University of Strathclyde Glasgow

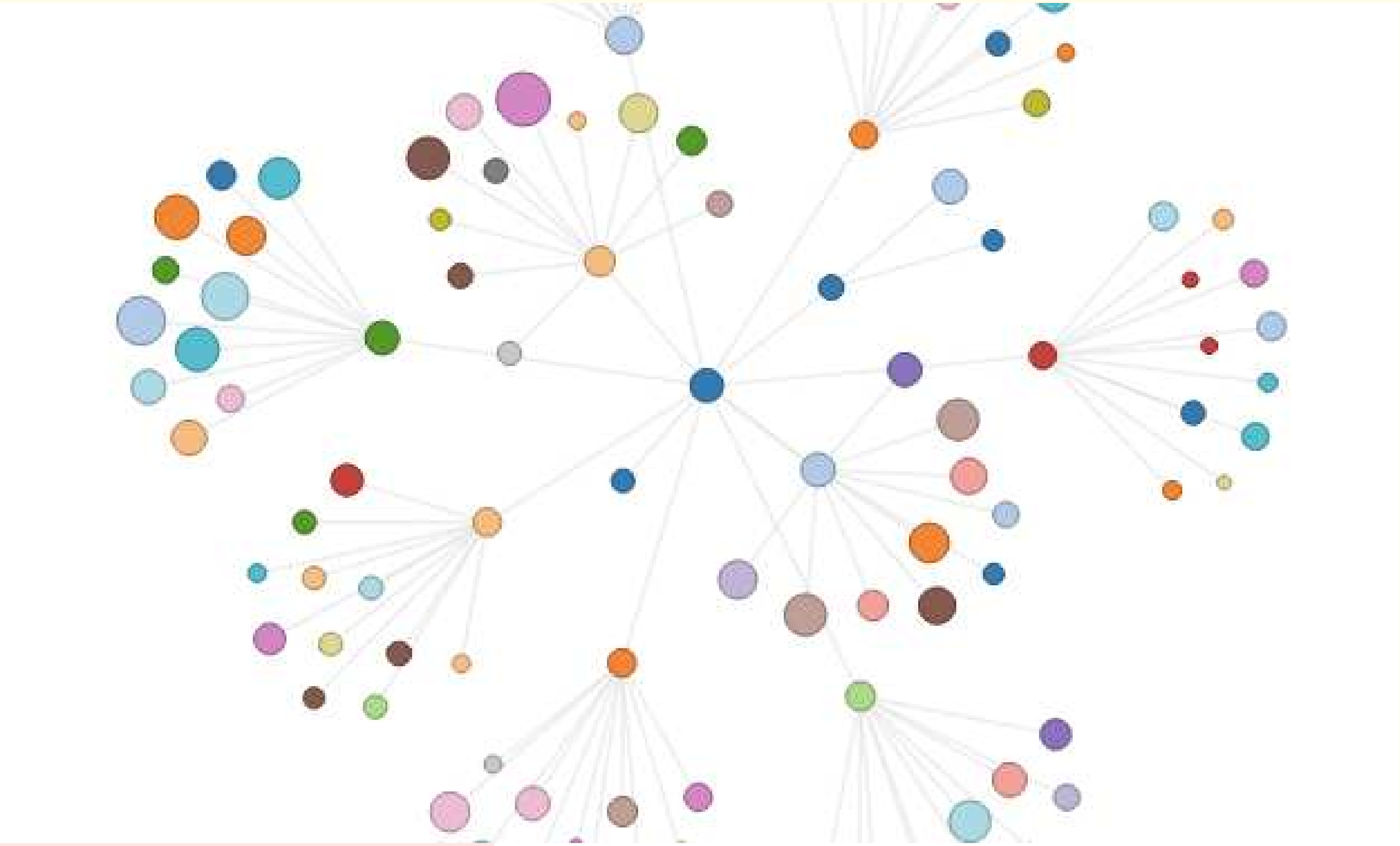
Combinatorics of nonbacktracking and non-cycling walks and their applications to network science

F. Arrigo, D. J. Higham and V. Noferini

WACA 2021

May 25 – 28, 2021

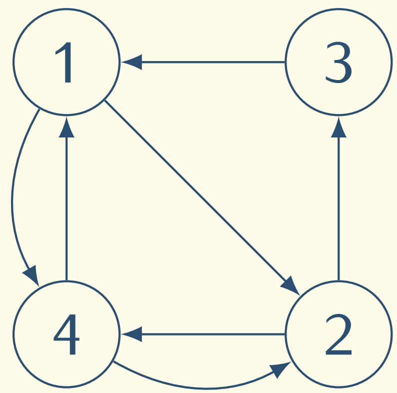
- Graph $G = (V, E)$ with n nodes and m edges, unweighted and w/o loops.
- Every undirected edge $i - j$ is regarded as: $i \to j$ and $j \to i$.



The importance of counting walks

A **walk** is any traversal of the network.

Adjacency matrix: $A = (a_{ij}) \in \mathbb{R}^{n\times n}$, $a_{ij} = 1$ iff $i \to j$.



$$A = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix}$$

Easy: $(A^2)_{ij} = \sum_{k=1}^{n} a_{ik}a_{kj}$ counts walks of length two.
Generally: entries of A^r count walks of length r.

Counting walks is at the heart of several clustering and centrality algorithms!

Walk-based centrality

Let $f(x) = \sum_{r=0}^{\infty} c_r x^r$, then, within the radius of convergence:

$$f(A) = \sum_{r=0}^{\infty} c_r A^r$$

Looking closely at $(f(A))_{ij}$:

$$(f(A))_{ij} = \sum_{r=0}^{\infty} c_r (A^r)_{ij}$$

- it tells us how many walks (up to infinite length) originate at node i and end at node j
- if $c_r \geq 0$ and $c_r \to 0$ as $r \to \infty$, longer walks are given less weight.

Walk-based centrality

Centrality measures assign scores to nodes in graphs to quantify their importance.

