Centrality Measures

Network Science '21: Session 3

Carlo Campajola

Blockchain & Distributed Ledger Technologies Group

Claudio J. Tessone

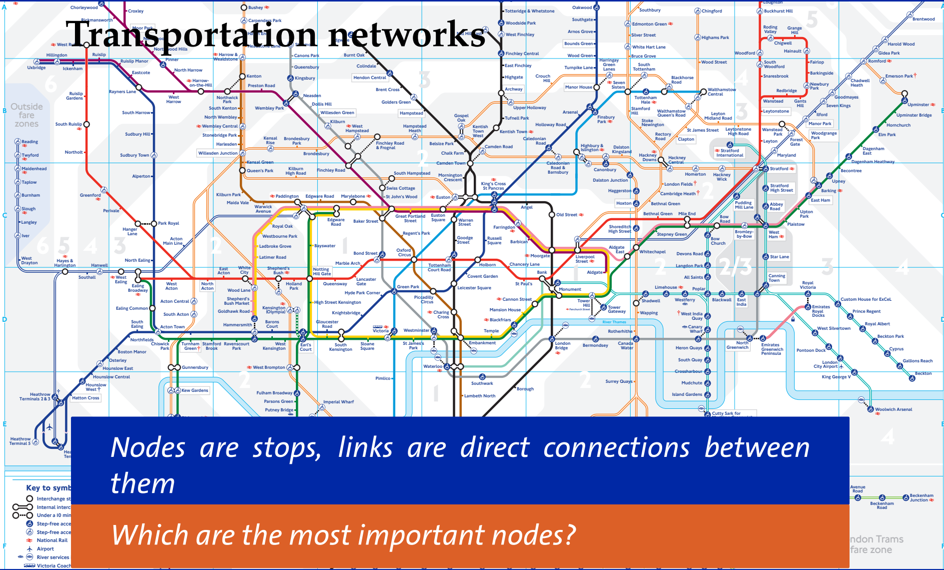
Lecture Objectives

1. From a structural point of view, what makes nodes important?
2. The multiple notions of centrality
3. How to rank important nodes?

Outlook

Outlook

Transportation networks



Nodes are stops, links are direct connections between them

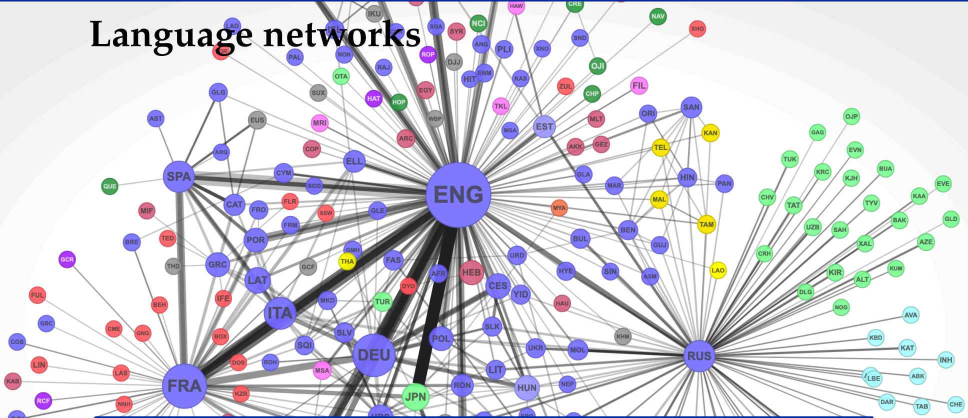
Which are the most important nodes?

What does most important mean?

Do we mean...

- Most connected
- Closest to others
- Can perform meaningful commutes
- ...

Language networks



Nodes are languages, links are likelihood to be co-spoken (Ronen, et al. (2014))

Which are the most important nodes?

What does most important mean?

Do we mean...

- Most popular?
- Best bridge to other languages?
- Co-spoken by other important languages?
- ...

Foreword on notation

- $\mathcal{G}$: a graph with sets of nodes $\mathcal{V}$ and edges $\mathcal{E}$
- $\mathbb{A}$: Adjacency matrix with elements $a_{i,j}$
- N: Number of nodes
- $\mathcal{N}_i$: Is the set of neighbours of i
- $|\mathcal{X}|$: represents the number of elements of $\mathcal{X}$

One of the primary uses of graph theory in network analysis is the identification of the "most important" actors in the network

adapted from Wasserman and Faust (1995)

Centrality

How can we rank node's importance?

Important actors (elements, nodes) are also located in strategic locations within the network

- J. Moreno (1934) already talked about *sociometric stars*
- Bavelas (1948) was talking about centrality of nodes as an indicator of their prominence

Degree centrality

Intuition

The degree centrality is equivalent to say that the importance of a node is given by *how much it is connected to others in the network*

The **degree centrality** of a node is just measured by the number of links it has to others in a network

$$k_i = \sum_{j=1}^{N} a_{ij}$$

It is a local property. It depends on node and its neighbours

Degree centrality

Intuition

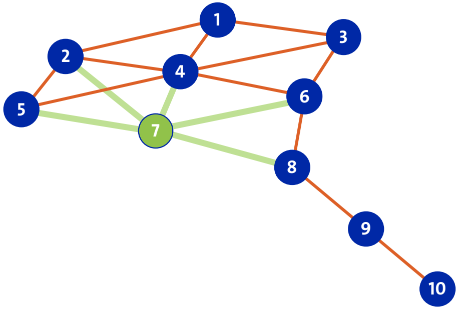
The degree centrality is equivalent to say that the importance of a node is given by *how much it is connected to others in the network*

The **degree centrality** of a node is just measured by the number of links it has to others in a network

$$k_i = \sum_{j=1}^{N} a_{ij}$$

It is a local property. It depends on node and its neighbours

Degree centrality example



In this network, node 7 has a degree centrality $k_7 = 5$

Degree centrality normalisation

It is customary to normalise the importance of nodes between 0 and 1. This can be achieved in two ways:

- **Normalise by the maximum observed degree:**

$$CK_i = \frac{k_i}{\max_j(k_j)}$$

- **Normalise by the maximum possible degree:**

$$CK_i = \frac{k_i}{N-1}$$

Degree centrality normalisation

It is customary to normalise the importance of nodes between 0 and 1. This can be achieved in two ways:

- **Normalise by the maximum observed degree:**

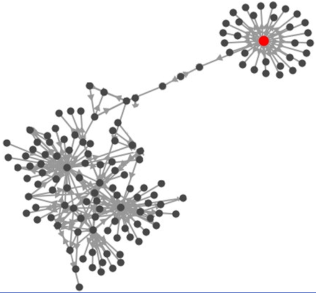
$$CK_i = \frac{k_i}{\max_j(k_j)}$$

- **Normalise by the maximum possible degree:**

$$CK_i = \frac{k_i}{N-1}$$

Degree centrality: pros and cons

- Pro: a measure of popularity in a social network
- Pro: simple and often correlated with more complex metrics
- Con: does not tell anything about "where" is the node
- Con: does not discriminate "quality" of connections



Closeness centrality

- If $d_{i,j}$ is the distance (i.e. the shortest path length) between nodes i and j
- Then,

$$\sum_{j=1}^{N} d_{i,j}$$

is the sum of all distances from node i to all other nodes

Closeness centrality

Intuition

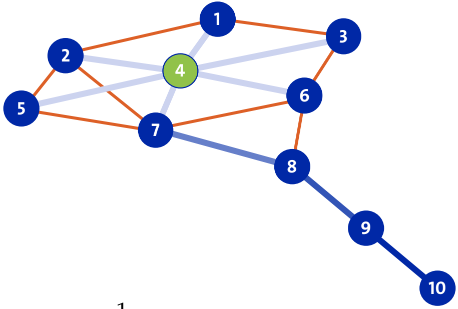
Closeness centrality is stating that the importance of a node is given by *how short are the distances from it to all others*

$$CC_i = \frac{N-1}{\sum_{j=1}^{N} d_{i,j}}$$

- It measures how fast the information from one node can reach the whole network

It is a global property. It depends on the whole network

Closeness centrality example

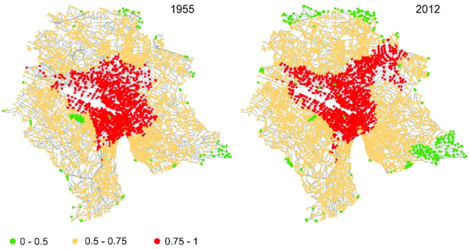


$$CC_4 = \left(\frac{\sum_j d_{4,j}}{N-1}\right)^{-1} = \left(\frac{1+1+1+1+1+1+2+3+4}{9}\right)^{-1}$$

$$= \frac{3}{5}$$

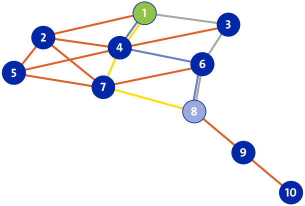
Closeness centrality: pros and cons

- Pro: identifies top sources of spreading - great for information, diseases, public transport, ...
- Pro: "topologically" meaningful concept of centrality
- Con: requires computing $\binom{N}{2}$ distances - very demanding for not small N
- Con: sources are central, not bottlenecks



