Matrix methods in the analysis of complex networks

Spectral centralities

Dario Fasino

Rome, Univ. “Tor Vergata”, November 22–24, 2022

Spectral centralities

The different centrality measures formalize what it means for a node to be “important” or “influential”. Some (like degree) only look at local influence, others (like eigenvector centrality and PageRank) emphasize long-range effects, and others (like subgraph and Katz centrality) try to attain a balance between short- and long-range effects.

Spectral centralities

The different centrality measures formalize what it means for a node to be "important" or "influential". Some (like degree) only look at local influence, others (like eigenvector centrality and PageRank) emphasize long-range effects, and others (like subgraph and Katz centrality) try to attain a balance between short- and long-range effects.

Remark

For directed graphs it is often necessary to distinguish between hubs and authorities. Indeed, in a directed graph a node can play two roles: broadcaster and receiver of information.

Crude measures of hub and authority are provided by the out-degree and in-degree, respectively.

Spectral centralities

The different centrality measures formalize what it means for a node to be "important" or "influential". Some (like degree) only look at local influence, others (like eigenvector centrality and PageRank) emphasize long-range effects, and others (like subgraph and Katz centrality) try to attain a balance between short- and long-range effects.

Spectral centralities are based on eigenvalues and eigenvectors of adjacency matrices and implement a "mutual reinforcement" concept:
A node is "important" if it is connected to (or pointed by) other "important" nodes.

Eigenvector centrality

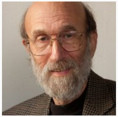
Mutual reinforcement concept

The "importance" of a node depends on the "importance" of nodes linked to it.

Perron eigenvector of symmetric adjacency matrix A:

$$\lambda x = Ax \quad \iff \quad \lambda x_i = \sum_{j=1}^{n} A_{ij} x_j = \sum_{j:i\sim j} x_j.$$

Existence, uniqueness, positivity: due to Perron–Frobenius thm., if the graph is strongly connected.



P. Bonacich.
Power and Centrality: A Family of Measures.
Amer. J. Sociology, 92 (1987) 1170–1182.

Eigenvector centrality

Mutual reinforcement concept

The "importance" of a node depends on the "importance" of nodes linked to it.

Perron eigenvector of symmetric adjacency matrix A:

$$\lambda x = Ax \quad \Longleftrightarrow \quad \lambda x_i = \sum_{j=1}^{n} A_{ij} x_j = \sum_{j:i\sim j} x_j.$$

$$x = \lim_{k\to\infty} A^k x_0 / \|A^k x_0\|.$$

$\|\cdot\| = \|\cdot\|_1 \rightsquigarrow$ The eigenvector centrality x_i of node $i \in V$ is the limit as $k \to \infty$ of the percentage of walks of length k which visit node i among all walks of length k on $\mathcal{G}$.

Thus, the eigenvector centrality measures the global influence of node i.

Eigenvector centrality

Mutual reinforcement concept

The "importance" of a node depends on the "importance" of nodes linked to it.

Perron eigenvector of symmetric adjacency matrix A:

$$\lambda x = Ax \quad \Longleftrightarrow \quad \lambda x_i = \sum_{j=1}^{n} A_{ij} x_j = \sum_{j:i\sim j} x_j .$$

If the graph is directed then we have two definitions:

Left eigenvector

$$A^T x = \lambda x$$

A node is important if it points to important nodes.

Right eigenvector

$$Ax = \lambda x$$

A node is important if it is pointed by important nodes.

Eigenvector centrality - Complements

Theorem

Let $A, \bar{A}$ be irreducible, nonegative matrices,
let $Ax = \rho x$ and $\bar{A}\bar{x} = \bar{\rho}\bar{x}$ be their Perron eigenpairs.
Let $\mathcal{I}$ be the index set of unchanged rows:

$$\mathcal{I} = \{i : A_{i,:} = \bar{A}_{i,:}\}.$$

Then,

$$\forall i \in \mathcal{I}, \qquad \frac{\bar{x}_i}{x_i} \leq \frac{\rho}{\bar{\rho}} \max_k \frac{\bar{x}_k}{x_k}.$$

In particular, if $\bar{\rho} > \rho$ then $\max_{k \in \mathcal{I}} \frac{\bar{x}_k}{x_k} < \max_k \frac{\bar{x}_k}{x_k}$.

E. Dietzenbacher. Perturbations of matrices: a theorem on the Perron vector and its applications to input-output models. *Z. Nationalökonom.*, 48 (1988), 389–412.

