Visualizing Graph Neural Networks

with CorGIE : Corresponding a Graph to Its Embedding

Zipeng Liu
University of British Columbia (UBC)

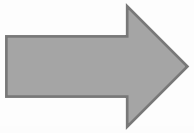
Yang Wang
Facebook

Jürgen Bernard
University of Zurich

Tamara Munzner
UBC

Paper under review
ChinaVis 2021 at Wuhan
July 2021

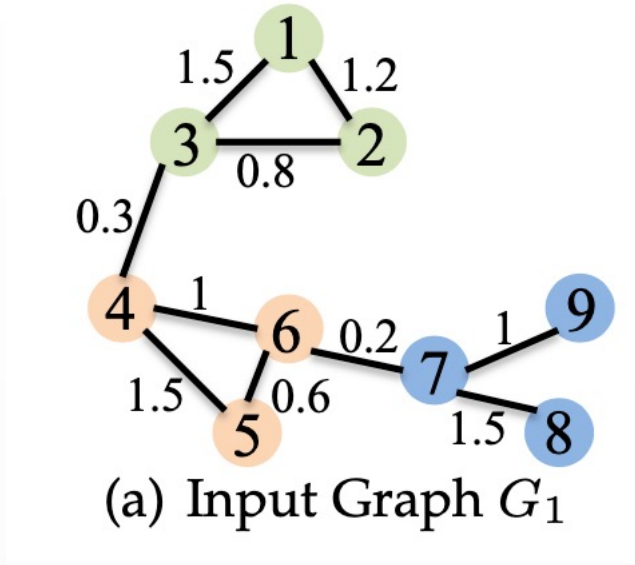
Outline



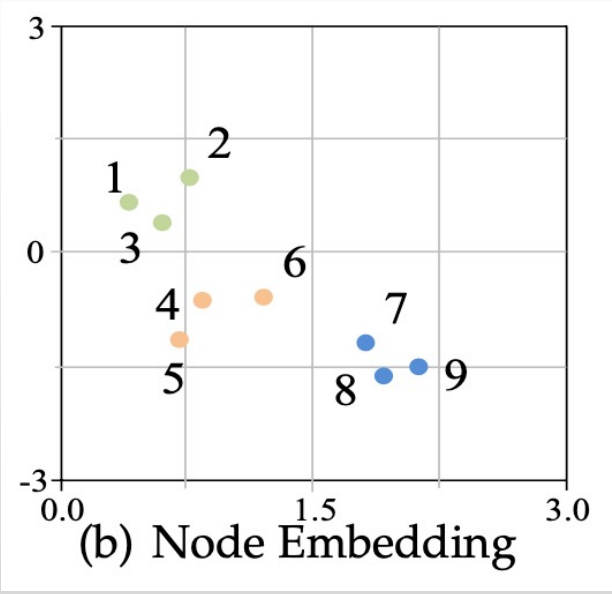
- Introduction of GNN
- Visual evaluation of GNN
  - Previous
  - Ours: CorGIE
    - Overview
    - Data & tasks
    - CorGIE interface
    - Reflections

Graph neural network (GNN)

- Machine learning (ML) models for graph
  - Like CNN for images
  - Like Transformer for text
- Many real-world graph-related applications
  - Node classifications
    - e.g. fraud detection, disease classification
  - Link prediction
    - e.g. recommendation of products, protein interactions



(a) Input Graph G_1

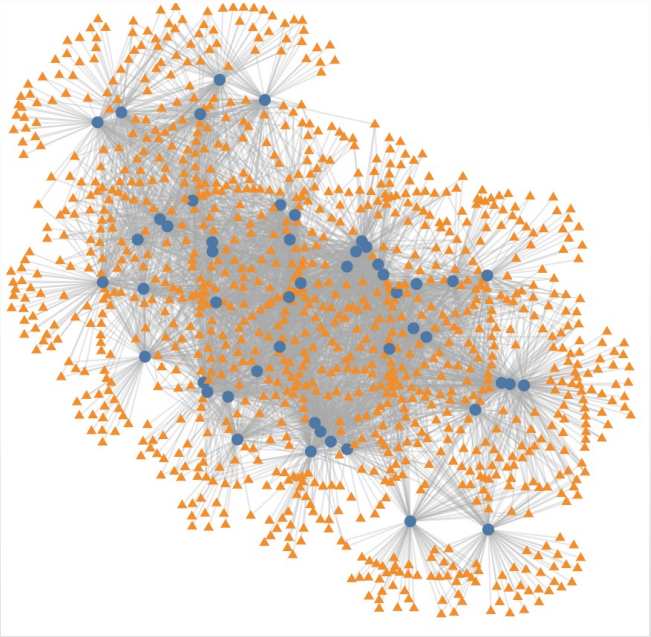


(b) Node Embedding

Cai et al. TKDE'18

Graph neural network (GNN)

Input graph → Graph neural Network (GNN) → High Dim latent space: Node embedding → Downstream ML applications → Predictions (node labels, links)