Among the most popular centrality measures:

- *f-subgraph centrality (SC)*

$$\mathbf{x}(t)_i = \mathbf{e}_i^T \left(\sum_{r=0}^{\infty} c_r t^r A^r \right) \mathbf{e}_i = f(A)_{ii}$$

- *f-total (node) communicability (TC)*:

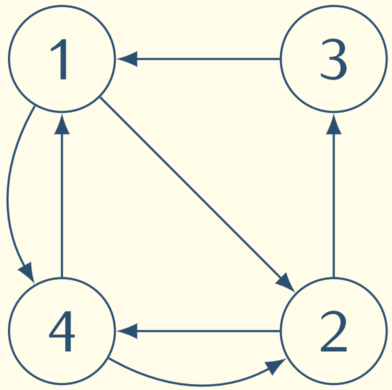
$$\mathbf{y}(t)_i = \sum_{j=1}^{n} \sum_{r=0}^{\infty} c_r t^r (A^r)_{ij} = \sum_{j=1}^{n} (f(A))_{ij} = (f(A)\mathbf{1})_i$$

The importance of counting walks

Katz centrality: row sums of $I + tA + t^2A^2 + t^3A^3 + \cdots$, i.e.:

$$\mathbf{k} = (I - tA)^{-1}\mathbf{1}, \qquad \text{for } 0 < t < \rho(A)^{-1}.$$

$\Longrightarrow$ Solve a sparse linear system!



Same importance to these walks of length four:

$1 \to 4 \to 2 \to 3 \to 1$

$1 \to 4 \to 1 \to 4 \to 1.$

... but the first visits the network more thoroughly.

Katz, *"A new status index derived from sociometric analysis"*, Psychometrika (1953).

A **walk (of length $r-1$)** is any traversal of the network:

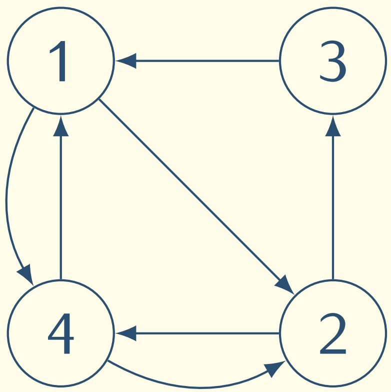
$$\underline{i} = (i_1, i_2, \ldots, i_r) \quad \text{or} \quad i_1 \to i_2 \to \cdots \to i_r.$$

A walk is said to be **non-backtracking (NBT)** if it <ins>does not</ins> contain any instance of the form $i \to j \to i$.

A **walk (of length $r-1$)** is any traversal of the network:

$$\underline{i} = (i_1, i_2, \ldots, i_r) \quad \text{or} \quad i_1 \to i_2 \to \cdots \to i_r.$$

A walk is said to be **non-backtracking (NBT)** if it does not contain any instance of the form $i \to j \to i$.



$1 \to 4 \to 2 \to 3 \to 1$ is NBT
$1 \to 4 \to 1 \to 4 \to 1$ is backtracking.

Question:

- Can we define NBTW-based subgraph centrality and total communicability indices, using entries or sum of entries of

$$c_0p_0(A) + c_1tp_1(A) + c_2t^2p_2(A) + c_3t^3p_3(A) + \cdots$$

where $p_r(A)$ generalizes A^r to the NBT setting, i.e., it counts NBTWs of length r, and $t > 0$ is selected so that the series converges.

Non-backtracking walks have been suggested and analyzed in a wide range of fields including (but not limited to!)

- spectral graph theory [Angel, Friedman, Hoory, Horton, Stark, Terras, ...],
- number theory [Terras, ...],
- discrete mathematics [Bowen, Lanford, Stark, Terras, ...],
- stochastic analysis [Alon, Benjamini,Lubetzky, Sodin, ...],
- applied linear algebra [Tarfulea, Perlis, ...],
- computer science [Zaade, Krzakala, Zdeborová],
- ...

Motivation

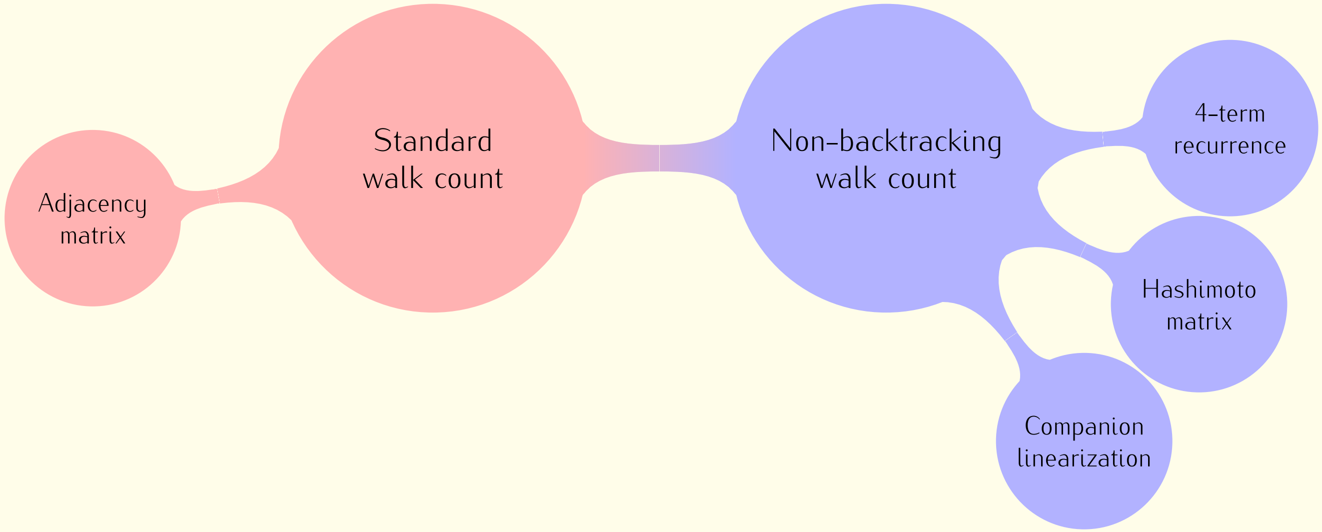
Non-backtracking walks have been suggested and analyzed in a wide range of fields including (but not limited to!)

- spectral graph theory [Angel, Friedman, Hoory, Horton, Stark, Terras, ...],
- number theory [Terras, ...],
- discrete mathematics [Bowen, Lanford, Stark, Terras, ...],
- stochastic analysis [Alon, Benjamini,Lubetzky, Sodin, ...],
- applied linear algebra [Tarfulea, Perlis, ...],
- computer science [Zaade, Krzakala, Zdeborová],
- ...

