# Graph Ricci Flow and Applications in Network Analysis and Learning

Jie Gao

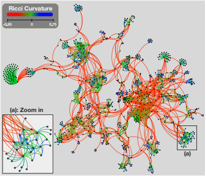
Computer Science, Rutgers University
http://sites.rutgers.edu/jie-gao

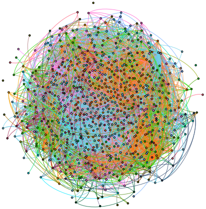
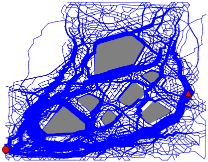
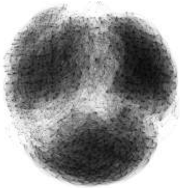
Jan 4th 2024

# Graph Data and Complex Networks

Complex networks in nature: social networks, biological networks, the Internet, WWW, mobility data.

- Small world phenomena
- Power law degree distribution
- Community structures (clustered, closely knit groups).



# Analyzing Graph Data and Complex Networks

Understand a single network:

- Community detection.
- Graph learning (label propagation & prediction)

# Analyzing Graph Data and Complex Networks

Understand a single network:

- Community detection.
- Graph learning (label propagation & prediction)

Understand a family of networks:

- Network alignment.
- Graph generative model.

# Analyzing Graph Data and Complex Networks

Understand a single network:

- Community detection.
- Graph learning (label propagation & prediction)

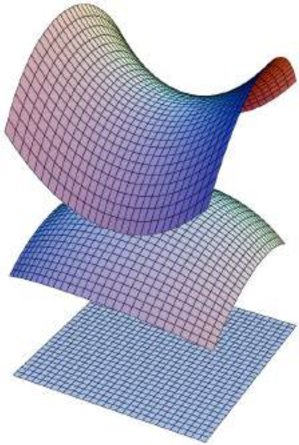
Understand a family of networks:

- Network alignment.
- Graph generative model.

Our project: use geometric tools, Ollivier Ricci curvature flow, to analyze complex networks.

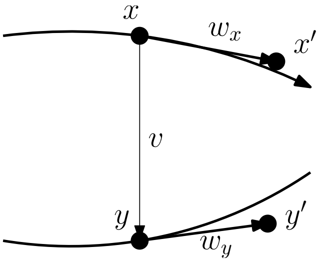
# Curvature in Geometry

Sphere: positive curvature; Plane: zero curvature; Hyperbolic plane: negatie curvature.



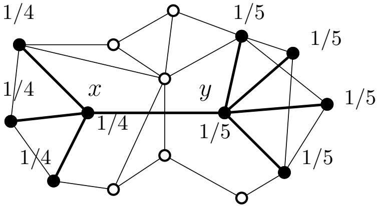
# Sectional Curvature in Geometry

Consider a tangent vector $v = xy$ and another tangent vector $w_x$ at $x$. Transport $w_x$ along $v$ to be a tangent vector $w_y$ at $y$.
If $|x'y'| < |xy|$, then sectional curvature is positive.



# Ollivier Ricci Curvature

Take the analog: for an edge $xy$, consider the "distances" from $x$**'s neighbors** to $y$**'s neighbors** and compare it with the length of $xy$.




How to compute the "distances" between two neighborhoods? Use the optimal transport distance.