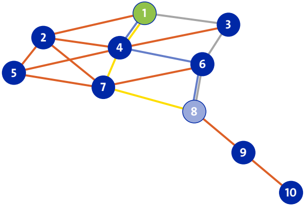
Casali, Y. & Heinimann, H., A topological analysis of growth in the Zurich road network, *Comput Environ Urban Syst*, 2019

Betweenness centrality: Concepts



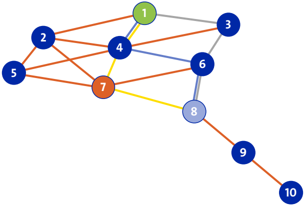
- $\mathcal{P}_{j,k}$ is the set of shortest paths between nodes j and k... *e.g. from nodes 1 and 8 there are three shortest paths*
- $\mathcal{P}^{7}_{j,k}$ is the set of shortest paths between nodes j and k that pass through i... e.g. from nodes 1 and 8 there is one shortest path that goes through 7
- $\left|\mathcal{P}_{j,k}\right|$ and $\left|\mathcal{P}^{i}_{j,k}\right|$ are the number of these shortest paths

Betweenness centrality: Concepts



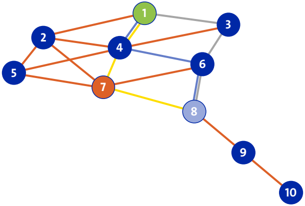
- $\mathcal{P}_{j,k}$ is the set of shortest paths between nodes j and k... *e.g. from nodes 1 and 8 there are three shortest paths*
- $\mathcal{P}^{7}_{j,k}$ is the set of shortest paths between nodes j and k that pass through i... e.g. from nodes 1 and 8 there is one shortest path that goes through 7
- $\left|\mathcal{P}_{j,k}\right|$ and $\left|\mathcal{P}^{i}_{j,k}\right|$ are the number of these shortest paths

Betweenness centrality: Concepts



- $\mathcal{P}_{j,k}$ is the set of shortest paths between nodes j and k... *e.g. from nodes 1 and 8 there are three shortest paths*
- $\mathcal{P}^{7}_{j,k}$ is the set of shortest paths between nodes j and k that pass through i... e.g. from nodes 1 and 8 there is one shortest path that goes through 7
- $\left|\mathcal{P}_{j,k}\right|$ and $\left|\mathcal{P}^{i}_{j,k}\right|$ are the number of these shortest paths

Betweenness centrality: Concepts



- $\mathcal{P}_{j,k}$ is the set of shortest paths between nodes j and k... *e.g. from nodes 1 and 8 there are three shortest paths*
- $\mathcal{P}^{7}_{j,k}$ is the set of shortest paths between nodes j and k that pass through i... e.g. from nodes 1 and 8 there is one shortest path that goes through 7
- $\left|\mathcal{P}_{j,k}\right|$ and $\left|\mathcal{P}^{i}_{j,k}\right|$ are the number of these shortest paths

Betweenness centrality

Intuition

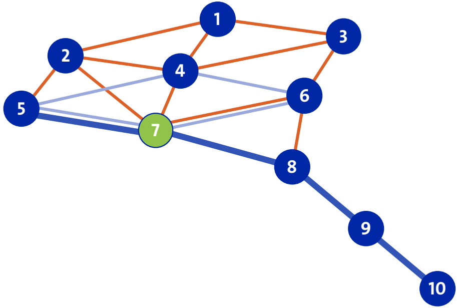
The importance of each node is given by how much they can route / channel information flow within the network

$$CB_i = \sum_{j,k \in \mathcal{V}} \frac{\left|\mathcal{P}^{i}_{j,k}\right|}{\left|\mathcal{P}_{j,k}\right|}$$

- Nodes with high betweenness have a large influence on the flow of information through the network
- Brokerage role

It is a global property. It depends on the whole network

Betweenness centrality example

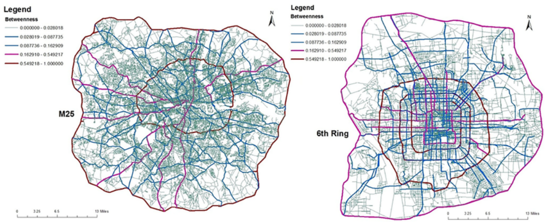


Consider all pairs of nodes;

- Add 1 for every **unique** shortest path going through 7
- Add $\frac{1}{2}$ for every shortest path going through 7, if there is another one not going through 7

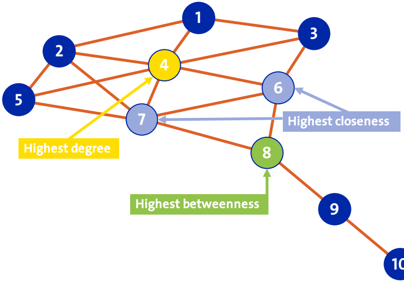
Betweenness centrality: pros and cons

- Pro: identifies high-throughput nodes - great for vulnerability analysis
- Pro: “topologically” meaningful centrality
- Con: again, $\binom{N}{2}$ shortest paths to compute
- Con: does not identify sources, only bottlenecks



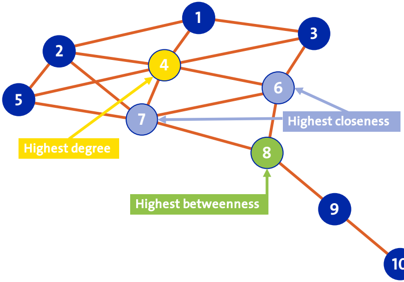
Wang, J., Resilience of Self-Organised and Top-Down Planned Cities - A Case Study on London and Beijing Street Networks, *PLoS ONE*, 2015

Different centrality measures



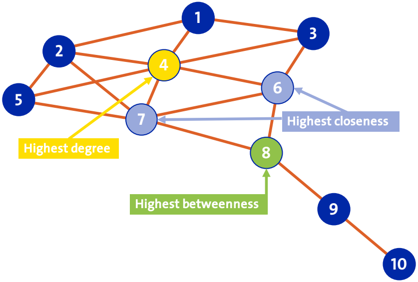
- The **degree centrality** represents **popularity**"
- The **closeness centrality** shows how easily a node can **reach all others**
- The **betweenness centrality** measures **control** of information flow

Different centrality measures



- The **degree centrality** represents **popularity**”
- The **closeness centrality** shows how easily a node can **reach all others**
- The **betweenness centrality** measures **control** of information flow

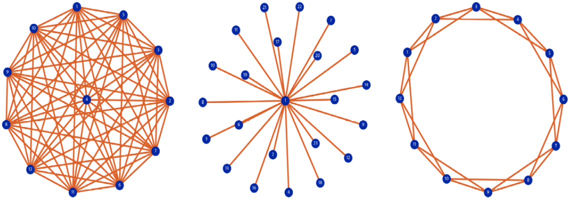
Different centrality measures



- The **degree centrality** represents **popularity**"
- The **closeness centrality** shows how easily a node can **reach all others**
- The **betweenness centrality** measures **control** of information flow

Depending on the measure, different nodes may be the most important ones. The ranking order depends on the network structure

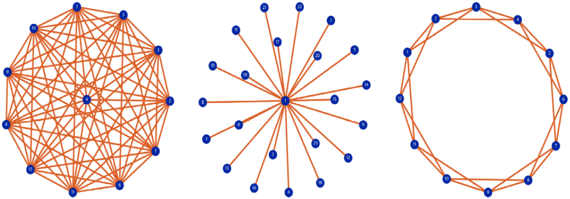
Centralisation



Which one is the most *centralised*?