Movie – user graph

Node 0	-1.98	0.74	-0.51	1.19	-1.20	0.97	-1.66	0.90	0.06	-1.89	-1.33	-0.77	0.37	1.60	-0.13	1.47
Node 1	-0.21	0.11	-0.08	0.17	-0.16	0.14	-0.21	0.15	-0.03	-0.20	-0.18	-0.15	0.09	0.15	-0.01	0.16
Node 2	-0.23	0.12	-0.09	0.20	-0.17	0.16	-0.23	0.17	-0.04	-0.22	-0.21	-0.17	0.10	0.16	-0.02	0.18
Node 3	-0.27	0.15	-0.11	0.23	-0.20	0.19	-0.27	0.21	-0.05	-0.26	-0.25	-0.21	0.13	0.18	-0.02	0.21
Node 4	-0.30	0.17	-0.12	0.27	-0.23	0.23	-0.31	0.24	-0.07	-0.29	-0.29	-0.25	0.15	0.20	-0.03	0.23
Node 5	-0.19	0.09	-0.06	0.14	-0.13	0.11	-0.17	0.11	-0.01	-0.18	-0.15	-0.11	0.06	0.15	-0.01	0.15
Node 6	-0.28	0.16	-0.11	0.26	-0.22	0.22	-0.30	0.23	-0.07	-0.28	-0.28	-0.24	0.14	0.19	-0.03	0.22
Node 7	-0.30	0.17	-0.12	0.27	-0.23	0.22	-0.31	0.24	-0.07	-0.29	-0.28	-0.25	0.15	0.20	-0.03	0.23
Node 8	-0.23	0.12	-0.08	0.18	-0.16	0.14	-0.21	0.15	-0.02	-0.22	-0.19	-0.15	0.09	0.17	-0.01	0.18
Node 9	-0.31	0.18	-0.12	0.28	-0.24	0.24	-0.33	0.25	-0.08	-0.30	-0.30	-0.26	0.16	0.20	-0.03	0.24
Node 10	-0.33	0.19	-0.13	0.30	-0.26	0.25	-0.35	0.27	-0.08	-0.32	-0.32	-0.28	0.17	0.22	-0.03	0.26
Node 11	-0.21	0.11	-0.07	0.17	-0.16	0.14	-0.20	0.15	-0.03	-0.20	-0.18	-0.14	0.09	0.16	-0.01	0.17
Node 12	-0.20	0.10	-0.07	0.16	-0.15	0.13	-0.19	0.14	-0.03	-0.19	-0.17	-0.13	0.08	0.15	-0.01	0.16
Node 13	-0.26	0.14	-0.10	0.23	-0.20	0.19	-0.26	0.20	-0.05	-0.25	-0.24	-0.20	0.12	0.18	-0.02	0.20
Node 14	-0.19	0.08	-0.06	0.13	-0.13	0.11	-0.17	0.11	-0.01	-0.18	-0.15	-0.10	0.06	0.14	-0.01	0.15
Node 15	-0.16	0.06	-0.04	0.09	-0.10	0.07	-0.13	0.07	0.01	-0.14	-0.11	-0.06	0.03	0.13	-0.00	0.12