# Ollivier Ricci Curvature

## Definition (Ollivier)

Let $(X, d)$ be a metric space and let $m_1, m_2$ be two probability measures on $X$. For any two distinct points $x, y \in X$, the (Ollivier-) Ricci curvature along $xy$ is defined as

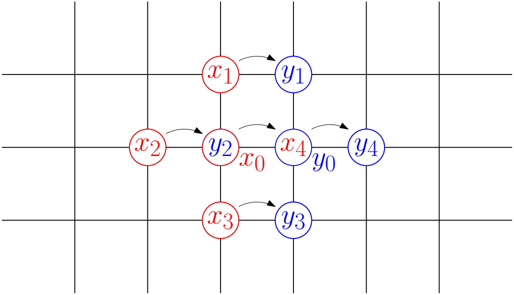
$$\kappa(x, y) := 1 - \frac{W_1(m_x, m_y)}{d(x, y)},$$

where $m_x$ ($m_y$) is a probability distribution defined on $x$ ($y$) and its neighbors, $W_1(\mu_1, \mu_2)$ is the $L_1$ **optimal transportation distance** between two probability measure $\mu_1$ and $\mu_2$ on $X$:

$$W_1(\mu_1, \mu_2) := \inf_{\psi \in \Pi(\mu_1, \mu_2)} \int_{(u,v)} d(u, v) d\psi(u, v)$$

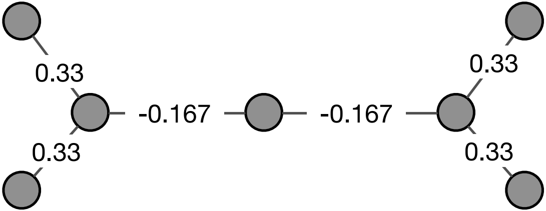
# Examples

Zero curvature: 2D grid.



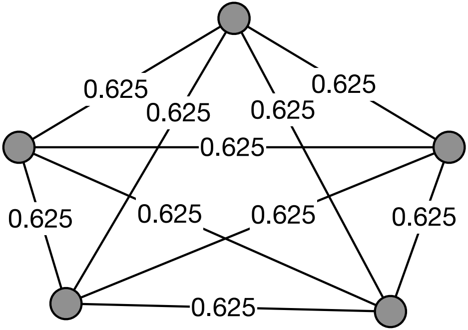
# Examples

Negative curvature: tree: $\kappa(x, y) = 1/d_x + 1/d_y - 1$, $d_x$ is degree of $x$.



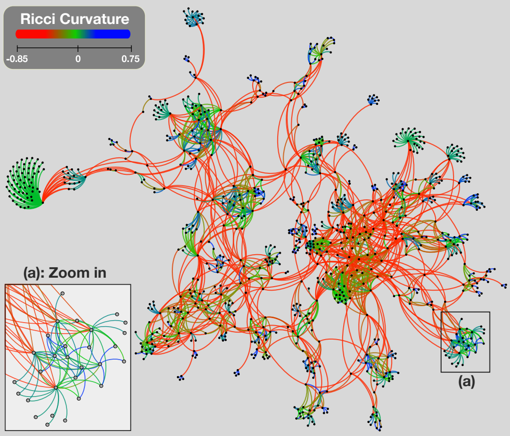
# Examples

Positive curvature: complete graph.



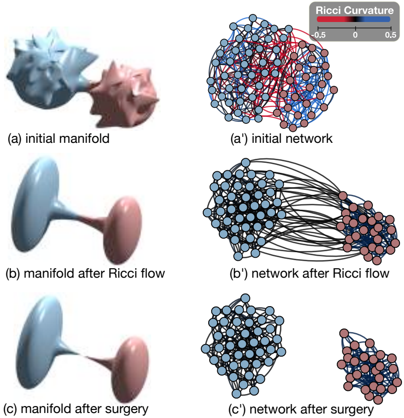
# Curvature Distribution

Negatively curved edges are like “backbones”, maintaining the connectivity of clusters, in which edges are mostly positively curved.



# Ricci Flow on Manifold vs. on Networks

Hamilton introduced Ricci flow, a curvature guided process.



