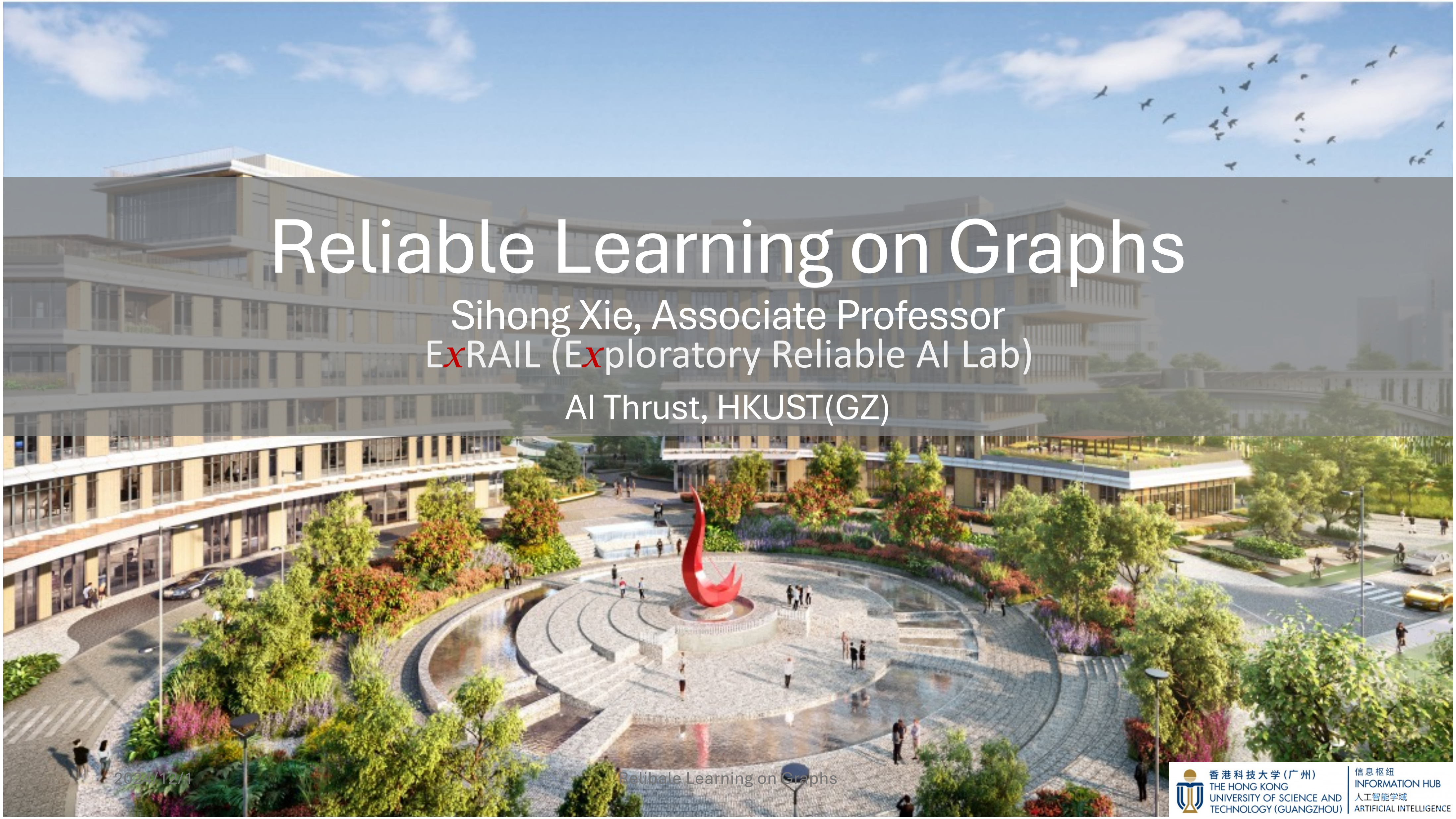


Reliable Learning on Graphs

Sihong Xie, Associate Professor

ExRAIL (Exploratory Reliable AI Lab)

AI Thrust, HKUST(GZ)

香港科技大学(广州)
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)

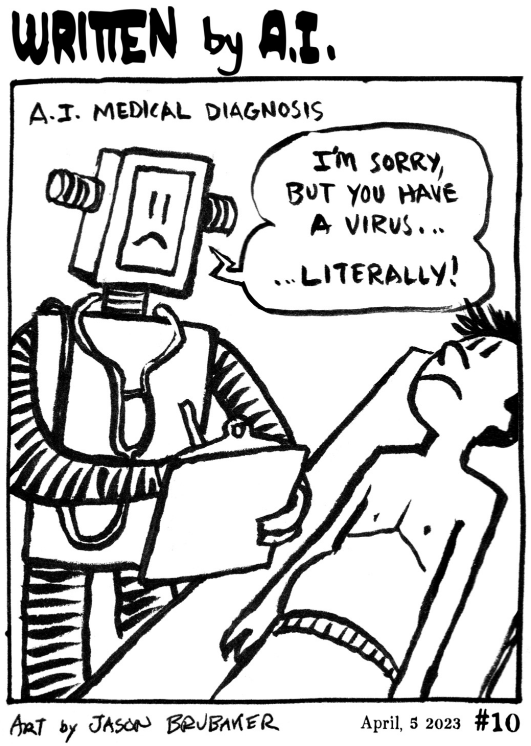
信息枢纽 INFORMATION HUB

人工智能学域 ARTIFICIAL INTELLIGENCE

Reliable machine learning

THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)

When applying AI to clinical medicine



Do you trust the AI medical diagnosis?

What makes you feel better

- Explanation
- Stability
- Confidence

Reliable machine learning

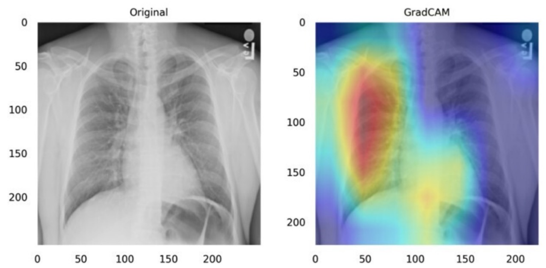
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)

Make ML reliable to humans

-- E*x*RAIL

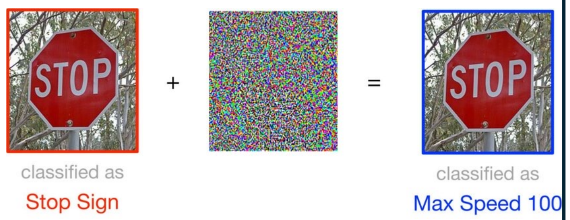
What makes you feel better about AI prediction

Explainability



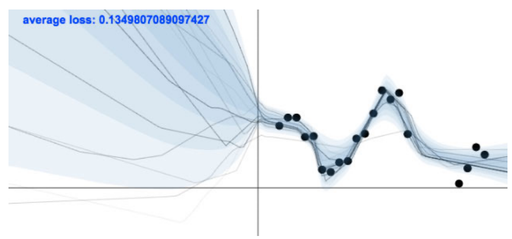
What causes the prediction?

Robustness



Prediction stable to noise?

Confidence



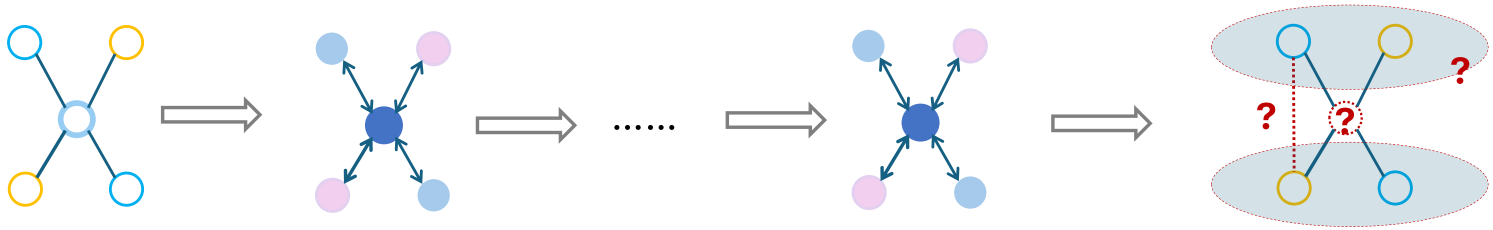
Is the prediction confident?

Reliable learning on graph

THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)

Graph learning pipeline

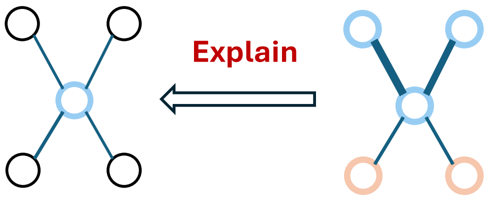
Input | Message propagation and aggregation | Readout | Downstream tasks



Node classification
Link prediction
Community detection

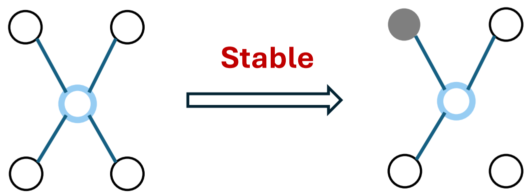
Reliale learning on graph

Explainability



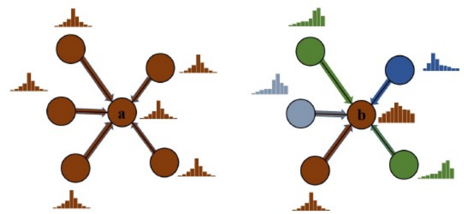
What **edges and nodes** causes the prediction?

