INTERNATIONAL JOINT CONFERENCE ON NEURAL NETWORKS
IJCNN2025
30 JUNE - 5 JULY 2025 | ROME, ITALY
INTERNATIONAL NEURAL NETWORK SOCIETY

TopoLink: Topology-enhanced Graph Transformer with Extended Persistent Homology For Link Prediction

Lizhi Liu

China UnionPay

UnionPay 银联

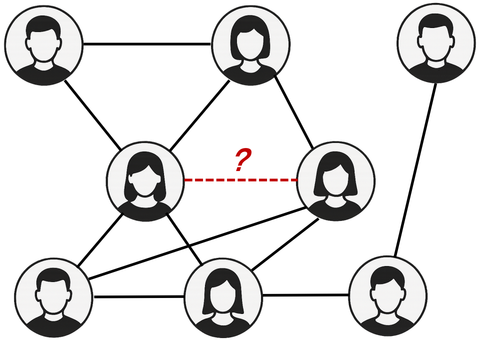
云闪付 UnionPay

The Challenge of Link Prediction

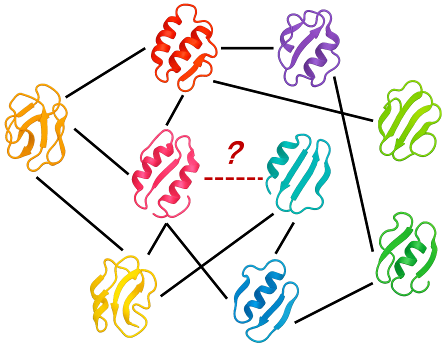
- **Link prediction** is a crucial task in graph machine learning, aiming to uncover unobserved relationships in networks.
- Graph Convolutional Network (**GCN**), while successful in many graph tasks, falters in link prediction.
- **Why do GCNs struggle?**
  - **Limited Expressive Power**: GCNs follow a message-passing scheme equivalent to the Weisfeiler-Leman (WL) test, which are provably incapable of counting 3-cycles (triangles) and consequently of counting Common Neighbors.

    Incorporating higher-order topological patterns
  - **Graph Sparsity**: In link prediction, many edges are inherently missing. GCNs restrict information flow only to existing edges, which hinders their ability to learn ideal node representations from the sparse data.