In Network Science:

- finding communities [Kawamoto, Krzakala, Moore, Mossel, Zhang,...];
- assigning centrality values to nodes [Martin, Zhang, Newman, Pastor-Satorras, Castellano, Flores, Criado, and many others].

- NBTWs visit the network more thoroughly;
- NBTW-based measures avoid localization;
- They can be computed at the same (or lower) computational costs;
- Larger radius of convergence;
- High versatility from the viewpoint of Linear Algebra



Problem setting

Let $p_r(A) \in \mathbb{R}^{n\times n}$ be such that

$$(p_r(A))_{ij} = |\{\text{NBTW s of length } r \text{ from node } i \text{ to node } j\}|.$$

Problem setting

Let $p_r(A) \in \mathbb{R}^{n\times n}$ be such that

$$(p_r(A))_{ij} = |\{\text{NBTW s of length } r \text{ from node } i \text{ to node } j\}|.$$

Given $f(A) = \sum_{r=0}^{\infty} c_r t^r A^r$, let the non-backtracking counterpart of the matrix function $f(A)$ be:

$$F(A) := \sum_{r=0}^{\infty} c_r t^r p_r(A).$$

Let $p_r(A) \in \mathbb{R}^{n\times n}$ be such that

$$(p_r(A))_{ij} = |\{\text{NBTW s of length } r \text{ from node } i \text{ to node } j\}|.$$

Given $f(A) = \sum_{r=0}^{\infty} c_r t^r A^r$, let the non-backtracking counterpart of the matrix function $f(A)$ be:

$$F(A) := \sum_{r=0}^{\infty} c_r t^r p_r(A).$$

Define the *NBTW-based SC* of node i as:

$$\mathbf{x}_2(i) = (F(A))_{ii}$$

and the *NBTW-based TC* of node i as:

$$\mathbf{y}_2(i) = (F(A)\mathbf{1})_i.$$

Theorem: Let A be the adj. matrix of an unweighted digraph, $D = \mathrm{diag}(\mathrm{diag}(A^2))$, and $N = A - A \circ A^T$. Then,

$$\begin{aligned} p_0(A) &= I, \\ p_1(A) &= A, \\ p_2(A) &= A^2 - D \end{aligned}$$

and

$$p_r(A) = p_{r-1}(A)A + p_{r-2}(A)\,(I - D) - p_{r-3}(A)N, \quad \forall r \geq 3.$$

Tarfulea & Perris, *"An Ihara formula for partially directed graphs"*, LAA (2009).

Four term recurrence

$$p_{r+3}(A) = p_{r+2}(A)A + p_{r+1}(A)(I - D) - p_r(A)N.$$

Entrywise:

$$\begin{array}{r|c}
(p_{r+3}(A))_{ij} = & \overbrace{i \to \cdots \to \star \to \star \to \star \to j}^{NBTW \text{ of length } r+3} \\
\hline
+(p_{r+2}(A)A)_{ij} & i \to \cdots \to \star \to \star \to \ell \to j
\end{array}$$

$$p_{r+3}(A) = p_{r+2}(A)A + p_{r+1}(A)(I - D) - p_r(A)N.$$

Entrywise:

$(p_{r+3}(A))_{ij} =$	$\overbrace{i \to \cdots \to \star \to \star \to \star \to j}^{NBTW \text{ of length } r+3}$
$+(p_{r+2}(A)A)_{ij}$	$i \to \cdots \to \star \to \star \to \ell \to j$
$-(p_{r+1}(A)D)_{ij}$	$i \to \cdots \to \star \to j \to \ell \to j$

Four term recurrence

$$p_{r+3}(A) = p_{r+2}(A)A + p_{r+1}(A)(I - D) - p_r(A)N.$$

Entrywise:

$(p_{r+3}(A))_{ij} =$	$\overbrace{i \to \cdots \to \star \to \star \to \star \to j}^{NBTW \text{ of length } r+3}$
$+(p_{r+2}(A)A)_{ij}$	$i \to \cdots \to \star \to \star \to \ell \to j$
$-(p_{r+1}(A)D)_{ij}$	$i \to \cdots \to \star \to j \to \ell \to j$
$+(p_{r+1}(A))_{ij}$	$i \to \cdots \to \ell \to j \to \ell \to j$

$$p_{r+3}(A) = p_{r+2}(A)A + p_{r+1}(A)(I-D) - p_r(A)N.$$

Entrywise:

$(p_{r+3}(A))_{ij} =$	$\overbrace{i \to \cdots \to \star \to \star \to \star \to j}^{NBTW \text{ of length } r+3}$
$+(p_{r+2}(A)A)_{ij}$	$i \to \cdots \to \star \to \star \to \ell \to j$
$-(p_{r+1}(A)D)_{ij}$	$i \to \cdots \to \star \to j \to \ell \to j$
$+(p_{r+1}(A))_{ij}$	$i \to \cdots \to \ell \to j \to \ell \to j$
$-(p_r(A)N)_{ij}$	it must hold $j \leftrightarrow \ell$

Example: *Katz centrality*

The non-backtracking analogue of Katz centrality is:

$$\mathbf{k}_{NBT} := F(A)\mathbf{1} = \sum_{r=0}^{\infty} t^r p_r(A)\mathbf{1}.$$

Theorem: Let A, N, and D be defined as before. Moreover, if we use $c_r = 1$ and $t > 0$ is such that $F(A) = \sum_{r=0}^{\infty} t^r p_r(A)$ converges, then

$$[I - At + (D - I)t^2 + Nt^3]F(A) = (1 - t^2)I$$

Example: *Katz centrality*

The non-backtracking analogue of Katz centrality is:

$$\mathbf{k}_{NBT} := F(A)\mathbf{1} = \sum_{r=0}^{\infty} t^r p_r(A)\mathbf{1}.$$

Theorem: Let A, N, and D be defined as before. Moreover, if we use $c_r = 1$ and $t > 0$ is such that $F(A) = \sum_{r=0}^{\infty} t^r p_r(A)$ converges, then

$$[I - At + (D - I)t^2 + Nt^3]F(A) = (1 - t^2)I$$

$$\boxed{\mathbf{k}_{NBT} = (1 - t^2)[I - At + (D - I)t^2 + Nt^3]^{-1}\mathbf{1}}$$

$\Longrightarrow$ same cost as Katz.

