A CRITICAL, HISTORICAL AND MATHEMATICAL REVIEW OF CENTRALITY SCORES

Sebastiano Vigna
Università degli Studi di Milano

Work in progress with Paolo Boldi
supported by the EU-FET grant NADINE (GA 288956).

What do these people have in common?

What do these people have in common?

What do these people have in common?

What do these people have in common?

What do these people have in common?

PageRank (believe it or not)

PageRank (believe it or not)

* These are the top-8 actors of the Hollywood graph according to PageRank

PageRank (believe it or not)

* These are the top-8 actors of the Hollywood graph according to PageRank

*

Who is going to tell

that is better?

PageRank (believe it or not)

* These are the top-8 actors of the Hollywood graph according to PageRank

*

Who is going to tell

that

is better?

PageRank (believe it or not)

* These are the top-8 actors of the Hollywood graph according to PageRank

*

Who is going to tell

that

is better?

PageRank sucks!

(or NOT?)

PageRank sucks!

(or NOT?)

* The Hollywood graph we used contains 2,000,000 nodes. Most of them are *completely unknown*!

PageRank sucks!

(or NOT?)

- The Hollywood graph we used contains 2,000,000 nodes. Most of them are *completely unknown*!
- PageRank is not singling out the best actors...

PageRank sucks!

(or NOT?)

- The Hollywood graph we used contains 2,000,000 nodes. Most of them are *completely unknown*!
- PageRank is not singling out the best actors...
- ...but still it is not pointing to random individuals, is it?

The grand plan

The grand plan

- *Centrality indices / link analysis* has been with us for 60+ years

The grand plan

- *Centrality indices / link analysis* has been with us for 60+ years
- Sociology, psychology, bibliometrics, information retrieval all have given their contributions...

The grand plan

- *Centrality indices / link analysis* has been with us for 60+ years
- Sociology, psychology, bibliometrics, information retrieval all have given their contributions...
- ...turning it into a *jungle*

The grand plan

- *Centrality indices / link analysis* has been with us for 60+ years
- Sociology, psychology, bibliometrics, information retrieval all have given their contributions...
- ...turning it into a *jungle*
- Use some machetes to sort it out somehow

Centrality in social sciences

a historical account

Centrality in social sciences

a historical account

* First works by Bavelas at MIT (1948)

Centrality in social sciences

a historical account

- First works by Bavelas at MIT (1948)
- This sparked countless works (Bavelas 1951; Katz 1953; Shaw 1954; Beauchamp 1965; Mackenzie 1966; Burgess 1969; Anthonisse 1971; Czapiel 1974...) that Freeman (1979) tried to summarize concluding that:

Centrality in social sciences

a historical account

- First works by Bavelas at MIT (1948)
- This sparked countless works (Bavelas 1951; Katz 1953; Shaw 1954; Beauchamp 1965; Mackenzie 1966; Burgess 1969; Anthonisse 1971; Czapiel 1974...) that Freeman (1979) tried to summarize concluding that:

several measures are often only vaguely related to the intuitive ideas they purport to index, and many are so complex that it is difficult or impossible to discover what, if anything, they are measuring

The 1990s revival

Link Analysis Ranking

The 1990s revival

Link Analysis Ranking

- With the advent of search engines, there was a strong revamp of centrality (LAR in this context)

The 1990s revival

Link Analysis Ranking

- With the advent of search engines, there was a strong revamp of centrality (LAR in this context)
- New scenarios

The 1990s revival

Link Analysis Ranking

- With the advent of search engines, there was a strong revamp of centrality (LAR in this context)
- New scenarios
  - graphs are directed mainly

The 1990s revival

Link Analysis Ranking

- With the advent of search engines, there was a strong revamp of centrality (LAR in this context)
- New scenarios
  - graphs are directed mainly
  - they are huge

The 1990s revival

Link Analysis Ranking

- With the advent of search engines, there was a strong revamp of centrality (LAR in this context)
- New scenarios
  - graphs are directed mainly
  - they are huge
  - new attention to efficiency

What a mess!

What a mess!

- Only few, brave guys tried to make some order in this mess!

What a mess!

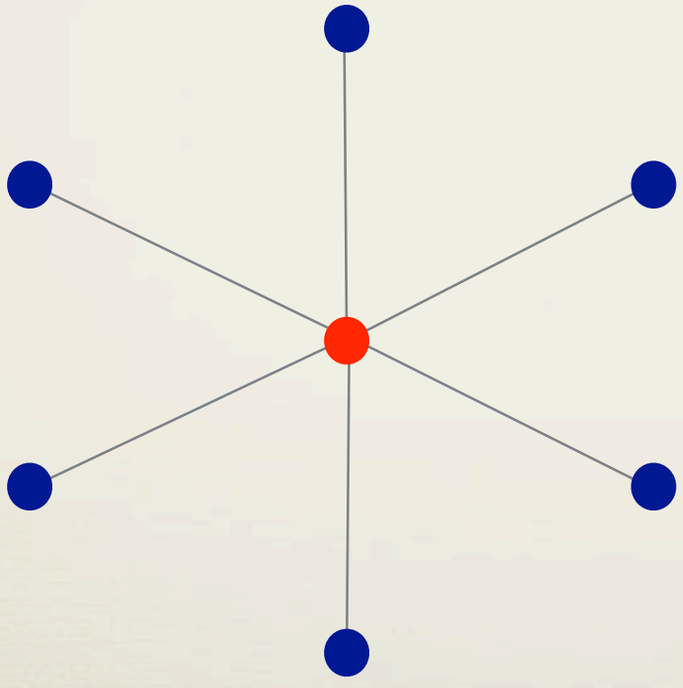
- Only few, brave guys tried to make some order in this mess!
- Noteworthy (in the IR context): Craswell, Upstill, Hawking (ADCS 2003); Najork, Zaragoza, Taylor (SIGIR 2007); Najork, Gollapudi, Panigrahy (WSDM 2009)

What a mess!

- Only few, brave guys tried to make some order in this mess!
- Noteworthy (in the IR context): Craswell, Upstill, Hawking (ADCS 2003); Najork, Zaragoza, Taylor (SIGIR 2007); Najork, Gollapudi, Panigrahy (WSDM 2009)
- Guess that the others were scared by computational burden of classical measures

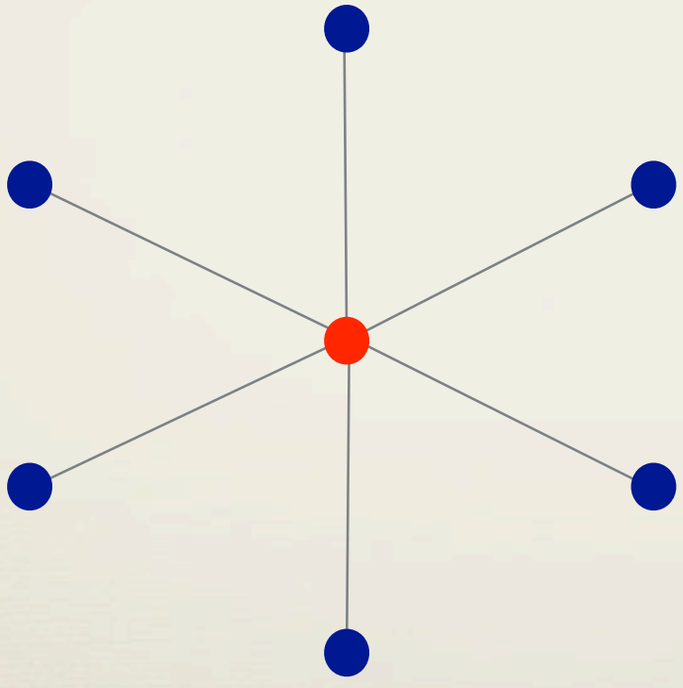
Begin the begin

Begin the begin



The only point on which everybody seems to agree: the center of a star is more important than the other nodes

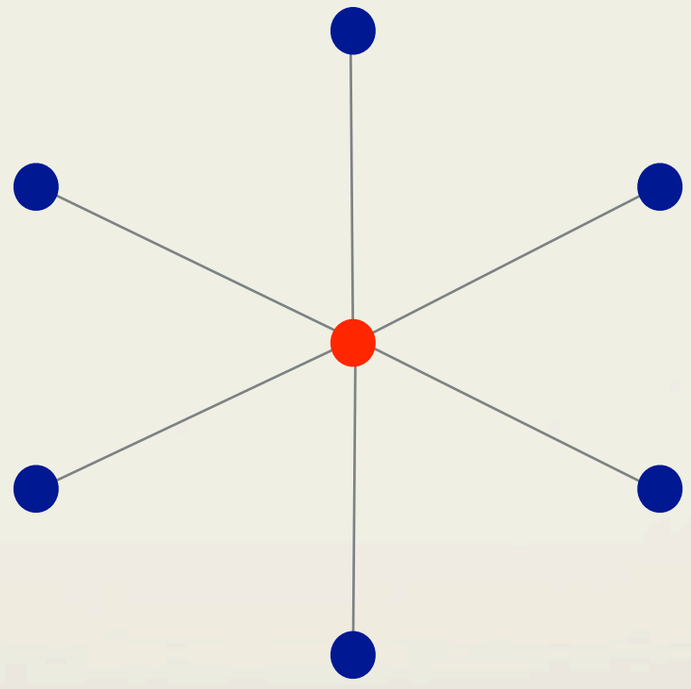
Begin the begin



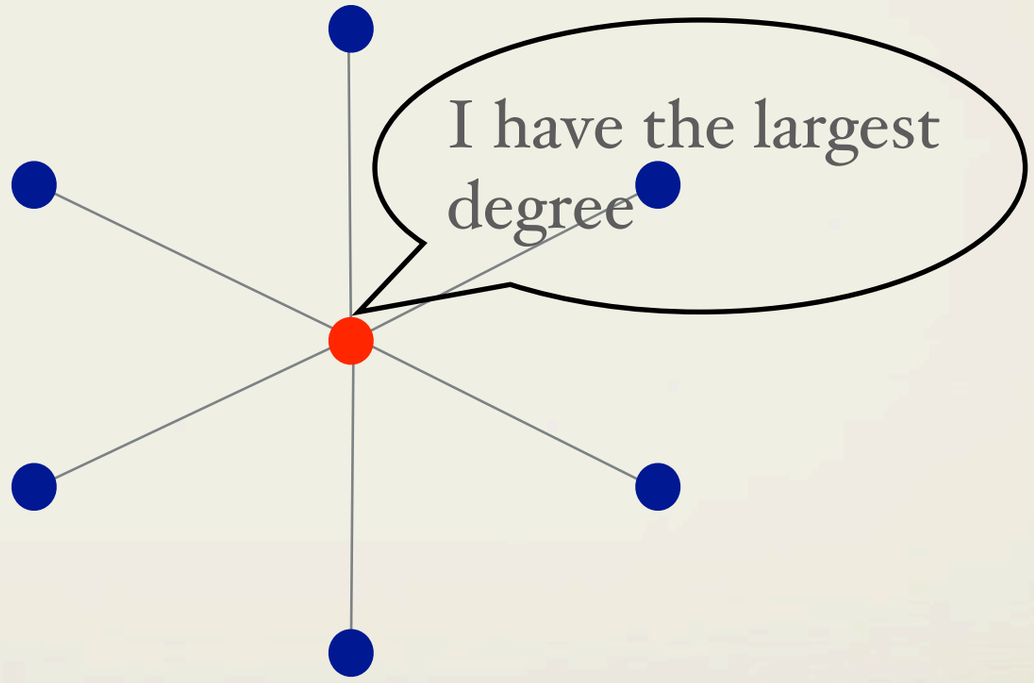
The only point on which everybody seems to agree: the center of a star is more important than the other nodes

But what does it make more important?

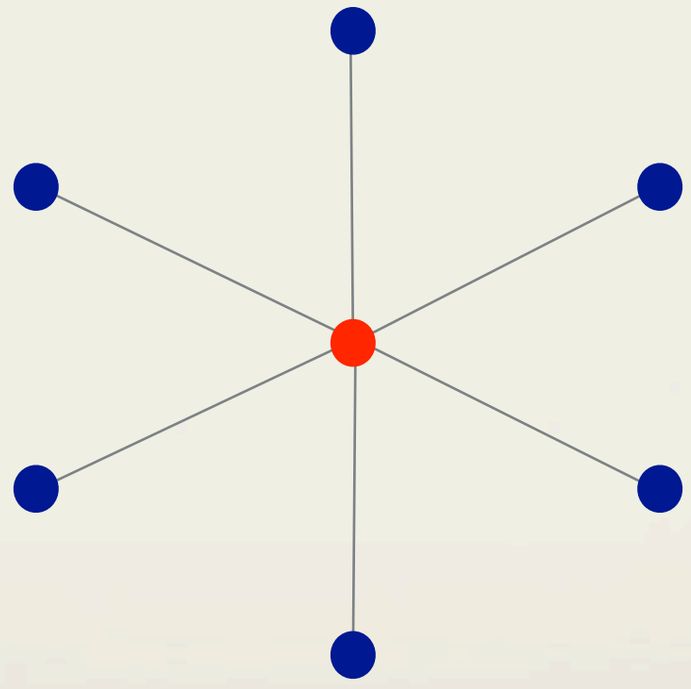
Begin the begin



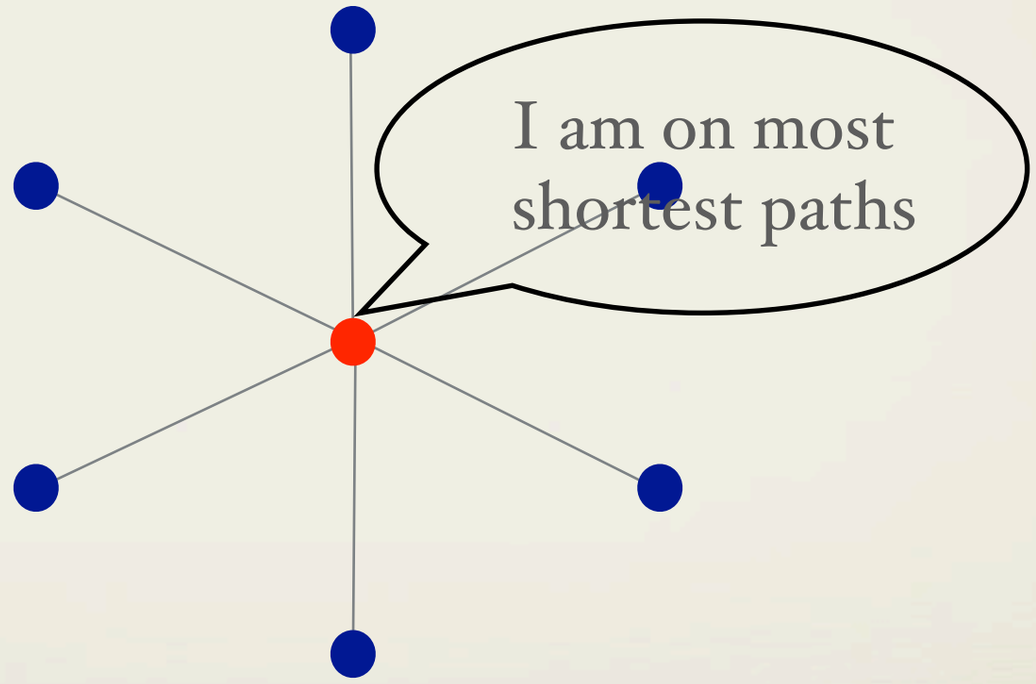
Begin the begin



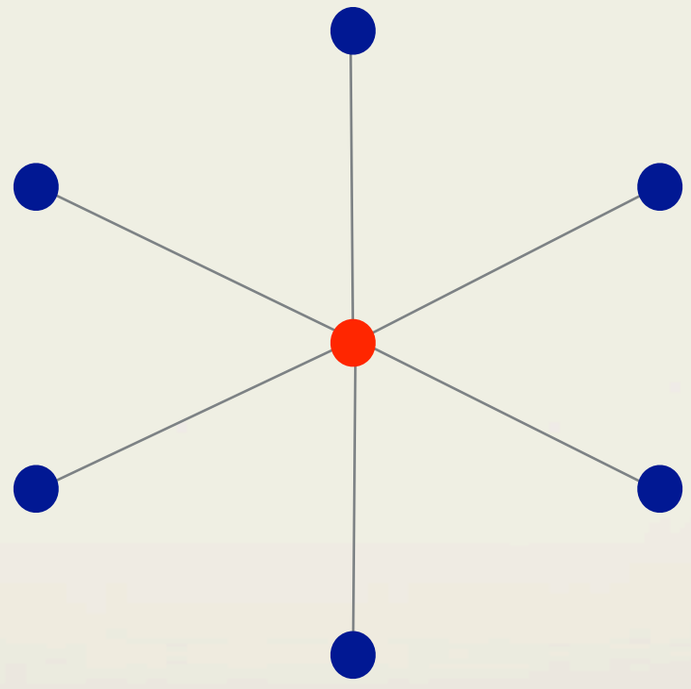
Begin the begin



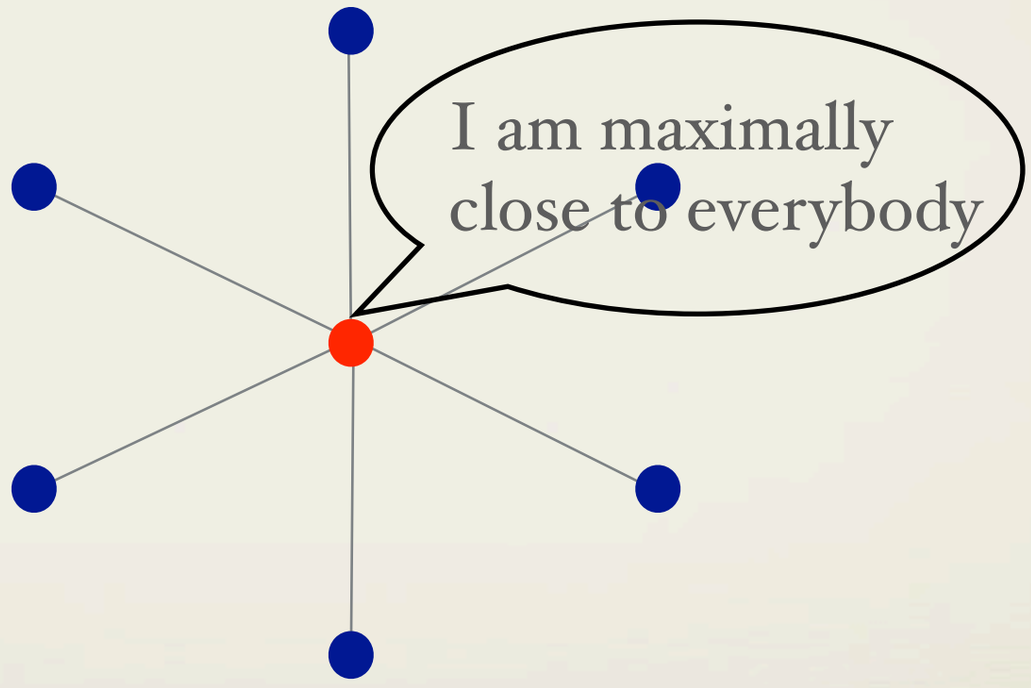
Begin the begin



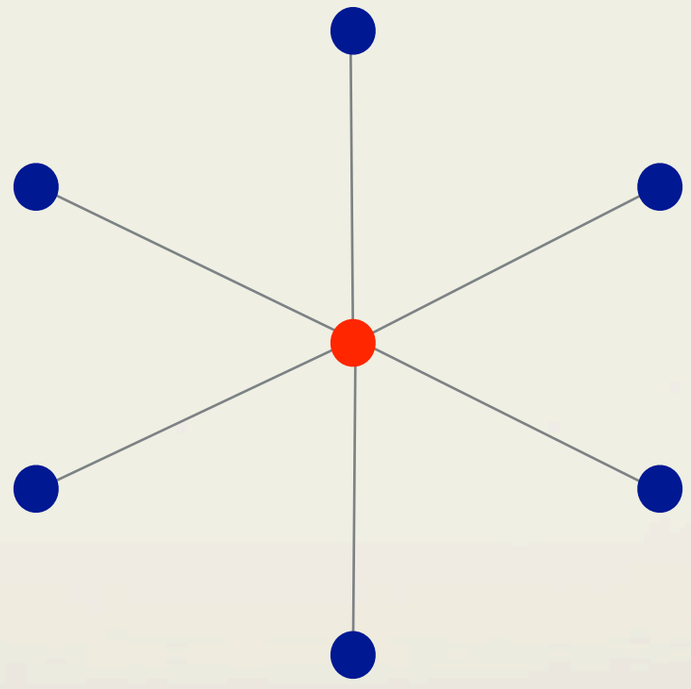
Begin the begin



Begin the begin



Begin the begin



A tale of three tribes

A tale of three tribes

* *Spectral indices*, based on some linear-algebra construction

A tale of three tribes

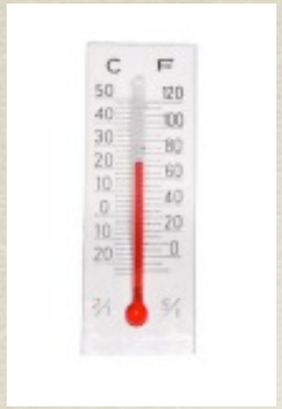
- *Spectral indices*, based on some linear-algebra construction
- *Path-based indices,* based on the number of *paths* or shortest paths (geodesics) passing through a vertex