# Ricci Flow Metric

Intuition: flatten the network – shrink an edge if it is within a well connected community; stretch an edge if otherwise, s.t., the network curvature is uniform everywhere.

$$d_{i+1}(x,y) = (d_i(x,y) - \varepsilon \cdot \kappa_i(x,y) \cdot d_i(x,y)) \cdot N$$

# Ricci Flow Metric

Intuition: flatten the network – shrink an edge if it is within a well connected community; stretch an edge if otherwise, s.t., the network curvature is uniform everywhere.

$$d_{i+1}(x,y) = (d_i(x,y) - \varepsilon \cdot \kappa_i(x,y) \cdot d_i(x,y)) \cdot N$$

Distribution on a node $x$:

- Uniform distribution.
- $\exp(-d(x,x_i)^p)$, for a constant $p$.

# Theory on Discrete Ricci Flow

Q: Does Ricci flow converge? Does it generate a unique solution?

- Classical manifold setting: contributes to the proof of the Poincare conjecture.
- Discrete Gaussian curvature on a triangulation: established [Gu, Luo, Wu 2019; Gu, Luo, Sun, Wu 2018 I, II].

# Theory on Discrete Ricci Flow

Q: Does Ricci flow converge? Does it generate a unique solution?

- Classical manifold setting: contributes to the proof of the Poincare conjecture.
- Discrete Gaussian curvature on a triangulation: established [Gu, Luo, Wu 2019; Gu, Luo, Sun, Wu 2018 I, II].
- Discrete curvature on graphs: largely unknown.

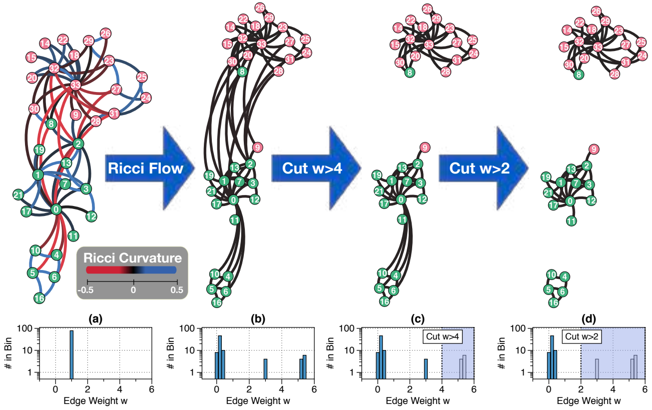
Ollivier Ricci flow:

- Analysis of a very special case. [Ni, Lin, Luo, Gao, 2019]
- Continuous flow, assumption that the edge $uv$ is the shortest path from $u$ to $v$. [Bai, Lin, Lu, Wang, Yau, 2021]

# Applications of Discrete Ricci Flow

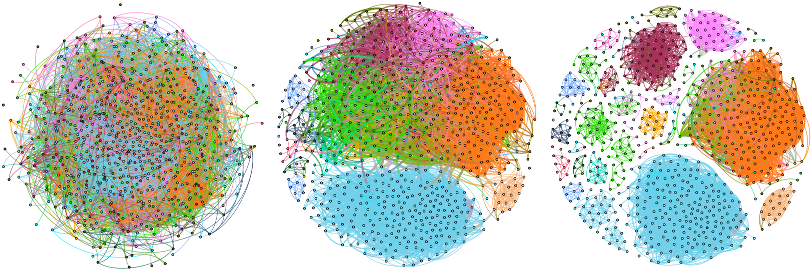
- Community detection
- Network alignment
- Graph neural network

# Community Detection: Karate Club Network



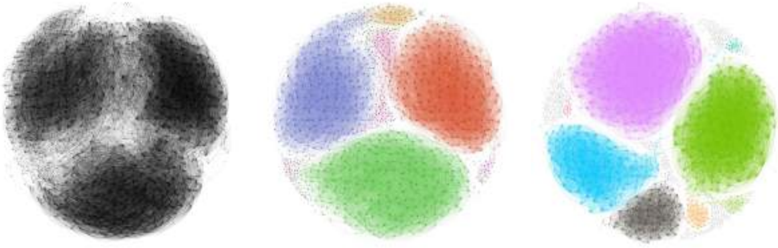
# Community Detection: Facebook Ego Network

792 friends and 14025 edges. The colors represent 24 different friend circles (communities).



