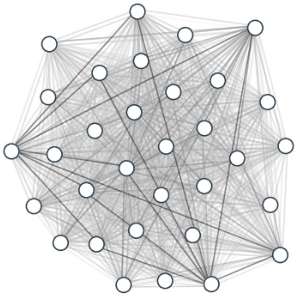
A Topological Centrality Measure for Directed Networks

Linda Fenghuan He
Mentor: Lucy Yang

Commonwealth School

October 16, 2021
MIT PRIMES Conference



Motivation & Background

Motivation & Background

Networks model complex systems as (directed) graphs

Motivation & Background

Networks model complex systems as (directed) graphs

Node Centrality

Motivation & Background

Networks model complex systems as (directed) graphs

Node Centrality

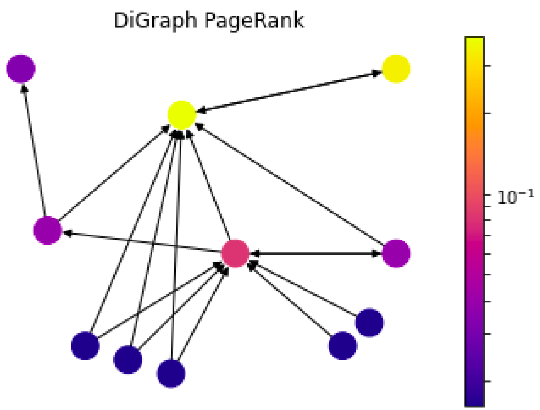
- Betweenness centrality in social networks (J.Lee, 2021)
- Eigenvector centrality in temporal networks (D.Taylor, 2016)

Motivation & Background

Networks model complex systems as (directed) graphs

Node Centrality

- Betweenness centrality in social networks (J.Lee, 2021)
- Eigenvector centrality in temporal networks (D.Taylor, 2016)

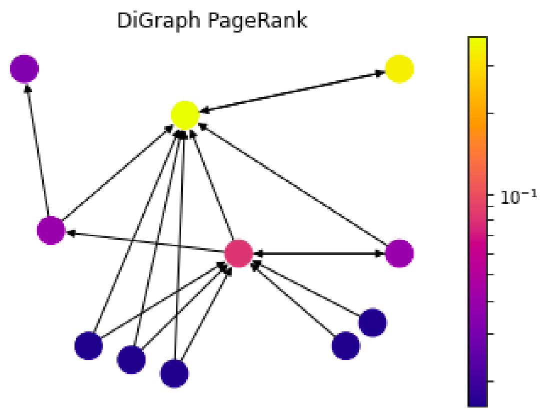


Motivation & Background

Networks model complex systems as (directed) graphs

Node Centrality

- Betweenness centrality in social networks (J.Lee, 2021)
- Eigenvector centrality in temporal networks (D.Taylor, 2016)



Goal

Define a centrality measure that captures non-local propagating effects and directedness.

Networks

Networks

Definition

A **network** G is a pair (X, w_X) where X is a finite set and $w_X : X \times X \to \mathbb{R}$ is called the weight function.

Networks

Definition

A **network** G is a pair (X, w_X) where X is a finite set and $w_X : X \times X \to \mathbb{R}$ is called the weight function.

We will be restricted to networks G where $w_X(x, x) = 0$ for all $x \in X$.

Networks

Definition

A **network** G is a pair (X, w_X) where X is a finite set and $w_X : X \times X \to \mathbb{R}$ is called the weight function.

We will be restricted to networks G where $w_X(x,x) = 0$ for all $x \in X$.

Definition (F.Iannelli, 2017)

Let $G = (X, w_X)$ be a network, define $\gamma(G)$ to be (X, m_X) where $m_X : X \times X \to \mathbb{R}$ is given by:

$$m(x,y) = \begin{cases} 1 - \log \frac{w(x,y)}{\sum_{z \neq y} w(x,z)} \geq 1 & \text{if } y \neq x \\ 0 & \text{if } y = x \end{cases}$$

Two nodes that interact a lot ($w(x,y) \gg 0$) will be closer ($m(x,y) \sim 1$).

Goal

Goal

Definition

Given a network G and x a node in G, define $f(G,x) = \left(X \setminus \{x\}, w_X|_{X\setminus\{x\}}\right)$, i.e. the sub-network induced by deleting x and all edges incident to x in G.

Goal

Definition

Given a network G and x a node in G, define $f(G,x) = \left(X \setminus \{x\}, w_X|_{X\setminus\{x\}}\right)$, i.e. the sub-network induced by deleting x and all edges incident to x in G.