A tale of three tribes

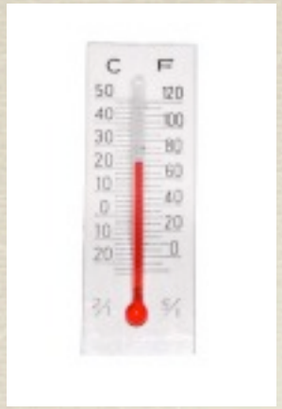
- *Spectral indices*, based on some linear-algebra construction
- *Path-based indices,* based on the number of *paths* or shortest paths (geodesics) passing through a vertex
- *Geometric indices,* based on *distances* from a vertex to other vertices

A tale of three tribes

- *Spectral indices*, based on some linear-algebra construction
- *Path-based indices,* based on the number of *paths* or shortest paths (geodesics) passing through a vertex
- *Geometric indices,* based on *distances* from a vertex to other vertices
- (Actually, the first two families are largely the same, even if that wasn't understood for a long time)



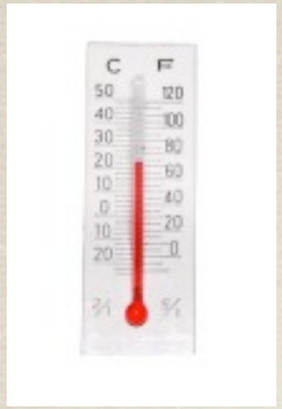
Degree



Degree

* (In-)**Degree centrality**: the number of incoming links

$$c_{\text{deg}}(x) = d^{-}(x)$$

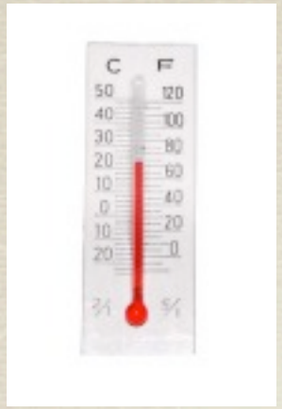


Degree

- (In-)**Degree centrality**: the number of incoming links

$$c_{\text{deg}}(x) = d^{-}(x)$$

- Or number of nodes at distance one



Degree

- (In-)**Degree centrality**: the number of incoming links

$$c_{\deg}(x) = d^{-}(x)$$

- Or number of nodes at distance one
- Careful: when dealing with *directed* networks, some indices present two variants (e.g., in-degree vs. out-degree), the ones based on incoming paths being usually more interesting

The path tribe

The path tribe

- **Betweenness** centrality (Anthonisse 1971):

$$c_{\mathrm{betw}}(x) = \sum_{y,z \neq x} \frac{\sigma_{yz}(x)}{\sigma_{yz}}$$

The path tribe

Fraction of shortest paths from y to z passing through x

- **Betweenness** centrality (Anthonisse 1971):

$$c_{\mathrm{betw}}(x) = \sum_{y,z \neq x} \frac{\sigma_{yz}(x)}{\sigma_{yz}}$$

The path tribe

- **Betweenness** centrality (Anthonisse 1971):

$$c_{\mathrm{betw}}(x) = \sum_{y,z \neq x} \frac{\sigma_{yz}(x)}{\sigma_{yz}}$$

The path tribe

- **Betweenness** centrality (Anthonisse 1971):

$$c_{\mathrm{betw}}(x) = \sum_{y,z \neq x} \frac{\sigma_{yz}(x)}{\sigma_{yz}}$$

- **Katz** centrality (Katz 1953):

The path tribe

- **Betweenness** centrality (Anthonisse 1971):

$$c_{\mathrm{betw}}(x) = \sum_{y,z \neq x} \frac{\sigma_{yz}(x)}{\sigma_{yz}}$$

- **Katz** centrality (Katz 1953):

$$c_{\mathrm{K}atz}(x) = \sum_{t=0}^{\infty} \alpha^t \Pi_x(t) = \mathbf{1} \sum_{t=0}^{\infty} \alpha^t G^t$$

The path tribe

- **Betweenness** centrality (Anthonisse 1971):

$$c_{\mathrm{betw}}(x) = \sum_{y,z \neq x} \frac{\sigma_{yz}(x)}{\sigma_{yz}}$$

- **Katz** centrality (Katz 1953):

$$c_{\mathrm{K}atz}(x) = \sum_{t=0}^{\infty} \alpha^t \Pi_x(t) = \mathbf{1} \sum_{t=0}^{\infty} \alpha^t G^t$$

# of paths of length t ending in x

The path tribe

- **Betweenness** centrality (Anthonisse 1971):

$$c_{\mathrm{betw}}(x) = \sum_{y,z \neq x} \frac{\sigma_{yz}(x)}{\sigma_{yz}}$$

- **Katz** centrality (Katz 1953):

$$c_{\mathrm{K}atz}(x) = \sum_{t=0}^{\infty} \alpha^t \Pi_x(t) = \mathbf{1} \sum_{t=0}^{\infty} \alpha^t G^t$$

The distance tribe

The distance tribe

- **Closeness** centrality (Bavelas 1950):

$$c_{\text{clos}}(x) = \frac{1}{\sum_y d(y,x)}$$

The distance tribe

- **Closeness** centrality (Bavelas 1950):

$$c_{\text{clos}}(x) = \frac{1}{\sum_y d(y,x)}$$

Distance from y to x

The distance tribe

- **Closeness** centrality (Bavelas 1950):

$$c_{\text{clos}}(x) = \frac{1}{\sum_y d(y,x)}$$

The distance tribe

- **Closeness** centrality (Bavelas 1950):

$$c_{\text{clos}}(x) = \frac{1}{\sum_y d(y,x)}$$

- **Lin** centrality (Lin 1976):

$$c_{\text{Lin}}(x) = \frac{\text{canReach}(x)^2}{\sum_y d(y,x)}$$

The distance tribe

- **Closeness** centrality (Bavelas 1950):

$$c_{\text{clos}}(x) = \frac{1}{\sum_y d(y,x)}$$

- **Lin** centrality (Lin 1976):

$$c_{\text{Lin}}(x) = \frac{\text{canReach}(x)^2}{\sum_y d(y,x)}$$

The distance tribe

- **Closeness** centrality (Bavelas 1950):

$$c_{\text{clos}}(x) = \frac{1}{\sum_y d(y,x)}$$

- **Lin** centrality (Lin 1976):

$$c_{\text{Lin}}(x) = \frac{\text{canReach}(x)^2}{\sum_y d(y,x)}$$

- The summation is over all y such that $d(y,x)<\infty$

The distance tribe

- **Closeness** centrality (Bavelas 1950):

$$c_{\text{clos}}(x) = \frac{1}{\sum_y d(y,x)}$$

- **Lin** centrality (Lin 1976):

$$c_{\text{Lin}}(x) = \frac{\text{canReach}(x)^2}{\sum_y d(y,x)}$$

Completely neglected by the literature

- The summation is over all y such that $d(y,x)<\infty$

The distance tribe

a new member

Monday, July 23, 2012

The distance tribe

a new member

- Give a warm welcome to **Harmonic centrality**:

$$c_{\text{harm}}(x) = \sum_{y \neq x} \frac{1}{d(y,x)}$$

The distance tribe

a new member

- Give a warm welcome to **Harmonic centrality**:

$$c_{\text{harm}}(x) = \sum_{y \neq x} \frac{1}{d(y,x)}$$

- The denormalized reciprocal of the *harmonic mean of all distances* (even ∞)

The distance tribe

a new member

- Give a warm welcome to **Harmonic centrality**:

$$c_{\text{harm}}(x) = \sum_{y \neq x} \frac{1}{d(y,x)}$$

- The denormalized reciprocal of the *harmonic mean of all distances* (even ∞)

- Inspired by the use the the harmonic mean in (Marchiori, Latora 2000)

The distance tribe

a new member

- Give a warm welcome to **Harmonic centrality**:

$$c_{\text{harm}}(x) = \sum_{y \neq x} \frac{1}{d(y,x)}$$

- The denormalized reciprocal of the *harmonic mean of all distances* (even ∞)
- Inspired by the use the the harmonic mean in (Marchiori, Latora 2000)
- Probably already appeared somewhere (e.g., quoted for undirected graphs in Tore Opsahl's blog)

The spectral tribe

The spectral tribe

- All based on the eigenstructure of some graph-related matrix

The spectral tribe

- All based on the eigenstructure of some graph-related matrix
- Most obvious: left or right dominant eigenvector of some matrix derived from the graph adjacency matrix

The spectral tribe

- All based on the eigenstructure of some graph-related matrix
- Most obvious: left or right dominant eigenvector of some matrix derived from the graph adjacency matrix
- All share the same issues of unicity and computability, mainly solved using Perron-Frobenius theory and the power method or more sophisticated approaches.

Seeley index
(Seeley 1949)

Seeley index
(Seeley 1949)

- Basic idea: in a group of children, a child is as popular as the sum of the popularities of the children who like him, but popularities are divided evenly among friends:

Seeley index
(Seeley 1949)

- Basic idea: in a group of children, a child is as popular as the sum of the popularities of the children who like him, but popularities are divided evenly among friends:

$$c_{Seeley}(x) = \sum_{y \to x} \frac{c_{Seeley}(y)}{d^+(y)}$$

Seeley index
(Seeley 1949)

- Basic idea: in a group of children, a child is as popular as the sum of the popularities of the children who like him, but popularities are divided evenly among friends:

$$c_{Seeley}(x) = \sum_{y \to x} \frac{c_{Seeley}(y)}{d^+(y)}$$

- In general it is a left dominant eigenvector of G_r

Wei index

(Wei 1953)

Wei index
(Wei 1953)

* Basic idea: if a matrix represents whether a team defeated another team, we can define a general score by iteratively computing the sum of the scores of the teams that have been defeated:

Wei index
(Wei 1953)

* Basic idea: if a matrix represents whether a team defeated another team, we can define a general score by iteratively computing the sum of the scores of the teams that have been defeated:

$$c_{\mathrm{W}ei}(x) \leftarrow \sum_{x \to y} c_{\mathrm{W}ei}(y)$$

Wei index
(Wei 1953)

* Basic idea: if a matrix represents whether a team defeated another team, we can define a general score by iteratively computing the sum of the scores of the teams that have been defeated:

$$c_{\mathrm{W}ei}(x) \leftarrow \sum_{x \to y} c_{\mathrm{W}ei}(y)$$

* In general the fixpoint is a right dominant eigenvector of G

Wei index
(Wei 1953)

* Basic idea: if a matrix represents whether a team defeated another team, we can define a general score by iteratively computing the sum of the scores of the teams that have been defeated:

$$c_{\mathrm{W}ei}(x) \leftarrow \sum_{x \to y} c_{\mathrm{W}ei}(y)$$

* In general the fixpoint is a right dominant eigenvector of G

* Often called “Perron-Frobenius ranking” in the literature about ranking for sport teams

PageRank

(Brin, Page, Motwani, Winograd 1999)

PageRank

(Brin, Page, Motwani, Winograd 1999)

* The idea is to start from Seeley's equation and add an adjustment to make it have a unique solution (and more)

PageRank

(Brin, Page, Motwani, Winograd 1999)

* The idea is to start from Seeley's equation and add an adjustment to make it have a unique solution (and more)

$$c_{\mathrm{pr}}(x) = \alpha \sum_{y \to x} \frac{c_{\mathrm{pr}}(y)}{d^+(y)} + \frac{1-\alpha}{n}$$

PageRank

(Brin, Page, Motwani, Winograd 1999)

* The idea is to start from Seeley's equation and add an adjustment to make it have a unique solution (and more)

$$c_{\mathrm{pr}}(x) = \alpha \sum_{y \to x} \frac{c_{\mathrm{pr}}(y)}{d^+(y)} + \frac{1-\alpha}{n}$$

* It is the dominant eigenvector of

PageRank

(Brin, Page, Motwani, Winograd 1999)

* The idea is to start from Seeley's equation and add an adjustment to make it have a unique solution (and more)

$$c_{\mathrm{pr}}(x) = \alpha \sum_{y \to x} \frac{c_{\mathrm{pr}}(y)}{d^+(y)} + \frac{1-\alpha}{n}$$

* It is the dominant eigenvector of $\alpha G_r + (1-\alpha)\mathbf{1}^T\mathbf{1}/n$
* Or just

PageRank

(Brin, Page, Motwani, Winograd 1999)

* The idea is to start from Seeley's equation and add an adjustment to make it have a unique solution (and more)

$$c_{\mathrm{pr}}(x) = \alpha \sum_{y \to x} \frac{c_{\mathrm{pr}}(y)}{d^+(y)} + \frac{1-\alpha}{n}$$

* It is the dominant eigenvector of $\alpha G_r + (1-\alpha)\mathbf{1}^T\mathbf{1}/n$
* Or just $\mathbf{1}\sum_{t=0}^{\infty} \alpha^t G_r^t$

PageRank

(Brin, Page, Motwani, Winograd 1999)

* The idea is to start from Seeley's equation and add an adjustment to make it have a unique solution (and more)

$$c_{\mathrm{pr}}(x) = \alpha \sum_{y \to x} \frac{c_{\mathrm{pr}}(y)}{d^+(y)} + \frac{1-\alpha}{n}$$

* It is the dominant eigenvector of $\alpha G_r + (1-\alpha)\mathbf{1}^T\mathbf{1}/n$
* Or just $\mathbf{1}\sum_{t=0}^{\infty} \alpha^t G_r^t$
* Katz is the dominant eigenvector of $\alpha G + (1-\alpha)\mathbf{e}^T\mathbf{1}/n$

PageRank

(Brin, Page, Motwani, Winograd 1999)

* The idea is to start from Seeley's equation and add an adjustment to make it have a unique solution (and more)

$$c_{\mathrm{pr}}(x) = \alpha \sum_{y \to x} \frac{c_{\mathrm{pr}}(y)}{d^+(y)} + \frac{1-\alpha}{n}$$

* It is the dominant eigenvector of $\alpha G_r + (1-\alpha)\mathbf{1}^T\mathbf{1}/n$
* Or just $\mathbf{1}\sum_{t=0}^{\infty} \alpha^t G_r^t$
* Katz is the dominant eigenvector of $\alpha G + (1-\alpha)\mathbf{e}^T\mathbf{1}/n$
* Or just

PageRank

(Brin, Page, Motwani, Winograd 1999)

* The idea is to start from Seeley's equation and add an adjustment to make it have a unique solution (and more)

$$c_{\mathrm{pr}}(x) = \alpha \sum_{y \to x} \frac{c_{\mathrm{pr}}(y)}{d^+(y)} + \frac{1-\alpha}{n}$$

* It is the dominant eigenvector of $\alpha G_r + (1-\alpha)\mathbf{1}^T\mathbf{1}/n$
* Or just $\mathbf{1}\sum_{t=0}^{\infty} \alpha^t G_r^t$
* Katz is the dominant eigenvector of $\alpha G + (1-\alpha)\mathbf{e}^T\mathbf{1}/n$
* Or just $\mathbf{1}\sum_{t=0}^{\infty} \alpha^t G^t$

PageRank

(Brin, Page, Motwani, Winograd 1999)

* The idea is to start from Seeley's equation and add an adjustment to make it have a unique solution (and more)

$$c_{\mathrm{pr}}(x) = \alpha \sum_{y \to x} \frac{c_{\mathrm{pr}}(y)}{d^+(y)} + \frac{1-\alpha}{n}$$

* It is the dominant eigenvector of $\alpha G_r + (1-\alpha)\mathbf{1}^T\mathbf{1}/n$
* Or just $\mathbf{1}\sum_{t=0}^{\infty} \alpha^t G_r^t$
* Katz is the dominant eigenvector of $\alpha G + (1-\alpha)\mathbf{e}^T\mathbf{1}/n$
* Or just $\mathbf{1}\sum_{t=0}^{\infty} \alpha^t G^t$
* See my note "Spectral Ranking" for details

HITS

(Kleinberg 1997)

HITS
(Kleinberg 1997)

* The idea is to start from the system:

$$c_{\mathrm{Hauth}}(x) = \sum_{y \to x} c_{\mathrm{Hhub}}(y)$$

$$c_{\mathrm{Hhub}}(x) = \sum_{x \to y} c_{\mathrm{Hauth}}(y)$$

HITS
(Kleinberg 1997)

* The idea is to start from the system:

$$c_{\mathrm{Hauth}}(x) = \sum_{y \to x} c_{\mathrm{Hhub}}(y)$$

$$c_{\mathrm{Hhub}}(x) = \sum_{x \to y} c_{\mathrm{Hauth}}(y)$$

* HITS centrality is defined to be the “authoritativeness” score

HITS
(Kleinberg 1997)

* The idea is to start from the system:

$$c_{\mathrm{Hauth}}(x) = \sum_{y \to x} c_{\mathrm{Hhub}}(y)$$

$$c_{\mathrm{Hhub}}(x) = \sum_{x \to y} c_{\mathrm{Hauth}}(y)$$

* HITS centrality is defined to be the "authoritativeness" score
* It is a dominant eigenvector of $G^T G$, so it coincides with the dominant eigenvector on symmetric graphs

SALSA

(Lempel, Moran 2001)

SALSA

(Lempel, Moran 2001)

* The idea is to start from the system:

$$c_{\mathrm{Sauth}}(x) = \sum_{y \to x} \frac{c_{\mathrm{Shub}}(y)}{d^+(y)}$$

$$c_{\mathrm{Shub}}(x) = \sum_{x \to y} \frac{c_{\mathrm{Sauth}}(y)}{d^-(y)}$$

SALSA

(Lempel, Moran 2001)

* The idea is to start from the system:

$$c_{\mathrm{Sauth}}(x) = \sum_{y \to x} \frac{c_{\mathrm{Shub}}(y)}{d^+(y)}$$

$$c_{\mathrm{Shub}}(x) = \sum_{x \to y} \frac{c_{\mathrm{Sauth}}(y)}{d^-(y)}$$

* SALSA centrality is defined to be the "authoritativeness" score

SALSA

(Lempel, Moran 2001)

* The idea is to start from the system:

$$c_{\mathrm{Sauth}}(x) = \sum_{y \to x} \frac{c_{\mathrm{Shub}}(y)}{d^+(y)}$$

$$c_{\mathrm{Shub}}(x) = \sum_{x \to y} \frac{c_{\mathrm{Sauth}}(y)}{d^-(y)}$$

* SALSA centrality is defined to be the "authoritativeness" score

* It is a dominant eigenvector of $G_c^T G_r$

HITS (, SALSA etc.)