+ In the fully connected and regular networks, the nodes are homogeneous: All have the same number of opportunities to select resources from partners
+ In the star network, one node has full power on all of the others

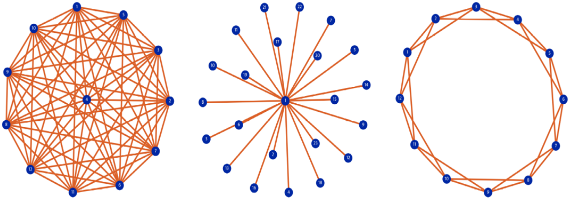
Centralisation



Which one is the most *centralised*?

- In the fully connected and regular networks, the nodes are homogeneous: All have the same number of opportunities to select resources from partners
- In the star network, one node has full power on all of the others

Centralisation



Which one is the most *centralised*?

- In the fully connected and regular networks, the nodes are homogeneous: All have the same number of opportunities to select resources from partners
- In the star network, one node has full power on all of the others

Centralisation

Freeman (1978) proposed a global measure to calculate the degree of centralisation of a particular network:

$$C = \frac{\sum_{i \in \mathcal{V}} C(\star) - C(i)}{\max \sum_{i \in \mathcal{V}} C(\star) - C(i)}$$

Where

- $C(\star)$ is the maximum centrality in a node of the network
- $\max \sum_{i \in \mathcal{V}} C(\star) - C(i)$ represents the maximum level of centralisation that can be attained in a network of the same size

Centralisation

Freeman (1978) proposed a global measure to calculate the degree of centralisation of a particular network:

$$C = \frac{\sum_{i \in \mathcal{V}} C(\star) - C(i)}{\max \sum_{i \in \mathcal{V}} C(\star) - C(i)}$$

Where

- $C(\star)$ is the maximum centrality in a node of the network
- $\max \sum_{i \in \mathcal{V}} C(\star) - C(i)$ represents the maximum level of centralisation that can be attained in a network of the same size

Eigenvector Centrality

Intuition

A node that is more important if it is connected to important nodes

If $\mathbb{A}$ is the adjacency matrix of the network $\mathcal{G}$ and λ is a constant, we define the centrality of node i by

$$v_i = \frac{1}{\lambda} \sum_{j \in N_i} v_j = \frac{1}{\lambda} \sum_{j=1}^{N} a_{ij} v_j$$

Or, in matrix notation

$$\mathbb{A} \cdot \mathbb{v} = \lambda \mathbb{v}$$

Eigenvector Centrality

Intuition

A node that is more important if it is connected to important nodes

If $\mathbb{A}$ is the adjacency matrix of the network $\mathcal{G}$ and λ is a constant, we define the centrality of node i by

$$v_i = \frac{1}{\lambda} \sum_{j \in \mathcal{N}_i} v_j = \frac{1}{\lambda} \sum_{j=1}^{N} a_{ij} v_j$$

Or, in matrix notation

$$\mathbb{A} \cdot \mathbb{v} = \lambda \mathbb{v}$$

Eigenvector Centrality

Intuition

A node that is more important if it is connected to important nodes

If $\mathbb{A}$ is the adjacency matrix of the network $\mathcal{G}$ and λ is a constant, we define the centrality of node i by

$$v_i = \frac{1}{\lambda} \sum_{j \in \mathcal{N}_i} v_j = \frac{1}{\lambda} \sum_{j=1}^{N} a_{ij} v_j$$

Or, in matrix notation

$$\mathbb{A} \cdot \mathbb{v} = \lambda \mathbb{v}$$

Refresher - Eigenvalues and Eigenvectors

Given a $N \times N$ matrix $\mathbb{A}$, **eigenvalues** are the zeros of its **characteristic polynomial**, i.e. λ $s.t.$ $\det(\lambda\mathbb{I} - \mathbb{A}) = 0$

If there are N distinct eigenvalues, then $\mathbb{A}$ is **diagonalizable**, meaning there exists an invertible matrix $\mathbb{V}$ s.t.

$$\mathbb{V}^{-1}\mathbb{A}\mathbb{V} = \Lambda$$

where Λ is a diagonal matrix with the eigenvalues on the diagonal. **Eigenvectors** are the columns of $\mathbb{V}$.

If the matrix is diagonalizable and positive, the **Perron-Frobenius theorem** applies

Eigenvector Centrality

Perron-Frobenius theorem

The theorem states that if $\mathbb{A}$ is a $N \times N$ non-negative irreducible matrix, then

- its maximum eigenvalue λ_{max} is positive and simple
- the corresponding eigenvector $\mathbb{v}^{max}$ is positive
- $\mathbb{v}^{max}$ is the only non-negative eigenvector

Eigenvector Centrality

Perron-Frobenius theorem

The theorem states that if $\mathbb{A}$ is a $N \times N$ non-negative irreducible matrix, then

- its maximum eigenvalue λ_{max} is positive and simple
- the corresponding eigenvector $\mathbb{v}^{max}$ is positive
- $\mathbb{v}^{max}$ is the only non-negative eigenvector

Hence, **Eigenvector Centrality** of node i is the i-th element of $\mathbb{v}^{max}$, possibly normalized as

$$EC_i = \frac{\mathbb{v}_i^{max}}{\sum_j \mathbb{v}_j^{max}}$$

Directed networks

Introduction

How is WWW different from any other document collection?

- WWW is hypertext and provides information on top of the text (link structure, link text);
- Web pages are free of control or publishing costs;
- Huge number of pages can be created, artificially inflating citation counts;
- We need global "importance" ranking of every webpage to assess the *quality* of a page.

Introduction

How is WWW different from any other document collection?

- WWW is hypertext and provides information on top of the text (link structure, link text);
- Web pages are free of control or publishing costs;
- Huge number of pages can be created, artificially inflating citation counts;
- We need global "importance" ranking of every webpage to assess the *quality* of a page.

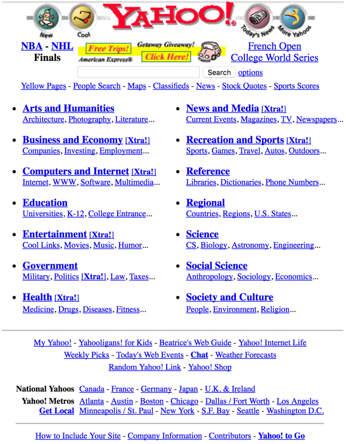
Introduction

How is WWW different from any other document collection?

- WWW is hypertext and provides information on top of the text (link structure, link text);
- Web pages are free of control or publishing costs;
- Huge number of pages can be created, artificially inflating citation counts;
- We need global "importance" ranking of every webpage to assess the *quality* of a page.

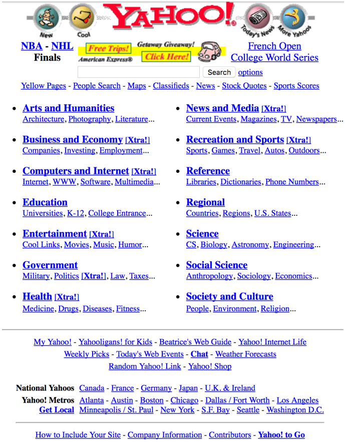
The dawn of ranking web pages

+ **Yet Another Hierarchical Officious Oracle** (developed by Jerry Yang and David Filo from Stanford)
+ When the number of web-pages was limited…
+ Hierarchical organisation with man-made taxonomy
+ Yahoo! *was there not to search the* web, but to explore the directory



The dawn of ranking web pages

- **Y**et **A**nother **H**ierarchical **O**fficious **O**racle (developed by Jerry Yang and David Filo from Stanford)
- When the number of web-pages was limited...
- Hierarchical organisation with man-made taxonomy
- Yahoo! *was there not to search the* web, but to explore the directory



When the number of pages is unbearable

[L. Page, S.Brin, R. Motwani and T. Winograd, *The PageRank Citation Ranking: Bringing Order to the Web* (1998)]

As of 1998: *The crawlable Web has roughly 150 million nodes (pages) and 1.7 billion edges (links).*

Main Idea

Use topology of the network to assess page importance, do not question the page content.