Robustness



Prediction stable to **edge and node addition/deletion**?

Confidence



Is the prediction confident given **depedent nodes**?

Framework of reliable graph learning

THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)

Explainability

- Discrete space
- Multi-objective
- Topology-aware
- Dynamic

ICDM'19
- Scalability of explanation search
- Cycle handling
- Multiple explanations

ICDM'21a
- Topological importance
- Multi-objective explanation
- User-study

ICLR'23
- Dynamic graph explanation
- Manifold of explanations
- Convex optimization

NeurIPS'24a
- Stable explanations
- Ranking-based robustness metric
- Bounding for optimization

Robustness

- Game on graph
- Topology constraint
- Unknown attacker

KDD'20
- Multple players
- Markov Game & RL
- Nash equilibria

ICDM'21b
- Siamese network
- Self-supervised
- Constrained optimization

NeurIPS'24b
- Multiple attackers and strategies
- attacking action only
- Inverse RL

Confidence

- Dependent uncertainty

Survey
- Graph uncertainty sources, representation, handling.

NeurIPS'24c
- Uncertainty propagation
- Convergence & bias-var decomp.

Table of Content

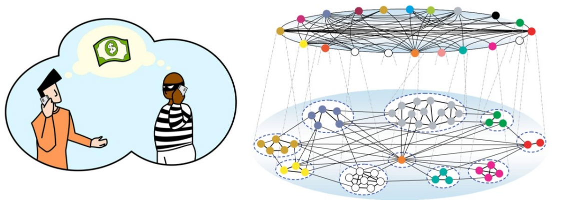
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)

Dynamic Graphs: background

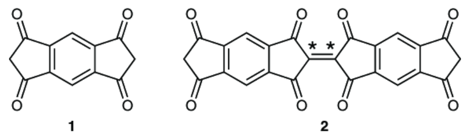
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)

- Graph G can be constantly changing on the **node/edge/attribute** levels.

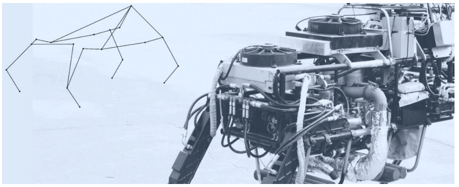
Finance: fraud detection



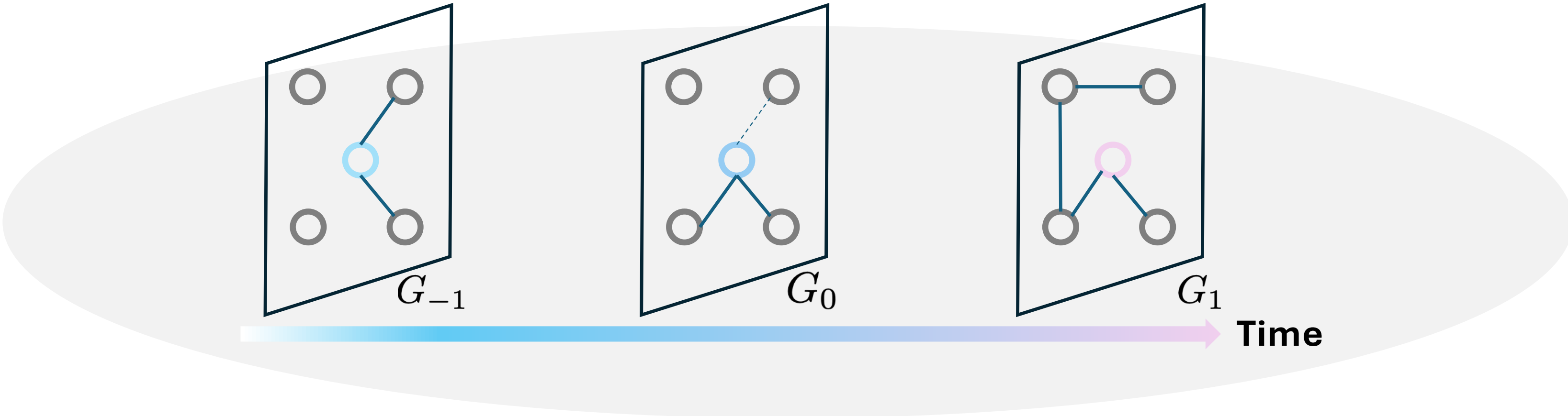
AI4S: molecule design



Robotics: state estimation

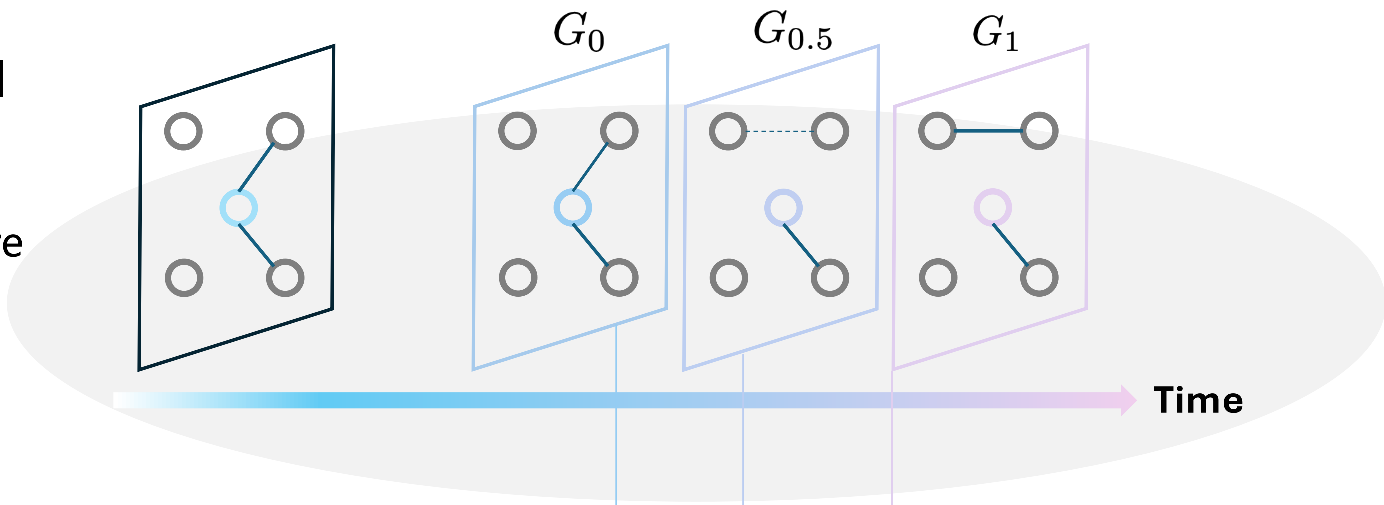


- Predictions $\Pr(Y|G;\boldsymbol{\theta})$ on G changes too

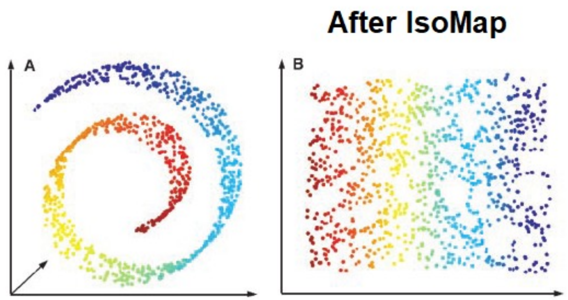


Dynamic Graphs: modeling

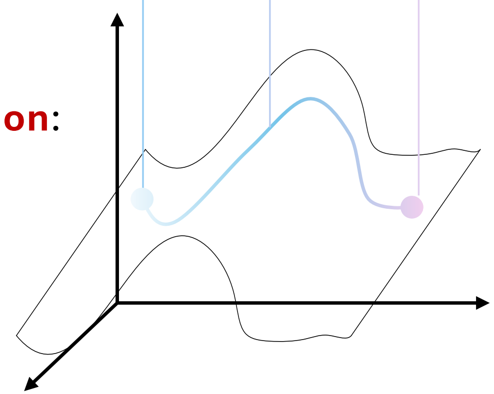
- How a parametric model responses to graph evolution?
  - ➢ Node/edge changes are insufficiently accurate.
  - ➢ What if changes are infinitesimal small?



Manifold: manifolds are smooth mapping, and can reveal intrinsic properties (e.g., distance) of the data.