Setup
The importance of counting walks
Walk-based centrality
Question 1
Motivation
Problem setting
Four term recurrence
From A to B
Small example
Counting NBT walks
Problem setting
question 2
$ABC\dots$ easy as $123\dots$
Counting non–k–cycling walks
What else can we do?

- $i - j \quad \Leftrightarrow \quad i \to j$ and $j \to i$
- $G = (V, E)$ with $n = m_1$ nodes and m_2 edges;
- **Adjacency matrix:** $A = W_1 = P_1$;
- **Source and target matrices:** $L_1, R_1 \in \mathbb{R}^{m_2 \times m_1}$ of P_1:

$$(L_1)_{\underline{i}\ell} = \begin{cases} 1 & \text{if } i_1 = \ell \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad (R_1)_{\underline{i}\ell} = \begin{cases} 1 & \text{if } i_2 = \ell \\ 0 & \text{otherwise} \end{cases}$$

- $W_2 = R_1 L_1^T$;
- **Hashimoto matrix:**

$$B = P_2 = W_2 - \Delta_2,$$

where $\Delta_2 = W_2 \circ W_2^T$ and $\circ$ entry-wise product.

- $i - j \quad \Leftrightarrow \quad i \to j$ and $j \to i$
- $G = (V, E)$ with $n = m_1$ nodes and m_2 edges;
- **Adjacency matrix:** $A = W_1 = P_1$;
- **Source and target matrices:** $L_1, R_1 \in \mathbb{R}^{m_2 \times m_1}$ of P_1:

$$(L_1)_{\underline{i}\ell} = \begin{cases} 1 & \text{if } i_1 = \ell \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad (R_1)_{\underline{i}\ell} = \begin{cases} 1 & \text{if } i_2 = \ell \\ 0 & \text{otherwise} \end{cases}$$

- $W_2 = R_1 L_1^T$;
- **Hashimoto matrix:**

$$B = P_2 = W_2 - \Delta_2,$$

where $\Delta_2 = W_2 \circ W_2^T$ and $\circ$ entry-wise product.

- $i - j \quad \Leftrightarrow \quad i \to j$ and $j \to i$
- $G = (V, E)$ with $n = m_1$ nodes and m_2 edges;
- **Adjacency matrix:** $A = W_1 = P_1$;
- **Source and target matrices:** $L_1, R_1 \in \mathbb{R}^{m_2 \times m_1}$ of P_1:

$$(L_1)_{\underline{i}\ell} = \begin{cases} 1 & \text{if } i_1 = \ell \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad (R_1)_{\underline{i}\ell} = \begin{cases} 1 & \text{if } i_2 = \ell \\ 0 & \text{otherwise} \end{cases}$$

- $W_2 = R_1 L_1^T$;
- **Hashimoto matrix:**

$$B = P_2 = W_2 - \Delta_2,$$

where $\Delta_2 = W_2 \circ W_2^T$ and $\circ$ entry-wise product.

From A to B

- $i - j \quad \Leftrightarrow \quad i \to j$ and $j \to i$
- $G = (V, E)$ with $n = m_1$ nodes and m_2 edges;
- **Adjacency matrix:** $A = W_1 = P_1$;
- **Source and target matrices:** $L_1, R_1 \in \mathbb{R}^{m_2 \times m_1}$ of P_1:

$$(L_1)_{\underline{i}\ell} = \begin{cases} 1 & \text{if } i_1 = \ell \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad (R_1)_{\underline{i}\ell} = \begin{cases} 1 & \text{if } i_2 = \ell \\ 0 & \text{otherwise} \end{cases}$$

- $W_2 = R_1 L_1^T$;
- **Hashimoto matrix:**

$$B = P_2 = W_2 - \Delta_2,$$

where $\Delta_2 = W_2 \circ W_2^T$ and $\circ$ entry-wise product.

From A to B (cont.)

We want to concatenate $\underline{i} = (i_1, i_2)$ and $\underline{j} = (j_1, j_2)$ without them backtracking:

$$i_1 \to i_2 = j_1 \to j_2 \quad \text{and} \quad i_1 \neq j_2$$

From A to B (cont.)

We want to concatenate $\underline{i} = (i_1, i_2)$ and $\underline{j} = (j_1, j_2)$ without them backtracking:

$$i_1 \to i_2 = j_1 \to j_2 \quad \text{and} \quad i_1 \neq j_2$$

Entrywise it holds:

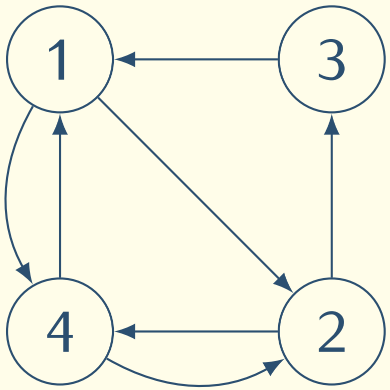
$$(W_2)_{\underline{ij}} = (R_1 L_1^T)_{\underline{ij}} = \delta_{i_2 j_1}$$

$$(P_2)_{\underline{ij}} = (W_2 - W_2 \circ W_2^T)_{\underline{ij}} = \delta_{i_2 j_1}(1 - \delta_{i_1 j_2})$$

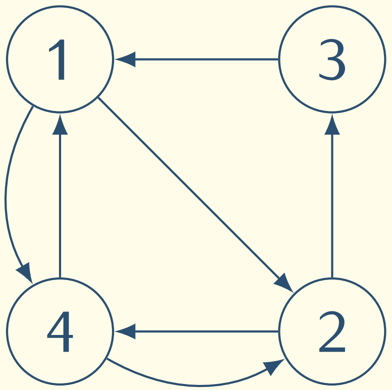
since

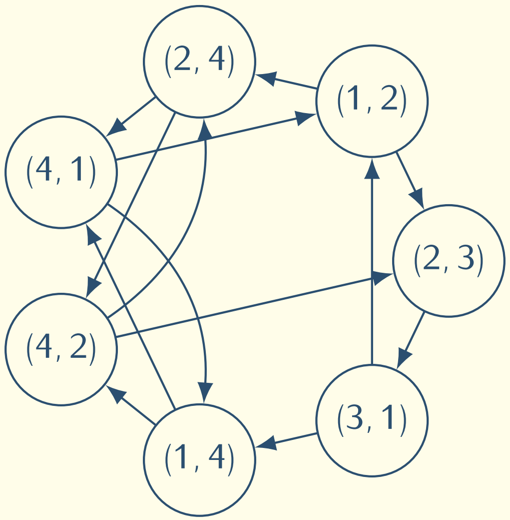
$$(W_2 \circ W_2^T)_{\underline{ij}} = (W_2)_{\underline{ij}}(W_2)_{\underline{ji}} = \delta_{i_2 j_1}\delta_{j_2 i_1}$$

where δ is the Kronecker delta.