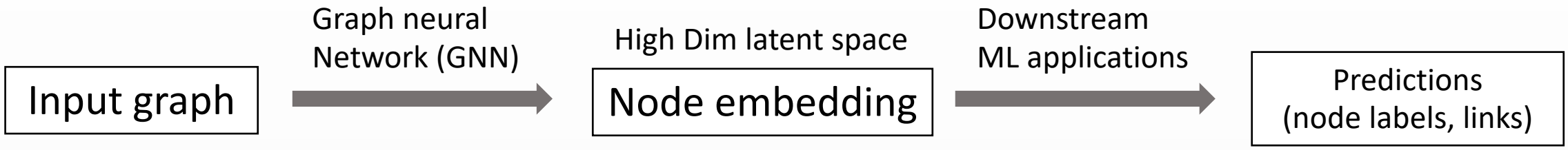
A vector for each node

Node 0: Alice

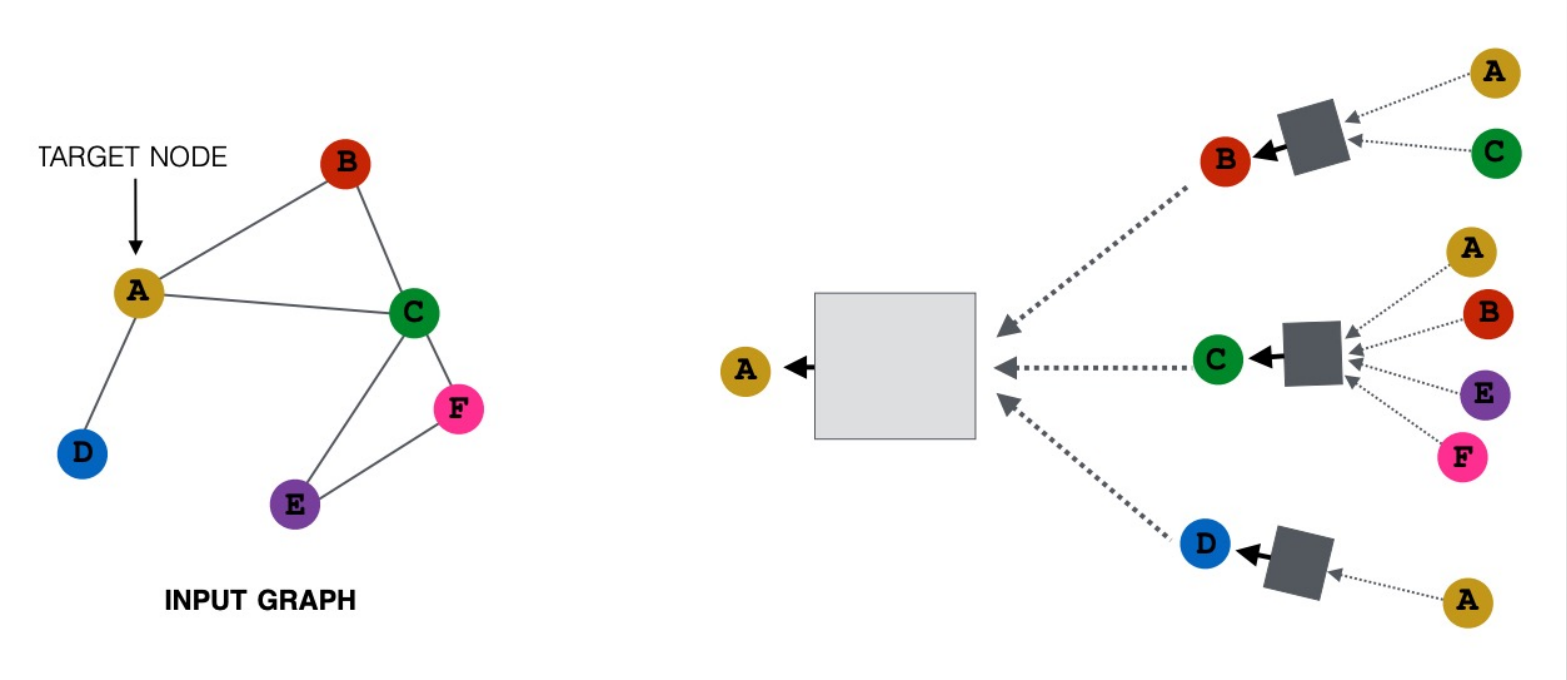
Movie recommendation

Node 12: Lord of the Rings

GNN: neighborhood aggregation



Node features are aggregated / passed through **topological neighborhood**



https://snap-stanford.github.io/cs224w-notes/machine-learning-with-networks/graph-neural-networks

Evaluate GNN

Two big-picture questions:

- “Are we there yet?”: should we train / tune more?
- “Are we lost?”: does it behave as we expect?

Outline

- Introduction of GNN
- Visual evaluation of GNN
  - Previous
  - Ours: CorGIE
    - Overview
    - Data & tasks
    - CorGIE interface
    - Reflections

Evaluate GNN: previous approaches

Topology
Node features

Input graph

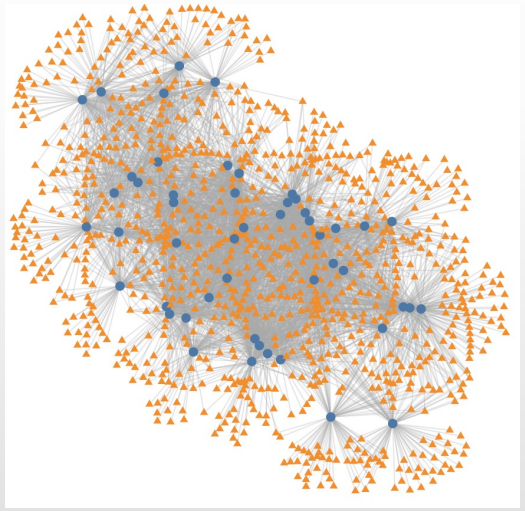
Graph neural
Network (GNN)

High Dim latent space

Node embedding

Downstream
ML applications

Predictions
(node labels, links)

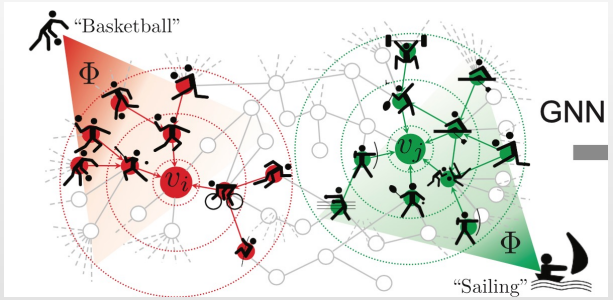


Dimensional reduction
(simple inspection)

Quant metrics

e.g. accuracy, hit rate (how many predicted links are true)

Influential nodes



Ying et al. GNNExpaliner. NeurIPS'19.

Evaluate GNN: previous approaches

Topology
Node features

Input graph

Graph neural
Network (GNN)

High Dim latent space

Node embedding

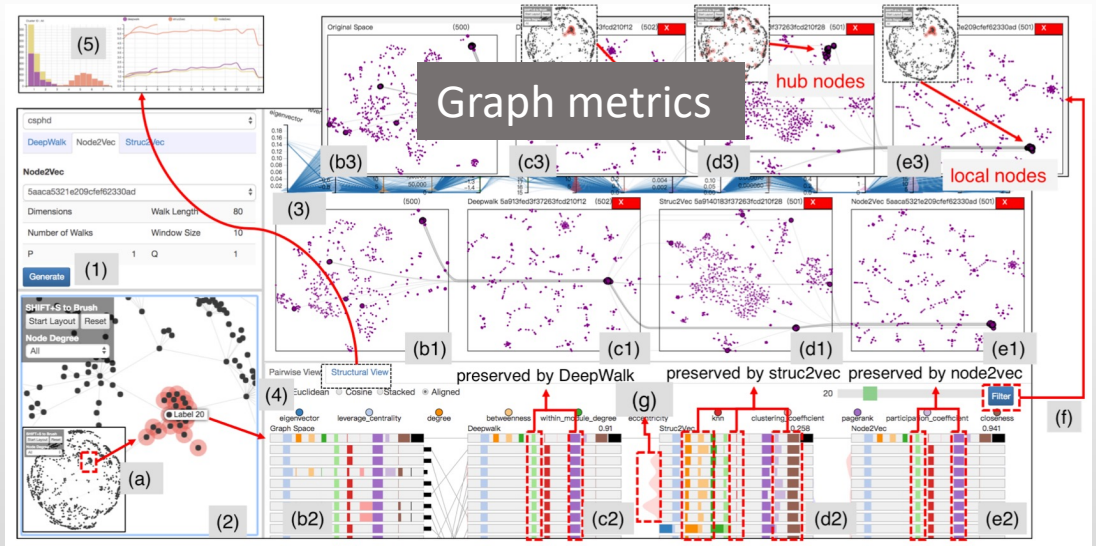
Downstream
ML applications

Predictions
(node labels, links)