Node classification: with c classes locally, $\Pr(Y|G;\boldsymbol{\theta})$ is on a (c-1)-dim manifold



Advantages:

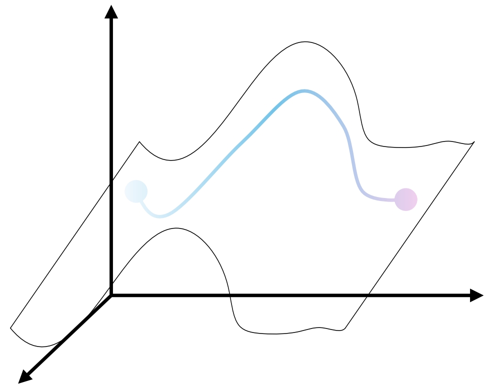
- Smooth manifold
- Fill in the gap
- Differentiable
- Nonlinearity (via. Fisher Information)

Dynamic Graphs: modeling

THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)

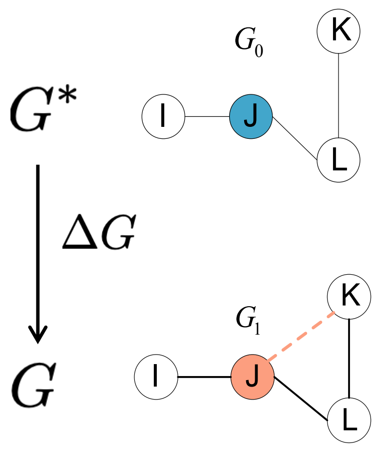
- Information geometry provides a manifold of exponential family.

$$\{\Pr(Y|G) = \mathrm{softmax}(z_1, \ldots, z_c) : \mathbf{z} = \mathrm{logit}(G), \forall G\}$$

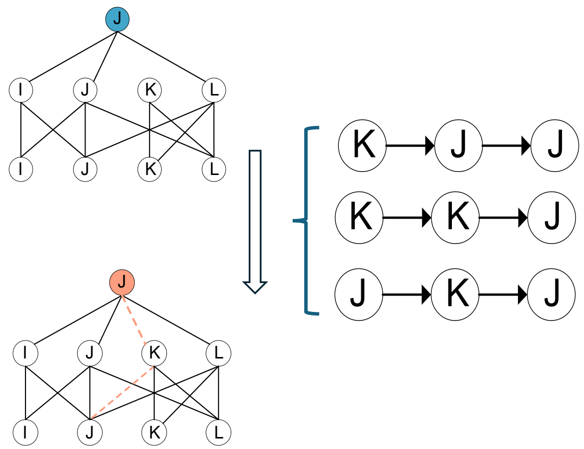


Coordinates	$\mathbf{z} = [z_1, \ldots, z_c]$	A	Path contributions
Gen. linear model	$\mathrm{softmax}(z_1, \ldots, z_c)$	N/A	$\mathrm{softmax}(\mathbf{z}_J(G^*) + \mathbf{1}^\top C_J(G))$
Extrinsic dim	c	$\lvert V\rvert \times \lvert V\rvert$	$m \times c$

Input Graph View



Computation Graph View



Coordinate View

$$[z_1(G^*), \ldots, z_c(G^*)] \qquad \mathbf{z}^*$$

$$+$$

$$\left[\sum_{p=1}^{m} C_{p,1}(G), \ldots, \sum_{p=1}^{m} C_{p,c}(G)\right] \qquad \Big\downarrow \Delta\mathbf{z}$$

$$\parallel$$

$$[z_1(G), \ldots, z_c(G)] \qquad \mathbf{z}$$

Dynamic Graphs: modeling

THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)

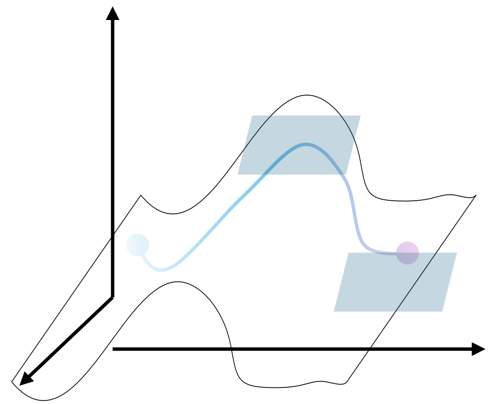
Properties

- The logits and therefore the log-probability is differentiable with respect to the coordinate (path contributions). Define the Fisher Information Matrix

$$I(\mathbf{vec}(C_J(G_1))) = (\nabla_{\mathbf{vec}(C_J(G_1))}\mathbf{z}_J(G_1))^\top \mathbb{E}_{Y\sim \Pr(Y|G_1)}[s_{\mathbf{z}_J(G_1)} s_{\mathbf{z}_J(G_1)}^\top](\nabla_{\mathbf{vec}(C_J(G_1))}\mathbf{z}_J(G_1))$$

- The distance metric on the manifold is curved (non-Euclidean) and adaptive to the local curvature.

$$\mathbf{vec}(\Delta C_J(G_1, G_0))^\top I(\mathbf{vec}(C_J(G_1)))\mathbf{vec}(\Delta C_J(G_1, G_0))$$



- Given $G_0 \to G_1$, define a curve on the manifold

$$\{\Pr(Y|G(s)) : s \in [0,1], \Pr(Y|G(0)) = \Pr(Y|G_0), \Pr(Y|G(1)) = \Pr(Y|G_1)\}$$

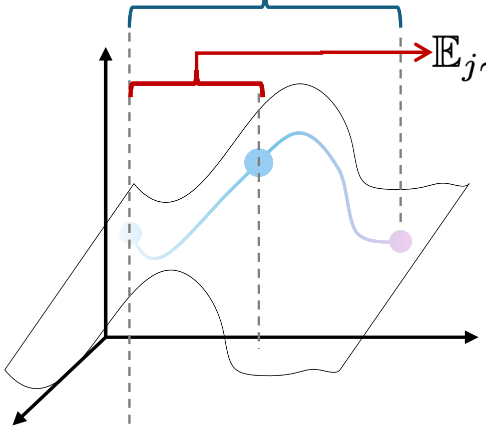
where $\Pr(Y|G(s))$ is differentiable w.r.t. the time variable $s \in [0,1]$

Explanation evoluting graphs

The distance between two distributions is $D_{\mathrm{KL}}(\Pr(Y|G_1)||\Pr(Y|G(s)))$

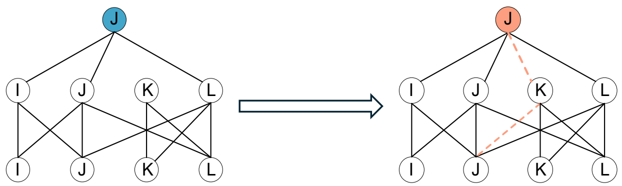
$$\mathbb{E}_{j\sim\Pr(Y|G_1)}[\mathbf{1}^\top(C_{:j}(G_1)-C_{:j}(G_0))]-\log Z(G_1)+\log\sum_{j=1}^{c}\exp\{z_j(G^*)+\mathbf{1}^\top C_{:j}(G_0)\}$$

$$\mathbb{E}_{j\sim\Pr(Y|G_1)}[\mathbf{1}^\top(C_{:j}(G_1)-C_{:j}(G_0))]-\log Z(G_1)+\log\sum_{j=1}^{c}\exp\{z_j(G^*)+\mathbf{1}^\top \color{red}{C_{:j}(G(s))}\}$$

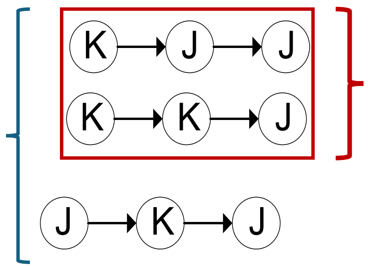


$$\boxed{\min_{\color{red}{G(s)}} D_{\mathrm{KL}}(\Pr(Y|G_1)||\Pr(Y|\color{red}{G(s)}))}$$