HITS (, SALSA etc.)

* **WARNING:** These measures were proposed exactly for ranking results in hyperlinked collections

HITS (, SALSA etc.)

- **WARNING:** These measures were proposed exactly for ranking results in hyperlinked collections
- Should be applied *not* to the **whole graph**, but to a *suitable* **subgraph** derived from the query

HITS (, SALSA etc.)

- **WARNING:** These measures were proposed exactly for ranking results in hyperlinked collections
- Should be applied *not* to the **whole graph**, but to a *suitable* **subgraph** derived from the query
- How the subgraph is derived is *very relevant* for effectiveness (Najork, Gollapudi, Panighray 2009)

HITS (, SALSA etc.)

- **WARNING:** These measures were proposed exactly for ranking results in hyperlinked collections
- Should be applied *not* to the **whole graph**, but to a *suitable* **subgraph** derived from the query
- How the subgraph is derived is *very relevant* for effectiveness (Najork, Gollapudi, Panighray 2009)
- Not really the central point here, though...

How?

How to assess centrality

How?

How to assess centrality

- *Axiomatic approach*

How?

How to assess centrality

- *Axiomatic approach*
- *Ground-truth approach*

How?

How to assess centrality

- *Axiomatic approach*
- *Ground-truth approach*
- *IR approach*

How?

How to assess centrality

- *Axiomatic approach*
- *Ground-truth approach*
- *IR approach*
- *Computational feasibility approach*

Axiomatic lens

Axiomatic lens

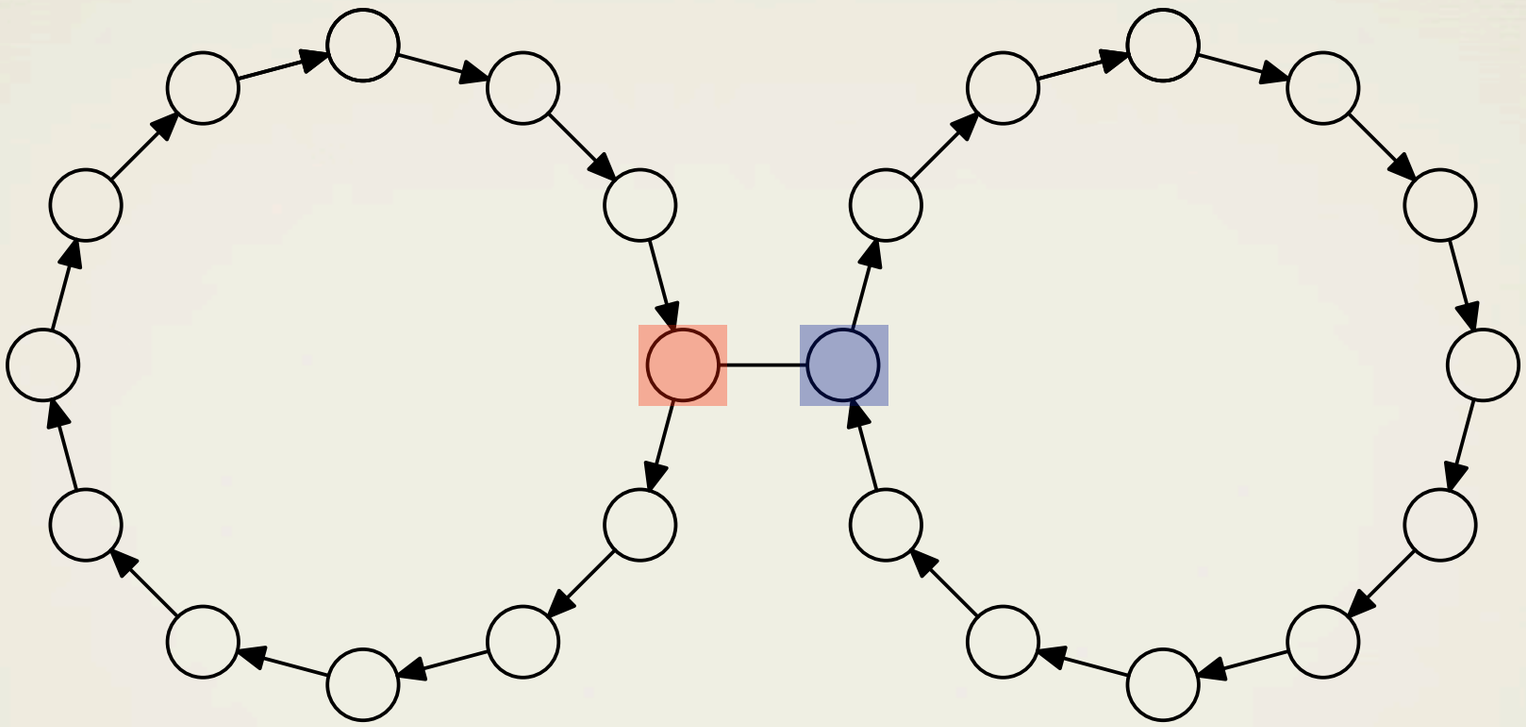
start from some minimal mathematical requirements

Axiomatic

Sensitivity to density

Axiomatic

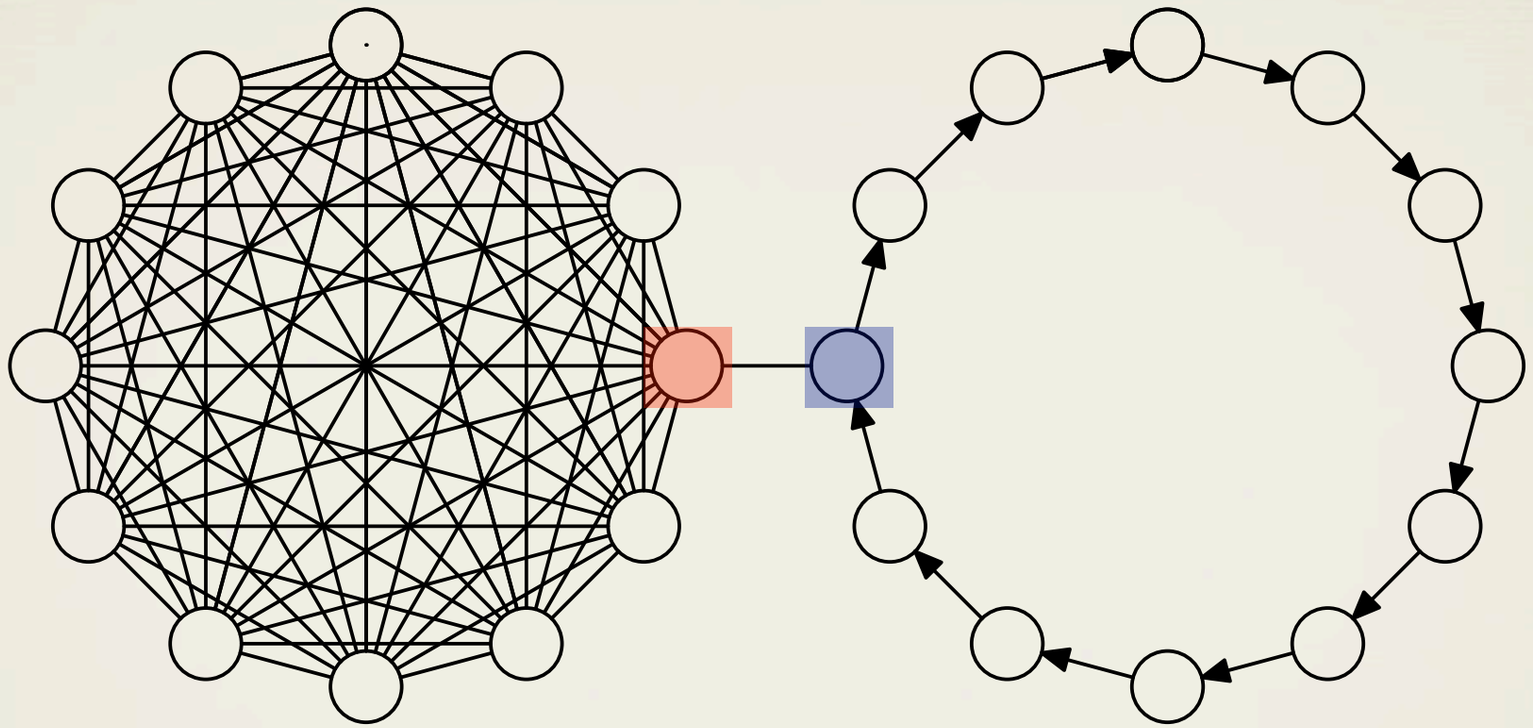
Sensitivity to density



The blue and the red node have the same importance (the two rings have the same size!)

Axiomatic

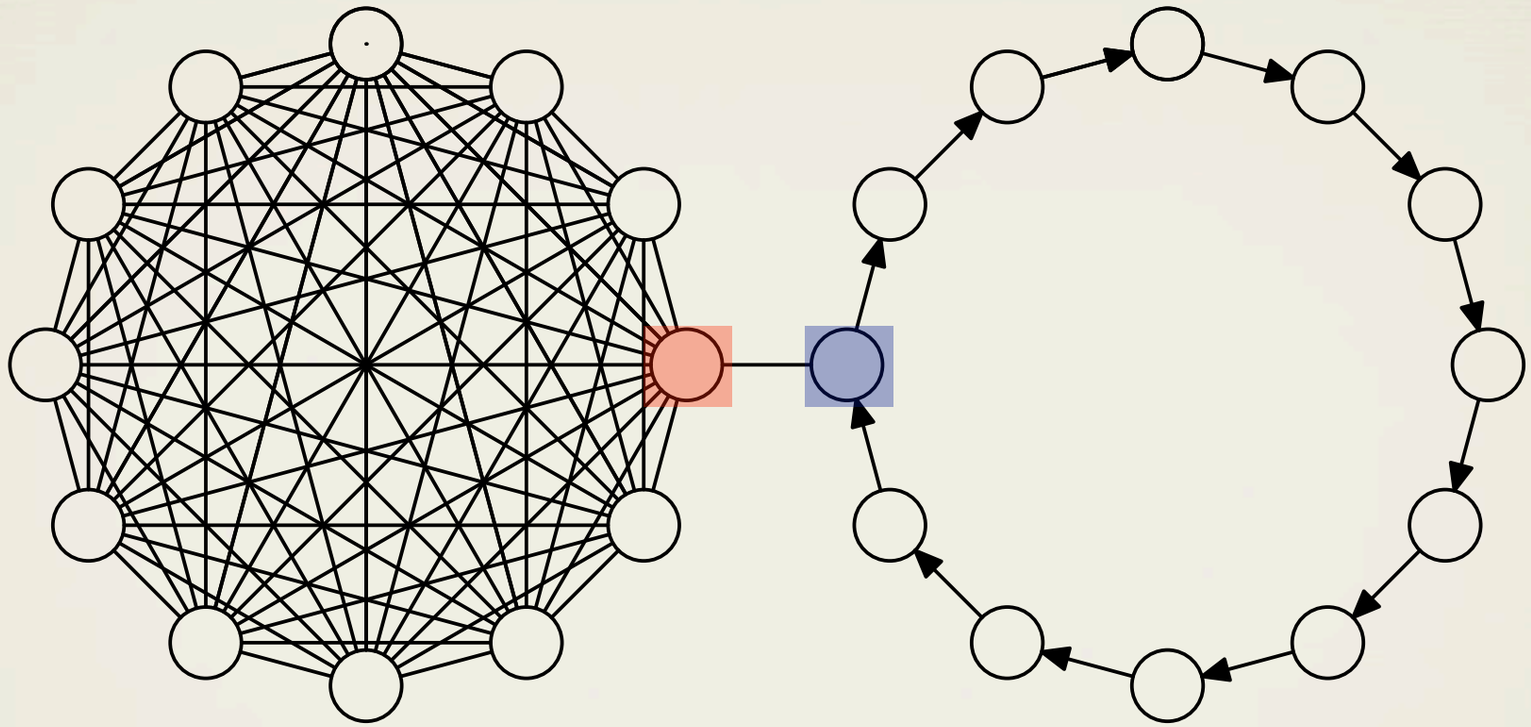
Sensitivity to density



The blue and the red node have the same importance (the two rings have the same size!)

Axiomatic

Sensitivity to density



Densifying the left-hand side, we expect the red node to become more important than the blue node

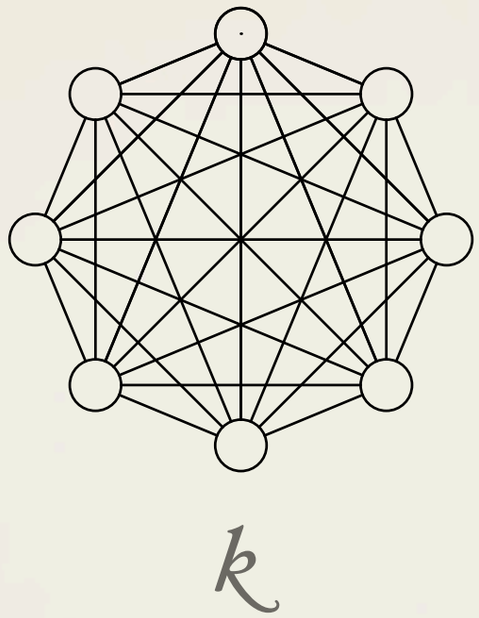
Axiomatic

Sensitivity to size

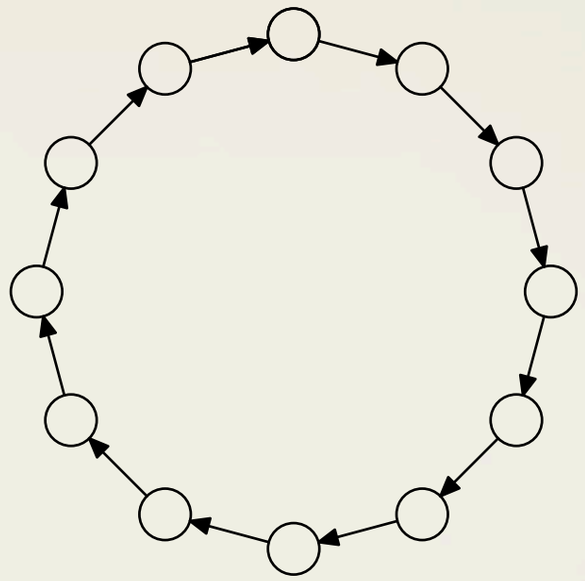
Axiomatic

Sensitivity to size

Two disjoint (or very far) components of a single network



k

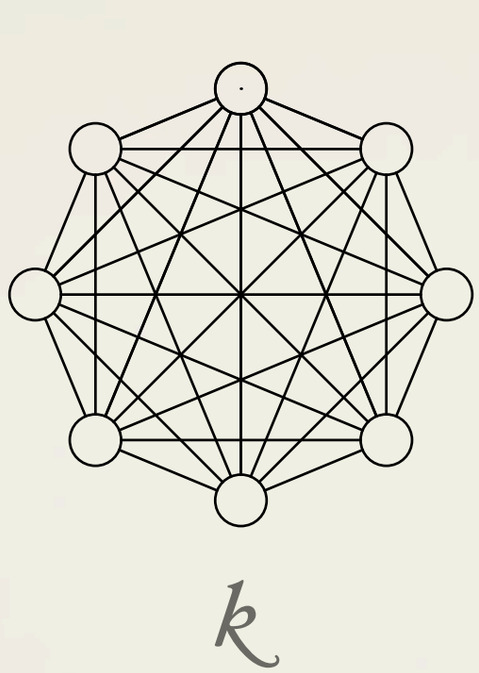


p

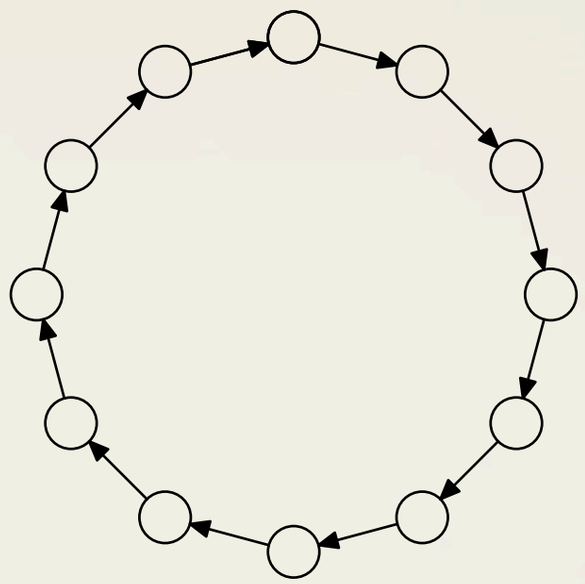
Axiomatic

Sensitivity to size

Two disjoint (or very far) components of a single network



k



p

When k or p goes to ∞, the nodes of the corresponding subnetwork must become more important

An axiomatic slaughter

An axiomatic slaughter

An axiomatic slaughter

Degree		
Betweenness		
Katz		
Closeness		
Lin		
Harmonic		
PageRank		
Seeley		
HITS		
SALSA		

An axiomatic slaughter

	Density	
Degree	yes	
Betweenness	no (!)	
Katz	yes	
Closeness	no	
Lin	no	
Harmonic	**yes**	
PageRank	yes	
Seeley	yes	
HITS	yes	
SALSA	yes	

An axiomatic slaughter

	Density	Size
Degree	yes	only k
Betweenness	no (!)	only p
Katz	yes	only k
Closeness	no	no (!)
Lin	no	only k
Harmonic	**yes**	**yes**
PageRank	yes	no
Seeley	yes	no
HITS	yes	no
SALSA	yes	no

Ground-truth lens
(mostly: anecdotal / comparative)

Ground-truth lens
(mostly: anecdotal / comparative)

check them against real (social?) networks on which you have some ground truth about importance/centrality/...

