Geometer: Graph Few-Shot Class-Incremental Learning via Prototype Representation

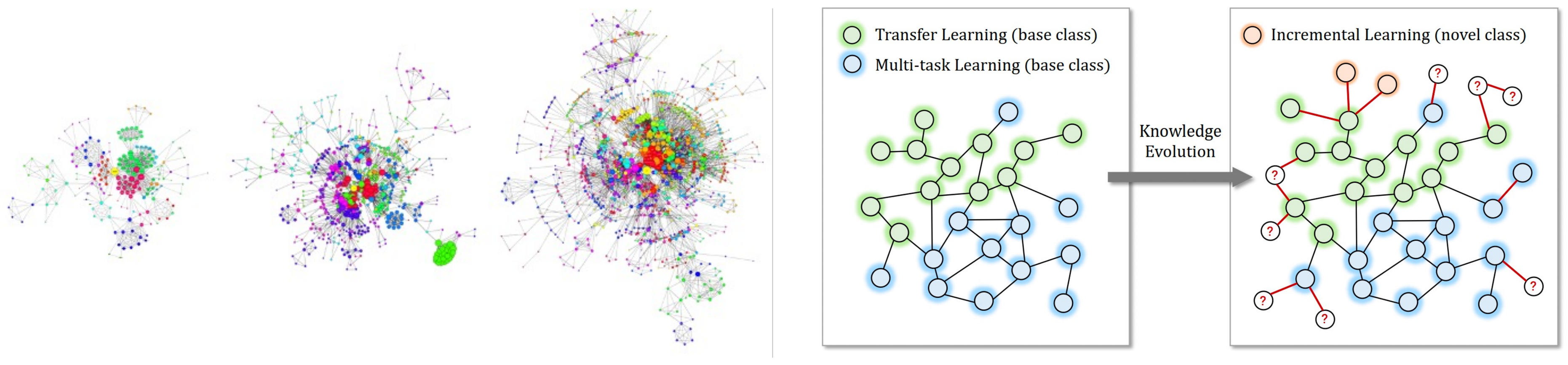
Bin Lu, Xiaoying Gan, Lina Yang, Weinan Zhang, Luoyi Fu, Xinbing Wang

Shanghai Jiao Tong University Shanghai, China

Background & Introduction

What is "Graph Few-Shot Class-Incremental Learning" problem?

- Graph evolves with emergence of new nodes and edges.
- Novel classes appear incrementally along with few labeling.
- How to classify unlabeled nodes into base class or novel class?



Challenges:

- Q1: How to find a way out of "forgetting old"?
- Q2: How to overcome unbalanced labeling between base classes and novel classes?
- Q3: How do we capture the dynamic structure as the network evolves?

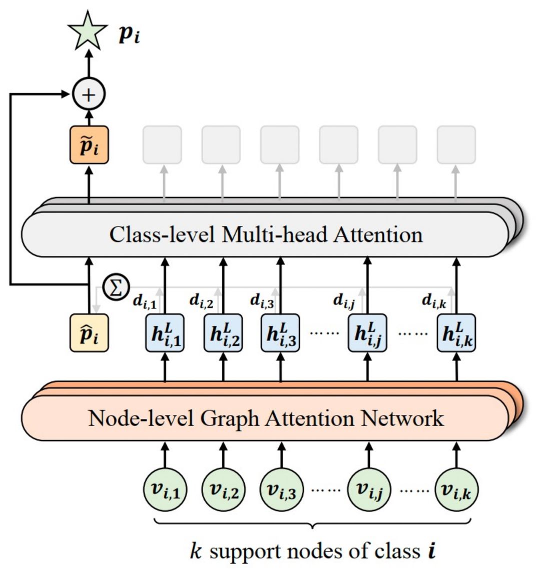
Methodology

Problem Definition

- Base stage $\mathcal{G}^{\text{base}}$ and T snapshots of $\mathcal{G}^{\text{stream}} = \{\mathcal{G}^1, \cdots, \mathcal{G}^T\}$
- Sets of classes $\{C^{\text{base}}, C^1, \cdots, C^T\}$, $C^t = C^{\text{base}} + \sum_i \Delta C^i$
- ΔC^t-way K-shot GFSCIL problem