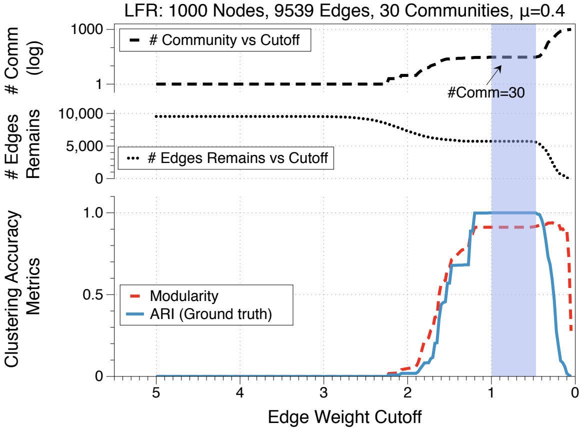
# Community Detection: Brain Connectome Network

Brain network from resting-state (rs-fMRI) data, where edges with cross-correlation less than a threshold are removed.



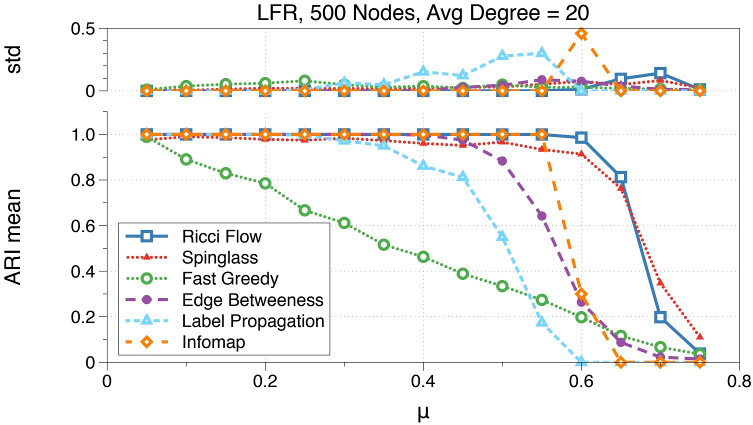
# Cutoff Threshold vs Modularity

Adjusted Rand index (ARI) on Lancichinetti-Fortunato-Radicch (LFR) benchmark network (community size $\sim$ power law).



# Performance Comparison

Adjusted Rand index (ARI) on Lancichinetti-Fortunato-Radicch (LFR) benchmark network (community size $\sim$ power law).



# Ricci Flow Metric: Quantify the Network Distance

In a network, what is a proper distance between two nodes?

# Ricci Flow Metric: Quantify the Network Distance

In a network, what is a proper distance between two nodes?

- On the Internet, measure the delay – Trouble: time-consuming, traffic dependent.

# Ricci Flow Metric: Quantify the Network Distance

In a network, what is a proper distance between two nodes?

- On the Internet, measure the delay – Trouble: time-consuming, traffic dependent.
- On a social network, use tie strength – Trouble: not easy to measure.

# Ricci Flow Metric: Quantify the Network Distance

In a network, what is a proper distance between two nodes?

- On the Internet, measure the delay – Trouble: time-consuming, traffic dependent.
- On a social network, use tie strength – Trouble: not easy to measure.
- Count # hops on the shortest paths– Trouble: small world property;

# Ricci Flow Metric: Quantify the Network Distance

In a network, what is a proper distance between two nodes?

- On the Internet, measure the delay – Trouble: time-consuming, traffic dependent.
- On a social network, use tie strength – Trouble: not easy to measure.
- Count # hops on the shortest paths– Trouble: small world property;
- Distances from some geometric embedding (spectral embedding, Tutte embedding) – Trouble: sensitivity to noises.