Idea

Given x a node in G, we compare the difference in the "[dis]connectivity" of $\gamma(G)$ and $\gamma(f(G,x))$.

TDA in Networks

Question

How to quantify [dis]connectivity of a graph G?

TDA in Networks

Question

How to quantify [dis]connectivity of a graph G?

Algebraic topology measures the "holes" in a "shape" using "homology".

Idea

We use the "size" of the homology of a "shape" built from G as a proxy for disconnectivity.

TDA in Networks

Question

How to quantify [dis]connectivity of a graph G?

Algebraic topology measures the "holes" in a "shape" using "homology".

Idea

We use the "size" of the homology of a "shape" built from G as a proxy for disconnectivity.

Recall a *simplicial complex* is a set of tetrahedrons of any dimension "glued together in a nice way".

TDA in Networks

Question

How to quantify [dis]connectivity of a graph G?

Algebraic topology measures the "holes" in a "shape" using "homology".

Idea

We use the "size" of the homology of a "shape" built from G as a proxy for disconnectivity.

Recall a *simplicial complex* is a set of tetrahedrons of any dimension "glued together in a nice way".

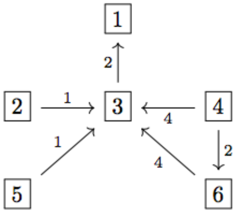
Definition (F.Memoli and S.Chowdhury, 2016)

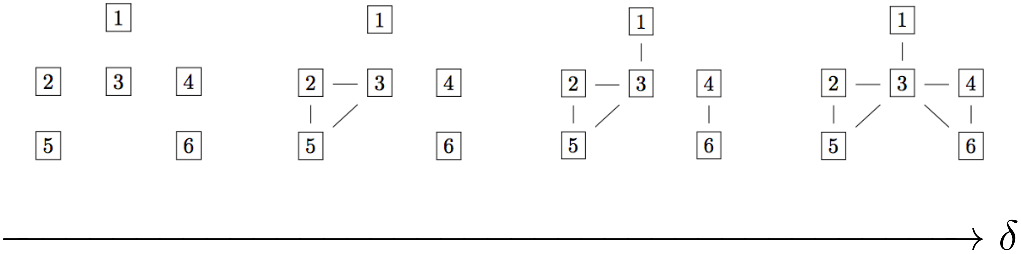
Given a network $G = (X, w_X)$ and $\delta \in \mathbb{R}$, the **Dowker Complex** $\mathcal{D}_{\delta,G}$ is the simplicial complex given by:

$$\mathcal{D}_{\delta,G} := \{\sigma \subseteq X : \exists p \in X \text{ s.t. } w(x,p) \leq \delta \ \forall\, x \in \sigma\}.$$

Example

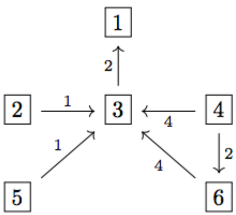
$$\mathcal{D}_{\delta,G} := \{\sigma \subseteq X : \exists p \in X \text{ s.t. } w(x,p) \leq \delta \; \forall \, x \in \sigma\}.$$

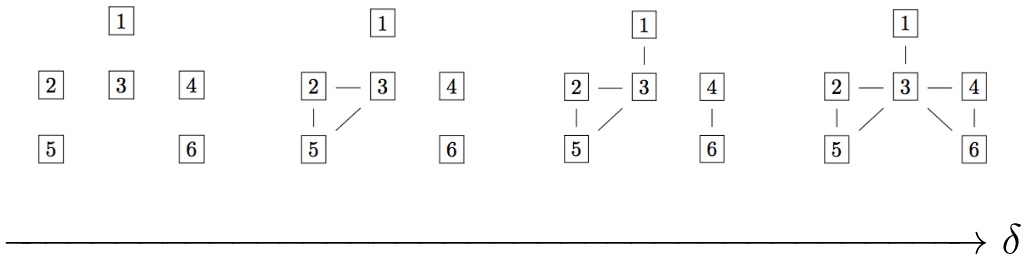




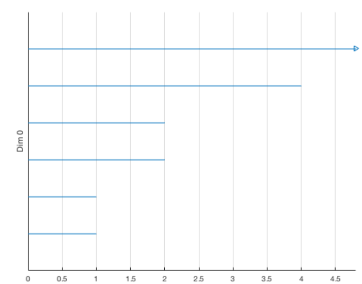
Example

$$\mathcal{D}_{\delta,G} := \{\sigma \subseteq X : \exists p \in X \text{ s.t. } w(x,p) \leq \delta \; \forall \, x \in \sigma\}.$$





1. As $\delta \nearrow$, number of path components $\searrow$.
2. This data is recorded on a **persistence diagram**.
3. We denote $\mathbf{P}_0(G)$ as the set of 0-dimensional barcodes for the Dowker complex $\mathcal{D}_{\cdot,G}$.