Eigenvector centrality - Complements

Theorem

Let $A, \bar{A}$ be irreducible, nonegative matrices,
let $Ax = \rho x$ and $\bar{A}\bar{x} = \bar{\rho}\bar{x}$ be their Perron eigenpairs.
Let $\mathcal{I}$ be the index set of unchanged rows:

$$\mathcal{I} = \{i : A_{i,:} = \bar{A}_{i,:}\}.$$

Then,

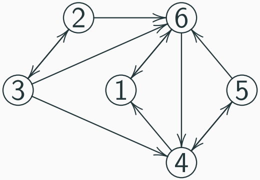
$$\forall i \in \mathcal{I}, \qquad \frac{\bar{x}_i}{x_i} \leq \frac{\rho}{\bar{\rho}} \max_k \frac{\bar{x}_k}{x_k}.$$

Corollary

If we add an edge to a graph then the largest relative increase in eigenvector centrality is attained by one of its nodes.

Katz centrality

A popularity contest among six people. Arrows represent votes. Who is the leader?



i	1	2	3	4	5	6
d_i^{in}	2	1	1	3	1	4
k_i	13	1	1	11.4	6.2	12.6

The in-degree is the received vote count, but it does not take into account the "status" of who casts the votes.

L. Katz. A new status index derived from sociometric analysis. *Psychometrica*, 18 (1953), 39–43.

Katz centrality

Idea: Take into account not only the votes cast directly but also the votes received by those who vote, and so on.

A person is prestigious if he is endorsed by prestigious people.

Katz centrality of node i:

$$k_i = \underbrace{\alpha \sum_j A_{ji}}_{\text{direct votes}} + \underbrace{\alpha^2 \sum_j [A^2]_{ji}}_{\text{1-indirect votes}} + \underbrace{\alpha^3 \sum_j [A^3]_{ji}}_{\text{2-indirect votes}} + \cdots$$

Matrix-vector notation: Let $k = (k_1, \ldots, k_n)^{\mathrm{T}}$. Then,

$$\begin{aligned} k &= \alpha A^{\mathrm{T}} e + (\alpha A^{\mathrm{T}})^2 e + (\alpha A^{\mathrm{T}})^3 e + \cdots \\ &= \sum_{i=1}^{\infty} (\alpha A^{\mathrm{T}})^i e. \end{aligned}$$

Theorem

If $|\alpha| < 1/\rho(A)$ then

$$I + \alpha A + \alpha^2 A^2 + \cdots = \sum_{k=0}^{\infty} \alpha^k A^k = (I - \alpha A)^{-1}.$$

Katz centrality

If $|\alpha| < 1/\rho(A)$ then

$$k = (I - \alpha A)^{-1} e - e.$$

Apart of the constant term e, the vector k of Katz indices is the solution of the linear system $(I - \alpha A^T)x = e$. Note that $I - \alpha A$ is a nonsingular M-matrix, in particular $(I - \alpha A)^{-1} \geq 0$.

Katz centrality (cont.)

In the case of a directed network one can use the solution vectors of the linear systems

$$(I - \alpha A)x = e \quad \text{and} \quad (I - \alpha A^{T})y = e$$

to rank hubs and authorities, respectively. Iterative methods are normally used for this task. Note that for an undirected graph the condition number of $I - \alpha A$ is bounded by $2/(1 - \alpha\rho(A))$. For α not too close to $1/\rho(A)$, convergence is usually fast.

Theorem

- For $\alpha \to 0$ Katz ranking reduces to degree ranking.
- For $\alpha \to 1/\rho(A)$ Katz centrality reduces to eigenvector centrality.

M. Benzi, C. Klymko. *SIMAX* 36 (2015), 686–706.

Katz centrality — example

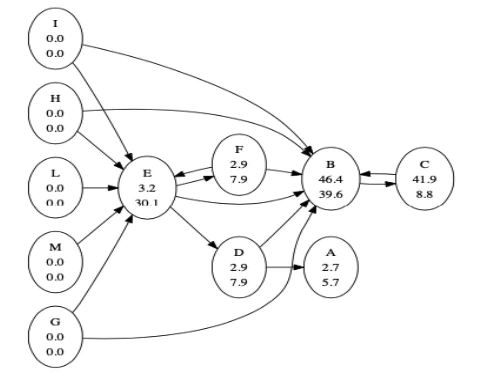


Figure 1: An example of the Katz model using $\alpha = .9$ (top) and $\alpha = .1$ (bottom). Here $\rho(A) = 1$.