Computation Graph View



$G_0 \to G_1$



$G_0 \to \color{red}{G(s)}$

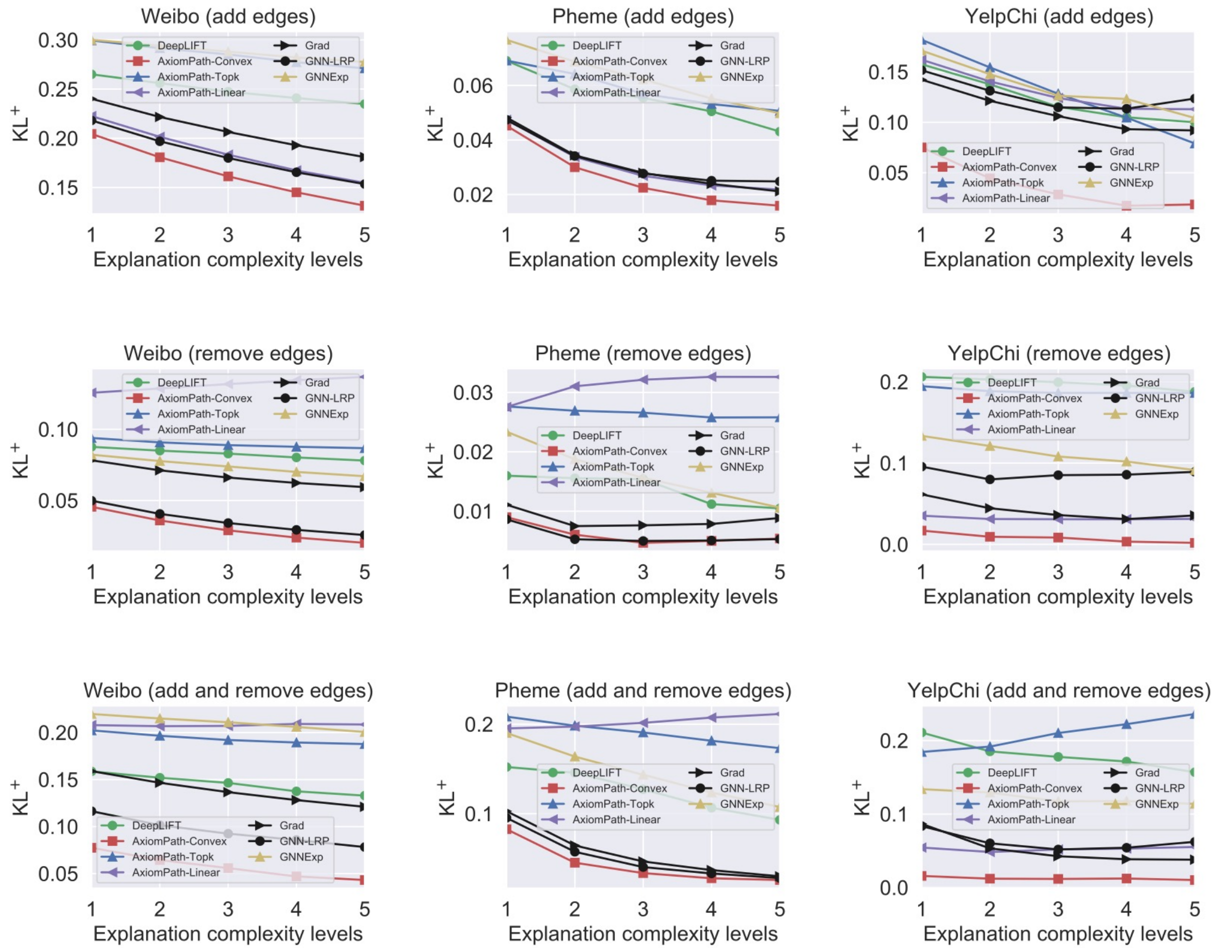
Explanation evoluting graphs

THE HONG KONG
UNIVERSITY OF SCIENCE AND
TECHNOLOGY (GUANGZHOU)

Verified on node classification, link prediction, and graph classification tasks.

8 graph datasets.

Metric: explanation faithfulness (KL⁺) ↓



See the paper
A Differential Geometric View and Explainability of GNN on Evolving Graphs (ICLR 2023)
for more details.

Table of Content

THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)

- Dynamic graph explanation (ICLR’23)
- Robust graph explanation (NeurIPS’24a)
- Learn about attacker on graph (NeurIPS’24b)
- Uncertainty quantification on graph (NeurIPS’24c)

Robust explanation: motivation

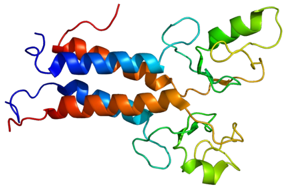
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)

- Robustness of explanations: an explanation won't change due to irrelevant perturbations.

> *Stability has to do with the extent to which a relationship holds across diverse segments of the population (or across various circumstances).*
>
> -- Nadya Vasilyeva, Thomas Blanchard, Tania Lombrozo. *"Stable Causal Relationships Are Better Causal Relationships"*. Cognitive Science 42 (2018) 1265–1296

- A mental experiment about robustness/stability

Gene BRCA1 Mutation



⇨ causes cancer

Gene Gabrb1 Mutation

⇨ causes alcoholism

	Income and social status	
	High	Low
Gene BRCA1 Mutation ⇨ causes cancer	True	True
Gene Gabrb1 Mutation ⇨ causes alcoholism	False	True

Which causal relationship do you trust?

- Many empirical studies: stable relationship under different background is trusted more.

Robust explanation: motivation

- Gradient-based explanation is sensitive to ***irrelevant*** perturbations?

Credit card approval

x

Feature	Value
Age	25
Gender	**Male**
Education	Bachelor
House	Rental
Deposits	Below $5,000
Active cards	3
…	…

$f(x)$: Rejected

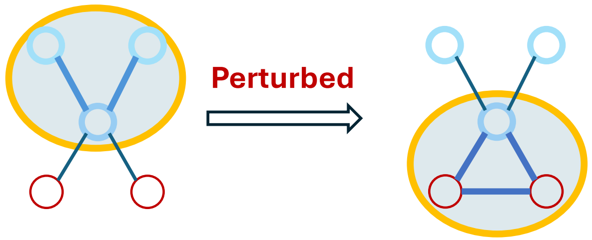
Perturbed ⟹

x'

Feature	Value
Age	25
Gender	**Female**
Education	Master
House	Rental
Deposits	Below $5,000
Active cards	3
…	…

$f(x')$: Rejected

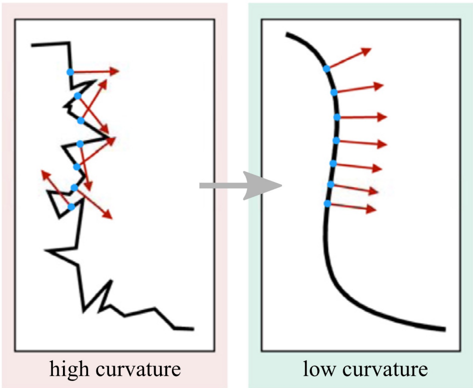
Node classification

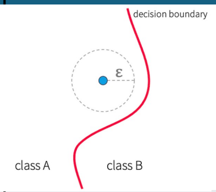
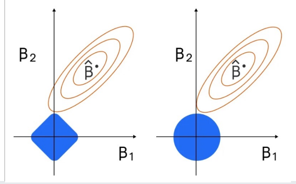
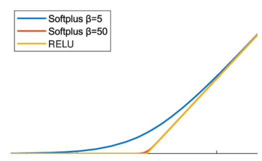


Prediction =

Robust explanation: existing work

- Gradient-based explanations are vulnerable.



Prior work	Adversarial Training	Regularization	Weight decay	Soft-plus
Technique		$\min_{\theta} \mathbb{E}_{(x,y)\sim\mathcal{D}}[L((x,y);\theta) + \beta \max_{i} \xi_i]$		
Weakness	Distance/norms fail to measure explanation robustness			

Features	$g(x)$	$g(x')$	$g(x'')$	$g(x''')$
Age	0.10	0.01	0.25	3.3
Gender	0.06	0.02	0.06	2.0
Education	0.05	0.05	0.05	2.8
House	0.30	0.20	0.15	41.5
Deposits	0.33	0.63	0.33	57.7
Active cards	0.16	0.09	0.16	10.3

- Distance/norm do not reflect ranking invariance.

$$|g(x) - g(x')| > |g(x) - g(x'')|$$

- Distance/norm are sensitive to the scale of the gradient.

$$|g(x) - g(x''')| \gg |g(x) - g(x'')|$$

Robust explanation: a novel metric

THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)

Features	$g(x)$	$g(x')$
Age	0.10	0.01
Gender	0.06	0.02
Education	0.05	0.05
House	0.30	0.20
Deposits	0.33	0.63
Active cards	0.16	0.09

Definition: the distance between the importance scores of features i and j

$$h(x, i, j) = g_i(x) - g_j(x)$$

Example: deposits (i) is more indicative than Gender (j), then $h(x, i, j) > 0$

Invariant 1) the gap remains positive to perturbations

$$\int_0^1 h(x(t), i, j)\, dt > 0$$

Invariant 2) and the gap remains positive for all input

$$\Theta(i, j) = \mathbb{E}_{x' \sim D}\left[\int_0^1 h(x(t), i, j)\, dt\right]$$

Focused on top k important features

Top-k Thickness

$$\Theta(k) = \frac{1}{k(n-k)} \sum_{i=1}^{k} \sum_{j=k+1}^{n} \Theta(i, j)$$

Robust explanation: optimization

THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)

- Thickness is **bounded** by:

$$h(x,i,j) - \frac{\epsilon}{2}\left|H_i(x) - H_j(x)\right|_2 \leq \Theta(i,j)$$
$$\leq h(x,i,j) + \epsilon\left(L_i + L_j\right),$$

where $H_i(x)$ is the i-th row of the Hessian matrix, and $L_i = \max_{x' \in B(x,\epsilon)} |H_i(x')|_2$.

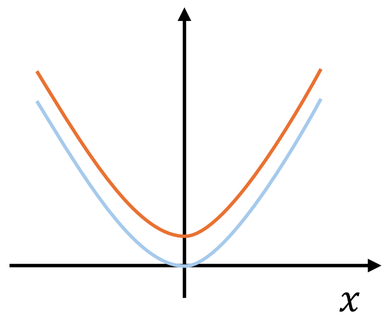
- **R2ET**: train a prediction model, while encouraging a larger gap and smaller Hessian norm.

$$\min_{\theta} \mathcal{L}_{cls} - \lambda_1 \mathbb{E}_x \left[\sum_{i}^{k}\sum_{j}^{n} h(x,i,j)\right] + \lambda_2 \mathbb{E}_x |H(x)|_2$$

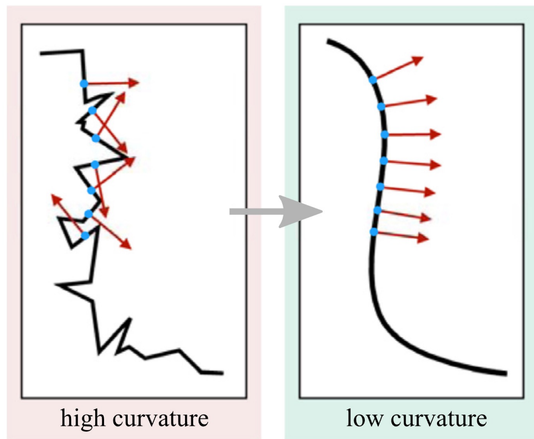
Max the gap locally

0.01
0.02
0.05
0.20
0.63
0.09

Same speed of gradient Change to maintain the positive gap



Smooth the curve



Experiments

- Experimental results on image and graphs (with ***many*** features)

Maintaining *all* "important" features on the top is difficult when the explanation functions are not smooth.