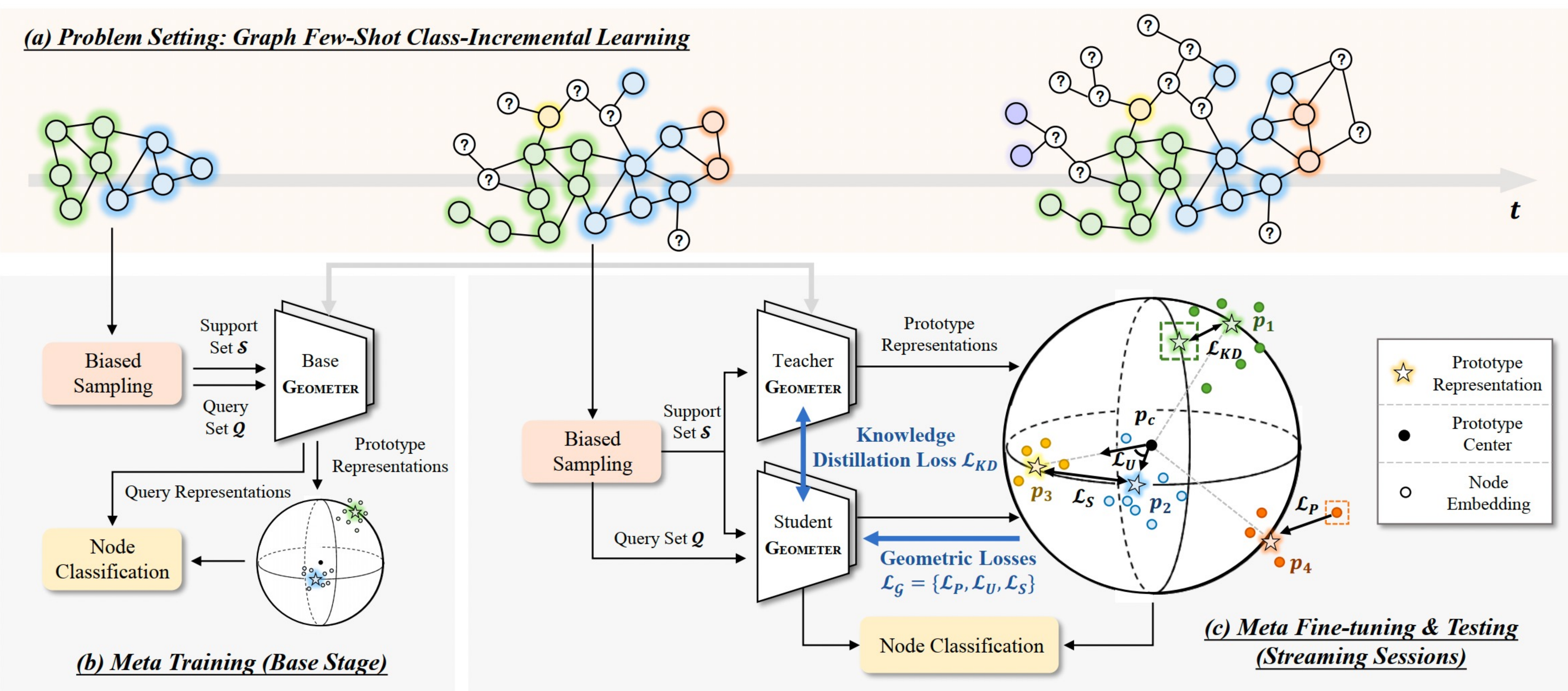
Attention-based Prototype Representation

- Node-level graph attention network
- Class-level multi-head attention



Geometric Metric Learning

- Intra-Class Proximity $\mathcal{L}_P$: Nodes of same classes should be closely clustered
- Inter-Class Uniformity $\mathcal{L}_U$: Uniformity of different prototypes in metric space
- Inter-Class Separability $\mathcal{L}_S$: Prototypes of novel classes and old classes should keep a distance



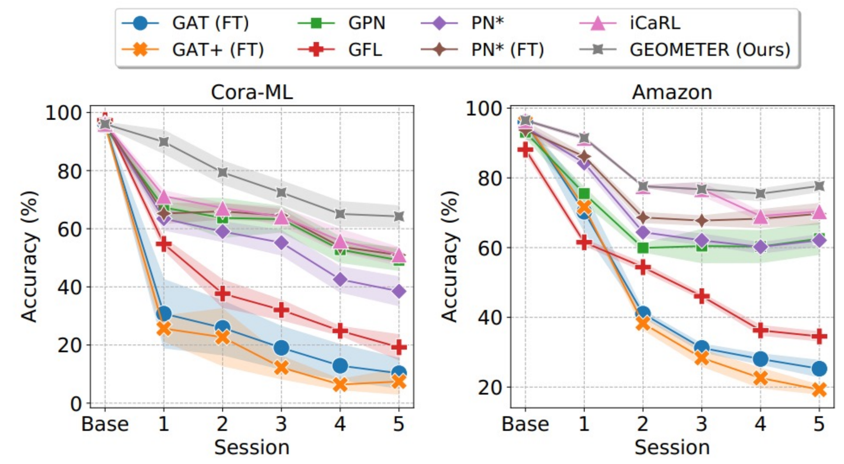
Overview of the proposed Geometer for Graph Few-Shot Class-Incremental Learning.