Hollywood: PageRank

Ron Jeremy

Adolf Hitler

Lloyd Kaufman

George W. Bush

Ronald Reagan

Bill Clinton

Martin Sheen

Debbie Rochon

Hollywood: PageRank

Ron Jeremy

Adolf Hitler

Lloyd Kaufman

George W. Bush

Ronald Reagan

Bill Clinton

Martin Sheen

Debbie Rochon

Hollywood: Degree

William Shatner

Bess Flowers

Martin Sheen

Ronald Reagan

George Clooney

Samuel Jackson

Robin Williams

Tom Hanks

Hollywood: Degree

William Shatner

Bess Flowers

Martin Sheen

Ronald Reagan

George Clooney

Samuel Jackson

Robin Williams

Tom Hanks

Hollywood: Betweenness

Adolf Hitler

Lloyd Kaufman

Ron Jeremy

Tony Robinson

Olu Jacobs

Max von Sydow

Udo Kier

George W. Bush

Hollywood: Betweenness

Adolf Hitler

Lloyd Kaufman

Ron Jeremy

Tony Robinson

Olu Jacobs

Max von Sydow

Udo Kier

George W. Bush

Hollywood: Katz

William Shatner

Martin Sheen

Tom Hanks

Robin Williams

George Clooney

Ronald Reagan

Bruce Willis

Samuel Jackson

Hollywood: Katz

William Shatner

Martin Sheen

Tom Hanks

Robin Williams

George Clooney

Ronald Reagan

Bruce Willis

Samuel Jackson

Hollywood: Closeness

Lina Tjeng

Anh Loan Nguyen Thi

Ryan Villapoto

Chad Reed

Bjorn van Wenum

J.P. Ramackers

Herbert Sydney

R.D. Nicholson

Hollywood: Closeness

Lina Tjeng

Anh Loan Nguyen Thi

Ryan Villapoto

Chad Reed

Bjorn van Wenum

J.P. Ramackers

Herbert Sydney

R.D. Nicholson

Hollywood: Closeness

Lina Tjeng

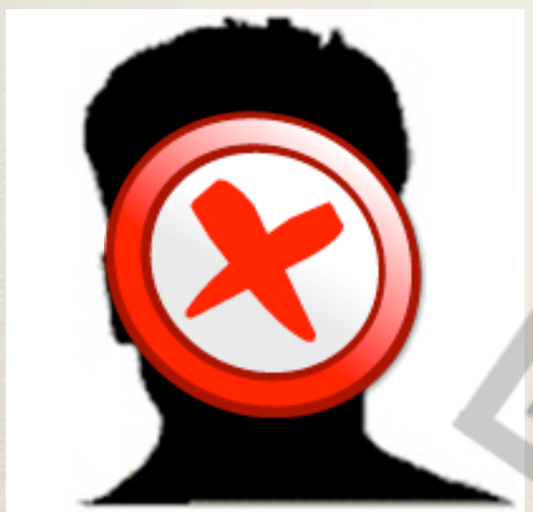
Anh Loan Nguyen Thi

Ryan [illegible]oto

Chad Reed

have largest centrality

Isolated nodes ha



Bjorn van Wenum

J.P. Ramackers

Herbert Sydney

R.D. Nicholson

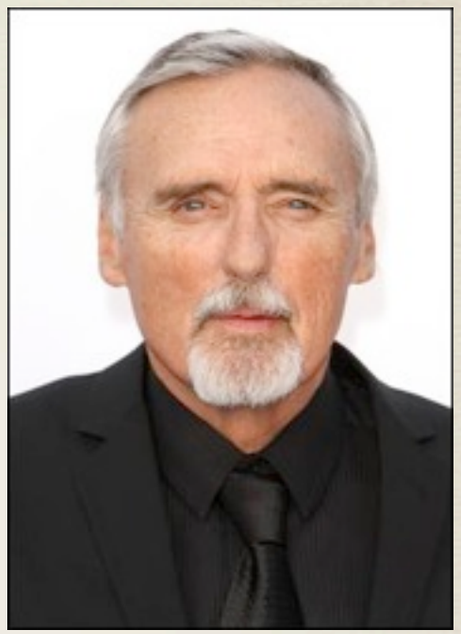
Hollywood: Lin

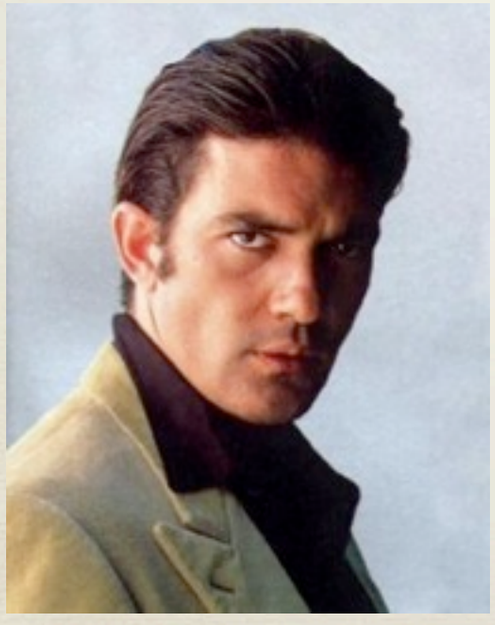
George Clooney

Samuel Jackson

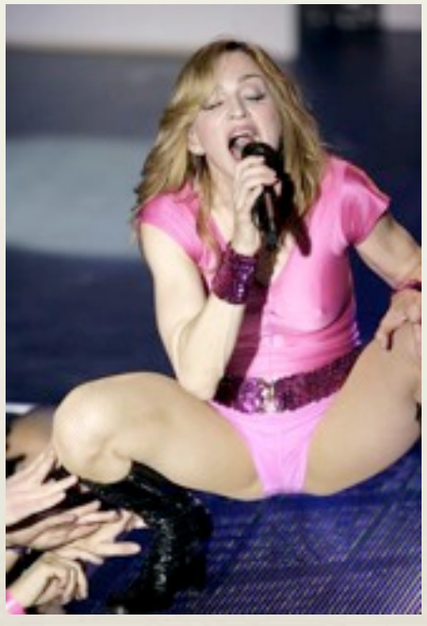
Martin Sheen

Dennis Hopper



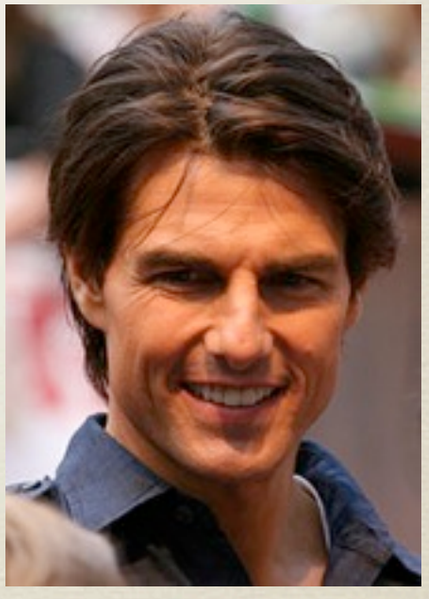


Antonio Banderas



Madonna

Michael Douglas



Tom Cruise

Hollywood: HITS

Tom Hanks

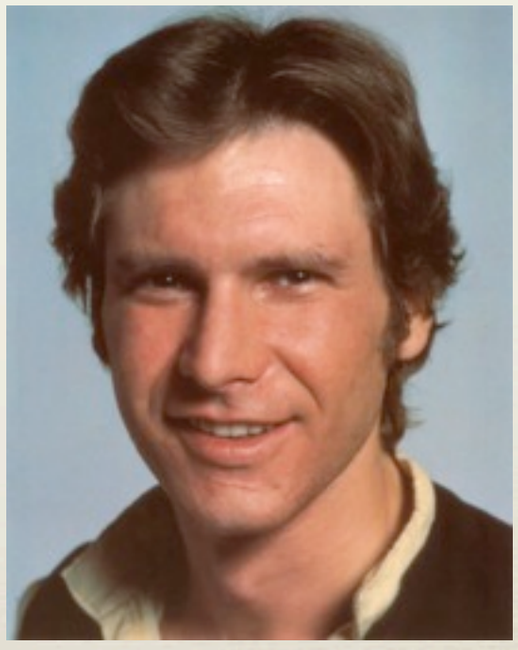
William Shatner

Robin Williams

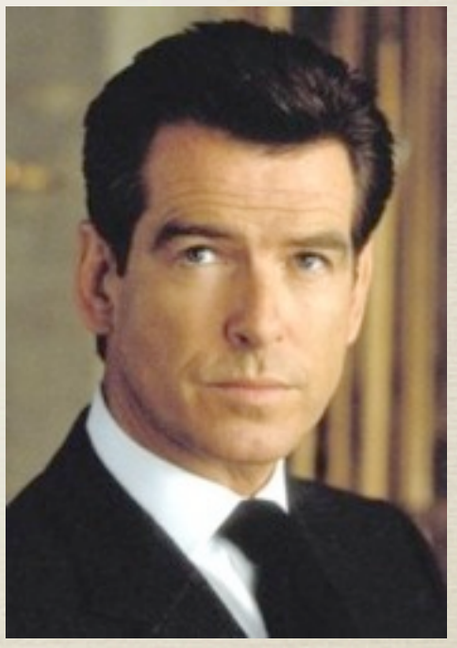
Bruce Willis

Michael Douglas

Cameron Diaz



Harrison Ford



Pierce Brosnan

Hollywood: **Harmonic**

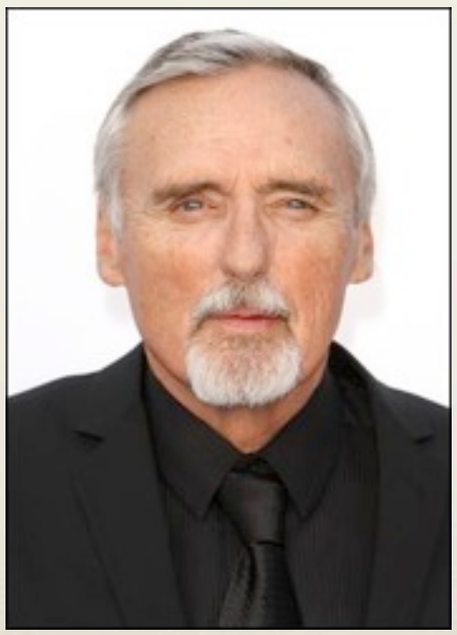
George Clooney

Samuel Jackson

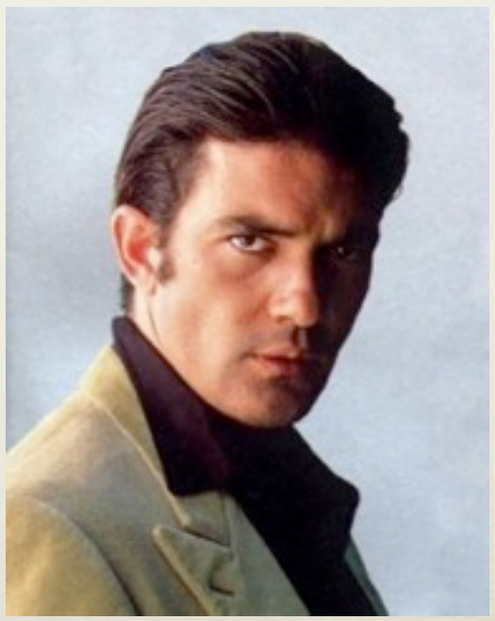
Sharon Stone

Tom Hanks

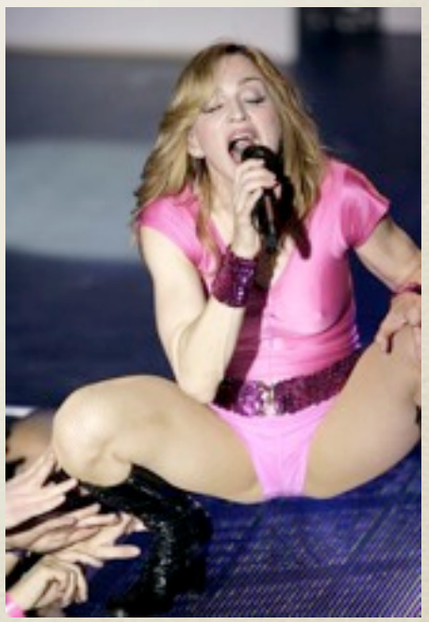
Martin Sheen



Dennis Hopper



Antonio Banderas



Madonna

What about the web?
.uk Top Ten

What about the web?
.uk Top Ten

What about the web?
.uk Top Ten

PageRank			
http://www.direct.gov.uk/			
http://www.direct.gov.uk/en/index.htm			
http://www.names.co.uk/			
http://www.names.co.uk/hosting.html			
http://www.names.co.uk/email.html			
http://www.names.co.uk/controlpanel.html			
http://www.scdc.org.uk/			
http://www.freelyricsearch.co.uk/index.html			
http://www.247partypeople.co.uk/login.asp			
http://www.becs.co.uk/catalog/cookie_usage.php			

What about the web?
.uk Top Ten

PageRank	Katz		
http://www.direct.gov.uk/	http://www.direct.gov.uk/		
http://www.direct.gov.uk/en/index.htm	http://www.kelkoo.co.uk/		
http://www.names.co.uk/	http://www.kelkoo.co.uk/b/a/kc_top_searches_charts.html		
http://www.names.co.uk/hosting.html	http://www.kelkoo.co.uk/b/a/co_2765_128501-company-information-pages.html		
http://www.names.co.uk/email.html	http://www.kelkoo.co.uk/b/a/sm_site-map.html?displayType=alpha		
http://www.names.co.uk/controlpanel.html	http://www.kelkoo.co.uk/b/a/co_5199_128501-how-to-use-kelkoo.html		
http://www.scdc.org.uk/	http://www.kelkoo.co.uk/b/a/co_2120_128501-shopping-guides-price-comparison-on-kelkoo-uk.html		
http://www.freelyricsearch.co.uk/index.html	http://www.ebay.co.uk/		
http://www.247partypeople.co.uk/login.asp	http://www.top50scrappers.co.uk/		
http://www.becs.co.uk/catalog/cookie_usage.php	http://cgi1.ebay.co.uk/aw-cgi/eBayISAPI.dll?TimeShow		

What about the web?
.uk Top Ten

PageRank	Katz	Lin	
http://www.direct.gov.uk/	http://www.direct.gov.uk/	http://www.direct.gov.uk/	
http://www.direct.gov.uk/en/index.htm	http://www.kelkoo.co.uk/	http://www.bbc.co.uk/accessibility/	
http://www.names.co.uk/	http://www.kelkoo.co.uk/b/a/kc_top_searches_charts.html	http://news.bbc.co.uk/	
http://www.names.co.uk/hosting.html	http://www.kelkoo.co.uk/b/a/co_2765_128501-company-information-pages.html	http://www.dti.gov.uk/	
http://www.names.co.uk/email.html	http://www.kelkoo.co.uk/b/a/sm_site-map.html?displayType=alpha	http://www.google.co.uk/	
http://www.names.co.uk/controlpanel.html	http://www.kelkoo.co.uk/b/a/co_5199_128501-how-to-use-kelkoo.html	http://www.guardian.co.uk/	
http://www.scdc.org.uk/	http://www.kelkoo.co.uk/b/a/co_2120_128501-shopping-guides-price-comparison-on-kelkoo-uk.html	http://www.homeoffice.gov.uk/	
http://www.freelyricsearch.co.uk/index.html	http://www.ebay.co.uk/	http://www.statistics.gov.uk/	
http://www.247partypeople.co.uk/login.asp	http://www.top50scrappers.co.uk/	http://www.bbc.co.uk/privacy/	
http://www.becs.co.uk/catalog/cookie_usage.php	http://cgi1.ebay.co.uk/aw-cgi/eBayISAPI.dll?TimeShow	http://www.bbc.co.uk/info/	

What about the web?
.uk Top Ten

PageRank	Katz	Lin	Harmonic
http://www.direct.gov.uk/	http://www.direct.gov.uk/	http://www.direct.gov.uk/	http://www.direct.gov.uk/
http://www.direct.gov.uk/en/index.htm	http://www.kelkoo.co.uk/	http://www.bbc.co.uk/accessibility/	http://news.bbc.co.uk/
http://www.names.co.uk/	http://www.kelkoo.co.uk/b/a/kc_top_searches_charts.html	http://news.bbc.co.uk/	http://www.dti.gov.uk/
http://www.names.co.uk/hosting.html	http://www.kelkoo.co.uk/b/a/co_2765_128501-company-information-pages.html	http://www.dti.gov.uk/	http://www.direct.gov.uk/en/index.htm
http://www.names.co.uk/email.html	http://www.kelkoo.co.uk/b/a/sm_site-map.html?displayType=alpha	http://www.google.co.uk/	http://www.google.co.uk/
http://www.names.co.uk/controlpanel.html	http://www.kelkoo.co.uk/b/a/co_5199_128501-how-to-use-kelkoo.html	http://www.guardian.co.uk/	http://www.bbc.co.uk/accessibility/
http://www.scdc.org.uk/	http://www.kelkoo.co.uk/b/a/co_2120_128501-shopping-guides-price-comparison-on-kelkoo-uk.html	http://www.homeoffice.gov.uk/	http://www.homeoffice.gov.uk/
http://www.freelyricsearch.co.uk/index.html	http://www.ebay.co.uk/	http://www.statistics.gov.uk/	http://www.statistics.gov.uk/
http://www.247partypeople.co.uk/login.asp	http://www.top50scrappers.co.uk/	http://www.bbc.co.uk/privacy/	http://www.bbc.co.uk/privacy/
http://www.becs.co.uk/catalog/cookie_usage.php	http://cgi1.ebay.co.uk/aw-cgi/eBayISAPI.dll?TimeShow	http://www.bbc.co.uk/info/	http://www.bbc.co.uk/info/

How do they compare?