Intuition: the ranking changes a lot, leading to many local optima.

Method	MNIST	CIFAR-10	ROCT	ADHD	BP
# of features	28*28	32*32	771*514	6555	3240
Vanilla	59.0 / 64.0	66.5 / 68.3	71.9 / 77.7	45.5 / 81.1	69.4 / 88.9
WD	59.1 / 64.8	64.2 / 65.6	77.2 / 68.9	47.6 / 79.4	69.4 / 88.6
SP	62.9 / 66.9	67.2 / 71.9	73.9 / 69.5	42.5 / 81.3	68.7 / 90.1
Est-H	85.2 / 90.2	77.1 / 78.7	78.9 / **78.0**†	58.2 / 83.7	(~~75.0 / 91.4~~)*
Exact-H	- / -	- / -	- / -	- / -	- / -
SSR	- / -	- / -	- / -	- / -	- / -
AT	56.0 / 63.9	61.6 / 66.8	78.0 / 72.9	59.4 / 81.0	72.0 / 89.0
R2ET$_{\backslash H}$	82.8 / 89.7	67.3 / 72.2	**79.4** / 70.9	60.7 / 86.8	70.9 / 89.5
R2ET-mm$_{\backslash H}$	81.6 / 89.7	77.7 / **79.4**†	77.3 / 60.2	64.2 / 88.8	72.4 / 91.0
R2ET	**85.7** / 90.8	75.0 / 77.4	79.3 / 70.9	**71.6**† / **91.3**†	71.5 / 89.9
R2ET-mm	85.3 / **91.4**†	**78.0**† / 79.1	79.1 / 68.3	58.8 / 87.5	**73.8**† / **91.1**†

Optimizing Hessian-related terms make ranking easier (smoother) to find better optima.

Experiments

THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)

- Experimental results on tabular data (with ***fewer*** features)

Minimizing Hessian norm may be harmful.

An experimental observation: smaller Hessian norm more likely to result in smaller gradient magnitude (and gap).

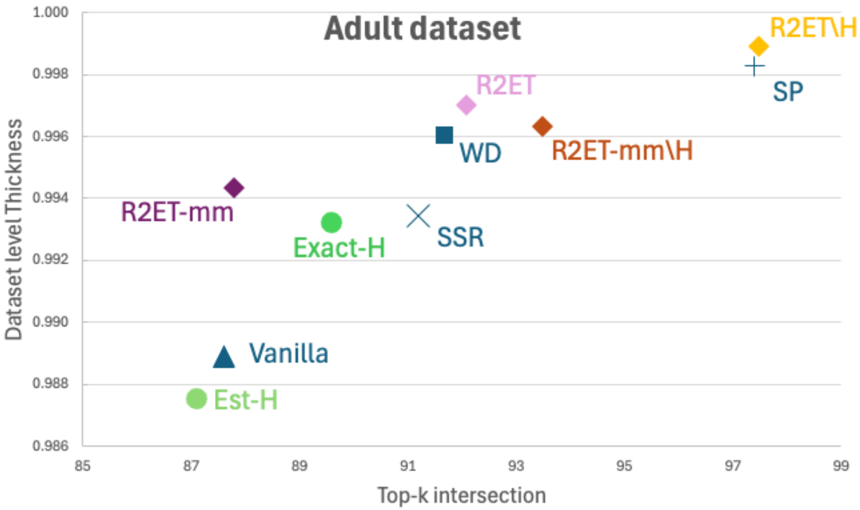
Broadening gaps only is good enough.

Method	Adult	Bank	COMPAS
# of features	28	18	16
Vanilla	87.6 / 87.7	83.0 / 94.0	84.2 / 99.7
WD	91.7 / 91.8	82.4 / 85.9	87.7 / 99.4
SP	97.4 / 97.5	95.4 / 95.5	**99.5**† / **100.0**
Est-H	87.1 / 87.2	78.4 / 81.8	82.6 / 97.7
Exact-H	89.6 / 89.7	81.9 / 85.6	77.2 / 96.0
SSR	91.2 / 92.6	76.3 / 84.5	82.1 / 97.2
AT	68.4 / 91.4	80.0 / 88.4	84.2 / 90.5
R2ET$_{\setminus H}$	**97.5** / **97.7**	**100.0**† / **100.0**†	91.0 / 99.2
R2ET-mm$_{\setminus H}$	93.5 / 93.6	95.8 / 98.2	95.3 / 97.2
R2ET	92.1 / 92.7	80.4 / 90.5	92.0 / 99.9
R2ET-mm	87.8 / 87.9	75.1 / 85.4	82.1 / 98.4

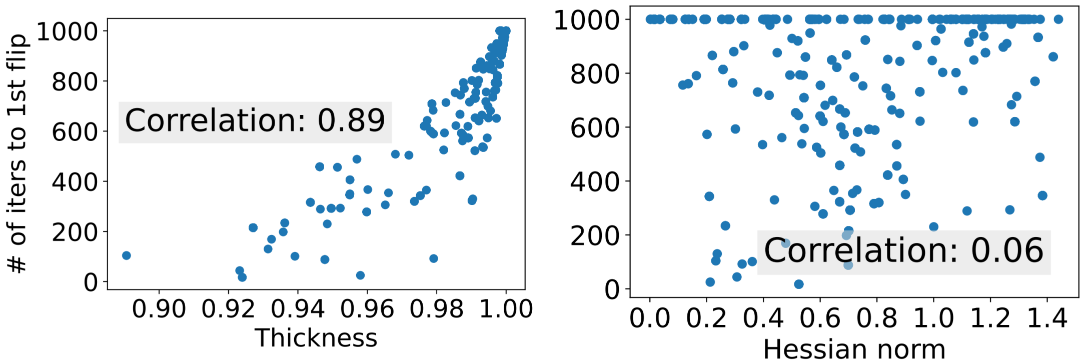
Experiments

- **Thickness** pinpoints the fundamental metric for explanation robustness.

Why R2ET *may not be* the best?



- Each dot is a sample data
- x-axis: thickness (*left*) or hessian norm (right).
- y-axis: the number of iterations needed to **manipulate** any ranking.



Higher thickness leads to better robustness.

R2ET does not have the highest thickness.

Compared with **Hessian norm**, **thickness** shows significantly closer relationship to explanation robustness.

See ***Training for Stable Explanation for Free*** (NeurIPS 2024) for more details.

Table of Content

THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)

- Dynamic graph explanation (ICLR’23)
- Robust graph explanation (NeurIPS’24a)
- Learn about attackers on graph (NeurIPS’24b)
- Uncertainty quantification on graph (NeurIPS’24c)

Secure learning by learning the opponents

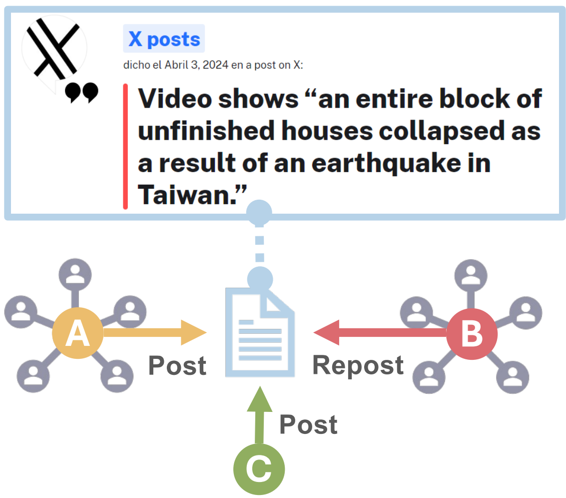
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)

「知己知彼，百戰不殆；不知彼而知己，一勝一負；不知彼，不知己，每戰必殆。」
--《孫子兵法. 謀攻篇》

Know yourself and your enemy, and you will be victorious in every battle.
-- Sun Tzu's Art of War

The underlying strategy is a mixture and unknown to us

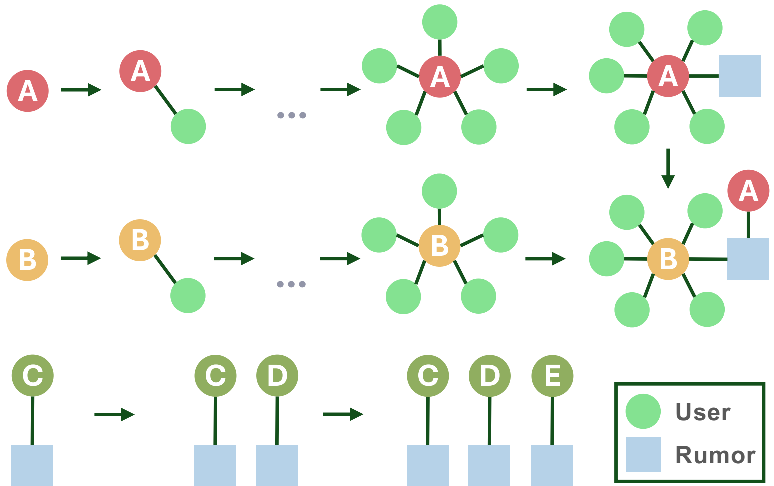
An X Rumor in the Social Graph



Attack Styles

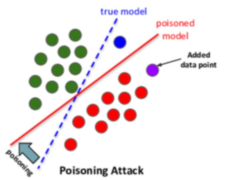
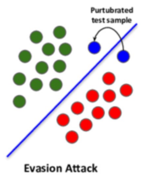
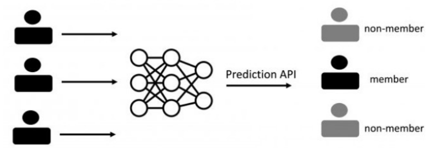
Observable trajectory data