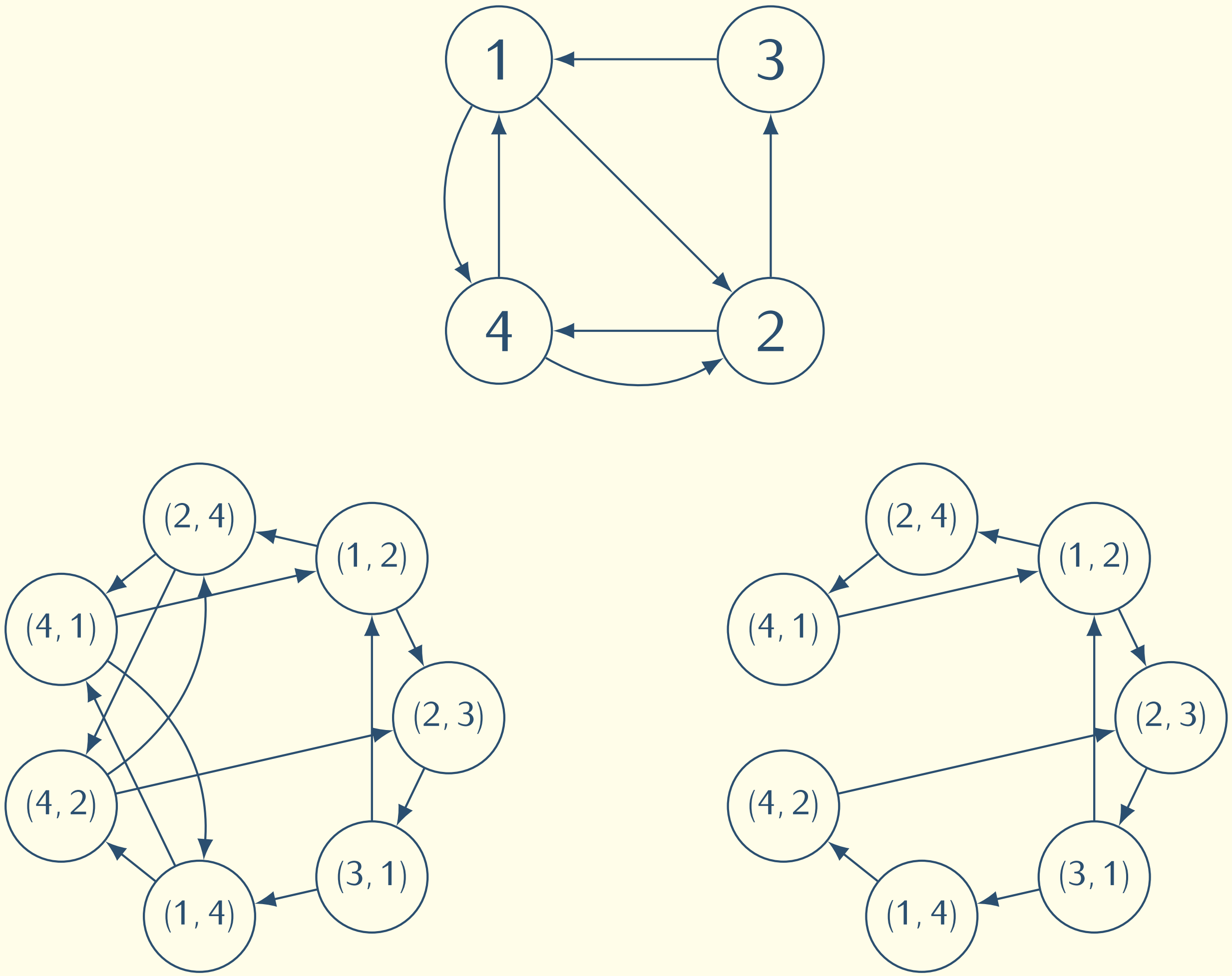

1
3
4
2

Small example

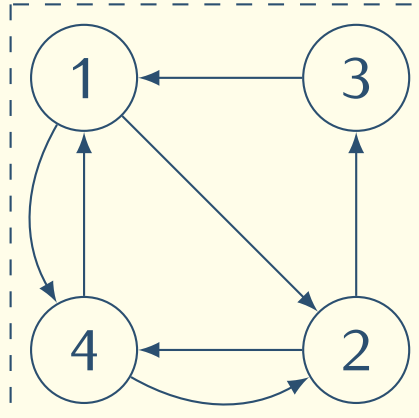


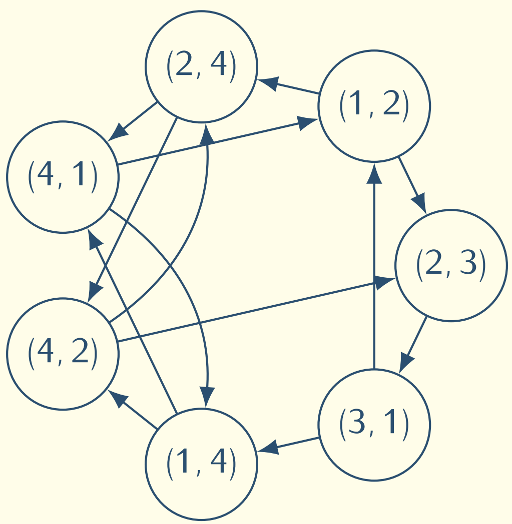


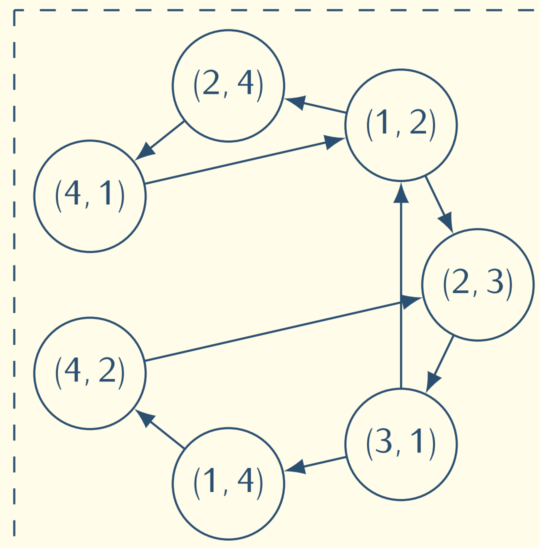
Small example



Small example







It can be shown that

$$L_1^T(B^r)R_1 = L_1^T(P_2^r)R_1 = p_{r+1;2}(A)$$

It can be shown that

$$L_1^T(B^r)R_1 = L_1^T(P_2^r)R_1 = p_{r+1;2}(A)$$

and moreover, within the radius of convergence,

$$\mathbf{x}_2(t)_i = (L_1^T \partial f(tP_2)R_1)_{ii} \quad \text{and} \quad \mathbf{y}_2(t)_i = (L_1^T \partial f(tP_2)\mathbf{1})_i$$