Topology of the network

- The web is the result of a collective action
- Every page has some number of forward links (out- edges) and backlinks (in- edges)
- *A link is like a vote for a page*

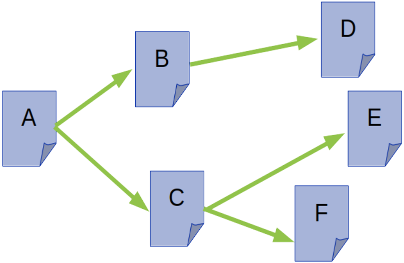


Figure: A is backlink of B and C

Topology of the network

- The web is the result of a collective action
- Every page has some number of forward links (out- edges) and backlinks (in- edges)
- *A link is like a vote for a page*

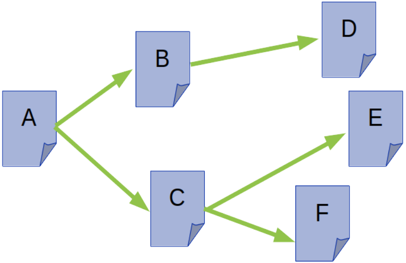
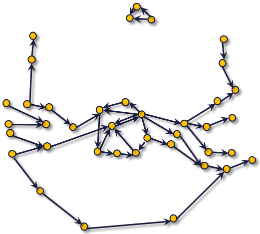


Figure: A is backlink of B and C

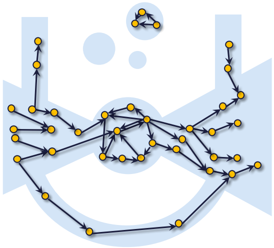
The bow tie structure

The silhouette of the WWW



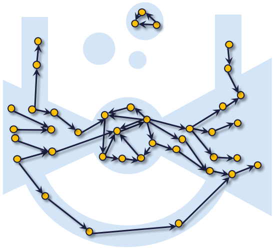
- Nodes: web pages (or sites / domains) // Edges: hyperlinks
- edge origin: **linking page**, edge end: **linked page**
- It has a *bow-tie* structure
- Heterogeneous cluster sizes
- Size of the components non-trivial

The silhouette of the WWW



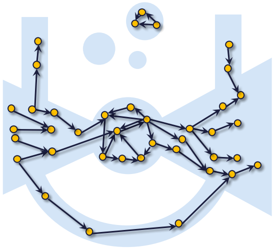
- Nodes: web pages (or sites / domains) // Edges: hyperlinks
- edge origin: **linking page**, edge end: **linked page**
- It has a *bow-tie* structure
- Heterogeneous cluster sizes
- Size of the components non-trivial

The silhouette of the WWW



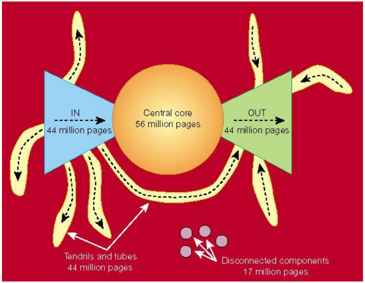
- Nodes: web pages (or sites / domains) // Edges: hyperlinks
- edge origin: **linking page**, edge end: **linked page**
- It has a *bow-tie* structure
- Heterogeneous cluster sizes
- Size of the components non-trivial

The silhouette of the WWW



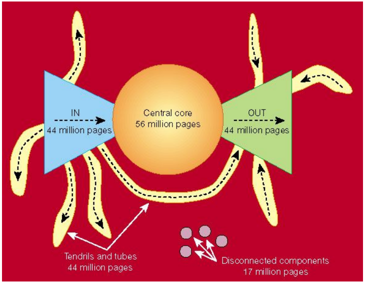
- Nodes: web pages (or sites / domains) // Edges: hyperlinks
- edge origin: **linking page**, edge end: **linked page**
- It has a *bow-tie* structure
- Heterogeneous cluster sizes
- Size of the components non-trivial

The bow-tie structure



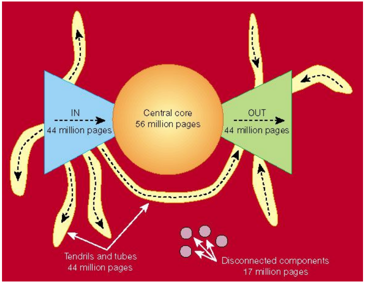
- **Central Core – SCC**
  - There is a path between every pair of pages

The bow-tie structure



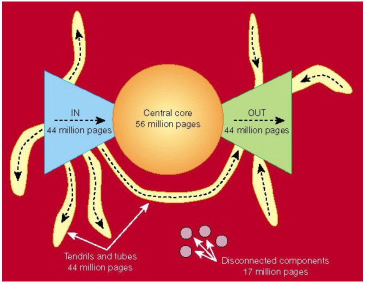
- **IN**
  - There is a path between these pages and the SCC
  - They cannot be reached from the SCC

The bow-tie structure



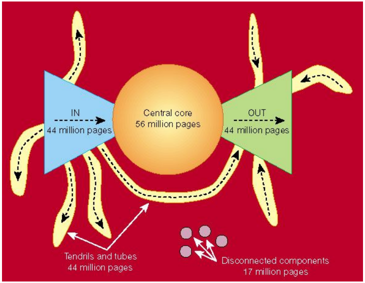
- **OUT**
  - There is a path between the SCC and these pages
  - The SCC cannot be reached from it

The bow-tie structure



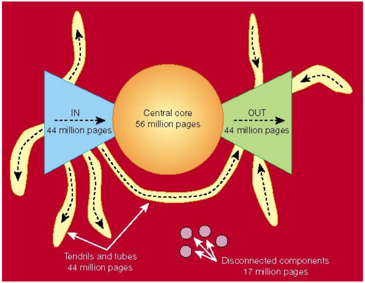
- **Tubes**
  - There is a path starting from the IN and finishing in the OUT that goes through the nodes in this region
  - The SCC cannot be reached from it
  - They cannot be reached from the SCC

The bow-tie structure



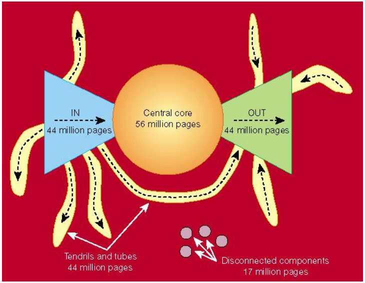
- **Tendrils IN**
  - There is a path starting from the IN and finishing in the Tendrils IN
  - The SCC cannot be reached from it

The bow-tie structure



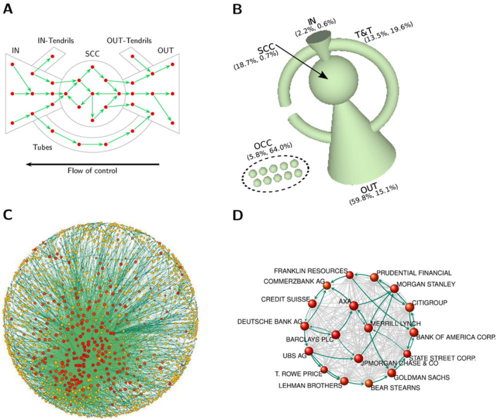
- **Tendrils OUT**
  - There is a path starting from the Tendrils OUT and finishing in OUT
  - They cannot be reached from the SCC

The bow-tie structure



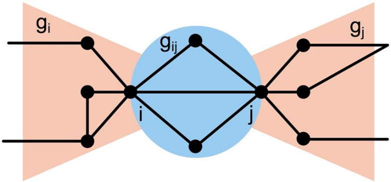
- **Disconnected components**
  - There is no path connecting these to/from other components

Where else? Corporate control



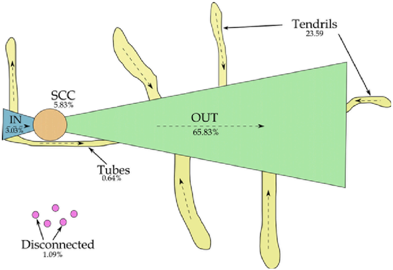
Nodes: public companies - Edges: share ownership

Where else? Social networks



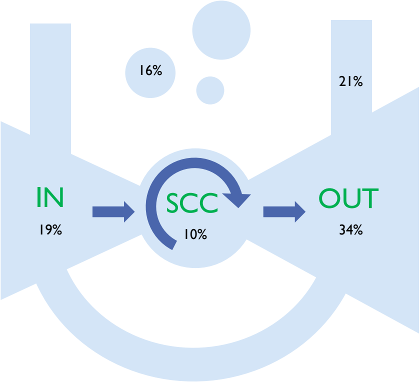
- Local bow-tie: Built from pairs of social interactions
- The largest the overlap in IN / OUT, the higher the friendship tie

Where else? Commercial relations



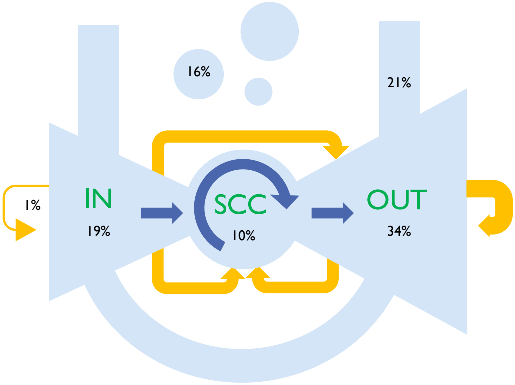
- Asymmetric bow-tie

Where else? Cryptocurrencies transaction



- The structure evolves slowly
- Very large WCC
- Relatively small SCC
- Small IN, very large OUT (hoard!), stable SCC