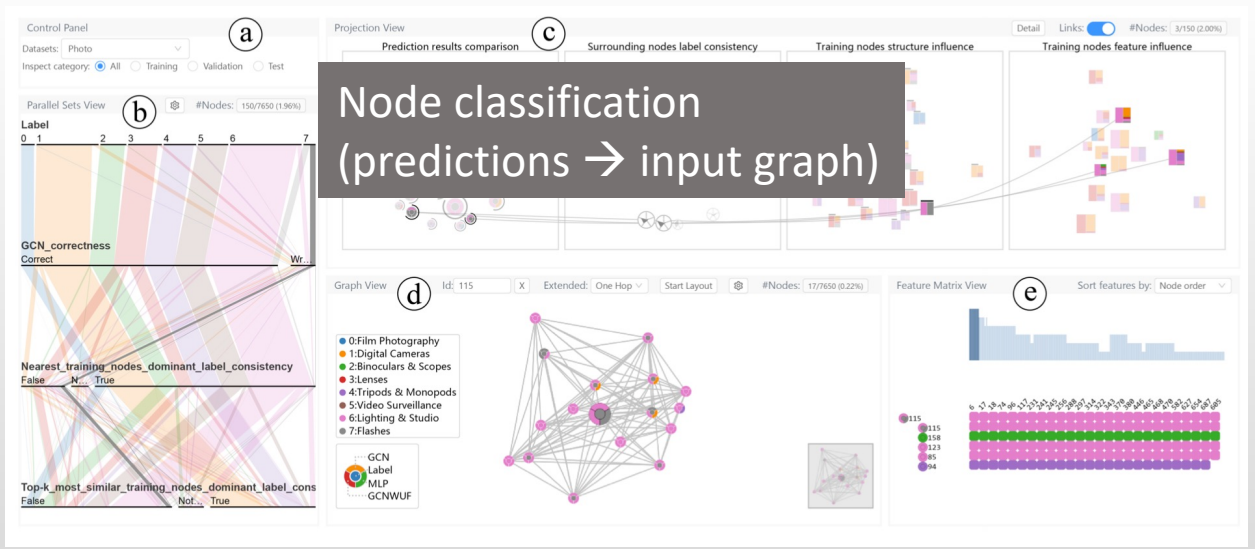
Dimensional reduction
(simple inspection)

Quant metrics

Influential nodes



Li et al. EmbeddingVis. VAST'18.



Jin et al. GNNVis. Arxiv'20.

Evaluate GNN: previous approaches

Topology
Node features

Input graph

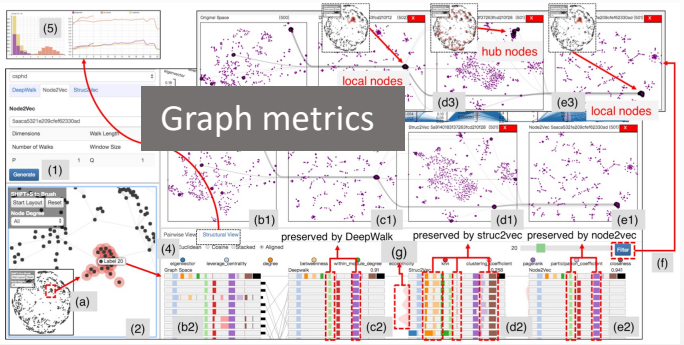
Graph neural
Network (GNN)

High Dim latent space

Node embedding

Downstream
ML applications

Predictions
(node labels, links)

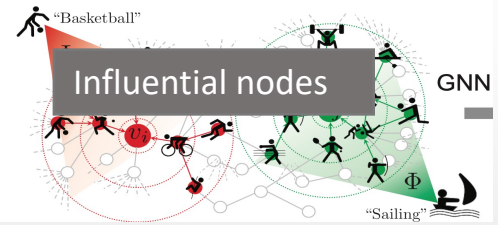


Li et al. EmbeddingVis. VAST'18.

Dimensional reduction
(simple inspection)

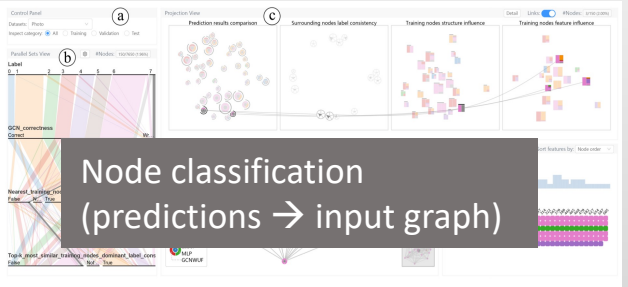
Quant metrics

e.g. accuracy, hit rate (how many predicted links are true)



Ying et al. GNNExpaliner.

Input graph & node embedding
under-used!



Jin et al. GNNVis. Arxiv'20.

Outline

- Introduction of GNN
- Visual evaluation of GNN
  - Previous
  - Ours: CorGIE
    - Overview
    - Data & tasks
    - CorGIE interface
    - Reflections

Evaluate GNN: CorGIE idea

Topology
Node features

Input graph

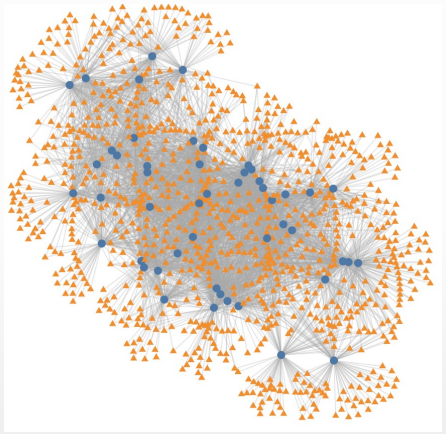
Graph neural
Network (GNN)

High Dim latent space

Node embedding

Downstream
ML applications

Predictions
(node labels, links)



Shared topo neighbors
Similar node features

GNN

Explore correspondences

Nearby positions

Evaluate GNN: CorGIE idea

Topology
Node features

Input graph

Graph neural
Network (GNN)

High Dim latent space

Node embedding

Downstream
ML applications

Predictions
(node labels, links)

Shared topo neighbors
Similar node features

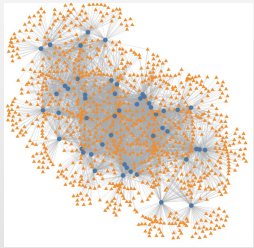
Explore correspondences

Nearby positions

Examples of correspondences:

Check [Similar topology? Similar node features?]