where

$$f(z) = \sum_{r=0}^{\infty} c_r z^r \quad \text{and} \quad \partial f(z) := \sum_{r=0}^{\infty} c_{r+1} z^r$$

Let $p_{r;k}(A)$ entrywise defined as

$$(p_{r;k}(A))_{ij} = \#\{(i_1 = i, \ldots, i_{r+1} = j) \,|\, \text{w/o cycles of length} \leq k\}$$

Let $p_{r;k}(A)$ entrywise defined as

$$(p_{r;k}(A))_{ij} = \#\{(i_1 = i, \ldots, i_{r+1} = j) \,|\, \text{w/o cycles of length} \leq k\}$$

Given $f(A) = \sum_{r=0}^{\infty} c_r t^r A^r$, let the non-$k$-cylcing counterpart of the matrix function $f(A)$ be:

$$F_k(A) := \sum_{r=0}^{\infty} c_r t^r p_{r;k}(A).$$

Let $p_{r;k}(A)$ entrywise defined as

$(p_{r;k}(A))_{ij} = \#\{(i_1 = i, \dots, i_{r+1} = j) \,|\, \text{w/o cycles of length} \leq k\}$

Given $f(A) = \sum_{r=0}^{\infty} c_r t^r A^r$, let the non-$k$-cylcing counterpart of the matrix function $f(A)$ be:

$$F_k(A) := \sum_{r=0}^{\infty} c_r t^r p_{r;k}(A).$$

Define the *non-k-cycling SC* of node i as:

$$\mathbf{x}_k(i) = (F_k(A))_{ii}$$

and the *non-k-cycling TC* of node i as:

$$\mathbf{y}_k(i) = (F_k(A)\mathbf{1})_i.$$

Question: Can we define define non–k–cycling subgraph centrality and total communicability indices, using entries or sum of entries of

$$c_0 p_{0;k}(A) + c_1 t p_{1;k}(A) + c_2 t^2 p_{2;k}(A) + c_3 t^3 p_{3;k}(A) + \cdots$$

where $p_{r;k}(A)$ generalizes A^r to the non–k–cycling setting and $t > 0$ is selected so that the series converges.

$$\mathbf{x}_k(t)_i = \left(\sum_{r=0}^{\infty} c_r t^r p_{r;k}(A) \right)_{ii}$$

and

$$\mathbf{y}_k(t)_i = \left(\sum_{r=0}^{\infty} c_r t^r p_{r;k}(A)\mathbf{1} \right)_{i}$$

$ABC\ldots$ easy as $123\ldots$

- For $k = 2, 3, \ldots$, given $P_{k-1} \in \mathbb{R}^{m_{k-1} \times m_{k-1}}$
- Source and target matrices $L_{k-1}, R_{k-1} \in \mathbb{R}^{m_k \times m_{k-1}}$ of P_{k-1}:

$$(L_{k-1})_{\underline{i}\underline{\ell}} = \begin{cases} 1 & \text{if } i_j = \ell_j \text{ for } j = 1, \ldots, k-1 \\ 0 & \text{otherwise} \end{cases}$$

$$(R_{k-1})_{\underline{i}\underline{\ell}} = \begin{cases} 1 & \text{if } i_{j+1} = \ell_j \text{ for } j = 1, \ldots, k-1 \\ 0 & \text{otherwise} \end{cases}$$

- $W_k = R_{k-1} L_{k-1}^T$ (kth order De Bruijn graphs of paths in the network.)
- **Non-k-cycling matrix**

$$P_k = W_k - \Delta_k,$$

where $\Delta_k = W_k \circ (W_k^T)^{k-1}$

$ABC\ldots$ easy as $123\ldots$

- For $k = 2, 3, \ldots,$ given $P_{k-1} \in \mathbb{R}^{m_{k-1} \times m_{k-1}}$
- Source and target matrices $L_{k-1}, R_{k-1} \in \mathbb{R}^{m_k \times m_{k-1}}$ of P_{k-1}:

$$(L_{k-1})_{\underline{i}\underline{\ell}} = \begin{cases} 1 & \text{if } i_j = \ell_j \text{ for } j = 1, \ldots, k-1 \\ 0 & \text{otherwise} \end{cases}$$

$$(R_{k-1})_{\underline{i}\underline{\ell}} = \begin{cases} 1 & \text{if } i_{j+1} = \ell_j \text{ for } j = 1, \ldots, k-1 \\ 0 & \text{otherwise} \end{cases}$$

- $W_k = R_{k-1} L_{k-1}^T$ (kth order De Bruijn graphs of paths in the network.)
- **Non-k-cycling matrix**

$$P_k = W_k - \Delta_k,$$

where $\Delta_k = W_k \circ (W_k^T)^{k-1}$

$ABC\ldots$ easy as $123\ldots$

- For $k = 2, 3, \ldots$, given $P_{k-1} \in \mathbb{R}^{m_{k-1} \times m_{k-1}}$
- Source and target matrices $L_{k-1}, R_{k-1} \in \mathbb{R}^{m_k \times m_{k-1}}$ of P_{k-1}:

$$(L_{k-1})_{\underline{i}\underline{\ell}} = \begin{cases} 1 & \text{if } i_j = \ell_j \text{ for } j = 1, \ldots, k-1 \\ 0 & \text{otherwise} \end{cases}$$

$$(R_{k-1})_{\underline{i}\underline{\ell}} = \begin{cases} 1 & \text{if } i_{j+1} = \ell_j \text{ for } j = 1, \ldots, k-1 \\ 0 & \text{otherwise} \end{cases}$$