Hollywood
Kendall's τ

	degree	Katz 1/4λ	Katz 1/2λ	Katz 3/4λ	SALSA	closeness	harmonic	Lin	HITS	between	PR 1/4	PR 1/2	PR 3/4
degree	1.0000	0.9709	0.9287	0.8627	0.9005	0.4357	0.5526	0.5512	0.5170	0.5034	0.3699	0.4225	0.5074
Katz 1/4λ	0.9709	1.0000	0.9609	0.8957	0.8719	0.4638	0.5816	0.5801	0.5476	0.5026	0.3448	0.3964	0.4801
Katz 1/2λ	0.9287	0.9609	1.0000	0.9369	0.8291	0.4965	0.6157	0.6139	0.5849	0.4952	0.3108	0.3605	0.4416
Katz 3/4λ	0.8627	0.8957	0.9369	1.0000	0.7630	0.5488	0.6697	0.6676	0.6478	0.4811	0.2633	0.3098	0.3865
SALSA	0.9005	0.8719	0.8291	0.7630	1.0000	0.5371	0.4519	0.4504	0.4185	0.4692	0.4496	0.5042	0.5924
closeness	0.4357	0.4638	0.4965	0.5488	0.5371	1.0000	0.8503	0.8508	0.7366	0.3293	0.1529	0.1813	0.2319
harmonic	0.5526	0.5816	0.6157	0.6697	0.4519	0.8503	1.0000	0.9925	0.8694	0.3929	0.0752	0.1041	0.1549
Lin	0.5512	0.5801	0.6139	0.6676	0.4504	0.8508	0.9925	1.0000	0.8680	0.3916	0.0753	0.1041	0.1546
HITS	0.5170	0.5476	0.5849	0.6478	0.4185	0.7366	0.8694	0.8680	1.0000	0.3645	0.0518	0.0780	0.1249
between	0.5034	0.5026	0.4952	0.4811	0.4692	0.3293	0.3929	0.3916	0.3696	1.0000	0.4852	0.4909	0.4923
PR 1/4	0.3699	0.3448	0.3108	0.2633	0.4496	0.1529	0.0752	0.0753	0.0518	0.4852	1.0000	0.9317	0.8276
PR 1/2	0.4225	0.3964	0.3605	0.3098	0.5042	0.1813	0.1041	0.1041	0.0780	0.4909	0.9317	1.0000	0.8952
PR 3/4	0.5074	0.4801	0.4416	0.3865	0.5924	0.2319	0.1549	0.1546	0.1249	0.4923	0.8276	0.8952	1.0000

Hollywood
Kendall's τ

	degree	Katz 1/4λ	Katz 1/2λ	Katz 3/4λ	SALSA	closeness	harmonic	Lin	HITS	between	PR 1/4	PR 1/2	PR 3/4
degree	1.0000	0.9709	0.9287	0.8627	0.9005	0.4357	0.5526	0.5512	0.5170	0.5034	0.3699	0.4225	0.5074
Katz 1/4λ	0.9709	1.0000	0.9609	0.8957	0.8719	0.4638	0.5816	0.5801	0.5476	0.5026	0.3448	0.3964	0.4801
Katz 1/2λ	0.9287	0.9609	1.0000	0.9369	0.8291	0.4965	0.6157	0.6139	0.5849	0.4952	0.3108	0.3605	0.4416
Katz 3/4λ	0.8627	0.8957	0.9369	1.0000	0.7630	0.5488	0.6697	0.6676	0.6478	0.4811	0.2633	0.3098	0.3865
SALSA	0.9005	0.8719	0.8291	0.7630	1.0000	0.5371	0.4519	0.4504	0.4185	0.4692	0.4496	0.5042	0.5924
closeness	0.4357	0.4638	0.4965	0.5488	0.5371	1.0000	0.8503	0.8508	0.7366	0.3293	0.1529	0.1813	0.2319
harmonic	0.5526	0.5816	0.6157	0.6697	0.4519	0.8503	1.0000	0.9925	0.8694	0.3929	0.0752	0.1041	0.1549
Lin	0.5512	0.5801	0.6139	0.6676	0.4504	0.8508	0.9925	1.0000	0.8680	0.3916	0.0753	0.1041	0.1546
HITS	0.5170	0.5476	0.5849	0.6478	0.4185	0.7366	0.8694	0.8680	1.0000	0.3645	0.0518	0.0780	0.1249
between	0.5034	0.5026	0.4952	0.4811	0.4692	0.3293	0.3929	0.3916	0.3696	1.0000	0.4852	0.4909	0.4923
PR 1/4	0.3699	0.3448	0.3108	0.2633	0.4496	0.1529	0.0752	0.0753	0.0518	0.4852	1.0000	0.9317	0.8276
PR 1/2	0.4225	0.3964	0.3605	0.3098	0.5042	0.1813	0.1041	0.1041	0.0780	0.4909	0.9317	1.0000	0.8952
PR 3/4	0.5074	0.4801	0.4416	0.3865	0.5924	0.2319	0.1549	0.1546	0.1249	0.4923	0.8276	0.8952	1.0000

most geometric indices and HITS
are rather correlated to one another

Hollywood
Kendall's τ

	degree	Katz 1/4λ	Katz 1/2λ	Katz 3/4λ	SALSA	closeness	harmonic	Lin	HITS	between	PR 1/4	PR 1/2	PR 3/4
degree	1.0000	0.9709	0.9287	0.8627	0.9005	0.4357	0.5526	0.5512	0.5170	0.5034	0.3699	0.4225	0.5074
Katz 1/4λ	0.9709	1.0000	0.9609	0.8957	0.8719	0.4638	0.5816	0.5801	0.5476	0.5026	0.3448	0.3964	0.4801
Katz 1/2λ	0.9287	0.9609	1.0000	0.9369	0.8291	0.4965	0.6157	0.6139	0.5849	0.4952	0.3108	0.3605	0.4416
Katz 3/4λ	0.8627	0.8957	0.9369	1.0000	0.7630	0.5488	0.6697	0.6676	0.6478	0.4811	0.2633	0.3098	0.3865
SALSA	0.9005	0.8719	0.8291	0.7630	1.0000	0.5371	0.4519	0.4504	0.4185	0.4692	0.4496	0.5042	0.5924
closeness	0.4357	0.4638	0.4965	0.5488	0.5371	1.0000	0.8503	0.8508	0.7366	0.3293	0.1529	0.1813	0.2319
harmonic	0.5526	0.5816	0.6157	0.6697	0.4519	0.8503	1.0000	0.9925	0.8694	0.3929	0.0752	0.1041	0.1549
Lin	0.5512	0.5801	0.6139	0.6676	0.4504	0.8508	0.9925	1.0000	0.8680	0.3916	0.0753	0.1041	0.1546
HITS	0.5170	0.5476	0.5849	0.6478	0.4185	0.7366	0.8694	0.8680	1.0000	0.3645	0.0518	0.0780	0.1249
between	0.5034	0.5026	0.4952	0.4811	0.4692	0.3293	0.3929	0.3916	0.3696	1.0000	0.4852	0.4909	0.4923
PR 1/4	0.3699	0.3448	0.3108	0.2633	0.4496	0.1529	0.0752	0.0753	0.0518	0.4852	1.0000	0.9317	0.8276
PR 1/2	0.4225	0.3964	0.3605	0.3098	0.5042	0.1813	0.1041	0.1041	0.0780	0.4909	0.9317	1.0000	0.8952
PR 3/4	0.5074	0.4801	0.4416	0.3865	0.5924	0.2319	0.1549	0.1546	0.1249	0.4923	0.8276	0.8952	1.0000

Hollywood
Kendall's τ

	degree	Katz 1/4λ	Katz 1/2λ	Katz 3/4λ	SALSA	closeness	harmonic	Lin	HITS	between	PR 1/4	PR 1/2	PR 3/4
degree	1.0000	0.9709	0.9287	0.8627	0.9005	0.4357	0.5526	0.5512	0.5170	0.5034	0.3699	0.4225	0.5074
Katz 1/4λ	0.9709	1.0000	0.9609	0.8957	0.8719	0.4638	0.5816	0.5801	0.5476	0.5026	0.3448	0.3964	0.4801
Katz 1/2λ	0.9287	0.9609	1.0000	0.9369	0.8291	0.4965	0.6157	0.6139	0.5849	0.4952	0.3108	0.3605	0.4416
Katz 3/4λ	0.8627	0.8957	0.9369	1.0000	0.7630	0.5488	0.6697	0.6676	0.6478	0.4811	0.2633	0.3098	0.3865
SALSA	0.9005	0.8719	0.8291	0.7630	1.0000	0.5371	0.4519	0.4504	0.4185	0.4692	0.4496	0.5042	0.5924
closeness	0.4357	0.4638	0.4965	0.5488	0.5371	1.0000	0.8503	0.8508	0.7366	0.3293	0.1529	0.1813	0.2319
harmonic	0.5526	0.5816	0.6157	0.6697	0.4519	0.8503	1.0000	0.9925	0.8694	0.3929	0.0752	0.1041	0.1549
Lin	0.5512	0.5801	0.6139	0.6676	0.4504	0.8508	0.9925	1.0000	0.8680	0.3916	0.0753	0.1041	0.1546
HITS	0.5170	0.5476	0.5849	0.6478	0.4185	0.7366	0.8694	0.8680	1.0000	0.3645	0.0518	0.0780	0.1249
between	0.5034	0.5026	0.4952	0.4811	0.4692	0.3293	0.3929	0.3916	0.3696	1.0000	0.4852	0.4909	0.4923
PR 1/4	0.3699	0.3448	0.3108	0.2633	0.4496	0.1529	0.0752	0.0753	0.0518	0.4852	1.0000	0.9317	0.8276
PR 1/2	0.4225	0.3964	0.3605	0.3098	0.5042	0.1813	0.1041	0.1041	0.0780	0.4909	0.9317	1.0000	0.8952
PR 3/4	0.5074	0.4801	0.4416	0.3865	0.5924	0.2319	0.1549	0.1546	0.1249	0.4923	0.8276	0.8952	1.0000

Katz, degree and SALSA are also highly correlated

Hollywood
Kendall's τ

	degree	Katz 1/4λ	Katz 1/2λ	Katz 3/4λ	SALSA	closeness	harmonic	Lin	HITS	between	PR 1/4	PR 1/2	PR 3/4
degree	1.0000	0.9709	0.9287	0.8627	0.9005	0.4357	0.5526	0.5512	0.5170	0.5034	0.3699	0.4225	0.5074
Katz 1/4λ	0.9709	1.0000	0.9609	0.8957	0.8719	0.4638	0.5816	0.5801	0.5476	0.5026	0.3448	0.3964	0.4801
Katz 1/2λ	0.9287	0.9609	1.0000	0.9369	0.8291	0.4965	0.6157	0.6139	0.5849	0.4952	0.3108	0.3605	0.4416
Katz 3/4λ	0.8627	0.8957	0.9369	1.0000	0.7630	0.5488	0.6697	0.6676	0.6478	0.4811	0.2633	0.3098	0.3865
SALSA	0.9005	0.8719	0.8291	0.7630	1.0000	0.5371	0.4519	0.4504	0.4185	0.4692	0.4496	0.5042	0.5924
closeness	0.4357	0.4638	0.4965	0.5488	0.5371	1.0000	0.8503	0.8508	0.7366	0.3293	0.1529	0.1813	0.2319
harmonic	0.5526	0.5816	0.6157	0.6697	0.4519	0.8503	1.0000	0.9925	0.8694	0.3929	0.0752	0.1041	0.1549
Lin	0.5512	0.5801	0.6139	0.6676	0.4504	0.8508	0.9925	1.0000	0.8680	0.3916	0.0753	0.1041	0.1546
HITS	0.5170	0.5476	0.5849	0.6478	0.4185	0.7366	0.8694	0.8680	1.0000	0.3645	0.0518	0.0780	0.1249
between	0.5034	0.5026	0.4952	0.4811	0.4692	0.3293	0.3929	0.3916	0.3696	1.0000	0.4852	0.4909	0.4923
PR 1/4	0.3699	0.3448	0.3108	0.2633	0.4496	0.1529	0.0752	0.0753	0.0518	0.4852	1.0000	0.9317	0.8276
PR 1/2	0.4225	0.3964	0.3605	0.3098	0.5042	0.1813	0.1041	0.1041	0.0780	0.4909	0.9317	1.0000	0.8952
PR 3/4	0.5074	0.4801	0.4416	0.3865	0.5924	0.2319	0.1549	0.1546	0.1249	0.4923	0.8276	0.8952	1.0000

Hollywood
Kendall's τ

	degree	Katz 1/4λ	Katz 1/2λ	Katz 3/4λ	SALSA	closeness	harmonic	Lin	HITS	between	PR 1/4	PR 1/2	PR 3/4
degree	1.0000	0.9709	0.9287	0.8627	0.9005	0.4357	0.5526	0.5512	0.5170	0.5034	0.3699	0.4225	0.5074
Katz 1/4λ	0.9709	1.0000	0.9609	0.8957	0.8719	0.4638	0.5816	0.5801	0.5476	0.5026	0.3448	0.3964	0.4801
Katz 1/2λ	0.9287	0.9609	1.0000	0.9369	0.8291	0.4965	0.6157	0.6139	0.5849	0.4952	0.3108	0.3605	0.4416
Katz 3/4λ	0.8627	0.8957	0.9369	1.0000	0.7630	0.5488	0.6697	0.6676	0.6478	0.4811	0.2633	0.3098	0.3865
SALSA	0.9005	0.8719	0.8291	0.7630	1.0000	0.5371	0.4519	0.4504	0.4185	0.4692	0.4496	0.5042	0.5924
closeness	0.4357	0.4638	0.4965	0.5488	0.5371	1.0000	0.8503	0.8508	0.7366	0.3293	0.1529	0.1813	0.2319
harmonic	0.5526	0.5816	0.6157	0.6697	0.4519	0.8503	1.0000	0.9925	0.8694	0.3929	0.0752	0.1041	0.1549
Lin	0.5512	0.5801	0.6139	0.6676	0.4504	0.8508	0.9925	1.0000	0.8680	0.3916	0.0753	0.1041	0.1546
HITS	0.5170	0.5476	0.5849	0.6478	0.4185	0.7366	0.8694	0.8680	1.0000	0.3645	0.0518	0.0780	0.1249
between	0.5034	0.5026	0.4952	0.4811	0.4692	0.3293	0.3929	0.3916	0.3696	1.0000	0.4852	0.4909	0.4923
PR 1/4	0.3699	0.3448	0.3108	0.2633	0.4496	0.1529	0.0752	0.0753	0.0518	0.4852	1.0000	0.9317	0.8276
PR 1/2	0.4225	0.3964	0.3605	0.3098	0.5042	0.1813	0.1041	0.1041	0.0780	0.4909	0.9317	1.0000	0.8952
PR 3/4	0.5074	0.4801	0.4416	0.3865	0.5924	0.2319	0.1549	0.1546	0.1249	0.4923	0.8276	0.8952	1.0000

Betweenness does not correlate to anything

Hollywood
Kendall's τ

	degree	Katz 1/4λ	Katz 1/2λ	Katz 3/4λ	SALSA	closeness	harmonic	Lin	HITS	between	PR 1/4	PR 1/2	PR 3/4
degree	1.0000	0.9709	0.9287	0.8627	0.9005	0.4357	0.5526	0.5512	0.5170	0.5034	0.3699	0.4225	0.5074
Katz 1/4λ	0.9709	1.0000	0.9609	0.8957	0.8719	0.4638	0.5816	0.5801	0.5476	0.5026	0.3448	0.3964	0.4801
Katz 1/2λ	0.9287	0.9609	1.0000	0.9369	0.8291	0.4965	0.6157	0.6139	0.5849	0.4952	0.3108	0.3605	0.4416
Katz 3/4λ	0.8627	0.8957	0.9369	1.0000	0.7630	0.5488	0.6697	0.6676	0.6478	0.4811	0.2633	0.3098	0.3865
SALSA	0.9005	0.8719	0.8291	0.7630	1.0000	0.5371	0.4519	0.4504	0.4185	0.4692	0.4496	0.5042	0.5924
closeness	0.4357	0.4638	0.4965	0.5488	0.5371	1.0000	0.8503	0.8508	0.7366	0.3293	0.1529	0.1813	0.2319
harmonic	0.5526	0.5816	0.6157	0.6697	0.4519	0.8503	1.0000	0.9925	0.8694	0.3929	0.0752	0.1041	0.1549
Lin	0.5512	0.5801	0.6139	0.6676	0.4504	0.8508	0.9925	1.0000	0.8680	0.3916	0.0753	0.1041	0.1546
HITS	0.5170	0.5476	0.5849	0.6478	0.4185	0.7366	0.8694	0.8680	1.0000	0.3645	0.0518	0.0780	0.1249
between	0.5034	0.5026	0.4952	0.4811	0.4692	0.3293	0.3929	0.3916	0.3696	1.0000	0.4852	0.4909	0.4923
PR 1/4	0.3699	0.3448	0.3108	0.2633	0.4496	0.1529	0.0752	0.0753	0.0518	0.4852	1.0000	0.9317	0.8276
PR 1/2	0.4225	0.3964	0.3605	0.3098	0.5042	0.1813	0.1041	0.1041	0.0780	0.4909	0.9317	1.0000	0.8952
PR 3/4	0.5074	0.4801	0.4416	0.3865	0.5924	0.2319	0.1549	0.1546	0.1249	0.4923	0.8276	0.8952	1.0000

Hollywood
Kendall's τ

	degree	Katz 1/4λ	Katz 1/2λ	Katz 3/4λ	SALSA	closeness	harmonic	Lin	HITS	between	PR 1/4	PR 1/2	PR 3/4
degree	1.0000	0.9709	0.9287	0.8627	0.9005	0.4357	0.5526	0.5512	0.5170	0.5034	0.3699	0.4225	0.5074
Katz 1/4λ	0.9709	1.0000	0.9609	0.8957	0.8719	0.4638	0.5816	0.5801	0.5476	0.5026	0.3448	0.3964	0.4801
Katz 1/2λ	0.9287	0.9609	1.0000	0.9369	0.8291	0.4965	0.6157	0.6139	0.5849	0.4952	0.3108	0.3605	0.4416
Katz 3/4λ	0.8627	0.8957	0.9369	1.0000	0.7630	0.5488	0.6697	0.6676	0.6478	0.4811	0.2633	0.3098	0.3865
SALSA	0.9005	0.8719	0.8291	0.7630	1.0000	0.5371	0.4519	0.4504	0.4185	0.4692	0.4496	0.5042	0.5924
closeness	0.4357	0.4638	0.4965	0.5488	0.5371	1.0000	0.8503	0.8508	0.7366	0.3293	0.1529	0.1813	0.2319
harmonic	0.5526	0.5816	0.6157	0.6697	0.4519	0.8503	1.0000	0.9925	0.8694	0.3929	0.0752	0.1041	0.1549
Lin	0.5512	0.5801	0.6139	0.6676	0.4504	0.8508	0.9925	1.0000	0.8680	0.3916	0.0753	0.1041	0.1546
HITS	0.5170	0.5476	0.5849	0.6478	0.4185	0.7366	0.8694	0.8680	1.0000	0.3645	0.0518	0.0780	0.1249
between	0.5034	0.5026	0.4952	0.4811	0.4692	0.3293	0.3929	0.3916	0.3696	1.0000	0.4852	0.4909	0.4923
PR 1/4	0.3699	0.3448	0.3108	0.2633	0.4496	0.1529	0.0752	0.0753	0.0518	0.4852	1.0000	0.9317	0.8276
PR 1/2	0.4225	0.3964	0.3605	0.3098	0.5042	0.1813	0.1041	0.1041	0.0780	0.4909	0.9317	1.0000	0.8952
PR 3/4	0.5074	0.4801	0.4416	0.3865	0.5924	0.2319	0.1549	0.1546	0.1249	0.4923	0.8276	0.8952	1.0000