    Message-passing beyond hardwired interactions



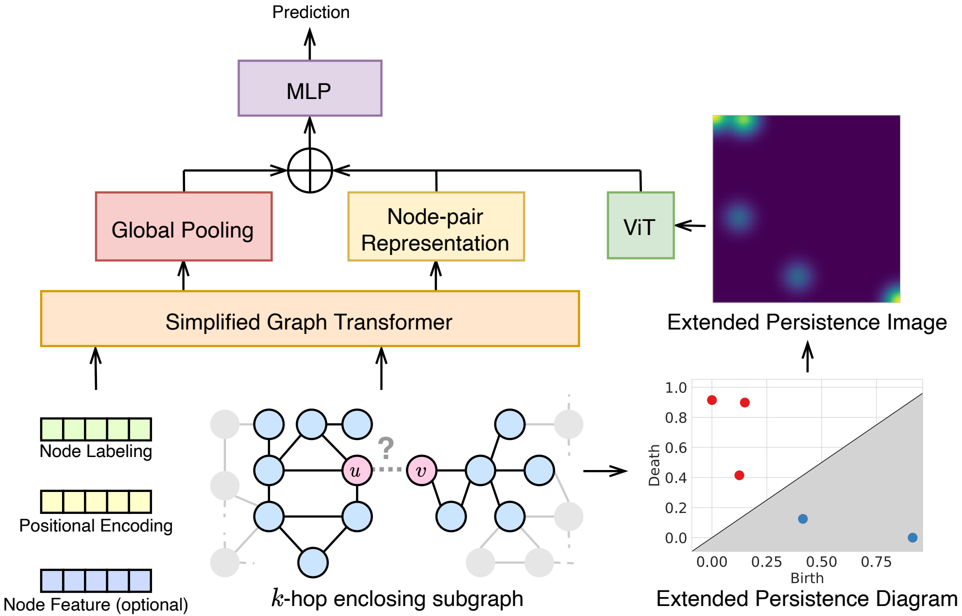
Social Network



Protein-Protein Interaction Network

Our Solution: TopoLink

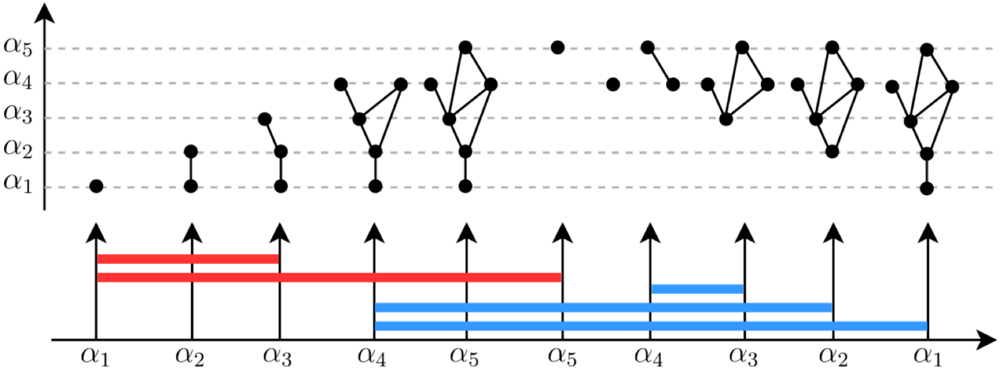
- We propose **TopoLink**, a novel topology-enhanced graph Transformer with extended persistent homology designed specifically for link prediction.
- **Core Contributions:**
    - A **Simplified Graph Transformer (SGT)** that integrates a classical GCN with a linear-complexity global attention module to capture both local and global information.
    - Introduction of **Extended Persistent Homology (EPH)** to detect topological features like cycles, enhancing the expressive power.
    - A novel approach to process topological data: We transform EPH into persistence images and use a **Vision Transformer (ViT)** to create powerful vectorized topological fingerprints.

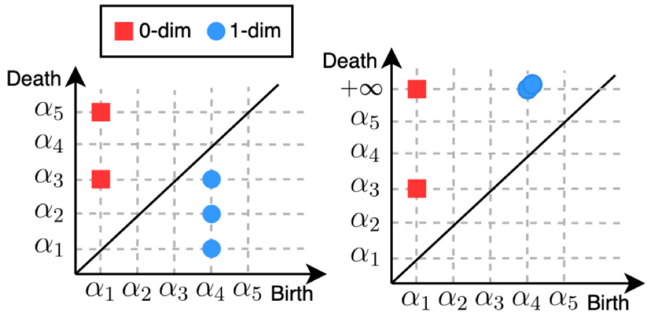


Overall model architecture of TopoLink

Background: Persistent Homology (PH)

- PH is a tool from Topological Data Analysis (TDA) that quantifies multi-scale topological features ("**holes**"), like connected components (0D holes) and loops (1D holes).
- It works by analyzing a **filtration**, which is a sequence of nested subgraphs.
- The "birth" and "death" of topological features during the filtration are recorded in a **persistence diagram**.
- **Problem with ordinary PH:** In graphs, essential features like the main connected component or large loops might never "die", leading to information loss.
- **Extended Persistent Homology (EPH)**: Solves this by using a complementary descending filtration, ensuring all features have a finite lifespan. This captures richer topological information.





Extended Persistent Diagram

Persistent Diagram

Methodology: Simplified Graph Transformer (SGT)

- **1. Graph Convolution** Local Information
  - Aggregates information from the local 1-hop neighborhood.
  - Follows the classic message-passing mechanism of GCNs:

$$\mathbf{X}_G^{(l+1)} = \tilde{\mathbf{A}}\mathbf{X}^{(l)}\mathbf{W}^{(l)}$$

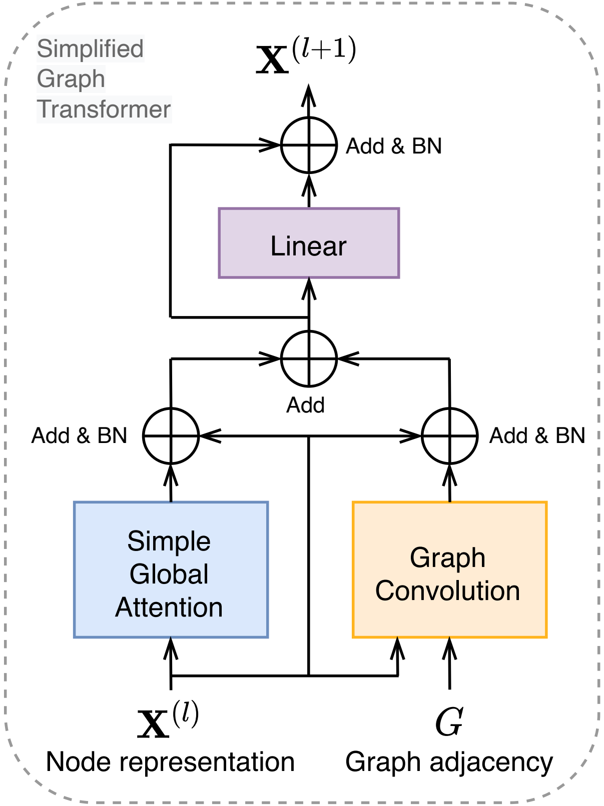
- **2. Simple Global Attention (SGA)** Global Information
  - Captures global information and implicit dependencies between all node pairs.

