUC	Acc	F1	AUC
supervised	GAT	74.45	73.21	82.97	83.00	83.20	89.33	55.64	62.03	65.38	79.00	73.42	97.81	75.00	77.56	79.72
	GCN	77.55	77.45	83.71	88.00	81.79	94.79	54.38	52.47	56.82	95.36	93.99	96.23	53.64	66.67	69.89
	GT	74.25	75.21	82.04	86.33	85.62	90.13	61.50	61.38	65.56	85.50	86.01	93.01	51.50	67.34	71.91
pre-train + fine-tune	GraphCL+GAT	76.05	76.78	81.96	87.64	88.40	89.93	57.37	66.42	67.43	78.67	72.26	95.65	76.03	77.05	80.02
	GraphCL+GCN	78.75	79.13	84.90	87.49	89.36	90.25	55.00	65.52	74.65	96.00	95.92	98.33	69.37	70.00	74.74
	GraphCL+GT	73.80	74.12	82.77	88.50	88.92	91.25	63.50	66.06	68.04	94.39	93.62	96.97	75.00	78.45	75.05
	SimGRACE+GAT	76.85	77.48	83.37	90.50	91.00	91.56	56.59	65.47	67.77	84.50	84.73	89.69	72.50	68.21	81.97
	SimGRACE+GCN	77.20	76.39	83.13	83.50	84.21	93.22	58.00	55.81	56.93	95.00	94.50	98.03	77.50	75.71	87.53
	SimGRACE+GT	77.40	78.11	82.95	87.50	87.05	91.85	66.00	69.95	70.03	79.00	73.42	97.58	70.50	73.30	74.22
prompt	GraphCL+GAT	76.50	77.26	82.99	88.00	90.52	91.82	57.84	67.02	75.33	80.01	75.62	97.96	77.50	78.26	83.02
	GraphCL+GCN	79.20	79.62	85.29	88.50	91.59	91.43	56.00	68.57	78.82	96.50	96.37	98.70	72.50	72.64	79.57
	GraphCL+GT	75.00	76.00	83.36	91.00	91.00	93.29	65.50	66.08	68.86	95.50	95.43	97.56	76.50	79.11	76.00
	SimGRACE+GAT	76.95	78.51	83.55	93.00	93.14	92.44	57.63	66.64	69.43	95.50	95.43	97.56	73.00	74.04	81.89
	SimGRACE+GCN	77.85	76.57	83.79	90.00	89.47	94.87	59.50	55.97	59.46	95.00	95.24	98.42	78.00	78.22	87.66
	SimGRACE+GT	78.75	79.53	85.03	91.00	91.26	95.62	69.50	71.43	70.75	86.00	83.72	98.24	73.00	73.79	76.64
	IMP (%)	1.47	1.94	1.10	3.81	5.25	2.05	3.97	5.04	6.98	4.49	5.84	2.24	8.81	4.55	4.62
	Reported Acc of GPPT (Label Ratio 50%)	77.16	–	–	65.81	–	–	92.13	–	–	86.80	–	–	72.23	–	–
	appr. Label Ratio of our 100-shot setting	∼ 25%			∼ 18%			∼ 1.7%			∼ 7.3%			∼ 1.5%		

Experiments

➢ Efficiency

- Tunable parameters.
- Token numbers.

For more evaluation tasks, please see our paper.

Table 5: Tunable parameters comparison. RED (%): average reduction of the prompt method to others.

Methods	Cora	CiteSeer	Reddit	Amazon	Pubmed	RED (%)
GAT	~ 155K	~ 382K	~ 75K	~ 88K	~ 61K	95.4↓
GCN	~ 154K	~ 381K	~ 75K	~ 88K	~ 61K	95.4↓
GT	~ 615K	~ 1.52M	~ 286K	~ 349K	~ 241K	98.8↓
prompt	~ 7K	~ 19K	~ 3K	~ 4K	~ 3K	–

