



---

# PaGE-Link: Path-based Graph Neural Network Explanation for Heterogeneous Link Prediction

---

Shichang Zhang<sup>1</sup>, Jiani Zhang<sup>2</sup>, Xiang Song<sup>2</sup>,  
Soji Adeshina<sup>2</sup>, Da Zheng<sup>2</sup>, Christos  
Faloutsos<sup>2,3</sup>, Yizhou Sun<sup>1,2</sup>

<sup>1</sup>University of California, Los Angeles (UCLA)

<sup>2</sup>Amazon <sup>3</sup>Carnegie Mellon University

July 2023

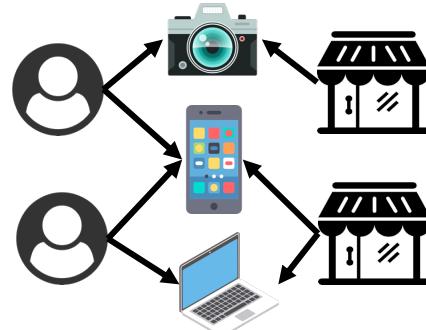
# Outline

---

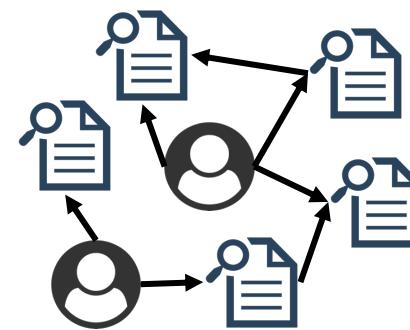
- Machine Learning on Graphs
  - Graph Neural Networks (GNNs)
  - GNNs for Link Prediction
- Model Explainability
- PaGE-Link
  - Main Idea
  - Path-Enforcing Mask
  - Experiments

# Machine Learning on Graphs

Graphs are a general language for modeling entities with relations



E-commerce graphs



Citation graphs



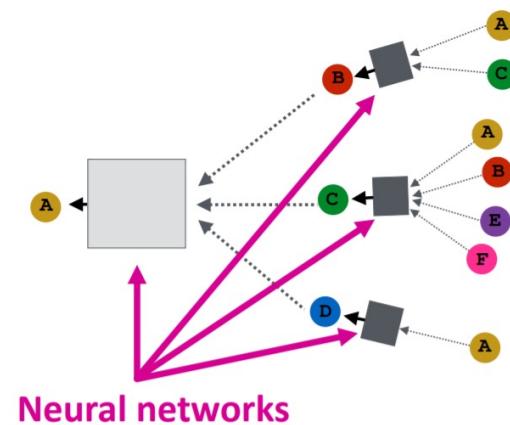
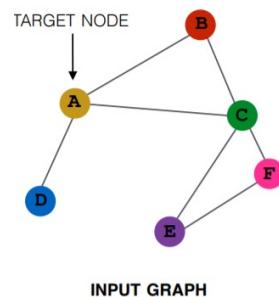
Transportation graphs

(Image Credit: [www.visitlondon.com](http://www.visitlondon.com))

Molecule graphs, code graphs, scene graphs, and many more ...

# Graph Neural Networks

GNNs: a family of neural-network-based models for machine learning on graphs



(Hamilton, W. L., Ying, R., & Leskovec, J. 2017)