$$\mathbf{X}_A^{(l+1)} = \left(\mathbf{\Lambda}^{(l)}\right)^{-1}\left[\mathbf{V}^{(l)} + \frac{1}{N}\hat{\mathbf{Q}}^{(l)}\left(\left(\hat{\mathbf{K}}^{(l)}\right)^{\top}\mathbf{V}^{(l)}\right)\right]$$

where $\mathbf{Q}^{(l)} = \mathcal{F}_q^{(l)}\left(\mathbf{X}^{(l)}\right)$, $\mathbf{K}^{(l)} = \mathcal{F}_k^{(l)}\left(\mathbf{X}^{(l)}\right)$, $\mathbf{V}^{(l)} = \mathcal{F}_v^{(l)}\left(\mathbf{X}^{(l)}\right)$,

$$\mathbf{\Lambda}^{(l)} = \operatorname{diag}\left(\mathbf{1} + \frac{1}{N}\hat{\mathbf{Q}}^{(l)}\left(\left(\hat{\mathbf{K}}^{(l)}\right)^{\top}\mathbf{1}\right)\right).$$

  - Alters the standard attention computation to achieve linear complexity, making it scalable.
  - Helps overcome the message-passing limitations imposed by sparse, unobserved edges.
- **3. Fusion**
  - Local and global representations are combined, passed through normalization and linear layers to produce the final node embeddings.



Methodology: Detecting Topological Structure

- **Filtration**: We use a filtration based on the **Ollivier-Ricci curvature**, which measures local connectivity and geometry in the graph.

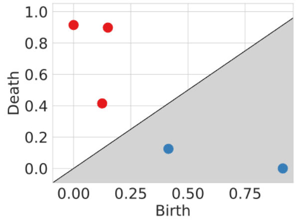
Filtration function $f : u \mapsto \sum_{v \in \mathcal{N}(u)} \kappa_\alpha(u, v)$

Ollivier-Ricci curvature $$\kappa_\alpha(u, v) := 1 - \frac{W_1(\mu_u^\alpha, \mu_v^\alpha)}{d(u, v)}$$

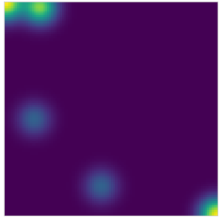
$$\mu_u^\alpha(x) := \begin{cases} \alpha, & \text{if } x = u, \\ (1-\alpha)/\deg(u), & \text{if } x \in \mathcal{N}(u), \\ 0, & \text{otherwise.} \end{cases}$$

Lazy random walk-based probability measure

- **Extended Persistence Image (EPI)**: The resulting EPH diagram is transformed into a 2D EPI. Each point in the diagram is represented by a Gaussian kernel on the image plane.



EPH diagram



EPI

- **Vision Transformer (ViT) Descriptor**: Instead of simply flattening the image, we use ViT to process EPI.
  - It allows for a deep abstraction and refinement of the visual patterns in the EPI.
  - The ViT's output serves as a sophisticated **topological fingerprint** for the subgraph.

Experimental Results

- **Key Result:** TopoLink achieves new state-of-the-art performance across all 15 datasets, outperforming runner-up methods by **0.55-18.84%**.
- Performance gains are particularly pronounced on non-attributed graphs, highlighting our model's strength in leveraging purely topological information when node features are absent.

RESULTS ON LINK PREDICTION BENCHMARKS. THE RESULTS ON OGB BENCHMARK COLLAB ARE TAKEN FROM THE OGB LEADERBOARD. THE BOLDFACE ITEMS REPRESENT THE BEST PERFORMANCE, AND THE RUNNER-UPS ARE UNDERLINED.