Attack Sequences

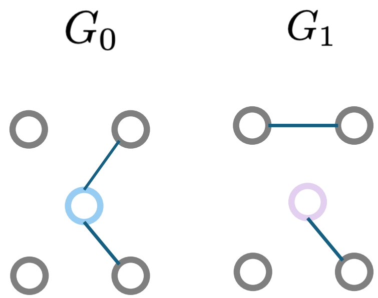


Secure learning on graph

- Attackers know about and can edit the graph
  - ✓ Add reviews to a product;
  - ✓ Friend an account;
  - ✓ Create new accounts;
  - ✓ Modify account profile.
- Knowledge about attackers
  - ✓ Generate attacking samples for adversarial training;
  - ✓ Help humans understand weaknesses of the algorithm.
- Attacking a model

<table>
<tr><th>Attack</th><th>Poisoning</th><th>Evasion</th><th>Backdoor</th><th>Membership Inference</th></tr>
<tr><td>Technique</td><td>
</td><td>
</td><td>
</td><td>
</td></tr>
<tr><td>Weakness</td><td colspan="4">Cannot handle discrete attacking on graphs</td></tr>
</table>

RL is useful for graph security

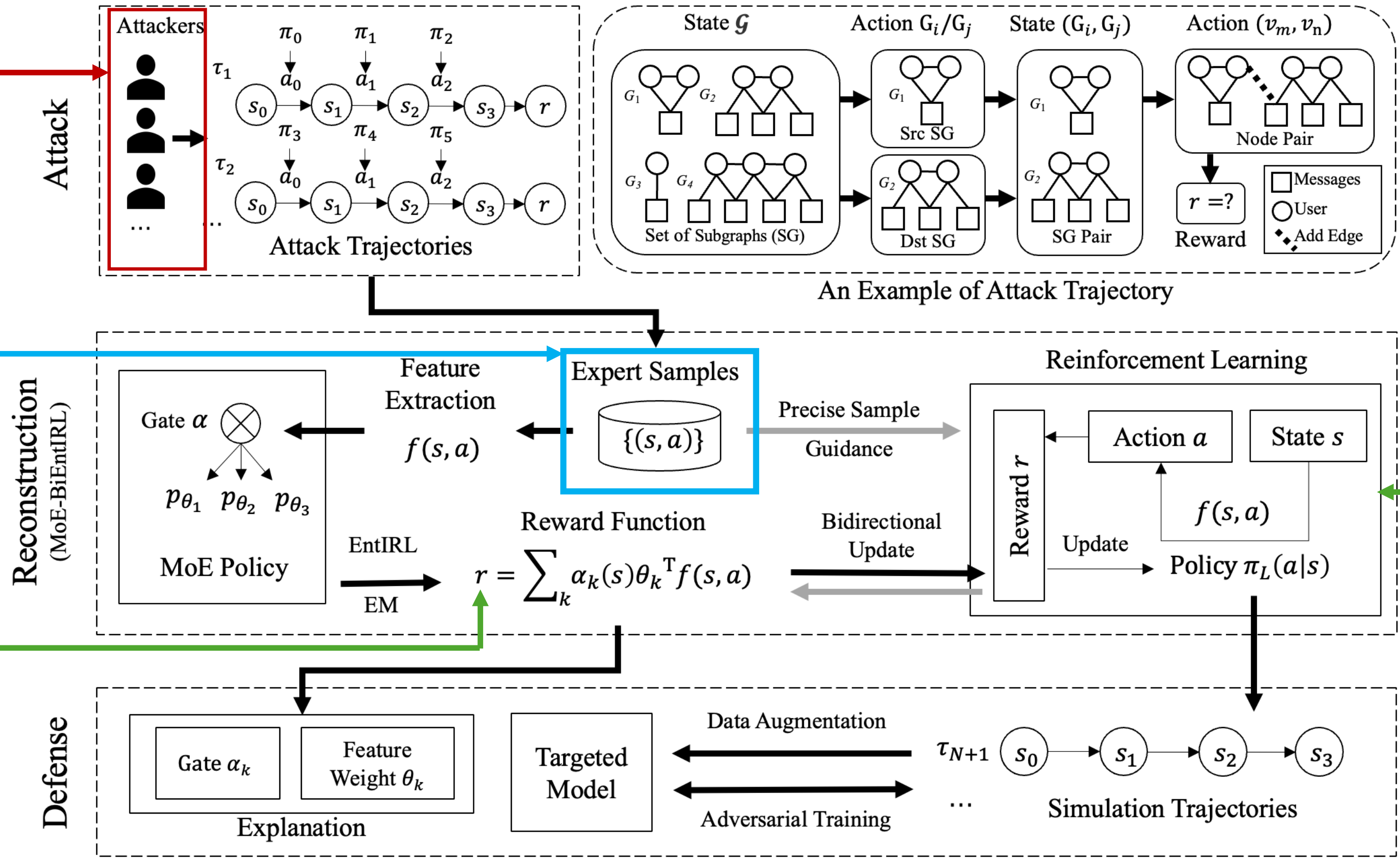


Framework

Their underlying strategy is a mixture and unknown to us.

Attacking trajectories are observable.

Learn the reward



Mimic attacker policy

Details

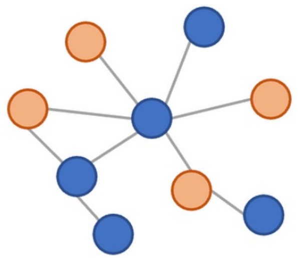
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)

IRL method	MaxEntIRL	MoE
Policy	$p(a\|s) = \frac{1}{Z}\exp(r_\theta(s,a)),$	$p(a^{(t)}\|s^{(t)},\theta) = \sum_{k=1}^{K}\alpha_k(s^{(t)},\varphi)p(a^{(t)}\|s^{(t)},\theta_k),$ $p(a^{(t)}\|s^{(t)},\theta_k) = \frac{\exp(\theta_k^\top f(s^{(t)},a^{(t)}))}{\sum_{a\in\mathcal{A}_{s,t}}\exp(\theta_k^\top f(s^{(t)},a))},$
Reward	$r_\theta(s,a) = \theta^\top f(s,a)$	$r_\theta(s,a) = \sum_{k=1}^{K}\alpha_k(s)\theta_k^\top f(s,a)$
Learning algorithm	$\max_\theta \sum_{s,a} \log p(a\|s,\theta)$	$\hat{\gamma}_{jkt} = P(\gamma_{jkt}=1\|a_j^{(t)},s_j^{(t)},\theta^{(i)}) = \frac{\alpha_k(s_j^{(t)})p\left(a_j^{(t)}\|s_j^{(t)},\theta_k^{(i)}\right)}{\sum_{k=1}^{K}\alpha_k(s_j^{(t)})p(a_j^{(t)}\|s_j^{(t)},\theta_k^{(i)})}$ $L_{gate}(\varphi) = \sum_{t=0}^{T-1}\sum_{k=1}^{K}\sum_{j=1}^{N}\hat{\gamma}_{jkt}\log\alpha_k(s_j^{(t)}),$ $L_{ex}(\theta_k) = \sum_{t=0}^{T-1}\sum_{j=1}^{N}\hat{\gamma}_{jkt}\log p(a_j^{(t)}\|s_j^{(t)},\theta_k).$

EM algorithm: Latent variables indicating which expert generates which action.

Experiment results

THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)



- Training IRL
- Testing resulting attacking π

- Trajectory generating methods

Expert Samples

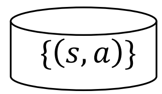


Table 1: Dataset statistics.

	Weibo	Pheme
Nodes	10,280	2,708
Edges	16,412	4,401
Rumors	1,538	284
Non-rumors	1,849	859
Users	2,440	1,008
Comments	4,453	557

	Inverse RL baselines	High-Cost Attack			Low-Cost Attack		
		PRBCD	*AdRumor*	Mixture	*PageRank*	*GC-RWCS*	Mixture
Weibo T=5	Expert	4.865	4.877	-	3.000	3.000	-
	Apprenticeship	1.275	0.788	0.704	0.850	0.763	1.071
	EntIRL	4.650	4.770	4.550	**5.000**	**4.950**	**4.950**
	MoE-BiEntIRL	**4.989**	**4.990**	**4.929**	4.860	4.900	4.900
Weibo T=20	Expert	19.521	19.854	-	5.449	5.160	-
	Apprenticeship	1.142	3.066	3.945	0.030	0.040	0.020
	EntIRL	19.030	19.749	19.199	19.830	**20.000**	**20.000**
	MoE-BiEntIRL	**19.876**	**19.936**	**19.979**	**19.970**	19.700	18.749
Pheme T=5	Expert	4.804	5.947	-	2.991	3.990	-
	Apprenticeship	1.788	3.387	2.619	0.000	0.000	0.000
	EntIRL	0.000	0.018	0.010	0.000	0.062	0.000
	MoE-BiEntIRL	**2.205**	**4.965**	**4.277**	**1.488**	**2.105**	**1.549**

- Evaluation metric
  - increase in rumor detection error ↑

See "***Enhancing Robustness of Graph Neural Networks on Social Media with Explainable Inverse Reinforcement Learning***", *NeurIPS'24* for more details.

