Association for the Advancement of Artificial Intelligence

GraphMoRE: Mitigating Topological Heterogeneity via Mixture of Riemannian Experts

Zihao Guo, **Qingyun Sun***, **Haonan Yuan**, **Xingcheng Fu**, **Min Zhou**, **Yisen Gao**, **Jianxin Li**

Email: guozh@act.buaa.edu.cn

北京航空航天大學 BEIHANG UNIVERSITY

廣西師範大學 GUANGXI NORMAL UNIVERSITY

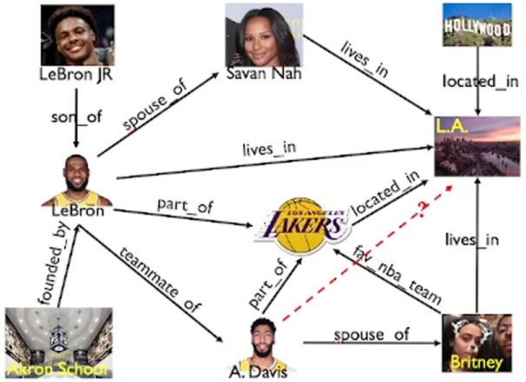
HUAWEI

Paper Code

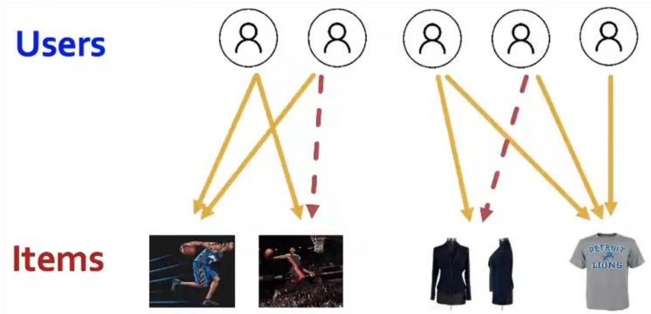
■ Graphs are everywhere in the Real-world