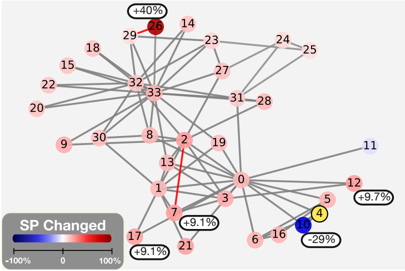
# Ricci Flow Metric: Quantify the Network Distance

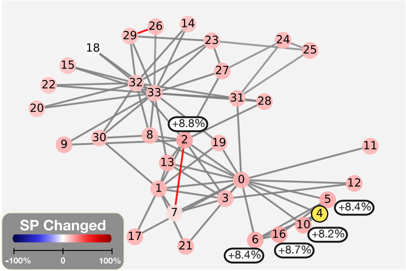
In a network, what is a proper distance between two nodes?

- On the Internet, measure the delay – Trouble: time-consuming, traffic dependent.
- On a social network, use tie strength – Trouble: not easy to measure.
- Count # hops on the shortest paths– Trouble: small world property;
- Distances from some geometric embedding (spectral embedding, Tutte embedding) – Trouble: sensitivity to noises.

# Robustness of Ricci Flow Metric: Remove Two Edges

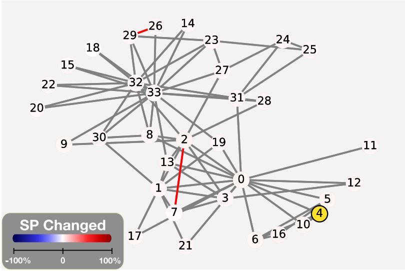
Left: Spectral embedding; Right: Tutte/Spring embedding.

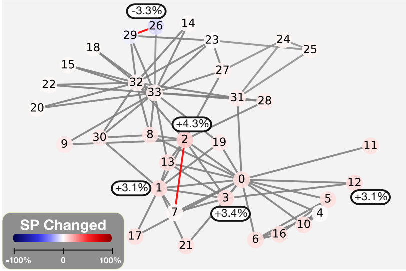


# Robustness: Remove Two Edges

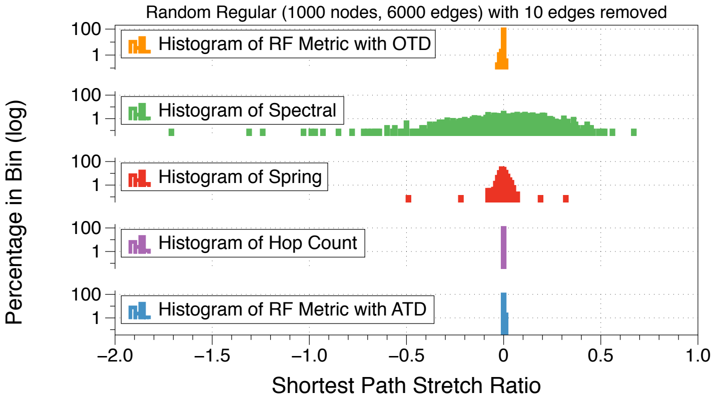
Left: Hop count; Right: our metric.



# Evaluation on Resilience

Randomly remove 10 edges in a random regular graph.

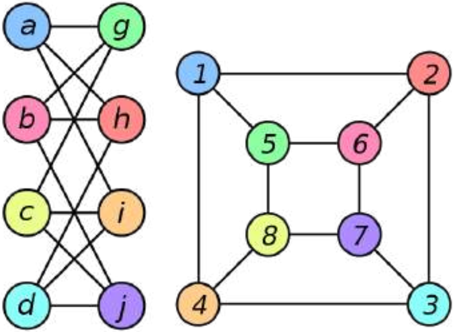


# Graph Isomorphism

Given a pair of graphs $G_1, G_2$, find a one-to-one correspondence of the vertices in $G_1$ to vertices in $G_2$ such that $(u, v)$ is an edge in $G_1$ if and only if their corresponding nodes $f(u), f(v)$ are connected in $G_2$.