Dataset	Attributed Graph							Non-attributed Graph							
	Pubmed	CS	Physics	Computers	Photo	Wiki	Collab	USAir	NS	PB	Yeast	C.ele	Power	Router	E.coli
Metric	Hits@20	Hits@50	Hits@50	Hits@50	Hits@50	Hits@50	Hits@50	Hits@50	Hits@50	Hits@50	Hits@50	Hits@50	Hits@50	Hits@50	Hits@50
CN	27.91	53.84	66.83	33.29	43.50	76.96	61.37	87.26	82.85	50.81	79.38	65.89	18.51	13.76	60.16
GCN	49.55	62.49	72.98	35.98	47.33	80.79	44.75	92.45	46.72	54.58	79.81	65.89	35.81	61.60	79.40
SEAL	51.87	66.55	76.79	37.12	46.03	84.29	64.74	95.28	96.72	48.59	89.39	75.23	61.15	72.96	85.95
SIEG	56.24	81.11	77.28	39.23	55.67	89.21	-	95.28	97.08	59.72	89.48	80.37	44.01	74.72	84.11
ELPH	54.20	71.34	69.65	39.08	55.69	92.94	66.36	95.75	86.81	58.17	82.81	85.98	37.63	39.36	79.40
NCNC	54.91	74.58	75.67	40.01	53.48	91.69	66.61	94.81	96.35	57.69	86.91	74.30	38.24	73.44	81.51
BScNets	43.75	69.41	70.16	36.63	51.79	81.29	-	95.28	95.62	52.32	79.56	78.88	44.61	58.88	81.38
TLC-GNN	40.46	68.54	72.29	36.84	50.37	76.93	-	86.79	91.97	52.74	81.09	79.72	37.78	54.72	83.45
PLH-GNN	46.46	75.15	74.14	37.38	55.95	82.08	-	93.87	88.32	55.72	87.94	85.98	41.73	38.88	87.31
TopoLink	**58.91**	**86.16**	**78.88**	**40.99**	**58.18**	**93.45**	**67.92**	**99.53**	**98.90**	**63.61**	**93.58**	**92.52**	**67.07**	**88.80**	**90.86**

Experimental Results

- TopoLink is the **first PH-based method** to rank at the top of the competitive Stanford **OGB Leaderboard** for the Collab dataset.

OGB   Get Started   Updates   Large-Scale Challenge   Datasets   Leaderboards   Papers   Team   Github

Leaderboard for ogbl-collab

The Hits@50 score on the test and validation sets. The higher, the better.

Package: >=1.2.1

Rank	Method	Ext. data	Test Hits@50	Validation Hits@50	Contact	References	#Params	Hardware	Date
1	**HyperFusion**	No	0.7129 ± 0.0018	0.7385 ± 0.0099	Xinwei Zhang (Tsinghua University)	Paper, Code	1,064,446,212	RTX 3080	Feb 24, 2024
2	**GIDN@YITU**	No	0.7096 ± 0.0055	0.9620 ± 0.0040	Zixiao Wang, Yu Zhang(ZhejiangLab, HUST)	Paper, Code	60,449,025	DepGraph@SCTS/CGCL	Oct 10, 2022
3	**PLNLP + SIGN**	No	0.7087 ± 0.0033	1.0000 ± 0.0000	Liang Yao (Tencent)	Paper, Code	34,980,864	Tesla-P40 (24G GPU)	Apr 7, 2022
9	**TopoLink**	No	0.6792 ± 0.0074	0.6771 ± 0.0083	Lizhi Liu (China UnionPay)	Paper, Code	483,363,845	RTX 4090 (24GB GPU)	Sep 9, 2024

Ranked No. 9

Source: https://ogb.stanford.edu/docs/leader_linkprop/#ogbl-collab (Time: Jun 2025)

Ablation Study on Persistent Homology

- **Finding #1:** Removing the EPH module entirely (w/o EPH) causes a significant performance drop, demonstrating the necessity of topological features.
- **Finding #2:** EPH provides better results than ordinary PH, confirming its ability to capture more useful information.
- **Finding #3:** Using the ViT to process the persistence image is superior to simply flattening it (w/ Flatten PI) or using other vectorization methods like Persistence Landscapes (w/ EPL). This supports our novel use of computer vision techniques.

ABLATION ANALYSIS OF EXTENDED PERSISTENT HOMOLOGY.

Dataset	Pubmed	CS	PB	Yeast
TopoLink	58.91	86.16	63.61	93.58
w/o EPH	55.70	83.55	60.98	92.66
w/ Ordinary PH	57.96	84.93	61.62	92.74
w/ Flatten PI	57.05	84.39	58.77	92.73
w/ PL	55.39	84.98	60.38	92.21
w/ EPL	55.96	85.83	61.28	92.90

Conclusion

- We introduce **TopoLink**, a new SOTA topological model for link prediction.
- **Key Innovations:**
    - A **Simplified Graph Transformer (SGT)** that combines local GCN-based message passing with an linear global attention mechanism, overcoming issues of graph sparsity.
    - The integration of **Extended Persistent Homology (EPH)** to capture higher-order topological structures like cycles, boosting the model's expressive power.
    - The first use of a **Vision Transformer (ViT)** to deeply analyze and vectorize persistence images, creating powerful topological fingerprints.
- **Limitations**:
    - The **computation time** of persistent homology is significantly longer compared to other components in the model, becoming a bottleneck.
    - This hinders the **scalability** of the methodology to large-scale graphs.
- **Future Work**:
    - Improve the **computational efficiency** of persistent homology to enable scalability to large graphs.
    - Generalize the framework to other graph types like **directed and bipartite graphs**.

Thank you!

GitHub

Homepage