Experiment

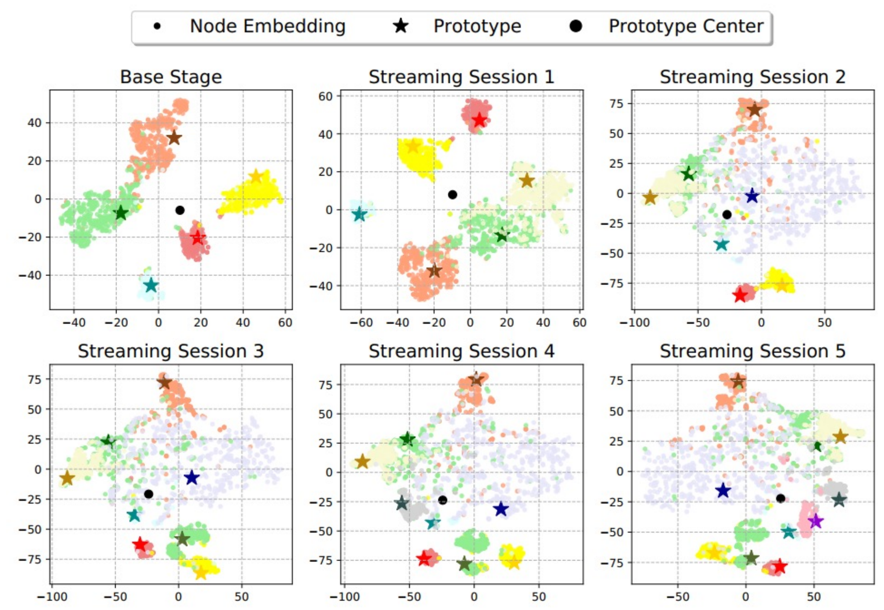
Dataset

Dataset	Field	Nodes	Edges	Features	Class
Cora-ML	Academic	2,995	16,316	2,879	7
Flickr	Social Network	7,575	479,476	12,047	9
Amazon	E-commerce	13,752	491,722	767	10
Cora-Full	Academic	19,793	126,842	8,710	70