PageRank stands alone

.uk (May 2007 snapshot)

Kendall's τ

	degree	Katz 1/4λ	Katz 1/2λ	Katz 3/4λ	SALSA	closeness	harmonic	Lin	HITS	PR 1/4	PR 1/2	PR 3/4
degree	1.0000	0.9053	0.9024	0.9000	0.9114	0.1950	0.2060	0.2060	0.2853	0.6449	0.6161	0.5784
Katz 1/4λ	0.9053	1.0000	0.9957	0.9922	0.8141	0.2059	0.2268	0.2265	0.2773	0.5917	0.5820	0.5595
Katz 1/2λ	0.9024	0.9957	1.0000	0.9966	0.8112	0.2078	0.2289	0.2286	0.2776	0.5914	0.5827	0.5611
Katz 3/4λ	0.9000	0.9922	0.9966	1.0000	0.8089	0.2094	0.2307	0.2303	0.2778	0.5911	0.5832	0.5622
SALSA	0.9114	0.8141	0.8112	0.8089	1.0000	0.1782	0.1617	0.1619	0.1917	0.6445	0.6146	0.5747
closeness	0.1950	0.2059	0.2078	0.2094	0.1782	1.0000	0.8592	0.8566	0.3817	0.1518	0.1746	0.2004
harmonic	0.2060	0.2268	0.2289	0.2307	0.1617	0.8592	1.0000	0.9694	0.4253	0.1503	0.1770	0.2072
Lin	0.2060	0.2265	0.2286	0.2303	0.1619	0.8566	0.9694	1.0000	0.4272	0.1503	0.1768	0.2069
HITS	0.2853	0.2773	0.2776	0.2778	0.1917	0.3817	0.4253	0.4272	1.0000	0.1529	0.1484	0.1415
PR 1/4	0.6449	0.5917	0.5914	0.5911	0.6445	0.1518	0.1503	0.1503	0.1529	1.0000	0.9182	0.8289
PR 1/2	0.6161	0.5820	0.5827	0.5832	0.6146	0.1746	0.1770	0.1768	0.1484	0.9182	1.0000	0.9088
PR 3/4	0.5784	0.5595	0.5611	0.5622	0.5747	0.2004	0.2072	0.2069	0.1415	0.8289	0.9088	1.0000

.uk (May 2007 snapshot)

Kendall's τ

	degree	Katz 1/4λ	Katz 1/2λ	Katz 3/4λ	SALSA	closeness	harmonic	Lin	HITS	PR 1/4	PR 1/2	PR 3/4
degree	1.0000	0.9053	0.9024	0.9000	0.9114	0.1950	0.2060	0.2060	0.2853	0.6449	0.6161	0.5784
Katz 1/4λ	0.9053	1.0000	0.9957	0.9922	0.8141	0.2059	0.2268	0.2265	0.2773	0.5917	0.5820	0.5595
Katz 1/2λ	0.9024	0.9957	1.0000	0.9966	0.8112	0.2078	0.2289	0.2286	0.2776	0.5914	0.5827	0.5611
Katz 3/4λ	0.9000	0.9922	0.9966	1.0000	0.8089	0.2094	0.2307	0.2303	0.2778	0.5911	0.5832	0.5622
SALSA	0.9114	0.8141	0.8112	0.8089	1.0000	0.1782	0.1617	0.1619	0.1917	0.6445	0.6146	0.5747
closeness	0.1950	0.2059	0.2078	0.2094	0.1782	1.0000	0.8592	0.8566	0.3817	0.1518	0.1746	0.2004
harmonic	0.2060	0.2268	0.2289	0.2307	0.1617	0.8592	1.0000	0.9694	0.4253	0.1503	0.1770	0.2072
Lin	0.2060	0.2265	0.2286	0.2303	0.1619	0.8566	0.9694	1.0000	0.4272	0.1503	0.1768	0.2069
HITS	0.2853	0.2773	0.2776	0.2778	0.1917	0.3817	0.4253	0.4272	1.0000	0.1529	0.1484	0.1415
PR 1/4	0.6449	0.5917	0.5914	0.5911	0.6445	0.1518	0.1503	0.1503	0.1529	1.0000	0.9182	0.8289
PR 1/2	0.6161	0.5820	0.5827	0.5832	0.6146	0.1746	0.1770	0.1768	0.1484	0.9182	1.0000	0.9088
PR 3/4	0.5784	0.5595	0.5611	0.5622	0.5747	0.2004	0.2072	0.2069	0.1415	0.8289	0.9088	1.0000

.uk (May 2007 snapshot)

Kendall's τ

	degree	Katz 1/4λ	Katz 1/2λ	Katz 3/4λ	SALSA	closeness	harmonic	Lin	HITS	PR 1/4	PR 1/2	PR 3/4
degree	1.0000	0.9053	0.9024	0.9000	0.9114	0.1950	0.2060	0.2060	0.2853	0.6449	0.6161	0.5784
Katz 1/4λ	0.9053	1.0000	0.9957	0.9922	0.8141	0.2059	0.2268	0.2265	0.2773	0.5917	0.5820	0.5595
Katz 1/2λ	0.9024	0.9957	1.0000	0.9966	0.8112	0.2078	0.2289	0.2286	0.2776	0.5914	0.5827	0.5611
Katz 3/4λ	0.9000	0.9922	0.9966	1.0000	0.8089	0.2094	0.2307	0.2303	0.2778	0.5911	0.5832	0.5622
SALSA	0.9114	0.8141	0.8112	0.8089	1.0000	0.1782	0.1617	0.1619	0.1917	0.6445	0.6146	0.5747
closeness	0.1950	0.2059	0.2078	0.2094	0.1782	1.0000	0.8592	0.8566	0.3817	0.1518	0.1746	0.2004
harmonic	0.2060	0.2268	0.2289	0.2307	0.1617	0.8592	1.0000	0.9694	0.4253	0.1503	0.1770	0.2072
Lin	0.2060	0.2265	0.2286	0.2303	0.1619	0.8566	0.9694	1.0000	0.4272	0.1503	0.1768	0.2069
HITS	0.2853	0.2773	0.2776	0.2778	0.1917	0.3817	0.4253	0.4272	1.0000	0.1529	0.1484	0.1415
PR 1/4	0.6449	0.5917	0.5914	0.5911	0.6445	0.1518	0.1503	0.1503	0.1529	1.0000	0.9182	0.8289
PR 1/2	0.6161	0.5820	0.5827	0.5832	0.6146	0.1746	0.1770	0.1768	0.1484	0.9182	1.0000	0.9088
PR 3/4	0.5784	0.5595	0.5611	0.5622	0.5747	0.2004	0.2072	0.2069	0.1415	0.8289	0.9088	1.0000

Betweenness could not be computed because of graph size (106M nodes)

.uk (May 2007 snapshot)

Kendall's τ

	degree	Katz 1/4λ	Katz 1/2λ	Katz 3/4λ	SALSA	closeness	harmonic	Lin	HITS	PR 1/4	PR 1/2	PR 3/4
degree	1.0000	0.9053	0.9024	0.9000	0.9114	0.1950	0.2060	0.2060	0.2853	0.6449	0.6161	0.5784
Katz 1/4λ	0.9053	1.0000	0.9957	0.9922	0.8141	0.2059	0.2268	0.2265	0.2773	0.5917	0.5820	0.5595
Katz 1/2λ	0.9024	0.9957	1.0000	0.9966	0.8112	0.2078	0.2289	0.2286	0.2776	0.5914	0.5827	0.5611
Katz 3/4λ	0.9000	0.9922	0.9966	1.0000	0.8089	0.2094	0.2307	0.2303	0.2778	0.5911	0.5832	0.5622
SALSA	0.9114	0.8141	0.8112	0.8089	1.0000	0.1782	0.1617	0.1619	0.1917	0.6445	0.6146	0.5747
closeness	0.1950	0.2059	0.2078	0.2094	0.1782	1.0000	0.8592	0.8566	0.3817	0.1518	0.1746	0.2004
harmonic	0.2060	0.2268	0.2289	0.2307	0.1617	0.8592	1.0000	0.9694	0.4253	0.1503	0.1770	0.2072
Lin	0.2060	0.2265	0.2286	0.2303	0.1619	0.8566	0.9694	1.0000	0.4272	0.1503	0.1768	0.2069
HITS	0.2853	0.2773	0.2776	0.2778	0.1917	0.3817	0.4253	0.4272	1.0000	0.1529	0.1484	0.1415
PR 1/4	0.6449	0.5917	0.5914	0.5911	0.6445	0.1518	0.1503	0.1503	0.1529	1.0000	0.9182	0.8289
PR 1/2	0.6161	0.5820	0.5827	0.5832	0.6146	0.1746	0.1770	0.1768	0.1484	0.9182	1.0000	0.9088
PR 3/4	0.5784	0.5595	0.5611	0.5622	0.5747	0.2004	0.2072	0.2069	0.1415	0.8289	0.9088	1.0000

.uk (May 2007 snapshot)

Kendall's τ

	degree	Katz 1/4λ	Katz 1/2λ	Katz 3/4λ	SALSA	closeness	harmonic	Lin	HITS	PR 1/4	PR 1/2	PR 3/4
degree	1.0000	0.9053	0.9024	0.9000	0.9114	0.1950	0.2060	0.2060	0.2853	0.6449	0.6161	0.5784
Katz 1/4λ	0.9053	1.0000	0.9957	0.9922	0.8141	0.2059	0.2268	0.2265	0.2773	0.5917	0.5820	0.5595
Katz 1/2λ	0.9024	0.9957	1.0000	0.9966	0.8112	0.2078	0.2289	0.2286	0.2776	0.5914	0.5827	0.5611
Katz 3/4λ	0.9000	0.9922	0.9966	1.0000	0.8089	0.2094	0.2307	0.2303	0.2778	0.5911	0.5832	0.5622
SALSA	0.9114	0.8141	0.8112	0.8089	1.0000	0.1782	0.1617	0.1619	0.1917	0.6445	0.6146	0.5747
closeness	0.1950	0.2059	0.2078	0.2094	0.1782	1.0000	0.8592	0.8566	0.3817	0.1518	0.1746	0.2004
harmonic	0.2060	0.2268	0.2289	0.2307	0.1617	0.8592	1.0000	0.9694	0.4253	0.1503	0.1770	0.2072
Lin	0.2060	0.2265	0.2286	0.2303	0.1619	0.8566	0.9694	1.0000	0.4272	0.1503	0.1768	0.2069
HITS	0.2853	0.2773	0.2776	0.2778	0.1917	0.3817	0.4253	0.4272	1.0000	0.1529	0.1484	0.1415
PR 1/4	0.6449	0.5917	0.5914	0.5911	0.6445	0.1518	0.1503	0.1503	0.1529	1.0000	0.9182	0.8289
PR 1/2	0.6161	0.5820	0.5827	0.5832	0.6146	0.1746	0.1770	0.1768	0.1484	0.9182	1.0000	0.9088
PR 3/4	0.5784	0.5595	0.5611	0.5622	0.5747	0.2004	0.2072	0.2069	0.1415	0.8289	0.9088	1.0000

The same correlations as with Hollywood,
but even more emphasized

.uk (May 2007 snapshot)

Kendall's τ

	degree	Katz 1/4λ	Katz 1/2λ	Katz 3/4λ	SALSA	closeness	harmonic	Lin	HITS	PR 1/4	PR 1/2	PR 3/4
degree	1.0000	0.9053	0.9024	0.9000	0.9114	0.1950	0.2060	0.2060	0.2853	0.6449	0.6161	0.5784
Katz 1/4λ	0.9053	1.0000	0.9957	0.9922	0.8141	0.2059	0.2268	0.2265	0.2773	0.5917	0.5820	0.5595
Katz 1/2λ	0.9024	0.9957	1.0000	0.9966	0.8112	0.2078	0.2289	0.2286	0.2776	0.5914	0.5827	0.5611
Katz 3/4λ	0.9000	0.9922	0.9966	1.0000	0.8089	0.2094	0.2307	0.2303	0.2778	0.5911	0.5832	0.5622
SALSA	0.9114	0.8141	0.8112	0.8089	1.0000	0.1782	0.1617	0.1619	0.1917	0.6445	0.6146	0.5747
closeness	0.1950	0.2059	0.2078	0.2094	0.1782	1.0000	0.8592	0.8566	0.3817	0.1518	0.1746	0.2004
harmonic	0.2060	0.2268	0.2289	0.2307	0.1617	0.8592	1.0000	0.9694	0.4253	0.1503	0.1770	0.2072
Lin	0.2060	0.2265	0.2286	0.2303	0.1619	0.8566	0.9694	1.0000	0.4272	0.1503	0.1768	0.2069
HITS	0.2853	0.2773	0.2776	0.2778	0.1917	0.3817	0.4253	0.4272	1.0000	0.1529	0.1484	0.1415
PR 1/4	0.6449	0.5917	0.5914	0.5911	0.6445	0.1518	0.1503	0.1503	0.1529	1.0000	0.9182	0.8289
PR 1/2	0.6161	0.5820	0.5827	0.5832	0.6146	0.1746	0.1770	0.1768	0.1484	0.9182	1.0000	0.9088
PR 3/4	0.5784	0.5595	0.5611	0.5622	0.5747	0.2004	0.2072	0.2069	0.1415	0.8289	0.9088	1.0000

.uk (May 2007 snapshot)

Kendall's τ

	degree	Katz 1/4λ	Katz 1/2λ	Katz 3/4λ	SALSA	closeness	harmonic	Lin	HITS	PR 1/4	PR 1/2	PR 3/4
degree	1.0000	0.9053	0.9024	0.9000	0.9114	0.1950	0.2060	0.2060	0.2853	0.6449	0.6161	0.5784
Katz 1/4λ	0.9053	1.0000	0.9957	0.9922	0.8141	0.2059	0.2268	0.2265	0.2773	0.5917	0.5820	0.5595
Katz 1/2λ	0.9024	0.9957	1.0000	0.9966	0.8112	0.2078	0.2289	0.2286	0.2776	0.5914	0.5827	0.5611
Katz 3/4λ	0.9000	0.9922	0.9966	1.0000	0.8089	0.2094	0.2307	0.2303	0.2778	0.5911	0.5832	0.5622
SALSA	0.9114	0.8141	0.8112	0.8089	1.0000	0.1782	0.1617	0.1619	0.1917	0.6445	0.6146	0.5747
closeness	0.1950	0.2059	0.2078	0.2094	0.1782	1.0000	0.8592	0.8566	0.3817	0.1518	0.1746	0.2004
harmonic	0.2060	0.2268	0.2289	0.2307	0.1617	0.8592	1.0000	0.9694	0.4253	0.1503	0.1770	0.2072
Lin	0.2060	0.2265	0.2286	0.2303	0.1619	0.8566	0.9694	1.0000	0.4272	0.1503	0.1768	0.2069
HITS	0.2853	0.2773	0.2776	0.2778	0.1917	0.3817	0.4253	0.4272	1.0000	0.1529	0.1484	0.1415
PR 1/4	0.6449	0.5917	0.5914	0.5911	0.6445	0.1518	0.1503	0.1503	0.1529	1.0000	0.9182	0.8289
PR 1/2	0.6161	0.5820	0.5827	0.5832	0.6146	0.1746	0.1770	0.1768	0.1484	0.9182	1.0000	0.9088
PR 3/4	0.5784	0.5595	0.5611	0.5622	0.5747	0.2004	0.2072	0.2069	0.1415	0.8289	0.9088	1.0000

Exception: HITS used to be correlated with the geometric indices, while now it is alone

.uk (May 2007 snapshot)

Kendall's τ

	degree	Katz 1/4λ	Katz 1/2λ	Katz 3/4λ	SALSA	closeness	harmonic	Lin	HITS	PR 1/4	PR 1/2	PR 3/4
degree	1.0000	0.9053	0.9024	0.9000	0.9114	0.1950	0.2060	0.2060	0.2853	0.6449	0.6161	0.5784
Katz 1/4λ	0.9053	1.0000	0.9957	0.9922	0.8141	0.2059	0.2268	0.2265	0.2773	0.5917	0.5820	0.5595
Katz 1/2λ	0.9024	0.9957	1.0000	0.9966	0.8112	0.2078	0.2289	0.2286	0.2776	0.5914	0.5827	0.5611
Katz 3/4λ	0.9000	0.9922	0.9966	1.0000	0.8089	0.2094	0.2307	0.2303	0.2778	0.5911	0.5832	0.5622
SALSA	0.9114	0.8141	0.8112	0.8089	1.0000	0.1782	0.1617	0.1619	0.1917	0.6445	0.6146	0.5747
closeness	0.1950	0.2059	0.2078	0.2094	0.1782	1.0000	0.8592	0.8566	0.3817	0.1518	0.1746	0.2004
harmonic	0.2060	0.2268	0.2289	0.2307	0.1617	0.8592	1.0000	0.9694	0.4253	0.1503	0.1770	0.2072
Lin	0.2060	0.2265	0.2286	0.2303	0.1619	0.8566	0.9694	1.0000	0.4272	0.1503	0.1768	0.2069
HITS	0.2853	0.2773	0.2776	0.2778	0.1917	0.3817	0.4253	0.4272	1.0000	0.1529	0.1484	0.1415
PR 1/4	0.6449	0.5917	0.5914	0.5911	0.6445	0.1518	0.1503	0.1503	0.1529	1.0000	0.9182	0.8289
PR 1/2	0.6161	0.5820	0.5827	0.5832	0.6146	0.1746	0.1770	0.1768	0.1484	0.9182	1.0000	0.9088
PR 3/4	0.5784	0.5595	0.5611	0.5622	0.5747	0.2004	0.2072	0.2069	0.1415	0.8289	0.9088	1.0000

.uk (May 2007 snapshot)