Table of Content

THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)

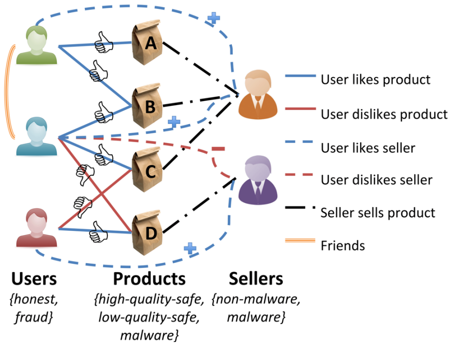
- Dynamic graph explanation (ICLR’23)
- Robust graph explanation (NeurIPS’24a)
- Learn about attacker on graph (NeurIPS’24b)
- Uncertainty quantification on graph (NeurIPS’24c)

Uncertainty quantification on graphs

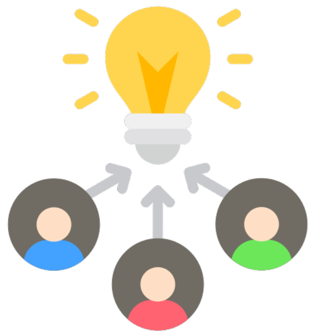
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)

- Quantify the uncertainty of graph inference results can be useful.
  - Graph inference can be applied to link prediction, node classification, label denoising.

Social Network Modeling
don't recommend a friend if the inferred mutual interest is not confident.



Fraud Detection[1]
suspicious users or sellers with high confidence should be filtered.

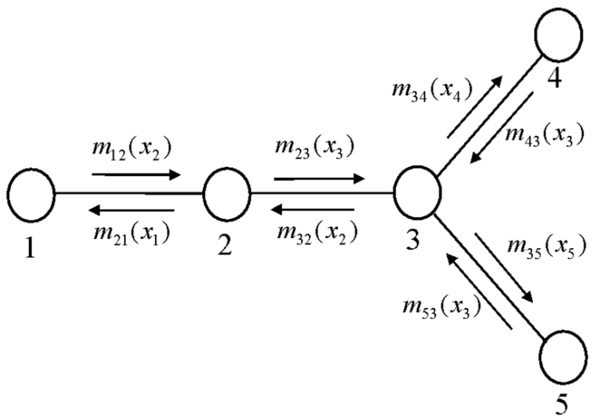


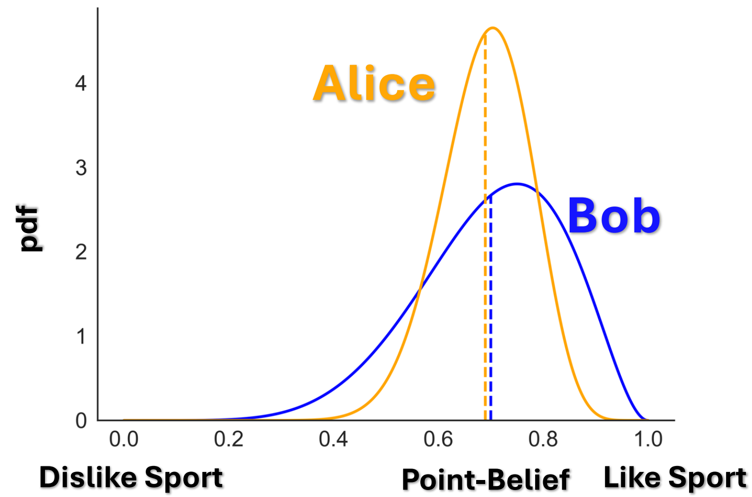
Crowdsourcing
ask for human labeling if the crowdsourcing workers are not confident in the annotation.

Github Homepage

Uncertainty quantification on graphs

- Graphical model inference
  - Belief Propagation (BP) estimates the posterior probability of a node's classes.
  - BP only provides point estimates, failing to capture uncertainty in predictions.



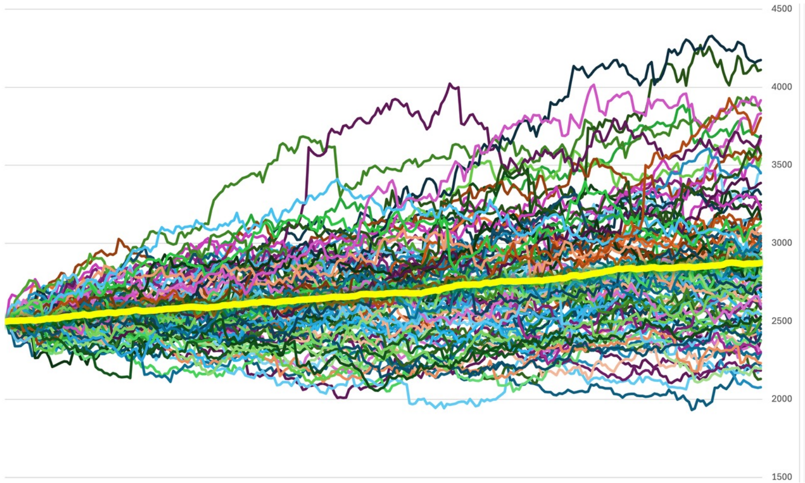


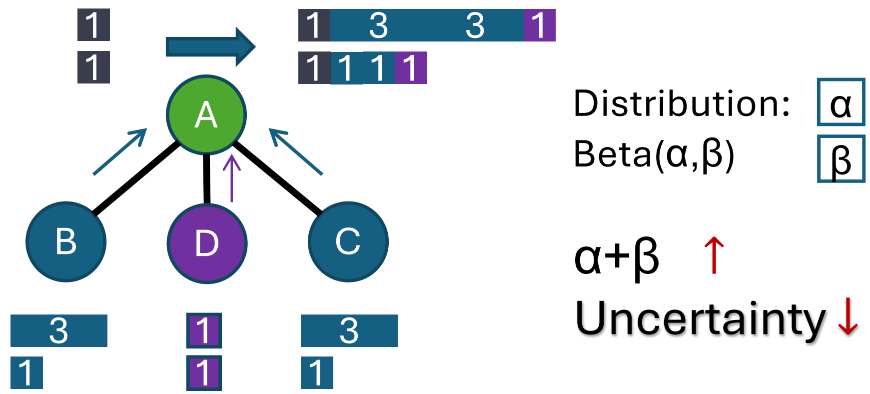
- Why there is uncertainty
  - Imagine: the nodes prior is only a sample from a distribution.
  - Sampling the priors multiple times can result in different poster distributions.

Existing work

THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)

- Monte Carlo sampling
  - Pros: Unbiased uncertainty estimates, general, easy to implement
  - Cons: Time-consuming for large-scale graphs, and **no convergence guarantee**.
- Propagation of uncertainty
  - Pros: Bayesian point of view with rigorous proof, likely to **converge**.
  - Cons: make assumption about the dist. form (multi-nomial) and can be biased.

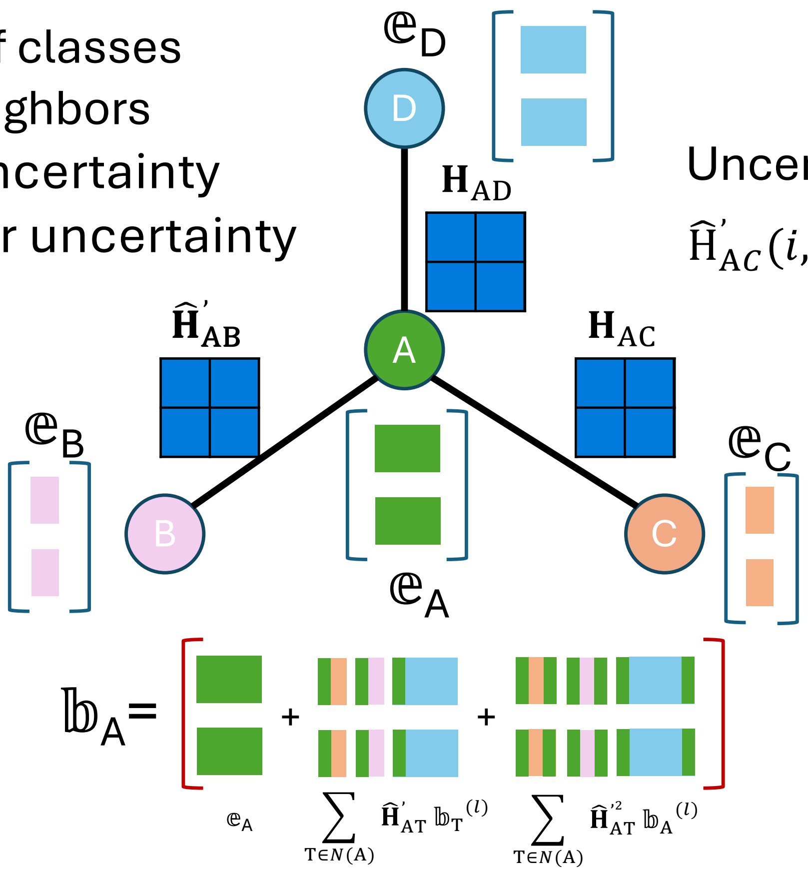




Can we have the best of both worlds?