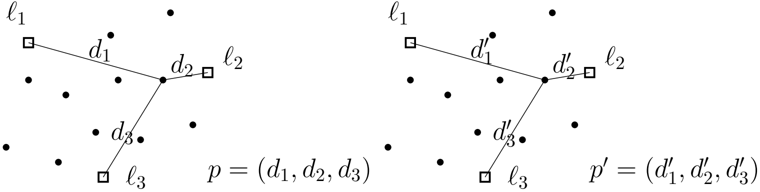
# Graph Isomorphism

Given a pair of graphs $G_1, G_2$, find a one-to-one correspondence of the vertices in $G_1$ to vertices in $G_2$ such that $(u, v)$ is an edge in $G_1$ if and only if their corresponding nodes $f(u), f(v)$ are connected in $G_2$.



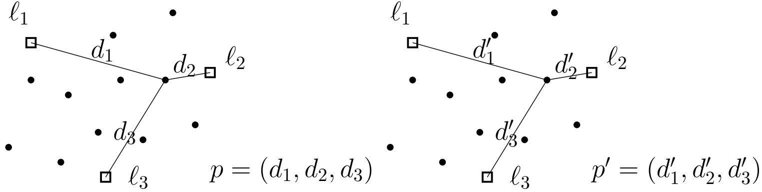
# Our Solution: A Geometric Approach

How to align two sets of points in the plane, assuming that some landmarks $\ell_i$ are already aligned?



# Our Solution: A Geometric Approach

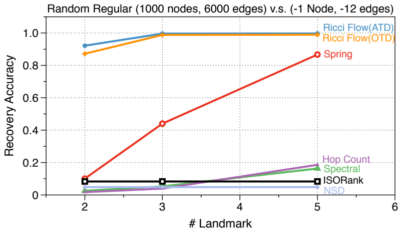
How to align two sets of points in the plane, assuming that some landmarks $\ell_i$ are already aligned?

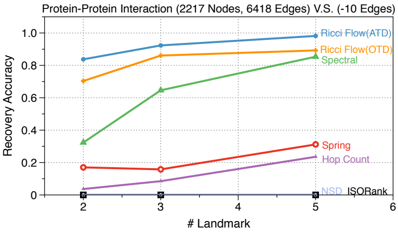



- Any point $p$ can be represented by the barycentric coordinates $(d_1, d_2, d_3)$, $d_i$ is distance to $\ell_i$.
- If the barycentric coordinates of $p$ and $p'$ are similar, we match $p$ and $p'$.

# Evaluation on Matching Performance

- Randomly remove one node in a random regular graph w/ degree 12.
- Right: remove randomly 10 edges in a protein protein network.



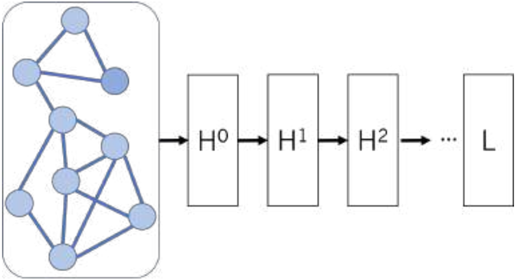



# Graph Neural Network

Graph Neural Network for node classification: given labels of a subset of nodes, predict the labels of the rest.