Kendall's τ

	degree	Katz 1/4λ	Katz 1/2λ	Katz 3/4λ	SALSA	closeness	harmonic	Lin	HITS	PR 1/4	PR 1/2	PR 3/4
degree	1.0000	0.9053	0.9024	0.9000	0.9114	0.1950	0.2060	0.2060	0.2853	0.6449	0.6161	0.5784
Katz 1/4λ	0.9053	1.0000	0.9957	0.9922	0.8141	0.2059	0.2268	0.2265	0.2773	0.5917	0.5820	0.5595
Katz 1/2λ	0.9024	0.9957	1.0000	0.9966	0.8112	0.2078	0.2289	0.2286	0.2776	0.5914	0.5827	0.5611
Katz 3/4λ	0.9000	0.9922	0.9966	1.0000	0.8089	0.2094	0.2307	0.2303	0.2778	0.5911	0.5832	0.5622
SALSA	0.9114	0.8141	0.8112	0.8089	1.0000	0.1782	0.1617	0.1619	0.1917	0.6445	0.6146	0.5747
closeness	0.1950	0.2059	0.2078	0.2094	0.1782	1.0000	0.8592	0.8566	0.3817	0.1518	0.1746	0.2004
harmonic	0.2060	0.2268	0.2289	0.2307	0.1617	0.8592	1.0000	0.9694	0.4253	0.1503	0.1770	0.2072
Lin	0.2060	0.2265	0.2286	0.2303	0.1619	0.8566	0.9694	1.0000	0.4272	0.1503	0.1768	0.2069
HITS	0.2853	0.2773	0.2776	0.2778	0.1917	0.3817	0.4253	0.4272	1.0000	0.1529	0.1484	0.1415
PR 1/4	0.6449	0.5917	0.5914	0.5911	0.6445	0.1518	0.1503	0.1503	0.1529	1.0000	0.9182	0.8289
PR 1/2	0.6161	0.5820	0.5827	0.5832	0.6146	0.1746	0.1770	0.1768	0.1484	0.9182	1.0000	0.9088
PR 3/4	0.5784	0.5595	0.5611	0.5622	0.5747	0.2004	0.2072	0.2069	0.1415	0.8289	0.9088	1.0000

A larger correlation between PageRank and Katz & degree

Orkut (2007 snapshot)
Kendall's τ

Orkut (2007 snapshot)

Kendall's τ

	degree	Katz 1/4λ	Katz 1/2λ	Katz 3/4λ	SALSA	closeness	harmonic	Lin	HITS	PR 1/4	PR 1/2	PR 3/4
degree	1.0000	0.9522	0.8982	0.8242	1.0000	0.5521	0.5391	0.5513	0.3508	0.7265	0.7596	0.8016
Katz 1/4λ	0.9522	1.0000	0.9489	0.8750	0.9522	0.5972	0.5875	0.5967	0.3984	0.6868	0.7179	0.7577
Katz 1/2λ	0.8982	0.9489	1.0000	0.9275	0.8982	0.6382	0.6338	0.6380	0.4491	0.6400	0.6690	0.7067
Katz 3/4λ	0.8242	0.8750	0.9275	1.0000	0.8242	0.6839	0.6910	0.6842	0.5213	0.5742	0.6005	0.6355
SALSA	1.0000	0.9522	0.8982	0.8242	1.0000	0.5521	0.5391	0.5513	0.3505	0.7265	0.7596	0.8016
closeness	0.5521	0.5972	0.6382	0.6839	0.5521	1.0000	0.9458	0.9830	0.6539	0.3862	0.4040	0.4268
harmonic	0.5391	0.5875	0.6338	0.6910	0.5391	0.9458	1.0000	0.9471	0.7090	0.3612	0.3777	0.3992
Lin	0.5513	0.5967	0.6380	0.6842	0.5513	0.9830	0.9471	1.0000	0.6546	0.3852	0.4030	0.4257
HITS	0.3508	0.3984	0.4491	0.5213	0.3505	0.6539	0.7090	0.6546	1.0000	0.1689	0.1778	0.1917
PR 1/4	0.7265	0.6868	0.6400	0.5742	0.7265	0.3862	0.3612	0.3852	0.1689	1.0000	0.9520	0.8889
PR 1/2	0.7596	0.7179	0.6690	0.6005	0.7596	0.4040	0.3777	0.4030	0.1778	0.9520	1.0000	0.9363
PR 3/4	0.8016	0.7577	0.7067	0.6355	0.8016	0.4268	0.3992	0.4257	0.1917	0.8889	0.9363	1.0000

Orkut (2007 snapshot)

Kendall's τ

	degree	Katz 1/4λ	Katz 1/2λ	Katz 3/4λ	SALSA	closeness	harmonic	Lin	HITS	PR 1/4	PR 1/2	PR 3/4
degree	1.0000	0.9522	0.8982	0.8242	1.0000	0.5521	0.5391	0.5513	0.3508	0.7265	0.7596	0.8016
Katz 1/4λ	0.9522	1.0000	0.9489	0.8750	0.9522	0.5972	0.5875	0.5967	0.3984	0.6868	0.7179	0.7577
Katz 1/2λ	0.8982	0.9489	1.0000	0.9275	0.8982	0.6382	0.6338	0.6380	0.4491	0.6400	0.6690	0.7067
Katz 3/4λ	0.8242	0.8750	0.9275	1.0000	0.8242	0.6839	0.6910	0.6842	0.5213	0.5742	0.6005	0.6355
SALSA	1.0000	0.9522	0.8982	0.8242	1.0000	0.5521	0.5391	0.5513	0.3505	0.7265	0.7596	0.8016
closeness	0.5521	0.5972	0.6382	0.6839	0.5521	1.0000	0.9458	0.9830	0.6539	0.3862	0.4040	0.4268
harmonic	0.5391	0.5875	0.6338	0.6910	0.5391	0.9458	1.0000	0.9471	0.7090	0.3612	0.3777	0.3992
Lin	0.5513	0.5967	0.6380	0.6842	0.5513	0.9830	0.9471	1.0000	0.6546	0.3852	0.4030	0.4257
HITS	0.3508	0.3984	0.4491	0.5213	0.3505	0.6539	0.7090	0.6546	1.0000	0.1689	0.1778	0.1917
PR 1/4	0.7265	0.6868	0.6400	0.5742	0.7265	0.3862	0.3612	0.3852	0.1689	1.0000	0.9520	0.8889
PR 1/2	0.7596	0.7179	0.6690	0.6005	0.7596	0.4040	0.3777	0.4030	0.1778	0.9520	1.0000	0.9363
PR 3/4	0.8016	0.7577	0.7067	0.6355	0.8016	0.4268	0.3992	0.4257	0.1917	0.8889	0.9363	1.0000

The same correlation as in Hollywood,
even if this time SALSA is also pretty correlated
with PageRank as well

Orkut (2007 snapshot)

Kendall's τ

	degree	Katz 1/4λ	Katz 1/2λ	Katz 3/4λ	SALSA	closeness	harmonic	Lin	HITS	PR 1/4	PR 1/2	PR 3/4
degree	1.0000	0.9522	0.8982	0.8242	1.0000	0.5521	0.5391	0.5513	0.3508	0.7265	0.7596	0.8016
Katz 1/4λ	0.9522	1.0000	0.9489	0.8750	0.9522	0.5972	0.5875	0.5967	0.3984	0.6868	0.7179	0.7577
Katz 1/2λ	0.8982	0.9489	1.0000	0.9275	0.8982	0.6382	0.6338	0.6380	0.4491	0.6400	0.6690	0.7067
Katz 3/4λ	0.8242	0.8750	0.9275	1.0000	0.8242	0.6839	0.6910	0.6842	0.5213	0.5742	0.6005	0.6355
SALSA	1.0000	0.9522	0.8982	0.8242	1.0000	0.5521	0.5391	0.5513	0.3505	0.7265	0.7596	0.8016
closeness	0.5521	0.5972	0.6382	0.6839	0.5521	1.0000	0.9458	0.9830	0.6539	0.3862	0.4040	0.4268
harmonic	0.5391	0.5875	0.6338	0.6910	0.5391	0.9458	1.0000	0.9471	0.7090	0.3612	0.3777	0.3992
Lin	0.5513	0.5967	0.6380	0.6842	0.5513	0.9830	0.9471	1.0000	0.6546	0.3852	0.4030	0.4257
HITS	0.3508	0.3984	0.4491	0.5213	0.3505	0.6539	0.7090	0.6546	1.0000	0.1689	0.1778	0.1917
PR 1/4	0.7265	0.6868	0.6400	0.5742	0.7265	0.3862	0.3612	0.3852	0.1689	1.0000	0.9520	0.8889
PR 1/2	0.7596	0.7179	0.6690	0.6005	0.7596	0.4040	0.3777	0.4030	0.1778	0.9520	1.0000	0.9363
PR 3/4	0.8016	0.7577	0.7067	0.6355	0.8016	0.4268	0.3992	0.4257	0.1917	0.8889	0.9363	1.0000

IR lens

(using TREC .gov2)

IR lens
(using TREC .gov2)

use centrality (in isolation or combined with textual features)
to rerank query results and see how good (bad) they do

TREC .gov2

TREC .gov2

* 150 queries (query title words, in AND; with stemming, no stopword elimination)

TREC .gov2

* 150 queries (query title words, in AND; with stemming, no stopword elimination)
* Generated the result graph using the method described by Najork et al. 2009 (a variant of Kleinberg's HITS graph, taking *a* in-links and *b* out-links)

TREC .gov2

- 150 queries (query title words, in AND; with stemming, no stopword elimination)
- Generated the result graph using the method described by Najork et al. 2009 (a variant of Kleinberg's HITS graph, taking *a* in-links and *b* out-links)
- Considered many combinations: here I present only the cases *a*=*b*=*0* (i.e., subgraph induced by the result set)

TREC .gov2

- 150 queries (query title words, in AND; with stemming, no stopword elimination)
- Generated the result graph using the method described by Najork et al. 2009 (a variant of Kleinberg's HITS graph, taking *a* in-links and *b* out-links)
- Considered many combinations: here I present only the cases *a*=*b*=*0* (i.e., subgraph induced by the result set)
- With or without intra-host links

P@10 and NDCG@10

P@10 and NDCG@10

P@10 and NDCG@10

Betweenness				
Closeness				
PageRank (best)				
Degree				
SALSA				
Katz (best)				
Lin				
HITS				
Harmonic				

P@10 and NDCG@10

	All links			
	P@10			
Betweenness	0.0584			
Closeness	0.1101			
PageRank (best)	0.1107			
Degree	0.1208			
SALSA	0.1221			
Katz (best)	0.1242			
Lin	0.1295			
HITS	0.1349			
Harmonic	0.1430			

P@10 and NDCG@10

	All links			
	P@10	NDCG@10		
Betweenness	0.0584	0.0595		
Closeness	0.1101	0.1061		
PageRank (best)	0.1107	0.1078		
Degree	0.1208	0.1091		
SALSA	0.1221	0.1194		
Katz (best)	0.1242	0.1228		
Lin	0.1295	0.1308		
HITS	0.1349	0.1364		
Harmonic	0.1430	0.1449		

P@10 and NDCG@10

	All links		Inter-host links only	
	P@10	NDCG@10	P@10	
Betweenness	0.0584	0.0595	0.0577	
Closeness	0.1101	0.1061	0.1121	
PageRank (best)	0.1107	0.1078	0.1295	
Degree	0.1208	0.1091	0.1248	
SALSA	0.1221	0.1194	0.1282	
Katz (best)	0.1242	0.1228	0.1262	
Lin	0.1295	0.1308	0.1248	
HITS	0.1349	0.1364	0.1107	
Harmonic	0.1430	0.1449	0.1262	

P@10 and NDCG@10

	All links		Inter-host links only	
	P@10	NDCG@10	P@10	NDCG@10
Betweenness	0.0584	0.0595	0.0577	0.0588
Closeness	0.1101	0.1061	0.1121	0.1168
PageRank (best)	0.1107	0.1078	0.1295	0.1347
Degree	0.1208	0.1091	0.1248	0.1283
SALSA	0.1221	0.1194	0.1282	0.1384
Katz (best)	0.1242	0.1228	0.1262	0.1297
Lin	0.1295	0.1308	0.1248	0.1286
HITS	0.1349	0.1364	0.1107	0.1179
Harmonic	0.1430	0.1449	0.1262	0.1293

Intra-host links?

Intra-host links?

* Keep them or throw them away?

Intra-host links?

* Keep them or throw them away?
* Most indices get **better** if you throw them away...

Intra-host links?

- Keep them or throw them away?
- Most indices get **better** if you throw them away...
- Throwing such links away injects a lot of information, but apparently harmonic doesn't need it!

Intra-host links?

- Keep them or throw them away?
- Most indices get **better** if you throw them away...
- Throwing such links away injects a lot of information, but apparently harmonic doesn't need it!
- ...but **harmonic** is better (and best of all) with the whole thing!

.uk (May 2007 snapshot)

Kendall's τ with no intra-host links

	degree	Katz 1/4λ	Katz 1/2λ	Katz 3/4λ	SALSA	closeness	harmonic	Lin	HITS	PR 1/4	PR 1/2	PR 3/4
degree	1.0000	0.9995	0.9995	0.9995	0.9965	0.9883	0.9984	0.9984	0.8767	0.9956	0.9956	0.9955
Katz 1/4λ	0.9995	1.0000	1.0000	1.0000	0.9960	0.9870	0.9986	0.9978	0.8763	0.9952	0.9953	0.9953
Katz 1/2λ	0.9995	1.0000	1.0000	1.0000	0.9960	0.9870	0.9986	0.9978	0.8763	0.9952	0.9953	0.9953
Katz 3/4λ	0.9995	1.0000	1.0000	1.0000	0.9960	0.9870	0.9986	0.9978	0.8763	0.9952	0.9953	0.9953
SALSA	0.9965	0.9960	0.9960	0.9960	1.0000	0.9904	0.9950	0.9950	0.8718	0.9959	0.9959	0.9958
closeness	0.9883	0.9870	0.9870	0.9870	0.9904	1.0000	0.9859	0.9876	0.8714	0.9894	0.9893	0.9892
harmonic	0.9984	0.9986	0.9986	0.9986	0.9950	0.9859	1.0000	0.9984	0.8759	0.9944	0.9945	0.9946
Lin	0.9984	0.9978	0.9978	0.9978	0.9950	0.9876	0.9984	1.0000	0.8759	0.9945	0.9946	0.9946
HITS	0.8767	0.8763	0.8763	0.8763	0.8718	0.8714	0.8759	0.8759	1.0000	0.8727	0.8727	0.8727
PR 1/4	0.9956	0.9952	0.9952	0.9952	0.9959	0.9894	0.9944	0.9945	0.8727	1.0000	0.9998	0.9997
PR 1/2	0.9956	0.9953	0.9953	0.9953	0.9959	0.9893	0.9945	0.9946	0.8727	0.9998	1.0000	0.9999
PR 3/4	0.9955	0.9953	0.9953	0.9953	0.9958	0.9892	0.9946	0.9946	0.8727	0.9997	0.9999	1.0000

.uk (May 2007 snapshot)

Kendall's τ with no intra-host links

	degree	Katz 1/4λ	Katz 1/2λ	Katz 3/4λ	SALSA	closeness	harmonic	Lin	HITS	PR 1/4	PR 1/2	PR 3/4
degree	1.0000	0.9995	0.9995	0.9995	0.9965	0.9883	0.9984	0.9984	0.8767	0.9956	0.9956	0.9955
Katz 1/4λ	0.9995	1.0000	1.0000	1.0000	0.9960	0.9870	0.9986	0.9978	0.8763	0.9952	0.9953	0.9953
Katz 1/2λ	0.9995	1.0000	1.0000	1.0000	0.9960	0.9870	0.9986	0.9978	0.8763	0.9952	0.9953	0.9953
Katz 3/4λ	0.9995	1.0000	1.0000	1.0000	0.9960	0.9870	0.9986	0.9978	0.8763	0.9952	0.9953	0.9953
SALSA	0.9965	0.9960	0.9960	0.9960	1.0000	0.9904	0.9950	0.9950	0.8718	0.9959	0.9959	0.9958
closeness	0.9883	0.9870	0.9870	0.9870	0.9904	1.0000	0.9859	0.9876	0.8714	0.9894	0.9893	0.9892
harmonic	0.9984	0.9986	0.9986	0.9986	0.9950	0.9859	1.0000	0.9984	0.8759	0.9944	0.9945	0.9946
Lin	0.9984	0.9978	0.9978	0.9978	0.9950	0.9876	0.9984	1.0000	0.8759	0.9945	0.9946	0.9946
HITS	0.8767	0.8763	0.8763	0.8763	0.8718	0.8714	0.8759	0.8759	1.0000	0.8727	0.8727	0.8727
PR 1/4	0.9956	0.9952	0.9952	0.9952	0.9959	0.9894	0.9944	0.9945	0.8727	1.0000	0.9998	0.9997
PR 1/2	0.9956	0.9953	0.9953	0.9953	0.9959	0.9893	0.9945	0.9946	0.8727	0.9998	1.0000	0.9999
PR 3/4	0.9955	0.9953	0.9953	0.9953	0.9958	0.9892	0.9946	0.9946	0.8727	0.9997	0.9999	1.0000

Everything becomes correlated. Why???