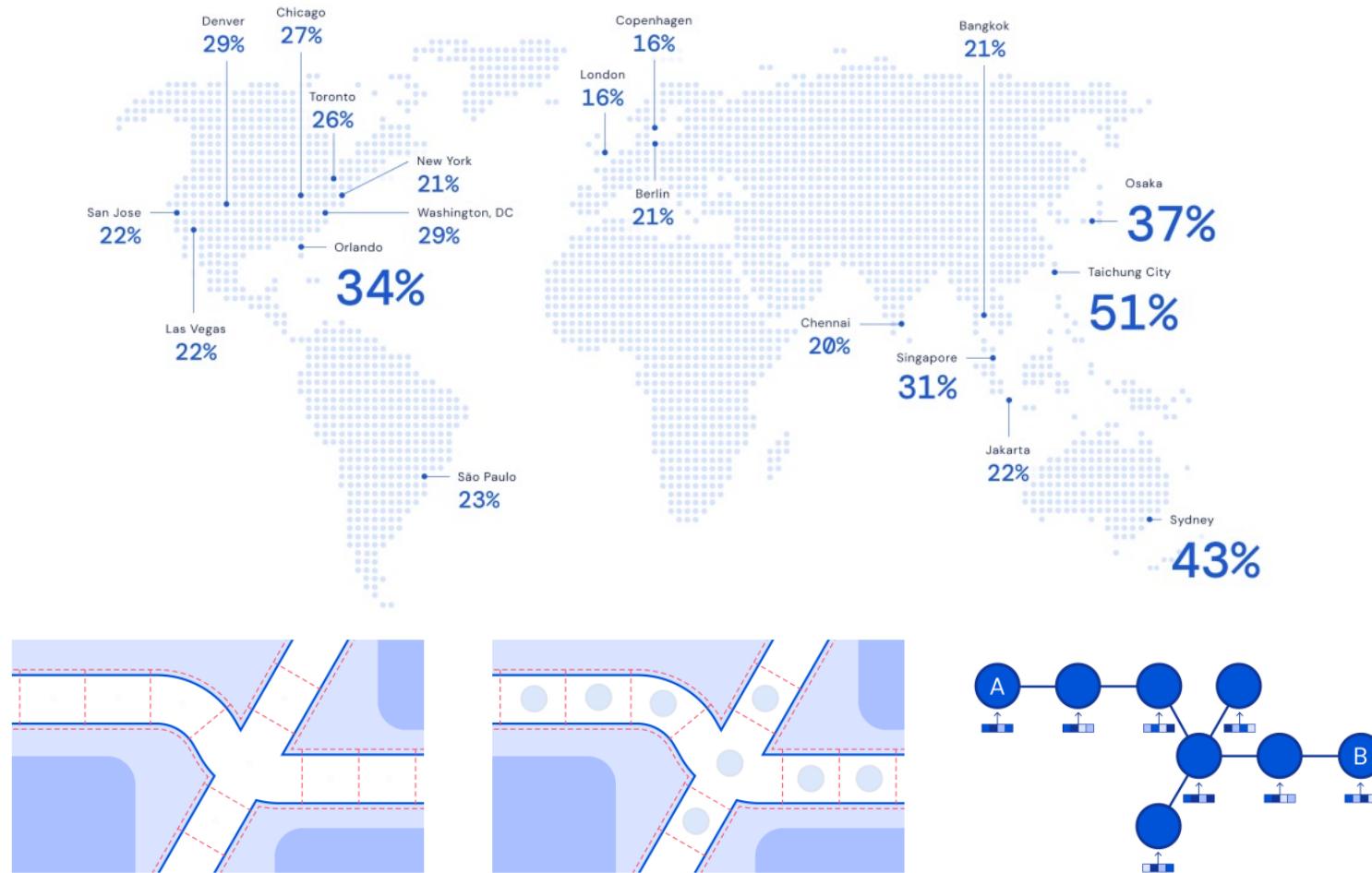
*Message passing*

Each node aggregates messages from its neighbors and recursively extends to multi-hop neighbors

Node A with neighbors  $\mathcal{N}(A)$  aggregates messages at step  $l$ :

$$\mathbf{h}_A^{(l)} = \text{AGGR}(\mathbf{h}_A^{(l-1)}, \{\mathbf{h}_i^{(l-1)} | i \in \mathcal{N}(A)\})$$

