

Community Detection on Networks with Ricci Flow

Jie Gao

Stony Brook University
<http://www.cs.stonybrook.edu/~jgao>

October 22nd 2019

Community Detection in Complex Networks

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

- Community structures (clustered, closely knit groups).

Community Detection in Complex Networks

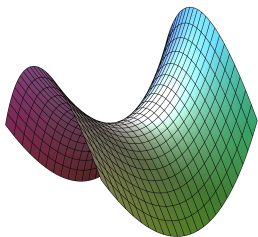
Complex networks in nature: social networks, biological networks, the Internet, WWW.

- Community structures (clustered, closely knit groups).

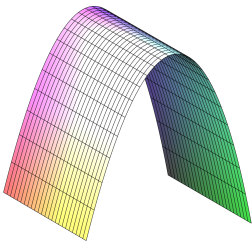
Our project: use geometric tools, Ollivier Ricci curvature flow, to analyze complex networks.

- Community detection.

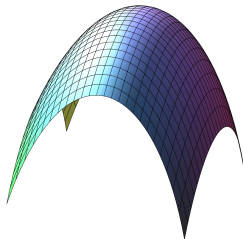
Curvatures on Surfaces vs. on Networks



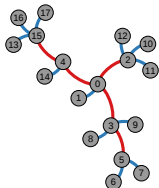
(a) Surface of Negative Curvature



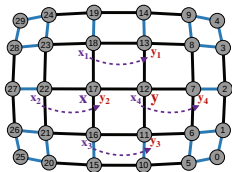
(b) Surface of Zero Curvature



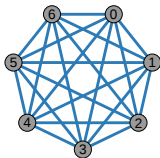
(c) Surface of Positive Curvature



(d) Negative Curvature



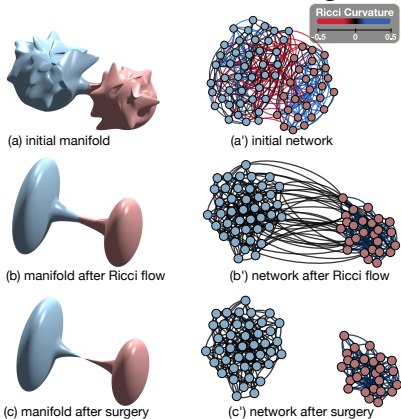
(e) Zero Curvature



(f) Positive Curvature

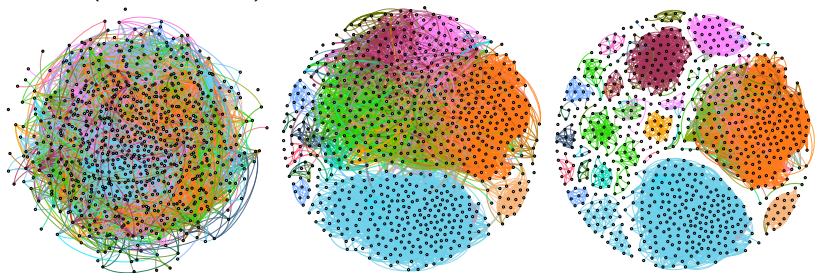
Ricci Flow on Manifold vs. on Networks

Hamilton introduced Ricci flow, a curvature guided process.



Facebook Ego Network

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



Discrete 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 μ_1 and μ_2 on X :

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

Ricci flow for Community Detection

Given a graph G in which $d(x, y)$ is the weight of the edge xy and $\kappa(x, y)$ is the discrete Ricci curvature, we run

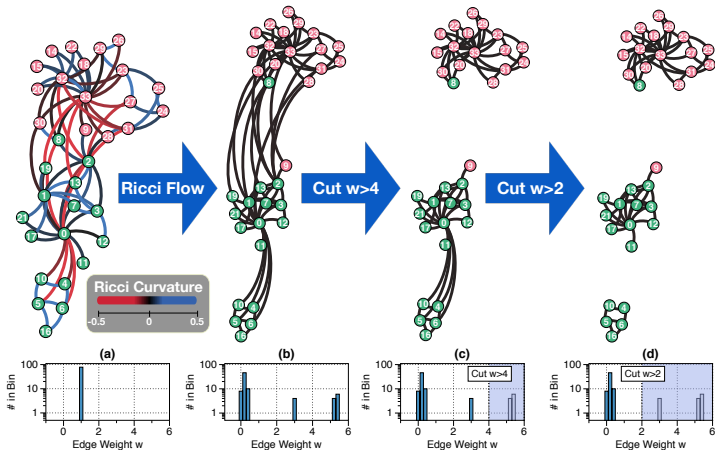
$$d_{i+1}(x, y) = (d_i(x, y) - \varepsilon \cdot \kappa_i(x, y) \cdot d_i(x, y)) \cdot N$$

Distribution on the neighbors of a node x :

- All prior work: uniform distribution.
- This work: $\sim \exp(-d(x, x_i)^p)$, for a constant p .