M. Franceschet. PageRank: Standing on the shoulders of giants. *Comm. ACM*, 54 (2011), 92–101.

PageRank

A Web page is important if it is pointed to by other important pages.

Let $\mathcal{G}$ be a directed graph where $d_i^{out} > 0$ for all $i \in V$.
For any $0 < \alpha < 1$ the equation

$$p_i = 1 - \alpha + \alpha \sum_{j:j \to i} \frac{p_j}{d_j^{out}},$$

defines the PageRank centrality of node $i \in V$.
Originally introduced by S. Brin and L. Page (1999) to rank web pages in the Google search engine.

D. Gleich. PageRank beyond the web.
SIAM Rev. 57 (2015), 321–363.

A list of more than 20 PageRank-related centrality indices currently used within different domains including bibliometry, social networks, literature, biology...

PageRank

A Web page is important if it is pointed to by other important pages.

Let $\mathcal{G}$ be a directed graph where $d_i^{out} > 0$ for all $i \in V$.
For any $0 < \alpha < 1$ the equation

$$p_i = 1 - \alpha + \alpha \sum_{j:j \to i} \frac{p_j}{d_j^{out}},$$

defines the PageRank centrality of node $i \in V$.
In matrix form, $(I - \alpha M)p = (1-\alpha)e$ where $M_{ij} = A_{ji}/d_j^{out}$ is column stochastic.

Basic iterative method

The iteration

$$p^{(k+1)} = \alpha M p^{(k)} + (1-\alpha)e$$

is convergent:

$$\|p^{(k+1)} - \bar{p}\| \leq \alpha \|p^{(k)} - \bar{p}\|.$$

PageRank

A Web page is important if it is pointed to by other important pages.

Let $\mathcal{G}$ be a directed graph where $d_i^{out} > 0$ for all $i \in V$.
For any $0 < \alpha < 1$ the equation

$$p_i = 1 - \alpha + \alpha \sum_{j:j \to i} \frac{p_j}{d_j^{out}},$$

defines the PageRank centrality of node $i \in V$.
Equivalently, $p = (p_1, \ldots, p_n)^{\mathrm{T}}$ is a Perron eigenvector of the Google matrix

$$G = \alpha \underbrace{A^{\mathrm{T}} \,\mathrm{Diag}(1/d_1^{out}, \ldots, 1/d_n^{out})}_{M} + \frac{1-\alpha}{n} ee^{\mathrm{T}}.$$

Apart of a normalization, the basic iterative method coincides with the power method for G.

A Web page is important if it is pointed to by other important pages.

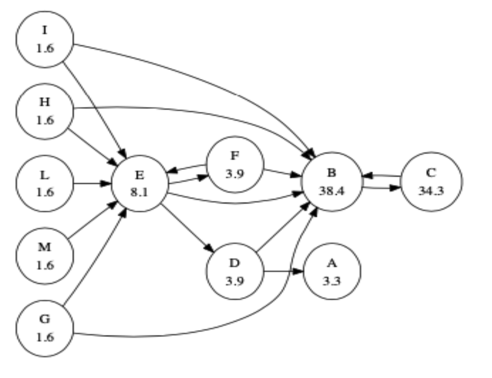


Figure 2: A PageRank instance using $\alpha = .85$. Scores normalized so that $\sum_i p_i = 100$.

Hyperlinked Induced Topics Search (HITS)

J. Kleinberg. Authoritative sources in a hyperlinked environment, *J. ACM*, 48 (1999), 604–632.

Internet is represented by a digraph $\mathcal{G}$ where nodes represent web pages and directed edges represent hyperlinks.

Kleinberg's idea

Each node in a directed network is associated to two scores:

- The hub score: quantifies the goodnes of that node as a "portal" or access point to informative nodes
- The authority score: quantifies the "informative quality" of that node.

HITS (cont.)

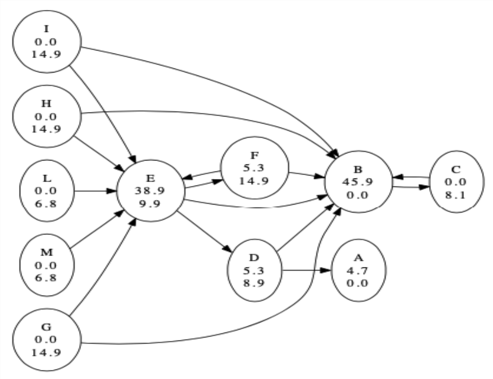
Mutual reinforcement concept

A good hub is a node that points to good authorities.