Pick [a cluster]



Evaluate GNN: CorGIE idea

Topology
Node features

Input graph

Graph neural
Network (GNN)

High Dim latent space

Node embedding

Downstream
ML applications

Predictions
(node labels, links)

Shared topo neighbors
Similar node features

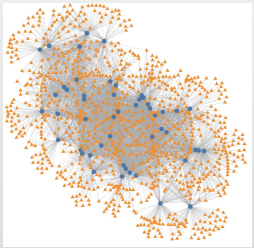
Explore correspondences

Nearby positions

Examples of correspondences:

Check [Similar topology? Similar node features?] ◂- - - - - Pick [a cluster]

Check [Different topology? Different node features?] ◂- - - - - Pick [two far-away clusters]



Evaluate GNN: CorGIE idea

Topology
Node features

Input graph

Graph neural
Network (GNN)

High Dim latent space

Node embedding

Downstream
ML applications

Predictions
(node labels, links)

Shared topo neighbors
Similar node features

Explore correspondences

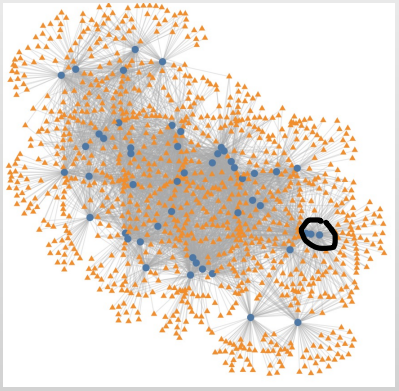
Nearby positions

Examples of correspondences:

Check [Similar topology? Similar node features?] ← - - - - Pick [a cluster]

Check [Different topology? Different node features?] ← - - - - Pick [two far-away clusters]

Pick [two nodes sharing many topo neighbors] - - - - → Check [how close the nodes are compared to others?]



Evaluate GNN: CorGIE idea

Topology
Node features

Input graph

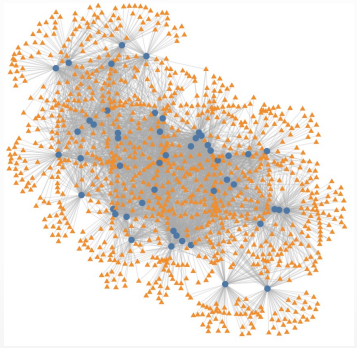
Graph neural
Network (GNN)

High Dim latent space

Node embedding

Downstream
ML applications

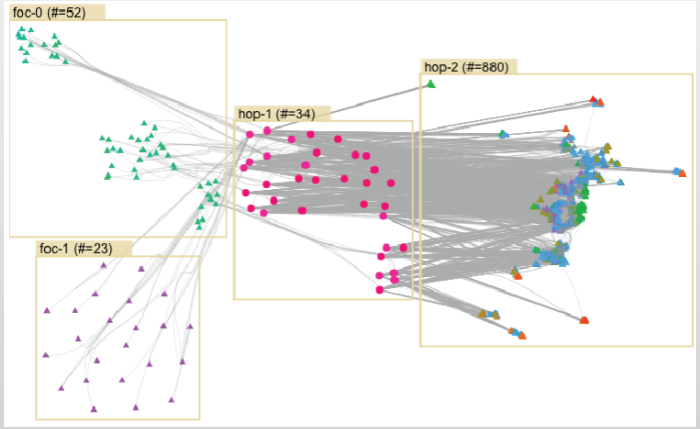
Predictions
(node labels, links)



Explore correspondences

Shared topo neighbors
Similar node features

Nearby positions



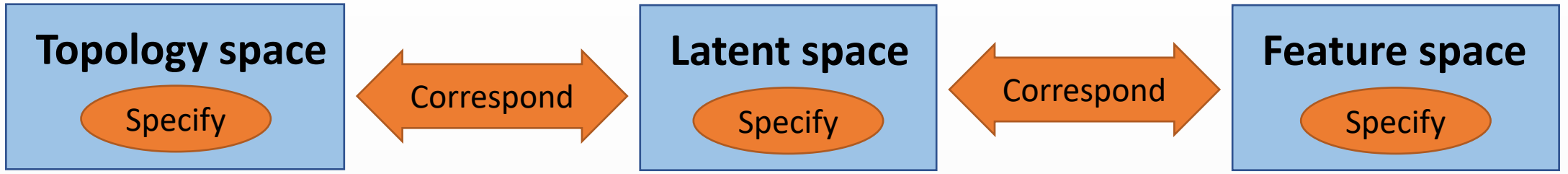
K-hop layout

- Topo neighbors in hops
- Clustering structure within hop

Data (sub-)spaces

	Topology space	**Latent space**	**Feature space**
Targets:	Neighbors; connections	Clustering; relative positions	Feature distribution

Tasks

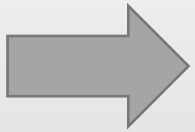


Targets: Neighbors; connections | Clustering; relative positions | Feature distribution

- **Specify** nodes in space
  - Properties of the targets
    - E.g. tight clusters in latent space, disconnected nodes in topology space
- **Correspond** them between spaces
  - Targets should tell the same story between spaces
    - E.g. nodes in tight clusters in latent space are expected to share neighbors in topo space
- Iterative process
  - Refine specification
  - Inspire new specification