Social network

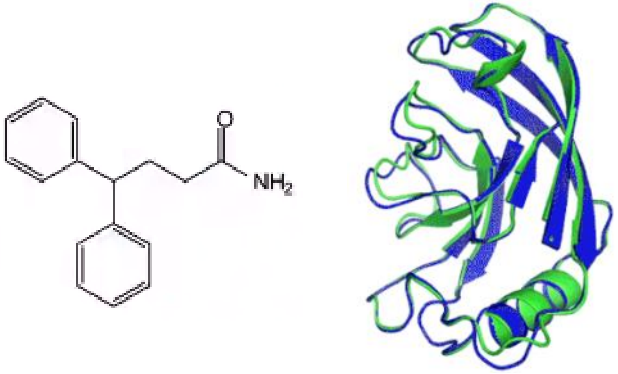
Traffic network



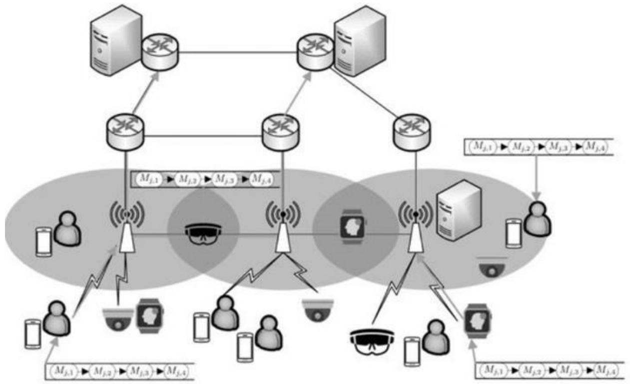
Knowledge graph



Co-purchase network

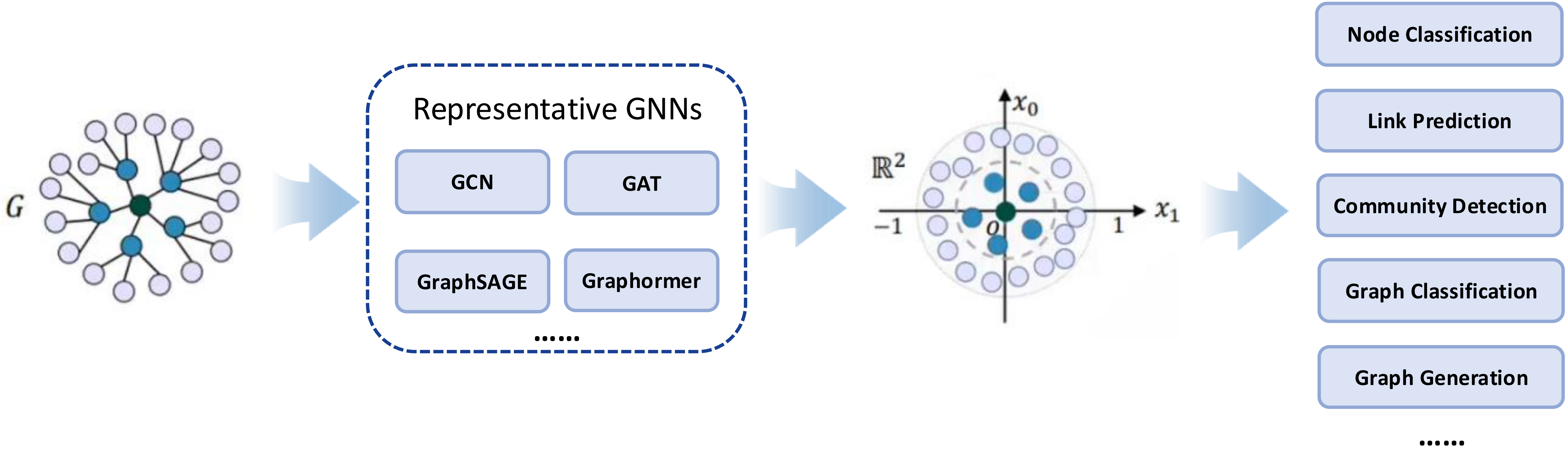


Chemical molecules



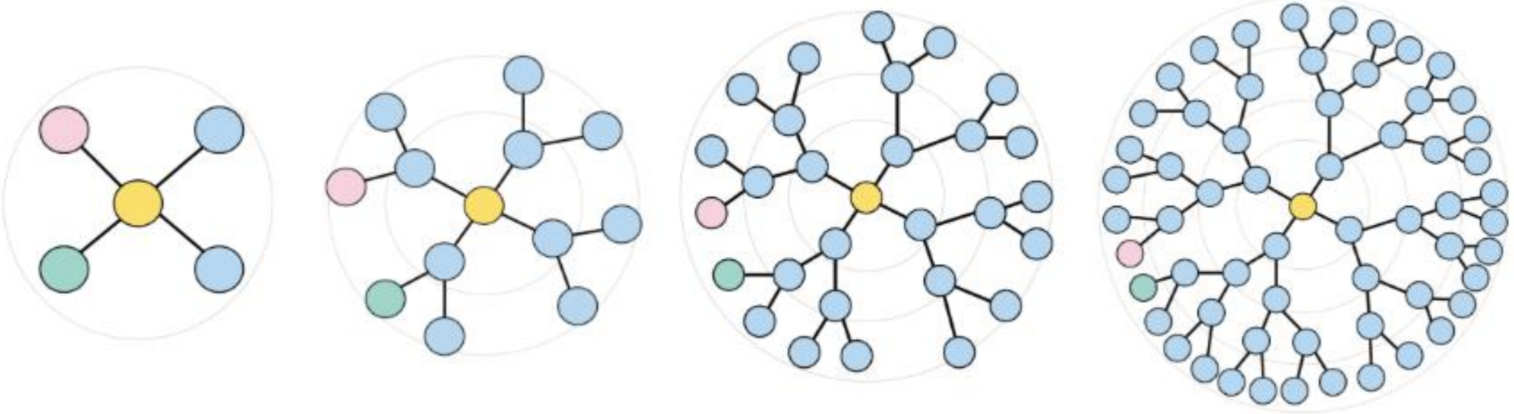
Communication network

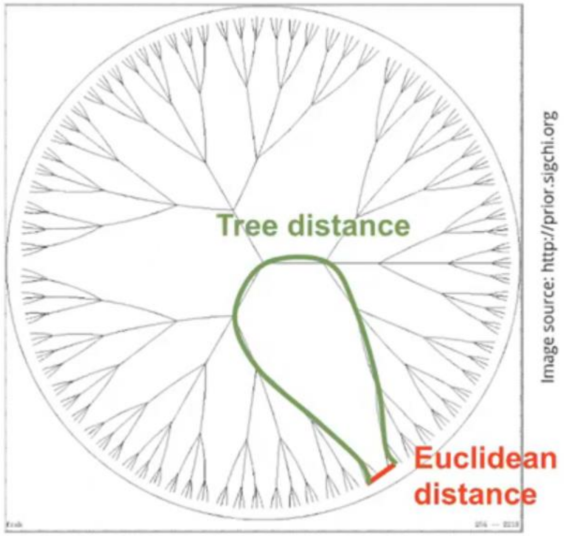
Graph Learning in Euclidean Space



Most existing graph representation learning methods **embed graphs into Euclidean space**. However, ...

■ Distortion in Euclidean Space



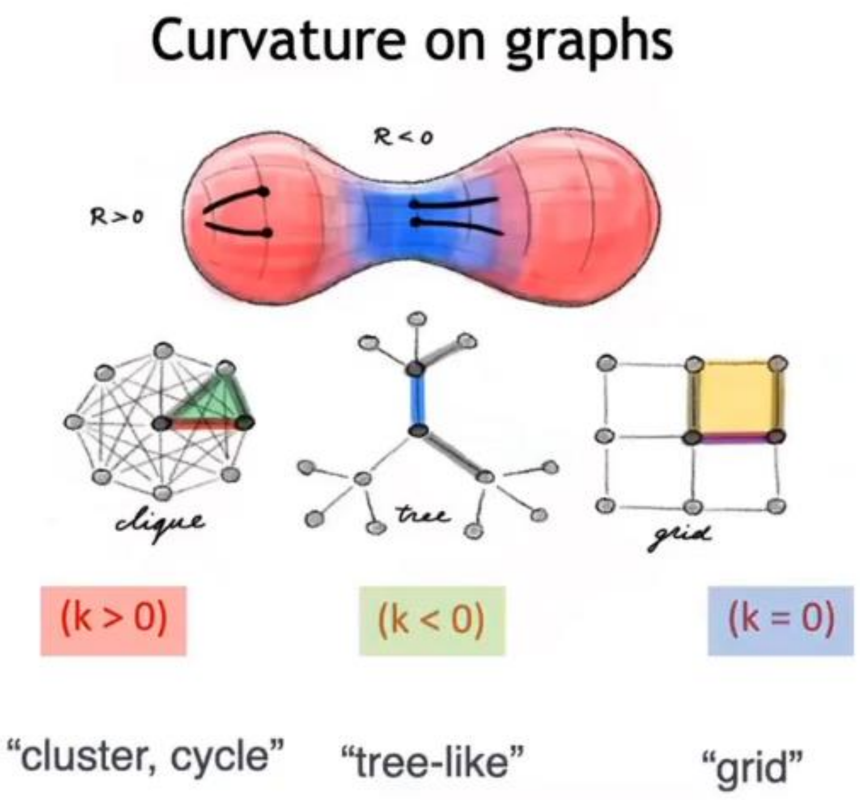


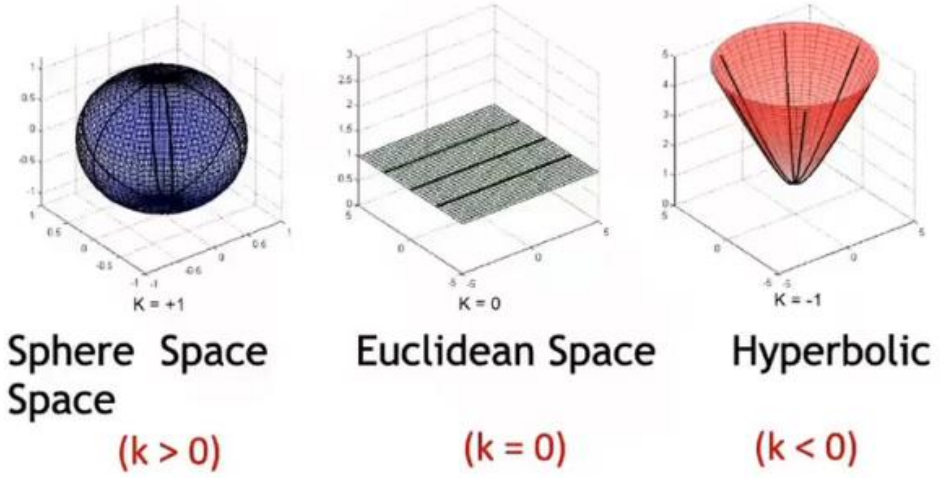
A tree: the number of nodes grows **exponentially** with the tree depth!