# GNNs for Google Map ETA Prediction

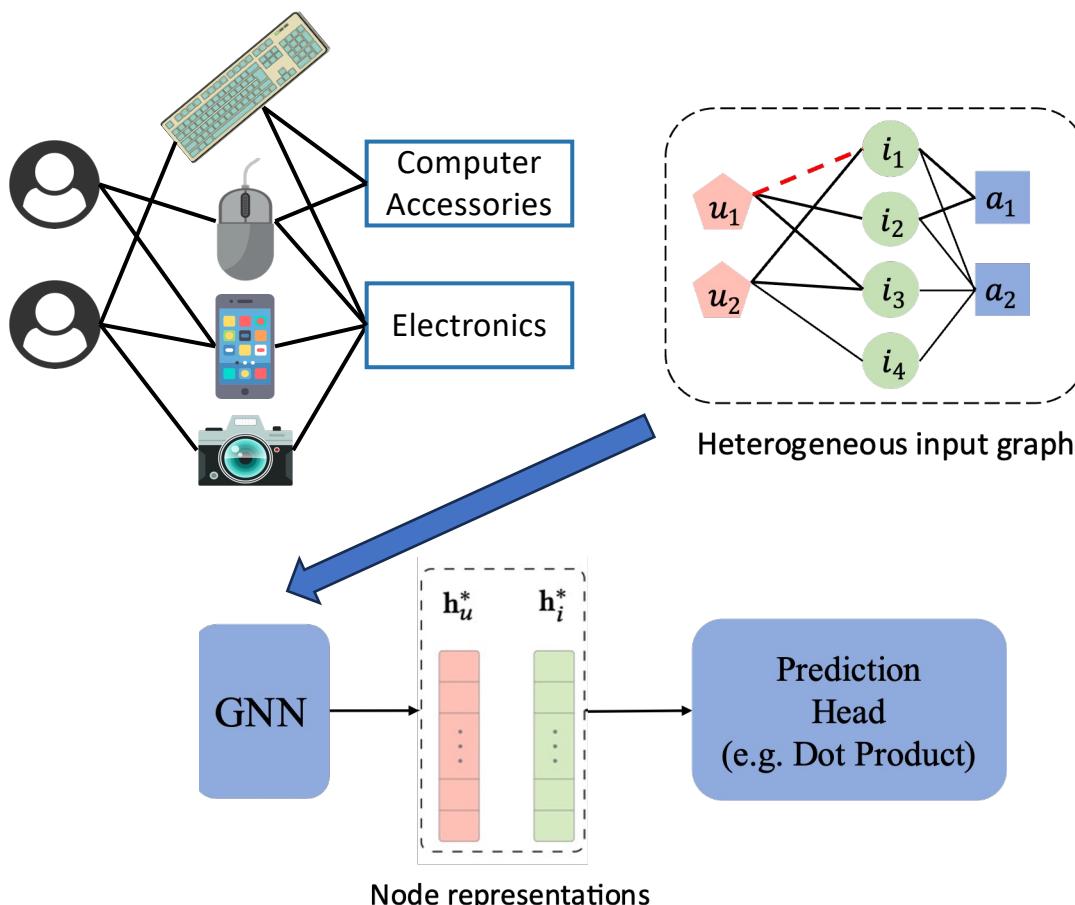


Derrow-Pinion, Austin, et al. "ETA prediction with graph neural networks in google maps." CIKM. 2021.

Zhang, Shichang, et al. "PaGE-Link: Path-based Graph Neural Network Explanation for Heterogeneous Link Prediction" WWW 2023

# GNNs for Link Prediction

- Link Prediction: Recommend items to users



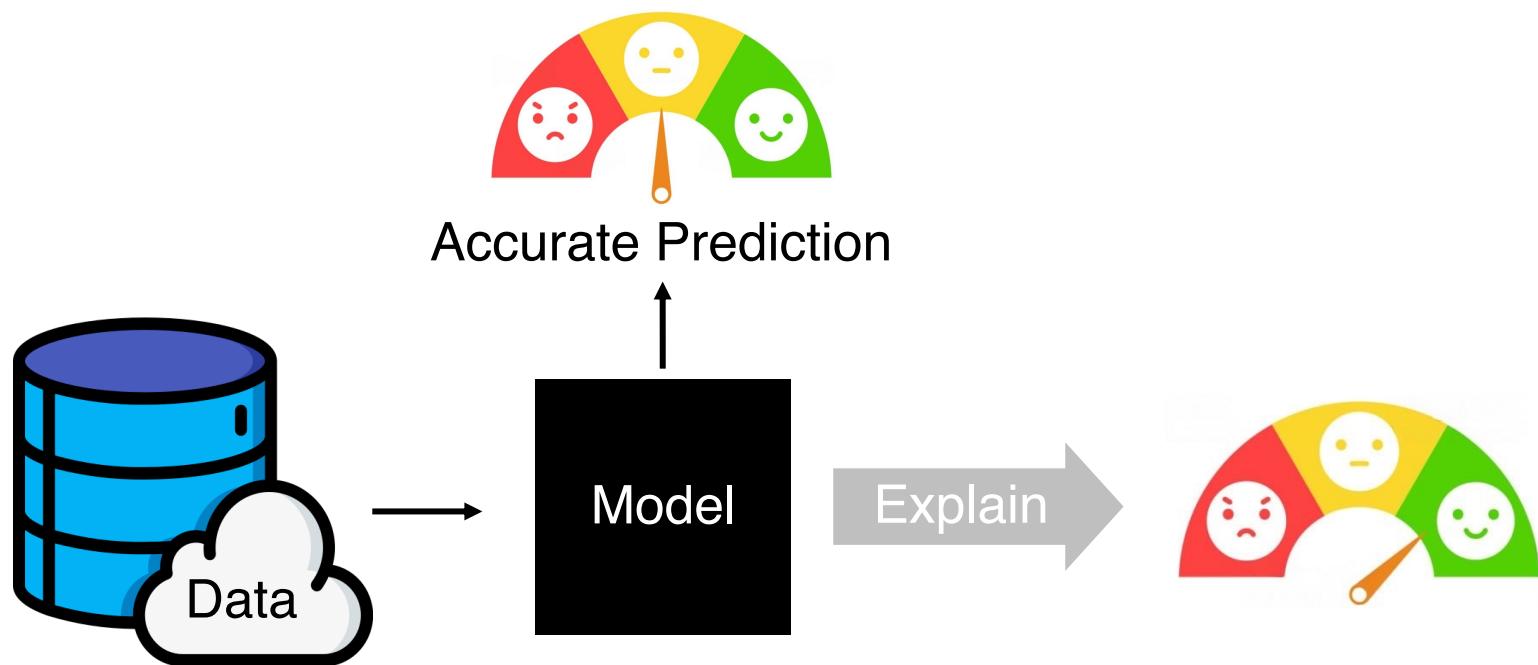
Traditional models  
(matrix factorization & shallow embedding)      **GNNs**