Where else? Cryptocurrencies transaction

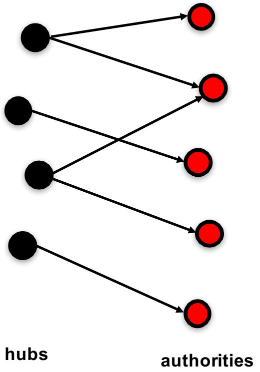


- The structure evolves slowly
- Very large WCC
- Relatively small SCC
- Small IN, very large OUT (hoard!), stable SCC

Idea of Hypertext Induced Topic Search

HITS: Hypertext Induced Topic Search

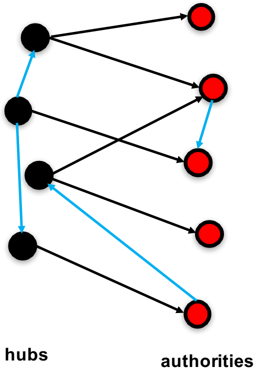
[Jon M. Kleinberg, *Authoritative Sources in a Hyperlinked Environment* (1997)]



Idea of Hypertext Induced Topic Search

HITS: Hypertext Induced Topic Search

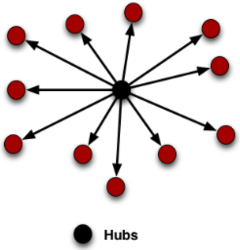
[Jon M. Kleinberg, *Authoritative Sources in a Hyperlinked Environment* (1997)]



Idea of Hypertext Induced Topic Search

HITS: Hypertext Induced Topic Search

[Jon M. Kleinberg, *Authoritative Sources in a Hyperlinked Environment* (1997)]

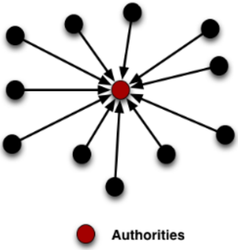


- Divide all pages into: **Hubs** and **Authorities**
- **Hubs**: no particular information; a series of addresses to find detailed description of various topics (e.g.: Yahoo! web page)
- **Authorities**: pages with few links but large information content on specific topics (e.g.: train timetable, series of paintings, recipes, etc.)

Idea of Hypertext Induced Topic Search

HITS: Hypertext Induced Topic Search

[Jon M. Kleinberg, *Authoritative Sources in a Hyperlinked Environment* (1997)]



- Divide all pages into: **Hubs** and **Authorities**
- **Hubs**: no particular information; a series of addresses to find detailed description of various topics (e.g.: Yahoo! web page)
- **Authorities**: pages with few links but large information content on specific topics (e.g.: train timetable, series of paintings, recipes, etc.)

Hubness and authoritativeness

Assumption of HITS

For every page we can assign the *authoritativeness* x_i and *hubness* y_i of every vertex i in a given graph

$$x_i = \sum_{j \to i} y_j \qquad y_i = \sum_{i \to j} x_j$$

$j \to i$ corresponds to all *ingoing* edges of i,
$i \to j$ corresponds to all *outgoing* edges of i

Matrix notation

Previous equations can be written in matrix notation:

$$\mathbf{x} = \mathbb{A}^T \cdot \mathbf{y}$$
$$\mathbf{y} = \mathbb{A} \cdot \mathbf{x}$$

where $\mathbb{A}$ is the **directed** adjacency matrix of the webpages graph

Further developing the equations:

$$\mathbf{x} = \mathbb{A}^T \mathbb{A} \cdot \mathbf{x}$$
$$\mathbf{y} = \mathbb{A}\,\mathbb{A}^T \cdot \mathbf{y}$$

This is an eigenvalue problem!

Matrix notation

Previous equations can be written in matrix notation:

$$\mathbf{x} = \mathbb{A}^T \cdot \mathbf{y}$$
$$\mathbf{y} = \mathbb{A} \cdot \mathbf{x}$$

where $\mathbb{A}$ is the **directed** adjacency matrix of the webpages graph

Further developing the equations:

$$\mathbf{x} = \mathbb{A}^T \mathbb{A} \cdot \mathbf{x}$$
$$\mathbf{y} = \mathbb{A}\,\mathbb{A}^T \cdot \mathbf{y}$$

This is an eigenvalue problem!

Matrix notation

Previous equations can be written in matrix notation:

$$\begin{aligned} \mathbf{x} &= \mathbb{A}^T \cdot \mathbf{y} \\ \mathbf{y} &= \mathbb{A} \cdot \mathbf{x} \end{aligned}$$

where $\mathbb{A}$ is the **directed** adjacency matrix of the webpages graph

Further developing the equations:

$$\begin{aligned} \mathbf{x} &= \mathbb{A}^T \mathbb{A} \cdot \mathbf{x} \\ \mathbf{y} &= \mathbb{A}\, \mathbb{A}^T \cdot \mathbf{y} \end{aligned}$$

This is an eigenvalue problem!

Solution

Ways of finding the solution: exact or recursive methods

Possible iterative algorithm:

$\forall i : x_i^{(0)}, y_i^{(0)} \leftarrow$ initial values

while $\delta > \epsilon$:

$\mathbf{x}^{(k+1)} \leftarrow \mathbb{A}^T \mathbb{A} \mathbf{x}^{(k)}$

$\mathbf{y}^{(k+1)} \leftarrow \mathbb{A} \mathbb{A}^T \mathbf{y}^{(k)}$

normalise

$$\mathbf{x}^{(k+1)} \leftarrow \mathbf{x}^{(k+1)} / \left\|\mathbf{x}^{(k+1)}\right\|$$

$$\mathbf{y}^{(k+1)} \leftarrow \mathbf{y}^{(k+1)} / \left\|\mathbf{y}^{(k+1)}\right\|$$

$$\delta \leftarrow max\left(\left\|\mathbf{x}^{(k)} - \mathbf{x}^{(k+1)}\right\|, \left\|\mathbf{y}^{(k)} - \mathbf{y}^{(k+1)}\right\|\right)$$

Solution

Ways of finding the solution: exact or recursive methods

Possible iterative algorithm:

$\forall i : x_i^{(0)}, y_i^{(0)} \leftarrow$ initial values

while $\delta > \epsilon$:

$$\mathbf{x}^{(k+1)} \leftarrow \mathbb{A}^T \mathbb{A} \mathbf{x}^{(k)}$$

$$\mathbf{y}^{(k+1)} \leftarrow \mathbb{A}\, \mathbb{A}^T \mathbf{y}^{(k)}$$

normalise

$$\mathbf{x}^{(k+1)} \leftarrow \mathbf{x}^{(k+1)} / \left\|\mathbf{x}^{(k+1)}\right\|$$

$$\mathbf{y}^{(k+1)} \leftarrow \mathbf{y}^{(k+1)} / \left\|\mathbf{y}^{(k+1)}\right\|$$

$$\delta \leftarrow max\left(\left\|\mathbf{x}^{(k)} - \mathbf{x}^{(k+1)}\right\|, \left\|\mathbf{y}^{(k)} - \mathbf{y}^{(k+1)}\right\|\right)$$

Mathematical properties

- Matrices $\mathbb{A}^T \mathbb{A}$, $\mathbb{A} \mathbb{A}^T$ are symmetric, PSD (positive semi-definite)
- Therefore, their eigenvalues are real and non-negative; they can be sorted as $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n \geq 0$
- We are interested in the **principal eigenvector**, corresponding to the largest eigenvalue
- Speed of convergence depends on how much $\lambda_1 > \lambda_2$
- **One requirement for convergence: normalisation of the vectors**

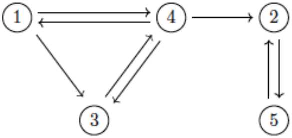
Mathematical properties

- Matrices $\mathbb{A}^T \mathbb{A}$, $\mathbb{A}\,\mathbb{A}^T$ are symmetric, PSD (positive semi-definite)
- Therefore, their eigenvalues are real and non-negative; they can be sorted as $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_n \geq 0$
- We are interested in the **principal eigenvector**, corresponding to the largest eigenvalue
- Speed of convergence depends on how much $\lambda_1 > \lambda_2$
- One requirement for convergence: normalisation of the vectors