Linear Uncertainty Propagation (LinUProp)

k: number of classes
$N(\mathrm{A})$: A's neighbors
$\mathbb{e}_{\mathrm{A}}$: initial uncertainty
$\mathbb{b}_{\mathrm{A}}$: posterior uncertainty



Uncertainty Dependency Matrices

$$\widehat{\mathrm{H}}'_{\mathrm{A}C}(i,j) = |\mathrm{H}_{\mathrm{AC}}(i,j) - 1/k\,|$$

Posterior uncertainty

$$\mathbb{b}_{\mathrm{A}}{}^{(l+1)} = \mathbb{e}_{\mathrm{A}} \qquad \text{(initial uncertainty)}$$

$$+ \sum_{\mathrm{T}\in N(\mathrm{A})} \widehat{\mathbf{H}}'_{\mathrm{AT}}\, \mathbb{b}_{\mathrm{T}}{}^{(l)} \qquad \text{(first-order Propagated uncertainty)}$$

$$+ \sum_{\mathrm{T}\in N(\mathrm{A})} \widehat{\mathbf{H}}'^{2}_{\mathrm{AT}}\, \mathbb{b}_{\mathrm{A}}{}^{(l)} \qquad \text{(second-order Propagated uncertainty)}$$

$\mathbb{b}_{\mathrm{T}}{}^{(0)}$ and $\mathbb{b}_{\mathrm{A}}{}^{(0)}$ can be set to $\mathbb{e}_{\mathrm{T}}{}^{(0)}$ and $\mathbb{e}_{\mathrm{A}}{}^{(0)}$, respectively

Linear Uncertainty Propagation (LinUProp)

THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)

Theoretical properties

- Matrix form: $\mathrm{vec}(\mathbb{B}) = (\mathbf{I} - \underbrace{(\mathbf{\Psi}_1' + \mathrm{Diag}(\mathbf{\Psi}_2'\mathbf{Q}))}_{\mathbf{T}})^{-1} \cdot \mathrm{vec}(\mathbb{E})$
- Convergence: LinUProp converges $\Leftrightarrow \rho(\mathbf{T}) < 1$
- Interpretability: $\mathrm{vec}(\mathbb{B}) = (\mathbf{I} + \mathbf{T} + \mathbf{T}^2 + \ldots) \cdot \mathrm{vec}(\mathbb{E})$

$$c_{w \to v} = \mathbf{T}_{v,w}\,\mathrm{vec}(\mathbb{E})_w + (\mathbf{T}^2)_{v,w}\,\mathrm{vec}(\mathbb{E})_w + (\mathbf{T}^3)_{v,w}\,\mathrm{vec}(\mathbb{E})_w + \cdots$$

- Bias–variance decomposition: $\mathbb{E}\left[\left(h(\hat{\mathbf{E}}) - \mathrm{vec}(\hat{\mathbf{B}})_v\right)^2\right] = \underbrace{\left(h(\hat{\mathbf{E}}) - \mathbb{E}\left[\mathrm{vec}(\hat{\mathbf{B}})_v\right]\right)^2}_{(\mathrm{Bias})^2} + \underbrace{\mathbb{E}\left[\left(\mathrm{vec}(\hat{\mathbf{B}})_v - \mathbb{E}\left[\mathrm{vec}(\hat{\mathbf{B}})_v\right]\right)^2\right]}_{\mathrm{Variance}}$

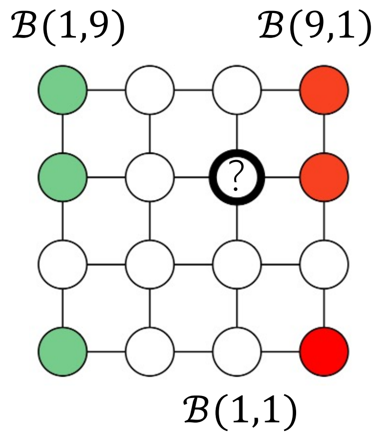
True uncertainty

LinUProp uncertainty

Experiments

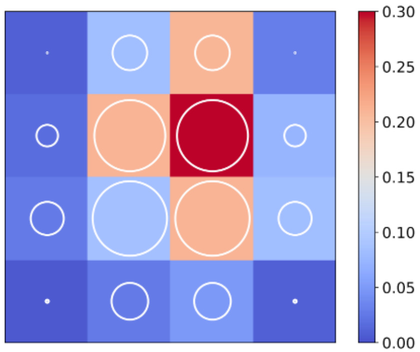
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)

Toy example graph



$\mathcal{B}(\alpha, \beta)$: Prior Beta Distribution with Parameters α and β

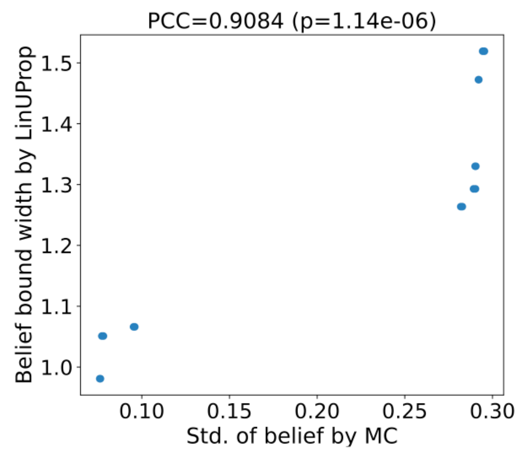
Inferred uncertainty



Circle size:
Uncertainty Bound

Color:
Contribution to Bolded

Correlation(LinUProp, MC)

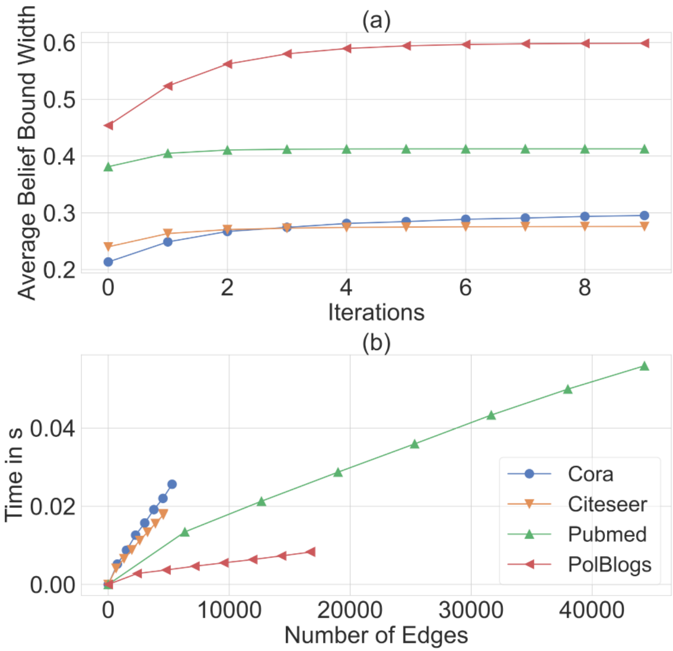


Experiments

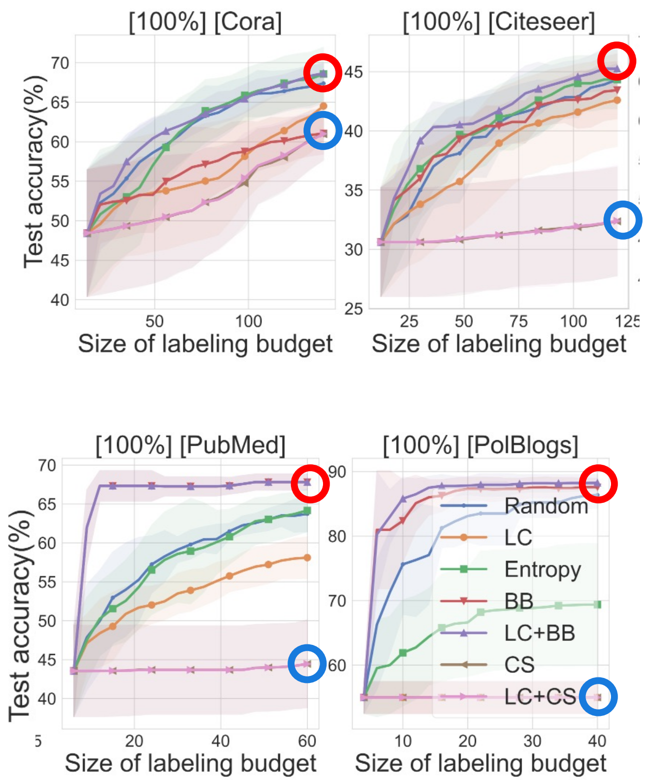
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)

Convergence

Scalability



Active learning



Conclusions and future work

- Conclusions
  - Reliability = {Explainability, Robustness, Confidence, …}.
  - Graphs provide a research avenue with many problems.
  - Dependencies make reliability harder to achieved.
- Future work
  - LLM and graph foundation model have more obstacles.
  - Embodied AI that uses graph required reliability.
  - Multi-modality: graph+X

Thank you!