Quasi-centrality

For a node $x \in X$, let $\mu(x) := d$ to be the value of δ for which x merges into another connected component.

Quasi-centrality

For a node $x \in X$, let $\mu(x) := d$ to be the value of δ for which x merges into another connected component.

Definition

Let G be a network. The **quasi-centrality** $C(x)$ for node $x \in X$ is:

$$C(x) = \sum_{c \in \mathbf{P}_0(f(\gamma(G),x))} length(c) - \sum_{c \in \mathbf{P}_0(\gamma(G))} length(c) + d$$

Quasi-centrality

For a node $x \in X$, let $\mu(x) := d$ to be the value of δ for which x merges into another connected component.

Definition

Let G be a network. The **quasi-centrality** $C(x)$ for node $x \in X$ is:

$$C(x) = \sum_{c \in \mathbf{P}_0(f(\gamma(G), x))} length(c) - \sum_{c \in \mathbf{P}_0(\gamma(G))} length(c) + d$$

Theorem

For a network $G = (X, w_X)$, $C(x)$ is nonnegative for all $x \in X$.

Applications

Goals

- Demonstrate that C is a valid measure of centrality
- Use quasi-centrality to assess the influence of a node in a real-world network.

Applications

Goals

- Demonstrate that C is a valid measure of centrality
- Use quasi-centrality to assess the influence of a node in a real-world network.

Trade networks

- Interdependency between far-flung communities
- Trade networks are fragile (Y.Korniyenko, 2017)
- Economic perturbations originated in a single country can propagate elsewhere

Applications

Goals

- Demonstrate that C is a valid measure of centrality
- Use quasi-centrality to assess the influence of a node in a real-world network.

Trade networks

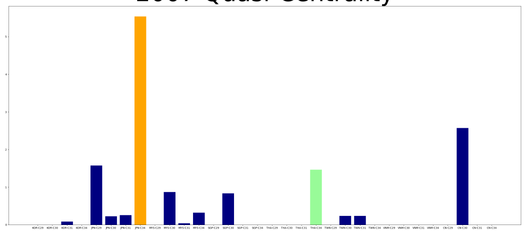
- Interdependency between far-flung communities
- Trade networks are fragile (Y.Korniyenko, 2017)
- Economic perturbations originated in a single country can propagate elsewhere

Data

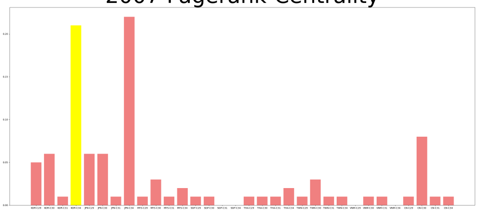
- OECD Inter-Country Input-Output (ICIO) Tables
- Machinery production network in Asia
- Industries: machinery equipment, computer and electronics, electrical machinery, auto machinery

Results

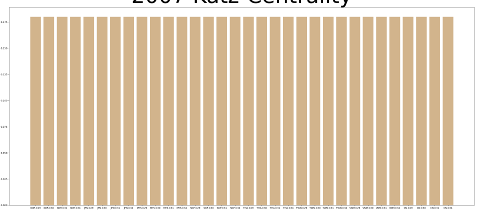
2007 Quasi-Centrality



2007 Pagerank-Centrality

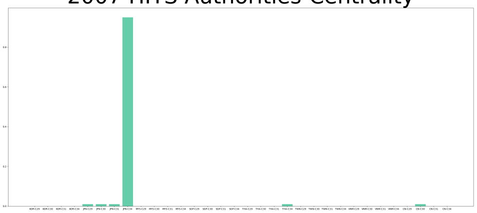


2007 Katz-Centrality



2007 HITS Hubs-Centrality

2007 HITS Authorities-Centrality



Future directions

- Compute the quasi-centrality measure for other asymmetric networks
  - biological networks
  - airflight networks
- Relate higher dimensional homological features in directed networks to real-world phenomena
  - trade flows
  - embargo
- Define other measures in network analysis using TDA
  - connectivity
  - robustness
  - efficiency

Acknowledgements

- My mentor, Lucy Yang
- Prof. Memoli of the Ohio State University
- Dr. Slava Gerovitch
- Prof. Pavel Etingof
- Dr. Tanya Khovanova
- MIT PRIMES
- My family