A good authority is a node that is pointed by good hubs.

$$\sigma h_i = \sum_{j:i \rightarrow j} a_j$$

$$\sigma a_i = \sum_{j:j \rightarrow i} h_j$$



HITS (cont.)

Mutual reinforcement concept

A good hub is a node that points to good authorities.

A good authority is a node that is pointed by good hubs.

$$A = \text{adj. matrix} \quad \Rightarrow \quad \begin{cases} \sigma h = Aa & \text{(hub scores)} \\ \sigma a = A^{\mathrm{T}} h & \text{(auth. scores)} \end{cases}$$

$$\sigma^2 h = M_{\text{hub}} h, \quad M_{\text{hub}} = AA^{\mathrm{T}} = \text{hub matrix};$$
$$\sigma^2 a = M_{\text{auth}} a, \quad M_{\text{auth}} = A^{\mathrm{T}} A = \text{authority matrix}.$$

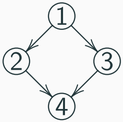
- $M_{\text{hub}}, M_{\text{auth}}$ are often reducible (Perron eigenpairs not unique)
- If $A = A^T$ then $h = a =$ eigenvector centrality.

HITS (cont.)

We say that HITS behaves fairly on a digraph $\mathcal{G}$
when the following two conditions are met:
(1) hub/auth scores are unique, apart of normalization;
(2) $d_i^{\mathrm{out}} > 0 \Rightarrow h_i > 0$ or, equivalently, $d_i^{\mathrm{in}} > 0 \Rightarrow a_i > 0$.

An example of unfair behaviour:

$\mathcal{G}$:



$$\rightsquigarrow \qquad M_{\mathrm{hub}} = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Any vector $h = (\alpha, \beta, \beta, 0)^{\mathrm{T}}$ with $\alpha, \beta \geq 0$ is a valid hub vector.

HITS (cont.)

We say that HITS behaves fairly on a digraph $\mathcal{G}$
when the following two conditions are met:
(1) hub/auth scores are unique, apart of normalization;
(2) $d_i^{\mathrm{out}} > 0 \Rightarrow h_i > 0$ or, equivalently, $d_i^{\mathrm{in}} > 0 \Rightarrow a_i > 0$.

Theorem

Let $\widehat{M}_{hub}$ be the matrix obtained by removing null rows and columns from M_{hub}. HITS behaves fairly if and only if $\widehat{M}_{hub}$ is irreducible.
(Analogous statement with *auth* in place of *hub*.)

HITS (cont.)

We say that HITS behaves fairly on a digraph $\mathcal{G}$
when the following two conditions are met:
(1) hub/auth scores are unique, apart of normalization;
(2) $d_i^{\mathrm{out}} > 0 \Rightarrow h_i > 0$ or, equivalently, $d_i^{\mathrm{in}} > 0 \Rightarrow a_i > 0$.

If $\mathcal{G}$ is weakly connected then adding a (weighted) loop on every node makes HITS behave fairly ($M_{\mathrm{hub}}, M_{\mathrm{auth}}$ become irreducible)

Subgraph centrality

Subgraph centrality measures the centrality of a node by taking into account the number of subgraphs the node "participates" in.

This is done by counting, for all $k = 1, 2, \dots$ the number of closed walks in $\mathcal{G}$ starting and ending at node i, with longer walks being penalized (given a smaller weight).

It is sometimes useful to introduce a tuning parameter $\beta > 0$ to simulate external influences on the network, for example, increased tension in a social network, financial distress in the banking system, etc.

E. Estrada, J. A. Rodríguez-Velásquez, Phys. Rev. E, 2005.

Subgraph centrality (cont.)

Recall that

- $(A^k)_{ii} = \#$ of closed walks of length k based at node i,
- $(A^k)_{ij} = \#$ of walks of length k that connect nodes i and j.

Using $\beta^k/k!$ as weights leads to the notion of subgraph centrality:

$$\begin{aligned} SC(i,\beta) &= \left[I + \beta A + \frac{\beta^2}{2!}A^2 + \frac{\beta^3}{3!}A^3 + \dots \right]_{ii} \\ &= \left[\mathrm{e}^{\beta A} \right]_{ii}. \end{aligned}$$

Note that $SC(i,\beta) \geq 1$. The weights are needed to “penalize” longer walks, and to make the power series converge.