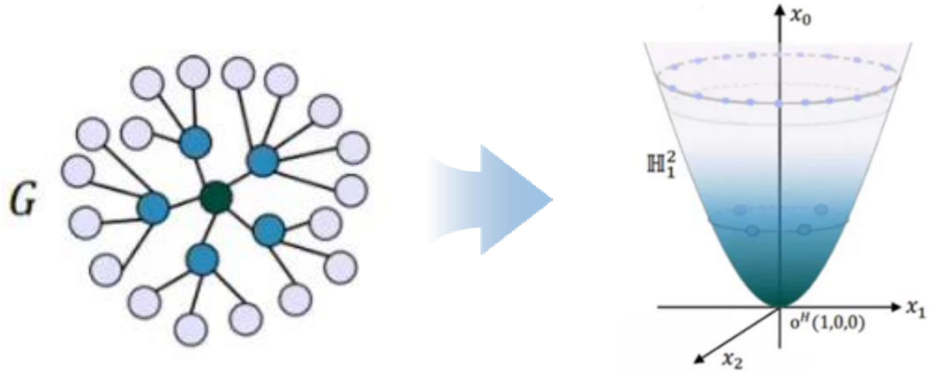
However, embedding in Euclidean space causes **serious distortion for certain structures.**

[1] Constant curvature graph convolutional networks. Bachmann, B´ecigneul, and Ganea. ICML 2020.

Riemannian Manifold and Constant Curvature Spaces

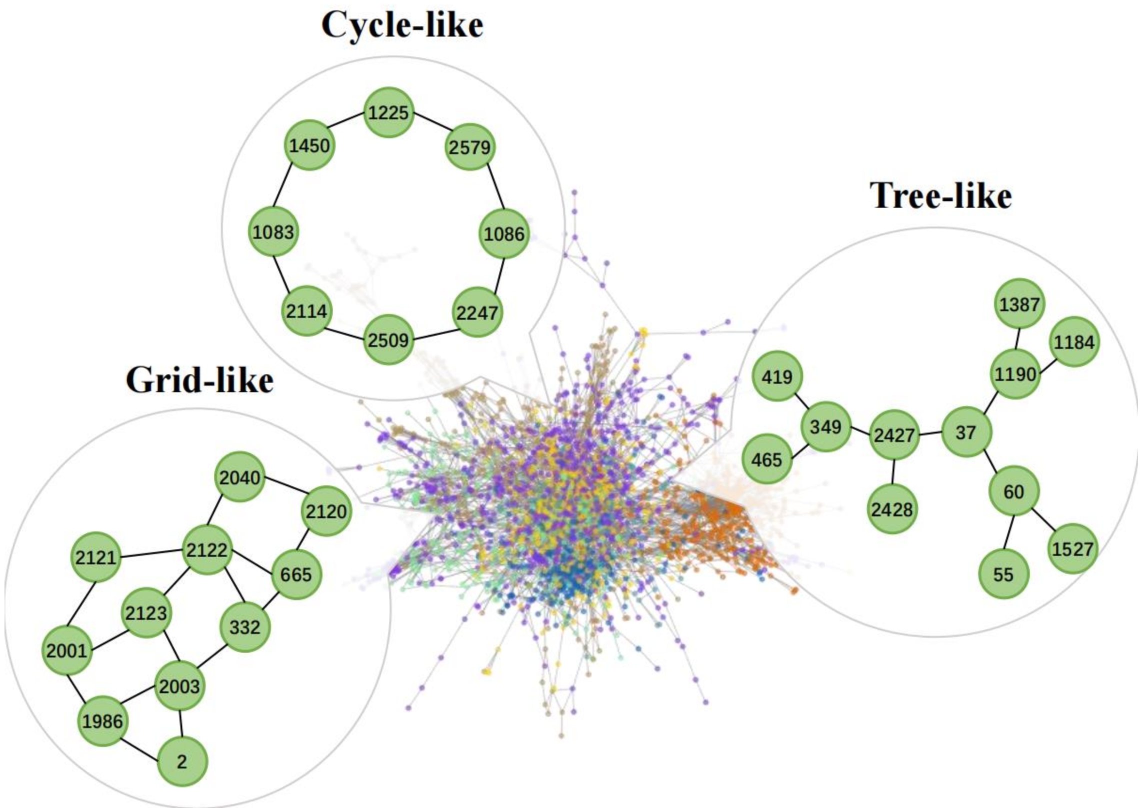






Riemannian representation learning has the ability to naturally represent different topological structures.