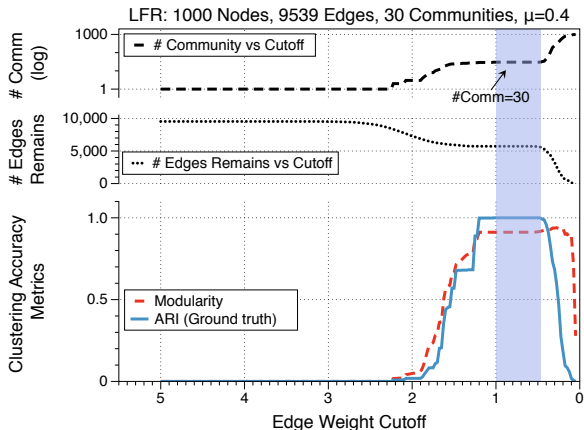
Perform surgery (remove long edges) to reveal hierarchical structure.

Karate Club Network



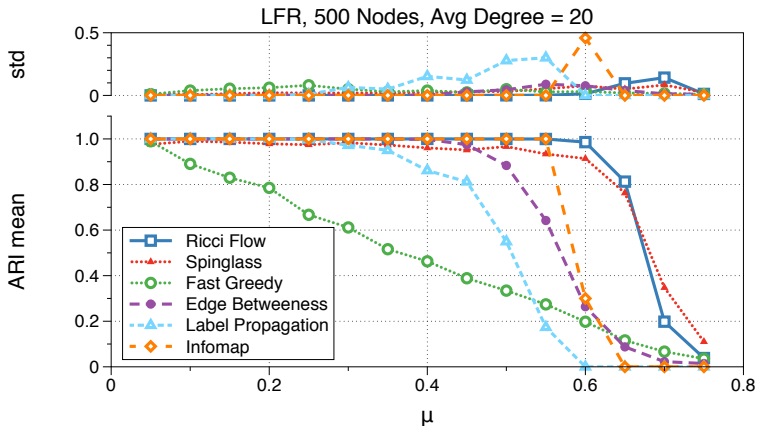
Cutoff Threshold vs Modularity

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



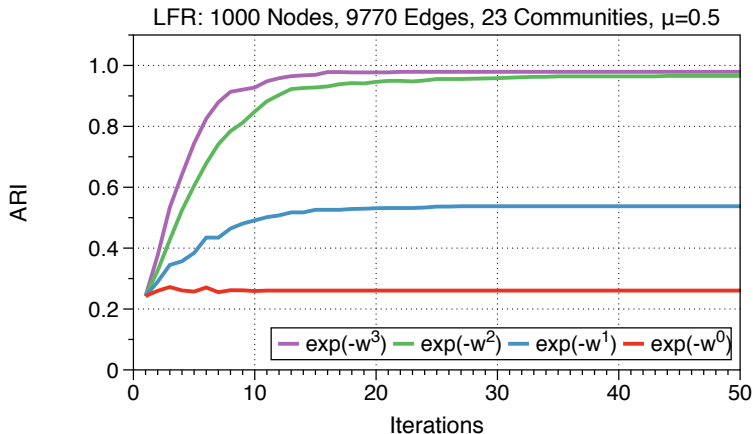
Performance Comparison

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



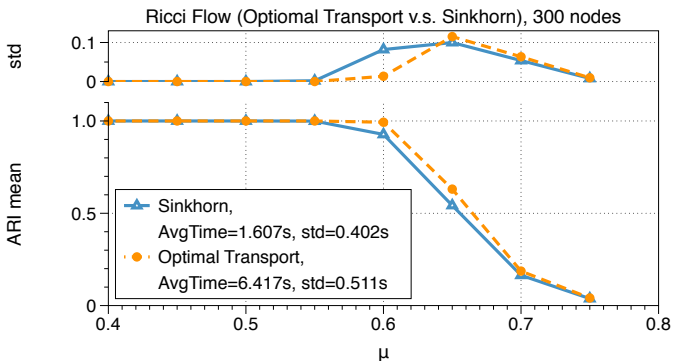
Parameter on the Exponent

Distribution on the neighbors of a node x : $\sim \exp(-d(x, x_i)^p)$, for a constant p .



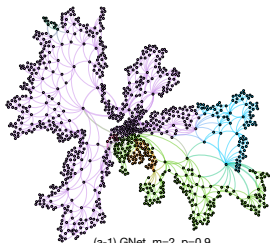
Computation

Use Sinkhorn Algorithm (approximate optimal transport distance) to speed up computation by 4X.

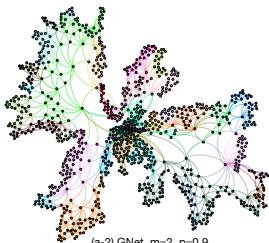


Hierarchical Community Structure

Emergent geometrical network model (GNet): a growing network with high clustering coefficient using the triadic closure property.



(a-1) GNet, $m=2$, $p=0.9$
Cutoff: 3.9, #Comm.: 6



(a-2) GNet, $m=2$, $p=0.9$
Cutoff: 3.2, #Comm.: 38



(a-3) GNet, $m=2$, $p=0.9$
Cutoff: 1, #Comm.: 108

Acknowledgement

- Chien-Chun Ni, Yu-Yao Lin, Jie Gao, Feng Luo, Community Detection on Networks with Ricci Flow, Nature Scientific Reports 9, Article number 9984, published 10 July 2019.
<https://www.nature.com/articles/s41598-019-46380-9>
- Github code:
<https://github.com/saibalmars/GraphRicciCurvature>