Mathematical properties

- Matrices $\mathbb{A}^T \mathbb{A}$, $\mathbb{A}\ \mathbb{A}^T$ are symmetric, PSD (positive semi-definite)
- Therefore, their eigenvalues are real and non-negative; they can be sorted as $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n \geq 0$
- We are interested in the **principal eigenvector**, corresponding to the largest eigenvalue
- Speed of convergence depends on how much $\lambda_1 > \lambda_2$
- **One requirement for convergence: normalisation of the vectors**

Mathematical properties

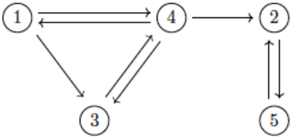


+ **Another requirement for convergence: irreducibility of the webpage graph**

Irreducibility

Matrix $\mathbb{A}$ is irreducible $\iff$ the corresponding graph is strongly connected, i.e. *it is possible to reach every vertex from every other*

Mathematical properties



- **Another requirement for convergence: irreducibility of the webpage graph**

Irreducibility

Matrix $\mathbb{A}$ is irreducible $\iff$ the corresponding graph is strongly connected, i.e. *it is possible to reach every vertex from every other*

Problems of the HITS Algorithm

- Hard to solve the eigenvalue problem for large networks
- **The web is not an irreducible graph**, *the calculation does not converge*

Problems of the HITS Algorithm

+ Hard to solve the eigenvalue problem for large networks
+ **The web is not an irreducible graph**, *the calculation does not converge*

PageRank

PageRank: first approximation

PageRank of a page is weighted sum over PageRanks of its backlinks:

$$pr(i) = \sum_{j \to i} \frac{pr(j)}{k_j^O}$$

- Any webmaster of a page can vote by pointing to another page
- In this way the page authority is passed to pointed pages
- The rank of a page is divided among its forward links evenly

PageRank: first approximation

PageRank of a page is weighted sum over PageRanks of its backlinks:

$$pr(i) = \sum_{j \to i} \frac{pr(j)}{k_j^O}$$

- The more pages the webmaster links, the lower is the voting power of each link
- The equation is recursive but it may be computed by starting with any set of ranks and iterating the computation

A webpage has high rank if the sum of the ranks of the pages that link to it is high

Matrix notation

PageRank Spectrum

The PageRank equation is an eigenvalue problem

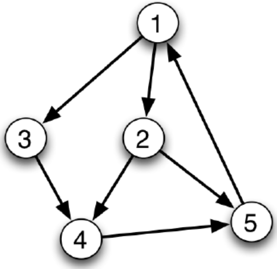
The equation can be rewritten as:

$$pr = \mathbb{N} \cdot pr$$

where $\mathbb{N}$ is a square matrix with the entries:

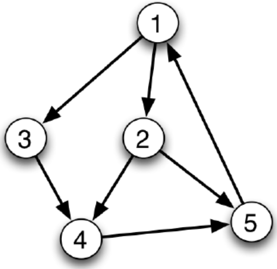
$$N_{ij} = \begin{cases} \frac{1}{k_i^O} & \text{, if there is an edge from } i \text{ to } j \\ 0 & \text{otherwise} \end{cases}$$

Simple example



$$\begin{cases} pr(1) = pr(5) \\ pr(2) = \frac{1}{2}pr(1) \\ pr(3) = \frac{1}{2}pr(1) \\ pr(4) = pr(3) + \frac{1}{2}pr(2) \\ pr(5) = pr(4) + \frac{1}{2}pr(2) \end{cases}$$

Simple example: matrix notation



$$\begin{pmatrix} pr(1) \\ pr(2) \\ pr(3) \\ pr(4) \\ pr(5) \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ \frac{1}{2} & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} pr(1) \\ pr(2) \\ pr(3) \\ pr(4) \\ pr(5) \end{pmatrix}$$

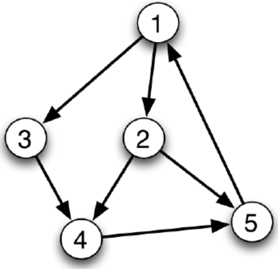
Three matrices

The same graph is associated to three different matrices:

$\mathbb{A}$	$\mathbb{N}$	$\mathbb{P}$
$\begin{pmatrix} 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ \frac{1}{2} & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 1 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}$

- $\mathbb{A}$ - adjacency matrix,
- $\mathbb{N} = \left(\frac{1}{k_i}\mathbb{A}\right)^T$
- $\mathbb{P} = \mathbb{N}^T$, Markov process

Random surfer interpretation



$$\begin{pmatrix} pr(1) \\ pr(2) \\ pr(3) \\ pr(4) \\ pr(5) \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ \frac{1}{2} & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} pr(1) \\ pr(2) \\ pr(3) \\ pr(4) \\ pr(5) \end{pmatrix}$$

The process is equivalent to a "random surfer" who simply keeps clicking on successive links at random

Random surfer interpretation

- Suppose she starts on vertex 2. Then initial state is $X_0 = [0, 1, 0, 0, 0]$;
- The surfer moves along the edges of the graph randomly;
- Markov process: $X_{t+1} = X_t \mathbb{P}$
- GOAL: find a **steady state** $X^* = X^* \mathbb{P}$

Looks familiar?

Random surfer interpretation

- Suppose she starts on vertex 2. Then initial state is $X_0 = [0, 1, 0, 0, 0]$;
- The surfer moves along the edges of the graph randomly;
- Markov process: $X_{t+1} = X_t \mathbb{P}$
- GOAL: find a **steady state** $X^* = X^* \mathbb{P}$

Looks familiar?

*Of course! It's **Eigenvector Centrality** again!*

Perron-Frobenius, again

Perron-Frobenius theorem

Every primitive and irreducible stochastic matrix $\mathbb{P}$ has a steady state, and the largest eigenvalue is always 1.

$\mathbb{P}$ must be:

+ non-negative: $\mathbb{P}_{ij} \geq 0, \forall i, j$
+ primitive: $\exists m : \mathbb{P}^m$ is positive definite
+ irreducible: $Pr(X_{n_{ij}} = j | X_0 = i) = p_{ij}^{n_{ij}} > 0$

For the WWW

In this prescription, the "random surfer" **does not satisfy the irreducibility condition**

Perron-Frobenius, again

Perron-Frobenius theorem

Every primitive and irreducible stochastic matrix $\mathbb{P}$ has a steady state, and the largest eigenvalue is always 1.

$\mathbb{P}$ must be:

- non-negative: $\mathbb{P}_{ij} \geq 0, \forall i, j$
- primitive: $\exists m : \mathbb{P}^m$ is positive definite
- irreducible: $Pr(X_{n_{ij}} = j | X_0 = i) = p_{ij}^{n_{ij}} > 0$

For the WWW

In this prescription, the “random surfer” **does not satisfy the irreducibility condition**

Perron-Frobenius, again

Perron-Frobenius theorem

Every primitive and irreducible stochastic matrix $\mathbb{P}$ has a steady state, and the largest eigenvalue is always 1.

$\mathbb{P}$ must be:

- non-negative: $\mathbb{P}_{ij} \geq 0, \forall i, j$
- primitive: $\exists m : \mathbb{P}^m$ is positive definite
- irreducible: $Pr(X_{n_{ij}} = j | X_0 = i) = p_{ij}^{n_{ij}} > 0$

For the WWW

In this prescription, the "random surfer" **does not satisfy the irreducibility condition**

Perron-Frobenius, again

Perron-Frobenius theorem

Every primitive and irreducible stochastic matrix $\mathbb{P}$ has a steady state, and the largest eigenvalue is always 1.

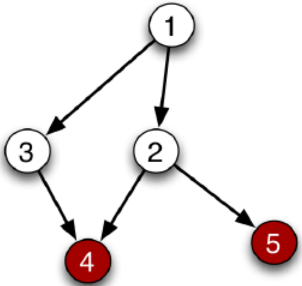
$\mathbb{P}$ must be:

- non-negative: $\mathbb{P}_{ij} \geq 0, \forall i, j$
- primitive: $\exists m : \mathbb{P}^m$ is positive definite
- irreducible: $Pr(X_{n_{ij}} = j | X_0 = i) = p_{ij}^{n_{ij}} > 0$

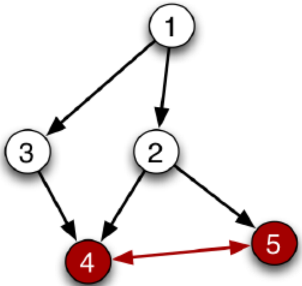
For the WWW

In this prescription, the “random surfer” **does not satisfy the irreducibility condition**

Issues with irreducibility



Dangling ends



Trapping states

The solution to the problem

Teleportation rule

To force the matrix to be irreducible we allow to jump out from a page to any other chosen at random, unformily

Random surfer interpretation: the surfer periodically "gets bored" and jumps to a random page chosen based on a certain distribution (here - uniform over all possible pages).