- $W_k = R_{k-1} L_{k-1}^T$ (kth order De Bruijn graphs of paths in the network.)
- Non-k-cycling matrix

$$P_k = W_k - \Delta_k,$$

where $\Delta_k = W_k \circ (W_k^T)^{k-1}$

- For $k = 2, 3, \ldots$, given $P_{k-1} \in \mathbb{R}^{m_{k-1} \times m_{k-1}}$
- Source and target matrices $L_{k-1}, R_{k-1} \in \mathbb{R}^{m_k \times m_{k-1}}$ of P_{k-1}:

$$(L_{k-1})_{\underline{i}\underline{\ell}} = \begin{cases} 1 & \text{if } i_j = \ell_j \text{ for } j = 1, \ldots, k-1 \\ 0 & \text{otherwise} \end{cases}$$

$$(R_{k-1})_{\underline{i}\underline{\ell}} = \begin{cases} 1 & \text{if } i_{j+1} = \ell_j \text{ for } j = 1, \ldots, k-1 \\ 0 & \text{otherwise} \end{cases}$$

- $W_k = R_{k-1} L_{k-1}^T$ (kth order De Bruijn graphs of paths in the network.)
- **Non-k-cycling matrix**

$$P_k = W_k - \Delta_k,$$

where $\Delta_k = W_k \circ (W_k^T)^{k-1}$

$ABC\ldots$ easy as $123\ldots$

We want to concatenate $\underline{i} = (i_1, i_2, \ldots, i_k)$ and $\underline{j} = (j_1, j_2, \ldots, j_k)$ without them closing a k-cycle:

$$i_1 \to i_2 = j_1 \to i_3 = j_2 \to \cdots \to i_k = j_{k-1} \to j_k$$

and

$$i_1 \neq j_k$$

We want to concatenate $\underline{i} = (i_1, i_2, \ldots, i_k)$ and $\underline{j} = (j_1, j_2, \ldots, j_k)$ without them closing a k-cycle:

$$i_1 \to i_2 = j_1 \to i_3 = j_2 \to \cdots \to i_k = j_{k-1} \to j_k$$

and

$$i_1 \neq j_k$$

$$(W_k)_{\underline{i}\underline{j}} = (R_{k-1}L_{k-1}^T)_{\underline{i}\underline{j}} = \begin{cases} 1 & \text{if } i_{\ell+1} = j_\ell \text{ for } \ell = 1, \ldots, k-1 \\ 0 & \text{otherwise} \end{cases}$$

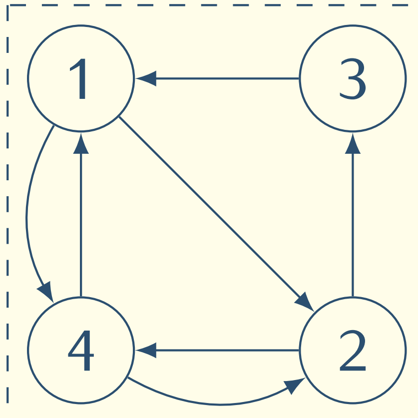
$$(P_k)_{\underline{i}\underline{j}} = (W_k - W_k \circ (W_k^T)^{k-1})_{\underline{i}\underline{j}} = \prod_{\ell=1}^{k-1} \delta_{i_{\ell+1}, j_\ell}(1 - \delta_{i_1 j_k})$$

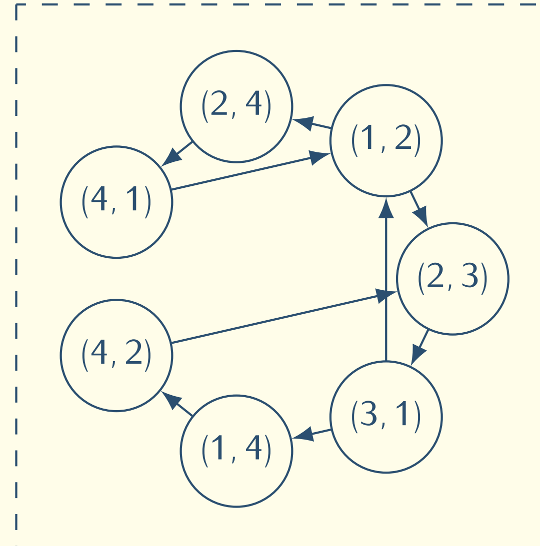
since

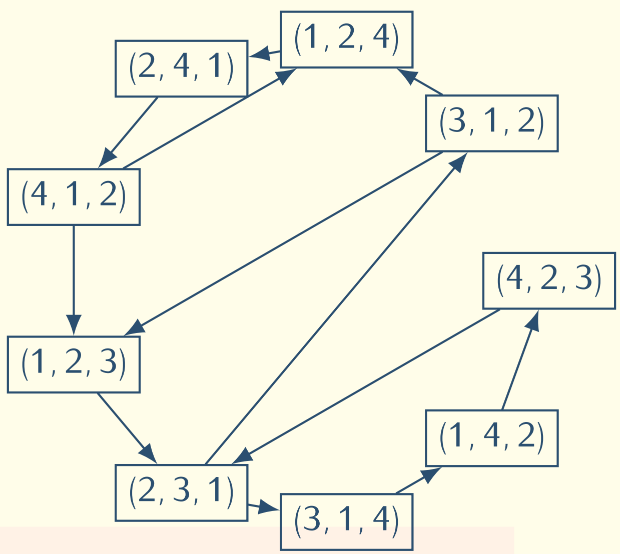
$$(W_k^T)^{k-1}_{\underline{i}\underline{j}} = \delta_{i_1 j_k}$$

where δ is the Kronecker delta.