Model	Amazon-Books	
	Recall@20	NDCG@20
MF-BPR	0.0338	0.0261
CML	<u>0.0522</u>	<u>0.0428</u>
ENMF	0.0359	0.0281
DeepWalk	0.0346	0.0264
LINE	0.0410	0.0318
Node2Vec	0.0402	0.0309
NGCF	0.0344	0.0263
NIA-GCN	0.0369	0.0287
LR-GCCF	0.0335	0.0265
LightGCN	0.0411	0.0315
DGCF	<u>0.0422</u>	<u>0.0324</u>
UltraGCN <sub>Base</sub>	0.0504	0.0393
UltraGCN	<b>0.0681</b>	<b>0.0556</b>

GNN achieve SOTA link prediction/recommendation results (Mao et al. CIKM 2021).

# Model Explainability

- Many start-of-the-art AI models are black boxes.
- Explainability helps to increase user satisfaction and improve model design.



# A Real Amazon Recommendation

Query



Recommendation without explanation

More items to explore

Mouse Pads 11 x 8.7 inches- Stitched Edges Premium-Textured Large Mouse Pads Mat Natural...  
★★★★★ 2,729  
\$4.98  
Get it as soon as Thursday, Jun 30

Sponsored Rii RK907 Ultra-Slim Compact USB Wired Keyboard for Mac and PC, Windows 10/8 / 7 / ...  
★★★★★ 3,554  
\$9.99  
List: \$11.99 (17% off)  
Get it as soon as Thursday,

Sponsored Computer Keyboard Wired, Plug Play USB Keyboard, Low Profile...  
★★★★★ 119  
[Amazon's Choice](#) in Computer Keyboards  
\$14.59  
Get it as soon as Thursday,

Sponsored Verbatim Slimline Wired Keyboard and Mouse Combo, Optical Wired...  
★★★★★ 5,989  
[Amazon's Choice](#) in Computer Mice  
\$14.99  
Get it as soon as Thursday,

ALTEC Lansing Computer Wired Mouse Ergonomic 6 Button USB Mice with...  
★★★★★ 215  
[Amazon's Choice](#) in Computer Mice & Mouse Combos  
\$14.99  
Get it as soon as Thursday,

Recommendation with explanation

4 stars and above

Rii RK907 Ultra-Slim Compact USB Wired Keyboard for Mac and PC, Windows 10/8 / 7 / ...  
\$9.99 [prime](#)  
★★★★★ 3,551

Computer Keyboard Wired, Plug Play USB Keyboard, Low Profile Chiclet Keys, Large Nu...  
\$14.59 [prime](#)  
★★★★★ 119

Verbatim Slimline Wired Keyboard and Mouse Combo, Optical Wired Mouse, Full-Size Ke...  
\$13.99 [prime](#)  
★★★★★ 5,989

BTO USB Wired Keyboard, 104 Keys with Numeric Pad, Anti Spill and Dust Proof, Slim ...  
\$9.99 [prime](#)  
★★★★★ 604

Basic Keyboard and Mouse, Rii RK203 Ultra Full Size Slim USB Basic Wired Keyboard and K...  
\$14.99 [prime](#)  
★★★★★ 57

Products related to this item

Rii RK907 Ultra-Slim Compact USB Wired Keyboard for Mac and PC, Windows 10/8 / 7 / ...  
★★★★★ 3,551  
\$9.99 [prime](#)

2020 Apple MacBook Air Laptop: Apple M1 Chip, 13" Retina Display, 8GB RAM, 512GB SSD...  
★★★★★ 15,214  
\$1,149.99 [prime](#)

Cordless Electric Air Duster, 2-Gear 43000RPM Portable Air Blower, Replaces...  
★★★★★ 16  
\$59.99 [prime](#)

Manhattan USB Wired Computer Keyboard - with 4.5 ft USB-A Cable, 104-keys, Full-Siz...  
\$15.99 [prime](#)