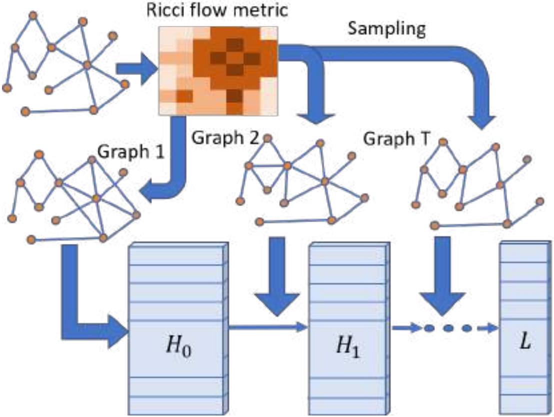
- Graph topology $G = (V, E)$
- Vertex features $H$

Vulnerability: Removal/insertion of fake edges can dramatically hurt model performance.

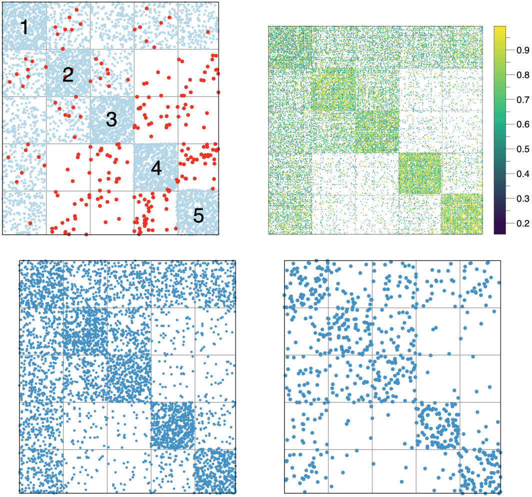


# Robust Graph Neural Network through Resampling

- Recover the underlying metric of $G$ using Ricci flow.
- Re-sample an ensemble of graphs for training.



# Robust Graph Neural Network through Resampling



# Robustness to Graph Topology Attacks

| Dataset | Ptb rate% | GCN | GAT | RGCN | GCN-Jaccard | GCN-SVD | Pro-GNN | Ricci-GNN |
|---|---|---|---|---|---|---|---|---|
| Cora | 0 | 83.50±0.44 | **83.97**±0.65 | 83.09±0.44 | 82.05±0.51 | 80.63±0.45 | 82.98±0.23 | 83.03 ±0.59 |
| | 5 | 76.55±0.79 | 80.44±0.74 | 77.42±0.39 | 79.13±0.59 | 78.39±0.54 | 82.78 ±0.45 | **82.80** ±0.43 |
| | 10 | 70.39±1.28 | 75.61±0.59 | 72.22±0.38 | 76.16±0.76 | 71.47±0.83 | 77.91±0.86 | **79.70** ±0.43 |
| | 15 | 65.10±0.71 | 69.78±1.28 | 66.82±0.39 | 71.03±0.64 | 66.69±1.18 | 76.01 ±1.12 | **77.28**±0.50 |
| | 20 | 59.56±2.72 | 59.94±0.92 | 59.27±0.37 | 65.71±0.89 | 58.94±1.13 | 68.78 ±5.84 | **74.10** ±0.63 |
| | 25 | 47.53±1.96 | 54.78±0.74 | 50.51±0.78 | 60.82±1.08 | 52.06±1.19 | 56.54 ±2.58 | **71.73**±0.60 |
| Citeseer | 0 | 71.96±0.55 | 73.26±0.83 | 71.20±0.83 | 72.10±0.63 | 70.65±0.32 | 73.26 ±0.69 | **73.95**±0.53 |
| | 5 | 70.88±0.62 | 72.89±0.83 | 70.50±0.43 | 70.51±0.97 | 68.84±0.72 | 73.09 ±0.34 | **73.39**±0.52 |
| | 10 | 67.55±0.89 | 70.63±0.48 | 67.71±0.30 | 69.64±0.56 | 68.87±0.62 | 72.43 ±0.75 | **72.51**±0.62 |
| | 15 | 64.52±0.62 | 69.02±0.62 | 65.69±0.62 | 65.95±0.62 | 63.26±0.62 | 70.82 ±2.38 | **71.99**±0.71 |
| | 20 | 62.03±3.49 | 61.04±1.52 | 62.49±1.22 | 59.30±1.40 | 58.55±1.09 | 66.19 ±2.57 | **68.40**±0.51 |
| | 25 | 56.94±2.09 | 61.85±1.12 | 55.35±0.66 | 59.89±1.47 | 57.18±1.87 | 66.40 ±2.57 | **68.84**±0.51 |
| Polblogs | 0 | 95.69±0.38 | 95.35±0.20 | 95.22±0.14 | - | 95.31±0.18 | 93.20 ±0.64 | **95.72**±0.24 |
| | 5 | 73.07±0.80 | 83.69±1.45 | 74.34±0.19 | - | 89.09±0.22 | **93.29** ±0.18 | 90.54 ±0.47 |
| | 10 | 70.72±0.62 | 76.32±0.62 | 71.04±0.62 | - | 81.24±0.62 | **89.42** ±1.09 | 86.88±0.85 |
| | 15 | 64.96±1.91 | 68.80±1.14 | 67.28±0.38 | - | 68.10±3.73 | 86.04 ±2.21 | **86.10**±0.96 |
| | 20 | 51.27±1.23 | 51.50±1.63 | 59.89±0.34 | - | 57.33±3.15 | 79.56 ±5.68 | **81.37**±1.24 |
| | 25 | 49.23±1.36 | 51.19±1.49 | 56.02±0.56 | - | 48.66±9.93 | 63.18 ±4.40 | **79.95** ±1.89 |