■ Topological Heterogeneity in Graphs

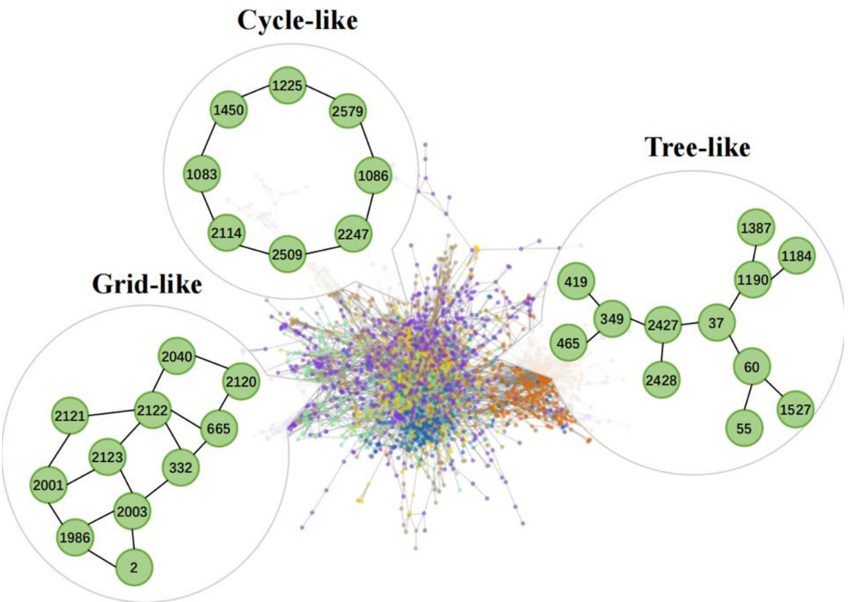


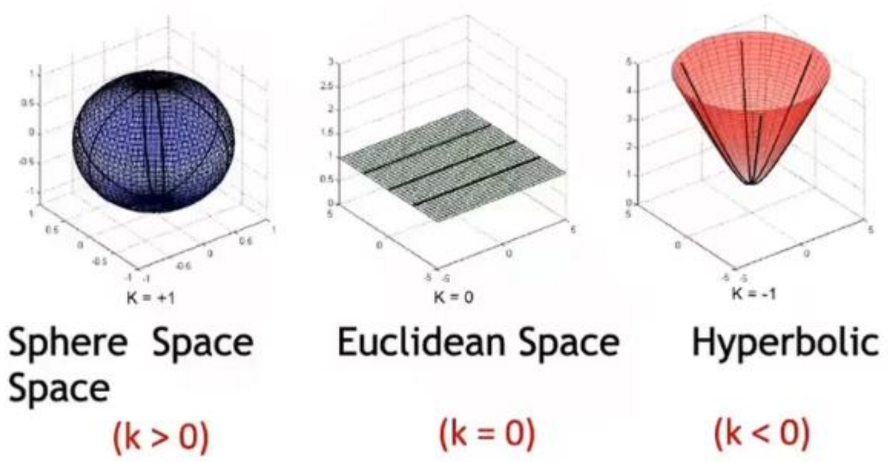
- ➢ **Single constant curvature space**
  - insufficient to adapt to topological heterogeneity
- ➢ **Mixed-curvature space (i.e. product manifold)**
  - still homogeneous with the globally uniform curvature [2]
- ➢ **Some other methods**
  - depend on the global curvature to some extent
- ➢ **Our method**
  - directly constructs personalized heterogeneous manifold for different geometric properties

[2] Heterogeneous manifolds for curvature-aware graph embedding. Di Giovanni, Luise, and Bronstein. ICLR 2022 Workshop on GTRL.

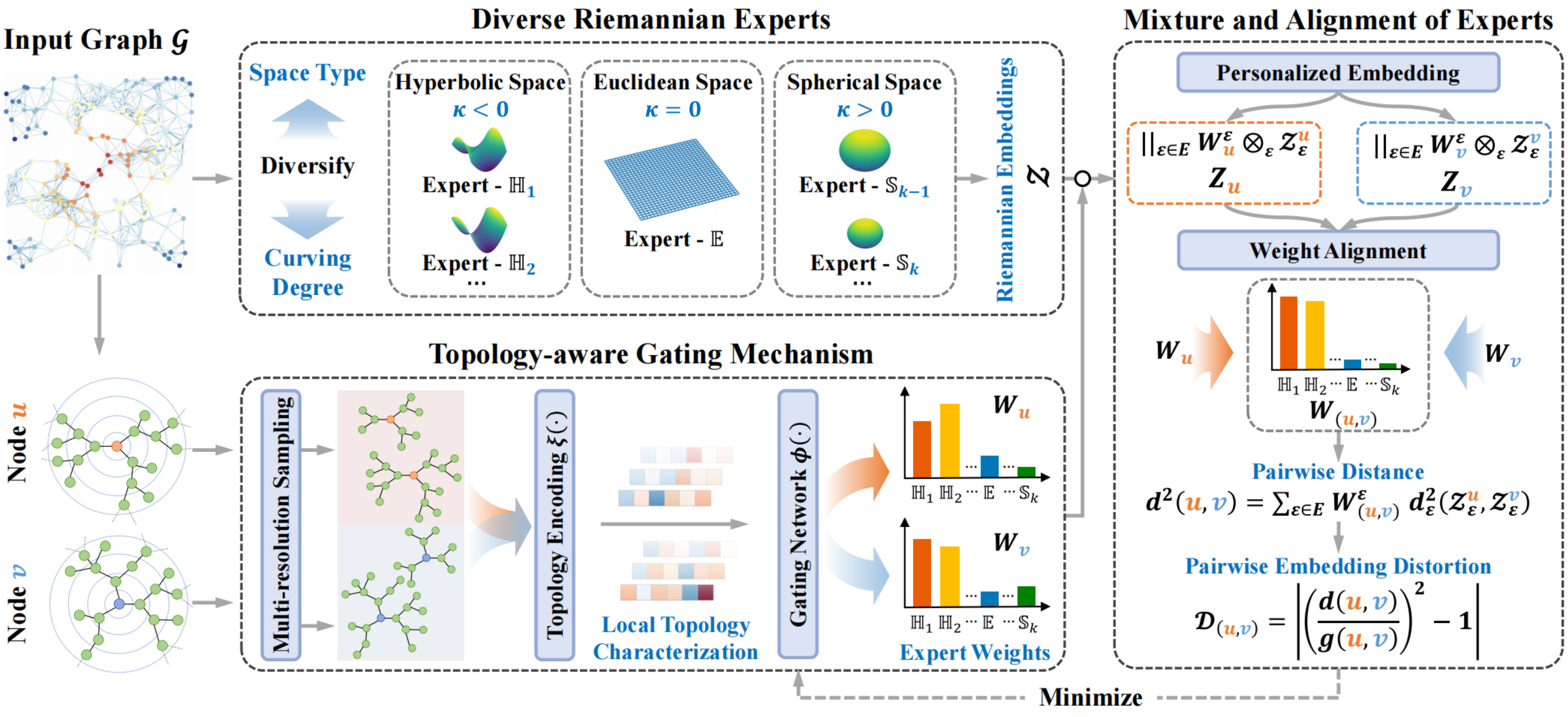
Major Challenges

- **Topology pattern identification issue**
  - The **topological patterns of substructures** are often **not explicitly expressed and vary with resolution.**
- **Optimal embedding space selection issue**
  - The **appropriate type of Riemannian space and the optimal curvature vary from different topological properties**. Moreover, the **substructure is still complex** and difficult to classify as a single type.