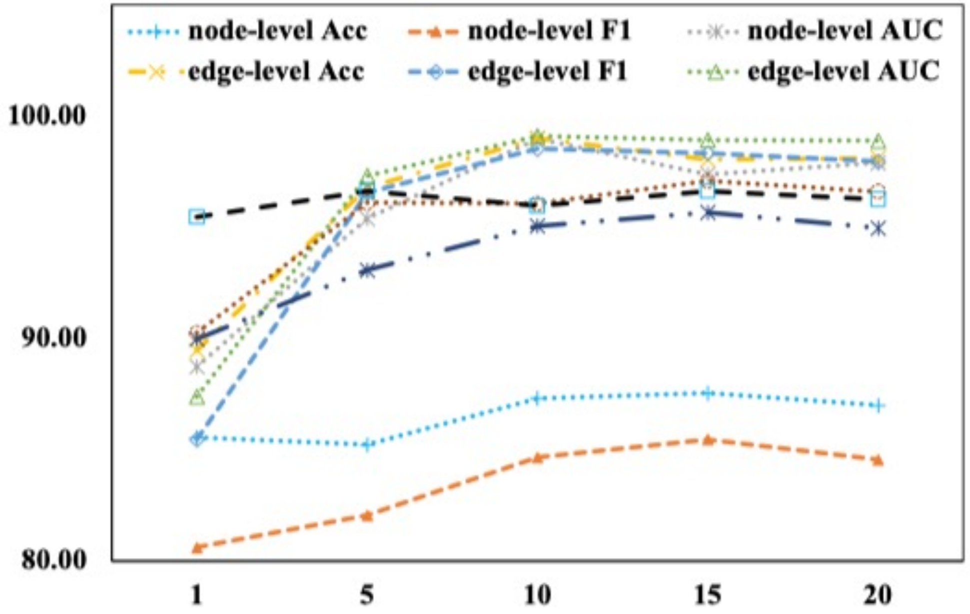
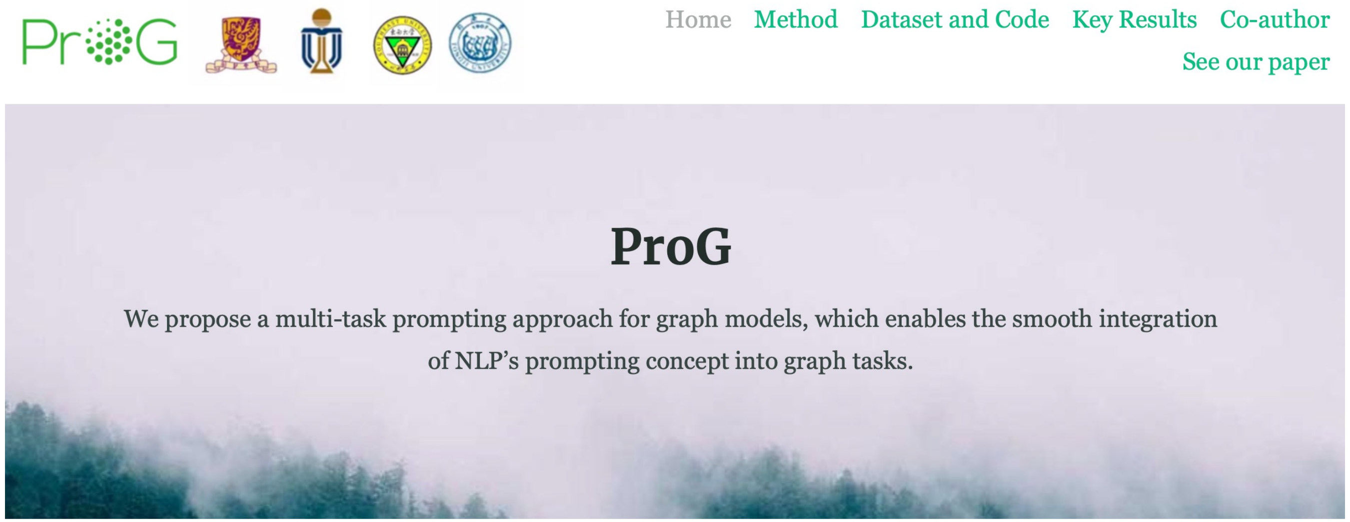


Figure 6: Impact of token numbers

Open Resources

KDD2023
AUGUST 6-10
29TH ACM SIGKDD CONFERENCE ON KNOWLEDGE DISCOVERY & DATA MINING



The website of this paper
https://graphprompt.github.io/

ProG (*Prompt Graph*) is a library built upon PyTorch to easily conduct single or multiple task prompting for a pre-trained GNN.
https://github.com/sheldonresearch/ProG

ProG

Latest version v0.1.1 | docs in progress | pypi package in progress | PyTorch v1.13.1 | license MIT

| Website | Paper | Video | Raw Code |

Summary

- A Brief Introduction to Online Social Networks
- Social Relation Analysis
- Sociological Analysis via Deep Learning Technique
- From LLM to LGM: Prompting graphs like ChatGPT

QA

香港中文大學
The Chinese University of Hong Kong

The End!

Online Social Network Representation: A Brief Introduction and Its Future Direction

Xiangguo Sun

The Chinese University of Hong Kong