2021 Apple MacBook Pro (14-inch, Apple M1 Pro chip with 10-core CPU and 16-core GPU)...  
★★★★★ 869  
\$2,349.00 [prime](#)

Customers who searched for "keyboard" ultimately bought

Verbatim Slimline Full Size Wired Keyboard USB Plug-and-Play - Compatible with PC, Laptop - Black  
\$8.95  
List: \$10.99 (11% off)  
Get it as soon as Thursday,

Logitech MK270 Wireless Keyboard and Mouse Combo For Windows, 2.4 GHz Wireless, Compact...  
\$11.99  
#1 Best Seller in Computer Keyboard & Mouse Combos  
\$27.97

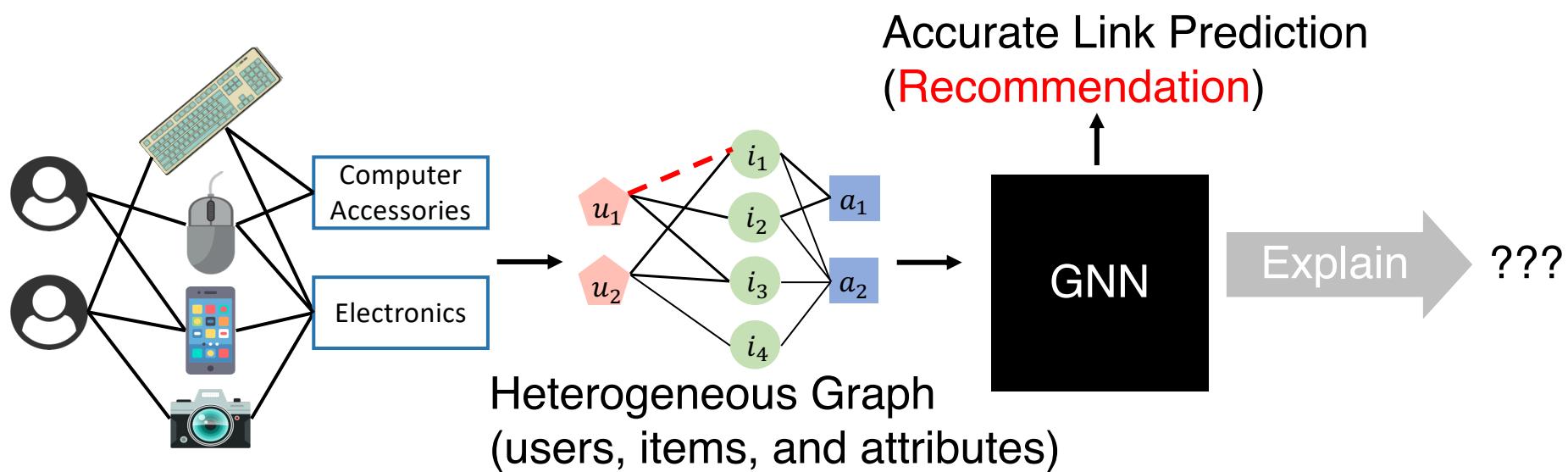
Manhattan Wired Computer Keyboard, Black - Basic Keyboard - with 5ft USB-A Cable, 104-keys, Foldable...  
\$11.99  
Get it as soon as Thursday, Jun 30

Computer Keyboard Wired, Plug Play USB Keyboard, Low Profile Chiclet Keys, Large Number Pad, Caps...  
\$14.59  
Get it as soon as Thursday,

Redragon S101 Wired Gaming Keyboard and Mouse Combo RGB Backlit Gaming Keybo...  
\$35.98  
#1 Best Seller in PC Gaming Keyboards  
Get it as soon as Thursday,