.uk (May 2007 snapshot)

Kendall's τ with no intra-host links

	degree	Katz 1/4λ	Katz 1/2λ	Katz 3/4λ	SALSA	closeness	harmonic	Lin	HITS	PR 1/4	PR 1/2	PR 3/4
degree	1.0000	0.9995	0.9995	0.9995	0.9965	0.9883	0.9984	0.9984	0.8767	0.9956	0.9956	0.9955
Katz 1/4λ	0.9995	1.0000	1.0000	1.0000	0.9960	0.9870	0.9986	0.9978	0.8763	0.9952	0.9953	0.9953
Katz 1/2λ	0.9995	1.0000	1.0000	1.0000	0.9960	0.9870	0.9986	0.9978	0.8763	0.9952	0.9953	0.9953
Katz 3/4λ	0.9995	1.0000	1.0000	1.0000	0.9960	0.9870	0.9986	0.9978	0.8763	0.9952	0.9953	0.9953
SALSA	0.9965	0.9960	0.9960	0.9960	1.0000	0.9904	0.9950	0.9950	0.8718	0.9959	0.9959	0.9958
closeness	0.9883	0.9870	0.9870	0.9870	0.9904	1.0000	0.9859	0.9876	0.8714	0.9894	0.9893	0.9892
harmonic	0.9984	0.9986	0.9986	0.9986	0.9950	0.9859	1.0000	0.9984	0.8759	0.9944	0.9945	0.9946
Lin	0.9984	0.9978	0.9978	0.9978	0.9950	0.9876	0.9984	1.0000	0.8759	0.9945	0.9946	0.9946
HITS	0.8767	0.8763	0.8763	0.8763	0.8718	0.8714	0.8759	0.8759	1.0000	0.8727	0.8727	0.8727
PR 1/4	0.9956	0.9952	0.9952	0.9952	0.9959	0.9894	0.9944	0.9945	0.8727	1.0000	0.9998	0.9997
PR 1/2	0.9956	0.9953	0.9953	0.9953	0.9959	0.9893	0.9945	0.9946	0.8727	0.9998	1.0000	0.9999
PR 3/4	0.9955	0.9953	0.9953	0.9953	0.9958	0.9892	0.9946	0.9946	0.8727	0.9997	0.9999	1.0000

.uk (May 2007 snapshot)

Kendall's τ with no intra-host links

	degree	Katz 1/4λ	Katz 1/2λ	Katz 3/4λ	SALSA	closeness	harmonic	Lin	HITS	PR 1/4	PR 1/2	PR 3/4
degree	1.0000	0.9995	0.9995	0.9995	0.9965	0.9883	0.9984	0.9984	0.8767	0.9956	0.9956	0.9955
Katz 1/4λ	0.9995	1.0000	1.0000	1.0000	0.9960	0.9870	0.9986	0.9978	0.8763	0.9952	0.9953	0.9953
Katz 1/2λ	0.9995	1.0000	1.0000	1.0000	0.9960	0.9870	0.9986	0.9978	0.8763	0.9952	0.9953	0.9953
Katz 3/4λ	0.9995	1.0000	1.0000	1.0000	0.9960	0.9870	0.9986	0.9978	0.8763	0.9952	0.9953	0.9953
SALSA	0.9965	0.9960	0.9960	0.9960	1.0000	0.9904	0.9950	0.9950	0.8718	0.9959	0.9959	0.9958
closeness	0.9883	0.9870	0.9870	0.9870	0.9904	1.0000	0.9859	0.9876	0.8714	0.9894	0.9893	0.9892
harmonic	0.9984	0.9986	0.9986	0.9986	0.9950	0.9859	1.0000	0.9984	0.8759	0.9944	0.9945	0.9946
Lin	0.9984	0.9978	0.9978	0.9978	0.9950	0.9876	0.9984	1.0000	0.8759	0.9945	0.9946	0.9946
HITS	0.8767	0.8763	0.8763	0.8763	0.8718	0.8714	0.8759	0.8759	1.0000	0.8727	0.8727	0.8727
PR 1/4	0.9956	0.9952	0.9952	0.9952	0.9959	0.9894	0.9944	0.9945	0.8727	1.0000	0.9998	0.9997
PR 1/2	0.9956	0.9953	0.9953	0.9953	0.9959	0.9893	0.9945	0.9946	0.8727	0.9998	1.0000	0.9999
PR 3/4	0.9955	0.9953	0.9953	0.9953	0.9958	0.9892	0.9946	0.9946	0.8727	0.9997	0.9999	1.0000

Because **most** nodes have degree 0 (and hence they also have all the other scores at a tie)

P@10 and NDCG@10

P@10 and NDCG@10

	All links		Inter-host links only	
	P@10	NDCG@10	P@10	NDCG@10

P@10 and NDCG@10

	All links		Inter-host links only	
	P@10	NDCG@10	P@10	NDCG@10
Weighted degree	0.1356	0.1373	0.1262	0.1295

P@10 and NDCG@10

$d^-(x) \cdot \text{canReach}(x)$

	All links		Inter-host links only	
	P@10	NDCG@10	P@10	NDCG@10
Weighted degree	0.1356	0.1373	0.1262	0.1295

P@10 and NDCG@10

	All links		Inter-host links only	
	P@10	NDCG@10	P@10	NDCG@10
Weighted degree	0.1356	0.1373	0.1262	0.1295

P@10 and NDCG@10

	All links		Inter-host links only	
	P@10	NDCG@10	P@10	NDCG@10
Weighted degree	0.1356	0.1373	0.1262	0.1295
SALSinA	0.1349	0.1357	0.1255	0.1318

P@10 and NDCG@10

	All links		Inter-host links only	
	P@10	NDCG@10	P@10	NDCG@10
Weighted degree	0.1356	0.1373	0.1262	0.1295
SALSinA	0.1349	0.1357	0.1255	0.1318

$d^-(x) \cdot \text{connected}(x)$

P@10 and NDCG@10

	All links		Inter-host links only	
	P@10	NDCG@10	P@10	NDCG@10
Weighted degree	0.1356	0.1373	0.1262	0.1295
SALSinA	0.1349	0.1357	0.1255	0.1318

P@10 and NDCG@10

	All links		Inter-host links only	
	P@10	NDCG@10	P@10	NDCG@10
Weighted degree	0.1356	0.1373	0.1262	0.1295
SALSinA	0.1349	0.1357	0.1255	0.1318
Degree	0.1208	0.1091	0.1248	0.1283
SALSA	0.1221	0.1194	0.1282	0.1384
HITS	0.1349	0.1364	0.1107	0.1179
Harmonic	0.1430	0.1449	0.1262	0.1293

P@10 and NDCG@10

	All links		Inter-host links only	
	P@10	NDCG@10	P@10	NDCG@10
Weighted degree	0.1356	0.1373	0.1262	0.1295
SALSinA	0.1349	0.1357	0.1255	0.1318
Degree	0.1208	0.1091	0.1248	0.1283
SALSA	0.1221	0.1194	0.1282	0.1384
HITS	0.1349	0.1364	0.1107	0.1179
Harmonic	0.1430	0.1449	0.1262	0.1293

P@10 and NDCG@10

	All links		Inter-host links only	
	P@10	NDCG@10	P@10	NDCG@10
Weighted degree	0.1356	0.1373	0.1262	0.1295
SALSinA	0.1349	0.1357	0.1255	0.1318
Degree	0.1208	0.1091	0.1248	0.1283
SALSA	0.1221	0.1194	0.1282	0.1384
HITS	0.1349	0.1364	0.1107	0.1179
Harmonic	0.1430	0.1449	0.1262	0.1293
BM25	0.5644	0.5842	0.5644	0.5842

Computational feasibility lens

Computational feasibility lens

which indices are computable effectively on *large* networks; consider also parallelizability / distributability...

Best algorithms so far

Best algorithms so far

Best algorithms so far

	How?		Progr.	

Best algorithms so far

	How?		Progr.	
Degree	Trivial O(1)	-	-	+++

Best algorithms so far

	How?		Progr.	
Degree	Trivial O(1)	-	-	+++
Betweennes	O(nm) [Brandes 2001]	-	no	-

Best algorithms so far

	How?		Progr.	
Degree	Trivial O(1)	-	-	+++
Betweennes	O(nm) [Brandes 2001]	-	no	-
Katz	Iterative (Gauss-Seidel)	Fast convergence	yes	+

Best algorithms so far

	How?		Progr.	
Degree	Trivial O(1)	-	-	+++
Betweennes	O(nm) [Brandes 2001]	-	no	-
Katz	Iterative (Gauss-Seidel)	Fast convergence	yes	+
Closeness	By diffusion or sampling	No unbiased estimator	no	++

Best algorithms so far

	How?		Progr.	
Degree	Trivial O(1)	-	-	+++
Betweennes	O(nm) [Brandes 2001]	-	no	-
Katz	Iterative (Gauss-Seidel)	Fast convergence	yes	+
Closeness	By diffusion or sampling	No unbiased estimator	no	++
Lin	By diffusion	No unbiased estimator	no	+

Best algorithms so far

	How?		Progr.	
Degree	Trivial O(1)	-	-	+++
Betweennes	O(nm) [Brandes 2001]	-	no	-
Katz	Iterative (Gauss-Seidel)	Fast convergence	yes	+
Closeness	By diffusion or sampling	No unbiased estimator	no	++
Lin	By diffusion	No unbiased estimator	no	+
Harmonic	By diffusion	Unbiased estimator	yes	+++

Best algorithms so far

	How?		Progr.	
Degree	Trivial O(1)	-	-	+++
Betweennes	O(nm) [Brandes 2001]	-	no	-
Katz	Iterative (Gauss-Seidel)	Fast convergence	yes	+
Closeness	By diffusion or sampling	No unbiased estimator	no	++
Lin	By diffusion	No unbiased estimator	no	+
Harmonic	By diffusion	Unbiased estimator	yes	+++
PageRank	Iterative (PM, GS, Jacobi...)	Fast convergence	yes	+++

Best algorithms so far

	How?		Progr.	
Degree	Trivial O(1)	-	-	+++
Betweennes	O(nm) [Brandes 2001]	-	no	-
Katz	Iterative (Gauss-Seidel)	Fast convergence	yes	+
Closeness	By diffusion or sampling	No unbiased estimator	no	++
Lin	By diffusion	No unbiased estimator	no	+
Harmonic	By diffusion	Unbiased estimator	yes	+++
PageRank	Iterative (PM, GS, Jacobi...)	Fast convergence	yes	+++
Seeley	Iterative (PM)	Slow convergence	no	-

Best algorithms so far

	How?		Progr.	
Degree	Trivial O(1)	-	-	+++
Betweennes	O(nm) [Brandes 2001]	-	no	-
Katz	Iterative (Gauss-Seidel)	Fast convergence	yes	+
Closeness	By diffusion or sampling	No unbiased estimator	no	++
Lin	By diffusion	No unbiased estimator	no	+
Harmonic	By diffusion	Unbiased estimator	yes	+++
PageRank	Iterative (PM, GS, Jacobi...)	Fast convergence	yes	+++
Seeley	Iterative (PM)	Slow convergence	no	-
HITS	Iterative (PM)	Slow convergence	no	-

Best algorithms so far

	How?		Progr.	
Degree	Trivial O(1)	-	-	+++
Betweennes	O(nm) [Brandes 2001]	-	no	-
Katz	Iterative (Gauss-Seidel)	Fast convergence	yes	+
Closeness	By diffusion or sampling	No unbiased estimator	no	++
Lin	By diffusion	No unbiased estimator	no	+
Harmonic	By diffusion	Unbiased estimator	yes	+++
PageRank	Iterative (PM, GS, Jacobi...)	Fast convergence	yes	+++
Seeley	Iterative (PM)	Slow convergence	no	-
HITS	Iterative (PM)	Slow convergence	no	-
SALSA	Direct O(sum sq. degr.)	-	no	+++

…and the winner is…

…and the winner is…

- Almost everybody (or you wouldn't be listening to this talk…)

…and the winner is…

- Almost everybody (or you wouldn't be listening to this talk…)
- Lin is more problematic, though:

$$c_{\mathrm{Lin}}(x) = \frac{\mathrm{canReach}(x)^2}{\sum_y d(y,x)}$$

…and the winner is…

- Almost everybody (or you wouldn't be listening to this talk…)
- Lin is more problematic, though:

$$c_{\mathrm{Lin}}(x) = \frac{\mathrm{canReach}(x)^2}{\sum_y d(y,x)}$$

- The numerator makes it difficult…

…and the winner is…

- Almost everybody (or you wouldn't be listening to this talk…)
- Lin is more problematic, though:

$$c_{\mathrm{Lin}}(x) = \frac{\mathrm{canReach}(x)^2}{\sum_y d(y,x)}$$

- The numerator makes it difficult…
  - to obtain an unbiased estimator

…and the winner is…

- Almost everybody (or you wouldn't be listening to this talk…)
- Lin is more problematic, though:

$$c_{\mathrm{Lin}}(x) = \frac{\mathrm{canReach}(x)^2}{\sum_y d(y,x)}$$

- The numerator makes it difficult…
  - to obtain an unbiased estimator
  - to have monotone convergence

…and the winner is…

…and the winner is…

* On the contrary, harmonic…

…and the winner is…

- On the contrary, harmonic…
- But how easy is it to compute?

$$c_{\mathrm{harm}}(x) = \sum_{y \neq x} \frac{1}{d(y,x)}$$

...and the winner is...

- On the contrary, harmonic...
- But how easy is it to compute?

$$c_{\text{harm}}(x) = \sum_{y \neq x} \frac{1}{d(y,x)}$$

- ...or...

$$c_{\text{harm}}(x) = \sum_{t=1}^{\infty} \frac{|B_t(x)| - |B_{t-1}(x)|}{t}$$

...and the winner is...

* On the contrary, harmonic...
* But how easy is it to compute?

$$c_{\text{harm}}(x) = \sum_{y \neq x} \frac{1}{d(y,x)}$$

* ...or...

$$c_{\text{harm}}(x) = \sum_{t=1}^{\infty} \frac{|B_t(x)| - |B_{t-1}(x)|}{t}$$

t: Ball of radius t about x

...and the winner is...

- On the contrary, harmonic...
- But how easy is it to compute?

$$c_{\text{harm}}(x) = \sum_{y \neq x} \frac{1}{d(y,x)}$$

- ...or...

$$c_{\text{harm}}(x) = \sum_{t=1}^{\infty} \frac{|B_t(x)| - |B_{t-1}(x)|}{t}$$

Computing by diffusion

Computing by diffusion

- Clearly

$$B_0(x) = \{x\}$$

Computing by diffusion

- Clearly

$$B_0(x) = \{x\}$$

- Moreover

$$B_{t+1}(x) = \{x\} \cup \bigcup_{x \to y} B_t(y)$$

Computing by diffusion

- Clearly

$$B_0(x) = \{x\}$$

- Moreover

$$B_{t+1}(x) = \{x\} \cup \bigcup_{x \to y} B_t(y)$$

Computing by diffusion

- Clearly

$$B_0(x) = \{x\}$$

- Moreover

$$B_{t+1}(x) = \{x\} \cup \bigcup_{x \to y} B_t(y)$$

- So one needs just one single sequential scan of the graph to compute the balls at the next iteration

Computing by diffusion

- Clearly

$$B_0(x) = \{x\}$$

- Moreover

$$B_{t+1}(x) = \{x\} \cup \bigcup_{x \to y} B_t(y)$$

- So one needs just one single sequential scan of the graph to compute the balls at the next iteration
- And you don't really compute balls, but very small *probabilistic sketches* that represent them

And balls are important!

And balls are important!

* HyperANF (http://webgraph.dsi.unimi.it/) uses the same technique to compute distance distributions

And balls are important!

* HyperANF (http://webgraph.dsi.unimi.it/) uses the same technique to compute distance distributions


HOME PAGE | TODAY'S PAPER | VIDEO | MOST POPULAR | TIMES TOPICS

The New York Times — Business Day Technology

WORLD | U.S. | N.Y. / REGION | BUSINESS | TECHNOLOGY | SCIENCE | HEALTH | SP

Separating You and Me? 4.74 Degrees

By JOHN MARKOFF and SOMINI SENGUPTA
Published: November 21, 2011

The world is even smaller than you thought.

RECOMMEND
TWITTER

And balls are important!

* HyperANF (http://webgraph.dsi.unimi.it/) uses the same technique to compute distance distributions


HOME PAGE | TODAY'S PAPER | VIDEO | MOST POPULAR | TIMES TOPICS

The New York Times — Business Day Technology

WORLD | U.S. | N.Y. / REGION | BUSINESS | TECHNOLOGY | SCIENCE | HEALTH | SP

Separating You and Me? 4.74 Degrees

By JOHN MARKOFF and SOMINI SENGUPTA
Published: November 21, 2011

The world is even smaller than you thought.

RECOMMEND
TWITTER

* Adapted for computing closeness, Lin, harmonic

And balls are important!

* HyperANF (http://webgraph.dsi.unimi.it/) uses the same technique to compute distance distributions


HOME PAGE | TODAY'S PAPER | VIDEO | MOST POPULAR | TIMES TOPICS

The New York Times — Business Day Technology

WORLD | U.S. | N.Y. / REGION | BUSINESS | TECHNOLOGY | SCIENCE | HEALTH | SP

Separating You and Me? 4.74 Degrees

By JOHN MARKOFF and SOMINI SENGUPTA
Published: November 21, 2011

The world is even smaller than you thought.

RECOMMEND
TWITTER

* Adapted for computing closeness, Lin, harmonic
* Diffusion processes are easily parallelizable!

Did we shade some light?

Did we shade some light?

* We have three independent and unrelated methodologies giving very consistent results

Did we shade some light?

- We have three independent and unrelated methodologies giving very consistent results
- Some things have been **ruled out**, at least

Did we shade some light?

* We have three independent and unrelated methodologies giving very consistent results
* Some things have been **ruled out**, at least
* Some others, that were neglected, have found **revenge**

Did we shade some light?

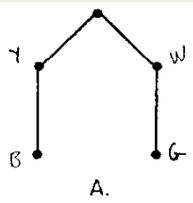
- We have three independent and unrelated methodologies giving very consistent results
- Some things have been **ruled out**, at least
- Some others, that were neglected, have found **revenge**
- Some new ones seem to be promising

Did we shade some light?

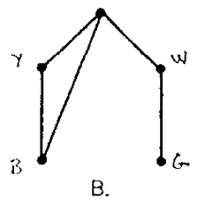
- We have three independent and unrelated methodologies giving very consistent results
- Some things have been **ruled out**, at least
- Some others, that were neglected, have found **revenge**
- Some new ones seem to be promising
- Next time you need a centrality index... try harmonic!

…when size matters…

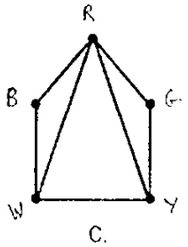
(Freeman 1979)



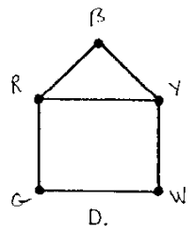
A.



B.



C.

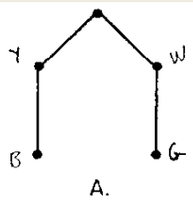


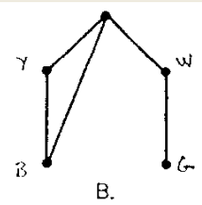
D.

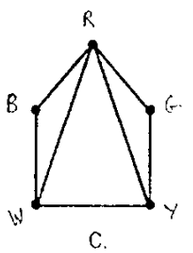
Subjects were 100 volunteers from among the student body at Lehigh University. They ranged from freshmen to graduate students; 51 were male and 49 were female. None had previously participated in a group problem-solving experiment.

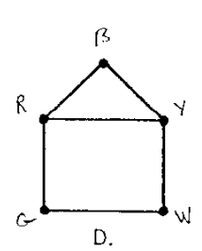
…when size matters…

(Freeman 1979)







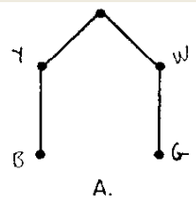


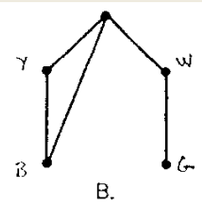
Subjects were 100 volunteers from among the student body at Lehigh University. They ranged from freshmen to graduate students; 51 were male and 49 were female. None had previously participated in a group problem-solving experiment.

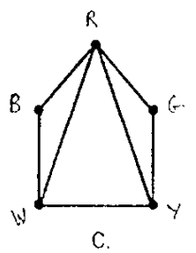
Betweenness, then, seems to be the key to understanding choice as leader

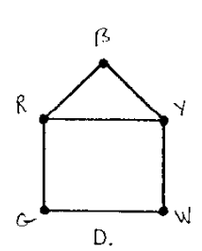
…when size matters…

(Freeman 1979)









Subjects were 100 volunteers from among the student body at Lehigh University. They ranged from freshmen to graduate students; 51 were male and 49 were female. None had previously participated in a group problem-solving experiment.

But on these four networks,
other centrality measures (e.g., harmonic)
give exactly the same result!