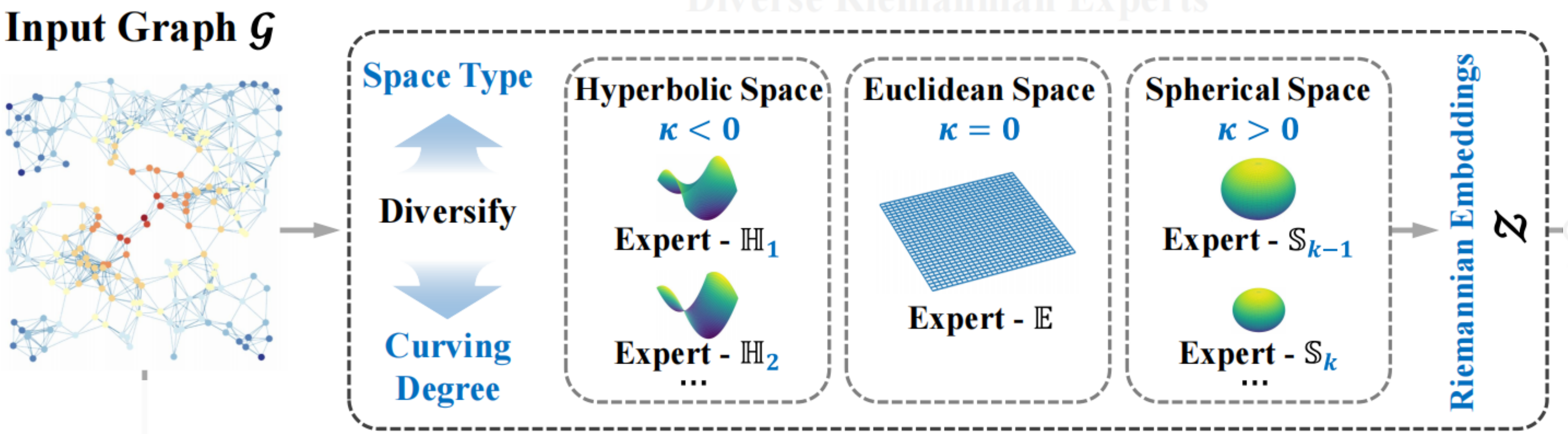




Framework of GraphMoRE (Graph Mixture of Riemannian Experts)



■ Framework of GraphMoRE

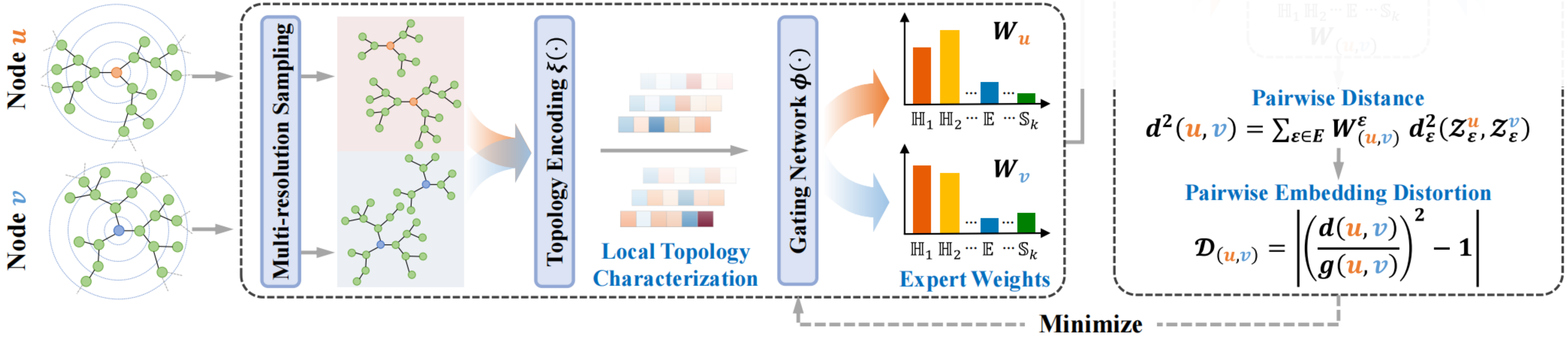


- **Diverse Riemannian Experts**
  - **introduce MoE to model the complex geometric properties**
  - **the experts can naturally correspond to different curvature spaces** and construct personalized mixed curvature spaces for different nodes
  - **consider the diversity** of experts in terms of both **the type of curvature spaces and the degree of curving in space**

Framework of GraphMoRE

➢ Topology-aware Gating Mechanism

- from the **perspective of local topology characterization**
- **estimating the geometric properties** around each node **through multi-resolution local topology sampling and encoding** $\mathcal{S}_v = \{\text{Sampler}(v, r), r \in \mathcal{R}\}, \quad \mathcal{T}_v = \| \text{Pooling}(\xi(s), s \in \mathcal{S}_v),$
- **the distortion guided gating module** can adaptively **route different nodes to the appropriate embedding space** $\mathbf{W}_v = \text{Softmax}(\phi(\mathcal{T}_v)),$



Framework of GraphMoRE

- **Mixture and Alignment of Experts**
  - **fuse the output of Riemannian experts with the expert weights** which are assigned based on the local geometric properties