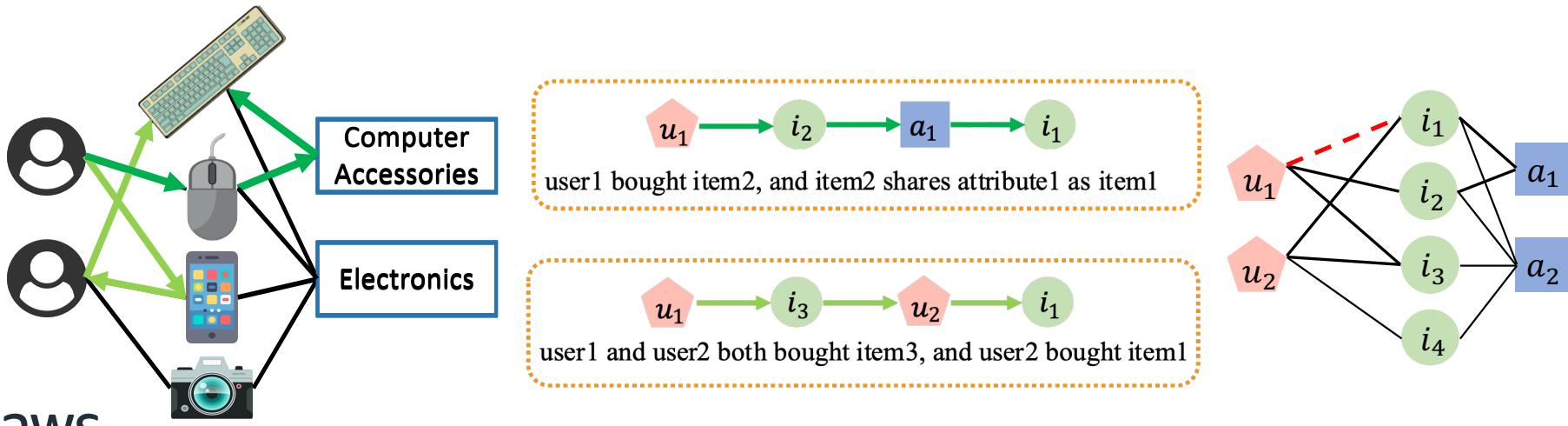
# GNN Explainability

- Data: Heterogeneous graphs.
- Model: GNNs for link prediction.
- Explanation: Why recommend an item to a user? (Why predict a user-item link?)



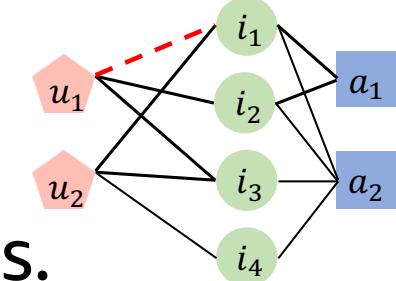
# Main Idea: Paths As Explanations

- Natural human-interpretable explanations boil down to paths.
- Paths form a much smaller search space compared to general subgraphs.
- Define explanations as *concise* and *informative* paths that are *influential to the prediction*.



# PaGE-Link: Path-Enforcing Mask

- Challenges for finding good paths.
  - Many path candidates.
  - Criterion for selecting good paths.
- Learn an edge mask to select meaningful edges.
  - Edges form short paths with low-degree nodes.



$$\mathcal{L}_{path}(\mathcal{M}) = - \sum_{r \in \mathcal{R}} (\alpha \sum_{\substack{e \in \mathcal{E}_{path} \\ \tau(e)=r}} \mathcal{M}_e^r - \beta \sum_{\substack{e \in \mathcal{E}, e \notin \mathcal{E}_{path} \\ \tau(e)=r}} \mathcal{M}_e^r)$$

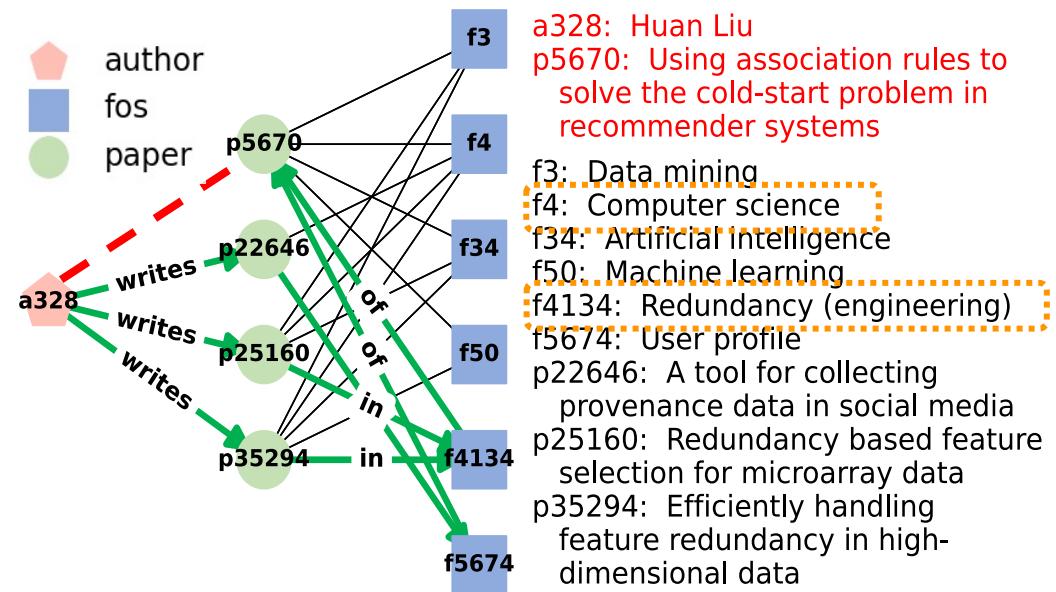
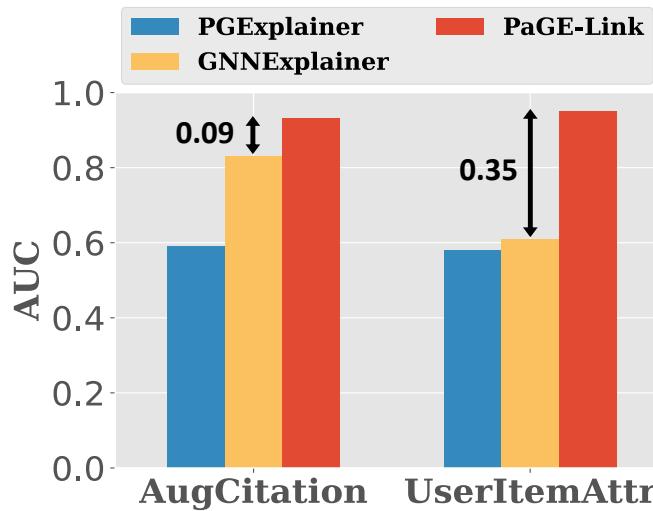
- Edges maximize the mutual information.