Subgraph centrality (cont.)

It is sometimes desirable to normalize the subgraph centrality of a node by the sum

$$EE(\mathcal{G}) = \sum_{i=1}^{n} SC(i) = \sum_{i=1}^{n} [\mathrm{e}^{\beta A}]_{ii} = \mathrm{trace}(\mathrm{e}^{\beta A})$$

of all the subgraph centralities. The quantity $EE(\mathcal{G})$ is known as the Estrada index of the graph $\mathcal{G}$. This index has powerful discrimination in measuring the robustness of complex networks to link removal.

Communicability

Communicability measures how "easy" it is to send a message from node i to node j in a network by a weighted sum of walks $i \rightsquigarrow j$:

- $C(i,j) = \left[\mathrm{e}^{\beta A}\right]_{ij} = \sum_{k=0}^{\infty} \beta^k (A^k)_{ij}/k!$
- $C(i,j) = \left[(I - \alpha A)^{-1}\right]_{ij} = \sum_{k=0}^{\infty} \alpha^k (A^k)_{ij}$

Communicability has been successfully used to identify bottlenecks in networks and for community detection.

E. Estrada, N. Hatano. Communicability in complex networks. *Phys. Rev. E*, 77 (2008), 036111;

E. Estrada, D. J. Higham. Network properties revealed through matrix functions. *SIAM Rev.*, 52 (2010), 696–714.

Communicability

Communicability measures how "easy" it is to send a message from node i to node j in a network by a weighted sum of walks $i \rightsquigarrow j$:

- $C(i,j) = \left[\mathrm{e}^{\beta A}\right]_{ij} = \sum_{k=0}^{\infty} \beta^k (A^k)_{ij}/k!$
- $C(i,j) = \left[(I - \alpha A)^{-1}\right]_{ij} = \sum_{k=0}^{\infty} \alpha^k (A^k)_{ij}$

The total communicability of node $i \in V$ is defined as

$$TC(i) = \sum_{j=1}^{n} C(i,j).$$

Highly connected networks, such as small-world networks without bottlenecks, can be expected to have a high total communicability.

Communicability

Communicability measures how “easy” it is to send a message from node i to node j in a network by a weighted sum of walks $i \rightsquigarrow j$:

- $C(i,j) = \left[\mathrm{e}^{\beta A}\right]_{ij} = \sum_{k=0}^{\infty} \beta^k (A^k)_{ij}/k!$
- $C(i,j) = \left[(I - \alpha A)^{-1}\right]_{ij} = \sum_{k=0}^{\infty} \alpha^k (A^k)_{ij}$

Communicability distance

Let $\mathcal{G}$ be undirected. Then

$$\mathrm{Cdist}(i,j) = \sqrt{C(i,i) + C(j,j) - 2C(i,j)}$$

is a distance on $\mathcal{G}$. Indeed,

$$\mathrm{Cdist}(i,j)^2 = (e_i - e_j)^{\mathrm{T}}\, \mathrm{e}^{\beta A}(e_i - e_j) = \|e_i - e_j\|_E^2$$

where $E = \mathrm{e}^{\beta A/2}$ is symmetric and positive definite.

Communicability

Communicability measures how "easy" it is to send a message from node i to node j in a network by a weighted sum of walks $i \rightsquigarrow j$:

- $C(i,j) = \left[\mathrm{e}^{\beta A}\right]_{ij} = \sum_{k=0}^{\infty} \beta^k (A^k)_{ij}/k!$
- $C(i,j) = \left[(I - \alpha A)^{-1}\right]_{ij} = \sum_{k=0}^{\infty} \alpha^k (A^k)_{ij}$

The average total communicability of $\mathcal{G}$,

$$TC(\mathcal{G}) = \frac{1}{n} \sum_{i,j=1}^{n} C(i,j) = \frac{1}{n} e^{\mathrm{T}} \mathrm{e}^{\beta A} e$$

(usually with $\beta = 1$) provides a global measure of how "well-connected" a network is, and can be used to compare different network structures.

This is another spectral measure that can be computed very efficiently, even for large graphs, since it involves computing linear functionals of $\mathrm{e}^{\beta A}$ or $(I - \alpha A)^{-1}$, and Krylov subspace methods are very good at this!