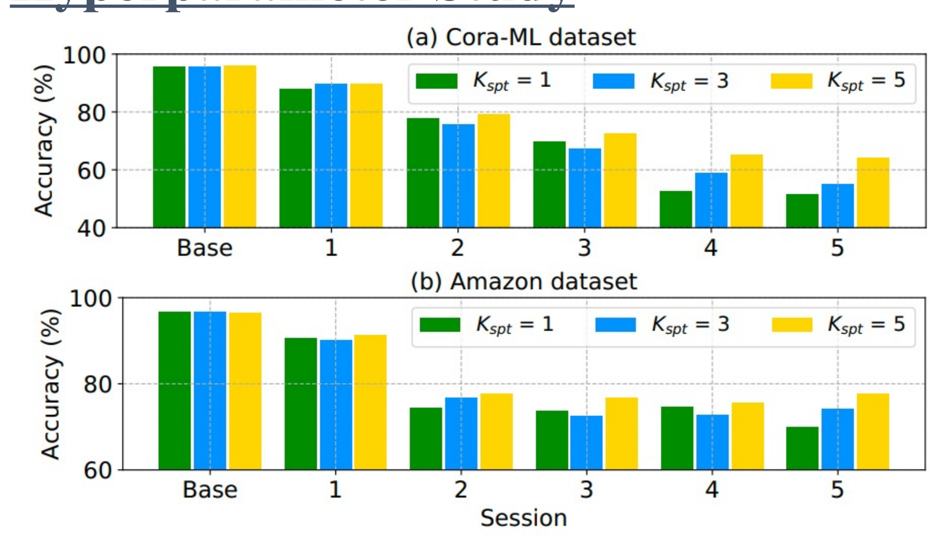
Performance Comparison



Case Study: t-SNE visualization of the query node embeddings and prototypes



Hyperparameter Study

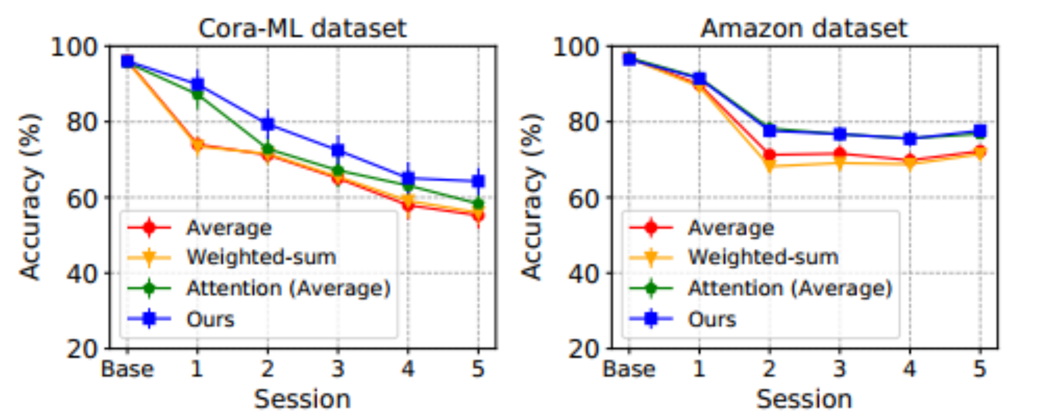


Ablation Study: Loss functions

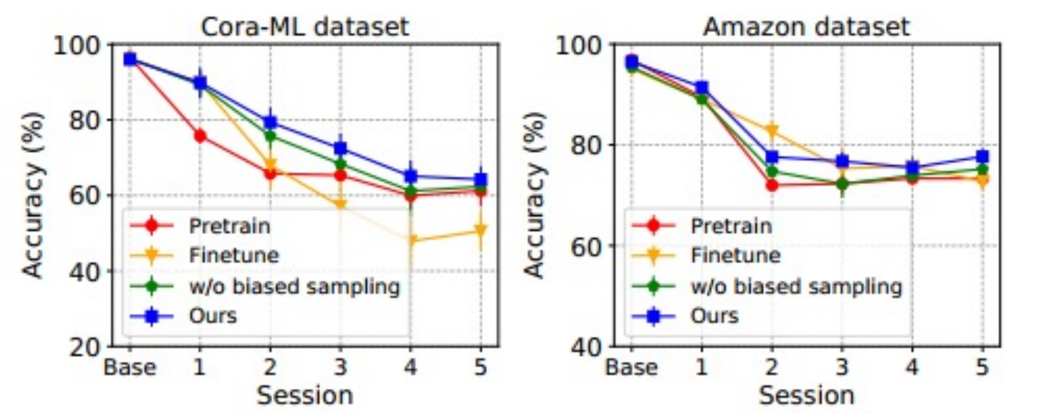
$\mathcal{L}_P$	$\mathcal{L}_U$	$\mathcal{L}_S$	$\mathcal{L}_{KD}$	Cora-ML (1-way 5-shot GFSCIL setting) Base Classes	Session 1	Session 3	Session 5
✓			✓	**96.21±0.67%**	88.25±3.99%	64.89±2.53%	56.21±5.55%
✓	✓		✓	95.85±0.56%	**90.41±3.86%**	68.26±3.65%	58.72±4.66%
✓		✓	✓	95.71±0.55%	89.21±2.88%	69.57±2.71%	54.28±3.90%
✓	✓	✓		95.74±0.61%	90.40±4.82%	68.16±1.45%	62.41±2.37%
✓	✓	✓	✓	96.01±0.92%	89.89±3.97%	**72.45±4.01%**	**64.25±3.60%**

$\mathcal{L}_P$	$\mathcal{L}_U$	$\mathcal{L}_S$	$\mathcal{L}_{KD}$	Amazon (1-way 5-shot GFSCIL setting) Base Classes	Session 1	Session 3	Session 5
✓			✓	96.72±0.28%	90.91±0.59%	74.74±2.33%	73.73±3.01%
✓	✓		✓	96.72±0.22%	91.39±0.56%	76.55±1.94%	73.97±1.90%
✓		✓	✓	96.83±0.32%	91.15±0.35%	75.08±2.61%	73.92±2.60%
✓	✓	✓		**96.86±0.35%**	91.17±0.38%	75.36±1.28%	74.51±2.53%
✓	✓	✓	✓	96.50±0.29%	**91.44±0.46%**	**76.74±1.89%**	**77.66±1.58%**

Ablation Study: Prototype representation



Ablation Study: Biased sampling strategy



Conclusion & Future Work

- We are the first to investigate this novel problem: graph few-shot class-incremental learning (GFSCIL).
- With the novel classes popping up, Geometer learns and adjusts the attention-based prototypes based on the geometric relationships of proximity, uniformity and separability of representations.
- Geometer proposes teacher-student knowledge distillation and biased sampling strategy to further mitigate the catastrophic forgetting and unbalanced labeling.

Contact Us

- Bin Lu: robinlu1209@sjtu.edu.cn
- Intelligent IoT Research Center: http://iiot.sjtu.edu.cn/