$$\mathcal{L}_{pred}(\mathcal{M}) = - \log P_\Phi(Y=1 | \mathcal{G} = (\mathcal{V}, \mathcal{E} \odot \sigma(\mathcal{M})), (s, t))$$

- Pruning: more informative paths and efficiency.

# Experiments

- ROC-AUC: 9%-35% improvement over baselines.
- Concise paths without generic nodes.
- Human evaluation: 78.79% responses selected our method as the best compared to baselines.



# Thank you!

## Q & A

---

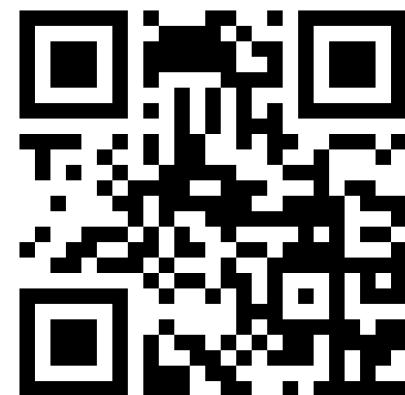
Paper



Code



Contact author



# Appendix

---

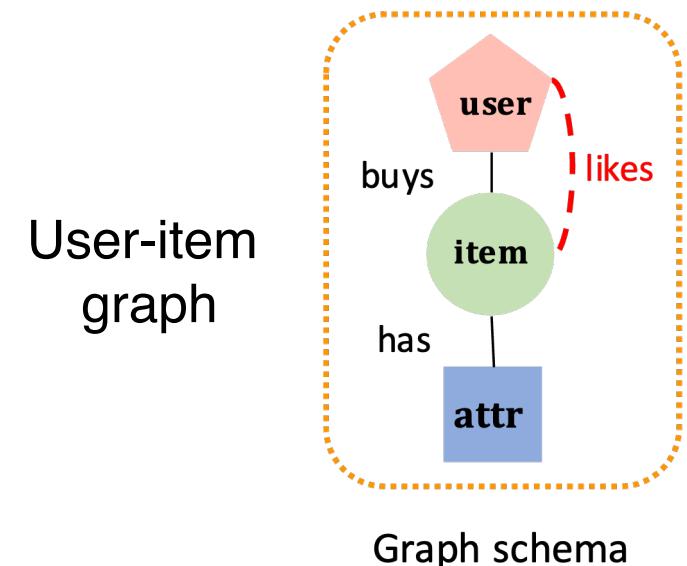
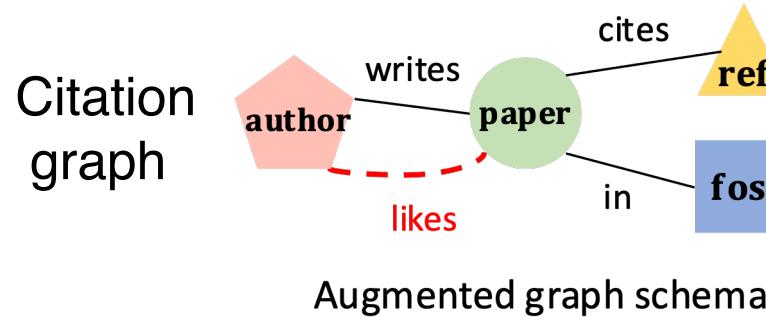
# Experiments: Dataset Generation