Krylov subspace methods in a nutshell

The main idea behind Krylov methods is the following:

- A nested sequence of suitable low-dimensional Krylov subspaces $\langle b, Ab, A^2 b, \ldots, A^{k-1} b\rangle$ is generated, $0 < k \ll n$.
- The original problem is projected onto these subspaces.
- The (small) projected problems are solved "exactly".
- The approximate solution is projected back to the original n-dimensional space.

Krylov subspace methods are examples of polynomial approximation methods, where $f(A)v$ is approximated by $p(A)v$, where p is a (low-degree) polynomial. This polynomial approximation is optimal in some sense. If f is a smooth function, convergence is usually fast; for entire functions it is superexponential in the degree of the polynomial.

Limiting behaviors

- $SC(i, \beta) = (\mathrm{e}^{\beta A})_{ii}$
- $TC(i, \beta) = (\mathrm{e}^{\beta A} e)_i$

Theorem

Let $\mathcal{G}$ be a simple graph.

- For $\beta \to 0$ the rankings produced by $SC(i, \beta)$ and $TC(i, \beta)$ reduce to degree ranking.
- For $\beta \to \infty$ the centralities $SC(i, \beta)$ and $TC(i, \beta)$ reduce to eigenvector centrality.

M. Benzi, C. Klymko. On the limiting behavior of parameter-dependent network centrality measures. *SIMAX* 36 (2015), 686–706.

Other matrix functions

Other matrix functions of interest are

$$\cosh(A) = \tfrac{1}{2}(e^{A} + e^{-A}), \qquad \sinh(A) = \tfrac{1}{2}(e^{A} - e^{-A}),$$

which contain the even part and the odd part of the power series of e^{A}, respectively:

$$e^{A} = I + A + \tfrac{1}{2!}A^2 + \tfrac{1}{3!}A^3 + \dots$$

Thus $\cosh(A)$ and $\sinh(A)$ correspond to considering only walks of even and odd length, respectively.

Other matrix functions

Other matrix functions of interest are

$$\cosh(A) = \frac{1}{2}(\mathrm{e}^{A} + \mathrm{e}^{-A}), \qquad \sinh(A) = \frac{1}{2}(\mathrm{e}^{A} - \mathrm{e}^{-A}),$$

which contain the even part and the odd part of the power series of e^{A}, respectively:

$$\mathrm{e}^{A} = I + A + \frac{1}{2!}A^{2} + \frac{1}{3!}A^{3} + \dots$$

A bipartitiveness measure

In a bipartite graph, all closed walks have odd lengths and the eigenvalues of A occur in $\pm$-pairs. Hence $\mathrm{trace}(\sinh(A)) = 0$. Thus the quantity

$$B(\mathcal{G}) = \frac{\mathrm{trace}(\cosh(A))}{\mathrm{trace}(\mathrm{e}^{A})}$$

provides a measure of how "close" a graph $\mathcal{G}$ is to being bipartite.

Other matrix functions

Other matrix functions of interest are

$$\cosh(A) = \frac{1}{2}(\mathrm{e}^{A} + \mathrm{e}^{-A}), \qquad \sinh(A) = \frac{1}{2}(\mathrm{e}^{A} - \mathrm{e}^{-A}),$$

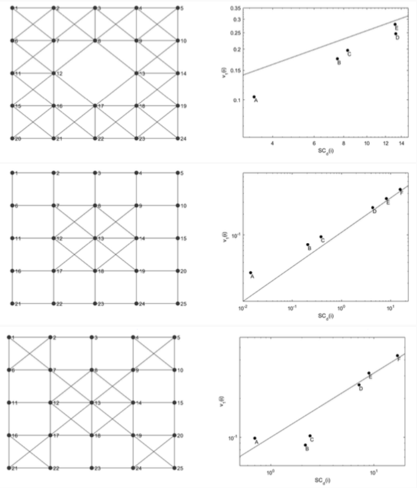
which contain the even part and the odd part of the power series of e^{A}, respectively:

$$\mathrm{e}^{A} = I + A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \dots$$

Odd subgraph centrality of node $i \in V$: $SC^{odd}(i) = (\sinh(A))_{ii}$.

Let $v = (v_1, \dots, v_n)^{\mathrm{T}}$ be the eigenvector centrality.
Plotting $SC^{odd}(1), \dots, SC^{odd}(n)$ vs. $v_1, \dots, v_n$ allows to grasp essential features of $\mathcal{G}$.

Other matrix functions



E. Estrada. Topological structural classes of complex networks. *Phys. Rev. E*, 75 (2007), 016103.