# Robustness to Graph Topology Attacks

Is the gain coming from graph ensembling or from the choice of metrics?

| Ptb rate % | Cora | | | Citeseer | | | Polblogs | | |
|---|---|---|---|---|---|---|---|---|---|
| | HC | Spectral | Ricci | HC | Spectral | Ricci | HC | Spectral | Ricci |
| **5** | 81.4 | 76.1 | **82.8** | 73.0 | 71.3 | **73.4** | 85.6 | 88.2 | **90.5** |
| **10** | 77.5 | 73.1 | **79.7** | 71.0 | 69.3 | **72.5** | 85.6 | **88.0** | 86.9 |
| **15** | 75.5 | 65.1 | **77.3** | 70.7 | 67.6 | **72.0** | 70.6 | 80.3 | **86.1** |
| **20** | 72.8 | 60.2 | **74.1** | 67.9 | 67.8 | **68.4** | 65.8 | 79.4 | **81.4** |
| **25** | 65.4 | 54.7 | **71.7** | 65.4 | 65.6 | **68.8** | 62.0 | 74.6 | **80.0** |

Cora: sparse

Polblogs: dense, diameter $=4$.

# Conclusion and Discussion

- Classical geometric notions for discrete graph analysis.
- Network embedding: Euclidean, hyperbolic, hybrid?
- Network evolution: why?
- More applications due to robustness of the metric?

# Acknowledgement


- Chien-Chun Ni, Yu-Yao Lin, Jie Gao, Xianfeng Gu, Emil Saucan, INFOCOM'15.
- Chien-Chun Ni, Yu-Yao Lin, Jie Gao, Xianfeng Gu, Network Alignment by Discrete Ollivier-Ricci Flow, Symposium on Graph Drawing and Network Visualization (GD'18).
- Chien-Chun Ni, Yu-Yao Lin, Jie Gao, Feng Luo, Community Detection on Networks with Ricci Flow, Scientific Reports 9, 9984, July 2019.
- Ye Ze, Kin Sum Liu, Tengfei Ma, Jie Gao and Chao Chen, Curvature Graph Network, ICLR, 2020.
- Ze Ye, Chien-Chun Ni, Tengfei Ma, Chao Chen, Jie Gao, Ricci-GNN: Defending Against Structural Attacks Through a Geometric Approach, under submission.

# Github Code

- `https://github.com/saibalmars/GraphRicciCurvature`

Questions and comments?