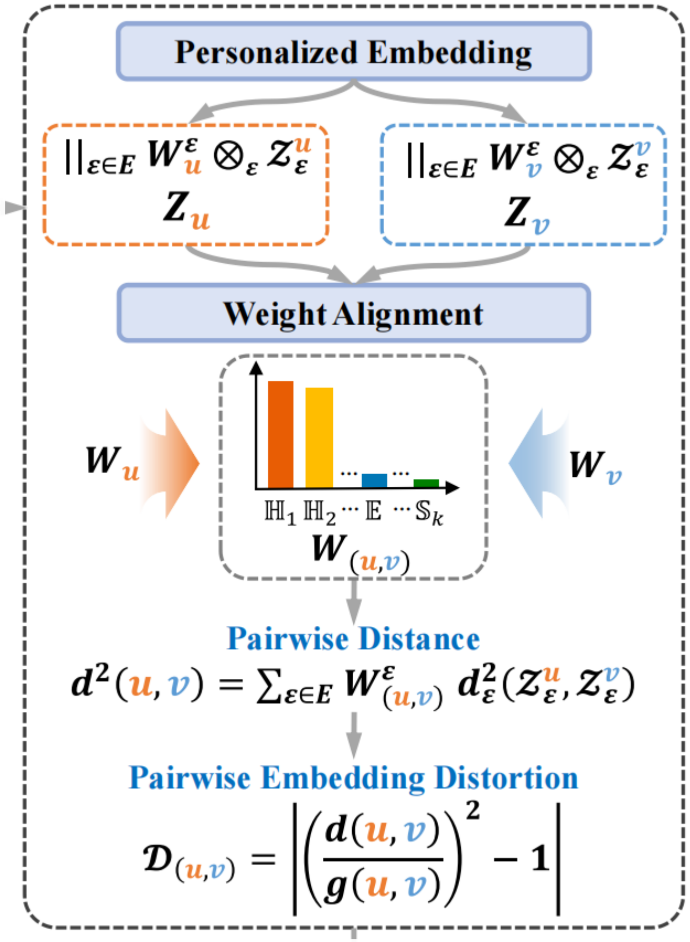
$$\mathbf{Z}_v = \|_{\varepsilon \in E} \mathbf{W}_v^{\varepsilon} \otimes_{\varepsilon} \mathcal{Z}_{\varepsilon}^{v},$$

  - **integrate the expert weights** of nodes for embedding alignment

$$\mathbf{W}_{(u,v)} = \mathrm{Softmax}(\mathbf{W}_u \cdot \mathbf{W}_v),$$

  - **calculate the pairwise distance** in an aligned embedding space

$$d^2(u,v) = \sum_{\varepsilon \in E} \mathbf{W}_{(u,v)}^{\varepsilon} d_{\varepsilon}^2(\mathcal{Z}_{\varepsilon}^{u}, \mathcal{Z}_{\varepsilon}^{v}).$$



Overall Training Pipeline of GraphMoRE

Algorithm 1: The overall training process of GraphMoRE

Input: Graph $\mathcal{G}$; Number of experts K; Initial curvature set C; Number of training epochs T.
Output: Predicted result of the downstream task.

```
Initialize Riemannian experts E with C;
Sample local topology subgraphs S for nodes by Eq. (4);
for t = 1, 2, ..., T do
    for expert ε in E do
        Get the output Z_ε of expert ε by Eq. (1) and (2);
    end
    // Topology-aware Gating Mechanism
    Calculate local topology characterizations 𝒯 by Eq. (5);
    Get the expert weights W for each node by Eq. (6);
    // Mixture and Alignment of Experts
    for each node pair (u, v) in G do
        Get the aligned expert weights W_(u,v) by Eq. (9);
        Calculate embedding distance d^2(u, v) by Eq. (10);
    end
    // Update all parameters
    Update model parameters by minimizing L by Eq. (11);
end
```

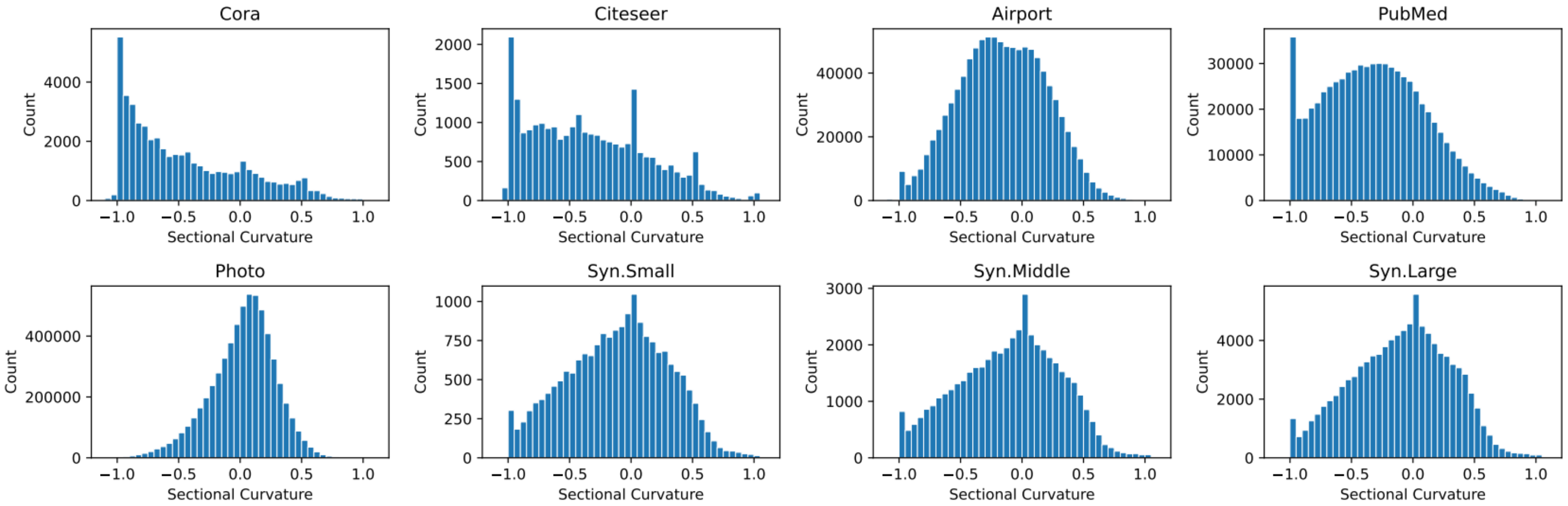
Experimental setup

Datasets

- **Real-world graphs** (Cora, Citeseer, PubMed, Airport, Photo)

Dataset	# Nodes	# Edges	# Features	# Classes
Cora	2,708	5,429	1,433	7
Citeseer	3,327	4,732	3,703	6
PubMed	19,717	44,338	500	3
Airport	3,188	18,631	4	4
Photo	7,650	119,081	745	8