The solution to the problem

Teleportation rule

To force the matrix to be irreducible we allow to jump out from a page to any other chosen at random, unformily

Random surfer interpretation: the surfer periodically "gets bored" and jumps to a random page chosen based on a certain distribution (here - uniform over all possible pages).

Final definition

PageRank

$$pr(i) = \alpha \sum_{j \to i} \frac{pr(j)}{k_j^O} + (1 - \alpha)\frac{1}{N}$$

where N - number of vertices.

Typically $\alpha \simeq 0.85$.

Matrix notation

Define matrix $\mathbb{E} : e_{ij} = \frac{1}{N}$. Then:

$$pr = \left[\alpha \left((K^O)^{-1} \mathbb{A} \right)^T + (1-\alpha)\mathbb{E}\right] pr$$

that is,

$$\mathbb{N} = \alpha \left((K^O)^{-1} \mathbb{A} \right)^T + (1-\alpha)\mathbb{E}$$

Iterative solution

PageRank Spectrum

The PageRank equation is an eigenvalue problem and can be solved exactly or iteratively.

Iterative algorithm:

$$pr_0 \leftarrow S$$

while $\delta > \epsilon$:

$$\begin{aligned} pr_{i+1} &\leftarrow N \cdot pr_i \\ d &\leftarrow \|pr_i\|_1 - \|pr_{i+1}\|_1 \\ pr_{i+1} &\leftarrow pr_{i+1} + dE \\ \delta &\leftarrow \|pr_i - pr_{i+1}\|_1 \end{aligned}$$

Convergence of the algorithm

[L. Page, S.Brin, R. Motwani and T. Winograd, *The PageRank Citation Ranking: Bringing Order to the Web* (1998)]

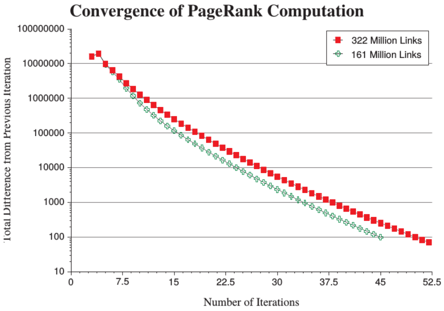
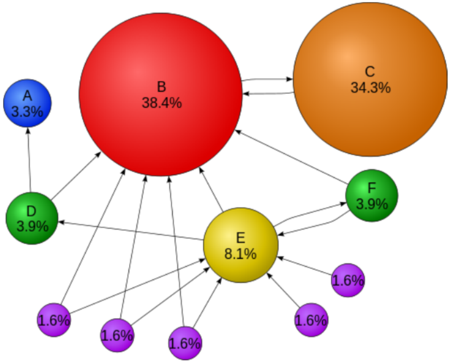


Figure: Rates of Convergence for Full size and Half Size Link Databases

Visualisation



PageRank top ’96

Web Page	PageRank (average is 1.0)
Download Netscape Software	11589.00
http://www.w3.org/	10717.70
Welcome to Netscape	8673.51
Point: It's What You're Searching For	7930.92
Web-Counter Home Page	7254.97
The Blue Ribbon Campaign for Online Free Speech	7010.39
CERN Welcome	6562.49
Yahoo!	6561.80
Welcome to Netscape	6203.47
Wusage 4.1: A Usage Statistics System For Web Servers	5963.27
The World Wide Web Consortium (W3C)	5672.21
Lycos, Inc. Home Page	4683.31
Starting Point	4501.98
Welcome to Magellan!	3866.82
Oracle Corporation	3587.63

Advantages of PageRank

- Can be used for personalised search by manipulating the $\mathbb{E}$ matrix
- *More immune to manipulations*, such as bought advertisements, fake websites, etc.
- Captures the concept of **collaborative trust**: **if a page was mentioned by a trustworthy or authoritative source, it is more likely to be trustworthy or authoritative**
- Uses information that is external to the Web-pages (their backlinks), providing kind of peer review
- A good way to find representative pages to display for a cluster centre

Advantages of PageRank

- Can be used for personalised search by manipulating the $\mathbb{E}$ matrix
- *More immune to manipulations*, such as bought advertisements, fake websites, etc.
- Captures the concept of **collaborative trust**:
  if a page was mentioned by a trustworthy or authoritative source, it is more likely to be trustworthy or authoritative
- Uses information that is external to the Web-pages (their backlinks), providing kind of peer review
- A good way to find representative pages to display for a cluster centre

Advantages of PageRank

+ Can be used for personalised search by manipulating the $\mathbb{E}$ matrix
+ *More immune to manipulations*, such as bought advertisements, fake websites, etc.
+ Captures the concept of **collaborative trust**:
 if a page was mentioned by a trustworthy or authoritative source, it is more likely to be trustworthy or authoritative
+ Uses information that is external to the Web-pages (their backlinks), providing kind of peer review
+ A good way to find representative pages to display for a cluster centre

Advantages of PageRank

+ Can be used for personalised search by manipulating the $\mathbb{E}$ matrix
+ *More immune to manipulations*, such as bought advertisements, fake websites, etc.
+ Captures the concept of **collaborative trust**:
  if a page was mentioned by a trustworthy or authoritative source, it is more likely to be trustworthy or authoritative
+ Uses information that is external to the Web-pages (their backlinks), providing kind of peer review
+ A good way to find representative pages to display for a cluster centre

Web applications of PR

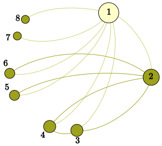
- PageRank is **one of the tools** used for Google searching engine (together with standard IR measures, proximity, anchor text);
- Great for underspecified queries;
- Estimation of web-traffic;
- Backlink prediction;
- User navigation;

Nestedness

What is network nestedness?

Definition of nestedness

A network is **nested** if for any pair of nodes i, j
If $|\mathcal{N}_i| \geq |\mathcal{N}_j|$ then $\mathcal{N}_i \supseteq \mathcal{N}_j$



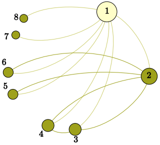
$$\mathbf{A} = \begin{pmatrix} 0&1&1&1&1&1&1&1 \\ 1&0&1&1&1&1&0&0 \\ 1&1&0&1&0&0&0&0 \\ 1&1&1&0&0&0&0&0 \\ 1&1&0&0&0&0&0&0 \\ 1&1&0&0&0&0&0&0 \\ 1&0&0&0&0&0&0&0 \\ 1&0&0&0&0&0&0&0 \end{pmatrix}$$

Nested network

What is network nestedness?

Definition of nestedness

A network is **nested** if for any pair of nodes i, j
If $k_i \geq k_j$ then $\mathcal{N}_i \supseteq \mathcal{N}_j$



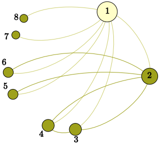
$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Nested network

What is network nestedness?

Definition of nestedness

A network is **nested** if for any pair of nodes i, j
If $k_i \geq k_j$ then $\mathcal{N}_i \supseteq \mathcal{N}_j$



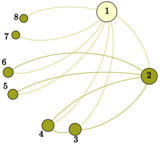
$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Nested network

What is network nestedness?

Definition of nestedness

A network is **nested** if for any pair of nodes i, j
If $k_i \geq k_j$ then $\mathcal{N}_i \supseteq \mathcal{N}_j$



$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

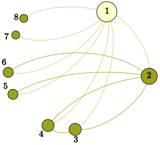
Nested network

- Core-periphery structure, dissortative, *step-wise* adjacency matrix

What is network nestedness?

Definition of nestedness

A network is **nested** if for any pair of nodes i, j
If $k_i \geq k_j$ then $\mathcal{N}_i \supseteq \mathcal{N}_j$



$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Nested network

What is network nestedness?