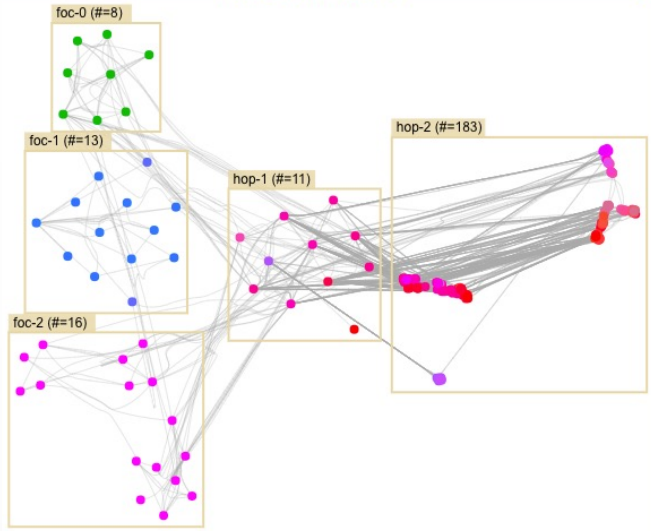
Outline

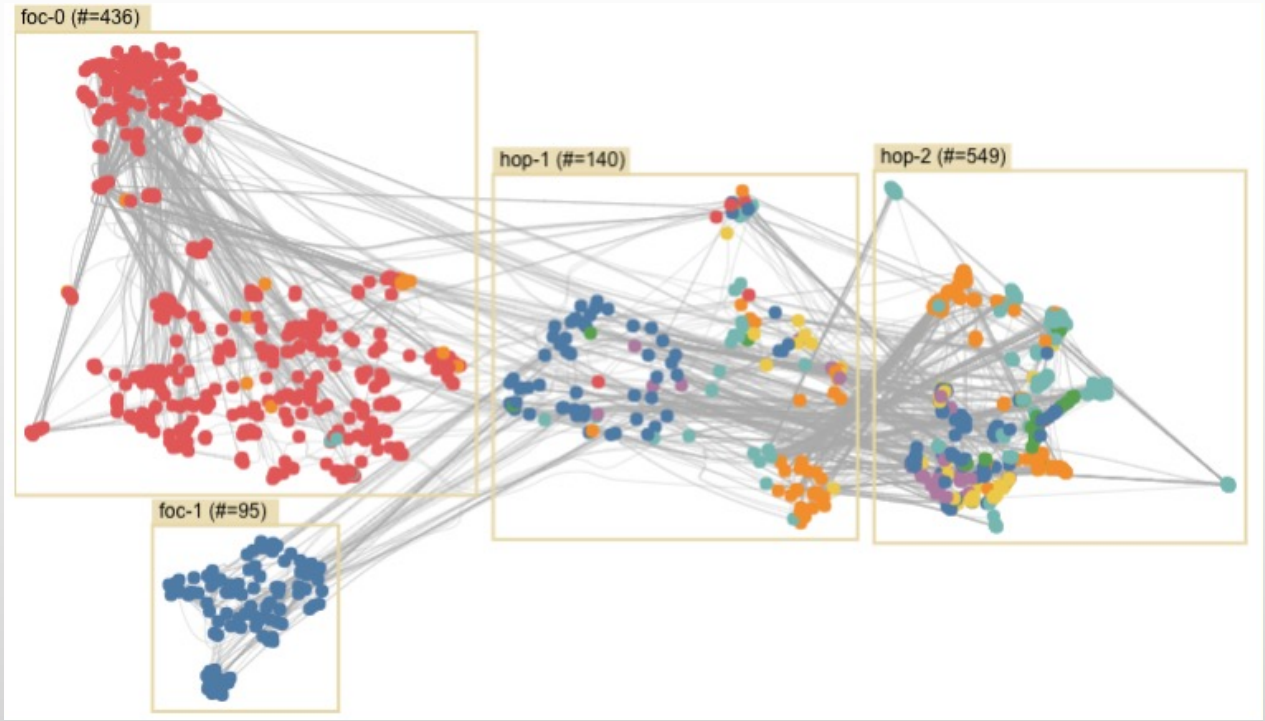
- Introduction of GNN
- Visual evaluation of GNN
  - Previous
  - Ours: CorGIE
    - Overview
    - Data & tasks
    - CorGIE interface
    - Reflections