- **Synthetic graphs** (Syn.Small, Syn.Middle, Syn.Large)



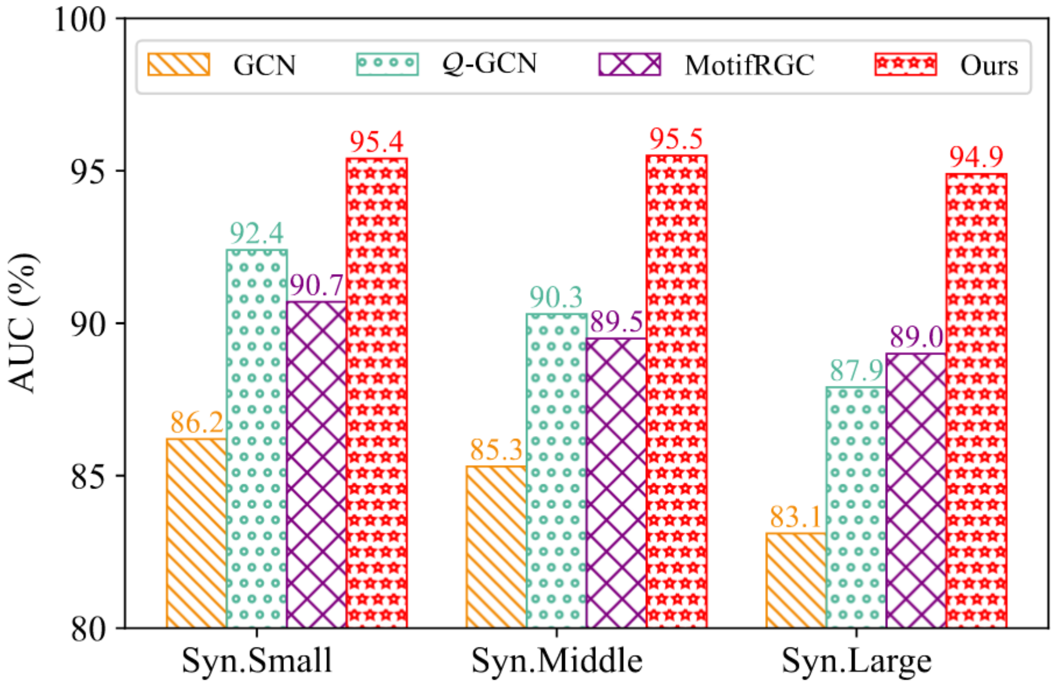
■ Performance on Real-world Graphs

	Method	Cora		Citeseer		Airport		PubMed		Photo	
		AUC	AP	AUC	AP	AUC	AP	AUC	AP	AUC	AP
Euclidean methods	GCN	91.88±0.33	92.44±0.36	90.10±0.65	90.66±0.57	91.95±0.91	91.35±0.69	93.79±2.91	93.49±2.73	92.17±2.57	90.42±2.71
	GAT	90.63±0.30	91.13±0.31	88.83±0.64	88.81±0.18	89.55±1.22	89.25±1.34	88.87±1.57	87.96±1.30	94.06±0.40	93.15±0.50
	SAGE	89.29±0.59	90.03±0.69	86.93±0.53	88.47±0.69	89.04±0.76	85.44±1.36	89.97±0.40	90.76±0.31	92.87±2.28	91.28±2.51
Riemannian methods	HNN	90.85±0.44	89.88±0.49	95.27±0.48	95.41±0.52	92.63±0.42	92.96±0.21	93.04±0.19	92.06±0.14	92.77±0.31	90.12±0.32
	HGCN	93.62±0.25	93.73±0.27	94.87±0.41	95.28±0.37	93.50±0.36	94.13±0.28	95.20±0.15	95.16±0.16	96.96±0.93	96.12±0.94
	κ-GCN	93.81±1.95	94.97±1.55	96.76±1.20	97.39±0.88	93.87±0.34	93.18±0.23	97.17±0.12	97.05±0.13	97.29±0.09	96.72±0.12
	LGCN	93.75±0.30	94.47±0.24	95.50±0.20	95.99±0.14	94.20±0.26	94.43±0.20	96.20±0.11	96.25±0.16	97.50±1.10	96.96±1.20
	$\mathcal{Q}$-GCN	93.66±0.17	92.96±0.12	94.41±0.25	93.93±0.15	94.99±0.33	94.49±0.21	95.21±0.05	95.14±0.11	97.33±0.04	96.53±0.06
	MotifRGC	95.90±0.71	96.51±0.54	96.04±0.39	96.33±0.43	96.86±0.39	96.65±0.28	97.26±0.76	97.20±0.78	OOM	
	GraphMoRE	**97.91±0.10**	**98.42±0.09**	**98.70±0.20**	**98.84±0.14**	**97.53±0.28**	**96.71±0.36**	**99.18±0.05**	**99.22±0.05**	**98.83±0.04**	**98.53±0.06**