Definition of nestedness

A network is **nested** if for any pair of nodes i, j
If $k_i \geq k_j$ then $\mathcal{N}_i \supseteq \mathcal{N}_j$

decreasing nestedness $\longrightarrow$

$$\begin{pmatrix} 0&1&1&1&1&1&1&1\\ 1&0&1&1&1&1&0&0\\ 1&1&0&1&0&0&0&0\\ 1&1&1&0&0&0&0&0\\ 1&1&0&0&0&0&0&0\\ 1&1&0&0&0&0&0&0\\ 1&0&0&0&0&0&0&0\\ 1&0&0&0&0&0&0&0 \end{pmatrix} \quad \begin{pmatrix} 0&1&1&1&0&1&1&1\\ 1&0&1&1&1&1&0&0\\ 1&1&0&1&0&0&0&0\\ 1&1&1&0&0&0&1&0\\ 0&1&0&0&0&0&0&0\\ 1&1&0&0&0&0&0&0\\ 1&0&0&1&0&0&0&0\\ 1&0&0&0&0&0&0&0 \end{pmatrix} \quad \begin{pmatrix} 0&0&1&1&0&1&1&1\\ 0&0&0&0&1&1&0&0\\ 1&0&0&1&0&0&0&0\\ 1&0&1&0&1&0&1&0\\ 0&1&0&1&0&0&0&0\\ 1&1&0&0&0&0&0&1\\ 1&0&0&1&0&0&0&0\\ 1&0&0&0&0&1&0&0 \end{pmatrix}$$

+ Decreasing nestedness is given by increasing number of violations

What is network nestedness?

Definition of nestedness

A network is **nested** if for any pair of nodes i, j
If $k_i \geq k_j$ then $\mathcal{N}_i \supseteq \mathcal{N}_j$

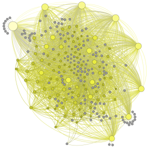
decreasing nestedness $\longrightarrow$

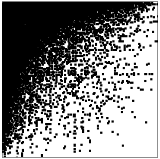
$$\begin{pmatrix} 0&1&1&1&1&1&1&1\\ 1&0&1&1&1&1&0&0\\ 1&1&0&1&0&0&0&0\\ 1&1&1&0&0&0&0&0\\ 1&1&0&0&0&0&0&0\\ 1&1&0&0&0&0&0&0\\ 1&0&0&0&0&0&0&0\\ 1&0&0&0&0&0&0&0 \end{pmatrix} \quad \begin{pmatrix} 0&1&1&1&0&1&1&1\\ 1&0&1&1&1&1&0&0\\ 1&1&0&1&0&0&0&0\\ 1&1&1&0&0&0&1&0\\ 0&1&0&0&0&0&0&0\\ 1&1&0&0&0&0&0&0\\ 1&0&0&1&0&0&0&0\\ 1&0&0&0&0&0&0&0 \end{pmatrix} \quad \begin{pmatrix} 0&0&1&1&0&1&1&1\\ 0&0&0&0&1&1&0&0\\ 1&0&0&1&0&0&0&0\\ 1&0&1&0&1&0&1&0\\ 0&1&0&1&0&0&0&0\\ 1&1&0&0&0&0&0&1\\ 1&0&0&1&0&0&0&0\\ 1&0&0&0&0&1&0&0 \end{pmatrix}$$

- Decreasing nestedness is given by increasing number of violations

Socio-economic systems

World-trade network



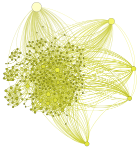


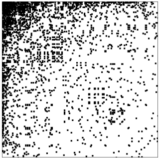
import-export trade relations (2000)

$n = 196, m = 4138$

$C = 0.73, \ell = 2.25, \gamma = -0.40$

R&D alliance network



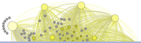


SIC: 737 *Computer Software* (1994-1997)

$n = 325, m = 2632$

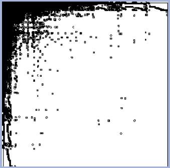
Socio-economic systems

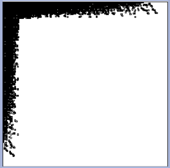
World-trade network



import-export trade relations (2000)

Interbank networks

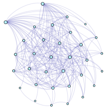
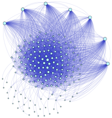
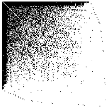




Bank for international settlements (Fedwire) and Austrian interbank network

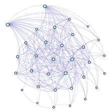
7)

Organisation communication networks

<table>
<tr><td colspan="3">International consulting company (Zürich branch)</td></tr>
<tr><td></td><td></td><td>inter-employee communication (30d, 2013)
$n = 33, m = 229$
$C = 0.74, \ell = 1.57, \gamma = -0.39$</td></tr>
<tr><td colspan="3">Mid-size manufacturing company</td></tr>
<tr><td></td><td></td><td>e-mail communication (2010)
$n = 67, m = 3251$
$C = 0.59, \ell = 1.96, \gamma = -0.29$</td></tr>
</table>

Organisation communication networks

International consulting company (Zürich branch)

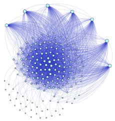
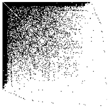



inter-employee communication (30d, 2013)

$n = 33, m = 229$

$C = 0.74, \ell = 1.57, \gamma = -0.39$

Mid-size manufacturing company

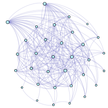
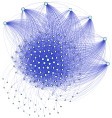
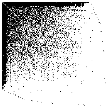



e-mail communication (2010)

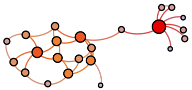
$n = 67, m = 3251$

$C = 0.59, \ell = 1.96, \gamma = -0.29$

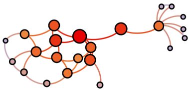
Organisation communication networks

International consulting company (Zürich branch)		
		inter-employee communication (30d, 2013) $n = 33, m = 229$ $C = 0.74, \ell = 1.57, \gamma = -0.39$
Mid-size manufacturing company		
		e-mail communication (2010) $n = 67, m = 3251$ $C = 0.59, \ell = 1.96, \gamma = -0.29$

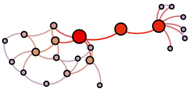
Centrality: one term, several concepts



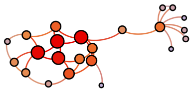
degree centrality



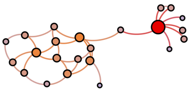
closeness centrality



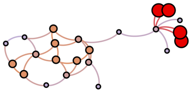
betweeneness centrality



eigenvector centrality



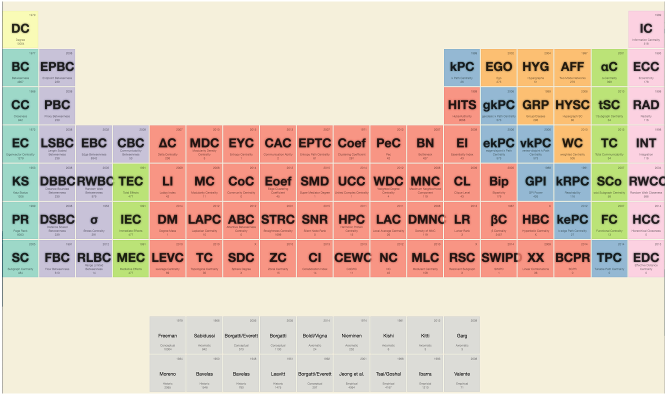
pagerank



clustering coefficient

+ Different measures for centrality, depending on the context

But which one to use?

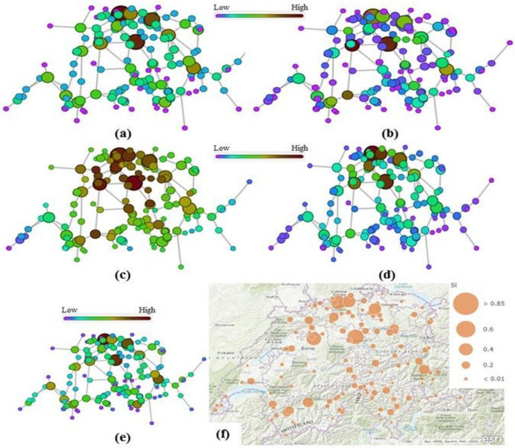


http://schochastics.net/sna/periodic.html

The importance of nodes (and therefore, the centrality measure to be used) depends on the context of the application

In nested networks, all centrality measures are ranked in the same way

Quick recap: centrality metrics



(a) Degree
(b) Betweenness
(c) Closeness
(d) Eigenvector
(e) PageRank
(f) Combined

Abedi, A. & Romerio, F., Systemic Vulnerability of Swiss Power Grid to Natural Events, *MATEC Web of Conferences*, 2019

Carlo Campajola

Blockchain & Distributed Ledger Technologies Group

campajola@ifi.uzh.ch
https://www.ifi.uzh.ch/en/bdlt.html