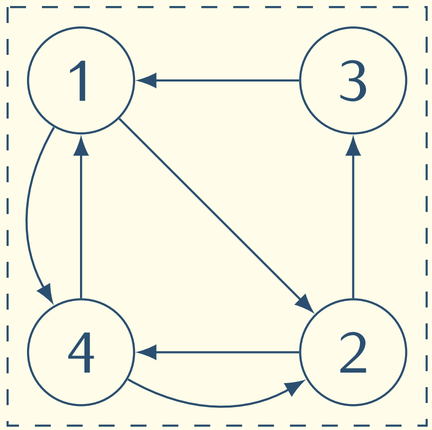
Small example

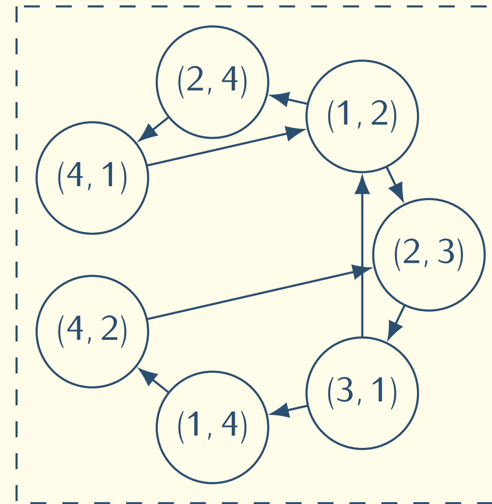


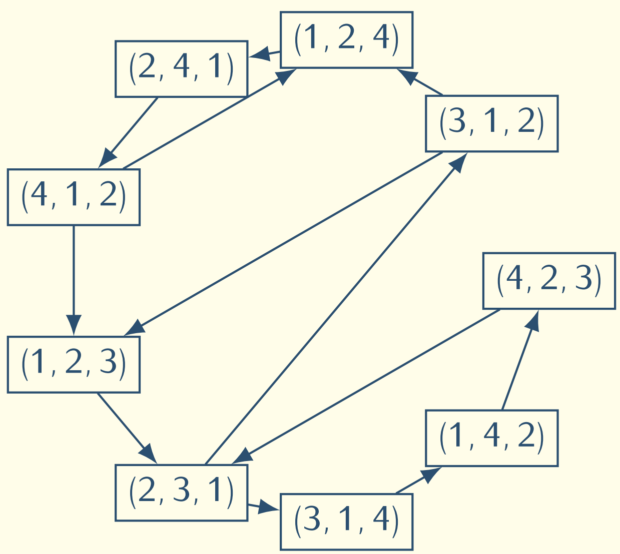


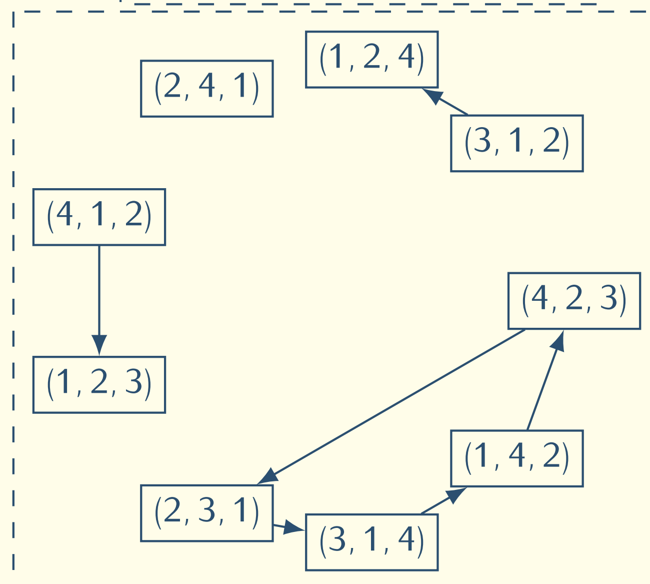


Small example









Counting non-k-cycling walks

It can be shown that

$$\mathcal{L}_{k-1}^T(P_k^r)\mathcal{R}_{k-1} = p_{r+k-1;k}(A)$$

where

$$\mathcal{L}_{k-1} = L_1L_2\cdots L_{k-1} \quad \text{and} \quad \mathcal{R}_{k-1} = R_1R_2\cdots R_{k-1}$$

and moreover, within the radius of convergence,

$$\mathbf{x}_k(t)_i = (\mathcal{L}_{k-1}^T\partial^{k-1}f(tP_k)\mathcal{R}_{k-1})_{ii} \quad \text{and} \quad \mathbf{y}_k(t)_i = (\mathcal{L}_{k-1}^T\partial^{k-1}f(tP_k)\mathbf{1})_i$$

Remark: The behaviour of the $\mathbf{x}_k(t)$ is very different from that of $\mathbf{y}_k(t)$: $\mathbf{x}_k(t)$ lacks "memory".

What else can we do?

- Generalize PageRank to the NBT framework.
- Downweight backtracking rather than completly eliminating it.
- Described the limiting behavior of

$$\mathbf{x}_k(t)_i = \left(\sum_{r=0}^{\infty} c_r t^r p_{r;k}(A)\right)_{ii} \quad \text{and} \quad \mathbf{y}_k(t)_i = \left(\sum_{r=0}^{\infty} c_r t^r p_{r;k}(A)\mathbf{1}\right)_{i}$$

  as $t \to 0$ and $t \to \bar{t}_k$, for $A^T = A$ and for all $k \geq 2$.
- *Ongoing work:* general limiting results, temporal networks and ad hoc numerical techniques.

Thank you

- F. A., D. J. Higham, V. Noferini, *"Non-backtracking walk centrality for directed networks"*, Journal of Complex Networks, 6 (2018), pp. 54–78.
- F. A., D. J. Higham, V. Noferini, *"Beyond non-backtracking: non-cycling network centrality measures*, Proc. R. Soc. A 476: 20190653 (2020).
  - M. Benzi, C. Klymko, *"On the limiting behavior of parameter-dependent network centrality measures*, SIAM J. Matrix Anal. Appl. 36, pp. 686–706 (2015).
  - M. Benzi and C. Klymko, *Total communicability as a centrality measure*, Journal of Complex Networks, 1 (2013), pp. 124–149.
  - E. Estrada and D. J. Higham, *Network properties revealed through matrix functions*, SIAM Review, 52 (2010), pp. 696–671.
  - P. Grindrod, D. J. Higham, and V. Noferini, *The deformed graph Laplacian and its applications to network centrality analysis*, SIAM Journal on Matrix Analysis and Applications, 39 (2018), pp. 310–341.
  - T. Martin, X. Zhang, M. E. J. Newman, *"Localization and centrality in networks"*, Phys. Rev. E90, 052808 (2014).
  - L. Torres, T. Eliassi-Rad, *"Non-backtracking spectrum: unitary eigenvalues and diagonalizability*, Preprint. arXiv:2007.13611 (2020).

LEVERHULME TRUST