Method	Cora		Citeseer		Airport		PubMed		Photo	
	W-F1	M-F1	W-F1	M-F1	W-F1	M-F1	W-F1	M-F1	W-F1	M-F1
GCN	79.41±1.25	78.92±1.04	65.92±2.13	62.46±1.73	81.38±0.82	77.44±0.82	75.95±0.40	75.74±0.34	91.68±0.70	89.77±1.11
GAT	77.49±1.11	76.98±1.01	65.43±1.46	61.87±1.38	83.06±1.02	80.33±1.03	75.33±0.89	75.04±0.84	92.63±0.59	90.96±0.76
SAGE	77.47±0.71	76.61±0.87	64.03±1.11	60.24±1.56	84.07±2.19	81.27±2.17	73.97±0.90	73.90±0.87	88.09±1.56	85.25±1.90
HNN	58.98±0.52	57.15±0.70	59.52±0.51	57.38±0.68	70.71±1.47	54.49±1.81	69.56±0.71	68.85±0.49	90.90±0.42	89.34±0.53
HGCN	77.78±0.63	73.94±0.61	67.51±0.81	60.56±0.83	84.82±1.46	81.23±2.06	78.69±0.59	77.78±0.50	90.78±0.28	88.10±0.42
κ-GCN	78.67±1.00	78.01±0.73	63.88±0.69	60.28±0.60	87.40±1.64	84.33±2.08	77.31±1.71	76.37±1.54	92.31±0.45	90.52±0.76
LGCN	80.19±0.98	79.05±0.81	67.94±1.78	61.41±4.25	89.03±1.24	84.50±2.00	77.19±0.41	76.92±0.44	92.97±0.32	90.53±1.23
$\mathcal{Q}$-GCN	75.97±0.97	74.03±0.96	68.89±1.70	62.68±1.48	89.66±0.67	85.46±0.86	**80.61±0.81**	**79.55±0.91**	91.85±0.47	90.30±0.63
MotifRGC	80.91±0.64	80.19±0.63	68.31±1.41	64.63±1.39	84.53±2.32	83.96±2.38	79.13±1.37	77.71±1.34	OOM	
GraphMoRE$_{\text{GCN}}$	80.51±0.91	79.44±0.77	**69.73±0.70**	**65.98±0.67**	91.75±0.93	91.49±0.98	77.16±0.61	76.50±0.76	93.18±0.31	91.86±0.34
GraphMoRE$_{\text{GAT}}$	79.81±0.71	78.53±0.85	68.59±1.64	64.70±1.74	91.06±1.52	90.50±1.78	77.30±0.74	76.46±0.75	93.57±0.29	92.20±0.33
GraphMoRE$_{\text{SAGE}}$	**81.42±0.68**	**80.32±0.56**	69.40±0.82	65.71±1.12	**92.32±0.70**	**91.33±0.46**	77.30±0.81	76.53±0.73	**94.33±0.37**	**92.99±0.49**

➤ GraphMoRE achieves **superior performance** among 9 baselines.

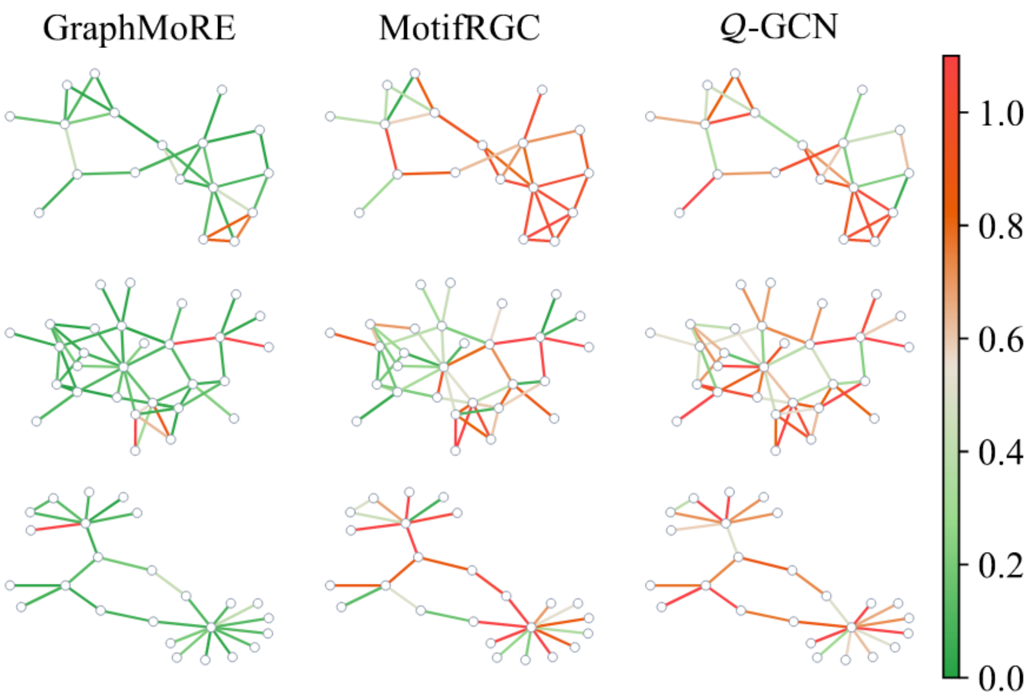
■ Performance on Synthetic Graphs & Ablation Study



Method	Cora		Airport	
	AUC	W-F1	AUC	W-F1
GraphMoRE	**97.91**	**80.51**	**97.53**	**91.75**
GraphMoRE (*w/o distortion*)	97.30	79.41	97.07	91.07
GraphMoRE (*w/o gating*)	96.96	79.17	96.86	90.87
GraphMoRE (*w/o diverse*)	94.34	79.69	94.69	90.46
GraphMoRE (*w/o align*)	97.11	79.07	96.34	90.18

- GraphMoRE significantly outperforms other baselines on synthetic graphs.
- **As the scale of datasets increases**, the performance of other baselines declines to varying degrees, while GraphMoRE consistently **maintains excellent performance.**

■ Comparison and Visualization of Embedding Distortion



Method	Cora	Citeseer	Airport	PubMed	Photo
GraphMoRE	**0.22**	**0.32**	**0.55**	**0.21**	**0.59**
MotifRGC	0.69	0.78	0.64	0.58	OOM
$\mathcal{Q}$-GCN	0.80	0.91	0.70	0.75	0.78

➢ GraphMoRE has **the lowest average embedding distortion**, indicating the excellent expressive ability for topological heterogeneity.

Paper Highlights

- We propose **analyzing topological heterogeneity from the perspective of local topology characterization**, **construct personalized embedding spaces for nodes**, and provide an alignment strategy to calculate pairwise distances, to minimize the distortion of heterogeneous topologies.
- To the best of our knowledge, **we are the first to introduce the MoE into Riemannian representation learning** to address the problem of topological heterogeneity.

Future Works

- Topological heterogeneity also constitutes a critical challenge for graph foundation models, which are expected to uniformly handle a wide variety of diverse graph data. Graph Riemannian MoE **provide a novel architectural perspective for GFMs.**
- Future work will **further explore the potential and broader impact of Riemannian MoE in enhancing the ability of GFMs** to uniformly handle diverse graph data from a wide variety of types and domains.

GraphMoRE: Mitigating Topological Heterogeneity via Mixture of Riemannian Experts

Zihao Guo, Qingyun Sun*, Haonan Yuan, Xingcheng Fu, Min Zhou, Yisen Gao, Jianxin Li

Email: guozh@act.buaa.edu.cn

Paper

Code