Generate datasets new evaluation

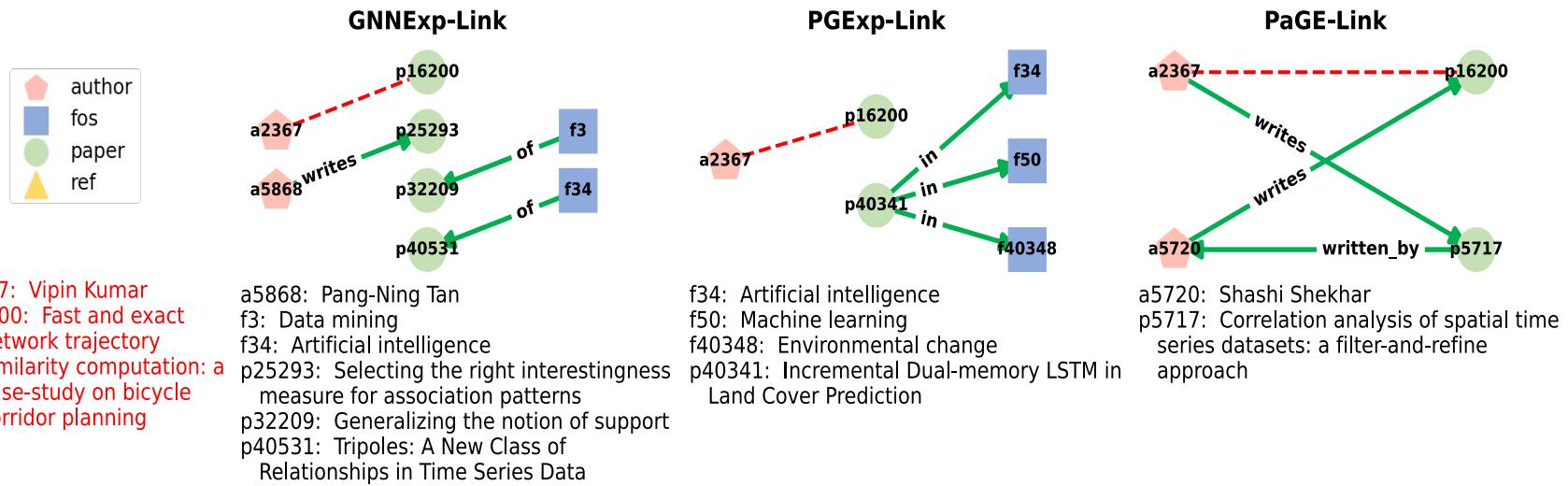
- Create a new edge  $s-t$  if they are connected by a concise and informative path  $p$

$$\mathcal{P} = \{p | p \text{ is a } s-t \text{ path with max length } l_{max} \text{ and max node degree } D_{max}\}$$

- Use  $p$  as the ground truth for evaluating the prediction of  $(s, t)$



# Experiments: Visualization



# Proposition and Theorems

---

- Paths form a much smaller search space

**Proposition 4.1.** *Let  $\mathcal{G}(n, d)$  be a random graph with  $n$  nodes and density  $d$ , i.e., there are  $m = d\binom{n}{2}$  edges chosen uniformly randomly from all node pairs. Let  $Z_{n,d}$  be the expected number of paths between any pair of nodes. Let  $S_{n,d}$  be the expected number of edge-induced subgraphs. Then  $Z_{n,d} = o(S_{n,d})$ , i.e.,  $\lim_{n \rightarrow \infty} \frac{Z_{n,d}}{S_{n,d}} = 0$ .*

- Asymptotic normality of the k-core

**Theorem 5.1** (Pittel, Spencer and Wormald [28]). *Let  $\mathcal{G}(n, d)$  be a random graph with  $m$  edges as in Proposition 4.1. Let  $\mathcal{G}^k(n, d) = (\mathcal{V}^k(n, d), \mathcal{E}^k(n, d))$  be the nonempty  $k$ -core of  $\mathcal{G}(n, d)$ . Then  $\mathcal{G}^k(n, d)$  will contain  $\delta_{\mathcal{V}}(n, d, k)n$  nodes and  $\delta_{\mathcal{E}}(n, d, k)m$  edges with high probability (w.h.p.) for large  $n$ , i.e.,  $|\mathcal{V}^k(n, d)|/n \xrightarrow{p} \delta_{\mathcal{V}}(n, d, k)$  and  $|\mathcal{E}^k(n, d)|/m \xrightarrow{p} \delta_{\mathcal{E}}(n, d, k)$  with  $\xrightarrow{p}$  stands for convergence in probability.*