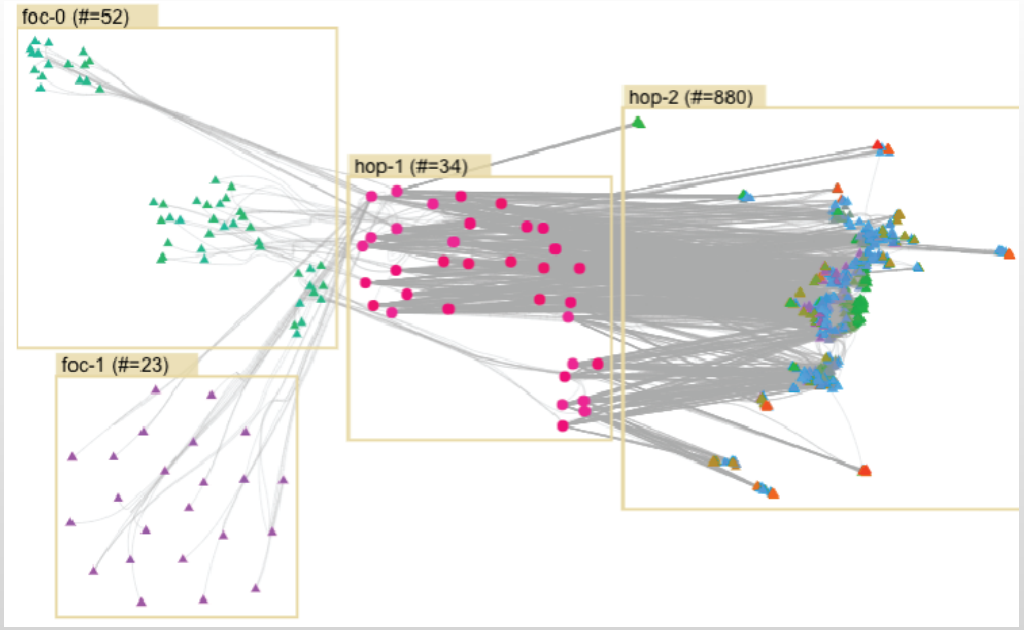


CorGIE interface: K-hop layout

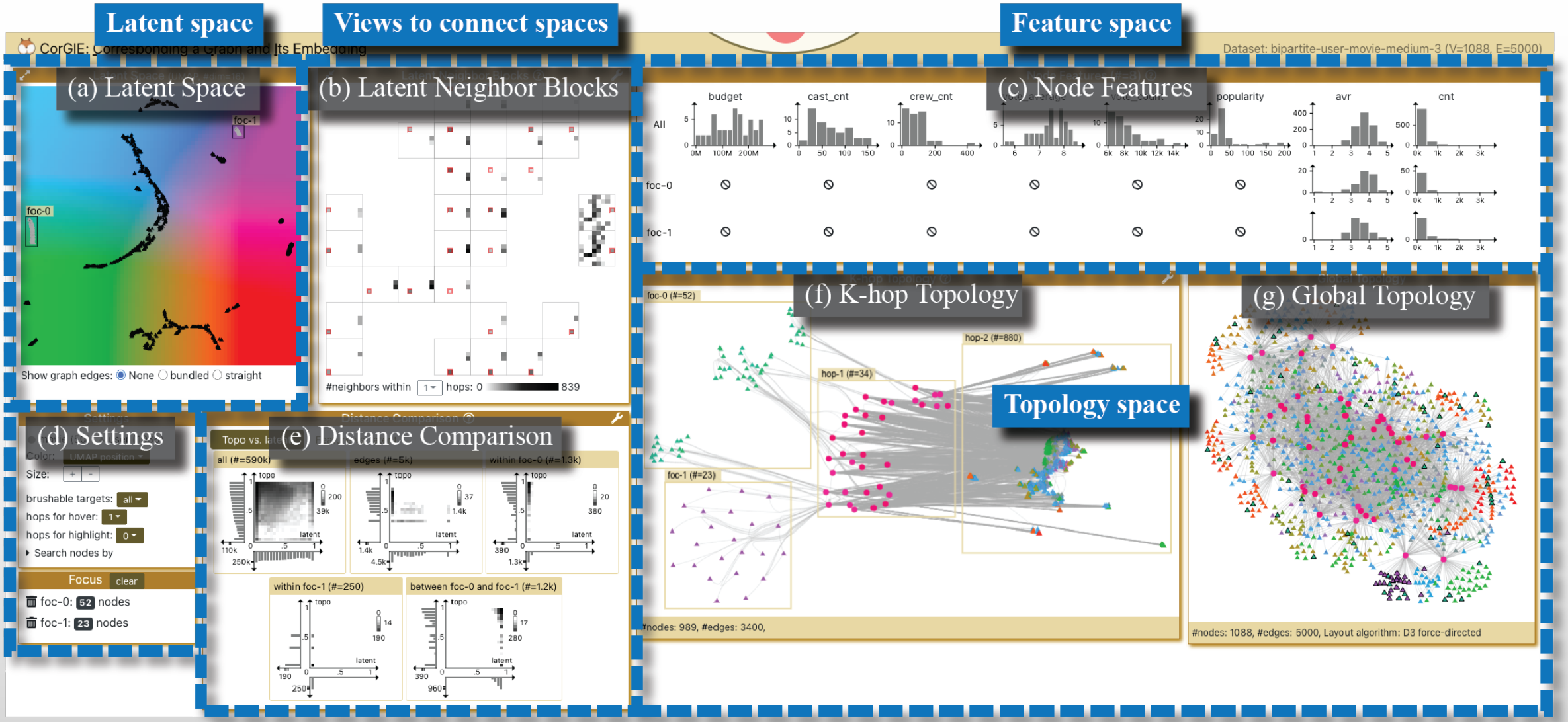
- Show topo neighbors of user-specified node sets
  - Mimic how info is aggregated in the GNN
    - Boxes from left to right: Focal nodes, hop-1, hop-2, …
  - Within box, cluster neighbors using their topo connections





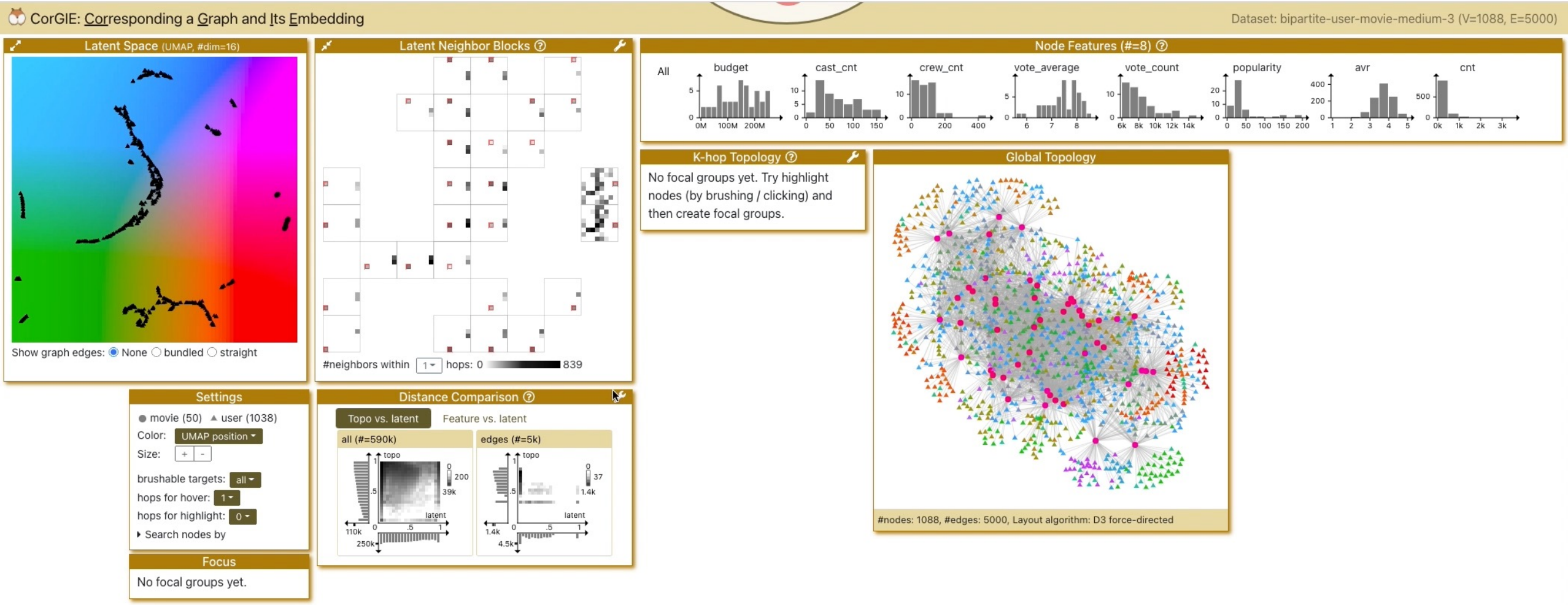


Multiple views for data spaces and connecting them



CorGIE: Corresponding a Graph and Its Embedding

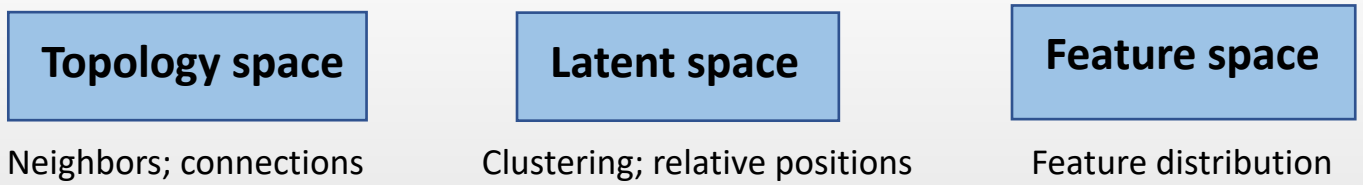
Dataset: bipartite-user-movie-medium-3 (V=1088, E=5000)



Reflections

- Correspondences between input, output, middleware
  - Grey-box approach
    - Works for many GNN models
  - Generalizable to other types of models

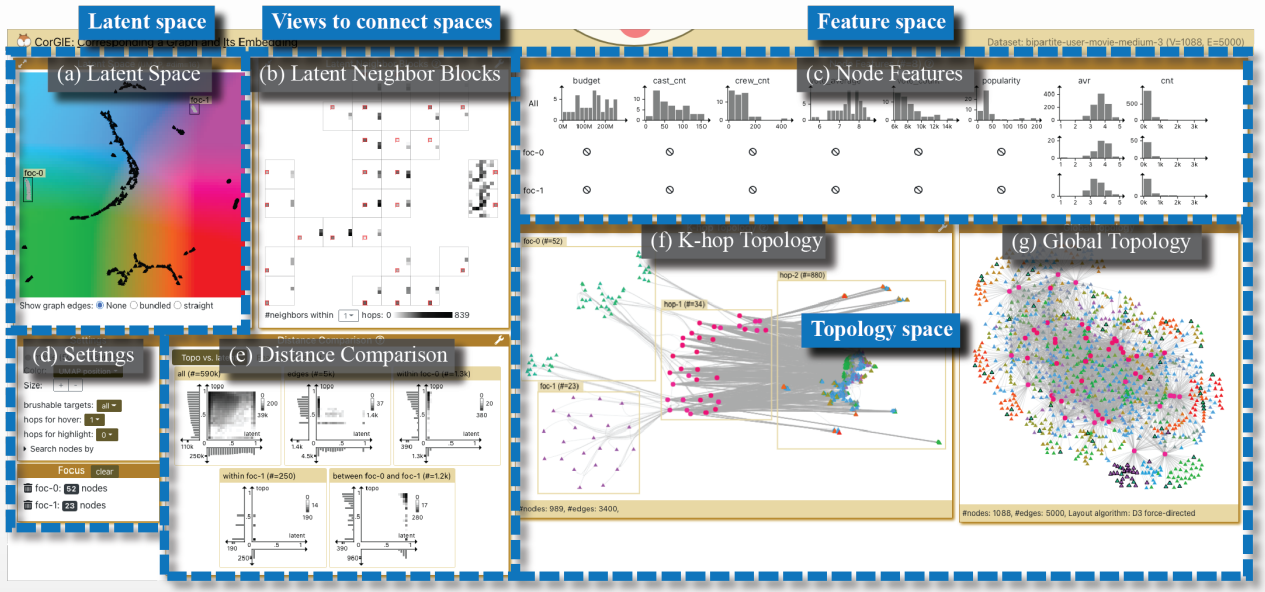
- Data space notion



  - Useful to think about connecting data spaces
  - New spaces for future extension
    - e.g., geospatial spaces for graphs dealing with traffic

Visualizing Graph Neural Networks with CorGIE: Corresponding a Graph to Its Embedding

Zipeng Liu, Yang Wang, Jürgen Bernard, Tamara Munzner
Paper under review
Presented at ChinaVis 2021
http://www.cs.ubc.ca/group/infovis/pubs/2021/corgie/



Take-away

- Evaluate GNN visually by exploring **correspondences** between graph & its embedding
  - Abstraction: connecting data spaces
- Reveal graph topology used in GNN with **K-hop layout**